Version of Record: https://www.sciencedirect.com/science/article/pii/S0165783620300205
Manuscript_74c2787d0256c4198c0f3fdc55973088

A review of approaches to quantifying uncertainty in fisheries stock assessments

KRISTIN M. PRIVITERA-JOHNSON* AND ANDRÉ E. PUNT
School of Aquatic and Fishery Sciences, University of Washington, Box 355020, Seattle, WA 98195-5020, USA
** Corresponding author*

A review of approaches to quantifying uncertainty in fisheries stock assessments

KRISTIN M. PRIVITERA-JOHNSON* AND ANDRÉ E. PUNT
School of Aquatic and Fishery Sciences, University of Washington, Box 355020, Seattle, WA 98195-5020, USA
** Corresponding author*

Abstract

Scientific uncertainty affects all parts of the fisheries management process. This study reviews methods for quantifying scientific uncertainty for presentation as part of the scientific advice to fisheries managers. We surveyed stock assessment scientists to a) identify the methods commonly used to quantify uncertainty, b) describe how method use has changed over time, c) investigate the factors that influence which methods are used, and d) characterize how scientific uncertainty is presented to fisheries managers. We found that scientific uncertainty is being quantified and included in scientific advice across multiple fishery management systems. Frequentist approaches for quantifying uncertainty are used more broadly than Bayesian approaches, and the survey did not detect this changing over time. Time restrictions and methodology requests during the scientific review process were commonly reported as factors influencing the use of uncertainty methods. Uncertainty in estimates of management targets (e.g., fishing mortality or biomass), projections, and catch limits were the quantities most frequently included in the scientific advice presented to fisheries managers. Methods for quantifying uncertainty and their incorporation into management advice are quickly advancing, and our approaches for reviewing progress towards clearly and explicitly communicating the sources, treatment, and impacts of uncertainty in management processes must keep pace.

Key words: scientific uncertainty; stock assessment; scientific advice

1. Introduction

Communicating uncertainty is an inescapable component of providing scientific advice for fisheries management. The advent of national commitments to a precautionary approach for the conservation and management of ecological resources was championed by organizations such as the Food and Agriculture Organization (FAO) in the 1990s and inspired a thorough review of methods for quantifying uncertainties in stock size, stock productivity, reference points, and fishing mortality by Patterson et al. (2001). The efficacy of fisheries management is influenced by at least five types of uncertainty: 1) observation uncertainty, the uncertainty in measurement of observable quantities such as biomass from surveys, catch or sizes-at-age; 2) process uncertainty, the uncertainty due to underlying stochasticity in stock dynamics such as recruitment or variation in the growth of a fish stock; 3) model uncertainty, the misspecification of model parameters or structure (e.g., assuming the incorrect form for selectivity as a function of size); 4) estimation uncertainty, the inaccuracy and imprecision associated with estimated model parameters; 5) and implementation uncertainty, the variability in the implementation of management strategies (Holland and Herrera, 2009; Rosenberg and Restrepo, 1994). These uncertainties occur in all fishery systems, and affect the interpretation of data, analysis results, ranking of management options, and the efficacy of those options (Peterman, 2004). The resulting impact of uncertainty on scientific advice is critical because both overemphasis and understatement of uncertainty can undermine scientific credibility and ultimately progress towards management goals (Dankel et al., 2012). Failure to effectively account for uncertainty can lead to overshooting management targets, failing to rebuild depleted stocks, and missing opportunities to take advantage of sustainable fishing opportunities (Cadrin et al., 2015). Rosenberg (2007) suggested those who produce scientific advice for fisheries management navigate the pitfalls of blanket generalizations about uncertainty by discerning “the almost certain from the less certain”.

It is convenient to consider two classes of uncertainty when discussing the quantification of uncertainty: scientific uncertainty (i.e., observation, process, model, and estimation uncertainties) and management uncertainty (i.e., implementation uncertainty). The focus of this paper is on methods for and applied examples of quantifying scientific uncertainty, as these dominate the literature and are general across jurisdictions and taxa. Identifying the widely used tools and methods for quantifying scientific uncertainty and the frequency of use over time, fish stocks, and regions can contribute to the continued development of best practices. Understanding the factors influencing the use of a tool or method can inform the allocation and development of resources to better quantify uncertainty.

During this exploration of methods and tools for quantifying uncertainty, we will use the following definitions of key concepts. The *fisheries management process* consists of data collection, analysis, scientific review, provision of scientific advice, decision-making, setting of catch limits, and enforcement (FAO, 1997). *Jurisdictions* are the organizations (e.g., single governmental, multi-national governmental, and non-governmental) designing and implementing the fisheries management process. A *stock assessment* is a process that includes the activities, analyses, and reports related to the data collection, analysis, and scientific review components of the fishery management process (PFMC, 2018). More specifically, the *analysis* process of a stock assessment applies statistical and mathematical models to use different data sources (e.g., survey, fishery, biological) to make quantitative predictions about the abundance and trends of fish stocks and of fishing intensity (Hilborn and Walters, 1992). Scientific uncertainty can be quantified using frequentist and Bayesian paradigms of statistical inference (hereby deemed as *uncertainty methods*). *Sensitivity analyses* elucidate how these uncertainties propagate through an assessment model and can be apportioned to sources of uncertainty in the model inputs and parameter values (Satelli, 2002; Steel et al., 2009). The modeling frameworks designed for stock assessment analysis, the uncertainty methods, and sensitivity analyses can be assembled into *packages* (i.e., well-documented software repositories) to be downloaded and installed on a computer for reproducible analyses.

Dichmont et al. (2016a) assert that assembling stock assessment modeling frameworks into packages is integral for increasing access to tools for quantifying uncertainty. The advantages of assessment packages include that open access to such packages facilitates exploration of multiple assessment configurations and strengthens the peer-review process. However, implementing a new model for a stock using packages developed for different, specific stocks presents challenges such as dealing with the "black box" effect when debugging potential errors and the steep learning curve for packages with many options. This meta-analytic approach to characterizing package use in U.S. fisheries management lends itself well to exploration of other analysis components such as methods for quantifying scientific uncertainty.

We apply a similar meta-analytic approach to Dichmont et al. (2016a) to summarize the methods that produce model outputs used to communicate scientific uncertainty to fisheries managers. Specifically, we are interested in the methods used for quantifying uncertainty within a given assessment framework and across such frameworks. We surveyed stock assessment scientists to investigate the following: 1) what methods for quantifying uncertainty are used?; 2) how have methods changed over time?; 3) what are the most common factors that influence the use of a specific method?; and 4) how are scientific uncertainties presented to fisheries managers?

2. Methods

The survey addressed each research question through the use of multiple choice and free response questions (see Supplementary Figs 1- 7 for survey questions). Participants in the survey were asked to state the assessment tools (e.g., packages) they have used, the approaches used for quantifying scientific uncertainty while conducting assessments, and the quantities of interest used in sensitivity analyses (Supplementary Figs 2-4). To characterize how method and tool use has changed over time, participants were asked to provide the tools, analyses, and approaches used in a (subjective) representative sample of the assessments they have conducted (Supplementary Figs 6-7). To identify factors that may influence method and tool use, analysts were asked which available methods for quantifying uncertainty were not used, which quantities of interest could have been considered for uncertainty evaluation but were not, and why (Supplementary Figs 3-4). Finally, participants were asked how they have presented uncertainties to fishery managers (Supplementary Fig. 5).

The survey was distributed to scientists who have conducted stock assessments and provided scientific advice to management. Survey participants (N=68) have provided scientific advice for many organizations around the world. Respondents self-defined the numbers of years worked as a stock assessment scientist, and these ranged from 1 to 38 years (Fig. 1).

We asked survey respondents to provide information for some representative stock assessments they have conducted over the last 5-10 years. This included the common and scientific names of the stock, the agency for which the assessment was conducted, the year the assessment was conducted, the packages used, the data types used, the sensitivity analyses conducted, and the uncertainty methods used. Survey respondents were invited to list the uncertainty methods, sensitivity analyses, and packages featured in the survey and any additional analyses and analysis methods. The resulting time series covered 1997 to 1999 (ranging from 1 to 3 assessments each year) and 2002 to 2018 (ranging from 1 to 64 assessments each year). Originally, there were 372 assessments reported. However, there were cases of repeat assessments because multiple assessment authors who have worked on the same assessments were surveyed. The information was collated across respondents and resulted in 353 individual assessments.

3. Results

3.1 Representativeness of the survey results

Our survey reviewing methods for quantifying scientific uncertainty has notable limitations. The representation of agencies and regions is not evenly distributed, as 35 of the 68 respondents were based at the U.S. National Marine Fisheries Service. However, the assessment scientists and the stocks they reported working with fall within 17 of the 18 regions used by the RAM Legacy Stock Assessment Database [Ricard et al., 2012] to aggregate assessment summaries (Fig. 1). The patterns in method use presented have low sample sizes (e.g., 1-2 stocks) for assessments in the early part of the time series.

3.2 Software and model framework

Identifying the packages and thus the modeling frameworks used to conduct assessments allows us to draw connections between the tools available and the methods for quantifying uncertainty—i.e., how scientists are discerning "the almost certain from the less certain". Survey respondents were asked to identify software they use (and have used) in the process of conducting a stock assessment. The provided list of available software featured 23 options (Table 1; Supplementary Fig. 2) and the survey responses identified an additional 23 (Table 2). The packages used in the provided assessments (N=353) were sorted into the following modeling frameworks: surplus production models (N=2), virtual population analyses (VPA; N=17), age-structured models (N=244), length-structured models (N=19), depletion models (N=9), depletion-based stock reduction analyses (DB-SRA; N=3), and not specified (N=59). Not specified consisted of responses with package descriptions that did not indicate the model framework used for an assessment (e.g., "User-written ADMB code"). The use of the VPA model framework decreased over time, the use of length-structured models increased in most recent years, and age-structured models were used consistently throughout the time period surveyed (Fig. 2, panel a). Not enough information was provided to describe trends in surplus production, depletion, or DB-SRA models over time (Fig. 2, panel a). Assessments developed using age-structured models used the most methods for quantifying uncertainty (Fig. 3). Frequentist uncertainty methods were used across all frameworks except DB-SRA, with asymptotic methods being the most used frequentist approach (Fig. 3).

3.3 Structural models and estimation methods

The patterns in use of sensitivity analyses and statistical inference paradigms (and the drivers of such patterns) can influence the type and complexity of information to present with scientific advice to management.

Sensitivity analysis help understand some aspects of model uncertainty. When asked if they utilize sensitivity analyses to directly quantify uncertainty, 38 respondents provided a response: 20 respondents reported yes, 16 reported no, and 2 reported sometimes. The participants stated that they used sensitivity analyses to qualitatively characterize uncertainty, i.e., as a "2nd tier of uncertainty" to be used in conjunction with other methods (e.g., management strategy evaluation and Bayesian methods). Sensitivity tests can be used to capture some aspects of model uncertainty when providing management advice; for example, when defining states of nature and bracketing ranges of plausible outcomes when important elements of uncertainty cannot be incorporated directly into a model (e.g., if one cannot estimate natural mortality, steepness, or catch uncertainty). These responses also highlighted the strength of sensitivity testing as a qualitative tool useful for representing extremes to demonstrate model behavior and assess the robustness of model results to baseline assumptions and assumed values for model parameters.

Participants were asked if they used frequentist (i.e., asymptotic methods, bootstrapping, jackknife, and likelihood profiles) and Bayesian (i.e., Adaptive Importance Sampling (AIS), Markov Chain Monte Carlo (MCMC), Sample-Importance-Resample (SIR)) methods to quantify process and estimation uncertainty. For the frequentist methods, asymptotic methods and likelihood profiles were selected most frequently, followed by bootstrapping, and jackknife (Fig. 4, panel a). The dominant Bayesian approach was MCMC and its many variants (N=48) (Fig. 4, panel b). The survey did not detect a substantial change in estimation method use over time (Fig. 5). Additional methods for quantifying uncertainty provided by survey respondents were decision tables, ensemble modeling, retrospective analyses, and the Approximate Bayesian Computation.

Respondents referred to assessing the “performance” of models using retrospective analyses of base models and previous assessments of the same stock, and models of various levels of complexity (e.g., fitting a production model as well as a model that includes all of the data).

3.4 Model specification and sensitivity analyses

Evaluating the sensitivity of the outcomes of an assessment to the specifications of the model on which it is based is integral for the prevention of overemphasis or understatement of uncertainty and maintaining progress toward management goals and scientific credibility. We asked respondents if they routinely conduct sensitivity analyses based on alternative catch streams, and on assumptions about catchability, growth, maturity, natural mortality, recruitment (e.g., fixed values for stock-recruitment steepness), selectivity parameterization, the stock-recruit relationship (e.g., a Beverton-Holt or Ricker parameterization) and data set choice, and data weighting. The quantities of interest investigated for sensitivity analyses did not change over time (Fig. 6). Of the provided list, data weighting was the most selected option and maturity was the least (Fig. 7). Additional sensitivities fell into two categories: data processing and changes to structural assumptions. Sensitivities involving data processing (related in part to observation uncertainty) included the range of years of data used for specific data sets and how they are used (e.g., a survey using different sampling methods in different years), the binning of length compositions, alternative survey indices (e.g., design- vs. model-based), use of tagging data, how survey data are aggregated over space, area-stratified vs. spatially lumped, and alternative assumptions regarding ageing imprecision and time-varying selectivity. Model structure sensitivity analyses (i.e., model uncertainty) involved comparing results using different stock assessment packages (e.g., personalized ADMB model, Stock Synthesis, and SAM) [see Table 2 for example references], the number of growth morphs (in a Stock Synthesis assessments), whether the model is single- or two-sex, the number of areas, fleet structure, temporal step, alternative time ranges for the assessment movement/migration assumptions, likelihood distribution assumptions, proportional vs. non-proportional relationships between catch-rate and abundance, amount of fishing prior to the start of the data series, cetacean depredation, and illegal, unreported, and unregulated fishing trends.

3.5 Presentation of uncertainty to fishery managers

Determining the most common assessment outputs used for producing scientific advice for management may reveal how the methods for quantifying uncertainty and the information requested by fisheries managers overlap. Survey participants were given eight options for assessment outputs used to communicate scientific uncertainty to fishery managers: estimates of fishing mortality and/or biomass; estimates of fishing mortality and/or biomass relative to reference points; the results of simulation testing; the results of management strategy evaluations; decision tables; values for catch limits (e.g., Total Allowable Catch (TAC), Acceptable Biological Catch (ABC), Overfishing Limit (OFL)); projections under uncertainty; and other (Fig. 8). The precision of estimates of incoming year class strengths (i.e., recruitment) and the results of ensemble models were suggested as additional ways to communicate uncertainty by the respondents. Several respondents reported that while they have presented many of these model outputs to fisheries managers and their scientific review bodies, there are cases when the information (and its associated estimates of uncertainty) have not been used in fisheries management.

4. Discussion

Scientific uncertainty is being quantified and included in scientific advice across multiple fisheries management systems. Frequentist approaches for quantifying process and estimation uncertainty are used more broadly than Bayesian approaches, and the survey did not detect this trend changing over time. This is also reflected in the prolific use of packages using asymptotic methods for estimating uncertainty, which has qualitatively increased over time (in particular, those based on ADMB). Similarly, there has been little change in the quantities of interest investigated for sensitivity analyses over time, supporting Maunder and Piner's (2015) statement that successful interpretation of data requires knowledge of growth, recruitment, natural mortality, selectivity, and sampling processes for the stock—knowledge that remains incomplete for most stocks and regions. Time restrictions and methodology requests during the scientific review process were commonly reported as factors influencing the use of uncertainty methods (more below). Uncertainty in estimates of management targets (e.g., fishing mortality or biomass), projections, and catch limits were the quantities most frequently presented to managers. Survey respondents also expressed that not all uncertainties that are quantified are presented and not all those presented are used by managers in the decision-making process.

Ultimately, asking assessment scientists what factors influence their use of specific approaches to quantifying scientific uncertainty revealed a common theme: the design of the fisheries management system. The priorities of jurisdictions designing each component of this cycle vary with their respective values, economic structures, and political traditions (Marchal et al., 2016). At the heart of the fisheries management system lies the mission to have a transparent process operating with the utmost integrity to strengthen stakeholder confidence in the decisions being informed by scientific advice. Failure to explicitly define the roles and responsibilities of managers and scientists presents opportunities for certain sources of uncertainty to not be properly identified (Cadrin et al., 2015). The definition and communication of these roles and responsibilities is a dynamic process that changes as the fishery management system encounters new situations and experiences changes in decision-making participants and government structures (Francis and Shotton, 1997).

Many jurisdictions create and implement review protocols to meet their goals and avoid the above pitfalls, which directly influence the methods used to quantify uncertainty. This can manifest as specific methodology requests for conducting assessments, quantifying uncertainty, and presentation of scientific advice. The ICES uses a Generic Terms of Reference for many of its stock assessment working groups. Each stock assessment working group applies an assessment model framework that is either analytical, forecast, or based on trend indicators, and the final report is requested to address the following: input data and data quality; catch misreporting; percent of total catch taken in a regulatory area; if applicable, estimates of maximum sustainable yield proxy reference points; the status of the stock relative to reference points; projected catch scenarios; and historical and analytical performance of the assessment and catch options (ICES, 2018). In the U.S., the Pacific Fishery Management Council has specific requests for the evaluation of uncertainty in assessment results for U.S. West Coast groundfish and coastal pelagic species stocks: model specification uncertainty; parameter uncertainty (including likelihood profiles); retrospective analysis; historical analysis; probability statements for ranges of model runs; and for groundfish at least three states of nature for model ranges (i.e., most probable, lower biomass trajectory, and high biomass trajectory) (PFMC, 2018).

The time allocated to conduct and review a stock assessment varies by jurisdiction and directly influences the methods used to quantify uncertainty. The amount of time available to conduct and review a stock assessment to prepare scientific advice depends on resource availability (e.g., external reviewers) and the timetable for making short-term management decisions (e.g., setting catch limits). In some regions of the U.S., assessments are conducted over the course of a few months and are reviewed over a short time period (e.g., 5 days for U.S. West Coast) and any additional model requests must be performed within this time frame (PFMC, 2018). In other management systems such as New Zealand, fish-stock assessment groups meet daily over the course of weeks or months to conduct the assessment and respond to scientific review feedback (Marchal et al., 2009). ICES stock assessment working groups meet for 5-10 days to complete and review assessments (Marchal et al., 2009). The combination of requested methodology and the time available for conducting and reviewing assessments may not leave stock assessment scientists with enough time to run full Bayesian analyses for quantifying process and estimation uncertainty. However, advances in optimization approaches in software such as ADMB (e.g., Hamiltonian No U-Turn Samplers) may reduce this bottleneck in analysis run time enough to influence the use of uncertainty methods for management advice in the future (Monnahan et al., 2019).

The frequency of assessment for a stock (i.e., stock prioritization) also influences time restrictions and may indirectly impact the use of specific uncertainty methods. Stock prioritization generally relates to the total number of stocks and species assessed by a jurisdiction, relative commercial importance of the stock, and data availability (Marchal et al., 2009; Methot, 2015). Depending on the stock, assessment frequency may range from once a year to once every 10 or more years. Jurisdictions with many stocks may have longer gaps between assessments for a single stock because limited resources (e.g., number of available assessment scientists) may restrict the number of assessments that can be conducted annually. The stocks and species assessed may also rotate over time. Stock prioritization procedures (informally and formally defined) decide how this stock rotation occurs and fisheries scientists and managers should collaborate to design procedures that “focus limited resources where they are most needed to reduce uncertainty” (Cadrin et al., 2015). Given these constraints, incorporating major changes to model structure and uncertainty methodology may not be feasible every assessment cycle, especially if there are long gaps since the last assessment for a stock, new data considerations, and requested methodology from scientific review committees.

Summarizing the influence of management design highlights opportunities to expand the repertoire for quantifying uncertainty for use in the development of scientific advice. Survey respondents reported that the uncertainty methods and sensitivity analyses they employ often differ between assessment reports for tactical management and research publications. Scientists should continue to explore and test alternative hypotheses in a research context and integrate the reliable approaches into the packages (new and existing) and requested methods used to inform management. Using packages that have been previously reviewed and approved by the scientific review committees has the potential to alleviate some of the burden of the review process and may promote more effective communication of results (Dichmont et al., 2016a). The expansion of current packages (e.g., Stock Synthesis and CASAL) can include the addition of spatially-structured population dynamics models, incorporation of non-traditional data types (e.g., tagging data, and habitat information), and integration of economic models (e.g., Australian fisheries requiring management advice related to maximum economic yield) (Dichmont et al., 2016b). Continued and expanded focus on cooperative research opportunities such as courses (e.g., Advanced School on Multispecies Modelling Approaches for Ecosystem Based Marine Resource Management in the Mediterranean Sea (AMARE-ED), www.echo.inogs.it/amare-med/; ICES training courses, www.ices.dk/news-and-events/Training/) and workshops (e.g., National Stock Assessment Workshops, www.st.nmfs.noaa.gov/stock-assessment/workshops; Center for the Advancement of Population Assessment Methodology (CAPAM), www.capamresearch.org) are integral for "leveling the playing field" by fostering environments for the development and dissemination of new methods for quantifying scientific uncertainty for use in scientific advice across jurisdictional boundaries (Cadrin et al., 2015; Dichmont et al., 2016b).

Using meta-analytic approaches to characterize how uncertainty permeates through fisheries management systems around the world can also be further developed. The set of methods for quantifying management uncertainty for use in advice for fisheries management has increased in the last decade (e.g., Dichmont et al., 2006; Fulton et al., 2011; Sethi et al., 2005) and exploring how factors such as fisheries management system design influences the development and implementation of these methods is a promising area of future research. Investigating how uncertainties are presented in scientific advice across jurisdictions (e.g., the Kobe framework [Kell et al., 2016]) may provide insight about how to progress effective communication of uncertainty in a field producing increasingly complex and multidimensional management advice to a broad audience of stakeholders. Survey approaches *sensu* Levontin et al. (2017) complement this effort by evaluating the reliability of our visualization of modeling approaches. The authors suggest that repositories of stock assessment results begin routinely storing uncertainty measures in addition to point estimates. Methods for quantifying uncertainty and their incorporation into management advice is quickly advancing and our approaches for reviewing our progress towards clearly and explicitly communicating the sources, treatment, and impacts of uncertainty in our management processes must keep pace.

Acknowledgements

This research has been supported by the University of Washington School of Aquatic and Fishery Sciences and the National Science Foundation. The authors warmly thank the 68 stock assessment scientists for their responses to our survey. We also thank J. Cope (NOAA, NWFSC), M. Haltuch (NOAA, NWFSC), and J. Valero for their feedback during the development of the survey. J. Cope (NOAA, NWFSC) and T. Essington (University of Washington) are thanked for their mentorship and early manuscript review. Two anonymous reviewers and the editor Steven Cadrin (U Mass, Dartmouth) are thanked for their comments on an earlier version of this paper. KPJ was funded by the National Science Foundation Graduate Research Fellowship Program. AEP was partially funded by the Joint Institute for the Study of the Atmosphere and Ocean (JISAO) under NOAA Cooperative Agreement NA15OAR4320063, Contribution No. _____

References

Brodziak, J., Rago, P., Conser, R., 1998. A general approach for making short-term stochastic projections from an age-structured fisheries assessment model. *In* Proceedings of the International Symposium on Fishery Stock Assessment Models for the 21st Century. Edited by F. Funk, T. Quinn II, J. Heifetz, J. Ianelli, J. Powers, J. Schweigert, P. Sullivan, and C.-I. Zhang. Alaska Sea Grant College Program, University of Alaska, Fairbanks. pp. 933–954.

Bull, B., Francis, R.I.C.C., Dunn, A., McKenzie, A., Gilbert, D.J., Smith, M.H., Bain, R., Fu, D., 2012. CASAL (C++ algorithmic stock assessment laboratory): CASAL user manual v2.30-2012/03/21. NIWA Technical Report 135. 280 p.

Cadrin, S., Henderschedt, J., Mace, P., Mursalski, S., Powers, J., Punt, A.E., Restrepo, V., 2015. Addressing Uncertainty in Fisheries Science and Management. National Aquarium. http://www.fao.org/3/a-bf336e.pdf

Collie, J.S., Sissenwine, M.P., 1983. Estimating population size from relative abundance data measured with error. Can. J. Fish. Aquat. Sci., 40, 1871–1879.

Dankel, D.J., Aps, R., Padda, G., Röckmann, C., van der Sluijs, J.P., Wilson, D.C., Degnbol, P., 2012. Advice under uncertainty in the marine system. ICES J. Mar. Sci. 69, 3–7.

Darby, C.D., Flatman, S., 1994. Virtual population analysis: Version 3.1 (Windows/DOS) User Guide. MAFF Directorate of Fisheries Research IT Report 1. 85 pp.

Davies, N.M., Gilbert, D.J., McKenzie, J.R., 2001. Length-based growth estimates for application in an integrated age and length structured population model. Final research report for Ministry of Fisheries Research Project SNA1999/01 Objective 1. National Institute of Water and Atmospheric Research. https://fs.fish.govt.nz/Page.aspx?pk=113&dk=22515

de la Mare, W.K., and Cooke, J.G. (1993). BALEEN II: The population model used in the Hitter-Fitter programs. Manuscript available from the IWC Secretariat.

De Oliveira, J., Darby, C.D., Roel, B.A., 2010. A linked separable-ADAPT VPA assessment model for western horse mackerel (*Trachurus trachurus*), accounting for realized fecundity as a function of fish weight. ICES J. Mar. Sci. 67, 916–930.

Dichmont, C.M., Deng, R.A., Punt, A.E., Venables, W., Haddon, M., 2006. Management strategies for short-lived species: the case of Australia's northern prawn fishery: 1. Account for multiple species, spatial structure and implementation uncertainty when evaluating risk. Fish. Res. 82, 204–220.

Dichmont, C.M., Deng, R.A., Punt, A.E., Brodziak, J., Chang, Y., Cope, J.M., Ianelli, J.N., Legault, C.M., Methot, R.M., Porch, C.E., Prager, W.H., Shertzer, K.W., 2016a. A review of stock assessment packages in the United States. Fish. Res. 183, 447–460.

Dichmont, C.M., Deng, R.A., Punt, A.E., 2016b. How many of Australia's stock assessments can be conducted using stock assessment packages? Mar. Policy. 74, 279–287.

Dick, E. J., and MacCall, A.D., 2011. Depletion-Based Stock Reduction Analysis: A catch-based method for determining sustainable yields for data-poor fish stocks. Fish. Res. 110, 331–341.

FAO., 2007. Fisheries management. Technical Guidelines for Responsible Fisheries. 4: 1-82. FAO, Rome. http://www.fao.org/3/a-w4230e.html

FAO., 2013. Fisheries and aquaculture software. FISAT II - FAO-ICLARM Stock Assessment Tool. In: FAO Fisheries and Aquaculture Department. Rome. Updated 28 November 2013. http://www.fao.org/fishery/topic/16072/en

Fournier, D.A., Hampton, J., Sibert, J.R., 1998. MULTIFAN-CL: a length-based, age-structured model for fisheries stock assessment, with application to South Pacific albacore, *Thunnus alalunga*. Can. J. Fish. Aquat. Sci., 55, 2105–2116.

Fournier, D.A., Skaug, H.J., Ancheta, J., Ianelli, J., Magnusson, A., Maunder, M.N., Nielsen, A., Sibert, J., 2012. AD Model Builder: using automatic differentiation for statistical inference of highly parameterized complex nonlinear models. Optim. Methods Softw. 27, 233–249.

Fox, W., 1975. Fitting the generalized stock production by least-squares and equilibrium approximation. Fish. Bull., 73, 23–37.

Francis, R.I.C.C., Shotton, R., 1997. "Risk" in fisheries management: a review. Can. J. Fish. Aquat. Sci. 54, 1699–1715.

Fulton, E.A., Smith, A.D.M., Smith, D.C., van Putten, I.E., 2011. Human behavior: the key source of uncertainty in fisheries management. Fish. Fish. 12, 2–17.

Gadget. 2020. Gadget development repository. http://hafro.github.io/gadget/. (Accessed 11 January 2020).

Gayanilo, F., Sparre, P., Pauly, D., 1994. The FAO-ICLARM stock assessment tools (FISAT) User's guide. FAO computerised information series: Fisheries. 1048.

Goodyear, C., 2004a. FSIM-a simulator for forecasting fish population trends and testing assessment methods. Collective Volume Scientific Papers ICCAT. 56, 120–131.

Goodyear, C., 2004b. A data simulator for testing alternative longline CPUE standardization methods. Col. Vol. Sci. Pap. ICCAT. 56, 132–145.

Haddon, M., 2011. Modelling and quantitative methods in fisheries. CRC Press. Boca Raton, FL.

Hilborn, R., Walters, C., 1992. Quantitative fisheries stock assessment: Choice, dynamics and uncertainty. Chapman and Hall.

Hilborn, R., Mangel, M., 1997. The ecological detective. Princeton University Press.

Hilborn, R., Maunder, M., Parma, A., Ernst, B., Payne, J., Starr, P., 2003. Coleraine: A generalized age-structured stock assessment model. User's manual version 2.0. University of Washington Report SAFS-UW-0116. http://fish.washington.edu/research/coleraine/coleraine.pdf.

Hoggarth, D. D., Abeyasekera S., Arthur R.I., 2006. FAO Fish. Tech. Pap. Stock assessment for fishery management. FAO, Rome, 487p.

Holland, D.S., Herrera, G.E., 2009. Uncertainty in the management of fisheries: contradictory implications and a new approach. Mar. Resour. Econ. 24, 289–299.

ICES., 2018. 2019 ACOM and ACOM Expert Group Terms of Reference. ICES AC. November 6, 2018.

Jardim, E., Millar, C. P., Mosqueira, I., Scott, F., Osio, G. C., Ferretti, M., Alzorriz, N., Orio, A., 2014. What if stock assessment is as simple as a linear model? The a4a initiative. ICES Journal of Marine Science, 72: 232-236.

Kell, L.T., Levontin, P., Davies, C.R., Harley, S., Kolody, D.S., Maunder, M.N., Mosqueira, I., Pilling, G.M., Sharma, R., 2016. The quantification and presentation of risk. Management Science in Fisheries: An Introduction to Simulation-Based Methods. Oxford, UK: Earthscan (Routledge), 348.

Levontin, P., Baranowski, P., Leach, A.W., Bailey, A., Mumford, J.D., Quetglas, A., Kell, L.T., 2017. On the role of visualization in fisheries management. Mar. Policy. 78, 114–121.

MacCall, A.D., 2009. Depletion-corrected average catch: a simple formula for estimating sustainable yields in data-poor situations. ICES J. Mar. Sci., 66, 2267–2271.

Magnusson, A., Punt, A.E., Hilborn, R., 2013. Measuring uncertainty in fisheries stock assessment: the delta method, bootstrap, and MCMC. Fish. Fish. 14, 325–342.

Majkowski, J., 1982. Usefulness and applicability of sensitivity analysis in a multispecies approach to fisheries management. 149–165. In Pauly, D. and G.I. Murply (eds) Theory and management of tropical fisheries. ICLARM Conference Proceedings 9, 360 p. International Center for Living Aquatic Resource Management, Manila, Philippines and Division of Fisheries Research, Commonwealth Scientific and Industrial Research Organization, Cronulla, Australia.

Marchal, P., Lallemand, P., Stokes, K., Thébaud, O., 2009. A comparative review of the fisheries resource management systems in New Zealand and in the European Union. Aquat. Living Resour. 22, 463–481.

Marchal, P., Andersen, J.L., Aranda, M., Fitzpatrick, M., Goti, L., Guyader, O., Haraldsson, G., Hatcher, A., Hegland, T.J., Le Floc'h, P., Macher, C., Malvarosa, L., Maravelias, C.D., Mardle, S., Murillas, A., Nielsen, J.R., Sabatella, R., Smith, A.D.M., Stokes, K., Thoegersen, T., Ulrich, C., 2016. A comparative review of the fisheries management experiences in the European Union and in other countries worldwide: Iceland, Australia, and New Zealand. Fish. Fish. 17, 803–824.

Martell, S.J., Schweigert, J.F., Haist, V., Cleary, J.S., 2012. Moving towards the sustainable fisheries framework for Pacific herring: data, models, and alternative assumptions; Stock Assessment and Management Advice for the British Columbia Pacific Herring Stocks: 2011 Assessment and 2012 Forecasts. DFO Can. Sci. Advis. Sec. Res. Doc. 2011/136. xii + 151 p.

Maunder, M.N., Piner, K.R., 2015. Contemporary fisheries stock assessment: many issues still remain. ICES J. Mar. Sci. 72, 7–18.

McAllister, M., Babcock, E., 2006. Bayesian surplus production model with the sampling importance resampling algorithm (BSP): a user's guide.

Methot, R.D., Wetzel, C.R., 2013. Stock synthesis: a biological and statistical framework for fish stock assessment and fishery management. Fish. Res. 142, 86–99.

Methot, R.D., 2015. Prioritizing fish stock assessments. U.S. Dep. Commer., NOAA Tech. Memo. NFMS-F/SPO-152, 31 p.

Monnahan, C.C., Branch, T.A., Thorson, J.T., Stewart, I.J., Szuwalski, C.S., 2019. Overcoming long Bayesian run times in integrated fisheries stock assessments. ICES J. Mar. Sci. 00, 00–00.

Muppet. 2020. Muppet development repository. http://github.com/hafro/muppet. (Accessed 11 January 2020).

Nielsen, A., Berg, C.W., 2014. Estimation of time-varying selectivity in stock assessments using state-space models. Fish. Res. 158, 96–101.

NOAA Toolbox, 2016. NOAA Fisheries Toolbox. https://www.nefsc.noaa.gov/nft/ (accessed 14 August 2019).

Pacific Fishery Management Council, 2018. Terms of Reference for the groundfish and coastal pelagic species stock assessment review process for 2019-2020. Pacific Fishery Management Council, 7700 Ambassador Place NE, Suite 200, Portland, Oregon.

Patterson, K., Cook, R., Darby, C., Gavaris, S., Kell, L., Lewy, P., Mesnil, B., Punt, A., Restrepo, V., Skagen, D.W., Stefansson, G., 2001. Estimating uncertainty in fish stock assessment and forecasting. Fish. Fish. 2, 125-157.

Payá, I., 2019. Stock Assessment Programs. https://sites.google.com/site/fishassessement/stock-assessment-programs (accessed 14 August 2019).

Pedersen, M.W., Casper, W.B., 2017. A stochastic surplus production model in continuous time. Fish. Fish. 18, 226–243.

Peterman, R.M., 2004. Possible solutions to some challenges facing fisheries scientists and managers. ICES J. Mar. Sci. 61, 1331–1343.

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

Porch, C., 2017. NOAA Tech. Memo. NMFS-SEFSC-708 PRO-2BOX 3.0 User Guide. 10.13140/RG.2.2.12258.17604.

Prager, M.H., 1994. A suite of extensions to a nonequilibrium surplus–production model. Fish. Bull. 92, 374–389.

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

Punt, A.E., Walker, T.I., 1998. Stock assessment and risk analysis for the school shark (*Galeorhinus galeus*) off southern Australia. Mar. Freshwater Res. 49, 719–731.

Punt, A.E., 2015. Strategic management decision-making in a complex world: quantifying, understanding, and using trade-offs. ICES J. Mar. Sci., 74, 499–510.

Ricard, D., Minto, C., Jensen, O.P. and Baum, J.K. 2012. Evaluating the knowledge base and status of commercially exploited marine species with the RAM Legacy Stock Assessment Database. Fish. Fish. 13, 380-398.

Roel, B.A., De Oliveira, J.A.A., Beggs, S., 2009. A two-stage biomass model for Irish Sea herring allowing for addition-al variance in the recruitment index caused by mixing of stocks. ICES J. Mar. Sci, 66, 1808–1813.

Rosenberg, A.A., Restrepo, V.R., 1994. Uncertainty and risk evaluation in stock assessment advice for U.S. marine fisheries. Can. J. Fish. Aquat. Sci. 51, 2715–2720.

Rosenberg, A.A., 2007. Fishing for certainty. Nature. 449: 989.

Satelli, A., 2002. Sensitivity analysis for importance assessment. Risk Analysis. 22, 1–12.

Sethi, G., Costello, C., Fisher, A., Hanemann, M., Karp, L., 2005. Fishery management under multiple uncertainty. J. Environ. Econ. Manag. 50, 300–318.

Steel, A.E., McElhany, P., Yoder, N.J., Purser, M.D., Malone, K., Thompson, B.E., Avery, K.A., Jensen, D., Blair, G., Busack, C., Bowen, M.D., Hubble, J., Kantz, T., 2009. Making the best use of modeled data: Multiple approaches to sensitivity analysis of a fish-habitat model. Fisheries. 34, 330–339.

Stewart, I.J., Hicks, A.C., Taylor, I.G., Thorson, J.T., Wetzel, C., Kupschus, S., 2013. A comparison of stock assessment uncertainty estimates using maximum likelihood and Bayesian methods implemented with the same model framework. Fish. Res. 142, 37–46.

Thorson, J.T., Cope, J.M., 2015. Catch curve stock-reduction analysis: an alternative solution to the catch equation. Fish. Res. 171, 33–41.

Walmsley, S.F., Howard, C.A., Medley, P.A., 2005. Participatory fisheries stock assessment (ParFish) guidelines. London: MRAG.

Williams, E.H., Shertzer, K.W., 2015. Technical documentation of the Beaufort assessment model (BAM). U.S. Department of Commerce, NOAA Tech. Memo. NMFS-SEFSC-671. 43 p.

Figure captions

Figure 1. The distribution of respondents by RAM Legacy Stock Assessment Database region (panel a) and the number of years each survey respondent has worked as a stock assessment scientist (panel b).

Figure 2. Frequencies of model framework used over time (panel a) and the number of reported assessments conducted in each assessment year (i.e., the sample size for each year of panel a) (panel b).

Figure 3. Frequencies of uncertainty methods used by model framework pooled over the 20 years of available assessments. Note that the assessment sample size is less than total number of available assessments ($N=294$ vs. $N_{total}=353$) because not all package descriptions provided by respondents indicated the model framework used (e.g., “User-written ADMB code”).

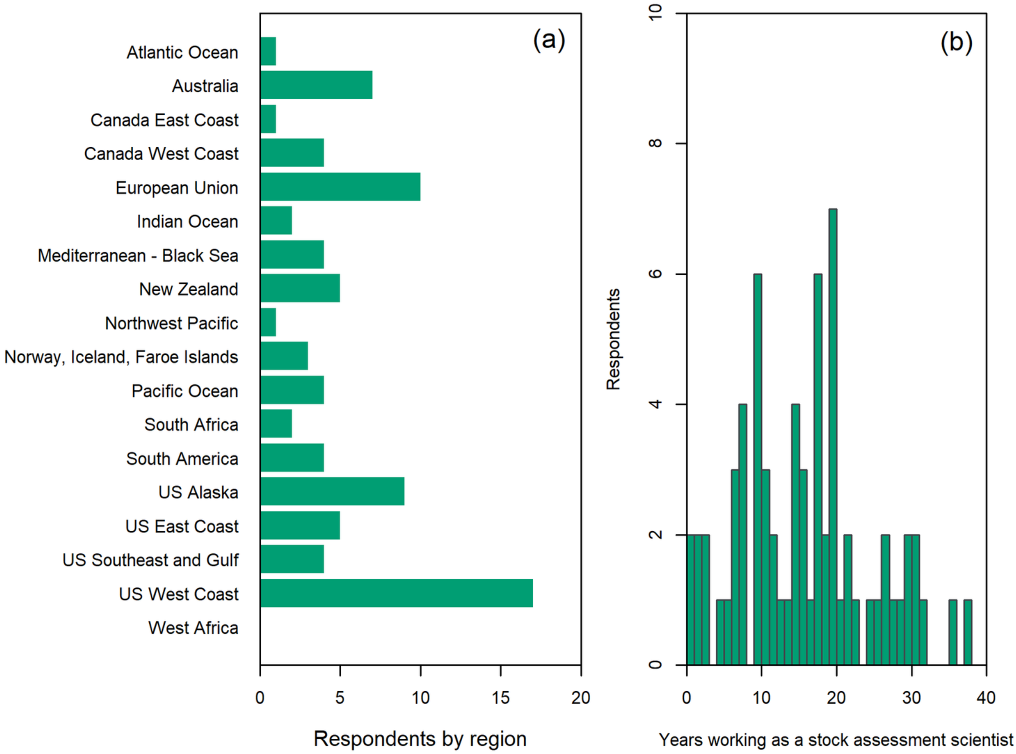
Figure 4. Frequency of use of frequentist (panel a, shades of gray) and Bayesian approaches (panel b, shades of green) for computing measures of uncertainty in assessment model outputs.

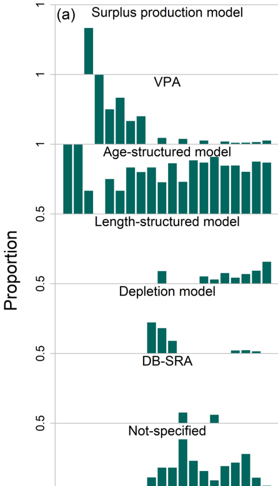
Figure 5. Frequencies of uncertainty methods used over time (panel a) and the number of reported assessments conducted in each assessment year (i.e., the sample size for each year of panel a) (panel b). Frequentist methods are the gray bars and Bayesian methods are the dark green bars.

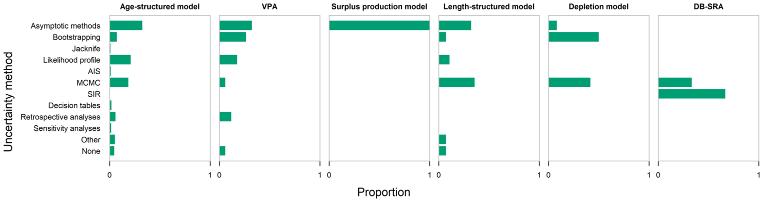
Figure 6. Frequencies of sensitivity analyses used over time (panel a) and the number of reported assessments conducted in each assessment year (i.e., the sample size for each year of panel a) (panel b).

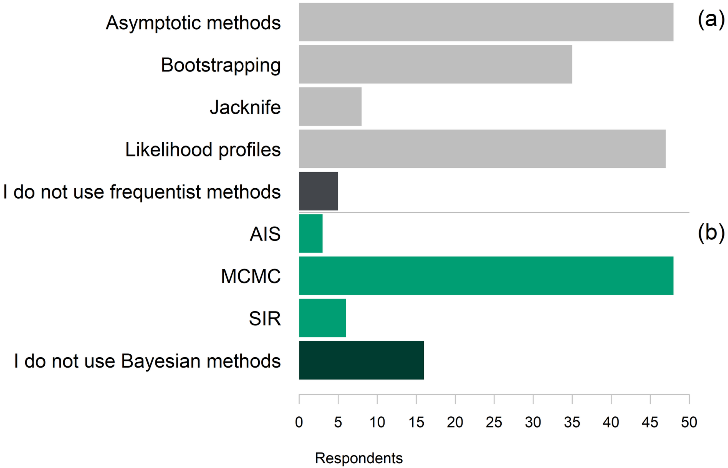
Figure 7. Frequency of sensitivity analyses not conducted during routine stock assessments (left column) and the reasons for not doing so (right column). Note that the respondents may have selected multiple reasons for not conducting sensitivity analyses and thus the bars on the right column do not sum to the bars on the left.

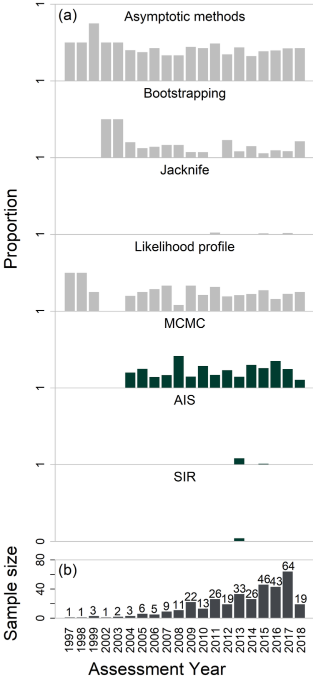
Figure 8. Frequency of quantities presented to managers (panel b). F and B represent fishing mortality and biomass, respectively.


(a)
Atlantic Ocean
Australia
Canada East Coast
Canada West Coast
European Union
Indian Ocean
Mediterranean - Black Sea
New Zealand
Northwest Pacific
Norway, Iceland, Faroe Islands
Pacific Ocean
South Africa
South America
US Alaska
US East Coast
US Southeast and Gulf
US West Coast
West Africa
0
5
10
15
20
Respondents by region
(b)
Respondents
0
2
4
6
8
10
0
10
20
30
40
Years working as a stock assessment scientist

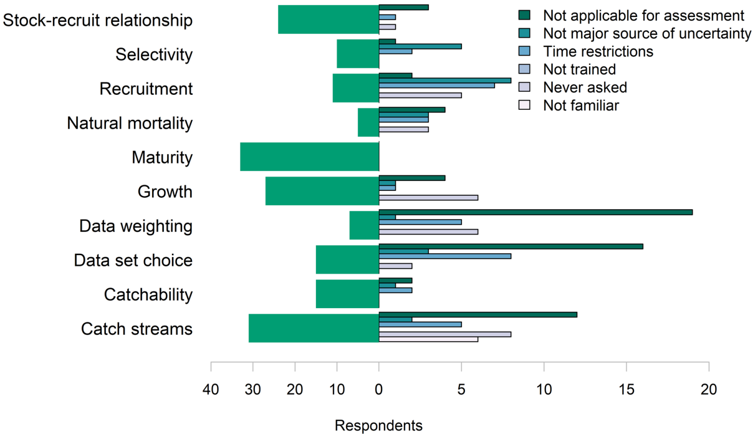

(a)
Surplus production model
VPA
Age-structured model
Length-structured model
Depletion model
DB-SRA
Not-specified
1
1
1
0.5
0.5
0.5
0.5
0
Proportion
(b)
Sample size
80
40
0
1
1
3
1
2
3
6
5
9
11
22
13
26
19
33
26
46
43
64
19
1997
1998
1999
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
Assessment Year

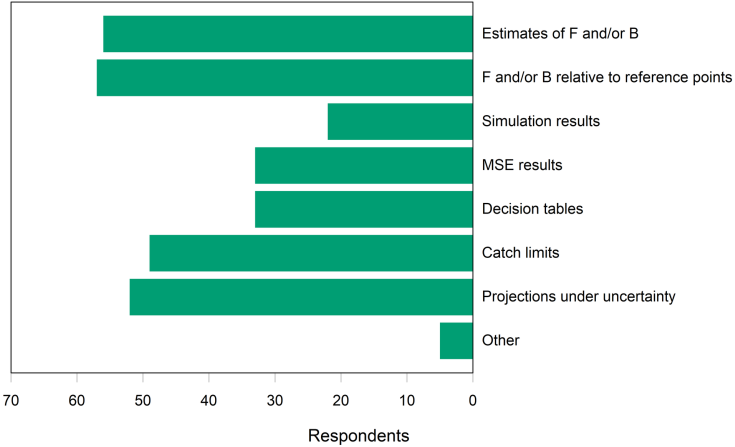

Age-structured model
VPA
Surplus production model
Length-structured model
Depletion model
DB-SRA
Uncertainty method
Asymptotic methods
Bootstrapping
Jacknife
Likelihood profile
AIS
MCMC
SIR
Decision tables
Retrospective analyses
Sensitivity analyses
Other
None
0
1
Proportion


Asymptotic methods
Bootstrapping
Jacknife
Likelihood profiles
I do not use frequentist methods
(a)
AIS
MCMC
SIR
I do not use Bayesian methods
(b)
0
5
10
15
20
25
30
35
40
45
50
Respondents


(a)
Asymptotic methods
Bootstrapping
Jacknife
Likelihood profile
MCMC
AIS
SIR
Proportion
1
1
1
1
1
1
1
0
(b)
Sample size
80
40
0
1
1
3
1
2
3
6
5
9
11
22
13
26
19
33
26
46
43
64
19
1997
1998
1999
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
Assessment Year


(a)
Catch streams
Catchability
Data set choice
Data weighting
Growth
Maturity
Natural mortality
Recruitment
Selectivity
Stock-recruit relationship
Proportion
1
0
(b)
Sample size
60
40
20
0
1 1 3 1 2 3 6 5 9 11 22 13 26 19 33 26 46 43 64 19
1997 1998 1999 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018
Assessment Year


Stock-recruit relationship
Selectivity
Recruitment
Natural mortality
Maturity
Growth
Data weighting
Data set choice
Catchability
Catch streams
Not applicable for assessment
Not major source of uncertainty
Time restrictions
Not trained
Never asked
Not familiar
40
30
20
10
0
5
10
15
20
Respondents


Estimates of F and/or B
F and/or B relative to reference points
Simulation results
MSE results
Decision tables
Catch limits
Projections under uncertainty
Other
70
60
50
40
30
20
10
0
Respondents

Tables

Software for assessments

Option	Description	Example reference
ADMB-based model	Auto-differentiation Model Builder	Fournier (2012)
ASPIC	A Stock Production Model Incorporating Covariates	Prager (1994)
AMAK	Age/Age-size Models Assessment Method for	NOAA Toolbox
ASAP	Age-structured assessment program	NOAA Toolbox
BSP	Bayesian surplus production model	McAllister and Babcock (2006)
CASAL	C++ Algorithmic Stock Assessment Laboratory	Bull et al. (2012)
CEDA	Catch Effort Data Analysis	Hoggarth et al. (2006)
Coleraine	A Generalized Age-Structured Stock Assessment	Hilborn et al. (2003)
CSA	Catch-Survey Analysis	Collie et al. (1983)
FiSAT	FAO-ICLARM Stock Assessment Tools	Gayanilo et al. (1994)
FiSAT II	FAO-ICLARM Stock Assessment Tools II	FAO (2013)
FSIM	Forecasting simulator	Goodyear (2004a)
LFDA	Length Frequency Distribution Analysis	Hoggarth et al. (2006)
ParFISH	Participatory Fisheries Stock Assessment	Walmsley et al. (2005)
PRO-2Box	Project future abundance and mortality	Porch (2017)
PRODFIT	Surplus production model	Fox (1975)
SCALE	Statistical Catch-At-Length	NOAA Toolbox
SEEPA	Simulates longline catch and effort data	Goodyear (2004b)
SS	Stock Synthesis	Methot and Wetzel (2013)
STATCAM	Statistical-Catch-At-Age Model	NOAA Toolbox
VPA	Virtual Population Analysis	Pope (1972)
VPA-2BOX	Dual Zone Virtual Population Analysis model	NOAA Toolbox
Yield	Calculates fishery yields & stock biomasses	Hoggarth et al. (2006)

Table 1. Software used in the process of conducting a stock assessment; provided to the survey respondents (i.e., the de facto options featured in survey Section 1). All methods can be used to conduct sensitivity analyses.

Modeling framework	Description	Example reference
Data limited assessment models		
CC-SRA	Catch curve stock reduction analysis	Thorson and Cope (2015)
DCAC	Depletion-corrected average catch	MacCall (2009)
DB-SRA	Depletion-Based Stock Reduction Analysis	Dick and MacCall (2011)
Other assessment models		
a4a	Statistical catch-at-age model	Jardim et al. (2014)
ADBAYECOLA	Age structured production model for trawling and longline catches	Payá (2019)
Baleen II	Age structured production model	de la Mare and Cooke (1993)
BAM	Statistical catch-at-age model	Williams et al. (2015)
BATOOTHFISH	Age structured production model with trawling and longline catches	Payá (2019)
CALEN	Catch at length model	Davies et al. (2001)
CHOSAM	Age structured production model with trawling and longline catches	Payá (2019)
CHUSmodel	Chilean Humboldt Squid Depletion Model	Payá (2019)
F-ADAPT	A custom statistical catch-at-age spatial model	Brodiak et al. (1998)
GADGET	Globally applicable Area Disaggregated General Ecosystem Toolbox	Gadget (2020)
Grenadier model	Age structured production model with swept area biomass and length composition	Payá (2009)
iSCAM delay-difference model	Integrated statistical catch age model	Martell (2012)
Modified Punt-Walker model	Spatially aggregated age- and sex- structured population dynamics model	De Oliveira et al. (2013) Punt and Walker (1998)
MULTIFAN-CL	Statistical, length-based, age-structured model	Fournier et al. (1998)
MUPPET	Age structured production model	MUPPET (2010)
SAD	A linked separable ADAPT VPA model	De Oliveira et al. (2010)
SAM	Age-structured state-space model	Nielsen and Berg (2014)
SPiCT	Surplus production in continuous-time	Pedersen and Berg (2017)
Two-stage biomass model (custom)	Stage-structured production model	Roel et al. (2009)
XSA	Extended survivor analysis	Darby and Flatman (1994)

Table 2. Software used in the process of conducting a stock assessment; provided by the survey respondents.