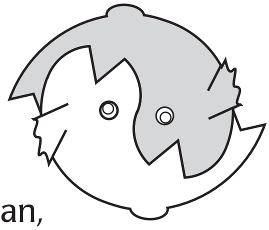
FISHERY STOCK ASSESSMENT MODELS



Edited by F. Funk,
T.J. Quinn II, J. Heifetz,
J.N. Ianelli, J.E. Powers,
J.F. Schweigert, P.J. Sullivan,
and C.-I. Zhang

Proceedings of the International Symposium on
Fishery Stock Assessment Models for the 21st Century,
October 8-11, 1997, Anchorage, Alaska

Lowell Wakefield Fisheries Symposium

University of Alaska Sea Grant College Program
AK-SG-98-01 1998 Price $40.00

Elmer E. Rasmuson Library Cataloging-in-Publication Data

International Symposium on Fishery Stock Assessment Models for the 21st Century (1997 : Anchorage, Alaska).

Fishery stock assessment models : proceedings of the International Symposium on Fishery Stock Assessment Models for the 21st Century, October 8-11, 1997, Anchorage, Alaska.

(Lowell Wakefield Fisheries Symposium) (University of Alaska Sea Grant College Program ; AK-SG-98-01)

1. Fish stock assessment—Alaska—Congresses. 2. Fishery management—Alaska—Congresses. I. Alaska Sea Grant College Program. II. Title. III. Series: Lowell Wakefield fisheries symposium series. 15th. IV. Series: Alaska Sea Grant College Program report : 98-01.

SH329.F56F57 1997

ISBN 1-56612-057-8

Citation for this volume is Fishery Stock Assessment Models, edited by F. Funk, T.J. Quinn II, J. Heifetz, J.N. Ianelli, J.E. Powers, J.F. Schweigert, P.J. Sullivan, and C.-I. Zhang, Alaska Sea Grant College Program Report No. AK-SG-98-01, University of Alaska Fairbanks, 1998.

Acknowledgments

This book is the result of work sponsored by the University of Alaska Sea Grant College Program, which is cooperatively supported by the U.S. Department of Commerce, NOAA National Sea Grant Office, under grant no. NA86RG-0050, project A/75-01; and by the University of Alaska with state funds. The University of Alaska is an affirmative action/equal opportunity institution.

Sea Grant is a unique partnership with public and private sectors combining research, education, and technology transfer for public service. This national network of universities meets changing environmental and economic needs of people in our coastal, ocean, and Great Lakes regions.

Contents

Keynote Address

Data Issues and Model Specification

Stage-Structured and Size-Based Applications

Ocean Ecosystems

Harvest Policy

About the Symposium

The International Symposium on Fishery Stock Assessment Models for the 21st Century: Combining Multiple Information Sources is the fifteenth Lowell Wakefield symposium. The program concept was suggested by Fritz Funk of the Alaska Department of Fish and Game. The meeting was held October 8-11, 1997, in Anchorage, Alaska. Eighty-five presentations were made, with a keynote address by Keith Sainsbury, CSIRO (Commonwealth Scientific and Industrial Research Organisation) Division of Marine Research, Hobart, Australia.

The symposium was organized and coordinated by Brenda Baxter, Alaska Sea Grant College Program, with the assistance of the organizing and program committees. Organizing committee members are: Fritz Funk, co-chair, Alaska Department of Fish and Game; Terrance J. Quinn II, co-chair, Fisheries Division, University of Alaska Fairbanks; Jane DiCosimo, North Pacific Fishery Management Council; James N. Ianelli, U.S. National Marine Fisheries Service; and Patrick J. Sullivan, International Pacific Halibut Commission. Program planning committee members are: Alajandro Anganuzzi, Inter-American Tropical Tuna Commission; Rick Deriso, Inter-American Tropical Tuna Commission; Stratis Gavaris, Department of Fisheries and Oceans Canada; John Hampton, South Pacific Commission; Tore Jakobsen, Institute of Marine Research, Norway; Peggy Merritt, Alaska Department of Fish and Game; Rick Methot, U.S. National Marine Fisheries Service; Steven A. Murawski, U.S. National Marine Fisheries Service; Peggy Murphy, Alaska Department of Fish and Game; Tony J. Pitcher, University of British Columbia, Canada; Joseph Powers, U.S. National Marine Fisheries Service; Jacob F. Schweigert, Department of Fisheries and Oceans Canada; John Sibert, University of Hawaii; Stephen J. Smith, Department of Fisheries and Oceans Canada; Tony Smith, Division of Fisheries, CSIRO, Australia; Gunnar Stefansson, Marine Research Institute, Iceland; and Chang-Ik Zhang, National Fisheries University of Pusan, Korea.

Symposium sponsors are: Alaska Department of Fish and Game; International Pacific Halibut Commission; North Pacific Fishery Management Council; U.S. National Marine Fisheries Service, Alaska Fisheries Science Center; and Alaska Sea Grant College Program, University of Alaska Fairbanks.

About This Proceedings

This publication has 55 symposium papers. Each paper has been reviewed by two peer reviewers.

Peer reviewers are: Milo Adkison, Jim Anderson, John Anderson, Manuela Azevedo, Timothy Baker, Rex Baleña, Bill Bechtol, Dave Bernard, Andrew Betlehem, Höskuldur Björnsson, Tony Booth, Barry Bracken, Jon Brodziak, Joan Browder, Robert Campbell, David Carlile, Lew Coggins, Keith Criddle, Paul Crone, Nancie Cummings, Nick Davies, José De Oliveira,

Richard Deriso, Martin Dorn, Emmanis Dorval, Lowell Fair, Chris Francis, Steven Fried, Caihong Fu, Jeff Fujioka, Fritz Funk, Maria Gasalla, Stratis Gavaris, Harold Geiger, Gudmundur Gudmundsson, Vivian Haist, John Hampton, Stuart Hanchet, David Hankin, Jonathan Heifetz, Ray Hilborn, Ryan Hill, Mark Huiskes, Victoria Isaac, Larry Jacobson, Tore Jakobsen, Christopher Jones, Dan Kimura, Tom Kline, Cliff Kraft, Gordon Kruse, Hanlin Lai, Patricia Livingston, Chris Lunsford, Alec MacCall, Bob Marshall, Michele Masuda, Ole Mathisen, Mark Maunder, Sandy McFarlane, Peggy Merritt, Rick Methot, Steven Murawski, Peggy Murphy, Ken Newman, Jorge Flores Olivares, Ted Otis, William Overholtz, Ana Parma, A.J. Paul, Daniel Pauly, Jerry Pella, Ian Perry, Tony Pitcher, Tom Polacheck, Clay Porch, Ann Preece, Andre Punt, Terry Quinn, Steve Ralston, Victor Restrepo, Jake Rice, Laura Richards, Gregg Rosenkranz, Saul Saila, David Sampson, Mark Saunders, Jon Schnute, Jake Schweigert, John Sibert, Michael Sigler, Alan Sinclair, Stephen Smith, Tony Smith, Kevin Stokesbury, Zhenming Su, Jack Tagart, Joe Terry, Grant Thompson, Sachiko Tsuji, Clive Turnbull, Jack Turnock, Douglas Vaughan, Jay Ver Hoef, Erik Williams, George Winters, Dave Witherell, Chris Wood, Douglas Woodby, and Jie Zheng.

Editing, layout, format, and proofing are by Carol Kaynor, Sue Keller, and Brenda Baxter of University of Alaska Sea Grant. Cover design is by Susan Gibson and David Brenner.

The Lowell Wakefield Symposium Series

The University of Alaska Sea Grant College Program has been sponsoring and coordinating the Lowell Wakefield Fisheries Symposium series since 1982. These meetings are a forum for information exchange in biology, management, economics, and processing of fish species and complexes, as well as an opportunity for scientists from high latitude countries to meet and discuss their work.

The symposium series had its origin in the Americanization of the fisheries off Alaska in the late 1970s. At that time a lack of information on target species impeded the ability to make good management decisions. In 1979 the North Pacific Fishery Management Council Scientific and Statistical Committee (SSC) recommended that scientists meet to look at the pandalid shrimp resource, in circumpolar countries where the shrimp was commercially important. The meeting was held that year and a proceedings with 42 papers was published by Alaska Sea Grant. In 1980 the SSC suggested a meeting to provide information on herring. The meetings evolved into a series named in honor of Lowell Wakefield, who founded the Alaska king crab industry.

Wakefield recognized the two major ingredients necessary for the fishery to survive were ensuring that a quality product was available to the consumer, and that a viable fishery could be maintained only through sound management practices based on the best scientific data available. Lowell

Wakefield and Wakefield Seafoods played important roles in the development and implementation of quality control legislation, in the preparation of fishing regulations for Alaska waters, and in drafting international agreements for the high seas. Toward the end of his career, Lowell Wakefield joined the faculty of the University of Alaska as an adjunct professor of fisheries, where he influenced the early directions of the university's Sea Grant Program. Four Wakefield symposia are planned for 1999-2001.

Review of the Symposium

Reviews in Fish Biology and Fisheries 8:113-115, 1998. Reprinted by permission of the editorial team of *Reviews in Fish Biology and Fisheries.*

Conference Report

15th Lowell Wakefield Fisheries Symposium

Fishery Stock Assessment Models for the 21st Century: Combining Information from Multiple Sources.

Having worked with fish stock assessment models for more than 20 years, I often ponder the limitations of current techniques for resolving the real problems of fishery management. Frequent reports of declining fish stocks do not inspire confidence. Obviously mindful of these problems, colleagues in Alaska organized the 15th Lowell Wakefield Fisheries Symposium on the topic: "Fishery stock assessment models for the 21st century: combining information from multiple sources." During four sunny October days in Anchorage, 167 participants from 19 countries debated the merits of a wide spectrum of assessment techniques.

The Lowell Wakefield symposia have a distinguished history of gathering fishery expertise on specialized topics. Proceedings have been published for each meeting, and the 1992 symposium on "Management strategies for exploited fish populations" went through a second printing to meet the demand. Brenda Baxter, at the University of Alaska Fairbanks, deserves particular credit for coordinating this run of 15 successful meetings. She explained to me that her approach stems from an innate Alaskan sense of hospitality. Living in splendid isolation among mountains and glaciers, Alaskans seek perspective by inviting outsiders to visit and share their knowledge. She's right. The Alaskans make great hosts, and the isolation stimulates an intense scientific exchange.

This symposium was organized into four theme sessions: data conflicts and model specification, stage structured populations, ocean ecosystems, and harvest policy. Talks never ran concurrently. Thus, the opportunity to hear them all was limited only by a participant's mental energy and preoccupation with interesting hallway conversations. Central issues came up repeatedly in each of the four sessions. Is stock assessment really possible; that is, do the available data really allow us to infer stock abundance and potential yield? Do ancillary data add or detract from the assessment results? In particular, can the influence of climate data be detected convincingly in modern assessment models? Recognizing that ecosystems are complex, do simple or complex models lead to better understanding? Above all, what aspects of the analysis contribute to sound, useful advice for managers and stakeholders? How can this advice be communicated in ways that draw attention to risk and the need for risk management?

The opening keynote talk by Keith Sainsbury (Hobart, Tasmania) focused particularly on management issues, challenging many of the classical approaches to assessment. For example, fishery scientists need to examine broader control options than mere quota recommendations. In a later talk, Chris Francis (Wellington, New Zealand) reviewed the concept of risk. He observed that fisheries management objectives are often conflicting. Thus, the measure of risk varies with the objective. Chris also provided summary comments at the close of the meeting, after an exhausting four days. Even then, had it not been for flight departures and other personal reasons, I'm sure that participants would have peppered him with questions. Such was the enthusiasm and intensity of the meeting.

Not surprisingly, the discussion often focused on statistical methods. Bayesian inference clearly dominates the modern analyses, in contrast with typical analyses 20 (or even 10) years ago. Furthermore, Bayes posterior sampling methods, such as the Markov Chain Monte Carlo (MCMC), have captured some people's attention. At one point in the discussion, I mentioned a very readable book on MCMC methods (Gilks et al. 1996). To my surprise, Ken Newman (University of Idaho, Moscow, Idaho) actually had a copy in hand at that moment. The book's cover, with distinct shades of red, yellow, and green, could easily be recognized across the room. This small example illustrates the common thread of new analytical techniques being pursued independently by fishery scientists worldwide. On the other hand, Saul Saila (University of Rhode Island, Narragansett, Rhode Island) demonstrated his usual flair for complete originality by presenting analyses based on fuzzy logic and fuzzy regressions. Who can say what statistical methods will dominate analyses 20 years in the future?

I take a particular interest in statistical theory, and an encounter at the meeting shed some light on its practical consequences. During a talk I presented, I cited an example in which frequentist and Bayes approaches give different interpretations of the risk of a low stock biomass. Later, at an evening gathering for wine and pizza, Kevin Sullivan (Wellington, New Zealand) pointed out that this seeming anomaly actually motivates two schools of thought in managing New Zealand fisheries.

Computer software came up repeatedly in the talks. Several speakers reported positively on the use of AD Model Builder, a generic model development tool available from Otter Research Ltd. The author of this software, David Fournier (Nanaimo, British Columbia), attended the meeting and appeared as coauthor on three of the presentations. Other speakers had used Stock Synthesis, a software package designed explicitly for catch-age analyses with auxiliary data. Richard Methot, the author of Stock Synthesis, also attended the meeting and appeared as coauthor on one of the talks. Still other speakers preferred relatively simple analyses that could be implemented in spreadsheets. It wasn't difficult to get involved in passionate discussions about software preferences, which correlate to some extent with views about model complexity.

Numerous talks described modern complex models designed to capture the information from multiple data sources. Examples included: (1) eastern Bering Sea walleye pollock data linked with oceanographic data, (2) nine data sources for California bocaccio, in which the age composition data are ultimately rejected, and (3) British Columbia sablefish catch-age data linked with tag recovery data. One speaker (David Sampson, Newport, Oregon) conducted an experiment on the Stock Synthesis model itself, using simulated data to test the model's sensitivity to various sources of error.

Not everyone, however, proposed models with high complexity. For example, Daniel Pauly (University of British Columbia, Vancouver) discussed a mass balance approach to food web ecosystems. Based on data from various species, the mass balance concept implies a relatively straightforward calculation of biomass requirements to sustain the web. Participants were certainly sympathetic to "back of the envelope" calculations. In fact, a few fishermen attended the meeting and made serious efforts to follow the bookkeeping in complex catch-age models.

In this brief report, I've only sampled topics that were discussed. Interested readers can consult the full published proceedings planned for the fall of 1998. Perhaps a final story can portray the candid spirit that characterized the meeting. Patrick Sullivan (Seattle, Washington) gave a thought-provoking talk describing recent changes in the model for Pacific halibut. These have resulted in substantially revised estimates of stock status. His analysis illustrated the variety of interpretations possible for most fishery data. He closed with a quotation from Edmund Burke, "Nobody makes a greater mistake than he who does nothing because he could only do a little." Not surprisingly, I found participants ready to debate even that point.

References

Gilks, W.R., S. Richardson, and D.J. Spiegelhalter. (1996) Markov Chain Monte Carlo in Practice. London, UK, Chapman & Hall, 486 pp.

—Jon Schnute
Pacific Biological Station, Nanaimo, British Columbia

Living Marine Resource Assessment for the 21st Century: What Will Be Needed and How Will It Be Provided?

Keith Sainsbury
CSIRO (Commonwealth Scientific and Industrial Research Organisation)
Division of Marine Research, Hobart, Australia

Introduction

It is not often that one gets the opportunity to look broadly at the likely circumstances and needs of the next generation of fishery assessment practitioners, and to consolidate the lessons of our experience. The twenty-first century is only three years away, but in this talk I will look mainly to the circumstances and needs of about 50 years hence in the mid-twenty-first century. By that time even the youngest of the present assessment practitioners are unlikely to be still practicing, but we will have left our mark—a legacy of assessment methods and assessment approaches, and the health of the world's living marine resources.

Here I will be drawing a long and speculative bow, with the aim of providing food for thought going into this symposium [Fishery Stock Assessment Models for the 21st Century, Anchorage, Alaska, Oct. 8-11, 1997]. It is a personal view, recognizing that this is just one view among many and that the track record of such views is not good. Thomas Huxley (in Smith 1994), for example, stated in 1883 that "…all the great sea-fisheries are inexhaustible; that is to say nothing we can do seriously affects the number of fish." Huxley was right for about thirty years, but this view did not prepare the fishing industry, fishery managers, or fishery scientists well for what was to follow.

In this paper I will first look at the circumstances and expectations of marine living resource assessments in the mid-twenty-first century, the issues that are likely to be important, and their effects on the assessment themselves. This is a mixture of some things that I hope will happen and some things that I fear will happen. I will then look at whether there will have to be a paradigm shift from present approaches, based on a brief examination of how well present approaches have performed. And finally I will address the question: If there is to be a paradigm shift, what will it be toward? In answering this question I will identify the major developments needed to address the likely circumstances and expectations of assessments in the mid-twenty-first century, and will illustrate developing approaches in present assessments.

Circumstances and Expectations of Assessments in the Mid-Twenty-First Century

More change in the external circumstances affecting assessments can be expected in the next 50 years than occurred in the last 50 years. Expected key changes in circumstances are

Globally an increased requirement for food and other environmental support for human life and activities. This will be driven both by increased human population and increased aggregate economic affluence. It is also likely that there will be an increased disparity in the distribution of wealth both within and between countries. The economically affluent groups will create a strong market demand for seafood products and for marine environmental use (e.g., recreation, coastal urban development, and effluent disposal). The less advantaged groups will variously make demands for subsistence food; increasingly this will involve "resource mining" for short-term advantage, illegal fishing, and use of black markets. The two groups will compete strongly for fish and environmental use, and activities such as resource mining and black marketing will be indirectly driven by the affluent groups. The increased economic power of the affluent groups, combined with increased desperation among the less affluent, will greatly increase the difficulty of implementing management measures for some national and most internationally shared marine living resources.

Increased use of the marine environment and the effects of non-fisheries activities. This includes increased coastal zone development and urbanization, increased nutrient and sediment loads resulting from catchment land use patterns, and increased waste and effluent disposal to the sea. In addition the marine system will be significantly used by industries such as tourism and recreation, aquaculture, oil and gas production, and minerals production. This increased use will generate competition between alternate users for space and resources, will increase levels of nutrients, toxins and heavy metals, and will extensively modify some marine habitats—with these effects being particularly strong in coastal and continental shelf regions. These changes will significantly affect fisheries production.

Increased role of business interests in political structures and in achieving sustainability. There will be a much stronger use of property rights in the management of fisheries (and other marine-based industries), with rights including access to space as well and to the marine living resources themselves. A range of industries, not just the commercial fishing industry, and some non-commercial interest groups such as recreational users, indigenous peoples and conservation interests, will have such property rights. There will be an increase in the influence and participation of business and companies in research, management decision-making, and implementation of management measures. With this increase will come a close link between assessment of performance at the industry level and management decisions. Performance accreditation and standards will be commonly used by both industry and management groups. Product "ecolabeling" and accreditation regimes such as the International Standards Organisation (ISO)14000 will be common and powerful tools for marketing and achieving sustainability in commercial fishing, through their ability to inform affluent consumers about the sustainability of the industry (i.e., its biological production base, ecological effects, and the management process).

Greatly increased industry access to the marine system. Increased technology has arguably been the major factor driving fishery impacts and resource assessments in the last 50 years. This will remain a major driver of change in the next 50 years. Greatly improved sensors, deployed everywhere from vessels to satellites, will provide extensive data on the ocean, seabed habitats, biological productivity, and the spatial distributions of fishery target species. Real-time analysis and assimilation of these data will routinely be used by commercial fishing vessels. Capture technology will improve. Product processing technology will also improve, to the point where commercial products can be made from most species. The combined effect will be extremely effective and flexible targeting of fishing effort, plus the ability to economically harvest both further down the food chain and on highly dispersed resources (e.g., small to medium sized oceanic fish and squid).

Greatly increased capacity to monitor the marine system. The same technological developments used to improve fishery targeting will also revolutionize fishery and environmental monitoring. Developments will include use of automated vessel tracking and catch monitoring methods, extensive electronic tagging experiments, and direct observations of the marine system by remote sensing, vessels of opportunity, and moored sensor arrays. Very diverse and extensive data sets on the marine system and human activities will be available for resource assessments.

Recognition that the marine system does not follow a stationary process. Systematic change and quasi-periodic fluctuations in global climate and oceanography will occur and be recognized during the next 50 years. These changes will greatly affect oceanographic patterns of spatial interconnection and interannual variability, and consequently the spatial and temporal patterns of biological production. These changes will have significant effects on regional fishery resources and industries. In addition there will be significant recoveries of some top predator populations, including whales, seals, some sharks, and seabirds. The large tunas could be in this category of significantly recovered populations, but I doubt that the necessary catch controls and reductions will be made in time to give that level of recovery by the mid-twenty-first century.

With these changes in circumstances will come several major changes in the social and political expectations of the fishery assessment and management process. Expected changes in expectations include

A demand for greater social responsibility, and in particular a demand for greater equity in decision-making processes and a lower risk to food supply, jobs, and generation of economic wealth.

Increased concern for maintenance of biodiversity. This will be motivated both by increased recognition of the intrinsic value of biodiversity and by use of biodiversity as an indicator of ecosystem health and resilience, and the sustainability of fishery production. Some fish species will join the ranks of the "charismatic megafauna" and so become the subject of significant public interest and concern in their own right.

Increased range of interest groups and users in assessment and decision-making processes. This will come from increased interaction among different uses of the marine environment and more frequent use of property rights. The impact of decisions made in one use sector on the rights of other user sectors will result in an expectation of broader equity and participation in the fishery management decision-making process. The interest groups and sectors that are likely to be involved are commercial fisheries, recreational use, indigenous subsistence use, aquaculture, tourism, mineral extraction, oil and gas production, conservation groups, animal welfare groups, and governments.

The first two of these changed expectations will support adoption of precautionary approaches in management decision-making, increased consideration of species that are dependent or associated with the harvested species, and decreased acceptance of high bycatch and discard rates in fisheries. Unfortunately these changed expectations of fishery assessment and management will come the hard way—as a result of repeated fishery failures, large and small, contrasted with a few successes—rather than from foresight and inspired leadership.

These changed expectations will have a major effect on the focus of scientific fishery assessment in the mid-twenty-first century. The assessments will have a strong emphasis on delivery of sustainable economic and social benefits through maintenance of ecosystem integrity. They will need to evaluate use of spatially zoned property rights for environmental access and resource use, including marine protected areas or reserves. They will increasingly be conducted through integrated regional management structures and arrangements, including arrangements that integrate domestic with international management and arrangements that integrate watershed with coastal marine ecosystem management.

These changed expectations will cause some major changes to the scientific advisory groups within the management structures. The scientific advisory groups will contain very diverse interest-stakeholder group representation, will be under increased political scrutiny and pressure, and will provide assessments that are very contentious and contestable. They will use large and diverse data inputs, and will assess a wider range of environmental impacts than just fishery impacts.

The changed management circumstances and expectations will also cause major changes in the analyses undertaken. Expected key changes in analyses will be

Stationarity or constancy will not be an acceptable default assumption in population and observation process models. This will apply to processes such as growth, productivity, recruitment, natural mortality, population spatial dynamics, fishing efficiency and fishing selectivity.

Commercial catch per unit of fishing effort will not be accepted as a default simple index of abundance that is proportional to stock size. Catch per unit effort will still be used in assessments, but default interpretations will be based on a range of possible relationships between fish abundance and catch per unit effort.

Assessments will be spatially explicit so as to deal with the increased use of spatial zoning and management controls. Spatial signals and statistical contrasts in the observational data will become important in model parameter estimation.

Data sets available for analysis will contain orders of magnitude more observations.

Ecosystem, food-chain, and habitat effects will be frequently included in the assessment models.

Precautionary management will require the identification and use of precautionary interpretations of the available data, based on hypotheses that may not be specifically identifiable or justifiable in the available data.

The combination of these changes will mean that assessment models will be considerably more complex, and many alternative hypotheses about system dynamics and the observation process will have to be carried into the assessments.

Will There Have to Be a Paradigm Shift?

Given the likely major changes in the circumstances and expectations of assessments of marine living resources, the question arises as to how well current assessment approaches would deal with them and whether there will need to be a shift in the assessment paradigm.

The paradigm of the last about half century has been based on six major elements:

1. A predictive approach, based on simple models of biological processes and a simple interpretation of objectives such as maximum sustainable yield.
2. An adaptive approach in which assessments are updated regularly as additional data becomes available. The structure of this adaptive system is ad hoc, however, in that no specific attention is usually given to the design and performance of the adaptive feedback loop that is being implicitly used in management decision-making.
3. Assessments are highly spatially aggregated, to the extent that space is usually not explicitly included in analyses other than via a broad statement indicating the area occupied by the stock and across which statistics are aggregated.
4. Assessments are focused on single species, and population dynamics are internally driven (i.e., population abundance or density is the principal control variable, with external ecosystem and environmental effects being ignored or internalized through heuristic functions for natural mortality and recruitment).
5. Assessments are focused on a single use of marine living resources (i.e., commercial fisheries).
6. Assessments assume stationarity of many dynamical processes (i.e., that the parameters of the dynamical process models are assumed to not vary through time).

It is clear that changes in the approach to items 3-6 above will be required in resource assessments to meet the circumstances and expectations of assessments in the mid-twenty-first century. Items 1 and 2 above sound reasonable—they describe in principle a predictive and adaptive assessment scheme—but how well has this actually performed?

Much has been written lately about the poor performance of fisheries management, and I do not intend to belabor the point here. However, the judgment, based on management outcomes, is that the fishery assessment and management process is arguably unsuccessful or at best weakly successful. There is considerable evidence for this.

For example:

- Fishery status reports by country and internationally. About 40% of U.S. fisheries are overexploited or recovering. Almost all ICES managed fisheries are being harvested well above their maximum sustainable levels. About 20% of Australian fisheries are overexploited, and Australian fisheries have a relatively short history of commercial exploitation. As a whole the global catch is at or beyond maximum sustainable yield (Garcia and Newton 1997).

- Ludwig et al. (1993) question whether sustainable use, including the conclusions of the Bruntland Report and the UNCED (WCED 1989) which resulted in the widespread acceptance of sustainable development as an achievable goal, can be achieved even in principle, and question our ability to reliably estimate sustainable harvest levels. Like Rosenberg et al. (1993), I agree with their conclusions about the difficulty in estimating sustainable harvest levels, but disagree with their criticism of the Brutland Report and their conclusion that sustainable resource use is unachievable in principle. There is a level of living resource use that is sustainable—it may not be a stable or constant level of use as has often been assumed in the past, it may not be sensible to try to maximize our use and extract the absolute most from the resource that is possible, and we may have difficulty estimating what a sustainable harvest level is at any point in time—but a sustainable level of resource use does exist.

- Holmes (1994), Hutchings and Myers (1995), Walters and Maguire (1996), and several others have examined the causes of the recent collapse of the northwestern Atlantic cod stocks, and while several factors were involved it is clear that incorrect stock assessment advice was at least partly responsible.

- Mangel et al. (1996) provided a detailed review of the performance and principles of management of wild living resources generally, and revealed widespread failure to achieve sustainable use. In most cases the main difficulties are the demands of human population and economic expansion that require permanent changes to populations, habitats, and ecosystems. For example the clearing of forests for agriculture, in which the forestry harvest supports a transitional industry. In this example the aim is alteration rather than sustainable use. However, they also cite many cases where sustainable use was intended but not achieved.

- Johannes (1998) and Munro (1987) examined the performance of sustainable use of tropical inshore fisheries, particularly in Southeast Asia and the Pacific Island nations. They concluded that performance is dismal. Johannes goes on to conclude that a major contribution to this poor performance is resource managers wait-ing for quantitative resource analysis to underpin their decisions—the kind of assessments expounded in text books and development courses. Johannes calls for a stop to the wait for quantitative resource analysis for these fisheries. He points out that such analyses cannot be expected because of the complexity of the resources, the complexity of the usage patterns, and the paucity of the monitoring and analytic capacity that is necessary for the scientific assessment-based approach to be successful. He recommends instead an immediate start on the application of management that is strongly precautionary and that can succeed in data-free (although not information-free) situations. This call is essentially for management that is robust to the constraints on scientific observation, scientific analysis, and management capacity in a particular situation. A similar conclusion was reached by Isaac et al. 1998 (this symposium) from consideration of South American fisheries assessment and management.

- Mace (1996) provided an excellent review of the performance of the world's fisheries and their management, and found that there is a general failure to achieve sustainable harvest levels for many individual fisheries. She pointed particularly to the gross overcapitalization of the world's catching capacity, achieved through extensive government subsidy, as providing a major impediment to implementing sustainable harvest levels. However, she also recognized that incorrect scientific assessments of the productive capacity of fish resources have contributed significantly to the failure to achieve sustainability.

- Perhaps the most significant evidence for failure of the recent approaches to assessment and management of marine living resources, and the one that will move the social and political expectations for improvement, is the increased frequency of negative public media articles, documentaries, and reports. They describe an irresponsible fishing industry that is managed inadequately on the basis of insufficient understanding. This "bad press" is a worldwide phenomenon, and the present approaches to assessment (and also the present approaches to management and industry decision-making) will not be sustainable for very long under such exposure.

However, the picture is not all doom and gloom. There are successes. There are fisheries such as the western Australian rock lobster, the Falkland Islands squid, the Pacific halibut, and the North Pacific walleye pollock which have been sustained for long periods in comparison to the lifetime of the target species. A significant fraction of the world's fisheries are reported to be at or near their sustainable levels (e.g., 20% for the U.S.; 30% for Australia). And there have been stock recoveries—for example some whales, seals, and U.S. Atlantic striped bass—although it is notable that very major changes in the harvest levels have been necessary to achieve these recoveries.

Furthermore, it is also clear that often it is not the scientific assessments that are at fault. Management decision-making processes are often dominated by short-term considerations, which in effect favors development over conservation goals. Management decision-making also often shows a remarkable asymmetry in the vigor of response to good and bad scientific news. A scientific recommendation to increase catches typically receives little scientific scrutiny and is rapidly enacted, while a recommendation to reduce catches is likely to be scientifically reviewed or judged in need of further research support. It may be acted upon some years later. Two good examples of this delay in management action can be provided from Australia.

The first concerns the southern bluefin tuna. In the late 1970s Murphy (1977) recommended an increase in the Australian catch of juvenile tuna on the basis of a yield-per-recruit argument that assumed recruitment would not be affected by fishing. Within two years the industry had expanded its catch and capacity by more than an order of magnitude. Murphy then called for restraint, raising the prospect that catches at that level could affect recruitment. It took about ten years to reduce the catches again, and to this day many in the fishing industry cite incorrect scientific advice as the cause of the Australian contribution to overfishing of southern bluefin tuna.

The second example concerns the Australian southern shark fishery. In 1989 scientific agreement within the stock assessment process was reached that catch levels were unsustainable and needed to be reduced. Catch reductions were resisted by industry and the management agency agreed to an external scientific review to be conducted by an industry consultant (Prof. Carl Walters). The review disagreed with the original findings and recommended further research to clarify matters. After seven years of additional research, at a cost of about A$250,000 per year, it has recently been concluded that catch cuts of almost exactly the magnitude originally recommended are needed. These catch cuts are in the process of being implemented. The main, and critical, difference between the present and past assessments is that now there is much greater support for the decision among fishing industry participants.

To return to the question “will there need to be a paradigm shift in living marine resource assessment?” The answer is yes. Under continued use of the current paradigm there will be too many fishery failures and too many indications of broad deterioration in resources, ecosystems, and social benefits to maintain the social and political (let alone scientific) support for this approach. The scientific assessments will not be solely responsible for the failures. However they will be in part responsible for the failures, and will be recognized as being inadequate to the task. As mentioned earlier, the assessment task will get more complex as issues of food-chain dynamics, spatial dynamics, multiple use interactions, and non-stationarity need to be addressed. It is debatable whether or not improved data and technology will arrive in time to support improved scientific assessments to meet these additional demands, but it seems to me that the additional demands will arrive first. And these additional demands will make the inadequacy of the present approach even more obvious than it is now.

So there will be development of a new paradigm. But before looking at what the new paradigm might contain, there is a need to look more carefully at why we are having such difficulty with the old one. What are the lessons and what needs to be changed?

What Are the Fishery Assessment Lessons of the Late 20th Century?

I think there are four main scientific lessons, or reasons for failure, in the recent assessment paradigm.

1. The assessment models do not adequately represent the real world.
2. The inability to correctly identify models and to accurately estimate parameters.
3. The uncertainties in assessments are not well recognized or treated by science or management.
4. Assessments are not evaluated and designed in the context of their use in the management system.

1. The Assessment Models Do Not Adequately Represent the Real World

It is widely recognized by assessment practitioners that the models used are simplifications of the real world, yet there is little emphasis in stock assessments and stock projections on the resultant effects on the management advice given. There is usually very little explicit recognition or analysis based on a broader range of more complex dynamics, or even analysis based on variants of the simple model structures commonly used. Sensitivity tests and confidence intervals are routinely calculated, but these are highly constrained by the simplified underlying assessment model and do not represent the uncertainty due to model specification. For example a retrospective analysis shows that the actual changes in southern bluefin tuna abundance over the last ten years has been consistently outside the range of predictions made on the basis of simple models after as little as two years after the prediction is made (Klaer et al. 1996). The common result is that the assessment models used cannot capture many of the very interactions and complexities that have caused assessment failures in the past. Usually this weakness due to use of oversimplified models is not even “hedged” by examination of a range of alternative simple models.

In addition to these ecological process simplifications, there are two important aspects of the fishery management system that are treated inadequately in assessments and predictions. The first of these is the process by which observations (data) are collected, interpreted via intermediate analyses, and reported. This is a complex process in itself, much moreso than the simple statistical treatment provided by most assessments, which can greatly affect assessment results. The fact that this symposium has the treatment of conflicting data sources as a major subtitle is simply an indication that we often do not view and model the observation process properly. There is only one reality and if we have conflicting signals from it then we have not correctly specified the observation process model. The second is implementation uncertainty. The difficulty in implementing management decisions and controls as intended is legend in fisheries management, and has been a major contributor to stock collapses; and considerable scientific effort goes into trying to correct fishery data from implementation failures of the past. Yet projections of future population size in stock assessments usually assume perfect management implementation and very rarely include management implementation as a source of uncertainty.

2. The Inability to Correctly Identify Models and to Accurately Estimate Parameters

This point could almost be two separate points, one focusing on the difficulty of learning about the dynamics of nonlinear systems generally and the other focused on the data available from fisheries in particular. Both are major issues in marine living resource assessment, and they conspire together to make scientific assessment of fisheries very difficult.

First it is widely recognized that ecological systems are highly nonlinear systems, and that linear approximations do not perform well. Nonlinear in this sense means that the dynamics of a population of interest cannot be reasonably approximated by a linear function of its own abundance, the abundance of other observable populations, or environmental conditions. This nonlinearity sets some very strong and unpleasant constraints on stock assessments. In particular the results obtained in one circumstance do not generalize well—the observations made at one place, at one population size, under one set of environmental conditions, and under one composition of the surrounding ecological community are not easily generalized to accurately predict the effects of a change in any of these factors. And in fisheries assessment there is usually a lack of the controls and replicates that would provide the usual scientific hope for dealing with such circumstances. The result is that, to a major extent, each fishery is tackled as a separate and new learning experience. Consequently, for each new species or population encountered, empirical observations are needed across a range of population sizes (and environmental conditions and community compositions) to characterize the dynamics of that species or population. In practice the state of the environment and community composition are usually very poorly observed while observations of the target population are being made, which may well be a major cause of the non-stationarity seen in simple population models and the poor success of generalizations.

The extent to which nonlinear processes and interactions influence living marine resource assessment is yet to be fully determined. It is clear that they provide a major influence and constraint on generalizing results, but at a coarse level there are some features that can be extracted from the collective experience. For example the basic population parameters, food consumption rates, and general ecological role can be coarsely predicted from species taxonomy, morphology, and geographic location. There is also some indication that the same may be true in a broad sense for stock-recruitment parameters (e.g., Myers et al. 1995). These "meta-analyses" are very useful and address the real nature of the scientific understanding that experts bring to a particular problem. But so far they indicate that, while very broad patterns may be generalizable, the predictions made are not sufficiently precise to be used as the basis of a resource assessment. We are left to a large extent with the curse of the nonlinear system, and the resulting need to establish and track events locally. Ironically it seems that aspects of management system may be more reasonably generalized than those of the biological system. It may transpire that the structure and parameters of management feedback control procedures generalize better than do the structure and parameters of the nonlinear systems they are designed to control.

The second major issue is the data available from which to infer the state of the ecological system. Marine living resource assessment is essentially an observation-limited activity at present, with the population abundance and dynamical processes being close to unobservable. In most cases nothing is measured either accurately or precisely.

In most fishery assessments, data from the fishery operations themselves (e.g., the catch, fishing effort, fishery discards, and the nature and location of fishing operations) is the major source of information about stock size. Unfortunately fisheries usually do not provide accurate and precise data on their own operations. For some fisheries there are also independent surveys of resource abundance, and there is no doubt that the availability of such surveys greatly improves the rigor and confidence in resource assessments. However these surveys are often greatly limited in their coverage by cost (usually survey effort is many orders of magnitude lower than fishery effort) and the competency of survey fishers to catch fish can be questioned. And so overall survey precision is low, bias in catch results remains an issue, hypotheses relating to potentially important small-scale spatial-temporal events in which survey and industry data contradict cannot be resolved, and the fishery catch data remains the most abundant source of data.

As a result considerable scientific effort goes into trying to correct fishery and survey data for bias—and in attempting to draw conclusions, despite the high level of measurement error. Inability to correct biased and imprecise data from the fishery has been a significant factor in some assessment failures. The quality of information from a fishery is potentially under management control. However exercising this control by top-down control methods is very costly, and attempts to improve the data quality by education and collaboration with the fishing industry can be expected to take many years of effort (and probably a generation or so in some cases) to bring results.

To make matters worse the information content, with respect to statistical properties of interest, of the data available from fisheries is also low. This is partly a reflection of the low precision and accuracy of measurements, but is also a result of fisheries not providing a good experimental design with respect to the properties of interest. For example, a fishery operates so as to generate strong correlation between fishing effort and the abundance of the target and associated species in the ecosystem. Consequently there is often low statistical contrast in the data across the supposed independent variables of interest in an analysis, and so the data available from fisheries is often not informative for the identification of cause and effect or for the estimation of key parameters in process models.

It is becoming widely recognized that the information required for scientific assessment of marine living resources is considerably greater than that currently being provided. Interestingly this is being recognized simultaneously and separately in both developed and developing countries. The reality is that in both situations the dynamics of the marine resources are very weakly observable, and the accuracy and precision of resource assessments are very weak and inadequate as a result. A corollary of this is that only large assessment and management errors are detectable in both developed and developing nations, and even large errors are usually detected only when they start to cause serious social or ecological disruption. If this global generality is to be changed then the quality of information available for scientific assessment of marine resources will need to be improved, and improved in a specifically targeted and intelligent manner.

3. The Uncertainties in Assessments Are Not Well Recognized or Treated by Science or Management

To date the call for management to be based on the "best scientific advice," a phrase repeated in many policy and international convention documents, has been interpreted as meaning a consensus point estimate or "base case" analysis decided by the scientists involved. Under this approach uncertainty is not well reflected in the scientific advice provided to management, if is reflected at all, despite there usually being a considerable range of scientifically valid interpretations that could be drawn from the data. Typically a huge number of decisions are made by the assessment practitioners in the process of selecting the point estimate or base case to consider or present. Often these decisions are essentially arbitrary restrictions on model structure, parameter estimates, or the domain and form of Bayesian prior distributions. The application of age-structured assessment models involves scores of such decisions. In the majority of cases these restrictions are made without fully examining their consequences or documenting the selections made. The result is that uncertainties in the assessments, and the consequences of those uncertainties, are poorly understood both by those providing and receiving the scientific management advice.

Recently there has been increased effort to provide a more objective and transparent treatment of the uncertainties in the stock assessment advice. This has caused difficulties in both the scientific and management arenas, as scientists try to determine what information to provide and managers try to develop appropriate ways of responding to the strange world of scientifically analyzed uncertainty. This is an area that I expect will provide developments with great impact on the science of living resource management in the next few decades.

The response of science to this challenge has been mixed to date. In some cases the uncertainties continue to be ignored and point estimates or base case analyses are provided (perhaps with a few variants to meet questions of sensitivity). In other cases distributions of predictions are provided, leading to risk analysis and decision tables which aim to provide probability of achieving outcomes of management interest given certain management decisions. Risk analysis and decision tables are provided in the hope that they clearly outline the consequences of different management actions across the uncertainties in the assessment, so that the managers can select the balance of risk across the (inevitably conflicting) management objectives. As discussed below the experience with this approach has generally been disappointing.

An increasingly recognized point is that scientific assessment analyses and management advice cannot be totally objective—scientific assessments are not value-free analyses even given the best intentions and effort (see Sainsbury et al. 1997b). No analysis can consider all possible interpretations, no matter how good the software and how dedicated the analyst, and selection of the hypotheses to include involves some level of subjective selection. Within many assessments considerable effort is put into formulating the uncertainties so that they are as accessible as possible to statistical and other observation-based methods of analysis. But this cannot deal with all issues and alternative interpretations, and there remain subjective judgments about the interpretations to include and favor (i.e., give greater or lesser weight) in an analysis.

Subjectivity in a scientific process is not surprising, and this is well recognized in the general philosophy of science (e.g., Chalmers 1976). However, it carries with it several important implications in the scientific assessment process that have not been well recognized or acted upon by scientists in their search for the single best answer to provide to resource managers. The most obvious implication is that the receivers of the scientific assessment advice should be made aware of the subjective judgments made by scientists in conducting the assessment. Less obvious are a number of fundamental implications that arise from the use of the scientific advice in a decision context. These implications:

- Blur the separation usually sought between the scientific and management input to decision-making.

- Challenge the common use of Occam's razor in scientific selection of hypotheses (i.e., acceptance of the simplest, most parsimonious model that can explain the data).

- Call into question the scientific focus on type I error as the most appropriate criterion for accepting or rejecting a hypothesis. (Type I error is the probability of rejecting a true hypothesis and type II error is the probability of accepting a false hypothesis.)

In a decision context it can be very important to carry a hypothesis through the assessment analysis and examine its management implications—even if that hypothesis has a low probability of being correct. A hypothesis that is rejected as being improbable on the basis of the usual scientific criteria could and should remain in a decision analysis if it has particularly important management implications. And the judgment to include or exclude the hypothesis from analysis is made on a combination of the probability of the hypotheses being correct (the stuff of science) and the importance of the management implications (the stuff of management utilities). The scientific and management considerations are intimately connected. They cannot be treated separately if scientific uncertainty in assessments is to be treated seriously in a decision-making context, and any attempt to separate them is flawed in the logic of decision theory.

With this, two major difficulties come into play. The first is that the management utilities that determine the importance of the management implications, and hence influence which hypotheses should be included in the analysis, are not readily accessible to description and analysis. These utilities are complex, mostly qualitative, change with economic and political circumstances, and vary across individuals. The second difficulty is that potentially important hypotheses may not be identifiable in the data. This difficulty will increase as the non-stationarity of ecological system dynamics is increasingly accepted, and management increasingly accepts the need for precautionary approaches to confront uncertainty. But this will raise difficult questions for assessment analysts. Should the analysis include hypotheses with no data support, or hypotheses that are not observable even in principle, but that have dramatic management implications? Should the analysis include complex elaboration of hypotheses, with potentially important management implications, when simple interpretations are sufficient to explain the available data? Some scientists would be inclined to conclude that an analysis using such hypotheses and hypothesis selection methods is not a scientific one. But such selection methods are necessary to reflect uncertainty in the scientific assessment and to evaluate the management consequences of the hypothesis selections made by scientists. The scientific expedient of Occam's razor is not appropriate in a decision context. Put in Bayesian terms, Occam's razor amputates the Bayesian priors carried into an analysis, and once the domain of a Bayesian prior distribution has been restricted then strictly speaking it cannot be recovered in a posterior distribution no matter what any subsequent data indicate. Zero values in a Bayesian prior distribution should only be used if there is sufficient information to conclude that the relevant hypothesis is *impossible* and can be completely dismissed from all future consideration. This would require a very high level of scientific proof, which is lacking in most resource assessments. Yet individual resource assessments usually contain many such arbitrary restrictions on the domain of the hypotheses considered—usually made on the basis of the expertise of the analyst and with little consideration of the management implications of the restrictions made. Type II error is highly relevant in the context of decision analysis, but it is not the main focus of research science or resource assessment.

Application of Occam's razor and the focus on type I error in scientific hypothesis testing brings a very strong and largely unrecognized conditioning into the scientific assessments. But if these tools cannot be used, then where does one stop? If hypotheses can be constructed, elaborated, and carried into the analysis without data support, and if the criteria for inclusion of hypotheses extend to the inaccessible and varying utilities of managers, then how does the analysis avoid becoming arbitrary? The scientific assessment approaches of the past 50 years have been very mixed in their response to these issues. There is an indecision apparent between embracing uncertainty and its implications as fully and objectively as possible, on the one hand, through to using informal scientific discussion and personal judgments to select a single base case or central point estimate on the other. Current approaches are not satisfying in either a scientific or management decision context.

The response of managers to scientific uncertainty in the assessments has been similarly mixed, but three general responses are apparent.

1. Under one management response the uncertainty essentially has no effect on management decision-making, with only the central point estimates or base case being used. This approach relies on the skill of the assessment scientists to come up with advice that will achieve the

desired balance of management objectives given the analysis available—with that balance being essentially decided among scientists in the selection of what is reported to managers. This is the "ignorance is bliss" approach by managers to decision-making under uncertainty (although more formally and impressively described as the certainty equivalent approach). The ignorance is indeed bliss, until something goes wrong—at which time the ignorance is recognized for what it is but there is always the option of blaming the problem on the (by then) demonstrably incorrect selections made within the scientific assessment process. However, the underlying cause of the failure is a management process that does not use, or even actively filters out, uncertainty in the scientific assessment.

2. A different management response is for uncertainty to trigger a conservative and precautionary approach to decision-making, in which the risk of error due to uncertainty is borne primarily by constraints on the resource user groups rather than by the resource. Catch is forgone because of the uncertain effects of the harvest. This approach is probably most fully developed by the processes adopted by the Commission for the Conservation of Antarctic Marine Living Resources (CCAMLR); it is a weak developed element in the deliberations of many fisheries management bodies, but it appears to be increasingly considered. For example it is a key element in precautionary and responsible fishing as recommended by FAO (FAO 1995a,b), and the North Pacific Fishery Management Council has very strong limits on total allowable catches that relate to generic considerations of uncertainty.

3. Another, very common, management response to uncertainty in resource assessments is to regard the science as providing broad bounds of possible outcomes, with the management decisions then being negotiated on other grounds. The outcomes selected by managers may or may not be within the bounds provided by the assessment. Generally, however, outcomes at the optimistic end of the assessment range will be selected, because the high level of uncertainty is interpreted as meaning that the scientific analysis provides little basis for selecting management actions across a wide range of possible actions. In the perceived absence of a scientific basis for decision other criteria dominate decisions.

The responses of scientific groups to providing management advice under uncertainty is similarly mixed, and ranges from ignoring uncertainty through to providing detailed risk and scenario assessments. However a very common response relates closely to management approach (1) described above. This view is that scientists have both a right and a responsibility to attempt to integrate the risk across competing hypotheses and uncertainties in the assessment, and that scientists are in the best position to understand and to do this integration. Under this view failure to provide strong direction in the advice to managers is seen as an abdication of responsibility.

A key point is that while scientific and management groups must deal with uncertainty in the resource assessments, and are increasingly attempting to do so, the responses of the two groups are largely disconnected. Amazingly this applies even to the scientific and management groups operating within the same fishery management arrangement. Very little attention has been paid to how the approaches developed and used by the scientists and the managers relate to one another. And some truly awful combinations have occurred; combinations in which delivery and expectations on both sides do not match, giving destructive combinations that actually prevent one group or the other from making progress in the treatment of uncertainty where otherwise the will and capacity for progress exists.

And this is a fitting introduction to the fourth lesson from assessments of the recent past.

4. Assessments Are Not Evaluated and Designed in the Context of Their Use in the Management System

The main methods that are currently used to communicate uncertainty from the assessment process to the management processes are risk analysis and decision table analysis (e.g., Hilborn and Walters 1992, FAO 1995b). These approaches typically present the consequences of different prospective management actions across a range of hypotheses about resource dynamics. The consequences are expressed in terms of the measures of management interest, such as the long-term yield or the size of the resource population. Risk assessments and decision table analyses are usually presented after each annual resource assessment, and the managers use these to select management measures for the following year.

These methods are clearly a step in the right direction, in that something of the uncertainty in the assessment is transmitted to the management process, but it does not appear to be particularly effective. There are several problems, including mutual understanding of the analysis, and the complexity of presentation for all but a very truncated range of hypotheses and management options. An important issue is the rather abstract nature of the risk and consequences presented to managers in risk analysis and decision tables. This is in stark contrast to the tangible "here and now" consequences of management decisions on political and economic interests. A 50% chance of the population falling below the nominated limit reference point (intended to reflect a biological safety level) in the next ten years may seem of less concern than the Minister on the phone expressing a view or a group of industry members at the door demanding continuation of their jobs. This combination of immediate pressures and abstract measures of risk tends to result in decision-making being dominated by short-term interests, and in population limit reference points effectively becoming targets rather than population levels to be strenuously avoided.

Risk assessment and decision table presentations appear to actually encourage short-term tradeoff and ad hoc decision-making. This is probably because these analyses emphasize that there is uncertainty in the outcomes, and so the situation may not get bad and if it does then the problem will be confronted some time in the future. But probably the main issue is that where risk analyses and decision tables are used they are used repeatedly and almost independently year after year in an ad hoc manner. A high risk catch level may be chosen in some years and a low risk catch level may be chosen in others, in a mix that depends on the pressures on the management process. But there is little understanding, among either the scientists or the managers, of the ultimate effects of this repeated ad hoc annual decision-making where the degree of risk in individual decisions varies according to fluctuating political and economic circumstances. Ironically, as they are presently used, risk assessments and decision tables appear to encourage ad hoc and short-term decision-making without improving understanding of the long-term consequences of the ad hoc management strategy that is implied.

This raises the question of whether a particular management strategy (that is the combination of monitoring, assessment and analysis, and management decision and implementation) can succeed even in principle. Based on the performance of just the assessment methods, as one component of the strategy, it is clear that we should not expect 100% success. For example the $F_{0.1}$ strategy is commonly used to provide scientific recommendations from a stock assessment, and is widely regarded as being risk averse with respect to the resource conservation. However, a review by Mace and Sissenwine (1993) found that $F_{0.1}$ was higher than the replacement fishing mortality (as measured by F_{rep}) for 12% of the stocks they examined, and so use of an $F_{0.1}$ strategy could result in declining stock abundance in a significant fraction of fisheries just because of the performance of the assessment method.

Figure 1 describes a framework of the management system that can be used to evaluate the performance of a management strategy in the context of the whole management system (see Sainsbury 1993). There are three broad components within the evaluation: the management objectives from which measurable performance criteria are derived, the biological system dynamics (coupled with the economic system dynamics, which for simplicity are not shown in Fig. 1), and the management process (involving information collection, assessment and analysis, the management decision, and the management implementation of the decision).

The whole system in Fig. 1 can be simulated to examine its likely performance, where performance is measured by the extent to which policy objectives are achieved. Uncertainty is treated through the use of one or more “deep” or operating models that reflect the uncertainties of concern in various parts of the system. For example, a range of operating

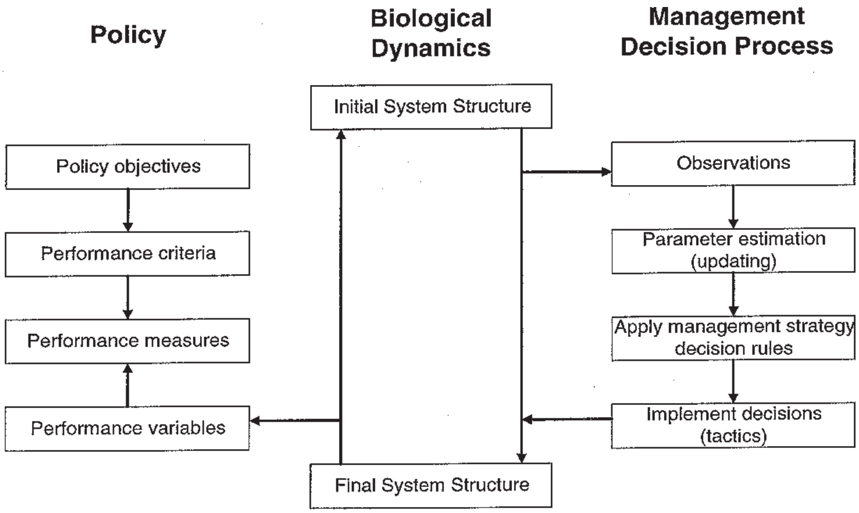


Figure 1. A general framework of the management system that can be used for evaluation of the performance of resource assessments in the context of the broader management objectives and processes (Sainsbury 1993).

models could be used that reflect different hypotheses about resource dynamics, the observation process, and the implementation process. The assessment model within the framework need not (and usually will not) be the same as the biological dynamics model; one question of interest is the performance of the assessment method across the range of stock dynamics models, in the context of the whole management processes. The simulation results are then used to give the range of management outcomes that could occur from use of any particular management strategy, and the extent to which the strategy is likely to succeed despite the uncertainties in resource dynamics and management implementation. Management strategies that robustly achieve the management objectives despite the uncertainties can be sought. The simulation results can also be used to identify hypotheses of particular management importance under the chosen management strategy, and so to provide a focus for targeted research or monitoring to determine whether these troublesome hypotheses occur in the real world.

This framework has been used by a few researchers to examine whether the management strategies being used could be reasonably expected to succeed using common assessment approaches. The results have not been encouraging. In several fisheries it was found that under quite reasonable (common) levels of uncertainty some assessments did not provide the outcomes they had been explicitly designed to give. For example for the

Australian eastern gemfish population, Smith et al. (1996) evaluated a number of common assessment and management strategies. They found that an F_{MSY} strategy results in overfishing and the population biomass being below B_{MSY} in well over 50% of simulations. They also found that a $F_{0.2}$ strategy is needed to achieve the target biomass of an $F_{0.1}$ strategy in this fishery; that is, catches had to be set according to an $F_{0.2}$ strategy so as to achieve the stock reduction levels theoretically associated with $F_{0.1}$. Similar behavior of the $F_{0.1}$ strategy has been found in simulations of other fisheries (Punt 1993). In another example Francis, Smith and Wayte (Pers. comm., NIWA, N.Z.) have evaluated an orange roughy management strategy that was intended to be risk averse. The strategy was based on a repeated annual assessment in which forward population projections under a number of catch levels are calculated, and the catch for the year is selected to give a low estimated probability of violating a limit reference point. This is an approach used in many fishery assessments. Repeated simulation of this strategy across some reasonable uncertainties gave much more frequent violation of the limit reference point than was intended and that was specifically built into the method of selecting the annual catch levels from the projection results. The desired level of risk aversion could be achieved relatively easily by use of a modified (and somewhat more complex) catch control rule. However, the important points from this example are that the bias was not intuitively apparent before the simulation testing was done and such simulation testing of fisheries management strategies as a whole is very rarely done.

The exact reasons for the poor performance of the tested management strategies vary between examples and are not totally clear in some examples. However, the failures generically relate to uncertainties in the assessment, and how these interact with the flow of data from the fishery and the incremental catch decisions applied to the fishery.

The International Whaling Commission (IWC) used an approach and framework very similar to that in Fig. 1 to evaluate its New Management Procedure (NMP, de la Mare 1996). The NMP was adopted by the IWC in the 1970s; it represented the pinnacle of the classical assessment approach and included use of a catch control rule and reference points. The assessment was based on a maximum sustainable yield (MSY) concept, and required estimation of three parameters: the current level of depletion, the MSY catch level, and the stock carrying capacity (*K* of the logistic population growth model; the population size at MSY was assumed to be 0.6 *K*). The catch control rule set the total allowable catch to zero if the population was estimated below 0.54 *K*, and to the MSY catch level if the population was estimated to be above 0.6 *K* (Fig. 2). At the time of its development there was confidence that the parameters could be estimated from available data with sufficient accuracy to allow reasonable application of the approach. However, two related difficulties quickly emerged. First it was not possible in the adversarial processes of the scientific committee of the IWC to obtain scientific consensus on a value for these parameters,

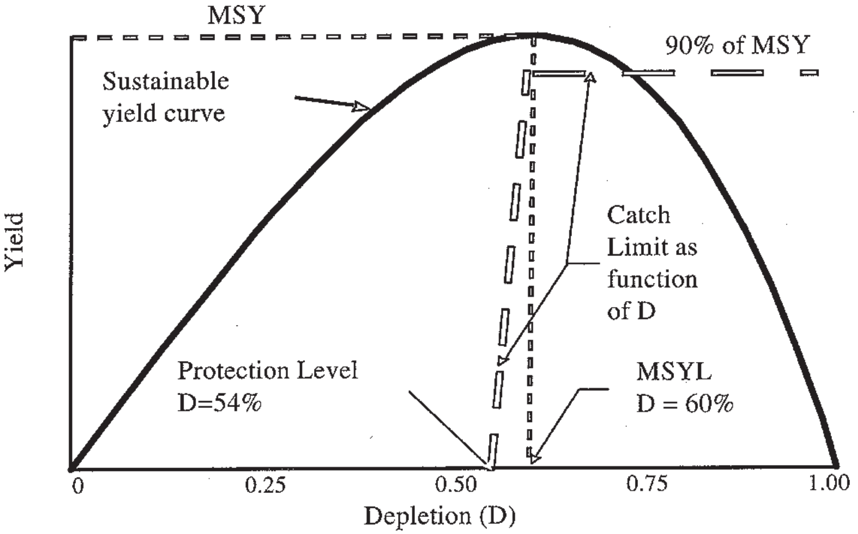


Figure 2. *The catch control rule and stock production model used by the International Whaling Commission during the 1960s and 1970s. The total allowable catch was zero if the stock was estimated to be depleted to below 0.54 of the carrying capacity, and set equal to the Maximum Sustainable Yield if the stock was estimated to be above 0.6 of the carrying capacity. This was found to perform poorly (de la Mare 1996).*

from either biological or catch history information. Second, simulation tests of the management strategy across a reasonable range of uncertainties about the data and stock dynamics showed that the strategy produced total allowable catches which varied greatly from year to year and gave a high chance of low stock size (de la Mare 1996).

The lesson from the IWC evaluation, and the other simulation tests mentioned, is that it is not sufficient simply to have reference points and catch control rules that seem reasonable. There appears to be a common misconception at present that just having defined and quantitative reference points and catch control rules is enough to ensure resource sustainability and other management objectives. It is not. There is a need to specifically evaluate the performance of the management strategy across a reasonable range of uncertainties about resource dynamics and in the context of the particular monitoring and management arrangements of the fishery. It is worth observing that the problem with the NMP was not with the much (and I think mistakenly) criticized concept of maximum sustainable yield—rather that the particular approach chosen to implement the concept was neither achievable nor robust across key uncertainties.

Overall it is very likely that commonly applied management strategies have much weaker performance, in principle, than is presently recognized. To date most simulation studies to examine the performance of management strategies have considered only a few of the uncertainties and constraints that really exist and so are likely to underestimate the extent of the problem. For example they have not fully reflected the real uncertainties in the coupled nonlinear and stochastic dynamics of biological systems in a non-stationary world, the uncertainties and asymmetric constraints on management decisions, and uncertainties and biases in implementation of management decisions. These are the issues that will have to be grappled with for success in the assessment of fisheries for the twenty-first century.

The New Paradigm?

Given the four major sources of problem with the present and recent approaches to fisheries assessment in the previous section of this paper, what new paradigm for the mid-twenty-first century could overcome them? The present problems are caused by a constellation of issues, rather than a single weakness, and the new paradigm will need to encompass all of them.

The new paradigm will integrate and be based on developments in three major areas:

1. Concepts and methods for treatment of uncertainty.
2. Modeling and statistical methods to assimilate large and diverse data sets.
3. Design of robust feedback strategies for the whole fishery management system.

1. Concepts and Methods for Treatment of Uncertainty

This is the methodology and philosophy for science in social decision-making. It is facing some critical trials at present as socially undesirable outcomes are seen, there is a demand for greater transparency in scientific input to management decisions, and the precautionary approach to management develops. Occam's razor and the focus on type I error have served science well. However, in decision-making type II error is highly relevant, but its consideration makes necessary the specification of alternatives to the null hypothesis, including the more complex hypotheses excluded by Occam's razor and hypotheses for which there are at present no pertinent data.

What hypotheses should be included in an analysis, and what are the objective criteria for their selection? To what extent should an analysis include hypotheses without specific data support, or more complex elaborations of simpler hypotheses with specific data support, that have important management implications if those hypotheses were true? How much precaution is reasonable? These are the questions that will be the focus in the new assessment paradigm.

These issues go to the limit of what is scientifically knowable. In practice the interaction between the range of hypotheses to include and their social utility is usually explored through an informal trial and error approach, in which decision makers are made aware of the implications of hypotheses (often through scientific reports and sometimes via the media) and they provide feedback on the areas of particular concern and requirements for more detail. In this way an accepted balance is established between the scientific and management groups, but it is not a balance based on an objective or scientific method.

Approaches to development of more objective approaches can be seen in Bayesian decision theory and in the Bayesian-frequentist debate in applied statistics (for example Punt and Hilborn 1997 and the papers in *Ecological Applications* 6(4) give a good coverage of the views). Bayesian methods, and especially the empirical Bayes methods, provide one way of achieving objectivity, but strictly interpreted they contain an underlying philosophy and interpretation that is not acceptable to some scientists. Others, including myself, are using the methodology to address the need without necessarily accepting all of the underlying philosophy. The Bayesian methods are capturing something of what is needed even if the basis of their derivation is not fully accepted by all. But the Bayesian-frequentist debate is engaging the issues and providing a crucible for paradigm change. A more coherent and accepted body of theory and methods can be expected to emerge. I expect a major development in this area for the new paradigm of the twenty-first century.

For the present my personal preference is to use empirical Bayesian type methods for the treatment of hypothesis credibility and updating, to document the process of hypothesis selection and weighting fully, to define prior distributions from meta-analyses that summarize previous studies as far as possible, and to ensure close interaction between scientists, managers, and resource users in the selection of alternative hypotheses. A useful hierarchical checklist to aid the selection and documentation of hypotheses (see Sainsbury et al. 1997b) is:

a. The strength of support for the hypothesis in the data from the species or fishery in question.

b. The strength of support for the hypothesis from similar species or fisheries.

c. The strength of support for the hypothesis from any species or fishery.

d. The strength of support for the hypothesis on theoretical grounds.

2. Modeling and Statistical Methods to Assimilate Large and Diverse Data Sets

This is probably the most active area of development at present, and many of the papers presented at this symposium relate to these developments. The development of computer technology is allowing analyses that were computationally prohibitive just a few years ago. Modern computers have made possible the development and application of algorithms that allow efficient estimation of large numbers of parameters (e.g., AD Model Builder, Fournier and Hampton 1996), that provide numerical estimates of complex Bayesian posteriors (e.g., Rubin 1988, Hastings 1970), that allow structured time series modeling of processes such as natural mortality and fishery selectivity, and that allow data assimilating models of physical and coupled physical-biological systems (e.g., Morrow and De Mey 1995). These developments will have a major impact on modeling for resource and environmental assessment, in that they will provide greatly improved tools for analysis of complex systems. The growing ability to model the coupling between physical and biological oceanography will be particularly important.

At present most of the development of improved modeling and statistical techniques for resource assessment is focused on the dynamics of single species. However, this focus will increasingly expand to include spatial dynamics (and oceanographic processes in particular), habitat dynamics, and food-chain dynamics. Even with advances in modeling and statistical methods the analysis of these aspects of resource dynamics are strongly limited by poorly developed theory and observational data. There will need to be significant developments in the observational technology employed in marine research and assessment before the modeling and statistical advances can be fully exploited—technological developments that can accurately and cheaply measure key features of the marine biological system. The development of these technologies will have the same impact on ecological modeling as satellite observations have had on oceanographic modeling.

Even with improved modeling and statistical methods, it is likely that resource assessment models used in decision-making in the mid-twenty-first century will be of intermediate complexity, rather than of high complexity, because of the basic limitations of the observational data. This point can be illustrated by considering the treatment of food-chain dynamics. A suggested food web structure surrounding the toothfish population at Macquarie Island is shown in Fig. 3a. It is very unlikely that models of this complexity will be used in fishery assessments and decision-making even by the mid-twenty-first century. However, simplified “minimum realistic models” of such food webs will be used in decision-making, and a few are already being used (e.g., Punt and Butterworth 1995). An important feature of these models will be treatment of uncertainty in the structure and linkages in the simplified food webs. So for example Fig. 3b gives a

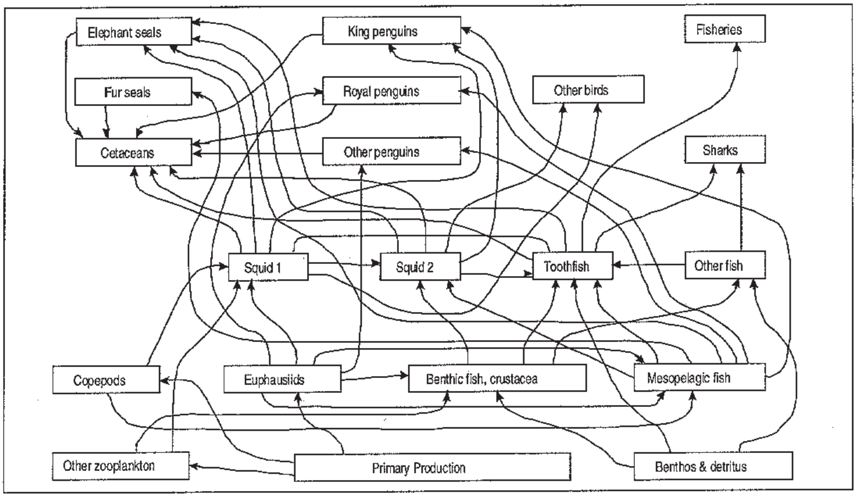


Figure 3a. The suggested food web for the marine system around Macquarie Island, in the sub-Antarctic, which is subject to the developing fishery for toothfish (Dissostichus eleginoides).

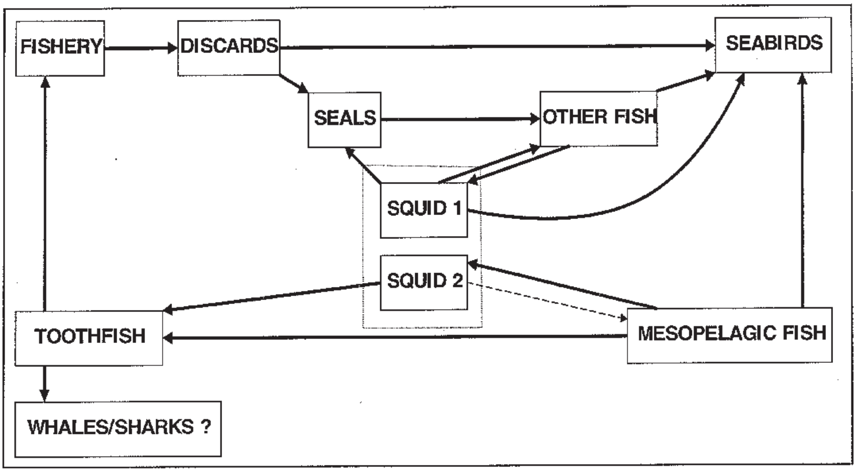


Figure 3b. A simplified food web thought to capture the gross dynamics of the effects of the imposition of the fishery. The major uncertainties are expressed by construction of alternative models through two changes to the web—treating the squid as either one or two compartments, and inclusion or exclusion of consumption of squid by mesopelagic fish.

simplified “minimum realistic model” of the Macquarie Island food web. In this simplification alternative models are created, through two changes to the web, that are thought to encompass the main uncertainties in predicting the gross effects of the fishery—the first change being treating the squid as either one or two compartments and the second being inclusion or exclusion of consumption of squid by mesopelagic fish.

Such approaches, combining simplified food webs with uncertainty in their structure, will be increasingly common in fishery assessment. The motivation will be both to improve the assessment of the fished stocks and to address the broader ecosystem effects of fishing. These considerations will quickly lead to the development and use of precautionary reference points and strategies designed to protect marine food-chain integrity. An early example of this is the adoption by CCAMLR of a strategy based on precautionary reference points to protect the predators or prey of the fishery target species (de la Mare 1996). The approach is pragmatic and motivated by a desire to limit reduction of the target species, to protect ecosystem integrity, and to leave sufficient escapement to support the food chain. By recognizing that the information to accurately determine these requirements will not be easily obtained, the strategy consists of a precautionary reference point and decision rule. The decision rule operates on the results of two calculations. The first calculation (Fig. 4a) addresses the level of reduction of the target population associated with different exploitation rates. It involves calculating the fixed proportion of the population that can be caught each year while ensuring that the spawning biomass has less than a 10% chance of being below 20% of the median unexploited level of spawning biomass during a 20-year period. The second calculation (Fig. 4b) addresses the amount of escapement of the target species from the fishery; escapement that is then available to support food chain integrity and dependent predator populations. It involves calculating the fixed proportion of the population that can be caught each year while ensuring that the median spawning biomass over a 20-year period is equal to either 50% or 75% of the unexploited median level; 75% is used for designated key prey species and 50% is used for all other species. The fixed proportions obtained from these two calculations are then multiplied by an estimate of the biomass to give a corresponding catch level, and the precautionary decision rule then selects the lower of the resulting two catch levels. This precautionary strategy was tested by simulation. It attempts to account for predator-prey interactions without engaging the full complexity, or requiring extensive information on mechanisms, of the real marine food web.

This strategy was initially developed by CCAMLR for the Antarctic krill resource, but has now also been applied to some fish stocks, so as to protect dependent predator populations.

This approach will become more common during the twenty-first century. Specifically it will involve development and use of default and pre-

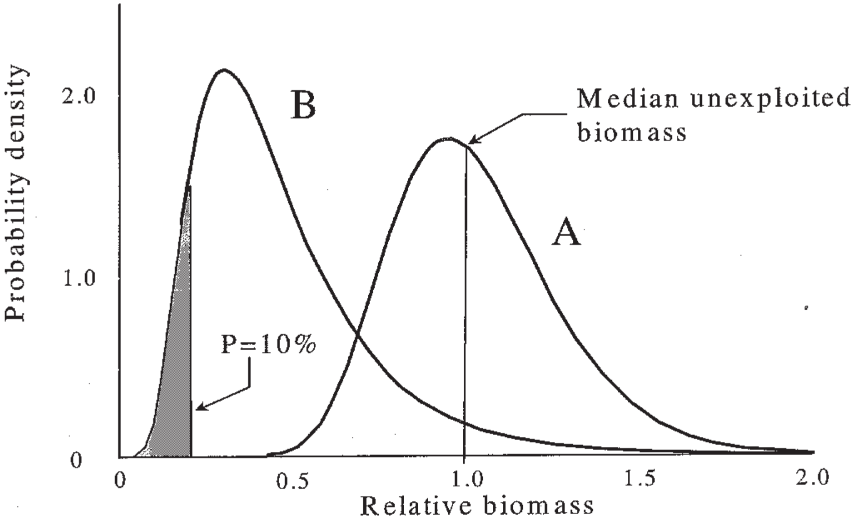


Figure 4a. The first part in calculation of the CCAMLR precautionary catch level for krill intended to protect dependent predator species. Simulation is used to generate the distribution of population sizes by year without fishing (distribution A) as a result of such things as recruitment variability, and the distribution of population sizes with a certain fishery exploitation rate (distribution B). The exploitation rate that gives a 10% probability of the population being below 20% of the unexploited median level in a 20-year period is calculated and is the first part of the decision rule to determine the precautionary catch level (from de la Mare 1996).

cautionary reference points and decision rules to protect habitats, food chains, and ecosystem integrity. Initially these will be based on simulation tests that use relatively simple models of spatial dynamics, habitat dynamics, and food-chain dynamics. Improvements in the ability to effectively use complex models with large and diverse data sources are already clear. The element that is lagging in its development, and which will be necessary to realize the potential of the analytic developments, is improved technology and methods for observing the ocean and its biological system.

3. Design of Robust Feedback Strategies for the Whole Fishery Management System

The third major area of paradigm shift will be increased involvement of science and scientists in the design and evaluation of management strategies that incorporate the whole fishery management system. Presently this task is seen primarily as the domain of the policy and management

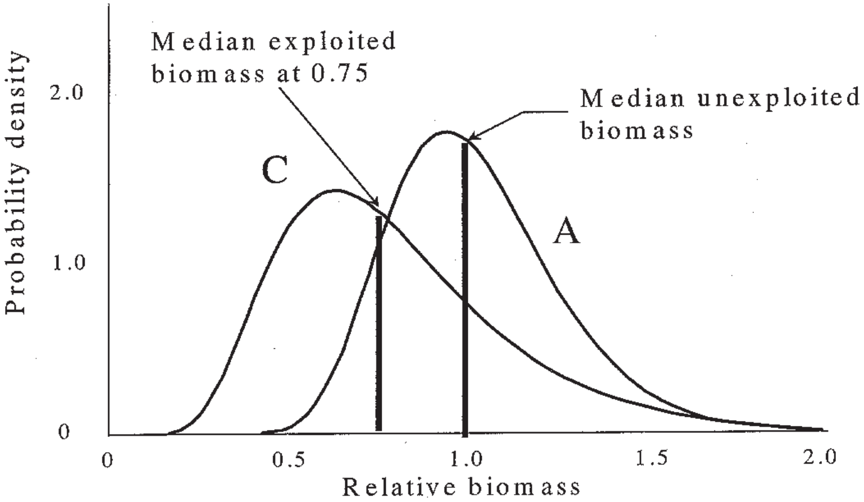


Figure 4b. The second part in calculation of the precautionary catch level for krill. The distribution of krill abundance after 20 years of fishing at a certain exploitation rate is calculated by simulation (distribution C). The exploitation rate that results in the median of the distribution being 75% of the median of the unfished distribution (distribution A) is calculated. The precautionary catch level for krill is then calculated using the lower of the exploitation rates from Figs. 4a and 4b (de la Mare 1996).

specialists in government bureaucracies, with science having a role in the implementation of the selected strategies but not being central to strategy design and evaluation. If there is to be significant progress in the effectiveness of living marine resource assessments in the twenty-first century this omission will have to change, and there will be need for change among policy, management, and science specialists so as to accommodate and integrate their different skills. Mace (1996) identifies the key elements for future success in managing marine resources as being stakeholder development of global conservation standards and guidelines, holistic national policies that integrate environmental and economic sustainability, and strengthened fishery management plans based on the precautionary approach. I totally agree with her in this conclusion, but amplify the point that scientists are stakeholders with a contribution to make to each of these three key elements and a responsibility to evaluate whether management plans are likely to achieve their intended objectives.

A key attitudinal shift required is acceptance that it is appropriate for science to play a central role in the design and performance of the whole management system. Scientists have a unique perspective and skill base for comment—not the only one and not necessarily the most relevant one at all times, but one that should be utilized. It is appropriate for scientists to comment on the design of the resource management system as a whole, and the strategies it employs, and to engage in debate about the performance and implications of the whole system and its strategies (e.g., Kerr and Ryder 1997, Hutchings et al. 1997). This can be a difficult and controversial task as the comments from scientists are not always welcome, particularly if they arise during crises with major political or social impact. It is an appropriate and valuable input nonetheless, particularly if it makes use of basic scientific procedures—documenting the observations used, documenting the methods and logic used to draw conclusions, and peer review. Organizational and methodological difficulties aside, resource assessment scientists of the twenty-first century will have to look beyond their "assessment box" of Fig. 1 to be effective. Their assessments must examine whether what is in the assessment box, when combined with the rest of the management system, can meet the overall management objectives. This will include consideration of social and economic outcomes, in addition to the state of the ecological resources, and evaluation of various economic and social instruments for resource management.

Paralleling and supporting this attitudinal shift is development of two practical and tangible areas of scientific activity, and the beginnings of development in each of these closely related areas are already in evidence.

The first is development of practical and measurable objectives and success indicators for marine resource management. It is common for fishery policy and management objectives to be stated very vaguely, making the qualitative evaluation of the likely success of different implementation strategies difficult and scientific evaluation impossible. The development of international standards and default objectives and reference points was the starting point of Mace's (1996) suggestion for improved fishery management. Increasingly scientific input is being used alongside that of other groups in the development of practical and measurable objectives. Some recent examples are the definitions of overfishing for U.S. fisheries, the development of the FAO Code of Conduct for Responsible Fishing (FAO 1995a), the FAO guidelines for use of the precautionary approach in fisheries (FAO 1995b), and recent efforts to develop coherent environmental sustainability indicators for assessment of both resource management and state of the environment. There is a long way to go before such developments give internationally accepted and measurable conservation standards for marine resources and ecosystems, but scientific development in this area is necessary for the new fisheries management paradigm and it will receive increasing attention.

The second is the design and evaluation of robust feedback strategies for management as a whole. This brings together and addresses a number of issues and failings of the present approaches to resource assessment, and will be the core of the new paradigm. It will use elements of the adaptive management framework of Walters, Hilborn, Ludwig and others (e.g., Walters 1986, Hilborn and Walters 1992), and has been referred to as a

Management Oriented Paradigm (a MOP for untidy fisheries) by de la Mare (in press). A key change in the focus of the science under the new paradigm is in the attitude to uncertainty; the focus moves from a preoccupation with finding the "one true model" and instead attempts to quantify or bound the uncertainty through explicit identification of alternative hypotheses. These alternative hypotheses are then used to examine the distribution of outcomes likely under any given management strategy. The scientific assessment process under this new paradigm will focus on the identification and weighting of alternative hypotheses, on determining system performance as measured by the achievement of management objectives, and on the design of feedback strategies which can use observations from targeted real world monitoring to detect departures from intended management outcomes and guide their correction.

The new paradigm is based on examining the management strategy as a whole, where the strategy consists of:

- Operational management objectives.
- Quantifiable performance measures based on these objectives.
- Specification of the measurements and monitoring that will be made and how they will be used in the resource assessment.
- Specification of how the results of the assessment will be used in management (usually via a "decision rule") and how decisions will be implemented.

This evaluation is usually by simulation trials across a range of alternative hypotheses that are thought to encompass the range of dynamics the real world might show in response to resource harvesting.

This approach makes a very fundamental change to the central questions being addressed through scientific assessment for resource management. The first step is the requirement that scientists demonstrate the assessment methods used can, in the context of the data available and the management structure, meet the management goals that have been identified for the fishery. It is notable that this essential first step is not addressed in most fishery assessment processes under most present fishery management arrangements; instead effort is usually strongly focused on trying to reach consensus on such things as the present state of the resource. That is, activity is strictly within the resource assessment box of Fig. 1.

The focus under the new paradigm will be on the robustness of the management strategy in the context of the available information, on monitoring that provides for detection and correction of errors, and on methods to successfully implement of management decisions.

In some regions of the world assessments will be almost data free, and assessments of robustness will be based on simulation testing across hypotheses of resource dynamics based on experience from elsewhere. In these situations robust management strategies will be found that extensively use time-space limitations to resource access (such as closed areas and seasons), and under these strategies considerable catch will be forgone to reasonably ensure sustainability under limited information. These strategies will be close to the traditional management strategies employed by many indigenous cultures. The scientific questions will be reasonably easily soluble this way, but there will be a significant management challenge in effective implementation of the necessary strategy under the social and economic pressures of the twenty-first century.

In other regions assessments will be data rich and robust strategies will use complex controls, including marine protected areas and spatial zoning of resource use. However, it is likely that strategies which reasonably ensure sustainability will be found to involve higher levels of protection and forgone catch than is the case in most presently used management strategies. Resource assessment and management in the twenty-first century will routinely see this reduced average catch as simply an expected cost of stability and sustainability of catches, and a consequence of the low information content of fisheries, slow and imperfect management implementation, and uncertainty in fishery assessments.

There are several examples where starts have already been made on the development and application of the methods of the new paradigm. These include South African hake (Punt 1992) and anchovy (Butterworth et al. 1993), Australian orange roughy (Smith 1993), Australian gemfish (Smith et al. 1996), an Australian tropical trawl fishery (Sainsbury 1991), Icelandic cod (Baldursson and Steffansson 1993), and the IWC-revised management procedure (Kirkwood 1993). All these examples used simulation trials of the performance of the whole management system, across a range of alternative hypotheses about resource dynamics and the observation process, and all judged performance using quantifiable measures derived from the management objectives. The IWC-revised management procedure was one of the earliest to develop and use the approach, and it provides a good example (see de la Mare 1996 for details).

The IWC revised management procedure was developed from extensive simulation testing of a management a framework similar to that in Fig. 1. Observations about the stock could come from the fishery itself and from scientific surveys, and the simulation tests were conducted across alternative hypotheses about features such as the dynamics of the stock (including time varying carrying capacity and population growth rate, and episodic disease outbreaks), stock structure, bias in historical catch records, variability in survey estimates of abundance, and time varying bias in survey estimates of abundance. The performance measures used to evaluate prospective management strategies reflected the IWC management objectives to have a high total catch, to have a low interannual variability in catches, and to avoid serious depletion of the stock at any time. The task then was to find a management strategy (i.e., a combination of the observations made, the analysis or stock assessment methods applied, and catch control rule for determining the catch from the stock assessment results) that scored well for all performance measures across all of the alternative hypotheses about stock dynamics and observation errors. The observations used in the strategy finally adopted were the annual catch and regular scientific survey estimates of stock abundance (fishery catch rate was not used to measure abundance). The stock assessment uses a very simple two parameter difference equation model of the population, with the parameters being carrying capacity and productivity. The catch control rule gives a zero catch if the population is estimated to be lower than 54% of the carrying capacity, and gives a linear increase in the catch limit for increasing abundance estimates above this. The rate of linear increase in the catch limit increases as the estimate of the stock productivity increases (see Fig. 5a). A Bayesian procedure is used to generate an updated distribution of catch limits, by combining the catch and abundance survey estimates with broad priors on the level of population depletion, the productivity and the bias in survey estimates. The distribution of catch limits represents the uncertainty in catch limit that results from uncertainty about the present population depletion and productivity, and selection of catches at higher or lower percentiles along that distribution would give less or more conservative management strategies. The IWC was presented with the implications of different selections to each of the performance measures and selected the forty-first percentile as giving the desired balance across the different management objectives (see Fig. 5b).

The key features of the approach are that considerable effort was put into identifying the key uncertainties in the resource dynamics and observation processes, and into developing and simulation testing prospective strategies across these uncertainties with respect to quantifiable management performance measures. Science contributed greatly in helping to pose the management questions in a way that allowed effective scientific analysis, in conducting that analysis, and in clearly showing the implications that different strategies had on achievement of management objectives. Selection of the desired balance across potentially competing management objectives was not considered to be a scientific task, but rather to be in the role of the managers.

The IWC example illustrates the evaluation of a passively adaptive management strategy. In passively adaptive management strategies, the fishery controls (such as catch) are altered in response to perceived changes in the resource toward or away from the desired resource state. Continued passively adaptive management will result in some empirical learning and resolution of uncertainty about the dynamics of the resource, but no specific changes in the fishery controls are taken to increase this rate of learning. Consequently the rate of learning can be very slow because often a fishery provides a weak experimental design for discrimination of alternative hypotheses about population regulation. There is often weak statistical contrast between alternative population control variables, such

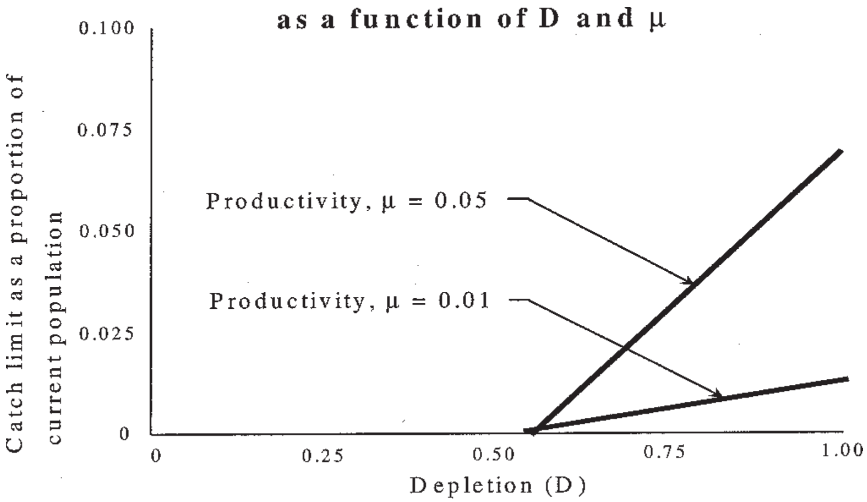


Figure 5a. The catch control rule for the IWC's revised management procedure. The proportion of the population caught increases linearly above a threshold level, and the gradient of this straight line increases as the estimated productivity of the stock increases (de la Mare 1996).

as population size, recruitment level, environmental variables, fishing intensity, and the size of other populations in the ecosystem; thus even after many years of observation there is not definitive discrimination between alternative hypotheses. If it is sufficiently important to separate some of these alternative hypotheses, as for example if they had significantly different management implications, then specific management experimentation may be justified. In such actively adaptive or experimental management strategies the fishery controls are altered specifically to improve the rate of learning about some important alternative hypotheses about the fishery. The design of an actively adaptive management strategy contains elements of experimental design to scientifically distinguish alternative ecological hypotheses (and in principle also hypotheses about economic and social dynamics). The performance of prospective actively adaptive strategies is evaluated using the same performance measures, derived from the management objectives, as would be used for passively adaptive strategies. The resolution of the alternative hypotheses is not an end in itself, and use of an actively adaptive strategy is justifiable only if it leads to an improvement in these performance measures that is acceptable to the managers and other relevant stakeholders.

There are two well developed examples of actively adaptive management in Australian fisheries; one dealing with species composition changes in a tropical multispecies trawl fishery on the North West Shelf (Sainsbury

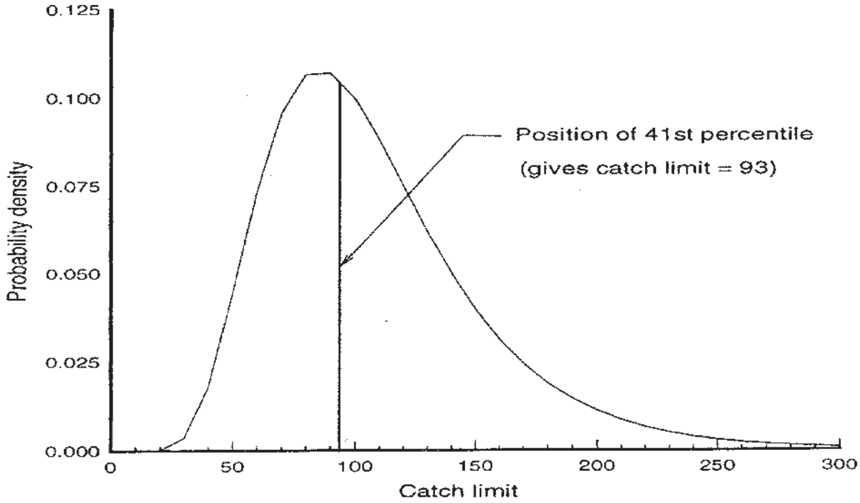


Figure 5b. A Bayesian method is used to calculate a posterior distribution of the catch limit from data on catches and resource surveys combined with prior distributions on the uncertainty in stock depletion, stock productivity, and bias in resource surveys. Selection of any particular catch limit has associated with it a distribution of outcomes for each of the management performance measures, which relate to management objectives such as maintaining the stock and catches. The IWC selected the forty-first percentile as providing the desired balance of outcomes across its management objectives.

1991, Sainsbury et al. 1997a) and the other dealing with the effects of line fishing on the Great Barrier Reef (Mapstone et al. 1996). Development of both strategies involved identification of uncertainties in understanding of the dynamics of the resource and formulation of explicit alternative hypotheses intended to reflect the range of responses to fishing thought to be reasonably likely. Both also involved detailed comparison of the performance of various possible strategies, including passively adaptive strategies, and extensive interaction between scientists, managers, and other stakeholders. The North West Shelf strategy was designed to discriminate among four hypotheses that could explain undesirable changes in species composition following the introduction of trawling. Field implementation of the North West Shelf strategy involved closure of some large areas to trawling. This has been under way since 1986 and has provided the expected resolution between the alternative hypotheses of resource dynamics. The Great Barrier Reef strategy was designed to provide estimates of the parameters of key population and fishery interaction parameters in a complex and spatially interconnected system of reefs. The experiment began in 1997, and the experimental design involves the use of four clusters, each of six reefs, and a sequence of opening and closing some reefs to fishing while others remain constantly open or closed.

Conclusions

So where do these thoughts leave us in speculating about the living marine resource assessments of the mid-twenty-first century?

The resource assessments will have to deal with more resource uses and users, more data and more diverse data sources (including new technologies for observing the marine system), more conflict over resource sharing, and complex spatial, habitat, and food-chain effects. There will still be a high level of uncertainty in the appropriate structure and parameter values for the models used in assessments. However, the assessments and related management structure will operate under a different paradigm and process. The central tenets of this paradigm will be precautionary management and the use of management strategies that have been scientifically demonstrated to robustly lead to achievement of management objectives despite the key uncertainties faced. The associated resource assessments will explicitly recognize model uncertainty, and will focus on evaluating the performance of prospective strategies. The key issues will be the identification and weighting of alternative hypotheses, the robustness of management strategies to uncertainty, and the design and performance of the management system. The management environment will include extensive stakeholder participation, and the management system will explicitly search for and use suitable social and economic incentives and management instruments to achieve its goals.

If living marine resource management is to be more successful in the twenty-first century than it has been in the twentieth century, then scientific resource assessments will need to adopt a new paradigm something like that outlined, and we will need to be more successful at linking science into the management system in a practical and effective manner. If this is not done then the future for marine living resources, and the industries and people dependent on them, is gloomy indeed. However, the outline of some features of the new paradigm appear to be forming now and, even if there are still some surprises in store, the appropriate way forward is there. I believe and hope that we will develop and adopt a new paradigm for marine resource assessment, and that we will be more effective in linking the science with practical fishery management. But we have a lot of work to do.

References

Baldursson, F.M., and G. Stefansson. 1993. On the rational utilization of the Icelandic cod stocks. ICES C.M. 1993/g:56.

Butterworth, D.S., J.A.A. De Oliveria, and K.L. Cochrane. 1993. Current initiatives in refining the management procedure for the South African anchovy resource. In: G. Kruse, D.M. Eggers, R.J. Marasco, C. Pautzke, and T.L. Quinn (eds.), Proceedings of the International Symposium on Management Strategies for Exploited Fish Populations. University of Alaska Sea Grant, AK-SG-93-02, Fairbanks, pp. 439-473.

Chalmers, A.F. 1976. What is this thing called science? An assessment of the nature and status of science and its methods. University of Queensland Press, St. Lucia, Qld., Australia.

de la Mare, W.K. 1996. Some recent developments in the management of marine living resources. In: R.B. Floyd, A.W. Shepherd, and P.J. De Barro (eds.), Frontiers of population ecology. CSIRO Publishing, Melbourne, Australia, pp. 599-616.

de la Mare, W.K. In press. Tidier fisheries management requires a new MOP (management oriented paradigm). Rev. Fish Biol. Fish.

FAO. 1995a. Code of Conduct for Responsible Fisheries. FAO, Rome.

FAO. 1995b. Precautionary approach to fisheries. Part I: Guidelines on the precautionary approach to capture fisheries and species introductions. FAO Fish. Tech. Pap. 350/1.

Fournier, D.A., and J. Hampton. 1996. Preliminary analyses of South Pacific albacore catch, effort and length frequency data, using an age structured, length-based model with spatial structure. Paper presented to ICCAT 25th Anniversary Tuna Symposium, June 10-18, 1996, Azores, Portugal.

Garcia, S.M., and C. Newton. 1997. Current situation, trends and prospects in world capture fisheries. In: E.K. Pikitch, D.D. Huppert, and M.P. Sissenwine (eds.), Global trends: Fisheries management. Amer. Fish. Soc. Symp. 20:3-27.

Hastings, W.K. 1970. Monte Carlo sampling methods using Markov Chains and their applications. Biometrica 57:97-109.

Hilborn, R., and C.J. Walters. 1992. Quantitative fisheries stock assessment. Chapman and Hall, New York.

Holmes, R. 1994. Biologists sort the lessons of fisheries collapse. Science 264:1252-1253.

Hutchings, J.A., and R.A. Myers. 1995. The biological collapse of Atlantic cod off Newfoundland and Labrador: An exploration of historical changes in exploitation, harvest technology, and management. In: R. Arnason and L.F. Felt (eds.), The North Atlantic fishery: Strengths, weaknesses and challenges. Inst. Island Stud., University of Prince Edward Island, Charlottetown, Prince Edward Island, pp. 37-93

Hutchings, J.A.C., C.J. Walters, and R.L. Haedrich. 1997. Is scientific inquiry incompatible with government information control? Can. J. Fish. Aquat. Sci. 54:1198-1210.

Isaac, V.J., M.L. Ruffino, and D. McGrath. 1998. In search of a new approach to fisheries management in the middle Amazon region. In: F. Funk, T.J. Quinn II, J. Heifetz, J.N. Ianelli, J.E. Powers, J.F. Schweigert, P.J. Sullivan, and C.-I. Zhang (eds.), Fishery stock assessment models. University of Alaska Sea Grant, AK-SG-98-01, Fairbanks (this volume).

Johannes, R.E. 1998. The case for data-less marine management: Examples from tropical nearshore fisheries. Trends Ecol. Evol. 13(6):243-245.

Kerr, S.R., and R.A. Ryder. 1997. The Laurentian Great Lakes experience: A prognosis for the fisheries of Atlantic Canada. Can. J. Fish. Aquat. Sci. 54:1190-1197.

Kirkwood, G.P. 1993. Incorporating allowance for risk in management: The revised management procedure of the International Whaling Commission. ICES C.M. 1993/N:11.

Klaer, N., K.J. Sainsbury, and T. Polacheck. 1996. A retrospective examination of southern bluefin tuna VPA and projection results 1982-1995. Commission for the Conservation of Southern Bluefin Tuna Scientific Committee document, CCSBT/SC/96/31.

Ludwig, D., R. Hilborn, and C. Walters. 1993. Uncertainty, resource exploitation, and conservation: Lessons from history. Science 260:17-36.

Mace, P.M. 1996. Developing and sustaining world fisheries resources: The state of the science and management. In: D.A. Hancock, D.C. Smith, A. Grant, and J.P. Beumer (eds.), 2nd World Fisheries Congress, CSIRO Publishing, Collingwood, Australia, pp. 1-20.

Mace, P.M., and M.P. Sissenwine. 1993. How much spawning biomass per recruit is enough? In: S.J. Smith, J.J. Hunt, and D. Rivard (eds.), Risk evaluation and biological reference points for fisheries management. Can. Spec. Publ. Fish. Aquat. Sci. 120:101-118

Mangel, M., et al. 1996. Principles for the conservation of wild living resources. Ecol. Appl. 6:339-362.

Mapstone, B.D., R.A. Campbell, and A.D.M. Smith. 1996. Design of experimental investigations of the effects of line and spear fishing on the Great Barrier. Reef CRC Reef Research Centre, Technical Report No. 7, Townsville.

Morrow, R., and P. De Mey. 1995. An adjoint assimilation of altimetric, surface drifter and hydrographic data in a QG model of the Azores Current. Journal of Physical Oceanography 100(C12):25007-25025.

Myers, R.A, J. Barrowman, H.J.A. Hutchins, and A.A. Rosenberg. 1995. Population dynamics of exploited stocks at low population levels. Science 269:1106-1108.

Munro, J.L. 1987. Workshop synthesis and directions for future research. In: J.J. Polovina and S. Ralston (eds.), Tropical snappers and groupers: Biology and fisheries management. Westview Press, Boulder, CO, pp. 639-659

Murphy, G.I. 1977. New understanding of southern bluefin tuna. Aust. Fish. 36(1):2-6.

Punt, A.E. 1992. Selecting management methodologies for marine resources, with an illustration for the South African hake. S. Afr. J. Mar. Sci. 12:943-958.

Punt, A.E. 1993. The comparative performance of production-model and ad hoc tuned VPA based feedback-control management procedures for the stock of Cape hake off the west coast of South Africa. In: S.J. Smith, J.J. Hunt, and D. Rivard (eds.), Risk evaluation and biological reference points for fisheries management. Can. J. Fish. Aquat. Sci. Spec. Publ. 120:283-299.

Punt, A.E., and D.S. Butterworth. 1995. The effects of future consumption by the Cape fur seal on catches and catch rates of the Cape hakes. 4. Modelling the biological interaction between Cape fur seals *Arctocephalus pusillus pusillus* and the Cape hake *Merluccius capensis* and *M. paradoxus.* S. Afr. J. Mar. Sci. 16:255-285.

Punt, A.E., and R. Hilborn. 1997. Fisheries stock assessment and decision analysis: The Bayesian approach. Rev. Fish Biol. Fish. 7:35-63.

Rosenberg, A.A., M.J. Fogarty, M.P. Sissenwine, J.R. Beddington, and J.G. Shepherd. 1993. Achieving sustainable use of renewable resources. Science 262:828-9.

Rubin, D.B. 1988. Using the SIR algorithm to simulate posterior distributions. In: J.M. Bernardo, M.H. Degroot, D.V. Lindley, and A.M. Smith (eds.), Bayesian statistics 3: Proceedings of the Third Valencia International Meeting, 1-5 June 1987. Clarendon Press, Oxford, pp. 385-402.

Sainsbury, K.J. 1991. Application of an experimental approach to management of a tropical multispecies fishery with highly uncertain dynamics. ICES Mar. Sci. Symp. 193:31-320.

Sainsbury, K.J. 1993. The use of simulation to evaluate the performance of stock rebuilding strategies, including the use of reference points. In: A.A. Rosenberg (ed.), Defining overfishing-defining stock rebuilding. NOAA Tech. Memo. NMFS-F/SPO-8.

Sainsbury, K.J., R.A. Campbell, R. Lindholm, and A.W. Whitelaw. 1997a. Experimental management of an Australian multispecies fishery: Examining the possibility of trawl-induced habitat modification. In: E.K. Pikitch, D.D. Huppert, and M.P. Sissenwine (eds.), Global trends: Fisheries management. Amer. Fish. Soc. Symp. 20:107-112.

Sainsbury, K.J., D. Butterworth, C. Francis, N. Klaer, T. Polacheck, A.E. Punt, and A.D.M. Smith. 1997b. Incorporating uncertainty into stock projections. Report of the Scientific Meeting 3-7 April 1995. CSIRO Marine Laboratories, Hobart, Australia.

Smith, A.D.M. 1993. Risks of over- and under-fishing new resources. In: S.J. Smith, J.J. Hunt, and D. Rivard (eds.), Risk evaluation and biological reference points for fisheries management. Can. Spec. Publ. Fish. Aquat. Sci. 120:261-267.

Smith, A.D.M., A.E. Punt, S.E. Wayte, and N.L. Klaer. 1996. Evaluation of harvesting strategies for eastern gemfish *(Rexea slandri)* using Monte Carlo simulation. In: A.D.M. Smith (ed.), Evaluation of harvesting strategies for Australian fisheries at different levels of risk from economic collapse. CSIRO, Australia, pp. 120-160.

Smith, T. 1994. Scaling fisheries. Cambridge University Press, Cambridge, U.K.

Walters, C.J. 1986. Adaptive management of renewable resources. Macmillan, New York.

Walters, C., and J.J. Maguire. 1996. Lessons for stock assessment from the northern cod collapse. Rev. Fish Biol. Fish. 6:125-137.

WCED. 1989. Our common future. World Commission on Environment and Development. Oxford University Press, New York.

Evaluation of the Status of Fisheries Data Collection and Stock Assessment Problems in São Paulo, Southeastern Brazil

Maria de los Angeles Gasalla and
Acácio Ribeiro Gomes Tomás
Marine Fisheries Research Center, Instituto de Pesca, Santos, Brazil

Abstract

While the modernization of fishery management tools is increasing throughout the world, developing countries are still trying to build appropriate data collection mechanisms with limited facilities and in fisheries critical condition. Despite the depletion of resources, our fishery scientists usually find an inappropriate context for assessment, expressed in particular by uncertain or nonexistent data and poor estimation of spatial and temporal scales. This is the case in Brazil, which has always found it logistically very difficult and expensive to collect data along its 8,000 km long coast where artisanal activities account for more than 50% of the activity. However, São Paulo state fisheries data have been considered one of the most complete and robust, especially due to the Instituto de Pesca collection efforts that have carefully archived historical landing data in a classical approach since 1968.

This paper gives an overview of the situation with São Paulo State fisheries and databases with the aim of evaluating the possible kinds of data inputs and outputs for stock assessment. Critical points will be discussed. The industrial fleet is composed mainly of sardine seiners, shrimp and paired bottom trawlers, gillnetters and tuna longliners. Fleet size and discards are scarcely known. Besides data collection based on interviews, vessel owners also contribute information about total fish production. Unfortunately, fishery effort data has been obtained only for part of the landings, and provide inaccurate estimates. Recently, with the increase in landing sites, the decreasing number of official data collectors, and poor research budgets, complete data collection is becoming more of a chal-

lenge. However, the lack of information on other fisheries that share the same stocks is the most critical assessment problem, resulting in poor estimates of total stock catch and effort. The need to improve data acquisition and archive strategies is clear. Nevertheless, information currently available is essential and valuable if we wish to discuss how to use and optimize our resources properly.

Introduction

Several problems concerning stock assessment can often be the result of failed data collection practices. An evaluation of the condition of fishery data can be a starting point to make possible the implementation of stock assessment procedures in places where these practices are not systematic. São Paulo state industrial fishery landing control historically seems to be one of the most robust and consistent in Brazil. The most challenging problems concern catch and effort data collection. However, at present, these limitations can seriously compromise fishery production estimates for stock assessment.

This work aims to show a synthesis of the present status of fisheries and a review of data collection practices and stock assessment methods in the state of São Paulo. A fishing activity overview and a comprehensive description of data sources are the focus of this report so that a search for alternative solutions to future assessment may be reached.

Methods

The information used in the evaluation was obtained from several published articles, official statistics, reports, personal communications, fisherman and researcher interviews, and field work.

Sources included: SUDEPE (no date), Ripley (1956), Richardson and Moraes (1960), Braga (1962), Vazzoler and Vazzoler (1969), Neiva and Moura (1977), SUDEPE/PDP (1985), Bendazoli and Rossi-Wongtschowski (1990), Valentini et al. (1991a, b, c), Valentini and Cardoso (1991), and Zavala-Camin and Antero da Silva (1991), IBAMA (1992, 1993, 1995), Lin (1992), Cergole (1993), Arfelli and Amorim (1994), Tiago et al. (1994), Amorim et al. (1995), Castro and Castro (1995), Costa et al. (1995), Rossi-Wongtschowski et al. (1995), Dias and Dornelles (1996), SEMA (1996), Silva (1996), Tomás (1996), Arfelli et al. (1997), Gasalla et al. (1997), REVIZEE (1997).

Field work consisted in personally checking the number of operational fishing boats of the region and interviewing fishermen. This process allows a fair estimate for fleet size.

Information on the present situation of fishery databases was obtained from the Instituto de Pesca Marine Fisheries Research Center.

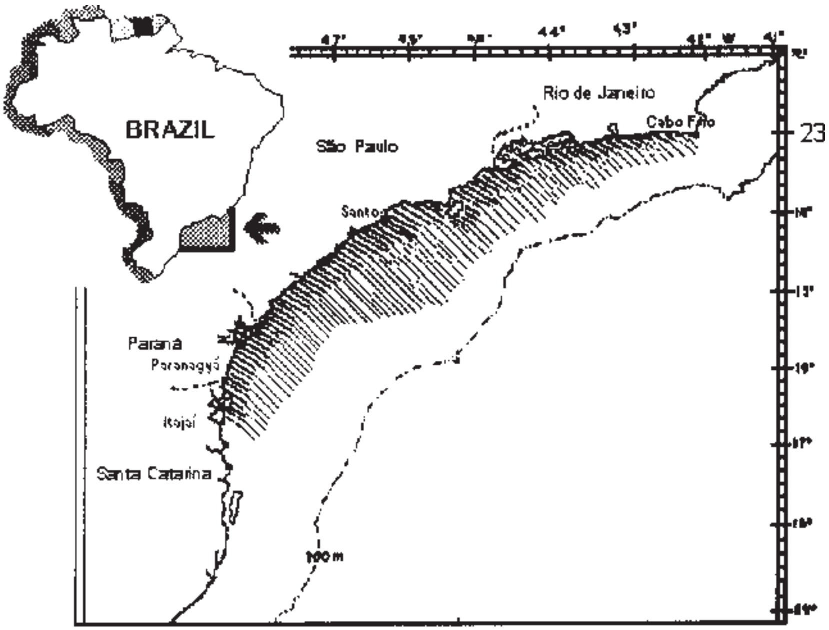


Figure 1. Location of fishery area of São Paulo fleet.

Brief Historical and Socioeconomic Background

The state of São Paulo is located in southeastern Brazil, the most economically developed region of the country, which hosts the largest South American metropolis, São Paulo City, with about 16 million people. Crossed by Tropic of Capricorn and extending between 23°33′ and 25°18′28″S and 44°43′24″ and 48°14′18″W, its coastal zone is about 700 km long (northeast-southwest direction) (Fig. 1). São Paulo coastal zone shelters Atlantic tropical rainforest remnant regions, with great biological diversity and geomorphological complexity, including estuaries, mangroves, caverns, rivers, rocky and sandy shorelines and several islands. Santos region, located 70 km from São Paulo City, is the most urbanized part of the coast, and supports the biggest South American harbor and an industrial petrochemical complex. It is the most environmentally impacted as well.

In 1991, 1,647,235 inhabitants were living in São Paulo state's coastal zone, depending on tourism, agriculture and fisheries (SEMA 1996) for economic support. As this marine coast represents an extensive area for fisheries, its landing harbors provide a significant part of the big São Paulo City commercial fish for market. Even though the fishery sector seems to play a minor role in the economy when compared with other agricultural activities, the economic importance of fisheries tends to be highly localized and concentrated. Fish, as food, is important for people living along the coastal zone and fisheries directly or indirectly represent an important source of employment and livelihood. The fisheries sector provides different kinds of employment varying with the type of exploitation and level of commercialization. Small-scale fisheries are labor intensive and families benefit greatly from this type of employment. No public regulation (such as credits, minimum price control, or government incentives) is presently available to this sector.

The São Paulo total coastal fleet fisheries area encompasses 5,308 square km and experiences high fishing effort from other states' fleets. Compared with other South American continental shelves, São Paulo state's shelf is narrow (with a maximum width of 110 miles, at Santos), with a sandy and muddy bottom especially favorable for trawling activities.

São Paulo total landings represent about 10% of total Brazilian catches and showed a downward trend especially after 1984. Fig. 2 shows the evolution of São Paulo State total landings from 1947 to 1995. In 1984, São Paulo's harvest of fish from capture fisheries reached a level of 131,000 tons. In 1995, official data indicates values of 35,000 tons. On average, fishes contributed 87% annually, and crustaceans and mollusks, 13%. The most important regional fishery resource is the Brazilian sardine, whose industrial landings contribute about 50% of total annual landings, followed by demersal fishes and shrimp. Figs. 3 and 4 show the makeup of the fishing gear and species resources.

It is difficult to identify the line between subsistence and commercial, or between industrial and small-scale fisheries in this region. Small-scale fisheries have a great historical and socioeconomic importance for the traditional coastal population (called "caiçaras"). They have been losing traditions and lands due to intensification of tourism and real estate expansion in coastal areas. To date, no restricted area is guaranteed to small-scale fishermen. They have been adapting several typical artisanal fishing gears, such as set traps, gillnets, beach and haul seines, trammel, fyke and hoop nets, squid hook-and-line, and others. No up-to-date fleet size estimates are known for these fisheries.

Present Situation of the Fishery

A complete picture of São Paulo state's main fisheries is shown in Tables 1, 2 and 3, including target species, fleet characteristics, evolution, and a synthesis of regulation procedures and constraints. These tables present the context necessary for understanding the general problems of the fishery. The evolution of the main stock landings in the state of São Paulo and in the entire landing area is shown in Figs. 5, 6, 7, 8, 9, and 10.

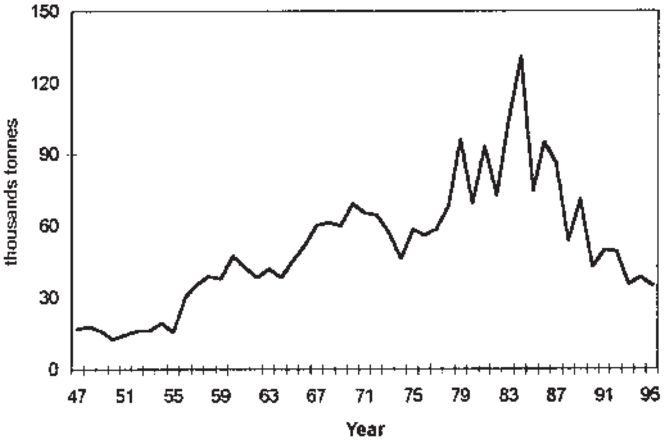


Figure 2. Evolution of São Paulo landings.

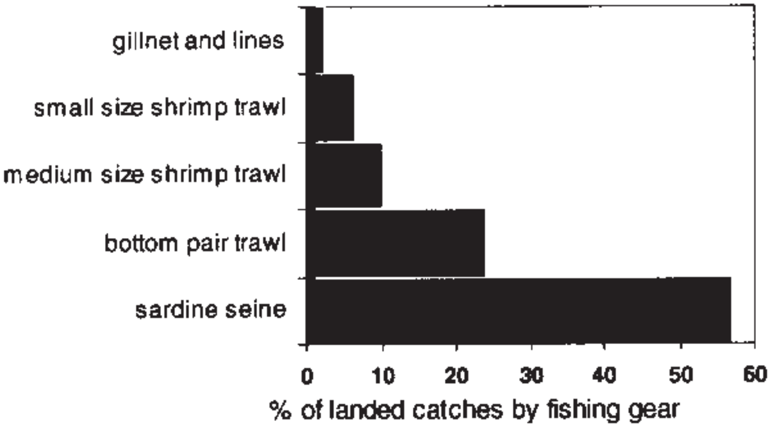


Figure 3. Relative importance of fishing gear in São Paulo 1995 landings (tuna longlining not included).

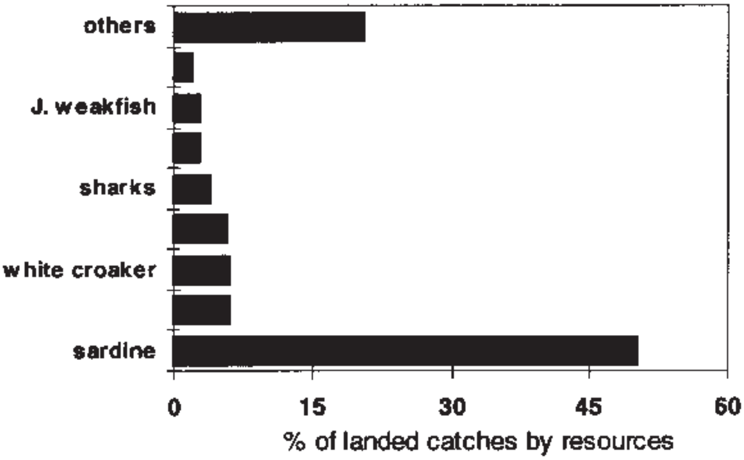


Figure 4. Relative importance of main resources in São Paulo 1995 landings.

The Federal government, through legal norms implemented by the Brazilian Institute for Environment and Natural Resources (IBAMA) regulates fisheries in Brazil. Decisions from this organization are often based on studies and scientific work group technical reports. Beyond these results, stock assessment is not a systematic practice, nor are management decisions based on specific reference points.

São Paulo State's Fishery Data Collection and Stock Assessment

Background

São Paulo landings statistic control program began in 1942, and was conducted until 1946 by federal agencies (Braga 1962). This data was gathered by a local research group coordinated by FAO technicians in the 1950s and 1960s, such as Ripley (1956), Richardson and Moraes (1960), and Vazzoler and Vazzoler (1969). All later sampling strategies were based on this approach. Data were based on interviews of fleet captains in which qualified collectors registered data on logbooks. The fishermen's own record prints had several problems due to the captain's education level. The sample unit was Santos commercial fleet, which in 1958 landed at only three well-controlled points.

The Instituto de Pesca, a state public institution for fishery research, made fishery collection efforts after 1968 using the previous methodology. The objective was to verify stock tendencies and not to survey total

Table 1. São Paulo state's fisheries: target species, fleet, and characteristics.

Stocks	Target species	Fishing gear	Operation depth	Fleet size[a]	Technical characteristics	Bycatch species	Observations
Brazilian sardine	*Sardinella brasiliensis*	Purse seine	To 70 m, mostly common to 50 m	112	Wooden or steel hulls Average length: 19.96 m Average GRT: 54.73 Average HP: 264.01 Mesh size: 12 mm	*Trachurus lathami, Mugil platanus, Katswonus pelamis, Scomber japonicus, Trichiurus lepturus* and over 20 more species.	Seasonal bycatches
Pink shrimp	*Penaeus brasiliensis* and *P. paulensis*	Double otter trawl	30-90 m	236	Average length: 19.5 m Average GRT: 62 (21% less than 20) Average HP: 264 Mesh size: 7 mm Average effort per boat: 19 trips/y, 13 days/trip, 3 hauls/day, and 4 hr/haul	Around 165 fish species, 35 crustaceans, and 25 mollusks. After 1984, toadfish (*Lophius gastrophysis*), rays (*Raja* and *Dasyatis*) and small sharks (*Mustelus*, *Squatina*, *Squalus*, and carcharhinids and sphirnid juveniles) had been landed.	Target species represent less than 20% of total landings; bycatch species represent profit as byproduct since middle of 1980s. Biggest boats often operate in greater depths searching for deep lobster *Metapenaeus rubellus*.
Sea-bob shrimp	*Xiphopenaeus kroyeri*	Double otter trawl	30 m	~1500	Average length: 8.9 m Average GRT: 4.9 (84% less than 5) HP: 18 to 180 Mesh size: 4 mm Average effort per boat: 50 trips/y, 4 days/trip, 5 hauls/day, and 3 hr/haul	80 fish species, more than 20 crustaceans (seasonally also *Penaeus schmitti*, *Artemesia longinaris* and exceptionally *Pleoticus muelleri*) and mollusk species.	Target species represent 90% of total landings High discard ratios (mainly on finfishes)
Tuna-like fishes	*Xiphias gladius*	Shallow monofilament longline	50 m	16	Average length: 27 m Average GRT: (100-150) HP: 580 Average effort per boat: 11 trips/year, 15-25 days/trip, 1,500 hooks/haul-day	Tunas and sharks (*Thunnus alalunga, T. albacares, Prionace glauca, Isurus oxyrinchus, Carcharhinus* spp., *Tetrapturus albidus, Sphyrna* spp., *Alopias superciliosus, Istiophorus albicans*), and more than 10 other species.	Target species effort changed from tuna to sharks in early 1980s, and from sharks to swordfish after 1994.

[a] Estimated size

Table 1. (Continued.) São Paulo state's fisheries: target species, fleet, and characteristics.

Stocks	Target species	Fishing gear	Operation depth	Fleet size[a]	Technical characteristics	Bycatch species	Observations
Demersal fishes	*Micropogonias furnieri*, *Macrodon ancylodon*, *Cynoscion jamaicensis*, *C. guatucupa*, and	Bottom pair trawl	40 m	48	Average length: 21.3 m Average GRT: 70.66 Average HP: 291 Mesh size: 90 mm (body) 50 mm (codend)	77 species from 25 families	Bottom pair trawling captures more than 80% of demersal fish catches
	Balistes capriscus	Bottom gillnet		15	Average length: 15 m Average HP: 210 Mesh size: 10-40 mm	Unknown	Target species represent 82% of catches, and bycatches include many commercial species High selective gillnetting capture mainly adults
Groundfish (groupers, tilefish, and sandperch)	*Lopholatilus villarii*, *Epinephelus* spp. and *Pseudopercis* spp.	Bottom longline	500 m	6	Steel or wooden boats Average length: 21 m Average HP: 380	*Urophycis cirrata*, *Helicolenus lahillei*, *Genypterus brasiliensis*, *Squalus* spp. and 15 more species	New fishery
Anchovy	*Anchoviella lepidentostole*	Seine	3-4 m	525	Wooden canoes 4-5 m length 4-6 HP small engine powered motors	None	Anadromous fishery is in the "Ribeira do Iguape" river during spawning
Sharks	*Sphyrna lewini* and *S. zygaena*	Surface driftnet	30 m	0 (1997) 60 (1990) and 5 (1996)	Average length: 17 m Mesh size: 14-40 cm	*Carcharhinus* spp., *Isurus oxyrinchus*, other sharks and small cetaceans and turtles	Important fishery introduced in late 1980s. Presently no driftnet operates in the state of São Paulo. Boats have changed to bottom longlining and trawling gear, due to low yield. Changes may be seasonal.

[a] Estimated size

Table 2. São Paulo state's fisheries evolution.

Main fisheries	Degree of exploitation	Attributed causes
Brazilian sardine, purse seining	Stock hardly overexploited, no recuperation indication but a possible equilibrium step after 1994 due to reduction of effort.	1. Effort increased, fishing mortality increase. 2. Spawning stock biomass and recruitment decrease. 3. Recruitment failure hypothesis attributed to oceanographic anomalies during spawning period (weak penetration of cold mass water in coastal zone influenced larval survival).
Pink shrimp, trawling	Overexploitation. However, a possible equilibrium step is being reached at a low population size, due to the reduction of effort (some boats have changed to other fisheries).	1. Excessive effort. 2. Reduction of number of recruits incorporated into parental stock.
Sea-bob shrimp, trawling	Overexploitation, possible recovering.	Excessive effort.
Tuna and tuna-like, longlining	Increasing effort, swordfish juvenile catches, some shark species may be endangered.	Changing effort to swordfish after a long period with more than 50% of shark catch.
Demersal fishes, pair-trawling	Some species (such as king weakfish) are hardly overexploited; white croaker and *Cynoscion jamaicensis* seem to be in equilibrium. Average length of many species are in downward trend, upward trend of non-past value species as triggerfish, *Balistes capriscus*, which acquired market importance.	1. High effort but present reduction of fleet number caused by migration to other states' harbors. 2. Non-past value species (as triggerfish, *Balistes capriscus*) gain importance in last decade. 3. Non-effective control of minimum length catches. 4. Small individuals are caught by shrimp fleet.
Groundfish, bottom longlining	Unknown	Fishing gear introduced in 1994.
Anchovy	After an overexploitation period, stock recovered and landing values fishery increased twice.	1. Overfishing was associated with catches over age 1 and 2 groups, also over-mature individuals. 2. High fishing mortality. 3. Recuperation may reflect successful management procedures.
Sharks, surface driftnet	Some shark species may be endangered, accidental catches of small cetaceans and turtles.	1. Great increase in last ten years. 2. Lately it has been reduced.

Table 3. São Paulo state's fisheries regulation.

Main fisheries	Application of management procedures (federal legal basis)	Studies recommendation	Present situation of regulation	Present problems
Brazilian sardine, purse seining	1976 to 1990: 3 month summer closures to protect spawning season, forbidden catches of individuals less than 17 cm length with 15% of weight tolerance, effort control through fleet licenses. 1991: 45 day summer closure and 15 day winter closure to protect recruitment, reduction to 5% of weight tolerance for fish under 17 cm. 1995-1997: only 45 day summer spawning period closure, to avoid sardine biomass reduction "Bonito" pole-and-line boats must catch its own live-bait other than sardine.	Protection of spawning stock; effort control (not to allow new licenses); use of alternative species as live-bait for "bonito" pole-and-line fishery; need further investigation on sardine stock, as acoustic biomass work, biological monitoring and oceanographic conditions survey by remote sensing.	1. Reduction of closure period due to fishing trade sector pressure after the first sign of resource recuperation. Increased landings in 1994 (80,000 ton) could be related to the reduction of effort in early years (partially due to better prices offered by imported sardine), 2. Reduction of illegal landings even though weak inspection, 3. Nonrespected minimum length, 4. Fleet size not controlled, 5. First government intention of cooperative management.	1. Intensive fishery over juveniles—difficult effective control; 2. Necessity for reproductive parameter monitoring (Does L_{50} remain at 17cm?); 3. Excessive fishery effort (number of operating boats) retracted in 1991 due to low productivity, many of them changing to other fishing gear; 4. Real risk of capture industries going bankrupt.
Pink shrimp, trawling	3 month closures, to protect recruitment; 7 mm codend mesh size limitation. Fleet bigger than 10 GRT should operate 1.5 miles from the coast.	Increase of samples to biological studies and intensification of data handling program with methodology standardization. Studies on bycatch suggested reduction necessity.	Fleet size limitation (no more licenses were given) since 1984.	Nonrespect of legal norms. Exportation difficulties (caused by money currency); fishers complain that TED reduces net effectiveness.
Sea-bob shrimp, trawling	4 mm codend mesh size limitation.	Continued study of population structure and biological parameter estimates. Recommendation to maintain fishery effort level. Studies suggested bycatch reduction necessity.	Fleet size limitation (no more licenses were given) since 1984.	Variation of fleet size, weak control over forbidden catches until 1.5 nmi offshore. High fluctuation on landings may be explained by the detour of part of catches to clandestine wharves.
Tuna and tuna-like, longlining	None for tunas, minimum catch size and dressed weight for swordfish.	ICCAT recommendations: reduction of juvenile catches.	Minimum sizes and weights to YFT (in 1973) and BET (in 1981).[a]	Variation of fleet size.

Table 3. (Continued.)

Main fisheries	Application of management procedures (federal legal basis)	Studies recommendation	Present situation of regulation	Present problems
Demersal fishes, pair-trawling	Allowed minimum length to white croaker (30 cm), king weakfish (25 cm) and stripper weakfish (30 cm) (legal norm of 1984). Mesh size control (90 mm). Fleet bigger than 10 GRT should operate 1.5 nmi offshore.	Biological parameter estimates to southeastern populations. Applied effort should not be more than last 15 years average (13,700 hauls/year). Studies showed fleet actually operates with lower mesh size.	Weak control.	Drastic biomass reduction of some demersal species (as weakfish). Fishing over nursery grounds. Weak control over forbidden catches until 1.5 nmi offshore. Small individuals caught by shrimp fleet.
Groundfish, bottom longlining	None	None	None	Long-life target species; lack of biological parameter estimates.
Anchovy	After 1990: 90 day fishing season limitation and restricted areas closure, to protect spawning period. Standardization of fishing gear.	Warning on excessive effort, which may result in possible recruitment failure.	Well regulated; effort values near optimum.	
Sharks	No more fishery licenses are given.	Need for specific regulation, better monitoring, regulation of maximum net size (2.5 km), necessity to stop fin extraction and to limit licenses, non-commercial live shark liberation.	Effort down due to strong reduction in catches.	Potential risk of reduction of several shark species (as *Sphyrna lewini*) and also small cetaceans and turtles.

Abbrev.: TED: Turtle excluder device; YFT: Yellowfin tuna *(Thunnus albacares);* BET: Bigeye tuna *(Thunnus obesus).*

[a] A regulation of 1981 forbids seine for tuna fishery.

fish landed in the state of São Paulo. At that time, a minimum condition for data collection and handling was ensured and information was carefully archived. Over time, as budgets, agreements, and vacancies were cut, the statistics control system lost accuracy and collecting and interpreting fishery data became a big challenge. The public resource reduction period was coincidental with the beginning of the first shrimp and sardine fishery crisis. After 1988, the only public fishing warehouse was privatized causing social upheaval in the fishery sector as a result of increasing taxes. It included fleet landing moving to clandestine points and to other state's harbors. This disruption of fleet dynamics and landings made it difficult for the fishery inspection and survey.

Stock assessment has not been a regular practice systematically applied to this fishery; however, the scientific community always made efforts to make recommendations to the federal management.

Present Situation

Fishery-Dependent Data

The Instituto de Pesca has an ongoing interview-logbook database program for the state of São Paulo. Presently, there are only five official data recorders, regularly collecting landing reports on local fishing wharves. Reports follow the traditional approach (as shown in Braga 1962), considering 60 fish categories and a fishing area divided in blocks of 10° on a side. Information is collected on field data sheets, but some industries also provide commercial bills or fishing logs. Improved in 1994, the present sample unit is Santos, Cananeia and Ubatuba region landings.

Fishery data include fishing area, depth, sampled landing catches, and randomly sampled information about fishery effort, which is used to derive CPUE estimates. Effort is measured in number of boats, trips, hauls, fishing hours and hooks, and often number of small-scale fishermen. Table 4 shows, for the main fisheries, the percentage of landed catches for which effort data are computed. Landed catch information is periodically obtained by collectors, interviews, logbooks, and industry commercial bulletins. Logbooks are filled out and sent in only by tuna fleet captains.

Compared with other periods, there was an increase in landing sites and a reduction of data collectors. Table 5 shows the sampling coverage on observed landing sites. These data indicate poor estimates of small-scale activities.

The main problems observed in the evaluation are the sampling strategy due to the increase in landing points and the reduction of research budget and fieldwork personnel.

Onboard discards are unknown, and catchability and selectivity estimates cannot be derived nor can the real fishery impact on the ecosystem be determined.

Table 4. Percentage of landings with known effort information, in the last 5 years.

Sardine purse seine	Pink shrimp trawl	Sea-bob shrimp trawl	Tuna longline
100% (in number of boats)	78% (in fishing hours)	53% (in fishing hours)	100% (in number of hooks)

Demersal fishes trawl	Groundfishes bottom longline	Anchovy seine	Sharks gillnet
85% (in fishing hours)	100% (in number of hooks)	100% (only in number of fishermen)	0%

Table 5. Present sampling coverage in the state of São Paulo.

	Number of licensed boats	Number of estimated operational boats	Number of known landing sites
Fishery			
Industrial	401	421	9
Small-scale	621	>2,500	32

	Landing sites covered by data collectors (%)		Landing sites where biological sampling is done (%)
	Daily	Sporadically	
Fishery			
Industrial	11.1	100	71.3
Small-scale	12.5	15.7	3.1

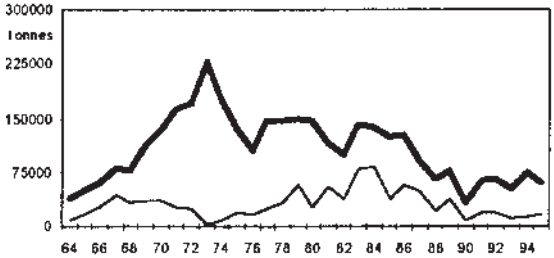


Figure 5. *Brazilian sardine 1964-1994 landings. Dark line = total landings (TL); weak line = São Paulo landings (SP).*

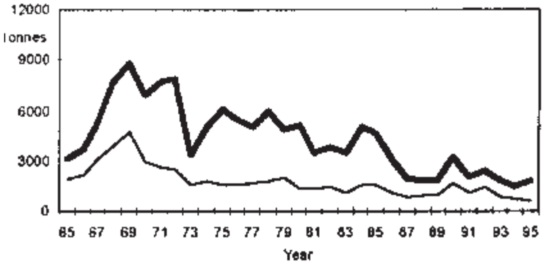


Figure 6. *Pink shrimp 1965-1995 landings, showing total (TL) and São Paulo (SP) landings.*

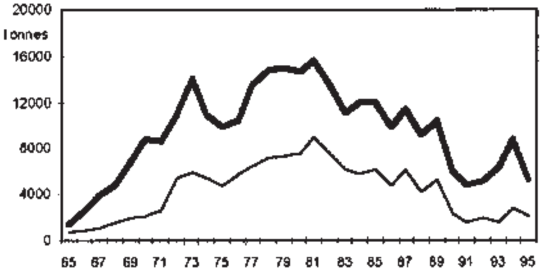


Figure 7. *Sea-bob shrimp 1965-1995 landings, showing total (TL) and São Paulo (SP) landings.*

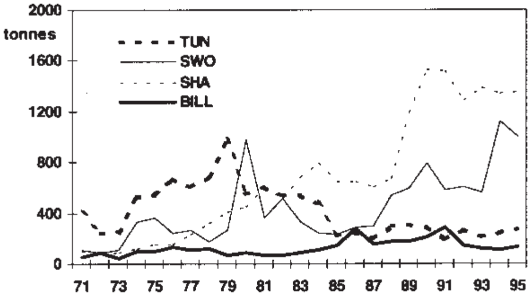


Figure 8. Tuna and tuna-like São Paulo 1971-1995 landings by species/category (TUN = tunas; SWO = swordfish; SHA = sharks, and BILL = billfishes).

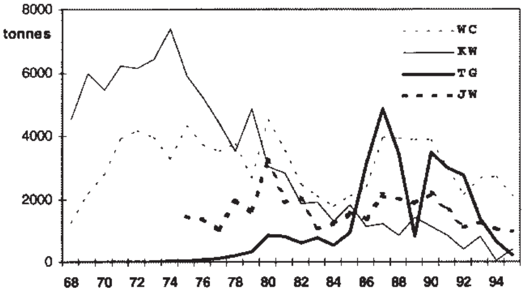


Figure 9. Demersal fish 1968-1995 landings in São Paulo by species (WC = white croaker; KW = king weakfish; TG = triggerfish; JW = Jamaican weakfish).

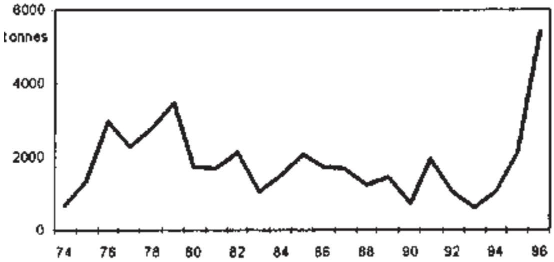


Figure 10. Broadband anchovy 1974-1996 landings in São Paulo.

Data archiving on computers has begun in the last few years. Recently, there is a database project aiming to include complete fleet landing and biological sampling information in Microsoft Access 7.0® framework.

Fishery-Independent Data

Oceanographic cruises also gather fishery survey reports and environmental data, and are conducted by universities (University of São Paulo and University of Rio Grande). Effort is concentrated only on a small part of the stock at any given time. The Instituto de Pesca has carried out several exploratory cruises using several gear types, such as traps, light attraction, sliding net, and tuna and bottom longlining.

Principal Stock Assessment Methods

Some assessment techniques were applied to some stocks from southeastern Brazil, landed in the state of São Paulo, such as sardine, shrimps, croakers, weakfish and swordfish. A Yield Isopleths Analytic Model (Beverton and Holt 1966), a Cohort Analysis (Pope 1972), VPA and a Length Cohort Analysis (Jones 1974) were used to the brazilian sardine stock (SUDEPE/PDP 1985; Cergole 1993). A Surplus Production Model (Schaeffer 1954) and a generalized Production Model (Fox 1970) was applied to the main shrimp and demersal fish stocks, as pink and sea-bob shrimp, croaker and weakfish. Recently, another study from Arfelli (1996) applied a Yield Analysis for Polycohortical Populations (Santos 1992) to the swordfish stock.

Discussion

The historical data collection on the São Paulo fishery can be used to study the evolution of regional fishery problems; however, some critical points must be observed to improve present and future data sampling for stock assessment.

The reduction of research budgets and technical personnel focusing on fishery data collection and the multiplication of landing sites can further damage the São Paulo database program. As Braga pointed out in 1962, the uniformity of future statistical surveys depends on centralized landing operations. At that time, there were only three landing points; present work estimates 9 industrial plus 32 small-scale, excluding extra clandestine and unknown points of landing. For instance, part of the shrimp production is landed at points where there are no data collectors, resulting in underestimates of total catches. Sardine can also be often underestimated.

On the other hand, São Paulo state's fleet dynamics is very complicated and unstable. Studies on fleet behavior have shown that, after one or two years, the situation can completely change (Tiago et al. 1994, Gasalla et al. 1997). Sporadic better conditions or higher prices in other regions can decrease landing values at controlled points, without indicating a downward trend in the stock. The limited knowledge of total fleet size is another limiting factor to fishery analyses; there is poor or nonexistent fleet characterization and effort standardization, as well as fishing power. Not having selectivity or catchability estimates available can result in indices of abundance that poorly reflect stock size.

Another criticism of these data sources for stock assessment are the poor estimation of spatial and temporal scales and inaccurate effort measures. Uncertain total stock catch estimates are due to the lack of data from other states, where often no database is implemented or available. It is also necessary to evaluate precision and accuracy of different data sources and also biased sampling practices.

The development of new data acquisition and archiving strategies should be seriously recommended. We suggest as possible solutions to these problems the implementation of a new sampling program able to define a spatial strategy scheme to derive fleet analysis studies including fluctuation of catchability, better species composition analysis, and effort-related factors, such as efficiency of targeting the resource. To provide useful information for stock assessment and to monitor the health of the stocks, it will be necessary to have this process standardized in the entire area, including the other states of the region. Funding will be necessary to provide education to the fishery sector as well as the implementation of a statistics program.

New Insights

Despite the difficulties observed in fishery science and managing fisheries, we can still set goals for future work on fishery stock assessment:

1. Acknowledge the uncertainty of historical data.
2. Consider new statistical tools to treat uncertainties and probabilities.
3. Search for alternative regional methods, considering the fisheries as a predator and studying its local strategies.
4. Search for alternative management options, such as cooperative management or co-management.
5. Improve ecological knowledge of fisheries, relating target species to other species in the environment, aiming towards broadly based management of the ecosystem.

"We have at present only very limited ability to predict the influence of various factors on the catch of fish; hence we cannot be sure what new kinds of information will improve our predictions, or make management schemes more effective and efficient." (Dickie 1978)

Acknowledgments

We acknowledge the University of Alaska Sea Grant and the two anonymous referees for careful suggestions. This article was based on information from a joint group of researchers from the Instituto de Pesca, without whom it would never have been possible.

References

Amorim, A.F., F.E.S. Costa, L. Fagundes, R. Assumpção, and C.A. Arfelli. 1995. Analysis on albacore *Thunnus alalunga* caught by Santos longliners off south and southeast of Brazil (1971-1994). Coll. Vol. Sci. Pap. ICCAT 59:47-61.

Arfelli, C.A. 1996. Estudo da pesca e aspectos da dinamica populacional do espadarte *Xiphias gladius* L., 1758 no Atlantico Sul. Ph.D. thesis, Universidade Estadual Paulista, Instituto de Biociencias. 175 pp.

Arfelli, C.A., and A.F. Amorim. 1994. Shark fishery from Santos-SP longliners off south and southeastern of Brazil. Proceedings of 4th Indo-Pacific Fish Conference, 28 November to 4 December 1993, Bangkok, Thailand.

Arfelli, C.A., A.F. Amorim, and L. Fagundes. 1997. Sharks and other species composition in the yield of regular and shallow longliners from Santos, São Paulo (1971-95). Anais da I Reunião da Sociedade Brasileira de Estudos de Elasmobrânquios, Ilhéus, BA, Brasil.

Bendazoli, A., and C.L.D.B. Rossi-Wongtschowski (eds.). 1990. A manjuba (*Anchoviella lepidentostole*) no Rio Ribeira de Iguape: Biologia, comportamento e avaliação de estoque. IBAMA/IOUSP/IP-SA/SEMA. 125 pp.

Beverton, R.J.H., and S.J. Holt. 1966. Manual of methods for fish stock assessment. Part 2. Tables of yield fuctions. FAO Fish. Tech. Pap. 38 Rev. 1. 67 pp.

Braga, A.S. 1962 Métodos de compilação e computação de dados estatísticos de desembarque de pescado no porto de Santos. Bolm. Inst. Oceanogr. 12(2):39-64.

Castro, L.A.B., and P.M.G. Castro. 1995. Variação do comprimento médio da pescada foguete, *Macrodon ancylodon*, capturada na região Sudeste do Brasil, 23°S a 29°S, desembarcada em Santos, de 1976 a 1990. Publ. Esp. Inst. Oceanogr. S. Paulo 11:187-196.

Cergole, M.C. 1993. Avaliação do estoque da sardinha-verdadeira, *Sardinella brasiliensis*, na costa sudeste do Brasil período 1977-1990. Ph.D. thesis, Univ. São Paulo, Instituto Oceanográfico. 245 pp.

Costa, F.E.S., F.M.S. Braga, A.F. Amorim, and C.A. Arfelli. 1995. Fishery analysis on shortfin mako, *Isurus oxyrinchus*, off southeast and south of Brazil (Elasmobranchii, Lamnidae). Arq. Ciênc. Mar 30(1-2):5-12.

Dias, J., and L.D.C. Dornelles. 1996. Diagnostico da pesca marítima do Brasil. Col. Meio Ambiente, Sér. Estudos Pesca 20. 163 pp.

Dickie, L.M. 1978. Mathematical models for growth. In: T.B. Bagenal (ed.), Methods for assessment of fish production in fresh waters. IBP Handbook 3, 3rd edn. Blackwell, Oxford, pp. 130-136.

Fox, W.W. 1970. An exponential surplus-yield model for optimizing exploited fish populations. Trans. Am. Fish. Soc. 99:80-88.

Gasalla, M.A., P.M.G. Castro, and M.H. Carneiro. 1997. Comportamento da frota de arrasteiros de parelha de Santos: Uma análise da CPUEs entre 1993 e 1995. XII Encontro Brasiliero de Ictiologia, 24 to 28 of February, São Paulo, SBI/IOUSP (abstracts), pp. 315.

IBAMA. 1992. Sardinha. Col. Meio Ambiente, Sér. Estudos Pesca 4. 41 pp.

IBAMA. 1993. Peixes Demersais. Col. Meio Ambiente, Sér. Estudos Pesca 8. 93 pp.

IBAMA. 1995. Peixes Demersais. Col. Meio Ambiente, Sér. Estudos Pesca 16. 126 pp.

Jones, R. 1974. Assessing the long term effects of changes in fishing effort and mesh size from length composition data. ICES CM 1974/F:33. 13 pp.

Lin, C.F. 1992. Atuns e afins: Estimativa da quantidade de isca viva utilizada pela frota atuneira. Col. Meio Ambiente, Sér. Estudos de Pesca 6. 80 pp.

Neiva, G.S., and S.J.C. Moura. 1977. Sumário sobre a explotação de recursos marinhos do litoral brasileiro: Situação e perspectivas. Sér. Documentos Ocasionais SUDEPE/PDP 27:1-44.

Pope, J.G. 1972. An investigation of the accuracy of virtual population analysis using cohort analysis. Res. Bull. ICNAF 9:65-74.

REVIZEE. 1997. Estatística da Pesca 1995. MMA/IBAMA/CEPENE. 33 pp.

Richardson, I.D., and M.N. Moraes. 1960. A first appraisal of landing and mechanism of the Santos fishery. Bolm. Inst. Oceanogr. 9(1):5-85.

Ripley, W.E. 1956. Relatório ao Governo do Brasil sobre biologia da pesca. Relatório FAO/ETAP 494. 23 pp.

Rossi-Wongtschowski, C.L.D.B., S.A. Saccardo, and M.C. Cergole. 1995. Situação do estoque da sardinha no litoral sudeste e sul do Brasil. Col. Meio Ambiente, Sér. Estudos Pesca 17. 44 pp.

Santos, E.P. 1992. Um modelo matemático em dinamica de populacions I. Bolm. Inst. Pesca 19:97-102.

Schaefer, M. 1954. Some aspects of the dynamics of populations important to management of the commercial marine fisheries. Bull. I. ATTC/Bolm. CIAT 1(2):22-56.

SEMA/SP. 1996. Plano de Gerenciamento Costeiro do Estado de São Paulo. Secretaria de Meio Ambiente, São Paulo. 129 pp.

Silva, A.O.A. 1996. Idade, crescimento, mortalidade e aspectos reprodutivas do pargo, *Pagrus pagrus* (Teleostei: Sparidae), na costa do Estado de São Paulo e adjacências. M.Sc. dissertation, Univ. São Paulo, Instituto Oceanográfico. 116 pp.

SUDEPE. (no date). III Plano Nacional de Desenvolvimento da Pesca 1975-1979. Ministerio da Agricultura, Brasil. 181 pp.

SUDEPE/PDP. 1985. Relatório da segunda Reunião do Grupo de Trabalho e Treinamento (GTT) sobre Avaliação de Estoques, realizada em Tamandaré/PE, de 29 de junho a 24 de julho de 1981. Sér. Documentos Técnicos 34. 439 pp.

Tiago, G.G., S.L.S. Tutui, R.W. von Seckendorff, R.T.B. Grassi, and M.L.S. Inácio. 1994. Análise da frota pesqueira sediada em Ubatuba, Estado de São Paulo. Bolm. Inst. Pesca 22(2):71-83.

Tomás, A.R.G. 1996. Biologia e Pesca do Falso Congro Rosa, *Ophiodon holbrooki* (Putnam, 1874) no Sudeste do Brasil. M.Sc. dissertation, UNESP/Inst. Biociencias. 274 pp.

Valentini, H., and R.D. Cardoso. 1991. Análise da pesca da sardinha-verdadeira, *Sardinella brasiliensis,* na costa sudeste-sul do Brasil. Atlantica 13(1):45-54.

Valentini, H., P.M.G. Castro, G.J.M. Servo, and L.A.B. Castro. 1991a. Evolução da pesca das principais espécies demersais da costa sudeste do Brasil, pela frota de arrasteiros de parelha baseada em São Paulo, de 1968 a 1987. Atlantica 13(1):87-95.

Valentini, H., F. D´Incao, L.F. Rodrigues, J.E. Rebelo Neto, and L.G. Domit. 1991b. Análise da pesca do camarão sete-barbas (*Xiphopenaeus kroyeri*) nas regiões sudeste e sul do Brasil. Atlantica 13(1):171-177.

Valentini, H., F. D´Incao, L.F. Rodrigues, J.E. Rebelo Neto, and E. Rahn. 1991c. Análise da pesca do camarão-rosa (*Penaeus brasiliensis* e *P. paulensis*) nas regiões sudeste e sul do Brasil. Atlantica 13(1):143-157.

Vazzoler, A.E.A., and G. Vazzoler. 1969. Relatório sobre os conhecimentos atuais referentes à pesca marítima no Estado de São Paulo. Contrs. Inst. Oceanogr. Univ. S. Paulo, Sér. Ocean. Biol. 16:1-28

Zavala Camin, L.A., and J.N. Antero da Silva. 1991. Histórico da pesquisa e pesca de atuns com espinhel no Brasil. Atlantica 13(1):107-114.

Evaluation of CPUE Estimates for the 1995 Crab Fisheries of the Bering Sea and Aleutian Islands Based on Observer Data

Laurence C. Byrne and Douglas Pengilly
Alaska Department of Fish and Game, Kodiak, Alaska

Abstract

Observers on board fishing vessels have gathered data from Bering Sea, Aleutian Islands, and western Gulf of Alaska crab fisheries since 1988. That data has provided estimates of catch per unit effort (CPUE) for both the directed catch and bycatch during the commercial fisheries. Because observers can sample only a small portion of the total effort from an observed boat and because observers are present only on the catcher-processor component of the commercial fleet in some fisheries, an evaluation of the precision and accuracy of fishery-wide CPUE estimates derived from observer pot samples is needed. We present such an evaluation for observed crab fisheries prosecuted in 1995. CPUE estimates and their standard errors were generated from data collected by observers from 27,627 pots sampled in 1995 fisheries and stratified by vessel-days. Estimates of mean CPUE for retained legal crabs were within 8% of CPUEs generated from confidential interviews in 10 of the 13 fisheries. Estimates of mean CPUE for retained legal crabs were within 8% of CPUEs generated from fish ticket delivery records in two of the five fisheries with less than 100% observer coverage. Differences larger than 8% between observer-based estimates and confidential interviews or fish ticket data were largely attributed to the small number of pots sampled by observers in those fisheries. Standard errors were in general very small, indicating the CPUE estimates were precise. Standard errors of CPUE estimates generated from bootstrap simulations were all equal to or less than the standard errors computed analytically. Bootstrapping also indicates that the sampling distributions for CPUE estimates of retained-legal crabs and commonly encountered bycatch are unimodal and symmetric.

Introduction

Observers on board fishing vessels have gathered data from Bering Sea, Aleutian Islands, and western Gulf of Alaska crab fisheries since 1988. Development of a mandatory observer program for crab fisheries in Alaska was initiated primarily to assure that the sex and size-limit restrictions in king (Lithodidae) and Tanner *(Chionoecetes bairdi)* crab fisheries were enforced on at-sea processing vessels. Whereas Alaska Department of Fish and Game (ADFG) dockside samplers could assess the crabs delivered to shore-based processors, prior to 1988 ADFG had no authority or program to sample the crabs that were processed at sea. By 1990, observers were required on all catcher-processor vessels (C/Ps) or floating-processor vessels processing king crab or any species of *Chionoecetes* at sea in Alaska.

Collection of biological or fishery data was a secondary motivation at the inception of Alaska's crab fishery observer program. However, as well as collecting data from the catch retained for processing, crab fishery observers afforded the opportunity to record information from the non-retained catch by sampling the catch of randomly selected pot lifts from C/Ps. The utility of observers in collecting biological and fishery data was soon apparent and by the mid-1990s observers were also required on all vessels fishing in the Korean hair crab *(Erimacrus isenbeckii)* fishery; all vessels targeting scarlet king crab *(Lithodes couesi),* grooved Tanner crab *(Chionoecetes tanneri),* or triangle Tanner crab *(C. angulatus);* and all vessels participating in the red king crab *(Paralithodes camtschaticus)* and brown (or "golden") king crab *(Lithodes aequispina)* fisheries in the Aleutian Islands. Data collected from those fisheries by observers provided information that could not otherwise be obtained, including information on the geographic and temporal distribution of fishery effort, on levels and distribution of directed catch and bycatch, on the biological characteristics (e.g., size, sex, and reproductive condition) of discarded bycatch, on fishing methods (e.g., soak times and escape mechanism employed), and on the effectiveness of gear modifications in reducing bycatch.

Data on catch rates and sex-size-species composition of discarded bycatch crabs has been of particular interest to fishery regulators and researchers. Effort during an Alaska crab fishery season is typically directed on only one species and participants are permitted to retain only one or a few species. Only males may be lawfully harvested and a minimum size for legal retention is established for each exploited species in an area. All females, undersized males and non-permitted species must be immediately returned unharmed to the sea. Total harvest and catch rates (crabs per pot lift, CPUE) of retained legal males has been well measured in the Alaska commercial crab fisheries due to the state's requirement that catch and effort (pot lifts) for each landing be recorded on an ADFG fish ticket, coupled with the state's program of sampling catch and interviewing skippers at deliveries. Prior to the initiation of the crab observer program, however, information on discarded bycatch was not available to fishery managers.

CPUE and composition of the bycatch is of interest for two reasons. The first is that it provides the information necessary to address Alaskan and federal government intentions to minimize bycatch. The second is that it provides information on female and undersized males that is not available from some unsurveyed crab stocks. Since 1993, virtually every consideration by the Alaska Board of Fisheries (BOF) of regulations pertaining to area closures, season openings, size limits, gear restrictions, and gear conflicts in Bering Sea and Aleutian Islands crab fisheries has been directly informed by observer data on bycatch. CPUE of bycatch as estimated from observer data has, for example, been used to assess the effectiveness of area closures in reducing red king crab bycatch in the Bering Sea Tanner crab fishery (Tracy and Pengilly 1996) and as input to fishery models used to assess harvest strategies in the Bristol Bay red king crab fishery (Zheng et al. 1997a, 1997b). As observer-based estimates of bycatch CPUE become increasingly relied upon in the development of regulations, management decisions, and management strategies for Alaska crab fisheries, the reliability of those estimates needs to be examined. That is particularly true in cases where observers are present on only a distinct component of the fleet, the C/Ps.

We present information on the variability (standard errors and coefficients of variation) and sampling distribution of crab CPUE estimates that were based on observer pot sampling from fisheries prosecuted in the Bering Sea, Aleutian Islands, and western Gulf of Alaska during 1995. The goal here is to help fishery managers, researchers, and regulatory bodies judge the reliability of the historic bycatch CPUE estimates that they use. We also examine the legal crab CPUE estimates, their estimated standard errors, and estimated sampling distribution in comparison with the CPUE of legal crab computed from fish tickets and skipper interview data. We assume that the CPUE, computed from fish tickets and skipper interview data, is the actual CPUE against which the bias and sampling error of the observer-based estimate can be directly examined. Although the CPUE estimates of bycatch crab are generally of interest to users of the observer data, there are no independent estimates of bycatch CPUE that allow for an assessment of bias and sampling error. The results for the legal crab estimates should, however, provide some guidance on the conditions that affect the reliability of the bycatch CPUE estimates and assist in interpreting the estimates of crab bycatch presented in annual crab observer database reports (e.g., Boyle et al. 1996). Our results also have relevance to the planning and designing of observer pot-sampling protocols in the future. Nonetheless, our goal here is not to determine optimal pot-sampling procedures for observers, because that determination must consider other factors and other priorities for observer deployments that are beyond the scope of this paper.

Table 1. Vessel participation, observer coverage, total pot lifts, and number of pot lifts sampled by observers in Aleutian Islands, Bering Sea, and western Gulf of Alaska crab fisheries during 1995.

Fishery	Total vessels	Observed vessels	Total potlifts	Sampled potlifts
Bering Sea snow crab[a]	253	19	506,802	1,530
Bering Sea Tanner crab[a]	196	11	247,853	421
St. Matthew blue king crab[a,b]	90	2	48,560	47
Pribilof red and blue king crab[a]	119	1	107,521	34
Adak brown king crab[a,c]	34	4	319,006	1,430
Adak red king crab[d]	6	6	2,265	263
Dutch Harbor brown king crab[d]	17	17	65,732	2,512
Bering Sea Korean hair crab[d]	21	21	447,555	10,117
Bering Sea grooved Tanner crab[d]	8	8	60,069	4,407
Eastern Aleutian grooved Tanner crab[d]	7	7	71,931	4,216
Western Aleutian grooved Tanner crab[d]	5	5	11,324	1,096
South Peninsula grooved Tanner crab[d]	7	7	7,248	925
Eastern Aleutian triangle Tanner crab[d]	1	1	Confidential	629

[a] Observer coverage mandatory only on catcher-processors (C/Ps).

[b] Observer also deployed on one catcher-only vessel.

[c] Observers also deployed on two catcher-only vessels.

[d] 100% observer coverage mandatory.

Methods

Data Collection

ADFG managed 13 Bering Sea, Aleutian Islands, and western Gulf of Alaska crab fisheries during 1995 in which observers were present on at least some fishing vessels (Table 1). Effort in those 13 fisheries was targeted on eight species: red king crab, blue king crab *(Paralithodes platypus),* brown king crab, Tanner crab, snow crab *(Chionoecetes opilio),* grooved Tanner crab, triangle Tanner crab, and Korean hair crab. In aggregate, those 13 fisheries covered an area extending 1,600 km east-west and 1,000 km north-south (Fig. 1).

Prior to deployment of observers, the ADFG observer coordinator established a target daily number of pots to sample and observers were instructed to randomly select and sample that number of pots each day from all periods during which fishing occurred. The target daily number of pots to sample was determined for each fishery on the basis of the anticipated time available for pot sampling given other duties and the anticipated time required to sample a pot. For example, a target of only

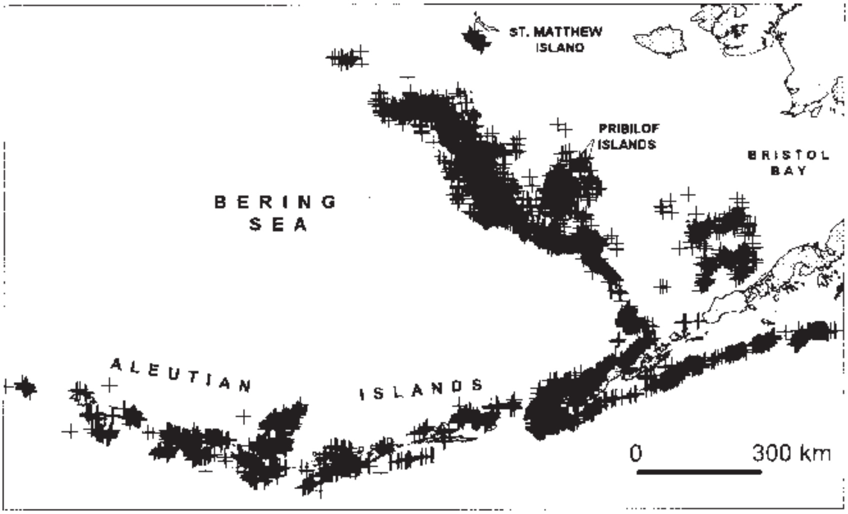


Figure 1. Location of pots sampled by observers in Aleutian Islands, Bering Sea, and western Gulf of Alaska crab fisheries during 1995.

four pots per day was established for observers stationed on C/Ps in the Bering Sea snow crab fishery, because pot sampling in that fishery is time consuming due to the typically high CPUE and because observers on C/Ps have numerous other duties, such as sampling crabs retained for processing. On the other hand, a daily target of 40 pots was established for observers in the Bering Sea Korean hair crab fishery, because vessels in that fishery typically perform a large number of pot lifts with low CPUE each day and because the observers had no duties associated with monitoring the processing of crabs. The number of pots sampled did occasionally vary from the daily target due to, for example, weather conditions or limited effort for the day.

Observers categorized and enumerated by species the catch from each sampled pot. All commercial crab species were also categorized and enumerated by sex and legal status. Although observers collect data on all species captured and on other biological attributes of captured crabs, in this paper we present only the results for data collected on CPUE of legal and bycatch crabs from the targeted species.

Data Analysis

We use CPUE to denote the average number of crabs caught per pot lift for an entire fishery season for some defined component of the fleet. We also make a distinction between the "observed fleet" and the "total fleet" for fisheries in which observer coverage is less than 100%. In no fishery considered here with less than 100% observer coverage (Table 1), can the observed fleet be considered a random sample of the total fleet. The observed fleets in the Bering Sea snow crab, Bering Sea Tanner crab, and Pribilof red and blue king crab fisheries were made exclusively of the relatively few participating C/Ps. In the St. Matthew blue king crab and Adak brown king crab fisheries, the observed fleets also included catcher-vessels that voluntarily carried observers. In cases where observer coverage is less than 100%, there are two sources of error if CPUE estimated from observer pot samples is used to estimate the actual CPUE for the total fleet: differences between the actual CPUEs of the observed fleet and the total fleet; and, as exists with 100% observer coverage, differences between the actual CPUE of the observed fleet and that estimated from the pot samples.

Because a target number of pots to sample each day from each observed vessel was stipulated in the observer sampling procedures, the estimates of CPUE for the observed fleet and their standard errors (SE) were computed according to a stratified sampling model for estimation of a population mean (Cochran 1977, equations 5.1 and 5.12) in which each vessel-day combination is a stratum. Details on application of the stratified sampling model in computation of estimated CPUE and its standard error in this application are provided in Boyle et al. (1996). Coefficients of variation (CV, the estimate divided by its standard error and expressed as a percentage) were computed as standardized measures of the estimator's variability.

The sampling distribution of the CPUE estimates for the observed fleet was assessed by resampling the data with replacement ("bootstrapping"; Efron and Tibshirani 1993). Resampling was performed in accordance with the stratified sampling model by resampling at the level of the vessel-day strata. That is, each bootstrapped estimate of CPUE required resampling with replacement from the data on catch per pot lift within each vessel-day up to the number of pots actually sampled on the vessel-day. The bootstrap process of computing CPUE was replicated 500 times to provide an estimated sampling distribution, estimated mean, and estimated standard error that could be compared with the CPUE estimate and standard error computed analytically from the stratified sampling model.

Daily confidential interviews with vessel skippers performed by onboard observers provided data on daily effort and catch of retained legal crabs for each observed vessel. The CPUE of retained legal crabs computed from confidential interviews served as the standard for assessing error and bias in the observer-based estimates of CPUE for the observed fleet. We refer to the retained legal CPUE estimated from confidential interview data as the "actual observed fleet CPUE" (AOF CPUE). In fisheries with 100% observer coverage, the AOF CPUE also provides an estimate for the total fleet. In those fisheries with less than 100% observer coverage (Table 1), ADFG fish tickets, supplemented with confidential interviews of skippers by dockside samplers, provided data on CPUE of retained legal crabs for the total fleet. We refer to the CPUE of retained legal crabs estimated from the fish ticket data for all fishery landings as the "actual total fishery CPUE" (ATF CPUE). We expressed the difference between the actual CPUE and that estimated from observer data as a percent error:

$$Percent\ error = \frac{CPUE_{ObserverData} - CPUE_{Actual}}{CPUE_{Actual}} \times 100\%$$

where Actual denotes either the AOF or ATF CPUE estimate.

Results

Retained legal crab CPUEs based on observer data were within 8% of the AOF CPUE value in all but one (the Adak red king crab fishery) of the eight fisheries with 100% observer coverage and in two (the Bering Sea Tanner crab and Bering Sea snow crab fisheries) of the five fisheries with less that 100% observer coverage (Fig. 2). Of the 13 fisheries sampled, the three fisheries showing the largest discrepancies between the AOF CPUE and the observer-based estimates were those with the lowest numbers of potlifts sampled by observers (34 pots in the Pribilof red and blue king crab fishery, 47 pots in the St. Matthew blue king crab fishery, and 263 pots in the Adak red king crab fishery). Not surprisingly, observer pot samples also provided a poor estimate of the ATF CPUE in the Pribilof blue king crab fisheries. Notably, however, the estimated CPUE for retained legal males in the Adak brown king crab fishery was within 3% of the AOF CPUE but differed by 37% from the ATF CPUE.

Coefficients of variation (CVs) for estimated CPUEs of retained legal crabs and sublegal males were generally 5% or less, indicating high precision (i.e., low variability) of estimates (Fig. 3). Precision of CPUE estimates of both legal and bycatch crabs tended to increase with increasing numbers of pots sampled. Within fisheries, CVs for estimated CPUE of females were generally much higher than for those of males and CVs for estimated CPUE of retained legal crabs tended to be lower than those for sublegal-sized male crabs.

Bootstrapped estimates of mean CPUE for retained legal males and their standard errors are presented in Table 2. Bootstrapped mean estimates differed from those computed analytically by no more than 0.08 crab per pot. Bootstrapped standard errors were lower than or equal to the standard errors computed analytically.

The sampling distributions of estimated CPUE based on 500 bootstrapped samples show three general patterns (Fig. 4). The first is typified by the south Peninsula grooved Tanner crab retained legal males data (Fig. 4a). The sampling distribution in this case is approximately normal and is

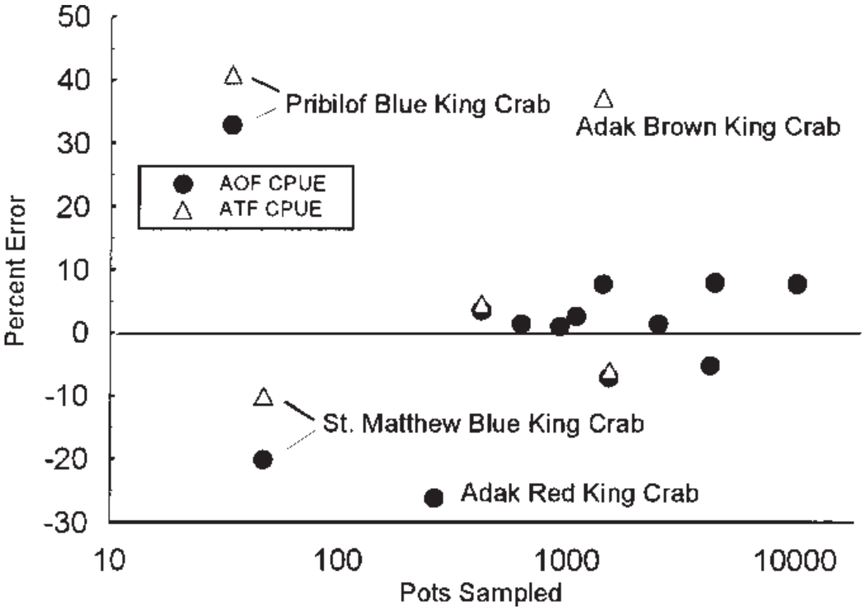


Figure 2. Percent error between CPUEs for retained legal males as estimated from observer pot-sampling data and the actual values as computed from confidential interviews and fish ticket data in 13 Aleutian Islands, Bering Sea, and western Gulf of Alaska crab fisheries during 1995. Black dots denote percent errors relative to actual observed fleet CPUEs (AOF CPUE). Open triangles denote percent errors relative to actual total fishery CPUE (ATF CPUE) for those fisheries in which observer coverage was less than 100%. Fisheries in which the error was greater than 8% are labeled.

well characterized by the estimated CPUE and its standard error. In addition, the AOF CPUE is within two standard errors of the estimated value. The second pattern is exemplified by the Bering Sea snow crab retained legal males data (Fig. 4b). The sampling distribution in this example is approximately normal and well characterized by the estimated CPUE and its standard error; however, the AOF CPUE is more than four standard errors from the estimated value and well outside of the estimated sampling distribution. Out of the 13 observed fisheries, the AOF CPUE fell outside of the estimated sampling distribution in five cases and was more than two standard errors from the estimated value in six cases. The estimated sampling distributions of CPUE were closely approximated by a normal distribution in all cases except that of female blue king crab CPUE in the Pribilof red and blue king crab fishery (Fig. 4c). The skewed, multimodal distribution in this example indicates limited inferential value of these data.

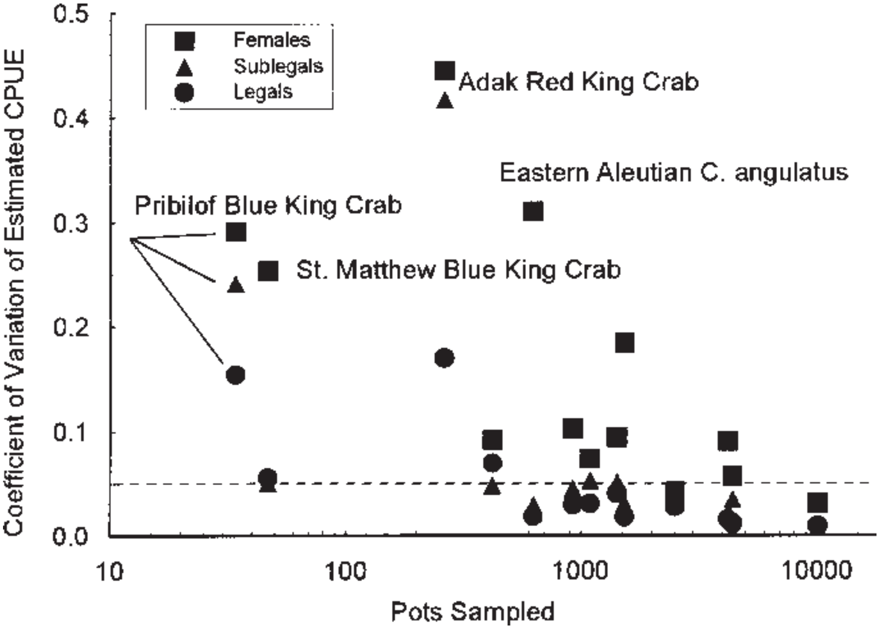


Figure 3. Coefficients of variation of observer-based CPUE estimates for female, sublegal, and legal crabs of the targeted species in 13 Aleutian Islands, Bering Sea, and western Gulf of Alaska crab fisheries in 1995.

Discussion

CPUE estimates computed from 1995 observer pot sample data for retained legal crabs during the Bering Sea snow crab, Bering Sea Tanner crab, Adak brown king crab, Dutch Harbor brown king crab, Bering Sea Korean hair crab, Bering Sea grooved Tanner crab, eastern Aleutian area grooved Tanner crab, western Aleutian area grooved Tanner crab, south Peninsula grooved Tanner crab, and the eastern Aleutian triangle Tanner crab fisheries were within 8% or less of the actual value for the observed portion of the fleet. That close agreement between the observer-based CPUE estimates for retained legal crab and the actual values in each of those 10 fisheries suggests that observer data should also provide reasonably accurate, low-bias estimates of bycatch CPUE for the observed portion of the fleet in those fisheries. CPUE estimates computed from observer pot sample data for retained legal crabs in each of the St. Matthew blue king crab, the Pribilof king crab, and the Adak red king crab fisheries differed from the actual observed fleet value by more than 20%, indicating lower reliability of observer data in providing catch rate estimates for the observed fleets in those fisheries. Not surprisingly, large relative errors in

Table 2. Comparison of observer-based CPUE estimates and their standard errors computed analytically from the stratified sampling model with bootstrapped estimates.

	Computed Analytically		Bootstrapped	
Fishery	CPUE	SE	CPUE	SE
Bering Sea snow crab	110.00	1.91	109.92	1.68
Bering Sea Tanner crab	7.92	0.55	7.93	0.46
Adak brown king crab	6.62	0.23	6.62	0.19
Adak red king crab	2.19	0.37	2.19	0.33
Dutch Harbor brown king crab	6.26	0.17	6.25	0.15
Bering Sea Korean hair crab	3.50	0.03	3.50	0.03
Bering Sea grooved Tanner crab	8.19	0.09	8.18	0.09
Eastern Aleutians grooved Tanner crab	6.73	0.10	6.74	0.09
Western Aleutians grooved Tanner crab	6.79	0.20	6.79	0.19
South Peninsula grooved Tanner crab	85.15	1.87	85.10	1.63

CPUE estimates are for retained legal males in the Aleutian Islands, Bering Sea, and western Gulf of Alaska fisheries with observer coverage during 1995. Estimates for the St. Matthew blue king crab, Pribilof blue and red king crab, and eastern Aleutian triangle Tanner crab fisheries are not presented due to confidentiality of data.

estimating retained legal CPUE were associated with the low numbers of sampled pots.

Observers were present on 100% of the vessels fishing in the Dutch Harbor brown king crab, Adak red king crab, Bering Sea Korean hair crab, Bering Sea grooved Tanner crab, eastern Aleutian grooved Tanner crab, western Aleutian grooved Tanner crab, south Peninsula grooved Tanner crab, and the eastern Aleutian triangle Tanner crab fisheries. Hence, the conclusions stated in the above paragraph on accuracy of CPUE estimates relative to the observed fleet in those eight fisheries also pertain to the total fleet for those fisheries. Observers were deployed only on C/Ps in the Bering Sea Tanner crab and snow crab fisheries, but their data were sufficient to provide reasonably accurate (within 6% of the actual value) estimates of retained legal CPUE for the entire fleet. Presumably, the observer data from those two fisheries also provides reasonably accurate estimates of the fleet-wide bycatch CPUEs. On the other hand, partial observer coverage in the Pribilof red and blue king crab and the Adak brown king crab fisheries provided very poor estimates of retained legal CPUE for the total fleet and can be assumed to have provided poor bycatch CPUE estimates for the total fleet in those fisheries. In fact, observer coverage in the Pribilof king crab fishery did not provide any information on

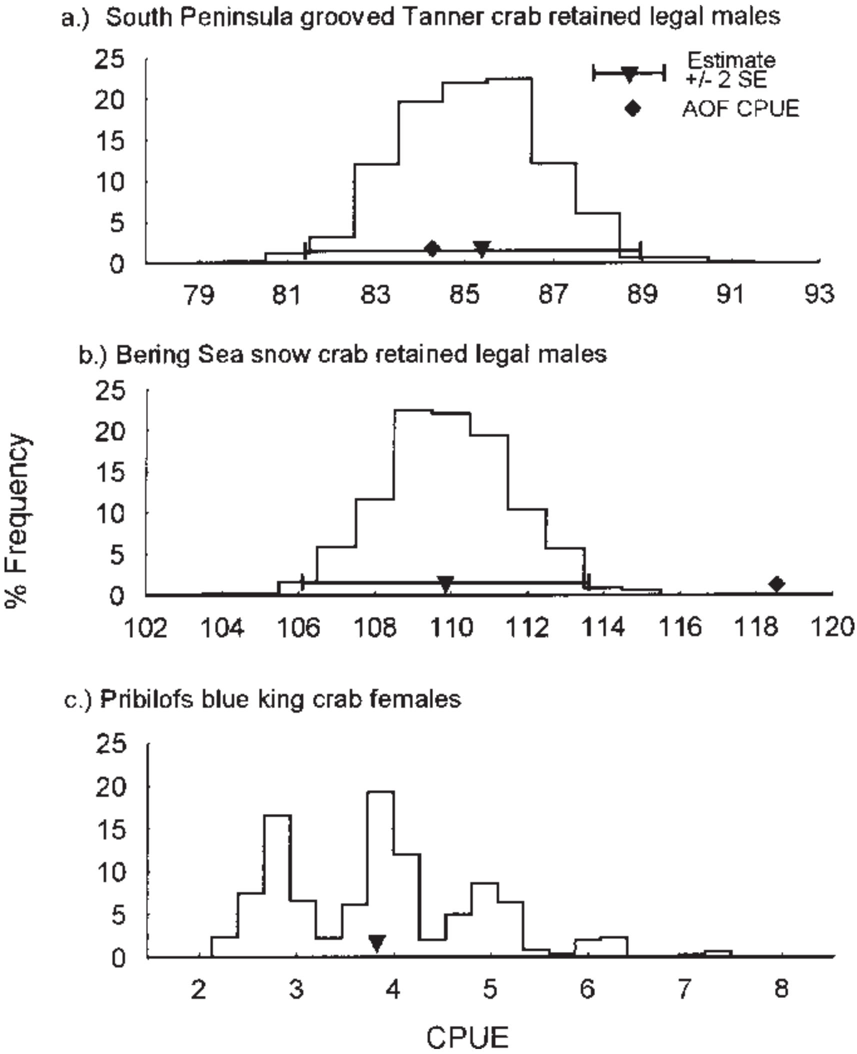


Figure 4. *Examples of bootstrapped sampling distributions of observer-based CPUE estimates for three fisheries selected from the 13 Aleutian Islands, Bering Sea, and western Gulf of Alaska fisheries with observer coverage in 1995: (a) South Peninsula grooved Tanner crab retained legal male CPUE; (b) Bering Sea snow crab retained legal male CPUE; and (c) Pribilof blue king crab female CPUE.*

the catch rate of red king crab, one of the two species of king crabs harvested during that fishery.

The poor performance of the observer data in estimating CPUE for the total fleet is not surprising in the case of the Pribilof king crab fishery. Only one vessel carried an observer during that fishery and the pot-sampling data provided a poor estimate of CPUE for that single vessel. The more interesting case is the Adak brown king crab fishery, in which the observer-based CPUE estimate differed from the actual CPUE for the total fleet by 37%. Although observer data was adequate to characterize the CPUE for the four vessels that constituted the observed portion of the fleet in the Adak area brown king crab fishery, the partial observer coverage in that fishery was not adequate to reliably characterize catch rates for the entire fishery. The percentage of vessels with observer coverage during the Adak brown king crab fishery was actually higher (12%) than that for either of the Bering Sea snow crab (8%) and Tanner crab (6%) fisheries. The poor performance of partial observer coverage in the Adak brown king crab fishery most likely reflects the inability of partial observer to adequately sample a fishery prosecuted over a vast area during a season lasting 9.5 months.

The CVs of most observer-based CPUE estimates for the observed fleet were low, indicating high precision (low variability due to sampling error) of the estimates. As would be expected, higher CVs for CPUE estimates were associated with low numbers of pots sampled. In each fishery, the CV for the estimated CPUE of females were higher than those for the estimated CPUE of male crabs, particularly the retained legal males. Hence, users of the observer-based bycatch CPUE estimates should be aware that CPUE of bycatch females is estimated with less precision than for the targeted legal crabs. The generally higher CVs for female CPUE may be attributable to a more aggregated distribution of the incidentally captured females (e.g., Stone et al. 1992). The CVs presented in this paper are generally lower than those reported by Turnock and Karp (1997) who investigated observer-based estimates of the mean number of salmon bycatch per haul in the 1995 Aleutian Islands and Bering Sea trawl fisheries for pollock. Differences may be attributable to the fact that Turnock and Karp reported on bycatch species, whereas we limited our discussion to the bycatch of the targeted species.

Bootstrapped standard errors of CPUE were in close agreement with those computed analytically from the stratified sampling model, indicating that estimates of standard errors can be directly computed from the data. Also, the bootstrapped sampling distributions of CPUE were, with only one exception, well approximated by the a normal distribution, the estimated CPUE, and its standard error. Notably, however, the CPUE estimates based on observer data for retained legal crabs in the Bering Sea snow crab, Bering Sea Korean hair crab, Bering Sea grooved Tanner crab, and eastern Aleutian area grooved Tanner crab fisheries differed from the actual CPUE for the observed fleet by more than three standard errors. It remains unclear if that is due to some slight biases that may exist in the observer-based estimates or to slight errors in the confidential interviews that we assumed provided the actual CPUE values.

Acknowledgments

The reader may refer to this paper as contribution number PP-164 of the Alaska Department of Fish and Game, Commercial Fisheries and Management and Development Division, Juneau. This paper was funded in part by cooperative agreement #NA67FN0273 between the Alaska Department of Fish and Game and the National Oceanic and Atmospheric Administration. We thank Ivan Vining of the Alaska Department of Fish and Game, Kodiak, for his help with programming for the bootstrapping procedure.

References

Boyle, L., L.C. Byrne, H. Moore, and D. Pengilly. 1996. Summary of the 1995 mandatory shellfish observer program database. Alaska Department of Fish and Game, Regional Information Report 4K96-47, Kodiak.

Cochran, W. 1977. Sampling techniques. Wiley and Sons, New York.

Efron, B., and R.J. Tibshirani. 1993. An introduction to the bootstrap. Chapman and Hall, New York.

Stone, R.P., C.E. O'Clair, and T.C. Shirley. 1992. Seasonal migration and distribution of female red king crabs in a southeast Alaska estuary. J. Crustac. Biol. 12(4):546-560.

Tracy, D.A., and D. Pengilly. 1996. Report to the Alaska Board of Fisheries: A review of incidental red king crab bycatch in the Bering Sea Tanner fishery between 1990 and 1995. Alaska Department of Fish and Game, Regional Information Report 4K96-4, Kodiak.

Turnock, J., and W.A. Karp. 1997. Estimation of salmon bycatch in the 1995 pollock fishery in the Bering Sea/Aleutian Islands: A comparison of methods based on observer sampling and counts of salmon retained by fishing vessel and processing plant personnel. Report to North Pacific Fishery Management Council Observer Advisory Committee, June 1997. National Marine Fisheries Service, Seattle.

Zheng, J., M.C. Murphy, and G.H. Kruse. 1997a. Analysis of harvest strategies for red king crab*, Paralithodes camtschaticus*, in Bristol Bay, Alaska. Can. J. Fish. Aquat. Sci. 54(5):1121-1134.

Zheng, J., M.C. Murphy, and G.H. Kruse. 1997b. Alternative rebuilding strategies for the red king crab *Paralithodes camtschaticus* fishery in Bristol Bay, Alaska. J. Shellfish Res. 16(1):205-217.

Catch and Effort Analysis with Uncertain Stock and Effort Dynamics: Southern Bluefin Tuna Longline

Robert Campbell
CSIRO (Commonwealth Scientific and Industrial Research Organisation)
Division of Marine Research, Hobart, Australia

Abstract

Of interest in recent years has been how best to construct annual indices of stock abundance for the southern bluefin tuna (SBT) stock which adequately account for the spatial and temporal dynamics of the fishery. This fishery has been characterized by a continuous decline in the size of the spawning stock since the 1960s, the introduction of restrictive catch quotas in the mid-1980s with concomitant reductions in the spatial distribution of effort, and a recovery in the abundance of juveniles during the 1990s. In particular, the spatial and temporal distribution of the fishing effort for SBT has declined appreciably since the mid-1980s and this has resulted in a problem in the use of catch and effort data to calculate indices of SBT abundance. Basically there are no data to assess the present state of the stock in large areas in which the stock was fished historically. Investigations of the spatial distribution of fishing effort also indicate the possibility of increased spatial targeting of localized regions of higher catch rates. Nevertheless, due to the lack of fishery independent indices of abundance, many aspects of the underlying stock and effort dynamics remain uncertain. In order to account for these present uncertainties, a range of models, based on various hypotheses of the underlying stock and effort dynamics, have been developed for estimation of the CPUE-based annual indices of stock abundance.

Introduction

Commercial catch and effort data and the resulting catch-per-unit-effort (CPUE) are routinely used in assessing the impact of commercial fishing operations on stock abundance. However, many problems persist in the interpretation of catch rates and their use in estimating abundance indices for exploited fish populations (Paloheimo and Dickie 1964, Gulland 1974, Hilborn and Walters 1992). These problems include inadequate data coverage of the spatial extent of the stock and the more general issues of whether data from a commercial fishery, which attempts to maintain a high catch rate, truly reflects stock abundance.

There can be little doubt that fishing success is influenced markedly by a fisher's prior knowledge of the fishery, by the degree of variability in the spatial distribution of the fish stock, and possibly also by the fishing success of other vessels in surrounding areas. Many fish aggregate in localized areas due to currents, nutrient levels, water temperature, thermocline depth and proximity to land masses, and fishers use their knowledge of these areas to target their fishing activities. High catch rates are often maintained by fine-scale targeting of the resource that can mask declines in the stock. Continual improvements in fish finder technologies have no doubt exacerbated this problem.

While the use of fine-scale spatial catch and effort data can assist in overcoming the biases incumbent in the use of spatially aggregated data, the maintenance of high catch rates in localized areas may also mask a concomitant contraction in the spatial range of the stock. Uncertainty concerning the state of the stock is further increased if the spatial extent of the fishery is reduced over time. Fishing strategies may also be influenced by management decisions. For example, the introduction of competitive quotas often results in increased competition between vessels and an associated strategy of reducing risk by targeting areas of reliable catch rates. These issues raise serious questions about the adequacy of the catch and effort data alone to provide accurate indices of stock abundance.

Many of the problems found in the construction of indices of stock abundance from catch and effort data from a fishery can be illustrated by the problems that face scientists in assessing the status of the southern bluefin tuna (SBT) stock. This fishery has been characterized by a continuous decline in the size of the spawning stock since the 1960s, the introduction of restrictive catch quotas in the mid-1980s, reductions in the spatial and temporal distributions of fishing effort, and, more recently, a recovery in the abundance of juveniles during the 1990s. In particular, the spatial and temporal distribution of the fishing effort for SBT has declined appreciably since the fishery was first considered fully exploited in the early 1970s and this has resulted in a problem with the use of catch and effort data to calculate indices of SBT abundance. Basically there are no data to assess the present state of the stock in large areas in which the stock was historically distributed. In this paper an approach developed to deal with this problem is presented.

Abundance Indices in a Spatially Changing Fishery

Some of the problems encountered in the analysis of catch and effort data when the spatial distribution of a fishery changes over time can be clarified through a simple example. In a spatially expanding fishery there are no data in the early years for areas which are only fished in later years. If one only uses abundance estimates from the areas fished, then estimates of total size of the resource in the initial years will be underestimated. Prorating abundance estimates over the potential extent of the fishery may help to correct this problem, but if the areas fished in latter years are those with greater density, then the index will again underestimate total abundance. Alternatively, if a fishery contracts to the most productive areas, then the pro rata estimate of abundance for the total fishery will tend to overestimate the stock in the later years. Apart from the natural preference of fishers to target areas with higher returns, spatial contraction of effort could arise when quotas limit the total catch and fishing becomes limited to the most productive seasons and areas to reduce risks and maximize vessel quota share or profits.

As an illustrative example, consider the spatially expanding and contracting fishery displayed in Table 1. The fishing ground is spatially stratified into four distinct regions and an index of the number of fish in each region is also given. The regions fished each year are indicated in bold and the number of fish at the start of each year in each region is adjusted for the catch taken during the previous year. No movement between areas is assumed. The total abundance index of fish over all regions is the sum of the abundance indices in each of the separate regions. Based on the catch and effort data, assumed proportional to true abundance, together with assumptions about the stock numbers in the areas not fished each year, various indices of stock abundance can be calculated. Two possible indices are as follows:

1. Constant-Squares Index

It is assumed that the spatial extent of the stock is a constant from year to year and is equivalent to the union of the spatial extent of the fishery over all years. In any year the catch rates in the regions fished are assumed to be representative of all regions. As such, the catch rates in the regions not fished are assumed to be the weighted average (by area) of the catch rates in the regions fished. An index of abundance for the entire stock in any year is then:

$$I_{con\_sqs}(y) = \frac{A_{total}}{A_y} \times \sum_{i=1}^{N_y} A_i \times CPUE_{yi}$$

where $CPUE_{yi}$ is the observed CPUE in the ith region and yth year, A_i is the area of the ith region, N_y and A_y are the number and area of all regions fished in the yth year and A_{total} denotes the total area of all regions fished over all years.

2. Variable-Squares Index

It is assumed that the spatial extent of the stock varies from year to year and is the same as the spatial extent of the fishery in each year, i.e. $A_{total} = A_y$ in each year. This implies that the catch rates in the regions not fished are assumed to be zero. An index of abundance across for the stock in any year is then:

$$I_{var\_sqs}(y) = \sum_{i=1}^{N_y} A_i \times CPUE_{yi}$$

For the example given in Table 1 the values of each index, assuming $A_y = 1$ for all i and $CPUE_i$ is equivalent to the abundance index in each region, are given for each year.

As explained previously, when a fishery expands into more productive fishing regions, indices of abundance based on early catch and effort data will most likely underestimate the total size of the resource. Of course, had the fishery expanded into less productive regions, the Constant-Squares index would have overestimated the size of the resource in these early years. For the two years when all four regions are sampled, both of the above indices give unbiased estimates of the change in relative abundance. However, during the contraction phase of the fishery, since the fishery preferentially targets the areas of highest resource abundance, the Constant-Squares index is seen to overestimates the size of the resource. On the other hand, the Variable-Squares index continues to underestimates the size of the resource. Given the nature of the assumptions underlying the construction of the indices above, the true abundance in these later years can probably be assumed to be somewhere between the levels indicated by the two indices. This is because fishers are unlikely to have perfect knowledge of the areas of highest abundance, and the spatial extent of the stock is likely to be greater than the spatial extent of the fishery.

The two indices above are relatively simple, and it is possible to construct alternatives. For example, one can make use of the information in the years when the spatial extent of the fishery is greatest to model the expected catch rates in the regions not fished in other years. While there are many possible ways this can be done, one simple method is presented here. This method involves calculating the ratio of the catch rates in each region with the highest catch rate in that year. In the example above this gives 1.0, 0.675, 0.40, and 0.175 in the fourth year and 1.0, 0.667, 0.388, and 0.166 in the fifth year. If these distributions can be taken as representative of the distribution of catch rates in most years, then the mean dis-

Table 1. Schematic representation of a hypothetical spatially expanding and contracting fishery.

	Regions				Total	Constant	Variable	Ratio
Year	1	2	3	4	index	squares	squares	squares
1	**50**	100	150	200	500	200	50	
2	**45**	**100**	150	200	495	290	145	
3	**40**	**90**	**150**	200	480	373	280	
4	**35**	**80**	**135**	**200**	450	450	450	
5	**30**	**70**	**120**	**180**	400	400	400	
6	25	**60**	**105**	**160**	350	432	325	352
7	25	50	**90**	**140**	305	460	230	309
8	25	50	75	**120**	270	480	120	271

The numbers on the left give the average stock abundance in each region and year. The regions fished each year are indicated in bold. Calculated indices of stock abundance for each year are given in the columns on the right.

tribution (1.0, 0.671, 0.394, 0.171) can be used as indicative of the distribution of catch rates by area in the years when one or more of the regions is not fished. Finally, if it is assumed that the fishers always manage to fish in the regions with the highest catch rates, then the catch rates in the regions not fished can be calculated by multiplying the highest catch rate observed in a given year with the corresponding tail value of the average distribution. For example, the expected catch rate in the missing fourth region in the sixth year is $160 \times 0.171 = 27$, which is close to the true value. Again, the sum of the observed or expected catch rate in all regions for a given year gives an index of annual abundance. The value of this index, called the Ratio-Squares index, is shown in Table 1 for the later years.

Each of the three indices presented above are based on their own set of simple assumptions concerning the stock and fleet dynamics within the fishery. The Constant-Squares and Ratio-Squares models assume that the spatial extent of the stock remains constant between years. The allocation of fishing effort within each region then assumes that areas fished in any year are (i) chosen at random, or (ii) preferentially targeted at the areas with the highest catch rates. On the other hand, the Variable-Squares model assumes that the spatial extent of the stock can vary from year to year and that in any year fishing takes place in all areas where there are fish. This model therefore implies that there are no fish in the areas not fished. However, the Constant-Squares and Variable-Squares indices make predictions concerning the catch rates in unfished squares that make no use of the spatial distributions of catch rates in the squares that are fished. Alternatively, the predictions made by the Ratio-Squares index makes use of the information concerning the spatial distributions of catch rates in all squares in a number of years. It also attempts to more closely model the expected behavior of fishers who will attempt to maximize their economic returns by choosing to fish in areas of known high catch rates and forego areas with little or no fish. Indeed, after an initial expansion and learning phase, many fisheries may enter what may be regarded as a more mature phase characterized by the possible contraction of effort to more favorable areas. In particular, this may be the situation in those fisheries where total effort has declined, possibly due to the introduction of quotas or other management restrictions. While more complex models can be developed which more fully account for these factors, the Ratio-Squares index helps to encapsulate some of these ideas in a relatively simple manner.

The Longline Fishery for Southern Bluefin Tuna

Southern bluefin tuna (SBT) have a widespread but patchy distribution that is reflected in the spatial distribution of the fishing effort for this species. Like other tunas, SBT also tend to form transient aggregations in areas where oceanic thermal features favor local enrichment. The Japanese longline fishery for southern bluefin tuna has undergone remarkable changes since its inception as a major fishery in the early 1950s. Shingu and Hisada (1971) give a brief outline of the changes between 1957 and 1969 during which the area exploited by the longline fishery expanded by about nine times. While SBT spawn in the tropical waters south of Java the main fishing grounds since 1970 are located within the southern oceans between the mid-Atlantic and New Zealand (Fig. 1).

By 1970 the fishery was considered fully exploited, if not overexploited, with the abundance of the adult stock considered to be less than 10% of the initial level (Hayasi 1974). As a result, the Japanese initiated voluntary closures (begun in October 1971) of certain fishing grounds where young juveniles predominated (Suda 1974). Evaluation of these closures in 1974, however, showed they had not restricted effort with catch rates continuing to decrease and the proportion of small-sized fish in the catch continuing to increase (Warashina and Hisada 1974).

Since 1971 no major new fishing grounds for southern bluefin tuna have been discovered. However, there have been substantial contractions in the spatial distribution of effort in most of the major fishing areas since 1971 (Fig. 2) possibly due to the targeting of fishing effort to favorable areas. For example, the number of 1° squares fished in area 7 has more than halved since 1975. These changes are concurrent with larger scale changes in the amount of effort being expended within each statistical area as well as changes in the percentage of the total effort within each area. For example, the contraction in area 7 coincides with a general decrease in the effort within area 7, while the fluctuations in spatial

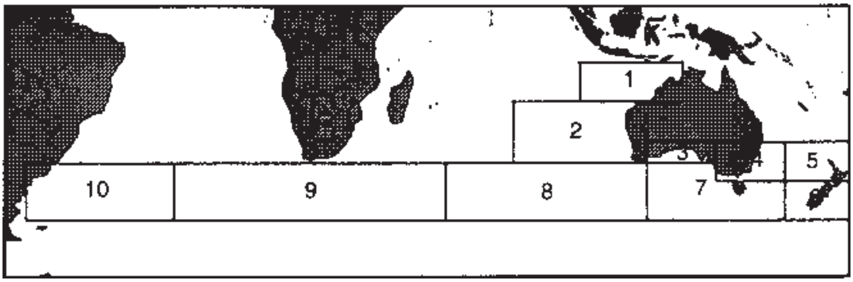


Figure 1. *Statistical areas used to provide coverage of the fishery for southern bluefin tuna.*

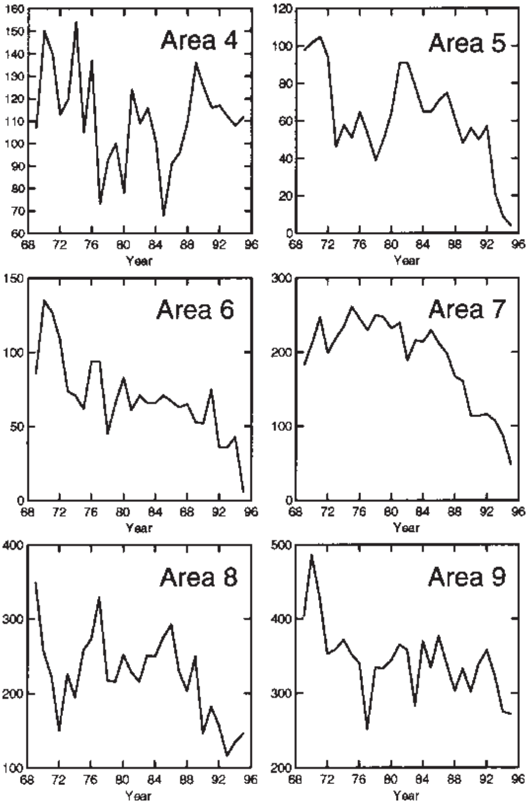


Figure 2. *Number of 1° squares fished by Japanese longliners within each of the SBT statistical areas 4 to 9 for the years 1969 to 1995.*

distribution in area 8 up to the mid-1980s coincide with a general increase in the total effort within that area during that time. Within area 9 a generally stable pattern in the spatial extent of the fishery is seen despite a general increase in effort up to the mid-1980s.

The significance of the contractions in the spatial distribution of effort indicated above will be increased if these shifts coincide with contraction of the fishing effort into regions with generally high catch rates. As indicated in the example above, this can mask the true extent of the decline in the stock.

Effort Dynamics with Statistical Area 7

The changes within statistical area 7 can be used to illustrate the nature of some of the changes across the entire fishery. The data used in the analyses described below come from two sources—Japanese longline data for SBT and joint-venture data for Japanese longline activities undertaken within the Australian Fishing Zone. In both instances the data describes the catch and effort for individual sets with the spatial resolution being 1° of latitude and longitude. The former are part of the extensive data base of tuna catch and effort held at the National Research Institute of Far Seas Fisheries in Shimizu, Japan, while the latter come from the data stored in the Australian Fishing Zone Information System database in Canberra, Australia.

In order to analyze the degree of spatial aggregation of fishing effort in each year, the 1° squares fished were ordered according to the amount of fishing effort (number of hooks) in them for each year. The percentage of the total annual effort that was expended in the top 50, 25, 10, and 5% of these squares was then calculated. The results are given in Table 2. For example, in 1971 fishing took place in 282 different 1° squares with 95% of that effort occurring in only 50% of these squares and 40% occurring in the top 10% of squares. Indeed, in all years the effort appears to be spatially aggregated. A comparison of the extent of this aggregation is provided for the three years 1971, 1980, and 1991 (Fig. 3a). There appears to be little change for the years 1971 and 1980, while the proportion of the effort expended in the top 10% of squares has increased significantly for the year 1991. Indeed, the increase in the level of aggregation appears to have been a feature of the fishery since 1986 and could indicate a shift in the fishing practices of the fleet since the introduction of quotas in the mid-1980s. These changes are clearly illustrated in the comparison of the spatial distribution of the Japanese longline effort for the years 1971, 1980, and 1991 (Fig. 4). Not only has the spatial extent of the fishery shrunk substantially, but the remaining effort is more concentrated. The greatest amount of effort in any single square during 1971 was less than 700,000 hooks, while during 1980 this had increased to 1,440,000 hooks and to just less than 2,000,000 hooks in 1991.

Table 2. Number of 1° squares fished each year in statistical area 7 and the percentage of the total effort within the top 50, 25, 10, and 5% of squares after ordering the squares by decreasing effort and catch rate.

Year	Squares fished	Ordered by effort				Ordered by catch rate			
		50%	25%	10%	5%	50%	25%	10%	5%
1971	282	94.8	73.4	39.9	23.1	46.0	22.1	5.9	3.1
1972	214	94.8	75.1	43.6	26.8	63.0	30.7	9.2	3.6
1973	231	94.7	79.4	50.2	31.0	53.1	24.3	7.0	2.8
1974	247	90.4	65.9	35.1	20.2	54.7	28.7	11.6	5.2
1975	271	94.5	81.0	53.1	33.6	66.0	18.3	4.8	1.0
1976	257	94.4	77.7	48.4	31.0	52.5	25.0	5.9	1.3
1977	244	92.9	73.1	48.0	32.2	63.7	24.6	7.9	2.6
1978	260	91.5	73.5	47.7	32.3	72.2	28.0	8.0	2.8
1979	248	92.3	71.3	41.9	25.4	72.5	28.2	4.2	0.6
1980	233	93.4	75.3	46.4	30.1	69.7	29.6	11.6	3.5
1981	243	94.1	78.1	52.1	37.4	67.7	31.6	6.8	4.9
1982	189	93.1	79.5	55.1	37.8	85.1	36.2	11.6	4.8
1983	216	90.1	68.6	43.8	30.5	69.4	21.0	8.4	1.8
1984	216	92.9	75.2	45.5	27.3	66.8	37.0	10.8	3.2
1985	231	90.1	67.1	39.6	25.6	70.0	35.8	21.3	10.0
1986	217	92.9	80.2	59.6	40.8	82.6	62.2	34.3	12.7
1987	199	93.4	76.9	52.1	35.9	66.8	42.7	11.4	6.4
1988	169	95.8	81.5	56.5	38.6	84.2	47.1	15.1	2.0
1989	165	95.6	86.9	73.0	60.7	86.4	76.7	56.7	1.3
1990	114	95.9	84.5	65.5	49.9	73.3	13.8	5.3	0.6
1991	120	93.4	82.0	63.9	47.3	79.2	63.5	10.9	2.3

In order to investigate the directed targeting of areas of higher stock densities, the relation between the distribution of effort and the stock density (using catch rates as a proxy) was analyzed by calculating the cumulative effort against area fished (after ranking the squares by catch rate). As before, the percentage of the total effort in the top 50, 25, 10, and 5% of these squares was calculated and the results are displayed in Table 2. For 1971 it is seen that 46% of the effort was expended in the top 50 of squares and only 6% of the effort in the top 10% of squares. In this situation there does not appear to be very effective targeting of the areas with higher nominal catch rates. However, a comparison of the degree of spatial targeting for the three years 1971, 1980, and 1991 (Fig. 3b) shows large changes between the years with the degree of targeting increasing

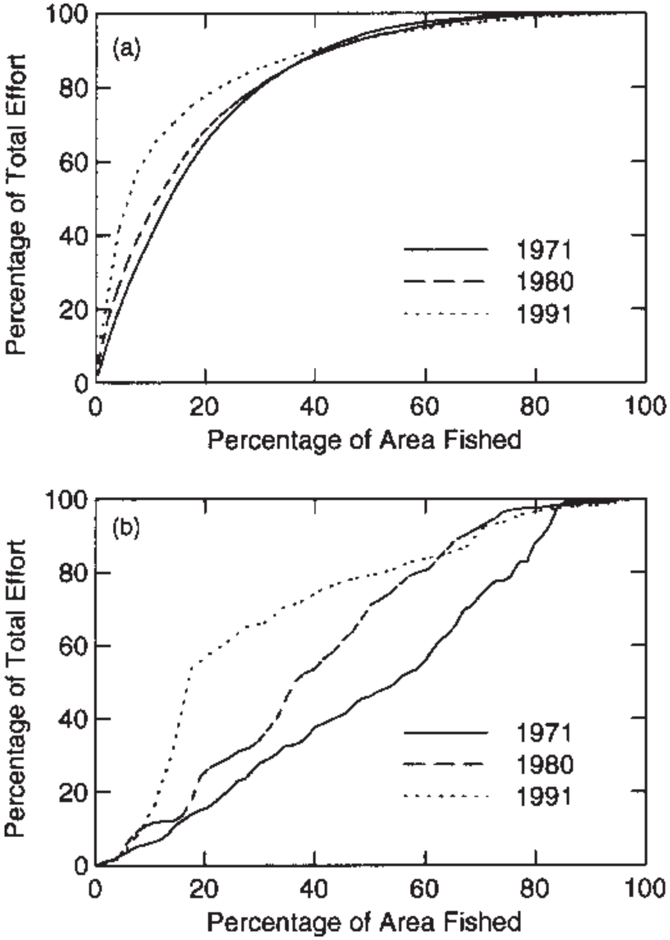


Figure 3. *Cumulative effort versus cumulative area fished (both expressed as a percentage of the respective annual totals) for the SBT fishery within statistical area 7 during the years 1971, 1980, and 1991 after ordering the 1° squares fished by (a) decreasing effort and (b) decreasing catch rates.*

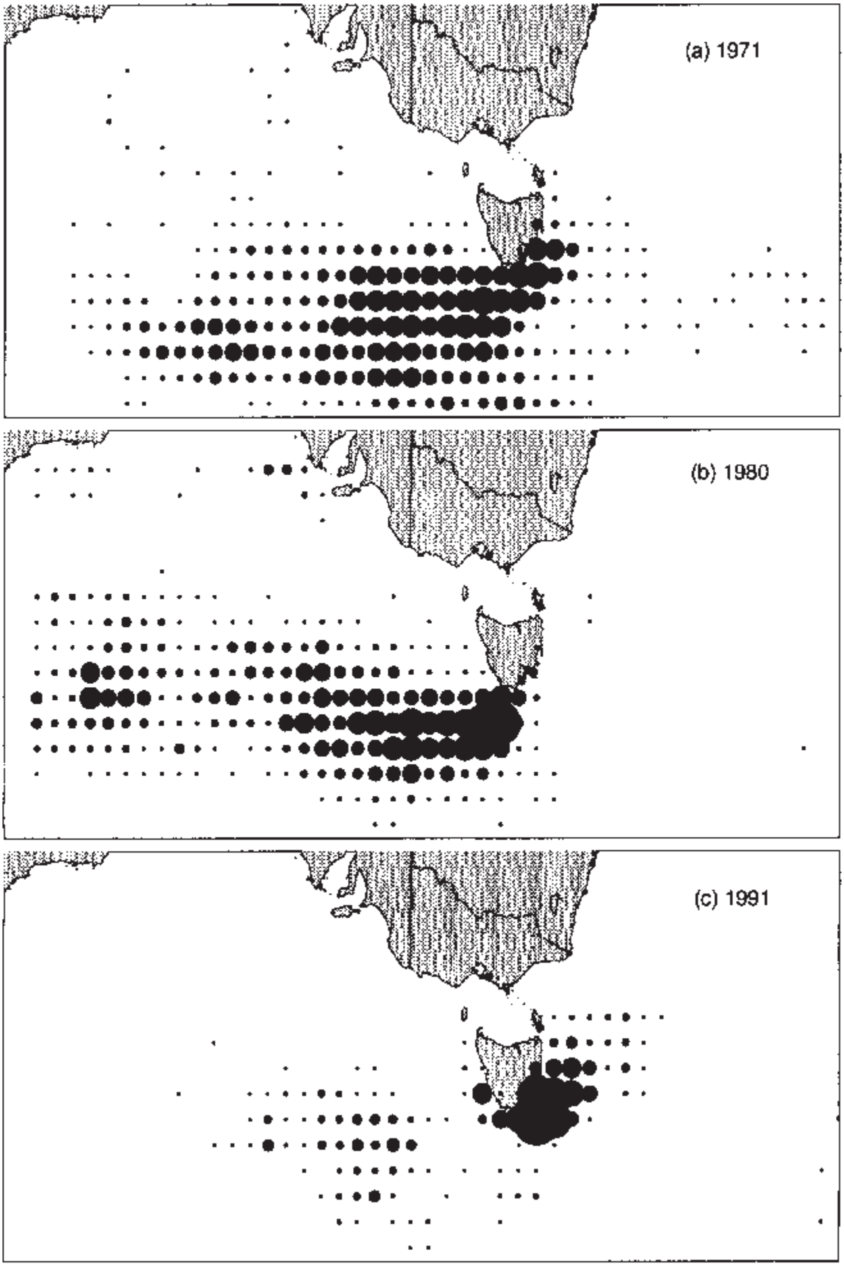


Figure 4. Spatial distribution of the Japanese longline fishing effort (number of hooks) within the SBT statistical area 7 for the years 1971, 1980, and 1991.

dramatically between 1980 and 1991. This increase in targeting appears to have increased steadily during the 1970s and more quickly in the late 1980s, reaching its greatest extent in 1989 when 57% of the effort was targeted in 10% of the squares.

It is evident that spatial targeting of fishing effort has always been a feature of the SBT longline fishery within area 7. However, in the early 1970s it appears that the effort was not preferentially targeted at areas of above average catch rates. This practice began to change in the mid-1970s and since 1986 there has been a significant increase in the spatial aggregation of the fishing effort with the areas with greatest effort increasingly being those with the highest nominal catch rates. This situation appears to be similar to that illustrated in the example given previously where in the later years fishing effort was reduced in the areas of low stock abundance. In both cases the increase in fishing effort in the areas of higher catch rates in comparison to areas of low catch rates has the potential to unduly weight the average catch rate upward, and hence causes an upward bias in the average index of abundance over the whole fishing ground. In order to overcome the potential for bias, one needs to ensure that the analysis of the data is undertaken on a spatial scale that matches the scale at which the effort is targeted in the fishery. For the SBT longline fishery this appears to be at the 1° scale. Although finer scale targeting is possible, the length of the longline (often up to 100 kilometers) and the manner in which the data is recorded results in this finer-scale information being lost.

Whether the changes since 1971 have been based on a deliberate strategy of reducing the risk of getting low catch rates is difficult, if not impossible, to discern from the data alone. However, such a strategy would seem reasonable from an economic perspective where catch rates need to be above a certain level in order to return a profit. While the actual fishing strategies of the fleet remain uncertain, it is generally seen to involve a searching component to find the areas of higher returns. However, the introduction of quotas in the mid-1980s and the concomitant reduction in fishing effort, competition between boats (and or components of the fleet) for quota share would necessarily involve a strategy of reducing the risk of returning low catches. The potential to reduce risk and concentrate fishing effort in productive areas would also be increased with the sharing of information within components of the fleet. Such an exchange of information appears to be a general practice among components of the Japanese fleets (Doulman 1987). Furthermore, increased concentration of effort in higher catch rate areas may also be due to greater targeting on smaller fish as returns from targeting the more widely dispersed larger fish diminish. The shift away from areas of low catch rates to areas of consistently high catch rates, as seen in the SBT fishery, would therefore be consistent with an economic practice of trying to maintain high profitability.

Stock Dynamics

Stock assessments indicate that the stock abundance of SBT has declined dramatically since the early 1960s, with the present spawning stock biomass estimated to be less than 10% of that in 1960 (Fig. 5). The spawning stock biomass in 1980 was 74% of the level in 1969, declining further to 29% of this level by 1991 (Polacheck et al. 1996).

Whether there has been a concomitant decrease in the spatial extent of the stock given the declines in stock abundance remains uncertain. Decreases in the spatial extent of the fishing effort may not be associated with decreases in the spatial extent of the stock, though variations in the spatial range of a decreasing fish population are consistent with observations from other animal populations and with the theory of density dependent habitat selection (MacCall 1990). Indeed, the collapse of the SBT within area 4 in the early 1980s may have been associated with a contraction in the distribution of the stock in this region (Caton et al. 1990).

Indices of SBT Stock Abundance

Indices of southern bluefin tuna abundance based on analysis of the Japanese longline catch and effort data have been calculated routinely as part of the annual assessment of the SBT stock. Due to the limited coverage for some statistical areas and quarters, the analyses are usually limited to statistical areas 4 to 9 and quarters 2 and 3 only. Total catch is converted to catch-at-age using size information aggregated by 5 × 10° and by quarter, as this was the usual scale at which the data were historically collected. The methods used for standardization of catch rates follow the work of Allen and Punsley (1984) and use a general linear model to estimate the parameter values in the fitted model. A full description of these analyzes is given elsewhere (Campbell et al. 1995, 1996). The discussion here will focus on the method adopted in light of the uncertainty in the stock and fleet dynamics described above.

Due to the variable spatial coverage of catch and effort data for the fishery from year to year, together with the uncertainty associated with the preferential allocation of effort, it has not been possible to calculate a single reliable unbiased index of stock abundance. To help overcome the potential for biased estimates due to the analysis of data with uncertain stock and effort dynamics, and to help bracket the uncertainty that arises with analysis of incomplete data, an approach has been developed which models the catch rates in the areas bypassed by the fishery under various assumptions concerning the spatial dynamics of the stock and the fleet. These assumptions are based on the concepts underlying the Constant-Squares, Variable-Squares, and Ratio-Squares indices presented in the illustrative example above. For brevity, the indices based on these three models are known as the B-avg, B-ratio, and B-min indices respectively

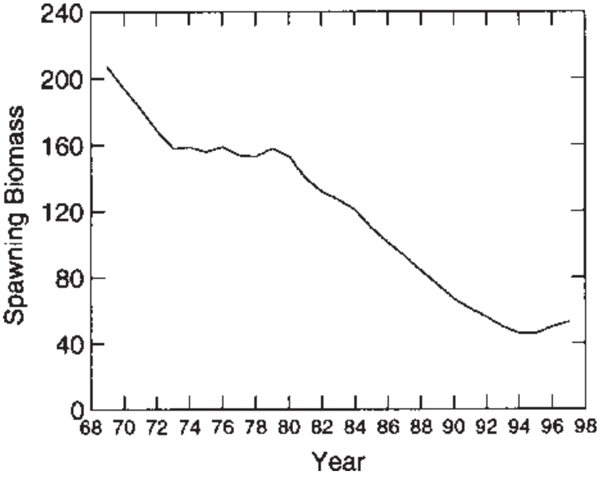


Figure 5. Annual estimates of southern bluefin tuna spawning biomass (thousand t). From Polacheck et al. 1996.

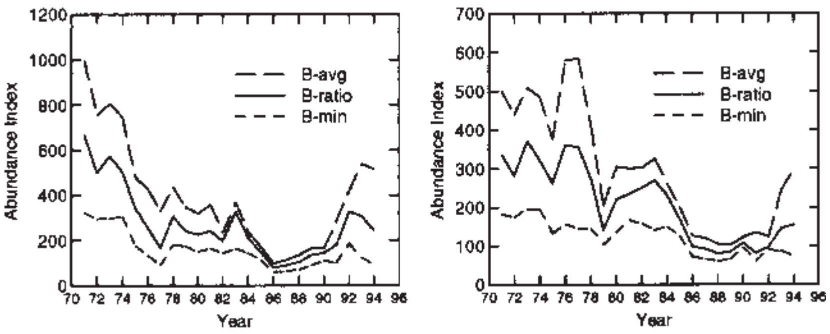
and the concepts underlying each of these models were implemented as follows:

B-avg: The spatial extent of the stock remains similar from year to year and is given by the union of all the 1° squares in which SBT have been caught over the time-frame of the available data. In any given year and quarter the 1° squares fished in each 5° block are assumed to be selected on a random basis. The expected catch rates in the 1° squares not fished in that block are assumed to be equal to the average of the catch rates in the fished 1° squares in that same block. (Note, where a 5° block is not fished in a particular year, the average catch rate of the 5° blocks which are fished within the same statistical area is used.)

B-ratio: The spatial extent of the stock is again assumed to be similar across years. However, for any given year and quarter the fishing effort in each 5° block is assumed to be preferentially targeted to those 1° squares with the highest catch rates. Catch rates in the 1° squares not fished are based on the tail of the average over all years of the distribution of the ratio of ordered catch rates in each 1° square fished in all 5° blocks. A more detailed description of the model is given in the Appendix.

B-min: The spatial extent of the stock is assumed to change from year to year and coincide with the spatial extent of the fishing effort. The catch rates in all 1° squares that are not fished in any year are therefore assumed to be zero.

(a) Ages 4-5 (b) Ages 6-7



(c) Ages 8-11 (d) Ages 12+

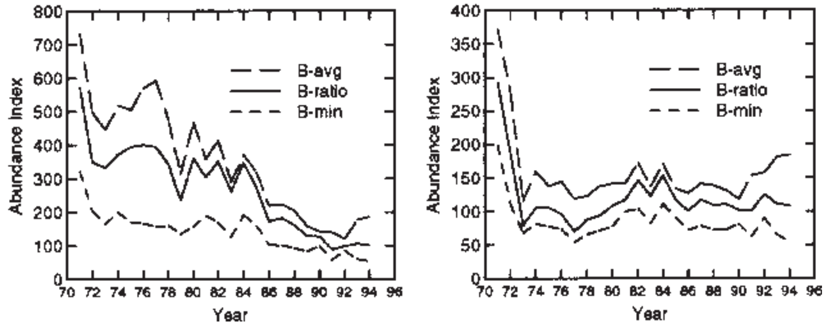


Figure 6. Annual indices of southern bluefin abundance for the different age classes.

The above models are used to augment the observed catch and effort data so that an extra catch and effort data record is created for each year-quarter-1° cell that is not fished. In this manner the data set to be analyzed by the general linear model becomes spatially balanced in the sense that for each year at least one data record exists for each quarter-1° cell. These augmented data sets are then analyzed by means of a general linear model and the results are used to calculate annual indices of stock abundance (Campbell et al. 1995, 1996). For the SBT assessments separate analyses are carried out on the individual age classes 4, 5, 6, 7, and 8 and on the combined age classes 4-5, 6-7, 8-11, and 12+. Annual indices of stock abundance since 1971 are then calculated for each of the three models (Fig. 6). All indices are incorporated as tuning indices in the subsequent VPA models (Polacheck et al. 1996) and by assigning statistical weights to each of the indices based on an assessment of the credibility of the model assumptions, probabilities can be assigned to the resulting projections (Klaer et al. 1996). An examination of these assumptions is given in Campbell and Tuck (1996).

An advantage of the approach just described is that instead of relying on just a single CPUE based tuning index, the assessment allows the results from several tuning indices to be combined and in so doing explicitly incorporates some of the uncertainties inherent in the catch and effort data. This uncertainty is clearly seen in the results presented. For example, differences in the resource levels indicated by the indices given in Fig. 6 are a reflection of the uncertainties in the information available. In particular, large differences between the indices in a given year are due to the uncertainties resulting from the lack of information in the areas which remain unfished, combined with the high uncertainty regarding possible catch rates in these areas due to a large range in the catch rates observed in those areas which are fished. Indeed, high catch rates in some areas in the early 1970s helps to explain the large differences noted between the different indices during this period. However, as the stock was fished down the range of catch rates observed in the fishery became smaller resulting in the differences between the indices also becoming smaller. During the 1990s, there has also been a substantial contraction in the spatial extent of the fishing effort and this has contributed to the increased differences between the individual indices during this later period.

While inferences about the status of the SBT stock do not concern us here, several features can be noted. First, for most age classes there was a substantial decline in abundance between 1970 and the mid-1980s followed by a recovery in the late 1980s or early 1990s. However, the size of this recovery varies significantly between the different indices. In some instances substantial increases in abundance are seen while in other instances the recovery is either small or a further decline continues. Again, these differences are a reflection of the uncertainties in our knowledge concerning the stock. Second, if one assumes the B-ratio indices best reflect the underlying dynamics of the fishery, then one sees a sequential recovery in individual age classes since the mid-1980s (Fig. 7). Furthermore, the large cohort which entered the fishery as 4-year-olds in 1992 is seen to persist as 5-year-olds in 1993 and as 6-year-olds in 1994. However, assuming comparable catchability, the diminishing size of the index for this cohort may be indicative of a sequential fishing-down over these years. Finally, the younger and older age classes are seen to respond differently in 1983, with the younger age classes having a relatively high index and the older age classes having a relatively low index. These deviations from the trend are possibly due to unusual environmental effects during 1983 when a strong El Niño event was recorded. The responses noted in the indices for this year may be indicative of changes in age-class availability in response to changing environmental conditions not adequately accounted for in the standardizing model used.

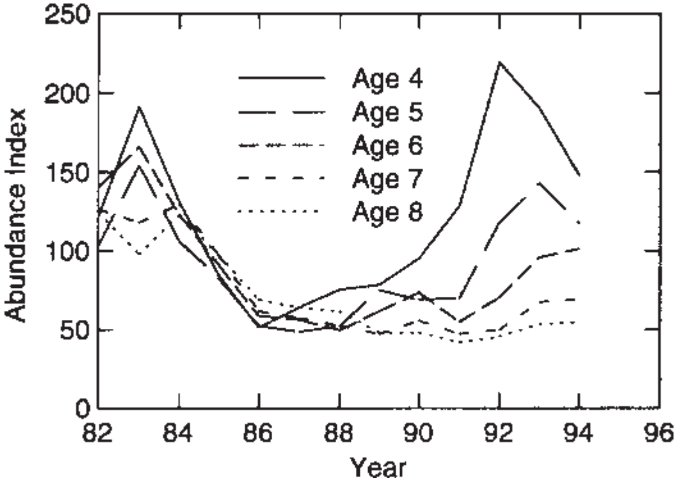


Figure 7. Annual B-ratio indices of southern bluefin abundance for age classes 4-8 since 1982.

While the results presented here are limited to the three models described above, which apply the same set of assumptions across all years for each index, it would be possible to have different sets of assumptions in different years. For example, one could perhaps use the results concerning the levels of targeting of high catch rate squares in each year (cf. Table 2) to weight the effort assumption used in the B-ratio index. Alternatively, one could combine the individual indices as calculated, using different weights for each in different years. For example, Hearn and Polacheck (1996) used the concepts behind the density dependent habitat models of MacCall (1990) to construct an index that gives different annual weights to two indices based on the Constant-Squares and Variable-Squares concepts described previously.

Discussion

Interpretation of catch rates and construction of indices of stock abundance should be based on an understanding of both the distribution of the stock and the distribution of the fishing effort, and ultimately on the relationship between these two. It is usually the case, however, that the fishery dependent catch and effort data alone will not allow one to adequately resolve competing hypotheses concerning the underlying stock and effort dynamics. In such cases fishery-independent data such as systematic research surveys and understanding of the decision rules by which fishing effort is spatially allocated are also needed if these hypotheses are to be resolved.

A number of issues related to the difficulties in interpretation of catch and effort data appear to have contributed to the recent collapse of the North Atlantic cod fishery. A summary of some of the events leading up to this crisis describes how the "fleet was fishing a smaller and smaller area of ocean" and "the fishermen were catching more fish per hour than the scientists because they were going to warmer patches where they knew cod were congregating" and "you get a false impression that there are lots of fish, while the surrounding ocean is empty" (New Scientist 1995).

The extent to which the SBT fishery has been able to concentrate fishing effort in areas with higher than average fish density has been investigated and instances of considerable spatial targeting have been found. This practice seems to have become more prevalent since the introduction of quotas in the mid-1980s. In order to overcome the potential biases that can result in using the catch and effort data from a fishery with a high degree of spatial targeting, analysis of the commercial catch and effort data to obtain indices of stock abundance has been carried out at the finest spatial scale possible. For the SBT longline fishery, catch and effort data are recorded to within one degree of latitude and longitude.

Related to the problems in analyzing catch and effort data for a spatially targeted fishery, are the additional biases and uncertainty which enter due to missing observations, e.g., historical areas of the fishery which are no longer fished. Due to spatial contractions in the fishing effort over the years, the extent to which the fishing grounds are known to overlap the entire stock area has become more uncertain. The changing nature of the SBT fishery from year to year makes it difficult to say whether the fish population in an area with no fishing effort should be considered unsampled or not present. Given this uncertainty, it has not been possible to agree on a single reliably unbiased index of stock abundance based solely on the commercial catch and effort data. Instead a number of indices based on various assumptions as to the underlying stock and effort dynamics of the fleet have been constructed. Support for or rejection of the assumptions underlying the calculation of the various indices can then be based on a spatial analysis of the data for the fishery itself and/or from behavior observed in other fisheries or from the ecological field in general.

Analyses undertaken for the 1996 assessments (Campbell and Tuck 1996) found disproportional changes in the distribution of the stock over time resulting in a greater portion of the stock being found in a smaller area. While this result may be due to several factors, such variations in the spatial range of a fish population are consistent with observations from other animal populations and with the theory of density-dependent habitat selection (MacCall 1990). The advantages of such behavior may include improved foraging, improved migration or spawning success, and protection from predation (Clark 1985). Whatever the natural advantages, this behavior is also important to the fishing fleets. However, whether or not the fishing fleets have been able to successfully target the main habitat areas containing the major proportion of the stock is less certain (and underlies the main differences in the indices described above), though spatial contraction of a fishery to high catch rate areas is consistent with the economic practice associated with competitive quotas.

The uncertainties in our ability to presently assess the status of the stock of southern bluefin tuna have led to calls for a number of experimental fishing programs to be undertaken (see Anon. 1996 for a summary). Given the contraction in the spatial extent of the commercial fishery and the resulting uncertainties incurred, attempts to gain a better understanding of the spatial distribution of the stock have been given a high priority, though the mechanism by which this can be carried out remains unresolved. At its simplest, one needs to sample the areas of the fishery which were fished historically but which are no longer fished. However, given the high costs of fishing, the high likelihood of low catch returns, and the large area to be sampled, this does not seem an economically viable option for the commercial fleet alone. The details of an appropriate sampling design and fishing strategy remain the subject of ongoing research.

Finally, while this paper has focused on the problems associated with the interpretation of commercial catch and effort data in a fishery with uncertain stock and effort dynamics, many other factors influence our ability to interpret commercial catch rates as indices of stock abundance. Many of these factors are well known (see Paloheimo and Dickie 1964, Gulland 1974, Hilborn and Walters 1992) and include improvements in the operational and technological aspects of the fishery, together with the influence of economic and management related decisions, all of which may change catchabilities over time. Attempts to document the changes in the SBT fishery and include them in future analyses are under way.

Acknowledgments

The author extends his thanks to the National Research Institute of the Far Seas Fishery in Shimizu, Japan, for permission to access the fine-scale data for the SBT fishery and for the use of its facilities. He also wishes to thank the staff for their assistance and friendliness. In particular, thanks go to Drs. Tom Nishida, Sachiko Tsuji, and Yoshio Ishizuka for their help in preparation and interpretation of the data files.

References

Allen, R.L., and R.G. Punsley. 1984. Catch rates as indices of abundance of yellowfin tuna, *Thunnus albacares*, in the eastern Pacific Ocean. Inter-Am. Trop. Tuna Comm. Bull. 18 (4):301-379.

Campbell, R.A., and G. Tuck. 1996. Spatial and temporal analysis of the SBT fine-scale catch and effort data. Working Paper No. 18 presented at the 2nd CCSBT (Commission for the Conservation of Southern Bluefin Tuna) Scientific Meeting held August 26-September 6, Hobart, Australia.

Campbell, R.A., G. Tuck, and T. Nishida. 1995. Indices of abundance for southern bluefin tuna from analysis of fine-scale catch and effort data. Working Paper No. 14 presented at the 1st CCSBT (Commission for the Conservation of Southern Bluefin Tuna) Scientific Meeting held July 10-19, Shimizu, Japan.

Campbell, R.A., G. Tuck, S. Tsuji, and T. Nishida. 1996. Indices of abundance for southern bluefin tuna from analysis of fine-scale catch and effort data. Working Paper No. 16 presented at the 2nd CCSBT (Commission for the Conservation of Southern Bluefin Tuna) Scientific Meeting held August 26-September 6, Hobart, Australia.

Caton, A., K. McLoughlin, and M.J. Williams. 1990. Southern bluefin tuna: Scientific background to the debate. Bureau of Rural Resources, Bulletin No. 3, Department of Primary Industries and Energy, Australian Government Publishing Service, Canberra. 41 pp.

CCSBT. 1996. Report of the workshop “Developing a framework for evaluating the impact of experimental fishing on the southern bluefin tuna stock” held May 27-June 1, Shimizu, Japan. Convened by the Commission for the Conservation of Southern Bluefin Tuna, Canberra, Australia.

Clark, C.W. 1985. Bioeconomic modelling and fisheries management. John Wiley and Sons, New York. 291 pp.

Doulman, D.J. (ed.). 1987. Tuna issues and perspectives in the Pacific Island region. Pacific Islands Development Programme, East-West Centre, Honolulu.

Gulland, J.A. 1974. Catch per unit effort as a measure of abundance. Inter. Comm. Cons. Atlan. Tunas, Coll. Vol. Sci. Pap. 3:1-5.

Hayasi, S. 1974. Stock Assessment of southern bluefin tuna based on information up to September 1973: Its conclusions and problems. Bull. Far Seas Fish. Res. Lab. 11:51-65.

Hearn, W.S., and T. Polacheck. 1996. Estimation of indices of southern bluefin tuna abundance by applying general linear models to CPUE. Working Paper No. 19 presented at the 2nd CCSBT (Commission for the Conservation of Southern Bluefin Tuna) Scientific Meeting held August 26-September 6, Hobart, Australia.

Hilborn, R., and C.J. Walters. 1992. Quantitative fisheries stock assessment: Choice, dynamics and uncertainty. Chapman and Hall, New York. 570 pp.

Klaer, N., T. Polacheck, K. Sainsbury, and A. Preece. 1996. Southern bluefin tuna stock and recruitment projections. Working Paper No. 30 presented at the 2nd CCSBT (Commission for the Conservation of Southern Bluefin Tuna) Scientific Meeting held August 26-September 6, Hobart, Australia.

MacCall, A.D. 1990. Dynamic geography of marine fish populations. University of Washington Press, Seattle. 153 pp.

New Scientist. 1995. The cod that disappeared. New Scientist, 16 September, pp. 24-29.

Paloheimo, J.E., and L.M. Dickie. 1964. Abundance and fishing success. Rapp. P.-V. Reun. Cons. Int. Explor. Mer 155:152-63.

Polacheck, T., A. Preece, A. Betlehem, and K. Sainsbury. 1996. Assessment of the status of the southern bluefin tuna stock using virtual population analysis. Working Paper No. 26 presented at the 2nd CCSBT (Commission for the Conservation of Southern Bluefin Tuna) Scientific Meeting held August 26-September 6, Hobart, Australia.

Shingu, C., and K. Hisada. 1971. Fluctuations on amount and age composition of catch of southern bluefin tuna in longline fishery, 1957-1969. Bull. Far Seas Fish. Res. Lab. 2:195-198.

Suda, A. 1974. Recent status of resources of tuna exploited by longline fishery in the Indian Ocean. Bull. Far Seas Fish. Res. Lab. 10:27-59.

Warashina, Y., and K. Hisada. 1974. Preliminary evaluation of effect of the voluntary regulation on stock of southern bluefin tuna and the longline fishery. Bull. Far Seas Fish. Res. Lab. 10:193-215.

Appendix: Protocol for estimating catch rates for the B-ratio index.

Consider a situation where fishing effort within a 5 × 5° block spatially contracts over time (Fig. A1). In the first year, 18 of the 1° squares are fished while in the second and third years this number decreases to 13 and 11 respectively. The data set across the three years is unbalanced in the sense that we have 18 squares sampled in the first year, but fewer in later years. Estimates of the catch rates in the squares not fished in these later years are based on the following protocol.

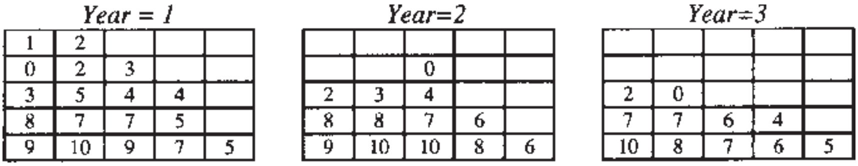
First, for each year the average catch rate in each of the 1° squares are sorted into descending order. These values are shown in the column headed CPUE in Table A1. Second, for each year *i* the ratio of the *j*th catch rate with the maximum catch rate is calculated. These values are in the column R_{ij}. Third, for each level the average of these ratios is calculated across all years for which there is data:

$$\overline{R}_j = \frac{1}{N_j} \sum_{i=1}^{N_j} R_{ij}$$

where N_j is the number of years for which the ratio value exists at the *j*th level. The values of $\overline{R}_j$ are listed in the last column in Table A1. Finally, the values for the missing catch rates in year *i* for the *j*th square in each block are calculated as follows:

$$CPUE_{ij} = \frac{1}{m} \sum_{k=1}^{m} CPUE_{ik} \times \frac{\overline{R}_j}{\overline{R}_k} = \frac{\overline{R}_j}{m} \sum_{k=1}^{m} \frac{CPUE_{ik}}{\overline{R}_k}$$

where *m* is the number of squares having non-zero catch rates in that block in the *i*th year. The calculated values are displayed in Table A1 in bold. Note that averaging over all the observed catch rates in the manner shown helps reduce the influence of outliers in the observed data.


Year = 1

1	2			
0	2	3		
3	5	4	4	
8	7	7	5	
9	10	9	7	5

Year=2

		0		
2	3	4		
8	8	7	6	
9	10	10	8	6

Year=3

2	0			
7	7	6	4	
10	8	7	6	5

Figure A1. Spatial distribution of fishing effort within a 5 × 5° block over three successive years. The numbers represent the observed catch rates within each 1 × 1° square. Blank squares indicate the absence of fishing effort.

Table A1. Tabulation of observed and modeled 1° catch rates used in the example for the calculation of the *B*-ratio indices.

	Year = 1		Year = 2		Year = 3		
Level j	$CPUE_{1j}$	R_{1j}	$CPUE_{2j}$	R_{2j}	$CPUE_{3j}$	R_{3j}	$\overline{R}_j$
1	10	1.0	10	1.0	10	1.0	1.00
2	9	0.9	10	1.0	8	0.8	0.90
3	9	0.9	9	0.9	7	0.7	0.83
4	8	0.8	8	0.9	7	0.7	0.77
5	7	0.7	8	0.8	7	0.7	0.73
6	7	0.7	8	0.8	6	0.6	0.70
7	7	0.7	7	0.7	6	0.6	0.67
8	5	0.5	6	0.6	5	0.5	0.53
9	5	0.5	6	0.6	4	0.4	0.50
10	5	0.5	4	0.4	2	0.2	0.37
11	4	0.4	3	0.3	0	0.0	0.23
12	4	0.4	2	0.2	**2.36**		0.30
13	3	0.3	0	0.0	**1.18**		0.15
14	3	0.3	**2.98**		**2.36**		0.30
15	2	0.2	**1.99**		**1.58**		0.20
16	2	0.2	**1.99**		**1.58**		0.20
17	1	0.1	**1.00**		**0.79**		0.10
18	0	0.0	**0.00**		**0.00**		0.00

The modeled catch rates are in bold.

Stock Production Models of Blue Marlin and White Marlin in the Atlantic Ocean: A Case History

Christopher D. Jones
University of Miami, Rosenstiel School of Marine and Atmospheric Science, Miami, Florida

Eric D. Prince, Gerald P. Scott, and Mark I. Farber
National Marine Fisheries Service, Southeast Fisheries Science Center, Miami, Florida

Abstract

Historically, stock assessments of Atlantic blue marlin *(Makaira nigricans)* and white marlin *(Tetrapturus albidus),* conducted under the auspices of the International Commission for the Conservation of Atlantic Tunas (ICCAT), have been restricted to production modeling approaches. Production models are used due to the unique fisheries and biological aspects of the species which result in a paucity of detailed information on the size or age structure of the catch. These analyses have evolved from single index equilibrium, to multiple index non-equilibrium models, as the ICCAT Enhanced Research Program for Billfish improved the Atlantic-wide data and new non-equilibrium multifishery production models became available. Analyses of these fisheries data have been conducted over the last two decades through a series of intersessional billfish workshops held by ICCAT, under various stock structure hypotheses, to provide estimates of historical relative biomass, fishing mortalities, and maximum sustainable yield. Among the difficulties in modeling stock biomass, there have been conflicting indices of abundance for several fisheries, difficulties in modeling the dynamics of precipitous drops in CPUE through non–age structured approaches, and typically flat solution surfaces which cause difficulty in searches for optimal, unconstrained model solutions. However, in each case various assumptions were made, or certain parameters

---

Current address for C.D. Jones is National Marine Fisheries Service, Southwest Fisheries Science Center, 8604 La Jolla Shores Dr., La Jolla, CA 92038

fixed, generally based on working group consensus, and solutions were achieved. A case history of multi-fishery Atlantic billfish assessments is presented, along with approaches that enabled specific problems to be addressed in the model fitting.

Introduction

International management of Atlantic blue marlin *(Makaira nigricans)* and white marlin *(Tetrapturus albidus)* falls under the auspices of the International Commission for the Conservation of Atlantic Tunas (ICCAT), which currently has 25 member countries and is headquartered in Madrid, Spain. The commission was originally formed in 1966 to maintain the populations of tunas and tunalike fishes (including billfishes, Istiophoridae, and swordfish, Xiiphidae) at levels that permit the maximum sustainable catch for food and other purposes. Assessing the population status of Atlantic blue marlin and white marlin, as well as other istiophorids, has historically been difficult because of the unique aspects of the fisheries (Furman 1989) and biology (Boggs 1989) of these species which hinder acquisition of information to assess the status of the stocks.

In general, billfishes are large highly mobile species, long-lived, and sparsely distributed predators with an extensive geographical range. The historical delineation of Atlantic billfish stock structure has had considerable uncertainties. Although there have been numerous instances of trans-Atlantic movements of both blue and white marlin (Jones and Prince 1996), and at least one documented trans-oceanic movement of a blue marlin to the Indian Ocean (NMFS 1994), tag recovery data provides limited insight into ocean-wide structure of the populations and questions on stock structure have not been fully resolved. As a result, in most assessments three possible scenarios have been contemplated: a North Atlantic stock, a South Atlantic stock, and a total Atlantic stock.

Recognizing that gains in scientific advice about the status of these resources could be realized if more detailed and comprehensive information was available, ICCAT initiated the Enhanced Research Program for Billfish (ERPB) in 1986. The ERPB elevated billfish research priorities to the international level and made funds (primarily from private recreational interests in the United States) available to reduce deficiencies in the ICCAT billfish database. The major objectives of the ERPB included: (1) Improve catch, effort, and landings statistics; (2) develop an Atlantic-wide tagging program for billfish; and (3) promote the advancement of age and growth studies. This data collection and research program continues in 1998.

Atlantic-wide estimates of billfish nominal landings and catch-per-unit-effort (CPUE) have undergone extensive revision since statistics were first compiled, as a result of research activities of ERPB and through a series of intersessional billfish workshops held by ICCAT (ICCAT 1981, 1994, 1996). The changes in data collection and estimation of abundance indices have been addressed in four billfish workshops, three of which were held under the auspices of ICCAT. The results of these workshops, along with the evolution of billfish stock assessment techniques from single to multi-fishery production models, is the primary focus of this paper.

Approaches to Population Assessment

Data Collection

Collection of data for stock assessments of blue and white marlin are difficult, due to several factors related to the fishery. In any given year, 70-90% of the Atlantic-wide billfish landings (representing about 1% of the total Atlantic-wide landings of tuna and tuna-like species) reported to ICCAT come from longline fisheries that target tuna and swordfish, in which billfish are incidentally caught. Billfishes caught in this manner are normally dressed at sea, with heads, spines, fins, tails, and viscera removed to permit efficient onboard storage. The carcasses are then frozen for long periods before they are off-loaded at transshipment ports. This method of processing may lead to species mis-identification, non-reporting of landings, and lumping two or more billfish species into "unclassified" billfish category. In addition, this can lead to increased difficulties in acquisition of size frequency data, determination of sex ratios, and collection of other landings statistics. These problems are not unique to Atlantic fisheries (Alverson et al. 1994), since non-target species are generally not accounted for in as much detail in terms of landing statistics. Estimating dead discards and incidental mortality is difficult for both commercial and recreational fisheries. In the recreational fishery, the majority of catches are released. Not only is the estimation of the total recreational landings (actual landings plus dead discards from fish released) extremely difficult due to few adequate surveys, but the proportion of released fish that die from the stress of capture and subsequent release is poorly understood in both commercial and recreational catches.

The collection of nominal landings data is further complicated because as many as four dozen nations (many not members of ICCAT) catch Atlantic billfishes but do not routinely report billfish landings directly to ICCAT, although catch reports are obtained through other international organizations (e.g., FAO). Prior to 1986, billfish had a lower research priority than tunas or swordfish for many ICCAT countries due to the incidental nature of most of the Atlantic billfish landings. Recognition of a large but generally unquantified socioeconomic value (Meyer 1989) for the recreational component of Atlantic billfish fisheries has elevated the priority somewhat.

Atlantic-wide blue marlin and white marlin nominal landings reported to ICCAT have shown great fluctuations over time (Fig. 1). The longline fishery in the Atlantic was established in the late 1950s, and very high landings occurred for both species into the 1960s. The landings peaked in

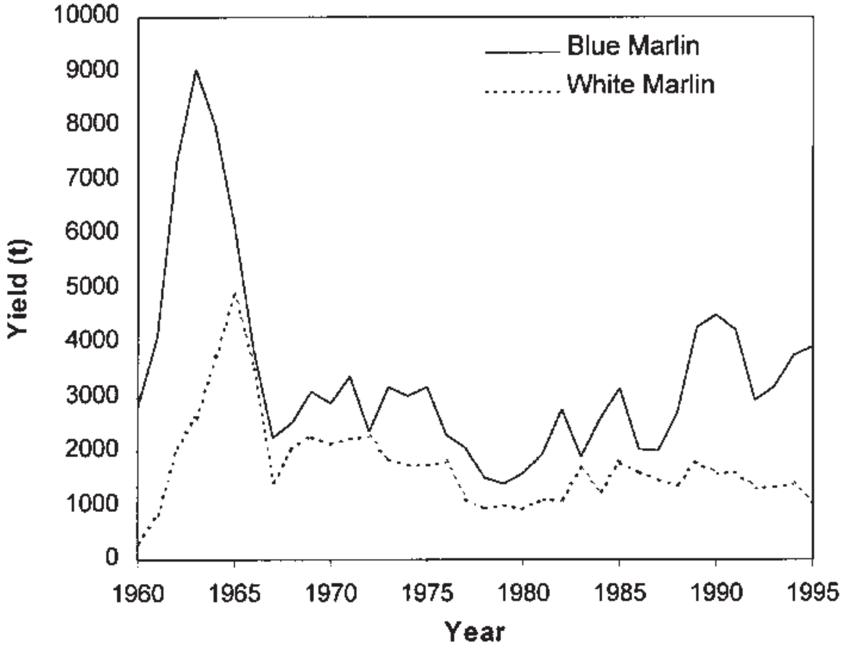


Figure 1. Historical total catch (t) of total Atlantic blue marlin and white marlin, 1960-1995.

the mid-1960s and have since remained well below those levels for both species. Trends in landings have generally followed longline fishing effort, particularly by the Japanese fleet. However, estimated landings have increased for blue marlin in the last decade, while remaining at fairly constant levels since the early 1970s for white marlin.

Population Production Models

The difficulties in compilation of landing statistics coupled with the lack of detailed information on biology of these species makes information very limited or unavailable. As a result, these restrictions hinder attempts to apply a range of age or size based stock assessment analyses. Because lumped biomass stock production models require only a time series of catch and index of abundance (i.e., CPUE) these models have been the quantitative method of choice for billfish stock assessments.

The simplest production models assume a logistic increase in the rate of change of a stock due to production, and often include an equilibrium assumption (Graham 1935, Schaefer 1954). Several variations on the simple stock production approach have been developed. One of the most flexible approaches was the generalized production model (Pella and Tomlinson 1969). This model adds an additional parameter to the logistic equation that allows various shapes in the production function, and does not require the equilibrium assumption. Though this model has increased flexibility, most implementations only allow fitting of a single index of abundance.

The introduction of ASPIC (a stock-production model incorporating covariates, Prager 1992) also extended the simple logistic population production model approaches by allowing for simultaneous analysis of multiple data series, and not requiring equilibrium assumptions. These characteristics allowed for a more extensive use of the ICCAT Atlantic-wide multi-fishery database. The detailed theory and mechanics of ASPIC are fully described in Prager (1992, 1995). The parameters estimated in the ASPIC model formulation are: *K,* the biological carrying capacity; i.e., maximum equilibrium stock size (t); B_1R, the ratio of biomass in the first year to *K*; *r,* the biological yearly intrinsic rate of increase of the stock; and $q(i)$, the catchability coefficients for each of the individual *i* data series. For the purposes of ICCAT, the derived quantities of most interest were: MSY, the maximum sustainable yield (t) per year (= $Kr/4$); B_{MSY}, the stock biomass (t) at MSY (= $K/2$); and F_{MSY}, the fishing mortality rate at MSY (= $r/2$), as well as the time trajectories of the relative statistics B/B_{MSY} and F/F_{MSY}, which provide information on the status of the resource over time. Because the quantities estimated most precisely by production models are MSY, effort at MSY, and biomass and fishing mortality levels relative to MSY (Prager 1992), estimates of biomass and fishing mortality trajectories presented are in terms of relative biomass (B/B_{MSY}) and relative fishing mortality (F/F_{MSY}).

Overview of Past Assessment Results

Early Assessments

Little modeling of Atlantic billfish population dynamics was attempted prior to 1977, due to limited catch and effort data and a poor understanding of the fishery and biology of the species. The billfish stock assessment workshop held in Hawaii in 1977 (NMFS 1978) was the first important forum that addressed data needs and direction of research efforts to better understand the world's populations of billfish. Conser and Beardsley (1979) built upon the work of Kikawa and Honma (1978), and assessed the status of stocks of blue marlin and white marlin in the Atlantic Ocean. Following the recommendations of the 1977 workshop, available data were analyzed under the two stock structure assumptions: separate North Atlantic and South Atlantic stocks; and a single total Atlantic-wide stock. In that assessment, Japanese longline data and estimates of Atlantic-wide catches were analyzed using the generalized stock production model (Pella and Thomlinson 1969). Although model fits to the data were generally

Table 1. Historical estimates of maximum sustainable yield (t) for blue marlin and white marlin under the north, south, and total Atlantic Ocean stock assumption.

Stock assumption	Source	Blue marlin	White marlin
North Atlantic	Kikawa & Honma 1978	2,300	1,700
	Conser & Beardsley 1979	2,884-3,136	1,877-2,042
	Farber & Conser 1983	2,232-2,623	2,092-3,776
	Cramer & Prager 1994	1,718-1,864	
	Farber & Jones 1994		388-921
	ICCAT 1996	1,741-2,133	85-771
South Atlantic	Conser & Beardsley 1979	2,516-2,871	
	Farber & Conser 1983	2,074-2,353	2,579-2,672
	Cramer & Prager 1994	704-1,278	
	Farber & Jones 1994		739-2,282
	Jones & Farber 1996	1,193-1,224	1,000
Total Atlantic	Conser & Beardsley 1979	4,768-5,333	
	Farber & Conser 1983	3,807-5,040	6,230-6,286
	Cramer & Prager 1994	3,517-3,623	
	Farber & Jones 1994		1,502-1,741
	ICCAT 1996	4,096-4,787	2,102-2,228

poor, some of the earliest estimates of maximum sustainable yield were developed as a result of the 1977 workshop (Table 1). Although the report by the Standing Committee on Research and Statistics (SCRS) (ICCAT 1980) concluded that it was not clear as to “whether the apparent over-fishing of the North Atlantic stock of marlin is growth overfishing or recruitment overfishing,” the analyses indicated that the stock was probably below the level which could produce MSY.

Soon after this workshop, Beardsley and Conser (1981) examined catch and effort data from the U.S. recreational fishery for billfishes to evaluate their usefulness in determining trends in abundance. Using a power model (Robson 1966), they were able to develop an index of relative abundance over the period 1971-1978. However, the usefulness of this series for Atlantic wide population assessment was limited, given that there were no analytical tools at the time for modeling multiple fishery data series simultaneously. However, the importance of accounting for separate CPUE series for the overall stock was recognized as a means of improving assessment advice.

During the First ICCAT Inter-Sessional Billfish Workshop (ICCAT 1981), blue marlin and white marlin catch statistics by country, compiled primarily from the ICCAT Statistical Bulletins, were reviewed, refined, and reestimated for the period 1957-1979. The only extensive standardized time series of CPUE was collected from the Japanese longline fishery. Farber and Conser (1981) followed the methodology of Conser and Beardsley (1979) and applied the generalized stock production model with equilibrium assumptions, fitting an index of abundance and a weighted average of effective fishing intensity using the PRODFIT program (Fox 1975). Farber and Conser (1981) found that under the North Atlantic stock assumption, "the models fit the data fairly well." Although there were still problems and uncertainties with the assessment, they concluded that the assumed north, south, and total Atlantic stock of blue marlin appeared "to have been overexploited in the early 1970s" and that if the most recent (i.e., 1977-1978) indices were reliable, "the North Atlantic stock of white marlin may be seriously overfished." Relatively uncertain results were found for the South Atlantic and for the total Atlantic stock assumptions. However, despite the uncertainties, they concluded that the total Atlantic stock was "at least fully exploited since 1970 and probably overexploited by 1977 and 1978."

Farber (1982) followed the methodology of Conser and Beardsley (1979) and Farber and Conser (1981) and attempted to assess the status of marlin stocks based on revised data. However, an inconsistency between CPUE and fishing effort was found. Concern was expressed that the Japanese longline fishery, used in the past to index abundance for all Atlantic marlins, represented a decreasingly smaller percentage of the total billfish catch—down to roughly 10% in 1979, compared to 95% over the period 1960-1964, possibly reflecting changes in the fishing strategies over time by the Japanese fleet. It was concluded that it could not "be determined if exploitation levels were above optimum," but "that high levels of effort and yield had been followed by declining yields," with a decline in CPUE over time. Farber and Conser (1983) updated the marlin assessments of Farber (1982) using the catch and effort data through 1980 and followed the same methodology and assumptions as previous assessments. In that analysis, the generalized production model did not fit the data well under either stock structure assumption. Nevertheless, separate estimates of maximum sustainable yield for the north, south, and total Atlantic Ocean were again provided (Table 1) for blue and white marlin.

Assessments in the 1990s

Available Indices of Abundance for Blue and White Marlin

The Second ICCAT Billfish Workshop, held in Miami in 1992 (ICCAT 1994), was significant in that existing indices of abundance were refined and newly standardized CPUE series were developed for several countries that catch significant amounts of blue and white marlin (Fig. 2). For example, a major problem hindering improved stock assessments for all billfish species was accounting for changes in fishing strategy for the Japanese Atlan-

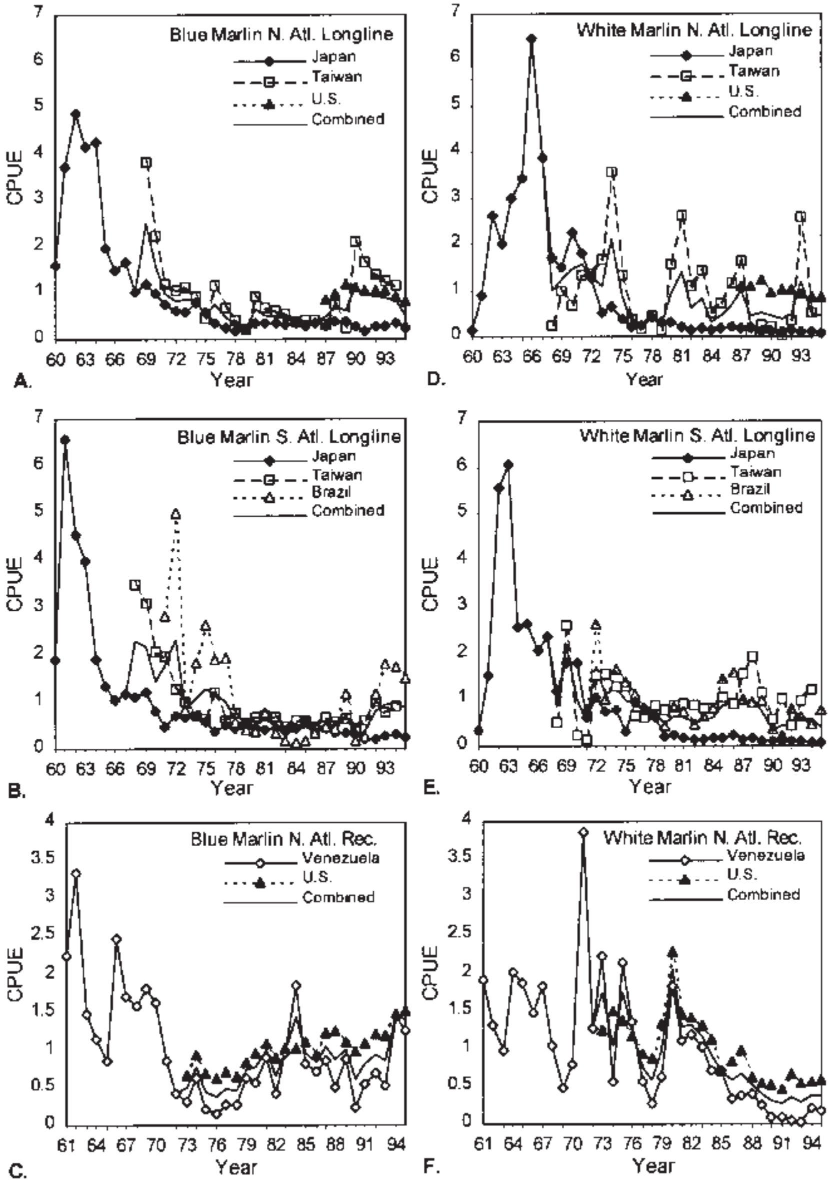


Figure 2. Standardized relative indices of abundance, including composite series, for fisheries used in the 1996 analysis. Fig. 4A contains the blue marlin North Atlantic longline CPUE trajectories; Fig. 4B is the blue marlin South Atlantic longline series; Fig. 4C is the blue marlin North Atlantic recreational series; Fig. 4D is the white marlin North Atlantic longline CPUE series; Fig. 4E is the white marlin South Atlantic longline series; Fig. 4F is the white marlin North Atlantic recreational series.

tic longline fleet in the mid- to late-1970s (ICCAT 1991). The historical Japanese longline data series was considered to represent two distinct fisheries due to changes in the tuna target species: the earlier 1961-1979 period using regular longline techniques, and the more recent 1980-1990 period using deep longline techniques. The differences between methods are detailed in Uozumi and Nakano (1994) and basically refer to the gear configuration and deployment that corresponded to changes in target species and spatial changes in effort. The changes were addressed for blue marlin by Nakano et al. (1994a) and for white marlin by Nakano et al. (1994b) by standardizing marlin CPUEs from the Japanese Atlantic longline fishery using a general linear model (GLM) and the Honma (1974) method. This permitted standardized CPUE indices to be presented for the entire historical time-series, 1960-1989, while accounting for shifts in fleet effort and deployment patterns.

Supplementary CPUE series in most cases were standardized to account for gear and geographical effects. These included the blue and white marlin U.S. recreational fishery in the North Atlantic for 1973-1991 (Farber et al. 1994), and the Venezuelan recreational fishery in the North Atlantic for 1961-1990 (Gaertner and Alio 1994). In addition to these series, standardized CPUE series were developed at the workshop for the Brazilian longline fishery in the South Atlantic (Amorim et al. 1994, Antero-Silva et al. 1994) and the Taiwanese longline fishery (ICCAT 1994). The Brazilian series was, however, based on limited data, and continued development was recommended at the time. The Taiwanese series was developed at the workshop without the advice of Taiwanese scientists (who were not present), and was therefore considered a tentative, and potentially unreliable, data series.

The Third ICCAT Billfish Workshop, held in Miami in 1996 (ICCAT 1996), provided greater opportunities for global cooperation in advancing the state of billfish assessments. This workshop marked the first time representatives from all major Asian longline fleets fishing the Atlantic attended an ICCAT Billfish intersessional meeting. Here, five countries submitted standardized indices of abundance for blue marlin and white marlin fisheries (Fig. 2). The standardized blue marlin and white marlin CPUEs from the Japanese Atlantic longline fishery for 1960 to 1995 were presented by Uosaki (1996) and Uozumi (1996), respectively. Taiwanese scientists attended the billfish workshop for the first time and provided corrected and updated standardized CPUE estimates for the Taiwanese longline fishery from 1967 to 1994 (Hsu 1996). A scientist from Korea also attended the billfish workshop for the first time and provided revisions for the Korean nominal billfish landings data. The U.S. longline standardized CPUE for the years 1987-1995 (Cramer 1996) was presented and the Brazilian longline data for 1971 to 1995 (Amorim et al. 1996) was substantially revised and updated at the workshop. In addition, the Venezuelan (Gaertner and Alio 1996) and U.S. (Jones et al. 1996) recreational CPUE data were revised and updated through 1995.

Blue Marlin

The 1992 Assessment. Cramer and Prager (1994) presented the first multifishery exploratory stock assessment analysis for blue marlin using the ASPIC approach. Total catches from 1960 to 1990 (Fig. 1), were compiled from ICCAT statistical tables. Longline catches were matched with their respective longline CPUE series, recreational catches with the recreational series, and any additional unallocated longline catches were indexed by the Japanese longline series. Because questions of stock delineation were still unresolved at this time, the assessment was analyzed using North Atlantic, South Atlantic, and total Atlantic stock structure assumptions.

In the North Atlantic runs, weighted and unweighted (inverse variance weighting) models were fitted with and without the Taiwanese data series. Where the Taiwanese series was not used, all Taiwanese catch was indexed by the Japanese CPUE. It was quickly determined that this series was needed to estimate model parameters, as fits without this index resulted in unrealistically large estimates of stock biomass. Hence, the Taiwanese data series was used in this assessment. The resulting estimates of MSY (Table 1) were lower than those estimated from previous assessments. The relative biomass trajectory (Fig. 3A) demonstrated a precipitous drop in the mid-1960s, with an upward trend in the late 1980s, but at levels below historical highs. The relative fishing mortality (Fig. 3A) had much greater fluctuations, at times greater than twice the optimum fishing mortality, though there was a slight decline in the late 1980s.

The addition of the Taiwanese data series had little effect on the estimated parameters in the South Atlantic. However, the weighted and unweighted model gave different results, with the unweighted model estimating a higher MSY and slightly less optimistic relative biomass. Nevertheless, both estimates resulted in relative biomass ratios less than half of that required to produce MSY in the terminal year (Fig. 3B).

Estimates of model parameters for the total Atlantic stock assumption were generated using all available data series. The computed MSY estimate (Table 1) was also lower than any previous assessment, though higher than most annual catches in the two decades prior to the assessment. The relative biomass (Fig. 3C) was very low throughout most of the time series. It was concluded that the stock could not support such a high level of harvest, and that relative fishing mortality well exceeded the optimum level for the total Atlantic (Cramer and Prager 1994).

The 1996 Assessment. The question of north/south blue marlin stock delineation was explored in much greater detail at the 1996 ICCAT Billfish workshop. One of the major advances in the 1996 assessment was that progress in genetic and tagging studies allowed the scientific working group to conclude that a total Atlantic stock structure assumption for both blue marlin and white marlin was most appropriate for stock assessment. However, the working group also concluded that biological evidence concerning stock structure was not totally definitive and therefore recom-

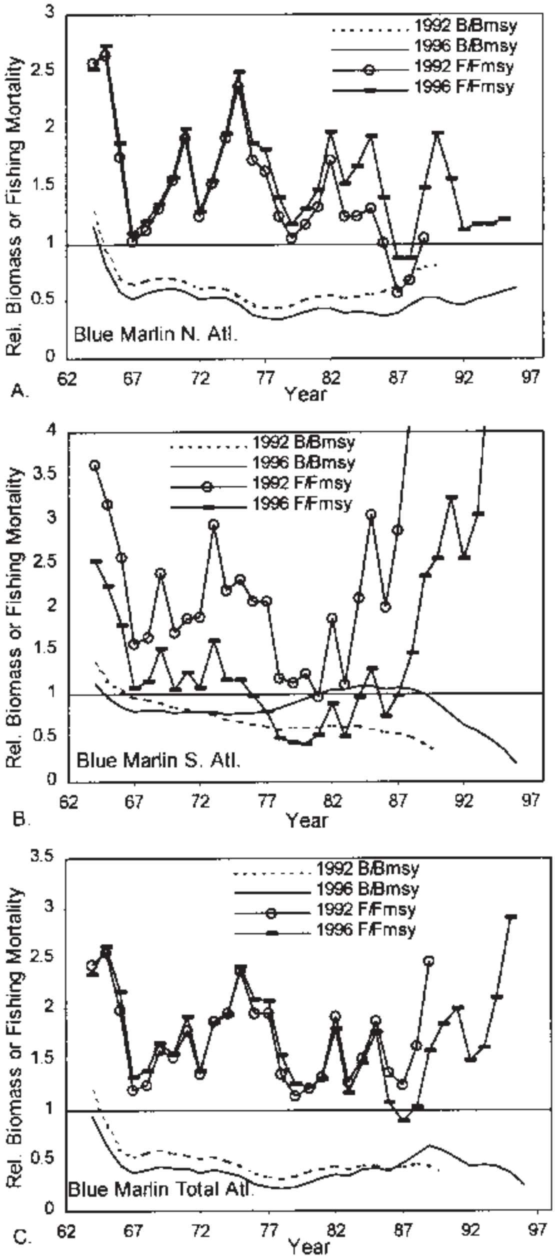


Figure 3. Annual relative biomass and relative fishing mortality trajectories from 1992 and 1996 assessments for blue marlin under the North Atlantic (A), South Atlantic (B) and total Atlantic stock assumption (C).

mended that north and South Atlantic runs also be done as a prudent approach to the assessments.

The initial fits of the North Atlantic model incorporated five available standardized CPUE series for the North Atlantic (Japanese longline, Taiwanese longline, U.S. longline, Venezuelan recreational, and U.S. recreational; Figs. 2A and 2C) for the entire time series (ICCAT 1996). These fits were unsuccessful, due to negative correlations among the longline CPUE series, as well as between the recreational and longline series. Although these negative correlation problems were not severe in terms of the trends across entire time series (Fig. 2), resolving this problem prior to model fitting presented a considerable challenge to the working group. Several alternative models were proposed, including using Japanese, or Taiwanese, CPUEs only to index all longline catches, and fitting models with separate catchability coefficients for the longline fisheries in the 1960s. Finally, composite CPUE indices were constructed for the longline and recreational series. This provided a way of incorporating all information, and confronting the negative correlation problem. The longline and recreational composite CPUE indices were constructed by first dividing all CPUEs for each series by the mean CPUE of the overlapping years. In years where there was only one CPUE available (e.g., Japanese longline 1960-1968, and Venezuelan recreational 1961-1972), only that estimate was used. In years with overlapping indices, averages of the adjusted CPUE estimates were used. The result was a separate composite CPUE for the longline fishery and for the recreational fishery (Figs. 2A and 2C). These composite indices were fairly well correlated ($r = 0.76$). The longline composite indices were matched to all longline catches and the recreational composite index was matched to all recreational catches.

Both longline and recreational composite indices of abundance in the North Atlantic for the 1960s exhibited initial rapid declines, and model runs using catch data from 1960 to 1995 failed to capture the dynamics of this stock response. The working group extensively discussed the issue of whether the rapid declines in CPUE and landings were real or an artifact of data collection. The working group concluded that because the peak in fishing effort during the mid-1960s was at about twice the level of MSY (as calculated in 1996), the rapid declines were likely accurate reflections of population responses to such high rates of fishing mortality. The group felt that such steep declines might well be expected from any fish populations exposed to such heavy fishing pressure. The solution to this problem was to use catches for blue marlin dating back to 1956, when the stock was believed relatively unexploited, fix the B_1R parameter to 2.0 (the equilibrium level for an unexploited stock), and estimate the populations rate of growth, r, and the separate catchabilities, $q(i)$, for the longline and recreational composite series. In addition, a 5% two-sided tail trim was applied to the residual distribution to mitigate the effects of possible statistical outliers during the fitting.

The resulting model fits for the North Atlantic stock assumption generated MSY estimates (Table 1) similar to the estimates of Cramer and Prager (1994), though with a slightly greater range. The relative biomass trajectory (Fig. 3A) was also very similar to that of Cramer and Prager (1994), for overlapping years, though this appears to be scaled down slightly. The relative fishing mortality (Fig. 3A) also tracked well with the 1992 assessment. As with previous assessments, the greatest impact on the stock biomass appears in the early 1960s.

Jones and Farber (1996) conducted the South Atlantic blue marlin assessment separately from the other 1996 assessments. The available CPUE series for this analysis included the Japanese, Taiwanese, and Brazilian longline data sets (Fig. 2B). Initially, models were attempted using all data sets, because attempting to fit models with a single composite CPUE series was unsuccessful. Due to difficulties finding minima during the model's search routine, it was necessary to fix the initial biomass, B_1R, similar to the North Atlantic model. The resulting model produced levels of MSY (Table 1) similar to that of Cramer and Prager (1994). However, biomass trajectories (Fig. 3B) for blue marlin were scaled considerably higher than that of any other trajectory for the mid-1980s. Nevertheless, models indicate there was a downturn of biomass through the 1990s, and the results suggest that the blue marlin stock is heavily overexploited under a South Atlantic stock assumption. This downturn in biomass corresponds to very high estimated relative fishing mortality levels in the 1990s (Fig. 3B).

The assessment for blue marlin under a total Atlantic stock assumption, considered the superior approach by the working group, combined all available information for Atlantic blue marlin. Along with freely estimated B_1R and r parameters, three separate q estimates were derived: a North Atlantic composite longline (Japanese, Taiwanese, U.S.); South Atlantic composite longline (Japanese, Taiwanese, Brazilian); and a North Atlantic composite recreational (Venezuelan, U.S.). The estimates of MSY from these model parameters were slightly higher than those of the 1992 assessment (Table 1). The relative biomass ratio (Fig. 3C) shows a similar drop, as has been the case with all models, in the 1960s, a general increase in the 1980s, followed by a decline from 1989 to 1996.

White Marlin

The 1992 Assessment. In preparation for the ASPIC production models for white marlin, Farber and Jones (1994) indexed white marlin abundance from 1961 to 1990 (Fig. 1) using standardized CPUE series for the Japanese and Taiwanese longline and the U.S. and Venezuelan recreational data series for the North Atlantic (Fig. 2D), and the Japanese and Brazilian longline for the South Atlantic (Fig. 2E). The total Atlantic analysis used all series (Fig. 2F). In the analysis, all ASPIC models were fits using yield (t) and fishing effort by fleet, *f*. Initially, models were fitted using calculated fishing effort, *f*, as catch/CPUE. This methodology proved inappropriate with these data, with the model either not converging to any solution or the fit being extremely poor with unreasonable parameter estimates. A plot of the average weight of individual fish (i.e., catch in weight/catch in number), for both the North Atlantic and South Atlantic data exhibited great variability and was unrealistically high for several years. An alternative method of estimating *f*, based on catch in weight (=catch [t]/CPUE [number/unit effort]) was assumed a more reliable statistic than the catch in number, given that the average weight of white marlin was presumed reasonably constant over the period considered. The values of *f* for all longline fisheries were derived using the CPUE series from the Japanese fleet.

For the North Atlantic model, large fluctuations were evident in these CPUE series from 1961 through 1965. This resulted in difficulties in model fitting (large residuals), and an inability to capture the dynamics of the large catch and effort fluctuations during this period. Beginning the longline series in 1966, along with U.S. recreational data for 1973-1990 and Venezuelan recreational data for 1966-1990, mitigated this problem. As a result, the models were successfully fit, with effort residuals reasonably balanced with no apparent trends or extreme values. The computed estimates of MSY (Table 1) were well below estimates from previous assessments. The North Atlantic relative biomass was less than 1.0 for all years after 1974 (Fig. 4A), and was estimated at the start of 1991 to be about 57% of the biomass that could produce MSY. The relative annual fishing mortality (Fig. 4A) showed considerable annual variability, and was estimated well above the optimum level for most years except during 1978-1980 and 1989-1990, when it was below optimum.

The South Atlantic white marlin analysis included the Brazilian longline CPUE series for the period 1971 to 1990 (Fig. 2E). Because there was considerably less fluctuation in the South Atlantic Japanese longline CPUE series for the years 1961 to 1965, this data series was retained. Estimates of MSY from this analysis were less than that of the previous assessment (Table 1), though not nearly as dramatic as the North Atlantic. The estimated relative biomass was less than that which could produce MSY for the entire time series and exhibits a declining trend to extremely low levels (Fig. 4B), with the estimated biomass at the start of 1991 only 3% of that which could produce MSY. Estimates of relative fishing mortality (Fig. 4B) were greater than 1.0, exhibiting variability without trend over the period 1962-1982, and then increasing sharply from 1.0 to very high levels over the period 1983-1990. Farber and Jones (1994) noted that during the mid-1970s, the Japanese longline catches of white marlin in the South Atlantic were extremely low, and questioned if the CPUEs for those years (used to calculate effort) were representative of the total South Atlantic longline catch.

For the total Atlantic analysis, all annual yields (t) were the arithmetic sums of the North Atlantic and South Atlantic data series. The point estimate of MSY (Table 1) was, like the north and south, well below the previous

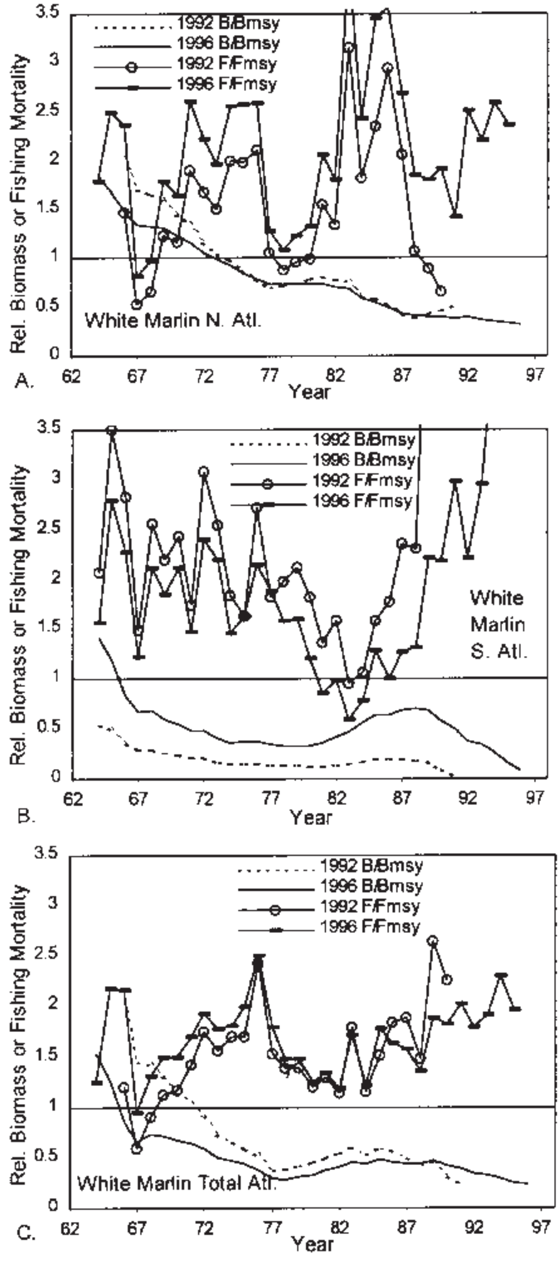


Figure 4. Annual relative biomass and relative fishing mortality trajectories from 1992 and 1996 assessments for white marlin under the North Atlantic (A), South Atlantic (B) and total Atlantic stock assumption (C).

assessment estimate. The relative biomass trajectory indicates a declining stock over the period 1966-1977, followed by an increase through the mid-1980s, and then decreasing again through 1991 (Fig. 4C). The estimated relative biomass was less than optimum for all years after 1972, with the estimated biomass at the start of 1991 at 25% of that which could produce MSY. The estimates of relative fishing mortality were greater than 1.0 for all years after 1969, with variability and periods of both increasing and decreasing trend (Fig. 4C). The conclusion for the total Atlantic from this assessment was that white marlin were "at least fully exploited with a strong possibility of substantial overexploitation during the last 17 to 20 years."

The 1996 Assessment. Catch and effort data available for white marlin were similar to blue marlin. Available CPUE series were from the Japanese, Taiwanese, and Brazilian longline fisheries, and the Venezuelan and U.S. recreational fishery (Figs. 2D,2E,2F). Similar to the blue marlin runs, the initial models were not fitted due to a lack of correlation between indices of abundance. Separate series of composite indices for the North Atlantic longline, the South Atlantic longline, and the North Atlantic recreational CPUE series were estimated to mitigate this problem. The procedures for constructing the combined CPUE series were identical to those used for the blue marlin. Further, an approximate 5% two-sided tail trim was applied to the composite indices to allay the potential effects of statistical outliers. As with blue marlin, the longline composite indices were matched to all longline catches and the recreational composite index was matched to all recreational catches.

The resulting MSY estimates from these model fits for the North Atlantic stock assumption (Table 1) were among the lowest of any previous assessment. However, relative biomass (Fig. 4A) and relative fishing mortality levels were very similar to that of the 1992 assessment, and demonstrate a continuous downward trend across the entire time series.

For the South Atlantic model, Jones and Farber (1996) were unable to fit models with single composite indices for Japanese, Taiwanese, and Brazilian fisheries due to problems with model convergence. They found it was necessary to fix the B_1R parameter, as well as the population intrinsic rate of increase, r to facilitate model convergence. Although there were several problems with the South Atlantic analysis, results were consistent with previous assessments, both in terms of MSY estimates (Table 1), and relative trajectories (Fig. 4B).

Data preparation for the total Atlantic analysis was similar to that of blue marlin, except the rod and reel (i.e., recreational) composite index was not used. Although rod and reel catches for white marlin represent a very low (about 5%) proportion of the total white marlin landings, the model could not capture the catch rate pattern observed in the earliest period, unless initial biomass was set at a biologically unrealistic low level (ICCAT 1996). The working group recommended that this series therefore be excluded. The resulting estimates of MSY from the fitted model were higher than the 1992 assessment for the total stock assumption (Table 1), though substantially less than the Farber and Conser (1983) estimate. Relative biomass levels (Fig. 4C) have declined since the mid-1980s, after a slight period of recovery in the early 1980s (ICCAT 1996). Relative fishing mortality (Fig. 4C) appears to be about twice the optimal level necessary to produce MSY.

Discussion

The early production model assessments of blue and white marlin relied exclusively on the Japanese longline CPUE to index abundance and did not adequately represent the complexity and dynamics of multiple countries fishing a widely distributed stock with different gear types, representing both directed and incidental bycatch fisheries. Nevertheless, both single and multiple fishery-based assessment approaches indicate levels of biomass for Atlantic blue marlin and white marlin have greatly declined over the time series. Assessments suggest that blue marlin and white marlin have been fully exploited since the 1960s, and are likely currently heavily overexploited.

One of the recommendations resulting from the workshop was to evaluate whether critical model assumptions were met. There was particular concern that the landings data was incomplete or systematically underreported. This is a difficult problem to detect and address, since evaluating the implications of errors in annual landings by a constant percentage of under-reporting only adjusts absolute levels of maximum sustainable yield and biomass levels. Thus, there remains much uncertainty about these absolute levels, though less so for relative trends.

Conclusion

The evolution of assessment approaches represents substantial progress in analyzing the status of the blue marlin and white marlin stocks in the Atlantic Ocean. Increased representation of fishing nations at workshops, as well as the working group approach, has been instrumental in advancing the state of billfish stock assessments. The application of the ASPIC model permits the incorporation of all available data, giving a more comprehensive representation of the stocks than previous assessments that relied exclusively on Japanese longline CPUE data. This is particularly important because the Japanese longline data represents an increasingly smaller percentage of white marlin landings during the most recent years (though Japanese longline landings for blue marlin have increased in the most recent years). The trends in the 1996 assessment for the total Atlantic are similar to those described in the 1992 assessment, and although there remains considerable uncertainty in absolute trends of abundance and fishing mortality, relative levels appear to be more precisely estimated. Most important, all assessments, using both single and multiple index approaches, indicate that Atlantic blue marlin and white marlin stocks are overexploited, and possibly severely so.

The strength of using production model techniques to examine trends in stock abundance depends on many factors, including degree of density-dependence on recruitment, recruitment variability, and possible age-structured lags in the population's response to fishing pressure. Further improvements in billfish assessment will also be needed, if we are to provide advice on the expected dynamics of these stocks under potentially heavier exploitation. Evaluations of alternative underlying production functions, relating to overexploited stock productivity may provide a basis for better understanding low biomass dynamics of these stocks. This could be accomplished by implementing a generalized stock production model that would allow fitting of multiple indices of abundance. Methods for better incorporation of mixed unit abundance indices (e.g., in numbers and in biomass) need to be applied as well. Finally, incorporation of information now being developed on the size structure of the catches could also be used to improve assessments, possibly to better accommodate the initial declines in catch rate indices observed in these fisheries.

This paper is Miami Laboratory Contribution MIA-97/98-01.

References

Alverson, D.L., M.H. Freeburg, S.A. Murawski, and J.G. Pope. 1994. A global assessment of fisheries bycatch and discards. FAO Fish. Tech. Pap. 339. 233 pp.

Amorim, A.F., C.A. Arfelli, J.N. Antero-Silva, L. Fagundes, F.E.S. Costa, and R. Assumpcao. 1996. Blue marlin *(Makaira nigricans)* and white marlin *(Tetrapturus albidus)* caught off the Brazilian Coast. Int. Com. Cons. Atl. Tunas, Coll. Vol. Sci. Pap. SCRS/96/101.

Amorim, A.F., C.A. Arfelli, F.H.V. Hazzin, J.N., Antero-Silva, R.P.T. Lessa, and R.R. Areas. 1994. Blue marlin *(Makaira nigricans)* fisheries off Brazilian coast from national and leased longliners (1971-91). Int. Comm. Cons. Atl. Tunas, Coll. Vol. Sci. Pap. 41:208-213.

Antero-Silva, J.N., A.F. Amorim, R.P.T. Lessa, F.H.V. Hazzin, and C.A. Arfelli. 1994. White marlin *(Tetrapturus albidus)* fisheries off Brazilian coast from national and leased longliner fleet. Int. Comm. Cons. Atl. Tunas, Coll. Vol. Sci. Pap. 41:189-198.

Beardsley, G.L., and R.J. Conser. 1981. An analysis of catch and effort data from the U.S. recreational fishery for billfishes (Istiophoridae) in the western North Atlantic Ocean and Gulf of Mexico, 1971-1978. Fish Bull. U.S. 79(1):49-68.

Boggs, C.H. 1989. Vital rate statistics for billfish stock assessment. Marine Recreational Fisheries 13:225-234.

Conser, R.J. 1989. Assessing the status of stock of Atlantic blue marlin and white marlin. Marine Recreational Fisheries 13:153-164.

Conser, R.J., and G.L. Beardsley. 1979. An assessment of the status of stocks of blue marlin *(Makaira nigricans)* and white marlin *(Tetrapturus albidus)* in the Atlantic Ocean. Int. Comm. Cons. Atl. Tunas, Coll. Vol. Sci. Pap. 8(2):461-489.

Cramer, J.L. 1996. Large pelagic logbook catch rates indices for billfish. Int. Comm. Cons. Atl. Tunas, Coll. Vol. Sci. Pap. SCRS/96/110.

Cramer, J.L., and M.H. Prager. 1994. Refinements in exploratory surplus-production analyses of Atlantic blue marlin. Int. Comm. Cons. Atl. Tunas, Coll. Vol. Sci. Pap. 41:565-571.

Farber, M.I. 1982. An assessment of the status of stocks of blue marlin *(Makaira nigricans)* and white marlin *(Tetrapturus albidus)* in the Atlantic Ocean through 1979. Int. Comm. Cons. Atl. Tunas, Coll. Vol. Sci. Pap. 17(2):395-414.

Farber, M.I., and R.J. Conser. 1981. An assessment of the status of stocks of blue marlin *(Makaira nigricans)* and white marlin *(Tetrapturus albidus)* in the Atlantic Ocean. Int. Comm. Consv. Atl. Tunas, Coll. Vol. Sci. Pap. 15(2):387-406.

Farber, M.I., and R.J. Conser. 1983. An update on the status of stocks of blue marlin and white marlin in the Atlantic Ocean. Int. Comm. Cons. Atl. Tunas, Coll. Vol. Sci. Pap. 18(3):674-692.

Farber, M.I., and C.D. Jones. 1994. An exploratory stock-production model analysis of white marlin *(Tetrapturus albidus)* in the Atlantic Ocean. Int. Comm. Cons. Atl. Tunas, Coll. Vol. Sci. Pap. 41:572-587.

Farber, M.I., J.A. Browder, and J.P. Contillo. 1994. Standardization of recreational fishing success for marlin in the western North Atlantic Ocean, 1973-1991, using generalized linear model techniques. Int. Comm. Cons. Atl. Tunas, Coll. Vol. Sci. Pap. 41:363-392.

Fox Jr., W.W., 1975. Fitting the generalized stock production model by least-squares and equilibrium approximation. Fish. Bull. U.S. 73:23-37.

Furman, R.S. 1989. The adequacy of current billfish fisheries statistics for stock assessment and management purposes. Marine Recreational Fisheries 13: 215-224.

Gaertner, D., and J. Alio. 1994. Changes in the apparent abundance indices of billfishes in the Venezuelan recreational fishery off Playa Grande (1961-1990), Central Venezuelan coast. Int. Comm. Cons. Atl. Tunas, Coll. Vol. Sci. Pap. 41:473-489.

Gaertner, D., and J. Alio. 1996. Trends in the recreational billfish fishery CPUE off Playa Grande (1961-1995), Central Venezuelan coast. Int. Comm. Cons. Atl. Tunas, Coll. Vol. Sci. Pap. SCRS/96/109.

Graham, M. 1935. Modern theory of exploiting a fishery, and application to North Sea trawling. J. Cons. Cons. Int. Explor. Mer 37:199-204.

Honma, M. 1974. Estimation of overall effective fishing intensity of tuna longline fishery. Bull. 10, Far Seas Fish. Res. Lab. Shimizu 424, Japan, pp. 63-86.

Hsu, C.C. 1996. A review of catches, fishing efforts and catch per unit effort of the Taiwanese longline fishery for billfishes in the Atlantic. Int. Comm. Cons. Atl. Tunas, Coll. Vol. Sci. Pap. SCRS/96/100.

ICCAT. 1980. Report for biennial period, 1978-79. Int. Comm. Cons. Atl. Tunas. 280 pp.

ICCAT. 1981. Report of the ICCAT inter-sessional workshop on billfish. Int. Comm. Cons. Atl. Tunas, Coll. Vol. Sci. Pap. 16. 375 pp.

ICCAT. 1991. Report for biennial period, 1990-91. Int. Comm. Cons. Atl. Tunas. 445 pp.

ICCAT. 1994. Report of the Second ICCAT Billfish Workshop. Int. Comm. Cons. Atl. Tunas, Coll. Vol. Sci. Pap. 41. 587 pp.

ICCAT. 1996. Report of the Third ICCAT Billfish Workshop. Int. Comm. Cons. Atl. Tunas, Coll. Vol. Sci. Pap. COM-SCRS/96/19.

Jones, C.D., and M.I. Farber. 1996. Exploratory stock-production model analysis of blue marlin and white marlin in the South Atlantic Ocean. Int. Comm. Cons. Atl. Tunas, Coll. Vol. Sci. Pap. SCRS/96/159.

Jones, C.D., and E.D. Prince. 1996. The Cooperative Tagging Center mark recapture database for istiophoridae (1954-1995), with an analysis of the West Atlantic ICCAT billfish tagging program. Int. Comm. Cons. Atl. Tunas, Coll. Vol. Sci. Pap. SCRS/96/96.

Jones, C.D., M.T. Judge, and M. Ortiz. 1996. Standardization of recreational CPUE for blue and white marlin in the western North Atlantic Ocean 1973-1995. Int. Comm. Cons. Atl. Tunas, Coll. Vol. Sci. Pap. SCRS/96/104.

Kikawa, S., and M. Honma. 1978. Status of the white and blue marlins caught by the longline fisheries in the North Atlantic Ocean, 1956-76. Int. Comm. Cons. Atl. Tunas, Coll. Vol. Sci. Pap. SCRS/78/100.

Meyer, P.A. 1989. Socio-economic trends in recreational billfish fisheries. Marine Recreational Fisheries 13:103-108.

Nakano, H., Y. Uozumi, and M. Honma. 1994a. The CPUE trend for Atlantic blue marlin caught by Japanese longline fishery. Int. Comm. Cons. Atl. Tunas, Coll. Vol. Sci. Pap. 41:408-424.

Nakano, H., Y. Uozumi, and M. Honma. 1994b. The CPUE trend for Atlantic white marlin caught by Japanese longline fishery. Int. Comm. Cons. Atl. Tunas, Coll. Vol. Sci. Pap. 41:393-407.

NMFS. 1978. Report of the Billfish Stock Assessment Workshop, Atlantic Session. Dec. 1977, Honolulu, HI. NOAA/NMFS, SEFC, Miami, FL 33149. 47 pp.

NMFS. 1994. Cooperative game fish tagging program annual newsletter: 1992. NOAA Tech. Memo. NMFS/SEFSC-346. 23 pp.

Pella, J.J., and P.K. Tomlinson. 1969. A generalized stock production model. Bull. Inter-Amer. Trop. Tuna Comm. 13:419-496.

Prager, M.H. 1992. ASPIC: A surplus-production model incorporating covariates. Inter. Comm. Conserv. Atl. Tunas, Coll. Vol. Sci. Pap. 37:218-229.

Prager, M.H. 1995. A suite of extensions to a nonequilibrium surplus-production model. Fish. Bull. U.S. 92:374-389.

Schaefer, M.B. 1954. Some aspects of the dynamics of populations important to the management of the commercial marine fisheries. Inter-Am. Trop. Tuna Comm. Bull. 1(2):27-56.

Robson, D.S. 1966. Estimation of the relative fishing power of individual ships. Int. Comm. Northwest Atl. Fish. Res. Bull. 3:5-14.

Uosaki, K. 1996. The CPUE trend for Atlantic blue marlin caught by the Japanese longline fishery. Int. Comm. Cons. Atl. Tunas, Coll. Vol. Sci. Pap. SCRS/96/93.

Uozumi, Y.H. 1996. Standardization of CPUE for white marlin caught by the Japanese longline fishery in the Atlantic. Int. Comm. Cons. Atl. Tunas, Coll. Vol. Sci. Pap. SCRS/96/94.

Uozumi, Y.H., and H.J. Nakano. 1994. A historical review of Japanese longline fishery and billfish catches in the Atlantic Ocean. Int. Comm. Cons. Atl. Tunas, Coll. Vol. Sci. Pap. 41:233-243.

Evaluation of Multiple Survey Indices in Assessment of Black Sea Bass from the U.S. South Atlantic Coast

Douglas S. Vaughan
National Marine Fisheries Service, Beaufort Laboratory, Beaufort, North Carolina

Boxian Zhao, Mark R. Collins, John C. McGovern, and H. Scott Meister
South Carolina Department of Natural Resources, Charleston, South Carolina

Abstract

The age structure and status of the U.S. southern Atlantic stock of black sea bass were examined using recorded and estimated landings and size frequencies of fish taken from commercial, recreational, and headboat fisheries during 1979-1995. Fishery-independent data were obtained from the Marine Resources Monitoring, Assessment, and Prediction (MARMAP) program for calibrating virtual population analysis (VPA). MARMAP collected black sea bass with four gear types: hook-and-line (1979-1995), blackfish trap (1979-1989), Florida snapper trap (1980-1989), and chevron trap (1988-1995). Reduced effort with the hook-and-line gear in 1988 coincided with reduced catch per effort (CPE). The size fraction of fish captured with chevron trap was similar to the Florida snapper trap, but dissimilar to the blackfish trap. Hence, an extended time series of CPE was developed combining the Florida snapper and chevron traps. We examined the effect of different combinations of CPE indices (individually and in combination) on estimates of fully recruited fishing mortality, spawning stock biomass, and spawning potential ratio using the FADAPT VPA approach.

Introduction

Black sea bass, *Centropristis striata*, also called blackfish, is a serranid that inhabits continental shelf waters in depths of 2-120 m, predominantly between Cape Canaveral, FL, and Cape Cod, MA (Mercer 1989). Two populations are thought to occur along the Atlantic coast, separated by Cape Hatteras, NC (Mercer 1989, Shepherd 1991). This study is concerned only with the southern population.

Spawning of black sea bass occurs during January through June along the U.S. southern Atlantic coast, peaking during March to May (Wenner et al. 1986, Mercer 1989). Black sea bass are protogynous hermaphrodites, but mature males occur in all age groups (Vaughan et al. 1995, Table 1). Sex ratios and female maturity schedules were summarized by Vaughan et al. (1995). Because fish undergo sexual transition from female to male during a short time (weeks to a few months), the transitional stage was combined with males for calculation of sex ratios used in the analyses that follow.

In this study, we investigated the use of fishery-independent indices from the Marine Resources Monitoring, Assessment, and Prediction (MARMAP) program for calibrating virtual population analyses. The MARMAP program collects fishery-independent data from the South Atlantic Bight using a stratified random design. Because specific trapping gear types were not consistent over the entire study time period, a greater sampling duration was attained by extending backward the current trap gear (chevron). We then analyzed the black sea bass catch matrix (from Vaughan et al. 1996) using a calibrated virtual population analytic (VPA) approach with various combinations of the MARMAP gear indices. VPA output included estimates of fully recruited fishing mortality (ages 4-7), spawning stock biomass (both sexes combined), and spawning potential ratio (SPR). Because black sea bass are protogynous (transforming from females to males) and the effect of changes in population abundance on sex transformations is unknown, spawning potential ratio was based on total mature biomass.

MARMAP Gear Comparisons

The MARMAP program has collected fishery-independent reef fish data off the southeastern U.S. Atlantic coast since 1979. Evolution of the MARMAP sampling design was thoroughly described by Harris and McGovern (1997). Briefly, standard MARMAP sampling was conducted during daylight from May through August between Cape Fear, NC and Cape Canaveral, FL. Gear used included blackfish traps (1979-1989), Florida snapper traps (1980-1989), chevron traps (1988-present), and hook-and-line (1979-present) (Collins and Sedberry 1991). Fishing effort (soak time and bait type/amount for traps; angling time, bait type/amount, and terminal tackle for hook-and-line) was standardized for all gear types. Efficiencies of the three trap

Table 1. Maximum differences (observed and bootstrap) and Smirnov test statistics for length frequency comparisons among gear (blackfish, Florida snapper, and chevron traps) and between year (1988-1989) from MARMAP data.

		Maximum differences			Smirnov statistic[c]	
Gears	Year	Observed	Boot[a]	Boot[b]	$\alpha = 0.005$	$\alpha = 0.05$
Blackfish	1988-89 (243,176)	0.060	0.087	0.056	0.161	0.121
Florida snapper	1988-89 (401,152)	0.135*	0.146*	0.052	0.155	0.116
Chevron	1988-89 (974,669)	0.059	0.062*	0.028	0.082	0.061
Blackfish-chevron:						
(243,974)	1988	0.371**	0.372**	0.331**	0.117	0.087
(176,669)	1989	0.278**	0.287**	0.331**	0.138	0.103
FL Snapper-chevron:						
(401,974)	1988	0.105**	0.109**	0.097**	0.097	0.072
(152,669)	1989	0.063	0.083	0.104	0.147	0.110

Sample sizes listed in parentheses under comparison variable. Note: * indicates Smirnov test significant at 0.05 level, and ** indicates Smirnov test significant at 0.005 level for individual test.

[a] Median from 1,000 bootstrap simulations with years separate.

[b] Median from 1,000 bootstrap simulations with years pooled, then reconstructed.

[c] Individual test based on $\alpha = 0.005$, overall test based on seven independent tests (Bonferroni's theorem) is $\alpha = 0.035$.

types were compared during 1988-1989, and use of blackfish and Florida snapper traps was discontinued in 1990 because the chevron traps sampled a greater species diversity (Collins 1990, design and construction of the traps also presented). Samples were randomly collected from four broad areas of live bottom (identified with underwater TV) during 1979-1987. Since then, sampling has been on a stratified (by latitude) random basis, with 300-600 sites/yr randomly chosen from a data base of over 2,500 known live bottom locations.

Black sea bass were measured (total and standard lengths), and sagittal otoliths were removed and stored dry. For each year (1979-1995) the commonly sampled sizes (150-259 mm TL) were divided into 10 mm TL classes and 16-26 individuals were randomly subsampled. All individuals larger than 259 mm TL or smaller than 150 mm TL were examined since these individuals were infrequently encountered. The number of black sea bass aged for each year was approximately 400. Otoliths were placed in water and read whole with transmitted light using a Nikon SMZ-2T dissecting microscope. Aging was done without prior knowledge of the size of the fish or the date of capture. If the two readers disagreed on an age or considered the otolith unreadable, that fish was deleted from analyses. Wenner et al. (1986) used marginal increment analysis to validate the annual nature of increments on black sea bass otoliths.

Declines in catch per effort (CPE) were noted in indices based on MARMAP sampling using hook-and-line and traps (Fig. 1a-b). There was a precipitous decline in hook-and-line CPE during 1987 and 1988, when sample sizes per gear were reduced for comparison sampling while introducing new trapping gear (chevron). Less drastic declines were noted in blackfish trap and Florida snapper trap CPEs during the middle to late 1980s. Unfortunately, both of these indices were discontinued after 1989, while the use of the chevron trap began in 1988. All three trap gear types were fished synoptically during 1988-1989.

Development of an extended trap gear to represent the whole time period was needed for greater comparison of a calibration index with the earlier, converged portion of the virtual population analysis. Recent indices of short duration can have a great impact on VPA estimates, because they occur primarily during the unconverged portion of the VPA (Vaughan et al. 1988). However, without demonstrating some correspondence to the converged portion of the VPA, it is difficult to judge the usefulness of the index as a calibration tool.

The question arises as to which historical trap gear (blackfish or Florida snapper trap) should be used to extend backward the duration of the chevron trap? The cumulative total length frequencies for the two years of overlap (1988-1989) for the three trapping gear types (Fig. 2a-b), showed considerable similarity between the chevron trap and the Florida snapper trap, but not between the chevron trap and the blackfish trap.

Kolmogorov-Smirnov-Type nonparametric tests (Conover 1971) allow comparisons among the maximal difference between cumulative (length) distribution functions. The observed maximal differences among gear types and between years were summarized (Table 1). To compare these differences to the appropriate values of the Smirnov Test Statistic (Conover 1971, Table 17), it was necessary to apply Bonferroni's correction, because of inflation of α (level of significance) with the conduct of multiple "independent" tests ($\alpha' = \alpha/n$, where n is the number of independent test). Since seven tests were conducted, an overall Type I error (α) of about 0.035 resulted from seven individual test statistics (α') based on (0.005). The large sample approximation for the Smirnov Test Statistic was $s[(m+n)/mn]^{1/2}$ where $s = 1.63$ for $\alpha' = 0.005$ and $s = 1.22$ for $\alpha' = 0.05$. These comparisons suggested that the chevron length frequency was different from the blackfish length frequency for each year, and different from the Florida snapper length frequency during 1988 (Table 1). All between-year comparisons (1988-1989) indicated no significant differences within gear types at $\alpha = 0.005$, but at $\alpha = 0.05$ differences were suggested between years for the Florida snapper trap.

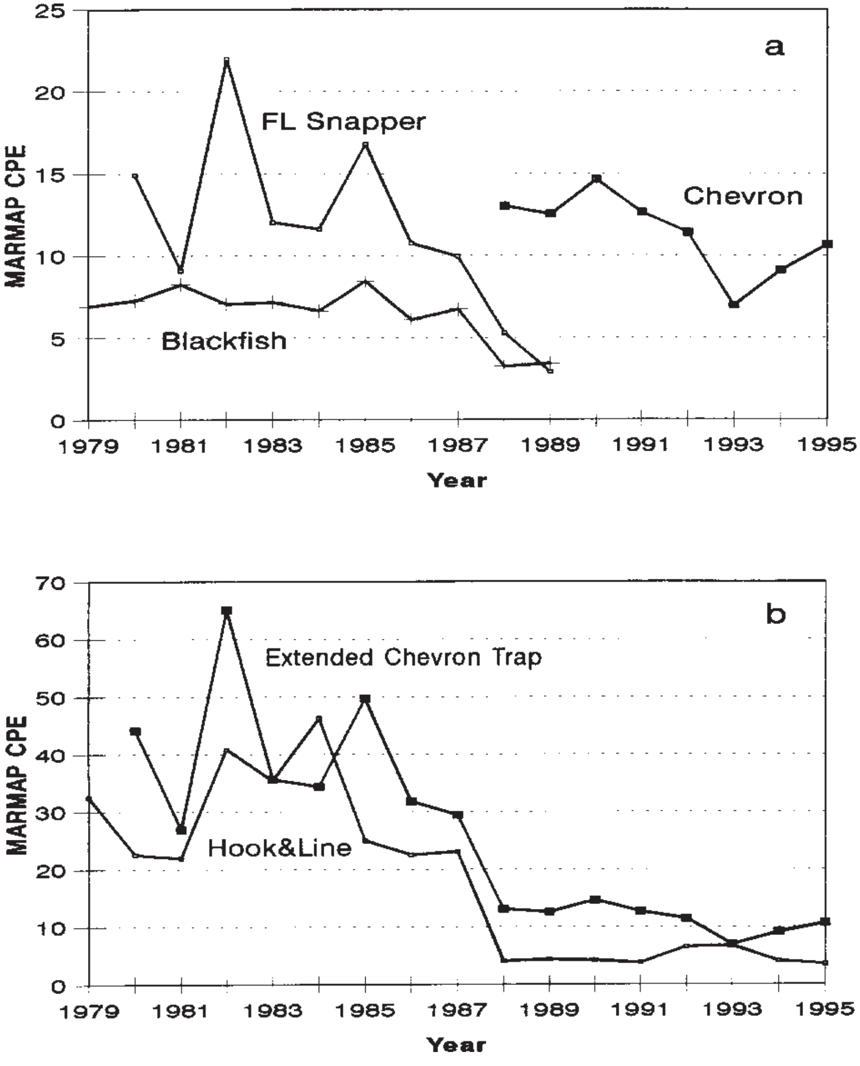


Figure 1. Comparison of CPE from MARMAP indices of black sea bass abundance from (a) trap gears and (b) extended chevron trap and hook-and-line gears.

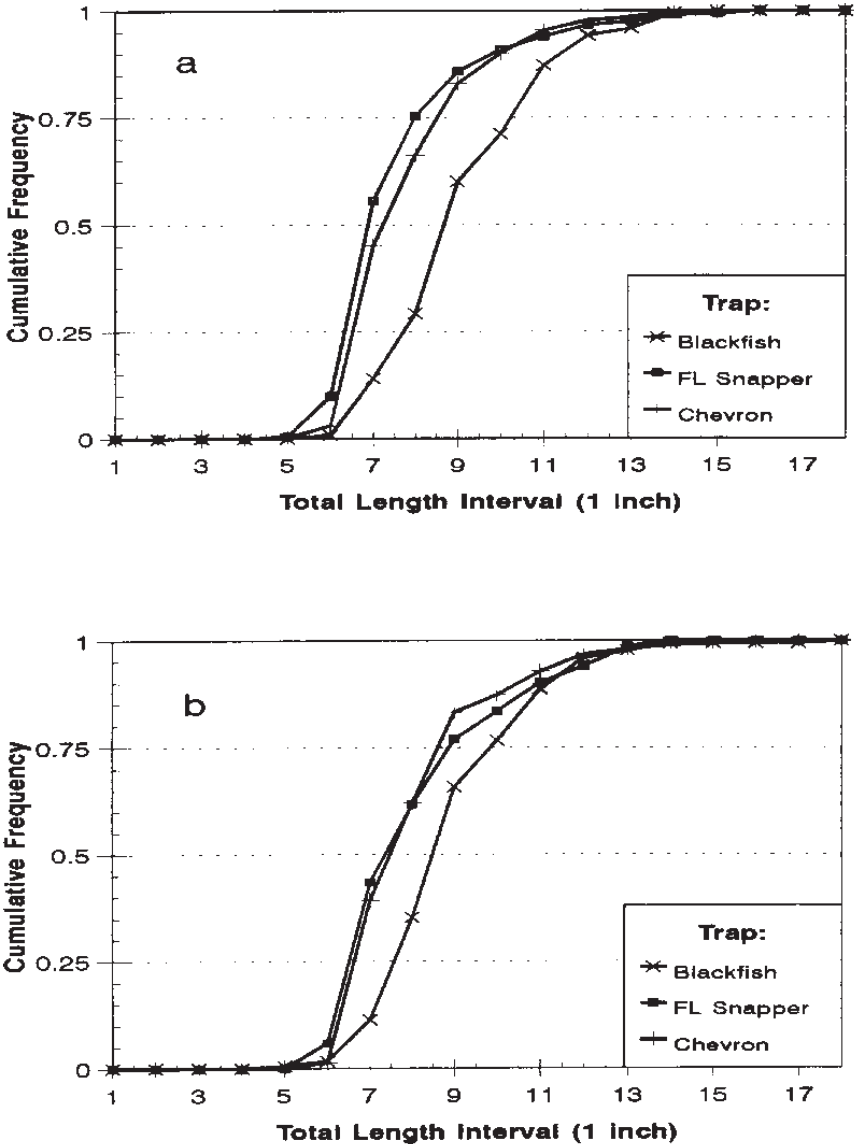


Figure 2. Comparison of black sea bass cumulative length frequency data from MARMAP traps (blackfish [BFT], Florida snapper [FST], and chevron [CHV]) for (a) 1988 and (b) 1989.

All of the between-gear differences may not have been biologically significant when comparing between gear differences to within gear (year-to-year) differences. Relatively large sample sizes available for length frequency comparisons allow one to detect small, but statistically significant, differences between length frequency distributions. Analogous to Helser (1996), we used a bootstrap approach to determine whether differences between different gear types exceeded differences from natural variability between the two years for which concurrent data were available. By sampling repeatedly with replacement (1,000 bootstrap experiments), we reconstructed the original data sets. The median values of maximal differences from the 1,000 bootstrap experiments are summarized (Table 1). A second series of simulations (of 1,000 bootstrap experiments) was conducted for which sampling with replacement was from gear data pooled across years, and then randomly resorted between the two years for each gear. Again, the median values of maximal differences from the 1,000 bootstrap experiments are summarized (Table 1). We then visually compared the maximal difference between cumulative length distributions between years of the same gear to maximal differences between gear types in the same year (Fig. 3a-b; blackfish-chevron traps and Florida snapper-chevron traps, respectively). We noted a large separation between within-gear maximal differences (between years for same gear) and between-gear maximal differences when comparing chevron with blackfish trap, suggesting real differences between the sampling properties of these two gears. However, the large overlap among maximal differences when comparing chevron with Florida snapper trap suggested that differences between years for the Florida snapper and chevron traps exceeded differences between the two traps for each year.

Hence, we concluded that the Florida snapper trap CPE could serve as the basis for extending the chevron trap CPE to include the earlier period of 1980-1987 (Fig. 1b). The conversion was accomplished by multiplying Florida snapper trap CPE in numbers at age by 2.96 (Collins 1990, with additional samples).

Calibrated Virtual Population Analyses

The catch-at-age matrix was developed from landings (commercial, recreational, and headboat) and corresponding length frequency sampling by fishery and gear for 1979 through 1995 using annual age-length keys developed from MARMAP sampling of similar gear types (traps and hook-and-line) as described in Vaughan et al. (1996) (Table 2).

The analytic program calibrates the virtual population analysis (VPA) to fishery-independent indices of abundance (Pope and Shepherd 1985). Specifically, FADAPT (modified from ADAPT [Gavaris 1988] and described in Restrepo [1996]) was the calibration-based VPA program used to analyze the above described catch-at-age matrix with various combinations of the MARMAP CPE indices:

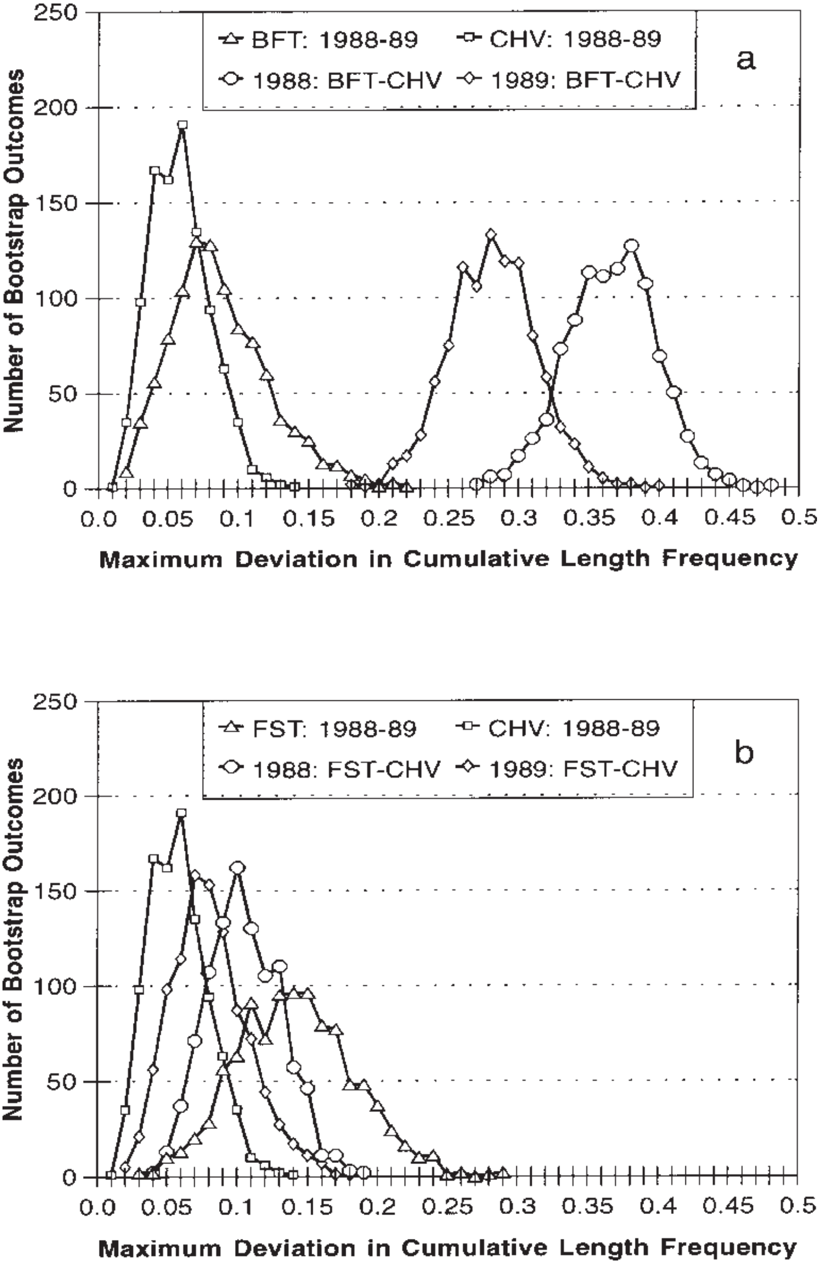


Figure 3. *Comparison of bootstrapped maximum differences in black sea bass cumulative length frequency data from MARMAP traps for 1988-89 between (a) chevron-blackfish traps, and (b) chevron-Florida snapper traps. Year identity is maintained in these bootstraps.*

Table 2. U.S. southern Atlantic black sea bass catch-in-numbers-at-age (in thousands) matrices for ages 1 through 8+ (and total numbers) and years 1979 through 1995.

Year	Age (yr)									Number	Weight
	0	1	2	3	4	5	6	7	8+	(1,000)	(1,000 lbs)
1979	2.4	303.2	**1381.4**	1250.2	710.5	360.2	229.2	54.7	19.1	4310.9	2433.2
1980	7.8	160.3	799.1	**1715.1**	1016.9	321.3	149.5	34.3	30.6	4234.9	3585.6
1981	6.4	232.4	632.2	**819.7**	729.3	278.2	83.9	29.9	11.6	2823.4	2285.0
1982	6.0	246.8	1213.9	**1539.7**	1236.1	264.0	95.6	17.5	3.9	4623.6	2888.4
1983	3.0	293.0	795.7	**837.5**	522.5	206.8	22.6	3.8	1.2	2686.0	1853.5
1984	5.8	54.5	388.6	1098.2	**1298.4**	436.3	139.7	10.6	1.9	3434.0	2802.1
1985	23.8	350.4	616.7	**984.6**	791.4	240.7	42.5	3.0	1.0	3054.3	2136.6
1986	0.0	126.5	**891.6**	547.4	505.4	270.5	52.6	9.2	2.2	2405.4	1746.3
1987	0.0	228.0	686.0	**1012.7**	512.6	252.3	67.5	23.6	1.9	2784.7	2252.8
1988	3.3	71.8	727.4	1024.1	**1122.3**	397.4	124.0	47.2	5.1	3522.5	2918.4
1989	4.4	83.0	470.5	**1191.6**	1010.5	395.5	74.9	17.6	9.1	3257.0	2483.2
1990	0.0	38.9	556.9	**1134.3**	685.1	329.2	124.8	32.5	6.8	2908.5	2021.9
1991	0.0	52.2	637.9	**1001.5**	505.6	275.9	72.2	14.0	6.0	2565.1	2002.1
1992	0.0	11.2	489.1	**893.9**	491.8	316.2	74.7	4.8	8.6	2290.3	1659.5
1993	0.1	25.0	383.0	**829.5**	344.9	184.4	59.5	6.0	6.8	1839.3	1440.3
1994	0.0	10.4	183.3	397.3	**832.8**	217.6	116.8	15.8	7.2	1781.3	1567.7
1995	0.0	8.4	219.8	**411.5**	336.0	309.1	111.8	8.7	5.0	1410.3	1324.6

Total catch in weight (thousands of pounds) are also presented.

Note that 5% of catch-release recreationally caught fish (type B2 fish from MRFSS) are included in estimates by number (modal age bold).

1. Four gear CPE indices.
2. Hook-and-line CPE only.
3. Three trap CPE indices.
4. Chevron trap CPE only.
5 Extended chevron trap CPE only.
6. Extended chevron trap and hook-and-line CPE.

The first combination uses all four indices separately in a single run. The second combination considers only the hook-and-line gear, while the following three combinations consider only trap gear. The final combination considers the extended chevron trap gear and hook-and-line together.

Virtual population analysis sequentially estimates population size and fishing mortality rates for younger ages of a cohort from a starting value of fishing mortality for the oldest age. An estimate of natural mortality, usually assumed constant across years and ages, is also required. We used the FORTRAN program for separable virtual population analysis (SVPA) developed by Clay (1990) to estimate the selectivity pattern for recent years to serve as input to the FADAPT runs. For the SVPA runs, starting values of F were based on the mean of the final three year class (1987-1989) estimates of Z (1.0 yr^{-1}) and final F obtained by subtracting M from Z.

The output from the calibrated FADAPT VPA considered in this study (Figs. 4-5) included full F (weighted mean over ages 4-7), spawning stock biomass (SSB), and spawning potential ratio (SPR). Most biological reference points used as targets or overfishing thresholds in fishery management in the United States are based on these variables (Mace and Sissenwine 1993, Mace 1994, Rosenberg et al. 1994).

The age-specific estimates of F for ages 4-7 were weighted by catch in numbers at age to obtain full F. Spawning stock biomass was calculated from mean weight at age by year multiplied by the number of fish and by the percent mature males and females. Spawning potential ratio (SPR) compares spawning stock biomass per recruit with and without fishing. All other life history parameters are usually held constant (e.g., maturity schedule and age-specific sex ratios) in calculating SPR, so the estimate of SPR increases as fishing mortality decreases.

Full F, SSB, and SPR from the FADAPT runs with different sets of CPE indices were compared (Figs. 4-5). As noted above, the starting partial recruitment vector for the FADAPT runs was based on an SVPA run for the period 1990-1995 (this minimized the coefficient of variation based on several SVPA runs of varying duration with final year of 1995).

First, we compared FADAPT runs calibrated to all four indices (hook-and-line and 3 separate unexpanded traps), three separate unexpanded trap indices, and the hook-and-line index only (Fig. 4). The FADAPT run calibrated to the hook-and-line index suggested a considerable increase in full F and moderate decline in SPR in 1995 as compared to FADAPT runs calibrated with the trap indices. Although full F and SPR showed no temporal trend, SSB declined after 1987. The slight increase in SSB in 1992

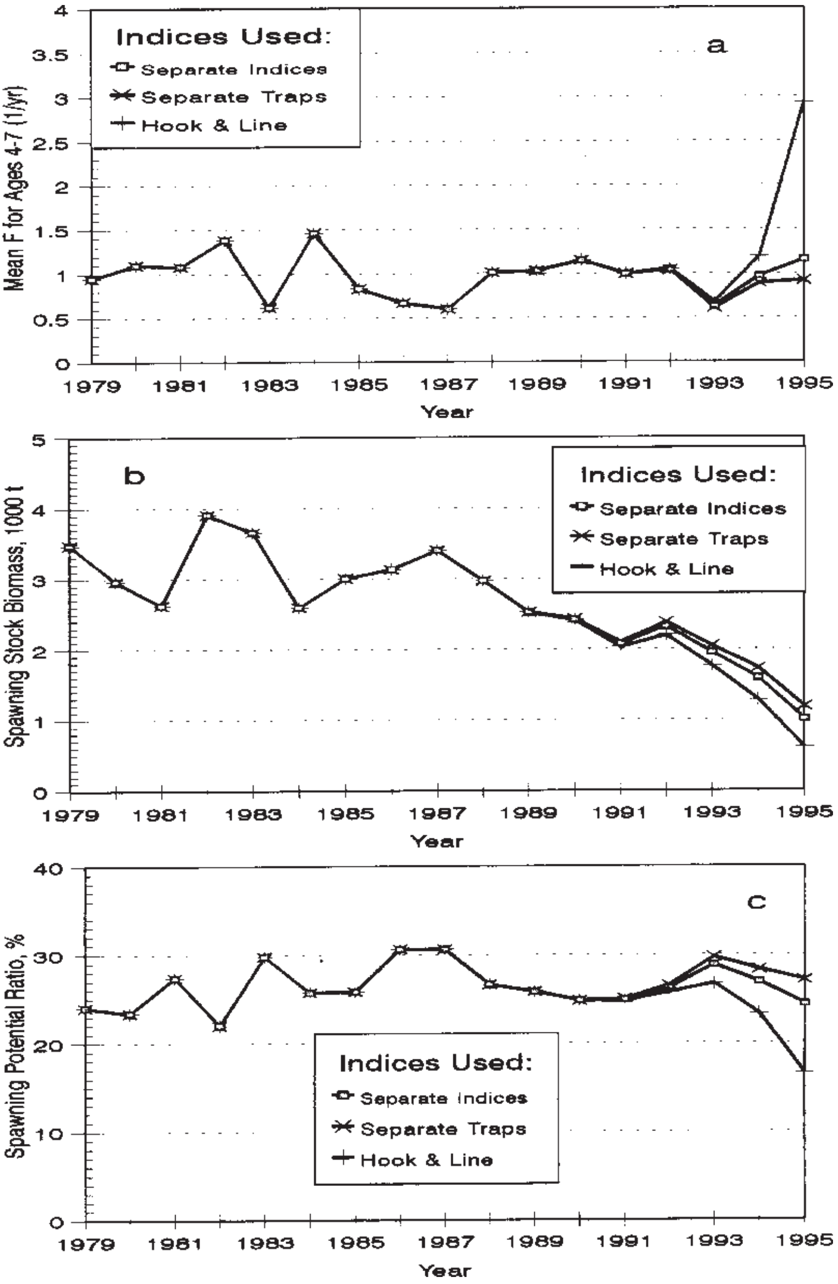


Figure 4. Comparison of full F, SSB, and SPR for black sea bass for MARMAP CPE indices: (a) hook and line and 3 individual unexpanded traps, (b) 3 individual unexpanded traps, and (c) hook and line using FADAPT method.

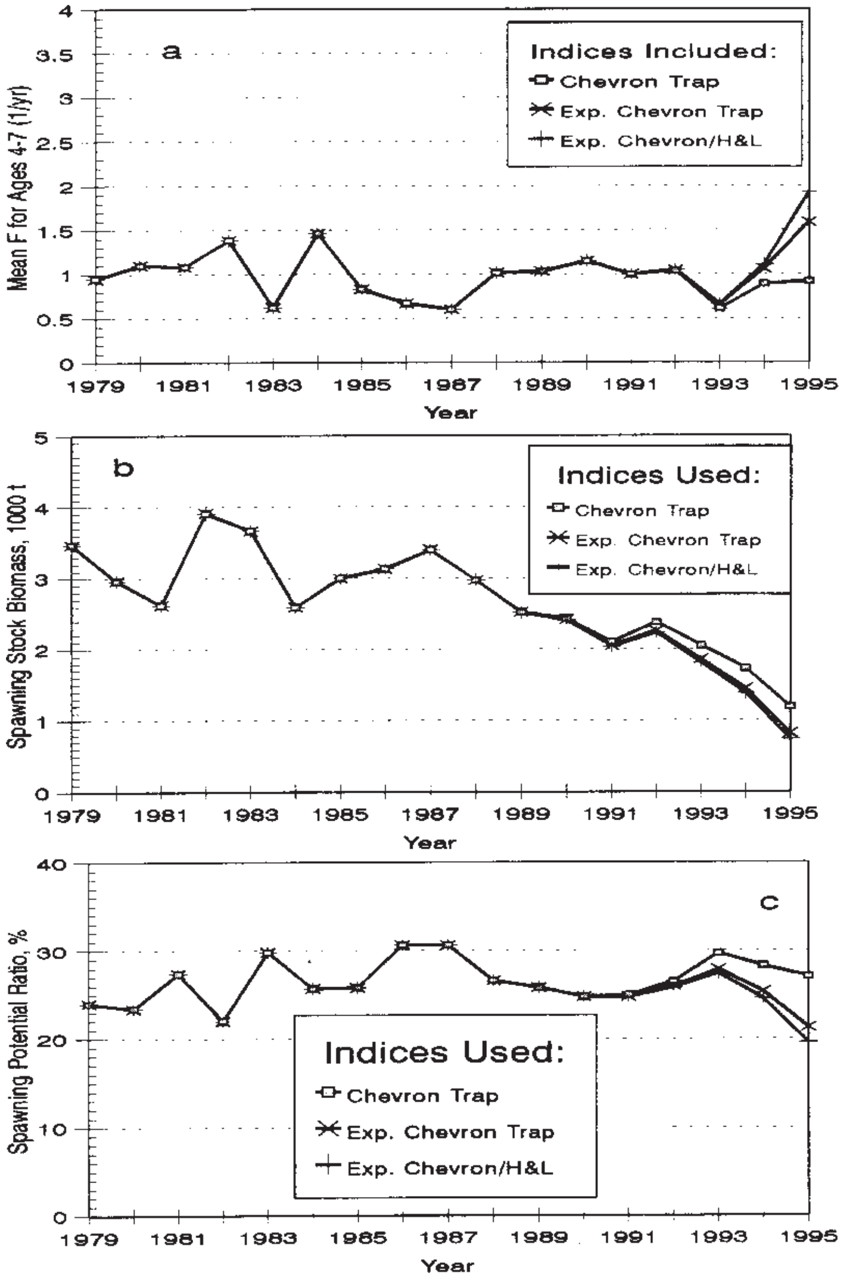


Figure 5. Comparison of full F, SSB, and SPR for black sea bass for MARMAP CPE indices: (a) unextended chevron trap, (b) extended chevron trap, and (c) hook and line and chevron trap using FADAPT method.

Table 3. Sample size (*n*), residual sum of squares (RSS), and mean squared error (MSE) from FADAPT runs for black sea bass with various combinations of calibration indices.

FADAPT Run	Sample size	RSS	MSE
(1) All gears	46	3.098	0.067
(2) Hook-and-line only	17	11.823	0.695
(3) All trap gears	29	2.072	0.071
(4) Chevron trap only	8	1.066	0.133
(5) Extend chevron trap only	16	5.076	0.317
(6) Hook-and-line and extended chevron trap	33	7.295	0.221

Weightings for MARMAP indices estimated by FADAPT: (1) 0.093 for hook-and-line, 0.303 for blackfish trap, 0.138 for Florida snapper trap, and 0.466 for chevron trap; (3) 0.327 for blackfish trap, 0.149 for Florida snapper trap, and 0.523 for chevron trap; and (6) 0.314 for hook-and-line and 0.686 for extended chevron trap.

may have been associated with the introduction of a minimum size limit that year.

Next, we compared FADAPT runs calibrated to the chevron trap indices (unexpanded, expanded, and combined with hook-and-line) (Fig. 5). The FADAPT run calibrated to the unexpanded chevron index suggested much lower full *F* and higher SPR in 1995 as compared to FADAPT runs calibrated with the expanded chevron trap index (with or without hook-and-line). Again, although full *F* and SPR showed no temporal trend, SSB declined since 1987.

Model error (MSE) from FADAPT was minimized by use of all gear separately (0.067), with similar model error using only trapping gear (0.071) (Table 3). Downweighting by FADAPT of the hook-and-line CPE was apparent when all gear CPEs were used (9.3%) and relatively high weight given to the recent, short chevron CPE (46.6% for all four gears and 52.3% for all three trap gears). Some of the spiky aspect of the Florida snapper trap CPE (Fig. 1) probably resulted in its smaller weighting when all four gear types were used (13.8%) and when all three trap types were used (14.9%). Using only hook-and-line CPE gave highest model error (0.695). As expected, model error was greater for the extended chevron trap CPE (0.317) than for the chevron trap CPE alone (0.133). As the chevron trap CPE was extended backward using the Florida snapper trap, any lack of agreement with the converged portion of the VPA would tend to increase MSE. Combining the extended chevron trap CPE (68.6%) and hook-and-line CPE (31.4%) in the calibration gave a value for MSE (0.221) less than the two indices used alone, although the residual sum of squares was intermediate.

Discussion

The blackfish trap CPE was inappropriate for extending the chevron trap CPE, because it had different selectivity properties from the other gear types. Differences in selectivity between the Florida snapper trap CPE and chevron trap CPE were less than or of the same magnitude as annual differences in selectivity by these gear types. Hence, the Florida snapper trap CPE provided a useful surrogate for extending the chevron trap CPE, because it had essentially the same selectivity properties.

The hook-and-line CPE suggested greater full *F*, and smaller SSB and SPR than indicated by the trap indices. Furthermore, the extended chevron trap CPE (1980-1995) demonstrated greater full *F*, and smaller SSB and SPR than indicated by the chevron trap CPE (1988-1995). The longer period for calibration permitted better historical comparison between the catch matrix (as represented by the converged portion of the VPA) and the CPE index. The trend in the extended chevron trap CPE agreed well with the trend noted in the hook-and-line CPE, and, in particular, confirming the sharp decline in CPE between 1987 and 1988 noted earlier in the MARMAP hook-and-line CPE index for black sea bass.

Although model error (mean squared error) was lowest for the FADAPT run using all indices separately (without extending chevron trap CPE), it was not necessarily the preferred run. Some of the reduction in model error was obtained by giving greater weighting to the recent, short duration chevron CPE which has fewer years to compare to the converged portion of the VPA. Hence, the true uncertainty associated with the chevron CPE index may have been underestimated. By using the extended chevron CPE with the hook-and-line CPE, uncertainty associated with the longer CPE index better represents the uncertainty inherent with the model fit and the information content of the MARMAP program was more fully used in analyzing the status of the black sea bass stock.

Although no major trends were noted for full *F* (and hence SPR) in all FADAPT VPA runs with different calibration indices, all FADAPT VPA runs suggested a significant decline in SSB since 1987.

Acknowledgments

The authors wish to thank the following for the fishery-dependent data obtained through their efforts for this analysis: Robert Dixon and Michael Burton (headboat data); Gerry Gray and Dave Van Voorhees (MRFSS data); and Laura Bishop, Guy Davenport, and Linda Hardy (commercial data). This is contribution 405 of the South Carolina Marine Resources Center.

References

Clay, D. 1990. TUNE: A series of fish stock assessment computer programs written in FORTRAN for microcomputers (MS DOS). Int. Comm. Conserv. Atl. Tunas, Coll. Vol. Sci. Pap. 32:443-460.

Collins, M.R. 1990. A comparison of three fish trap designs. Fish. Res. 9:325-332.

Collins, M.R., and G.R. Sedberry. 1991. Status of vermilion snapper and red porgy stocks off South Carolina. Trans. Am. Fish. Soc. 120:116-120.

Conover, W.J. 1971. Practical nonparametric statistics. Wiley, New York. 462 pp.

Gavaris, S. 1988. An adaptive framework for the estimation of population size. Can. Atl. Fish. Sci. Adv. Comm. (CAFSAC) Res. Doc. 88/29. 12 pp.

Harris, P.J., and J.C. McGovern. 1997. Changes in the life history of red porgy, *Pagrus pagrus*, from the southeastern United States, 1972-1994. Fish. Bull. U.S. 95:732-747.

Helser, T.E. 1996. Growth of silver hake within the U.S. continental shelf ecosystem of the northwest Atlantic Ocean. J. Fish Biol. 48:1059-1073.

Mace, P.M. 1994. Relationships between common biological reference points used as thresholds and targets of fisheries management strategies. Can. J. Fish. Aquat. Sci. 41:110-122.

Mace, P.M., and M.P. Sissenwine. 1993. How much spawning per recruit is enough? Can. Spec. Publ. Fish. Aquat. Sci. 120:101-118.

Mercer, L.P. 1989. Species profile: Life histories and environmental requirements of coastal fishes and invertebrates (South Atlantic): Black sea bass. U.S. Fish Wildl. Serv. Biol. Rep. 82(11.99); and U.S. Army Corps Eng., TR EL-82-4. 16 pp.

Pope, J.G., and J.G. Shepherd. 1985. A comparison of the performance of various methods for tuning VPAs using effort data. J. Cons. Cons. Int. Explor. Mer 42:129-151.

Restrepo, V.R. 1996. FADAPT Version 3.0, A guide. University of Miami, RSMAS, 4600 Rickenbacker Causeway, Miami, FL 33149. 21 pp.

Rosenberg, A., P. Mace, G. Thompson, G. Darcy, W. Clark, J. Collie, W. Gabriel, A. MacCall, R. Methot, J. Powers, V. Restrepo, T. Wainwright, L. Botsford, J. Hoenig, and K. Stokes. 1994. Scientific review of definitions of overfishing in U.S. fishery management plans. NOAA Tech. Memo. NMFS-F/SPO-17.

Shepherd, G. 1991. Meristic and morphometric variation in black sea bass north of Cape Hatteras. N. Am. J. Fish. Manage. 11:139-148.

Vaughan, D.S., M.R. Collins, and D.J. Schmidt. 1995. Population characteristics of the black sea bass *Centropristis striata* from the southeastern U.S. Bull. Mar. Sci. 56:250-267.

Vaughan, D.S., S.C. Turner and R. Conser. 1988. Preliminary investigation of the effects of weighting CPUE indices on estimates of fishing mortality from the CAL VPA program. Internat. Comm. Conserv. Atl. Tunas, Coll. Vol. Sci. Pap. 30(2):318-326.

Vaughan, D.S., M.R. Collins, B. Zhao, and H.S. Meister. 1996. Population characteristics of the black sea bass *Centropristis striata* from the U.S. southern Atlantic coast. Report to South Atlantic Fishery Management Council, Charleston, SC. 59 pp.

Wenner, C.A., W.A. Roumillat, and C.W. Waltz. 1986. Contributions to the life history of black sea bass, *Centropristis striata*, off the southeastern United States. Fish. Bull. U.S. 84:723-741.

Fishery Stock Assessment Models
Alaska Sea Grant College Program • AK-SG-98-01, 1998

An Index for Weighting Results in Catch-at-Age Models Based on Diagnostic Tests for Lack of Fit

T. Polacheck
CSIRO (Commonwealth Scientific and Industrial Research Organisation)
Division of Marine Research, Hobart, Tasmania

K.A. Haskard
CSIRO Mathematical and Information Sciences, Hobart, Tasmania

N. Klaer, A. Betlehem, and A. Preece
CSIRO Division of Marine Research, Hobart, Tasmania

Abstract

As part of the VPA assessments of southern bluefin tuna, agreed-on objective diagnostic tests designed to detect significant lack of fit are reported for each VPA analysis. However, in summarizing the current status of the stock and in the estimation of the probabilities of recovery, VPA results in which the diagnostic tests suggest significant lack of fit have been considered no differently than ones for which the test do not indicate any lack of fit. A procedure is developed whereby the results of these diagnostic tests can automatically and objectively be incorporated as an additional weight into the procedures used to combine results from different VPAs. The procedure takes into account that multiple tests are performed for a single tuning index and that possible correlation among the tests may exist. It yields a scale that accounts for both the relative lack of fit in a single index and the total number of tuning indices in which there is lack of fit. A survey was carried out using simulated assessment results to provide an indication of the performance of the procedure relative to human perceptions of what constitutes significant lack of fit. The results indicate that the two are generally consistent. The procedure is not meant as an alternative to developing and using likelihood approaches in assessments, but as an additional weighting term to be used in conjunction with both an output “likelihood” weight and an input prior weight to provide an overall relative weight when combining alternative assessment results.

Introduction

Management decisions that regulate catch levels for fish stocks are often based on scientific assessment of the current state of the stock relative to historical levels and projections of the future stock sizes under different catch scenarios. The estimations of both current and future stock sizes are based on population models in which biological data from the stock and catch data from the fisheries are used to estimate the parameters of the models. Large uncertainties exist in the current knowledge and understanding of the population dynamics of fish stocks. In addition, the sampling error and biases associated with the basic data used for estimating the catch and biological processes are frequently unknown, but potentially large. Consequently, there is a correspondingly large uncertainty about both the most appropriate model structure for assessing specific stocks and the estimates of the input parameters to use in these models. It is not uncommon for different model structures or parameterizations within the full range of uncertainties to yield substantially different results with respect to stock status and management implications. As such, there is increasing recognition that evaluation of the consequences of model and input uncertainties needs to be incorporated into the scientific stock assessment process and that results and advice should be presented in a probabilistic framework.

Bayesian approaches provide one framework for developing probabilistic assessments and the approach is being adopted in a number of stock assessments (Hilborn and Walters 1994). Bayesian approaches require that the assessment model be framed as a maximum likelihood problem and that prior weights or probabilities (i.e., degree of belief) be assigned to alternative inputs into the model. The Bayesian analyses use observed data and estimates of the likelihood of the different parameters given the observed data to provide updated probabilities for the parameter values of the model. However, there are limitations to Bayesian approaches including specification of the appropriate distributional form and the variance-covariance structure for the data inputs into the assessment. Also, a Bayesian approach requires that all formulations of specification models or hypotheses can be specified in a nested structure and that all the different nested models use the same input data. The process of updating the probabilities assigned to the different input parameters is based completely on the likelihood of specific parameter values given the data inputs. There is no scope within the formal Bayesian framework for testing and taking into account whether there is any fundamental lack of fit to the basic model (e.g., nonlinearity in assumed linear relationship; temporal trends in residuals, etc.). The only measure of fit is the estimated overall likelihood of the data.

Since 1992, all VPA assessments of the southern bluefin tuna (SBT) stock have been based on the ADAPT (Gavaris 1988) framework (Ishizuka and Tsuji 1990; Polacheck et al. 1992, 1997; Tsuji and Takeuchi 1997).

These integrated approaches allow for a wide variety of different sources of information to be used in the tuning process based on a user-defined objective function. By putting the VPA tuning process into a statistical framework, these integrated approaches also allow for examination of the residuals for lack of fit (e.g., nonlinearity, time trends, etc.). Differential weighting of the various sources of information in the objective function can then be used to evaluate consistency, sensitivities, and sources of lack of fit, and to identify areas of uncertainty in the overall assessment results.

Within the current SBT stock assessment, there are multiple sources of identified uncertainty. These include different interpretations of the catch rate data, the natural mortality rates, the catch-at-age matrices, the estimation of the plus group, and specification of the population-fishery dynamic model underlying the VPA (e.g., selectivity assumptions). For each of these sources of uncertainty, a range of different models and/or parameterizations have been developed to represent different plausible hypotheses for the observed data and underlying stock and fishery dynamics.

The process of incorporating uncertainties has been an evolving one. Initially the number of dimensions and the range within each dimension considered was small. Under this situation, all combinations of results could be individually evaluated, and simple (e.g., visual) comparisons formed the basis for synthesizing the various results into an overall assessment of the status of the stock and providing management advice. As the process has evolved to incorporate a more realistic and comprehensive range of the actual uncertainties, the number of individual VPA assessments with corresponding stochastic projections increased to over several hundred. More automated and objective processes were needed for integrating the results.

The process evolved to providing a weighted average over an agreed-upon range of assessments. Different national delegations of scientists involved in the assessment process independently develop sets of weights and the results across the different sets of weights are compared. The process of setting weights has elements of an “informal” Bayesian process in that each group of scientists is meant to decide on the relative plausibility of the various hypotheses and specific parameter values, based on prior knowledge and beliefs updated by review and examination of the VPA results presented up to the point when the weights are specified. Such an informal process can, and has, led to very divergent sets of weights with different implications for stock status and management advice. The process is further confounded in that the weights are only assigned to each of the uncertainty dimensions independently (i.e., only the marginal distributions are considered). When constructing the overall average weighted results, each dimension is treated independently and the weights are simply multiplied together. There is no consideration of the joint probabilities. This is particularly problematic because once the weights have been assigned there is no process for further updating based on the “plausibility" or "likelihood" of the actual results. As such, this process ignores the power of these integrated VPA approaches both to provide some overall measure of goodness of fit to the input data and to examine the residuals for systematic or unacceptable lack of fit.

There has been agreement among all involved in the SBT assessment process that an objective approach is needed that not only takes into account prior information but also allows for updating of the weights based on the VPA outputs. Ideally such an approach would update the input prior weights based on some measure of goodness of fit related to the objective function (e.g., the likelihood estimates of the data given the fitted parameters) and also take into account any substantial lack of fit. The current stage of development of the SBT assessment does not provide an obvious measure for goodness of fit as a result of the non-nested structure of the alternative models and because not all of the data are used in the different models. Moreover, for some of the major sources of uncertainty (e.g., natural mortality rates, interpretation of CPUE), the catch-at-age data and fitted VPA model would not be expected to be informative.

The SBT stock assessments are reviewed by the Scientific Committee of the Commission for the Conservation of Southern Bluefin Tuna (CCSBT). At a scientific workshop held by the CCSBT to examine CPUE and VPA models (CCSBT 1996), the participants agreed on a set of diagnostic tests that should be performed and reported with each individual VPA. These diagnostic tests are designed to detect significant temporal trends in the residuals and non-linearity between observed and predicted terms in the objective function. Probability levels for each test can be generated using statistical methods. While results of these diagnostic tests are reported, their results have not been formally incorporated into the assessment process. Currently, VPA results in which the diagnostic tests suggest significant model lack of fit are considered no differently than ones for which the test do not indicate any lack of fit (e.g., Klaer et al. 1996). In general, as long as the number of alternative models and parameterizations being considered in an assessment is relatively small, questions of model lack of fit have been dealt with by using visual examinations of residuals along with the results from diagnostic tests to disqualify those results that clearly do not fit the data. However, such an approach is not feasible when the number of alternative models and parameterizations increases. Also, such an approach only results in a zero or one weighting, which makes it difficult to deal with non-clear-cut situations where there is apparently lack of fit to some portions of the input data. As such, an approach is needed that can handle assessments with multiple tuning indices in which large numbers of alternative models and parameterizations are being considered.

The purpose of the present paper is to develop a procedure whereby the results of such diagnostic tests can automatically and objectively be incorporated into weighting procedures used to combine results from different VPAs. The procedure has been developed in the context of the SBT assessment within the CCSBT framework, but could be used in other situations. It should be emphasized that the approach developed here is not meant as a replacement for the need to develop a "likelihood" weighting approach that would update the prior weights based on the overall goodness of fit to the data. Instead, we would see the approach being used in conjunction with a "likelihood" measure of fit in order to provide a method for objectively downweighting results when diagnostic tests indicate substantial and significant model misspecification in a manner analogous to a penalized likelihood approach.

Diagnostic Tests

Within the CCSBT framework, initially five diagnostic tests were defined and the results were calculated for each time series used in tuning the VPA (CCSBT 1996). However, these have subsequently reduced to the following three tests:

1. A runs test for the time series of residuals (i.e., a test based on the number of runs of positive and negative residuals for the tuning index versus time).
2. A hypothesis test of whether the slope of the regression through the origin of the predicted versus observed values for a tuning index differs from 1.0.
3. A runs test based on the residuals from the regression through the origin of the predicted versus observed values assuming a slope of 1 (i.e., a test based on the number of runs of positive and negative residuals for the tuning index versus the value of the predicted index).

As initially defined in 1996, the runs tests were conducted based on the normal approximation to the expected number of runs in a series with n_1 positive residuals and n_2 negative residuals (see Draper and Smith 1966). However, in the most recent assessment, the Durbin-Watson test for serial correlation has been substituted for the runs test in 1. This test provides an indication of whether there is an unusually low or high number of runs. The regression diagnostic test (test 2) was performed using standard least squares regression methods assuming normal and additive errors. The first test is meant to detect if there are large temporal trends in the residuals, and the other two tests were meant to detect non-zero intercepts or nonlinear relationships between the VPA estimates and the observed tuning indices. These diagnostic tests have been selected by the CCSBT scientific committee and are used here to illustrate the process we have developed for combining multiple diagnostic tests applied to multiple tuning indices. Alternative suites of tests may be more appropriate in other situations and assessment forums.

Development of a Lack of Fit Weighting Based on Diagnostic Test Results

When the probability levels for one or more of the diagnostic tests are low, this suggests that there is substantial lack of fit of the model to the data. Thus, the model as structured in combination with the data inputs is not fully consistent. As such, the results and predictions from such a model have a degree of implausibility. When some VPA results indicate substantial lack of fit, it would seem appropriate to use the probability levels from these tests to derive a relative weighting scale. However, there is no well developed statistical theory for doing this, although in more straightforward statistical analyses a low probability level (zero weight) would usually lead to rejection of the model. It should also be noted that in the application of these diagnostic tests to the residuals from the VPA estimates, the probability levels from these tests are not fully accurate measures of the true probability. This is because of the multiple number of tests being conducted and uncertainty about the actual form of the probability distribution for these tests, particularly within the context in which they are being used. This suggests that, in developing a weighting scale based on these tests, the scale should be conservative; i.e., only downweight VPA results in those cases where the diagnostic test indicates a very significant lack of fit.

Given the above, the following procedure has been proposed as a basis for assigning relative weights to a particular VPA result based on the probability levels from the diagnostic tests:

1. For each tuning index, determine the number of the three diagnostic tests for which the probability level is less than 0.05. Then assign a value to that index according to the number of tests with a probability level (p) less than 0.05:
    a. If the probability levels for all tests are above 0.05, then this index gets a value of 1.
    b. If the probability level for only one test is less than 0.05 then this index gets assigned the value if it is less than 0.017; otherwise it is assigned the value 1.
    c. If the probability levels for only two tests are less than 0.05, then this index gets assigned the value of the lowest probability value for the two tests if both are less than 0.025; otherwise it is assigned the value 1.
    d. If the probability level for all three tests are less than 0.05, then this index gets assigned the lowest probability value for the three tests.
2. If the value assigned to an index is less than 0.005, then assign it the value of 0.005.

3. The geometric mean of the values assigned to all indices in steps 1 and 2 is assigned as the relative weight for this VPA run.

This procedure attempts to take into account both the fact that multiple diagnostic tests are being performed on a single tuning index and that the tests among the different tuning indices cannot be considered completely independent. The latter is because of the correlation among the indices, the catch-at-age data, and the VPA parameter estimates. The rationale for deciding when to downweight an index in step 1 is developed in detail in the Appendix based on considerations of the multiple number of tests and the possible correlation among them. Setting a value of 0.005 as the minimum that can be assigned to a single index means that there is a limit to which the VPA results will be downweighted as the result of a lack of fit for a single index. Thus, if there are five indices, the minimum weight when only one index was found to have significant lack of fit would be 0.35, which is the geometric mean of 0.005 and 1,1,1,1.

Taking the geometric mean across the different tuning indices means that the amount of downweighting that occurs depends both on how severe the lack of fit for any of the indices is and how many indices had substantial lack of fit (i.e., if there were only five indices, the minimum weight due to lack of fit in 1, 2, 3, 4, or 5 indices would be 0.35, 0.12, 0.042, 0.014, and 0.005 respectively). The use of the geometric mean gives more weight to having multiple indices with lack of fit than to having severe lack of fit in a single index. This was seen as an important property since the former suggests a general overall lack of fit (see Discussion).

The above procedure has an element of arbitrariness. Other procedures or values for the probability levels and minimum values could be used (see Discussion). However, the procedure does provide an objective approach for assigning relative weights based on the results of the diagnostic tests. It takes into account the multiple testing and possible correlation among the tests. It also yields a scale that takes into account both the relative lack of fit in a single index and the number of tuning indices in which there is lack of fit. Finally, it provides a weighting scale in which there has to be severe lack of fit (as measured by the diagnostic tests) in multiple indices for a VPA to assign a negligible weight. This latter seems a desirable property since the particular model structure and parameterization were presumably selected based on a judgment that they were a plausible representative for the underlying stock dynamics.

Application of Lack of Fit Weightings to Simulated Assessment Results

A survey was designed to compare the weights assigned using the objective procedure described above with the subjective evaluation of lack of fit provided by fisheries scientists. Each survey consisted of a presentation of simulated results for 12 assessments that were fitted to three tun-

Table 1. Diagnostic test values for simulated assessment results used in the survey.

	Runs test by time			Runs test by regression			Significance of slope other than 1		
S.A.	Index 1	Index 2	Index 3	Index 1	Index 2	Index 3	Index 1	Index 2	Index 3
1[a]	0.99	0.23	0.98	**0.00**	0.23	**0.00**	**0.00**	0.82	**0.00**
2	**0.00**	**0.01**	0.19	**0.00**	**0.01**	0.19	**0.00**	**0.00**	0.29
3[b]	**0.00**	**0.02**	**0.00**	**0.00**	**0.02**	**0.00**	**0.00**	0.75	0.79
4[c]	0.99	0.23	**0.00**	0.99	0.23	**0.00**	0.50	0.82	**0.00**
5	**0.02**	0.23	0.19	**0.02**	0.23	0.19	0.35	0.82	0.29
6[c]	0.99	**0.01**	0.19	0.99	**0.01**	0.19	0.50	**0.00**	0.29
7	0.99	0.23	0.19	0.99	0.23	0.19	0.50	0.82	0.29
8[b]	**0.02**	**0.01**	**0.00**	**0.02**	**0.01**	**0.00**	0.35	**0.00**	0.79
9	**0.00**	**0.02**	**0.00**	**0.00**	**0.02**	**0.00**	**0.00**	0.75	**0.00**
10[a]	0.99	0.92	0.19	**0.00**	**0.01**	0.19	**0.00**	**0.00**	0.29
11	0.29	0.23	0.19	**0.02**	0.23	0.19	0.35	0.82	0.29
12	0.99	0.23	0.98	0.99	0.23	**0.00**	0.50	0.82	**0.00**

Results that failed the diagnostic test at the 0.05 level are displayed in bold. The letters a, b, and c indicate simulated assessments that are essentially duplicates.

S.A. = simulated assessment.

ing indices. For each tuning index, a plot of observed versus predicted and residuals versus time were provided. In all cases, the coefficient of variation of the residuals was fixed at 0.2, and residuals were chosen at random using a normal distribution and the mean of the simulated observed values. This means that the sum of squares values (or likelihood value given the data) for the fit for each assessment were exactly the same. Trends in some residuals were produced by the introduction of a nonlinear (V-shaped) relationship or a non-zero intercept for observed versus predicted values. Individuals completing the survey were requested to assign a value between 0 and 10 to each simulated assessment according to how well they considered the tuning indices fitted the VPA results. They were not supplied with diagnostic test results, or any further information other than the residual plots. Surveys were distributed during this Lowell Wakefield symposium (International Symposium on Fishery Stock Assessment Models for the 21st Century, Oct. 8-11, 1997, Anchorage, Alaska). Completed surveys were returned during and after the symposium.

The results for each of the diagnostic tests for the simulated assessments used in the survey are shown in Table 1. The diagnostic test results for simulated assessments 1, 3, and 4 were duplicated in the survey by assessments 10, 8, and 6 respectively, except that the order the indices were presented was changed and there were small differences in the diagnostic test results due to different randomizations. Table 1 shows that the rates of failure of the diagnostic tests at the 0.05 level for the duplicate

Table 2. Standardized survey scores assigned to each of the simulated sets of assessment results by people who participated in the survey, that were judged to be consistent.

	Standardized survey scores									
S.A.	1	2	3	4	5	6	7	8	9	10
1	0.89	0.38	0.63	0.33	0.60	0.20	0.75	0.67	0.50	0.33
2	0.33	0.25	0.38	0.17	0.40	0.20	0.50	0.17	0.50	0.33
3	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
4	0.44	0.50	0.63	0.50	0.40	0.60	0.50	0.67	0.75	0.67
5	0.56	0.63	0.63	0.67	0.40	0.60	0.25	0.83	0.75	0.83
6	0.44	0.50	0.63	0.50	0.40	0.60	0.25	0.67	0.75	0.67
7	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
8	0.22	0.13	0.00	0.00	0.00	0.00	0.00	0.17	0.00	0.17
9	0.11	0.25	0.00	0.00	0.20	0.00	0.00	0.00	0.00	0.00
10	0.67	0.63	0.63	0.17	0.80	0.20	0.75	0.67	0.75	0.33
11	0.67	0.75	0.88	0.83	0.80	0.60	0.75	0.83	0.75	0.83
12	0.78	0.88	0.88	0.50	0.60	0.80	1.00	0.67	0.75	0.67

S.A. = simulated assessment.

assessments are the same. Duplication was included to allow judgment of whether individuals completing the survey were doing so in a consistent manner.

Simulated assessments 3, 8, and 9 fail a large number of diagnostic tests, and were included to represent "bad" fits, whereas assessment 7 passed all diagnostic tests and represents a "good" fit. Survey respondents assigned values in the range of 0 to 10 to each assessment, but the range of individual survey scores varied from 0-9 to 3-7. To eliminate these range differences and to convert the scores to weightings, they were standardized to a range of 0 to 1.0.

A survey was judged to be inconsistently completed if the summed differences of standardized results for the duplicate assessments was greater than 0.5. On this basis, 3 of the 13 completed surveys were discarded, leaving 10 for further analysis. The summed difference values for the surveys discarded were 0.75, 0.75, and 1.0, and for those retained were 0.44, 0.38, 0.00, 0.17, 0.20, 0.00, 0.25, 0.17, 0.25, and 0.17. Standardized survey weightings for consistently completed surveys are shown in Table 2. It should be noted that 10 surveys is a small sample for making general inferences. We would have preferred more, but had difficulty promoting a higher level of participation.

The objective procedure produces relative weights that are not evenly spread; e.g., there is a gap immediately below 1 where no weight values can fall, and small weights are more likely than larger weights. However, in the survey it would be reasonable to expect that people would assign their weightings based on a continuous linear scale from 0 to 1.0 for the standardized results. In order to account for this, a simple log transformation and re-scaling of the calculated weightings $\{[\log(x) + 2.1175] / 2.1175\}$ compared to survey weights was performed (2.1175 was the value that produced the best fit between the two weightings) (Fig. 1).

After this transformation, there is a good correlation between the values produced by the objective procedure and those resulting from the survey. This suggests that the objective procedure performs well in comparison with subjective human judgments when consideration is given to the nonlinear scale of the objective procedure. The objective procedure was designed to be conservative and only assign weights less than 1 when there was significant indication of lack of fit in at least one index. The procedure appears to provide a reasonable probabilistic approach for assigning weights to VPA results that corresponds with subjective weightings assigned "by eye" on examination of assessment output.

Discussion

It is important to emphasize that the procedure developed here is not meant as an alternative to developing and using a "likelihood" approach in assessments. The index from the procedure was envisioned as an additional weighting term to be used in conjunction with both an output "likelihood" weight and an input "prior" weight to provide an overall relative weight to assign to alternative models and parameterizations.

We think that there is a real need for such a lack of fit measure when the number of alternative models and parameterizations becomes large, as it almost inevitably will when one tries to capture the full range of uncertainty contained in most fishery stock assessments. With the increase in the number of alternatives, it becomes impossible to individually determine which combinations of alternative options are inconsistent with the basic input data. From our experience of using this index and examination of diagnostics in the SBT context, we have found that for many single options (e.g., a particular natural mortality rate vector or a specific CPUE interpretation) that they do not fall into the simple categorizations of being always consistent or inconsistent with the input data. Instead, it is the combinations of specific alternative options for two or more areas of uncertainty that can result in substantial and unacceptable lack of fit. As such, some form of automated procedure becomes essential.

We also realize that any such procedure cannot replace logic and common sense in the initial selection of the alternative models and parameterizations that are to be used in an assessment. Problems with lack of fit can always be circumvented by increasing the number of parameters and by detailed modeling of the residuals. Clearly, one needs to avoid overparameterization. Likelihood ratio tests are one tool that can assist with this. We see the procedure developed here, or any similar such procedure,

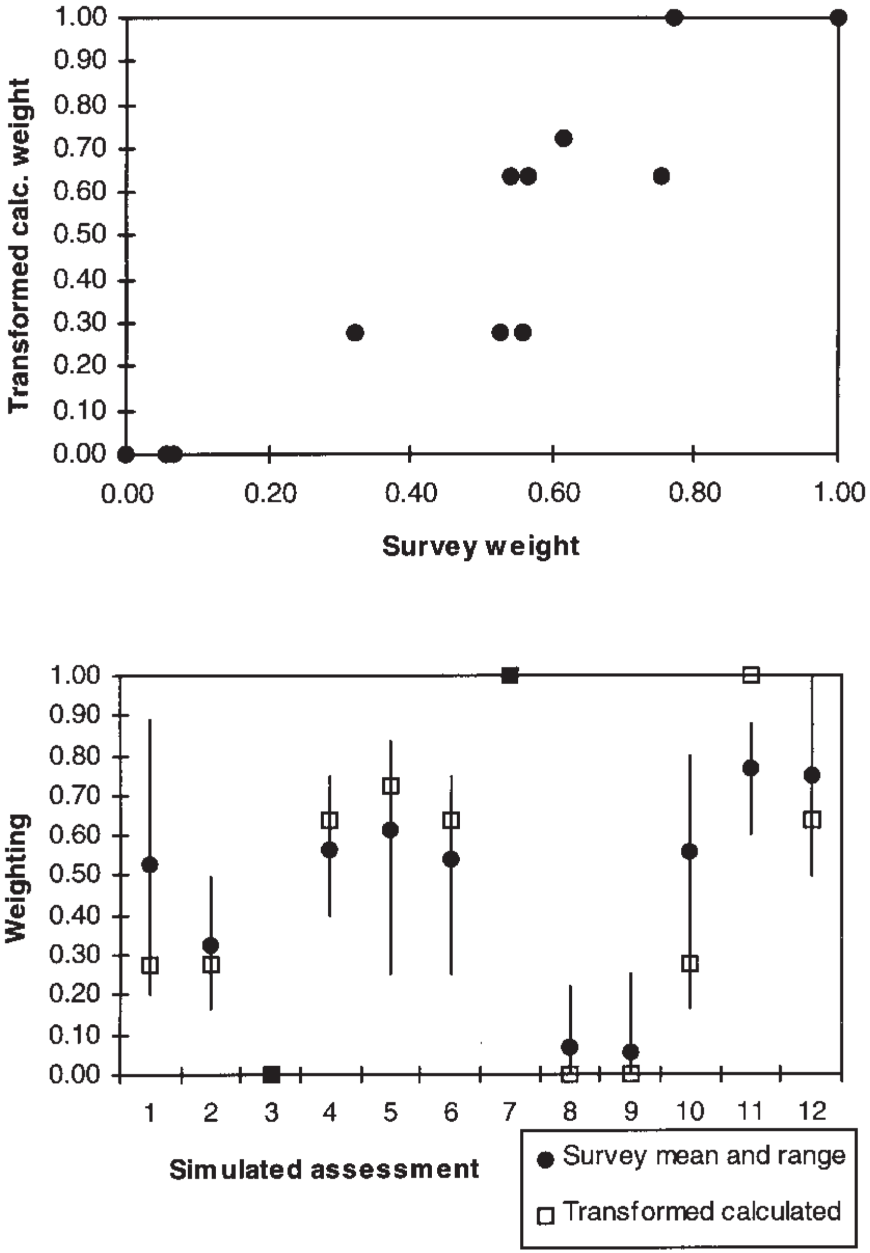


Figure 1. Comparison of the mean and range of weights assigned to simulated assessment results from the survey and those calculated using the objective procedure after log-transformation.

as an additional tool to assist in the model selection and model weighting process. It is not meant as a replacement for existing ones.

The exact procedure developed here is only one out of a large number of possible alternative approaches that could be constructed for assigning weight based on the results of a set of diagnostic tests. We think there is scope for improving the procedure. We would hope that there would be further developments and comparative evaluation of alternative approaches. Any approach has to deal with two basic issues: (1) combining results from several diagnostic tests applied to the same tuning index, and (2) combining the results when two or more tuning indices are used in the assessment.

With respect to (1), a well-developed statistical basis exists for combining multiple tests. The major problem in directly applying this is having to determine the covariance among the different tests. We have attempted to uses this statistical basis in step 1 of our procedure (see Appendix) in the absence of knowledge about the covariance among the tests and given a predefined significance level. Because the covariances are unknown, there is no unique solution. The choice of how to combine the results from several tests will depend on how much belief one has on the degree of independence and the relative amount of conservativeness desired with respect to the tests being correlated. We have taken a fairly conservative approach. There would be scope for improving how this combining across tests is done, particularly if additional research determines the covariances among the different diagnostic tests being used.

The question of how best to combine the results for a set of different indices is less straightforward. There is the issue of how to weight the different indices, particularly if they are of different length (with respect to the years to which they apply) and have different degrees of reliability (i.e., variance and bias). We have chosen to give all of the indices equal weight, mainly because the main tuning indices in the SBT assessment are of equal length and would be expected to have approximately equal variances. If the indices were of different lengths or had substantially different variances, equally weighting may not be appropriate (although this should be accounted for, at least to some extent by the *p*-values generated by the diagnostic tests).

A more difficult question is how to deal with the fact that the diagnostic tests can indicate lack of fit for varying number of indices. If one considered that the diagnostic tests for each of the indices provide an independent measure of the lack fit for the overall assessment, then it would be appropriate to treat the *p*-values as separate overall probabilities. They could then be combined as the resultant product and the resulting *p*-value could be used as a rejection criteria after taking into account the multiple testing. This would tend to result in very severe downweighting when lack of fit was detected even in a single index. We did not think this was appropriate, particularly when the different indices correspond to different age groups or time periods. In this situation, lack of fit to several indices would indicate a general lack of fit in the overall assessment, while lack of fit in only one index would suggest a more restricted problem, perhaps related to that specific index. We felt that this latter case should not be downweighted as severely. This was the reason for using the geometric mean when combining the p-values across different indices. While this is somewhat ad hoc and scope exists for further development, the use of the geometric mean provided the desired general behavior (e.g., giving a substantially lower weight in the situation when the p-value for two indices is 0.05 compare to the situation where one index had a value of 0.025).

The procedure developed here results in a non-continuous weighting scheme. Thus, it is impossible to get a value between one and some threshold that depends upon the number of tuning indices. As pointed out by one reviewer, such thresholds can cause problems for those dependent upon the scientific advice for decision making over a number of years as the potential exists for the scientific advice to change radically with the addition of one more year's data. We do not see any simple way around this problem in dealing with the question of lack of fit. As long as the basic model and parameterization are not a priori implausible, a basic lack of fit can only become manifest when the time series becomes sufficiently long. Thus, no matter what the criteria used for judging lack of fit, there will be a minimum number of observations required in which lack of fit is manifested before the overall fit could be judged inadequate. By using a procedure like the one developed here, it is possible to consider an uncertainty space spanning a wide range of combinations of models and parameterizations. This should yield a fair amount of robustness in the overall general results even when one portion of the space is judged by the procedure to have significant lack of fit as the result of an additional point in the time series.

References

CCSBT. 1996. Report of the CCSBT Workshop on VPA and CPUE Modelling, 4-15 March 1996. Hobart, Tasmania. (Commission for the Conservation of Southern Bluefin Tuna.)

Draper, N., and H. Smith. 1966. Applied regression analysis. John Wiley, New York.

Gavaris, S. 1988. An adaptive framework for the estimation of population size. Can. Atl. Fish. Sci. Adv. Comm. (CAFSAC) Res. Doc. 88/29.

Hilborn, R. and C.J. Walters. 1994. Quantitative fisheries stock assessment. Chapman and Hall.

Ishizuka, Y., and S. Tsuji. 1990. Assessment of the southern bluefin tuna stock. 9th Trilateral Scientific Meeting on SBT, Hobart, Australia. Document SBFWS/90/2.

Klaer, N., T. Polacheck, K. Sainsbury, and A. Preece. 1996. Southern bluefin tuna stock and recruitment projections. Second CCSBT Scientific Committee Meeting 1996, Hobart, Australia. Document CCSBT/SC/96/36

Polacheck, T., A. Betlehem, and A. Preece. 1997. Documentation of the virtual population analysis methods and model inputs used for estimation of current and historical stock sizes of southern bluefin tuna population. Third CCSBT Scientific Committee Meeting 1997, Canberra, Australia. Document CCSBT/SC/97/24.

Polacheck, T., K. Sainsbury, and N. Klaer. 1992. Assessment of the status of the southern bluefin tuna stock using virtual population analysis, 1992. 12th Trilateral Scientific Meeting on SBT, Shimizu, Japan. Document SBFWS/92/7.

Tsuji, S., and Y. Takeuchi. 1997. Stock assessment and future projection of southern bluefin tuna, 1997. Third CCSBT Scientific Committee Meeting 1997, Canberra, Australia. Document CCSBT/SC/97/17.

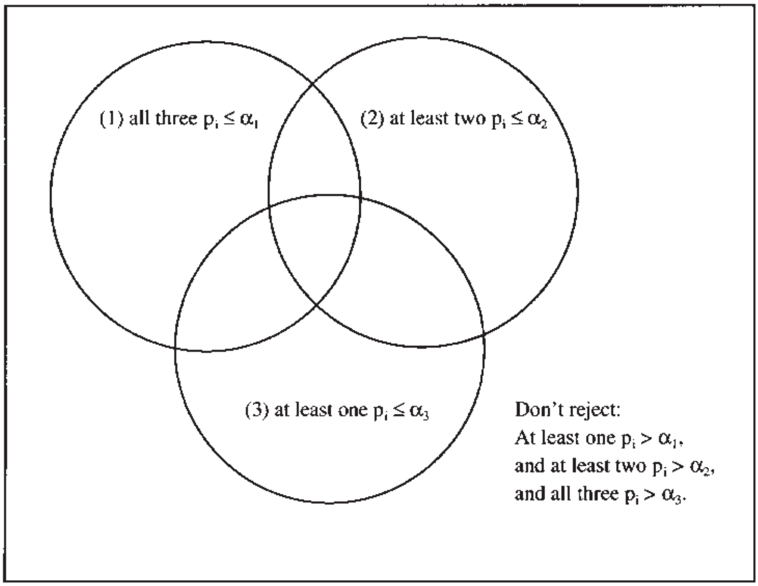
Appendix. A Composite Rejection Rule Based on Three Goodness of Fit Tests

Given three different (but not necessarily independent) goodness of fit tests for a given model with a given set of data, each producing a *p*-value, we wish to formulate a procedure to accept or reject the model using all three *p*-values.

Suppose the three *p*-values are p_1, p_2, and p_3. Choose α_1, α_2, and α_3 with $0 < \alpha_3 < \alpha_2 < \alpha_1 < 1$, such that we reject H_0: the model fits, under any of the following conditions:

1. All three $p_i \le \alpha_1$, i.e., $\max(p_i) \le \alpha_1$.
2. At least two $p_i \le \alpha_2$, i.e., $\mathrm{median}(p_i) \le \alpha_2$.
3. At least one $p_i \le \alpha_3$, i.e., $\min(p_i) \le \alpha_3$.

These three conditions are not mutually exclusive, nor, in general, is any one totally included in any other. In any given situation, in general we can be in *any* region of the following Venn diagram.



This composite rejection rule can be re-expressed as the three mutually exclusive conditions

a. At least one $p_i \le \alpha_3$.

b. $\alpha_3 <$ all three $p_i \le \alpha_1$.

c. $\alpha_3 <$ two of the $p_i \le \alpha_2$ and one $p_i > \alpha_1$.

We propose the following values for the α_i, using the nominal level of $\alpha = 0.05$ for examples throughout.

Let $\alpha_1 = \alpha = 1 - (1 - \alpha) = 0.05$
$\alpha_2 = 1 - (1 - \alpha)^{1/2} = 0.02532$
$\alpha_3 = 1 - (1 - \alpha)^{1/3} = 0.01695$

The value 0.05 for α_1 means we reject H_0 if all three tests are significant at the conventional level. The lower value 0.017 for α_3 means we reject H_0 if any one test is particularly bad (highly significant), even if the others are good (not significant).

The intermediate value 0.025 for α_2 allows one large (non-significant) p-value, but rejects H_0 if two are significant at a more stringent level than the conventional 0.05.

The value for α_3 was calculated such that if the three tests are mutually independent, then the probability of rejecting a true null hypothesis by rule (3) or (a), at least one $p_i \le \alpha_3$, namely $1 - \Pr(\text{all } p_i > \alpha_3) = 1 - (1 - \alpha_3)^3$, is set equal to α, e.g., the conventional level 0.05. In the case of mutually independent tests, there are additional positive probabilities of conditions (b) and (c), namely $(\alpha_1 - \alpha_3)^3$ and $3(1 - \alpha_1)(\alpha_2 - \alpha_3)^2$ respectively. These are both small and bring the total rejection probability to 0.050236, or, in general, after some algebra, $1 + 3(1 - \alpha)^2 - (1 - \alpha)^3 + 3(1 - \alpha)^{7/3} - 6(1 - \alpha)^{11/6}$. Thus these values of the α_i work well for mutually independent p_i, giving an overall rejection probability very close to 0.05, when $\alpha = 0.05$.

At the other extreme, all three tests would be equivalent, giving identical p-values, and we effectively have only one test. Then the overall probability of rejection of a true null hypothesis would be exactly $\alpha_1 = 0.05$. If conditions (2) or (3) are true, then so is condition (1). Alternatively, conditions (a) or (b) can apply, but not (c).

Thus α_1 was set to ensure the "correct" significance level with three identical goodness of fit tests; and α_3 ensures the correct significance level under condition (3) or (a) with three mutually independent goodness of fit tests. The value for α_2 was chosen to ensure a correct significance level under a specific intermediate scenario. Consider the situation where two of the goodness of fit tests are identical and the third independent. Then we effectively have only two independent tests. We chose α_2 such that under this scenario, the probability of at least one of those independent tests having $p_i \le \alpha$ is 0.05, or α in general. This does not directly relate to condition (2), but fits in with our values for α_1 and α_3 as follows:

α_1 identified by 0.05 or $\alpha = \Pr(\text{at least one } p_i \le \alpha_1$, when only have one test).

α_2 identified by 0.05 or $\alpha = \Pr(\text{at least one } p_i \le \alpha_2$, when have two independent tests).

α_3 identified by 0.05 or $\alpha = \Pr(\text{at least one } p_i \le \alpha_3$, when have three independent tests).

$$\begin{aligned}&\Pr(\text{at least one } p_i \le \alpha_2\text{, from two independent tests})\\ &\quad = 1 - \Pr(\text{both } p_i > \alpha_2)\\ &\quad = 1 - (1 - \alpha_2)^2\end{aligned}$$

Equating this to α gives $\alpha_2 = 1 - (1 - \alpha)^{1/2}$.

The total rejection probability, under the null hypothesis, for this scenario of two equivalent tests and the third test independent, is found by summing the probabilities of mutually exclusive conditions (a), (b), and (c). These are, after some algebra:

$$\begin{aligned}\Pr[\text{condition (a)}] &= 1 - (1 - \alpha_3)^2 = 1 - (1 - \alpha)^{2/3}\\ \Pr[\text{condition (b)}] &= (\alpha_1 - \alpha_3)^2 = [-(1 - \alpha) + (1 - \alpha)^{1/3}]^2\\ &= (1 - \alpha)^2 - 2(1 - \alpha)^{4/3} + (1 - \alpha)^{2/3}\\ \Pr[\text{condition (c)}] &= (1 - \alpha_1)(\alpha_2 - \alpha_3) = (1 - \alpha)^{4/3} - (1 - \alpha)^{3/2}\end{aligned}$$

and the total rejection probability is the sum of these, which comes to

$$1 + (1 - \alpha)^2 - (1 - \alpha)^{4/3} - (1 - \alpha)^{2/3}$$

and for $\alpha = 0.05$, this is 0.0427, again reasonably close to the ideal 0.05.

Bounds on Overall Probability of Rejection in the General Case

We can obtain bounds on the overall probability of rejection of a true null hypothesis in the general case, using the following elementary probability rules. For any events A, B, and C,

$0 \le \Pr(\text{A and B and C}) \le \min[\Pr(\text{A}),\Pr(\text{B}),\Pr(\text{C})]$;

$\max[\Pr(\text{A}),\Pr(\text{B}),\Pr(\text{C})] \le \Pr(\text{A or B or C}) \le \Pr(\text{A}) + \Pr(\text{B}) + \Pr(\text{C})$;

and for mutually exclusive events A, B, and C, $\Pr(\text{A or B or C}) = \Pr(\text{A}) + \Pr(\text{B}) + \Pr(\text{C})$.

Using these rules, and assuming $\Pr(p_i \le \alpha_j) = \alpha_j$, i.e., that the p-values from the goodness of fit tests are meaningful, the three mutually exclusive conditions have bounds on probabilities:

a. Condition (3), at least one $p_i \le \alpha_3$, has probability between α_3 and $3\alpha_3$.

b. Condition (1) but not (3), $\alpha_3 <$ all three $p_i \le \alpha_i$, has probability between 0 and $\alpha_1 - \alpha_3$.

c. Condition (2) but neither (3) nor (1), $\alpha_3 <$ two $p_i \le \alpha_2$ and one $p_i > \alpha_1$, has probability between 0 and $3 \times \min(\alpha_2 - \alpha_3, 1 - \alpha_1) = 3(\alpha_2 - \alpha_3)$, at least for any $\alpha_1 < 0.5$.

Summing these gives overall bounds of

$$\begin{aligned}\alpha_3 \le \text{overall probability of rejection} &\le 3\alpha_3 + \alpha_1 - \alpha_3 + 3\alpha_2 - 3\alpha_3 \\ &= \alpha_1 + 3\alpha_2 - \alpha_3\end{aligned}$$

which for $\alpha = 0.05$ gives $0.01695 \le$ overall probability of rejection ≤ 0.19091 for any three goodness of fit tests, whether independent or related in any way. Note that these bounds may not be attainable, i.e., they may be wider than necessary.

Summary of Total Probability of Rejection of a True Null Hypothesis

Scenario	Total probability of rejection (Using α_1, α_2, and α_3 based on $\alpha = 0.05$)
Three mutually independent tests	0.0502
Two identical tests and a third independent	0.043
Three identical tests	0.05
General	Between 0.017 and 0.109

An Integrated Assessment of Southern Blue Whiting *(Micromesistius australis)*from New Zealand Using Separable Sequential Population Analysis

Stuart Hanchet
National Institute of Water and Atmospheric Research, Nelson, New Zealand

Vivian Haist
Department of Fisheries and Oceans, Pacific Biological Station, Nanaimo, B.C., Canada

Dave Fournier
Otter Consulting Ltd., Nanaimo, B.C., Canada

Abstract

The southern blue whiting *(Micromesistius australis)* fishery is one of the largest fisheries in New Zealand waters, with landings peaking at over 75,000 t in 1992. Early assessments of the Campbell Island stock were carried out in 1992 and 1993 using Virtual Population Analysis. There was concern over these assessments because diagnostics suggested that the survivors were poorly estimated due to the noisy CPUE data. In addition, estimates of recruited biomass from acoustic surveys could not be fitted and confidence intervals could not be estimated using the existing software.

A separable Sequential Population Analysis was developed to simultaneously analyze all the available data and to address these issues. The 17 years of catch-at-age data were fitted as a multinomial distribution with a median sample size of 100, weighted annually between years. Three acoustic indices of spawning stock biomass were fitted as relative abundance indices, with a coefficient of variation (CV) of 0.3. Eleven years of standardized CPUE indices were converted to relative effort and fitted with a CV of 0.5. Simulated data were used to estimate confidence limits, and included uncertainty in the annual catch, catch-at-age, CPUE, and acoustic data.

The model results suggest that the stock underwent a major decline during the 1980s and early 1990s but has since recovered, due mainly to the recruitment of the strong 1991 year class. However, the extent of the recovery is uncertain largely because of observation error in the tuning indices and the sensitivity of the model to the selectivity assumptions.

Introduction

The southern blue whiting (*Micromesistius australis* Norman) fishery is one of the largest fisheries in New Zealand waters. It was developed in the early 1970s by the Russian fleet and since then landings have fluctuated considerably averaging about 20,000 t, and peaking at over 75,000 t in 1992 (Fig. 1). The fishery targets southern blue whiting as they aggregate to spawn on the Bounty Platform, Pukaki Rise, and the Campbell Island Rise. Fish from these three grounds appear to form separate stocks (Hanchet 1998), and have been treated separately for stock assessment and management purposes. Before 1993, no catch or effort restrictions were in place in the fishery. In 1993, a total catch limit of 32,000 t was introduced with an area limit of 11,000 t on the Campbell Island Rise, 15,000 t on the Bounty Platform, and 6,000 t on the Pukaki Rise. Historically, most fishing has been carried out on the Campbell Island Rise (Fig. 1) and so the rest of this paper deals primarily with that stock.

The first quantitative assessment of the Campbell Island stock was carried out in 1991 using an age-structured stock reduction analysis following the method of Francis (1990). The model was fitted to CPUE data and estimates of Z from catch curve analysis, and assumed deterministic recruitment (Hanchet 1991). The following year a time series of catch-at-age data extending back to 1982 was developed, and this allowed the use of catch-at-age models to be used for the first time. Virtual Population Analysis, tuned using the Laurec-Shepherd method (Laurec and Shepherd 1983), was carried out using the Lowestoft suite of programs (Darby and Flatman 1994). However, the CPUE data used to tune the VPA were noisy, and the diagnostics from the tuning indicated that the terminal fishing mortalities on each age were poorly estimated with CVs ranging from 0.2 to 0.9 (Hanchet 1993). Several limitations were identified with the approach being used: (1) there was likely to be error in the catch-at-age, CPUE, and landings data, (2) a time series of biomass estimates from acoustic surveys was being developed, (3) confidence intervals had not been estimated. A separable Sequential Population Analysis was developed in 1994 to address these issues and to integrate all the available data in one stock assessment (Hanchet and Haist 1994). This paper summarizes the model, the four main sources of data used to assess this stock, and the results of the model fitting.

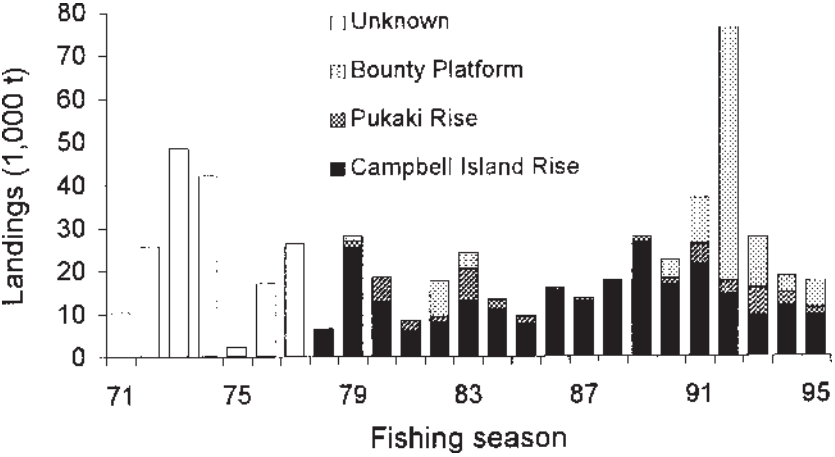


Figure 1. Annual landings of southern blue whiting (t) by the split October-September fishing year.

Model Inputs

Landings Data

Landings for the period 1971 to 1995 are shown in Fig. 1. Landings before 1979 were not split up by area, and are possibly unreliable, and so only the period 1979 to 1995 has been considered in the analysis. The landings since 1979 are based on summaries of estimated catches recorded on trawl catch and effort logbooks completed by all vessels operating in the fishery. The fleet is composed primarily of Japanese surimi trawlers and Russian head and gut trawlers using large midwater nets with codend mesh sizes of 60 mm. The size distribution of fish caught by the two types of vessels are virtually identical suggesting they have similar selectivity patterns. No effort or catch restrictions were in force in the fishery until the introduction of the total catch limit in 1993. Although the catch limit has been reached in two of the three years since 1993, many vessels carry scientific observers, and between 50% and 75% of the catch each year has been observed.

Catch-at-Age Data

Length frequency data and otoliths have been collected from commercial vessels working in the Campbell Island Rise fishery since 1979. Samples before 1986 came mainly from single vessels carrying out exploratory fishing in these waters, whereas samples since 1986 have come from scientific observers on board several commercial fishing vessels each year.

Length frequency data were stratified by time and area and scaled up to the total landings following Schweigert and Sibert (1983).

Between 180 and 780 otoliths were read each year using the validated methodology of Hanchet and Uozumi (1996). Otolith ages were used to construct an age-length key for each year which was combined with the scaled length frequency to obtain the annual catch-at-age, which was then converted to proportion-at-age (see Fig. 4). Between reader variability increased after age 10 (Hanchet and Uozumi 1996), so ages 2 to 10 were used with a plus group at age 11.

One particular feature of the fishery is the occurrence of two distinct spawning grounds on the north and south of Campbell Island Rise. In years when strong year classes are first recruiting to the fishery (as 3-year-olds) the proportion of new recruits is much higher on the southern ground than in the north (Hanchet 1998). In some years the fleet has carried out more fishing on the southern ground which may have caused an increase in the selectivity of the 3-year-old fish in those years.

Acoustic Data

Acoustic surveys of southern blue whiting were carried out in the spawning season in 1993, 1994, and 1995 (Ingerson and Hanchet 1996). The surveys were conducted using the random stratified parallel transect approach of Jolly and Hampton (1990). The survey area was stratified based on the location of tows made by commercial fishing vessels targeting southern blue whiting over the previous ten years, and covered 25,000 km^2. Although primarily designed as a spawning survey, transects were extended inshore to cover the principal nursery grounds so that both adult and prerecruit biomass estimates were obtained. Estimates of backscattering were converted to biomass using the target strength–fish length relationship derived for North Atlantic blue whiting (Ingerson and Hanchet 1996). Each year two acoustic surveys were completed and the results averaged (Table 1).

CPUE Data

Individual tow by tow data were analyzed for the Campbell Island fishery for the period 1986 to 1995 using a generalized linear model (Hanchet and Ingerson 1996). Because of the large number of zero tows (ranging from 2 to 20% per year), a gamma error distribution with a log-link function was used to model CPUE. Annual catch per unit effort indices were obtained after standardizing for the effects of vessel length, depth fished, time of day, tow position, season, and headline height. The resulting CPUE indices were converted to estimates of total annual effort by dividing them into the annual landings (Table 2).

Table 1. Biomass (t) and CV (%) of adult and prerecruit (mainly 2-year-old) southern blue whiting from acoustic surveys of the spawning and nursery grounds on the Campbell Island Rise.

	Adult		Prerecruit	
	Biomass	CV	Biomass	CV
1993	18,500	21	89,600	23
1994	161,400	36	22,400	38
1995	121,100	30	20,000	25

Table 2. Results of standardized CPUE analysis for the Campbell Island Rise.

Year	Number of tows	Percentage zero tows	Relative year effect	Standard error	Landings (t)	Relative effort
1986	893	4.6	1.00	–	15,252	15,252
1987	637	5.3	0.68	0.06	12,804	18,829
1988	843	7.1	0.52	0.04	17,422	33,504
1989	1008	4.7	0.53	0.04	26,611	50,209
1990	994	7.8	0.46	0.04	16,652	35,883
1991	1057	3.7	0.35	0.03	21,314	61,072
1992	1091	18.7	0.23	0.02	14,208	62,044
1993	411	10.7	0.61	0.06	9,316	15,272
1994	384	6.8	0.53	0.07	11,290	21,302
1995	170	2.4	0.88	0.13	9,750	11,080

Modeling

Model Specification

The model developed to analyze the fishery is an extension of the approach outlined by Fournier and Archibald (1982). Errors associated with the observed catch, the proportion-at-age, fishing effort, and the acoustic biomass indices are all explicitly considered in the model. The fit to the total catch and the catch-at-age data are kept separate following Fournier and Archibald (1982), which has allowed incorporation of variability into the age determination process. Because fishing takes place during a very short 2-3 week season in September, at the end of the fishing year, the catch equations assume that 95% of M occurs before fishing starts and the remaining 5% of M occurs during fishing.

The form of the catch equations used in the model is given by the following relationships.

$$C_{ij} = \frac{F_{ij}}{Z_{ij}}\left[1 - \exp(-Z_{ij})\right]N'_{ij}$$

$$C_{i\cdot} = \sum_j C_{ij}$$

$$Z_{ij} = F_{ij} + 0.05M$$

$$N'_{ij} = \exp(-0.95M)\,N_{ij}$$

$$N_{i+1,j+1} = \exp(-Z_{ij})N'_{ij}$$

$$N_{i+1,a} = \exp(-Z_{i,a-1})N'_{i,a-1} + \exp(-Z_{ia})N'_{ia}$$

where,

i	indexes year,
j	indexes age class,
a	is the number of age classes,
C_{ij}	is the catch (in numbers) of age class j fish in year i,
$C_{i\cdot}$	is the total catch (in numbers) in year i,
F_{ij}	is the instantaneous fishing mortality rate for age class j in year i,
M	is the instantaneous natural mortality,
Z_{ij}	is the instantaneous total mortality rate for age class j during the fishing period in year i,
N_{ij}	is the number of age class j fish in the population at the beginning of year i, and

N'_{ij} is the number of age class j fish in the population at the beginning of the fishing period in year i.

Fishing mortality is treated as a function of the observed fishing effort ($\tilde{E}_i$), catchability (q), and age-specific selectivity (s_{ij}). Preliminary analysis suggested asymptotic selectivity, completed by age 5, so selectivity of age 5 and older was assumed to be 1.0. Because of the potential for targeting new recruits, time dependent changes in age-selectivity were also estimated for ages 2 and 3. A single selectivity parameter is estimated for age-4 fish. The relationships describing observed fishing effort, expected fishing effort (E_i), and fishing mortality are:

$$E_i = \tilde{E}_i \exp(d_i^E)$$

$$\text{and } F_{ij} = q\, s_{ij}\, \tilde{E}_i \exp(d_i^E)$$

where,

$s_{ij} = \exp(l_j + d_{ij}^s)$

the d_{ij}^s represent deviations from average selectivity at age j for $j = 2$ and 3, and 0 for $j > 3$.

$l_j = 0$ for $4 < j \le a$

and the d_i^E represent deviations in the effort-fishing mortality relationship.

Following Fournier and Archibald (1982) we assume that the age-composition samples adhere to a multinomial sampling distribution and that annual total catch estimates are independent and lognormally distributed. The negative of the log-likelihood function for the catch-at-age model is then

$$-n_i \tilde{p}_{ij} \ln p_{ij} + \sigma_C \sum_i \left[\ln(\tilde{C}_i) - \ln(C_{i\cdot})\right]^2 + \sigma_E \sum_i (d_i^E)^2$$

where n_i is the sample size, $\tilde{p}_{ij}$ and p_{ij} are the observed and predicted proportion of fish of age j in year i respectively. $\tilde{C}_i$ is the observed catch in year i, and σ_C and σ_E are standard deviations for catch and effort which are described below. An additional component is added to the negative of the log-likelihood function to tune the model to the acoustic survey biomass estimates. These surveys are conducted during the spawning season and the model is fit to the estimates of adult abundance. The model estimates of spawning stock biomass (B_i) are

$$B_i = \sum_j l_j w_j \exp(-0.5 Z_{ij}) N'_{ij}$$

where w_j is the weight of fish in age class j. In this formulation we are assuming that the average selectivity l_j is synonymous with maturity, which is probably valid for this spawning fishery. We assume the survey abundance estimates are relative and they have a lognormal error distribution. The following term is added to the negative of the log-likelihood function.

$$\sigma_B \sum_{i=93}^{i=95} \left[\ln(\tilde{B}_i) - \ln(rB_i) \right]^2$$

where σ_B is the standard deviation of the biomass described below, $\tilde{B}_i$ is the adult acoustic biomass estimate in year *i*, and *r* is the abundance scalar.

The model was implemented using AD Model Builder software (Fournier 1994), which gave simple and ready access to minimization routines, and provided the ability to estimate the variance-covariance matrix for all dependent and independent parameters of interest. The parameters being estimated when minimizing the negative log-likelihood function are $\ln(N_{i1})$, $\ln(N_{1j})$, d_i^E, d_{ij}^s, l_j, q, and r.

Weightings

It was not possible to estimate the various standard deviations so they were fixed at values that represented our levels of confidence in the various data sets. A convenient way to do this was in terms of weights where $w_x = \sigma_x^{-2}$, for each variate *x*, and the corresponding CVs are given for each dataset considered. An estimate of the confidence came from a consideration of both the estimated variance and possible bias inherent in the data. Where appropriate, weights were assigned different values between years.

Annual landings appear to be well estimated in this fishery and so a weight equivalent to a CV of 5% was given to each year's catch.

Weights were assigned to the proportion-at-age based on the sample size in a multinomial distribution. The amount and precision of data collected from this fishery suggests a sample size of around 300 may be warranted (Hanchet and Ingerson 1996). However, the standard multinomial sampling process is not robust to violations of assumptions (Fournier and Archibald 1982). A number of factors including ageing error, sampling bias, and nonconformity with the model assumption of separability would all lead to the sample size being inflated relative to the true deviations of predicted versus observed proportions-at-age. The sample size was therefore reduced in the model to a value of 100.

The earlier data (pre-1986) are much less reliable than more recent data because only one vessel was sampled each year, and there were fewer length-frequencies taken and otoliths collected and read. Therefore, the sample size was adjusted between years by the proportion of tows made in that year compared to the median number of tows in the series following Cordue (1993):

$$n_i = n(t_i/t_{med})$$

where, n_i is the sample size in year i, n is the sample size, t_i is the number of tows sampled in year i, and t_{med} is the median number of tows in any year in the series.

A weight of 5 (equivalent to a CV of 30%) was used for each acoustic data point which is consistent with the average of the CVs from the surveys. There is no reason to believe that the biomass was better estimated in any of the years, since each estimate was based on two surveys, so the indices were given equal weighting for each year.

The standard errors of the CPUE indices estimated by the general linear model were quite low, suggesting a high weighting (Table 2). However, because of the highly aggregated nature of the fishery, it was considered that the CPUE series may not be accurately monitoring abundance and that it should have a lower weighting than the acoustic indices. Therefore, the weight for the effort data was assigned a value of 2, which is equivalent to a CV of 50%.

Other Model Inputs

Natural mortality (M) was estimated by dividing $\log_e 100$ by A_{max}, where A_{max} is the age reached by 1% of the population. Using aging data from 1986, A_{max} was calculated to be 22 years giving an estimate for M of 0.21 (Hanchet 1991). This has since been rounded down to 0.2 to reflect its imprecision. In the absence of data to the contrary, M is assumed to be constant for all ages and years.

Weight at age was calculated from the weight-length relationship and von Bertalanffy growth coefficients given in Hanchet (1991) and assumed to be constant for each year.

Estimation of confidence intervals

Simulated data were used to estimate confidence limits for the results. Four sources of uncertainty were included in the procedure. The individual otolith length-age data within individual years were resampled (with replacement) and then scaled up to catch-at-age using the weighted length frequency of the catch for that year. Uncertainty in the acoustics data, the effort data and annual catch was captured by assuming the data were lognormally distributed with CVs of 30%, 50% and 5%, respectively. For each of the 500 bootstrap runs data were randomly selected from each distribution. The percentile method was used to estimate confidence intervals (Effron 1981): the estimate of the 90% confidence interval was computed as the 5th and 96th points in the set of bootstrap estimates after sorting them into ascending order.

Sensitivity Analysis

A number of sensitivity analyses were carried out to examine the sensitivity of the model results to alternative model assumptions. These included the relative weightings given to the catch-at-age and effort data, *M*, selectivity assumptions, and whether the acoustic indices were treated as relative or absolute.

Results and Discussion

The biomass trajectory and 90% confidence intervals are plotted in Fig. 2. The results suggest that the stock underwent a large decline during the 1980s and early 1990s but has since recovered, due mainly to the recruitment of the strong 1991 year class (Fig. 3). The wide confidence limits suggest that estimates of current biomass and the 1991 year class are highly uncertain. However, independent supporting evidence of its size comes from the large prerecruit biomass in the 1993 acoustic survey (Table 1), and the very slow growth rate of this year class (Hanchet 1998). The model suggests that this year class is at least 3 times the size of the strong 1979 and 1980 year classes, and about 10 times the average.

The fit of the model to the adult acoustic biomass is shown in Fig. 2. The value of the abundance scalar *r* was 1.14, indicating reasonable agreement between the absolute adult acoustic estimate and population biomass. However, the model is unable to fit the decline in the index between 1994 and 1995, suggesting observation error in one or other of these indices. The predicted proportions-at-age fit the observed proportions-at-age very well for most year classes (Fig. 4). The model has problems fitting the observed effort data between 1989 and 1993, but overall there are no obvious trends in the residuals (Fig. 5).

The results of the sensitivity analysis are shown in Table 3. The results were relatively insensitive to the relative weightings examined for the catch-at-age and effort data. Estimates of historic biomass were most sensitive to the value of *M*. Estimates of current biomass and the size of the 1991 year class were most sensitive to the inclusion of time dependent age 3 selectivity, and to a lesser extent *M*.

Now the reason for freeing up the annual selectivity of the 2- and 3-year-olds is that the proportion of fish in these age classes caught in the commercial fishery can be quite variable. Targeting on 2-year-olds occurs by fishing shallower than usual, while targeting on 3-year-olds occurs when fishing the southern ground (Hanchet 1998). The model estimates a large deviation in the selectivity of 3-year-olds when the proportion of 3-year-olds in the catch-at-age data is inconsistent with the proportion of fish of that year class in other ages. A plot of the annual selectivity of 2- and 3-year-olds by year shows that targeting of the 3-year-olds often occurs when strong year classes are recruiting to the fishery for the first time (e.g., the 1979, 1980, 1988 year classes) (Fig. 6). In the catch-at-age data there was a high proportion of 3-year-olds of the 1991 year class relative to the proportion of 2- and 4-year-olds of this year class. Therefore when the selectivity was freed up the model estimated a selectivity of 0.85, whereas

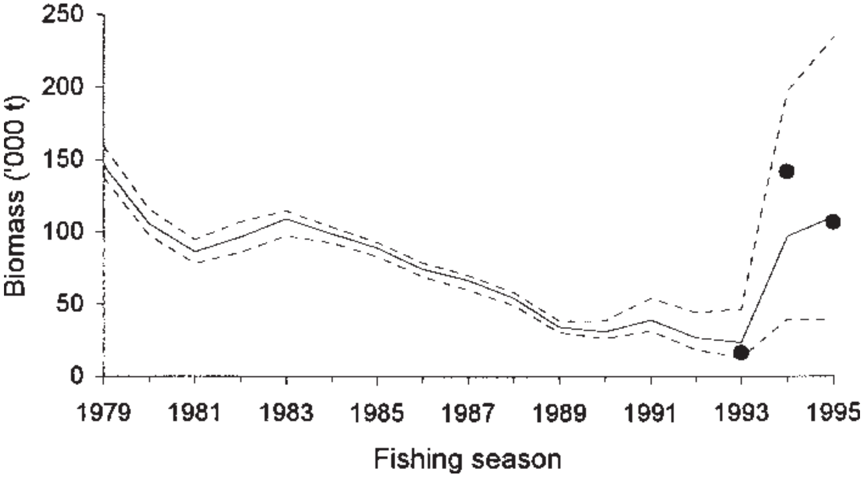


Figure 2. Mid-season spawning stock biomass (t) with 90% confidence intervals, showing the fit to the three adult acoustic survey indices.

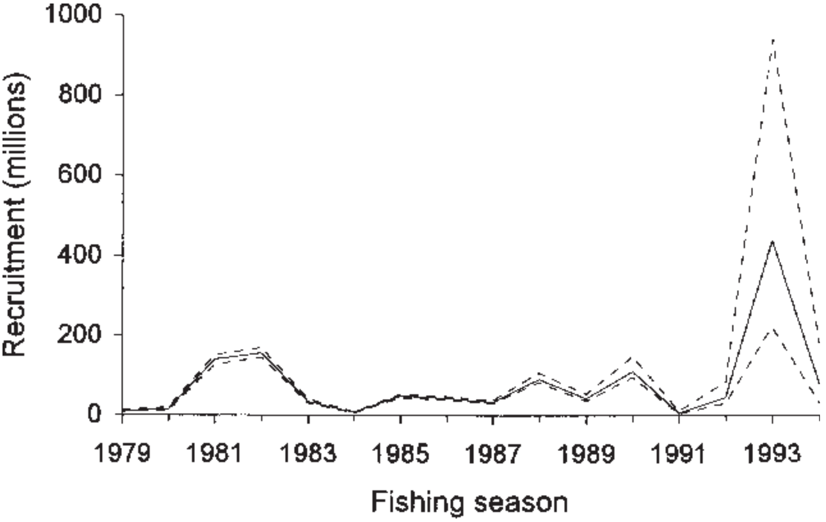


Figure 3. Estimated number of 2-year-olds in the population with 90% confidence intervals.

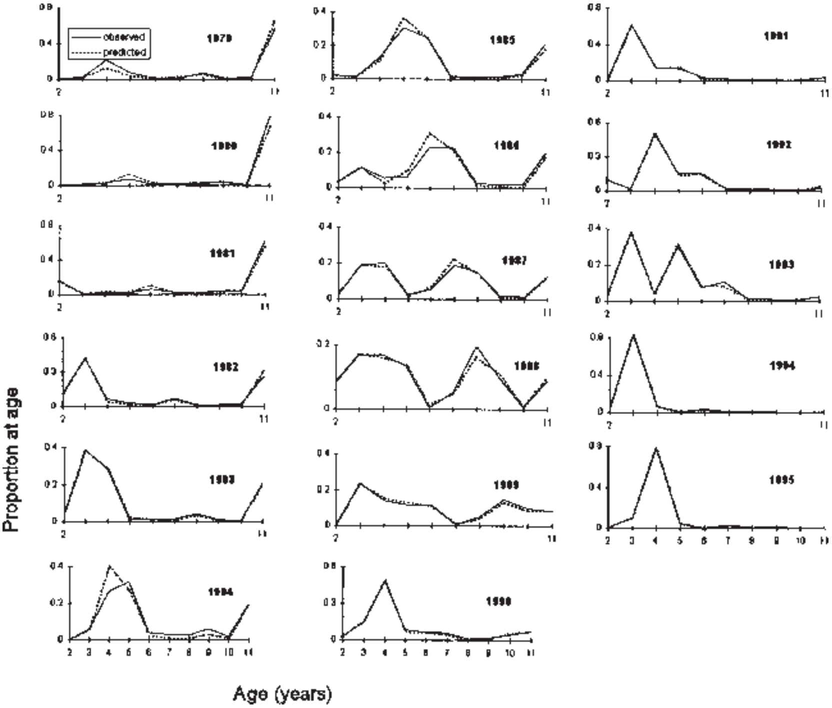


Figure 4. Observed and expected proportion-at-age in the commercial fishery.

when it was kept constant between years it was assigned the mean value (which was 0.58 in that run). Without knowledge of the extent of targeting, or an independent estimate of the abundance of 3-year-olds, it is not possible to discern which selectivity value is the most appropriate. This does, however, highlight the need to include the prerecruit acoustic indices in the fitting procedure in future assessments of the stock.

The model sSPA has overcome some of the limitations of the earlier models used to assess the stock (Hanchet 1991, 1993). Following Fournier and Archibald (1982) and Methot (1989), the model has been structured to integrate and simultaneously analyze data from fishery dependent and fishery independent data sources. Errors associated with the observed catch, the proportion-at-age, fishing effort, and the acoustic biomass indices are all explicitly considered in the model, and confidence intervals have been estimated for the first time. The Ad Model Builder software gives ready access to minimization routines, is flexible enough to allow the incorporation of future data as they become available, and is fast enough to enable a large number of sensitivity analyses and the estimation of

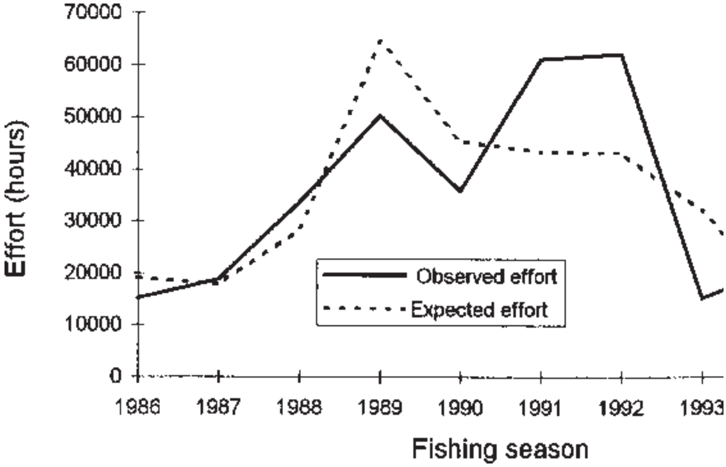


Figure 5. Observed and expected effort. The observed effort was calculated from CPUE data, the expected effort was estimated from the model.

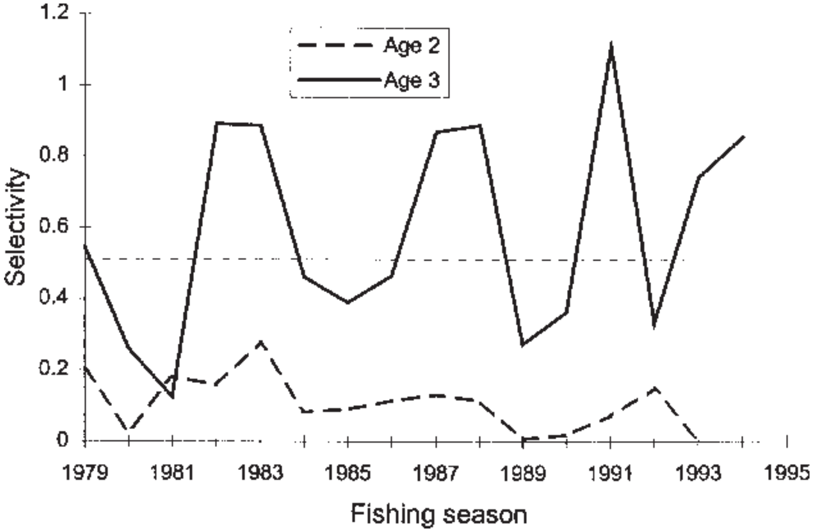


Figure 6. Estimated annual deviations in the selectivity of 2- and 3-year-olds and the 3-year-old mean.

Table 3. Relative changes (expressed as percentages) of selected parameter estimates as a result of alternative model assumptions for the stock. *B,* mid-season spawning stock biomass; R_{1991}, size of the 1991 year class.

Model	B_{1980}	B_{1988}	B_{1995}	R_{1991}
Sample size = 325	0.0	0.1	2.0	1.6
Sample size = 50	–0.1	–0.1	–1.8	–1.3
Effort CV = 70%	0.1	0.5	9.1	7.3
$M = 0.15$	–24.7	–9.2	–9.1	–20.0
$M = 0.25$	37.3	11.3	11.2	26.4
Acoustic index absolute	0.3	0.9	11.8	9.5
Age 3 selectivity constant between years	3.4	0.8	29.5	32.9

confidence intervals to be carried out on a routine basis. Future improvements to the model will include the use of an annual weight-at-age matrix, the incorporation of aging error, and the fitting of age structured acoustic indices.

Acknowledgments

Thanks to C. Francis, A. Betlehem, and an anonymous referee for constructive comments on earlier drafts of the manuscript. This project was funded by the New Zealand Ministry of Fisheries project number SBW9701.

References

Cordue, P.L. 1993. MIAEL estimation of biomass and fishery indicators for the 1993 hoki stock assessment. N.Z. Fisheries Assessment Research Document 93/15. 33 pp.

Darby, C.D., and S. Flatman. 1994. Virtual population analysis: Version 3.1 (Windows/ DOS) user guide. MAFF [Ministry of Agriculture, Food and Fisheries] Information Technology 1. 85 pp.

Effron, B. 1981. Nonparametric estimates of standard error: The jackknife, the bootstrap and other methods. Biometrica 68(3):589-599.

Fournier, D. 1994. An introduction to AD Model Builder, for use in nonlinear modelling and statistics. Otter Research Ltd. Canada. 51 pp.

Fournier D., and C.P. Archibald. 1982. A general theory of analysing catch-at-age data. Can. J. Fish. Aquat. Sci. 39:1195-1207.

Francis, R.I.C.C. 1990. A maximum likelihood stock reduction method. N.Z. Fisheries Assessment Research Document 90/4. 12 pp.

Hanchet, S.M. 1991. Southern blue whiting fishery assessment for the 1991-92 fishing year. N.Z. Fisheries Assessment Research Document 91/7. 48 pp.

Hanchet, S.M. 1993. Southern blue whiting *(Micromesistius australis)* fishery assessment for the 1993-94 fishing year. N.Z. Fisheries Assessment Research Document 93/17. 56 pp.

Hanchet, S.M. 1998. A review of southern blue whiting *(Micromesistius australis)* stock structure. N.Z. Fisheries Assessment Research Document 98/8. 28 pp.

Hanchet, S.M., and V. Haist. 1994. Southern blue whiting *(Micromesistius australis)* fishery assessment for the 1994-95 fishing year. N.Z. Fisheries Assessment Research Document 94/22. 41 pp.

Hanchet, S.M., and J.K.V. Ingerson. 1996. Southern blue whiting *(Micromesistius australis)* fishery assessment for the 1995-96 fishing year. N.Z. Fisheries Assessment Research Document 96/12. 34 pp.

Hanchet, S.M., and Y. Uozumi. 1996. Age validation and growth of southern blue whiting, *Micromesistius australis* Norman, in New Zealand. N.Z. J. Mar. Freshw. Res. 30:57-67.

Ingerson, J.K.V., and S.M. Hanchet. 1996. Acoustic biomass estimates of southern blue whiting *(Micromesistius australis)* from the Bounty Platform, Pukaki Rise, Campbell Island Rise, and Auckland Island Shelf, August-September 1995. N.Z. Fisheries Assessment Research Document 96/18. 29 pp.

Jolly, G.M., and I. Hampton. 1990. A stratified random transect design for acoustic surveys of fish stocks. Can J. Fish. Aquat. Sci. 47:1282-1291.

Laurec, A., and J.G. Shepherd. 1983. On the analysis of catch and effort data. J. Cons. Cons. Int. Explor. Mer. 41:81-84.

Methot, R. 1989. Synthetic estimates of historical abundance and mortality for northern anchovy. Amer. Fish. Soc. Symp. 6:66-82.

Schweigert, J.F., and J.R. Sibert. 1983. Optimising survey design for determining age structure of fish stocks: An example from British Columbia Pacific herring *(Clupea harengus pallasi).* Can. J. Fish. Aquat. Sci. 40:588-597.

Statistics, Software, and Fish Stock Assessment

Jon T. Schnute, Laura J. Richards, and Norm Olsen
Fisheries and Oceans Canada, Pacific Biological Station, Nanaimo, B.C., Canada

Abstract

Classical regression, analysis of variance, and experimental design fail to provide adequate tools for building modern fishery models. In the last decade, however, both statistical theory and software have progressed in directions that make realistic stock assessment models possible. We review the modern theory and show how to apply it to complex fishery analyses, in which dual simulation and estimation models relate parameters to available data. In particular, we demonstrate that state space models in a Bayesian framework provide useful design prototypes. More generally, the theoretical framework leads to probability distributions that must be translated into computer code. We discuss software requirements for implementing this framework. We also present benchmark results from a catch-at-age model rendered in four current software environments.

Introduction

A typical fish stock assessment uses historical data to evaluate the impact of fishing on past, present, and future stock status. As preparation for this work, most assessment scientists have taken at least a few university courses in statistics. Such courses generally include a thorough discussion of regression theory, analysis of variance, and experimental design. Unfortunately, even a fairly detailed knowledge of classical statistics fails to equip the scientist with all the tools needed to build realistic assessment models. For example, a regression model cannot generally explain the link between available data and underlying stock biomass.

In this paper, we examine recent statistical approaches to designing models for complex data sets, such as those encountered in stock assessments. We show that these new methods offer an integrated framework in which complex biological processes can be interpreted in the light of available data. We also investigate various software environments for implementing the analyses required by these complex models, and we provide benchmark comparisons for evaluation.

Both statistical theory and computer software have experienced rapid changes over the past decade. Problems in stock assessment that once seemed unique to fisheries science have now been addressed in a wider framework, with applications in many fields. Furthermore, these concepts allow a systematic approach to model development, in which assumptions range from simple to complex. Model design starts with two obvious questions. First, what biological and fishery processes govern fish population dynamics in space and time? Second, how do the observed data relate to these underlying processes? These questions can each be given deterministic or probabilistic answers. The combination of answers defines a *state space model* (SSM; Harvey 1989, Schnute 1994) that describes data collection from a dynamic population. A corresponding Bayes posterior distribution can be used for inferring unknown parameters from available data.

Stock Assessment

Figure 1 portrays the essential elements of stock assessment. The process begins with “Data” compiled from fisheries and field studies. A leftward arrow (marked “EDA”) from “Data” to “Model” represents the exploratory data analysis required to identify a model appropriate for the data. Thus, the model describes underlying processes that theoretically produce the observed data. Unknown “Parameters” relate to particular processes, such as survival, growth, and selectivity by the fishery. The model involves “Controls,” such as catch removed by the fishery. Environmental variables, such as water temperature, might also act as controls. We regard the controls as known quantities that contribute directly to the available data (arrow from “Controls” to “Data”). Similarly, controls directly influence system dynamics (arrow from “Controls” to “Model”). The upward arrow marked “Policy” indicates that a harvest policy algorithm in the model might also dictate some of the controls. Thus, the model might be used to design an appropriate control policy.

Given parameters and controls, the model can be used to generate simulated data, and this relatively easy simulation process can be instructive for understanding model characteristics. A much harder analysis deals with the inference problem: estimating parameters from the available data. Ideally, probability distributions in the simulation model should, via “Statistical Theory,” imply an “Estimation Algorithm” for resolving the inference problem. For example, a classical linear regression model with additive normal error implies a least-squares algorithm for estimating the intercept and slope of a line. In the present context, where the simulation model might be highly complex, an appropriate estimation algorithm may be less obvious. Thus, we require that the statistical theory developed here should be general enough to resolve the inference problem for a broad

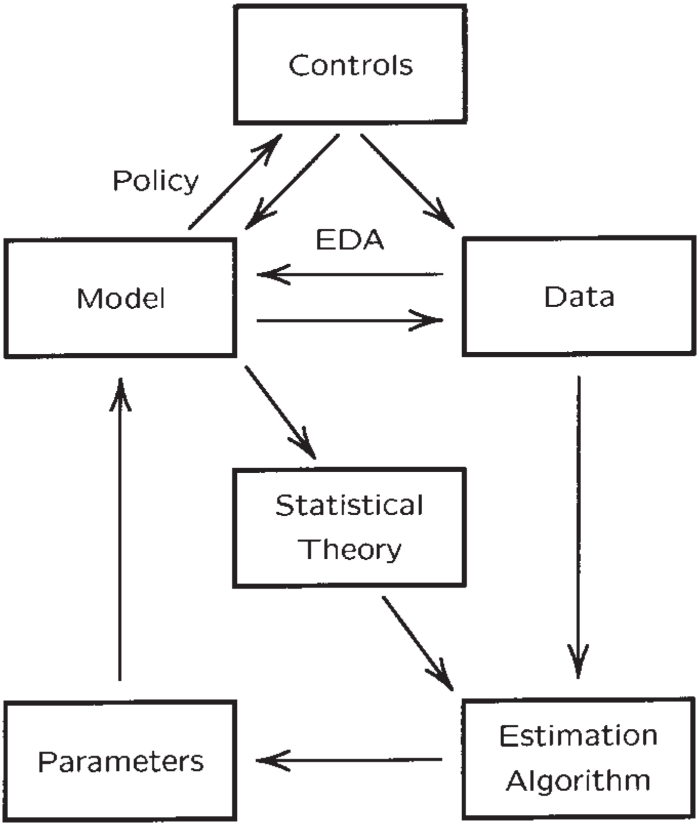


Figure 1. Logical relationships among components of the stock assessment process.

class of stochastic models that could potentially produce the observed data.

Figure 1 suggests an iterative process that can be used to test both model design and statistical theory. Start with a hypothetical model, specified parameters, and a given control policy. Run the model to produce simulated data, and use the estimation algorithm to obtain parameter estimates from these data. Essentially, this process starts with "Parameters" and cycles through "Model," "Data," "Estimation Algorithm," and back to "Parameters." Repeating this process many times gives a distribution of estimates conditional on the original parameters.

These cycles of simulation and estimation sometimes show that the available data cannot be used reliably to reconstruct the population dynamics. Furthermore, the control policy may influence how well parameters can be estimated. Typically, controls that vary substantially from year to year produce contrasting data that better reveal the system dynamics. Schnute and Richards (1995) illustrate this principle in the analysis of catch-at-age data. An adaptive management policy (Walters 1986, Hilborn and Walters 1992) seeks to optimize information about the system, as well as optimal return from the resource.

State Space Models

Analysis from Fig. 1 begins by identifying the contents of the boxes "Data," "Controls," and "Model." For example, consider a fishery in which available data consist of the annual catch C_t and an abundance index I_t in each year t. Assume that the catch C_t also constitutes the only known control. Within the population, let R_t and S_t denote numbers of new recruits and survivors from past years, respectively, at the start of each year t. An appropriate model for this fishery could include the equations

$$S_t = (R_{t-1} + S_{t-1} - C_{t-1})\, e^{-M+\sigma\delta_t} \tag{1}$$

$$I_t = q(R_t + S_t - \tfrac{1}{2} C_t)\, e^{\tau\varepsilon_t}, \tag{2}$$

where δ_t and ε_t are independent standard normal variates (mean 0, variance 1). The dynamic process equation (1) describes survival from year $t-1$ to year t, where removals by the fishery precede natural mortality. Survival includes a lognormal process error with natural mortality parameter M and associated standard error σ. The observation equation (2) relates the index I_t to the population midway through the fishery, when half the catch has been removed. The parameter q acts as a proportionality constant, and an associated lognormal measurement error is scaled by the standard error τ.

This simple example illustrates the concept of a state space model (SSM). Hidden states $X_t = (R_t, S_t)$ are governed by a dynamic process equation (1), and observations $Y_t = (I_t)$ result from these states by an observation equation (2). Both equations (1) and (2) are influenced by the controls $Z_t = (C_t)$ and the model parameters $\Theta = (M, q, \sigma, \tau)$. More generally, an SSM is characterized by three probability distributions:

$$P(X_1 \mid Z_1, \Theta); \tag{3}$$

$$P(X_t \mid X^{t-1}, Z^t, \Theta);\ t \geq 2; \tag{4}$$

$$P(Y_t \mid X^t, Z^t, \Theta);\ t \geq 1; \tag{5}$$

where superscripts denote cumulative vectors $X^t = (X_1, \cdots, X_t)$ and $Z^t = (Z_1, \cdots, Z_t)$. Thus, the distribution (3) of the initial state vector X_1 depends on the initial control Z_1 and the parameters Θ. Similarly, for $t \geq 2$, the distribution (4) of X_t is conditional on all previous states X^{t-1} and all controls Z^t up to the current year t. Finally, the observations Y_t in (5) depend on all states and controls up to the current year.

Any model (3)-(5) explicitly meets the simulation requirements discussed above in connection with Fig. 1. Start with the model, known parameters, and given controls. Use (3) to generate X_1; apply (4) iteratively to simulate the population dynamics; and thus obtain the complete population vector X^T for a fixed time $t = 1, \cdots, T$. Finally, use (5) to simulate the data Y^T that would result from the simulated population.

Statistics

We have so far established an explicit correspondence between an SSM and four boxes in Fig. 1: "Parameters" ↔ Θ; "Model" ↔ (3)-(5); "Controls" ↔ Z_t; "Data" ↔ Y_t. It remains to explain how "Statistical Theory" produces an "Estimation Algorithm." Given the data and controls for years in the range $1 \leq t \leq T$, unknown quantities to be estimated include model parameters Θ and hidden states X^T. A function that involves all these quantities is the product of probabilities (3)-(5):

$$L(\Theta, X^T) = P(X_1 \mid Z_1, \Theta) \prod_{t=2}^{T} P(X_t \mid X^{t-1}, Z^t, \Theta) \prod_{t=1}^{T} P(Y_t \mid X^t, Z^t, \Theta). \tag{6}$$

The first two components of the product in equation (6) represent the joint distribution of all states X^T, where this is expressed as an iterative product of conditional distributions. The third component of the product corresponds to the distribution of the data Y^T, given the states X^T. From a Bayes perspective, $L(\Theta, X^T)$ comprises a likelihood function for the unknown quantities, based on the model in equations (3)-(5). The posterior distribution is then given by the proportionality

$$P(\Theta, X^T) \propto L(\Theta, X^T)\ Q(\Theta, X^T), \tag{7}$$

where $Q(\Theta, X^T)$ denotes a prior distribution.

The posterior (6)-(7) exploits the fact that a joint probability distribution can be expressed as a product of conditional probabilities. In the SSM paradigm, these dependencies come primarily from transitions between time steps $t - 1$ and t, although the approach can be extended to cover a broader range of conditional distributions. For example, fishery selectivities might be modeled as random variates from an underlying distribution independent of t. Such structural assumptions introduce extra factors in the product (7). Thus, if $\Theta = (\Theta_1, \Theta_2)$ and Θ_1 is conditional on Θ_2, then the distribution of Θ can be expressed as the product $P(\Theta) = P(\Theta_1 \mid \Theta_2)\ P(\Theta_2)$.

In general, a simulation model defines a sequential calculation of quantities relevant to the analysis, such as the states X_t above. At each step of this calculation, some quantities are already known from previous computations, and others remain to be computed. As illustrated by equations (3)-(5), the stochastic steps involve probability distributions, in which quantities to be simulated depend conditionally on quantities already computed. The Bayes approach treats all unknown quantities as parameters, and the "Statistical Theory" in Fig. 1 consists of defining a Bayes posterior distribution proportional to the product of all these conditional distributions. For simulation purposes, the model involves drawing random variates from the proposed distributions. For estimation purposes, the model prescribes a Bayes posterior distribution. Standard Bayes inference techniques then provide a description of uncertainty in all unknown quantities, based on the known data.

This Bayesian link between simulation and estimation offers an attractive universal approach to models with arbitrary complexity. Theoretically, almost any deterministic or random process can be contemplated in the fishery model and evaluated in the light of available data. Clifford (1994) describes the shift in thinking among statisticians as follows:

> The recent revival of interest in Bayesian methods arises not from a wholesale conversion among statisticians to Bayesian ideology, but from a need to deal with high dimensional parameter spaces in modern applications…. In these applications it is convenient to co-ordinate parameters by imposing probabilistic structures analogous to those introduced by Bayesian theory. The structures may involve relatively few hyperparameters which themselves are amenable to estimation by classical methods…. When the number of parameters goes into the thousands, it is less obvious how to proceed and the clear centralist dogma of Bayesianism becomes an attractive starting point.

Clifford's remarks have particular relevance to the SSM context discussed here. The unknown states X^T can be considered a high dimensional parameter vector "coordinated" by "hyperparameters" Θ. Thus, the stochastic model (3)-(5) prescribes the likelihood (6), in which states X^T are linked stochastically (coordinated) by relatively few underlying system parameters (hyperparameters) Θ. For example, in a model for catch-at-age data with A age classes and T years of data, the annual state vector X_t has A components; thus, X^T contains $A \times T$ unknown parameters. This number may seem large to scientists accustomed to historical analyses, such as virtual population analysis (VPA), in which population sizes are computed deterministically by back calculation. Nevertheless, even VPA produces a complete set of $A \times T$ population estimates, although the fact that the computed numbers are merely estimates is somewhat disguised by the deterministic calculation.

In summary, a state space model leads to the problem of estimating both parameters and states. To simplify our notation, we represent the complete set of unknowns as $\theta = (\Theta, X^T)$, and we denote the function on the right side of (7) as $p(\theta) = L(\theta)Q(\theta)$. Thus, $P(\theta) \propto p(\theta)$, where $p(\theta)$ can be computed directly from probability distributions in the model. If $p(\theta)$ has a mode (maximum value) at $\hat{\theta}$, then the second order Taylor expansion

$$\log p(\theta) \approx \log p(\hat{\theta}) - \tfrac{1}{2}(\theta - \hat{\theta})' H (\theta - \hat{\theta}) \tag{8}$$

corresponds to the normal approximation in which θ has mean $\hat{\theta}$ and covariance matrix H^{-1}. The prime symbol in (8) denotes transpose, and H is the Hessian matrix of second partial derivatives of the function $[-\log p(\theta)]$, evaluated at $\hat{\theta}$. This approximation leads to the first computational requirement for Bayes inference:

Problem 1. Given the function $p(\theta)$, compute $\hat{\theta}$ and H^{-1}.

A software environment for solving problem 1 must allow the distributions (6)-(7) to be converted easily into computer code. In addition, the environment must provide tools for maximizing a function, computing its second partial derivatives at the maximum, and inverting a matrix.

Unfortunately, the solution to problem 1 alone can give a poor representation of the Bayes posterior. Distributions of some parameters may be highly skewed; consequently, the symmetric confidence regions of the normal approximation can be deceptive. Perhaps more seriously, the posterior may have multiple modes, corresponding to different biological interpretations of the available data. Theoretically, information from the posterior should be available by computing its density in various regions of parameter space. However, the "curse of dimensionality" represents a significant impediment to a brute force approach. For instance, examining a range of 10 values for each of m parameters would require $10m$ cells, more than the available memory on most computers when $m > 8$. A more reasonable computational alternative is described in

Problem 2. Given the function $p(\theta)$ proportional to the posterior $P(\theta)$ and a large integer n, draw a random sample of n vectors θ_i $(1 \le i \le n)$ from $P(\theta)$. Then use properties of the sample to estimate properties of the posterior.

An approximate solution to problem 2 comes from solving problem 1 and applying a known algorithm for sampling a multivariate normal distribution. However, a general solution to problem 2 cannot depend on such specific algorithms, because $P(\theta)$ typically is much more complex than any classical distribution.

In recent years, numerous methods have been proposed for solving problem 2 (Tanner 1993, Gilks et al. 1996, Punt and Hilborn 1997). Some of the most popular can be categorized as Markov chain Monte Carlo (MCMC) methods. These techniques start at a trial point θ_1 and move in a Markov process to future points θ_i $(1 \le i \le n)$; thus, each new trial θ_i depends only on the previous point θ_{i-1}. Relative values $p(\theta_i)$, automatically proportional to $P(\theta_i)$, are used to decide whether to retain trial points from the Markov chain. An initial "burn in" segment of the sample is usually ignored. The solution to problem 1 can greatly assist a solution to problem 2, because the choice $\theta_1 = \hat{\theta}$ starts the chain at the point of highest posterior density. Furthermore, the inverse Hessian H^{-1} can be used to set an appropriate movement scale for the Markov process. If the solution to problem 1 is ignored, software requirements for solving problem 2 are minimal. For example, only a few lines of code are required to implement the popular MCMC Metropolis-Hastings algorithm (Gilks et al. 1996, p. 7).

The discovery of posterior sampling algorithms has had a profound effect on Bayesian statistics. Clifford (1993) summarized this effect as follows:

> It used to be that you could tell a Bayesian by his tendency to hold meetings in isolated parts of Spain and his obsession with coherence, self-interrogation, and other manifestations of paranoia. Things have changed.... What the authors [of MCMC sampling methods] have done today is to announce that from now on we can compare our data with the model that we actually want to use rather than a model which has some mathematically convenient form. This is surely a revolution.... Having opened up this Pandora's box of allowing people to use the models that they want to use, the question is whether or not statistics is ready to deal with what is inside. Will we support the widespread use of Monte Carlo methods or will we try to preserve our professional mystique by introducing jargon and acronyms?"

Software

A software environment for stock assessment should make it easy to accomplish all tasks outlined in Fig. 1. These include:

- Facilities for loading data from text files or databases ("Data").
- Graphical tools for exploratory data analysis ("EDA").
- Functions to generate random variates from a wide variety of distributions ("Model," used for simulation).
- A language for expressing the probability distributions (3)-(5) that define the inference function $p(\theta)$.
- Numerical routines for solving the inference problems 1 and 2, in particular, for maximizing $p(\theta)$ ("Estimation Algorithm").

The software should also support vectors, matrices, and arrays of arbitrary dimensions, along with the usual associated functions (e.g., sums, products, inverses, eigenvalues, decompositions) commonly used in modern applications. Ideally, the environment should offer an interactive com-

Table 1. Software products used in the benchmark, with associated version numbers and choice of optimizer used for solving problem 1.

Product	Company	Generation	Optimizer
AD Model Builder 2.0.1 (Borland C++ 5.01)	Otter Research Ltd.	3.5 GL	(Native)
GAUSS 3.2.28	Aptech Systems, Inc.	4 GL	Optmum
MATLAB 5.0.0.4069	The MathWorks, Inc.	4 GL	Fminu
S-PLUS 3.3 (Release 1)	MathSoft, Inc.	4 GL	Nlmib

mand line for investigative graphics and trial analyses. Software packages with the features discussed here are commonly called fourth generation languages (4GLs), to distinguish them from the classical third generations languages (3GLs), such as FORTRAN, PASCAL, C, and C++.

Table 1 lists four packages that we have tested in light of the above requirements. Three of them (GAUSS, MATLAB, S-PLUS) can be considered 4GLs, although the current version of GAUSS supports only two-dimensional matrices. Thus, GAUSS cannot easily deal with fishery models where populations have multiple indices (such as age, year, area, and species). Also, the graphical facilities currently available in GAUSS are relatively primitive compared with those in MATLAB or S-PLUS.

The AD Model Builder (ADMB) package, which we have classified as a 3.5 GL, essentially provides a C++ class library that extends the C++ language to address some of the 4GL requirements identified here. The library is distributed in binary form for a variety of popular C++ compilers, which must be obtained separately. The letters AD refer to automatic differentiation, a technique that allows the gradient of a function to be computed automatically from the computer code for the function itself. Efficient algorithms for locating a function's maximum generally require computed values of both the function and its gradient. Thus, as demonstrated in our benchmarks, automatic differentiation can dramatically assist the solution to problem 1. The ADMB package offers full matrix support, but no graphics or interactive command line. Programs are written as templates, which are converted to C++ code by a text-processing program. Code within a template compares favorably with similar code that might be written for a 4GL.

As a benchmark example from fisheries literature, we used the catch-at-age model proposed by Schnute and Richards (1995). We chose this reference partly because computer code can be written directly from tables listing the model equations and likelihood function, which are too lengthy to include here. We first applied S-PLUS to simulate data from a given parameter vector θ_0. We then attempted to estimate these parameters

Table 2. Benchmark trials for two cases of a catch-at-age model, in which *T* years of data are available for *A* age classes and the total number of parameters is *m* = *T* + *A* + 5.

Case	T	A	m	Product	ms[a]/call	Calls	Time	Scale
1	20	12	37	ADMB	29	161	4.6 s	1.0
				GAUSS	42	4,041	2.8 min	1.0
				MATLAB	178	1,936	5.8 min	1.0
				S-PLUS	1429	n/a	n/a	
2	80	15	100	ADMB	131	291	38 s	8.3
				GAUSS	167	23,365	1.08 hr	23.1
				MATLAB	639	18,360	3.25 hr	33.6
				S-PLUS	n/a	n/a	n/a	

[a] ms = millisecond

with each software package, based on the same simulated data set and the inference function $p(\theta)$ given by Schnute and Richards (1995). More precisely, we coded $p(\theta)$ in each software package and used an available optimizer (Table 1) to locate $\hat{\theta}$, starting from the initial point θ_0. Thus, each optimizer was faced with the same problem of moving from θ_0 to the modal estimate $\hat{\theta}$, where $\hat{\theta} \neq \theta_0$ due to stochastic error in the simulation model. With T years of data for A age classes, the total number of parameters (dimension of θ) in this model is $m = T + A + 5$.

To keep the benchmark simple, we confined our analysis to the estimation phase of problem 1. Thus, our results do not include the computation of H^{-1}. We used a Pentium 133 MHz computer with 48 MB of RAM, operating under Windows NT 4.0. We considered two cases, a small example with only 37 parameters and a larger one with 100 parameters (Table 2). We accepted the default settings for each product and optimizer shown in Table 1. For each run, we recorded the total time required (Time), the number of function calls (Calls), and the time in milliseconds for each call (ms/call).

It soon became clear that S-PLUS cannot handle problems of substantial size, although the language theoretically can express all mathematical concepts required for Fig. 1. In each case marked n/a in Table 1, the S-PLUS interpreter exhausted computer memory before finishing the calculation, apparently due to a memory management philosophy (Mathsoft, Inc. 1993, chapter 12) inappropriate for the problems considered here. Furthermore, in the small example where we could at least obtain the time for one evaluation of $p(\theta)$, S-PLUS was much slower than the other products.

ADMB performed much faster and required fewer function calls than its competitors in this benchmark. The reduced number of function calls stems directly from automatic differentiation. One call in ADMB evaluates both the function and its gradient. The other products approximate gradients numerically by making a small change to each component of θ; thus, they require $m+1$ function calls to achieve the same result. Furthermore, ADMB obtains gradients accurate to machine precision, whereas the other products lose precision in the subtraction needed to approximate derivatives. Consequently, the ADMB optimizer can perform more efficiently near the minimum, where gradients are small and the loss of precision from differencing is greatest. Also, ADMB uses the "reverse" method of automatic differentiation, which is known to be highly efficient (Griewank and Corliss 1991).

As problems grow in size, the required computer time increases. For example, cases 1 and 2 in Table 2 have 37 and 100 parameters, respectively. Thus, case 2 scales case 1 upward in size by the factor $2.7 = 100/37$. If case i requires computing time t_i ($i = 1,2$), then the scale factor t_i/t_1 measures the time increase for case i relative to case 1. For each product in Table 2, the scale factor (Scale) for case 2 is greater than 2.7, due partly to increased time for a function call in the more complex model. Nevertheless, ADMB appears more scalable than the other products, probably from the efficiencies of reverse automatic differentiation.

Discussion

Statistics has recently moved in new directions that can make stock assessment more systematic and rigorous. As discussed by Clifford (1994), we have adopted the Bayes approach, "not from a wholesale conversion," but from the need to work with high dimensional models where the "clear centralist dogma of Bayesianism" provides an adequate general framework. Within this framework, the "Statistical Theory" in Fig. 1 can be summarized by the inference problems 1 and 2, where an inference function $p(\theta)$ can be written explicitly for a broad range of models. Perhaps fishery analysts can finally use models with realistic complexity, rather than a limited range of classical models with "some mathematically convenient form" (Clifford 1993).

We have not attempted to address some of the ambiguities that typically accompany this approach. For example, even in the simple model (1)-(2), different results can be obtained for different choices of the measurement to process error ratio τ/σ. As models become more complex, the number of such choices increases, and the ambiguities may dominate the analysis. As discussed earlier, simulation-estimation cycles in Fig. 1 can be used to test how well available data identify the underlying population dynamics. In fishery data analysis, the results are often disappointing.

Like statistical theory, software tools have also seen recent substantial advances. The modeling framework in Fig. 1 and the computational requirements for solving problems 1 and 2 dictate certain minimal requirements, which we have itemized in the first paragraph of the preceding section. We chose the products benchmarked here for three reasons.

First, they represent a blend at the level of third and fourth generation languages essential for the minimal requirements identified here. Our discussion shows the strengths and limitations of each product from this point of view. Second, two products (ADMB and S-PLUS) already have been used extensively for building sock assessment models. We have personal experience with both (e.g., Schnute and Richards 1995, Richards et al. 1997); thus we have some confidence in our ability to write appropriate code. Third, the remaining products (GAUSS and MATLAB) have explicitly extended their 4GL capabilities to include routines for addressing the optimization issues inherent in solving problem 1. Other optimization languages and software (e.g., AMPL; Fourer et al. 1993) focus primarily on linear and nonlinear programming problems, which differ substantially from the statistical problems 1 and 2 cited above. For an extensive review of optimization software in various contexts, see Moré and Wright (1993) and the related link to their software guide at the Internet site http://www.mcs.anl.gov/otc/.

We do not think that the ideal software tool exists at present. We would like a rich language with an interactive command line for fast prototyping. It would connect easily to our data sets and produce revealing graphics for rapid insight into complex multivariate data. Vectors, matrices, and multidimensional arrays would be intrinsically available, along with all the usual functions and transformations of these objects. It would include every known mathematical function, and it would produce random variates drawn from all known univariate and multivariate distributions. Automatic differentiation of all orders would be supported; for example, an exact (to machine precision) calculation of the Hessian matrix of second derivatives would be routinely available. A fast optimizer would be linked to all these capabilities. An application module would implement all the concepts in Fig. 1; thus, we could supply only a model and data, and the rest of the analysis would follow automatically. Perhaps through hidden compilation, code would run as fast as if compiled and would be upwardly scaleable to large data sets. The language would contain a lucid help facility, and high quality documentation would be available from diverse authors and publishers. Finally, everyone would use it, so that code could always be shared.

We certainly do not consider our benchmark exhaustive of the possibilities or even definitive for the products considered here. Software and hardware change continuously, and we intend primarily to offer a framework for discussion. We recognize that languages and software environments require long learning curves to master. People naturally become committed to a tool they understand, and they experience frustration when it fails to extend easily to meet new requirements. We hope that our example encourages readers to test and document results from other software environments.

We also hope that our comments stimulate vendors to provide better software tools. For example, our tests convincingly demonstrate the advantage of automatic differentiation in the time required to estimate $\hat{\theta}$ in problem 1. The framework in Fig. 1 shows that this is an important issue. Many cycles of simulation and estimation may be required to test the effectiveness of a control policy or the potential for estimating a model's parameters. For example, consider the problem of investigating 4 control policies on 100 trajectories, each running 50 years into the future. All model parameters must be estimated for each future year on each trajectory. If the combined simulation-estimation cycle for each year takes about 40 seconds (Table 2, case 2, ADMB), then the entire run requires over 9 days ($4 \times 100 \times 50 \times 40$ seconds). If the cycle takes one hour (Table 2, case 2, GAUSS), the run time extends to more than 2 years! Obviously, run time can dictate whether an analysis is feasible or not.

The Babel tower myth offers some insight into the current software reality for the community of fishery scientists. After the people of Babel sought to build a tower to heaven, the Lord God devised a plan (Genesis 11: 4-7). "Behold the people is one; and they all have one language; and this they begin to do; and *now nothing will be restrained from them, which they have imagined to do*.... Let us go down, and there confound their language, that they may not understand one another's speech." Our italics highlight the prospects for accomplishment with a common language, if the scientific community could ever agree on one.

Although modern statistical and software tools have become quite sophisticated, no elaborate model should preclude a thoughtful examination of the facts. Fishers often accuse scientists of placing excessive faith in mathematical models and ignoring practical experience. We regard models as tools for investigating biological scenarios consistent with the facts, analogous with a detective reconstructing the crime scene from available clues. A credible model scenario can also be supported by an intuitive examination of the raw data (Richards et al. 1997). The final product must go beyond the mathematical paradigm of Fig. 1 to an explanation that policy makers can understand and communicate to a skeptical public.

Acknowledgments

We thank D. Fournier (Otter Research Ltd.), R. Schoenberg (Aptech Systems, Inc.), N. Miller (The MathWorks, Inc.), and S. Kaluzny (Data Analysis Products Division of MathSoft, Inc.) for helpful conversations regarding the software products benchmarked here. R. Deriso introduced us to the MCMC algorithm and drew our attention to the quotation by Clifford (1994). Cliff Kraft encouraged us with his keen interest in a historical and philosophical perspective on statistics and stock assessment. Two anonymous reviewers provided critical comments leading to an improved final draft.

References

Clifford, P. 1993. Page 53 in: Discussion on the meeting on the Gibbs sampler and other Markov chain Monte Carlo methods. Journal of the Royal Statistical Society, Series B, 55:53-102.

Clifford, P. 1994. Pages 34-35 in: M.A. Newton and A.E. Raftery. 1994. Approximate Bayesian inference with the weighted likelihood bootstrap, and discussion. Journal of the Royal Statistical Society, Series B, 56:3-48.

Fourer, R., D.M. Gay, and B.W. Kernighan. 1993. AMPL: A modeling language for mathematical programming. Boyd and Fraser Publishing Company, Danvers, MA.

Gilks, W.R., S. Richardson, and D.J. Spiegelhalter. 1996. Markov chain Monte Carlo in practice. Chapman and Hall, London, U.K.

Griewank, A., and G.F. Corliss (eds.). 1991. Automatic differentiation of algorithms: Theory, implementation, and application. Society for Industrial and Applied Mathematics, Philadelphia, PA.

Harvey, A.C. 1989. Forecasting, structural time series models and the Kalman filter. Cambridge University Press, Cambridge, U.K.

Hilborn, R., and C.J. Walters. 1992. Quantitative fisheries stock assessment: Choice, dynamics and uncertainty. Chapman and Hall, New York.

Mathsoft, Inc. 1993. S-PLUS programmers manual. StatSci, a Division of MathSoft, Inc. Seattle, WA.

Moré, J.J., and S.J. Wright. 1993. Optimization software guide. Society for Industrial and Applied Mathematics, Philadelphia, PA.

Punt, A.E., and R. Hilborn. 1997. Fisheries stock assessment and decision analysis: The Bayesian approach. Rev. Fish Biol. Fish. 7:35-63.

Richards, L.J., J.T. Schnute, and N. Olsen. 1997. Visualizing catch-age analysis: A case study. Can. J. Fish. Aquat. Sci. 54:1646-1658.

Schnute, J.T. 1994. A general framework for developing sequential fisheries models. Can. J. Fish. Aquat. Sci. 51:1676-1688.

Schnute, J.T., and L.J. Richards. 1995. The influence of error on population estimates from catch-age models. Can. J. Fish. Aquat. Sci. 52:2063-2077.

Tanner, M.A. 1993. Tools for statistical inference, 2nd edn. Springer-Verlag, New York.

Walters, C.J. 1986. Adaptive management of renewable resources. Macmillan, New York.

Fishery Stock Assessment Models
Alaska Sea Grant College Program • AK-SG-98-01, 1998

A Statistical Framework for Analysis of Limit Reference Points

Laura J. Richards, Jon T. Schnute, and Norm Olsen
Fisheries and Oceans Canada, Pacific Biological Station, Nanaimo, B.C., Canada

Abstract

Fisheries stock assessment typically involves an evaluation of historical stock status and a forecast of future status under one or more harvest policies. In this paper, we describe an integrated modeling framework both for estimation of historical stock status and for simulation of future status under a precautionary management regime. The paper extends our earlier work, where we used state-space models to conduct the estimation component of the analysis. Here we demonstrate how large uncertainties in historical biomass trends can be propagated into forecasts of future status. Following the precautionary approach, fisheries models must consider the probability that stock biomass will fall below preset limit (conservation) reference points within a specified time horizon. Because reference points are estimated from historical data, they are themselves highly uncertain. Our method uses estimates of the model parameters and covariance matrix to recreate plausible scenarios for the historical state dynamics and corresponding reference points. For each scenario, we then simulate future states under a fixed catch policy, allowing for process error in recruitment. The outcome for a given time horizon and harvest policy can be perceived as a bivariate scatterplot with forecasted biomass on one axis and the corresponding reference point on the other axis. Thus, our model framework provides an explicit method for evaluating the likelihood that future stock sizes will fall below the preset limit. We also suggest visualization tools for clearly portraying these risks to decision makers.

Introduction

The United Nations Straddling Stocks Agreement (United Nations 1995) provides explicit guidelines on application of the precautionary approach. In the context of fishery stock assessments, reference points are identified as the primary advisory tool. Reference points may be expressed in terms of fishing mortality rates or stock biomass (Caddy and Mahon 1995). Limit reference points define upper bounds to fishing mortality rates or lower bounds to biomass. Under the precautionary approach, harvest strategies are constrained within these safe biological limits. According to Annex II.7 of the Agreement, "For stocks which are not overfished, fishery management strategies shall ensure that fishing mortality does not exceed that which corresponds to maximum sustainable yield, and that the biomass does not fall below a predefined threshold."

In this paper, we describe a statistical framework that is explicitly designed to evaluate the risk of either approaching or exceeding the limit reference points. Specifically, the Agreements state that the fishing mortality F shall be maintained at a rate below F_{msy}, the F corresponding to maximum sustainable yield. Similarly, stock biomass B shall be maintained at a level above a limit biomass B_{lim}. These conditions can be stated mathematically as the probability

$$P(F > F_{msy}; B < B_{lim}) < \alpha, \tag{1}$$

where α is small. The key problem is to define harvest strategies such that condition (1) is met. The problem is complicated by uncertainties in the definition and measurement of F_{msy} and B_{lim}, as well as in stock status and underlying biological relationships.

We describe an integrated modeling approach for evaluating the risk that a condition similar to (1) is not met. Our approach allows estimation of uncertainties in historical stock status and other key stock assessment parameters, such as the natural mortality rate. Simulations of future stock status then incorporate these historical uncertainties, as well as process error in recruitments. We include a case study where we use such forecasts to evaluate alternative harvest control policies. In addition, we suggest practical visualization tools for clearly portraying risks in advisory contexts. We also demonstrate how advice can become less precautionary when the full uncertainties are not considered.

Methods

Reference Points and Harvest Control Laws

Several real problems are associated with evaluating the probability in (1). Both the fishing mortality rate F and the biomass B are measured with error. Furthermore, in the example here, F is not measured directly, but inferred from the estimate of B and fishery removals in the catch. Thus, if B is overestimated, then the historical F will be underestimated. Similarly, if B is underestimated, then the historical F will be overestimated.

Estimation of F_{msy} in (1) poses additional problems. The concept of F_{msy} is based on equilibrium conditions that are unlikely to be valid for stochastic environments. A stock-recruitment function is also required. Unfortunately, the relationship between recruitment and spawning stock biomass is poorly determined in most cases, particularly when uncertainty in the estimates of stock biomass and recruitment are acknowledged.

For these reasons, we prefer to illustrate our approach with control policies based on the catch rather than F and F_{msy}. Unlike F-based quantities, the catch can be directly observed and measured in most cases. Furthermore, F-based harvest policies are often translated into a catch or TAC for actual implementation in a management context. We suggest the condition

$$P\left(B_t < B_{lim} \mid \{C_s\}_{s=1}^{t-1}\right) < \alpha, \tag{2}$$

as an alternative to (1), where future years s $(1 \le s < t)$ lie within a time horizon of t years. In (2), B_t denotes the biomass at the start of the final year t, given a proposed sequence of annual catches $\{C_s\}_{s=1}^{t-1}$. In this paper, we envisage a fixed catch policy $C_s = C$, and we examine time horizons t in the range $1 \le t \le T$. Of course, other control policies could be adopted. For example, in a strictly precautionary approach, the catch would be set to 0 when $B_t < B_{lim}$.

An appropriate value for B_{lim} in (1) and (2) is obviously stock specific. Values of B_{lim} typically are set relative to historical estimates of biomass and other model parameters. For example, B_{lim} could be specified as a proportion of the estimated unfished biomass or as the lowest observed biomass. Thus, both B_{lim} and B_t in (2) can be random variables. We discuss this issue further in the context of our case study.

Given a definition of B_{lim} and a harvest control policy, (2) can be interpreted as a measure of risk; a higher value of α implies a higher risk to the stock. Similarly, the longest time horizon T provides a measure of sustainability. For example, for large values C_t, the condition in (2) may be met for short time horizons only. More generally, we define the probability

$$\alpha_t = P\left(B_t < B_{lim} \mid \{C_s\}_{s=1}^{t}\right) \tag{3}$$

as a measure of risk in a specific year t within a time horizon of $1 \le t \le T$ years.

Estimation and Simulation

To evaluate risk and sustainability from the probability in equation (3) for a particular fish stock, we conduct the following steps:

1. Develop a statistical fishery model, tailored to the observed data for the stock.
2. Apply the model to obtain estimates of the model parameters and the covariance matrix for these estimates.
3. Draw a sample parameter vector from the multivariate normal distribution estimated in step 2, and thus obtain a sample historical trajectory for the stock. Pick a large number *n*, and repeat this process *n* times to obtain *n* independent trajectories, each of which reflects the estimated correlations among model parameters.
4. Choose an appropriate definition for B_{lim}, a range of harvest control policies based on a constant catch, and a time horizon *T*.
5. Project forward *T* years from the final year in each historical trajectory using simulated recruitments and a constant catch.
6. Repeat step 5 with different choices of constant catch, but using the same set of *n* historical trajectories and future recruitment scenarios.

Our analysis could be applied to a wide range of fishery models that produce biomass and recruitment estimates. We illustrate the process here in the context of catch-age analysis (Fournier and Archibald 1982; Methot 1989, 1990; Schnute and Richards 1995). We assume for our case study that recruitment is unrelated to stock size, within the range of stock sizes of interest. Thus, we can redraw recruitments from the estimated recruitment series and apply the same set of recruitments to each simulated control policy. However, if $C_t \geq B_t$ in any future year *t*, we allow the population to crash, setting the remaining future recruitments to zero.

A feature of our approach that differs from other methods is that we explicitly include uncertainty in the model parameter estimates and in estimates of current biomass used as a starting point for forward projections. For example, our forward projections can include uncertainty in estimates of the natural mortality rate *M*. A more typical approach would be to start the forward simulations at the current biomass estimate and include in the projections only uncertainties in future process. We term this the "future process error" approach and apply it to demonstrate the importance of the "full model error" approach.

Case Study

To illustrate our methods, we choose data from a stock of Pacific Ocean perch *(Sebastes alutus)* from southern Queen Charlotte Sound, Canada. For this stock, longevity is up to 90 years, recruitment is sporadic with strong year classes about once a decade, and recruitment to the fishery and maturity occur between ages 6-15 years. A complete catch-age analysis is described by Richards et al. (1997), based on the model of Schnute and Richards (1995). These papers described stock reconstructions and their associated uncertainties; we did not consider forward biomass projections, harvest policies, or appropriate choices of reference points. Here, we extend our work to examine forward projections of stock biomass under different control policies.

Definition of B_{lim}

Rockfish (*Sebastes* spp.) stocks along the Pacific coast of North America were targeted by Soviet and Japanese fleets during the late 1960s and early 1970s before Canada extended seaward jurisdiction to 320 km. Several analyses indicated low levels of biomass by the late 1970s (Ketchen 1981, Kimura 1981, Archibald et al. 1983, Ianelli and Heifetz 1995). The fishing industry also agreed that our study stock was depleted at this time. Thus, we choose estimated stock biomass in 1977 (at the time of extended jurisdiction) as our measure of B_{lim}. Due to uncertainties in stock and recruitment estimates (Fig. 8 in Richards et al. 1997), we have no evidence for a stock-recruitment relationship that might provide an alternative value for B_{lim}.

Estimation

Our approach follows state-space design principles (Schnute 1994, Schnute et al. 1998). The state dynamic equations of our catch-age model are expressed in terms of numbers N_{at} at both age a and time t. A key feature of the stock reconstruction is that we assume separability of the proportion of the fish N_{at} selected by the fishery into age and time components. Specifically, we assume that fishery selectivity increases asymptotically with age and that this pattern is time independent. The initial states for the model must allow calculation of N_{a1} for each a; we accomplish this through estimation of historic recruitments, assuming $F = 0$ for $t < 1$. The observation equation then compares the proportions at age determined from the state dynamics and the observed proportions at age in the catch. We use a multivariate logistic function (Schnute and Richards 1995) for this comparison. A second observation equation uses a catchability constant to relate the predicted biomass at the time of the survey to the survey biomass index. Two index series were available, covering different periods with one year of overlap. We estimate a separate catchability constant for each, assuming lognormal error.

For Pacific Ocean perch, age proportion data cover the 34-year span from 1963 to 1996 and the model parameter vector contains 54 terms. Following the rationale in Richards et al. (1997), we fix one parameter, the ratio of recruitment process error to survey measurement error. All other parameters, including the natural mortality rate M, are treated as quantities to be estimated. Computationally, we require procedures to estimate the parameter vector by minimizing the negative log likelihood. We obtain these estimates with the software package AD Model Builder (Otter Research Ltd. 1994) which provides automatic derivatives (Griewank and Corliss 1991) for C++ code and a gradient based search algorithm. The software also facilitates the computation of covariances among parameter estimates and derived quantities based on normal approximations (Appendix B in Richards et al. 1997).

Biomass Reconstructions and Projections

Given estimates of the model parameter vector and covariance matrix, we generate multivariate normal random draws and obtain 300 samples from the multidimensional parameter space. If sample parameter values obtained by this process fall exterior to pre-determined parameter bounds used during estimation, we set the value at the bound. For example, we constrain M to the range (0.02,0.08). We also reject a few cases where the implied historical stock biomass becomes 0. For each random vector, we then apply the model equations to recreate the perceived population state dynamics. Thus, we obtain 300 plausible scenarios for the historical biomass trajectory and, consequently, 300 plausible scenarios for the current state.

The next step in our analysis requires forward projections of recruitment. Due to the apparent lack of a stock-recruitment relationship (Richards et al. 1997), we choose to project recruitments by re-sampling from the estimated 1963 to 1996 recruitment series. Because the historical recruitments are components of the parameter vector, each scenario provides us with a different series of recruitments for the forward projection.

For each set of 300 scenarios (past estimates and future recruitments), we apply a constant catch policy and project forward for $T = 50$ years. We examine constant catch policies ranging from 0 to 4,000 t. If the projected biomass in any scenario becomes smaller than the catch, the biomass is set to 0 and maintained at 0. We also repeat these simulations for the future process error approach, using one historical scenario represented by the maximum likelihood parameter estimates. Again, we re-sample from estimated recruitments on this trajectory to obtain 300 forward projections of 50 years each. We consider the same catch policies as in the full model error analysis.

Results

Sample trajectories from the full model error reconstructions consistently indicate that the stock declined during the late 1960s and early 1970s from historical highs in the 1960s (Fig. 1A). For most trajectories, biomass was relatively stable during the late 1970s and early 1980s and then increased through the early 1990s. The historical trend in biomass appears better determined than the absolute biomass level; trajectories starting with a high biomass in 1963 tend to lead to a high biomass in 1996, the final year of historical data. Forward projections, illustrated in Fig. 1A for

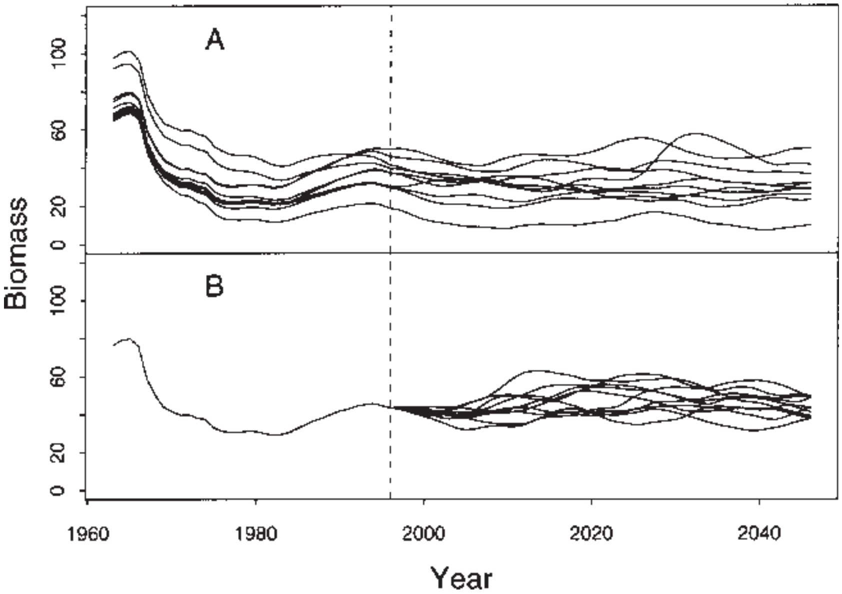


Figure 1. Ten sample trajectories of stock biomass (1,000 t) from model reconstructions and forward projections under a catch policy of 2,000 t for the (A) full model error approach and (B) future recruitment error approach. The vertical dashed line identifies the final year of reconstruction.

a catch policy of 2,000 t, appear more entangled and overlap to a greater degree. For the future process error approach, one historical trajectory identifies the maximum likelihood biomass estimates (Fig. 1B). The forward projections begin at the 1996 biomass estimate and diverge somewhat over the first decade of the simulation. However, even after 50 years, the projections span a narrower biomass range than the full model error projections.

Quantiles in Fig. 2 summarize the 300 reconstructions and projections for full model error under four catch policies. Given our sample size, the 5%, median and 95% quantiles of the biomass projection appear relatively stable over time. Under a "no fishing" policy, biomass increases to a level consistent with estimates of historical biomass. Biomass tends to increase from the 1996 level for catches smaller than about 2,000 t; otherwise, biomass tends to decrease. In fact, for a catch policy of 3,000 t, 12% of the simulated stocks become extinct by year $T = 50$.

As might be anticipated from Fig. 1B, the comparable quantile plots appear more certain for the future process error approach (Fig. 3). The range in biomass represented by the 5% to 95% quantiles spans less than

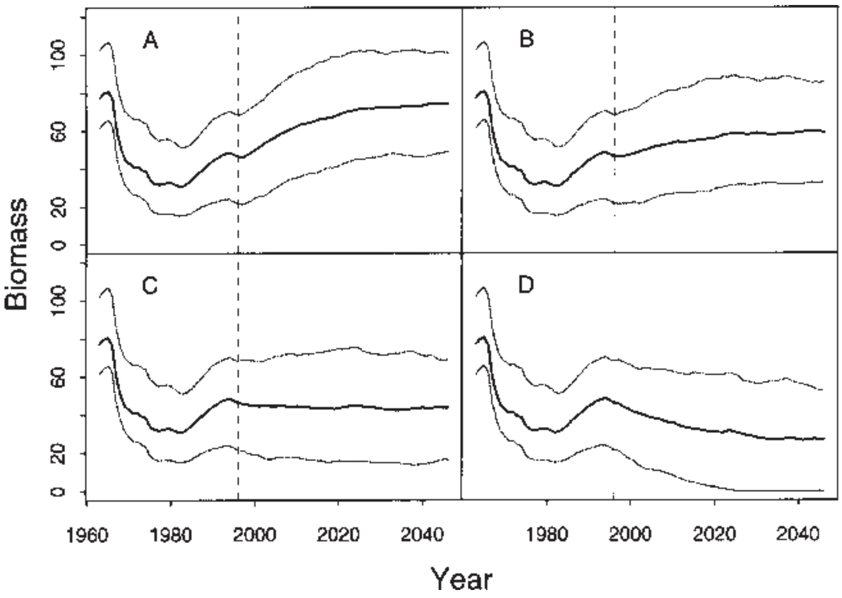


Figure 2. Median (heavy line) with 5% and 95% quantiles (light lines) of stock biomass (1,000 t) for each year in the reconstruction and projection for the full model error approach under catch policies of (A) no fishing, (B) 1,000 t, (C) 2,000 t, and (D) 3,000 t. The vertical dashed line identifies the final year of reconstruction.

half that for full model error in Fig. 2. In particular, the entire range in biomass does not include 0 for a catch policy of 3,000 t. Thus, in contrast to Fig. 2D, Fig. 3D suggests that a catch policy of 3,000 t is sustainable, although the stock is reduced to a low biomass level.

In order to apply the condition in equation (2) or the probability in equation (3) to quantify risk, we must relate B_t to B_{lim}. For full model error, Fig. 4 portrays the biomass B_t in the final projected year ($T = 50$) with our definition of B_{lim}, the estimated 1977 biomass. Estimates of B_{lim} differ for each historical trajectory, ranging from approximately 10,000 to 60,000 t. The final projected biomass is obviously correlated with B_{lim}; low estimates of final biomass tend to be associated with low estimates of B_{lim}. Thus, our interpretation of B_{lim} depends on the model parameter estimates and the corresponding uncertainty in the 1977 biomass estimate.

For the different catch policies in Fig. 4, points below the 45° line represent projections where $B_t < B_{lim}$. This condition never occurs under a no fishing policy. However, even under a low catch policy of 1,000 t, the condition is met for 3% of the simulations. The condition is more likely to

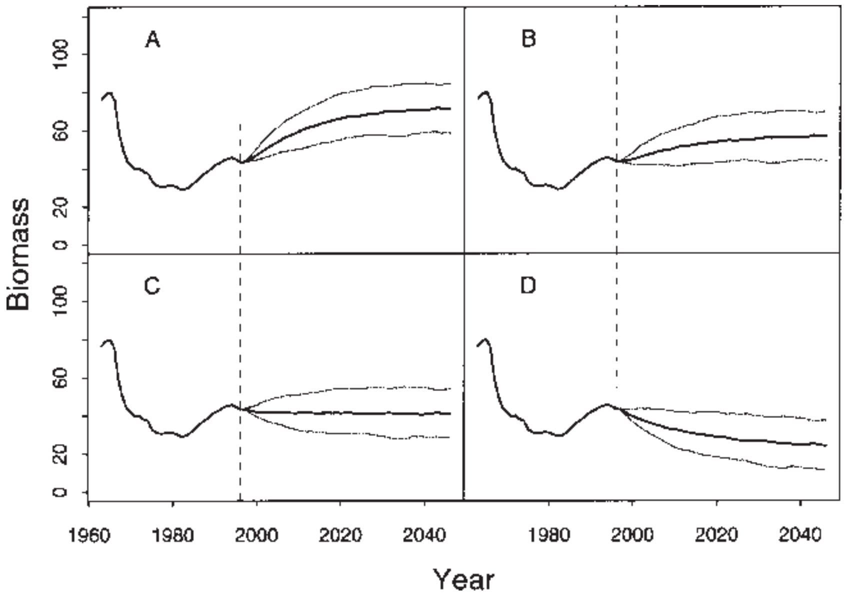


Figure 3. Historical estimates of stock biomass (1,000 t) with the median (heavy line) and 5% and 95% quantiles (light lines) for each year of the forward projections for the future process error approach under catch policies of (A) no fishing, (B) 1,000 t, (C) 2,000 t, and (D) 3,000 t. The vertical dashed line identifies the final year of estimation.

be met for higher catch policies. Under a catch policy of 3,000 t, 75% of the simulations lead to final biomass estimates smaller than B_{lim}.

For Fig. 4, we used the long term time horizon $T = 50$ years. To ensure sustainability, we must examine the condition in (2) under a range of time horizons. We compute α_t in equation (3) from the proportion of simulations for which $B_t < B_{lim}$ for each year t. When full model error is considered, $\alpha_1 = 0.12$ at the starting point for the projections, prior to any catch removals (Fig. 5). Under no fishing, α_t decreases to near 0 in the first few years of the projection. Similarly, under a catch policy of 1,000 t, α_t declines to a small, near constant value. Values of α_t increase over time for catch policies of 2,000 and 3,000 t. A catch policy of 3,000 t appears particularly risky; $\alpha_t > 0.5$ for future years $t > 16$.

For future process error, $\alpha_t = 0$ (no measurable risk) for almost all years under no fishing or a catch policy of 1,000 t (Fig. 5). Under higher catch policies, $\alpha_t = 0$ for at least the initial years in the projection. The apparent risk associated with a catch policy of 2,000 t is substantially smaller than that suggested for full model error. However, the two error

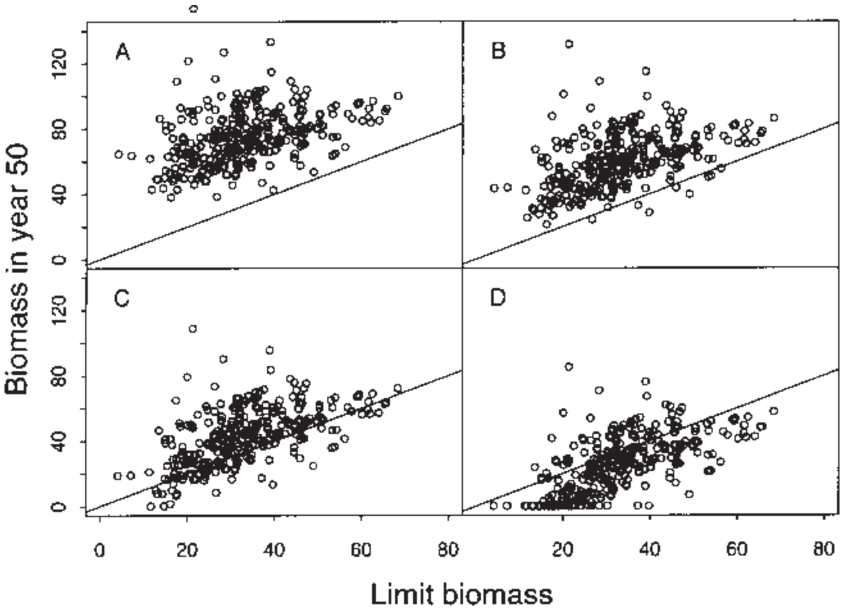


Figure 4. Projected biomass B_{50} (1,000 t) in relation to historical estimates of B_{lim} for the full model error approach under catch policies of (A) no fishing, (B) 1,000 t, (C) 2,000 t, and (D) 3,000 t. The solid line identifies the condition $B_{50} = B_{lim}$.

approaches lead to similar estimates of α_t for a catch policy of 3,000 t for future years removed from the effect of the initial conditions. Thus, even the future process error approach appears able to detect a high risk policy in the long term.

Figure 6 relates α_{50} to the catch policy. This choice of time horizon ($T = 50$) captures the overall trend for each policy in Fig. 5. Values α_{50} increase progressively with larger catches. The increase from 0 is more rapid for full model error than for future process error over the range of catch policies of potential interest to fishery managers. For example, $\alpha_{50} = 0.25$ for full model error and only 0.12 for future process error under a catch policy of 2,000 t.

Discussion

We have described an integrated framework for estimation of historical stock status and simulation of future states. This framework allows us to demonstrate how uncertainty in model parameter estimates influences our interpretation of a biomass limit reference point. We have also pro-

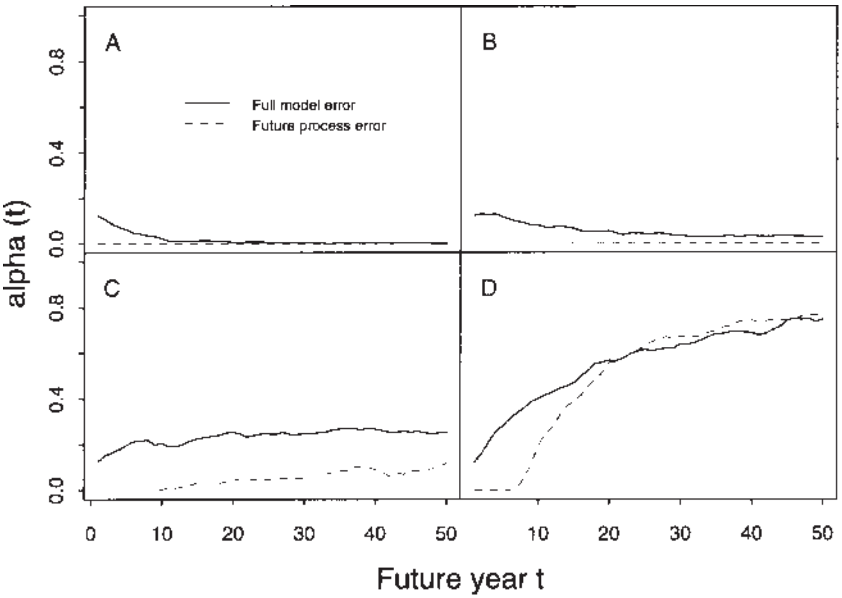


Figure 5. The measure of risk α_t as a function of future year t for both error approaches under catch policies of (A) no fishing, (B) 1,000 t, (C) 2,000 t, and (D) 3,000 t.

vided an explicit method for evaluating the risk that future stock sizes will fall below B_{lim}. In particular, we showed that risk can be underestimated substantially when full model error is not considered. Our approach could easily be extended to other reference points that can be defined in terms of model quantities (Schnute and Richards 1998).

A major difficulty in forecasting future states is the knowledge of future conditions. Because we forecasted recruitments by re-sampling from the estimated series, we implicitly assume that the future will be consistent with the past. We also assume that the model is correctly specified. Thus, our results should be interpreted as short-term risk projections, where the long-term simulations indicate potential trends.

Figure 6 suggests a practical method for portraying risk in an advisory context. For our case study, small catch policies lead to a measurable risk. Thus, if managers want to maintain $\alpha < 0.05$ in (2) for a fixed catch policy, then the catch must be smaller than 1,500 t ($\alpha_{50} = 0.06$). Furthermore, from Fig. 5, the condition $\alpha < 0.05$ cannot be met in the short term even with no fishing. Obviously, this analysis points to a compromise between the definition of B_{lim} and the choice of α. For example, a precau-

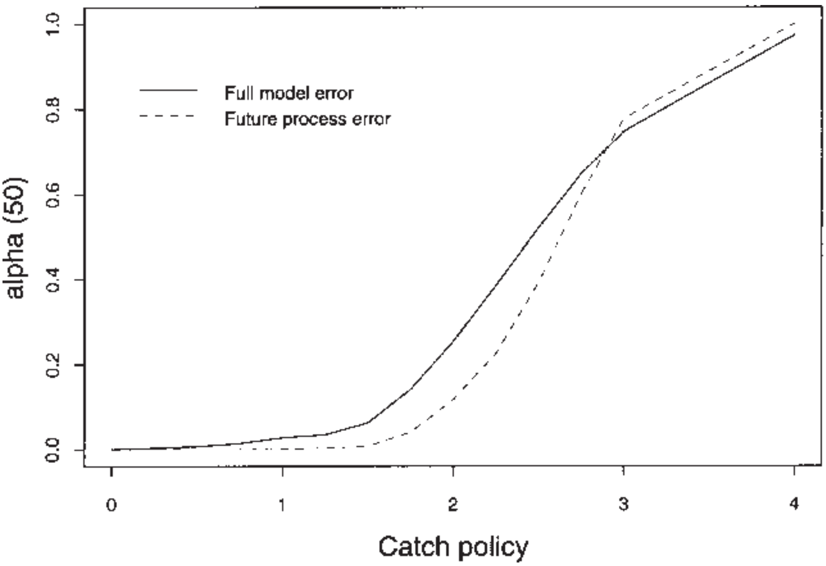


Figure 6. The risk α_t measured for the final simulation year $T = 50$ as a function of catch policy for both error approaches.

tionary choice of B_{lim} and a moderate value of α could result in the same catch policy as a less precautionary choice of B_{lim} and a low value of α.

Our results are obtained from a fish population with very low natural mortality rates and infrequent strong recruitment events. We anticipate somewhat different patterns in forward projections using data from other types of stocks. Furthermore, we have only examined constant catch policies. Potential rebuilding policies with varying catch in equations (2) and (3) will become an issue for stocks with a current biomass smaller than B_{lim}.

An obvious extension of our work would incorporate an adaptive management policy, based on future data. Because we have ignored the value of future information, we have probably overestimated the long-term risk. We have also assumed that a constant catch policy can be implemented perfectly. More realistic projections would acknowledge future data collection activities, adaptive control policies, and uncertainties associated with policy implementation. Our framework provides tools for addressing these issues in an integrated modeling environment.

Acknowledgments

We thank participants of the ICES Comprehensive Fishery Evaluation Working Group for their motivating discussions of the precautionary approach. Comments by John Hampton and an anonymous reviewer improved the manuscript.

References

Archibald, C.P., D. Fournier, and B.M. Leaman. 1983. Reconstruction of stock history and development of rehabilitation strategies for Pacific Ocean perch in Queen Charlotte Sound, Canada. N. Am. J. Fish. Manage. 3:283-294.

Caddy, J.F., and R. Mahon. 1995. Reference points for fisheries management. FAO Fish. Tech. Pap. No. 347.

Fournier, D., and C.P. Archibald. 1982. A general theory for analyzing catch at age data. Can. J. Fish. Aquat. Sci. 39:1195-1207.

Griewank, A., and G.F. Corliss (eds.). 1991. Automatic differentiation of algorithms: Theory, implementation, and application. Society for Industrial and Applied Mathematics, Philadelphia, PA.

Ianelli, J.N., and J. Heifetz. 1995. Decision analysis of alternative harvest policies for the Gulf of Alaska Pacific Ocean perch fishery. Fish. Res. 24:35-63.

Ketchen, K.S. 1981. Reconstruction of Pacific ocean perch *(Sebastes alutus)* stock history in Queen Charlotte Sound. Part II. Catch per unit of effort as a measure of abundance, 1959-79. Can. Manuscr. Rep. Fish. Aquat. Sci. 1599.

Kimura, D.K. 1981. Standardized measures of relative abundance based on modeling log(c.p.u.e.), and their application to Pacific Ocean perch *(Sebastes alutus).* J. Cons. Cons. Int. Explor. Mer 39:211-218.

Methot, R.D. 1989. Synthetic estimates of historical abundance and mortality for northern anchovy. Amer. Fish. Soc. Symp. 6:66-82.

Methot, R.D. 1990. Synthesis model: An adaptable framework for analysis of diverse stock assessment data. Int. North Pac. Fish. Comm. Bull. 50:259-277.

Otter Research Ltd. 1994. An introduction to AD Model Builder for use in nonlinear modeling and statistics. Otter Research Ltd., Nanaimo, B.C.

Richards, L.J., J.T. Schnute, and N. Olsen. 1997. Visualizing catch-age analysis: A case study. Can. J. Fish. Aquat. Sci. 54:1646-1658.

Schnute, J.T. 1994. A general framework for developing sequential fisheries models. Can. J. Fish. Aquat. Sci. 51:1676-1688.

Schnute, J.T., and L.J. Richards. 1995. The influence of error on population estimates from catch-age models. Can. J. Fish. Aquat. Sci. 52:2063-2077.

Schnute, J.T., and L.J. Richards. 1998. Analytical models for fishery reference points. Can. J. Fish. Aquat. Sci. 55:515-528.

Schnute, J.T., Richards, L.J., and N. Olsen. 1998. Statistics, software, and fish stock assessment. In: F. Funk, T.J. Quinn II, J. Heifetz, J.N. Ianelli, J.E. Powers, J.F. Schweigert, P.J. Sullivan, and C.-I. Zhang (eds.), Fishery stock assessment models. University of Alaska Sea Grant, AK-SG-98-01, Fairbanks (this volume).

United Nations. 1995. Agreement for the implementation of the provisions of the United Nations Convention on the Law of the Sea of 10 December 1982 relating to the conservation and management of straddling fish stocks and highly migratory fish stocks. United Nations General Assembly Doc. A/CONF.164/37.

Joint Time Series Analysis of Catch-at-Age and CPUE Data

Gudmundur Gudmundsson
Central Bank of Iceland, Reykjavik, Iceland

Abstract

Catch-at-age observations are subject to measurement errors and irregular variations of natural and fishing mortality rates. By modeling the fishing mortality rates as multivariate time series the stocks can be estimated from these data without reference to effort or catch per unit effort observations. Analysis of CPUE data without catch-at-age observations is less informative.

The paper describes the theoretical models and the application is demonstrated by actual examples. The estimation includes a likelihood function so that appropriate weighting of respective data sets is implicit.

Relative measurement errors in CPUE data from research vessel surveys are commonly much higher than for catches-at-age. Nevertheless, joint analysis with such data can reduce uncertainty in the estimation of stocks and fishing mortality rates. Changes in catchability in CPUE observations can be modeled and estimated and do not upset estimation of other quantities.

Estimation of constant rates of natural mortality or unreported catches is very inaccurate. The uncertainty about actual unobserved mortality amounts to a considerable uncertainty about stock size. Variations in unreported catches, different from variations in observed catches, can to some extent be estimated.

Introduction

In the analyses presented here fishing mortality rates are modeled as time series. The paper is concerned with the application of the method, with particular attention to detection and estimation of systematic errors in either catch-at-age or catch per unit effort (CPUE) data. A full description of the models will be presented, but not details of the estimation procedure. The basic aspects of the statistical methodology are described by

Harvey (1989) and modifications required for the present problem by Gudmundsson (1994).

An infinite number of solutions fit the basic equations of catch-at-age analysis exactly. In Virtual Population Analysis a selection from these solutions is made by finding the solution with best fit to effort or CPUE observations. In the present approach discrimination between different patterns of fishing mortality rates is achieved by requiring that they be generated by a time series model, determined by a small number of parameters. When the basic equations are combined with the time series model, catches are predictable and model parameters are estimated by the likelihood function of the catch prediction errors. The time series models do not determine the actual values of the series; stocks and fishing mortality rates are calculated by the Kalman filter according to the models and the observations.

This methodology enables fish stock assessment from catch-at-age data without introducing other observations. The quality of the estimation depends on the amount and accuracy of the observations and how well the actual variations in the fishing mortality rates can be represented by the model. The method can be extended to include CPUE or other observations which depend upon the unobserved stock values and thus contain information about them. Comparisons of results from time series analysis and other methods in fish stock assessment were presented by ICES (1991 and 1993).

CPUE data are often seriously deficient. The effect of technical change and regulations upon the effort of commercial fleets is difficult to assess and environmental changes, affecting catchability, can distort survey results. In the time series methodology permanent changes in catchability in the CPUE are modeled and estimated.

Unreported catches and discards are regarded as a major problem in some fisheries. We shall investigate possibilities of learning something about this by joint analysis of reported catches and CPUE data. From the point of view of analyzing such data this mortality is analogous to unobserved natural mortality.

All applications presented here are based on annual observations of catches at age of Icelandic cod and CPUE from a bottom trawl survey. The data were collected by the Marine Research Institution in Reykjavik (1997) and are reproduced in Tables 1 and 2. Schopka (1994) and Stefansson (1992) describe recent research on this stock.

Models

The time series analysis is based on the usual expressions of the relationship between catches, stocks and mortality rates:

$$N_{at} = N_{a-1,t-1}e^{-Z_{a-1,t-1}}, \tag{1}$$

Table 1. Catch-at-age of Icelandic cod, thousands of fish.

	Age							
Year	4	5	6	7	8	9	10	11
1984	31,553	19,420	15,326	8,082	7,336	2,680	512	538
1985	24,552	35,392	18,267	8,711	4,201	2,264	1,063	217
1986	20,330	26,644	30,839	11,413	4,441	1,771	805	392
1987	62,130	27,192	15,127	15,695	4,159	1,463	592	253
1988	39,323	55,895	18,663	6,399	5,877	1,345	455	305
1989	27,983	50,059	31,455	6,010	1,915	881	225	107
1990	12,313	27,179	44,534	17,037	2,573	609	322	118
1991	25,131	15,491	21,514	25,038	6,364	903	243	125
1992	21,708	26,524	11,413	10,073	8,304	2,006	257	46
1993	33,078	15,195	13,281	3,583	2,785	2,707	1,181	180
1994	24,136	19,661	6,966	4,392	1,257	599	508	283
1995	9,102	16,827	13,064	4,115	1,596	313	184	156
1996	14,874	7,366	12,297	9,422	2,155	836	208	76

Table 2. Indices of catch per unit effort.

	Age					
Year	4	5	6	7	8	9
1985	57,217	53,064	16,085	8,645	3,175	1,630
1986	24,405	21,293	21,342	6,782	2,217	745
1987	82,001	16,926	9,132	9,862	2,150	553
1988	101,089	58,467	7,236	4,099	3,564	429
1989	67,884	60,410	28,395	3,096	1,115	774
1990	10,059	22,290	29,004	12,412	1,130	334
1991	25,418	10,776	13,722	17,950	3,193	389
1992	15,037	12,072	3,824	4,768	4,605	734
1993	36,980	11,767	7,149	1,666	1,193	689
1994	23,676	18,092	3,881	2,445	456	291
1995	5,949	17,694	12,710	2,815	1,071	145
1996	23,850	8,501	11,113	12,008	2,465	549
1996	53,097	27,249	7,437	7,680	5,119	347

$$C_{at} = \frac{F_{at}}{Z_{at}}\left(1-e^{-Z_{at}}\right)N_{at} + \varepsilon_{at}, \tag{2}$$

where

N_{at} = number of fish of age a at the beginning of year t,
C_{at} = observed number of caught fish of age a in year t,
F_{at} = fishing mortality rate (corresponding to the observed catches),
Z_{at} = total mortality rate,
ε_{at} = measurement errors.

The total mortality rate is defined as

$$Z_{at} = F_{at} + H_{at}$$

where the hidden mortality rate, H_{at}, accounts for natural mortality and unreported commercial catches and discards.

We let $a = 1$ for the youngest age included in the analysis and $t = 1$ for the first year and call the number of ages and years included A and T respectively.

All measurement errors and residuals in our models are assumed uncorrelated between different years and with zero means. The covariance matrix of ε_{at} in year t is $\sigma^2\Sigma_t$ where Σ_t is a given $A \times A$ matrix.

Equation (1) is a multivariate time series model, connecting stocks at ages ≥ 2 with the previous year's stocks and mortality rates. For the youngest fish we use the model

$$N_{1t} = N_0 + \theta_1 r_t + \delta_{0t}, \tag{3}$$

where N_0 is a constant value, r_t a recruitment index, θ_1 a parameter and δ_{0t} residuals with variance σ_0^2.

The main purpose of the present analysis is to estimate stocks and fishing mortality rates during the period of observations up to the beginning of year T+1. For that purpose we need a time series model of fishing mortality rates with a manageable number of parameters, sufficiently flexible to represent the actual variations in F_{at}. We use the following model:

$$\log F_{at} = U_{a,t-1} + V_{t-1} + \delta_{1at} + \delta_{2t}.$$

Permanent variations in $\log F_{at}$ are generated by a multivariate random walk model where joint variations at all ages are represented by

$$V_t = V_{t-1} + b + \delta_{4t}$$

and selectivity by

$$U_{a,t} = U_{a,t-1} + \delta_{3at}, \; a \leq a_m < A,$$

$$U_{at} = U_{a_m t},\ a > a_m,$$

$$\text{with the constraint that } \sum_{a=1}^{a_m} \delta_{3at} = 0.$$

where δ_{1at}, δ_{2t}, δ_{3at} and δ_{4t} are residuals. The variances of δ_{2t} and δ_{4t} are σ_2^2 and σ_4^2 respectively and the covariance matrices of δ_{1at} and δ_{3at} are $\sigma_1^2\mathbf{G_1}$ and $\sigma_3^2\mathbf{G_3}$ where $\mathbf{G_1}$ and $\mathbf{G_3}$ are predetermined matrices. The parameters in this model are $\sigma_1 - \sigma_4$ and b. The value of a_m is predetermined, but the likelihood function can be applied to choose between values.

The residuals denoted by δ_{1at} represent transitory variations of individual values of $\log F_{at}$ in year t, and δ_{2t} similar joint variations at all ages in that year. These variations do not affect the Fs in subsequent years. Permanent variations in selectivity are produced by the residuals δ_{3at} and joint permanent variations in $\log F_{at}$ at all ages by a constant annual increase b and by δ_{4t}. If $\sigma_3 = 0$ the fishing mortality rates follow the separable model apart from the random variations denoted by δ_{1at}. If $b = \sigma_3 = \sigma_4 = 0$ the values of $\log F_{at}$ fluctuate around a fixed mean at each age.

The first difference of $\log F_{at}$,

$$\log F_{at} - \log F_{a,t-1} = \delta_{1at} - \delta_{1a,t-1} + \delta_{2t} - \delta_{2,t-1} + \delta_{3a,t-1} + \delta_{4,t-1} + b,$$

is a stationary series, with zero mean unless there is trend in F, and variance independent of time. If we know something, or discover during the analysis, that strongly contradicts this assumption the model must be adjusted accordingly.

The Kalman filter needs initial values of the fishing mortality rates in the first year. They are determined as follows: The initial $\log F_{1,1}$ is estimated separately. The initial values of $\log F_{a,1}$ for a from 2 to a_m are determined by a third order polynomial in a, with the constraint that the first derivative at a_m is zero. $\operatorname{Log} F_{a,1} = \log F_{a_m,1}$ for $a > a_m$. This specification is sufficiently flexible to approximate a large variety of variations in selectivity with age.

The remaining fishing mortality rates are calculated by the Kalman filter from the observations in accordance with equations (1) and (2) and the time series model.

Ideally we should like to estimate a similar model for the rate of hidden mortality. But as there are no observations of unreported catches or natural deaths, this is not practically possible and a simpler deterministic model will be adopted:

$$\log H_{at} = \log M_{at} + D_{at}\left[\alpha_1 + \alpha_2\left(t - \frac{T+1}{2}\right) + \alpha_3 \cos\left(\frac{2\pi t}{T}\right) + \alpha_4 \sin\left(\frac{2\pi t}{T}\right)\right]. \quad (4)$$

In this model M_{at} is a given function and could represent the rate of natural mortality which is often regarded as known. D_{at} accounts for presumed variations in mortality with age in the hidden catches and is defined a priori. For discards confined to the smallest fish a suitable definition could be $D_{1t} = 1$ and all other values zero.

Let us call CPUE observations, carried out at time τ each year, $CPUE_{at}$. Their relationship with the variables above is given by

$$CPUE_{at} = \psi_t \Phi_a e^{-\tau Z_{at}} N_{at} + \varepsilon_{CPUE,at}. \tag{5}$$

The residuals $\varepsilon_{CPUE,\,at}$ have covariance matrix $\sigma_{CPUE}{}^2 \Sigma_{CPUE,t}$. They represent measurement errors and irregular transitory variations in catchability. LogΦ_a is estimated in a similar manner as the initial values of logF_{at}. For joint variations of catchability with time we specify the model

$$\log\psi_t = \xi_{t-1} + \delta_{5t},$$

$$\xi_t = \xi_{t-1} + d + \delta_{6t}.$$

In this model permanent variations are produced by a constant annual change d and as random walk by the residuals δ_{6t} with variance σ_6^2. Joint transitory variations are represented by δ_{5t} with variance σ_5^2.

Glossary of Terms in the Time Series Models

ε_{at} = measurement errors of catch-at-age, covariance matrix $\sigma^2\Sigma_t$.
$\varepsilon_{CPUE,at}$ = measurement errors of CPUE observations, covariance matrix $\sigma_{CPUE}{}^2\Sigma_{CPUE,t}$.
r_t = recruitment index
δ_{0t} = residuals in recruitment equation, variance σ_0^2.
δ_{1at} = transitory variations in logF_{at}, covariance matrix $\sigma_1^2\mathbf{G_1}$.
δ_{2t} = joint transitory variations in logF_{at}, variance σ_2^2.
δ_{3at} = permanent variations in selectivity in logF_{at}, covariance matrix $\sigma_3^2\mathbf{G_3}$.
δ_{4t} = joint permanent variations in logF_{at}, variance σ_4^2.
b = annual linear increase in logF_{at}.
$\psi_t\phi_a$ = catchability in CPUE data

Applications

We present various examples of the application of these models. They are all based on two sets of annual observations for Icelandic cod; catch-at-age data from 1984 to 1996 and CPUE values from a bottom trawl survey from 1985 to 1997. The data are reproduced in Tables 1 and 2. We start by defining various aspects of the covariance structure of measurement errors and residuals that are not specified by estimated parameters.

Specifications

The matrices Σ_t, $\Sigma_{CPUE,t}$, $\mathbf{G_1}$ and $\mathbf{G_3}$ are first defined a priori and subsequently modified in accordance with examination of residuals. The magnitude of measurement errors obviously varies with the observed values of C_{at} and $CPUE_{at}$. This is accounted for by making the diagonal elements of Σ_t and $\Sigma_{CPUE,t}$ proportional to the squared value of a calculated average for respective age and cohort. Let us call the diagonal elements $k_a{}^2\tilde{C}_{at}^2$. The normal value of k_a is 1, but bigger values can be inserted if examination of the data or residuals indicates that measurements at respective age are relatively less accurate. In the present analysis k_a is 1.5 for 4-year-old fish and 2.5 and 3 for ages 10 and 11 years respectively.

In principle some parameters in the specification of the matrices Σ_t, $\Sigma_{CPUE,t}$, $\mathbf{G_1}$ and $\mathbf{G_3}$ could be estimated. In practice we cannot even estimate both σ and σ_1 and the ratio between them is fixed at 1. By this we ensure that neither measurement errors of catch-at-age data nor irregular transitory variation of fishing mortality rates are ignored, but some under- or overestimation is inevitable.

The observed values of C_{at} are obtained from reports of total catches and age and length measurements of samples. Errors in assessing the relative proportion of age groups in the total catch are negatively correlated and we assume a correlation of –0.2 between the errors of adjacent ages. Other correlations are defined as 0.

The random variations of $\log F_{at}$, denoted by δ_{at}, are defined as independent and with the same variance so that the diagonal elements are 1, except for the 4-year-old fish. The catchability of the young fish is sensible to variations in growth so that its diagonal element was defined as 2.25. Finally the matrix $\mathbf{G_3}$ is defined as diagonal with unit values on the diagonal except for correlations of 0.2 between the 4-year-olds and 6- to 8-year-olds

Outliers can have a strong influence on estimates of parameters and unobserved series. The best way to deal with outliers depends on whether they represent correct values that do not seem to fit the estimated model, or measurement errors. In the first case the best course might be to modify the model so that the values do not appear as outliers. In the present case I consider that the most practical approach is to assume that they are measurement errors and reduce their influence accordingly. This is easily achieved in Kalman filter calculations by increasing the variance of respective value of ε_{at} or $\varepsilon_{CPUE,at}$. This was done here with ε_{at} for 11-year-olds in 1988 and 1993, 9-year-olds in 1989, and 10-year-olds in 1993. The variance of $\varepsilon_{CPUE,at}$ was increased for 8 years in 1996 and 9 years in 1997.

In practice the specifications described here are carried out in two or more steps. An initial choice is made, based on visual examination of the data and possibly some other knowledge. Examination of residuals and statistics calculated from them usually indicates that some modifications are needed. (Gudmundsson 1994).

There was a distinct change in exploitation policy beginning with a substantial drop in catches in 1994. This is modeled by subtracting an estimated constant value from all $\log F_{at}$ after 1993. (If unaccounted for the change emerges clearly in the residual analysis.)

Mature cod migrates from Greenland to Icelandic waters. For most of the years examined here this was negligible because stocks in Greenland waters were small. An exception is 6-year-old cod in 1990 and this was estimated.

Analysis of the Original Observations

Joint analysis was carried out with the full model for both data sets. CPUE values for 3-year-old fish were used as a recruitment index. Hidden mortality was fixed at 0.2 year^{-1} for all ages.

The estimated value for the trend parameter b in the model for $\log F$ was 0.029. This corresponds to an increase of 0.35 in $\log F$ over the estimation period, which is far from negligible. But estimation with b fixed at zero produced practically the same estimates of stocks and fishing mortality rates and a negligible decrease in the likelihood function. There is no reason to expect a linear trend in $\log F$ for Icelandic cod during this period so that we leave this parameter out of the analyses presented here. The main reason why a trend parameter as large as this has negligible effect upon results and goodness of fit is that the Kalman filter can produce similar variations in $\log F$ by means of the random walk variations δ_{4t}.

In the model for CPUE, permanent variations are represented by the trend parameter d and the residual δ_{6t} with variance σ_6^2. The estimated value of σ_6 was practically zero, but the trend parameter d was –0.022. This implies a decrease of 0.26 in $\log\psi_t$ from 1985-1997. However, the values of ψ_t obtained by these model parameters are practically the same as those obtained when both are fixed at zero and the logarithm of the likelihood function only decreases by 0.3. The bottom trawl survey is designed and conducted with the aim of avoiding permanent variations in catchability. As there is no evidence against the hypothesis that this has been achieved it is appropriate to use estimates where σ_6 and d are fixed at 0. The estimated time series parameters when b, σ_6 and d were thus fixed at 0 are presented in the first row of Table 3.

The effect of the policy change in 1994 on $\log F$ was estimated as –0.48. Migration of 6-year-old fish from Greenland in 1990 was estimated as 27 million.

Estimation of standard deviations of parameters by means of the Hessian matrix is often unreliable in this large nonlinear problem, with several parameters close to the boundary of permissible values and these estimates are not reported here. The parameters σ and σ_1 are well determined under the restriction that they have the same value. The values of σ_2 and σ_3 are not significantly different from zero.

When CPUE observations are from surveys, catching negligible proportions of the total stock, measurement errors and irregular variations in

Table 3. Parameters in time series models.

	C	$CPUE$	r	H	$\sigma = \sigma_1$	σ_2	σ_3	σ_4	σ_{CPUE}	σ_5	σ_6	d
1	C	$CPUE$	$CPUE_{0,t-1}$	0.2	0.075	0.001	0.001	0.121	0.228	0.289		
2	C		t	0.2	0.071	0.007	0.022	0.133				
3		$CPUE$	$CPUE_{0,t-1}$	0.2	0.143	0.002	0.043	0.001	0.143	0.142		
4	C	$CPUE$	t	0.2	0.073	0.000	0.000	0.134	0.212	0.235	0.092	0.005
5	C	$h^1 CPUE$	t	0.2	0.072	0.000	0.000	0.132	0.214	0.139	0.172	−0.016
6	C	$h^2 CPUE$	t	0.2	0.072	0.001	0.001	0.135	0.215	0.123	0.238	0.044
7	C	$CPUE$	$CPUE_{0,t-1}$	H	0.074	0.001	0.000	0.126	0.222	0.281		
8	$G^1 C$	$CPUE$	$CPUE_{0,t-1}$	0.2	0.072	0.000	0.001	0.166	0.236	0.216		
9	$G^1 C$	$CPUE$	$CPUE_{0,t-1}$	H	0.072	0.000	0.001	0.152	0.227	0.272		
10	$G^2 C$	$CPUE$	$CPUE_{0,t-1}$	0.2	0.069	0.041	0.041	0.090	0.228	0.416		
11	$G^2 C$	$CPUE$	$CPUE_{0,t-1}$	H	0.066	0.000	0.060	0.163	0.235	0.321		
12	$G^3 C$	$CPUE$	$CPUE_{0,t-1}$	0.2	0.072	0.000	0.000	0.147	0.237	0.238		
13	$G^3 C$	$CPUE$	$CPUE_{0,t-1}$	H	0.074	0.001	0.001	0.148	0.229	0.215		

catchability are indistinguishable. However, in the present case it is obvious that measurement errors did not produce the large values of σ_{CPUE} and σ_5. The magnitude of the joint variations in catchability at all ages from 4 to 9 years is remarkable and indicates that environmental factors, affecting a substantial proportion of the area covered by the survey, have a major influence on the observed CPUE.

The estimated stocks and mortality rates are presented in Tables 4 and 5, together with the Kalman filter estimates of the standard deviations of the estimated values in the last year. These standard deviations provide a good assessment of the uncertainty caused by measurement errors and other random elements in the data. But they do not include errors caused by wrong model specification or choice of M. (Standard deviations in previous years are also calculated. They depend critically upon the ratio between σ and σ_1 which is arbitrarily fixed so that I do not present them here. But regardless of this ratio they are considerably lower than the values in the last year.)

With the time series method stocks and fishing mortality rates can be estimated from catch-at-age data without using any effort or CPUE observations. In the analysis above we used the survey values as $CPUE_{at}$ in equation (5) and also for the recruitment index r_t in equation (3). If we leave out the recruitment index, equation (3) entails fluctuations of recruitment around a constant value. Systematic variations in actual recruitment then produce bias in the estimation of stocks and mortality rates. In the present case stocks and recruitment are higher in the first half of the period than

Table 4. Estimated stock size of Icelandic cod, thousands of fish.

	Age							
Year	4	5	6	7	8	9	10	11
1984	184,206	77,606	40,313	18,574	13,535	5,271	1,094	1,072
1985	105,798	122,595	44,352	19,476	7,945	4,601	1,936	408
1986	109,932	64,938	67,489	20,743	7,914	2,944	1,636	687
1987	252,308	69,867	29,968	27,283	6,980	2,416	904	496
1988	231,182	151,214	33,530	11,596	8,937	2,051	702	263
1989	135,969	147,292	73,211	12,166	3,507	2,153	508	182
1990	64,265	84,317	103,234	31,912	4,517	1,117	687	165
1991	102,389	40,963	43,447	44,597	11,490	1,480	364	222
1992	74,797	60,818	19,697	16,398	14,445	3,458	427	100
1993	132,149	42,067	26,695	6,701	4,720	4,208	1,011	119
1994	106,498	78,940	20,268	10,110	2,178	1,274	1,088	271
1995	46,643	67,383	45,414	10,130	4,288	806	486	418
1996	95,717	30,592	39,944	24,443	4,831	1,915	359	213
1997	132,145	65,131	18,398	21,298	11,767	2,076	822	154
	(21,000)	(10,000)	(2,600)	(3,100)	(1,600)	(300)	(140)	(30)

Standard deviations in parentheses, 1997.

Table 5. Estimated fishing mortality rates of Icelandic cod.

	Age							
Year	4	5	6	7	8	9	10	11
1984	0.21	0.35	0.53	0.65	0.88	0.80	0.78	0.78
1985	0.28	0.40	0.56	0.70	0.79	0.83	0.84	0.83
1986	0.26	0.53	0.70	0.88	0.98	0.98	0.98	0.99
1987	0.31	0.53	0.74	0.92	1.02	1.03	1.03	1.03
1988	0.25	0.53	0.81	0.98	1.21	1.19	1.15	1.14
1989	0.27	0.46	0.63	0.78	0.94	0.94	0.92	0.93
1990	0.25	0.46	0.64	0.82	0.92	0.92	0.93	0.94
1991	0.32	0.53	0.76	0.92	1.00	1.04	1.08	1.06
1992	0.38	0.62	0.88	1.04	1.03	1.02	1.07	1.11
1993	0.32	0.53	0.77	0.92	1.11	1.15	1.11	1.13
1994	0.26	0.35	0.49	0.66	0.79	0.76	0.76	0.76
1995	0.22	0.32	0.41	0.53	0.60	0.60	0.62	0.62
1996	0.19	0.31	0.43	0.53	0.65	0.65	0.65	0.63
	(0.03)	(0.04)	(0.05)	(0.07)	(0.09)	(0.10)	(0.10)	(0.09)

Standard deviations in parentheses, 1996.

in the last years so we include t as a recruitment index in order to reduce bias from this.

The estimated time series parameters are presented in row 2 of Table 3. The estimated stocks and fishing mortality rates are practically the same as those obtained in the joint analysis with $CPUE_{at}$ but the standard deviations of stocks and fishing mortality rates obtained from the Kalman filter are higher; about 0.20 for $\log F_{at}$ in the last year.

The close agreement between the estimates could be a coincidence. "Retrospective analysis" is a useful and widely applied method to assess the accuracy of stocks and fishing mortality rates, estimated from catch-at-age data. Estimates in years $T-1$, $T-2,\ldots$, obtained from the analysis of the full data set, are compared with the last year's values when only $T-1$, $T-2,\ldots$ years are included in the analysis. The total biomass for ages 4-11 and average F for ages 5-9 obtained in retrospective analyses with the full data set and also with only catch-at-age data are presented in Fig. 1.

When the last observations of C_{at} are in 1995 or 1996 the results are practically the same for the full data set and catch-at-age data only. There are no major changes in the estimated selection patterns so that these summary measures are representative for the differences between the results. The biggest difference between the last year's estimate and the estimate obtained from the longest data sets is in 1994 when only catch-at-age data are included. This was the first year after the drop in fishing mortality rates, estimated by a separate parameter, which was overestimated in this case. But these results are in agreement with my experience that it is usually possible to obtain useful estimates of stocks and fishing mortality rates from catch-at-age data without effort or CPUE observations.

Let us now investigate what can be derived from CPUE data without a joint analysis with catch-at-age data. We can regard the series $CPUE_{at}$ for $t = 1,2,\ldots,T$ as an index for stock of age a at the time of the survey. However, this is not an optimal use of the data because it neither uses the relationships between stock values, given by equation (1), nor correlations between variations in catchability at different ages within the same year.

Fishing mortality rates do not appear separated from total mortality rates in equations (1) and (5). Nor can variation of total mortality rates with age be distinguished from the variations in selectivity, denoted by Φ_a. For analysis of CPUE data alone it is convenient to do a slight reformulation of equations (1) and (5). We introduce a calendar where the CPUE observations are made at the beginning of the year and define

$$n_{at} = n_{a-1,t-1}e^{-W_{a-1,t-1}},$$

$$CPUE_{at} = \psi_t\, n_{at} + \varepsilon_{CPUE,at},$$

$$\text{where } n_{at} = \Phi_a N_{at}$$

$$\text{and } W_{a-1,t-1} = Z_{a-1,t-1} - \log(\Phi_a/\Phi_{a-1}).$$

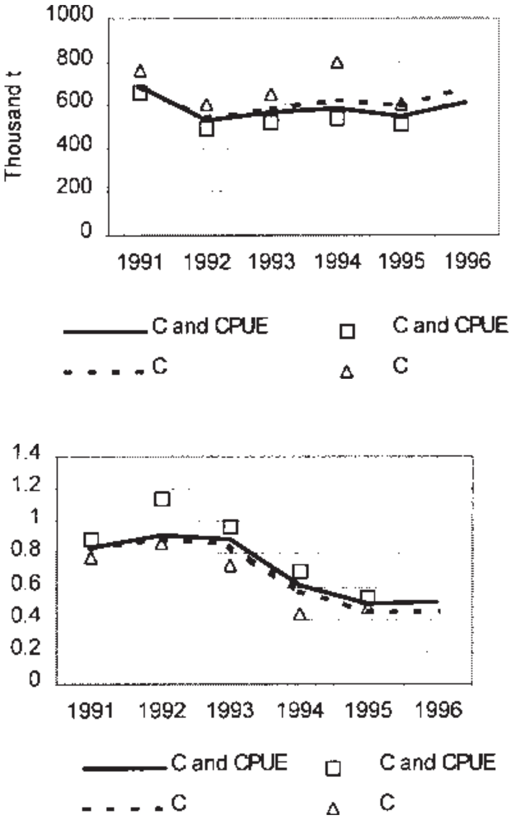


Figure 1. Retrospective analyses. Total biomass and average F 4-9 years. Full and broken lines show the results from the longest data sets. Squares and triangles show the results from the last years' estimates.

We use the same time series model for W_{at} as for F_{at} and assume that the CPUE data are free from systematic variations in catchability, i.e., $\sigma_6 = d = 0$. The estimated parameters are presented in row 3 of Table 3. The value of σ_1 is defined to be equal to σ_{CPUE}.

The random walk variations produce most of the estimated variations of F when they are estimated from catch-at-age data. But here σ_4 is practically zero and so is the parameter of joint transitory variations, σ_2. Apart from the drop in 1994, estimation based only on the CPUE data fails to reveal the variations in fishing mortality rates with time. Biomass indices cannot be calculated without knowing Φ_a.

Systematic Errors in CPUE Data

In virtual population analysis it is customary to assume that CPUE observations are proportional to stock size at respective age and time. Our model allows for systematic departure from this assumption by trend and random walk variations in catchability, but there was no evidence of such variations in the data. Let us now construct two new CPUE data sets by multiplying the observed CPUE values by the functions h_t^1 or h_t^2, presented in Fig. 2. (The CPUE values are used unchanged from 1985 to 1991 and then decrease/increase linearly to half/double observed values.) Joint analysis of catch-at-age observations and these data sets was carried out and compared with results from the unadulterated CPUE values. A linear trend is included in the recruitment equation (3), but no recruitment index. The estimated time series parameters are presented in rows 5 and 6 of Table 3. Row 4 shows similar estimation with the observed CPUE values. The effect of changing the CPUE data hardly affects the parameters of $\log F_{at}$ at all, but the magnitudes of random walk and trend parameters of $\log \psi_t$ are increased.

The estimated values of $\log \psi_t$ are presented in Fig. 3. With the altered CPUE data sets the time series models produce large variations in the second half of the period in accordance with the functions h_t^1 and h_t^2.

As the variations in catchability, produced by h^1 and h^2, are successfully estimated they have little effect upon the estimation of stocks and fishing mortality rates. With the results obtained by including the observed CPUE values as reference, the biomass is underestimated by 4% at the beginning of 1997 with $h^1 CPUE$ and overestimated by 10% by $h^2 CPUE$. The results obtained by $h^i CPUE$ converge rapidly towards the reference values as we go back in years.

According to the standard deviations, produced by the Kalman filter, stocks and fishing mortality rates are somewhat better estimated by including the deficient CPUE data than by only using the catch-at-age observations. But the estimated standard deviations are larger with the adulterated CPUE data sets than with the observed CPUE values because the random walk parameter, σ_6, is bigger.

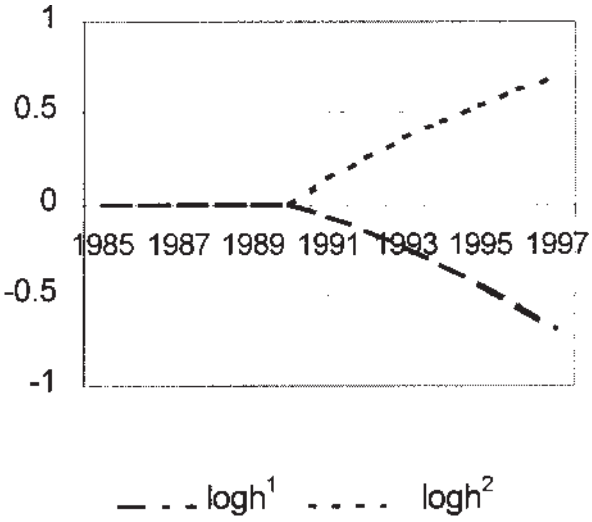


Figure 2. *Log values of two functions, used to produce systematic variations of catchability in CPUE data.*

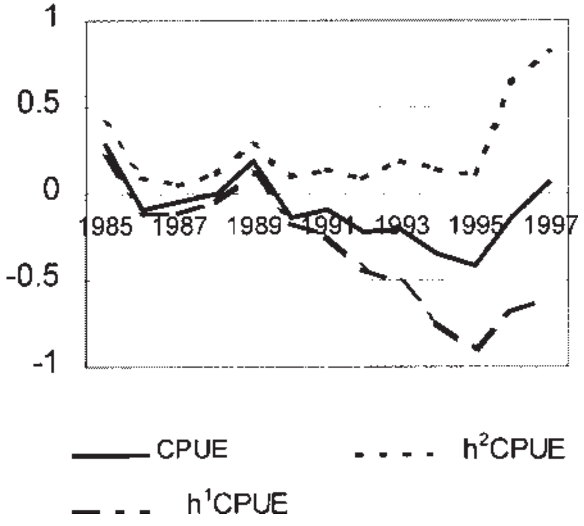


Figure 3. *Estimated variations in catchability, logψ_τ, from observed values and two altered data sets.*

Hidden Mortality

In principle the rate of natural mortality, which was fixed as 0.2 in the estimates reported so far, is a parameter that can be estimated. In practice its value has too little effect upon the likelihood function for joint estimation with the remaining parameters of the model, but the maximum likelihood was calculated for or a couple of fixed values of M:

M	0.05	0.10	0.20	0.30	0.40	0.50
$\text{Log}L(M) - \log L(0.2)$	0.00	0.02	0.00	–0.45	–1.40	–2.75

As there is negligible evidence against the conventional value of 0.2 year^{-1} we have kept it when a fixed value is used. However, values of M that provide practically the same goodness of fit for these observations entail substantial differences in estimated stocks and fishing mortality rates. The biomass estimates for three values of M are presented in Fig. 4.

Estimation of variations in natural or hidden mortality rates with age is also practically impossible; they cannot be distinguished from variations of F_{at} with age and the selectivity of the survey.

Unreported catches are a major problem for fish stock assessment in some fisheries. Estimation of unreported catches is analogous to natural mortality so that we have established that the possibilities of estimating a constant rate of unreported fishing mortality are poor.

All information about fishing mortality rates in catch-at-age and CPUE data is entailed in ratios between observations. Let us suppose that natural mortality is known and the unreported catches are a fixed proportion, k, of the actual catches so that

$$H_{at} = M + kF_{at}.$$

An estimation of the fishing mortality rates, assuming that $H_{at} = M$, would result in the sum of the reported and unreported fishing mortality rates, i.e. $(1+k)F_{at}$, and the stock estimates would underestimate the actual stocks as $N_{at}/(1+k)$.

Let us now investigate the possibilities of estimating stocks and mortality rates in the presence of substantial variations of hidden mortality rates with time, different from the variations of the reported catches. For this purpose we use the observed CPUE values unchanged but induce time-varying hidden mortality by using the catch-at-age values $g_t^i C_{at}$ where g_t^i are functions presented in Fig. 5.

The actual hidden mortality rate is

$$H_{at} = M_{at} + (1 - g_t^i)F_{at}.$$

Stocks and fishing mortality rates were estimated for the data sets $g_t^i C_{at}$ and $CPUE_{at}$, both by using the constant value 0.2 for hidden mortali-

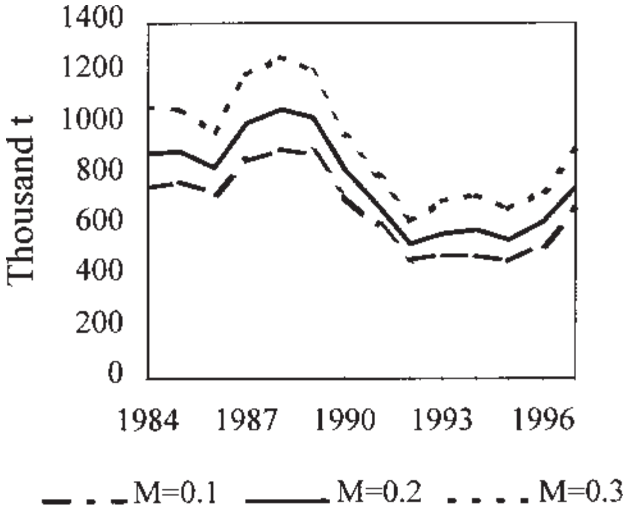


Figure 4. Biomass estimated with three values of M.

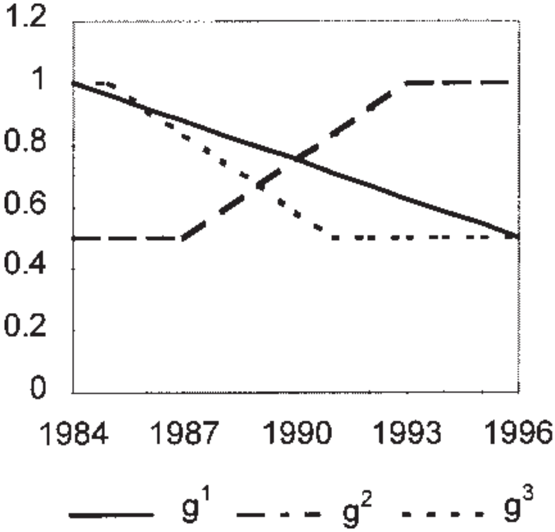


Figure 5. Three functions used to induce hidden mortality.

ty, and by the model of equation (4) with $M_{at} = 0.2$. In order to take into account that catchability is lower for the youngest fish we used:

$$D_{1t} = ⅓ \text{ and } D_{2t} = ⅔ \text{ when } \log H_{at} > \log M_{at},$$

$$D_{at} = 1 \text{ when } \log H_{at} \leq \log M_{at} \text{ and for all } a > 2.$$

Estimation of variations in hidden mortality is based on discrepancy between the observed catch-at-age data and the CPUE data. In order to produce discrepancy, requiring variations in H_{at}, we must assume that the catchability in the CPUE data set is free from permanent variations and fix the parameters σ_6 and d as zero; otherwise the discrepancy might be accommodated by random walk or linear trend in $\log\psi_t$.

Our model for hidden mortality rates is a misspecification of the actual hidden mortality produced by multiplying C_{at} by g_t^i. This is in accordance with reality; we would rarely be in a position to specify an almost correct parametric model of the rates of hidden fishing mortality. The improvement in goodness of fit achieved by estimating the model can be judged by comparing the likelihood functions. The results are presented below:

Catch-at-age data	C_{at}	$g_t^iC_{at}$	$G_t^2C_{at}$	$G_t^3C_{at}$
$\log L(H) - \log L(0.2)$	1.89	3.61	5.57	7.76

The model of H_{at} contains 4 estimated parameters. But some of these are always superfluous so that the improvement is significant for all $g_t^iC_{at}$ and we could in fact have detected somewhat smaller unreported catches of the kind presented by $(1–g_t^l)C_{at}$. For comparison we also estimated the model for hidden mortality with the unadulterated catch-at-age data, but here the increase in likelihood, 1.89, is not significant.

The data sets C_{at} and $g_t^iC_{at}$ all represent the same biomass and total mortality rates. The biomass estimates are presented in Fig. 6.

By estimating hidden mortality, instead of assuming a constant value of 0.2, the overestimation of the fishing mortality rates, associated with the observed catches, was reduced but not eliminated in each of our examples. Estimated values of H_{at} were too low. However, the pattern and magnitude of stocks and total mortality rates were much better estimated by including the model for hidden mortality than with a fixed value.

Conclusions

The results confirm previous experience that stocks and fishing mortality rates can be estimated from catch-at-age data by time series models, without reference to any effort or CPUE data. On the other hand, analysis of CPUE data without catch-at-age data could only provide an index of stock size at each age.

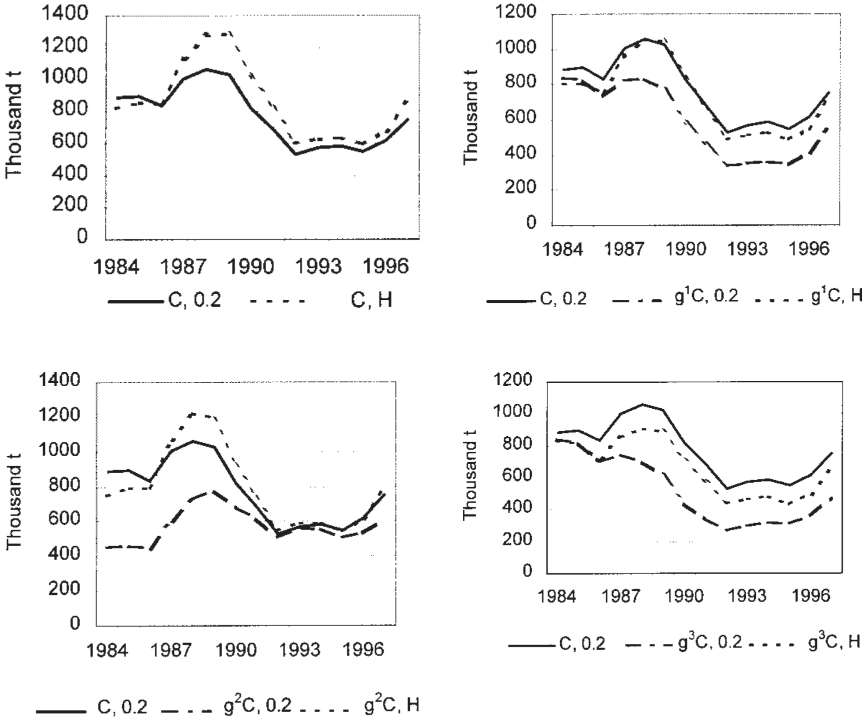


Figure 6. *Estimates with hidden mortality. Unbroken line shows estimate with the original data set and a fixed hidden mortality 0.2. In top left the broken line shows an estimate with the original data set and model H for hidden mortality. In the remaining figures broken lines show estimates from data with hidden mortality, estimated with fixed hidden mortality 0.2 and also with the model H.*

Joint analysis was carried out of catch-at-age and CPUE data with large systematic variations of catchability in the last years. The catchability variations were successfully detected and estimated and thus did not greatly affect the estimation of stocks and mortality rates.

Estimation of a constant rate of natural mortality is very inaccurate. The uncertainty about the actual level of natural mortality amounts to a considerable uncertainty about stock size.

In a joint analysis with CPUE data, free from systematic variations in catchability, substantial variations in unreported catches could be detected when they differed from variations in reported catches. The models that can be estimated for unreported catches are less flexible than the time series models for observed catches at age. This implies that unreported catches are likely to be underestimated in relation to reported catches. Estimation of the unreported catches improved the estimation of stocks and total mortality rates in the present experiments, but we have no further experience of this application of the time series methodology.

References

Gudmundsson, G. 1994. Time series analysis of catch-at-age observations. Appl. Statist. 43(1):117-126.

Harvey, A.C. 1989. Forecasting structural time series models and the Kalman filter. Cambridge University Press, Cambridge, U.K.

ICES. 1991. Report of the Working Group on Methods of Fish Stock Assessment. ICES Coop. Res. Rep. 199.

ICES. 1993. Report of the Working Group on Methods of Fish Stock Assessment. ICES C.M. 1993/Assess:12.

Marine Research Institute. 1997. State of marine stocks in icelandic waters 1996/1997: Prospects for the quota year 1997/98. Marine Research Institute, Reykjavik.

Marine Research Institute. 1997. Icelandic Groundfish Survey 1997. Marine Research Institute, Reykjavik.

Schopka, S.A. 1994. Fluctuations in the cod stock of Iceland during the twentieth century in relation to changes in the fisheries and environment. ICES Mar. Sci. Symp. 198:175-193.

Stefansson, G. 1992. Notes on the stock dynamics of the Icelandic cod. ICES C.M. 1992/G:71.

Dynamics of a Migratory Fish Population with Applications to the Management of Sablefish

Jonathan Heifetz
National Marine Fisheries Service, Alaska Fisheries Science Center, Auke Bay Laboratory, Juneau, Alaska

Terrance J. Quinn II
University of Alaska Fairbanks, Juneau Center, School of Fisheries and Ocean Sciences, Juneau, Alaska

Abstract

The reproduction, mortality, and migration of an age-structured fish population can be incorporated into a matrix model. The population projection matrix for such a model describes the dynamics of age classes within regions and the migration among regions. Such a model provides a flexible framework to examine the dynamics of a migratory fish population. The model can be parameterized to include regional specificity in the stock-recruitment relationship as well as larval dispersal that is decoupled from local reproduction. The subject of oriented movement, where young fish tend to migrate in the opposite direction of older fish is examined within the context of the sablefish (*Anoplopoma fimbria*) fishery in the northeast Pacific Ocean. We show that fishing at a common rate among regions may be appropriate for the pooled population but may be detrimental to the population in a given region. As an alternative, region-specific fishing strategies can be devised to meet fishery management objectives such as maintenance of regional spawning potential.

Introduction

The processes of reproduction, survival, and migration of an age-structured population can be formulated within the basic framework of matrix models described in Caswell (1989) and Getz and Haight (1989). Such a model is a natural extension of the Leslie (1945) matrix population model. First described by Rogers (1966, 1985), the model has been used extensively to analyze regional aspects of human populations (Liaw 1980) and has occasionally been applied to plants, fish, and wildlife (e.g., Fujioka 1978, Lande 1988). The underlying feature of such models is that all demographic features of a population can be formulated within a projection matrix.

In this paper, we first present a general structure for modeling the dynamics of a migratory age-structured fish population subject to exploitation. We focus on showing how the model can be used to efficiently track and project changes in abundance over time and space. After presenting the basic equations that describe the general model structure, we apply the model to the sablefish *(Anoplopoma fimbria)* population in the northeast Pacific Ocean. In addition, we examine how the model can be used for management of the sablefish fishery. In particular we show how a regional per-recruit analysis can be undertaken. Detailed derivations of the mathematical equations and methods to examine the equilibrium, stability, and transient properties of the model are provided in Heifetz (1996). Applications of the model with hypothetical data are in Heifetz (1996) and in Quinn and Deriso (1998). For fisheries that occur in different geographic areas, are prosecuted by different user groups, and in some instances may cross international boundaries, knowledge of how a fishery in one area interacts with another may lead to improved management.

Model Specification

The General Age-by-Region Model

The stage-classified matrix model described in Caswell (1989) forms the basis for this model. We translate the formulation into that commonly used for age-structured fish populations (e.g., Quinn and Szarzi 1993). In our notation variables in bold denote matrices or vectors. Let

$N_{a,t,i}$ = abundance at age a, year t, and region i,
$\theta_{a,t}(i \rightarrow k)$ = migration proportion from region i to k at age a and year t,
K = number of regions,
A = oldest age,
$f_{a,i}$ = net fecundity at age a and region i,
M = instantaneous rate of natural mortality,
$s_{a,i}$ = fishery selectivity at age a and region i,
$F_{t,i}$ = instantaneous rate of full-recruitment fishing mortality (i.e., $s_{a,i} = 1$) in year t and region i,
$Z_{a,t,i}$ = $M + s_{a,i}F_{t,i}$ = total instantaneous mortality at age a, year t, and region i,
$S_{a,t,i}$ = $\exp(-Z_{a,t,i})$ = survival at age a, year t, and region i, and
r = recruitment age and usually first appearance of fish in the database.

Note that $\sum_{k=1}^{K}\theta_{a,t}(i \to k) = 1$ and that the probability of staying in the same region $\theta_{a,t}(i \to i) = 1 - \sum_{k \neq i}^{K}\theta_{a,t}(i \to k)$. In our formulation, absence of age or region subscripts for abundance denotes summation over the missing subscript, and absence of a time subscript denotes equilibrium (i.e., $t \to \infty$). For other model parameters absence of a subscript indicates the parameter is constant over the missing subscript.

The assumptions for such a model are generally the same as that for many age-structured fishery models (e.g., Deriso et al. 1985) with the added dimension of region. Migration rates $\theta_{a,t}(i \to k)$ are assumed to depend on age, year, and region. The migration process is assumed to be Markovian because for a given year and region all fish of the same age have the same migration probabilities regardless of their previous history. Net fecundity $f_{a,i}$ is assumed to depend on age and region and is computed from the product of proportion of females mature at age and fecundity of a mature female at age. Alternate formulations of the model are conceptually possible such as age-, year-, and region-specific natural mortality, year-specific fecundity and selectivity, and migration rates that depend on year class, population density, or season. Migration is assumed to occur instantaneously just after the start of a year followed by natural and fishing mortality.

The basic recursive equation that describes the dynamics of the population is

$$\textit{New population} = \textit{survival}\,(\textit{old population} - \textit{emigration} + \textit{immigration})$$

$$N_{a+1,t+1,i} = S_{a,t,i}\left[N_{a,t,i} - N_{a,t,i}\sum_{k \neq i}^{K}\theta_{a,t}(i \to k) + \sum_{k \neq i}^{K} N_{a,t,k}\theta(k \to i)\right]. \tag{1}$$

Equation (1) simplifies to

$$N_{a+1,t+1,i} = S_{a,t,i}\sum_{k=1}^{K}\theta_{a,t}(k \to i)N_{a,t,k}. \tag{2}$$

From the Baranov catch equation, the exploitation fraction is

$$\mu_{a,t,i} = \frac{s_{a,t,i}F_{t,i}}{Z_{a,t,i}}\left[1 - \exp(-Z_{a,t,i})\right]. \tag{3}$$

Hence from (2), catch is

$$C_{a,t,i} = \mu_{a,t,i}\sum_{k=1}^{K}\theta_{a,t}(k \to i)N_{a,t,k}. \tag{4}$$

Egg production in year t and region i is

$$N_{0,t,i} = \sum_{a=1}^{A} f_{a,i} N_{a,t,i}, \tag{5}$$

which assumes that egg production takes place before migration at the start of the year. In the absence of migration during the early life stages, the relationship between recruitment and egg production can be written

$$N_{r,t+r,i} = l_{r,i} N_{0,t,i} \psi_i (N_{0,t,i}), \tag{6}$$

where $l_{r,i}$ is early life survival from age 0 up to recruitment age r without density dependence, and $\psi_i(N_{0,t,i})$ is a function representing density dependence in a region. If, however, recruitment is decoupled from local reproduction by movement (e.g., dispersal of planktonic larvae or movement of prerecruits), then

$$N_{r,t+r,i} = l_{r,i} \sum_{k=1}^{K} N_{0,t,k} \psi(N_{0,t,k}) \theta_{0,t}(k \to i), \tag{7}$$

which assumes that density dependence precedes dispersal, and dispersal occurs up to the recruitment age r followed by survival. Note that equation (7) is the same as (6) if $\theta_{0,t}(i \to k) = 0$ for $k \neq i$ and $\theta_{0,t}(i \to k) = 1$ for $k = i$. A more complex model that incorporates physical factors such as oceanographic currents can more realistically account for dispersion (e.g., Possingham and Roughgarden 1990), but the simpler model may be adequate.

Putting the model described in equations (1)-(7) in the form of a projection matrix is a convenient way of examining the dynamics of such a model (Caswell 1989). Given the vector of abundance at age and region $\boldsymbol{N}_t$, the matrix $\boldsymbol{P}_t$ in the equation $\boldsymbol{N}_{t+1} = \boldsymbol{P}_t \boldsymbol{N}_t$ is the projection matrix. As an example, consider the hypothetical example of three age groups and three regions, without density dependence, without migration during early life stages, and constant fecundity, survival and migration over time. The projection matrix $\boldsymbol{P}$ is

$$\boldsymbol{P} = \begin{pmatrix}
l_1 f_{1,1} & l_1 f_{2,1} & l_1 f_{3,1} & 0 & 0 & 0 & 0 & 0 & 0 \\
S_{1,1}\theta_1(1 \to 1) & 0 & 0 & S_{1,1}\theta_1(2 \to 1) & 0 & 0 & S_{1,1}\theta_1(3 \to 1) & 0 & 0 \\
0 & S_{2,1}\theta_2(1 \to 1) & 0 & 0 & S_{2,1}\theta_2(2 \to 1) & 0 & 0 & S_{2,1}\theta_2(3 \to 1) & 0 \\
0 & 0 & 0 & l_2 f_{1,2} & l_2 f_{2,2} & l_2 f_{3,2} & 0 & 0 & 0 \\
S_{1,2}\theta_1(1 \to 2) & 0 & 0 & S_{1,2}\theta_1(2 \to 2) & 0 & 0 & S_{1,2}\theta_1(3 \to 2) & 0 & 0 \\
0 & S_{2,2}\theta_2(1 \to 2) & 0 & 0 & S_{2,2}\theta_2(2 \to 2) & 0 & 0 & S_{2,2}\theta_2(3 \to 2) & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & l_3 f_{1,3} & l_3 f_{2,3} & l_3 f_{3,3} \\
S_{1,3}\theta_1(1 \to 3) & 0 & 0 & S_{1,3}\theta_1(2 \to 3) & 0 & 0 & S_{1,3}\theta_1(3 \to 3) & 0 & 0 \\
0 & S_{2,3}\theta_2(1 \to 3) & 0 & 0 & S_{2,3}\theta_2(2 \to 3) & 0 & 0 & S_{2,3}\theta_2(3 \to 3) & 0
\end{pmatrix},$$

where the population vector for abundance N_t can be written as

$$\boldsymbol{N_t} = \begin{pmatrix} N_{1,t,1} \\ N_{2,t,1} \\ N_{3,t,1} \\ N_{1,r,2} \\ N_{2,t,2} \\ N_{3,t,2} \\ N_{1,t,3} \\ N_{2,t,3} \\ N_{3,t,3} \end{pmatrix}.$$

The entries in the first, fourth, and seventh rows of $\boldsymbol{P}$ are termed the fertility coefficients and the other entries are termed the survival-migration coefficients. To introduce migration and/or density dependence at early life stages requires modification of the fertility coefficients. For example from equations (6) and (7), the element $P_{1,1}$ (row 1, column 1) is

$$\begin{aligned} P_{1,1} &= l_1 f_{1,1} \psi_1 \left(\sum_{a=1}^{A} f_{a,1} N_{a,t,1} \right) \theta_0 (1 \rightarrow 1) \\ &= l_1 f_{1,1} \psi_1 (N_{0,t,1}) \theta_0 (1 \rightarrow 1), \end{aligned}$$

and the element $P_{1,4}$ (row 1, column 4) is

$$P_{1,4} = l_1 f_{1,2} \psi_2 (N_{0,t,2}) \theta_0 (2 \rightarrow 1).$$

With the addition of density dependence, this is no longer a constant linear Leslie matrix because elements of $\boldsymbol{P}$ are now functions of abundance and ψ can be any function.

Equations (2) and (5) are equivalent to the traditional Leslie matrix model with migration when $\psi(x) = 1$, and survival, fecundity, and migration are constant over time. Note that Caswell (1989, Chapter 4) describes this model extensively. For many populations such a model is not adequate. A more realistic model can be formulated by inclusion of density dependence to account for nonlinearity in the relationship between egg production and recruitment. Density dependence can result in a stationary population and stable age-by-region structure, or in complex behavior such as limit cycles and chaotic fluctuations (Nisbet and Gurney 1982, Caswell 1989). Details of alternative ways that regional density dependence can be formulated are in Heifetz (1996).

Application of the Model to Sablefish

Traditional Leslie Matrix with Migration

Sablefish, one of the most valuable commercial fish in waters off Alaska, is used to illustrate the application of the model. Migration rates based on tagging experiments were estimated by Heifetz and Fujioka (1991). Population parameters and age-specific migration rates are shown in Tables 1 and 2, respectively. There are five regions which correspond to the five North Pacific Fishery Management Council (NPFMC) regulatory areas (Fig. 1). Note that the population parameters are assumed to be the same in reach region. Estimates of egg production and number of recruits at age 3, and early life survival, are in Table 3. These were derived by Heifetz (1996) based on the results of an age-structured stock assessment model of Alaska's sablefish population (Sigler 1993, updated by pers. comm., M. Sigler, Nat. Mar. Fish. Serv., Juneau, AK 99801, Oct. 1995). The median early life survival of 1.00×10^{-6} from Table 3 can be used as an estimate of l_r (Quinn and Szarzi 1993). To initialize the population we used the regional age composition of the population in 1991 estimated by Heifetz (1996). The most recent year where age data were collected from all NPFMC regulatory regions was 1991.

During 1992-1995 the transition matrix $\boldsymbol{P}$ is applied with F rates from Sigler (Pers. comm., October 1995). These F rates were 0.088, 0.094, 0.111, and 0.105 for 1992, 1993, 1994, and 1995, respectively. Projection of future abundance starting in 1996 is made with the F value (F_{st}) that results in stationary population abundance. The F_{st} value of 0.143 was determined with a nonlinear search procedure.

The projected population reaches a stable age-by-region equilibrium in about 25 years (Fig. 2). The initial dip in numbers of age > 10 sablefish is due to the poor recruitment in recent years. The regional distribution stabilized in about 5 years which was much faster than the age distribution. Thus for sablefish, the amount of mixing among regions is rapid relative to the progression of a cohort through the age classes. This is partly due to sablefish being long lived, the considerable time for the initial population to die out, and that additions to the overall population are only made at the youngest age. On the other hand, movement is rapid; annually 19-72% of the population in a region migrates to other regions (found by subtracting each diagonal term in Table 2 from 1), and exchange occurs among all regions.

Density Dependence

To introduce density dependence, egg production and subsequent recruits from Table 3 can be used to estimate a stock-recruitment relationship. When plotted, the data are suggestive of a Ricker type relationship (Fig. 3) of the form $N_{r,t+r} = \alpha N_{0,t} \exp(-\beta N_{0,t})$. Parameters for this equation were estimated using a nonlinear search procedure.

Table 1. Estimates of population parameters for sablefish in the northeastern Pacific Ocean.

Age (*a*)	3	4	5	6	7	8	9	10	11	12	≥13
Natural mortality, M_a	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10
Maturity, m_a	0.002	0.032	0.232	0.629	0.868	0.949	0.977	0.988	0.993	0.998	1.00
Selectivity, s_a	0.15	0.32	0.48	0.63	0.73	0.82	0.88	0.93	0.97	1.00	1.00

Female von Bertalanffy length-age model	Female allometric weight-length model	Fecundity model
$L_a(\text{cm}) = L_\infty\{1 - \exp[-k(a - t_0)]\}$	$\text{weight}_a(\text{kg}) = cL_a^b$	$f_a(\text{eggs}) = m_a 0.02349(L_a)^{3.88}$
$L_\infty = 81.4$	$c = 3.15 \times 10^{-6}$	
$k = 0.249$	$b = 3.290$	
$t_0 = -0.770$		

Fraction mature at age, selectivity at age, and growth parameters are from Sigler and Fujioka (1993). Fecundity parameters are adapted from Lowe et al. (1991).

Table 2. Annual movement rates from one area to another for different age groups of sablefish.

From area *(i)*	To area *(k)* EG	CG	WG	BS	AL
Age 3-4					
EG	0.489	0.378	0.109	0.012	0.012
CG	0.194	0.484	0.229	0.047	0.046
WG	0.078	0.322	0.308	0.145	0.147
BS	0.010	0.073	0.160	0.710	0.047
AL	0.005	0.040	0.093	0.049	0.813
Age 5-8					
EG	0.712	0.227	0.054	0.003	0.004
CG	0.273	0.476	0.199	0.023	0.029
WG	0.134	0.409	0.285	0.072	0.100
BS	0.029	0.165	0.246	0.494	0.066
AL	0.013	0.077	0.128	0.034	0.748
Age ≥ 9					
EG	0.751	0.200	0.044	0.002	0.003
CG	0.472	0.426	0.074	0.011	0.017
WG	0.262	0.116	0.517	0.046	0.059
BS	0.095	0.181	0.259	0.287	0.178
AL	0.047	0.103	0.155	0.018	0.677

Adapted from Heifetz and Fujioka (1991).

EG, Eastern Gulf of Alaska; CG, Central Gulf of Alaska; WG, Western Gulf of Alaska; BS, Eastern Bering Sea; AL, Aleutians.

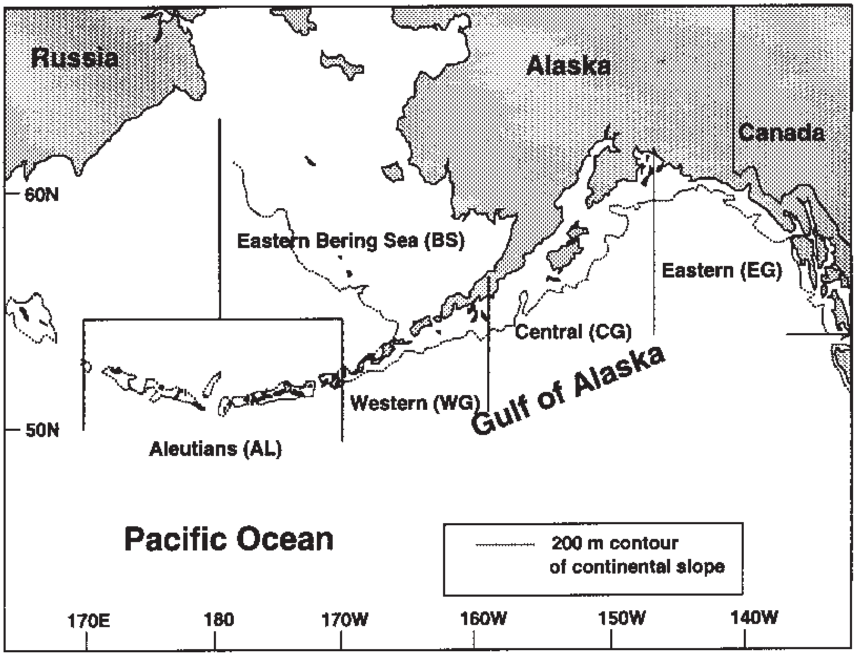


Figure 1. Northeast Pacific ocean showing North Pacific Fishery Management Council (NPFMC) regulatory areas. Abbreviations used throughout the text are in parentheses.

The best fit to the data is with parameter values of 2.214×10^{-5} for α and 2.037×10^{-7} for β. Such a fit indicates an extremely strong level of density dependence and is probably unrealistic (Fig. 3). Thus, the stock-recruitment relationship is mainly used for illustrative purposes. Given the large uncertainty in this relationship, alternative levels of density dependence that are reasonably consistent with the data are based on values of β one-half (medium density dependence) and one-quarter (weak density dependence) of the best fit estimate of β (Fig. 3).The results with medium density dependence (i.e., $\alpha = 6.86 \times 10^{-6}$; $\beta = 1.02 \times 10^{-7}$) are used to illustrate the application of the model.

We assume that density-dependent effects within a region are from the population as a whole, with the same stock-recruitment relationship among regions (i.e., "Case 1" in Heifetz 1996). In this case the stock-recruitment relationship represents the entire population pooled over regions so that early life survival in the absence of density dependence (α)

Table 3. Stock and recruitment data for sablefish in the northeast Pacific Ocean.

Year	Eggs (10^{12})	Recruitment at age 3 (10^6)	Survival to age 3 (10^{-6})
1976	7.09	23.09	3.26
1977	7.05	44.19	6.27
1978	7.61	40.60	5.33
1979	8.57	31.09	3.62
1980	8.95	38.01	4.25
1981	9.50	41.70	4.39
1982	10.87	13.82	1.27
1983	13.41	15.25	1.14
1984	16.41	14.27	0.87
1985	19.23	10.10	0.53
1986	21.94	5.05	0.23
1987	23.64	5.16	0.22
1988	23.78	4.15	0.17
1989	22.70	4.49	0.20
1990	21.35	8.69	0.41
1991	19.63	4.57	0.23
		median =	1.00

Estimated by Heifetz (1996).

and the level of density dependence (β) are independent of region. Such a model can be written as

$$N_{r,t+r,i} = \alpha N_{0,t,i} \exp\left(-\beta \sum_{k=1}^{K} N_{0,t,k}\right).$$

The term $N_{0,t,i}$ insures that recruits are apportioned to region i in direct proportion to the number of eggs produced in that region.

The population was initialized with the estimated 1991 population and projected with $F = 0.1$ (near the current level). Figure 4 shows the trajectories of population number by region for a medium level of density dependence. In general the approach to the equilibrium population was a succession of damped oscillations.

Characteristic of the Ricker stock-recruitment relationship is that certain parameter values can cause instability of the equilibrium population. Such populations may exhibit stable limit cycles or chaotic fluctuations. For example in Fig. 5, early life survival in the absence of density depen-

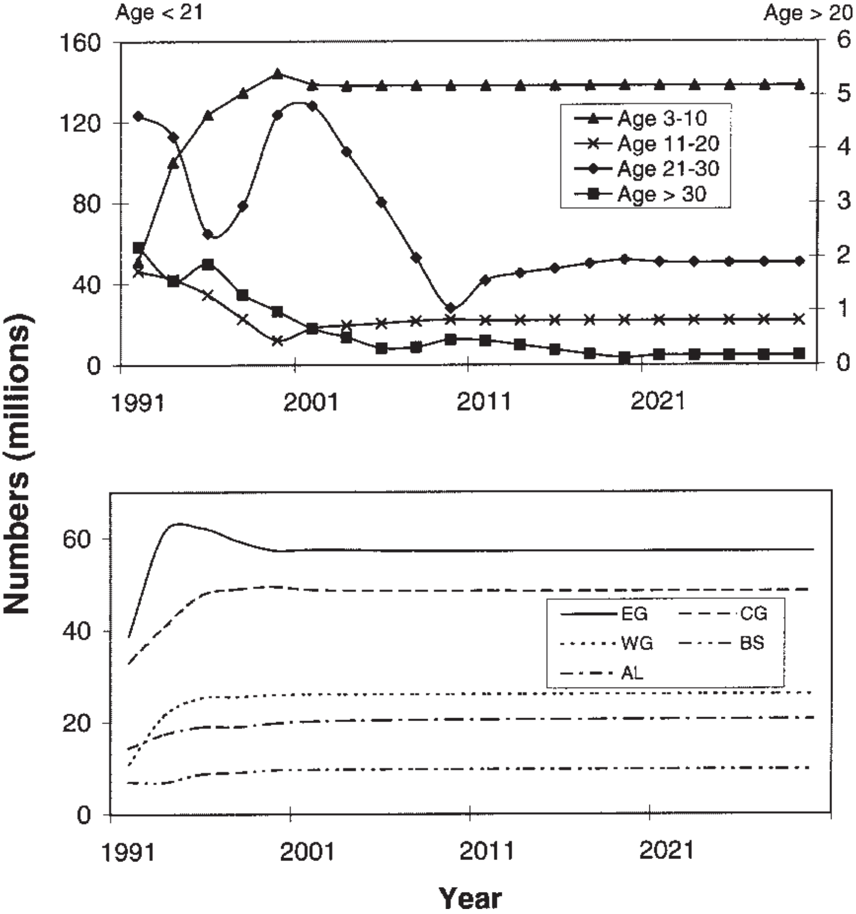


Figure 2. Estimated abundance of sablefish by age and region from 1991 to 1995 and projected abundance from 1996 to 2031 based on the traditional Leslie matrix approach with migration, median early life survival of 1.00×10^{-6}, and $F_{st} = 0.143$. EG, Eastern Gulf of Alaska; CG, Central Gulf of Alaska; WG, Western Gulf of Alaska; BS, Eastern Bering Sea; AL, Aleutians.

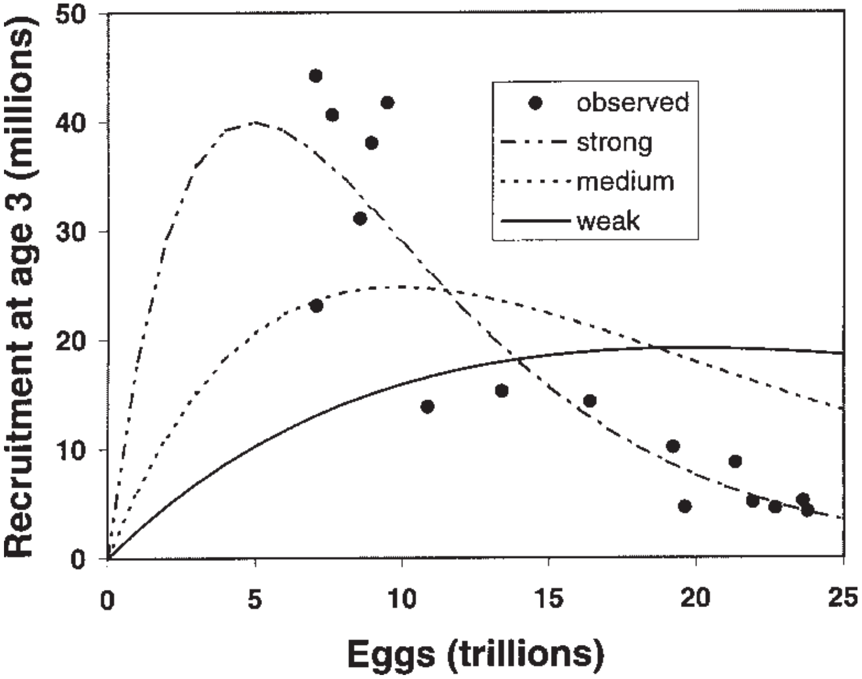


Figure 3. Observed recruitment versus egg production and three Ricker stock-recruitment curves for sablefish in the Northeast Pacific Ocean. Strong, medium, and weak density dependence are defined with (α, β) = (2.21 × 10^{-5}, 2.04 × 10^{-7}), (6.86 × 10^{-6}, 1.02 × 10^{-7}), and (2.65 × 10^{-6}, 5.09 × 10^{-8}).

dence (i.e., α) in one region (EG) was increased from 6.86 × 10^{-6} to 1.00 × 10^{-3}. The resulting regional population numbers exhibit undamped oscillations with periodic shifts in the proportion of the population in each region. An interesting cyclical occurrence is the periodic, nearly equivalent population numbers in EG and CG accompanied by the peak in the proportion of the population in the other regions. Detailed examination of the mathematical properties that govern the dynamic behavior exhibited by stock recruitment relationships is provided in May (1974) and Clark (1976).

Management Application: Per-Recruit Analysis

The age-by-region model can be easily adapted to examine harvest strategies. For most commercial fisheries there is great difficulty in reliably defining the stock-recruitment relationship. Per-recruit analysis is frequently used when there is uncertainty in the stock-recruitment relationship

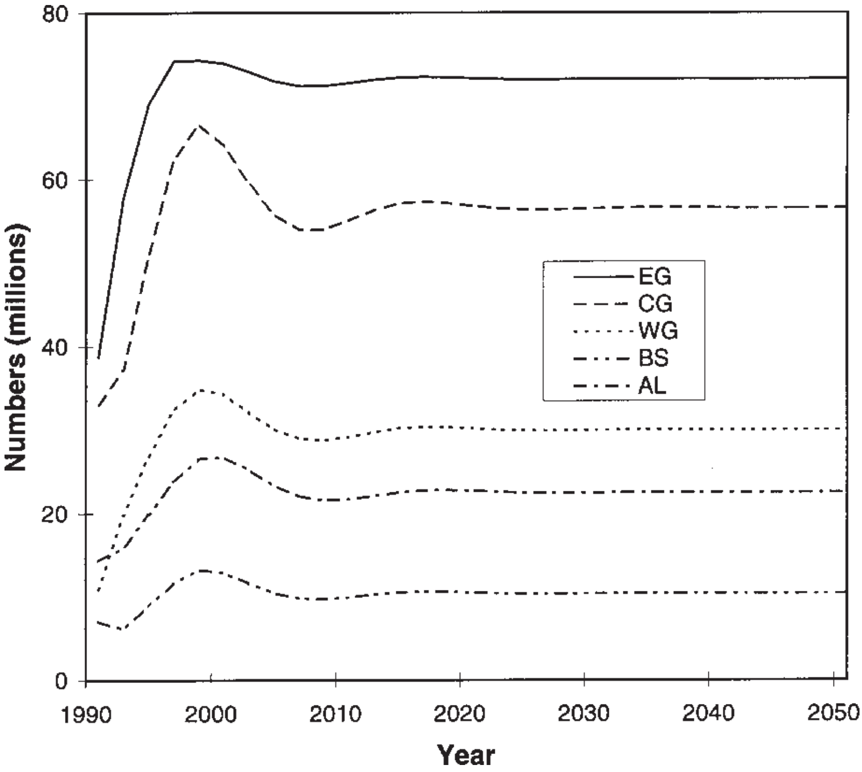


Figure 4. Estimated regional abundance and regional distribution of sablefish from 1991 to 1995 and projected abundance from 1996 to 2050 with $F = 0.1$ and density dependence defined by a Ricker stock-recruitment relationship with $(\alpha, \beta) = (6.86 \times 10^{-6}, 1.02 \times 10^{-7})$. EG, Eastern Gulf of Alaska; CG, Central Gulf of Alaska; WG, Western Gulf of Alaska; BS, Eastern Bering Sea; AL, Aleutians.

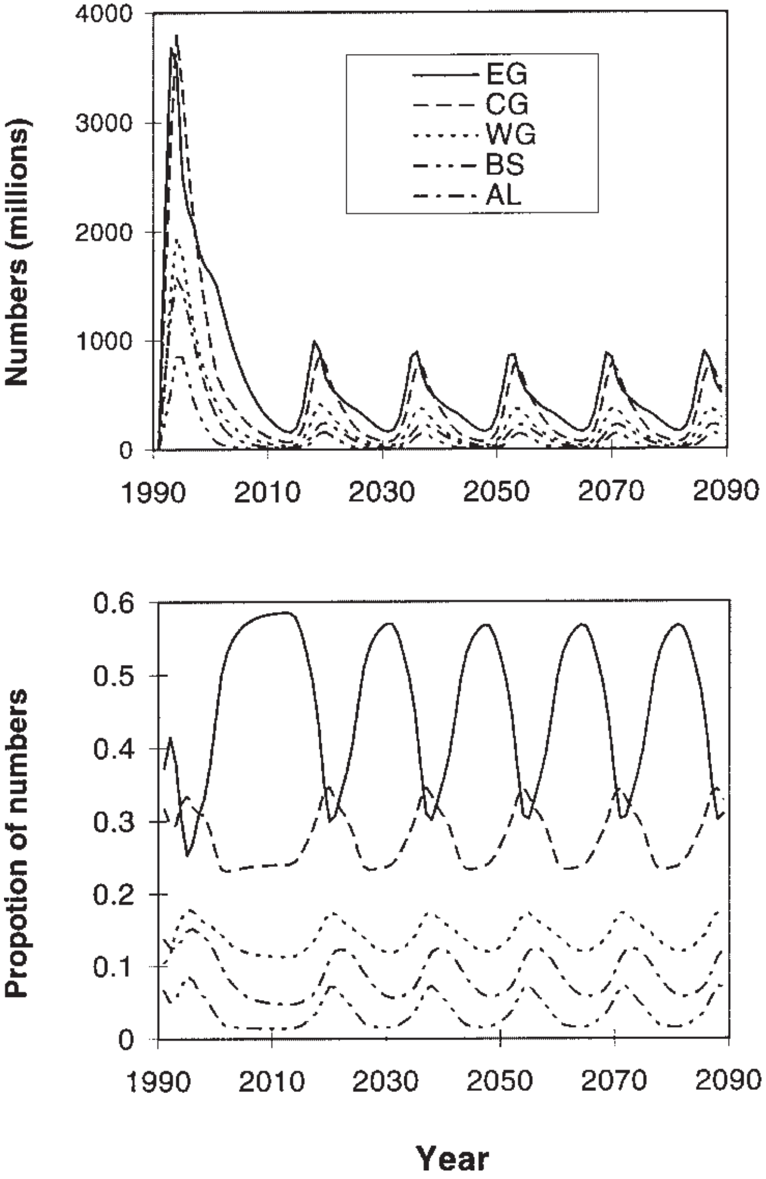


Figure 5. Estimated regional abundance and regional distribution of sablefish from 1991 to 1995 and projected abundance from 1996 to 2090 with $F = 0.1$ and density dependence defined by a Ricker stock-recruitment relationship with $(\alpha, \beta) = (6.86 \times 10^{-6}, 1.02 \times 10^{-7})$ and $\alpha_{EG} = 1.00 \times 10^{-3}$. EG, Eastern Gulf of Alaska; CG, Central Gulf of Alaska; WG, Western Gulf of Alaska; BS, Eastern Bering Sea; AL, Aleutians.

(Clark 1991). Thus our management application of the model focuses on per-recruit analysis.

In per-recruit analysis, the consequences of fishing are surmised from following a cohort as it passes through the fishery during its life span. By rewriting equations (2) and (4) as a series of equations to be updated, per-recruit analyses can be undertaken. Given initial and perhaps arbitrary recruitment numbers by region ($N_{r,i}$), the abundance for each age and region after migration can then be written as

$$N'_{a,i} = \sum_{k=1}^{K} \theta_a(k \to 1) N_{a,k} \quad \text{for } a > r.$$

The catch-at-age and region is

$$C_{a,i} = \mu_{a,i} N'_{a,i},$$

and the updated abundance-at-age after movement and survival is

$$N_{a+1,i} = S_{a,i} N'_{a,i}.$$

One example of an initial recruitment vector is to start with 1,000 recruits per region. Another is to distribute them in proportion to estimates from a survey or stock assessment. Catch, yield (Y), and spawning biomass (SB) per recruit can be computed from

$$\frac{C_i}{N_r} = \frac{\sum_{a=r}^{A} \mu_{a,i} N'_{a,i}}{N_r}$$

$$\frac{Y_i}{N_r} = \frac{\sum_{a=r}^{A} \mu_{a,i} N'_{a,i} w_{a,i}}{N_r}$$

$$\frac{SB_i}{N_r} = \frac{\sum_{a=r}^{A} m_{a,i} N'_{a,i} w_{a,i}}{N_r}$$

where $w_{a,i}$ is weight at age and region and $m_{a,i}$ is the proportion mature at age and region. The total number of recruits for all regions combined, N_r, is used because recruitment from all regions may contribute to yield and spawning biomass within a given region. As shown in Heifetz et al. (1997), per-recruit estimates can be sensitive to assumptions about the regional distribution of recruitment.

Commonly used reference fishing mortality rates such as F_{max}, $F_{0.1}$, and $F_{35\%}$ (Clark 1991), can be determined for each region and all regions combined. The use of F_{max}, the F value that maximizes yield per recruit, has mostly passed out a favor because this rate is typically very high and will likely deplete spawning biomass too drastically (Clark 1991). We present F_{max} values, however, because they are useful benchmarks for comparing among regions. $F_{0.1}$ is the rate where the slope of the yield per recruit curve as a function of F falls to 10% of the value at the origin. $F_{35\%}$ is the rate required to keep spawning biomass per recruit at 35% of the unfished level. For sablefish $F_{0.1}$ is close to $F_{35\%}$, and $F_{35\%}$ is close to the current harvest strategy (Fujioka 1995). (At the time our analysis was done, $F_{35\%}$ was the harvest strategy for sablefish in Alaska.) Thus, only results for F_{max} and $F_{35\%}$ are shown.

We examine two management applications of the model based on per-recruit analysis. The implications for fishery management are first examined with a constant fishing rate among regions. Second, we show how regional specific fishing rates can be devised that result in the population in a given region meeting a specific criterion.

F Common among Regions

To illustrate per-recruit analysis we use the distribution of regional recruitment estimated with $F = 0$ based on the traditional Leslie matrix with migration. Recruitment was distributed in the following proportions by region 0.439 in the EG, 0.270 in the CG, 0.146 in the WG, 0.039 in the BS, 0.106 in the AL.

We examine yield per recruit (YPR) and spawning biomass per recruit (SBPR) as a function of a common F among regions. YPR and SBPR expressed relative to the maximum within a region are shown in Fig. 6. The value of F that maximized YPR (F_{max}) differed dramatically by region. These F_{max} values are 0.20 for the EG, 1.02 for the CG, 0.86 for the WG, 1.44 for BS and 0.74 for the AL. Spawning biomass per recruit (SBPR) values also differed as a function of F by region. The values of F that reduced the SBPR to 35% of its unfished level ($F_{35\%}$) in a region were 0.09 for the EG, 0.12 in the CG, 0.13 in the WG, 0.20 in the BS, and 0.16 in the AL. The pooled $F_{35\%}$ value of 0.112 is close to the currently used exploitation strategy for sablefish in waters off Alaska. Fishing at the pooled rate will result in SBPR maintained at ≥ 35% of the unfished level for all regions except the EG.

The result that F values that maximize yields and maintain spawning biomass differ by region can be explained from the combination of two factors. The first factor is from differences in the rate at which abundance of a cohort changes with time due to “natural survival” (i.e., the combination of natural mortality and migration). All other parameters being equal, in traditional YPR and SBPR analyses that do not consider migration, F_{max} and $F_{35\%}$ values are correlated with natural mortality M. The population with the greatest M will have the greatest values of F_{max} and $F_{35\%}$. For the age-region model, migration rates and natural mortality determine the level of natural survival of a cohort in a given region. This phenomenon is exemplified in Fig. 7 which traces the relative abundance of a cohort through

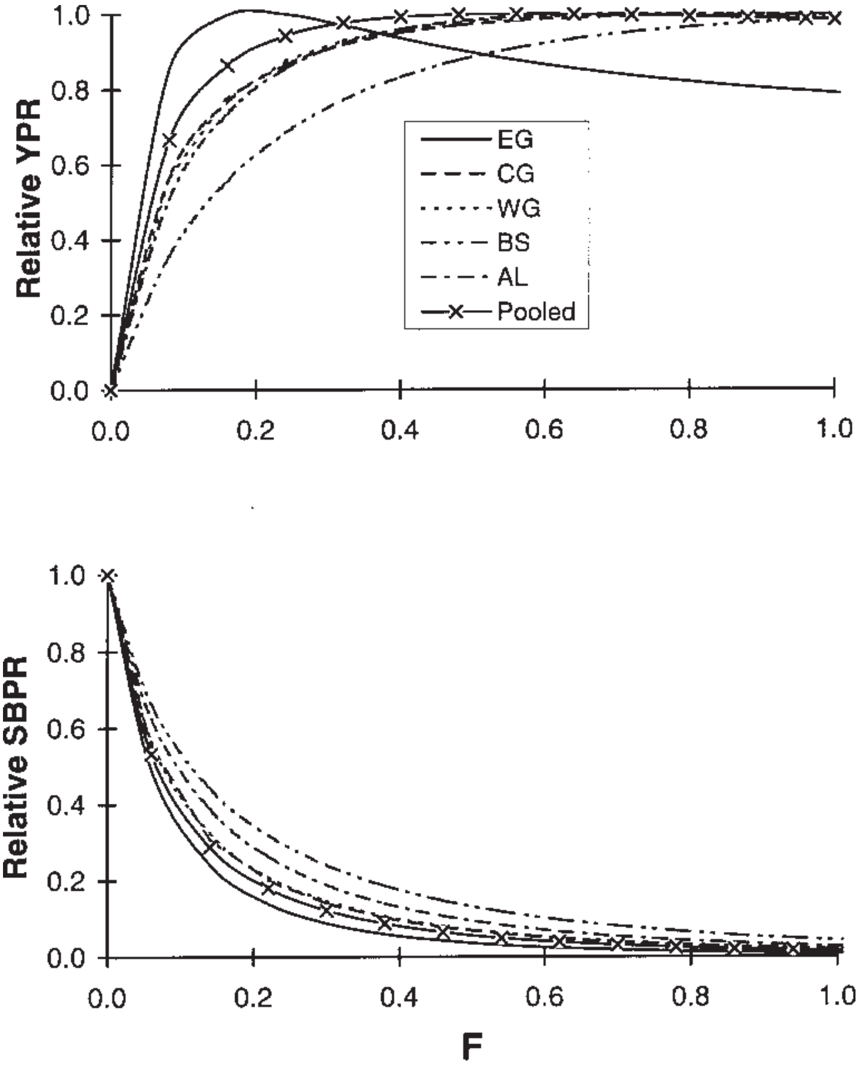


Figure 6. *Yield per recruit (YPR) and spawning biomass per recruit (SBPR) for sablefish relative to the maximum within a region for F common to all regions. EG, Eastern Gulf of Alaska; CG, Central Gulf of Alaska; WG, Western Gulf of Alaska; BS, Eastern Bering Sea; AL, Aleutians.*

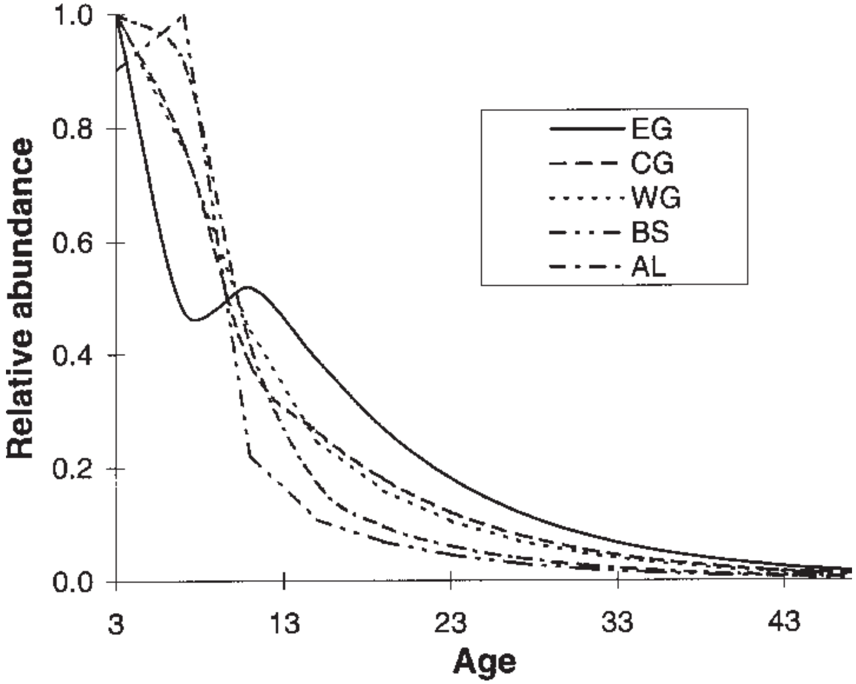


Figure 7. Abundance of a cohort relative to the maximum in a region with F = 0 for sablefish in the Northeast Pacific Ocean. EG, Eastern Gulf of Alaska; CG, Central Gulf of Alaska; WG, Western Gulf of Alaska; BS, Eastern Bering Sea; AL, Aleutians.

time in each region in the absence of fishing. In general, the quicker the abundance declines in a region, the higher the F_{max} and $F_{35\%}$ values.

The second factor that may influence per-recruit rates is fishery interceptions. For example, the lower value of F_{max} and $F_{35\%}$ in the EG compared to the other regions can in part be explained by the relationship between *F* and the regional distribution of yield and spawning biomass. As *F* increases an increasingly greater proportion of the yield is taken in other regions, especially in the CG (Fig. 8). This phenomenon is paralleled for spawning biomass where an increasingly greater proportion of the spawning biomass is present in the CG as *F* is increased. Young sablefish (age 3-5) mostly migrate north and westward from the EG to other regions but return as they get older (Heifetz and Fujioka 1991).

Region-Specific F

The results of the preceding section indicate that fishing at a constant rate which may be appropriate for the pooled population may be detri-

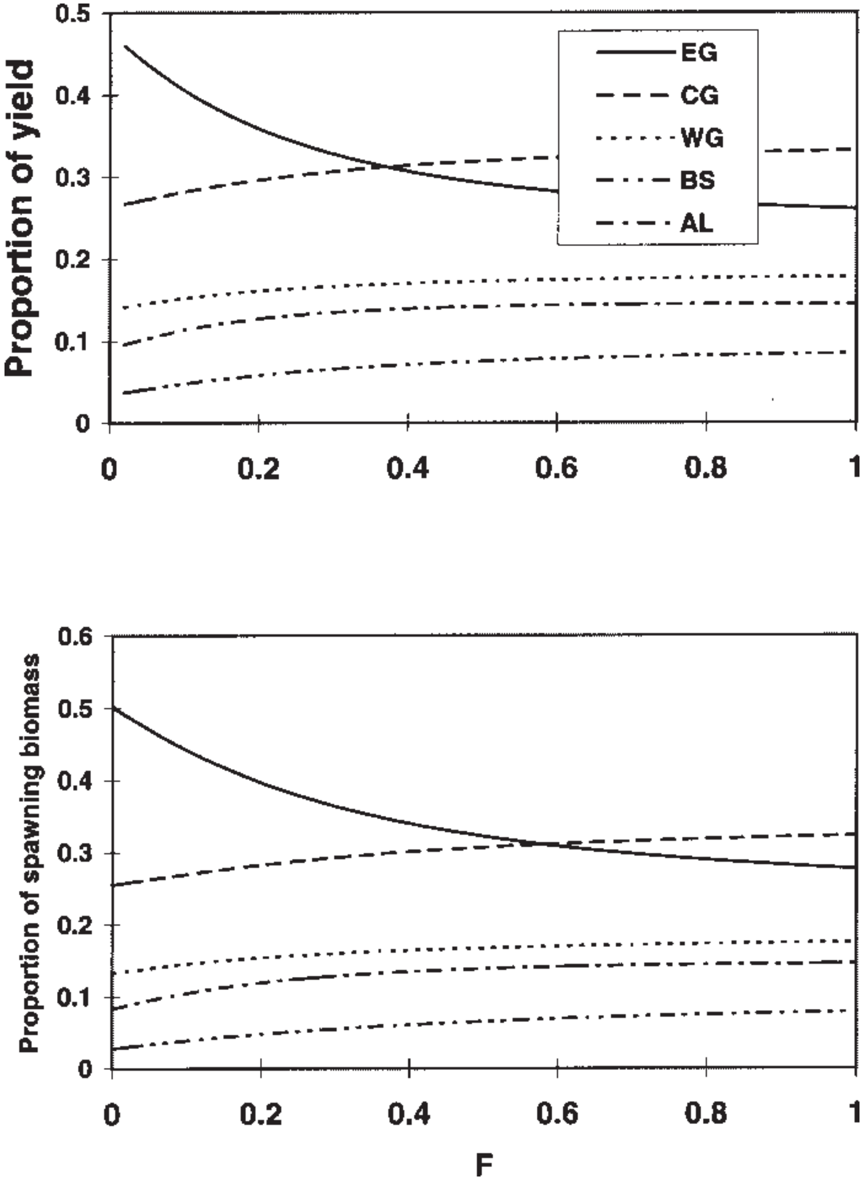


Figure 8. Proportion of yield and proportion of spawning biomass by region versus F for sablefish in the Northeast Pacific Ocean. EG, Eastern Gulf of Alaska; CG, Central Gulf of Alaska; WG, Western Gulf of Alaska; BS, Eastern Bering Sea; AL, Aleutians.

mental to the population in a given region. Alternatively there may be sets of region-specific fishing rates that meet fishery management goals. A fishery management goal can be formulated as an objective function to be optimized. For example, a reasonable goal of fishery management might be to choose a set of region-specific *F* values that maintain the regional distribution of spawning potential at unfished levels and at the same time maintain pooled spawning biomass per recruit at a specified target level. Just as with other harvest policies that are commonly used to manage fish populations (Quinn and Deriso 1998) this goal is based on equilibrium conditions. Departures from the equilibrium population are expected because of variation in recruitment, migration, and other population parameters. The objective function for such a goal comprises two parts: (1) the sum over regions of the squared deviations between regional-specific proportions of spawning biomass (SBP) with and without fishing and, (2) the squared deviation between the reduction in pooled spawning biomass per recruit (SBPR) relative to unfished levels and a target level of reduction (*x*),

$$\min\left|\sum_{i=1}^{I}\left(SBP_{F_i}-SBP_{F_i=0}\right)^2+\left(\frac{\sum_{i=1}^{I}SBPR_{F_i}}{\sum_{i=1}^{I}SBPR_{F_i=0}}-x\right)\right|^2.$$

To illustrate this approach for sablefish, we set the target level of reduction for SBPR at 35% of the unfished level (ie., $x = 0.35$; Case 1 in Table 4). We also determined sets of *F* values that minimized the objective function given a fixed fishing rate (0.04-0.16) in a given region (EG) (Cases 2-5; Table 4). We used the distribution of regional recruitment based on $F = 0$ as determined from the traditional Leslie matrix with migration. Recruitment was distributed in the following proportions by region 0.439 in the EG, 0.270 in the CG, 0.146 in the WG, 0.039 in the BS, 0.106 in the AL. Based on this distribution of recruitment, equilibrium spawning biomass was distributed in the following proportions by region 0.502 in the EG, 0.254 in the CG, 0.133 in the WG, 0.028 in the BS, 0.083 in the AL.

A target level of spawning biomass per recruit of 35% of the unfished level and regional preservation of the distribution of spawning biomass at unfished levels is obtained at regionally specific *F* values (Case 1) that differ substantially from the constant regional $F_{35\%}$ value of 0.112 (Case 6; Table 4). Pooled yield was nearly identical for the constant *F* values and regional *F* values. The constant *F* value results in pooled spawning biomass per recruit meeting the target but disruption of the distribution of spawning biomass. An exact solution to the objective function was found at regional specific *F* values of 0.042 for the EG, 0.164 for the CG, 0.128 for the WG, 0.284 for the BS, and 0.180 for the AL. Thus, relative to the constant *F* value, the level of fishing would need to be reduced considerably in the EG, moderately increased in the CG, WG, and AL, and increased

Table 4. Regional *F* values, sum of squared deviations (SSQ), spawning biomass relative to unfished levels (SBPR/SBPR$_{F=0}$), spawning biomass (SB), proportion of spawning biomass (SBP), and yield.

Region	F	SSQ (10^6)	SBPR/SBPR$_{F=0}$	SB	SBP	Yield	Case
EG	0.042	0.01	0.350	32,784	0.502	1,886	1
CG	0.164		0.350	16,646	0.255	4,519	
WG	0.128		0.350	8,670	0.133	1,829	
BS	0.294		0.350	1,820	0.028	1,169	
AL	0.180		0.350	5,406	0.083	1,755	
Pooled			0.350	65,325		11,158	
EG	0.040	6.19	0.351	32,853	0.504	1,781	2
CG	0.168		0.348	16,525	0.253	4,623	
WG	0.130		0.348	8,619	0.132	1,855	
BS	0.291		0.351	1,824	0.028	1,157	
AL	0.179		0.350	5,408	0.083	1,745	
Pooled			0.350	65,228		11,161	
EG	0.080	1,231.51	0.330	30,915	0.475	3,435	3
CG	0.108		0.377	17,932	0.276	3,122	
WG	0.097		0.372	9,219	0.142	1,449	
BS	0.338		0.337	1,749	0.027	1,319	
AL	0.200		0.340	5,256	0.081	1,919	
Pooled			0.349	65,072		11,244	
EG	0.120	4,907.24	0.310	29,072	0.447	4,965	4
CG	0.060		0.403	19,166	0.295	1,810	
WG	0.070		0.395	9,794	0.151	1,091	
BS	0.355		0.338	1,756	0.027	1,394	
AL	0.216		0.336	5,184	0.080	2,059	
Pooled			0.348	64,973		11,318	
EG	0.160	10,495.05	0.291	27,263	0.421	6,363	5
CG	0.022		0.425	20,189	0.312	684	
WG	0.049		0.416	10,309	0.159	792	
BS	0.346		0.351	1,827	0.028	1,396	
AL	0.226		0.336	5,185	0.08	2,162	
Pooled			0.347	64,773		11,397	
EG	0.112	5,652.76	0.303	28,407	0.435	4,521	6
CG	0.112		0.372	17,710	0.271	3,216	
WG	0.112		0.387	9,573	0.147	1,738	
BS	0.112		0.504	2,618	0.040	558	
AL	0.112		0.452	6,984	0.107	1,311	
Pooled			0.350	65,292		11,344	

substantially in the BS to meet regional management goals. Analogous to the analysis of constant *F* values in the preceding section, differences in *F* values among regions can be explained from the influence of different natural survival rates among regions and fishery interceptions.

Fixing *F* for a variety of values for the EG and solving for the *F* values in the other regions shows the relationship among the regional *F* values (Cases 2-5; Table 4). In general as F_{EG} is increased, *F* values in the CG and WG decrease substantially while *F* values in the BS and AL remain relatively constant. While resulting in larger deviations in regional spawning distribution than the *F* values that optimized the objective function (Case 1), these sets of *F* values give some alternatives and consequences of adopting regionally specific harvest strategies.

Discussion

We have presented a general framework to model the regional dynamics of an age-structured fish population. The model has its origin in the generic stage-classified matrix model described in Caswell (1989). Compared to age-structured models that aggregate the populations in different regions, at a minimum the additional parameters needed to construct an age-by-region model are estimates of age-specific migration rates among regions. The model can accommodate dispersal patterns at early life stages and regional differences in the stock-recruitment relationship and demographic parameters. Data needed to incorporate such complexity are often not available for fish populations or may be difficult to collect, although for some species dispersal of early life stages is becoming better understood. Temporal and stochastic variation in model parameters can be incorporated to account for variability due to environmental or other extrinsic causes (e.g., Caswell 1989, Chapter 8). In cases where estimates of age-specific migration rates are not available, size rather than age can be used for stage classification (Caswell 1989), or techniques based on stock-production models may be better suited (e.g., Die et al. 1990, MacCall 1990).

The age-by-region model can be modified to include a more complicated representation of migration dynamics such as density-dependent migration patterns. Such density-dependent migration patterns can result in damping of population cycles (Emlen 1984). The “basin model” of MacCall (1990) incorporates density-dependent migration patterns but has not been extended to age-structured populations. An appealing feature of the basin model is the incorporation of possible interactions between population density, the geographic range of distribution of a population, and habitat suitability. Exploration of such an ecologically based model within the context of our model formulation might be warranted.

For some fish species inclusion of seasonal migration patterns may be required to adequately describe population dynamics. Examples of seasonal migrations are the latitudinal shifts of many Mid-Atlantic Ocean fishes (e.g., bluefish [*Pomatomus saltatrix*]) related to seasonal changes in water temperature (Tyler 1971, Colvocoresses and Musick 1983) and the seasonal spawning migrations of Pacific whiting (*Merluccius productus*; Swartzman et al. 1987), Atlantic cod (*Gadus morhua*; Rose et al. 1995), and Pacific cod (*Gadus macrocephalus*; Shimada and Kimura 1994). Whether sablefish undergo seasonal migrations is not known. To evaluate the interaction between seasonal migration and exploitation patterns may require that such migrations be included in the model structure. Accounting for seasonal changes in the fraction of the overall population that is available to fishing in a given region would require incorporation of an additional dimension to the model structure and a shorter time step.

Although Beverton and Holt (1957) and more recent studies such as Fujioka (1978), Polacheck (1990), Die et al. (1990), and Deriso et al. (1991) included spatial dynamics into their population models, most models used to describe the dynamics of exploited fish populations are based on the assumptions of a stock that is spatially homogeneous with uniform fishing mortality. We describe methods to examine the regional dynamics of an exploited fish population. These methods parallel those commonly used for a single stock or populations pooled over regions (Getz and Haight 1989). First we assume there is a proportional relationship between stock size and recruitment. Then we modified the assumption of a proportional relationship between stock size and recruitment by inclusion of a stock-recruitment relationship that incorporates density dependence.

Our application of the model focused on population projection and evaluation of harvest strategies. Other uses for the model include estimation of population abundance and parameters, environmental impact studies, and general ecological studies. An essential component of fisheries science is estimation of population abundance (i.e., stock assessment). For many species the method of obtaining regional estimates of abundance is by first obtaining the abundance of the population pooled over regions through age-structured modeling and then partition this estimate by regional specific estimates of relative abundance (i.e., fishery or survey catch per unit effort). Quinn et al. (1990) describe an alternative to this method called migratory catch-age analysis. The underlying equations that describe the dynamics of the population for their method are analogous to equations (1) and (2) in our formulation. Migration parameters were assumed to be known and were not estimated within the model. While generally improving the precision of regional estimates of abundance, a problem encountered was that negative abundance estimates occurred for some areas due to inconsistency of data sources or model deficiencies. To alleviate this problem, perhaps migration rates could be determined within the estimation procedure if sufficient data are available. Although adding complications to the estimation procedure, incorporation of estimation of migration rates into a stock assessment may improve estimates of stock abundance as was the case for Pacific whiting (Dorn et al. 1991). For sablefish incorporation of tag data in the estimation procedure along with survey estimates of relative abundance may assure that regional estimates of abundance are consistent with the migratory character of the population.

Our management application of the model was used to show how a regional per-recruit analysis can be undertaken that enables estimates of reference harvest rates such as $F_{35\%}$ to be computed. Our method is a logical extension of the fishing strategy advocated by Clark (1991) that is used to recommend catch quotas for many groundfish species in the North Pacific Ocean and elsewhere. In Clark's procedure, a fishing mortality rate that reduces the spawning biomass per recruit to 35% of the unfished level was shown to obtain a high proportion of maximum sustainable yield over a range of life-history parameters and stock-recruit relationships. The approach to selecting a region-specific harvest strategy requires the same life-history parameters used in Clark's procedure with the addition of age-specific migration rates.

We applied the regional per-recruit analysis to management of sablefish in the northeast Pacific Ocean. Although the application developed for sablefish was designed primarily as an exploration of the approach, some implications for fishery management were demonstrated. Fishing at a rate that may be appropriate for the pooled population may differentially affect the populations in each region. Region-specific harvest rates that are designed to maintain regional spawning potential are a major departure from the harvest strategy that is currently used to manage sablefish in Alaska. At present, the sablefish fishery is managed by individual fishing quotas (IFQs) which are allocated by region. Adoption of region-specific harvest rates makes biological sense; however, there is the potential for more conflicts among user groups if such fishing rates are used because individual quota shares for sablefish are directly related to fishing rates. Use of the regional specific fishing rates would likely be met with strong resistance by some fisherman and embraced by others.

The region-specific fishing rates that we compute are only appropriate for the assumption that regional recruitment strengths depend on region-specific spawning biomass. Given the high migration rates exhibited by pre-recruit sablefish (McFarlane and Saunders 1997) such an assumption may not be appropriate. In addition, we did not incorporate uncertainty in model parameters, especially movement rates (Heifetz and Fujioka 1991). The analysis is thus an initial step in identifying a region-specific harvest strategy. A more in-depth study would need to consider alternative hypothesis about regional recruitment dynamics, perhaps within a decision analysis framework. These hypotheses could consider larval and juvenile dispersal patterns and incorporate oceanographic conditions.

For sablefish as with many other species, migration is not directly considered in the stock assessment process. For sablefish, catch is apportioned among regions based on an exponentially weighted average of survey estimates of regional-specific relative biomass over time (Sigler et al. 1997). Such a policy adapts to current information, does not require estimates of migration rates for implementation, and reduces the effects of measurement errors associated with regional biomass estimates (Heifetz et. al 1997). As long as reliable estimates of region-specific relative abundance indices are available, then this method is adequate. This method however, does not foster understanding about why changes in regional population abundance occur. In addition, integration of the migratory nature of a population into the stock assessment process enables estimation of regional recruitment strengths and may help assure that regional estimates of abundance are consistent with the migratory character of the population. Thus, to account for annual changes in regional abundance the migratory character of a population should be considered in the stock assessment and management decision process.

In conclusion, recognition of the importance of migration to the dynamics of exploited fish populations will enhance our understanding of how fisheries in different regions may interact with each other. We have presented a model framework that enables analysis of such populations. While not without pitfalls, especially with regard to data requirements, use of such a model may permit a reasonably realistic description of a migratory fish population that can reveal dynamic processes that will otherwise be obscured.

Acknowledgements

We thank Drs. J. Fujioka, L. Haldorson, J. Collie, and B. Norcross for helpful comments on an earlier version of this manuscript. Several suggestions from two anonymous referees improved the manuscript.

References

Beverton, R.J.H., and S.J. Holt. 1957. On the dynamics of exploited fish populations. Fish. Invest. Ser. II, 19. 533 pp.

Caswell, H. 1989. Matrix population models. Sinauer Associates, Inc., Sunderland, MA. 328 pp.

Clark, C.W. 1976. Mathematical bioeconomics. Wiley, New York. 352 pp.

Clark, W.G. 1991. Groundfish exploitation rates based on life history parameters. Can. J. Fish. Aquat. Sci. 48:734-750.

Colvocoresses, J.A., and J.A. Musick. 1983. Species associations and community composition of middle Atlantic bight continental shelf demersal fishes. Fish. Bull. U.S. 82:295-312.

Deriso, R.B., W.H. Bayliff, and R.G. Punsley. 1991. A Markov movement model of yellowfin tuna in the eastern Pacific Ocean and some analyses for international management. Fish. Res. 11:375-395.

Deriso, R.B., T.J. Quinn II, and P.R. Neal. 1985. Catch-age analysis with auxiliary information. Can. J. Fish. Aquat. Sci. 42:815-824.

Die, D.J., V.R. Restrepo, and W.W. Fox. 1990. Equilibrium production models that incorporate fished area. Trans. Amer. Fish. Soc. 119:445-454.

Dorn, M.W., R.D. Methot, E.P. Nunnallee, and M.E. Wilkins. 1991. Status of the Pacific whiting resource in 1990. NOAA Tech. Memo. NMFS F/NWC-204. 97 pp.

Emlen, J.M. 1984. Population biology. Macmillan Publishing Co., New York. 547 pp.

Fujioka, J.T. 1978. A simulation study of dogfish shark populations *(Squalus acanthias)* in Puget Sound. Ph.D. dissertation. University of Washington, Seattle. 108 pp.

Fujioka, J.T. 1995. Sablefish. In: Gulf of Alaska Plan Team (eds.), Stock assessment and fishery evaluation report for the groundfish resources of the Gulf of Alaska as projected for 1996, Section 4. North Pacific Fishery Management Council, P.O. Box 103136, Anchorage, AK 99510.

Getz, W.M., and R.G. Haight. 1989. Population harvesting: Demographic models of fish, forest, and animal resources. Princeton University Press, Princeton, NJ. 391 pp.

Heifetz, J. 1996. Dynamics of a migratory fish population with application to the management of sablefish in the Northeast Pacific Ocean. Ph.D. dissertation, University of Alaska Fairbanks, Juneau. 127 pp.

Heifetz, J., and J.T. Fujioka. 1991. Movement dynamics of tagged sablefish in the northeastern Pacific. Fish. Res. 11:355-374.

Heifetz, J., J.T. Fujioka, and T.J. Quinn II. 1997. Geographic apportionment of sablefish harvest in the northeastern Pacific Ocean. In: M.E. Wilkens and M.W. Saunders (eds.), International Symposium on the Biology and Management of Sablefish. NOAA Tech. Report NMFS 130:229-237.

Lande, R. 1988. Demographic models of the northern spotted owl *(Strix occidentalis caurina).* Oecologia (Berlin) 75:601-607.

Leslie, P H. 1945. On the use of matrices in certain population mathematics. Biometrika 33:183-212.

Liaw, K-L. 1980. Multistate dynamics: The convergence of an age-by-region population system. Environment and Planning A. 12:589-613.

Lowe, S.A., J.T. Fujioka, and J.M. Terry. 1991. Bioeconomic analysis of a minimum size limit for Gulf of Alaska sablefish using a yield per recruit model. Fish. Res. 11:307-320.

MacCall, A.D. 1990. Dynamic geography of marine fish populations. University of Washington Press, Seattle. 153 pp.

May, R.M. 1974. Biological populations with nonoverlapping generations: Stable points, stable cycles, and chaos. Science 186:645-647.

McFarlane, G.A., and M.W. Saunders. 1997. Dispersion of juvenile sablefish, *Anoplopoma fimbria* as indicated by tagging in Canadian waters. In: M.E. Wilkens and M.W. Saunders (eds.), International Symposium on the Biology and Management of Sablefish. NOAA Tech. Report NMFS 130:137-150.

Nisbet, R.M., and W.S.C. Gurney. 1982. Modeling fluctuating populations. Wiley, New York. 379 pp.

Polacheck, T. 1990. Year around closed areas as a management tool. Natural Resource Modeling 4:327-354.

Possingham, H.P., and J. Roughgarden. 1990. Spatial population dynamics of a marine organism with a complex life cycle. Ecology 7:973-985.

Quinn II, T.J., and R.B. Deriso. 1998. Quantitative fish dynamics. Oxford University Press, New York. (in press).

Quinn II, T.J., and N.J. Szarzi. 1993. Determination of sustained yields in Alaska's recreational fisheries. In: G. Kruse, D.M. Eggers, R.J. Marasco, C. Pautzke, and T.J. Quinn II (eds.), Proceedings of the International Symposium on Management Strategies for Exploited Fish Populations. University of Alaska Sea Grant, AK-SG-93-02, Fairbanks, pp. 61-84.

Quinn II, T.J., R.B. Deriso, and P.R. Neal. 1990. Migratory catch-age analysis. Can. J. Fish. Aquat. Sci. 47:2315-2327.

Rogers, A. 1966. The multiregional matrix growth operator and the stable interregional age structure. Demography 3:537-544.

Rogers, A. 1985. Regional population projection models. Sage Publications, Beverly Hills, CA. 96 pp.

Rose, G.A., B. deYoung, and E.B. Colbourne. 1995. Cod (*Gadus morhua* L.) migration speeds and transport relative to currents on the north-east Newfoundland shelf. ICES J. Mar. Sci. 52:903-914.

Shimada, A.M., and D.K. Kimura. 1994. Seasonal movements of Pacific cod *(Gadus macrocephalus)* in the eastern Bering Sea and adjacent waters based on tag-recapture data. Fish. Bull. U.S. 92:800-816.

Sigler, M.F. 1993. Stock assessment and management of sablefish *Anoplopoma fimbria* in the Gulf of Alaska. Ph.D. dissertation, University of Washington, Seattle. 188 pp.

Sigler, M.F., and J.T. Fujioka. 1993. A comparison of policies for harvesting sablefish *(Anoplopoma fimbria)* in the Gulf of Alaska. In: G. Kruse, D.M. Eggers, R.J. Marasco, C. Pautzke, and T.J. Quinn II (eds.), Proceedings of the International Symposium on Management Strategies for Exploited Fish Populations. University of Alaska Sea Grant, AK-SG-93-02, Fairbanks, pp. 7-20.

Sigler, M.F., J.T. Fujioka, and S.A. Lowe. 1997. Sablefish. In: North Pacific Fishery Management Council, Stock assessment and fishery evaluation report for the groundfish resources of the Gulf of Alaska. 605 W 4th Ave., Suite 306, Anchorage, AK 99501.

Swartzman, G.L., W.M. Getz, and R.C. Francis. 1987. Binational management of Pacific hake *(Merluccius productus):* A stochastic modeling approach. Can. J. Fish. Aquat. Sci. 44:1053-1063.

Tyler, A.V. 1971. Periodic and resident components in communities of Atlantic fishes. J. Fish. Res. Board Can. 28:935-946.

Stock Assessment of Gemfish *(Rexea solandri)* in Eastern Australia Using Maximum Likelihood and Bayesian Methods

Anthony D.M. Smith and André E. Punt
CSIRO (Commonwealth Scientific and Industrial Research Organisation)
Division of Marine Research, Hobart, Australia

Abstract

Maximum likelihood and Bayesian assessments for the stock of gemfish off eastern Australia are described. The values of the parameters of an age- and sex-structured population dynamics model are estimated using information on catches, catch rates, the fraction of winter fishery catches consisting of females, and the age-composition and length-frequency of catches. Estimated parameters include virgin biomass, recruitment steepness, annual recruitment anomalies, selectivity, and natural mortality. For the Bayesian assessment, priors were constructed as part of the stock assessment process by a working group involving fishermen, managers, and scientists. The working group also reviewed all data inputs, and determined base-case analyses and sensitivity tests. The latter included sensitivity to weightings on data, levels of historical catches, and the possibility of auto-correlated recruitment anomalies. The results of the Bayesian analyses are used as the basis for risk assessment and evaluation of future harvest strategies for the stock.

Introduction

The gemfish *(Rexea solandri)* resource off eastern Australia is fished mainly by demersal trawling on the upper continental slope in depths ranging from 350 to 500 meters. The fishery commenced in the early 1960s and catches peaked in the early 1980s (Fig. 1). Quota management was introduced in 1988, initially as a competitive quota, and subsequently based on individual transferable quotas. Monitoring of the age and size composition of the catches indicated a series of weak year classes spawned in

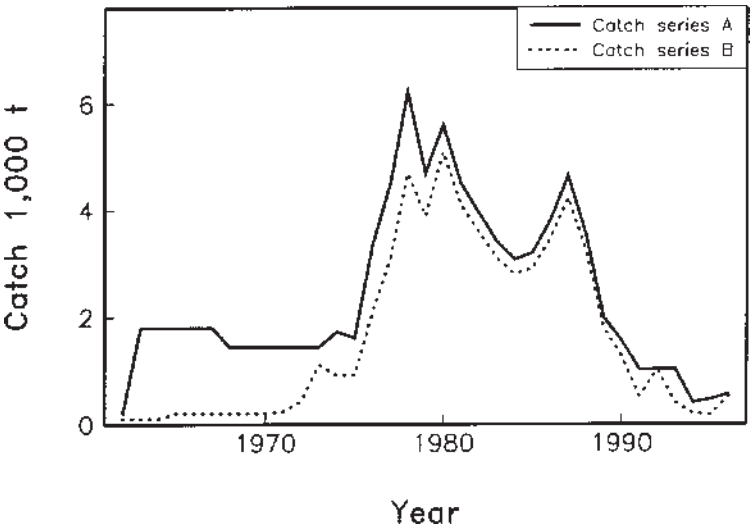


Figure 1. Alternative historical catch series for the winter fishery.

the late 1980s, and a zero TAC was set in 1993 and maintained until 1996. The gemfish fishery is part of the multispecies South East Trawl Fishery, and is managed by the Australian Fisheries Management Authority (AFMA). AFMA set as a criterion for reopening the gemfish fishery that there be a greater than 50% chance that the current biomass (of 5+ males and 6+ females) exceeds 40% of the corresponding biomass in 1979.

Stock assessment for eastern gemfish is currently undertaken by the Eastern Gemfish Assessment Group (EGAG) which was set up in early 1996. This group comprises fishery managers, industry (catching and processing sectors), scientists (government and independent), an economist, and a conservation member. EGAG's principle objectives are to undertake stock assessments, evaluate future harvest strategies, and set research priorities. Previous assessments of eastern gemfish have used a variety of age-structured assessment techniques, including cohort analysis (Allen 1989), CAGEAN (c.f. Deriso et al. 1985), ad hoc tuned VPA (c.f. Pope and Shepherd 1985) and integrated analysis (e.g., Methot 1989, 1990; Punt 1996; Bax 1996). EGAG has adopted an assessment approach based on maximum likelihood and Bayesian methods, as described in this paper. Industry and management members of EGAG, together with the scientific members, have played a critical role in the assessment process through review of data, identification of key assumptions and uncertainties, and in constructing prior distributions for the parameters of the Bayesian assessment.

The assessment is conducted using an age- and sex-structured population dynamics model. The values for the parameters of this model which are not determined from auxiliary information are obtained using data on catch rates, the fraction of females in the winter fishery catches and information about the age-composition and length-frequency of the catches. The results of the maximum likelihood analyses are used to (a) identify the sensitivity of important management-related outputs (e.g., current depletion) to the data and model specifications, (b) understand the behavior of the model, (c) assess whether the model fits the data adequately, and (d) assess the impact of the various data sources on the assessment results. The Bayesian results provide a more suitable basis for evaluating the consequences of alternative harvest strategies because the Bayesian approach takes greater account of the uncertainty about the fit of the model to the data.

The methodology upon which the maximum likelihood analyses are based is similar to the “stock synthesis” approach (Methot 1989, 1990) while the Bayesian analyses use an extension of the methodology described in McAllister et al. (1994), Punt et al. (1994) and Punt (1996). Although Bayesian stock assessment methods are relatively new in fisheries (for instance, see Hilborn et al. 1994, McAllister et al. 1994, Walters and Ludwig 1994, Givens et al. 1995, Punt and Hilborn 1997, McAllister and Ianelli 1997), they allow the analyst to incorporate (subjective) prior distributions for model parameters, and provide a consistent framework for decision analysis.

Methods

Population Dynamics Model

The population dynamics model considered in this analysis (Appendix I) is age- and sex-structured and takes account of ages up to 12 years, with ages over 12 being pooled into a plus-group. The trawl fishery for gemfish comprises a summer and a winter component, the latter targeting a prespawning run northward along the east coast. The summer and winter fisheries are modeled as pulses at the start of December and at the end of June respectively. The “year” in this paper runs from July to June because spawning occurs during the winter fishery. The number of 0-year-olds at the start of the year (July) is taken to be related to the egg production at the end of the preceding year by a Beverton and Holt stock-recruitment relationship, which allows for lognormally distributed recruitment anomalies. (The term “recruitment anomaly” is used to refer to the logarithm of the relative difference between the estimated number of births [also referred to as the year-class strength] and the value predicted from the [deterministic] stock-recruitment relationship.) The sex ratio at birth is taken to be 1:1.

The model assumes that the population was distributed about its unexploited equilibrium level at the start of 1962 when the first substantial catches are assumed to have been taken during the winter spawning run fishery (see Punt et al. 1997). The selectivity pattern for the winter fishery is assumed to follow a logistic curve, while that for the summer fishery is assumed to be dome-shaped. The assumption of sex-specific natural mortality rates is needed to mimic the age-composition of the catches by sex.

Parameter Values

Table 1 lists the values for the fixed model parameters (i.e. those taken to be known exactly). The values for the growth parameters were estimated from data collected during the winter spawning run (Rowling 1990) and so pertain (roughly) to the start of July. In order to calculate length-at-age for the summer fishery, the value for a in equation (I.6) is replaced by $a - 0.583$ (0.583 is the fraction of a year between the start of December [the assumed middle of the summer fishery] and the end of June [the assumed middle of the winter fishery]). The CVs for length-at-age are computed by linear interpolation from those listed in Table 1. Table 2 lists the priors chosen by EGAG for the remaining parameters (EGAG 1996).

Data and the Likelihood Function

The data used in the stock assessment include catches by fishery since 1962, the fraction of the winter catch (in number) which consists of females, catch rate based indices of relative abundance, length-frequency data (in 2 cm size-classes), and catch-at-age data (Punt et al. 1997). Catch rates are assumed to be proportional to abundance, based on previous results supporting this assumption (Allen 1989).

A major uncertainty in the assessment identified by EGAG is the level of historical catches, particularly catches in the early period of development of the fishery. EGAG identified two historical catch series which represent plausible upper and lower bounds (Fig. 1), but was unable to agree on the relative credibility of these catch series. In this paper, the higher catch series (series A) is used for the reference analysis. The two catch series, and other data inputs to the analysis, are described in Punt et al. (1997).

Length-frequencies are available for some of the years for which age-length keys are missing as well as being available for all of the years for which age-composition data are available. If age-composition data are available, these are used in preference to length-frequency data and length-frequencies by sex are used in preference to unsexed length-frequencies. Appendix II details the contributions of the various data sources to the negative of the logarithm of the likelihood function.

When fitting the model, it is necessary to provide specifications for the catchability coefficient, q, and the residual standard deviation for catch rate, σ_c. For the maximum likelihood analyses, the estimate of q is obtained

Table 1. The values for the fixed parameters of the population dynamics model.

Parameter	Male	Female	Both sexes
L_∞(cm)[a]	97.5	109.4	
κ (yr^{-1})[a]	0.212	0.180	
t_0 (yr)[a]	–0.54	–0.63	
CV of length-at-age[b]	0.066 (age 3)	0.050 (age 4)	
	0.046 (age 9)	0.044 (age 11)	
b_1 (gm cm^{-3})[b]			0.143×10^{-5}
b_2[b]			3.39
plus-group = x (years)[c]			12
σ_r[c]			0.6

[a] Rowling and Reid (1992).

[b] K. Rowling, NSW Fisheries, Sydney, Australia.

[c] EGAG.

Table 2. The prior distributions specified by EGAG for the model parameters.

Parameter	Male	Female	Both sexes
B_0			U[5,000 t, 50,000 t]
M (yr^{-1})	U[0, 0.8]	U[0, 0.8]	
Steepness = h			See Fig. 2
L_{full} = summer fishery			$\text{U}[L_2^1, L_4^1]$
$L_{50\%}$ = summer fishery			$\text{U}[L_2^1, L_{10}^1]$
L_{50} = winter fishery	$\text{U}[L_4^{2,m}, L_6^{2,m}]$	$\text{U}[L_4^{2,f}, L_6^{2,f}]$	
L_{95} = winter fishery	$\text{U}[L_4^{2,m}, L_8^{2,m}]$	$\text{U}[L_4^{2,f}, L_8^{2,f}]$	
$\ell\text{n}q$			U[$-\infty$, ∞]

analytically (equation II.3). The prior for q is taken to be uniform on a log-scale for the Bayesian analyses. This choice of prior can be shown to be non-informative for B_0 (Pikitch et al. 1993) although it may not be non-informative for other quantities of management interest (Cordue and Francis 1994). The value for σ_c is pre-specified instead of being assigned a prior as has been the case in several other Bayesian assessments (e.g., McAllister et al. 1994, Givens et al. 1995).

The values assumed in the Bayesian analysis for the residual standard deviations are 0.15 for σ_c and 0.1 for σ_f, σ_a, σ_s, and σ_l. These choices are similar to the maximum likelihood estimates for the corresponding residual standard deviations based on preliminary fits to the data. The contributions to the likelihood function are multiplied by "weighting factors" to account for relative (a priori) perceptions of the reliability of the various data sources. The weights for the age-composition, fraction female, and winter length-frequency data are set equal to 1 while that for the summer length-frequency data is set equal to 0.1. This lower weight is based on the observation that selectivity to the summer fishery varies markedly from year to year. This violates the assumption underlying equation (II.7) that all of the error is observational. The assumption of a lower weight for the summer length-frequency data accounts for this to some extent. The weight assigned to the catch rate data is varied until the estimated residual standard deviation for these data is 0.15. For the reference analysis, this weight is 3.0.

The contributions to the likelihood function by the age-composition and length-frequency data are based on the assumption that the observed data are lognormally distributed about their expected values with a CV which depends on the inverse of the square root of the expected proportion (see Section C of Appendix II). The justification for this approach is dicussed in Punt (1997a). This specification differs from the possibly more conventional assumption of a multinomial distribution (e.g., McAllister and Ianelli 1997).

The age-composition data included in the likelihood function are restricted to males older than 4 years and females older than 5 years. The length-frequency information for males is restricted to animals of 56 cm and larger while the length-frequency information for females is restricted to animals larger than 64 cm. The 64 cm restriction also applies to unsexed length-frequencies for the winter fishery and the fraction female data are defined in terms of animals larger than 64 cm. The model-estimates of the fraction of the catch falling into each age-class take account of the possibility of age-reading error (see equation II.9). The length-frequency and age-composition data that are included in the likelihood function are restricted to large and/or old animals, because early maturation may have occurred during the period of poor recruitment, and because the length-frequencies for some recent years represent bycatch, which may be biased toward small fish.

The performance of the estimator deteriorates markedly if the likelihood includes contributions for ages and lengths that constitute only a small fraction of the catch. The age-composition data for animals age 10 years and older are pooled at age 10 and the winter length-frequency data for animals larger than 80 cm (males) and 94 cm (females and unsexed) are pooled to remove this problem. The length-frequency data for the summer fishery are restricted to the range 36-78 cm for the same reason. The data for the summer fishery are restricted further by omitting those years in which the sample size is less than 1,000 (1976, 1985, 1987, 1990, 1991, and 1992). This restriction is imposed because the likelihood in Section C of Appendix II gives the same weight to each year irrespective of the number of animals measured.

Estimation Procedures

The prior distributions listed in Table 2 are incorporated into the maximum likelihood analyses to enhance the comparability of these and the Bayesian analyses. This is achieved by using the limits of the uniform distributions in Table 2 as bounds for the estimates of the model parameters; if a parameter is outside its bounds (for example, if B_0 exceeds 50,000 t), the negative log-likelihood is set to a very large value. The prior distribution for steepness was included in the maximum likelihood analyses by subtracting the logarithm of the prior probability for the value of steepness (Fig. 2). The lognormal prior distributions for the recruitment anomalies (both those which constitute the initial conditions and those which apply to births after 1962) are included in the maximum likelihood analyses by adding the following component to the negative of the log-likelihood function:

$$\sum_y \frac{\varepsilon_y^2}{2\sigma_r^2} \tag{1}$$

where the summation is taken over the 46 recruitment anomalies, and σ_r is assumed to be 0.6 (see Table 1). The choice $\sigma_r = 0.6$ was made because 0.6 is close to the median of the estimates of σ_r for teleost species obtained by Beddington and Cooke (1983).

This treatment of the recruitment anomalies is equivalent to treating the year class strengths (in absolute terms) as parameters while at the same time allowing for a stock-recruitment relationship (e.g., Fournier and Archibald 1982; Methot 1989, 1990). The implication of the inclusion of equation (1) in an analysis is that a recruitment anomaly will be assumed to be zero (i.e. recruitment is equal to the prediction from the deterministic component of the stock-recruitment relationship) unless there are data which indicate otherwise. Another way of interpreting the results from the maximum likelihood analyses is that they are Bayesian modal

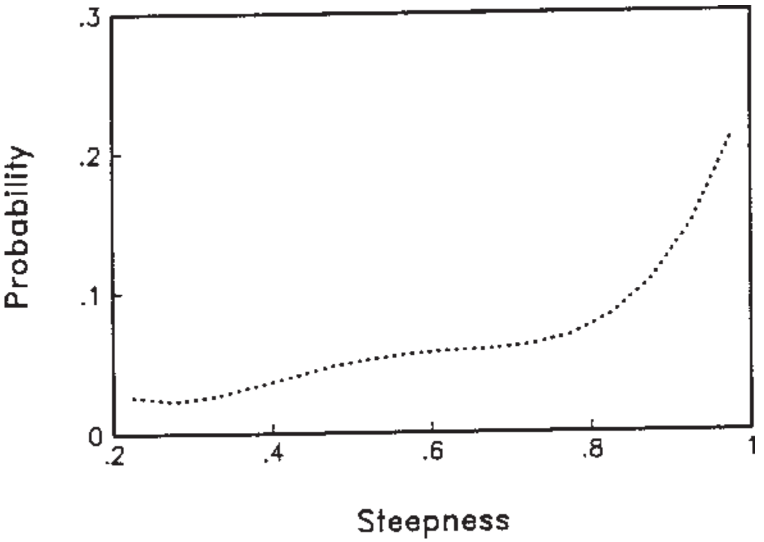


Figure 2. The prior distribution assumed for the steepness of the Beverton and Holt stock-recruitment relationship (Pers. comm., J. Ianelli, National Marine Fisheries Service, Seattle).

results. This is because the parameter estimates from the maximum likelihood analyses correspond to the maximum of the posterior distribution of the Bayesian analysis.

In summary then, for the reference maximum likelihood analysis, the values for 57 parameters (see Table 3) are obtained by minimizing the sum of the negative of the log-likelihood function, the negative of the prior probability for steepness, and equation (1), while the posterior distributions for these parameters are obtained by integrating the likelihood function over the prior distributions. The Sample-Importance-Resample (SIR) method (see Appendix III) is used to perform the numerical integration needed to calculate the posterior distribution for the Bayesian calculations.

Alternative Analyses

A large number of sensitivity tests was specified by EGAG and undertaken for the maximum likelihood analyses (Table 4). The results are presented elsewhere (see Punt 1997b). The number of Bayesian analyses has been restricted to five due to the computational demands of conducting the Bayesian integrations numerically. Catch series A is used for the reference analysis and for three of the four other analyses.

Table 3. The estimated parameters of the model.

Parameter type	Details	# Parameters
Population-model-related	B_0, M^f, M^m, h	4
Selectivity	L_{full}, $L_{50\%}$, L^s_{50}, L^s_{95}	6
Catchability	q	1
Recruitment anomalies	$\{\varepsilon_a : a = 0,1,2,\ldots,x-1\}$	12
	$\{\varepsilon_y : y = 1963,64,\ldots,95,96\}$	34

Table 4. Hypotheses considered as part of the assessment of the eastern stock of gemfish.

No	Description
	Data-related sensitivity tests
D1	Changing the weight assigned to the CPUE data
D2	Changing the time series of historical catches
D3	Changing the weight assigned to the age composition data
D4	Changing the weight assigned to the length-frequency data
D5	Changing the vessel used when standardizing the catch and effort information
D6	Ignoring the age-composition data
D7	Ignoring the summer length-frequency data
D8	Ignoring the winter length-frequency data
D9	Ignoring the fraction female information
D10	Analyzing the catch and effort data using a GLM approach
	Model-related sensitivity tests
M1	Changing the form of the stock-recruitment relationship to a Ricker or the Thompson (1993) form
M2	Changing the assumed extent of variation in the recruitment anomalies, σ_r
M3	Allowing for autocorrelation in the recruitment anomalies
M4	Allowing for density-dependent selectivity
M5	Allowing for density-dependent growth
M6	Ignoring age-reading error

The five Bayesian analyses are:

a. The reference analysis (abbreviation “Catch series A”).

b. Replacing catch series A by catch series B (abbreviation “Catch series B”).

c. Decreasing the weight assigned to the daily catch rate data from 3.0 to 1.5 (abbreviation “$w_{CPUE} = 1.5$”).

d. Increasing the weight assigned to the daily catch rate data from 3.0 to 4.5 (abbreviation “$w_{CPUE} = 4.5$”).

e. Allowing for auto-correlation among the recruitment anomalies (abbreviation “Correlated recruitment”).

The weight assigned to the catch rate data for catch series B is altered from 3.0 (the value for the reference analysis) so that the fit to the catch rate data is comparable to that for catch series A. This is to enhance comparability between these analyses. Analyses (c) and (d) explore the effect of decreasing and increasing the weight on the catch rate data for the reference analysis.

The reference analysis suggests that there is considerable covariance between successive recruitment anomalies—this is hardly surprising if the reason for the recruitment anomalies is the impact of some (autocorrelated) environmental variable (or variables). However, the (joint) prior distribution for the recruitment anomalies (equation 1) assumes that these anomalies are independent and this could potentially bias the results. A sensitivity test examines the consequences of replacing the contribution of the recruitment anomalies to the likelihood function (equation 1) by an expression which allows for the possibility of inter-annual correlation in recruitment anomalies (see Appendix IV for details):

$$\frac{1}{2\sigma_r^2}\begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_i \\ \vdots \\ \varepsilon_{47} \end{pmatrix}^T \begin{pmatrix} 1 & \tau & \dots & \tau^{45} & \tau^{46} \\ \tau & 1 & \ddots & \tau^{44} & \tau^{45} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \tau^{45} & \tau^{44} & \ddots & 1 & \tau \\ \tau^{46} & \tau^{45} & \dots & \tau & 1 \end{pmatrix} \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_i \\ \vdots \\ \varepsilon_{47} \end{pmatrix} \qquad (2)$$

where τ is the correlation between recruitment anomalies for adjacent years.

The extent of autocorrelation for the “Correlated recruitment” analysis was set equal to the estimate of autocorrelation among the recruitment anomalies for the reference maximum likelihood case (which is based on the assumption $\tau = 0$). This was estimated to be $\tau = 0.5$.

Exploration of the consequences of future harvest strategies is limited to making projections for the stock under future fixed levels of catch. These projections are based on 200 draws from the posterior distributions for the Bayesian analyses. Future recruitments are based on the lognormal distribution of anomalies about the levels predicted by the deterministic stock recruitment relationship. In the case of autocorrelated recruitment anomalies, the autocorrelation structure also applies to future recruitment anomalies.

Results

Three different biomass series are computed:

a. *The winter biomass* is the biomass of selected animals in the middle of the winter fishery (see equation I.17).

b. *The spawner biomass* is the selected biomass of females at the end of the year (i.e. after the winter fishery) (see equation I.4).

c. *The 5+/6+ biomass* is the biomass at the end of the year of all males aged 5 and older and all females aged 6 and older:

$$E_y = \sum_{a=5}^{x} w_a^m \tilde{N}_{y,a}^m + \sum_{a=6}^{x} w_a^f \tilde{N}_{y,a}^f \tag{3}$$

The management-related statistics used to summarize the results are:

B_0	the unexploited equilibrium winter biomass.
M^m	the (age-independent) rate of natural mortality for males.
M^f	the rate of natural mortality for females aged 3 and older (see equation I.7).
h	the "steepness" of the stock-recruitment relationship.
B_{96}/B_0	the ratio of the winter biomass during 1996 to the unexploited equilibrium level (expressed as a percentage).
E_{96}/E_{79}	the ratio of the 5+/6+ biomass at the end of 1996 to that at the end of 1979 (expressed as a percentage).

The closeness of fit of the model to the data is summarized by the estimates of the residual standard errors for each data type.

The largest weight assigned to a single parameter combination in the posterior distribution is 0.066 (catch series A) and 0.061 (catch series B) which implies that an adequate numerical representation of the posterior has been achieved. Of the 10,000 parameter combinations used to construct Figs. 7-9, 3,769 were unique for catch series A and 4,738 for catch series B. The largest weights are larger than those obtained by McAllister and Ianelli (1997) for the assessment of the yellowfin sole *(Limanda aspera)* resource in the eastern Bering Sea and notably larger than those obtained by McAllister et al. (1994) for the assessment of the New Zealand's western stock of hoki *(Macruronus novaezelandiae).* This is not unexpected because the current assessment is based on considerably more data (the yellowfin sole assessment was based on relative abundance data and sex-aggregated age composition data while the hoki assessment was based solely of relative abundance data).

Maximum Likelihood Analyses

Results for the maximum likelihood analyses are shown in Tables 5 and 6 and in Figs. 3 to 6. As noted earlier, these are not true maximum likelihood estimates, but are posterior modal estimates for the Bayesian analyses. They incorporate the prior distributions used in the Bayesian analyses, to enhance comparability between estimation methods. Table 5 shows summary results for the reference analysis (catch series A) and for the catch series B sensitivity test. The model fits the data reasonably well. The weightings on the CPUE data have been set to achieve a residual CV for those data of approximately 15%. The fit to the fraction female data is good ($\sigma =$ 0.07), while the age data are better fitted for females than males. The poorest fit is to the summer size data ($\sigma = 0.17$), which are given a low weighting in the analysis because of the variable nature of this mainly bycatch fishery. The fit to the catch rate data is shown in Fig. 4. The fits to the two catch series are very similar.

Table 5 and the upper panels of Fig. 3 indicate that the biomass was reduced rapidly from 1962 as a consequence of the catches (particularly for the reference analysis) and a string of weak year classes. Lower catches between 1968 and 1973 combined with a series of good year classes from 1971 led to an increase in biomass after 1972. After peaking in 1977, the winter biomass declined under the impact of catches in excess of 3,000 t. Although the catches after 1983 are somewhat lower than those between 1977 and 1982, there is no increase in biomass owing to another sequence of poor year classes. The winter biomass is estimated to have declined continuously until 1994 and to have recovered somewhat primarily as a result of the strong 1990 year class. The current (winter) biomass is estimated to be at about 22% of the unexploited equilibrium level for the reference analysis. This corresponds to about 55% of the 1979 5+/6+ biomass, which is higher than the management target of 40% for reopening of the fishery. For catch series B, the estimated level of unfished biomass B_0 is considerably lower than for catch series A. This is not unexpected given the lower historical catches. The estimated level of depletion is not as great (32% compared with 22% for catch series A), but the level of 5+/6+ biomass relative to 1979 is similar.

For both catch series, the natural mortality rate for males is larger than that for females. This is a consequence of the age-composition data in which males dominate the catch of the younger (<5-year-old) individuals while females comprise the bulk of the catch of older (>7-year-old) animals. The estimate of steepness for the reference analysis is 0.504.

Table 5. Summary statistics corresponding to the mode of the posterior distribution (maximum likelihood).

Quantity	Catch series A	Catch series B
Estimates of precision		
CPUE	0.148	0.151
Fraction female	0.070	0.068
Age data (males)	0.131	0.131
Age data (females)	0.097	0.096
Size data (summer)	0.170	0.169
Size data (winter, unsexed)	0.090	0.090
Size data (winter, male)	0.108	0.107
Size data (winter, females)	0.096	0.095
B_0	20337	14555
M^m	0.486	0.593
M^f	0.317	0.409
h	0.504	0.379
B_{96}/B_0	0.225	0.321
E_{96}/E_{79}	0.547	0.514

Results are shown for analyses based on the two alternative series of catches.

Table 6. Posterior correlations among four of the parameters of the population dynamics model based on a normal approximation to the posterior.

Parameter	B_0	M^m	M^f	h
B_0		–0.061	–0.075	–0.238
M^m	0.089		0.951	–0.598
M^f	0.077	0.940		–0.614
h	–0.257	–0.497	–0.507	

Results are shown for analyses based on catch series A in the upper triangular and for catch series B in the lower triangular.

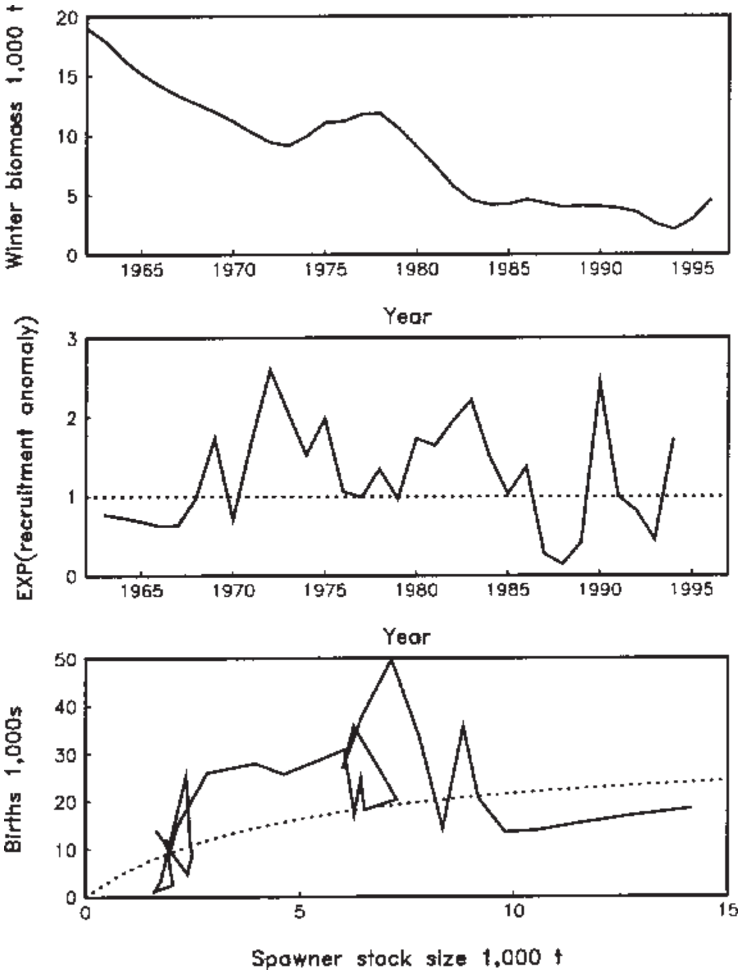


Figure 3. Time trajectories of winter biomass, EXP(recruitment anomaly), the spawner-birth information, and the estimated stock-recruitment relationship, corresponding to the mode of the posterior distribution (the maximum likelihood analysis). Results are shown for catch series A (this page) and for catch series B (facing page).

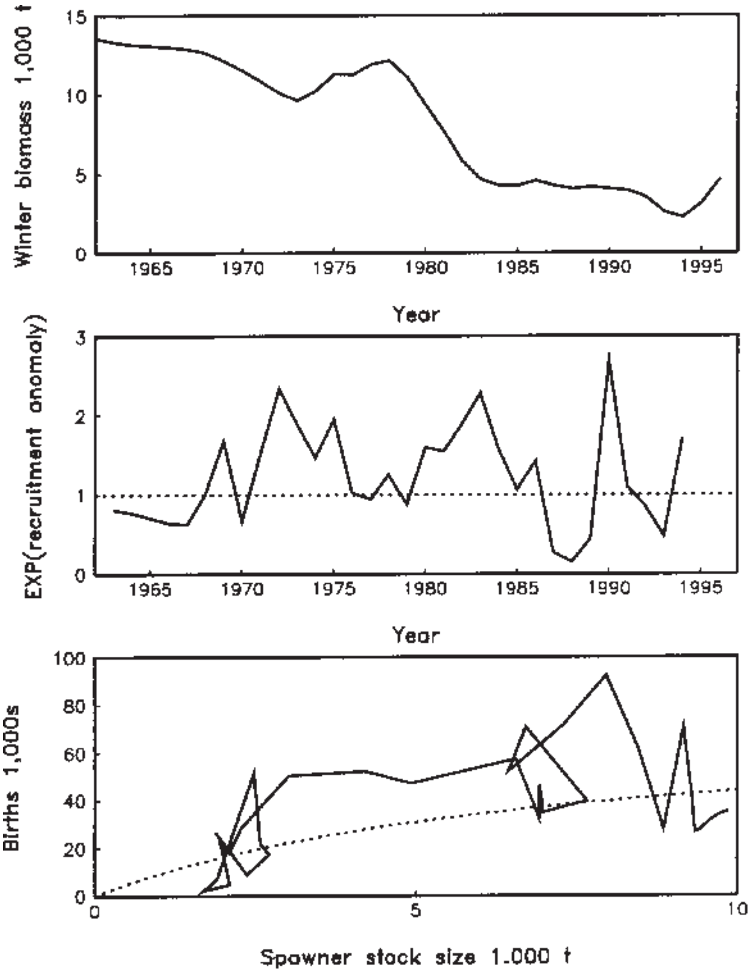


Figure 3. (Continued.)

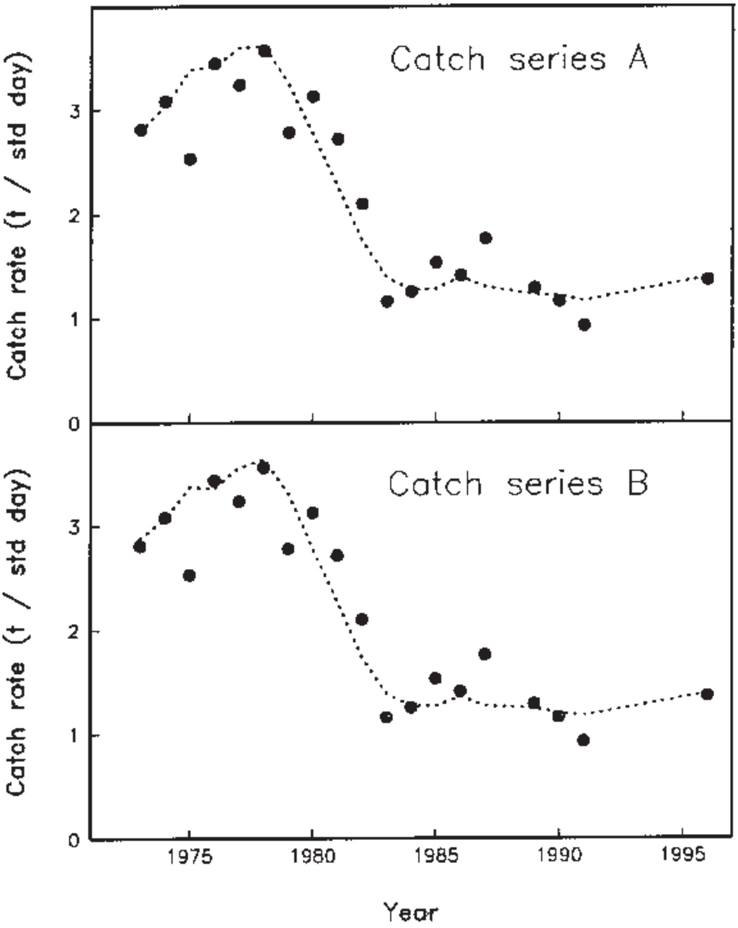


Figure 4. Observed (solid dots) and model-predicted (dotted lines) catch rates.

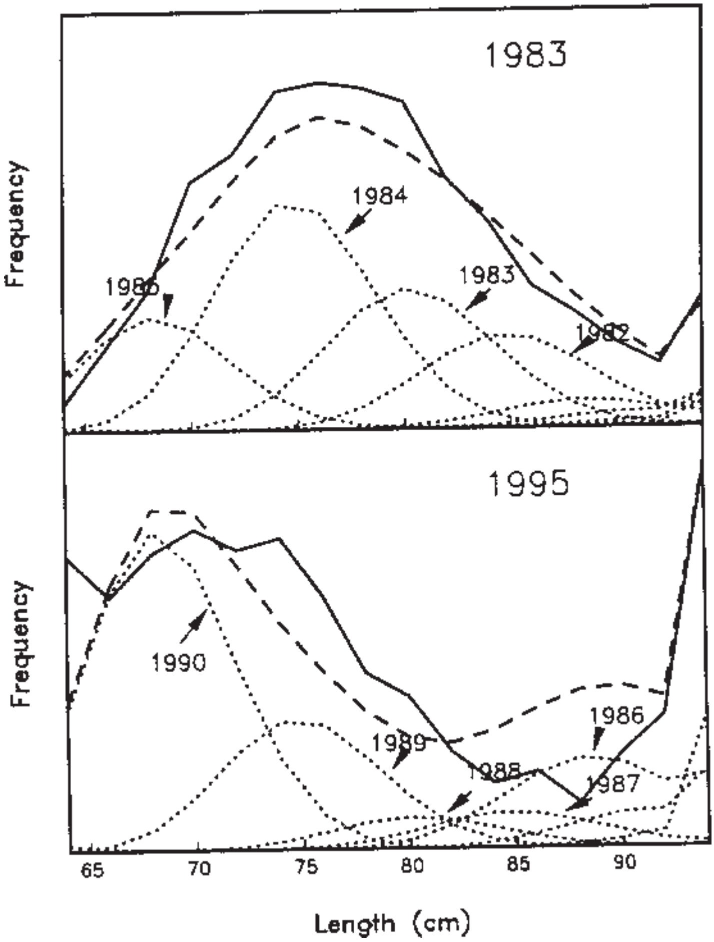


Figure 5. Observed (solid lines) and model-predicted (dashed lines) female length-frequency distributions for the winter fishery for 1983 and 1995. The catch length-frequency distributions for the year classes in the population are also shown. The model predictions are based on the posterior mode for catch series A.

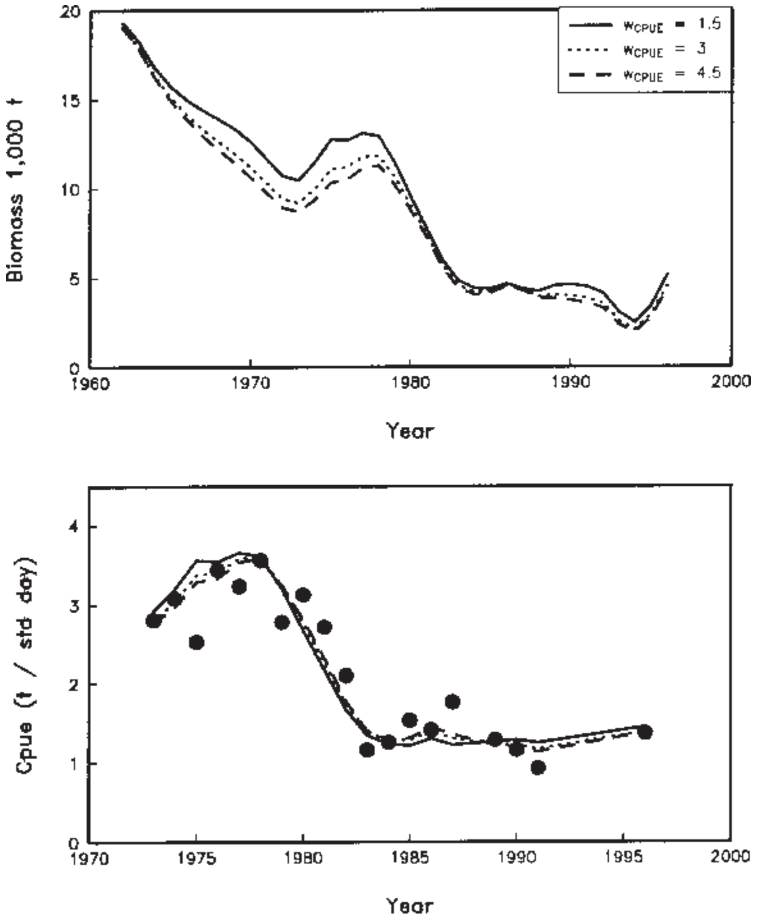


Figure 6. Time-trajectories of winter biomass, and fits to the catch rate information, for analyses based on catch series A which differ in the weight assigned to the CPUE data.

This is quite a low level considering the prior in Fig. 2 (which has a mean of 0.735) and suggests that this is a relatively unproductive stock. For catch series B, the estimate of steepness is even lower, but the estimates for male and female natural mortality are higher. Table 6 shows the correlations among four of the parameters. There is a very high positive correlation between the sex-specific natural mortalities, and a high negative correlation between natural mortality and steepness.

The middle panels of Fig. 3 show the time series of recruitment anomalies. One feature evident in these time series is the sequences of below average and then above average recruitment anomalies. Of particular note are the runs of above expected recruitments from the early 1970s to the mid-1980s, and the extremely poor recruitments in the late 1980s. The 1990 year class is estimated to be very much larger than that predicted by the stock-recruitment relationship. It currently forms the mode of the length-frequency distribution for the winter fishery (Fig. 5). The current size distribution of the catch is notably different from that in the early 1980s (contrast the 1995 and 1983 size frequency distributions in Fig. 5). Note that, although the 1990 year class is much larger than expected, it is no stronger in absolute terms than the average of the sizes of the year classes during the 1970s and early 1980s (Fig. 3 lower panel). The 1992 and particularly the 1993 year classes are estimated to be poorer than expected while the 1994 year class is estimated to be better than expected.

The effects on the reference analysis of increasing and decreasing the weight put on the catch rate data are shown in Fig. 6. Decreasing the weight tends to result in higher predicted stock sizes and vice versa. Not surprisingly, increasing the weight assigned to the catch rate data improves the fit to these data. Qualitatively, however, the results are little changed from those for the reference analysis.

Bayesian Analyses

The results of the Bayesian analyses are shown in Table 7 and Figs. 7 to 9 for the reference analysis and for catch series B. Other sensitivities are considered in Table 8.

Table 7 lists for several parameters the posterior means, 90% intervals, and the ratios of the posterior variances to those of the priors. For most of the parameters of the model, the prior distributions are updated markedly by the data. The main exception is the selectivity parameter a_{full} for the summer (bycatch) fishery, for which the age composition data are largely non-informative. In terms of means, the results are qualitatively (and quantitatively) similar to those for the reference maximum likelihood analysis shown in Table 5. This implies that the results of the Bayesian analysis are dominated by the data rather than by the choice of priors. Figure 7 shows the prior, post-model–pre-data, and posterior distributions for B_0 and the implicit prior and posterior for B_{96}/B_0. The post-model–pre-data distribution is different from the prior because model projections

Table 7. Results for the Bayesian analysis showing posterior mean and lower and upper 5% points corresponding to several parameters and model-derived quantities. Also shown are the ratios of posterior to prior variances for parameters.

(a) Catch series A

Quantity	Posterior mean	Variance ratio (%)	Lower 5% point	Upper 5% point
B_0	18920	7.7	14360	25340
B_{96}	4944	–	4148	5938
B_{62}/B_0	1.071	–	0.754	1.474
B_{96}/B_0	0.269	–	0.185	0.355
E_{96}/E_{79}	0.544	–	0.458	0.631
h	0.479	15.6	0.350	0.631
M^m	0.534	4.3	0.450	0.600
M^f	0.355	2.6	0.290	0.403
a_{full}	2.590	42.4	2.051	3.263
$a_{50\%}$	6.245	2.9	4.992	7.798
a_{50} (males)	4.760	2.0	4.635	4.888
a_{95} (males)	6.325	4.8	5.943	6.762
a_{50} (females)	5.596	1.6	5.475	5.713
a_{95} (females)	6.839	4.1	6.489	7.253

(b) Catch series B

Quantity	Posterior mean	Variance ratio (%)	Lower 5% point	Upper 5% point
B_0	14390	5.4	10270	19370
B_{96}	5009	–	4069	5891
B_{62}/B_0	1.104	–	0.752	1.582
B_{96}/B_0	0.362	–	0.240	0.481
E_{96}/E_{79}	0.517	–	0.429	0.615
h	0.370	11.9	0.238	0.506
M^m	0.620	2.4	0.564	0.687
M^f	0.431	1.6	0.385	0.480
a_{full}	2.758	69.7	2.066	3.528
$a_{50\%}$	7.169	4.8	5.527	9.184
a_{50} (males)	4.799	2.0	4.655	4.932
a_{95} (males)	6.332	4.6	5.969	6.796
a_{50} (females)	5.606	2.0	5.457	5.744
a_{95} (females)	6.810	4.9	6.335	7.211

Table 8. Results for the Bayesian analyses of 20-year projections at fixed catch levels, for the reference assessment and four variants thereof.

(a) The probability that the winter biomass exceeds B_{MSY} at the end of the projection period.

	Catch level			
Scenario	0 t	500 t	1,000 t	1,500 t
Reference	0.970	0.760	0.395	0.110
$w_{CPUE} = 1.5$	0.890	0.665	0.305	0.070
$w_{CPUE} = 4.5$	0.970	0.700	0.305	0.030
Catch series B	0.865	0.510	0.150	0.030
Correlated recruitment	0.940	0.700	0.315	0.115

(b) Lowest value that the 5+/6+ biomass drops to over the 20-year projection period, expressed as the median for E_y/E_{79}.

	Catch level			
Scenario	0 t	500 t	1,000 t	1,500 t
Reference	0.546	0.546	0.376	0.011
$w_{CPUE} = 1.5$	0.549	0.490	0.366	0.011
$w_{CPUE} = 4.5$	0.543	0.493	0.317	0.002
Catch series B	0.521	0.403	0.083	0.001
Correlated recruitment	0.549	0.549	0.423	0.002

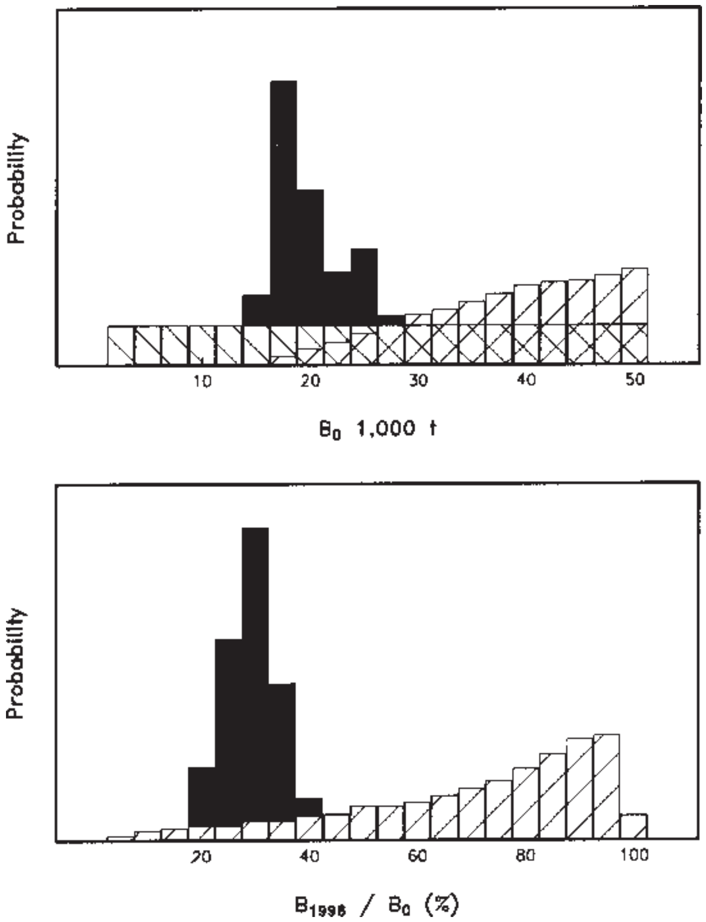


Figure 7. Post-model–pre-data (open bars) and posterior (solid bars) distributions for virgin biomass and current (1996) depletion. The prior distribution for B_0 is indicated in the upper panels. Results are shown for catch series A (this page) and for catch series B (facing page).

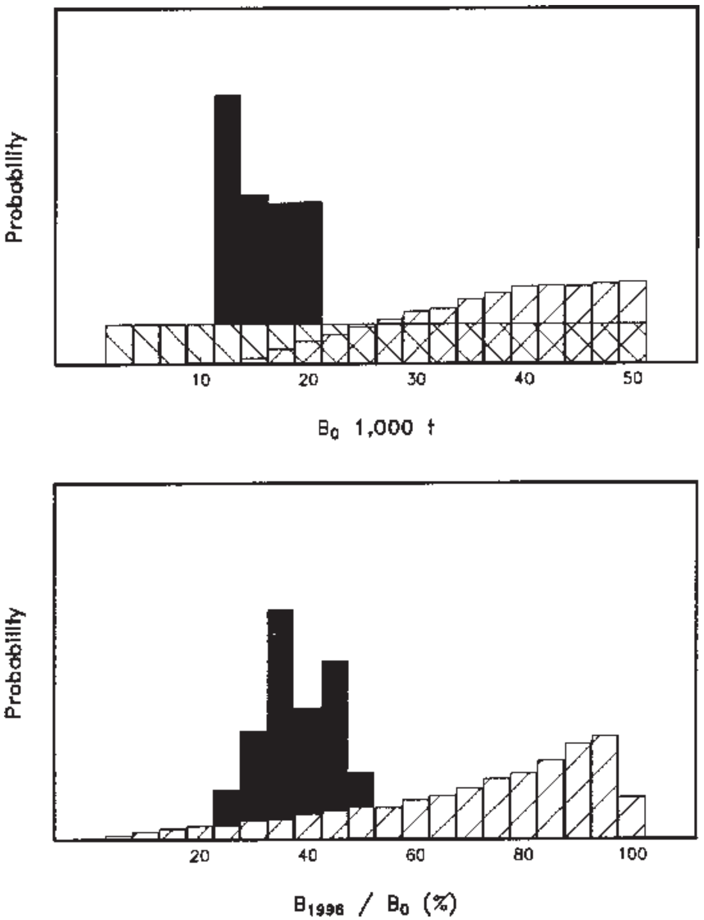


Figure 7. (Continued.)

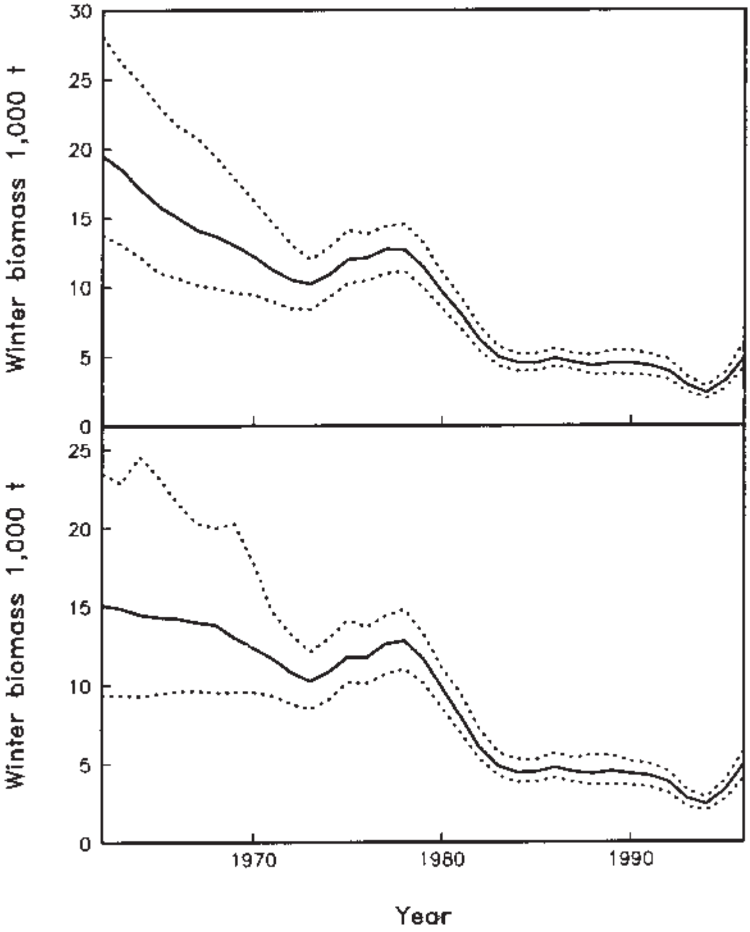


Figure 8. Medians and 90% probability intervals for winter biomass for catch series A (upper panel) and catch series B (lower panel).

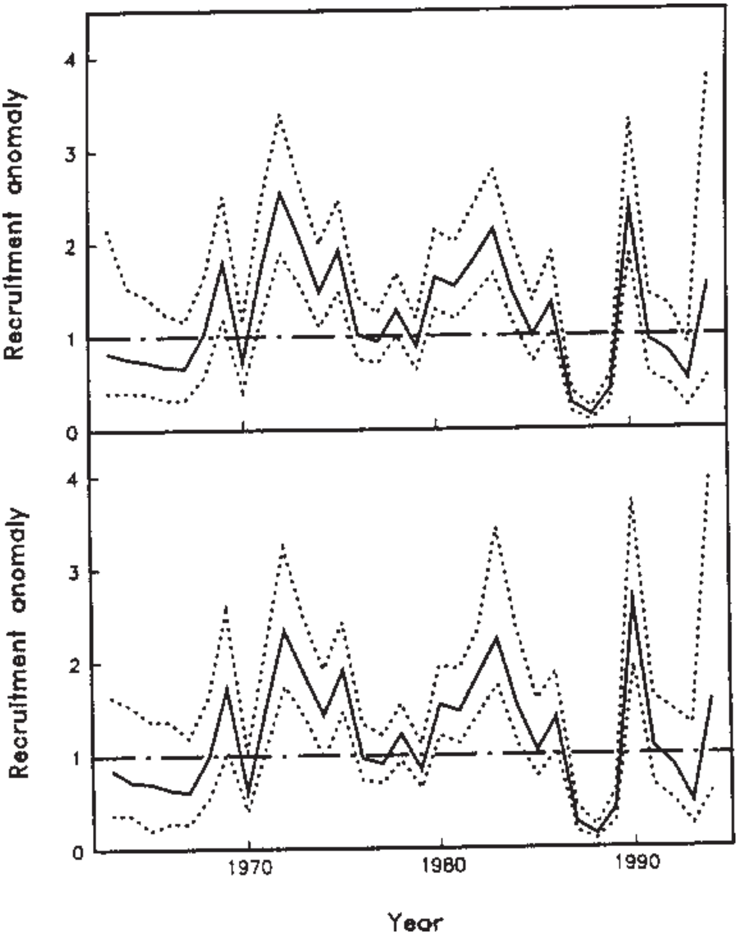


Figure 9. Medians and 90% probability intervals for the time series of EXP(recruitment anomaly) for catch series A (upper panel) and catch series B (lower panel).

from 1962 to 1996 which result in stock extinction are given zero likelihood.

Figure 8 shows the time series for winter biomass for catch series A and B for the Bayesian analysis. Note that what are shown are not trajectories in the normal sense of the word, but rather the medians and 90% probability intervals of the posterior distributions for the population size for each successive year. The winter biomass for the pre-1970 period is much more uncertain than for the later years. This is because catch rate data are only available from 1973, and catch composition data from 1975. Figure 9 shows time series for the expected (multiplicative) recruitment anomalies, again as posterior distributions for each year rather than as trajectories. The patterns are very similar to the maximum likelihood results (Fig. 3), while the 90% probability intervals for the early and most recent recruitment anomalies are very wide. This is because the available data do not provide information about these recruitments, so that the posteriors approximate the priors in these years. However the posterior distributions for the 1970s and 1980s suggest that the sequences of strong and then weak year classes are significant.

Table 8 shows the results of 20-year fixed catch level projections for the reference analysis and four variants. Part (a) of Table 8 shows the probability that the biomass at the end of the 20 years exceeds B_{MSY}. B_{MSY} is the biomass corresponding to the maximum yield assuming no recruitment variability, and is computed for each projection. In general, the reference analysis is the most optimistic, while the most pessimistic outcomes occur if catch series B is assumed to be correct. At higher levels of catch, the correlated recruitment scenario is the most optimistic. Levels of catch in excess of 500 t achieve satisfactory outcomes for this indicator for all variants. Part (b) of Table 8 indicates the lowest level to which the biomass is reduced over the 20 years, expressed as the median for the ratio E_y/E_{79} for 5+/6+ biomass. Most of the variants do not meet the management criterion even at 500 t levels of catch. This is because the stock is predicted to decline further during 1997 as the strong 1990 year class starts to die out, and is followed by several more weak year classes.

Discussion

The Bayesian analysis described and developed in this paper is similar to the approach developed by McAllister and Ianelli (1997). It incorporates a wider set of data types than that analysis, including length frequency and proportion female, but uses similar methods for the Bayesian integration and estimates a similar set of parameters. A novel feature of this analysis is the incorporation of autocorrelation in the recruitment residuals for one of the sensitivity tests. There is good evidence that this stock exhibits extended periods of strong recruitment, as well as multiple successive years of weak recruitment, even after taking account of a relatively significant underlying relationship between stock and recruitment.

In general, the model fits the data well and there are no major inconsistencies or conflicts among separate data sources. The results are not notably sensitive to either selection of data or the weighting placed upon alternative sets of data. The only exception is that the assessment of stock status is sensitive to uncertainty in the historical time series of catches, a source of uncertainty not often addressed in stock assessment. The results do not seem sensitive to the selection of priors, given the low variance ratios of posteriors to priors for most parameters. One surprising feature of the results is the low estimates for stock recruitment steepness. These may be a consequence of a temporal change in the environmental determinants of recruitment rather than reflecting a true density dependent effect.

The stock assessment methods which have been applied to the eastern gemfish resource have not been examined by simulation to assess their likely levels of bias. Such analyses (e.g., Punt 1989, Cordue and Francis 1994) could indicate that the methods which have been applied to this stock are either markedly biased or highly imprecise.

The results of the Bayesian analyses form the basis for inferring the current status of the eastern gemfish population, and also provide a basis for evaluating risks and benefits associated with alternative future management strategies for this resource. In terms of AFMA's current reference point, all of the Bayesian analyses have posterior means for the ratio E_{96}/E_{79} that exceed 0.4. The status of the population relative to the unexploited equilibrium biomass is much less certain and depends very much on the catch series selected. The posterior mean for the ratio B_{96}/B_0 is 0.27 for catch series A and 0.36 for catch series B. However, the results for all of the Bayesian analyses indicate that some recovery has occurred since 1991 owing to the strong 1990 cohort recruiting to the fishable biomass.

One consequence of the relatively high level of recruitment variability in this stock, coupled with autocorrelation between recruitment anomalies in successive years, is that constant harvest strategies are likely to perform particularly poorly, relative to more adaptive strategies, in relation to risks of overfishing, and longer term average catches. This tradeoff was recognized in the process of determining a quota for the stock in 1997. Since all analyses suggested that the criterion for reopening the fishery had been exceeded, a non-zero quota was set for this stock in 1997 for the first time since 1992. However it was also recognized that the current recovery was based mainly around one strong year class. A decision was made to adopt a very flexible strategy for setting quota, involving fishing the stock while there were strong year classes present, and being prepared to close the fishery again at short notice if the reference point would be exceeded.

One of the notable features of the assessment for eastern gemfish is the process by which it was developed. This assessment was developed and undertaken by the Eastern Gemfish Assessment Group (EGAG) which, as noted in the introduction, comprises a diverse membership including fishermen, scientists, and managers. It took about 18 months and six meetings for EGAG to develop the current assessment. One key to its success was that EGAG initiated an industry-sponsored trawl survey in 1996, which provided the first opportunity for targeted fishing of gemfish since 1992, and the first reliable index of relative abundance since 1991. Another key to success was the opportunity for the fishing industry to have real input into the stock assessment for the first time for this fishery. One of the direct consequences of this involvement was the identification of the uncertainties surrounding historical catches, previously unidentified, which turned out to be the major uncertainty in the assessment. A very important consequence of this involvement was also a much better acceptance by the fishing industry of both the science and the management process. The EGAG structure and process seems to have been a particularly effective one, and is currently being adopted more widely as a model for the stock assessment process in federally managed fisheries in Australia.

The next step for EGAG is to further explore alternative future harvest strategies for eastern gemfish, including feedback strategies. The Bayesian assessments will form the basis for such an exploration and evaluation (e.g., Smith et al. 1996). The medium term aim is to develop an agreed management procedure (a combination of monitoring strategy, assessment method, and harvest control law), along the lines of those developed in the International Whaling Commission and in some South African fisheries (Donovan 1989, Butterworth and Bergh 1993).

Acknowledgments

The specifications of the analyses were provided by the Eastern Gemfish Assessment Group. All the members of EGAG are thanked for their contibutions to this analysis. In particular, Kevin Rowling (NSW Fisheries, Sydney, Australia) is thanked for providing the catch per unit effort and length-frequency data, and for critical comment on the analyses. Jim Ianelli (AFSC, NOAA, Seattle, USA) is thanked for providing the prior distribution assumed for steepness.

References

Allen, K.R. 1989. A cohort analysis of the gemfish stock of southeast Australia. Document submitted to the Demersal and Pelagic Fish Research Group, 29 Nov. 1989, Canberra, ACT. 51 pp.

Bax, N.J. 1996. Preliminary assessment of eastern gemfish *(Rexea solandri)* using stock synthesis. In: A.D.M. Smith (ed.), Evaluation of harvesting strategies for Australian fisheries at different levels of risk from economic collapse. Fisheries Research and Development Corporation, Australia, Report T93/238, pp. 68-119.

Beddington, J.R., and J.G. Cooke. 1983. The potential yield of fish stocks. FAO Fish. Tech. Pap. 242:1-47.

Butterworth, D.S., and M.O.Bergh. 1993. The development of a management procedure for the South African anchovy resource. In: S.J. Smith, J.J. Hunt, and D. Rivard (eds.), Risk evaluation and biological reference points for fisheries management. Can. Spec. Publ. Fish. Aquat. Sci. 120:83-99.

Cordue, P.L., and R.I.C.C. Francis. 1994. Accuracy and choice in risk estimation for fisheries assessment. Can. J. Fish. Aquat. Sci. 51:817-829.

Deriso, R.B., T.J. Quinn II, and P.R. Neal. 1985. Catch-age analysis with auxiliary information. Can. J. Fish. Aquat. Sci. 42:815-824.

Donovan, G.P. (ed.). 1989. The comprehensive assessment of whale stocks: The early years. Rep. Int. Whaling Comm. (Special Issue 11). 210 pp.

EGAG. 1996. Report of the 23-24 October 1996 meeting of EGAG (Eastern Gemfish Assessment Group). Canberra, ACT. 13 pp.

Fournier, D., and C.P. Archibald. 1982. A general theory for analyzing catch at age data. Can. J. Fish. Aquat. Sci. 39:1195-1207.

Francis, R.I.C.C. 1992. Use of risk analysis to assess fishery management strategies: A case study using orange roughy *(Hoplostethus atlanticus)* on the Chatham Rise, New Zealand. Can. J. Fish. Aquat. Sci. 49:922-930.

Givens, G.H., J.E. Zeh, and A.E. Raftery. 1995. Assessment of the Bering-Chukchi-Beaufort Seas stock of bowhead whales using the Baleen II model in a Bayesian synthesis framework. Rep. Int. Whaling Comm. 45:345-364.

Hilborn, R., E.K. Pikitch, and M.K. McAllister. 1994. A Bayesian estimation and decision analysis for an age-structured model using biomass survey data. Fish. Res. 19:17-30.

McAllister, M.K., and J.N. Ianelli. 1997. Bayesian stock assessment using catch-age data and the sampling-importance resampling algorithm. Can. J. Fish. Aquat. Sci. 54:284-300.

McAllister, M.K., E.K. Pikitch, A.E. Punt, and R. Hilborn. 1994. A Bayesian approach to stock assessment and harvest decisions using the sampling/importance resampling algorithm. Can. J. Fish. Aquat. Sci. 51:2673-2687.

Methot, R. 1989. Synthetic estimates of historical abundance and mortality for northern anchovy. Amer. Fish. Soc. Symp. 6:66-82.

Methot, R. 1990. Synthesis model: An adaptable framework for analysis of diverse stock assessment data. Int. North Pac. Fish. Comm. Bull. 50:259-277.

Pikitch, E.K., R. Hilborn, M.K. McAllister, and A.E. Punt. 1993. Stock assessment and decision analysis for the western stock of hoki in 1992. N.Z. Fisheries Assessment Research Document 93/10, Wellington, N.Z. 46 pp.

Pope, J.G., and J.G. Shepherd. 1985. A comparison of the performance of various methods for tuning VPAs using effort data. J. Cons. Cons. Int. Explor. Mer 42:129-151.

Punt, A.E. 1989. A further preliminary comparative investigation into the performance of catch-effort model estimation procedures used in the management of the ICSEAF hake stocks. Collection of Scientific Papers of the International Commission for Southeast Atlantic Fisheries 16(II):63-122.

Punt, A.E. 1996. Preliminary stock assessments of eastern gemfish *(Rexea solandri)* using Bayesian and maximum likelihood methods. In: A.D.M. Smith (ed.), Evaluation of harvesting strategies for Australian fisheries at different levels of risk from economic collapse. Fisheries Research and Development Corporation, Canberra, Australia, Report T93/238, pp. 6-67.

Punt, A.E. 1997a. The performance of VPA-based management. Fish. Res. 29:217-243.

Punt, A.E. 1997b. Bayesian and maximum likelihood assessments of the eastern stock of gemfish for 1997. Document submitted to the 10-12 March 1997 meeting of EGAG (Eastern Gemfish Assessment Group), Canberra, ACT. 75 pp.

Punt, A.E., and D.S. Butterworth. 1996. Further remarks on the Bayesian approach for assessing the Bering-Chukchi-Beaufort Seas stock of bowhead whales. Rep. Int. Whaling Comm. 46:481-191.

Punt, A.E., and R. Hilborn. 1997. Fisheries stock assessment and decision analysis: The Bayesian approach. Rev. Fish Biol. Fish. 7:35-63.

Punt, A.E., K. Rowling, and J. Prince. 1997. Summary of the data for use in the assessments of the eastern stock of gemfish. Document submitted to the 10-12 March 1997 meeting of EGAG (Eastern Gemfish Assessment Group), Canberra, ACT. 16 pp.

Punt, A.E., M.K. McAllister, E.K. Pikitch, and R. Hilborn. 1994. Stock assessment and decision analysis for hoki *(Macruronus novaezelandiae)* for 1994. N.Z. Fisheries Assessment Research Document 94/13, Wellington, N.Z. 47 pp.

Rowling, K.R. 1990. Changes in the stock composition and abundance of spawning gemfish *Rexea solandri* (Cuvier), Gempylidae, in south-eastern Australian waters. Aust. J. Mar. Freshw. Res. 41:145-63.

Rowling, K.R., and D.D. Reid. 1992. Effect of temporal change in size composition on estimates of von Bertalanffy growth parameters for gemfish, *Rexea solandri* (Cuvier), Gempylidae. Aust. J. Mar. Freshw. Res. 43:1229-1239.

Rubin, D.B. 1987. Comment: The calculation of posterior distributions by data augmentation. J. Am. Statist. Assoc. 82:543-546.

Smith, A.D.M., A.E. Punt, S.E. Wayte, and N.L. Klaer. 1996. Evaluation of harvest strategies for eastern gemfish *(Rexea solandri)* using Monte Carlo simulation. In: A.D.M. Smith (ed.), Evaluation of harvesting strategies for Australian fisheries at different levels of risk from economic collapse. Fisheries Research and Development Corporation, Canberra, Australia, Report T93/238, pp. 120-164.

Thompson, G.G. 1993. A proposal for a threshold stock size and maximum fishing mortality rate. In: S.J. Smith, J.J. Hunt, and D. Rivard (eds), Risk evaluation and biological reference points for fisheries management. Can. Spec. Publ. Fish. Aquat. Sci. 120:303-320.

Van Dijk, H.K., J.P. Hop, and A.S. Louter. 1987. An algorithm for the computation of posterior moments and densities using simple importance sampling. The Statistician 36:83-90.

Walters, C.J., and D. Ludwig. 1994. Calculation of Bayes posterior probability distributions for key population parameters. Can. J. Fish. Aquat. Sci. 51:713-722.

Appendix I: The Population Dynamics Model

The model considered in this appendix is age- and sex-structured, takes account of two pulse fisheries, and assumes that the number of births is related to the total egg production of the population by means of a Beverton and Holt stock-recruitment relationship.

A. Basic Population Dynamics

The resource dynamics are modeled using the equations:

$$N^s_{y+1,a} = \begin{cases} N^s_{y+1,0} & \text{if } a = 0 \\ \tilde{N}^s_{y,a-1} & \text{if } 1 \le a \le x-1 \\ \tilde{N}^s_{y,x} + \tilde{N}^s_{y,x-1} & \text{if } a = x \end{cases} \tag{I.1}$$

where

$N^s_{y,a}$ is the number of fish of sex s and age a at the start of year y,

$\tilde{N}^s_{y,a}$ is the number of fish of sex s and age a at the end of year y,

$$\tilde{N}^s_{y,a} = (N^s_{y,a} e^{-t_1 M^s_a} - C^{1,s}_{y,a}) e^{-t_2 M^s_a} - C^{2,s}_{y,a} \tag{I.2}$$

$N^s_{y,0}$ is the number of 0-year-olds of sex s at the start of year y,

M^s_a is the rate of natural mortality on fish of sex s and age a,

$C^{1,s}_{y,a}$ is the catch (in number) of fish of sex s and age a during the summer fishery of year y,

$C^{2,s}_{y,a}$ is the catch (in number) of fish of sex s and age a during the winter fishery of year y,

t_1 is the time between the start of the year and the mid-point of the summer fishery,

t_2 is the time between the mid-point of the summer fishery and that of the winter fishery ($t_1 + t_2 = 1$ because the winter fishery is assumed to occur at the end of the year), and

x is the maximum age considered (taken to be a plus-group).

B. Births

$$N^s_{y,0} = \frac{0.5 SB_{y-1} e^{\varepsilon_y}}{[\alpha + \beta SB_{y-1}]} \tag{I.3}$$

where

SB_y is proportional to the total egg production at the end of year y:

$$SB_y = \sum_{a=1}^{x} w_a^f S_a^{2,f} \tilde{N}_{y,a}^f \tag{I.4}$$

$S_a^{2,s}$ is the selectivity of the fishing gear used during the winter fishery on fish of sex s and age a, and

w_a^s is the mass of a fish of sex s and age a during the winter fishery:

$$w_a^w = b_1^s (L_a^s)^{b_2^s} \tag{I.5}$$

$$L_a^s = L_\infty^s \left[1 - e^{-\kappa^s (a - t_0^s)}\right] \tag{I.6}$$

ε_y is the recruitment anomaly for year y [$\varepsilon_y \sim N(0;\sigma_r^2)$],

σ_r is the standard deviation of the logarithm of the multiplicative fluctuations in births (approximately the coefficient of variation of these fluctuations), and

α,β are the stock-recruitment relationship parameters.

This formalism assumes that egg production is determined by the biomass of females which survives the winter spawning run fishery and that maturation is the same as selectivity to the winter fishery. Some of the catches during the winter fishery occur after spawning (the “back run”). However, these catches are generally a small fraction of the total winter catch (~5-10%, K. Rowling, pers. comm.) so this complication has been ignored here.

C. Natural Mortality

The rate of natural mortality for animals of age a and sex s, M_a^s, is given by:

$$M_a^s = \begin{cases} M^m & \text{if } a < 3 \text{ or } s = m \\ M^f & \text{otherwise} \end{cases} \tag{I.7}$$

This formalism implies that, in the absence of exploitation, the size of the male and female cohorts are the same at age 3 and diverge thereafter due to different rates of natural mortality. The specification that the rate of natural mortality for animals aged 2 and less is the same as that of males is based on the assumption that the rate of natural mortality for males is larger than that for females. However, the algorithm used to specify natural mortality for animals aged 2 years and less is largely irrelevant because the fishery hardly impacts animals younger than 3 years.

D. Initial Conditions

Were there no fluctuations in recruitment, the resource would be assumed to be at its unexploited equilibrium level, with the corresponding age-structure, at the start of exploitation (1962). Instead, because of historical fluctuations in recruitment, the sizes of the cohorts at the start of 1962 are drawn from distributions that allow for this fluctuation, and the initial biomass is thus similarly distributed about the corresponding deterministic equilibrium level. The initial numbers-at-age are given by the equations:

$$\begin{aligned} N^s_{1962,a} &= 0.5 R_0 \exp(-\sum_{a'=0}^{a-1} M^s_{a'}) e^{\varepsilon_a} && \text{if } 0 \le a \le x-1 \\ N^s_{1962,x} &= 0.5 R_0 \exp(-\sum_{a'=0}^{x-1} M^s_{a'}) / \{1 - \exp(-M^s_x)\} && \text{if } a = x \end{aligned} \tag{I.8}$$

where

R_0 is the number of 0-year-olds at the deterministic equilibrium that corresponds to an absence of harvesting, and

ε_a is a random variable from $N(0;\sigma_r^2)$.

Note that the equation for the plus-group does not incorporate a recruitment variability term because this group comprises a large number of age classes which will largely damp out this effect.

A value for R_0 is calculated from the value for the unexploited equilibrium biomass at the end of the year, B_0 (where this biomass is defined using the selectivity pattern for the winter fishery and the mass-at-age vector is calculated from data collected during the winter spawning run), using the equation:

$$R_0 = \frac{B_0}{\sum_s \left\{ \sum_{a=0}^{x-1} w^s_a S^{2,s}_a \exp(-\sum_{a'=0}^{a} M^s_{a'}) + w^s_x S^{2,s}_x \dfrac{\exp(-\sum_{a'=0}^{x} M^s_{a'})}{1 - \exp(-M^s_x)} \right\}} \tag{I.9}$$

Values for the stock-recruitment function parameters α and β are calculated from the values of R_0 and the "steepness" of the stock-recruitment relationship (h). The "steepness" is the fraction of R_0 to be expected (in the absence of recruitment variability) when the egg production is reduced to 20% of its pristine level (Francis 1992), so that:

$$\beta = \frac{5h-1}{4hR_0}$$

$$\alpha = \frac{(1-h)\tilde{B}_0^S}{4h} \tag{I.10}$$

$$\tilde{B}_0^S = \frac{\sum_{a=1}^{x-1} w_a^f S_a^{2,f} \exp(-\sum_{a'=0}^{a} M_{a'}^f) + w_a^f S_x^{2,f} \exp(-\sum_{a'=0}^{x} M_{a'}^f)}{1-\exp(-M_x^f)}$$

E. Catches

The catch (in number) of fish of sex *s* and age *a* during the summer fishery of year *y*, $C_{y,a}^{1,s}$, is calculated from $\tilde{C}_y^1$, the catch (in mass) during the summer fishery, using the equation:

$$C_{y,a}^{1,s} = S_a^1 F_y^1 N_{y,a}^s e^{-t_1 M_a^s} \tag{I.11}$$

where

S_a^1 is the selectivity of the gear used during the summer fishery on fish of age *a* (assumed invariant of sex), and

F_y^1 is the exploitation rate on fully selected fish during the summer fishery of year *y*:

$$F_y^1 = \frac{\tilde{C}_y^1}{\sum_s \sum_{a=0}^{x} w_{a-t_2}^s S_a^1 N_{y,a}^s e^{-t_1 M_a^s}} \tag{I.12}$$

The catch (in number) of fish of sex *s* and age *a* during the winter fishery of year *y*, $C_{y,a}^{2,s}$, is calculated from $\tilde{C}_y^2$, the catch (in mass) during the winter fishery, using the equation:

$$C_{y,a}^{2,s} = S_a^{2,s} F_y^2 (N_{y,a}^s e^{-t_1 M_a^s} - C_{y,a}^{1,s}) e^{-t_2 M_a^s} \tag{I.13}$$

where

F_y^2 is the exploitation rate on fully selected fish during the winter fishery of year *y*:

$$F_y^2 = \frac{\tilde{C}_y^2}{\sum_s \sum_{a=0}^{x} w_a^s S_a^{2,s} (N_{y,a}^s e^{-t_1 M_a^s} - C_{y,a}^{1,s}) e^{-t_2 M_a^s}} \tag{I.14}$$

F. Selectivity

The selectivity functions are sex- and fishery-specific. The selectivity function for the winter fishery is assumed to have the logistic form:

$$S_a^{2,s} = \begin{cases} 0 & \text{if } a < 2 \\ \left(1 + \exp\left[-\ell n(19)(L_a^{2,s} - L_{50}^{2,s}) / (L_{95}^{2,s} - L_{50}^{2,s})\right]\right)^{-1} & \text{otherwise} \end{cases} \qquad \text{(I.15)}$$

where

$L_{50}^{2,s}$ is the length-at-50%-selectivity for fish of sex *s* during the winter fishery, and

$L_{95}^{2,s}$ is the length-at-95%-selectivity for fish of sex *s* during the winter fishery.

The logistic form was chosen because selectivity to the winter fishery is equivalent to being mature and the probability of a fish being mature increases with size and age. The selectivities for ages 0 and 1 are set equal to zero because fish aged less than two years have never been encountered during the winter spawning run.

The length-frequencies for the summer fishery suggest that the selectivity function for this fishery is dome-shaped. The simplest functional form with this shape is a normal distribution:

$$S_a^1 = \begin{cases} 0 & \text{if } a = 0 \\ \exp[-0.69(L_a^1 - L_{\text{full}}^1)^2 / (L_{50\%}^1 - L_{\text{full}}^1)^2] & \text{otherwise} \end{cases} \qquad \text{(I.16)}$$

where

L_{full}^1 is the length corresponding to maximum selectivity during the summer fishery, and

$L_{50\%}^1$ is the length at which selectivity during the summer fishery drops to 50% of the maximum.

G. Data Series

The catch rate data for the winter fishery are assumed to be proportional to the exploitable biomass in the middle of the winter fishery:

$$B_y^2 = \sum_s \sum_{a=0}^{x} w_a^s S_{y,a}^{2,s} \left[(N_{y,a}^s e^{-t_1 M_a^s} - C_{y,a}^{1,s}) e^{-t_2 M_a^s} - C_{y,a}^{2,s} / 2\right] \qquad \text{(I.17)}$$

The age-structure information is taken to be proportional to the model-predicted catches-at-age (i.e. $\{C_{y,a}^{1,s}$ and $C_{y,a}^{2,s}\}$). The estimate of the fraction of the catch in length-class L (where length-class L ranges from $28 + 2L$ to $30 + 2L$ cm and $L = 1,2,\ldots,43$) in year y (for a given sex and fishery) is calculated using the equation:

$$C_{y,L} = \sum_a C_{y,a}\Phi(a,L) \tag{I.18}$$

where

$\Phi(a,L)$ is the probability that a fish of age a lies in length-class L:

$$\Phi(a,L) = \int_{\ell n(28+2L)}^{\ell n(30+2L)} \frac{1}{\sqrt{2\pi}\phi_a} e^{-\frac{(\ell n\tilde{L}-\ell nL_a)^2}{2(\phi_a)^2}} \, \mathrm{d}\,\ell n\tilde{L} \tag{I.19}$$

L_a is the mean length of a fish of age a (computed using the von Bertalanffy growth equation, see equation I.6), and

ϕ_a is the standard deviation of the logarithm of the length of a fish of age a (approximated here by the CV of L_a).

Appendix II: The Likelihood Function

A. The Catch Rate Data

The catch rate indices are assumed to be lognormally distributed about their expected values with a constant (and known) coefficient of variation, i.e.:

$$\ell\mathrm{n}O_y = \ell\mathrm{n}(qB_y^2) + \eta_y \qquad \eta_y \sim N(0;\sigma_c^2) \tag{II.1}$$

where

O_y is the catch rate for year y,

q is the catchability coefficient, and

σ_c is (approximately) the coefficient of variation of the noise about the catch rate-abundance relationship.

The contribution of the catch rate information to the likelihood function is:

$$\lambda_1 = \prod_y \frac{1}{O_y\sqrt{2\pi\sigma_c^2}} \exp\left(-\tfrac{1}{2\sigma_c^2}\left\{\ell\mathrm{n}O_y - \ell\mathrm{n}(qB_y^2)\right\}^2\right) \tag{II.2}$$

where the product is taken over all years for which catch rates are available.

The maximum likelihood estimate for q is given by:

$$\hat{q} = \exp\left(\tfrac{1}{n_c}\sum_y \ell n\,(O_y / B_y^2)\right) \tag{II.3}$$

where

n_c is the number of catch rate data points.

The integral of the likelihood over the prior assumed for q (see Walters and Ludwig 1994, Punt and Butterworth 1996) is:

$$\int_{-\infty}^{\infty} \lambda_1(q)\, p(\ell\mathrm{n}q)\, d\ell\mathrm{n}q \propto \exp\left(-\tfrac{1}{2\sigma_c^2}\sum_y\left\{\ell\mathrm{n}O_y - \ell\mathrm{n}(\hat{q}B_y^2)\right\}^2\right) \tag{II.4}$$

where $p(\ell\mathrm{n}q) = 1$.

B. Fraction of the Winter Catch Which is Female

The estimates of the fraction of the catch during the winter fishery (in numbers) consisting of females are assumed to be normally distributed about the model-estimates with a standard deviation of σ_f. The contribution of the female fraction data to the likelihood is therefore:

$$\lambda_2 = \prod_y \frac{1}{\sqrt{2\pi\,\sigma_f^2}} \exp\left(-\frac{\left(Q_y - \hat{Q}_y\right)^2}{2\sigma_f^2}\right) \tag{II.5}$$

where

Q_y is the fraction of the winter catch during year y (in numbers) of animals larger than 64 cm which consists of females, and

$\hat{Q}_y$ is the model-estimate of Q_y:

$$\hat{Q}_y = \sum_{L=64}^{114} C_{y,L}^{2,f} \Big/ \sum_s \sum_{L=64}^{114} C_{y,L}^{2,s} \tag{II.6}$$

The product in equation (II.5) is taken over the number of years for which estimates of Q_y are available.

C. The Age-Composition Data

The age-structure data contain information about the total number of fish landed as well as about the age-composition of the catch. The former is used (implicitly) when projecting the model forward (equations I.12 and I.14) so only the information about the age-composition of the catches is included in the likelihood function. The observed fraction of the catch (by number) taken in year y composed of fish of age a is assumed to be log-normally distributed about its expected value with a CV of $\sigma_a^s / \sqrt{\hat{p}_{y,a}^s}$, so the contribution to the likelihood function of the catch of animals of sex s by the winter fishery is:

$$\lambda_3^s = \prod_y \prod_{a \geq a_{a,y}^s} \frac{1}{\rho_{y,a}^{s,obs} \sqrt{2\pi(\sigma_a^s)^2 / \hat{\rho}_{y,a}^s}} \exp\left(-\frac{\hat{\rho}_{y,a}^s}{2(\sigma_a^s)^2}(\ell\mathrm{n}\rho_{y,a}^{s,obs} - \ell\mathrm{n}\hat{\rho}_{y,a}^s)^2\right) \tag{II.7}$$

where

$\rho_{y,a}^{s,obs}$ is the observed fraction (by number) of the winter fishery catch during year y of sex s which animals of age a make up of the catch of animals not younger than age $a_{a,y}^s$:

$$\rho_{y,a}^{s,obs} = C_{y,a}^{2,s,obs} \Big/ \sum_{a'=a_{a,y}^s}^{x} C_{y,a'}^{2,s,obs} \tag{II.8}$$

$\hat{\rho}_{y,a}^s$ is the model-predicted value of $\rho_{y,a}^{s,obs}$:

$$\hat{\rho}_{y,a}^s = \sum_{a'=0}^{x} \chi_{a',a}\, C_{y,a'}^{2,s} \Big/ \sum_{a'=a_{a,y}^s}^{x} \sum_{a''=0}^{x} \chi_{a'',a'}\, C_{y,a''}^{2,s} \tag{II.9}$$

$a^s_{a,y}$ is the lowest age included in the likelihood for sex s and year y,

$C^{f,s,obs}_{y,a}$ is the observed catch during year y of animals of sex s and age a by fishery f (summer or winter), and

$\chi_{a',a}$ is the probability that an animal of (real) age a' will be assigned age a when its age is determined (the age-reading error matrix).

The products in equation (II.7) are taken over all years and ages for which age-composition data are available. For improved numerical stability, the fractions for ages 10+ are pooled.

D. The Size-Composition Data

The treatment of the size-composition data is analogous to that of the age-composition data except that no account needs to be taken of age-reading error. When fitting to the unsexed length-frequency data, the model-estimated proportion of the catch falling into each 2 cm length interval is computed by combining the model-estimates of the catches-at-length for the two sexes.

Appendix III: The Sample Importance Resample (SIR) Algorithm

The sample-importance-resample (SIR) algorithm (Rubin 1987, Van Dijk et al. 1987) is a method for approximating the posterior distribution for high dimensional problems. The purpose of the SIR algorithm is to sample a set of parameter combinations $\{\underline{\theta}_i : i = 1, 2, \ldots\}$ from the posterior distribution, $\pi(\underline{\theta}) \propto g(\underline{\theta}) = L(D \mid \underline{\theta}) p(\underline{\theta})$. This sample can then be used as input to a decision analysis or for inference about the values of quantities of interest to management (Punt and Hilborn 1997). Evaluation of $L(D \mid \underline{\theta})$ involves projecting the age- and sex-structured population dynamics model forward, using known catches, to predict stock biomasses and then calculating the likelihood for the projection. If the population becomes extinct before the most recent year, the likelihood is set equal to zero. This is equivalent to placing an improper prior distribution on the current depletion which is uniform above 0 (Punt and Hilborn 1997).

The SIR algorithm requires an approximation to the posterior distribution to generate a sample from $\pi(\underline{\theta})$. This approximation, $\tilde{\pi}(\underline{\theta})$, referred to as the importance function, must have non-zero probability wherever $\pi(\underline{\theta})$ has non-zero probability and must be easy to sample from. The simplest choice for $\tilde{\pi}(\underline{\theta})$ is the prior distribution, $p(\underline{\theta})$, although this choice may not be very efficient if the likelihood supports only a small part of $p(\underline{\theta})$. For the purposes of the present study, the importance function has been taken to be a multivariate normal distribution with mean given by the values which maximize $g(\underline{\theta})$—the mode of the posterior distribution—and variance-covariance matrix determined by fitting a quadratic surface to $g(\underline{\theta})$ at its mode. The SIR algorithm proceeds as follows.

a. Generate a vector, $\underline{\theta}_i$, from the distribution $\tilde{\pi}(\underline{\theta})$.

b. Calculate $Y_i = g(\underline{\theta}_i)$ and the importance weight $w_i = Y_i / \tilde{\pi}(\underline{\theta}_i)$ where $\tilde{\pi}(\underline{\theta}_i)$ is the probability of generating the vector $\underline{\theta}_i$.

c. Repeat steps (a) and (b) a very large number of times (20,000,000 for the calculations of this paper).

d. Select 10,000 parameter combinations from the 20,000,000 with replacement, assigning a probability of selecting a particular parameter combination proportional to its importance weight.

Appendix IV: The Derivation of Equation (2)

The correlated recruitment sensitivity test is based on the assumption that the recruitment anomaly for year y, ε_y, is correlated with those for all previous years. For the purposes of the analyses of this paper, it is assumed that this correlation structure can be modeled by an ARMA-1 process:

$$\varepsilon_y = \tau\varepsilon_{y-1} + \chi\varepsilon'_y \tag{IV.1}$$

where

ε'_y is a independent random variate from $N(0;1^2)$.

The variance of ε_y: $y = 1,2,\ldots,\infty$ is defined to be equal to σ_r^2 so

$$\begin{aligned}\sigma_r^2 &= E(\varepsilon_y\varepsilon_y)\\ &= E\left[(\tau\varepsilon_{y-1} + \chi\varepsilon'_y)(\tau\varepsilon_{y-1} + \chi\varepsilon'_y)\right]\\ &= E(\tau^2\varepsilon_{y-1}^2) + 2E(\tau\chi\varepsilon_{y-1}\varepsilon'_y) + E(\chi^2\varepsilon'^2_y)\\ &= \tau^2E(\varepsilon_{y-1}^2) + 2\tau\chi E(\varepsilon_{y-1}\varepsilon'_y) + \chi^2E(\varepsilon'^2_y)\\ &= \tau^2\sigma_r^2 + 0 + \chi^2\end{aligned} \tag{IV.2}$$

The term $E(\varepsilon_{y-1}\varepsilon'_y)$ is equal to zero because ε_{y-1} and ε'_y are independent and both have expectation zero. Solving for σ_r^2 leads to an expression for χ:

$$\chi = \sqrt{1-\tau^2} \tag{IV.3}$$

Equation (IV.1) can therefore be rewritten as:

$$\varepsilon_y = \tau\varepsilon_{y-1} + \sqrt{1-\tau^2}\,\varepsilon'_y \tag{IV.4}$$

As the recruitment anomalies are not uncorrelated, their joint prior is not simply the product of the individual (independent) priors; equation (1) of the main text is the negative of the logarithm of such a prior after the removal of constants. It is necessary to construct a joint multivariate normal prior for the recruitment anomalies which incorporates the correlation structure of equation (IV.4). The generic form for such a prior is given by:

$$P(\underline{\varepsilon}) = \frac{1}{(2\pi)^{n/2}\sqrt{\det \mathbf{V}}}\exp\left(-\tfrac{1}{2}\,\underline{\varepsilon}^T\,\mathbf{V}^{-1}\,\underline{\varepsilon}\right) \tag{IV.5}$$

where

n is the number of elements in $\underline{\varepsilon}$

V is the variance-covariance matrix, i.e.:

$$\mathbf{V} = \begin{pmatrix} E(\varepsilon_1\,\varepsilon_1) & E(\varepsilon_1\,\varepsilon_2) & E(\varepsilon_1\,\varepsilon_3) \\ E(\varepsilon_2\,\varepsilon_1) & E(\varepsilon_2\,\varepsilon_2) & E(\varepsilon_2\,\varepsilon_3) \\ E(\varepsilon_3\,\varepsilon_1) & E(\varepsilon_3\,\varepsilon_2) & E(\varepsilon_3\,\varepsilon_3) \end{pmatrix} \tag{IV.6}$$

for case of $n = 3$.

The elements of **V** are determined as follows (for $i > j$):

$$\begin{aligned} E(\varepsilon_i\varepsilon_j) &= E(\varepsilon_j\varepsilon_i) \\ &= E\left[\varepsilon_j\left(\tau\varepsilon_{i-1} + \sqrt{1-\tau^2}\varepsilon'_i\right)\right] \\ &= E\left\{\varepsilon_j\left[\left(\tau^2\varepsilon_{i-2} + \sqrt{1-\tau^2}\varepsilon'_{i-1}\right) + \sqrt{1-\tau^2}\varepsilon'_i\right]\right\} \\ &= E\left[\varepsilon_j\left(\tau^{(i-j)}\varepsilon_j + \sum_{k=j+1}^{i}\tau^{i-k}\sqrt{1-\tau^2}\varepsilon'_k\right)\right] \\ &= \tau^{(i-j)}E(\varepsilon_j\,\varepsilon_j) \\ &= \tau^{(i-j)}\sigma_r^2 \end{aligned} \tag{IV.7}$$

Substituting equation (IV.7) into equation (IV.5), taking logarithms, negating, and dropping the term under the radical (because given fixed values for σ_r and τ it is a constant) leads to equation (2).

Forecast Methods for Inseason Management of the Southeast Alaska Chinook Salmon Troll Fishery

Jerome J. Pella and Michele M. Masuda
National Marine Fisheries Service, Auke Bay Laboratory, Juneau, Alaska

D.G. Chen
Department of Fisheries and Oceans, Pacific Biological Station, Nanaimo, B.C., Canada

Abstract

Under an agreement among U.S. commissioners of the Pacific Salmon Commission (Allen 1996), the allowable catch of chinook salmon in Southeast Alaska for 1997 depended on abundance in the fishery. Historically, abundance has been forecast prior to the summer fishery and evaluated later, from numerous information sources, by the Chinook Technical Committee of the Pacific Salmon Commission. Here, early catch rates during the summer fishery are shown to reflect abundance and are used to update the preseason forecast for use by the commissioners in assessing the allowable catch. The Bayesian calibrative distribution for the updated abundance provides the inseason statistical description of its uncertainty arising from unknown parameter values and random variables of underlying models. Cross validation shows that the mode of the Bayesian calibrative distribution would have produced more accurate forecasts than either the preseason forecast or a maximum likelihood forecast from the early catch rates alone.

Introduction

Allowable catch to be taken from a stock is commonly assessed in advance of the fishing season. Computation of a recommended catch presumably includes a forecast of abundance and possibly some measure of its precision. In some cases, the recommended catch could be updated within the fishing season by correcting the abundance from that forecast. The revision would be based on new information on abundance derived from inseason monitoring of escapement, the fishery, or both. Therefore, instead of specifying a fixed recommended catch prior to the season, the fishery manager would more reasonably provide an allowable catch subject to revision as incoming information better delimits abundance. Many salmon fisheries in Alaska are conducted under similar programs (e.g., Minard and Meacham 1987; Mundy et al. 1993; Geiger et al. 1997).

The Pacific Salmon Treaty of 1985 provided the opportunity for cooperative research and management of Canadian and American stocks that intermingle in coastal fisheries. Among the five species, none surpasses chinook salmon for allocation and management problems because of its high value and migratory nature. In general, stocks of chinook salmon migrate northward varying distances from their home streams along the North American coast and into the Gulf of Alaska; they mature most commonly from two to four years later and encounter coastal fisheries during the return migration. As a result, many stocks of the American Pacific Northwest are caught in both British Columbia and Alaska, and many stocks of British Columbia are caught in Alaska. At the time of the treaty signing, most chinook stocks were judged in need of rebuilding which was to be implemented through negotiated harvest ceilings.

To aid in monitoring the condition and rebuilding of chinook stocks, the Chinook Technical Committee (CTC) of the Pacific Salmon Commission (PSC) developed a numerical computer model (CTC model) to represent the life history and population dynamics of all chinook salmon stocks, using detailed information regarding stocks and fisheries. Information sources comprising catches, escapements, recoveries of coded-wire-tagged hatchery releases, recruitment forecasts, and auxiliary experiments were integrated by the CTC model. The structure of the CTC model continues to evolve, accommodating members' concerns.

Forecasts of relative abundance in regional fisheries are currently produced from the CTC model prior to the fishing season. These forecasts are based in part on predictions of magnitudes of component recruitments from stocks or stock groups provided by participating government agencies. Further, the regional forecasts require predicted geographical distributions of the component recruitments, as well as updated estimates of magnitudes and areal distributions of components for older age groups. The estimates for older ages are generated by the CTC model from catch and escapement counts of tagged and untagged fish subsequent to the forecasts. When new information is processed using the CTC model, the outcome is called a calibration and denoted by the year followed by the order within that year, e.g., 9702 is the second CTC model calibration in 1997 that produced the CTC model values used in this study. The estimated relative abundance in each regional fishery is updated annually through the calibrations. At present, calibrations are not updated during the fishing season for any regional fishery.

The sports and commercial fisheries of Southeast Alaska (SEAK) are one of the regional fisheries identified by the model. The commercial troll fleet takes the preponderance of the SEAK catch. From 1985 through 1996, negotiated harvest ceilings were set for the SEAK fishery under the Pacific Salmon Treaty and its coastwide rebuilding program. The ceilings were based on long-term forecasts of abundance and allowable harvest rates.

In 1995, the Alaska Department of Fish and Game (ADFG) proposed to assess the allowable catch with near-term forecasts of abundance and a baseline harvest rate (Koenings et al. 1995). The abundance was to be estimated from inseason catch rates. The ADFG plan was critically reviewed and found wanting in its implementation, particularly the assessment of abundance (CTC 1995). The plan was challenged successfully in U.S. Federal court, and the fishery was limited by court order. In the fall of 1996, the states of Alaska and Washington signed a letter of agreement directing that SEAK catch would be varied in relation to abundance provided an acceptable method of abundance estimation could be developed. Canada did not participate in the agreement and remains opposed to the separate implementation of abundance-based management in the SEAK fishery without attending to the remaining fisheries.

This report describes and compares methods of abundance estimation for revising SEAK abundance forecasts from the CTC model by fishery performance data within the season. The benefits of combining the two kinds of information are contrasted with their separate use. Bayesian and maximum likelihood calibration methods (Aitchison and Dunsmore 1975) are used in the updating, and the future relative abundance index from the CTC model for the 1997 SEAK fishery is assessed. This assessment was used to set the SEAK catch ceiling for 1997.

Methodology

Problem Statement and Assumptions

Annual relative abundance (abundance for brevity) in the SEAK fishery is assessed from historical records by the CTC model. Practically, the assessment is subject to error so that the values for annual abundance are estimates. These historical values for abundance are the best available and no attempt to improve them is considered here. Rather, current abundance in the summer of 1997 cannot be assessed in the same manner as past abundance because all information is not yet available, but a preseason forecast has been made from the CTC model. Annual forecasts (z) and abundances (τ) are available for 1987 through 1996, from which a forecast error distribution can be derived (Table 1). For future reference, the sequence of their paired logarithms will be denoted as $D_f = \{[\ln(z_i), \ln(\tau_i)], i = 1, 2, \ldots, 10\}$, and the subsequence composed of the first j pairs, by $D_{f[j]}$. (Let $D_f \equiv D_{f[10]}$.) Further, historical series of annual abundance and corresponding measures of fishery performance—inseason indices called the power troll statistic (x)—are available for 1980 through 1996 (Table 1),

Table 1. Historical series for the Southeast Alaska chinook salmon fishery.

Year	Abundance index (τ)	CTC forecast (z)	Power troll statistic (x)
1980	1.03	–	20.55
1981	0.91	–	22.39
1982	1.21	–	26.55
1983	1.29	–	46.53
1984	1.30	–	24.34
1985	1.17	–	32.39
1986	1.30	–	29.68
1987	1.51	1.89	45.83
1988	1.78	2.01	67.20
1989	1.73	1.83	78.61
1990	1.81	1.99	65.74
1991	1.90	1.91	97.11
1992	1.75	1.71	74.92
1993	1.87	1.76	79.24
1994	1.60	1.82	71.31
1995	0.95	0.81	41.38
1996	0.90	0.79	42.33

The abundance index and its forecast from the CTC model, and the power troll statistic (PTS), or numbers caught per power troll permit landing catches during the first five days of the summer season.

and the power troll statistic covaries with abundance. The sequence of their paired logarithms will be denoted as $D_c = \{[\ln(x_i),\ln(\tau_i)],\ i = 1, 2,\ldots,17\}$. After the power troll statistic for the current 1997 season was observed, what could be said about current year abundance in light of the forecast and the historical records? Several answers are compared. First, the current power troll statistic is ignored by adhering to the CTC model forecast. Second, the current CTC model forecast is ignored by assessing abundance with a maximum likelihood method (maximum likelihood calibration) applied to the current power troll statistic and historical data, D_c. Third and last, both pieces of information are combined with the historical background of D_f and D_c by using Bayesian methods (Bayesian calibration) to quantify the plausibility of possible values for current abundance. (The maximum likelihood and Bayesian methods used both share "calibration" in their designation, which will not cause confusion with the more recent name given by the CTC to its model updates.) The first tack needs no further description, but the other two do: notation that is introduced for combining the information by Bayesian calibration covers that needed subsequently for separate assessment from the power troll statistic by maximum likelihood calibration.

Bayesian Calibration

Several assumptions underlie the application of the Bayesian method. Annual records of abundance (τ_i in the ith year) are statistically independent, given the corresponding forecasts (z_i) and the preceding forecasts and abundances, $D_{f[i-1]}$,and their logarithms have a probability distribution described by a parametric model, $p[\ln(\tau_i) \,|\, \ln(z_i), D_{f[i-1]}, \boldsymbol{\psi}]$. Annual records of the power troll statistic (x_i) are statistically independent, given corresponding abundances (τ_i), and their logarithms have a probability distribution described by a parametric regression model, $p[\ln(x_i) \,|\, \ln(\tau_i), \boldsymbol{\beta}]$. The initial sources of information regarding parameters $\boldsymbol{\psi}$ and $\boldsymbol{\beta}$ are independent, so $p(\boldsymbol{\psi},\boldsymbol{\beta}) = p(\boldsymbol{\psi})p(\boldsymbol{\beta})$. The current and unknown logarithm of abundance, $\ln(\tau_{now})$, given the corresponding forecast value, $\ln(z_{now})$, and D_f, is drawn from the same density, $p[\ln(\tau_{now}) \,|\, \ln(z_{now}), D_f, \boldsymbol{\psi})$, as were the historical records.

Given the assumptions, Aitchison and Dunsmore (1975) show that the posterior predictive distribution for $\ln(\tau_{now})$ (a predictive distribution includes the uncertainty from unknown parameters and unobserved random variables; see sec. 1.3 of Gelman et al. 1996), which they call its calibrative distribution, viz.,

$$p[\ln(\tau_{now}) \,|\, \ln(x_{now}), \ln(z_{now}), D_f, D_c] = \frac{p[\ln(\tau_{now}), \ln(x_{now}) \,|\, \ln(z_{now}), D_f, D_c]}{p[\ln(x_{now}) \,|\, \ln(z_{now}), D_f, D_c]}, \tag{1a}$$

can be computed as

$$\frac{p[\ln(\tau_{now}) \,|\, \ln(z_{now}), D_f] \times p[\ln(x_{now}) \,|\, \ln(\tau_{now}), D_c]}{\int_{-\infty}^{+\infty} p[\ln(\tau_{now}) \,|\, \ln(z_{now}), D_f] \times p[\ln(x_{now}) \,|\, \ln(\tau_{now}), D_c] \, d\ln(\tau_{now})}, \tag{1b}$$

where

$$p[\ln(\tau_{now}) \,|\, \ln(z_{now}), D_f] = \int_{\psi} p[\ln(\tau_{now}) \,|\, \ln(z_{now}), D_f, \boldsymbol{\psi}] p(\boldsymbol{\psi} \,|\, D_f) \, d\boldsymbol{\psi} \tag{2}$$

is the predictive conditional probability density of current logarithm of abundance given the current CTC model forecast and historical data,

$$p[\ln(x_{now}) \,|\, \ln(\tau_{now}), D_c] = \int_{\beta} p[\ln(x_{now}) \,|\, \ln(\tau_{now}), \boldsymbol{\beta}] \; p(\boldsymbol{\beta} \,|\, D_c) \, d\boldsymbol{\beta} \tag{3}$$

is the predictive conditional density for the current logarithm of the power troll statistic given current abundance and historical data,

$p(\psi | D_f)$ and $p(\beta | D_c)$ are probability assessments of ψ and β with the historical data, viz.,

$$p(\psi \mid D_f) = \frac{p(\psi)p(D_f|\psi)}{\int_\psi p(\psi)p(D_f|\psi)d\psi}, \tag{4}$$

$$p(\beta | D_c) = \frac{p(\beta)p(D_c|\beta)}{\int_\beta p(\beta)p(D_c|\beta)d\beta}, \tag{5}$$

$p(\psi)$ and $p(\beta)$ are the prior densities for ψ and β without the historical data,

$p(D_f|\psi) = \Pi\, p[\ln(\tau_i)\,|\,\ln(z_i), D_{f[i-1]}, \psi]$ is the likelihood function for ψ given the historical data,

$p(D_c|\beta) = \Pi\, p[\ln(x_i)\,|\,\ln(\tau_i), \beta]$ is the likelihood function for β given the historical data,

$p[\ln(\tau_{now})\,|\,\ln(z_{now}), D_f, \psi]$ is the density function for the current logarithm of abundance given the current forecast, historical data, and forecast model parameters, ψ, and

$p[\ln(x_{now})\,|\,\ln(\tau_{now}), \beta]$ is the density function for the current logarithm of the power troll statistic given the logarithm of current abundance and inseason regression parameters, β,

$p[\ln(\tau_{now}), \ln(x_{now})\,|\,\ln(z_{now}), D_f, D_c]$ is the joint predictive density of the current logarithms of abundance and power troll statistic given the forecast and historical data, and

$p[\ln(x_{now})\,|\,\ln(z_{now}), D_f, D_c]$ is the marginal predictive density of the current logarithm of power troll statistic given the forecast and historical data.

Three statistical models, $p(D_f|\psi)$, with increasing temporal dependence of CTC model forecast errors are constructed from D_f: independent, autocorrelated, and linear trending errors. Several aims underlie this extended modeling: to reflect to some extent the apparent diversity of initial beliefs of CTC members about the reliability of the CTC model forecasts; to focus attention on the importance of understanding the statistical behavior of the CTC model forecast; to illustrate the flexibility of Bayesian calibration for dealing with this behavior; and to provide a check on robustness of the current year abundance estimation. These models are used with the current forecast and noninformative prior densities for parameters, $p(\psi)$, to derive three corresponding predictive densities, $p[\ln(\tau_{now})\,|\,\ln(z_{now}), D_f]$. These predictive densities can be employed before the fishing season and are called the independence, autocorrelation, and trend priors for current abundance.

A regression model, $p(D_c|\boldsymbol{\beta})$, relating the power troll statistic to the abundance index is developed from D_c. This model is used with a noninformative prior density for its parameters, $p(\boldsymbol{\beta})$, to derive the predictive conditional density for the current logarithm of the power troll statistic given abundance, $p[\ln(x_{now})|\ln(\tau_{now}),D_c]$, which hereafter is called the predictive power troll density.

Corresponding to each of the three abundance priors, the joint predictive density at equation (1a), $p[\ln(\tau_{now}),\ln(x_{now})|\ln(z_{now}),D_f,D_c]$, hereafter called the log-log bivariate density, is estimated by kernel smoothing from a Monte Carlo sample of 10,000 pairs, $[\ln(\tau_{now})^*,\ln(x_{now})^*]$. This sample is drawn using equation (1b) and the method of composition (sec. 3.3.2 of Tanner 1996) with the appropriate abundance prior and the predictive power troll density (Appendix). The estimated log-log bivariate density is transformed by probability calculus to the original scales of measurement, i.e.,

$$\begin{aligned} &p[\tau_{now},x_{now}\,|\,\ln(z_{now}),D_f,D_c] \\ &= p[\ln(\tau_{now}),\ln(x_{now})\,|\,\ln(z_{now}),D_f,D_c]\,/\,e^{\ln(\tau_{now})}\times e^{\ln(x_{now})}], \end{aligned} \tag{6}$$

and the result, $p(\tau_{now},x_{now})|\ln(z_{now}),D_f,D_c)$, is called the retransformed bivariate density. A numerical approximation to a normalized section through this retransformed bivariate density at x_{now} is the reported Bayes calibrative density for abundance measured in its original scale, i.e.,

$$p[\tau_{now}\,|\,x_{now},\ln(z_{now}),D_f,D_c] = \frac{p[\tau_{now},x_{now}\,|\,\ln(z_{now}),D_f,D_c]}{\int_{-\infty}^{+\infty} p[\tau_{now},x_{now}\,|\,\ln(z_{now}),D_f,D_c]d\tau_{now}} . \tag{7}$$

Maximum Likelihood Calibration

The maximum likelihood method was originally proposed by the CTC for a comparative study of possible forecast methods (CTC 1997). This classical solution to the calibration problem is to maximize a likelihood function for the logarithm of current abundance, $\ln(\tau_{now})$, and the power troll regression parameters, $\boldsymbol{\beta}$, viz., $p[\ln(x_{now})|\ln(\tau_{now}),\boldsymbol{\beta}] \times p(D_c|\boldsymbol{\beta}) = p[\ln(x_{now})|\ln(\tau_{now}),\boldsymbol{\beta})\times\Pi p[\ln(x_i)|\ln(\tau_i),\boldsymbol{\beta})$. Here, the likelihood function arises from a linear regression model relating the logarithm of the power troll statistic to logarithm of abundance. Therefore, the maximum likelihood estimate for $\boldsymbol{\beta}$ is the least squares estimate from D_c and that for $\ln(\tau_{now})$ is obtained by solving the least squares regression line for $\ln(\tau_{now})$ corresponding to $\ln(x_{now})$ (sec. 10.7 of Aitchison and Dunsmore 1975). The forecast for abundance (τ_{now}) is set equal to the exponentiated maximum likelihood estimate for $\ln(\tau_{now})$, i.e., to the estimated median value in the original abundance scale. The maximum likelihood method does not use the prior density for ln(τ_{now}).

Cross Validation

The CTC model forecast, maximum likelihood estimator, and Bayes estimator of abundance are compared by cross validation (Mosteller and Tukey 1977, CTC 1997) for years 1987 through 1996. The CTC performed 10 separate calibrations using the current structure of the CTC model to hindcast its forecast for SEAK abundance from data available prior to each fishing season (CTC 1997). Annual cross validation estimates for the maximum likelihood and Bayes calibration methods are computed by reserving test year data from D_f and D_c, i.e., only data from years preceding and succeeding the test year are used.

Box-Cox Power Transformation

The lognormal distribution is assumed for both forecast errors and sampling variation of the power troll statistic. The assumption is more readily acceptable for forecast errors in view of its standard usage to describe recruitment variability. Appropriateness of the assumption for the power troll statistic is checked by data analysis. The profile likelihood method (Aitkin et al. 1994) is used to estimate the appropriate power parameter, λ, of the Box-Cox (Box and Cox 1964) transformation for the power troll statistic by minimizing $-2 \times pl(\lambda)$, where $pl(\lambda)$ is the profile log-likelihood function. Minimization of $-2 \times pl(\lambda)$ with respect to λ is performed with the Fortran program NPSOL (Gill et al. 1986) in the range [–5, 5] with an initial guess, 1.0.

Kernel Smoothing

The log-log bivariate density is estimated with the function bkde2D of KernSmooth (Wand 1997). Bandwidths are chosen with the function dpik using the biweight kernel. Contours of density values from bkde2D for 300 × 300 equally-spaced points covering the Monte Carlo sample's range are plotted with MathSoft's *S*-plus contour function.

Results

The general calibration methodology is applied to estimation of chinook salmon abundance in the SEAK fishery during the summer of 1997. The earliest that the CTC will evaluate this abundance (τ_{now}) will be later in the spring of 1998 when all the various data sources are available; the CTC model calibrations are published as PSC reports. The CTC had forecast from its model a point estimate of abundance (z_{now}) of 1.33 in advance of the initial fishery opening in July of 1997; this abundance would be intermediate among values since 1980.

Predictive Prior Densities for 1997 Abundance Index

Hindcasts of CTC model forecasts (z) agree well ($\hat{\rho} = 0.93$, $n = 10$) with the abundance indices (τ) for 1987-1996 (Fig. 1). In the logarithm scale, forecasts standardized to corresponding abundance ($r = z/\tau$) are additive forecast errors, i.e., $\ln(r) = \ln(z) - \ln(\tau)$: negative values result from low forecasts; zero, from a correct forecast; and positive values, from high forecasts. These logarithms appear weakly and positively autocorrelated (lag-1 autocorrelation coefficient = +0.33, $n = 9$, one-tailed $p = 0.19$) if at all (Fig. 2). However, their time series seems to trend downward (slope = -0.032 yr^{-1}, $t = -3.72$, two-tailed $p < 0.01$) with interruptions in 1990 and especially in 1994 (Fig. 3). Although predictions from an ideal forecaster would be unbiased with independent annual errors, in reality the CTC model predictions tend to be too high in early years and too low in recent years. Nonetheless, the forecast series is short and exceptions to the trend in forecast error occur, so projecting future forecast error is problematic.

Next, CTC forecast errors are modeled under three assumptions regarding their stochastic nature: the $\ln(r)$ are assumed to be normally distributed and either (1) independent, (2) autocorrelated, or (3) trending. The distributional assumption is adopted in view of its routine use for describing recruitment variability about expected recruitment (Hilborn and Walters 1992); forecasts are projections of future abundance from approximations to expected abundance. Each model is used to derive a predictive prior density for 1997 abundance, $\ln(\tau_{now})$, given the 1997 CTC forecast, $\ln(z_{now})$, and possibly other historical data of D_f: the steps include defining the probability model for current $\ln(r)$, $p[\ln(r_{now}) | D_f, \boldsymbol{\psi}]$, from the basic assumption; transforming it to a probability model for current abundance, $p[\ln(\tau_{now}) | \ln(z_{now}), D_f, \boldsymbol{\psi}]$, by the relationship that $\ln(\tau_{now}) = \ln(z_{now}) - \ln(r_{now})$; and integrating this result by the Monte Carlo method (see equation [2] and Appendix) over uncertainty in $\boldsymbol{\psi}$ as described by the probability model, $p(\boldsymbol{\psi} | D_f)$. In the following model descriptions, notice that when the density for $\ln(r_{now})$ is $N(a,b)$ for any a and b, then the density for $\ln(\tau_{now})$ is $N[\ln(z_{now}) - a, b]$ due to the linear relationship between these random variables and fixed $\ln(z_{now})$. The local definition of parameter $\boldsymbol{\psi}$ at equations (2) and (4) is provided in the heading for each model.

1. *Independence prior, $\psi = \sigma_f$: unbiased forecasts with independent lognormal errors.*

Logarithms of the standardized forecasts are assumed to be normally distributed [$\ln(r) \sim N(\mu_f, \sigma_f^2)$]. If true, the expected value of r is $\exp(\mu_f + ½\sigma_f^2)$. Then unbiasedness of forecasts, i.e., $E(r) = 1$, is equivalent to the equality constraint, $\mu_f = -½\sigma_f^2$. Therefore, a single parameter (e.g., σ_f) will describe the error distribution. If a noninformative, improper prior is used for σ_f, viz., $p(\sigma_f) \propto 1$, the posterior for σ_f (if n forecasts were available) is its likelihood function,

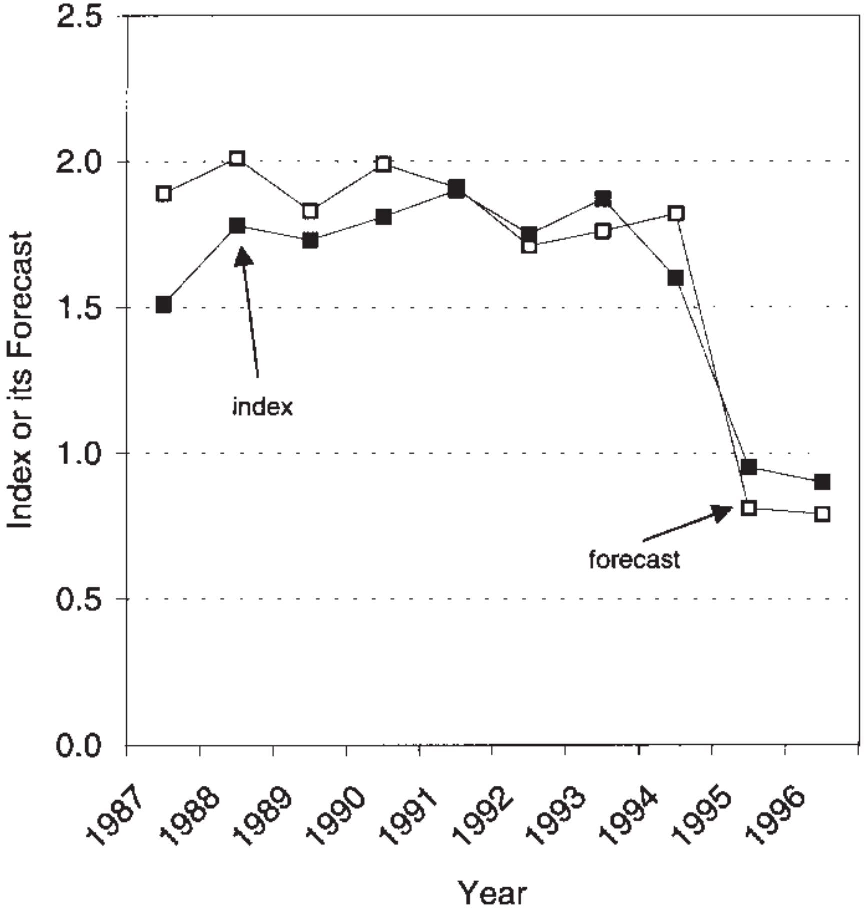


Figure 1. Time series of the abundance index (■) and its hindcasted CTC model forecast (□), 1987-1996.

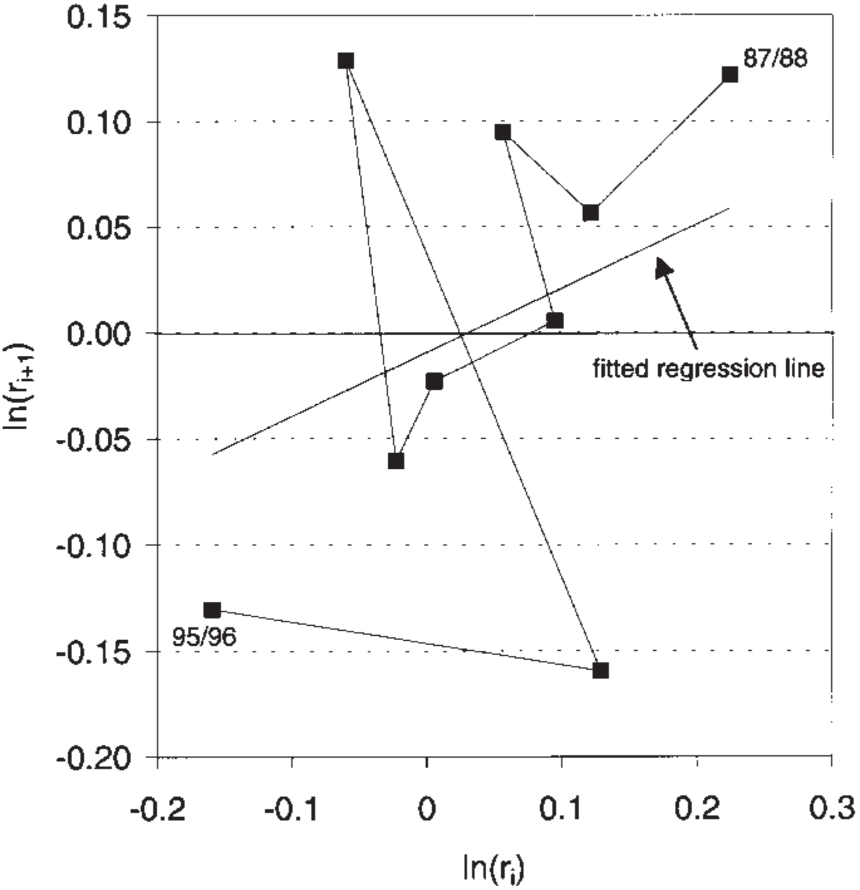


Figure 2. Logarithms of standardized forecasts in successive years for 1987-1996 ($\ln[r_{i+1}]$ vs. $\ln[r_i]$) joined in temporal order (1987-1988 to 1995-1996), and the estimated mean from fitting a linear regression model.

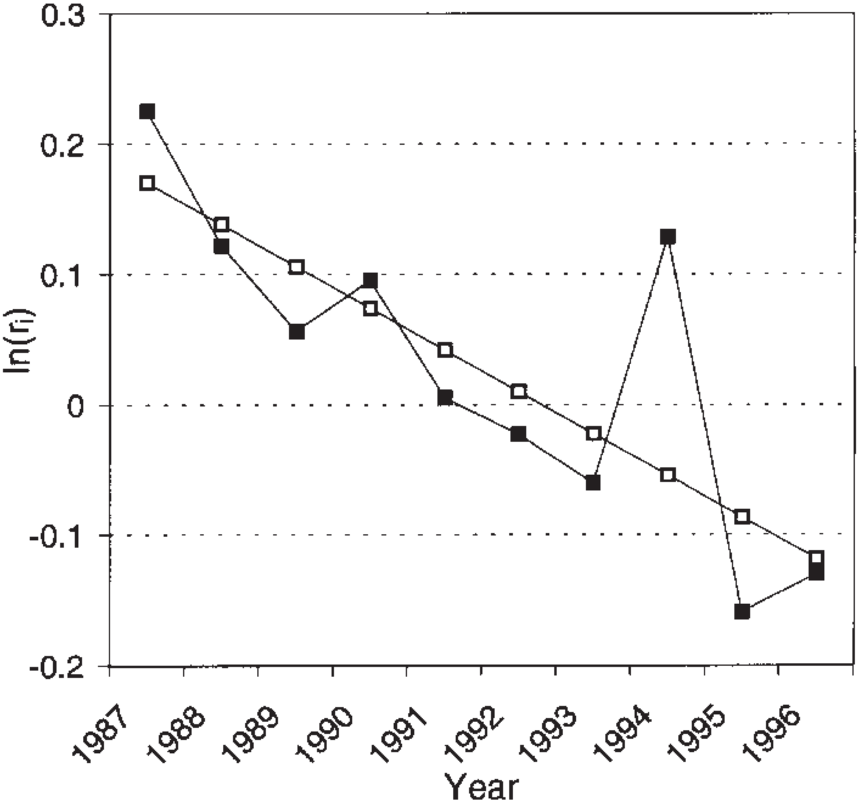


Figure 3. Time series of logarithms of standardized forecasts (■), ln(r_i), and fitted values from a linear trend regression model (□), 1987-1996.

$$p(\sigma_f \mid D_f) \propto \left(\frac{1}{\sigma_f} e^{-\frac{1}{8}\sigma_f^2} \right)^n e^{-\frac{1}{2\sigma_f^2}\Sigma[\ln(r_i)]^2} \; (\sigma_f > 0). \tag{8}$$

The normalized version of the posterior for σ_f is a skewed distribution with maximum occurring at the maximum likelihood estimate for σ_f, 0.118345 (Fig. 4). The conditional density, $p[\ln(\tau_{now}) \mid \ln(z_{now}),\sigma_f]$, is $N[\ln(z_{now}) + \frac{1}{2}\sigma_f^2, \sigma_f^2]$. To obtain the predictive density, $p[\ln(\tau_{now}) \mid \ln(z_{now}),D_f]$, the product of this conditional density and the posterior density for σ_f is integrated over the range of σ_f as in equation (2).

2. Autocorrelation prior, $\psi = (\phi_0, \phi_1, \sigma_f^2)'$: autocorrelated, lognormal forecast errors.

The time series of ln(r) from 1987 onward is next modeled as a first-order autoregressive process,

$$\ln(r_{i+1}) = \phi_0 + \phi_1 \ln(r_i) + \varepsilon_{i+1}, \quad i = 1,2,\ldots,9 \tag{9}$$

where

$r_i = z_i / \tau_i$ is the standardized forecast for the ith year,

ϕ_0 and ϕ_1 are the autoregressive intercept and slope, respectively, and

ε_{i+1} is a normal random annual disturbance with mean of zero and variance of σ_f^2.

If the annual disturbances are independent and the random process begins with a fixed value at 1987, then the maximum likelihood estimators for the parameters are those from ordinary least squares regression of $\ln(r_{i+1})$ on $\ln(r_i)$ (sec. 3.2 of Harvey 1994) (Table 2, Fig. 2). The estimated slope, $\hat{\phi}_1$, is positive but not significant (one-tailed $p \approx 0.19$), and the model only explains about 11% of the variation in the ln(r) of 1987-1996.

Using the noninformative prior density, $p(\phi_0,\phi_1,\sigma_f^2) \propto \sigma_f^{-2}$, the posterior density for the autoregressive parameters, $p(\phi_0,\phi_1,\sigma_f^2 \mid D_f)$, equals the product of an inverse chi square and conditional bivariate normal density, $p(\sigma_f^2 \mid s_f^2) \times p(\phi_0,\phi_1 \mid \hat{\phi}_0,\hat{\phi}_1,\sigma_f^2)$ (pp. 17-19 of Tanner 1996),

where

$p(\sigma_f^2 \mid s_f^2) = 7\, s_f^2\, \chi_7^{-2}$ and $p(\phi_0, \phi_1 \mid \hat{\phi}_0, \hat{\phi}_1, \sigma_f^2) = N[(\hat{\phi}_0, \hat{\phi}_1)', \sigma_f^2 (D_\phi' D_\phi)^{-1}]$,

$\hat{\phi}_0$ and $\hat{\phi}_1$ are the intercept and slope estimates from the least squares fit of the autoregressive model to the 9 (= $n - 1$) data pairs, [$\ln(r_i)$, $\ln(r_{i+1})$],

s_f^2 is the mean square error with 7 (= $n - 3$) degrees of freedom, and

D_ϕ is the matrix of independent variables for these data.

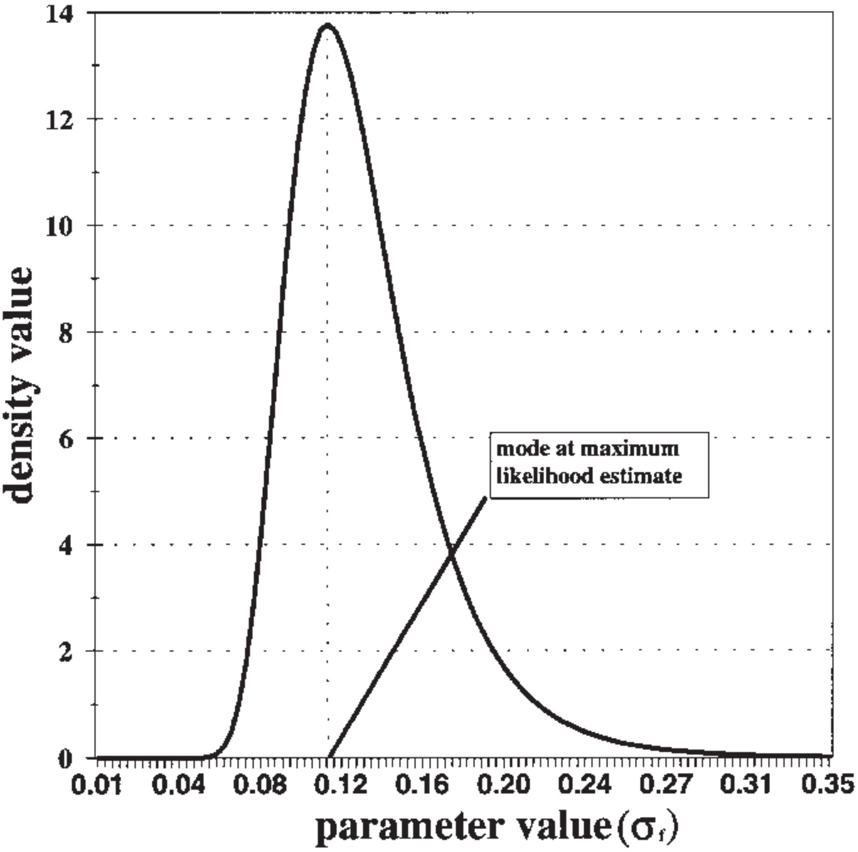


Figure 4. Posterior density for lognormal parameter, σ_f, when CTC model forecasts are unbiased and errors are independent. The mode occurs at the maximum likelihood estimate.

Table 2. Statistics from fitting a linear autoregression model to logarithms of standardized forecasts in successive years.

Autoregressive model:

$$\ln(r_{i+1}) = \phi_0 + \phi_1 \ln(r_i) + \varepsilon_{i+1},$$

where

r_i = standardized forecast in year i, and
ε_i = random error for year i, *iid* $N(0, \sigma_f^2)$.

Parameter	Estimate	Standard error	*t*-statistic	*P*-value
Intercept (ϕ_0)	−0.0093	0.0384	−0.24	0.82
Slope (ϕ_1)	0.3016	0.3277	0.92	0.39
Variance (σ_f^2)	$(0.1071)^2$ on 7 degrees of freedom			

$R^2 = 0.11$

The conditional density, $p[\ln(\tau_{now}) | \ln(z_{now}), D_f, \phi_0, \phi_1, \sigma_f^2]$, is $N[\ln(z_{now}) - \phi_0 - \phi_1 \ln(r_{1996}), \sigma_f^2]$. To obtain the predictive density, $p[\ln(\tau_{now}) | \ln(z_{now}), D_f)$, the product of this conditional density and the posterior for the parameters is integrated over the range of ϕ_0, ϕ_1, and σ_f^2 as in equation (2).

3. *Trend prior, $\psi = (\kappa_0, \kappa_1, \sigma_f^2)'$: linearly trending, lognormal forecast errors.*

The time series of ln(r) from 1987 onward is modeled as a linear process in time,

$$\ln(r_i) = \kappa_0 + \kappa_1 i + v_i \quad i = 8, 9, \ldots, 17, \tag{10}$$

where

$r_i = z_i/\tau_i$ is the standardized forecast for the ith year, after 1979,

κ_0 and κ_1 are the linear process intercept and slope, respectively, and

v_i is a normal random annual disturbance with mean of zero and variance of σ_f^2.

If the annual disturbances are independent, then the maximum likelihood estimators for the parameters are those from ordinary least squares regression of $\ln(r_i)$ on i (Table 3, Fig. 3). The estimate of the slope parameter, κ_1, is negative and significant (two-tailed $p < 0.01$), and the model explains about 63% of the variation in the logarithms of standardized forecasts of 1987-1996.

Table 3. Statistics from fitting a linear trend regression model to logarithms of standardized forecasts.

Linear trend regression model:

$$\ln(r_i) = \kappa_0 + \kappa_1\, i + v_i, \quad i = 8, 9, \ldots, 17,$$

where

r_i = standardized forecast in year i, and
v_i = random error for year i, *iid* $N(0, \sigma_f^{\,2})$.

Parameter	Estimate	Standard error	*t*-statistic	*P*-value
Intercept (κ_0)	0.4266	0.1106	3.86	0.005
Slope (κ_1)	–0.0321	0.0086	–3.72	0.006
Variance (σ_f^2)	$(0.0783)^2$ on 8 degrees of freedom			

$R^2 = 0.63$

Using the noninformative prior density, $p(\kappa_0, \kappa_1, \sigma_f^2) \propto \sigma_f^{-2}$, the posterior density for the linear process parameters, $p(\kappa_0, \kappa_1, \sigma_f^{\,2} \mid D_f)$, equals $p(\sigma_f^2 \mid s_f^2) \times p(\kappa_0, \kappa_1 \mid \hat{\kappa}_0, \hat{\kappa}_1, \sigma_f^2)$ (pp. 17-19 in Tanner, 1996) where

$p(\sigma_f^2 \mid s_f^2) = 8\, s_f^2\, \chi_8^{-2}$ and $p(\kappa_0 \kappa_1 \mid \hat{\kappa}_0, \hat{\kappa}_1, \sigma_f^2) = N[(\hat{\kappa}_0, \hat{\kappa}_1)', \sigma_f^2 (D_\kappa' D_\kappa)^{-1}]$,

$\hat{\kappa}_0$ and $\hat{\kappa}_1$ are the intercept and slope estimates from the least squares fit of the linear trending model to the 10 (= n) data pairs, $[\ln(r_i), i]$,

s_f^2 is the mean square error with 8 (= $n - 2$) degrees of freedom, and

D_κ is the matrix of independent variables for these data.

The conditional density, $p[\ln(\tau_{now}) \mid \ln(z_{now}), \kappa_0, \kappa_1, \sigma_f^2]$, is $N[\ln(z_{now}) - \kappa_0 - 18\kappa_1, \sigma_f^{\,2}]$. To obtain the predictive density, $p[\ln(\tau_{now}) \mid \ln(z_{now}), D_f]$, the product of this conditional density and the posterior for the parameters is integrated over the range of κ_0, κ_1, and σ_f^2 as in equation (2).

Predictive Density for 1997 Power Troll Statistic as Related to Abundance

A survey of potential inseason indices using years 1980 through 1996 (CTC 1997) selected the power troll statistic (PTS), or the number of chinook salmon caught per power troll permit that landed catches during the first five days of the summer season (Table 1). Abundance trended upward for much of the era (1980-1993), but in recent years (1994-1996) declined sufficiently to expose an apparent changing relationship with the PTS (Fig. 5). Specifically, the PTS for given abundance in recent years appears greater than in earlier years: in 1995 and 1996, the PTS was nearly twice that of 1980-1981 even though abundance in 1995 and 1996 was

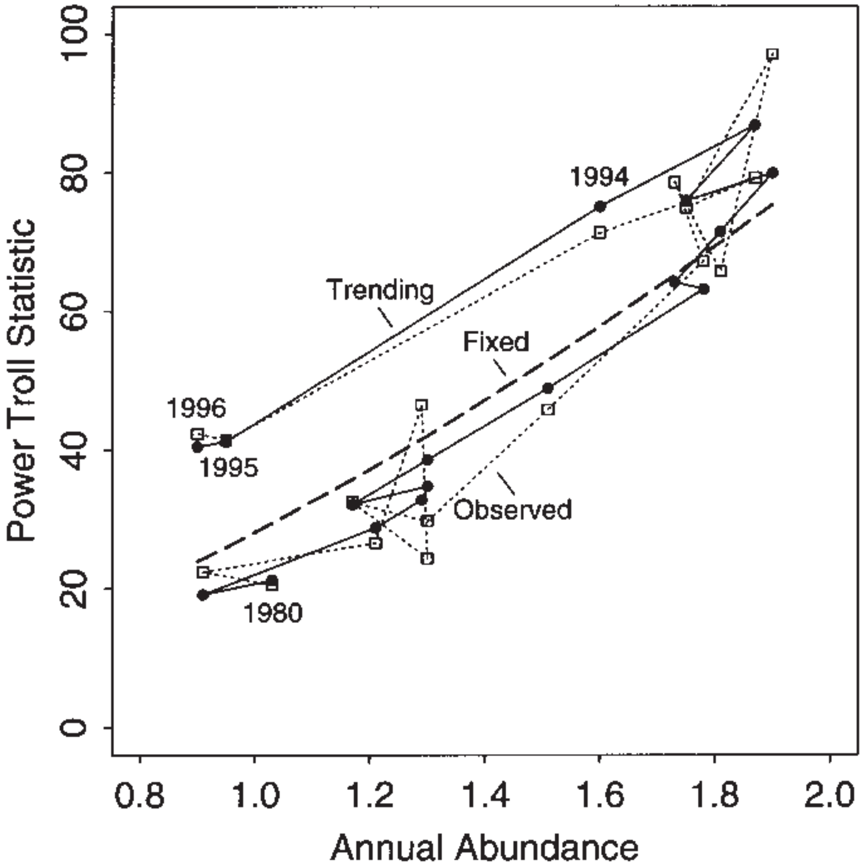


Figure 5. Observed inseason index or power troll statistic, vs. annual abundance index from CTC model (□) and estimated median values of the inseason index for completely fixed (dashed line) and trending intercept (●) models, 1980-1996 (Note that observed- and trending-intercept fitted values are serially connected with dotted and solid lines, respectively).

about as low as in 1980-1981. Any proposed model relating the PTS to abundance must accommodate probable trends in its expected value at fixed abundance. A linear regression model relating logarithms of the PTS and abundance is next considered in which both the intercept and slope are allowed to trend in order to approximate the dynamic relationship.

The relationship of the power troll statistic and abundance is examined for the 17 years of available data from 1980 to 1996 (Table 1, Fig. 5) using the model,

$$y_i = \beta_0 + \beta_1 i + \beta_2 \ln(\tau_i) + \beta_3 i \times \ln(\tau_i) + \eta_i, \quad i = 1, 2, \ldots, 17 \tag{11}$$

where

y_i is the logarithm of the PTS in the ith year,

τ_i is the abundance from the CTC model for the ith year,

β_0, β_1, β_2, and β_3 are constants, and

η_i is the normal random annual error with mean of zero and variance of σ_η^2.

The combination, $\beta_0+\beta_1 i$, is a trending intercept and the combination, $\beta_2+\beta_3 i$, is a trending slope for the relationship between ln(PTS) and ln(τ). The fit to the data, D_c, indicates slope does not trend detectably (i.e., $\beta_3 = 0$), but the intercept does ($\beta_1 \neq 0$). The reduced model with intercept (β_0) ($p < 0.0001$), intercept trend (β_1) ($p < 0.0001$) and fixed slope (β_2) ($p < 0.0001$) accounts for 89% of the variation in the dependent variable (Table 4). The estimate of β_1 implies that the PTS at fixed abundance increased at roughly 5% per annum, i.e., $\exp(\hat{\beta}_1) \cong 1.05$. Assuming the actual relationship changed with time, possible causes currently contemplated by the CTC that require further study include increased susceptibility to SEAK fishing, improvements in vessel efficiency, or changes in abundance before the summer fishery induced by earlier fishing activity (CTC 1997). Median model values for untransformed PTS from a completely fixed model (fixed intercept and fixed slope) and trending intercept model (fixed slope) are obtained by exponentiating expected values from equation (11) with estimated values (Table 4) replacing unknown parameters (Fig. 5). Median values from the completely fixed model tend to overestimate the PTS in early years and underestimate the PTS in recent years; those from the trending intercept model are more accurate.

The logarithmic transformations of both the power troll statistic and abundance for the inseason regression model at equation (11) are justified by the outcome of a search for a good power transformation and the reasonable model that results. The estimated power from the Box-Cox analysis, $\hat{\lambda} = 0.05$, corresponds approximately to the natural logarithm transformation, i.e., $\lambda = 0$. The fit of the regression model to the power-transformed inseason index with either value for λ produces negligible

Table 4. Inseason regression models and statistics of their fit.

PTS regression model:

$$\ln(x_i) = \beta_0 + \beta_1 \mathrm{i} + \beta_2 \ln(\tau_i) + \beta_3 \, i \times \ln(\tau_i) + \eta_i, \; i = 1, 2, \ldots, 17,$$

where

x_i = PTS in year *i*,
τ_i = abundance index in year *i*, and
η_i = random error for year *i*, *iid* $N(0, \sigma_\eta^2)$.

Parameter	Estimate	Standard error	*t*-statistic	*P*-value
Intercept (β_0)	3.3351	0.1188	28.08	2.20×10^{-14}
(not included) (β_1)	0	0		
Slope (β_2)	1.5374	0.2937	5.23	0.0001
(not included) (β_3)	0	0		
Variance (σ_η^2)	$(0.3089)^2$ on 15 degrees of freedom			
Model 1: Fixed intercept and fixed slope, $R^2 = 0.65$				
Intercept (β_0)	2.9806	0.1206	24.72	2.59×10^{-12}
Intercept trend (β_1)	0.0500	0.0108	4.62	0.0005
Slope (β_2)	1.1240	0.5829	1.93	0.0759
Slope trend (β_3)	0.0109	0.0470	0.23	0.8210
Variance (σ_η^2)	$(0.1874)^2$ on 13 degrees of freedom			
Model 2: Trending intercept and trending slope, $R^2 = 0.89$				
Intercept (β_0)	2.9653	0.0972	30.52	3.29×10^{-14}
Intercept trend (β_1)	0.0511	0.0094	5.45	0.0001
Slope (β_2)	1.2515	0.1798	6.96	6.69×10^{-6}
(not included) (β_3)	0	0		
Variance (σ_η^2)	$(0.1810)^2$ on 14 degrees of freedom			
Model 3: Trending intercept and fixed slope, $R^2 = 0.89$				

difference in the multiple correlation coefficient (R): $R^2 = 0.8867$ ($\hat{\lambda} = 0$) and $R^2 = 0.8865$ ($\hat{\lambda} = 0.05$). The retransformed (from the logarithmic fit) relationship between the inseason index and abundance approaches the origin as abundance approaches zero, thereby agreeing with common sense.

Using the noninformative prior density, $p(\beta_0,\beta_1,\beta_2,\sigma_\eta{}^2) \propto \sigma_\eta{}^{-2}$, the posterior density, $p(\beta_0,\beta_1,\beta_2,\sigma_\eta{}^2 | D_c)$, equals $p(\sigma_\eta^2 \mid s_\eta^2) \times p(\beta_0, \beta_1, \beta_2 \mid \hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2, \sigma_\eta^2)$ (pp. 17-19 of Tanner 1996) where

$p(\sigma_\eta^2 \mid s_\eta^2) = 14\, s_\eta^2\, \chi_{14}^{-2}$ and

$p(\beta_0, \beta_1, \beta_2) \mid \hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2), \sigma_\eta^2) = N[(\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2)', \sigma_\eta^2 (\mathrm{D}_\beta{}', \mathrm{D}_\beta)^{-1}]$,

$\hat{\beta}_0, \hat{\beta}_1,$ and $\hat{\beta}_2$ are the estimated coefficients from the least squares fit of the inseason regression model to the 17 (= n) annual observations,

$\sigma_\eta{}^2$ is the mean square error with 14 (= n–3) degrees of freedom, and

D_β is the matrix of independent variables for these data.

The conditional density, $p[\ln(x_{now}) | \ln(\tau_{now}), \beta_0, \beta_1, \beta_2, \sigma_\eta{}^2]$ is $N[\beta_0 + 18\beta_1 + \beta_2 \ln(\tau_{now}), \sigma_\eta{}^2]$. To obtain the predictive power troll density, $p[\ln(x_{now}) | \ln(\tau_{now}), D_c]$, the product of this conditional density and the posterior for the parameters is integrated by the Monte Carlo method (see equation [3] and Appendix) over the range of the parameters, β_0, β_1, β_2, and $\sigma_\eta{}^2$.

Kernel Smoothed Density Estimates

The estimated log-log bivariate densities are roughly symmetric with the direction of the long axis reflecting the positive dependence of the inseason index on abundance (Fig. 6, left side). Their centers of mass are to the right of (greater than) the logarithm of the forecast, and they shift progressively to the right and upward (toward greater values of both variables) with increase in dependence of forecast error assumed in the priors for abundance, i.e., from independence to autocorrelation to trend priors. The center of mass on the abundance axis is determined from that of the abundance prior, and the abundance prior results at Equation (2) from the weighted averaging of the conditional density, $p[\ln(\tau_{now}) | \ln(z_{now}), D_f, \psi]$, for uncertainty in the parameters, ψ. The expected value for this conditional density, i.e., $E[\ln(\tau_{now}) | \ln(z_{now}), D_f, \psi] = \ln(z_{now}) - E[\ln(r_{now})]$, increases with temporal dependence of the prior from $\ln(z_{now}) + \frac{1}{2}\sigma_f^2$ (independence), to $\ln(z_{now}) - \phi_0 - \phi_1 \ln(r_{1996})$ (autocorrelation), to $\ln(z_{now}) - \kappa_0 - 18\kappa_1$ (trend). Recent forecasts have been below corresponding abundance ($r < 1$, $\ln(r) < 0$) (Figs. 1-3) and are projected to remain so or to become more biased from the models underlying the abundance priors.

The retransformed bivariate densities in the original scales of measurement are asymmetric and their centers of mass are shifted similarly

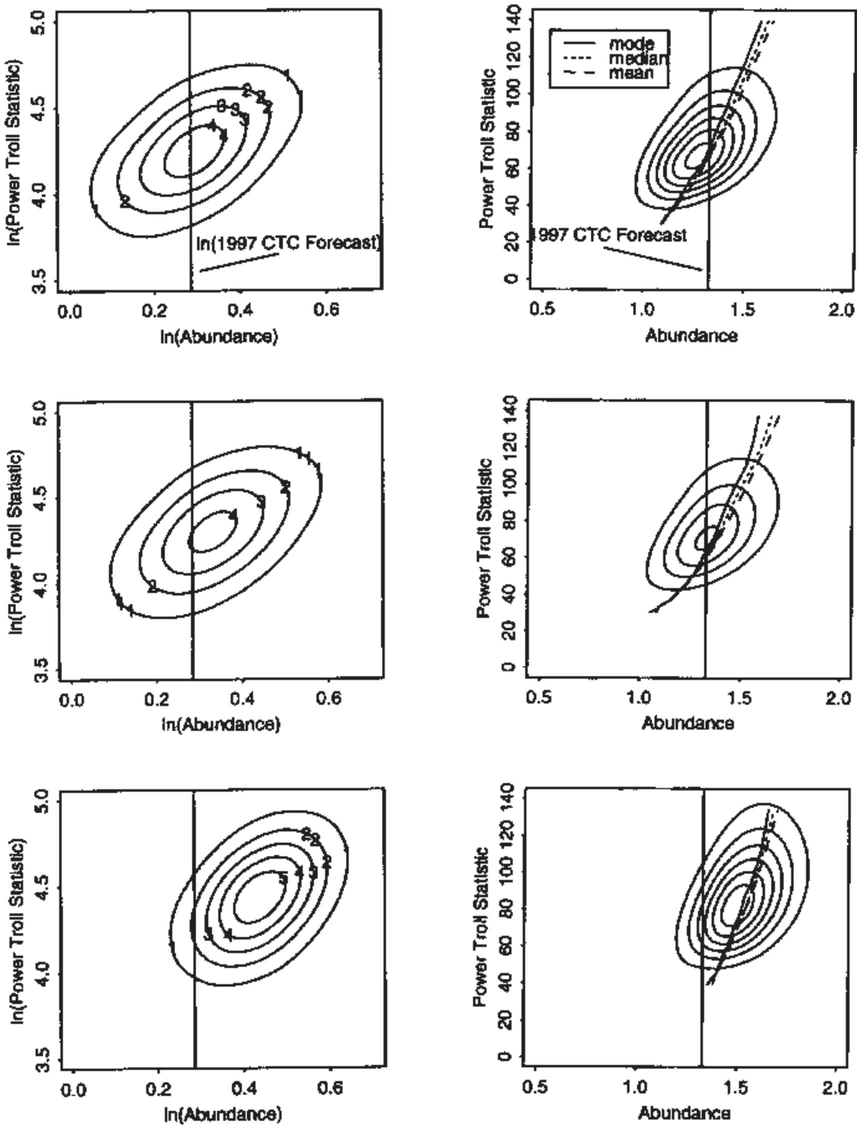


Figure 6. *Bivariate predictive densities—logarithm scale (left) and original scale (right)—for 1997 SEAK abundance index and inseason index resulting from prior densities of abundance based on increasing temporal dependence of forecast errors: independent (top), autocorrelated (middle), and trending (bottom).*

(Fig. 6, right side). The relative positions of the mode, median, and mean of the calibrative density, i.e., normalized section of the retransformed bivariate density, for any value of the power troll statistic are maintained: the mode is smallest, the median is intermediate, and the mean is largest (Fig. 6, right side). If these parameters of the calibrative density are used to estimate abundance, the mode is most conservative, i.e., estimated abundance is smallest, and the mean is least conservative, i.e., estimated abundance is largest. The influence of an increasing PTS on estimated abundance continuously diminishes as can be seen from the accelerating curvature of mode, median, and mean lines toward the vertical (Fig. 6, right side); ultimately, increase in the PTS has no further effect on estimated abundance. The abundance priors limit estimated abundance.

In this initial application of the method, the CTC model in spring 1997 forecast a middling abundance of 1.33, but the power troll statistic from the summer fishery of early July (106.5 fish per permit) was the largest on record. As a result, the probability mass of each of the three prior densities for 1997 abundance shifts rightward and mildly concentrates in the corresponding calibrative density (Fig. 7). The calibrative distributions under the independence, autocorrelation, and trend priors (Fig. 8) have modes at 1.46, 1.51, and 1.60, respectively. Corresponding calibrative variances equal 0.041, 0.047, and 0.027, down 5.2%, 10.4%, and 17.5% from the prior variances, respectively. Using the mode as an estimate of the abundance from future postseason analysis with the CTC model, that record PTS increases estimated abundance 10% to 20% from the forecast value, depending on the prior.

Cross Validation

The cross validation experiment shows that in retrospect Bayesian calibration is moderately superior to CTC model forecasts or maximum likelihood calibration if judged by mean percent error or mean absolute percent error (Table 5). The Bayesian calibration from the trend prior performs well (<9% error) in nine of ten years; its annual absolute percent errors are less than or equal to those of the CTC model forecasts in eight of ten years, and those of maximum likelihood calibration in seven of ten years. Bayesian calibrations from the independence or autocorrelation priors have smaller maximum percent errors (independence prior, 15%; autocorrelation prior, 19%) than from the trend prior (23%) or the other methods (CTC model, 25%; maximum likelihood calibration, 22%).

Afterword

The calibrative density integrates the information on forecast error by the CTC model, the regression relationship between the inseason index and abundance, the current forecast from the CTC model, and the current summer fishery performance. In doing so, it accounts for uncertainty in the unknown parameters of specified models comprising the predictive

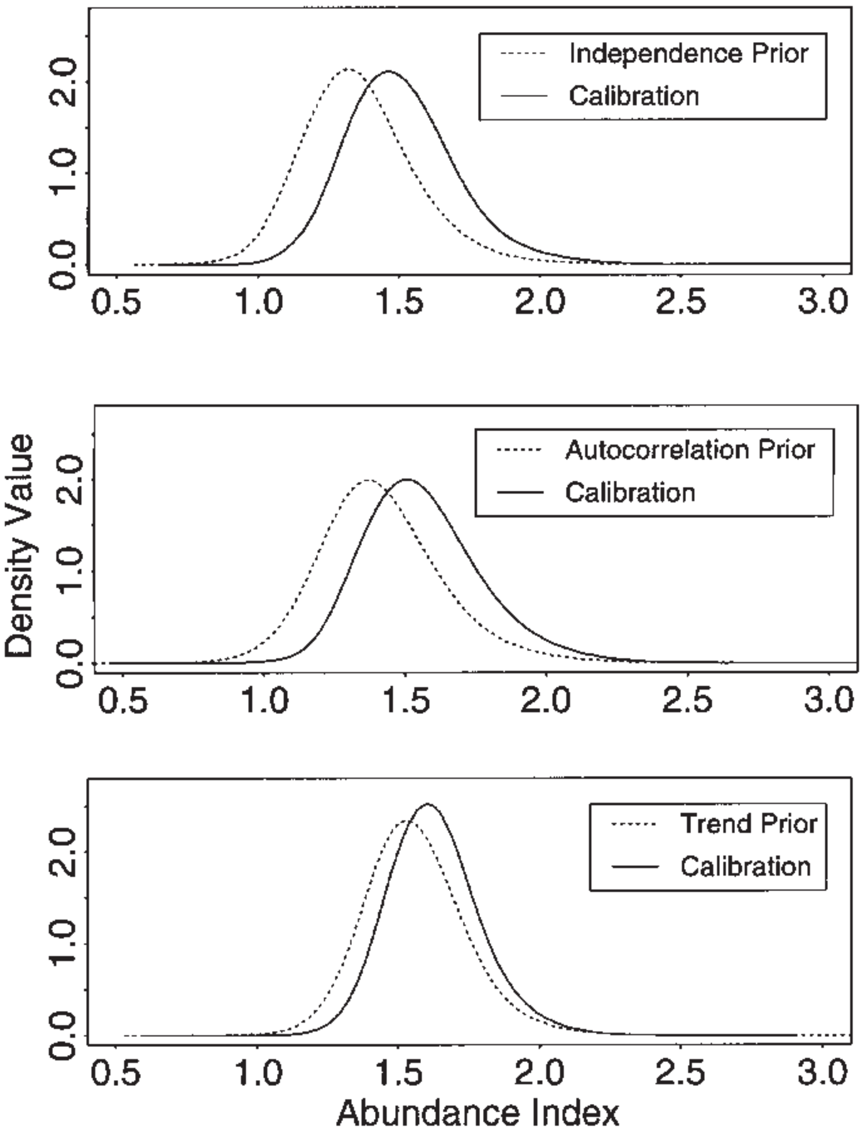


Figure 7. The prior and calibrative densities for 1997 SEAK abundance given the inseason index of 106.5 fish per permit, corresponding to increasing temporal dependence of forecast errors: independent (top), autocorrelated (middle), and trending (bottom).

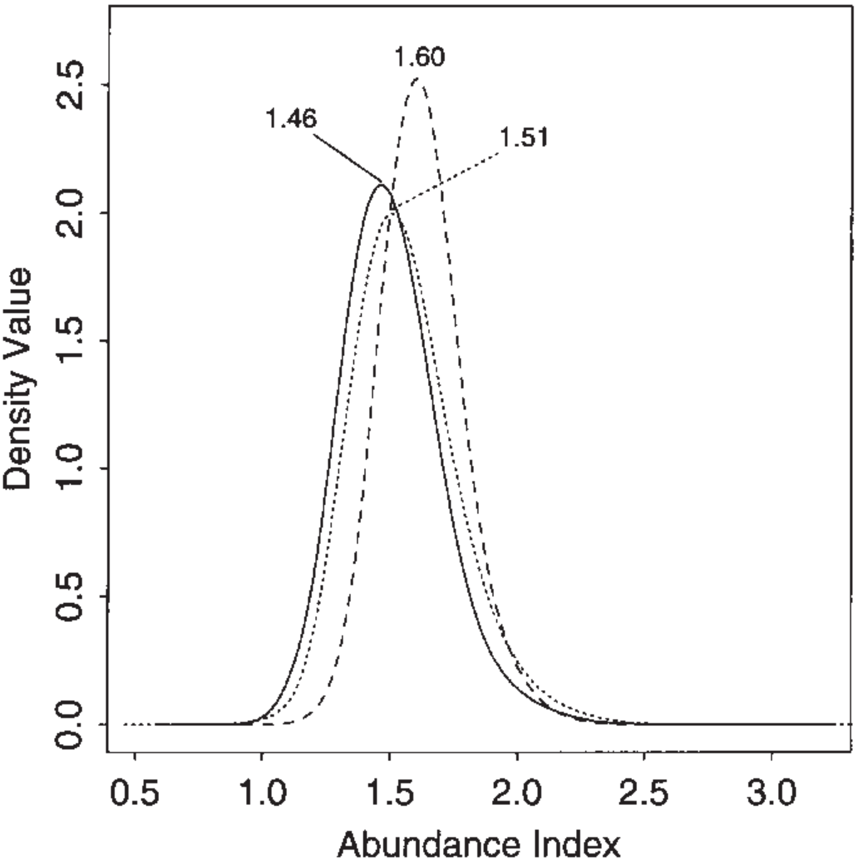


Figure 8. The calibrative densities and modes for 1997 SEAK abundance given the observed inseason index of 106.5 fish per permit, corresponding to prior densities of increasing temporal dependence of forecast errors: independent (solid line), autocorrelated (dotted line), and trending (dashed line).

Table 5. **Annual abundance indices from the CTC model and their forecasts from the CTC model, maximum likelihood calibration of the inseason indices, and Bayesian calibration combining CTC model forecasts and inseason indices.**

Year	Abundance index	CTC model forecast Projected	CTC model forecast % Error	Maximum likelihood Projected	Maximum likelihood % Error
1987	1.51	1.89	+25	1.43	−5
1988	1.78	2.01	+13	1.88	+6
1989	1.73	1.83	+6	2.09	+21
1990	1.81	1.99	+10	1.68	−7
1991	1.90	1.91	+1	2.31	+22
1992	1.75	1.71	−2	1.73	−1
1993	1.87	1.76	−6	1.72	−8
1994	1.60	1.82	+14	1.52	−5
1995	0.95	0.81	−15	0.96	+1
1996	0.90	0.79	−12	0.96	+7
MPE			+3		+3
MAPE			10		8

		Bayesian calibration					
	Abundance	Independence		Autocorrelation		Trend	
Year	Index	Projected	% Error	Projected	% Error	Projected	% Error
1987	1.51	1.73	+15	–	–	1.57	+4
1988	1.78	1.95	+10	1.88	+6	1.74	−2
1989	1.73	1.88	+9	1.84	+6	1.68	−3
1990	1.81	1.86	+3	1.86	+3	1.80	−1
1991	1.90	1.98	+4	1.94	+2	1.88	−1
1992	1.75	1.70	−3	1.69	−3	1.67	−5
1993	1.87	1.73	−7	1.71	−9	1.76	−6
1994	1.60	1.71	+7	1.81	+13	1.97	+23
1995	0.95	0.83	−13	0.77	−19	0.87	−8
1996	0.90	0.81	−10	0.80	−11	0.88	−2
MPE			+1		−1		−0.1
MAPE			8		8		5

Bayesian calibration includes three priors for 1997 SEAK abundance based on independent and unbiased errors, autocorrelated errors, or time trending errors. MPE is the mean percent error, and MAPE is the mean absolute percent error.

priors for abundance and the predictive power troll density as well as in random variables depending on the parameter values, i.e., all the densities are Bayes predictive densities. At present, the Bayesian calibrative distribution has been accepted by the CTC as a satisfactory technical description of uncertainty in chinook salmon abundance in the SEAK fishery.

Future application of the Bayesian methodology to abundance estimation in regional fisheries may well depend on interplay with CTC model development. The CTC may wish to incorporate Bayesian estimation in their model: at present, only point values are computed for abundance without measures of their precision or accuracy. The CTC has chosen the mode as a point estimate of abundance, but recommends that the PSC consider fuller use of the calibrative density (CTC 1997). A risk or utility function is customarily specified by Bayesians for selecting an estimator, and this approach uses the complete calibrative density. An acceptable function for this purpose might require predicting long-term consequences of errors in abundance estimates on the chinook stocks and their subsequent abundance in fisheries. Such prediction would require detailed modeling of stock and fishery dynamics similar to that by the CTC model.

Acknowledgments

The helpful discussions with members of the CTC committee guided our development of the Bayesian methodology. The suggestion by Jim Scott to perform the cross-validation experiment was first-rate.

References

Allen, W.R., D. Benton, and R. Turner. 1996. Letter of agreement regarding an abundance-based approach to managing chinook fisheries in Southeast Alaska. U.S. Section Pacific Salmon Commission agreement, June 24, 1996.

Aitchison, J., and I.R. Dunsmore. 1975. Statistical prediction analysis. Cambridge University Press, London. 273 pp.

Aitkin, M., D. Anderson, B. Francis, and J. Hinde. 1994. Statistical modeling in GLIM. Oxford University Press, New York. 374 pp.

Box, G.E.P., and D.R. Cox. 1964. An analysis of transformations. Journal of the Royal Statistical Society, Series B 26:211-243.

CTC (Chinook Technical Committee). 1995. A review of the ADFG proposal for an abundance-based chinook management regime. Pacific Salmon Commission Report TCChinook (95)-1. 37 pp. + appendices.

CTC (Chinook Technical Committee). 1997. Evaluation of three methods for predicting the abundance index for chinook salmon available to the Southeast Alaska troll fishery. Pacific Salmon Commission Report TCChinook (97)-3. 39 pp. + appendices.

Geiger, H.J., B. Frenette, and D. Hart. 1997. Run forecasts and harvest projections for 1997 Alaska salmon fisheries and review of the 1996 season. Commercial Fisheries Management and Development Division, Regional Information Report 5J97-10, Juneau, AK. 33 pp. + appendices.

Gelman, A., J.B. Carlin, H.S. Stern, and D.B. Rubin. 1996. Bayesian data analysis. Chapman and Hall, London. 526 pp.

Gill, P.E., W. Murray, M.A. Saunders, and M.H. Wright. 1986. User's guide for NPSOL (Version 4.0): a Fortran package for nonlinear programming. Technical Report SOL 86-2. 54 pp.

Harvey, A.C. 1994. Time series models. 2nd edn. The MIT Press, Cambridge, MA. 308 pp.

Hilborn, R., and C. J. Walters. 1992. Quantitative fisheries stock assessment. Choice, dynamics and uncertainty. Chapman and Hall, New York. 570 pp.

Koenings, J.P., D. Gaudet, S. Marshall, and J.E. Clark. 1995. The 1995-1998 fishery regime for Southeast Alaska chinook salmon fisheries. Regional Information Report No. 5J95-14. Alaska Department of Fish and Game, Commercial Fisheries Management and Development Division. P.O. Box 25526, Juneau, AK.

Minard, R.E., and C.P. Meacham. 1987. Sockeye salmon *(Oncorhynchus nerka)* management in Bristol Bay, Alaska. In: H.D. Smith, L. Margolis, and C.C. Wood (eds.). Sockeye salmon *(Oncorhynchus nerka)* population biology and future management. Can. Spec. Publ. Fish. Aquat. Sci. 96:336-342.

Mosteller, F., and J.W. Tukey. 1977. Data analysis and regression: A second course in statistics. Addison-Wesley, Reading, MA. 588 pp.

Mundy, P.R., K.K. English, W.J. Gazey, and K.E. Tarbox. 1993. Evaluation of the harvest management strategies applied to sockeye salmon populations of upper Cook Inlet, Alaska, using run reconstruction analysis. In: G. Kruse, D.M. Eggers, R.J. Marasco, C. Pautzke, and T.J. Quinn II (eds.). Proceedings of the International Symposium on Management Strategies for Exploited Fish Populations. University of Alaska Sea Grant Report 93-02, Fairbanks.

Press, S.J. 1989. Bayesian statistics: Principles, models, and applications. John Wiley and Sons, New York. 237 pp.

Press, W., B. Flannery, S. Teukolsky, and W. Vetterling. 1989. Numerical recipes: The art of scientific computing. FORTRAN Version. Cambridge University Press, New York. 702 pp.

Tanner, M.A. 1996. Tools for statistical inference: Methods for the exploration of posterior distributions and likelihood functions. 3rd edn. Springer, New York. 207 pp.

Wand, M.P. 1997. KernSmooth 2.2 reference manual. Australian Graduate School of Management, University of New South Wales, Sydney, Australia. 12 pp.

Appendix

Ten thousand pairs of pseudorandom values are sampled from each of three log-log bivariate densities corresponding to the independence, autocorrelation, and trend priors. These pairs are obtained by Monte Carlo sampling from the appropriate abundance prior and the predictive power troll density, using the method of composition (pp. 52-54 of Tanner 1996). In this sampling, draws from posterior densities for parameters of normal linear models are performed (pp. 17-19 of Tanner 1996). To draw samples from the posterior density at Equation (8), the transformation method (Press et al. 1989) is used. The distribution function is evaluated at 500 equispaced points in the range of σ_f between 0.001215 and 0.607500. The integrations for evaluating the distribution function are performed with subroutine QSIMP (parameters eps = 10^{-5} and $J_{max} = 20$) (Press et al. 1989). At $\sigma_f = 0.001215$, the distribution function is approximately zero; and at 0.607500, the distribution function is 0.99999. For each draw, a uniform random number between 0 and 1 is generated and the sampled value of σ_f is set equal to that point among the 500 equispaced points with distribution function value nearest the uniform number.

The *b*th pair of values from a log-log bivariate density was computed as follows:

Step 1. Sample a pseudorandom value, $\ln(\tau_{now,b})^*$, from (1), (2), or (3) below:

(1) Independence prior—draw random $\sigma_{f,b}{}^*$ from the posterior distribution, $p(\sigma_f|D_f)$, at equation (8) using the transformaton method. Next, draw $\ln(r_{now,\ b})^* \sim N(-½\sigma_{f,b}{}^{*2}\ ,\ \sigma_{f,b}{}^{*2})$. Compute $\ln(\tau_{now,b})^* = \ln(z_{now}) - \ln(r_{now,\ b})^*$.

(2) Autocorrelation prior—draw $w_b{}^*$ from the chi-square distribution with 7 degrees of freedom and compute $\sigma_{f,b}{}^{2*} = (7s_f^2)/w_b{}^*$. Next, draw $\boldsymbol{\phi}_b{}^* = (\phi_{0,b}{}^*, \phi_{1,b}{}^*)'$ from the bivariate normal distribution with mean $\hat{\phi} = (\hat{\phi}_0, \hat{\phi}_1)'$ and covariance matrix, $\sigma_{f,b}{}^{2*}(D_\phi{}' D_\phi)^{-1}$. Finally, draw $\varepsilon_b{}^* \sim N(0, \sigma_{f,b}{}^{2*})$. Compute $\ln(r_{now,b})^* = \phi_{0,b}{}^* + \phi_{1,b}{}^* \ln(r_{1996}) + \varepsilon_b{}^*$, then $\ln(\tau_{now,b})^* = \ln(z_{now,b}) - \ln(r_{now,b})^*$.

(3) Trend prior—draw $w_b{}^*$ from the chi-square distribution with 8 degrees of freedom and compute $\sigma_{f,b}{}^{2*} = (8s_f^2)/w_b{}^*$. Next, draw $\boldsymbol{\kappa}_b{}^* = (\kappa_{0,b}{}^*, \kappa_{1,b}{}^*)'$ from the bivariate normal distribution with mean $\hat{\kappa} = (\hat{\kappa}_0, \hat{\kappa}_1)'$ and covariance matrix, $\sigma_{f,b}{}^{2*}(D_\kappa{}' D_\kappa)^{-1}$. Finally draw $v_b{}^* \sim N(0, \sigma_{f,b}{}^{2*})$. Compute $\ln(r_{now,b})^* = \kappa_{0,b}{}^* + 18\kappa_{1,b}{}^* + v_b{}^*$, then $\ln(\tau_{now,b})^* = \ln(z_{now}) - \ln(r_{now,\ b})^*$.

Step 2. Sample a pseudorandom value, $\ln(x_{now,b})^*$ given the draw, $\ln(\tau_{now,b})^*$, at step 1: Draw $w_b{}^*$ from the chi-square distribution with 14 degrees of freedom and compute $\sigma_{\eta,b}{}^{2*} = (14s_\eta{}^2)/w_b{}^*$. Next, draw $\boldsymbol{\beta}_b{}^* = (\beta_{0,b}{}^*, \beta_{1,b}{}^*, \beta_{2,b}{}^*)'$ from the trivariate normal distribution with mean $\hat{\beta} = (\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2)'$ and covariance matrix, $\sigma_{\eta,b}{}^{2*}(D_\beta{}' D_\beta)^{-1}$. Finally, draw $\eta_b{}^* \sim N(0, \sigma_{\eta,b}{}^{2*})$. Compute $\ln(x_{now,b})^* = \beta_{0,b}{}^* + 18\beta_{1,b}{}^* + \beta_{2,b}{}^* \ln(\tau_{now,b})^* + \eta_b{}^*$.

A Monte Carlo Evaluation of the Stock Synthesis Assessment Program

David B. Sampson and Yanshui Yin
Oregon State University, Newport, Oregon

Abstract

Stock assessments for many U.S. Pacific coast groundfish stocks are developed using the catch at age method known as Stock Synthesis. The Stock Synthesis computer program attempts to reconstruct the demographic history of a stock from observed changes in fish age or size distributions, coupled with auxiliary information such as an index of stock biomass developed from a research survey or an index of fishing mortality based on fishing effort. In this study Monte Carlo simulation techniques were used to generate fishery and survey data with known characteristics. The simulated data were then analyzed with the age-structured version of the Stock Synthesis program and results from the program were compared with the true values to evaluate the influence of measurement errors on the accuracy of the Stock Synthesis results. Data sets were constructed with low and high levels of random error in each of four types of sample data (fishery age composition, a fishing effort index, survey age composition, and a survey index of stock biomass). A series of experiments, based on a fractional factorial design, was conducted to examine the importance of eight factors: low versus high rates of natural mortality; constant versus variable annual recruitment; low versus high rates of increase in fishing mortality; dome-shaped versus asymptotic fishery selectivity; short versus long data series; low versus high variability in the fishing effort index; low versus high variability in the survey biomass index; and small versus large samples for age composition. On average the Stock Synthesis estimates for total biomass, exploitable biomass, recruitment, and fishing mortality in the ending year were slightly positively biased (3.5-6.1%) but less variable than the input data. In general, the number of years in the data series and the size of the age samples were the most influential factors, with increased amounts of data producing less biased and less variable estimates.

Introduction

Although fisheries managers and the fishers they regulate generally appreciate that exploited fish stocks have limited productive capabilities and therefore cannot support unconstrained harvests, the fisheries scientists who advise the managers are rarely able to provide highly accurate estimates of the size or potential productivity of any stock. The patchiness and inaccessibility of most fishes, coupled with natural variability and an uncontrolled environment, make it particularly difficult to measure, let alone predict, the stock size and dynamics of wild fish populations. Inaccurate perceptions of stock size and dynamics can be costly, however. Estimates that are too small can result in lost fishing opportunities due to regulations that are too restrictive; estimates that are too liberal can result in overfishing, overcapitalization of the fishing fleet, and decreased long-term productivity of the fish stock.

Fisheries managers rely on fish stock assessments to provide them with a biological foundation for their management decisions. These assessments supply fundamental information about the status of the stocks, whether the stocks are increasing and can support increased levels of harvest, or whether they are decreasing and require stricter control of harvests. Often the stock assessments recommend annual catch quotas to act as upper bounds on the rates of fish removals. Although fisheries scientists are normally aware of the tremendous variability associated with fisheries data, stock assessment documents often do not state explicitly that the estimates of the current stock size can be highly inaccurate, and that recommended catch quotas may consequently fail to meet management objectives.

Determining a stock assessment's level of accuracy is a very involved process. In many modern stock assessments age-structured population models (Megrey 1989) are applied to catch at age data collected from multistage sampling schemes (Gavaris and Gavaris 1983, Quinn et al. 1983). It is a daunting challenge to estimate sampling variances for catch at age data. To trace the consequences of these sampling errors on the final stock assessment estimates is an even more formidable problem. However, there have been some attempts to do so (Pope and Garrod 1975, Mohn 1983, Rivard 1983, Sampson 1987, Prager and MacCall 1988, Kimura 1989, Rivard 1989, Pelletier and Gros 1991, Restrepo et al. 1992). If stock assessments could routinely provide estimates of their accuracy and measures of the relative importance of different sources of uncertainty, then it would be possible to identify which inputs were most in need of improvement. Appropriate modifications to the sampling programs could then be specified.

The stock assessments for many U.S. Pacific Coast groundfish stocks are developed using a catch at age method known as Stock Synthesis (Methot 1990, Pacific Fishery Management Council [PFMC] 1996). This program, which elaborates on the methods of Fournier and Archibald (1982) and

Deriso et al. (1985), attempts to reconstruct the demographic history of a stock from observed changes in fish age distributions, coupled with auxiliary information such as survey indices of stock biomass. Stock Synthesis uses time series of catch biomass and age (or size) composition to derive maximum likelihood estimates of numerical stock abundance, stock biomass, and related parameters. It is a very powerful tool for examining fisheries data; it offers considerable flexibility in the types of data that it can accommodate.

In this study we evaluated the influence of random sampling errors in fishery and survey data on the accuracy of a suite of estimates output by the age-structured version of the Stock Synthesis assessment program. We used Monte Carlo simulation techniques (Rubinstein 1981) to generate fishery data with known characteristics. The random data sets were then analyzed using the Stock Synthesis program and estimates from Stock Synthesis were compared with the true values. There have been few published Monte Carlo evaluations of the Stock Synthesis program.

Stock Synthesis Approach

The Stock Synthesis program uses standard deterministic equations to describe the dynamics of an age-structured population (Methot 1990). The number of fish in a given year class follows an exponential decay function,

$$N_{ya} = N_{y-1,a-1} \times \exp[-(M + S_{a-1} \times F_{y-1})],$$

where N_{ya} denotes the number of fish at the start of year y that are a years old, M is the instantaneous rate of natural mortality, S_a is the selectivity coefficient for age a fish, and F_y is the instantaneous rate of fishing mortality in year y for fully selected ages. For simplicity here, M is constant and the S_a do not vary from year to year. The number of fish in the oldest (terminal, T) age class is given by

$$N_{yT} = N_{y-1,T} \times \exp[-(M + S_T \times F_{y-1})] + N_{y-1,T-1} \times \exp[-(M + S_{T-1} \times F_{y-1})].$$

The Synthesis program hosts a range of methods for modeling the selectivity coefficients. We used the four parameter "double-logistic" function, which produces a selectivity versus age relationship that is sometimes described as being "dome-shaped,"

$$S_a = \frac{\left\{1 + \exp\left[-b1 \times (a - a1)\right]\right\}^{-1} \times \left\{1 + \exp\left[b2 \times (a - a2)\right]\right\}^{-1}}{\mathrm{Max}S},$$

where $a1$ controls the inflection age and $b1$ controls the slope for the ascending portion of the curve, and $a2$ and $b2$ exert similar controls over the descending portion of the curve. Max$S$, which is the maximum value of the numerator for integer values of age up to the terminal age, scales the selectivity coefficients so that at least one age class has a selectivity coefficient of one and thus suffers the full rate of fishing mortality. If parameter $b2$ is zero, the selectivity function degenerates to a so-called "asymptotic" selectivity versus age relationship.

The catch at age (in numbers of fish) is given by the following catch equation,

$$C_{ya} = \frac{N_{ya} S_a F_y \times \left\{1 - \exp\left[-\left(M + S_a F_y\right)\right]\right\}}{M + S_a F_y} \ .$$

The yield at age (catch in weight) is given by

$$Y_{ya} = C_{ya} W_a \ ,$$

where W_a denotes the average weight of age a fish in the fishery, here assumed to be time-independent.

As it is typically configured, the Synthesis program assumes that the data for catch biomass are measured with perfect accuracy and it adjusts the estimates for F_y so that the estimate of total catch biomass,

$$Y_y = \sum_a Y_{ya} \ ,$$

equals the observed total catch biomass.

All methods for analyzing catch at age data require additional auxiliary information for tuning the analysis (Pope and Shepherd 1982, Shepherd and Nicholson 1991). It is otherwise impossible to distinguish a stock that is increasing from one that is decreasing. For tuning a Stock Synthesis analysis one can use survey indices of stock biomass or numerical abundance, or data series for fishing effort or catch per unit effort. If survey biomass data are used and the survey is conducted at the beginning of the year, the expected value of the survey biomass index is given by

$$E\left[B'_y\right] = Q' \sum_a N_{ya} W'_a S'_a \ ,$$

where Q' denotes the survey catchability coefficient, W'_a is the average fish weight at age in the survey, S'_a is the survey selectivity coefficient for age a fish. If fishing effort data (f) are used, the expected value of the effort is related to the rate of fishing mortality by

$$F_y = QE\left[f_y\right] \ ,$$

where Q denotes the fishery catchability coefficient.

The Stock Synthesis program is very flexible in allowing the user to either fix or estimate the various parameters that define the population structure and dynamics. For parameters that are estimated rather than prespecified, the values are maximum likelihood estimates and satisfy a total likelihood function of the form

$$L_{total} = \sum_i L_i e_i \ ,$$

where L_i denotes the log-likelihood value for likelihood component i and e_i is the so-called "emphasis value" for component i. An application with the Synthesis program will include a likelihood component and emphasis value for each type of input data that contains observation error, e.g., the fishery age composition data, the fishing effort data, the survey age composition data, and the survey biomass index. If the entire model has been correctly defined, which includes properly specifying the measurement error associated with each type of observation, the emphasis values should all be equal to one.

By assuming that age determination is exact and that simple random samples of fish are obtained (either from the fishery or the survey), then the age composition data are distributed as multinomial random variables and the likelihood component for these data is

$$L_{age} = \sum_y J_y \times \sum_a \left\{ p_{ya} \times \log\left(E\left[p_{ya}\right]\right) - p_{ya} \times \log\left(p_{ya}\right) \right\} \ ,$$

where J_y is the number of fish in the sample for year y, p_{ya} is the observed proportion at age in the sample for year y, and $E[p_{ya}]$ is the true proportion at age in the sample for year y. If the observed proportions at age are all equal to the expected proportions at age, the L_{age} likelihood component will be zero.

If the survey estimates of biomass are distributed as a lognormally distributed random variable, then the likelihood component for these data is

$$L_{survey} = -\log(\sigma_{Sy}) - \sum_y \frac{1}{2\sigma_{Sy}^2} \times \left(\log \frac{B'_y}{E\left[B'_y\right]} \right)^2 \ ,$$

where B'_y is the observed survey estimate of biomass in year y and $E[B'_y]$ is its expected value, and σ_{Sy} is the true, log-scale standard deviation for these data,

$$\sigma_{Sy} = \sqrt{\log\left(1 + cv_{Sy}^2\right)} \ .$$

The term cv_{Sy} is the true, arithmetic-scale coefficient of variation of the survey biomass estimate in year y.

If the fishing effort data also follow a log-normal distribution, the likelihood component for these data is

$$L_{effort} = -\log(\sigma_F) - \sum_y \frac{1}{2\sigma_F^2} \times \left(\log \frac{f'_y}{E\left[f'_y\right]} \right)^2 ,$$

where f'_y is the observed fishing effort in year y and $E[f'_y]$ is its expected value, and σ_F is the true, log-scale standard deviation for these data,

$$\sigma_F = \sqrt{\log\left(1 + cv_F^2\right)} .$$

The term cv_F is the true, arithmetic-scale coefficient of variation for the fishing effort series.

One can configure the Stock Synthesis program to include likelihood components that tend to constrain the estimates of annual recruitment to conform to a spawner recruit relationship. We specified zero emphasis on the spawner recruit components.

Methods

We used a brute-force approach for evaluating the performance of the Stock Synthesis program. We generated random data sets with known characteristics, analyzed the data sets using the Stock Synthesis program, and then compared estimates from Synthesis with the true values. We used the Windows 95 version of Stock Synthesis, which was provided to us in June 1997 by the program's author (Richard Methot, National Marine Fisheries Service, Northwest Fisheries Science Center, Seattle, WA). We produced the random data sets using two C++ programs developed by Yanshui Yin. The first program simulates the dynamics of an age-structured population using the same deterministic equations (described above) that underlie Methot's Stock Synthesis program. The second program takes the true demographic data produced by the first program and generates random data sets that can be analyzed directly by the Stock Synthesis program.

Stock Synthesis Configuration for this Study

We simulated a simple fishery system: a single fishery with data available annually on total catch, age composition, and nominal fishing effort; and a single survey that provided estimates of annual stock biomass and age composition. In all cases we generated the simulated random data in such a manner that they would be unbiased, and we gave the Synthesis program the true parameter values as the initial values with which to begin its iterative search for the set of maximum likelihood parameter estimates. We conducted a small sensitivity study to examine the influence of using the true parameter values as the initial values. For all simulated stocks we generated a set of errorless data that we analyzed with Stock Synthesis and thereby verified an exact correspondence between the deterministic population equations of the data simulator and those of Stock Synthesis.

The Stock Synthesis program assumes that the data on catch biomass are exact, but in our simulations the catch biomass data were estimates that were lognormally distributed with a 10% relative accuracy (on the arithmetic scale). The age composition data were generated as simple random samples from the true catch and were not roughened to mimic randomness due to age-reading errors, nor were they distorted to mimic age-reading bias. The Stock Synthesis program was configured to treat the age composition data as being without reading error but with multinomial sampling error, and the program was given the true sample size. The fishing effort data for a given simulation were generated as lognormally distributed random variables with expected values equal to the true values (F/Q) and with a fixed coefficient of variation (cv_F) for all years. The Stock Synthesis program was configured to treat the effort data as being lognormally distributed and was given the true log-scale standard deviation for these data (σ_F). The survey estimates of biomass for a given simulation were generated as lognormal random variables with expected values equal to the true values ($Q' \times \sum_a N_{ya} W'_a S'_a$) and with a fixed coefficient of variation (cv_S) for all years. The Stock Synthesis program was configured to treat the survey biomass estimates as being lognormally distributed and was given the true log-scale standard deviation for these data (σ_S).

For each of the types of simulated data that were lognormally distributed (the catch biomass, fishing effort, and survey biomass) the mean and standard deviation on the arithmetic scale ($E[Y]$, $V[Y]$) were related to the mean and standard deviation on the log scale (μ,σ) by the following:

$$E[Y] = \exp\left(\mu + 0.5\sigma^2\right) ,$$

$$V[Y] = \exp\left(2\mu + \sigma^2\right) \times \left[\exp\left(\sigma^2\right) - 1\right] ,$$

i.e., the parameters defining lognormal random data were adjusted so that the generated random data did not have transformation bias.

The fish stock in all the simulations was unfished prior to the start of the simulated period and suffered an instantaneous rate of fishing mortality (F) of 0.07/yr during the first year, with F increasing a fixed amount at the start of each year thereafter. Also, the instantaneous rate of natural mortality (M) was constant with age and through time, and the Stock Synthesis program was configured with M fixed at its correct value. The true fishery catchability coefficient and the survey catchability coefficient were constant throughout each simulated period ($E[Q] = 0.003$, $E[Q'] = 0.1$); the selectivity coefficients for the fishery and survey were also constant throughout each simulated period. There was no sexual dimorphism in the simulated stock. The two sexes had identical growth and identical vulnerability to the fishing and survey gear. The weight at age data were generated from the following deterministic equation,

$$W_a = 0.0001\left\{10\left[1-\exp\left(-0.2a\right)\right]\right\}^3 ,$$

and the Synthesis program was given exact information on weight at age. In our simulations and in the Synthesis configuration weight at age changed abruptly with age at the start of each year.

Experimental Design

Random data sets were generated in accordance with a one-eighth fraction of a 2^8 factorial design (Table 1). For each experimental treatment 200 random data sets were generated and then analyzed with Stock Synthesis. The eight control variables (and their low versus high levels) were: (1) the rate of natural mortality (0.2 versus 0.4/yr); (2) the annual increase in the rate of fishing mortality (0.01 versus 0.02/yr); (3) the number of years in the data series (8 versus 16 years); (4) annual recruitment (constant versus variable); (5) fishery selectivity (domed versus asymptotic); (6) the size of the annual age composition samples (100 versus 400 fish per annual sample); (7) the coefficient of variation of the annual fishing effort data (20% versus 80%, arithmetic scale); and (8) the coefficient of variation of the annual survey biomass estimates (20% versus 80%, arithmetic scale).

Several stock parameters were coupled with the level of the natural mortality coefficient. When M was 0.2/yr, the initial and terminal age classes were 4 and 20 years, the true ascending and descending inflection ages ($a1$ and $a2$) for the fishery selection function were 6 and 16 years, and the true ascending inflection age for the survey selection function was 5 years. When M was 0.4/yr, the initial and terminal age classes were 2 and 10 years, the true ascending and descending inflection ages for the fishery selection function were 4 and 8 years, and the true ascending inflection age for the survey selection function was 3 years. The ascending slope coefficients ($b1$) for the fishery selection function and for the survey selection function were 1.0 and 1.5/yr for all simulated stocks. When the fishery selection function was domed, the descending slope coefficient ($b2$) was 1.0/yr.

For experimental treatments with constant recruitment, the annual recruitment was 3,000 fish (in thousands) and the initial age composition at the start of the first year was at equilibrium with this level of recruitment, and Stock Synthesis was configured to estimate the initial equilibri-

Table 1. Fractional factorial experimental design.

Treatment	FishSel	NatMort	RecVar	FishMort	NumYrs	SmplSize	FishCV	SurvCV
1	–1	–1	–1	–1	–1	–1	–1	1
2	1	–1	–1	–1	–1	1	1	1
3	–1	1	–1	–1	–1	1	1	–1
4	1	1	–1	–1	–1	–1	–1	–1
5	–1	–1	1	–1	–1	1	–1	–1
6	1	–1	1	–1	–1	–1	1	–1
7	–1	1	1	–1	–1	–1	1	1
8	1	1	1	–1	–1	1	–1	1
9	–1	–1	–1	1	–1	–1	1	–1
10	1	–1	–1	1	–1	1	–1	–1
11	–1	1	–1	1	–1	1	–1	1
12	1	1	–1	1	–1	–1	1	1
13	–1	–1	1	1	–1	1	1	1
14	1	–1	1	1	–1	–1	–1	1
15	–1	1	1	1	–1	–1	–1	–1
16	1	1	1	1	–1	1	1	–1
17	–1	–1	–1	–1	1	–1	–1	–1
18	1	–1	–1	–1	1	1	1	–1
19	–1	1	–1	–1	1	1	1	1
20	1	1	–1	–1	1	–1	–1	1
21	–1	–1	1	–1	1	1	–1	1
22	1	–1	1	–1	1	–1	1	1
23	–1	1	1	–1	1	–1	1	–1
24	1	1	1	–1	1	1	–1	–1
25	–1	–1	–1	1	1	–1	1	1
26	1	–1	–1	1	1	1	–1	1
27	–1	1	–1	1	1	1	–1	–1
28	1	1	–1	1	1	–1	1	–1
29	–1	–1	1	1	1	1	1	–1
30	1	–1	1	1	1	–1	–1	–1
31	–1	1	1	1	1	–1	–1	1
32	1	1	1	1	1	1	1	1

Factors:
FishSel = Fishery selectivity (domed vs. asymptotic).
NatMort = Natural mortality (0.2 vs. 0.4/yr).
RecVar = Recruitment variability (constant vs. variable).
FishMort = Fishing mortality trend (0.01 vs. 0.02/yr).
NumYrs = Number of years of data (8 vs. 16 yrs).
SmplSize = Sample size for age compositions (100 vs. 400 fish).
FishCV = Fishing effort variability (20% vs. 80% CV).
SurvCV = Survey biomass variability (20% vs. 80% CV).

um age composition. For treatments with variable recruitment, the average annual recruitment was also 3,000 fish, but the annual recruitment values followed the sequence 3,500, 4,000, 1,200, 4,200, 3,000, 3,200, 1,700, 3,200 during the simulation period (with repetition as necessary) and for the years preceding the simulation, and Stock Synthesis was configured to estimate the initial non-equilibrium age composition.

Because we used a fractional factorial design, not all effects and interactions were separately estimable (Table 2). For example, the main effect for fishery selection was "aliased" with two four-way interactions and with two three-way interactions (natural mortality × recruitment variability × sample size) and (natural mortality × fishing mortality × effort variability), meaning that the value estimated for the fishery selectivity effect included these additional interactions (Box et al. 1978).

For any given data set the Stock Synthesis program produces a wide variety of outputs including ones for the annual series of biomass, fishing mortality, and recruitment. In our analyses we focused on the bias and variability of the Synthesis estimates for the last year for total biomass ($\Sigma N_a W_a$), exploitable biomass ($\Sigma S_a N_a W_a$), rate of fishing mortality, and recruitment, and on estimates for the first year for total biomass. To reduce potential non-normality of these outputs and yet provide replicates for subsequent analyses, for each experimental treatment and output type we separated the 200 sets of estimates from Synthesis into four groups of 50 and then calculated the relative bias and relative variability for each group. We measured relative bias within each group of 50 estimates using the average of

$$\frac{(estimated\ value - true\ value)}{true\ value},$$

and measured relative variability within each group using the coefficient of variation. To summarize results for each experimental treatment and output type we calculated the mean relative bias and mean coefficient of variation by averaging the four replicate measures. For each of the ten output types (the dependent variables) we conducted separate fractional analyses of variance using the Minitab statistics program (release 11.12 for Windows).

Sensitivity to Initial Parameter Values

Likelihood functions can have multiple maxima, in which case the choice of initial values for the parameters can influence whether or not the search algorithm finds a local rather than the global maximum. Also, when the search algorithm follows certain paths it may get stuck in regions where the likelihood surface is very flat. With each of the 200 replicates for each treatment in our main experiment we started the Stock Synthesis program with the true parameter values. To examine the influence of initial parameter values on the performance of Stock Synthesis we generated 50 additional random data sets for two experimental treatments: treatment 6, which had large relative variability in the output estimates; and treatment 30, which had small relative variability. For each of these 100 data sets we ran the Synthesis program 20 times, each time using a different set of randomized initial parameter values, with each parameter varying uniformly within ±50% of its true value.

Results

The five types of Stock Synthesis estimates that we examined varied greatly in relative bias and relative variability among the 32 experimental treatments (Table 3), with the estimates of ending exploitable biomass on average showing the largest negative bias (–4.8%) and the estimates of the ending fishing mortality coefficient showing the largest positive bias (31.6%). The mean relative variability ranged from a low of 3.6% for the estimates of starting biomass to a high of 92% for the estimates of the ending fishing mortality coefficient. For all 32 treatments the estimates of starting biomass were less variable than the other four types of estimates.

The Stock Synthesis estimates of ending biomass, ending exploitable biomass, ending recruitment, and starting biomass were in general skewed to the right, whereas the estimates of the ending fishing mortality coefficient were reasonably symmetric (e.g., Fig. 1). Because the analyses of variance were applied to averages of 50 values, the residuals from the analyses were reasonably well approximated by normal distributions. For the variables that measured relative bias, diagnostic plots of the residual versus fitted values indicated little evidence of heterogeneous variability, but similar plots for the variables that measured relative variation showed some tendency for residual variability to increase with the magnitude of the fitted values.

In the analyses of variance the main effects and two-way interactions were highly significant ($P < 0.01$) for all ten dependent variables (Table 4). The three-way interactions were not significant ($P < 0.05$) for any of the variables, however, except for the relative bias of the starting biomass estimate, which was significant at the $P = 0.05$ level.

Effects on Relative Bias

On average across all levels of the eight factors the five types of estimates that we examined had slight but statistically significant ($P < 0.05$) positive bias, ranging from a low of 1.7% for the estimates of starting biomass to a high of 6.1% for the estimates of the ending fishing mortality coefficient (Table 5). The factor for age composition sample size was the most influential main effect for four of the five types of Synthesis estimates, and the estimated coefficients for this factor were negative for all five types of estimates, indicating that larger samples produced less biased estimates.

Table 2. Alias structure of the fractional factorial design.

Alias Structure (up to order 4)
Grand mean + ABCF + ABDG + CDFG
A + BCF + BDG + CEGH + DEFH
B + ACF + ADG + CDEH + EFGH
C + ABF + DFG + AEGH + BDEH
D + ABG + CFG + AEFH + BCEH
E + ACGH + ADFH + BCDH + BFGH
F + ABC + CDG + ADEH + BEGH
G + ABD + CDF + ACEH + BEFH
H + ACEG + ADEF + BCDE + BEFG
AB + CF + DG
AC + BF + EGH + ADFG + BCDG
AD + BG + EFH + ACFG + BCDF
AE + CGH + DFH + BCEF + BDEG
AF + BC + DEH + ACDG + BDFG
AG + BD + CEH + ACDF + BCFG
AH + CEG + DEF + BCFH + BDGH
BE + CDH + FGH + ACEF + ADEG
BH + CDE + EFG + ACFH + ADGH
CD + FG + BEH + ABCG + ABDF
CE + AGH + BDH + ABEF + DEFG
CG + DF + AEH + ABCD + ABFG
CH + AEG + BDE + ABFH + DFGH
DE + AFH + BCH + ABEG + CEFG
DH + AEF + BCE + ABGH + CFGH
EF + ADH + BGH + ABCE + CDEG
EG + ACH + BFH + ABDE + CDEF
EH + ACG + ADF + BCD + BFG
FH + ADE + BEG + ABCH + CDGH
GH + ACE + BEF + ABDH + CDFH
ABE + CEF + DEG + ACDH + AFGH + BCGH + BDFH
ABH + CFH + DGH + ACDE + AEFG + BCEG + BDEF
ACD + AFG + BCG + BDF + ABEH + CEFH + DEGH

Factor codes:
A = Fishery selectivity.
B = Natural mortality.
C = Recruitment variability.
D = Fishing mortality trend.
E = Number of years of data.
F = Sample size for age compositions.
G = Fishing effort variability.
H = Survey biomass variability.

Table 3. Relative bias and relative variability for the 32 experimental treatments.

	Mean relative bias					Mean relative variability				
Treatment	end Bio	end exB	end *F*	end Rec	start Bio	end Bio	end exB	end *F*	end Rec	start Bio
1	0.1015	0.0853	0.0159	0.1258	0.0557	0.3611	0.4201	0.3220	0.5804	0.2548
2	−0.0426	−0.0436	0.2274	−0.0260	−0.0238	0.4029	0.3756	0.4961	0.4971	0.2019
3	0.0439	0.0468	0.0121	0.0411	0.0338	0.2124	0.2282	0.2986	0.2522	0.1670
4	0.0775	0.0796	0.0493	0.0990	0.0487	0.3731	0.4292	0.3495	0.5258	0.2743
5	−0.0001	−0.0027	0.0463	−0.0050	−0.0037	0.2135	0.2156	0.2592	0.2867	0.1330
6	0.1659	0.1571	0.0321	0.2094	0.0825	0.4972	0.5871	0.4280	0.7682	0.3283
7	−0.0242	−0.0222	0.3160	−0.0023	−0.0225	0.4011	0.4002	0.9220	0.4869	0.2914
8	−0.0401	−0.0483	0.1472	−0.0337	−0.0279	0.2827	0.2615	0.3504	0.3398	0.1447
9	0.0991	0.0959	0.0241	0.1080	0.0585	0.3212	0.3775	0.3486	0.5367	0.2068
10	0.0628	0.0602	0.0019	0.0872	0.0282	0.2699	0.2810	0.2575	0.4018	0.1285
11	0.0197	0.0204	0.0360	0.0257	0.0138	0.2106	0.2194	0.2658	0.2720	0.1294
12	0.0485	0.0380	0.2878	0.0915	0.0166	0.4852	0.4941	0.7910	0.6657	0.2405
13	0.0633	0.0589	0.0136	0.0796	0.0370	0.2773	0.2943	0.3112	0.3949	0.1599
14	0.1414	0.1261	0.0345	0.1725	0.0580	0.4643	0.5370	0.3863	0.6951	0.2556
15	−0.0061	−0.0069	0.1290	0.0274	−0.0011	0.3341	0.3311	0.4285	0.4226	0.2774
16	0.0421	0.0434	0.0292	0.0563	0.0227	0.2832	0.2920	0.2880	0.3514	0.1498
17	0.0141	0.0009	0.0135	0.0279	0.0044	0.1461	0.1490	0.1702	0.3349	0.0743
18	0.0094	0.0108	0.0183	−0.0048	0.0047	0.1744	0.1705	0.2129	0.2506	0.0426
19	0.0236	0.0236	0.0467	0.0271	0.0027	0.2706	0.2763	0.3148	0.3379	0.0948
20	0.0295	0.0255	0.0257	0.0768	0.0050	0.2349	0.2466	0.2371	0.3963	0.1155
21	0.0070	0.0024	0.0098	0.0284	0.0036	0.1557	0.1608	0.1720	0.2409	0.0597
22	0.1275	0.1139	0.1960	0.1720	0.0178	0.5265	0.5542	0.7680	0.8750	0.0936
23	0.0508	0.0623	0.0178	0.0607	0.0732	0.2557	0.2753	0.3092	0.3467	0.2243
24	0.0246	0.0228	0.0034	0.0318	0.0014	0.1381	0.1466	0.1667	0.1896	0.0647
25	−0.0046	−0.0111	0.0884	0.0060	−0.0037	0.2575	0.2635	0.3053	0.4238	0.0591
26	0.0165	0.0149	0.0043	0.0178	0.0044	0.1625	0.1575	0.1698	0.2286	0.0358
27	0.0184	0.0197	−0.0121	0.0275	0.0056	0.1443	0.1534	0.1623	0.1920	0.0754
28	0.0170	0.0269	0.0132	0.0041	0.0127	0.1773	0.2013	0.2502	0.2895	0.0825
29	0.0054	0.0014	0.0159	0.0093	0.0021	0.1314	0.1338	0.1873	0.2078	0.0423
30	0.0287	0.0237	0.0110	0.0690	0.0044	0.1259	0.1260	0.1605	0.3254	0.0463
31	0.0092	0.0218	0.0578	0.0103	0.0357	0.2463	0.2601	0.3038	0.3238	0.1811
32	0.0839	0.0831	0.0252	0.1040	0.0044	0.3068	0.3278	0.3553	0.4042	0.0526
Min	−0.0426	−0.0483	−0.0121	−0.0337	−0.0279	0.1259	0.1260	0.1605	0.1896	0.0358
Max	0.1659	0.1571	0.3160	0.2094	0.0825	0.5265	0.5871	0.9220	0.8750	0.3283

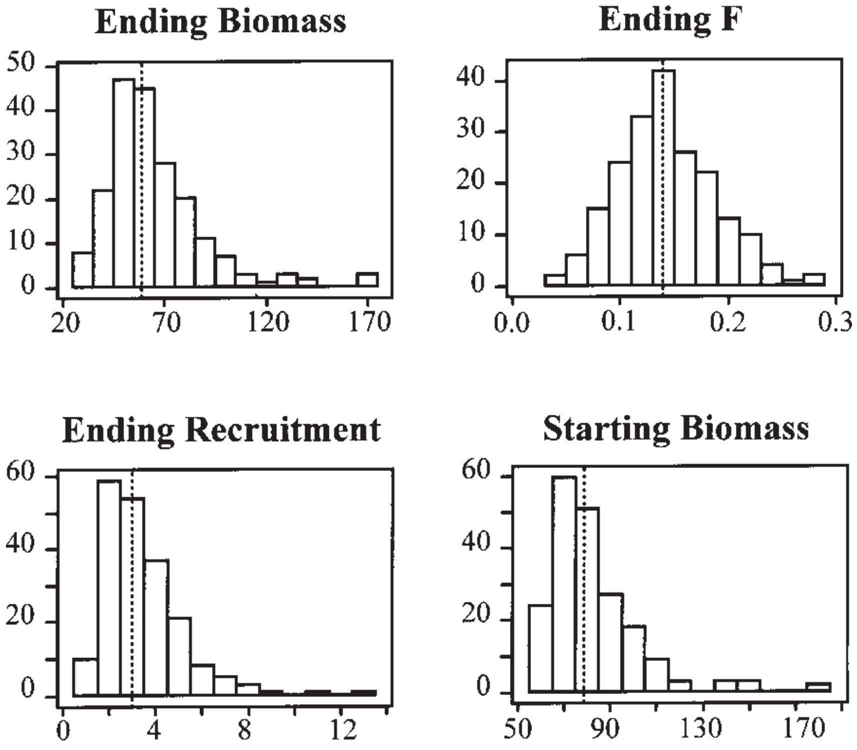


Figure 1. Example histograms from experimental treatment 1 of variables output by the Stock Synthesis program and used as dependent experimental variables. The dashed vertical lines indicate the true values. The units for the biomass and recruitment axes are in thousands.

The number of years in the data series was the only other factor that produced a significant ($P < 0.05$) main effect for all five types of estimates, and in all cases an increase in the number of years produced less biased estimates. For the estimates of ending biomass, ending exploitable biomass, and ending recruitment, however, the interactions between these two factors were significantly ($P < 0.05$) positive, indicating that these two effects were not additive. Recruitment variability was the only factor that did not produce a significant ($P < 0.05$) main effect for any of the estimate types, but there were significant interactions with this factor for several estimate types. For all five estimate types there were significant ($P < 0.05$) interactions between the natural mortality coefficient and the number of years, between the recruitment variability and the fishing effort variability, and between the number of years and the survey biomass variability.

Effects on Relative Variability

The overall average relative variability of the five types of Stock Synthesis estimates ranged from a low of 14.6% for the estimates of starting biomass to a high of 40.1% for the estimates of ending recruitment (Table 6). The factors for the number of years in the data series and for the age composition sample size produced the first or second most influential main effects for all five types of estimates, and the estimated coefficients for these factors were always negative, indicating that larger samples and longer data series produced less variable estimates. The interactions between these two factors were significantly ($P < 0.05$) positive in all cases, however, indicating that these two effects were not additive. The factors for fishing effort variability and survey biomass variability were significant ($P < 0.05$) and had great influence over the variability in the estimates of ending biomass, ending exploitable biomass, ending fishing mortality, and ending recruitment, but not over the estimate of starting biomass. All eight factors produced significant ($P < 0.05$) main effects for at least three of the five types of estimates.

Sensitivity to Initial Parameter Values

For treatment 30, which in the main experiment produced output values with little variability, using randomized initial parameter values had essentially no effect on the final Stock Synthesis output values. For example, across the 50 random data sets the coefficients of variation for the estimates of ending biomass (calculated from 20 replicates for each data set) ranged from 0.06% to 0.26% and the coefficients of variation for the log-likelihood values ranged from less than 0.0001% to 0.14%. For treatment 6, however, which in the main experiment produced output values with large variability, the Synthesis program did not always converge to the same output values when started with randomized parameter values. Although the final log-likelihood values were consistent across the 50 ran-

Table 4. ANOVA tables from fractional factorial experiment.

Source	DF	SS	MS	*F*	*P*
Relative bias in ending total biomass.					
Main effects	8	0.0916	0.0114	7.08	0.000
2-way interactions	20	0.2066	0.0103	6.39	0.000
3-way interactions	3	0.0065	0.0022	1.33	0.269
Residual error	96	0.1552	0.0016		
Total	127	0.4599			
Relative bias in ending exploitable biomass.					
Main effects	8	0.0757	0.0095	5.73	0.000
2-way interactions	20	0.1967	0.0098	5.96	0.000
3-way interactions	3	0.0051	0.0017	1.03	0.383
Residual error	96	0.1584	0.0017		
Total	127	0.4359			
Relative bias in ending *F*.					
Main effects	8	0.4512	0.0564	19.18	0.000
2-way interactions	20	0.4188	0.0209	7.12	0.000
3-way interactions	3	0.0109	0.0036	1.23	0.303
Residual error	96	0.2822	0.0029		
Total	127	1.1631			
Relative bias in ending recruitment.					
Main effects	8	0.1669	0.0209	8.23	0.000
2-way interactions	20	0.2526	0.0126	4.98	0.000
3-way interactions	3	0.0119	0.0040	1.56	0.203
Residual error	96	0.2434	0.0025		
Total	127	0.6748			
Relative bias in starting biomass.					
Main effects	8	0.0269	0.0034	7.37	0.000
2-way interactions	20	0.0603	0.0030	6.62	0.000
3-way interactions	3	0.0045	0.0015	3.29	0.024
Residual error	96	0.0437	0.0005		
Total	127	0.1354			

Table 4. (Continued.)

Source	DF	SS	MS	*F*	*P*
Relative variability in ending total biomass.					
Main effects	8	1.2957	0.1620	59.14	0.000
2-way interactions	20	0.3174	0.0159	5.80	0.000
3-way interactions	3	0.0091	0.0030	1.11	0.350
Residual error	96	0.2629	0.0027		
Total	127	1.8851			
Relative variability in ending exploitable biomass.					
Main effects	8	1.5774	0.1972	47.57	0.000
2-way interactions	20	0.4643	0.0232	5.60	0.000
3-way interactions	3	0.0102	0.0034	0.82	0.485
Residual error	96	0.3979	0.0041		
Total	127	2.4498			
Relative variability in ending *F*.					
Main effects	8	2.8859	0.3607	39.13	0.000
2-way interactions	20	1.2001	0.0600	6.51	0.000
3-way interactions	3	0.0485	0.0162	1.75	0.161
Residual error	96	0.8850	0.0092		
Total	127	5.0195			
Relative variability in ending R.					
Main effects	8	2.8630	0.3579	52.36	0.000
2-way interactions	20	0.6966	0.0348	5.10	0.000
3-way interactions	3	0.0280	0.0093	1.36	0.259
Residual error	96	0.6561	0.0068		
Total	127	4.2437			
Relative variability in starting biomass.					
Main effects	8	0.7888	0.0986	68.17	0.000
2-way interactions	20	0.1335	0.0067	4.62	0.000
3-way interactions	3	0.0010	0.0003	0.22	0.880
Residual error	96	0.1388	0.0014		
Total	127	1.0621			

Table 5. Analysis of relative bias.

Factor	End Bio	End exBio	End *F*	End Rec	Start Bio
Grand mean	0.03792[a]	0.00359[a]	0.06055[a]	0.05388[a]	0.01734[a]
FishSel	0.01163[a]	0.01055[a]	0.00862	0.01655[a]	–0.00111
NatMort	–0.01178[a]	–0.00806[a]	0.01346[a]	–0.01343[a]	–0.00329
RecVar	0.00453	0.00447	0.00726	0.00797	0.00063
FishMort	0.00240	0.00319	–0.01305[a]	0.00212	0.00137
NumYrs	–0.00911[a]	–0.00766[a]	–0.02711[a]	–0.01214[a]	–0.00619[a]
SmplSize	–0.01682[a]	–0.01572[a]	–0.02146[a]	–0.02475[a]	–0.01053[a]
FishCV	0.00639	0.00750[a]	0.02470[a]	0.00461	0.00258
SurvCV	–0.00292	–0.00478	0.03523[a]	0.00084	–0.00629[a]
FishSel×NatMort	–0.00238	–0.00396	–0.01001[a]	–0.00327	–0.00249
FishSel×RecVar	0.01768[a]	0.01487[a]	–0.01659[a]	0.01927[a]	0.00355
FishSel×FishMort	0.00317	0.00297	–0.00521	0.00275	0.00132
FishSel×NumYrs	0.00171	0.00199	–0.00492	0.00055	–0.00319
FishSel×SmplSize	–0.01316[a]	–0.01225[a]	0.00942	–0.01662[a]	–0.00394[a]
FishSel×FishCV	0.00054	0.00032	0.00980[a]	0.00077	–0.00162
FishSel×SurvCV	–0.00104	–0.00239	0.01413[a]	0.00060	–0.00314
NatMort×NumYrs	0.01508[a]	0.01610[a]	–0.02471[a]	0.01448[a]	0.00974[a]
NatMort×SurvCV	–0.00446	–0.00477	0.00855	–0.00386	–0.00428[a]
RecVar×FishMort	0.00113	0.00095	–0.01522[a]	0.00207	0.00106
RecVar×NumYrs	0.00880[a]	0.00929[a]	0.00142	0.01096[a]	0.00603[a]
RecVar×FishCV	0.01549[a]	0.01494[a]	–0.01177[a]	0.01967[a]	0.00659[a]
RecVar×SurvCV	0.00647	0.00695	–0.00301	0.00367	0.00157
FishMort×NumYrs	–0.00941[a]	–0.00831[a]	0.00508	–0.01287[a]	–0.00431[a]
FishMort×SurvCV	0.00982[a]	0.01029[a]	–0.01427[a]	0.00658	0.00836[a]
NumYrs×SmplSize	0.01159[a]	0.01040[a]	0.00197	0.01314[a]	0.00299
NumYrs×FishCV	0.00392	0.00371	–0.00544	0.00094	0.00051
NumYrs×SurvCV	0.01068[a]	0.01139[a]	–0.01193[a]	0.01273[a]	0.00389[a]
SmplSize×SurvCV	–0.00179	–0.00090	–0.01053[a]	–0.00212	0.00126
FishCV×SurvCV	–0.00697	–0.00796[a]	0.02967[a]	–0.00284	–0.01007[a]
FishSel×NatMort×NumYrs	–0.00432	–0.00472	0.00097	–0.00244	–0.00492[a]
FishSel×NatMort×SurvCV	0.00350	0.00264	–0.00906	0.00835	0.00278
FishSel×RecVar×FishMort	–0.00442	–0.00325	–0.00136	–0.00415	–0.00179
Coefficient Std Deviation	0.00355	0.00359	0.00479	0.00445	0.00189

Factor coefficients (relative bias) from fractional factorial experiment.

[a] Coefficients with t-statistics significant at the $P = 0.05$ level.

Table 6. Analysis of relative variability.

Factor	End Bio	End exBio	End *F*	End Rec	Start Bio
Grand mean	0.27637[a]	0.29207[a]	0.33586[a]	0.40137[a]	0.14649[a]
FishSel	0.03020[a]	0.03216[a]	0.01834[a]	0.04887[a]	–0.00543
NatMort	–0.00410	–0.00813	0.02620[a]	–0.03911[a]	0.01385[a]
RecVar	0.01362[a]	0.01437[a]	0.02638[a]	0.01480[a]	0.01004[a]
FishMort	–0.01401[a]	–0.01398[a]	–0.02513[a]	–0.01792[a]	–0.01380[a]
NumYrs	–0.06049[a]	–0.06691[a]	–0.07054[a]	–0.06595[a]	–0.06246[a]
SmplSize	–0.04911[a]	–0.06119[a]	–0.06913[a]	–0.09841[a]	–0.04137[a]
FishCV	0.03492[a]	0.03615[a]	0.07578[a]	0.04165[a]	0.00585
SurvCV	0.03901[a]	0.03597[a]	0.06854[a]	0.04626[a]	0.00165
FishSel×NatMort	–0.01730[a]	–0.01622[a]	–0.03185[a]	–0.01585[a]	–0.01434[a]
FishSel×RecVar	0.00791	0.01538[a]	–0.01771[a]	0.02852[a]	–0.00918[a]
FishSel×FishMort	–0.00817	–0.00820	0.00329	–0.01212	–0.00333
FishSel×NumYrs	–0.01526[a]	–0.01601[a]	0.00637	–0.01442	–0.01192[a]
FishSel×SmplSize	–0.00489	–0.01150[a]	0.00200	–0.01895[a]	0.00286
FishSel×FishCV	0.01522[a]	0.01494[a]	0.01871[a]	0.02079[a]	0.00205
FishSel×SurvCV	0.01266[a]	0.00905	0.02149[a]	0.01619[a]	–0.00020
NatMort×NumYrs	0.00996[a]	0.01889[a]	–0.02911[a]	0.01367	0.01349[a]
NatMort×SurvCV	–0.00651	–0.00917	0.01191	–0.00521	–0.00575
RecVar×FishMort	–0.00482	–0.00474	–0.03449[a]	–0.00760	0.00289
RecVar×NumYrs	0.00631	0.00854	0.01112	0.01393	0.00150
RecVar×FishCV	0.00998[a]	0.01548[a]	0.00804	0.02154[a]	0.00539
RecVar×SurvCV	0.00359	0.00702	0.01527	0.00762	–0.00337
FishMort×NumYrs	–0.00787	–0.00827	–0.00336	–0.01812[a]	0.00166
FishMort×SurvCV	–0.00007	0.00511	–0.01817[a]	–0.00370	0.00491
NumYrs×SmplSize	0.01870[a]	0.02685[a]	0.02144[a]	0.01943[a]	0.01582[a]
NumYrs×FishCV	0.01172[a]	0.01401[a]	–0.00326	0.01483[a]	–0.00341
NumYrs×SurvCV	0.01522[a]	0.01971[a]	–0.00564	0.02210[a]	0.00084
SmplSize×SurvCV	–0.00764	–0.00771	–0.03087[a]	–0.00980	0.00306
FishCV×SurvCV	0.01569[a]	0.00904	0.05275[a]	0.02138[a]	–0.00477
FishSel×NatMort×NumYrs	–0.00509	–0.00528	–0.00296	–0.00871	–0.00086
FishSel×NatMort×SurvCV	–0.00292	–0.00325	–0.01699[a]	–0.00104	0.00204
FishSel×RecVar×FishMort	–0.00605	–0.00644	–0.00901	–0.01190	–0.00164
Coefficient Std Deviation	0.00463	0.00569	0.00849	0.00731	0.00336

Factor coefficients (relative variability) from fractional factorial experiment.

[a] Coefficients with *t*-statistics significant at the $P = 0.05$ level.

dom data sets, with coefficients of variation ranging from 0.0013% to 0.28%, the estimates of ending biomass were sometimes quite variable, with coefficients of variation ranging from 0.71% to 27%. Scatterplots of the ending biomass estimates versus the log-likelihood values (e.g., Fig. 2) indicated that the likelihood functions for these data sets were flat and did not have well-defined maxima; the Synthesis program sometimes stopped prematurely. However, we saw no evidence of likelihood functions having multiple maxima.

Also, when the Synthesis program started with the true parameter values for the 50 data sets from treatment 30, it produced final ending biomass estimates that were within –0.49% to 0.52% of the estimates associated with the observed maximum likelihood values; for 44 of the data sets the final ending biomass estimates were within the observed ranges of estimates produced by starting with random initial parameter values (20 replicates per data set). When the Synthesis program started with the true parameter values for the 50 data sets from treatment 6, it produced final ending biomass estimates that were within –25% to 8.7% of the observed maximum likelihood estimates; for 44 of the 50 data sets the final ending biomass estimates were within the observed ranges produced by starting with random initial parameter values. Hence, it appears that the results in the main experiment would not have differed substantially had we used a different strategy for choosing sets of initial parameter values.

Discussion

Results from our experiments were generally in accord with what we anticipated, but there were some surprises. The magnitude of the trends in the fishing mortality coefficient had no significant ($P < 0.05$) effect on the bias of the Synthesis estimates and relatively small influence on their variability. This seems to contradict the well known fact from Virtual Population Analysis that estimates of stock size converge more rapidly to their true values when there are high rates of fishing mortality (Pope 1972). Also, we found that bias and variability in the estimates of ending biomass, ending exploitable biomass, ending fishing mortality, and ending recruitment were larger when the fishery selectivity curve was asymptotic rather than domed. In contrast, Bence et al. (1993) in a Monte Carlo investigation of the Stock Synthesis program found that biomass estimates were more accurate if derived using data from a survey that had an asymptotic rather than domed selectivity curve. In their simulated populations the fishery selection curve was always domed. In our study the survey selection curve was always asymptotic.

The study reported here is only the first in a series of experiments. We intend to conduct additional experiments that will examine the influence of other forms of input data error including random variation in discards and variability in estimates of average weight at age. Also, there usually is

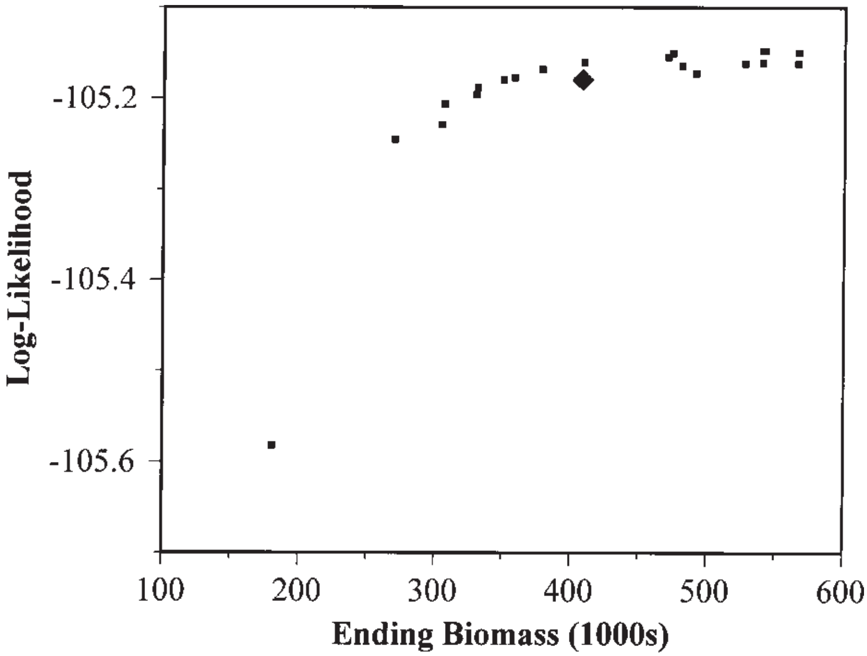


Figure 2. Stock Synthesis estimates can be sensitive to the choice of initial parameter values. Twenty sets of randomized initial parameter values here produced highly variable (CV = 27%) estimates of ending biomass (shown as small squares) when applied to one particular random data set generated using experimental treatment 6. Several other random data sets produced similar, but less extreme, patterns. In the main experiment the Stock Synthesis program always started with the true parameter values. The large diamond shows the biomass and likelihood values obtained by starting with the true parameter values. The true ending biomass for this experimental treament was 54,500.

considerable sample to sample variation in age composition data (Crone 1995, Smith and Maguire 1983) and it is unrealistic to treat age composition data as simple multinomial random variables. We intend to explore the influence on Stock Synthesis estimates of a compound multinomial error structure for age composition data and evaluate the suggestion (Fournier and Archibald 1982) that age sample sizes in the likelihood specification be limited to 400 fish per sample.

Besides data errors, there are other potentially important sources of uncertainty in assessments based on the Stock Synthesis program. For example, structural errors, such as an incorrect assumption that fishery selectivity has been constant from year to year, can produce seriously inaccurate estimates of stock size (Sampson 1993). Other important structural errors that we want to investigate include how measurement errors are specified (e.g., normal versus lognormal) and whether or not mortality effects and recruitment include process error in addition to measurement error.

Stock Synthesis estimates of ending exploitable biomass form the basis for the annual catch quotas for many groundfish stocks on the U.S. Pacific coast (PFMC 1996). With respect to variability in the input data, the results of this study suggest that Synthesis estimates of exploitable biomass are only slightly biased and are relatively less variable than the input data. The ANOVA model for ending exploitable biomass predicts that the most variable estimates of ending exploitable biomass will occur for a treatment with asymptotic fishery selectivity, a low natural mortality coefficient (0.2/yr), variable recruitment, a low trend in the fishing mortality coefficient (0.01/yr), a short data series (8 years), small fishery and survey age composition samples (100 fish per annual sample), and high variability (80%) in the fishing effort and survey biomass indices. For the particular set of parameter values that we examined for this worst-case scenario the ANOVA model predicts a 70.0% coefficient of variation in the estimate of ending exploitable biomass, and it predicts that a four-fold increase in the size of the annual age composition samples (to 400 fish per sample) would reduce the relative variability to 48.5%. In contrast, a four-fold decrease in the relative variability of the survey biomass index would only reduce the relative variability to 58.7%. This implies that increased port sampling to obtain more age composition data would produce greater improvements in assessment precision than increased research surveys to obtain more precise biomass indices. Furthermore, it would very likely be less expensive to collect the additional age composition data. For example, a four-fold improvement in the precision of a survey biomass estimate would require roughly a sixteen-fold increase in sampling because, under simple random sampling, the standard error of a biomass estimate is inversely proportional to the square root of the number of samples.

Acknowledgments

This research was partially supported by grant no. NA36RG0451 (Project no. R/OPF-50) from the National Oceanic and Atmospheric Administration to the Oregon Sea Grant College Program and by appropriations made by the Oregon State legislature. The views expressed herein are those of the authors and do not necessarily reflect the views of NOAA or any of its subagencies.

References

Bence, J.R., A. Gordoa, and J.E. Hightower. 1993. Influence of age-selective surveys on the reliability of stock synthesis assessments. Can. J. Fish. Aquat. Sci. 50:827-840.

Box, G.E.P., W.G. Hunter, and J.S. Hunter. 1978. Statistics for experimenters: An introduction to design, data analysis, and model building. Wiley and Sons, New York. 653 pp.

Crone, P.R. 1995. Sampling design and statistical considerations for the commercial groundfish fishery of Oregon. Can. J. Fish. Aquat. Sci. 52:716-732.

Deriso, R.B., T.J. Quinn II, and P.R. Neal. 1985. Catch-age analysis with auxiliary information. Can. J. Fish. Aquat. Sci. 42:815-824.

Fournier, D., and C.P. Archibald. 1982. A general theory for analyzing catch at age data. Can. J. Fish. Aquat. Sci. 39:1195-1207.

Gavaris, S., and C.A. Gavaris. 1983. Estimation of catch at age and its variance for groundfish stocks in the Newfoundland region. Can. Spec. Publ. Fish. Aquat. Sci. 66:178-182.

Kimura, D.K. 1989. Variability in estimating catch-in-numbers-at-age and its impact on cohort analysis. Can. Spec. Publ. Fish. Aquat. Sci. 108:57-66.

Megrey, B.A. 1989. Review and comparison of age-structured stock assessment models from theoretical and applied points of view. Amer. Fish. Soc. Symp. 6:8-48.

Methot, R.D. 1990. Synthesis model: An adaptable framework for analysis of diverse stock assessment data. Int. North Pac. Fish. Comm. Bull. 50:259-277.

Mohn, R.K. 1983. Effects of error in catch and effort data on tuning Cohort Analysis, with a postscript on logistic production models. Can. Spec. Publ. Fish. Aquat. Sci. 66:141-150.

Pacific Fishery Management Council. 1996. Status of the Pacific Coast groundfish fishery through 1996 and recommended acceptable biological catches for 1997. Pacific Fishery Management Council, Portland, OR.

Pelletier, D., and P. Gros. 1991. Assessing the impact of sampling error on model-based management advice: Comparison of equilibrium yield per recruit variance estimators. Can. J. Fish. Aquat. Sci. 48:2129-2139.

Pope, J.G. 1972. An investigation of the accuracy of Virtual Population Analysis using cohort analysis. Int. Comm. Northwest Atl. Fish. Res. Bull. 9:65-74.

Pope, J.G., and D.J. Garrod. 1975. Sources of error in catch and effort quota regulations with particular reference to variations in the catchability coefficient. Int. Comm. Northwest Atl. Fish. Res. Bull. 11:17-30.

Pope, J.G., and J.G. Shepherd. 1982. A simple method for the consistent interpretation of catch-at-age data. J. Cons. Cons. Int. Explor. Mer 40:176-184.

Prager, M.H., and A. MacCall. 1988. Sensitivities and variances of Virtual Population Analysis as applied to the mackerel, *Scomber japonicus*. Can. J. Fish. Aquat. Sci. 45:539-547.

Quinn II, T.J., E.A. Best, L. Bijsterveld, and I.R. McGregor. 1983. Sampling Pacific halibut *Hippoglossus stenolepis* landings for age-composition: History, evaluation, and estimation. Int. Pac. Halibut Comm. Sci. Rep. 68. 56 pp.

Restrepo, V.R., J.M. Hoenig, J.E. Powers, J.W. Baird, and S.C. Turner. 1992. A simple simulation approach to risk and cost analysis, with applications to swordfish and cod fisheries. Fish. Bull. U.S. 90:736-748.

Rivard, D. 1983. Effects of systematic, analytical, and sampling errors on catch estimates: A sensitivity analysis. Can. Spec. Publ. Fish. Aquat. Sci. 66:114-129.

Rivard, D. 1989. Overview of the systematic, structural, and sampling errors in cohort analysis. Amer. Fish. Soc. Symp. 6:49-65.

Rubinstein, R.Y. 1981. Simulation and the Monte Carlo Method. John Wiley and Sons, New York. 278 pp.

Sampson, D.B. 1987. Variance estimators for Virtual Population Analysis. J. Cons. Cons. Int. Explor. Mer 43:149-158.

Sampson, D.B. 1993. The assumption of constant selectivity and the stock assessment for widow rockfish. Fish. Bull. U.S. 91:676-689.

Shepherd, J.G., and M.D. Nicholson. 1991. Multiplicative modelling of catch-at-age data, and its application to catch forecasts. J. Cons. Cons. Int. Explor. Mer 47:284-294.

Smith, S.J., and J.J. Maguire. 1983. Estimating the variance of length composition samples. Can. Spec. Publ. Fish. Aquat. Sci. 66:165-170.

Alaska Sea Grant College Program • AK-SG-98-01, 1998

Fuzzy Regression in Fisheries Science: Some Methods and Applications

Saul B. Saila
University of Rhode Island, Narragansett, Rhode Island

Scott Ferson
Applied Biomathematics, Setauket, New York

Abstract

Methodology for fuzzy regression is described and then illustrated with some applications to published data from the fisheries literature. Fuzzy regression should be a useful alternative or complement to conventional statistical regression whenever the relationship between variables is imprecise, data are imprecise, or sample sizes are very small. Formulations of fuzzy regression models are described, and goodness-of-fit criteria are briefly investigated and compared. The fuzzy regression technique is first illustrated by applications to data from the relationship between the proportion of chinook salmon jacks entering a river and the proportion of jacks in the ensuing cohort. A second example showing the relation between chinook salmon yearlings and water flow is also investigated. The third example deals with the relationship between an abundance index and virtual population analysis results for a short-time series of Atlantic mackerel data. Fuzzy regression methodology seems relatively straightforward, but it usually involves a constrained minimization problem, which may require mathematical programming methods. On the basis of this review and applications, we believe that fuzzy regression has utility in some fisheries-related applications, but further evaluation is suggested.

Introduction

Regression methods have been widely used and critiqued for some time in fishery science related applications, as is evident from a review by Ricker (1973). Sparre and Venema (1992) describe 18 equations from methods used in fishery science based on transformation and ordinary least square regression models of the form: $y = a + bx$. Prairie (1996) indicated that nearly 45% of the articles in the 1990 volume of the *Canadian Journal of Fisheries and Aquatic Sciences* contained at least one regression analysis. Clearly, regression models are ubiquitous in fishery science today. Ryan (1997) provides an excellent description of recent developments, as well as background in regression methodologies.

In general, a necessary characteristic of problems suitable for statistical regression analysis is that sufficient data should be available for drawing a valid statistical relation between the dependent and independent variable. Another important characteristic of a linear regression model in practical applications is the appropriateness of the statistical model, including the linearity assumption about the underlying functional relationship but also the assumptions of error-free x-values, homescedasticity, and independence. Fuzzy linear regression was first described by Tanaka et al. (1982), and it provides a means for addressing problems which fail to satisfy one or both of these characteristics. Relationships based on relatively few experimental data are common in the fisheries literature. For example, sparse data and an imprecise relationship between variables are illustrated by Fig. 1, redrawn from Cada et al. (1997, Fig. 2, page 136). Unwin and Glova (1997) contained several similar Fig.s. In many instances, scientists have used the relationship based on experience, judgment, or theoretical grounds, in spite of the fact that the statistical regression analysis alone may not lead to a strong conclusion.

Statistical regression analysis provides, under certain assumptions, a means to analyze the dependence among variables using statistical properties of the relationship. Consider a linear regression with one independent variable x and one dependent variable y for which the assumed form $y = a + bx$ represents a straight line. Given a set of observed data (x_1, y_1), (x_2, y_2), ... , (x_n, y_n) for the pair of variables (x, y), familiar ordinary least squares (OLS) techniques are used to find values of a and b for which the total error of the estimated points on the straight line with respect to the observed points is minimized.

In the case of large data sets, statistical techniques can be effectively used to assess measurement or modeling error. However, in many practical fisheries applications, it often occurs that there are so few data available that statistical regression analysis cannot or should not be used. In such cases, the fishery scientist is faced with two alternatives. One is to disregard the relationship even when it is believed that it really exists. For example, most fishery scientists would agree that Pacific salmon smolt survival and travel time to the sea are related in spite of sparse data and ambiguous evidence about the relation between variables. The other alternative is to construct a regression line and calculate indices, such as the coefficient of determination (r^2) which roughly evaluates the fit, or to show confidence intervals. This latter approach was used in Fig. 1.

The techniques of fuzzy regression may provide another alternative which has not been explored in a fisheries context to date and which may

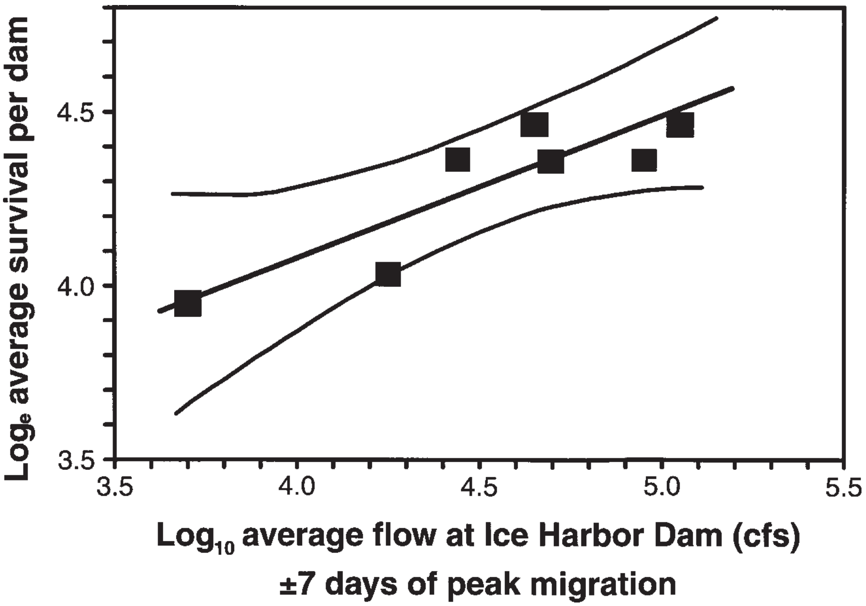


Figure 1. Ordinary least squares regression of river flow at Ice Harbor Dam versus yearling chinook salmon survival. Redrawn from Cada et al. (1997).

contribute to regression problems lacking an abundance of data or having a vague relationship between the variables. Briefly, fuzzy regression is a method of calibrating a fuzzy numerical relationship embodied in an equation involving fuzzy numbers (Kaufmann and Gupta 1985). The regression parameters may be fuzzy numbers and so describe the degree of acceptance of values for a parameter. As shown by Bárdossy et al. (1990), if the regression parameters are fuzzy numbers, the dependent variable is also a fuzzy number. The objective of fuzzy regression is to minimize some measure of "vagueness" of the dependent variable. Bárdossy et al. (1990) provided a good introduction to fuzzy regression in the context of hydrology using an example describing the relationship between soil electrical resistance and hydraulic permeability. Our intention is to provide a similar introduction to fishery scientists using fishery-related examples.

Johnson and Ayyub (1996) used fuzzy regression in a civil engineering application to model uncertainty in the prediction of bridge pier scour. They illustrate the use of fuzzy factors describing the bias between observed field data and estimates derived from laboratory or preliminary field data. This bias exists because of the use of small-scale experiments to model large-scale and real-world problems. The use of small-scale experiments to describe large-scale phenomena is also a problem in fisheries applications such as projecting the results of small-scale net behavior from flume studies to real-world trawl behavior. The above-mentioned study may be helpful in resolving similar problems in fishing gear technology research.

In summary, there seem to be at least two major motivations for applying fuzzy regression. The first one results from questions concerning the validity of a linear model for the given variables. The second motivation results from the nature of the data, which in some applications are inherently vague, very sparse, or both. These two motivations lead to two types of fuzzy regression analyses. One involves conventional crisp data and fuzzy parameters, and the other involves crisp parameters and fuzzy data, where the term "crisp" denotes ordinary, single values from the real number line. Fuzzy numbers possess a natural capability to express observation and measurement uncertainties that traditional crisp numbers do not have. In the material below, we will employ the approach of Klir and Yuan (1995). The next section describes an analysis under the first motivation—namely, the validity of the regression form used.

Linear Regression with Fuzzy Parameters (LRFP)

In this case, the dependence of an output variable Y on input variables is expressed as:

$$Y = C_1x_1 + C_2x_2 + \dots C_mx_m, \tag{1}$$

where $C_1, C_2, \dots, C_m$ are fuzzy numbers, and $x_1, x_2, \dots, x_m$ are crisp (real-valued) input variables. For each m-tuple of input variable values, the output variable defined by Y in equation (1) is a fuzzy number. Other definitions related to fuzzy regression are defined in Appendix A of Bárdossy et al. (1990). The objective of the regression analysis is to find fuzzy parameters $C_1, C_2, \dots, C_m$ for which equation (1) provides the best agreement with the data, according to a goodness-of-fit criterion, such as minimizing the sum of the widths of the fuzzy regression coefficients.

Assume the parameters in equation (1) are symmetric triangular fuzzy numbers defined by

$$C_j(c) = \begin{cases} 1 - \dfrac{|c - c_j|}{s_j} & \text{when } c_j - s_j \le c \le c_j + s_j \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

where c_j is the point for which $C_j(c_j) = 1$, and $s_j > 0$ is the spread of C_j. Each C_j can therefore be denoted by (c_j, s_j) for $j = 1, 2, \dots, m$. It can then be proved by the extension principle (Zadeh 1965) that Y in equation (1) is also a symmetric fuzzy number, which is given by

$$Y(y) = \begin{cases} 1 - \dfrac{|y - \boldsymbol{x}^{\mathrm{T}}\boldsymbol{c}|}{\boldsymbol{s}^{\mathrm{T}}|\boldsymbol{x}|} & \text{when } \boldsymbol{x} \neq 0 \\ 1 & \text{when } \boldsymbol{x} = 0, y \neq 0 \\ 0 & \text{when } \boldsymbol{x} = 0, y = 0 \end{cases} \tag{3}$$

for all $y \in \Re$, where T denotes the transposition and

$$\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ . \\ . \\ . \\ x_m \end{pmatrix}, \ \boldsymbol{c} = \begin{pmatrix} c_1 \\ c_2 \\ . \\ . \\ . \\ c_m \end{pmatrix}, \ \boldsymbol{s} = \begin{pmatrix} s_1 \\ s_2 \\ . \\ . \\ . \\ s_m \end{pmatrix}, \ |\boldsymbol{x}| = \begin{pmatrix} |x_1| \\ |x_2| \\ . \\ . \\ . \\ |x_m| \end{pmatrix}$$

Finding the fuzzy parameters C_1, C_2, ..., C_m can be converted to the problem of finding the vectors **c** and **s** such that the $Y(y)$ functions given by equation (3) fit the data as well as possible. Each data point (x_i, y_i) consists of a vector of observed values of the independent variables x_i, and the corresponding observed value of the dependent variable y_i. Let $Y_i(y)$ denote the fuzzy number defined by equation (3) for the ith vector of independent variables. The fuzzy parameters are fitted so that they are as narrow as possible while permitting the values of $Y_i(y)$ to be large for the values y_i represented in the observed data. In particular, for each vector of independent variables x_i, the associated y_i should belong to the corresponding fuzzy number Y_i with a grade that is greater than or equal to some given value $h \in [0, 1]$. That is, $Y_i(y_i) \geq h$ for each $i = 1,2, ..., n$. The linear regression with fuzzy parameters problem can be expressed in terms of a classical linear programming problem illustrated by Klir and Yuan (1995, page 456) as

$$\text{Minimize } \sum_{j=1}^{m} s_j ,$$

$$\text{Subject to} \tag{4}$$

$$(1 - h)\mathbf{s}^{\mathrm{T}} | \mathbf{x}_i | - | y_i - \mathbf{x}_i^{\mathrm{T}}\mathbf{c} | \geq 0, \qquad i = 1,2,\ldots,n$$

$$s_j \geq 0, \ j = 1,2,\ldots,m.$$

The minimization tightens the spread of the fuzzy parameters so that the model is as specific as possible. Without this minimization, it would always be possible to get $Y(y)$ to be larger than h for any value of the dependent variable, simply by widening the spread of the fuzzy parameters. The positivity constraint insures that the fitted parameters are feasible fuzzy numbers.

An example of linear regression with fuzzy parameters is provided using the data in Table 1, which we extracted from Unwin and Glova (1997, Fig. 3). We assume the form $Y = Cx$, where $C = (c, s)$ is a fuzzy parameter expressed by a symmetric fuzzy number. This linear programming problem has the form

$$\text{Minimize } s,$$

$$\text{Subject to}$$

$$8(1-h)s - |24 - 8c| \geq 0,$$

$$9(1-h)s - |8 - 9c| \geq 0,$$

$$\vdots$$

$$47(1-h)s - |23 - 47c| \geq 0,$$

$$s \geq 0, \text{ and}$$

$$h \in [0,1] \text{ is a fixed number.}$$

Some justification for the use of this form is provided by comparing of the results from fitting an ordinary least squares regression with intercept versus a regression through the origin. It was evident that an improvement in the r^2 value occurred for the latter case. Of course, as Ryan (1997) notes, r^2 values for these two forms of linear regression are not strictly comparable, and, in any case, Prairie (1996) pointed out that r^2 does not necessarily provide an intuitive measure of the predictive power of regressions. Nevertheless, the difference was large, and it also seemed intuitively evident that $Y = 0$ when $x = 0$, since no returns are expected if there are no spawners.

We used the optimization modeling software LINGO™, which includes simultaneous linear and nonlinear solver-optimizers, to make the computations for this example. The mathematical programming problem is linear in this case. For the data from the Table 1 and the grade of membership set at 0.5, we found $C = (0.872, 1.32)$.

Fuzzy numbers $Y = Cx$ can be calculated for various values of x. In this case, C is a fuzzy parameter expressed by a symmetric triangular fuzzy

Table 1. Raw data on chinook salmon spawners versus returns extracted from Unwin and Glova (1997) with predictions based on ordinary least squares regression (OLS), linear regression with fuzzy parameters (LRFP), and linear regression with fuzzy data (LRFD).

	Raw data		OLS		LRFP		LRFD	
Case	Spawners *X*	Returns *Y*	Predicted returns	Residuals	Predicted returns	Residuals	Predicted returns	Residuals
1	8	24	7.03	16.97	6.98	17.02	6.91	17.09
2	9	8	7.90	0.10	7.85	0.15	7.78	0.22
3	11	15	9.66	5.34	9.59	5.41	9.50	5.50
4	11	16	9.66	6.34	9.59	6.41	9.50	6.50
5	15	18	13.17	4.83	13.02	4.92	12.96	5.04
6	22	19	19.32	–0.32	19.18	–0.18	19.01	–0.01
7	25	33	21.96	11.04	21.80	11.20	23.60	5.40
8	26	32	22.84	9.17	22.67	9.33	22.46	9.54
9	30	26	26.35	–0.35	26.16	–0.16	25.96	0.08
10	33	24	28.98	–4.98	26.78	–4.78	28.51	–4.51
11	29	52	25.47	26.53	25.29	26.71	25.51	26.94
12	39	28	34.25	–6.25	34.01	–6.01	33.70	–5.70
13	44	31	38.64	–7.64	38.37	–7.37	38.02	–7.02
14	47	23	41.28	–18.28	40.98	–17.98	40.60	–17.60

For the fuzzy methods, the tabled values for predicted returns are the values where the fuzzy numbers produced by the regression equations have their peaks. This is analogous to using the means of the predictions from OLS (which are also distributions). Also shown are the residuals for each analysis.

number. Although mathematical operations for fuzzy arithmetic are relatively straightforward, they can become tedious. Risk Calc (Ferson et al. 1998) is a convenient environment for evaluating fuzzy mathematical expressions. These can include scalars representing known or mathematically defined integers and real numbers, intervals representing uncertain numbers about which bounds can be established, and fuzzy numbers representing uncertain numbers for which, in addition to having a range of possible values, one can say that some are more plausible than others. All these kinds of numbers can be mixed and used together in mathematical expressions. In our case, a real-valued independent variable is multiplied by a symmetric triangular fuzzy number (the slope of the regression) to obtain fuzzy predicted values of the dependent variable. The results are summarized numerically in Table 1 and also illustrated in Fig. 2. The figure depicts the original data points, the regression line and, in their own graph, the predicted values $Y_i = Cx_i$. Each of these predictions is a symmetric triangular fuzzy number whose peak is located at the best estimate and whose breadth indicates the reliability of the estimate.

Linear Regression with Fuzzy Data (LRFD)

In this type of regression, the dependence of an output variable Y on input variables X_i is expressed by the following:

$$Y = a_1 X_1 + a_2 X_2 + \ldots + a_m X_m \tag{5}$$

where the values of the input and output variables are fuzzy numbers which are assumed to be triangular and symmetric, and a_1, a_2, ..., a_m are real-valued parameters. Let $X_j = (x_j, s_j)$ for $j = 1,2, ..., m$. Then

$$Y(y) = \begin{cases} 1 - \dfrac{|y - \mathbf{a}^{\mathrm{T}}\mathbf{x}|}{\mathbf{s}^{\mathrm{T}}|\mathbf{a}|} & \text{when } \mathbf{a} \neq 0 \\ 1 & \text{when } \mathbf{a} = 0, y \neq 0 \\ 0 & \text{when } \mathbf{a} = 0, y = 0 \end{cases}$$

for all $y \in \Re$, where

$$\boldsymbol{a} = \begin{pmatrix} a_1 \\ a_2 \\ . \\ . \\ . \\ a_m \end{pmatrix}, \boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ . \\ . \\ . \\ x_m \end{pmatrix}, \boldsymbol{s} = \begin{pmatrix} s_1 \\ s_2 \\ . \\ . \\ . \\ s_m \end{pmatrix}.$$

Data are provided as pairs (X_i, Y_i) for $i=1, 2, ..., n$, where each X_i is an m-tuple of symmetric triangular fuzzy numbers, and Y_i is a symmetric triangular fuzzy number. The pair of real numbers (x_{ij}, s_{ij}) will specify the fuzzy number for the jth independent variable of the ith data point. The pair of real numbers (y_i, s_i) will specify the fuzzy number for the corresponding dependent variable for the ith data point. Let Y_i denote the observed fuzzy number associated with X_i, and let $\hat{Y}_i$ denote the predicted fuzzy number obtained from fuzzy linear function (5) when X_i is used as its input. The objective of the regression problem is to find parameters a_1, a_2, ..., a_m such that the fit between the fuzzy prediction $\hat{Y}_i$ and the fuzzy data Y_i is as good as possible. According to Klir and Yuan (1995), two simultaneous criteria are usually used for goodness of fit. First, minimize the total differences between the areas of the fuzzy numbers $\hat{Y}_i$ and Y_j for $i = 1, ..., n$. Second, make the fuzzy numbers $\hat{Y}_i$ and Y_j compatible to some given de-

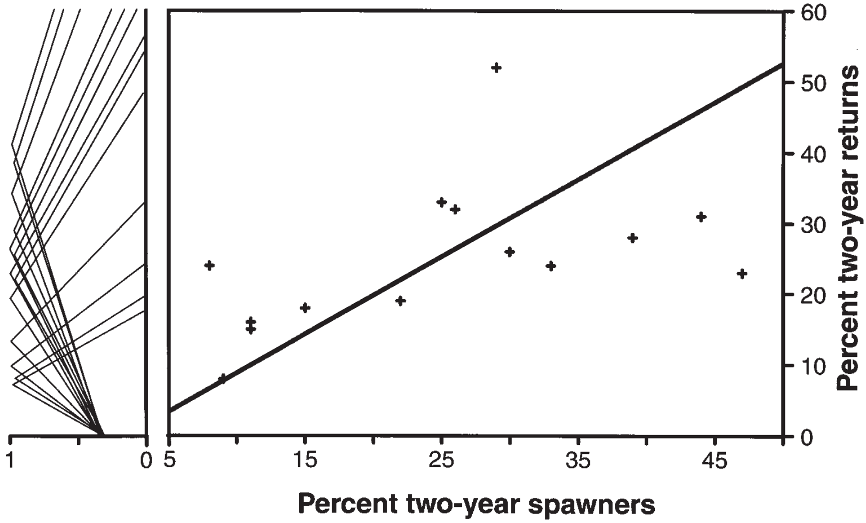


Figure 2. Linear regression with fuzzy parameters for the relationship between chinook salmon spawners and returns. The regression line is depicted in the graph on the right along with the original data from Unwin and Glova (1997) depicted as crosses. The small graph on the left is a rotated display of the fuzzy numbers predicted for returns by the regression for each x-value observed in the original data.

gree $h \in [0,1]$. This can be formulated as the optimization problem (Klir and Yuan 1995, page 456)

$$\text{Minimize } \sum_{i=1}^{n} \left| s_i - \Sigma_{j=1}^{m} \left| a_j \right| s_{ij} \right|,$$

$$\text{Subject to} \tag{6}$$

$$-\sum_{j=1}^{m} \left| a_j \right| s_{ij} + \sum_{j=1}^{m} a_j x_{ij} \le y_i - s_i, \text{ for all } i = 1,\dots,n,$$

$$\sum_{j=1}^{m} \left| a_j \right| s_{ij} + \sum_{j=1}^{m} a_j x_{ij} \ge y_i - s_i, \text{ for all } i = 1,\dots,n,$$

where $a_j \in \Re$ for all $j = 1,\dots, m$.

To illustrate the application of linear regression with fuzzy data, we reanalyzed a subset of the data used in the previous example. For this example, however, the independent and dependent variables are now presumed to be fuzzy numbers which characterize the measurement error or

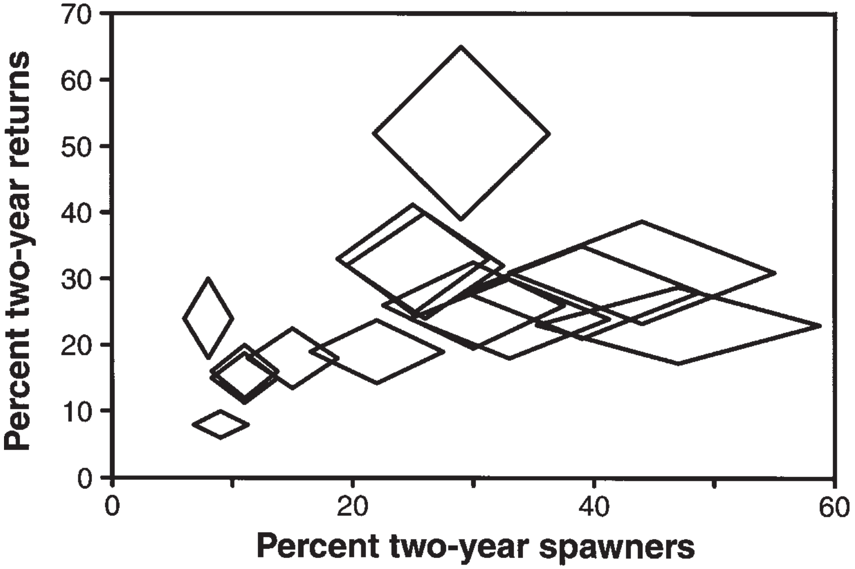


Figure 3. Data to be used in a linear regression with fuzzy data. Based on the chinook salmon data of Unwin and Glover (1997). Each diamond represents a pair of fuzzy numbers.

ambiguity of the values. Although the regression technique can accept fuzzy numbers with arbitrary uncertainties, for the sake of simplicity, we assumed the uncertainties to be proportional to the estimated magnitude. The data, which are displayed in Fig. 3, are therefore bivariate symmetric triangular fuzzy numbers, each with a breadth in each dimension of ±25% the original value. The LINGO optimization modeling language yielded the solution (according to the chosen criteria) $Y = 0.864X$.

Tables 1 and 2 summarize and compare the results from three kinds of linear regression on the Unwin and Glova data set. Table 1 gives the (central) predicted values and the residuals for ordinary least squares regression (OLS), linear regression with fuzzy parameters (LRFP), and linear regression with fuzzy data (LRFD). Predicted values are presented for each x-value observed in the data set. Residuals are the differences between the observed and predicted values for the Y variable. Values are displayed to two decimal places for the benefit of those who attempt to repeat these calculations, although this may not reflect the precision of the data set.

Table 2 compares the three regressions using quality coefficients defined by Pop and Sarbu (1996). $QC1$ is a coefficient used in analytical chemistry and elsewhere to judge the goodness of fit to a regression line. $QC5$ relativizes residuals by the maximum of absolute residuals and has a

Table 2. Values of quality coefficients described by Pop and Sarbu (1996) for the three regressions on the chinook salmon data of Table 1.

	Model type		
Coefficient type	OLS	LRFP	LRFD
$QC1$	0.81	0.82	0.82
$QC5$	1.6	1.6	1.5
$NQC5$	0.21	0.21	0.18
$QC6$	5.0	5.0	5.1
$NQC6$	0.12	0.12	0.13

$$QC1 = \left[\frac{1}{n-1} \sum_{i=1}^{n} \left(\frac{r_i}{\hat{y}_i} \right)^2 \right]^{1/2}$$

$$QC5 = \left[\sum_{i=1}^{n} \left(\frac{r_i}{\max|r_i|} \right)^2 \right]^{1/2} \qquad NQC5 = \frac{QC1 - 1}{\sqrt{n} - 1}$$

$$QC6 = \left[\sum_{i=1}^{n} \left(\frac{r_i}{r} \right)^2 \right]^{1/2} \qquad NQC6 = \frac{QC6 - \sqrt{n}}{n - \sqrt{n}}$$

In the formulas for the coefficients given at the bottom of the table, n is the number of data points, the r_i are the residuals $(y_i - \hat{y}_i)$, where y_i and $\hat{y}_i$ are observed dependent values at each datum and those predicted by the model respectively, and r is the mean of the r_i.

range [1, $n^{1/2}$]. $QC6$ uses the mean of absolute residuals and has a range [$n^{1/2}$, n]. These coefficients were normalized by Pop and Sarbu (1996) as $NQC5$ and $NQC6$ which take values between [0,1] as more practical ranges for evaluative purposes. Table 2 shows that both fuzzy regressions compare reasonably well with the traditional least squares regression.

Reexamination of Figure 1 Data

The seven data points of Fig. 1 were analyzed by fitting a fuzzy regression (FR) with intercept to the data because it is the most commonly used form of regression. The fuzzy regression procedure used for this example was developed from an algorithm provided by Pop and Sarbu (1996) termed the Modified Fuzzy 1-Lines. Results obtained by using this algorithm were compared by the above authors with ordinary least squares and weighted least squares, as well as several robust regression methods. This fuzzy regression procedure was found to exceed the least squares method and equal or exceed all other methods tested including two previously proposed fuzzy methods. The criteria for comparison included the maximum of absolute residuals and the mean of absolute residuals as well as other preference measures. The Pop and Sarbu algorithm was converted into a FORTRAN 77 program and compared with the results from examples provided by the above-mentioned authors. Our results were virtually identical to those reported. We also implemented the quality coefficients used by Pop and Sarbu to evaluate the goodness of fit of the regression lines. The software is available from the authors.

The fuzzy regression program (FR) developed from the Modified Fuzzy 1-Lines algorithm produces the fuzzy number associated with the classical set X to the membership threshold $\alpha (\alpha = 0.05$ in our applications) together with its linear representation. Our program also estimates the five quality coefficients used by Pop and Sarbu, as well as a membership function describing the grade of membership for each observed point in the data set which may be useful in detecting outliers in the data.

Table 3 shows the results of comparing the OLS regression with intercept to the new fuzzy regression (FR) in terms of the five quality coefficients. Note that most of the quality coefficients using the fuzzy regression method were smaller than those derived from the OLS regression. This indicates that the performance of the fuzzy regression method was superior using these several quality coefficients.

Calibration Program

The problem addressed in this example is the so-called calibration problem which has received attention by statisticians as well as by scientists in many disciplines, including fishery scientists (ICES 1993). This problem can be loosely defined as a prediction from past data on the relationship between a precise and an imprecise measurement of the same thing, using a future imprecise observation to predict the more precise value. An example is the estimation of year class strength (recruitment) from a survey index, based on calibration against virtual population analysis (VPA) estimates of past year class strengths. We believe that the recent results from calibration studies in analytical chemistry using fuzzy regression (Pop and Sarbu 1996) may be useful in contributing to a better understanding of the calibration problem in a fisheries context.

We consider the estimation of year class strength from a survey index against VPA estimates of past year class strength for the northwest Atlantic mackerel. The data set is from Clark (1979), and it has also been used by the 1984 ICES Working Group on Methods of Fish Stock Assessments. We compare the results obtained by the Modified Fuzzy 1-Lines regression with results obtained by OLS regression with intercept based on log-transformed variables. The transformed data and the results obtained from the two regressions are shown in Table 4. From an examination of the residuals by year between the OLS and the FR methods, it is evident that the FR method had smaller residual values in five of the eight years. How-

Table 3. **Comparison of OLS versus FR regressions on the flow versus survival data from Cada et al. (1997) shown in Fig. 1.**

Quality coefficients	Method	
	OLS	FR
*QC*1	0.026	0.025
*QC*5	1.9	1.4
*NQC*5	0.55	0.25
*QC*6	3.5	2.7
*NQC*6	0.20	0.23

ever, the quality coefficients indicate slightly higher (i.e., worse) values for the FR method. This is not altogether surprising. By design, the OLS regression yields the line of best fit as measured by the sum of squared residuals. The quality coefficients are closely related to this criterion. The FR method may be useful as another form of robust regression that is insensitive to large residuals and therefore appropriate for data sets which may contain outliers.

Discussion and Conclusions

Fuzzy regression has been described and illustrated with some simple examples from recent fisheries literature describing important relationships based on relatively limited data. The results suggest that fuzzy regression compares favorably with standard statistical regression procedures in terms of performance as measured by certain quality coefficients. We believe fuzzy regression can overcome some of the weaknesses of OLS. Justification for these statements is provided by the following observations.

In the case of applying linear regression with fuzzy parameters (LRFP) and linear regression with fuzzy data (LRFD) to the chinook salmon returns example (Table 1), the residuals from both types of fuzzy regression are very similar to the least squares residuals. The comparison of quality coefficients (Table 2) also indicates that the performances for all methods are quite similar.

The quality coefficients (Table 3) also have very similar values for traditional and fuzzy regressions of the flow versus survival relationship depicted in Fig. 1. Four out of the five coefficients, however, suggest the fuzzy regression is superior.

Table 4. Raw data on northwest Atlantic mackerel survey indices versus VPA results based on Clark (1979) comparing fuzzy regression (FR) and ordinary least squares (OLS) regression for log-transformed data.

	Transformed data		OLS		FR	
Case	Index	VPA Recruits	VPA Predicted	Residuals	VPA Predicted	Residuals
1	0.000	6.68	6.85	–0.17	6.90	–0.05
2	0.406	7.31	6.97	0.35	7.02	0.30
3	1.610	7.35	7.31	0.04	7.35	–0.00
4	1.792	7.09	7.36	–0.27	7.40	–0.31
5	1.946	7.17	7.40	–0.23	7.44	–0.27
6	2.526	7.74	7.57	0.17	7.60	0.14
7	2.603	7.70	7.59	0.10	7.63	0.07
8	3.091	7.74	7.73	0.01	7.76	0.02

Quality coefficients		
*QC*1	0.030	0.030
*QC*5	1.6	1.9
*NQC*5	0.34	0.47
*QC*6	3.4	3.5
*NQC*6	0.10	0.12

For the Atlantic mackerel data, the results are again very similar (Table 4), although the OLS results seem to be slightly but consistently higher. The FR intercept value for the fuzzy regression was 6.85 versus 6.90 for the least-squares regression. The slope of the fuzzy regression was 0.278 versus 0.288 for the OLS model. This suggests that the FR model weights large values of residuals less than the OLS model.

It should be recalled that OLS regression is based on the assumption of an independent and normal error distribution with uniform variance (homoscedasticity). In fisheries applications such as the calibration problem, the y-direction error is dependent on the survey index value and on the presence of outliers which lead to heteroscedastic results. In practice, we don't usually know the shape of the error distribution function and its variance. A consequence is that the least squares method does not lead to the maximum likelihood estimate. If the tails of the experimental error distribution contain a much larger percentage of the total area than the normal distribution, the best linear unbiased estimator may not be very good. The least squares method is well known to be sensitive to the effect of large residuals, so results are distorted if there are large differences between the observed data and the model predictions. Fuzzy regression appears to overcome some of these difficulties.

A basic issue relates to the interpretation of fuzzy regression results. These cannot be interpreted in the same way as the results from classical regression. The prediction from a fuzzy regression is a fuzzy number for each value of the independent variable(s). From a practical point of view, we believe that the results from fitting a fuzzy regression to uncertain or imprecise data may provide a fishery manager with a more conservative but realistic indication of the variability and uncertainty associated with a predicted response. This is in keeping with the so-called precautionary approach to fisheries management.

This work has only briefly introduced the use of fuzzy regression. It can be extended to multiple independent variables (e.g., Heshmaty and Kandel, 1985) and can also be applied to quadratic and other polynomial relationships (Bárdossy 1990). The results of this work can be briefly summarized as follows:

1. Fuzzy regression leads to a mathematical programming problem which can be solved in a straightforward manner.

2. Rules of fuzzy mathematics are used to define the regression parameters on fuzzy numbers.

3. Results of fuzzy regression applications indicate that it compares favorably with ordinary least squares regression.

4. Our limited experience in applying fuzzy regression to some fisheries-related data suggests that it may be useful in overcoming some difficulties associated with heteroscedastic and uncertain data.

References

Bárdossy, A. 1990. Note on fuzzy regression. Fuzzy Sets and Systems 37:65-75.

Bárdossy, A., I. Bogardi, and L. Duckstein. 1990. Fuzzy regression in hydrology. Water Resour. Res. 26(7):1497-1508.

Cada, G.F., M.D. Duncan, S.V. Mitz, and M.S. Bavelhimer. 1997. Effects of water velocity on the survival of downstream-migrating juvenile salmon and steelhead: A review with emphasis on the Columbia River Basin. Reviews in Fisheries Science 5(2):131-183.

Clark, S.N. 1979. Application of bottom-trawl survey data to fish stock assessment. Fisheries 4(3):9-15.

Ferson, S., W. Root, and R. Kuhn. 1998. RAMAS Risk Calc: Risk assessment with uncertain numbers. Applied Biomathematics, Setauket, NY.

Heshmaty, B., and A. Kandel. 1985. Fuzzy linear regression and its applications to forecasting in uncertain environments. Fuzzy Sets and Systems 15:159-191.

ICES. 1993. Reports of the working group on methods of fish stock assessments. ICES Coop. Res. Rep. 191.

Johnson, P.A., and B.A. Ayyub. 1996. Modeling uncertainty in prediction of pier scour. Journal of Hydraulic Engineering, 1 February 1996:66-76.

Kaufmann, A., and M.M. Gupta. 1985. Introduction to fuzzy arithmetic: Theory and applications. Van Nostrand Reinhold, New York.

Klir, G.J., and B. Yuan. 1995. Fuzzy sets and fuzzy logic: Theory and applications. Prentice Hall, Upper Saddle River, NJ.

Pop, H.F., and C. Sarbu. 1996. A new fuzzy regression algorithm. Anal. Chem. 66:771-778.

Prairie, Y.T. 1996. Evaluating the predictive power of regression models. Can. J. Fish. Aquat. Sci. 53:490-492.

Ricker, W.E. 1973. Linear regressions on fishery research. J. Fish. Res. Board Can. 30:409-434.

Ryan, T.R. 1997. Modern regression methods. J. Wiley and Sons, New York.

Sparre, P., and S.C. Venema. 1992. Introduction to tropical fish stock assessment. FAO Fish. Tech. Pap. 306/1, Rev. 1.

Tanaka, H., S. Uijima, and K. Asai. 1982. Linear regression analysis with fuzzy model. IEEE Trans. on Systems, Man, and Cybernetics 12(6):903-907.

Unwin, M.J., and G.J. Glova. 1997. Changes in life history parameters on a naturally spawning population of chinook salmon (*Oncorhynchus tshawytscha*) associated with release of hatchery-raised fish. Can. J. Fish. Aquat. Sci. 54:1235-1245.

Zadeh, L.A. 1965. Fuzzy sets. Information and Control 8(3):338-353.

Evaluation of Assumed Error Structure in Stock Assessment Models That Use Sample Estimates of Age Composition

Paul R. Crone
National Marine Fisheries Service, Northwest Fisheries Science Center, Hatfield Marine Science Center, Newport, Oregon

David B. Sampson
Oregon State University, Hatfield Marine Science Center, Newport, Oregon

Abstract

The sampling error associated with estimates of age composition for five groundfish species commercially landed at Oregon ports is used to examine the ability of age-structured stock assessment models to adequately describe the stochastic properties of actual catch-at-age data. Specifically, estimated coefficients of variation associated with samples of catch-at-age are presented graphically to evaluate a theoretical consideration involved in stock assessment models widely used in marine fishery management. Results presented here indicate that a multinomial probability error structure, included in models that are based on maximum likelihood estimation, more closely follows the variability associated with the sampled landing data than does a lognormal error structure used in models based on least squares estimation. Weighted nonlinear regression analysis is used to determine the specific multinomial distribution (sample size *n*) that provides the most accurate description of the actual variability associated with the sample estimates of age composition. Implications for stock assessment modeling are discussed. Finally, a linear regression model is derived that describes the relationship between multinomial sample size and the number of boat trips sampled, in efforts to provide an adequate error structure for the models without having to rely on the relatively complex and tedious sampling estimators and subsequent analytical techniques.

Introduction

In recent years, stock assessment models have gained widespread application in fisheries management. In particular, several age-structured assessment methods have become the primary tools used to derive estimates of fishery parameters, such as fish population abundance and exploitation rates, in many fisheries throughout the world (Megrey 1989). The motivation behind these assessment methods is that a time series of catch-at-age data for a fishery (i.e., estimates of age composition of the landings for a particular stock) can be utilized in models containing certain biological processes of fish populations, such as growth, mortality, and reproduction. The extent to which these models can describe or predict the inherent stochastic properties of an animal population is based largely on the validity and reliability associated with their parameters and assumptions (Pielou 1977, Gulland 1983, Gottelli 1995).

It is generally agreed that estimates of catch-at-age alone are insufficient to reliably determine the status of exploited fish stocks (Doubleday 1976, Pope 1977, Megrey and Wespestad 1988, Quinn and Collie 1990). The types and function of auxiliary data in fishery models need to be examined rigorously to safeguard against inappropriate application in management situations. In particular, it is critical that important model assumptions be reviewed and tested to ensure that generated results are interpreted appropriately (Gudmundsson 1986, Edwards and Megrey 1989, Schnute 1989, Sampson 1993).

To date, one of the most important advances in model development has been the inclusion of an error structure to address the variability associated with the: (1) separate biological processes that influence fish population abundance, such as reproduction and mortality; and (2) input data on which the models are based, namely the sample estimates of catch-at-age (see Megrey 1989) for an extensive review of age-structured stock assessment models). The age-structured assessment models that accommodate stochastic data can be broadly classified into two groups based on the statistical estimation technique that is used (Kimura 1989): (1) the method of least squares (e.g., Doubleday 1976; Pope and Shepherd 1982; Deriso et al. 1985, 1989; Kimura 1989); or (2) the method of maximum likelihood based on multinomial distribution probabilities (e.g., Fournier and Archibald 1982; Dupont 1983; and Methot 1989, 1990). Note that the least squares and maximum likelihood methods generate equivalent solutions if the error terms are assumed to be distributed as normal random variables (Bain and Engelhardt 1987). In addition, although the two groups of models above are most often defined by distinct error structures, each estimation method (nonlinear least squares and maximum likelihood) is capable of fitting lognormally and multinomially distributed catch-at-age data (Kimura 1990). If least squares residuals are weighted inversely by their expected variances, the least squares method can be used in an iterative manner to generate maximum likelihood estimates (McCullagh and Nelder 1989).

An important theoretical consideration when choosing an appropriate stock assessment model involves defining the correct sampling distribution for estimates of age composition (Kimura 1990). The focus of the work described here is the assumed error structure for age composition data relied upon in the different models, which is one criterion that can be used to differentiate the two groups of models discussed above. That is, the objective of this research does not directly involve a critical examination of the methods of estimation utilized by the different models. Also, note that there exist other sources of error (uncertainty) associated with age-structured stock assessment models that are not evaluated in this paper, including measurement of: (1) aging error (i.e., error associated with determining the age of a fish specimen); (2) effort error; (3) spawner-recruit error; (4) stochastic mortality; and (5) total catch, including discard and bycatch estimates (Megrey 1989).

The suite of models that is based on least squares estimation generally assumes that observation errors in catch-at-age data are lognormally distributed and the models use $\log_e$ transformed estimates of catch-at-age. The least squares estimators assume constant variance of the log transformed estimates of catch-at-age, which dictates that the coefficients of variation (CVs) associated with the untransformed catch-at-age estimates be approximately equal (Kimura 1989, 1990; Methot 1990).

In the models that use maximum likelihood estimation, the error structure for catch-at-age data is based on multinomial probabilities, which implies that the CVs associated with the estimates of proportion-at-age are distributed in a multinomial fashion (Methot 1990). That is, the magnitude of the CV decreases steadily as the size of the proportion increases (Fig. 1). The sampling variance of a proportion derived from a multinomial distribution is treated as a binomially defined parameter and calculated as $V(P) \cong [P(1-P)/n]$, where P is the population proportion and n is the number of units in the sample. The CV of P is calculated as $(\sqrt{V(P)}/P)$. In probability theory, n can be practically interpreted as an index that defines a particular distribution from an infinite number of possible probability distributions. In stock assessment models, n represents the sample size associated with an estimated age composition. It is a weighting factor, such as J_y below, that adjusts the theoretical CV versus P curve upward or downward to reflect the estimated variability associated with the age-composition sample. Henceforth, the term “multinomial index” is used to identify a specific n from the family of possible distributions.

The general form of the objective function used in the two groups of models to derive fishery-related parameters is as follows, summation being over $y = 1,\ldots,Y$ for years and $a = 1,\ldots,A$ for ages: (1) least squares (Deriso et al. 1985), choose parameters that minimize

$$\sum_{y=1}^{Y}\sum_{a=1}^{A}[\log_e(c_{ya}) - \log_e(\hat{c}_{ya})]^2; \text{ and}$$

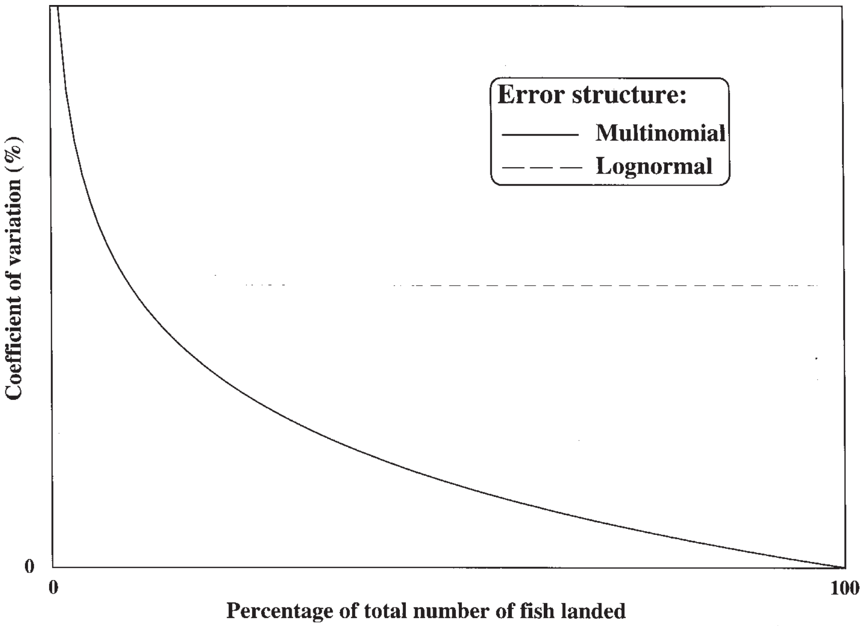


Figure 1. Distributions of estimated coefficients of variation (%) associated with estimates of age composition (percentage of total number of fish landed) for two different error structure assumptions used in fish stock assessment models. Lines depict the general form of the distributions and do not reflect the inherent variability associated with samples of age composition.

(2) maximum likelihood (Methot 1990), choose parameters that maximize

$$\sum_{y=1}^{Y} \sum_{a=1}^{A} J_y p_{ya} \log_e(\hat{p}_{ya}),$$

where c_{ya} is the observed catch-at-age (in number), $\hat{c}_{ya}$ is the predicted catch-at-age (in number), J_y is a weighting factor that reflects the number of fish in the sample if the fish were selected as a single simple random sample (i.e., if the multinomial probability distribution was strictly correct), p_{ya} is the observed proportion-at-age, and $\hat{p}_{ya}$ is the predicted proportion-at-age. The assumption regarding the pattern of variability exhibited by the actual sample estimates of catch-at-age is generally different between these two groups of models (Fig. 1).

The primary objective of this research was to examine the statistical properties associated with estimates of age composition for groundfish landings in Oregon from 1989 to 1994. Specifically, we graphically present the distributions of estimated CVs associated with estimates of age composition for five species of groundfish to evaluate the appropriateness of the error structure assumed in fishery models to describe the uncertainty associated with the catch-at-age sample data. In addition, we present a nonlinear regression technique and generally discuss other methods that can be used to determine a specific multinomial index (n) to apply to particular age-composition data sets. Finally, a simple linear regression model is developed that describes the relationship between the multinomial indices and the number of boat trips sampled.

Methods

Estimates of age composition (in number and percent) and their errors were calculated for five species of groundfish commercially landed at Oregon ports from 1989 to 1994: widow rockfish (*Sebastes entomelas*), yellowtail rockfish *(Sebastes flavidus),* canary rockfish *(Sebastes pinniger),* English sole *(Pleuronectes vetulus),* and Dover sole *(Microstomus pacificus).* Age-composition data sets for 22 combinations of species and year were analyzed in this study, i.e., age-composition samples for 8 species-year combinations were not collected or had specimens that had not been analyzed for age determination at the time of this study. The analyses of age composition for this study were based on a stratified two-stage random sampling design combined with poststratification (Fig. 2). The sampling design used for 1989 and 1990 landings incorporated a single sampling unit (basket of fish) at the second stage and utilized straightforward ratio estimation techniques to derive landing statistics (Sen 1986). A “complete” multistage design was used for 1991-94 landings, which incorporated replicate sampling units at the second stage and used standard two-stage estimators to generate landing estimates (Crone 1995).

A coefficient of variation (CV), (standard error / estimate), was used to describe the variability associated with the individual landing estimates of age composition. This statistic is also referred to as a relative standard error (Som 1973) and a coefficient of variation of the estimate (Cochran 1977). Weighted nonlinear least squares regression (Neter et al. 1990) was applied to estimates of age composition and their estimated CVs to determine multinomial indices (n) for combinations of species and year, with estimated proportions-at-age used as the weights. Simple linear regression methods (e.g., Draper and Smith 1981; Neter et al. 1990) were applied to relate the multinomial indices with the number of boat trips sampled. The number of boat trips represents the first-stage sample size in estimators

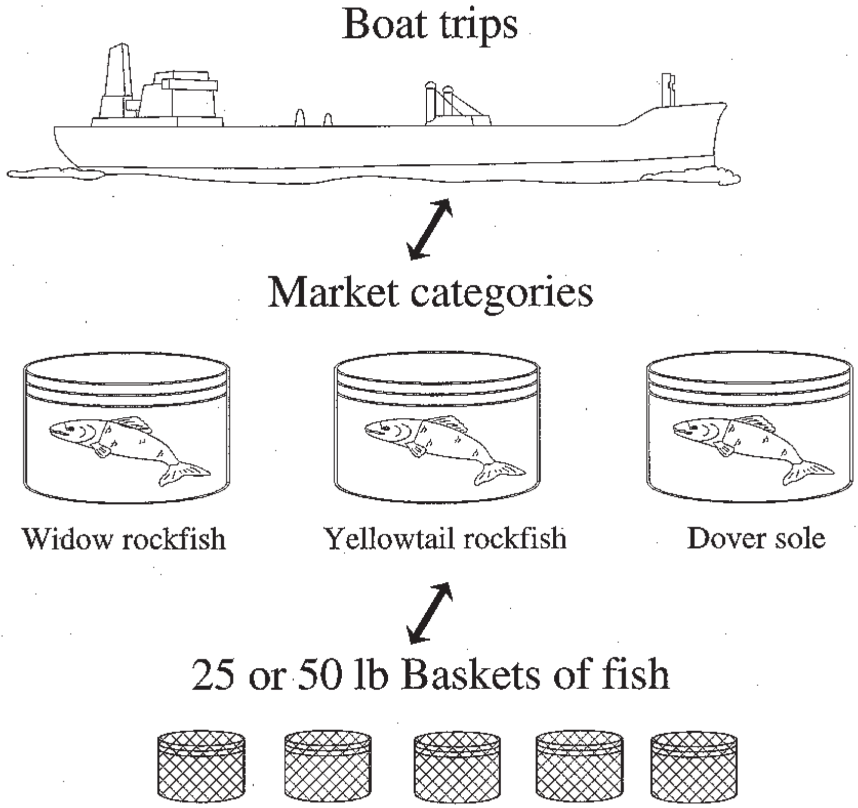


Figure 2. Two-stage sampling design used to monitor age compositions of commercially landed groundfish in Oregon (1989-1994). Port and quarter combinations were treated as strata. Three examples of market categories (presorted groups of fish) are presented. Top: primary sampling units; middle: poststratification units; bottom: secondary sampling units.

associated with the multistage sampling design used to generate age-composition statistics (Fig. 2).

Results

Results from age-composition analyses were very consistent for the five species evaluated in this study. Selected species-year combinations are presented as examples that generally describe the statistical properties of the age compositions.

The relationship between the individual estimates of age composition, presented as percentages of the total number of fish landed, and their associated CVs was negatively curvilinear for each species-year combination (e.g., 1991 results are presented in Fig. 3). The actual ages associated with the data points are not included on the graphs; however, particular data points (ages) are identified to highlight general patterns exhibited in the age compositions. Also, for purposes of graphical clarity, we have omitted age groups that constituted less than 0.07% of the total landings; the CVs associated with these age groups were very consistent across species and years, ranging from 70 to 110%. In general, estimates that composed large percentages of the total landings were measured with higher relative precision (i.e., had smaller CVs) than estimates of age composition that constituted small percentages of the total landings.

For example, 851,558 age-6 English sole were landed in 1991, which was roughly 34% of the total number of English sole landed for the year, and the CV associated with this estimate was 13%; whereas, 16,071 age-16 fish were landed, which was approximately 1% of the total number landed, and this estimate had a CV of 87% (star-filled squares, middle right panel of Fig. 3).

The CVs associated with the landing estimates clearly mimicked the curves generated from the theoretical multinomial distributions, presented in the figures for $n = 100$, 400, and 1,000. The role of n as weighting factors in stock assessment models that incorporate an error structure based on multinomial probabilities is discussed below. Although the patterns of variation that characterized the estimated age compositions were generally similar between the five species, the amount of statistical "noise" associated with each set of estimates (species-year combinations) was not identical. For example, ages 7, 14, and 15 made up roughly 5% (approximately 47,000 fish) of the total landings of canary rockfish in 1991; however, the CVs associated with these similar landing estimates ranged from 17 to 33% (star-filled squares, middle left panel of Fig. 3). For Dover sole, ages 15 and 16 individually composed approximately 5% (roughly 1,000,000 fish) of the total landings and these estimates had CVs that were less than 12% (star-filled squares, lower left panel of Fig. 3).

The theoretical distributions more accurately reflected the variation associated with the estimates of age composition that contributed significantly to the total landings than they did for estimates that constituted small percentages of the total, particularly estimates that composed less than 1% of the total. For example, the estimate of 759,068 (approximately 18%) age-7 widow rockfish landed in 1991 had a CV of 8% (star-filled square, upper right panel of Fig. 3), with CVs of 21, 11, and 7% associated with the analogous percentage defined by the multinomial indices for $n = 100$, 400, and 1,000, respectively. Whereas, the estimate of 3,634 (roughly 0.09%) age-31 widow rockfish had a CV of 59% (star-filled square, upper right panel of Fig. 3), compared with CVs of 333, 167, and 105%, for the three theoretical distributions defined by $n = 100$, 400, and 1,000, respectively.

In general, at least three-fourths of the total landings of a species and year combination was composed of a relatively small range of consecu-

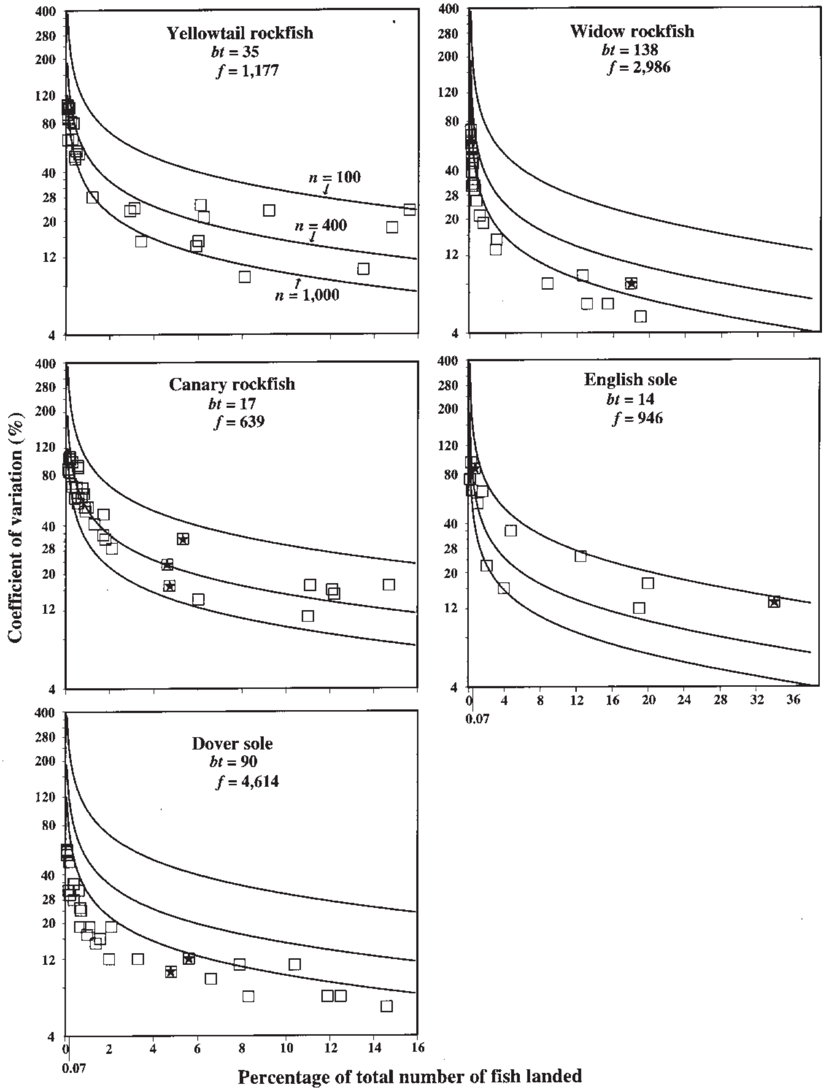


Figure 3. Distributions of coefficients of variation (%) associated with estimates of age composition (percentage of total number of fish landed denoted by squares) for five species of groundfish landed at Oregon ports in 1991. See Results section for description of star-filled squares. Coefficients of variation derived from percentage estimates for three multinomial indices are presented, n = 100, 400, and 1,000. Estimates for ages that composed at least 0.07% of the total landings are included. The Y-axis has been logarithmically scaled. For each species, the sample sizes (number of boat trips) used to derive age-composition statistics are denoted as bt and the total number of fish collected across all boat trips is denoted as f.

tive ages that individually contributed at least 5% to the total and these estimates were relatively precise, with CVs less than 25%. The remaining approximately one-fourth of each age composition included comparatively more ages that individually composed less than 5% of the total and these estimates were more variable than for the ages that constituted the three-fourths majority, with CVs generally greater than 30% and most often between 50 and 100%. Detailed landing statistics for yellowtail rockfish in 1990 illustrate the general properties of the age-composition sample data collected in Oregon from 1989 to 1994 (Table 1).

Weighted nonlinear regression analysis was used to identify the multinomial index (n) that most accurately characterized the actual variation associated with estimates of age composition for combinations of species and year, with weights equal to the estimated proportions-at-age. In effect, ages that were landed more frequently and reflected relatively large proportions of the total landings were given more importance in the fitting procedures than those ages that were associated with small proportions. For example, using this straightforward analytical technique, the age composition of yellowtail rockfish landed in 1991 (upper left panel of Fig. 3) was characterized by a multinomial index with an n of 688.

Following the determination of multinomial indices for age compositions by species and year, a linear regression model was developed that describes the relationship between the multinomial indices and the corresponding number of boat trips sampled for the species-year data sets. The amount of variation associated with a species-year age composition is directly influenced by the number of boat trips sampled (Fig. 2; Crone 1995). A relatively simple model was developed that accounted for a significant portion of the variability ($r^2 = 0.90$) in the dependent variable (n), where the number of boat trips sampled was statistically related ($P < 0.001$) to the multinomial index (Fig. 4). The total number of fish that constituted a species-year age-composition sample was another sampling descriptor that was considered as a potential independent variable in the model. Although this variable was found to be a statistically significant ($P < 0.001$) term when included in the regression model along with the number of boat trips sampled, we felt its presence in the model was not applicable, given: (1) it did not substantially increase the amount of total variation in multinomial index (n) accounted for by the regression equation ($r^2 = 0.94$); and more important, (2) the strong correlation (multicollinearity) that existed between the two independent variables ($r = 0.80$) confounds the interpretation and use of a fitted regression model (Neter et al. 1990).

Discussion

Results presented here indicate that stock assessment models that utilize maximum likelihood estimation techniques with a multinomial probability error structure (e.g., Fournier and Archibald 1982; Methot 1989, 1990) more adequately address the variability associated with observed catch-

Table 1. Age-composition estimates for yellowtail rockfish landings in Oregon (1990).

Age	Landing estimate	Percent of total landings	CV (%)
4	236	<1	102
5	23,759	2	45
6	140,419	11	33
7	132,329	10	18
8	81,400	6	18
9	135,648	10	15
10	142,835	11	10
11	96,412	7	15
12	101,234	8	14
13	89,401	7	18
14	102,838		25
15	74,107	6	13
16	62,956	5	22
17	37,099	3	26
18	21,001	2	32
19	13,310	1	28
20	6,461	<1	39
21	10,613	1	32
22	1,142	<1	55
23	8,852	1	42
24	1,000	<1	78
25	1,457	<1	86
26	1,464	<1	71
27	6,556	<1	44
28	6,479	<1	77
29	3,134	<1	52
30	2,911	<1	81
31	2,634	<1	77
32	90	<1	102
34	1,638	<1	78
36	701	<1	100
37	393	<1	101
38	392	<1	101
40	43	<1	103
41	2,786	<1	75
42	1,638	<1	78
46	1,245	<1	98
Total	1,316,613	100	

Landing estimates are in number of fish and are based on a sample of 35 boat trips.

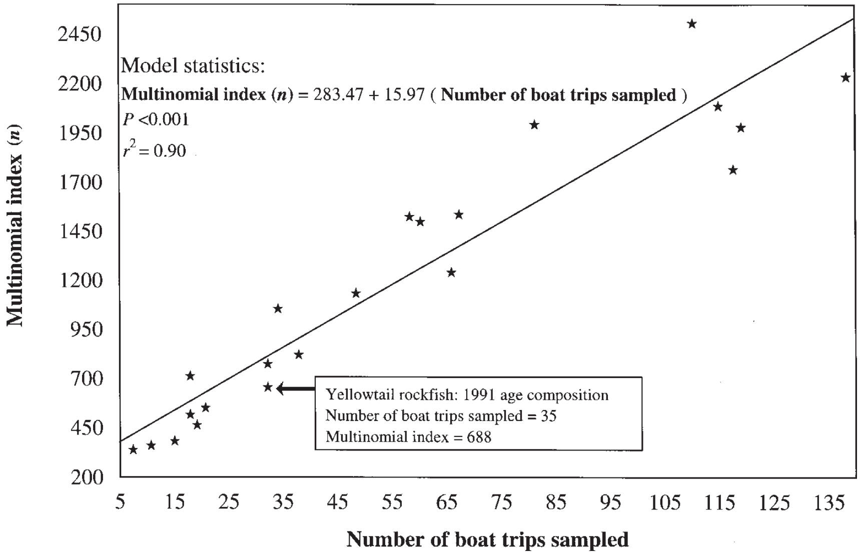


Figure 4. Multinomial index (n) as a linear function of the number of boat trips sampled. The multinomial index (dependent variable) represents a theoretical sample size (n) associated with multinomially distributed proportion-at-age estimates. The number of boat trips sampled (independent variable) represents the actual sample size associated with sampling estimators used to derive the age-composition landing statistics. Data points (★) represent multinomial indices associated with age-composition data sets for combinations of species and year; an example is shown for the age composition of yellowtail rockfish landed in 1991 (see upper left panel of Fig. 3 for estimates of age composition and their CVs for yellowtail rockfish landings in 1991).

at-age data than models based on lognormal measurement errors. The statistics generated from these analyses show clearly that the CVs associated with individual estimates of age composition are not constant, or even approximately so, but rather follow the general properties of a multinomially distributed variable.

The choice of the most appropriate multinomial index (*n*) to use in a model is inherently problematic. The sampling designs generally employed to collect fishery-related data, along with the selection protocols utilized in the field, generate age-composition estimates that necessarily depart, to some degree, from a strict theoretical probability distribution, multinomial or otherwise. That is, the catch-at-age data used in stock assessment models are most often based on a sample that consists of many boat trips, which have been selected over a period of time using nominally random

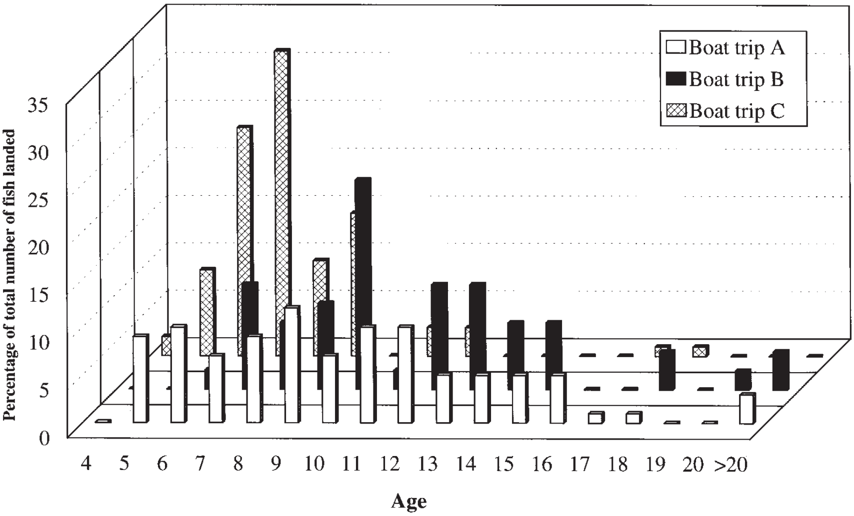


Figure 5. Estimated age compositions (in percent) for landings (in number) of Dover sole for three boat trips (A, B, and C) sampled at Oregon ports in 1994. Sampling design used to generate estimates is described in Crone (1995).

selection protocols (e.g., see Tomlinson 1971, Crone 1995), rather than a sample that was obtained in a strictly random manner from a single trip.

For example, multinomial distribution theory could be applied to a single random sample from a boat trip in a generally straightforward fashion, say to calculate the variance associated with the estimated proportion of age-6 yellowtail rockfish in the landings. However, the actual estimated variances associated with the landing estimates (in number or percent) were necessarily derived from appropriate sample estimation techniques that required no assumptions regarding distribution properties of the measurement variables. Kimura (1990) also argues that although rigorous multinomial sampling for catch-at-age data may be convenient from an intuitive or modeling standpoint, it is most often an impractical sample selection approach, given the expanse and dynamics of commercial fisheries.

Determination of the most appropriate multinomial index (*n*) becomes problematic because the age compositions estimated from the individual trips sampled are often considerably different from one another (Fig. 5), which confounds the use of explicit theory (e.g., a multinomial distribution supposition) to summarize the actual variability associated with the sample estimates. In the context of generalized linear modeling, this sampling scenario often results in "overdispersion" in the data, i.e., the variability associated with the landing estimates is greater than what is expected from a simple multinomial variable.

Models that account for overdispersion, sometimes called "mixed" models, can be adapted for multistage sampling designs, which requires estimating the scale of overdispersion from a residual deviance statistic (Francis et al. 1993). For example, a comparative study showed that a compound multinomial model was found to be more flexible than a simple multinomial model for defining the variance structure of combined length frequency samples from commercial cod landings; however, the authors advised that this mixed model be applied with caution until further research has fully established the estimation theory (Smith and Maguire 1983). The beta-binomial distribution has also proved useful for specifying and fitting a model that accounts for overdispersion (Crowder 1978). McCullagh and Nelder (1989) suggest an ad hoc approach for dealing with overdispersion that utilizes residual deviance statistics and is directly related to standard regression procedures.

We have shown that the specific form of the multinomial distribution error structure can be estimated using some of the general properties of commercial fishery sampling designs. The motivation behind the techniques we have presented is generally addressed in Shepherd and Nicholson (1986, 1991). That is, Shepherd and Nicholson (1991) present an intuitively attractive ad hoc method to determine an error structure for multiplicative modeling of catch-at-age data, but this method attempts to define a variance (error) structure for landing estimates of age composition in general and does not allow for an objective determination of an appropriate weighting factor. Although general methods, such as the approach proposed by Shepherd and Nicholson, do provide results that are broadly appropriate, the actual catch-at-age data may be much more or less precise than the results indicate, due largely to the effectiveness of the sampling design. We present methods that can be utilized to determine the most appropriate weighting factor (n) to describe the variation associated with a specific age composition(s), thus these methods require that the actual variance measures associated with the sample estimates be available.

Research objectives that address the impact of various error structures on model results have received sparse attention in fishery science (Megrey 1989). Intuitively, a model should be designed in a manner that allows the correct stochastic properties of the catch-at-age data to be incorporated, otherwise additional sources of bias are inherently introduced into the analytical processes. Methot (1990) suggested that the multinomial error structure is a preferred model feature because it emphasizes the variation associated with landing estimates that reflect large proportions of the total landings, which are documented here as being comparatively more precise than the estimates that constituted small proportions of the total. An experiment conducted to examine the sensitivity of assessment results to the assumption of constant selectivity showed that the assumed error structure could have a large impact on the final estimates generated from two different modeling approaches, namely stock synthesis analysis and Catch AGE ANalysis or CAGEAN (Sampson 1993).

In contrast, Deriso et al. (1985) demonstrated that a stock assessment model that utilized a least squares estimator (CAGEAN) generated similar results in a comparative study of three different theoretical distributions, based on lognormal measurement error, multinomial measurement error, and process error, applied individually to the model to address the stochastic properties of hypothesized catch-at-age data. Kimura (1990) simulated catch-at-age data using lognormal and multinomial error structures and then analyzed the data using nonlinear least squares and multinomial maximum likelihood estimation, and showed that the results from the overall analyses were similar. The author did however recommend that caution be used when interpreting his findings, given that the results from the simulation experiments may depend strongly on the population and constraints utilized in the model.

Further research is needed that focuses on the relationship between departures from assumptions and model output to critically examine the issue of statistical robustness of fishery models. Results presented here indicate that more complex error structures may be needed to account for the variability in catch-at-age sample data from commercial fisheries, and subsequently, methods need to be developed for incorporating these complicated error structures in the modeling process. The results and discussion presented here are used to evaluate the appropriateness of an assumption used in current age-structured assessment models to address the stochastic properties of catch-at-age data and should not be interpreted as broad recommendations of the overall performance of a model, given that these models incorporate a host of estimated parameters and other assumptions that were not investigated in this study.

Acknowledgments

This research was supported by: the Oregon Department of Fish and Wildlife, Marine Resources Division, Newport, Oregon (ODFW); grant no. NA89AA-D-SG108 (project no. R-ES-16) from the National Oceanic and Atmospheric Administration (NOAA) to the Oregon State University Sea Grant College Program; and by appropriations made by the Oregon State Legislature. The views expressed herein are those of the authors and do not necessarily reflect the views of NOAA or any of its subagencies. We thank Nick Davies, Terry Quinn, and an anonymous referee for their comments on the manuscript.

References

Bain, L.J., and M. Engelhardt. 1987. Introduction to probability and mathematical statistics. Duxbury Press, Boston. 565 pp.

Cochran, W.G. 1977. Sampling techniques, 3rd edn. John Wiley & Sons, New York. 428 pp.

Crone, P.R. 1995. Sampling design and statistical considerations for the commercial groundfish fishery of Oregon. Can. J. Fish. Aquat. Sci. 52:716-732.

Crowder, M.J. 1978. Beta-binomial ANOVA for proportions. Applied Statistics 27:34-37.

Deriso, R.B., P.R. Neal, and T.J. Quinn II. 1989. Further aspects of the theory of catch-age analysis. Can. Spec. Publ. Fish. Aquat. Sci. 108:127-135.

Deriso, R.B., T.J. Quinn II, and P.R. Neal. 1985. Catch-age analysis with auxiliary information. Can. J. Fish. Aquat. Sci. 42:815-824.

Draper, N.R., and H. Smith. 1981. Applied regression analysis, 2nd edn. John Wiley & Sons, New York. 709 pp.

Doubleday, W.G. 1976. A least squares approach to analysing catch at age data. Int. Comm. Northwest Atl. Fish. Res. Bull. 12:69-81.

Dupont, W.D. 1983. A stochastic catch-effort method for estimating animal abundance. Biometrics 39:1021-1033.

Edwards, E.F., and B.A. Megrey. 1989. Mathematical analysis of fish stock dynamics: Rationale and organization of this volume. Amer. Fish. Soc. Symp. 6:1-2.

Fournier, D., and C.P. Archibald. 1982. A general theory for analyzing catch at age data. Can. J. Fish. Aquat. Sci. 39:1195-1207.

Francis, B., M. Green, and C. Payne (eds.). 1993. GLIM 4: The statistical system for generalized linear interactive modeling. Clarendon Press, Oxford. 821 pp.

Gottelli, N.J. 1995. A primer of ecology. Sinauer Associates, Inc., Sunderland, MA. 206 pp.

Gudmundsson, G. 1986. Statistical considerations in the analysis of catch-at-age observations. J. Cons. Cons. Int. Explor. Mer 43:83-90.

Gulland, J.A. 1983. Fish stock assessment: A manual of basic methods. John Wiley-Interscience, New York. 223 pp.

Kimura, D.K. 1989. Variability, tuning, and simulation for the Doubleday-Deriso catch-at-age model. Can. J. Fish. Aquat. Sci. 46:941-949.

Kimura, D.K. 1990. Approaches to age-structured separable sequential population analysis. Can. J. Fish. Aquat. Sci. 47:2364-2374.

McCullagh, P., and J.A. Nelder. 1989. Generalized linear models, 2nd edn. Chapman and Hall, London. 511 pp.

Megrey, B.A. 1989. Review and comparison of age-structured stock assessment models from theoretical and applied points of view. Amer. Fish. Soc. Symp. 6:8-48.

Megrey, B.A., and V.G. Wespestad. 1988. A review of biological assumptions underlying fishery assessment models. In: W.S. Wooster (ed.), Lecture notes on coastal and estuarine studies: Fishery science and management 28. Springer-Verlag, New York, pp. 31-69.

Methot, R.D. 1989. Synthetic estimates of historical abundance and mortality for northern anchovy. Amer. Fish. Soc. Symp. 6:66-82.

Methot, R.D. 1990. Synthesis model: An adaptable framework for analysis of diverse stock assessment data. Int. North Pac. Fish. Comm. Bull. 50:259-277.

Neter, J., W. Wasserman, and M.H. Kutner. 1990. Applied linear statistical models. Richard D. Irwin, Inc., Boston, MA. 1,181 pp.

Pielou, E.C. 1977. Mathematical ecology. John Wiley & Sons, New York. 385 pp.

Pope, J.G. 1977. Estimation of fishing mortality, its precision and implications for the management of fisheries. In: J.H. Steele (ed.), Fisheries mathematics. Academic Press, London, pp. 63-76.

Pope, J.G., and J.G. Shepherd. 1982. A simple method for the consistent interpretation of catch-at-age data. J. Cons. Cons. Int. Explor. Mer 40:176-184.

Quinn II, T.J., and J.S. Collie. 1990. Alternative population models for eastern Bering Sea pollock. Int. North Pac. Fish. Comm. Bull. 50:243-257.

Sampson, D.B. 1993. The assumption of constant selectivity and the stock assessment for widow rockfish, *Sebastes entomelas*. Fish. Bull. U.S. 91:676-689.

Schnute, J. 1989. The influence of statistical error on stock assessment: Illustrations from Schaefer's model. Can. Spec. Publ. Fish. Aquat. Sci. 108:101-109.

Sen, A.R. 1986. Methodological problems in sampling commercial rockfish landings. Fish. Bull. U.S. 84:409-421.

Shepherd, J.G., and M.D. Nicholson. 1986. Use and abuse of multiplicative models in the analysis of fish catch-at-age data. The Statistician 35:221-227.

Shepherd, J.G., and M.D. Nicholson. 1991. Multiplicative modeling of catch-at-age data, and its application to catch forecasts. J. Cons. Cons. Int. Explor. Mer 47:284-294.

Smith, S.J., and J.J. Maguire. 1983. Estimating the variance of length composition samples. Can. Spec. Publ. Fish. Aquat. Sci. 66:165-170.

Som, R.K. 1973. A manual of sampling techniques. Heinemann Educational Books Ltd., London. 384 pp.

Tomlinson, P.K. 1971. Some sampling problems in fishery work. Biometrics 27:631-641.

A Parametric Bootstrap of Catch-Age Compositions Using the Dirichlet Distribution

Erik H. Williams and Terrance J. Quinn II
University of Alaska Fairbanks, Juneau Center, School of Fisheries and Ocean Sciences, Juneau, Alaska

Abstract

Stock assessment models that incorporate catch or survey age compositions and indices of abundance are commonly used for managing fisheries. Obtaining error estimates for estimated parameters in these models can be difficult. Many models incorporate a bootstrap procedure for estimating errors. In order to perform a bootstrap procedure, age compositions can be modeled by some type of statistical distribution and are often assumed to follow a multinomial process. However, the multinomial distribution is based on the assumption that a simple random sample of ages is taken from the catch, which in practice is rarely true and can lead to underestimation of error. In some applications, a smaller "effective sample size" is substituted for the actual sample size in a subjective manner. We developed an a posteriori method to model observed age compositions with the Dirichlet distribution, using the estimated age compositions from the stock assessment model. A parametric bootstrap procedure using the fitted Dirichlet distribution provided error estimates for estimated stock assessment parameters. We illustrate this technique by application to an age-structured model for Pacific herring (*Clupea pallasi*) from Norton Sound, Alaska. The Dirichlet procedure properly accounted for the variation in observed age compositions compared to those estimated by the stock assessment model. The effective sample sizes from the Dirichlet procedure were lower than the actual sample sizes, suggesting that a comparative multinomial bootstrap procedure underestimated the variance.

Introduction

Fisheries managers commonly utilize stock assessment models which incorporate catch or survey age compositions and indices or estimates of abundance (Hilborn and Walters 1992, Quinn and Deriso, in press). These catch-age or age-structured assessment models are among the best techniques for analysis available to fisheries managers. To fully interpret the parameter estimates from catch-age analysis, estimates of parameter variability are required. Obtaining error estimates of estimated parameters in these models can be awkward and sometimes subjective. Many catch-age models incorporate a bootstrap procedure for estimating errors (Efron 1982, Efron and Tibshirani 1993, Quinn and Deriso, in press).

A nonparametric bootstrap procedure, as applied to catch-age analysis, involves randomly sampling age composition residuals with replacement and adding them to the original estimated values, giving a new set of observed values (Deriso et al. 1985). The new observed values are then used to estimate the parameters of the model resulting in a bootstrap replicate. The standard deviation of the bootstrap replicates is an estimate of the standard error of the estimate (Efron 1982, Efron and Tibshirani 1993). In the case of age composition data, large negative residuals relative to the smallest estimated value results in unrealistic negative "observed" values during the bootstrap procedure.

An alternative method of performing a bootstrap of age composition residuals in catch-age analysis which may be more accurate is to fit a parametric model, usually in the form of some type of statistical distribution, to the observed age compositions. The most commonly employed model for age composition data is the multinomial distribution (Fournier and Archibald 1982, Kimura 1990). However, incorporating a multinomial process into a bootstrap procedure may require manipulation of the sample size and assumes simple random sampling with known sample sizes. Fournier and Archibald (1982) and Methot (1986) suggested that there is variability other than that due to multinomial sampling alone. For this reason, they suggested the sample size for estimating age composition be limited to $n = 400$, even though the actual sample size is often much larger. In order to accurately apply the multinomial sampling distribution to catch-age analysis, the sample size may be adjusted in an ad hoc fashion so that the resulting error matches the expected level of variance.

Alternatively, an objective method is to determine the error structure empirically by examination of the observed and estimated age compositions. We propose to fit a posteriori the Dirichlet distribution to the set of observed age compositions for implementation in a parametric bootstrap procedure for estimation of standard errors of parameter estimates. We apply this technique and compare it to the multinomial bootstrap procedure using an age-structured model for Pacific herring (*Clupea pallasi*) from Norton Sound, Alaska (Williams and Quinn 1998).

Methods

The Dirichlet distribution is a multivariate statistical distribution for a set of random variables that take on values between 0 and 1, and sum to 1

(Evans et al. 1993). The beta distribution, a special case of the Dirichlet distribution, for a single variable (or more precisely, two variables summing to 1), is a flexible statistical distribution with two parameters v and w. It is appropriate for modeling proportions, because it takes on values between 0 and 1 and its density function can assume many shapes: U-shaped ($v = w$, $v < 1$, $w < 1$), J-shaped ($v >$ or $< w$), or unimodal ($v = w$, $v > 1$, $w > 1$). The probability density function of the Dirichlet distribution for a set of k proportions $\{\theta_a,\ a = 1,\ldots,\ k\}$ utilizes the gamma function, $\Gamma(\gamma) = \int_0^\infty \exp(-u)u^{\gamma-1}du$, and is given by

$$\frac{\Gamma(\gamma)}{\prod_{a=1}^{k}\Gamma(\gamma_a)}\prod_{a=1}^{k}\theta_a^{\gamma_a-1}\text{, such that } \sum_{a=1}^{k}\theta_a = 1 \text{ and } \gamma = \sum_{a=1}^{k}\gamma_a \tag{1}$$

We identify γ and $\{\gamma_a,\ a = 1,\ldots,\ k-1\}$ as the parameters of the distribution and obtain γ_k by subtraction as $\gamma_k = \gamma - \sum_{a=1}^{k-1}\gamma_a$. Note that γ_k is not a free parameter of the distribution, as a consequence of the proportions of the distribution summing to 1.

Let $\theta_{a,t}$ = true proportional age composition at age a in year t, $\theta'_{a,t}$ = observed proportional age composition from a sample and $\hat{\theta}_{a,t}$ = estimated proportional age composition from the stock assessment model. Because age composition data are available over several years, the Dirichlet distribution is initially parameterized for each year t. The observed proportions, $\theta'_{a,t}$, are assumed to follow the Dirichlet distribution with parameters γ_t and

$$\gamma_{a,t} = \gamma_t\theta_{a,t},\ a = 1,\ldots,\ k-1, \text{ where } \sum_{a=1}^{k}\theta_{a,t} = 1 \tag{2}$$

(For the last age group, $\gamma_{k,t} = \gamma_t - \sum_{a=1}^{k-1}\gamma_{a,t}$.)

The parameters $\{\gamma_{a,t}\}$ are assumed proportional to $\{\theta_{a,t}\}$ in order to connect the data to common parameters across years. The variance of $\theta'_{a,t}$ is given by

$$\text{var}(\theta'_{a,t}) = \gamma_{a,t}(\gamma_t - \gamma_{a,t}) / [\gamma_t^2(\gamma_t + 1)] = \gamma_t\gamma_{a,t}[\gamma_t(1 - \theta_{a,t})] / [\gamma_t^2(\gamma_t + 1)]$$

$$= \theta_{a,t}(1 - \theta_{a,t}) / (\gamma_t + 1).$$

Thus, the Dirichlet distribution variance is related to the variance of proportions, given by $p(1-p)/n$, with $\gamma_t = n - 1$ and $\theta_{a,t} = p$, where n is the sample size (Zar 1984). Therefore, the parameter γ_t of the Dirichlet distribution represents an "effective sample size." Intuitively, the γ_t parameter is inversely related to the amount of variance in the age composition data.

In this application, the sets of observed and estimated age compositions, $\{\theta'_{a,t}\}$ and $\{\hat{\theta}_{a,t}\}$, are the only information sources available for estimating the variance. The fit of the Dirichlet distribution to the age composition data is based on the probability density function (1). The observed age compositions, $\{\theta'_{a,t}\}$, from the age-structured model are used in a log likelihood function, for k age groups and j years, given by

$$\ln L(\{\gamma_t\} \mid \{\theta'_{a,t}\}) = \sum_{t=1}^{j}\left[\ln\Gamma(\gamma_t) - \sum_{a=1}^{k}\ln\Gamma(\gamma_{a,t}) + \sum_{a=1}^{k}(\gamma_{a,t}-1)\ln(\theta'_{a,t})\right] \tag{3}$$

with $\gamma_{a,t} = \gamma_t\hat{\theta}_{a,t}$ from equation (2). By maximizing (3), estimates of the Dirichlet parameters $\hat{\gamma}_t$ and $\{\hat{\gamma}_{a,t},\ a = 1, \ldots, k\text{-}1\}$ are obtained. (For the last age group, $\hat{\gamma}_{k,t} = \hat{\gamma}_t - \sum_{a=1}^{k-1}\hat{\gamma}_{a,t}$.) The larger the discrepancies between the observed and estimated age compositions, the smaller $\hat{\gamma}_t$ should be. A possibly more parsimonious model may be warranted in which a single effective sample size $\gamma_t = \gamma$ is used.

The parametric bootstrap analysis is performed by using the maximum likelihood estimates, $\{\hat{\gamma}_{a,t}\}$, to generate Dirichlet distributed random age compositions. Some methods of generating Dirichlet random age compositions are to use the $\hat{\gamma}_{a,t}$ Dirichlet parameters to calculate (1) random gamma variates (using a common scale parameter and $\{\hat{\gamma}_{a,t}\}$ for the shape parameters), (2) beta variates (with parameters $\{\hat{\gamma}_{a,t}\}$ and $\{\hat{\gamma}_t - \hat{\gamma}_{a,t}\}$) or (3) chi-squared variates (with $\{2\hat{\gamma}_{a,t}\}$ degrees of freedom) (Evans et al. 1993). In order to follow the Dirichlet distribution, the random age compositions must be adjusted to sum to one by dividing by the sum of the variates across ages for that year.

The Dirichlet random age compositions are then used in place of the observed age compositions in the model and the parameters re-estimated to produce one bootstrap sample. Efron and Tibshirani (1993) suggest a bootstrap sample size of no less than 1,000 for accurate standard errors and confidence intervals. Figure 1 illustrates the algorithm for fitting the Dirichlet distribution to age composition data and the procedure for bootstrapping the model for error estimates.

The nature of this log-likelihood function does not allow for observed age compositions of 0 or 1 to be used, as seen in the log-likelihood equation (3). If not too many observed proportions of zero occur, then essentially they can be ignored in the likelihood estimation of the γ_ts. Note that if $\gamma_{a,t} = 1$, then its likelihood contribution is zero; this is one of the few cases of the Dirichlet distribution with a nonzero probability for $\theta_a = 0$. In the bootstrap replications, there is no problem with zeros, because the $\hat{\gamma}_{a,t}$s are based on estimated proportions $\{\hat{\theta}_{a,t}\}$ that are usually nonzero. If there are quite a few zeros, then it might be best to pool ages to eliminate the zeros.

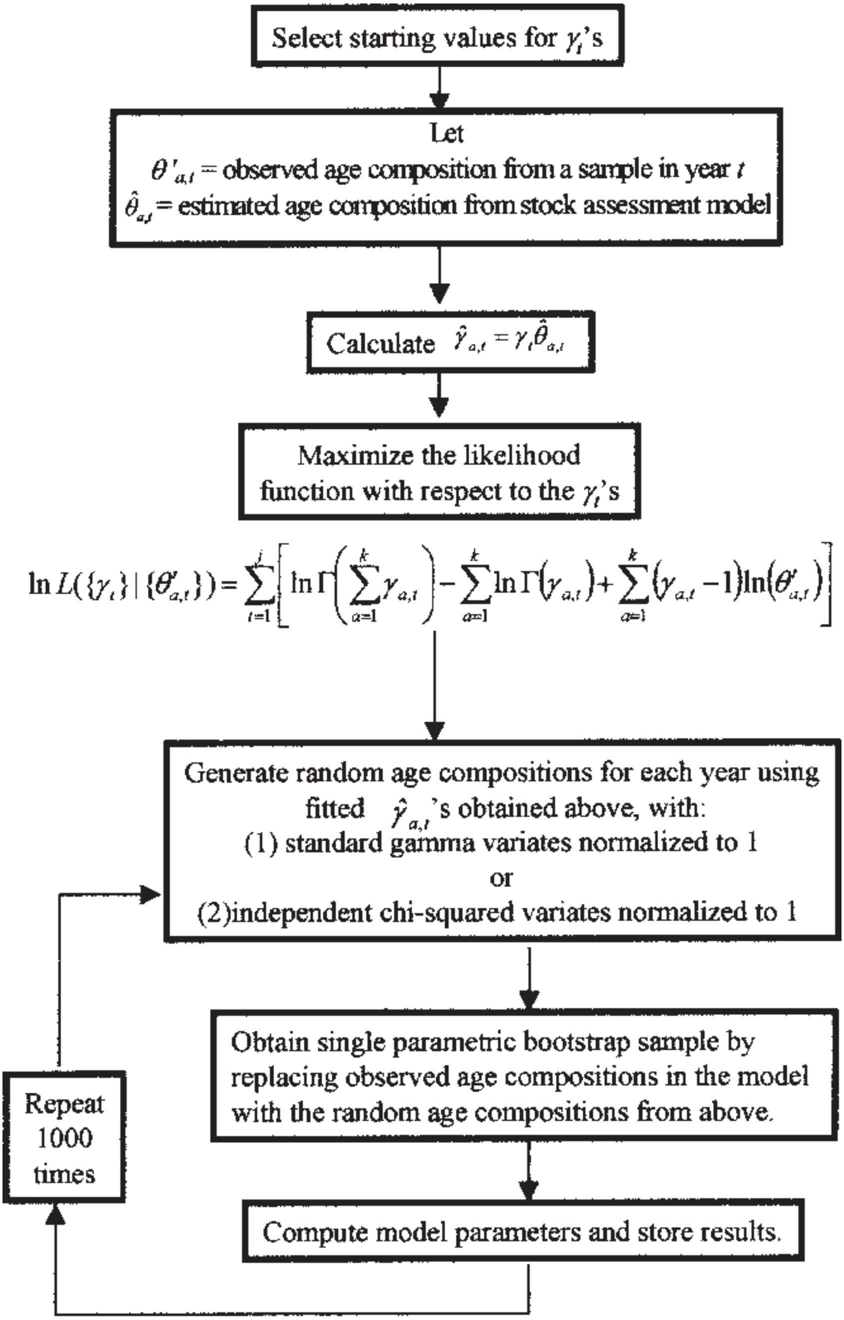


Figure 1. Flow chart of the methodology for applying the Dirichlet distribution bootstrap procedure to an age-structured model with age composition data.

Table 1. Annual sample sizes for age composition of Norton Sound, Alaska herring.

Year	Gillnet fishery	Total run
1981	1,084	4,244
1982	637	2,163
1983	994	3,570
1984	671	2,056
1985	1,265	3,550
1986	927	2,981
1987	408	1,595
1988	388	1,100
1989	622	1,593
1990	555	2,586
1991	914	3,704
1992	0[a]	1,197
1993	1,410	3,968
1994	789	2,626
1995	1,645	6,159
1996	1,372	4,231

[a] No fishery occurred in 1992.

Application to Pacific Herring

We used an age-structured model for Pacific herring from Norton Sound, Alaska developed by Williams and Quinn (1998). This model incorporates catch and total-run age composition data and aerial survey estimates of abundance for years 1981 to 1996 and ages 3 to 10+, similar to other age-structured models for Pacific herring (Funk et al. 1992, Brannian et al. 1993, Yuen et al. 1994). Sample sizes for the gillnet and total run age compositions were usually large, ranging from about 400 to over 6,000 (Table 1). A method for determining the variability associated with age-structured model estimates for Pacific herring could prove useful for herring management, because no variance estimates are currently calculated by the Alaska Department of Fish and Game (F. Funk, Alaska Department of Fish and Game, personal communication).

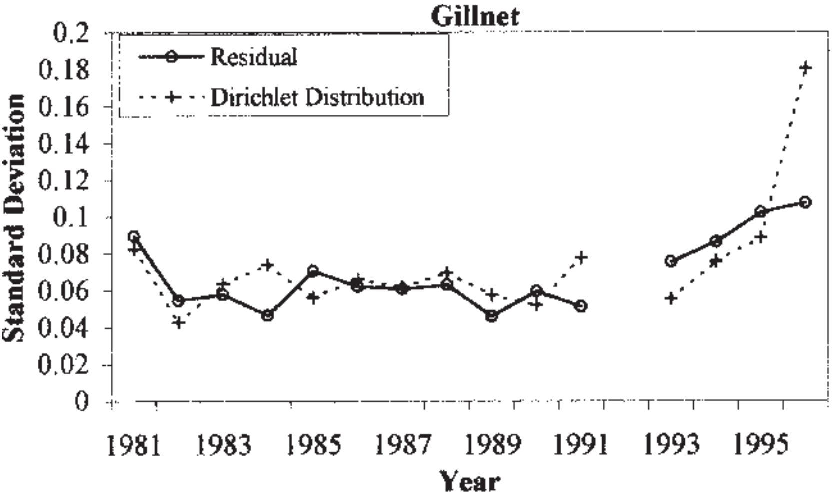
We fitted two Dirichlet distributions to the observed catch and total-run age-composition data: one with constant γ for each age composition data set, the other with annual γ_t. Estimates were obtained by maximizing the log likelihood function (3) in an Excel spreadsheet (available from the authors), which utilizes a quasi-Newton optimizer. Estimates of γ and γ_t for the gillnet and total run age compositions are given in Table 2, along with

Table 2. Maximum likelihood estimates γ and γ_t for gillnet and total run age compositions, along with maximum log-likelihood values.

	Gillnet	Total run
$\hat{\gamma}$	21.57	26.33
$\hat{\gamma}_t$		
1981	7.39	17.11
1982	41.38	17.23
1983	21.52	24.95
1984	22.42	45.24
1985	29.78	25.17
1986	11.66	42.92
1987	22.13	19.93
1988	19.61	18.87
1989	22.12	91.77
1990	34.07	80.41
1991	86.22	20.27
1992		18.71
1993	279.01	54.35
1994	129.28	11.57
1995	150.37	23.74
1996	11.51	20.23
Log likelihood		
$\hat{\gamma}$	118.11	177.42
$\hat{\gamma}_t$	137.59	182.00

log likelihood statistics. The fit using annual γ_t parameters for the gillnet data set was significantly better than using a constant γ (likelihood ratio test: $\chi^2 = 38.97$, $P < 0.001$), while the annual γ_t parameterization for the total run data set was not significantly different ($\chi^2 = 9.16$, $P = 0.87$). Note that the estimates of γ in Table 2 are smaller than the sample sizes in Table 1, showing that effective sample size is smaller than actual sample size.

In order to evaluate the annual γ_t parameterization fits of the Dirichlet distribution, we calculated the empirical standard deviation for each year in each data set. These values are compared to the analytical standard deviation based on the Dirichlet variance formulae given above. Fig. 2 indicates the Dirichlet standard deviation based on the computed effective sample sizes is comparable to the empirical standard deviation from the age-structured model.



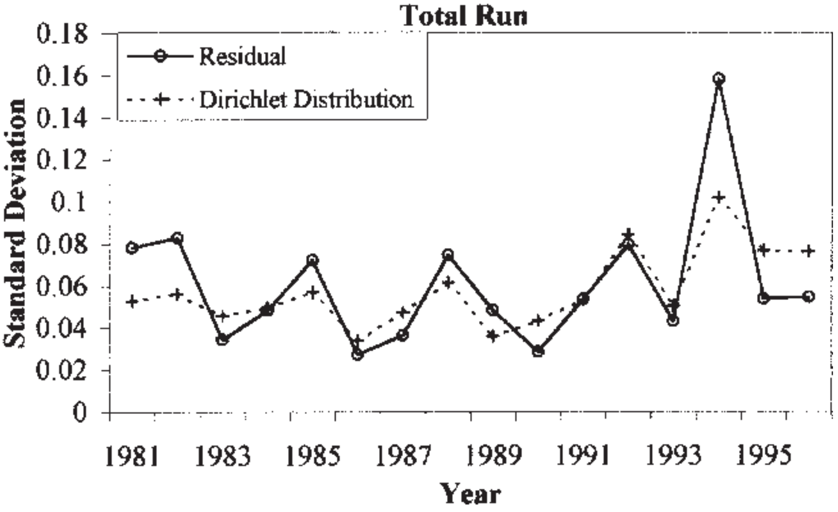


Figure 2. Standard deviations from the residuals and annual γ_t case of the Dirichlet distribution fit of the Norton Sound herring age-structured model.

As mentioned earlier, zero-valued observed age compositions cannot be estimated by the Dirichlet distribution. For this application there were some unrepresented (zero) ages, due to a closed fishery in 1992 and selective fishing pressure on older fish. In the bootstrap procedure these values were left at zero. A standard non-parametric bootstrap procedure was used for the aerial survey biomass data. A bootstrap sample of 1,000 was chosen for this exercise.

For comparison to the Dirichlet distribution bootstrap procedure, a multinomial bootstrap procedure was performed with the Norton Sound herring age-structured model using the sample sizes in Table 1. The objective function of the Norton Sound herring model is a simple combination of sums of squares, which does not follow a multinomial process. The multinomial bootstrap procedure was only intended as a comparison to the Dirichlet distribution bootstrap procedure and therefore changing the objective function is not warranted.

Bootstrapping Results

One of the 1,000 bootstrap samples for the annual γ_t case of the Dirichlet distribution failed to converge by settling on unrealistic, negative valued estimates. This failure suggests some evidence of model instability but is inconsequential to the final results, so we ignored this replication. All the bootstrap samples for the constant γ case of the Dirichlet distribution and for the multinomial distribution appeared to be realistic (no zero or negative valued estimates).

From the bootstrap samples the coefficient of variation (*cv*, bootstrap standard deviation divided by original estimate), a relative measure of uncertainty, was calculated for estimated total biomass (Fig. 3), pre-fishery abundance estimates for recruitment age herring (age 3) (Fig. 4), and selectivity and catchability parameters in the model (Fig. 5). In general, the multinomial distribution resulted in lower *cv* estimates as compared to the two Dirichlet distribution estimates, suggesting that use of the multinomial understates actual variability. The error estimates for the total biomass and aerial survey parameters are similar for all three bootstrap procedures, probably because these quantities are determined more from aerial survey information than from age composition information. The *cv*s for the fishery and sampling gear selectivities indicate larger *cv*s for the two Dirichlet bootstrap procedures, particularly for the sampling gear selectivity parameters for the last six years. The recruitment *cv*s indicate the difference among the three bootstrap procedures increases in more recent years. The annual γ_t case for the Dirichlet bootstrap procedure seems to result in the highest *cv*s in the most recent years.

Discussion

The residuals from an age-structured model include both measurement and process error. The true sample size in a multinomial distribution boot-

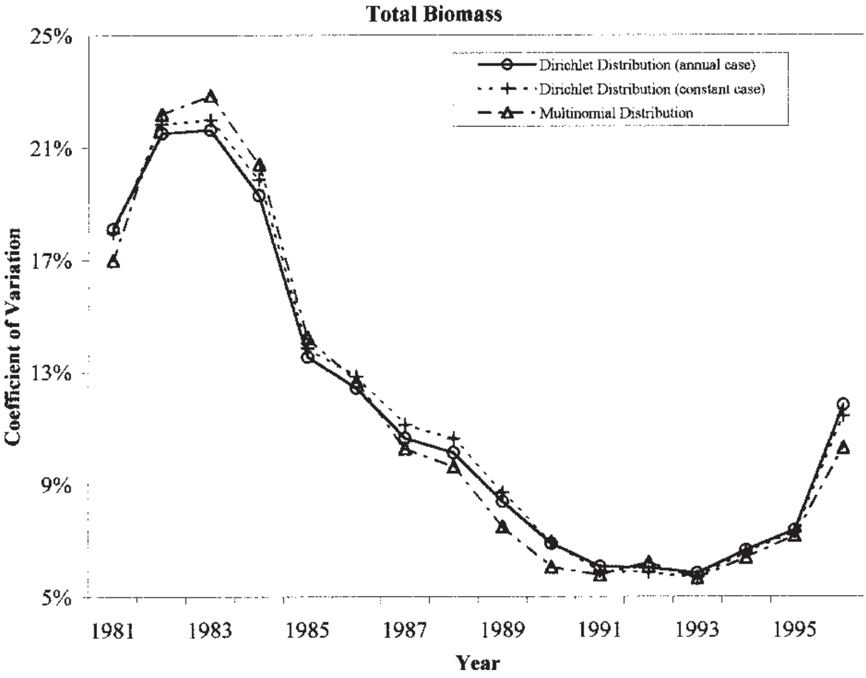


Figure 3. Annual percent error estimates for total biomass from 1,000 bootstrap samples of three different procedures for the Norton Sound herring age-structured model.

strap procedure is limited to representing the measurement error in the model under simple random sampling and has been suggested to be an inappropriate representation of the true estimation error (Fournier and Archibald 1982, Methot 1986). The multinomial bootstrap error estimates in this application are useful for comparison to the Dirichlet distribution bootstrap error estimates. Since the Dirichlet distribution fit is based on the observed and estimated age compositions from the model, the resulting *cv*s for most cases were higher than those for the multinomial. This result indicates that the Dirichlet distribution is capturing variation in the estimates that is unaccounted for by the multinomial distribution. Otherwise the estimates of effective sample size would have been much closer to the actual ones. The multivariate logistic model is an alternative approach for incorporating greater variability into stock assessment models (Schnute and Richards 1995).

The annual γ_t case produced higher *cv*s than the constant γ case, unexpectedly, since the maximum likelihood result in Table 2 indicated a better fit for the γ_t case. The reason for this is probably due to the lower

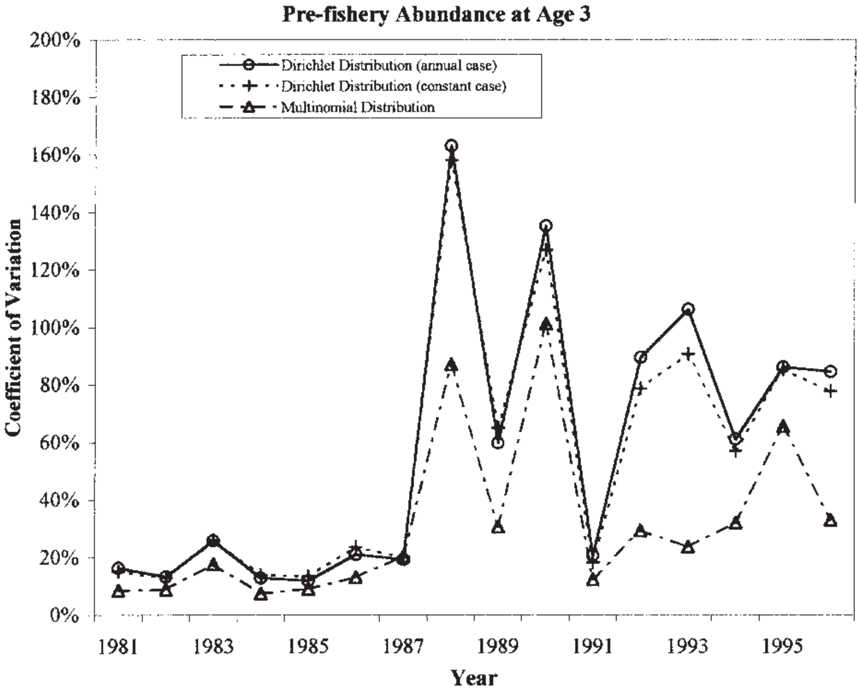


Figure 4. Annual percent error estimates for pre-fishery abundance at age 3, the age of recruitment, from 1,000 bootstrap samples of three different procedures for the Norton Sound herring age-structured model.

effective sample sizes in the last three years for the total run data and the last year for the gillnet data for the annual γ_t case relative to the constant γ case. Since the nature of age-structured analysis is to track cohorts through time, poor fits in the last few years will affect estimates in earlier years. So, even though the constant γ case involves less parameterization, it resulted in lower *cv*s in later years, because it smoothed over the high variability in the last years of the age-structured model.

Coefficients of variation for total biomass and the aerial survey catchability parameters from the Dirichlet method were much closer to the multinomial estimates as compared to the abundance and selectivity parameter estimates. Since the total biomass estimates and catchability parameters depend primarily on the aerial survey abundance data, the similarity of *cv*s for all three bootstrap methods suggests these portions of the model are relatively unaffected by changes in sample size for the age composition data. The similarity of *cv*s is consistent with the aerial survey abundance component in the all three bootstrap procedures being computed by the same non-parametric procedure. This result suggests

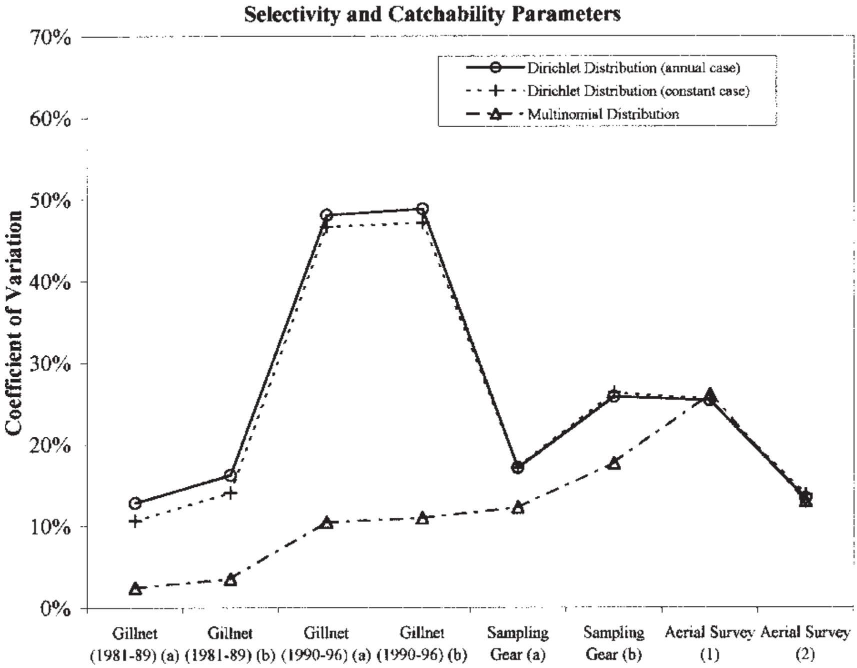


Figure 5. Percent error estimates for gillnet fishery, sampling gear, and aerial surveyability parameters from 1,000 bootstrap samples of three different procedures for the Norton Sound herring age-structured model.

that the recommendation to limit age composition samples to 400 observations (Fournier and Archibald 1982, Methot 1986) in order to prevent age composition from dominating the likelihood does not have a strong rationale. Indeed, since effective sample size can be much smaller than actual sample size, a re-evaluation of sample size requirements in age-structured models is needed.

Overall, the Dirichlet distribution appears to be a statistical distribution which allows for fairly easy and objective bootstrap estimates of error for stock assessment models which utilize age composition data. It provides a means to utilize the observed variation in the model without having to resort to making unrealistic assumptions about the error distribution. An implicit assumption in this approach is that the underlying stock assessment model is correct. If not, the estimated proportions $\{\hat{\theta}_{a,t}\}$ may be biased, which would lead to bias in the Dirichlet parameter estimates.

We also note that this approach is an approximation to a true bootstrap, in which the original objective function for age composition would be based on the Dirichlet distribution. Equations (1) and (3) for the Dirichlet distribution would then be incorporated directly into the objective function of the stock assessment model. Dirichlet parameters would be estimated jointly with the other model parameters, which would avoid the necessity of fitting the Dirichlet distribution a posteriori as we did. Further study is necessary to determine if the additional model complexity presents difficulties in estimation, but such an approach would lend greater consistency to the stock assessment. We did not do so here because we were looking for a quick and simple procedure to generate standard errors after being satisfied with the basic stock assessment results.

Indeed, the major strength of bootstrap procedures in general is their ability to provide measures of uncertainty for all key variables of interest in a straightforward manner and with a minimum of assumptions. Our procedure should be generally applicable to stock assessment models with observed and estimated proportions.

Finally, the concept of effective sample size need not invoke the Dirichlet distribution. An empirical estimate of effective sample size can be made from the observed and estimated age compositions for a given year and gear type (McAllister and Ianelli 1997). If θ'_a and $\hat{\theta}_a$ are the observed and estimated age compositions for a given year (omitting the time subscript for clarity), then an estimate of the mean squared error is $\sum_a^k (\theta'_a - \hat{\theta}_a)^2 / k$. The average estimated variance of observed proportions from the multinomial distribution is $\sum_a^k \theta'_a (1 - \theta'_a) / (kn)$, where n is sample size. (Alternatively, the estimated $\hat{\theta}_a$s could be used.) By equating these two quantities and solving for n, the estimated effective sample size is $n = \sum_a \theta'_a (1 - \theta'_a) / \sum_a (\theta'_a - \hat{\theta}_a)^2$. A bootstrap procedure could be performed utilizing these effective n values across years with the multinomial distribution. However, the approach utilizing the Dirichlet distribution is more rigorous, and is based on simultaneous analysis of all the information. The Dirichlet approach tends to give lower effective sample sizes, because it explicitly accounts for the extra variance due to the use of observed or estimated proportions.

Acknowledgments

This publication is the result of research sponsored by Alaska Sea Grant with funds from the National Oceanic and Atmospheric Administration, Office of Sea Grant, Department of Commerce, under grant no. NA90AA-D-SG066, project number R/07-16, from the University of Alaska with funds appropriated by the state, the Alaska Department of Fish and Game, Herring Natal Habitats project under *Exxon Valdez* Oil Spill restoration funds, and a Rasmuson Fisheries Research Center fellowship. The U.S. Government is authorized to produce and distribute reprints for government purposes, notwithstanding any copyright notation that may appear hereon.

We would like to thank Dr. Grant G. Thompson for an excellent review of the beta version of this paper and Dr. Jim Ianelli for helpful suggestions.

Fritz Funk and Helen Hamner provided the Norton Sound herring fishery and sampling data upon which the age-structured model is based.

References

Brannian, L.K., K.A. Rowell, and F. Funk. 1993. Forecast of the Pacific herring biomass in Togiak District, Bristol Bay, 1993. Alaska Department of Fish and Game, Regional Information Report 2A93-42, Anchorage.

Deriso, R.B., T.J. Quinn II, and P.R. Neal. 1985. Catch-age analysis with auxiliary information. Can. J. Fish. Aquat. Sci. 42:815-824.

Efron, B. 1982. The Jacknife, the Bootstrap and other resampling plans. SIAM Conf. Ser. 38.

Efron, B., and R.J. Tibshirani. 1993. An Introduction to the Bootstrap. Chapman and Hall, New York. 436 pp.

Evans, M., N. Hastings, and B. Peacock. 1993. Statistical distributions. John Wiley and Sons, Inc., New York. 170 pp.

Fournier, D., and C.P. Archibald. 1982. A general theory for analyzing catch at age data. Can. J. Fish. Aquat. Sci. 39:1195-1207.

Funk, F., L.K. Brannian, and K.A. Rowell. 1992. Age-structured assessment of the Togiak herring stock, 1978-1992, and preliminary forecast of abundance for 1993. Alaska Department of Fish and Game, Regional Information Report 5J92-11, Juneau.

Hilborn, R., and C.J. Walters. 1992. Quantitative fisheries stock assessment: choice, dynamics, and uncertainty. Chapman and Hall, New York. 570 pp.

Kimura, D.K. 1990. Approaches to age-structured separable sequential population analysis. Can. J. Fish. Aquat. Sci. 47:2364-2374.

McAllister, M.K., and J.N. Ianelli. 1997. Bayesian stock assessment using catch-age data and the sampling-importance resampling algorithm. Can. J. Fish. Aquat. Sci. 54:284-300.

Methot, R.D. 1986. Synthetic estimates of historical abundance and mortality for northern anchovy, (*Engraulis mordax*). Southwest Fish. Center Administrative Rep. LJ 86-29. 85 pp.

Quinn II, T.J., and R.B. Deriso. In press. Quantitative fish dynamics. Oxford Univ. Press, New York.

Schnute, J.T., and L.J. Richards. 1995. The influence of error on population estimates from catch-age models. Can. J. Fish. Aquat. Sci. 52:2063-2077.

Williams, E.H., and T.J. Quinn II. 1998. Age-structured analysis of Pacific herring from Norton Sound, Alaska. Alaska Fish. Res. Bull. 4(2):87-109.

Yuen, H.J., L.K. Brannian, and F. Funk. 1994. Forecast of the Kamishak herring stock in 1994. Alaska Department of Fish and Game, Regional Information Report 2A94-12, Anchorage.

Zar, J.H. 1984. Biostatistical analysis. Prentice-Hall, Inc., Englewood Cliffs, NJ. 718 pp.

Some Intrinsic Limitations of Sample Variances in Stock Assessment Models

Clay E. Porch
National Marine Fisheries Service, Southeast Fisheries Science Center, Miami, Florida

Abstract

Most modern stock assessment algorithms incorporate multiple catch and abundance estimates, with differing levels of precision. Often these estimates are weighted by their corresponding variances, but estimates of variance are even less precise than estimates of expectation. For this reason, transformations and weighting schemes based on a presumed variance to expectation relationship have often been suggested. Since the true variance to expectation relationship is seldom known, it is usually inferred from regressions of sample variance on sample mean. Unfortunately, the sample variances of non-negative variates like catch per unit effort lie within bounds that are functions of the sample size, sample mean, and maximum possible catch (density). Furthermore, the sample mean and sample variance are self-correlated because the sample mean appears in the expression for sample variance. This paper demonstrates that these, and other mathematical artifices, can force a strong spurious correlation between the sample variance and sample mean regardless of the statistical distribution of the data. Therefore, plots of sample variance against sample mean are unlikely to reveal the true variance to mean relationship for any given population. Moreover, stock assessments that use sample variance to weight the input observations will tend to be biased. Underestimates will tend to receive too much weight and overestimates too little, particularly if the resource in question is uncommon and the sample sizes are small.

Introduction

Most fishery stock assessments depend on estimates of historical abundance derived from resource surveys and samples of the fishery catch. Such data are often expensive to collect, but because of their highly variable nature, tend to yield rather imprecise estimates unless the number of sample observations (n) is large. For this reason, it is important to develop efficient sampling programs that maximize precision under the limited resources available. Of course historical data cannot be remedied by changes in future sampling. One common solution has been to incorporate multiple data sources into the stock assessment, thereby increasing the effective n. In that case the different estimates should be weighted according to their precision, either implicitly by a suitable transformation or explicitly by modeling the variances.

Whether one wishes to develop an efficient sampling scheme or properly weight the estimates in a stock assessment model, a central issue is estimating variance. It is well known that if n observations $\{x_1, x_2, \ldots, x_n\}$ are sampled at random from an infinite population, then the sample mean ($\bar{x}$) and sample variance,

$$s^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})}{n-1},$$

are unbiased estimators of the population mean μ and population variance σ^2. The variance of $\bar{x}$ is σ^2/n; however, σ^2 is seldom known so s^2 is often used in its place. Unfortunately, s^2 is very inefficient (Table 1). More efficient estimators of σ^2 can be derived if the statistical distribution of the population is known, but their performance may be questionable where the data depart from the assumed distribution (see Smith 1990, Myers and Pepin 1991, Pennington 1996). Theoretical arguments supporting one distribution or another abound in the fisheries literature (Taylor 1953, Houser and Dunn 1967, Pella and Psaropulos 1975, Mangel and Beder 1985, Deriso and Parma 1987, Porch and Fox 1990, Pennington 1996), but none apply universally. In any case, neither s^2 nor its distribution-specific counterparts will in general be a useful proxy for σ^2 unless $n > 100$.

An alternative method of specifying σ^2 has its roots in the work of Curtis (1943) and Bartlett (1947), who showed that transformations to homogenize variance can be determined from the relationship between σ^2 and μ. Inasmuch as the σ^2 to μ relationship is seldom known, it is frequently inferred from regressions of s^2 on $\bar{x}$ (Green 1979, Taylor 1980, Taylor et al. 1983). This practice is quite common in fisheries work because s^2 and $\bar{x}$ are often found to be highly correlated (Taylor 1953, Small and Downham 1985, Porch and Fox 1990, Lester et al. 1991, Cyr et al.

Table 1. Coefficients of variation for $\bar{x}$ and s^2 with selected probability densities and sample sizes (each computed from 2,000 replicate samples drawn at random with replacement).

Density		Sample size (n)			
(μ, σ^2)	Statistic	5	10	20	100
Uniform	$\bar{x}$	3	2	2	1
(100, 57)	s^2	51	32	21	9
Normal	$\bar{x}$	16	11	8	4
(20, 50)	s^2	72	46	33	14
Lognormal	$\bar{x}$	24	18	12	5
(10, 30)	s^2	132	101	68	29
Poisson	$\bar{x}$	14	10	7	3
(10, 10)	s^2	71	47	33	14
Neg. binomial	$\bar{x}$	46	31	22	10
(5, 25)	s^2	125	82	61	26

1992, Warwick and Clarke 1993, Welch and Ishida 1993, Dong and Restrepo 1996). However, the relationship between s^2 and $\bar{x}$ is largely spurious.

Downing (1989) pointed out that the s^2 of non-negative variates must be less than or equal to $n\bar{x}^2$. Tokeshi (1995) argued that "ecological common sense" also dictates $s^2 \geq \bar{x}$ and that the restricted sampling space imposed by these two constraints greatly facilitates fits of regression curves to any set of (s^2, $\bar{x}$) pairs. Tokeshi's presumption of $s^2 \geq \bar{x}$ is misleading inasmuch as values less than $\bar{x}$ are quite feasible even if on average $s^2 > \bar{x}$, but his basic premise is correct. Moreover, s^2 and $\bar{x}$ are self-correlated because $\bar{x}$ appears in the computational expression for s^2. Therefore, one might expect strong correlations between s^2 and $\bar{x}$ even when there is little or no relationship between σ^2 and μ.

The Mathematical Artifices

Constraints on Sample Variance

Consider n non-negative numbers $\{x_1, x_2, \ldots, x_n\}$ with mean $\bar{x}$. The maximum possible s^2 for a given $\bar{x}$ is achieved when one number equals $n\bar{x}$ and the rest are zero, i.e.,

$$s^2_{MAX} = \frac{(n\bar{x} - \bar{x})^2 + (n-1)(0-\bar{x})^2}{n-1} = n\bar{x}^2 \qquad (1)$$

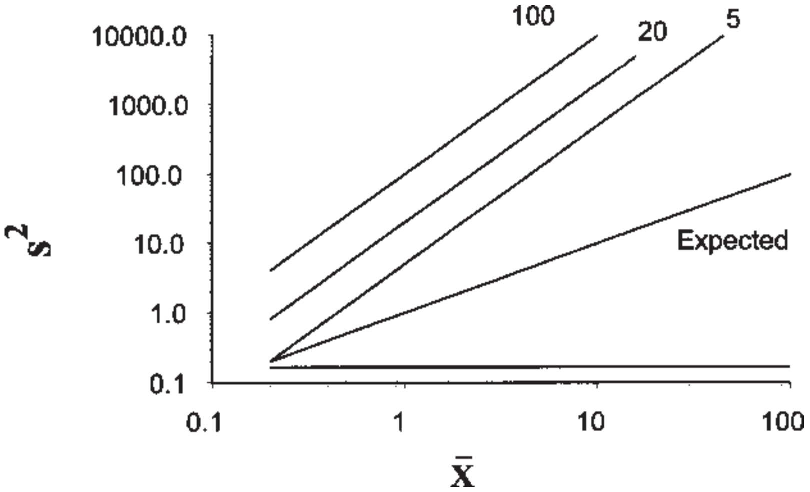


Figure 1. The feasible region of variance-mean regressions for sample sizes of 5, 20, and 100 in log-space. The bottom line is the average value of s^2_{MAX} (⅙) and the line labeled "Expected" is the expected value of s^2 given $\bar{x}$ for sampling units (trips) with the same odds of catching fish (equation 3).

The minimum s^2 is zero when all the numbers have the same value, but for discrete variates this is only possible when $\bar{x}$ is an integer. Otherwise the minimum s^2 is

$$s^2_{MIN} = (\bar{x} - i - 1)(i - \bar{x})\frac{n}{n-1} \tag{2}$$

where i is the integer part of $\bar{x}$ (see the Appendix). Notice that neither equation (1) nor (2) depend at all on the population parameters σ^2 and μ. Thus, if $\sigma^2 > s^2_{MAX}$, then the value of s^2 from that particular sample must underestimate σ^2. Likewise, if $\sigma^2 < s^2_{MIN}$, then s^2 must overestimate σ^2.

The expected relationship of s^2 to $\bar{x}$ depends on the underlying pattern of the population, but lies somewhere between s^2_{MIN} and s^2_{MAX} (Fig. 1). Suppose n trips combined to catch $T(=n\bar{x})$ fish and that each trip had the same (constant) ability to catch fish. Then the number caught per trip ought to be binomial distributed with probability $p = 1/n$ that a given fish will be assigned to a given trip. The expected variance of the binomial distribution given T is $Tp(1 - p)$. Substituting $1/n$ for p we obtain the expected variance given $\bar{x}$:

$$E[s^2 \mid \bar{x}, n] = \frac{n-1}{n}\bar{x}. \tag{3}$$

Of course different fishing units can have different abilities to catch fish, in which case $E[s^2 \mid \bar{x}, n]$ can exceed equation (3). In general however, if $\sigma^2 > E[s^2 \mid \bar{x}, n]$, then the value of s^2 for that particular sample will probably (but not necessarily) underestimate σ^2.

In practice, the s^2 of continuous variates may take on any value between s^2_{MIN} and s^2_{MAX}, some values being more likely than others depending on the underlying pattern of the data. The s^2 of discrete variates, however, can have only a finite number of values for a given $\bar{x}$ and n. Consider again the case where n fishing trips were sampled from a population of fishing trips that are equally able to catch fish. The probability of obtaining a particular distribution of catch per trip $\{x_1, x_2, \ldots, x_n\}$ in any order is

$$P(x_1, x_2, \ldots, x_n) = \frac{T!}{x_1!x_2!\ldots x_n!}\frac{n!}{n_0!n_1!\ldots n_T!}N^{-T} \tag{4}$$

where n_x is the number of variates with value x and $T = \sum x_i$ (see pages 38-40 in Feller 1968). The probability distribution (pdf) of s^2 is obtained from equation (4) by stepping through all possible distributions and adding the probabilities for those with the same variances. As Fig. 2 demonstrates, the pdf described by equation (4) is dominated by fine structure when n is small. Some of the possible values of s^2 will rarely be observed despite the fact that both smaller and larger values may be common because there happen to be very few combinations that lead to them (a statistical artifact common to ratios of discrete quantities, see Johnston et al. 1995). The pdf becomes smoother and more values of s^2 become possible as n increases, but the odds of s^2 being very much greater than its expected value, $\bar{x}(n-1)/n$, remain small. Thus, if $\sigma^2 >> \bar{x}(n-1)/n$, then s^2 is likely to underestimate σ^2 even when n is large. As $\bar{x}$ increases, however, the pdf of s^2 becomes less skewed, making it more likely that s^2 will attain a value near σ^2 and therefore less likely that s^2 will underestimate σ^2. Of course the pdf in Fig. 2 reflects the fact that the catching power of each trip was the same. Had the catching power differed, the expectation of s^2 would increase and the spread of the pdf would be different.

Self-Correlation between Sample Variance and Sample Mean

As mentioned earlier, s^2 and $\bar{x}$ are self-correlated because $\bar{x}$ appears in the formula for s^2. Kenney (1982) develops a general expression for the expected self-correlation between two variables, which for s^2 and $\bar{x}$ reduces to

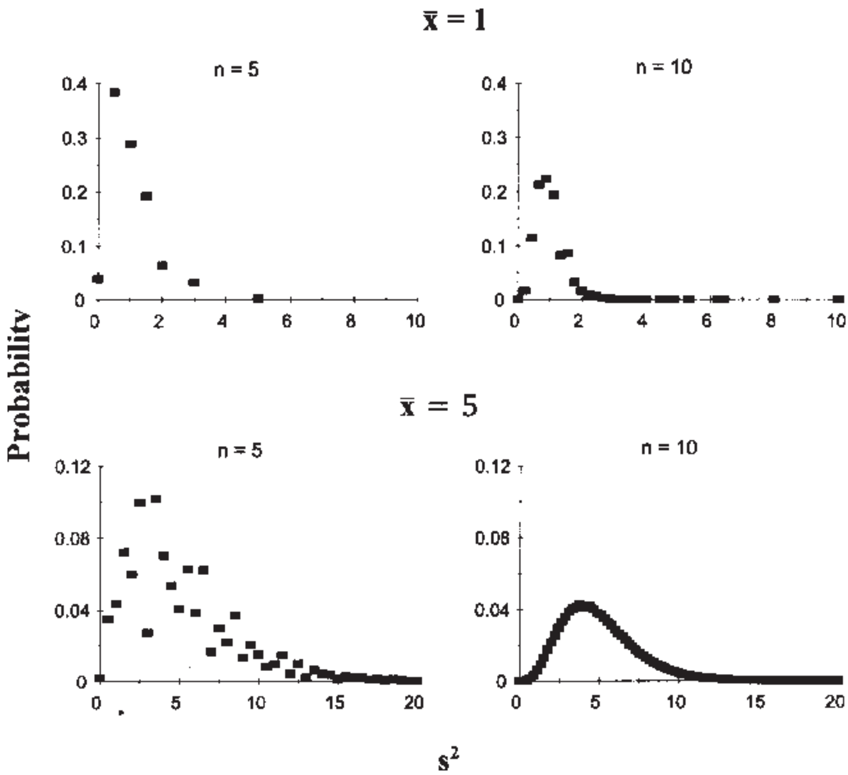


Figure 2. The pdf of the variances of discrete variates when the sampling units (trips) have the same odds of catching fish.

$$r_{s^2,\bar{x}} = \frac{r_{p\bar{x}}(\sigma_p/\sigma_q) - r_{q\bar{x}}}{\sqrt{1+(\sigma_p/\sigma_q)^2 - 2r_{pq}(\sigma_p/\sigma_q)}}, \tag{5}$$

where $s^2 = p - q$, $p = \sum x^2/(n-1)$ and $q = n\bar{x}^2/(n-1)$. The correlation between q and $\bar{x}$ is always close to 1 for non-negative variates. If it were possible for p and $\bar{x}$ to be uncorrelated ($r_{p\bar{x}} = 0$), then also $r_{pq} = 0$ and equation (5) would reduce to

$$r_{s^2,\bar{x}} = \frac{-r_{q\bar{x}}}{\sqrt{1+(\sigma_p/\sigma_q)^2}} \approx \frac{-1}{\sqrt{1+(\sigma_p/\sigma_q)^2}}. \tag{6}$$

Table 2. Linear correlations of s^2 and $p = \sum x^2/(n-1)$ with $\bar{x}$ for selected probability densities and sample sizes (computed from 2,000 replicate samples drawn at random with replacement).

Density (μ, σ^2)	Statistic	Sample size (n) 2	5	20	100
Lognormal	p	0.87	0.82	0.80	0.79
(5, 25)	s^2	0.74	0.72	0.67	0.65
Lognormal	p	0.99	0.99	0.99	0.99
(100, 500)	s^2	0.35	0.39	0.38	0.43
Neg. bin.	p	0.90	0.89	0.89	0.90
(5, 25)	s^2	0.62	0.68	0.66	0.70
Neg. bin.	p	0.99	0.99	0.99	0.99
(100, 500)	s^2	0.17	0.25	0.24	0.27
Poisson	p	0.97	0.96	0.96	0.96
(5, 5)	s^2	0.21	0.25	0.30	0.29
Poisson	p	1.00	1.00	1.00	1.00
(100, 100)	s^2	0.01	0.02	0.04	0.04

This implies that including $\bar{x}$ on both sides of the equation imparts a strong negative correlation unless $\sigma_p >> \sigma_q$ (which depends on the skewness of the distribution of x).

Some Numerical Illustrations

This section uses simulated data to illustrate how the mathematical artifices discussed above can influence one's perception of the variance to mean relationship.

Example 1: Unchanging Population (μ and σ^2 Constant)

Six data sets of 2,000 (s^2, $\bar{x}$) pairs each were simulated by randomly drawing n observations with replacement from three types of statistical distributions (Poisson, negative binomial and lognormal) and two parameter sets ($\mu = 5$, $\sigma^2 = 25$ and $\mu = 100$, $\sigma^2 = 500$). The linear correlation between $p = \sum x^2/(n-1)$ and $\bar{x}$ was 0.8 or larger in all six cases, but the correlation between s^2 and $\bar{x}$ depended on the distribution (Table 2, Fig. 3). Sample size (n) was relatively unimportant; however, the magnitudes of μ and σ^2 were very important. Samples drawn from the distributions with $\mu = 5$ had much higher (s^2, $\bar{x}$) correlations than those drawn from the distributions with $\mu = 100$. Likewise, the samples drawn from distributions with high σ^2

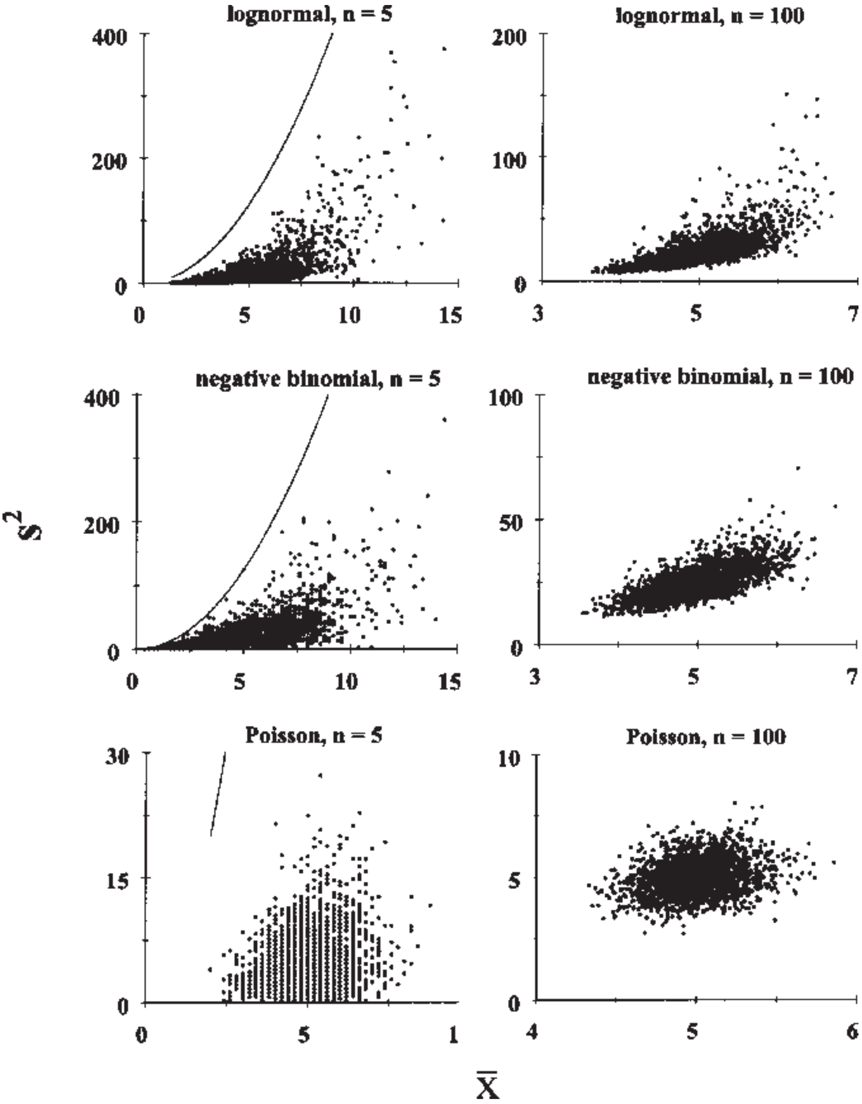


Figure 3. *Scatter plots of (s^2, $\overline{x}$) pairs from samples of 5 or 100 drawn from the lognormal and negative binomial densities with mean 5 and variance 25 and the Poisson density with mean 5 and variance 5. The adjacent curves represent* s^2_{MAX} .

(lognormal and negative binomial) had much higher (s^2, $\bar{x}$) correlations than those drawn from the Poisson distribution (with $\sigma^2 = \mu$). Not surprisingly, the power function ($\log_e s^2 = \log_e a\bar{x}^b$) fit the data generated from the lognormal and negative binomial distributions with $\mu = 5$ and $\sigma^2 = 25$ quite well, yielding coefficients of determination (r^2) in excess of 0.5, but provided a very poor fit to the data generated from the remaining distributions.

Example 2: Changing μ, but Constant σ^2

Now consider a situation where samples are drawn from a stock that has changed substantially over time, but for which the variance remains constant. Data sets consisting of 20 (s^2, $\bar{x}$) pairs were constructed by randomly choosing 20 values (Y_j) between 0 and 10 and then sampling n times with replacement from lognormal distributions with $\sigma^2 = 25$ and $\mu = Y_j$. The 20 sample (s^2, $\bar{x}$) pairs in each simulated data set were then fitted with the power function. The average r^2 from 1,000 such data sets were surprisingly high (0.59, 0.78, 0.75 and 0.69 for $n = 2$, 5, 10, and 20). The values of s^2 generally fluctuated around σ^2 (25) when $\bar{x}$ was large, but severely underestimated σ^2 when $\bar{x}$ was small (Fig. 4).

Example 3: All Possible Sample Variances Equally Likely

In this example s^2 was allowed to assume any value between the minimum and maximum feasible values (equations 1 and 2) with equal probability. This is equivalent to a situation where the underlying pattern of a continuous variate is completely inconsistent. Artificial data sets were generated by randomly selecting 20 $\bar{x}$ values and assigning to each $\bar{x}$ a corresponding s^2 drawn at random from the feasible region. The power function was then fitted to each data set by least squares after taking the logarithms of s^2 and $\bar{x}$. The median r^2 was about 0.8 and values under 0.5 were observed for only 10% of the simulated data sets. Sample size had little impact on the r^2 because the possible variances were uniformly distributed across the sampling space and the sampling space for $n = 5$ is similar to that for $n = 100$ on a logarithmic scale (Fig. 1). Thus, it is clear that the restricted feasible region for (s^2, $\bar{x}$) by itself can account for a large fraction of the strong correlations between the two statistics.

Discussion

Although s^2 is an unbiased estimator of σ^2, the value of s^2 conditional on $\bar{x}$ is not. The statistics s^2 and $\bar{x}$ are spuriously correlated because $\bar{x}$ appears in the computation of s^2 and because the probability space of s^2 is restricted, especially when n or $\bar{x}$ are small. As the examples show, it is possible to find a strong correlation between s^2 and $\bar{x}$ even when σ^2 and

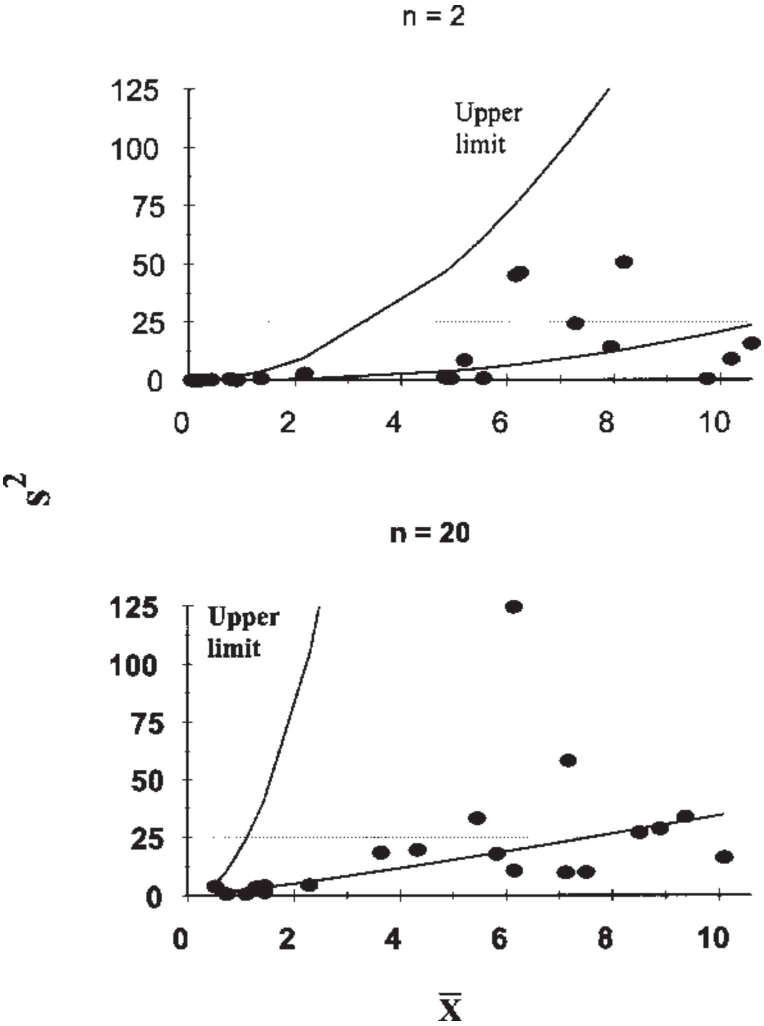


Figure 4. Typical fits of the power function to the simulated data (dots) in example 2. The dashed line indicates the true variance (25).

μ are unrelated. Similarly, where a relationship between σ^2 and μ does exist, regressions of s^2 and $\bar{x}$ may not accurately reflect it.

Other factors may further confound the s^2 to $\bar{x}$ relationship. Downing (1986) showed empirically that estimates of the exponent parameter of the power function vary systematically with the number of (s^2, $\bar{x}$) pairs used as well as n. Additional bias may occur if n varies among the (s^2, $\bar{x}$) pairs since the pdf of s^2 varies with n. The method used to fit the data can also bias the results. Ross (1990) pointed out that the logarithm of s^2 tends to be biased low, particularly for highly skewed data with high $\bar{x}$ and low n (<30). This tends to cause the curvature parameters of the regression equation to be underestimated (Clark and Perry 1994). On the other hand, failing to use logarithms typically reduces the precision of the estimates because the variance of s^2 tends to increase with $\bar{x}$ and because anomalous points are allowed excessive leverage (Taylor et al. 1978).

Still another problem arises when the maximum value the variate can attain is finite. The number of fish kept per trip, for example, might be limited to L owing to regulations or gear saturation. In such cases the maximum possible s^2 for a given $\bar{x}$ is

$$s^2_{MAX} = \frac{i(L-\bar{x})^2 + (n-i-1)\bar{x}^2 + (R-\bar{x})^2}{n-1}, \tag{7}$$

where i is the integer part of $n\bar{x}/L$ and R is the remainder. Clearly s^2_{MAX} decreases to zero as $\bar{x}$ approaches L, therefore the s^2 to $\bar{x}$ relationship will exhibit a dome-shaped trend regardless of the statistical distribution of the population or the relationship between σ^2 and μ. For example, if all of the observations happen to be either 0 or 1, then s^2 always equals $n(\bar{x} - \bar{x}^2)/(n-1)$ and there is absolutely no possibility of deducing the relationship between σ^2 and μ. One could attempt to derive a suitable transformation from the fit of a quadratic equation to the (s^2, $\bar{x}$) pairs, but this would be subject to the same caveats discussed above.

In summary, it is clear that the s^2 and $\bar{x}$ of non-negative variates are spuriously correlated. The examples suggest that, when μ is small and σ^2 is large, s^2 tends to underestimate σ^2 in some proportion to the degree that $\bar{x}$ underestimates μ. Inasmuch as uncommon resources with high variance are frequently encountered in fisheries work, this statistical artifice is likely to pose a significant challenge. In general, weighting estimates from scarce resources by s^2/n may bias the assessment by overemphasizing estimates that happen to be too low. This is especially likely to be a problem when n is small because $\bar{x}$ is more likely to be poorly estimated and improperly weighted; therefore it may be prudent to simply ignore statistics from small samples. It must be emphasized, however, that the strong spurious correlation between s^2 and $\bar{x}$ persists even with large n and that regressions of s^2 and $\bar{x}$ may be highly significant even when there is no relationship between σ^2 and μ as long as some of the $\bar{x}$ values used in the regression are small. Accordingly, weighting schemes or transformations that are based on the empirical relationship between s^2 and $\overline{x}$ are apt to be meaningless unless μ is large (>> 5) or σ^2 is small (on the order of μ). Perhaps a better alternative might be to employ transformations or objective functions that are known to be robust to the type of deviations one is likely to observe in fisheries data (e.g., Otter Research Ltd. 1994, Schnute and Richards 1995), but more work is needed.

Acknowledgments

I thank F. Funk, V. Restrepo, D. Vaughan, J. Zweifel, and two anonymous reviewers for their helpful comments.

References

Bartlett, M.S. 1947. The use of transformations. Biometrics 3:39-52.

Clark, S.J., and J.N. Perry. 1994. Small sample estimation for Taylor's power law. Environ. Ecol. Stat. 1:287-302.

Curtis, J.H. 1943. On transformations used in analysis of variance. Ann. Math. Stat. 14:10.

Cyr, H., J.A. Downing, S. Lalonde, S.B. Baines, and M.L. Pace. 1992. Sampling larval fish populations: Choice of sample number and size. Trans. Am. Fish. Soc. 121:356-368.

Deriso, R.B., and A.M. Parma. 1987. On the odds of catching fish with angling gear. Trans. Am. Fish. Soc. 116:244-256.

Dong, Q., and V.R. Restrepo. 1996. Notes on the Poisson error assumption made to estimate relative abundance of West Atlantic bluefin tuna. Int. Comm. Conserv. Atl. Tunas, Coll. Vol. Sci. Pap. 45(2):158-161.

Downing, J.A. 1986. Spatial heterogeneity: Evolved behavior or mathematical artefact. Nature 323:255-257.

Downing, J.A. 1989. Precision of the mean and the design of benthos sampling programmes: Caution revised. Mar. Biol. 103:231-234.

Feller, W. 1968. An introduction to probability theory and its applications. Volume 1. John Wiley and Sons, New York. 509 pp.

Green, R.H. 1979. Sampling design and statistical methods for environmental biologists. John Wiley and Sons, New York. 257 pp.

Houser, A., and J.E. Dunn. 1967. Estimating the size of the threadfin shad population in Bull Shoals Reservoir from midwater trawl catches. Trans. Am. Fish. Soc. 96:176-184.

Johnston, R.G., S.D. Schroder, and R.A. Mallawaaratchy. 1995. Statistical artifacts in the ratio of discrete quantities. The American Statistician 49(3):285-291.

Kenney, B.C. 1982. Beware of spurious self-correlations! Water Resour. Res. 18(4):1041-1048.

Lester, N.P., M.M. Petzold, and W.I. Dunlop. 1991. Sample size determination in roving creel surveys. Amer. Fish. Soc. Symp. 12:25-39.

Mangel, M., and J.H. Beder. 1985. Search and stock depletion: Theory and applications. Can. J. Fish. Aquat. Sci. 42:150-163.

Myers, R.A., and P. Pepin. 1991. Reply to: On testing the robustness of lognormal-based estimators. Biometrics 47:1623-1624.

Otter Research Ltd. 1994. AUTODIF: A C++ array language extension with automatic differentiation for use in nonlinear modeling and statistics. Otter Research Ltd., Nanaimo, B.C.

Pella, J.J., and C.T. Psaropulos. 1975. Measures of tuna abundance from purse-seine operations in the eastern Pacific Ocean adjusted for fleet-wide evolution of increased fishing power, 1960-1971. Inter-Am. Trop. Tuna Comm. Bull. 16:283-354.

Pennington, M. 1996. Estimating the mean and variance from highly skewed marine data. Fish. Bull. U.S. 94:498-505.

Porch, C.E., and W.W. Fox. 1990. Simulating the dynamic trends of fisheries regulated by small daily bag limits. Trans. Am. Fish. Soc. 119:836-849.

Ross, G.J.S. 1990. Incomplete variance functions. Journal of Applied Statistics 17:3-8.

Schnute, J.T, and L.J. Richards. 1995. The influence of error on population estimates from catch-age models. Can. J. Fish. Aquat. Sci. 52:2063-2077.

Small, I., and D.Y. Downham. 1985. The interpretation of angler's records (trout and seatrout, *Salmo trutta* L., and *Salmo salar* L.) Aquacult. Fish. Manage. 16:151-169.

Smith, S.J.1990. Use of statistical models for the estimation of abundance from groundfish trawl survey data. Can. J. Fish. Aquat. Sci. 47:894-903.

Taylor, C.C. 1953. Nature of variability in trawl catches. Fish. Bull. U.S. 54:145-166.

Taylor, L.R. 1980. New light on the variance/mean view of aggregation and transformation: Comment. Can. J. Fish. Aquat. Sci. 37:1330-1332.

Taylor, L.R., and I.P. Woiwod. 1982. Comparative synoptic dynamics. I. Relationships between inter- and intra-specific spatial and temporal variance/mean population parameters. J. Anim. Ecol. 51:879-906.

Taylor, L.R., I.P. Woiwod, and J.N. Perry. 1978. The density-dependence of spatial behavior and the rarity of randomness. J. Anim. Ecol. 47:383-406.

Taylor, L.R., R.A. Taylor, I.P. Woiwod, and J.N. Perry. 1983. Behavioural dynamics. Nature 303:801-804.

Tokeshi, M. 1995. On the mathematical basis of the variance-mean power relationship. Resource Population Ecology 37(1):43-48.

Warwick, R.M., and K.R. Clarke. 1993. Increased variability as a symptom of stress in marine communities. J. Exp. Mar. Biol. Ecol. 172:215-225.

Welch, D.W., and Y. Ishida. 1993. On the statistical distribution of salmon in the sea: Application of the negative binomial distribution, and the influence of sampling effort. Can. J. Fish. Aquat. Sci. 50:1029-1038.

Appendix: Derivation of the Minimum Possible Variance

Taylor and Woiwood (1982) show that, if the total number of objects T is less than n, then $s^2_{MIN} = n(\bar{x} - \bar{x}^2)/(n-1)$. This expression can be generalized to any value of T by noting that the minimum variance is obtained when the number of objects is distributed as evenly as possible among the two integer categories i and $i + 1$, where i is the integer part of $\bar{x}$. Designating the number of sampling units containing i objects as n_i, we have

$$T = (i+1)n_{i+1} + in_i$$
$$n = n_{i+1} + n_i \,.$$

Solving for n_{i+1} and n_i and then substituting them into the expression for sample variance yields

$$s^2_{MIN} = \frac{(T - in)(i + 1 - \bar{x})^2 + [(i+1)n - T](i - \bar{x})^2}{n-1}$$

which simplifies to

$$s^2_{MIN} = (\bar{x} - i - 1)(i - \bar{x})\frac{n}{n-1}.$$

Diagnosing Systematic Errors in Reported Fishery Catch

Stratis Gavaris and Lutgarde Van Eeckhaute
Department of Fisheries and Oceans, Biological Station, St. Andrews, N.B., Canada

Abstract

Age structured assessment methods integrate information from the commercial fishery catch and from indices of abundance. Systematic errors in the catch at age, often caused by mis-reporting of landings or discarding of small fish, could introduce appreciable bias to the results of stock assessments. Important inconsistencies were revealed when results from a base assessment employing catch at age were compared to results from a model that employs proportion caught at age but estimates total numbers caught each year in addition to the other model parameters, terminal year abundance and index catchability. For eastern Georges Bank haddock, the diagnostic model suggested that the base model had underestimated the year-class abundance for years prior to 1977. It also indicated that the mid-1980s year classes might have been overestimated by the base model. Though not recommended for estimation, the diagnostic model subjects data to different assumptions, making it useful as an investigative tool for evaluating the robustness of results.

Introduction

Fish stock assessment methods for age-structured analyses integrate auxiliary information from population abundance trend indices with data from catch removals by the fisheries to estimate stock status (e.g., Fournier and Archibald 1982, Deriso et al. 1985, Gavaris 1988). Deviations between observed and predicted quantities for these models are assumed to be random. Abundance indices are often derived from controlled research survey data or from a selected subset of fishery catch and effort records. Therefore, appropriate measures, such as random selection of survey locations, maintaining consistent fishing practices on surveys from year to year or factoring out gear, seasonal, and spatial effects from catch rates, can be taken to prevent the introduction of systematic deviations. Commercial fishery catch data, however, may be suspect of containing errors which display a non-random pattern. Fishery catch removals by age can be obtained by applying the age composition derived from samples for length composition and aging material to the total number caught (Gavaris and Gavaris 1983). The total amount caught is typically based on a complete census of fishery landings and ancillary information on discards, if available. Systematic errors in the catch data are commonly due to mis-reporting of total amount caught, although the samples for length composition may also be biased if size-selective culling has occurred prior to observation.

Systematic errors in the catch-at-age could introduce appreciable bias to the results of stock assessments. A constant mis-reporting rate may not pose problems but trends in the total amount or the age composition of mis-reported catch result in complex patterns of bias (Sinclair et al. 1991). Employing assessment models that accommodate random error in the catch-at-age would not be an appropriate treatment for this problem. Concern regarding the reliability of catch data has stimulated development of methods that employ only survey data. One approach proposed by Cook assumes separable age and year effects on fishing mortality rate (ICES 1997) while Sinclair and Chouinard (1992) employed a multiplicative model and assumed a constant fishing mortality rate over some period. Patterson retained the assumption of separable fishing mortality rate and extended the method to include sampling data from the commercial fishery (ICES 1997).

Many assessments are conducted without invoking the separable fishing mortality rate assumption because year-to-year variation in the exploitation pattern by age has been observed and is considered meaningful. We compare results from a non-separable assessment employing catch-at-age and abundance indices to those from a non-separable model that did not use total number caught but employed proportion caught-at-age and abundance indices. The total number caught each year was estimated in addition to the other model parameters. We will refer to the latter as the diagnostic model and the former as the base model. These models were applied to data for eastern Georges Bank haddock. We do not promote the diagnostic model for estimation purposes. The purpose of this study was to explore the utility of the diagnostic model, which places less reliance on suspect fishery catch data, as a tool for investigating inconsistencies which might be caused by systematic errors.

Methods

The commercial fishery catch-at-age is used to reconstruct the past using virtual population analysis (VPA). Due to the error propagation properties of VPA (Pope 1972, Sampson 1988), random errors in the catch-at-age typically have negligible impact on assessment calibration methods compared to random errors in the abundance indices (Sinclair et al. 1991). Capitalizing on this characteristic can greatly reduce the number of unknown parameters. It is not uncommon, therefore, to encounter assessments which ignore the random error in catch-at-age. Both the base and diagnostic models adopt this tact. In its general form, the VPA calculations require an estimate of population abundance for each year class. With suitable assumptions about fishing mortality rate patterns on older ages, the calculations can be done if we have estimates of population abundance for the terminal year of the analysis. The abundance indices are the auxiliary data used to obtain those estimates by a process referred to as calibration of the VPA. The calibration process involves fitting abundance index data and population abundance results from VPA to a catchability model to estimate the parameters of that relationship, referred to as the calibration constants. A search is conducted to find the terminal year-class abundance estimates for the VPA that result in the best fit to the catchability model.

For eastern Georges Bank haddock, the sampling data from the commercial fishery were used to derive estimates of annual proportion caught at age, $P_{a,t}$, for ages $a = 0, 1, 2, \ldots, 8$ and periods $t = 1969, 1970, 1971, \ldots,$ 1996. The time period label signifies the beginning of the period during which the catch occurred. The commercial fishery catch information may be used with the samples to derive estimates of annual numbers caught at age, $C_{a,t}$, for the same ages and periods. The procedures described in Gavaris and Gavaris (1983) were used to combine samples and to apply the catch information. The available abundance indices, $I_{s,a,t}$, were from bottom trawl research surveys conducted by the Canadian Department of Fisheries and Oceans (DFO) and the U.S. National Marine Fisheries Service (NMFS):

for s = DFO spring survey, ages $a = 1,2,\ldots,8$, times $t = 1986.16,$ 1987.16,…, 1997.16

s = NMFS spring survey, ages $a = 1,2,\ldots,8$, times $t = 1969.29, 1970.29, \ldots,$ 1996.29

s = NMFS fall survey, ages $a = 0,1,\ldots,5$, times $t = 1969.69, 1970.69, \ldots,$ 1996.69.

The fall survey captures young of the year and that information is included as 0 group, but older haddock appear less available during this season. Survey indices were not included for older ages where catches were sparse and there were frequent occurrences of zero catches. Accordingly, a plus group was not included in the VPA as it would have no impact on the calibration. If desired, the abundance of the plus group in the population could be computed by making a suitable assumption about its fishing mortality rate and applying the catch equation (equation 3 below). Zero observations for abundance indices were treated as missing data. Further specific details of the data used for the assessment are provided in Gavaris and Van Eeckhaute (1997).

The model formulation employed for the base model assumed that the random error in the catch at age was negligible and the annual total number caught, L_t, were reported accurately. The catch at age then can be expressed as $C_{a,t} = L_t P_{a,t}$. The errors in the abundance indices were assumed independent and identically distributed after taking natural logarithms of the values. The annual natural mortality rate, M, was assumed constant and equal to 0.2. A model formulation, using as parameters the natural logarithm of population abundance at the beginning of the year, was considered because of close to linear behavior for such a parameterization (Gavaris 1993). The following model parameters were defined: $\theta_{a,t'}$ = ln population abundance for ages a = 1, 2,…, 8 at time t' = 1997.25; $\kappa_{s,a}$ = ln calibration constants for each abundance index source s and relevant ages a. We solved for the parameters by minimizing the sum of squared differences between the natural logarithm observed abundance indices and the natural logarithm population abundance adjusted for catchability by the calibration constants. The objective function for minimization was defined as

$$\Psi_{s,a,t}\left(\hat{\theta},\hat{\kappa}\right) = \sum_{s,a,t}\left[\psi_{s,a,t}\left(\hat{\theta},\hat{\kappa}\right)\right]^2 = \sum_{s,a,t}\left\{\ln I_{s,a,t} - \left[\hat{\kappa}_{s,a} + \ln N_{a,t}\left(\hat{\theta}\right)\right]\right\}^2 \qquad (1)$$

For convenience, the population abundance $N_{a,t}(\hat{\theta})$ is abbreviated by $N_{a,t}$. At time t', the population abundance was obtained directly from the parameter estimates, $N_{a,t'} = e^{\hat{\theta}_{a,t'}}$. For all other times, the population abundance was computed using the virtual population analysis algorithm, which incorporates the common exponential decay model (Beverton and Holt 1957)

$$N_{a+\Delta t,t+\Delta t} = N_{a,t} e^{-\left(F_{a,t}+M_a\right)\Delta t} \qquad (2)$$

Year was used as the unit of time; therefore ages were expressed as years and the fishing and natural mortality rates were annual instantaneous rates. The fishing mortality rate $F_{a,t}$ exerted during the time interval t to $t + \Delta t$, was obtained by solving the catch equation

$$C_{a,t} = \frac{F_{a,t}\Delta t N_{a,t}\left(1 - e^{-\left(F_{a,t}+M_a\right)\Delta t}\right)}{\left(F_{a,t} + M_a\right)\Delta t} \qquad (3)$$

using a Newton-Raphson algorithm. The fishing mortality rate for the oldest age in the last time interval of each year was assumed equal to the weighted average for ages fully recruited to the fishery during that time interval

$$F_{8,t} = \frac{\sum_{a=4}^{7} N_{a,t} F_{a,t}}{\sum_{a=4}^{7} N_{a,t}} \tag{4}$$

The diagnostic model was similar in every respect to the base model except that we rejected the assumption that the total number caught each year by the fishery was reported accurately. Therefore, in addition to the model parameters $\theta_{a,t}$, and $\kappa_{s,a}$, the total number caught each year were estimated, λ_t = total number caught for time period t = 1969, 1970,..., 1996. The catch-at-age was obtained from $C_{a,t} = \lambda_t p_{a,t}$ and the computations thereafter were the same as for the base model. To maintain comparability, all other assumptions were the same as those for the base model. Both the base model and the diagnostic model were implemented within ADAPT (Gavaris 1988), which employs a Marquardt nonlinear least squares algorithm.

With both models, the covariance matrix of the parameters was estimated using the common linear approximation (Kennedy and Gentle 1980, p. 476)

$$\mathrm{Cov}\left(\hat{\theta},\hat{\kappa}\right) = \hat{\sigma}^2\left[J^T\left(\hat{\theta},\hat{\kappa}\right)J\left(\hat{\theta},\hat{\kappa}\right)\right]^{-1} \tag{5}$$

where $\hat{\sigma}^2$ is the mean square residual and $J(\hat{\theta},\hat{\kappa})$ is the Jacobian matrix. The bias of the parameters was estimated using Box's (1971) approximation, which assumes that the errors are normally distributed

$$\begin{aligned}\mathrm{Bias}\left(\hat{\theta},\hat{\kappa}\right) = \frac{-\hat{\sigma}^2}{2}\left[\sum_{s,a,t} J_{s,a,t}\left(\hat{\theta},\hat{\kappa}\right)J^T_{s,a,t}\left(\hat{\theta},\hat{\kappa}\right)\right]^{-1}\left[\sum_{s,a,t} J_{s,a,t}\left(\hat{\theta},\hat{\kappa}\right)\right] tr \\ \left\{\left[\sum_{s,a,t} J_{s,a,t}\left(\hat{\theta},\hat{\kappa}\right)J^T_{s,a,t}\left(\hat{\theta},\hat{\kappa}\right)\right]^{-1} H_{s,a,t}\left(\hat{\theta},\hat{\kappa}\right)\right\}\end{aligned} \tag{6}$$

where $J_{s,a,t}(\hat{\theta},\hat{\kappa})$ are vectors of the first derivatives for each $\psi_{s,a,t}(\hat{\theta},\hat{\kappa})$ and $H_{s,a,t}(\hat{\theta},\hat{\kappa})$ are the Hessian matrices for each $\psi_{s,a,t}(\hat{\theta},\hat{\kappa})$.

Population quantities of interest for management advice are functions of the estimated parameters. Denote an arbitrary quantity by $\hat{\alpha} = g(\hat{\theta},\hat{\kappa})$ where g is the transformation function. The variance and bias were estimated using the methods described in Ratkowsky (1983)

$$\mathrm{Var}\left(\hat{\alpha}\right) = tr\left[GG^T \mathrm{cov}\left(\hat{\theta},\hat{\kappa}\right)\right] \tag{7}$$

$$\text{Bias}(\hat{\alpha}) = G^T \text{Bias}(\hat{\theta}, \hat{\kappa}) + \frac{tr\left[W \operatorname{cov}(\hat{\theta}, \hat{\kappa})\right]}{2} \tag{8}$$

where G is the vector of first derivatives of g with respect to parameters and W is the matrix of second derivatives of g with respect to parameters.

Results

Proportion caught at age and abundance indices do not contain any information on the absolute magnitude of the population. Therefore, to make the diagnostic model determinate, it is necessary to set one of the parameters to a fixed value. We arbitrarily assigned the abundance of the 1992 year class at time $t' = 1997.25$. In retrospect, we suggest that it might be better to select one of the estimated catches to be assigned a fixed value instead of one of the year-class estimates as we did. This implementation modification would not alter the basic results but it would avoid unduly restricting a single year-class estimate, as we observed in the sensitivity trials.

The first attempt to estimate all the remaining parameters in the diagnostic model failed. The abundance indices for this resource, which were derived from bottom trawl surveys, were variable and occasionally indicated an increase in the abundance of a year class (Fig. 1) when mortality processes dictate a decline. In these instances, the model estimates of total numbers caught tend to zero, an unrealistic result. Seven of the 29 years displayed this behavior, so we assigned fixed values to these parameters equal to the reported total number caught in those years. We investigated the sensitivity of the results to these assumed values and report on that later.

Parameter estimates from nonlinear models are typically biased, although the bias may be negligible if the parameter behavior is close to linear. The estimates of population abundance needed to compute the VPA for eastern Georges Bank haddock have substantial bias and the VPA computed with bias adjusted estimates is used as the basis for resource evaluation (Gavaris and Van Eeckhaute 1997). Further, the nature of the bias for the base model and for the diagnostic model displayed different characteristics. Consequently, all comparisons were based on bias adjusted results.

Since the absolute magnitude of the population abundance was arbitrarily set by the fixed parameters, we compared patterns of relative year-class abundance. The base and diagnostic models both identified the two largest year classes, 1975 and 1978, and the 1983, 1985, 1987 and 1992 year classes which were of intermediate magnitude (Fig. 2). The diagnostic model indicated that the 1971 and 1972 year classes were relatively stronger than the estimates obtained from the base model. In general, the rankings of year classes from the two models were comparable (rank cor-

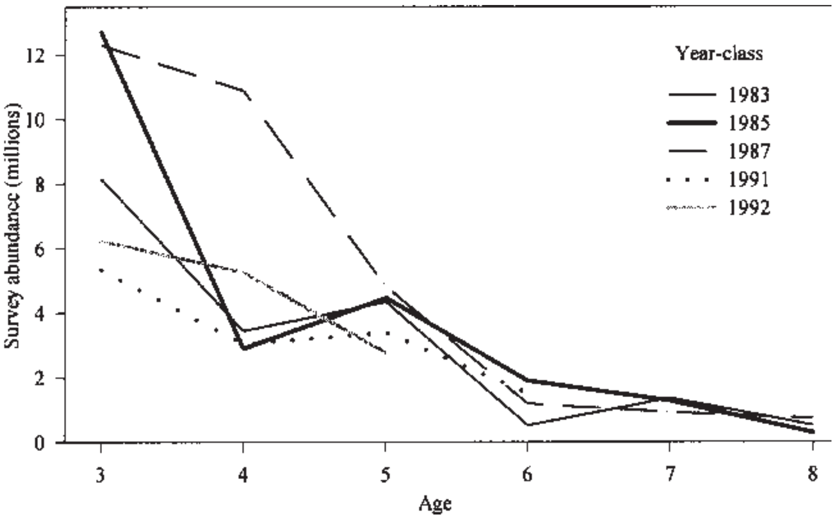


Figure 1. As year classes increase in age, mortality processes dictate a decline in abundance with each successive year. However, year-class abundance estimates from bottom trawl surveys for Georges Bank haddock at ages for which they are fully recruited to the survey sometimes show an increase in abundance due to sampling variability.

relation = 0.98). Better resolution for comparison of the trends is obtained by comparing the ratio of diagnostic model estimates to the base model estimates. The ratio was typically about two but there was a tendency for somewhat higher values in the earlier years. The largest deviations occurred for the 1971 and 1972 year classes. The deviation for the large 1975 year class was among the larger ones at 50% greater than the modal value. There also appears to be a short period during the mid-1980s when the ratios were smaller. The relative errors for year-class abundance from the two models show markedly different patterns. The base model results, on the merit of assumed known catch-at-age, show a pattern of rapidly diminishing error for earlier year classes while the diagnostic model results remain constant at about 50% over the entire period (Fig. 2).

Comparison of the fishing mortality rates for ages 4 to 8, which were considered to be fully recruited to the fishery, indicated that the results from the diagnostic model were much more erratic (Fig. 3). In the diagnostic model, the exploitation pattern is derived principally from the abundance indices, which, as has been noted, are quite variable. This type of erratic behavior is not uncharacteristic of mortality rates derived from abundance indices alone. Further, the fishing mortality pattern from the diagnostic model does not correspond well with that from the base model

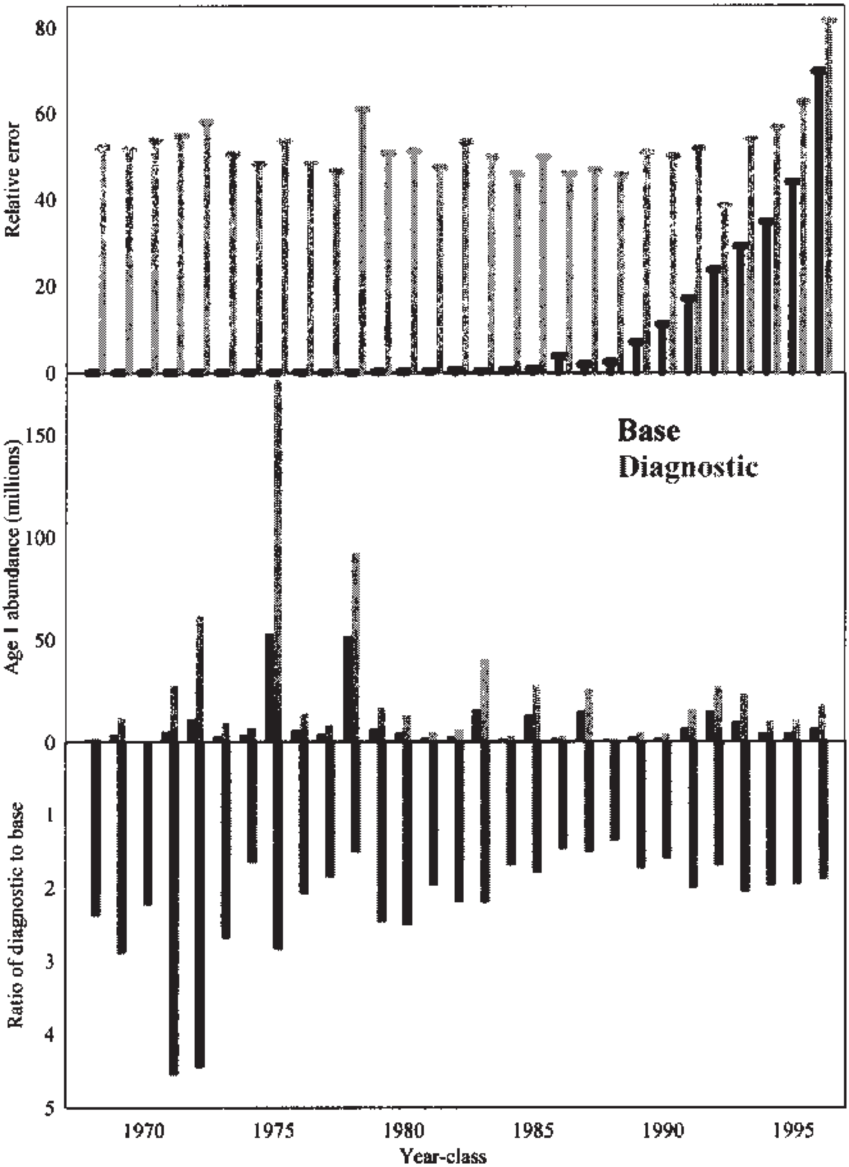


Figure 2. Absolute (middle) and relative (bottom) year-class abundance at age 1 from base and diagnostic assessment models for Georges Bank haddock data, and relative error (standard error/mean) (top) associated with abundance estimates. The 1971 and 1972 relative estimates from the diagnostic model are proportionally much larger than the average ratio. Relative errors for the diagnostic model are about 50% for all estimates while the error in the base model diminishes rapidly for earlier year classes.

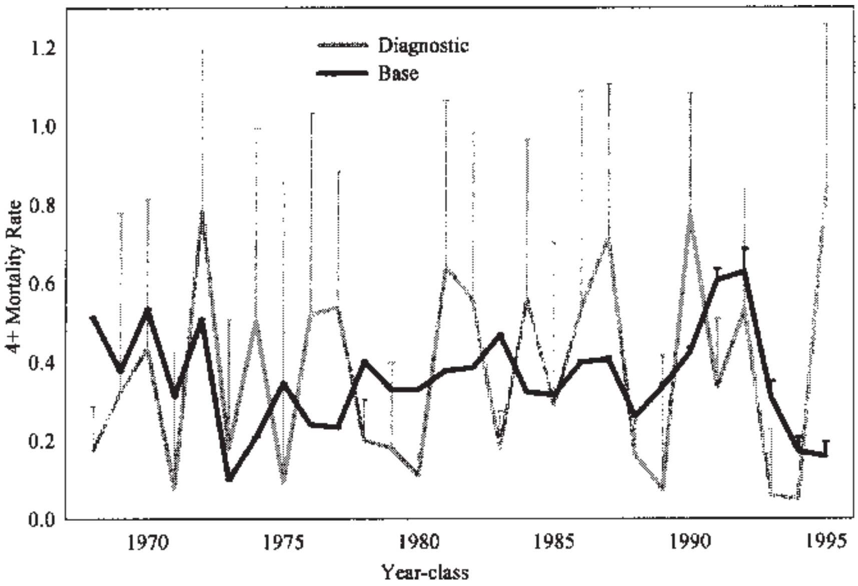


Figure 3. Georges Bank haddock fishing mortality rates for ages which are fully recruited to the fishery, ages 4 to 8, are much more variable for the diagnostic than for the base model. The relative error (standard error/ mean) on the base model estimates are small and diminish rapidly for earlier year classes while the error on the diagnostic model estimates remain high throughout the time series.

(rank correlation = 0.40), though the general magnitude over the series was comparable for the two. As for the population abundance, the standard errors for the fishing mortality rate from the base model diminish in earlier years but those from the diagnostic model are much larger throughout the time series.

To investigate the sensitivity of the results to the assumed values for the number caught in the seven years where the parameters were assigned fixed values, we conducted two sets of trials. In one set of trials, we halved all the assumed values simultaneously and then we doubled all the assumed values simultaneously. In the former case, all the year-class abundance estimates decreased proportionately while in the latter, they all increased proportionately. Therefore, the relative year-class strengths did not change. In the second set of trials, we selected random multipliers between one half and two and applied them to the fixed values for number caught. The results from 10 such replicates indicated that the year-class abundance patterns were similar (Fig. 4). Examining the ratio of these 10 trials to the original estimates from the diagnostic model revealed bet-

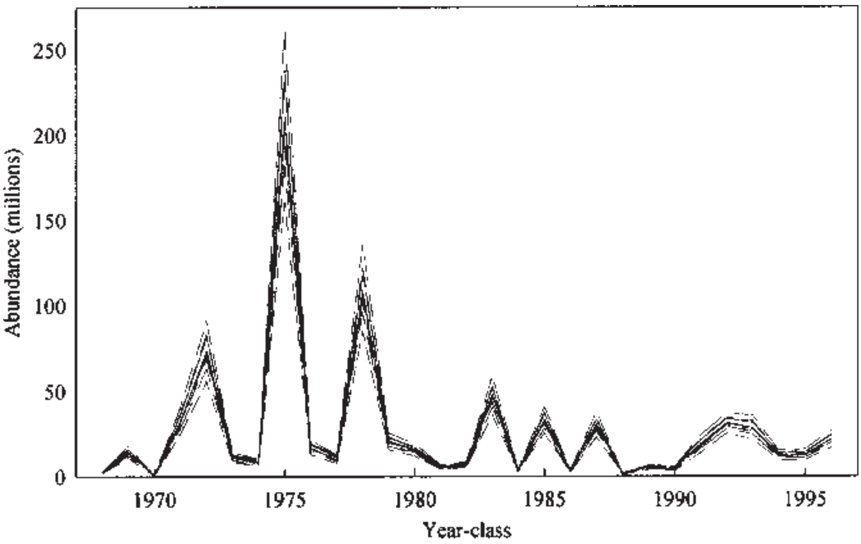


Figure 4. To test the sensitivity of the diagnostic model to fixing the values for numbers caught for several years, the values were varied randomly over a range of half to double the original values. Georges Bank haddock year-class abundance patterns were similar for the 10 trials shown.

ter resolution of the deviations. The ratio lines are essentially horizontal with no persistent slopes (Fig. 5) suggesting that the relative year-class strengths were practically unchanged. Recall that the abundance of the 1992 year class was arbitrarily set to a constant at the beginning of 1997. Therefore, there is limited scope for the abundance of this year class to adjust, and this is evident in the comparison. We concluded from these results that the relative year-class pattern was fairly robust to the assumed fixed values for number caught in those seven years, rendering the comparisons valid.

Discussion

Statistical models that accommodate random variation do not adequately handle systematic errors in data. Consequently, model results may be inaccurate and unreliable. For fishery stock assessment models, mis-reporting of catches is generally considered to be the most serious source of systematic errors. However, fishery catches play a very important role in the estimation of stock status. In the absence of catch information, absolute estimates of abundance generally cannot be obtained and the relative

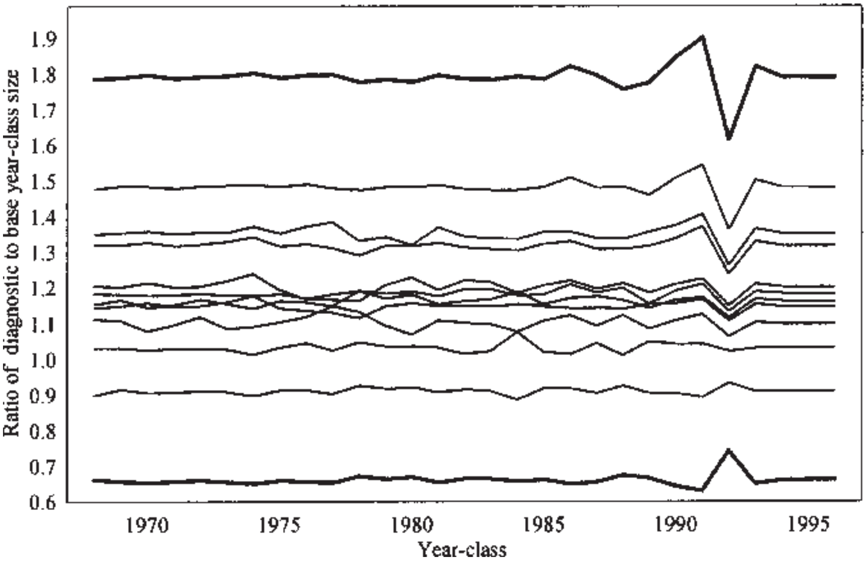


Figure 5. Relative size of Georges Bank haddock year-class abundance estimates from the diagnostic model when the fixed catch values were varied randomly through one half to double the values used in the base model. The ratio of year-class size, diagnostic to base, show that the relative year-class pattern was fairly robust to the fixed number caught in those select years. The thicker top and bottom lines are the results from doubling (top) and halving (bottom) the base values.

estimates may display large estimation error unless the abundance indices are very precise.

We employed a diagnostic model to identify inconsistencies between results when fishery catches were accepted as accurate and reliable compared to results when those data were not used. We do not advocate use of the diagnostic model for estimation of the status of resources. As anticipated, application of this model to the Georges Bank data demonstrated that the solution might be indeterminate or unstable unless additional assumptions are made. However, the comparison of results suggested that the base model had smaller estimates of the year-class abundance for earlier years. It also indicated that the mid-1980s year classes might have been overestimated by the base model. These observations are useful for focusing attention on specific periods where further investigation of data reliability or robustness to violation of assumptions may be pursued.

For the Georges Bank haddock fishery, 1977 was the last year this resource was fished under an international quota management system.

Subsequently, the fishery was restricted to coastal states. Differences in pre-1977 and post-1977 monitoring and reporting practices by fleets or possibly intentional mis-reporting by fleets to circumvent restrictive quotas are likely possibilities. The strong 1975 year class would have been susceptible to heavy exploitation at age 2 in 1977 under the international regime. Though estimates of discards for the U.S. fishery have been included, no such adjustments have been attempted for any of the other fleets. The reported catch for this year class at younger ages is probably deficient. Similarly, discard estimates for the U.S. fishery only are included for the 1972 year class, suggesting that the discrepancy for the 1971 and 1972 year classes may have been due to unreported culling and discarding of small fish. Immediately following the establishment of the Canada-U.S. boundary on Georges Bank in 1985, the Canadian fishery in that area was relatively unrestricted. Quotas in the adjacent management unit however, were often restrictive at that time. Fish caught in the adjacent management unit may have been reported as originating on Georges Bank. To prevent this from occurring, fishing vessels now have to choose which grounds they will fish during a trip and they may not fish in both management units on the same trip.

A further avenue that may be pursued is to apply the base and diagnostic models to data for older fish only. For example, in the haddock case, we could apply the models and compare results for ages 3 and older. This may allow us to distinguish between overall misreporting of catches and adequate reporting of landed catch but non-reporting of culled and discarded fish catches.

When the accumulated evidence indicates that fishery catch data may include systematic errors and the assessment results are significantly affected by inclusion of those data, corrective measures should be pursued. In extreme cases where persistent trends are apparent over the time series, it may be prudent to truncate the range of years considered in the analysis to avoid possible non-stationary processes, e.g., decreasing discard rates. For the haddock example, we might investigate the results from an analysis using only the data subsequent to the late 1970s. Another option would be to exclude suspect data points from the objective function. For example, with haddock, we might exclude all data for the 1971, 1972, and 1975 year classes in the minimization. A less severe variant of this would be to include the data but arbitrarily assign a lower weight to those observations. A further option might be to use a hybrid model where the number caught are estimated in only a few years which were suspect while the reported catches were used in all other years. We might use such a model for haddock and estimate the number caught for the years 1985-1988 only.

We have focused here on biases in the fishery catch data because we have assumed that appropriate measures were taken to avoid systematic deviations in the abundance indices. Differences between the diagnostic model and the base model could be caused by systematic patterns in the indices, such as time trends in catchability. Such possibilities should be considered and investigated before the results can be interpreted.

The inability of models, which do not use fishery catch data, to estimate absolute abundance and the high imprecision for estimates obtained from these models limits their use for practical fisheries management advice. It is good practice however, to subject data to models with differing assumptions to evaluate the robustness of results. In this regard, we recommend the diagnostic model as a useful investigative tool, which can uncover important inconsistencies.

Acknowledgments

We are grateful to the anonymous reviewers for many useful comments and suggestions.

References

Beverton, R.J.H., and S.J. Holt. 1957. On the dynamics of exploited fish populations. Fish. Invest. Ser. II, 19. 533 pp.

Box, M.J. 1971. Bias in nonlinear estimation. J. R. Statist. Soc., Ser. B 33:171-201.

Deriso, R.B., T.J. Quinn II, and P.R. Neal. 1985. Catch at age analysis with auxiliary information. Can. J. Fish. Aquat. Sci. 42:815-824.

Fournier, D., and C.P. Archibald. 1982. A general theory for analyzing catch at age data. Can. J. Fish. Aquat. Sci. 39:1195-1207.

Gavaris, S. 1988. An adaptive framework for the estimation of population size. Can. Atl. Fish. Sci. Adv. Comm. (CAFSAC) Res. Doc. 88/29. 12 pp.

Gavaris, S. 1993. Analytical estimates of reliability for the projected yield from commercial fisheries. In: S.J. Smith, J.J. Hunt, and D. Rivard (eds.), Risk evaluation and biological reference points for fisheries management. Can. Spec. Publ. Fish. Aquat. Sci. 120:185-191.

Gavaris, S., and C.A. Gavaris. 1983. Estimation of catch at age and its variance for groundfish stocks in the Newfoundland region. In: W.G. Doubleday and D. Rivard (eds.), Sampling commercial catches of marine fish and invertebrates. Can. Spec. Publ. Fish. Aquat. Sci. 66:178-182.

Gavaris, S., and L. Van Eeckhaute. 1997. Assessment of haddock on eastern Georges Bank. DFO Canadian Stock Assessment Secretariat Res. Doc. 97/54. 72 pp.

ICES. 1997. Report of the comprehensive fishery evaluation working group. ICES C.M. 1997/Assess:15. 140 pp.

Kennedy Jr., W.J., and J.E. Gentle. 1980. Statistical computing. Marcel Dekker, New York. 591 pp.

Pope, J.G. 1972. An investigation of the accuracy of virtual population analysis using cohort analysis. Int. Comm. Northwest Atl. Fish. Res. Bull. 9:65-74.

Ratkowsky, D.A. 1983. Nonlinear regression modeling. Marcel Dekker, New York. 276 pp.

Sampson, D.B. 1988. The stability of virtual population analysis cohort size estimates. J. Cons. Cons. Int. Explor. Mer. 44:135-142.

Sinclair, A., and G. Chouinard 1992. Application of a multiplicative model to research survey data for two cod stocks. Can. Atl. Fish. Sci. Adv. Comm. (CAFSAC) Res. Doc. 92/66. 25 pp.

Sinclair, A., D. Gascon, R. O'Boyle, D. Rivard, and S. Gavaris. 1991. Consistency of some northwest Atlantic groundfish stock assessments. Northwest Atl. Fish. Organ. Sci. Counc. Stud. 16:59-77.

Fishery Stock Assessment Models
Alaska Sea Grant College Program • AK-SG-98-01, 1998

Model Misspecification or Data Bias: The Assessment of Prince Rupert Herring

Jacob F. Schweigert
Fisheries and Oceans Canada, Pacific Biological Station, Nanaimo, B.C., Canada

Abstract

Pacific herring stock assessments have been conducted with two analytical models since the early 1980s. Correspondence between the estimates of abundance provided by these independent assessments of stock size has been good for four of the five major stock groups. In the remaining area, the Prince Rupert assessment region, estimates of abundance from these models have diverged widely, particularly in recent years. An investigation of the biological sampling data for this area revealed limited information on age structure from the fall food and bait fisheries of the late 1970s and early 1980s. In addition, there were shifts in the pattern of sample collection throughout the region with time which could be responsible for a bias in the estimates from the age-structured or catch-age model. Deletion of the fall samples from the input data resulted in an improvement in model fit. An investigation of possible bias in the spawn index, used to tune the age-structured model, due to changes in survey methodology yielded minor improvements in model fit and may be artifactual rather than real. Analyses based on the reduced data series described here appear to provide a more realistic reconstruction of stock size and trajectory than in the past and should be used in future assessments of the herring population. This paper provides an example of the benefits of exploring the value and consistency of data inputs or sources.

Introduction

The assessment of Pacific herring stocks in British Columbia has relied on two models to estimate and forecast stock size since the early 1980s. The first model is a modification of the escapement model described by

Schweigert and Stocker (1988) and reconstructs abundance from an assessment of egg deposition, age composition, and weight-at-age data. The second is a catch-at-age or age-structured model (Fournier and Archibald 1982). These two models have provided similar estimates of absolute abundance and stock trajectories for four of the five major herring assessment regions in British Columbia. In the remaining assessment region, Prince Rupert, the age-structured model estimate of abundance has deviated substantially from the escapement model estimate of stock size in recent years. The age-structured model estimate greatly exceeds the stock size determined from egg deposition data, as well as from the qualitative in-season sonar and sounder estimates of abundance, to the extent that this estimate is not considered in the decision making process for determining annual harvestable surplus. The objective of this study is to revisit the biological data available for this assessment region to determine whether the observed discrepancy in stock trends is a function of model misspecification or bias in the underlying catch-at-age data, with the aim of providing a more realistic assessment of abundance for this area.

Methods

Stock Considerations

The geographic scope of the Prince Rupert herring assessment region is shown in Fig. 1 detailing the herring sections which are used to summarize biological characteristics of the stock at finer spatial scales within the region (Haist and Rosenfeld 1988). The major spawning sites for herring in this region occur near Port Simpson in section 42 and in Kitkatla in section 52. In recent years there has been an increase in spawning at Port Simpson and a decrease at Kitkatla. This has led to speculation that the two sites represent distinct herring stocks. However, two separate tagging studies have demonstrated considerable movement of fish between these and other areas (Stevenson 1954, Haegele 1991). Also, suggestions have been made that some of this stock may move north across the British Columbia-Alaska border to mix with the Kah Shakes herring stock. The available biological data do not indicate any large scale unidirectional movement of herring north or south, although tagging has indicated some small scale movement between these areas. For the purposes of this study, the Prince Rupert assessment region is assumed to represent a distinct biological entity.

Assessment Models

The two assessment models, an “escapement” model and an “age-structured” model, used to assess British Columbia herring stocks, have been described in detail by Schweigert et al. (1997). The escapement model, used since 1984, relies on an assessment of the escapement from the fishery plus total catch to estimate the pre-fishery spawning stock biomass

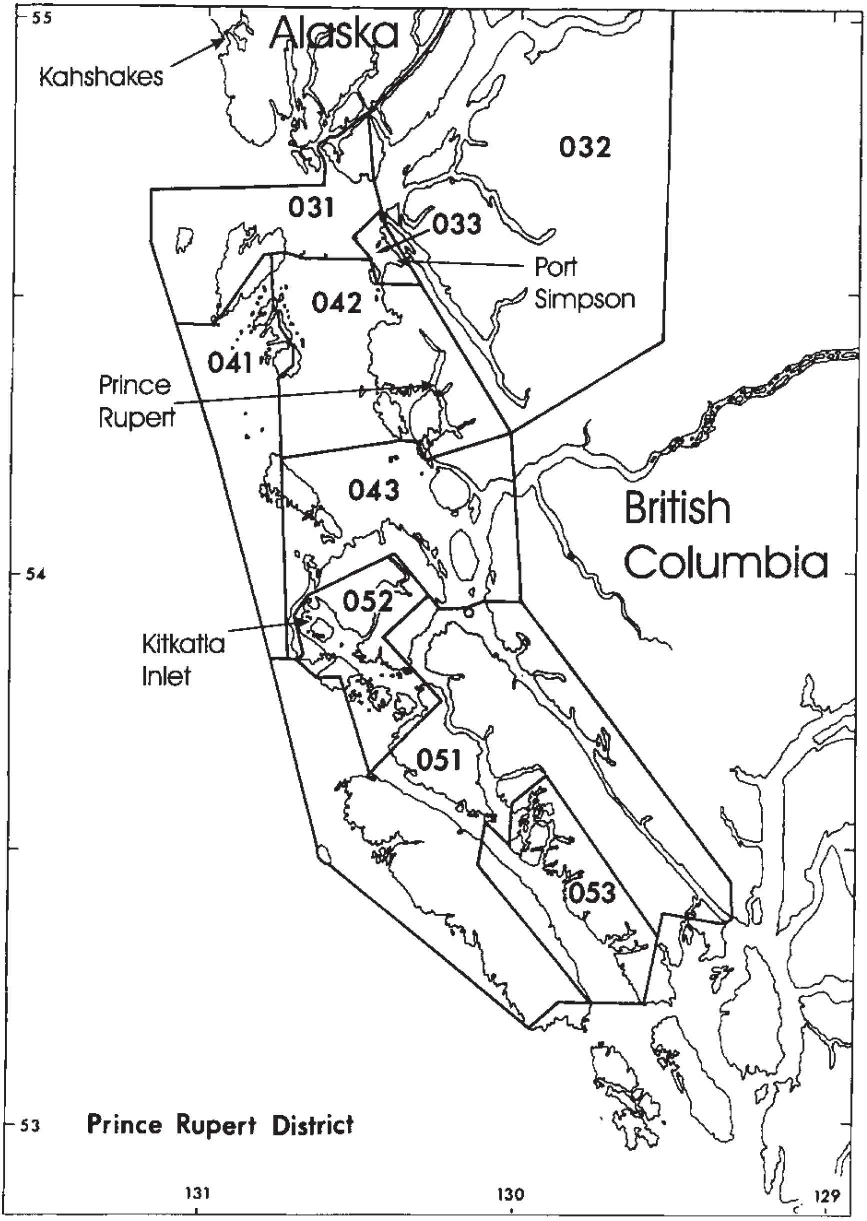


Figure 1. Herring sections in the Prince Rupert District which define the assessment region and delineate the geographic subunits for data aggregation.

for each assessment region. The estimated escapement to each region is derived from information on spawn deposition. Total egg deposition for each spawning bed or location is calculated as the product of: total length of the bed, an estimate of the observed or adjusted width of the bed, and an estimate of the egg density determined from a visual assessment of the number of layers of eggs on the available vegetation substrates (Schweigert and Stocker 1988, Schweigert 1993). Prior to 1988 all estimates of these parameters of the egg beds were determined from the surface of the water either from a skiff using a grapple or on foot at low tide. Subsequently, scuba surveys of spawning locations have been used to quantify the egg deposition and historical surface observations have been adjusted to emulate more recent diving data. Egg deposition estimates for all spawn beds are summed within each assessment region and the total egg deposition is converted to tons of spawning fish based on an estimate of 100 eggs per gram of herring (Hay 1985).

The age-structured model has been used to assess B.C. herring stocks since 1982. Ongoing revisions to the model have made it more consistent with the life history of herring and the fisheries. The current version uses auxiliary information on spawning escapement, separates catch and age composition data by gear type, and includes availability parameters to estimate partial recruitments to the spawning stock. Model parameters are estimated simultaneously using a maximum likelihood method. The model is implemented in the C++ programming language using AUTODIF for derivative calculations (Otter Research Ltd. 1992).

Two types of fishing gear are used commonly in B.C. herring fisheries. Seine nets are assumed to be nonselective while gillnets are selective for larger, older fish. Herring fisheries have concentrated primarily on fish which are on, or migrating to, the spawning grounds. Therefore, the relative availability of age classes to nonselective gear should be equivalent to the partial recruitment of age classes to the spawning stock. The age-structured model explicitly separates availability (partial recruitment) and gear selectivity. Seine and gillnet fisheries are temporally separate so catch and age composition are partitioned into fishing periods, separating data for the different gears. Three fishing periods are modeled. The first period encompasses all catch prior to the spring roe herring fisheries. This includes reduction fishery catches prior to 1968 and the winter food and bait fisheries since 1970. Most of this catch was taken by seine gear although small amounts were caught with trawl nets (which are also assumed to be non-size selective). The second fishing period includes all seine roe herring catch and the third period includes all gillnet roe herring catch.

Briefly, assume T_{ij} is the total number of fish in age class j at the beginning of season i, where season is equivalent to year, and λ_{ij} is the proportion of age j fish which are available to the fishery. Then, N_{ij1} the total number of age class j fish which are available at the start of period 1 in season i is given by

$$N_{ij1} = \lambda_{ij} T_{ij} \tag{1}$$

To model the fishing process a form of the catch equations, which models fishing and natural mortality as continuous processes over time period *r*, is used:

$$C_{ijr} = \frac{F_{ijr}}{F_{ijr} + M_r}\left[1 - \exp(-F_{ijr} - M_r)\right]N_{ijr}, \tag{2}$$

and, for $r < p$

$$N_{ijr+1} = N_{ijr}\exp(-F_{ijr} - M_r), \tag{3}$$

where

C_{ijr} is the catch of age class *j* in season *i* for period *r*,

F_{ijr} is the fishing mortality of age class *j* in season *i* for period *r*,

M_r is the natural mortality for period *r*,

N_{ijr} is the number of fish in age class *j* in season *i* for period *r*,

p is the number of fishing periods ($p = 3$),

n is the number of seasons ($n = 47$),

k is the number of age classes ($k = 9$).

$N_{i+1,j+1,1}$ is defined by equation (4) where for $j + 1 < k$

$$T_{i+1,j+1} = N_{ijp}\exp(-F_{ijp} - M_p) + T_{ij}(1 - \lambda_{ij})\exp\sum_r -M_r\,, \tag{4}$$

In the model the last age class, *k*, accumulates all fish aged *k* and older. To reduce the number of parameters to be estimated assumptions are made about the form of the availabilities and mortalities. The availabilities are formulated to increase with age and are set to 1 for age 6+ and older. For age 3+ to 5+ the availabilities are assumed to be constant between years. Because the proportion of age 2+ fish that are mature appears to vary among years (Haist and Stocker 1985) and some reduction fisheries targeted on immature 1+ fish, the availabilities for these two age classes are estimated for each year for which there is age-composition data with the exception of the final year.

For the selective gillnet fishery (i.e., fishing period 3), fishing mortality is separated into age selectivity and fishing intensity components. For nonselective fisheries (i.e., fishing periods 1 and 2) only fishing intensity parameters are estimated. A natural mortality parameter, $M_\bullet$, is also estimated. Natural mortality for the three fishing periods is modeled as,

$$\begin{aligned} M_1 &= 0.95M_\bullet \\ M_2 = M_3 &= 0.025M_\bullet \end{aligned} \tag{5}$$

Additional structure is built into the model through the inclusion of annual spawn data (spawn index, I_i). Spawning occurs at the end of the season so the number of spawners at age j in season i (G_{ij}) is estimated by

$$G_{ij} = N_{ijp} \exp(-F_{ijp} - M_r) \tag{6}$$

and the spawning stock biomass, which is assumed to be equivalent to egg production, in season i (R_i) is

$$R_i = \sum_j w_{ij} G_{ij}, \tag{7}$$

where w_{ij} is the average weight-at-age j in season i. The errors in the spawn index observations (I_i) are assumed to be multiplicative so that

$$I_i = qR_i \exp(\xi_i), \tag{8}$$

where q is a spawn conversion factor and ξ_i is a normally distributed random variable with mean 0 and variance σ_1^2.

Data Considerations

Biological data on total catches, weight at age, age structure, and egg deposition are available from 1951 to present. From these data, catch at age and total egg deposition data were determined and total stock abundance calculated from age-structured and escapement model analyses as described above. The results of applying the age-structured model to three alternative sets of input data are presented below. First, consider all the available data assuming no data biases which is the current model. Second, accept that the catches and associated biological samples collected from fall fisheries since the beginning of the roe fishery in 1972 are limited and may not be representative of the selectivity pattern that occurred during the earlier reduction fishery or of the current roe fishery, and therefore remove these data from the analysis. Third, delete potentially biased samples as above but also adjust the spawn index or egg deposition data since 1988 for a possible 25% overestimate of spawn deposition from diving surveys relative to previous surface based spawn assessments due to increased survey effort and efficiency.

Results

Trends in abundance for the age-structured model estimate of stock size and for the escapement model or the egg deposition index are presented in Fig. 2. The results indicate good agreement between the escapement and the age-structured estimates for all areas except the Queen Charlotte Islands and Prince Rupert. It is generally agreed that because of its remoteness, some spawnings in the Queen Charlotte Islands were not completely surveyed in the 1970s and early 1980s. However, with the beginning of diving surveys in 1988, spawn surveys are felt to be representative of the total spawn deposition in the area and these results agree with the age-structured assessment since that time. The same is not true of the Prince Rupert assessment where there is poor agreement with the escapement estimate from the early 1970s through the mid-1980s and again in recent years. Because of its proximity to the city of Prince Rupert it is assured that no significant spawning events would have been missed during the annual spawn assessments throughout this period (Pers. comm., L. Gordon, Dept. Fisheries and Oceans, Pt. Alberni, B.C.). Hence, the discrepancy between the two abundance indicators has been difficult to reconcile.

The results of conducting a retrospective analysis for these data from the age-structured model further supports these findings (Fig. 3). The assessment of stock size for all areas except Prince Rupert has been consistent over time in a retrospective sense although there has been a tendency to underestimate abundance in the two southern assessments. For Prince Rupert the agreement between the models has become worse for the period from 1970 to 1985 as new data have been added to the analysis, and somewhat better thereafter as the estimate of stock size by the age-structured model has declined.

In an effort to reconcile the two estimates of abundance the biological sampling data were reviewed. It was noted that the spatial and temporal distribution of the sampling effort had changed substantially over time. During the reduction fishery prior to 1972, sampling effort was fairly evenly distributed throughout the assessment region with most samples coming from sections 41 or 42 and 51 (Fig. 1). With the advent of the roe fishery in the 1970s, which focused on the spawning grounds, sampling effort became more concentrated spatially in sections 33 or 42 and 52 with the majority coming from the latter area. In the current analysis it is assumed that all available biological samples are representative of the mature population and so contribute equally to the estimate of catch at age. Therefore, to identify possibly unrepresentative or biased samples, a variety of ordination methods were investigated as a means to detect outliers in the data series for each year that might be contributing to the poor fit of the age-structured model to the catch-at-age data. However, because there was no objective basis for rejecting any particular sample in the data series this approach was not very useful for identifying biased biological

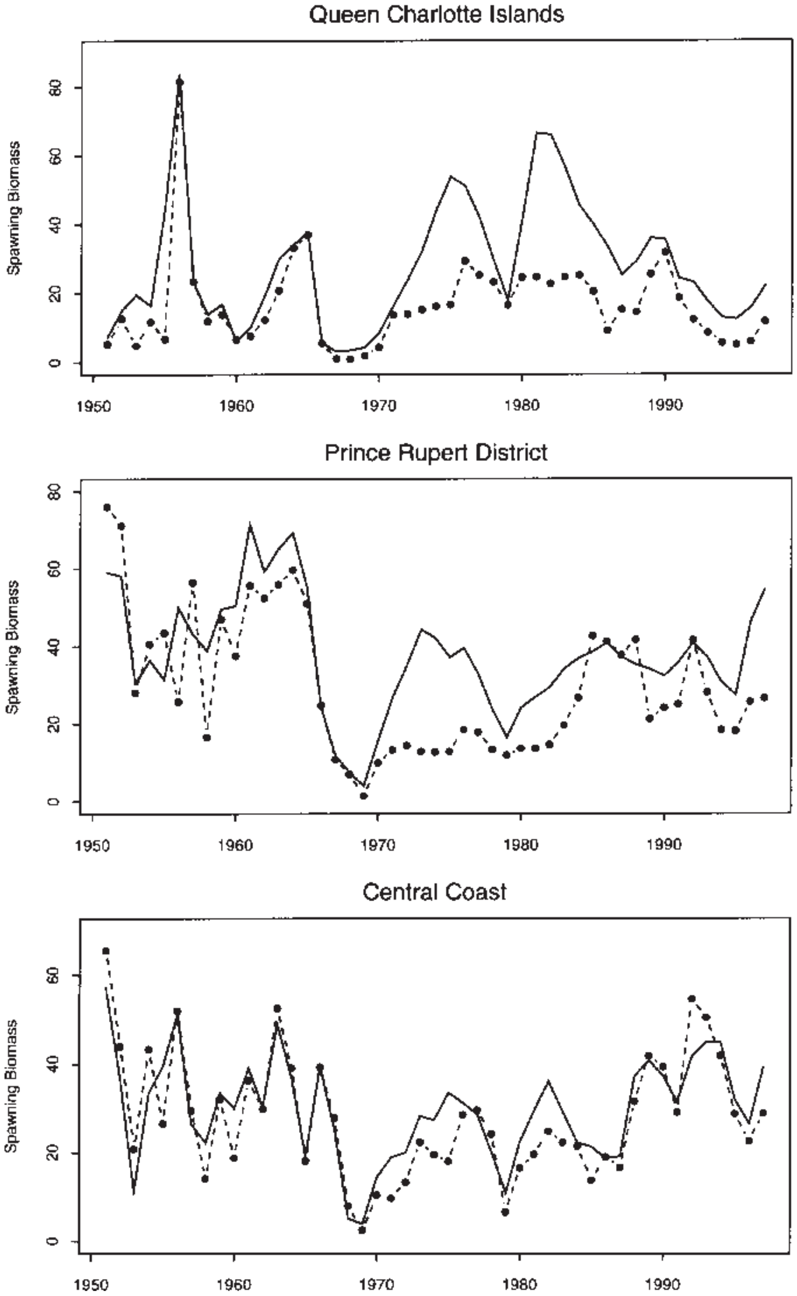
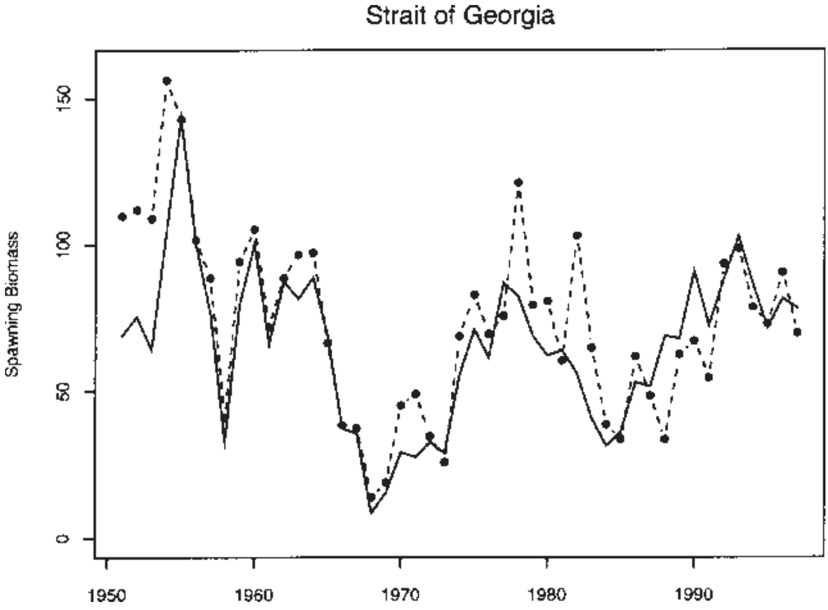


Figure 2a. Estimates of pre-fishery spawning stock biomass (1,000 t) from age-structured (solid line) and escapement model (dashed dotted line) analyses for northern B.C. herring stock assessment regions, 1951-1997.



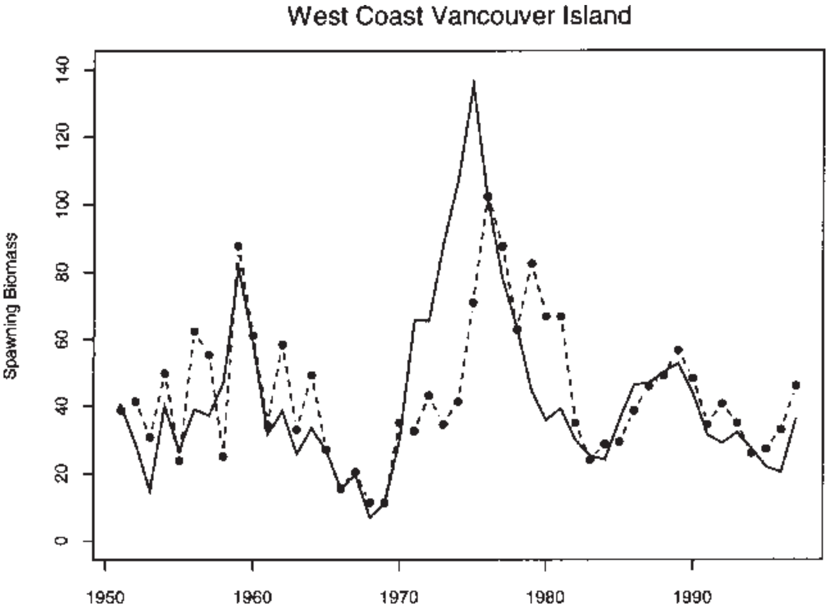


Figure 2b. Estimates of pre-fishery spawning stock biomass (1,000 t) from age-structured (solid line) and escapement (dashed dotted line) model analyses for southern B.C. herring stock assessment regions, 1951-1997.

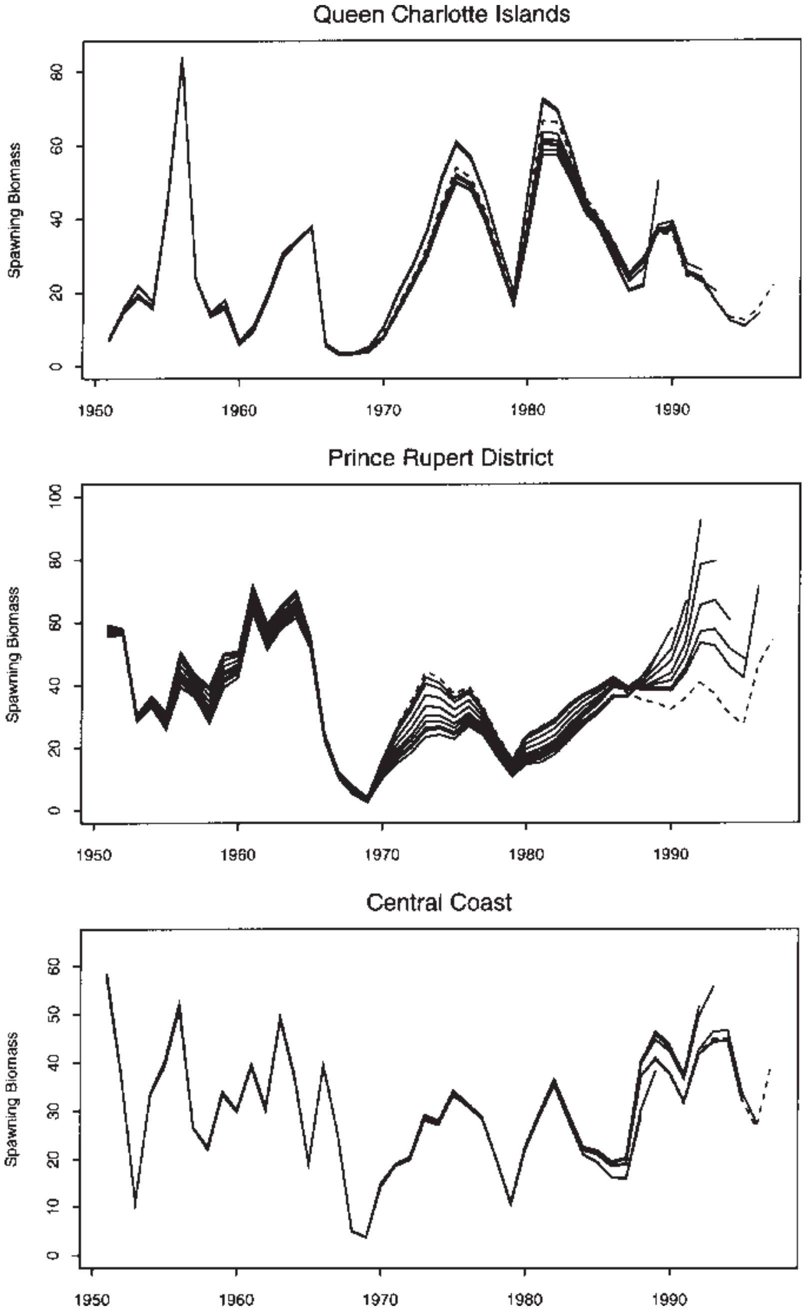


Figure 3a. Retrospective analysis of estimated spawning biomass (1,000 t) from age-structured analysis for northern B.C. herring stocks, 1951-1997.

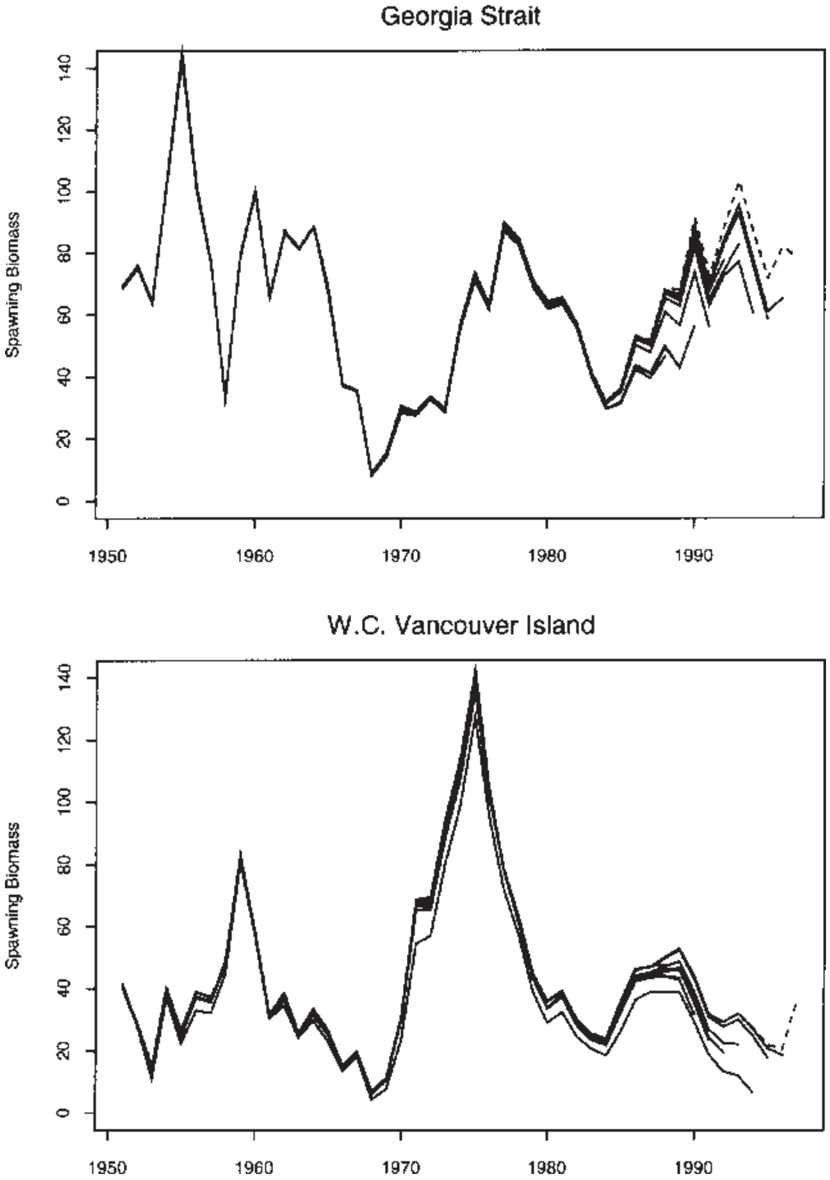


Figure 3b. Retrospective analysis of estimated spawning biomass (1,000 t) from age-structured analysis for southern B.C. herring stocks, 1951-1997.

samples. Instead, the estimated age compositions from all the available samples for each of the three fishing periods (reduction seine, roe seine, and roe gillnet) were examined. These data highlighted significant differences between the estimates of age structure of the catch depending on when and where the samples were collected (Fig. 4). It is evident that, for a number of the years for which samples are available from both a fall food or bait fishery and a subsequent spring seine roe fishery, the same population is not being sampled (1973, 1976-1979, 1982, 1985, 1986). In fact, the graphs indicate that the fall food and bait samples do not reveal the dominant cohorts of fish that are found in the spawning run the following spring and as a result are apparently biasing the age-structured assessment of stock abundance and trend.

To examine the impact of the fall samples, these data were dropped from the data set and the catches taken in these fisheries were combined with the roe catches the following spring. The roe fishery sampling data then became the only source of catch-at-age information for the age-structured analysis. The result of this analysis is shown in Fig. 5 where the abundance time series is compared to the current analysis and the escapement model estimate. The resulting fit to the catch-age data from 1970 to 1980 remains similar although abundance is markedly lower from 1972 to 1974. However, from 1981 to 1990 the estimate of abundance increases significantly and subsequently more closely approximates the magnitude and trend in the escapement model estimate of stock size.

In addition to this analysis, the impact of a possible systematic bias in the spawn index due to a change in data collection techniques was investigated. Beginning in 1988, the assessment of egg deposition has increasingly relied on diving surveys since much of the spawn occurs subtidally and was likely underestimated by earlier surface based surveys. While historical data have been adjusted to account for this effect (Schweigert et al. 1997), it is conceivable that the use of scuba based surveys has discovered more egg beds than was the case previously, and as a consequence the spawn index is biased upward due to increased survey effort and efficiency. To evaluate this possibility the spawn index data used in the age-structured analysis were arbitrarily deflated by 25% to render them more similar to what might have been observed using historical survey procedures, and the analysis was repeated on the reduced sampling data set (no food and bait samples after 1972). The biomass trajectory for this analysis is also presented in Fig. 5 and closely follows that for the reduced sampling data set until about 1985 when the fit more closely approximates the escapement model abundance series.

The residuals from the current age-structured analysis for the fit to the age composition data are shown in Fig. 6a. For comparison, the residuals from the fit to the reduced sampling data set are presented in Fig. 6b. A similar plot for the adjusted spawn analysis does not result in any detectable difference from Fig. 6b and so is not presented here. It is clear from Fig. 6b that removal of the food and bait samples from the analysis

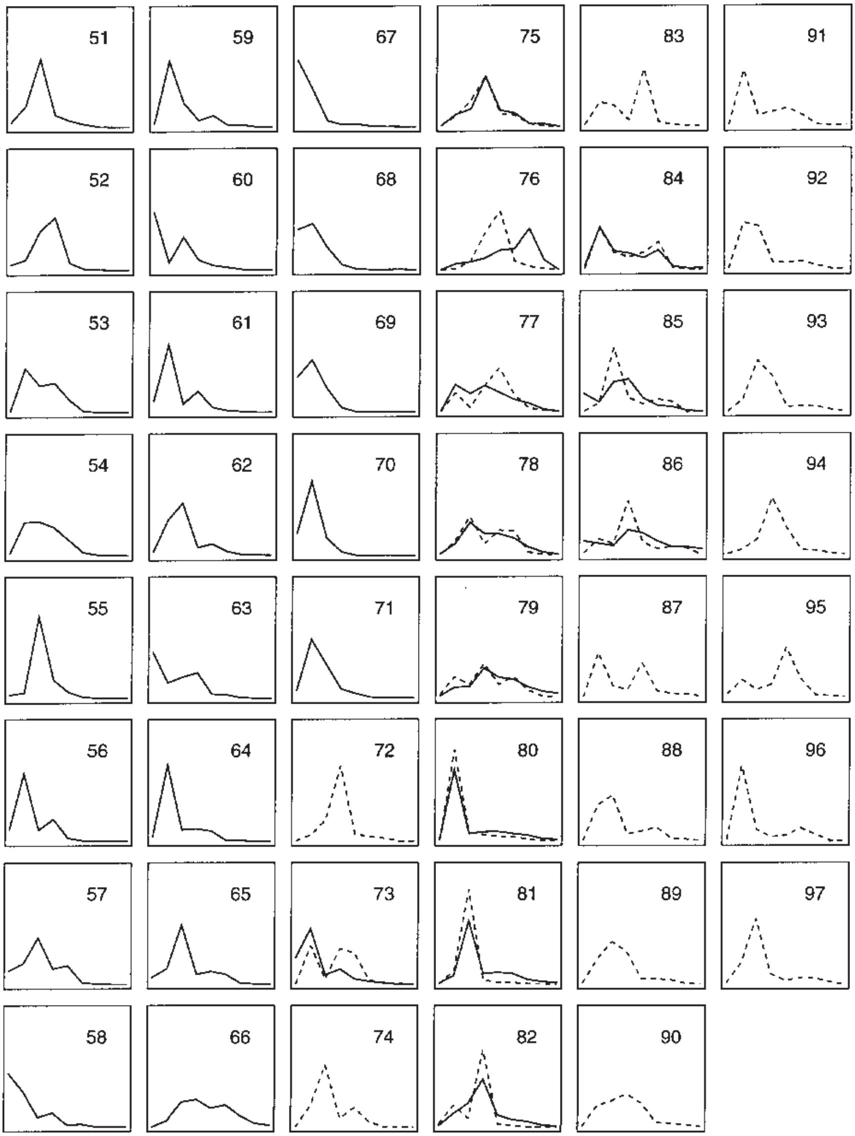


Figure 4. Estimated age composition by year for period 1 (fall fisheries = solid line) and period 2 (spring fisheries = dashed line) in the Prince Rupert region, 1951-1997. Ordinate ranges from zero to one and abscissa from ages 1 through 10.

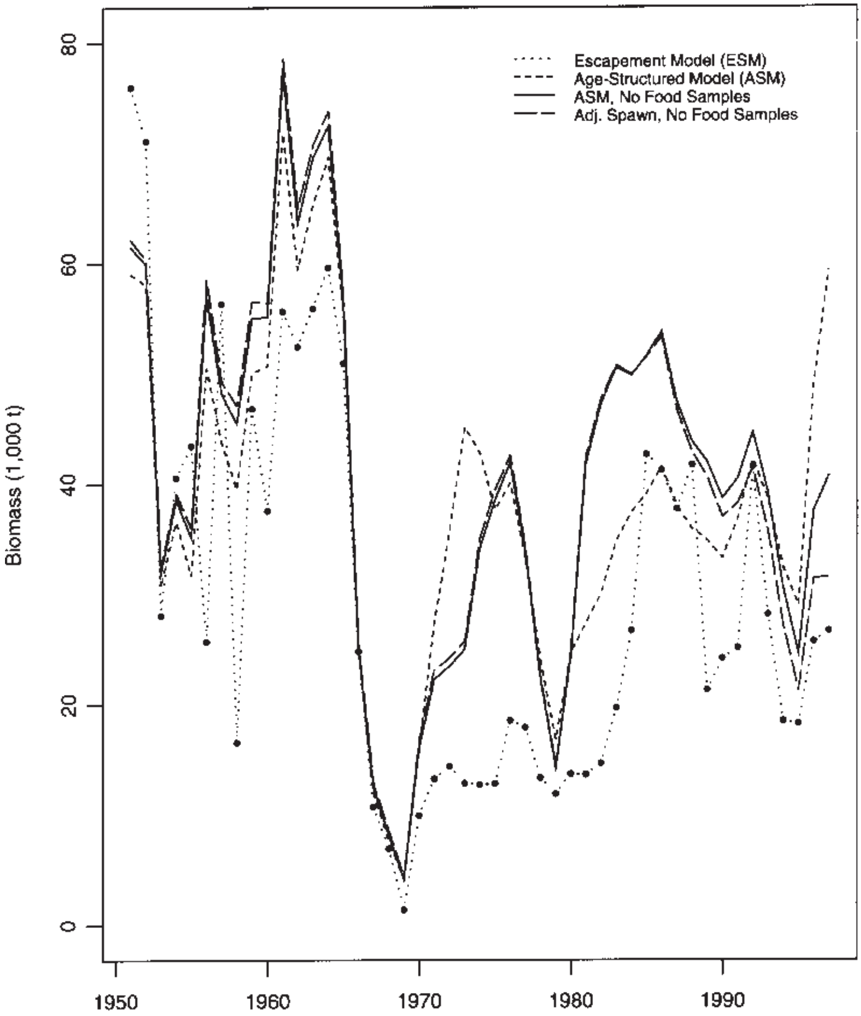


Figure 5. Estimates of mature biomass for age-structured model analysis, for an adjusted data set removing fall food fishery samples since 1972, and for an adjusted spawn index without fall food samples, relative to the escapement model estimate for 1951-1997.

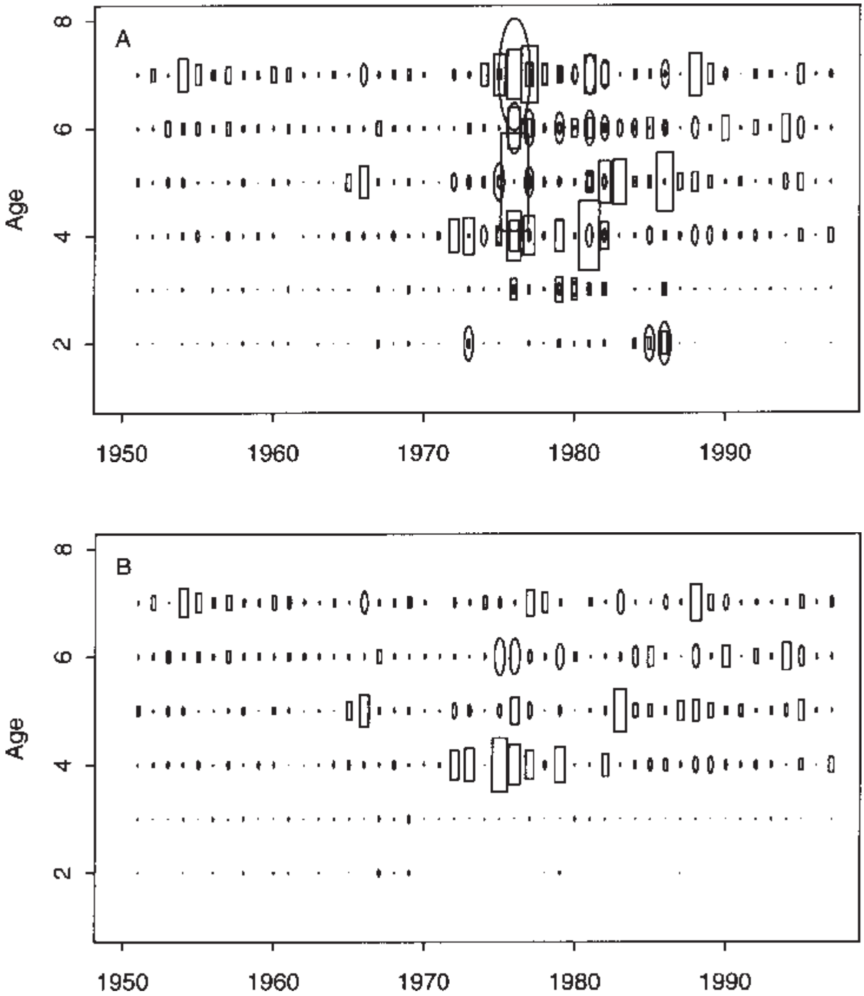


Figure 6. Standardized residuals from the age-structured analysis using (A) the current model and (B) the reduced data set without fall food fishery samples after 1972. Positive residuals are elliptical and negative values appear as rectangles. Multiple residuals for year and age combinations reflect the presence of both fall and spring samples in A, from which fall samples are removed in B.

has markedly improved the fit to the age structure data for this assessment region.

The lognormal residuals from the fit of the age-structured model estimate of egg production (equation 8) and the observed spawn index data are plotted in Fig. 7. Results indicate a generally improved fit to the spawn index over much of the time series, except the period 1981-1984, for the reduced sampling data set relative to the base analysis. There is also a marked deviation of the model estimates from the spawn index for the 1950s which may have other explanations such as incomplete spawn surveys, poor catch reporting, or fisheries targeting on mixtures of fish destined for other areas. The adjustment of the spawn index to assess the impact of changes in survey methodology is equivocal since there is no marked improvement in the fit to the spawn index; in fact, the fit is actually poorer in many of the recent years.

The estimated minimum function value determined for each analysis is presented in Table 1 for all assessment regions and for the two alternate analyses described here. It is apparent that the current assessment for Prince Rupert has provided a significantly poorer fit to the available data than has the same model in any of the other areas. The fit of the Prince Rupert assessment region without the food and bait sampling data is significantly better than the base assessment and in the range of that observed for other stocks. The fit of the model to the reduced data set and the adjusted spawn is marginally better than to the reduced data set alone. For comparison, the fit to the current data series without the roe seine sampling data from section 52 also shows a very marginal improvement in the fit for this stock relative to the base assessment. The modified analysis was presented in a recent assessment to adjust for apparent overweighting of sampling from section 52 relative to other sections in the region in recent years (Schweigert et al., unpubl. manuscript).

Discussion

An assessment of the status of any resource relies on the development of a quantitative model of the processes that describes changes in the dynamics of the system with time. The model may then be tested against the available data to estimate the parameters of interest. An important question that frequently arises in the evaluation of the adequacy of the model is whether the mathematical structure or relationships hypothesized in constructing the model are supported by the available data or if perhaps there are errors or biases in these data or there are insufficiencies in the model structure. Generally, it is not possible to decide between these alternatives although there are subjective approaches for evaluating alternative competing models, such as comparing the function values of a class of models at their respective maxima or comparing the parameter estimates from competing models against independent information that was not used in constructing the model. Often, however, there are systematic

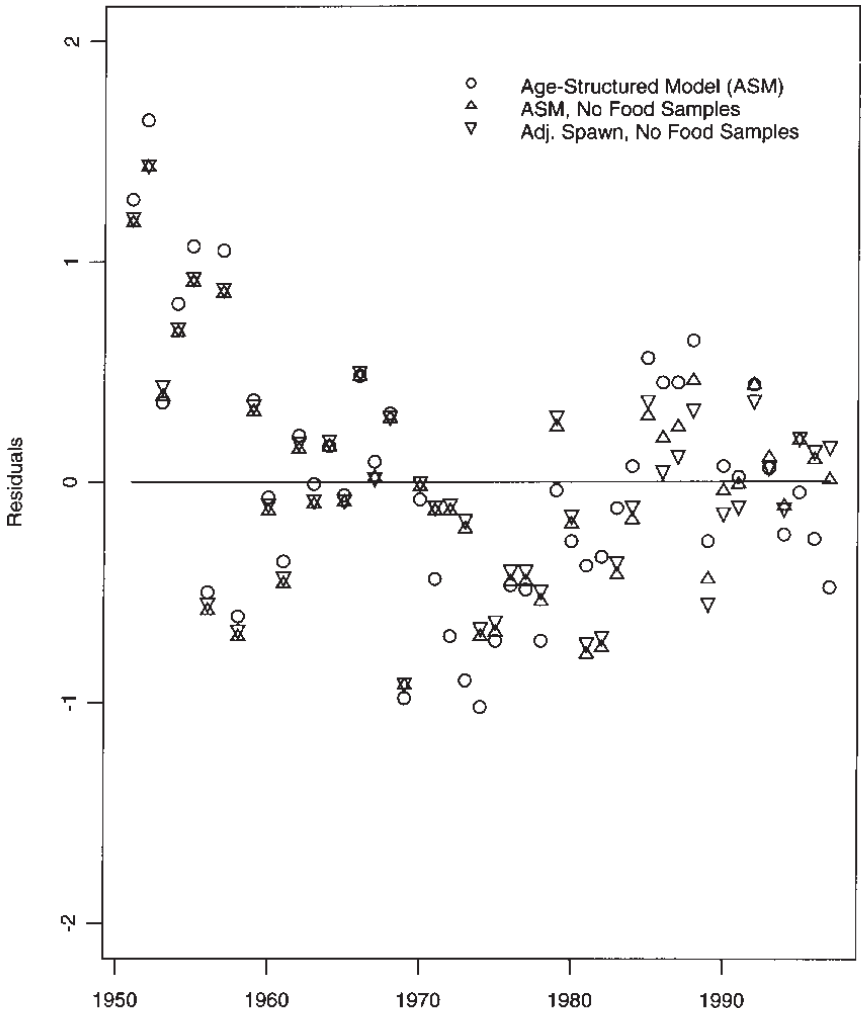


Figure 7. Estimated lognormal residuals from the relationship between the observed spawn index and the total egg production derived by the current age-structured model, for a reduced model without fall food fishery samples since 1972, and for an adjusted spawn index without fall food samples.

Table 1. **Estimated minimum function values from age-structured model analysis for British Columbia herring stocks from 1951 to 1997.**

	Function value	No. of parameters
Queen Charlotte Islands	582	199
Prince Rupert base model	2,652	233
Reweight sect. 52 samples	2,222	233
No fall food samples (NFFS)	1,152	221
Adjusted spawn (NFFS)	1,148	221
Central coast	520	222
Georgia Strait	1129	249
W.C. Vancouver Is.	935	207

biases in the available biological data used to assess the resource that have effects on model performance but are difficult to interpret or to adjust for in model restructuring. The present study attempts to discern some of these effects for one of the British Columbia herring stocks using a combination of methods.

The age-structured model used in the assessment of British Columbia herring stocks evolved over a period of years through a trial and error process which attempted to encapsulate the important population dynamics processes of herring populations and the associated fisheries (Haist et al. 1988). The available data on catch at age and relative stock size in the form of spawn index information are fitted to the model to estimate population parameters of interest and deviations or residuals from the fitted model examined to attempt to understand the process and measurement errors of the model. In the application of this model to the five major herring populations in British Columbia it was evident that an acceptable fit to the spawn index occurred in the three southern populations. A reasonable fit also occurred for the Queen Charlotte Islands for all but the early roe fishery period when there was apparently limited effort devoted to assessing spawn deposition. The fit is improved in recent years where extensive diving surveys have provided good spawn survey coverage. Taken together this result suggests that the current model structure adequately represents the population dynamics of British Columbia Pacific herring and that the lack of fit to the Prince Rupert area data is indicative of data anomalies which are inconsistent with the underlying model assumptions about herring population dynamics.

Richards et al. (1997) recently commented on the applicability of graphical approaches for interpreting catch-at-age analyses and some of these suggestions are adopted here. For example, a simple time series plot of the estimated age structure of the population each year as determined by samples from the various fisheries revealed anomalies between the population structure as evidenced in the fall and that observed the following spring (Fig. 4). While it might theoretically be possible to alter the structure of the age-structured model to describe the change in age composition from fall to spring as a function of changes in distribution or migration, the underlying biology is not well understood. It is also possible that the measurement error in the samples which were collected in the fall is so large that these samples do not accurately reflect the true underlying population age structure which is caught in the spring roe fishery, or in fact yields a biased estimate of these parameters. The retrospective analysis conducted for all five herring stocks lends some support to this interpretation because it demonstrates that for all but the Prince Rupert population the stock reconstructions have remained relatively stable, as new data are included in the data set and stock trajectories are re-estimated (Fig. 3). For Prince Rupert the stock reconstruction varies widely suggesting that there are serious inconsistencies in the underlying age composition and consequently catch at age which make it impossible to consistently determine relative cohort strengths over time. As a consequence the stock reconstruction varies widely depending on which new set of data are added to the model and how they alter the estimated fishing intensity parameters. The approach adopted here to deal with this problem was an expedient one since it was determined that the number of samples available for estimating age composition for the years of greatest discrepancy were minimal. Hence, it was simplest to ignore these samples and rely on the estimates of age composition determined from the more extensive sampling program during the spring roe fisheries, to estimate the catch at age for the entire food and roe catch throughout the roe fishery period beginning in 1972.

It is also recognized that this simple adjustment to the data series, although significantly improving model fit, has not accounted for all of the residual variation in the fit to the data series. For example, it is evident that as the fisheries have focused on roe, sampling has become more concentrated geographically which has had the effect of differentially weighting the age structure and consequently the catch-at-age data to those areas or components of the stock that are most readily available to the sampling gear. Such a practice could also be introducing a bias into the sampling data and the effects of differentially weighting some of the sampling data should be further investigated. It has also been demonstrated that the older fish spawn first so the age structure of the herring run changes during the season (Hay 1985) which may bias the sampling data depending on the synchronicity of the herring spawning times and the sampling program. An attempt to adjust for this effect was presented in a previous assessment for this stock and involved removing all seine fishery samples collected from section 52 since 1970 from the data set (Schweigert et al., unpubl. manuscript). The effect of such a simple reweighting of the catch-at-age data was not particularly effective, providing only a marginally better fit to the data than the base assessment (Table 1). This also raises the question of how important an effect this might be in other assessment regions. The only other area in which a significant fall food and bait fishery occurs is the Strait of Georgia, and it will be necessary to examine these data in a similar manner to determine whether adjustments to the basic data might also improve the model fit for that stock.

An interesting question that remains is why the age-structured and escapement model biomass estimates differ so significantly in the early roe fishery period if survey effort and coverage were thorough and consistent throughout this period. Prior to 1970 harvest rates were generally very large and removed upward of 50% of the available stocks in most areas of the coast (Hourston 1980). As a result most of the biomass estimated during this period was determined by the catch since there was relatively less spawning escapement, and so the stock abundance estimate should be fairly well determined. Throughout the roe fishery period, harvest rates have been maintained close to the 20% target (Schweigert et al. 1997) and as a result the biomass estimate is heavily dependant on the estimate of egg deposition. A possible explanation remains that significant herring spawning occurred in unusual areas or at unusual times and as a consequence went unreported and unsurveyed. Another possibility is that the procedures used to convert these data to estimates of spawning biomass are inappropriate for this area, although they seem to work quite well in the other assessment regions. The only other possibility seems to be that there are outstanding biases in the age structure information during the earliest years of the roe fishery that need to be resolved or the model structure altered to account for them if a reliable catch age analysis for this stock is to be obtained.

In summary, based on the compare and contrast strategy adopted here to evaluate model process error, there is no evidence of any serious failure to meet the structural assumptions of the existing catch-age model used in British Columbia herring stock assessments. Removal of a few apparently biased samples of population age structure resulted in a markedly improved fit of the model to catch-age data and spawn index information without the need to alter model formulation for the Prince Rupert assessment region. The base assessment model applied to a reduced data series provides a comparable fit to that for the other B.C. herring stocks and should be used in the future assessment of this resource.

References

Fournier, D., and C.P. Archibald. 1982. A general theory for analyzing catch at age data. Can. J. Fish. Aquat. Sci. 39:1195-1207.

Haegele, C.W. 1991. Returns from anchor taggings of herring in British Columbia, 1979 to 1985. Can. Data Rep. Fish. Aquat. Sci. 582. 129 pp.

Haist, V., and L. Rosenfeld. 1988. Definitions and codings of localities, sections, and assessment regions for British Columbia herring data. Can. Manuscr. Rep. Fish. Aquat. Sci. 1994. 123 pp.

Haist, V., and M. Stocker. 1985. Growth and maturation of Pacific herring *(Clupea harengus pallasi)* in the Strait of Georgia. Can. J. Fish. Aquat. Sci. 42(Suppl. 1):138-146.

Haist, V., J.F. Schweigert, and D. Fournier. 1988. Stock assessments for British Columbia herring in 1987 and forecasts of the potential catch in 1988. Can. Tech. Rep. Fish. Aquat. Sci. 1990. 63 pp.

Hay, D.E. 1985. Reproductive biology of Pacific herring *(Clupea harengus pallasi).* Can. J. Fish. Aquat. Sci. 42(Suppl. 1):111-126.

Hourston, A.S. 1980. The decline and recovery of Canada's Pacific herring stocks. Rapp. P.-V. Reun. Cons. Int. Explor. Mer 177:143-153.

Otter Research Ltd. 1992. AUTODIF: A C++ array extension with automatic differentiation for use in nonlinear modeling and statistics. Otter Research Ltd., Nanaimo, B.C.

Richards, L.J., J.T. Schnute, and N. Olsen. 1997. Visualizing catch-age analysis: A case study. Can. J. Fish. Aquat. Sci. 54:1646-1658.

Schweigert, J.F. 1993. Review and evaluation of methodology for estimating Pacific herring egg deposition. Bull. Mar. Sci. 53(2):818-841.

Schweigert, J.F., and M. Stocker. 1988. A new method for estimating Pacific herring stock size from spawn survey data and its management implications. N. Am. J. Fish. Manage. 8:63-74.

Schweigert, J.F., C. Fort, and L. Hamer. 1997. Stock assessments for British Columbia herring in 1996 and forecasts of the potential catch in 1997. Can. Tech. Rep. Fish. Aquat. Sci. 2173. 73 pp.

Stevenson, C.J. 1954. The movement of herring in British Columbia waters as determined by tagging, with a description of tag and tag recovery methods. In: Herring tagging and results. Cons. Int. Explor. Mer Spec. Sci. Meeting 55. 36 pp.

A Stochastic Implementation of an Age-Structured Production Model

Victor R. Restrepo and Christopher M. Legault
University of Miami, Rosenstiel School of Marine and Atmospheric Science, Miami, Florida

Abstract

An age-structured production model (ASPM) has been used by the International Commission for the Conservation of Atlantic Tunas for the assessments of some stocks. That model is sometimes preferable to traditional biomass-based surplus production models because it can accommodate age-structured indices of relative abundance. The traditional ASPM estimates a deterministic stock-recruitment relationship, a property that may result in inconsistencies between the estimated level of recruitment and the observed level of catches. In this work we relax the deterministic assumption by incorporating stochasticity in recruitment around the deterministic predictions as a first-order, autoregressive time-series process. We use data for western Atlantic bluefin tuna to contrast the deterministic and stochastic model fits.

Introduction

Age-structured production models (ASPM) have been used in assessments carried out by the International Commission for the Conservation of Atlantic Tunas (ICCAT) in the past, particularly for albacore tuna *(Thunnus alalunga)* in the South Atlantic and for bluefin tuna *(Thunnus thynnus)* tuna in the Western Atlantic. Conceptually, ASPMs fall somewhere between simple biomass-based production models (e.g., Schaefer 1957, Prager 1994) and the more data-demanding sequential age-structured population analyses (Megrey 1989). Simple production models estimate parameters related to carrying capacity, rate of productivity, biomass at the start of the time series, and coefficients that scale indices of abundance to the absolute magnitude of biomass. ASPMs estimate similar parameters but make

---

Current address for Christopher M. Legault is National Marine Fisheries Service, Southeast Fisheries Science Center, 75 Virginia Beach Drive, Miami, FL 33149

explicit use of age-structured computations, rather than lumped-biomass ones, and directly estimate parameters of a stock-recruitment relationship. Their main advantage over simpler production models is that they can make use of age-specific indices of relative abundance.

During the 1996 assessments of Atlantic bluefin tuna (ICCAT 1997), an implementation of an ASPM (Restrepo 1997) was criticized because its estimates of recruitment in recent years could not be reconciled against available catch estimates for extant cohorts. This could be in part attributed to the fact that the ASPM implementation estimated a deterministic stock-recruitment relationship and, therefore, estimated levels of recruitment would vary smoothly over time at the deterministic predictions. The main objective of this paper is to extend the ASPM formulation in order to account for stochastic recruitment. The same approach to modeling recruitment can easily be incorporated into the more general "integrated approaches" (e.g., Fournier and Archibald 1982, Deriso et al. 1985, Methot 1990) or into tuned sequential population analyses as was done by Porch (in press).

Model Formulations and Estimation

Throughout this paper, a Beverton and Holt (1957) type of stock recruitment relationship (SRR) is assumed.

Deterministic Formulation

The deterministic model is similar to that of Punt (1994), which was based on ideas presented by Hilborn (1990), with modifications to account for multiple fisheries. Details of the fitting procedure can be found in Punt (1994) and only a brief summary is presented here, along with some concepts needed for the presentation.

The resource dynamics are modeled by a forward population projection using standard fishery equations. The projection includes a "plus" group (ages p and older) and each year's recruitment is obtained from a deterministic stock-recruitment relationship. The fishing mortality values needed to project the population forward are computed based on total yields and selectivities that are input and assumed exact. The Restrepo (1997) formulation allows for gear-specific yields and selectivities because, in multi-gear cases, it is easier to obtain annual input selectivities by gear, rather than for all gears combined. The estimation consists of finding the values of the stock-recruitment relationship parameters that result in stock size trajectories which best explain observed indices of relative abundance by minimizing the negative log-likelihood

$$-\ln(L_1) = \sum_i \left[\frac{n_i}{2} \sum_t \ln\left(\sigma_{i,t}^2\right) + \sum_t \frac{1}{2\sigma_{i,t}^2} \left(I_{i,t} - \hat{I}_{i,t}\right)^2 \right] \qquad (1)$$

where t denotes year, i denotes each available series of relative abundance comprised of n_i observations, the last term is for the squared differences between observed (I) and predicted ($\hat{I}$) indices of abundance (these could be in logarithmic units if a lognormal error is assumed), and $\sigma_{i,t}^2$ are abundance index variances that are either input or estimated. Additional parameters include catchability coefficients that scale the indices to absolute abundance (or biomass), and possibly a parameter related the initial conditions. In many ASPM applications, the initial conditions are not reliably estimated and the initial age structure is fixed, e.g. by assuming that the stock was in a virgin state at the start of the time series if the series extends back to the onset of fishing.

The Beverton and Holt SRR is usually described by the equation

$$R_{t+1} = \frac{\alpha S_t}{\beta + S_t}, \tag{2}$$

where R is the number of recruits and S is the product of numbers, maturity and fecundity, summed over all ages. For simplicity, we refer to S as "spawning biomass," which is often used as a proxy for reproductive output. During estimation, the ASPM uses a different parameterization, following Francis (1992). It consists of defining a "steepness" parameter, τ, which represents the fraction of the virgin recruitment (R_0) that is expected when S has been reduced to 20% of its virgin level: $R = \tau R_0$ when $S = \gamma/5$, where γ is the virgin spawning biomass. The SRR is thus defined in terms of steepness and virgin biomass, two parameters for which initial values are somewhat easier to guess than α and β. (For a Beverton-Holt relationship, virgin biomass should generally be of similar magnitude to the largest observed yields, while steepness should fall somewhere between 0.2 and 1.0). R_0 is computed as the ratio of virgin spawning biomass to spawning biomass per recruit in the absence of fishing, $(S/R)_{F=0}$,

$$R_0 = \frac{\gamma}{(S/R)_{F=0}}, \tag{3}$$

and α and β are given by

$$\alpha = \frac{4\tau R_0}{5\tau - 1}, \text{ and} \tag{4}$$

$$\beta = \frac{\gamma(1-\tau)}{5\tau - 1} \tag{5}$$

The computation of statistics such as maximum sustainable yield (MSY) and related benchmarks (e.g., S_{MSY}, F_{MSY}) follows the procedure summarized by Shepherd (1982). Conditional on a given F (including an overall selectivity pattern), equilibrium spawning biomass, recruitment and yield are computed as (for the Beverton and Holt SRR),

$$S_F = \alpha(S/R)_F - \beta, \tag{6a}$$

$$R_F = \frac{S_F}{(S/R)_F}, \text{ and} \tag{6b}$$

$$Y_F = R_F(Y/R)_F\ , \tag{6c}$$

where $(S/R)_F$ and $(Y/R)_F$ are the spawning biomass per recruit and yield per recruit values resulting from exploitation at F. F_{MSY} would be computed by searching for the F that maximizes equation (6c). Note that, if the selectivity pattern changes over time, then the computed *MSY*-related values will also change as a result of changes in the per-recruit computations.

A set of useful benchmarks for management is based on the so-called "spawning potential ratio," *SPR*, defined as the spawning biomass per recruit obtained under a given F, divided by that under $F = 0$ (Goodyear 1993). An important benchmark is the *SPR* corresponding to the slope of the SRR at the origin, i.e., at the point when the stock is expected to "crash." From equations (3) to (5) it follows that this SPR_{crash} is a function of steepness:

$$SPR_{crash} = \frac{(S/R)_{crash}}{(S/R)_{F=0}} = \frac{\beta/\alpha}{\gamma/R_o} = \frac{1-\tau}{4\tau}. \tag{7}$$

Deterministically, any fishing mortality that results in an *SPR* lower than SPR_{crash} is not sustainable.

Stochastic Formulation

A stochastic ASPM requires that a recruitment value be estimated for every year. In this work, we have chosen to constrain the (log) recruitment deviations from the equilibrium SRR to follow a first-order autoregressive (AR[1]) process. The population projection equations are as in the deterministic model, except that recruitment is estimated as

$$N_{1,t} = R_0 e^{\upsilon_t}, \tag{8}$$

where the notation $N_{a,t}$ denotes population numbers for age a in year t. Thus, recruitment is estimated as deviations from a virgin level. Instead of estimating γ and τ directly as parameters, the model estimates γ and all the υ_t. R_0 is computed from equation (3). Besides the initial population size, these are all the parameters that are needed to project the population forward. The AR[1] process is incorporated by assuming that the recruitment estimates thus obtained vary around the expected SRR as

$$N_{1,t+1} = R_{t+1}e^{\varepsilon_{t+1}} = \frac{\alpha S_t}{\beta + S_t} e^{\varepsilon_{t+1}} \tag{9}$$

with $\varepsilon_{t+1} = \rho\varepsilon_t + \eta_{t+1}$, where $|\rho| < 1$, and the η have zero expectation and variance equal to σ_η^2. In equation (9) we distinguish between recruitment values estimated as parameters in the search ($N_{1,t}$) and those predicted from the estimated stock-recruitment relationship (R_t).

The negative log-likelihood for these recruitment "residuals" is (Seber and Wild 1989):

$$-\ln(L_2) = \frac{n_t}{2}\ln(\sigma_\eta^2) - \frac{1}{2}\ln(1-\rho^2) + \frac{1}{2\sigma_\eta^2}\left[(1-\rho^2)\varepsilon_1^2 + \sum_{t=2}^{n_t}\left(\varepsilon_t - \rho\varepsilon_{t-1}\right)^2\right] \tag{10}$$

where n_t is the number of years in the analysis. It follows from equation (9) that the residuals are computed as

$$\varepsilon_{t+1} = \ln\left(N_{1,t+1}\right) - \ln\left(\frac{\alpha S_t}{\beta + S_t}\right). \tag{11}$$

Computation of the first residual would depend on the initial conditions. For example, in a virgin state, it would be $\varepsilon_1 = \ln(N_{1,1}) - \ln(R_0)$. Note that α and β in equations (9) and (11) could be computed from knowledge of virgin biomass and steepness (see equations 4 and 5). However, only the former is being estimated directly as a parameter. To include steepness as an additional parameter to be directly estimated by the search would confound the information contained in R_0 and γ. Our approach is to replace α and β in the SRR of equation (11) by a function of those parameters being estimated in the search, and steepness. From equations (4) and (5) it follows that deterministic recruitment can also be predicted by

$$R_{t+1} = \frac{4R_0 S_t \tau}{\tau(5S_t - \gamma) - S_t + \gamma}, \text{ such that} \tag{12}$$

$$\varepsilon_{t+1} = \ln\left(N_{1,t+1}\right) - \ln\left(\frac{4R_0 S_t \tau}{\tau(5S_t - \gamma) - S_t + \gamma}\right). \tag{13}$$

We use these relationships in order to solve for τ, noting that, for a given ρ and σ_η^2, equation (10) will be at a minimum when

$$\sum_{t=2}^{n_t-1}\left\{\ln\left(N_{1,t+1}\right)-\ln\left[\frac{4R_0S_t\tau}{\tau(5S_t-\gamma)-S_t+\gamma}\right]-\rho\ln\left(N_{1,t}\right)+\rho\ln\left[\frac{4R_0S_{t-1}\tau}{\tau(5S_{t-1}-\gamma)-S_{t-1}+\gamma}\right]\right\}^2 \quad (14)$$

is also at a minimum. Thus, in every iteration in the search, a subprocedure is invoked to find the τ that minimizes (14). Having thus calculated the steepness (and, consequently, α and β), the log-likelihood of equation (10) is added to the overall objective function.

It remains to be mentioned what to do about the parameters ρ and σ_η^2. In theory, there is a potential for these to also be estimated. In practice, however, it is unlikely that the data will contain so much information as to determine the relative contribution from recruitment variability with respect to the variability in the index values. In our limited experience with this model, it appears that these values should be controlled by the analyst in much the same way as contributions to the likelihood from different data sources are weighted externally in other assessment methods (e.g., Deriso et al. 1985). Lower σ_η^2 values will result in lower stochasticity in recruitment, while higher σ_η^2 values will allow recruitment to fluctuate more widely. A value of $\rho = 0$ would assume no autocorrelation between successive recruitment deviations. Empirical studies such as those of Beddington and Cooke (1983) and Myers et al. (1990) may yield information about likely ranges of values for ρ and σ_η^2 for species groups.

Estimating the initial conditions for the stochastic model can be problematic, as with the deterministic model. Estimating the age structure in year $t = 1$ would not generally be an option unless there were age-specific relative abundance data for the start of the series. Thus, using a long time series of data extending to the onset of fishing, and assuming an initial equilibrium state at $S = \gamma$, remains a useful option. In this paper we calculate a stable age structure resulting from a pre-series recruitment that is fixed. We fix υ_0 and set the starting population sizes as

$$N_{2,1} = R_0 e^{\upsilon_0} e^{-M_1}, \quad (15a)$$

$$N_{a,1} = N_{a-1,1}e^{-M_{a-1}} \text{ for ages } a = 3 \text{ to } p\text{-}1, \text{ and} \quad (15b)$$

$$N_{p,1} = \frac{N_{p-1,1}e^{-M_{p-1}}}{1-e^{-M_p}} \text{ for the plus group, } p. \quad (15c)$$

This alternative allows the initial age structure to be either higher or lower than that corresponding to an equilibrium virgin state. The parameter υ_0 could be estimated in the search procedure as well. If it is, it may be desirable to place a penalty on how much it can alter the initial biomass away from γ. This could be accomplished with the term

$$-\ln(L_3) = \frac{\ln(\sigma_\upsilon^2)}{2} + \frac{\left[\ln(S_1) - \ln(\gamma)\right]^2}{2\sigma_\upsilon^2} \tag{16}$$

where σ_υ^2 is fixed by the analyst.

Estimation of the stochastic model parameters for any given data set then requires several choices associated with how much recruitment can fluctuate around its deterministic predictions and about the initial conditions. In addition to choices about variances (σ_η^2, σ_υ^2 and possibly $\sigma_{i,t}^2$), the log-likelihood components could be given different emphases (λ) to obtain model estimates by minimizing:

$$-\ln(L_T) = -\ln(L_1) - \lambda_2 \ln(L_2) - \lambda_3 \ln(L_3). \tag{17}$$

Projections using the stochastic model results incorporate the autoregressive process in recruitment by first generating normally distributed random deviates, η_t, with mean zero and variance σ_η^2, and then setting the autocorrelated residuals as

$$\varepsilon_{t+1} = \rho\varepsilon_t + \eta_{t+1} \tag{18}$$

for use in equation (9). The value of the first residual to "seed" the projections is available from the model fit, i.e. from equation (11).

An Example: Western Atlantic Bluefin Tuna

We used the stochastic ASPM model with the same data set that was used for western Atlantic bluefin tuna during the 1996 ICCAT assessment (ICCAT 1997). The input data consists of yields for four fisheries starting in 1950 (Fig. 1), and of nine indices of abundance assumed to be lognormally distributed, with associated coefficients of variation (available in ICCAT 1997). Fishery selectivities were fixed by the assessment working group for different time periods when regulations substantially affected fishing operations and were based on average patterns in the results from an age-structured VPA (Gavaris 1988, Powers and Restrepo 1992). The biological parameters used in the ASPM runs are given in Table 1.

We carried out a series of estimations making different assumptions about some of the inputs that are fixed (Table 2), or estimating different numbers of parameters, in order to examine the sensitivity of the results to these choices. Run 1 is a deterministic one, conducted for comparison to the stochastic model. Runs 2 to 5 examine the effect of estimating different numbers of recruitment deviations from R_0 (equation 8). In these, the υ_t were fixed at values obtained from Run 1 for years when they were not being estimated, thus forcing the initial pattern of recruitment to be smooth, like in the deterministic fits. Runs 6 and 7 allowed the initial

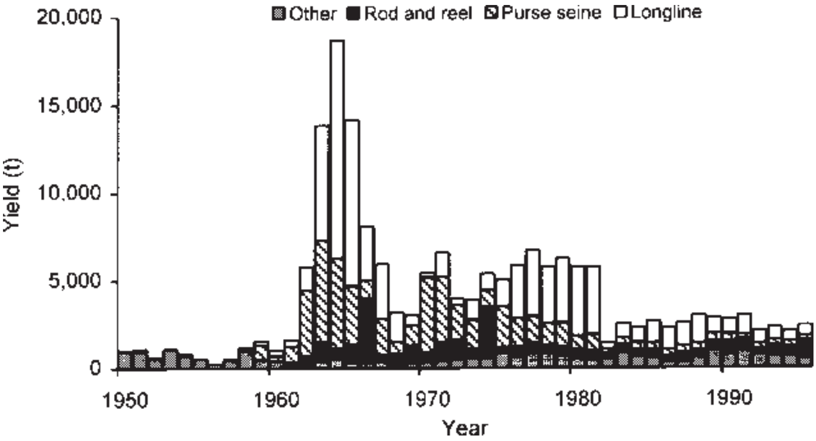


Figure 1. Reported landings for western Atlantic bluefin tuna, by fishery.

biomass to be arbitrarily smaller or larger than the virgin one. Run 8 was performed to examine the effect of assuming no autocorrelation in recruitment residuals ($\rho = 0$), and Runs 9 and 10 used different choices for the assumed variance in recruitment residuals. In all runs, we set σ_v^2 equal to 0.1, and set λ_2 and λ_3 equal to 1.

For Runs 1 and 2, we approximated the variance of outputs of interest by bootstrapping (Efron 1982) the indices of abundance from their residual distributions. The algorithm used for estimation was that of Nelder and Mead (1965), with multiple restarts.

Results and Discussion

A summary of outputs for Runs 1-10 is given in Table 3, where we focus on some quantities that are of interest to management. Some general conclusions can be derived from these results:

1. Maximum sustainable yield (*MSY*) resulting from recent (1990-1995) fishery selectivity patterns is estimated at 6,000-7,000 t, or about 2.5 times recent yield levels. The potential *MSY* has changed considerably over time as a result of changes in the overall selectivity effected by the various gears (Fig. 2). (Note that long-term changes in *MSY* for a stock could be due to a number of other factors, including changes in survival or productivity. In this case, due to the model's assumptions and use of a constant natural mortality rate, the estimated shifts in *MSY* levels can only be attributed to changes in selectivity). Recent *MSY* levels are higher than those attainable with the 1960s-1970s selectivities, probably as a result of ICCAT minimum size recommendations and subsequent demise of purse seine harvest of small bluefin tuna (Fig. 1).

Table 1. Biological parameters assumed for bluefin tuna, held constant for all years.

Age	Midyear weight (kg)	Spawning biomass (kg)	Natural mortality (yr^{-1})
1	4.4	0.0	0.14
2	10.6	0.0	0.14
3	20.5	0.0	0.14
4	37.5	0.0	0.14
5	56.6	0.0	0.14
6	80.6	0.0	0.14
7	111.3	0.0	0.14
8	143.0	131.3	0.14
9	176.1	164.8	0.14
10+	292.8	267.2	0.14

Table 2. Age-structured production model runs performed.

Run	Type	υ_t estimated	$\upsilon_{t=0}$	ρ	σ_η^2
1	Deterministic	n/a	n/a	n/a	n/a
2	Stochastic	1950-1995	0	0.5	0.2
3	Stochastic	1960-1995	0	0.5	0.2
4	Stochastic	1970-1995	0	0.5	0.2
5	Stochastic	1980-1995	0	0.5	0.2
6	Stochastic	1950-1995	–0.1	0.5	0.2
7	Stochastic	1950-1995	0.1	0.5	0.2
8	Stochastic	1950-1995	0	0	0.2
9	Stochastic	1950-1995	0	0.5	0.05
10	Stochastic	1950-1995	0	0.5	0.4

See text for description of symbols.

Table 3. Results from the age-structured production model runs (defined in Table 2). R^2 is the coefficient of determination for the nine fitted indices of abundance.

Run	R^2	MSY (t)	S_{96}/γ	S_{96}/S_{MSY}	$R_{92\text{-}95}/R_0$	$F_{8+,92\text{-}95}$	$\frac{SPR_{92\text{-}95}}{SPR_{\text{crash}}}$
1	0.23	6,079 (3.2)	0.028 (14.7)	0.083 (13.9)	0.21 (7.6)	0.35 (12.1)	0.75
2	0.28	6,502 (8.2)	0.035 (23.5)	0.107 (21.3)	0.16 (22.7)	0.37 (19.3)	1.30
3	0.27	7,290	0.029	0.091	0.170	0.46	1.04
4	0.26	6,665	0.033	0.101	0.181	0.44	1.03
5	0.25	6,369	0.036	0.105	0.150	0.35	0.88
6	0.28	6,776	0.033	0.099	0.156	0.38	0.91
7	0.28	6,490	0.036	0.105	0.152	0.35	0.88
8	0.27	7,357	0.028	0.087	0.198	0.46	1.07
9	0.26	7,224	0.030	0.092	0.212	0.41	1.05
10	0.30	5,943	0.040	0.113	0.131	0.31	0.82

The results present estimates of maximum sustainable yield (*MSY*); 1996 biomass (S_{96}) relative to virgin (γ) and *MSY* levels; recent recruitment (R) relative to virgin levels (R_0); recent average fishing mortality rate for ages 8+; and equilibrium spawning potential ratio (*SPR*) expected under current exploitation relative to the limit at which the stock is expected to collapse (SPR_{crash}). Numbers in parentheses are coefficients of variation (percent).

2. The resource is perceived to be severely depleted, as has been suggested by ICCAT assessments for well over one decade. Current spawning biomass is estimated to be at 3-4% of virgin levels, or 8-11% of the level that can sustain *MSY*. Recent recruitment is also low, estimated to be 13-21% of virgin levels for the most recent years.

3. The estimated annual fishing mortality rate for the spawners (ages 8 and older) in recent years ranges from 0.31 to 0.46 for the various runs. The ratio of current equilibrium spawning potential ratio (*SPR*) to SPR_{crash} (equation 7) suggests that this level of fishing mortality is very high or even unsustainable (ratio values smaller than 1.0 are not sustainable).

Figure 3 contrasts various trajectories between Runs 1 and 2. The recruitment tendencies are generally similar although, as expected, the stochastic fit shows more variability. Tendencies in spawning biomass and fishing mortality relative to *MSY* levels are very similar, with the stochastic model estimating a larger decline in biomass between the early 1960s and early 1980s. The stock-recruitment trajectories are in general agreement between the two models (Fig. 3).

Figure 4 compares trends in estimated recruitment and spawning biomass since 1970 for the deterministic (Run 1) and stochastic (Run 2) ASPM fits, and for the VPA used in the ICCAT assessment of 1996. The 1970s biomass estimated by the VPA is substantially higher than that from either ASPM fit. However, the VPA estimates were obtained with an algorithm that is sensitive to the choices made to model fishing mortality in the plus group (Hiramatsu 1992). Thus, the estimates of biomass from the VPA

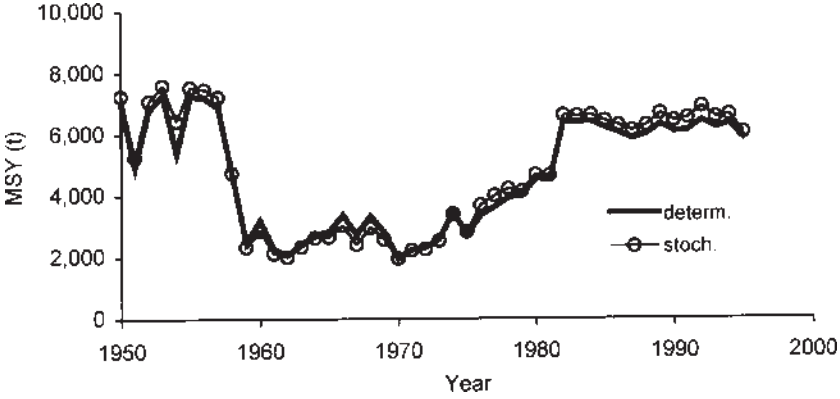


Figure 2. Estimates of potential maximum sustainable yield (MSY) conditional on year-specific selectivity patterns for the deterministic (Run 1) and stochastic (Run 2) ASPM fits.

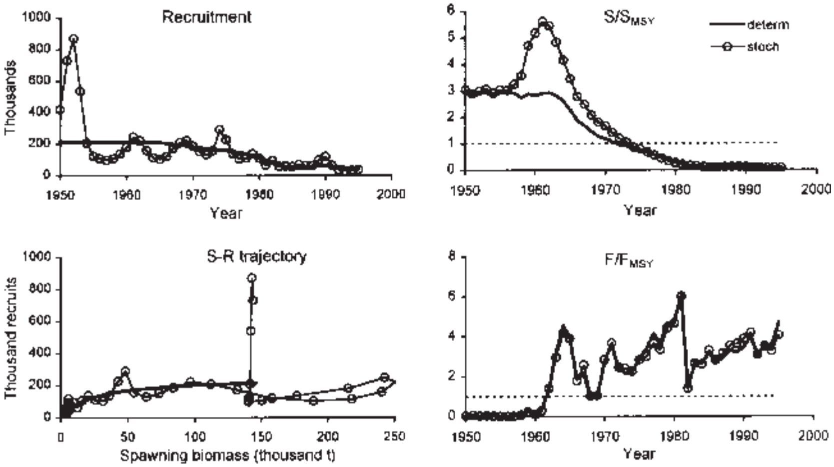


Figure 3. Estimated trajectories of recruitment (upper left), stock-recruitment pairs (lower left), the ratio of spawning biomass to that at MSY (S/S_{MSY}, upper right), and the ratio of fishing mortality to that resulting in MSY (F/F_{MSY}, lower right). Results are shown for the deterministic (Run 1) and stochastic (Run 2) ASPM fits.

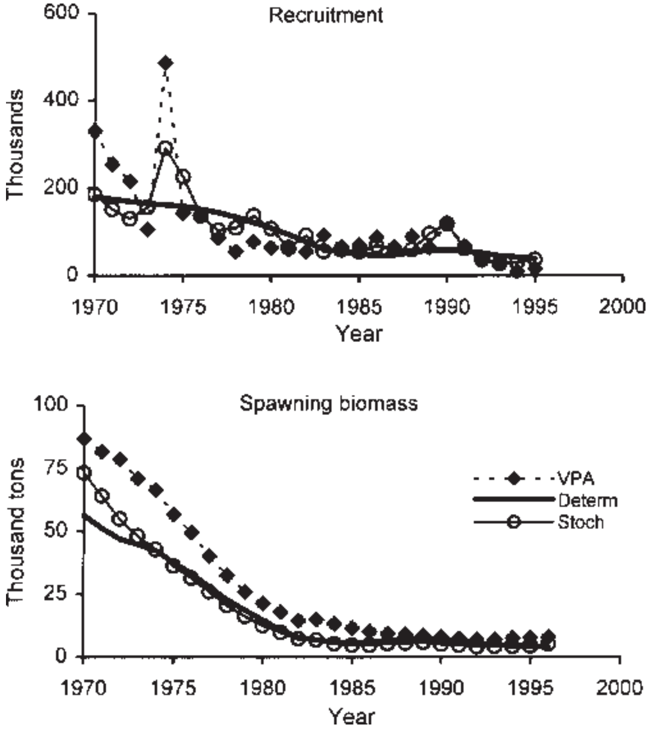


Figure 4. Comparison of recruitment and spawning biomass estimates since 1970 for the VPA assessment, and the deterministic (Run 1) and stochastic (Run 2) ASPM fits.

could change substantially if the plus group were handled differently. Similarly, biomass estimates from the ASPM could be affected by changes in inputs that are assumed to be known exactly (e.g., weights at age, selectivities).

Recruitment levels for the two most recent years estimated with the ASPM are about 3 times larger than those from the VPA (Fig. 4), a difference that is potentially important in making projections into the future. We conducted deterministic projections with the results from Run 1, and stochastic projections with the results from 100 bootstraps of Run 2. In the latter case, for each bootstrap set of starting conditions, the population was projected forward at different catch levels using 100 realizations of future stochastic recruitments (equation 18). The results of these projections are shown in Fig. 5 as the median, 10th and 90th percentiles of spawning biomass and recruitment (also shown are the deterministic

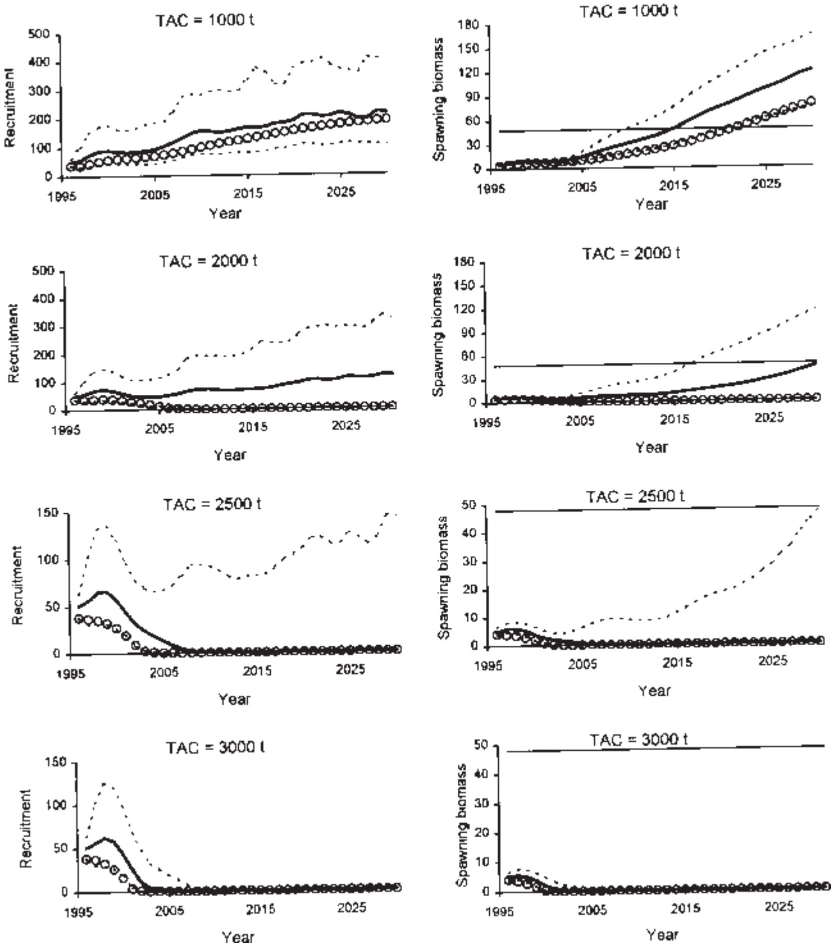


Figure 5. Projected recruitment (in thousands) and spawning biomass (in thousand t) under four different constant catch levels (TAC). The thick solid and dashed lines give the median and approximate 80% confidence intervals from the stochastic (Run 2) projections. The circles correspond to the deterministic (Run 1) projections. The horizontal solid line is the spawning biomass at MSY level (S_{MSY}), used as a recovery target.

projections from Run 1). We note that these projections are not directly comparable to those of ICCAT because some of the choices we made differ from those made by the assessment working group (see ICCAT 1997 for details).

Results of the stochastic ASPM projections suggest that constant catch levels above 3,000 t are not sustainable and that catch levels below 1,000 t are (Fig. 5). Projections made with intermediate levels of catch indicate that yields of 2,500 t are probably not sustainable, but 2,000 t may be (as indicated by the median, Fig. 5). However, there is great uncertainty in these projections as suggested by the approximate 80% confidence intervals. This uncertainty is due to the variance of estimated 1996 stock sizes as well as future recruitment. The deterministic projection is clearly more pessimistic than the stochastic one. The difference can be attributed in part to the deterministic model's failure to estimate relatively large 1988-1989 year classes (Fig. 4), which is reflected into lower adult stock sizes at the start of the projection period.

The results presented above indicate that the general perception about the depleted status of bluefin tuna obtained from application of the ASPM is quite similar to that reached by ICCAT using other methods. The main methodological focus of our work, to extend the basic ASPM in order to better model interannual changes in recruitment, also makes the stock projections more compatible with those of ICCAT (1997) (although there are differences in the median estimates of projected stock trajectories, some of the choices we made for making the projections differ from ICCAT's). We do not propose that the stochastic ASPM presented here be used as the primary assessment method for the bluefin stock because the ASPM requires several stringent assumptions, such as known selectivities for the fishing gears. However, we recommend that the ASPM be used as an auxiliary tool to examine the stock's trajectory in relation to MSY levels.

Acknowledgments

Support for this study was provided by the University of Miami's (UM) Cooperative Unit for Fisheries Education and Research through National Oceanic and Atmospheric Administration Cooperative Agreement NA67RJ0149, and by the National Marine Fisheries Service (NMFS) Office of Science and Technology. The conclusions presented here are entirely the authors' and are not necessarily endorsed by UM or NMFS. We are grateful to Clay Porch, Joe Powers, Gerry Scott, Steve Turner, and two anonymous reviewers for helpful comments on an earlier draft.

References

Beddington, J.R., and J.G. Cooke. 1983. The potential yield of fish stocks. FAO Fish. Tech. Pap. 242. 50 pp.

Beverton, R.J.H., and S. Holt. 1957. On the dynamics of exploited fish populations. Reprinted in 1993 by Chapman and Hall, London. 533 pp.

Deriso, R.B., T.J. Quinn II, and P.R. Neal. 1985. Catch-age analysis with auxiliary information. Can. J. Fish. Aquat. Sci. 42:815-824.

Efron, B. 1982. The jackknife, the bootstrap and other resampling plans. Society for Industrial and Applied Mathematics Monograph 38. 92 pp.

Fournier, D., and C.P. Archibald. 1982. A general theory for analyzing catch at age data. Can. J. Fish. Aquat. Sci. 39:1195-1207.

Francis, R.I.C.C. 1992. Use of risk analysis to assess fishery management strategies: A case study using orange roughy *(Hoplostethus atlanticus)* on the Chatham Rise, New Zealand. Can. J. Fish. Aquat. Sci. 49:922-930.

Gavaris, S. 1988. An adaptive framework for the estimation of population size. Can. Atl. Fish. Sci. Adv. Comm. (CAFSAC) Res. Doc. 88/29. 12 pp.

Goodyear, C.P. 1993. Spawning stock biomass per recruit in fisheries management: foundation and current use. In: S.J. Smith, J.J. Hunt, and D. Rivard (eds.), Risk evaluation and biological reference points for fisheries management. Can. Spec. Publ. Fish. Aquat. Sci. 120:67-81.

Hilborn, R. 1990. Estimating the parameters of full age structured models from catch and abundance data. Int. North Pac. Fish. Comm. Bull. 50:207-213.

Hiramatsu, K. 1992. Possible biases in the VPA estimates of population sizes of the plus group. Int. Comm. Cons. Atl. Tunas, Coll. Vol. Sci. Pap. 39:497-502.

ICCAT. 1997. 1996 SCRS detailed report on bluefin tuna (Report of the 1996 SCRS Bluefin Tuna Stock Assessment Session). Int. Comm. Cons. Atl. Tunas, Coll. Vol. Sci. Pap. 46(1):1-186.

Megrey, B.A. 1989. Review and comparison of age-structured stock assessment models from theoretical and applied points of view. Amer. Fish. Soc. Symp. 6:8-48.

Methot, R.D. 1990. Synthesis model: An adaptable framework for analysis of diverse stock assessment data. Int. North Pac. Fish. Comm. Bull. 50:259-277.

Myers, R.A., W. Blanchard, and K.R. Thompson. 1990. Summary of North Atlantic fish recruitment 1942-1987. Can. Tech. Rep. Fish. Aquat. Sci. 1743. 108 pp.

Nelder, J.E., and R. Mead. 1965. A simplex method for function minimization. The Computer Journal 7:308-313.

Porch, C.E. In press. Autocorrelated stock-recruitment penalties applied to the 1996 SCRS assessment of West Atlantic bluefin tuna. Int. Comm. Cons. Atl. Tunas, Coll. Vol. Sci. Pap. (working document SCRS/97/60).

Powers, J.E., and V.R. Restrepo. 1992. Additional options for age-sequenced analysis. Int. Comm. Cons. Atl. Tunas, Coll. Vol. Sci. Pap. 39:346-354.

Prager, M.H. 1994. A suite of extensions to a nonequilibrium surplus-production model. Fish. Bull. U.S. 92:374-389.

Punt, A.E. 1994. Assessments of the stocks of Cape hakes, *Merluccius* spp. off South Africa. S. Afr. J. Mar. Sci. 14:159-186.

Restrepo, V.R. 1997. An implementation of the age-structured production model with application to West Atlantic bluefin tuna fisheries. Int. Comm. Cons. Atl. Tunas, Coll. Vol. Sci. Pap. 46(2):348-353.

Schaefer, M.B. 1957. A study of the dynamics of the fishery for yellowfin tuna in the eastern tropical Pacific Ocean. Inter-Am. Trop. Tuna Comm. Bull. 2:247-285.

Seber, G.A.F., and C.J. Wild. 1989. Nonlinear regression. J. Wiley, New York. 768 pp.

Shepherd, J.G. 1982. A versatile new stock-recruitment relationship for fisheries, and the construction of sustainable yield curves. J. Cons. Cons. Int. Explor. Mer 40:67-75.

When Lengths Are Better Than Ages: The Complex Case of Bocaccio

Stephen Ralston
National Marine Fisheries Service, Southwest Fisheries Science Center, Tiburon, California

James N. Ianelli
National Marine Fisheries Service, Alaska Fisheries Science Center, Seattle, Washington

Abstract

Bocaccio *(Sebastes paucispinis)* has historically been the most important rockfish harvested in the California groundfish fishery. A stock assessment of bocaccio in the Eureka-Monterey-Conception area indicated that in 1996 spawning biomass was 5-10% of that present in 1970. This finding was based on the application of the Stock Synthesis model to a split-sex population, assuming length-dependent gear selectivities for four distinct fisheries. A variety of fishery-dependent and fishery-independent data sources were used to model population biomass, including (1) landings from the trawl, setnet, hook-and-line, and recreational fisheries, (2) trawl catch-at-age data for the period 1980-1985 using surface otolith ages, (3) trawl catch-at-age data for 1988, 1991, and 1994 using break-and-burn otolith ages, (4) a probability transition matrix for conversion of age types, (5) length composition data from each fishery over the period 1980-1994, (6) an effort index in the recreational fishery, (7) triennial shelf trawl survey CPUE and length-frequency data, (8) a spawning biomass index derived from larval abundance in CalCOFI surveys, and (9) an index of year-class strength from a midwater trawl survey of young-of-the-year pelagic juvenile abundance.

An evaluation of these diverse sets of information indicated that the age composition data were in fundamental disagreement with all other data sources. This discrepancy was apparently due to bias and imprecision in bocaccio ages, which resulted in uninformative age composition data that were incapable of resolving a highly variable pattern of recruitment to the fishery. For this purpose, length composition data were much more useful, especially including those from the trawl and recreational fisheries.

Introduction

The need to incorporate diverse sets of information into the statistical analysis of fish population dynamics has led to the development and evolution of flexible stock-assessment models (e.g., Fournier and Archibald 1982; Deriso et al. 1985; Methot 1989, 1990). Along the West Coast of the United States, Methot's Stock Synthesis model has become the standard analytical tool for estimating the population status of groundfish stocks. Within the framework offered by the synthesis model, the inclusion of catch-at-age data has been the cornerstone of most groundfish assessments. Like Fournier and Archibald's (1982) model, Stock Synthesis plausibly treats age composition data as measured with a multinomial error structure, but it is unique in that errors attributable to reader mis-aging can also be included in the model.

In a broader context, the use of catch-at-age data in fish stock assessments has been reviewed by Megrey (1989). It is widely presumed that estimates of the age composition of the catch are the most informative and useful data one can obtain when modeling the effects of fishing on a stock. This is particularly true of species that show a variable pattern of recruitment, as is typical of the rockfishes (*Sebastes*). Even so, a number of studies have highlighted the benefits of including "auxiliary" data in age-structured stock assessments, especially in terms of constraining the fits of population models to catch-at-age data (e.g., Bence et al. 1993; Hightower 1996).

Here we report on certain findings from a recently completed stock assessment of bocaccio *(Sebastes paucispinis)* which employed the Stock Synthesis model (detailed results available in Ralston et al. 1996). The assessment was notable in that a large number of fishery-dependent and fishery-independent data sources were involved in the analysis. Two of the fishery-independent data sources were new and had not been used previously in groundfish stock assessments conducted on the West Coast. We also included new break-and-burn age-frequency distributions in the assessment, as recommended by Bence and Rogers (1992), and evaluated the relationship between those data and the surface age composition information that had been used in the last stock assessment.

Bocaccio is an important species of rockfish that has a long history of exploitation in California (Fig. 1; Lenarz 1987, Ralston et al. 1996). It is most abundant off southern and central California and is uncommon between Cape Mendocino and Cape Blanco. A second population center exists near the Oregon-Washington border, and extends north to Cape Flattery (Gunderson and Sample 1980, Ralston et al. 1996). Bocaccio frequents an exceptional diversity of habitats, including kelp forests, rocky reefs, midwater, and open, low relief bottoms (Eschmeyer 1983). Even though subadult growth can be very rapid in absolute terms (24 cm at age 1), adults grow slowly ($K = 0.11$-0.13 yr^{-1} [Wilkins 1980]). Moreover, growth is sexually dimorphic, with females reaching much larger sizes than males (i.e., 90 versus 70 cm).

Sources of Data

Fishery-Dependent Data

The assessment was restricted to California because Oregon landings of bocaccio are virtually nil and few biological samples were collected for the small catches taken from the Washington subpopulation. California commercial landings statistics for the period 1980-1995 were summarized using procedures outlined in Erwin et al. (1997) and Pearson and Erwin (1997). Estimates of recreational landings, which have been significant, were extracted from the national Marine Recreational Fishery Survey Statistics (MRFSS) database. The catch time series was lengthened to encompass the 1950-1995 period using information and methods detailed in Ralston et al. (1996). During the last half century, bocaccio have been harvested in four distinct sectors, i.e., the trawl, hook-and-line, setnet, and recreational fisheries (Fig. 1). Each fishery has been characterized by a distinct exploitation pattern, and in the assessment each was modeled independently of the others.

Bocaccio age composition data from the trawl fishery were available for use in the assessment, although all data for the 1980-1985 period were based on surface ages (Table 1), which were thought to be biased low for older fish (Beamish 1979). To estimate the age bias of these fish, 612 bocaccio that had been surface-aged in 1983-1984 were re-aged using the break-and-burn method. That study showed that at a break-and-burn age of 10-yr the mean surface age of bocaccio was 8-yr, while at a break-and-burn age of 20-yr, surface age averaged 14-yr. These results were further analyzed and a probability transition matrix was developed to transform model age composition vectors to predicted surface age composition vectors. The matrix was included in the Stock Synthesis model as a means of generating predicted surface age data from the underlying dynamics of the model. The 1980-1985 surface age data were also supplemented with break-and-burn age data from the trawl fishery for the years 1988, 1991, and 1994 (Table 1). Last, the precision of break-and-burn ages was evaluated by re-examining 25% of all the aged fish. Based on these 275 otoliths, percent agreement between readings declined from ~90% for age-1 fish to ~10% agreement for age-20 fish. The pattern of decline reflected an exponential decay in the precision of age estimates with increasing age (Fig. 2).

Sex-specific length compositions were also available for each year and each of the three commercial fisheries for the period 1980-1994 (Pearson

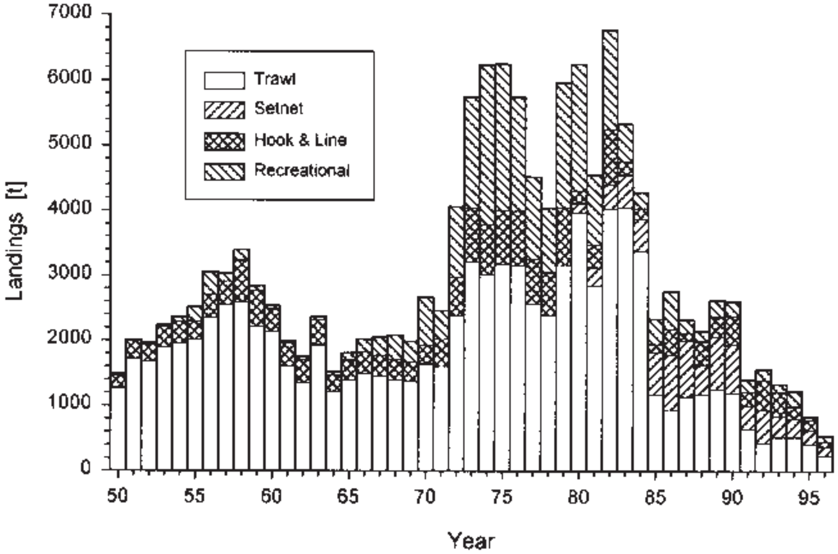


Figure 1. Estimated landings of bocaccio in the Eureka-Monterey-Conception INPFC areas during the last half century.

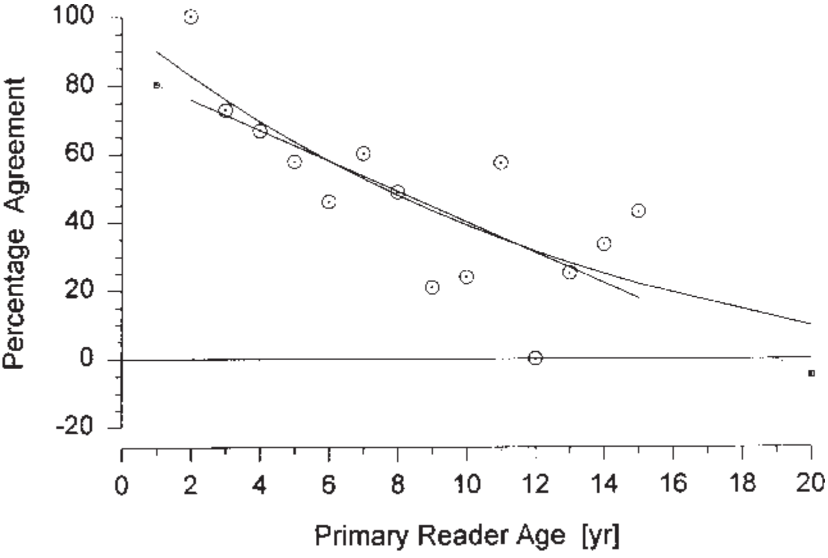


Figure 2. Precision of bocaccio break-and-burn age data, as measured by percent agreement to the year among re-examined otoliths. Lines represent linear and exponential fits.

Table 1. **Sources of data and information used in the Stock Synthesis model of the Eureka-Monterey-Conception bocaccio stock during the 1969-1996 period.**

	Year																											
	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96
Fishery Dependent																												
Trawl landings[a]	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	
Trawl length composition[a]												●	●	●	●	●	●	●	●	●	●	●	●	●	●	●		
Trawl surface age composition[c]												●	●	●	●	●	●											
Trawl break-and-burn age composition[c]																				●			●			●		
Hook-and-line landings[a]	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	
Hook-and-line length composition[b]												●	●	●	●	●	●	●	●	●	●	●	●	●	●	●		
Setnet landings[a]	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	
Setnet length composition[b]												●	●	●	●	●	●	●	●	●	●	●	●	●	●	●		
Recreational landings[a]	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	+	+	+	●	●	●	
Recreational length composition[a]													●	●	●	●	●	●	●	●	●				●	●	●	
Recreational fishing effort[a]													●	●	●	●	●	●	●	●	●				●	●	●	
Fishery Independent																												
Triennial trawl survey CPUE[a]									●			●			●			●			●			●			●	
Triennial trawl survey length composition[b]									●			●			●			●			●			●			●	
Pelagic juvenile rockfish survey[a]																●	●	●	●	●	●	●	●	●	●	●	●	●
CalCOFI larval abundance index[a]	●	●		●			●	●		●			●			●												

[a] Primary data types (see Model Results section).

[b] Secondary data types.

[c] Tertiary data types (i.e., age composition data)

Note that recreational landings from 1990-92 were interpolated (+).

and Erwin 1997). In contrast, length composition data for the combined-sex recreational fishery were available from the MRFSS database for 1981-1989 and 1993-1995, as was a recreational fishing effort series (Table 1).

Fishery-Independent Data

Three sources of auxiliary survey information were used in the bocaccio assessment. These included the Alaska Fisheries Science Center's triennial shelf survey, the Southwest Fisheries Science Center's (SWFSC) pelagic juvenile rockfish midwater trawl survey, and the California Cooperative Oceanic Fisheries Investigation's (CalCOFI) ichthyoplankton surveys. A brief description of each follows.

The triennial bottom trawl survey has been completed once every three years since 1977 (Table 1). The survey, which has found very widespread use in Pacific coast groundfish stock assessments, samples continental shelf habitats in the 55-366 m depth range using a high-opening Nor'eastern bottom trawl equipped with bobbin roller gear (Wilkins 1996). In this study, only standard trawls conducted in the southern area (Eureka, Monterey, and Conception International North Pacific Fisheries Commission areas) were used. Although the survey is often used to provide swept-area estimates of absolute biomass, we treated the survey as a relative index of bocaccio abundance (Fig. 3). The triennial survey indicates that a substantial reduction in bocaccio biomass has occurred over the last two decades. Catch-weighted estimates of year-specific and sex-specific length compositions from the survey were also used as input data to the model.

The pelagic juvenile rockfish midwater trawl survey is designed to estimate the relative year-class strength of a group of 10 rockfish species, including bocaccio. The survey has been conducted every year since 1983 and uses a modified Cobb midwater trawl. A series of 36 standard stations are sampled during three repetitive occupations of a 110-mile study area along the central California coast. Stratified means are calculated that represent the average number of 100-day-old fish taken during a standard trawl, with the maximum value among the three occupation means providing an estimate of year-class strength (see Ralston and Howard 1995). In this instance the time series was shifted forward by one year, representing the relative abundance of age-1 bocaccio recruits (Table 1, Fig. 4). Note that the relatively low abundance of age-1 fish in 1984 and 1993 was due to the adverse effects of the 1983 and 1992 El Niños on rockfish reproductive success.

Within the California Current ecosystem, CalCOFI data have been collected over a grid of north-south lines and onshore-offshore stations since 1951. Within that time period cruises are typically identified by the year and the principal month of sampling. At occupied stations, plankton samples have been collected using both bongo and ring nets; samples are later sorted in the laboratory. The ichthyoplankton are identified and

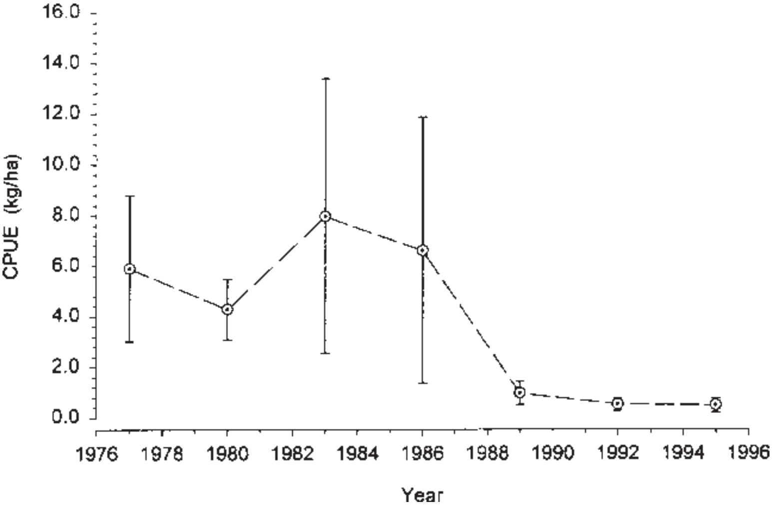


Figure 3. Alaska Fisheries Science Center's triennial shelf trawl survey catch-per-unit-effort (CPUE) of bocaccio in the Eureka-Monterey-Conception INPFC areas. Error bars represent ± 1.0 standard error.

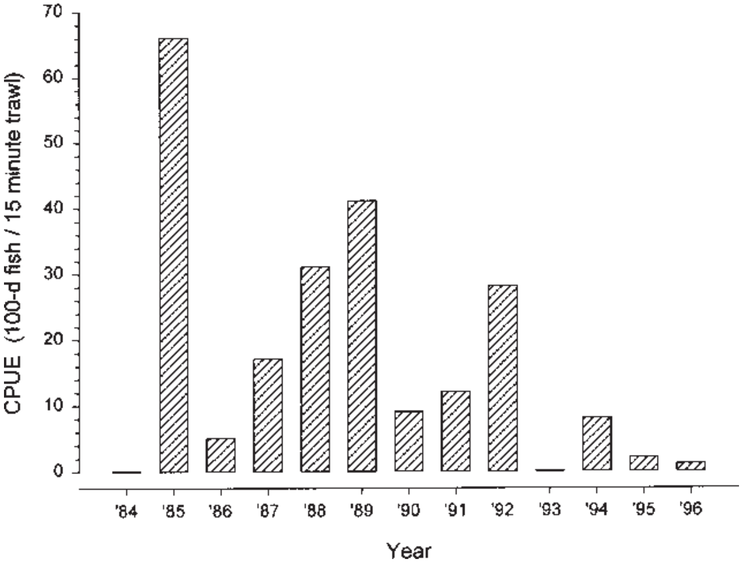


Figure 4. Relative year-class strength of age-1 bocaccio in year t based on the abundance in year t-1 (data from the pelagic juvenile rockfish midwater trawl survey). Note that low survey catches in 1983 and 1992 were associated with El Niño events.

enumerated when possible and the information entered into the CalCOFI database (Moser et al. 1993). Bocaccio larvae are relatively easy to identify, but have not been sorted from the entire time series of CalCOFI collections (see Table 1 for available years). Jacobson et al. (1996) describe the use of the log-transformed data to index the abundance of bocaccio larvae using a General Linear Model (GLM), which included terms for year, month, line, station, and all non-year interaction terms. Because the survey primarily samples very young larvae, year effects from the GLM can be used to provide an index of spawner abundance (Fig. 5).

Model Structure

The Stock Synthesis model is a forward-projecting, separable, age-structured population model. The separability assumption requires that the fishing mortality rate experienced by fish of age a in year t ($F_{a,t}$) is defined by the product of a year-specific full-selection instantaneous fishing mortality rate (F_t) and an age-specific value of selectivity (s_a), i.e., $F_{a,t} = F_t s_a$. Key features of the model are that it incorporates a multinomial error structure for both age and length composition data, it explicitly models aging errors when constructing predicted age composition data, and it conveniently allows a variety of data elements to be combined and evaluated under one umbrella formulation. In particular, all data types are combined in a total $\log_e$-likelihood equation of the form:

$$\ell_{Total} = \sum_{i=1}^{m} \ell_i \lambda_i$$

where ℓ_{Total} is the total $\log_e$-likelihood of the model and the ℓ_i are the individual $\log_e$-likelihoods for each of the m data components used by the model. These are weighted by "emphasis" factors (λ_i), such that in combination the various data sources used by the model can be controlled. To reduce the influence of one data type the particular λ_i can be reduced to a nil emphasis (e.g., 0.001).

The model is typically configured to treat observations of age composition data to be measured with a multinomial sampling error structure. In particular, a $\log_e$-likelihood component for the ith type of age data takes the form:

$$\ell_i(p \mid \hat{p}) = \sum_t n_{i,t} \sum_a p_{i,a,t} \log_e(\hat{p}_{i,a,t})$$

where $p_{i,a,t}$ is the observed proportion of fish that are age a in samples collected in year t, $\hat{p}_{i,a,t}$ is the model's prediction of that proportion, and $n_{i,t}$ is the year-specific sample size upon which the observed proportions are based. The model then performs an iterative search for values of $\hat{p}_{i,a,t}$

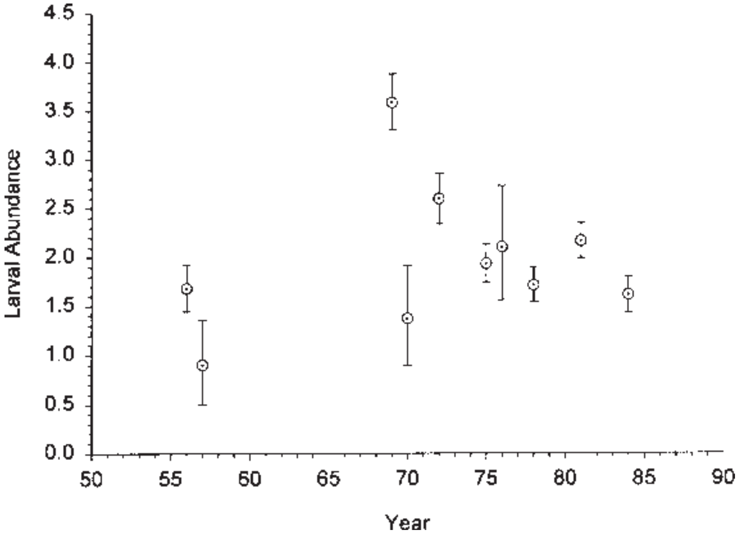


Figure 5. Long-term patterns in the abundance of larval bocaccio, as estimated by a General Linear Model (GLM) applied to the CalCOFI data base. Error bars indicate 95% confidence intervals.

that will maximize ℓ_i. Length composition data are fitted in a similar manner. Survey data, however, are usually modeled with a lognormal error term, i.e.,

$$\ell_i = -\sum_t \left[\log_e(\sigma_{i,t}) + \frac{\log_e(I_{i,t} / \hat{I}_{i,t})^2}{2\sigma_{i,t}^2} \right]$$

where ℓ_i is the $\log_e$-likelihood component for the ith survey, $I_{i,t}$ is the observed value of the survey index in year t, $\hat{I}_{i,t}$ is the model's prediction of the index value, and $\sigma_{i,t}$ is the standard error of the statistic on $\log_e$-scale.

During the development of a baseline model for bocaccio, we explored a number of different configurations. Since the level of data and model complexity was very high (four fisheries, split sexes, unbiased and biased ages, three surveys, etc.), we decided some simplification was needed. First, because there were only estimates of total catch for the period 1950-1968, we elected to exclude that period and we modeled the population from 1969 to the present. Even so, data from the omitted period were used to establish historic catch levels. Although some CalCOFI data were

Table 2. Summary of parameters estimated in the baseline Stock Synthesis model of bocaccio (M was fixed at 0.15 yr^{-1}).

Model element		Number of parameters
Trawl selectivity	Stationary	8
	Time-varying	8
Setnet selectivity		9
Hook-and-line selectivity		9
Recreational selectivity	Stationary	5
	Time-varying	21
Triennial survey selectivity		3
Recruit survey selectivity		1
CalCOFI survey selectivity		0
Recreational fishery effort		1
Growth		5
Recruitments		28
Total		98

available from the mid-1950s (Fig. 5), preliminary analyses indicated they had little or no effect on the model's final estimate of current stock size.

In the prior assessment, Bence and Rogers (1992) showed that bocaccio selectivity patterns were inadequately described by the age-based Stock Synthesis model (see Methot 1990). Like them, we used the length-based implementation of the model, although the selectivity curve for the pelagic juvenile rockfish survey was modeled as full vulnerability at age-1 and no vulnerability at any other age. Similarly, in the previous assessment, a component of the trawl fishery selectivity function was allowed to vary with time. We initially explored a model with constant selectivity and compared this to models where the ascending inflection point of the selectivity curve was allowed to vary. When interannual variation in that parameter was fully expressed (i.e., a value was estimated for each year), a gain of 49 log-likelihood units was realized at the cost of eleven new parameters, representing a significant improvement in fit. We were able to further simplify the model and reduce the total number of parameters by pooling five of these, without substantially affecting the total log-likelihood of the model. A similar procedure was used for the selectivity curve in the recreational fishery, except that two parameters were allowed to vary with time (i.e., the ascending inflection point and the initial selectivity).

Thus, the final "baseline" version of the bocaccio model included eleven $\log_e$-likelihood components (i.e., $m = 11$) that together controlled the fit of the model to the data. These were components for: (1) the trawl fishery length composition data, (2) trawl fishery surface age composition

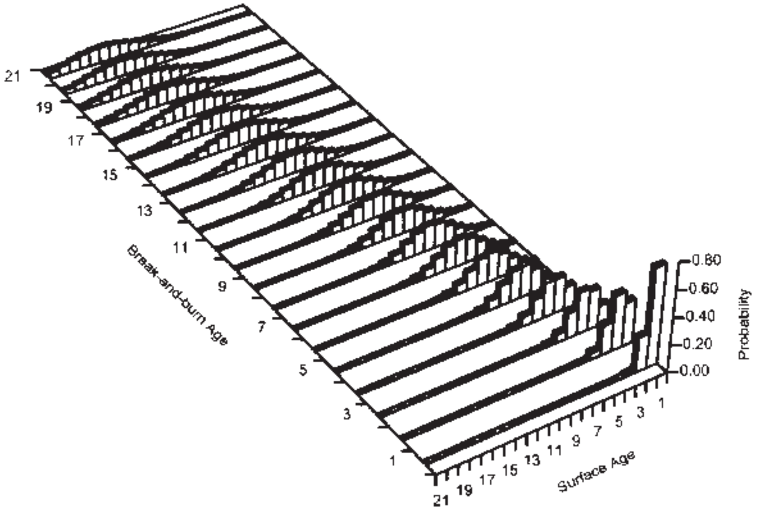


Figure 6. Probability transition matrix used to generate predicted surface age distributions.

data, (3) trawl fishery break-and-burn age composition data, (4) hook-and-line fishery length-frequency data, (5) setnet fishery length-frequency data, (6) recreational fishery length-frequency data, (7) a recreational fishing effort series, (8) the triennial trawl survey time series of catch rate, (9) triennial survey length compositions, (10) the pelagic juvenile rockfish survey, and (11) the CalCOFI larval abundance survey (assumed $\propto$ spawning biomass). To fit the baseline bocaccio model a total of 98 parameters were estimated (Table 2). Also note that, once converged, the model solves for the fishing mortality rate that is required to produce an exact match to the observed landings, subject to the specific constraints imposed by the estimated parameter set.

Model Results

All of the age data used in the previous assessment (Bence and Rogers 1992) were based on surface ages. In our assessment we attempted to incorporate new age data that were derived from the break-and-burn method. This had the unfortunate effect of adding a new level of complexity in interpreting the age data. First, it became evident that, given the spread at older ages in the transition matrix, very little information could be extracted from surface ages (Fig. 6). Second, the percent agreement between reexamined break-and-burn samples was rather poor (Fig. 2). Finally, specific age determination criteria for bocaccio had not been rigorously validated and the two age readers indicated that, at least relative to other species of rockfish, bocaccio otoliths were difficult to decipher. The lack of validation implied that the break-and-burn ages could be biased. These issues led us to question the fundamental reliability of the different data sources, particularly the age composition data.

Concerns over the age composition data component were verified when the model was fit with (1) all the age data fully emphasized, (2) reliance on break-and-burn ages only, and (3) nil emphasis on any of the age composition data (Fig. 7). In the last case the model indicated a strong recruitment event of age-1 fish occurred in 1978 and that minor but above average events occurred in 1985 and 1989. These patterns were also evident in the trawl length composition data (Fig. 8), the recreational length composition data, and in the recruit survey (Fig. 4). The strong 1977 year class was, moreover, well known to California Department of Fish and Game biologists based on a tremendous influx of small fish in the 1977 and 1978 nearshore recreational fishery. In contrast, in the first two cases, wherein the age data were allowed to influence the fit of the model, the estimated time series of recruitments was blended in the 1980s and no dominant 1977 year class was evident.

Based on these findings, we classified the data into three general groups: primary, secondary, and tertiary (Table 1). This was done to improve the robustness of model outputs and to avoid model sensitivity to data that were noisy or were otherwise questionable. The primary data types were used in the estimation of growth, year-class strength, and population trend. The secondary data were used only to estimate selectivity patterns for the different gear types and the tertiary data were effectively omitted from the analysis except to highlight their deficiencies. To fit the model to the different data classes, we followed a simple, iterative scheme, i.e., (1) fitting the model with only the primary data types emphasized, (2) fixing parameters estimated in the first step and estimating selectivity parameters for secondary data types, and (3) repeating steps one and two until the model showed no further tendency to change. The resulting baseline model was characterized by trajectories of bocaccio summary biomass and spawning output that showed severe declines over the course of the modeled period, with terminal year values in the range of 5-10% of their maxima, which occurred in 1969 (Fig. 9). The obvious “bump” in these downward trends represents the strong 1977 year-class passing through the population.

Discussion

The assessment of bocaccio was complicated by a number of factors. In particular, we modeled the effects of four distinct fisheries on the abundance of male and female bocaccio. In addition, the trawl and recreational fisheries were marked by significant time-varying effects on selectivity.

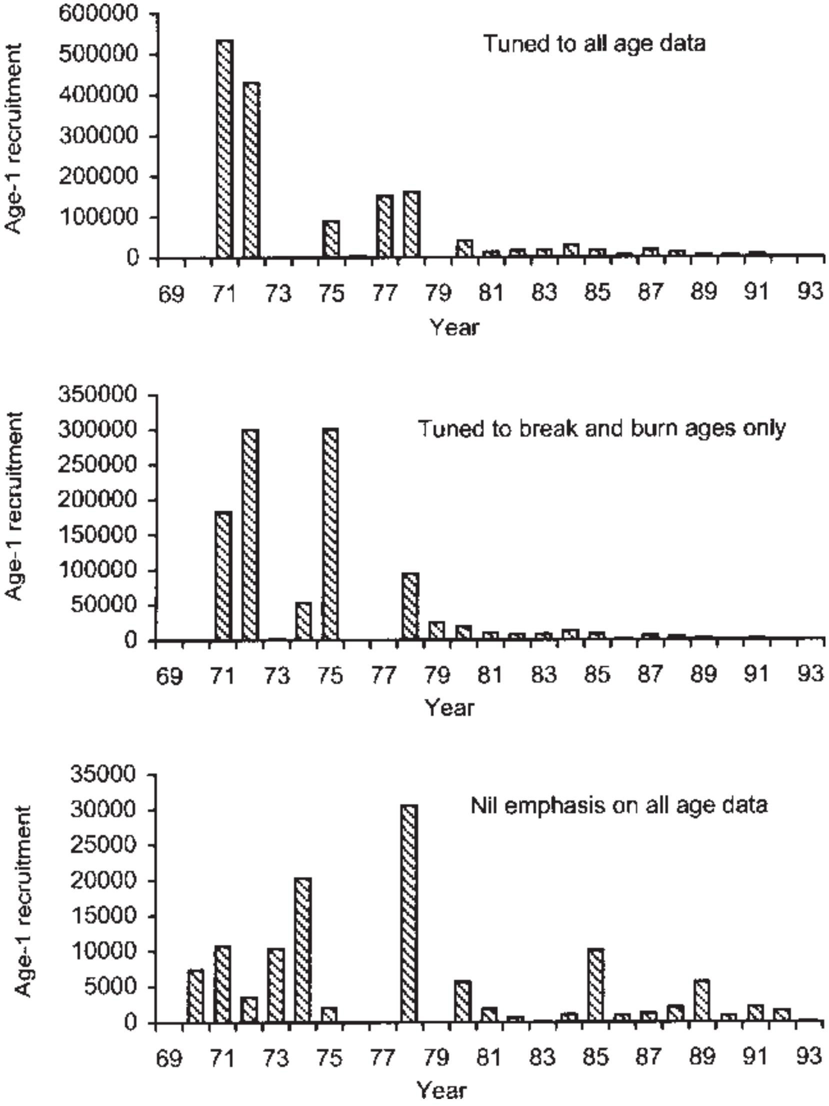


Figure 7. Time series of age-1 recruitments estimated with different emphasis levels on the age composition data.

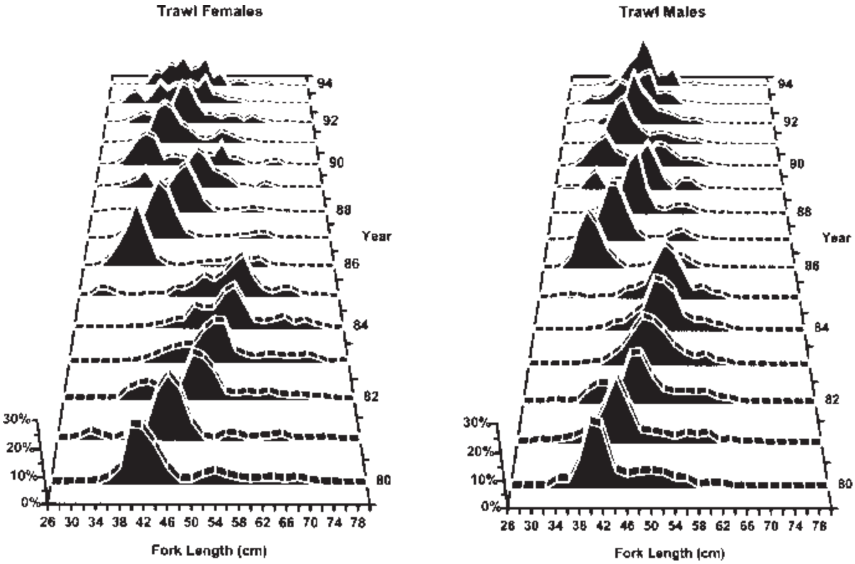


Figure 8. Relative length composition distributions of bocaccio caught in the trawl fishery (1980-1994).

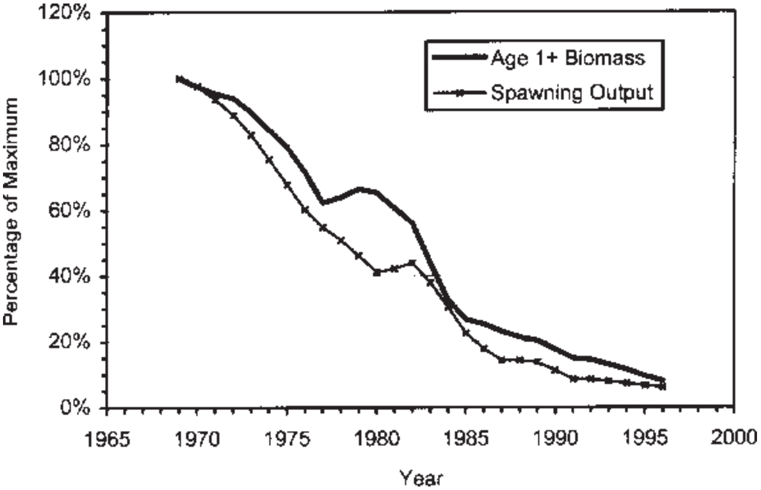


Figure 9. Estimated trends in summary biomass and spawning potential from the baseline bocaccio Stock Synthesis model.

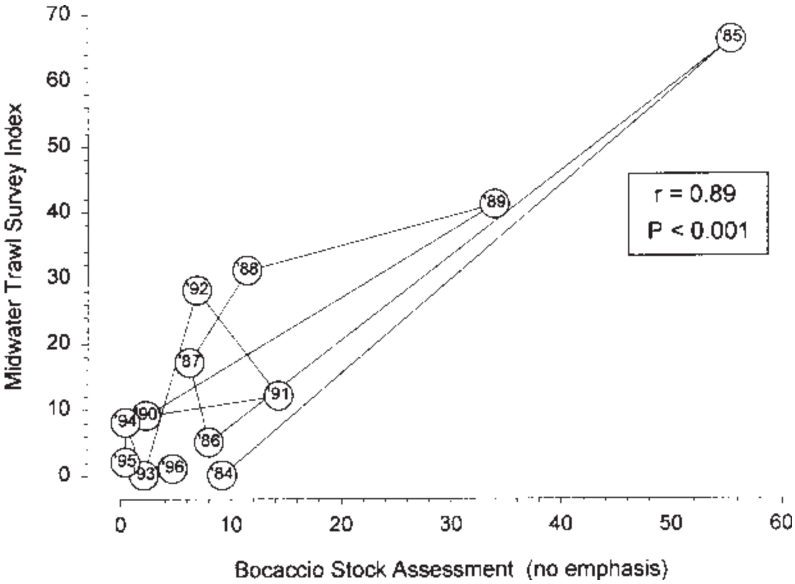


Figure 10. Relationship between the Southwest Fisheries Science Center's pelagic juvenile rockfish midwater trawl survey index of bocaccio abundance and year-class strength estimated from the Stock Synthesis model with no emphasis on those survey data.

We also included three separate sources of fishery-independent information in the model. Perhaps most complex, however, was the inclusion of surface and break-and-burn age information in the assessment and unraveling the conflicting signals of those data with other information in the model.

In the case of bocaccio, the age data were of questionable utility, apparently due to problems with bias and imprecision (Figs. 2 and 6). In contrast, the length composition data were very informative and carried clear, unambiguous signals of strong year-classes passing through the fisheries (e.g., Fig. 8). Biologically, we believe that these two observations are related. The clear progression of modes in the length data was due to the rapid absolute growth of young bocaccio and the relatively brief seasonal expression of spawning. However, these two features exacerbated the interpretation of bocaccio otoliths. Rapid growth of subadult fish resulted in the proliferation of false annuli and accessory check marks in the otoliths, which were difficult to interpret, resolve, and validate through the application of marginal increment analysis.

We also utilized two new sources of information in a groundfish stock assessment, i.e., the SWFSC midwater trawl survey of pelagic juvenile rockfish abundance and the CalCOFI database of larval bocaccio abundance.

These complemented the traditional triennial groundfish trawl survey, which has been used extensively in previous stock assessments. To substantiate and validate the midwater trawl survey index as a relative index of recruitment, we also fitted our final model with zero weight on the recruit survey (Fig. 10). In this independent comparison the survey index was in full agreement with the model's interpretation of year-class strength, which was largely based on the available length-frequency data, particularly those from the trawl and recreational fisheries. It is our belief that in future applications this index will provide a reliable basis for projecting trends in biomass into the near future. Likewise, the CalCOFI database was used for the first time to index the relative spawning biomass of a groundfish. Except for the obvious outlier in 1970, it tended to closely follow the overall pattern of decline in spawning output observed in the base run model (Figs. 5 and 9). Based on our experience with these information sources, we encourage and advocate the use of both these fishery-independent auxiliary data sets in future groundfish stock assessments. However, one should always closely examine models for assumptions made about data quality.

Acknowledgments

We would like to acknowledge the collaboration and assistance of a number of people who contributed significantly to this work, including Brian Culver, Jerry Gray, Kevin Hill, Larry Jacobson, Jerry Kobylinski, Bill Lenarz, Alec MacCall, Rick Methot, Rachael Miller, Don Pearson, Jean Rogers, Dave Thomas, Wade Van Buskirk, and Mark Wilkins.

References

Beamish, R.J. 1979. New information on the longevity of Pacific ocean perch (*Sebastes alutus*). J. Fish. Res. Board Canada 36:1395-1400.

Bence, J.R., and J.B. Rogers. 1992. Status of bocaccio in the Conception/Monterey/Eureka INPFC areas in 1992 and recommendations for management in 1993. In: Appendices to the Status of the Pacific Coast groundfish fishery through 1992 and recommended acceptable biological catches for 1993. Pacific Fishery Management Council, 2000 SW First Ave., Portland, OR 97201.

Bence, J.R., A. Gordoa, and J.E. Hightower. 1993. Influence of age-selective surveys on the reliability of stock synthesis assessments. Can. J. Fish. Aquat. Sci. 50:827-840.

Deriso, R.B., T.J. Quinn II, and P.R. Neal. 1985. Catch-age analysis with auxiliary information. Can. J. Fish. Aquat. Sci. 42:815-824.

Erwin, B.A., D.H. Thomas, G.J. Kobylinski, and J.R. Bence. 1997. Groundfish data collection in California, Chapter 4. In: D.B. Sampson and P.R. Crone (eds.), Commercial fisheries data collection procedures for U.S. Pacific Coast groundfish. NOAA Tech. Memo. NMFS, NOAA-TM-NMFS-NWFSC-31, pp. 105-140.

Eschmeyer, W.N. 1983. A field guide to Pacific coast fishes of North America. Houghton Mifflin Company, Boston. 336 pp.

Fournier, D., and C.P. Archibald. 1982. A general theory for analyzing catch at age data. Can J. Fish. Aquat. Sci. 39:1195-1207.

Gunderson, D.R., and T.M. Sample. 1980. Distribution and abundance of rockfish off Washington, Oregon, and California during 1977. Mar. Fish. Rev. 42(3-4):2-16.

Hightower, J.E. 1996. Comparison of reliability of catch-at-age analyses using auxiliary survey estimates of relative or absolute recruitment. Can. J. Fish. Aquat. Sci. 53:70-79.

Jacobson, L.D., S. Ralston, and A.D. MacCall. 1996. Historical larval abundance indices for bocaccio (*Sebastes paucispinis*) from CalCOFI data. NOAA, NMFS, SWFSC Admin. Rep., La Jolla, LJ-96-06. 30 pp.

Lenarz, W.H. 1987. A history of California rockfish fisheries. In: Proceedings of the International Rockfish Symposium. University of Alaska Sea Grant, AK-SG-87-02, Fairbanks, pp. 35-41.

Megrey, B.A. 1989. Review and comparison of age-structured stock assessment models from theoretical and applied points of view. Amer. Fish. Soc. Symp. 6:8-48.

Methot, R.D. 1989. Synthetic estimates of historical abundance and mortality for northern anchovy. Amer. Fish. Soc. Symp. 6:66-82.

Methot, R.D. 1990. Synthesis model: An adaptable framework for analysis of diverse stock assessment data. Int. North Pac. Fish. Comm. Bull. 50:259-277.

Moser, H.G., R.L. Charter, P.E. Smith, D.A. Ambrose, S.R. Charter, C.A. Meyer, E.M. Sandknop, and W. Watson. 1993. Distributional atlas of fish larvae and eggs in the California Current region: Taxa with 1000 or more total larvae, 1951 through 1984. Calif. Coop. Ocean. Fish. Investig. Atlas 31. 233 pp.

Pearson, D.E., and B. Erwin. 1997. Documentation of California's commercial market sampling data entry and expansion programs. NOAA Tech. Memo. NMFS, NOAA-TM-NMFS-SWFSC-240. 62 pp.

Ralston, S., and D.F. Howard. 1995. On the development of year-class strength and cohort variability in two northern California rockfishes. Fish. Bull. U.S. 93:710-720.

Ralston, S., J.N. Ianelli, R.A. Miller, D.E. Pearson, D. Thomas, and M.E. Wilkins. 1996. Status of bocaccio in the Conception/Monterey/Eureka INPFC areas in 1996 and recommendations for management in 1997. In: Status of the Pacific Coast groundfish fishery through 1996 and recommended acceptable biological catches for 1997, stock assessment and fishery evaluation, Appendix B. Pacific Fishery Management Council, 2130 SW Fifth Ave., Suite 224, Portland, OR 97201.

Wilkins, M.E. 1980. Size composition, age composition, and growth of chilipepper, *Sebastes goodei*, and bocaccio, *S. paucispinis*, from the 1977 rockfish survey. Mar. Fish. Rev. 42(3-4):48-53.

Wilkins, M.E. 1996. Long term trends in abundance: Results of triennial bottom trawl surveys of west coast groundfish resources between 1977 and 1995. In: Status of the Pacific Coast groundfish fishery through 1996 and recommended acceptable biological catches for 1997, stock assessment and fishery evaluation, Appendix F. Pacific Fishery Management Council, 2130 SW Fifth Ave., Suite 224, Portland, OR 97201.

Integrated Catch-Age Mark-Recapture Model: Application to B.C. Sablefish Stocks

Vivian Haist
Fisheries and Oceans Canada, Pacific Biological Station, Nanaimo, B.C., Canada

Abstract

An integrated catch-age mark-recapture model is described along with its application to two British Columbia sablefish stocks. The model can be implemented with explicit migration or not, and the model structure includes age- and sex-specific depth stratification. Model parameters estimated through the analyses include recruitments, fishery parameters, emigration rates, tag loss rates, and tag reporting rates. The parameters of the catch-age and mark-recapture components of the model are estimated simultaneously. The results indicate a high degree of correlation between tag loss and reporting rate but the estimated values for these parameters are precise and robust. The tag attrition rate is greater than would be expected from natural and fishing mortality rates alone. The model accounts for this with high values for the instantaneous tag loss rate (0.26-0.27). These values are significantly higher than would be expected if tag shedding were the major component of tag loss. Implementation of the migration component of the model does not provide a satisfactory fit to the observed tag recovery data, again because the tag attrition rate for migrant fish is greater than model predictions. A multi-stock model that explicitly accounts for immigration as well as emigration may resolve some of the remaining questions.

Introduction

The use of catch-age models in fisheries stock assessments has become common practice for stocks where a reasonable time-series of catch and age composition data is available. However, catch and age composition data alone are generally inadequate to reconstruct the history of a fish population and catch-age analyses require auxiliary information to accurately estimate absolute abundance (Doubleday 1976, Pope and Shepherd 1982, Deriso et al. 1985). Survey-based relative or absolute abundance indices are commonly used to tune catch-age analyses. In the absence of fishery-independent auxiliary data, catch-effort statistics from fisheries are often used. This requires the assumption that the catch per unit effort is proportional to stock abundance, an assumption which is often violated (Hilborn and Walters 1992, Chapter 5).

Alternate methods to catch-age analysis for estimating fishing mortality rates and stock abundance include various forms of mark-recapture analysis (Ricker 1975, Seber 1982, Pollock 1991). Mark-recapture analyses generally require independent estimates of model parameters such as tag loss rates and tag reporting rates because these parameters are confounded with fishing and total mortality estimates (e.g., Kleiber et al. 1983, Myers et al. 1997). Ricker (1975, pp. 113-120) suggests that tag loss can be calculated as the difference between the total mortality rate for the population as estimated from catch curve analysis and the total mortality estimates for the marked fish. However, catch curve analysis requires assumptions of regularity in recruitment and constant fishing mortality rates, assumptions that are rarely met in commercial fishery situations. These assumptions can be eliminated by integrating mark-recapture analysis in a catch-age model.

For British Columbia sablefish stocks, time-series of catch and age composition data are available; however, commonly used auxiliary data are either unavailable or are considered unreliable indices of stock abundance. Tagging studies were conducted on these stocks in the early 1980s and again during the 1990s, so a reasonable time-series of mark and recapture data is available to provide auxiliary information for catch-age analysis. In this paper I describe an integrated catch-age mark-recapture model that was developed to assess the abundance of B.C. sablefish stocks. Parameters of the catch equations and the mark-recapture equations are estimated simultaneously. The model can be implemented as a migration model, with explicit immigration and emigration parameters, or as a nonmigration model. Analyses of the B.C. sablefish data are presented for both implementations of the model.

Model Description

Population Dynamics and Catch Equations

Biological data indicate that sablefish are both age- and sex-stratified. Younger fish are relatively more abundant in shallow waters and the proportion of females in samples collected in shallow waters is higher than in samples from deeper waters (Saunders et al. 1997). The model developed for the sablefish stock reconstructions includes age- and sex-specific segregation of the population by depth strata. The model structure assumes partial recruitment of younger age classes and that all recruited fish are vulnerable to the fisheries. The catch equations, which relate the numbers of fish in the population to the numbers of fish in the catch, are described by the following relationships:

$$C_{ijks} = \frac{F_{ik}}{Z_{iks}}\left[1-\exp\left(-Z_{iks}\right)\right] p^{d}_{jks} p^{r}_{js} N_{ijs}$$

$$N_{i+1,j+1,s} = \sum_{k} \exp\left(-Z_{iks}\right) p^{d}_{jks} p^{r}_{js} N_{ijs} + \exp\left(-M_{s}\right)\left(1-p^{r}_{js}\right) N_{ijs} \qquad \text{for } j<a$$

$$\begin{aligned} N_{i+1,a,s} = {} & \sum_{k} \exp\left(-Z_{iks}\right) p^{d}_{a-1,ks} p^{r}_{a-1,s} N_{i,a-1,s} + \exp\left(-M_{s}\right)\left(1-p^{r}_{a-1,s}\right) N_{i,a-1,s} + \\ & \sum_{k} \exp\left(-Z_{iks}\right) p^{d}_{aks} p^{r}_{as} N_{ias} + \exp\left(-M_{s}\right)\left(1-p^{r}_{as}\right) N_{ias} \qquad \text{for } j=a \end{aligned}$$

$$Z_{iks} = F_{ik} + M_{s} + G$$

$$C_{i\cdot k\cdot} = \sum_{j}\sum_{s} C_{ijks} w_{js}$$

$$E_{ik} = \sum_{j}\sum_{s} p^{d}_{jks}\, p^{r}_{js}\, N_{ijs}$$

where,

i	indexes year,
j	indexes age,
a	is the number of age classes,
k	indexes depth stratum,
s	indexes sex,
p^{d}_{jks}	is the proportion age class j and sex class s fish in stratum k, where $\sum_{k} p^{d}_{jks} = 1$ and $p^{d}_{jks} \geq 0$,
p^{r}_{js}	is the proportion age class j and sex class s fish that are recruited to the exploitable population, where $0 \leq p^{r}_{js} \leq 1$
C_{ijks}	is the catch (in numbers) of age class j and sex s fish in stratum k in year i
$C_{i\cdot k\cdot}$	is the total catch (in biomass) in stratum k in year i,
F_{ik}	is the instantaneous fishing mortality rate in stratum k in year i,
M_{s}	is the instantaneous natural mortality rate for sex s fish,
G	is the instantaneous net migration rate,
Z_{iks}	is the instantaneous total mortality rate for sex s fish in stratum k in year i,
N_{ijs}	is the number of fish in age class j and sex class s in year i,
E_{ik}	is the number of exploitable fish in stratum k in year i,
w_{js}	is the average weight of fish of sex s and age j.

Parameterizing Partial Recruitment, Depth Stratification, and Mortalities

The partial recruitments are defined in terms of parameters c_{ms} such that they are constrained between 0 and 1, are non-decreasing with age, and are fixed at 1.0 for age classes a' and older. That is,

$$p^r_{js} = \sum_{m=1}^{j} c_{ms} \Big/ \sum_{m=1}^{a'} c_{ms} \qquad \text{for } j \le a'$$

$$p^r_{js} = 1.0 \qquad \text{for } j > a'$$

The parameter a' is fixed at 9 for all analyses presented here. For each sex, the proportion of an age class occurring in a depth stratum is parameterized as a linear trend from the first age class through the penultimate age class. The trend parameters, t_{jks},

$$t_{jks} = b^1_{ks} + \frac{j}{a-1} b^2_{ks}$$

are re-scaled so that for each sex the sum of the proportions for each age class are 1.0,

$$p^d_{jks} = t_{jks} \Big/ \sum_k t_{jks}$$

The proportions by depth strata for the final age class are independent of those for younger age classes.

$$p^d_{aks} = b^3_{ks} \Big/ \sum_k b^3_{ks}$$

The estimated instantaneous natural mortality rate for B.C. sablefish stocks has been estimated as 0.08 (Sunders et al. 1994). For this analysis, the average instantaneous fishing mortality rate (M) is fixed at 0.08, but a sex-specific deviation (m) from the average rate is estimated. That is,

$$M_1 = M + m \quad \text{and} \quad M_2 = M - m,$$

where M_1 and M_2 are the instantaneous natural mortality rates for males and females respectively.

The fundamental parameters, i.e., those which are estimated when minimizing the log-likelihood function, are the $\ln(N_{i11})$, $\ln(N_{1j1})$, $\ln(F_{ik})$, c_{ms}, b^1_{ks}, b^2_{ks}, b^3_{ks}, G, and m. The number of females (sex 2) are assumed to be equal to the number of males in the population for the first age class

($N_{i,11} = N_{i,1,2}$). For the first year, the numbers of females at each age are scaled relative to the number of males based on an estimate of their cumulative total mortality. The cumulative total mortality for an age and sex class is calculated under the assumption that fishing mortality rates for years prior to the first year of the analysis are equal to the fishing mortality estimates for the first year and that the estimated partial recruitment and age-sex specific depth stratification hold for earlier years.

Likelihood of Observed Catch-at-Age

For fitting the model to the observed catch and age composition data I assume the error structure assumptions proposed by Fournier and Archibald (1982). That is, the proportion-at-age result from taking a sample of size S_{ik} from a population that has a multinomial distribution and the estimates of total catch result from sampling a lognormal distribution. The negative log-likelihood function for the data observations is then

$$-\sum_i\sum_j\sum_k\sum_s S_{ijks}\ln(p_{ijks}) + w_C\left\{\sum_i\sum_k\left[\ln\left(C_{i\cdot k\cdot}\right)-\ln(\tilde{C}_{ik})\right]^2\right\} \qquad (1)$$

where S_{ijks} is the number of fish taken from samples of the catch in depth stratum k in year i that were sex s and age j, $\tilde{C}_{ik}$ is the observed catch in depth stratum k in year i, and the p_{ijks} are the predicted proportions at age $\left(p_{ijks} = C_{ijks}\Big/\sum_j\sum_s C_{ijks}\right)$. The term w_C is determined by the prior assumption made about the accuracy of the observed catch data. For the sablefish analysis, I assumed $w_C = 5000$, which is consistent with a coefficient of variation of about 0.01. The sample sizes, S_{ik}, were scaled relative to the actual number of samples collected with a maximum value of 200.

Relative Abundance Data

For the model as described there is little information in the catch and age composition data to determine the proportions of the population in each depth stratum. To provide some stability to these parameters, the model has a component for the fit to survey estimates of the relative abundance of fish in each depth stratum ($\tilde{E}_{ik}$). Because there appear to be environmentally induced anomalies in the survey data for some years, the model is fit to the average abundance in each depth stratum rather than the time-series of observations. The average observed relative abundance ($\tilde{x}_k$) and predicted absolute abundance (x_k) are

$$\tilde{x}_k = \sum_i \tilde{E}_{ik}\Big/n_k$$

$$x_k = \sum_i E_{ik}\Big/n_k$$

where n_k is the number of years with survey observations for depth stratum k. Assuming a lognormal error distribution for the survey abundance data, the contribution to the negative log-likelihood equation is

$$w_E \left\{ \sum_k \left[\ln\left(q\,x_k\right) - \ln\left(\tilde{x}_k\right) \right]^2 \right\} \tag{2}$$

where q is an abundance scalar and the constant w_E reflects the prior assumptions about the accuracy of the survey relative abundance data. For the sablefish analysis, values of w_E ranging from 5 to 500 were investigated. This range is consistent with coefficients of variation from 0.03 to 0.32.

Mark-Recapture Equations

The mark-recapture component of the integrated model tracks tag cohorts, i.e., all the fish tagged in a year. The age and sex of tagged fish is estimated based on the number of tags applied in each depth stratum and the age and sex composition estimated for the stratum. The model explicitly accounts for emigration from the tagged population but not for movement between depth strata. Movement between depth strata is implicit in the model in that the probability of a fish being in a depth stratum is related to its age. Note that the population migration rate is different from the tag cohort migration rate because the population parameter accounts for net movement (i.e., immigration plus emigration) whereas for the tag cohorts only emigration can be estimated. The following equations describe the relationships for the non-migrant component of the tag cohorts, where tag cohorts are referenced by the year in which they were tagged.

$$T_{lijs} = \sum_k \frac{p^d_{jks}\,p^r_{js}\,N_{lks}}{E_{lk}} (1 - s\ )\,\tilde{U}_{lk} \qquad \text{for } i = l$$

$$C^T_{lik} = \sum_j \sum_s \frac{F_{ik}}{Z^T_{iks}} \left[1 - \exp\left(-Z^T_{iks}\right)\right] p^d_{jks} T_{lijs} \qquad \text{for } i > l$$

$$Z^T_{iks} = F_{ik} + M_s + G^T + L \qquad \text{for } i > l$$

$$T_{l,i+1,j+1,s} = \sum_k \exp(-Z^T_{iks}) p^d_{jks} T_{lijs} \qquad \text{for } i > l,\ j < a$$

$$\begin{aligned} T_{l,i+1,a,s} &= \sum_k \exp(-Z^T_{iks}) p^d_{a-1,ks} T_{li,a-1,s} \\ &\quad + \sum_k \exp(-Z^T_{iks}) p^d_{aks} T_{lias} \end{aligned} \qquad \text{for } i > l$$

$$I_{li} = \sum_j \sum_k \sum_s \frac{G^T}{Z^T_{iks}} \left[1 - \exp\left(-Z^T_{iks}\right)\right] p^d_{jks} T_{lijs} \qquad \text{for } i > l$$

where,

T_{lijs}	is the number of tag cohort *l* fish of age *j* and sex *s* remaining in year *i*
s	is the rate of tag loss that occurs immediately after tagging
$\tilde{U}_{lk}$	is the number of fish of tag cohort *l* (i.e., tagged in year *l*) tagged in depth stratum *k*
G^T	is the instantaneous rate of emigration for the tagged fish
L	is the instantaneous rate of tag loss
Z^T_{iks}	is the total instantaneous mortality for tagged fish of sex *s* in year *i*
C^T_{lik}	is the predicted catch of tagged fish from tag cohort *l* in year *i* and depth stratum *k*
I_{li}	is the number of number of fish from tag cohort *l* migrating out of the population in year *i*

A number of the relationships described above are defined for a tag cohort only for the years following the tagging year ($i > l$). During the same year that tags are applied, the tagged fish are susceptible to only a fraction of the annual mortalities, dependent on the time of year tagging takes place. For the situation, $i = l$, the quantities, F_{ik}, Z^T_{iks}, and G^T in the preceding mark-recapture equations are replaced with the following quantities; $F'_{ik} = t_{ik}F_{ik}$, $Z^{T'}_{iks} = h_i Z_{ik}$, and $G^{T'} = h_i G^T$. The t_{ik} are the proportions of the annual catch in year *i* and depth stratum *k* which are taken after tagging occurs and h_i is the proportion of year *i* remaining after tagging occurs. The values of the t_{ik} and h_i are calculated from the data and are fixed model inputs.

Because migration, as defined, is continuous throughout the year, not all fish that migrate during a year will be vulnerable to the fisheries in the area they migrate to. The model structure assumes that the fish that migrate during a year will incur only half of the fishing mortality for that year. The relationships that describe the dynamics for the migrant fish are

$$V_{l,i+1} = \exp(-0.5\, h_i\, F_i^{TM})\, I_{li} \qquad \text{for } i = l$$

$$C_{li}^{TM} = \left[1 - \exp(-0.5 h_i F_i^{TM})\right] I_{li} \qquad \text{for } i = l$$

$$V_{l,i+1} = \exp(-Z_i^{TM})\, V_{li} + \exp(-0.5 F_i^{TM}) I_{li} \qquad \text{for } i > l$$

$$Z_i^{TM} = F_i^{TM} + M + L \qquad \text{for } i > l$$

$$C_{li}^{TM} = \frac{F_i^{TM}}{Z_i^{TM}}\left[1 - \exp(-Z_i^{TM})\right] V_{li} + \left[1 - \exp(-0.5 F_i^{TM})\right] I_{li} \qquad \text{for } i > l$$

where,

V_{li} is the number of tagged fish from tag cohort l in the migrant pool in year i

F_i^{TM} is the instantaneous fishing mortality rate for tagged migrant fish in year i

Z_i^{TM} is the instantaneous total mortality rate for tagged migrant fish in year i

C_{li}^{TM} is the predicted catch of migrant fish from tag cohort l fish in year i

I do not attempt to estimate all the annual fishing mortality rate parameters for the migrant tagged fish, because the information relative to these parameters is limited (i.e., only tag recovery data) and the parameter values will be highly confounded with other model parameters. Rather, two parameters are estimated, one for the first year and one for the last year (F_{80}^{TM} and F_{96}^{TM}), and a linear trend between these two is assumed. More complex formulations than the linear time trend could be developed; however, preliminary analyses suggested that there is not enough information in the data to estimate even an average level of fishing mortality for the migrant tagged fish.

Likelihood of Tag Recoveries

There are two observation error structure models commonly used when fitting tag recovery observations in mark-recapture analysis (Seber 1982, Hilborn 1990). These are the Poisson and the multinomial distribution. I have implemented both in the integrated sablefish model and as reported by Hilborn (1990), found that results are virtually identical for the two. Results presented in this paper are based on analyses conducted with the multinomial implementation. The quantities required to fit the multinomial model are the total tagged for a tag cohort ($\tilde{U}_{l\cdot}$),

$$\tilde{U}_{l\cdot} = \sum_k \tilde{U}_{lk} ,$$

the total number of tags in a tag cohort which are not recovered ($\tilde{W}_l$),

$$\tilde{W}_l = \tilde{U}_{l\cdot} - \left(\sum_i \sum_k \tilde{C}_{lik}^T + \sum_i \tilde{C}_{li}^{TM} \right),$$

($\tilde{C}_{lik}^T$ and $\tilde{C}_{li}^T$ are the number of tags from tag cohort l returned in year i from depth stratum k and from outside the assessment region, respectively), and the predicted proportion of tag releases recovered in each recovery stratum (g_{lik} and g_{li}^M) and the predicted proportion not recovered (g_l^W),

$$g_{lik} = r_i \; C_{lik}^T \big/ \tilde{U}_{l\cdot} ,$$

$$g_{li}^{M} = r_i \; C_{li}^{TM} / \tilde{U}_{l\cdot},$$

$$g_{l}^{W} = 1.0 - \left(\sum_i \sum_k g_{lik} + \sum_i g_{li}^{M} \right).$$

Note that additional mark-recapture parameters are introduced in these equations; r_i is the proportion of tagged fish that are reported in year *i*. Two alternate hypotheses regarding the reporting rates are explored. The first is that reporting rates have been constant over the time-series. In this case the parameters, r_i are replaced with a single parameter (i.e., $r_i = r^*$). Alternately, the r_i have an autoregressive form, that is

$$r_1 = r^*$$

$$r_i = \exp\left[\ln\left(r_{i-1}\right) + \delta_i\right] \qquad \text{for } i > 1$$

where the δ_i are independent normally distributed random variables.

The negative of the log-likelihood for the mark-recapture component of the model, assuming a multinomial error structure for tag recovery data is then,

$$\begin{aligned} & - w_T \sum_l \left\{ \sum_i \sum_k \left[\tilde{C}_{lik}^{T} \ln(g_{lik}) \right] + \sum_i \left[\tilde{C}_{li}^{TM} \ln(g_{li}^{M}) \right] + \tilde{W}_l \ln(g_l^{M}) \right\} \\ & + w_R \sum_{i=2}^{n} \delta_i^2 \end{aligned} \tag{3}$$

where the constants w_T and w_R are adjusted to reflect prior assumptions about the variance of the tag recovery data and the trends in reporting rates, respectively. The additional fundamental parameters for the mark-recapture component of the model, i.e., ones that are estimated through the minimization, are G, G^T, F_{80}^{TM}, F_{96}^{TM}, L, r^*, and δ_i. The full objective function is the sum of equations 1, 2, and 3. Note that the model as described can also be implemented as a model without migration if the value of the parameters; G, G^T, F_{80}^{TM}, and F_{90}^{TM} are fixed at 0. This implementation of the model will be termed the "non-migration" model, as opposed to the "migration" model where all the mark-recapture related parameters are estimated.

Model Implementation

The model, as described, is implemented using AD Model Builder software (Otter Research Ltd. 1994). This software uses automatic differentiation for calculating the derivatives needed for finding the posterior mode via a quasi-Newton function minimization routine. The software provides estimates of the variance-covariance matrix for all dependent and specified independent variables. Likelihood profiles of key quantities of interest are also produced. The following list shows model parameters whose values are either fixed or constrained, except where explicitly stated otherwise.

Fixed parameters	Constrained parameters
$w_C = 5{,}000$	$G \geq 0$
$w_E = 50$	$G^T \geq 0$
$w_T = 0.5$	$s = 0.15$
$w_R = 20$	$0 \leq L \leq 1.0$
$\mathrm{M} = 0.08$	$0 \leq r \leq 1.0$
$\delta_i = 0.0$	

Data Sources

Analyses using the integrated catch-age mark-recapture model were conducted for the northern B.C. and the southern B.C. sablefish stocks with the data segregated into three depth strata (<500 m, 500-800 m, >800 m). The commercial sablefish fishery is composed of three gear types (trap, longline, and trawl), with the trap fishery accounting for 80 to 85% of the total annual landings. There is no program to obtain biological samples from the commercial catch; however, since 1980 biological data have been collected during research trips onboard commercial trap fishing vessels using commercial gear. In some cases the commercial vessels were chartered specifically for research purposes (e.g., tagging and abundance surveys) and in other cases research activities, primarily tagging, occurred during regular commercial fishing operations. Because these research activities used commercial trap gear and fished in commercial fishery locations, age composition samples collected during the research operations should reflect fishing selectivities similar to the commercial trap fishery. However, age and sex-specific fishing selectivity for the longline and trawl fisheries probably differs from the trap fishery. The estimates of the age and sex composition obtained during the research operations are used to approximate the age and sex composition of the entire commercial catch. It is possible that these are biased estimates of the removals from the stock, particularly for the longline and trap components. A potential bias arising from the assumption of common fishery selectivities for the different gear types would be worth exploring, but is beyond the scope of the current analysis. Annual catch data, by stock and depth strata, are shown in Table 1 for the 1980 to 1996 period. Annual age and sex composition data is plotted in Fig. 1 and summarized in Table 2.

Annual relative abundance indices were calculated from survey CPUE (catch in numbers per trap) data for each depth strata. The mean annual CPUEs are assumed to be indices of fish density and these estimates were weighted by the relative size of the depth strata (Saunders and McFarlane

Table 1. Catch in metric tons by depth strata for two B.C. stocks.

	Southern B.C.				Northern B.C.			
Year	<500 m	500-800 m	>800 m	Total	<500 m	500-800 m	>800 m	Total
1980	422	1,466	302	2,192	384	896	321	1,602
1981	303	924	106	1,334	660	971	775	2,407
1982	260	1,116	149	1,526	541	932	864	2,338
1983	234	1,073	163	1,471	217	1,062	1,308	2,588
1984	141	1,487	10	1,639	250	1,399	327	1,977
1985	172	1,813	0	1,986	439	1,496	111	2,047
1986	930	1,181	70	2,182	742	1,380	143	2,265
1987	913	1,508	27	2,449	440	1,212	481	2,133
1988	1,609	1,101	117	2,828	291	1,797	565	2,654
1989	1,135	1,532	193	2,861	478	2,001	150	2,631
1990	1,188	949	182	2,319	1,252	1,382	157	2,793
1991	689	594	84	1,368	754	3,181	164	4,100
1992	628	245	142	1,016	592	3,462	300	4,355
1993	457	1,254	54	1,765	297	2,526	519	3,343
1994	571	1,360	48	1,980	515	2,372	274	3,162
1995	401	1,236	232	1,870	283	1,416	605	2,305
1996	564	736	33	1,334	213	1,615	260	2,089

1993, Table 5.6) to generate relative abundance estimates for each stratum. The depth strata area measurement were based on slightly different geographical units from those used for the current analyses, so the abundance indices may not provide accurate information on the relative abundance by depth.

Sablefish were tagged in British Columbia during the late 1970s through early 1980s and from 1991 to the present. Only tag release data since 1980 and recovery data from these taggings are used. Summaries of tag and recapture data are presented in Table 3. Of the tags applied to fish in the southern B.C. stock assessment region, approximately 84% were recovered in the southern region, with the remainder of the recoveries primarily in the northern B.C. assessment region and in Alaska. For the northern B.C. assessment region, 79% of recoveries were in northern B.C. and 12% were in Alaska.

Stock Reconstructions

A series of stock reconstructions were conducted using both the migration and the non-migration implementations of the sablefish model. For the migration version of the model all recoveries of tags which occurred

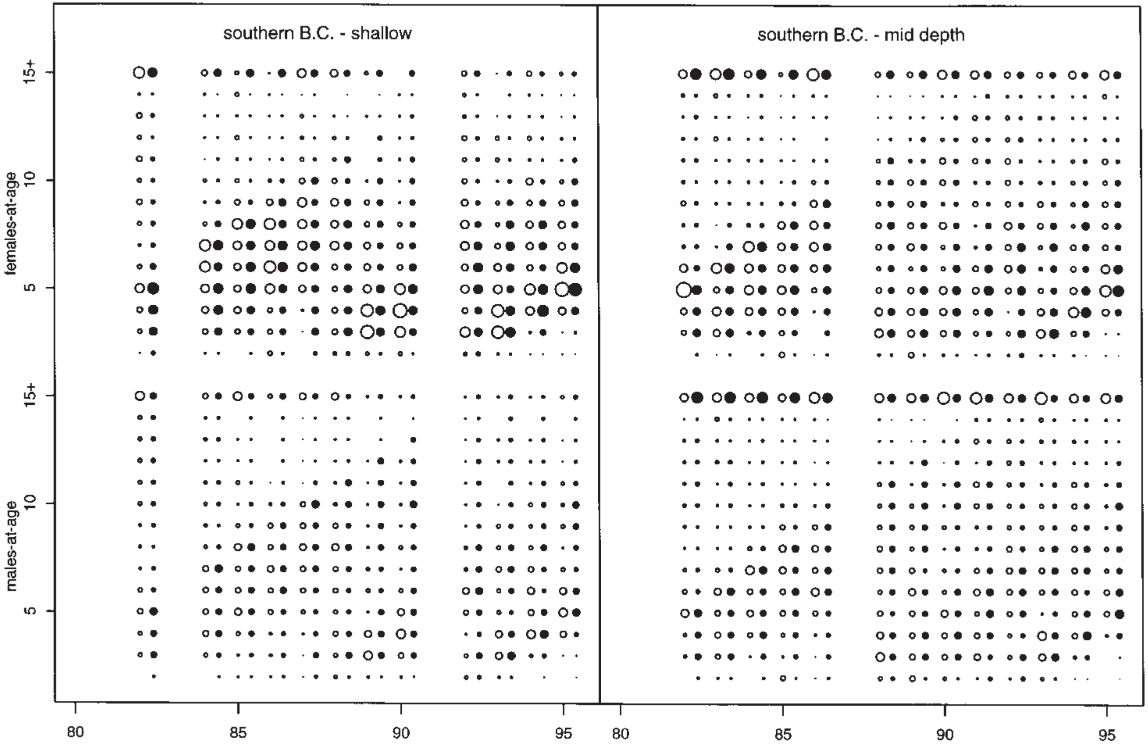


Figure 1. *Observed (open circles) and predicted (filled circles) proportion-at-age and sex for northern and southern B.C. depth zones, 1980-1996. Circle areas are proportional to the proportion-at-age and sum to one across age and sex.*

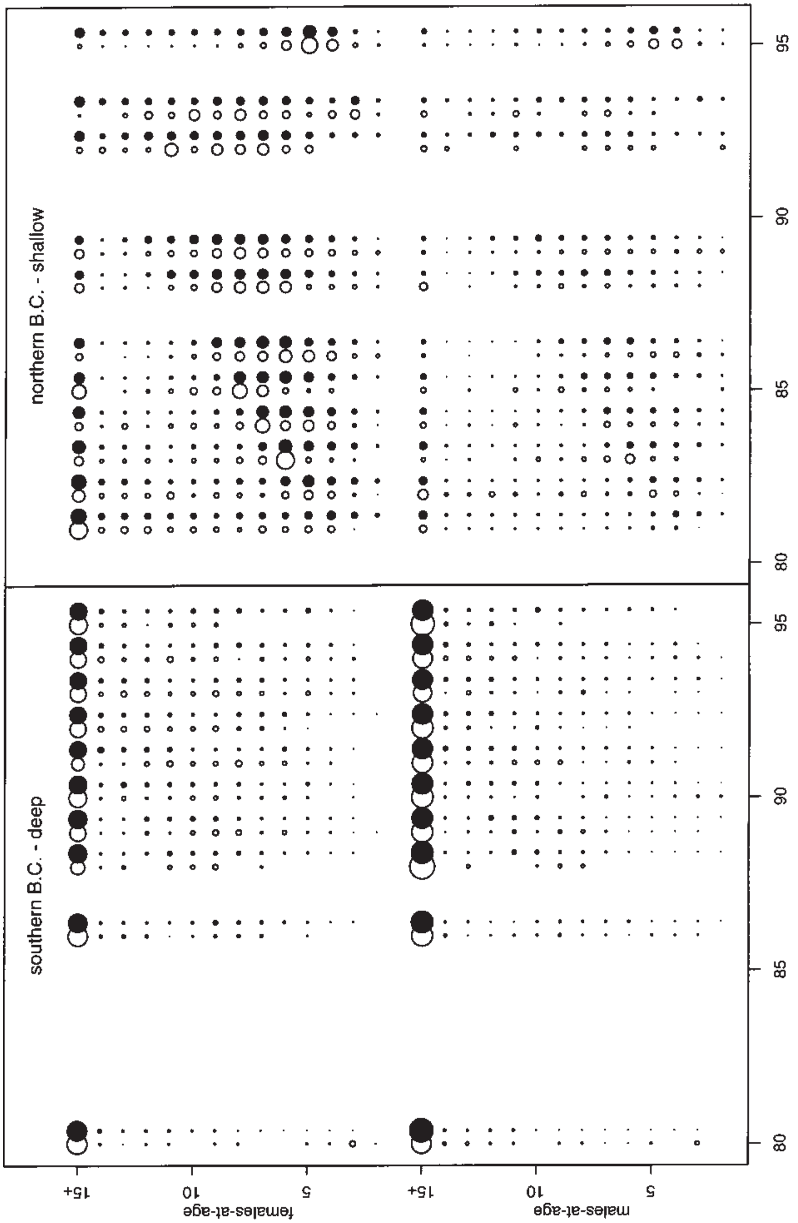


Figure 1. *(Continued.)*

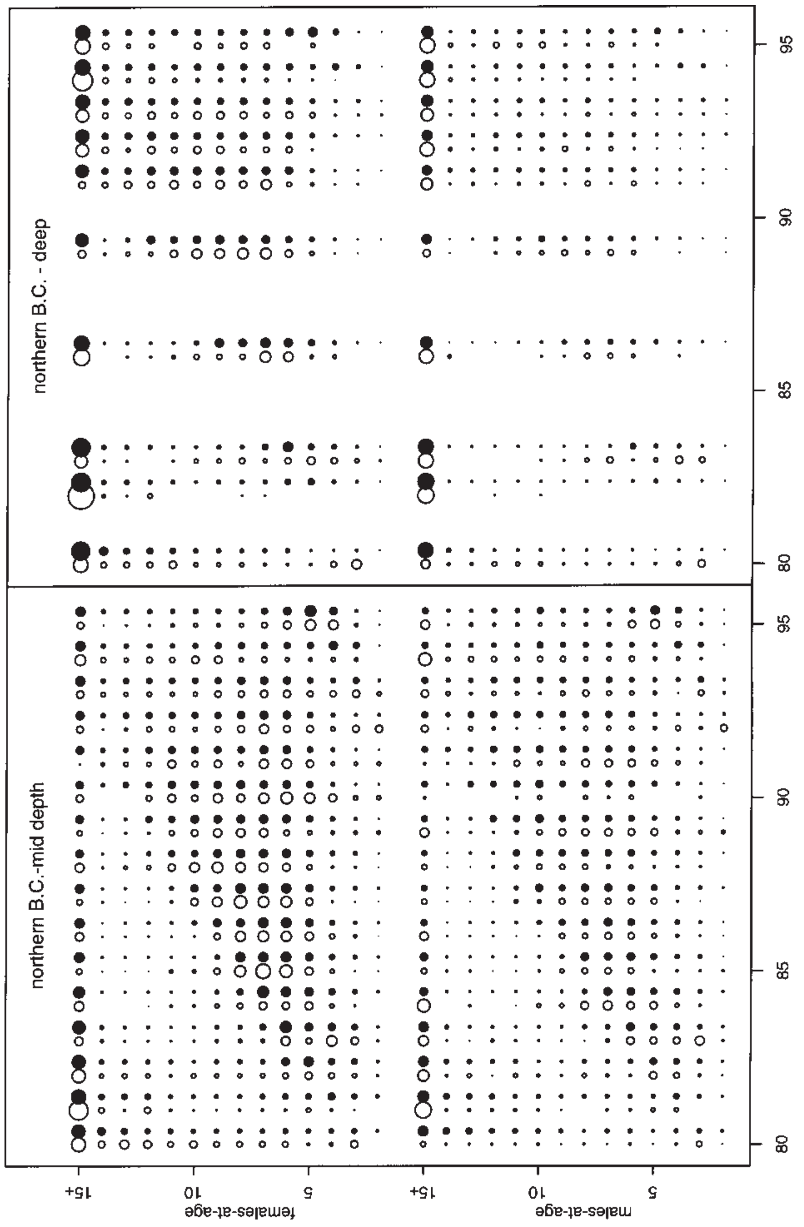


Figure 1. (Continued.)

Table 2. Average (1980-1996) observed percent-at-age and sex, by stock and depth strata.

	Observed															Predicted	
Strata	Sex	2	3	4	5	6	7	8	9	10	11	12	13	14	15+	Sex ratio	Sex ratio	15+
Southern B.C.																		
Shallow	M	0.8	3.0	4.6	4.3	3.9	3.2	3.0	2.4	1.7	1.1	0.8	0.8	0.4	3.7	34	40	3.1
	F	1.1	7.1	8.7	10.1	7.7	7.8	6.9	4.5	2.8	1.5	1.7	1.2	0.8	4.4			4.7
Middle	M	1.3	3.2	3.9	4.3	4.5	4.1	3.0	2.3	2.1	1.8	1.2	1.5	1.1	11.4	46	46	8.4
	F	1.3	4.2	6.0	8.2	6.0	5.0	4.2	3.3	2.4	2.3	1.3	1.1	1.0	8.0			7.6
Deep	M	0.1	0.6	0.4	0.4	0.4	0.5	0.9	1.2	0.9	0.9	1.0	1.4	1.0	43.6	53	54	42.1
	F	0.1	0.7	0.4	0.8	0.9	1.3	1.8	2.9	2.0	2.0	1.7	2.0	1.9	28.3			29.5
Northern B.C.																		
Shallow	M	0.7	0.7	2.1	3.1	2.8	2.7	1.8	1.3	1.0	1.5	0.7	0.7	0.5	4.4	24	26	3.7
	F	0.7	2.1	5.7	8.1	10.7	10.0	8.6	5.6	4.3	4.2	2.4	2.0	1.6	10.1			11.9
Middle	M	0.8	1.9	2.2	3.4	3.9	3.5	3.4	2.2	1.7	1.5	1.2	1.1	0.8	8.7	36	41	6.1
	F	1.1	2.7	3.9	6.0	7.7	7.8	6.3	5.2	3.8	3.0	2.6	1.9	1.7	10.0			10.5
Deep	M	0.1	1.2	1.3	0.6	1.4	2.2	1.8	1.5	1.7	1.3	1.5	0.8	1.2	18.2	35	34	14.6
	F	0.2	1.3	1.3	2.3	3.7	5.4	4.7	4.7	4.4	3.5	3.5	3.0	3.3	24.3			22.8

Percentages sum to 100 across age and sex for each depth stratum.

Table 3. Number of fish tagged by stock, year, and depth strata and number of tag recoveries by tag year and area.

	Number tagged by depth				Number of recoveries by area					
Year	<500 m	500-800 m	>800 m	Total	S.U.S.	S.B.C.	N.B.C.	N.U.S.	Total known	Total
Southern B.C.										
1980	0	4,410	275	4,685	17	732	84	30	863	930
1982	1,632	1,091	0	2,723	12	411	25	41	489	517
1991	0	525	963	1,488	3	49	2	4	58	73
1992	326	1,030	920	2,276	4	79	12	16	111	164
1993	530	2,045	1,957	4,532	12	190	34	9	245	331
1994	605	618	759	1,982	1	96	18	4	119	161
1995	1,925	2,258	965	5,148	1	200	9	6	216	284
Total	5,018	11,977	5,839	22,834	50	1,757	184	110	2,101	2,460
					(0.024)	(0.836)	(0.088)	(0.052)	(1.00)	
Northern B.C.										
1980	641	3,555	1,963	6,159	3	55	752	98	908	966
1981	123	1,901	95	2,119	4	17	227	53	301	320
1982	472	2,242	0	2,714	8	23	345	122	498	523
1991	0	555	403	958	1	3	107	27	138	173
1992	45	755	508	1,308	1	5	93	27	126	170
1993	170	1,552	765	2,487	1	8	218	43	270	371
1994	874	3,282	900	5,056	0	22	289	40	351	462
1995	577	9,959	257	10,793	0	73	633	68	774	1,060
Total	2,902	23,801	4,891	31,594	18	206	2,664	478	3,366	4,045
					(0.005)	(0.061)	(0.791)	(0.142)	(1.00)	

Proportion of total recoveries by area shown in brackets.

outside of the assessment (and tagging) region were treated as a single migrant pool. For the non-migration version of the model, tags that were recovered in areas outside the assessment area were treated as if they had been recovered within the assessment area. This treatment of the data would be consistent with the stock dynamics if there were no net stock migration (i.e., immigration is equal to emigration) and if the fishing mortality rates on the tagged migrant population were the same as the rates on the tagged non-migrant population. The first series of results presented are from analyses using the migration model.

Migration Model

Initial runs using the migration implementation of the model indicated that values for the migrant fishing mortality parameters (F_{80}^{TM}, F_{96}^{TM}) were unrealistically high so a series of runs were conducted where the value of these parameters were constrained to maximum levels ranging from 0.08 to 0.45. Estimates of the mark-recapture model parameters for this series of analyses are shown in Table 4. For the southern stock the 1980 fishing

Table 4. Parameter estimates from the migration model for the southern and northern B.C. sablefish stocks.

Constraint on F^{TM}	G	G^T	F_{80}^{TM}	F_{96}^{TM}	L	r	Function value	1997 biomass
Southern B.C.								
0.08	0.00	0.10	0.08	0.07	0.29	1.00	1173.8	27.6
0.15	0.00	0.06	0.15	0.13	0.31	1.00	1164.4	28.0
0.25	0.00	0.04	0.25	0.21	0.32	1.00	1157.2	28.5
0.35	0.00	0.03	0.35	0.30	0.32	1.00	1152.4	28.9
0.45	0.00	0.03	0.45	0.39	0.32	1.00	1148.9	29.2
Northern B.C.								
0.08	0.03	0.22	0.04	0.08	0.29	1.00	1595.2	25.6
0.15	0.01	0.19	0.06	0.15	0.29	0.76	1583.0	19.5
0.25	0.00	0.15	0.10	0.25	0.29	0.66	1575.8	17.4
0.35	0.00	0.12	0.13	0.35	0.29	0.63	1573.7	16.7
0.45	0.00	0.11	0.16	0.45	0.29	0.61	1573.2	16.2

mortality parameter is at the upper limit for all runs, while for the northern stock the 1996 estimates are consistently at the upper limit. Results from these runs show a high correlation between the fishing mortality parameters and the tag migration rate parameter, although estimates of 1997 exploitable biomass are relatively insensitive to these parameters. The value of the population migration parameter was 0 for most runs, indicating that the best fit to the observed catch and age composition data is obtained with no immigration to the population.

The residuals (standard normal deviates) of the predicted versus observed proportion of tag recoveries by recovery stratum are shown in Fig. 2, for the runs with the maximum fishing mortality parameters fixed at 0.15. For the migrant tag recoveries, the patterns of residuals follow a strong nonrandom pattern with primarily negative residuals for 0 and 1 year-at-large followed by positive residuals for 2 and greater years-at-large. That is, the model is unable to fit the observed attrition of tag recoveries for migrant fish that occurs over time. The analyses with high F values on migrant fish and low migration rates provide a somewhat better fit to the data observations because under this scenario the migrant pool of tagged fish decreases at a faster rate. With higher migration rates and lower F values the ongoing emigration of tagged fish maintains the numbers in the migrant pool. However, the nonrandom pattern of residuals holds even for the runs where the maximum value for the migrant fishing mortality parameter was fixed at 0.45. Clearly, the pattern in the residuals suggests that the model formulation for the migration implementation is inconsistent with the data observations.

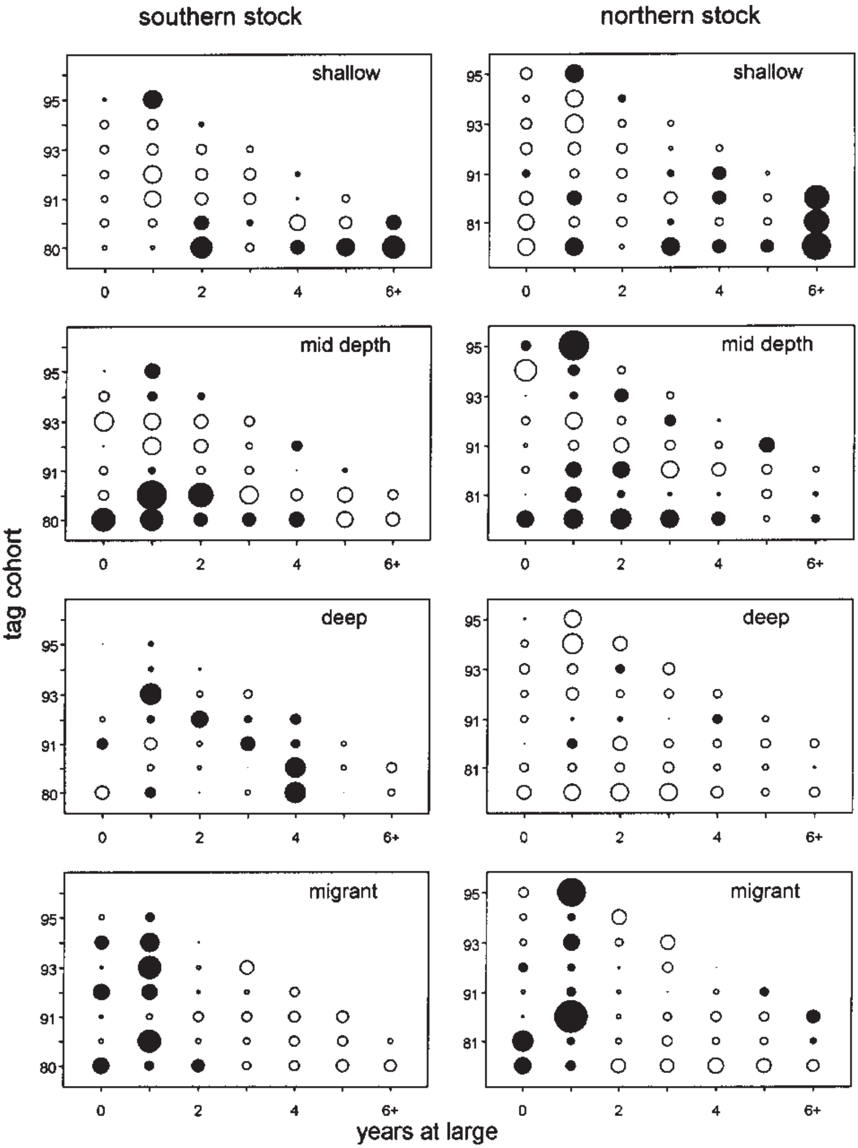


Figure 2. Tag recovery residuals by tag cohort and years-at-large for the migration model. Positive residuals are represented by open circles and negative residuals by closed circles. The circle area is proportional to the absolute value of the residual.

Non-Migration Model

A series of runs were conducted with the non-migration implementation of the integrated model to evaluate the sensitivity of parameter estimates to the penalty weights for the tag data and the survey abundance data. The weightings for the tag data are 1.0, 0.5, and 0.1, and for the survey data they are 500, 50, and 5. Results from this series of runs are shown in Table 5. For the runs with a high weight on the survey data, the model estimates of the proportions of the population in each depth stratum fit the observations almost exactly. With lower weighting on the fit to the survey data the model estimates a higher proportion of the population in the middle depth stratum for the northern population and a higher proportion in the shallow stratum for the southern population. The tag loss, reporting rate, and 1997 exploitable biomass estimates are relatively insensitive to the weightings, with the exception of analyses for the northern stock with low weight on the fit to the tagging data. The likelihood profiles indicate relatively precise estimates for these model parameters, given the model structure and data observations (Fig. 3). The fit to the age composition data is similar for all weighting combinations, and there is reasonable agreement between predicted and observed proportions-at-age (Fig. 1). For further analyses the penalty weights used are 0.5 for the fit to the tagging data and 50 for the fit to the survey relative abundance data.

Although the estimated values for the tag loss parameter appear to be relatively precise, they are high relative to expected values if this parameter accounts primarily for tag shedding. A series of runs were conducted with the tag loss parameter fixed at values ranging from 0.05 to 0.35 (Table 6). The model fits to the data observations, as measured by the objective function value, deteriorates significantly when the tag loss parameter value is less than 0.20. Also, the reporting rate parameter is highly correlated with the tag loss parameter.

For the analyses described so far, the value of the initial tag loss parameter was fixed at a value of 0.15, as reported by Beamish and McFarlane (1988). This parameter is intended to account for the immediate losses of tags from the tag cohort resulting from tag shedding and tag-induced mortality. Table 7 shows the results of model runs where the value of this parameter was fixed at levels ranging from 0.05 to 0.35. As the results for the northern stock show, this parameter is completely confounded with the reporting rate parameter. That is, as the initial tag loss value is changed, the values for the tag loss parameter, the 1997 exploitable biomass, and the function value remain the same and only the value of the reporting rate parameter changes. For the southern stock, the estimate for the reporting rate parameter is at its upper bound (1.0) so the correlation with the initial tag loss parameter is not apparent.

Because of concerns that the reporting rate for recaptured tags has changed over time, a final set of analyses was conducted with year-dependent

Table 5. Parameter estimates from the non-migration model resulting from alternative weightings for the tag data *(W_T)* and relative survey abundance data *(w_E)*.

		Northern B.C. Proportion at depth					
w_T	w_E	<500	500-800	>800	*L*	*r*	B
1.0	5	0.09	0.79	0.12	0.26	0.59	13.0
	50	0.16	0.72	0.12	0.26	0.59	12.9
	500	0.22	0.65	0.13	0.26	0.59	12.9
0.5	5	0.08	0.81	0.11	0.26	0.59	13.6
	50	0.15	0.73	0.11	0.26	0.59	13.6
	500	0.21	0.65	0.13	0.26	0.60	13.8
0.1	5	0.06	0.83	0.11	0.27	0.63	16.3
	50	0.07	0.82	0.10	0.23	0.51	11.9
	500	0.11	0.77	0.11	0.16	0.35	7.9
Observed:		0.22	0.64	0.13			

		Southern B.C. Proportion at depth					
w_T	w_E	<500	500-800	>800	*L*	*r*	B
1.0	5	0.72	0.11	0.17	0.30	1.0	29.7
	50	0.54	0.20	0.26	0.28	1.0	28.0
	500	0.45	0.22	0.33	0.27	1.0	28.4
0.5	5	0.74	0.09	0.17	0.29	1.0	29.4
	50	0.52	0.21	0.27	0.27	1.0	26.6
	500	0.45	0.22	0.33	0.26	1.0	26.8
0.1	5	0.67	0.14	0.19	0.28	1.0	27.8
	50	0.48	0.21	0.31	0.26	1.0	25.5
	500	0.45	0.22	0.34	0.25	1.0	25.6
Observed:		0.44	0.22	0.34			

Model parameters are: tag loss rate *(L)*, tag reporting rate *(r)*, and 1997 exploitable biomass (B).

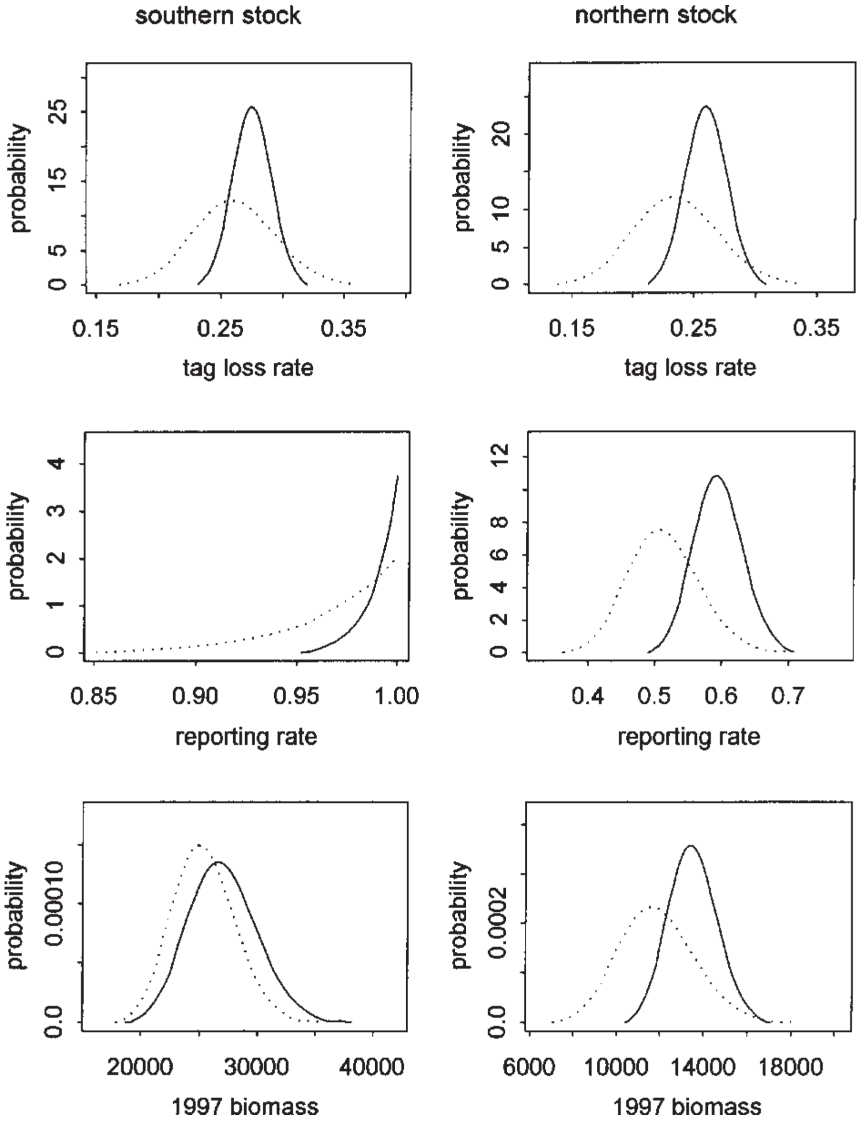


Figure 3. Likelihood profiles for tag loss rates, tag reporting rates, and 1997 biomass estimates. The solid lines and dashed lines are from analyses with penalty weights of 0.5 and 0.1 on the tag data, respectively. The weighting for the survey data is 50 in both cases.

Table 6. Parameter estimates for alternative runs of the non-migration model with the tag loss parameter fixed at various levels.

	Southern B.C.			Northern B.C.		
Tag loss	Reporting rate	Function value	1997 biomass	Reporting rate	Function value	1997 biomass
0.05	0.59	1,272.8	22.5	0.30	1,718.9	6.9
0.10	0.82	1,227.0	29.5	0.36	1,673.5	8.4
0.15	1.00	1,188.1	33.5	0.43	1,643.3	10.0
0.20	1.00	1,162.8	30.1	0.50	1,625.9	11.6
0.25	1.00	1,151.2	27.6	0.58	1,619.4	13.3
0.26				0.59	1,619.2	13.6
0.27	1.00	1,149.9	26.6			
0.30	1.00	1,151.2	25.6	0.66	1,622.1	14.9
0.35	1.00	1,160.8	23.9	0.74	1,632.8	16.5

The best model fits are underlined.

reporting rate parameters. The time series of estimated reporting rates is shown in Fig. 4. The estimated reporting rates are similar for the northern and southern B.C. stocks for the 1986 to 1996 period, but are higher for the southern stock in the earlier years. Under the variable reporting rate scenario, the estimated stock trajectories change somewhat with a lower 1997 exploitable biomass estimate for the southern stock and a higher terminal biomass estimate for the northern stock.

Discussion

The inability of the migration version of the integrated model to fit observed recoveries of migrant tags is of concern because significant numbers of tags are recovered outside the tagging area, suggesting migration is an important aspect of stock dynamics. The lack of model fit may be caused by a number of factors. The structure of the integrated model assumes that tag migration is a permanent event. That is, after tagged fish leave the population they do not return. There is some evidence suggesting that this assumption is not appropriate for sablefish in the eastern Pacific. That is, analysis of tag return data shows a tendency for smaller fish to move in a northwesterly direction and larger fish to move in a south and easterly direction (Heifetz and Fujioka 1991, Maloney and Heifetz 1997). This pattern suggests that there may be different phases to the ontogenetic dispersion of sablefish where the direction of migration is age-or size-dependent. If the sablefish migration pattern is more complex

Table 7. Parameter estimates for alternative runs of the non-migration model with the initial loss parameter fixed at various levels.

	Southern stock				Northern stock			
Initial loss	Tag loss	Reporting rate	Function value	1997 biomass	Tag loss	Reporting rate	Function value	1997 biomass
0.05	0.29	1.00	1,140.6	29.6	0.26	0.53	1,619.2	13.6
0.15	0.27	1.00	1,149.9	26.6	0.26	0.59	1,619.2	13.6
0.25	0.26	1.00	1,161.7	23.5	0.26	0.67	1,619.2	13.6
0.35	0.23	1.00	1,177.1	20.3	0.26	0.77	1,619.2	13.6

than what has been modeled, for example, age- or sex-specific migration or seasonal migration, model parameter estimates will likely be biased.

Another model assumption that may account for the lack of fit of the migration implementation is that tag reporting rates are the same for the migrant and non-migrant fish. Beamish and McFarlane (1988) suggest that most tags recaptured in British Columbia are reported, and based on this Heifetz and Fujioka assume a tag reporting rate of 0.90 for British Columbia. For Alaska fisheries, reporting rate estimates range from 0.25 (Heifetz and Fujioka 1991) to 0.50 (Bracken 1983). Results from the analyses presented here suggest higher tag return rates for British Columbia than those estimated for Alaska. If the U.S. reporting rate is lower than the B.C. rate, the integrated model will underestimate migration rates and thus overestimate the numbers of tagged fish at large in both the migrant and non-migrant components of the population.

An extension to the single-stock analytical approach presented here, to a multi-stock model that explicitly allows movement between all stock components may provide more consistent fits to the tagging data. For the single stock model, the migration rate parameters are highly correlated with the fishing mortality rates for migrant fish. In a multi-stock analysis, the additional information from catch-age data for each stock may allow relatively independent estimates for these parameters. Alternately, treatment of all sablefish in the eastern Pacific as a single population, as suggested by Beamish and McFarlane (1988), may be the appropriate geographic scale for stock assessment, given the apparent high migration rates between stock units.

For the non-migration implementation of the integrated model, the parameters representing ongoing tag loss and reporting rates appear to be well determined, given the model structure and data observations. That is, although the parameters are correlated, their estimates are precise and robust to alternative model assumptions. The parameter representing the immediate loss of tags from the population is completely confounded

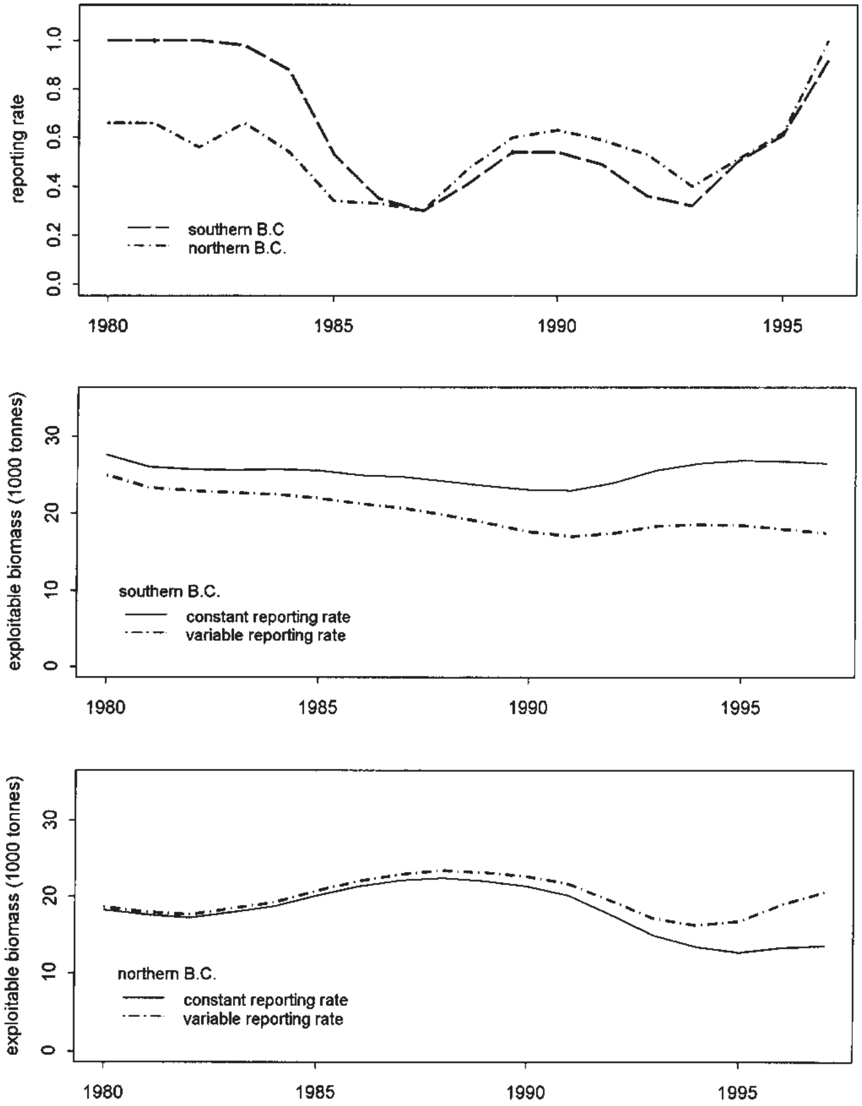


Figure 4. Estimates of time-dependent reporting rates and resulting exploitable biomass estimates for northern and southern B.C. stocks.

with reporting rates, so either one parameter must be fixed at an assumed value or a single parameter accounting for the joint effect can be estimated. The similarity between the estimates of the annual reporting rates for the northern and southern B.C. stocks suggests that there may be adequate information in the tag recovery data to obtain reasonable estimates for these parameters. The B.C. sablefish fleet fishes both of these assessment regions, so there is no reason to believe that reporting rates should be different in the two areas. If the model estimates of the reporting rates were driven by random noise rather than an underlying signal in the tag return data the high level of agreement in the estimates for the two stocks would not be expected.

The estimates for the ongoing tag loss parameter are unexpectedly high and suggest that there may be aspects of sablefish population dynamics that are not accounted for in the model. In mark-recapture analysis the parameter that I have termed “tag-loss” is generally attributed to tag shedding. However, the values of the tag loss parameter obtained in these analyses (0.26 – 0.27) are substantially higher than those estimated from double tagging experiments. Beamish and McFarlane (1988) estimated immediate tag shedding at 10% and ongoing tag shedding at 2% per year for sablefish tagged in British Columbia. Similar, low tag shedding rates were estimated by Lenarz and Shaw (1997) for sablefish tagged in the southern U.S. zone. Their estimates were 5% for immediate tag shedding and instantaneous ongoing shedding rates ranging from 0.03 to 0.07. Thus, it is likely that the tag loss parameter in the integrated model is accounting for the disappearance of tagged fish resulting from mechanisms in addition to tag shedding. Other potential sources of tag loss include ongoing tagging-induced mortality and migration to areas where there are no fisheries. Alternately, the model estimates of fishing mortality rates may be biased (i.e., underestimated).

References

Beamish, R.J., and G.A. McFarlane. 1988. Resident and dispersal behavior of adult sablefish (*Anoplopoma fimbria)* in the slope waters off Canada’s west coast. Can. J. Fish. Aquat. Sci. 45:152-164.

Bracken, B. 1983. Sablefish migration in the Gulf of Alaska based on tag recoveries. In: Proceedings of the International Sablefish Symposium, University of Alaska Sea Grant Report AK-SG-83-08, Fairbanks, pp. 185-190.

Deriso, R.B., T.J. Quinn, and P.R. Neal. 1985. Catch-age analysis with auxiliary information. Can. J. Fish. Aquat. Sci. 42:815-824.

Doubleday, W.G. 1976. A least squares approach to analyzing catch at age data. Int. Comm. Northwest Atl. Fish. Res. Bull. 12:69-81.

Fournier, D. and C.P. Archibald. 1982. A general theory for analyzing catch at age data. Can. J. Fish. Aquat. Sci. 39:1195-1207.

Heifetz, J. and J.T. Fujioka. 1991. Movement dynamics of tagged sablefish in the northeastern Pacific. Fish. Res. 11:355-374.

Hilborn, R. 1990. Determination of fish movement patterns from tag recoveries using maximum likelihood estimators. Can. J. Fish. Aquat. Sci. 47:635-643.

Hilborn, R., and C.J. Walters. 1992. Quantitative fisheries stock assessment: Choice, dynamics and uncertainty. Chapman and Hall, New York. 570 pp.

Kleiber, P., A.W. Argue, and R.E. Kearnew. 1983. Assessment of skipjack *(Katsuwonus pelamis)* resources in the central and western Pacific by estimating standing stock and components of population turnover from tagging data. Tuna and Billfish Assessment Programme, South Pacific Commission, Noumea, New Caledonia, Tech. Rep. No. 8. 38 pp.

Lenarz, W.H. and F.R. Shaw. 1997. Estimates of tag loss from double-tagged sablefish, *Anoplopoma fimbria.* Fish. Bull. U.S. 95:293-299.

Maloney, N.E. and J. Heifetz. 1997. Movements of tagged sablefish, *Anoplopoma fimbria,* released in the eastern Gulf of Alaska. In: M. Wilkens and M. Saunders (eds.), Biology and management of sablefish, *Anoplopoma fimbria.* NOAA Tech. Rep. NMFS 13:115-122.

Myers, R.A., N.J. Barrowman, and J.A. Hutchings. 1997. Inshore exploitation of Newfoundland Atlantic cod (*Gadus morhua)* since 1948 as estimated from mark-recapture data. Can. J. Fish. Aquat. Sci. 54(Suppl. 1):224-235.

Otter Research Ltd. 1994. An introduction to AD Model Builder for use in nonlinear modeling and statistics. Otter Research Ltd., Nanaimo, B.C.

Pollock, K.H. 1991. Modeling capture, recapture, and removal statistics for estimation of demographic parameters for fish and wildlife populations: Past, present, and future. J. Am. Stat. Assoc. 86:225-238.

Pope, J.G., and J.G. Shepherd. 1982. A simple method for the consistent interpretation of catch-at-age data. J. Cons. Cons. Int. Expl. Mer 40:176-184.

Ricker, W.E. 1975. Computations and interpretation of biological statistics of fish populations. Bull. Fish. Res. Board Can. 191. 383 pp.

Saunders, M.W., and G.A. McFarlane. 1993. Sablefish. In: B.L. Leaman and M. Stocker (eds.), Groundfish stock assessments for the west coast of Canada in 1992 and recommended yield options for 1993. Can. Tech. Rep. Fish. Aquat. Sci. 1919. 407 pp.

Saunders, M.W., B.M. Leaman, and G.A. McFarlane. 1997. Influence of ontogeny and fishing mortality on the interpretation of sablefish, *Anoplopoma fimbria,* life history. In: M. Wilkens and M. Saunders (eds.), Biology and management of sablefish, *Anoplopoma fimbria.* NOAA Tech. Rep. NMFS 130:81-92.

Saunders, M.W., G.A. McFarlane, M. Stocker, and B.M. Leaman. 1994. Sablefish. In: M. Stocker (ed.), Groundfish stock assessments for the west coast of Canada in 1993 and recommended yield options for 1994. Can. Tech. Rep. Fish. Aquat. Sci. 1975. 352 pp.

Seber, G.A.F. 1982. The estimation of animal abundance and related parameters. 2nd edn. Charles Griffin, London. 654 pp.

Integrating Ecosystem Studies: A Bayesian Comparison of Hypotheses

Milo D. Adkison, Brenda Ballachey, James Bodkin, and Leslie Holland-Bartels
U.S. Geological Survey, Anchorage, Alaska

Abstract

Ecosystem studies are difficult to interpret because of the complexity and number of pathways that may affect a phenomenon of interest. It is not possible to study all aspects of a problem; thus subjective judgment is required to weigh what has been observed in the context of components that were not studied but may have been important. This subjective judgment is usually a poorly documented and ad hoc addendum to a statistical analysis of the data. We present a Bayesian methodology for documenting, quantifying, and incorporating these necessary subjective elements into an ecosystem study. The end product of this methodology is the probability of each of the competing hypotheses. As an example, this method is applied to an ecosystem study designed to discriminate among competing hypotheses for a low abundance of sea otters at a previously oiled site in Prince William Sound, Alaska.

Introduction

Ecosystem approaches are increasingly advocated as a way of improving the science and management of natural systems (Lackey 1998). For instance, studies of the effects of anthropogenic stressors on a species can be misleading if they ignore possible indirect effects acting through predator or prey populations (Higashi and Patten 1989). Further, natural changes in these other components of the ecosystem may cause changes in the focal population, masking or exaggerating the effects of the stressor (Piatt and Anderson 1996). Many studies of the impacts of human actions on a particular species now include research on other components of the ecosystem thought to be important to the focal species.

Current address for Milo Adkison is Juneau Center, School of Fisheries and Ocean Sciences, University of Alaska Fairbanks, 11120 Glacier Highway, Juneau, AK 99801

Nonetheless, there are practical limitations to an ecosystem approach. Because of cost and logistical constraints, not all ecosystem components can be studied and therefore some indirect impacts may be missed. Experimentation or replication may not be possible, and it may thus be difficult to unambiguously assign causes to any observed differences in populations between impacted and non-impacted sites, or before versus after an impact at a single site. It is also highly likely that among the suite of studies, some will give results that are to some degree contradictory.

For these reasons, interpreting the results of an ecosystem study requires some degree of expert judgment. Synthesizing the results of numerous studies of parts of a complex problem is difficult, and it may thus be difficult for investigators to reach conclusions in a rational fashion. Further, different scientists faced with the same evidence may arrive at different conclusions. As the subjective interpretation of results tends to be an ad hoc and poorly documented process, the sources of disagreement may be difficult to uncover and resolve. This paper presents a structured method for documenting and quantifying the expert interpretation of the results of an ecosystem study.

Proposed Methodology

The methodology presented here is designed for testing ecosystem-level hypotheses. It integrates studies of diverse components of the ecosystem, summarizing the results as the relative evidence for each hypothesis from each study and the overall evidence for each hypothesis from the ensemble of studies. Its Bayesian features consist of incorporating and quantifying the subjective step of interpreting results, and calculating a probability that each hypothesis is true.

The method consists of the following steps:

1. Generate hypotheses
2. Summarize the experiments and their results
3. Create a table of the expected results under each hypothesis if each experiment were ideal
4. Calculate the probability of the observed result under each hypothesis using statistical considerations
5. Adjust probabilities by considering potential violations of statistical assumptions
6. Adjust probabilities to account for differences between the hypotheses tested and the hypotheses of interest
7. Summarize the evidence for each hypothesis, accounting for dependencies among experiments

Table 1. Hypothetical results of a set of ideal experiments.

	Hyp. 1	Hyp. 2	Hyp. 3	Hyp. 4
Study A	Positive	Negative	Negative	Positive
Study B	Negative	Negative	Positive	Negative
Study C	Positive	Positive	Positive	Negative

Steps 3-6 deal with eliciting statements of probability from experts. Such elicitations can be problematic if experts are unfamiliar with translating their experiments into numerical probabilities (Morgan and Henrion 1990, Ch. 7). Our sequence of steps is designed to overcome such problems by sequentially considering several sources of uncertainty, progressing from the most to least familiar. At each of the seven steps, in particular those where subjective judgment is required, the rationale leading to the decision should be thoroughly documented.

Step 1. Generate Hypotheses. The first step is to have the experts identify the hypotheses that are the competing explanations for the phenomenon under investigation. It is important that the hypotheses be both exhaustive and mutually exclusive. If not, the confidence assigned to some hypotheses will be overstated, as the evidence for them will in some respects be counted twice.

Often, there will be reason to believe that several of the hypothesized phenomena might act simultaneously. There are two principal ways of constructing mutually exclusive hypotheses if this is a possibility. The first is to consider a "multiple causes" hypothesis. The second is to redefine the hypotheses to allow minor effects of other factors. For instance, the two hypotheses "effect is produced by factor A" and "effect is produced by factor B" can be made mutually exclusive by redefinition as "effect is principally produced by factor A" and "effect is principally produced by factor B."

Step 2. Summarize the Available Data. In this step, the studies and their results are summarized. For clarity, it is often more useful to use a short verbal description of the results. For instance, a study of differences in prey abundance between control and treatment might be summarized as "much greater abundance found at the control site."

Step 3. Consider Ideal Studies. The third step in this process is to lay out a table with the different hypotheses as the top row and the different experiments as the left-most column (Table 1). Then, have the experts fill out this table as if each study were an ideal experiment; i.e., there was no possibility of either false positive or false negative results.

Table 2. Table of likelihoods.

	Hypothesis 1	Hypothesis 2
Study A	*P*(Result of *A*|Hyp. 1)	*P*(Result of *A*|Hyp. 2)
Study B	*P*(Result of *B*|Hyp. 1)	*P*(Result of *B*|Hyp. 2)

P(Result of *A*|Hyp. 1) means the probability of getting the observed result of Study A if Hypothesis 1 were true.

In the hypothetical example in Table 1, Study A would distinguish between Hypotheses 1 or 4 and Hypotheses 2 or 3. In combination, the three studies would be able to determine which hypothesis was true.

Step 4. Statistical Considerations. While ideally the three studies would determine which hypothesis was true with 100% accuracy, in the real world misleading results may be obtained. One of the ways this may happen is through random sampling error. Often, almost any result is possible under any of the hypotheses. Nonetheless, the observed result will be more probable under some hypotheses than others.

The objective of this step is to calculate these relative probabilities, otherwise known as the *likelihoods* of each of the hypotheses (Gelman et al. 1995, Ch. 1). Often, with continuously distributed variables, the likelihood is a probability density rather than a probability per se. Likelihoods (Table 2) are usually obtained from standard statistical distributions such as the normal or binomial. The exact distribution used depends upon the assumptions made about the experimental data, such as whether each point is independent and identically distributed, whether the sampling variance is constant, etc.

Table 2 shows the first of a series of steps in which experts are asked to assign probabilities to the competing hypotheses. Some experts are unfamiliar with quantitative probability statements and scientists in particular are often uncomfortable making assertions about the relative merits of competing hypotheses without conclusive evidence. This step is important in that it introduces experts to assigning probabilities to the hypotheses, yet does so in a rigorous way using familiar statistical calculations.

Step 5. Account for Possible Biases in the Test or Experimental Results. The assumptions of statistical tests are rarely exactly met. Samples may not be completely independent, important sources of error may not be included in the statistical model (e.g., ignoring error in the measurement of the independent variable), and measurements may have some unknown biases. Historically, statistical confidence tends to overstate the certainty of scientific results (Henrion and Fischoff 1986).

In constructing the table of likelihoods of results, this overconfidence needs to be accounted for. Generally, the effect of such errors is to make the probabilities of the result under each hypothesis more similar. Based on their knowledge of the experiment, experts should determine which assumptions of the test are likely to be violated, and to what degree. These judgments are to some extent subjective, but once made the statistical literature or computer simulations can provide guidance on their likely effects. In consultation with a statistician, the experts should adjust the table of probabilities to account for such violations.

Step 6. Account for Differences Between the Statistical Hypothesis Being Tested and the Biological Hypothesis That Is Actually of Interest. Often, an experiment to test a hypothesis tests it only indirectly. The results may thus be ambiguous if the indirect indicator could occur in several ways, some of which are not related to the hypothesis.

For example, if the hypothesis were that some population was affected by an environmental contaminant, an investigator might test the environment for the presence of the contaminant and test individuals for signs of poor health. A positive result in either case would not necessarily implicate the contaminant; the contaminant might be present yet not be causing health effects, or poor health might be due to causes other than the contaminant.

As in step 5, the effect of a difference between the hypothesis tested and the hypothesis of interest is to even further equalize the probabilities of the observed results under each hypothesis. The appropriate amount of adjustment of the table entries depends on the probability of other (possibly unknown) alternative explanations for the test results.

Such assessments are unavoidably subjective and require the judgment of experts. Hopefully, by this point in the process the experts are comfortable with assessing the relative probability of the data under each hypothesis and how violations of assumptions may result in misleading experimental results. It is crucial that they consider alternative explanations for their data yet not be paralyzed by such possibilities. They should be willing to examine data that seems to strongly favor one hypothesis and consider whether there are other, possibly unstudied ecosystem pathways that could produce similar results *and* state how probable they feel such pathways are.

Step 7. Summarize the Evidence. In this step, the table of probabilities is summarized to derive the overall weight of evidence for each hypothesis provided by the ensemble of studies. If the studies are independent, then elementary statistical theory says the joint likelihood of each hypothesis is simply the multiplication of its probability under each study (equation 1). The overall likelihood of each hypothesis is then simply the product of its column of probabilities (here $R1$, $R2$, and $R3$ signify the results of experiments 1, 2, and 3, respectively).

$$\text{Likelihood of hypothesis} = P(R1|\text{hyp.}) \times P(R2|\text{hyp.}) \times P(R3|\text{hyp.}) \quad (1)$$

The different hypotheses can then be compared in terms of their relative likelihoods. This comparison is easier if the likelihoods are re-scaled so that the sum of all of the likelihoods is 1. From a Bayesian perspective, each re-scaled likelihood could then be interpreted as the probability that a hypothesis was true.

Complication A. Dependencies among Results. There are two ways that experimental results might not be independent. First, the data from two experiments may have been taken from the same random sample. Second, two experiments may measure the same ecological phenomenon two different ways. In either case, it is not appropriate to treat the results as providing independent evidence bearing on the alternative hypotheses; i.e., simply multiplying the probabilities of the two experiments together will overweight the evidence.

There are several possible methods to account for dependencies among experimental results. If experiments are highly interdependent, they should be lumped and a single probability of each hypothesis calculated for the ensemble results. If experiments are only partially dependent, the correlation of results must be accounted for. If the correlation can be calculated, probability theory provides methods for calculating a joint probability. If not, a value must be obtained from experts, although experts have been found to perform poorly at providing a numerical value for correlation coefficients (Morgan and Henrion 1990, Ch. 7).

A more intuitive method for dealing with partially correlated results is to ask investigators to provide an estimate of the "effective" number of experiments. For instance, investigators may feel that dependence between two experiments is such that they jointly provide only as much evidence as 1.5 independent experiments. Then, the appropriate adjustment would be to raise each of the probabilities to the 0.75 power (e.g., equation 2). In general, if N experiments are correlated so that the effective number is E, probabilities for hypotheses for each experiment should be adjusted by raising them to the E/N power.

$$\text{Likelihood of hypothesis} = P(R1|\text{hyp.})^{0.75} \times P(R2|\text{hyp.})^{0.75} \quad (2)$$

Complication B. Prior Probabilities. Bayesian statistics involves multiplying the likelihoods by a set of prior weights (the prior probabilities) for the hypotheses before re-scaling to calculate the posterior probabilities. In the Bayesian approach, these prior probabilities reflect the weight accorded each hypothesis *before* the experiments were conducted. Assuming the probability of each hypothesis to be proportional to the joint likelihoods treats each hypothesis as being equally likely a priori, thus letting the data determine the relative probability of each hypothesis. While this is intuitively appealing, it may not be appropriate.

For instance, if the analysis were being used in a legal proceeding, it might be appropriate to give the benefit of the doubt to the defendant by assigning small prior weights to hypotheses implicating the defendant. Similarly, in investigating current scientific theory a high prior weight might be assigned to the currently accepted paradigm, so that a novel competing theory would not get much credence unless the evidence for it was overwhelming. An alternative to using prior weights is to calculate probabilities only from likelihoods, but require a very high probability that a hypothesis is true before acting on it. Whatever the prior weights, if data strongly support one hypothesis over the others the final probabilities will reflect this.

Standard Bayesian practice is to compare the evidence for competing hypotheses using Bayes factors (Kass and Raftery 1995). The Bayes factor is simply the ratio of the posterior probabilities of two competing hypotheses divided by the ratio of the prior probabilities assigned before the experiments were conducted. When the prior probabilities of the hypotheses are equal, this is simply the ratio of the posterior probabilities.

An Example: Sea Otters after the *Exxon Valdez* Oil Spill

On March 4, 1989, the supertanker *Exxon Valdez* spilled nearly 42 million liters of crude oil in Prince William Sound, Alaska (Spies et al. 1996). This spill is hereafter referred to with the acronym EVOS. Sea otter populations in oiled areas suffered high mortality (Loughlin et al. 1996). Other components of the ecosystem were likewise severely affected. Five years after the spill, residual oil was present in sediments and mussel beds in some areas of the spill (Spies et al. 1996). Even today, residual oil is found in some areas.

The Nearshore Vertebrate Predator (NVP) project (Holland-Bartels et al. 1996), a multi-university and agency investigation funded by the EVOS Trustee Council, is aimed at determining whether top predators in Prince William Sound are still suffering the effects of the oil spill. The question is difficult to answer unambiguously because of the complicated nature of the ecosystem and the lack of data from the period before EVOS. The NVP project studies predator populations from several points of view, and also looks at other components of the ecosystem on which these predators depend. If a population is still being affected by EVOS, the study is designed to ascertain whether the effects are due to the continuing toxic effects of oil, a slow rate of recovery from past mortality, or an indirect effect on some critical ecosystem component.

With limited resources and such an intensive approach, few populations can be studied. Sea otter abundance at Knight Island, which was oiled in 1989, is lower than at Montague Island, which was not. The NVP sea otter study has focused on these two populations, trying to find the reason for these differences in abundance. The principal hypotheses are:

1. **Direct toxicity of residual oil.** Residual oil is present and reducing the fecundity and/or survival of otters at the oiled site.
2. **Reduced forage due to oil effects.** The initial impact of oil or residual oil is reducing prey available to sea otters.
3. **Slow recovery due to demographic limitations.** Aside from the initial otter mortality from EVOS, residual oil is absent or does not affect otters or their food. However, limitations on the maximum growth rate of the population have prevented the population from reaching capacity yet.
4. **Natural differences in capacity.** The oiled site has poorer or less abundant otter habitat.

A variety of studies have been undertaken to determine which hypothesis is the most likely. These include:

1. **Demographic comparisons.** Population abundance, age structure, and reproductive rates were compared between islands.
2. **Individual health.** Otters were captured at both locations. Individuals were weighed and measured, and blood samples taken. In particular, blood cells and serum chemistry were examined for signals of poor health, and a specific signal of exposure to oil (the enzyme P450) was tested for.
3. **Prey abundance and foraging success.** The abundance and size distribution of major prey items of sea otters were compared among islands. In addition, foraging sea otters were observed to determine relative rates of success in obtaining prey items.

Statistical hypothesis tests were performed for many of the studies but are not reported here. We chose not to calculate likelihoods based solely on statistical distributions—step 4 of our methodology—because the limitations imposed by the design of the study tended to emphasize the considerations dealt with in steps 5 and 6. There are multiple predictions from each of the hypotheses, not all of which are distinct. Any particular study result may eliminate some hypotheses but leave several others. More likely, any particular study result would be ambiguous, as there is a small likelihood of almost any result from each hypothesis. In particular, the detection of a phenomenon does not necessarily imply that this was the cause of the difference in abundance between the two islands. For instance, oil could be present but yet not greatly affect survival. Likewise, prey abundance could differ between one site and another but be unrelated to the difference in otter abundance.

Table 3. First attempt at integrating studies.

Experiment and (result)	"A" Demographic limitation	"B" Food limitation	"C" Oil persistence	"D" Recovery has occurred
Otter density (K << M)	0.9	0.9	0.9	0.3
Repro. rates (equal)	0.9	0.5	0.7	0.9
Blood chemistry (equal)	0.9	0.7	0.3	0.9
P450 (equal)	0.7	0.7	0.1	0.9
Prey abundance (M < K)	0.9	0.1	0.1	0.1
Foraging success (M < K)	0.9	0.1	0.7	0.1
Joint likelihood	0.4133	0.0022	0.0013	0.0022
Probability of hypotheses	98.6%	0.53%	0.32%	0.52%

Top row gives hypotheses, and left column gives experiments with the results in parentheses. "M" refers to Montague Island (control), and "K" to Knight Island (oiled). The main body of the table gives the probability of obtaining each experimental result under each hypothesis. The bottom two rows summarize the result as the product of the probabilities for each hypothesis (i.e. the joint likelihood) and the probability products re-scaled to sum to 100%.

Thus, the interpretation of the results of the studies required some judgment. Our chief tool was to ask ourselves, "What is the probability we would get the result we observed from Study ___ if Hypothesis ___ was true?" We attempted to quantify our impression of the strength of each piece of evidence by filling out the table of probabilities, sequentially considering what the result would mean in an ideal world, what the statistical tests implied, how the assumptions of the tests might be violated, and what mechanisms might cause the results to be misleading.

We felt our ability to discriminate among probability levels was fairly coarse. Accordingly, we initially filled in the table of probabilities verbally, using the categories "high," "moderate-high," "moderate," "low-moderate," and "low," which we later replaced with 0.9, 0.7, 0.5, 0.3, and 0.1, respectively (Table 3).

The result of our first analysis was to assign more than a 98% probability to the hypothesis that the population differences were due to a demographic limitation in the rate of recovery of the Knight Island population from spill mortality. All other hypotheses combined had less than a 1.5%

probability of being true. We were unhappy with this result, as this high degree of confidence did not reflect our personal higher degree of uncertainty. We felt that the evidence for this hypothesis was not that strong.

In examining the reasons for this initial result, we identified three principal sources of error. First, we overstated the power of the studies to discriminate among hypotheses. For instance, we assigned a 0.90 probability of seeing greater prey abundance at the oiled site if demography was limiting recovery, but only a probability of 0.10 under any of the other hypotheses. We did not adequately address step 6 of our methodology; for instance, there would be a fairly good chance of seeing higher prey abundance at the oiled site under several alternative hypotheses.

Second, the range of hypotheses we considered was too narrow. In retrospect, we felt there was a strong possibility that all of the hypotheses might be incorrect, and some other factor might be responsible for differences between areas. This resulted in an unrealistically high probability for the hypothesis most consistent with the data.

Third, we did not adequately account for dependencies among experimental results (step 7, complication A). While we lumped most blood chemistry measures into one result, we kept the assay for the enzyme P450 (a more direct measure of exposure to oil) as a separate experiment. Since this assay could indicate the same phenomenon, and was measured on the same sample of animals, we felt the two results were effectively equivalent to only 1.5 experiments. Similarly, measures of prey size, prey abundance, and foraging success to some extent measured the same phenomenon. In retrospect, we decided to consider them as equivalent to two experiments.

We therefore revised the tabled probabilities, taking what we hoped was a more realistic look at the power of the studies and adding another alternative hypothesis to those we had listed. While we were able to think of several specific alternatives, we felt the true explanation for population differences might be something we hadn't considered. Therefore, we added only one hypothesis; an "unknown causes" category. Meanwhile, the completion of analyses of blood chemistry and the enzyme P450 suggested that residual oil might be present at the oiled site, and new information became available about the size distribution of prey species (Table 4).

The revised table again supports the hypothesis that the populations differ because the population in the oiled area has not had the time to recover fully from the losses due to the oil spill. However, it shows even greater support for the hypothesis that residual oil is still affecting the population. The hypothesis that some unknown factor accounts for the difference between populations is also quite probable.

Two hypotheses were eliminated from consideration, principally because of the forage abundance studies. Forage was more abundant and foraging success higher at the oiled site. These results were not at all

Table 4. Second attempt at integrating studies.

Experiment and (result)	"A" Demogr. limit	"B" Food limit	"C" Oil persist	"D" Recovered	"E" Unknown causes
Otter density (K << M)	0.9	0.9	0.9	0.3	0.9
Repro rates (equal)	0.9	0.5	0.7	0.9	0.9
Blood CBCs & chemistry (weak indication of liver damage at K)	0.5	0.5	0.7	0.3	0.5
P450 (M < K)	0.3	0.3	0.9	0.3	0.3
Prey abundance (M < K)	0.9	0.1	0.5	0.3	0.5
Prey size (M < K)	0.9	0.1	0.7	0.3	0.7
Foraging success (M < K)	0.9	0.1	0.7	0.3	0.7
Joint likelihood	0.1581	0.0011	0.1744	0.0040	0.0764
Probability of hypotheses	38.2%	0.3%	42.1%	1.0%	18.5%

Top row gives hypotheses, and left column gives experiments with the results in parentheses. "M" refers to Montague Island (control), and "K" to Knight Island (oiled). The main body of the table gives the probability of obtaining each experimental result under each hypothesis. The bottom two rows summarize the result as the product of the probabilities for each hypothesis (i.e. the joint likelihood) and the probability products re-scaled to sum to 100%.

consistent with the food limitation hypothesis, and were also unlikely if the population at the oiled site had recovered to its carrying capacity. However, it should be noted that the "unknown causes" hypothesis, which has a fairly high probability of being true, is not necessarily related to the spill. Thus it would be inappropriate to say the probability that the population is no longer suffering effects of the spill is only 0.01.

We will refine and expand this analysis as more data become available and more experts are consulted. These results are not our final interpretation, and should be viewed as a preliminary analysis. We provided this example solely to illustrate the use of the methodology.

Discussion

The Bayesian aspects of the proposed methodology are (1) use of subjective expert judgment in interpreting indirect tests of hypotheses, and (2) integration of experimental results and expert judgment into an overall probability for each hypothesis using Bayesian probability calculations. A large literature exists on using Bayesian methods to compare hypotheses (Kass and Raftery 1995).

Bayesian methods have been criticized from a variety of standpoints (e.g., Dennis 1996). The principal criticism is that Bayesian methods inject subjectivity into scientific analyses that should be objective. However, in extrapolating from the results of diverse studies on small aspects of a larger question, subjectivity in the form of expert judgment is unavoidable. We propose a methodology that formalizes the intuitive process experts use in interpreting the results of ecosystem studies. This approach clearly distinguishes subjective interpretation from experimental results, and clearly shows the reasoning used.

Our methodology provides a tool for investigators to organize their thinking. The ecosystem and the results of the numerous studies may be too complex to be readily grasped in their entirety. By allowing investigators to approach the synthesis of the studies one element at a time, our method increases the tractability of the process.

The methodology also facilitates openness and discussion, since subjective components of the synthesis of the studies are documented and quantified. It clearly shows why a particular conclusion was reached, and what evidence investigators felt was ambiguous or particularly strong. Areas of disagreement among investigators are also easily identified.

Our methodology is based on principles derived from other methods widely used for eliciting probabilities from experts (summarized in Morgan and Henrion 1990, Ch. 7). Examples of such methods include the Stanford/SRI protocol (Spetlzer and Stael von Holstein 1975, Merkhofer 1987) and the Wallsten/EPA protocol (Wallsten and Whitfield 1986). We've tailored our methodology to the specific goal of summarizing the relative support for alternative hypotheses from an interrelated but necessarily incomplete set of studies.

Most methods for probability elicitation pay great attention to getting experts comfortable with the idea of translating their knowledge and judgment into probability statements, and to overcoming a tendency of experts to give probabilities that overstate the level of certainty (Tversky and Kahneman 1982; Morgan and Henrion 1990, Ch. 7). Our solution to these difficulties is to take experts through a specific sequence of probability elicitation steps. These start with specifying deterministic outcomes, then progress through familiar specifications of probability (likelihood calculations) to less familiar probability specifications (the effects of violation of statistical assumptions and of not directly testing the hypothesis of interest). This sequence gradually introduces the process of making probability statements. It also sequentially introduces more and more forms of uncertainty, continually forcing the expert to reflect on whether the degree of confidence he's previously expressed is appropriate.

Our example illustrates both the utility and limitations of the methodology. The summary table lists the hypotheses and the experimental results. Probabilities within the table explicitly document the experts' interpretation of the consistency of the results of each experiment with each hypothesis. The summary probabilities excluded two hypotheses but retained three others, one of which appears to be only half as probable as the other two.

However, the 18.5% probability assigned to the "Unknown Causes" hypothesis makes interpretation of the other probabilities somewhat ambiguous. Much of the probability assigned to this hypothesis may indicate that recovery has occurred, and the differences we found are caused by some unknown factor(s) unrelated to the spill. It is also possible that "unknown causes" represents effects related to the spill such as cascading ecological effects. In either case, the results do provide guidance for further research; they suggest that continuing studies should focus on hypotheses "A," "C," and "E."

The necessity for re-evaluating our initial analysis because of unrealistic results in instructive. It reinforces the experience of others who have found that numerical statements of probability given by experts tend to be overly confident (Tversky and Kahneman 1982, Henrion and Fischoff 1986). Our second try produced a result that we felt better reflected the strength of the evidence provided by the experiments.

There is a danger that allowing such reanalysis could result in investigators juggling numbers to arrive at a result that reflected their preconceptions. However, an honest reappraisal of each element in the table is not inappropriate. Most methods for probability elicitation do recommend that assessors return to an earlier phase in the process whenever questioning reveals that the probabilities elicited clearly don't reflect the expert's judgment (Kadane et al. 1980; Morgan and Henrion 1990, Ch. 7; Laskey 1995). We found the reanalysis of the table caused us to re-examine the basis of our interpretations; rather than reinforcing our preconceptions, it tended to make us change them.

Use of our methodology will make it easier to examine the source of differences in interpretation of a study. For example, a scientist who disagreed with our conclusions might find that the basis of his difference was the weight placed on the blood chemistry results. A sensitivity analysis to alternative interpretations would be easy to perform by replacing the disputed probability with an alternative value to see if this affected the conclusions.

This method is not proposed as a substitute for good experimentation. With scarce, poor quality, and ambiguous data the conclusion reached after applying this method will be that considerable uncertainty remains. However, in such situations this methodology may identify areas of major uncertainty and suggest fruitful lines of investigation. The major benefit of this approach is the explicit documentation and quantification of the unavoidable subjective interpretation of ambiguous results that arise in many ecosystem investigations. In contrast, when strong experimental designs are available that produce clear evidence, subjective interpretation will be minimized and investigators should reach consensus.

Acknowledgments

The authors wish to thank Tom Dean, Jennifer DeGroot, George Esslinger, Steve Jewett, Dan Monson, Chuck O'Clair, Alan Rebar, Paul Snyder, and Glenn VanBlaricom for their contributions. The EVOS Trustee Council provided financial support for this study.

References

Dennis, B. 1996. Discussion: Should ecologists become Bayesians? Ecol. Appl. 6:1095-103.

Gelman, A., J.B. Carlin, H.S. Stern, and D.B. Rubin. 1995. Bayesian data analysis. Chapman and Hall, New York.

Henrion, M., and B. Fischoff 1986. Assessing uncertainty in physical constants. Amer. J. Physics 54:791-798.

Higashi, M., and B.C. Patten. 1989. Dominance of indirect causality in ecosystems. Am. Nat. 133:288-302.

Holland-Bartels, L., B. Ballachey, J. Bodkin, T. Bowyer, T. Dean, L. Duffy, D. Esler, S. Jewett, L. McDonald, C. O'Clair, A. Rebar, D. Roby, P. Snyder, and G. VanBlaricom. 1996. Mechanisms of impact and potential recovery of nearshore vertebrate predators. *Exxon Valdez* Oil Spill Restoration Project Annual Report. (Restoration Project 95025), National Biological Service, Anchorage, AK.

Kadane, J.B., J.M. Dickey, R.L. Winkler, W.S. Smith, and S.C. Peters. 1980. Interactive elicitation of opinion for a normal linear model. J. Amer. Stat. Assoc. 75:845-854.

Kass, R.E., and A.E. Raftery. 1995. Bayes factors. J. Amer. Stat. Assoc. 90:773-795.

Lackey R.T. 1998. Seven pillars of ecosystem management. Landscape and Urban Planning 40:21-30.

Laskey, K.B. 1995. Sensitivity analysis for probability assessments in Bayesian networks. IEEE Trans. on Systems, Man, and Cybernetics 25:901-909.

Loughlin, T.R., B.E. Ballachey, and B.A.Wright. 1996. Overview of studies to determine injury caused by the *Exxon Valdez* oil spill to marine mammals. In: S.D. Rice, R.B. Spies, D.A. Wolfe, and B.A. Wright (eds.), Proceedings of the *Exxon Valdez* oil spill symposium. Amer. Fish. Soc. Symp. 18:798-808.

Merkhofer, M.W. 1987. Quantifying judgmental uncertainty: Methodology, experiences, and insights. IEEE Trans. on Systems, Man, and Cybernetics 17:741-52.

Morgan, M.G., and M. Henrion. 1990. Uncertainty: A guide to dealing with uncertainty in quantitative risk and policy analysis. Cambridge University Press, New York.

Piatt, J.F., and P. Anderson. 1996. Response of common murres to the *Exxon Valdez* oil spill and long-term changes in the Gulf of Alaska marine ecosystem. In: S.D. Rice, R.B. Spies, D.A. Wolfe, and B.A. Wright (eds.), Proceedings of the *Exxon Valdez* oil spill symposium. Amer. Fish. Soc. Symp. 18:720-737.

Spetzler, C.S., and C.-A.S. Stael von Holstein. 1975. Probability encoding in decision analysis. Management Sci. 22:340-352.

Spies, R.B., S.D. Rice, D.A. Wolfe, and B.A. Wright. 1996. The effects of the *Exxon Valdez* oil spill on the Alaskan coastal environment. In: S.D. Rice, R.B. Spies, D.A. Wolfe, and B.A. Wright (eds.), Proceedings of the *Exxon Valdez* oil spill symposium. Amer. Fish. Soc. Symp. 18:1-16.

Tversky, A., and D. Kahneman. 1982. Judgment under uncertainty: Heuristics and biases. In: D. Kahneman, P. Slovic, and A. Tversky (eds.), Judgment under uncertainty: Heuristics and biases. Cambridge University Press, Cambridge, U.K., pp. 3-20.

Wallsten, T.S., and R.G. Whitfield. 1986. Assessing the risks to young children of three effects associated with elevated blood-lead levels. ANL/AA-32, Argonne National Laboratory, Argonne, IL.

Assessment of Southeast Alaska Pink Salmon Abundance Based on Commercial Catch and Effort and Sex Ratio Data

Jie Zheng
Alaska Department of Fish and Game, Juneau, Alaska

Ole A. Mathisen
University of Alaska Fairbanks, School of Fisheries and Ocean Sciences, Juneau, Alaska

Abstract

Pink salmon *(Oncorhynchus gorbuscha)* is the most abundant salmon species in Southeast Alaska and supports an important commercial fishery. Like most salmon fisheries in Alaska, pink salmon fisheries in Southeast Alaska are managed by a fixed escapement policy. To achieve a targeted escapement, managers must know the abundance of the incoming spawning run. The accuracy of abundance information acquired inseason substantially affects the manager's ability to achieve management objectives. To improve accuracy of inseason forecasts of southern Southeast Alaska pink salmon runs, we incorporated sex ratio information into inseason forecast models to annually adjust timing and shape of the run timing curves. First, we developed a sex ratio index and subsequently evaluated three inseason forecast models—linear, nonlinear, and combined—using this index and cumulative catch of all gears or cumulative catch per unit effort of the seine fishery from 1983 to 1997. Based on a cross-validation evaluation of forecast accuracy, the nonlinear model outperformed the linear and combined models. Cumulative catch per unit effort was a better predictor than cumulative catch in the first three weeks (weeks 28-30) of a fishing season, and vice versa in the remaining five weeks. Inseason abundance estimations greatly improved the preseason forecasts. Incorporating sex ratios into inseason forecast models correctly adjusted the run timings during a large majority of years and thus improved overall forecasts starting in the second week. In weeks 29-32, the best performing model using sex ratios improved forecasts more than 30% over the best model without using sex ratios; improvements included averages of relative forecast errors, absolute deviations, or squared residuals. Averages of relative forecast errors in weeks 29-34 were less than 24% for the best performing model using sex ratios and less than 38% for the best model without using sex ratios, compared to 51% for preseason forecasts. Average relative forecast errors from the best model were less than 20% before the run midpoint and less than 14% after the run midpoint.

Fishery Stock Assessment Models
Alaska Sea Grant College Program • AK-SG-98-01, 1998

Progress and Problems in the Application of Management Procedures to South Africa's Major Fisheries

José A.A. De Oliveira
Sea Fisheries Research Institute, Cape Town, South Africa

Douglas S. Butterworth and Susan J. Johnston
University of Cape Town, Rondebosch, South Africa

Abstract

Management procedures have formed the basis of the regulation of two of the three major fisheries of South Africa (the demersal trawl fishery for hake, and the purse-seine fishery for anchovy and sardine) since 1990, and have recently been developed and implemented in the third (west coast rock lobster). Essentially, these procedures comprise a set of pre-agreed and possibly simple rules, tested by simulation to give an appropriate catch vs. risk tradeoff in the medium term, that routinely translate data from the fishery into a TAC (total allowable catch) each year. Uncertainty is dealt with directly, by requiring rules that provide robust performance over a range of plausible scenarios for resource status and dynamics. This can circumvent issues such as appropriate weightings of different data sources, which can prove problematic if TACs are to be based on an annual "best assessment" coupled to some biological reference point. The paper discusses some key experiences in developing and implementing management procedures for the three fisheries above, specifically: that robustness to model structure uncertainty is of greater importance than "optimal" estimation; that feedback-control procedures do indeed self-correct in practice; and that short-term sociopolitical considerations undercut longer-term objectives in selecting between alternative candidate procedures when quota holders do not have established long-term rights.

Introduction

Management procedures (MP) are described by Butterworth et al. (1997) as a set of clearly defined, possibly quite simple rules, which translate data from the fishery into a TAC (or other regulatory mechanism, e.g., maximum fishing effort) each year. These rules are first tested by simulation to ensure reasonably robust performance in terms of expected catches and risk to the resource, given the prevailing uncertainties about resource status and dynamics. The rules should be agreed upon by all parties concerned (scientists, industry, managers) before being implemented, and should specify exactly how the regulatory mechanism is to be calculated and what data are to be collected and used for this purpose. Once implemented, MPs should be left to operate "automatically" for a period of 3-5 years (i.e., scientists should not seek to alter the recommendations which they provide unless very strong evidence pointing to such a need becomes available). After such a period, MPs should be reviewed and modified as necessary in the light of any changes in understanding of the resource or fishery which may have occurred in the interim (Butterworth et al. 1997, Cochrane et al. 1998).

A key difference between the MP and the conventional fisheries management approach is that the latter requires a regular and time-consuming re-evaluation of data, assessment methodology and the process for setting the regulatory mechanism as well as an update of the actual assessment results, typically on an annual basis. MPs require a more comprehensive re-evaluation but at a less frequent interval (typically only once every 3-5 years), hence arguably leading to a more time-efficient process, and a better basis to prioritize and pursue longer term research aimed at resolving key uncertainties. They also allow for the fact that a proper appraisal of risk for most fisheries cannot be made for a management decision that applies to a single year only, but needs to be based on the repeated application of a decision-making formulation over a number of years.

MPs have formed the basis of the regulation of two of South Africa's three major fisheries since the early 1990s, namely the demersal trawl fishery for hake (two species of hake are caught: *Merluccius capensis* and *M. paradoxus*), and the purse-seine fishery for anchovy *(Engraulis capensis)* and sardine *(Sardinops sagax).* They have recently been developed and implemented in the third major fishery, that for west coast rock lobster *(Jasus lalandii).* MPs are in the process of being formally incorporated into the South African Fisheries Act as a requirement for management of the larger South African fisheries. The following extract is taken from the government's White Paper on marine fisheries policy for South Africa (Department of Environmental Affairs and Tourism 1997):

> Long-term management plans, which include operational management procedures, will be developed and published to ensure optimal utilization of all significant marine resources. They will:

- be developed through a cooperative process involving all interested parties
- be binding, though with procedures to allow amendments
- include appropriate and cost-effective monitoring and control programs and strict enforcement of fishing regulations
- consider the socioeconomic implications of altered levels of utilization (e.g. the effect of a reduced TAC on employment).

> Operational management procedures will be based on scientific principles recognizing the inherent variability of resources and the interdependence of the components of marine ecosystems.
>
> Finally and importantly, conservative interim approaches will have to be adopted where insufficient scientific information is available; research will be dedicated to filling the gaps in knowledge.

There has been ongoing debate about whether MPs as opposed to conventional annual “best assessment” type approaches are a better way to manage fisheries (Butterworth et al. 1997, Cochrane et al. 1998). It is not the intention of this paper to repeat these arguments, nor to provide a comprehensive account of the MPs used in South African fisheries. Rather, a few examples of their application in the South African context are provided to illustrate three key points:

1. Robustness to uncertainty about model structure is a more important concern than the development of an “optimal” estimator (assessment procedure) for a given model and assumed error structure for input data.

2. MPs involving feedback control mechanisms can indeed appropriately adjust TACs as new data show past perceptions to have been in error, though this does not obviate the need to eventually revise the MP to better reflect such changed perceptions.

3. Short-term sociopolitical considerations often undercut longer-term objectives when there is an immediate threat to the long-term security of quota holders.

These three points are illustrated by considering the South African sardine, hake and west coast rock lobster fisheries respectively.

Sardine

The South African pelagic fishery has recorded average annual landings of some 400,000 t for all species over the last ten years. Catches currently comprise two species for which TACs are set, namely anchovy and sardine, and four likely underutilized "non-quota" species of which round herring *(Etrumeus whiteheadi)* is the most important in terms of landed mass. MPs were first applied in the pelagic fishery to anchovy in 1991 and to both anchovy and sardine together in 1994. These procedures are well documented elsewhere (Butterworth and Bergh 1993, Butterworth et al. 1993, Cochrane et al. 1998, De Oliveira et al. 1998). The sardine procedure of 1994 is multispecies in nature because it accounts for operational interactions between anchovy and sardine, and also between round herring and sardine (De Oliveira et al. 1998). These interactions take the form of mixed shoaling, such that when either anchovy or round herring is targeted, a bycatch of sardine is unavoidable (0 aged fish with the former, 1+ aged fish with the latter). The 1994 sardine procedure was therefore a complex one which provided both a directed catch TAC (1+ fish only because 0 aged fish are not targeted for an operation aimed primarily at canning) and a bycatch TAC (both 1+ and 0 aged fish) whose size depended, inter alia, on the size of the anchovy TAC (primarily converted to fishmeal). When anchovy abundance fell to a very low level at the end of 1996 so that the directed anchovy fishery was suspended, a simpler MP for sardine was developed and implemented in 1997. This accounted for directed adult catches and a fixed bycatch of adult sardine with the round herring fishery. It is this procedure which is used here to illustrate the importance of structural uncertainty considerations.

The process of developing the MP for sardine had two phases. The first consisted of an assessment of the sardine resource to estimate its "status" and productivity under a variety of assumptions. The second involved the development and simulation testing of an array of alternative MPs. These tests were based on the results of the first phase, so that the alternative scenarios considered remained consistent with the available information. Finally, all parties concerned (scientists, managers, industry), participated in selecting an MP whose performance was suitably robust to key uncertainties about model structure and which achieved an appropriate tradeoff in satisfying conflicting objectives such as maximizing catch and minimizing risk to the resource. Figure 1 illustrates how these phases are linked.

Population parameter values were obtained by maximum likelihood estimation in the assessment procedure of the first phase, and were used as input into the testing process of the second phase (see Butterworth et al. 1993 and De Oliveira 1995 for a description of the methodology). The same age-structured population model was used for both phases; its key assumptions were 6 age-classes, pulse fishing for all ages in the middle of the fishing season and a year-invariant natural mortality whose value for the 0 age group might differ from that for the 1+ age group. The stock-recruit relationships required by the testing process of the second phase to provide levels of future recruitment were obtained from the assessments of the first phase by fitting stock-recruit curves to the estimates of

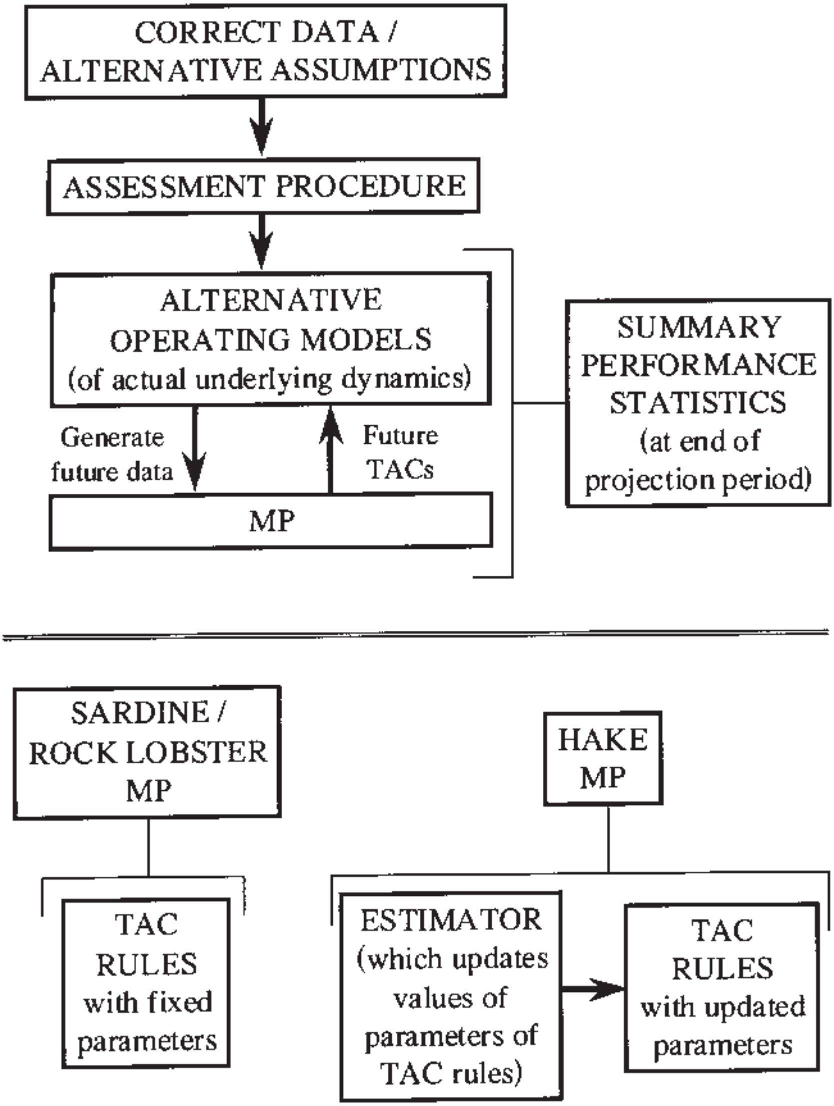


Figure 1. A flow diagram of the overall testing process for MPs. The structure of the MP component is the same for sardine and rock lobster, but differs for hake (see main text for details).

recruitment and 1+ biomass provided (Fig. 2). The data used in the assessment were

a. 1+ biomass estimates (B_{1+}) and sampling CVs. from hydroacoustic spawner biomass surveys conducted in November each year.

b. Recruit estimates and sampling CVs. from hydroacoustic recruit surveys held in May/June each year.

c. Age composition information from the surveys mentioned in (a).

d. Commercial catch-at-age data.

Data series (a)-(c) contributed to the likelihood function, which related model predictions to observations, where the model assumed data series (d) to be without error.

In the development of the 1997 sardine MP, the following key uncertainties for which it was considered a selected MP should show adequate robustness, were identified.

i. Levels of natural mortality for the 0 and 1+ age group (M_0 and M_{1+} respectively)—different values (reflecting various assumptions about the "true" values for M_0 and M_{1+}) had to be assumed, as the available data hardly admitted their direct estimation.

ii. Bias in the B_{1+} estimate resulting, inter alia, from possible bias in hydroacoustic target strength estimation.

iii. Additional variance (λ_{1+} and λ_r respectively) not explained by the sampling (i.e., inter-survey-transect) variability of the B_{1+} and recruit survey estimates—the λ and sampling variability summed together reflect the overall precision of the corresponding survey estimate in relation to the true abundance (increasing λ_{1+} or λ_r implies down-weighting the associated data in the likelihood function).

iv. Appropriate weighting (γ_{age}) in the likelihood function of the age composition data (decreasing γ_{age} implies down-weighting the age composition data in the likelihood function).

v. Choice of an appropriate stock-recruit (S/R) curve.

The testing process of the second phase essentially encompassed projecting the sardine population into the future (a 20-year projection period was used) from estimates of the current level of the resource provided by the assessment procedure of the first phase, by using the stock-recruit relationship fitted to recruitment and 1+ biomass estimates from the assessment results to provide future recruitment values. Two sources of variability in the data were taken into account, namely observation error

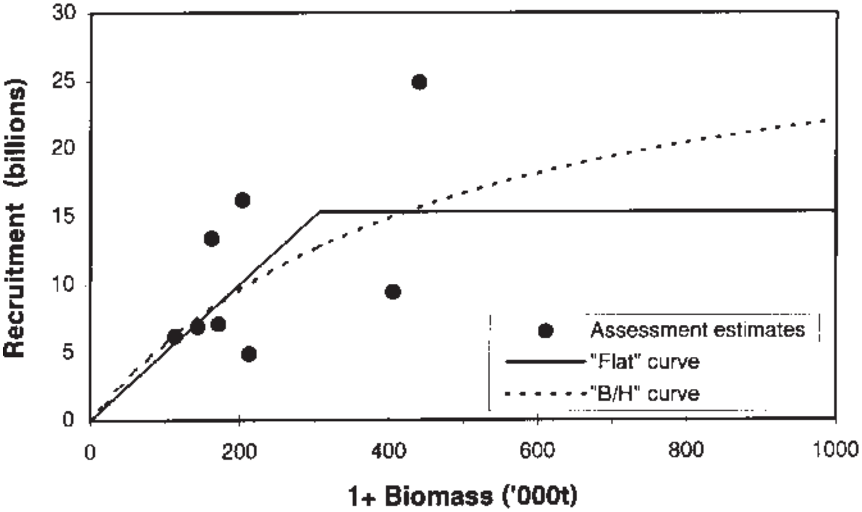


Figure 2. The stock-recruit curves considered for sardine. The "flat" curve decreases linearly from an estimated median recruitment level to zero below 20% of the average pre-exploitation 1+ biomass value, but reflects no dependence of recruitment on spawning stock size above this value. The "B/H" curve reflects a Beverton-Holt type stock-recruit relationship. The dots are estimates of 1+ biomass and recruitment from the assessment procedure.

in the form of sampling variability in simulated future 1+ biomass survey estimates (which are used in the formula which provides TAC recommendations—see below), and process error in the form of recruitment variability, which was based on the assumption that recruitment is lognormally distributed around the stock-recruit relationship. The future catches in these tests were as specified by the MP being tested. Selection between alternative MPs was carried out by considering the summary statistics described in the footnote to Table 1, and robustness to the key uncertainties listed in (i)-(v) above. For ease of presentation, only the MP that was eventually selected and implemented in 1997 is discussed below.

The 1997 sardine MP (intended for implementation in the absence of any anchovy fishery) is described by the following equations and constraints:

$$\begin{aligned} TAC_{dir}(y) &= 0.175\, B_{1+}(y-1) \\ TAC_{byc}(y) &= 10{,}000 \text{ t} \end{aligned} \tag{1}$$

Table 1. Results for the robustness trials for the 1997 sardine management procedure (MP) described by equations (1) and constraints (2) in the main text.

	Uncertainty							Summary statistics			
	M_{1+}	M_0	B_{1+} mult. bias	λ_{1+}	λ_r	γ_{age}	S/R curve	$\overline{C}_{tot}$	$\overline{g}_{1+}$	*risk*	$\overline{V}$
1	0.6	1.0	1.5	0.0	0.4	0.1	Flat	164	1.18	0.030	22%
2	0.4							186	1.37	0.022	16%
3		0.6						166	1.24	0.016	22%
4		1.4						156	1.11	0.048	23%
5			1.0					169	1.18	0.006	21%
6			2.0					153	1.17	0.052	24%
7				0.2				163	1.17	0.036	23%
8					1,000			147	1.19	0.012	25%
9						0.3		152	1.20	0.028	24%
10							B/H	179	1.61	0.104	17%

"Uncertainty" reflects key uncertainties about model structure or optimal estimation, as defined in points (i)-(v) in the sardine subsection of the main text. "Summary statistics" reflect the sensitivity of the 1997 sardine MP to the robustness trials with respect to the given performance attributes.

The summary statistics are defined as follows: $\overline{C}_{tot}$ is the average annual total catch (directed and bycatch combined) over the 20 year projection period; $\overline{g}_{1+}$ is the average 1+ biomass at the end of the projection period as a proportion of the 1+ biomass at the start of the projection period; *risk* is the probability that the 1+ biomass falls below 20% of its average value in the absence of exploitation at least once during the projection period; $\overline{V}$ is the average annual change in directed TAC as a percentage of the directed catch portion of $\overline{C}_{tot}$. The "flat" and "B/H" stock-recruit (S/R) curves are detailed in the caption to Fig. 2. Catch values are in thousand t, and empty cells acquire the corresponding value in robustness trial 1.

subject to

$$\begin{aligned} 70{,}000\text{ t} \le TAC_{dir}(y) \le 210{,}000\text{ t} \\ TAC_{dir}(y) \ge 0.75\, TAC_{dir}(y-1) \end{aligned} \qquad (2)$$

where

$TAC_{dir}(y)$ is the directed sardine TAC (1+ fish only) implemented at the start of the fishing season (January) in year y;

$TAC_{byc}(y)$ is the bycatch sardine TAC (1+ fish only) to accommodate the round herring fishery, implemented at the start of the fishing season (January) in year y; and

$B_{1+}(y-1)$ is the 1+ biomass estimate from the hydroacoustic spawner biomass survey held in November of year $y-1$.

Robustness of this procedure to the key uncertainties (i)-(v) is reflected in Table 1. Essentially, the results in this table are obtained by keeping equations (1) and constraints (2) unchanged, but inputting sets of population parameter estimates from the assessment procedure to the testing process that differ for each robustness trial, the appropriate changes having been made to the model structure of both the assessment procedure and testing process.

When the results of Table 1 were discussed with the industry, an important concern from a conservation point of view was that the *risk* level was contained below an acceptable limit across the range of model structure uncertainties considered, while at the same time some scope for growth of the resource was consistently ensured (an accepted goal in the sardine fishery since the early 1980s has been to rebuild the resource, Cochrane et al. 1998). The actual criteria used were that *risk* should not much exceed 0.1, and that the resource growth measure $\bar{g}_{1+}$ should exceed 1 (the definitions of *risk* and $\bar{g}_{1+}$ are given in the footnote to Table 1).

These two criteria were met for the 1997 sardine MP for all the robustness trials of Table 1, and hence led to its selection. However, it is clear that robustness trial 10, which assumes a different though still quite plausible shape for the stock-recruit relationship (given the data, see Fig. 2), has by far the worst effect on *risk*. In particular, this effect is much larger than that arising from statistical estimation considerations, such as the values used for λ_{1+}, λ_r and γ_{age} which amount to different weightings given to different sources of data in the likelihood. This highlights the point that considerations of structural uncertainty, such as the form of the stock-recruit relationship, are the much more important concern in this case.

Hake

The fishery for Cape hakes, with annual catches over the last ten years averaging 143,000 t, forms the backbone of the demersal trawl fishery, and is currently the most valuable fishery in South Africa (Crawford et al. 1987, Department of Environmental Affairs and Tourism 1997). Two species of hake occur in the fishery, and because they are morphologically very similar, the catch and effort statistics collected from the fishery have not been able to distinguish between them (Punt 1993). However, Punt (1992) found that aggregation of data for the two species should not compromise assessment results seriously. In the analyses described below, the two species have been treated as one, and only the west coast component of the hake fishery is considered.

The MP used for hake differs in one important respect from that just described for sardine (see Fig. 1). The latter reflects a “Myopic Bayes” approach, i.e., the MP is based on perceptions of resource parameter values at the time of its development, and no attempt is made to update these values as more data become available. This process is left to the 3-5 year review of the procedure itself. In contrast, the MP for hake encompasses feedback control; whereas the sardine MP sets TACs as a fixed fraction of the most recent survey biomass estimate, the hake MP updates population dynamics parameters given new data, and uses these updates in computing its TAC recommendations. Although this re-estimation appears to mimic the conventional management process, it differs in two important aspects (Butterworth et al. 1997). First, the combination of the estimation and the TAC calculation is an automatic process repeated every year and does not necessarily correspond to the "best" assessment of the resource possible at any one time. Second, as in the selection of the sardine MP, this procedure was chosen over other possibilities after extensive tests had been carried out of the anticipated performance of the fishery and its robustness to prevailing uncertainties in the medium term (Punt 1992, 1993). These characteristics shift the process into the realm of a "management procedure" (Butterworth et al. 1997).

The current MP for hake comprises a process for setting a TAC which includes an observation error estimator based on a dynamic age-aggregated surplus production model assuming a Schaefer form for the production function (Payne and Punt 1995). The estimator (Fig. 1) utilizes the following data:

a. Total biomass estimates (with CVs) from swept area surveys treated as an index of abundance in relative terms.

b. Commercial CPUE data (standardized using power factor analysis).

c. Total annual commercial catches.

A TAC is calculated directly by using an $f_{0.2}$ harvesting strategy (the $f_{0.2}$ harvesting strategy is a constant effort strategy, where the effort level is calculated from a surplus production model and is that for which the slope of the equilibrium yield vs. effort plot is 20% of the slope of this curve at the origin; Payne and Punt 1995 document the reasons for the selection of the $f_{0.2}$ strategy). Note that although an age-aggregated estimation model is used, its selection was based on superior performance in simulation tests conducted using a fully age-structured model, with its greater biological realism, to better mimic the actual underlying dynamics.

The plot in Fig. 3a illustrates the model fit of this estimator to observed CPUE data up to 1990, when this MP was first implemented. It also shows the corresponding fits in 1993 and 1996, after 3 and 6 years additional data respectively. Figure 3b shows the sequence of future TACs predicted by projecting the model fit in 1990 forward (using data up to 1990 only) under an $f_{0.2}$ harvesting strategy, and compares these to the TACs subsequently set under the MP. As time proceeded, the CPUE did not show as rapid an increase as expected, suggesting that resource productivity as assessed in 1990 was too high. However, the feedback nature of the MP (which automatically refitted the model each year to the data series updated by one year) allowed it to compensate for this, by indicating catch levels less than predicted earlier (see Fig. 3b).

Recent research suggests that these features in the recent CPUE data may be a consequence of a change in the age-specific selectivity pattern in the fishery away from younger fish. This is possibly the result of a reduction in the illegal use of net liners as the industry's economic "need" for this to enhance catch rates diminished as the resource size recovered. Therefore, the poorer fit to the data in the most recent years (Fig. 3a) may be a consequence of model misspecification. Such a possibility was not considered when the tests for the original MP were conducted, but is being taken into account in its re-evaluation which is currently in progress.

West Coast Rock Lobster

South Africa's west coast rock lobster resource has yielded an average commercial catch of 2,600 t over the past ten years. The resource was heavily fished earlier this century, and the current biomass of exploitable animals (>75mm carapace length) is estimated to be only 5% of its pre-exploitation level. Furthermore, following a period of low somatic growth that commenced in the late 1980s, this component of the biomass has dropped to about half of its mid-1980s level.

An MP for the west coast rock lobster fishery has only recently been developed and was implemented for the first time in 1997. The process followed in its development is similar to that for sardine and hake, namely that an assessment procedure determines the status of the resource, and forms the basis for the testing process from which an MP is selected by the parties concerned (Fig. 1). The MP selected, which is to apply for three years before revision, is strictly Myopic Bayes without feedback of new data to update estimates of population dynamics parameters. Nevertheless the manner in which these data are used to calculate TAC recommendations does have some self-correcting properties.

The assessment procedure uses maximum likelihood estimation to fit a size-structured sex-disaggregated model to observed data which include the following:

a. CPUE.

b. Catch-at-size frequencies for male and female lobsters.

c. Percentage females caught.

d. Somatic growth for males, females and juveniles.

e. Total commercial catch.

Data series (a)-(c) are available from two sources (one fishery dependent, the other fishery independent) and are used directly in the likelihood function, whereas series (d) and (e) are assessment model inputs.

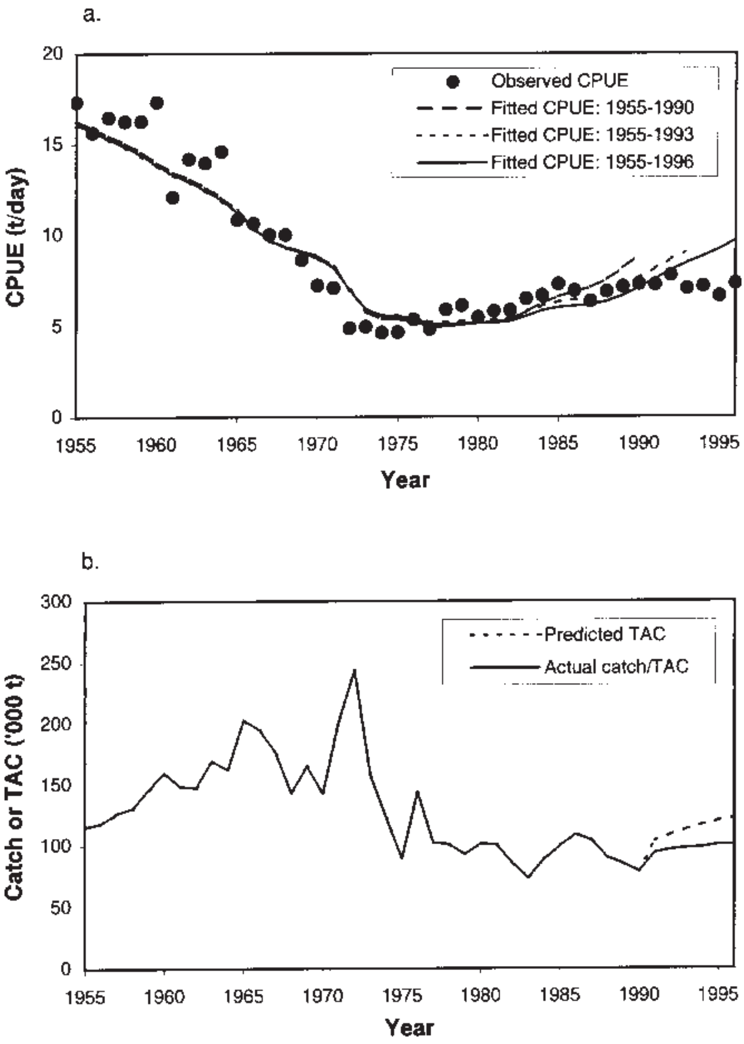


Figure 3. *Results of applying the hake estimator (Fig. 1) to the west coast component of the hake fishery. Figure 3a reflects the model fit of this estimator to observed CPUE data at three-year intervals from 1990. Figure 3b shows the observed catch history on the west coast up to 1990, the TAC values predicted by the model fit in 1990 using an* $f_{0.2}$ *harvesting strategy, and the TACs subsequently awarded as the model fit was updated from 1991 onwards. (Formally, a single TAC is awarded each year for the hake fishery, being the sum of TACs calculated by the management procedure for each of the west and south coast components. Advice is then that ⅔ of the total TAC should be taken on the west coast. Accordingly from 1991, Fig. 3b shows ⅔ of the total TAC awarded annually.)*

The MP for rock lobster comprises a relatively simple formula which uses three indices of resource status, namely the two CPUE series and the somatic growth data given in (a) and (d) respectively above (Sea Fisheries Research Institute 1997). The formula is such that an increase in any one of the indices tends to adjust the TAC upward, while a decrease has the opposite effect. In the process of making an appropriate choice of MP for this rock lobster fishery, 12 candidate MPs were presented to the Sea Fisheries Advisory Council (SFAC), a statutory body providing advice to the Government Minister responsible for fisheries management. Each of these had been tested for robustness across a wide range of uncertainties (as was done for sardine), and each was considered by scientists to have displayed an adequate level of resilience to these uncertainties in achieving management goals (Sea Fisheries Research Institute 1997). These management goals included maintaining a reasonable chance of achieving rebuilding targets and ensuring low inter-annual variations in TAC recommendations.

The 12 candidate management procedures presented to the SFAC are shown in Table 2. They differ in two respects, namely the level to which they are intended to rebuild the exploitable biomass (i.e., biomass of animals >75mm carapace length) over a period of 10 years, and the immediate effect on the TAC of applying the selected management procedure. Figure 4a illustrates catch trajectories (expectation in terms of the baseline assessment) associated with the four of these MPs which correspond to an immediate 20% increase in TAC for 1997 (compared to that for 1996), but differing recovery levels over a 10 year period, while Fig. 4b shows the corresponding biomass trajectories. The greater the extent of recovery desired, the lesser the increases in TAC over the period.

On being asked to choose one of the 12 MPs, SFAC members selected MP3 (Table 2). This choice amounted to selecting an MP which was the least conservative of the 12, providing the maximum immediate TAC gain and the lowest level of rebuilding. The probable reasons for this choice are discussed below.

1. Because of the changing political circumstances in South Africa, with redistribution of wealth and opportunity being a focal point in many spheres of public life, and stressed in the White Paper on marine fisheries policy (Department of Environmental Affairs and Tourism 1997), existing quota holders in this fishery are uneasy about their immediate and medium-term future, given much talk of “restructuring” (i.e. introduction of new entrants) to the fishery. These existing quota holders see their interests best served, given the present recovery trend in the resource following TAC cutbacks during the period of low somatic growth, by a rapid increase in the TAC giving them a better chance of securing larger rights in the long term; in contrast, they see the risk of the slow TAC increases associated with higher biomass recovery targets as likely to lead to all the consequent lesser TAC increases being allocated to new entrants. Furthermore, why sacrifice TAC increases

Table 2. The 12 candidate management procedures (MP1-MP12) for the west coast rock lobster resource presented to the SFAC for different "rebuilding strategy–immediate TAC increase" combinations.

10-year	Immediate TAC increase		
rebuilding strategy	3% increase	10% increase	20% increase
20% recovery	MP1	MP2	MP3
30% recovery	MP4	MP5	MP6
40% recovery	MP7	MP8	MP9
50% recovery	MP10	MP11	MP12

The "rebuilding strategy" options reflect the increase in biomass (of animals above 75 mm carapace length) at the end of a 10-year period when the corresponding MP is applied for the whole period. The "immediate TAC increase" options reflect the immediate increase in the TAC relative to the TAC in 1996 when the corresponding MP is applied in 1997. Catch and biomass trajectories for the MPs in the "20% increase" column are shown in Fig. 4.

now for greater levels of future biomass recovery when it will be others who reap the benefits of these sacrifices in the future.

2. The SFAC includes members of the fishing industry, though these persons are appointed in their personal rather than a representative capacity. Because all 12 candidate MPs were deemed "scientifically" acceptable in terms of risk to the resource, differing only in the target level of recovery and the amount of immediate TAC increase, non-industry members probably had no strong basis on which to object to any of the 12 MPs. Opinion within the SFAC was therefore likely heavily influenced by the members of the industry driven to some large extent by the considerations above.

Discussion

The above examples of the application of MPs in the South African context show that the MP approach provides a natural way to deal with uncertainties about model structure in a direct manner. This is achieved by considering the effect on key performance statistics of the MP of changes in model assumptions which reflect these uncertainties, and seeing which uncertainties have the greatest impact on these statistics. For example, in the case of sardine, robustness trials revealed that the form of the actual stock-recruit relationship is a far greater concern in terms of *risk* than the appropriate weighting of data used for estimation in the likelihood function. In a formal sense, evaluation of the results of a potentially large number of these different trials would entail taking weighted averages, where the weight corresponds to the relative likelihood of the scenario

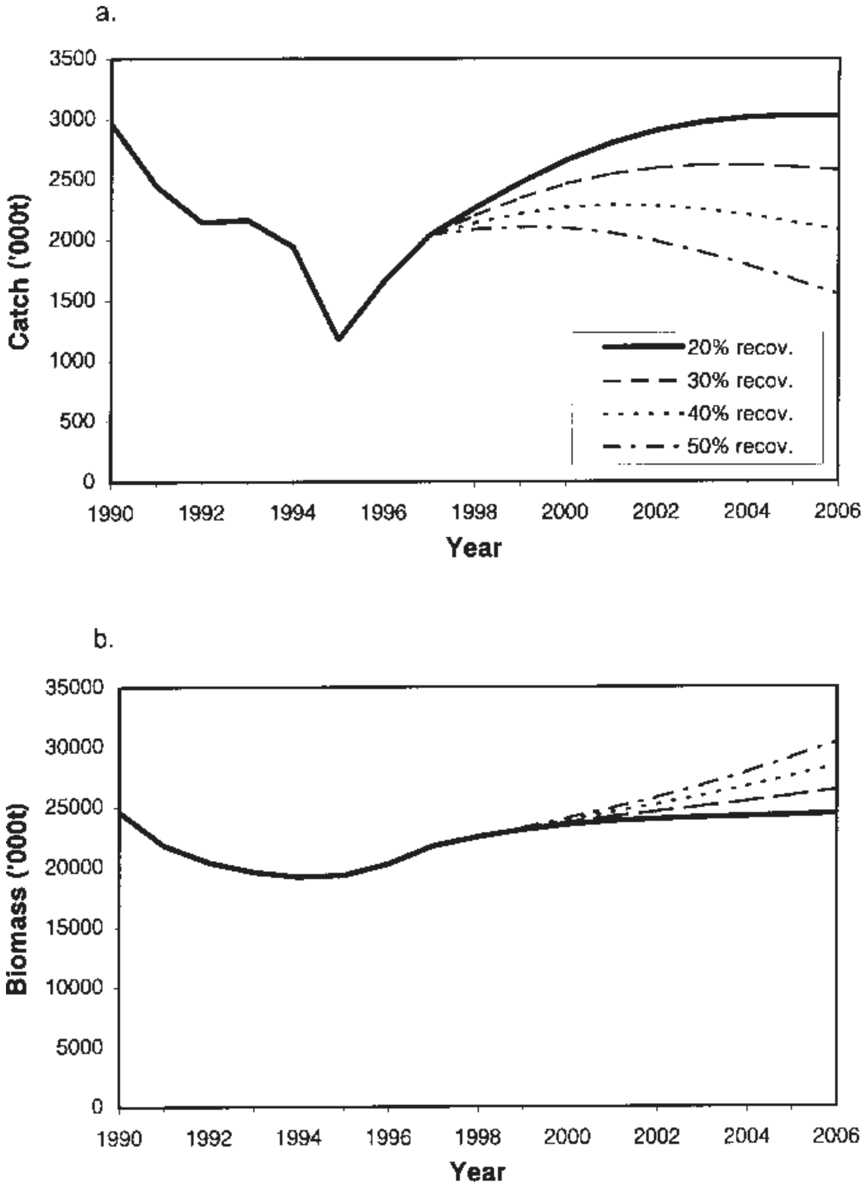


Figure 4. Predicted catch and biomass trajectories for the four candidate MPs for west coast rock lobster corresponding to an immediate 20% increase in TAC for 1997 compared to 1996 (Table 2), but for differing recovery levels of exploitable biomass (biomass of animals >75mm carapace length) over a 10-year period. Figure 4a shows the catch trajectories, while Fig. 4b shows the exploitable biomass trajectories.

reflected by the trial representing reality (Butterworth et al. 1996). However, experience has shown that, as in the sardine example, usually only one or two factors dominate in terms of their impact on MP performance statistics. Therefore, the problem of assigning appropriate weights effectively disappears as only a relatively small set of trials need be considered when comparing the anticipated performances of different candidate management procedures.

The hake MP is appealing in its ability to self-correct by directly updating parameter estimates as it encompasses new data through a process of feedback-control. This type of approach may be useful in situations where there is unease (not necessarily justifiable from a scientific point of view) about letting the MP operate "automatically" for a number of years without carrying out a new assessment each year. This type of unease was experienced with the MP implemented recently for west coast rock lobster, where considerable resistance from the industry was encountered. The difficulty of using estimators in the rock lobster (or sardine) MPs which are similar to the assessment procedures for those resources is that the latter involve complex age/size structured models, for which estimation is highly computer intensive. Full-scale simulation testing of MPs involving such assessments thus becomes problematic because of computing time requirements. Whether or not these assessment procedures could be adequately mimicked by simpler models in the estimator components of the MPs (as, for example, in the hake MP for which the estimator involves a relatively simple age-aggregated model) is a matter still to be investigated for the sardine and west coast rock lobster resources.

Short-term sociopolitical considerations carry considerable weight when industry and managers are faced with the selection of one of many candidate MPs in a situation of uncertain long-term security of quotas. This was probably the major concern that drew the SFAC to make the choice they did when selecting an MP for the west coast rock lobster. Cochrane et al. (1998) highlight establishment of long-term rights as one of three important criteria to be met if MPs are to be fully able to achieve objectives expected of them. Nevertheless, the approach used in the rock lobster example of presenting only those MPs that met more essential criteria, such as achieving at least "some" recovery while still being robust to key uncertainties, is important. Therefore, even though the SFAC opted for the least conservative MP for rock lobster, it remained acceptable from a conservation point of view.

At root, the MP approach to fisheries management is philosophically very different from the traditional approach in primary use throughout the world of the combination of an annual "best assessment" coupled with a regulatory mechanism (e.g., a catch control law) based upon some biological reference point (Butterworth et al. 1997, Cochrane et al. 1998). The scientific motivation in South Africa to move towards an MP approach to fisheries management arose from involvement in the process of development of the Revised Management Procedure (RMP) in the International

Whaling Commission's (IWC) Scientific Committee that took place over the 1987-1992 period (Kirkwood 1992, 1997; IWC 1994). The particular appeal of the approach was the manner in which it took uncertainties into account. For reasons essentially outside the scientific domain, the IWC has yet to apply the RMP in actuality. The South African experiences referenced above therefore provide some of the few available examples that address the question of whether or not this is indeed a viable basis to manage fisheries in practice.

Acknowledgments

Financial support from Sea Fisheries Research Institute, South Africa, to enable the first author to attend the Alaska Sea Grant College Program symposium and the comments of two anonymous reviewers and Dr. J.F. Schweigert are gratefully acknowledged. Dr. R.W. Leslie is also acknowledged for his advice and for providing the data for Fig. 3.

References

Butterworth, D.S., and M.O. Bergh. 1993. The development of a management procedure for the South African anchovy resource. In: S.J. Smith, J.J. Hunt, and D. Rivard (eds.), Risk evaluation and biological reference points for fisheries management. Can. Spec. Publ. Fish. Aquat. Sci. 120:83-99.

Butterworth, D.S., K.L. Cochrane, and J.A.A. De Oliveira. 1997. Management procedures: A better way to manage fisheries? The South African experience. In: E.K. Pikitch, D.D. Huppert, and M.P. Sissenwine (eds.), Global trends: Fisheries management. Amer. Fish. Soc. Symp. 20:83-90.

Butterworth, D.S., J.A.A. De Oliveira, and K.L. Cochrane. 1993. Current initiatives in refining the management procedure for the South African anchovy resource. In: G. Kruse, D.M. Eggers, R.J. Marasco, C. Pautzke, and T.J. Quinn II (eds.), Proceedings of the International Symposium on Management Strategies for Exploited Fish Populations. University of Alaska Sea Grant, AK-SG-93-02, Fairbanks, pp. 439-473.

Butterworth, D.S., A.E. Punt, and A.D.M. Smith. 1996. On plausible hypotheses and their weighting, with implications for selection between variants of the Revised Management Procedure. Rep. Int. Whal. Comm. 46:637-640.

Cochrane, K.L., D.S. Butterworth, J.A.A. De Oliveira, and B.A. Roel. 1998. Management procedures in a fishery based on highly variable stocks and with conflicting objectives: Experiences in the South African pelagic fishery. Rev. Fish Biol. Fish. 8:177-214.

Crawford, R.J.M., L.V. Shannon, and D.E. Pollock. 1987. The Benguela ecosystem. Part IV. The major fish and invertebrate resources. In: M. Barnes (ed.), Oceanography and marine biology: An annual review 25. University Press, Aberdeen, pp. 353-505.

De Oliveira, J.A.A. 1995. Assessment and management of the South African anchovy resource. M.Sc. thesis, University of Cape Town, South Africa. 185 pp.

De Oliveira, J.A.A., D.S. Butterworth, B.A. Roel, K.L. Cochrane, and J.P. Brown. 1998. The application of a management procedure to regulate the directed and bycatch fishery of the South African sardine *Sardinops sagax.* In: S.C. Pillar, C.L. Moloney, A.I.L. Payne, and F.A. Shillington (eds.), Benguela dynamics: Impact of variability on shelf-sea environments and their living resources. S. Afr. J. Mar. Sci. 19:449-469.

Department of Environmental Affairs and Tourism. 1997. A marine fisheries policy for South Africa. White paper, Department of Environmental Affairs and Tourism, Cape Town. 46 pp.

International Whaling Commission (IWC). 1994. Report of the Sub-Committee on Management Procedures. Rep. Int. Whal. Comm. 44:74-92.

Kirkwood, G.P. 1992. Background to the development of revised management procedures. Rep. Int. Whal. Comm. 42:236-243.

Kirkwood, G.P. 1997. The revised management procedure of the International Whaling Commission. In: E.K. Pikitch, D.D. Huppert, and M.P. Sissenwine (eds.), Global trends: Fisheries management. Amer. Fish. Soc. Symp. 20:91-99.

Payne, A.I.L., and A.E. Punt. 1995. Biology and fisheries of South African Cape hakes (*M. capensis* and *M. paradoxus*). In: J. Alheit and T.J. Pitcher (eds.), Hake: Biology, fisheries and markets. Chapman and Hall, London, pp. 15-47.

Punt, A.E. 1992. Management procedures of Cape hake and baleen whale resources. Rep. Benguela Ecol. Progm. S. Afr. 23. 689 pp.

Punt, A.E. 1993. The comparative performance of production-model and *ad hoc* tuned VPA based feedback-control management procedures for the stock of cape hake off the west coast of South Africa. In: S.J. Smith, J.J. Hunt, and D. Rivard (eds.), Risk evaluation and biological reference points for fisheries management. Can. Spec. Publ. Fish. Aquat. Sci. 120:283-299.

Sea Fisheries Research Institute. 1997. Recommendations to the Sea Fisheries Advisory Committee for management of the west coast rock lobster resource in 1997/1998 and the following two years. Unpublished report, Sea Fisheries Research Institute, South Africa. 12 pp.

A Length-Based Population Model for Hard-to-Age Invertebrate Populations

Terrance J. Quinn II
University of Alaska Fairbanks, Juneau Center, School of Fisheries and Ocean Sciences, Juneau, Alaska

Clive T. Turnbull
Northern Fisheries Centre, Cairns, Australia

Caihong Fu
University of Alaska Fairbanks, Juneau Center, School of Fisheries and Ocean Sciences, Juneau, Alaska

Abstract

Invertebrates such as shrimp and prawns are notoriously difficult to age, but much information about their length, growth, and harvest is routinely collected. We construct a length-based population model that utilizes and explains such harvest and survey information. The model contains population parameters for recruitment, growth, and mortality, and the data are used in conjunction with the model to estimate parameters. Our model is an extension of the model of Deriso and Parma (1988), which determines probability distributions of abundance and catch as a function of length. Our extension of their model is to "discretize" the length distribution, to allow more general selectivity and natural mortality representations, and to generalize the model for two sexes. These enhancements provide a more flexible approach to length-based modeling, although the solutions are recursion equations that are more difficult to compute. We illustrate application of the model to the Torres Strait prawn (*Metapenaeus* sp.) fishery in Australia. For this invertebrate population, dynamics occur rapidly over the course of a year, so the time step for the data and the model is monthly. Relative trends in estimated recruitment and abundance are similar among model configurations, which depend on the model parameters to be estimated, but absolute estimates differ substantially. While the model is able to fit the harvest and length frequency data, there is not enough information to jointly estimate all model parameters, especially catchability and natural mortality. However, estimates of growth and selectivity parameters are robust and differ from those obtained from tagging data collected from a different time period. Further resolution of model parameters should be possible with additional research survey information.

Introduction

Aging of invertebrates is notoriously difficult, because molting of the exoskeleton prevents the establishment of annuli. Thus, for invertebrate populations, length information is the primary data source for stock assessment, and a length-based population model is the primary analytical approach for estimating population parameters. A length-based model that utilizes harvest and survey information would be useful to integrate various sources of available information and to estimate population parameters such as recruitment, abundance, growth, and mortality. We propose such a model based on an extension of the length-based model of Deriso and Parma (1988).

There is a long history of development of length-based models, which are exceedingly important in tropical climes; a summary of length-based methods is contained in the volume edited by Pauly and Morgan (1987). Many of the methods are for analysis of growth information and variations of cohort analysis using length data. Parametric uncertainty induces inaccuracy in estimates of abundance from length cohort analysis, which weakens the utility of such procedures (Lai and Gallucci 1988). Several more flexible and more complicated length-based models have emerged in the late 1980s and 1990s (see the discussion section). Yet Hilborn and Walters (1992) pointed out: “Analysis of length data ... is a very fashionable topic at the moment, but we are very skeptical that much can be learned from length data even in principle .... We believe that not enough is known about the reliability of most of the new length-based methods ... at present.” Since then, more familiarity with these methods has resulted in wider use, but Hilborn and Walter’s cautions are still relevant.

One important length-based model (Deriso and Parma 1988, Parma and Deriso 1990) incorporates stochastic variation in growth. This stochastic population model uses a recurrence relationship (developed by Cohen and Fishman 1980) to generate growth using a von Bertalanffy or Gompertz growth curve and an underlying normal or lognormal distribution for statistical variability. Because growth is stochastic, probability distributions of length at age for abundance and catch are generated. By assuming a piece-wise quadratic polynomial function for selectivity, Deriso and Parma (1988) obtained closed-form solutions for these distributions. The model incorporates recruitment, growth, natural mortality, selectivity, and fishing mortality and includes stochastic elements for the mortality process and gear selectivity as a function of length.

Our extension of the length-based model of Deriso and Parma (1988) is designed to be more applicable for hard-to-age invertebrate and fish populations. This model characterizes the probability distributions of abundance and catch as a function of length in the presence of a general gear selectivity function. Our extension of their model is to "discretize" the length distribution, to allow more general selectivity and natural mortality representations, and to generalize the model for two sexes. These enhancements provide a more flexible approach to length-based modeling, although the solutions are recursion equations more difficult to compute. However, it has not been difficult to use a spreadsheet or computer program to perform the necessary calculations.

We illustrate application of the model to the Torres Strait prawn (*Metapenaeus* sp.) population in Australia. Its fishery is jointly managed by Australia and Papua New Guinea as required under the Torres Strait Treaty, and has been the focus of a major research program since 1985. The model is applied to length frequency data by sex from sampling commercial catches and total catch information from fishery logbooks.

Model

One-Sex Model

Our length-based model determines the length frequencies and corresponding probability density functions (PDFs) of the catch and the population for each cohort present. Starting with the PDF and total abundance of recruitment for a cohort, recursion equations are developed to obtain the PDF and total abundance of the cohort at each subsequent age during its lifetime. Following Deriso and Parma (1988), mortality is assumed to occur first, followed by growth.

Suppose that recruitment of N_r individuals occurs at age r, where age can have any time unit (e.g. daily, monthly, yearly). The length x of each individual is considered to be a discrete variable. Assume that the distribution of lengths x for age r is discrete normal ($\mathrm{N_D}$) with mean μ_r and variance σ_r^2; the probability density function (PDF) can be written

$$f_r(x) \sim \mathrm{N_D}(\mu_r, \sigma_r^2) = e^{-\frac{1}{2\sigma_r^2}(x-\mu_r)^2} / \xi_r, \quad \text{where} \quad \xi_r = \sum_x e^{-\frac{1}{2\sigma_r^2}(x-\mu_r)^2}. \tag{1}$$

The constant ξ_r is a normalizing constant so that the sum of equation (1) over x equals 1. The abundance of age r individuals at each length is then

$$N_r(x) = N_r f_r(x). \tag{2}$$

To obtain abundance $N_{a+1}(x)$ at age $a+1$ from abundance $N_a(x)$ at age a (starting at age r), the processes of growth and mortality must be considered. Natural mortality and gear selectivity are assumed to be functions of length, not age. Let

x, L = subscripts for length,

M_x = natural mortality,

S_x = gear and/or survey selectivity,

F_a = full-recruitment fishing mortality, and

$F_{a,x}$ = fishing mortality for partially selected lengths.

Fishing mortality is assumed separable into selectivity and full-recruitment fishing mortality:

$$F_{a,x} = S_x F_a. \tag{3}$$

Total mortality is the sum of natural and fishing mortality:

$$Z_{a,x} = M_x + F_{a,x}. \tag{4}$$

Model parsimony can be enhanced by specifying functional forms for some parameters. For example, natural mortality can be parameterized as a linear function $M_x = M_a + M_b x$ or an exponential function $M_x = M_a \exp(-M_b x)$, where M_a and M_b are parameters. Zheng et al. (1995) give a U-shaped mortality function, for species in which mortality increases at older ages. Similarly, selectivity S_x can be parameterized as the logistic function

$$S_x = 1\Big/\left(1 + e^{-\gamma(x - L_{50\%})}\right),$$

where $L_{50\%}$ is the size where 50% of the fish are vulnerable and γ is a shape parameter for the logistic function. Finally, if fishing effort data $\{E_a\}$ are available, then full-recruitment fishing mortality may be approximately related to fishing effort as $F_a = qE_a$ where catchability q is assumed constant. Fishing effort and mortality are written as functions of age that are equivalent to functions of time, because as a cohort ages, time elapses as well.

For a given PDF $f_a(x)$ at the start of age a, the relative distribution of lengths in the population after mortality occurs (at the end of age a) is

$$p_{a,Z}(x) = f_a(x)e^{-(M_x + S_x F_a)} = f_a(x)e^{-Z_{a,x}}. \tag{5}$$

The absolute number of individuals at length x after mortality is given by

$$N_{a,Z}(x) = N_a p_{a,Z}(x), \tag{6}$$

and the PDF for lengths after mortality is then

$$f_{a,Z}(x) = p_{a,Z}(x) / \sum_x p_{a,Z}(x). \tag{7}$$

From Deriso and Parma (1988), the relative distribution of lengths in the catch follows from the Baranov catch equation $C = N(F/Z)[1 - \exp(-Z)]$ and is given by

$$p_{a,C}(x) = f_a(x) \frac{F_{a,x}}{Z_{a,x}} (1 - e^{-Z_{a,x}}). \tag{8}$$

Absolute catch at age a and its PDF follow from equations (6) and (7) by replacing Z with C, which yields

$$C_a(x) = N_a p_{a,C}(x) \tag{9}$$

and

$$f_{a,C}(x) = p_{a,C}(x) / \sum_x p_{a,C}(x). \tag{10}$$

Total catch at age a from equation (9) is $C_a = \sum_x C_a(x)$.

If a survey is conducted with gear similar to the fishery, then the resulting length distribution will be influenced by gear selectivity. In essence, it is the surveyable or exploitable population that is measured. Even if a survey is not conducted, the exploitable population is often a primary quantity of interest. The mathematical definition of the surveyable or exploitable population for length x at age a is

$$EN_a(x) = S_x \frac{1 - e^{-Z_{a,x}}}{Z_{a,x}} N_a(x) \approx S_x N_a(x),$$

where the approximation holds for small Z. Thus, the relative distribution of lengths in the exploitable population for age a (after mortality occurs) is given by

$$p_{a,S}(x) = S_x p_{a,Z}(x). \tag{8b}$$

Exploitable abundance at age a and its PDF follow from equations (6) and (7) by replacing subscript Z with S, which yields

$$EN_a(x) = N_a p_{a,S}(x) \tag{9b}$$

and

$$f_{a,S}(x) = p_{a,S}(x) / \sum_x p_{a,S}(x). \tag{10b}$$

Total exploitable abundance at age a from equation (9b) is $EN_a = \sum_x EN_a(x)$. If the survey has a different selectivity pattern than the fishery, then a separate set of selectivity parameters can be used.

To account for growth, it is assumed that an individual of length x will grow to length L in one time step according to a stochastic growth model. One useful model is the von Bertalanffy (LVB) model with stochastic error, which was first derived in Cohen and Fishman (1980) and utilized in the length-based model of Deriso and Parma (1988). The deterministic LVB model is written

$$L_a = L_\infty(1 - e^{-\kappa(a-a_0)}),$$

where L_∞ is asymptotic length, κ is a growth parameter, and a_0 is the age corresponding to length 0. An equivalent formulation for size L_{a+1} at age $a+1$ as a function of previous size L_a with the inclusion of a stochastic term is

$$L_{a+1} = L_\infty(1-\rho) + \rho L_a + \varepsilon_a, \tag{11}$$

where the Brody coefficient ρ is $\exp(-\kappa)$ and ε_a is an independent, normally distributed random variable with mean zero and variance σ^2. From Cohen and Fishman (1980), the expected length and variance at age $a+1$ for an individual of length x at age a, considering that it was recruited at age r, are

$$\mu_{a+1}(x) = L_\infty(1-\rho) + \rho x \tag{12a}$$

and

$$\sigma_{a+1}^2 = \sigma^2 \frac{1-\rho^{2(a+1-r)}}{1-\rho^2} + \rho^{2(a+1-r)}\sigma_r^2. \tag{12b}$$

Alternative models include a gamma distribution for variance (Sullivan 1992, Zheng et al. 1995), more general growth models (e.g., Gompertz, Schnute), and letting σ_{a+1}^2 be a function of length x as well as age; we use the stochastic LVB model for simplicity. For the LVB model, mean length at age $a+1$ for an individual at length x is not a function of age a, but for other models, it would be, which is why we write the subscript in the equation. Equation (12b) may not be exactly correct due to selectivity effects, in which case different variance parameters $\{\sigma_a\}$ for individual ages could be used in place of relationship (12b), in which variance is solely as function of parameters σ_r, σ, and ρ.

The normal PDF for the length distribution after one growth increment for an individual originally of size x is given by

$$f_{a+1,G}(L \mid x) \sim \mathrm{N_D}(\mu_{a+1}(x), \sigma^2_{a+1}) = e^{-\frac{1}{2\sigma^2_{a+1}}[L-\mu_{a+1}(x)]^2} \Big/ \xi_{a+1,x}\,, \tag{13}$$

where

$$\xi_{a+1,x} = \sum_L e^{-\frac{1}{2\sigma^2_{a+1}}[L-\mu_{a+1}(x)]^2}\,.$$

The relative distribution of lengths at the start of age a+1 is then obtained from the relative distribution of lengths after mortality equation (5) and the PDF for growth equation (13), which results in

$$\begin{aligned} p_{a+1}(L) &= \sum_x f_{a+1,G}(L \mid x) p_{a,Z}(x) \\ &= \sum_x f_a(x) e^{-Z_{a,x}} e^{-\frac{1}{2\sigma^2_{a+1}}[L-\mu_{a+1}(x)]^2} \Big/ \xi_{a+1,x}\,. \end{aligned} \tag{14}$$

The absolute numbers at length at age a+1 and corresponding PDF are then

$$N_{a+1}(L) = N_a\, p_{a+1}(L) \tag{15}$$

and

$$f_{a+1}(L) = p_{a+1}(L) \,/ \sum_L p_{a+1}(L)\,. \tag{16}$$

The total number of individuals at age a+1 from equation (15) is then

$$N_{a+1} = \sum_L N_{a+1}(L). \tag{17}$$

Conversion of abundance (in numbers) to biomass (in weight) for either the population or the catch is easily accomplished by assuming an allometric weight-length relationship

$$W(L) = \alpha L^{\beta}, \tag{18}$$

and then multiplying length-specific values by the corresponding weight. Thus biomass (abundance in weight) and yield (catch in weight) are found by multiplying length-specific abundance and catch, respectively, by weight at that length and summing over length, or

$$B_a = \sum_L N_a(L) W(L) \quad \text{and} \quad Y_a = \sum_L C_a(L) W(L). \tag{19}$$

This formulation covers the progress of a single year class through its lifespan. Identical equations can be constructed for all year classes and referenced to the passage of time. Thus, the length distributions of the population at a given time from this model can be assembled by summing over all age classes present.

Two-Sex Model

This model can be easily generalized to account for sex. We denote sex by subscript y where $y \in \{m,f\}$ (male, female). All previous parameters, variables, and PDFs that are to be a function of sex are then written with subscript y, such as $s_{x,y}$, $F_{a,y}$, ρ_y, $f_{a,y}(x)$, and $N_{a+1,y}(L)$. The first step in constructing a two-sex model is to first consider each sex at recruitment. If χ_y is the proportion of sex y at recruitment (such that $\chi_m + \chi_f = 1$), then the initial abundance for each sex is given by

$$N_{r,y} = N_r \chi_y. \tag{20}$$

An added complexity would be to make the abundance of each sex a random variable using, say, the binomial distribution, but for large populations only a miniscule amount of variation would be added.

The dynamics of each sex are then followed separately according to the one-sex model with $N_{r,y}$ in place of N_r, resulting in separate relative length distributions $p_{a+1,y}(L)$ from equation (14), abundance distributions $N_{a+1,y}(L)$ from equation (15), and PDFs $f_{a+1,y}(L)$ from equation (16). Analogous distributions for catch and exploitable abundance follow as above. Finally, corresponding distributions combined over sex are found by summing abundance over sex and then converting to relative distributions.

Estimation

The appropriate objective function to construct for fitting data to the length-based model depends on the type and amount of data available. Either sums of squares or likelihood objective functions can be constructed to fit observed data, as explained in Deriso et al. (1985), Deriso and Parma (1988), Methot (1990), Sullivan et al. (1990), and Zheng et al. (1995). Let subscripts L, y, and t denote length, sex, and time, respectively. For use in the illustration below, we construct an objective function for the situation where length frequency data by sex $\{f'_{L,y,t}\}$ and total yields $\left\{Y'_t\right\}$ (catch in biomass) for combined sexes are available. Their counterparts without primes denote equivalent quantities obtained from the length-based model. We estimate the length frequencies $\{f_{L,y,t}\}$ from the model by multiplying the monthly sample size by the proportion of prawns of a particular sex and length in the model catch.

By assuming an underlying root normal distribution (Quinn 1985), a possible objective function is a combination of residual sums of squares, given by

$$\mathrm{RSS} = \sum_t \sum_y \sum_L \left(\sqrt{f'_{L,y,t}} - \sqrt{f_{L,y,t}}\right)^2 + \lambda \sum_t \left(\sqrt{Y'_t} - \sqrt{Y_t}\right)^2, \tag{21}$$

where λ is a weighting term governing how close the yield data and model values match up. The square root transformation in equation (21) is not as aggressive as the logarithmic transformation, handles 0 values readily, and in our experience provides more reliable convergence and stable parameter estimates than no transformation.

Model parameters that might be estimated include year-class strength parameters $\{N_r\}$, fishing mortality parameters $\{F_t\}$, selectivity parameters which permit calculation of $\{S_x\}$, growth parameters $\{L_\infty, \rho, \sigma_r^2, \sigma^2, \alpha$, and $\beta\}$, natural mortality parameters which permit calculation of $\{M_x\}$, and auxiliary catchability parameter(s) if fishing effort data are used (as in $F_t = qE_t$). It is usually desirable to specify as many parameters as possible from auxiliary information, because many combinations of parameters can explain the observed data equally well. In our experience, estimating parameters on a square root or log scale is advantageous, in that resulting transformed estimates are always positive and stable.

Illustration—Torres Strait Prawns

Brief Description

The Torres Strait prawn fishery is estimated to be worth AU$18-23 million annually and is jointly managed by Australia and Papua New Guinea (PNG) as required under the Torres Strait Treaty. The fishery operates in the eastern section of the Torres Strait Protected Zone (Fig. 1) and is managed by both seasonal and spatial closures. Although the brown tiger prawn *(Penaeus esculentus)* is the primary target species, the blue endeavour prawn *(Metapenaeus endeavouri)* examined in this study composes approximately 60% of the catch. These species are fast-growing and short-lived.

The fishing fleet is composed of about 90 Australian prawn trawlers that operate in Australian waters within Torres Strait and on the Queensland East Coast. Several PNG licensed trawlers that mainly fish the Gulf of Papua also fish PNG waters of Torres Strait in the later part of the season. The fleet is highly mobile and most vessels fish Torres Strait for only part of the fishing season. Fishing licenses are transferable and have an allocation of fishing access days attached to them. The fleet has rapidly adopted new electronic navigational systems (GPS), which may have increased the effective effort of the fleet in recent years.

Although the commercial prawn species in Torres Strait exhibit a typical penaeid life history, the seagrass meadows utilized by the juvenile phase occur on coral reef platforms (Turnbull and Mellors 1990). In other penaeid fisheries these seagrass meadows occur along sheltered coastlines. Juvenile prawns grow and migrate from seagrass habitat into shallow waters that are mainly located to the west of the Warrior Reefs (Fig. 2).

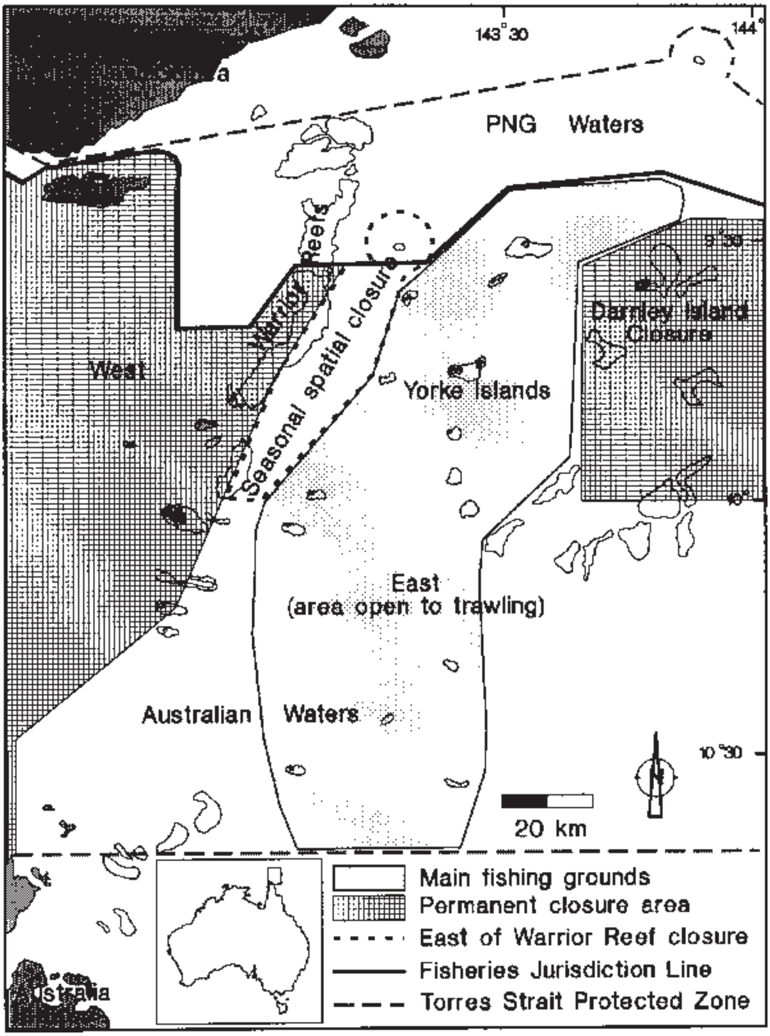


Figure 1. The location of the prawn fishery within the Torres Strait Protected Zone.

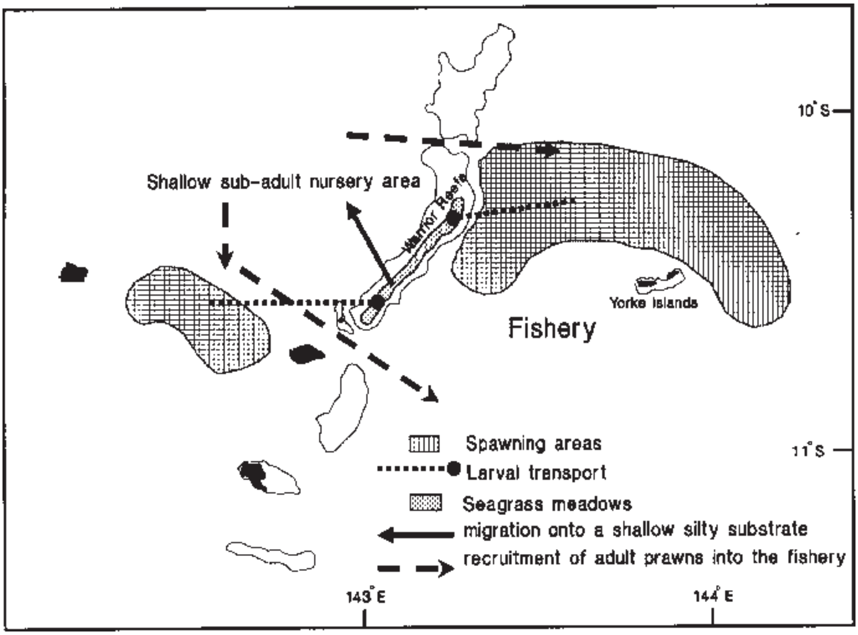


Figure 2. Life history of Torres Strait prawns.

At 4 to 6 months of age they move into deeper waters (18 to 50 m) that are largely found east of the Warrior Reefs. These species reach maturity at about six months of age and have a high fecundity, which generally increases with age. Although prawns migrate only short distances (an average of 30 to 40 km) their movement is significant to management strategies. Tagging studies (Watson and Turnbull 1993) and analysis of length frequency data indicate that prawns in the Torres Strait fishery generally migrate from around the Warrior Reef Complex to the east and southeast.

Observed Data

The model covers the period between March 1992 to November 1994; no fishery data are available from December to February due to a seasonal closure of the fishery. We selected this period because it was the most recent period for which length frequency data were available (see below). A maximum age of 18 months is used in the model, because catch of animals older than that comprise less than 1% of the highest monthly catch from a particular cohort. We use data from the fishery in 1993-1994 and also ancillary fishing effort and research survey information from other periods to develop and test the model. Since 1989 the entire fleet has been required to complete daily catch and effort returns (logbooks);

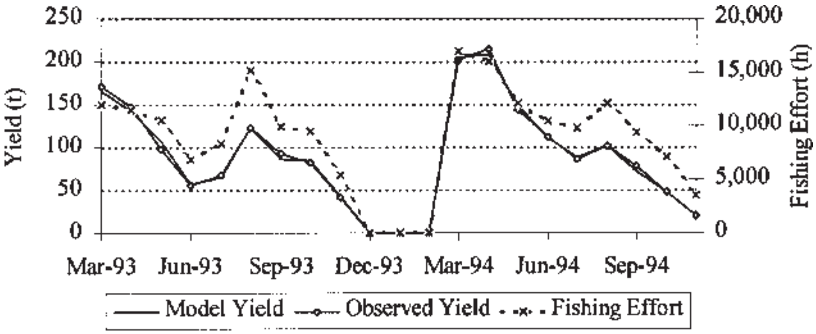


Figure 3. Observed yield and effort data, and yield estimated from case 3.

summary statistics of yield and effort are shown in Fig. 3 with corresponding values of catch-per-unit-effort (CPUE) shown in Fig. 4. Effort data from 1992 are used to determine fishing mortality in 1992, so that estimates of recruitment can be made for year-classes originating in 1992.

During the late 1980s and early 1990s the Queensland Department of Primary Industries investigated the basic life history of the tiger and endeavour prawn species in Torres Strait using monthly research trawl surveys and prawn tagging. This information provides initial estimates of parameters for growth, selectivity, and the length distribution at recruitment, and an understanding of the life history of endeavour prawns against which the model results can be verified. In 1991, an area-swept estimate of exploitable biomass from the surveys of about 800 t was obtained (Turnbull and Watson 1995). This value is useful for comparing results of different model configurations, because annual CPUE was nearly constant from 1991-1994 (Turnbull, unpubl. data).

A commercial catch sampling program was conducted bimonthly during 1993-1994, starting in March. A researcher on board commercial trawlers measured length and sex of prawns in subsamples of the catch, which provided estimates of the commercial catch size distributions for females (Fig. 5) and males (Fig. 6). There were 464 length frequencies (29 length intervals between 13 and 41 mm × 8 months sampled × 2 sexes) and 18 non-zero observations (9 months that the fishery is open per year × 2 years) of total catch for a total of 482 observations.

Initial Parameters

The fixed parameters given in Table 1 are obtained from previous research to reduce the number of parameters estimated. The length-weight parameters come from samples collected during the research trawl program. The recruit parameters (μ_r, σ_r) in Table 1 come from examination of beam trawl samples on the seagrass meadows, based on the size at which post-

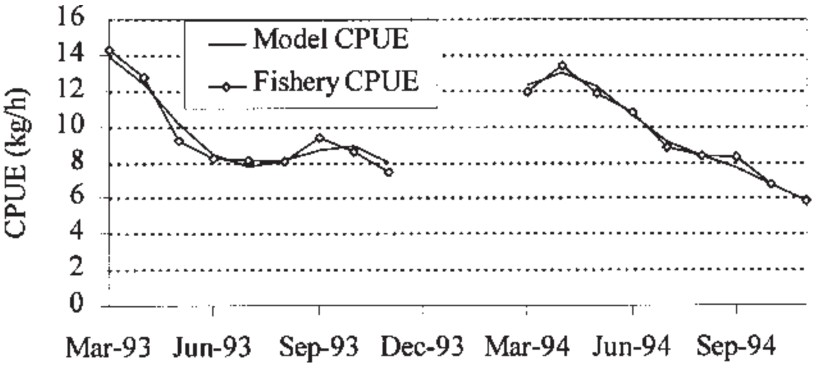


Figure 4. Observed and estimated catch per unit effort from case 3.

larvae settle at about $r = 1$ month. The sex ratios of prawns from commercial catch sampling data and from research surveys conducted during the late 1980s and early 1990s were near 50/50 pooled over all size classes, so it is assumed that the sex ratio at recruitment (χ_y) must be near 50/50.

Initial values for growth parameters (L_∞, κ, σ) by sex given in Table 2 are obtained by fitting the Cohen-Fishman growth model to research tagging data. Commercial trawl selectivity parameters (γ, $L_{50\%}$) in Table 2 come from a comparison of the catches of fine mesh and commercial mesh trawls. Initial values for parameters common to both sexes are given in Table 3. The natural mortality (M_a, M_b) and catchability (q) parameters are based on values in the literature for other penaeid species and an assumption that mortality decreases with age. Initial model runs were made with a constant (flat) recruitment pattern.

Fitting Procedure

Estimates of model parameters were initially obtained by minimizing the objective function (equation 21) using an Excel spreadsheet, but more stable and quicker results were obtained with a Fortran program using a Gauss-Marquardt search algorithm (either is available from the authors). The Fortran program made it easy to obtain standard errors using the Hessian.

The model is simultaneously fitted to yield data combined over sex and to size data classified by sex, corresponding to how the data are collected. Various starting values based on ancillary data as explained above were used until a stable fit was obtained to both data sets. Recruitment parameters are estimated for each month starting in November 1992 and ending in November 1994 when the fishery closed and no more data were available for that year. Recruitments in March-October 1992 are assumed proportional to those in 1993, with proportionality constant ψ, resulting

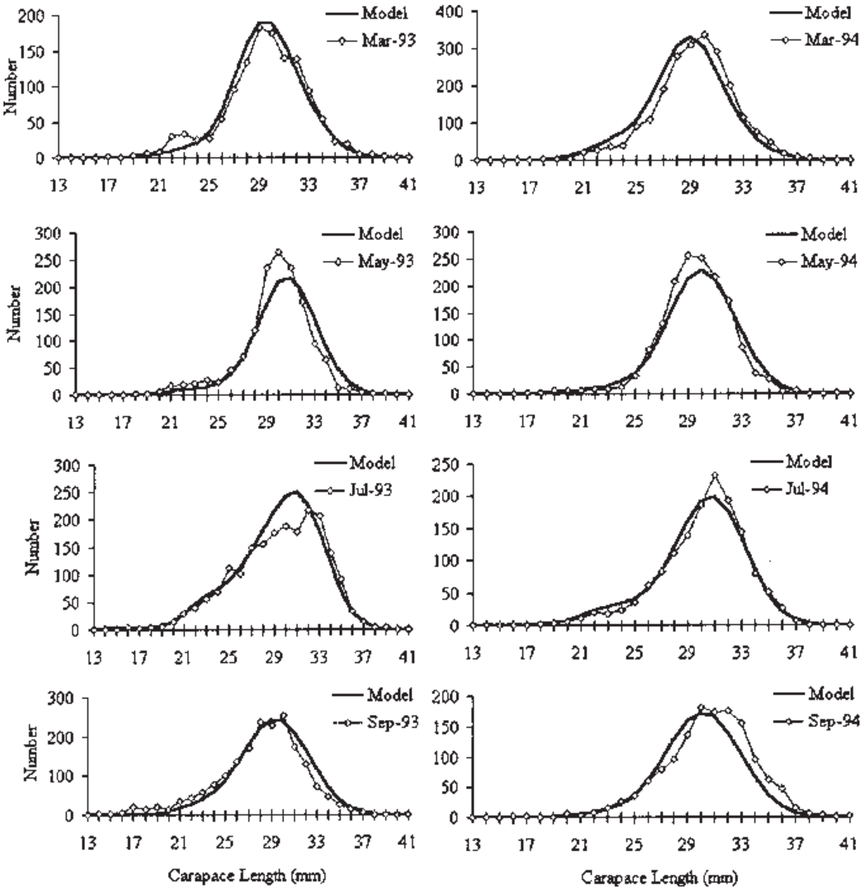


Figure 5. Female size data and corresponding estimates from the model (case 3).

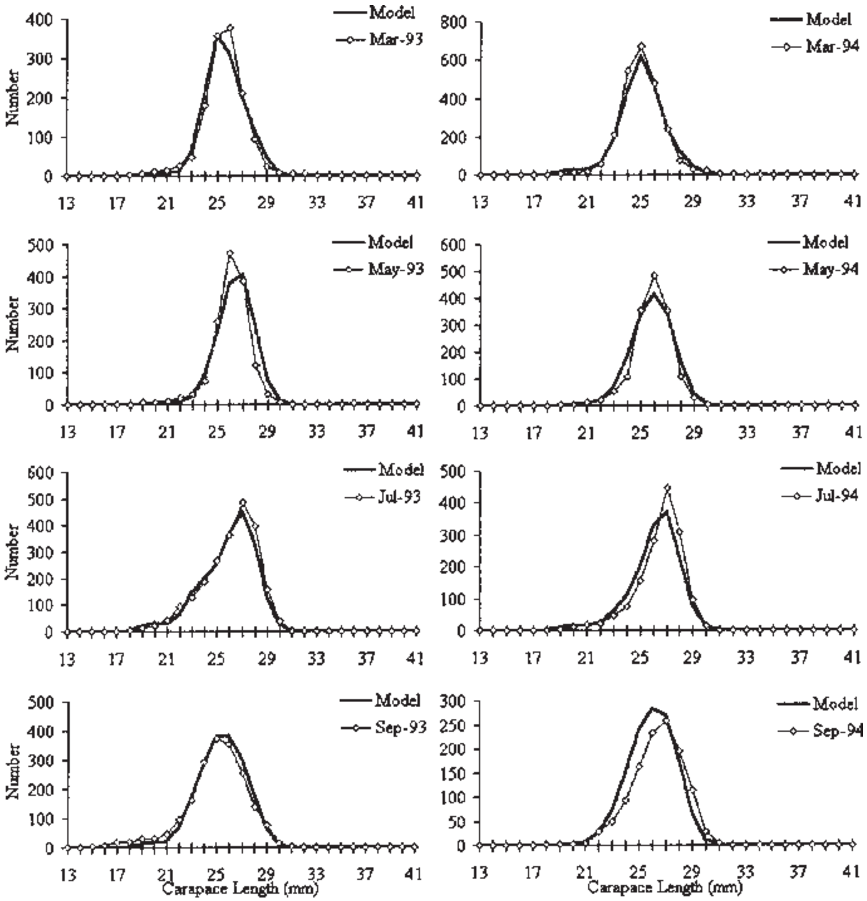


Figure 6. Male size data and corresponding estimates from the model (case 3).

Table 1. **Fixed parameters in the length-based model.**

Parameter	Female	Male
Length-weight parameter α (g)	0.0015	0.0017
Length-weight parameter β	2.81	2.79
Mean length of recruits μ_r (mm)	2	2
Sigma of recruits σ_r (mm)	0.5	0.5
Sex ratio χ_y	0.5	0.5

Table 2. **Estimated growth (L_∞, κ, σ) and selectivity ($L_{50\%}$, γ) parameters by sex and their coefficients of variation (CV).**

	Females				Males			
	Initial	Case 1	Case 2	Case 3	Initial	Case 1	Case 2	Case 3
L_∞ (mm)	36.87	31.64	31.58	31.93	27.75	27.43	27.35	27.53
CV		1%	1%	1%		1%	<1%	1%
κ (mo)$^{-1}$	0.30	0.52	0.53	0.52	0.39	0.58	0.58	0.56
CV		4%	4%	4%		2%	2%	2%
σ (mm)	0.98	1.57	1.69	1.54	1.08	0.82	0.81	0.78
CV		7%	6%	7%		7%	6%	8%
$L_{50\%}$ (mm)	19.9	27.23	26.37	27.49	19.9	23.43	22.88	23.95
CV		1%	2%	2%		1%	1%	1%
γ (mm)$^{-1}$	0.4	0.44	0.46	0.46	0.4	0.88	0.96	0.80
CV		7%	8%	6%		7%	9%	6%

"Initial" indicates initial values (see text). Case 1 fixes natural mortality and catchability parameters to initial values shown in Table 3. Case 2 fixes M_a at 0.33 (mo)$^{-1}$ and M_b at 0.02 (see Table 3). Case 3 fixes catchability q at 1.0×10^{-5} (h/mo)$^{-1}$ and natural mortality parameter M_b at 0 (see Table 3). Remaining parameters are estimated in cases 1 to 3.

Table 3. Estimated parameters common to both sexes and their coefficients of variation (CV).

Parameter	Initial	Case 1	Case 2	Case 3
Natural Mortality M_a (mo)$^{-1}$	0.42	0.42*	0.33*	0.36
CV				9%
Natural Mortality M_b (mm)$^{-1}$	0.03	0.03*	0.02*	0.00*
Catchability q (h/mo)$^{-1}$	2.5×10^{-5}	2.5×10^{-5}*	1.7×10^{-5}	1.0×10^{-5}*
CV			19%	
Recruitment ψ	1.00	1.49	1.77	2.00
CV		12%	12%	15%

An asterisk denotes a fixed value. See Table 2 for definition of cases.

in 26 recruitment parameters to be estimated. In addition there are 6 growth (3 per sex), 4 selectivity (2 per sex), 2 natural mortality (M_a and M_b) and 1 catchability (q) parameters, for a total of 39 parameters to be estimated.

We conducted a study of the model's sensitivity to various choices of the model parameters to be estimated. The most optimistic situation is that all parameters can be estimated precisely. However, it might be possible that the model can fit the data well but that there is not sufficient information to estimate all parameters, such as catchability and natural mortality. We estimate parameters for several alternative model configurations, in which one or more of the model parameters (selected from catchability, natural mortality, growth) is fixed at its initial value or a value suggested by an alternate model configuration. Results are compared in terms of their fit to the data and the plausibility of the estimates of exploitable biomass compared to the 1991 estimated exploitable biomass of about 800 t.

Results

The single sex model fitted the size data poorly compared to the two-sex model, because the two-sex model allows for differences between the sexes in the length-age and weight-length parameters. Consequently, further analysis involved only the two-sex model.

Several different model configurations were investigated for the two-sex model. The most important result is that widely different estimates of catchability, mortality, and abundance occurred depending on the set of fixed parameters and their specific values. The best fit to the data was obtained by letting all parameters be estimated, but the model went through several hundred iterations, which indicates difficulty in converging. Estimates of exploitable biomass were implausible by the earlier mentioned criterion of being near 800 t. Recruitment and natural mortality tended to

Table 4. Values of the objective function and its components for the three cases.

Objective function	Case 1	Case 2	Case 3
Total RSS	684	665	606
λ Catch RSS	86	91	78
Catch RSS	865	911	785
Size RSS	598	574	527
Female Size RSS	276	267	257
Male Size RSS	322	307	270
λ	0.1	0.1	0.1

See Table 2 for definition of cases.

increase as catchability tended to decrease. Even fixing one of the three parameters $\{q, M_a, M_b\}$ did not result in stable convergence or plausible biomass estimates.

Convergence was usually obtained for model configurations in which 2 or 3 of $\{q, M_a, M_b\}$ were fixed. We present three cases that illustrate the range of results obtained. Case 1 fixes natural mortality and catchability parameters to initial values shown in Table 3. Case 2 fixes M_a at 0.33 $(\text{month})^{-1}$ and M_b at 0.02 (see Table 3) corresponding to an intermediate set of natural mortalities that provided plausible results. Case 3 fixes catchability q at 1.0×10^{-5} and natural mortality parameter M_b at 0 (see Table 3). The choices for case 3 were made based on the tendency of different model runs to produce an estimate of q near that value (or even less) and the tendency of M_b to go to zero when estimated (but not always).

The objective function values decrease from case 1 to case 2 to case 3, showing that case 3 fitted the data best (Table 4). The improvements in fit occurred uniformly among the data components for total catch, female length frequency, and male length frequency (Table 4). We illustrate the model fit using case 3; the graphs for the other two cases are similar but have greater deviations. Model yield fitted the observed yield well (Fig. 3) and both reflect the level of fishing effort. (The increase in effort and catches in August was due to opening the spatial closure area to the east of the Warrior Reefs.) Consequently, model CPUEs (Fig. 4) also closely fitted the observed CPUEs and tended to increase around September-October indicating a second wave of recruitment to the fishery. This trend in CPUE, however, also occurred in the years before the closure existed and does not occur in tiger prawns, so it is not just a result of opening the spatial closure area.

The fits of case 3 to the length frequency data for females (Fig. 5) and males (Fig. 6) are generally good. Some data conflicts do exist between the yield and size data. Using $\lambda = 0.5$ improves the fit to the yield data while degrading the fit to the size data. The choice of $\lambda = 0.1$ was made to balance residual variability in the sources of information.

Absolute estimates of total and exploitable biomass (Fig. 7) vary greatly among the three cases, but relative trends are the same. Exploitable biomass has the same trend as CPUE (Fig. 4) as expected. Even more extreme results than these were obtained, suggesting that determination of absolute abundance is not possible given these data on total yield and length frequency. Estimates of recruitment parameters for March 1992 to August 1994 (Fig. 8) also show large absolute differences but small relative differences among the cases. The first 8 recruitments in the series are forced to have the same trend as the following year, but the other recruitments have their own parameters. There appears to be an annual pattern in recruitment with a large peak in November and a smaller peak about May. The November 1993 recruitment peak produces a peak in biomass (Fig. 7) in March 1994 and in exploitable biomass in April 1994. Similarly the smaller March-July 1993 recruitments produce a slight peak in biomass during July-September 1993 and in exploitable biomass during August-November 1993. These results are in accord with auxiliary information that percentage of female prawns with ovaries at the last stage before spawning peaks in August-September and to a less extent in January-March (Turnbull, unpubl. information). Settlement of post-larvae onto seagrass beds, which is represented as recruitment in the model, occurs 2-3 weeks after spawning.

There were some major differences between initial parameters and final parameter estimates from the model. Female L_∞ decreased while the male L_∞ stayed the same. Female and male κ increased to near the same value, whereas female σ increased while male σ decreased (Table 2). Since initial growth parameters were obtained from tagging data collected at another time, it is not surprising that estimates from the model would differ from the initial parameters. The higher L_∞ and σ for females combined with a similar κ compared to males result in females growing faster than males on average and with greater variability. This phenomenon explains why the length frequencies of females in the commercial catch (Fig. 5) are farther to the right and have a broader distribution compared to males (Fig. 6).

Estimates of selectivity parameters also differed from initial values that were based on a comparison of fine and commercial mesh net samples collected during research surveys (Table 2). The fitted $L_{50\%}$ values agree with the mean sizes for each sex of prawns moving into fishable areas (Turnbull, unpubl. survey data), and estimating selectivity parameters separately for each sex considerably improved the fit to the size data. It is interesting that selectivity parameters differ by sex, even though selectivity is a function of length. This result suggests that selectivity repre-

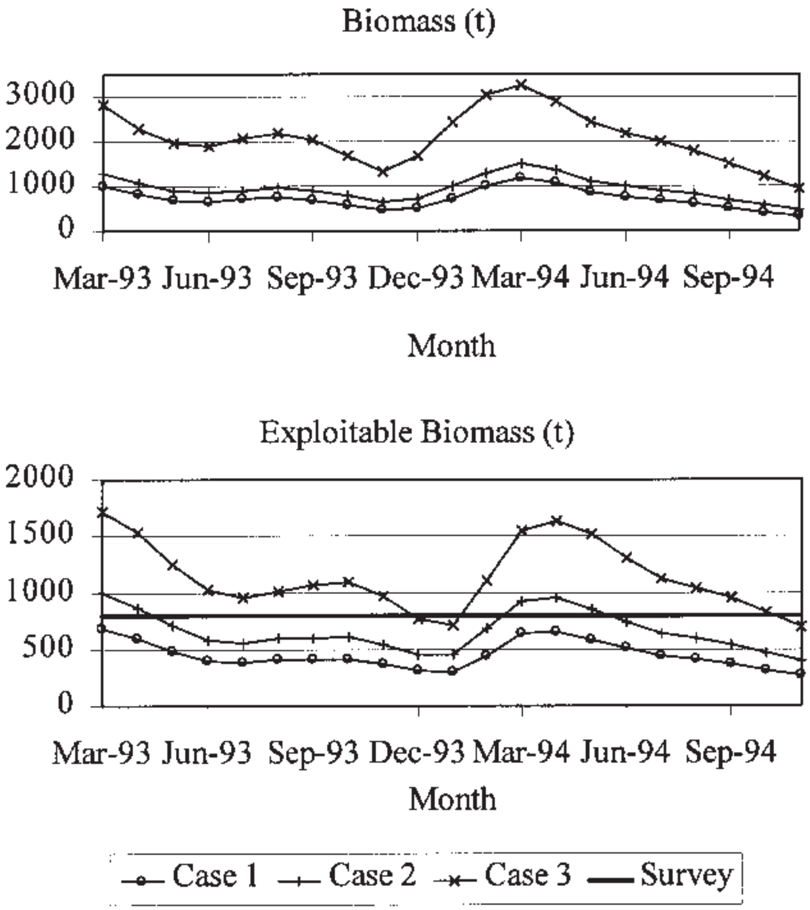


Figure 7. Estimated absolute and exploitable biomass from the model. The area-swept exploitable biomass estimate from the 1991 survey is also shown.

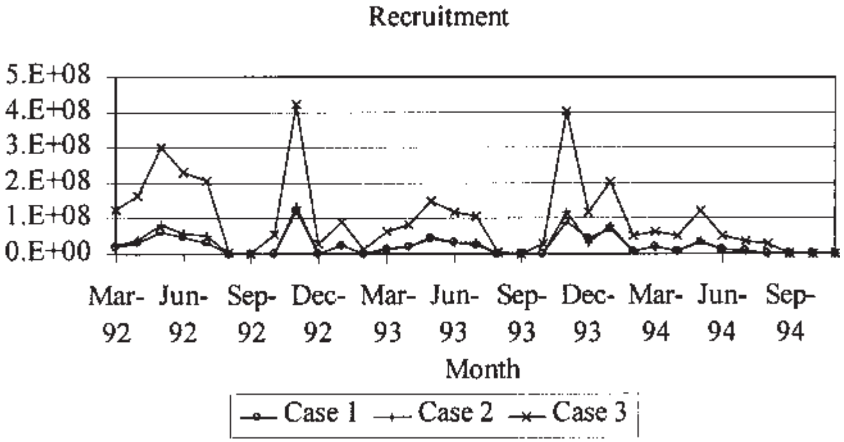


Figure 8. Estimated recruitment parameters from the model.

sents more than just the properties of the gear harvesting the prawns. Because smaller prawns are not necessarily available to the fleet due to the spatial closure, selectivity in the model serves as an availability function as well as representing gear selectivity. Further support comes from the fact that the ratio $L_{50\%}/L_\infty$ is similar between sexes, so that the faster growth of females to a larger size is also accompanied by a shift to larger lengths in selectivity. The average age at which 50% of each sex are selected can be resolved from the selectivity and growth parameters, which results in an average age of about 4.5 months, or about 3.5 months after post-larval settlement (or recruitment to the population) for both sexes. This result is consistent with the hypothesis that females and males move from the closure areas to the fishable areas at about the same time. The ratio of female to male totals in the length frequency samples is about 0.97, suggesting that females and males experience similar mortality throughout their lives despite differences in size. This is because fishing mortality is low compared to natural mortality; estimated full-recruitment fishing mortality is at most ½ of natural mortality and usually much less (as can be seen by comparing yield in Fig. 3 with biomass in Fig. 7).

Neither growth nor selectivity parameters varied much across the different model configurations, and their coefficients of variation in Table 2 are small, suggesting that there is sufficient information in the combined data for estimating these parameters. We also investigated a model configuration that had growth parameters fixed at the initial values. The fit to the length frequency data was very poor, reinforcing the conclusion that estimated growth parameters are better than those obtained from tagging data.

The model tracks the size distribution of individual monthly cohorts, so that the size structure of a particular month can be examined by cohort. Fig. 9 shows the main cohorts that make up the March 1994 female endeavour prawn population from the case 3 fit. The strong November 1993 (age 5) and January 1994 (age 3) cohorts are obvious. Examination of the PDFs of size from the model for ages 2, 3, 5, and 8 (Fig. 10) shows that the spread of the cohort increases with age. This figure also shows that the size distribution of the exploitable population is shifted to the right of the population, particularly for the younger ages, which are not fully selected. The difference between the exploitable and the total portion of a cohort lessens with age as most members grow to vulnerable size (Fig. 10).

Discussion

The application to the prawn data set shows that our model is an effective tool to investigate the interaction between available data on harvest and size (by sex) and to examine the ability of estimating growth, selectivity/availability, recruitment, and absolute abundance by size. Growth, selectivity, and relative condition of the population could be determined from the available information, but absolute estimates of biomass and recruitment are not well determined. The estimates obtained from the three cases are plausible with respect to other information collected from this population, but this constitutes a fairly wide range of choices. Coefficients of variation for the estimated parameters are underestimates, because no estimates could be obtained from a model with all parameters estimated. We determined that natural mortality and catchability are the most sensitive of the model parameters and those that have the greatest impact. The results suggest that catchability is the prime quantity for converting CPUE into biomass and that natural mortality is the prime quantity for establishing absolute estimates of recruitment.

We attempted to fit natural mortality parameters because earlier research provided little information. The initial natural mortality parameters were based on literature values and the assumption that natural mortality is higher for juvenile animals. There was little information in the combined data for estimating natural mortality. Model configurations with different values of natural mortality fitted the data about the same, and there was no clear indication that natural mortality varied by age or size. The fairly small coefficient of variation (CV) for M_a with case 3 (Table 3) is certainly an underestimate, because q and M_b were fixed. The range of values for natural mortality across the cases in Table 3 is probably a better indication of the true range of variability. Other studies have suggested that parameter confounding may prevent good estimation of natural mortality (e.g., Deriso et al. 1985, Zheng et al. 1995).

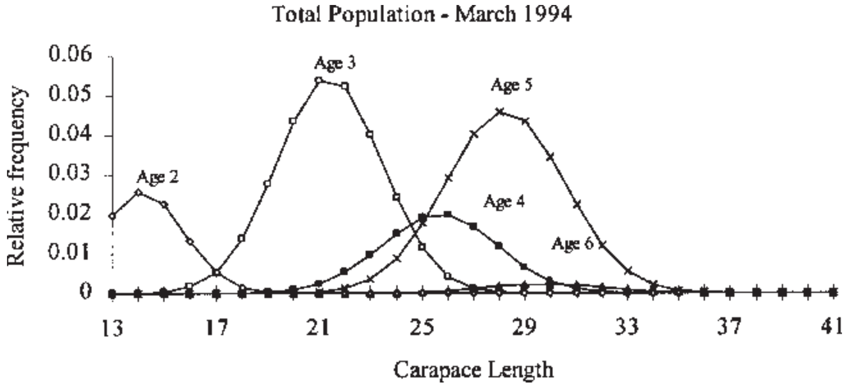


Figure 9. Relative size distributions {$p_a(x)$} for female cohorts in March 1994 from the model (case 3).

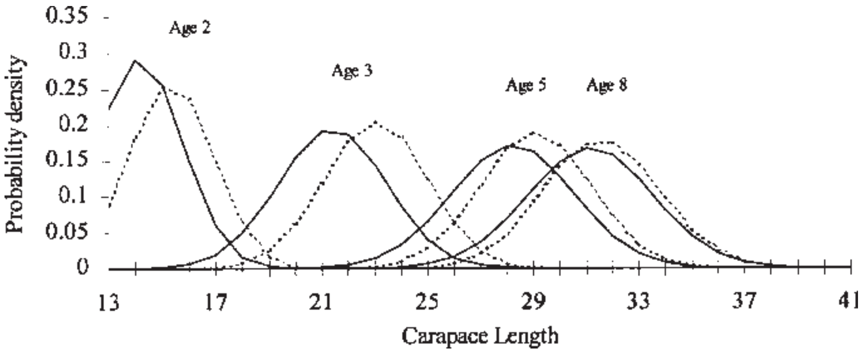


Figure 10. Probability density functions {$f_a(x)$} of the total female population (solid lines) and the exploitable population (dotted lines) in March 1994 from the model (case 3).

Estimated catchability *q* was the most variable of the model parameters (Table 3). The common confounding between catchability and natural mortality found in the literature affirms our conclusion that the ability to reliably estimate both factors is limited. We did not expect that catchability would tend toward lower values than the initial value, because we also introduced selectivity parameters that would tend to raise catchability to that for fully recruited prawns. Perhaps the previous estimate was based on a restricted population area, which would tend to elevate the value.

The parameter confounding we found is a recurring problem in stock assessment with both age- and length-based models. We found that the additional complexity in a length-based model increased the amount of confounding and that it was more difficult to understand when it occurred. Thus, length-based class of models should not be viewed as a panacea to not collect aging information but rather a tool to be used when aging information is not credible or available.

The best solution to reduce parameter confounding is to collect better data. For example, if estimates of recruitment or biomass could be made from a survey, then it would be possible to better isolate combinations of model parameters to provide a particular level of biomass. Essentially this additional data set(s) could be incorporated into the estimation procedure by adding a term to the objective function equation (21) that accounts for deviations between observed and modeled abundance. We could have done this with the survey biomass estimate of 800 t, but for the purposes of this study, we were more interested in examining whether commercial fishery information alone was sufficient, and it was not.

There are alternative length-based models such as catch-length analysis (Sullivan et al. 1990, Sullivan 1992, Zheng et al. 1995) and stock synthesis analysis (Methot 1990), which are derived from catch-age analysis (Deriso et al. 1985) in different ways. Catch-length analysis involves a probability transition matrix for growth based on the von Bertalanffy growth curve and an underlying gamma distribution for statistical variability, and stock synthesis analysis in its length-based form utilizes a length-age probability transition matrix rather than a stochastic growth model. Both models have estimation procedures associated with them.

Our length-based model is similar to these models. Its primary distinction is in the way that cohorts are individually tracked throughout their life with the possibility of age-specific differences in growth through parameters $\{\mu_a\}$ and $\{\sigma_a\}$. Both catch-length analysis and stock synthesis analysis utilize transition matrices that obviate the need to follow individual cohorts but are thus less general than our model. Our model also has the capability to utilize sex-specific information but this feature could easily be incorporated in the others.

Further improvements to the model and estimation procedure could be made. Alternative growth functions could be considered. Fishing mortality (F) could be treated as a lognormal random variable about the fishing mortality under constant catchability (qE), as is done in the models of

Deriso et al. (1985) and Sullivan et al. (1990). We did not do that at this stage of model development to avoid further confounding caused by estimating 18 fishing mortality parameters in addition to the others, but there is the possibility that seasonal differences in catchability do exist. One possible solution to this problem is to use the approach in stock synthesis, wherein total yield is assumed to be measured without error so that fishing mortality can be calculated by solving a nonlinear equation (Methot 1990). Combined process-measurement error models using Kalman filter methods (Sullivan 1992) could also be developed. With the variety of different information sources available for the endeavour prawn population, it might be useful to use some of it to construct prior distributions for population parameters within the context of a Bayesian analysis (McAllister and Ianelli 1996).

The additional information about population parameters provided by this model has already been useful. An earlier version of this model was used to generate simulated data to test various methods of isolating recruitment patterns (Watson et al. 1994). Further, it may be possible to use the model to explore the effects of seasonal closures, because the spatial dynamics of prawns are related to growth. Other uses for the model include the investigation of the impacts of alternative harvest strategies by varying fishing effort or mortality and examination of recruitment estimates in relation to environmental and biological variables. Therefore, we view this model as a necessary component of a rigorous stock assessment process that, combined with contemporaneous collection of commercial fishery and survey data including length frequency by sex, should be used for management of the Torres Strait prawn population. Currently, the process relies on CPUE information from the fishery combined with periodic surveys, and we believe that the model should be used to integrate these data sources. We also plan to apply this length-based model to other species.

Acknowledgments

We appreciate the support and comments of Dr. Gordon Kruse, Alaska Department of Fish and Game, for helping to make this project possible. This publication is the result of research sponsored by Alaska Sea Grant with funds from the National Oceanic and Atmospheric Administration, Office of Sea Grant, Department of Commerce, under grant no. NA90AA-D-SG066, project number R/07-21, and from the University of Alaska with funds appropriated by the state. The U.S. Government is authorized to produce and distribute reprints for government purposes, notwithstanding any copyright notation that may appear hereon. The Department of Industry Science and Technology (Canberra, Australia) provided travel funding for Turnbull to work on the development of the model and the application of the model to the Torres Strait prawn fishery. Jie Zheng, an anonymous referee, and editor Jake Schweigert provided helpful comments that improved this paper.

References

Cohen, M.D., and G.S. Fishman. 1980. Modeling growth-time and weight-length relationships in a single year-class fishery with examples on North Carolina pink and brown shrimp. Can. J. Fish. Aquat. Sci. 37:1000-1011.

Deriso, R.B., and A.M. Parma. 1988. Dynamics of age and size for a stochastic population model. Can. J. Fish. Aquat. Sci. 45:1054-1068.

Deriso, R.B., T.J. Quinn II, and P.R. Neal. 1985. Catch-age analysis with auxiliary information. Can. J. Fish. Aquat. Sci. 42:815-824.

Hilborn, R., and C.J. Walters. 1992. Quantitative fisheries stock assessment: Choice, dynamics and uncertainty. Chapman and Hall, New York. 570 pp.

Lai, H.-L., and V.F. Gallucci. 1988. Effects of parameter variability on length cohort analysis. J. Cons. Cons. Int. Explor. Mer 45:82-92.

McAllister, M.K., and J.N. Ianelli. 1997. Bayesian stock assessment using catch-age data and the sampling-importance sampling algorithm. Can. J. Fish. Aquat. Sci. 54:284-300.

Methot, R.D. 1990. Synthesis model: An adaptable framework for analysis of diverse stock assessment data. Int. North Pac. Fish. Comm. Bull. 50:259-277.

Parma, A., and R.B. Deriso. 1990. Dynamics of age and size in a population subject to size-selective mortality: Effects of phenotypic variation. Can. J. Fish. Aquat. Sci. 47:274-289.

Pauly, D., and G.R. Morgan (eds.). 1987. Length-based methods in fisheries research. ICLARM, Manila, Philippines. 468 pp.

Quinn II, T.J. 1985. Catch-per-unit-effort: A statistical model for Pacific halibut. Can. J. Fish. Aquat. Sci. 42:1423-1429.

Sullivan, P.J. 1992. A Kalman filter approach to catch-at-length analysis. Biometrics 48:237-257.

Sullivan, P.J., H.-L. Lai, and V.F. Gallucci. 1990. A catch-at-length analysis that incorporates a stochastic model of growth. Can. J. Fish. Aquat. Sci. 47:184-198.

Turnbull, C.T., and J.E. Mellors. 1990. Settlement of juvenile *Penaeus esculentus* (Haswell, 1879) on nursery grounds in Torres Strait. In: J.E. Mellors (ed.), Torres Strait prawn project: A review of research 1986-1988. Queensland Dept. Primary Industries Information Series, Brisbane, Australia, QI90018, pp. 29-37.

Turnbull, C.T., and R.A. Watson. 1995. The Torres Strait prawn fishery, 1994. In: Torres Strait Fisheries Assessment Group (eds.), Fisheries Assessment Report. Australian Fisheries Management Authority, Canberra, Australia.

Watson, R.A., and C.T. Turnbull. 1993. Migration and growth of two tropical penaeid shrimps within Torres Strait, northern Australia. Fish. Res. 17:353-368.

Watson, R.A., C.T. Turnbull, and K.J. Derbyshire 1994. Identifying tropical penaeid recruitment patterns. Proc. workshop on spawning stock-recruitment relationships in Australian crustacean fisheries, Bribie Is., Queensland, June 1994.

Zheng, J., M.C. Murphy, and G.H. Kruse. 1995. A length-based population model and stock-recruit relationship for red king crab, *Paralithodes camtschaticus*, in Bristol Bay, Alaska. Can. J. Fish. Aquat. Sci. 52:1229-1246.

Coping with a Challenging Stock Assessment Situation: The Kamishak Bay Sac-Roe Herring Fishery

Edward O. Otis, William R. Bechtol, and Wesley A. Bucher
Alaska Department of Fish and Game, Homer, Alaska

Abstract

Aerial biomass estimates and commercial-catch and total-run age compositions were the principal components of an integrative age-structured model used to forecast spawning biomass of Pacific herring *Clupea pallasi* returning to Kamishak Bay, Alaska. Poor weather in Kamishak Bay created gaps in aerial survey coverage, and when surveys could be flown, the ability of surveyors to observe herring schools was often limited by poor water clarity. Resultant aerial biomass estimates were influenced by survey effort and conditions. Although we used standardized procedures to quantify the surface area of observed herring schools and convert them to biomass estimates, these standards could not compensate for poor survey conditions or extended gaps in survey coverage. Aerial survey difficulties highlighted the importance of maximizing the quality and influence of other data sources in the integrative model. Total-run age-composition data, derived from fishery-independent samples, were instrumental in tracking cohorts and evaluating the strength of recruiting year classes. Historical and current data indicate that age structure was not static throughout the Kamishak Bay run. In most years the age structure of herring returning to the spawning grounds shifted from older to younger fish around early May. This created the potential for a temporal sampling bias that was mitigated by collecting age data from both early- and late-spawning components of the run and weighting the data according to the relative biomasses associated with these temporally distinct spawning aggregations. Uncertainty about current abundance is increasing and warrants a conservative harvest strategy.

Introduction

Mathematical models that integrate multiple information sources to assess fish stocks have been used for many years (Deriso et al. 1985, Megrey 1989, Hilborn and Walters 1992). Integrative models, although representing a marked improvement over assessments using a single information source such as fishery catch-per-unit-effort (CPUE), are still limited by the quality of their input data and the model's ability to relate estimated parameters to some aspect of stock dynamics. Problems that compromise the quality of individual parameter estimates limit the model's ability to accurately represent natural phenomena. While this recognition is intuitive, approaches to mitigate the problems are not. Our paper uses the 1998 forecast of the sac-roe fishery for Pacific herring *Clupea pallasi* in Kamishak Bay as a case study to illustrate some practical assessment challenges and strategies to mitigate them. Although the integrative model used to assess Kamishak herring is briefly described, our discussion focuses on the assessment and management processes: adjusting sampling strategies to maximize data quality, and considering sensitivity analyses and auxiliary information when forecasting.

Study Area

Kamishak Bay is located at about 59°10′N latitude, 153°50′W longitude, along the western shore of lower Cook Inlet in southcentral Alaska. Located 150 km from Homer, Kamishak Bay is characterized by extensive rocky reefs that create navigational hazards to mariners but provide abundant spawning habitat for herring. Kamishak Bay is vulnerable to weather fronts emanating from several directions; high winds and strong currents resulting from tidal fluctuations in excess of 8 meters frequently produce rough seas in the area. These conditions stir up sediments along the beaches and shallow bays that, combined with glacial silt from several freshwater drainages, create persistent turbidity in the marine environment. This turbidity hinders the ability of aerial surveyors to observe and quantify herring schools in the 12 index areas composing the Kamishak Bay District (Fig. 1).

Methods

Fishery and Assessment History

Pacific herring were first commercially harvested in lower Cook Inlet in 1914; however, not until 1973 did spotter pilots and pioneering fishermen first locate and exploit herring in Kamishak Bay (Schroeder 1989). Frequent storms, treacherous reefs, and the relatively remote location were responsible for the Kamishak Bay fishery's delayed development. These same characteristics created challenges for assessing and managing the fishery. The fishery developed rapidly from 220 t harvested by just a few

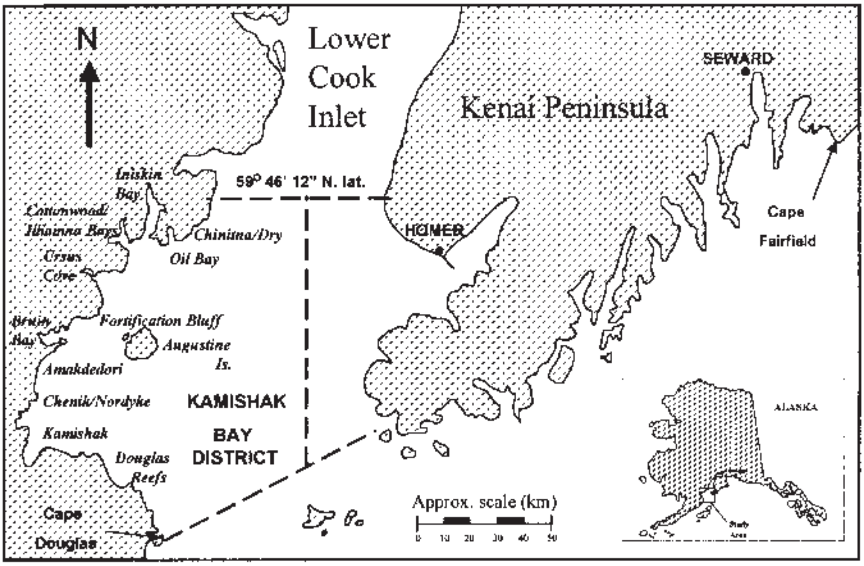


Figure 1. *Lower Cook Inlet Alaska and the Kamishak Bay District. Names of index areas are italicized; they indicate where aerial surveys of herring schools occur.*

permit holders in 1973 to nearly 4,400 t landed by 66 permit holders in 1976. To abate further expansion of effort, a ceiling of 73 permits was established when limited entry was imposed on the Kamishak Bay sac-roe herring fishery in 1977.

The Alaska Department of Fish and Game (ADFG) began collecting herring scales to estimate catch-age composition in 1973 and aerial surveys to estimate spawning biomass were initiated in 1978 (Fig. 2). The aerial estimate of total spawning biomass in a year was used to set the harvest guideline for the succeeding year. However, harvests rapidly declined from 4,393 t in 1976 to 376 t in 1979, and the fishery was closed from 1980 through 1984 to allow stock rebuilding. Limited age, weight, length, and sex data were collected, and aerial surveys were conducted to monitor stock status during the closed period. The fishery reopened in 1985 with a revised assessment program and more conservative harvest strategies. Catch-age analysis was used to develop age-specific estimates of natural mortality and recruitment. These data were integrated with aerial estimates of spawning biomass to track stock status and forecast the following year's return (Yuen et al. 1990). A stepwise harvest strategy implemented by the Alaska Board of Fisheries in 1993 set maximum exploitation rates ≤10% of the spawning biomass if that biomass was projected to be

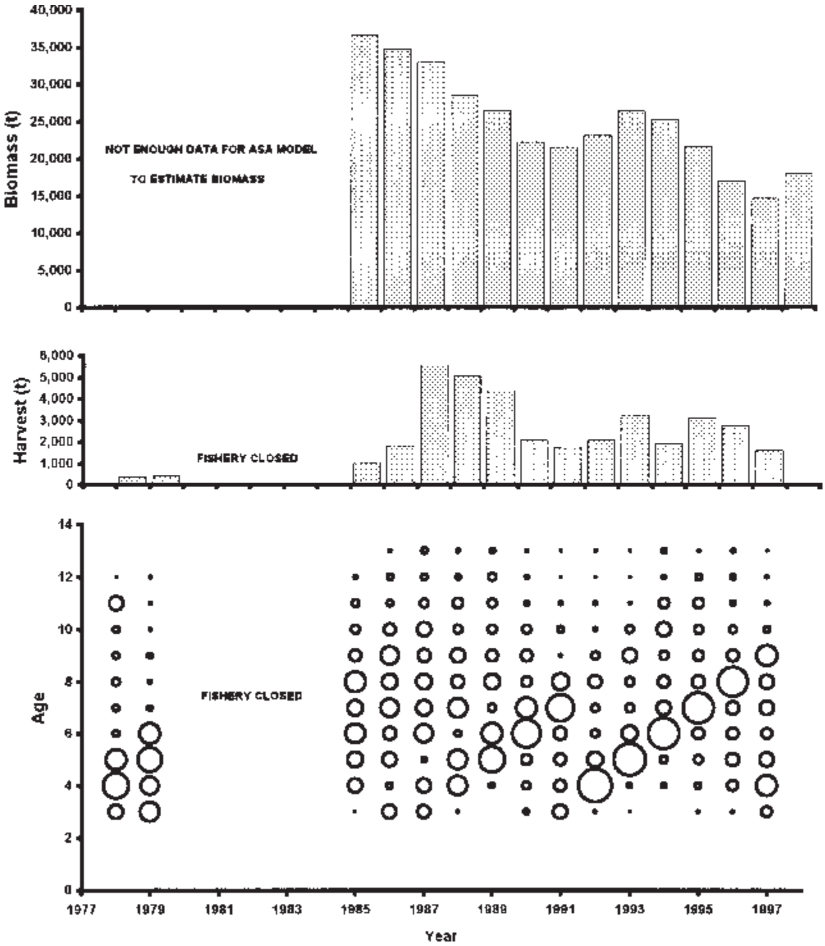


Figure 2. Kamishak Bay sac-roe herring fishery history: biomass trend, commercial harvest trend, and commercial catch age composition by year. Circle size represents the proportion that cohort contributed to the overall age composition in a given year.

7,257-18,144 t, ≤15% for biomasses of 18,144-27,216 t, or ≤20% for biomasses over 27,216 t. No harvest is allowed from a spawning biomass that is projected to be < 7,257 t (ADFG 1996, 5 AAC 27.465.e [Alaska Administrative Code]). Twenty percent is a relatively conservative maximum exploitation rate for herring (Zheng et al. 1993, Schweigert 1993, Funk and Rowell 1995). Reduced exploitation rates facilitate increased roe production since more old fish survive to be harvested in subsequent fisheries (Funk 1991). Conservative harvest rates are also prudent when stock abundance is low and/or uncertain; our stepwise harvest strategy guards against overexploitation of low biomasses.

Beginning in 1990, observer subjectivity in aerial surveys was reduced by more rigidly quantifying the biomass estimation process and documenting survey effort and conditions. In this process, the surface area of a herring school observed from a fixed distance above the water surface was converted to a biomass estimate based on a correlation developed for herring in Bristol Bay (Lebida and Whitmore 1985). Conversion ratios were stratified by water depth. Aerial observations in Bristol Bay were periodically calibrated when a commercial seine vessel captured an estimated school of herring and the catch was pumped aboard a tender vessel to determine the actual biomass. The 95% confidence interval (CI) around Bristol Bay biomass estimates based on these conversion ratios has ranged from ±14% to ±37% (L. Brannian, ADFG, Anchorage, unpublished data). Adoption of these quantitative procedures helped Kamishak Bay managers standardize aerial surveys and estimates of observed herring schools, but could not compensate for the poor visibility and large temporal gaps that hamper survey coverage in Kamishak Bay. As a result, estimates of total annual spawning biomass from aerial surveys were frequently compromised. To reduce the dependence on annual aerial surveys, Yuen et al. (1994) developed an age-structured assessment (ASA) model for Kamishak Bay.

ASA Model

The ASA model was developed in the Microsoft Excel spreadsheet using the Solver function (Microsoft 1993). The 2-dimensional format of the spreadsheet facilitates a more intuitive understanding of the model's mechanics. This feature has been very useful when explaining the model's basis to fishery managers and the Alaska Board of Fisheries, two very important links in the management chain in Alaska.

Herring are first accounted for in our model at age 3 when they begin to recruit to sexual maturity and are vulnerable to our sampling program. Herring older than age 12 do not compose a significant component of the stock and are pooled as age 13+. The ASA model integrates three heterogeneous sources of stock information: commercial catch age composition, total run age composition (i.e., total spawning biomass), and aerial survey estimates of spawning biomass. Although some fishery and stock information is available since 1978, we believed the data and collection procedures prior to 1985 were too inconsistent for inclusion in the ASA model. Thus, we truncated the model's source information in 1985, the year the fishery reopened and ADFG began consistently collecting stock information. The model estimates values for age-specific maturity, age-specific fishery selectivity, and initial cohort abundance. Differences between predicted and observed values for the three input information sources are minimized using a nonlinear optimization function (Excel Solver). Further details on model mechanics are provided in Otis and Bechtol (1997).

Because we did not believe sufficient data were available for the model to estimate survival (S), we fixed S for all years at 0.67, a relatively conservative rate that falls within the range of survival estimates reported in the literature (Funk and Sandone 1990).

Although the model updated estimates of historical abundance from 1985 to present, our primary goal was to generate a 1-year-ahead forecast of the herring spawning biomass. A forecast allows area managers to set harvest guidelines for the following year's commercial sac-roe fishery and allows the industry to plan accordingly. Unless in-season assessment data dramatically conflict with the pre-season forecast, the fishery is managed based on the pre-season forecast.

Forecast Procedure

Model inputs for the 1998 forecast included commercial harvests, catch age compositions, total run age compositions, weights-at-age, and aerial survey biomass estimates from 1985 to 1997. Final parameter estimates for initial cohort abundance and age-specific maturity and fishery selectivity from the 1997 forecast were used as starting values for the 1998 forecast. Several scenarios were constructed whereby data inputs were slightly modified. For instance, the 1997 aerial survey data were of low value due to weather-induced gaps in survey coverage and a decision had to be made whether to ultimately include those data in the model. Trials were run with and without the 1997 aerial survey data to evaluate their effect on results. Once all viable data input scenarios had been identified, they were run through the model several times while incrementally increasing the weighting factor attributed to each model component. The resulting sum of squares (SSQ) values and historical biomass trends were then plotted relative to these incremental changes to evaluate their respective effects on select model results. This sensitivity analysis identified weighting scenarios yielding unstable or unrealistic model results. The primary goal of data weighting is to allow the model to explore options to incorporate the different data sets, such that the deviation between predicted and observed values is minimized among all available data. Finally, in a process that involved some qualitative decisions based on all available auxiliary information, a single model run was selected to represent the current stock status and provide a 1-year-ahead forecast.

Results

Model Inputs

Since 1987, the commercial fishery has typically occurred over a very brief period in late April. The 1997 fishery similarly consisted of short openings on April 29 and 30 and May 1. ADFG staff collected catch samples from as many fishing vessels as possible while the catch was pumped aboard tenders. Age, weight, length, and sex information were later compiled in our Homer laboratory. Each sampled vessel's contribution to the total catch was determined from harvest tickets, and the data were weighted accordingly to estimate the age composition for the total catch.

Collecting age composition data to accurately represent the total run was more problematic. Fishing vessels on the grounds prior to the commercial opening collected early season data. Because ADFG will not announce fishery openings until adequate samples are collected to estimate the age composition of the biomass on the grounds, the fleet had a vested interest in collecting those data. ADFG's ability to collect age samples diminished when the fishery closed and the fleet departed the grounds around 2 May. Successive waves of herring continued to enter the spawning grounds until early June. Historical data collected in Kamishak Bay indicated the age structure shifted from older to younger fish around the first week of May (Yuen 1994). Samples collected from two temporally distinct spawning aggregations in 1997 revealed a similar shift in age structure (Fig. 3), and chi-square analysis indicated the shift was significant at $\alpha = 0.005$ ($\chi^2 = 1711$; 10 d.f.). The mean weight of early returning fish ($\bar{x}$ = 202.8 g, SD = 65.6; $n = 2{,}883$) also differed significantly from late-returning fish ($\bar{x}$ = 122.4 g, SD = 38.2; $n = 2{,}203$; two-tailed, two-sample t-test, $p < 0.0001$). Differences in mean weights between components of the run did not just result from their significantly different age structures; herring mean weights-at-age were also consistently greater for early returning fish (Fig. 3). These results revealed the potential for a temporal sampling bias when estimating the total run age composition, herring mean weight-at-age, and overall mean weight.

Although we have not yet evaluated the impact a temporal sampling bias would have on the model's estimate of stock level, we chose to guard against potential impacts by selecting a transition date to represent the shift between early and late returning components of the run. We then weighted the age samples by the respective aerial-survey biomasses estimated before and after the transition date. Based upon historical observations of age shift occurrence (Yuen 1994), we selected 8 May as the transition date for 1997. By this criterion, early run fish composed 41% and late-run fish 59% of the season's total biomass. We used the same ratio to weight our overall mean weight and mean weight-at-age data prior to their inclusion in the model.

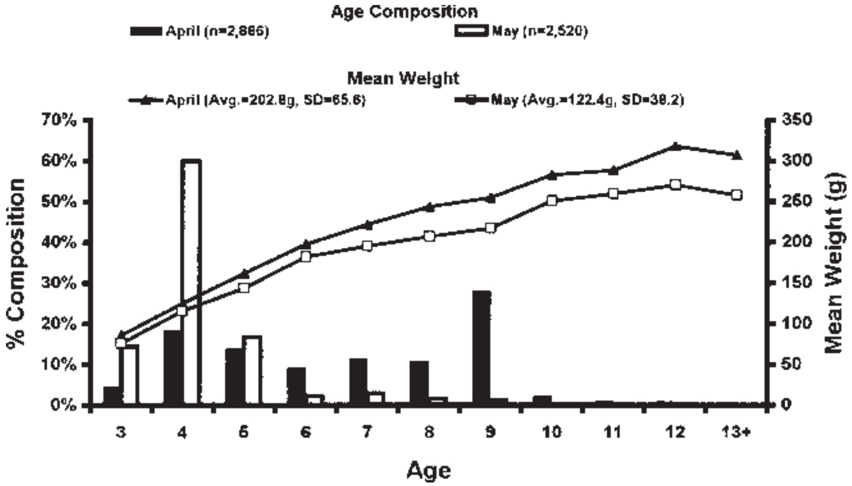


Figure 3. Age compositions and mean weights-at-age of two temporally distinct spawning aggregations of herring sampled in Kamishak Bay, Alaska, 1997.

Since the aerial survey program was restructured in 1990, only 1990 and 1992 were judged to have adequate survey coverage and conditions during peak biomass periods to yield realistic total spawning biomass estimates. An example of what we considered to be unrealistic aerial survey results occurred in 1991, a year characterized by good early season survey coverage, but poor survey conditions. The peak counts of herring biomass in 1991 amounted to less than 3,000 t. This contrasts with the roughly 23,000 t estimated during 1990 and 1992. Ancillary stock abundance information (e.g., fleet hydroacoustic observations and achievement of the harvest quota after only 1 hour of fishing) also indicated that the 1991 aerial survey biomass estimate grossly under-represented the actual abundance of spawning herring.

Evaluation of the 1990 survey year illustrates how important "good timing" is, along with adequate survey effort and conditions, in order to estimate total spawning biomass in Kamishak Bay (Figs. 4a and 4b). The observed biomass increased rapidly from 22 to 25 April in 1990. As evidenced by miles of spawn observed from the air, significant spawning also occurred throughout that period. Evidence of spawn in Kamishak Bay disappeared concurrent with the abrupt departure of the spawning biomass between 25-27 April 1990. This suggests the first spawning wave of herring had a brief period of residency and immediately left the area after spawning. About 4,500 t of spawning herring would not have been detected if flyable weather and good water visibility had not coincided during that brief 4-day period in early April 1990. Similarly, another 4,500 t would

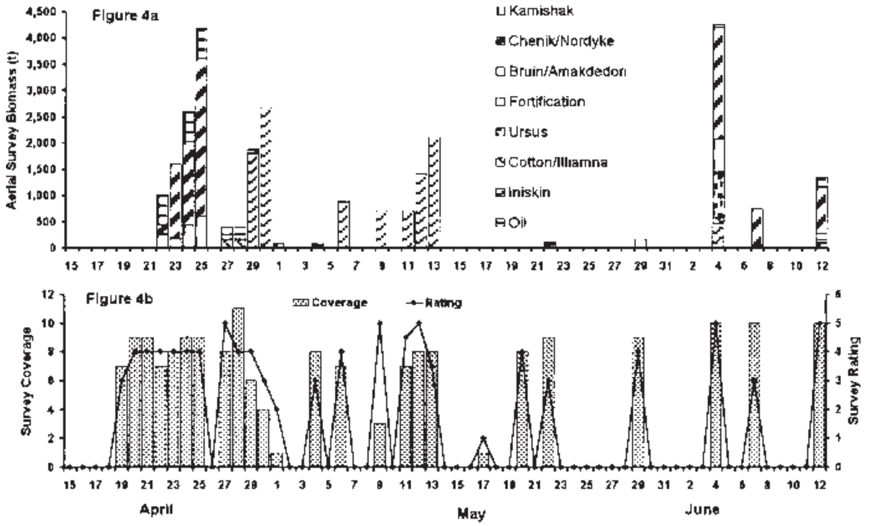


Figure 4. (a) Herring biomass estimates by index area from aerial surveys flown in Kamishak Bay, Alaska, 1990; and (b) Aerial survey coverage and survey condition ratings for aerial surveys flown in Kamishak Bay, Alaska, 1990. Survey coverage indicates the number of index areas flown on a given day; rating refers to the survey conditions: 0 = no survey, 5 = excellent.

have been unaccounted for had the 4 June survey not been successful. The combined biomass from these two events, which easily could have been missed, composed nearly 40% of the total observed spawning biomass for 1990.

Weather-induced gaps in survey coverage hampered our ability to estimate total spawning biomass in 1997. Because of the effect that poor survey coverage can clearly have on biomass estimates, we decided not to include the 1997 aerial surveys in the model. Instead, we again used only the survey years in which comprehensive coverage and adequate conditions facilitated realistic total spawning biomass estimates.

Sensitivity Analysis and Forecast

Aerial survey weights >0.01 stabilized the forecast biomass (Fig. 5a). Thus, we selected an aerial survey weight ≥0.01 that was strong enough to draw the historical biomass trend through the 1990 and 1992 aerial survey data points. We determined that an aerial survey weight around 0.1 was sufficient to achieve this result without compromising the fit to the age composition data (Fig. 5b).

Similar techniques were used to filter out unstable weightings for catch and total-run age composition data. The catch and total-run age composition

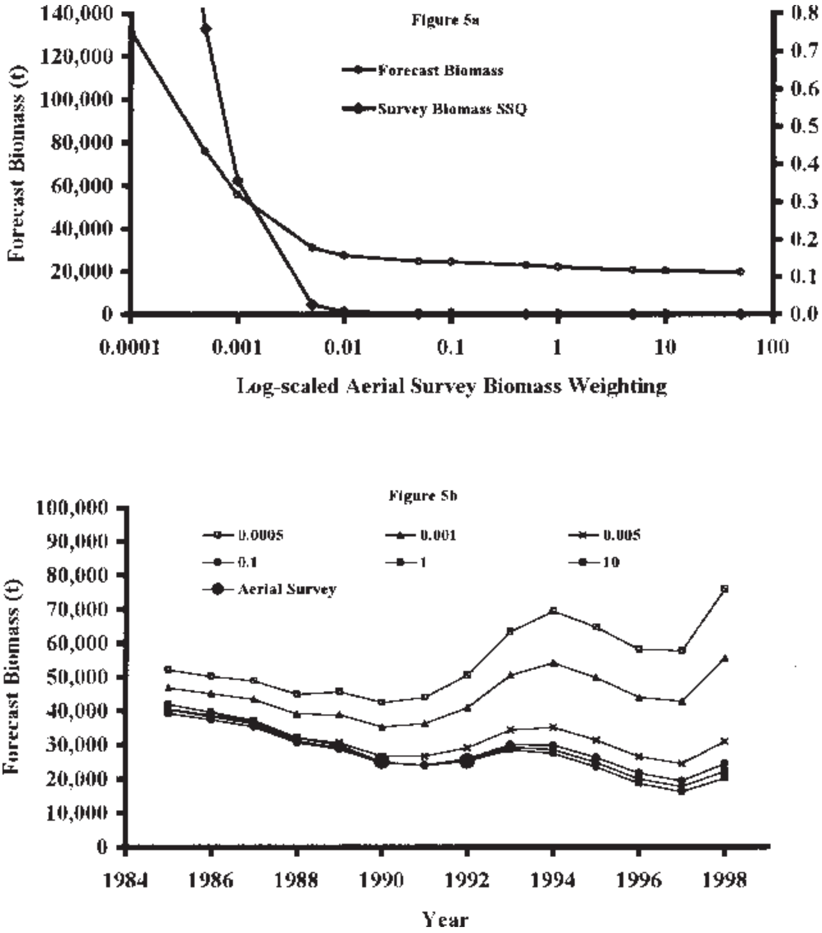


Figure 5. (a) ASA forecast biomass and aerial survey sum of squares (SSQ) trends relative to increases in the log-scaled aerial survey weighting; and (b) Historical biomass trend relative to increases in the log-scaled aerial survey weighting.

sums of squares conflicted in their reactions to increases in weighting (Fig. 6). Catch and total-run weightings in the range of 0.5-1.0 provided relatively stable results without pushing the forecasted biomass into what we perceived to be unrealistic ranges. The forecasted biomasses from these alternative weightings ranged from 14,250 to 32,830 t. This wide forecast range suggested there was some uncertainty, and perhaps even data conflicts, associated with the various model inputs. Given these uncertainties, a relatively conservative forecast was the only prudent alternative. We selected a final weighting schedule that yielded a forecast of 17,870 t and a harvest guideline of 1,787 t.

Other auxiliary information was also instrumental in the final qualitative decision to select a conservative forecast and harvest strategy. Our 1997 field samples indicated strong recruitment from the 1993 and 1994 cohorts; these year classes combined are projected to make up over 55% of the 1998 return. However, these two cohorts are just beginning to appear in our field samples and their absolute abundance is still highly uncertain. In addition, the regulatory Kamishak Bay Herring Management Plan contains a provision to limit exploitation on herring age 5 and younger. This provision necessitates a conservative harvest strategy for 1998 because recruit classes are projected to compose the majority of our 1998 return. Forecast uncertainty is also increasing because of the years elapsed since the last quality aerial survey biomass estimate in 1992.

Discussion

Aerial surveys generally provide the most direct means for assessing herring spawner abundance. However, poor weather in Kamishak Bay creates gaps in the aerial survey time series and high water turbidity frequently limits the ability of surveyors to observe herring schools when surveys can be flown. Resultant aerial-survey biomass estimates are sometimes unrealistically low and highly influenced by a given year's survey effort (i.e., spatial and temporal coverage) and conditions. Our efforts to mitigate these effects by standardizing survey effort across years have been largely unsuccessful. Attempts to standardize aerial biomass estimates for survey effort by documenting the number of survey hours flown and index areas surveyed does not account for temporal effects and variable survey conditions. Surveys that are missed, incomplete, or of poor quality during periods of peak abundance compromise the estimate of total spawning biomass more than surveys missed between spawning waves. It is difficult to build this effect into a model (e.g., by fitting a run-timing curve to aerial survey estimates) when the run is divided into waves of returning herring whose spawning peaks exhibit inter-annual temporal variability (H. Yuen, ADFG, Homer, unpublished manuscript). It is doubtful that a long-term average run-timing curve will accurately represent the actual run timing within a given year. Developing a baseline of run-timing curves relative to varying physical (e.g., tide series, water temperature: Wespestad

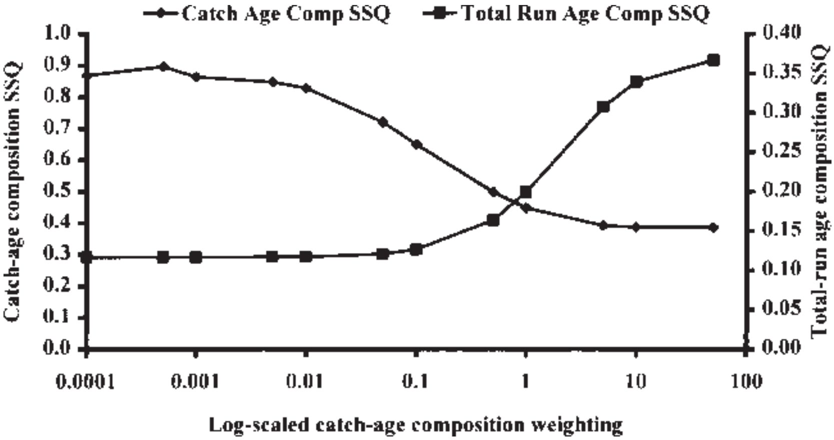


Figure 6. Catch- and total-run age composition sum of squares relative to increases in the log-scaled catch-age composition weighting.

1991) and biological (e.g., age composition, fat content, and mean weight-at-age of returning fish: Ware and Tanasichuk 1989, Rajasilta 1992) relationships may allow us to incorporate these parameters in future models.

Many herring are unaccounted for by our aerial survey assessments, even during good survey years. For instance, in 1990 the miles of observed spawn peaked on 21 April, prior to any herring schools being observed from the air, despite good survey effort and conditions. Nøttestad et al. (1996) found that Norwegian spring-spawning herring schools immigrating to the spawning grounds swam much deeper than spawning, feeding, or emigrating schools. This behavioral trait may make immigrating schools less susceptible to aerial assessment and partially explain why no herring schools were observed in 1990 until spawning began in earnest.

More tangible evidence of unsurveyed herring biomasses was consistently documented during the years 1992-1996. In each of these 5 years, virtually no herring were observed from the air by management staff prior to the date of the first commercial landings. An average of 2,050 t of unassessed herring was caught in these initial openings, and over 3,200 t of unassessed herring were caught in 1993. These data, along with anecdotal hydroacoustic observations by fishermen, have led us to be very critical of our aerial survey results, particularly when poor weather inhibits visibility or creates lengthy gaps in survey coverage.

By using ancillary stock-abundance information and sensitivity analyses to identify weak data sets, we can mitigate their impact on forecasts by down-weighting them or excluding them from the model. While this treatment of weak data is intuitively appealing, it can create side effects. Although the ASA model does not require an abundance-scaling index for every year, moderate amounts of auxiliary information are required to stabilize stock estimates generated by catch age analysis (Deriso et al. 1985). The greater the period between the last abundance-index (i.e., good survey year) and the current forecast, the more uncertain the forecast becomes.

We have tried to enhance the quality of other model inputs to compensate for uncertainties caused by our lack of consistent auxiliary information. Our expectation is that the effects of not having a recent abundance index could be mitigated if we had a continuous time series of data to annually represent the age composition of the total run. Collection of these data has only been possible during 8 of the past 13 years. Lacking a continuous time series of total run age composition data, the model cannot effectively scale upcoming recruit classes to strong cohorts it has tracked for several years. Recognizing the importance of this limitation to the model, ADFG has now begun to re-emphasize late-season sampling with the addition of an annually chartered test-fishing vessel. As this continuous time series grows and incorporates more years in which upcoming recruit classes overlap with strong older age classes, the model's ability to compensate for infrequent aerial survey estimates should be enhanced. The model's ability to scale upcoming recruits to older year classes will be particularly improved if a good aerial assessment occurs at least once during the life span of strong year classes. Maintaining continuous time series for both catch and total run age compositions also should improve our estimates of age-specific maturity and fishery selectivity and allow the model to estimate survival. These improvements may reduce the forecast uncertainty resulting from our lack of a consistent abundance index.

Although maintaining a continuous time series of catch and total run age compositions should improve the model's performance, a reliable abundance-scaling index remains its most integral component. Given the inconsistency with which aerial surveys have been able to provide this index we are considering other methods. Compact airborne spectrographic imaging (CASI) equipment has been used successfully to digitally remote-sense and quantify forage fish schools in Canada (Borstad et al. 1992, Nakashima and Borstad 1997) and Alaska (Funk et al. 1995). CASI works by discriminating between the spectral signatures of fish schools and their natural background; the surface areas of observed herring schools are then calculated from the digital images captured by CASI. Unfortunately, the turbid waters and inhospitable flying conditions characteristic of Kamishak Bay would likely inhibit CASI's ability to improve upon our current aerial survey program (Pers. comm., Gary Borstad, G.A. Borstad Assoc. Ltd., B.C. V8L 3S1, February 1998.).

Miles-of-milt indices have been used to scale spawner-abundance in Prince William Sound herring assessments (Funk 1994). We have consistently documented miles of milt observed during aerial surveys in Kamishak Bay since 1990. However, the same conditions that limit our aerial assessment of herring schools in Kamishak Bay also reduce our ability to observe and quantify evidence of spawn. Consequently, there is a high level of uncertainty regarding the completeness of our miles-of-milt data that has made us reluctant to use them in our model. In the future, we may investigate the utility of including miles-of-milt indices from years with good survey coverage and conditions. Developing technology, such as high resolution images from low orbit satellites (Pers. comm., Ron Brooks, remote sensing consultant, Fairbanks, AK, March 1998.), may facilitate locating and quantifying spawning events (i.e., miles of milt) more consistently in the future.

Egg deposition surveys have been used in Alaska (Funk 1994) and elsewhere in the Pacific Northwest (Schweigert and Stocker 1988, Burton 1991) to calculate estimates of total spawning biomass. In 1991, ADFG conducted foot surveys of the intertidal reefs and shoreline around Chenik Lagoon in Kamishak Bay to determine the feasibility of estimating egg deposition (Yuen 1993). These data allowed a rough calculation (i.e., 95% CI was ±100% of the estimate) of spawning biomass for the small area surveyed; however, several factors limit the feasibility of pursuing spawn deposition surveys in Kamishak Bay. Most of them involve the fact that the timing and distribution of spawning is protracted and would require costly, labor-intensive surveys to obtain a viable sample. Akenhead et al. (1993) investigated the feasibility of using CASI to quantify exposed egg biomass at low tide. Remotely sensing herring eggs is a relatively expensive and uncertain process that we decided not to pursue given the unique characteristics of Kamishak Bay (e.g., poor flying conditions, relatively sparse and widely distributed spawn, high potential for subtidal spawning to be missed; pers. comm., Gary Borstad, G.A. Borstad and Associates, Ltd. B.C., V8L 3S1, February 1998.).

Another consideration that reduces the feasibility of spawn deposition surveys in Kamishak Bay involves establishing the extent of subtidal spawning. To estimate spawning biomass from egg deposition with any accuracy, one must estimate the total number of eggs deposited. Annual scuba surveys would be required to estimate the contribution subtidal spawning makes to total egg deposition. Along with their expense, high energy beaches, swift tidal currents, sea lions and kelp forests make scuba diving in Kamishak Bay an undesirable assessment method. These practical and budgetary limitations restrained our interest in pursuing egg deposition surveys during an era when weather permitted the periodic success of aerial surveys. Now that 5 years have passed since our last "good survey" year, we are reconsidering alternative methods to obtain consistent abundance indices to stabilize our age-structured herring assessment model. Despite the challenges associated with securing accurate and consistent input data, we believe our use of an integrative model has resulted in improved management of the Kamishak Bay sac-roe herring fishery.

Acknowledgments

Contribution PP-166 of the Alaska Department of Fish and Game, Commercial Fisheries Management and Development Division, Juneau. Henry Yuen, Fritz Funk, and Linda Brannian developed the ASA model presently used to assess herring in Kamishak Bay. Fritz Funk also provided advice regarding sensitivity analyses and summarizing assessment and fishery information for in-season management of the fishery. We thank Fritz Funk, Robert Wilbur, Jake Schweigert, and an anonymous reviewer for improving our manuscript through their insightful comments.

References

ADFG. 1996. 1996-1997 Commercial herring fishing regulations. Alaska Department of Fish and Game, Juneau. 75 pp.

Akenhead, S.A., G.A. Borstad, R.C. Kerr, and D.A. Hill. 1993. Airborne remote sensing for herring egg biomass, littoral zone mapping, and herring fisheries management. Final report submitted to Jake Schweigert, Pacific Biological Station, DFO, Canada by G.A. Borstad Associates Ltd., Sydney, B.C. 46 pp.

Borstad, G.A., D.A. Hill, R.C. Kerr, and B.S. Nakashima. 1992. Direct digital remote sensing of herring schools. Int. J. Remote Sens. 13(12):2191-2198.

Burton, S.F. 1991. Comparison of Pacific spawner herring biomass estimates from hydroacoustic-trawl and spawning ground escapement surveys in Puget Sound, Washington. In: Proceedings of the International Herring Symposium. University of Alaska Sea Grant, AK-SG-91-01, Fairbanks, pp. 209-221.

Deriso, R.B., T.J. Quinn II, and P.R. Neal. 1985. Catch-age analysis with auxiliary information. Can. J. Fish. Aquat. Sci. 42:815-824.

Funk, F.C. 1991. Harvest policy implications of yield per recruit models for Pacific herring in Alaska. In: Proceedings of the International Herring Symposium. University of Alaska Sea Grant, AK-SG-91-01, Fairbanks, pp. 453-476

Funk, F.C. 1994. Forecast of the Pacific herring biomass in Prince William Sound, Alaska, 1993. Alaska Department of Fish and Game, Regional Information Report 5J94-04, Juneau. 39 pp.

Funk, F.C., and K.A. Rowell. 1995. Population model suggests new threshold for managing Alaska's Togiak fishery for Pacific herring in Bristol Bay. Alaska Fishery Research Bulletin 2(2):125-136.

Funk, F.C., and G.J. Sandone. 1990. Catch-age analysis of Prince William Sound, Alaska, herring, 1973-1988. Alaska Department of Fish and Game, Fishery Research Bulletin 90-01, Juneau. 36 pp.

Funk, F.C., G.A. Borstad, and S.A. Akenhead. 1995. Imaging spectrometer detects and measures the surface area of Pacific herring schools in the Bering Sea. In: Proceedings of the Third Conference on Remote Sensing of Marine and Coastal Environments, Seattle, WA, 18-20 September, 1995, pp. 833-844.

Hilborn, R., and C.J. Walters. 1992. Quantitative fisheries stock assessment: Choice, dynamics, and uncertainty. Chapman and Hall, New York.

Lebida, R.C., and D.C. Whitmore. 1985. Bering Sea herring aerial survey manual. Alaska Department of Fish and Game, Bristol Bay Data Report 85-2, Anchorage. 31 pp.

Megrey, B.A. 1989. Review and comparison of age-structured stock assessment models from theoretical and applied points of view. Amer. Fish. Soc. Symp. 6:8-48.

Microsoft Corporation. 1993. Microsoft Excel user's guide. Microsoft Corporation, Redmond, WA.

Nakashima, B.S., and G.A. Borstad. 1997. Abundance and behavior of capelin schools *(Mallotus villosus)* from aerial surveys. In: Forage fishes in marine ecosystems. University of Alaska Sea Grant, AK-SG-97-01, Fairbanks, pp. 383-393.

Nøttestad, L., M. Aksland, A. Beltestad, A. Fernö, A. Johannessen, and O.A. Misund. 1996. Schooling dynamics of Norwegian spring spawning herring (*Clupea harengus* L.) in a coastal spawning area. Sarsia 80:277-284.

Otis, E.O., and W.R. Bechtol. 1997. Forecast of the Kamishak herring stock in 1997. Alaska Department of Fish and Game, Regional Information Report 2A97-03, Anchorage. 37 pp.

Rajasilta, M. 1992. Relationship between food, fat, sexual maturation, and spawning time of Baltic herring *(Clupea harengus membras)* in the Archipelago Sea. Can. J. Fish. Aquat. Sci. 49:644-654.

Schroeder, T.R. 1989. A summary of historical data for the lower Cook Inlet, Alaska, Pacific herring sac roe fishery. Alaska Department of Fish and Game, Regional Information Report 89-04, Anchorage. 114 pp.

Schweigert, J.F. 1993. Evaluation of harvesting policies for the management of Pacific herring stocks, *Clupea pallasi,* in British Columbia. In: Proceedings of the International Symposium on Management Strategies for Exploited Fish Populations. University of Alaska Sea Grant, AK-SG-93-02, Fairbanks, pp. 167-190.

Schweigert, J.F., and M. Stocker. 1988. Escapement model for estimating Pacific herring stock size from spawn survey data and its management implications. N. Am. J. Fish. Manage. 8:63-74.

Ware, D.M., and R.W. Tanasichuk. 1989. Biological basis of maturation and spawning waves in Pacific herring *(Clupea harengus pallasi).* Can. J. Fish. Aquat. Sci. 46:1776-1784.

Wespestad, V.G. 1991. Pacific herring population dynamics, early life history, and recruitment variation relative to eastern Bering Sea oceanographic factors. Ph.D. dissertation, University of Washington, Seattle. 237 pp.

Yuen, H.J. 1993. Intertidal survey of herring spawn in Lower Cook Inlet. Alaska Department of Fish and Game, Regional Information Report 2A93-29. 52 pp.

Yuen, H.J. 1994. A model to predict pacific herring age composition in early and late spawning migrations in Kamishak Bay, Alaska. Alaska Fishery Research Bulletin 1:35-54.

Yuen, H.J., L.K. Brannian, and F. Funk. 1994. Forecast of the Kamishak herring stock in 1994. Alaska Department of Fish and Game, Regional Information Report 2A94-12, Anchorage. 42 pp.

Yuen, H.J., W.A. Bucher, and R. Morrison. 1990. Methods for the 1991 Kamishak herring stock projection. Alaska Department of Fish and Game, Regional Information Report 2H90-16. 24 pp.

Zheng, J., F.C. Funk, G.H. Kruse, and R. Fagen. 1993. Evaluation of threshold management strategies for Pacific herring in Alaska. In: Proceedings of the International Symposium on Management Strategies for Exploited Fish Populations. University of Alaska Sea Grant, AK-SG-93-02, Fairbanks, pp. 141-165.

Abundance Estimation of St. Matthew Island Blue King Crabs Using Survey and Commercial Catch and Effort Data

J. Zheng, M.C. Murphy, and G.H. Kruse
Alaska Department of Fish and Game, Juneau, Alaska

Abstract

We conducted a modified catch-survey analysis incorporating commercial catch per unit effort and bottom-trawl survey data for the blue king crab stock, *Paralithodes platypus*, near St. Matthew Island in the eastern Bering Sea. Although the general area was surveyed to estimate abundance of crabs, the commercial fleet fished over a more limited and generally untrawlable portion of the surveyed area. Therefore, the current abundance estimation procedure based on assessment surveys was modified to include fishery performance data to explore the utility of both types of data in estimating abundance. Total annual catch and observed relative survey abundance of legal-sized crabs were classified into recruits and postrecruits by shell age and size, and commercial fishing effort was defined as total potlifts. The analysis separated measurement errors from the true changes in population abundance and converted the relative abundance to an absolute value. The analysis also provided smoothed annual estimates of the legal blue king crab abundance. Based on a probable range of instantaneous natural mortality (0.16 to 0.35) and fishing effort weighting factors (0.5 and 1.0), we estimated survey catchability to be 0.91 to 1.25. Weighting the fishing effort heavier or assuming higher natural mortality resulted in lower estimates of survey catchability. Because of confounded effects among survey catchability and natural mortality, additional information is needed to reliably estimate these parameters.

Introduction

In the eastern Bering Sea, blue king crabs, *Paralithodes platypus*, support commercial fisheries off St. Matthew Island and the Pribilof Islands. The

St. Matthew Island fishery developed subsequent to baseline ecological studies associated with oil exploration (Otto 1990); 10 U.S. vessels harvested 545 t in 1977. Harvest peaked in 1983 when 164 vessels landed 4,288 t. From 1986 to 1991 the fishery was fairly stable, harvesting a mean of 731 t by <70 vessels. Participation increased from 68 vessels in 1991 to 174 vessels in 1992. Since 1993, the St. Matthew and Pribilof Islands blue king crab fisheries have been opened concurrently, dividing vessel effort between both fisheries and initially stabilizing vessel participation at about 90 vessels. To reduce total fishing effort and improve manageability of the relatively small allowable harvests, maximum limits of 60 pots and 75 pots were set in 1993 for vessels <38.1 m and ≥38.1 m, respectively. Those limits reduced the number of pots registered by a third from 1992 to 1993 (Morrison et al. 1997). However, the number of potlifts in the fishery increased slightly because the season length doubled and pot turnover rates increased. In 1996 and 1997 participation increased to an average of 120 vessels per year and the average number of potlifts increased 56% from 1992.

St. Matthew Island blue king crabs inhabit rugged bottom typified by patches of cobble, rocks, and boulders. This habitat is well suited to fishing with crab pots. However, assessment surveys of the stock have traditionally been conducted with a bottom trawl, and area-swept estimates of abundance have been used to set annual guideline harvest levels (Otto et al. 1997). Area-swept abundance estimates are affected by survey measurement errors, which for blue king crabs have been attributed to limited trawlable terrain impeding gear performance and to potential changes in crab distribution across fixed survey stations as effected by possible biennial mating and molting migrations of large female crabs (Somerton and MacIntosh 1985; Jensen and Armstrong 1989; Blau 1996). Poor assessments of female and juvenile crabs by the trawl survey may be due to preferences for rocky habitat (Otto et al. 1997) that probably offers protection from predation and cannibalism.

Wide swings in annual survey abundance estimates affect interpretations of stock status, application of the fishery management strategy, and stability in annual harvests. Underestimation causes reduced harvest opportunity; overestimation increases risk of overfishing. To separate survey measurement errors from the true changes in population abundance, Zheng et al. (1997) constructed a catch-survey analysis (CSA, Collie and Kruse 1998) of the St. Matthew blue king crab stock. CSA is particularly well suited to this stock because of uncertainty in survey catchability, large measurement errors, and limited size-frequency data needed for more detailed length-based analyses (Zheng et al. 1996). Previously, we used only survey and commercial catch data to analyze this stock (Zheng et al. 1997). The survey area and the area of commercial harvest are not coincident in part due to untrawlable habitat. The commercial fishery occurs on grounds about 1,000 square nautical miles (nm^2) whereas the survey stations cover a broad area and area-swept estimates are extrapolated to approximately 10,000 nm^2 (Pers. comm., D. Pengilly, Alaska Department of Fish and Game). In this study, we incorporated commercial catch per unit effort (CPUE) data into the catch-survey model to account for the disparity between survey and commercial fishing grounds and to examine the effect of fishery performance on estimates of survey catchability and absolute population abundance. The application of the model is limited to estimating abundance of only legal-sized males because the analysis relies on landings data.

Methods

Survey, Catch, and Effort Data

We obtained survey data from the National Marine Fisheries Service (Otto et al. 1997) and catch and effort data collected by Alaska Department of Fish and Game from 1980 to 1997 (Morrison et al. 1997). We updated the area-swept estimates of crab abundance from the 1980-1997 trawl survey data using the stratification method developed by Zheng et al. (1997). Prior to 1980, trawl survey data are incomplete and few commercial vessels targeted this stock. For convenience we refer to our area-swept estimates of abundance as the "observed relative abundance."

We defined nominal CPUE as total season catch of legal males divided by potlifts. Because the nominal CPUE is greatly affected by total number of vessels in the fishery and the season length, with a larger number of vessels or a longer season resulting in a lower nominal CPUE, we standardized the CPUE as

$$U_t = C_t / E_t^r \tag{1}$$

where U_t is standardized CPUE in year t, C_t is annual catch, E_t is total potlifts, and r is a parameter with a value < 1. Because pot limits were initiated in 1993, two r parameters, $r1$ and $r2$, were estimated for 1980-1992 and 1993-1997, respectively. For comparison to observed potlifts, we also estimated total potlifts as

$$E_t = \frac{C_t}{[q_c(A_t e^{(-T_t M)} - 0.5C_t)]^{1/r}} \tag{2}$$

where q_c is commercial pot catchability, A_t is absolute legal crab abundance in year t, T_t is the time lag from the survey to the midpoint of the fishery in year t, and M is natural mortality.

Catch-Survey Analysis

We estimated abundance of legal blue king crabs by a modified CSA that incorporates commercial CPUE data. Kruse and Collie (1991) and Collie and Kruse (1998) developed the two-stage CSA for red king crab populations, *Paralithodes camtschaticus*, and Zheng et al. (1997) applied it to the two blue king crab stocks off the Pribilof and St. Matthew islands. A useful feature of CSA, in addition to accounting for measurement errors separately from true changes in stock abundance, is that the survey catchability coefficient can be estimated given an estimate of natural mortality.

Legal male blue king crabs have a carapace length (CL) ≥120 mm in the St. Matthew Island fishery. The average growth increment per molt for legal male crabs is about 14 mm CL (Otto and Cummiskey 1990). Legal blue king crabs were categorized into recruits (newshell, 120-133 mm CL) and postrecruits (all other legal males). The model links postrecruits in year t+1 to recruits, postrecruits, and catch in year t through natural mortality and catchability of the survey gear:

$$p_{t+1} = (p_t + r_t)e^{-M} - q_s C_t e^{-M(1-T_t)} \tag{3}$$

where p_t and r_t are relative abundances of postrecruits and recruits in year t, and q_s is the survey catchability for legal crabs. For simplicity, we assumed that recruits and postrecruits have the same survey catchability. These relative abundances divided by the survey catchability are equal to corresponding absolute abundances of recruits and postrecruits by year. Thus, for a given M, the CSA model will estimate q_s, which scales relative abundance to absolute abundance for the entire stock in both trawlable and untrawlable areas.

Parameter Estimation

Measurement errors of survey estimates of recruit and postrecruit relative abundances and fishing effort were assumed to follow lognormal distribution, and the least squares approach was used to minimize residual sum of squares (RSS):

$$\begin{aligned} RSS = \sum_t \{&[\log(p_t+\kappa)-\log(\tilde{p}_t+\kappa)]^2 + [\log(r_t+\kappa)-\log(\tilde{r}_t+\kappa)]^2 \\ &+ w[\log(E_t+\kappa)-\log(\tilde{E}_t+\kappa)\} \end{aligned} \tag{4}$$

where $\tilde{p}_t$ and $\tilde{r}_t$ are observed relative abundances of postrecruits and recruits in year t, $\tilde{E}_t$ is observed fishing effort in year t, κ is a small constant set as 0.001, and w is an effort weighting factor. We examined w values ranging from 0 to 5.0.

Estimated parameters include M, q_s, q_c, $r1$, $r2$, recruits each year except the last, and postrecruits in the first year. Ideally, survey catchability or natural mortality would first be estimated through research studies. However, uncertainty exists about both parameters for the St. Matthew Island blue king crab stock. For this reason, we used three different approaches to estimate survey catchability and natural mortality, which are confounded and difficult to estimate simultaneously. First, we assumed $q_s = 1$ and estimated M. Second, we estimated q_s for M ranging from 0.1 to 0.5 at intervals of 0.05. Finally, we used a 2-phase approach to estimate both q_s and M by (1) setting $q_s = 1$ and estimating M, and (2) estimating both q_s and M simultaneously.

Using AD Model Builder (Otter Research Ltd. 1994), we estimated parameters using the quasi-Newton method to minimize *RSS*. The Builder calculates standard deviations for all estimated parameters and any other desired values and calculates correlations among them.

Results

The results were quite robust in terms of sensitivity to the initial parameter values and are summarized in Table 1 for $w = 0$, 0.5, and 1.0. M is about 0.3 when $q_s = 1$ and $w \leq 1.0$; q_s decreased from 1.2 to 0.95 and M increased from 0.21 to 0.33 under the 2-phase approach when w increased from 0 to 1.0 (Table 1). The 2-phase approach resulted in slightly smaller *RSS* than fixing q_s at 1.0. As expected, $r1$ and $r2$ were always <1.0.

The observed relative abundances of recruit, postrecruit, and all legal crabs were fitted well for $w = 0$, 0.5, and 1.0 (Fig. 1). A higher w resulted in a slightly poorer fit of the observed relative abundances. Large positive bias in observed relative abundances occurred under all scenarios in 1982 and 1993. When q_s was fixed at 1, there was no consistent trend in absolute abundance estimates of all legal crabs as w was increased from 0 to 1.0 (Table 1). However, under the 2-phase approach, absolute abundance estimates increased as w was increased (Table 1).

The relationship between survey abundance and commercial CPUE is illustrated in Fig. 2. Because standardized CPUEs under different values of w were in different scales, we scaled each series to a maximum value of 1. Unlike survey abundance, there were no obvious temporal trends in nominal CPUE, except perhaps a slight decline from 1980 to 1987. The flat nominal CPUE was caused in part by changes in total potlifts (effort), which decreased 94% from 1982 to 1988 and then increased 800% from 1982 to 1996. An extremely high number of potlifts in 1982 and extremely low numbers in 1988 and 1986 caused the nominal CPUEs in these three years to differ greatly from the observed survey abundances. Abundance in 1982 was also overestimated by the trawl survey. After standardization, CPUE generally followed survey abundance trends, particularly for low values of w.

The *RSS*s were inverted dome-shaped functions of estimated natural mortality or survey catchability (Fig. 3). As expected, natural mortality was negatively correlated with survey catchability, and both survey and commercial fishery catchabilities (not shown) were positively correlated. The response surface of objective function values was quite flat; different combinations of M and q_s resulted in similar *RSS*s. Increasing w reduced q_s and increased M.

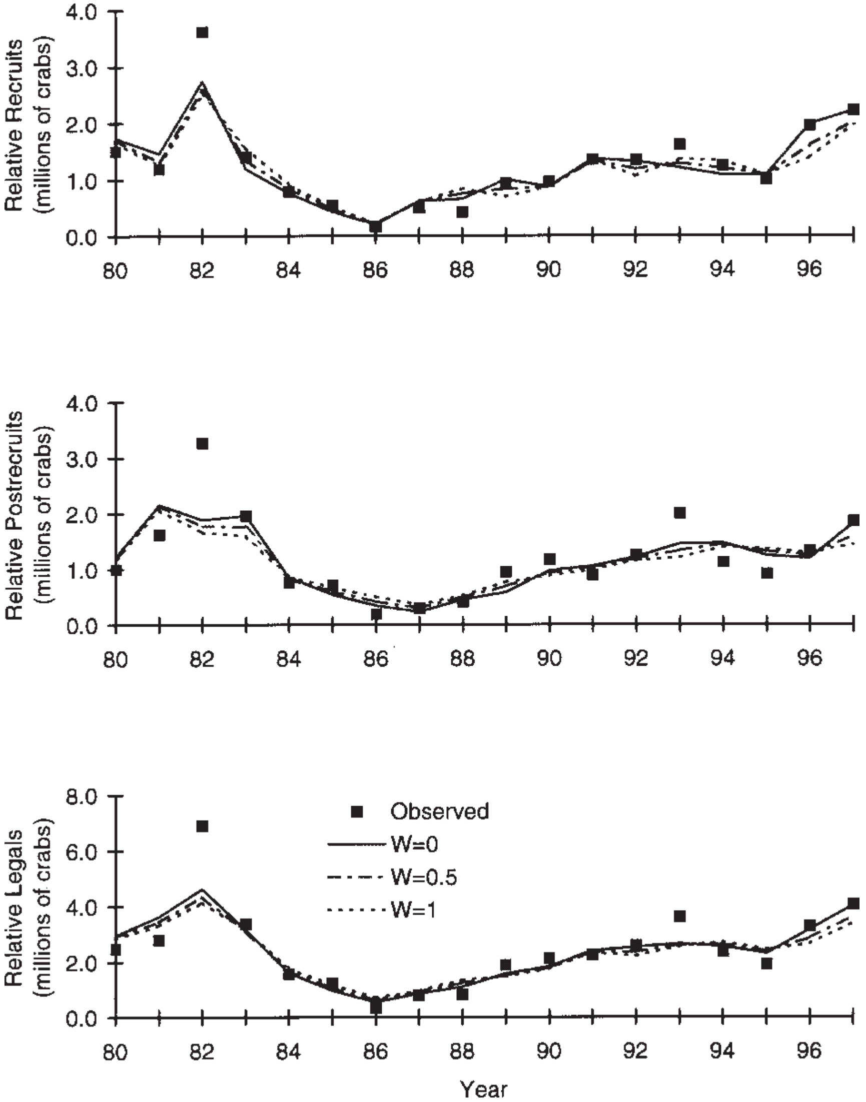


Figure 1. Comparisons of observed and estimated relative abundances of recruit, postrecruit, and legal blue king crabs in St. Matthew Island waters with q fixed at 1 and with three effort weighting factors w = 0, 0.5, and 1.0.

Table 1. **Estimates and associated standard deviations (SD) of parameters and absolute legal crab abundance (in millions) for a catch-survey analysis of the St. Matthew Island blue king crab stock under two approaches with effort weighting factor equal to 0, 0.5, and 1.**

	Weight = 0				Weight = 0.5				Weight = 1.0			
	Fixed $q = 1$		2-phase		Fixed $q = 1$		2-phase		Fixed $q = 1$		2-phase	
	Value	SD	Value	SD	Value	SD	Value	SD	Value	SD	Value	SD
M	0.302	0.138	0.207	0.287	0.299	0.134	0.267	0.283	0.307	0.133	0.328	0.288
q_s	1.000	NA	1.201	0.537	1.000	NA	1.070	0.551	1.000	NA	0.952	0.579
q_c	NA	NA	NA	NA	0.019	0.023	0.020	0.028	0.026	0.023	0.024	0.028
$r1$	NA	NA	NA	NA	0.811	0.322	0.808	0.323	0.723	0.243	0.725	0.244
$r2$	NA	NA	NA	NA	0.682	0.299	0.681	0.300	0.614	0.228	0.615	0.228
RSS	2.108	NA	2.033	NA	4.679	NA	4.671	NA	6.529	NA	6.525	NA
Legals:												
1980	2.946	1.172	2.387	1.551	2.902	1.093	2.687	1.883	2.836	1.026	3.002	2.356
1981	3.618	0.996	3.127	1.369	3.449	0.877	3.266	1.581	3.312	0.778	3.453	1.943
1982	4.628	1.057	4.135	1.403	4.352	0.981	4.166	1.653	4.150	0.918	4.299	2.103
1983	3.149	0.592	2.860	0.788	3.113	0.588	2.996	1.012	3.164	0.625	3.265	1.441
1984	1.599	0.345	1.417	0.466	1.668	0.376	1.587	0.679	1.772	0.411	1.845	1.019
1985	0.966	0.242	0.835	0.331	1.066	0.275	1.004	0.516	1.179	0.295	1.235	0.774
1986	0.542	0.162	0.454	0.221	0.627	0.185	0.585	0.350	0.712	0.193	0.749	0.513
1987	0.849	0.323	0.704	0.408	0.915	0.319	0.854	0.535	0.954	0.295	1.003	0.689
1988	1.096	0.368	0.911	0.502	1.244	0.389	1.160	0.717	1.357	0.387	1.431	1.013
1989	1.570	0.552	1.268	0.793	1.515	0.474	1.407	0.905	1.460	0.418	1.539	1.081
1990	1.834	0.580	1.515	0.848	1.790	0.523	1.673	0.992	1.752	0.480	1.841	1.226
1991	2.415	0.742	1.996	1.090	2.360	0.658	2.206	1.285	2.336	0.592	2.454	1.605
1992	2.534	0.777	2.054	1.220	2.362	0.673	2.197	1.366	2.210	0.581	2.326	1.590
1993	2.658	0.768	2.180	1.223	2.614	0.730	2.435	1.485	2.561	0.693	2.697	1.860
1994	2.550	0.720	2.124	1.110	2.639	0.731	2.466	1.442	2.721	0.744	2.862	1.931
1995	2.315	0.707	1.918	1.053	2.391	0.686	2.230	1.351	2.444	0.664	2.574	1.773
1996	3.166	1.137	2.605	1.549	2.852	0.892	2.657	1.666	2.661	0.757	2.803	1.948
1997	4.037	1.793	3.403	2.080	3.619	1.384	3.387	2.154	3.414	1.156	3.588	2.501

NA – not available.

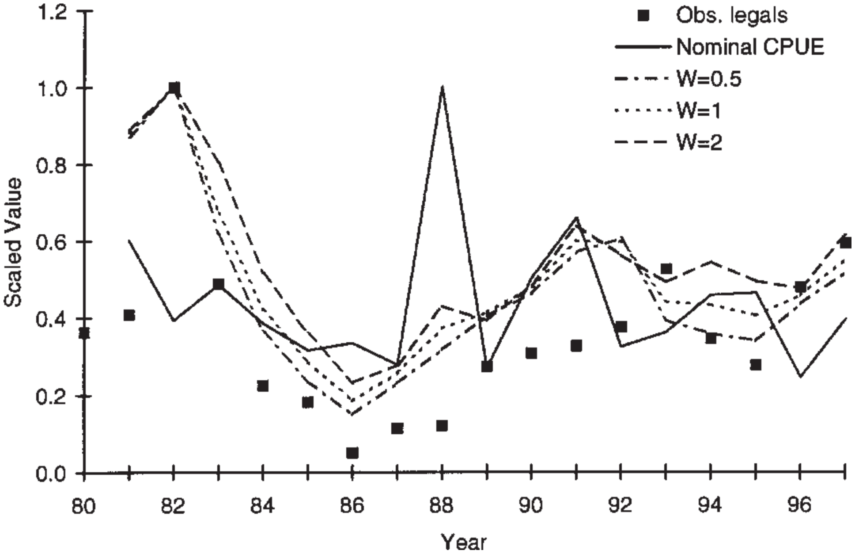


Figure 2. Comparisons of time series of observed legal crab abundance, observed nominal catch per unit effort (CPUE), and estimated standardized CPUE with three effort weighting factors w = 0.5, 1, and 2. Each time series was scaled to a maximum value of 1.

Absolute abundance and fishing effort estimates were affected by w, although their trends were similar over time with all five w values examined (Fig. 4). Abundance estimates were similar with $w = 0$ and 0.5 and generally were the highest with $w = 2.0$. Abundance estimates were smoother with $w = 0.5$ and 1.0 than with other w values and were least smooth with $w = 5.0$. Estimated potlifts with $w = 5.0$ were similar to the observed values, and as expected, a decrease in w caused further departure of estimated values from the observed values.

Estimates of absolute abundance and harvest rates depend on q_s and M. When weighting the effort data half as much as survey abundance and fixing survey catchability at 1.0, harvest rates increased from 0.01 in 1980 to 0.62 in 1983, declined steadily to 0.16 in 1989, and fluctuated between 0.22 and 0.32 from 1990 to 1997 (Fig. 5). Catch in 1997 was more than 3 times as high as catch in 1989 because of increased crab abundance.

Discussion

Traditional stock assessment models assume a unit stock without significant internal spatial structure by sex, size, or age. This assumption is reasonable if animals are distributed relatively uniformly across the space

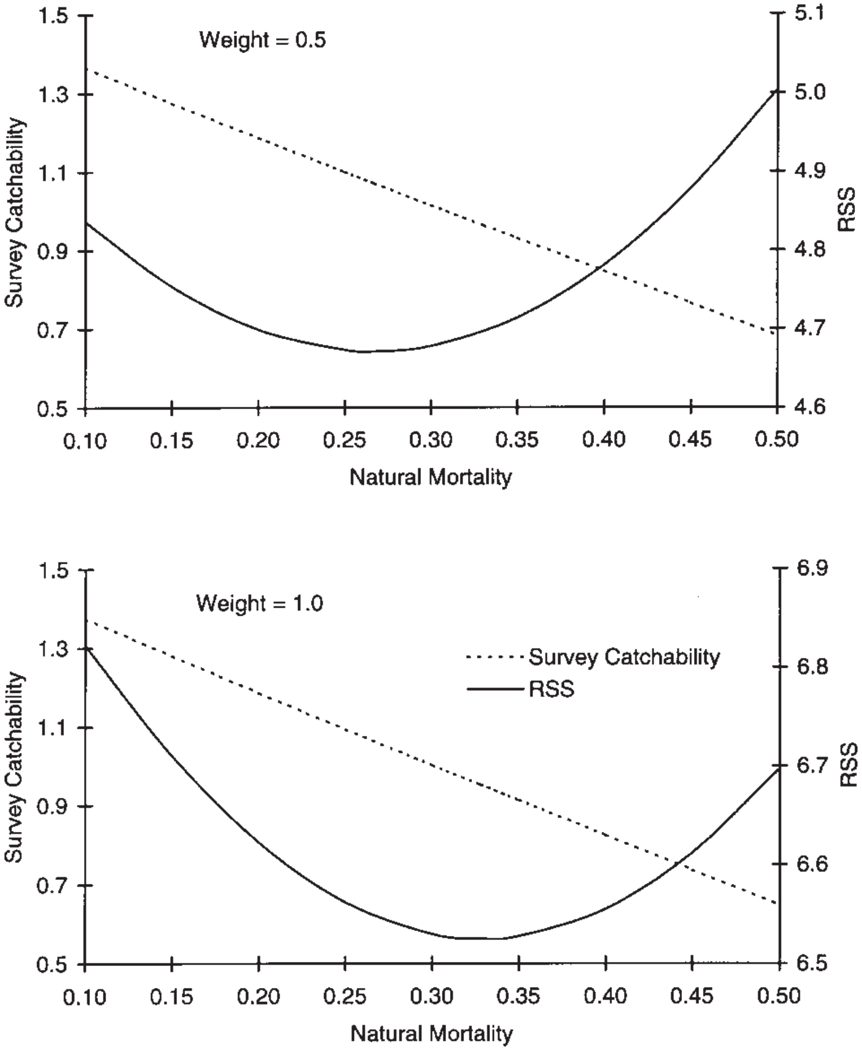


Figure 3. Relationships among survey catchability, natural mortality, and residual sum of squares (RSS) for the St. Matthew Island blue king crab stock from the catch-survey analysis using effort weighting factors w = 0.5 and 1.

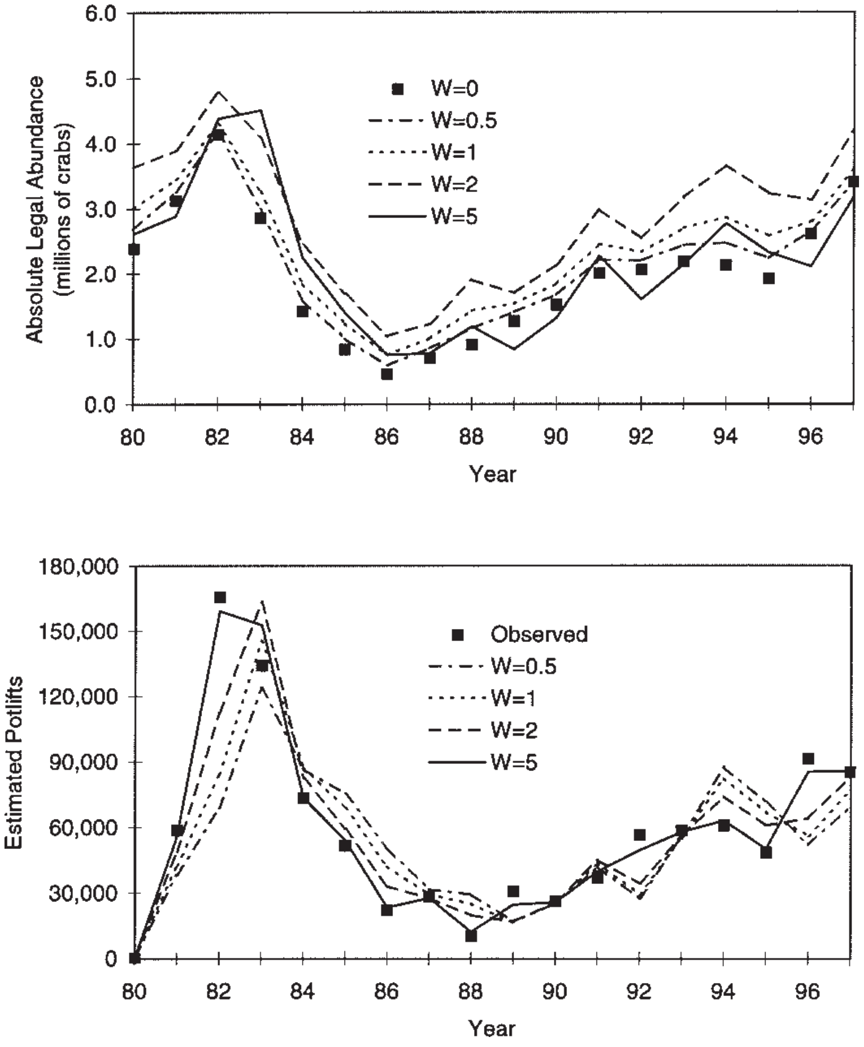


Figure 4. Comparisons of time series of absolute legal crab abundance estimates and potlifts under the 2-phase approach with effort weighting factors w = 0, 0.5, 1, 2, and 5.

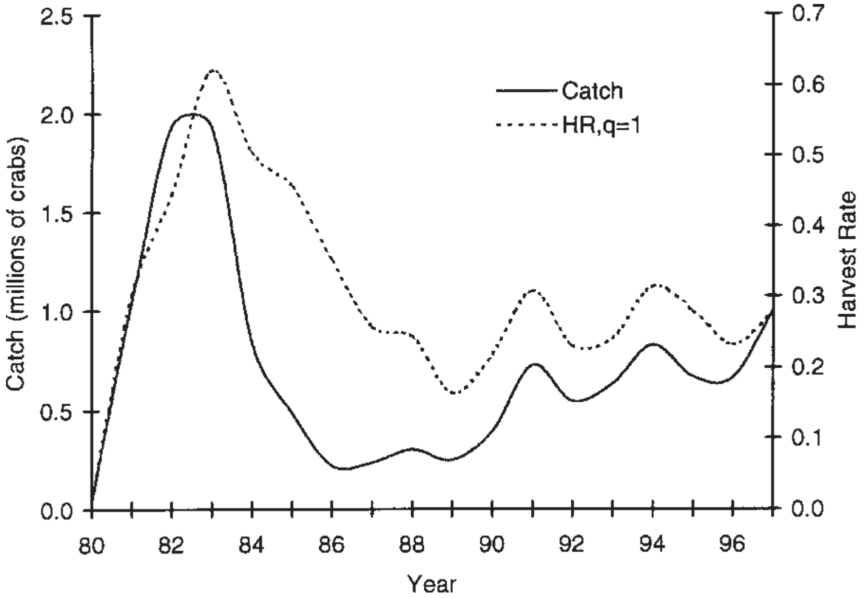


Figure 5. Time series of catches (solid line) and legal male harvest rates (dotted line) for the St. Matthew Island blue king crab stock with q fixed at 1 and w = 0.5.

they occupy, or if survey and commercial fishing effort equally represent all occupied habitats and the gear is not selective. Under such conditions, gear catchability is expected to be independent of exploitable stock abundance. Schooling behavior, common in many pelagic fish stocks (e.g., MacCall 1976), violates this assumption. Schooling creates sharp density differences in different habitats, and commercial fishing effort typically concentrates in areas of high density, causing a nonlinear relationship between commercial fishing catchability and exploitable stock abundance (MacCall 1976). The traditional stock assessment assumption may also be violated when areas of survey and commercial fishing do not completely overlap, such as when the survey or fishery is concentrated in a small part of the stock's range. For such stocks, trends of survey abundance and commercial CPUE may conflict, both of which may differ somewhat from true population changes. Temporary local depletions may also occur in areas where the commercial fishing effort concentrates.

In the case of St. Matthew blue king crabs, even though the trawl survey approximately covers the area in which crabs occur, each 1-nautical mile tow is expanded to a very large area (20 × 20 nm^2), and few tows are located close to the island where the commercial fishing fleet concentrates

(Pers. comm., D. Pengilly, Alaska Department of Fish and Game). Because the population is small and fishing effort is concentrated in a small area, CPUE is strongly affected by the total number of potlifts or fishing effort, which is a function of season length, pot limits, and number of vessels. A longer season or more vessels will effect a lower nominal CPUE because either vessels continue to fish crabs in the local area with a significantly reduced density or localized depletions cause vessels to fish less productive waters. We dealt with these problems by taking a power transformation of fishing effort. After the transformation, the trends of standardized CPUE and survey abundance were similar, even after weighting the effort 5 times heavier than the survey abundance.

Because of incomplete overlap between the survey area and commercial fishing area and uncertainty in the natural mortality estimate, we have low confidence in the estimate of survey catchability. Our earlier CSA model without commercial fishing effort (Zheng et al. 1997), which is analogous to the current model with a weighting factor of 0 on fishing effort, estimated survey catchability to be >1 under a biological feasible range of natural mortality. However, with our current analysis under the two-phase approach, when effort was weighted as heavy as or heavier than survey abundance, the survey catchability estimate is <1, which suggests the commercial fishing area has higher crab density than the survey area. This is supported by the results of a pot survey on St. Matthew blue king crabs in 1995 (Blau 1996). High natural mortality also results in estimates of survey catchability <1. Natural mortality of 0.3 or higher would generate a survey catchability estimate for legal crabs close to or <1 when the ratio of recruit to postrecruit catchabilities is fixed at 1 and the effort weighting factor is 0. From a mark-recapture study, natural mortality for St. Matthew blue king crabs was estimated at 0.78 to 0.97 for crabs 120-134 mm CL (Otto and Cummiskey 1990). These estimates of natural mortality appear to be too high; the return rate of tagged crabs may have been underestimated because of tag loss, underreporting, and migration. By comparison, overall mean natural mortality for legal (≥135 mm CL) red king crabs in Bristol Bay of the eastern Bering Sea was estimated to be about 0.3 (Zheng et al. 1995a, b). If natural mortality is indeed as high as that estimated by the tagging experiments, the survey catchability will be much smaller than 1. If the survey catchability of blue king crabs is indeed less than 1, then absolute abundance is probably underestimated.

Another uncertainty of survey catchability estimates is whether recruit and postrecruit crabs are subject to the same catchability. This uncertainty cannot be resolved by the current data. For simplicity, we assumed they have the same survey catchability in this study. Generally, this uncertainty does not affect abundance estimates very much because a higher or lower recruit catchability is offset by a lower or higher postrecruit catchability. However, different survey catchabilities influence natural mortality estimates greatly; a higher recruit catchability results in a lower natural mortality estimate (Zheng et al. 1997).

Survey catchability >1 would result if higher density of legal crabs occurred in trawlable habitats than in untrawlable habitats. A large bias in catchability could also occur for the CSA model if errors in shell aging were large and there was a difference of magnitude between recruits and postrecruits, or if aging errors were very asymmetric (Collie and Kruse 1998). If the survey catchability is underestimated, then absolute abundance estimates of blue king crabs are inflated and harvest rates are higher than managers now assume. Fortunately, much lower harvest rates were used during the last 10 years than during the early 1980s. Based on our upper-end estimate of survey catchability of 1.2, actual harvest rates realized from the St. Matthew Island fishery were still ≤37% of the legal male crabs during 1987-1997. However, the harvest rates on the targeted portion of the stock were likely much higher than harvest rates on the surveyed stock. Although the harvest rates on the surveyed stock are still below the maximum legal male rate of 60% (Pengilly and Schmidt 1995), precaution is advisable because of the possibility that survey catchability is greater than 1.

Catchability of the trawl survey, used to estimate abundance and a subsequent harvest level, has implications for management. If the current harvest strategy is optimal for the current population abundance estimated by area-swept methods, then lower population estimates corresponding to a survey catchability estimate >1 will result in underharvest of the stock. Conversely, higher population estimates as a result of a survey catchability estimate <1 will lead to overharvest. Unfortunately, only one pot survey has been conducted in a limited area near St. Matthew Island (Blau 1996). Multiple years of tagging studies and pot surveys in both untrawlable and trawlable habitats would provide information germane to estimating natural mortality and trawl survey catchability. The harvest strategy needs formal reevaluation when survey catchability is better known. However, to avoid overfishing, it may be prudent to act sooner and consider modifications in harvest strategy that would take into account uncertainty in the assessments.

Concentration of commercial fishing effort in small areas has additional implications for management. Because the catch quota is currently based on the whole survey area, fishing effort concentrated in small areas results in higher actual harvest rates on the targeted portion of the stock than would be the case if effort was spread over the entire stock. Furthermore, bycatch of female and juvenile crabs is probably higher in the untrawlable habitat, where the fishery concentrates, than further offshore. Unfortunately, crab movements between the summer survey and the fall fishery are unknown, as are timing variations in possible biennial mating and molting cycle of large females (Jensen and Armstrong 1989). Information on the spatial and temporal structures and migration patterns of the stock, once available, needs to be incorporated into the harvest strategy.

Use of commercial fishery CPUE as a measure of population abundance is well known to be problematic (e.g., Hilborn and Walters 1992).

Gear developments, improvements in vessel positioning equipment, and regulations affecting harvesting efficiency can have significant impacts on CPUE independent of crab population size. Although we have not adjusted CPUE for such factors, except for the institution of pot limits in 1993, we believe that there is merit to including commercial CPUE data in the annual assessment for St. Matthew blue king crabs because of the disparity in survey and fishing locations. In addition, expanding the 2-stage model to include a third stage for prerecruit abundance (crabs with CL one growth increment below legal size) will improve estimates of legal crab abundance in the terminal year. A 3-stage model is currently being developed. However, further meaningful improvements in the assessment of the stock will probably have to wait for better information on seasonal migrations, natural mortality, and crab distributions among trawlable and untrawlable habitats.

Acknowledgments

This is contribution PP-163 of the Alaska Department of Fish and Game, Commercial Fisheries Management and Development Division, Juneau.

References

Blau, S.F. 1996. The 1995 St. Matthew Island blue king crab survey. Alaska Dept. Fish and Game, Regional Information Report 4K96-27, Kodiak. 36 pp.

Collie, J.S., and G.H. Kruse. 1998. Estimating king crab *(Paralithodes camtschaticus)* abundance from commercial catch and research survey data. In: G.S. Jamieson and A. Campbell (eds.), Proceedings of the North Pacific Symposium on Invertebrate Stock Assessment and Management. Can. Spec. Publ. Fish. Aquat. Sci. 125:73-83.

Hilborn, R., and C.J. Walters. 1992. Quantitative fisheries stock assessment: Choice, dynamics and uncertainty. Chapman and Hall, Inc., New York. 570 pp.

Jensen, G.C., and D.A. Armstrong. 1989. Biennial reproductive cycle of blue king crab, *Paralithodes platypus*, at the Pribilof Islands, Alaska and comparison to a cogener, *P. camtschatica*. Can. J. Fish. Aquat. Sci. 46:932-940.

Kruse, G.H., and J.S. Collie. 1991. Preliminary application of a population size estimation model to the Bristol Bay stock of red king crabs. Alaska Dept. Fish and Game, Regional Information Report 5J91-09, Juneau. 25 pp.

MacCall, A.D. 1976. Density dependence of catchability coefficient in the California Pacific sardine, *Sardinops sagax caerulea*, purse seine fishery. Calif. Coop. Ocean. Fish. Investig. Rep. 18:136-148.

Morrison, R., R.K. Gish, and N.H. Sagalkin. 1997. Annual management report for the shellfish fisheries of the Bering Sea. In: Annual management report for the shellfish fisheries of the westward region, 1996. Alaska Dept. Fish and Game, Regional Information Report 4K97-41, Kodiak, pp. 120-210.

Otter Research Ltd. 1994. An introduction to AD Model Builder: For use in nonlinear modeling and statistics. Otter Research Ltd., Nanaimo, B.C. 51 pp.

Otto, R.S. 1990. An overview of eastern Bering Sea king and Tanner crab fisheries. In: Proceedings of the International Symposium on King and Tanner Crabs. University of Alaska Sea Grant, AK-SG-90-04, Fairbanks, pp. 9-26.

Otto, R.S., and P.A. Cummiskey. 1990. Growth of adult male blue crab *(Paralithodes platypus).* In: Proceedings of the International Symposium on King and Tanner Crabs. University of Alaska Sea Grant, AK-SG-90-04, Fairbanks, pp. 245-258.

Otto, B.G., J.A. Haaga, and R.A. MacIntosh. 1997. Report to industry on the 1996 eastern Bering Sea crab survey. NMFS, Alaska Fisheries Science Center Processed Report 97-02, Kodiak. 52 pp.

Pengilly, D., and D. Schmidt. 1995. Harvest strategy for Kodiak and Bristol Bay red king crab and St. Matthew Island and Pribilof blue king crab. Alaska Dept. Fish and Game, Special Publication 7, Juneau. 10 pp.

Somerton, D.A., and R.A. MacIntosh. 1985. Reproductive biology of the female blue king crab *Paralithodes platypus* near the Pribilof Islands, Alaska. J. Crust. Bio. 5(3):365-376.

Zheng, J., G.H. Kruse, and M.C. Murphy. 1996. Comparisons of abundance estimation methods for red king crabs in Bristol Bay and Kodiak. In: High latitude crabs: Biology, management and economics. University of Alaska Sea Grant, 96-02, Fairbanks, pp. 283-294.

Zheng, J., M.C. Murphy, and G.H. Kruse. 1995a. A length-based population model and stock-recruitment relationships for red king crab, *Paralithodes camtschaticus*, in Bristol Bay, Alaska. Can. J. Fish. Aquat. Sci. 52:1229-1246.

Zheng, J., M.C. Murphy, and G.H. Kruse. 1995b. An update of the length-based population model and stock-recruitment relationships for red king crab, *Paralithodes camtschaticus*, in Bristol Bay, Alaska. Alaska Fishery Research Bulletin 2(2):114-124.

Zheng, J., M.C. Murphy, and G.H. Kruse. 1997. Application of a catch-survey analysis to blue king crab stocks near Pribilof and St. Matthew Islands. Alaska Fishery Research Bulletin 4(1):62-74.

Using Multiple Data Sets to Assess Red King Crab in Norton Sound, Alaska: Length-Based Stock Synthesis

Jie Zheng and Gordon H. Kruse
Alaska Department of Fish and Game, Juneau, Alaska

Lowell Fair
Alaska Department of Fish and Game, Anchorage, Alaska

Abstract

Red king crab, *Paralithodes camtschaticus*, in Norton Sound, Alaska, support three fisheries: summer commercial, winter commercial, and winter subsistence fisheries. Four types of surveys have been conducted periodically during the last two decades: summer trawl, summer pot, winter pot, and preseason summer pot, but none of these surveys were conducted every year. To improve abundance estimates, we developed a length-based stock synthesis model of male crab abundance that combines multiple sources of survey, catch, and mark-recovery data from 1976 to 1996. A maximum likelihood approach was used to estimate abundance, recruitment, and catchabilities of the commercial pot gear. The model yielded a time series of abundance that is smooth and consistent over time. Estimates of parameters and legal crab abundance are not very sensitive to weighting factors for survey abundances and fishing effort, and maximum effective sample size. Assumed natural mortality influenced recruitment estimates, but had limited impact on estimates of legal abundances. The model can be used to improve fisheries management by reducing annual variations in catch and the risk of overfishing.

Introduction

Red king crab, *Paralithodes camtschaticus*, in Norton Sound, Alaska, support two small, but locally important commercial fisheries, summer and winter, and a subsistence fishery. The commercial fisheries began in 1977.

The summer commercial catch peaked in 1979 at 970,962 crabs and has been under 130,000 crabs since 1987 (Lean and Brennan 1997). The winter commercial catch has not exceeded 10,000 crabs and was less than 2,000 crabs during most years. The subsistence fishery mainly occurs during winter, and the annual catch generally ranges from 1,000 to 12,000 crabs (Lean and Brennan 1997). Although annual ex-vessel values for the commercial catch were under US$1.0 million during the last 15 years (Lean and Brennan 1997), Norton Sound red king crab provide an important economic opportunity and food source to local communities.

Composition of the fishing fleet has changed over time. During most years prior to 1993 the fleet consisted of a few catcher/processors (large vessels) but in some years the fleet consisted of up to 24 catcher and catcher/processor vessels (C. Lean, ADFG, personal communication). Due to the establishment of a superexclusive registration area (i.e., if a vessel participates the Norton Sound red king crab fishery, the vessel cannot participate in any other red king crab fisheries in Alaska in the same year) and a pot limit, large fishing vessels have been replaced by small fishing vessels since 1993. The pot limit started in 1992 with 100 per vessel. Since 1993, vessels longer than 125 ft (38.1 m) have been limited to 50 pots and smaller vessels to 40 pots per vessel.

The Norton Sound summer commercial red king crab fishery is currently managed with an exploitation rate ≤10%, which is half of rates common in other Bering Sea commercial king crab fisheries (5 AAC [Alaska Administrative Code] 34.915). Estimates of annual abundance of legal crabs (≥104 mm carapace length [CL]) are used to set the summer catch quota each year. Because of low catch rates, catch levels for the winter commercial and subsistence fisheries are not limited.

To assess abundance, four types of surveys have been conducted periodically during the last two decades by the National Marine Fisheries Service (NMFS) and Alaska Department of Fish and Game (ADFG): summer trawl, summer pot, winter pot, and preseason summer pot. No individual survey type, however, was conducted annually on this stock. Estimated abundances fluctuate greatly from survey to survey because of changes in stock size, survey methods, and annual measurement errors. Survey measurement errors were large; for example, in 1979 commercial catch in the summer fishery exceeded estimated abundance of harvestable (legal) male crabs from the trawl survey. To help distinguish true changes in stock size from measurement errors, we constructed a length-based population model that synthesizes multiple sources of survey, fisheries, and mark-recovery data and links population abundances by length classes in adjacent years. The model is similar to length-based models developed by Zheng et al. (1995) for crab populations. A maximum likelihood approach was used to estimate male abundances, recruitment, molting probabilities, catchabilities of the commercial pot gear, and selectivities of different survey and commercial fishing gears.

Table 1. Summary of available data for Norton Sound male red king crab.

Data set	Years	Data types
Summer trawl survey	79,82,85,88,91,96	Abundance and proportion by length and shell condition
Summer pot survey	80-82,85	Abundance and proportion by length and shell condition
Winter pot survey	81-86,88-90,92,94-96	Proportion by length and shell condition
Summer preseason survey	95	Proportion by length and shell condition
Summer commercial fishery	76-90,92-96	Catch, effort, and proportion by length and shell condition
Observer data	87-90,92,94	Proportion by length and shell condition
Winter commercial fishery	76-96	Catch
Subsistence fishery	76-96	Catch
Tagging data	80-91	Mean and standard deviation of growth increment

Methods

Data

Available data are summarized in Table 1. NMFS conducted trawl surveys every three years from 1976 to 1991 (Stevens and MacIntosh 1986), and ADFG conducted a trawl survey in 1996 (Fair 1997). Total population abundances and length and shell compositions for males >73 mm CL were estimated by "area-swept" methods from the trawl survey data (Alverson and Pereyra 1969). The compositions consisted of six 10-mm length classes and two shell conditions (new and old shells). If multiple hauls were conducted for a single station (10×10 nmi) during a survey, then the average of abundances from all hauls within the station was used. If a station was not surveyed during a given year, its abundance was assumed to be equal to the average of abundances in the four adjacent survey stations. If more than two adjacent stations were not surveyed, the abundance of the non-surveyed station was assumed as zero. Because the non-surveyed stations are located in areas with very low density of red king crabs, esti-

Table 2. Means and standard deviations (SD) of growth increments per molt and growth matrix (proportion of crabs molting from a given premolt carapace length range into postmolt length ranges) for Norton Sound male red king crab.

Premolt length		Growth increment		Postmolt length range					
Class	Range	Mean	SD	74-83	84-93	94-103	104-113	114-123	>123
1	74-83	14.50	3.34	0.01	0.54	0.45	0	0	0
2	84-93	14.50	3.34	0	0.01	0.54	0.45	0	0
3	94-103	14.09	2.69	0	0	0.01	0.58	0.41	0
4	104-113	13.35	2.80	0	0	0	0.01	0.65	0.35
5	114-123	11.35	2.19	0	0	0	0	0.03	0.97
6	>123	11.35	2.19	0	0	0	0	0	1.00

Length is measured as mm CL. Results are derived from mark-recapture data from 1980 to 1991.

mates of total crab abundance were little affected by these assumptions. The complete data file from the survey in 1979 could not be located, so we used the published population estimate for that year (Lean and Brennan 1997) and estimated length and shell condition compositions from the partial survey file. Some trawl surveys occurred during September, the molting period. To make survey abundances comparable with premolt abundances, we adjusted trawl survey abundances by subtracting average growth increment of each length class (Table 2) from the length of each soft-shell crab (molting within the past two months).

Four summer pot surveys were conducted by ADFG (Table 1), and total male crab abundances were estimated using Petersen mark-and-recapture methods (Brannian 1987). ADFG also conducted 13 winter pot surveys and one preseason pot survey in the summer of 1995 (Table 1); total crab abundances were not estimated for these pot surveys because of unreliable catch per unit effort (CPUE) data due to change in environmental conditions over time and lack of tagging data. For all pot surveys, length and shell condition compositions were estimated.

Red king crab catches from the summer fishery were sampled by ADFG from 1976 to 1996 to determine length and shell condition. Bycatch of sublegal males (observer data) from the summer fishery in 1987-1990, 1992, and 1994 were also sampled by observers to determine length and shell condition. Total catch from all fisheries and effort (potlifts) from the summer fishery were obtained from Lean and Brennan (1997). Red king crabs were tagged and released during 1980-1991 (Powell et al. 1983, Brannian 1987); 222 tagged male crabs were recovered after spending at least one molting season at liberty. These tagging data were used to estimate a growth matrix and molting probabilities by premolt length.

Population Model

The summer fishery for Norton Sound red king crab usually occurs during July and August, and crabs mostly molt after August. Thus, we modeled male summer population abundance during July and winter population abundance in January. Because few crabs <74 mm CL were caught during surveys or fisheries and sample sizes for trawl and winter pot surveys were relatively small, we modeled abundances for only male crabs ≥74 mm CL at 10-mm length intervals.

Observations on shell condition are commonly used to classify crabs into two categories: newshell crabs, those that molted within the past year, and oldshell crabs, those that did not. We modeled newshell and oldshell male crabs separately but assumed they have the same molting probability and natural mortality. Summer crab abundances are the survivors of crabs from the previous winter:

$$\begin{aligned} N_{s,l,t+1} &= (N_{w,l,t} - C_{w,t}P_{w,n,l,t} - C_{p,t}P_{p,n,l,t})e^{-0.5M_l}, \\ O_{s,l,t+1} &= (O_{w,l,t} - C_{w,t}P_{w,o,l,t} - C_{p,t}P_{p,o,l,t})e^{-0.5M_l}, \end{aligned} \tag{1}$$

where $N_{s,l,t}$ and $O_{s,l,t}$ are summer abundances of newshell and oldshell crabs in length class l in year t, $N_{w,l,t}$ and $O_{w,l,t}$ are winter abundances of newshell and oldshell crabs in length class l in year t, $C_{w,t}$ and $C_{p,t}$ are total winter and subsistence catches in year t, $P_{w,n,l,t}$ and $P_{p,n,l,t}$ are length compositions of winter and subsistence catches for newshell crabs in length class l in year t, $P_{w,o,l,t}$ and $P_{p,o,l,t}$ are length compositions of winter and subsistence catches for oldshell crabs in length class l in year t, and M_l is instantaneous natural mortality in length class l, which, for simplicity, we assumed to be constant (M) for all sizes and shell conditions except for the last length class where $M_6 = 1.6M$. Higher natural mortality for the last length class was justified by the sharp decrease of observed length frequency, and factor 1.6 resulted in a good fit of the data.

Winter abundance of newshell crabs is the combined result of growth, molting probability, mortality, and recruitment from the summer population:

$$N_{w,l,t} = \sum_{l'=1}^{l'=l} \left\{ G_{l',l} \left[N_{s,l',t} + O_{s,l',t} - C_{s,t}\left(P_{s,n,l',t} + P_{s,o,l',t}\right) \right] m_{l',t} e^{-0.5M_l} \right\} + R_{l,t}, \tag{2}$$

where $G_{l',l}$ is a growth matrix representing the expected proportion of crabs molting from length class l' to length class l, $C_{s,t}$ is total summer catch in year t, $P_{s,n,l,t}$ and $P_{s,o,l,t}$ are length compositions of summer catch for newshell and oldshell crabs in length class l in year t, $m_{l,t}$ is molting probability in length class l in year t, and $R_{l,t}$ is recruitment into length class l in year t. Winter abundance of oldshell crabs is the non-molting portion of survivors of crabs from summer:

$$O_{w,l,t} = \left[N_{s,l,t} + O_{s,l,t} - C_{s,t}(P_{s,n,l,t} + P_{s,o,l,t})\right](1 - m_{l,t})e^{-0.5M_l}. \tag{3}$$

Males >123 mm CL were grouped together to form the last length class. Sublegal males (<104 mm CL) are not legally retained in the commercial catch but are sorted, discarded, and subjected to handling mortality. Because of lack of data and complexity, we did not model handling mortality. Instead for simplicity we considered handling mortality to be included in the M estimates.

Following Balsiger's (1974) findings, we used a reverse logistic function to fit molting probabilities as a function of length and time:

$$m_{l,t} = \left(1 - \frac{1}{1 + e^{-\alpha_t(\iota - \beta_t)}}\right)\delta_t, \tag{4}$$

where α_t, β_t, and δ_t are parameters, and ι is the mean length of length class l. The sample size for the mark-recapture data is too small to estimate annual molting probabilities. However, residuals of length and shell compositions under a single molting probability function suggest molting probabilities for Norton Sound red king crab were different during three periods: 1976-1981, 1982-1990, and 1991-1997. Accordingly, three logistic functions were used to describe molting probabilities during different periods, resulting in three sets of parameters: α_1 and β_1, α_2 and β_2, and α_3 and β_3. δ_t was set to 1 for all years except in 1993 when it was set to 0.35 to closely fit the low molting probability. The unusually long ice coverage in 1993 may have caused most crabs to skip molt.

We modeled recruitment, R_t, as a stochastic process about the median, R_0:

$$R_t = R_0 e^{\tau_t}, \tau_t \sim N(0, \sigma_R^2). \tag{5}$$

R_t was assumed only to enter length classes 1 and 2 (Table 2); thus, $R_{l,t} = 0$ when $l \geq 3$. The recruits belonging to the first two length classes are:

$$R_{l,t} = rR_t, R_{2,t} = (1 - r)R_t, \tag{6}$$

where r is a parameter ≤ 1.

Estimated length and shell compositions of winter commercial catch were derived from the winter population, winter selectivity for pot, and proportion of legal crabs for each length class:

$$P_{w,n,l,t} = \frac{N_{w,l,t} S_{w,l} L_l}{\sum_l \left[(N_{w,l,t} + O_{w,l,t}\mu) S_{w,l} L_l \right]},$$
$$P_{w,o,l,t} = \frac{O_{w,l,t} \mu S_{w,l} L_l}{\sum_l \left[(N_{w,l,t} + O_{w,l,t}\mu) S_{w,l} L_l \right]}, \tag{7}$$

where L_l is proportion of legal crabs for length class l estimated from the observer data, $S_{w,l}$ is winter selectivity for pot for length class l, and μ is an additional selectivity parameter for oldshell crabs. Based on winter pot survey data, winter selectivities for length classes 3-5 were assumed to be 1, $S_{w,1}$, $S_{w,2}$, $S_{w,6}$, and μ were estimated as parameters, and μ was greater than one.

The subsistence fishery does not have a size limit, but crabs smaller than length class 3 are generally not retained. So, we estimated length compositions of subsistence catch when $l > 2$ as follows:

$$P_{p,n,l,t} = \frac{N_{w,l,t} S_{w,l}}{\sum_l \left[(N_{w,l,t} + O_{w,l,t}\mu) S_{w,l} \right]},$$
$$P_{p,o,l,t} = \frac{O_{w,l,t} \mu S_{w,l}}{\sum_l \left[(N_{w,l,t} + O_{w,l,t}\mu) S_{w,l} \right]}. \tag{8}$$

Estimated length compositions of winter pot survey for newshell and oldshell crabs $P_{sw,n,l,t}$ and $P_{sw,o,l,t}$ were also based on equation (7) except that $l \geq 1$.

Estimated length and shell compositions of summer commercial catch were based on summer population, selectivity, and legal abundance:

$$P_{s,n,l,t} = \frac{N_{s,l,t} S_{s,l} L_l}{A_t},$$
$$P_{s,o,l,t} = \frac{O_{s,l,t} v_l S_{s,l} L_l}{A_t}, \tag{9}$$

where v_l are additional selectivity for oldshell crabs, $S_{s,l}$ is pot selectivity for the summer commercial fishery, and A_t is exploitable legal abundance in year t. Based on catch data, v_l were equal to 1 except the last length class, v_6, which was estimated as a parameter (<1), and $S_{s,l}$ was described by a logistic function with parameters ϕ and ω:

$$S_{s,l} = \frac{1}{1+e^{-\phi(\iota-\omega)}}. \tag{10}$$

$S_{s,l}$ was scaled such that $S_{s,5} = 1$ and $S_{s,6} \leq 1$. Exploitable abundance was estimated as

$$A_t = \sum_l \left[(N_{s,l,t} + O_{s,l,t}v_l)S_{s,l}L_l\right]. \tag{11}$$

Summer fishing effort (f_t, measured as the number of potlifts) was estimated as total summer catch, C_t, divided by the product of catchability q and mean exploitable abundance such that

$$f_t = \frac{C_t}{\left[q(A_t - 0.5C_t)\right]}. \tag{12}$$

Because of the change in the fishing fleet and pot limit in 1993, q was replaced by q_1 for fishing efforts before 1993 and by q_2 after 1992. Estimated length and shell compositions of bycatch were:

$$\begin{aligned} P_{b,n,l,t} &= \frac{N_{s,l,t}S_{s,l}(1-L_l)}{\sum_l \left[(N_{s,l,t} + O_{s,l,t})S_{s,l}(1-L_l)\right]}, \\ P_{b,o,l,t} &= \frac{O_{s,l,t}S_{s,l}(1-L_l)}{\sum_l \left[(N_{s,l,t} + O_{s,l,t})S_{s,l}(1-L_l)\right]}. \end{aligned} \tag{13}$$

Replacing v_l and L_t in equations (9) and (11) with 1 resulted in estimated length compositions of the summer preseason survey for both newshell and oldshell crabs, $P_{sf,n,l,t}$ and $P_{sf,o,l,t}$.

Estimated length and shell condition compositions of summer pot survey abundance were

$$\begin{aligned} P_{sp,n,l,t} &= \frac{N_{s,l,t}S_{sp,l}}{\sum_l \left[(N_{s,l,t} + O_{s,l,t})S_{sp,l}\right]}, \\ P_{sp,o,l,t} &= \frac{O_{s,l,t}S_{sp,l}}{\sum_l \left[(N_{s,l,t} + O_{s,l,t})S_{sp,l}\right]}. \end{aligned} \tag{14}$$

where $S_{sp,l} = 1$ when $l \geq 3$, and $S_{sp,1}$ and $S_{sp,2}$ were estimated as two parameters. Similarly, length and shell condition compositions of summer trawl survey abundance were estimated with selectivity $S_{st,l} = 1$ when $l \geq 3$, and $S_{st,1}$ and $S_{st,2}$ were two parameters. Because some trawl surveys occurred during the molting period, we combined the length compositions of newshell and oldshell crabs as one single shell condition, $P_{st,l,t}$.

Parameter Estimation

We used tagging data to estimate mean growth increment per molt and standard deviation for each premolt length class (Table 2). The growth matrix was derived from normal distributions generated with estimated mean growth increments per molt and standard deviations (Table 2). Observed growth increments per molt are approximately normally distributed.

A maximum likelihood approach was used to estimate 41 parameters, which include catchability, parameters for selectivities of survey and fishing gears and for molting probabilities, recruits each year except the last, and total abundance in the first year. Under assumptions that measurement errors of annual total survey abundance and summer commercial fishing effort estimates follow lognormal distributions and length and shell composition estimates from each data set have a multinomial error structure (Fournier and Archibald 1982, Methot 1989), the log-likelihood function is

$$\begin{aligned}
&\sum_{i=1}^{i=6}\sum_{t=1}^{t=n_i}\left\{K_{i,t}\sum_{l=1}^{l=6}\sum_{j=1}^{j=2}\left[\hat{P}_{i,j,l,t}\ln(P_{i,j,l,t}+\kappa)\right]\right\}\\
&-\sum_{i=1}^{i=2}\left\{W_i\sum_{k=1}^{k=2}\sum_{t=1}^{t=n_i}\left[\ln(\hat{B}_{i,k,t}+\kappa)-\ln(B_{i,k,t}+\kappa)\right]^2\right\}\\
&-W_f\sum_{t=1}^{t=22}\left[\ln(\hat{f}_t+\kappa)-\ln(f_t+\kappa)\right]^2-W_R\sum_{t=1}^{t=21}\tau_t^2,
\end{aligned}\tag{15}$$

where i stands for a data set (1 for summer trawl survey, 2 for summer pot survey, 3 for winter pot survey, 4 for summer preseason survey, 5 for summer fishery, and 6 for observer data during the summer fishery); j represents shell condition (1 for newshell and 2 for oldshell); n_i is the number of years in which data set i is available; $k = 1$ stands for legal crabs and $k = 2$ for non-legal crabs; $K_{i,t}$ is the effective sample size of length and shell compositions for data set i in year t; $\hat{P}_{i,j,l,t}$ and $P_{i,j,l,t}$ are observed and estimated length compositions for data set i, shell condition j, length class l, and year t; κ is a constant equal to 0.001; W_i is the weighting factor of annual total survey abundance for data set i; $\hat{B}_{i,k,t}$ and $B_{i,k,t}$ are observed and estimated annual total abundances for data set i and year t; W_f is the weighting factor of the summer fishing effort; $\hat{f}_t$ and f_t are observed and estimated summer fishing efforts; and W_R is the weighting factor of recruitment. We did not impose measurement error on total annual catch. Variances for total survey abundance and summer fishing effort estimates were not estimated; rather, we used weighting factors to reflect these variances.

Crabs usually aggregate and, thus, are not completely independently sampled. To reduce the non-independent effect, annual total sample sizes of length and shell compositions for all data sets were reduced by 50%. In addition, annual effective sample sizes were capped at 400 to avoid overweighting the data with a large sample size (Fournier and Archibald 1982). Weighting factors represent prior assumptions about the accuracy or the variances of the observed data or random variables. W_i was set as 200 for all survey abundances, W_f was set to be 100, or 50% of W_i, and W_R was set to be 0.01. According to the fishery manager (C. Lean, ADFG, personal communication), the CPUE in 1992 was not as reliable as in the other years. Therefore, we weighted the effort in 1992 half as much as in the other years. Sensitivity of estimated legal abundance to changes in W_i, W_f, and maximum effective sample size was investigated.

Because of confounding effects among natural mortality, catchability, and recruitment, we did not estimate M in the model. Based on Low (1991) and Kruse et al. (1996), M was assumed to be 0.3. We examined sensitivity of estimated legal abundance to different assumptions about M.

We estimated parameters with AD Model Builder (Otter Research Ltd. 1994) using the quasi-Newton method to minimize –1 times likelihood values. To reduce the number of parameters, we assumed that length and shell compositions, from the first year (1976) summer trawl survey data, approximate true compositions. Abundances by length and shell condition in all other years were computed recursively from abundances by length and shell condition in the first year and by annual recruitment, catch, and model parameters. Initial parameter estimates were an educated guess based on observation and current knowledge. The results were quite robust in terms of sensitivity to the initial parameter values; the final estimates are the same with the initial parameter values within ±30% of the educated guess.

Results

The model fit well to observed total and legal male trawl abundances except in 1979 when the trawl survey greatly underestimated the crab abundance (Fig. 1). Observed pot survey abundances were not fit as well as observed trawl survey abundances, especially for the first two surveys. Estimated fishing effort for the summer commercial fishery was very similar to, but smoother than, observed fishing effort in most years (Fig. 1). This close fit between the observed effort and the model effort, which is calculated from catch and abundance data, indicates that the CPUE of the summer commercial fishery is closely associated with the estimated legal abundance.

The residuals of length and shell compositions were generally large, except for the summer pot survey (Fig. 2, 3). The large residuals for the trawl survey are probably due to small sample sizes; all trawl surveys except in 1976 caught fewer than 200 legal crabs. The large residuals for

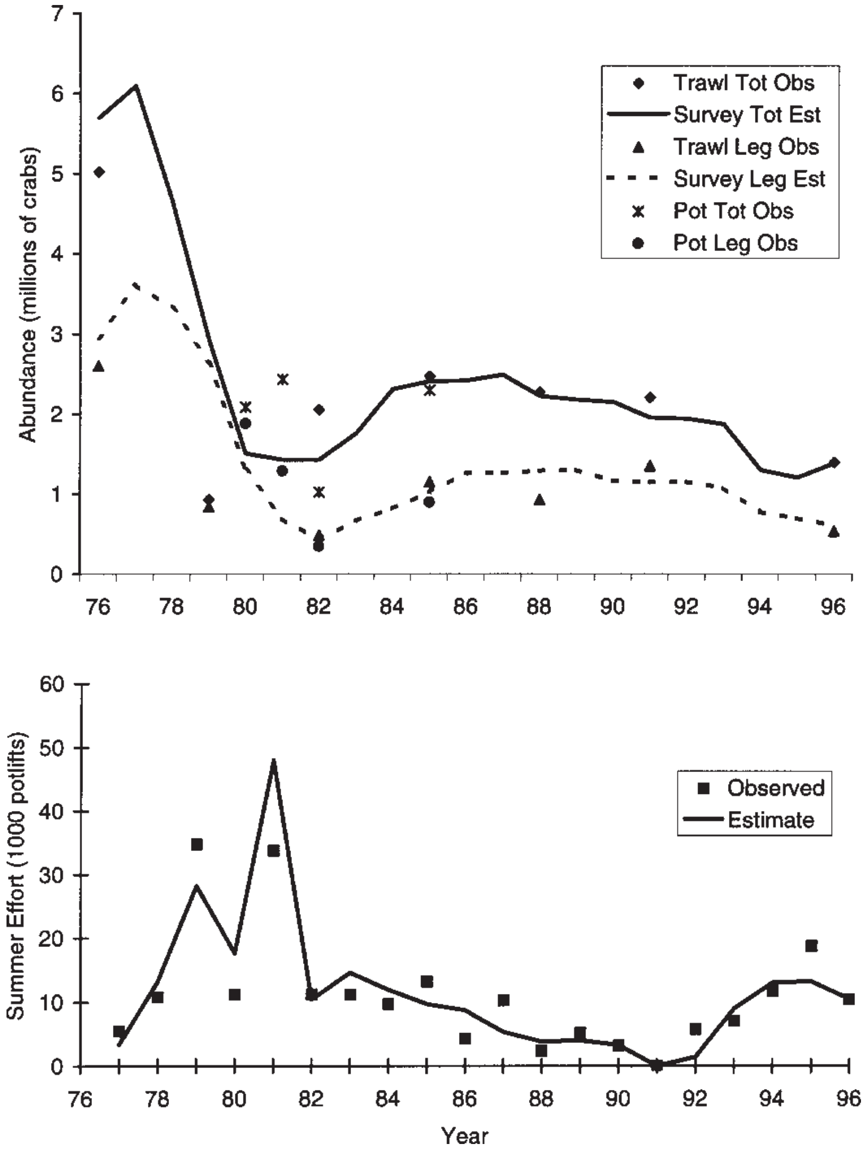


Figure 1. Comparison of observed and estimated Norton Sound red king crab abundances (legal and total males) by summer trawl and pot surveys (upper plot) and observed and estimated summer fishing efforts (lower plot). "Survey Tot Est" and "Survey Leg Est" are total and legal catchable male abundances estimated by the model, respectively. Catchable abundance is equal to population abundance times survey selectivities.

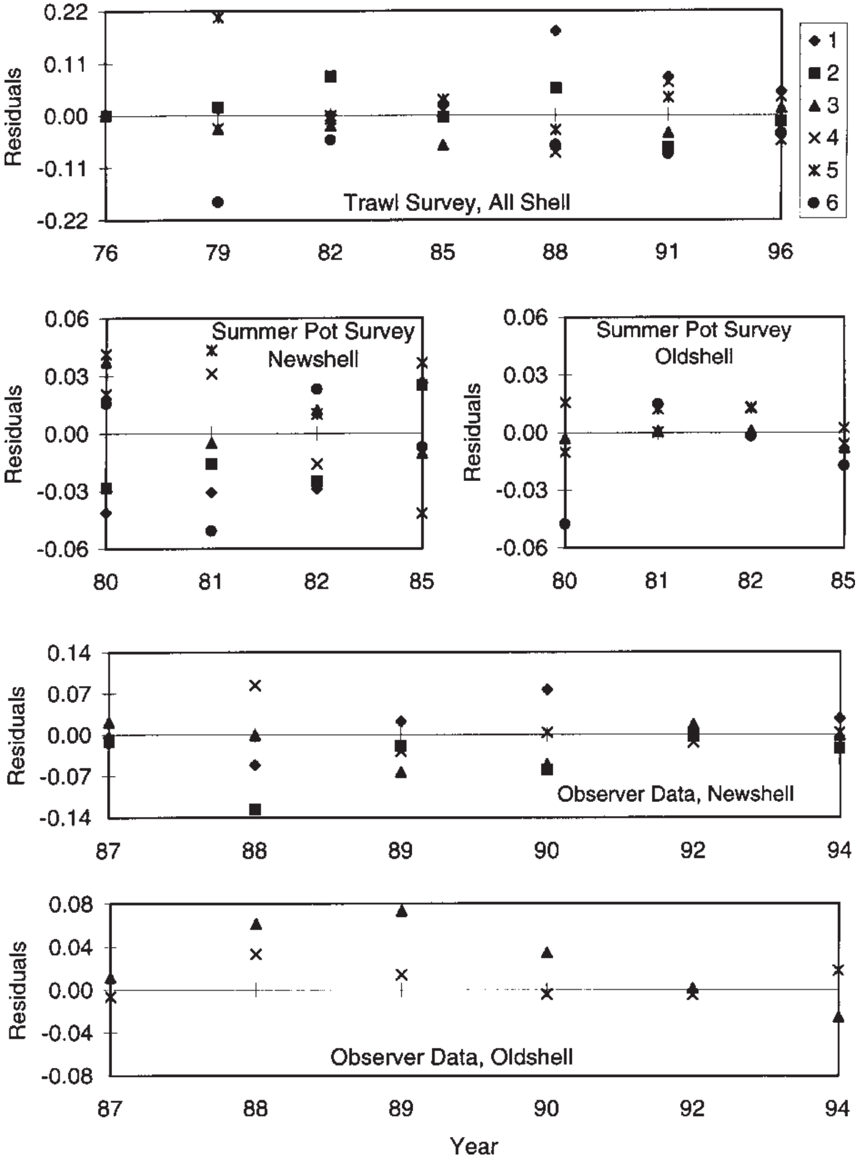


Figure 2. Residuals of length and shell compositions by year for summer trawl and pot surveys and observer data for Norton Sound red king crab. Numbers in the legend represent length classes. All plots have the same legend.

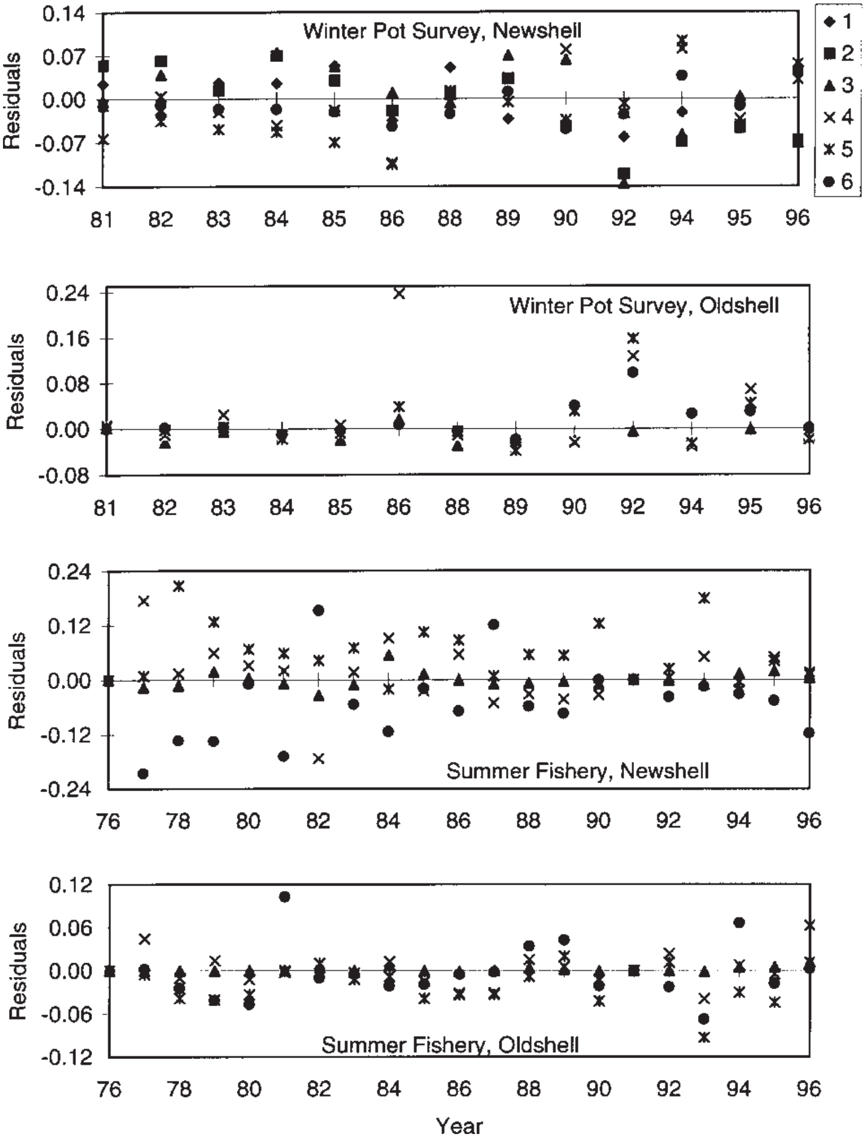


Figure 3. Residuals of length and shell compositions by year for winter pot surveys and summer fishery for Norton Sound red king crab. Numbers in the legend represent length classes. All plots have the same legend.

the winter pot surveys and observer data also occurred in those years with a small sample size. The likelihood function placed less weight to those data with a small sample size. The sample sizes for the summer commercial fishery were large for most years; the large residuals may indicate a large sampling error. Residuals were generally uncorrelated among years, length class, and shell condition with two exceptions: (1) residuals of newshell length classes for the winter pot surveys were generally negative for large length classes and positive for small length classes from 1981 to 1986, and (2) residuals of newshell length class 5 for the summer commercial fishery were mostly positive. These patterns could be modeled by increasing selectivity parameters. However, because the population abundance estimates are unaffected, we chose not to increase the number of model parameters to account for them.

Selectivities for both summer trawl and pot surveys were essentially identical; both were higher than for the summer commercial pot fishery (Table 3). The winter pot surveys caught a small number of crabs in the first and last length classes. A small proportion of crabs belonged to legal crabs in length class 3, and almost all crabs in the last three length classes were legal crabs (Table 3). Here the proportion of legal crabs was only used to separate retained catch in the observer data. For the purpose of this study, legal crab abundance was the sum of abundances in the last three length classes. Molting probabilities were the highest during 1982 to 1990 and lowest during 1976 to 1981 (Table 3), but the differences were small.

Population abundances were very high in the late 1970s and low in the early 1980s and mid-1990s (Fig. 4). Recruitment fluctuated greatly during the past two decades. An exceptionally strong recruitment in 1976 was followed by three years of very weak ones, and then recruitment increased to 1982 with a slightly downward trend from 1983 to 1993. After failing in 1993, recruitment increased during the last three years. High harvest rates (>25%) from the summer fishery occurred from 1979 to 1981, and since then harvest rates have been below 22% (Fig. 4).

Standard deviations of estimated parameters and abundances were artificially small except for those of recruitment estimates. Coefficients of variation for recruitment estimates were from 14% to 72% for 16 out of 21 years, whereas coefficients of variation for other parameter and legal crab abundance estimates were below 11%. Such small standard deviations may partially be caused by the assumptions made in the model and a small number of survey abundances available to estimate catchabilities of the commercial fishing gear. AD Model Builder may also underestimate the standard deviations.

Estimates of parameters and legal crab abundance were not very sensitive to weighting factors for survey abundances and fishing effort, and maximum effective sample size (Table 4; Fig. 5). Increasing weight for fishing effort resulted in decreasing estimated legal abundance from 1976 to 1979 and from 1989 to 1997, whereas increasing weight for survey

Table 3. Estimated selectivities, molting probabilities, and proportions of legal crabs by length (mm CL) class for Norton Sound male red king crab.

			Selectivities				Molting probabilities		
Length class	Length range	Proportion of legals	Summer trawl	Summer pot surv	Winter pot surv	Summer fishery	1976-1981	1982-1990	1991-1996
1	74-83	0.00	0.60	0.60	0.31	0.22	1.00	1.00	1.00
2	84-93	0.00	0.65	0.65	0.61	0.42	1.00	1.00	1.00
3	94-103	0.15	1.00	1.00	1.00	0.66	0.83	0.89	0.88
4	104-113	0.92	1.00	1.00	1.00	0.87	0.61	0.74	0.70
5	114-123	1.00	1.00	1.00	1.00	1.00	0.40	0.57	0.49
6	>123	1.00	1.00	1.00	0.31	1.00	0.24	0.41	0.30

abundances increased estimated abundances. Without survey abundance data, estimated legal crab abundances would be much smaller. Maximum effective sample sizes ranging from 100 to 1,500 hardly affected estimated legal abundances except during the first three years. Assumed natural mortality influenced recruitment estimates but had limited impact on estimates of legal abundances (Fig. 5; Table 4). Higher natural mortality resulted in higher recruitment estimates for all years and higher estimated legal crab abundances during the 1970s and lower abundances after 1988.

Discussion

The advantage of a length-based approach to estimate crab population abundance is that it synthesizes survey, fisheries, and tagging data, and links abundances from multiple years together. Thus, the random component of measurement errors in abundances estimated directly from survey or fisheries data can be filtered out, and the abundance can be projected in years without survey data. The length-based synthesis model is particularly useful for Norton Sound red king crab, for which multiple types of survey data are sporadically available, but none individually provides enough information to estimate abundances each year. The model fit the survey and fisheries data well, and population abundance estimates from the model were much smoother than those estimated by survey data directly.

Weights for different types of data sets that are contradictory can affect the population estimates substantially (Hilborn and Walters 1992; Schnute and Hilborn 1993). Although our data sets contradicted each other in some years (e.g., the catch was higher than the legal population abundance estimated by the trawl survey in 1979, and proportions of oldshell crabs were much higher from the winter surveys than from the commercial

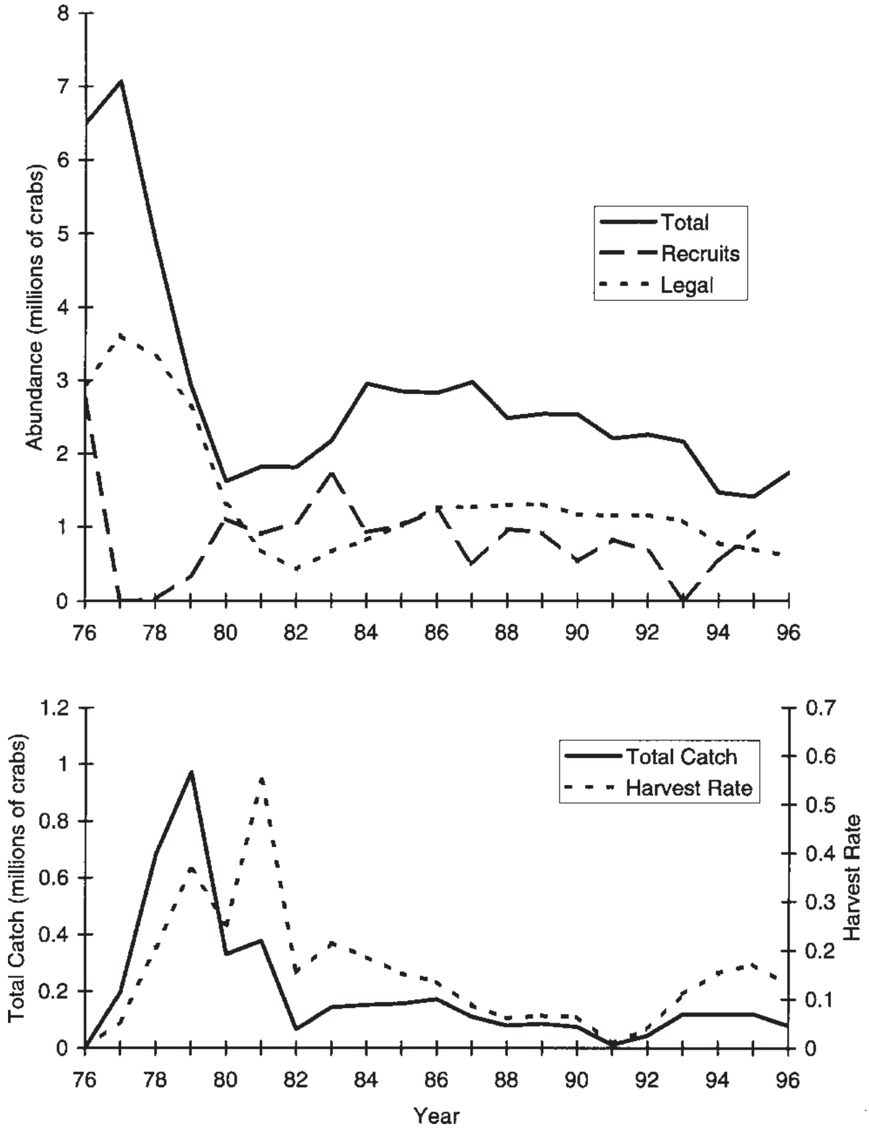


Figure 4. *Estimated total and legal male abundances and recruits (upper plot), and total summer catches and harvest rates (lower plot) of Norton Sound red king crab from July 1, 1976, to June 30, 1996.*

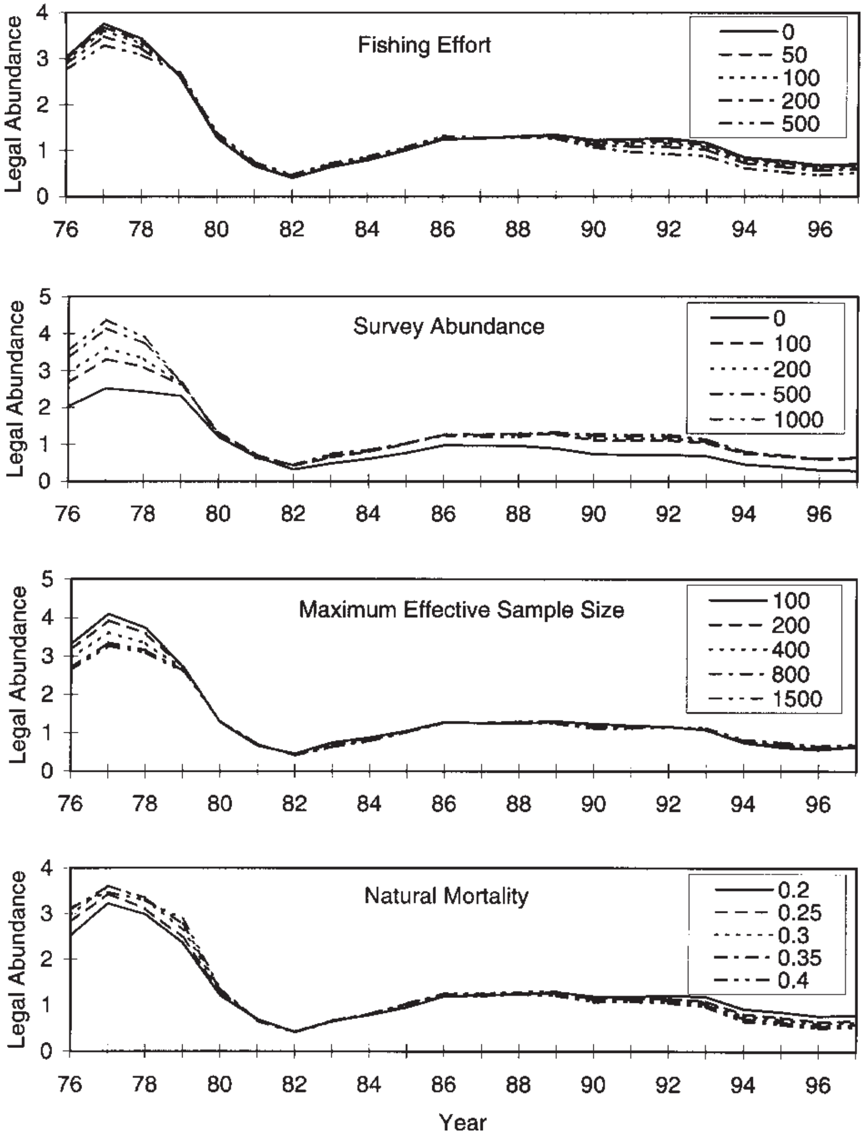


Figure 5. Comparison of estimated legal crab abundances (millions of crabs) under different weighting factors applied to summer fishing effort and survey abundances and under different natural mortalities and maximum effective sample sizes.

Table 4. Summary of parameter estimates for a length-based stock synthesis population model of Norton Sound red king crab.

Parameter	Base	$W_f = 50$	$W_f = 200$	$W_i = 100$	$W_i = 500$	$M = 0.25$	$M = 0.35$	$K_\infty = 200$	$K_\infty = 800$
N_{76}	5.69	5.76	5.56	5.15	6.45	5.44	5.94	6.11	5.21
R_{76}	2.75	2.57	3.05	3.13	1.92	2.05	4.00	2.30	3.02
R_{77}	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
R_{78}	0.02	0.02	0.01	0.06	0.00	0.01	0.02	0.00	0.06
R_{79}	0.34	0.34	0.34	0.29	0.38	0.30	0.38	0.37	0.29
R_{80}	1.10	1.09	1.12	1.07	1.25	0.95	1.21	1.22	1.04
R_{81}	0.90	0.89	0.92	0.91	0.86	0.72	1.05	0.92	0.89
R_{82}	1.06	1.07	1.04	1.07	1.01	0.89	1.26	1.00	1.11
R_{83}	1.71	1.69	1.74	1.76	1.67	1.39	1.92	1.71	1.71
R_{84}	0.93	0.94	0.89	0.91	0.94	0.71	1.12	0.92	0.93
R_{85}	1.01	1.02	1.00	1.07	0.85	0.81	1.20	0.93	0.95
R_{86}	1.22	1.26	1.17	1.10	1.41	0.94	1.46	1.25	1.18
R_{87}	0.49	0.51	0.45	0.43	0.61	0.38	0.59	0.65	0.47
R_{88}	0.96	1.03	0.86	0.95	0.98	0.78	1.15	0.91	1.00
R_{89}	0.92	0.98	0.82	0.88	0.96	0.71	1.11	0.89	0.95
R_{90}	0.53	0.52	0.53	0.53	0.53	0.43	0.58	0.56	0.56
R_{91}	0.82	0.87	0.75	0.83	0.77	0.68	1.00	0.73	0.92
R_{92}	0.66	0.70	0.61	0.66	0.64	0.54	0.76	0.54	0.70
R_{93}	0.01	0.01	0.02	0.01	0.02	0.01	0.05	0.08	0.00
R_{94}	0.53	0.55	0.49	0.53	0.53	0.45	0.57	0.55	0.58
R_{95}	0.94	0.96	0.90	0.89	1.04	0.76	0.99	0.98	0.94
r	0.60	0.61	0.59	0.64	0.58	0.64	0.55	0.59	0.61
q_1	$1.85e{-5}$	$1.85e{-5}$	$1.84e{-5}$	$1.88e{-5}$	$1.81e{-5}$	$1.86e{-5}$	$1.89e{-5}$	$1.81e{-5}$	$1.88e{-5}$
q_2	$1.52e{-5}$	$1.44e{-5}$	$1.66e{-5}$	$1.56e{-5}$	$1.49e{-5}$	$1.42e{-5}$	$1.66e{-5}$	$1.57e{-5}$	$1.43e{-5}$
α_1	0.08	0.08	0.08	0.08	0.07	0.07	0.09	0.08	0.08
β_1	137.08	137.04	137.07	136.92	137.43	138.53	136.10	137.91	135.54
α_2	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.07
β_2	117.84	117.65	118.21	118.05	117.47	117.48	118.29	117.60	120.49
α_3	0.08	0.08	0.08	0.08	0.07	0.07	0.08	0.08	0.07
β_3	115.45	115.68	115.04	115.79	114.73	114.86	115.92	114.79	116.38
$S_{st,1}$	0.60	0.60	0.60	0.60	0.60	0.60	0.60	0.60	0.60
$S_{st,2}$	0.65	0.61	0.70	0.60	0.60	1.00	0.60	0.60	0.60
$S_{sp,1}$	0.60	0.60	0.60	0.57	0.60	0.60	0.58	0.60	0.60
$S_{sp,2}$	0.65	0.65	0.65	0.68	0.60	0.81	0.60	0.61	0.67
$S_{w,1}$	0.31	0.30	0.31	0.28	0.32	0.31	0.32	0.34	0.28
$S_{w,2}$	0.61	0.61	0.60	0.62	0.61	0.68	0.60	0.61	0.61
$S_{w,6}$	0.31	0.31	0.30	0.30	0.31	0.26	0.36	0.31	0.29
m	1.43	1.43	1.42	1.45	1.40	1.30	1.62	1.51	1.59
ϕ	0.07	0.07	0.07	0.07	0.07	0.08	0.07	0.07	0.07

Table 4. (Continued.)

Parameter	Base	$W_f = 50$	$W_f = 200$	$W_i = 100$	$W_i = 500$	$M = 0.25$	$M = 0.35$	$K_\infty = 200$	$K_\infty = 800$
ω	108.22	107.65	109.34	106.53	109.49	99.63	115.72	110.63	107.18
v_6	0.50	0.50	0.50	0.50	0.50	0.50	0.57	0.50	0.50
Log-likelihood	–28,691	–28,524	–29,007	–28,257	–29,784	–28,737	–28,717	–16,871	–43,957

Recruits *R* are in millions of crabs. Unless indicated, the weighting factors are 200 for survey abundances (W_i) and 100 for summer fishing effort (W_f), the maximum effective sample size (K_∞) is 400, and natural mortality (*M*) is 0.3.

catches in 1986 and 1992), the overall trends of all data sets are basically consistent. Therefore, our results are not very sensitive to different weighting factors. Because the commercial harvest is limited to legal crabs and the surveys are designed to estimate the legal component of the population, weighting legal and non-legal crab abundances separately improves estimates of catchability parameters of the commercial fishing gear and helps stabilize population estimates. Estimates of legal crab abundances are more sensitive to the weighting factors if legal and non-legal abundances are not weighted separately.

Different data sets for Norton Sound red king crab complement each other. Winter surveys are inexpensive to conduct and provide length and shell condition information to gauge relative recruitment strength before the coming summer fishery. Currently, winter surveys are the only surveys that are economically feasible to conduct each year. However, because winter conditions vary each year and limit the area sampled, winter surveys do not provide reliable information on absolute or relative abundance indices. Summer surveys extensively cover crab habitats and, combined with tagging studies, provide crucial information on absolute population abundances. But high cost of the surveys in this remote region of Alaska makes them infeasible each year. The preseason survey during summer, which was voluntarily sponsored by the fishermen, is rarely conducted. For most years, summer fishery data provide information only on legal crabs because of lack of observer data on small vessels. The synthesis of all data sets provides information each year to estimate population abundance and dynamics.

In our case, fitting fishing effort is equivalent to fitting CPUE because catches are fairly accurately estimated. The model fit summer fishing effort quite well, and the trend of fishery CPUE basically agreed with survey data. Wide fluctuation of CPUE over time was partially caused by alternation of high and low fishing effort on an annual basis. The changes in size of fishing vessels, pot limit, and fishing patterns in 1993 and 1994 probably caused change in catchability. Small vessels tend to fish in shallower waters than large vessels. Our results indicate that catchability after 1992 was lower than that in the earlier years. However, there is only one summer survey that can be used to scale catchability of summer fishing effort after 1992. More summer surveys are needed to derive a reliable estimate of recent catchability.

Natural mortality is one of the most important parameters affecting population abundance estimates by a length-based stock synthesis model. Unfortunately, little information is available to estimate M for Norton Sound red king crab. Following Low (1991) and Kruse et al. (1996), we assumed M as 0.3, and in addition, the natural mortality for the last length class was assumed as $1.6M$ based on the interactive analysis of residuals. It is intuitive that large-sized, slow-growing old crabs suffer a higher natural mortality caused by senescence. Likewise, because molting probability declines with increasing size, oldshell crabs tend to be older than newshell crabs and may also suffer a higher natural mortality. However, data are not available to estimate natural mortality for oldshell and newshell male crabs separately, and residual patterns do not suggest different natural mortalities. Overall, the robustness of legal abundance estimates to M suggests that uncertainty about natural mortality is not a concern as long as it ranges from 0.25 to 0.35.

Our results suggest that molting probabilities were variable over time, but the difference was small for most years except 1993. Variable molting probabilities over time are reported for red king crabs in Bristol Bay and off Kodiak Island (Balsiger 1974; Zheng et al. 1995, 1996). The extremely low molting probability in 1993 was probably caused by colder than normal temperatures associated with a protracted period of ice cover. Length frequency data from the summer fishery and winter survey confirmed this low molting probability. Variable molting probabilities over time will create uncertainty affecting future population projections. Fortunately, winter surveys are conducted after molting is completed, so they provide information to assess unusually high or low molting probabilities for 1-year-ahead projections.

Our model results can be used to improve the management of the Norton Sound red king crab fisheries. The current constant harvest rate strategy requires estimates of legal crab abundance before the summer fishery. The current practice is to directly use the most-current survey estimate, regardless of measurement errors and how long ago the survey was conducted. This approach could result in much higher or lower harvest rates than expected; for example, estimated harvest rates were consistently lower from 1988 to 1992 and much higher from 1994 to 1996 than the targeted rate of 10%. With the length-based synthesis model, legal abundance can be projected before the summer fishery each year regardless of whether a survey is conducted. In addition, the model filters out random measurement errors and thus provides more reliable abundance estimates than observed abundances. Therefore, the model approach reduces not only the uncertainty on legal abundance estimates but also annual variation in catch. With improved abundance estimates and the current conservative management strategy, risks of overfishing the stock will be greatly reduced.

Acknowledgments

C. Lean of Alaska Department of Fish and Game provided the Norton Sound red king crab fisheries data and reviewed the manuscript. B.G. Stevens of National Marine Fisheries Service (NMFS) provided the survey data collected by NMFS. This study was funded in part by cooperative agreement from the National Oceanic and Atmospheric Administration (NOAA). The views expressed herein are those of the authors and do not necessarily reflect the views of NOAA or any of its sub-agencies. This is contribution PP-162 of the Alaska Department of Fish and Game, Commercial Fisheries Management and Development Division, Juneau.

References

Alverson, D.L., and W.T. Pereyra. 1969. Demersal fish in the northeastern Pacific Ocean: An evaluation of exploratory fishing methods and analytical approaches to stock size and yield forecasts. J. Fish. Res. Board Can. 26:1985-2001.

Balsiger, J.W. 1974. A computer simulation model for the eastern Bering Sea king crab. Ph.D. thesis, Univ. Washington, Seattle, WA. 197 pp.

Brannian, L.K. 1987. Population assessment survey for red king crab (*Paralithodes camtschatica*) in Norton Sound, Alaska, 1985. Alaska Dept. Fish and Game, Tech. Data Rep. 214. 51 pp.

Fair, F.L. 1997. Analysis of red king crab data from the 1996 ADFG trawl survey of Norton Sound. Alaska Dept. Fish and Game, Commercial Fisheries Management and Development Division, Regional Information Report 3A97-10. 16 pp.

Fournier, D., and C.P. Archibald. 1982. A general theory for analyzing catch at age data. Can. J. Fish. Aquat. Sci. 39:1195-1207.

Hilborn, R., and C.J. Walters. 1992. Quantitative fisheries stock assessment: Choice, dynamics and uncertainty. Chapman and Hall, Inc., New York. 570 pp.

Kruse, G.H., F.C. Funk, and J. Zheng. 1996. Were Alaskan red king crabs overfished? In: High latitude crabs: Biology, management, and economics. University of Alaska Sea Grant, AK-SG-96-02, Fairbanks, pp. 295-300.

Lean, C., and B. Brennan. 1997. 1996 Norton Sound district shellfish report to the Alaska Board of Fisheries. Alaska Dept. Fish and Game, Commercial Fisheries Management and Development Division, Regional Information Report 3A97-09. 31 pp.

Low, L.-L. (ed.). 1991. Status of living marine resources off Alaska as assessed in 1991. NOAA Tech. Memo. NMFS F/NWC-211. 95 pp.

Methot, R.D. 1989. Synthetic estimates of historical abundance and mortality for northern anchovy. Amer. Fish. Soc. Symp. 6:66-82.

Otter Research Ltd. 1994. An introduction to AD Model Builder: For use in nonlinear modeling and statistics. Otter Research Ltd., Nanaimo, B.C. 44 pp.

Powell, G.C., R. Peterson, and L. Schwarz. 1983. The red king crab, *Paralithodes camtschatica* (Tilesius), in Norton Sound, Alaska: History of biological research and resource utilization through 1982. Alaska Dept. Fish and Game, Inf. Leafl. 222. 103 pp.

Schnute, J.T., and R. Hilborn. 1993. Analysis of contradictory data sources in fish stock assessment. Can. J. Fish. Aquat. Sci. 50:1916-1923.

Stevens, B.G., and R.A. MacIntosh. 1986. Analysis of crab data from the 1985 NMFS survey of the northeast Bering Sea and Norton Sound. Northwest and Alaska Fisheries Science Center, Processed Rep. 86-16. NMFS, 7600 Sand Point Way NE, Seattle, WA 98115. 35 pp.

Zheng, J., M.C. Murphy, and G.H. Kruse. 1995. A length-based population model and stock-recruitment relationships for red king crab, *Paralithodes camtschaticus*, in Bristol Bay, Alaska. Can. J. Fish. Aquat. Sci. 52:1229-1246.

Zheng, J., M.C. Murphy, and G.H. Kruse. 1996. A catch-length analysis for crab populations. Fish. Bull. U.S. 94:576-588.

Treatment of Data and Model Uncertainties in the Assessment of Southern Bluefin Tuna Stocks

Tom Polacheck, Ann Preece, Andrew Betlehem, and Neil Klaer
CSIRO (Commonwealth Scientific and Industrial Research Organisation)
Division of Marine Research, Hobart, Australia

Abstract

Southern bluefin tuna (SBT) is a highly migratory pelagic species with reasonably complete catch and effort time series data for the entire history of its exploitation. The information base for assessing its status is rich compared to that available for most large pelagic fisheries. The current SBT stock assessment is based on information from catch per unit effort (CPUE), qualitative fishery indicators, and virtual population analyses (VPA) tuned to CPUE and tagging data using the ADAPT framework (Gavaris 1988). Within this framework conflicting results about the most recent trends and future projections occur, depending on hypotheses used regarding natural mortality rates, selectivity, and methods for modeling the older aged fish. A full cross of the parameter space representing hypotheses for the major uncertainties results in thousands of VPAs, which can no longer be individually assessed. Instead, a sample of the parameter space is taken and a weighted mean estimate of the status of the stock is produced. The approaches and methods for identifying, weighting, sampling, and incorporating uncertainties into the assessment process for SBT are described and discussed. The VPA estimates of the current status of the SBT spawning stock relative to historical levels exhibit a remarkable degree of robustness to the range of hypotheses while the estimates of absolute stock size exhibit a much wider range of variability. Not all hypotheses for the major uncertainties were equally compatible with the data and we explore the effects on the results of weighting them by an objective lack of fit index. We also look at the effects on the results of different prior weightings. The results emphasize the need for an integrated and balanced research strategy if substantial improvements to the assessment are to be achieved.

Overall, this paper outlines one approach for integrating and evaluating the large number of hypotheses across a number of dimensions representing some of the major sources of uncertainty in the SBT assessment. The method has provided a useful approach for synthesizing results, determining the sensitivity and importance of uncertainties to different inputs and model structures, comparing the results across different user specified input prior weights, and examining the compatibility of the data with various hypotheses in terms of lack of fit.

Introduction

In this paper we present a method for addressing multidimensional uncertainty in the assessment of the southern bluefin tuna (SBT) stock. We explore a large parameter space that incorporates alternative hypotheses used to address these uncertainties. We look at the robustness of the results, and the effects on the results of different prior weightings for the different hypotheses. We also explore the effects on the results of weighting them by an objective lack of fit index.

Southern bluefin tuna is a highly migratory pelagic species for which the information base for assessing the status of the stock is rich compared to that available for fisheries on most large pelagic species. There is relatively complete catch and effort data for almost the entire history of exploitation as well as extensive auxiliary biological and tagging data.

Advice on the status and recovery of the SBT stock is based on information from catch per unit effort (CPUE), qualitative fishery indicators (e.g., presence or absence of fish in areas where they were abundant historically) and virtual population analyses (VPA) using the ADAPT framework (Gavaris 1988). Within this framework conflicting results about the most recent trends and future projections occur, depending upon hypotheses used regarding natural mortality rates, selectivity, and methods for modeling the older aged fish (Polacheck et al. 1992, 1996, 1997b; Ishizuka and Tsuji 1990; Tsuji and Takeuchi 1997; Klaer et al. 1996, 1997).

To address these conflicts there has been increased research effort into many biological and modeling assumptions, which has yielded improved information on changing growth rates, direct aging, longevity, age of maturity, and age-specific mortality. Many plausible hypotheses for the observed data and underlying stock and fishery dynamics have been developed.

The VPA tuning process is based on a user-defined objective function, which allows for integration and evaluation of these hypotheses and data. Differential weighting of parts of the objective function allows us to examine and evaluate the consistency, sensitivities, residuals, and sources of lack of fit and areas of uncertainty in the overall assessment results. Examination of the compatibility of the data and various hypotheses in terms of lack of fit, has led to the development of a lack of fit index which can be used to weight model results (Polacheck et al. 1997c, 1998).

In the past a small number of VPAs was examined and the results synthesized by simple visual analysis. The increased research effort of recent years has resulted in consideration of a larger number of dimensions of uncertainty and range of hypotheses within each dimension as a basis for the stock assessment and management advice. A full cross of the parameter space results in thousands of VPAs, which no longer can be individually assessed. Instead, a sample of the parameter space is taken and a weighted mean estimate of the status of the stock is produced. In addition, the stock assessment procedure includes weighting of the different hypotheses by the three national delegations of scientists involved in the Commission for the Conservation of SBT (CCSBT). The model results can then be combined to give a mean weighted estimate of the status of the stock for each delegation.

The specification and application of the delegation weights is ad hoc. The procedure is meant to incorporate each delegation's judgment of the prior likelihood of the different hypotheses plus an update of these priors based on their assessment of the consistency of the VPA results with the basic available data (i.e., in a very informal and subjective sense mimicking a Bayesian process). It is recognized by all the scientists involved in the assessment process that there is a need to develop a more objective and statistically based procedure for weighting the different VPA results (see Polacheck et al. [this volume] for a discussion of the problems involved). The results discussed in this paper reveal the value of integrating all available data into a single assessment framework that can be used to test for consistency of the information. They also demonstrate the power of the ADAPT framework to perform this integration within the VPA context and the importance of examining the implication of unresolved inconsistencies and uncertainties for both the assessment and the management of the stock. The approaches and methods for identifying, weighting, sampling, and incorporating uncertainties into the assessment process for SBT are described and discussed.

Material and Methods

The basic catch and age data used in the analyses are documented in Polacheck et al. (1997e) and the details of the VPA model structure and inputs are documented in Polacheck et al. (1997a).

In the analyses presented here, uncertainties in the following dimensions were initially identified as important for evaluation:

1. Estimates of the total catch by age and year.
2. Interpretation of catch rates as indices of abundance.
3. The age-specific natural mortality rates.
4. Parameters to estimate.

Table 1. The range of hypotheses considered for the different uncertainties.

Uncertainty	Options
1. Catch at age	1. Standard transition model 2. Standard plus recent additional catches
2. CPUE interpretation	1. geo_a: spatially correlated errors; area and latitude fixed effects (Thomas and Tosca 1997) 2. geo_c: spatially correlated errors; no fixed effects (Thomas and Tosca 1997) 3. Variable square: aus1var (Hearn and Polacheck 1997) 4. Constant square: aus5con (Hearn and Polacheck 1997) 5. Quasi variable: jpnqv (Nishida and Tsuji 1997)
3. Natural mortality	Mortality rate vectors 1-10 (see Table 3)
4. Parameters estimated	1. All terminal age Fs (i.e., $F_{11,1969}$ to $F_{11,1996}$) and $F_{4,1996}$, $F_{5,1996}$, $F_{7,1996}$ 2. Every 5th year terminal age Fs with linear interpolation between them (i.e., $F_{11,1969}$, $F_{11,1974}$, $F_{11,1979}$, $F_{11,1984}$, $F_{11,1989}$, $F_{11,1994}$, $F_{11,1996}$), and $F_{4,1996}$, $F_{5,1996}$, $F_{7,1996}$
5. Validity of effort tuning index	1. weight = 0 2. Weight = 1
6. Validity of 1990s tagging data for providing estimates of fishing mortality rates (see Polacheck 1997d for description of tagging data and reporting rate models)	1. Weight = 0 2. Weight = 1.0; reporting rate model 1 3. Weight = 1.0; reporting rate model 5 4. Weight = 1.0; reporting rate model 8

Table 1. (Continued.)

Uncertainty	Options
7. PLUS group options	
a. Estimator for the initial size of the plus group[a]	1. $\lambda = 1.00$ 2. $\lambda = 0.75$ 3. $\lambda = 1.50$ 4. $\lambda =$ estimable parameter 5. Alpha method (Restrepo and Legault 1995)
b. Terminal age class	1. Terminal age = 11 2. Terminal age = 7
c. Validity of $F_{\text{continuity 11-12+}}$ term (Tsuji and Takeuchi 1997)[b]	1. Weight = 0 2. Weight = 1
d. Validity of plus group CPUE index	1. Weight = 0 2. Weight = 1.0

[a] For some plus group estimation procedures the initial size of the plus group (in numbers) is estimated as a scaler of the stable age distribution given the estimated natural and fishing mortality rates for the terminal age class (see Polacheck et al. 1997a). λs the parameter specifying the scaler.

[b] $F_{\text{continuity 11-12+}}$ is a term added to the objective function which attempts to maintain to the extent possible a proportional relationship for selectivity within each year between ages 11 and the plus group. It has the form $\Sigma[(\ln F_{10,y} - \ln F_{11,y}) - (\ln F_{11,y} - \ln F_{12+,y})]^2$.

5. The validity of the effort tuning index.
6. The validity of the 1990s tagging data to provide information on fishing mortality rates and the uncertainty in the reporting rates.
7. Method of estimating the plus group.
    a. The estimator for the initial size of the plus group.
    b. The terminal age class (i.e., the final age class before the plus group) used in the assessment.
    c. Structural assumptions regarding selectivity.
    d. Validity of the plus group CPUE as an index of abundance.

Table 1 provides a list of the main alternative models or hypotheses that have been developed to represent the identified uncertainties. Polacheck et al. (1997a) and references therein provide details on the basis for the various alternative models. A large number of different models and hypotheses have been proposed for estimating the size of the plus group. These involve the hypotheses under uncertainty 7 in Table 1. They result from different assumptions and hypotheses about selectivity and the reliability with which the age of the larger fish can be estimated from their lengths (see Polacheck et al. 1996, 1997a). At the 1997 CCSBT scientific committee meeting, these plus group estimation methods and structural hypotheses were condensed into a single, combined "plus group" uncertainty dimension. This was because some combinations of models and hypotheses within uncertainty 7 in Table 1 cannot be crossed (although a conditional approach could also have been used). At the meeting, the options within this plus group uncertainty were arbitrarily limited to the 10 listed in Table 2 due to time and computing resource constraints. The 10 options chosen were based on the national delegations' assessment of the most critical ones to consider. Only these 10 options are considered here. Ten age-specific natural mortality vectors were used (Table 3).

The number of individual VPA estimates required to produce a full cross of all the options listed in Table 1 for the first six uncertainties plus the 10 plus group options in Table 2 is extremely large (16,000). It is not feasible to perform all of these analyses within the time frame of the "normal" assessment process, or to examine in detail a reasonable proportion of them. Instead a large random sample (1,000) of VPA estimates was generated across all of the dimensions and the results summarized. The selection of hypotheses/models across the parameter space for each individual run is made at random such that within each of the dimensions for the first six uncertainties in Table 1 and for the 10 plus group options in Table 2 all individual hypotheses/models are considered and have equal probability of being selected. The randomization process is done with replacement.

In this sampling of the parameter space, weights are assigned to all of the VPA estimates and then summary statistics are calculated as a weighted mean across all of the VPAs run in the sample set. Four individual weights are assigned to each VPA and then a total weight for a VPA is calculated based on the product of these four weights. The four weights are

1. An input user specified prior weight.
2. A convergence weight.
3. A lack of fit output calculated weight.
4. A surplus yield weight.

The input user specified weights are meant to be selected taking into account information that supports the different hypotheses within each of the dimensions and taking into account the consistency of the VPA results with the basic available data (i.e., an examination of the VPA results is undertaken prior to the assignment of the weights). Only marginal weights for each dimension are specified. Weightings assigned by three delegations of scientists at the most recent scientific committee meeting

Table 2. Plus group options.

	1	2	3	4	5	6	7	8	9	10
Terminal age	11	7	11	11	11	11	11	11	11	11
CPUE 12+ index	On	On	On	Off	On	Off	Off	On	Off	On
$F_{\text{cont 11-12+}}$ or Alpha[a]	On	Off	Alpha	Off	Off	Alpha	Off	Off	Off	Off
λ	Est	Est	Est	1	1	1	0.75	0.75	1.5	1.5

[a] Alpha refers to the method of estimating the plus group that has been adopted by ICCAT for assessing Atlantic bluefin tuna (Restrepo and Legault 1995).

Table 3. Natural mortality vectors.

Age	v1	v2	v3	v4	v5	v6	v7	v8	v9	v10
0	0.500	0.500	0.500	0.500	0.400	0.400	0.400	0.300	0.300	0.200
1	0.483	0.450	0.417	0.400	0.383	0.350	0.333	0.283	0.267	0.200
2	0.467	0.400	0.333	0.300	0.367	0.300	0.267	0.267	0.233	0.200
3	0.450	0.350	0.250	0.200	0.350	0.250	0.200	0.250	0.200	0.200
4	0.367	0.300	0.233	0.200	0.300	0.233	0.200	0.233	0.200	0.200
5	0.283	0.250	0.217	0.200	0.250	0.217	0.200	0.217	0.200	0.200
6	0.200	0.200	0.200	0.200	0.200	0.200	0.200	0.200	0.200	0.200
7	0.175	0.175	0.175	0.175	0.175	0.175	0.175	0.175	0.175	0.175
8	0.150	0.150	0.150	0.150	0.150	0.150	0.150	0.150	0.150	0.150
9	0.125	0.125	0.125	0.125	0.125	0.125	0.125	0.125	0.125	0.125
10	0.100	0.100	0.100	0.100	0.100	0.100	0.100	0.100	0.100	0.100
11	0.100	0.100	0.100	0.100	0.100	0.100	0.100	0.100	0.100	0.100
12	0.100	0.100	0.100	0.100	0.100	0.100	0.100	0.100	0.100	0.100

of the Commission for the Conservation of SBT are considered (Table 4; CCSBT 1997), plus an option in which all hypotheses are assigned the same weight (referred to as the equal weighting option).

The convergence weight is set to one if the quasi-Newton procedure used for estimating the VPA parameters satisfactorily met all of the convergence criteria, otherwise zero. Zero convergence occurs where the best estimate is at a constraint. Two constraints are imposed; all estimates of *N* need to be positive and terminal *F*s need to be less than 2.0. Zero convergence also occurs for runs in which the minimization procedure failed to converge (i.e., the number of iterations exceeded the maximum allowed within the computer code). The maximum number of iterations were exceeded in cases where there was a sharp gradient in the objective function at a constraint and/or a tradeoff among which parameters to set at a constraint value. In an automated procedure, the best way to handle unconverged results and results which converge at a constraint is not clear. The approach adopted here of assigning these zero weights was based partially on pragmatic considerations that it does not make sense to include nonconverged estimates, while recognizing that additional effort ideally should be expended on finding the converged results. With respect to results that converge at a constraint, there is no clear best approach. On the one hand the results are "real" estimates for the specified model given that one considers that the constraints as part of the model. On the other hand, the constraints are there because values in excess of them are considered to be "unrealistic." As such, the estimate which includes an estimated value for a constraint is at the limit of being "realistic" and as such highly unlikely. This would suggest that such results should be downweighted, if not completely discounted. On balance, the decision was made to discount such results.

The lack of fit output weights is calculated using the procedure developed in Polacheck et al. (1997c, 1998). This lack of fit weight is based on results of diagnostic tests on regression residuals, i.e., the difference between the tuning indices and the VPA estimate of stock size (assuming a lognormal error structure). It is designed to downweight VPA runs in which there is very substantial and significant lack of fit among the residuals in terms of either temporal trends or nonlinearity. The lack of fit weight was introduced to provide an approach for addressing the question of model misspecification. Note that it does not contain any measure of likelihood or overall general fit to the data. Also, in the current procedure there is no updating of the delegation weights based on the basic fit to the data as measured by the objective function.

Surplus yield, in this context, is defined as the proportional difference between the spawning stock biomass (SSB) that produced a particular recruitment and the cumulative SSB that would result from that recruitment over its life span in the absence of fishing. The total surplus yield for a particular VPA run is the sum of the surplus yield for each year

Table 4. Delegation weightings for each hypothesis.

Uncertainty	Options	Delegation 1 weights	Delegation 2 weights	Delegation 3 weights
Catch at age	Standard	0.35	1	0.4
	Recent additional catches	0.65	0	0.6
CPUE interpretation	geo_a	0.3	0.35	0.05
	geo_c	0.175	0.1	0.05
	Variable square	0.25	0.05	0.6
	Constant square	0.1	0.3	0.2
	Quasi variable	0.175	0.1	0.1
Natural mortality	v1	0.05	0.08	0.15
	v2	0.12	0.08	0.15
	v3	0.12	0.08	0.15
	v4	0.12	0.08	0.15
	v5	0.15	0.12	0.1
	v6	0.15	0.12	0.1
	v7	0.1	0.12	0.1
	v8	0.08	0.12	0.05
	v9	0.08	0.12	0.05
	v10	0.03	0.08	0
Tuning parameters	All terminal age Fs	0.7	0.5	0.5
	Every 5th year Fs	0.3	0.5	0.5
Effort tuning	Off	0.35	0.95	0.5
	On, weight = 1	0.65	0.05	0.5
1990 tagging data	Tagging index off	0.2	0.6	0.2
	Reporting rate 1	0.15	0.1	0.2
	Reporting rate 5	0.25	0.1	0.2
	Reporting rate 8	0.4	0.2	0.4
Plus group options	1	0.03	0.2	0
	2	0.27	0.05	0
	3	0	0.1	0
	4	0.2	0.1	0.25
	5	0.06	0.1	0
	6	0	0.05	0.25
	7	0.15	0.1	0.25
	8	0.04	0.15	0
	9	0.2	0.05	0.25
	10	0.05	0.1	0

in the VPA. The surplus yield weight is defined to be zero or one such that if the VPA results predict surplus production less than one, then the surplus yield weight is set to zero, otherwise one. This surplus yield weight eliminates VPA runs in which the net productivity for stock has on average been negative for the past 27 years. It ensures that the stock and recruitment rates conform to the basic requirements of a viable population and is based on the assumption that on average the recruitment to the population must be sufficient to at least replace the loss in parental stock due to natural mortality.

The surplus yield weight is a form of a prior assumption, which assumes that the SBT stock would be self-sustaining in the absence of a fishery. As a prior it acts indirectly as a constraint which eliminates estimates with very high initial spawning biomass. There is always some danger in assuming a prior which eliminates possibilities which are not "physically" or "logically" impossible. However, in the current situation, this prior was considered appropriate for a number of reasons besides the long-term biological one for the stock to be ecologically and evolutionarily viable. The SBT fishery started in the 1950s on the spawning grounds off Indonesia and grew rapidly. The largest catches ever taken from the stock occurred over a period of a few years in the early 1960s on the spawning grounds. These catches were accompanied by very rapid and steep declines in catch rates. Lack of adequate size data and tuning indices have prevented the current VPA assessment from extending back an additional nine years to cover this initial period of heavy exploitation. However, extremely high initial spawning biomass would be incompatible with a stock that had been greatly reduced nine years earlier as a result of removals on the spawning ground. (Note that recruitment in the intervening period is not a likely explanation as it takes eight years for fish to enter the spawning stock. Moreover, an indication of the levels of recruitment in the 1958-1968 period can be had by extending backward the VPA estimates for incompletely exploited cohorts in 1969. There is no indication of any extremely large year classes [see Polacheck et al. 1997a,b].) Moreover, within the period covered by the VPA, recruitment trends in all cases are relatively constant for the period up to 1978-1980 indicating a high level of compensation over what would have been a period of very rapidly declining spawning biomass. This would not seem consistent with model results predicting no surplus productivity over the same period.

In order to get the combined mean weighted estimate for the output statistics from individual VPAs, a weighted mean is calculated as follows:

$$\overline{X}_w = \sum_i X_i \prod_j W_{ij} \Big/ \sum_i \prod_j W_{ij}$$

$\overline{X}_w$ = combined mean weighted estimate for a statistic of interest (e.g., spawning stock biomass [SSB] in 1996);

X_i = the estimate from the ith VPA model;

W_{ij} = the individual possible j weightings being considered in the analysis (i.e., weightings 1 to 4 above).

This paper focuses on how the defined uncertainties translate into uncertainty in the estimates of current stock sizes, historical stock sizes, and the ratio of current stock sizes to historical levels. Stock size is defined in terms of SSB, which is the total weight of the sexually mature individuals. SSB in several key years for the fishery are used to assess the status of the stock. These include SSB in 1960 which is the earliest year for which the SSB is estimated (1960 is near the beginning of exploitation on this stock, and it is also the year when the largest catches were taken), SSB in 1962 (the earliest year for which an SSB estimate is available for all models, estimates of SSB prior to 1962 are not possible within the current model framework for models in which the terminal age is 7), SSB in 1980 which is the current target level identified by the CCSBT for recovery of the stock, and SSB in 1996 which is the latest year in the models. Relative measures of the status of the stock are also examined (e.g., SSB96/62 ratio, SSB96/80 ratio, SSB96/95 ratio, SSB95/94 ratio) to provide a measure of depletion and recent relative changes in SSB.

Results

The shape of the histograms for absolute stock size estimates in 1996 show a long tail of outlying estimates (Fig. 1). Note that in all the figures, results for the equal weighting option are used together with the convergence and surplus yield weights, and the lack of fit weight has not been included. In Fig. 2 a similar wide range in model results for the absolute SSB values for 1962 are evident. The factors that give the higher SSB values are the terminal age 7 plus group model 2 and the variable squares CPUE series especially in combination with use of the tagging tuning data.

The relative stock status is examined via the ratio of SSB in 96 to SSB in 62 (Fig. 3). The general trends in the VPA results are remarkably similar across most of the ranges of uncertainty examined. It is the absolute magnitude of the estimates which varies (this is a scaling issue within the VPAs). The current status of the stock is consistently estimated to be around 10-20% of historical stock size. The importance of the early years in assessing the current status of the stock can be seen in the comparison of the ratio of SSB96/SSB62 and SSB96/SSB60 in Fig. 3. The comparisons indicate that across these VPAs the mean of the 1996 SSB is 10% of the 1960 level and 14% compared with the 1962 SSB. The recovery target for the stock is the 1980 SSB. The mean estimate of the current status of the stock relative to the 1980 level is 39% (Fig. 4).

To examine the effects of the different weightings, the mean and coefficient of variation of the estimates are presented in Table 5. In all cases, the results have been weighted by the convergence weight and surplus

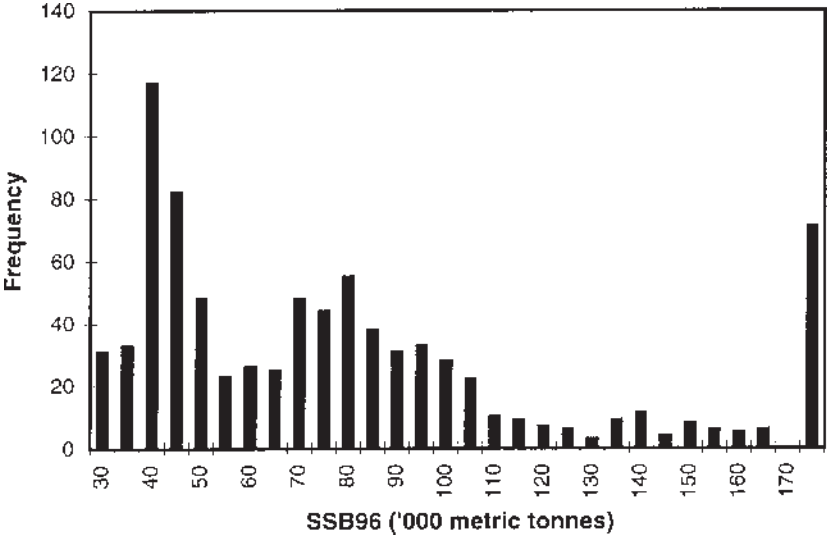


Figure 1. A histogram of VPA estimates of SSB in 1996.

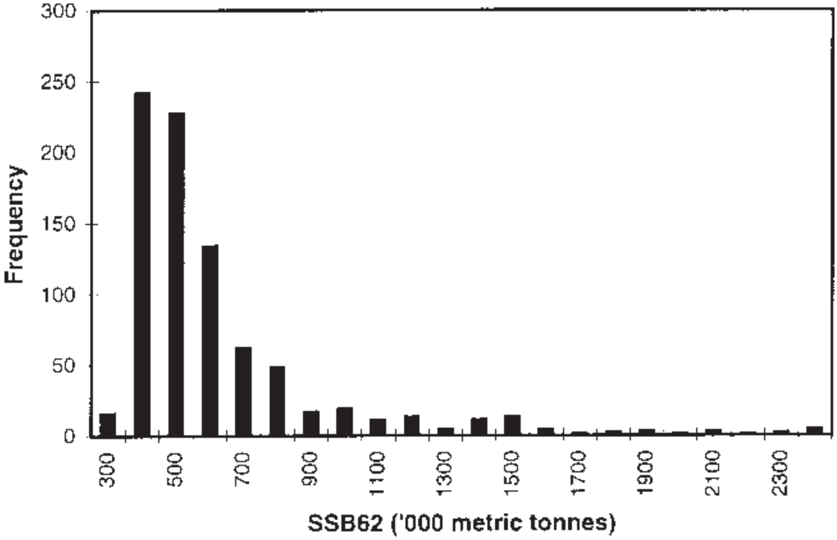


Figure 2. A histogram of VPA estimates of SSB in 1962.

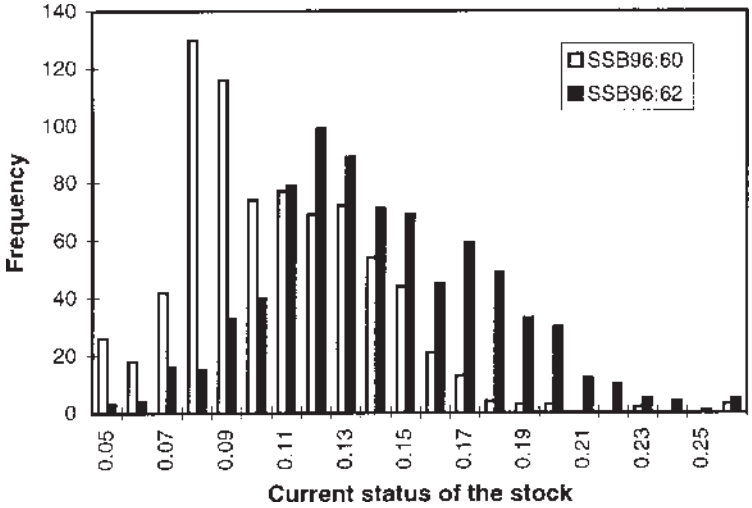


Figure 3. A comparison of the SSB in 1996 relative to SSB in 1962 and relative to SSB in 1960.

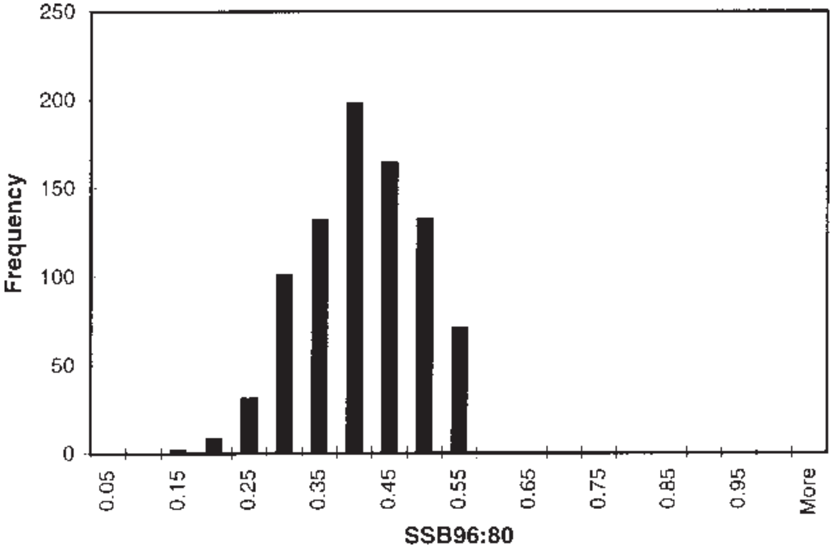


Figure 4. A histogram of the VPA estimates of the 1996 status of the stock relative to the target level SSB, i.e., the ratio of SSB in 1996 to SSB in 1980.

Table 5. The average estimate of SSB and levels of depletion, weighted and unweighted by delegation and lack of fit weightings.

Convergence weight	Surplus yield weight	Delegation weight	Lack of fit weight	SSB62 mean	CV	SSB80 mean	CV	SSB96 mean	CV
Y	Y	1	No	660.23	0.83	263.66	0.86	99.14	0.91
Y	Y	1	Yes	814.78	0.78	344.92	0.82	135.51	0.88
Y	Y	2	No	429.97	0.57	139.69	0.97	59.51	0.85
Y	Y	2	Yes	516.14	0.61	201.16	0.92	76.85	1.05
Y	Y	3	No	571.57	0.44	225.17	0.54	67.88	0.72
Y	Y	3	Yes	562.38	0.41	223.77	0.53	65.96	0.78
Y	Y	None	No	579.65	0.71	226.59	0.87	86.18	0.92
Y	Y	None	Yes	761.29	0.77	326.63	0.94	123.37	1.06

Convergence weight	Surplus yield weight	Delegation weight	Lack of fit weight	SSB96/62 mean	CV	SSB80/62 mean	CV	SSB96/80 mean	CV	SSB96/95 mean	CV	SSB95/94 mean	CV
Y	Y	1	No	0.140	0.296	0.382	0.194	0.366	0.210	1.058	0.058	0.992	0.056
Y	Y	1	Yes	0.154	0.311	0.406	0.191	0.372	0.184	1.053	0.053	0.974	0.060
Y	Y	2	No	0.133	0.228	0.307	0.317	0.443	0.208	1.112	0.075	1.049	0.067
Y	Y	2	Yes	0.137	0.267	0.365	0.187	0.375	0.173	1.095	0.065	1.026	0.057
Y	Y	3	No	0.116	0.385	0.386	0.169	0.296	0.375	1.001	0.090	0.964	0.063
Y	Y	3	Yes	0.113	0.416	0.389	0.184	0.286	0.403	0.997	0.093	0.961	0.066
Y	Y	None	No	0.140	0.294	0.367	0.241	0.387	0.212	1.055	0.070	1.000	0.059
Y	Y	None	Yes	0.146	0.354	0.402	0.234	0.356	0.209	1.047	0.066	0.979	0.061

Table 6. The mean lack of fit weight and average surplus yield for each of the three delegations.

Delegation	Mean lack of fit weight	Average surplus yield
1	0.455	289
2	0.167	565
3	0.380	245

yield weight which effectively excludes 161 of the 1,000 VPAs. Estimates are presented for each of the four input weighting schemes (each delegation's weights plus the equal weighting option) with and without applying the output lack of fit weight. The lack of fit weight for all of the input weighting schemes has only a slight effect on the estimate of the current status of the stock and increases the estimate of the absolute stock size for all but the delegation 3 weighting scheme. The mean estimates of the current status of the stock relative to 1962 when the lack of fit weight is ignored is 14% for delegation 1, 13.3% for delegation 2, 11.6% for delegation 3, and 14% for the equal weighting option.

For each of the delegations the mean lack of fit weight and mean surplus yield are presented in Table 6. Delegation 1 had the highest lack of fit and delegation 2 had the lowest.

Table 7 shows the mean lack of fit weight for the main factors affecting the VPA results. (Note that the equal weighting option is used in Tables 7-9.) The plus group models 1, 3, 5, 6, 8, and 10 had low lack of fit weights (<0.15) which indicates that these models do not fit the data adequately. Plus group model 2 had the highest mean lack of fit weight and model 4 also had a reasonable overall lack of fit weight. Plus group model 2 examines the hypothesis that the terminal age should be set to 7. Plus group model 4 excludes the CPUE for age 12+ index based on hypotheses that there have been major changes in selectivity for older fish due to changes in targeting and fishing practices. The mean lack of fit results for the different CPUE series show that the VPAs were best able to fit the variable squares and geo_a CPUE series.

Mean estimates of SSB, depletion levels, surplus yield and fit weights highlight the differences in the results that the plus group models can produce (Table 8). Table 9 shows the range of mean depletion level estimates for the set of hypotheses considered for each of the main uncertainties.

The range of the mean estimates of the current stock size relative to 1962 for each of the individual hypotheses within any of the main uncer-

Table 7. Average lack of fit weight for the plus group models and CPUE series.

Plus group option	Average lack of fit weight	N
1	0.039	98
2	0.886	69
3	0.137	34
4	0.483	88
5	0.100	109
6	0.135	78
7	0.398	84
8	0.088	93
9	0.443	101
10	0.075	86
CPUE	**Average lack of fit weight**	**N**
Variable squares	0.359	164
Constant squares	0.111	166
geo_a	0.377	163
geo_c	0.243	172
Quasi variable	0.259	175

tainties ranged from 0.11 to 0.18 (Table 9). Overall, across all combinations of hypotheses, 80% of the estimates of this ratio were between 0.10 and 0.20 and only 0.4% were below 0.05 and another 0.65% were above 0.25 (Fig. 3). Estimates of the mean of this ratio were relatively insensitive to the disparate individual delegation weightings and were similar to the mean unweighted by delegation (Table 5).

The estimates of absolute stock size exhibited a much wider range of variability and the absolute magnitude of the estimates were scaled up or down depending on combinations of hypotheses. The scale of the biomass estimates, however, is not unbounded. The data and the model (through the magnitude of the catches in combination with the CPUE trends) constrain the estimates of the absolute range in two ways: (1) by requiring that the more recent stock size estimates are sufficient to have sustained the recent catches without unrealistically high *F*s, and (2) by requiring that the historical population levels were not so great as to result in a long

Table 8. Mean estimates of SSB, depletion levels and lack of fit weights for the different plus group models.

Plus group model	Mean SSB62	Mean SSB96	Mean SSB96/62	Mean SSB96/80	Mean surplus yield	Mean lack of fit weight
1	370	42	0.11	0.42	666	0.04
2	1419	239	0.18	0.41	78	0.89
3	731	110	0.15	0.36	145	0.14
4	450	57	0.12	0.33	350	0.48
5	533	81	0.15	0.43	411	0.10
6	588	95	0.16	0.39	240	0.14
7	488	81	0.14	0.33	361	0.40
8	533	82	0.15	0.43	406	0.09
9	491	58	0.12	0.32	299	0.44
10	523	79	0.15	0.44	452	0.07

Table 9. The range of mean depletion level estimates for the range of individual hypotheses considered for each of the main uncertainties within the SBT assessment.

Uncertainty	SSB96/62	SSB96/80
Plus group model	0.113-0.177	0.320-0.439
CPUE interpretation	0.122-0.147	0.288-0.441
Natural mortality rate	0.118-0.170	0.364-0.410
Effort index	0.139-0.141	0.386-0.387
Tagging estimates	0.127-0.157	0.371-0.398
Tuning parameters	0.131-0.152	0.379-0.397
Catch at age	0.140-0.140	0.374-0.040

time trend of historical stock sizes with negative surplus productivity. These two requirements resulted in 16% of the VPA results being given zero weight and indicate that some combination of hypotheses for the range of uncertainties are inconsistent with the basic data being used in the assessment. These tended to be the plus group model 3 and the variable squares CPUE series. However, within any of the individual uncertainties considered, none of the hypotheses could be considered to be completely inconsistent on these grounds (i.e., for any given hypothesis within any one of the uncertainties, there were always some VPAs which gave realistic results).

Plus group model 2 tended to yield substantially higher estimates of absolute spawning stock levels than the other methods (e.g., its mean for the 1962 SSB level was nearly twice that of any of the other models [Table 8]). Model 2 attempts to reconcile the inconsistency noted above with the plus group CPUE by assuming that the source of the problem is in the aging of older fish, and thus it sets the terminal age in the VPA at 7. In pooling all of the ages past age 7, model 2 assumes that selectivities past age 7 are equal when interpreting the CPUE tuning series. However, the models with a terminal age of 11 suggested that selectivities between ages 7 and 11 have not been equal and that there has been large changes over time. Model 2 resulted in the highest mean lack of fit weight (Table 7), which is perhaps not surprising given that it does not have to deal with either the inconsistencies in the CPUE time series or these selectivity changes. However, a relatively high proportion of the estimates it provides suggest negative net productivity for the stock. These factors combined with the fact that the estimates of the spawning biomass is composed entirely of the plus group raises questions about the reliability of the results from this model.

In contrast, model 1 tends to yield the lowest absolute estimates. Model 1 is based on a terminal age of 11 and tries to minimize differences in selectivity between ages 10 and 11 and the plus group. These differences highlight the uncertainty and importance in the SBT assessments of the different models for the plus group and related selectivity assumptions (see Polacheck et al. 1997b for further discussion).

It is important to note that the different plus group models are not equally compatible with the input data. In particular, models 1, 3, 5, and 6 always yielded VPA results in which there was significant lack of fit to at least one of the tuning indices based on the diagnostic tests and these models also tended to have low mean fit weights (i.e., less than 0.15) as did model 8 and 11 (Table 7). Model 1, especially, appears to be incompatible with the basic data, as the fit weight for any of its VPA results never exceeded 0.17. For the other uncertainty dimensions, similar incompatibility was not apparent for any of individual hypotheses (i.e., there was always some VPA results for which the fit weight equaled 1). However, some combinations of hypotheses for these other uncertainties also appear to be incompatible with the input data. For example, some combinations

Table 10. The total number of VPAs given zero convergence weight for each of the plus group models and for the CPUE series.

Plus group option	Number of VPAs with zero convergence weight
1	0
2	32
3	69
4	12
5	3
6	18
7	12
8	3
9	6
10	3
CPUE	**Number of VPAs with zero convergence weight**
Variable squares	48
Constant squares	26
geo_a	38
geo_c	19
Quasi variable	27

of natural mortality rates and CPUE hypotheses always yielded relatively low fit weights.

Of the random sample of 1,000 VPAs, 161 VPAs did not meet the convergence and surplus yield criteria (see above) and were given a weight of zero. An analysis of the factors which led to VPAs having a convergence or surplus yield weight of zero show that the variable squares and geo_a CPUE series and plus group model 3 were the main uncertainties involved (see Table 10).

Discussion

The VPA estimates of the current status of the SBT spawning stock relative to historical levels exhibited a remarkable degree of robustness to the wide range of hypotheses that have been developed for the main uncertainties. For example, the mean estimates of the current stock size relative to almost pre-exploitation levels (1962) ranged from 0.11 to 0.18 (Table

9). A similar robustness was also exhibited in the estimates of the current level of depletion relative to the CCSBT management defined rebuilding target of the 1980 SSB for this stock (Fig. 4, Table 5).

As noted in the results, the estimates of absolute stock size exhibited a much wider range of variability. For many of the hypotheses the absolute magnitude of the estimates are scaled up or down without having substantial effects on the relative trends in the VPA time series of estimates (see Polacheck et al. 1997b for more detail). This appears to stem from the fact that most of the tuning information only provides relative indices and does not directly relate to absolute abundances or actual fishing mortality rates. The CPUE and effort indices, being relative indices, are reasonably compatible with a wide range of absolute stock size. The catch-at-age matrix in itself is not sufficient to determine the absolute scale, i.e., the least square fit to these indices is flat relative to the absolute stock size (see Polacheck et al. 1997b). In the absence of adequate other data, some form of a structural assumption in terms of selectivity appears to be required and these are the source of the various models for the plus group (see Polacheck et al. 1997b).

The different hypotheses for modeling the plus group are the major source of the large variability in the absolute estimates of spawning stock biomass. The prominent role of the plus group estimates in this assessment is of concern since the plus group estimates utilize no age and cohort information from the catch. This is the dominant source of information within the VPA and catch at age estimation approaches. Thus, the mean of the 1962 spawning stock biomass estimates vary by a factor of nearly 4 over the range of different plus group hypotheses. In some respects, this is not surprising as the plus group comprises a substantial fraction of the spawning age fish (i.e., in these assessments, SBT are assumed to mature on average at age 8 and the plus group is composed of fish aged 8-40 for plus group model 2 and 12 to 40 for the remaining plus group models). The modeling of the plus group is problematic in the SBT assessment because of the need for a selectivity assumption to help establish a scale for the VPAs and because the estimated catch-at-age and CPUE series appear to be inconsistent (CCSBT 1996). The latter reflects the fact that the CPUE series for the plus group when set at age 12 tends to be relatively flat over an extended time period which is inconsistent with the large and continuous decline observed for the other age classes (see CCSBT 1996; Polacheck et al. 1996, 1997b).

The large differences in the fit weights across the range of plus group models demonstrate a major potential pitfall in attempting to do automated analyses of uncertainty in a high dimensional space across a wide range of hypotheses. In such cases, it is not possible to examine individually each result to ensure that the input data are compatible with the combination of hypotheses used to generate a particular result. In such automated analyses, there clearly needs to be a mechanism for screening or weighting the outputs in terms of lack of fit.

It is also important to note that while there is a large amount of robustness in the estimates of current relative stock sizes or depletion levels that this does not necessarily translate into a similar degree of robustness when estimating the probability of recovery under constant catch scenarios (CCSBT 1996; Klaer et al. 1996, 1997). The latter is the current basis for the management advice within the CCSBT. Both the absolute estimates of absolute stock size and relative depletion levels affect the projection results. Perhaps somewhat paradoxically, VPA results in which the current stock sizes are less depleted relative to historical levels, tend to have lower probabilities of recovery. This is because VPA results for which the estimated depletion level is low tend to result in higher levels of compensation in estimates of the stock recruitment relationship (i.e., greater steepness at the origin). This can be seen in the relationship between the estimated mean surplus yield and the relative depletion levels (Fig. 5). There is even a stronger relationship between surplus yield and estimated spawning stock size in 1962 (Fig. 6). High initial stock sizes are incompatible with SBT being a highly productive stock.

The delegation-specified input weightings had little effect on the mean estimates of the relative measures of stock size. This is not surprising given the robustness noted above in the relative measures. However, the differences in the delegation weighting had substantial effects on the mean estimates of absolute stock sizes with the means between delegations varying by a factor of 1.9 for the SSB in 1962. Given the general relationship between the estimates of absolute stock size and surplus yield, there are also large differences in the estimates of the mean surplus yield. There were also substantial differences in the mean fit across the different delegations (Table 6). These large differences in both the mean absolute abundance estimates, surplus yield and fit weight statistics largely reflects the differential weighting that the delegations placed on the models for the plus group (Tables 4 and 8). Thus, delegation 2 placed a high weight on plus group model 1, which as noted above always had low value for the fit weight while delegation 1 placed a relatively high weight on model 2 which tended to have high fit weight values. However, additionally weighting the delegation weighted results by the fit weights did not markedly reduce the discrepancy between the delegation-weighted estimates (Table 5). This is because the input weights from the delegations were highly divergent for the plus group models. (i.e., models weighted relatively high by one delegation were given low or zero weights by another, Table 4).

A major focus of the discussions of the uncertainty in the SBT stock assessments has been the differences in the interpretations of the CPUE time series (Hearn and Polacheck 1997). However, the results presented here clearly demonstrate that the uncertainty in the CPUE is only one of the large sources of uncertainty in the current SBT assessment (Table 9). The uncertainty associated with modeling the plus group and natural mortality rates induce similar or greater levels of uncertainty than the different interpretation of CPUE in terms of estimates of absolute stock

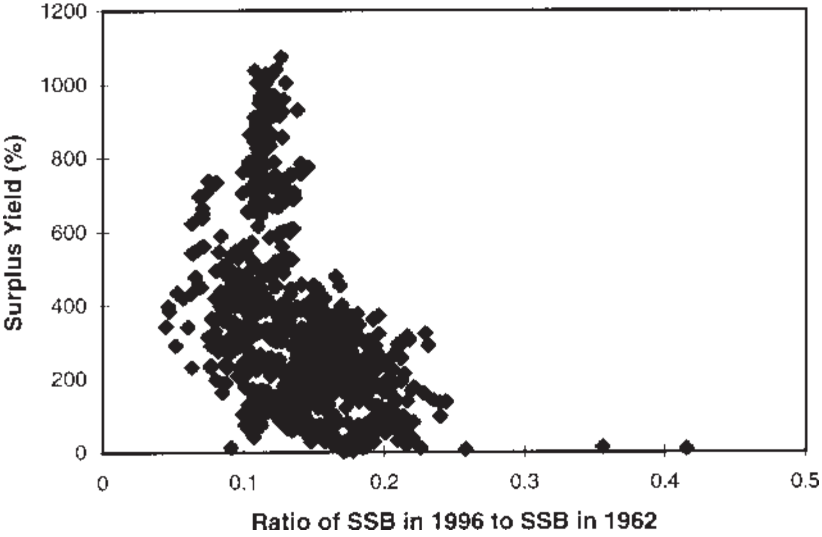


Figure 5. Surplus yield versus relative depletion levels.

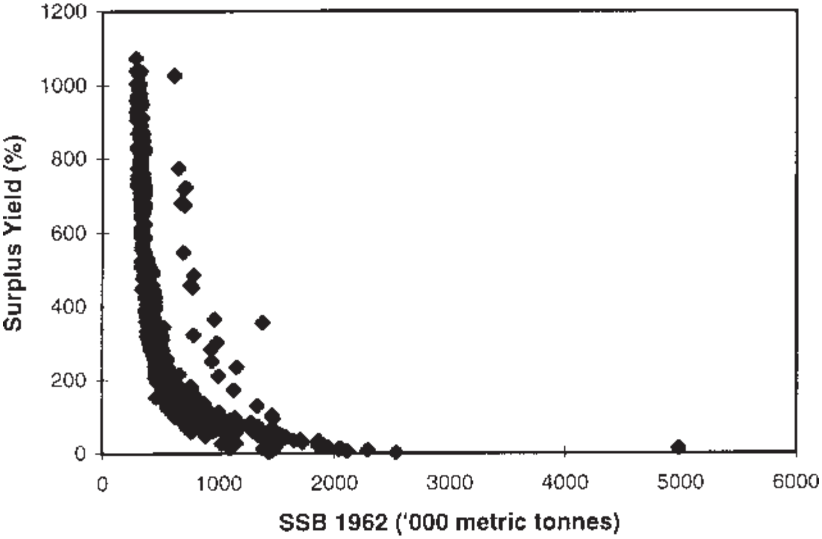


Figure 6. Surplus yield versus SSB in 1962.

size, depletion levels, and mean productivity. Even if the interpretation of CPUE could be narrowed to a single hypothesis, the overall range of uncertainties in the current assessments would not be substantially reduced. This emphasizes the need for an integrated and balanced research strategy if substantial improvements to the assessment are to be achieved.

It must be noted that some important areas of uncertainty have not been included in the current analyses and for some dimensions that the range of alternative hypotheses was limited. These included:

1. Changes in catchability or effective effort over time due to gear and other technological improvements (e.g., Whitelaw and Barron 1995).

2. The reliability and representativeness of the size distributional data and their conversion to age.

3. Levels of total catch (including discards) both within and outside of the CCSBT fisheries, particularly in recent years when the official fisheries have been limited by restrictive quotas.

4. The mean age of maturity which recent results suggest may be substantially older than the currently used value of 8 (Gunn et al. 1996).

While the above areas are recognized as large areas of uncertainty with potentially important consequences, development and/or implementation of meaningful alternative hypotheses is difficult because of a paucity of data from which to parameterize and evaluate them. There are clearly large uncertainties that exist both within and outside of the VPA models as currently structured. These are making it difficult both to provide assessment advice on short-term changes in stock sizes and to measure the uncertainty in the actual assessments (Polacheck et al. 1997b). However, the VPA results appear to be providing reasonably robust conclusions about the current status of the stock relative to their historical and pre-exploitation sizes.

Overall, this paper outlines one approach for integrating and evaluating the large number of hypotheses across a number of dimensions representing some of the major sources of uncertainty in the SBT assessment. The method has provided a useful approach for synthesizing the overall results, determining the sensitivity and importance of uncertainties to different inputs and model structures, comparing the results across different user-specified input prior weights, and examining the compatibility of the data with various hypotheses in terms of lack of fit. The results also demonstrate the importance of developing automated measures for evaluating the lack of fit with any such approach.

References

CCSBT. 1996. Report of the CCSBT workshop on VPA and CPUE modelling. 4-15 March 1996. Hobart, Tasmania.

CCSBT. 1997. Report of the third scientific committee meeting. Commission for the Conservation of Southern Bluefin Tuna. 28 July-8 August 1997, Canberra, Australia.

Gavaris, S. 1988. An adaptive framework for the estimation of population size. Can. Atl. Fish. Sci. Adv. Comm. (CAFSAC) Res. Doc. 88/29.

Gunn, J., T. Davis, J. Farley, N. Clear, and K. Haskard. 1996. Preliminary estimates of the age structure of the southern bluefin tuna spawning stock (including a revised estimate of the age at first spawning). Second CCSBT Scientific Committee Meeting, Hobart, Australia. CCSBT/SC/96/10.

Hearn, W.S., and T. Polacheck 1997. Applying general linear models to CPUE to estimate indices of SBT abundance. Third CCSBT Scientific Committee Meeting, Canberra, Australia. CCSBT/SC/97/19.

Ishizuka, Y., and S. Tsuji. 1990. Assessment of the southern bluefin tuna stock. Ninth Trilateral Scientific Meeting on SBT, Hobart, Australia. SBFWS/90/2.

Klaer N., T. Polacheck, K. Sainsbury, and A. Preece. 1996. Southern bluefin tuna stock and recruitment projections. Second CCSBT Scientific Committee Meeting, Hobart, Australia. CCSBT/SC/96/17.

Klaer N., T. Polacheck, A. Preece, A. Betlehem, and T. Jones. 1997. Assessment of the probability of recovery under different catch scenarios and the effect of resolving uncertainties in SBT assessment. Third CCSBT Scientific Committee Meeting, Canberra, Australia. CCSBT/SC/97/28.

Nishida, T., and S. Tsuji, 1997. Estimation of abundance indices for southern bluefin tuna *(Thunnus maccoyii)* based on the coarse scale Japanese longline fisheries data (1969-96). Third CCSBT Scientific Committee Meeting, Canberra, Australia. CCSBT/SC/97/14.

Polacheck, T., A. Betlehem, and A. Preece. 1997a. Documentation of the virtual population analysis methods and model inputs used for estimation of current and historical stock sizes of southern bluefin tuna population for the third CCSBT scientific committee meeting, 1997. Third CCSBT Scientific Committee Meeting, Canberra, Australia. CCSBT/SC/97/24.

Polacheck, T., K. Sainsbury, and N. Klaer. 1992. Assessment of the status of the southern bluefin tuna stock using virtual population analysis, 1992. Eleventh Trilateral Scientific Meeting on SBT, Shimizu, Japan. SBFWS/92/7.

Polacheck, T., A. Betlehem, A. Preece, and N. Klaer. 1997b. Estimates of current and historical stock sizes for use in the evaluation of the impact of experimental fishing programs on the southern bluefin tuna population. Third CCSBT Scientific Committee Meeting, Canberra, Australia. CCSBT/SC/97/23.

Polacheck, T., A. Preece, A. Betlehem, and K. Sainsbury. 1996. Assessment of the status of the southern bluefin tuna stock using virtual population analysis, 1996. Second CCSBT Scientific Committee Meeting, Hobart, Australia. CCSBT/SC/96/26.

Polacheck, T., K. Haskard, N. Klaer, A. Betlehem, and A. Preece. 1998. An index for weighting results in catch-at-age models based on diagnostic tests for lack of fit. In: F. Funk, T.J. Quinn II, J. Heifetz, J.N. Ianelli, J.E. Powers, J.F. Schweigert, P.J. Sullivan, and C.-I. Zhang (eds.), Fishery stock assessment models. University of Alaska Sea Grant, AK-SG-98-01, Fairbanks (this volume).

Polacheck, T., K. Haskard, N. Klaer, A. Preece, and A. Betlehem. 1997c. A procedure for assigning output weights to VPA results based on diagnostic tests for the tuning indices. Third CCSBT Scientific Committee Meeting, Canberra, Australia. CCSBT/SC/97/29.

Polacheck, T., W.S. Hearn, C. Millar, W. Whitelaw, and C. Stanley. 1997d. Updated estimates of mortality rates for juvenile SBT from multi-year tagging of cohorts. Third CCSBT Scientific Committee Meeting, Canberra, Australia. CCSBT/SC/97/26.

Polacheck, T., A. Preece, A. Betlehem, N. Klaer, and W.S. Hearn. 1997e. Estimation of the historical catch and its age distribution for southern bluefin tuna. Third CCSBT Scientific Committee Meeting, Canberra, Australia. CCSBT/SC/97/20.

Restrepo, V.R., and C.M. Legault. 1995. Approximations for solving the catch equation when it involves a “plus group.” Fish. Bull. U.S. 93:308-314.

Tsuji, S., and Y. Takeuchi. 1997. Stock assessment and future projection of southern bluefin tuna, 1997. Third CCSBT Scientific Committee Meeting, Canberra, Australia. CCSBT/SC/97/17.

Thomas, M., and P. Tosca. 1997. Spatial analysis of southern bluefin tuna catch per unit effort data. Third CCSBT Scientific Committee Meeting, Canberra, Australia. CCSBT/SC/97/22.

Whitelaw, W., and M. Baron. 1995. Recent changes in Japanese longline fishing gear and techniques which may affect cpue. First CCSBT Scientific Committee Meeting, Shimizu, Japan. SBFWS/95/10.

Virtual Population Analysis with the Adjoint Method

M.J. Huiskes
Wageningen Agricultural University, Wageningen, The Netherlands

Abstract

In this paper we investigate an adjoint data assimilation technique for fisheries data analysis. This technique is described for general models and applied to the model underlying the well-known Virtual Population Analysis. It provides a systematic and efficient procedure to obtain stock parameters from a time series of catch-by-age and effort data that can be extended in several directions as desired. The procedure has been implemented in a C++ program which performs well on data from the Pacific halibut fishery *(Hippoglossus stenolepis).* Attention is also paid to the sensitivity of the estimated parameters to various sources of error.

Introduction

In this paper we investigate an adjoint data assimilation technique with the purpose of using it in a fisheries data analysis context. In particular we will focus our attention on age-structured stock assessment methods. Currently there is a large number of these age-structured methods available and we show how the adjoint technique can be used for these methods. For a review of age-structured stock assessment methods we refer to Megrey (1989).

We have applied the adjoint technique to the model underlying Virtual Population Analysis (VPA). This age-structured method for estimating stock sizes and fishing mortalities uses a model that describes the cohort size dynamics backward in time.

Adjoint data assimilation provides a systematic and efficient procedure for the estimation of parameters in complex models by comparing predictions from the model to real world observations.

In the section "Data Assimilation Using the Adjoint Method" we present a general introduction to the adjoint method and include some thoughts on how to perform a sensitivity analysis.

Next, in the section "Virtual Population Analysis," we apply the adjoint method to the model underlying VPA. We have implemented the method in a C++ program which requires a time series of catch-by-age and effort data as input and returns an estimation of stock sizes and a number of other fishery and stock parameters as output. This program is applied to data from the well-documented Pacific halibut *(Hippoglossus stenolepis)* fishery in the "Results" section. In "Concluding Remarks" we discuss a number of possible extensions and ideas for future work.

Data Assimilation Using the Adjoint Method

Introduction

The aim of data assimilation is to combine models and data as efficiently as possible. The most prominent areas of application are meteorology and oceanography, where data assimilation is used to combine numerical models and large data sets (from multiple sources) to improve the forecasts that can be obtained by these models. For reviews see for instance Ghil and Malanotte-Rizzoli (1991) and Lorenc (1986).

Data assimilation methods can roughly be divided into sequential and variational methods. Sequential methods, such as Successive Correction, Optimal Interpolation and Kalman Filtering, update the state of a model sequentially, i.e., for every new observation that becomes available. Variational methods, which include the adjoint method, achieve their aim of assimilating data by fitting the model as closely as possible to the observational data by variation of the model parameters. This is usually done by minimization of a function, called the cost or penalty function. This cost function quantifies the misfit between model and observations.

A cost function can be constructed in several ways. One could, for instance, use Bayesian analysis to derive a function proportional to the likelihood of certain parameters given the observational data. In this paper we directly compare observations and their corresponding model predictions by using a least-squares approach.

The adjoint method is specifically aimed at efficiently computing the gradient of the cost function for large models. Using this gradient, the parameters corresponding to the best fit can be found by a first derivative unconstrained minimization method, e.g., a Quasi-Newton or Conjugated Gradient method.

To compute the gradient a so-called adjoint system for the model is derived. This can be done at several stages of the modeling process; see for instance Giering and Kaminski (1998) for a discussion. We will follow the approach described in Lawson et al. (1995) where the adjoint system is constructed from a model description that is close to computer code. The adjoint system is composed in reverse order compared to the system for the model itself. The computation of the gradient then basically is equivalent to solving the adjoint system. This has two advantages compared to approximations using finite differences. First of all it saves a considerable amount of run-time. It was shown in Baur and Strassen (1983) and Griewank (1989) that evaluation of the adjoint model takes only 2-5 times the computation of the cost function. Approximation of the gradient by finite differences takes at least $n+1$ computations of the cost function, where n is the dimension of the control space, i.e., the number of parameters being estimated. Second, the computed gradient is exact. A final advantage of using an adjoint method for data assimilation is that it makes no assumptions regarding linearity of the model. All that is required is differentiability of the cost function in its arguments and a model that can be represented as a sequence of computations using differentiable functions. The entire data assimilation process is summarized in the scheme of Fig. 1.

Adjoint Systems

Let us consider a data vector d from some space D. It is assumed that each element of this vector corresponds to an observation of a certain quantity. Our aim is to estimate a vector of parameters c (from some control space C with dimension N_c) for a model represented by a system that iteratively produces a vector $x(c)$. We assume that this vector $x(c)$ can be used to compute predictions for the observed quantities in d.

In order to find the most suitable values for the parameter vector we assume that we have a cost function $J[d,x(c),c]$ mapping into R_0^+ of which the values are smaller for parameters corresponding to predictions in better agreement with the observations.

Given the data variables we want to minimize J with respect to the model parameters, i.e., our problem will be

$$\min_{c \in C} J\left[d, x(c), c\right], \tag{1}$$

where d is provided by the observational data. Since $x(c)$ depends on the parameter vector through an entire system, it is difficult to obtain the derivative of this cost function. An approach based on Lagrange multipliers is used to tackle this problem.

As mentioned we use a general model structure that is close to computer code. It is assumed that the model can be represented by the system

$$x_n = f_n\,(x_1, \ldots, x_{n-1}, c),\ n = 1, \ldots, N, \tag{2}$$

i.e., the model is represented as a sequence of calculations where at each step a new scalar quantity is computed which may depend upon all previously calculated quantities and the parameter vector c. The calculations can be considered as corresponding to assignment statements in a computer program.

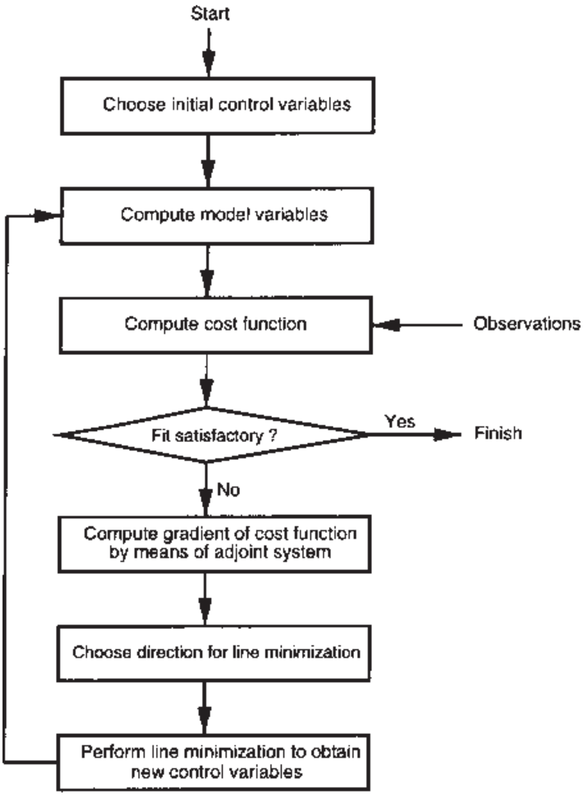


Figure 1. Data assimilation using the adjoint method.

Now define the Lagrangian function

$$L[x(c),c,\lambda] = J[x(c),c] - \sum_{n=1}^{N} \lambda_n [x_n - f_n(x_1,\ldots,x_{n-1},c)], \tag{3}$$

where λ_1, ..., λ_N are the *adjoint variables* and we have dropped the dependency of J on d from our notation since d is assumed to be constant. If x_1, ..., x_N are chosen according to the equations (2) it follows that

$$D_c L[x(c),\ c,\ \lambda] = D_c J[x(c),c] \tag{4}$$

The operator D_c denotes the total derivative with respect to c. This relation holds for all λ, since the functions are in fact the same for all λ. Notice that this is the case even if we choose the adjoint variables dependent on the parameter vector. This is exactly what we will do, i.e., we use $\lambda = \lambda(c)$. Since we are only interested in $D_c J[\mathrm{c}, x(c)]$ we now are free to choose the adjoint variables in such a way that it becomes easier to calculate $D_c L$ which is the same as $D_c J$. Observe that

$$(D_c L)_k = \frac{\partial L}{\partial c_k} + \sum_{n=1}^{N} \left(\frac{\partial L}{\partial x_n} \frac{\partial x_n}{\partial c_k} + \frac{\partial L}{\partial \lambda_n} \frac{\partial \lambda_n}{\partial c_k} \right) \tag{5}$$

for $k = 1, \ldots, N_c$ and that $\partial L / \partial \lambda_n = 0$ if and only if the system equations (2) are satisfied. It follows that if we choose the adjoint variables such that $\partial L / \partial x_n = 0$ for $n = 1, \ldots, N$, we then obtain

$$(D_c J)_k = (D_c L)_k = \frac{\partial L}{\partial c_k}. \tag{6}$$

We have

$$\frac{\partial L}{\partial x_N} = \frac{\partial J}{\partial x_N} - \lambda_N \tag{7}$$

and

$$\frac{\partial L}{\partial x_n} = \frac{\partial J}{\partial x_n} - \lambda_n + \sum_{i=n+1}^{N} \frac{\partial f_i}{\partial x_n} \lambda_i, \qquad n = (N-1), \ldots, 1. \tag{8}$$

So it follows that $\partial L / \partial x_n = 0$ if and only if we choose $\lambda_1, \ldots, \lambda_N$ according to the following *adjoint equations*

$$\lambda_N = \frac{\partial J}{\partial x_N}, \tag{9}$$

$$\lambda_n = \frac{\partial J}{\partial x_n} + \sum_{i=n+1}^{N} \frac{\partial f_i}{\partial x_n} \lambda_i, \qquad n = (N-1), \ldots, 1, \tag{10}$$

that constitute the *adjoint system*. Notice that these adjoint variables depend on the parameters since all partial derivatives are evaluated at the parameter vector *c* for which the derivative is being computed. Using equation (6) we find the following expression for the derivative of the cost function

$$(D_c J)_k = \frac{\partial J}{\partial c_k} + \sum_{n=1}^{N} \frac{\partial f_n}{\partial c_k} \lambda_n, \qquad k = 1, \ldots, N_c, \tag{11}$$

which can be evaluated when we have computed the adjoint variables by means of the adjoint equations.

Implementation

We now present a construction procedure for the cost function gradient that implicitly uses the adjoint system described above. Its description allows an easy translation into computer code.

For every model variable x_n we have an adjoint variable λ_n. We also introduce for every parameter c_k a corresponding variable γ_k. These variables γ_k should contain the corresponding components of the cost function gradient at the end of the procedure, i.e., we want

$$\gamma_k = (D_c J)_k = \frac{\partial J}{\partial c_k} + \sum_{n=1}^{N} \frac{\partial f_n}{\partial c_k} \lambda_n, \qquad k = 1, \ldots, N_c. \tag{12}$$

We first initialize λ_n and γ_k by setting

$$\lambda_n := \frac{\partial J}{\partial x_n}, \qquad n = 1, \ldots, N \tag{13}$$

and

$$\gamma_k := \frac{\partial J}{\partial c_k}, \qquad n = 1, \ldots, N_C. \tag{14}$$

We now consider equations (2) one by one in reverse direction, updating the variables at each step. When finished, λ_n and γ_k will satisfy adjoint equations (9), (10), and (12), respectively. First notice that λ_N already satisfies equation (9). Furthermore it can be seen that each λ_n depends only on model equations and adjoint variables with a larger index. Thus, when arriving at equation n we use the following construction rules:

If $\partial f_n / \partial x_i$ is not equal to zero, then introduce an update statement that adds $\partial f_n / \partial x_i \lambda_n$ to λ_i, i.e.

$$\lambda_i := \lambda_i + \frac{\partial f_n}{\partial x_i} \lambda_n, \qquad i = 1, \ldots, n-1, \tag{15}$$

where the := assignment operator is used to indicate that a variable is being updated, in this case using its own previous value.

If $\partial f_n / \partial c_k$ is not equal to zero, then introduce an update statement that adds $\partial f_n / \partial c_k \lambda_n$ to γ_k, i.e.

$$\gamma_k := \gamma_k + \frac{\partial f_n}{\partial c_k}\lambda_n, \qquad i = 1,\ldots,N_C. \tag{16}$$

Notice the similarity in form of these two construction rules. After proceeding in this manner for N steps, λ_n and γ_k will satisfy equations (9), (10), and (12) as required.

To summarize, each model assignment will correspond to a number of assignments that update adjoint variables of smaller index and variables γ_k. Usually a model assignment depends on only a few of the system variables and parameters, thereby reducing the number of update assignments required. The sequence of initialization and update statements together constitute the procedure to obtain the cost function gradient.

Sensitivity Analysis

To investigate the sensitivity of the resulting parameters to the input data, we adopt a novel approach in fisheries analysis based on the implicit function theorem.

The optimal parameters $\tilde{c}$ for some data vector $\tilde{d}$ satisfy

$$D_c J\left[\tilde{d}, x(\tilde{c}), \tilde{c}\right] = 0. \tag{17}$$

By the implicit function theorem there exists a function $c = g(d)$ in the neighborhood of $\tilde{d}$ which satisfies $\tilde{c} = g(\tilde{d})$ and for which

$$D_d g(\tilde{d}) = -D_c^2 J\left[\tilde{d}, x(\tilde{c}), \tilde{c}\right]^{-1} D_{dc}^2 J\left[\tilde{d}, x(\tilde{c}), \tilde{c}\right]. \tag{18}$$

Using this linear mapping one can approximate how changes δd in the data vector influence the parameter vector corresponding to the optimal fit

$$\delta c = D_d g(\tilde{d})\delta d. \tag{19}$$

To find the directions in the data space D which are particularly important in their effect on the control variables a singular value decomposition of $D_d g(\tilde{d})$ should be made. Notice that to obtain $D_d g(\tilde{d})$ the inverse of the Hessian of J, i.e. $D_c^2 J[\tilde{d}, x(\tilde{c}), \tilde{c}]^{-1}$, with respect to c must be computed.

Virtual Population Analysis

Standard VPA

The model to which we apply the adjoint method is the model underlying the well-known VPA (Gulland 1965). We will first give a short overview of this model and the way in which this model is used in VPA to obtain estimates for fishing mortalities and population sizes. It is assumed that the dynamics of the size of a single cohort, a subpopulation of fish born in the same year, is described by

$$N(y+1,\ a+1) = N(y,a) - C(y,a) - D(y,a). \tag{20}$$

This equation states that the number of fish in a cohort at the beginning of year $y+1$ is equal to the number of fish at the beginning of year y minus the number of fish caught during that year, $C(y,a)$, and minus the number of fish that died from natural causes, $D(y,a)$.

Fishing mortality during a year is modeled by assuming for each fish a fixed probability of being caught per unit of time depending on its age. Natural mortality is modeled in the same way with a fixed probability of dying from natural causes. These probabilities are denoted by $F(y,a)$ and $M(y,a)$, respectively. Usually it is assumed that natural mortality has a fixed known value M independent of year and age. We then have

$$\frac{dC}{dt} = F(y,a)N(t) \quad \text{and} \quad \frac{dD}{dt} = MN(t) \tag{21}$$

and consequently,

$$\frac{dN}{dt} = -\left[F(y,a) + M\right]N(t), \tag{22}$$

where t is time during year y with $N(0) = N(y,a)$ and $N(1) = N(y+1,\ a+1)$. We obtain

$$N(t) = N(0)e^{-[F(y,a)+M]t}, \tag{23}$$

so

$$N(y,a) = e^{[F(y,a)+M]}N(y+1, a+1). \tag{24}$$

From equations (21) and (23) it follows that catch during year y for a certain age group satisfies

$$\frac{dC}{dt} = F(y,a)e^{-[F(y,a)+M]t}N(y,a) \tag{25}$$

with $C(0) = 0$ giving

$$C(t) = \frac{F(y,a)}{F(y,a)+M}(1 - e^{-[F(y,a)+M]t})N(y,a). \tag{26}$$

Using $C(1) = C(y,a)$ and assuming that relation (24) holds, we obtain the *catch equation*:

$$C(y,a) = \frac{F(y,a)}{F(y,a)+M}\left(e^{[F(y,a)+M]} - 1\right)N(y+1,a+1). \qquad (27)$$

It is assumed that data for the catch $C(y,a)$ are available in a catch-at-age table, i.e., for years between y_{min} and y_{max} and ages between a_{min} and a_{max}. If we have an estimate for $N(y+1,\ a+1)$ then we can use the catch equation to obtain the fishing mortality $F(y,a)$. Once $F(y,a)$ is computed, the population size $N(y,a)$ follows from equation (24). Virtual Population Analysis consists of a sequence of such computations that ends when no more catch data are available for the cohort under consideration. This happens when the defined age of recruitment a_{min} or the minimum year y_{min} is reached. To obtain a starting estimate for the cohort size at the maximum age or maximum year for which catch $C(y,a)$ is available, one usually assumes that the terminal fishing mortality F_{term}, i.e., the fishing mortality for that maximum age or year, is known. The final cohort size is then given by

$$N(y_{max},a) = \frac{(F_{term}+M)C(y_{max},a)}{F_{term}\left[1-e^{-(F_{term}+M)}\right]},\ \text{or}\ N(y,a_{max}) = \frac{(F_{max}+M)C(y,a_{max})}{F_{term}\left[1-e^{-(F_{term}+M)}\right]} \qquad (28)$$

depending on whether the cohort under consideration is limited by a maximum age or a maximum year, respectively. Notice that in general population sizes for the maximum year y_{max} are especially interesting, since they correspond to estimates for the most recent stock size. However, as we see here, in standard VPA the estimates for these stock sizes are based on only one catch observation $C(y_{max},a)$ and a terminal fishing mortality F_{term}, so they cannot be considered to be very reliable.

Adjoint VPA

In our approach, which for convenience we shall call adjoint VPA, we will also use equation (24) to describe the cohort size dynamics. However, parameters are now obtained by minimization of a cost function instead of by a sequential procedure. In fact, the method is quite similar to statistical catch-at-age methods that maximize likelihood functions as described in for instance Deriso et al. (1985) and Megrey (1989). We assume that besides catch-at-age data for years y_{min} to y_{max} and age groups a_{min} to a_{max}, an indicator of fishing effort is also available for every year. We make the common assumption that the fishing mortality is a product of an age-dependent term and a year-dependent term:

$$F(y,a) = q(a)E(y), \qquad (29)$$

where $E(y)$ is the indicator of fishing effort during year y and $q(a)$ is an age-dependent catchability coefficient. This catchability coefficient determines the relation between the effort and the actual fishing mortality of the different age groups. The product of equation (29) is well-defined if we assume that the effort indicator has a fixed value, say 100, for y_{min} and that values for subsequent years are expressed in an index relative to the first year. Instead of using a fixed terminal fishing mortality as in VPA we now introduce the final cohort sizes as additional parameters $N_{fin}(y,a_{max})$ and $N_{fin}(y_{max},a)$. Thus, we try to estimate parameters

$q(a)$, for $a = a_{min}, \ldots, a_{max}$,
$E(q)$, for $y = y_{min} + 1, \ldots, y_{max}$,
$N_{fin}(y, a_{max})$, for $y = y_{min}, \ldots, y_{max}$,
$N_{fin}(y_{max}, a)$, for $a = a_{min}, \ldots, a_{max} - 1$,

such that the difference between model predictions and available observations is as small as possible. This is achieved by minimizing the cost function

$$J = \frac{1}{2}\left\{\sum_{y,a}\left[C_{data}(y,a) - C_{model}(y,a)\right]^2 + \kappa\sum_{y}\left[E_{data}(y) - E_{model}(y)\right]^2\right\}, \qquad (30)$$

where κ is a weighting factor that indicates the relative weight attributed to effort observations compared to catch observations. Based on the ideas presented in the section “Data Assimilation Using the Adjoint Method” we proceed with the following algorithm:

1. Take an initial guess of the parameters to be estimated.
2. Apply the cohort dynamics model represented by equation (24).
3. Compare the observations to the model predictions using the cost function equation (30). Model predictions for $C(y, a)$ are obtained from equation (26).
4. Compute the gradient of the cost function for the chosen set of parameter values using the adjoint method.
5. Use this derivative information to find a direction in which to minimize the cost function.
6. Use a line minimization routine to find a new set of parameters corresponding to a better fit to the observations.
7. Repeat step 2 through 6 until the best has been found.

We have implemented this algorithm in C++ code using the procedure described in the “Implementation” section to compute the gradient of the

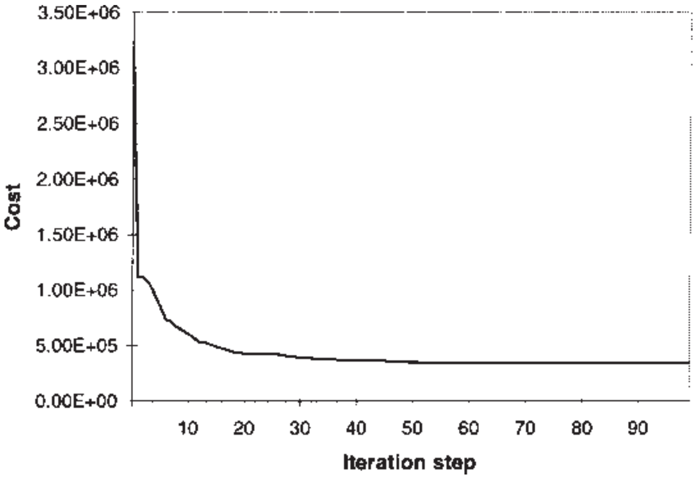


Figure 2. Cost versus iteration number. In 100 iterations 108 variables are estimated. The cost function value for the optimal parameters is about 15% of the cost for the initial parameters.

cost function. To find a direction in parameter space in which to perform a line minimization of the cost function the Polak-Ribiere conjugated gradient model is used, see Polak and Ribiere (1969). Line minimization is achieved by means of Brent's method (Brent 1973). The program computes the optimal parameters $q(a)$, $E(y)$, $N_{fin}(y,a_{max})$ and $N_{fin}(y_{max},a)$ for all years and ages under consideration. It also computes the matrices required to perform the sensitivity analysis described in the "Sensitivity analysis" section.

Results for the Pacific Halibut Fishery

We apply our program for adjoint VPA to data from the Pacific halibut *(Hippoglossus stenolepis)* fishery and compare the results to those from standard VPA. We use catch-at-age data of both the setline and the trawl fishery in International Pacific Halibut Commission regulatory Area 2 from 1935 to 1976 and age classes 8 to 20 as reported in Hoag and McNaughton (1978). Effort data for these years is taken from Hoag et al. (1983). As natural mortality, $M = 0.2$ is chosen for all classes. It is not chosen as a parameter to be varied. In general this parameter is hard to estimate, as our study will confirm.

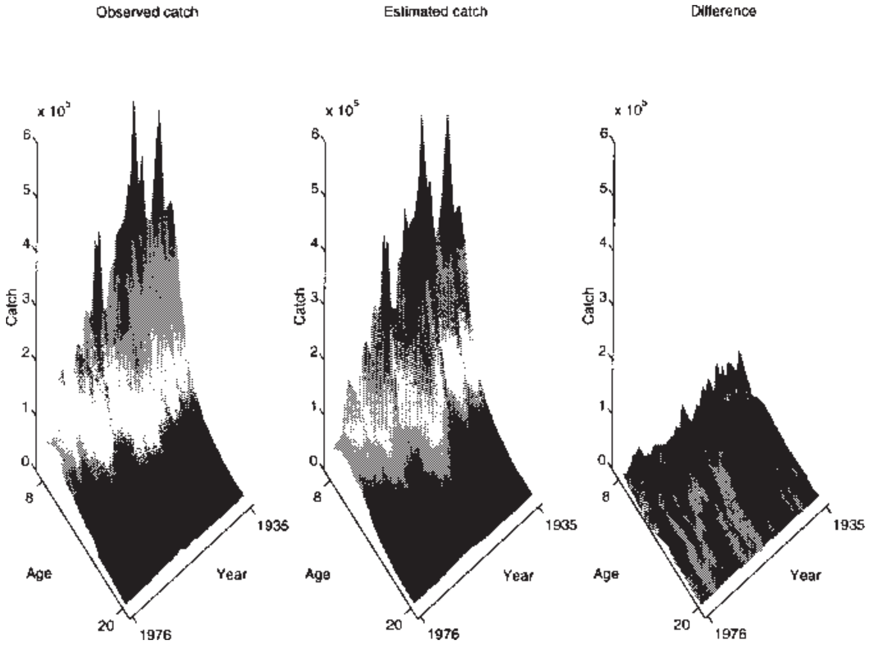


Figure 3. Fit to catch observations for $\kappa = 0$. The average absolute difference relative to the observed catch is 18%.

Performance of Adjoint VPA

The optimization procedure performs well as can be seen in Fig. 2. The procedure converges to the optimal parameter vector in about 100 iterations. The cost function is reduced to approximately 15% of its value for the initial parameters. The fit of the model predictions to the catch observations depends on the value of the weighting factor κ. For small values of κ the effort parameters $E(y)$ can freely be chosen without substantially affecting the cost function. This results in a good fit to the catch data. The average absolute error of the catch predictions compared to the observations is 18%. Figure 3 shows the fit to the catch data for $\kappa = 0$. Since the effort observations are not used in this analysis, it can be expected that the effort values found for the optimal fit differ substantially from the observed values for the effort indicators. This is shown in Fig. 4. Choosing larger values of κ will make the optimal effort variables come closer to the observed values, see again Fig. 4. For very large values of κ the effort variables are forced to correspond to their observed values, leaving less freedom to fit to the catch data. This is shown in Fig. 5. The average absolute error of the catch predictions compared to the observations is now 20%.

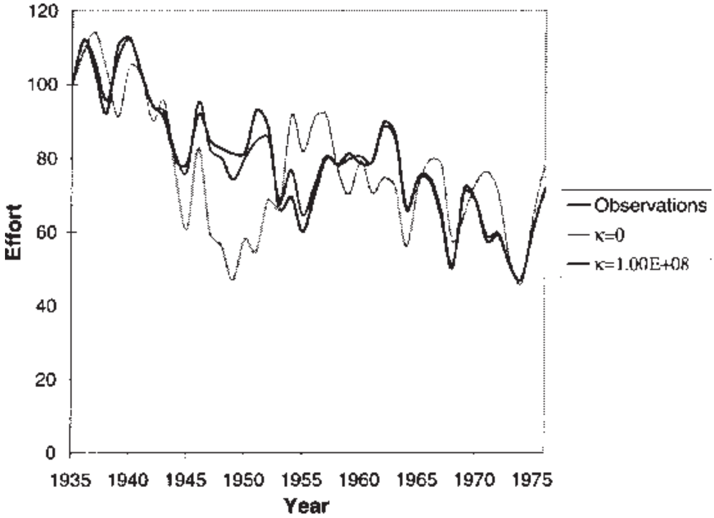


Figure 4. *Effort comparison between observations and model predictions. For $\kappa = 0$ deviation from observations is largest. For larger values of κ predictions and observations are in better agreement. For $\kappa = 1.0 \times 10^{10}$ predictions and observations are the same.*

Comparison to Standard VPA

Estimated population sizes are close to those obtained by standard VPA. Due to separation of fishing mortality into effort and catchability coefficients, adjoint VPA cannot predict observed catches exactly. To get some idea of the performance of the catch fitting procedure, we compare our results to those obtained for standard VPA which is also submitted to a separability assumption equation (29). First standard VPA is performed. To obtain catchability coefficients every fishing mortality $F(y,a)$ is divided by its corresponding effort indicator $E(y)$. The resulting catchability coefficients depend on both year and age. To obtain catchability coefficients which depend on age only, they are averaged over time. These coefficients can be used to obtain new fishing mortalities using equation (29). The catches corresponding to these fishing mortalities are no longer the same as the observed catches.

It turns out that the cost function for this set of parameters is about six times as large as the value obtained by adjoint VPA. The average absolute error between predicted catches and observations is now 35% which is about twice as large as the average absolute error relative to the observed catch for adjoint VPA.

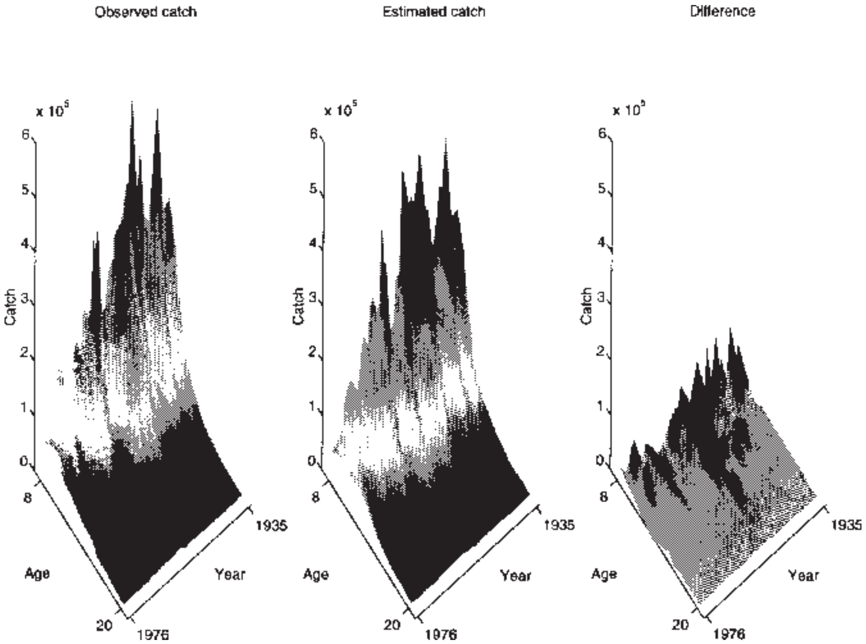


Figure 5. Fit to catch observations for $\kappa = 1.0 \times 10^{10}$. The average absolute difference relative to the observed catch is 20%.

In Fig. 6 stock size estimates for a typical cohort are compared for standard VPA, adjoint VPA, and the approach described above using averaged catchability coefficients.

Sensitivity Considerations

We have repeated our procedure for various values of the natural mortality parameter and found no significant change in the minimum value of the cost function; see Table 1. We conclude that additional information is needed to estimate this parameter or that it should be estimated by other means.

We have also performed the sensitivity analysis described in the "Sensitivity Analysis" section. To test which catch-at-age table entries are particularly important in their effect on the optimal parameters, all parameters are first scaled to the same order of magnitude. Next we use the linear mapping from equation (18) to investigate the change in the parameters resulting from a change in the catch data. For each catch-at-age table entry the norm of the difference in the optimal parameters as a consequence of a unit change in that entry is plotted. See Figs. 7 and 8 for $\kappa = 0$ and $\kappa = 1.0 \times 10^{10}$, respectively.

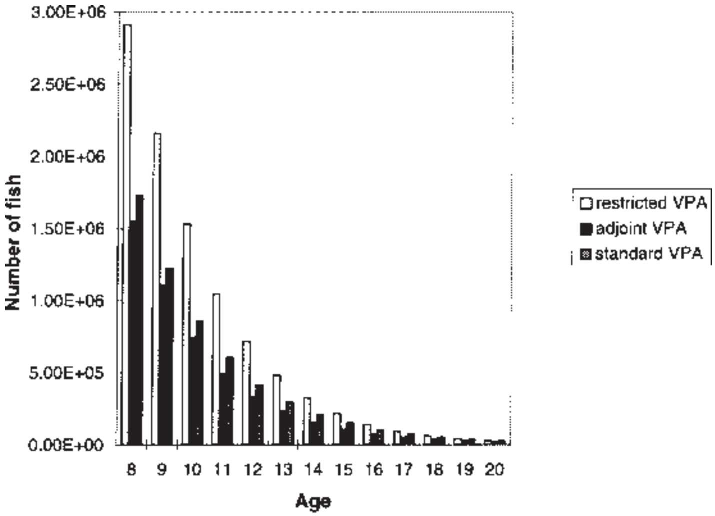


Figure 6. Comparison between restricted VPA and adjoint VPA. Large relative errors in predicted catches for restricted VPA also result in large deviations in predicted population sizes from those obtained by standard and adjoint VPA.

Table 1. Optimal cost function values for various natural mortality coefficients.

M	Minimum cost
0.05	341,582
0.1	339,790
0.15	340,219
0.2	341,172
0.25	342,215
0.3	343,020

Relative changes in the optimal cost function values are small for variations of the natural mortality coefficient in the value region of interest.

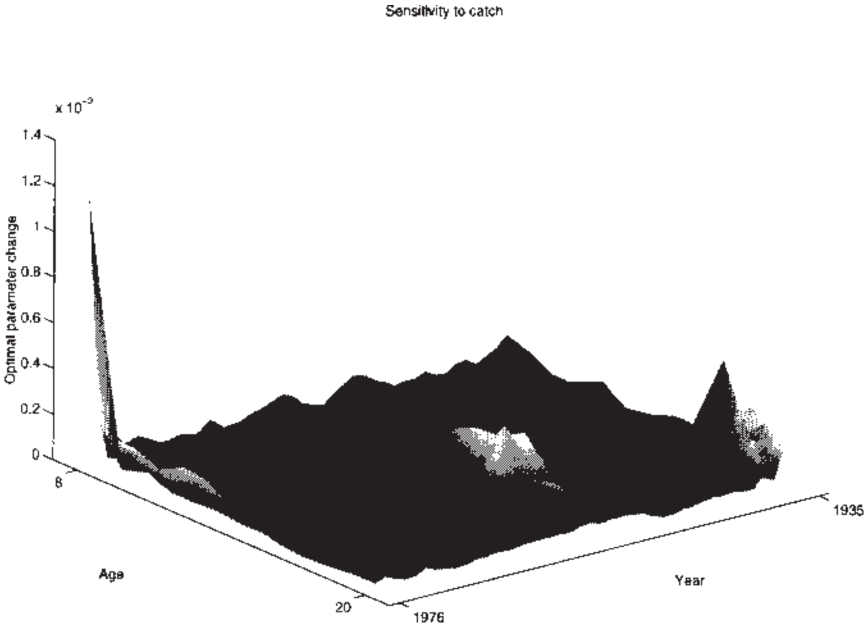


Figure 7. *Sensitivities for $\kappa = 0$. For each observation the norm of the vector change, in the scaled optimal parameters due to a unit change in the observation value, is plotted. The entries in the catch-at-age table corresponding to incomplete cohorts have the largest impact on the optimal parameters. For $\kappa = 0$ changes in the effort observations have no effect.*

Notice that catch measurement errors for table entries corresponding to incomplete cohorts cause the largest change in optimal parameter values. A similar analysis is performed for the effort time series. Measurement errors in effort observations have a negligible effect on the optimal parameters for $\kappa = 0$. They do, however, become important for large values of κ. See Fig. 8. The optimal parameter that is mainly affected by a change in an effort observation is the effort system variable corresponding to this observation.

Next, as another test, all catch observations are increased by 1%. We can compute the resulting relative change in the optimal parameter vector. It turns out that most parameters change approximately 1%. Some of the final cohort sizes change up to 3%. For $\kappa = 1.0 \times 10^{10}$ the changes are somewhat smaller.

Using a singular value decomposition of the linear mapping of equation (18) we can find the directions in data space that result in the largest change of the parameters. If we use an observation error vector of length

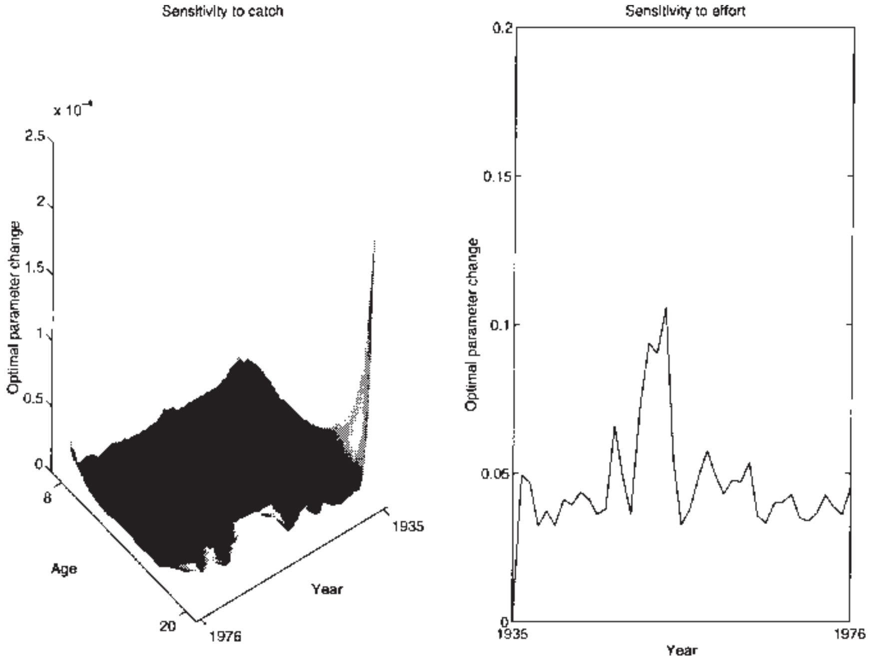


Figure 8. Sensitivities for $\kappa = 1.0 \times 10^{10}$. Impact of a unit change in the catch entries is now smaller than for $\kappa = 0$. Especially the system effort variables are now hardly affected by changes in catch observations, since they are determined by the effort observations. Changes in effort variables now have a substantial effect.

equal to the vector corresponding to a 1% increase in catches, the changes in the parameters are very large. This is as expected, however, since this error observation vector corresponds to large relative changes in some of the catch observations.

Changing the effort observations by 1% has no effect for $\kappa = 0$ and results in a change of exactly 1% for $\kappa = 1.0 \times 10^{10}$.

Concluding Remarks

We have implemented a procedure for estimating fishery parameters using the same hindcasting model as in VPA. Our procedure is similar to a number of statistical age-structured methods that are currently employed for stock assessment. For these methods the adjoint method can be a useful technique for the estimation of the parameters. The main advantage of the adjoint method is its efficient and reliable computation of the cost function gradient which allows the simultaneous estimation of a large number of parameters.

This paper describes how to implement the adjoint method for relatively simple models. For more complex models it is more practical to use software packages for reverse automatic differentiation, such as ADOL-C for C++. Also for Fortran several packages are available. For an overview see Griewank (1989).

Adjoint VPA performed well on data from the Pacific halibut fishery. For these data a total of 108 parameters are estimated in about 100 iterations. The results are used in Grasman and Huiskes (1997) for a stochastic model of halibut recruitment. We would like to mention again that adjoint VPA is implemented as an example and the approach can be used for estimating parameters in any model that is based on a differentiable algorithm.

The procedure can be extended in a number of ways. Additional observational data, e.g., from recruitment studies, can be easily incorporated into the cost function. Also cost functions based on different assumptions regarding the error structure of the observations can be used.

Another interesting possibility lies in the fact that the code for minimizing a cost function can be adapted to find extremes of arbitrary functions depending on the system and model parameters. One could for instance define a function which describes total catch and maximize this function in the effort variables. This would only require a change in the initialization part of the adjoint method code.

In future work we intend to apply a similar procedure to physiologically structured population models. In these models knowledge about mechanisms at the individual level is translated to effects on the population level. For example, this knowledge can be used to model the effects of food availability, length structure of the population, and fecundity. Although this approach requires data from a large number of sources, such a study may provide more realistic estimates of fishery parameters and may even lead to reliable natural mortality and recruitment estimates.

Acknowledgment

The author thanks Johan Grasman, Maarten de Gee, and Jon Schnute for valuable comments.

References

Baur, W., and V. Strassen. 1983. The complexity of partial derivatives. Theoretical Computer Science 22:317-330.

Brent, R.P. 1973. Algorithms for minimization without derivatives. Prentice Hall, Englewood Cliffs.

Deriso, R.B., T.J. Quinn II, and P.R. Neal. 1985. Catch-age analysis with auxiliary information. Can. J. Fish. Aquat. Sci. 42:815-824.

Ghil, P.E., and P. Malanotte-Rizzoli. 1991. Data assimilation in meteorology and oceanography. Adv. Geophys. 33:141-266.

Giering, R., and T. Kaminski. 1998. Recipes for adjoint code construction. Association for Computing Machinery, Transactions on Mathematical Software. In press.

Grasman, J., and M.J. Huiskes. 1997. A stochastic model of recruitment and critical stock size for a halibut population. Wageningen Agricultural University, Wageningen, The Netherlands. Technical Note 97-08.

Griewank, A. 1989. On automatic differentiation. In: M. Iri and K. Tanabe (eds.), Mathematical programming: Recent developments and applications. Kluwer Academic Publishers, Amsterdam, pp. 83-108.

Gulland, J.A. 1965. Estimation of mortality rates. Annex to Arctic Fisheries Working Group report, ICES C.M. Doc. 3, Copenhagen.

Hoag, S.H., and R.J. McNaughton. 1978. Abundance and fishing mortality of Pacific halibut. Cohort Analysis 1935-1976, Int. Pac. Halibut Comm., Seattle.

Hoag, S.H., J.R. Myhre, G. St-Pierre, and D.A. McCaughran. 1983. The Pacific halibut resource and fishery in regulatory Area 2, I. Management and biology. Int. Pac. Halibut Comm. Sci. Rep. 67, Seattle.

Lawson, L.M., Y.H. Spitz, E.E. Hofmann, and R.B. Long. 1995. A data-assimilation technique applied to a predator-prey model. Bull. Math. Biol. 57(4):593-617.

Lorenc, A.C. 1986. Analysis methods for numerical weather prediction. Q. J. R. Meteorol. Soc. 112:1177-1194.

Megrey, B.A. 1989. Review and comparison of age-structured stock assessment models from theoretical and applied point of view. In: B.A Megrey and E.F. Edward (eds.), Mathematical analysis of fish stock dynamics. Amer. Fish. Soc. Symp. 6:8-48.

Polak, E., and G. Ribiere. 1969. Note sur la convergence de methodes de directions conjugées. Revue Française d'Automatique, Informatique, Recherche Operationelle 16-R1:35-43.

Examination of a Length-Based Population Analysis

Takashi Matsuishi
Hokkaido University, Faculty of Fisheries, Hokkaido, Japan

Virtual population analysis (VPA) and cohort analysis are widely used for estimating fish population size from catch-at-age data. Generally age determination is so laborious that catch-at-age is often determined from catch-at-length data. Errors in this conversion, however, can be significant.

Fishery managers often control the selectivity of fishing gear. To manage fish stocks, the managers need population-at-length estimates instead of population-at-age. Another conversion is needed to estimate population-at-length from population-at-age, which is calculated from catch-at-age using VPA. Estimation errors in this conversion can also be significant.

If population at length can be estimated directly from catch-at-length, these two conversions can be omitted, and thus the estimation accuracy will be improved.

In this study, I utilize a method of population estimation based on catch-at-length called Length-based Population Analysis or LPA. LPA estimates population-at-length from the largest to the smallest length class, by avoiding the estimation errors of recruitment. I also examine the validity and accuracy of LPA using computer simulation, and present the result of two applications.

Data

Data required to estimate the population-at-length by LPA are natural mortality M, catch-at-length C_{ij} for length class j at year i, and matrix $\boldsymbol{P}$.

Matrix $\boldsymbol{P}$ is a matrix in which the (j, k) component is the probability that the individual at year $t+1$ in length class k belonged to the length class j. If the history of growth is available, this probability is estimated directly from growth data (method 1). In the population analysis of abalone mentioned later, growth history can be determined from shell rings.

If growth history is not available, this probability is estimated from the average and variance of length at each age (method 2). In the population analysis of Walleye pollock, also described later, this method was used.

Procedure

The LPA procedure is as follows. First, give initial values for yearly fishing mortality, $\tilde{f}_j$ and for selectivity $\tilde{s}_i$. Selectivity is assumed to be constant and independent from the year. Fishing mortality $\tilde{F}_{i,j}$ is given as the product of $\tilde{f}_j$ and $\tilde{s}_i$. Total mortality $\tilde{Z}_{i,j}$ is given as the sum of $\tilde{F}_{i,j}$ and natural mortality M.

Calculate the population of all length classes for the most recent year data was collected, and for the largest length class for every year data was collected from observed catch-at-length $C_{i,j}$, fishing mortality $\tilde{F}_{i,j}$, and total mortality $\tilde{Z}_{i,j}$. Then calculate the population-at-length $\tilde{N}_{i,j}$ for the previous year using matrix $\boldsymbol{P}$ and $\tilde{Z}_{i,j}$ sequentially. From $\tilde{N}_{i,j}$, $\tilde{F}_{i,j}$, and $\tilde{Z}_{i,j}$, calculate the expected catch-at-length for each year $\tilde{C}_{i,j}$. To compare estimates with the observed catch, calculate the sum of square residual SS and minimize SS using a nonlinear optimization method, such as the simplex search method. When SS is minimized, the parameters $\hat{f}_j$, $\hat{s}_i$, and $\hat{F}_{i,j}$, $\hat{N}_{i,j}$ are the final estimators.

Validation of LPA Using the Simulation Method

Factors such as natural mortality and growth will fluctuate in relation to biological and physical conditions. Fishing pressure and fishing-gear selectivity will also fluctuate. To examine the effect of these fluctuations on the estimators, I developed a data generator to generate sham catch-at-length data with various fluctuations. I examined the validity by estimating population-at-length from the sham data using LPA, and compared the original population set to the generator and estimators from LPA.

First, I generated sham catch data without any fluctuation, and estimated population-at-length by LPA. The estimate of population-at-length was very close to the estimate of the original population. When fluctuations are added to yearly fishing mortality and selectivity, the estimators are not largely biased. For fluctuations in natural mortality, the estimators are robust.

These results suggest that LPA is robust for various fluctuations that will occur in real catch data.

Comparison with the Former Methods

Formerly, to estimate catch-at-age from catch-at-length, age length key (ALK) or length frequency analysis (LFA) has been used. From the estimated catch-at-age, population-at-age has been estimated using VPA. Then, population-at-length has been estimated by converting the population-at-age to population-at-length.

To compare these two former methods with LPA, populations at length were estimated from sham catch-at-length data using ALK and VPA, LFA and VPA, and LPA.

From the sham data that included a CV 5% fluctuation in all parameters, the 50% point of the relative error in the estimate of the population for the fifth length class in the sixth year was smallest for LPA (0.0095) followed by LFA and VPA (–0.0987), and ALK and VPA (0.142).

These results suggest that LPA performs equally well or better than the former methods.

Application to Abalone Fishery in Akita Prefecture

To examine the validity and effectiveness of LPA for an actual population, I applied LPA to the abalone diving fishery in Konoura, Akita prefecture.

Annual rings of 356 shells of abalone caught in this area were measured and matrix $\boldsymbol{P}$ was calculated using method 1. I also calculated the matrix using method 2. Catch-at-length data was generated from the length composition data recorded by the Institute for Fisheries and Fisheries Management in Akita Prefecture and catch numbers in 1986 to 1993. Natural mortality was set at 0.102, as estimated by the Institute for Fisheries and Fisheries Management.

Using method 2, the estimated annual number of abalone with a shell length greater than 10 cm ranged from 44,000 (1987) to 68,000 (1990).

The population dynamics are quite similar to the estimates of Zhao et al. (1993), who used ALK and VPA in the same area.

The results of retrospective analysis indicate the estimates were stable.

Application to Walleye Pollock Fishery in Funka Bay

I next examined the gillnet and trawl fishery of walleye pollock, *Theragra chalcogramma*, whose spawning ground is in Funka Bay in the southwestern part of Hokkaido. I used the length composition data recorded by the Muroran branch of the Hokkaido Hakodate Fisheries Experimental Station from 1971 to 1986. Matrix $\boldsymbol{P}$ was calculated using method 2 from the age and length data of 4,460 individuals measured in 1984. Natural mortality was set to 0.3, the value calculated by VPA for the walleye pollock in this area.

The estimated annual population size ranged from 195,000,000 (1978) to 513,000,000 (1971). The estimated parameter of the gear selectivity increased rapidly at 35 cm, which agrees to the real selectivity of the fishing gear used this area.

For a population with an unknown growth history, it is suggested that LPA is applicable.

Conclusion

From the results of both simulation and application, it is suggested the LPA is LPA is a valid and effective method to estimate population-at-length from catch-at-length data. For populations that are difficult to age, it is possible to estimate population size using LPA.

Acknowledgment

I thank Dr. Makoto Shimizu and Dr. Hirohisa Kishino of Tokyo University for useful discussions. This work is funded in part by a Grants-in-Aid for Scientific Research from the Ministry of Education, Science, Sports and Culture of Japan. This study is in preparation to submit to *Fisheries Science*.

Reference

Zhao, B., E. Tanaka, and J. Yamada. 1993. Stock assessment of abalone in Kisakata, Akita prefecture. Nippon Suisan Gakkaishi 59:1823-1830.

Incorporation of Predation into a Population Assessment Model of Eastern Bering Sea Walleye Pollock

Patricia A. Livingston
National Marine Fisheries Service, Alaska Fisheries Science Center, Seattle, Washington

Richard D. Methot
National Marine Fisheries Service, Northwest Fisheries Science Center, Seattle, Washington

Abstract

The effect of predation by adult walleye pollock *(Theragra chalcogramma),* Pacific cod *(Gadus macrocephalus),* and northern fur seals *(Callorhinus ursinus)* was incorporated into a stock assessment model of walleye pollock in the eastern Bering Sea. Previous studies showed these three species are the primary predators of young walleye pollock. The stock synthesis model was used to incorporate the three predators as sources of mortality in a single-species catch-at-age analysis. Predators were entered into the model by defining (1) a time series of predator abundance (i.e., effort) over the entire modeled period, (2) a series of pollock consumption data (i.e., catch per unit effort) for each predator for each year when food habits data were available, and (3) the age composition of pollock consumed for each year where food habits data were available. The instantaneous annual predation mortality rates estimated from cod and fur seal predation were small, ranging from about 0.02 to 0.04. However, mortality rates of age-1 pollock due to cannibalism were high.

Adding cannibalism on age-1 fish increased model estimates of the number of age-1 pollock in recent years, which changed the apparent stock-recruitment relationship at age 1 from one with declining recruitment at high spawning stock sizes to one where recruitment is more asymptotic at high spawning stock sizes. Our analysis suggests that mortality due to cannibalism on pollock may be an important factor in determining

recruitment into the fishery at age 3. Adding predation, and particularly cannibalism, to the model improved the fit of the stock recruitment relationship at age 1. The largest outliers from the estimated relationship were years with generally warmer temperatures and higher inshore transport of surface waters. Oceanographic conditions and predation both appear to play an important role in pollock year-class success. Simultaneously accounting for both these factors in future models of pollock population dynamics may improve our understanding of the relative importance of those factors in determining pollock recruitment.

Introduction

Predation is considered to be an important source of mortality influencing the dynamics of some groundfish populations. However, in most traditional stock assessment models, predation in the form of natural mortality is entered as a constant rate across time. Multispecies methods that include predation, such as multispecies virtual population analysis (MSVPA), are being explored by stock assessment biologists. However, these models are not yet widely used to provide management advice because of the large number of species interactions that are included and the difficulty in validating model estimates of these interactions (Sissenwine and Daan 1991, Brugge and Holden 1991). The stock synthesis model (Methot 1989, 1990) offers a framework to incorporate predators as sources of mortality in a single-species catch-at-age analysis and to simultaneously examine the model fit to predation data sources along with other more traditional data sources used to assess population size. This approach provides not only a more objective basis to evaluate model estimates related to predation but it also provides a way to examine predation in a standard single-species stock assessment that may be more acceptable to stock assessment scientists and managers eager to move toward ecosystem-based management but in an incremental fashion. Walleye pollock *(Theragra chalcogramma)* is an important groundfish species in the eastern Bering Sea which supports a large commercial fishery and also serves as an important prey of other groundfish, marine mammals, and birds (Kajimura and Fowler 1984; Livingston 1989a, 1991, 1993). Here we investigate the effect of including predation by adult walleye pollock, Pacific cod *(Gadus macrocephalus),* and northern fur seals *(Callorhinus ursinus)* into a population model of walleye pollock in the eastern Bering Sea. These predators are the primary sources of predation mortality for young walleye pollock (Livingston 1993).

Methods

Population Model

The stock synthesis model was used for this study of population dynamics of walleye pollock. This model simulates the dynamics of an age-structured exploited population and the processes by which the fisheries and surveys affect and observe this population. The model estimates parameters (such as annual recruitment, fishing mortality, selectivity patterns) to simultaneously fit various kinds of data (survey biomass, fishery effort or catch per unit effort (CPUE), fishery and survey size and age composition) by maximizing a log-likelihood function which includes components for each type of data. In practice, the various types of available data may indicate patterns that are inconsistent with each other. Here the model is perhaps even more useful by providing a framework for identifying and exploring these discrepancies.

The age-structured version of stock synthesis can be briefly described by the following parameters and equations:

N_{ya} = population numbers in year y and age a (Note that multiple periods within year can be defined.)

M = instantaneous rate of natural mortality (may be age-specific)

W_{yat} = body weight at age observed for fishery (or predator) of type t

S_{at} = fishery (or predator) and age-specific selectivity (Selectivity patterns can change over time.)

E_{yt} = fishing mortality rate in year y, fishery (predator) t for fully available ages (e.g., ages for which $S_{at} = 1.0$)

$F_{yat} = E_{yt}S_{at}$ = fishing mortality rate by year, fishery (or predator), and age

$Z_{ya} = M + \sum_t F_{yat}$ = total mortality rate

$c_{yat} = N_{ya}\left(\frac{F_{yat}}{Z_{ya}}\right)\left(1-e^{-Z_{ya}}\right)$ = catch at age for fishery (predator) t

$C_{yt} = \sum_a c_{yat}W_{yat}$ = catch biomass for fishery (predator) t

$N_{y+1,a+1} = N_{ya}e^{-Z_{ya}}$ = (with accumulation in the terminal age)

Usually, the model is given a time series of catch biomasses, C_{yt}, which are assumed to be known with high precision, and the model estimates the E_{yt} which will exactly match these observed catch biomasses. If effort data are available for a particular fishery, these data are assumed proportional to the E_{yt} by a catchability coefficient Q_t. Equivalently, if CPUE data are available for a fishery, these data are assumed proportional to the mean available biomass. The model estimates Q_t as a nuisance parameter; it simply scales the estimated E_{yt} to the units of the observed fishing effort.

Incorporating predators as additional, independent sources of removals similar to a fishery is a logical extension of the stock synthesis model. The abundance of a predator is analogous to the amount of fishing effort expended by a fishery. An individual predator's consumption of pollock at different pollock densities is analogous to the CPUE and allows the estimation of the predator's functional feeding response curve. The age composition of the pollock in the guts of the predators is the catch age composition, and the model's comparison of this age composition to the estimated age composition of the pollock population enables calculation of the predator's selectivity curve for pollock of different ages.

The procedure for estimating the predators' E_{yt} was modified because the total removals (e.g., amount eaten) by each predator is not known over the whole time series being modeled. Instead, the model was provided with estimates of predator abundance, P_{yt}, over the whole time series because these estimates are usually available. Observations of each predator's consumption rate of pollock, R_{yt}, (prey consumed per predator) are entered for years in which food habits data for that predator are available. The model then estimates the Q_t for each predator as the constant of proportionality that maximizes the log-likelihood for the deviations between the pollock biomass available to the predator (i.e., the pollock biomass obtained after accounting for the age-specific selectivity of pollock by the predator) and the predator's R_{yt}. The estimate of predation mortality rate, E_{yt}, is then equal to $P_{yt}Q_t$.

This process models each predator's feeding response to changes in pollock density as a linear function with a slope of Q, analogous to Holling's (1959) Type I predator functional feeding response curve without an upper asymptote. This means that one must check model outputs to determine that the prey biomass above which a predator would be satiated has not been attained over the period modeled. An extra parameter, which modifies Q as a power function of prey biomass, is also available if there is enough contrast in the data for its estimation. This extra parameter would produce a functional feeding response curve similar to a Type II curve. The software does not yet have the ability to fit a Type III sigmoidal-shape response curve, although it may be modified in the future to accommodate this function. Also, the effects of alternate prey abundance on feeding response is not presently considered. Predators are presently entered as a single biomass value in each year, which has been standardized to account for predator age-dependent factors in consumption efficiency of pollock. Explicitly including age-specific predator information would be a further improvement to explore in future versions of this type of model.

Model Parameterization and Data Sources

The modeled period is from 1964 to 1995 and explicitly includes pollock ages 1 through 14, with a plus group for ages 15 and older.

The fisheries data sources used were total annual catch biomass from all nations and vessel classes, age composition of the catch biomass, and effort data from the Japanese trawl fishery during 1964 to 1973 (Table 1). Auxiliary sources of information from surveys included estimates of the biomass and age composition of age 3+ pollock from the combined bottom trawl-hydroacoustic surveys conducted on a triennial basis since 1979 and an index of age-1 pollock numbers from the bottom trawl surveys conducted each summer.

Fishery and predator selectivity-at-age was modeled as the product of an ascending logistic curve for young pollock and a descending logistic curve for older pollock (Methot 1990). For the fishery, the descending curve was allowed to have different parameters between two periods: 1964 to 1972 when the fishery was developing, and 1973 to 1991 when pollock was the main fishery target. The inflection age for the fishery ascending curve was allowed to have a year-specific value to track changes in the fishery's impact on young pollock. For the predator selectivity curves, only the descending curve was used, i.e., young pollock were defined to have a selectivity of 1.0 for the predators. Aging error was incorporated using percent agreement between readers from Kimura and Lyons (1991).

Predation data included an estimate of predator abundance over the modeled period (1964 to 1995) (Fig. 1), predator consumption of pollock (per unit predator) for years where food habits data were available (Table 2), age composition of pollock consumed during years where food habits data were available, and the mean weight at age of pollock consumed by each predator. Predator abundance for cod and fur seal was input to the model, while abundance on cannibalistic age 5+ pollock is generated within the model.

Cannibalism data were obtained from Livingston (1991) and Livingston et al. (1993). These data indicate that only age-5 and older fish were significantly cannibalistic on young pollock. Cannibalism on age-2 and older pollock occurs but at such a low level that it was not included in this analysis. Because cannibalism estimates included consumption of age-0 and age-1 fish, the estimates were revised downward to include cannibalism on age-1 fish only. Further, biomass of age-1 pollock consumed per 1,000 metric tons of adult pollock for 1985 to 1989 (Table 2) was converted to the number of age-1 pollock consumed per unit weight of cannibals. This strategy was employed for implementing cannibalism because it allowed for interannual differences in mean weight at age of cannibalized pollock. Stock synthesis uses inputs of pollock weight at age in each data source (e.g., fishery, survey, cod, fur seal, cannibals), which are typically kept constant across all modeled years, to convert from biomass to number. Therefore, the mean weight of age-1 pollock consumed by cannibals was set equal to one kilogram to keep pollock consumption per unit cannibal estimates in terms of numbers .

Northern fur seal abundance (Fig. 1) for the modeled time period was considered to be the biomass of age-4 and older females (standardized for differences in energy consumption of lactating versus non-lactating animals per Perez and Mooney, 1986) derived from annual estimates of pups born on St. Paul and St. George Islands in the Bering Sea (Antonelis et al. 1996, Lander and Kajimura 1982). For the period from 1975 to 1995, pup abundance was estimated in years without counts by taking the average of the survey counts from the years before and after the missing count. Pup counts on St. George Island from 1964 to 1974 were estimated as 0.25

Table 1. Data sources, coefficients of variation (CV) and assumed error distributions used in the population models of walleye pollock in the eastern Bering Sea.

Data source	Years	CV	Error distribution
Annual catch biomass	1964-1995	0.03	Lognormal
Age composition of the catch	1964-1995	0.15	Multinomial
Japanese trawl fishery effort	1964-1973	0.30	Lognormal
Age 3+ biomass from combined bottom trawl and hydroacoustic surveys	1979, 1982, 1985, 1988, 1991, 1994	0.20	Lognormal
Age 3 + age composition from combined bottom trawl and hydroacoustic surveys	1979, 1982, 1985, 1988, 1991, 1994	0.15	Multinomial
Index of age 1 number from bottom trawl survey	1979-1996	0.50	Lognormal
Pacific cod biomass	1976-1995	0.03	Lognormal
N. fur seal biomass (age 4+ females)	1964-1995	0.03	Lognormal
Pacific cod consumption of pollock	1984-1989	0.60	Lognormal
N. fur seal consumption of pollock	1973-1974, 1982, 1985	0.60	Lognormal
Pollock consumption of age-1 pollock	1985-1989	0.60	Lognormal
Age composition of pollock consumed by cod	1984-1989	0.15	Multivariate
Age composition of pollock consumed by fur seals	1973-1974, 1982, 1985	0.15	Multivariate

times the count on St. Paul Island for each year, as advised by York (1987). Biomass of pollock consumed per unit mass of fur seals (corrected for consumption of age-0 pollock) (Table 2) and the age composition of pollock in fur seals were estimated for 1973, 1974, 1982, and 1985 because these were years where fur seal food habits and pollock prey age composition data were available (AFSC 1974, 1975; McAlister and Perez 1976; Sinclair 1988).

Pacific cod abundance estimates (Fig. 1) were derived from mid-year biomass estimates at age from stock synthesis (Thompson 1996). Because cod exhibit size-related changes in their consumption of pollock, total cod biomass was standardized for size differences by deriving relative

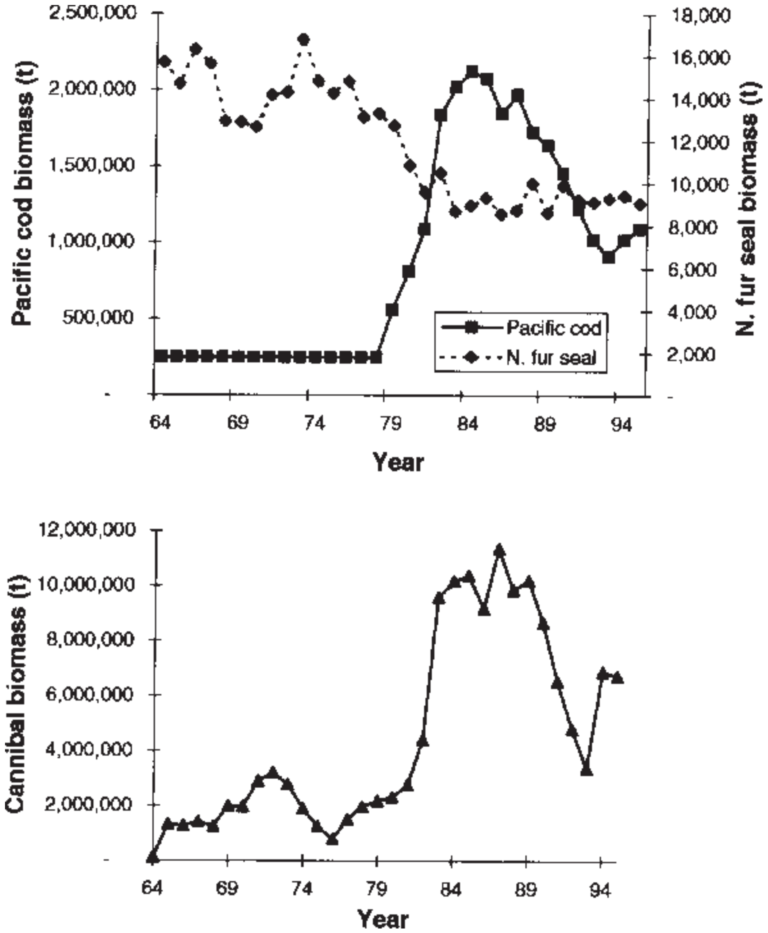


Figure 1. Pacific cod, female northern fur seal (top) and pollock cannibal (bottom) biomass estimates in the eastern Bering Sea used in the pollock population model with predation.

consumption efficiency factors for each size group relative to the largest size group using analysis of variance. Biomass of pollock consumed per unit mass of cod (corrected for consumption of age-0 pollock) (Table 2) and the size composition of pollock consumed by cod were modeled using data from Livingston (1991) and Livingston et al. (1993).

Error levels for each type of data are necessary in order to calculate the log-likelihood function, but these error levels are not known for the predator data. Presumably, the levels of variability for the predator data are greater than for the fishery and survey data. We set error levels for the predator data by examining the residual error in preliminary model runs with nominal levels of error for these data. This resulted in the nominal sample size for fishery age composition data to be set at 200, 130 for survey age composition, and 10 for predator age composition data. Error levels in predator catch per unit data were explored, with the coefficient of variation ranging from 0.15 to 0.75. A level of 0.6 was chosen to best reflect the uncertainty in this data source. Error levels and assumed error structure for each data source are shown in Table 1. Preliminary sensitivity analyses were previously performed on the predator data sources, examining the effects of different emphasis levels on predator data sources to model fits of other model data sources (Livingston and Methot 1996). These runs showed that cannibalism data were compatible with other data sources because increasing the emphasis level of cannibalism data up to 5 did not substantially change model fits and outputs. In fact, increasing emphasis on cannibalism data sources enhanced model fits to some other data sources. However, increasing emphasis on cod or fur seal data inputs degraded model fits to other more certain data sources. This indicates that further improvements in cod and fur seal predation estimates may be required through additional sampling.

In the baseline model without predation, natural mortality was 0.6 yr^{-1} for age-1 pollock, 0.45 yr^{-1} for age-2 and 0.3 yr^{-1} for age-3 and older pollock (Wespestad 1996). After exploring levels of baseline natural mortality ranging from 0.1 to 0.6 yr^{-1} in the model with predation, the level of residual natural mortality was set at 0.3 for all ages, with the assumption that the remaining natural mortality consists of predation mortality from the predators included in the model.

Results

The predator model improved the fit to the fishery data and slightly degraded the fit to the survey information (Table 3). Estimated numbers of pollock at age 1 were much higher in the model with predators than in the standard (no predator) model, particularly for 1979 to 1995 (Fig. 2). Estimates of number at age 3, however, were similar between the two models. This is not unexpected because both models were tuned to the same age 3+ survey data. Exploitable biomass estimates (Fig. 2) were slightly higher in the model including predation.

As mentioned previously, availability of pollock to cannibals was fixed at 1 for age-1 pollock and 0 for older pollock. The model estimates of pollock availability to Pacific cod and northern fur seals were similar between the two predators with availability declining from 1.0 for age-1 pollock to less than 0.1 for age 15+ pollock. Pollock availability with increasing pollock age declined somewhat faster for fur seals than for cod.

The combined predation mortality of age-1 pollock due to cannibals, cod, and fur seals ranged from 0.1 to over 1.4 yr^{-1} (Fig. 3), with the majority of age-1 predation mortality consisting of cannibalism. Mortality of

Table 2. **Rates of walleye pollock consumption per unit predator (1,000 t pollock per 1,000 t predator) used as input to the model containing predators**

Year	N. fur seal	Pacific cod	Walleye pollock
1973	8.25		
1974	13.90		
1982	7.61		
1984		0.20	
1985	7.97	0.15	0.013
1986		0.40	0.020
1987		0.29	0.004
1988		0.44	0.007
1989		0.41	0.007

Table 3. **Log-likelihood and change in log-likelihood of major likelihood components between two population models of walleye pollock in the eastern Bering Sea.**

Likelihood component	Log-likelihood		
	No predator model	Predator model	Change
Fishery age composition (1964-1976)	–156.2	–155.3	+0.9
Fishery age composition (1977-1991)	–53.2	–52.7	+0.5
Survey biomass (age 3+)	7.4	5.7	–1.7
Survey age composition	–32.3	–33.3	–1.0
Age 1 number (survey index)	0.49	–0.45	–0.9
Total			–2.2

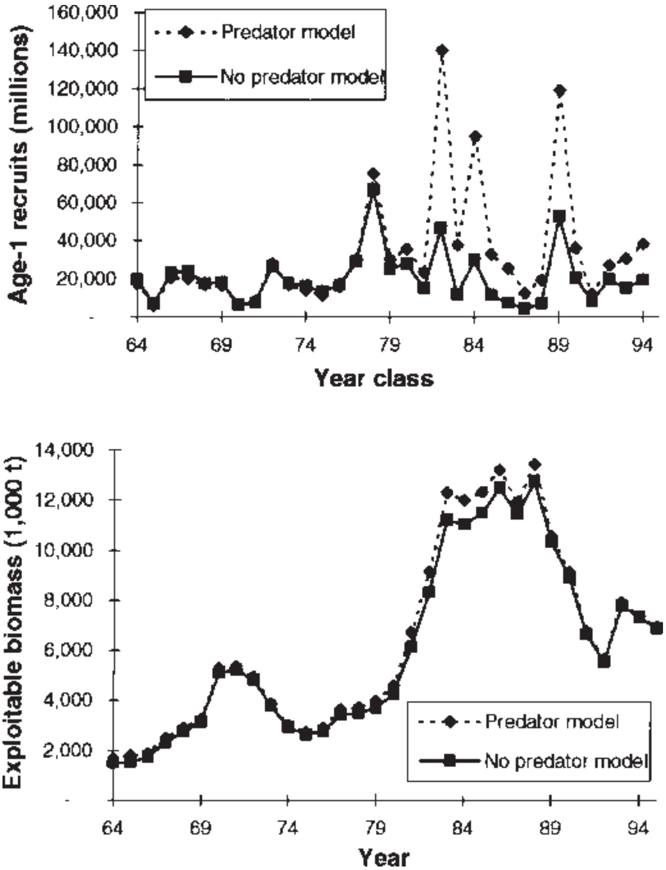


Figure 2. Estimates of number of age 1 (top) and exploitable biomass (bottom) of walleye pollock from the pollock population models with and without predators.

this age group increased substantially in 1983 when the large 1978 year class became cannibalistic at age 5, and the time trend in age-1 predation mortality mirrors the trend in cannibal biomass (Fig. 1). Predation mortality at age 2 was relatively low over the modeled period, ranging from 0.04 yr^{-1} to 0.09 yr^{-1} (Fig. 2). Mortality at this age was due to predation by cod and fur seals. Fur seals were the dominant source of mortality in the early part of the modeled period up to 1978 while cod was the dominant predator of age-2 pollock in the subsequent period .

The relationship between age-1 recruits and female spawning biomass for the standard model without predators shows a downward-bending Ricker curve ($r^2 = 0.08$) (Fig. 4) with declining recruits with increasing spawn-

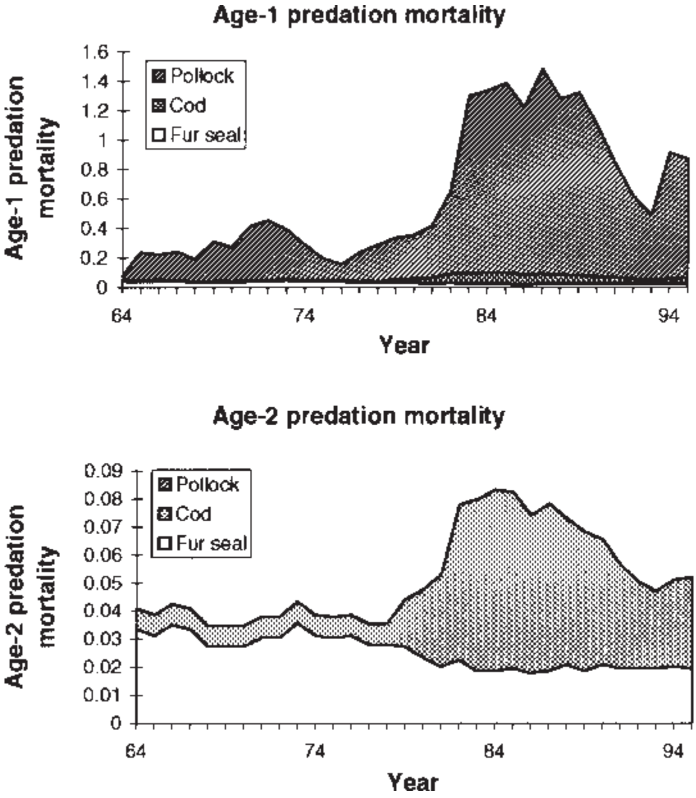


Figure 3. Predation mortality of age 1 (top) and age 2 (bottom) walleye pollock from 1963 to 1995 estimated by the eastern Bering Sea pollock population model with predation by adult pollock, Pacific cod, and northern fur seals.

ing biomass. However, the shape of the relationship when predators are added shows a more asymptotic relationship between spawners and recruits at age 1 ($r^2 = 0.12$).

Discussion

The model was able to fit predator data inputs without substantially degrading the fit to the fishery and survey data. The exploitable biomass of pollock in the predation model was very similar to that in the model without predators. This is different from some previous efforts by Honkalehto (1989) and Mito (1990), which incorporated cannibalism into cohort analysis models. In those efforts, estimated biomass was much higher and

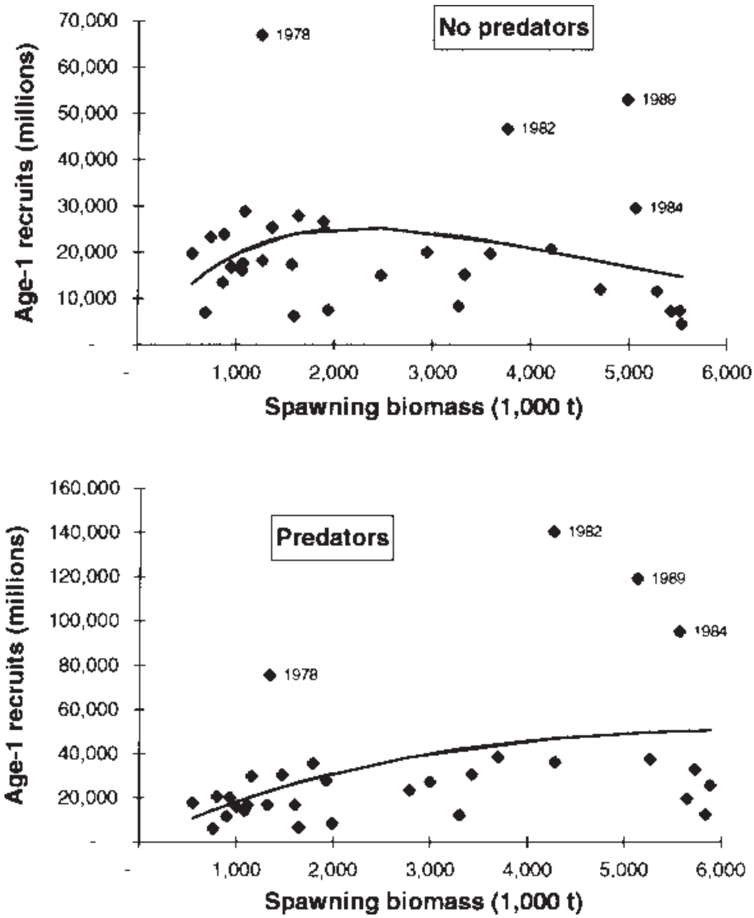


Figure 4. Estimates of age-1 recruitment of pollock recruits versus pollock spawning biomass from the pollock population models without (top) and with (bottom) predators and the fitted Ricker stock recruitment curves.

there was no attempt to fit model results to abundance estimates from surveys as was done in the present study. Also, Mito (1990) estimated very high amounts of cannibalism on age-2 pollock, a phenomenon which has not been observed in pollock stomach samples taken annually from the eastern Bering Sea shelf since 1985 (Livingston et al. 1993). Mito's report indicates differences in estimated size-at-age for young pollock relative to the sizes-at-age from U.S. data collections used in this modeling exercise, which possibly explain some of this difference. Honkalehto's model used fixed estimates of pollock consumption per unit cannibal over the whole modeled period using consumption rates derived from a study performed in the early 1970s which showed very high rates of cannibalism relative to rates presently observed (Livingston and Lang 1996). The combination of fixed, high consumption rates in that study likely produced the large biomass estimates, which are not supported by present-day survey estimates.

Estimates of total natural mortality of age-1 pollock obtained from this study are comparable to estimates from other studies. Livingston (1993) obtained estimates ranging from 1.1 to 1.9 in the mid-1980s using groundfish and marine mammal predation estimates. A multispecies virtual population analysis model of the eastern Bering Sea, which contains many more predators than considered in the present study, produced average rates of age-1 mortality in recent years of 1.85 (Livingston and Jurado-Molina, in press). Mito (1990) estimated age-1 predation mortality for the shelf pollock stock at around 1.0 yr^{-1} for the period 1970-1985. Total natural mortality at age 2 from the present study, which ranged from 0.34 to 0.38 yr^{-1}, was lower than the annual rate of 0.45 used in the standard population model and was much lower than estimates of Mito (1990), which ranged from 2.0 to over 4.0. Other groundfish, such as arrowtooth flounder which has been increasing in abundance in recent years, are predators of age-2 pollock and should possibly be incorporated into future population models of pollock with predation.

The asymptotic relationship between age-1 recruits and spawning biomass implies that mortality rates at age 0 are density-dependent. A similar conclusion was reached by Livingston (1989b, 1993). The change in the shape of the recruitment curve from an asymptotic relationship for age-1 recruits to a curve with an even steeper decline in number of age-3 recruits at high spawning stock sizes highlights the importance of cannibalism on age-1 fish in reducing even further the number that recruit to the fishery. Adding cannibalism to the model explained some of the variation in the spawner-recruit relationship. However, the largest outliers from the estimated stock-recruitment curve were years of above-average recruitment (1978, 1982, 1984, and 1989). Physical environmental data in Niebauer and Day (1989) showed that in 1978 and 1982 there was above-average first quarter sea surface temperatures and below-average first quarter northerly wind speeds. Possibly, these physical factors produced an environment during these years that enhanced normal survival rates of pollock larvae. Quinn and Niebauer (1995) have also linked variation in pollock recruitment with physical factors, with the strongest relationships observed between age-2 recruitment and air temperature and ice cover variables with a lag of 1, indicating mechanisms influencing recruitment beyond the larval stage . Wespestad et al. (submitted) presented an analysis suggesting that physical factors influencing transport of larvae into areas away from cannibalistic adults may be important. Echeverria (1995) hypothesized that sea–ice cover provides a mechanism to separate juvenile pollock from cannibalistic adults, thus influencing juvenile pollock survival rates between cold and warm years. Apparently, there are several mechanisms that may influence recruitment, each operating at different life history stages and possibly changing in importance depending on the climatic factors operating in a certain year.

Including predation mortality into a single-species stock synthesis model of walleye pollock appears to produce results that are consistent with existing data sources on population levels and fishing removals. This is a first step in improving our predictive capabilities in a multispecies system with a changing physical environment. It appears for pollock that the debate is not about which factor, physical environment or predation, determines the amount of recruitment into the fishery. Both factors play an important role, and defining the conditions under which each factor dominates should be the focus. Development of spatially explicit models that include both climate and predation may be needed to improve understanding of recruitment variability in walleye pollock.

References

AFSC. 1974. Fur seal investigations, 1973. NMFS, Alaska Fisheries Science Center, Seattle, WA 98115.

AFSC. 1975. Fur seal investigations, 1974. NMFS, Alaska Fisheries Science Center, Seattle, WA 98115.

Antonelis, G.A., A.E. York, B.W. Robson, R.G. Towell, and C.W. Fowler. 1996. Population assessment, Pribilof Islands, Alaska. In: E.H. Sinclair (ed.), Fur seal investigations 1994. NOAA Tech. Memo. NMFS-AFSC-69. 144 pp.

Brugge, W.J., and M.J. Holden. 1991. Multispecies management: A manager's point of view. ICES Mar. Sci. Symp. 193:353-358.

Echeverria, T.W. 1995. Sea-ice conditions and the distribution of walleye pollock *(Theragra chalcogramma)* on the Bering and Chukchi Sea shelf. In: R.J. Beamish (ed.), Climate change and northern fish populations. Can. Spec. Publ. Fish. Aquat. Sci. 121:131-136.

Holling, C.S. 1959. The components of predation as revealed by a small-mammal predation of the European pine sawfly. Can. Entomol. 91:293-320.

Honkalehto, T. 1989. A length-cohort analysis for walleye pollock based on empirical estimation of cannibalism and predation. In: Proceedings of the International Symposium on the Biology and Management of Walleye Pollock. University of Alaska Sea Grant, AK-SG-89-01, Fairbanks, pp. 651-666.

Kajimura, H., and C.W. Fowler. 1984. Apex predators in the walleye pollock ecosystem in the eastern Bering Sea and Aleutian Islands regions. In: D.H. Ito (ed.), Proceedings of the workshop on walleye pollock and its ecosystem in the eastern Bering Sea. NOAA Tech. Memo. NMFS F/NWC-62, pp. 193-234.

Kimura, D.K., and J.A. Lyons. 1991. Between-reader bias and variability in the age determination process. Fish. Bull. U.S. 89:53-60.

Lander, R.H., and H. Kajimura.1982. Status of northern fur seals. In: Mammals in the seas. FAO Fish. Ser. 5(IV):319-345.

Livingston, P.A. 1989a. Key fish species, northern fur seals, *Callorhinus ursinus*, and fisheries interactions involving walleye pollock, *Theragra chalcogramma*, in the eastern Bering Sea. J. Fish. Biol. 35(Suppl. A):179-186.

Livingston, P.A. 1989b. Interannual trends in walleye pollock, *Theragra chalcogramma*, cannibalism in the eastern Bering Sea. Proceedings of the International Symposium on the Biology and Management of Walleye Pollock, University of Alaska Sea Grant, AK-SG-89-1, Fairbanks, pp. 275-296.

Livingston, P.A. 1991. Executive summary. In: P.A. Livingston, (ed.), Groundfish food habits and predation on commercially important prey species in the eastern Bering Sea from 1984 to 1986. NOAA Tech. Memo. NMFS F/NWC-207.

Livingston, P.A. 1993. Importance of predation by groundfish, marine mammals, and birds on walleye pollock, *Theragra chalcogramma*, and Pacific herring, *Clupea pallasi*, in the eastern Bering Sea. Mar. Ecol. Progr. Ser. 102:205-215.

Livingston, P.A., and J. Jurado-Molina. In press. A multispecies virtual population analysis model of the eastern Bering Sea. ICES J. Mar. Sci.

Livingston, P.A., and G.M. Lang. 1996. Interdecadal comparisons of walleye pollock, *Theragra chalcogramma*, cannibalism in the eastern Bering Sea. NOAA Tech. Rep. NMFS 126:115-124.

Livingston, P.A., and R.D. Methot. 1996. Stock synthesis model of walleye pollock, *Theragra chalcogramma*, including predation (abstract). NOAA Tech. Rep. NMFS 126:217.

Livingston, P.A., A. Ward, G.M. Lang, and M-S. Yang. 1993. Groundfish food habits and predation on commercially important prey species in the eastern Bering Sea from 1987 to 1989. NOAA Tech. Memo. NMFS-AFSC-11. 207 pp.

McAlister, W.B., and M.A. Perez. 1976. Ecosystem dynamics of birds and marine mammals, Part I: Preliminary estimates of pinniped-finfish relationships in the Bering Sea. Final Rep. RU-77, Environmental Assessment of the Alaskan Continental Shelf. Alaska Fisheries Science Center, 7600 Sand Point Wy. NE, Seattle, WA 98115.

Methot, R.D. 1989. Synthetic estimates of historical abundance and mortality for northern anchovy. In: E. Edwards and B. Megrey (eds.), Mathematical analysis of fish stock dynamics: Reviews and current applications. Amer. Fish. Soc. Symp. 6:66-82.

Methot, R.D. 1990. Synthesis model: An adaptable framework for analysis of diverse stock assessment data. Int. North Pac. Fish. Comm. Bull. 50:259-277.

Mito, K-I. 1990. Estimation of consumption number by cannibalism of pollock in the eastern Bering Sea under assumption of multi-stocks. Int. North Pac. Fish. Comm. Bull. 50:179-206.

Niebauer, H.J., and R.H. Day. 1989. Causes of interannual variability in the sea ice cover of the eastern Bering Sea. GeoJournal 18.1:45-49.

Quinn II, T.J., and H.J. Niebauer. 1995. Relation of eastern Bering Sea walleye pollock *(Theragra chalcogramma)* recruitment to environmental and oceanographic variables. In: R.J. Beamish (ed.), Climate change and northern fish populations. Can. Spec. Publ. Fish. Aquat. Sci. 121:497-507.

Perez, M.A., and E.E. Mooney. 1986. Increased food and energy consumption of lactating northern fur seals, *Callorhinus ursinus*. Fish. Bull. U.S. 84:371-381.

Sinclair, E. 1988. Feeding habits of northern fur seals *(Callorhinus ursinus)* in the eastern Bering Sea. M.S. thesis, Oregon State University. 94 pp.

Sissenwine, M.P., and N. Daan. 1991. An overview of multispecies models relevant to management of living resources. ICES Mar. Sci. Symp. 193:6-11.

Thompson, G.G. 1996. Pacific cod. In: Stock assessment and fishery evaluation report for the groundfish resources of the Bering Sea/Aleutian Islands regions as projected for 1997. North Pacific Fishery Management Council, P.O. Box 103136, Anchorage, AK 99510.

Wespestad, V.G. 1996. Walleye pollock. In: Stock assessment and fishery evaluation report for the groundfish resources of the Bering Sea/Aleutian Islands regions as projected for 1997. North Pacific Fishery Management Council, P.O. Box 103136, Anchorage, AK 99510.

Wespestad, V.G., L.W. Fritz, W.J. Ingraham Jr., and B.A. Megrey. Submitted. On relationships between cannibalism, climate variability, physical transport and recruitment success of Bering Sea walleye pollock, *Theragra chalcogramma*. ICES J. Mar. Sci.

York, A.E. 1987. Northern fur seal, *Callorhinus ursinus*, eastern Pacific population (Pribilof Islands, Alaska, and San Miguel Island, California). In: J.P. Croxall and R.L. Gentry (eds.), Status, biology, and ecology of fur seals: Proceedings of an international symposium and workshop, Cambridge, England, 23-27 April 1984. NOAA Tech. Rep. NMFS 51:9-21.

Problems with an Environmental-Based Recruitment Index: Examples from a New Zealand Snapper Assessment *(Pagrus auratus)*

Mark N. Maunder
University of Washington, Fisheries Research Institute, Seattle, Washington

Abstract

This paper summarizes the different methods used to incorporate annual recruitment variation into the assessment of a New Zealand snapper stock and presents the implications of these methods. Recruitment in the Hauraki Gulf snapper stock has been related to water temperature and relationships with temperature are used to determine year-class strengths used in stock assessment models. Using a recruitment series derived from environmental factors has two major problems. (1) Environmental factors are used to predict recruitment from outside of the period that the relationship was derived and small changes in the relationship can influence the estimates of current biomass, its relative position to optimal biomass, and the age-structure. (2) Assumptions about the period of historic recruitment used to estimate future recruitment influences forward projections and yield calculations. The most defensible approach is to limit the time-frame of the model to recent history and assume that recruitment in the near future is similar to recent history. The problem with limiting the time-frame of the model is that the analysis becomes less model-dependent and more data-dependent, increasing the uncertainty and providing little management guidance. Catch-at-age data have been used as an alternative method to estimate recruitment and have also indicated other problems such as incomplete recruitment due to annual variation in growth rates.

Introduction

Snapper is the most researched inshore species in New Zealand due to its importance commercially, recreationally, and traditionally. Despite the large amount of research, a number of unresolved problems remain. As a result, a number of different methods have been used to assess this stock. One controversial set of data is a recruitment index calculated from a correlation between abundance of juveniles in prerecruit trawl surveys and temperature (Francis 1993). This index is used to represent year-class strengths and scale average recruitment to give annual absolute recruitment levels used in stock assessment models. Incorporation of this index into the assessment and modeling procedure has highlighted many inadequacies and these are discussed in this paper. First I describe the fishery and the methods used to assess its status; then I describe the temperature recruitment relationship and how it is used in the assessment models. Next I discuss three areas related to the use of the recruitment index: (1) the time frame of the stock assessment model and which years of the recruitment index are used, (2) the time frame of historic recruitment estimates used in predicting future recruitment and yield, and (3) the alternative use of catch-at-age data to determine year-class strength and what it tells us about the recruitment index.

Introduction to the Fishery

Snapper (*Pagrus auratus*) is the most valuable commercial inshore finfish in New Zealand. The total population of snapper surrounding New Zealand is divided into six management areas. Only two areas are of any "significant" commercial value, SNA1 on the east coast of the North Island (Fig. 1) and SNA8 on the west coast of the North Island. The total commercial catch of snapper in the 1995-96 fishing year (October 1, 1995 to September 31, 1996) was 6,937 t; 4,959 t in SNA1 and 1,558 t in SNA8 (Annala and Sullivan 1997). The fishery has been exploited since the mid 1800s and catches were greatest from the mid 1960s to the mid 1980s, averaging around 8,000 t. The quota management system was introduced in 1986 and the TACC (Total Allowable Commercial Catch) for SNA1 was set at 4,710 t for the 1986-87 fishing year. The reduced catch was implemented to allow for stock rebuilding. Decisions from the Quota Appeal Authority allowed the TACC to increase over time to 6,010 t in the 1991-92 fishing year. For the 1992-93 fishing year the TACC was reduced to 4,900 t and the TACC was again cut in the 1995-96 fishing year, but an appeal to the high court by commercial fishers reversed the decision.

SNA1 is the most popular recreational fishery in New Zealand due to its proximity to Auckland, the largest city in New Zealand, and the high abundance of snapper in this area. "The 1987 National Marine Recreational Survey showed that snapper was the most important finfish species sought by recreational fishers" (Annala 1995). Recreational catch in SNA1 was estimated to be 1,600 t in 1984-85, 2,794 t in 1993-94, and 2,052 t in

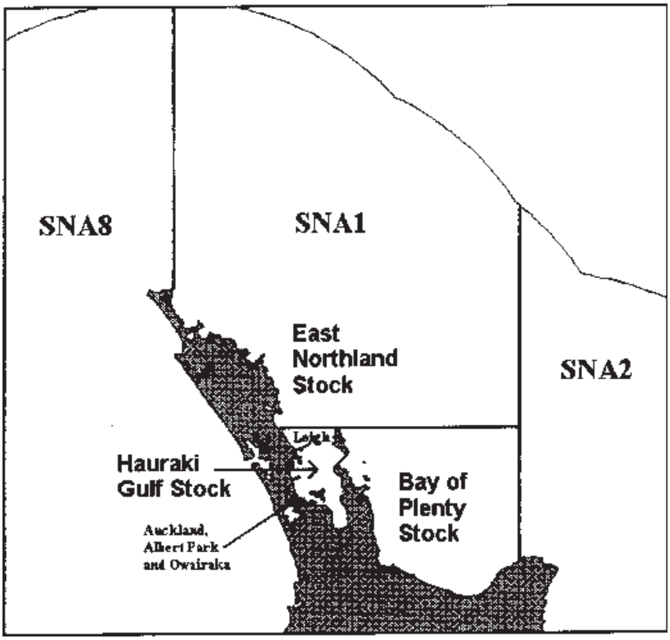


Figure 1. A map of the northern third of New Zealand's North Island showing the SNA1 snapper stock and its three substocks.

1995-96 (Annala and Sullivan 1997). There was an increase in the minimum legal size for recreational-caught fish from 25 cm to 27 cm in 1994 and a reduction in the bag limit from 15 to 9 in 1995. The commercial minimum legal size has remained at 25cm.

For assessment purposes, the SNA1 fishery has historically been broken up into three substocks based on biological considerations (Fig. 1); East Northland (EN), Hauraki Gulf (HG), and Bay of Plenty (BoP). The biological considerations included limited movement of individuals between populations, different size structures of the populations, and identifiable spawning areas (Annala and Sullivan 1997). Of these three substocks HG is the largest and the most heavily researched. The current assessments combine the HG and BoP substocks into one stock and assess EN as a separate stock. The combination of the HG and BoP substocks is based on movement between these two populations and the limited data for BoP. These assessments indicate that the HG-BoP stock is well below the level that will support maximum sustainable yield (B_{MSY}) but current removals are probably sustainable and the EN stock is at B_{MSY} but will decline under current removals (Annala 1995, Annala and Sullivan 1997).

The SNA1 fishery has a high political profile because of its value to both recreational and commercial sectors. In addition to conflicts between recreational and commercial fishers there are also conflicts with environmental groups, charter operators, and between different sectors within the commercial fishery (gear conflicts). This conflict causes political pressure directing a high proportion of government research into this fishery making snapper the most researched New Zealand fish species. The fishing industry (through the New Zealand Fishing Industry Board, NZFIB, which has recently been reorganized into the New Zealand Seafood Industry Council) also invests a significant amount of research into this fishery. Despite this research, there is still uncertainty in current assessments.

Assessments

There have been four recent model-based assessments of the SNA1 fishery. All four assessments use an age-structured model and two tagging biomass estimates (1984 and 1993). The first two models have been presented by government scientists and incorporate a recruitment index that depends on prerecruit trawl surveys and a correlation with temperature. The two models differ in the period over which they model the population. The third model was presented by industry scientists (NZFIB) and uses catch-at-age data to estimate recruitment rather than the recruitment index. The fourth, most recent model, combines the total catch history and catch-at-age models and is a collaboration of government and industry scientists. The models are as follows:

1. Total catch history model (Gilbert 1994). This method models the fishery from its virgin (unexploited) equilibrium state in 1850. The only parameter estimated is the average recruitment. Annual recruitment strength is determined by the recruitment index, and the model is fit to the two biomass estimates.

2. Short-term model (Gilbert and Sullivan 1994). This method models the fishery starting at a known (or assumed) biomass size and age-structure in 1985. The only parameter estimated is the average recruitment. Annual recruitment strength is determined by the recruitment index and the model is fit to the 1993 biomass estimate.

3. Short-term catch-at-age model (Maunder and Starr 1998). This model is similar to the short-term model (2), but differs by using catch-at-age data to estimate annual recruitment strength and a limited number of annual selectivity parameters for the youngest age-classes.

4. Long-term catch-at-age model (reported in Annala and Sullivan 1997). This model combines the catch-at-age model (3) with the total catch history model (1). In addition to average recruitment, annual recruitment residuals and selectivity parameters are estimated. The model

is fit to catch-at-age data, the recruitment index, and the two biomass estimates.

Temperature Recruitment Relationship

Recruitment to the Hauraki Gulf snapper stock is correlated with water temperature and relationships between abundance estimates of juveniles from prerecruit trawl surveys, water temperature, and air temperature have been generated. This relationship carries through to recruitment into the fishery at around age 4. Abundance estimates of 1+ snapper in Hauraki Gulf trawl surveys have a positive relationship with sea surface temperatures recorded at the Leigh Marine Laboratory (Francis 1993, Francis et al. 1997). A similar relationship has been made between the trawl surveys and air temperatures recorded at Albert Park and Owairaka (Gilbert 1994). Catch-at-age analysis shows that the strong and weak year classes seen in the prerecruit trawl surveys show up in the commercial catch data and can be seen for a number of years (Maunder and Starr 1998).

The temperature recruitment relationship is used to derive a recruitment index that is used in stock assessment models (Gilbert 1994, Gilbert and Sullivan 1994). The index is used as a relative measure of the strength of the year class recruiting to the fishery at age 4. The recruitment index is multiplied by an estimated average recruitment to give annual absolute recruitments. The recruitment index uses the observed data from the prerecruit trawl survey for years where the data is available (1983 to 1989 and 1991 to 1993). Missing years are estimated using the relationship between recruitment and sea surface temperature where sea surface temperatures are available (1967-1982, 1990, 1994-present), otherwise the relationship between recruitment and air temperature is used (1910-1966). Because no temperature data is available before 1910, recruitment for this period is assumed to be constant and equal to the average recruitment.

Time Frame of the Stock Assessment Model and Recruitment Index

The time frame of models used to assess the SNA1 stock have been influenced by scientists' faith in the temperature recruitment relationship used to generate the recruitment index. As the time frame of the estimation model increases, stronger assumptions about the applicability of the temperature recruitment relationship have to be made. An alternative to using the full recruitment index is to assume constant recruitment for the earliest years of the model time frame. Using constant average recruitment also includes strong assumptions in the possibility of regime shifts, global warming, and habitat degradation. Short-term models have been used to reduce the assumptions about historic levels of recruitment. The following section describes the benefits and inadequacies of the short- and long-term models.

Total Catch History Model

Total catch history models have been used to assess the SNA1 stock (Gilbert 1994, Gilbert et al. 1996). These models project the population forward in time from an unexploited population size. This requires the knowledge of the total catch history of the fishery. The SNA1 fishery is one of New Zealand's earliest and the current assessment uses a catch history reconstructed back to 1850 (Annala and Sullivan 1997). Catch from 1983 onward is based on area of capture while the catch from 1931 to 1982 is based on port of landing. Catch history before 1931 is unknown and has been assumed to follow a gradually increasing trend, totaling 150,000 t from 1850 to 1930 (Gilbert 1994). Fortunately, the total catch history model is insensitive to this historical catch (Gilbert 1994). This insensitivity is due to fish during this period dying of natural mortality before reaching the time when the model predictions are fit to data (Gilbert 1994) and the lack of a stock recruitment relationship in the model. The model is more sensitive to assumptions about unrecorded Japanese catch in the 1960s and 1970s (Annala and Sullivan 1997).

A total catch history model relies highly on the assumptions of the model rather than on the data used to fit the model. Generally the method assumes that the model structure and most parameters are known, and the parameter values are constant over the full history of the fishery. These assumptions make the assessment model dependent rather than data dependent. If parameter uncertainty and uncertainty in the model structure are ignored, then very precise estimates of the average recruitment (Fig. 2, top) and consequently maximum sustainable yield (MSY) and its associated biomass (B_{MSY}) are produced. The precision is a consequence of small changes in estimated average recruitment accumulating over a number of years transferring into large changes in current biomass. Confidence intervals calculated for these parameters give an unreliable estimate of uncertainty. In an attempt to describe uncertainty some researches present alternative models. A base case model or the average of all models is used to determine point estimates, and individual models with extreme results are used as bounds. The differences between alternative models can be in parameter values (i.e. rate of natural mortality) or structural differences (i.e. the functional form of the stock recruitment relationship). Presentation of alternative models is very informative if results are similar for all models. If results differ between models, the information is only useful if the models can be assigned relative probabilities.

A problem with using the total catch history model is that it requires a recruitment index over a time outside the range of the observed recruitments. In the SNA1 assessment this involves using recruitment relationships with both sea surface and air temperature. These relationships have to predict recruitment from temperatures outside the range of those used in forming the relationships. Recruitment has been observed to vary by a factor of 8.2 times in the trawl surveys and 27.6 times as predicted from

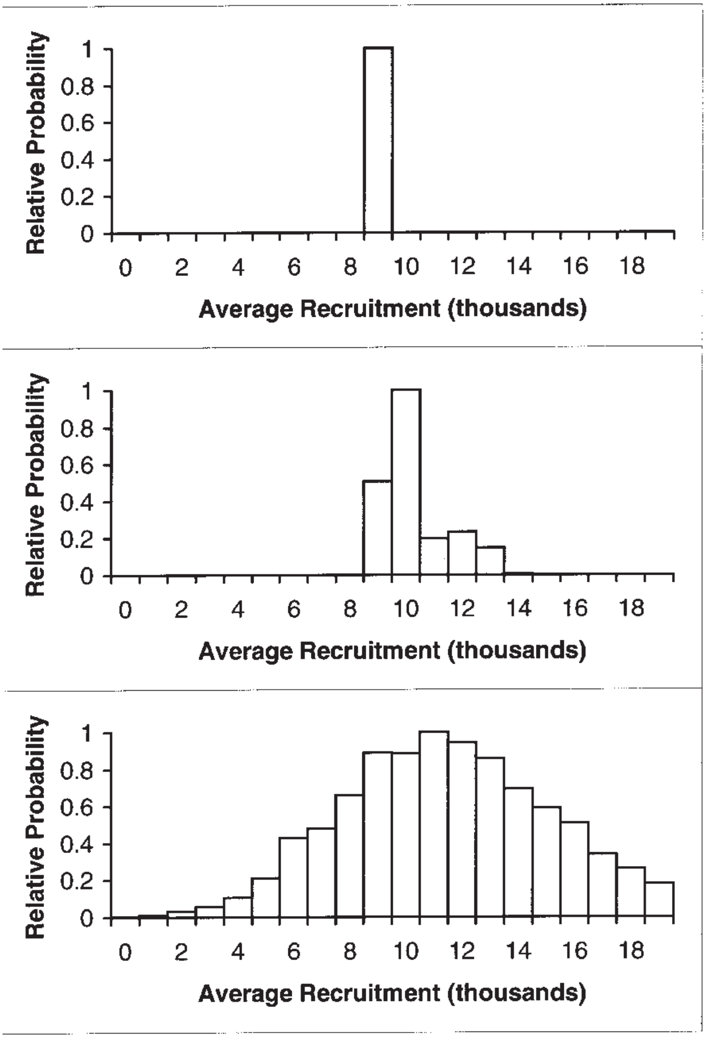


Figure 2. Posterior distributions of virgin recruitment (R_0) from a simple age-structured model using only two biomass estimates as data. Top is the posterior from a total catch history model with a uniform prior on R_0. Middle is the posterior from a total catch history model with a uniform prior on R_0 and a uniform prior on the stock-recruitment steepness (h, the proportion of the virgin recruitment that is realized at a spawning biomass level 20% of the virgin spawning biomass) from 0.5 to 1. Bottom is the posterior from a short-term model with a uniform prior on R_0 and a normal distributed prior for the starting biomass based on the biomass estimate. Note none of these models incorporate recruitment variability or a recruitment index.

the temperature recruitment relationship (Francis et al. 1995). The recruitment index calculated from air temperatures shows an increasing trend from 1910 to the present (Gilbert 1994) which may be caused by an urbanization effect around the temperature recording stations or global warming. This increase in the air temperature readings has an influence on the estimation of model parameters.

To avoid assumptions about the temperature recruitment relationship the recruitment was assumed to be constant before 1971 in the current assessment (Annala and Sullivan 1997). This is an alternative to reducing the time frame of the estimation model. The time frame of the recruitment index used is reduced by making the assumption that in an early part of the estimation models time frame, recruitment was constant at some average level. Unfortunately, this also involves strong assumptions about historic recruitment.

Density dependent effects are important in total catch history models because the population can go through large changes in population size. When the population is modeled from its virgin state, its size will vary considerably and the stock recruitment relationship will play a major role in determining MSY and in particular the biomass at MSY in relation to the virgin biomass (B_{MSY}/B_0). New Zealand fishery management requires that a stock should move to a biomass level that will support MSY, therefore B_{MSY}/B_0 is an important management parameter. The current assumption for the SNA1 stock is that recruitment is independent of stock size (Annala and Sullivan 1997). This is based on the evidence that environmental factors play a much larger role in determining year-class strength than stock size and is a consequence of discovering the correlation between temperature and recruitment. Other density-dependent factors (i.e. density-dependent growth and natural mortality) may also have a large influence on estimates of management parameters.

Bayesian analysis (Punt and Hilborn 1997) is a method that provides a framework for incorporating model or parameter uncertainty in the assessment. The dependence of total catch history models on model or parameter assumptions makes it important to use Bayesian or similar techniques to capture the uncertainty in estimation of management parameters. For example, uncertainty in the stock recruitment steepness (the proportion of the virgin recruitment that is realized at a spawning biomass level 20% of the virgin spawning biomass) increases the variance of the posterior distribution for the virgin recruitment parameter (compare Fig. 2 top and middle). This will allow decisions to be made that can incorporate the uncertainty. Bayesian analysis has been used in the assessments of the SNA1 fishery to give researches an idea of the uncertainty in the assessments, but has yet to be presented to management.

Short-Term Model

Due to the problems of the total catch history model, a more defensible approach may be to limit the time frame of the model to recent history ignoring historical trends in recruitment. The problem with limiting the time frame of the model is that the analysis becomes less model-dependent and more data-dependent, increasing the uncertainty and providing little management guidance. For the SNA1 stock a short-term model (starting in 1985) has been used to project the population between two biomass estimates. The estimates of average recruitment from this model are very uncertain with large confidence bounds (Fig. 2, bottom).

Assumptions about recruitment used in the estimation model are reduced for the short-term model because most estimates are taken from the trawl survey data, and values estimated from the temperature relationship are within the temperature range used to generate the relationship. Also in a short-term model there is very little change in stock size over the period modeled and density-dependent assumptions about recruitment and other process do not need to be made in the estimation model.

Yield Estimates with a Recruitment Index

Yield estimates for snapper in SNA1 are sensitive to the choice of years over which the assessment is made and to the choice of years used to represent average future recruitment. This sensitivity is due to recruitment of snapper being highly influenced by sea-surface temperatures in the first summer-fall of the birth year and the use of the temperature recruitment index in the models. Therefore, the yields will be a function of whether the chosen years are particularly warm or cold. For example the mean recruitment index between 1910 and 1930 is 0.40 (the average from 1931 to 1994 is one) and there has been a trend of increasing temperatures over time (Gilbert 1994).

The approach taken for SNA1 in previous assessments (Annala 1995) has been to use as long a period as possible, for both the estimation period and calculation of future average recruitment, to avoid introducing short-term biases into the yield estimates. However, there are insufficient prerecruit trawl survey data available to do more than a few years of high quality assessment for snapper. Therefore, a time series of temperature data have been used to standardize the period over which snapper recruitment is estimated. The period chosen was from 1931 to 1994 and all recruitments have been scaled so that the average of the recruitment index in this period is one. The approach assumes that this period represents the long term average of SNA1 snapper recruitment. Due to factors introduced in the preceding section, this may not be the best approach.

Two important management parameters are maximum sustainable yield (MSY) and current surplus production (CSP). These two parameters are at the extremes of management planning horizons; MSY is a long-term objective, while CSP is a short-term indicator. These two parameters have different requirements for the calculation of recruitment, MSY requires a long term average and CSP requires the best estimate of the recruiting

Table 1. Results using different periods to calculate average recruitment (R_0) from Gilbert (1994).

Period of average	R_0 (10^6)	B_{cur}/B_{MSY}	CSP (kt)	MSY (kt)
1910-1991	8.7	0.74	6.4	6.6
1931-1991	10.1	0.62	7.3	7.6
1967-1991	13.0	0.46	8.8	9.8

R_0 = average recruitment in numbers, B_{cur} = current biomass, B_{MSY} = biomass at maximum sustainable yield, CSP = current surplus production, MSY = maximum sustainable yield, and kt = kilotons.

cohort. Gilbert (1994) showed that using different periods to calculate average recruitment had a large influence on CSP, MSY and B_{cur}/B_{MSY} (Table 1). Gilbert (1994) calculated the average recruitment over the period 1931 to 1991 for his base case which implies that the initial virgin biomass, recruitment before 1910, projections and yield estimates are calculated from this average.

Maximum Sustainable Yield (MSY)

Calculations for MSY are dependent only upon the estimate of average recruitment when assuming no stock recruitment relationship. This will depend on what years are used to take as an average. It can be argued that the average should be taken over the longest possible period. It is also possible, due to a degraded or changing environment, that the carrying capacity, and therefore MSY, may have changed and that recent recruitment should be used for determining MSY. Also, current evidence for regime shifts (Francis and Hare 1994) means that there is likely to be different MSY values for different regimes, and MSY for the current regime can be determined more accurately from current recruitment.

MSY is dependent on assumptions about density dependence. This is true independent of the estimation model being short or long term. Recruitment to the population will be less when the population is at a size that supports MSY (B_{MSY}) compared to the unexploited population if there is a Beverton-Holt type stock recruitment relationship.

Current Surplus Production (CSP)

Current surplus production is the amount of the stock's production that can be removed allowing the population to remain at its current level. Therefore, CSP relies on an estimate of the recruiting cohort. Prediction of future recruitment is important when recruitment variation is large and the recruiting cohort represents a high proportion of the total available population or there is strong autocorrelation in recruitment. These recruitment estimates require information on both the average recruitment and the recruitment strength.

As with MSY, estimation of average recruitment for CSP depends on the time frame of the estimation model and the time frame used to determine average future recruitment. Density dependent processes will also influence CSP in the same manner as MSY.

Calculation of the recruiting cohort also requires an estimate of recruitment strength. The prerecruit trawl surveys in SNA1 allow prediction of recruitment to the fishery three years in advance and the temperature recruitment relationship four years in advance. The Southern Oscillation Index may also be used to predict high or low years classes before they are spawned giving an even longer lead time (Francis et al. 1995). However, when making these types of recruitment predictions precaution has to be taken because, as in snapper, variations in growth rate may mean that a recruiting cohort may not reach the legal size when expected, affecting yields available to the fishery (Maunder and Starr 1998). Also, the above estimates of recruitment strength are not without error. This error should be acknowledged by introducing a distribution for the recruitment strength allowing the estimation of uncertainty in CSP.

If there is no model independent estimate of recruitment strength for the recruiting cohort then a distribution of possible values should be used. If estimates of historic recruitment strength are available then the distribution of possible recruitment strengths can either be a parametric fit to these recruitment estimates or from nonparametric sampling of these estimates. If no historic estimates of recruitment strength are available then some assumptions about their distribution needs to be made. Recruitment strength is often assumed to be lognormally distributed with a variance taken from examining data from similar species.

Catch-at-Age Data or Environmental Index

An alternative to the recruitment index is estimation of annual recruitment from catch-at-age data. The relative abundance of an age-class in the catch gives information of cohort strength. Each year the age-class appears in the catch, more information is gained about its relative size. Therefore, catch-at-age analysis allows for the estimation of recruitment for cohorts that have been observed in the catch data only. Maunder and Starr (1998) showed that there is a high correlation between the recruitment index and recruitment strengths estimated from catch-at-age data (correlation coefficient = 0.88). Despite this high correlation there are still differences in recruitment strengths from the two methods.

There are advantages for using the recruitment index and also for using estimates from catch-at-age data. For most cohorts there is more than one catch-at-age data point but only one reliable trawl survey data point (age 1+). Estimation of recruitment strength for cohorts that have been seen in the catch-at-age data multiple times will have less bias due to random sampling errors. The catch-at-age data measures recruitment to the fishery at age 4, whereas the trawl surveys measure relative abundance of age 1+ juveniles. There are several factors that may change the relative abundance of a cohort before they recruit to the fishery, weakening the correlation between the prerecruit trawl surveys and recruitment to the fishery. In the SNA1 fishery the prerecruit trawl surveys are only available for the HG stock, but catch-at-age data are available for all stocks and indicate differences in the recruitment patterns between HG and EN (Annala and Sullivan 1997).

A disadvantage of using catch-at-age data is that estimates will be dependent on selectivity parameters. Maunder and Starr (1998) showed that there was incomplete recruitment to the fishery at age 4 and this varied annually. Estimates of recruitment for the last year of their model were the proportion of recruits vulnerable to the gear and not the size of the cohort. On the other hand, using a recruitment index alone may overestimate the numbers vulnerable to the gear if annual selectivity is not estimated. Catch-at-age data can only reliably estimate recruitment once the cohort has been in the fishery a number of years, but in contrast an environmental index can give future projections.

The observed annual variation in growth rates suggests it is important to estimate annual selectivity, particularly for the younger age classes. Catch-at-age data can be used to estimate selectivity. Recruitment and selectivity parameters are confounded and often, as is the case for the current SNA1 assessment, optimization routines fail. In these cases it is necessary to fix either recruitment or selectivity. If fixing recruitment gives similar results to fixing selectivity for the desired management parameters, scientists are usually happy. Otherwise, more analysis or data collection is needed. The current assessment of the SNA1 stock investigated fixing either recruitment (based on the recruitment index) or selectivities (based on tagging data) and a stepwise process that estimated both. Fortunately, the results from all models were very similar (Annala and Sullivan 1997).

Discussion

The evolution of the SNA1 stock assessment has developed around estimating historic recruitment. Incorporation of a recruitment index based on a correlation with temperature introduced concerns that a total catch history model may produce biased yield estimates due to decadal scale auto-correlated temperatures. The bias is due to assumptions regarding historic recruitment in the model-fitting process and the choice of years to use in the calculation of average recruitment when determining yield estimates. An alternative short term model avoided the problems of the total catch history model but produced imprecise estimates providing little management guidance.

Neither a short- or long-term model seems appropriate to model the SNA1 stock. It is possible that an estimation model with a time frame between the short- and long-term models will provide better management advice. The time frame should be chosen to avoid using the recruitment index based on air temperature (1910-1966) and the unknown quantity of Japanese longline catch in the 1960s and 1970s. This mid-term model will require at least one addition parameter to be estimated: the starting biomass. Both the recruitment index and catch-at-age data should be included in the estimation procedure, with appropriate error structure, to provide information on recruitment and selectivity. Methods should be investigated to enable simultaneous estimation of recruitment strengths and selectivity and the associated uncertainty.

Bayesian analysis provides a formal way to include uncertainty in the assumptions made about an assessment model and posterior distributions can be used to present information about uncertainty to decision makers. These methods can incorporate the uncertainty in all data, including the recruitment index, and include additional prior information from other sources. Bayesian methods may also be more appropriate because they can cope with confounding between recruitment strengths and selectivity parameters.

In New Zealand the management goal is to move the population to a size that will support MSY. In the context of decadal shifts in temperature and possible degraded environments, this seems to be an inadequate goal. In addition, the yield vs. biomass curve is flat with a high proportion of MSY available over a wide range of biomass levels. Therefore, management should focus on conserving the stock and providing a stable local economy rather than trying to squeeze the last fish out of the system. This ideology follows the trend of moving away from optimization toward non-optimal conservative goals. Current surplus production is one possible guide for total catch allocation. Estimation of the current surplus production needs the incorporation of uncertainty in average conditions and interannual variability. A prerecruit trawl survey or temperature correlation can help reduce uncertainty due to interannual variability, but many factors such as incomplete recruitment due to variation in growth need to be addressed in order to prevent incorrect decisions being economically costly. Due to recruitment variation, estimation of CSP will cause variation in annual catch allocation which may not be desirable. Recruitment in SNA1 is not a high proportion of the biomass and annual recruitment variation is not as much a concern as it might be for species like salmon.

Acknowledgments

I thank the SNA working group, particularly Nick Davies, Malcolm Francis, Dave Gilbert, Paul Starr, and Kevin Sullivan, for useful discussions and analysis on the SNA1 stock. I thank Greg Meikle for the map of the SNA1 stock. Financial support for this paper and research was provided by the New Zealand Fishing Industry Board.

References

Annala, J.H. 1995. Report from the Special Fishery Assessment Plenary, 29 May 1995: Stock assessments and yield estimates for snapper. 36 pp. (Unpublished report held in NIWA library, Wellington.)

Annala, J.H., and K.J. Sullivan. 1997. Report from the Fishery Assessment Plenary, May 1997: stock assessments and yield estimates. 381 pp. (Unpublished report held in NIWA library, Wellington.)

Francis, M.P. 1993. Does water temperature determine year class strength in New Zealand snapper (*Pagrus auratus,* Sparidae)? Fish. Oceanogr. 2:65-72.

Francis, M.P., A.D. Langley, and D.J. Gilbert. 1995. Snapper recruitment in the Hauraki Gulf. New Zealand Fisheries Assessment Research Document 95/17. 27 pp.

Francis, M.P., A.D. Langley, and D.J. Gilbert. 1997. Prediction of snapper *(Pagrus auratus)* recruitment from sea surface temperature. In: D.A. Hancock, D.C. Smith, A. Grant, and J.P. Beumer (eds.), Developing and sustaining world fisheries resources: The state of science and management. Second World Fisheries Congress, Brisbane 1996. CSIRO Publishing, Melbourne, pp. 67-71.

Francis, R.C., and S.R. Hare. 1994. Decadal-scale regime shifts in the large marine ecosystems of the North-east Pacific: A case for historical science. Fish. Oceanogr. 3(4):279-291.

Gilbert, D.J. 1994. A total catch history model for SNA 1. New Zealand Fisheries Research Document 94/24, 16 pp.

Gilbert, D.J., and K.J. Sullivan. 1994. Stock assessment of snapper for the 1992-93 fishing year. New Zealand Fisheries Research Document 94/3. 37 pp.

Gilbert, D.J., K.J Sullivan, N.M. Davies, J.R. McKenzie, M.P. Francis, and P.J. Starr. 1996. Population modelling of the SNA 1 stock for the 1995-96 fishing year. New Zealand Fisheries Research Document 96/15. 39 pp.

Maunder, M.N., and P.J. Starr. 1998. Validating the Hauraki Gulf snapper prerecruit trawl surveys and temperature recruitment relationship using catch at age analysis with auxiliary information. New Zealand Fisheries Research Document 98/15. 23 pp.

Punt, A.E., and R. Hilborn.1997. Fisheries stock assessment and decision analysis: The Bayesian approach. Rev. Fish Biol. Fish. 7:1-29.

Cross-Validation of Trophic Level Estimates from a Mass-Balance Model of Prince William Sound Using $^{15}N/^{14}N$ Data

Thomas C. Kline Jr.
Prince Willliam Sound Science Center, Cordova, Alaska

Daniel Pauly
University of British Columbia, Fisheries Centre, Vancouver, B.C., Canada

Abstract

Trophic mass-balance models of ecosystems constructed using the Ecopath approach and software include the diet composition of functional groups as model inputs, and trophic level estimates for these same groups as a model outputs. The well-documented 0.34% enrichment of $^{15}N/^{14}N$ that occurs at each feeding step in food webs can be used to determine trophic level as well. This contribution is the first to ever examine the relation between trophic levels estimated by these two independent methods. This was achieved by using a published Ecopath model of Prince William Sound (PWS) as reference, i.e., estimating $^{15}N/^{14}N$ ratios for each of the model's functional groups. Re-expression of theses ratios as absolute estimate of trophic levels (TL) was done following calibration using the herbivorous copepod *Neocalanus cristatus*, for which TL = 2. The correlation between both sets of TL values ($n = 7$) was extremely high ($r = 0.986$), with the points evenly distributed about the 1:1 line. Also, the magnitude of the standard errors of the TL estimates based on $^{15}N/^{14}N$ data was similar to those of the Ecopath estimates.

Applying $^{15}N/^{14}N$ data from PWS to an Ecopath model of the Alaska Gyre System resulted in a reduced correlation ($r = 0.755$, for $n = 16$), suggesting that TL estimates may be transferred between ecosystems, though at the cost of reduced precision. These encouraging results warrant further exploration.

Introduction

While the trophic level concept existed since the beginning of ecosystem research (Golley 1993), controversy has raged as to its operational validity. Particularly, some ecologists could not reconcile this concept, articulated in form of integers (primary producers = 1; first order consumers = 2, etc.), with the observation that many organisms derive their food from widely different parts of food webs (e.g., Rigler 1975). This problem was overcome by the introduction, through Odum and Heald (1975), of fractional trophic levels (*TL*s). These are computed as weighted means from disparate diet compositions, and their variance can be interpreted as an omnivory index (*OI*), in agreement with Pimm (1982), who defined omnivory as "feeding on more than one trophic level" (Pauly et al. 1993, Pauly and Christensen 1995). Until the late 1980s, however, estimation of trophic levels continued to be largely definitional for lower levels (see above), or based on crude, and often grossly erroneous guesses for higher levels. (See Pauly 1996 for the a discussion of such guesses by Ryther 1972 and other authors.)

In recent years, two methods have emerged that are capable of reliably estimating *TL* and related statistics: (1) the $^{15}N/^{14}N$ method (DeNiro and Epstein 1981, Fry 1988, Wada et al. 1991, and see below), and (2) the construction of mass-balance trophic models of ecosystems (Christensen and Pauly 1992, and see below). The former estimates are identified as TL_N, the latter as TL_E. In this study, we present a first comparison of results obtained by these two approaches, using a preliminary mass balance model of the Prince William Sound (PWS) ecosystem (Dalsgaard and Pauly 1997) as a starting point. Another model, describing the Alaska Gyre System (AGS; Pauly and Christensen 1996) is then used to test whether estimates of *TL* and *OI* may be transferred from one ecosystem to the other.

Materials and Methods

The trophic mass-balance model of PWS used here was constructed using the Ecopath approach of Polovina (1984) and Christensen and Pauly (1992); this is based on the system of linear equations

$$B_i \times (P/B)_i \times EE_i = Y_i \times \sum B_j \times (Q/B)_j \times DC_{ij} \tag{1}$$

where, for any conventional period without massive change of system structure:

B_i is the mean biomass of functional group *i* (e.g., a group of species with similar vital statistics, diet compositions and consumers); B_j is the mean biomass of the consumers of *i*; $(P/B)_i$ is the production/biomass ratio of *i* (equivalent to its instantaneous rate of mortality; Allen 1971); EE_i is the fraction of production $\{P_i = B_i \times (P/B)_i\}$ that is consumed within the system; Y_i is the catch of *i* by the fishery, if any; $(Q/B)_j$ the consumption per unit biomass of *j*; and DC_{ij} is the contribution of *i* to the diet of *j*.

The Ecopath software (Christensen and Pauly 1992) was used to solve this system of equations, after estimation of values of *B*, *P/B*, etc, from the literature on PWS and related systems. The assumptions made when estimating the inputs, and for their subsequent adjustment when establishing mass-balance, are documented in Dalsgaard and Pauly (1997). The other trophic mass-balance model used here, representing the Alaska Gyre system, was constructed in similar fashion and is documented in Pauly and Christensen (1996).

For both models, estimation of the *TL* values was performed by Ecopath, based on

$$TL_{Ei} = (1 + \text{mean trophic level of prey}) \qquad (2)$$

Omnivory indices (OI_{Ei}) were computed as the variance of the TL_{Ei} estimates (Christensen and Pauly 1992); the square root of the OI_{Ei} values was then treated as standard error of the TL_{Ei} estimates (S.E._{Ei}).

Further details on Ecopath, including its ability to account for uncertainty in input values through a Monte Carlo resampling scheme interpreted in a Bayesian context, are provided in Walters (1996) and other contributions in Pauly and Christensen (1996).

Stable isotope measurements are unique in that they trace assimilated material. Nitrogen stable isotope ratios provide excellent definition of relative trophic level (Fry 1988, Wada et al. 1991, Hobson and Welch 1992, Kiriluk et al. 1995). The heavy isotope of nitrogen, ^{15}N, is enriched by about 0.34 % (or 3.4 ‰ in conventional delta units) with each trophic level (DeNiro and Epstein 1981, Minagawa and Wada 1984) and has been shown to accurately indicate the "realized" trophic level of species within an ecosystem (Kling et al. 1992, Cabana and Rasmussen 1994). Thus, the set of TL_N values used for comparison was derived from samples collected in 1994-1995 in PWS (Kline 1997), and based upon the trophic bioconcentration of ^{15}N.

The conventional delta notation used to express a stable isotope ratio is reported relative to international standards (air for N) and defined by the following expression:

$$\delta^{15}N = (R_{sample} / R_{standard} - 1) \times 1000‰ \qquad (3)$$

where $R = {}^{15}N/{}^{14}N$. The isotope standard has a delta value of zero by definition, i.e., $\delta^{15}N = 0$ for atmospheric N_2. Naturally occurring $\delta^{15}N$ values observed in biota range from ~0 to ~ +20.

TL_N is estimated by relating observed $\delta^{15}N$ values to a reference value (TL_{ref}), and to the trophic enrichment factor, 3.4 (Minagawa and Wada 1984, Kline 1997), via:

$$TL_{Ni} = 1 + (\delta^{15}N_i - \delta^{15}N_{ref})/3.4 \tag{4}$$

Where TL_{Ni} is the trophic level of functional group *i*, $\delta^{15}N_i$ is the mean $\delta^{15}N$ of functional group *i*, and $\delta^{15}N_{ref}$ is the mean $\delta^{15}N$ value of the herbivorous copepod (with $TL = 2$, by definition) used as reference, here *Neocalanus cristatus* (Kline 1997). The standard deviations of the $\delta^{15}N_i$, $S.D._{Ni}$ were multiplied by the trophic enrichment factor, 3.4 (Minagawa and Wada 1984, Kline 1997), to estimate $\delta^{15}N$-based omnivory indices, OI_{Ni}. Given their relative magnitudes, we compared the OI_{Ni} estimates with the $S.E._{Ei}$, rather than with the OI_{Ei}.

Isotope sampling and data acquisition procedures were described elsewhere (Kline et al. 1993, Kline 1997). To account for the different definitions of the functional groups in the two ecosystem models (PWS and AGS), the available isotopic databases were arranged differently for each comparison. For example, in the PWS model, herring are considered separately (Table 1), while they are included in the small pelagics group of the AGS (Table 2).

Results and Discussion

Table 1 presents our results for PWS, i.e., the estimates of *TL* and omnivory index (*OI*) derived by the two methods under comparison. As might be seen, the fit between the estimates of TL_E and TL_N is very tight, even if one omits the reference data point, pertaining to *Neocalanus*, and for which $TL_E = TL_N = 2$ by definition (see arrow in Fig. 1A). Moreover, the data points are close to the 1:1 line, as they should if TL_E and TL_N measure the same underlying quantity. Figure 1B shows that the magnitude of the omnivory index estimates are similar for both methods compared here, except for the reference group (see arrow), which cannot, by definition, take a value of $S.E._E$ other that zero, while the estimate of OI_N can be quite large. This feature precludes correlation analysis.

The good match between the two types of *TL* estimates for functional groups in the PWS model is due, we believe, to the taxonomic correspondence between the Ecopath groups and the groups for which $\delta^{15}N$ were available. Another factor is scope of the sampling for stable isotope data, which involved numerous samples (Table 1) and which was very broad for all groups except birds. (Bird isotopic data were collected within a limited spatial and temporal range and included few of the common duck-like shorebirds; pers. comm., M.A. Bishop, U.S. Forest Service).

The Ecopath model of PWS was based almost exclusively on data collected prior to the 1989 *Exxon Valdez* oil spill, while the $\delta^{15}N$ data were collected after the spill. The good match between the two data sets implies that the basic structure of the food web has not been modified by the spill, at least as far as can be detected within the scope of this comparison. On the other hand, $\delta^{15}N$ data are not available for marine mammals

Table 1. Comparison of trophic level (*TL*) and omnivory index (S.E.$_E$ and OI_N) estimates for seven functional groups in the Prince William Sound ecosystem.

Ecopath group[a]	Corresponding taxa with $\delta^{15}N$ data	Data source[b]	N[c]	TL_E[d]	TL_N	S.E.$_E$[e]	OI_N
Mesozooplankton	*Neocalanus cristatus*[f]	1	938	2	2	0	0.54
Macrozooplankton	Euphausiids, amphipods, chaetognaths, etc.	1,2	329	2.8	2.81	0.43	0.49
Salmon fry	Young-of-the-year chum, sockeye and pink salmon	3	285	3.2	3.18	0.32	0.31
Herring	*Clupea pallasi* (Pacific herring)	1,3	385	3.3	3.33	0.37	0.21
Small pelagics	Smelts, juvenile gadids	3	273	3.3	3.08	0.34	0.32
Demersal fishes	Flatfish, rockfishes, greenlings, gadids, cottids	3	459	3.9	3.88	0.53	0.40
Birds	Seabirds	4	191	4.1	3.81	0.45	0.44

[a] Model is described in Dalsgaard and Pauly (1997), including details of species included in each functional group.

[b] Numbers in this column indicate sources of $\delta^{15}N$ data, viz. 1 = Kline (in press); 2 = Kline (unpublished data) ; 3 = Kline (1997); 4 = M.A. Bishop (U.S.F.S., Cordova, unpublished data).

[c] Number of $\delta^{15}N$ measurements.

[d] As given on Table 23 of Dalsgaard and Pauly (1997).

[e] From Ecopath file of PWS, available from second author.

[f] Reference group for $^{15}N/^{14}N$ as discussed in text.

and many bird species that experienced large mortalities, and this conclusion may thus change when the issue is revisited.

Table 2 and Fig. 1C show that using $\delta^{15}N$ data from one system (PWS) to estimate trophic levels in another, adjacent system (AGS), leads to predictions that are less precise than when inferences are drawn within the same system. Moreover, the omnivory index values become completely uncorrelated (Fig. 1D), partly because of overaggregation of functional prey groups, which leads to predators that appear to feed at only one trophic level.

A likely reason for the differences between TL_E in AGS and TL_N estimates from PWS is the shoe-horning of taxa with $\delta^{15}N$ data into the "boxes" of the AGS model. For example, the TL_N estimate for demersal fish in PWS is compared with a TL_E estimate for "large fish" in the AGS, though these consist of large pelagic fish (Pauly and Christensen 1996).

Table 2. **Comparison of estimates of TL_E and S.E.$_E$ for functional groups in the Alaska Gyre system with TL_N and OI_N estimates derived from Prince William Sound organisms.**

Ecopath group[a]	Corresponding taxa with δ^{15}N data	Data source[b]	N[c]	TL_E[d]	TL_N	S.E.$_E$	OI_N
Small herb. zoopl.	*Neocalanus cristatus*[e]	1	938	2	2	0	0.54
Microzooplankton	Bulk net samples	1	195	2	2.17	0	0.34
Carnivorous zoopl.	Amphipods, chaetognaths, decapod larvae, etc.	1,2	159	3	2.89	0	0.52
Jellies[f]	Gelatinous zooplankton	2	4	3	2.35	0	0.35
Krill	Euphausids	1	170	2.05	2.74	0.22	0.45
Squids	Squids	3	104	3.15	3.44	0.37	0.25
Small pelagics	Herring, smelts, juvenile gadids	1,3	545	3.16	3.24	0.36	0.27
Sockeye salmon	Sockeye salmon	4	118	3.91	2.88	0.45	0.13
Chum salmon	Chum salmon	2	1	4	3.05	0	
Pink salmon	Pink salmon	2	4	3.99	2.94	0.38	0.18
Steelhead	Dolly Varden[g]	2	6	4.12	3.46	0.20	0.31
Mesopelagics	Myctophids, smoothtongue, glass shrimp, snailfish	3	73	3.04	3.12	0.05	0.39
Large fish	Flatfishes, rockfishes, greenlings, gadids, etc.	2	459	4.04	3.88	0.30	0.40
Sharks	Dogfish, salmon shark	2	5	4.49	3.66	0.45	0.40
Pinnipeds	Copper R. Delta Harbor seals & Steller sea lions	5	22	4.33	4.92	0.36	0.08
Marine birds	Seabirds	6	191	4.08	3.81	0.28	0.44

[a] Model described in Pauly and Christensen (1996), including details of species included in each functional group.

[b] Numbers in this column indicate sources of δ^{15}N data, viz. 1 = Kline (in press); 2 = Kline (unpublished data) ; 3 = Kline (1997); 4 = Kline and Ewald (unpublished data); 5 = Hobson et al. (1997); 6 = M.A. Bishop (U.S.F.S., Cordova, unpublished data).

[c] Number of δ^{15}N measurements.

[d] From Ecopath file of Alaska gyre, available from the second author.

[e] Reference group for δ^{15}N.

[f] Diet composition, omitted in Pauly and Christensen (1996), consisted of 88% herbivorous zooplanton, and 12% microzooplanton.

[g] Presumed analogous to steelhead because both are iteroparous salmonids; see text.

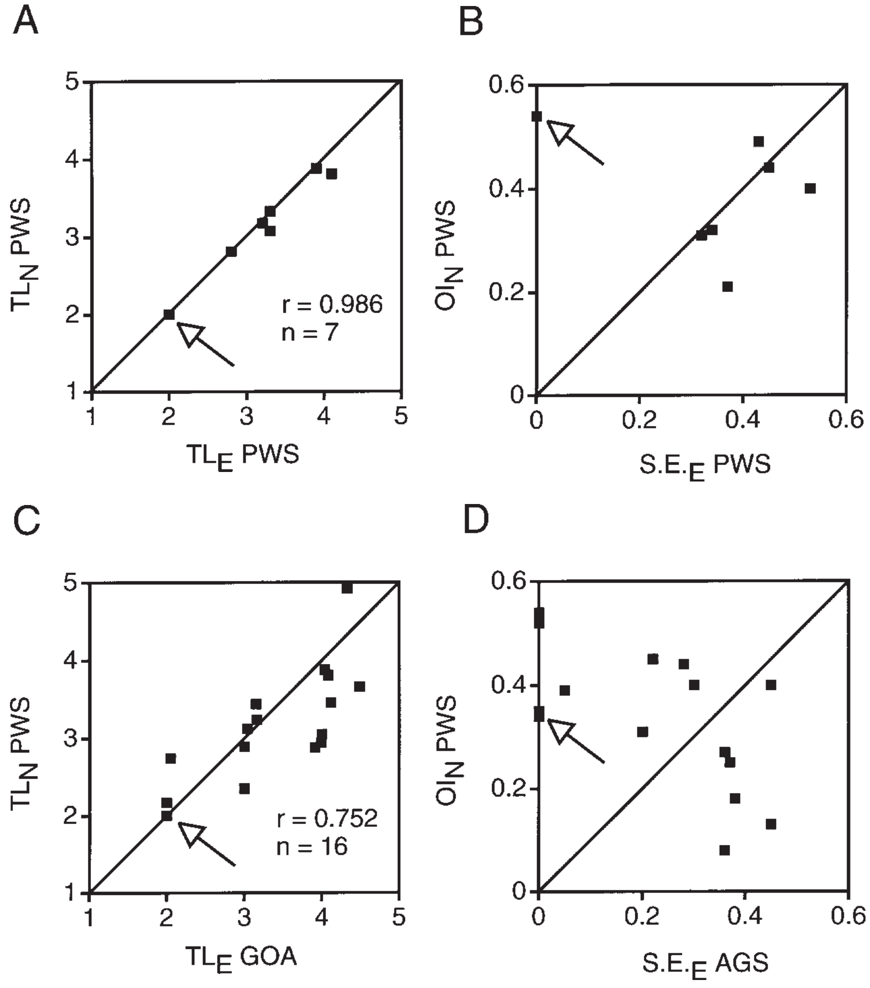


Figure 1. Relationships between estimates of trophic level derived from $\delta^{15}N$ data (TL_N) and Ecopath modeling (TL_E), and between the corresponding standard errors, for functional groups in Prince William Sound (PWS) and the Alaska Gyre System (AGS). (Arrows indicate reference group [Neocalanus cristatus]; see Tables 1 and 2 for details.) (A) Correlation between TL_N and TL_E estimates from PWS; (B) Relationship between OE_N and $S.E._E$ estimates from PWS; (C) Correlation between TL_N estimates from PWS and TL_E estimates from AGS; (D) Lack of relationship between OI_N estimates from PWS and $S.E._E$ estimates from AGS.

The $\delta^{15}N$ data on salmonids in Table 2 refer to adult specimens collected during their return migration through PWS and/or the Copper River, of which all would have done most of their feeding on the Gulf of Alaska shelf and/or in the AGS. However, having broadly similar life histories may not be sufficient for their *TL* values to be similar: the iteroparous salmonid, Dolly Varden (*Salvelinus malma*) had an estimated TL_N much lower than another iteroparous salmonid, steelhead (*Oncorhynchus mykiss*; Table 2), but close to the value of $TL_N = 3.54 \pm 0.21$ estimated for semelparous coho salmon (*O. kisutch*; $n = 12$). Thus, caution needs to be used when selecting "analog" species for cross-validation.

Still, we view the cross-validation exercise presented here as encouraging, in that the data we assembled led to coherent results, fully validating the independent, within-system estimates of *TL.* We also think that the lower correlation between the *TL* values from PWS and AGS were largely due to the lack of correspondence between the species included in the functional groups that were compared.

The next step is to refine our analyses, based on ecosystem models that are more detailed, thus requiring less "shoehorning," making better use of the available $\delta^{15}N$ data, and reducing the occurrence of misleadingly low estimates of $S.E._E$. Also, an important test will be to verify the high trophic level ($TL_E \sim 5$; Pauly and Christensen 1996) estimated for transient killer whales. Future work by both authors will be devoted to these and related issues, important in view of the recent demonstration (Pauly et al. 1998) that the trophic levels of global fisheries catches have been steadily declining in recent decades.

Acknowledgments

This project was funded by the *Exxon Valdez* Oil Spill Trustee Council through the Sound Ecosystem Assessment (SEA) and the Mass Balance Modeling (MBM) projects. However, the findings presented by the authors are their own and do not necessarily represent the Trustee Council's position. John Williams assisted with the translation of incompatible computer files between DP and TCK. Two anonymous reviewers provided useful improvements to our first draft.

References

Allen, K.R. 1971. Relation between production and biomass. J. Fish. Res. Board Can. 28:1573-1581.

Cabana, G., and J.B. Rasmussen. 1994. Modeling food chain structure and contaminant bioaccumulation using stable nitrogen isotopes. Nature 372:255-257.

Christensen, V., and D. Pauly. 1992. Ecopath II: A system for balancing steady-state ecosystem models and calculating network characteristics. Ecol. Model. 61: 169-185.

Dalsgaard, J., and D. Pauly. 1997. A tentative mass-balance trophic model of Prince William Sound, Alaska, for the period 1980-1989, prior to the *Exxon Valdez* Oil Spill. University of British Columbia Fisheries Centre Research Reports. 5(2). 33 pp.

DeNiro, M.J., and S. Epstein. 1981. Influence of diet on the distribution of nitrogen isotopes in animals. Geochim. Cosmochim. Acta. 45:341-353.

Fry, B. 1988. Food web structure on the Georges Bank from stable C, N, and S isotopic compositions. Limnol. Oceanogr. 33:1182-1190.

Golley, F.B. 1993. A history of the ecosystem concept in ecology: More than the sum of its part. Yale University Press, New Haven. 254 pp.

Hobson, K.A., and H.E. Welch. 1992. Determination of trophic relationships within a High Arctic marine food web using $\delta^{13}C$ and $\delta^{15}N$ analysis. Mar. Ecol. Progr. Ser. 84:9-18.

Hobson, K.A., J.L. Sease, R.L. Merrick, and J.F. Piatt. 1997. Investigating trophic relationships of pinnipeds in Alaska and Washington using stable isotope ratios of nitrogen and carbon. Mar. Mammal Sci. 13:114-13.

Kiriluk, R.M., M.R. Servos, D.M. Whittle, G. Cabana, and J.B. Rasmussen. 1995. Using ratios of stable nitrogen and carbon isotopes to characterize the biomagnification of DDE, mirex, and PCB in a Lake Ontario pelagic food web. Can. J. Fish. Aquat. Sci. 52:2660-2674.

Kline, T.C. 1997. Confirming forage fish food web dependencies in Prince William Sound using natural stable isotope tracers. In: Forage Fishes in Marine Ecosystems. Proceedings of the International Symposium on the Role of Forage Fishes in Marine Ecosystems. University of Alaska Sea Grant Report 97-01, Fairbanks, pp. 257-269.

Kline, T.C. In press. Scales and patterns of temporal and spatial variability of $^{13}C/^{12}C$ and $^{15}N/^{14}N$ in pelagic biota. In: S. Brandt and D. Mason (eds.), Space, time and scale: New perspectives in fish ecology and management. Can. J. Fish. Aquat. Sci.

Kline, T.C. Jr., J.J. Goering, O.A. Mathisen, P.H. Poe, P.L. Parker, and R.S. Scalan. 1993. Recycling of elements transported upstream by runs of Pacific salmon: II. $\delta^{15}N$ and $\delta^{13}C$ evidence in the Kvichak River watershed, southwestern Alaska. Can. J. Fish. Aquat. Sci. 50:2350-2365.

Kling, G.W., B. Fry, and W.J. O'Brien. 1992. Stable isotopes and plankton trophic structure in arctic lakes. Ecology 73:561-566.

Minagawa, M., and E. Wada. 1984. Stepwise enrichment of ^{15}N along food chains: Further evidence and the relation between $\delta^{15}N$ and animal age. Geochim. Cosmochim. Acta 48:1135-1140.

Odum, W.E., and E.J. Heald. 1975. The detritus-based food web of an estuarine mangrove community. In: L.E. Cronin (ed.), Estuarine research, Vol. 1. Academic Press, New York, pp. 265-286.

Pauly, D. 1996. One million tonnes of fish and fisheries research. Fish. Res. 25(1):25-38.

Pauly, D., and V. Christensen. 1995. Primary production required to sustain global fish catches. Nature 374:255-257.

Pauly, D. and V. Christensen (eds.). 1996. Mass-balance models of north-eastern Pacific ecosystems. University of British Columbia Fisheries Centre Reports 4(1). 131pp.

Pauly, D., M. Soriano-Bartz, and M.L. Palomares. 1993. Improved construction, parameterization and interpretation of steady-state ecosystem models. In: D. Pauly and V. Christensen (eds.), Trophic models of aquatic ecosystems. ICLARM Conf. Proc. 26:1-13.

Pauly, D., V. Christensen, J. Dalsgaard, R. Froese, and F.C. Torres Jr. 1998. Fishing down marine food webs. Science 279:860-863.

Pimm, S.L. 1982. Food webs. Chapman and Hall, London. 219 pp.

Polovina, J.J. 1984. Model of a coral reef system. I. The Ecopath model and its application to French Frigate Shoals. Coral Reefs 3:1-11.

Rigler, F.H. 1975. The concepts of energy flow and nutrient flow between trophic levels. In: W.H. Van Dobben and R.H. Lowe-McConnel (eds.), Unifying concepts in ecology. Dr. W. Junk Publishers, The Hague, pp. 15-26.

Ryther, J.H. 1972. Photosynthesis and fish production in the sea. Science 166:72-76.

Wada, E., H. Mizutani, and M. Minagawa. 1991. The use of stable isotopes for food web analysis. Crit. Rev. Food Sci. Nutr. 30:361-371.

Walters, C. 1996. Suggested improvements for the Ecopath model. In: D. Pauly and V. Christensen (eds.), Mass-balance models of north-eastern Pacific ecosystems. University of British Columbia Fisheries Centre Reports 4(1):82-87.

Calculating Capelin Consumption by Icelandic Cod Using a Spatially Disaggregated Simulation Model

Höskuldur Björnsson
Marine Research Institute, Reykjavík, Iceland

Abstract

BORMICON (BOReal MIgration and CONsumption model) is a simulation model including multispecies interaction and migrations. In this paper the model is used to investigate interactions between the Icelandic cod and capelin stocks. The main emphasis is on the spatial distribution of consumption of capelin by cod, cod growth, and the ability of the model to follow available data. The estimated effects of variations of capelin availability on the growth of cod are also investigated and compared to results obtained by other authors. Consumption according to model results is compared to consumption calculated directly from stomach content data using an evacuation rate model.

Introduction

Pálsson (1983) analyzed the feeding habits of cod, haddock, saithe, catfish, redfish, long rough dab, and plaice based on stomach samples taken in the period 1976 to 1981. His results showed the dominant role of the cod stock as a predator, especially with regard to commercially important prey species. The most important prey stock was capelin with northern shrimp *(Pandalus borealis)* second. Further studies on feeding of cod and on cod-capelin interactions (Magnusson and Pálsson 1989, 1991; Pálsson 1994; Marine Research Institute 1997a) have verified this conclusion. However, calculating the amount of capelin consumed by the cod stock directly from stomach samples, gives results that are in conflict with acoustic measurements of capelin if they are considered to be an absolute measure of stock size (Magnusson and Pálsson 1989).

Relating the growth of a predator to abundance of a prey species can help in estimating the amount of prey consumed by the predator. However, this only applies if the prey is a major part of the predator's diet. In the case of Icelandic cod, only capelin is important enough to impact growth significantly (Magnusson and Pálsson 1991, Steinarsson and Stefánsson 1996).

Since around 1980 the size of the cod stock has decreased with the lowest catchable biomass in 1992 to 1995 (Marine Research Institute 1997b). At the same time the stock size and commercial importance of many of its prey types, including capelin and northern shrimp, has increased (Marine Research Institute 1997b). This contrast in cod biomass facilitates analysis of feeding characteristics of cod through abundance indices of its prey types.

Models have been developed to investigate the effect of the cod stock on capelin and northern shrimp. Some of these models have been used in attempts to find optimum harvesting policies using economic criteria (Stefánsson et al. 1994).

It has been pointed out (ICES 1993) that spatial information must be incorporated to enable estimation of predation on certain prey types. In the case of Icelandic cod and capelin, the stocks do not occupy the same areas during large part of the year. When separate, the capelin is located farther north than the cod but the areas of distribution overlap near the edge of the continental shelf north of Iceland (Vilhjálmsson 1994). The same also applies to predation of northern shrimp, since part of the shrimp stock is distributed further north than the cod stock.

BORMICON (BOReal MIgration and CONsumption model) is a multispecies simulation model, developed at the Marine Research Institute in Reykjavík and described in Stefánsson and Pálsson (1997). As the name indicates, the model is designed to describe boreal ecosystems. It has been shown (ICES 1994) that boreal ecosystems have many features in common, among them variability in growth and long spawning migrations. Four features were considered essential in the model:

1. Spatial disaggregation and migrations.
2. Ability to calculate growth from feeding.
3. Ability to use data from many different sources for the estimation of parameters.
4. Maturation model.

The third requirement infers an ability to use predator growth, prey abundance indices, and stomach content data simultaneously for parameter estimation. A maturation model was considered essential due to differences in migrations of the mature and immature part of many stocks (Jónsson 1996, Vilhjálmsson 1994).

The operations implemented in the model are:

1. Migrations.
2. Consumption by predators (commercial fleets included).

3. Natural mortality.
4. Growth.
5. Maturation.
6. Spawning.
7. Recruitment to the stock.

The above list reflects the order in which the operations are executed at each time step. The model is in many ways similar to the Norwegian MULTSPEC model (Bogstad 1997).

Stefánsson and Pálsson (1997) tested a number of different alternative cod-capelin relationships, involving different migration patterns of mature cod and different treatments of stomach content data in the objective function. The conclusion was that the proportion of capelin in the diet of cod was on average 27%, varying from 26% to 30%.

Method

In this paper the cod-capelin relationship (Stefánsson and Pálsson 1997) is investigated with a special emphasis on the consumption of capelin by cod and on cod growth.

Description of Cod-Capelin Example

The simulation is based on the area division shown in Fig. 1. The areas are numbered from 1 to 16, but only 13 areas are used. Cod lives in areas 1 to 10, mature capelin in areas 1 to 3, 5 to 12, and 16, and immature capelin in areas 2, 3, 5, 6, 11, and 12.

The modeled age of cod is 3 to 11 years but 0 to 4 years for capelin. Recruitment to the cod stock is in March at age-3, using area distribution and length distributions from the groundfish survey (Pálsson et al. 1989). Recruitment of capelin is in August at age-0, using area distribution and length distributions from the 0-group survey.

Consumption is calculated by using the feeding level concept from Anderson and Ursin (1977) and type II feeding function (Magnússon and Pálsson 1991). The total consumption by a predator is the feeding level multiplied by maximum consumption which is both length and temperature dependent.

Stocks in the model are generally split into substocks. In this example cod and capelin were split into immature and mature parts. The entities stored for each substock are the number and mean-weight of each length and age group in each area. Calculation of growth is based on weight; growth in length is subsequently calculated from the growth in weight and the condition of the fish.

Growth of capelin was calculated by von Bertalanffy's equation (Beverton and Holt 1957), based on the current weight of the fish, and growth

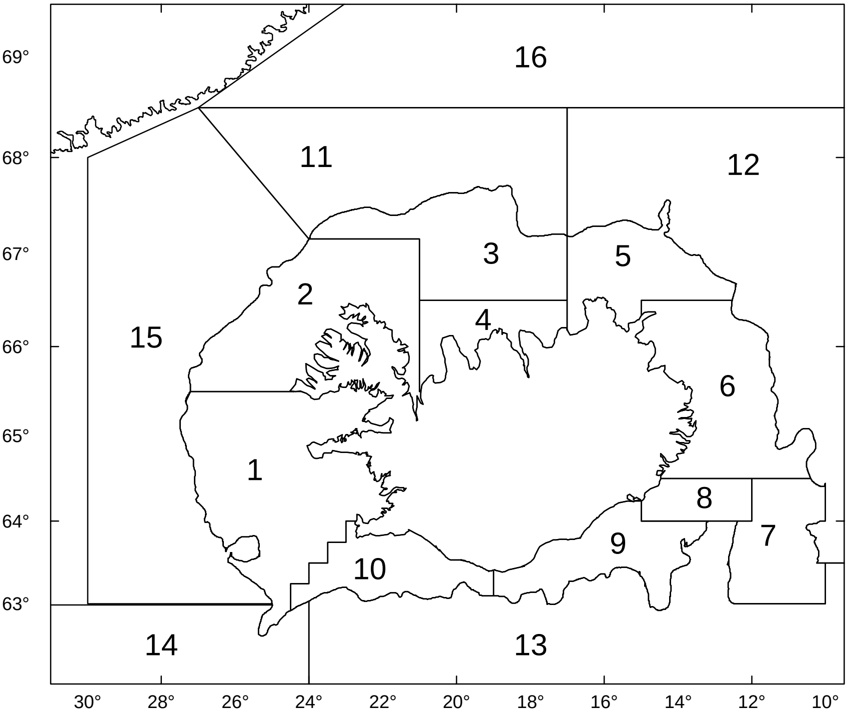


Figure 1. Area division used in cod-capelin simulation.

of cod was calculated by the equation of Jones and Hislop (1978), based on the amount of food consumed by the fish.

In the simulation capelin does not feed, while cod consumes two types of food, capelin and "other food." "Other food" is a prey which does not develop in the model but the amount in each area at each time step is predefined. Ideally, the amount of "other food" should be year-, month-, and area-dependent. This is of course not practical and in the simulation "other food" was divided into three spatio-temporal divisions.

1. North area summer.
2. North area winter.
3. South area summer and winter.

The amount of "other food" was abundant and its suitability low so there was no danger of depletion. Dividing other food in only three parts as done here may be too restrictive.

The data used in the objective function to be minimized were:

- Mean length at age of cod in survey and commercial catch.
- Age and area disaggregated survey indices of cod.
- Length and age distribution and age-length keys of cod in survey and commercial catch.
- Age and area disaggregated indices of capelin based on acoustic measurements.
- Stomach samples.
- Understocking of cod. (Not enough cod exists in an area to cover the commercial catch).
- Understocking of capelin.

The total objective function to be minimized is a weighted sum of the different components.

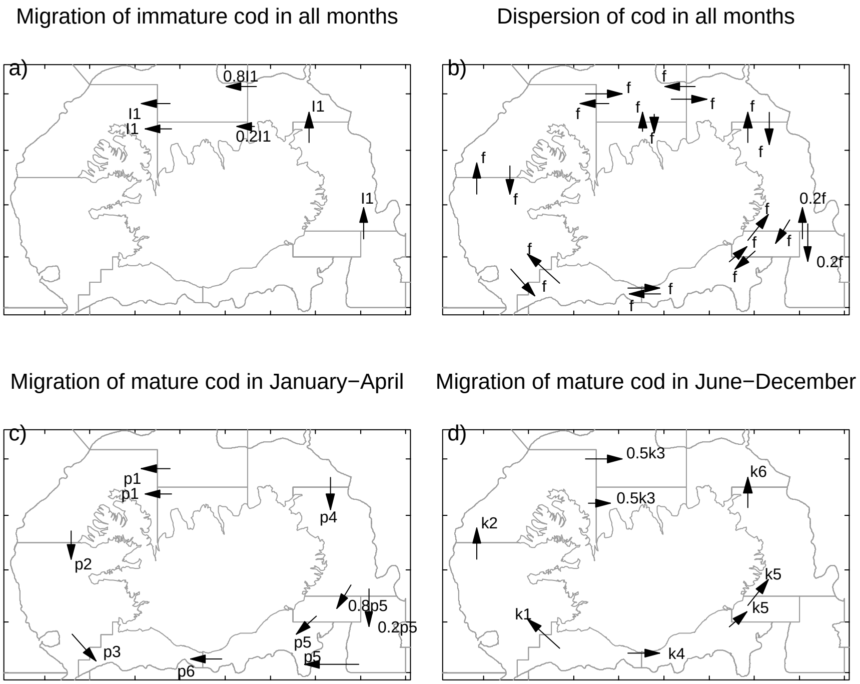
The estimated variables were:

- Recruitment of cod at age 3 and capelin at age 0.
- Initial number of cod and capelin.
- Parameters describing the suitability of capelin as food for cod.
- Suitability of "other food" as food for cod.
- Migration parameters for capelin.
- Selection pattern of fleets catching cod.
- Parameters describing maturation of cod.

In most of the alternatives of the cod-capelin example of Stefánsson and Pálsson (1997) abundance indices of age-3 capelin in autumn and age-4 in January were treated as absolute measures of stock size while indices of younger capelin were proportional.

Migration is calculated using migration matrices. Element (**i,j**) in a migration matrix describes the proportion of fishes in area **i** that will migrate to area **j** in next time step. Constructing the migration matrices directly quickly becomes difficult as the number of areas increases. In the cod-capelin example, cod lives in 10 areas (1-10) so the number of elements in each migration matrix is 100, reduced to 36 if migration is only allowed between adjacent areas.

To simplify the migrations and allow migrations to be estimated the migration matrices can be calculated from few migration parameters. The migration patterns used for cod in the cod-capelin example are shown in Fig. 2 and the migration matrix for mature cod in January-April in Table 1.


Migration of immature cod in all months

Dispersion of cod in all months

Migration of mature cod in January−April

Migration of mature cod in June−December

Figure 2. Migration pattern used for cod in the cod-capelin simulation.

Distribution of Capelin Consumption According to Stomach Data

Figure 3 shows the distribution of the amount of capelin in cod stomachs as a percentage of body weight. The figure is based on all available stomach content data at the Marine Research Institute (MRI), spanning the years 1979 to 1992. The sampling in March is much more extensive than in other months. As may be seen from the figure, capelin consumption occurs in all areas in March but is most extensive in area 1 and the southern part of area 2. The capelin consumed in these areas is about to spawn or has already spawned (Vilhjálmsson 1994). In July and November most of the capelin consumption occurs near the edge of the continental shelf north of Iceland. When considering the proportion of capelin in the diet of cod it must be borne in mind that some of the best fishing areas in summer and autumn are near the edge of the continental shelf northwest of Iceland, where capelin can be found nearby most of the year.

Table 1. Migration matrices for mature cod in January-April constructed from the migration parameters in Fig. 2b and 2c.

	From									
To	1	2	3	4	5	6	7	8	9	10
1	1-p3-2f	p2+f	0	0	0	0	0	0	0	f
2	f	1-p2-2f	p1+f	p1	0	0	0	0	0	0
3	0	f	1-p1-2f	f	f	0	0	0	0	0
4	0	0	f	1-p1-f	0	0	0	0	0	0
5	0	0	0	0	1-p4-2f	f	0	0	0	0
6	0	0	0	0	p4+f	1-p5-2.2f	f	f	0	0
7	0	0	0	0	0	0.2f+0.2p5	1-0.2f-p5	0	0	0
8	0	0	0	0	0	f+0.8p5	0	1-p5-2f	f	0
9	0	0	0	0	0		p5	p5+f	1-2f-p6	f
10	p3+f	0	0	0	0				f+p6	1-2f

Calculating the Capelin Consumption Directly from Stomach Content Data

Many authors have attempted to use stomach content data to calculate consumption of various prey types (e.g., Bogstad and Mehl 1992, Magnússon and Pálsson 1989). One of the main difficulties with this kind of work is to know the distribution of the predator. Results obtained from the BORMICON model provide information on the distribution of the cod stock which can be used in connection with stomach samples to calculate the predation of capelin and other preys.

The consumption will be calculated from stomach samples for March, July and November 1992. The annual consumption in 1992 is then the mean of the consumption in these three months multiplied by 12.

The stomach evacuation rate model used is that derived by Magnússon and Pálsson (1989)

$$C = 3.09 \times L^{1.15} 1.09^{T-6} \sqrt{Q} \tag{1}$$

Where Q is the total stomach content , C total consumption, L the length of the predator and T temperature.

Assuming the same digestion rate of all preys the consumption of each prey is calculated from

$$C_{prey} = C \frac{Q_{prey}}{Q} \tag{2}$$

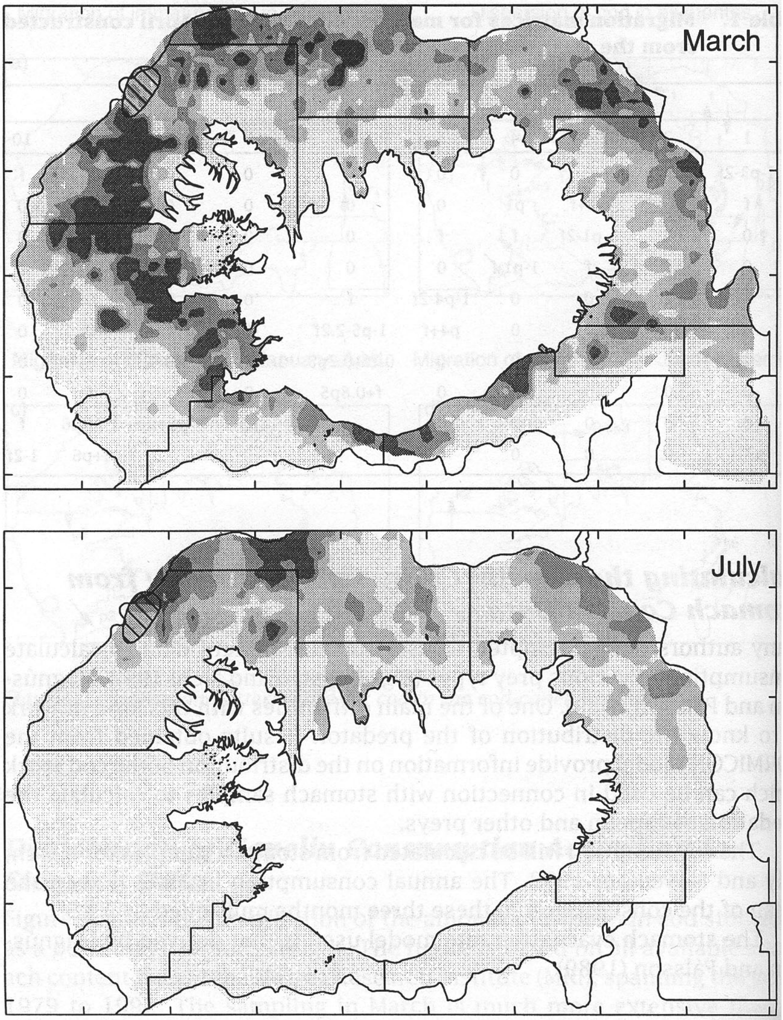


Figure 3. Amount of capelin as % body weight in cod stomachs according to stomach sample.

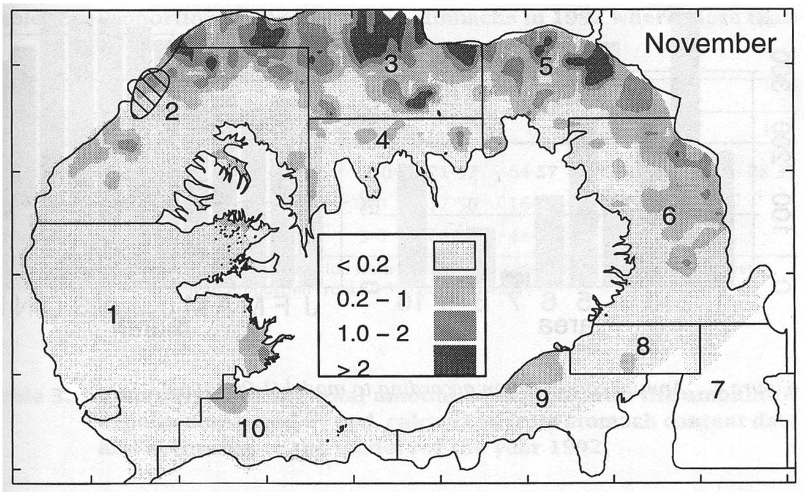


Figure 3. (Continued.)

Where C_{prey} is the consumption of prey and Q_{prey} the amount of prey in each stomach.

Results

In the following sections, results from the model are investigated. The emphasis is on spatial distribution of capelin consumption, cod growth, and its relationship to the size of the capelin stock. Results from the model will be compared to measurements which in this case are mean length at age and stomach content data.

Consumption of Capelin and Comparison to Stomach Content Data

Figures 4 and 5 show the total amount and amount of capelin consumed according to the model, disaggregated by area and month. Proportion of capelin in the diet of cod exceeds 50% in areas 2, 3, 5, and 6, highest during the winter or 60 to 80% but lowest in spring and early summer. In the southern area, capelin are only available during the spawning migration.

Table 2 compares the proportion of capelin in the diet of cod according to model results and stomach content data, showing an underestimate of capelin consumption in areas 1 and 2 in March by the model. Reason for the underestimate is that modeled migration of spawning capelin is

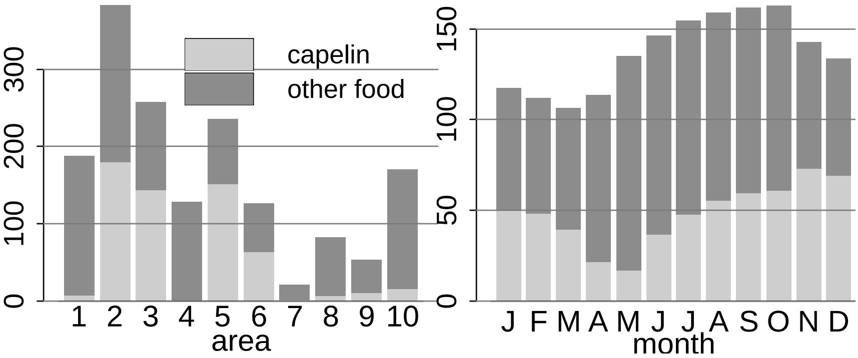


Figure 4. Annual consumption according to model (1,000 tons).

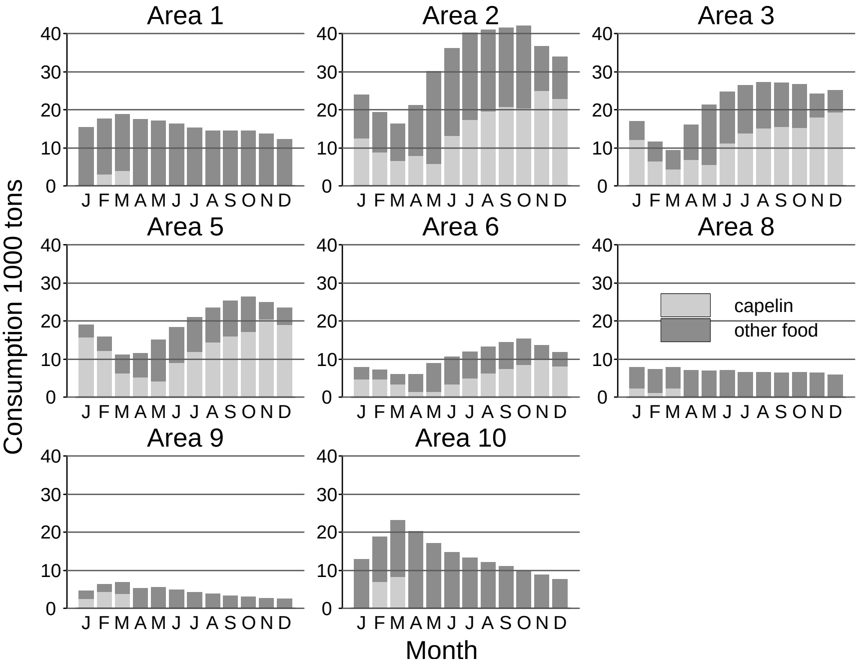


Figure 5. Annual consumption according to model disaggregated by area and month.

Table. 2 Proportion of capelin in cod stomachs in 1992 where more than 50 stomachs were sampled.

	Area								
	1	2	3	4	5	6	8	9	10
March	72-21	81-41	44-47	11-0	71-57	64-57	56-31	42-57	28-35
July	0-0	37-43	41-52	1-0	17-56	16-40	5-0	0-0	
November	6-0	39-68	54-74	3-0	68-82	48-72	1-0		

The first number in each cell is the proportion from cod stomach samples but the second number is derived from the model. Mean proportion from 30-90 cm cod used.

Table 3. Comparison of the total amount of capelin, and the amount of capelin consumed by cod, calculated from stomach content data and according to the model for the year 1992.

	Stomach content data			Model		
	Total	Capelin	% capelin	Total	Capelin	% capelin
March	142	84	59.4	106	39	36.8
July	120	24	20.5	155	48	31
November	123	42	35.1	143	73	51
Total	1,545	604	39.1	1,646	578	35.1

too “slow” so too few capelin make it to area 1 in March. In addition, migration of spawned capelin from area 1 to area 2 is not modeled. In March the highest capelin consumption per cod is in area 1 and the southern part of area 2 (Fig. 3).

Table 3 shows consumption calculated from stomach content data for March, July, and November 1992 as described in “Calculating the Capelin Consumption Directly from Stomach Content Data.” The results are compared to consumption obtained from the model. Stomach content data indicate a higher proportion of capelin in March but lower in July and November. As mentioned above most of the difference in March is due to problems in modeling the spawning migration of mature capelin.

The difference between stomach content data and model results in July and November 1992 is due both to problems with stomach samples and to compensate for the underestimate of capelin consumption in March caused by problems with the spawning migrations. In this period interaction between cod and capelin occurs in a narrow area near the edge of the continental shelf. Capelin in these areas are pelagic and stomach samples taken from cod caught with demersal trawl could lead to underestimation

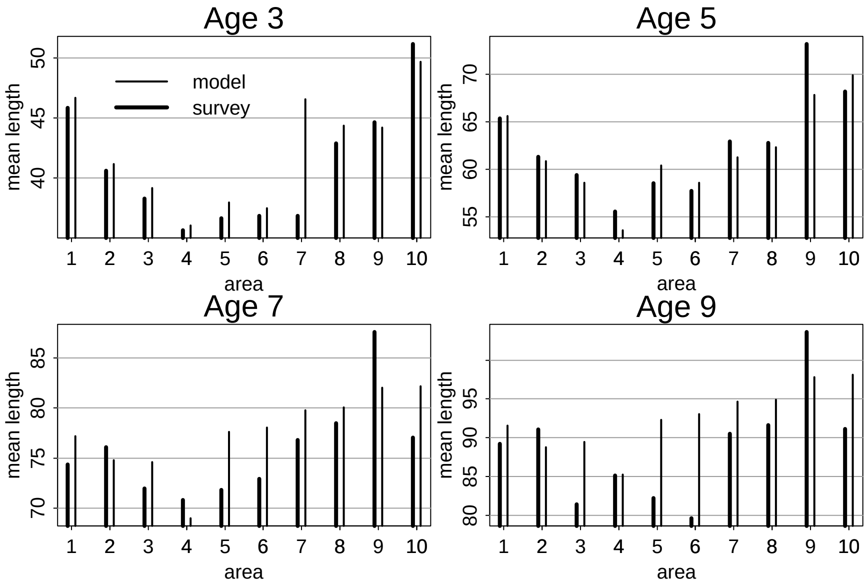


Figure 6. Mean length at age for some selected age groups in 1989 to 1996.

of capelin consumption in the area. Data from commercial fisheries indicate an abundance of cod in these areas but the distribution is patchy so obtaining a reliable estimate of the distribution is difficult.

Stefánsson (1998) examines a series of model results in which one component of the total objective function dominates in each estimation. The results appear reasonable except when emphasis is put on stomach content data, which is not surprising in view of the results obtained here where the model does not follow the stomach content data well.

Growth of Cod

Direct measurements of growth are difficult to obtain except from tagging data. Of the components used in the objective function, mean length at age is the best measure of growth. Relating growth to mean length at age is not difficult for the youngest fish but things get more complicated when maturation of a cohort begins and the mature individuals migrate from north to south.

In Fig. 6, mean length at age according to the model is compared to data from the groundfish survey. The model appears to predict the main features of the data. The most notable discrepancy between model and survey is age-7 in areas 3 to 6 where mean length according to the model is higher. Other cases with large discrepancy involve few individuals and are of little importance.

Table 4. Mean weight in March 1992 and March 1993 with capelin stock as in the model and no capelin stock from March 1992 to March 1993.

	No length dependent mortality of cod from April 1992 to March 1993					All mortality of cod included
Year class	Weight in March 92	Weight in March 93 no capelin	Weight in March 93 capelin	Growth with no capelin	Growth with capelin	Weight in March 1993 with capelin
1989	0.52	0.93	1.21	0.41	0.69	1.15
1988	1.25	1.97	2.43	0.73	1.18	2.23
1987	1.76	2.55	3.01	0.79	1.25	2.8
1986	3.21	4.2	4.78	1	1.58	4.61
1985	4.04	5.16	5.79	1.12	1.75	5.66
1984	5.69	7.12	7.83	1.43	2.14	7.79
1983	8.6	10.61	11.38	2.01	2.78	11.4

Relating cod growth to capelin stock size is of much interest because a large part of the commercial value of the capelin stock is as food for cod. Steinarson and Stefánsson (1996) derived a relationship between growth (in weight) of cod and biomass of capelin using mean weight at age from commercial catches and acoustic estimates of the capelin stock. According to their results cod growth in the absence of capelin is 75% of the growth when the capelin stock is large (at 1986 level).

Magnússon and Pálsson (1991) simulated the effect of capelin biomass on cod growth. The growth equation used is the same as used in BORMICON (Jones and Hislop 1978). According to their results the growth in weight in the absence of capelin is only 60% of the growth with the capelin stock at the 1986 level.

To investigate the effect of the capelin stock on cod growth three alternative simulations were run. The period was set from 1982 to 1993. All the simulations were identical to the cod-capelin example until in April 1992. From April 1992 to March 1993 the alternatives were as follows.

1. No capelin from April 1992 to March 1993. No cod caught in that period.
2. Normal development of the capelin stock from April 1992 to March 1993. No cod caught in that period.
3. Normal simulation from April 1992 to March 1993.

Table 4 shows the mean weight of cod in March 1996 and March 1997 according to the alternatives. As may be seen the capelin stock has the largest effect on the growth of younger cod. The mean annual growth of age 4 cod (1988 year class) with no capelin is 0.73 kg or 61% of the growth with normal development of the capelin stock. This agrees with the results from Magnússon and Pálsson (1991) which is expected as both studies use the same growth equation (Jones and Hislop 1978).

The data with highest weight in the objective function is cod data (mean length at age, length distribution, age length keys, etc.). This data relates to the feeding of cod through the growth model used. In the cod-capelin example, growth of cod is probably most important in estimating capelin consumption by the cod stock. The annual consumption of capelin is set to get the right contrast in cod growth between years. Underestimate of capelin consumption early in the year will therefore be followed by an overestimate in the latter part of the year as discussed in the section “Consumption of Capelin and Comparison to Stomach Content Data.”

Selection curves of a fleet (predator) in the model are length based, so a fleet can affect the mean length (and weight) at age. Comparison of columns 4 and 7 in Table 4 shows the effect of the commercial fleets on the mean weight at age of cod.

Using length distributions and fixed length-weight relationship as done here will probably underestimate the contrast in growth in weight between years with abundant and little capelin, as the condition of fish is better when more food is available. This problem will probably lead to underestimation of the importance of capelin in the diet of cod. The model has the potential to let the reference length-weight curve move with the feeding level (Stefánsson and Pálsson 1997), improving condition of the fish when the feeding level is high. Use of that possibility and inclusion of weighting of cod in the objective function could help in investigating this problem.

Conclusions

In this paper an attempt has been made to investigate some aspects of a multi-area simulation involving two species, cod and capelin. The items investigated are capelin consumption and cod growth, comparing model results to stomach content data, mean length at age from survey, and earlier work on the effect of the capelin stock on cod growth.

Results from the investigation indicate that the model results do not follow stomach content data well but the main features in the growth of cod are correctly predicted.

The extensive area disaggregation used has some major advantages especially with respect to the treatment of stomach samples. However, there are also many problems encountered. Implementing the migrations becomes more difficult and interannual variability increases.

Is the use of so much spatial disaggregation advantageous? The answer to this question is not obvious. Spatially disaggregated models are more data demanding and difficult to implement but can provide information about the ecosystem which more aggregated models cannot.

References

Anderson, K.P., and E. Ursin. 1997. A multispecies extension to the Beverton and Holt theory of fishing with accounts of phosphorus circulation and primary production. Meddelelser fra Danmarks Fiskeri og Havundersögelser 7:319-435.

Beverton, R.J.H., and S.J. Holt. 1957. On the dynamics of exploited fish populations. Chapman and Hall, London. Reprint edn. 1993.

Bogstad, B. 1997. Multispecies models for the Barents Sea. Ph.D. thesis, University of Bergen, Norway.

Bogstad, B., and S. Mehl. 1992. The north-east Arctic cod stock's consumption of various prey species. Proceedings of the Fifth IMR-PINRO symposium, Murmansk, 12-16 August 1991. Institute of Marine Research, Bergen, pp. 59-72.

ICES. 1993. Report of the Atlanto-Scandian herring and capelin working group. ICES Doc. C.M.1993/Assess 6. 177 pp.

ICES. 1994. Report of the multispecies assessment working group. ICES Doc. C.M. 1994/Assess 9. 177 pp.

Jones, R., and J.R.G. Hislop. 1978. Further observations on the relation between food intake and growth of gadoids in captivity. J. Cons. Cons. Int. Explor. Mer 38(2):244-251.

Jónsson, J. 1996. Tagging of cod *(Gadus morhua)* in Icelandic waters 1948-1986 and tagging of haddock *(Gadus aeglefinus)* in Icelandic waters 1953-1965. Rit Fiskideildar (Journal of the Marine Research Institute, Reykjavík). XIV(1). 108 pp.

Magnússon, K.G., and Ó.K. Pálsson. 1989. Trophic ecological relationships of Icelandic cod. Rapp. P.-V. Réun. Cons. Int. Explor. Mer 188:206-224.

Magnússon, K.G., and Ó.K. Pálsson. 1991. Predator-prey interactions of cod and capelin in Icelandic waters. ICES Mar. Sci. Symp. 193:153-170.

Marine Research Institute. 1997a. Fjölstofnarannsóknir 1992-1995 (Multispecies research 1992-1995). Marine Research Institute, Reykjavík.

Marine Research Institute. 1997b. Nytjastofnar sjávar og Aflahorfur fiskveidiárid 1997/1998 (State of marine stocks and prospects for the quota year 1997/ 1998). Unpublished, Marine Research Institute, Reykjavík.

Pálsson, Ó.K. 1983. The feeding habits of demersal fish species in Icelandic waters. Rit Fiskideildar (Journal of the Marine Research Institute, Reykjavík). VII(1), 60 pp.

Pálsson, Ó.K. 1994. A review of the trophic interaction of cod stocks in the North Atlantic. ICES Mar. Sci. Symp. 198:553-575.

Pálsson, Ó.K., E. Jónsson, S.A. Schopka, G. Stefánsson, and B.Æ. Steinarsson. 1989. Icelandic groundfish survey data used to improve precision in stock assessments. J. Northwest Atl. Fish. Sci. 9:53-72.

Stefánsson, G. 1998. Comparing different information sources. In: F. Funk, T.J. Quinn II, J. Heifetz, J.N. Ianelli, J.E Powers, J.F. Schweigert, P.J. Sullivan, and C.-I. Zhang (eds.), Fishery stock assessment models. University of Alaska Sea Grant, AK-SG-98-01, Fairbanks (this volume).

Stefánsson, G., and Ó.K. Pálsson. 1997. BORMICON. A boreal migration and consumption model. Unpublished, Marine Research Institute, Reykjavík.

Stefánsson, G., F.M. Baldursson, and K. Thórarinsson. 1994. Utilization of the Icelandic cod stock in a multispecies context. ICES C.M. 1994/T:43.

Steinarsson, B., and G. Stefánsson. 1996. Factors affecting cod growth in Icelandic waters and the resulting effect on potential yield of cod. ICES C.M.

Vilhjálmsson, H. 1994. The Icelandic capelin stock. Capelin, *Mallotus villosus* (Mueller), in the Iceland-Greenland-Jan-Mayen area. Rit Fiskideildar (Journal of the Marine Research Institute, Reykjavík). XIII(1). 281 pp.

Spatial Analysis of Fish Distribution and Abundance Patterns: A GIS Approach

Anthony J. Booth
Rhodes University, Grahamstown, South Africa

Abstract

A heterogeneous Geographical Information System (GIS) was developed to analyze fish distribution and abundance patterns on the Agulhas Bank, South Africa. The sparid fish, *Pterogymnus laniarius,* was chosen as a candidate species for analysis because of its commercial importance and high biomass. The results using Generalized Additive Modeling within the GIS revealed that there was a nursery area for immature fish on the central Agulhas Bank. After sexual maturation, approximately 40% of the biomass migrated eastward, colonizing large areas of the eastern Agulhas Bank. The use of GIS methodologies in integrated fishery management is discussed as it appears to be a developing trend in analyzing fish distribution and abundance data.

Introduction

Data collected from field surveys, be they biological, geological, or socioeconomic, have a spatial component. Depending on the observer and the purpose of the survey, this component is often ignored or given minor consideration. These data, abundance estimates in particular, often form the basis of many forms of fisheries assessments inter alia VPA (Pope and Shepherd 1985), age-structured production modeling (Booth and Punt 1998), and integrated analysis (Deriso et al. 1985).

Spatial analysis deals with data that is geo-referenced, in that it has a spatial component. As survey data is geo-referenced with respect to longitude and latitude it is an ideal candidate for spatial analysis. A suitable approach, which incorporates various qualitative and statistically robust methods to interpret this georeferenced and multivariate data, uses a Geographical Information System (GIS). A GIS is an information system-specifically designed to work with data referenced by spatial or geographic coordinates. In other words, a GIS is both a database system with specific capabilities for spatially referenced data, as well as a platform for analytical operations for working with the data (Star and Estes 1990). It can allow for the analysis of both qualitative and quantitative data types, identify associations between components, and therefore build a "living database" with exploratory data analysis, interpretative, and mapping capabilities.

The panga *Pterogymnus laniarius* (Cuvier 1830), a commercially important sparid fish species, was chosen to illustrate the development and application of a marine fish GIS. While aspects of its general biology are relatively well known (Booth and Buxton 1997), little is known about its distribution and abundance apart from the general descriptive work by Uozumi et al. (1981, 1985), Hatanaka et al. (1983), Badenhorst and Smale (1991) and Smale et al. (1993) There is, therefore, a clear need for a better understanding of the spatial dynamic of this species as it can highlight nursery areas, areas of high spawner biomass, and possible movements.

Materials and Methods

Biomass Surveys

Biomass survey data used in the analysis was collected on the South African Cape south coast between Cape Agulhas (34°50′S, 20°00′E) and Port Alfred (33°26′S, 26°54′E) during biannual biomass assessment surveys between 1988 and 1995 (Fig. 1). The sampling methodology used is described by Smale et al. (1993).

Fish Life History Stages

A life history stage classification proposed by Booth and Buxton (1997) was adopted to disaggregate the biomass of the stock and provide results on each stanza of its life history. Juvenile fish (<13 cm TL) were characterized by immature gonadal development and subadult fish (13-23 cm TL) were predominantly immature; few fish matured later in the life history stage and adult fish (>23 cm TL) were sexually mature.

Geographical Information System Design

A simple Windows NT (registered trademark of Microsoft Corporation) based Geographical Information System was developed using ArcView (registered trademark of ESRI Inc.). All statistical analysis were conducted using S-PLUS (registered trademark of Mathsoft). The statistical results in the form of summaries and coverages were exported back into ArcView for further manipulation and graphical presentation.

The starting location of each trawl was used to represent the spatial location, in decimal degrees, of the biological and physical data. The coastline and bathymetry was digitized from available nautical maps.

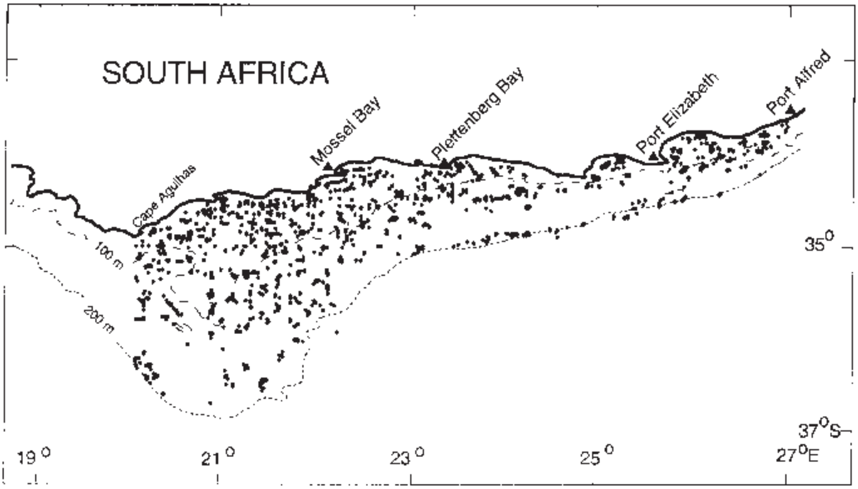


Figure 1. Map of Agulhas Bank, South Africa, with all stations trawled during 1988-1995.

Qualitative Analysis

Using overlaying procedures, characteristic of all GISs, relationships between covariates were investigated. Both point and polynomial interpolated physical data (to allow for better visual interpretation) were overlaid with fish density (fish per nm^2) in each life-history stage. Trends were identified using Boolean logic to select alternative combinations of values within the different coverages to identify zones of physical preference.

Quantitative Analysis

Generalized Additive Modeling (GAM) was used to model the spatial distribution and abundance of *P. laniarius* on the Agulhas Bank in response to various predictors (covariates) such as latitude, longitude, temperature, depth, and dissolved oxygen. From general observation it was noted that distribution of different life history stages were not homogenous over their distributional range (Fig. 2). It was also assumed that the density of life history stages was determined by an underlying Poisson process with the magnitude of the response variable (in this case fish per nm^2), being a function of at least two covariates—latitude and longitude. The use of GAM allows for changes in abundance to be related to spatial covariates without a restriction in the functional form of the relationship. This method allows for the incorporation of nonlinear (and possibly nonparametric) trends

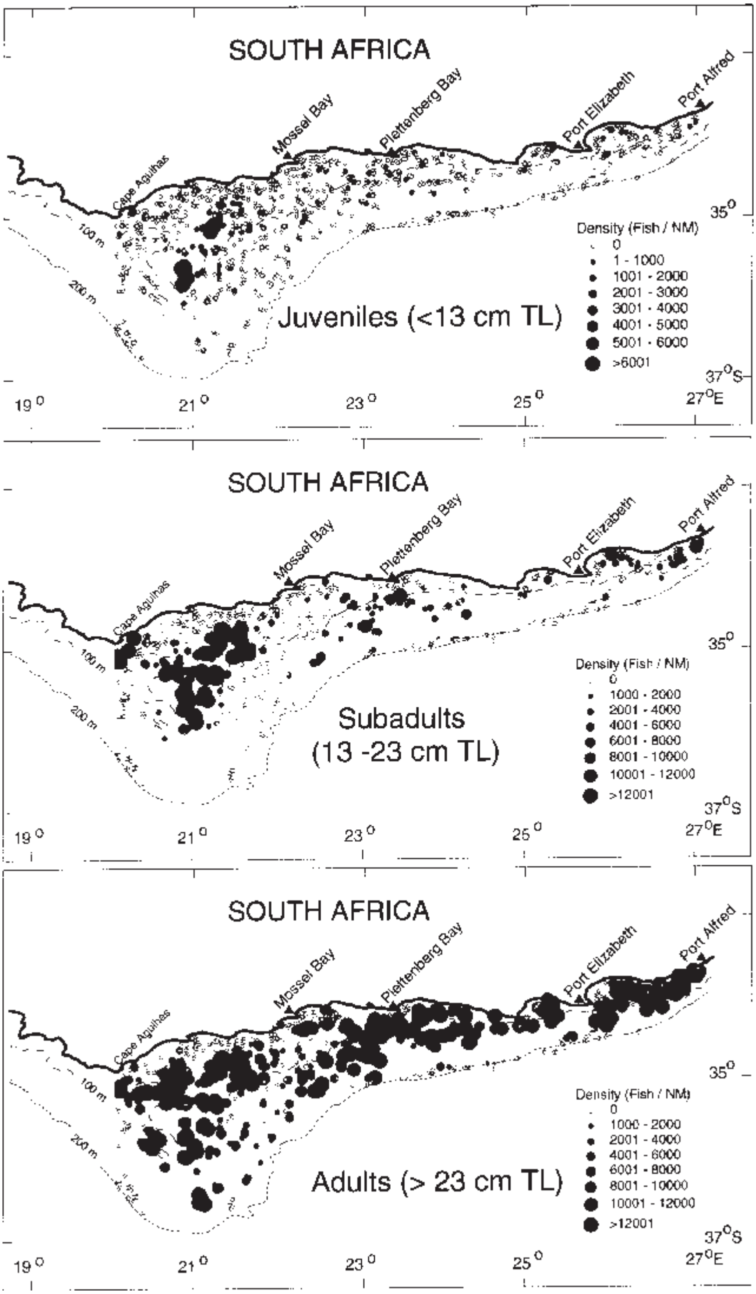


Figure 2. Distribution of three life history stages of Pterogymnus laniarius on the Agulhas Bank, South Africa. Data was collected between 1988 and 1995.

into the working model and to include covariates which could potentially determine spatial distribution. Details regarding the use of GAMs are described by Hastie and Tibshirani (1986, 1990) and Swartzman et al. (1992).

Optimum Fishing Areas

Once spatial trends were estimated, these were used to estimate biomass. That area of *P. laniarius* distribution was identified by isolating the area where fish have been sampled on the research surveys over the past decade. A polygon was drawn around the extreme points of the known panga distribution, forming a polygon hull. A polynomial extrapolation-interpolation algorithm (Press et al. 1997) was used on the predicted GAM point estimates to predict fish density at points not sampled, but occurred within the polygon hull on a $5 \times 5'$ square grid. The point density estimates were then multiplied by the average mass per fish in each life history stage, multiplied by 25 (the area in each interpolated block) and all life history stage biomasses summed.

Areas with consistently high adult and low immature biomass were determined by using Boolean logic within ArcView using the averaged GAM predicted biomass estimates. Those areas with at least 80% or 85% spawner biomass were selected. Total *P. laniarius* biomass was estimated for the central (20°-23°E) and eastern (23°-27°E) portions of the Agulhas Bank. These areas were chosen as they represent important fishing grounds used by two distinct trawl fisheries operating from Mossel Bay and Port Elizabeth.

Results

A summary of all data collected and utilized in the analyses is summarized in Table 1 and the trawl positions illustrated in Fig. 1. The importance of panga is highlighted on the Agulhas Bank as this species makes up about 7.5% of the demersal biomass and is the third most abundant demersal species by mass (excluding the pelagic species exhibiting diel migrations) surveyed since 1988 (Japp et al. 1994).

The polynomial interpolated point estimates of bottom temperature (Fig. 3) and dissolved oxygen levels (Fig. 4) provide an adequate description of the physical conditions on the Agulhas Bank on a seasonal basis. While only two surveys during spring and autumn 1993 have been illustrated in this study, a description of all the surveys follows. During autumn, the eastern Agulhas Bank was the warmest, with a small coastal intrusion of warmer water on the central Agulhas Bank. This situation was similar to the dissolved oxygen distribution with higher levels of dissolved oxygen, also being distributed on the eastern Agulhas Bank. During winter, the intrusion on the central Agulhas Bank from the coast to deeper areas became evident, and by spring the warmest water with the highest dissolved oxygen was widely distributed over this area. Generally, eastern Agulhas

Table 1. A summary of south coast biomass cruises within South African waters between 20° and 27°E.

Cruise number	Date	Total trawls	Trawls with *P. laniarius*	Trawls for analysis	Depth (m)	Temperature (°C)	Dissolved oxygen (ml/L)
SC 063	May/Jun 1988	93	67	84	117.36±79.61 (30-450)	11.13±1.48 (9.28-18.47)	3.83±0.92 (1.97-5.53)
SC 072	May 1989	62	49	55	100.59±39.02 (32-185)	10.59±1.59 (8.8-16.05)	3.70±0.89 (1.85-5.19)
SC 082	May/Jun 1990	58	54	57	100.88±58.91 (30-480)	11.85±2.14 (9.54-17.17)	3.64±1.13 (1.64-5.42)
SC 086	Sep 1990	91	47	72	79.56±43.22 (18-224)	12.44±2.23 (8.11-16.21)	4.47±0.67 (3.25-5.91)
SC 093	Jun/Jul 1991	91	70	82	109.69±66.41 (33-362)	11.54±2.20 (8.13-16.25)	4.07±0.68 (2.47-5.56)
SC 095	Sep/Oct 1991	75	60	68	81.64±24.87 (31-144)	13.78±1.69 (9.97-17.44)	4.85±0.70 (3.76-5.99)
SC 102	Apr 1992	82	60	44	110.17±60.71 (30-400)	11.04±0.78 (9.29-13.16)	3.75±0.78 (2.02-4.90)
SC106	Sep 1992	87	67	69	80.32±26.64 (25-124)	13.63±2.14 (9.33-16.42)	4.68±0.68 (3.36-5.87)
SC 111	Apr/May 1993	104	89	101	109.89±43.52 (29-237)	10.53±1.68 (7.70-18.16)	4.05±0.45 (2.72-4.79)
SC 116	Sep 1993	105	85	67	90.30±30.73 (29-186)	12.58±2.41 (8.40-16.30)	4.53±0.60 (3.55-5.98)
SC 122	Jun/Jul 1994	88	65	78	123.74±87.02 (35-500)	11.75±2.88 (6.59-2.88)	4.28±0.83 (1.36-5.65)
SC 125	Sep/Oct 1994	92	66	66	82.93±31.31 (30-200)	12.04±1.57 (9.50-15.13)	4.39±0.83 (3.12-9.2)
SC 129	Apr/May 1995	95	64	75	138.18±98.09 (29-483)	10.25±1.83 (8.30-16.20)	3.86±0.83 (1.80-6.02)
SC 131	Sep/Oct 1995	96	73	81	90.51±37.75 (28-193)	12.06±2.24 (9.20-16.62)	4.26±0.65 (3.03-6.19)

The mean (± standard deviation) and range (in parentheses) of the physical data were sampled using a CTD rosette sampler within 10 m of the substratum. Spring cruises occurred between September and October while autumn/winter cruises occurred between April and July.

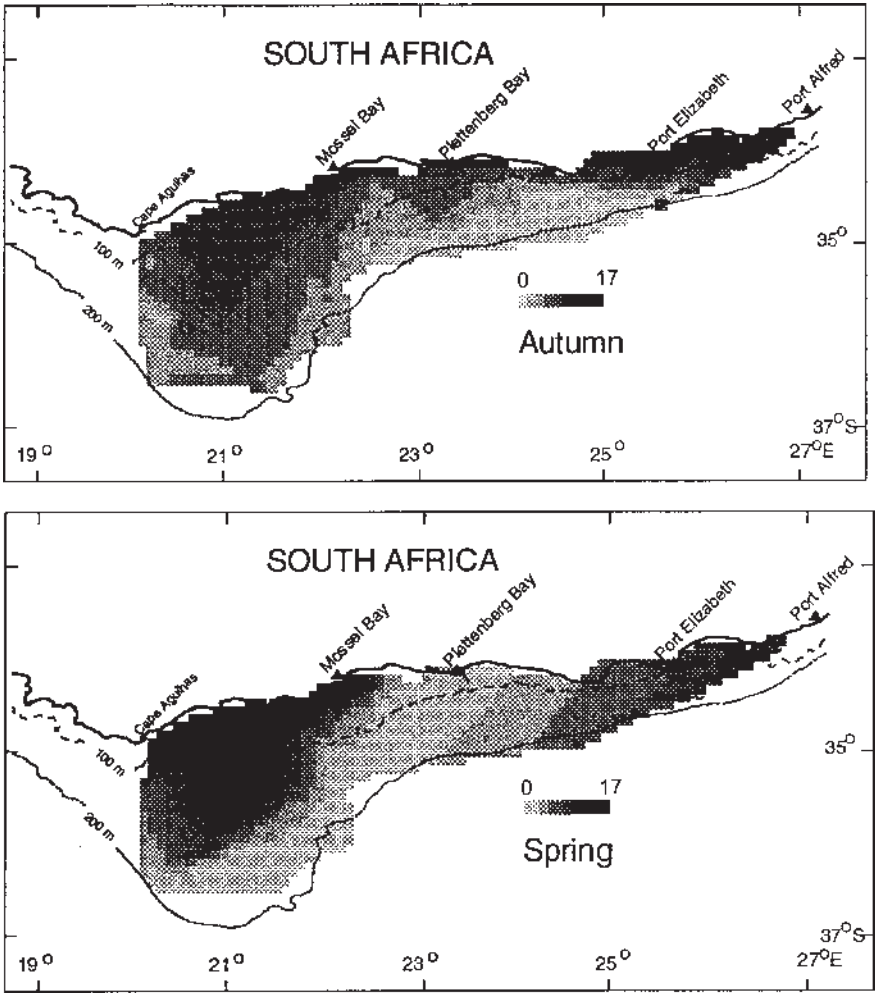


Figure 3. Bottom temperature profiles (°C) for spring and autumn 1993 on the Agulhas Bank, South Africa. The polygon coverage, represented as $5 \times 5'$ squares, was generated by using polynomial interpolation of the raw point data.

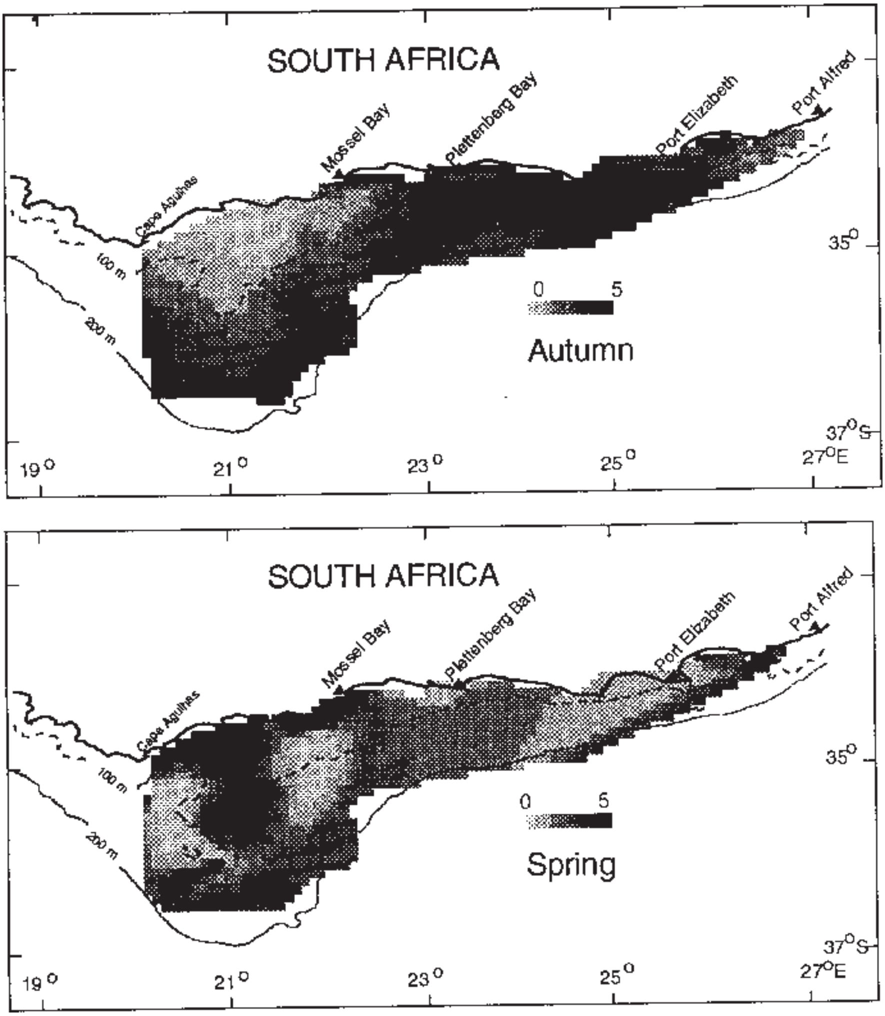


Figure 4. Dissolved oxygen profiles (ml/L) for spring and autumn 1993 on the Agulhas Bank, South Africa. The polygon coverage, represented as 5 × 5′ squares, was generated by using polynomial interpolation of the raw point data.

Bank water remained relatively warm with relatively high dissolved oxygen levels.

Qualitative Appraisal

Distinct patterns of distribution were evident within the various life history stages of *P. laniarius* when all trawl surveys were pooled (Fig. 2). Juvenile fish were distributed in a narrow area over the central Agulhas Bank within the 60-90 m isobath forming a nursery area. While small patches of juveniles were noticeable off Mossel Bay and Plettenberg Bay, these were of low abundance. Subadult fish showed an intermediary distribution pattern between the juvenile and adult life history stages, becoming more widely distributed than the juvenile fish yet more restricted than adult fish. They were also distributed predominantly over the central Agulhas Bank up to Plettenberg Bay, although at depths greater than juvenile fish. Adult fish were widely distributed, inhabiting deeper waters. Three distinct areas of high adult biomass were noticed, off the Central Agulhas Bank, west of Plettenberg Bay and east of Port Elizabeth. This confirms anecdotal evidence from fisherman who report high spawner biomass in these areas. Changes with respect to seasonality were also evident, with fish of all life history stages inhabiting areas with both higher temperatures and dissolved oxygen levels during the spring surveys (Table 2).

Quantitative Appraisal

The application of GAM allows for the investigation of the effect selected covariates have on fish density. The use of preliminary Generalized Linear Modeling (McCullagh and Nelder 1989), also with an underlying Poisson distribution, highlighted linear effects. Despite being statistically significant, the GAMs incorporating the linear fits provided higher deviations with respect to the null model, and nonlinear cubic splines were used in all further analyses.

A summary of the GAM scatterplot smooths, incorporating all survey data for the three life history stages, are presented in Figs. 5-7 for juveniles, subadults, and adults, respectively. In all three life history stages analyzed, fish density was shown to increase with longitude and decrease with respect to latitude. In juvenile fish, there was a steady increase in abundance with increased dissolved oxygen concentrations. This became less evident as fish grew larger, having little effect on the GAM smooths in the subadult and adult life history stages (Tables 3-5). Temperature was shown to have little effect on fish density in all life history stages. A preferred depth was clearly evident with fish moving into deeper water as they grew larger. Highest abundance occurred at approximately <90 m for juveniles, at approximately 90 m for subadults, and at approximately 100 m for adult fish. The incorporation of a seasonal factor into the GAM highlighted the fact that abundance was greatest during autumn in all life

Table 2. Qualitative analysis of preferred zones for three physical variables obtained by visual inspection of overlaying physical and fish density coverages.

Year	Temperature (°C)	Dissolved oxygen (ml/L)	Depth (m)	Temperature (°C)	Dissolved oxygen (ml/L)	Depth (m)
			Juveniles (< 13 cm TL)			
		Spring			Autumn	
1988	–	–	–	10.0-13.0	3.0-4.5	60-95
1989	–	–	–	>10.0	2.0-4.0	80-90
1990	10.0-15.0	>3.5	60-100	>10.0	2.0-4.0	50-100
1991	>14.0	>4.0	80-100	9.5-13.0	>3.0	70-100
1992	>13.0	>4.5	40-90	10.5-11.5	3.0-4.0	80-120
1993	>14.0	>4.8	60-95	10.0-16.0	3-4.3	60-95
1994	>12.0	>4.0	50-95	>10.0	3.0-4.0	70-90
1995	[a]	[a]	60-100	>10.0	3.0-4.0	60-95
			Subadults (13-23 cm TL)			
		Spring			Autumn	
1988	–	–	–	10.0-13.0	2.5-5.0	60-120
1989	–	–	–	>9.0	2.0-5.0	40-180
1990	9.0-16.0	>3.5	60-200	>9.0	2.0-5.0	40-180
1991	>13.0	>4.0	70-110	9.5-13.0	>2.5	70-100
1992	>10.5	>4.0	50-130	10.5-11.5	2.5-5.0	60-130
1993	>9.5	>4.2	50-130	6.0-11.2	3.0-4.3	80-140
1994	[a]	[a]	60-160	>9.5	2.5-4.5	60-180
1995	>10.0	3.5-5.0	60-120	9.0-11.0	3.0-5.0	60-160
			Adults (> 23 cm TL)			
		Spring			Autumn	
1988	–	–	–	10.0-15.0	2.5-5.0	40-180
1989	–	–	–	>9.0	2.0-5.0	40-180
1990	9.0-16.0	>3.5	60-200	>9.0	2.0-6.0	50-180
1991	>10.5	>3.5	40-140	>9.0	>2.5	60-160
1992	>10.0	>3.9	50-130	10.0-12.0	2.5-5.0	60-180
1993	>10	>4.0	50-180	>9.5	3.0-4.5	80-140
1994	[a]	[a]	60-180	>9.5	2.5-4.5	60-180
1995	>9.5	3.5-5.0	60-120	9.0-11.0	3.0-5.0	60-180

Analysis was conducted for different life history stages of *Pterogymnus laniarius* on the Agulhas Bank, South Africa during spring and autumn biomass surveys between 1988 and 1996.

[a] Physical data too sparse for analysis.

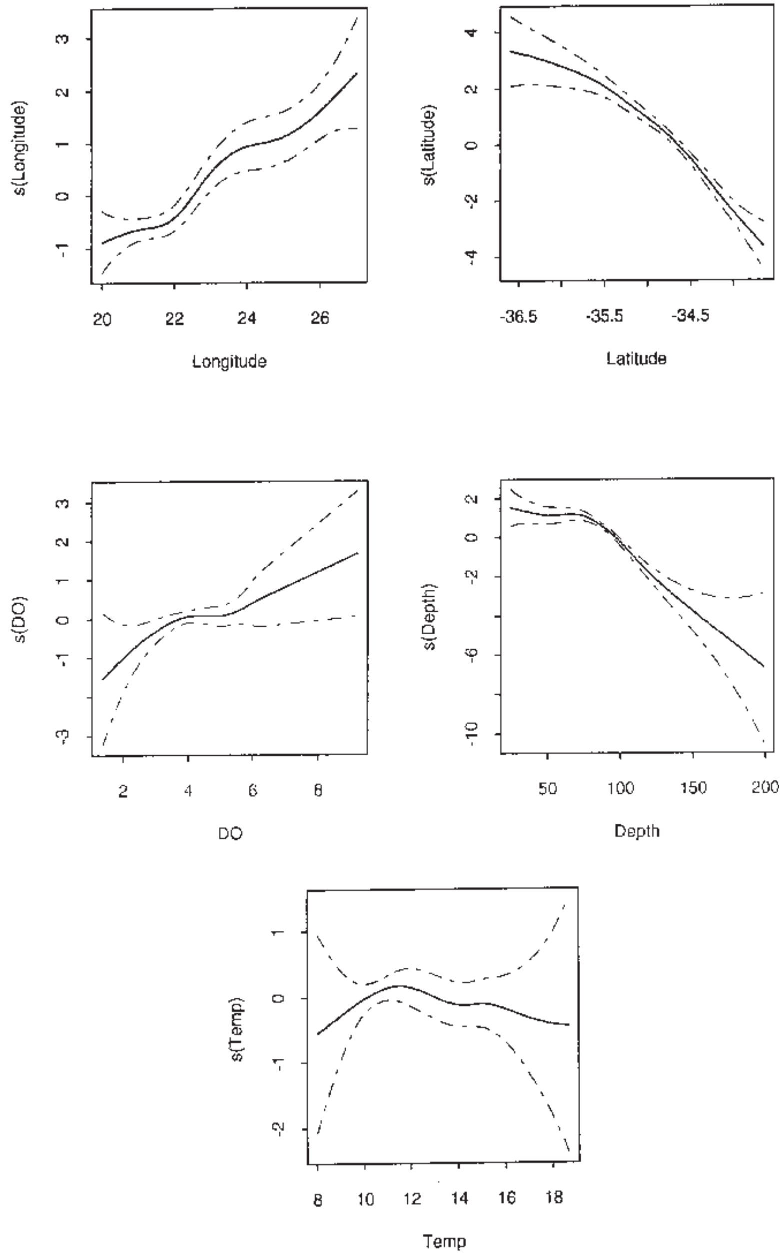


Figure 5. *GAM scatterplot smooths for juvenile Pterogymnus laniarius (<13 cm TL) density in response to various covariates for all surveys analyzed between 1988 and 1995 on the Agulhas Bank, South Africa. Dashed lined represent twice-standard error bands.*

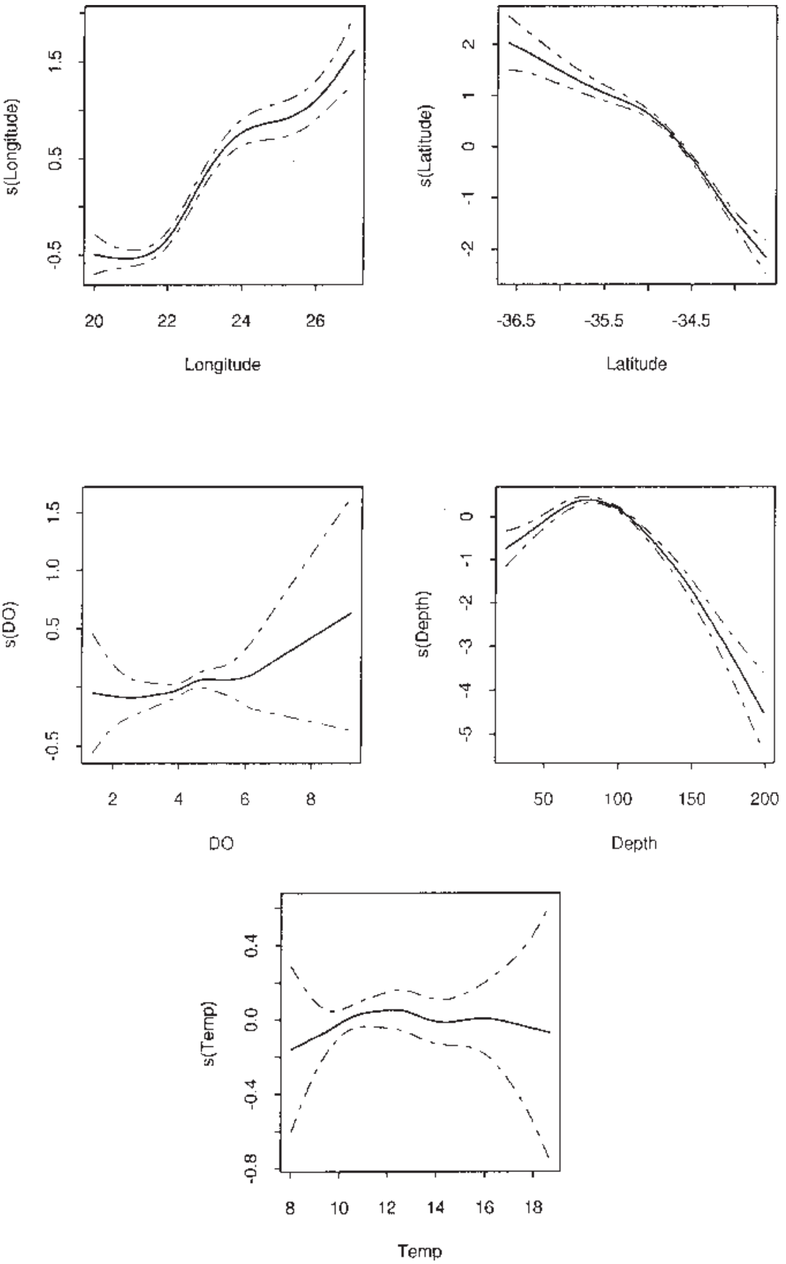


Figure 6. GAM scatterplot smooths for subadult Pterogymnus laniarius (13-23 cm TL) density in response to various covariates for all surveys analyzed between 1988 and 1995 on the Agulhas Bank, South Africa. Dashed lines represent twice-standard error bands.

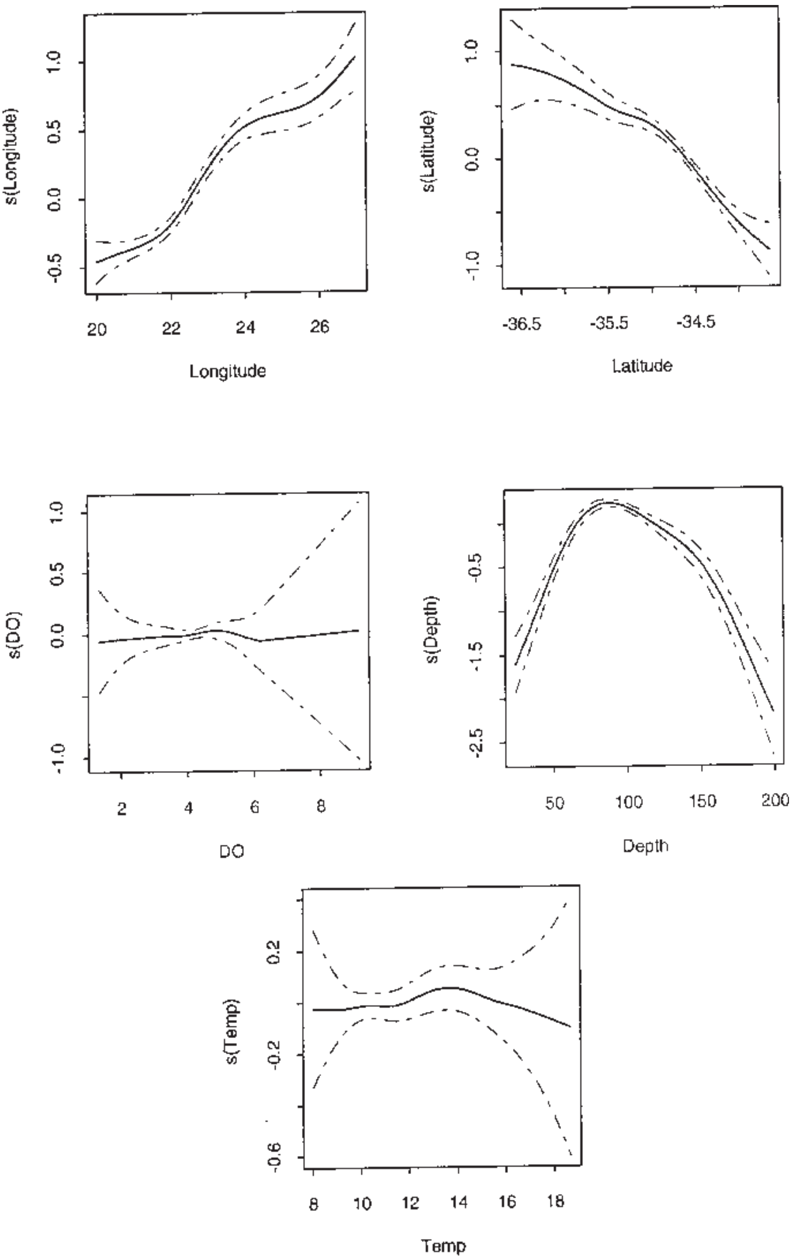


Figure 7. GAM scatterplot smooths for adult Pterogymnus laniarius (>23 cm TL) density in response to various covariates for all surveys analyzed between 1988 and 1995 on the Agulhas Bank, South Africa. Dashed lines represent twice-standard error bands.

Table 3. Nonparametric significance of various covariates to the GAM fitted to juvenile *Pterogymnus laniarius* (<13 cm TL) density.

Year	Longitude	Latitude	Temperature	Dissolved oxygen	Depth
Spring					
1990	<0.001	0.015	<0.001	<0.001	<0.001
1991	<0.001	<0.001	0.162	0.003	0.022
1992	0.006	<0.001	0.281	0.053	0.055
1993	0.005	0.009	0.004	0.093	<0.001
1994	0.014	<0.001	0.008	0.019	<0.001
1995	0.005	0.007	0.048	0.074	<0.001
Autumn					
1988	0.983	0.367	0.484	0.610	0.306
1989	0.022	0.249	0.001	0.467	<0.001
1990	<0.001	<0.001	0.005	0.087	0.212
1991	0.137	<0.001	<0.001	0.137	<0.001
1992	[a]	[a]	[a]	[a]	[a]
1993	<0.001	<0.001	0.089	0.151	<0.001
1994	0.004	<0.001	<0.001	0.025	0.490
1995	<0.001	<0.001	<0.001	0.003	0.001
All years	<0.001	<0.001	0.003	<0.001	<0.001

[a] Physical data was too sparse for analysis.

history stages analyzed. This is possibly an effect of sampling deeper areas during the autumn surveys.

Nonparametric statistical significance of the various covariates using an appropriate χ^2 statistic (Hastie and Tibshirani 1990) for each life history stage and survey analyzed is summarized in Tables 3-5 . In the juvenile data set, four of the five covariates used in the analysis showed a consistent significant effect on the GAM. While latitude and longitude consistently illustrated the presence of a nursery area on the central bank, the significant effects of these covariates appear to be merely descriptive. Dissolved oxygen generally had an insignificant effect in determining the density of juveniles, with one third of the surveys analyzed being significant. In contrast, both depth and temperature were significant in most surveys. A similar trend was evident with the subadult fish. Both dissolved oxygen and temperature were generally insignificant in the adult GAM. The eurytopy of the adult fish was stressed, and mature fish appear to be able to tolerate a wide range of physical conditions. This, therefore, provides the mechanism for adult fish to be able to move eastward and colonize vast areas of the eastern bank. In all life history stages, depth was found to be highly significant in all the surveys analyzed.

Table 4. Nonparametric significance of various covariates to the GAM fitted to subadult *Pterogymnus laniarius* (13-23 cm TL) density.

Year	Longitude	Latitude	Temperature	Dissolved oxygen	Depth
Spring					
1990	<0.001	<0.001	0.002	0.003	<0.001
1991	0.004	<0.001	<0.001	0.097	<0.001
1992	<0.001	0.002	0.177	0.003	0.010
1993	0.016	0.005	0.071	0.277	<0.001
1994	0.038	<0.001	0.001	0.082	<0.001
1995	0.102	<0.001	0.002	0.013	<0.001
Autumn					
1988	0.557	<0.01	0.097	<0.001	<0.001
1989	0.009	0.006	0.009	0.001	<0.001
1990	0.028	0.002	0.185	0.397	<0.001
1991	<0.001	0.225	0.215	0.089	0.001
1992	[a]	[a]	[a]	[a]	[a]
1993	<0.001	0.145	0.225	0.273	<0.001
1994	0.149	0.009	0.008	0.454	<0.001
1995	<0.001	0.002	0.002	0.149	<0.001
All years	<0.001	<0.001	0.043	0.029	<0.001

[a] Physical data too sparse for analysis.

Optimum Fishing Areas

During spring and autumn, at least 60% of the spawner biomass was situated over the central Agulhas Bank (Table 6). Those areas suitable for commercial exploitation, composing at least 80% of spawner biomass, are presented in Fig. 8. In both the scenarios presented, the eastern Agulhas Bank consistently contained the highest proportion of spawner biomass due to its distance from the nursery area on the central Agulhas Bank. If the proportion of spawner biomass was specified to be >85%, the model only selected the eastern Agulhas Bank and western edge of the central Agulhas Bank. When the criteria for area selection was reduced to contain only those areas with >80% spawner biomass, a larger portion of the Agulhas Bank was chosen. In both scenarios, the high juvenile and subadult abundance on the mid-central Agulhas Bank excluded these areas as suitable for fishing effort.

Discussion

From the data presented, it appears that *P. laniarius* has a distinct ontological shift in distribution with respect to size (and age). This is noticeable

Table 5. Nonparametric significance of various covariates to the GAM fitted to adult *Pterogymnus laniarius* (>23 cm TL) density.

Year	Longitude	Latitude	Temperature	Dissolved oxygen	Depth
Spring					
1990	0.104	<0.001	<0.001	0.212	<0.001
1991	0.458	0.001	0.387	0.047	<0.001
1992	0.136	0.177	0.124	0.165	<0.001
1993	0.030	0.231	0.020	0.047	<0.001
1994	0.110	0.063	<0.001	0.025	<0.001
1995	0.071	<0.001	0.005	0.614	<0.001
Autumn					
1988	0.097	0.009	0.132	0.839	<0.001
1989	0.353	<0.001	0.008	0.195	<0.001
1990	0.015	<0.001	0.324	0.007	<0.001
1991	0.257	0.040	0.247	0.177	<0.001
1992	[a]	[a]	[a]	[a]	[a]
1993	0.009	0.229	0.084	0.121	<0.001
1994	0.024	0.723	0.238	0.255	<0.001
1995	<0.001	0.008	0.012	0.003	<0.001
All years	<0.001	<0.001	0.227	0.321	<0.001

[a] Physical data too sparse for analysis.

Table 6. Proportion of spawner biomass that is situated over the central (20°-23°E) and eastern (23°-27°E) portions of the Agulhas Bank.

	Spring / summer surveys		Autumn / winter surveys	
	Central Bank	Eastern Bank	Central Bank	Eastern Bank
1988	–	–	0.35	0.65
1989	–	–	0.64	0.36
1990	0.63	0.37	0.50	0.50
1991	0.57	0.43	0.63	0.37
1992	0.57	0.43	[a]	[a]
1993	0.66	0.34	0.67	0.33
1994	[b]	[b]	0.66	0.34
1995	0.56	0.44	0.65	0.35
Average	0.60	0.40	0.59	0.41

Estimates were calculated using the GAM predicted estimates of biomass.

[a] No convergence in GAM; [b] Not enough coverage of Agulhas Bank for effective interpolation.

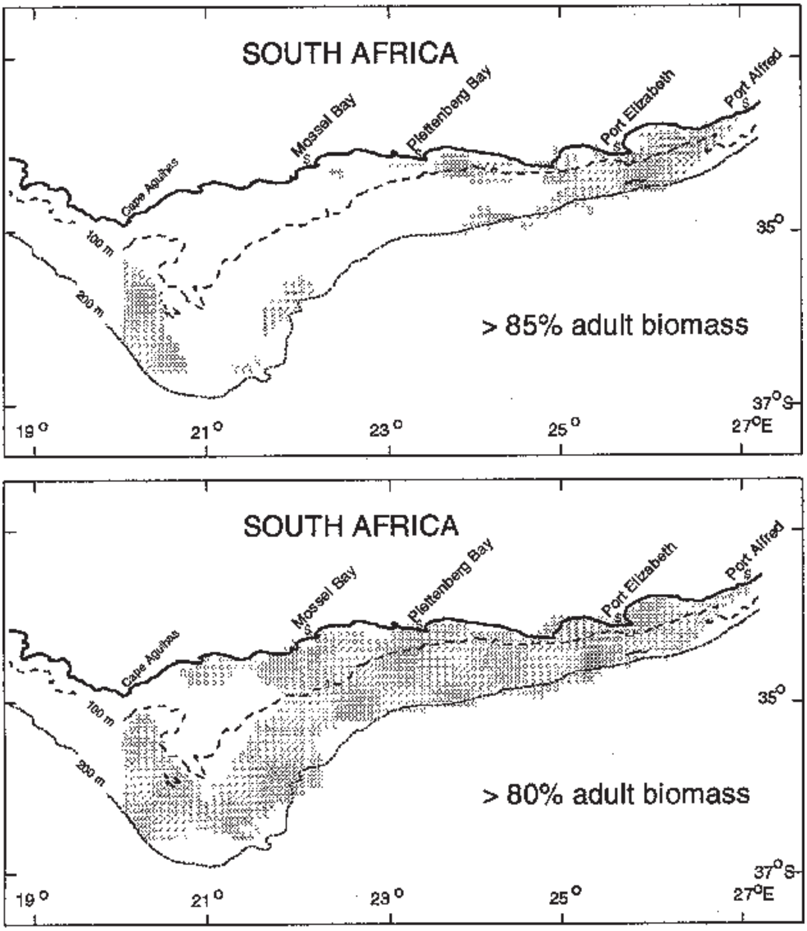


Figure 8. *Optimum areas to fish Pterogymnus laniarius which minimizes the impact on juvenile fish while directing effort toward spawner biomass. The hatched areas have total biomass composed of at least 85% adults (top) and at least 80% adults (bottom). Data used in the analysis were GAM biomass estimates which were pooled and averaged for all surveys between 1988 and 1995.*

in the distinct nursery area for immature fish over the central Agulhas Bank. After sexual maturation, corresponding to the stage when fish feed predominantly on harder shelled, soft-stratum prey (Booth and Buxton 1997), a large proportion of the adult population migrated eastward.

Results obtained from both the qualitative and quantitative methods were similar. The combined trends, therefore, provide a reasonable indication of the functional relationship between the abundance of fish and its predictive covariates. It was noted that in immature fish, density was significantly affected by both temperature and depth. In mature fish, only depth provided a significant effect (with the exception of the spatial covariates) to the GAM fit. Dissolved oxygen was found to be insignificant in the subadult and adult life history stages and was probably due to the relatively high dissolved oxygen levels over the entire Agulhas Bank. This is in contrast to the upwelling dominated Benguela system where anoxic conditions can prevail (Smale et al. 1993, Roberts and Sauer 1994). While there appears to be some autocorrelation between temperature and depth (with temperature decreasing with depth), depth was consistently a significant effect in all the GAM fits. It is, therefore, suggested that depth is a primary factor determining the abundance of *P. laniarius* on the Agulhas Bank. Similarly, depth was the covariate that was found to determine flatfish abundance in the Bering Sea (Swartzman et al. 1992) and gadoid and flatfish abundance on the Georges Bank (Murawski and Finn 1988).

The question still arises why small, immature fish are distributed over the central Agulhas Bank. If depth were the principal factor in determining abundance, their distribution should include the entire Agulhas Bank. A similar question arises as to why the larger fish move eastward after sexual maturation. These questions stress the use of exploratory data analysis using generalized statistical and spatial methods which allow for trends to be observed before hypotheses are formulated.

If both surface and bottom current profiles are investigated (see review by Boyd and Shillington 1994), an alternate scenario can be illustrated. *P. laniarius* has been shown to have a nonseasonal pattern with all mature individuals spawning throughout their range all year (Booth and Buxton 1997). This is possibly a response to the relatively stable oceanographic environment which they inhabit. Spawned eggs and pre-metamorphosed planktonic larvae from eastward distributed spawning adults would, therefore, be carried westward by the strong Agulhas Current before being advected to the coast by the anti-cyclonic gyre over the central Agulhas Bank. Similarly, eggs and larvae from central Agulhas Bank spawning adults would merely be advected toward the coast. This is similar to the conclusion presented by Werner et al. (1993) using a three-dimensional circulation model which in part explained the distribution of cod and haddock larvae on Georges Bank. The decrease in surface current strength over the central Agulhas Bank could provide a deposition zone for post-metamorphosed larvae, and the decreased bottom current strength would prevent fish from being moved far from their deposition site and lost to the West Coast Benguela system. As the fish grow and mature, they become large enough to tolerate stronger currents and also have the ability to feed on a greater variety of hard-shelled prey. As a density-dependent response, adult fish would reduce conspecific competition and move eastward to colonize larger areas of suitable habitat. Overall, *P. laniaris* would restrict themselves to a specific depth range to avoid stress associated with shallow, coastal wind-driven upwelling and the strong Agulhas current with its shelf break–induced upwelling (Lutjeharms et al. 1996).

It has been shown in a variety of studies that the distribution of populations is strongly determined by some underlying habitat association, preferred physical conditions, proximity to food sources, or areas where spawning success will be maximized (Murawski and Finn 1988, Mahon and Smith 1989, D'Amours 1993, Perry and Smith 1994, Le Clus et al. 1996). In South Africa, there is paucity of knowledge on the habitat types and the habitat preferences of associated fish assemblages on the Agulhas Bank. While it is known that there are extensive areas of hard, low-profile reef (Badenhorst and Smale 1991, Smale et al. 1993, Le Clus et al. 1996) which appears to be the preferred habitat for *P. laniarius*, there is a distinct need for this information. Data that is available is restricted to surficial sedimentary deposits. In this regard, a concerted effort is needed in this area before the spatial distribution of any species inhabiting the Agulhas Bank can be fully understood.

GIS technology offers the facility to combine many biological data sets and relate them to system-wide observations of habitat. Furthermore as habitat is not the resource of value, relating biological consequences to the habitat allows for the ease of management of the habitat via their biological indicators. In the case of the Agulhas Bank, as more habitat information becomes available together with knowledge of their associated fish assemblages, it can be added as a new coverage to the GIS. The updated GIS can then easily be reanalyzed and interpreted. As a result, the spatial distribution and abundance of *P. laniarius* will become more fully understood. Understanding the influence of the marine environment, a knowledge of the habitat types and preferences can therefore facilitate better management of *P. laniarius* and other sympatric species.

Optimal management areas can be obtained by selecting those areas which are the most suitable for fishing. Using a GIS, the biomass of the stock can be easily disaggregated by region or life history stage, thereby facilitating and improving age-structured modeling (Booth, in prep.). The most suitable areas would therefore include those in which fishing effort (and hence fishing mortality) is reduced on vulnerable life history stages. These would include immature fish which have discrete nursery areas or annual migrations of spawner biomass to specific spawning areas. In the case of *P. laniarius*, only the former constraint applies as adult fish spawn throughout the year throughout their range (Booth and Buxton 1997). As *P. laniarius* is also caught predominantly as bycatch in the demersal trawl fishery (Booth and Punt 1998), optimal areas for harvesting should also include most of the areas fished by the existing demersal trawl fleet. This includes most of the Agulhas Bank as the trawl fleets are based in Mossel Bay and Port Elizabeth. In both optimum fishing scenarios presented in this study, the fleet should merely be restricted from fishing on the mid-central Agulhas Bank area.

The use of spatial analysis to identify trends in fish distribution and abundance and to incorporate trends in assessment models and fisheries management is obvious. Similarly, GIS has enormous value for organizing biological data and discovering their underlying relationships. This is a result of each biological data set being only one temporary value because it is a specific sample drawn from a population or community that is highly variable in both time and space. Despite noticing the potential of GIS, Meaden and Kapetsky (1991), Simpson (1992), and Meaden (1996) noted that they have not been fully utilized; instead, they are restricted to terrestrial or freshwater applications. This is primarily a response to high costs associated with the collection of marine biological, physical-chemical, sediment, and reef data and in synthesizing the vast databases available into a compatible and comparable format (Caddy and Garcia 1986, Meaden 1996). GIS includes a complex technology that can enhance the utility of diverse data at both regional or larger scales and can facilitate improved fisheries management decisions by including the introduction of best management practices.

Acknowledgments

The authors wish to thank the director of the Sea Fisheries Research Institute and Sarah Walmsley, Clifford Hart, and Sharon du Plessis for extracting the South Coast biomass survey data. Trevor Hastie, Stanford University, and David Borchers, St. Andrews University, are thanked for giving assistance on the application of Generalized Additive Modeling. Thomas Hecht and the two reviewers are thanked for their comments on previous drafts of the manuscript. This project was funded by the International Foundation for Science (Grant A/2776-1) and the Benguela Ecology Programme.

References

Badenhorst, A., and M.J. Smale. 1991. The distribution and abundance of seven commercial trawlfish from the Cape south coast of South Africa, 1986-1990. S. Afr. J. Mar. Sci. 11:377-393.

Booth, A.J., and C.D. Buxton. 1997. The biology of the panga *Pterogymnus laniarius* (Teleostei: Sparidae) on the Agulhas Bank, South Africa. Environ. Biol. Fishes 49:207-226.

Booth, A.J., and A.E. Punt. 1998. Evidence for rebuilding in the panga stock on the Agulhas Bank, South Africa. Fish. Res. 34:403-421.

Boyd, A.J., and F.A. Shillington. 1994. Physical forcing and circulation patterns on the Agulhas Bank. S. Afr. J. Sci. 90:114-122.

Caddy, J.F., and S. Garcia. 1986. Fisheries thematic mapping: a prerequisite for intelligent management and the development of fisheries. Oceanogr. Trop. 21:31-52.

D'Amours, D. 1993. The distribution of cod (*Gadus morhua*) in relation to temperature and oxygen level in the Gulf of St. Lawrence. Fish. Oceanogr. 2:24-29.

Deriso, R.B., T.J. Quinn, and P.R. Neal. 1985. Catch-age analysis with auxillary information. Can. J. Fish. Aquat. Sci. 42:815-824.

Hastie, T., and R. Tibshirani. 1986. Generalized additive models. Stat. Sci. 1:297-318.

Hastie, T., and R. Tibshirani. 1990. Generalized additive models. Chapman and Hall, London.

Hatanaka, H., T. Sato, J. Augustyn, A. Payne, and R. Leslie. 1983. Report on the Japan/South African joint trawling survey on the Agulhas Bank in November/December 1980. Far Seas Fish. Res. Lab. Rep. 12. 73 pp.

Japp, D.W., P. Sims, and M.J. Smale. 1994. A review of the fish resources of the Agulhas Bank. S. Afr. J. Sci. 90:123-134.

Le Clus, F., H.F-K.O. Hennig, and J. Rodgers. 1996. Bathymetry and sediment type effects on catch rates of *Austroglossus pectoralis* (Soleidae) on the inner central Agulhas bank. S. Afr. J. Mar. Sci. 17:79-92.

Lutjeharms, J.R.E., A.A. Meyer, I.J. Ansorge, G.A. Eagle, and M.J. Orren. 1996. The nutrient characteristics of the Agulhas Bank. S. Afr. J. Mar. Sci. 17:253-274.

Mahon, R., and R.W. Smith. 1989. Demersal fish assemblages on the Scotian Shelf, Northwest Atlantic: Spatial distribution and persistence. Can. J. Fish. Aquat. Sci. 46(Suppl. 1):134-152.

McCullagh, P., and J. A. Nelder. 1989. Generalized linear models. 2nd edn. Chapman and Hall, London.

Meaden, G.J. 1996. Potential for geographical information systems (GIS) in fisheries management. In: B.A. Megrey and E. Moksness (eds.), Computers in fisheries research. Chapman and Hall, London, pp. 41-79.

Meaden, G.J., and J.M. Kapetsky. 1991. Geographical Information Systems and remote sensing in inland fisheries and aquaculture. FAO Fish. Tech. Pap. 318. 262 pp.

Murawski, S.A., and J.T. Finn. 1988. Biological bases for mixed-species fisheries: Species co-distribution in relation to environmental and biotic variables. Can. J. Fish. Aquat. Sci. 45:1720-1735.

Perry, R.I., and S.J. Smith. 1994. Identifying habitat associations of marine fishes using survey data: An application to the Northwest Atlantic. Can. J. Fish. Aquat. Sci. 51:589-602.

Pope, J.G., and J.G.Shepherd. 1985. A comparison of the performance of various methods for tuning VPAs using effort data. J. Cons. Perm. Int. Explor. Mer 42:129-151.

Press, W.H., S.A. Teukolsky, W.T. Vettering, and B.P. Flannery. 1997. Numerical recipes in C: The art of scientific computing. 2nd edn. Cambridge University Press, Cambridge U.K. 994 pp.

Roberts, M.J., and W.H.H. Sauer. 1994. Environment: The key to understanding the South African Chokka squid (*Loligo vulgaris reynaudii*) life cycle and fishery. Antarctic Sci. 6:249-258.

Simpson, J.J. 1992. Remote sensing and geographical information systems: Their past, present and future in global fisheries management. Fish. Oceanogr. 1:238-280.

Smale, M.S., B.A. Roel, A. Badenhorst, and J.G. Field. 1993. Analysis of the demersal community of fish and cephalopods on the Agulhas Bank, South Africa. J. Fish. Biol. 43(Suppl. A):169-191.

Star, J.L., and J.E. Estes. 1990. Geographical Information Systems: An introduction. Prentice Hall, NJ.

Swartzman, G., C. Huang, and S. Kaluzny. 1992. Spatial analysis of Bering Sea groundfish survey data using Generalized Additive Models. Can. J. Fish. Aquat. Sci. 49:1366-1378.

Uozumi, Y., H. Hatanaka, T. Sato, A. Payne, and J. Augustyn. 1985. Report on the Japan/South African joint trawling survey on the Agulhas Bank in June 1982. Far Seas Fish. Res. Lab. Rep. 13:1-77.

Uozumi, Y., H. Hatanaka, T. Sato, J. Augustyn, A. Payne, and R. Leslie. 1981. Report on the Japan/South African joint trawling survey on the Agulhas Bank in November/December 1981. Far Seas Fish. Res. Lab. Rep. 11:1-91.

Werner, F.E., F.H. Page, D.R. Lynch, J.W. Loder, R.G. Lough, R.I. Perry, D.A. Greenberg, and M.M. Sinclair. 1993. Influences of mean advection and simple behaviour on the distribution of cod and haddock early life stages on the Georges Bank. Fish. Oceanogr. 2:43-64.

Comparing Different Information Sources in a Multispecies Context

Gunnar Stefánsson
Marine Research Institute, Reykjavík, Iceland

Abstract

Different data sources tend to give apparently incompatible information on the state of an ecosystem. Various likelihood functions can be used to incorporate many data sources in a stock estimate. Such likelihoods should be viewed as components in an overall likelihood function. This paper compares the estimates of stock trends and values of likelihood components as a function of what data source is used as the primary criterion in a stock estimation procedure. A multispecies model, Bormicon, is used as the vehicle for these comparisons, based on indications on stock size, growth, migration, and consumption from a variety of different data sources. The model accommodates data from acoustic surveys, groundfish surveys, mean length at age in catches and surveys, and stomach content data. This is an area-based, multispecies, multifleet model which includes growth, consumption, and migration in addition to the effects of fishing. It is seen that different data sources may give considerably different views of the world in this complex modeling scenario, but in fact this is also the case even in the simplest of fisheries models. Most of the differences are seen to stem from model errors, where common fisheries models exclude certain important factors which are needed when considering many data sources.

Introduction

With the advent of multispecies models which are orientated toward modeling several aspects of an ecosystem comes a need to consider many more sources of information than has been necessary earlier. Thus, the simplest stock-assessment models such as Virtual Population Analysis, VPA, (Gulland 1965) and catch-curve analysis, (Beverton and Holt 1957) require only catches in numbers at age for historical analysis and a single survey index is often used to estimate the fishing mortality in the last year. An early multispecies model, MSVPA (Helgason and Gíslason 1979), also included stomach content data but assumed these to be without error, at least after initial smoothing. It was further realized that MSVPA, as implemented, did not predict the effects of consumption and food availability on growth and did not incorporate spatial variation in any way. The first extensions to accommodate this, such as MULTSPEC (Tjelmeland and Bogstad 1989), did not do so through a model-based comparison of model growth to data but rather through assumptions, and the only comparison to data in a statistical sense was through the use of a likelihood-type component involving stomach data (Bogstad et al. 1992).

These problems have been alleviated to some extent in the implementation of Bormicon, a BOReal MIgration and CONsumption model. The background of Bormicon was defined in general terms in the form a quest for an appropriate description of an arcto-boreal ecosystem such as the Icelandic waters (Stefánsson and Pálsson 1998) based on available knowledge on food chains and trophic levels (Magnússon and Pálsson 1989, 1991; Pálsson 1983). The model has been described in detail from a biological and general modeling viewpoint with a case study (Stefánsson and Pálsson 1997), with a user's manual (Stefánsson et al. 1997a) and a programmer's manual (Stefánsson et al. 1997b). This model is an area-based multispecies, multifleet model which incorporates predation along with growth as a function of food consumed and can be used to estimate parameters based on a collection of very general log-likelihood functions.

The following describes some of the issues which arise when complicated models like Bormicon are used. This includes problems due to apparent data conflicts as well as concerns on how data can best be viewed for diagnostic (and presentation) purposes when the output volume is more than can be handled easily in traditional graphs and tables.

Conflicting Evidence in Simple Models and Data Sets

Given the multitude of data sources used in Bormicon, one may expect some difficulties in reconciling all of them since they are linked through highly complicated processes. Such difficulties may appear as apparently conflicting information in the various sources.

The issue of conflicting evidence can be seen in much simpler models, however, as will first be demonstrated. It is further argued that a more appropriate view of the conflict is that there are errors in the model definition and that the apparent conflict should be used to redefine the model under consideration.

Example: Conflicting Indications within a Survey Data Set—Model Error

A groundfish survey typically provides information on the abundance of several species in terms of numbers caught per unit effort, by age and year. The resulting table can be used as it stands as a sole indicator of various stock attributes. Simple models for this purpose include multiplicative models (Shepherd and Nicholson 1991), which are particularly useful for obtaining a quick view of the inherent variability in the data due to their simplicity and ability to incorporate most of the structure in the data. The model is simply an analysis of variance of the logged survey catches in numbers using age, year, and year class as factors. The model parameters are not all estimable but this is of no consequence when the emphasis is on the estimation of residual variance.

An example of a survey data set is given in Table 1. This data set includes ages 1-9 and it is a priori quite possible that the variances of the survey indices vary from one age to the next. According to well-known statistical theory (Scheffé 1959), an incorrect variance specification is usually not overly important. It is, however, quite important to obtain a reasonable specification of the linear predictor, i.e., the mean function.

Fitting the Shepherd-Nicholson model to all age groups yields an estimate of the standard error (on log-scale, i.e., a CV approximation) of some 28%, when estimated as a single value across all ages. In fact, when this model is used and the same CV is computed only across ages 3-7, it is found to be 27% and thus the number of fish of these ages does not appear to be much more accurately known than the older and younger ages (Table 2). It is therefore somewhat curious that, when only ages 3-7 are used in the criterion for the estimation of the model, the estimated CV on these ages drops to 21%. An important use of this type of model is precisely to estimate variances (Pope 1993) and hence it is of some concern that these estimates can be quite dependent on the age range as in this case. In this particular case, however, the individual parameter estimates are fairly stable except for the year effects, which are considerably more variable when ages 3-7 are used, i.e., the reduced age range results in a portion of the error variability being explained by the year effect (Fig. 1).

The most likely explanation of this apparent data conflict is that the model in question is far too simple to capture the dynamics of the age-year-cohort system and that the built-in restrictions on, e.g., selectivity or effort changes, are not valid for the data set at hand.

Example: Conflicting Data Sources in Single-Species VPA Tuning—Model Error

An assessment of a fish stock typically involves a number of different data sets, common ones being catches, survey, and (commercial) CPUE indices in numbers by age and year. Many methods of assessment can include several "tuning fleets," i.e., survey and CPUE time series. Although quite variable, the usual reason for including a CPUE series in the first place is the lack of survey information, or high variability in the survey index. In this case the CPUE data may be an important component in the assessment and may drive the assessment results.

Table 1. **Indices by age (1-9) for cod in Icelandic waters from annual groundfish survey (1985-1995).**

	Age								
Year	1	2	3	4	5	6	7	8	9
85	16,775	36,310	34,532	54,825	51,526	14,296	6,292	1,832	783
86	13,289	53,926	82,614	22,073	18,257	18,298	4,876	1,420	389
87	2,653	25,221	91,880	79,754	17,747	10,333	8,510	1,592	364
88	1,699	5,466	65,640	90,562	50,341	6,908	3,025	2,535	274
89	1,933	15,072	16,735	69,564	56,152	19,589	1,695	571	326
90	3,505	11,627	18,147	8,067	16,026	20,228	6,613	463	140
91	1,750	14,537	15,943	21,716	7,589	9,237	9,393	1,076	99
92	233	20,685	34,874	16,063	10,386	3,737	2,369	1,366	161
93	1,319	2,330	26,851	33,630	10,065	6,815	1,145	703	349
94	8,579	10,968	5,782	17,996	12,779	3,094	1,544	229	100
95	492	28,643	17,393	4,173	11,233	7,075	1,379	436	54

Table 2. **Log-scale standard errors (×100) resulting from fitting Shepherd-Nicholson models using different criteria, i.e., fitting to different age-portions of the survey index table.**

Estimated CV for each age range	Criterion (ages used in sum of squares)		
	1-9	2-9	3-7
1-9	28		
2-9	25	22	
3-7	27	25	21

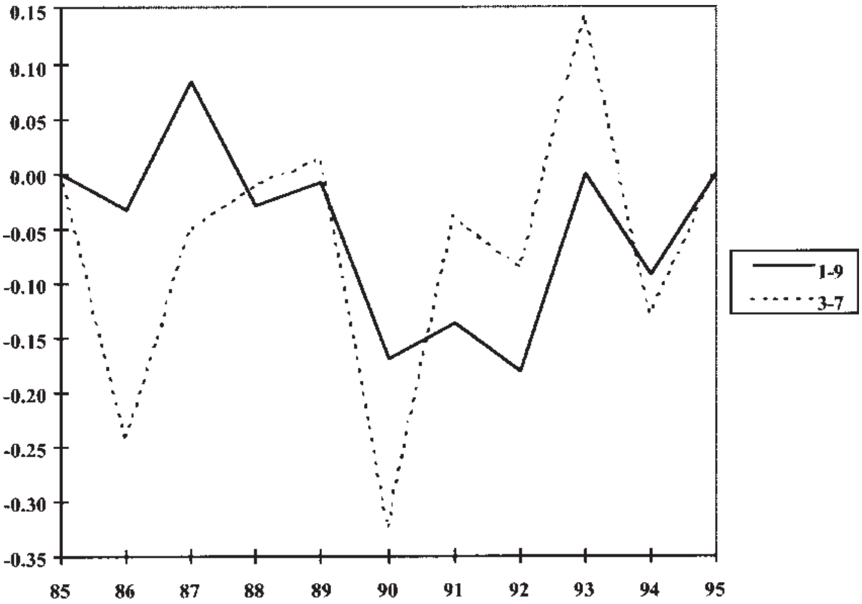


Figure 1. Year effects when fitting Shepherd-Nicholson models to groundfish survey data for cod in Icelandic waters, using data for age groups 1-9 or for ages 3-7.

Fleets typically increase their catchability by some percent per annum. An example of this effect is seen in Fig. 2, where the CPUE and VPA biomass values from Table 3 are plotted together. For this particular data set it is seen that the increase in fleet efficiency is some 4.7% per annum. It follows that a stock assessment based heavily on this index may lead to a correspondingly biased fishing mortality.

Of course the real problem with the commercial CPUE data is not the data themselves but rather the model used: It is assumed that there are no changes in catchability when these exist in reality. On the other hand, attempts to incorporate catchability changes in the assessment model leads to poorly determined estimates of stock size in the final year in most cases (ICES 1984).

Another problem which repeatedly occurs in "tuning" concerns the influence of the youngest (or, potentially, the oldest) ages on the assessment of intermediate age groups. This effect is best seen through the use of ADAPT (Gavaris 1988) or similar techniques, which have been reinvented on a number of occasions both in special cases (Halldórsson et al. 1986, de la Mare 1989) and the general case (Stefánsson 1988).

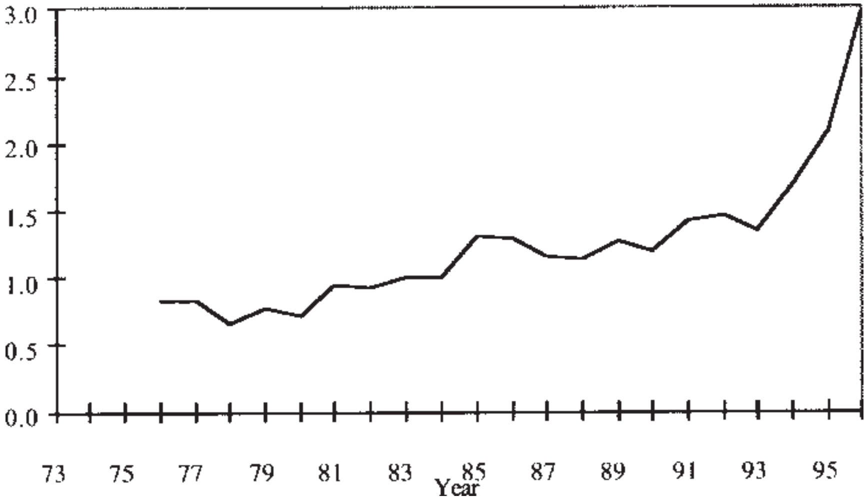


Figure 2. Ratio of catch per unit effort of bottom trawling fleet to VPA biomass for Iceland cod, indicating trend in catchability.

The catches are from ages 3-14 for this cod data set, whereas the survey includes ages 1-9. The ADAPT framework can be simplified so that the selection pattern is assumed known and equal to a VPA-based average of the first few years. The analysis then consists of a simple estimation of a single fishing mortality multiplier in the terminal year. Assuming the survey indices to be proportional to abundance and logarithmic errors, allows the estimation of the multiplier. If all ages (1-9) are used in the analysis by assuming that the age 1 index also provides an estimate of later recruitment at age 3, then the terminal fishing mortality is estimated at over 2 (Table 4). If, however, only ages 4 and older are used, then the estimate is around 1.25. This is a considerable difference and, for example, the estimate of the size of the last (1992) year class varies almost by a factor of 2 depending on which estimation criterion is used.

It follows that care needs to be exercised in how the various terms in the sum of squares are weighted. A common practice of either including or excluding data is equivalent to assigning either zero or full weight to each term and this is not "optimal" in any sense and hence one might want to try to estimate the appropriate weights. This particular model (ADAPT) is unlike many other statistical models, such as standard ANOVA, in that it is formally possible to obtain an unbiased estimate of each variance by eliminating all other variances from the estimation. These can subsequently be used as (inverse) weighting factors.

Table 3. Uncorrected catch per unit effort (U) from trawler logbooks and VPA biomass (B) for cod in Icelandic waters, 1973-1996.

Year	U	B
73	700	
74	800	
75	800	
76	800	950
77	1,000	1,214
78	800	1,205
79	1,000	1,290
80	1,100	1,548
81	1,200	1,263
82	900	979
83	800	795
84	900	901
85	1,200	921
86	1,100	854
87	1,200	1,035
88	1,200	1,056
89	1,300	1,024
90	1,000	836
91	1,000	700
92	800	547
93	800	591
94	1,100	650
95	1,300	620
96	2,000	675

Variance estimation can be done by minimizing the sum of squares for each age group separately to obtain each unbiased variance estimate and then proceeding to weight the sum to be minimized. When this is done the terminal fishing mortality varies from 1.03 to 1.71, where the results are clearly much more internally consistent than before. It follows that the choice of weights is crucial and attention needs to be given to how they are chosen. It is, however, also clear that the inclusion of highly variable age groups 1 and 2 leads to results that are considerably different from those indicated for age groups 3-9, even in the case when these young ages are down weighted in accordance with their apparent variance.

Table 4. Terminal fishing mortality (F) and recruitment of 1992 year class (R3) of Iceland cod based on various weighting of sums of squares of errors for fitted survey indices from selected age groups.

Age range	Unweighted F	Unweighted R3	Weighted F	Weighted R3
1-9	2.33	72	1.71	96
3-9	1.58	104	1.46	111
4-9	1.25	129	1.05	153
4-8	1.23	131	1.03	155
4-7	1.26	128	1.03	156

Although much more consistent than before, these results are quite variable, giving recruitment estimates which vary from 96 million to 156 million for the 1992 year class as 3 years old, depending on which age groups are included in the minimization.

A part of the reason for the inconsistency is of course variability in the data. Also important, however, is the model misspecification. In this case the assumption of a constant selection pattern links the abundance estimation of the different age groups. Survey data on a single age group can in fact be used to estimate the abundance of all year classes. It follows that data on one age group may drive the abundance estimate of another age group. Problems of the same nature are also seen in "standard" stock assessments when the selection pattern is free, but these inherent conflicts should be alleviated to some extent when the number of parameters is increased.

Multiple Data Sources in a Multispecies Model

The likelihood approach allows the combination of several sources of data. This approach has been re-invoked for combining multiple data sources in fisheries (Methot 1989) but has earlier been used for combining catch, CPUE, and effort information (Deriso et al. 1985) or for stock assessments where survey data plays a major role (Gavaris 1988).

Bormicon was designed as a formal method for combining several data sets into a comprehensive multispecies and multifleet assessment through the likelihood principle. Although this sounds idyllic in principle, it is clear that all the problems in the single-species settings above will multiply in the Bormicon setting.

A further complication is that the data sources vary more widely than is currently common in fish stock assessments. As a result the negative log likelihoods are often not simply sums of squares but are of more variable form.

Reconciling and Consolidating Data Sources in a Multispecies Model

A sample Bormicon model with an objective function which consists of several negative log-likelihood components for cod and capelin is given in Björnsson et al. (1997). Although quite complex in detail, the components available in the objective function are simply descriptions of proportions at age from biological sampling (agedist), mean length at age (avelen), the adequacy of a multiplicative fit to age-disaggregated survey indices (ind), similarly for capelin acoustic survey, although without the multiplier (cap) and a descriptor of how much biomass is lacking in each area in order to fulfill the need for enough catches by month and area (understock).

Thus, in principle any type of datum can be placed into a component of the objective function and can thus become a part of the analysis. The design of Bormicon has intentionally included considerable flexibility in this regard and the intent is to make it as trivial as possible to include new data sources with likelihoods different from those previously considered, but in the following only the above components will be considered.

Each of the components needs to be given a weight after which they are added up and minimized over a number of unknown parameters, the most important parameters being the ones describing the annual recruitment of each and a few basic migration parameters.

The crucial issue is how the underlying model links the models for the various data. If growth is area-dependent, then the estimation of migration parameters may be seriously affected due to incorrect specification of the basic growth model. Careful analysis needs to be the basis for reconciling the various data sources so that they can be firm building blocks in a consolidating model.

Conflicting Data Sources in Multispecies Models—Model Error

As in the simple models, giving emphasis to different components of the objective function tends to push the solution in somewhat different directions. This can be illustrated by multiplying one of the components by a large number (100) and then estimating the parameters over again.

Figure 3 gives model-predicted mean length at age for a selected year class of cod when each component is emphasized in this way. It is seen that the trends in mean length at age differ considerably depending on which objective is used and in particular, the stomach content data indicates that cod are smaller than implied by other criteria.

Similarly Fig. 4 depicts the trends in cod biomass. Results based on emphasizing the age-length keys are omitted from this figure as they are far off the scale, indicating quite unreasonably high biomass levels, much higher than indicated from other sources.

Age readings are commonly used to obtain age-length keys which are then used with length distributions to provide catches in numbers at age which then drive single species assessments. It has often been observed that length distributions alone do not provide enough information to obtain the same accuracy in the assessment (ICES 1991). The present results indicate that one should not overemphasize the age-length keys.

Further, it is again seen that although the various sets give somewhat conflicting information, the stomach content data are considerably different from the other main data sets. Thus, the stomach content data appear to indicate much greater variation in the cod biomass than do the other data.

This is somewhat worrisome since stomach data plays a central role in other models such as MSVPA where they drive the analysis, and in MULTSPEC, where they form the core of the likelihood function. In MSVPA, however, the stomach content data do not drive the estimates of population abundance of the predator as in this case (and MULTSPEC) and thus this particular problem may not be an issue for MSVPA. In the present setting it would seem quite possible that Bormicon is vulnerable to poorly sampled stomach content data.

The interactions between these various assumptions can also be seen in Table 5. Each column indicates the value for the individual terms given the particular emphasized component. The table illustrates a number of inherent problems in the interaction between these data sets and the model under consideration.

Thus, the lowest SSE-value obtained for the indices is 12, but when indices are used with less weight the resulting SSE becomes 18. Although it is not obvious how many degrees of freedom should be associated with this sum of squares, it is clear that there is a fair number, since there is one term for each year, age, and area giving on the order of a hundred observations and the total number of parameters estimated is around 30. It therefore appears that in terms of indices, the final parameter estimates obtained are significantly worse than the “best” ones.

Although issues such as the intra-haul correlation (Pennington and Völstad 1989, Pennington 1991) invalidate a formal statistical analysis of the above difference, it is clear that there are indications that the present model is somewhat too “stiff,” i.e., cannot adequately twist and turn to accommodate the different data sets.

Of course the ultimate aim is to obtain a model which fits the various data sources and is “correct.” If the model is wrong in form, then the first issue at hand would be to modify the basic form before continuing. In the

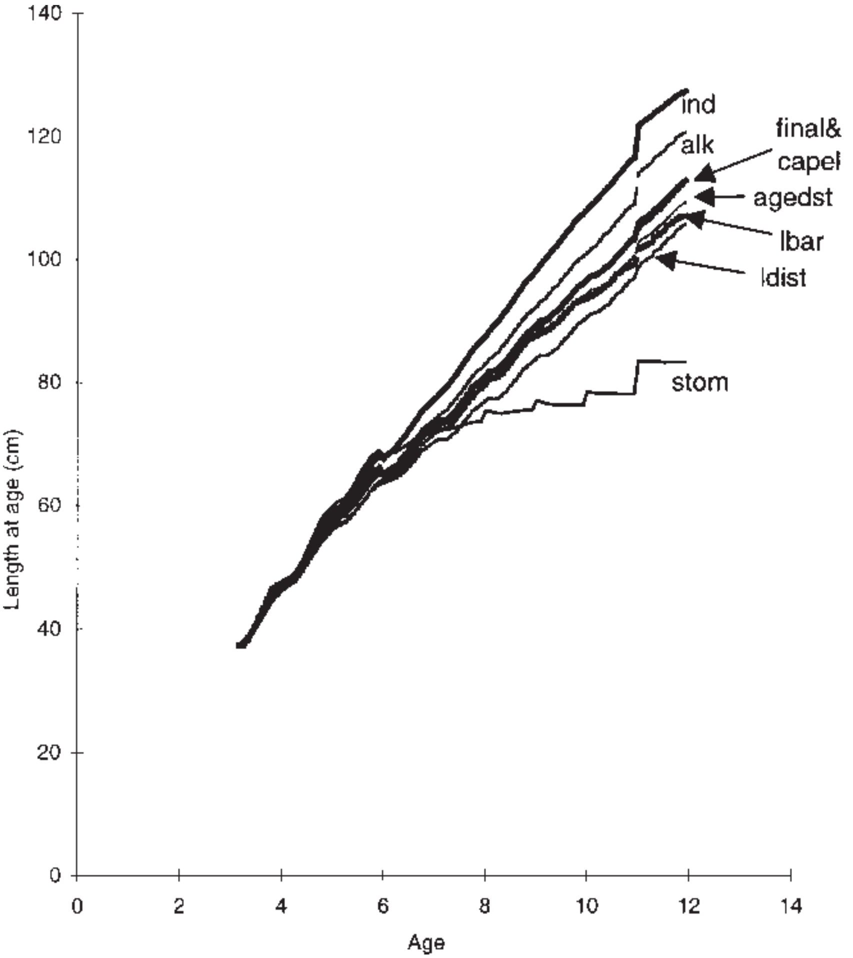


Figure 3. Mean length at age of 1984 year class based on emphasizing different components of the objective function during minimization. Survey indices (ind) and age-length keys (alk) result in highest values, whereas increased weight to stomach content data results in low values.

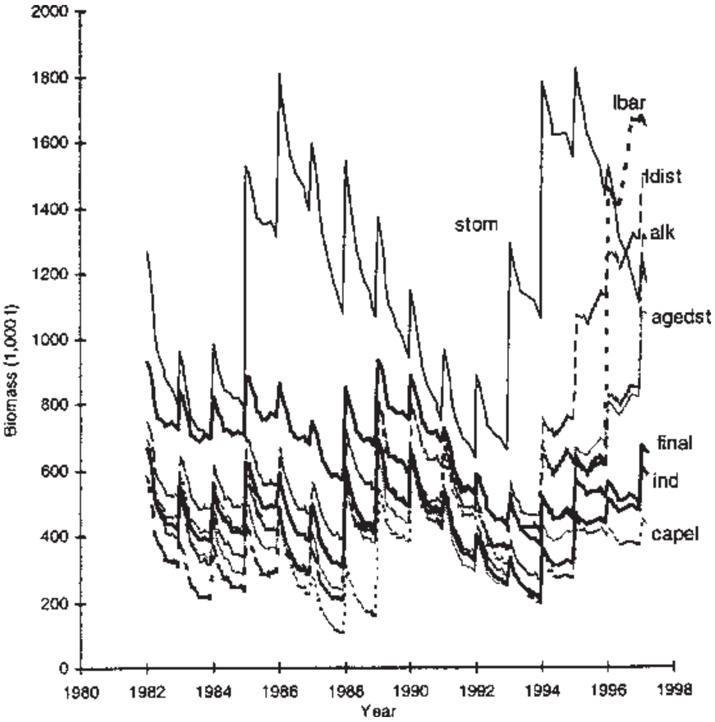


Figure 4. Trend in cod biomass based on emphasizing different components of the objective function during minimization.

present setting, however, it is quite possible that the primary problem is not with form but rather that the model is too parsimonious in terms of the number of parameters and there is probably room to estimate more parameters. In particular it is likely that a more flexible approach to migration modeling (e.g., by including interannual variation in migration) might be a useful addition to the model. Thus, the apparent lack of fit can be taken as a useful indication of required model extension. The precise extensions to be implemented need further study, but the migration parameters form an obvious choice for research since they affect the indices, predicted stomach content, growth, etc.

In addition to these model extensions, there is considerable room for diagnostics in this, like all other likelihood models. In particular, data exploration should be accompanied with mapping of the likelihood function for various parameters. This may well lead to some information on model reduction or sensitivity.

Table 5. Likelihood component values resulting from putting emphasis on different components (across) during minimization.

	data	agedst	alk	lbar	cap	ind	ldst	stom	rerun	neww
Negative log likelihood components										
agedst	44	41	49	57	75	55	50	84	46	14
alk	61	64	38	72	72	63	67	95	61	18
avelen	42	55	44	30	51	51	55	95	42	17
cap	88	99	94	137	67	99	144	200	116	23
ind	18	30	22	50	54	12	37	37	15	15
ldst	75	78	105	98	121	115	63	116	81	16
stom	81	80	79	108	93	80	105	53	80	16
underst	19	20	35	52	42	18	32	33	5	5
Percentage compositions of negative log likelihoods for each criterion										
agedst	10	9	10	9	13	11	9	12	10	11
alk	14	14	8	12	13	13	12	13	14	15
avelen	10	12	9	5	9	10	10	13	9	13
cap	20	21	20	23	12	20	26	28	26	19
ind	4	6	5	8	9	2	7	5	3	12
ldst	18	17	23	16	21	23	11	16	18	13
stom	19	17	17	18	16	16	19	7	18	13
underst	4	4	8	9	7	4	6	5	1	4
total	100	100	100	100	100	100	100	100	100	100
Status of each component in percent of minimum possible in each case										
agedst	108	100	119	140	182	134	121	204	111	33
alk	160	168	100	187	188	163	176	247	160	48
avelen	141	185	147	100	172	172	187	321	140	56
cap	130	147	140	203	100	147	214	297	172	34
ind	146	245	182	413	446	100	307	304	123	123
ldst	118	123	166	155	191	181	100	182	127	25
stom	154	152	150	205	177	152	200	100	152	30
underst	101	107	191	283	227	100	173	180	28	28

Each column represents a run with one emphasized component and a column entry gives the value for an individual term in that particular run. Base run values (Björnsson et al. 1997) are given in the data column. Values in the columns agedst, alk, lbar, cap, ind, ldst, and stom are negative log likelihood results from estimating model parameters while multiplying the corresponding component by 100. The weight is reset to the original value after estimation, in order to obtain comparable likelihood values in the table. Also given are values based on using parameters from reweighting (rerun and neww), assuming the minima obtained in previous columns to be unbiased variance estimates. The rerun column presents the likelihood components using the same weights as in the previous columns, but the neww column is based on the revised weights.

Presentation of Results and Diagnostics from a Spatially Disaggregated Multispecies Model

The use of spatially disaggregated multispecies models involves a number of large data tables and the output from these models often stretches into tens of megabytes. These data need to be analyzed for internal consistency and presented in a comprehensible manner. Concise graphical presentation is needed in order to illustrate in a lucid manner the essentials of the output but at the same time to avoid unnecessary detail.

One approach to such presentations is to use plots of the type given in Fig. 5. This figure consists of 5 sub-panels. Each sub-panel illustrates the spatial distribution of a single species, its catches, or its consumption of another species. One such plot is available per month for the period under consideration.

Several such plots can be joined together in an animation in order to illustrate how the different species migrate between the different regions. Such animations are not only a diagnostic tool, but also provide a useful way of presenting all of the important results from a complicated model such as Bormicon.

Discussion

Once the principle has been adopted to use several likelihood components in an overall likelihood function it becomes less important how data are analyzed prior to their inclusion in the stock assessment, as the assessment method itself can accommodate the raw data directly rather than only after a preliminary analysis. Thus, age-disaggregated survey indices are not really needed in order to “tune” an assessment. Rather, the otolith samples can be used for each area along with the total indices as two likelihood components. The resulting analysis becomes more comprehensive and more statistically valid than one where the prior analysis is largely hidden and involves extensively correlated use of age-length keys across many areas and even years.

On the other hand it is also abundantly clear that great care needs to be exercised when these data sources are combined since model misspecification of an individual data source may lead to that data source pulling the overall analysis in incorrect and unforeseen directions. Diagnostic methods such as testing different weightings to likelihood components are crucial for finding such tendencies in the model. Similarly, it is important to investigate likelihood profiles and in general investigate whether predicted abundance by area conforms more or less to that expected a priori. In some cases it may be possible to alleviate some apparent consistencies through increasing the number of parameters. In other cases one may want to change the form of model component, even to the extent of making certain relationships nonparametric, e.g., through using GAM models (Hastie and Tibshirani 1990).

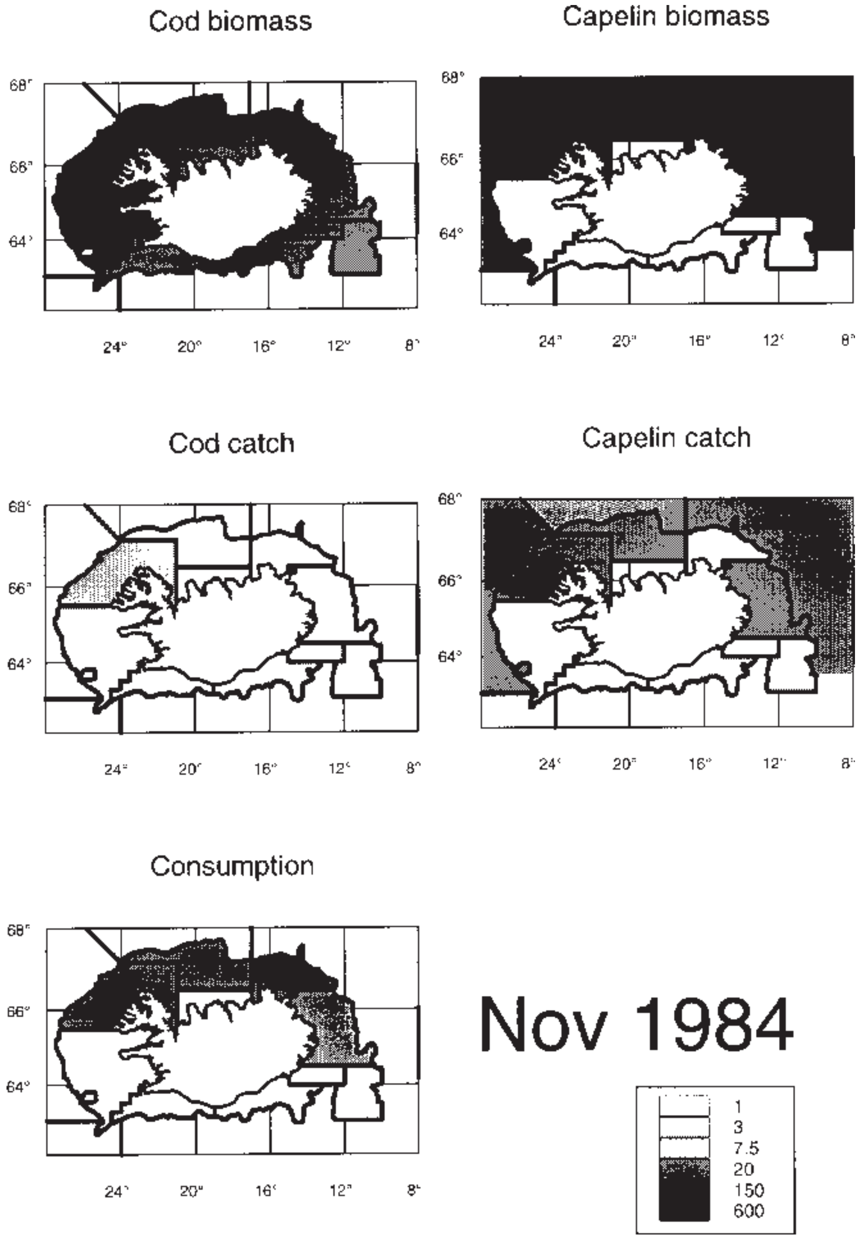


Figure 5. Snapshot output from Bormicon in November 1984. (a) Distribution of cod biomass across Bormicon regions, (b) distribution of capelin biomass, (c) cod catches, (d) capelin catches, (e) consumption of capelin by cod.

The likelihood principle, in the form of sums of squares and the least squares method of fitting, is the same principle as generally used in ADAPT (Gavaris 1988) and similar approaches (Halldórsson et al. 1986, Stefánsson 1988, de la Mare 1989). Further, dynamic stock-production models (Punt 1990, ICES 1995, Polacheck et al. 1993) often use the same principles. However, the present implementation includes an underlying multispecies, multifleet model with enough flexibility to incorporate any of these models.

One possibility not explored within the current implementation is to include a time-series approach, e.g., in a similar way as the method of Collie and Sissenwine (1983), further developed and implemented by Conser (1991). It is unlikely that a formal Kalman filter implementation as in Gudmundsson (1994) is feasible given the complexity of the model, but incorporating process error as in Collie and Sissenwine (1983) may well alleviate some of the "stiffness" in the usual parametric approaches.

Although the likelihood principle as presented in this paper will continue to be important well into the next century, another possibly fruitful way forward is to include Bayesian prior distributions (Hilborn and Walters 1992). This makes possible the incorporation of the uncertainty in various parameters and the inclusion of this uncertainty into future projections. In a huge model such as Bormicon the estimation procedure is not only time consuming but the total likelihood function may in fact have many maxima, and it is not clear whether the overall function is differentiable and thus the estimation of the variances and covariances between parameters is not trivial. When a Bayesian approach is used, some of these issues become less relevant, but of course other problems appear in their place, notably questions concerning the specification of priors and considerably increased computational complexity.

References

Beverton, R.J.H., and S.J. Holt. 1957. On the dynamics of exploited fish populations. Chapman and Hall. Reprint edn. 1993.

Björnsson, H., G. Stefánsson, H.N. Stefánsson, and H. Sigurgeirsson. 1997. Extended Bormicon examples. In: G. Stefánsson and Ó.K Pálsson, Bormicon: A Boreal migration and consumption model. Marine Research Institute Report 58, Reykjavík. 250 pp.

Bogstad, B., S. Tjelmeland, T. Tjelda, and Ö. Ulltang. 1992. Description of a multispecies model for the Barents Sea (MULTSPEC) and a study of its sensitivity to assumptions on food preferences and stock sizes of minke whales and harp seals. Int. Whaling Comm., Scientific Committee 44/O9.

Collie, J.S., and M.P. Sissenwine. 1983. Estimating population size from relative abundance data measured with error. Can. J. Fish. Aquat. Sci. 40:1871-1879.

Conser, R.J. 1991. A DeLury model for scallops incorporating length-based selectivity of the recruiting year-class to the survey gear and partial recruitment to the commercial fishery. Northeast Regional Stock Assessment Workshop Report, Woods Hole, MA. Res. Doc. SAW 12/2. Appendix to CRD-91-03. 18 pp.

de la Mare, W.K. 1989. The model used in the Hitter and Fitter program. Rep. Int. Whal. Comm. 39:150-151.

Deriso, R.B., T.J. Quinn II, and P.R. Neal. 1985. Catch-age analysis with auxiliary information. Can. J. Fish. Aquat. Sci. 42:815-824.

Gavaris, S. 1988. An adaptive framework for the estimation of population size. Can. Atl. Fish. Sci. Adv. Comm. (CAFSAC) Res. Doc. 88/29. 12 pp.

Gudmundsson, G. 1994. Time series analysis of catch-at-age observations. Appl. Statist. 43(1):117-126.

Gulland, J. 1965. Estimation of mortality rates. Annex to Arctic Fisheries Working Group report. Gadoid Fish Commission, ICES C.M. Doc. 3.

Halldórsson, Ó., P. Reynisson, and G. Stefánsson. 1986. A method for estimating terminal F's from a series of acoustic surveys. ICES C.M. 1986/H:62

Hastie, T.J., and R.J. Tibshirani. 1990. Generalized additive models. Chapman and Hall, New York. 335 pp.

Helgason, Th., and H. Gislason. 1979. VPA-analysis with species interaction due to predation. ICES C.M.1979/G:52.

Hilborn, R., and C.J. Walters. 1992. Quantitative fisheries stock assessment: Choice, dynamics and uncertainty. Chapman and Hall, New York. 570 pp.

ICES. 1984. Report of the Working Group on Methods of Fish Stock Assessments. ICES Coop. Res. Rep. 129.

ICES. 1991. Report of the Working Group on Methods of Fish Stock Assessments. ICES C.M. 1991/Assess:25.

ICES. 1995. Report of the Working Group on Methods of Fish Stock Assessments. ICES Coop. Res. Rep. 199.

Magnússon, K.G., and Ó.K. Pálsson. 1989. Trophic ecological relationships of Icelandic cod. Rapp. P.-V. Réun. Cons. Int. Explor. Mer 188:206-224.

Magnússon, K.G., and Ó.K. Pálsson. 1991. The predatory impact of cod on shrimps in Icelandic waters. ICES C.M. 1991/K:31

Methot, R.D. 1989. Synthetic estimates of historical abundance and mortality for. northern anchovy. Amer. Fish. Soc. Symp. 6:66-82.

Pálsson, Ó.K. 1983. The feeding habits of demersal fish species in Icelandic waters. Rit Fiskideildar 7(1):1-60.

Pennington, M. 1991. Assessing the effect of intra-haul correlation and variable density on population estimates from marine surveys. ICES C.M.1991/D:14.

Pennington, M., and J.H. Volstad. 1989. Optimum size of sampling unit for estimating the density of marine populations. ICES C.M.1989/D:22, Ref. K.

Pope, J.G. 1993. A reappraisal of the 1992 assessment of the Icelandic cod stock. MAFF [Ministry of Agriculture, Food and Fisheries], Fisheries Laboratory, Lowestoft.

Polacheck, T., R. Hilborn, and A.E. Punt. 1993. Fitting surplus production models: Comparing methods and measuring uncertainty. Can. J. Fish. Aquat. Sci. 50:2597-2607.

Punt, A.E. 1990. Is $B_1 = K$ an appropriate assumption when applying an observation error production-model estimator to catch-effort data ? S. Afr. J. Mar. Sci. 9:249-259.

Scheffé, H. 1959. The analysis of variance. John Wiley and Sons, Inc., New York.

Shepherd, JG., and M.D. Nicholson. 1991. Multiplicative modeling of catch-at-age data, and its application to catch forecasts. J. Cons. Cons. Int. Explor. Mer 47:284-294.

Stefánsson, G. 1988. A statistical analysis of Icelandic trawler reports, 1973-1987. ICES C.M.1988/D:13, Stat. Committee/Ref. G.

Stefánsson, G., and Ó.K. Pálsson (eds.). 1997. Bormicon: A boreal migration and consumption model. Marine Research Institute, Tech. Report 58, Reykjavik.

Stefánsson, G., and Ó.K. Pálsson. 1998. A framework for multispecies modeling of boreal systems. Rev. Fish Biol. Fish. 8:101-104.

Stefánsson, H.N., H. Sigurgeirsson, and H. Björnsson. 1997a. Bormicon: User's manual. Marine Research Institute, Report 59, Reykjavik.

Stefánsson, H.N., H. Sigurgeirsson, and H. Björnsson. 1997b. Bormicon: Programmer's manual. Marine Research Institute, Report 60, Reykjavik.

Tjelmeland, S., and B. Bogstad. 1989. Multspec: The manual. Sponsored by the Norwegian Fisheries Research Council, Bergen.

Rapid Appraisal of the Status of Fisheries for Small Pelagics Using Multivariate, Multidisciplinary Ordination

Tony Pitcher, Steven Mackinson, and Marcelo Vasconcellos
University of British Columbia, Vancouver, B.C., Canada

Leif Nøttestad
University of Bergen, Bergen, Norway

David Preikshot
University of British Columbia, Vancouver, B.C., Canada

Abstract

This paper explores a new method for rapidly evaluating the relative status of fisheries using information from multiple sources in the ecological, technological, economic and social fields. Within each discipline, each member of a set of fishery attributes is scored using published sources or information from experts. First, using multivariate data description, fisheries are ordinated within the four disciplinary areas using Multidimensional Scaling. Second, a single interdisciplinary ordination is performed by using the ordination scores from each discipline together. Reference points for the diagnoses are provided by constructing hypothetical fisheries that are assigned "good" or "bad" scores, defined in terms of sustainability for each discipline, or that have a random assignment of attribute values.

Evaluations of status are made among 29 fisheries for sardine, Atlantic herring, Pacific herring, and anchovy in a wide range of upwelling and coastal systems, including time series for three major herring fisheries. In general, Pacific herring fisheries ordinate in "better" positions than Atlantic herring fisheries; Western Atlantic better than Eastern Atlantic, Alaskan better than British Columbia for Pacific herring, sardines and anchovies better than herring fisheries. Fixed gears, bycatch reduction devices, and socially integrated fishing communities increase evaluated status. For most

individual fisheries and for the trajectories from the herring fishery time series, the ordinations provide evaluations of relative status that is in concordance with what we know from historical documentation and conventional assessments.

Although results are encouraging, the technique needs further refinement. It is not intended to replace conventional stock assessments, but it shows promise for rapid appraisal and monitoring, triage in the face of scarce management resources, and a more objective way of conflating diagnoses from a range of disciplines.

Introduction

This paper contributes to the development of a novel technique (Pitcher et al. 1998) aimed at providing a rapid, cost-effective appraisal of the status of a fishery. Multivariate ordinations of scored attributes evaluate fishery status separately within ecological, technological, economic and social areas, and then provide a combined, multidisciplinary evaluation. Ordination may include a set of individual fisheries, or the trajectory in time of a single fishery. Status is assessed relative to the best and worst possible fisheries that may be constructed using a given set of attributes.

The aim of this paper is an evaluation of fisheries for small pelagic zooplanktivorous shoaling fish of the clupeid family, such as herring, anchovies and sardines. This group of species constitute over 40% of the world fish catch (Garcia and Newton 1994), and their fisheries exhibit considerable changes in status over time, including dramatic and economically damaging stock collapses. Small pelagic fish perform extensive migrations driven by behavioral tradeoffs between food, predators and reproduction that optimize feeding, spawning, and overwintering during the life history (e.g., herring: Fernö et al. 1998) and these traits render them vulnerable to environmental fluctuation and change. Characteristically, they also exhibit great volatility of recruitment (Bailey 1992, Zheng 1996), probably as a hedge against ocean changes affecting larval retention areas (Gagne et al. 1991, Bakun 1996). In fisheries, the catchability of small clupeids may be increased by social aggregation into schools (Pitcher and Parrish 1993, Pitcher 1995), and by a rapid reduction in geographical range ("range collapse") in the face of environmental change and fishing. These factors can engender a dramatic fishery collapse (Pitcher 1997, Mackinson et al. 1997a). Many of these small clupeids have pivotal trophic roles in their ecosystems (Mackinson et al. 1997b), so that collapse can have consequences for other fishery sectors. Moreover, in some cases, particularly for herring fisheries, there are long historical databases going back into the past century that may be used for long-term analysis (e.g., Cushing 1988). In addition, small pelagic fish are harvested with a wide range of gear, such as pair trawls, purse seines, driftnets, gillnets, and traps. All of these features challenge the new method, but help by providing considerable contrast in the data.

Methods

Attributes Used in the Analysis

Our method ordinates fisheries in four disciplinary areas that are critical to long term viability of a fishery:

1. Ecological (including fish population parameters and environment).
2. Technological (including gear and fishing characteristics).
3. Economic (including both micro- and macroeconomic factors).
4. Social (including social and anthropological factors).

Within each ordination, a set of 9-10 attributes have been defined. Attribute numbers are designed to maximize discriminating power in the ordination technique, where a rule of thumb is to have three times as many fisheries as attributes used to ordinate them (Stalans 1995). Criteria for choosing attributes are that they were easily and objectively scored, that extreme values are easily ascribed to “good” and “bad” in relation to sustainability, and that scores are available for all the fisheries and time periods in the analysis.

Scores for each fishery were determined from the literature or from correspondence with experts. Some values for economic and social areas were obtained from the CIA World factbook (CIA 1995). Full details of the attributes are given in Table 1; most are scored on a 3- or 4-point ranked scale that makes it relatively easy both to obtain a score in the absence of precise surveys and interviews, and for a group of experts to agree. Moreover, ranked scores can easily be refined later by inserting precise values without disrupting the main features of an ordination.

Fisheries Used in the Analysis

Table 2 lists sources of information for the 29 individual fisheries that we have used. These comprised seven fisheries for Atlantic herring, nine for Pacific herring, six sardine fisheries and seven anchovy fisheries. Three herring fisheries have been followed through time from 1950 to the present day; the Norwegian spring spawning herring, the North Sea herring, and the British Columbia herring.

In order to provide fixed reference points for status, two hypothetical fisheries were simulated by choosing “good” or “bad” scores as the extremes for each attribute in each discipline, where good and bad were evaluated in terms of sustainability of the fishery. In addition, to show if status evaluations were meaningful, twenty random sets of attribute scores (“random” fisheries) were simulated for each discipline.

Ordination Method

We have used non-parametric multidimensional scaling (MDS) (Kruskal and Wish 1978, Schiffman et al. 1981, Stalans 1995), an ordination technique that can produce unbiased distance maps of relative location

Table 1. List and definitions of attributes used in the analysis, in four disciplinary areas, and showing the "good" and "bad" scores.

Attribute	Scoring	Good	Bad	Notes
ECOLOGICAL				
1 Catch/fisher	Metric tons	Lo	Hi	Metric tons per fisher per year
2 Exploitation status	0; 1; 2; 3	0	3	FAO-like scale; under, fully, heavily; overexploited to collapse
3 Recruitment variability	0; 1; 2	0	2	Low recruitment COV <40%; medium COV 40-100%; high COV >100%
4 Trophic level	number	Hi	Lo	Average trophic level of species in catch
5 Migratory range	0; 1; 2	0	2	1-2, 3-4, >4 jurisdictions encountered during migration
6 Catch < maturity	0; 1; 2	0	2	None; some (>30%); lots (>60%) caught before maturity
7 Discarded bycatch	0; 1; 2	0	2	Low 0-10%; med 10-40%; hi >40% of target catch
8 Species caught	0; 1; 2	0	2	Low 1-10; med 10-100; hi >100 species
9 Primary production	0; 1; 2; 3	3	0	$gCm^{-2}y^{-1}$; low=0-50 ; med=50-90; high=90-160; very high=160+
ECONOMIC				
1 Price	US$/t	Hi	Lo	US$/t of landed product for time of data point
2 Fisheries in GNP	0; 1; 2	2	0	Importance of fisheries sector in country: low; medium ; high
3 GNP/person	US$/capita	Hi	Lo	In country of fishery
4 Limited entry	0; 1; 2	2	0	Almost none; some; most (includes informal limitation)
5 Marketable right	0; 1; 2	2	0	Marketable right/quota/share right: none; some; full ITQ
6 Other income	0; 1; 2	0	2	Fishing mainly casual work; part time; full time
7 Sector employment	0; 1; 2	0	2	<10%; 10-20%; >20% employment in formal fishery sector
8 Ownership	0; 1; 2	0	2	Profit from fishery mainly to locals; mixed; foreigners
9 Market	0; 1; 2	0	2	Principally local/national; national/regional; global/international

Table 1. (Continued.)

Attribute	Scoring	Good	Bad	Notes
SOCIOLOGICAL				
1 Socialization of fishing	0; 1; 2	2	0	Fishers work as individuals, families, or community groups
2 Fishing comm. Growth	0; 1; 2	0	2	Growth over 10 years pre-data point: <10%; 10%-20%; >20%
3 Fisher sector	0; 1; 2	0	2	Households in fishing in the community: $<{}^{1}/_{3}$; ${}^{1}/_{3}$-${}^{2}/_{3}$; $>{}^{2}/_{3}$
4 Education level	0; 1; 2	2	0	Below; same; above population average
5 Conflict status	0; 1; 2	0	2	Level of conflict with other sectors
6 Information sharing	0; 1; 2	2	0	None; some; lots
7 Fisher influence	0; 1; 2	2	0	Strength of fisher direct influence on actual fishery regulations
8 Fishing income	0; 1; 2	2	0	Fishing income % of total family income: <50%; 50-80%; >80%
9 Kin participation	0; 1	1	0	Do kin sell family catch and/or process fish: no (0) or yes (1)
TECHNOLOGICAL				
1 Trip length	Days	Lo	Hi	Average days at sea per fishing trip
2 Landing sites	0; 1; 2	0	2	Landing sites dispersed; some centralization; heavily centralized
3 Pre-sale processing	0; 1; 2	2	0	None; some; lots (e.g. gutting, filleting) before sale
4 Use of ice	0; 1; 2	2	0	None; some ice; sophisticated (flash freeze, champagne ice)
5 Gear	0; 1	0	1	Passive gear = 0; active gear = 1
6 Selective gear	0; 1; 2	2	0	Device(s) in gear to increase selectivity: few; some; lots
7 Power gear	0; 1	0	1	No power assistance to gear = 0; power assisted gear = 1
8 FADS	0; 1	0	1	Fish Aggregation Devices (= FADs): not used = 0; used = 1
9 Sonar	0; 1	0	1	Sonar (acoustics) used to aid catch; no = 0; yes = 1
10 Vessel size	0; 1; 2	0	2	Average length of vessels <8m (0); 8–17m (1); >17m (2)

Lo = minimum from data, Hi = maximum from data.

Table 2. The herring, sardine and anchovy fisheries included in the analysis showing the symbols used in the plots and sources of information.

Species	Fishery location	Symbol on plots	Period	References/personal communications
South African sardine	South Africa	SA sar	Recent	T. Akkers,[a] R. Crawford,[a] K. De Swardt,[a] J. Van Zyl,[a] J. Van der Weshuizen[a]
Adriatic sardine	North Adriatic, pair trawl (volante)	Adsar vol	1989	Pitcher et al. 1992
Adriatic sardine	Central Adriatic, purse +l ight (lampara)	Adsar lam	1989	Pitcher et al. 1992
Brazilian sardine	Southeastern Bight, Brazil	Bra sar	Average 1985-1990	Cergole 1993, Diegues 1995, Valentini & Cardoso 1991, Rossi-Wongtschowski et al. 1995, Matsuura 1981,1990
Mexican sardine	Sea of Cortez	MexCZ sar	1995	Y. Harada[c]
Japanese sardine	Sea of Japan	Jap sar	1995	Y. Harada[c]
Peruvian anchovy	Peru	Peranc69, Peranc75	1969 1975	D. Pauly[h]
Adriatic anchovy	North Adriatic, pair trawl (volante)	Adanc vol	1989	Pitcher et al. 1992
Adriatic anchovy	North Adriatic, purse+light (lampara)	Adanc lamp	1989	Pitcher et al. 1992
South African anchovy	South Africa	SA anc	pecent	K. De Swardt,[a] J. Van Zyl,[a] D.W. Schutte,[b] R.J. Crawford,[a] J. Van der Weshuizen[a]
Japanese anchovy	Sea of Japan	Jap anc	1995	Y. Harada[c]
Japanese anchovy	Seto Island Sea, Japan, purse seine	Jap SRN anc,	1995	Y. Harada[c]
	Seto Island Sea, Japan, fixed trap	Jap SET anc	1995	Y. Harada[c]
Atlantic herring	North Sea	NS50-NS97	1950-present	Whitmarsh et al. 1995, Saville & Bailey 1980, Coull 1988, Burd 1974
	Norway-Norwegian Spring Spawning	Nor50-Nor96	1950-present	L. Nøttestad,[l] F. Asche,[m] R.E. Svarstad,[n] P. Sandberg,[o] Myrstad 1996
	Icelandic-Norwegian Spring Spawning	ICss	Present	H. porValtysson[h]
	Icelandic Summer Spawning	ICsum	Present	H. porValtysson[h]
	Baltic Sea	Balt	Present	F. Arrhenius[i]

Table 2. (Continued.)

Species	Fishery location	Symbol on plots	Period	References/personal communications
Atlantic herring (continued)				
	Bay of Fundy, seine	BFunS,	Present	R. Stephenson[j]
	Bay of Fundy, weir	BFunW	Present	R. Stephenson[j]
Pacific herring	British Columbia, Canada	BC50-BC96	1950-Present	D. Hay,[g] S. Mackinson[h]
	Southcentral Alaska	ScAk	Present	T. Otis[e]
	Bering Sea; Togiak; Bristol Bay, Alaska	Tog	Present	K. Rowell,[f] Taylor 1964,1985
	Security Cove, Alaska	Sec	Present	T. Cappiello[k]
	Nunivak Island, Alaska	Nun	Present	T. Cappiello[k]
	Nelson Island, Alaska	Nel	Present	T. Cappiello[k]
	Good News Bay, Alaska	News	Present	T. Cappiello[k]
	Cape Avinof, Alaska	Avin	Present	T. Cappiello[k]
	Prince William Sound, Alaska	PWS	Present	E. Brown[d]

References are listed in the reference list.

Addresses for personal communications:

[a] Sea Fisheries Research Institute, Private Bag X2, Rogge Bay 8012, Cape Town, South Africa.

[b] Human Sources Research Council, Western Cape Regional Office, Private Bag X5, Rogge Bay 8012, South Africa.

[c] Laboratory of Fish Population Dynamics, Faculty of Bioresources, Mie University, 1515 Kamihama, Tsu 514, Japan.

[d] University of Alaska Fairbanks, SFOS IMS, Box 757220, Fairbanks, AK 99775-7220.

[e] LCI Research Biologist, Alaska Dept. of Fish and Game, 3298 Douglas Pl., Homer, AK 99603.

[f] Alaska Dept. of Fish and Game, 333 Raspberry Road, Anchorage, AK 99518.

[g] Pacific Biological Station, Department of Fisheries and Oceans, 3190 Hammond Bay Road, Nanaimo, BC, Canada V9R 5K6.

[h] Fisheries Centre, University of British Columbia, Vancouver, BC, Canada.

[i] Cornell University Biological Field Station, 990 Shakelton Point Road, Bridgeport, NY 13030.

[j] Department of Fisheries and Oceans, St. Andrews, New Brunswick, Canada.

[k] Alaska Dept. of Fish and Game, P.O. Box 1467, Bethel, AK 99559.

[l] Department of Fisheries and Marine Biology, University of Bergen, Bergen, Norway.

[m] Centre for Social and Economic Research (SNR), NHH, Bergen, Norway.

[n] Fisheries Advisory Services in the County of Sogn and Fjordane, Norway.

[o] The Norwegian Directorate of Fisheries, Bergen, Norway.

(Clarke 1993) which may be rotated and shifted linearly with minimal disruption (Clarke and Warwick 1997). We employed a squared Euclidean distance matrix with attribute scores normalized using *Z*-values. MDS in two dimensions was implemented for our 44 fishery points, the "good," "bad," and 20 "random" fisheries using the SPSS statistical package (SPSS 1996) and the PRIMER package (Carr 1997). Fishery scores on the two MDS axes were plotted against each other for each of the four disciplinary analyses and goodness-of-fit evaluated using stress values (values below 0.25 are considered acceptable by Clarke and Warwick 1997). MDS scores for the simulated "good" and "bad" fisheries, and the average and 95% confidence limits from the "random" fisheries, were included on the plot. The two MDS scores from each analysis, making eight scores in all, were used as input data for the final interdisciplinary ordination. After ordination, we adopted a convention to rotate plots (to a least squares criterion) so that "good" appeared at top left (azimuth 315°) and "bad" at lower right (azimuth 135°); in all cases these two points fell very close to a straight line through the plot origin. The simulated random fisheries, which were in each case approximately normally distributed with SD ≈ 0.4, have been represented by their means and 95% confidence limits as a rotated cross on the plot.

To examine which attributes most influenced an ordination, we again rotated the plots using least squares until "good" and "bad" lay at 90 and 270°. We then took the *X*-axis as the dependent variable in a multiple regression with the normalized attributes as the independent variables.

To check that the method would ordinate fisheries monotonically, we simulated some fisheries whose status moved in single steps from "bad" to "good," scored on 10 ordinal attributes from 0 to 4. The resulting ordinations of two such simulated fisheries (Fig. 1a) are encouragingly monotonic. Figure 1b shows an ordination of a fishery exhibiting periodic larger steps in status (3 steps), which are reasonably linearly preserved relative to the reference fishery, although movement at the edges occupies more space than at the center. As might be expected, in both cases the "random" fisheries lie close to center of the plot, and this justifies our re-centering the fishery plots to the zero from the "random" fisheries. In answer to a referee, in a real ordination, an excess of points to one side or the other of this zero does not represent analytical bias, but a greater or lesser number of fisheries above or below the median status score that lies in the middle of our fixed scale from "good" to "bad."

Results

MDS ordinations are presented for each discipline and the combined analysis in Figs. 2 to 6, stress values are given in the figure captions, and symbols used for fisheries are listed in Table 2 and in Fig. 2.

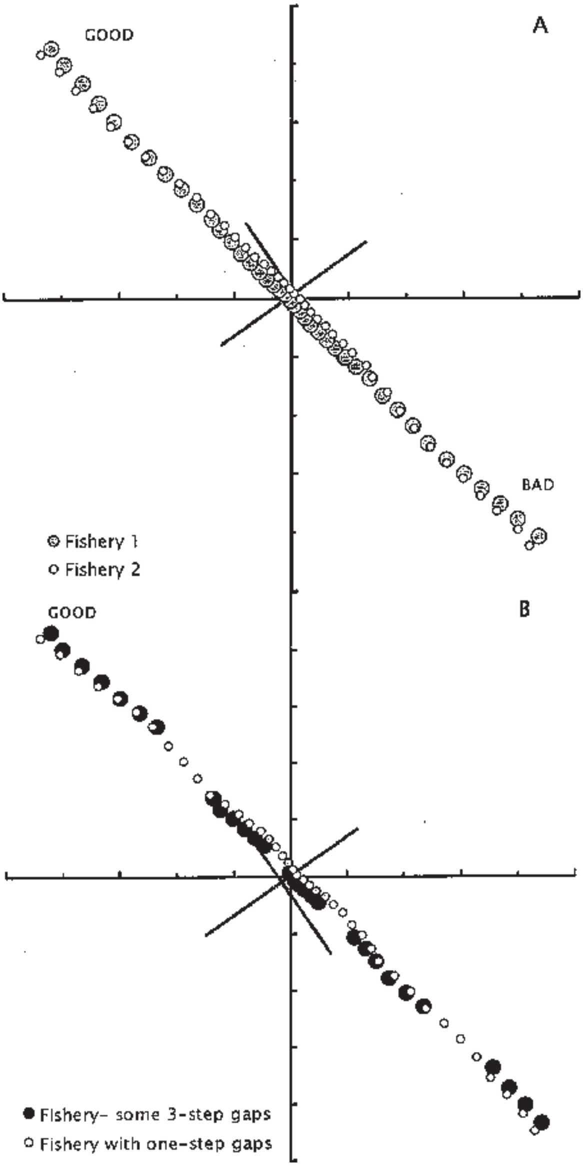


Figure 1. A. MDS ordination of simulated fisheries moving in single steps from "bad" status (lowest possible scores on ten attributes) to "good" (highest possible scores on each of ten attributes). Many alternative routes are possible: two fisheries are shown together with fisheries constructed from random scores (central cross). B. As Fig. 1A, but with fishery 2 showing periodic 3-step improvements in status.

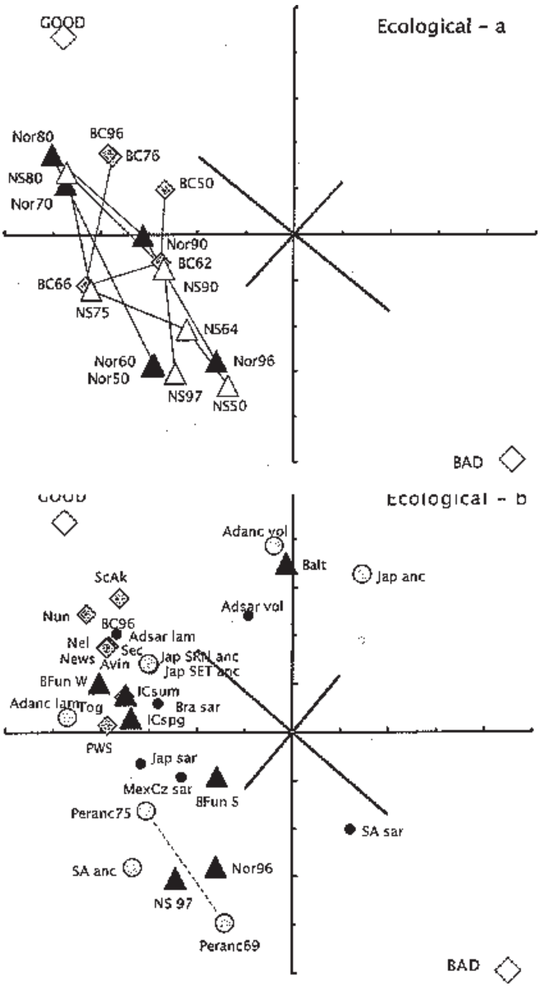


Figure 2. *Multidimensional scaling ordination (MDS) of the small pelagic fisheries using scores of ecological attributes in Table 1. Axes 1 and 2, ticked in standard deviations, represent the rotated axes as described in the text. "Good" and "bad" locations were constructed using extremes scores for each attribute. Cross indicates the 95% confidence limits on the original ordination axes, for 20 hypothetical fisheries constructed from random scores; large triangles = Atlantic herring; diamonds = Pacific herring; small circles = sardines; large circles = anchovy. Codes for individual fisheries are listed in Table 2. Stress value for this ordination = 0.25. 2a.Trajectories of ecological status for three herring fisheries through time. 2b. Ordinations of individual herring, sardine, and anchovy fisheries. The most recent points for three fisheries in Fig. 2a are shown.*

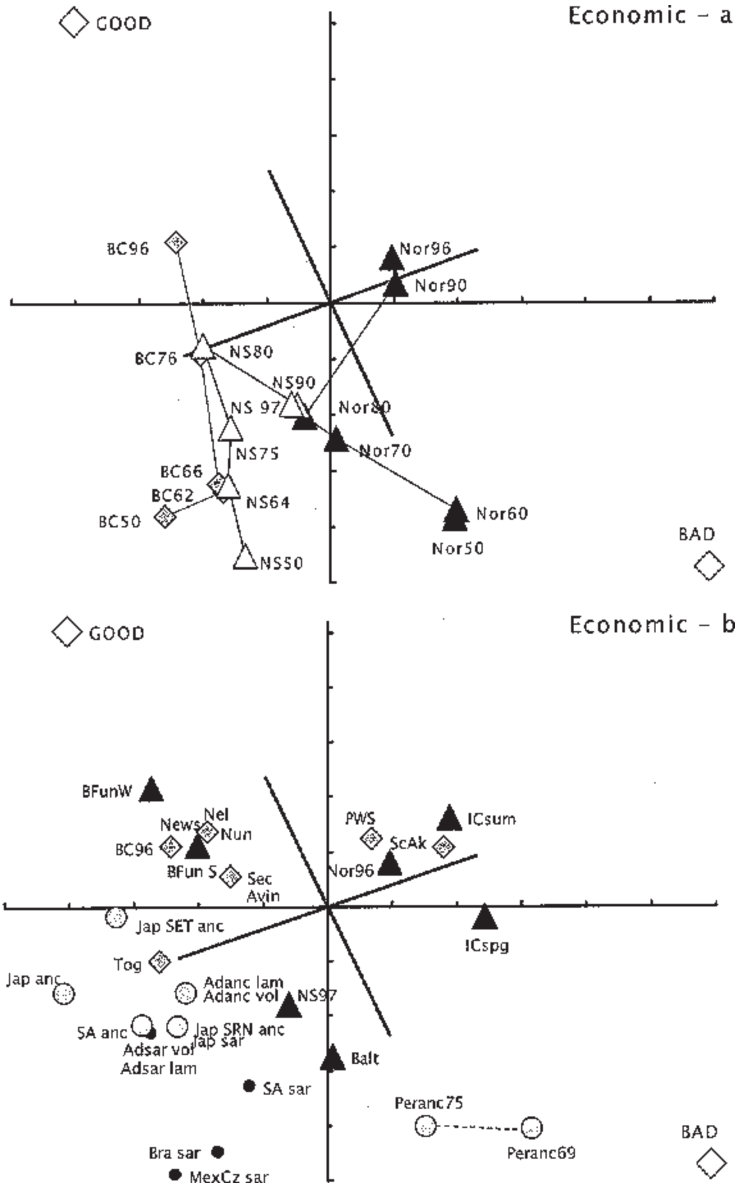


Figure 3. As Fig. 1, for MDS ordination using economic attributes. Stress value = 0.24.

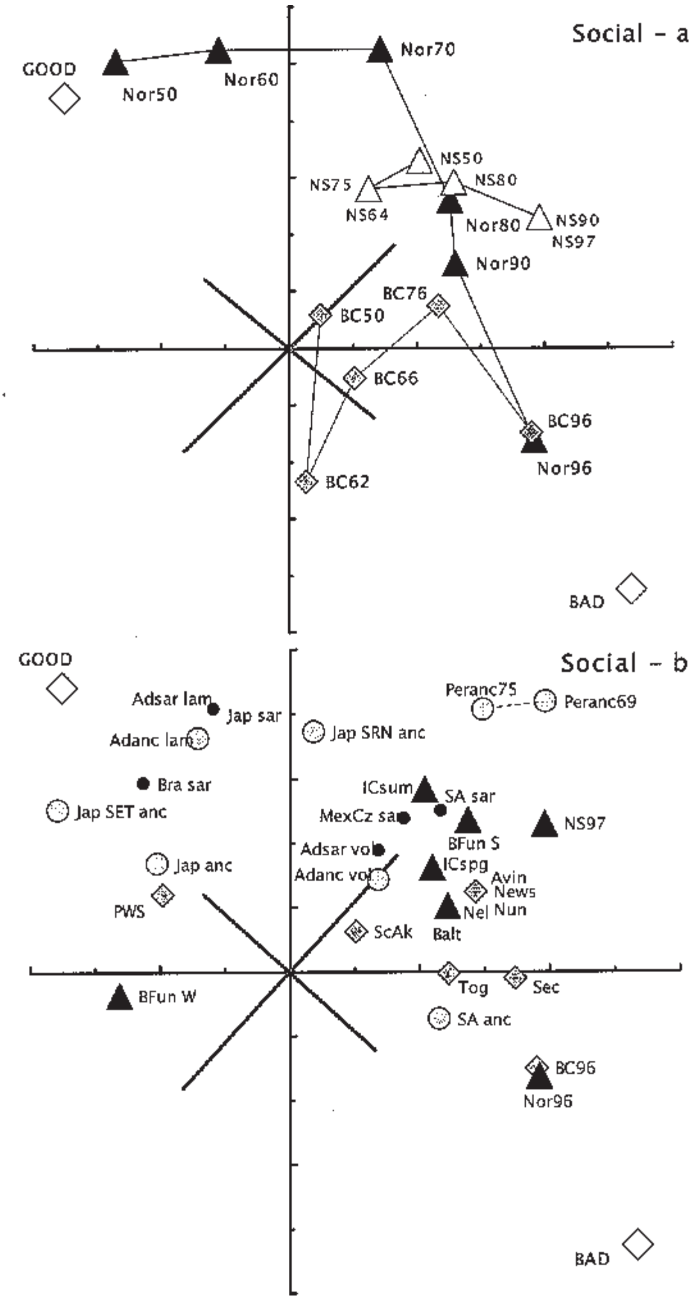


Figure 4. As Fig. 1, for MDS ordination using social attributes. Stress value = 0.28, somewhat high for a reliable ordination.

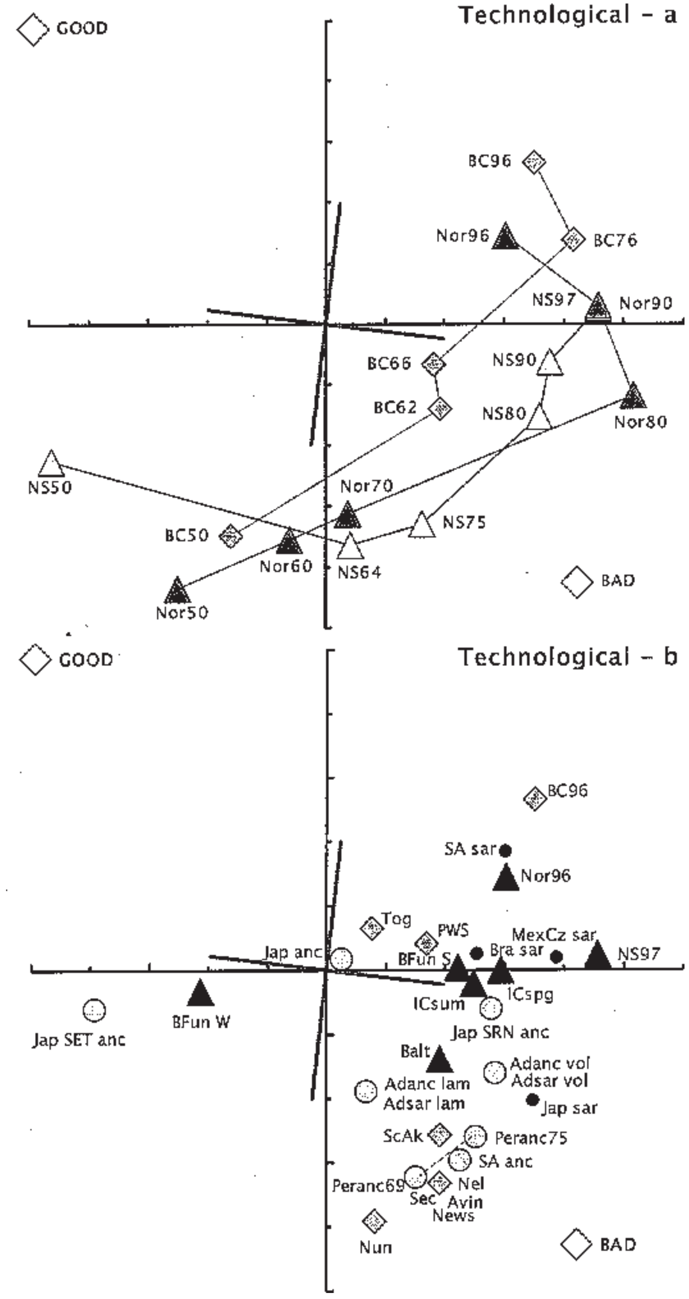


Figure 5. As Fig. 1, for MDS ordination using technological attributes. Stress value = 0.23.

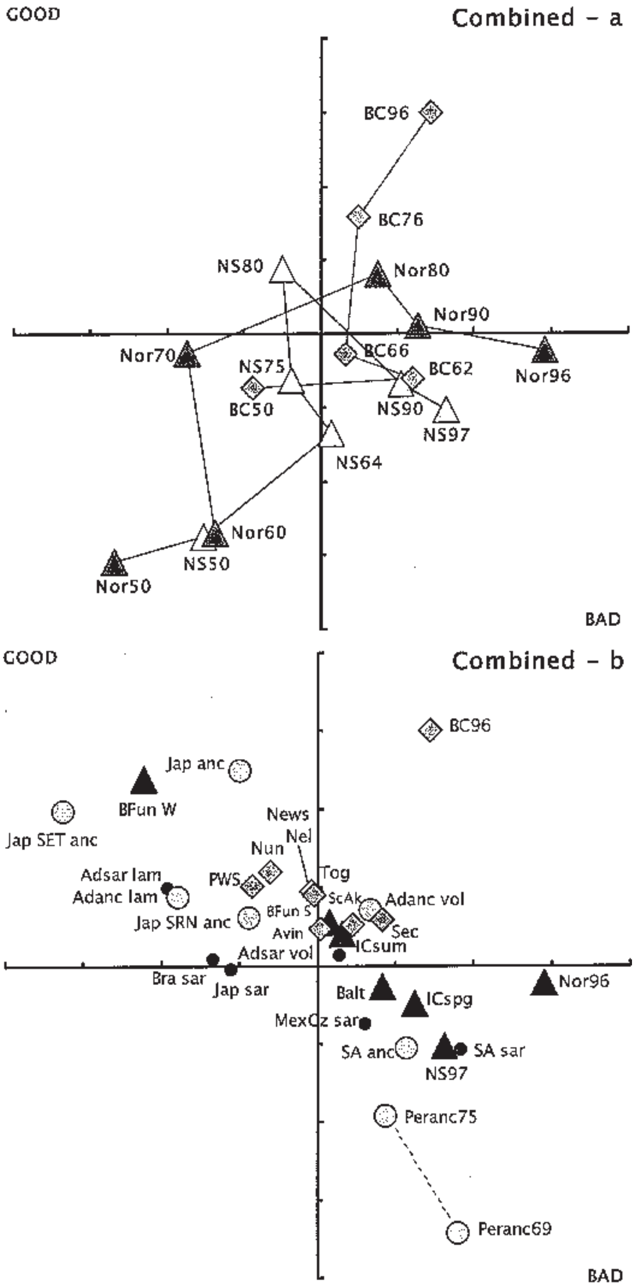


Figure 6. Combined interdisciplinary MDS ordination of fisheries using ordination scores on axes 1 and 2 from the four disciplinary analyses shown in Figs. 2 to 5. Stress value = 0.23.

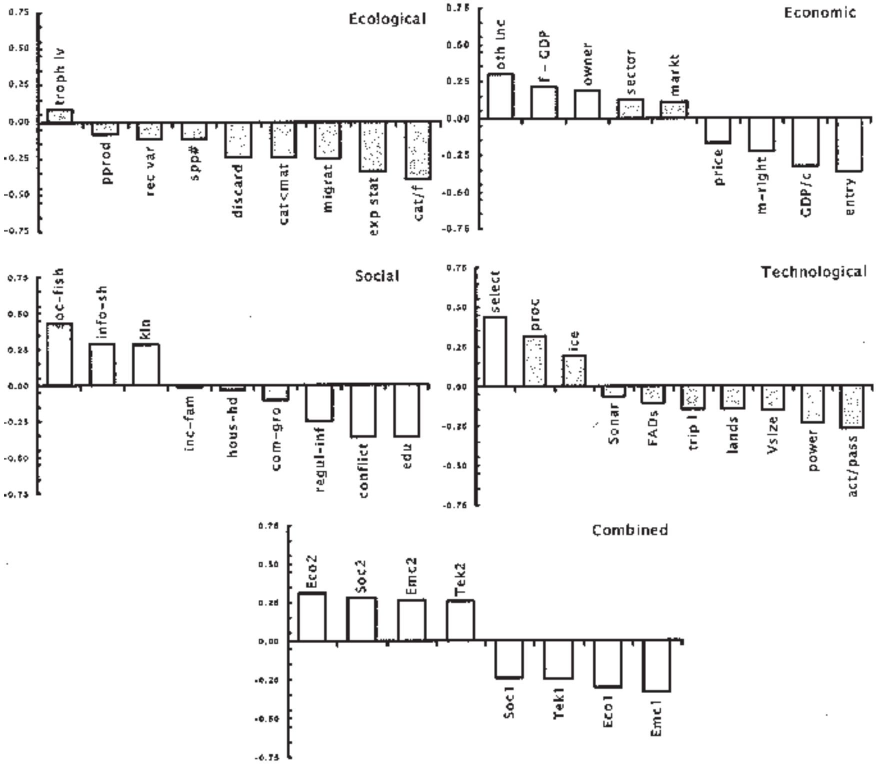


Figure 7. Multiple regression coefficients for normalized attributes in each discipline against the rotated "good" to "bad" axis. Shaded bars indicate non-significant coefficients. Multiple R squares were: Ecological = 0.16; Economic = 0.88; Social = 0.91; Technological = 0.28; Combined = 0.96.

Figure 7 shows the multiple regression coefficients for the attributes on the rotated "good" to "bad" axis. For the ecological analysis, catch per fisher, migratory range, exploitation status, discards, and immature catch all play a large but (non-significant) negative role in the separation of points. Seven of the nine economic attributes are significant in determining changes, fishers having other income, importance of fisheries in the country, and local ownership are positive factors, while limited entry, marketable rights, GDP per capita and price are, perhaps surprisingly for economists, negative. Devices to improve the selectivity in the gear have the largest significant positive influence on the technological ordination. Information sharing, socialization of fishing, and kinship are the largest significant positive factors in the social ordination, while negative factors are education level, conflict status, and, surprisingly, fisher influence on regulations. For the combined ordination, all eight axes are almost equally important, the first of each axis pair positively, and the second negatively related to the ordination, as should be expected from the original rotation that placed "good" at 315 and "bad" diagonally at 135° azimuth.

Time Series: British Columbia Herring

Like the North Sea, the B.C. herring fishery has experienced radical changes since the 1950s. Traditionally, herring were taken for reduction to meal in a fishery that intercepted fish on pre-spawning migrations from offshore feeding grounds in fall and winter (Outram and Humphreys 1974), and this fishery reached the point of collapse in 1967. However, since the early 1970s the fishery has changed entirely to a sac-roe product. The 1976 and 1996 points (roe fishery years) are clearly separated from the earlier years by a movement toward "good" quadrant on the ecological ordination (Fig. 2), a pattern that is repeated in all of the other disciplines.

Since the start of the sac-roe fishery, herring has become the highest landed value species in B.C. In 1991, on average, herring contributed about Can$220,000 to the gross income of a herring boat, while other species, mainly salmon, accounted for Can$146,500 per vessel (SOE 1994). Hence the economic ordination trajectory (Fig. 3) has headed rapidly toward "good."

As with the North Sea fishery, the social trajectory for B.C. herring (Fig. 4) does not provide obvious indicators as to its change of status measured in this discipline. With the exception of the most recent 1996 point which has clearly moved toward "bad," the B.C. fishery ordinates in the same area as fisheries with randomly chosen attribute scores.

During the development of the reduction fishery from 1950s to 1967 a number of changes occurred that help to explain the technological trajectory (Fig. 5). Radio-telephones were widely used by 1950 (Hourston and Haegle 1980) and during that decade the power block and sonar were introduced. Catches rose to 190,000 t in 1965, but the fishermen had more difficulty in locating quantities of fish (Hourston 1980). Drum seining became popular and the use of arc lights enabled fishermen to capture the more thinly distributed dwindling stocks, together with other species attracted by the light. The proportion of immature fish in the catch rose significantly, and despite the prohibition of lights in 1966, the stock collapsed in 1967 (Hourston 1980, Pearse 1982). Since arc lights are not used in the sac-roe fishery and net meshes are more stringently enforced than in the reduction era, these provide a "good" influence on the technological trajectory, and we see the 1976 and especially the 1996 points moving leftward. It is noteworthy that technological change in the Norwegian and North Sea fisheries produces a similar trajectory.

Time Series: North Sea Herring

On the ecological ordination (Fig. 2), the 1950 North Sea herring corresponds with a period of high catch below the age at first maturity during a rapidly expanding juvenile fishery. By 1974 both stock and range collapse (Whitmarsh et al. 1995), ascribed retrospectively by Saville and Bailey (1980) to overexploitation, were so serious that the fishery was closed, and so the 1975 point has moved toward "good," a trend continued to the 1980 point, which reflects stock rebuilding during closure. In more recent years the rebuilt North Sea herring stock has again suffered serious declines so that between 1990 and 1997 the fishery has continued to move back in the "bad" direction almost to where it started in 1950.

On the economic ordination (Fig. 3), the North Sea fishery clearly improved from 1950 to 1980, and some of this improvement was maintained until 1997.

Technological change was rapid in the North Sea during the 1960s as the fleet shifted from drifters to trawlers and purse seines. Moreover, primarily as a result of the decline in the Atlanto-Scandian herring fishery, purse seine fleets moved to the North Sea, increasing from 16 vessels in 1963 to 326 in 1967, by which time the proportion of trawlers dropped, and drift and ring-netting had all but disappeared. The geographic mobility of vessels and crew resulted in a high degree of transfer from innovators (Whitmarsh et al. 1995) so that, concurrent with advances in gear and storage of fish, came the ability to increase trip length. These changes are reflected in this fishery's ordination trajectory (Fig. 5) which shifts toward "bad" by 1975 and 1980. More recent changes, indicated by shifts on the plot from 1990 to 1997, have made little improvement. Note that in this discipline we have scored attributes indicative of highly mechanized, capital-intensive industrial fisheries as intrinsically "bad" in terms of sustainability.

On the social ordination (Fig. 4), little change is discernible except for a small movement toward the "bad" direction in recent years.

The combined, interdisciplinary ordination (Fig. 6) exhibits the trajectory that most accurately reflects what we know of the history of this fishery, remaining poor from 1950 to 1964, improving after closure but then shifting back toward "bad" in recent years.

Time Series: Norwegian Spring Spawning Herring

On the ecological ordination (Fig. 2), the Norwegian spring spawning herring stock is remarkably similar to the North Sea trajectory. It shows clear improvement from overfishing periods in the 1950s and 1960s to the fishery closure period during the 1970s and 1980s (Dragesund et al. 1980). But the trend then reverses and by 1996 the location moves back near to the 1960s points when commercial catches commence again after the closure (Fiskeridirektoratet 1996, ICES 1996, Slotte and Johannessen 1997).

Economically, the Norwegian herring fishery during the 1950s and 1960s lies in the poorest position of the ordination (Fig. 3), reflecting a particularly low economic yield. The situation improved while the fishery was closed in the 1970s and 1980s and the current economic position is clearly much improved, reflecting a greater quantity destined for direct human consumption (Bjørndal et al. 1996). The present location is, however, within the "random" attribute score area so a large movement would be required to verify any further improvement or deterioration.

On the social ordination (Fig. 4), the fishery exhibits the greatest movement of any of our time series. It plots amazingly close to the "best" possible fishery scorable on these attributes in the 1950s, 1960s, and into the 1970s, when many fishermen were employed in the fishing sector and supported families along the coast (Myrstad 1996). But when the fishery closed in the 1970s the social impact was understandably negative, and the ordination moves in the "bad" direction. The negative effect worsened by the 1980s after 20 years without any herring fishery. After the fishery re-opened in the late 1980s, the last two points show large negative movements toward "bad" and it is now as poor socially as the North Sea. Compared to previous decades, this may be connected with drives to increase economic efficiency such as centralization, rationalized fleet structure, a dramatic reduction in the number of fisherman, less family fishing, and fewer employees in the sector (Bjørndal et al. 1996).

Like most fisheries in the world, technologically the herring fishery in Norway has undergone dramatic changes (Fig. 5). The introduction of sonar, the power block, and a change in the fleet structure toward very large, efficient, and sophisticated purse-seiners (Myrstad 1996) have all affected the herring fishery quite dramatically, and this is seen in the large movement across the plot from 1950 to 1980, changes nevertheless viewed by our present analysis as neutral in terms of sustainability. Since 1980, recent points have moved in the "good" direction, perhaps driven by the increased use of gear selectivity devices to reduce bycatch.

The combined ordination linking the ecological, economic, social, and technological disciplines shows a clear movement toward good from 1970 to 1980, but progressive shift back to "bad" since then.

Individual Fisheries

The Peruvian anchovy fishery is plotted just before and just after the major collapse caused by El Niño and overfishing in 1970-1971. On the combined ordination (Fig. 6) the pre-collapse fishery has a location worse than any other point, and improvement by 1975 is also apparent on the economic and ecological plots (Figs. 2 and 3). Social and technological changes are minor (Figs. 4 and 5). This fishery is in the middle range socially, but the worst economically and on the combined plot, and among the worst ecologically and technologically.

The four Adriatic fisheries may be compared (Figs. 2–5). On all except the economic ordination, the unselective pair trawls "volante" score in worse locations than the night-operated selective purse seines "lampara," and, for both gear types, the anchovy, which is heavily fished and underwent a stock collapse the mid-1980s, scores worse than the less valuable and less heavily fished sardine fishery. Both "lampara" fisheries score very high on the social ordination, and are among the best ecologically and in the combined analysis. In general, fisheries for both sardines and anchovies in the Adriatic tend to score higher than their counterparts elsewhere.

On the combined economic and social ordinations, anchovy fisheries with the exception of the Peruvian fishery, tend to fall in better positions than sardines and herring. Among the anchovy fisheries, the Japanese fisheries, especially the fixed gear, consistently ordinate in "better" locations. Among the sardine fisheries, the South African is generally among the worst, especially on the ecological ordination, and we are unsure of the reason for the Brazilian often scoring better than Mexican, Japanese, and South African fisheries. Japanese and Brazilian sardines score very high on the social ordination.

The Alaska herring fisheries usually ordinate very close together, but the Nunivak Island and Prince William Sound fisheries seem to score better than Security Cove, Good News Bay, and Cape Avinoff, and generally the Alaska herring fisheries score better than the 1996 British Columbia herring fishery.

Among the present-day Atlantic herring fisheries, the fixed gear Bay of Fundy herring fishery is by far the best, falling among the best three on the combined plot (Fig. 6). The Icelandic and Bay of Fundy fisheries tend to score better than the Norwegian or North Sea fisheries, whereas the Baltic fishery often falls between the two. The relatively poor location of the two Icelandic fisheries in the economic ordination is a surprise, and suggests that other factors outweigh the high scores given to individual transferable quotas (ITQs).

Pacific herring fisheries, with smaller migration ranges that subject them to fewer jurisdictions, tend to ordinate in better locations than Atlantic herring fisheries (Fig. 6). Finally, with exceptions like the Peruvian anchovy and South African sardine, herring fisheries in general appear to score less well than most sardine or anchovy fisheries, perhaps reflecting their location in northern developed nations, a longer history of heavy exploitation, and a higher degree of capitalization.

On the combined plot (Fig. 6), the fixed gear fisheries are the best in terms of sustainability and the Peruvian anchovy the worst, with a wide spread of fisheries in between. The ecological ordination suggests that the best present day fisheries are for Alaskan and B.C. herring, the Adriatic lampara sardine, and the Bay of Fundy fixed gear herring (Fig. 2). The worst are the Peruvian anchovy, South African sardine, and the North Sea and Norwegian herring. The best fisheries socially (Fig. 4) are the Brazilian and Japanese sardines, and the Adriatic lampara fishery, while the worst are the B.C. and Norwegian herring. Economically the fixed gear Bay of Fundy herring fishery scores high (Fig. 3), along with B.C. herring. The Peruvian anchovy is by far the worst on the economic ordination. Figure 5 suggests that the best fisheries technologically are the fixed gears for anchovy and herring, while there is a large group of sardine and Pacific herring fisheries near to "bad." The preponderance of points to the lower right in this ordination suggest there is great scope for technological improvement in sustainability status.

Discussion

These results are encouraging, but it is clear that this technique needs refinement before it can be used on a new fishery with confidence, and before it can fully meet the aims set out in the introduction. Further simulations similar to those reported here will be required to validate the method. It is, however, promising that MDS simulation trajectories are monotonic and largely preserve relative changes in status. We are fortunate in having access to the considerable insight of non-parametric ordination that has been developed recently in marine ecology (Clarke and Warwick 1997).

In this paper, the ordinations for the three herring fisheries mirror in general what we know of the ecology, economics, and sociology associated with their histories of collapse and rebuilding, and their relative status and characteristics. Here, the technique appears to be giving the right signals about shifts in sustainability status, which might be loosely defined as fisheries "health." Among individual fisheries, such as the four fisheries for Adriatic pelagics, the ordinations appear to reflect what we know of their relative status. Status evaluations based on this technique might therefore serve to help set priorities for the allocation of limited survey and formal assessment management resources, so aiding a triage of fisheries.

In the present analysis we took attribute scores only for years when significant changes were documented. Ideally, individual fishery trajectories should be plotted as time series year by year but many technological, social, and economic attribute scores are not readily available on this basis. Care also has to be taken that correspondents who supply data interpret the scoring system in the same way. On the other hand, once it is begun, an analysis should be easily updated year by year. Some of the herring ordination trajectories could be extended back to the era of sail and steam drifters as fisheries became industrialized in the nineteenth century. Data for such an analysis could likely be found in historical records and documents from the North Sea, Scandinavia, and elsewhere. Additional ordination disciplines might be added effectively, for example management instruments, detailed changes in a gear over time, or ecosystem impacts.

A critical question is whether this technique can be used to diagnose key problems, (such as environmental change, overcapitalization, or recruitment overfishing) early enough to give warning of impending trouble. Ideally, all analyses would use the same universal set of attributes so that direct comparison can be made. At present, however, we find that attributes are likely to be tuned to a particular type of fishery. Attributes that more clearly define the period immediately preceding a documented collapse would increase the power of this method. A more formal analysis of the leverage of individual attributes would also be valuable.

Despite these problems, we feel that the method has promise in that it is robust in several senses. First, it can learn not only from refinement of historical analyses but also from more formal stock assessment science about what are "good" and "bad" attributes that may be scored. Second, users can make their own choice of which discipline to concentrate on; the biological, social, or economic analyses can be used alone or in combination with the technological area if required. It also provides a quantitative way of conflating interdisciplinary evaluations, a process considered essential in the management of many fisheries (Lane and Stephenson 1997). The present method avoids arbitrary weighting of disciplines in that the final ordination will reflect statistically the original choice of attribute values running from "bad" to "good" as the two fixed extremes. Third, the new method is robust in the sense that criticism of attributes from within each of the disciplines such as economics, anthropology, ecology, or stock assessment can serve to improve the power of the ordinations within each field, rather than invalidate the method.

For multivariate analysis, ordination techniques have a benefit in that they provide a "statistical common currency" for multivariate measures that are free both of the underlying models of stock assessment or modeling, and of the underlying assumptions that relate utility scores to measured attributes in decision analysis techniques. Moreover, updated ordinations can be performed each year on the basis of approximate information, minimizing the need for frequent surveys. We do not suggest that this method replace conventional stock assessment, but we do offer the technique as way of a cost-effective, heuristic, and rapidly updated appraisal of the exploitation status of fisheries and their sustainability.

Acknowledgments

We would like to thank all of the respondents listed in the Table 1 for generously providing data about their fisheries. Jake Rice (DFO, Canada), Peter Gehrke, Geoff Gordon, and Doug Ferrell, (NSW Fisheries, Australia), André Punt and Dennis Heinemann (CSIRO, Australia), and two referees have all made valuable suggestions that have improved the method.

References

Alaska Department of Fish and Game. 1996. From Web site www.state.ak.us/local/akpages/fish.game/notebook/herring.htm.

Bailey, R.S. 1992. The global pelagic fish resource and its biological potential. In: J.R. Burt, R. Hardy, and K.J. Whittle (eds.), Pelagic Fish: The resource and its exploitation. Fishing News Books, Oxford, U.K. 352 pp.

Bakun, A. 1996. Patterns in the Ocean: Ocean processes and marine population dynamics. California Sea Grant College System, La Jolla. 323 pp.

Bjørndal, T., A.D. Hole, W.M. Slinde, and F. Asche. 1996. Norwegian spring spawning herring: Some biological and economic issues. Working paper no. 29/1997, Institute of Marine Research, Bergen.

Burd, A.C. 1974. The Northeast Atlantic herring fishery and the failure of an industry. In: F.R. Harden Jones (ed.), Sea fisheries research. Elek Science, London, pp. 167-191.

Carr, M.C. 1997. PRIMER user manual: Plymouth routines in multivariate ecological research. Plymouth Marine Laboratory, U.K. 27 pp.

Central Intelligence Agency. 1995. World factbook 1995-1996. Brassey, Washington, DC. 512 pp.

Cergole, M.C. 1993. Avaliacao do estoque da sardinha-verdadeira, *Sardinella brasiliensis*, na costa sudeste do Brasil, periodo 1977-1990. PhD. dissertation. Universidade de São Paulo. 245 pp. (In Portuguese.)

Clarke, K.R. 1993. Multivariate analyses of changes in community structure. Aust. J. Ecol. 19:117-143.

Clarke, K.R., and R.M. Warwick. 1997. Change in marine communities: An approach to statistical analysis and interpretation. Plymouth Marine Laboratory, Plymouth, U.K. 144 pp.

Coull, J. 1988. The North Sea herring fishery in the 20th century. In: E.M. Borgese, N. Gindburg, and J.R. Morgan (eds.), Ocean yearbook 7. University of Chicago Press, pp. 115-131.

Cushing, D.H. 1988. The provident sea. Cambridge University Press, U.K. 329 pp.

Diegues, A.C.S. 1995. Povos e mares: Leituras em socio-antropologia maritima. NUPAUB-USP (Nucleo de Aporo a Perquisa valore Populaces Humanas e Areas Umidas Brasilienas-USP), São Paulo, Brazil. 296 pp. (In Portuguese.)

Dragesund, O., J. Hamre, and Ø. Ulltang. 1980. Biology and population dynamics of the Norwegian spring-spawning herring. Rapp. P.-V. Reun. Cons. Int. Explor. Mer. 177:43-71.

Fernö, A., T.J. Pitcher, W. Melle, L. Nøttestad, S. Mackinson, C. Hollingworth, and O.A. Misund. 1998. The challenge of the herring in the Norwegian Sea: Making optimal collective spatial decisions. Sarsia 83:149-167.

Fiskeridirektoratet. 1996. Beskatningsstrategi for norsk vårgytende sild. Rapporter og Meldinger nr.1. Fiskeridirektoratet, November 1996. (In Norwegian.)

Gagne, J.A., J. Lambert, F. Mongeau, and L. Fortier. 1991. Population dynamics of herring in the St. Lawrence estuary: A test of the member/vagrant hypothesis. In: Proceedings of the International Herring Symposium. University of Alaska Sea Grant, AK-SG-91-01, Fairbanks, pp. 103-106.

Garcia, S.M., and C. Newton. 1994. Current situation, tends and prospects in world capture fisheries. FAO, Global Trends in Fisheries Conference, Seattle, June 1994.

Hourston, A.S. 1980. The decline and recovery of Canada's Pacific herring stocks. Rapp. P.-V. Reun. Cons. Int. Explor. Mer. 177:143-153.

Hourston, A.S., and C.W. Haegle. 1980. Herring on Canada's Pacific coast. Can. Spec. Publ. Fish. Aquat. Sci. 48. 23 pp.

ICES. 1996. Report of the Atlanto-Scandian Herring, Capelin and Blue Whiting Assessment Working Group. Bergen 12-18 October 1995. ICES C.M. 1996/Assess:9.

Kruskal, J.B., and M. Wish. 1978. Multidimensional scaling. Sage Publications, Beverly Hills, CA.

Lane, D.E., and R.L. Stephenson. 1997. Fisheries management science: Integrating the roles of science, economics, sociology and politics into effective fisheries management. In: D.A. Hancock, D.C. Smith, A. Grant, and J.P. Beumer (eds.), Developing and sustaining world fisheries resources: The state of science and management. CSIRO, Collingwood, Australia, pp. 177-182.

Mackinson, S., R. Sumaila, and T.J. Pitcher. 1997a. Bioeconomics and catchability: Fish and fishers behaviour during stock collapse. Fish. Res. 31:11-17.

Mackinson, S., M. Vasconcellos, T.J. Pitcher, C.J. Walters, and K. Sloman. 1997b. Ecosystem impacts of harvesting small pelagic fish using a dynamic mass-balance model. In: Forage fishes in marine ecosystems. University of Alaska Sea Grant, AK-SG-97-01, Fairbanks, pp. 731-750.

Matsuura, Y. 1981. Analise economica da producao da sardinha na regiao sudeste do Brasil. Bol. Inst. Oceanogr., S. Paulo, 30(1):57-64. (In Portuguese.)

Matsuura, Y. 1990. Rational utilization of coastal ecosystem in tropics: Integrated investigation of coastal ecosystem in Ubatuba region. In: Simposio de ecossistemas da costa sul e sudeste brasileira: Estrutura, funcao e manejo, 2. ACIESP [Academia de Ciências do Estado de São Paulo] V.1, São Paulo, pp. 47-52. (In Portuguese.)

Myrstad, B. 1996. The Norwegian fishing industry. Ministry of Foreign Affairs. UDA245ENG, Oslo.

Outram, D.N., and R.D. Humphreys 1974. The Pacific herring in British Columbia waters. Fisheries and Marine Service, Pacific Biological Station, Nanaimo, B.C. Circular No. 100.

Pearse, P.H. 1982. Turning the tide: A new policy for Canada's Pacific fisheries. The Commission on Pacific Fisheries Policy, Vancouver, Canada. 292 pp.

Pitcher, T.J. 1995. The impact of pelagic fish behaviour on fisheries. Sci. Mar. 59:295-306.

Pitcher, T.J. 1997. Fish shoaling behaviour as a key factor in the resilience of fisheries: Shoaling behaviour alone can generate range collapse in fisheries. In: D.A. Hancock, D.C. Smith, A. Grant, and J.P. Beumer (eds.), Developing and sustaining world fisheries resources: The state of science and management. CSIRO, Collingwood, Australia, pp. 228-232.

Pitcher, T.J., and J. Parrish. 1993. The functions of shoaling behaviour. In: T.J.Pitcher (ed.), The behaviour of teleost fishes, 2nd edn. Chapman and Hall, London, pp. 363-439.

Pitcher, T.J., M. Esseen, and N. Cingolani. 1992. Analysis of Adriatic pelagic fish stocks. Marine Resources Assessment Group, London. 55 pp.

Pitcher, T.J., A. Bundy, D. Preikshot, T. Hutton, and D. Pauly. 1998. Measuring the unmeasurable: A multivariate interdisciplinary method for determining the health of fisheries. In: T.J. Pitcher, P.J.B. Hart, and D. Pauly (eds.), Reinventing fisheries management. Chapman and Hall, London, pp. 31-54.

Rossi-Wongtschowski, C.L.D.B., S.A. Saccaro, and M.C. Cergole. 1995. Situação do estoque da sardinha na região sudeste-sul do Brasil. Coleção meio ambiente. Série Estudos de Pesca. Ibama (Itajaí, SC). M17.

Saville, A., and R.S. Bailey. 1980. The assessment and management of the North Sea herring stocks in the North Sea and the west coast of Scotland. Rapp. P.-V. Reun. Cons. Int. Explor. Mer 177:112-142.

Schiffman, S.S., M.L. Reynolds, and F.W. Young. 1981. Introduction to multidimensional scaling: Theory, methods and applications. Academic Press, London.

Slotte, A., and A. Johannessen. 1997. Exploitation of Norwegian spring spawning herring (*Clupea harengus* L.) Before and after a stock decline: Toward a size selective fishery. In: D.A. Hancock, D.C. Smith, A. Grant, and J.P. Beumer (eds.), Developing and sustaining world fisheries resources: The state of science and management. CSIRO, Collingwood, Australia, pp. 103-108.

SOE. 1994. Sustaining marine resources: Pacific herring fish stocks. Environment Canada, Ottawa. State of the Environment Bulletin No. 94-5.

SPSS Inc. 1996. SPSS for Windows, release 7.5. SPSS Inc.

Stalans, L.J. 1995. Multidimensional scaling. In: L.G. Grimm and P.R. Yarnold (eds.), Reading and understanding multivariate statistics. American Psychological Association, Washington, DC. 373 pp.

Taylor, F.H.C. 1964. Life history and present status of British Columbia herring stocks. Bull. Fish. Res. Board Can. 143. 81 pp.

Taylor, F.H.C. 1985. Introduction to the international Symposium on the Biological Characteristics of Herring and Their Implications for Management. Can. J. Fish. Aquat. Sci. 42(Suppl. 1):2.

Valentini, H., and R. de D. Cardoso. 1991. Analise da pesca da sardinha-verdadeira, *Sardinella brasiliensis*, na costa sudeste-sul do Brasil. Atlantica, Rio Grande, 13:45-54.

Whitmarsh, D.J., C.A. Reid, C. Gulvin, and M. Dunn. 1995. Natural resource exploitation and the role of new technology: A case history of the UK herring industry. Environ. Conserv. 22:103-110.

Zheng, J. 1996. Herring stock-recruitment relationships and recruitment patterns in the North Atlantic and Northeast Pacific oceans. Fish. Res. 26:257-277.

Development of a Simple Biomass Analysis Model

Rex Baleña
University of the Philippines in the Visayas, College of Fisheries, Iloilo, Philippines

Abstract

A simple oceanographic mass conservation model is formulated with the capability to generate rich information on new and dynamic indices of biomass analysis solely from catch data. The model utilizes modal data decomposition and features some advantages: control of data errors, reduction of uncertainties due to multi-parameterization, economy of computation, and simple but robust interpretation. A case application is discussed along with the potential refinements of the model and its implications to modern stock assessment.

Introduction

Procedures in modern stock analysis may be thought of generally to have evolved from two types of complementary models: the Schaefer (1954) model and the Beverton and Holt (1957) model. Although the earlier is relatively unpopular and has been criticized as inflexible (Gulland 1969), apparently it is simple and has a low demand on input data. This desirable feature is exploited in the present conceptualization of a new tool for biomass analysis.

Coincidentally, this paper discusses the type of equation used by Schaefer, except differentiating an alternative interpretation of that equation; i.e., while Schaefer utilized only the usual prognostic interpretation, the equation is interpreted presently as an oceanographic mass continuity model (MCM from here on). Consequently, new indices useful to dynamic biomass analysis are derived: biomass divergence, which acts as center of biomass redistribution, and biomass flux or “migration” of biomass from or to the centers of redistribution.

MCM is an integral concept and complements the popular “black box” (Sparre 1985) on its inherent problem of differentiating the biomass flux into poorly known components. Specifically, because it appears that the analysis of biomass has only one piece of solid information to begin with, which is the catch data, it is this information alone that MCM exploits to describe biomass exchanges as an integral process. In contrast to the latter model, MCM produces new and meaningful results relative only to this single given information.

Methodology

MCM—The Theory

A philosophical basis for formulating MCM is the fact that the real ocean is a contiguous integral environment. Hence, though organisms may be found in "patches," they are not truly confined within arbitrary boundaries; they can move around (at will or in response to environmental changes) and, in so doing, redistribute biomass. Like many phenomena in nature, this redistribution process is governed by a well-founded law called the conservation of mass. The human role in this process represents an intervention which may modify only the process but not the basic conservation principle. The robustness of this principle is appealing to biomass modeling.

In fisheries, a relatively solid indicator of human disruption of the ocean biomass exchanges is the catch data, which is defined here simply as the amount of biomass obtained through fishing. The following demonstrates that the use of this data alone, in conjunction with the mass conservation principle, results in a simple but useful model that is devoid also of parameterization.

Consider a chunk of oceanic water as in Fig. 1a. Invoking mass continuity, or mass conservation in oceanographic parlance, a balance of biomass fluxes for the volume can be formulated as

$$\mathrm{d}_t B = F_{in} - F_{out} = F_{net} \tag{1}$$

where $\mathrm{d}_t B$ is defined as the temporal change of biomass B in the volume that is associated with F, the flux of biomass into or out of the same volume. Particularly, F comprises fluxes defined to occur at open lateral boundaries. In general, these fluxes need not balance and, as expressed by (1), a net flux occurs. Consequently, because fluxes at solid boundaries (e.g., the coastline) and ocean bottom may be considered nil, $\mathrm{d}_t B$ assumes this form of (net) flux that must occur at the surface. Thus, we use also the term surface divergence (or simply, divergence) for this flux because of its resemblance to an oceanographic analogue which is the surface divergence of water in a region associated with subsurface horizontal water fluxes converging to the same region. (The opposite phenomenon is called surface convergence and is associated with subsurface horizontal outfluxes.)

MCM is thus a simple statement of the balance of biomass fluxes at the ocean surface and at open horizontal boundaries. Particularly, the flux at the surface (divergence) is due mainly to fishing and, at the latter boundaries, due to "migration" of biomass (toward the region if there is surface

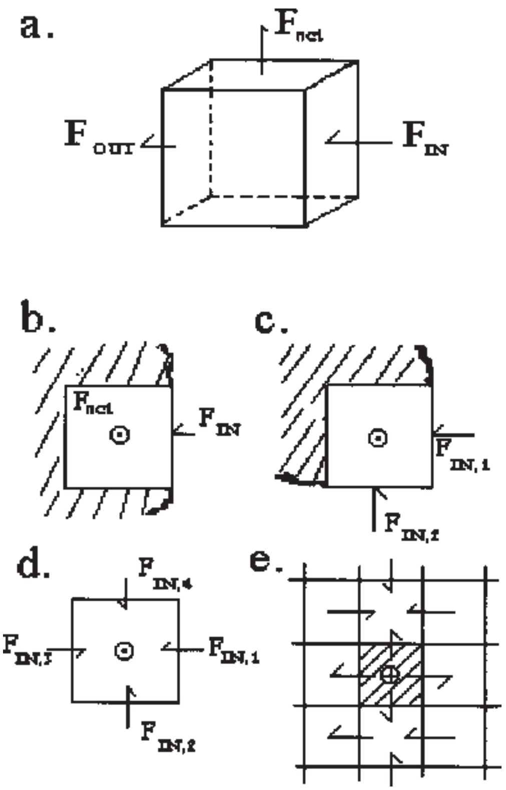


Figure 1. (a) Schematic of the mass balance in a given volume. (b) The simplest balance $F_{net} = F_{in}$ *(plan view). (c) An additional boundary to (b) makes the problem underdetermined relative to two influxes; the balance is* $F_{net} = F_{in,1} + F_{in,2}$*. (d) The most underdetermined condition wherein all boundary fluxes are unknown;* $F_{net} = F_{in,1} + F_{in,2} + F_{in,3} + F_{in,4}$*. (e) Sharing of boundary fluxes between cells. Rectangular cells had been assumed in all diagrams for convenience.*

divergence, as in Fig. 1b). Note that the contrasting processes of immigration and emigration contribute to the latter process but, to avoid confusion, these terms are not used in lieu of horizontal fluxes which, generally, refer to the horizontal movement of biomass or the tendency toward that movement. To be specific, this new quantity is the unknown in the present model.

Figures 1b-1d illustrate the relation of divergence and horizontal flux(es). The simplest case in Fig. 1b shows how in a semi-enclosed region a single horizontal influx explains totally the divergence (due to catch). Therefore, given the catch data as a function of time, *B(t),* both quantities, notably the influx, are obtained at once from (1). Certainly, a similar straightforward calculation cannot be done for the case shown in Fig. 1c, because an additional open boundary makes the problem underdetermined relative to the given data. Moreover, the worst case can happen in the open ocean where all horizontal fluxes are unknown, as in Fig. 1d. However, regardless of the configuration, a pragmatic method is available to determine these fluxes making use of only the catch data.

In general, catch data is a function of space and time (i.e., $B[\underline{x},t]$) and may not be evenly distributed along either dimensions. In particular, biomass has an observed "patchy" distribution in space. However, these are mere sampling problems such that, in the absence of dense data, one may resort to interpolation, if only to proceed with the necessary computations. Certainly, interpolated values can be meaningless and have to be interpreted with caution but, unless someone can finally claim to have sampled every square meter of the ocean at all times, these values remain as useable estimates.

To proceed with spatial interpolation, it is helpful to grid arbitrarily an area of interest into uniform cells (e.g., rectangles) such that, if the observations, $B(\underline{x},t)$, represent only certain cells, data-void cells may be supplied with appropriate values following, e.g., the procedure of optimal interpolation (Levitus 1982). A map of $B(\underline{x},t)$ will help describe its spatial distribution. A gridded domain is helpful also to depict, as in Fig. 1e, how neighboring cells share fluxes at their adjacent boundaries. This map of horizontal fluxes is a desirable product of the analysis but is rather not straightforward to obtain.

Once $B(\underline{x},t)$ is gridded, divergence, $\mathrm{d}_t B(\underline{x},t)$, can be computed directly through simple cellwise temporal differentiation. This is a critical stage of the analysis for two reasons. First, a contour map of the latter quantity will reveal features in the domain resembling as either "sinks" or "sources" of biomass in the sense that, in these places, a "deficit" (due to divergence or catch) or "excess" (due to negative divergence or convergence) of biomass has occurred, respectively. Specifically, divergence implies an increased catch and, consequently, an increased tendency for the horizontal convergence of biomass flux, e.g., what is expected normally when a region responds to or begins recovering from fishing stress. Likewise, convergence implies decrease in catch and, consequently, an increased tendency for biomass outflux. Divergences and convergences are significant features because they act as dynamic centers of biomass redistribution and are, therefore, important starting points in biomass or stock analysis. Second, the potential pathways associated with the biomass redistribution process (i.e., the trajectory of the fluxes) can be traced simply from

$$\boldsymbol{P}(\underline{x},t) = \nabla_h \mathrm{d}_t B(\underline{x},t) \tag{2}$$

where ∇_h is a horizontal gradient operator. A vector plot of this function will indicate orthogonals to the contours of $\mathrm{d}_t B(\underline{x},t)$ (i.e., $\mathrm{d}_t B(\underline{x},t) = c$), linking regions with excess biomass to regions of deficit biomass (i.e., the dynamic centers). Therefore, one obtains also an indication of the path of biomass movement—a very valuable guide at least to migration studies. Finally, it remains optional to deduce from these pathways the detailed fluxes across cell boundaries, as in Fig. 1e, if one desires to utilize these fluxes as boundary estimates to complete cell balances. Particularly, biomass budget studies of specific areas with relatively more comprehensive data will benefit from this procedure.

Modal Decomposition

Since $B(\underline{x},t)$ is the sole input to MCM, its quality is crucial to the success of the model. Hence, it is desirable to remove its inconsistencies before these are aggravated by differentiation. For this purpose, empirical orthogonal functions or EOF (e.g., Kutzbach 1967, Weare 1976, Wenzel 1984) are invoked as a natural means of re-expressing the observations into a noise-free model input. Specifically, the novel application of these functions in estimating errors jointly as a function of space and time is introduced. This procedure may be outlined briefly.

In the case of the catch data, the simultaneous determination of data modes or quality data, $\mathrm{D}(\underline{x},t)$, and error, $\xi(\underline{x},t)$, proceeds by assuming that

$$B(\underline{x},t) = \mathrm{D}(\underline{x},t) + \xi(\underline{x},t) \tag{3}$$

The usual EOF procedure computes for $\mathrm{D}(\underline{x},t)$ through a decomposition of $B(\underline{x},t)$, which is assumed a separable function of independent space and time components. An eigen-manipulation of the data covariance matrix finally yields $\mathrm{D}(\underline{x},t)$, after truncating these components using an appropriate criterion (e.g., Catell 1966). The general outcome demonstrates the power of modal decomposition to extract only a few most dominant modes of the data. Moreover, the procedure is most conveniently applied to time series data (below).

It is not usually done but, at this point, $\xi(\underline{x},t)$ is obtainable immediately from equation (3), simply as the residual of two known quantities. (A computer shareware to do this calculation and the eigen-decomposition is in preparation.) In this paper, we make use of this function to quantify the errors in catch data. Intrinsically, it possesses the desirable property of incorporating those observational inconsistencies with respect to both space and time.

An Application

An attempt at elucidating the application MCM is necessary. For this purpose, the Philippine archipelago is used because it typifies a complex topography to test the procedures (Fig. 2). The archipelago is constituted by some 7,107 islands of various shapes and sizes forming an intricate network of small seas, bays, and waterways. Yearly catch data for 10 of the 12

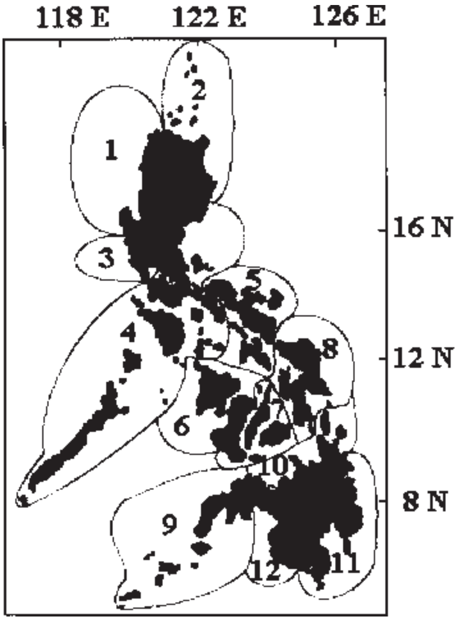


Figure 2. The Philippine archipelago and its 12 fishing regions. Only Regions 1 to 10 have complete data for analysis. (Source: BFAR 1972, 1973, and 1975.)

designated fishing regions were obtained for the period 1968 to 1975. This dataset represents one of the more complete time series that could be formed from a compilation of landing statistics (BFAR 1972, 1973, and 1975). Data decomposition as outlined previously was performed on this dataset in order to isolate the "error-free" component which was then used in the subsequent calculations.

The archipelagic domain was gridded and the processed data decimated proportionately over the respective regions according to area. Optimal interpolation was then performed to provide smooth transition between data-rich and data-void regions. This artificial procedure was necessary to avoid discontinuity problems in calculating derivatives. Afterward, separate modal decompositions were done for the data and divergence fields.

Results

Looking first at the observations, a comparative plot of the fish production for Regions 1 to 10 is shown in Fig. 3 along with the quality data, mean data, and error estimates from modal decomposition. The numeri-

Table 1. Fish production from10 regions of the Philippine archipelago with corresponding computed errors. Quality data (not printed) can be derived simply as the difference between the raw observations and these errors.

	Observations (× 10^6 kg)							
Region	1968	1969	1970	1971	1972	1973	1974	1975
1	1.51048	0.91868	1.07820	0.75824	0.82013	1.29546	1.71028	1.37997
2	0.87892	1.33180	0.94888	0.60224	1.41164	1.97223	1.30679	1.41635
3	310.66856	280.80096	298.19234	288.99692	296.62104	2.44386	1.85641	1.38272
4	9.64972	7.76212	7.76272	10.31332	24.49505	319.68174	235.19567	207.78613
5	63.96904	58.32484	54.57200	62.46500	57.83424	23.28600	14.38693	31.67572
6	5.99764	4.95292	5.20578	4.33540	13.94168	80.85922	151.64648	205.38829
7	7.57148	6.26700	5.44280	3.11260	20.52914	8.72095	17.56756	12.88235
8	4.97804	6.92444	7.27652	10.16788	7.46540	5.44624	9.18218	9.10980
9	0.73144	0.62772	0.63456	0.50976	0.86860	10.49140	16.77019	14.21560
10	0.83856	0.81616	0.76332	1.01428	0.76720	0.79052	8.19169	1.07833

	Errors (× 10^6 kg)							
Region	1968	1969	1970	1971	1972	1973	1974	1975
1	0.44861	–0.03678	0.07324	–0.23500	–0.27830	–0.26888	0.33948	–0.01467
2	–0.18980	0.37061	–0.06166	–0.39772	0.29834	0.25718	–0.19638	–0.11213
3	–0.21980	–0.35140	1.04422	–0.77421	0.28913	0.62039	1.23403	–1.92952
4	–0.13392	–0.22415	0.18883	0.57104	–0.20174	42.03657	–8.66193	–38.76615
5	1.60781	2.00458	–4.87586	4.28457	–2.96669	–1.63250	–7.30546	9.21291
6	0.23337	0.22050	0.68752	–1.38490	–0.12549	–73.98976	15.64256	67.87884
7	–1.35442	–1.76141	–2.99838	–5.23864	11.24085	–5.42317	5.17124	0.27630
8	–2.65294	0.04937	0.03648	3.03642	–0.27252	–2.98219	1.80254	1.58531
9	0.17157	0.16724	0.19399	–0.04524	–0.48286	–4.28051	3.79604	1.09775
10	–0.00331	0.06563	–0.01903	0.22178	–0.22814	–2.69784	5.13011	–2.02329

cal values are in Table 1. The largest variations in the observations are due only to Regions 3 to 6. Altogether, Figs. 3a to 3c reveal that the extracted error field has the largest variance beginning only in 1973 (coincidentally, right after the onset of the 1972 El Niño and the declaration of Philippine Martial Law) and due mainly to Regions 4 and 6. The mean curve for the archipelago in Fig. 3d shows a low catch from 1969 to 1971 and a rapid increase afterward. The cause of this behavior is evident in the regional plots. The low catch is attributed to Regions 1,2,4, and 6 to 10, while the recovery is due mainly to Regions 4 and 6, i.e., an apparent indication of a regional shift in catch effort. The overall trend in catch is increasing, suggesting that exploitation had increased. In fact, this trend had become obvious since 1956 (Fig. 4). Thus far, the observations have twofold indications. First, regardless of the cause of the increase in the volume of catch, e.g., population growth or demand for export (Fig. 5), the archipel-

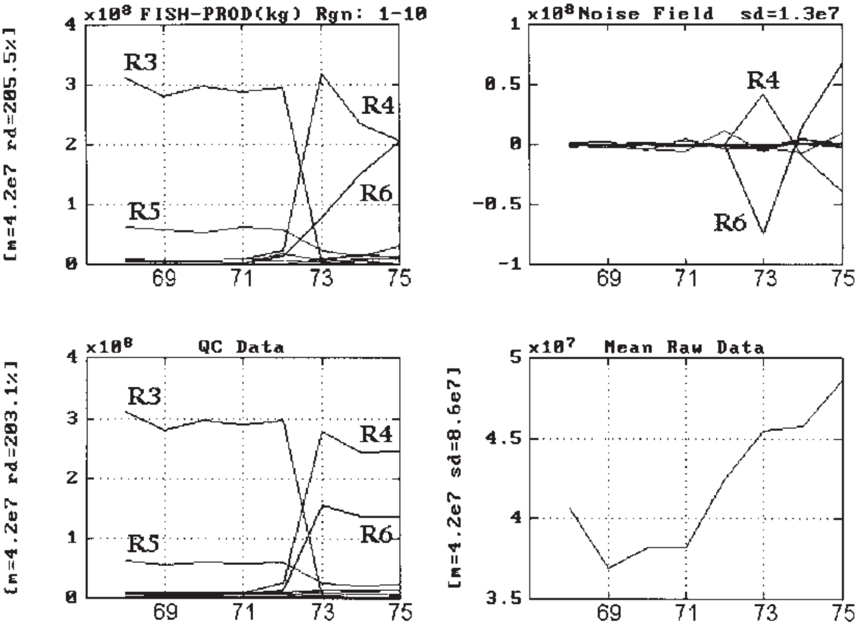


Figure 3. *Processing of catch data using modal decomposition: (a) raw data from10 regions, (b) calculated errors, (c) quality data (QC, quality control), and (d) mean raw data. Numerical values for (a) and (b) are found in Table 1. Relevant statistical parameters are shown: m = mean, sd = standard deviation, and rd = relative dispersion.*

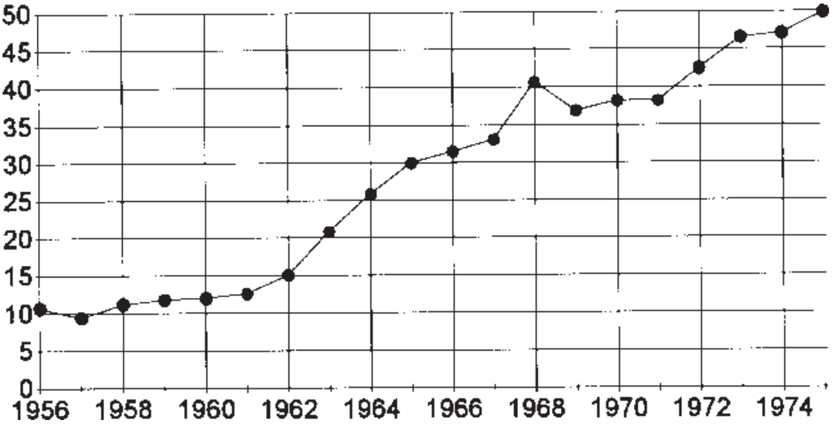
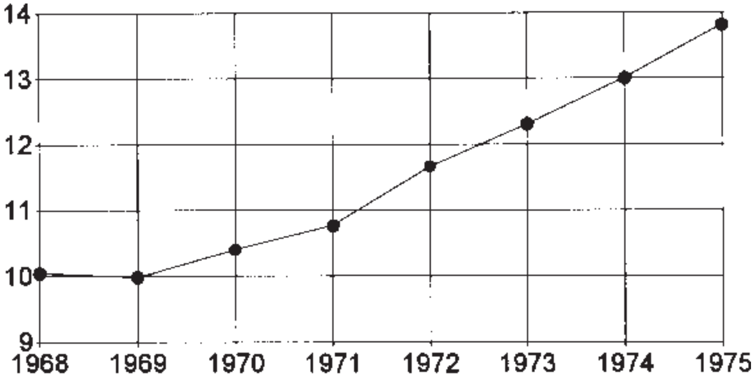


Figure 4. *Increasing trend in catch from 1956 to 1975 in 10^7 kg.*



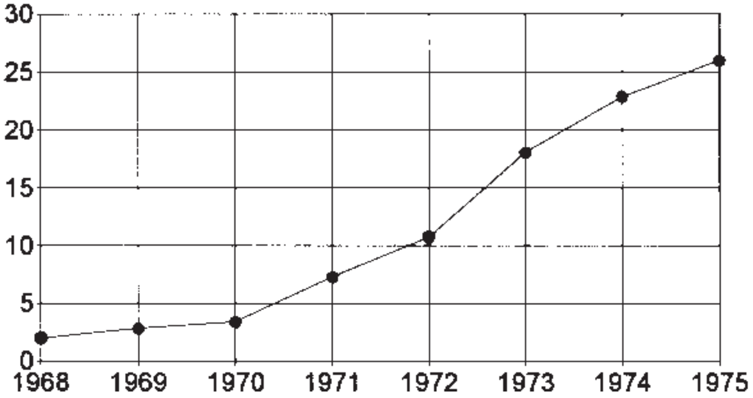


Figure 5. Similar trend as in Fig. 4, except for (a) consumption (related to population growth) and (b) export and for the period 1968-1975. Plotted values are in 10^8 kg and 10^9 kg, respectively.

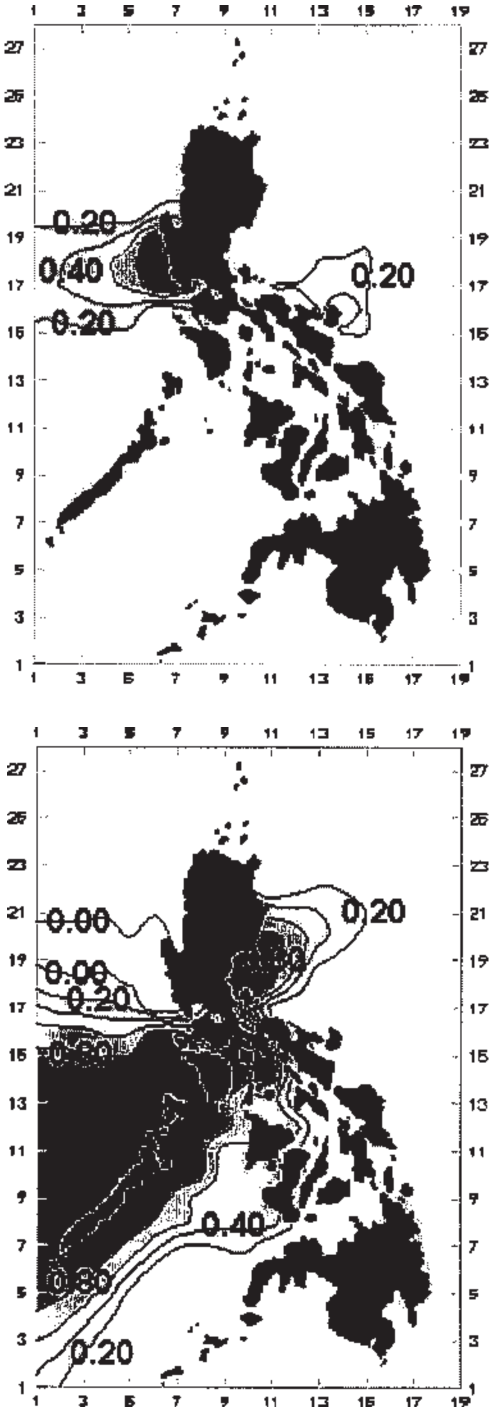


Figure 6. (a) and (b). (See caption on facing page.)

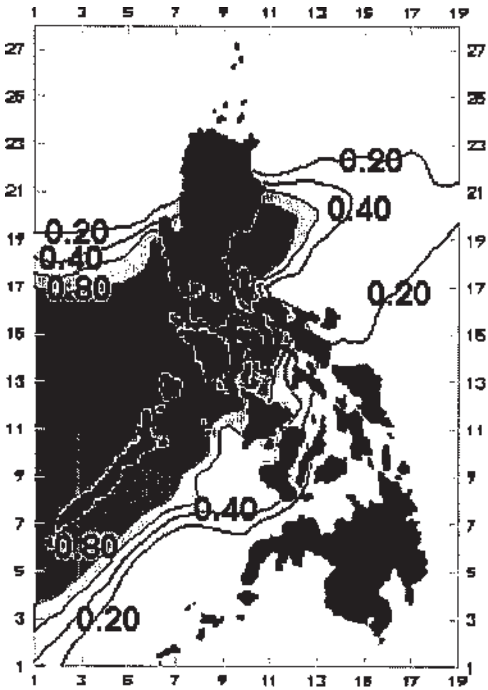


Figure 6. Dominant data modes: (a) (facing page top) the first mode (62.64% variance) showing peak values at Regions 3 and 5, (b) (facing page bottom) second mode (35.42% variance) showing a large area of high values at Regions 4 and 6, and (c) (above) the sum of the first and second modes (98.06% variance). (Axes labels indicate gridding of the domain using 19 × 28 cells.)

ago cannot be suspected as depleted of stocks. Second, it follows that some unidentified biomass influxes must be coming in from somewhere to sustain the increasing trend in catch. Precisely, this existence of biomass exchange, in conjunction with biomass "sources" or "sinks," is the main thesis of this paper.

The dominant modes of the data resulting from modal decomposition are plotted in Fig. 6. The first mode, capturing 62.64% of the data variance, characterizes the eight-year catch record as due mainly to relatively high catch in Regions 3 and 5 (Fig. 6a). The second mode, capturing some 35.42% of the variance, is due to relatively high catch in Regions 4 and 6 (Fig. 6b). The time coefficients of the decomposition show that these modes had opposing tendencies; the tendency of the first mode decreased in 1972, while that of the second mode increased (Fig. 7a). By 1973, both modes settled more or less to a new equilibrium. These results indicate a

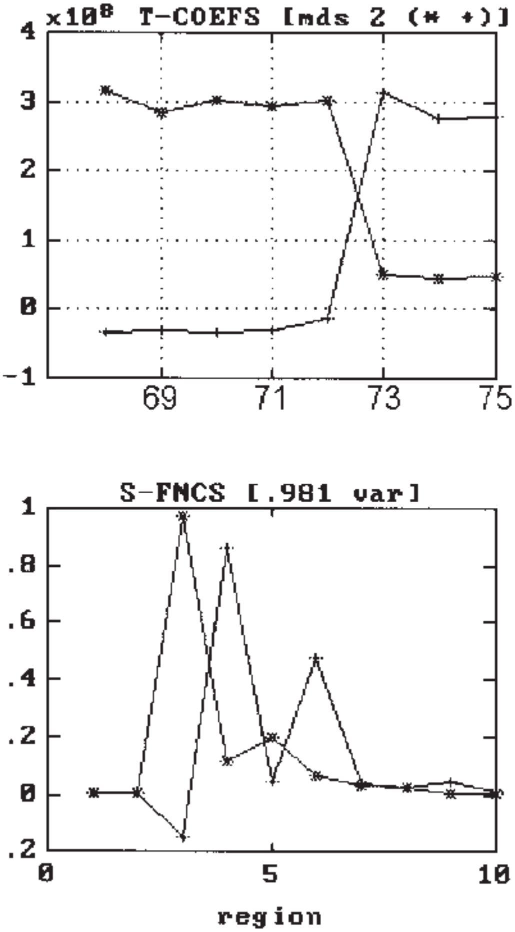


Figure 7. Plots of (a) (top) time coefficients and (b) (bottom) the spatial functions. Catch or fish production for a given area and time is the product of the time coefficient and space function for each of the eigenmodes.

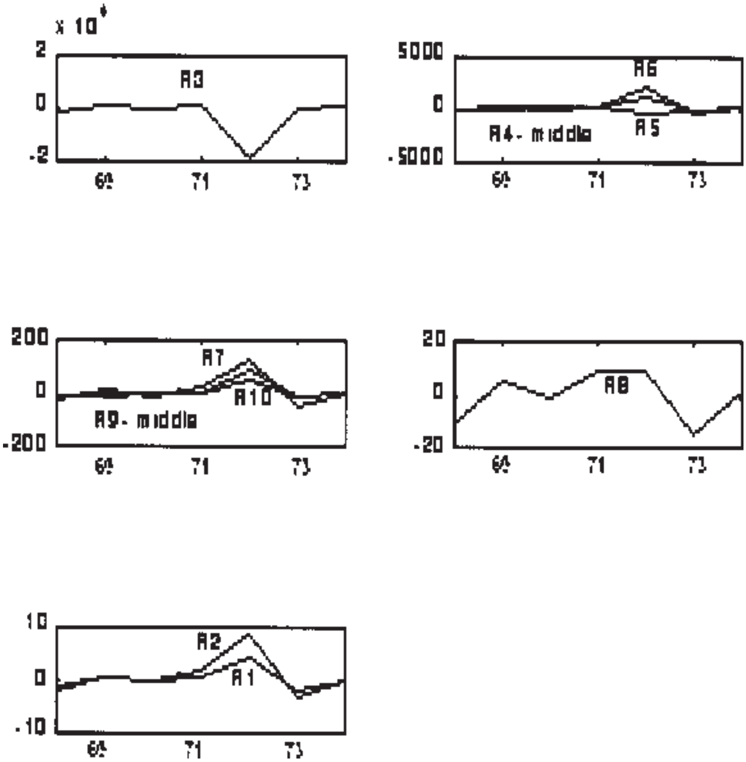


Figure 8. Time series of divergence for the fishing regions, clustered according to decreasing magnitude of variations. Notice that the process of differentiation caused a loss of one data point for 1975.

clear shift in concentration of catch from Regions 3 and 5 (Manila Bay area, San Miguel and Lagonoy gulfs) to Regions 4 and 6 (mainly Palawan and west-central Visayan waters) which are now a relatively wide coverage. In particular, a closer look at the point values of the space function in Fig. 7b reveals the large swing of catch at Region 3 which was synchronously positive with the rest of the regions for the first mode, but asynchronously negative (almost with Region 5) for the second mode. Notice that, in support of the above findings, the overall mode shows a preponderance of positive catch for the given period of observations (Fig. 6c).

Computed divergence for the regions is shown in Fig. 8. The plots reveal the overwhelming magnitude at Region 3 (about an order of magnitude), where the largest loss of biomass in 1972 occurred. (Manila Bay area, representing Region 3, is a well-known "overexploited" fishing ground.)

Basically, there is no indication of convergence in the region during the entire period. The variation is mirrored almost at Region 5 but only at extremely negligible magnitude. Region 8 exhibits a peculiar variation of divergence during 1968, 1970, and 1973. (Region 8 is at the terminus of the North Equatorial Current or NEC.) The rest of the regions show convergence during 1972, with Regions 4 and 6 having relatively higher magnitudes. Overall, it appears that the biggest variations happened in 1972, resulting to a net divergence for the archipelago mainly at Region 3. This is a possible aftermath of the heavy catch in this region as evident in the original data. (Landing bias could be a real cause of error here but, for convenience, this is presumed as one of the limitations of the data, as discussed in the following section.)

Another outcome of modal decomposition for the entire period from 1968 to 1975 is the dominant mode of divergence shown in Fig. 9a. Here the use of quality (processed) observations paid off in reducing error variance. The plot confirms that for the given period, Regions 3 and 5 had acted asynchronously with the rest of the regions as a center of divergence. Therefore, these active regions had acted as “sinks” of biomass which, according to equation (1), could be linked to convergence or biomass influx from other regions of the archipelago. Indeed, a schematic of this influx is shown in Fig. 9b. Notice that some pathways were so drawn to conform to the complex topography. It is a simple matter to verify the detailed origin and pathways of the fluxes using equation (2), especially with available boundary flux information (flux across a coastline is assumed nil). This task is considered beyond the scope of the present study.

It is emphasized that, in general, analysis of time series is subject to resolution restrictions based on its finite length. In the present case, the findings were derived from a record length of only eight years from 1968 to 1975. The addition of old or new data may influence these findings.

Discussion

The modeling introduced here focused on the dynamic concept involving biomass redistribution centers as exemplified by divergence and the associated horizontal fluxes. The model is simple, utilizes only one piece of information (catch data), and is devoid of excessive parameterization. The use of modal decomposition to control the quality of the input data proved crucial to the results. Further, because the well-established mass conservation law is utilized, the results, interpreted within the context of the model, are likely robust.

The coherent variations exhibited by the various regions of the Philippine archipelago suggest that biomass exchanges did occur simply because these regions share common boundaries. Fluctuations at the redistribution centers, resulting in either divergence or convergence, could then be linked to the fluxes according to equation (1). Further, the variations of regional divergence showed that this linkage was time-dependent.

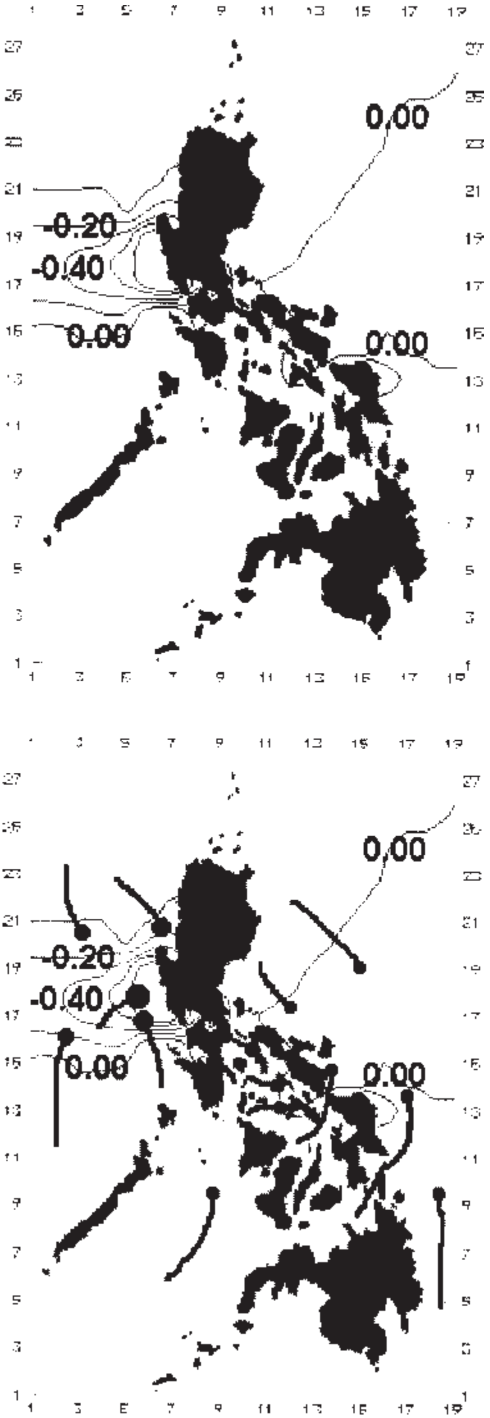


Figure 9. (a) (top) The dominant surface divergence mode (99.90% variance) showing Regions 3 and 5 as centers of biomass redistribution, particularly as centers of strong divergence. Computations utilized the quality data of Fig. 3. (b) (bottom) A schematic of the horizontal fluxes deduced from the distribution of divergence. Size of the circles reflects the relative strength of flux. Some pathways are shown to conform to topography.

Generalizing, it can be the case, therefore, that certain regions of the oceanic environment act as either "sinks" or "sources" during the process of biomass redistribution, and that we might have unknown biomass fluxes that continuously link these regions. This dynamic viewpoint appears fundamental and suggests that the existence of biomass fluxes or "migration" could be of central significance to modern stock analysis.

Indeed, the existence of fluxes could complicate currently upheld assumptions and may lead to intriguing queries. For instance, can Regions 3 and 5 be labeled as depleted or overexploited? The results of this study indicate that these regions appeared divergent relative only to the rest of the archipelago and because biomass were being removed from these regions more excessively than the rest of the archipelago which, in turn, served as a "source" of biomass. Hence, divergence at Regions 3 and 5, in fact, was being counteracted by influxes. How then can one define sustainable yield for these regions in the presence of changing fluxes? Further, how does managing these regions fare if one insists on strategies based on the usual concept of sustainable yield?

Therefore, it is conceivable that managerial indices like overexploitation and maximum sustainable yield (MSY) may have to be redefined as dynamical quantities because, unless the biomass fluxes are known constantly, these indices will remain largely arbitrary. For example, if one considers a region as overexploited (under-exploited) only in the sense that it continually imports (exports) stock as suggested by its biomass influx (outflux) then, sustainable yield may be re-defined to be proportional to some threshold or "equilibrium" value of the flux gradient:

$$MSY \propto \boldsymbol{P}(\underline{x},t)_{max} \tag{4}$$

Moreover, absolute biomass is unknown generally and deriving reliable information on its variation may require integration of a suitably large area. Hence, the complications due to biomass fluxes imply also that management indices are truly meaningful only if an integral area is considered because one cannot isolate for analysis a small area that interacts dynamically with neighboring areas through its boundaries. To be precise, how can one declare an open area as overexploited if exchanges across its boundaries are unknown? In order to drive home this point, it is tempting to relate an allegory at this juncture.

One can imagine a huge livestock population of some ten million chickens confined (squeezed) in an area of 16 km^2 (you have roughly one chicken per 2 m^2). (The ocean is not really this small, but this suffices for an illustration.) Suppose a hunter subsists on this livestock, consuming a maximum of three chickens per day, and is able to hunt within his perimeter of only 0.25 km^2. We make sure that the hunter does not consume all the chickens by limiting his hunting to a finite duration of, say 10 weeks (by then he would not have consumed more than a miniscule number of 210 chickens). It is then reasonable to expect that, not long after he begins shooting down his prey, the hunter will complain of a diminishing number of chickens within his perimeter. He will complain also of finding more of the young than the adult. (Obviously, mature animals respond quicker to gun shots than young animals.) Eventually, he will oblige himself to conclude that his place no longer has chickens to hunt—a very similar and familiar complaint heard about our fishing grounds. To finish, unknown to the hunter, when he is fast asleep plenty of adults pass by his open perimeter, possibly in search for their own food.

In Fig. 9b, the fluxes indicate a general northward and southward "migration," converging biomass towards the vicinity of Regions 3 and 5. The association of these fluxes with the mean and seasonal circulation patterns in the archipelago is another interesting subject to explore, perhaps in relation to migrating species. Indeed, although equation (1) was exemplified as a bulk biomass formula, nothing precludes a finer analysis with an available catch composition data. In general, individual, group, or bulk analysis of species is plausible with MCM. For instance, the model is at the same time useable for multispecies (or multilayer/depth) analysis, if each term in the mass balance is generalized to mean as the sum of its respective components, i.e.,

$$\Sigma_i \mathrm{d}_t B = \Sigma_i (F_{in} - F_{out});\; i = 1,2,3,\ldots,n \text{ species (or layer/depth)} \qquad (5)$$

Knowledge on the vertical distribution or migration of species may be inferred if there is ample data as a function of depth. Virtual population analysis (VPA) or intricate models which attempt to quantify the biomass/energy flow in an ecosystem via diagnostic or prognostic (simulation) means (e.g., Christensen and Pauly 1992 and 1996) might derive insights from varied applications of MCM. In purely prognostic mode, MCM may form a component of existing fishery simulation models (e.g., King 1995). Another interesting possibility is the coupling with a water circulation model to explore fully the time-dependent biophysical dynamics of biomass distribution. Further, MCM is capable potentially of producing reliable mass balances which are essential to the successful management of resources involving a wide area. An example of a national management scheme is the so-called Fisheries Sector Program (FSP), which attempts to implement complex and comprehensive management plans throughout the Philippines. At a staggering cost of US$155 million, borrowed from the Asian Development Bank (ADB) and the Overseas Economic Cooperation Fund (OECF), the program cannot possibly afford to risk implementing management policies devoid of integral information on the archipelago (DA 1993).

On the other hand, certain caveats are in order. One must be careful in interpreting the seemingly mechanical fashion that the horizontal fluxes are associated to the centers of redistribution. One way to imagine this dynamics is to treat the latter as some "forcing function" which generates (or responds to) the potential fluxes. This potential exists simply because it fulfills the conservation law stated in equation (1). Thus, any discrepancy found between computed fluxes and the actual movement of biomass (if ever this can be measured accurately) must not be taken to mean that it nullifies the dynamics. Discrepancy can be the result of a temporal lag in the interaction between redistribution centers and fluxes. Unfortunately, information on lag may only be determined with adequately sampled data, as it is largely pegged also on one complex factor—organismal behavior. There is no means to ascertain lag in the present data which uses a coarse interval of a year. However, it is conceivable that at this and suitably longer temporal scales, significant biomass adjustments (e.g., migration cycles) are allowed to minimize lags. Certainly, one other real cause of discrepancy is poor data resolution itself. Inadequate data, at the same time, limit the temporal as well as spatial scales of analysis. Moreover, the gradient in equation (2) is only a means of reckoning the magnitude of the forcing potential, indicating the tendencies for biomass movement. Again, these tendencies exist only because of the "forcing" due to the dynamic centers. It is essential that these centers are not misunderstood necessarily as places where you have catch data but rather those determined by the data.

Certain factors neglected in MCM require some clarifications. First, insignificant surface fluxes like the addition of biomass by artificial restocking are ignored relative to catch. Second, growth, biological recruitment, and mortality are processes considered as internal to the system. They are not boundary fluxes by definition and thus cannot be incorporated easily in the calculations. However, it is pointed out that any imbalance among these factors may result in a biomass flux as suggested also by the following mass balance formulation derived from Christensen and Pauly (1993):

$$\text{Export} = \text{Production} - \text{Mortalities} \tag{6}$$

In this equation, the delicate balance between production and mortalities (predation and nonpredatory) suggests that export/import (or the flux) in a given system is bound to happen and, when it does, it has the potential to induce biomass changes. These changes are transparent to the MCM and hard to measure because both estimates of production and mortalities are unreliable. Further, this implies also that the difference between these estimates could be erratic. These are the same reasons why the initial development of MCM must exclude these factors. As a new model, it is desired that MCM's consistency be verified as data becomes more available prior to modifying its dynamical properties. Third, it must be clarified that MCM is not concerned with the absolute measure of but rather the increase or decrease of biomass with time. (In the first place, biomass is not straightforward to calculate from equation [1].) For example, a huge biomass does not necessarily imply strong divergence because changes in its size can be small. Finally, one must be aware of the serious limitations of MCM: scarce and questionable reliability and accuracy of reported catch data, inadequate sampling of the water column and, as pointed out above, the presence of yet unexplained internal biomass redistribution. These are interesting subjects for future publication.

The human role in biomass redistribution, i.e., through fishing activities, is much easier to quantify than the complex role of nature's forcing functions. For instance, one may imagine the extreme difficulty of modeling changes in the environment coupled with responses from both the fishery and consumers. Therefore, one may appreciate MCM as a preliminary attempt at dynamic biomass modeling. The concept of center of biomass redistribution is potentially useful in defining an indicator or predictor of biomass concentration in an area. The model likewise establishes the relevance of an even more interesting (but difficult to observe) quantity—the biomass flux. Indeed, clearcut dynamics, as stated by equation (1), is easy to imagine and is intuitively appealing; it brings one closer to a quantitative grasp of biomass movement without having to observe directly that movement. More than that, the direction of flow, origin, and fate of biomass can be ascertained given only an ample spatial coverage of the catch data.

Ideas in this paper are developing and, indeed, some may not appeal at all to a fishery scientist. However, with the current deplorable state of fishery management strategies, one has no brighter option than to devise alternatives (cf. Maclean1996). Perhaps, having known the mass conservation as a robust natural principle, it may not be surprising to discover that the relevant ideas may work in the real world.

Acknowledgments

Many reading materials on stock assessment were handed down to me by Rommel Lao. Dr. C. Zhang and two anonymous reviewers suggested helpful revisions to the manuscript. I thank Eric Dalisay for working patiently on some figures. Most especially, I am grateful to my wife, Rochelle, who typed and laid out a seemingly endless stream of manuscripts.

References

Beverton, R., and S. Holt. 1957. On the dynamics of exploited fish populations. Fish. Invest. Ser. II. 19. 533 pp.

Bureau of Fisheries and Aquatic Resources (BFAR). 1972. Fishing statistics of the Philippines. Manila, Philippines. 104 pp.

Bureau of Fisheries and Aquatic Resources (BFAR). 1973. Fishing statistics of the Philippines. Manila, Philippines. 100 pp.

Bureau of Fisheries and Aquatic Resources (BFAR). 1975. Fishing statistics of the Philippines. Manila, Philippines. 115 pp.

Catell, R. (ed.) 1966. The meaning and strategic use of factor analysis. In: Handbook of multivariate experimental psychology. Rand McNally, Chicago.

Christensen, V., and D. Pauly. 1992. ECOPATH II: A software for balancing steady-state ecosystem models and calculating network characteristics. Ecol. Model. 61:169-185.

Christensen, V., and D. Pauly. 1993. Trophic models of aquatic ecosystems. ICLARM Conf. Proc. 26. 390 pp.

Christensen, V., and D. Pauly. 1996. Ecological modeling for all. Naga, ICLARM Q. 19(2):25-26.

Department of Agriculture. 1993. Fisheries Sector Program. Published by the Fisheries Sector Program, Program Management Office (FSP-PMO) Information Division, Quezon City, Philippines.

Gulland, J. 1969. Manual of methods for fish stock assessment. Part 1. Fish population analysis. FAO, Rome, Italy. 154 pp.

King, M. 1995. FISHIM: A generalized fishery simulation model. Naga, ICLARM Q. 18(4):34-37.

Kutzbach, J. 1967. Empirical eigenvectors of sea level pressure, surface temperature and precipitation complexes over North America. J. Appl. Meteor. 6:791-802.

Levitus, S. 1982. Climatological atlas of the World Ocean. NOAA Professional Paper 13. 173 pp.

Maclean, J. (ed.) 1996. Fisheries management. Naga, ICLARM Q. 19(3). 71 pp.

Schaefer, M. 1954. Some aspects of the dynamics of populations important to the management of the commercial marine fisheries. Inter-Am. Trop. Tuna Comm. Bull. 1:27-56.

Sparre, P. 1985. Introduction to tropical fish stock assessment. FAO/DANIDA Project Training in Fish Stock Assessment, GCP/INT/392/DEN. 338 pp.

Weare, B. 1976. Empirical orthogonal analysis of Pacific sea surface temperature. J. Phys. Oceanogr. 6:671-678.

Wenzel, J. 1984. Empirical orthogonal function analysis of sea level fluctuations in the tropical Pacific ocean. M.S. thesis, University of Hawaii, Honolulu, HI. 70 pp.

Multivariate Interdisciplinary Assessment of Small-Scale Tropical Fisheries

David Preikshot and Daniel Pauly
University of British Columbia, Fisheries Centre, Vancouver, B.C., Canada

Abstract

This paper presents a multidisciplinary approach for the comparative evaluation of small-scale tropical fisheries. Previously, these most often have been studied within single disciplines of the social or natural sciences, rarely leading to broad comparisons. Here, multidimensional scaling, canonical correlation, and cluster analysis were used to examine and contrast the information contained in multivariate data sets originating from the disciplines of biology, economics, social science, and technology of seventeen real tropical small-scale fisheries and two simulated ones. Multivariate analysis was performed within the four disciplines and the resultant outputs were compared and contrasted. The relative location of the fisheries in these derived scores suggests relations among social, economic and biologic factors.

Introduction

One of the most debated topics in fisheries science has been the incorporation of multidisciplinary information into what has chiefly been a biological discipline. With recent developments it is now possible, even with modest budgets, to purchase a computer that can store and synthesize gigabytes of data. Multivariate statistics provides the crucial tool, though, in accessing multidisciplinary studies. That potential is explored here.

There can be little doubt that there is presently a crisis in the management of many of the world's fisheries. This has been brought about by phenomena such as subsidization of overcapacity in fishing fleets (Pauly 1997), fish habitat destruction (Hagler 1995), unregulated bycatch (Hagler 1995 and Pitcher 1993), overly optimistic predictions by managers (Pauly 1996, Christensen and Pauly 1995, and Walters 1995), and institutional inability to deal with the complex issues created by the conflicts between players in fisheries policy development (McGoodwin 1990).

These problems are aggravated in the case of tropical small-scale fisheries, wherein the resources are being depleted, while there is an increase in the number of participants, a phenomenon now called “Malthusian overfishing” (Pauly 1997). This situation is catalyzed by the marginality of the participants in the fishery relative to the apparatus of political decision-making. This marginality is increased by a number of factors: isolation from power, geographic and cognitive remoteness, and illiteracy (Pauly 1997). The mechanism leading to the overexploitation of the fishery is an increase in the number of fishers resulting from natural increase and the immigration, into coastal areas, of outsiders displaced, by landlessness, from traditional activities such as farming. These outsiders often turn to fishing as a vocation, as it provides them with the potential of earning some money, or at least the capacity to get food (Pauly 1997). The introduction of these new entrants to the fisheries often upsets established local traditional management practices. This may be facilitated by the creation and imposition of new fisheries rules by central governments all too ready to find a place to settle the migrants (Pauly 1997). Blaikie (1985) writes of similar problems with terrestrial environments in developing countries.

By contrasting factors from ecological, economic, social science, and technological attribute sets, it may be possible to compare what the data sets, in particular and in combination, say about a fishery. This knowledge would be of great value to fisheries researchers since it would enable the rapid identification of “at risk” fisheries if based on easy to identify attributes. Further, if we can determine that information from one discipline reinforces information from another, it may be unnecessary to always initiate expensive research in all fisheries science disciplines at all times, benefiting developing countries. Last, by identifying which fishery in an area or country is at risk, it may be easier for governmental and non-governmental organizations to decide where aid may be most needed.

Methods

Attribute data were collected for seventeen real, and created for two model, small-scale tropical fisheries in four main categories: ecology, economics, sociology, and technology. See Table 1 for the attributes and their definitions. Attributes for the four groups were chosen such that they would describe as much distinct information as possible in each subject. The hypothetical fisheries A and B were scored such as to reflect our preconception of Malthusian overfishing effects through early, young, mature, and old stages. These were scored, with declining relative and absolute economic standards, collapsing social structures, and decreasing use of selective gears. The real fisheries that were examined were scored from available literature or based on interviews with fisheries scientists who had studied the fishery. The score table for this study is available from the authors via the U.B.C. Fisheries Centre WWW page: http://fisheries.com. The fisheries that were analyzed and their codes for graphing can be seen in Table 2.

The four groups were subjected to multidimensional scaling (MDS) analysis with the Statistical Package for the Social Sciences (SPSS 1995) to extract two principal coordinates. The distance matrix for scaling was of the Euclidean distance type and measured scores as intervals. To construct the distance matrix and make all variables of the same magnitude all data were transformed to a range from zero to one. This implied equality in the potential contribution of all the attributes to the final ordinations. In all of the MDS ordinations the axes were flipped where necessary to reflect the association of positive characteristics on the *x* axes on the left and the *y* axes to the top. Thus for all ordinations fisheries with favorable characteristics should tend to the upper left quadrant, while fisheries with less favorable characteristics should tend to the lower right quadrant.

After the first two principle coordinates were generated canonical correlation (CC) was used in the Statistica program (Statsoft 1996) to analyze relationships between the first two principal coordinates of each group and the data set from which they were derived. This allowed an investigation of the data most influential to each axis (Stalans 1995). Last, cluster analysis (CA), in the Statistica program (Statsoft 1996) was performed with the first two principal coordinates for each attribute group to help in the classification of the fisheries.

Results

The MDS ordinations produced in this analysis can be seen in Figs. 1, 2, 3, and 4. MDS produced distinct groupings for the fisheries in all four disciplines. Also, the fisheries stayed close to each other in the different subject groups, although these associations were weaker in the economic analysis. Statistics used with SPSS to analyze the derived MDS space were Young's S-stress formula 1, Kruskal's stress formula 1, and squared correlation (RSQ). For the Ecological MDS the S-stress after two iterations was 0.27733. A third iteration, producing a third dimension, was not deemed necessary since the S-stress improvement was less than 0.01% of the second iteration's S-stress value. Indeed, for all of the derived MDS ordinations two iterations were used since in none of them did a third iteration yield any better than a 2% improvement on the second iteration's S-stress. For the economic, sociological, and technological MDS ordinations, S-stress after two iterations was 0.32857, 0.38416, and 0.29281 respectively. Kruskal's stress formula 1 yielded values of 0.22219, 0.24091, 0.29693, and 0.23723 respectively for the ecological, economic, sociological, and

Table 1. Disciplines and attributes used in the multivariate analyses.

	Attribute	Ecological scoring	Notes
1	Catch/fisher	Metric tons	Person/year
2	Exploitation status	0, 1, 2	FAO scale, low, full, over
3	Trophic level	Number	Average trophic level of species in catch
4	Migratory range	0, 1, 2	1, 2-3, >3 jurisdictions encountered during migration
5	Catch < maturity	0, 1, 2	None, some, lots caught before maturity
6	Discarded bycatch	0, 1, 2	Low 0-10%, med 10-40%, hi >40% of target catch
7	Species caught	0, 1, 2	Low 1-10, med 10-100, hi >100 species
8	Primary production	0, 1, 2	$gCm^{-2}y^{-1}$, low = 0-50, med = 50-150, high = 150+

	Attribute	Economic scoring	Notes
1	Price	US$/t	US$/t of landed product for analysis time
2	Fisheries in GNP	0, 1, 2	Importance of fisheries sector in country: low, med, high
3	GNP/person	US$/capita	In country of fishery
4	Limited entry	0, 1, 2	Almost none, some, most (includes informal limitation)
5	Other income	0, 1, 2	Mainly casual, part-time, full-time fishers
6	Earnings by fishers	0, 1, 2	Below, same, above national average for workers
7	Market	0, 1, 2	Principally local, national, international

	Attribute	Sociological scoring	Notes
1	Fishing socialization	0, 1, 2	Fishing by individuals, families, or community groups
2	Fishing comm. growth	0, 1, 2	Over past 10 years: <10%, 10%-20%, >20%
3	Fisher sector	0, 1, 2	Community households fishing: <⅓, ⅓-⅔, >⅔
4	Education level	0, 1, 2	Below, same, above population average
5	Conflict status	0, 1, 2	Level of conflict with other sectors
6	Information sharing	0, 1, 2	None, some, lots
7	Fisher influence	0, 1, 2	Strength of fisher direct influence on fishery regulations
8	Fishing income	0, 1, 2	Family income from fishing: <50%, 50-80%, >80%
9	Kin participation	0,1	Do kin sell family catch and/or process fish?: no or yes

Table 1. (Continued.)

	Attribute	Technological scoring	Notes
1	Trip length	Days	Average days at sea per fishing trip
2	Landing sites	0, 1, 2	Dispersed, some centralization, heavily centralized
3	Processing	0, 1, 2	None, some, lots of gutting etc. before sale
4	Use of ice	0, 1, 2	None, some, lots
5	Gear	0, 1	Passive = 0, active = 1
6	Mesh	0, 1	Net meshes not in gear = 0, net mesh in gear = 1
7	Selective gear	0, 1, 2	Device(s) in gear to increase selectivity: few, some, lots
8	FADS	0, 1	Fish aggregation devices not used = 0, are used = 1

technological data sets. This is somewhat larger than what Stalans (1995) terms a "good fit" but acceptable. RSQ values for the four distance matrices were 0.76676, 0.72863, 0.55201, and 0.76798. RSQ values measure the proportion of variance of the scaled data in the distance matrix accounted for by their corresponding distances (SPSS 1995).

CC was used to help describe what general properties were implied by the dimensions produced by the MDS, see Table 3. Each attribute's association with either axis was proportional to the correlation (Stalans 1995). When the correlation was positive a high score in an attribute tends to give a high score on the MDS dimension. A negative correlation implied high scores for an attribute resulting in low scores on the associated MDS axis.

The final statistical procedure used on the data was CA to produce objective mathematical groupings, offsetting the tendency of the human mind to impose preconceived ones. Groupings for all cases analyzed were done for all four attribute sets. The groupings were determined by the complete distance Euclidean distance rule. This technique promotes "clumpiness" since distances between clusters are determined by the distances of their furthest neighbors (Statsoft 1995.) As Cooper and Weekes (1983) point out, there is much controversy over what constitutes a cluster and there are no rules for their identification. Therefore, it was decided to use CA to identify the first three or four groups created. This was done to help in the assessment aspect of the analysis by creating a manageable number of groups.

Table 2. Fisheries and codes used for graphing in the analyses.

Fishery	Graph name
Belize small scale fisheries, 1996	Belize 1996
Bolinao Reef, Philippines, 1992	Bolinao 1992
Diani-Kinondo reef, Kenya, 1995	Diani-Kinondo 1995
Federated States of Micronesia, *Trochus,* 1995	FSM 1995
Lake Kariba, Zambia, 1995	Kariba, Zam 1995
Lake Kariba, Zimbabwe, 1995	Kariba, Zim 1995
Lake Kivu, Rwanda, 1993	L. Kivu 1993
Lake Malombe, Malawi, 1993	L. Malombe 1993
Lake Victoria, Kenya, 1985	L. Vic Ken 1985
Lake Victoria, Tanzania, 1985	L. Vic Tan 1985
Lake Victoria, Uganda, 1985	L. Vic Ugan 1985
Model “A”	A early, young, mature, old
Model “B”	B early, young, mature, old
Sakumo Lagoon, Ghana, 1971	Sakumo 1971
Sakumo Lagoon, Ghana, 1994	Sakumo 1994
SE Lake Malawi, 1993	SE L. Malawi 1993
Ubolratana Reservoir, Thailand, 1984	Ubolratana 1984
Zanzibar, coral reef demersals, 1985	Zanzibar 1985
Zanzibar, coral reef demersals, 1995	Zanzibar 1995

Discussion

In the ecological component of the analysis the first, i.e. x, axis was most influenced by characteristics associated with the size and amount of fish caught, trophic level, migratory range, and catch per fisher. The second, or y, axis was most correlated with exploitation status, catch before maturity, and discarded bycatch, properties that described the overall state of an ecosystem. Therefore, in ecologically assessing these fisheries the “best” should be those in which relatively large fish are caught with little impact on other parts of the ecosystem. Indeed, the young and early stages of fishery A, modeled as a “pre-Malthusian” fishery, are seen in the upper left corner of Fig. 1. The only real world fishery in the upper left quadrant is Lake Kivu in Rwanda. Significantly, attributes for this fishery were collected before the civil war of 1994 and it is also the youngest analyzed, having only been established in 1980, after the introduction of the sardine

Limnothrissa miodon to the lake two decades earlier (de Iongh et al. 1995). The fisheries located in the lower right corner of the ordination would appear to be more problematic since they would be associated with poor environmental conditions and small catches of small fish.

The fisheries of the other two quadrants seem to be representative, then, of two types of intermediate fisheries: those with low environmental quality yet otherwise maintaining fair fishing characteristics (lower left), and those with diminished fish quality within a relatively more stable ecosystem (upper right). It should be noted that the latter group was almost entirely made up of modeled fisheries, except for the *Trochus* harvest in the Federated States of Micronesia (FSM), and the small-scale fisheries of Belize. Both derive all or most of their value from the harvesting of a few valuable invertebrates (Clarke and Ianelli 1995; pers. comm., Gillett, U.B.C. Fisheries Centre, 2204 Main Mall, Vancouver, B.C. V6T 1Z4, 1997). Thus, small-scale tropical fisheries appear to suffer environmental declines followed by declines in quality of the fish caught. This ecological "trajectory" is supported by the fact that tropical ecosystems are often species rich and the targeted organism of the associated fishery can be switched if the original species is/are exhausted. Four groups were identified by CA, and we can describe them as: upper left "favorable," upper right "fishery decline," lower left "environmental decline," and lower right "unfavorable."

In the economic ordination fisheries in the upper left quadrant were associated with localized markets, low GNP per capita yet relatively high incomes, and closed access on the *x* axis, and with high value fish and fisheries as important constituents of the national economy on the *y* axis. Note how the Belize and FSM fisheries again are segregated from the other real fisheries. Note too that of the thirteen real fisheries in the two ecological groups exhibiting less favorable characteristics, eight are found in the least favorable economic groups.

In the sociological ordination the left of the *x* axis correlated with fishing in community groups, high degrees of information sharing, high fisher influence on regulations, and lots of kin participation. The upper part of the *y* axis was also associated with lots of kin participation, but also with fishers as a large proportion of the community. If the two groups in the social ordination are considered to be at the bottom and right side of the graph then it is significant that of the thirteen least favorable ecological fisheries, nine are in the two least favorable sociological groups.

The technological ordination had the left side of the *x* axis associated with decentralization of landing sites, use of net meshes in fishing gear, use of selective gear, and absence of fish aggregating devices. The upper part of the *y* axis was correlated with lots of processing and use of ice by fishers. Of the three main groups identified by using CA, the group entirely to the right of the *y* axis appears to be least favorable. Of the thirteen least favorable fisheries from the ecological analysis, nine were found in the least favorable technological grouping.

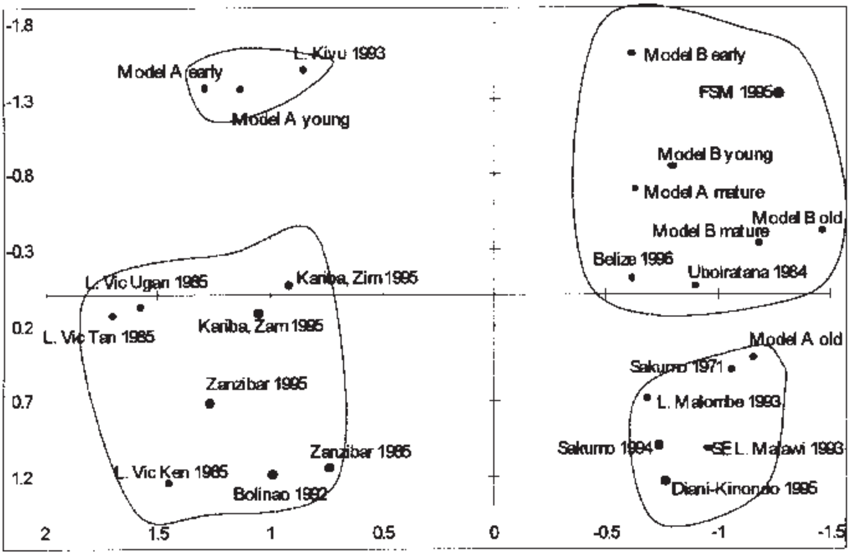


Figure 1. Ecological MDS ordination of small-scale fisheries of developing nations. Fisheries tending to the upper left are characterized by more favorable characteristic, while those to the lower right typically have less favorable characteristics.

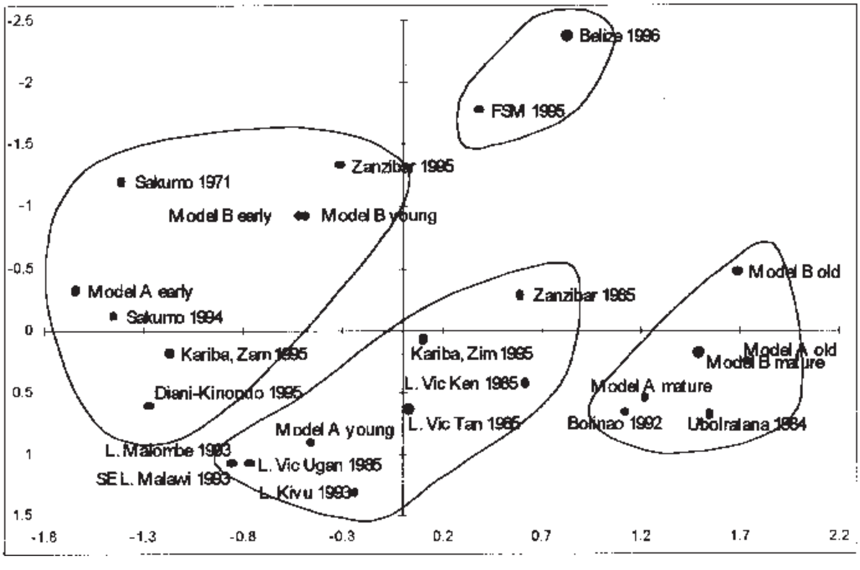


Figure 2. Economic MDS ordination, with favorable fisheries to the upper left, less favorable to the lower right.

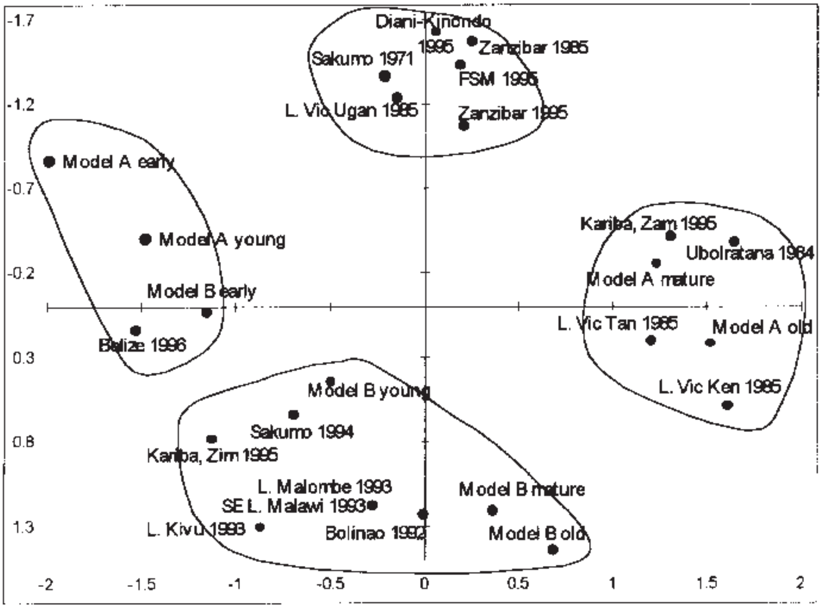


Figure 3. Sociological MDS ordination, with favorable fisheries to the upper left, less favorable to the lower right.

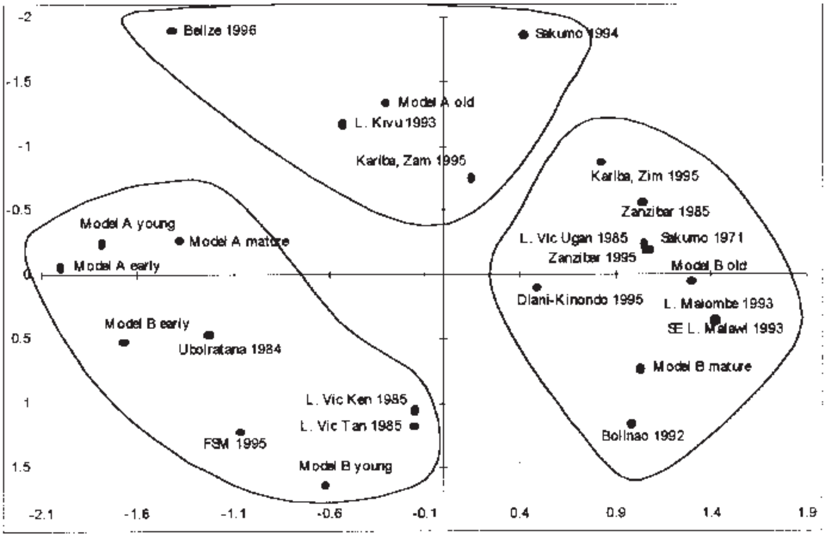


Figure 4. Technological MDS ordination, with favorable fisheries to the upper left, less favorable to the lower right.

Table 3. Correlations of original attributes to associated MDS dimensions 1 and 2.

Ecological	Dim. 1	Dim. 2	Economic	Dim. 1	Dim. 2
Catch/fisher	0.55	0.36	Price	0.16	–0.55
Exploitation status	–0.42	0.74	Fisheries in GNP	–0.05	–0.80
Trophic level	0.77	–0.05	GNP/person	0.58	0.02
Migratory range	0.97	0.04	Limited entry	–0.64	–0.26
Catch < maturity	–0.10	0.90	Other income	–0.48	0.58
Discarded bycatch	–0.33	–0.62	Earnings by fishers	–0.69	–0.27
Species caught	0.22	0.16	Market	0.72	–0.48
Primary production	0.28	0.36			

Sociological	Dim. 1	Dim. 2	Technological	Dim. 1	Dim. 2
Socialization of fishing	–0.63	–0.23	Trip length	–0.02	–0.46
Fishing comm. growth	0.28	0.28	Landing sites	–0.61	0.02
Fisher sector	–0.11	–0.60	Processing	–0.45	–0.62
Education level	–0.18	–0.26	Use of ice	–0.08	–0.57
Conflict status	0.30	0.09	Gear	0.51	0.43
Information sharing	–0.59	–0.48	Mesh	0.76	–0.52
Fisher influence	–0.62	–0.28	Selective gear	–0.70	0.26
Fishing income	–0.48	–0.39	FADS	–0.80	–0.35
Kin participation	–0.70	0.68			

The large degree of overlap among the least favorable fisheries in the attribute groups analyzed here indicates a linkage between them. This linkage is consistent with the Malthusian overfishing hypothesis (Pauly 1997). The relations between the ordinations of the fisheries in the four attribute groups would be improved if the sample size were doubled and time series data were included. A wider analysis would enable better identification of pathological tendencies in small scale fisheries of the developing world, and perhaps warn of problems.

Acknowledgments

The authors would like to thank Trevor Hutton, Tony Pitcher, and Alida Bundy for helping develop the methodology of this study. We would also like to express our gratitude to Vincent Gillett, Narriman Jiddawi, and the many researchers who provided data for this study. Last, our thanks to Dr. Chang Ik Zhang and two anonymous referees for their helpful comments.

References

Blaikie, P. 1985. The political economy of soil erosion. Longman Scientific and Technical, London. 188 pp.

Christensen, V., and D. Pauly. 1995. Fish production, catches and the carrying capacity of the world oceans. ICLARM Quarterly 18(3):34-40.

Clarke, R.P., and J.N. Ianelli. 1995. Current paradigms in *Trochus* management and opportunities to broaden perspectives. In: P. Dalzell and T.J.H. Adams (eds.), South Pacific Commission and Forum Fisheries Agency Workshop on the Management of South Pacific Inshore Fisheries. Manuscript collection of country statements and background papers, Vol. 1. South Pacific Commission, Noumea.

Cooper, R.A., and A.J. Weekes. 1983. Multivariate analysis, Section 10. In: R.A. Cooper and A.J. Weekes. Data, models and statistical analysis. Philip Allan Pubs. Inc., Deddington, Oxford, pp. 240-259.

de Iongh, H.H., P.C. Spliethoff, and F. Roest. 1995. The impact of an introduction of sardine into Lake Kivu. In: T.J. Pitcher and P.J.B. Hart (eds.), The impact of species changes in African lakes. Chapman and Hall, London, pp. 277-298.

Hagler, M. 1995. Deforestation of the deep: Fishing and the state of the oceans. The Ecologist 25(2/3):74-79.

McGoodwin, J.R. 1990. Crisis in the world's fisheries: People, problems, and policies. Stanford Univ. Press, Stanford. 235 pp.

Pauly, D. 1996. One hundred million tonnes of fish and fisheries research. Fish. Res. 25:25-38.

Pauly, D. 1997. Small-scale fisheries in the tropics: Marginality, marginalization, and some implications for fisheries management. In: E.K. Pikitch, D.D. Huppert, and M.P. Sissenwine (eds.), Global trends: Fisheries management. Amer. Fish. Soc. Symp. 20.

Pitcher, A.J. 1993. Stewardship and sustainability of Pacific fishery resources: The need for critical insight and an encyclopedia of ignorance. In: Our common shores and our common challenge: Environmental protection in the Pacific. Proceedings of the Fourth International Symposium of the Conference of Asian and Pan-Pacific University Presidents. University of Alaska Sea Grant, Fairbanks. 136 pp.

SPSS Inc. 1995. SPSS for Windows, release 7.0.

Stalans, L.J. 1995. Multidimensional scaling. In: L.G. Grimm and P.R. Yarnold (eds.), Reading and understanding multivariate statistics. American Psychological Assoc., Washington, DC. 373 pp.

Statsoft Inc. 1995. Cluster analysis. In: Statsoft. Statistica Vol. III: Statistics II. Statsoft Inc., Tulsa, OK, pp. 3165-3195.

Statsoft Inc. 1996. Statistica for Windows, release 5.1 D.

Tabachnick, B.G., and L.S. Fidell. 1996. Multivariate statistics: Why? In: B.G. Tabachnick and L.S. Fidell (eds.), Using multivariate statistics. Harper Collins College Pubs., New York. 880 pp.

Walters, C. 1995. Fish on the line: The future of Pacific fisheries. The David Suzuki Foundation, Vancouver. 82 pp.

Mass-Balance Food Web Ecosystem Models as an Alternative Approach for Combining Multiple Information Sources in Fisheries

Daniel Pauly, Johanne Dalsgaard, and Thomas A. Okey
University of British Columbia, Vancouver, B.C., Canada

Robert Powell and Stuart Pimm
The University of Tennessee, Knoxville, Tennessee

Highly parameterized analytical single-species models offer a tempting framework for integrating data from different sources, e.g., survey biomass estimates, fishery catches, and catch composition data. We argue, however, that forcing data that usually cover a number of species into single-species models, however sophisticated, does not optimally use such data.

Rather, emphasis should be given to models that explicitly account for multispecies interactions, especially trophic models. While mathematically not complex, trophic models can be made complete, i.e., they can be made to include all groups in a system, and thus consider direct and indirect trophic impact on target species. Such completeness also, in itself, provides set limits on difficult-to-estimate stock sizes, production, and mortality rates, i.e., on processes directly relevant to fisheries resource management. In addition, these models lend themselves to answering questions about ecosystem dynamics and the responses of ecosystems to anthropogenic changes.

As an example, we discuss the properties and behavior of a mass-balance trophic model representing the Prince William Sound ecosystem from 1980 to 1989, i.e., prior to the *Exxon Valdez* oil spill, pending the construction, through collaboration with experts on the various ecosystem components, of a more comprehensive, consensus model to be used for answering questions such as those mentioned above.

Adapted from a presentation at the 15th Lowell Wakefield Symposium, Fishery Stock Assessment Models for the 21st Century, October 8-11, 1997, Anchorage, Alaska.

Estimates of Shrimp Trawl Bycatch of Red Snapper (*Lutjanus campechanus*) in the Gulf of Mexico

B.J. Gallaway
LGL Ecological Research Associates, Inc., Bryan, Texas

M. Longnecker
Texas A&M University, Department of Statistics, College Station, Texas

J.G. Cole
LGL Ecological Research Associates, Inc., Bryan, Texas

R.M. Meyer
Meyers Chuck, Alaska

Abstract

Estimation of red snapper bycatch in the shrimp trawl fishery of the Gulf of Mexico has been a contentious issue. Estimates are generated by the National Marine Fisheries Service (NMFS) using a general linear model which establishes a relationship between resource trawl survey data and catch data from the fishery obtained by observers on shrimp fishing vessels. The more complete time series of resource trawl data is then used to predict commercial vessel CPUE which is multiplied by total fishing effort to determine bycatch. The estimates are characterized by exceptionally low R^2 values and highly skewed residuals (70% of the catch observations were zeros). We have attempted to improve the estimates by using fewer and larger time-space cells, pooling catch and effort data to reduce the number of zeros contained in the analysis, incorporating significant interactions, and using epochs to guard against nonstationarity. The R^2 values for the revised models are 2 to 3 times higher than the R^2 for the base case, and the distribution of the residuals is greatly improved. The revised estimates in recent years average on the order of 30 to 47% lower than the NMFS estimates. Nevertheless, bycatch levels are high (26 to 32 million per year) and are increasing due to increasing abundance of juveniles. However, the age of structure of the bycatch may consist of a much larger fraction of age-0 fish and fewer age-1 fish than has been thought. Bycatch Reduction Devices (BRDs) do not effectively exclude age-0 red snapper. Thus, the existing stock recovery policy based on NMFS GLM bycatch estimates and using BRDs to reduce shrimp trawl mortality of juvenile red snapper may be ineffective.

Introduction

The fishery for red snapper *(Lutjanus campechanus)* in the Gulf of Mexico began over 150 years ago off Pensacola, Florida (Goodyear 1995) and by 1872 it had developed as a separate industry. The stocks offshore of Pensacola were greatly depleted between 1865 and 1883, causing the fishery to shift first to the Florida middle grounds (1883 to 1885), and then farther southward along the west Florida coast during 1885 to 1910 (Camber 1955). In 1892, the fishery also expanded to (1) newly discovered red snapper grounds in the Campeche Banks off Mexico, and (2) to the western Gulf of Mexico between the mouth of the Mississippi River to about Galveston, Texas. In the eastern U.S. gulf and Campeche, Mexico, subsequent geographic expansions of the fishery were driven by dwindling stocks in the areas previously fished.

The U.S. fishing fleet was excluded from Mexican waters in the early 1980s and as a result the effort was redirected at the remaining stocks in U.S. waters. By this time, the U.S. stock was essentially restricted to the western Gulf part of the range from Mississippi/Alabama to Texas. This area has been fished commercially since 1892, and also lies in the heart of the Gulf of Mexico shrimping grounds. Many juvenile (age-0 and age-1) red snapper are taken as bycatch in the shrimp fishery (Nichols et al. 1987, 1990; Nichols 1990; Nichols and Pellegrin 1992; Nichols 1996).

Although there is some debate (e.g., Rothschild et al. 1997), most consider the gulf red snapper stock to be, at present, severely overfished (Goodyear 1995, MRAG Americas, Inc. 1997). Management actions began in the mid-1980s and a stock rebuilding plan has been developed. Some of the management measures implemented between 1984 and 1996 include size and bag limits for the recreational fisheries, commercial and recreational quotas, prohibition of traps and longline gears in certain areas, and prohibition of commercial sale of red snapper from shrimp trawls. Also, Turtle Excluder Devices (TEDs) were mandated for use in the gulf shrimp fishery in 1990. These mechanical separation devices likely exclude large fish as well as turtles, and may have some effect on reducing take of juvenile red snapper (unpublished data).

Collectively, these actions appear to have had positive effects on the stock as reflected by increases in both stock and recruitment and, possibly, the increased size of harvested fish (Schirripi and Legault 1997, Rothschild et al. 1997). A key component of the stock rebuilding plan, yet to be implemented, is to reduce mortality from shrimp trawl bycatch of age-0 and, in particular, age-1 red snapper through the use of bycatch reduction devices (BRDs). BRDs are more effective at excluding age-1-sized than age-0-sized red snapper (Nichols et al. 1995, NMFS 1996).

Quantification of bycatch levels and the age-0 and age-1 fractions represented in this incidental catch are necessary for stock assessment and rebuilding evaluations. Estimates of bycatch are provided by the National Marine Fisheries Service (NMFS) as described in Goodyear (1995). In summary, the NMFS bycatch estimates are generated from a general linear model (GLM) applied to two datasets generated from shrimp trawl catches of red snapper. One dataset consists of catch-per-tow data which are provided from resource surveys conducted by NMFS, predominantly in summer and fall of each year (Nichols and Pellegrin 1989, Goodyear 1995). Features of this program since 1985 include a semi-synoptic sampling of the entire western Gulf of Mexico in summer and fall based on a random sampling design and the use of a standard shrimp trawl. Although the scope and design of the resource surveys have varied through time, continuous data are available for the fall season offshore Louisiana since 1972. For simplicity, we shall refer to these data as SEAMAP data, and include results from the Fall Groundfish and Summer SEAMAP programs.

The second dataset comes from records of finfish catch and fishing effort compiled on an individual tow basis by observers placed on shrimp fishing vessels specifically to quantify the bycatch including red snapper. Observer data are collected year round, but observer programs are not conducted every year. Even when conducted, only a small fraction of the fleet is sampled. Observer data are available for 1972 to 1982, and 1992 to 1996 periods. We shall refer to these data as Observer (OBSR) data.

The structure of the GLM model used by NMFS to estimate commercial catch per unit effort (CPUE) for a single net is:

$$\text{Log }(\text{CPUE}+1)_{ijklmn} = \text{mean} + \text{dataset}_i + \text{year}_j + \\ \text{season}_k + \text{area}_l + \text{depth}_m + e_{ijklmn}$$

for an array of space (4 areas × 2 depths) – time (three 4-mo seasons or trimesters) cells over the 24-y period, 1972 to 1996. In effect, the GLM "calibrates" shrimp vessel catch rates and resource trawl surveys during the periods and areas that had observations in common, and then uses the resource trawl data to index shrimp trawl bycatch (Nichols et al. 1987, 1990; Nichols and Pellegrin 1992). The GLM-based estimators of the mean log (CPUE+1) are transformed to an unbiased estimate of the commercial CPUE. These estimates are then multiplied by 2 (the assumed average number of nets) times the effort estimated for that time-space cell in hours fished. The catch estimates are then summed to provide an overall bycatch estimate. This GLM approach has been selected because observer programs have not been conducted in each space-time cell and the stock assessments require annual estimates of bycatch mortality. Use of the

SEAMAP data enables an estimate for each year as well as an update of the previous year's estimates.

We believe the structure of the NMFS GLM model is problematic and that improvements can be made. The basic problems are that (1) on the order of 70% of the tows in each dataset have zero catch of red snapper, (2) less than 50% of the space-time cells have been sampled overall (29% of the OBSR cells and 46% of the SEAMAP cells), (3) interactions between the five factors in the GLM model are ignored, and (4) there is no consideration of possible effects of nonstationarity (effects of explanatory variables changing over time). Our approach toward addressing these problems is to (1) combine catch and effort over several tows to reduce the number of zeros, (2) use fewer but larger strata to reduce the number of empty cells, (3) consider significant interactions, and (4) conduct separate models for early and late epochs to protect against nonstationarity.

Our hypothesis was that the above described changes would result in a better fit of the revised models (higher R^2 and lower residual error) as compared to models using the NMFS GLM structure, and a more normal distribution of residuals. The findings of the revised approach are presented and then discussed in terms of their ramifications with respect to management measures designed to recover red snapper stocks.

The Data

As noted above, the basic data used in the analysis are of three types: observer data, resource trawl survey data, and shrimp fishing effort data. The observer data are available for only two periods: 1972 to 1982 (historical) and 1992 to 1996 (modern). The historical observer data and what we have defined as the SEAMAP data were obtained from S. Nichols (NMFS, Pascagoula Laboratory) and the "modern" observer data were obtained from J. Nance (NMFS, Galveston Laboratory). Dr. Nance also provided the shrimp fishing effort data which is estimated by statistical reporting grid and depth zone within each grid. The spatial distributions of the OBSR tow data, historical and modern, are shown in Figs. 1 and 2, along with the distribution of the tows that contained red snapper. The historical (1972-1982) sample sizes are small as compared to the sample sizes obtained in the 1992 to 1996 program, especially when considered on an annual basis (Figs. 1 and 2).

The distribution of the SEAMAP data for 1972 to 1984 and tows containing red snapper during this period are shown in Fig. 3; the same data for 1985 to 1996 are shown in Fig. 4. The "fall groundfish" component of the SEAMAP dataset originated in 1972, and through 1984 sampling was mainly restricted to the so-called primary area off Louisiana (see Goodyear 1995). In 1985, the fall component of the NMFS resource surveys was expanded to encompass the entire geographic region from Pensacola, Florida to Brownsville, Texas. The summer component of the resource trawl surveys (SEAMAP) originated in 1982 and sampled the entire western Gulf

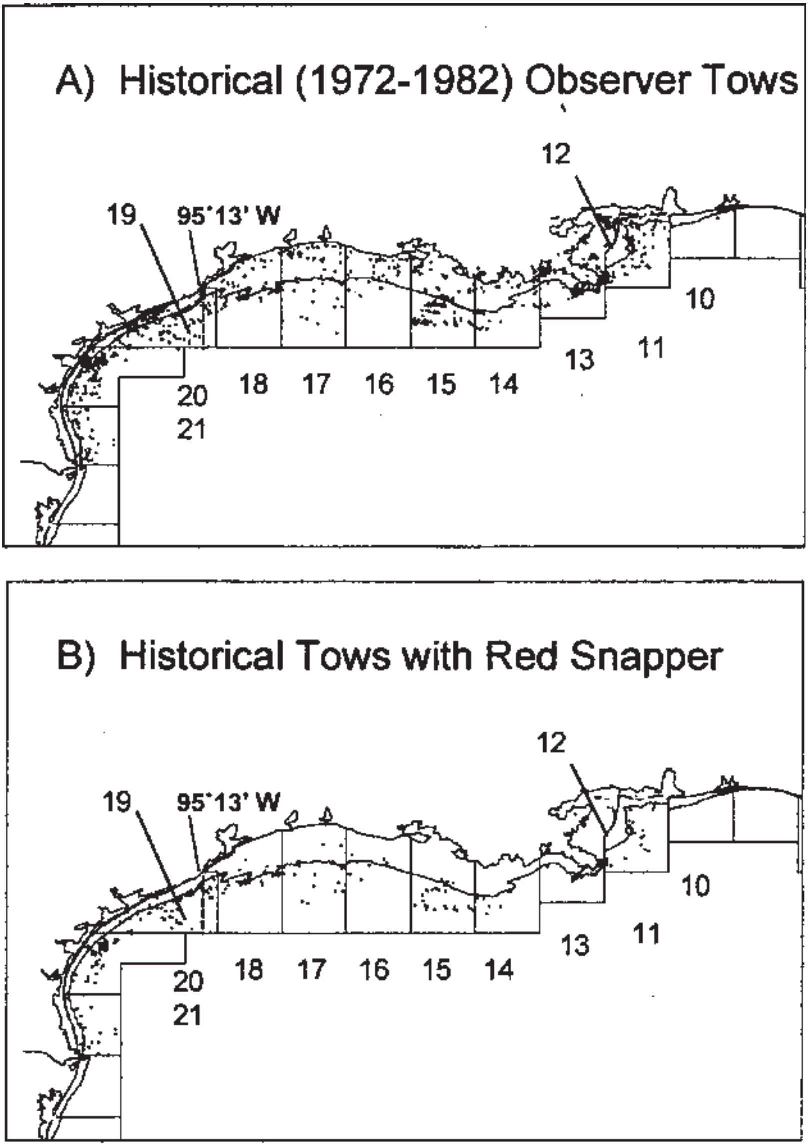


Figure 1. *Distribution of observer tows and tows containing red snapper, 1972 to 1982, in Gulf of Mexico statistical reporting grids 10-21. The 10-fathom contour is provided as a reference depth. Longitude 95°13′W provides the boundary between North and South Regions used in the GLM.*

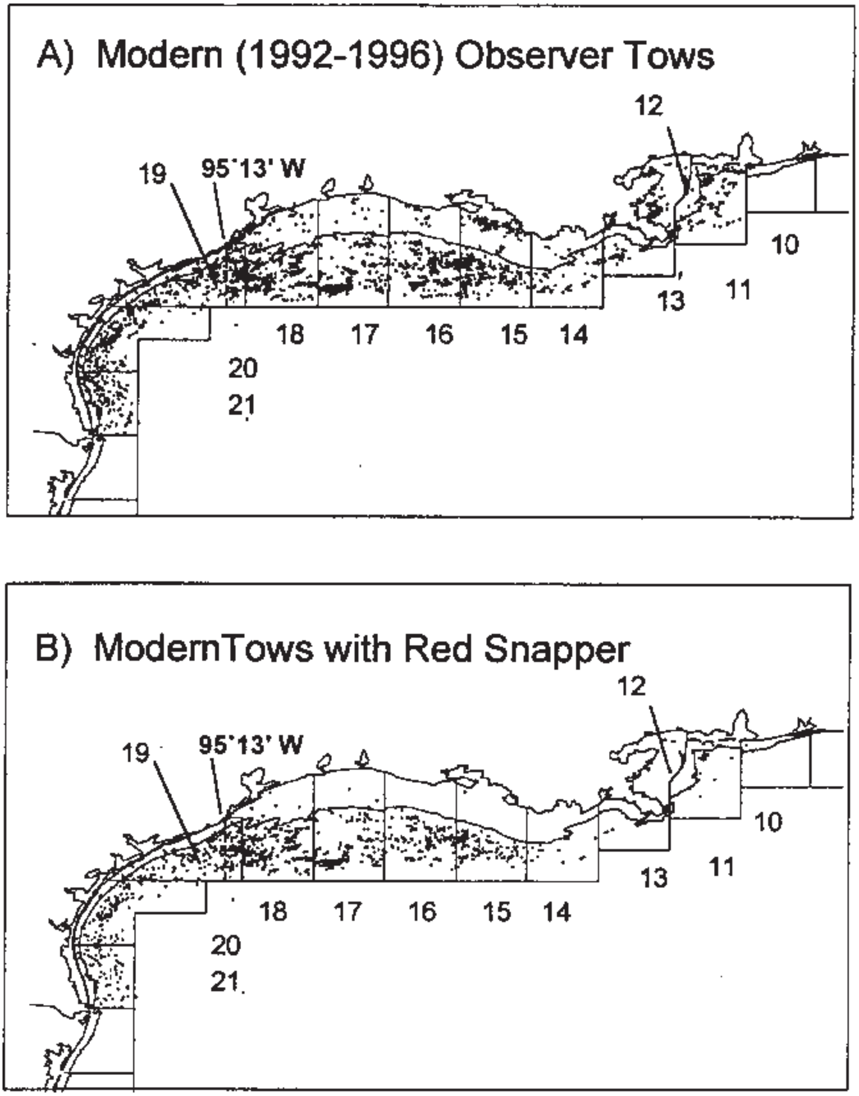


Figure 2. Distribution of observer tows and tows containing red snapper, 1992 to 1996, in Gulf of Mexico statistical reporting grids 10-21. The 10-fathom contour is provided as a reference depth. Longitude 95°13′W provides the boundary between North and South Regions used in the GLM.

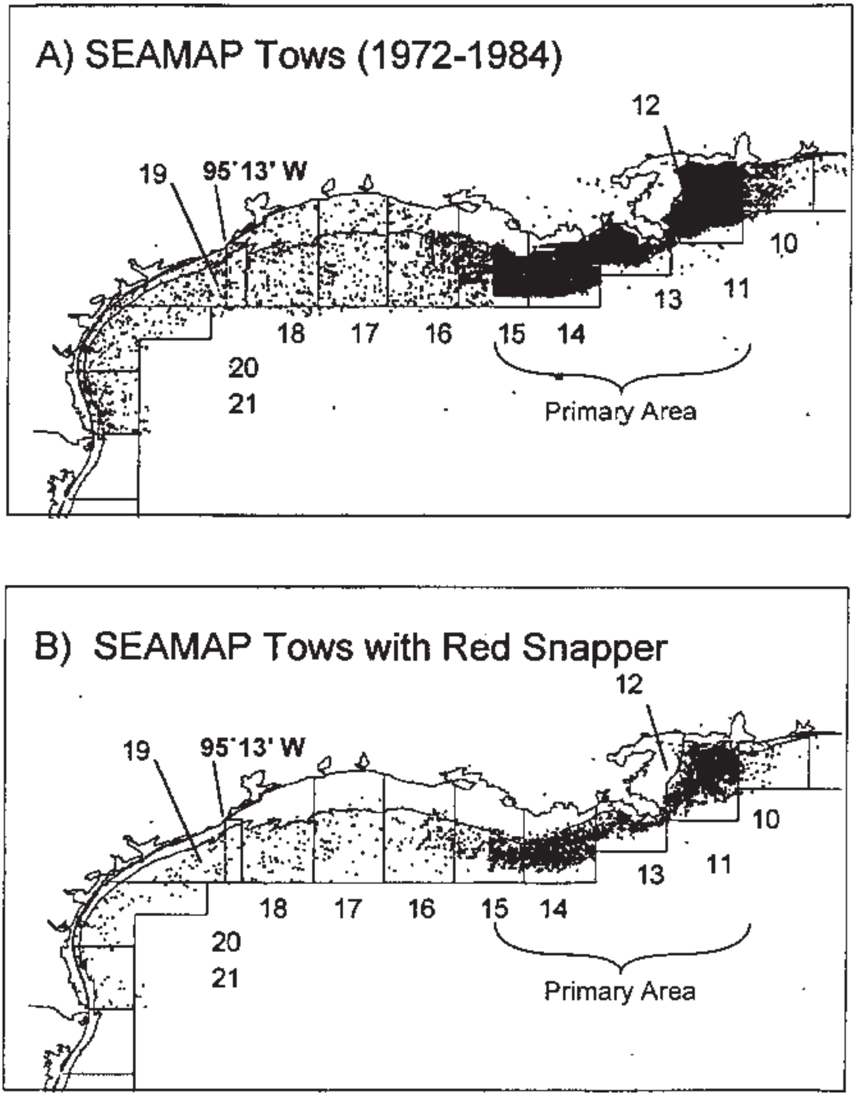


Figure 3. Distribution of SEAMAP tows and tows containing red snapper, 1972 to 1984, in Gulf of Mexico statistical reporting grids 10-21. The 10-fathom contour is provided as a reference depth. Longitude 95°13′W provides the boundary between North and South Regions used in the GLM.

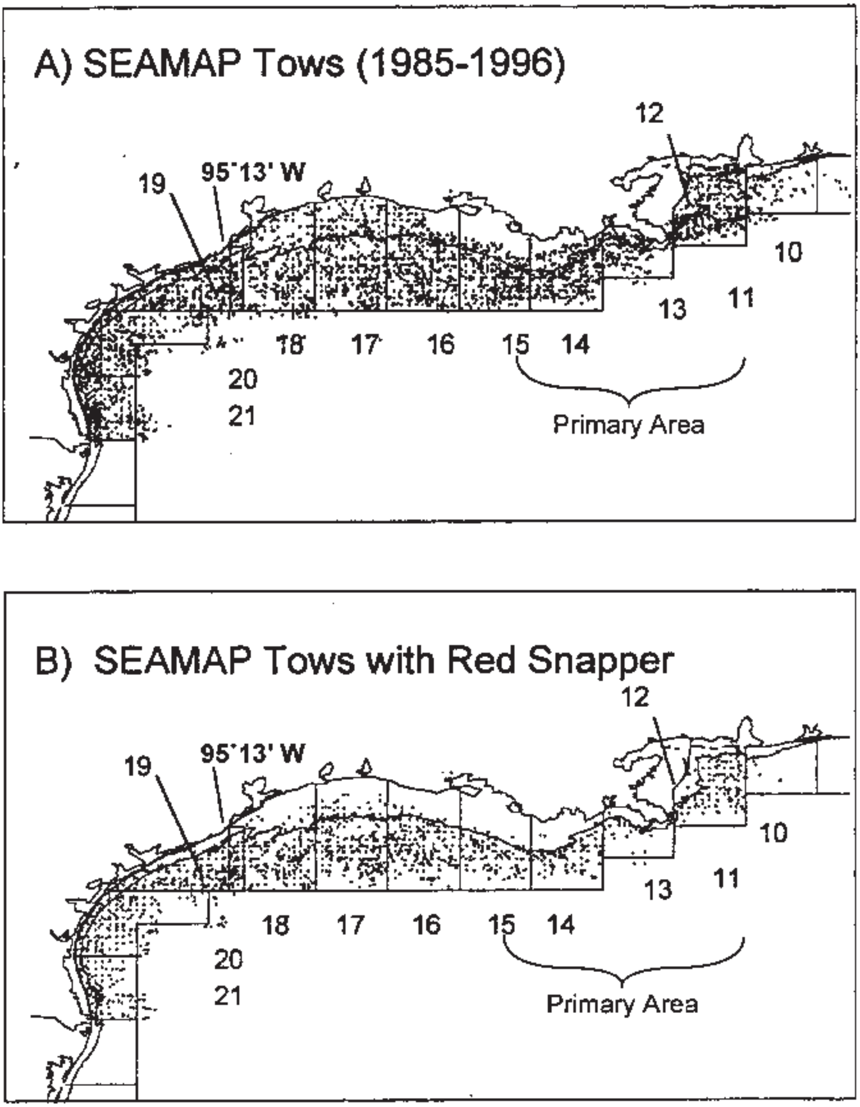


Figure 4. Distribution of SEAMAP tows and tows containing red snapper, 1985 to 1996, in Gulf of Mexico statistical reporting grids 10-21. The 10-fathom contour is provided as a reference depth. Longitude 95°13′W provides the boundary between North and South Regions used in the GLM.

from Pensacola, Florida to Brownsville, Texas. Summer and fall surveys thus became directly comparable in terms of spatial coverage in 1985, and shortly thereafter (1987) in terms of detailed sampling protocol. The data for years prior to 1985 are strongly dominated by fall samples taken in the primary area off Louisiana, but much better balance in seasonal and spatial sampling has been achieved since 1985 (compare Figs. 3 and 4).

Shrimp fishing effort data are estimated by statistical grid and depth zone based on a census of shrimp landings obtained from seafood dealers by port agents. The Gulf of Mexico is stratified into 221 spatial cells for each month. Port agents assign the landings to these cells based on their knowledge of the shrimp fishery and interviews. Interviews are conducted with a subset of the landings, targeting larger vessels, to obtain CPUE data of the shrimp fleet in each of these cells every month. Where data are missing, the CPUE of a cell is imputed using a GLM model. These estimates are then used to calculate an estimate of total shrimping effort in each of these cells every month. The values are summed to provide an estimate of total shrimp fishing effort.

Model Structure and Approach

We use only two spatial strata in our models, both restricted to the western gulf. The North Region consists essentially of statistical grids 10-18 and the South Region includes grids 19-21 (Figs. 1-4). Juvenile red snapper are scarce in the gulf region east of grid 10, and shrimp fishing effort there is small compared to effort in the modeled area. The South Region corresponds roughly to the Texas Transitional Faunal Province of Pulley (1952) and the Dry Sub-Humid Climatological Zone of Parker (1960). The North Region has been a major source of red snapper in the Gulf of Mexico since 1892. Some major differences between the North and South Regions are shown in Table 1.

Our models use only two periods—January to August and September to December—rather than three as used by NMFS. SEAMAP data for the January-April period were relatively sparse and winter shrimping effort is typically low compared to summer and fall. Further, age-1 juveniles predominate from January-August; age-0 fish are most abundant in the catches in September-December periods.

We did not incorporate separate depth cells in our models, but rather used two "cases." For Case I, the models were constructed using OBSR, SEAMAP, and shrimp fishing effort based on samples obtained for all depths greater than 5 fathoms for most of the North Region (as detailed below), and used data for all offshore depths for the remaining North Region and all of the South Region. For Case II, the models were based on data from depths greater than 10 fathoms. Examination of Figs. 1 and 2 shows that red snapper juveniles are infrequently encountered in commercial shrimp tows inside of 10 fathoms, and become even less frequent with proximity to the mainland. The historical (1972-1984) SEAMAP data reflect a similar

Table 1. Rationale for region determinations.

North Region (Grids 10-18)	South Region (Grids 19-21)
Positive flow estuaries	Neutral or negative flow estuaries
River influenced nearshore waters	More saline nearshore
Hypoxia prevalent	Hypoxia not prevalent
Currents predominantly to west in Zones 13-18	Currents predominantly to north
Bathymetry gradient gentle	Bathymetry gradient steep
Natural banks at shelf edge	Natural banks on mid-shelf
High density of petroleum platforms	Low density of petroleum platforms

The South Region corresponds to the Texas Transitional Faunal Province of Pulley (1952) and the Dry Sub-humid Climatological Zone of Parker (1960). Boundary between Stat Areas 18 and 19 displaced west to 95°13′W offshore to correspond with inshore boundary; i.e., “dogleg” removed (see Figs. 1-4).

trend (Fig. 3), but red snapper have occurred with some regularity inside of 10 fathoms in statistical grids 17-21, especially in recent years (Fig. 4). Gallaway and Cole (1997) have shown that high-value habitat for juvenile red snapper (and brown shrimp) in the western Gulf of Mexico to be largely restricted to depths (>10 fathoms based upon results of habitat suitability modeling [USFWS 1980, 1982]).

We constructed two temporal versions (epochs) for each model case. The first epoch covered the period 1972 to 1984; the second covered 1985 to 1996. In 1985, the spatial coverage of the fall component of the NMFS resource surveys was expanded from the so-called primary area off Louisiana (see Goodyear 1995) to encompass the entire geographic region from Pensacola, Florida to Brownsville, Texas. From this year, the summer and fall surveys became directly comparable in terms of spatial coverage and, shortly thereafter (1987), in terms of detailed sampling protocol. Also, the first red snapper management actions were taken in about that time frame (actually 1984), and it was not until about 1985 that good compliance with the mandatory dealer reporting of shrimp landings regulation was achieved (Pers. comm., M. Hightower, NMFS, Galveston). The latter should have resulted in more certain estimates of shrimping effort since effort is derived from the landings census.

Each model case thus has 100 cells for each dataset: 52 cells for epoch I (2 regions × 2 seasons × 13 years) plus 48 cells for epoch II (2 regions × 2 seasons × 12 years). Of these, 53% of the OBSR cells and 78% of the SEAMAP cells were sampled overall. This compares to the NMFS model structure which has 576 cells for each dataset (4 regions × 2 depths × 3 seasons × 24 years). Under this structure, only 29% of the OBSR cells and 46% of the SEAMAP cells were sampled overall.

We also reduced the percent of zero values by using catch per trip per week as the basic observation unit rather than catch per tow. In the SEAMAP data, the percent of zeros declined from 70.6 to 6.9%; and in the OBSR data from 70.9 to 33.2%. Further, since the resource survey tows are of short duration (e.g., 15 min), combining tows over weekly intervals reduced the imbalance between SEAMAP and OBSR sample sizes.

The data files used by NMFS to calculate their red snapper annual bycatch estimates were obtained from the NMFS Pascagoula Laboratory. This dataset included SEAMAP station and catch data for the years 1972-1995, and station and catch data from the early observer programs between 1972 and 1982 (Fig. 5). SAS® programs were run to merge the station and catch data into a new dataset (Mod_Data) which included the new variables period (NMFS trimester), year, week (SAS date divided by 7), region (North or South), and dataset (SEAMAP or OBSR).

The raw 1996 SEAMAP station, environment, and catch data, obtained from another department of the NMFS Pascagoula Laboratory were merged to form a new dataset (SEAMAP96) including the same variables from Mod_Data. The modern observer station and catch data were obtained from the NMFS Galveston Laboratory and were merged to create the "modern" Observer dataset, again including the same variables as the other sets (Fig. 5).

Assignment of the region fields in each of the data records was based on the combination of depth and longitude of the sampling station. All stations located west of 95°13′W were assigned to the south zone; stations between 87°W and 94°W and greater than 5 fathoms depth were assigned to the north region. Stations between 94°W and 95°13′W were also assigned to the north region, regardless of depth.

These three datasets were reduced to include only samples which contained a net operation code that indicated an undamaged tow, and these were then merged with a dummy dataset which included records for all year, period, and region cell combinations to be predicted by the SAS GLM (the catch per unit effort for each of the cell combinations was entered as a missing value) to form the model dataset.

The catch and effort data in model dataset were then summarized by year, period, region, week, and cruise number to create the dataset BYCRUISE. This catch and effort data was used to create a catch per unit of effort (CPUE) value for each year, period, region, week, and cruise number combination. The natural logarithm of the CPUE + 1 value was assigned to create the variable LNCPUE.

Data for years prior to 1985 were extracted from BYCRUISE to create the CRUISE_EARLY dataset; post 1984 data were used to create the CRUISE_LATE dataset. In each dataset, trimester 2 values were reassigned to trimester 1. The SAS procedure GLM was then run on each of the two new datasets, modeling LNCPUE as a function of the class variables dataset, year, period, and region, and their interactions. For the CRUISE_LATE dataset, the year by region interactions were found to be significant, and

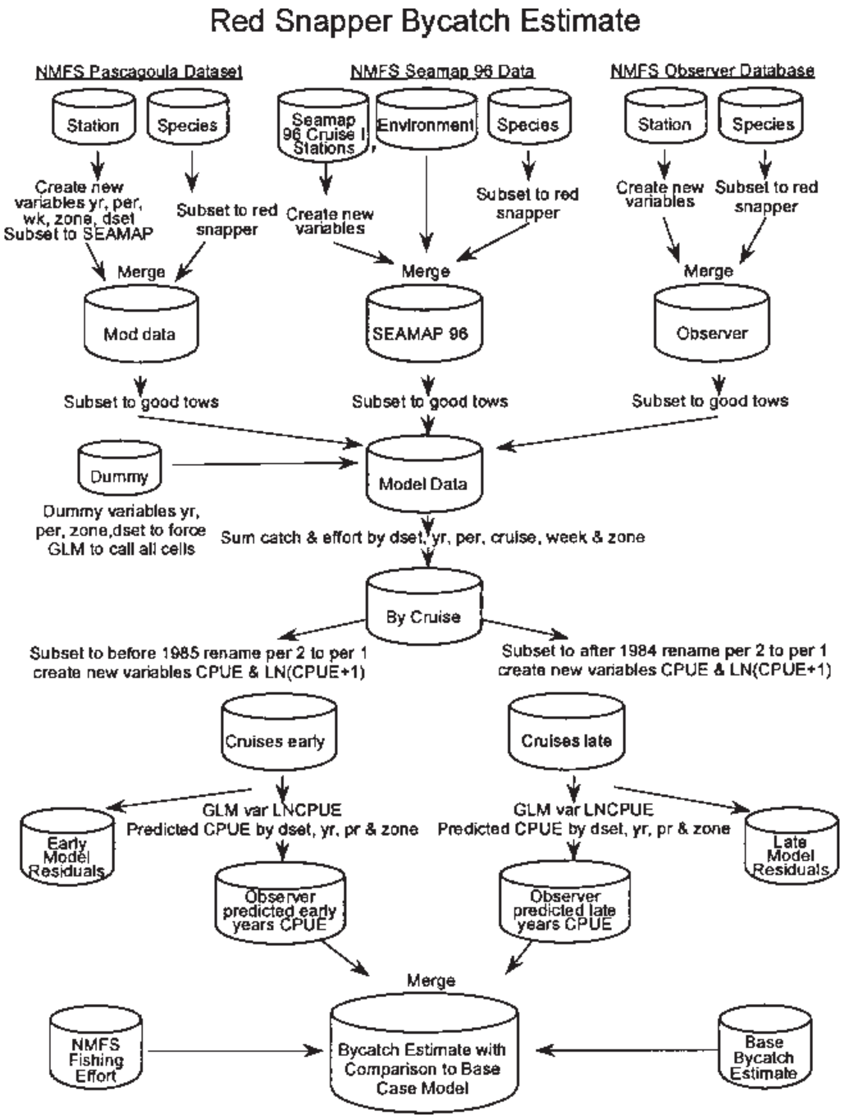


Figure 5. Schematic of the steps in the estimation of red snapper bycatch showing the sources and types of data, the modifications made to the data, and the sequence of the analysis.

were thus included in the analysis. No other significant interactions were found in either dataset.

The modified GLM results produced a pair of datasets with predicted LNCPUE for each year, period, region, dataset combination. The observer dataset values were translated by using the following formula:

$$e^{(\text{OBSR Predicted LNCPUE} + 0.5 \times \text{GLM MSE})} - 1$$

to create an unbiased estimate of the predicted observer cell values. These cell values were multiplied by the NMFS estimate of fishing effort for each cell (from the NMFS Galveston Laboratory dataset) to create a bycatch estimate for each year, period, and region combination. These estimates were then summed and compared to the annual bycatch estimates produced by the NMFS Pascagoula Laboratory (Fig. 5).

All of the bycatch in the January-August period was treated as if it was entirely age-1 fish. A few age-0 fish would be expected during June-August, but the fraction would be very low (less than 1 or 2% during June and July; on the order of 12% in August, Table 83 in Goodyear 1995). In fall, a substantial fraction of the catch (13 to 39%; Table 83, Goodyear 1995) can be age-1 fish although the catches are dominated by age-0 fish. The datasets provided to us included length-frequency data for juvenile red snapper in both the SEAMAP (1988 to 1995) and OBSR (1992-1995) datasets. We plotted length frequency histograms for each of our two regions. These were used to allocate the total fall bycatch for each of the two regions into age-0 and age-1 fractions.

Results and Discussion

The changes we instituted resulted in better performance of the models as compared to the NMFS estimates used as the base case (Table 2). Adjusted R^2 values for the historical period 1972 to 1984 and the more recent period (1985 to 1996) were 0.22 and 0.37, respectively, for Case I; and 0.22 and 0.38 for Case II. This compares to an R^2 of 0.13 for the NMFS base model. Likewise, skewness and kurtosis values of the modified GLM models were, considering sample size effects, greatly improved as compared to the base model, especially the model for the period 1985 to 1996 (Table 2). A part of the explanation for the better performance relates to improved balance between SEAMAP and OBSR data in recent years, particularly as compared to the historical period (compare Figs. 1-4). Using two models, one for each epoch, served to reduce the impact of the historical imbalances (which were severe in the early epoch) on the data for recent years. Pooling of effort and catch over several tows eliminated a large fraction of the zeros, and using larger spatial strata reduced the number of empty cells. The net result is better model performance, especially for the recent data.

Table 2. Results of ANOVA and univariate residuals analyses.

Basic ANOVA comparisons

Model structure	n	F	$P > F$	R^2	CV	Root MSE	LCPUE Mean
Case I 1972-1984 Epoch	446	7.87	0.0001	0.215	81.925	1.214	1.4819
Case I 1985-1996 Epoch	625	14.05	0.0001	0.370	68.377	0.867	1.2686
Case II 1972-1984 Epoch	350	5.87	0.0001	0.209	65.144	1.195	1.8338
Case II 1985-1996 Epoch	559	13.00	0.0001	0.379	63.360	0.858	1.3535
NMFS (1972-1995)[a]	25,390	125.29	0.0001	0.126	159.911	1.304	0.8156

Results of residual analysis

Model structure	Skewness	Kurtosis	Shapiro-Wilk statistics	$P < W$
Case I 1972-1984 Epoch	0.6445	–0.0487	0.9504	0.0001
Case I 1985-1996 Epoch	0.3065	0.0132	0.9690	0.0001
Case II 1972-1984 Epoch	0.3378	–0.2706	0.9714	0.0016
Case II 1985-1996 Epoch	0.2883	–0.1602	0.9749	0.0006
NMFS (1972-1995)[a]	1.3276	1.2746	0.1921[b]	<0.01[c]

[a] Calculated by LGL using stated model structure and data provided by NMFS.

[b] Kolmogorov *"D"*

[c] $P > D$

Bycatch Level

The ramifications of our changes in GLM structure to the bycatch estimates as compared to the NMFS results are shown by Table 3 and the bottom panel of Fig. 6. Our estimates for the period 1972 to 1984 vary around the NMFS base-case estimates showing an 11% reduction overall for Case I, and a 10% reduction for Case II. Four of the 13 annual estimates were actually higher in our revised models than annual estimates yielded by the NMFS model structure. In contrast, a net reduction of about 30% for Case I and 47% for Case II is seen for the recent epoch (1985 to 1995), and only one estimate from the revised models is higher (2%) than a corresponding estimate from the base case (Table 3).

Plots of the GLM estimates of annual CPUE and the corresponding levels of effort for each year illustrate the nature of the differences in

Table 3. **A comparison of shrimp trawl fishing bycatch estimates for juvenile red snapper in the western Gulf of Mexico, 1972 to 1996.**

GLM:		Case I Epoch I		Case II Epoch I	
Year	NMFS	This study	Reduction (%)	This study	Reduction (%)
1972	65,100,000	48,267,935	0.25856	39,181,831	0.39813
1973	23,200,000	14,370,909	0.38056	14,318,987	0.38280
1974	16,800,000	21,655,960	−0.28905	17,579,208	−0.04638
1975	15,200,000	8,338,273	0.45143	11,627,585	0.23503
1976	23,300,000	36,092,778	−0.54905	46,343,610	−0.98900
1977	24,800,000	27,219,142	−0.09755	28,203,783	−0.13725
1978	21,700,000	17,007,973	0.21622	20,564,819	0.05231
1979	23,000,000	16,593,497	0.27854	22,668,893	0.01440
1980	34,300,000	19,416,077	0.43393	21,804,067	0.36431
1981	34,300,000	57,943,153	−0.68930	62,250,591	−0.81489
1982	33,900,000	26,039,604	0.23187	30,439,811	0.10207
1983	21,300,000	12,853,086	0.39657	16,250,129	0.23708
1984	16,600,000	9,215,449	0.44485	11,126,182	0.32974
Totals	353,500,00	315,013,836	0.10887	342,359,597	0.09685

GLM:		Case I Epoch II		Case II Epoch II	
Year	NMFS	This study	Reduction (%)	This study	Reduction (%)
1985	20,100,000	11,712,627	0.41728	9,011,712	0.55166
1986	19,400,000	7,347,288	0.62127	6,500,352	0.66493
1987	24,400,000	14,450,654	0.40776	9,235,304	0.62150
1988	23,500,000	10,346,981	0.55970	9,386,562	0.60057
1989	28,400,000	13,101,080	0.53869	12,528,296	0.55886
1990	54,200,000	42,656,137	0.21299	22,959,556	0.57639
1991	48,100,000	28,266,626	0.41234	19,953,992	0.58516
1992	30,700,000	26,012,971	0.15267	24,551,139	0.20029
1993	32,900,000	24,354,494	0.25974	19,216,628	0.41591
1994	40,400,000	33,405,020	0.17314	23,292,520	0.42345
1995	42,000,000	43,184,149	−0.02819	35,895,796	0.14534
Totals	364,100,000	254,838,027	0.30008	192,531,856	0.47121

NMFS estimates are from Goodyear (1995).

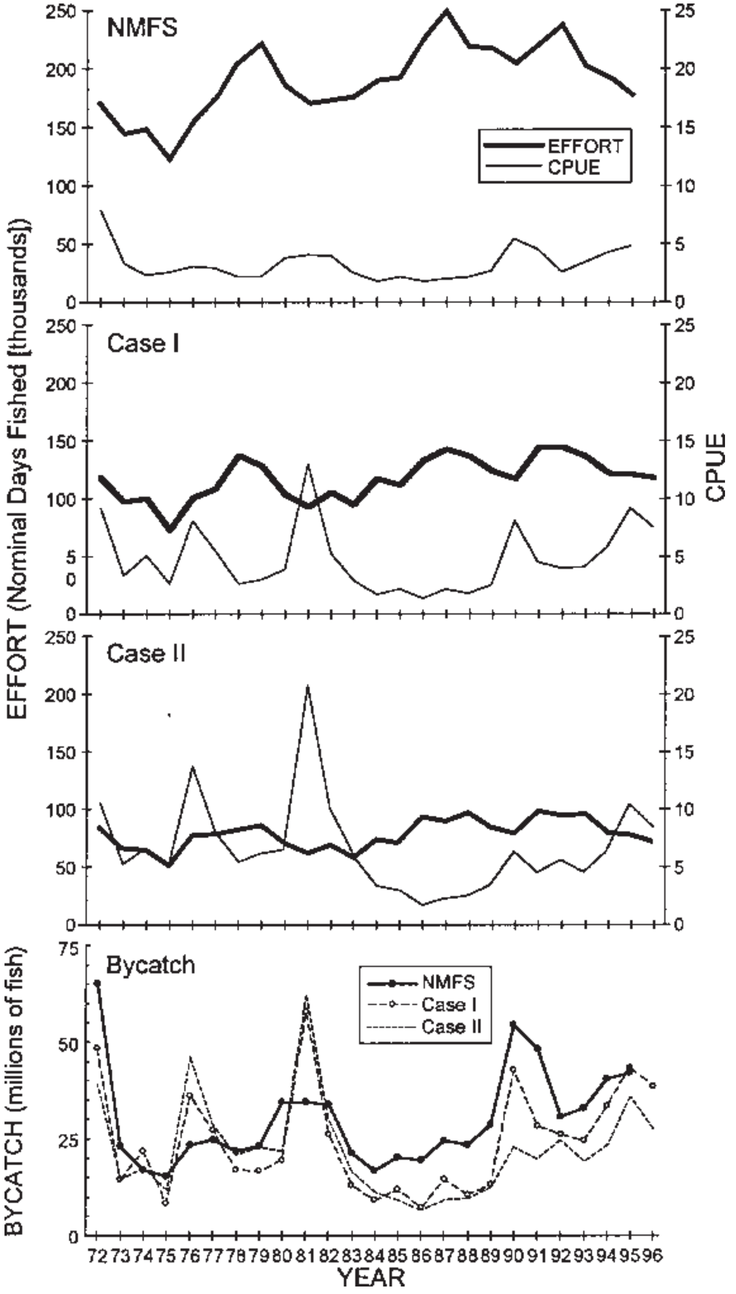


Figure 6. Annual trends in shrimp fishing effort, CPUE of red snapper in shrimp trawls, and red snapper bycatch in the western Gulf of Mexico, 1972 to 1996.

bycatch estimates among the three cases (NMFS; Case I and Case II of this study, Fig. 6). In the NMFS model, CPUE is low and shrimp fishing effort is high relative to our cases. This is because the NMFS model incorporates data for Florida where both juvenile abundance and shrimp fishing effort are low; and, more important, data from shallow nearshore areas in the western gulf where red snapper are either absent or low abundance, but shrimp fishing effort is high. Incorporating these data has a smoothing effect on the CPUE time series. In our cases, we excluded data for areas which, in our opinion, do not constitute suitable habitat for this species (Case I); and, in Case II, restricted the analysis to high-value habitats. As a result our values of CPUE are high relative to the NMFS case, but the corresponding effort multipliers are smaller (Fig. 6). With only one exception (1990), the annual CPUE values for Case II are higher than the values for Case I, reflecting the higher abundance of juvenile red snapper in areas greater than 10 fathoms in depth as compared to more shallow areas. The 1990 CPUE exception was largely attributable to a few large collections of red snapper which were obtained in the Fall SEAMAP sampling program just inside the 10 fathom contour in the North Region in the vicinity of the mouth of the Mississippi River where the bathymetry gradients are steep.

The results of the bycatch analyses show that the trend of increase in red snapper bycatch since the mid-1980s is attributable to increased CPUE of juvenile red snapper (Fig. 6). Effort has remained relatively stable, and, in the most recent years, has even declined. Fishery-independent recruitment indices for red snapper have exhibited a similar trend of increase over the same time frame and the range occupied by recruits appears to be expanding (Goodyear 1997, Schirripa and Legault 1997).

Age Composition of the Bycatch

The stock assessment requires that the bycatch be partitioned into age-0 and age-1 fractions for each year class beginning with 1982. Age-0 year class fish are abundant mainly in fall (September to December) and are treated as age-1 fish from January through the following December. The length frequency data for the fall collections of juvenile red snapper showed marked differences by region (Fig. 7). Juvenile fish of a size to suggest age 1 were not apparent in the samples from the South Region except in the 1992 SEAMAP collections (1 out of 8 years). In contrast, the length distributions were clearly bimodal in 6 of 8 fall collections from the North Region (Fig. 7). The fall bycatch totals for the South Region in 1992 to 1994 were allocated to age fractions based on the OBSR length frequencies, and the mean of the 1992 to 1994 OBSR bycatch frequency data for this region was used to allocate the bycatch to age fractions in years in which there were no OBSR size data (i.e., 1982 to 1991, 1995). The same approach was used in the North Region.

Based on the above, and using the more conservative Case I results, we estimate that the combined bycatch of the 1982 to 1992 year classes was on the order of 198 million fish, of which 65% were age-0 fish and 35%

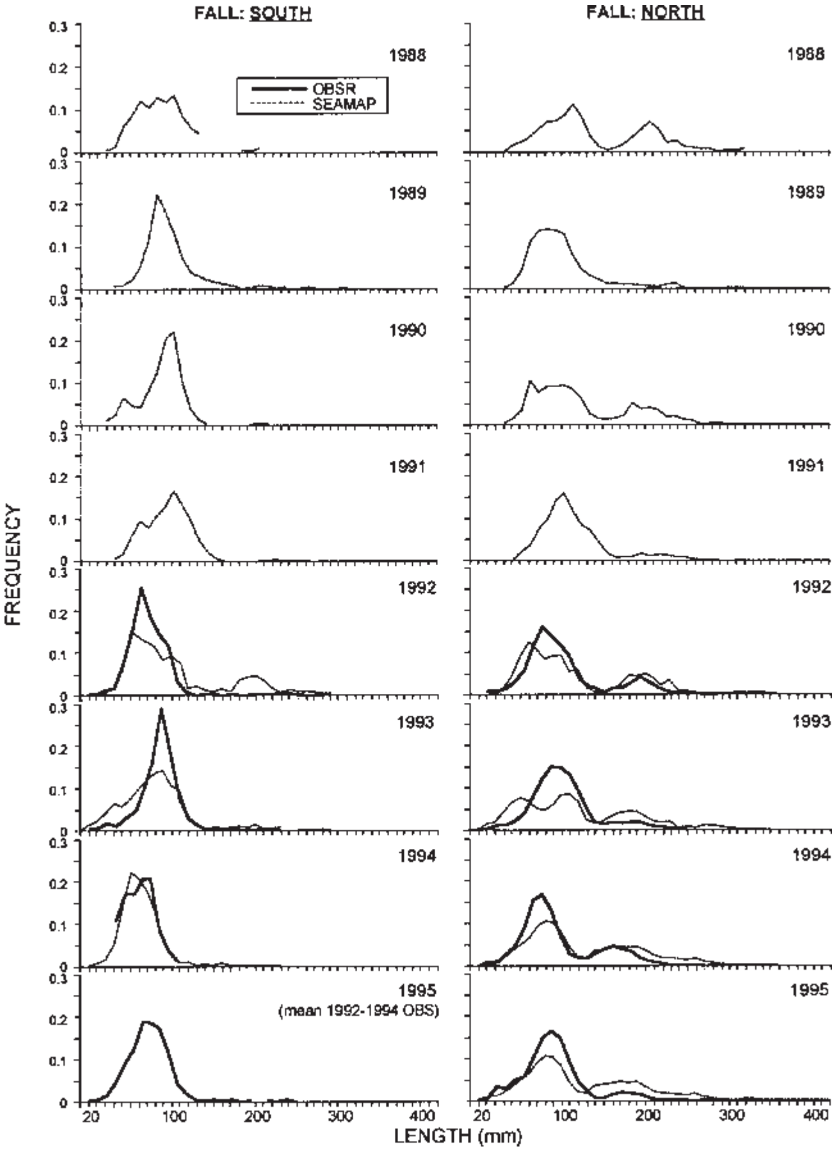


Figure 7. Length frequency of juvenile red snapper taken in the shrimp fishery observer program (OBSR) and National Marine Fisheries Service resource trawl surveys (SEAMAP) in fall in the North and South Regions of the Gulf of Mexico used in the analyses, 1988 to 1995.

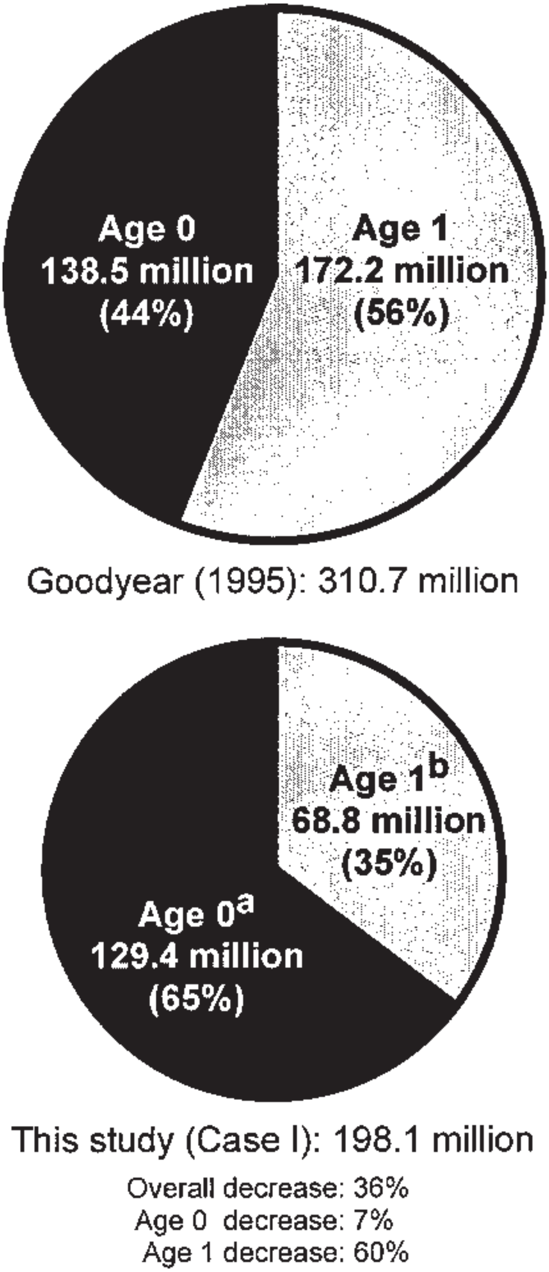


Figure 8. Estimated total bycatch and age composition of the bycatch for the 1982 to 1992 year classes based on comparisons of the results of Goodyear (1995) to results from this study.

were age-1 fish. The corresponding estimates used in the Goodyear (1995) stock assessment are 311 million fish, of which only 44% were age 0, and 56% were age 1 (Fig. 8). The difference appears to stem from NMFS using length data combined from both datasets as well as over all regions to estimate the age composition of the fall collections, whereas we partition the catch data to age on a region-specific basis. The overall reduction indicated for total bycatch from our estimates as compared to the NMFS estimates (36%) may not seriously affect the conclusions of the stock assessment (Goodyear 1996). However, the marked differences in age composition of the bycatch are a matter of concern since BRDs do not effectively exclude age-0-sized red snapper. Their mandated use thus may not be an effective means of lowering shrimp trawl bycatch mortality.

Independent Estimates

Rothschild et al. (1997) have independently estimated red snapper bycatch in the Gulf of Mexico shrimp fishery. Their approach was to use only the observer data and they initially focused on the depths (5-20 fathoms), locations (grids 16-21), and time of year where the great majority of red snapper bycatch is taken (May-December). They used standard ratio estimation techniques to derive CPUE, and multiplied these by the corresponding effort to obtain bycatch on a per net basis. To complete the time series for years with no OBSR data (1982-1991), the computed mean CPUE values for the available data were multiplied by the corresponding ratios of shrimp effort yielding estimates for all years. The correspondence between predicted and observed values in recent years was good, but there was less similarity between predicted and observed values for the historical period. They advised that the historical estimates should be viewed with reasonable caution.

The results obtained for recent years by Rothschild et al. (1997) indicated a bycatch level between 7 and 8 million juvenile fish per trawl net, which, assuming an average of two nets, provides a lower bound estimate of 15 million. Based on their professional opinion, this lower bound was inflated by 50% to account for multiple nets, catches during other times and places, etc. They estimated that total bycatch was perhaps on the order of 20 to 25 million fish (Rothschild et al. 1997). They suggested the age-0 component might be on the order of 15 million fish or about 75% of the total. Rothschild et al. (1997) also commented on the scarcity of age-1 fish off the Texas coast during fall, providing independent confirmation of our findings regarding regional differences in age composition of the bycatch.

Conclusions

The results of our analysis and those of Rothschild et al. (1997) would suggest that bycatch of the 1982 to 1992 year classes has been overestimated in the Goodyear (1995) stock assessment by 36%. Nevertheless, the number of juvenile red snapper taken incidental to shrimp trawling is large (in our study the 1992 to 1996 average was 26 million for Case II and 32 million for Case I), and bycatch is increasing as juveniles are becoming more abundant (Case I value was 43 million in 1995). The more important finding of this study is that the age-0 fraction of the catch is likely much larger and the age-1 fraction much lower than has been previously assumed. Since BRDs do not effectively exclude age-0 fish, the anticipated level of gain from BRD use may not be realized. Recruitment, the distributional range of recruits, and stock level of red snapper are apparently all increasing in the Gulf of Mexico, as is size of harvested fish (Schirripa and Legault 1997, Rothschild et al. 1997). The commercial fishery presently lasts only days before the quota is filled, and the recreational quota in 1997 was, for the first time ever, reached before the year ended. All of this is occurring in the face of unabated and growing bycatch levels.

Acknowledgments

We especially thank Scott Nichols and Ken Savastano of the NMFS Pascagoula Laboratory for the SEAMAP and historical OBSR data, and Jim Nance of the NMFS Galveston Laboratory for the modern OBSR data and the shrimp fishing effort data files. We have been involved in three peer reviews of the bycatch estimation procedures that are being used in the Gulf of Mexico shrimp fishery. The panelists who conducted the reviews helped shape our approach. The initial panel consisted of C.E. Gates, J.P. Geaghan, D.W. Hayne, J.M. Hoenig, and G.E. Lewis; the second panel consisted of M.C. Christman, L.P. Fanning, D.B. Hayes, and M.S. Kaiser; the third panel included M.K. McAllister, A. Sinclair, K.T. Stokes, J. Sutinen, and T.E. Target. All provided valuable insights, but the approach we used and the opinions we express herein do not necessarily represent the views of any of the listed panelists. We especially thank Wayne Swingle of the Gulf of Mexico Fishery Management Council and John Witzig of the NMFS Office of Science and Technology who organized and distributed results of the peer reviews. The Texas Shrimp Association, Inc. and the Biological Resources Division of the United States Geological Survey provided funding for this study on a collaborative basis.

References

Camber, C.I. 1955. A survey of the red snapper fishery of the Gulf of Mexico with special reference to the Campeche Banks. Technical Series of Florida State Board of Conservation 12:1-64.

Gallaway, B.J., and J.G. Cole. 1997. Cumulative ecological significance of oil and gas structures in the Gulf of Mexico: A Gulf of Mexico fisheries habitat suitability model—Phase II model description. U.S. Geological Survey, Biological Resources Division, USGS/BRD/CR, 1997-0009 and Minerals Management Service, Gulf of Mexico OCS Region, New Orleans, LA, OCS Study MMS 97-0044. 109 pp.

Goodyear, C.P. 1995. Red snapper in U.S. waters of the Gulf of Mexico. NMFS Southeast Fisheries Science Center, Miami Laboratory, Coastal Resources Division, Miami, FL. MIA-95/96-05. 171 pp.

Goodyear, C.P. 1996. Red snapper bycatch mortality: Implications of possible estimate bias on parameters of the recovery plan. Report to the Gulf of Mexico Fishery Management Council. 15 pp.

Goodyear, C.P. 1997. An evaluation of the minimum reduction in the 1997 red snapper shrimp bycatch mortality rate consistent with the 2019 recovery target. Report to the Gulf of Mexico Fishery Management Council. 14 pp.

MRAG Americas, Inc. 1997. Consolidated report on the peer review of red snapper *(Lutjanus campechanus)* research and management in the Gulf of Mexico. NMFS, Office of Science and Technology, December 1997.

Nichols, S. 1990. The spatial and temporal distribution of the bycatch of red snapper by the shrimp fishery in the offshore waters of the U.S. Gulf of Mexico. NMFS, Mississippi Laboratories, Pascagoula, MS. 52 pp.

Nichols, S. 1996. An update on some issues relating to the distribution of red snapper bycatch. NMFS, Southeast Fisheries Center, Mississippi Laboratories, Pascagoula, MS. 9 pp.

Nichols S., and G.J. Pellegrin Jr. 1989. Trends in catch per unit effort for 157 taxa caught in the Gulf of Mexico fall groundfish survey 1972 to 1988. NMFS, Mississippi Laboratories, Pascagoula, MS. 10 pp.

Nichols, S., and G.J. Pellegrin Jr. 1992. Revision and update of estimates of shrimp fleet bycatch 1972-1991. NMFS, Southeast Fisheries Center, Mississippi Laboratories, Pascagoula, MS. 17 pp.

Nichols, S., A. Shah, G. Pellegrin Jr., and K. Mullin. 1987. Estimates of annual shrimp fleet bycatch for thirteen finfish species in the offshore waters of the Gulf of Mexico. NMFS, Southeast Fisheries Center, Mississippi Laboratories, Pascagoula, MS. 29 pp.

Nichols, S., A. Shah, G.J. Pellegrin Jr., and K. Mullin. 1990. Updated estimates of shrimp fleet bycatch in the offshore waters of the U.S. Gulf of Mexico 1972-1989. NMFS, Mississippi Laboratories, Pascagoula, MS. 22 pp.

Nichols, S., J. Nance, C.P. Goodyear, A. Shah, and J. Watson. 1995 Some considerations in determining bycatch reduction requirements. NMFS, Southeast Fisheries Science Center, Miami, FL. 18 pp.

NMFS. 1996. Summary report on the status of bycatch reduction device development. NMFS, Southeast Fisheries Science Center, Mississippi Laboratory, Pascagoula, MS. 34 pp.

Parker, R.H. 1960. Ecology and distributional pattern of marine invertebrates, northern Gulf of Mexico. In: F.P. Shepard, F.B. Pheleger, and T.H. van Andel (eds.), Recent sediments, northwest Gulf of Mexico. American Association of Petroleum Geologists, Tulsa, OK, pp. 203-337.

Pulley, T.E. 1952. A zoogeographic study based on the bivalves of the Gulf of Mexico. Ph.D. thesis. Harvard University, Cambridge, MA.

Rothschild, B.J., A.F. Sharov, and A.Y. Bobyrev. 1997. Red snapper stock assessment and management for the Gulf of Mexico. Report submitted to NOAA, NMFS Office of Science and Technology, by University of Massachusetts, Center for Marine Science and Technology, North Dartmouth, MA. 173 pp.

Schirripa, M.J., and C.M. Legault. 1997. Status of the red snapper in U.S. waters of the Gulf of Mexico. NMFS, Southeast Fisheries Science Center, Miami Laboratory, MIA-97/98-05. 37 pp.

USFWS. 1980. Habitat evaluation procedures (HEP). U.S. Fish and Wildlife Service, Division of Ecological Services, ESM 102.

USFWS. 1982. Habitat suitability index models: Appendix A. Guidelines for riverine and lacustrine applications of fish HSI models with habitat evaluation procedures. U.S. Fish and Wildlife Service, Division of Ecological Services, FWS/OBS-82/10.A. 53 pp.

A Summary of Assessment Information for Managing Alaska Groundfish Stocks

Jane DiCosimo
North Pacific Fishery Management Council, Anchorage, Alaska

Abstract

The North Pacific Fishery Management Council (NPFMC) is responsible for effectively managing the groundfish fisheries in the Bering Sea (BS), Aleutian Islands (AI), and Gulf of Alaska (GOA). These fisheries target walleye pollock *(Theragra chalcogramma),* Pacific cod *(Gadus macrocephalus),* sablefish *(Anoplopoma fimbria),* Atka mackerel *(Pleurogrammus monopterygius)* and numerous flatfish (*Pleuronectes* sp. and *Hippoglossoides* sp.) and rockfish (*Sebastes* sp. and *Sebastolobus* sp.) species. The stocks are routinely evaluated by National Marine Fisheries Service (NMFS) scientists, who are relied on by the NPFMC to recommend harvest levels that will maintain healthy stocks. Alaska Department of Fish and Game scientists evaluate the stocks of three species of demersal shelf rockfish, for which the NPFMC has delegated management authority to the State of Alaska.

Introduction

Stock assessments have evolved since 1978 in response to changes in target species, data collection, and assessment methodology. Currently, biomass for most groundfish stocks is estimated using a stock synthesis model described by Methot (1990). The lack of age data, however, has prevented the traditional application of the synthesis model to some stocks of flatfish, rockfish, GOA Atka mackerel, and squid. Instead, stock biomass is estimated using an area-swept index from trawl survey data.

There also have been changes in methodology for estimating optimal harvest rates and overfishing rates. These harvest rates, when applied to estimated biomass for individual stocks, result in a preliminary recommendation for acceptable biological catch (ABC) and an overfishing level (OFL). For most stocks, ABC estimates are determined by calculating the fishing mortality which reduces the equilibrium level of spawning biomass per recruit to 40% of its unfished level ($F_{40\%}$) (Clark 1993). The OFL for most stocks is currently based on an $F_{30\%}$ rate. In the absence of maturity and growth information, ABC is based on an $F = 0.75M$ harvest strategy, and overfishing is based on $F = M$.

Historical Assessment Methodologies

The groundfish assessments are compiled into a stock assessment and fishery evaluation (SAFE) report, which is prepared and reviewed annually for each fishery management plan. These assessments undergo a thorough peer review process, first by the council's groundfish plan teams which review the stock assessments and then by the council's Scientific and Statistical Committee (SSC), both of which are composed of biologists and economists from state and federal agencies, and academia.

Conservation-based scientific assessment advice, along with strict adherence to scientific advice by managers, has resulted in relatively healthy stocks of groundfish in the North Pacific. Of 37 North Pacific groundfish stocks examined, three-year trends in relative abundance, based on stock assessment and survey catch-per-unit-effort (CPUE) trends, indicate 8 stocks are increasing, 8 stocks are stable, 13 stocks are decreasing, and 8 stocks are of unknown status (Table 1). Total groundfish harvests in 1996 relative to their respective ABCs are shown in (Table 2). Harvests for all species except GOA Pacific cod and BS/AI and GOA Pacific ocean perch (POP) *(Sebastes alutus)* were held at or below the respective ABC. Overruns occurred due to difficulty of in-season management of Pacific cod and POP.

The council's reliance on the plan teams and SSC is well-documented. During 1987-1997 (Table 3), the council exceeded the SSC ABC recommendation only twice in 344 ABC determinations. In 1980, the council set the AI sablefish ABC at 4,500 t, higher than the SSC recommendation of 3,700 t, but lower than the plan team recommendation of 9,600 t. In 1992, the council set the GOA pollock ABC midway between the SSC and plan team recommendations. The council set ABC lower than the SSC recommendation in five instances. At no time has the council set total allowable catch (TAC), or quota, higher than ABC. In fact, the council set TAC lower than ABC for 54% of the determinations and set TAC equal to ABC in 44% of determinations. ABCs were not specified by the council in 2% of cases; these occurred in the GOA between 1987 and 1990.

Based on current criteria in the Magnuson-Stevens Act, under which the council is authorized to manage North Pacific fish stocks, NMFS has determined that of 106 groundfish stocks under the council's jurisdiction, 0 are overfished or approaching overfished condition, 64 are not overfished, and 42 are of unknown status (NMFS 1997). Nationally, 86 stocks are overfished, 10 are approaching overfished, 183 are not overfished, and 448 are unknown. Stock status may change under revised guidelines that should be published in late 1998.

Table 1. Relative abundance (exploitable biomass in t) of Bering Sea (BS), Aleutian Islands (AI), and Gulf of Alaska (GOA) groundfish, 1996.

Species	Area	1996 biomass	Mean biomass	Reference years	Biomass relative to mean	3-year trend
Pollock	BS	6,200,000	8,540,000	1964-1996	Below	Down
	AI	230,000	688,000	1978-1996	Below	Down
	Bogoslof	680,000	1,210,000	1988-1996	Below	Stable
	GOA	574,000	1,495,600	1969-1996	Below	Down
Pacific cod	BSAI	1,106,000	1,364,000	1978-1996	Below	Up
	GOA	314,000	453,300	1978-1996	Above	Up
Deepwater flatfish	GOA	101,400	116,000	1978-1996	Below	Down
Yellowfin sole	BSAI	2,862,000	1,979,000	1960-1996	Above	Stable
Greenland turbot	BSAI	127,000	448,000	1960-1996	Below	Down
Arrowtooth flounder	BSAI	556,400	279,700	1975-1996	Above	Stable
	GOA	1,640,000	924,000	1961-1996	Above	Up
Rock sole	BSAI	2,183,000	1,187,000	1975-1996	Above	Stable
Rex sole	GOA	72,300	85,900	1990-1996	Below	Down
Flathead sole	BSAI	616,400	403,000	1975-1996	Above	Stable
	GOA	206,000	217,400	1990-1996	Stable	NK
Other flatfish	BSAI	589,500	595,000	1975-1996	Near	Stable
Shallow water flatfish	GOA	316,000	306,700	1990-1996	Stable	Down
Sablefish	BS	24,100	41,200	1979-1996	Below	Stable
	AI	24,100	54,200	1979-1996	Below	Down
	GOA	271,000	390,000	1979-1996	Below	Down
Pacific ocean perch	BS	72,500	97,300	1960-1996	Below	Down
	AI	324,000	234,000	1962-1996	Above	Up
	GOA	774,000	360,000	1984-1996	Above	Up
Sharpchin/northern	AI	96,800	96,646	1980-1996	Near	NK
Northern rockfish	GOA	85,000	73,100	1984-1996	Above	Up
Shortraker/rougheye	AI	45,600	45,616	1980-1996	Near	NK
	GOA	65,000	68,800	1984-1996	Near	Down
Pelagic shelf rockfish	GOA	78,000	56,300	1984-1996	Above	Up
Other slope rockfish	GOA	131,000	138,100	1984-1996	Near	Up
Other red rockfish	BS	29,700	NK	NK	NK	NK
Other rockfish	BS	7,100	5,200	1979-1996	NK	NK
	AI	13,600	13,600	1980-1996	NK	NK
Thornyheads	GOA	47,000	65,100	1967-1996	Below	Down
Atka mackerel	AI	576,000	699,000	1977-1996	Below	Down
	GOA	NK	NK	NK	NK	NK
Squid	BSAI	NK	NK	NK	NK	NK
Other species	BSAI	621,000	574,000	1975-1996	Above	Stable

NK = not known.

Table 2. Exploitable biomass and harvest specifications (t) of Bering Sea (BS), Aleutian Islands (AI), and Gulf of Alaska (GOA) groundfish, 1997.

Species	Area	Biomass	OFL	ABC	TAC
Pollock	BS	6,120,000	1,980,000	1,130,000	1,130,000
	AI	100,000	38,000	28,000	28,000
	Bogoslof	558,000	43,800	32,100	32,100
	GOA	1,097,830	111,270	79,980	79,980
Pacific cod	BSAI	1,590,000	418,000	306,000	270,000
	GOA	562,000	180,000	81,500	69,115
Deepwater flatfish	GOA	101,430	9,440	7,710	7,710
Yellowfin sole	BSAI	2,530,000	339,000	233,000	230,000
Greenland turbot	BSAI	118,000	22,600	12,350	9,000
Arrowtooth flounder	BSAI	587,000	167,000	108,000	20,760
	GOA	1,639,671	280,800	197,840	35,000
Rock sole	BSAI	2,390,000	427,000	296,000	97,185
Rex sole	GOA	72,330	11,920	9,150	9,150
Flathead sole	BSAI	632,000	145,000	101,000	43,500
	GOA	206,340	34,010	26,110	9,040
Shallow water flatfish	GOA	315,590	59,540	43,150	18,630
Other flatfish	BSAI	616,000	150,000	97,500	50,750
Sablefish	BS	17,900	2,750	1,308	1,100
	AI	18,600	2,860	1,367	1,200
	GOA	206,060	35,950	14,520	14,520
Other slope rockfish	GOA	103,710	7,560	5,260	2,170
Northern rockfish	GOA	83,370	9,420	5,000	5,000
Pelagic shelf rockfish	GOA	55,640	8,400	5,140	5,140
Demersal shelf rockfish	GOA	60,510	1,450	950	950
Pacific ocean perch	BS	72,500	5,400	2,800	2,800
	AI	324,000	25,300	12,8700	12,800
	GOA	242,300	19,760	12,990	9,190
Sharpchin/northern	AI	96,800	5,810	4,360	4,360
Shortraker/rougheye	AI	45,600	1,250	938	938
	GOA	65,380	2,740	1,590	1,590
Other red rockfish	BS	29,700	1,400	1,050	1,050
Other rockfish	BS	7,100	497	373	373
	AI	13,600	952	714	714
Thornyheads	GOA	46,110	2,400	1,700	1,700
Atka mackerel	AI	450,000	81,600	66,700	66,700
	GOA	NK	6,200	1,000	1,000
Squid	BSAI	NK	2,620	1,970	1,970
Other species	BSAI	688,000	138,000	25,800	25,800
	GOA	NK	NK	NK	13,470
TOTAL (all species)	BSAI	17,004,800	3,998,839	2,464,130	2,000,000
TOTAL (all species)	GOA	4,797,760	784,860	493,050	282,815
TOTAL (all species)	BOTH	21,802,560	4,783,699	2,957,180	2,282,815

OFL = overfishing level; TAC = total allowable catch; ABC = allowable biological catch; NK = not known.

Table 3. Comparison of allowable biological catch (ABC) and total allowable catch (TAC) recommendations for groundfish targets in the Gulf of Alaska (GOA) and Bering Sea/Aleutian Islands (BSAI) area.

			# of times council set:			
Year	# of species complexes	ABCs > SSC	ABCs < SSC	TACs > ABC	TACs = ABC	TACs < ABC
Gulf of Alaska						
1997	16	0	0	0	10	6
1996	16	0	1	0	8	8
1995	16	0	1	0	9	7
1994	16	0	1	0	8	8
1993	13	0	0	0	5	8
1992	13	1	0	0	7	6
1991	13	0	0	0	7	6
1990	10	0	1	0	5	4
1989	8	0	1	0	3	4
1988	8	0	0	0	4	3
1987	9	0	0	0	1	5
Total GOA	138	1	5	0	67	65
Bering Sea/Aleutian Islands						
1997	21	0	0	0	11	10
1996	21	0	0	0	4	17
1995	21	0	0	0	4	17
1994	20	0	0	0	14	6
1993	20	0	0	0	6	14
1992	20	0	0	0	13	7
1991	19	0	0	0	12	7
1990	17	1	0	0	6	11
1989	17	0	0	0	2	15
1988	15	0	0	0	5	10
1987	15	0	0	0	8	7
Total BSAI	206	1	0	0	85	121
TOTAL	344	2	5	0	152	186

From the North Pacific Fishery Management Council and its Scientific and Statistical Committee (SSC), 1987-1997, based on one allowable biological catch (ABC) per species or complex. Some rows do not sum as reported as a result of unspecified ABCs in some years for Atka mackerel, DSR, and other rockfish.

POP = Pacific ocean perch, OFD= overfishing definition; DSR = demersal shelf rockfish.

Table 4. Stock assessment models used by NMFS for North Pacific groundfish, 1986-1997.

BSAI	Pollock	Pacific cod	Sablefish	Atka mackerel	Arrowtooth	Pacific ocean perch	Greenland turbot	Rock sole	Other flatfish	Yellowfin sole
97	Ch, Ca, SS	SS	SS	SS	SS	SS	SS	SS	SS	SS
96	Ch, Ca, SS	SS	SS	SS	SS	SS	SS	SS	SS	SS
95	Ch, Ca, SS	SS	SRA	SS	SS	SS	SS	SS	None	SS
94	Ca, Ch	SS	SRA	SS	YPR	SS	SS	SS	None	SS
93	Ca, Ch	SRA	SRA	SS	YPR	SS	SRA	SS	YPR	SS
92	Ca, Ch	SRA	SRA	SS	YPR	SRA	SRA	S Index	YPR	SS
91	Ca, Ch	SRA	SRA	Ch	YPR	SRA	SRA	S Index	YPR	S Index
90	Ca, Ch	SRA	SRA	Ch	YPR	SRA	SRA	S Index	YPR	S Index
89	Ca, Ch	SRA	SRA	Ch	YPR	SRA	SRA	S Index	YPR	S Index
88	Ca, Ch	ASA	SRA	None	YPR	SRA	SRA	S Index	YPR	S Index
87	Ca, Ch	ASA	None	None	None	SRA	SRA	S Index	None	S Index
86	Ch	ASA	GPM	None	None	SRA	None	S Index	None	S Index

GOA	Pollock	P. cod	Sablefish	Atka mackerel	Arrowtooth	POP/slope rockfish	Thornyheads
97	SS	SS	SRA	None	SS	SS	SS
96	SS	SS	SRA	None	SS	SS	SS
95	SS	SS	SRA	None	SS	SS	SS
94	SS	SRA	SRA	None	In flatfish	SS	YPR
93	SS	SRA	SRA	None		SS	YPR
92	SS	SRA	SRA	In other Species		SS	YPR
91	SS	SRA	SRA			SRA	YPR
90	SS	SRA	SRA			SRA	YPR
89	SRA	SRA	SRA			SRA	YPR
88	SRA	YPR	SRA			SRA	None
87	None	PYM	SPM			SRA	None
86	ASA	PYM	SPM			SRA	None

ASA = Age-structured analysis, Ca = CAGEAN, Ch = Cohort, GPM = Growth production model, PYM = Production yield model, S. Index = Survey index, SPM = Stock production model, SRA = Stock reduction analysis, SS = Stock Synthesis, YPR = Yield-per-recruit analysis.

Table 5. Current research surveys conducted for Bering Sea (BS), Aleutian Islands (AI), and Gulf of Alaska (GOA) groundfish, by species and area.

Species	Area	Trawl survey	Acoustic survey	Longline survey
Pollock	BS	Annual	Triennial	
	AI	Triennial		
	Bog	–	Annual	
	GOA	Triennial	Annual	
Pacific cod	BSAI	Annual		
	GOA	Triennial		
Yellowfin sole	BSAI	Annual		
	GOA	Triennial		
Greenland turbot	BSAI	Annual		
	GOA	Triennial		
Arrowtooth	BSAI	Annual		
	GOA	Triennial		
Rock sole	BSAI	Annual		
	GOA	Triennial		
Flathead sole	BSAI	Annual		
	GOA	Triennial		
Other flatfish	BSAI	Annual		
	GOA	Triennial		
Sablefish	BS	–	–	Biennial
	AI	–	–	Biennial
	GOA	–	–	Annual
P. ocean perch	BS	Annual		
	AI	Triennial		
	GOA	Triennial		
Sharp/northern	AI	Triennial		
	GOA	Triennial		
Short/rougheye	AI	Triennial		
	GOA	Triennial		
Red rockfish	BS	Annual		
Other rockfish	BS	Annual		
	AI	Triennial		
	GOA	Triennial		
Atka mackerel	AI	Triennial		
	GOA	Triennial		
Squid	BSAI	Annual		
Other species	BSAI	Annual		
	GOA	Triennial		

Table 6. Life history characteristics for BSAI/GOA groundfish used in 1997 stock assessments.

		Growth parameters			Maturity indicators		Weight parameters	
Species	Area	*M*	L_{inf}	*k*	$L_{50\%}$	$A_{50\%}$	Alpha	Beta
Pollock	BS	0.30	59.0	0.228	NK	NK	$1.14\mathrm{E}^{-5}$	2.877
	AI	0.30	52.8	0.368	NK	NK	$2.73\mathrm{E}^{-5}$	2.651
	Bogoslof	0.20	55.7	0.171	NK	NK	$1.29\mathrm{E}^{-6}$	3.436
	GOA	0.30	56.2	0.328	NK	NK	$1.27\mathrm{E}^{-5}$	2.885
Pacific cod	BSAI	0.30	98.2	0.227	67	5.7	$5.29\mathrm{E}^{-6}$	3.206
	GOA	0.37	120.0	0.119	67	6.6	$5.29\mathrm{E}^{-6}$	3.206
Rex sole	GOA	0.20	59.5	0.20	NK	NK	$4.46\mathrm{E}^{-3}$	3.471
Dover sole	GOA	0.10	NK	NK	NK	NK	NK	NK
Yellowfin sole	BSAI	0.12	35.8	0.147	30	10.5	$9.72\mathrm{E}^{-4}$	3.056
	GOA	0.20	34.0	0.18	NK	NK	$6.68\mathrm{E}^{-3}$	3.18
Greenland turbot	BSAI	0.18	NK	NK	60	9.0	$2.69\mathrm{E}^{-6}$	3.309
Arrowtooth	BSAI	0.20	59.0	0.170	NK	NK	$5.68\mathrm{E}^{-6}$	3.103
	GOA	0.20	59.6	0.17	47	NK	$3.92\mathrm{E}^{-3}$	3.223
Rock sole	BSAI	0.20	45.1	0.180	NK	NK	$7.61\mathrm{E}^{-3}$	3.120
	GOA	0.20	38.8	0.21	NK	NK	$9.98\mathrm{E}^{-3}$	3.047
Flathead sole	BSAI	0.20	42.6	0.165	NK	NK	$3.96\mathrm{E}^{-3}$	3.259
	GOA	0.20	29.9	0.49	NK	NK	$4.06\mathrm{E}^{-3}$	3.237
Other flatfish	BSAI	0.20	72.2	0.053	NK	NK	$8.84\mathrm{E}^{-3}$	3.111
Sablefish	BS	0.10	70.7	0.275	NK	4	$3.23\mathrm{E}^{-3}$	3.294
	AI	0.10	77.6	0.206	NK	4	$3.23\mathrm{E}^{-3}$	3.294
	GOA	0.10	89.3	0.142	65	4	$2.99\mathrm{E}^{-3}$	3.30
Pacific ocean perch	BS	0.05	39.9	0.135	NK	NK	$1.19\mathrm{E}^{-5}$	3.037
	AI	0.05	39.6	0.167	NK	NK	$1.22\mathrm{E}^{-5}$	3.030
	GOA	0.02-0.08	44.8	0.088	10.5	NK	$1.54\mathrm{E}^{-5}$	2.96
Sharpchin/ northern	AI	0.06	NK	NK	NK	NK	NK	NK
	GOA	0.05/0.06	34.9/35.6	.095/.190	NK	NK	$1.63\mathrm{E}^{-5}$	2.98
Shortraker/ rougheye	AI	0.03	NK	NK	NK	NK	NK	NK
	GOA	0.027-0.042	54.7	0.050	NK	NK	NK	NK
Other red rockfish	BS	NK	NK	NK	NK	NK	NK	NK
Demersal shelf rockfish	GOA	0.02	68.9	0.053	52	21	$4.35\mathrm{E}^{-6}$	3.396
Thornyheads	GOA	0.07	NK	NK	22	NK	$1.36\mathrm{E}^{-6}$	3.390
Other rockfish	BS	0.07	NK	NK	NK	NK	NK	NK
	AI	0.07	NK	NK	NK	NK	NK	NK
	GOA	0.01-0.07	NK	NK	NK	NK	NK	NK
Atka mackerel	AI	0.30	43.5	0.449	31.1	3.6	$5.05\mathrm{E}^{-6}$	3.240
	GOA	0.30	47.3	0.610	38.3	3.6	$1.55\mathrm{E}^{-5}$	2.979
Squid	BSAI	NK	NK	NK	NK	NK	NK	NK

Includes natural mortality rate *(M),* length and age at 50% maturity (females), growth parameters (L_{inf} and k or von Bertalanffy equation where $L = L_{inf}\{[_1 - \exp(-k(t - t_0)]\}$, and weight parameters ($W = \text{alpha} \times L^{\text{beta}}$) for both sexes combined. Length is measured in centimeters (cm) and weight in grams (g).

NK= not known.

Table 7. Tiers used to determine allowable biological catch (ABC) and overfishing level (OFL) for North Pacific groundfish stocks.

---

(1) Information available: Reliable point estimates of B and B_{MSY} and reliable pdf of F_{MSY}.

- (a) Stock status: $B/B_{MSY} > 1$
  - $F_{OFL} = m_A$, the arithmetic mean of the pdf
  - $F_{ABC} \le m_H$, the harmonic mean of the pdf
- (b) Stock status: $a < B/B_{MSY} \le 1$
  - $F_{OFL} = m_A \times (B/B_{MSY} - a)/(1 - a)$
  - $F_{ABC} \le m_H \times (B/B_{MSY} - a)/(1 - a)$
- (c) Stock status: $B/B_{MSY} \le a$
  - $F_{OFL} = 0$
  - $F_{ABC} = 0$

(2) Information available: Reliable point estimates of B, B_{MSY}, F_{MSY}, $F_{30\%}$, and $F_{40\%}$.

- (a) Stock status: $B/B_{MSY} > 1$
  - $F_{OFL} = F_{MSY} \times (F_{30\%}/F_{40\%})$
  - $F_{ABC} \le F_{MSY}$
- (b) Stock status: $a < B/B_{MSY} \le 1$
  - $F_{OFL} = F_{MSY} \times (F_{30\%}/F_{40\%}) \times (B/B_{MSY} - a)/(1 - a)$
  - $F_{ABC} \le F_{MSY} \times (B/B_{MSY} - a)/(1 - a)$
- (c) Stock status: $B/B_{MSY} \le a$
  - $F_{OFL} = 0$
  - $F_{ABC} = 0$

(3) Information available: Reliable point estimates of B, $B_{40\%}$, $F_{30\%}$, and $F_{40\%}$.

- (a) Stock status: $B/B_{40\%} > 1$
  - $F_{OFL} = F_{30\%}$
  - $F_{ABC} \le F_{40\%}$
- (b) Stock status: $\mathrm{a} < B/B_{40\%} \le 1$
  - $F_{OFL} = F_{30\%} \times (B/B_{40\%} - a)/(1 - a)$
  - $F_{ABC} \le F_{40\%} \times (B/B_{40\%} - a)/(1 - a)$
- (c) Stock status: $B/B_{40\%} \le a$
  - $F_{OFL} = 0$
  - $F_{ABC} = 0$

(4) Information available: Reliable point estimates of B, $F_{30\%}$, and $F_{40\%}$.

$F_{OFL} = F_{30\%}$
$F_{ABC} \le F_{40\%}$

(5) Information available: Reliable point estimates of B and natural mortality rate M.

$F_{OFL} = M$
$F_{ABC} \le 0.75 \times M$

(6) Information available: Reliable catch history from 1978 through 1995.

OFL = the average catch from 1978 through 1995, unless an alternative value is established by the SSC on the basis of the best available scientific information
ABC ≤ 0.75 × OFL

F = fishing mortality; M = natural mortality; B = biomass; MSY = maximum sustainable yield; OFL = overfishing level; ABC = allowable biological catch; SSC = scientific and statistical committee; pdf = probability density function.

The council has responded to plan team and SSC recommendations for enhanced groundfish management by separating or grouping appropriate species within the management unit of the FMPs. When a directed fishery for black rockfish began developing in the GOA, the GOA plan team informed the council that the NMFS bottom trawl survey did not adequately assess black *(Sebastes melanops)* and blue *(S. mystinus)* rockfishes and recommended that the council separate those rockfishes from the pelagic shelf rockfish (PSR) management grouping to prevent overexploitation under the large overall PSR TAC. In 1997, the council submitted a plan amendment to do so, and deferred their management to the State of Alaska (NPFMC 1997). This action was approved by the Secretary of Commerce and implemented in 1998. GOA arrowtooth flounder was separated from GOA flatfish in 1995; GOA Atka mackerel was separated from GOA other species in 1993.

The history of assessment models for the BSAI and GOA for species/complexes with sufficient biological information since 1986 is presented in Table 4. Trawl, longline (sablefish), and hydroacoustic (pollock) research surveys collect required biological information to implement the models (Table 5). Life history information is summarized in Table 6.

The council also employs a precautionary approach in applying stock assessment results to commercial fisheries. New assessment models and modifications to existing models are reviewed and phased-in only after considerable peer review by the plan teams and SSC. New models are generally introduced in one year, but do not replace the existing model until tested and reviewed. In its December 1996 meeting minutes, the SSC requested that beginning in 1997, stock assessment authors incorporate: (1) biomass and yield projections for an $F_{40\%}$ harvest strategy under varying assumptions regarding recruitment and for other relevant exploitation rates; (2) standard errors or confidence intervals for important model parameters; (3) sensitivity analyses for key parameters and input assumptions; (4) weightings given to individual data components with justifications; and (5) risk analyses.

In the GOA SAFE report for 1998, Hollowed et al. (1997) included an appendix to the walleye pollock chapter that incorporated predation mortality by arrowtooth flounder, Pacific halibut, and Steller sea lion as a fishery type into the stock synthesis model. A second appendix demonstrated the use of the AD Model Builder software in the GOA pollock stock assessment, which facilitates the rapid development of nonlinear statistical models. This modeling software may also be used in the near future for GOA sablefish and rockfish species.

Conclusions

In the future, assessments will continue to be further developed to more explicitly account for uncertainty and provide more precautionary advice, and management likely will become more precautionary and risk averse.

Risk averse management is most feasible where effective management practices are in place, for example, the North Pacific. Recent advances in management, such as the NPFMC quota tier system (Table 7), suggest that under a risk adverse policy, greater uncertainty in important model parameters will generally result in lower quotas. Increasing complexity in stock assessment models directly results from improved data collection and modeling software, along with incorporation of ecosystem interactions and uncertainty. Faced with increased complexity in the stock assessment modeling, management agencies will rely increasingly on its scientific advisors.

Acknowledgments

Appreciation is extended to Sandra Lowe, Nancie Cummings, and one anonymous reviewer for helpful comments on the manuscript.

References

Clark, W.G. 1993. The effect of recruitment variability on the choice target level of spawning biomass per recruit. In: G. Kruse, D.M. Eggers, R.J. Marasco, C. Pautzke, and T.J. Quinn (eds.), Management strategies for exploited fish populations. University of Alaska Sea Grant, AK-SG-93-02, Fairbanks. 825 pp.

Hollowed, A.B., E. Brown, J. Ianelli, P. Livingston, B. Megrey, and C. Wilson. 1997. Walleye pollock. Appendix 2. An alternative software program for modeling GOA pollock. In: Preliminary stock assessment and fishery evaluation report for the groundfish resources of the Gulf of Alaska. September 1997. North Pacific Fishery Management Council, 605 W. 4th Avenue, Suite 306, Anchorage, AK 99501. 261 pp.

Methot, R.D. 1990. Synthesis model: An adaptable framework for analysis of diverse stock assessment data. Int. North Pac. Fish. Comm. Bull. 50:259-277.

NMFS. 1997. Report to Congress. Status of fisheries of the United States. NMFS, OST, 1315 East-West Hwy., Silver Spring, MD 20910. 75 pp.

NPFMC. 1997. Environmental assessment/regulatory impact review for GOA amendment 46 to revise management authority of pelagic shelf rockfish. North Pacific Fishery Management Council, 605 W. 4th Avenue, Suite 306, Anchorage, AK 99501.

Harvest Control for Schooling Fish Stocks under Cyclic Oceanographic Regimes: A Case for Precaution and Gathering Auxiliary Information

Marcelo Vasconcellos and Tony Pitcher
University of British Columbia, Vancouver, B.C., Canada

Abstract

This paper examines the influence of three aspects of the population dynamics of small pelagic fish stocks on fisheries harvest control: (1) the non-stationarity in productivity driven by cyclic environmental regimes; (2) the collapse in distribution range following stock collapse; and (3) the consequent change in catchability with stock size. The suitability of three harvest control options, including effort limitation, minimum age in the catch, and constant escapement, was evaluated with a simulation model and according to contrasting levels of uncertainty on population parameters. Harvest control by constant effort is particularly inadequate in avoiding catchability-lead stock collapse for populations which exhibit cyclic annual production. Overall, minimum escapement policies offer the best tradeoff between mean catch, catch variability, and rebuilding capacity of the stock. In cases of highly uncertain production rates, the estimation of harvest control measures can be hampered by the non-stationarity of population parameters and by the lack of contrast in catch rates. Resolution of complications from variable catchability requires auxiliary information on abundance trends and, more important, demands the adoption of more cautious harvest control policies. For pelagic stocks dominated by environmental regimes, the protection of the first spawning age classes seems to be an appropriate measure to increase resilience to overfishing.

Introduction

The notion of regimes in marine ecosystems has become increasingly important for fisheries management advice as evidence of inter-decade changes in fish stock productivity and species switches accumulates (Lluch-Belda et al. 1989, Bakun 1996, Steele 1996). In the case of small pelagic fish populations, such as anchovies and sardines inhabiting major upwelling systems of the world, cyclic "boom and bust" dynamics is an widely recognized characteristic (Caddy and Gulland 1983, Csirke 1995) that may be explained by alternative non-exclusive hypotheses concerning climatic effects on recruitment, trophic interactions, and overfishing (Radovich 1982, Lluch-Belda et al. 1989, Bakun 1996).

Fisheries assessment research has based advice on yields and the risk of overfishing based on species life history characteristics, population dynamics parameters, and associated observation and process errors. Alternative harvest policies are normally evaluated in simulation models where average yield, catch variability, and the risk of crossing biological thresholds are compared using the best estimates of key population parameters (Punt and Hilborn 1997). However, little attention has been given to uncertainties on the degree of control over the harvest accompanying a particular management decision, i.e., on the implementation of harvest policies. Nevertheless, harvest control constitutes most of the regulatory work, demands a good proportion of management resources, and can ultimately determine the probabilities of stock collapse. Fisheries for small pelagic schooling fish stocks, for instance, are characterized by an increase in catchability with decrease in biomass as a result of shoaling behavior (Clark 1974), range collapse and highly efficient fishing fleets (Pitcher 1995, 1997). For such species the lack of control on effort and harvest rates alone can generate a pathological response in catches and bio-economics (Mackinson et al. 1997), and lead to stock collapse. These two sets of factors, loosely described as environmental and behavioral, have rarely been considered together.

In this paper we analyze the likely benefits of different harvest control policies for small pelagic schooling fish stocks in a cyclic ocean regime affecting recruitment. Specifically, we explore the case where changes in catchability with stock size following range collapse make harvest control by effort limitation potentially difficult. In this context we discuss the role that auxiliary information may have in developing more efficient fishing control systems, and the precautionary measures that should be adopted in the face of uncertainties associated with cyclic environmental regimes.

Methods

The Population Model

A deterministic age-structured model was used to simulate the number, biomass, and catches from a population with dynamics determined by the equations:

$$
\begin{aligned}
N_{0.5,t} &= R_t \\
N_{i,t+0.5} &= N_{i,t}e^{-Z_{i,t}} \\
Z_{i,t} &= F_{i,t} + M \\
F_{i,t} &= f_t q_t V_i \\
V_i &= \frac{i^b}{A_{50\%}^b + i^b} \\
C_{i,t} &= \frac{F_{i,t}}{Z_{i,t}}[1 - e^{-Z_{i,t}}]N_{i,t} \qquad (1) \\
Y_t &= \sum_{i=0.5}^{3.5} C_{i,t}W_i
\end{aligned}
$$

where $N_{i,t}$ is the number of individuals of age i in year t, $Z_{i,t}$ is the total instantaneous annual base mortality rate of fish of age i in year t, $F_{i,t}$ is the instantaneous annual fishing mortality rate. $F_{i,t}$ is a function of the fishing effort f in year t, the catchability coefficient q and V_i, the age-specific vulnerability to the fishery. V_i is a function of the age at 50% vulnerability ($A_{50\%}$) and a concentration parameter ($b = 10$). M is the constant instantaneous annual natural mortality rate set as 1.0 year^{-1}, $C_{i,t}$ is the catch of the number of fish of age i in year t and Y_t the total annual yield in weight, W_i.

$N_{0.5,t}$ represents the number of recruiting fish of age 0.5 in year t. Recruitment was made a function of spawning stock biomass using a Beverton and Holt type function:

$$R_t = \frac{aS_{t-1}}{1 + \frac{aS_{t-1}}{b}}$$

where a is the maximum recruits per spawning biomass S when S is very small and b is the maximum number of recruits when S is very large. To reflect a population with high intrinsic rate of growth ($k = 0.55$) the parameter a was adjusted to 0.2 and b scaled to reflect the recruitment history of a known stock, *Sardinella brasiliensis* (Cergole 1993). To simulate cyclic changes in recruitment, a was varied as a sine function of time with period of 20 years (Fig. 1). Data on mean body weight (W_i) and the percentage of individuals mature at age for *Sardinella brasiliensis* (Cergole 1993) (Table 1) was incorporated into the model to represent growth and maturity processes. Table 1 also lists the estimated reproductive output A of a given age class. The reproductive output was used to weight the relative reproductive

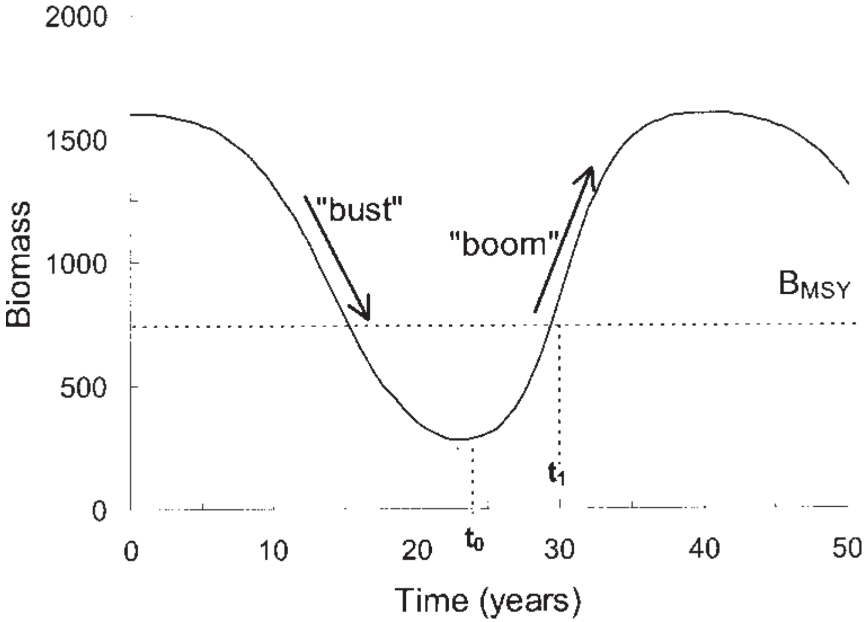


Figure 1. *Simulation of population with a cycle regime in productivity (period 20 years).* B_{MSY} *is the biomass at the maximum sustainable yield in equilibrium conditions. Time difference* $(t_1 - t_0)$ *is the time it takes for the stock to recover to* B_{MSY} *after the end of a "bust" phase.*

contribution of an age class by accounting for differences in relative fecundity and frequency of spawning. Parameter A was calculated using the equation proposed for the northern anchovy stock, *Engraulis mordax* (Parrish et al. 1986). These parameters were used in the analysis of spawning per recruit:

$$SPR_{A_{50\%},E} = \sum_{i=0}^{3.5} (l_i m_i A_i)$$

where $SPR_{A50\%,E}$ is the spawning per recruit value for a given combination of age at first capture ($A_{50\%}$) and exploitation rate [$E = F/Z$;(0,1)], l_i is the survivorship to age i [$l_i = l_0 s(1 - V_i E)$], l_0 is the survivorship to age 0(=1), $s = \exp^{(-Mi)}$, M is the instantaneous natural mortality rate, V_i is the age specific vulnerability from equation (1), and A_i the age-specific reproductive output. *SPR* is expected to decrease with E, and is expressed as the percentage of the value calculated for an unfished stock (%*SPR*) (Goodyear 1993).

Three classes of harvest strategies were tested: constant effort (= constant harvest rate with constant catchability), constant escapement, and

Table 1. Mean body weight, mature percentage and reproductive output (*A*) at age for a small pelagic stock.

Age	0	0.5	1	1.5	2	2.5	3	3.5
W (g)	8	32	42	45	58	70	84	110
% Mature	0	0	50	75	100	100	100	100
A	0	0	0.83	1.00	1.95	3.21	5.19	10.58

The reproductive output (*A*) was calculated using the equation proposed by Parrish et al. (1986) to *E. mordax*. The equation was adjusted so that the relationship $A_i = 0.0000432 W_i^{2.64}$ generates a factor $A = 1$ when $W = 45$g, the approximate mean weight at first maturity.

minimum age in the catch. Two harvest rate targets were tested under constant effort: the fishing mortality rate ($F_{Patterson}$) that produces an exploitation rate (F/Z) of 0.4, suggested for pelagics by Patterson (1992), and the fishing mortality rate (F_{opt}) that produces the maximum sustainable yield in equilibrium conditions. The later was calculated by running the model for a period long enough to stabilize yield and biomass at each of a range of *F* values, and then finding the maximum yield across *F* values. Constant escapement policies include the thresholds of 20% and 40% of the virgin stock biomass (B_{inf}) (Goodyear 1989), and assume that the stock is exploited at F_{opt} when biomass is above this threshold. The minimum age in the catch strategy considered the age at first capture ($A_{50\%}$) as the age at first maturity (1.5 years), and assume that the stock is exploited at high constant fishing rates.

To evaluate harvest control options we perform three analyses, in order of decreasing amount of information and increasing uncertainty:

1. First we evaluate the performance of the different management strategies when production parameters (*a*, *b*, and *M*) are known and harvest is efficiently controlled. Strategies are compared in terms of mean catch, catch variance, and the proportion of years with no catch in a 50-year period. Also, we analyze the time it takes for the stock to recover to B_{MSY} (biomass level that would produce the maximum sustainable yield in equilibrium conditions) after a bust phase (Fig. 1). Recovery time is considered a key aspect for managing spasmodic populations that have evolved in systems with periodic disturbances (Beddington and May 1977, Hilborn and Walters 1992).

2. Second we evaluate the effect of implementation uncertainties, due to changes in catchability with stock size, on the performance of the different harvest strategies. Production parameters are assumed known, and harvest is controlled by maintaining effort at the level that would produce the optimal harvest rates in situations of constant catchability. Shoaling and schooling behavior makes small pelagic fishes more

easily detected by modern purse seiners, and almost equally vulnerable to fishing gear regardless of stock density (Pitcher 1995). For these species, the catchability (q) can increase in inverse proportion to abundance:

$$q = q_1 B^{-q2} \tag{2}$$

where q_1 is a proportionality constant and q_2 is the degree to which catchability increases with declining stock size. Two values of q_2 were used to describe the pattern of change in catchability with stock size: 0.4, the value estimated by MacCall (1976) for the California sardine; and 1.0, approximately the value estimated by Csirke (1989) for the Peruvian anchoveta. The latter produces a more drastic scenario where catch rates in the fishery remain constant regardless of the actual fish stock size (Pitcher 1995, Mackinson et al. 1997).

3. Finally, we evaluate a situation where production parameters are not well known, and historical data on catch and effort are used to make inferences about the productivity parameters of the stock (k and B_{inf}) and the harvest control measures (optimum effort and escapement levels). An estimation procedure was designed to infer the population dynamics parameters of a Schaefer surplus production model of the type:

$$B_{t+1} = B_t + kB_t(1 - \frac{B_t}{B_{inf}}) - B_t q_t f_t \tag{3}$$

where B_t is the population biomass at time t, k is the intrinsic rate of growth, B_{inf} is the virgin stock biomass, q_t is the catchability coefficient at time t and f_t the fishing effort at time t. Catch and catch per unit of effort (U_t) data obtained from the age structured population model in a 20-year period of increasing effort (Fig. 5) were used to estimate the parameters k and B_{inf} with a maximum likelihood statistical procedure (Polacheck et al. 1993, Walters and Ludwig 1994). From initial inferences of B_{inf} and k, and the observed catches, the difference equation in (3) is used to produce predicted stock biomass values (B_t) for each year. With the predicted biomass values and the observed catch rates (U_t), from the simulation model with a lognormal error (i.e., $U_t = q_t B_t e\varepsilon$, $e \sim N[0;\sigma^2]$), a maximum likelihood estimator of q is calculated from

$$\hat{q} = \frac{\sum_{t=1}^{20} U_t B_t}{\sum_{t=1}^{20} B_t^{\,2}}$$

A time series of estimated catch rates U_t' is obtained assuming $U_t' = \hat{q}B_t$. The likelihood of a given combination of parameters B_{inf} and k is then calculated by

$$L_{Binf,k} = \left[\frac{\sum_{t=1}^{20} (U_t - \hat{q}B_t)^2}{n-1} \right]^{\frac{-(n-1)}{2}}$$

This expression provides the likelihood of B_{inf} and k by treating the variance term and the catchability (q) as nuisance parameters (Walters and Ludwig 1994). Bayesian posterior probability calculations were carried out assuming a uniform prior probability distribution for both parameters. The posterior probability distribution of a given parameter combination was then calculated as the corresponding $L_{Binf,\,k}$ value divided by the sum of $L_{Binf,k}$s for a grid of parameter combinations. Marginal posterior probability density for a single parameter was computed by integrating out the complementary parameter, i.e., by summing posterior probabilities for one parameter across all values of the other parameter. Two types of harvest control measures can be estimated from the surplus production model: optimal escapement levels (20 or 40% B_{inf}), and the optimal fishing effort f_{opt}, estimated as $k/2q$.

Results

Table 2 shows the performance of each strategy for a 50-year simulation period. For the minimum age strategy, simulation was carried out arbitrarily under high fishing pressure (Effort = 200 ~ $F = 2$ year^{-1} when q is constant). Protecting a minimum stock biomass by the constant escapement strategy yields a higher average catch, and allows a rapid recovery to B_{MSY} levels when a favorable period prevails. Nonetheless, maintaining a constant escapement produces a high probability of years with the fishery closed (on average, catches are null in 25 out of 50 years). Constant effort strategies considerably increase the time it takes for the stock to recover and produce relatively high long-term yield. Protecting the age at first maturity allows a relatively rapid stock recovery even under high fishing pressure, and produces the lower variance in catches among the strategies tested. Maintaining the stock under high fishing mortality, however, costs a noticeable decrease in mean catch.

When catchability changes with stock size (Table 3; Fig. 2), managing the stock at the constant effort levels previously used (Table 2) lead to collapse during the "bust" phase (Fig. 2c). The optimum constant fishing effort that could be sustained is less than 10% of the value calculated for

Table 2. Performance of three harvest control strategies in a constant catchability scenario (q = 0.01).

Strategies	Targets	Mean catch	C.V.	Recovery time	P
Const. effort	$F_{Patterson}$	121.9 (66.0)	1.06	17.0	0
	F_{opt}	122.0 (74.0)	1.11	17.5	0
Minimum age	$A_{50\%}$ 1.5	112.5 (200.0)	0.95	15.0	0
Const. escapement	$20\%B_{inf}$	134.7	1.03	14.0	0.46
	$40\%B_{inf}$	131.3	1.11	12.0	0.53

The values in brackets are the fishing efforts used in the simulations. C.V. is the coefficient of variation in catches and P is the proportion of the 50 year period with no catches. Recovery time measures the time (years) it takes the stock to rebuild its biomass to B_{MSY} levels after a "bust" phase (see Fig. 1 for explanation).

the constant catchability scenario (Table 2). In order to maintain a constant harvest rate, effort has to dramatically change from the "boom" to the "bust" phase so as to balance the increase in catchability (Fig. 3). For the extreme situation described by Csirke (1989) (q_2 = 1), where catch rates are independent of population size, the required annual change in effort to sustain the stock is proportional to the change in biomass between years, a dreadful scenario for any manager attempting to optimally control harvest every year without previous information on stock biomass. Higher average catches are obtained with the constant escapement strategy, at the cost of high variability in catches and a consequent high probability of years with the fishery closed (Table 3). The minimum age in the catch strategy is robust to catchability changes with stock size (Table 3). Nevertheless, under high fishing pressure, the overfished stock takes ca. 30% more time to recover to B_{MSY} and the mean catch is reduced by about 20%, compared to the same strategy in Table 2.

The robustness of minimum age policy to stock collapse is explained by the life history characteristics of the population. Figure 4a shows the reproductive potential of the stock (measured as %SPR) for different combinations of age at first capture ($A_{50\%}$) and exploitation rate (E). When the first spawning age class is effectively protected the total reproductive potential of a cohort is only slightly reduced with increasing fishing pressure. For A (reproductive output) constant with age (Fig. 4b), protecting the first spawning age class considerably decreases the impact of fishing mortality on the reproductive potential. The combination of high natural mortality rate and early maturity makes the stock highly dependent on the first spawning age classes, which make up the bulk of the reproduc-

Table 3. Performance of harvest control strategies in a variable catchability scenario (q_2 = 0.4 or 1.0, see equation 2).

Strategies	Targets	Mean catch		C.V.		Recovery time		P	
		q_2		q_2		q_2		q_2	
		0.4	1.0	0.4	1.0	0.4	1.0	0.4	1.0
Const. effort	F_{opt}	61.7 (6.0)	42.1 (2.2)	0.55	0.13	13.5	15.0	0.0	0.0
Minimum age	$A_{50\%}$ 1.5	92.7 (200)	85.6 (200)	1.16	1.30	19.0	20.5	0.0	0.0
Const. escapement	20%B_{inf}	143.6	139.7	1.08	1.24	15.5	20.0	0.50	0.61
	40%B_{inf}	140.8	153.8	1.16	1.21	12.5	16.0	0.57	0.63

Values in brackets are the fishing efforts used in the simulations. C.V. is the coefficient of variation in catches and P is the proportion of the 50 year period with no catches. Recovery time measures the time (years) it takes the stock to rebuild its biomass to B_{MSY} levels after a bust phase (see Fig. 1 for explanation).

tive capacity in many small pelagic stocks (Pauly and Soriano 1987) (Fig. 4c, *Sardinella brasiliensis*).

Parameter Uncertainties and Harvest Control

The estimation of population parameters and harvest control measures was evaluated in four different scenarios (Fig. 5). In each scenario catch rate data (cpue) was obtained with a simulated fishery consisting of 20 years of continuously increasing fishing effort. Scenarios A and B represent situations where the population is not subject to long-term cycles in productivity. In scenario B, catchability was considered a function of stock biomass (q_2 = 0.4; equation [2]). Scenarios C and D depict situations of long-term cycles in productivity, being fishery data acquired respectively at the declining ("bust") and recovery ("boom") phase of the cycle. In case A, data seems to provide some information about the virgin stock size (B_{inf}) and the population intrinsic growth rate (k) (Fig. 6), but, most important, any attempt to estimate parameter values will show that the standard deviation for each parameter is as large as the parameter estimate. The effect of including a density-dependence response for catchability (scenario B) and declining production rates (scenario C) cause a consistent underestimation of k and an overestimation of the virgin stock size, i.e., the decline in catch rates is, in both cases, best described as the result of a continuous depletion of a larger and unproductive stock. When the fishery develops under a favorable period ("boom" cycle, scenario D) catch rate data leads to an overestimation of both B_{inf} and k, depicting a large and very productive stock apparently not impacted at all by the continuous increase in fishing rate. Harvest control targets calculated from these

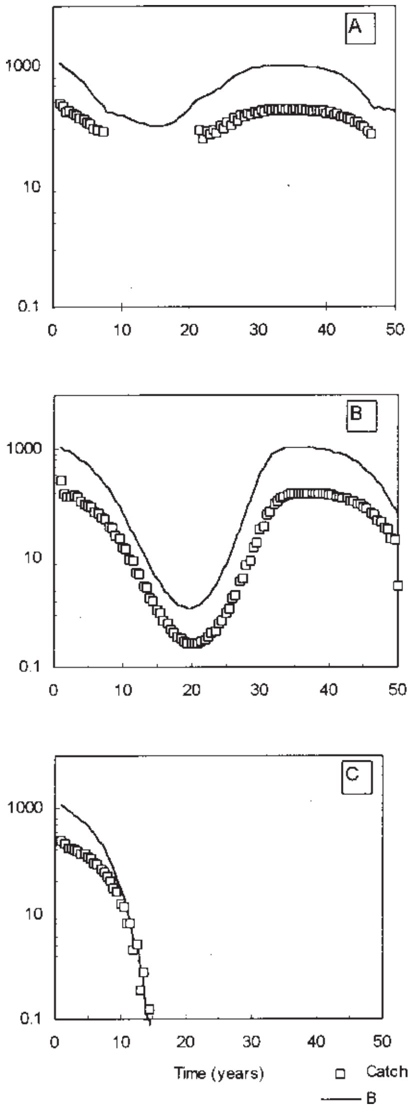


Figure 2. Catch and biomass for harvest control options when catchability is a function of stock size ($q_2 = 0.4$): (A) constant escapement; interruptions in catch history represent periods of fishery closed because of policy; (B) constant effort, age limit ($A_{50\%} = 1.5$); and (C) constant effort, no age limit. Left axis in arbitrary units.

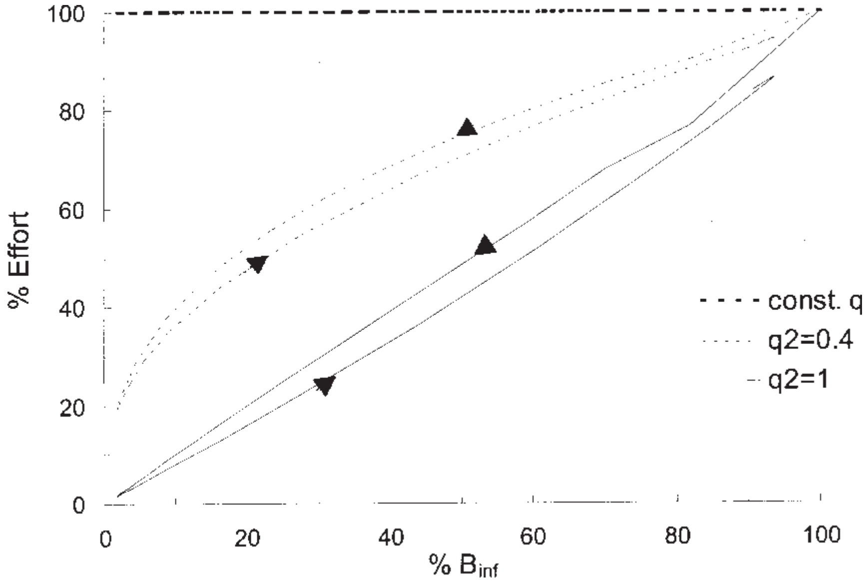


Figure 3. Adjustment in effort needed to balance the increase in catchability with stock size and maintain a constant harvest rate. Effort was calculated as (F/q_1B^{-q2}) in the simulation of a "bust" and "boom" cycle, and expressed as a percentage of the effort value at B_{inf}. Biomass is expressed as the percentage of B_{inf}.

parameters have therefore biases that involve, for instance, the overestimation of virgin stock size (scenarios B, C, and D) and hence of management escapement goals, and the underestimation (scenarios A, B, and C) and overestimation (scenario D) of optimal harvest rates. For all situations analyzed, the most useful information regarding harvest control measures is that they are highly uncertain.

Discussion

This work examines the influence of three aspects of the population dynamics of small pelagic fish stocks on fisheries harvest control:

1. Small pelagic stocks undergo multi-year regimes in productivity driven by environmental processes (Bakun 1996).

2. Schooling behavior mechanisms, presumably evolved to optimize tradeoffs among feeding, predation avoidance, and spawning in the pelagic environment, influence the mesoscale distribution of fish schools, which tend to maintain their size and inter-school distance

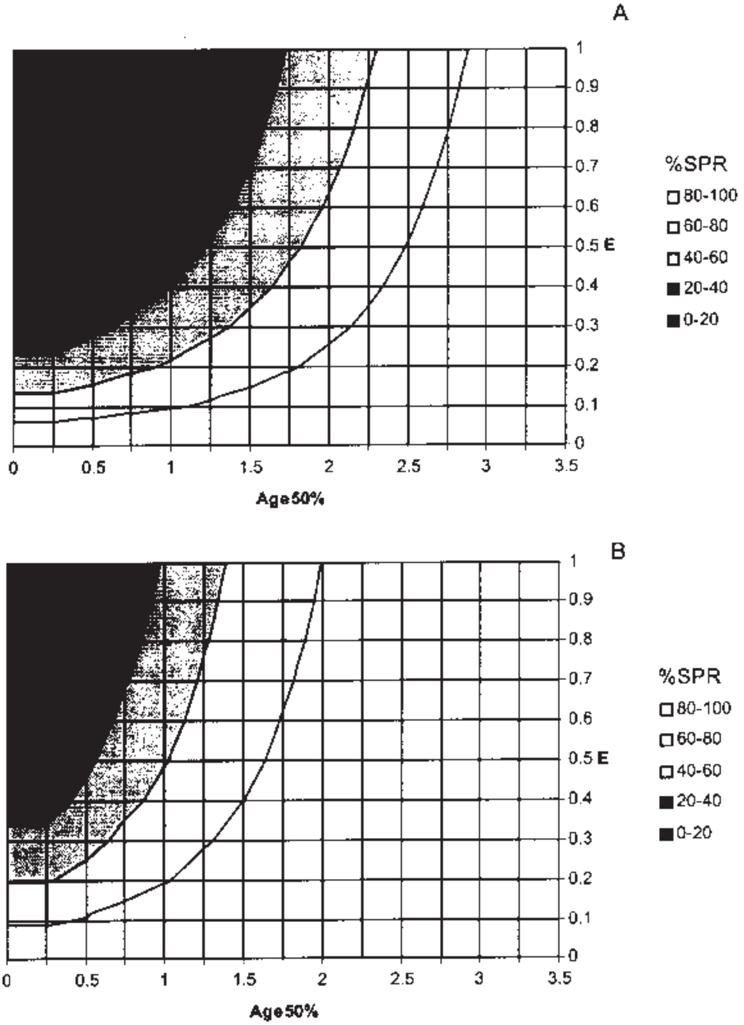


Figure 4. Percentage spawning per recruit (% SPR) for equilibrium exploitation rate (E) and age at first capture ($A_{50\%}$). (A) taking into account the age specific reproductive output (A); (B) if reproductive output were constant with age; (C) percentage contribution of sardine, Sardinella brasiliensis, first spawning age classes (ages 1 to 1.5) to the total reproductive potential of the stock. Data compiled from 1979 to 1989 (Cergole 1993).

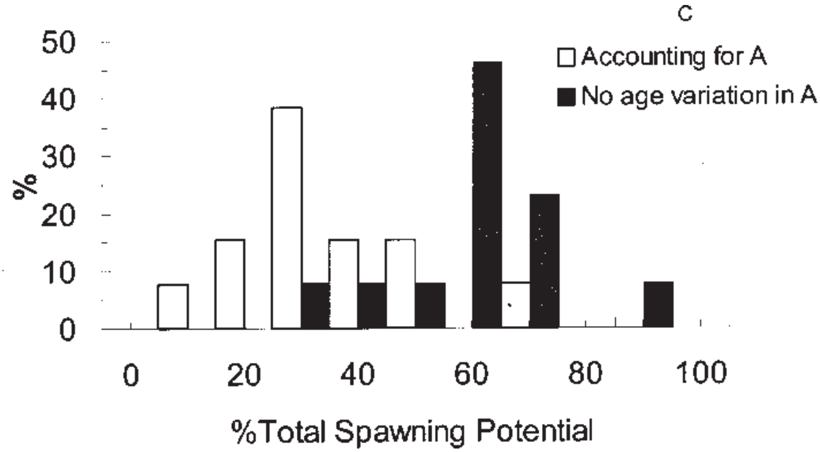


Figure 4. (Continued.)

regardless of the abundance of the stock (Pitcher and Parrish 1993). As a result, pelagic stocks concentrate their distribution range with declining biomass.

3. The mismatch between behavior adaptations and modern fishing techniques creates a pathological response of catch rates due to an increase in catchability with range collapse (Clark 1974, Winters and Wheeler 1985).

Each of the theories may have different consequences for fisheries harvest control decisions (Table 4). The simulation exercise carried out showed four general conclusions:

A. *Controlling harvest by constant effort is particularly inadequate in avoiding catchability lead stock collapse (CALSC) for populations with cyclic productivity.* The choice for harvest control by effort limitation involves tradeoffs between substantial decrease in fishing harvest and high monitoring and enforcement costs accompanying annual adjustments in fishing effort.

B. *Constant escapement policies offer the best tradeoff between mean catch, catch variability, and rebuilding capacity of the stock.* Management of small pelagic stocks should be targeted to favor recoveries when favorable environmental conditions prevail, rather than try to prevent the natural depletion accompanying an unfavorable regime. The objective of promoting stock rebuilding is intrinsic to management thresholds, defined by Quinn et al. (1990) as the population level below which the stock may be unable to rebuild its optimal level

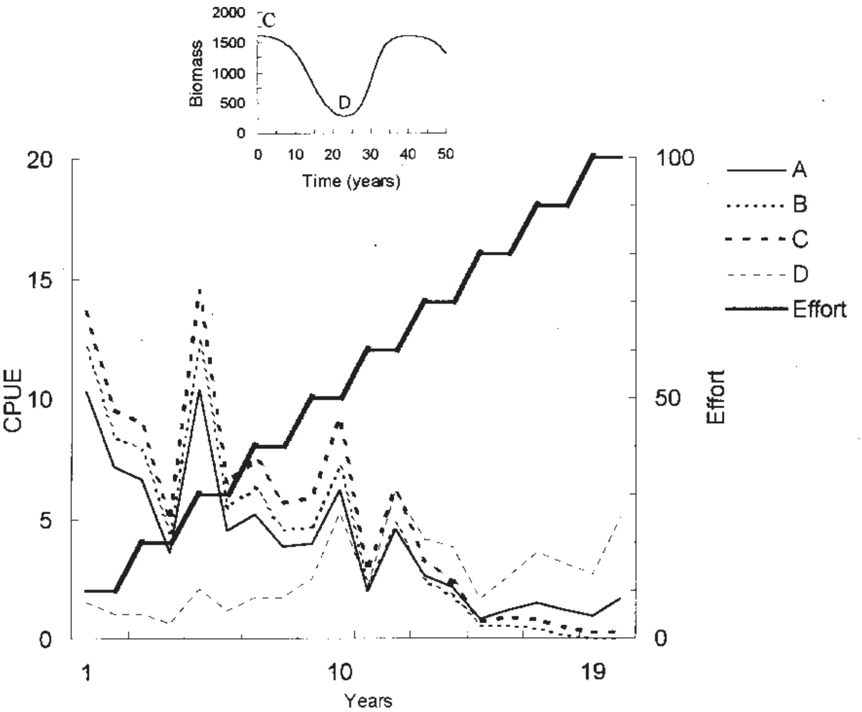


Figure 5. Fishing scenarios used in the estimation procedure for the Shaefer production parameters. Scenarios A and B show data acquired in a situation without cycles in productivity. Scenario B depicts a situation of catchability change with stock size (q_2 = 0.4). Scenarios C and D represent situations of long term cycles in productivity, being fishery data acquired respectively at the declining and recovery phase of the cycle (inset graph).

over an acceptable period of time. Escapement thresholds are, however, very often difficult to define, because of the poor information content of fisheries data, costly to implement since it requires independent biomass assessments, and further difficult to be accepted by fishery stakeholders since it may involve multi-year closures of the fishery.

C. *The estimation of population parameters and harvest control measures can be hampered by: (1) non-stationarity of parameters; (2) relatively short time series of data; (3) lack of independent abundance indices; and (4) lack of contrast in catch rates.* Because of concurrent changes in management actions and the environmental variables controlling productivity, it is often difficult to attribute causes of productivity changes to particular disturbances, as well as to correctly assign probability of population parameter values. Independent data on stock biomass

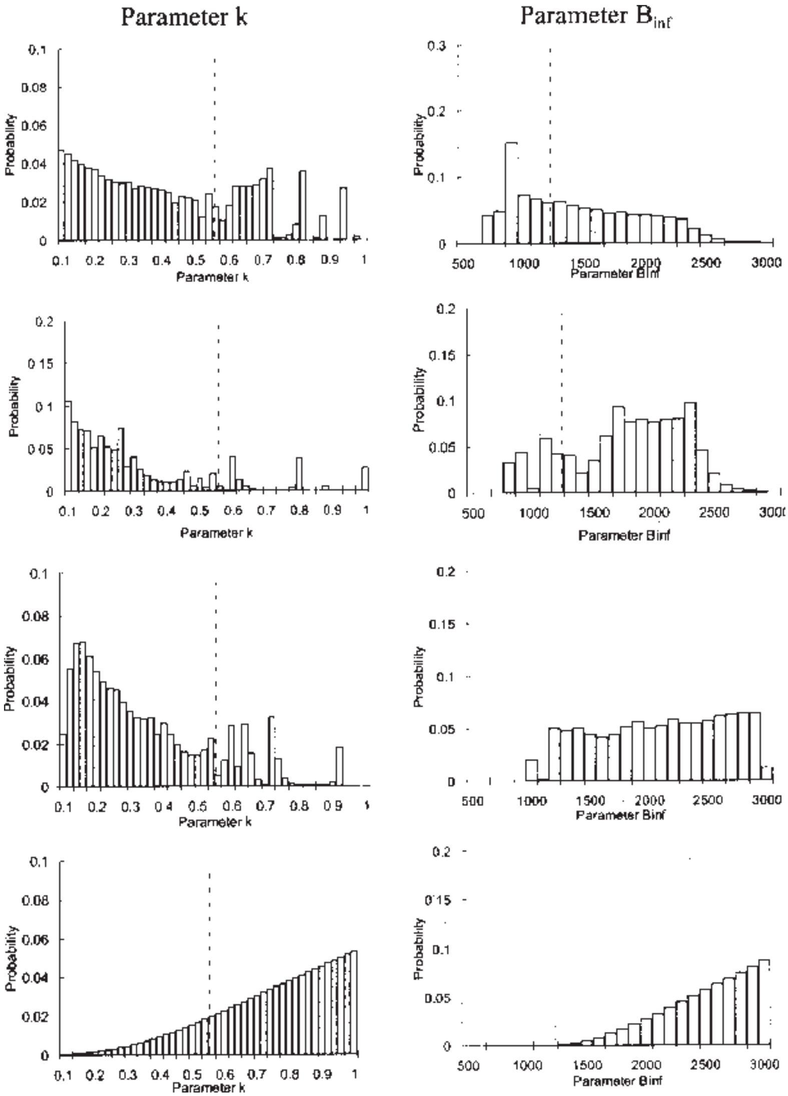


Figure 6. Marginal posterior probability distributions of parameters B_{inf} and k of a Schaefer production model fitted to data acquired in different simulation scenarios. The dashed line within each graph depicts the true values of the simulated parameters ($B_{inf} = 1,300$; $k = 0.55$). Data used for parameter estimation is shown in Fig. 5. Scenario A (top): No cycles in production. Scenario B (second from top): No cycles in production. Catchability change with stock size ($q_2 - 0.4$). Scenario C (second from bottom): cycles in production. Fishery data acquired during a "bust" phase in stock production cycle. Scenario D (bottom): Cycles in production. Fishery data acquired during a "boom" phase in stock production cycle.

Table 4. Analysis of the benefits, disadvantages and information needs for three different harvest control policies in cyclic regimes.

Harvest control	Benefits	Disadvantages	Information needs
$A_{50\%}$ or $L_{50\%}$	• Easy definition • Easy enforcement • Favor stock rebuilding • Protect spawning potential • Robust to catchability change	• May be difficult to implement (for non-selective gears/fisheries) • Decrease average yield for not preventing overfishing	• Size at first maturity • Gear selection • Size control at landing • Definition of nursery areas to be protected
Effort control	• May result in high average yield and low catch variance • Avoid the occurrence of “no-fishing” years.	• Requires in-season adjustments in order to prevent stock collapse • High monitoring + enforcement cost • Increase uncertainties for fishers and industry.	• Catchability estimates • *F* estimates • Fleet efficiency • Stock distribution • Schools attributes (density, inter-school distance, etc.)
Minimum escapement	• Favor stock rebuilding • May result in high average yield	• High uncertainty on harvest control targets (e.g. $B_{20\%,40\%}$) • High monitoring + enforcement cost • Increase probability of “no-fishing” years.	• Independent biomass estimates (e.g. surveys) • Monitoring of catches and landings • Catch at age/length

and density can provide the necessary information needed to efficiently control annual harvest. However, no prospect occurs that this information can reduce uncertainties on production relationships in systems markedly controlled by cyclic oscillations in the short term. For long-term management in non-stationary systems the optimum policies to reduce uncertainties may involve episodes of strong experimentation alternated with periods of more cautious, stabilizing management activity (Walters 1987). Or, alternatively, research should be focused on identifying and improving forecast of the natural processes involved in the generation of cyclic regimes in population dynamics, especially for the large small pelagic stocks inhabiting coastal upwelling systems (Bakun 1998). The biological, economical, and political impacts of such policies still need to be evaluated.

D. *Protecting the first spawning age classes can help prevent early collapses of the stock, and encourage more rapid responses to favorable environmental conditions.* Mace and Sissenwine (1993) analysis of replacement thresholds indicated that to persist, i.e., for successive generations replace each other on average, small pelagic populations must maintain an average 40 to 60% of their unfished spawning per recruit (%*SPR*). This relatively high %*SPR* led the authors to infer low resilience of these stocks to fishing mortality, since a small reduction in *SPR* would compromise the future replacement of the stock. Clearly, exploitation must be lower in order to maintain the stock for lower ages at first capture (see 40 to 60% bands, Fig. 4a,b). But protecting the first spawning age class (1 to 1.5 years) provides the stock with high resilience, in the sense that a broader range of fishing mortality rates could be sustained without substantially diminishing the capacity of the population to react to favorable oceanographic events. For stocks with a dominant controlling influence by environmental regimes, that seems to be the most appropriate strategy that could be supported on biological grounds (Winters et al. 1985). Yet, such strategy is probably untenable under most conditions for pelagic species due to the difficulty in controlling the age at entry in the fishery.

It is common ground in fisheries research that the problem of harvest strategies in stochastic environment can be adequately resolved with information on basic population parameters such as natural mortality, growth, and the steepness of stock-recruitment relationship (the analog of k in the Shaefer model) (Getz et al. 1987, Walters and Parma 1996, Patterson 1992). The vast literature and databases (e.g., ICLARM 1995, Myers et al. 1995) available today on fish life history and population parameters may offer at least prior information for the estimation of sustainable exploitation rates for most commercially important taxonomic groups. However, the prospect of reducing uncertainty in harvest control and hence in the probabilities of stock collapse for fisheries controlled by effort limitation is still restricted by the proper understanding of stock catchability (Arreguin-Sanchez 1996). Specifically for small pelagic schooling fish stocks, that will mean that more effort should be put on the analysis of stock spatial dynamics (spatial range) and its interaction with the fishing activity.

In this context, two concurrent lines of thought are developing: MacCall's (1990) approach describes the spatial dynamic of pelagic stocks according to a density dependent habitat selection model where a direct relationship between stock abundance and stock area is created by differences in habitat suitability (environmental conditions). His hypothesis has some corollaries useful for fisheries assessment, particularly that stock area could be a good indicator of the population state and thus used together with catch rates in the stock assessment. This approach has been successfully applied to the Pacific sardine, when the population was small and difficult to measure with conventional techniques (Barnes et al. 1992). A second approach (Pitcher 1997) suggests that shoaling behavior alone can cause range collapse in the absence of significant environmental gradients in space and time. If validated, this approach raises the prospect of obtaining cost-effective diagnostics of range collapse by monitoring behavioral parameters of shoaling fish. The inclusion of stock spatial attributes into conventional stock assessment research will require the gathering of auxiliary information not only from surveys but also from the fishery activity (e.g., time searching, school size, density of schools, etc.). We plan to explore the value of this auxiliary information in a future analysis.

Acknowledgments

We thank Reg Watson and two anonymous reviewers for suggestions and helpful comments. This study was conducted while the author held a sponsorship from the Conselho Nacional de Desenvolvimento Científico e Tecnológico-CNPq/Brazil.

References

Arreguin-Sanchez, F. 1996. Catchability: A key parameter for fish stock assessment. Rev. Fish Biol. Fish. 6:221-242.

Bakun, A. 1996. Patterns in the ocean: Ocean processes and marine population dynamics. California Sea Grant College System, La Jolla. 323 pp.

Bakun, A. 1998. The CEOS comparative analysis framework: Motivations and perceived opportunities. In: M.H. Durand, P. Cury, R. Mendelssohn, C. Roy, A. Bakun, and D. Pauly (eds.), Global versus local in upwelling systems. ORSTOM, Collection Colloques et séminaires, Paris, pp. 7-27.

Barnes, J.T., A.D. MacCall, L.D. Jacobson, and P. Wolf. 1992. Recent population trends and abundance estimates for the Pacific sardine (*Sardinops sagax*). Calif. Coop. Ocean Fish. Investig. Rep. 33:60-74.

Beddington, J., and R.M. May. 1977. Harvesting populations in a randomly fluctuating environment. Science 197:463-465.

Caddy, J.F., and J.A. Gulland.1983. Historical patterns of fish stocks. Marine Policy 10/83:267-278.

Cergole, M.C. 1993. Avaliação do estoque da sardinha-verdadeira, *Sardinella brasiliensis*, na costa sudeste do Brasil, período 1977-1990. Ph.D. thesis, Universidade de São Paulo. 245 pp.

Clark, C.W. 1974. Possible effects of schooling on the dynamics of exploited fish populations. J. Cons. Cons. Int. Explor. Mer 36(1):7-14.

Csirke, J. 1989. Changes in the catchability coefficient in the Peruvian anchoveta (*Engraulis ringens*) fishery. In: D. Pauly, P. Muck, and I. Tsukayama (eds.), The Peruvian upwelling ecosystem: Dynamics and interactions. ICLARM Conference Proceedings 18:207-219.

Csirke, J. 1995. Fluctuations in abundance of small and mid-size pelagics. Sci. Mar. 59(3-4):481-490.

Getz, W.M., R.C. Francis, and G.L. Swartzman. 1987. Managing variable fisheries. Can. J. Fish. Aquat. Sci. 44:1370-1375.

Goodyear, C.P. 1989. Spawning stock biomass per recruit: The biological basis for a 5 fisheries management tool. Int. Comm. Cons. Atl. Tunas working doc. SCRS/89/82. 10 pp.

Goodyear, C.P. 1993. Spawning stock biomass per recruit in fisheries management: Foundation and current use. In: S.J. Smith, J.J. Hunt, and D. Rivard (eds.), Risk evaluation and biological reference points for fisheries management. Can. Spec. Publ. Fish. Aquat. Sci. 120:67-81.

Hilborn, R., and C.J. Walters. 1992. Quantitative fisheries stock assessment and management: Choice, dynamics and uncertainty. Chapman and Hall, New York.

ICLARM. 1995. FishBase: A biological database on fish. Ver. 1.2. CD-ROM, ICLARM, Manila.

Lluch-Belda, D., R.J.M. Crawford, T. Kawasaki, A.D. MacCall, R.H. Parrish, R.A. Schwartzlose, and P.E. Smith. 1989. World-wide fluctuations of sardine and anchovy stocks: The regime problem. S. Afr. J. Mar. Sci. 8:195-205.

MacCall, A.D. 1976. Density dependence of catchability coefficient in the California Pacific sardine, *Sardinops sagax caerulea*, purse seine fishery. Calif. Coop. Ocean Fish. Investig. Rep. 18:136-148.

MacCall, A.D. 1990. Dynamic geography of marine fish populations. University of Washington Press, Seattle. 153 pp.

Mace, P.M., and M.P. Sissenwine. 1993. How much spawning per recruit is enough? In: S.J. Smith, J.J. Hunt, and D. Rivard (eds.), Risk evaluation and biological reference points for fisheries management. Can. Spec. Publ. Fish. Aquat. Sci. 120:101-118.

Mackinson, S., U.R. Sumaila, and T. Pitcher. 1997. Bioeconomics and catchability: Fish and fishers behaviour during stock collapse. Fish. Res. 31:11-17.

Myers, R.A., J. Bridson, and N.J. Barrowman. 1995. Summary of worldwide spawner and recruitment data. Can. Tech. Rep. Fish. Aquat. Sci. 2024. 54 pp.

Parrish, R.H., D. L. Mallicoate, and R.A. Klingbeil. 1986. Age dependent fecundity, number of spawning per year, sex ratio, and maturation stages in northern anchovy, *Engraulis mordax*. Fish. Bull. U.S. 84:503-517.

Patterson, K. 1992. Fisheries for small pelagic species: An empirical approach to management targets. Rev. Fish Biol. Fish. 2:321-338.

Pauly, D., and M. Soriano. 1987. Monthly spawning stock and egg production of Peruvian anchoveta (*Engraulis ringens*), 1953 to 1982. In: D. Pauly and I. Tsukayama (eds.), The Peruvian anchoveta and its upwelling ecosystem: Three decades of change. ICLARM Studies and Reviews 15:167-178.

Pitcher, T.J. 1995. The impact of pelagic fish behaviour on fisheries. Sci. Mar. 59:295-306.

Pitcher, T.J. 1997. Fish shoaling behaviour as a key factor in the resilience of fisheries: Shoaling behaviour alone can generate range collapse in fisheries. Proceedings of 2nd World Fisheries Congress, Brisbane, 1996.

Pitcher, T.J., and J. Parrish. 1993. The functions of shoaling behavior. In: T.J. Pitcher (ed.), The behaviour of teleost fishes. Chapman and Hall, London, pp. 363-349.

Polacheck, T., R. Hilborn, and A.E. Punt. 1993. Fitting surplus production models: Comparing methods and measuring uncertainties. Can. J. Fish. Aquat. Sci. 50:2597-2607.

Punt, A.E., and R. Hilborn. 1997. Fisheries stock assessment and decision analysis: The Bayesian approach. Rev. Fish Biol. Fish. 7: 35-63.

Quinn II, T.J., R. Fagen, and J. Zheng. 1990. Threshold management policies for exploited populations. Can. J. Fish. Aquat. Sci. 47:2016-2029.

Radovich, J. 1982. The collapse of the California sardine fishery. Calif. Coop. Ocean Fish. Investig. Rep. 23:56-78.

Steele, J.H. 1996. Regime shifts in fisheries management. Fish. Res. 25:19-23.

Walters, C.J. 1987. Nonstationarity of production relationships in exploited populations. Can. J. Fish. Aquat. Sci. 44(Suppl. 2):156-165.

Walters, C.J., and D. Ludwig. 1994. Calculation of Bayes posterior probability distribution for key population parameters. Can. J. Fish. Aquat. Sci. 51:713-722.

Walters, C.J., and A.M. Parma. 1996. Fixed exploitation rate strategies for coping with effects of climate changes. Can. J. Fish. Aquat. Sci. 53:148-158.

Winters, G.H., and J.P. Wheeler. 1985. Interaction between stock area, stock abundance, and catchability coefficient. Can. J. Fish. Aquat. Sci. 42:989-998.

Winters, G.H., E.L. Dalley, and J.A. Moores. 1985. Fortuity disguised as fisheries management: The case history of Fortune Bay herring. Can. J. Fish. Aquat. Sci. 42(Suppl. 1):263-274.

Implications of a Bayesian Approach for Simulating Salmon Population Dynamics

Ryan A. Hill
EVS Environment Consultants, North Vancouver, B.C., Canada

Brian J. Pyper
Simon Fraser University, School of Resource and Environmental Management, Burnaby, B.C., Canada

Abstract

Management options for Pacific salmon (*Oncorhynchus* spp.) populations are often analyzed using stochastic simulations that project spawner and recruit abundances over time. These forward simulations are commonly based on "best-fit" parameter estimates of the Ricker stock-recruitment relationship, despite the fact that such estimates are often highly uncertain. To examine if explicitly accounting for parameter uncertainty using Bayesian methods would affect the results of forward simulations, we compared simulations based on best-fit Ricker parameters to simulations where parameter uncertainty was incorporated through a Bayesian formulation with noninformative priors. Forward simulations using fixed harvest rates were conducted using data for two stocks of Fraser River sockeye *(O. nerka)* salmon. We found that expected values of spawner and recruit projections for the Bayesian simulations were consistently larger than projections for the best-fit case. Increases in the Bayes projections over those of the best-fit case typically ranged from 5 to 15% for the Stellako stock, and 20 to 60% for the Late Stuart stock. However, these differences were largely attributable to low values of the Ricker β parameter that were included in the Bayesian prior distribution. Results of Bayesian simulations were found to be highly sensitive to changes in the lower bound of the prior on β. We conclude that the common approach of specifying noninformative priors by extending uniform distributions across a broad range of parameter values can be inappropriate for the Ricker model, and recommend the use of informative priors whenever possible.

Introduction

Management of Pacific salmon (*Oncorhynchus* spp.) populations has relied heavily on stock-recruitment models relating abundances of spawners (escapements) to subsequent adult returns (recruits). Such models have been used as a basis for setting harvest policies for many salmon stocks (e.g., Minard and Meachem 1987) and for simulating salmon population dynamics over time (e.g., Welch and Noakes 1991, Emlen 1995). For example, Welch and Noakes (1991) used a form of the Ricker stock-recruitment model to simulate future rebuilding trajectories of the Adams River sockeye (*O. nerka*) salmon stock under alternative fixed harvest-rate policies. Similarly, Emlen (1995) used simulations based on the Ricker model and fixed harvest rates to estimate the probability that the Snake River chinook (*O. tshawytscha*) salmon stock would fall below some specified abundance in the future. However, such "forward" simulations of salmon population dynamics have been based almost exclusively on "best-fit" stock-recruitment relationships, even though these relationships are typically plagued with uncertainty.

Recently, Bayesian methods have been advocated for explicitly incorporating uncertainties inherent in fisheries models into stock assessments and decision making (reviewed by Punt and Hilborn 1997). Bayesian methods allow for probabilities to be placed on alternative hypotheses, such as different parameter combinations for stock-recruitment models (e.g., Ianelli and Heifetz 1995, Adkison and Peterman 1996). Accounting for uncertainties in fisheries models can have important implications for management (Thompson 1992, Frederick and Peterman 1995); however, there have been limited applications of Bayesian methods to analyses of stock-recruitment data for salmon populations (e.g., Geiger and Koenings 1991, Adkison and Peterman 1996, Robb and Peterman 1998). In particular, we are unaware of any papers that examine how incorporating uncertainty in salmon stock-recruitment relationships may affect results of forward simulations where fixed harvest rates are used. It is unclear to what extent projections of spawner and recruit abundances, like those of Welch and Noakes (1991) and Emlen (1995), may be sensitive to such uncertainty. Intuitively, we would expect results to differ somewhat when various, yet probable, shapes of stock-recruitment curve are considered in the simulations rather than just the best-fit relationship.

In this paper, we compare results of forward simulations based on the best-fit parameters of the Ricker model to results of simulations where uncertainty in the Ricker parameters is explicitly accounted for using Bayesian methods. In the Bayesian analyses, we use "noninformative" priors for the Ricker parameters, which is a common approach used in Bayesian formulations of fisheries models (e.g., Thompson 1992, Walters and Ludwig 1994, Ianelli and Heifetz 1995, Robb and Peterman 1998). Simulations are conducted using fixed harvest rates and data for two stocks of Fraser River sockeye salmon.

Methods

Data

We used abundances for age-4_2 spawners and recruits for two stocks of Fraser River sockeye salmon (the Stellako and Late Stuart stocks) over brood years 1948-1990 (Pers. comm., Allan Cass, Department of Fisheries and Oceans, Pacific Biological Station, Nanaimo, B.C.). We limited our analyses to age-4 fish, which account for the majority of recruits for both the Stellako and Late Stuart stocks (91% and 97%, respectively), to simplify the forward simulations described below. As we later discuss, our general findings are not affected by more complicated age structures.

Ricker Stock-Recruitment Model and Bayesian Approach

When fitting stock-recruitment relationships such as the Ricker model (Ricker 1975), two key statistical problems arise from (1) measurement error in spawner abundances (Walters and Ludwig 1981), and (2) the lack of independence between recruitment and future spawner abundances (i.e., time-series bias; Walters 1985). In our analyses, we assume that spawner and recruit abundances are measured accurately and we do not account for possible time-series bias in parameter estimates; however, as we later discuss, our general conclusions are not affected by either problem.

Given these caveats, the statistical properties of the Ricker stock-recruitment parameters can best be understood when the relationship is expressed as

$$R = S\exp(\alpha - \beta S + \nu), \tag{1}$$

where R and S denote recruits and spawners, respectively, and (α, β) are parameters. The stochastic error term ν is assumed to be normally distributed with standard deviation σ, reflecting the multiplicative, lognormal error structure common in salmon recruitment data (Peterman 1981). We refer to σ as the “process error.” Given this error structure, the Ricker relationship (equation 1) can be conveniently expressed as a linear, normal error model:

$$\log(R/S) = \alpha - \beta S + \nu, \tag{2}$$

where α is the intercept, representing productivity ($\log[R/S]$) at low spawner abundance, and β is the slope (with units $1/S$), representing density-dependent effects on productivity. Note that the quantity $1/\beta$ (with units S) corresponds to the spawner abundance at which recruitment is maximized (i.e., the peak of the stock-recruitment curve).

The parameter vector (α, β, σ), denoted as θ, can be estimated by linear regression or maximum likelihood (both methods yield the same estimates for [α,β]). In this paper, we refer to the linear regression estimates $\hat{\theta}$ as the "best-fit" parameters. With (α,β) defined as the intercept and slope, the statistical properties of their estimates $(\hat{\alpha},\hat{\beta})$ are easy to interpret (Neter et al. 1985). In some cases, the residual errors from the fit of the Ricker curve will contain temporal autocorrelation, which may be important to model in forward simulations (e.g., Welch and Noakes 1991, Emlen 1995). However, there was little evidence of autocorrelation in the error residuals for either the Stellako or Late Stuart sockeye stocks (e.g., lag-1 autocorrelation estimates were 0.12 and −0.05, respectively).

Recruitment data for salmon populations can be highly variable and, as a result, estimates of stock-recruitment parameters are often very imprecise. In the Bayesian approach, this uncertainty can be incorporated into the stock-recruitment analysis by computing probabilities for different combinations of (α, β, σ), each of which defines a curve of different shape. For a given parameter combination θ_i, the posterior probability of θ_i given the observed stock-recruitment data D is determined using Bayes' theorem:

$$P(\theta_i \mid D) = \frac{L(D \mid \theta_i)\mathrm{P}(\theta_i)}{\sum_j L(D \mid \theta_j)\mathrm{P}(\theta_j)}\,, \tag{3}$$

where $L(D|\theta_i)$ is the likelihood of the data given θ_i, $P(\theta_i)$ is the prior probability assigned to θ_i, and $\sum_j L(D|\theta_j)P(\theta_j)$ is the sum of the likelihood-prior products for all parameter combinations considered.

In our analyses, we used a "noninformative" prior for θ. Such a prior is intended to convey little information about θ so that the posterior distribution for θ is primarily determined by the data via the likelihood function. A noninformative prior recommended for the normal regression model equation (2) consists of uniform priors for both (α, β) and a prior probability on σ of $1/\sigma$ (Gelman et al. 1995, p. 236). To be consistent for both the Stellako and Late Stuart stocks, we used simple rules—based on the best-fit parameters $(\hat{\alpha},\hat{\beta})$ and their standard errors (SE)—for initially setting bounds on priors that would encompass most of the "biologically reasonable" parameter combinations that were likely given the data.

Specifically, upper and lower bounds on α were set to $\hat{\alpha} + 3\,\mathrm{SE}(\hat{\alpha})$ and $\hat{\alpha} - 3\,\mathrm{SE}(\hat{\alpha})$, respectively. Likewise, the upper bound on β was set to $\hat{\beta} + 3\,\mathrm{SE}(\hat{\beta})$. However, this approach was not used for the lower bound on β because negative values of β would have been considered for both sockeye stocks. Negative values of β imply density-independence or exponential population growth, and hence are not biologically reasonable. We therefore initially set the lower bound for β equal to $\hat{\beta}/10$, which implies an upper limit for the spawner abundance that maximizes recruitment (i.e., $1/\beta$) equal to 10 times that for the best-fit curve (i.e., for $\hat{\beta}$). For both (α, β), we divided the prior range into 30 values, with each value assigned equal probability (i.e., a uniform prior). For σ, we found that upper and lower bounds of $3\hat{\sigma}$ and $\hat{\sigma}/3$, respectively, were sufficient to cover the range of likely values. This prior range for σ was divided into 20 values. Thus, we considered a total of $30 \times 30 \times 20$ or 18,000 parameter combinations (θ_i) for each stock.

For a given θ_i, we computed the likelihood of the stock-recruitment data, across brood years t, using the normal likelihood function:

$$L(D \mid \theta_i) = \prod_t \frac{1}{\sigma_i \sqrt{2\pi}} \exp\left(-\frac{w_t^2}{2\sigma_i^2}\right), \tag{4}$$

where $w_t = \log(R_t/S_t) - (\alpha_i - \beta_i S_t)$. The prior probabilities and likelihoods of the various parameter combinations (θ_i) were then combined using Bayes' theorem (equation 3) to generate the posterior distribution for θ.

Forward Simulations

To compare results of stochastic simulations of stock-recruitment dynamics using only the best-fit Ricker parameters to results of simulations where parameter uncertainty was incorporated via Bayesian methods, we conducted the following baseline analyses for both the Stellako and Late Stuart sockeye stocks. For each stock, we used the Ricker model (equation 1) to simulate 5 generations (i.e., 20 years into the future) of spawner and recruit data using an initial escapement of 5,000 spawners and a fixed harvest rate of 0.7. For simulations using the best-fit parameters $\hat{\theta}$, we ran 500 Monte Carlo trials and computed the following performance measures using only data tabulated for the final (i.e., fifth) generation of each simulation:

A. mean of the 500 escapements, $\bar{S}$;

B. standard deviation of the 500 escapements, SD(S);

C. the proportion of the 500 escapements below (i) 5,000 spawners, (ii) 25,000 spawners, and (iii) 50,000 spawners; and

D. mean and standard deviation of the 500 catch values.

For the Bayesian case, we repeated the above procedure for each parameter combination (θ_i) considered in the Bayesian analysis. To allow for an accurate comparison between the best-fit and Bayes cases, the same sequence of random numbers that was used for the best-fit case was also used for each θ_i. The Bayes expected value for a given performance measure (PM) was then calculated across all parameter combinations by summing the values of that performance measure for each θ_i, weighted by its posterior probability, $P(\theta_i|D)$:

$$E[PM] = \sum_i PM_i P(\theta_i \mid D). \tag{5}$$

Table 1. Best-fit parameters of the Ricker stock-recruitment relationship for the Stellako and Late Stuart sockeye stocks.

Stock	$\hat{\alpha}$	SE ($\hat{\alpha}$)	$\hat{\beta}$ (10^{-6})	SE ($\hat{\beta}$) (10^{-6})	*P* value (H_o: $\beta = 0$)	$\hat{\sigma}$
Stellako	2.00	0.14	3.41	1.25	0.01	0.58
Late Stuart	2.11	0.23	1.56	1.34	0.25	1.29

P values are shown for a two-tailed *t* test. SE = standard error.

In addition, to assess the potential influence of the prior probability distribution on the results, we also computed expected values for the various performance measures using only the prior probability of each θ_i as the weight in the summation (i.e., the stock-recruitment data was ignored in this case).

We examined the sensitivity of the simulation results to the choice of harvest rate, the number of generations simulated, the initial spawner abundance, the number of Monte Carlo trials, and the sequence of random numbers. In addition, for the Bayes case, we examined the sensitivity of the results to the bounds placed on the prior distribution of each Ricker parameter (α, β, σ), as well as the number of discrete values considered for each parameter (i.e., grid fineness).

Results

The best-fit Ricker parameters for the Stellako and Late Stuart sockeye stocks are shown in Table 1. Estimates of α were fairly precise for both stocks (Table 1). However, estimates of β were less precise, particularly for Late Stuart sockeye, which also had a much larger estimate of process error σ. As indicated by the *P* values for the null hypothesis test of $\beta = 0$, there is much weaker evidence of density-dependence in the stock-recruitment data for Late Stuart sockeye than for Stellako sockeye (Table 1).

This weaker evidence of density-dependence for Late Stuart sockeye is also reflected in the (marginal) posterior distributions of (α, β) for the two stocks (Fig. 1). Recall that for both stocks, we placed a lower bound on β equal to $\hat{\beta}/10$. For Stellako sockeye, values of β near this lower bound have quite small posterior probabilities for all α values (Fig. 1A). In contrast, for Late Stuart sockeye, there are relatively large probabilities for many (α, β) combinations near the lower bound for β (Fig. 1B), where these values of β imply little density-dependence in the stock-recruitment curve. Aside from combinations of (α, β, σ) not considered for Late Stuart sockeye due to the constraint placed on low β values, the posterior distributions of each stock covered the majority of probable parameter combinations.

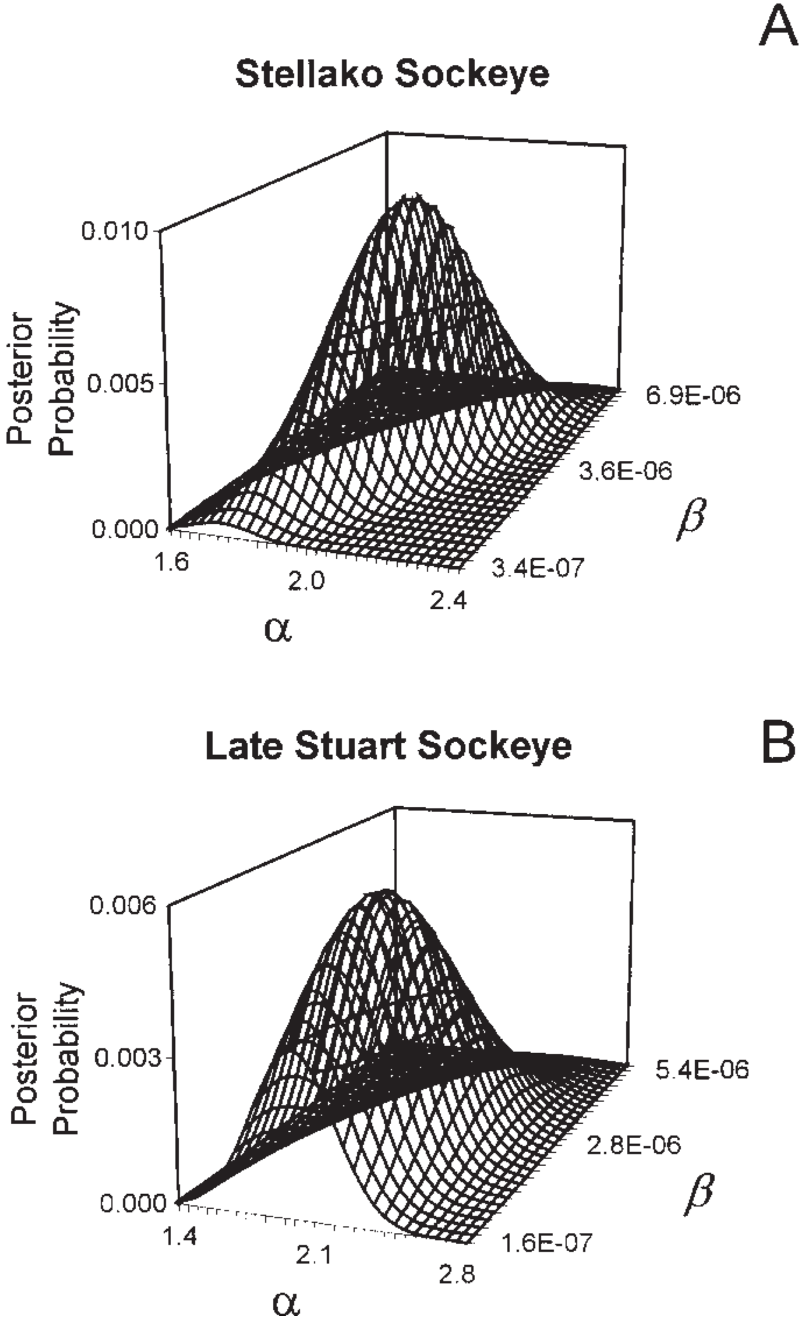


Figure 1. Marginal posterior distributions of (α, β), integrated over σ, for the (A) Stellako and (B) Late Stuart sockeye salmon stocks.

The effects of including parameter uncertainty in the forward simulations differed for the two sockeye stocks (Tables 2 and 3). For Stellako sockeye, the results of the baseline simulations, when integrated over the Bayes posterior distribution for (α, β, σ), were very similar to those using only the best-fit parameters (Table 2). For example, the expected value of mean escapement ($\bar{S}$) for the Bayes case was nearly identical to the best-fit $\bar{S}$ (Table 2). In addition, the standard deviation in escapement (SD[S]), as well as the proportions of escapements below 5,000, 25,000, and 50,000 spawners, were only slightly larger for the Bayes case (Table 2). In contrast, for Late Stuart sockeye, the Bayes $\bar{S}$ and SD(S) were larger than those for the best-fit case by 27.8% and 32.4%, respectively. Again, however, the proportion of escapements below a specified level was only slightly larger for the Bayes case (Table 3).

Note that we have not shown performance measures for catch in Tables 2 and 3 because the relative differences in these measures between the best-fit and Bayes cases were identical to those for escapement. This occurred because a fixed harvest rate (HR) was used in the simulations, so the ratio between catch and escapement was always constant (catch = escapement × $HR/[1\text{-}HR]$).

Results for simulations based on the Bayes prior distribution were considerably different from the best-fit results for both sockeye stocks (Tables 2 and 3). Values of $\bar{S}$ and SD(S) were roughly 30-40% larger for the Bayes priors. The proportion of escapements below a specified level also increased markedly for both stocks when the Bayes priors were used (Tables 2 and 3).

Sensitivity Analyses

We found that differences in performance measures between the best-fit and Bayes (posterior) cases were indeed sensitive to the harvest rate, the initial spawner abundance, and in particular, the number of generations used in the forward simulations. In all cases, however, values of $\bar{S}$ and SD(S) were larger for the Bayes case. For example, Fig. 2 shows percent differences in $\bar{S}$ between the Bayes and best-fit cases for various combinations of harvest rate and generations. For both stocks, the Bayes $\bar{S}$ became increasingly greater than the corresponding best-fit $\bar{S}$ as the number of simulated generations was increased, although the differences were much larger for Late Stuart sockeye (Fig. 2B). There were similar trends for differences in SD(S). However, in contrast to the results for $\bar{S}$ and SD(S), the proportion of escapements below a specified level varied little between the Bayes and best-fit cases over the range of simulation conditions we examined.

Changing the random number sequence or the number of Monte Carlo trails had little influence on the simulation results. Although the absolute values of a given performance measure could change somewhat, the

Table 2. Escapement-based performance measures computed for the baseline forward simulations of the Stellako sockeye stock using (i) the best-fit Ricker parameters, (ii) the Bayes posterior distribution for the Ricker parameters, and (iii) the Bayes prior distribution.

Performance measure	Simulations using best-fit parameters	Bayesian simulations (% difference from best-fit case)	
		Posterior	Prior only
Mean escapement ($\bar{S}$)	168.5	168.6 (+0.0%)	227.7 (+35.1%)
SD(S)	148.1	153.0 (+3.3%)	213.8 (+44.4%)
Proportion < 5,000	0.004	0.005	0.015
Proportion < 25,000	0.056	0.073	0.132
Proportion < 50,000	0.156	0.182	0.261

Expected values are shown for the Bayes cases. Values of mean escapement ($\bar{S}$) and standard deviation of escapement (SD[S]) are in thousands.

Table 3. Escapement-based performance measures computed for the baseline forward simulations of the Late Stuart sockeye stock using (i) the best-fit Ricker parameters, (ii) the Bayes posterior distribution for the Ricker parameters, and (iii) the Bayes prior distribution.

Performance measure	Simulations using best-fit parameters	Bayesian simulations (% difference from best-fit case)	
		Posterior	Prior only
Mean escapement ($\bar{S}$)	518	661 (+27.8%)	665 (+28.5%)
SD(S)	1394	1845 (+32.4%)	1801 (+29.2%)
Proportion < 5,000	0.074	0.087	0.145
Proportion < 25,000	0.194	0.217	0.304
Proportion < 50,000	0.288	0.301	0.408

Expected values are shown for the Bayes cases. Values of mean escapement ($\bar{S}$) and standard deviation of escapement (SD[S]) are in thousands.

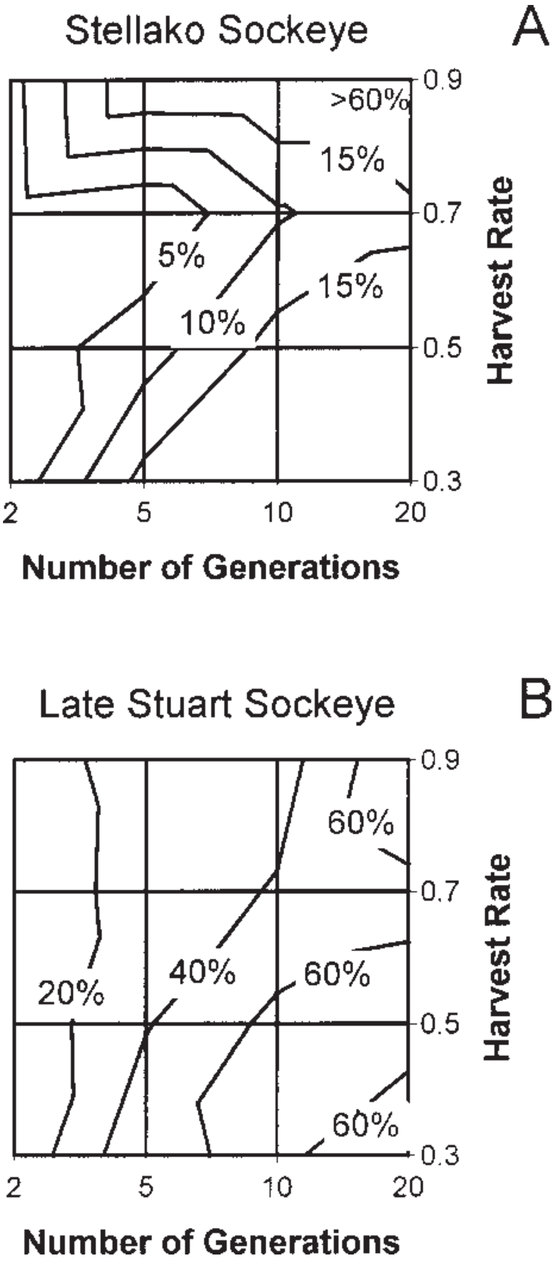


Figure 2. Percent differences between Bayes and best-fit estimates of mean escapement ($\bar{S}$) in the final generation for various combinations of harvest rate and number of generations used in the simulations.

relative difference between the Bayes and best-fit cases remained quite constant.

To examine if the simulation results for the Bayes cases were sensitive to the limited number of discrete parameter combinations we considered, we doubled the number of grid points for each parameter (α, β, σ), giving a total of $60 \times 60 \times 40$ combinations. This grid was used in the baseline simulation described above; however, for both stocks, results were nearly identical to those for the original grid.

Finally, we examined the sensitivity of the simulation results for the Bayes cases to changes in the bounds placed on the priors for (α, β, σ). For both stocks, results changed little for moderate increases or decreases in the bounds used for α (e.g., using bounds of $\hat{\alpha} \pm 4\,\mathrm{SE}[\hat{\alpha}]$ or $\hat{\alpha} \pm 2\,\mathrm{SE}[\hat{\alpha}]$). Results were especially insensitive to changes in the bounds for σ. Likewise, moderate changes in the upper bound for β (values that convey high density-dependence) had very little affect on the simulation results.

However, values of Bayes $\overline{S}$ and SD(S) were often very sensitive to changes in the lower bound for β (values that convey little density-dependence), depending on the stock and simulation conditions. Although results for Stellako sockeye varied little when the baseline simulations were used, results changed appreciably when simulations were conducted using 10 generations and a harvest rate of 0.5, for instance. In the latter case, extending the lower bound on β (i.e., allowing for values of β to approach zero) resulted in large increases in both the Bayes $\overline{S}$ and SD(S) (as shown for $\overline{S}$ in Fig. 3A). Values of Bayes $\overline{S}$ and SD(S) for Late Stuart sockeye were considerably more sensitive to the lower bound on β across all simulation conditions, as shown for $\overline{S}$ under the baseline conditions (Fig. 3B).

Discussion

Our results suggest that incorporating uncertainty in parameter estimates of the Ricker model can have profound effects on the results of forward simulations. Specifically, using a Bayesian formulation with noninformative priors over biologically reasonable bounds, we found that projections of future spawner abundances (and likewise for catch and recruitment) typically increased by 5-15% for Stellako sockeye and 20-60% for Late Stuart sockeye (Fig. 2). At first glance, these differences might be interpreted as strong evidence of the need to account for parameter uncertainty; however, the results of the Bayesian analyses may be misleading as they were highly dependent on the specification of the prior.

In particular, the differences in results of simulations between the Bayes and best-fit cases, or between the Stellako and Late Stuart stocks, were largely due to the low values of β considered in the Bayes analyses. For parameter combinations with low β values, which convey very little density-dependence, projections of spawner and recruit abundances typically increased over time to be 10 or more times greater than those of

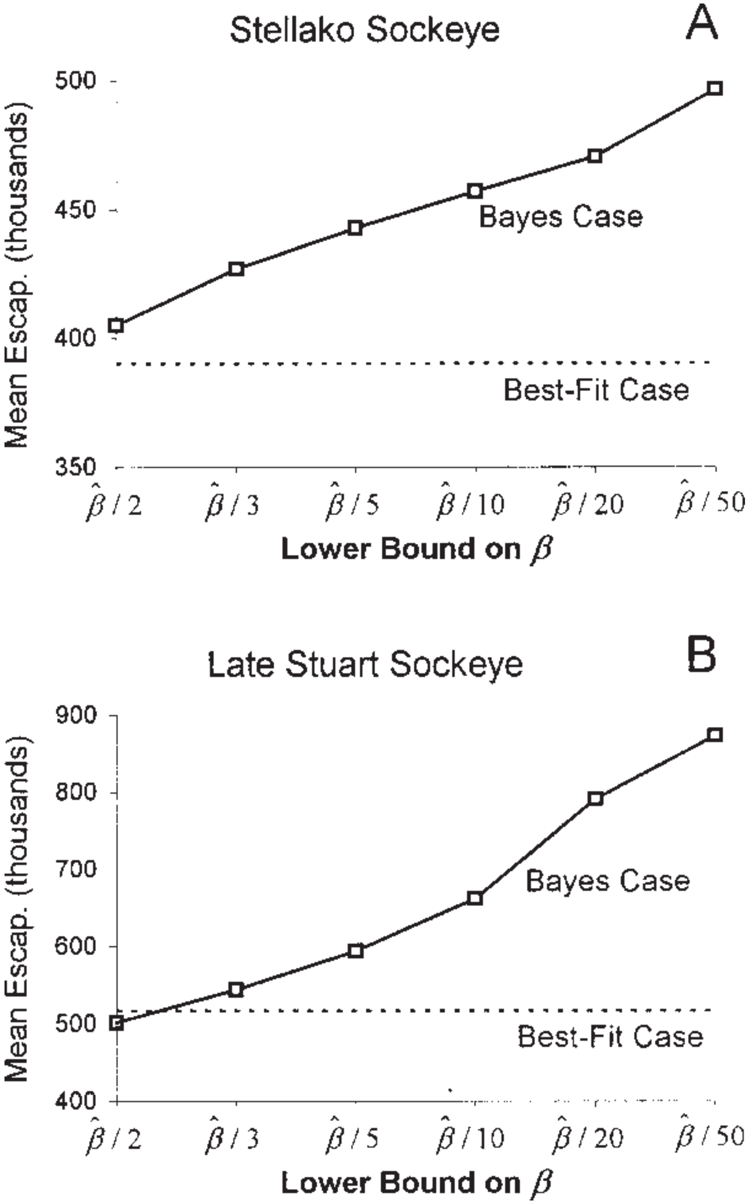


Figure 3. Effect of changes in lower bound of the prior distribution on beta on Bayes expected mean escapement ($\overline{S}$) for (A) the Stellako stock and (B) the Late Stuart stock. Simulations were based on 10 generations and a harvest rate of 0.5 for the Stellako stock, and 5 generations and a harvest rate of 0.7 for the Late Stuart stock.

the best-fit case. Thus, for Late Stuart sockeye, the Bayes $\overline{S}$ was considerably greater than the best-fit $\overline{S}$, especially when a large number of generations was simulated (Fig. 2B), because this stock had relatively large posterior probabilities for low β values (Fig. 1B). Indeed, for either stock, results for the Bayes case were highly sensitive to the lower bound placed on the prior for β (Fig. 3).

Moreover, although the prior distribution was designed to convey little information about the Ricker parameters themselves, the prior was quite informative in terms of the spawner-recruit projections. Recall that for both stocks, simulations based on only the prior distribution gave much larger values of $\overline{S}$ and SD(S) than the best-fit values (Table 2 and 3). Again, these differences increased considerably when the lower bound on β was extended. Note that extending this lower bound from $\hat{\beta}/10$ to $\hat{\beta}/50$ represents a relatively small change to the parameter grid, but in biological terms, it represents a 5-fold increase in the upper limit for the spawner abundance that maximizes recruitment (now 50 times that for the best-fit case).

Based on the statistical properties of the Ricker model and the simulation results for the prior distribution, we can generalize our findings to other stock-recruitment data sets. With a uniform prior on (α, β), the joint posterior will be a bivariate normal distribution (Gelman et al. 1995, p. 236). The Bayes expected value of recruits per spawner, when integrated over this joint posterior or over the marginal posteriors for either α or β, will always be greater than the best-fit value when liberal bounds on the prior are used. Moreover, this difference will increase for less precise estimates of α or β. Thus, we expect that spawner and recruit projections of the Bayes case will tend to be larger than the best-fit case for any stock-recruitment data set, with results that are often highly inconsistent with past data.

Given this expectation, the use of a broad-scale uniform prior for (α, β) appears to be inappropriate and potentially misleading. In the Bayesian approach, the prior distribution is meant to reflect the true "degree of belief" in alternative parameter combinations in the absence of stock-recruitment data. Punt and Hilborn (1997) recommend that noninformative priors be avoided on the basis that such priors likely ignore relevant biological information. It seems highly unlikely, for example, that a biologist would consider prior β values of $\hat{\beta}$ and $\hat{\beta}/10$ as equally probable given information on habitat constraints for spawning or juvenile rearing. Instead, a much more defensible and realistic approach would be to use such auxiliary information to place "informative" priors on alternative parameter combinations (Punt and Hilborn 1997). We would expect that using informative priors would greatly restrict the influence of low β values, in which case simulation results for the Bayes and best-fit cases would likely be quite similar. Note, however, that analysts should interpret results cautiously when informative priors contradict the stock-recruitment data (Adkison and Peterman 1996).

Several factors were not considered in our analyses. First, we used only a single age class of recruits. However, our general findings would not be affected by using more complicated age structures because the Ricker model relates spawner abundances to total recruits, regardless of their age. Second, we not did consider potential sources of parameter bias (Walters and Ludwig 1981, Walters 1985). Nevertheless, assuming that corrections for bias made to best-fit parameters could also be incorporated into the Bayesian framework, we would still expect simulations for Bayes and best-fit cases to show relative differences similar to those discussed above. Third, we did not consider other parameterizations of the Ricker model for which noninformative priors have been used (Adkison and Peterman 1996) or recommended (Schnute and Kronlund 1996). However, preliminary results (data not shown) suggest that these alternative priors lead to even larger discrepancies between the Bayes and best-fit cases. Finally, we did not consider uncertainty in the functional form of the stock-recruitment model. For example, there are three-parameter models for which the Ricker curve is but one limiting form (e.g., Schnute and Kronlund 1996). Again, however, we expect qualitatively similar results to those shown here for Bayesian analyses of these models that are based on noninformative priors which extend over parameter combinations that convey little density-dependence.

In conclusion, we have shown that incorporating uncertainty in parameter estimates of the Ricker model using Bayesian methods with noninformative priors can result in large increases in projections of spawner and recruit abundances. Such increases, if taken at face value, could obviously have important management implications. However, we believe these findings act more as a warning about the potential pitfalls of using noninformative priors in Bayesian analyses than as evidence of the importance of including parameter uncertainty in forward simulations of salmon dynamics.

Acknowledgments

We thank Andre Punt and an anonymous reviewer for their useful comments on the manuscript.

References

Adkison, M.D., and R.M. Peterman. 1996. Results of Bayesian methods depend on details of implementation: An example of estimating escapement goals. Fish. Res. 25:155-170.

Emlen, J.M. 1995. Population viability of the Snake River chinook salmon *(Oncorhynchus tshawytscha).* Can. J. Fish. Aquat. Sci. 52:1442-1448.

Frederick, S.W., and R.M. Peterman. 1995. Choosing fisheries harvest policies: When does uncertainty matter? Can. J. Fish. Aquat. Sci. 52:291-306.

Geiger, H.J., and J.P. Koenings. 1991. Escapement goals for sockeye salmon with informative prior probabilities based on habitat considerations. Fish. Res. 11:239-256.

Gelman, A., Carlin, J.B., Stern, H.S., and D.B. Rubin. 1995. Bayesian data analysis. Chapman and Hall, London. 526 pp.

Ianelli, J.N., and J. Heifetz. 1995. Decision analysis of alternative harvest policies for the Gulf of Alaska Pacific Ocean Perch Fishery. Fish. Res. 24:35-63.

Minard, R.E., and C.P. Meachem. 1987. Sockeye salmon *(Oncorhynchus nerka)* management in Bristol Bay, Alaska. In: H.D. Smith, L. Margolis, and C.C. Wood (eds.), Sockeye salmon population biology and future management. Can. Spec. Publ. Fish. Aquat. Sci. 96:336-342.

Neter, J., W. Wasserman, and M.H. Kutner. 1985. Applied linear statistical models. Richard D. Irwin. Homewood, IL.

Peterman, R.M. 1981. Form of random variation in salmon smolt-to-adult relations and its influence on production estimates. Can. J. Fish. Aquat. Sci. 38:1113-1119.

Punt, A.E., and R. Hilborn. 1997. Fisheries stock assessment and decision analysis: The Bayesian approach. Rev. Fish Biol. Fish. 7:35-63.

Ricker, W.E. 1975. Computation and interpretation of biological statistics of fish populations. Bull. Fish. Res. Board Can. 191.

Robb, C., and R.M. Peterman. 1998. Application of Bayesian decision analysis to management of a sockeye salmon fishery. Can. J. Fish. Aquat. Sci. 55:86-98.

Schnute, J.T., and A.R. Kronlund. 1996. A management oriented approach to stock recruitment analysis. Can. J. Fish. Aquat. Sci. 53:1281-1293.

Thompson, G.G. 1992. A Bayesian approach to management advice when stock-recruitment parameters are uncertain. Fish. Bull. U.S. 90:561-573.

Walters, C.J. 1985. Bias in the estimation of functional relationships from time series data. Can. J. Fish. Aquat. Sci. 42:147-149.

Walters, C.J., and D. Ludwig. 1981. Effects of measurement errors on the assessment of stock-recruit relationships from time series data. Can. J. Fish. Aquat. Sci. 38:704-710.

Walters, C.J., and D. Ludwig. 1994. Calculation of Bayes posterior probability distributions for key population parameters. Can. J. Fish. Aquat. Sci. 51:713-722.

Welch, D.W., and D.J. Noakes 1991. Optimal harvest rate policies for rebuilding the Adams River sockeye salmon *(Oncorhynchus nerka).* Can. J. Fish. Aquat. Sci. 48:526-535.

Alaska Sea Grant College Program • AK-SG-98-01, 1998

In Search of a New Approach to Fisheries Management in the Middle Amazon Region

Victoria J. Isaac
Museu Paraense Emílio Goeldi, Belém, Brazil

Mauro L. Ruffino
IBAMA/GOPA-GTZ, Santarém, Brazil

David McGrath
Núcleo de Altos Estudos Amazônicos-UFPA, Belém, Brazil

Abstract

Due to the complexity of Amazonian fisheries, the application of conventional models of stock assessment and, especially, of fisheries management do not work for Amazonian fisheries. Available information shows that regional artisanal fisheries, in addition to multispecific, are highly varied in their technology, geographically diffuse, and practiced with varying economic objectives. Two fisheries can be distinguished in the Middle Amazon. The first is more professionalized and highly seasonal in character, concentrates primarily on migratory catfish and operates primarily in the main river channels. The second fishery is more diffuse, focusing on smaller species in floodplain lakes, adapting gear, location and target species to the ecological characteristics of the stock. Classical assessment models indicate some stocks my be overfished, such as *Colossoma macropomum*. Nevertheless, implementation of efficient regulatory measures is almost impossible. Although there are measures regulating fisheries, these regulations are not obeyed. As a result of the inability of the government to regulate the fishery, conflicts and litigation have proliferated over the right of access to local floodplain lake fisheries. In response to these actions, government institutions are changing their approach and their fisheries management policies. This is evident in increased government support for research projects and is leading to changes in fisheries legislation. Recently, a more decentralized approach to fisheries management has

begun to take hold in government policy. In addition, social scientists have complained of the lack of attention to social dimensions of the fishery, questioning the validity of the bioeconomic paradigm for fisheries management. As a result of these questions, the participation of communities of fishers in discussions of regulatory measures is becoming the norm. Despite these changes, however, a mathematical model which takes into account the ecological, cultural, and social characteristics of the region has not yet been developed. On the other hand, the search for solutions cannot wait for the accumulation of data sets covering time periods long enough to permit more sophisticated approaches. Practical needs are imposing a system of resource evaluation which combines theoretical and empirical, conventional and alternative methods, and the development of multidisciplinary approaches, for the adoption of management measures.

Introduction

Fisheries of the Middle Amazonian region of Brazil (Fig. 1) have the characteristics of a typical tropical artisanal freshwater fishery in the initial stages of development. More than 200 species are captured utilizing a diversity of gear types, ranging from the bow and arrow to large gillnets. While the total catch may not be significant when compared to Brazil's industrial marine fisheries, this fishery does involve a large number of people and plays an important role in the local diet and in the regional economy (Isaac and Barthem 1995). Regional demand for fish is high. Studies indicate that consumption of fish is on the order of 360g per person per day in the Middle Amazon (Cerdeira et al. 1997) and can reach 500g per person per day in the Central Amazon (Batista et al., in press).

Dominant fishing techniques vary from region to region within the Amazon Basin. In part, this variability is due to the fact that until recently regulations of fishing gear varied from state to state and regionally, within individual states in the Amazon basin. Seines were prohibited in the inland waters of the state of Pará. However, they were permitted just upstream in the state of Amazonas. Trawls are only permitted offshore. These regulations have recently been revised. However the situation in the Middle Amazon reflects this earlier management regime. Here, the operational unit of fishing consists of small wooden canoes (3-5 meters) with one or two fishers. Fishers may travel up to a thousand kilometers or more from port to fish in the innumerable rivers and floodplain lakes of the region. The catch is purchased from local fishers and transported to market in larger vessels equipped with ice boxes of limited adequacy for preserving fish.

In addition to fishing, the majority of fishers also engage in other productive activities such as farming, animal husbandry, and temporary employment. Some fishers fish for subsistence. However, the large majority sell part of their catch in regional urban markets. A few in this latter group are more specialized, live in the city, have more capital and larger

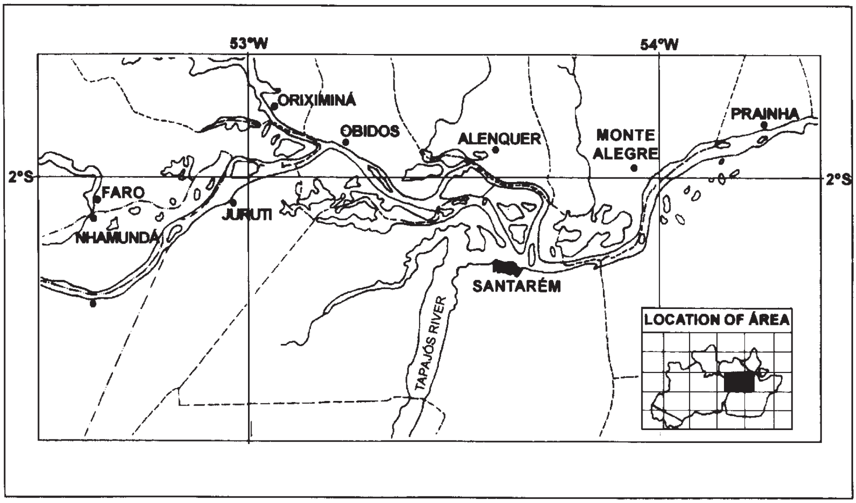


Figure 1. Map of the Middle Amazon.

vessels, which give them greater fishing capacity. This small group of large-scale fishers is responsible for a disproportionately large share of the total catch.

The lack of a continuous record of reliable information on fishing effort, catch, age structure, and life history characteristics of many fish stocks in the region makes it difficult to apply conventional methods for stocks assessment and management. Only in the last four or five years have research projects begun to systematically collect data. Although the necessary data are becoming available, the high biological and technological diversity of the fisheries makes it unlikely that conventional methods for stock assessment and regulation will be sufficient to manage fisheries in the Middle Amazon region.

As a consequence of this complexity, the evaluation of Middle Amazonian fisheries resources and the formulation and implementation of fisheries management policies has followed two alternative paths: (1) a classic scientific approach resulting from the analysis of catch and fishing effort data, as well as biological data collected in the last few years; and (2) an empirical approach, more social and almost anthropological in character, which considers the user groups' (residents and fishers) opinions, considerations and perceptions of the environment and of the resource. This paper summarizes existing information on the characteristics of Middle Amazonian fisheries and describes the results of these two approaches taken in the search for adequate methods for fisheries management.

Methods

Data on catch, fishing effort per species, characteristics of fisheries, and marketing of catch have been collected daily at landing sites in Santarém since 1992, through interviews with the person in charge of every boat landing at the city. Santarém receives fish from the entire Middle Amazon, a region extending from Parintins upstream to Prainha (Fig. 1). A description of the fisheries was done based on a review of a factor analysis of the landing data undertaken in Isaac et al. (1996). Later a general linear model (GLM) was calculated using as dependent variables the coefficients of the factors estimated earlier and as independent variables the characteristics of the fishery such as site, date, gear type, fishing effort, market, etc. In addition, growth curves and mortality rates were calculated for three species. Curves of yield per recruit were fitted in accordance with Beverton and Holt's (1957) model and compared with the similar result from the literature.

Information on the number of fishers, and the geographic, socioeconomic, and cultural characteristics of fishing communities was obtained through surveys conducted in all the riverside communities of the region surrounding Santarém. Observations regarding the causes and development of fishing conflicts in these communities were recorded through interviews with community residents, local leaders and others involved, as well as from reports received by the federal environmental agency, IBAMA (Instituto Brasileiro de Meio Ambiente e Recursos Renováveis).

Data on fishing strategies in community lake management were obtained through a two year study of two floodplain fishing villages. The fishing activity of approximately 25 households was monitored through interviews conducted on a daily basis during the last week of each month. Data on size and composition of the catch, fishing effort, consumption, gear repair, and marketing were recorded.

Results and Discussion

Description of the Fishery

The use of multivariate techniques (PCA and GLM) in the analysis of data on effort, catch, and marketing of fish in the city of Santarém made it possible to establish that commercial fishing activity can be characterized according to fishing capacity, target species, and fishing technology (Isaac et al. 1996).

The main commercial fishery has greater productivity and fishing capacity using longlines, and large drifting and fixed gillnets. The fishery concentrates primarily on stocks of migratory catfish (*Brachyplatystoma flavicans, B. filamentosum, B. vaillantii, Pseudoplatystoma fasciatus, P. tigrinus*, etc.) which move upstream in the main river channels during the low water season (September-October). During the off season, fishers shift to the floodplain lakes where they concentrate on the small catfish "ma-para," *Hypopthalmus* spp. Most of the commercial catfish catch is purchased by local processing plants and then transported to other states for consumption.

Sedentary species such as the tucunaré-*Cichla* spp., pirarucu-*Arapaima gigas*, and pescada-*Pagioscion* spp. or those which use lakes as trophic habitat during a part of their life cycle, such as the tambaqui-*Colossoma macroponum* and pacu-*Myleus* spp., are the focus of the less professionalized and smaller-scale fisheries undertaken by floodplain residents of the region. Using similar gear, they exploit the lakes and local channels of the floodplain near their homes. Fish derived from this lake fishery is destined for household subsistence and sale in regional markets.

Stock Assessment

On the basis of length data we obtained estimates of growth and mortality parameters for the period 1992-93. That permitted the calculation of a yield-per-recruit model and conclusions on the status of some commercially important fish species. The results (Table 1) indicate that *Colossoma macroponum* (Isaac and Ruffino 1996), *Pseudoplatystoma tigrinum,* and *Pseudoplatystoma fasciatum* are overexploited, being caught at very small sizes. In the same fashion, research on *Brachyplatystoma vaillantii*, based on data from the industrial trawl fishery in the Amazon estuary, indicated growth overfishing for this species (Barthem 1990, Barthem and Petrere 1995). Considering its long life cycle, its reproductive and parental care behavior (Neves 1995), and its intense exploitation for more than a century (Veríssimo 1895), we assume that *Arapaima gigas*, one of the largest species in the basin, also suffers growth overfishing. On the other hand, in species with more opportunistic ("r") life strategies, such as the pescada, (*Plagioscion* spp.) of the family Sciaenidae, it seems that environmental factors, such as the velocity and intensity of floods, explain the success or failure of recruitment, and consequently the total catch, better than the intensity of fishing effort (Annibal 1982, Merona 1993). Many authors argue that other impacts of human activity such as the construction of hydroelectric dams, deforestation of the margins of rivers and lakes, and pollution by heavy metals have a more devastating effect on fish stocks than does fishing activity itself, which in general can still be considered fairly moderate when one takes into account the great diversity of Amazonian fishes (Goulding 1983, Bayley and Petrere 1989, Pfeiffer et al. 1990, Leite 1993, Ribeiro et al. 1995).

Conventional Management Measures

Fisheries management in the Middle Amazon has until recently been conducted in the conventional centralized manner. The federal government instituted decrees, though not always based on scientific research, which regulated fishing activity on the basis of classic methods such as prohibition of fishing during reproductive migrations, limits on mesh size, minimum

Table 1. Populational parameters and stock assessment for some commercial exploited fish species in the Middle Amazon.

Species	Year	L_∞ (cm)	K (1/yr)	M (1/yr)	F (1/yr)	Z (1/yr)	L' (cm)	E	E_{max}
Colossoma	1992	121.2	0.229	0.45	0.95	1.40	32.5	0.68	0.47
macropomum [a]	1993	118.5	0.226	0.44	0.93	1.37	41.5	0.68	0.48
Pseudoplatystoma	1992	181.0	0.260	0.43	1.76	2.19	76.0	0.80	0.59
tigrinum [b]	1993	184.0	0.300	0.47	1.51	1.98	74.5	0.76	0.52
Pseudoplatystoma	1992	178.0	0.330	0.51	1.89	2.40	77.5	0.79	0.60
fasciatum [b]	1993	169.0	0.270	0.45	1.45	1.90	83.5	0.76	0.66
Brachyplatystoma	1992	161.0	0.220	0.40	0.47	0.87	73.5	0.54	0.60
flavicans [b]	1993	167.5	0.204	0.38	0.84	1.22	82.5	0.69	0.66

Source: [a] Isaac and Ruffino 1996; [b] Present study.

L_∞ = von Bertalanffy's asymptotic length

K = von Bertalanffy's growth constant

M = instantaneous rate of natural mortality

F = instantaneous rate of fishing mortality

Z = instantaneous rate of total mortality

L' = smallest length of fish fully recruited

E = observed explotation rate

E_{max} = optimum explotation rate, estimated according Beverton and Holt (1957)

t_0 = von Bertalanffy parameter for "age at 0 length"

Note: t_0 was fixed as 0

sizes, and the prohibition of especially noxious gear such as explosives and poisons (Isaac et al. 1993). The enforcement of this great variety of measures was not effective.

Based on the results of biological studies (Ruffino and Isaac 1995, Isaac and Ruffino 1997) and yield-per-recruit models, the legislation regulating fisheries has begun to be simplified and systematized, correcting inconsistencies and eliminating unnecessary measures (Isaac et al. 1993, Isaac and Barthem 1995). At present, legislation regulating minimum size of capture exists for only four species: *Arapaima gigas*, *Pseudoplatystoma tigrinus*, *P. fasciatus*, and *Colossoma macroponum*. In addition, fishing for certain species of characins is prohibited during the period of upstream reproductive migration, December through February. There are also restrictions on the number of vessels and the mesh size used in the industrial piramutaba *(Brachyplatystoma vaillantii)* fishery.

Informal Fisheries Management by Communities

Despite the existence of these regulations, government agencies have neither sufficient personnel, nor resources, to enforce compliance, so that commercial fisheries have largely developed in a regulatory vacuum. As a result, fishing conflicts between floodplain communities and outside commercial fishers have proliferated in the last two decades. Due to the limited presence of the state, many floodplain communities have begun to develop and implement their own management regulations, referred to as "fishing accords" (acordos de pesca), to serve their own interests. Rules regulating fishing activity are usually defined in regional meetings involving local residents who share the same lake system, usually without the participation of the state. In some cases, municipal fishers unions have taken an active role in support of community management initiatives.

A variety of different measures can be employed in the accords. Frequently, the lake system is divided into zones with different types of use (no fishing, subsistence, or commercial fishing). Local fishers may, in this case, restrict access to the lakes, which they consider private or common property (Ruffino, in press). In other cases, some gear types, such as gillnets, are prohibited, permanently or seasonally. Another measure concentrates on storage capacity and may involve restrictions on use of ice or limitations on the size of ice boxes.

Local management regulations tend to differ from those of IBAMA. For example, seasonal restrictions on fishing effort focus on the low water season, October through December, rather than the spawning season, December through February. The intent is to reduce pressure during the period when fish in the lakes are concentrated in small bodies of water and are therefore most vulnerable. Fishers know that rising water, from November on, provides natural protection during the spawning season.

Fishing accords are publicized locally by informal means. Enforcement is undertaken by local residents, and varies considerably in effectiveness. In some cases, due to the erratic often arbitrary nature of enforcement, fishing accords, which were created to reduce conflicts, can lead to higher levels of conflicts between floodplain communities and outside fishers. In these cases, the intervention of local governmental authorities is sought to mediate conflicts.

Over the last three years, IBAMA has come to recognize that the regulatory process developed by the local communities can be a useful tool for managing fisheries in the Amazon basin. Fishing accords are now considered potentially "legalizable." Criteria for the integration of fishing accords into the formal regulatory framework have been defined. These criteria prohibit the establishment of accords which exclude outsiders and favor local residents only, but promotes the adoption of measures which must be respected by both local and outside fishers.

The efficacy of this management model for lake fish stocks, though little studied, is questionable, since many commercially important species have complex migratory cycles, and use a variety of environments throughout their life cycle, some protected by "lake reserves" and others not. However, a comparison of fishing productivity in two lake systems near the city of Santarém, one with a system of community management and the other without it, demonstrated that the managed lake was approximately twice as productive as the unmanaged lake for all gear types used in common (McGrath et al. 1994).

The problem of the efficacy of lake fishing accords is exacerbated by their lack of legal legitimacy. Because most lake accords contradict the concept of "open access," which is the basis of existing fisheries legislation, and impose restrictions on fishing activity, such as the prohibition of gillnets, which are not supported by existing legislation, most "lake accords" cannot be legally recognized under the existing criteria.

In Search of a More Adequate Approach for Fisheries Management

Fisheries science and the classic methods of fisheries management are marked by uncertainty. This is a consequence of the variability of natural and social phenomena, be it spatio-temporal or economic. The sources of uncertainty are not always predictable and in many cases constitute "surprises" in models (Hilborn and Walters 1992).

The process of managing fisheries is essentially different from "pure science." Factors such as political interests, economic convenience, and the acceptability of certain measures are important factors. In the case of the artisanal fisheries of the Amazon, implementation of measures to regulate fisheries is a difficult task, due to the size of the basin, the difficulties of transport and communication, and the isolation of fishing communities. From interviews with fishers and community members it is apparent that the lack of participation of users in the process of planning and monitoring management, traditionally undertaken by the state, is a major factor in the lack of compliance with existing regulations. Through fishing accords civil society is developing alternatives to the conventional management model and its priorities in regulating fishing activity.

Taking into account the absence of historical data, the uncertainties of classic evaluation models, and the existence of traditional forms of management that have been implemented informally in regulating the fisheries, one of the most important conclusions of our work is the demonstration that any management policy can only be successful if created and implemented by the government, but with the active participation of the actors involved in utilizing the resource.

Despite the fact that communities are not allowed to impose sanctions on infractors, the recognition of fishing accords is an important landmark in the history of Brazilian fisheries management. Sen and Nielsen (1996) define a continuum of management approaches which ranges from instructive to informative depending on the relative roles of the state and user groups (Fig. 2). Until 1992, when the Program for Management of River Basins (Fischer et al. 1992) was implemented, existing management was based on the Instructive Model, in which the government simply informed users of the decisions that had been made. Since 1993, management has changed to the Consultive Model, where IBAMA consults users (municipal

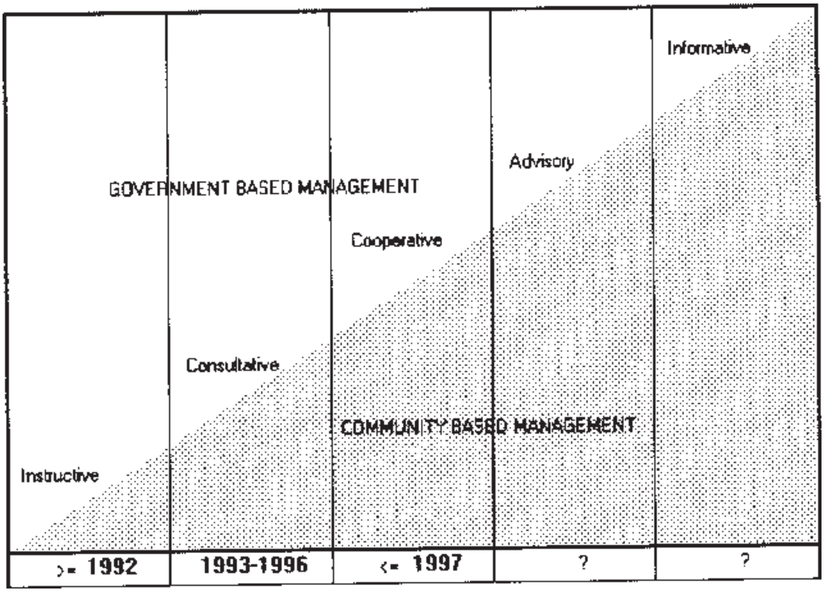


Figure 2. Development of the fisheries management in the Middle Amazon in the last years, according to the degree of involvement of user groups. (Adapted from Sen and Nielsen 1996.)

fishers unions) and researchers regarding the species to be included in the three-month spawning season prohibition on fishing activity. Finally, with the recent publication of an internal document (IBAMA 1997), government managers and user groups are beginning to cooperate in developing management strategies, although not yet as equal partners (Fig. 2).

This new management structure involves decentralizing decision-making from federal to state, municipal, and community levels. At the local level, inter-community councils (Conselhos Regionais de Pesca) are being organized which are composed of representatives of all communities that are located around the margins of individual lake systems. These councils are responsible for developing and implementing fisheries management policies in the form of "fishing accords." While most councils do not include local ranchers or outside commercial fishers, these stakeholders are usually represented in the meetings where local fishing accords are hammered out. (The support of local ranchers is critical to the success of these fishing accords, since they frequently claim ownership of much of the land surrounding floodplain lake systems.) Once the fishing accords have been approved by the fishing council, they are submitted to IBAMA for evaluation. If they meet IBAMA's criteria, they are formalized in a "portaria" or administrative law.

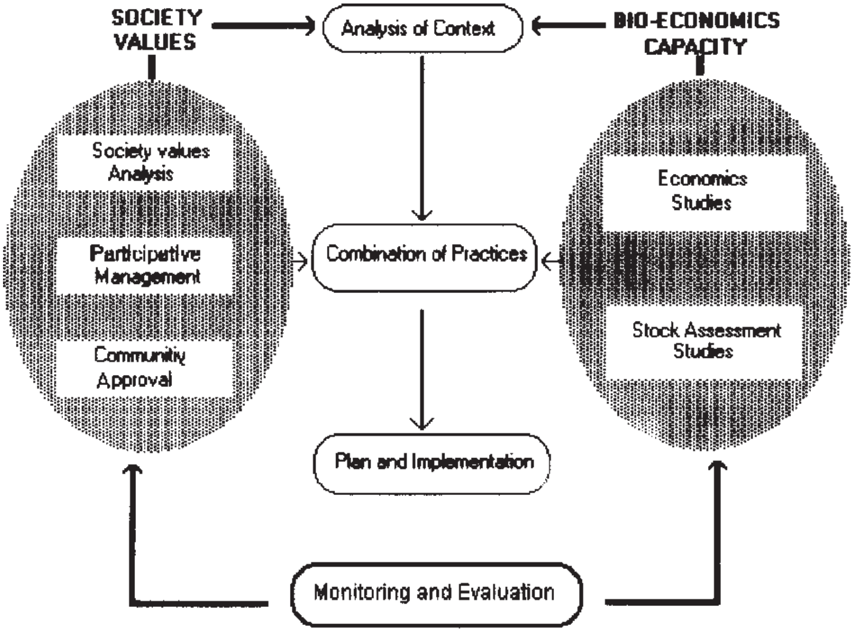


Figure 3. Proposed management strategies for the Middle Amazon fisheries, combining conventional (bio-economics studies) and alternative approaches (community values). (Adapted from Bormann et al. 1994.)

To assist in the enforcement of lake accords, IBAMA is training community environmental agents. These agents have the authority to cite, but not to arrest, infractors, and must call upon IBAMA or local police to make arrests or impose sanctions specified in local fishing regulations. In this fashion, IBAMA is incorporating local users groups and community institutions into the formal structure of fisheries management. At the same time, these grassroots management organizations are making possible more effective control of fishing activity throughout the Middle Amazonian floodplain.

Through this process, fisheries management is evolving toward the substitution of a centralized model with a new model in which both fishers and fisheries administrators will jointly manage the resource (Fig. 2). Communities are also being made aware that fishing accords must comply with existing legislation, conserving the resource, but without excluding use by some or privileging use by others. This new approach is based on the supposition that sustainability is possible if the resource is managed as a common property resource involving all local fishers and not as a resource for exclusive use.

At the same time, social scientists have been intensively engaged in strengthening organizations representing fishers, so that they will be prepared to assume greater responsibility in the process of comanagement. One example of this is the increase in the number of active members in the three most important municipal fishers unions in the region (Santarém, Alenquer, and Monte Alegre) which increased from 1,400 to 5,000 members.

Conclusions

The development of Amazonia over the last thirty years has been characterized by a high rate of demographic growth and the acceleration of extractive activities, especially logging, fishing, and hunting, combined with deforestation for farming and ranching. However, traditional extractive activities and farming on the floodplain are in crisis, threatening the economic base of the Middle Amazonian floodplain population (McGrath et al. 1993). These factors have caused floodplain smallholders to become increasingly dependent on fishing for income and subsistence (Furtado 1988).

Stock evaluation models and classic management measures focus attention on the biology and population dynamics of fisheries. The fisher, in this case, the "exploiter," is considered the principal threat to fisheries resources. In the development of a new perspective the environment that the fish use is a fundamental element for the maintenance of the resource (Goulding and Ferreira 1996). For the majority of fish species inundated areas play a key role in maintaining the productivity of the fishery. The floodplains provide food for fish growth, habitats for reproduction, and refuges for protection from predators (Junk 1984a,b; Goulding 1983; Bayley and Petrere 1989, Petrere 1989). In this case, the principal cause of the environmental transformations is associated with powerful economic interests (ranchers and loggers) generally unconcerned with the local population. Fishers who are floodplain residents understand the importance of floodplain habitat for fisheries production and seek to develop appropriate forms of control, but do not have the power to combat external influences.

Reconciling environmental conservation with development has been a constant in the environmental discourse of governments and Brazilian civic organizations in recent years. However, an acceptable scientific model has not yet been formulated which can take into account ecological, cultural, and situational dimensions of regional fisheries, or which provides a quantitative base for making predictions that can contribute to the process of developing more effective management policies and regulatory measures. Our evolving efforts are leading to the development of a multidisciplinary proposal which takes into account both the results of the application of classic methods (bio-economic studies), as well as the experience and perceptions of fishing communities (social values). At the moment the work concentrates on implanting measures for monitoring the process, through an iterative management system of the "adaptive" type, in which the correction of possible errors follows experimental attempts at consensus (Fig. 3).

The search for solutions cannot wait for the development of complex models of evaluation, nor the need to collect data over long periods. Practical necessity is imposing a system of resource evaluation which involves the combination of theoretical and empirical models, conventional and alternative, and the development of multidisciplinary approaches.

The Amazon floodplain ecosystem is based on the energy trilogy: Forest-Fish-Man. The close relations between them do not permit us to treat the problem of management from only one perspective. Understanding the relations between ecosystems and human activities seems to be the only way to achieve a process of development that reconciles the needs of human populations with those of environmental conservation. Resource use and its consequences need to be predicted and policies modified so that ecologically sound management strategies can be developed which seek to optimize returns from the fishery in the face of different and frequently competing interests (May 1995). Our path is uncertain, but this seems to the only viable option under the present circumstances.

Acknowledgment

The data which provide the basis for this paper were collected by research teams associated with two projects based in the city of Santarém in the state of Pará: (1) Project IARA, of the federal environmental agency, IBAMA (Instituto Brasileiro do Meio Ambiente e dos Recursos Naturais Renováveis) with technical support from the German government and the National Council for Scientific and Technical Development (CNPq), and (2) Project Varzea, of the Amazon Institute for Environmental Research (IPAM) and the Nucleus for Higher Amazonian Studies (NAEA) of the Federal University of Pará, with financial support from the World Wide Fund for Nature (WWF-UK) and the British government (Department for International Development).

References

Annibal, S.R.P. 1982. Avaliação bio-ecológica e pesqueira das “pescadas” (*Plagioscion squamosissimus* Heckel, 1840 e *P. montei* Soares, 1978) no sistema “Lago do Rei”- ilha do Carero - AM-Brasil. Tese de mestrado, INPA/FUA, Manaus.

Barthem, R.B. 1990. Ecologia e pesca da piramutaba *(Brachyplatystoma vaillantii).* Tese de Doutorado em Ciências. UNICAMP, Campinas-SP. 268 pp.

Barthem, R.B., and M. Petrere Jr. 1995. Fisheries and population dynamics of *Brachyplatystoma vaillantii* (Pimelodidae) in the Amazon Estuary. In: N.B. Armantrout (ed.), Condition of the world’s aquatic habitat. Proceeding of the World Fisheries Congress, Theme 1. Oxford and IBH Publishing Co. Pvt., New Delhi, pp. 329-340. (ISBN 81-204-0962-0)

Batista, V.S., A.J.I. Silva, C.E.C Freitas, and D. Freire-Brasil. In press. Characterization of the fishery in riverine communities in the Low-Solimões/High-Amazon region. Fisheries Management and Ecology.

Bayley, P.B., and M. Petrere Jr. 1989. Amazon fisheries: Assessment methods, current status and management options. Can. Spec. Publ. Fish. Aquat. Sci. 106:385-398.

Beverton, R.J.H., and S.J. Holt. 1957. On the dynamics of exploited fish populations. Fish Invest., Ser. II, 19. 533 pp.

Bormann, B.T., M.H. Broockes, E.D. Ford, A.R. Kiester, C.D. Oliver, and J.F. Weigand. 1994. A framework for sustainable ecosystem management. Gen. Tech. Rep. PNW-GTR-331. Portland, OR, USDA, Forest Service, Pacific Northwest Research Station. 61 pp.

Cerdeira, R.G.P., M.L. Ruffino, and V.J. Isaac. 1997. Consumo de pescado e outros alimentos pela população ribeirinha do Lago Grande de Monte Alegre, PA, Brasil. Acta Amazonica, Manaus 27(3):231-227.

Fischer, C.F.A., A.L. das G.A. Chagas, and L.D.C. Dornelles. 1992. Pesca de águas interiores. IBAMA. Coleção Meio Ambiente. Série Estudos de Pesca 2. 32 pp.

Furtado, L.G. 1988. Os cablocos pescadores do Baixo Rio Amazonas e o processo de mudança social e econômica, In: A.C. Diegues (ed.), Ciências sociasi e o Mar no Brasil. II Programa de Pesquisa e Conservação de Áreas Úmidas no Brasil, São Paulo, pp. 108-203.

Goulding, M. 1983. Amazon fisheries. In: E.F. Moran (ed.), The dilemma of Amazonian development. Westview Press, Boulder, CO, pp. 189-210.

Goulding, M., and E. Ferreira, 1996. Pescarias Amazônicas, proteção de habitats e fazendas nas várzeas: Uma visão ecológica e econômica. (Unpublished.)

Hilborn, R., and C.J. Walters. 1992. Quantitative fisheries stock assessment: Choice, dynamics and uncertainty. Chapman and Hall, New York. 570 pp.

IBAMA. 1997. Administração participativa: Um desafio à gestão ambiental. Instituto Brasileiro de Meio Ambiente e Recursos Renováveis. (Unpublished.)

Isaac, V.J., and R.B. Barthem. 1995. Os recursos pesqueiros da Amazônia brasileira: Estado atual, problemas e perspectivas de manejo. Bol. Mus. Para. Emilio Goeldi. Sér. Antropol. Belém 11(2):295-339.

Isaac, V.J., and M.L. Ruffino. 1996. Population dynamics of tambaqui, *Colossoma macropomum* Cuvier 1818, in the Lower Amazon-Brazil. Fish. Manage. Ecol. 3(4):315-333

Isaac, V.J., and M.L. Ruffino. 1997. Growth and mortality of *Pseudoplatystoma tigrinum* and *P. fasciatum* (Pimelodidae) in the Lower Amazon, Brazil. Abstracts of the International Symposium on Biology of Tropical Fish, Manaus, October 1997, p. 14.

Isaac, V.J., A. Milstein, and M.L. Ruffino. 1996. A pesca artesanal no Baixo Amazonas: Análise multivariada das capturas por espécie. Acta Amazonica 26(3):185-208.

Isaac, V.J., V.L.C. Rocha, and S. Mota. 1993. Considerações sobre a legislação da "piracema" e outras restrições da pesca da região do Médio Amazonas. In: L. Furtado, W. Leitão, and A.F. Mello (eds.), Povos das águas: Realidade e perspectivas na Amazônia. MCT/CNPq/MPEG, Belém, pp. 188-211.

Junk, W.J. 1984a. Ecology of the várzea, floodplain of Amazonian white-water rivers. In: H. Siolo (ed.), The Amazon: Limnology and landscape ecology of a mighty tropical river and its basin. Dr. W. Junk Publications, Netherlands, pp. 215-243.

Junk, W.J. 1984b. Ecology, fisheries and fish culture in Amazonia. In: H. Siolo (ed.), The Amazon: Limnology and landscape ecology of a mighty tropical river and its basin. Dr. W. Junk Publications, Netherlands, pp. 443-476.

Leite, R.A.N. 1993. Efeitos da Usina Hidroelétrica de Tucuruí sobre a composição da ictiofauna das pescarias experimentais de malhadeiras realizadas no baixo rio Tocantins (Pará). Tese doutoramento, INPA/FUA, Manaus. 133 pp.

May, P.H. 1995. Economia ecológica. Aplicações no Brasil. Ed. Campus. 179 pp.

McGrath, D., F. Castro, and C. Futema. 1994. Reservas de lago e o manejo comunitário da pesca no Baixo Amazonas: Uma avaliação preliminar. In: M.A. D'Incao and I.M. Silveira (eds.), A Amazônia e a crise da modernização. MPEG, Belém, pp. 389-402.

McGrath, D., J. Calabria, B. Amaral, C. Futema, and F. Castro. 1993. Varzeiros, geleiros e o manejo dos recursos naturais na várzea do Baixo Amazonas. Cadernos do NAEA (11):91-125.

Merona, B. 1993. Pesca e ecologia dos recursos aquáticos na Amazônia. In: L. Furtado, A.F. Mello, and W. Leitão (eds.), Povos das aguas realidade e perspectiva na Amazônia. Belém, MCT/CNPq/MPEG, pp. 159-185.

Neves, A.M.B. 1995. Conhecimento atual sobre o pirarucu *Arapaima gigas* (Cuvier, 1817). Bol. Mus. Para. Emilio Goeldi. Sér. Zool. 11(1):33-56.

Petrere Jr., M. 1989. Fish stock assessment in the Amazon. In: Amazonia: Facts, problems and solutions. USP Publications, São Paulo, pp. 391-404.

Pfeiffer, W.C., O. Malm, C.M.M. Souza, L.D. Lacerda, and E.G. Silveira. 1990. A ameaça do mercúrio nos garimpos. Ciência Hoje 11:10-20

Ribeiro, M.C.L.B., M. Petrere Jr., and A.F. Juras. 1995. Ecological integrity and fisheries ecology on the Araguaia-Tocantins river basin, Brazil. Regulated Rivers: Research and Management 10:31-45

Ruffino, M.L. In press. Desenvolvimento pesqueiro no Baixo Amazonas. In: C. Padoch, J.M. Ayres, A. Henderson, and M. Pinedo-Vasquez (eds.), The diversity, development and conservation of the Amazon floodplain.

Ruffino, M.L., and V.J. Isaac. 1995. Life cycle and biological parameters of several Amazon fish species. NAGA, The ICLARM Quarterly 18(4):41-45.

Sen, S., and J.R. Nielsen, 1996. Fisheries co-management: A comparative analysis. Marine Policy 20(5):405-418.

Veríssimo, J. 1895. A pesca na Amazônia. Livr. Alves, Rio de Janeiro. 137 pp.

From Fisheries Assessment Uncertainty to Risk Analysis for Immediate Management Actions

Stratis Gavaris
Department of Fisheries and Oceans, Biological Station, St. Andrews, N.B., Canada

Alan Sinclair
Department of Fisheries and Oceans, Gulf Fisheries Center, Moncton, N.B., Canada

Abstract

The translation of fisheries assessment uncertainty to a risk analysis for immediate management actions in relation to established harvest strategies has influenced the development of recent management plans in Atlantic Canada. Graphics depicting the risk of failing to achieve exploitation and biomass targets for a range of quotas were included with the scientific resource evaluations for Gulf of St. Lawrence cod and eastern Georges Bank haddock. The risk analysis framework for providing advice on forecast scenarios changed the emphasis from recommending the “best” action consistent with the harvest strategy to measuring the performance of alternate actions relative to the harvest strategy. Evolution away from a single option promoted communication between stakeholders and scientists and provided the opportunity for stakeholders to effectively participate in the decision process. Success was dependent on management institutions to cope with implementing a range of actions, timely use of current information, deliberate and thorough presentation of risk results, and previous exposure to risk analysis.

Introduction

There is a growing demand in fisheries management for information about uncertainties. Modern fish stock assessment methods can complement estimates of population state variables with statistics for those estimates. These statistics are increasingly being used for diagnostic purposes during resource evaluation. For fisheries management, however, the consequences of specific management actions are desired. For example, with the increased attention to the Precautionary Approach, fisheries managers need to know what the chances are that a specified management action might result in a fishing mortality rate exceeding that associated with maximum sustainable yield. The uncertainty in the results of fishery stock assessments propagates as risk for management actions. The decision process can be more effective when informed of these risks.

Risk analysis has received considerable attention in fisheries recently (Smith et al. 1993, Cordue and Francis 1994, McAllister et al. 1994, Megrey et al. 1994, Rosenberg and Restrepo 1994). Much of the attention has concentrated on risk analyses to assess harvest strategies while considering the uncertainties about population dynamics (e.g., Ianelli and Heifetz 1995a, 1995b). This approach may be used to conduct comprehensive probabilistic evaluations, generally through simulation, of the comparative performance of alternative strategies with respect to achieving stated long-term objectives. The strategies must be translated into well-defined management rules for determining actions and acceptable risk levels must be specified. Some fisheries management forums however, including those in Atlantic Canada, do not employ rigid management rules and permit variation in the acceptable level of risk for achieving harvest targets from year to year based on consideration of the prevailing circumstances. These management regimes do not lend themselves readily to evaluation of how well they perform in achieving long-term objectives, though the established harvest strategies have generally evolved from experience with the fishery dynamics. These regimes would benefit from quantification of the risk that contemplated actions might not achieve the established harvest strategies. This topic has received limited attention (Restrepo et al. 1992) and its application has not been well developed or promoted.

We present how the uncertainty in fisheries stock assessments was used to derive risk analysis for immediate management actions using two case studies, the southern Gulf of St. Lawrence cod and eastern Georges Bank haddock, to illustrate the approach. We highlight the important features of those risk analyses and how the results were communicated graphically for a range of actions. Finally, we discuss how this information was received and how it influenced the fisheries management decision process.

Methods

Risk analysis for immediate management decisions involves three components, a set of well defined actions, identification of one or more forecast event parameters, e.g. projected biomass, to be used as performance measures and harvest strategy targets for those forecast event parameters. We adopt a simple definition of risk as the probability that a forecast event parameter is less than or equal to the harvest strategy target for a given action, Pr{forecast event ≤ harvest strategy target | action}. Forecast event parameters are derived from the assessed fishery population state parameters, e.g. abundance or fishing mortality rate. Therefore, uncertainty in the stock assessment is accounted for as uncertainty in the state of the forecast event parameter. Given a harvest strategy then, the uncertainty in fisheries stock assessments may be translated into the risk associated with immediate fisheries management actions.

Harvest strategies may contain several elements. In Atlantic Canada, the customary target fishing mortality rate is equivalent to $F_{0.1}$ (Gulland and Boerema 1973), a yield per recruit strategy largely aimed at preventing growth overfishing. We capitalized on the general acceptance of this reference and thus included a fishing mortality rate of $F_{0.1}$ as one of the targets for the harvest strategy. The Atlantic Groundfish Management Plan has also recognized that recruitment overfishing considerations may at times be paramount, and espouses lower exploitation in these circumstances to promote rebuilding of resources. The two resources we examined were considered to be in a depleted state relative to historic levels and there has been concern regarding how this might impact their reproductive capacity. Accordingly, we included as a second element of the harvest strategy the requirement that the biomass should increase from current levels. For both the cod and haddock cases, probabilistic results were developed for these two forecast event parameters. The fishery for southern Gulf of St. Lawrence cod was closed to directed fishing in 1993 due to low stock biomass. The Fisheries Resource Conservation Council (FRCC) which advises the Canadian Minister of Fisheries and Oceans on fishery management measures, suggested in their discussion paper that a stock biomass of 115,000 t for ages 5 to 15 was as an appropriate reopening threshold for this stock (FRCC 1996). A third element of the harvest strategy for the cod stock then, was that the stock biomass should exceed the threshold.

In Atlantic Canada, management involves a variety of measures including catch quotas, gear regulations, area/season closures, and other tactics. Typically, most measures are established and retained indefinitely but catch quotas are adjusted annually. Consequently, management deliberations are dominated by discussion on the consequences of varying the catch quota. Accordingly, the risk analyses focused on the probability of satisfying the elements of the harvest strategy for a range of catch quotas.

Risk analysis follows from an evaluation of the uncertainty in the estimates of the forecast event parameters. The translation of assessment uncertainty to management risk is depicted by the graphical representation of the procedure (Fig. 1). A three dimensional graph of probability as a function of values of the forecast event parameter, projected fishing mortality in this example, and of management actions, quota in this example, can be constructed from the cumulative frequency distributions of the fishing mortality over the range of quota options examined. The intersection of this probability surface with an orthogonal plane at the harvest

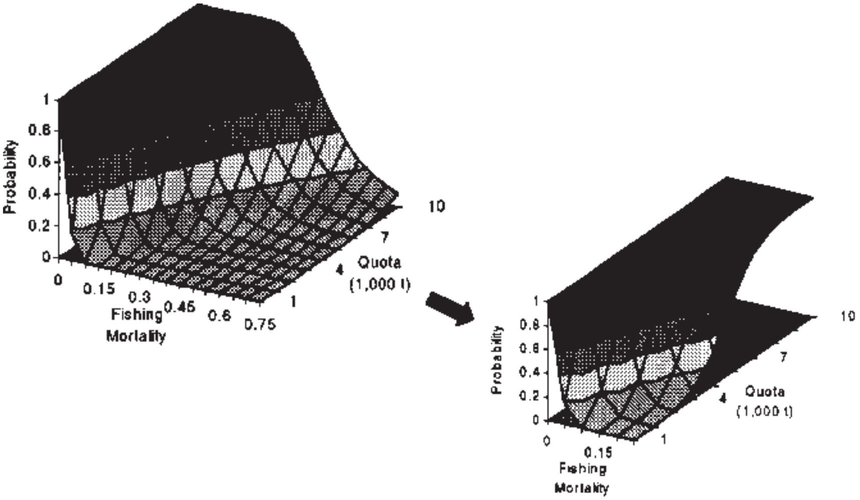


Figure 1. A graphical depiction of how a risk profile is derived from the probability surface that is obtained by assembling the complement of the cumulative frequency distributions for projected fishing mortality over a range of quotas. The intersection of the risk surface with an orthogonal plane at a fishing mortality rate corresponding to the harvest strategy target defines the risk profile.

strategy target, $F_{0.1} = 0.25$ in this example, on the forecast event parameter axis defines the risk profile. Though the risk graphic has the appearance of a cumulative frequency distribution, it conveys different information. On a cumulative frequency distribution graph, the ordinate represents the probability that the statistic is less than or equal to the value of the statistic at that point on the abscissa. On this risk graph, the ordinate represents the probability that the statistic is less than or equal to a harvest strategy target for the action specified by that point on the abscissa. These characteristics make the risk graph readily useable in the decision process.

Cumulative frequency distributions of forecast event parameters are derived from the statistical properties of the population state variables. The adaptive framework, ADAPT, (Gavaris 1988) was used to calibrate a sequential population analysis with independent abundance indices. For both stocks, the model formulation employed assumed that the error in the catch at age was negligible, the errors in the abundance indices were independent and identically distributed after taking natural logarithms of the values, and the annual natural mortality rate, M, was constant and equal to 0.2 over all years and ages. A model formulation using as parameters the natural logarithm of population abundance, $\theta_{a,t'}$ for specified ages *a* at time *t′*, where *t′* = beginning of year 1997 for the haddock case and beginning of year 1996 for the cod case, was considered because of close to linear behavior for such a parameterization (Gavaris 1993). The additional parameters in the model, $\kappa_{s,a}$, are associated with the catchability relationship between population abundance and each of the index sources, *s*. Further specific details on each of the assessments are provided in Sinclair et al. (1996) and in Gavaris and Van Eeckhaute (1997).

ADAPT employs a Marquardt algorithm to solve for the parameters of the nonlinear least squares objective function. The covariance matrix of the parameters was estimated using the common linear approximation (Kennedy and Gentle 1980, p. 476)

$$\mathrm{Cov}(\hat{\theta},\hat{\kappa}) = \hat{\sigma}^2\left[J^T(\hat{\theta},\hat{\kappa})J(\hat{\theta},\hat{\kappa})\right]^{-1} \tag{1}$$

where $\hat{\sigma}^2$ is the mean square residual and $J(\hat{\theta},\hat{\kappa})$ is the Jacobian matrix. The bias of the parameters was estimated using Box's (1971) approximation that assumes the errors are normally distributed

$$\begin{aligned}\mathrm{Bias}\left(\hat{\theta},\hat{\kappa}\right) = \frac{-\hat{\sigma}^2}{2}\left[\sum_{s,a,t} J_{s,a,t}\left(\hat{\theta},\hat{\kappa}\right)J^T_{s,a,t}\left(\hat{\theta},\hat{\kappa}\right)\right]^{-1}\left[\sum_{s,a,t} J_{s,a,t}\left(\hat{\theta},\hat{\kappa}\right)\right] tr \\ \left\{\left[\sum_{s,a,t} J_{s,a,t}\left(\hat{\theta},\hat{\kappa}\right)J^T_{s,a,t}\left(\hat{\theta},\hat{\kappa}\right)\right]^{-1} H_{s,a,t}\left(\hat{\theta},\hat{\kappa}\right)\right\}\end{aligned} \tag{2}$$

where $J_{s,a,t}(\hat{\theta},\hat{\kappa})$ are vectors of the first derivatives and $H_{s,a,t}(\hat{\theta},\hat{\kappa})$ are the Hessian matrices of second derivatives for each element of the objective function with respect to the parameters. The superscript *T* denotes transpose and the expression *tr* represents the trace of the matrix.

Population quantities of interest for management are functions of the estimated parameters. The variance and bias for an arbitrary function of estimated parameters, $\hat{\alpha} = g(\hat{\theta},\hat{\kappa})$, where *g* is the transformation function, were obtained using the methods described in Ratkowsky (1983)

$$\mathrm{Var}\left(\hat{\alpha}\right) = tr\left[GG^T \,\mathrm{cov}\left(\hat{\theta},\hat{\kappa}\right)\right] \tag{3}$$

$$\mathrm{Bias}\left(\hat{\alpha}\right) = G^T \mathrm{Bias}\left(\hat{\theta},\hat{\kappa}\right) + tr\left[W \,\mathrm{cov}\left(\hat{\theta},\hat{\kappa}\right)\right]\Big/2 \tag{4}$$

where *G* is the vector of first derivatives of *g* with respect to parameters and *W* is the matrix of second derivatives of *g* with respect to parameters.

For the cod stock, cumulative frequency distributions of the forecast event parameters were obtained from a parametric Monte Carlo experiment. Statistics for the population abundance, $N_{a,t'}$, at time *t′* were derived from the parameter estimates using equations (3) and (4) where the transformation function is simply $N_{a,t'} = e^{\hat{\theta}_{a,t'}}$. Replicate population abundance values were generated from a lognormal distribution characterized by the calculated variance and centered about the bias adjusted mean. The simulated population abundance replicates were used in standard fisheries projection procedures to obtain results of the three identified forecast event parameters. These were projected fishing mortality rate, F, projected change in ages 5 to 15 biomass, ΔB, and projected absolute biomass for ages 5 to 15, B. The outcomes of 300 generated replicates for each quota option over the range of 0 to 25,000 t in 1,000 t increments were used to construct the cumulative frequency distributions.

For the haddock stock, the forecast event parameters were defined as the projected inverse exploitation rate, $1/u_{t'}$, on ages 4 to 8 and the projected change in the beginning of year biomass, $\Delta B_{t'+1}$, for ages 3 to 9. Inverse exploitation rate was used rather than fishing mortality rate because inverse exploitation rate is a linear function of the population abundance and avoids complications of interpretation caused by transformation bias. For a range of catch quotas in 1997 from 1,000 t to 10,000 t, equations (3) and (4) were used to obtain statistics of these forecast event parameters directly from the model parameter estimates where the transformation functions were based on standard fishery projection procedures. Cumulative frequency distributions characterized by $\mathrm{Var}(1/u_{t'})$ and $\mathrm{Var}(\Delta B_{t'+1})$ and centered about $(1/u_{t'} - \mathrm{Bias}_{1/u_{t'}})$ and $(\Delta B_{t'+1} - \mathrm{Bias}_{\Delta B_{t'+1}})$ respectively were constructed assuming a normal distribution.

The uncertainty characterized by the cumulative frequency distributions was translated into risk by evaluating the probability that the identified forecast event parameters were less than the strategic references for each quota option in the range examined. For cod, the $\Pr\{F_{t'} \le 0.2\}$, the $\Pr\{B_{t'+1} \le 115{,}000\ t\}$ and the $\Pr\{\Delta B_{t'+1} \le 0\ t\}$ were computed while for haddock, the $\Pr\{(1/u_{t'} - \mathrm{Bias}_{1/u_{t'}}) \le 4.5\}$ and the $\Pr\{(\Delta B_{t'+1} - \mathrm{Bias}_{\Delta B_{t'+1}}) \le 0\ t\}$ were computed. Recall that for cod, the population abundance estimates were adjusted for bias prior to calculating projections.

The method used to derive the risk analysis from the assessment results differed somewhat for cod and haddock but Sinclair and Gavaris (1996) considered that both approaches led to similar conclusions. The method used for cod makes assumptions about the distribution of current bias adjusted population abundance estimates and projects those distributional qualities forward using Monte Carlo replicates. The method used for haddock projects the uncertainty of current model parameters to forecast event parameters using analytical approximations and then makes assumptions about the distributions of those forecast event parameters.

The utility of the risk analysis relies on adequate modeling of the uncertainties. The quantification of uncertainty involves various assumptions and approximations. The approach used here captured some of the uncertainty in the estimates of population abundance but did not include stochastic aspects of forecast weight at age, exploitation pattern by age,

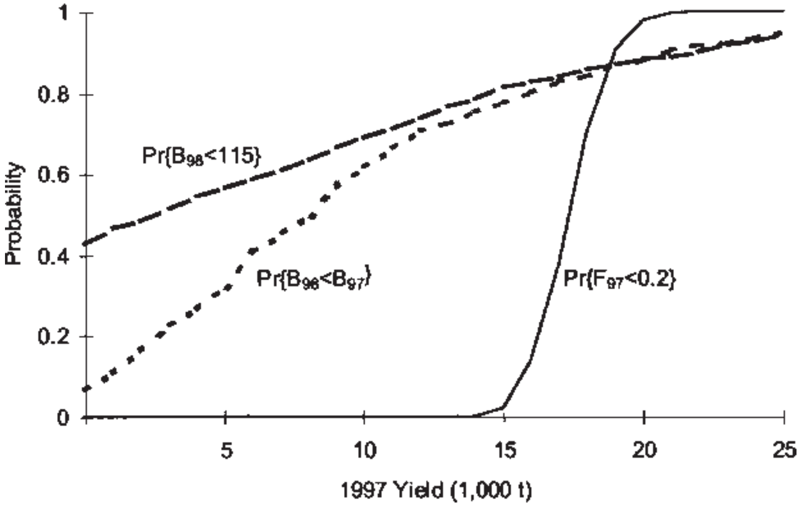


Figure 2. Risk analysis of 1997 quota options for southern Gulf of St. Lawrence cod using three harvest strategy targets indicated that, though fishing mortality rate is not likely to exceed the $F_{0.1}$ reference for yields less than about 17,000 t, the biomass rebuilding objectives dictated considerably lower yields.

and natural mortality. Neither was uncertainty arising from systematic errors in data reporting or model misspecification considered. In addition, assumptions about the distributions of population abundance or forecast event parameters were invoked to derive probabilities. Further research is needed to investigate practical approaches for incorporating additional stochastic elements and to explore the impacts of relaxing distributional assumptions, for example by using modern bootstrap techniques to derive empirical cumulative frequency distributions.

Results

The 1997 management advice for cod was based on an assessment conducted in the spring of 1996 (Sinclair et al. 1996). Forecast projections were computed for 1997 assuming a 1996 catch of 2,000 t to account for the catches by the sentinel survey, selectivity experiments, and bycatches in other fisheries. The probability that the 1998 biomass would be less than the 1996 biomass increased from 7% for a quota of 0 t in 1997 to about 50% for an 8,000 t quota and to 90% for a quota of 22,000 t (Fig. 2). There was a 50% probability that the 1998 biomass would be below the FRCC reopening threshold of 115,000 t for a 1997 catch of 3,000 t. The

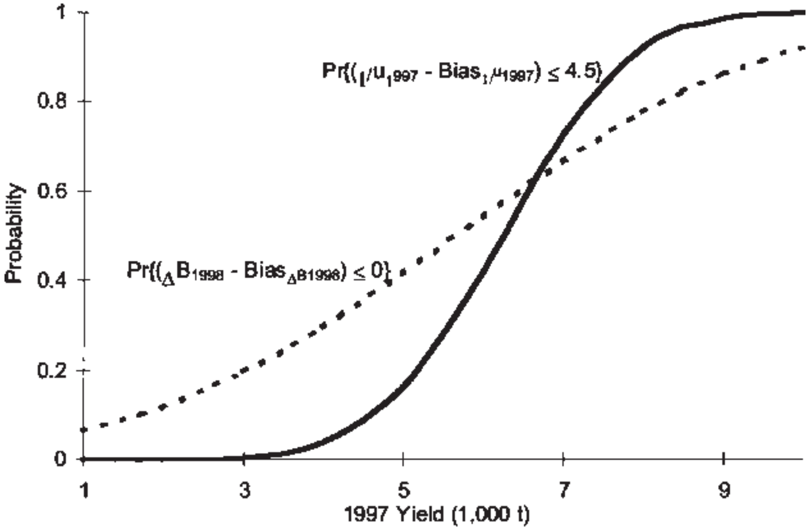


Figure 3. Risk analysis of 1997 quota options for eastern Georges Bank haddock indicated that, for yields less than about 3,000 to 4,000 t, fishing mortality rate was not likely to exceed the $F_{0.1}$ reference and there was a high probability for an increase in biomass.

probability that the 1997 *F* would be greater than 0.2 increased from 0% at 14,000 t to 50% at 17,000 t, and to 100% at 21,000 t.

The 1997 management advice for haddock was based on an assessment conducted in the spring of 1997 (Gavaris and Van Eeckhaute 1997). This was possible because the fishery was not opened until June of that year. Below about a 4,000 t quota there was virtually negligible probability that the inverse exploitation rate would be less than 4.5 (corresponding to an exploitation rate of 22% and $F_{0.1}$ of 0.27) and above about an 8,000 t quota there was virtual certainty that the inverse exploitation rate would be less than 4.5 (Fig. 3). On the other hand, the probability that the difference in biomass between 1998 and 1997 would be less than 0 t only increased from about 30% to about 80% over the same range of quotas. The quota corresponding to 50% probability that the inverse exploitation rate would be greater than 4.5 was about 6,200 t while the quota corresponding to 50% probability that the biomass would increase between 1997 and 1998 was about 5,500 t.

The risk profiles for exploitation rate were steeper and provided a relatively narrow range of quotas that satisfied the harvest strategy for forecast exploitation while the risk profiles for biomass were shallower and covered a larger range of quota options. This implies a higher degree of certainty in the range of quotas that would achieve the harvest strategy targets than those that would achieve the biomass objectives. Most of the difference in the steepness of the slopes is due to the age ranges used in the calculations. For haddock, the inverse exploitation rate did not include age 3 and for cod the fishing mortality did not include ages 5 and 6. The estimates for these younger ages are very imprecise and inflate the variance of biomass. Further, when considering biomass change, the correlation between the estimates from one year to the next can influence results.

Discussion

Prior to 1996, management consultations about the scientific evaluations for Atlantic Canadian groundfish generally focused on the acceptability to industry of quotas estimated from deterministic projections at $F_{0.1}$. With typically only one option to consider, stakeholders and managers felt disenfranchised in the decision process. The only perceived avenue for involvement was to question the validity of the scientific results. Attempts during the early 1990s to provide estimates of uncertainty for the $F_{0.1}$ projections only served to attract further criticism about the reliability of the advice. The desire by stakeholders and managers to participate in the decision process eventually led to the request that advice be provided for a range of options.

The request for advice over a range of options gave rise to considerable confusion. DFO scientists had been accustomed to providing advice for the "best" quota consistent with an $F_{0.1}$ strategy. It was natural to interpret the new request for advice as asking for provision of results on the "best" quota consistent with a range of options for F. This interpretation implied a departure from the established harvest strategy of $F_{0.1}$. Providing quota projections for a range of fishing mortality led to annual debates about the merits of alternative harvest strategies. This was an unproductive development as there was no basis to vary the harvest strategy from year to year.

In hindsight, clearly distinguishing between decisions about immediate actions and decisions about harvest strategy targets might have prevented the confusion. Although deterministic $F_{0.1}$ projections had come under criticism, the issue was not the suitability of $F_{0.1}$ as a harvest strategy target. Despite the lack of consistent and clearly defined conservation, social and economic objectives, and of associated rigorous risk analyses of alternate harvest strategies, there seemed to be general acceptance for an $F_{0.1}$ harvest strategy, which implied an exploitation rate in the range of 15% to 20% for most groundfish. The issue was how well alternate actions fared in achieving the strategy. It was necessary to develop a new view of projections as a tool for measuring performance of alternate actions relative to a strategy rather than as a tool for providing the "best" action for the strategy. This new interpretation led directly to risk analysis of immediate management decisions as measured against an established harvest strategy. The role of science was to quantify the consequences of alternate actions. It was incumbent on managers and stakeholders to consider these risks in the context of prevailing conditions. The choice of any particular quota option reflects their risk tolerance. Risk neutral strategies might be favored when the population abundance and recent exploitation rates are about average and would follow the deterministic catch projection results where the probability of achieving the harvest strategy target is 50%, assuming a symmetric distribution. Risk-averse strategies may be favored during rebuilding periods and would lean toward lower catches to lower the probability of exceeding exploitation reference levels and of not achieving desired biomass levels or increases. Risk prone strategies would tolerate higher quotas and accept probabilities lower than 50% of achieving harvest strategy objectives.

The three components of risk analyses, clear identification of available actions, establishment of suitable harvest strategies, and derivation of cumulative frequency distributions for forecast event parameters, needed to be refined and packaged for implementation. In the short period following the request to provide advice for a range of options, it had become customary to portray the consequences of alternate choices of fishing mortality rate. Explicitly recognizing that in Atlantic Canada the regulatory measure was catch quotas and not fishing mortality rate was a simple yet fundamental paradigm change which laid the foundation for subsequent development. In this paradigm, performance had to be measured against alternate management actions, therefore we evaluated the consequences for choices of catch quota. We capitalized on the existence of the established reference fishing mortality rate, $F_{0.1}$, as the harvest strategy target but recognized that this did not encompass recruitment overfishing concerns. Establishment of credible stock-recruit relationships for marine fish have been elusive and even agreement on minimum biomass levels have not been generally achieved. The generally accepted perception that the cod and haddock resources were in a depleted state offered the opportunity to introduce biomass increase as a useful harvest strategy target. Also, for cod, where the directed fishery was closed, the FRCC had already defined a threshold biomass as a reopening criterion and we advantageously included this aspect in the risk analysis. With growing acceptance of integrated statistical estimation procedures for fisheries assessment models, it was natural to extend results by applying available statistical practices to obtain cumulative frequency distributions for desired quantities. Recognizing that further refinements were required to incorporate all of the uncertainties, these methods attempted to encompass the most important sources of uncertainty.

The final required ingredient was to devise a succinct summary of complex probabilistic analyses. We used a graphic that depicted the relative performance measures against alternate actions. This risk analysis graphic proved to be easily interpreted and was quickly adopted by stakeholders and managers in their deliberations. Two or three performance

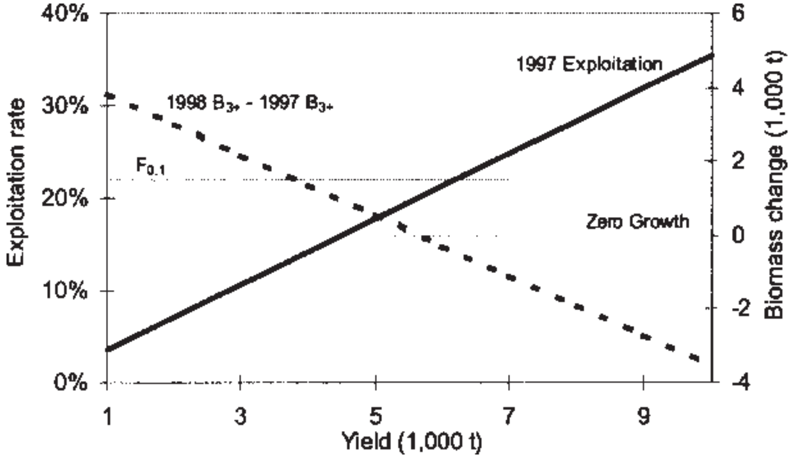


Figure 4. The graph of expected values of the forecast events for a range of quotas complemented the risk analysis and indicated that quotas in the range of 3,000 to 4,000 t should result in an exploitation rate of about half that corresponding to $F_{0.1}$ and an increase in biomass of almost 2,000 t. The $F_{0.1}$ and the zero growth harvest strategy references are indicated.

measures can be included on the same graph. Consideration of the merits of alternative actions was enhanced by a companion graphic which displayed the expected value of forecast exploitation rate and forecast change in biomass from the projections (Fig. 4). These two graphics appeared to encapsulate the information required to make informed fisheries management decisions. Such risk analysis graphics are practical for actions on a continuous scale or categorical actions that can be naturally ordered. It would also be possible to show a risk surface when decisions for two types of actions, for example quotas and gear mesh size, were contemplated. Examining the three-dimensional surface rather than marginal or conditional results could reveal interesting results. Considering decisions on more than two types of actions simultaneously greatly increases the complexity of the problem and may make complete graphical representation difficult.

The specific situation and environment may influence how readily stakeholders and managers embrace the risk analysis. There was a notable difference in the acceptance of the risk analysis results for the two cases examined here. During the FRCC public consultations, industry representatives contended that cod appeared to be more abundant than the stock assessment indicated and there were no specific discussions of the risk analysis, though it was included in the Stock Status Report (DFO 1996).

The FRCC recommended that the cod fishery be reopened in 1997 with a quota of 6,000 t and that the fishery be strictly controlled. In subsequent consultations between DFO and the fishing industry, there were serious difficulties reaching consensus and a sharing arrangement among user groups for this limited quota. At that time, the industry also indicated concern that the stock had not recovered sufficiently to support a re-opened fishery. The Minister of Fisheries and Oceans decided that the directed fishery should remain closed during 1997 and provision was made to allow a catch of 2,000 t in sentinel surveys, experimental fisheries, and bycatch. In this case the risk analysis had limited influence on the resource management decision.

There was careful consideration of the risk analysis included in the Stock Status Report for eastern Georges Bank haddock (DFO 1997), and because of the depleted state of the resource, attention focused on the consequences of quota options for biomass rebuilding. In general, industry recommended quota options in the range of 3,500 t to 4,500 t on the basis of a low probability that this would lead to a fishing mortality rate in excess of $F_{0.1}$ and a high probability that the biomass for ages 3 and older would increase. The advice submitted to the Minister by the FRCC (FRCC 1997a) highlighted four criteria to rebuild the Georges Bank groundfish resources. These were setting quotas below $F_{0.1}$, targeting an increase in biomass of 5% or more, keeping the risk of decline in biomass to 20% or less and maintaining an appropriate quota ratio of cod to haddock. For haddock specifically, the FRCC recommended a Canadian quota of 3,200 t, and with the expected U.S. catch, this quota would give a 73% probability of growth in biomass for ages 3 and older. The expected biomass increase was 7-8%. This quota level resulted in a probability greater than 95% that the inverse exploitation rate would exceed that associated with $F_{0.1}$ with the expected exploitation rate being about half that associated with $F_{0.1}$. The risk analysis, along with the evidence supporting rebuilding and the considerations regarding ratio of cod to haddock quota, were the essential factors forming the basis of the advice.

Technical enhancements in the construction of cumulative frequency distributions, such as investigations of bootstrap techniques which are under way, and advances in identification of suitable harvest strategy targets through comprehensive simulation studies will improve the risk analysis toolbox. However, this framework for translating assessment uncertainty to management risk has received acceptance and the results that can be obtained with existing technology can offer useful guidelines if the limitations are adequately qualified. In its consideration of re-opening criteria, the FRCC recognized the role of such risk analyses in the decision process (FRCC 1996). The latest report from the FRCC to the Minister titled “A Groundfish Conservation Framework for Atlantic Canada” (FRCC 1997b) endorsed the evaluation of uncertainty and recommended that “risk analyses should be carried out when feasible.”

Acknowledgments

We thank two anonymous reviewers for many useful suggestions.

References

Box, M.J. 1971. Bias in nonlinear estimation. J.R. Statist. Soc., Ser. B 33:171-201.

Cordue, P.L., and R.I.C.C. Francis. 1994. Accuracy and choice in risk estimation for fisheries assessment. Can. J. Fish. Aquat. Sci. 51:817-829.

Dept. of Fisheries and Oceans. 1996. DFO Atlantic fisheries stock status report 96/54. Available from Maritimes Regional Advisory Process, Dept. of Fisheries and Oceans, P.O. Box 1006, Stn. B105, Dartmouth, NS, Canada B2Y 4A2.

Dept. of Fisheries and Oceans. 1997. DFO science stock status report A3-08. See above for availability.

Fisheries Resource Conservation Council. 1996. Building the bridge. FRCC.96.R.2. Available from Fisheries Resource Conservation Council, P.O. Box 2001, Station D, Ottawa, ON, Canada K1P 5W3 or www.ncr.dfo.ca/frcc.

Fisheries Resource Conservation Council. 1997a. Georges Bank: 1997 conservation requirements for Georges Bank groundfish stocks. FRCC.97.R.2. See above for availability.

Fisheries Resource Conservation Council. 1997b. A groundfish conservation framework for Atlantic Canada. FRCC.97.R.3. See above for availability.

Gavaris, S. 1988. An adaptive framework for the estimation of population size. Can. Atl. Fish. Sci. Adv. Comm. (CAFSAC) Res. Doc. 88/29. 12 pp.

Gavaris, S. 1993. Analytical estimates of reliability for the projected yield from commercial fisheries. In: S.J. Smith, J.J. Hunt, and D. Rivard (eds.), Risk evaluation and biological reference points for fisheries management. Can. Spec. Publ. Fish. Aquat. Sci. 120:185-191.

Gavaris, S., and L. Van Eeckhaute. 1997. Assessment of haddock on eastern Georges Bank. DFO Canadian Stock Assessment Secretariat Res. Doc. 97/54. 72 pp.

Gulland, J.A., and L.K. Boerema. 1973. Scientific advice on catch levels. Fish. Bull. U.S. 71:325-335.

Iannelli, J.N., and J. Heifetz. 1995a. Decision analysis of alternative harvest policies for the Gulf of Alaska Pacific ocean perch fishery. Fish. Res. 24:35-63.

Iannelli, J.N., and J. Heifetz. 1995b. Erratum to Decision analysis of alternative harvest policies for the Gulf of Alaska Pacific ocean perch fishery. Fish. Res. 24:345-349.

Kennedy Jr., W.J., and J.E. Gentle. 1980. Statistical computing. Marcel Dekker, New York. 591 pp.

McAllister, M.K., E.K. Pikitch, A.E. Punt, and R. Hilborn. 1994. A Bayesian approach to stock assessment and harvest decisions using the sampling/importance resampling algorithm. Can. J. Fish. Aquat. Sci. 51:2673-2687.

Megrey, B.A., A.B. Hollowed, and R.T. Baldwin. 1994. Sensitivity of optimum harvest strategy estimates to alternative definitions of risk. Can. J. Fish. Aquat. Sci. 51:2695-2704.

Ratkowsky, D.A. 1983. Nonlinear regression modeling. Marcel Dekker, New York. 276 pp.

Restrepo,V.R., J.M. Hoenig, J.E. Powers, J.W. Baird, and S.C. Turner. 1992. A simple simulation approach to risk and cost analysis, with applications to swordfish and cod fisheries. Fish. Bull. U.S. 90:736-748.

Rosenberg, A.A., and V.R. Restrepo. 1994. Uncertainty and risk evaluation in stock assessment advice for U.S. marine fisheries. Can. J. Fish. Aquat. Sci. 51:2715-2720.

Sinclair, A., and S. Gavaris. 1996. Some examples of probabilistic catch projections using ADAPT output. DFO Atlantic Fisheries Res. Doc. 96/51. 12 pp.

Sinclair, A.F., G.A. Chouinard, and L. Currie. 1996. Assessment of the southern Gulf of St. Lawrence cod fishery, March 1996. DFO Atlantic Fisheries Res. Doc. 96/52. 75 pp.

Smith, S.J., J.J. Hunt, and D. Rivard (eds.). 1993. Risk evaluation and biological reference points for fisheries management. Can. Spec. Publ. Fish. Aquat. Sci. 120. 442 pp.

Fishery Stock Assessment Models
Alaska Sea Grant College Program • AK-SG-98-01, 1998

Consistency of Stock Assessment and Evaluation of Fisheries Management for European Fish Stocks, 1983-1995

D. Gascuel, G. Fontenelle, and C. Maguer
Laboratoire Halieutique, ENSAR, Rennes, France

A. Biseau
IFREMER DRV-RH, Lorient, France

Abstract

This paper looks at the actual catches of eight European demersal fish stocks back to 1983 in reference to the annual assessment and recommendations issued by the European Community management system. We used all estimates obtained at all steps of the process (catch data for any year *n,* VPA assessment, TACs for any (n+1) year, and respective actual catches). From this, we calculated ratios linking expected and actual values of demographic parameters of stocks. These ratios highlight discrepancies along the TAC setting process and their induced outcomes for exploitation and management. The scientific diagnoses obviously are pessimistic for saithe in West Scotland but not always pessimistic enough given the actual state of the stocks. Conversely, for whiting, the assessment and recommendations lead to an optimistic diagnosis but not optimistic enough. Additive uncertainty together with too much trust in sophistication of assessment methods, associated with bargaining along the whole process, may explain a low performance of the management system of some European fish stocks.

Introduction

During recent decades, the scientific assessments and the management framework that was set up within the European Union were unable to stop the decline of most fish stocks. Today, many stocks are overexploited, sometimes very heavily, in spite of technical measures taken for many years such as total allowable catches (TACs) and quotas (Biais 1993). By scrutinizing the history of the decision-making process, it seems possible to identify some reasons for this failure, to improve the assessment process, and to design more efficient management tools.

Northeastern Atlantic fish stocks are exploited by European Community fishing fleets. Many stocks are scientifically assessed on an annual basis by working groups of the ICES (International Council for the Exploration of the Sea). The assessment stage leads to management recommendations that are converted into TACs and fishing quotas through a complex process, which was described by Gueguen (1988). Our working assumption is that part of failure may come from this long and complicated process. Then, by using a retrospective analysis applied to some stocks exploited by Brittany fleets from 1983 to 1994, this paper evaluates the reliability of diagnoses issued from the ICES assessment groups, the discrepancies between proposed management recommendations, the adopted levels, and the recommendations that were implemented. The subsequent impact of actual catches on the stocks is also considered.

Material and Methods

Status of the Fish Stocks

High-seas French fleets based in Southern Brittany (Lorient, Concarneau, Guilvinec, Douarnenez, and Audierne) are composed of about 350 trawlers (artisanal, semi-industrial, and industrial) based in different harbors. They mainly exploit West Scotland zones (industrial trawlers from Lorient and semi-industrial from Concarneau), the Celtic Sea zone (semi-industrial from Lorient and artisanal fleet from Concarneau), and Bay of Biscaye (artisanal fleet). A fishing vessel is defined as a high-seas vessel if it makes more than one 96-hour fishing trip during a given year.

All quoted data come from Maguer (1996) and were extracted from databases held by the Administrative Center of Marine Affairs at St. Malo, France, and by IFREMER (Fishery Laboratory located at Lorient, France). As assessment refers to stock concept which associates a species with an ICES area, some species may include two distinct stocks. Thus, of the ten main species exploited by high-seas fishing fleets from southern Brittany, only eight are under scrutiny of the ICES and managed under TACs; this corresponds to nine stocks. Monkfish stocks *(Lophius piscatorius* and *L. budegassa)* are not considered in this paper because of their low level catch within West Scotland zone.

As a whole, catches of species under TACs decreased by a factor of two. During the mid-1980s, the total catch of the main eight species was about 100,000 metric tons from a total of 150,000 t caught by those fleets. Since 1989, overfishing and the decline of the industrial fishery led to a very severe depression in catch and spawning stock biomass. In 1994, they represent only 50,000 t (Fig. 1a). Similar patterns are observed for

Table 1. The eight main stocks exploited by high-seas fishing fleet based in Southern Brittany, under annual ICES working groups assessment.

Species	Name	Geographical zone	ICES zones	Stock level in 1995
Saithe	*Pollachius virens*	West Scotland	VI	Overexploited and biomass lower than MBAL
Cod	*Gadus morhua*	West Scotland	VIa	Overexploited and biomass lower than MBAL
Cod	*Gadus morhua*	Celtic Sea	VIIfgh	Overexploited and biomass close to MBAL
Hake	*Merluccius merluccius*	North Atlantic	IIIa IV VI VII VIIIab	Overexploited and biomass lower than MBAL
Whiting	*Merlangius merlangus*	Celtic Sea	VIIbk	Growth overfishing
Megrim	*Lepidorhombus* sp.	Celtic Sea and Bay of Biscaye	VIIbk + VIIIab	Growth overfishing
Norway lobster	*Nephrops norvegicus*	Celtic Sea and Bay of Biscaye	VIIIab	Growth overfishing (evaluation in 1992)
Haddock	*Melanogrammus aeglefinus*	West Scotland	VIa	Growth overfishing

The status level of exploitation is given for 1995 (from ICES 1996).

their relative spawning stock biomass which suggest a decline of some stocks (hake, saithe, and haddock) for the considered area. All these stocks are currently considered overexploited (Table 1) and three of them display biomass levels lower than the minimum considered under a precautionary approach (ICES 1996).

The MBAL (Minimum Biological Acceptable Level) is a critical empirical threshold which defines a precautionary approach. MBALs aim at avoiding overexploitation on recruitment by conserving a minimum biomass and spawning stock biomass to allow the stock to perpetuate in the long term. MBALs are estimated by referring to a fishing mortality, F_{high}, which corresponds to the mortality level inducing the decline of the biomass from one brood year to the next one for 90% of chance, if applied for the previous years (i.e., when the spawning biomass and recruitment would have been known).

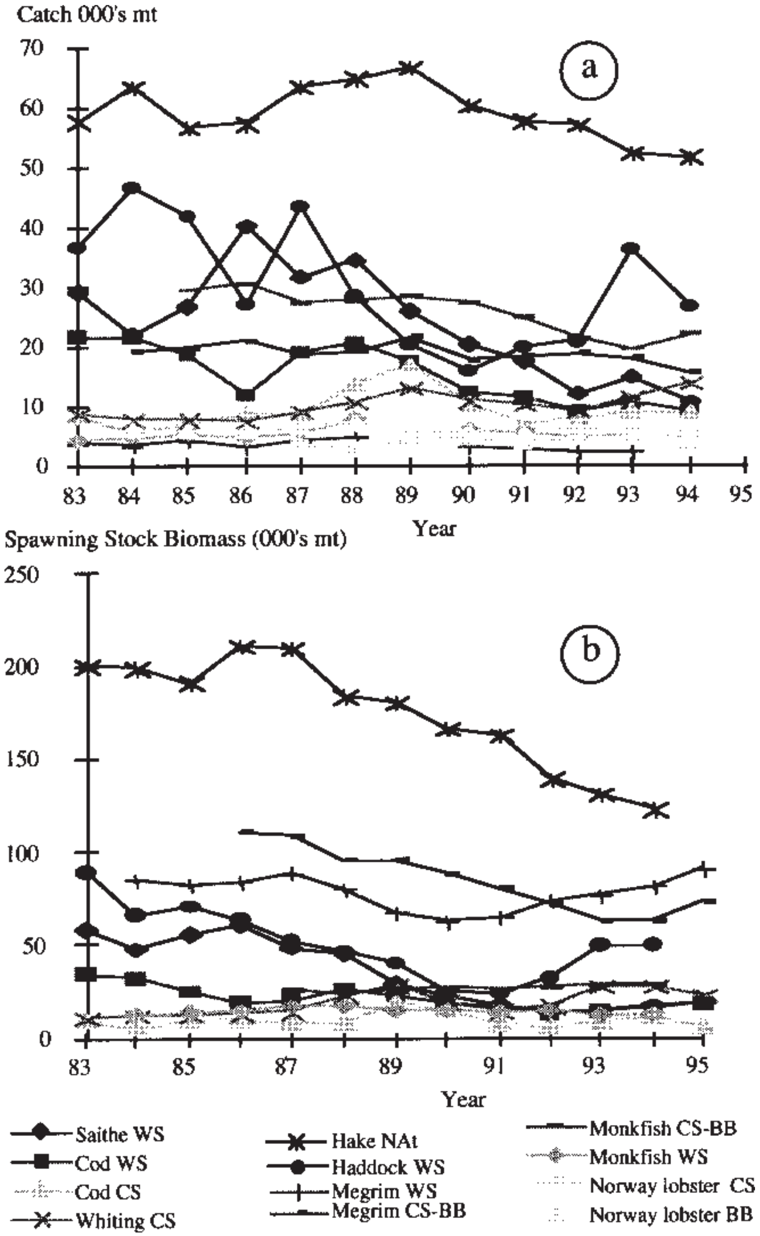


Figure 1. Evolution of (a) catches and (b) spawning stock biomass (in thousand metric tons) of the top eight species (12 ICES stocks) exploited within four ICES zones by high-seas fishing fleets based in Southern Brittany from 1983 to 1994. WS: West Scotland; CS: Celtic Sea; Nat: North Atlantic; BB: Bay of Biscaye.

A Backward Analysis of the Reliability of Diagnoses

Virtual population analysis (VPA) constitutes the basic tool used by the ICES working groups to elaborate their diagnoses. For any stock, this method leads to the estimation of fishing mortalities *F*, the biomass *B*, the spawning stock biomass *SSB*, and the recruitment *R*. For each year *n*, this evaluation applies for all years before up to (n–1). Because of the mathematical convergent properties of the VPA, the results can be considered more reliable for old data and more uncertain for recent data (Pope and Shepherd 1985). Results of the most recent period are also used to predict future catches. These results are subsequently used as input variables into simulation processes which constitute the basis of the management recommendations issued by the ACFM (Advisory Committee on Fishery Management). The crucial importance of recent year estimates to future predictions explains why the calibration process of VPA needs scrutiny. Calibration aims at refining the reliability of diagnoses after considering fishing effort data. However, this reliability can be appreciated only afterward.

Uncertainty of recent years' estimates remain a basic drawback of the VPAs. But for any past year, this is the most recent estimate that can be considered as the most probable value. Based on this background, we estimated the reliability of the scientific diagnoses with a retrospective analysis. As a convention in this paper, the parameters calculated in 1995 for the entire assessment period (back to 1983) served as the reference, and are considered as the "actual" values. Then, we compared estimates of fishing mortality (*F*), spawning stock biomass (*SSB*), and recruitment (*R*) made in each year prior to 1995 with estimates of the same parameters made using 1995 data (F^*, SSB^*, and R^*). The reliability for year *n* was estimated with three ratios of prior year estimates: the fishing mortality (average on main exploited age groups), the spawning stock biomass and the recruitment, which for year n–1 are noted as F_{n-1}, SSB_{n-1}, and R_{n-1}.

Given the estimate made in 1995 concerning any (n–1): F^*_{n-1}, SSB^*_{n-1}, R^*_{n-1}, the reliability of the diagnosis made the year *n* is appraised by calculating three ratios:

$$\frac{F_{n-1}}{F^*_{n-1}},\ \frac{SSB_{n-1}}{SSB^*_{n-1}},\ \text{and}\ \frac{R_{n-1}}{R^*_{n-1}}$$

If greater than 1, each parameter is overestimated; if less than 1, it is underestimated.

From the Diagnosis to the Management Decisions and Implementation

The annual estimations by ICES working groups are followed by simulations which give, for some fishing mortality values, the expected catch and biomass for the next years. Hence, a fishing mortality is recommended for the following year (n+1). In principle, this is the mortality that would allow the spawning stock to rebuild or at least to stop a declining trend. In order to reply to the European Commission's request, whose management system mainly lies on catch allocations, the proposed recommendation issued by the ACFM, implicitly and often explicitly, lies on defining a catch volume based on the "desired mortality." These scientific recommendations are noted: F^r_{n+1}, SSB^r_{n+1}, and R^r_{n+1}.

From the recommendations, the European Commission's Fishing Office (DG-XIV) sets up a TAC proposal. This proposal is conveyed to the Council of European Union through a complex administrative process in which various commissions are involved to consider social, economic, and political issues. Then, for the following year, the council adopts a TAC, so-called "approved TAC" noted as: TAC$_{n+1}$. We quantified the discrepancy between the scientific recommendation and the political choice by calculating the ratio:

$$\frac{TAC_{n+1}}{Y^r_{n+1}}$$

A ratio greater than 1 means the "approved TAC" exceeds the scientific recommendation

The implemented management measures can be more or less abided by, which leads to a discrepancy between actual catches (or landings) and the approved TACs. The catches made during the year n+1 are estimated by the working groups during the subsequent years. In particular, for the 1995 assessment, we have catch estimations for every year from 1983 to 1994. We calculated two ratios:

$$\frac{Y^*_{n+1}}{TAC_{n+1}}$$

which measures the lag between the management "approved TAC" by the Council and the "actual catch." A ratio greater than 1 means the TAC was exceeded, and:

$$\frac{Y^*_{n+1}}{Y^r_{n+1}}$$

which measures the lag between the initial scientific recommendation and the actual catch. A ratio greater than 1 means the recommendation was exceeded. Finally, we translated the catch into fishing mortality. This mortality also implies a given level of the surviving spawning stock biomass.

Table 2. Mean ratios (period 1983-1993) between estimated and actual values (*) of fishing mortality (*F*), spawning stock biomass (SSB) and recruitment for eight fish stocks in four ICES zones

Stock	Geographic zone	$\frac{F_{n-1}}{F^*_{n-1}}$	$\frac{SSB_{n-1}}{SSB^*_{n-1}}$	$\frac{R_{n-1}}{R^*_{n-1}}$
Saithe	West Scotland	0.88	1.18	1.45
Cod	West Scotland	0.88	1.06	1.37
Cod	Celtic Sea	0.92	0.81	0.81
Hake	North Atlantic	0.99	1.52	1.22
Whiting	Celtic Sea	1.33	0.65	0.64
Megrim	Celtic Sea and Bay of Biscaye	0.69	1.26	0.98
Norway lobster	Bay of Biscaye	1.43	0.87	0.97
Haddock	West Scotland	1.20	0.90	0.88

So, we compared the impact of actual catches on the stock with the recommended effect by calculating the two ratios:

$$\frac{F^*_{n+1}}{F^r_{n+1}}, \text{ and } \frac{SSB^*_{n+1}}{SSB^r_{n+1}}$$

These ratios account for the efficiency of the scientific work about limiting fishing pressure and preserving spawning potential. The closer to the value 1 they are, the closer to the biologist's advice the fishing practices are.

Results

Reliability of Diagnoses

The ratios between estimated fishing mortality and "actual" mortality range from 0.6 to 1.8, carrying errors through the whole process. Errors may sometimes be very important in estimates (Table 2 and Fig. 2). On average, for the considered period, the mortality is overestimated for three among nine stocks: the Norway lobster in Bay of Biscaye (but too few years are available for this species), haddock in West Scotland, and whiting in the Celtic Sea. In the three cases, they are considered as "moderately overexploited," but with a biomass greater than precautionary thresholds.

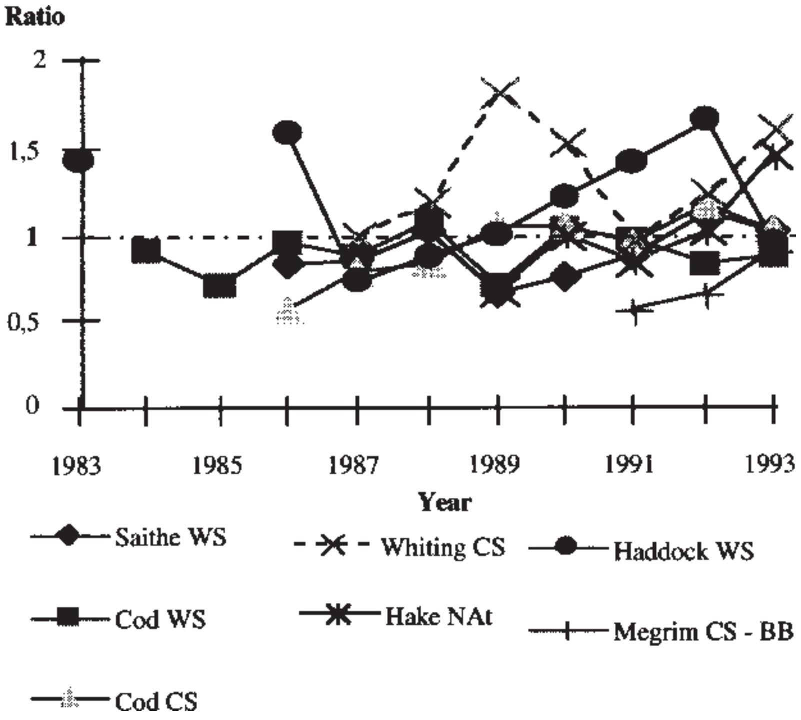


Figure 2. Evolution of ratios between estimated and actual annual fishing mortality values from 1983 to 1993 for ICES stocks exploited within four ICES zones. WS: West Scotland; CS: Celtic sea; Nat: North Atlantic; BB: Bay of Biscaye.

For most cases, however, the fishing mortality is underestimated. This is particularly true for the three stocks that are considered as heavily exploited (i.e., biomass lower than MBAL): hake in North Atlantic, saithe and cod in West Scotland. To sum up, diagnoses tend to underestimate the seriousness of the most critical situations. Moreover, this underestimation seems especially pronounced when fishing mortality is growing. For instance, this is the case for saithe in 1989 and for cod in the Celtic Sea in 1986.

Regarding the spawning stock biomass and recruitment, the ratios between estimated and actual values range from 0.4 to 2.4. The average values for the period strengthen the above results and lead to opposite situations:

- For the most exploited stocks such as saithe and cod in West Scotland, the diagnoses got more pessimistic as the stock declined. But

Table 3. Average ratios (1983-93) between recommended yield (Y^r), approved catches (TAC), and actual catches (Y^*) for nine fish stocks in four ICES zones.

Stock	Geographic zone	$\frac{TAC_{n+1}}{Y^r_{n+1}}$	$\frac{Y^*_{n+1}}{TAC_{n+1}}$	$\frac{Y^*_{n+1}}{Y^r_{n+1}}$
Saithe	West Scotland	1.37	0.95	1.24
Cod	West Scotland	1.17	0.78	0.90
Cod	Celtic sea	3.27	0.50	1.53
Hake	North Atlantic	1.25	0.94	1.31
Whiting	Celtic Sea	2.83	0.47	1.34
Megrim	Celtic Sea and Bay of Biscaye	1.03	0.98	0.80
Norway lobster	Bay Biscaye	1.04	0.77	–
Norway lobster	Celtic sea	1.19	0.18	0.25
Haddock	West Scotland	1.50	0.65	0.95

because the fishing mortality was steadily underestimated, whereas the spawning stock and recruitment were overestimated, the diagnoses remained not pessimistic enough.

- Conversely, the whiting stock in the Celtic Sea seems in a much better situation: the catches grew while mortality dropped. Here, the mortality remained overestimated whereas the spawning biomass and the recruitment were underestimated. The diagnoses became more and more optimistic, but not enough given the actual state of the stocks.

From the Recommended Catch to the Actual Catch

Through all the cases studied in this paper, the approved catches by the Council of Europe are higher than the scientific recommendations (Table 3 and Fig. 3). For the cod and the whiting in the Celtic Sea, the approved TACs reach about three times the recommendations (!). Such a result is misleading insofar as the recommendation deals with the only Celtic Sea zone (ICES zone VIIfgh), whereas the TACs concern a much wider area (ICES zone VII as a whole). Such a gap illustrates a lag between the assessment and the management stage which are set up for different spatial scales. The situation is different when they apply to the same zone, as the gaps between scientific recommendations and approved TACs often remain narrow; this is the case for the Norway lobster and the monkfish.

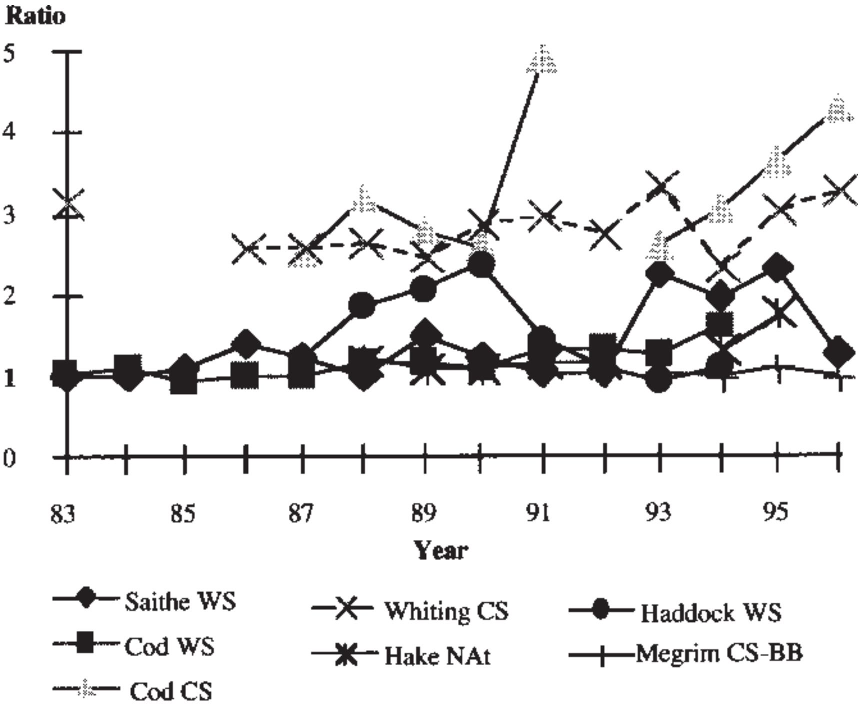


Figure 3. Evolution of ratios between annual recommended yields and approved catch for ICES stocks within four ICES zones from 1983 to 1996. WS: West Scotland; CS: Celtic sea; Nat: North Atlantic; BB: Bay of Biscaye.

However, wide gaps can be observed systematically when the working groups tend to recommend drastic decreasing catches for depleted stocks. This is the case for West Scotland stocks: haddock through years 1987-1990, saithe and cod through 1993-1995. Even when the approved catches are diminishing, they still exceed the values recommended by the ACFM. In other words, the political decisions follow rather well the biologists' recommendations but with a longer delay as the recommendations involve difficult choices. The ratios of actual catches vs. the "approved catches" are systematically lower than 1: i.e., quotas were never reached during the investigated 10-year period. Thus, for the studied stocks, the catch level was limited by fish availability, not by quotas; this is particularly obvious when the TACs exceed the scientific recommendations for cod and whiting in the Celtic Sea; the actual status of the stock (potential harvestable biomass) counterbalances the excessive optimism of the political decision. We see this pattern when TACs are close to recommendations (e.g., cod in West Scotland). In the latter case, the desirable regulation of

Table 4. Average ratios (1983-1993) between actual (*) and recommended values of fishing mortality (*F*) and spawning stock biomass (*SSB*) for fish stocks in four ICES zones.

Stock	Geographic zone	$\frac{F^*_{n+1}}{F^r_{n+1}}$	$\frac{SSB^*_{n+1}}{SSB^r_{n+1}}$
Saithe	West Scotland	1.60	0.78
Cod	West Scotland	1.44	0.71
Cod	Celtic Sea	1.37	1.21
Hake	North Atlantic	1.03	1.19
Whiting	Celtic Sea	0.83	2.39
Megrim	Celtic Sea and Bay of Biscaye	0.80	0.71
Haddock	West Scotland	1.07	1.28

catches seems to depend more on the weakness of the potential biomass than on wise political decisions with respect to scientific diagnoses. Finally, in many cases, the observed catch is greater than the recommendation. This pattern can be seen for overexploited stocks (saithe and hake) as well as for the stock of whiting in the Celtic Sea which might support more exploitation.

From the Real Catch to the Impact on the Stock

Because most of the actual catches were higher than the scientific recommendations, they were translated into a fishing mortality often higher than the initially recommended mortality (Table 4). The resulting spawning stock biomass became lower than the expected level. This situation can be observed for two of the most exploited stocks: saithe and cod in West Scotland. The too optimistic diagnoses were associated with failures in reaching the initially recommended objectives, which may have induced higher risks for the stocks to decline. Inversely, the less overexploited stock, whiting in the Celtic Sea, had lower mortality levels and higher spawning biomasses than the recommended levels. But if the actual catches were higher than the recommendations (as estimated in Table 3), they induced mortalities lower than expected because the initial diagnosis was too pessimistic. In short, the recommendation seems to have been too restrictive. In part, reducing the recommendation is warranted in this case. Indeed, catch limits of whiting in the Celtic Sea are proposed in order to protect the more fragile cod stock which is exploited together with whiting (ICES 1996).

Discussion and Conclusion

The lag between scientific evaluations and the real status of stocks are often addressed by assessment groups themselves (ICES 1996, 1997), as well as by specific studies on some stocks (e.g., Ould El Kettab 1993, for gadoids in the Celtic Sea). Our paper shows that the evaluations are delayed from the real status of a stock, and much more when changes in stock status are fast. They lead to overlay optimistic diagnoses in the most degraded situations. Moreover, the adopted TACs, often higher than biologists' recommendations, do not constrain the fishers enough and do not lead to expected lower mortalities.

For increasing overexploitation, the entire process from the assessment step to the catch step is questionable. Saithe and cod in West Scotland illustrate those situations. Neither the evaluation, nor the implemented management rules were able to prevent the declining of stocks from 1983 to 1992. Conversely, this degradation of the available resource can also explain the economic difficulties that the industrial fishing fleet from south Brittany is facing. Since the late 1980s, these difficulties have induced a strong decline in the industrial fishery. The related drop in fishing effort led to lower mortalities which we have noticed since 1990 for saithe and 1992 for cod. Recently the stock of saithe seems to be recovering with very recent increasing biomasses (ICES 1997). We may wonder if economic regulations, mainly declining rents for fishing companies, played a role over biological and statutory regulations by decreasing targeting on such depleted stocks. The successive lags that occur through the chain "assessment-decision-implementation-impact" have consequences on the stocks and decision makers must realize they have some responsibilities not to be minimized.

Exceeding the recommended catches can only partially explain the too high fishing mortality values. For instance, the actual fishing mortality for saithe was 60% higher than the recommended level whereas the excess was only 24% for catches. For cod in West Scotland, the recommended mortality was exceeded by 44% while the catches were 10% lower than recommendations. These differences may be linked to uncertainty of the diagnosis, in particular during the VPA stage. In both cases, the uncertainty led to an underestimation of past mortalities, and to an overestimation of available biomasses. Hence, even a strict implementation of scientific recommendations would not have induced the expected drop in mortalities. On the contrary, most of the biomass of cod and whiting in the Celtic Sea seem underestimated. When recommendations for catches were highly exceeded (+53% and +34% respectively), the effect on the related mortalities were slightly higher for cod (+37%) or lower than recommended for whiting (–17%).

The efficiency of scientific recommendations, to reduce fishing effort and conserve stock biomass, appears to depend as much on the reliability of stock assessments as on the way the administrative and political powers interpret them. The reliability of the scientific diagnosis relies on the VPA; however, we know that this method cannot be used to reliably estimate the actual status of stocks for recent years (Laurec and Shepherd 1983). This drawback is well acknowledged by scientists involved in assessment tasks (Ulltang 1977, Sampson 1988, ICES 1991). From this point of view, our analysis emphasizes that even the most powerful methods, as the ones used by ICES experts, cannot solve this difficulty. There has been much progress in the recent past to improve VPA methods (Sinclair et al. 1991, Angel et al. 1994). The growing sophistication of processes and software for VPAs (including statistical optimization processes of calibration, shrinkage, and others) were supposed to increase the reliability of estimations. So far, this improvement still remains to be demonstrated. Nevertheless, we wonder whether the more complex statistical processes do not lead scientists in charge of assessment to be set aside from the full methodological mastering of the evaluation. In spite of a rigorous procedure, there is a high risk that computing power may mask an oversimplification of some underlying assumptions, by minimizing the necessary critical analysis of methods and results. As an example, and without saying it could be the only or main causes of errors in evaluations, we suggest three methodological issues to address:

i. The commonly used calibration method known as XSA (eXtented Survivor Analysis, Darby and Flatman 1993, Laurec 1993) relies on the assumption that catchability for each age should be constant over the years; even if this assumption is routinely analyzed, we can sometimes wonder about its potential impacts, notably when fast changes in biomass induce changes in fishing strategies, and consequently in fishing power of vessels (Biseau 1996, Millischer and Gascuel 1997).

ii. The estimated fishing mortalities for the past year (n–1) and exploitation diagrams which are deduced for the last three years (from n–3 to n–1), are often used in the simulation process for the ongoing year n and the following year (n+1); which leads to the recommendation of catches. Of course, this process can contribute to all observed lags when the fishing effort gets higher.

iii. The impact of recruitment input for forecasting can also be important for some stocks. As an example, for whiting in the Celtic Sea, almost 40% of catches and more than 50% of spawning stock biomass that are predicted come from the recruitment input value.

Obviously, the prediction is the most uncertain operation. As a whole, the diagnoses on the past are very reliable; the deviations from reality are the most sensitive for the most recent years and often do not question the main stock trend. These deviations seem magnified through the prediction process the recommendation is based on. Of course, our paper does not aim at condemning the scientific assessment. Instead, we emphasize the major scientist's responsibility, the importance of the reliability of diagnoses, and the need for critical analyses given that sophisticated methods are not sufficiently reliable. This major issue has been raised by scientists Cadrin and Vaughan (1997) for the Atlantic menhaden and Starr et al. (1998), who ask for a more cooperative approach among players within the whole management process from revisited scientific assessments.

Acknowledgments

We would like to thank Jean-Claude Brèthes, Jon Heifetz, and two anonymous referees for their very helpful critical comments. Catherine Le Penven must also be acknowledged for her help in preparing the manuscript. Support for this study was provided by the French Ministry of Agriculture and Fisheries (DGER).

References

Angel, J.R., D.L. Burke, R.N. O'Boyle, F.G. Peacock, M. Sinclair, and K.C.Y.T. Zwanenburg. 1994. Report of the workshop on Scotia-Fundy groundfish management from 1977 to 1993. Can. Tech. Rep. Fish. Aquat. Sci. 1979. 81 pp.

Biais, G. 1993. La gestion de pêches par contingentement des captures dans le nord-ouest de l'Europe. In: D. Gascuel, J.-L. Durand, and A. Fonteneau (eds.), Les recherches françaises en évaluation quantitative et modélisation des ressources et des systèmes halieutiques. Premier Forum Halieumétrique, ORSTOM éditions, Paris, pp. 309-320.

Biseau, A. 1996. Definition of directed fishing effort, and its impact on CPUE trends: The case of the French demersal fishery of the west coast of Scotland. ICES C.M. 1996/G:17.

Cadrin, S.X., and D.S. Vaughan. 1997. Retrospective analysis of virtual population estimates for Atlantic menhaden stock assessment. Fish. Bull. U.S. 95(3):445-455.

Darby, C.C., and S. Flatman. 1993. Virtual population analysis: XSA user guide. MAFF [Ministry of Agriculture, Food and Fisheries] Directorate of Fisheries Research Ed., Lowestoft. 47 pp. Unpublished.

Gueguen, J. 1988. Les quotas, qui décide quoi? Equinoxe 19:4-9.

ICES. 1991. The working group on methods of fish stock assessments. ICES C.M. 1991/Assess: 25 Sess:S.

ICES. 1996. Report of the ICES Advisory Committee on Fishery Management 1995, 214, Part 1, 368 pp. and Part 2, 355 pp.

ICES. 1997. Report of the ICES Advisory Committee on Fishery Management 1996. 216 pp.

Laurec, A. 1993. Etalonnage de l'analyse des cohortes en halieutique. In: D. Lebreton and B. Asselain (eds.), Biométrie et environnement. Masson, Paris, p. 332.

Laurec, A., and J.G. Shepherd. 1983. On the analysis of catch and effort data. J. Cons. Cons. Int. Explor. Mer 41:81-84.

Maguer, C. 1996. De l'évaluation des stocks à l'activité des flottilles: Étude des flottilles hauturières de Bretagne Sud. Mémoire DAA Halieutique de l'ENSA-Rennes. 48 pp.

Millischer, L., and D. Gascuel. 1997. Estimation des puissances de pêche des navires: Étude des flottilles industrielles de Bretagne sud. 3e Forum Halieumétrique, Montpellier, 1-3 Juil 1997. 12 pp.

Ould El Kettab, M. 1993. La pêcherie de Gadidés de mer Celtique: Description, analyse de l'exploitation et évaluation des stocks. Tentative de gestion par un modèle économique. Ph.D. thesis, Université de Bretagne Occidentale, Brest. 207 pp.

Pope, J.G., and J.G. Shepherd. 1985. A comparison of the performance of various methods for tunings VPAs using effort data. J. Cons. Cons. Int. Explor. Mer 42:129-151.

Sampson, D.B. 1988. The stability of virtual population analysis cohort size estimates. J. Cons. Cons. Int. Explor. Mer 44:135-142.

Sinclair, A., D. Gascon, R. O'Boyle, D. Rivard, and S. Gavaris. 1991. Consistency of some northwest Atlantic groundfish stock assessments. NAFO Sci. Counc. Stud. 16:59-77.

Starr, P., J.H. Annala, and R. Hilborn. 1998. Contested stock assessment: Two case studies. Can. J. Fish. Aquat. Sci. 55:529-537.

Ulltang, O. 1977. Sources of errors in and limitations of virtual population analysis (cohort analysis). J. Cons. Cons. Int. Explor. Mer 37:249-260.

A General Approach for Making Short-Term Stochastic Projections from an Age-Structured Fisheries Assessment Model

Jon Brodziak
National Marine Fisheries Service, Northwest Fisheries Science Center, Newport, Oregon

Paul Rago
National Marine Fisheries Service, Northeast Fisheries Science Center, Woods Hole, Massachusetts

Ramon Conser
National Marine Fisheries Service, Northwest Fisheries Science Center, Newport, Oregon

Abstract

Catch-age analysis is a powerful tool for assessing the status of fisheries resources and catch-age analyses are routinely conducted for many commercially exploited fish stocks around the world. In this paper, we illustrate a general approach for making short-term stochastic projections from the ADAPT age-structured assessment model. Our approach uses the standard statistical techniques of bootstrapping and Monte Carlo simulation to project performance measures such as landings, discards, spawning biomass, and recruitment under alternative management policies. The key idea is to propagate variability in estimates of initial stock size forward in stochastic projections of future possibilities. We use bootstrap replicates of current population size from an age-structured ADAPT model combined with a stochastic stock-recruitment relationship to simulate population trajectories through the projection horizon. The approach is illustrated for a commercially important New England groundfish, Georges Bank yellowtail flounder, *Pleuronectes ferrugineus*. Although hypothetical, this illustration provides some general insight for the rebuilding of the Georges Bank yellowtail flounder stock.

Introduction

Catch-age analysis is a powerful tool for assessing the status of fisheries resources and catch-age analyses are routinely conducted for many commercially exploited fish stocks around the world. Diverse methods of catch-age analysis exist (Megrey 1989) and these include, but are not limited to, tuned virtual population analysis (Laurec and Shepherd 1983, Parrack 1986, Gavaris 1988), statistical catch-at-age analysis (Fournier and Archibald 1982, Deriso et al. 1985, Hilborn and Walters 1992), stock synthesis (Methot 1989, 1990), as well as state-space (Mendelssohn 1988, Schnute 1994) and Bayesian approaches (Hilborn et al. 1994, McAllister and Ianelli 1997). Despite underlying differences in model assumptions and statistical framework, these diverse approaches to age-structured assessment generally provide estimates of fishing mortality and stock size at age through the assessment time horizon. To provide quantitative advice to fisheries managers, these estimates of fishing mortality and stock size are often compared to biological reference points or to threshold stock sizes to evaluate whether current exploitation rates exceed target or limit reference points and whether stock size has been maintained above threshold levels (Caddy and Mahon 1995). However, while such analyses provide historic and current estimates of stock size and fishing mortality to evaluate previous harvest policies, they do not address the management question of projecting what would happen in the future under alternative harvest policies.

In this paper, we illustrate a general approach for making short-term stochastic projections from the ADAPT age-structured assessment model. Our approach extends the work of Mohn (1993), who used bootstrapping and simulation to project uncertainties in catch projections from ADAPT, and Smith and Gavaris (1993), who compared several methods of evaluating the accuracy of projected catch estimates from ADAPT, in several ways. First, in our approach, simulation is systematically applied to the entire bootstrap distribution of initial population sizes to propagate uncertainty. Second, our approach permits several types of future harvest policies to be evaluated including quota policies, target fishing mortality rate policies, or any mixture of these two. Third, we provide a standard set of performance measures, including spawning biomass, recruitment, landings, discards, and the probability of achieving spawning biomass thresholds, to evaluate the potential benefits of alternative harvest policies. Since 1994, this approach has been used to project the consequences of alternative management measures for several age-structured assessments in the northwest Atlantic including Atlantic mackerel, Georges Bank cod, haddock, and yellowtail flounder.

Our approach uses the standard statistical techniques of bootstrapping and Monte Carlo simulation to project performance measures such as landings, discards, spawning biomass, and recruitment under alternative management policies. The key idea is to propagate variability in estimates of initial stock size forward in stochastic projections of future possibilities. We use bootstrap replicates of current population size from an age-structured ADAPT model combined with a stochastic stock-recruitment relationship to simulate population trajectories through the projection horizon. As a consequence, uncertainties in both initial population abundance and future recruitment are directly incorporated into management advice. The implications of management decisions can be quantified and compared using empirically derived sampling distributions of catch, landings, discards, spawning biomass, recruitment, and, in the case of management under fixed catch quotas, fishing mortality. Estimates of the probability of exceeding biological reference points or achieving management targets are also quantified. The approach is illustrated for a commercially important New England groundfish, Georges Bank yellowtail flounder, *Pleuronectes ferrugineus*. This example is based on analyses of Amendment 7 to the Northeast Multispecies Fishery Management Plan (New England Fishery Management Council 1996) which was implemented to rebuild depleted New England groundfish stocks.

Materials and Methods

Initial Population Abundance

Initial population abundance is the first of two key sources of randomness in this approach to making stochastic projections, where $\boldsymbol{N}(1)$ is the vector of stock size at age in the current year which is also the first year of the projection horizon. Regardless of the age-structured assessment model used to estimate $\boldsymbol{N}(1)$, it is common practice to determine the sampling distribution of model estimates of current population size and other key outputs such as fishing mortality and spawning biomass. Uncertainties in key model outputs are often described with confidence intervals or probability profiles which express the chance that an output falls below a biological reference point (Rosenberg and Restrepo 1994).

The ADAPT method of sequential population analysis provides the initial abundance estimates for the projection approach described in this paper (Parrack 1986, Gavaris 1988, Conser and Powers 1990, Mohn and Cook 1993). This age-structured assessment approach combines deterministic virtual population analysis with a nonlinear least squares (NLS) objective function to estimate model parameters (θ) such as stock size at age through time. As generally implemented, the ADAPT method is a measurement error model in which observed indices of relative abundance (I_{iy}) are modeled as random deviations from the true values of the abundance indices. In particular, it is often assumed that the log-transformed measurement errors are independent and identically distributed random variables with zero mean and constant variance. The NLS objective function [$SS(\theta)$] typically consists of the weighted sum of squared differences between model predictions of indices of stock size at age and observed survey indices of stock size at age in natural logarithmic scale. That is, the objective function $SS(\theta)$ commonly has the form

$$SS(\theta) = \sum_{i=1}^{K} \sum_{y=1}^{Y_i} \lambda_{iy} \left(\log_e I_{iy} - \log_e I_{iy}^{*} \right)^2 \tag{1}$$

where I_{iy} is the observed value of the ith survey index in year y, I_{iy}^{*} is the predicted value of the ith survey index in year y, K is the number of age-specific survey indices, Y_i is the number of years of abundance data for the ith survey index, and λ_{iy} is the relative weighting factor for the ith survey and yth year term. In the absence of prior information on the accuracy and precision of survey indices, the λ_{iy} are often set to 1. Given initial estimates of model parameters and fixed weighting factors, an iterative minimization algorithm, such as the Marquardt algorithm, is used to compute the NLS estimate of model parameters, θ_{NLS}; this NLS estimate can be interpreted as a maximum likelihood estimate (conditioned on the relative weighting factors) if the measurement errors are independent and identically distributed lognormal random variables with zero mean and constant variance.

We use the nonparametric bootstrap (Efron 1982) to empirically determine the sampling distribution for the estimator of $\boldsymbol{N}(1)$. To accomplish this, the set of residuals $\{\varepsilon_{iy}\}$ from the observed and predicted values of survey indices of the best ADAPT model formulation are generated for resampling

$$\varepsilon_{iy} = \log_e I_{iy} - \log_e I_{iy}^{*} \tag{2}$$

The set of residuals is randomly resampled with replacement a large number of times (B) to create a family of new input data sets of relative abundance indices. In particular, the bth bootstrap data set $\{I_{iy}^{(b)}\}$ is constructed from the bth set of resampled residuals $\{\varepsilon_{iy}^{(b)}\}$ as

$$\log_e I_{iy}^{(b)} = \log_e I_{iy}^{*} + \varepsilon_{iy}^{(b)} \tag{3}$$

This resampling is conditioned on the correctness of the fitted model (Punt and Butterworth 1993) where it is assumed that residuals are independent and identically distributed. Parameters of the age-structured assessment model are re-estimated for each of the new data sets to generate a set of bootstrap replicates of initial population abundance, $\{\boldsymbol{N}^{(1)}(1), \boldsymbol{N}^{(2)}(1), \ldots, \boldsymbol{N}^{(B)}(1)\}$. This set gives the empirical (bootstrap) distribution function of the estimator of $\boldsymbol{N}(1)$ conditioned on the model. Bootstrap confidence intervals are derived from percentiles of the bootstrap distribution of $\boldsymbol{N}(1)$ (see Efron and Tibshirani 1993, Chapter 13) and point estimates of parameter bias can also be computed (Efron 1982, Smith and Gavaris 1993).

Age-Structured Population Model

An age-structured population model is the basis for making short-term projections. The model describes the dynamics of a closed, iteroparous fish population whose abundance changes due to recruitment, natural mortality, and fishing mortality. Recruitment to the population is measured at the beginning of each year. Population abundance changes throughout the year due to the concurrent forces of natural and fishing mortality so that the fishery is classified as a type-2 fishery (Ricker 1975).

Population Abundance, Survival, and Spawning Biomass

Population abundance is measured by the number of survivors within each recruited age class at the beginning of the year and is denoted by the vector $\boldsymbol{N}(t)$ with elements $N_a(t)$ where *"a"* indexes age class and *"t"* indexes year in the projection horizon. The youngest age class comprises the recruits and the age of recruitment (R) is typically age-1 or age-2. The oldest age class is a plus-group comprised of all fish that are at least as old as a chosen cutoff age (A).

Population survival at age from year to year is calculated in the standard manner using instantaneous mortality rates. To describe annual survival through mortality, $M_a(t)$ is the instantaneous natural mortality rate and $F_a(t)$ is the instantaneous fishing mortality rate on age-a fish. Population size at age changes according to

$$\begin{aligned} &N_a(t) = N_{a-1}(t-1) \times e^{-M_{a-1}(t-1)-F_{a-1}(t-1)} \text{ for } a = R+1 \text{ to } A-1 \\ &\text{and } N_A(t) = N_A(t-1) \times e^{-M_A(t-1)-F_A(t-1)} + N_{A-1}(t-1) \times e^{-M_{A-1}(t-1)-F_{A-1}(t-1)} \end{aligned} \tag{4}$$

Typically, age-specific values of natural mortality are not estimable due to a lack of data and natural mortality is parsimoniously modeled as a constant value across age classes. We adopt this convention and assume that natural mortality (M) is constant in what follows. Recruitment is modeled as a stochastic process and is not determined by equation (4).

Annual spawning biomass is determined from the population size vector $\boldsymbol{N}(t)$ and mortality rates as well as additional information concerning fish maturity and size at age. Population abundance at the start of the spawning season depends on the level of fishing and natural mortality prior to spawning; the number of age-a fish alive at the start of spawning ($N_{S,a}$) is

$$N_{S,a}(t) = N_a(t)e^{-T(t)[M+F_a(t)]} \tag{5}$$

where $T(t)$ is the fraction of total annual mortality that occurs from January 1 to the start of the spawning season. This fraction may need to be modeled as time-varying if seasonal patterns in fishing effort are altered by management measures. Because this approach focuses on short-term projections, growth and maturation rates are assumed to be independent of stock density during the projection horizon. Spawning biomass [$SB(t)$] is the sum over all age classes of the weight of mature female and male fish at the start of the spawning season

$$SB(t) = \sum_{a=R}^{A} W_{S,a} P_a N_a(t) e^{-T(t)[M+F_a(t)]} \tag{6}$$

where P_a is the average fraction of age-a fish that are mature and $W_{S,a}$ is the average spawning weight of an age-a fish.

Catch, Landings, and Discards

In this model, the entire population is available to the fishery. Annual catch at age [$C_a(t)$] is given by the standard catch equation (Baranov 1918)

$$C_a(t) = \frac{F_a(t)}{M + F_a(t)} \left(1 - e^{-[M+F_a(t)]}\right) N_a(t) \tag{7}$$

To account for age-specific discarding of fish, let $D_a(t)$ be the annual fraction of age-a fish that are discarded and die and $W_{L,a}$ and $W_{D,a}$ are the average weights of age-a fish that are landed and discarded, respectively. Here the age-specific discarding pattern may vary if fishery regulations are expected to change during the projection horizon. Total landed weight [$L(t)$] is

$$L(t) = \sum_{a=R}^{A} C_a(t) \left[1 - D_a(t)\right] W_{L,a} \tag{8}$$

and the total discarded weight [$D(t)$] is

$$D(t) = \sum_{a=R}^{A} C_a(t) D_a(t) W_{D,a} \tag{9}$$

Population Harvest

There are three options for determining the level of population harvest in each year of the projection horizon. The first option is a time series of annual fishing mortality rates (F-based management) and the second is a time series of annual landings quotas (quota-based management). Any mixture of the F-based and quota-based options can be combined so that F-based management is applied in some years and quota-based management in the other years. This mixed option can be useful for projecting stock status in periods between assessments. For example, if a full stock assessment is only conducted every 2 years while catch biomass data are updated annually, then population status can be monitored in years between assessments using the quota-based option to project the effects of an additional year of catch data.

Under F-based management, catch at age is determined by setting $F_a(t)$ in equation (7). Separability is assumed and fishing mortality for age-a fish is the product of fully recruited fishing mortality [$F(t)$] and partial recruitment for age-a fish [$S_a(t)$].

$$F_a(t) = F(t)\,S_a(t) \tag{10}$$

Typically, values of partial recruitment are based on recent estimates from the assessment and are modeled as constant unless fishery regulations are expected to change during the projection horizon.

Under quota-based management, the landings quota [$Q(t)$] corresponds to a fully recruited fishing mortality rate that depends on population size, partial recruitment pattern, and discard pattern, if applicable. This quota depends on the level of fully recruited fishing mortality, $Q = L(F)$, where L is given by equation (8), and the solution to $Q - L(F) = 0$ is determined numerically. After solving for the F that produces the landings quota, total catch at age and discard biomass are determined from equations (7), (9), and (10).

Stock-Recruitment Relationship

In many fisheries, stock-recruitment data are highly variable due to intrinsic variability in factors governing survival and measurement error in estimates of recruitment and spawning biomass. In our approach, the stock-recruitment relationship is the second key source of randomness for population projections because it ultimately defines the sustainable level of harvest and its variability over time. This follows from the model assumption that growth, maturation, and natural mortality are density-independent and adequately represented by expected (e.g., constant) values through the projection horizon.

In the application of this approach, several stochastic recruitment models have been used and two commonly used models are briefly described below. Because no single model will be suitable in all applications, we recommend that alternative stock-recruitment relationships be explored by the assessment practitioner. The first model is the empirical recruits per spawning biomass model. This is a non-parametric, density-dependent approach. In this model, random values of recruits per spawning biomass are generated from a uniform (0,1) variate by applying the probability integral transform to the empirical cumulative distribution function of recruits per spawning biomass from an appropriate time period. Here the choice of time period is important because this model requires stationarity in the distribution of recruits per spawning biomass. Recruitment is modeled as the product of spawning biomass and the random level of recruits per spawning biomass. That is, $N_R(t) = SB(t - R)\xi$ where ξ is a randomly chosen value of recruits per spawning biomass taken from the distribution of estimates of recruit per spawning biomass during the chosen period. This model exhibits no compensatory population response because average recruitment does not eventually decline as spawning biomass increases. The second model is the stock-recruitment curve with multiplicative lognormal error term. In this model, recruitment is determined as the output of a stock-recruitment curve multiplied by a random realization from a lognormal distribution. That is, $N_R(t) = \mathrm{g}[SB(t-R)]\exp(\xi)$ where g[] is a parametric stock-recruitment curve (typically a Beverton-Holt, Ricker, or Shepherd curve) and ξ is a normal random variable with zero mean and constant variance. This parametric model is stochastic, density-dependent, and exhibits compensatory population response due to the choice of stock-recruitment curve.

The age of recruitment determines whether additional information is needed to begin the projections with a density-dependent recruitment model. In particular, if $R \geq 2$, then assessment estimates of spawning biomass in the R–1 years prior to the first year of the projection horizon are needed for input to the recruitment model. These lagged estimates of spawning biomass must be available for each bootstrap replicate of initial population size, $\boldsymbol{N}^{(b)}(1)$ to use this approach to making stochastic projections.

Application to Georges Bank Yellowtail Flounder

Georges Bank yellowtail flounder *(Pleuronectes ferrugineus)* is a highly productive flatfish stock with a long-term potential yield of 16,000 t at a spawning biomass level of about 65,000 t (NEFSC 1995). This stock is currently at a low level of abundance due to overfishing (Fogarty and Murawski 1998). Based on the most recent assessment (NEFSC 1997), fishing mortality rates on Georges Bank yellowtail flounder exceeded the overfishing level ($F_{20\%} = 0.64$) during 1980-1994 (Fig. 1). Reductions in fishing effort achieved under Amendment 7 have reduced fishing mortality in recent years and spawning biomass levels have increased. A threshold level of spawning biomass of 10,000 t has been established as an interim rebuilding target for Georges Bank yellowtail flounder (New England Fishery Management Council 1996).

Georges Bank yellowtail flounder has been assessed with the ADAPT age-structured assessment model since 1991 (Conser et al. 1991). In the 1997 assessment (Cadrin et al. 1997), relative abundance indices taken from four research surveys on Georges Bank were used to calibrate a virtual population analysis for estimating population size at age for 6 age groups ($A = 6$) consisting of ages 1 to 5 and age 6+. The four surveys were the Northeast Fisheries Science Center (NEFSC) spring (age 1 to age 5+, 1968-1996) and autumn (age 1 to age 5+, 1963-1996) bottom trawl surveys, the NEFSC scallop survey (age 1 to age 4+, 1982-1996), and the Canadian spring bottom trawl survey (age 2 to age 5+, 1987-1996). Given the eighteen time series of age-specific indices of relative abundance, the objective function for the ADAPT assessment model was the sum of the

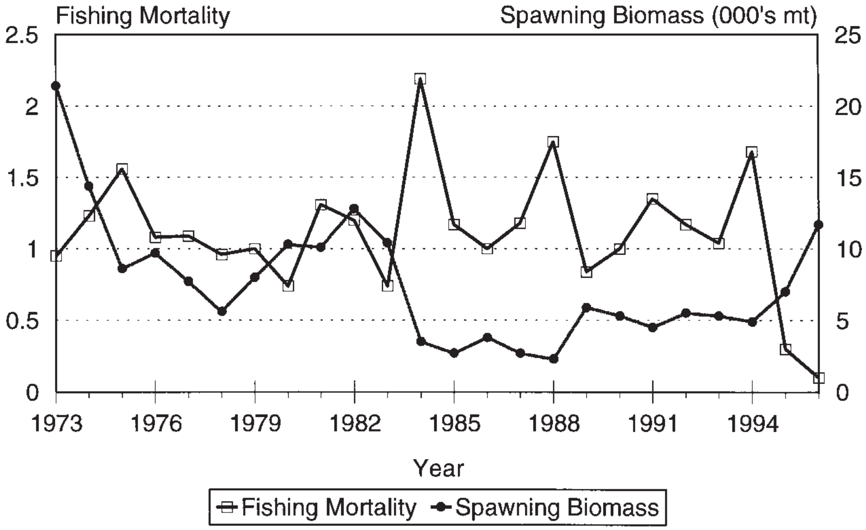


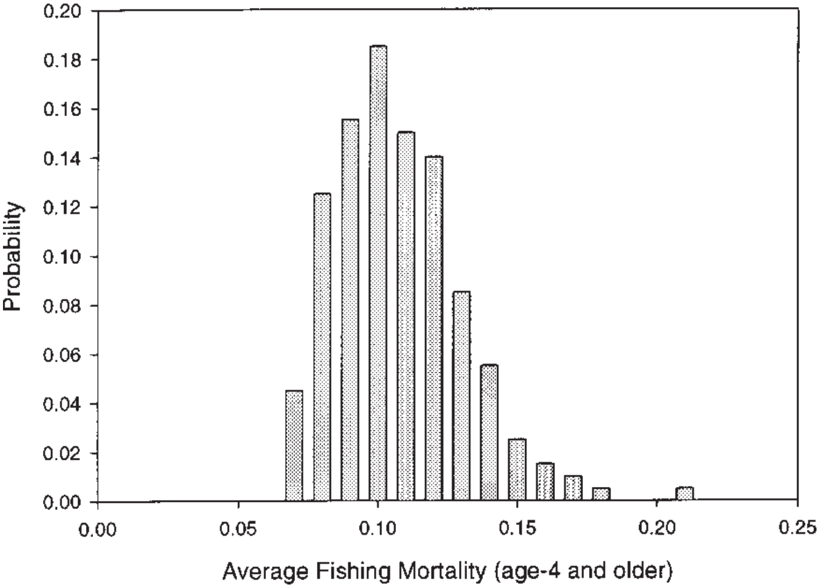
Figure 1. Fishing mortality (open square) and spawning biomass (solid circle) of Georges Bank yellowtail flounder, 1973-1996. Data from NEFSC (1997).

squared differences between predicted and observed survey indices as in equation (1) with $\lambda_{iy} = 1$ (NEFSC 1997). The non-parametric bootstrap was applied to resample the residuals of the fitted indices with replacement to generate a set of 200 bootstrap replicates of population abundance in 1996. This set determined the bootstrap distribution of population abundance and was used to characterize the sampling distribution of fishing mortality and spawning biomass in 1996. Bootstrap results gave 80% confidence intervals (CI) for $F(1996)$ of 0.08 to 0.14 (Fig. 2A) and for $SB(1996)$ of 9,800 to 14,600 t (Fig. 2B). In addition, bootstrap results indicated there was zero probability that $F(1996)$ exceeded the overfishing level of 0.64. Further the probability that $SB(1996)$ was below the interim rebuilding threshold of 10,000 t was about 0.12 (see Fig. 4) while the probability that $SB(1996)$ was below the long-term spawning biomass target of 65,000 t was 1.

A stochastic stock-recruitment relationship has been derived for Georges Bank yellowtail flounder based on a Beverton-Holt curve with a multiplicative lognormal error term (Beverton and Holt 1957). In this parametric model, survival to recruitment age is density independent and recruitment is subject to stochastic variation through

$$N_R(t) = \frac{a \times SB(t-R)}{b + SB(t-R)} \times e^w \qquad (11)$$

(A) Bootstrap distribution of fishing mortality in 1996



(B) Bootstrap distribution of spawning biomass in 1996

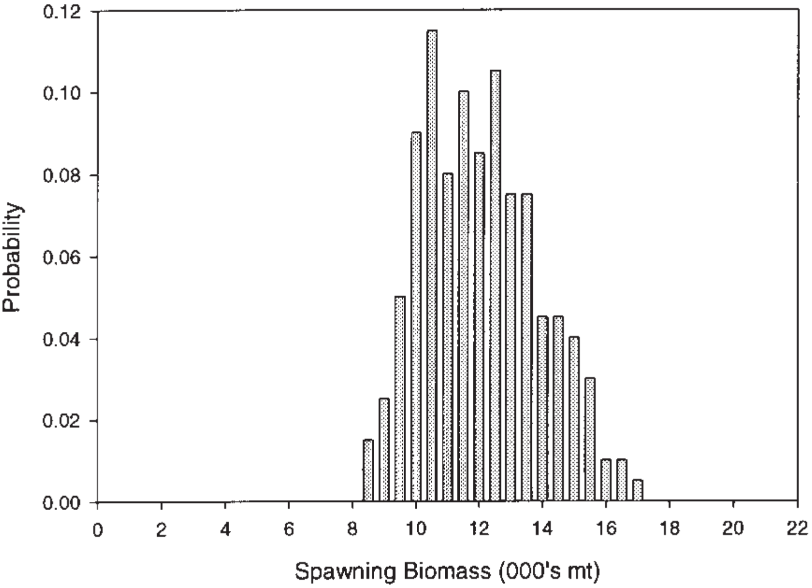


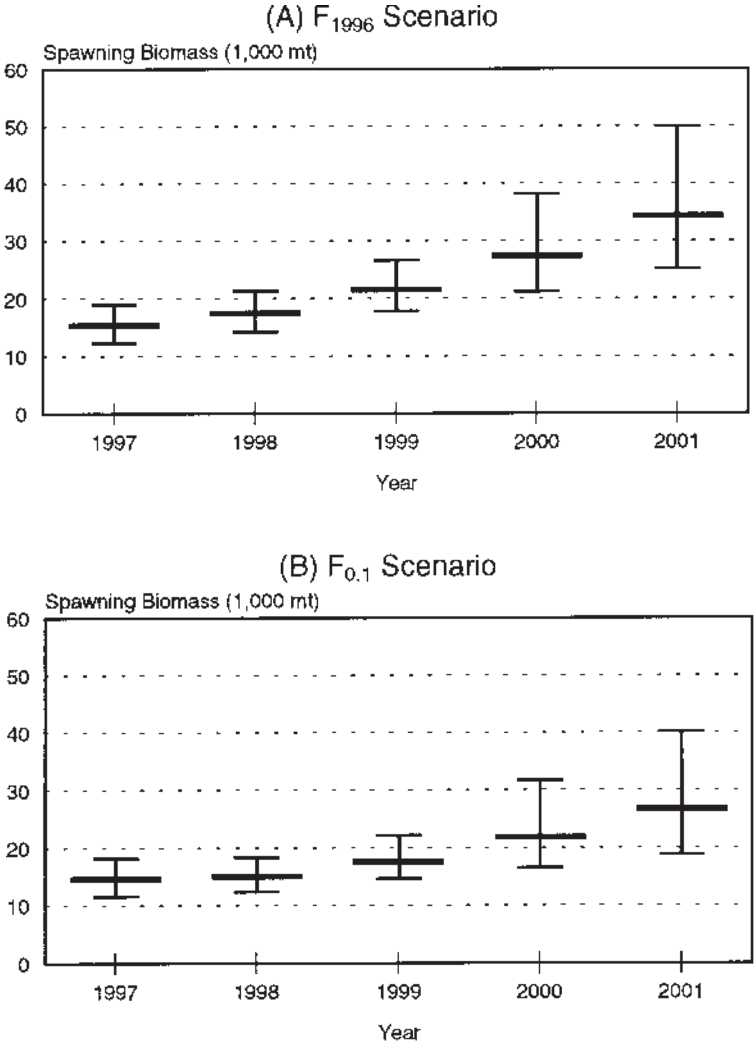
Figure 2. Bootstrap distributions of estimated fishing mortality and spawning biomass of Georges Bank yellowtail flounder in 1996.

where a and b are model parameters and where w is a normal random variable with zero mean and variance σ_w^2. Based on the most recent assessment of Georges Bank yellowtail flounder, Overholtz et al. (1997) estimated parameters of equation (11) with nonlinear regression and tested model residuals for normality and autocorrelation. Results indicated that an assumption of normally distributed errors could not be rejected and that autocorrelation was not significantly different from 0 for lags of 1 to 6 years. Point estimates of the parameters a and b were 50,090 and 10,737 while the residual variance σ_w^2 was 0.42.

Projections were conducted for 1997-2001 under three F-based management scenarios to illustrate the general approach to making stochastic projections. The three F-based scenarios were: constant F equal to the point estimate of fully recruited fishing mortality in 1996 (F_{1996}), constant F equal to $F_{0.1}$ (Gulland and Borema 1973), and constant F equal to F_{MAX} (Beverton and Holt 1957). These three alternative harvest policies represent a status quo scenario (F_{1996}), a target reference point scenario ($F_{0.1}$), and a limit reference point scenario (F_{MAX}). Values of F_{1996}, $F_{0.1}$, and F_{MAX} for yellowtail flounder were 0.10, 0.24, and 0.61, respectively. The three projections for 1997-2001 used biological and fishery parameters reported in Cadrin et al. (1997), and Overholtz et al. (1997). A total of 100 simulations were conducted for each of the 200 bootstrap replicates of initial population abundance to characterize the empirical distribution of spawning biomass, the probability of achieving the threshold level of spawning biomass, and the empirical distributions of recruitment, landings, and discards.

Projection results indicated that under the status quo scenario, median spawning biomass of Georges Bank yellowtail flounder in the year 2001 and its 80% CI would be 34,300 t with an 80% CI of 25,200 to 49,900 t (Fig. 3A). Under the target reference point scenario, median spawning biomass in 2001 would be 22% lower than under the status quo (Fig. 3B). In contrast, under the limit reference point scenario, median spawning biomass in 2001 would be 55% lower than under the status quo (Fig. 3C). Overall, levels of spawning biomass would likely be higher in 2001 under the status quo and target reference point scenarios in comparison to the limit reference point scenario, although the uncertainty in spawning biomass, as measured by its 80% CI, would also be greater. Further, the spawning biomass threshold of 10,000 t would be reached in 2001 under both the status quo and the target reference point scenarios and would have a good chance of being achieved under the limit reference point scenario (Fig. 4). In contrast, it was unlikely that the long-term spawning biomass target would be achieved in 2001 under any scenario. In particular, the probabilities of achieving the long-term spawning biomass target under the status quo, the target reference point, and the limit reference point scenarios were 0.02, 0.004, and 0, respectively.

In terms of recruitment, landings and discards, the status quo scenario would lead to median recruitment of about 38 million age-1 fish with an 80% CI of 16 to 88 million age fish (Fig. 5A). Median recruitment in 2001


(A) F_{1996} Scenario

(B) $F_{0.1}$ Scenario

Figure 3. Projections of median spawning biomass of Georges Bank yellowtail flounder and its 80% confidence interval during 1997-2001 under (A) F_{1996}, (B) $F_{0.1}$, and (C) F_{MAX} management scenarios.

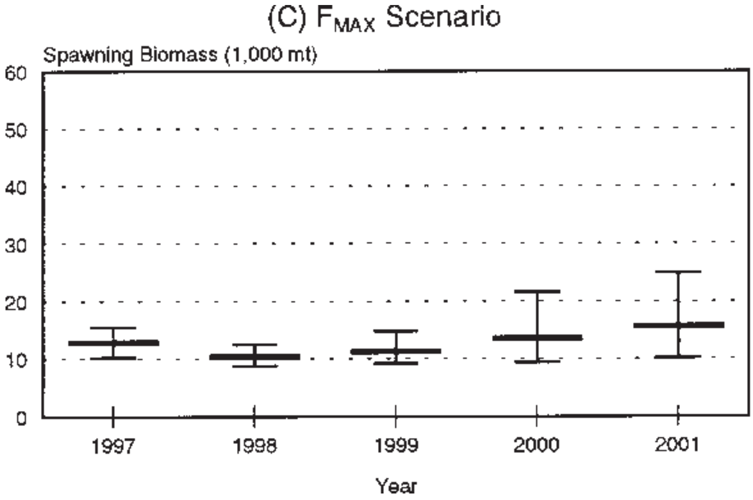


Figure 3. (Continued.)

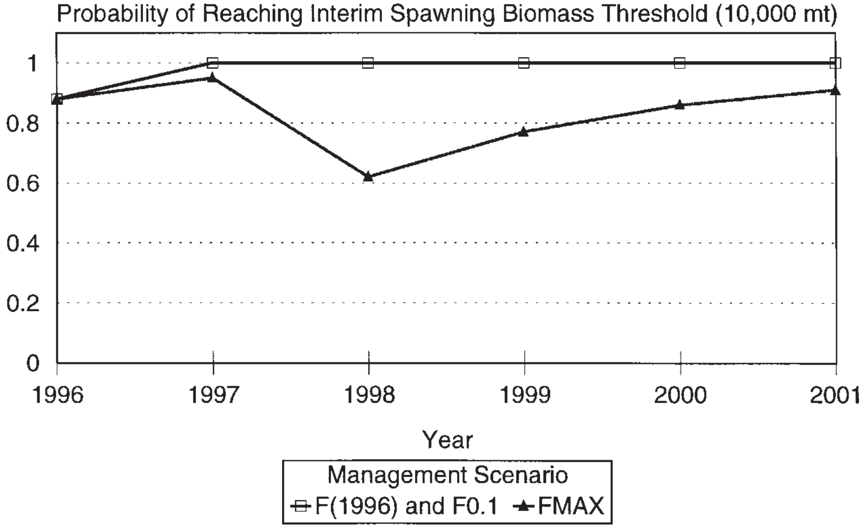
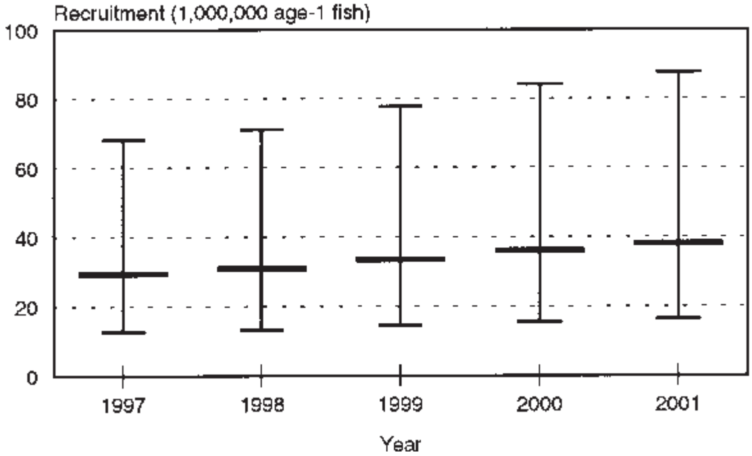


Figure 4. Probabilities of reaching the interim spawning biomass threshold for Georges Bank yellowtail flounder during 1996-2001 under F_{1996}, $F_{0.1}$, and F_{MAX} management scenarios.

(A) F_{1996} Scenario



(B) $F_{0.1}$ Scenario

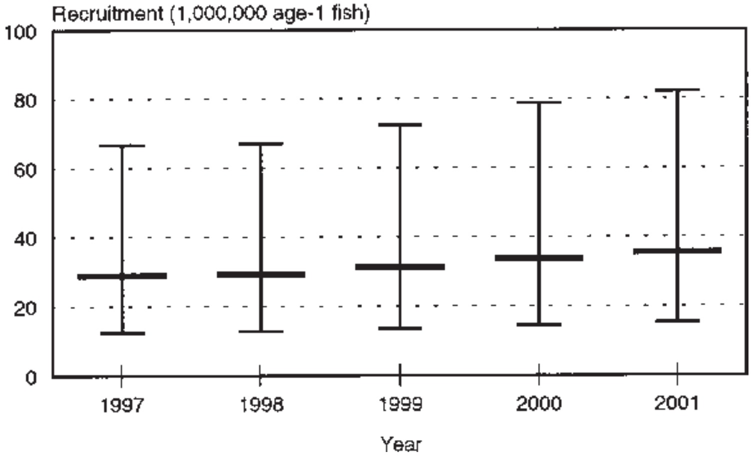


Figure 5. Projections of median recruitment of Georges Bank yellowtail flounder and its 80% confidence interval during 1997-2001 under (A) F_{1996}, (B) $F_{0.1}$, and (C) F_{MAX} management scenarios.

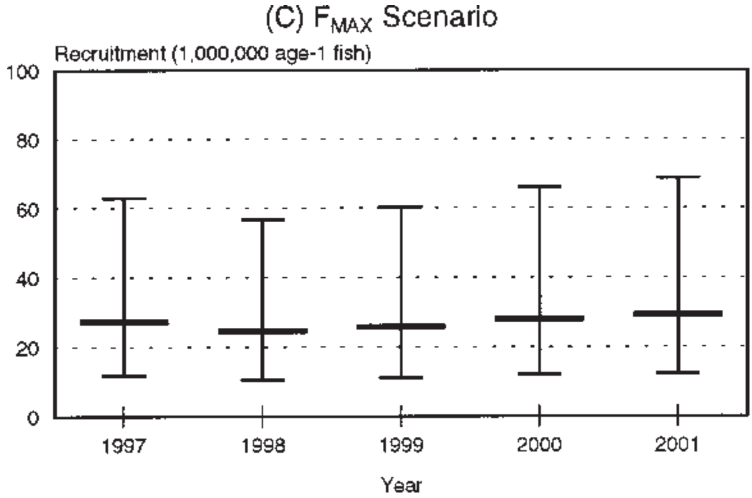


Figure 5. (Continued.)

would be about 6% lower under the target reference point (Fig. 5B) and 23% lower under the limit reference point (Fig. 5C). The status quo scenario would lead to median landings of 2,500 t in the year 2001 with an 80% CI of 1,900 to 3,500 t for yellowtail flounder (Fig. 6A). In comparison, median landings in 2001 under the target reference point would be 72% higher (Fig. 6B). Under the limit reference point scenario, median landings in 2001 would be 128% higher (Fig. 6C). Median discards of Georges Bank yellowtail flounder in 2001 would be 50 t under the status quo scenario with 80% CI of 30 to 80 t (Fig. 7A). In comparison, median discards in 2001 would be 114% and 314% higher under the target (Fig. 7B) and limit reference point scenarios (Fig. 7C), respectively. Overall, levels of landings and discards would likely be greater under the limit reference point scenario, but the uncertainty in these outputs, as measured by the 80% CI, would also be greater than under the status quo and target reference point scenarios. In contrast, recruitment would likely be lower under the limit reference point scenario but would have less uncertainty than under the status quo or target reference point scenarios.

Discussion

Although hypothetical, this illustration of the projection approach provides some general insight for the rebuilding of the Georges Bank yellowtail flounder stock. First, application of the limit reference point F_{MAX} as a harvest rate target is not precautionary for stock rebuilding in the short term and would not likely improve the reproductive potential of this depressed stock. Second,

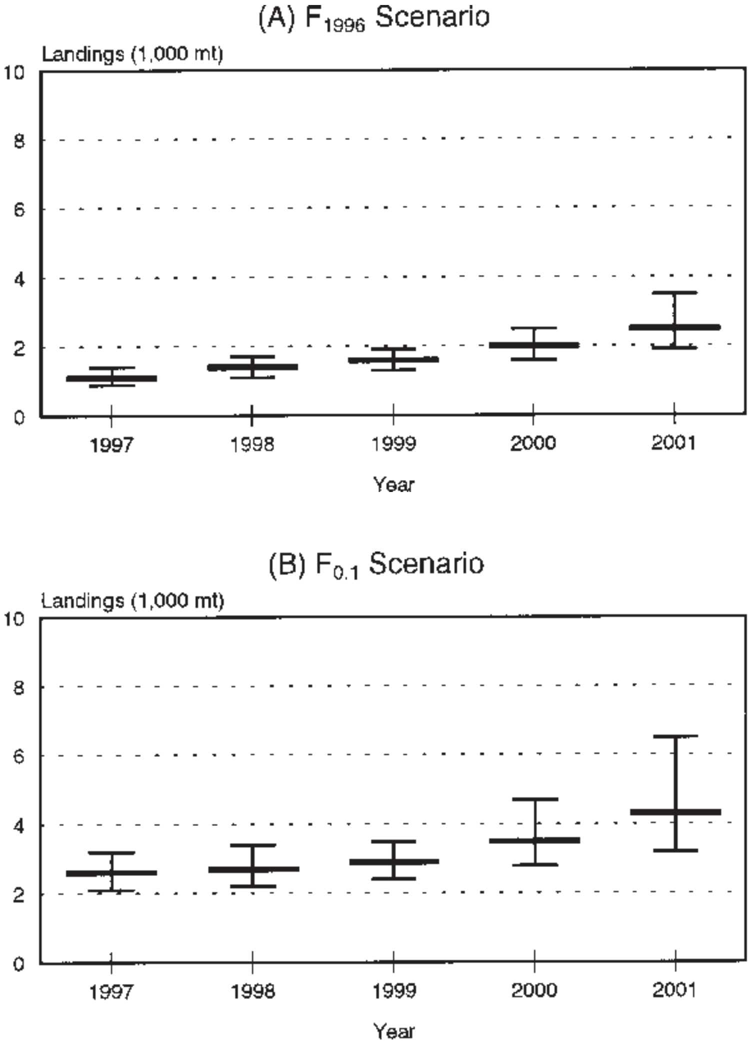


Figure 6. Projections of median landings of Georges Bank yellowtail flounder and its 80% confidence interval during 1997-2001 under (A) F_{1996}, (B) $F_{0.1}$, and (C) F_{MAX} management scenarios.

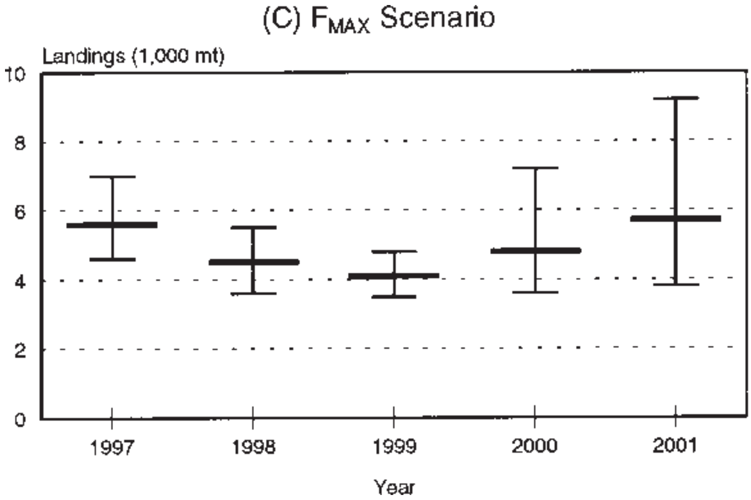


Figure 6. (Continued.)

projections showed that the expected level of variability in spawning biomass would be higher under the status quo scenario but variability in landings would be lower. Conversely, uncertainty in spawning biomass would be lower under the limit reference scenario but uncertainty in landings would be higher. Third, loss of yield due to discarding would be lowest under the status quo scenario and highest under the limit reference point scenario. Last, status quo levels of fishing effort appear sufficient to achieve the interim rebuilding target for Georges Bank yellowtail flounder during 1997-2001. Nonetheless, it is highly unlikely that the long-term target spawning biomass target would be achieved during the projection horizon, although the status quo scenario would lead to the greatest increase in spawning potential. Rebuilding the spawning potential of Georges Bank yellowtail flounder and other depressed New England groundfish stocks through effort control or other conservation measures is an important management consideration because rebuilding would lead to increased benefits (Edwards and Murawski 1993, Overholtz et al. 1995).

The projection approach presented here is not completely general because it does not consider model uncertainty or implementation uncertainty which may be important considerations in developing harvest policy (Rosenberg and Restrepo 1994). In some cases, for example, the model assumption that growth and maturation rates are constant through the projection horizon may not be tenable and density-dependence in growth, maturation, and natural mortality rates may need to be explicitly modeled (Overholtz 1993; Helser and Brodziak 1998). However, whether additional sources of uncertainty will affect the determination of harvest policy is

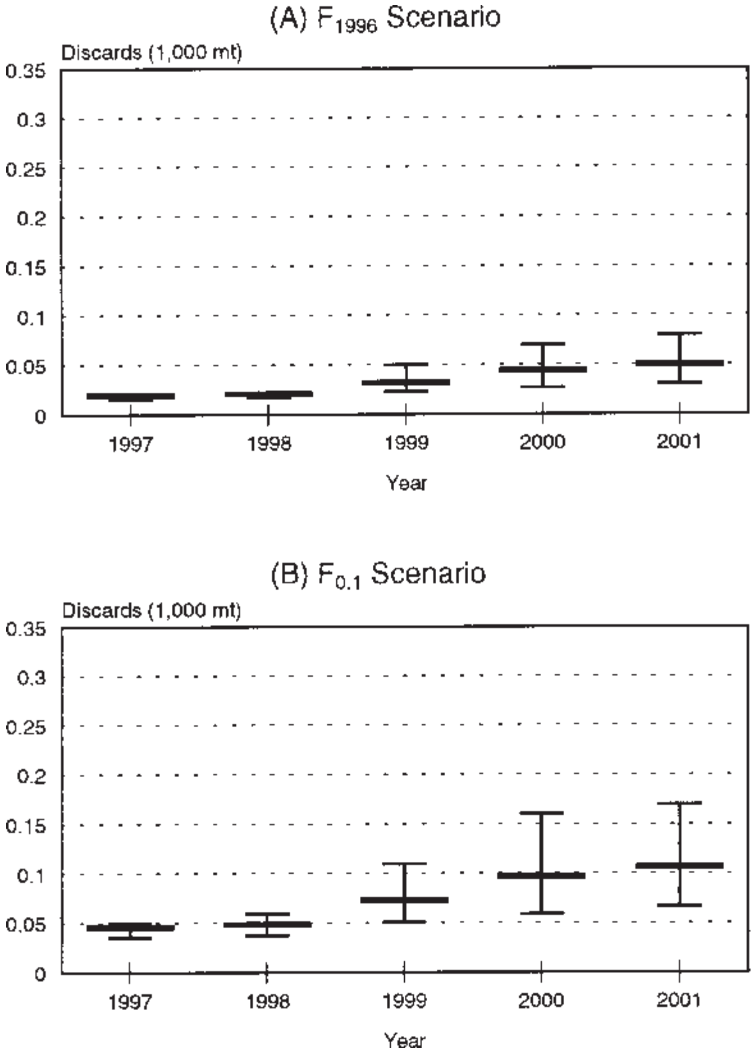


Figure 7. Projections of median discards of Georges Bank yellowtail flounder and its 80% confidence interval during 1997-2001 under (A) F_{1996}, (B) $F_{0.1}$, and (C) F_{MAX} management scenarios.

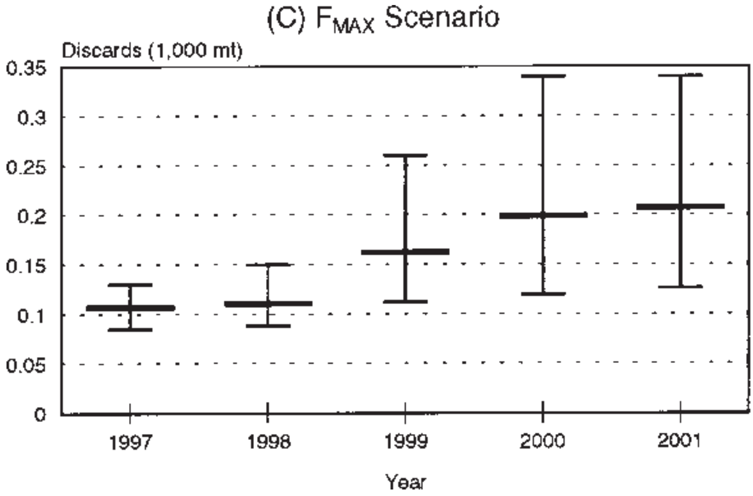


Figure 7. (Continued.)

application specific and dependent upon management objectives (Frederick and Peterman 1995). Although additional sources of random variation can readily be incorporated into this approach, uncertainty in initial population abundance and future recruitment will generally be primary factors to characterize for management advice.

Nonetheless, our approach to making short-term projections is a relatively simple algorithm based on the computationally intensive techniques of bootstrapping and Monte Carlo simulation. Others have used similar approaches to include uncertainty in initial conditions in projections (Mohn 1993, Smith and Gavaris 1993) and our purpose here is to emphasize the generality and flexibility of the approach as a way to evaluate the potential consequences of future management actions. In practice, this approach can readily be applied in any situation where residuals are independent and identically distributed (IID) errors conditioned on model structure or have a relatively simple composite error structure consisting of several IID components (cf. Mohn 1993). Furthermore, the approach is not restricted to age-structured assessment models. For example, this approach has been applied to analytical assessments of surf clams in the northwest Atlantic based on the modified DeLury assessment model (Conser 1994, Weinberg et al. 1995).

Parametric approaches to projecting the consequences of uncertainty in initial population abundance are, of course, also appropriate (Gavaris 1993) and may be computationally expedient in comparison to our resampling approach. However, any parametric approach requires an additional assumption to define the statistical distribution of errors and such assumptions are another potential source of error in projecting the consequences of alternative harvest policies. In many applications, the nonparametric bootstrap can be recommended because it is flexible and parsimonious and also because it has the desirable theoretical property of providing exact inference for nonparametric maximum likelihood estimation (Efron and Tibshirani 1993).

Acknowledgments

We thank J. Heifetz and two anonymous reviewers for their helpful comments on the draft manuscript. We also thank Steve Cadrin for providing data from the most recent Georges Bank yellowtail flounder assessment.

References

Baranov, F.I. 1918. On the question of the biological basis of fisheries. Nauchn. Issled. Ikhtiologicheskii Inst. Izv. 1:81-128.

Beverton, R.J.H., and S.J. Holt. 1957. On the dynamics of exploited fish populations. Fish and Fisheries Series 11, Chapman and Hall, New York. 533 pp.

Caddy, J.F., and R. Mahon. 1995. Reference points for fisheries management. FAO Fish. Tech. Pap. 347. 83 pp.

Cadrin, S.X., W.J. Overholtz, J.D. Neilson, S. Gavaris, and S.E. Wigley. 1997. Stock assessment of the Georges Bank yellowtail flounder stock for 1997. Northeast Fisheries Science Center Reference Document, Woods Hole, MA.

Conser, R.J. 1994. Stock assessment methods designed to support fishery management decisions in data-limited environments: Development and application. Ph.D. dissertation, Univ. of Washington, Seattle.

Conser, R.J., and J.E. Powers. 1990. Extension of the ADAPT VPA tuning method designed to facilitate assessment work on tuna and swordfish stocks. Int. Comm. Cons. Atl. Tunas, Coll. Vol. Sci. Pap. 32:461-467.

Conser, R.J., L. O'Brien, and W.J. Overholtz. 1991. An assessment of southern New England and Georges Bank yellowtail flounder stocks. Northeast Fisheries Science Center Reference Document 91-03, Woods Hole, MA.

Deriso, R.B., T.J. Quinn II, and P.R. Neal. 1985. Catch-age analysis with auxilliary information. Can. J. Fish. Aquat. Sci. 42:815-824.

Edwards, S.F., and S.A. Murawski. 1993. Potential economic benefits from efficient harvest of New England groundfish. N. Am. J. Fish. Manage. 13:437-449.

Efron, B. 1982. The jackknife, the bootstrap, and other resampling plans. SIAM (Soc. Ind. Appl. Math.) CBMS (Conf. Bd. Math. Sci.) Monograph 38. 92 pp.

Efron, B., and R. J. Tibshirani. 1993. An introduction to the bootstrap. Chapman and Hall, New York. 436 pp.

Fogarty, M.J., and S.A. Murawski. 1998. Large-scale disturbance and the structure of marine systems: Fishery impacts on Georges Bank. Ecol. Appl., Suppl. 8(1):S6-S22.

Fournier, D.A., and C.P. Archibald. 1982. A general theory for analyzing catch at age data. Can. J. Fish. Aquat. Sci. 39:1195-1207.

Frederick, S.W., and R.M. Peterman. 1995. Choosing fisheries harvest policies: When does uncertainty matter? Can. J. Fish. Aquat. Sci. 52:291-306.

Gavaris, S. 1988. An adaptive framework for the estimation of population size. Can. Atl. Fish. Sci. Adv. Comm. (CAFSAC) Res. Doc. 88/29.

Gavaris, S. 1993. Analytical estimates of the reliability of projected yield from commercial fisheries. In: S.J. Smith, J.J. Hunt, and D. Rivard (eds.), Risk evaluation and biological reference points for fisheries management. Can. Spec. Publ. Fish. Aquat. Sci. 120:185-191.

Gulland, J.A., and L.K. Boerema. 1973. Scientific advice on catch levels. Fish. Bull. 71:325-336.

Helser, T.E., and J.K.T. Brodziak. 1998. Impacts of density-dependent growth and maturation on assessment advice to rebuild depleted U.S. silver hake *(Merluccius bilinearis)* stocks. Can. J. Fish. Aquat. Sci. 55:882-892.

Hilborn, R., and C.J. Walters. 1992. Quantitative fisheries stock assessment. Chapman and Hall, New York. 570 pp.

Hilborn, R., E.K. Pikitch, and M.K. McAllister. 1994. A Bayesian estimation and decision analysis for an age-structured model using biomass survey data. Fish. Res. 19:17-30.

Laurec, A., and J.G. Shepherd. 1983. On the analysis of catch and effort data. J. Cons. Cons. Int. Explor. Mer 41:81-84.

McAllister, M.K., and J.N. Ianelli. 1997. Bayesian stock assessment using catch-age data and the sampling-importance resampling algorithm. Can. J. Fish. Aquat. Sci. 54:284-300.

Megrey, B.A. 1989. Review and comparison of age-structured stock assessment models from theoretical and applied points of view. In: E. Edwards and B. Megrey (eds.), Mathematical analysis of fish stock dynamics: Reviews and current applications. Amer. Fish. Soc. Symp. 6:8-48.

Mendelssohn, R. 1988. Some problems in estimating population sizes from catch-at-age data. Fish. Bull. U.S. 86:617-630.

Methot, R.D. 1989. Synthetic estimates of historical abundance and mortality for northern anchovy. In: E. Edwards and B. Megrey (eds.), Mathematical analysis of fish stock dynamics: Reviews and current applications. Amer. Fish. Soc. Symp. 6:66-82.

Methot, R.D. 1990. Synthesis model: An adaptable framework for analysis of diverse stock assessment data. Int. North Pac. Fish. Comm. Bull. 50:259-277.

Mohn, R.K. 1993. Bootstrap estimates of ADAPT parameters, their projection in risk analysis and their retrospective patterns. In: S.J. Smith, J.J. Hunt, and D. Rivard (eds.), Risk evaluation and biological reference points for fisheries management. Can. Spec. Publ. Fish. Aquat. Sci. 120:173-184.

Mohn, R.K., and R. Cook. 1993. Introduction to sequential population analysis. Northwest Atl. Fish. Organ. Sci. Counc. Studies 17.

New England Fishery Management Council. 1996. Amendment #7 to the Northeast Multispecies Fishery Management Plan. New England Fishery Management Council, Saugus, MA.

Northeast Fisheries Science Center. 1995. Status of the fishery resources off the northeastern United States for 1994. NOAA Tech. Memo NMFS-NE-108.

Northeast Fisheries Science Center. 1997. Report of the 24th Northeast Regional Stock Assessment Workshop. Northeast Fisheries Science Center Reference Document, Woods Hole, MA. 290 pp.

Overholtz, W.J. 1993. Harvesting strategies and fishing mortality reference point comparisons for the northwest Atlantic stock of Atlantic mackerel (*Scomber scombrus*). Can. J. Fish. Aquat. Sci. 50:1749-1756.

Overholtz, W.J., S.F. Edwards, and J.K.T. Brodziak. 1995. Effort control in the New England groundfish fishery: A bioeconomic perspective. Can. J. Fish. Aquat. Sci. 52:1944-1957.

Overholtz, W., S.A. Murawski, P.J. Rago, W.L. Gabriel, and M. Terceiro. 1997. Ten-year projections of landings, spawning biomass, and recruitment for five groundfish stocks considered at SAW 24. Northeast Fisheries Science Center Reference Document, Woods Hole, MA.

Parrack, M.L. 1986. A method of analyzing catches and abundance indices from a fishery. Int. Comm. Cons. Atl. Tunas, Coll. Vol. Sci. Pap. 24:209-221.

Punt, A.E. and D.S. Butterworth. 1993. Variance estimates for fisheries assessment: Their importance and how best to evaluate them. Can. Spec. Publ. Fish. Aquat. Sci. 120:145-162.

Ricker, W.E. 1975. Computation and interpretation of biological statistics of fish populations. Bull. Fish. Res. Board Can. 191.

Rosenberg, A.A., and V.R. Restrepo. 1994. Uncertainty and risk evaluation in stock assessment advice for U.S. marine fisheries. Can. J. Fish. Aquat. Sci. 51:2715-2720.

Schnute, J.T. 1994. A general framework for developing sequential fisheries models. Can. J. Fish. Aquat. Sci. 51:1676-1688.

Smith, S.J., and S. Gavaris. 1993. Evaluating the accuracy of projected catch estimates from sequential population analysis and trawl survey abundance estimates. In: S.J. Smith, J.J. Hunt, and D. Rivard (eds.), Risk evaluation and biological reference points for fisheries management. Can. Spec. Publ. Fish. Aquat. Sci. 120:163-172.

Weinberg, R.J., S.A. Murawski, R. Conser, J. Brodziak, L. Hendrickson, H.-L. Lai, and P. Rago. 1995. Analytical assessment of surfclam populations in the Middle Atlantic region of the United States. Northeast Fisheries Science Center Reference Document 95-05, Woods Hole, MA.

A Simulation Study of the Effects of Aging Error and Sample Size on Sustained Yield Estimates

Lewis G. Coggins Jr.
Alaska Department of Fish and Game, Kodiak, Alaska

Terrance J. Quinn II
University of Alaska Fairbanks, Juneau Center, School of Fisheries and Ocean Sciences, Juneau, Alaska

Abstract

A Monte Carlo simulation model of an exploited age-structured fish population was constructed to evaluate the effects of sampling and aging the catch on estimates of population parameters from catch-age analysis and resultant estimates of sustained yield. Aging error, error in the annual catch estimates, and error in the annual survey exploitable abundance estimates were included in the simulation. Three sets of simulation runs were conducted among small (100), medium (300), and large (900) sample sizes to estimate the catch-at-age composition. The simulation included five reader types with differing aging abilities resulting in a total of five catch-at-age compositions evaluated for each program replication. Sustained yield estimates from 1,000 replications of each reader type were compared to true sustained yield. Aging bias and imprecision had dramatic effects on estimated sustained yield: positive aging bias and imprecision generally caused underestimation of sustained yield, while negative aging bias caused overestimation of sustained yield. These results suggest that this component in stock assessment deserves greater attention in planning and practice. In addition, increases in sample size resulted in increased precision in sustained yield estimates only if the catch-at-age composition was estimated using a precise and unbiased reader type. While this evaluation was based on the Arctic grayling *(Thymallus arcticus)* population within the Gulkana River in Alaska, the simulation model can be

---

Current affiliation for L.G. Coggins Jr. is U.S.G.S. Alaska Biological Science Center, Anchorage, Alaska.

used for a variety of sport fish populations; a diskette and user manual are available.

Introduction

Fisheries management should be rooted in sound knowledge of all factors contributing to the dynamics of the exploited fish population (Summerfelt 1987). These factors include, but are not limited to, understanding the basic life history of the species, exploitation patterns of the prosecuting fisheries, and environmental factors that influence the abundance and distribution of the species. Once these basic biological parameters have been uncovered, various mathematical constructs to model population dynamics exist which can be used to define patterns and levels of exploitation to achieve management objectives. One model of particular utility in estimating sustained yield was proposed by Quinn and Szarzi (1993). Quinn and Szarzi's model (QS) assumes a constant fishing mortality harvest policy and relies on the results of catch-age analysis with auxiliary information (Deriso et al. 1985, 1989) to generate estimates of sustained fishing mortality and subsequent estimates of sustained yield in a Leslie matrix (Getz and Haight 1989) framework.

The concept of managing fisheries to achieve sustained yield is a management goal that is widely embraced but rarely explicitly defined. Various management objectives can result in a sustained yield policy but can have different implications with regard to the productivity of the stock, the expected age composition of the catch, and many other population parameters (Deriso 1987, Sissenwine and Shepard 1987, Clark 1991). Our study addressed one of the many management objectives which result in sustained yield, that is, the sustained yield (SY_{ST}) that results from applying a constant fishing mortality rate (F_{ST}) that tends to force the population to a steady long-term equilibrium abundance and age composition (Quinn and Szarzi 1993). This specific objective, conceptually equivalent to F_{rep} (Sissenwine and Shepard 1987), is applicable for many sport fishery management programs where the prevailing goal is "to maintain the historic size and age composition of the exploited population." Translating this broad goal into a specific management objective that is both quantifiable and defensible allows for evaluation of the successes and failures associated with management policies aimed at maintaining fish stocks in equilibrium. The QS model provides a structure to estimate F_{ST} and resultant SY_{ST} associated with a specific exploited fish population.

Like many complex age-structured analyses, the QS model relies on catch-at-age data which are subject to introduced error from sampling and age determination (aging error). Aging error occurs when a reader responsible for estimating the age of a structure assigns an age, called an age reading, that is different from the true age. Aging error can occur as a result of aging imprecision, defined as the variability among multiple readings of a single structure, aging bias, defined as the difference between the expected value of the observed age and the true age, or a combination of these two (Kimura and Lyons 1991).

Catch-age analysis with auxiliary information (CAGEAN) is the cornerstone to the QS model. CAGEAN belongs to a class of models generally known as age-structured stock assessment models (ASA; Megrey 1989). The effects of sample size and aging error on model estimates have been dealt with by various researchers in both ASA type models and in yield-per-recruit or biomass based models (Lai and Gunderson 1987, Bradford 1991). However, the effect of sample size and aging error has not been examined as related to sustained yield estimates from a combined ASA and Leslie matrix model like the QS model. The intent of this paper is to investigate the performance of the QS model given various sample sizes and types of aging error.

Methods

A FORTRAN computer program was constructed to model an exploited age-structured fish population and evaluate the effect of the processes of sampling and aging the catch on estimates of population parameters from catch-age analysis and resultant estimates of sustained yield. The program (AGEERR) contains four procedures: (1) generation of an exploited age-structured fish population with associated true catch- and abundance-at-age; (2) construction of observed catch-at-age data by incorporating measurement error due to sampling, aging error, and variability in the total catch; (3) estimation of population parameters through catch-age analysis with auxiliary information based on observed catch-at-age and survey data; and (4) estimation of sustained yield using the population parameter estimates from catch-age analysis with auxiliary information. Descriptions of the formulations of each of the procedures as well as the methods used to conduct the case study using Gulkana River grayling, an important sport fish population in Alaska, are presented in the following sections. The program can be used for a variety of sport fish populations; a diskette and user manual are available (Coggins 1997).

Generation of the True Population

To simulate the population under investigation, a time series of true catch- and abundance-at-age was generated using the typical recursion and Baranov catch equations (Baranov 1918). The relevant formulae are given as:

$$N_{a+1,\,t+1} = N_{a,t} e^{-Z_{a,t}}, \tag{1}$$

$$N_{A+,\,t+1} = N_{(A+)-1,t} e^{-Z_{(A+)-1,\,t}} + N_{A+,t} e^{-Z_{A+,\,t}}, \tag{2}$$

$$C_{a,t} = \mu_{a,t} N_{a,t}, \tag{3}$$

$$\mu_{a,t} = \frac{F_{a,t}}{Z_{a,t}}\left[1 - e^{-Z_{a,t}}\right], \tag{4}$$

$$Z_{a,t} = F_{a,t} + M, \tag{5}$$

$$F_{a,t} = s_a f_t, \tag{6}$$

where $N_{a,t}$ is the true abundance of age a fish in year t, $Z_{a,t}$ is the total instantaneous mortality rate, A+ is an aggregate age group, $C_{a,t}$ is the true catch, $\mu_{a,t}$ is the true exploitation rate, M is the true instantaneous natural mortality rate, $F_{a,t}$ is the true fishing mortality rate, s_a is the true gear selectivity coefficient, and f_t is the true full recruitment fishing mortality (Deriso et al. 1985, 1989). Equation (2) is a generalization of (1) appropriate when an aggregate age class (A+) is used. The common assumption of separable fishing mortality (Doubleday 1976, Pope 1977, Fournier and Archibald 1982, Pope and Shepherd 1982, Deriso et al. 1985) is implicit in equation (6).

Observed Catch-at-Age Data

Observed catch-at-age was calculated as the product of the observed catch-at-age proportion and the observed total catch as:

$$C'_{a,t} = C'_t\, \theta'_{a,t}, \tag{7}$$

where $C'_{a,t}$ is the observed catch of age a fish during year t, C'_t is the observed total catch, and $\theta'_{a,t}$ is the observed catch-at-age proportion. Variability was included in both the observed total catch and the observed catch-at-age composition. To mimic the uncertainty in the estimates of total catch, it was assumed that the observed total catch followed a normal distribution with expected value equal to the true total catch and a constant coefficient of variation (cv_C):

$$C'_t \sim N\left(C_t, \sigma_{C(t)}{}^2\right), \tag{8}$$

$$\sigma_{C(t)}{}^2 = \left(cv_C C_t\right)^2, \tag{9}$$

where C_t is the true total catch in year t, $\sigma_{C(t)}{}^2$ is the variance of the observed total catch, and cv_C is the constant coefficient of variation of the observed catch. This formulation allows one to specify the precision of the observed catch simply by an assumed constant coefficient of variation. Other forms of variance based on empirical sampling results could also be considered (Crone and Sampson 1998).

The observed age composition of the catch was generated by sampling and aging the true catch and incorporating error in the aging process. Catch sampling was conducted randomly without replacement such that the probability of selecting a fish of a particular age was equal to the true proportion of that age remaining in the catch.

Modeling Aging Error

The simulation uses multiple reader types to generate a set of aging error distributions that can be compared objectively. If a is the observed age, b is the true age, c is the constant bias of the inaccurate reader, σ is the constant standard deviation (SD), and $\sigma(b)$ is the SD as a function of true age, the five reader types are described as follows. Reader 0 (R0) is a perfect reader without aging error [$a = b$]. Reader 1 (R1) is characterized as the ideal "real life" reader, being both accurate (no bias) and precise (low variability) [$a \sim N(b, \sigma)$]. Reader 2 (R2) is accurate but less precise than reader R1 {$a \sim N[b, \sigma(b)], \sigma \le \sigma(b)$}. Reader 3 (R3) is inaccurate but precise [$a \sim N(b+c, \sigma)$]. Reader 4 (R4) is inaccurate and imprecise {$a \sim N[b+c, \sigma(b)]$}. Readers R1 and R3 always have identical constant and high precision relative to readers R2 and R4. Readers R2 and R4 have identical and variable low precision as a function of true age that is always defined to be less than or equal to the constant precision of readers R1 and R3.

Classification matrices were used to specify aging precision for each reader type. Richards et al. (1992) present the following formulation for constructing a classification matrix that is central to modeling aging error in this study. The classification matrix is a function of the parameter vector $\Phi = (\sigma_r, \sigma_A, \alpha)$ and specified as:

$$\sigma(b) = \begin{cases} \sigma_r + (\sigma_A - \sigma_r)\dfrac{1 - e^{-\alpha(b-r)}}{1 - e^{-\alpha(A-r)}}; & \alpha \neq 0 \\ \sigma_r + (\sigma_A - \sigma_r)\dfrac{b-r}{A-r}; & \alpha = 0 \end{cases}, \tag{10}$$

$$\chi_{ab}(\Phi) = \frac{1}{\sqrt{2\pi}\,\sigma(b)}\, e^{-\frac{1}{2}\left[\frac{a-b}{\sigma(b)}\right]^2}, \tag{11}$$

$$q(a|b,\Phi) = \frac{\chi_{ab}(\Phi)}{\sum\limits_{a=r}^{A} \chi_{ab}(\Phi)}, \tag{12}$$

where σ_r is the SD associated with recruit-age fish, σ_A is the SD associated with the oldest age fish, α is a parameter that governs the non-linearity of $\sigma(b)$, and $\chi_{ab}(\Phi)$ is the discrete normal probability density function of a

given b. The preceding three equations define the classification matrix [$\mathbf{Q}(\Phi)$] made up of the row vectors q, which are probabilities that a fish of true age b is assigned an observed age a.

There are two assumptions regarding the formulation of the classification matrix. The first assumption is:

$$q(b \mid b, \Phi) \geq q(a \mid b, \Phi),\ a \neq b\,.$$

This is the "modal" probability assumption and asserts that fish of true age b are assigned an observed age a equal to b with higher probability than any other observed age. The second assumption is:

$$\sum_{a=r}^{A} q(a \mid b, \Phi) = 1 \text{ for each } b.$$

This assumption assures that the probabilities of all the possible observed ages for any given true age sum to one.

Although the modal assumption assures that observed age a equals true age b with highest probability, observed age a cannot be said to be unbiased. This is because the observed age must lie within the lifespan of the species (or the true age range recruited to a fishery) and truncation of the observed age distribution occurs near the bounds of the true age range. For example, if the youngest fish of a given species ever captured had a true age of 3, the probability of assigning an observed age $a < 3$ would be zero. Therefore, the age of fish with a true age of 3 would tend to be overestimated. A similar argument can be made regarding the underestimation of the age of older fish.

Therefore, the simulation also allows for the inclusion of systematic bias by specifying a constant bias c. This was accomplished by adding the bias after the true age was modified with imprecision so that the imprecision was specified based on the true age. A negative value of the bias c caused under-aging of a structure while a positive value caused over-aging. It should be noted however, that the bias will not cause the observed age to traverse the true age range. For example, application of a negative bias will still result in an observed age no smaller than 2 if the lower bound of true age is equal to 2. Alternate formulations such as including bias as a function of true age may also be appropriate to consider in future investigations.

Catch-Age Analysis with Auxiliary Information

Catch-age analysis with auxiliary information was performed using a variant of CAGEAN (Deriso et al. 1985) called CAGEM. In CAGEM, a multinomial-like measurement error structure is assumed, appropriate given aging error in the observed catch-at-age data. We refer to measurement error as the deviation of an observed quantity from its true value in the course of sampling and aging, as is common in stock assessment (Deriso et al. 1985). The relevant formulae for the objective function for parameter estimation and survey exploitable abundance are:

$$\min\left\{\sum_{a,t}\frac{\left[C'_{a,t}-\hat{C}_{a,t}\right]^2}{\hat{C}_{a,t}}+\lambda_s\left[\ln\left(SUR_t\right)-\ln\left(\hat{EN}_t\right)\right]^2\right\}, \tag{13}$$

$$\hat{EN}_t=\sum_a \hat{N}_{a,t}\hat{s}_a\,, \tag{14}$$

where $C'_{a,t}$ is the observed catch of age a fish in year t, $\hat{C}_{a,t}$ is the estimated catch, SUR_t is the total exploitable abundance from an independent survey, $\hat{EN}_t$ is the estimated total exploitable abundance, λ_s is a weighting factor for the auxiliary survey information, $\hat{N}_{a,t}$ is the estimated abundance, and $\hat{s}_a$ is the estimated selectivity-at-age. The catch $\hat{C}_{a,t}$, abundance $\hat{N}_{a,t}$, and selectivity $\hat{s}_a$ were estimated within program CAGEM using equations (1)-(6).

For the simulation model, variability in the survey exploitable abundance was specified with a coefficient of variation as:

$$SUR_t \sim N\left[EN_t,\sigma_{S(t)}{}^2\right], \tag{15}$$

$$\sigma_{S(t)}{}^2 = \left[cv_S\left(EN_t\right)\right]^2, \tag{16}$$

where EN_t is the true exploitable abundance of age a fish in year t, $\sigma_{S(t)}{}^2$ is the variance of the survey total exploitable abundance around the true exploitable abundance, and cv_S is the assumed constant coefficient of variation of the survey total exploitable abundance.

Estimation of Sustained Yield

The estimation of sustained yield (SY_{ST}) assumes that the population is exploited with a constant fishing mortality (F_{ST}) and that early life survival, natural mortality and fecundity-at-age are constant. The management objective associated with SY_{ST} seeks to find F_{ST} that will take the population to its steady long-term equilibrium abundance and age composition. This was accomplished by finding F_{ST} that caused the net reproductive value of an r year-old fish (age at recruitment to the fishery) to equal 1. A simple interpretation of this equilibrium condition is that every fish recruited to the fishery must on average produce 1 recruit at age r. In the course of finding F_{ST}, it is necessary to estimate early life survival. This was accomplished by taking the mean of the annual estimates of early life survival. Once F_{ST} was found, SY_{ST} in the year following the last year of the analysis was estimated by projecting the abundance using fecundity-at-age, average early life survival, and survival-at-age. The relevant formulae for the estimation of F_{ST} and SY_{ST} are found in Quinn and Szarzi (1993).

Gulkana River Grayling Case Study

To examine the effect of sample size and aging error on estimates of sustained yield, AGEERR was used to simulate the estimation of sustained yield for the Arctic grayling *(Thymallus arcticus)* population of the mainstem Gulkana River (Bosch 1995). The population parameters estimated by Bosch (1995) were used to simulate the population among 24 runs of AGEERR, performed according to a factorial design with three sample sizes and eight levels of aging error (Table 1). Each of the runs from AGEERR used 1,000 Monte Carlo replications.

Bosch (1995) estimated sustained fishing mortality and yield (F_{ST}, SY_{ST}) using the QS model. The estimates of SY_{ST}, F_{ST}, and fecundity-at-age estimated by Bosch are reported in Table 2. Note that since Bosch used an aggregate age group of 7+ but considered the population to have a longevity of age 10, the plus group abundance was apportioned to the contributing age classes and the fecundity of the plus group was estimated (Coggins 1997). The parameter cv_S was set equal to the average annual coefficient of variation (22%) of the exploitable abundance estimates from mark-recapture studies of Gulkana River grayling (Bosch 1995). An estimate of cv_C was not available so an arbitrary value of 5% was used.

For the purposes of this study, we assumed that it was possible to select a single perfectly random sample of fish from the entire catch. Furthermore, three sample sizes (100, 300, and 900) were used to evaluate the effect of sample size on the sustained yield estimates from the QS model. The sample sizes were picked to generally reflect sampling rates which would be below, approximately equal to, or above acceptable limits to achieve age-class proportion estimates that were within 5% of the true value with 80% confidence assuming a strict multinomial error structure (Thompson 1987).

Two levels of aging precision (low and high cases) and 4 levels of bias were used to incorporate aging error in the simulations (Table 1). To specify aging precision germane to the R2 and R4 type readers [$\sigma(b)$] in what we call the low precision case, a data set was constructed containing estimated ages of scales collected from grayling before and after a known time-at-large (i.e., after capture and recapture in a mark-recapture experiment). These data were from grayling collected at Fielding Lake, and the Chatanika, Chena, Salcha, and Gulkana rivers in Alaska (Merritt and Fleming 1991, Bosch 1995). To estimate the standard deviation of observed age given a true age, we assumed that the first reading (at initial capture) was correct and the expected age of the second reading (at recapture) was the sum of the first reading and the time-at-large. The data were then

Table 1. Input parameters for each simulation run of AGEERR.

Run	Sample size	Aging precision	Aging bias (*c*)	σ_r[a]	σ_A[b]	α[c]	σ[d]	cv_s[e]	cv_c[f]	Age at full selectivity	Age at full maturity (*m*)	Pooling age
1	100	High	–1	0.41	0.64	–0.25	0.41	0.22	0.05	3	6	7
2	100	High	–2	0.41	0.64	–0.25	0.41	0.22	0.05	3	6	7
3	100	High	1	0.41	0.64	–0.25	0.41	0.22	0.05	3	6	7
4	100	High	2	0.41	0.64	–0.25	0.41	0.22	0.05	3	6	7
5	100	Low	–1	0.82	1.28	–0.25	0.82	0.22	0.05	3	6	7
6	100	Low	–2	0.82	1.28	–0.25	0.82	0.22	0.05	3	6	7
7	100	Low	1	0.82	1.28	–0.25	0.82	0.22	0.05	3	6	7
8	100	Low	2	0.82	1.28	–0.25	0.82	0.22	0.05	3	6	7
9	300	High	–1	0.41	0.64	–0.25	0.41	0.22	0.05	3	6	7
10	300	High	–2	0.41	0.64	–0.25	0.41	0.22	0.05	3	6	7
11	300	High	1	0.41	0.64	–0.25	0.41	0.22	0.05	3	6	7
12	300	High	2	0.41	0.64	–0.25	0.41	0.22	0.05	3	6	7
13	300	Low	–1	0.82	1.28	–0.25	0.82	0.22	0.05	3	6	7
14	300	Low	–2	0.82	1.28	–0.25	0.82	0.22	0.05	3	6	7
15	300	Low	1	0.82	1.28	–0.25	0.82	0.22	0.05	3	6	7
16	300	Low	2	0.82	1.28	–0.25	0.82	0.22	0.05	3	6	7
17	900	High	–1	0.41	0.64	–0.25	0.41	0.22	0.05	3	6	7
18	900	High	–2	0.41	0.64	–0.25	0.41	0.22	0.05	3	6	7
19	900	High	1	0.41	0.64	–0.25	0.41	0.22	0.05	3	6	7
20	900	High	2	0.41	0.64	–0.25	0.41	0.22	0.05	3	6	7
21	900	Low	–1	0.82	1.28	–0.25	0.82	0.22	0.05	3	6	7
22	900	Low	–2	0.82	1.28	–0.25	0.82	0.22	0.05	3	6	7
23	900	Low	1	0.82	1.28	–0.25	0.82	0.22	0.05	3	6	7
24	900	Low	2	0.82	1.28	–0.25	0.82	0.22	0.05	3	6	7

[a] σ_r = Standard deviation of observed age given recruit age fish.

[b] σ_A = Standard deviation of observed age given oldest age fish.

[c] α = Parameter governing the non–linearity of observed age as a function of true age.

[d] σ = Constant standard deviation of reader types R1 and R3.

[e] cv_s = Coefficient of variation of the survey data.

[f] cv_c = Coefficient of variation of the total catch data.

sorted by expected age and the standard deviation of observed age was estimated for each expected age class. Both formulations of equation (10) were then fit to the data to estimate σ_r, σ_A, and α. Although the three parameter case was not found to have a significantly better fit than the two parameter case ($p = 0.57$) using an *F*-test proposed by Schnute (1981), the three parameter case was used throughout the simulations so that precision of readers R2 and R4 could be specified non-linearly. The constant aging precision of readers R1 and R3 (σ) in the low precision case was set equal to the parameter estimate for σ_r.

As described above, the parameters of the function $\sigma(b)$ in equation (10) in the low precision case were estimated from available data, and the low precision case of σ set equal to the estimated σ_r . For the high precision case, the σ_r and σ_A parameters of the function $\sigma(b)$ were set equal to

Table 2. Estimated sustained yield (SY_{ST}), sustained fishing mortality (F_{ST}), and population parameters for the grayling stock of the Gulkana River in Alaska either contained in, or estimated from, Bosch (1995). Selectivity and fecundity are assumed to be constant over the time of the simulation.

Sustained fishing mortality	$F_{ST} = 0.402$
Sustained yield (number of fish)	$SY_{ST} = 29{,}867$
Projected abundance in 1992 (number of fish)	Proj. Abu. = 153,133

Fishing mortality

Year	f_t
1986	0.330
1987	0.232
1988	0.190
1989	0.154
1990	0.059
1991	0.081

Abundance (number of fish)

	Age					
Year	2	3	4	5	6	7+
1986	34,253	30,311	23,874	4,540	2,331	81
1987	41,271	23,852	16,127	12,702	2,416	1,283
1988	22,390	29,274	14,007	9,471	7,459	2,172
1989	46,004	16,010	17,946	8,587	5,806	5,905
1990	90,218	33,117	10,175	11,405	5,457	7,443
1991	16,940	66,100	23,125	7,105	7,964	9,008

	Age					
	2	3	4	5	6	7+
Selectivity	0.187	1	1	1	1	1
Net fecundity (eggs)	124	469	975	1,616	2,739	4,140

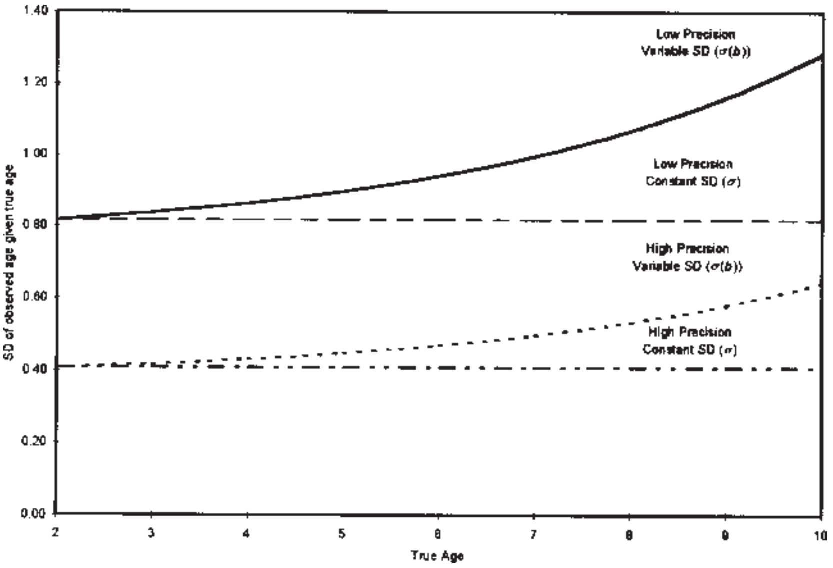


Figure 1. Standard deviation (SD) of observed age a given true age b for both the high and low precision cases of variable [σ(b)] and (σ) constant SD.

one half the values in the low precision case, and the constant value of σ was set equal to σ_r (Table 1 and Fig. 1). The bias levels were set equal to –1, –2, 1, and 2.

Lambda (λ_s), the weighting parameter that specifies the amount of influence the survey data has in parameter estimation (equation 13), was specified as 1,300 based on simulations of CAGEM during its development. To test the performance of $\lambda_s = 1{,}300$, a simulation set was run with no specified error in any of the input parameters. The parameter estimates given the R0 reader type and a sample size of 1,000 differed from the true parameter estimates by a negligible amount.

To evaluate the process of estimating sustained yield with the QS model and the variable input data detailed above, two statistics were examined. The first was the relative error, an indicator of bias, between the parameter estimate and the true value of that parameter. Relative error is calculated as:

$$RE = \frac{\hat{\theta} - \theta}{\theta}, \tag{17}$$

where *RE* is the relative error, θ is the true value of the parameter, and $\hat{\theta}$ is the parameter estimate. In a Monte Carlo framework, the parameter estimate is equal to the mean of the estimates of that parameter among the Monte Carlo replications.

Following Hightower (1996), the second statistic was the proportion of replicate parameter estimates among Monte Carlo replications that were within 10% of the true value. This statistic, like relative error, is an indicator of bias but has the additional property of revealing the precision of the process by estimating the probability of obtaining an "accurate" estimate (relative error < 10%).

Results

Overview of Sustained Yield Estimation

The estimation of sustained yield relies directly on estimates of sustained fishing mortality, projected abundance, and selectivity of age-2 fish. In general, an increase in any one of these parameters caused an increase in sustained yield. Moreover, negatively biased estimates of sustained fishing mortality and projected abundance resulted in negatively biased estimates of sustained yield. The exceptions occurred when negative aging bias caused extreme overestimation (relative error > 750%) of the selectivity of age-2 fish and offset the effect of underestimation from sustained fishing mortality and projected abundance. To show these general results, we averaged over all scenarios and expressed results in terms of relative error (Fig. 2).

The R0 reader type produced highly accurate (relative error < 5%) but slightly negatively biased estimates of SY_{ST}, F_{ST}, and projected abundance, as well as positively biased estimates of the selectivity of age-2 fish (relative error < 10%; Fig. 2). Under high aging precision, the R1 and R2 reader types produced accurate estimates of SY_{ST} (relative error < 10%), and fairly inaccurate estimates of SY_{ST} (relative error < 40%) under low aging precision. In all cases of the R1 and R2 reader types, F_{ST} and projected abundance were negatively biased while the selectivity of age-2 fish was positively biased.

Under both the high and low cases of aging precision, positive aging bias resulted in extreme underestimation of SY_{ST} (nearly 100% relative error) from both the R3 and R4 reader types. This was primarily a result of extreme underestimation of F_{ST} (roughly 100%), but also a result of underestimation of projected abundance. In addition, the positive aging bias caused the exclusion of age-2 (+1 bias) and age-2 and age-3 fish (+2 bias) from the analyses and therefore, selectivity of age-2 fish was not estimated. Under negative aging bias, SY_{ST} was overestimated for both high and low aging precision using the R3 and R4 reader types. This was a result of moderate overestimation of F_{ST} (relative error < 12%) and huge overestimation of the selectivity of age-2 fish (relative error > 750%). The low precision case produced more accurate estimates of SY_{ST} than the high precision case from the R3 and R4 reader types under negative aging bias.

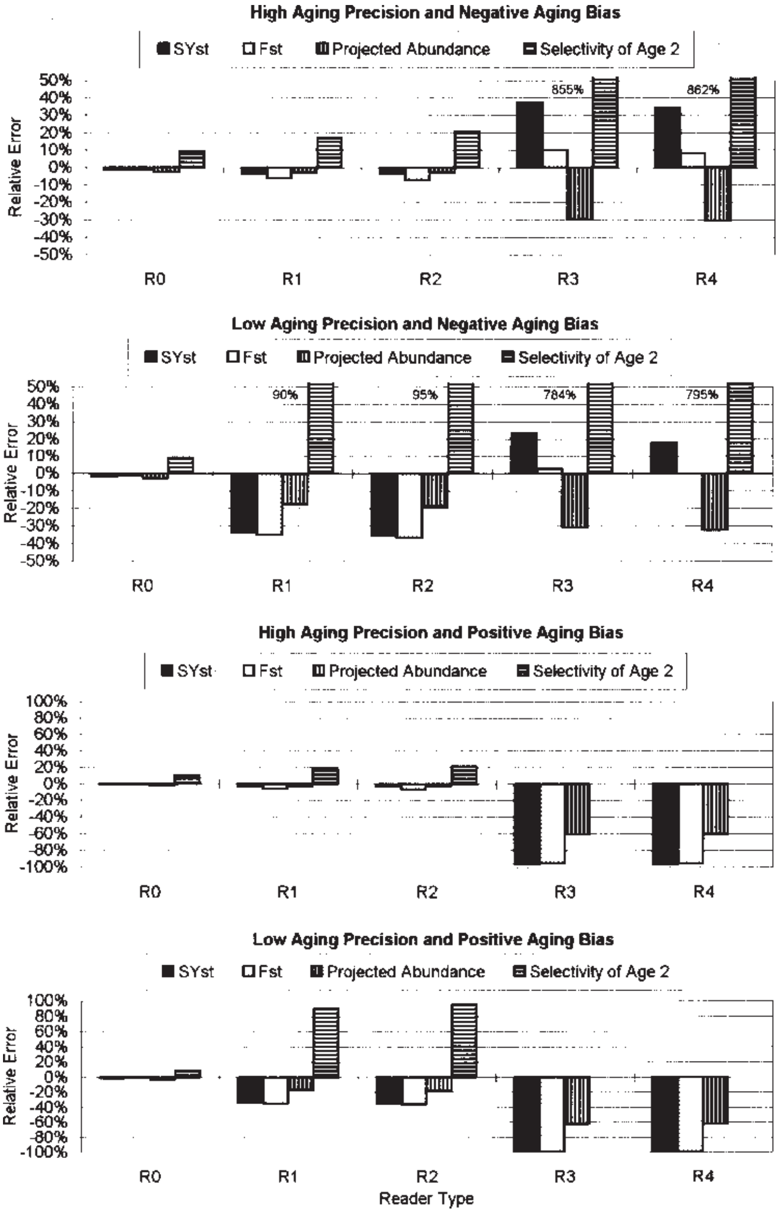


Figure 2. Average relative error in estimates of sustained yield (SY_{ST}), sustained fishing mortality (F_{ST}), projected abundance, and selectivity of age-2 fish under high and low aging precision and negative (two upper panels) and positive (two lower panels) aging bias, among all samples sizes and reader types (see Modeling Aging Error for descriptions of reader types R0-R4).

Sustained Yield by Reader Types and Sample Size

In this section, detailed examination of sustained yield estimates by reader type and sample size is described. The estimated replacement sustained yield in 1992 ranged from a low of 0 fish to a high of 42,959 fish corresponding to relative errors of –100% and 44%, respectively (Fig. 3). The R0 reader type produced estimates of SY_{ST} nearly identical to the true sustained yield but biased slightly low. This slight underestimate was due to the method of back transformation of the lognormal distribution used in AGEERR; which does not use the classic back-transformation defined as $\hat{\theta} = \exp(\mu + {}^{\sigma^2}\!/_2)$, but instead conducts the back-transformation as $\hat{\theta} = \exp(\mu)$. This method of back transformation inherently leads to slightly biased estimates, but is justified because there is one fewer parameters to be estimated. Under high precision, the SY_{ST} estimates from the R1 and R2 reader types are essentially equal to the R0 reader type and therefore, only slightly negatively biased. Under low precision, the R1 and R2 estimates of SY_{ST} are more negatively biased (relative error $\cong$ 35%) than under the high precision case. The R3 and R4 reader types produced estimates of SY_{ST} that are positively biased under negative aging bias and negatively biased under positive aging bias, regardless of aging precision. However, under low aging precision and negative aging bias, the magnitude of the bias in SY_{ST} is less than under high aging precision. Finally, the level of bias (i.e., relative error) is essentially unchanged for each reader type, level of aging precision, and direction/magnitude of aging bias among the three sample sizes.

While sample size did not affect the bias in estimates of sustained yield, it did influence the variability of the sustained yield estimates. This is shown by the proportion of replicate sustained yield estimates that are within 10% of the true sustained yield and is essentially the probability of obtaining an "accurate" estimate. The average proportion of replicate estimates of SY_{ST} that were within 10% of the true value ranged between 0% and 36% among all levels of specified aging error, reader types, and sample sizes (Fig. 4). As expected, the highest proportions were obtained with reader R0, with proportions ranging from approximately 24%, 31%, and 35% among sample sizes of 100, 300, and 900, respectively. For readers R1 and R2, the proportions were essentially identical to the R0 reader type under high aging precision, but dropped substantially under low aging precision. The proportions of accurate replicate estimates from the R3 and R4 reader types were uniformly less than or equal to the R0 reader type, identically equal to zero under positive aging bias, and greater than the R1 and R2 reader types only under low aging precision and negative bias. The general trend supported by the results is that increasing sample size results in increasing the probability of obtaining an accurate estimate only if the catch sample is aged by an accurate and precise reader.

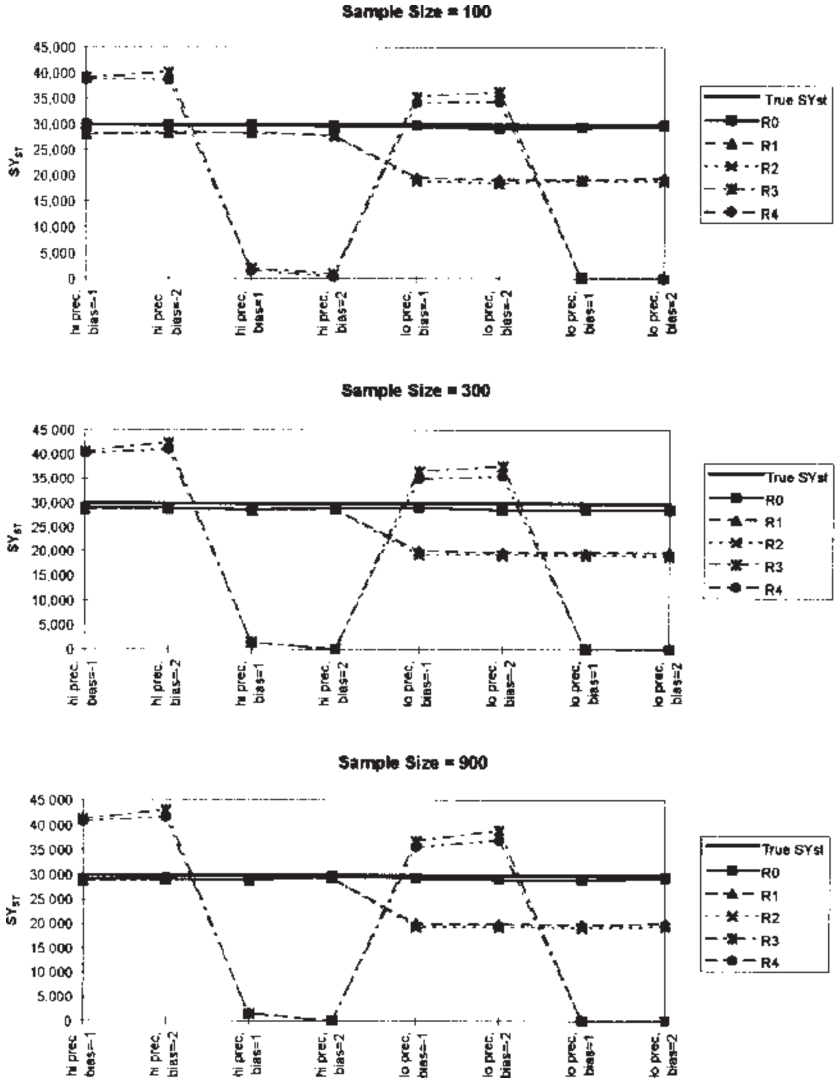


Figure 3. Estimated value of sustained yield (SY_{ST}) among reader types R0-R4 as compared to true sustained yield under different combinations of aging precision and bias for sample sizes of 100, 300, and 900 (see Modeling Aging Error for descriptions of reader types R0-R4).

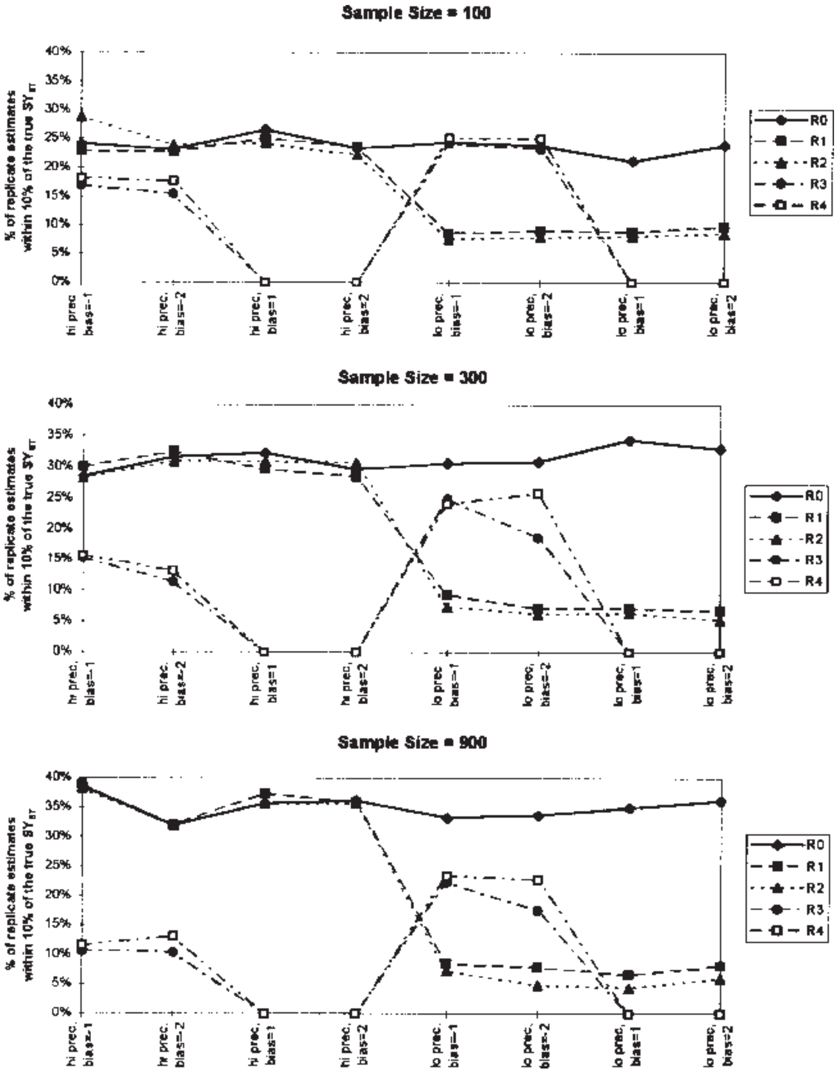


Figure 4. Proportion of Monte Carlo replicates that produced estimates of sustained yield (SY_{ST}) that were within 10% of the true value under different combinations of aging precision and bias for sample sizes of 100, 300, and 900 (see Modeling Aging Error for descriptions of reader types R0-R4).

Discussion

Effects of Sample Size on Estimates of Sustained Yield

The effect of sample size on estimates of SY_{ST} was of little consequence as demonstrated by the observation that estimates by reader type and aging error combination varied little among different sample sizes. This result suggests that attempting to mitigate bias in sustained yield estimates resulting from poor reader performance by increasing sample size is not an effective tactic.

In contrast, sample size does play a major role in determining the variability of sustained yield estimates. The probability of obtaining an estimate of sustained yield that is within 10% of the true value given a particular Monte Carlo replicate generally increases with increasing sample size. Kimura (1990) reported a similar trend in the precision of biomass estimates from ASA models as a function of sample size. This result was uniformly true among the R0 reader type and the R1 and R2 reader types under the high aging precision case. However, for the R1 and R2 reader types under the low precision case, the probability is either unchanged or decreasing as a function of increasing sample size. Similarly for the biased R3 and R4 reader types, an increasing sample size tended to have either no effect or a degrading effect on the probability of obtaining an accurate estimate.

The unexpected result that increasing sample size does not always increase the probability of obtaining an accurate estimate is probably also due to the influence other types of information have in catch-age analysis, namely estimates of relative or absolute abundance. The weighting procedures used for this other information can have a large effect on abundance estimation and, hence, obviate the influence of the aging information. Furthermore, the more imprecise the aging information is, the more deleterious an effect it may have with larger sample size, being in competition with the other abundance information.

Effects of Aging Error on Estimates of Sustained Yield

In comparison to the effects of sample size on the estimated value of SY_{ST}, the effects of aging error are more pronounced. Furthermore, aging bias tends to have greater effect on estimates of sustained yield than aging precision. In general, negative aging bias tended to produce positively biased estimates of SY_{ST} and positive aging bias tended to produce negatively biased estimates of sustained yield. This result is similar to the findings published by other researchers studying the effects of aging error on ASA and yield-per-recruit models (Rivard 1983, Lai and Gunderson 1987, Tyler et al. 1989, Kimura 1990). Also, this suggests that using bony structures that have a demonstrated tendency to negatively bias the age of fish (i.e. scales, whole otoliths; Pikitich and Demory 1988, Wilson and Boehlert 1990, Secor et al. 1995), could lead to ostensibly optimistic estimates of SY_{ST}. Furthermore, species that are extremely long lived and tend to approach asymptotic growth in mid-life, may also be subject to substantial negative aging bias.

The effect of decreasing aging precision on the expected value of SY_{ST} was to decrease the expected value of SY_{ST}. This result was particularly evident given the unbiased R1 and R2 reader types among all scenarios. This is a heartening result for fisheries managers, since it suggests that decreasing aging precision results in more conservative management strategies. Furthermore, in the case of the biased R3 and R4 reader types, the effect of aging imprecision caused the estimated value of SY_{ST} to be less biased given negative aging bias in the low precision case versus the high precision case, because of the offsetting trends of negative aging bias and low aging precision. However, these results should not be used to imply that managing fisheries with imprecise data is desirable.

Effects of Model Structure

By definition, the structure of a model affects output and, hence, implications drawn from its use. Recall that in AGEERR the introduction of aging error was a two step process: (1) incorporating imprecision as a function of true age using a classification matrix and (2) given a biased reader type, adding a constant bias. When incorporating aging error due to imprecision, the process prohibited the resultant observed age from traversing the true age range. Furthermore, when subsequently incorporating aging bias, the resultant observed age was again restricted to the true age range.

As a result, under negative aging bias, the biased R3 and R4 reader types produced age compositions ranging from 2 to 9 (bias of –1) or 2 to 8 (bias of –2). After the incorporation of the plus group, the age ranges considered in the catch-age analysis were 2 to 7+ for both –1 and –2 biases, and the resultant catch-at-age was highly skewed towards age-2 fish. The net effect on parameter estimates was to grossly overestimate the selectivity of age-2 fish (relative error > 750%), and to underestimate the projected abundance.

Under positive aging bias, the biased R3 and R4 reader types produced age compositions ranging from 3 to 10 (bias of +1) or 4 to 10 (bias of +2). After the incorporation of the plus group, the age ranges considered in the catch-age analysis were 3 to 7+ (+1 bias) and 4 to 7+ (+2 bias). Furthermore, since the simulation assumed that all ages > 2 were fully recruited, the catch-age analysis did not estimate any selectivity coefficients. The net effect was to cause even larger underestimation in the projected abundance than generated under negative aging bias.

Other strategies to incorporate aging bias could also be used, and we encourage their investigation. An alternative method for modeling aging bias is to specify bias directly in the transition matrix (Rivard 1983, Tyler et al. 1989, Bradford 1991). In addition, one could specify a bias as a function of age. A third strategy could be to specify a constant bias throughout the age range but allow observed age to traverse the age range. One could then disregard all observed ages outside the true age range. We suspect that all these methods would produce similar trends in the relative error of sustained yield, because the same changes in the distribution of ages would occur. The results using the methods incorporated in our study may be more dramatic than what would be found using alternate methods of modeling aging error. However, the focus of our study is to point out the relative trends caused by sampling and aging error rather than to predict absolute error. Indeed, to model all the vagaries found in estimating the age of fishes is nearly impossible. Moreover, the conclusions drawn from our study do provide fishery resource managers baseline information regarding how the inherent variability (uncertainty) surrounding age-composition data can influence fishery-related parameters routinely estimated in fish stock assessments.

Our study emphasizes the importance of careful validation of aging techniques. Common problems encountered in determining the age of fish are: (1) measurement error due to imprecision caused by poor aging judgment and ability (e.g. readers failing to recognize all the extant annuli in structures collected from older fish); and (2) process error due to annuli not being formed (e.g. missing first year annuli in some rainbow trout [Lentsch and Griffith 1987]). Since imprecision and bias can act in different directions, the resultant effect on estimates of sustained yield may be unclear. Careful age validation and development of aging protocols is needed to validate estimates of sustained yield when aging error is present.

Acknowledgments

Funding for this research was provided by the Alaska Department of Fish and Game, Sport Fish Division, Central Region. We thank Doug McBride for facilitating this study and members of Coggins' graduate committee, Lew Haldorson, Jim Hasbrouck, and Jim Reynolds, for their help. We also thank referee Paul Crone, an anonymous referee, and editor Jon Heifetz for helpful suggestions that led to a much improved manuscript.

References

Baranov, F.I. 1918. On the question of the biological basis of fisheries. Nauchn. Issled. Ikhtiologicheskii Inst. Izv. 1:81-128. (In Russian.)

Bradford, M.J. 1991. Effects of aging errors on recruitment time series from sequential population analysis. Can. J. Fish. Aquat. Sci. 48:555-558.

Bosch, D.E. 1995. Population dynamics and stock assessment of arctic grayling in the Gulkana River drainage, Alaska. Master's thesis, University of Alaska Fairbanks. 190 pp.

Clark, W.G. 1991. Groundfish exploitation rates based on life history parameters. Can. J. Fish. Aquat. Sci. 48:734-750.

Coggins Jr., L.G. 1997. The effects of sample size and aging error on estimates of sustained yield. Master's thesis, University of Alaska Fairbanks. 193 pp.

Crone, P.R., and D.B. Sampson. 1998. Evaluation of assumed error structure in stock assessment models that use sample estimates of age composition. In: F. Funk, T.J. Quinn II, J. Heifetz, J.N. Ianelli, J.E Powers, J.F. Schweigert, P.J. Sullivan, C.-I. Zhang (eds.), Fishery stock assessment models. Alaska Sea Grant College Program, University of Alaska Fairbanks, AS-SG-98-01 (this volume).

Deriso, R.B. 1987. Optimal $F_{0.1}$ criteria and their relationship to maximum sustained yield. Can. J. Fish. Aquat. Sci. 44(Suppl. 2):339-348.

Deriso, R.B., T.J. Quinn II, and P.R. Neal. 1985. Catch-age analysis with auxiliary information. Can. J. Fish. Aquat. Sci. 42:815-824.

Deriso, R.B., P.R. Neal, and T.J. Quinn II. 1989. Further aspects of catch-age analysis with auxiliary information. In: R.J. Beamish and G.A. McFarlane (eds.), Effects of ocean variability on recruitment and an evaluation of parameters used in stock assessment models. Can. Spec. Publ. Fish. Aquat. Sci. 108:127-135.

Doubleday, W.G. 1976. A least squares approach to analyzing catch at age data. Int. Comm. Northw. Atl. Fish. Res. Bull. 12:69-81.

Fournier, D.A., and C.P. Archibald. 1982. A general theory for analyzing catch at age data. Can. J. Fish. Aquat. Sci. 39:1195-1207.

Getz, W.M., and R.G. Haight. 1989. Population harvesting: Demographic models of fish, forest, and animal resources. Princeton University Press, Princeton, NJ. 391 pp.

Hightower, J.E. 1996. Comparison of reliability of catch-at-age analyses using auxiliary survey estimates of relative or absolute recruitment. Can. J. Fish. Aquat. Sci. 53:70-79.

Kimura, D.K. 1990. Approaches to age structured separable sequential population analysis. Can. J. Fish. Aquat. Sci. 47:2364-2374.

Kimura, D.K., and J.L. Lyons. 1991. Between-reader bias and variability in the age-determination process. Fish. Bull. U.S. 89:53-60.

Lai, H.L., and D.R. Gunderson, 1987. Effects of aging errors on estimates of growth, mortality, and yield per recruit for walleye pollock (*Theragra chalcogramma*). Fish. Res. 5:287-302.

Lentsch, L.D., and J.S. Griffith. 1987. Lack of first-year annuli on scales: Frequency of occurrence and predictability in trout of the western United States. In: R.C. Summerfelt and G.E. Hall, (eds.), Age and growth of fish. Iowa State University Press, Ames, IA, pp. 177-188.

Megrey, B.A. 1989. Review and comparison of age-structured stock assessment models from theoretical and applied points of view. Amer. Fish. Soc. Symp. 6:8-48.

Merrit, M.F., and D.F. Fleming. 1991. Evaluations of various structures for use in age determination of arctic grayling. Alaska Department of Fish and Game, Fishery Manuscript No. 91-6. Anchorage, AK. 21 pp.

Pikitch, E.K., and R.L. Demory. 1988. Assessment of scales as a means of aging dover sole. Trans. Am. Fish. Soc. 117:345-349.

Pope, J.G. 1977. Estimation of fishing mortality, its precision and implications for the management of fisheries. In: J.H. Steele (ed.), Fisheries mathematics. Academic Press, New York, pp. 63-76.

Pope, J.G., and J.G. Shepherd. 1982. A simple method for the consistent interpretation of catch-at-age data. J. Cons. Cons. Int. Explor. Mer 40:176-184.

Quinn II, T.J., and N.J. Szarzi. 1993. Determination of sustained yield in Alaska's recreational fisheries. In: G. Kruse, D.M. Eggers, R.J. Marasco, C. Pautzke, and T.J. Quinn II (eds.), Proceedings of the International Symposium on Management Strategies for Exploited Fish Populations. University of Alaska Sea Grant, AK-SG-93-02, Fairbanks, pp. 61-84.

Richards, L.J., J.T. Schnute, A.R. Kronlund, and R.J. Beamish. 1992. Statistical models for the analysis of aging error. Can. J. Fish. Aquat. Sci. 49:1801-1815.

Rivard, D. 1983. Effects of systematic, analytical, and sampling errors on catch estimates: A sensitivity analysis. In: W.G. Doubleday and D. Rivard (eds.), Sampling commercial catches of marine fish and invertebrates. Can. Spec. Publ. Fish. Aquat. Sci. 66:114-129.

Schnute, J. 1981. A versatile growth model with statistically stable parameters. Can. J. Fish. Aquat. Sci. 37:241-247.

Secor, D.H., T.M. Trice, and H.T. Hornick. 1995. Validation of otolith-based aging and a comparison of otolith and scale-based aging in mark-recaptured Chesapeake Bay striped bass, *Morone saxatilis*. Fish. Bull. U.S. 93:186-190.

Sissenwine, M.P., and J.G. Shepard. 1987. An alternative perspective on recruitment overfishing and biological reference points. Can. J. Fish. Aquat. Sci. 44:913-918.

Summerfelt, R.C. 1987. Preface. In: R.C. Summerfelt and G.E. Hall (eds.), Age and growth of fish. Iowa State University Press, Ames, IA, pp. xii-xiv.

Tyler, A.V., R.J. Beamish, and G.A. McFarlane. 1989. Implications of age determination error to yield estimates. In: R.J. Beamish and G.A. McFarlane (eds.), Effects of ocean variability on recruitment and an evaluation of parameters used in stock assessment models. Can. Spec. Publ. Fish. Aquat. Sci. 108:27-35.

Thompson, S.K. 1987. Sample size for estimating multinomial populations. American Statistician 41:42-46.

Wilson, C.D., and G.W. Boehlert. 1990. The effects of different otolith aging techniques on estimates of growth and mortality for the splitnose rockfish, *Sebastes diploproa*, and canary rockfish, *S. pinniger*. California Fish and Game 76(3):146-160.

The Effects of Artisanal Fisheries on Penaeid Shrimp Stocks in the Gulf of Mexico

Adolfo Gracia
Instituto de Ciencias del Mar y Limnología, Mexico City, México

Ana Rosa Vázquez Bader
Instituto de Ecología, Mexico City, México

Abstract

In the Gulf of Mexico, exploitation of the three main commercial penaeid species (brown shrimp *Penaeus aztecus*, white shrimp *P. setiferus* and pink shrimp *P. duorarum*) by artisanal fisheries is carried out in the nursery areas inside coastal lagoons. In addition, a new artisanal fishery on adult white shrimp has been developed. Present data on the offshore industrial fishery for white and pink shrimp shows a strong decrease to less than 20% of the mean maximum yield attained during 1970-1980 (1,700 and 7,000 metric tons whole weight, respectively). The sustained increase of artisanal fishing effort is one of the main factors that affected shrimp production. Available fishery catch-effort data of offshore fisheries showed a normal decline in catch associated with decreasing effort. However, this does not consider artisanal fishing effort on shrimp stocks causing growth overfishing on the three penaeid shrimp and recruitment overfishing on white shrimp. The lack of adequate artisanal fishing effort records does not allow assessment of its magnitude correctly. Both artisanal and industrial fisheries, besides affecting yield, may reduce shrimp reproductive potential and lead stocks to a collapse. Simulation exercises based on age-structured models are used to assess the effect of different fisheries and define management strategies. A multistage model connecting a series of stock-recruitment curves over main life-history stages could be adequate for defining broad strategies to keep penaeid shrimp stocks in good condition.

Introduction

Penaeid shrimp is a valuable resource, exploited mostly in tropical seas over a wide range of species (García 1989). Worldwide demand of this product is high and not satisfied. On the contrary, its demand has been growing steadily in the last decades. Due to the high price per unit of weight and the consumption of different sizes and ages, wild penaeid shrimp populations have been subjected to high exploitation in their juvenile and adult stages.

The worldwide shrimp industry, which started in the 1950s, developed very quickly and is presently at its maximal production limits, as wild shrimp populations seem to have reached their maximal production stage, and unexploited shrimp stocks are not likely to be found. Shrimp aquaculture has been considered a viable and important alternative to meet the demand deficit, and at the same time relieve some of the pressure exerted on the wild populations. However, sometimes shrimp aquaculture has exerted an additional pressure on life cycle stages not traditionally exploited due to the postlarval demand for pond stocking. Postlarval extraction plus sequential fishing of juvenile and adults has caused some species, such as the Pacific white shrimp *Penaeus vannamei*, to be exploited during almost their whole life cycle. The future of shrimp populations subjected to such intense fishing pressure seems uncertain (Gracia and Vázquez Bader, in press).

The objectives of our study are to examine the interactions of artisanal and industrial fisheries for white *(P. setiferus)* and pink shrimp *(P. duorarum)* in the Gulf of Mexico. We also examine the potential implications of exploitation of shrimp postlarvae for use in aquaculture. In our analysis we consider the diverse stages of the shrimp life cycle and examine the implications on the recruitment process and reproductive potential.

Catch History

The shrimp fishery in Mexican waters of the Gulf of Mexico (Fig. 1) started in the 1950s and, just as most penaeid fisheries, developed quickly, reaching its peak production in the 1970s. During this period, catches of over 30,000 t (total weight) per year, with a maximum of 39,500 t in 1972, were obtained with an associated fishery effort of 1,100 trawler boats (Fig. 2). Later on, in the 1980s, total annual production of shrimp declined to almost 20,000 t (Gracia and Vázquez Bader, in press). Coincidentally, during this period the fishing effort of trawler boats also declined approximately 40%; however, an increase of inshore artisanal fishing occurred. In the last decade, shrimp production was below the maximum recorded, and a negative tendency has been observed for some species (e.g., white *P. setiferus*, Fig. 3; and pink *P. duorarum* shrimps, Fig. 4).

The white shrimp, *P. setiferus*, is mostly fished in the southwestern Gulf of Mexico, in the Campeche Sound and the adjacent coastal region, mainly the Términos Lagoon. This fishery reached its maximum produc-

Figure 1. Fishing areas of shrimp in the Mexican Gulf of Mexico.

tion levels during the 1970s when catches ranged from 1,200 to 2,200 t (total weight). After this period, white shrimp catches declined to levels below 20% of the average maximal production (Fig. 3). According to Gracia (1989a), the maximum sustainable yield of this species was estimated in 1,630 t with an optimal fishery effort of 33,406 ship days.

Traditionally pink shrimp, *P. duorarum*, was the most important species for shrimp fishery at the Mexican coast of the Gulf of Mexico. Catches obtained at the Campeche Sound during the 1970s represented between 70% to 90% of the total shrimp catch in the south of the Gulf of Mexico. After this period, catch and fishing effort decreased, leveling off from 1980 to 1986. Afterward, catches presented an acute negative tendency

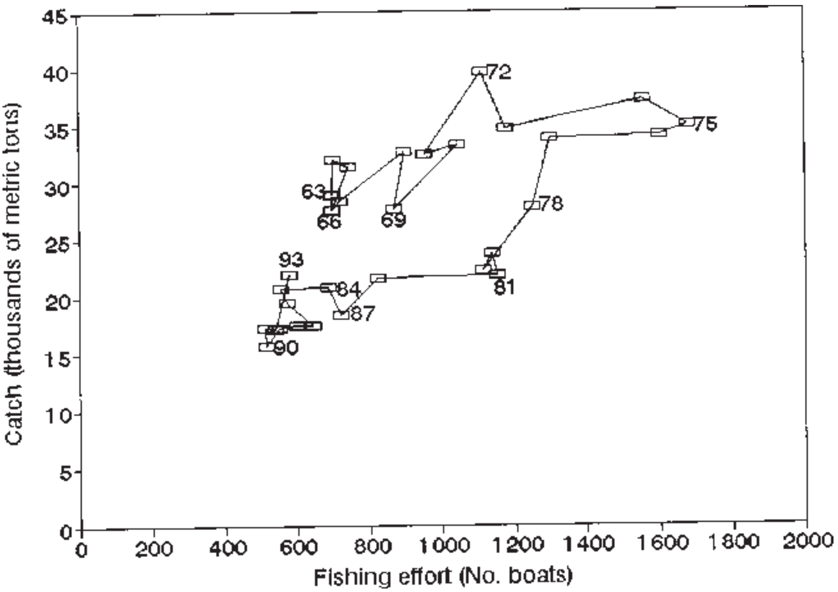


Figure 2. Diagram of catch and effort data of the shrimp fishery in the Gulf of Mexico (1963-1993).

that correlated with low abundance of the species. Pink shrimp yield recorded in one of the most important fishing ports (Carmen City) declined to almost 20% of the average maximal production reached in the 1970s (Fig. 4). According to estimations made by several authors (Arreguín-Sánchez and Chávez 1985, Navarrete and Uribe 1993), maximum sustainable yield varies from 4,300 to 11,000 t.

Artisanal fisheries target the three main shrimp species, *P. setiferus*, *P. duorarum* and *P. aztecus*, and include fishing in lagoons, estuaries, and coastal areas. These artisanal fisheries affect mainly juvenile stages, although recently a new artisanal fishery with drift nets was developed, which focuses mainly on the exploitation of the adult white shrimp in marine waters.

Material and Methods

Statistical data of the shrimp fisheries in the Gulf of Mexico were considered for the study. Analysis of shrimp industrial fisheries was based on catch and effort statistics for white and pink shrimp in the main fishing areas during 1973-1991 and 1973-1990, respectively. This information was available at government agencies (Fisheries Department of Mexico). Total catch per each of nine categories (10-15, 16-20...61-70 tails per pound)

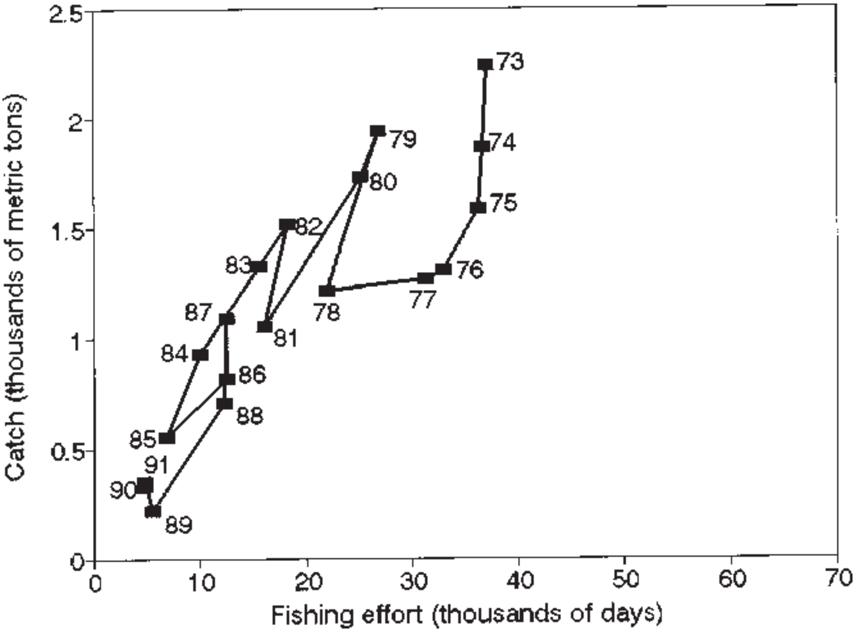


Figure 3. White shrimp fishery trajectory based on catch-effort data of Campeche Bank from 1973 to 1991 (redrawn from Gracia 1996).

by month were analyzed to estimate indices of shrimp abundances in open sea waters. Numbers of individuals by size and sex were estimated by the procedure of Gracia (1989a, 1991), based on sex ratio and the mean tail weight for each market category. Numbers were calculated by using a computer spreadsheet and assuming a normal distribution. Information of artisanal fishery total catch was also compiled (1985-1990), when available, from government statistics. Size composition and monthly shrimp abundance indices of the artisanal fishery were obtained from periodic samplings carried out in these fisheries (Gracia 1989a, 1991, 1995, 1996).

Population and fishery parameters were obtained from the literature or were estimated from the available statistics (Table 1). Size was converted to age according to the von Bertalanffy growth curve for each species. Age-specific fishing mortality was obtained through age-based virtual population analysis for white shrimp and pseudocohort analysis for pink shrimp of the composition of monthly shrimp catches (Sims 1985, Mesnil 1988). A mean life span of 16 months for the shrimp was considered based on size population structure and shrimp growth curves.

Analysis of the interaction among the diverse fisheries was based on a yield per recruit (Y/R) age-structured Thompson and Bell type model.

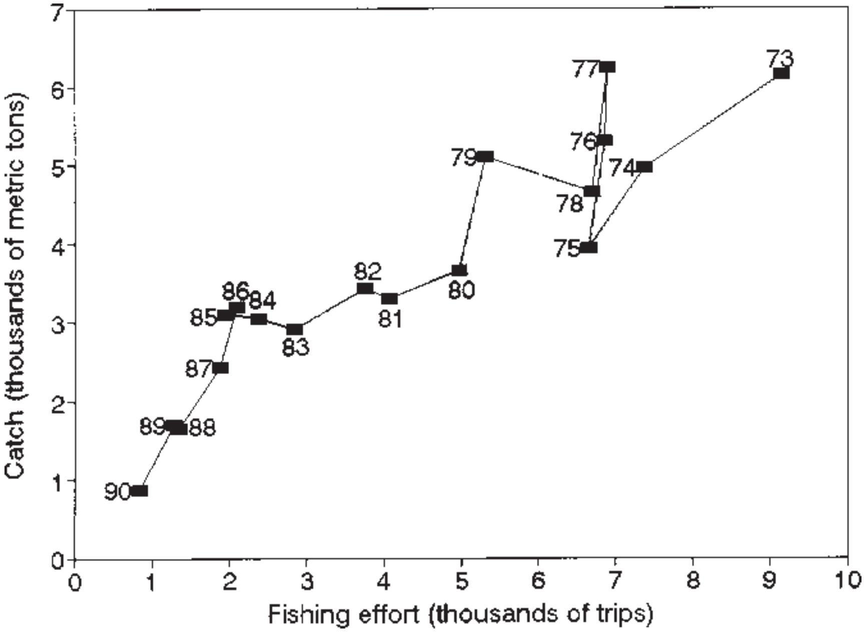


Figure 4. Pink shrimp fishery trajectory based on catch-effort data of Campeche Bank (1973-1990).

This model provides a simple description of non-steady state situations and has been used for sequential shrimp fisheries earlier by Willmann and García (1985). Basic equations used in the model are as follows:

$$Z_i = M + F_i$$

$$N(t_i + \mathrm{dt}) = N(t_i)\exp(-Z_i\mathrm{dt})$$

$$C_i = [N(t_i) - N(t_i + \mathrm{dt})]F_i/Z_i$$

$$Y_i = C_i w_i$$

$$B_i = Y_i/[F_i\mathrm{dt}]$$

Where Z_i is instantaneous coefficient of total mortality at age i, M is instantaneous natural mortality rate (monthly), F_i is instantaneous fishing mortality rate at age i, N_i is shrimp population size at the end of the age i, C_i is catch in numbers at age i, w_i is average shrimp weight, Y_i is yield, B_i is biomass, i is the age in months, t is time and dt is the time spent in the interval.

Table 1. White and pink shrimp parameters.

	White shrimp[a]	Pink shrimp[b]
Growth parameters		
L_α	209 mm	203 mm
W_α	72 g	68 g
K	0.2016	0.2167
t_0	0.3267	0.466
Estuarine natural mortality	0.6	0.7
Offshore natural mortality	0.25	0.27
Marine fishing mortality		
Age	F	F
4	0.012	0.018
5	0.063	0.053
6	0.090	0.089
7	0.178	0.125
8	0.245	0.225
9	0.492	0.338
10	0.355	0.388
11	0.409	0.393
12	0.412	0.445
13	0.386	0.464
14	0.454	0.306
15	0.364	0.366
16	0.330	0.315

Source: [a]Gracia 1989a, [b]this study.

For simplicity, it was assumed that recruitment, growth, and fishing pattern was constant during the life span of shrimps and were known without error. An average single-cohort, age-structured model was built for each species and was assumed that, at equilibrium, this yield was equivalent to that one resulting from all the cohorts present in the fishery during the year. To analyze the effect of inshore artisanal fisheries on the offshore industrial fishery different levels of artisanal mortality were used while maintaining offshore fishing rates constant. Simulation results were compared and expressed as percent change of the shrimp total yield.

Results and Discussion

White Shrimp

The proportion of white shrimp in inshore artisanal fisheries varies with the geographic location of the lagoons. According to fishery statistics, the white shrimp represents from 5% to 15% of the total artisanal shrimp catch in the coastal lagoons of the northeastern Gulf of Mexico. Juveniles are highly concentrated in Términos Lagoon in Campeche Sound (Gracia 1989a). This fact led to the development of fishery activities in this region during the 1970s. At present, this activity is prohibited but is performed clandestinely. Due to the illegal character of this activity, there are no formal statistical records of the effort and catch rates of white shrimp by the artisanal fishery; nonetheless, the available information allows analysis of the main aspects of its relationship to industrial fishery (Gracia 1989a).

Size composition of artisanal catch obtained using otter trawls ranges from 3.5 to 14.0 cm total length (TL), corresponding to 2- to 4-month-old organisms. However, the highest proportion is represented by organisms aged under 3 months with an average monthly size of 7.0 to 10.0 cm TL (Fig. 5) corresponding to a market category of over 70 tails per pound.

Artisanal catches vary with seasonal fluctuations in shrimp abundance in the nursery areas. According to Gracia (1989a), in Términos Lagoon they have high and low abundance periods. In general, the high abundance period goes from the end of spring to the fall (rainy season) and, on average, the maximal abundance of juveniles occurs between June and July. The lowest abundance usually occurs from December to March. The artisanal fishery adjusts to these variations; therefore the invested effort depicts proportional fluctuations to the availability of white shrimp in the nursery areas. These seasonal fluctuations in the abundance of juveniles and migrating sizes was taken into account for the simulation of interactions among white shrimp sequential fisheries based on the yield per recruitment model.

Analysis of the interaction among sequential fisheries in the Términos Lagoon–Campeche Sound area (illegal inshore–industrial offshore) based on population parameters, estimated for all stages of *P. setiferus* indicates that the decline in total production due to the illegal artisanal

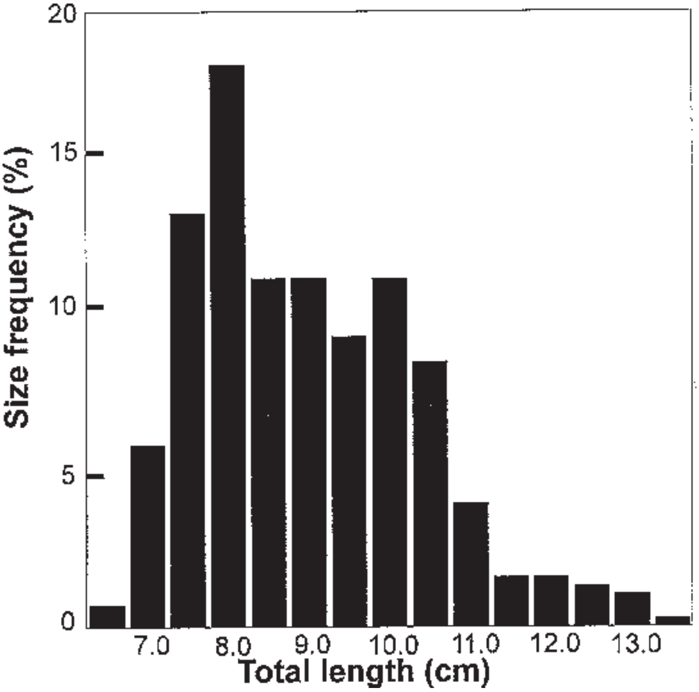


Figure 5. White shrimp size composition obtained by the inshore artisanal fishery using otter trawls.

fleet—estimated in 1984 to range from 30 to 300 boats (Smith 1985)—could be close to 20%. Based on this data, the estimated artisanal catch could only contribute between 3% to 6% of the total catch, resulting in a net reduction of shrimp production offshore. Although the lack of available statistics of artisanal fishing effort does not allow precise estimates, it is appreciated that the effect on total shrimp catch is directly related.

If the fishing effort in Términos Lagoon is increased to levels causing a monthly fishing mortality of 0.44-0.5, equivalent to using fixed fishing gears (Gracia and Soto 1986), the increase in artisanal catch would correspond to 14% of the total biomass, and the industrial fishery would decline 40% (Fig. 6). The effect is even more noticeable if it is analyzed in terms of the gross value of the catch, since the artisanal fishery catches smaller sizes of less market value. In this case gross value of the catch will reduce about 40% and the artisanal catch value will only represent about 2% of the total.

Interaction between growth rates and natural mortality, and even fishing mortality, results in a larger biomass of *P. setiferus* after migration,

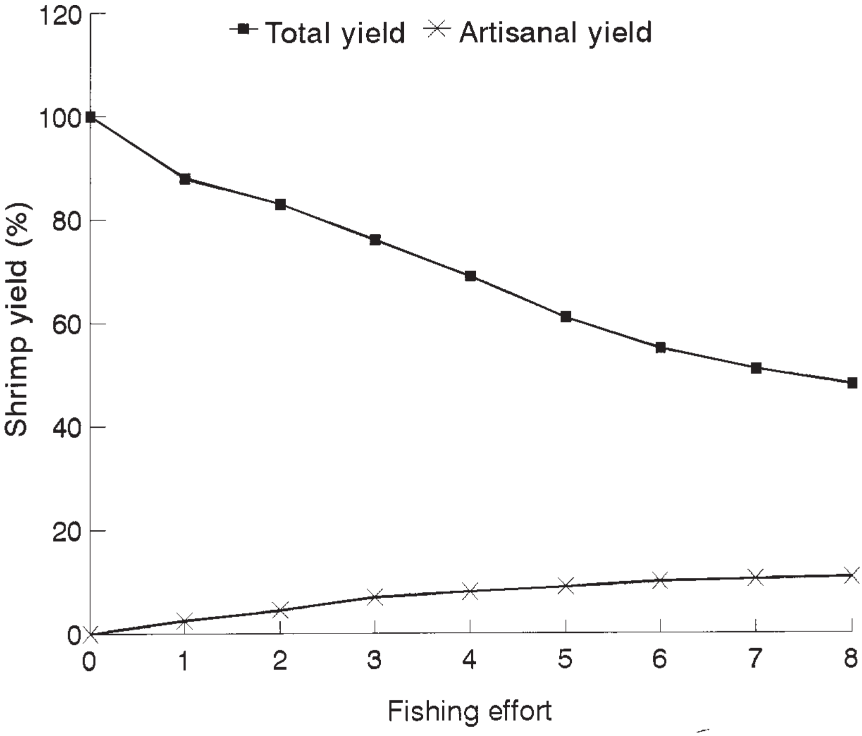


Figure 6. Impact of inshore artisanal fishery on the total yield of white shrimp.

once they escaped artisanal exploitation in the coastal lagoon environment. Each kilogram of white shrimp caught in the estuary decreases marine catch in a 1:2.8 proportion. The cumulative residual loss due to artisanal exploitation leads to a global reduction in the catch that correlates with increasing fishing effort.

Therefore, the magnitude of the impact caused by inshore artisanal fisheries on the catch levels of the following fisheries is directly proportional to the fishing effort exerted on the estuarine populations and has a negative exponential relationship with marine catch in terms of biomass (Fig. 11). If the number of organisms is considered this relationship is linear and directly proportional to the fishing effort.

Pink Shrimp

The greatest artisanal exploitation of pink shrimp is being taken in the protected coastal waters of the states of Campeche and Yucatán, which are the most important nursery areas of the pink shrimp in Mexican waters of the Gulf of Mexico.

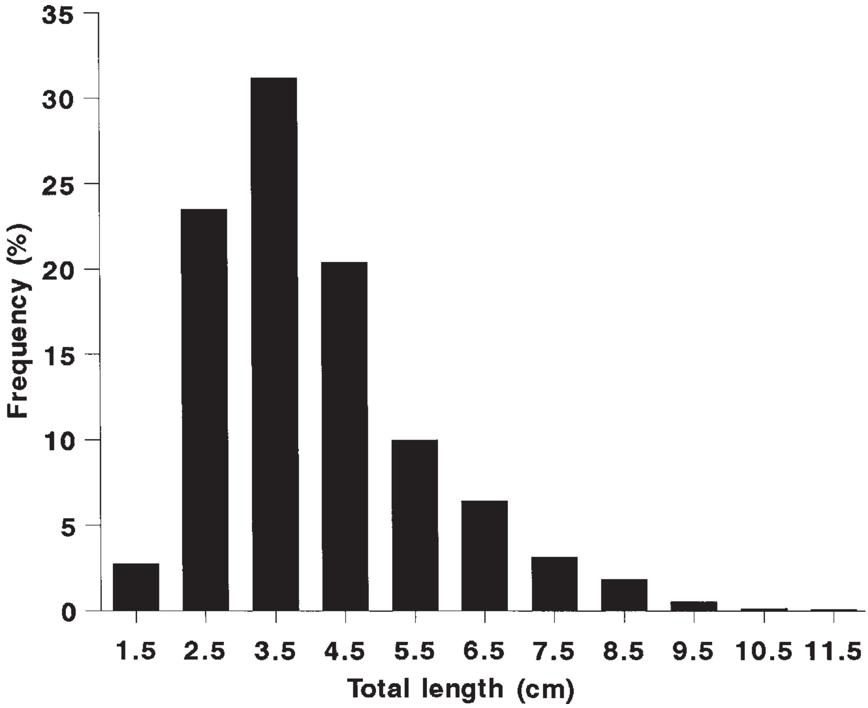


Figure 7. Pink shrimp size composition exploited by inshore artisanal fishery.

Artisanal fishery of pink shrimp in the coastal areas of Campeche is accomplished through the use of different types of fishing gears (Alarcón-Fuentes and Arreguín-Sánchez 1992). Two of them are fixed gears placed in the channels to catch shrimps during their migration to the sea. The other two are mobile and are operated manually in shallow waters near the coastline. Catches are basically made up of pink shrimp (more than 96%), although some palemonid shrimps have also been recorded in small proportion (4%).

The size range caught by this artisanal exploitation corresponds to very small shrimps, and although sizes go from 1.2 to 11.5 cm TL, the mode of the catch corresponds to 3.0-4.0 cm TL (Fig. 7). Most (85%) of the organisms caught through this artisanal fishery are only 25 to 60 days old (Fig. 7).

According to information gathered in the Campeche area, artisanal fishing of juvenile pink shrimp has been going on for more than 20 years, but exploitation used to be on a small scale just for household subsistence, i.e., catches were small and not recorded. At the beginning of the 1980s, exploitation increased to commercial levels and the first data of

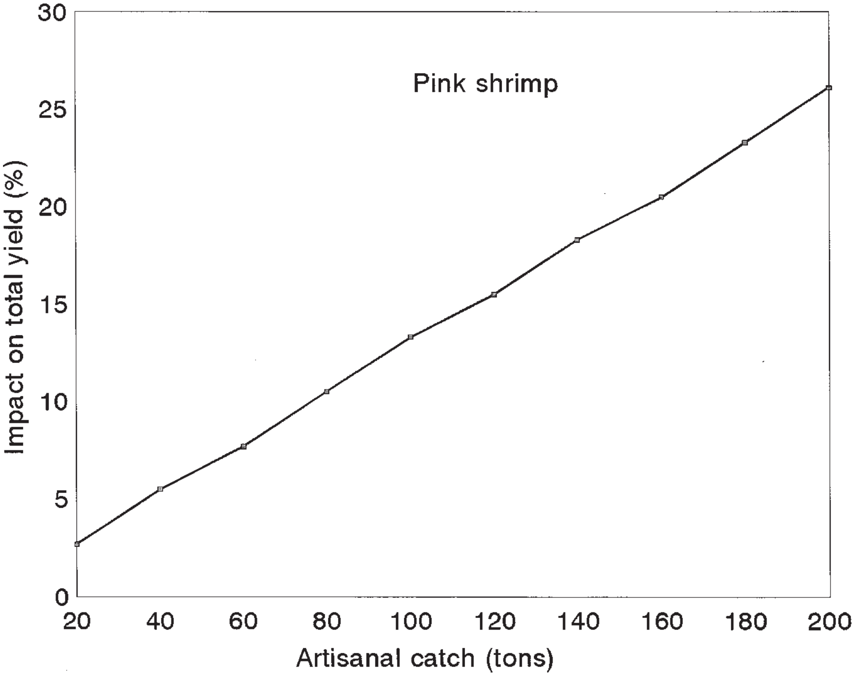


Figure 8. Impact of artisanal fishery on the pink shrimp total yield.

juvenile pink shrimp catches along the coast of Campeche was recorded. It is worthwhile mentioning that these records are variable and do not necessarily reflect the level of artisanal exploitation. The recorded catches of this juvenile stage revealed a steady increase from the first years of data recording until this activity was prohibited in 1992.

The average artisanal exploitation, according to the available statistical data, reveals a constant activity year-round with monthly variations that could coincide with abundance fluctuations of juveniles. The available records show that the maximum average catch was obtained in the months of March, May, September, and October, which agrees with periods of high abundance of juveniles in the area (Gracia and Soto 1990). However, it should be noted that no knowledge exists on the fluctuations in artisanal fishing effort that might introduce bias in the relationship of abundance-catch levels; for example, in the months of January and February, when shrimp abundance is low and juvenile catches do not follow this pattern (Gracia 1995).

The effect of artisanal fishery of juvenile pink shrimp on the industrial fishery is also directly related to the extracted volume (Fig. 8). The relationship is similar to that of the white shrimp, except that for the pink shrimp the impact is even greater on the total production. The loss of total biomass is greater since shrimp caught by the artisanal fishery are smaller. The exploitation of juvenile *P. duorarum* of such small sizes results in a waste of global biomass of 9 kg for each kilogram of juvenile organisms extracted from the nursery areas. In terms of commercial value, this proportion increases to an average of 28:1, since the value per unit weight increases with shrimp size.

The effect on the adult population, estimated through simulation models, indicates that a fishery activity similar to the one recorded could cause a 10%-20% decline in offshore production. The effect could be even greater if we consider that records of an artisanal fishery are usually incomplete. The annual production of artisanal pink shrimp catches in the nursery area of Yucatán state from 1985 to 1990 show fluctuations from 25 to 230 tons. Inclusion of precise catches in the model could account for most of the decline in pink shrimp production.

Other factors have been pointed out as causing the decline in pink shrimp production (Gracia 1995). Among them are (a) decrease in the fishing effort due to deterioration of the fishing fleet, (b) loss of fishing areas due to the expansion of the oil industry, (c) deterioration of the nursery areas, and (d) expansion of the artisanal fishery during the earlier stages of the pink shrimp. However, according to the results, the artisanal fishery seems to be the most important factor responsible for the decline in production. None of the other factors seem to be critical for the decline, since fishing effort has recovered in some cases without improving production, and the relationship between deterioration of nursery areas and decline in catches has not been proven.

Exploitation of Shrimp Postlarvae

At present there is no extraction of postlarval organisms to be used as “seeds” for shrimp culture in the Gulf of Mexico, since there is no aquaculture development with native species in the area. However, due to the problems posed by some diseases and the potential risk of exotic viruses transmitted by foreign species (NMFS 1997), culture of native species seems to be a good option for increasing shrimp production. Considering the potential involved in aquaculture, a postlarval fishery could be started with native species of the Gulf of Mexico. Nonetheless, this analysis could also be a reference point for other species, such as *P. vannamei* from the Pacific, where postlarvae are already being exploited.

Literature regarding the effect of postlarval extraction on shrimp population is scarce. Gracia (1989b) assessed such an activity for the white shrimp *P. setiferus*, a potential aquaculture species that occupies the third place in pond productivity after *P. vannamei* and *P. stylirostris* (Lawrence 1984). The impact varies according to the age of juveniles caught. The impact is lower when the postlarval organisms are caught before entering

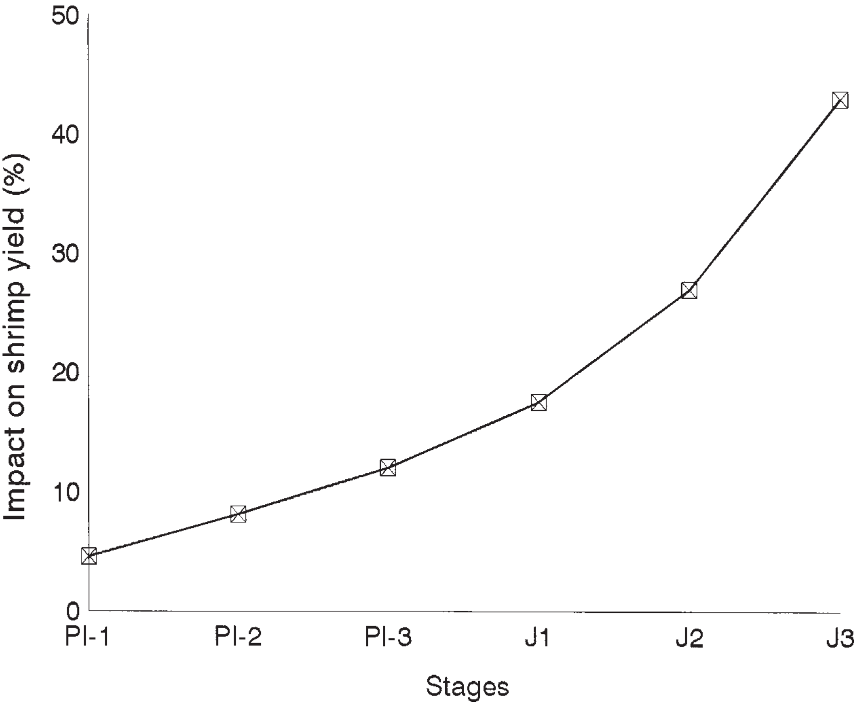


Figure 9. Impact of postlarvae exploitation of different ages on total white shrimp yield.

the coastal lagoons and increases exponentially with increasing age of the shrimp (Fig. 9). Indeed, catching of already settled postlarval-juvenile organisms in nursery areas is a variant of juvenile artisanal fishing as described for the pink and white shrimps with similar effects.

The impact on shrimp production increases proportionally with the exploitation level on postlarvae (Fig. 10). However, there is a marked difference if the catch is performed before the settlement of postlarval organisms in the estuarine benthos, or if the catch includes already established stages. This difference is due to the fact that postlarval settlement in nursery areas represents a critical point in the life history of the shrimp and in the regulation of the population size. The carrying capacity of the nursery areas, determined by diverse environmental factors (type of bottom, nutrients, salinity, depth, vegetation, etc.), is to a great extent responsible for the recruitment level in marine populations (García 1989). This explains why the curve slope elevation of postlarval exploitation-production impact is greater when organisms that have already surpassed this critical stage are exploited.

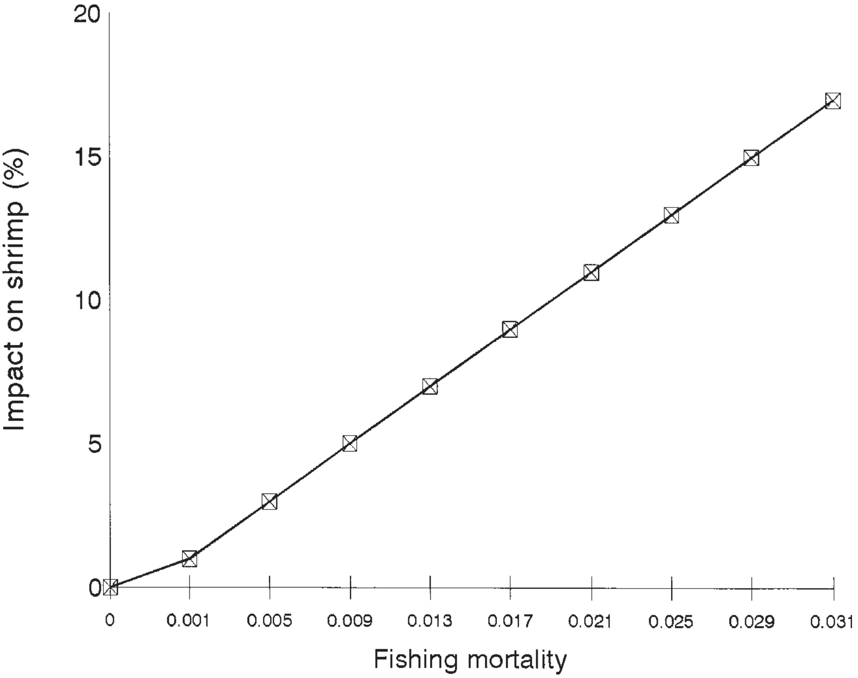


Figure 10. Relationship between postlarvae fishing mortality and shrimp yield decline.

Relationship among the Different Life Cycle Stages

White shrimp and pink shrimp have a similar life history. Both of them have an estuarine-dependent stage. Differences lie in the time they spend in the nursery areas. Pink shrimp tend to migrate out of the lagoons sooner than white shrimp. However, similarities allow common management strategies for both.

According to the performed analyses and the recorded data in the literature (Gracia 1989a, 1989b, 1991, 1997) it can be assumed that the relationships among the different stages of one shrimp generation subjected to a potential fishery vary between potential and linear, as shown in Fig. 11. In the presented model, it is assumed that there is a linear relationship between the number of spawners and the number of produced larvae. However, when relating spawners to settled postlarvae two weeks later (8-12 mm total length), a potential relationship is expected. Density dependent mortality mainly occurs during the estuarine stage, therefore the relationship between larvae and juveniles would be potential. The relationship between the number of resulting juveniles and the number of

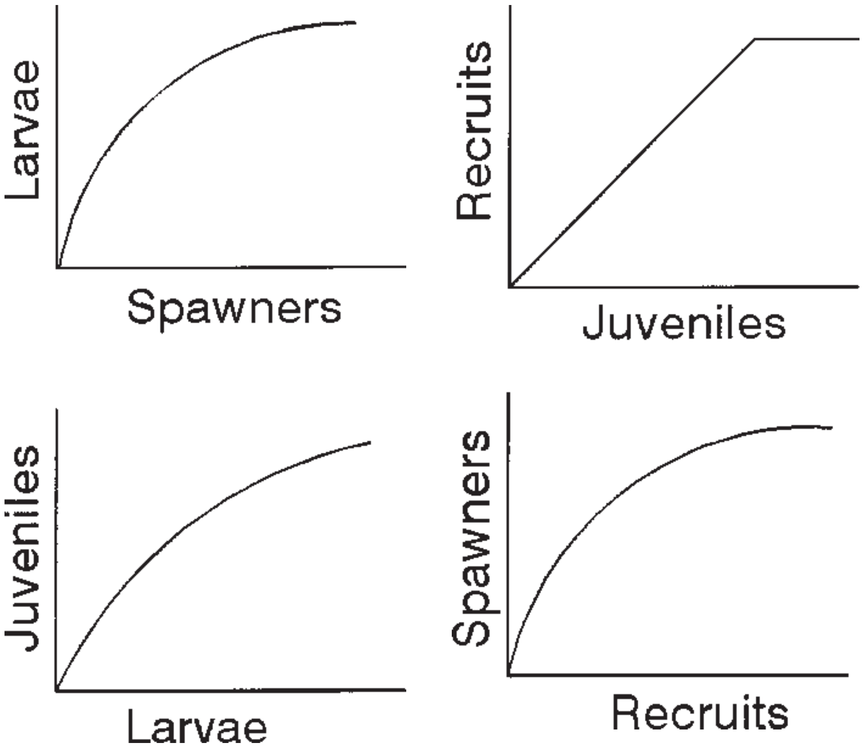


Figure 11. Relationship among the different shrimp life cycle stages.

recruits to the marine population is linear, although a limit would be expected for this relationship at very high abundance levels of juveniles in the nursery areas. The duration of this stage varies between an average of 2-3 months (6.5-8.5 cm TL) for pink shrimp and 3-4 months (8.5-9.5 cm TL) for white shrimp. In the marine environment the relationship between the number of recruits (3-4 months old) and spawners would have a potential form due to density-dependent factors. A stock-recruitment relationship of this type has been established for white shrimp by Gracia (1991).

These relationships can be affected by environmental changes, and could be represented by diverse levels on the curves limiting an area where different abundance values for each stage can be encountered. These values are influenced by the fishing mortality derived from the fishery over the different stages of the life cycle, and its impact is reflected on the subsequent stages, and eventually on the total production of shrimp.

General Discussion

Along the history of shrimp fisheries, small-scale artisanal fisheries were conducted before the industrial fishery developed, since exploiting this resource in coastal regions was easy. Profitability of the industrial fishery caused a shift of fishing efforts toward offshore fishing. During the 1980s, artisanal fishery efforts on several species of the Gulf of Mexico strongly increased, which in some instances even inverted the relationship of artisanal versus industrial fishing effort.

This increase in artisanal fishery effort directed to early shrimp stages of those species of greater market value caused overfishing at this growth stage, since exploitation occurs before the optimal biomass per organism is achieved. On the other hand, increases in artisanal fishery with drift nets has caused a recruitment overfishing of white shrimp (Gracia 1992a,b, 1996). Both factors led to a decline in total available biomass of the three species, affecting the optimal profitability of the resource and the reproductive potential of the species.

The critical problem in defining precisely the impact of artisanal fisheries on the exploitation of the main shrimp species and to differentiate this impact from other factors (such as environmental effects or deterioration of the nursery areas) is the lack of adequate records on artisanal fishing effort. Despite their importance, artisanal fishing effort is not given their due weight in assessing shrimp populations in the Gulf of Mexico. Information available at this moment suggests that most of the deterioration in shrimp populations is due to the uncontrolled increase in artisanal fisheries plus the already existing industrial fishery effort, which has already reached maximum allowable levels.

Fishing mortality in the juvenile and pre-adult stages must be added to the fishing mortality caused by industrial fishing, on which evaluations of shrimp stocks have traditionally been based. As a result, we now have total fishing efforts that had never been reached before in the history of shrimp exploitation. Disregard of the artisanal effort also hinders evaluation of the total accumulated fishing efforts in each of the exploitation stages, and impairs taking adequate measures to insure a healthy management of this resource.

The main characteristic of the artisanal fishery is in providing income to a relatively large number of fishermen with low-cost investments. This makes this activity quite attractive and fostered an uncontrolled increase in fishing effort. To a certain point, Mexican artisanal fisheries in the Gulf of Mexico are almost freely accessed without any control adversely affecting the resource. The established objective for shrimp fishing in the area are (1) that there must be a healthy balance between artisanal and industrial fisheries to allow for optimal exploitation of the resource without negatively affecting the reproductive population, and (2) to generate a global benefit. In this sense, it is necessary to limit the artisanal effort. This is particularly important for the present conditions of the resource, where (1) there is no possibility to increase catches of natural populations above the already reached levels, (2) global fishing effort have increased markedly, and (3) world aquaculture is in an expansion stage, which will increase supply of small shrimps competing with the shrimp obtained by artisanal fisheries.

An important point to be considered is the impact caused by all fisheries on the reproductive potential. According to the analyses made, the recruitment level can be affected, which in turn will affect the reproductive biomass. Garcia and Van Zalinge (1982) and Garcia (1989) emphasize that fecundity per recruit at a given exploitation level can be increased by protecting the juvenile population. Accumulation of fishing efforts on different stages of the shrimp's life cycle can exert an opposite effect and eventually lead to recruitment overfishing.

In most worldwide fisheries, recruitment overfishing has been questioned, since stock-recruitment (S-R) relationships are not apparent for most penaeid shrimps. However, these S-R relationships have been demonstrated in some penaeid species (Penn and Caputi 1986, Gracia 1991) and recruitment overfishing has been documented in some species (e.g., in *P. setiferus* resulting from the additional artisanal effort on reproductive populations, Gracia 1996).

Traditional management to prevent recruitment overfishing consists of maintaining a healthy relationship between the spawner stock and the recruitment produced by this stock, considering the different stages of the shrimp's life cycle as a black box. Such is the case for the white and pink shrimp in the Gulf of Mexico, although some protection is aimed to the juvenile stage. According to Hilborn and Walters (1992) the reasons to group the different life stages are (1) the spawning stock is the one that can be managed and not the other stages of the life cycle, and (2) it is important to have direct measurements of the consequences of this control. Although management through such a general relationship can be very useful, in the case of penaeid shrimps the multiple fishery conditions can lead to recurrent stock reductions, shortening of the fishing season, and/or depletion of the fishery. An increasing effort in artisanal fishery would tend to decrease the spawning stock independently of the protection measures taken at this stage. Reducing fishing effort only on spawners would affect mainly the offshore industrial fishery, which is at the end of the exploitation chain, and the results would not be very tangible. Eventually, the fishery could reach levels of smaller production due to growth overfishing and, if the fishing effort is excessive, become unprofitable at the marine stage.

To avoid damage to the fishery, it is necessary to implement adequate management strategies based on escape levels in each stage. Gracia (1996), based on a study in white shrimp, suggests some practical reference points for the management of the fishery in relation to the critical levels of the reproductive stock. In this case management strategies can be directed especially to the main annual cohorts. Based on the analysis of

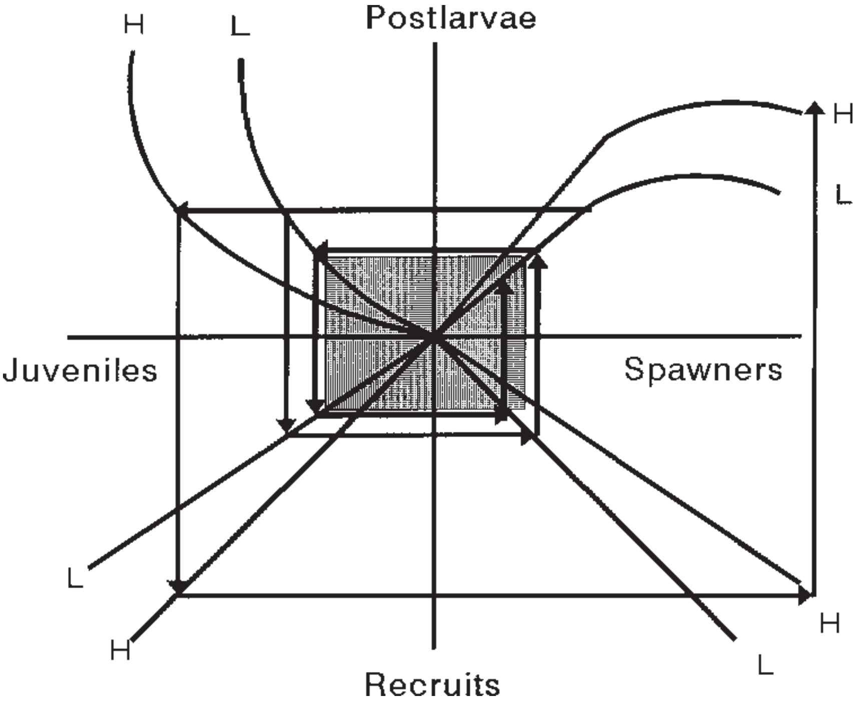


Figure 12. Base diagram to analyze the trajectory of one shrimp generation subjected to exploitation in its different life cycle stages. Shadow area represents the risk area for population renewal, H and L are "good" and "bad" environmental conditions. Arrows represent hypothetical population trajectories along varying conditions.

the S-R relation of the main *P. setiferus* cohorts, Gracia (1996) suggests that to preserve the population lying on the flat section of the stock-recruitment curve, the stock must be maintained at 25%-30% of the highest reproductive biomass value recorded in the fishery. This means that the spawning stock must be kept at levels higher than 17%-20% of virgin biomass.

These reference points can be linked with information on the different stages (Fig. 12) of the shrimp's life cycle through a multiple stage model as proposed by Paulik (1973). To design management strategies based on this type of models, it is necessary to have information on recruitment both in the growth areas and in the marine environment, and on effort levels in the artisanal and industrial fisheries, as well as to have critical reference points on the stock-recruitment relationship. Basically, escape of organisms in the different stages can be insured through catching quotas, reducing fishing effort by setting spatial and temporal closures, or monitoring catch per effort unit in the diverse fisheries. This should aim to leave a sufficient number of migrating organisms to allow for an adequate level of spawners and, thus, avoid deterioration or collapse of recruitment. Such a management strategy could be applied not only to penaeid shrimp of the Gulf of Mexico, but to the those which are subjected to similar exploitation levels. Establishing an escapement-based management strategy may, in first instance, increase total shrimp yield and gross fishery value by reducing growth overfishing, which would mainly benefit offshore fishery. An example of this can be found in the brown shrimp seasonal closure of the Gulf of Mexico directed to reduce growth overfishing which enhanced cohort yield by more than 300% (Gracia 1997). Benefits to the artisanal fisheries could be expected in the long term by increasing shrimp spawning potential and increasing average shrimp size in the nursery areas. However, the most important point will be that this strategy could allow the coexistence of the different fisheries while assuring high profitability in their exploitation.

Acknowledgments

The valuable help of M.A. Gómez Ponce in the development of this work is gratefully acknowledged.

References

Alarcón-Fuentes, T., and F. Arreguín-Sánchez. 1992. La pesca de camarón rosado en Sabancuy y Champotón, Campeche. JAINA Boletín Informativo 3(4):20.

Arreguín-Sánchez, F., and E.A. Chávez. 1985. Estado del Conocimiento de las pesquerías de camarón en el Golfo de México. Investig. Mar. CICIMAR 2(2):23-44.

García, S. 1989. The management of coastal penaeid shrimp fisheries. In: J.F. Caddy (ed.), Marine invertebrate fisheries: Their assessment and management. Wiley and Sons, New York, pp. 281-306.

García, S., and N.P. Van Zalinge. 1982. Shrimp fishing in Kuwait: Methodology for a joint analysis of artisanal and industrial fisheries. In: Assessment of the shrimp stocks off the west coast of the gulf between Iran and the Arabian Peninsula. FAO/UNDP Proj. FI/DP/RAB/80/015, pp. 119-142.

Gracia, A. 1989a. Ecología y pesquería del camarón blanco *Penaeus setiferus* (Linnaeus, 1767) en la Laguna de Términos-Sonda de Campeche. Tesis doctoral, Universidad Nacional Autónoma de México. 127 pp.

Gracia, A. 1989b. Impacto de la explotación de postlarvas sobre la pesquería del camarón blanco *Penaeus setiferus* (Linnaeus, 1767). An. Inst. Cienc. del Mar y Limnol., Universidad Nacional Autónoma de México 16(2):255-262.

Gracia, A. 1991. Spawning stock-recruitment relationships of white shrimp in the southwestern Gulf of México. Trans. Am. Fish. Soc. 120:519-527.

Gracia, A. 1992a. Explotación y manejo del recurso camarón. Ciencia y Desarrollo 18(106):82-95.

Gracia, A. 1992b. La sobrepesca del reclutamiento en las pesquerías de camarón. JAINA Boletín Informativo 3(1):11.

Gracia, A. 1995. Impacto de la pesca artesanal sobre la producción de camarón rosado *Penaeus farfantepenaeus duorarum* Burkenroad 1939. Cienc. Mar. 21(3):343-359.

Gracia, A. 1996. White shrimp *(Penaeus setiferus)* recruitment overfishing. Mar. Freshw. Res. 47:59-65.

Gracia, A. 1997. Simulated and actual effects of brown shrimp *Penaeus aztecus* closure in Mexico. Marine Fisheries Review 59(2):18-24.

Gracia, A., and L.A. Soto. 1986. Estimación del tamaño de la población, crecimiento y mortalidad de los juveniles de *Penaeus setiferus* (Linnaeus, 1767) mediante marcado-recaptura en la Laguna de Chacahito, Campeche, México. An. Inst. Cienc. del Mar y Limnol., Universidad Nacional Autónoma de México 13(3):217-230.

Gracia, A., and L.A. Soto. 1990. Population study of the penaeid shrimp of Términos Lagoon, Campeche, México. An. Inst. Cienc. del Mar y Limnol., Universidad Nacional Autónoma de México 17(2):241-255.

Gracia, A., and A.R. Vázquez Bader. In press. Shrimp fisheries in the south of the Gulf of Mexico: Present state and future management alternatives. In: The Gulf of Mexico large marine ecosystem: Assessment, sustainability and management. Blackwell Scientific Publications, Oxford.

Hilborn, R., and C.J. Walters. 1992. Quantitative fisheries stock assessment: Choice, dynamics and uncertainty. Chapman and Hall, New York. 570 pp.

Lawrence, A. 1984. Marine shrimp culture in the Western Hemisphere. In: P.C. Rothlisberg, B.J. Hill, and D.J. Staples (eds.), Second Australian National Prawn Seminar, Cleveland, Australia, pp. 327-336.

Mesnil, B. 1988. Logiciels pour l'évaluation des stocks de poissons. ANACO: Logiciel d'analyse des données de capture par classe d'age sur IBM PC et compatibles. FAO Doc. Tech. Pêches 101(Suppl. 3):1-73.

NMFS. 1997. An evaluation of potential shrimp virus impacts on cultured shrimp and wild shrimp populations in the Gulf of Mexico and Southeastern U.S. Atlantic Coastal waters. A report to the Joint Subcommittee on Aquaculture prepared by the JSA shrimp virus work group. NMFS-APHIS-NCEA-USFWS. 65 pp.

Navarrete, A., and J. Uribe. 1993. Evaluación de la pesquería industrial del camarón de altamar del Puerto de Campeche, México en el período de 1981 a 1990. Centro Regional de Investigaciones Pesqueras de Campeche, Camp., Instituto Nacional de la Pesca, pp. 1-15.

Paulik, G.J. 1973. Studies of the possible form of the stock and recruitment curve. Rapp. P.-V. Reun. Cons. Int. Explor. Mer 164:302-315.

Penn, J.W., and N. Caputi. 1986. Spawning stock-recruitment relationships and environmental influences on the tiger prawn *(Penaeus esculentus)* fishery in Exmouth Gulf, Western Australia. Aust. J. Mar. Freshw. Res. 37:49-51.

Sims, E. 1985. Selected computer programs in FORTRAN for fish stock assessment. FAO Fish. Tech. Pap. 259:1-189.

Smith, K. 1985. Investigación de pesquerías clandestinas y de las zonas y épocas de reproducción del camarón blanco en la zona de Cd. del Carmen. Centro Regional de Investigaciones Pesqueras de Cd. del Carmen, Instituto Nacional de la Pesca, SEPES, pp. 1-17.

Willmann, R., and S.M. García. 1985. A bioeconomic model for the analysis of sequential artisanal and industrial fisheries for tropical shrimp (with a case study on Suriname shrimp fisheries). FAO Fish. Tech. Pap. 270:1-49.

Fisheries Management Based on Reproductive Potential: An Example for Chub Mackerel

Toshio Katsukawa and Yoshiharu Matsumiya
University of Tokyo, Ocean Research Institute, Tokyo, Japan

Abstract

One of the main purposes of stock assessment is to evaluate the spawning stock size. Spawning stock biomass (SSB) has been generally used as an indicator of spawning stock size. However, SSB is a short-term indicator of spawning stock size, which represents the magnitude of the spawning in the coming spawning season. Because SSB cannot account for the long-term reproductive ability of the stock, fisheries management based on SSB may result in overfishing in the future. Katsukawa (1997) suggested the use of total reproductive value of the standing stock as an indicator of spawning stock size. We call this the reproductive potential. The reproductive potential can be used to evaluate the long-term reproductive ability of the standing stock, considering the biological parameters and the age composition. In this study, we have examined the performance of a fisheries management strategy based on reproductive potential. We compared the performance of constant reproductive potential strategy and that of constant SSB strategy. By deterministic analysis, we showed that fisheries management based on reproductive potential is more robust to the errors in the stock size than that based on SSB. A stochastic computer simulation also supports this result.

Introduction

One of the main purposes of fisheries management is to avoid recruitment overfishing. Spawning stock biomass (SSB) has been generally used as an indicator of the spawning stock size. Many reference points in fisheries management are based on SSB. For example, a level of 20% virgin or unexploited biomass has historically been adopted as a threshold for recruitment overfishing (Beddington and Cooke 1983). Myers et al. (1994)

and Zheng et al. (1993) examined a variety of methods for estimating the spawning stock biomass level as a threshold for recruitment overfishing.

Although a large number of studies has been done on the level of SSB needed to avoid recruitment overfishing, SSB presents some problems as an indicator of spawning stock size, especially for long-lived species. SSB is an indicator of short-term reproductive ability, representing only the magnitude of spawning in the coming spawning season. For long-lived species, the standing stock must sustain not only the next spawning but also the spawning of subsequent years. Ensuring sufficient SSB may be useful for avoiding recruitment overfishing in the coming spawning season. However, as SSB cannot account for spawning beyond the upcoming spawning season, a fisheries management strategy based on SSB may result in overfishing in the future. For example, immature fish have some potential value due to their future reproduction, which is not accounted for when using SSB.

In order to ensure the potential spawning ability of the stock, we must take into account the value of future spawning. Katsukawa (1997) developed an indicator of long-term spawning ability of the standing stock, called reproductive potential (*RP*). The objective of this paper is to examine the effectiveness of reproductive potential as an reference point of fisheries management. We examined the difference between reproductive potential and SSB by using an example of chub mackerel. We also compared the performance of fisheries management strategies based on *RP* (constant *RP*) and SSB (constant SSB).

Methods

Fisher's Reproductive Value

We could use the concept of Fisher's reproductive value (*RV*) for effectively evaluating the long-term productivity of the individual (Fisher 1930). *RV* can be expressed as follows:

$$\begin{aligned} RV_t &= \sum_{i=t+1}^{\infty} e^{-r(i-t)} E_i A_i \\ &= \sum_{i=t+1}^{\infty} \left[e^{-r(i-t)} E_i \exp\left\{ -\sum_{j=t}^{i-1} \left(F_j + M_j \right) \right\} \right] \end{aligned} \tag{1}$$

where RV_t is reproductive value of the individual at age t (just after the spawning season of age-t), E_i is fecundity of the individual at age i, A_i is survival rate till the spawning season at age i, F_j is the coefficient of fishing mortality at age j, and M_j is the coefficient of natural mortality at age j, and r is the instantaneous growth rate of the population size.

The instantaneous growth rate r is important when the stock size is exponentially increasing or decreasing. For example, when we think about

Table 1. Estimation of body weight (g), rate of maturity, selectivity, reproductive value (*RV*) (g), *RV* per weight by age, and stock number-at-age in 1970 (10^5 individuals) of chub mackerel as summarized from Wada et al. (1996).

Age	Body weight (w_i)	Rate of maturity (m_i)	Selectivity (q_i)	*RV* $f = 0$	*RV* $f = 1.73$	*RV* per weight $f = 0$	*RV* per weight $f = 1.73$	Number-at-age in 1970
1	252	0.0	0.25	1,120	147	4.29	0.583	52,766
2	434	0.2	0.44	1,584	251	3.55	0.578	31,833
3	610	1.0	0.62	1,753	191	2.80	0.313	14,411
4	672	1.0	0.79	1,944	160	2.83	0.238	5,370
5	811	1.0	1.00	2,089	126	2.53	0.155	3,328
6	912	1.0	1.00	2,204	146	2.37	0.160	2,560
7+	1,084	1.0	1.00	2,204	146	2.00	0.135	838

a stock rebuilding program, spawning of this year is more valuable than that of the next year. In this paper, we will discuss the management strategy to fix stock size at one level. Therefore, we assumed $r = 1$. In this case, *RV* of an age-t individual is the expected spawning of that individual from age t until the end of the individual's life.

Comparison between Reproductive Value and Body Weight

In order to examine if we can use body weight as an approximation of *RV*, we compared the reproductive value (*RV*) and the body weight of chub mackerel *(Scomber japonicus)*. The population parameters for chub mackerel were described by Wada et al. (1996) and they are summarized in Table 1. We assumed the natural mortality (M) = 0.4. The fishing mortality at age i (F_i) is expressed as the product of the selectivity at age i (q_i) and fishing effort (f) [$F_i = q_i f$]. The fecundity at age i (E_i) is expressed as the product of the rate of the maturity at age i (m_i) and the weight at age i (w_i) [$E_i = m_i w_i$].

When $f = 0$, *RV* and the body weight increase with age. As the increase in *RV* is much slower than that of the body weight, younger individuals have higher *RV* per body weight than older ones. This implies that young fish have higher future productivity than old fish with the same weight. If there is no fishing mortality, 1 kg of one-year-old chub mackerel will, with respect to long-term spawning, have an equal contribution as a 2.2 kg seven-year-old chub mackerel. Therefore, the reproductive ability of young fish is underestimated if we use biomass as an indicator of spawning stock size.

In Table 1, we show the *RV* corresponding to the average fishing effort between 1980-1989 ($f = 1.73$). An increase in fishing effort leads to a decrease in *RV*. The decreasing rate of *RV* is different for different ages. Because the selectivity of old individuals is higher than that of young ones, old individuals are generally more vulnerable to fishing pressure than young ones. One-year-old individuals are the only exception. Only 20% of two-year-olds are mature, and one-year-olds therefore take yet another year to fully mature. If the fishing pressure is high, the one-year-olds hardly survive next two years. Although one-year-olds have higher growth rate and lower fishing mortality than two-year-olds, they are more sensitive to fishing pressure. One-year-old individuals are as sensitive as three-year-old individuals.

Because the future production of the individual is affected by the age and fishing pressure, body weight is not a sufficient indicator to evaluate the future production of the individual. If we have enough biological information, the future production of an individual should be estimated by *RV*.

Reproductive Potential

Katsukawa (1997) suggested to use total *RV* of the standing stock as an indicator of spawning stock size. We have called this the reproductive potential (*RP*). *RP* can be expressed as follows:

$$\begin{aligned} RP &= \sum_{i=1}^{\infty} N_i RV_i \\ &= \sum_{i=1}^{\infty} \left\langle N_i \sum_{j=i+1}^{\infty} \left[\mathrm{e}^{-r(j-i)}\, E_i \exp\left\{ -\sum_{k=i}^{j-1} \left(F_k + M_k \right) \right\} \right] \right\rangle \qquad (2) \\ &= \sum_{i=1}^{\infty} \left\langle N_i \sum_{j=i+1}^{\infty} \left[\mathrm{e}^{-r(j-i)}\, E_j \exp\left\{ -\sum_{k=i}^{j-1} \left(q_k f + M_k \right) \right\} \right] \right\rangle \end{aligned}$$

where N_i is stock number of age *i*, q_k is selectivity at age *k*, and *f* is fishing effort. The reproductive potential can be used to evaluate the future production of the standing stock, considering the biological parameters and the age-composition of the stock.

We can use reproductive potential not only as an indicator for stock assessment but also as a reference point for fisheries management. If the number-at-age (N_i), fecundity (E_j), selectivity (q_k), and natural mortality (M_k) are estimated, reproductive potential (*RP*) can be expressed as a monotonic decreasing function of fishing effort (*f*). Therefore, we can change the value of *RP* by adjusting *f*. The maximum Reproductive potential is obtained when $f = 0$. As *f* increases, *RP* decreases toward zero. Consequently, we can obtain any *RP* value between $RP_{f=0}$ (the maximum *RP*) and zero by controlling *f*. Therefore, targeting a certain *RP* level will show us the *f* value needed to obtain that *RP* level. If the target level is higher than $RP_{f=0}$, no yield can be allowed because stock size is too low to achieve the target level. Otherwise, we can calculate the unique *f* needed to obtain the target reproductive potential level from equation (2). When the *RP* threshold is set, the *f* needed to obtain that threshold can be considered as the maximum acceptable *f*.

Let us think about the constant reproductive potential strategy (CRP). We used chub mackerel population as an example. We fixed the age-distribution as that in 1970. The stock size (α) is expressed relative to the stock size in 1970. Namely number-at-age *i* ($N_{i,t}$) equals $\alpha N_{i,1970}$. We employed relative stock size of $\alpha = 1$ (stock size in 1970) as the minimum stock size needed to reach the target level. In this case, fishing effort (f) is determined from the *RP* of the standing stock equal to the target level ($RP_{f=0,1970}$ = 1.55 × 10^7 t).

$$\begin{aligned}\sum_{i=1}^{7+}\left(N_{i,t}RV_i\right) &= \sum_{i=1}^{7+}\left(\alpha N_{i,1970}\sum_{j=i+1}^{\infty}\left[W_j m_j \exp\left\{-\sum_{k=i}^{j-1}\left(q_k f + M\right)\right\}\right]\right) \\ &= 1.55\times 10^7\,\mathrm{t}\end{aligned} \tag{3}$$

where $N_{i,t}$ is the number of age-*i* individuals in year *t*, α is the relative stock size $N_{i,t}/N_{i,1970}$, $N_{i,1970}$ is the number of age-*i* individuals in 1970. Figure 1 shows the relationship between *f* and relative stock size (α).

Figure 1 also contains *f* determined by the constant escapement strategy based on SSB (CSB). The target level of SSB is determined as the stock size in 1970 ($\alpha = 1$). The *f* of the CSB is determined as the SSB equals to the target level.

$$\sum_{i=1}^{7+}\left\{\alpha N_{i,1970}W_{i+1}m_{i+1}\exp\left(-q_i f - M\right)\right\} = 3.00\times 10^6\,\mathrm{t} \tag{4}$$

No fishing is allowed when $\alpha \le 1$, because stock size is too low to achieve the target stock size. On the other hand, when $\alpha > 1$, we can obtain an unique *f* that reaches the target level. The higher *f* is allowed as the stock size increases. When the relative stock sizes (α) is the same between CSB and CRP, the *f* of CSB is always higher than that of CRP. This is due to the difference in the time scale of the indicators. For the calculation of *RP*, we use the *f* value in the subsequent years. *RP* is calculated on the assumption that the stock will be exploited under the same fishing pressure throughout the remainder of its lifetime. By this assumption, we can determine the fishing pressure (f) that considers the future spawning of the standing stock. On the other hand, CSB allows exploitation of all surplus stock by the next spawning season. Let us think about a stock with low natural mortality. If the stock size is double of the target, CSB allows exploiting almost half of standing stock within one year. This exploitation

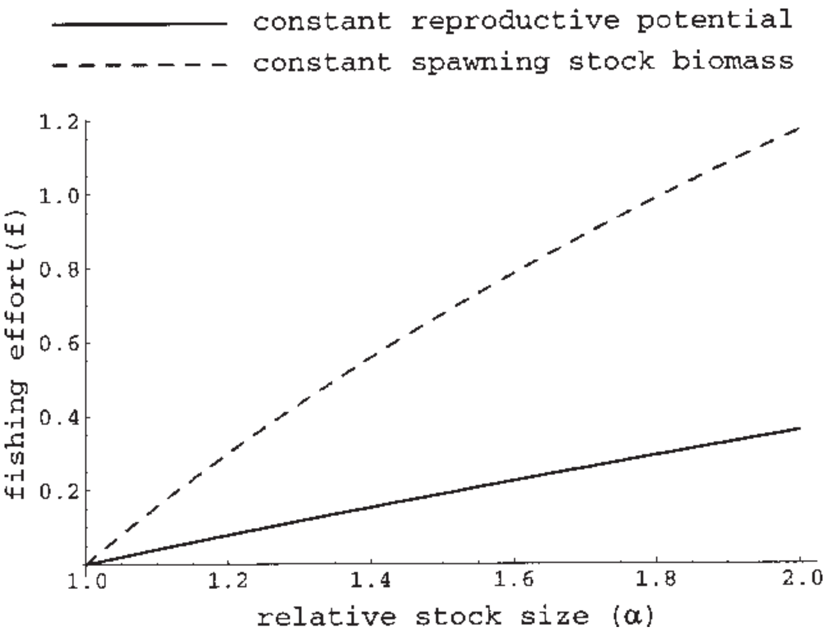


Figure 1. The fishing effort (f) determined by constant reproductive potential strategy and constant SSB strategy, for chub mackerel. We fixed age composition as that in 1970. Stock size (α) is expressed as the ratio to the stock size in 1970. We employed relative stock size of α = 1 (stock size in 1970) as the minimum stock size needed to reach the target level.

rate will decrease the stock size one forth in two years and one eighth in three years. Therefore, the future spawning of the standing stock is decreased to less than half. The *f* determined by CSB is not allowable from the viewpoint of long term productivity.

Simulation

By using a computer simulation, we compared the performances of the constant reproductive potential strategy (CRP) and the constant SSB strategy (CSB). We used chub mackerel for our simulation. This population is known to have considerable recruitment variation. Therefore, age composition is unstable. We can consider the age composition in light of reproductive potential. Since this stock is important for the Japanese fishing industry, the information needed for the calculation of *RP* is available.

We used a Beverton-Holt stock-recruitment relationship. The parameters of the Beverton-Holt curve were estimated from the stock-recruitment relationship during the period 1970-1991, by the maximum likelihood estimation to give

$$R = \frac{1.16 \times 10^{14} S}{5.32 \times 10^{15} + S} e^{N(0,1.01)} \tag{5}$$

where R is the number recruits (10^5 individuals), S is SSB (tons), and $N(0, 1.01)$ is the normal distribution, with mean 0 and standard deviation 1.01.

We used the number-at-age in 1970 as the initial stock size of the simulation. We used the same target levels as the deterministic analysis above. If the estimated stock size is lower than that in 1970, no fishing is allowed. We used a normal distribution with mean 1 and standard deviation 0.5 for the estimation error of the number-at-age. We assumed independent estimation error for each age class. Namely, not only the total stock size, but also the age composition were estimated with error. When an estimated number-at-age become negative, we used zero for the estimated number-at-age. We used the same target level of reproductive potential ($RP_{Target} = 1.55 \times 10^7$ t) as the deterministic analysis above. The fishing effort (f) under CRP is determined as the RP of the standing stock equals to the target level of RP:

$$\sum_{i=1}^{7+} \left(\varepsilon_i N_i \sum_{j=i}^{\infty} \left[W_j m_j \exp\left\{ -\sum_{k=i}^{j} \left(q_k f + M \right) \right\} \right] \right) = 1.55 \times 10^7 \text{ t} \tag{6}$$

where ε_j is the estimation error of the number-at-age j, W_j is weight at age j, and m_j is proportion mature at age j. Fecundity at age j (E_j) is expressed as the product of body weight at age j (W_j) and maturity at age j (m_j). If the estimated stock size is too low to reach RP_{Target}, we cannot satisfy equation (5). Therefore, in that case, f is set at zero.

The fishing pressure of CSB is determined as the SSB of the standing stock equals to the target level:

$$\sum_{i=1}^{7+} \left\{ \varepsilon_i N_i W_i m_i \exp\left(-q_i f - M \right) \right\} = 3.00 \times 10^6 \text{ t} \tag{7}$$

In this simulation, we updated f annually by the equations (6) and (7), using newly estimated number-at-age. The simulated time is 50 years. We repeated this simulation 100 times.

Results

The performance criteria of constant reproductive potential (CRP) and constant SSB (CSB) are presented in Table 2. Figure 2 shows fishing effort, SSB, and yield of one realization of the hundred replications.

Table 2. **Performance criteria of the constant reproductive potential (CRP) and constant SSB (CSB).**

Performance criteria	CRP	CSB
Target level (10^6 t)	15.5	3.00
Mean yield (10^6 t)	4.36	2.17
Variance in yield (10^{13} t^2)	4.08	1.18
Mean fishing effort (f)	0.76	0.82
Variance in fishing effort (f)	0.42	0.93
Mean *RP* level (10^7 t)	5.56	2.70
Mean SSB level(10^6 t)	6.20	2.99
Minimum *RP* (10^7 t)	1.69	0.96
Minimum SSB (10^6 t)	2.93	1.71
Overfishing years	0.69	29.96
Moratorium years	2.31	13.14

All criteria are expressed by the average of the 100 replications.

The average of fishing effort (f) of CSB was slightly higher than that of CRP, but the variance of f of CSB was almost four times bigger than that of CRP. Figure 2a shows one realization of fishing effort over time. As we used the same time series of estimation error and recruitment variation, the trends of fishing effort of both strategies were similar. But the peaks of CSB were sharper and higher than that of CRP. The f-curve for CSB is steeper than that for CRP (Fig. 1). This implies that the value of f determined by CSB suffer higher fluctuation by the error of stock size estimation than that determined by CRP. Numbers of the moratorium years were significantly different. The moratorium years of CSB was more than 25% of the simulated time. Such frequent moratorium is not desirable for the fishing industry.

Figure 2b shows the time series of SSB. SSB level of CRP was higher and more variable than that of CSB. As we used a highly variable stock-recruitment curve, there were some years with high recruitment. CSB exploit all of the surplus production within one season. Therefore, SSB level of CSB fluctuated near the target level. On the other hand, as CRP exploited the abundant cohort gradually, SSB level of CRP increased after recruitment success. This is the reason why SSB level of CRP was higher and more variable than that of CSB. The average minimum *RP* was higher than the target level. If the stock is productive enough, the target level of *RP* works as the lower limit. CRP may be effective for conservation of the stock.

More than half of the simulated time, CSB decreased SSB level below the target.

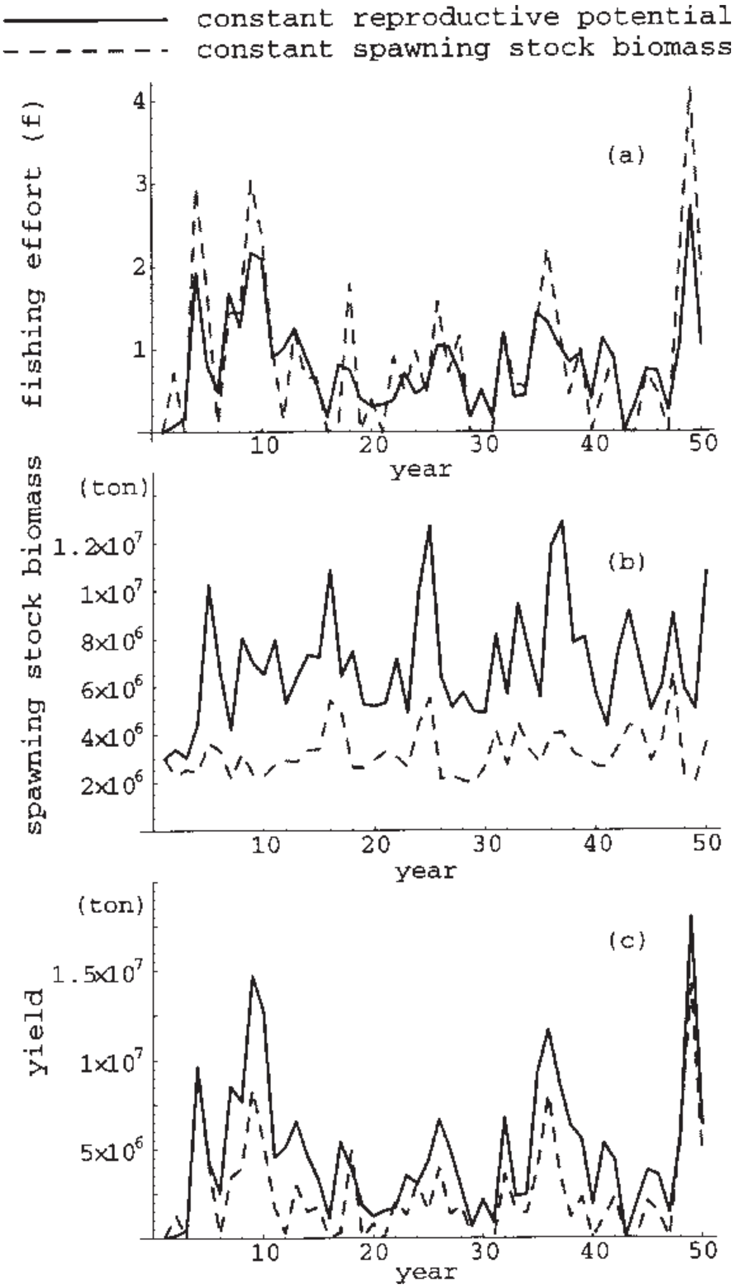


Figure 2. The time series of fishing effort, SSB and yield for constant reproductive potential and constant SSB. This represents one of the 100 replications.

As we used a stock with high productivity, the newly recruited cohort often compensated the overfishing. This is the reason why there was a moratorium on fishing only 13 years while overfishing occurred more than 30 years. The minimum SSB level of CSB is about half of the target level. For a stock with low productivity, such overfishing may result in long successive moratorium.

We examined the robustness of two strategies by a deterministic analysis. We assumed the same estimation error (ε) for all age classes. In this case, f is determined by the biased relative stock size ($\varepsilon\alpha$) instead of true relative stock size (α). Figure 3a shows the stock size after one year exploitation at f determined by CRP. Here we expressed the stock size as the maximum future spawning of survived stock, which is attained by the moratorium starting from the next year. The darker area indicates the smaller stock size. The areas surrounded by the dotted line indicates overfishing. In this area, the standing stock cannot achieve the target lifetime spawning (target *RP*), even if the standing stock will not be exploited from next year. Otherwise, we can continue exploitation and achieve RP_{Target} by tuning f. Under CRP, overestimating the stock size results in a fishing plan that leads to overfishing. But, we exploit only a part of surplus yield within one fishing season. Even when the stock size has been overestimated, we may still reach the target *RP* by decreasing f the next year.

Figure 3b shows the SSB after one year of exploitation at f determined by CSB. Because CSB exploits all surplus stock within a year, overestimating the stock number inevitably results in overfishing. The darkest areas indicates severe overfishing. In these areas, the escapement size is less than 75% of the target level. The darkest area of CSB is wider than that of CRP. Consequently, by applying CRP, we can reduce not only the probability of overfishing but also the intensity of overfishing.

By setting SSB at the most productive level, we can achieve the maximum sustainable yield (Reed 1979). As the recruitment of the next year relates to SSB, the fisheries management based on SSB has an advantage in terms of the maximizing recruitment. In this simulation, yield of CRP is about the double of CSB. This is mainly due to the difference in the biomass level. As the variance of SSB level is smaller for CSB, CSB may maximize yield if the target level is properly set. On the other hand, as *RP* represents the long-term productivity of the stock, the fisheries management based on *RP* is effective for the conservation of the stock.

The fisheries management based on *RP* and that based on SSB have different strengths depending on the nature of the indicators. Therefore, it is essential to clarify the objective of the fisheries management and to choose the appropriate indicator in view of that objective. SSB should be used for the maximizing yield, while *RP* should used for the conservation of the stock. We can use the advantage of both SSB and *RP* by setting the target level of SSB and the threshold level of SP at the same time. If the stock level is low, we can protect the stock by using the *RP* indicator.

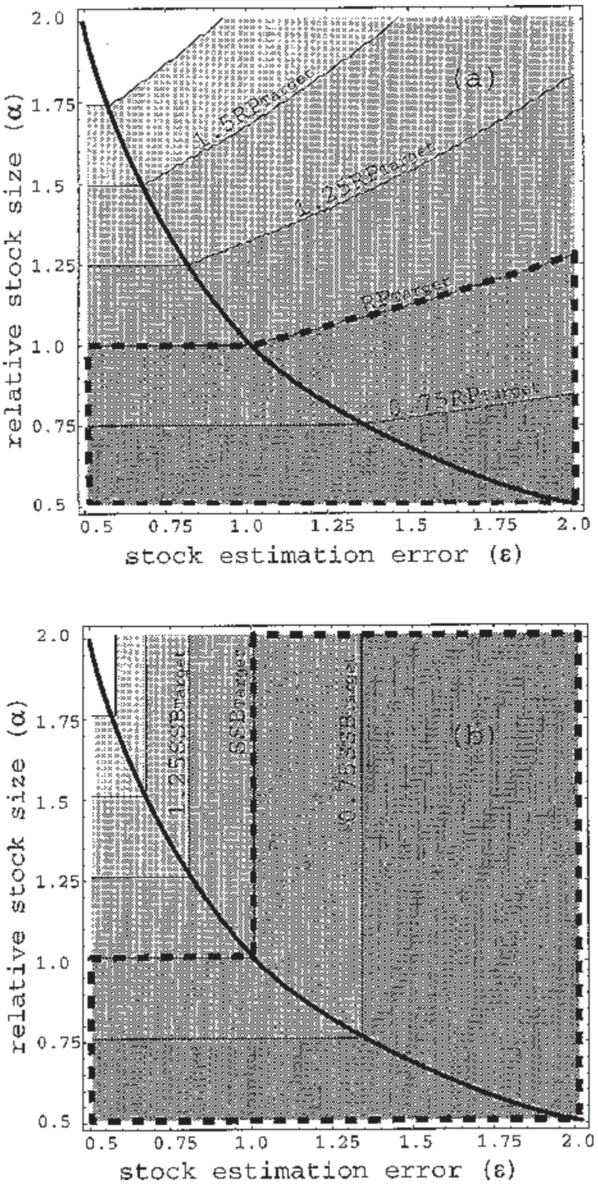


Figure 3. The distribution of escapement stock numbers after one year's exploitation under the constant escapement strategy based on the reproductive potential (a) and that based on SSB (b). The areas below the curved line in both diagrams represents the areas of moratorium on fishing. The area surrounded by the dotted line indicates overfishing.

Conversely, if the stock level is high enough, we can maximize the number recruits by using the SSB indicator.

Discussion

Feedback Management

In Rosenberg et al. (1994), 67% of the overfishing definitions are fishing mortality rate–based. The remaining 33% of the definitions are based on SSB. Although fishing mortality based management is widely used, sustainable fishing mortality rate is usually uncertain. Feedback from stock size is effective to avoid the risk from this uncertainty. Even if sustainable fishing mortality rate is overestimated, fishing mortality rate is automatically decreased by the feedback from stock decline. Feedback process works even when we don't have enough knowledge about the mechanism of the population dynamics.

Although theoretically appealing, constant escapement strategy is not widely used, because of highly vulnerability to the stock estimation error. Instead of CES, SSB threshold policy is commonly used (Sigler and Fujioka 1993, Ianelli and Heifetz 1995). Under threshold policy, harvesting occurs at a constant rate but reduced or ceased when stock size drops below a threshold. The objective of threshold policy is mainly conservation of the stock. Then *RP* is an appropriate indicator for threshold policy, because *RP* has advantages to evaluate long-term spawning ability. Furthermore, as we have seen, we can decrease risk of overfishing due to stock size estimation error by using a *RP* indicator.

Comparison between Reproductive Potential and Spawning per Recruit

The calculation method of reproductive value (*RV*) is similar to that of spawning per recruit, which is often used in fisheries management (Gabriel et al. 1989, Clark 1993, Goodyear 1993, Mace and Sissenwine 1993). Spawning per recruit is equivalent to the *RV* of an individual that has just been recruited. Although the basic equation of SPR and reproductive potential is related, the management concept of both are completely different. The objective of SPR analysis is to obtain the exploitation rate that will lead the stock to the sustainable equilibrium. SPR analysis is a tool for the constant harvesting rate strategy, under which the fishing rate is determined independent of the stock size. On the other hand, the objective of the management based on *RP* is directly escaping enough spawning stock. We can use *RP* as the target level or threshold level of the stock size. In this case, fishing pressure is determined by the estimated stock size.

Alee Effect

Alee (1931) pointed out that a population may fail to recover if the stock decreases below a certain threshold level. Some stocks have been suggested to exhibit such behavior, in the sense of experiencing a severe decline and subsequently failing to recover despite a reduction in the fishing mortality rate (Thompson 1993). In this study, CRP always led to a higher minimum stock level than CSB. Thus, fisheries management based on reproductive potential (*RP*) may effectively prevent the stock from falling below an undesirable point of no return.

Acknowledgments

This work owes much to the thoughtful and helpful comments of Drs. Yasushi Harada, Jonathan Heifetz, Kazuhiko Hiramatsu, Kjartan G. Magnusson, Hiroyuki Matsuda, and Shuhei Ohnishi. Anonymous reviewers helped to improve the manuscript. This research was supported by a Research Fellowship of the Japan Society for the Promotion of Science for Young Scientists.

References

Allee, W.C. 1931. Animal aggregations. University of Chicago Press, Chicago. 431 pp.

Beddington, J.R., and J.G. Cooke. 1983. The potential yield of fish stocks. FAO Fish. Tech. Pap. 242. 50 pp.

Clark, W.G. 1993. The effect of recruitment variability on the choice of a target level of spawning biomass per recruit. In: G. Kruse, D.M. Eggers, R.J. Marasco, C. Pautzke, and T.J. Quinn II (eds.), Proceedings of the International Symposium on Management Strategies for Exploited Fish Populations. University of Alaska Sea Grant, AK-SG-93-02, Fairbanks, pp. 233-246.

Fisher, R.A. 1930. The genetic theory of natural selection. Oxford, U.K. 291 pp.

Gabriel, W.L., M.P. Sissenwine, and W.J. Overholtz. 1989. Analysis of spawning stock biomass per recruit: An example for Georges Bank haddock. N. Am. J. Fish. Manage. 9:383-391.

Goodyear, C.P. 1993. Spawning stock biomass per recruit in fisheries management: Foundation and current use. In: S.J. Smith, J.J. Hunt, and D. Rivard (eds.), Risk evaluation and biological reference points for fisheries management. Can. Spec. Publ. Fish. Aquat. Sci. 120:67-81.

Ianelli, J.N., and J. Heifetz 1995. Decision analysis of alternative harvest policies for the Gulf of Alaska Pacific Ocean Perch Fishery. Fish. Res. 24:35-63.

Katsukawa, T. 1997. Introduction of spawning potential: Improvement in the threshold management theory. Rev. Fish Biol. Fish. 7:285-289.

Mace, P.M., and M.P. Sissenwine. 1993. How much spawning per recruit is enough? In: S.J. Smith, J.J. Hunt, and D. Rivard (eds.), Risk evaluation and biological reference points for fisheries management. Can. Spec. Publ. Fish. Aquat. Sci. 120:101-118.

Myers, R.A., A.A. Rosenberg, P.M. Mace, N. Barrowman, and V.R. Restrepo. 1994. In search of thresholds for recruitment overfishing. ICES J. Mar. Sci. 51:191-205.

Reed, W.J. 1979. Optimal escapement levels in stochastic and deterministic harvesting models. J. Environ. Econ. Manage. 6:350-363.

Rosenberg, A., P. Mace, G. Thompson, G. Darcy, W. Clark, J. Collie, W. Gabriel, A. MacCall, R. Methot, J. Powers, V. Restrepo, T. Wainwright, L. Botsford, J. Hoening, and K. Stokes. 1994. Scientific review of definitions of overfishing in U.S. fishery management plans. NOAA Tech. Mem. NMFS-F/SPO-17.

Sigler, M.F., and J.T. Fujioka. 1993. A comparison of policies for harvesting sablefish, *Anoplopoma fimbria*, in the Gulf of Alaska. In: G. Kruse, D.M. Eggers, R.J. Marasco, C. Pautzke, and T.J. Quinn II (eds.), Proceedings of the International Symposium on Management Strategies for Exploited Fish Populations. University of Alaska Sea Grant, AK-SG-93-02, Fairbanks, pp. 7-19.

Thompson, G.G. 1993. A proposal for a threshold stock size and maximum fishing mortality rate. In: S.J. Smith, J.J. Hunt and D. Rivard (eds.), Risk evaluation and biological reference points for fisheries management. Can. Spec. Publ. Fish. Aquat. Sci. 120:303-320.

Wada, T., C. Sato, and Y. Matsumiya. 1996. Fisheries management for the Pacific stock of chub mackerel, *Scomber japonicus*, based on spawning per recruit analysis. Bull. Jpn. Soc. Fish. Oceanogr. 60:363-371. (In Japanese.)

Zheng, J., T.J. Quinn II, and G.H. Kruse. 1993. Comparison and evaluation of threshold estimation methods for exploited fish populations. In: G. Kruse, D.M. Eggers, R.J. Marasco, C. Pautzke, and T.J. Quinn II (eds.), Proceedings of the International Symposium on Management Strategies for Exploited Fish Populations. University of Alaska Sea Grant, AK-SG-93-02, Fairbanks, pp. 267-289.

Participants

Adkison, Milo
University of Alaska Fairbanks
SFOS
11120 Glacier Highway
Juneau AK 99801
Ph: 907/465-6251; Fax: 907/465-6447
ffmda@uaf.edu

Anderson, Jim
University of Washington
School of Fisheries
Box 358218
Seattle WA 98195
Ph: 206/543-4772
jim@fish.washington.edu

Applegate, Andrew
New England Fishery Mgmt. Council
5 Broadway
Saugus MA 01906
Ph: 781/231-0422
aapplegate@nefmc.org

Aubone, Anibal
Inst. National de Invest. y
Desarrollo Pesquero (INIDEP)
Larrea 631
Mar del Plata Buenos Aires 7600
ARGENTINA
Ph: +54 23 673575
aaubone@mdp.edu.ar

Baleña, Rex
Ocean-Weather Laboratory
PO Box 249
Iloilo City 5000
PHILIPPINES
Ph: 63 3331 58378; Fax: 63 33 3158441

Bechtol, Bill
Alaska Dept. of Fish and Game
3298 Douglas St.
Homer AK 99603-7942
Ph: 907/235-8191; Fax: 907/235-2448
billb@fishgame.state.ak.us

Bernard, Dave
Alaska Dept. of Fish and Game
333 Raspberry Rd.
Anchorage AK 99518-1599
Ph: 907/267-2380
davidbe@fishgame.state.ak.us

Bingham, Allen
Alaska Dept. of Fish and Game
333 Raspberry Rd.
Anchorage AK 99518-1599
Ph: 907/267-2327
AllenB@fishgame.state.ak.us

Björnsson, Höskuldur
Marine Research Institute
PO Box 1390
Skulagata 4
101 Reykjavik
ICELAND
Fax: 354-5623790

Booth, Tony
Rhodes University
Dept. of Ichthyology & Fisheries
PO Box 94
Grahamstown 6140
SOUTH AFRICA
Ph: 27 461 318420; Fax: 27 461 24827
t.booth@ru.ac.za

Brannian, Linda
Alaska Dept. of Fish and Game
333 Raspberry Rd.
Anchorage AK 99518-1599
Ph: 907/267-2118
lindab@fishgame.state.ak.us

Brodziak, Jon
NOAA/NMFS/NWFSC
2030 SE. Marine Science Dr.
Newport OR 97365-5296
Ph: 541/867-0243; Fax: 541/867-0389
Jon.Brodziak@noaa.gov

Bromaghin, Jeff
Alaska Dept. of Fish and Game
333 Raspberry Rd.
Anchorage AK 99518
Ph: 907/267-2379; Fax: 907/267-2442
jeffbr@fishgame.state.ak.us

Brown, Evelyn
University of Alaska Fairbanks
Institute of Marine Science
PO Box 757220
Fairbanks AK 99775-7220
Ph: 907/474-5801
ebrown@ims.alaska.edu

Bruden, Dana
Alaska Dept. of Fish and Game
PO Box 240603
Anchorage AK 99524
Ph: 907/267-2440; Fax: 907/267-2442
danab@fishgame.state.ak.us

Bue, Fred
Alaska Dept. of Fish and Game
Box 1215
Nome AK 99762
Ph: 907/443-5167
fbue@fishgame.state.ak.us

Byrne, Larry
Alaska Dept. of Fish and Game
211 Mission Rd.
Kodiak AK 99615
Ph: 907/486-1875; Fax: 907/486-1824
lbyrne@fishgame.state.ak.us

Campbell, Robert
CSIRO Division of Marine Research
PO Box 1538
Hobart Tasmania 7001
AUSTRALIA
Robert.Campbell@marine.csiro.au

Carlile, David
Alaska Dept. of Fish and Game
PO Box 25526
Juneau AK 99802
Ph: 908/465-4216
davec@fishgame.state.ak.us

Clark, John
Alaska Dept. of Fish and Game
PO Box 240020
Douglas AK 99824-0020

Coggins, Lew
Alaska Biological Science Center
1101 E. Tudor Rd.
Anchorage AK 99503
Ph: 907/786-3576
Lewis_Coggins@usgs.gov

Conser, Ramon
NOAA/NMFS/NWFSC
2030 SE Marine Science Dr.
Newport OR 97365
Ph: 541/867-0196; Fax: 541/867-0389
rconser@sable.nwfsc-hc.noaa.gov

Correia, Steven
MA Division of Marine Fisheries
50A Portside Dr.
Pocasset MA 02559
Ph: 617/727-4809 x111
steven.correia@state.ma.us

Cox, Sean
University of British Columbia
Fisheries Centre
2204 Main Mall
Vancouver BC V6T 1Z4
CANADA
scox@fisheries.com

Crone, Paul
NOAA/NMFS/NWFSC
2030 SE Marine Science Dr.
Newport OR 97365
Ph: 541/867-0307; Fax: 541/867-0389
pcrone@nwfsc-hc.noaa.gov

Cummings, Nancie
NOAA/NMFS/SEFSC
75 Virginia Beach Dr.
Miami FL 33149
Ph: 305/361-4234; Fax: 305/361-4419
nancie.cumings@noaa.gov

Davies, Nick
National Inst. of Water and
Atmospheric Research Ltd (NIWA)
PO Box 1043
Whangarei
NEW ZEALAND
Ph: (09) 438 6578; Fax: (90) 438 0185
n.davies@niwa.cri.nz

De Oliveira, José
Sea Fisheries Research Institute
Private Bag X2
Rogge Bay 8012
Cape Town
SOUTH AFRICA
JDOLIVEI@sfri.wcape.gov.za

DiCosimo, Jane
North Pacific Fishery Mgmt. Council
605 W 4th Ave.
Anchorage AK 99501
Ph: 907/271-2809; Fax: 907/271-2817
Jane.DiCosimo@noaa.gov

Dorval, Emmanis
Old Dominion University (AMRL)
1034 W 45th St.
Norfolk VA 23529
Ph: 757/683-4195; Fax: 757/683-5293
dorval@estuary.amrl.odu.edu

Dressel, Sherri
University of Alaska Fairbanks
Institute of Marine Science
PO Box 757220
Fairbanks AK 99775-7220
Ph: 907/474-7918
dressel@ims.alaska.edu

Evenson, Matt
Alaska Dept. of Fish and Game
1300 College Rd.
Fairbanks AK 99701
Ph: 97/459-7207
mevenson@fishgame.state.ak.us

Fair, Lowell
Alaska Dept. of Fish and Game
Comm. Fisheries Mgmt. & Dev. Div.
333 Raspberry Rd.
Anchorage AK 99518-1599
LowellF@fishgame.state.ak.us

Fletcher, Rick
NSW Fisheries Research Institute
PO Box 21
Cronulla NSW 2230
AUSTRALIA
fletcher@fisheries.nsw.gov.au

Folmer, Ole
Greenland Inst. of Natural Res.
Box 570
Nuuk 3900
GREENLAND
Ph: 299 2 10 95
folmer@natur.gl

Fontenelle, Guy
Ecole Nationale Superieure
Agronomique de Rennes
65 Rue de St Brieuc
35042 Rennes cdx
FRANCE
Ph: 33 02 99 28 75 33; Fax: 33 02 99 28 75 35
fontenel@roazhon.inra.fr

Foster, Matt
University of Alaska Fairbanks
Fisheries Division
11450 Glacier Highway
Juneau AK 99801
ffmbf@uaf.edu

Fournier, David
Otter Consulting Ltd.
PO Box 265, Station A
Nanaimo BC V9R 5K9
CANADA
Ph: 604/756-0956; Fax: 604/756-0956
otter@island.net

Francis, Chris
National Institute of Water and
Atmospheric Research (NIWA)
PO Box 14901
Wellington
NEW ZEALAND
Ph: 64 4 386 0300; Fax: 64 4 386 0572
c.francis@niwa.cri.nz

Fu, Caihong
University of Alaska Fairbanks
SFOS
11120 Glacier Highway
Juneau AK 99801
Ph: 907/465-6372; Fax: 907/465-6320
ftcf@uaf.edu

Fujioka, Jeff
NOAA/NMFS Auke Bay Lab
11305 Glacier Highway
Juneau AK 99801-8626
Ph: 907/789-6026
Jeff.Fujioka@noaa.gov

Funk, Fritz
Alaska Dept. of Fish and Game
PO Box 25526
Juneau AK 99802
Ph: 907/465-6113; Fax: 907/465-6204
FritzF@fishgame.state.ak.us

Gallaway, Benny
LGL Alaska Research Associates
1410 Cavitt
Bryan TX 77801
Ph: 409/775-2000; Fax: 409/775-2002
cole_john@mns.com

Gasalla, Maria
Instituto de Pesca
Divisao de Pesca Maritima
Av. Bartolomeu de Gusmao, 192
Santos SP CEP: 11030-906
BRASIL
Ph: (013)227-5995; Fax: (013)236-1900
ipescapm@eu.ansp.br

Gavaris, Stratis
Dept. of Fisheries and Oceans
Biological Station
St. Andrews NB E0G 2X0
CANADA
Ph: 506/529-5912; Fax: 506/529-5862
GavarisS@mar.dfo-mpo.gc.ca

Gislason, Henrik
Danish Inst. for Fisheries Research
Charlottenlund Castle
Copenhagen DK 2920
DENMARK
Ph: 33963361
hg@dfu.min.dk

Gracia, Adolfo
Instituo de Ciencias del Mar y
Limnologia
Apartado Postal 70-305
Mexico D.F. 04510
MEXICO
Ph: 52 5 622 5779; Fax: 52 5 616 0748
gracia@mar.icmyl.unam.mx

Gudmundsson, Gudmundur
Central Bank of Iceland
Economics Department
Kalkofnsvegur 1
Reykjavik 150
ICELAND
Ph: 354 5699683; Fax: 354 5699608
gudmg@centbk.is

Gunnari, Gerald
Coos Bay Trawlers Assoc.
PO Box 5435
Charleston OR 97420
Ph: 541/888-8012
ctrawl@mail.coos.or.us

Gustafson, Richard
Alaska Dept. of Fish and Game
3298 Douglas St.
Homer AK 99603

Hagen, Peter
Alaska Dept. of Fish and Game
PO Box 25526
Juneau AK 99802-5526
Ph: 907/465-3054; Fax: 907/465-2765
peterh@fishgame.state.ak.us

Haist, Vivian
Dept. of Fisheries and Oceans
Pacific Biological Station
Hammond Bay Rd.
Nanaimo BC V9R 5K6
CANADA
Ph: 250/756-7205; Fax: 250/756-7138
haistv@pbs.dfo.ca

Hale, Jim
NOAA/NMFS
PO Box 21668
Juneau AK 99802
Ph: 907/586-7228
jhale@ptialaska.net

Hall, Norm
Western Australian Marine Research
Laboratories, Fisheries Department
Perth West. Aust.
AUSTRALIA
Ph: 618 9246 8441; Fax: 618 9447 3062
nhall@fish.wa.gov.au

Hamner, Helen
Alaska Dept. of Fish and Game
333 Raspberry Rd.
Anchorage AK 99518
Ph: 907/267-2120
HelenH@fishgame.state.ak.us

Hanchet, Stuart
National Inst. of Water and
Atmospheric Research (NIWA)
PO Box 893
Nelson
NEW ZEALAND
Ph: 03 548 1715; Fax: 03 548 1716
s.hanchet@niwa.cri.nz

Hankin, David
Humboldt State University
Dept. of Fisheries
Arcata CA 95521
Ph: 707/826-3683; Fax: 707/826-3682
dgh1@axe.humboldt.edu

Heifetz, Jonathan
NOAA/NMFS Auke Bay Lab
11305 Glacier Highway
Juneau AK 99801
Ph: 907/789-6054; Fax: 907/789-6094
jon.heifetz@noaa.gov

Hicks, Alan
PO Box 448
Avila Beach CA 93424
Fax: 707/825-7780
ach1@axe.humboldt.edu

Hill, Ryan
Triton Environmental Consultants
120-13511 Commerce Parkway
Richmond BC V6V 2L1
CANADA
Ph: 604/279-2093; Fax: 604/279-2047
rhill@triton-env.com

Hills, Sue
University of Alaska Fairbanks
Institute of Marine Science
PO Box 757220
Fairbanks AK 99775-7220
Ph: 907/474-5106; Fax: 907/474-5177
shills@ims.uaf.edu

Hiramatsu, Kazuhiko
Nat'l Research Inst. of Far Seas
Fisheries
5-7-1 Orido
Shimazu 424
JAPAN
Ph: 81-543-36-6014; Fax: 81-543-35-9642
hira@enyo.affrc.go.jp

Holland, Ken
Minerals Management Service
949 E 36th Ave., Rm. 308
Anchorage AK 99508-4302
Ph: 907/271-6684

Hosie, Mike
Coos Bay Trawlers Assoc.
PO Box 5050
Charleston OR 97420
Ph: 541/888-8012
ctrawl@mail.coos.or.us

Huiskes, Mark
Wageningen Agricultural University
Dept. of Mathematics
Dreijenlaan 4
Wageningen 6703 HA
THE NETHERLANDS
Ph: +31 317 482385; Fax: +31 317
483554
Mark.Huiskes@ztw.wk.wau.nl

Ianelli, Jim
NOAA/NMFS/AFSC
7600 Sand Point Way NE
Building 4
Seattle WA 98115
Ph: 206/526-6510; Fax: 206/526-6723
jim.ianelli@noaa.gov

Isaac, Victoria
Museu Paraense Emilio Goeldi
Dpto de Zoologia
C.P. 399
Belem-PA 66017-970
BRAZIL
Ph: 0055 91 2282897; Fax: 0055 91
2261615
victoria@amazon.com.br

Jakobsen, Tore
Inst. of Marine Research
PO Box 1870
N-5024 Bergen Nordnes
NORWAY
Ph: +47 55 23 8500; Fax: +47 55 23 8687
Tore.Jakobsen@imr.no

Jones, Christopher
University of Miami-RSMAS
4600 Rickenbacker Causeway
Miami FL 33149
Ph: 305/361-4288; Fax: 305/361-4499
cjones@rsmas.miami.edu

Jones, Cynthia
Old Dominion University
Appl. Marine Res. Lab.
1034 W 45th St.
Norfolk VA 23529-0456
Ph: 757/683-4497; Fax: 757/683-5293
jones@estuary.amrl.odu.edu

Katsukawa, Toshio
University of Tokyo
Ocean Research Institute
1-15-1, Nakano-ku
Tokyo 164
JAPAN
Ph: +81 3 5351 6495; Fax: +81 3 5351
6492
katukawa@ori.u-tokyo.ac.jp

Kirchner-Frankle, Carola
Ministry of Fisheries and Marine
Resources
Box 912
Swakopmund
NAMIBIA
Ph: 064 405744; Fax: 064 404385
carola@mail.fisheries.gov.na

Klaer, Neil
CSIRO Division of Marine Research
GPO Box 1538
Hobart Tasmania 7004
AUSTRALIA
Ph: +61 3 62 325465
neil.klaer@marine.csiro.au

Kline, Tom
Prince William Sound Science Center
PO Box 705
Cordova AK 99574
Ph: 907/424-5800; Fax: 907/424-5820
tkline@grizzly.pwssc.gen.ak.us

Knudsen, Eric
Alaska Biological Science Center
1011 East Tudor Rd.
Anchorage AK 99503
Ph: 907/786-3842; Fax: 907/786-3636
Eric_Knudsen@usgs.gov

Kraft, Cliff
University of Wisconsin
Sea Grant Institute
2420 Nicolet Dr. ES-105
Green Bay WI 54311-7001
Ph: 920/465-2795
kraftc@uwgb.edu

Kruse, Gordon
Alaska Dept. of Fish and Game
Comm. Fisheries Mgmt. & Dev. Div.
PO Box 25526
Juneau AK 99802
Ph: 907/465-6106; Fax: 907/465-2604
GordonK@fishgame.state.ak.us

Kvist, Trine
Technical University of Denmark
Bygning 321, DTU
Lyngby DK-2800
DENMARK
Ph: 4525 2525
tk@imm.dtu.dk

Lai, Han-lin
NOAA/NMFS/NEFSC
166 Water St.
Woods Hole MA 02543
Ph: 508/495-2238; Fax: 508/495-2393
Han-Lin.Lai@noaa.gov

Lane, Dan
University of Ottawa
Faculty of Administration
136 Jean-Jacques Lussier Privee
Ottawa ON K1N 6N5
CANADA
Ph: 613/562-5800 x479; Fax: 613/562-5164
dlane@uottawa.ca

Livingston, Patricia
NOAA/NMFS/AFSC
7600 Sand Point Way NE
Building 4
Seattle WA 98115
Ph: 206/526-4242; Fax: 206/526-6723
pliving@afsc.noaa.gov

Lloyd, Denby
Aleutians East Borough
211 4th St., # 314
Juneau AK 99801
Ph: 907/586-6655; Fax: 907/586-6644
denbyaeb@alaska.net

Lunsford, Chris
NOAA/NMFS Auke Bay Lab
11305 Glacier Highway
Juneau AK 99801
Ph: 907/789-6008
Chris.Lunsford@noaa.gov

Mackinson, Steven
University of British Columbia
Fisheries Centre
2204 Main Mall
Vancouver BC V6T 1Z4
CANADA
Ph: 604/822-1864; Fax: 604/822-8934
smackin@fisheries.com

March, Douglas
Canadian Groundfish Research
and Conservation Society
#300 15225 104th Ave.
Surrey BC V3R 6Y8
CANADA
Ph: 604/588-8585; Fax: 604/588-6708
SBCOCGA@aol.com

Marshall, Bob
Alaska Dept. of Fish and Game
PO Box 240020
Douglas AK 99824-0020
Ph: 907/465-4303
BobM@fishgame.state.ak.us

Masuda, Michele
NOAA/NMFS Auke Bay Lab
11305 Glacier Highway
Juneau AK 99801-8626
Michele.Masuda@noaa.gov

Mathisen, Ole
University of Alaska Fairbanks
SFOS
11120 Glacier Highway
Juneau AK 99801
ffoam@uaf.edu

Matsuishi, Takashi
Hokkaido University
Faculty of Fisheries
3-1-1 Minato-cho
Hakodate Hokkaido 041-8611
JAPAN
Ph: +81 138 40 8857; Fax: +81 138 40 8827
matuisi@fish.hokudai.ac.jp

Matsumiya, Yoshiharu
University of Tokyo
Ocean Research Institute
Minamidai 1-15-1, Nakano-ku
Tokyo 164
JAPAN
Ph: +81 3 5351 6495; Fax: +81 3 5351 6492
matumiya@ori.u-tokyo.ac.jp

Maunder, Mark
University of Washington
School of Fisheries/FRI
Box 357980
Seattle WA 98195-7980
Ph: 206/543-4744; Fax: 206/685-7471
mark@fish.washington.edu

McClure, Marianne
Columbia River Inter-Tribal Fish Commission
PO Box 82812
Portland OR 97282
Ph: 503/731-1254
marianne@pisces.fish.washington.edu

McComish, Tom
Ball State University
Dept. of Biology
Aquatic Biology and Fisheries
Muncie IN 47306
Ph: 765/285-8845
tmccomis@bsu.edu

Merritt, Peggy
Alaska Dept. of Fish and Game
1300 College Rd.
Fairbanks AK 99701
Ph: 907/459-7296
pmerritt@fishgame.state.ak.us

Methot, Rick
NOAA/NMFS/FAM
2725 Montlake Blvd. E
Seattle WA 98112
Richard.Methot@noaa.gov

Moffitt, Robert
NOAA/NMFS Honolulu Laboratory
2570 Dole St.
Honolulu HI 96822-2396
Ph: 808/943-1244
rmoffitt@honlab.nmfs.hawaii.edu

Moore, Holly
Alaska Dept. of Fish and Game
Box 920587
Dutch Harbor AK 99692-1239
Ph: 907/581-1239; Fax: 907/581-1572
hollymo@fishgame.state.ak.us

Müter, Franz
University of Alaska Fairbanks
Institute of Marine Science
Fairbanks AK 99775
Ph: 907/474-7839
mueter@ims.alaska.edu

Murawski, Steven
NOAA/NMFS/NEFSC
Woods Hole Lab, 166 Water St.
Woods Hole MA 02543
Ph: 508/548-5123; Fax: 508/548-1158
smurawsk@whsun1.wh.whoi.edu

Murphy, Mike
FL Dept. of Environmental Protection
Florida Marine Research Institute
100 Eighth Ave. SE
St. Petersburg FL 33701-5095
Fax: 813/823-0166

Murphy, Peggy
Alaska Dept. of Fish and Game
PO Box 25526
Juneau AK 99802
Ph: 907/465-6107; Fax: 907/465-2604
PeggyM@fishgame.state.ak.us

Nakashima, Brian
Dept. of Fisheries and Oceans
NW Atlantic Fisheries Centre
PO Box 5667
St. John’s NFLD A1C 5X1
CANADA
Ph: 709/772-4925; Fax: 709/772-4105
nakashima@athena.nwafc.nf.ca

Nelson, Patricia
Alaska Dept. of Fish and Game
211 Mission Rd.
Kodiak AK 99615
Ph: 907/486-1855

Newman, Ken
University of Idaho
Division of Statistics
Moscow ID 83843-1104
Ph: 208/885-6861; Fax: 208/885-5843
newman@uidaho.edu

Nottestad, Leif
University of British Columbia
Fisheries Centre
2204 Main Mall
Vancouver BC V6T 1Z4
CANADA
notten@ifm.uib.no

Okey, Tom
Center for Marine Conservation
580 Market St. #550
San Francisco CA 94104
Ph: 415/391-6204
tom@cenmarine.com

Olivares, Jorge Flores
Centro de Investigacion y
Graduados del mar
No 219-3 Centro
Mazatlan Sinaloa
MEXICO
cigmar@red2000.com.mx

Otis, Ted
Alaska Dept. of Fish and Game
3298 Douglas St.
Homer AK 99603-8027
Ph: 907/235-8191; Fax: 907/235-2448
tedo@fishgame.state.ak.us

Paine, Brent
United Catcher Boats
1711 W. Nickerson # B
Seattle WA 98119
Ph: 206/282-2599; Fax: 206/282-2414
ucb@eskimo.com

Parma, Ana
Int’l Pacific Halibut Commission
PO Box 95009
Seattle WA 98145-2009
Ph: 206/634-1838 x229
Ana@iphc.washington.edu

Parrack, Michael
NOAA/NMFS/SEFSC
75 Virginia Beach Dr.
Miami FL 33149
Ph: 305/361-4283

Pauly, Daniel
University of British Columbia
Fisheries Centre
2204 Main Mall
Vancouver BC V6T 1Z4
CANADA
Ph: 604/822-1201; Fax: 604/822-8934
pauly@fisheries.com

Pella, Jerry
NOAA/NMFS Auke Bay Lab
11305 Glacier Highway
Juneau AK 99801-8626
Ph: 907/789-6027; Fax: 907/789-6094
jpella@noaa.gov

Pengilly, Doug
Alaska Dept. of Fish and Game
211 Mission Rd.
Kodiak AK 99615
Ph: 907/486-2431; Fax: 907/486-1841
dougp@fishgame.state.ak.us

Perez, Marcelo
Inst. National de Invest. y
Desarrollo Pesquero (INIDEP)
Larrea 631
Mar del Plata Buenos Aires 7600
ARGENTINA
Ph: +54 23 897472
maperez@mdp.edu.ar

Pitcher, Tony
University of British Columbia
Fisheries Centre
2204 Main Mall
Vancouver BC V6T 1Z4
CANADA
Ph: 604/822-2731; Fax: 604/822-8934
TPitcher@fisheries.com

Porch, Clay
NOAA/NMFS/SEFSC
75 Virginia Beach Dr.
Miami FL 33149
Ph: 305/361-4232; Fax: 305/361-4499
clay.porch@noaa.gov

Powers, Joseph
NOAA/NMFS/SEFSC
75 Virginia Beach Dr.
Miami FL 33149
Ph: 305/361-4487; Fax: 907/361-4478
Joseph.Powers@noaa.gov

Preece, Ann
CSIRO Division of Marine Research
PO Box 1538
Hobart Tasmania 7001
AUSTRALIA
Ph: +61 3 6232 5336; Fax: +61 3 6232 5000
Ann.Preece@marine.csiro.au

Preikshot, Dave
University of British Columbia
Fisheries Centre
2204 Main Mall
Vancouver BC V6T 1Z4
CANADA
Ph: 604/822-0618; Fax: 604/822-8934
preikshot@fisheries.com

Quinn, Terry
University of Alaska Fairbanks
SFOS
11120 Glacier Highway
Juneau AK 99801
Ph: 907/465-5389; Fax: 907/465-6320
fftjq@aurora.alaska.edu

Radchenko, Vladimir
TINRO
4 Savchenko Alley
Vladivostock 690600
RUSSIA
Ph: 7 4232 25 95 04

Restrepo, Victor
Univesity of Miami-RSMAS
4600 Rickenbacker Causeway
Miami FL 33149
Ph: 305/361-4022; Fax: 305/361-4457
vrestrepo@rsmas.miami.edu

Richards, Laura
Dept. of Fisheries and Oceans
Pacific Biological Station
3190 Hammon Bay Rd.
Nanaimo BC V9R 5K6
CANADA
Ph: 250/756-7177; Fax: 250/756-7053
richardsl@pbs.dfo.ca

Roach, Don
Alaska Dept. of Fish and Game
1300 College Rd.
Fairbanks AK 99701
Ph: 907/459-7229
droach@fishgame.state.ak.us

Rosenkranz, Gregg
University of Alaska Fairbanks
SFOS
245 O'Neill Building
Fairbanks AK 99775-7220
Ph: 907/474-7594
fsger@uaf.edu

Rosing, Micheal
University of Washington
QERM, FRI
Box 357980
Seattle WA 98195-7980
Ph: 206/616-9420
rosing@cqs.washington.edu

Saelens, Mark
Oregon Dept. of Fish & Wildlife
2040 SE Marine Science Dr.
Newport OR 97365
Ph: 503/867-0300 x251
saelensm@ccmail.orst.edu

Sagalkin, Nick
Alaska Dept. of Fish and Game
211 Mission Rd.
Kodiak AK 99615
Ph: 907/486-1858
nsagalki@fishgame.state.ak.us

Saila, Saul
University of Rhode Island
Narragansett Bay Campus
Narragansett RI 02882-1197
Ph: 401/874-6211; Fax: 401/874-6486
ssaila@gsosun1.gso.uri.edu

Sainsbury, Keith
CSIRO Division of Marine Research
GPO Box 1538
Hobart Tasmania 7001
AUSTRALIA
Ph: 03 6232 5369; Fax: 06 6232 5199
Keith.Sainsbury@ml.csiro.au

Sampson, David
Oregon State University
Coastal OR Marine Experiment Sta. &
Dept. of Fisheries & Wildlife
Newport OR 97365
Ph: 541-867-0204; Fax: 503/867-0105
dsampson@slugo.hmsc.orst.edu

Sands, Norma Jean
Alaska Dept. of Fish and Game
PO Box 25526
Juneau AK 99802-5526
Ph: 907/465-4256; Fax: 907/465-2332
normas@fishgame.state.ak.us

Schirripa, Michael
NOAA/NMFS/SEFSC
Miami Laboratory
4600 Rickenbacker Causeway
Miami FL 33149-1099
Ph: 305/361-4221
michael.schirripa@noaa.gov

Schmidt, Graham
Canadian Groundfish Research and Conservation Society
#300 15225 104th Ave.
Surrey BC V3R 6Y8
CANADA
Ph: 604/588-8585; Fax: 604/588-6708
SBCOCGA@aol.com

Schnute, Jon
Dept. of Fisheries and Oceans
Pacific Biological Station
3190 Hammon Bay Rd.
Nanaimo BC V9R 5K6
CANADA
Ph: 250/756-7146; Fax: 250/756-7053
schnutej@pbs.dfo.ca

Schweigert, Jake
Dept. of Fisheries and Oceans
Pacific Biological Station
3190 Hammond Bay Rd.
Nanaimo BC V9R 5K6
CANADA
Ph: 250/756-7203
schweigertj@pbs.dfo.ca

Scott, Jason
Kalispel Tribe of Indians
PO Box 39
Usk WA 99180
Ph: 509/445-1147

Seifert-Ethier, Konstanze
CDS
100 International Dr.
Mount Olive NJ 07828
Ph: 973/691-3654
seifert@cdsusa.com

Shroyer, Steve
Ball State University
Dept. of Biology
Muncie IN 47306-0440
Ph: 765/285-8820; Fax: 765/285-8804
sshroyer@bsu.edu

Sigler, Michael
NOAA/NMFS Auke Bay Lab
11305 Glacier Highway
Juneau AK 99801
Ph: 907/789-6037; Fax: 907/789-6094
mike.sigler@noaa.gov

Smith, Tony
CSIRO Division of Fisheries
Castray Esplanade
GPO Box 1538
Hobart Tasmania 7001
AUSTRALIA
Ph: +61 3 6232-5372; Fax: +61 3 6232 5199
tony.smith@marine.csiro.au

Soiseth, Chad
National Park Service
PO Box 140
Gustavus AK 99826
Ph: 907/697-2230
chad_soiseth@nps.gov

Starr, Paul
NZ Seafood Industry Council
Private Bag 24901
Wellington
NEW ZEALAND
Ph: 001 644 3854005
paul@seafood.co.nz

Stefansson, Gunnar
Marine Research Institute
PO Box 1390
121 Reykjavik
ICELAND
Ph: +354-552-0240; Fax: +354-562-3790
gunnar@hafro.is

Su, Zhenming
University of Alaska Fairbanks
SFOS
11320 Glacier Highway
Juneau AK 99801
ftzs@uaf.eud

Sullivan, Kevin
Ministry of Fisheries
PO Box 1020
Wellington
NEW ZEALAND
Ph: 04 470 2692; Fax: 04 470 2686
SULLIVAK@fish.govt.nz

Sullivan, Patrick
Int'l Pacific Halibut Commission
PO Box 95009
Seattle WA 98145
Ph: 206/634-1838 x 22; Fax: 206/632-2983
pat@iphc.washington.edu

Swanton, Charles
Alaska Dept. of Fish and Game
211 Mission Rd.
Kodiak AK 99615
Ph: 908/486-1852
Swantonc@fishgame.state.ak.us

Tallman, Ross
Dept. of Fisheries and Oceans
Arctic Fish Stock Assessment
501 University Crescent
Winnipeg MB R3T 2N6
CANADA
Ph: 204/983-3362; Fax: 204/984-2403
TallmanR@DFO-MPO.GC.CA

Thera, Trevor
Alberta Conservation Association
Bag 900-26 Provincial Building
Peace River AB T8S 1T4
CANADA
Ph: 403/624-6405
tthera@env.gov.ab.ca

Tracy, Donn
Alaska Dept. of Fish and Game
211 Mission Rd.
Kodiak AK 99615
Ph: 907/486-1822
donnt@fishgame.state.ak.us

Tsuji, Sachiko
Nat'l Research Inst. of Far Seas Fisheries
5-7-1 Orido
Shimizu 424
JAPAN
Ph: 81 543 36 6042; Fax: 81 543 35 9642
tsuji@enyo.affrc.go.jp

Turnbull, Clive
Dept. of Primary Industries, Qnslnd
Northern Fisheries Centre
PO Box 5396
Cairns QLD 4870
AUSTRALIA
Ph: 61 70 529 888; Fax: 61 70 351 401
turnbuc@ensis.nth.dpi.qld.gov.au

Vasconcellos, Marcelo
University of British Columbia
Fisheries Centre
2204 Main Mall
Vancouver BC V6T 1Z4
CANADA
Ph: 604/822-1864
marcello@fisheries. com

Vaughan, Douglas
NOAA/NMFS Beaufort Laboratory
101 Pivers Island Rd.
Beaufort NC 28516-9722
Ph: 919/728-8761; Fax: 919/728-8784
dvaughan@hatteras.bea.nmfs.gov

Vining, Ivan
Alaska Dept. of Fish and Game
211 Mission Rd.
Kodiak AK 99615
IVANV@fishgame.state.ak.us

Wilcock, John
Alaska Dept. of Fish and Game
PO Box 669
Cordova AK 99574
Ph: 907/424-3212; Fax: 907/424-3235
johnwi@state.ak.us

Williams, Erik
University of Alaska Fairbanks
SFOS
11120 Glacier Highway
Juneau AK 99801
Ph: 907/465-5389; Fax: 907/465-6320
ftehw@aurora.alaska.edu

Witherell, Dave
North Pacific Fishery Mgmt. Council
605 W 4th Ave.
Anchorage AK 99501
Ph: 907/271-2809; Fax: 907/271-2817
David.Witherell@noaa.gov

Witzig, John
NOAA/NMFS F/ST2
1315 East-West Highway
Silver Spring MD 20910
Ph: 301/713-2363; Fax: 301/713-2313
john.witzig@noaa.gov

Woodby, Douglas
Alaska Dept. of Fish and Game
P.O Box 240020
Douglas AK 99824-0020
dougw@fishgame.state.ak.us

Xue, Qian-Li
The Johns Hopkins University
615 N. Wolfe St.
Baltimore MD 21205
Ph: 410/732-3865
qxue@athena.jhsph.edu

Yin, Yanshui
Oregon State University
Coastal OR Marine Experiment Sta. &
Dept. of Fisheries & Wildlife
Newport OR 97365
Ph: 541/737-4012; Fax: 503/867-0105
yiny@ucs.orst.edu

Zhan, Bing-Yi
Shanghai Fisheries University
334 Jungong Rd.
Shanghai 200090
CHINA
Ph: 86 21 65431090x38

Zhao, Boxian
SC Dept. of Natural Resources
217 Fort Johnson Rd.
Charleston SC 29412
Ph: 803/762-5653; Fax: 803/762-5110
zhaob@mrd.dnr.state.sc.us

Zheng, Jie
Alaska Dept. of Fish and Game
PO Box 25526
Juneau AK 99802-5526
Ph: 907/465-6102; Fax: 907/465-2604
jiez@fishgame.state.ak.us

Index