End-to-End Modeling as part of an Integrated Research Program in the Bering Sea

André E. Punt[1], Ivonne Ortiz[2,3], Kerim Y. Aydin[3], George L. Hunt, Jr[1], Francis K. Wiese[4]

1: School of Aquatic and Fishery Sciences, University of Washington, Seattle WA 98195-5020
2: Joint Institute for the Study of the Atmosphere and Ocean, University of Washington, Seattle, WA, 98195-5672
3: Alaska Fisheries Science Center, NOAA Fisheries, 7600 Sand Point Way NE, Seattle, WA 98115
4: Stantec Consulting Ltd., 2515 A Street, Anchorage AK 99503

Abstract

Traditionally, the advice provided to fishery managers has focused on the trade-offs between short- and long-term yields, and between future resource size and expected future catches. The harvest control rules that are used to provide management advice consequently relate catches to stock biomass levels expressed relative to reference biomass levels. There are, however, additional trade-offs. Ecosystem-based fisheries management (EBFM) aims to consider fish and fisheries in their ecological context, taking into account physical, biological, economic, and social factors. However, making EBFM operational remains challenging. It is generally recognized that end-to-end modeling should be a key part of implementing EBFM, along with harvest control rules that use information in addition to estimates of stock biomass to provide recommendations for management actions. Here we outline the process for selecting among alternative management strategies in an ecosystem context and summarize a Field-integrated End-To-End modeling program, or FETE, intended to implement this process as part of the Bering Sea Project. A key aspect of this project was that, from the start, the FETE included a management strategy evaluation component to compare management strategies. Effective use of end-to-end modeling requires that the models developed for a system are indeed integrated across climate drivers, lower trophic levels, fish population dynamics, and fisheries and their management. We summarize the steps taken by the program managers to promote integration of modeling efforts by multiple investigators and highlight the lessons learned during the project that can be used to guide future use and design of end-to-end models.

Keywords: Bering Sea; End-to-end Modeling; Ecosystem Based Fisheries Management; Management Strategy Evaluation

Corresponding author: A.E. Punt

a. *Email*: aepunt@uw.edu
b. *Phone*: 1-206-221-6319
c. *Fax*: 1-206-685-7471

1. **Introduction**

Progress on implementing ecosystem-based fisheries management (EBFM) [1] involves multiple facets, including a better understanding of the processes which characterize and control ecosystems. EBFM needs to be grounded by national and international legislation, which in the US is governed by the Magnuson–Stevens Fishery Conservation and Management Act (US Public Law 104–297). The Bering Sea Project (the combined Bering Ecosystem Study, BEST, and the Bering Sea Integrated Ecosystem Research Program, BSIERP) aimed to improve ecosystem understanding, and to support fisheries management in the eastern Bering Sea. It employed a combination of field studies and an end-to-end ecosystem model that included climate drivers, lower trophic levels and fish dynamics, which in turn could be driven by various fisheries (Wiese et al., 2012). Development and successful implementation of this project was a substantial undertaking that involved over a hundred principal investigators, with much of the historical data and fieldwork synthesized into the modeling. The Bering Sea Project has led to a better understanding of what it means to develop models for EBFM.

The primary focus of the Magnuson–Stevens Fishery Conservation and Management Act has been on single-species. However, there is an increasing recognition worldwide for the need to account for factors that are ignored when conducting single-species stock assessments. Likewise, there is growing recognition of the need to take into account the interactions among fisheries in scientific study, as well as in management decision making. This recognition has led to policy documents and statements of intent that fisheries management should move to a more ecosystem-based or ecosystem-focused approach.

In 1999, the National Research Council defined EBFM as "an approach that takes into account major ecosystem components and services, both structural and functional, in management of fisheries. It values habitat, embraces a multispecies perspective, and is committed to understanding ecosystem processes. Its goal is to achieve sustainability by appropriate fishery management". Several authors have since proposed alternative definitions for EBFM (e.g., Witherell et al., 2000; FAO, 2003; Sissenwine and Murawski, 2004; McLeod et al., 2005; Murawski and Matlock, 2006; Marasco et al., 2007; Francis et al. 2007). All of these definitions include reference to habitat and multi-species effects and more recently to climate impacts, and impacts of management on human as well as biological communities. For example, Marasco et al. (2007) provided the following definition for EBFM: "Ecosystem-based fishery management recognizes the physical, biological, economic and social interactions among the affected components of the ecosystem and attempts to manage fisheries to achieve a stipulated spectrum of societal goals, some of which are in conflict". This definition recognizes that socio-economic factors are core to an EBFM; this is supported by recent mathematical models evaluating trade-offs among management strategies that explicitly account for user responses to management regulations (e.g. Fulton, et al., 2011b). It also recognizes that management takes place within a legal management framework.

Several calls for the implementation of EBFM have been made (e.g. Pikitch et al., 2004). Section 406 of the 1996 US Sustainable Fisheries Act provided initial guidance on inclusion of ecosystem principles in management plans, and mandated the formation of the Ecosystems

---

[1] Several acronyms have been proposed for ecosystem-based fisheries management (EBFM), including EAF (ecosystem approach to fisheries). Conceptually, except for EBM (ecosystem based management) and EAM (ecosystem approach to management), which often envisage management of sectors in addition to fisheries, all these definitions have the same ultimate intent albeit their implementation may be at different management levels. We use EBFM in this paper for convenience.

Advisory Panel to the National Marine Fisheries Service, which reviews progress towards incorporation of ecosystem principles in Fishery Management Plans. However, balancing EBFM implementation with existing mandates for single-species catch limits has been challenging (see, for example, Moffitt et al., this issue).

While it has been recognized that quantitative ecosystem modeling will be a necessary component of EBFM, developing ecosystem models for fisheries management has been challenging, because: (1) field programs for EBFM are often "add-ons" to single-species surveys resulting in limited data for parameterizing ecosystem models; (2) ecosystem models, in part to ease complexity, often do not calculate quantities needed for management, such as age-structured spawning stock biomass; (3) resources often do not allow engagement of experts at all ecosystem levels during the course of a modeling project, possibly leading to misuse or misunderstanding of results; and (4) data requirements and computational complexity make it difficult to "certify" such models for management use given requirements for accuracy and the reporting of uncertainty.

The Bering Sea Project included an end-to-end model that would synthesize available data, incorporate new data from the parallel field program, and inform the ongoing research efforts. This project consequently required co-ordination of research activities by a diverse group of principal investigators to ensure that broad research goals would be achieved. Project goals included understanding biological and ecological processes, exploring various hypotheses related to the dynamics of the Bering Sea Ecosystem, and evaluating resource management options through a formal Management Strategy Evaluation (MSE).

The modeling project was designed to be tightly coupled to the fieldwork at all stages, with feedback and synthesis occurring at all levels. It required the development of standards for the ecosystem modeling efforts, and a different level of organizational guidance and regular feedback compared to 'traditional' projects. The combined organizational, modeling, and synthesis challenges were sufficiently unique from the process of "simply" constructing an end-to-end model from previously-available data that we describe the project using a new term, the Field-integrated End-To-End modeling program, or FETE.

Section 2 of this paper introduces the Bering Sea Project, and the concept and key components that constitute a FETE. Section 3 summarizes an approach (initially developed by Marasco et al. [2007]) for constructing management systems to implement EBFM based on the MSE approach and ecosystem modeling. While MSE was not the only focus of the modelling component of the project, it required the integration of all components of FETE. Section 4 outlines expectations for FETE models, guidelines established to ensure that the project was as statistically and ecologically rigorous as possible, and identifies progress against these expectations and guidelines. Section 5 summarizes best practices and future directions of integrated end-to-end modelling, i.e. what makes a successful FETE? Finally, Section 6 summarizes the legacy of the project.

2. The BEST, BSIERP and FETE

The development of BEST, and subsequently BSIERP, was initiated at an international planning workshop held in September 2002 to examine the feasibility and value of developing a large interdisciplinary study of the Bering Sea. A second planning workshop was convened in March 2003, the result of which was the development of the Bering Ecosystem Study Science Plan (2004). Contemporaneous with the development of the BEST Science Plan was the development of a long-term science plan for the North Pacific Research

Board (NPRB). Following the guidance of an *ad hoc* National Research Council panel which emphasized the importance of large-scale integrated studies of the marine ecosystems of the eastern North Pacific, similar to that being developed by BEST for the Bering Sea, NPRB developed a science and implementation plan for the BSIERP in 2005. After a limited field season funded by NSF in 2007, negotiations between NPRB and NSF resulted in a historic partnership for work in the Bering Sea, with NSF funding climate, ocean physics and lower trophic-level studies up through zooplankton, and NPRB funding work on large zooplankton through fish, seabirds, marine mammals and humans. The now combined Bering Sea Project launched its first field season in 2008 and included over one hundred principal investigators covering almost all disciplines of marine science (Wiese et al., 2012).

To aid in the development and evaluation of the modeling component in the proposals, the NPRB funded an Ecosystem Modeling Committee (EMC) in 2006, consisting of scientists not funded in the program, but experts in atmospheric and marine sciences, conceptual thinkers, as well as experienced modelers. The EMC was charged with designing modeling selection criteria to be used in proposal review and subsequent evaluations, providing advice to the funded modeling team, giving feedback to the funding agencies on the effort's progress, and helping the modelers obtain needed resources.

The resulting program, including modeling, field integration, and program review, made up the FETE. Key features included:

1. **End-to-end in scope and expertise**: Core modeling efforts and expertise were built around end-to-end research (climate, physics, plankton, fish, other animals, and humans). Critical here was the inclusion of expertise in the integration process, not merely the inclusion of "canned" results from other models and domains in the finished model.
2. **A priori and continuous integration between fieldwork and modeling**: Fieldwork and modeling were designed together from the start, with common end-goals in mind. Interactions between researchers occurred throughout the program and were structured (workshops or meetings) to allow for formal adjustments throughout the project as the field work informed the models and vice versa.
3. **Model outputs appropriate to stakeholder goals:** A priori consideration of stakeholder needs (as well as feedback from them during the program) was necessary to ensure models would produce adequate and useful results for management. For example, carbon is used in biogeochemical models concerned with climate change, but biomass may be used when examining fish foraging behavior, and numbers of fish-at-age is a key component to fisheries management.
4. **Modularity and "competition" in model design:** The structure of the FETE allowed individual components to be re-examined through "competitive" modeling; i.e. extracting the simplest component from the end to end model that captures the essence of or drivers of the interactions and using them in alternative less complex models.
5. **Centralized integration and steering:** To achieve this integration and have project goals useful to management, it was necessary to have strong project leadership, with a mandate to guide the FETE both scientifically and programmatically, including overseeing changes in scope or model design throughout the whole project.

Specific examples demonstrating how these key features were implemented in the Bering Sea Project, especially with respect to management strategy evaluation, are discussed in Sections 3-5.

2.1 FETE modeling program components

A central component of the FETE was the model[2] complex (Fig. 1) that formed the basis for exploring the impact of fishing and climate on both ecological processes and the performance of management strategies. It was used to run a 1970-2009 hindcast, and was set-up to run in forecast mode using input from selected Intergovernmental Panel on Climate Change climate models that performed well for the Eastern Bering Sea. These models are: i) the Coupled Global Climate Model, t47 grid, CGCM-t47 (low ice) from the Canadian Centre for Climate Modelling and Analysis, ii) the Hamburg Atmosphere-Ocean Coupled Circulation Model (ECHO-G; Legutke and Voss, 1999) ECHOG (high ice), from the Max Planck Institute in Germany, and iii) the Model for Interdisciplinary Research on Climate model, medium-resolution version (MIROC3.2-Medres) MIROCM (medium ice), developed by a consortium of agencies in Japan (Wang et al., 2010). The oceanography was based on the Regional Ocean Modeling System (ROMS)-Bering10K (10 km resolution), a coupled ocean-sea ice model whose spatial grid is a subset of the NEP5 model described and evaluated by Danielson et al. (2011), which itself was built on a model described by Curchitser et al. (2005) and Hermann et al. (2013). The lower trophic levels were modeled using a nutrient-phytoplankton-zooplankton detritus (NPZD) model coupled to the ROMS-Bering10K, specifically designed to incorporate the ice dynamics of the Bering Sea, and modeled nutrients, phytoplankton, copepods, euphausiids and detritus (Gibson and Spitz, 2011). Model coupling included feedback from the NPZD to the ROMS-Bering10K through phytoplankton density, which affects shortwave penetration (heat absorption) in the upper water column and between NPZD and the Forage Euphausiid Abundance in Space and Time (FEAST) model (Ortiz et a1., this issue) (functionally the fish module for this effort), through predation. A key design feature, unusual in many end-to-end models, was dynamic top-down coupling from fish to zooplankton. FEAST, thus coupled to both the NPZD and the ROMS-Bering 10K, was a multispecies bioenergetics model, with consumption as a function of length-based prey selection, prey preference and availability, and predator movement based on biomass gain optimization. Removals by fishery effort were based on spatially-explicit historical catches for the hindcast, and on a model of fishing effort allocation for model projections (FAMINE; Fishing effort Allocation Model In Nash Equilibrium).

3. Management Strategy Evaluation and EBFM

The Bering Sea Project used MSE to evaluate management strategies needed to achieve ecosystem objectives (*sensu* Sainsbury et al., 2000; Fulton et al., 2007; Dichmont et al., 2008; 2013). An MSE (Smith, 1994; Smith et al., 1999; Goodman et al., 2002; Butterworth, 2007; Punt et al., 2014b) involves assessing the performance of alternative candidate management strategies relative to performance measures that quantify the management (and legal) goals for the managed system. Thus, an MSE involves developing and parameterizing a model of the system to be managed. In the absence of data, it may also involve using hypotheses for how the system may change over time (Punt et al., 2014a).

An MSE (Fig. 2) aims to represent all key processes in system models and can provide performance metrics that relate to a broad range of goals. In the context of the Bering Sea Project, a key process was developing the scenarios regarding future climate. A concern with end-to-end models is the general inability to estimate the values for their parameters using

---

[2] It is important to distinguish the FETE modeling as a whole from any particular realization of the end-to-end model. A model in this group (e.g. "NPZD" or "FEAST") is referred to by its target trophic level, and may or may not include feedback to other components depending on the particular run. FETE as a whole refers to this suite, regardless of which components are being used for a particular result.

standard statistical models due to either lack of data or limits of computing time (Gaichas et al. [2010, 2011] being a noteworthy exception in this regard).

Which candidate management strategies are evaluated in an MSE depends in large part on the interests of the managers. Ideally, management strategies for EBFM should be based on the results of process studies, monitoring of ecosystem indicators, and ecosystem models, in addition to the outcomes of single-species stock assessments. In principle, management strategies for EBFM could involve monitoring a range of ecosystem indicators and modifying management practices based on whether the indicators are outside of acceptable limits, analogous to the types of management strategies used for single-species fisheries management. Management strategies for EBFM could be based on assessment methods that include multi-species considerations explicitly. However, to date the control rules that would underlie such management strategies have seldom been implemented or even fully defined (Moffitt et al., this issue).

To address this challenge, the FETE included a workshop with stakeholder groups to identify a preliminary set of management strategies (Fig. 3). In some cases, implementing the proposed strategies required modifications to the end-to-end model; these adjustments were made as the project progressed. The selected management strategies were based on three types of assessment methods: Ecosim, Climate-Enhanced Age-based model with Temperature-specific Trophic Linkages and Energetics (CEATTLE - the multispecies statistical model of Holsman et al. (this issue)) and the single-species assessment methods currently used to provide management advice to the North Pacific Fisheries Management Council. Each assessment method was linked to appropriate harvest control rules, which produced estimates of Total Allowable Catches. The workshop also recommended exploring a management strategy that did not implement the 2 million tonne cap on total harvest, which is currently written into regulation for the eastern Bering Sea (Fig. 3). The workshop also specified management scenarios based on the impact of climate change.

4. Guidelines and principles for the development of ecosystem models, and how to apply them towards end-to-end modeling

The questions the EMC developed to evaluate the proposals for the modeling component of the Bering Sea Project focused on what the various models were meant to produce and why, whether the outputs would be useful for management and would provide measures of uncertainty, how existing and future data could be integrated into the model, how the model could inform ongoing research, and whether the model could be validated. The questions and their rationale are discussed below and, even though they were developed for the Bering Sea Program, they provide a way to evaluate any model.

4.1 What is the model intended to predict?

This may seem like an extremely simple question. However, many models, particularly those of the end-to-end variety, claim to be able to predict many types of impacts. The aim of this question was to ensure that the models were designed given specific scientific and management questions, rather than having the models developed and subsequently retrofitted to address questions of scientific and management relevance.

The FEAST and NPZD models (effectively the biological component of the integrated model) were designed as predictive models responsive to long term climate variation and geared to address two basic purposes: (1) understand the underlying processes by which environmental variability affects biological processes such as primary and secondary production and fish recruitment and distribution, and (2) characterize the environmental effects on the distribution of fishing effort and hence the age structure in fish populations and recruitment to the fishery. This involved using FEAST as the system model for an MSE aimed at walleye pollock *Gadus chalcogrammus*, Pacific cod *Gadus macrocephalus*, and arrowtooth flounder *Atheresthes stomias*.

The ROMS model was designed to enable climate factors to be explicitly represented in the dynamics of the resources, while the FAMINE and MSE models were developed to represent management and how management actions translate into fishing effort and hence fishing mortality.

4.2 What specific aspect of the prediction is anticipated to be of direct value for fisheries management?

Many proposals for scientific research claim that their research will be of direct use for management purposes. The EMC envisaged that by explicitly stating how predictions would be used for management purposes, the modeling proposal and the subsequent research would be more likely to lead to predictions that would actually achieve this purpose.

Amongst the main goals was the ability to predict the responses of fish stocks and fishermen to long-term climate scenarios. The high resolution of ROMS (~10km) would provide maps that would allow detailed representation of fleet distributions. The full end-to-end model was geared to address expected changes in potential total allowable catches and fish availability to the catcher processors and catcher vessels, which have distinct spatial constraints. Each individual model had outputs that were linked, such that changes in climate would feed through the simulated ecosystem to impact how management strategies would be able to achieve the goals established for EBFM.

4.3 What measure of "accuracy" in the prediction is crucial to determining the usability of that prediction to fisheries management?

In principle, models can make predictions of virtually any quantity. However, the estimates may be very biased and/or imprecise. The EMC expected that the desired quality (or accuracy) of predictions would be evaluated before the modeling was to be conducted. This was perhaps one of the most challenging of the questions because establishing hard standards for model accuracy is difficult. Validations are time consuming to perform and can be computationally expensive. Some types of error are cumulative, and only emerge after multiple years into the simulation. In general, validations and performance assessments do not have a set level of accuracy. Rather, they have levels of conformance as measured by correlation, principal component analysis and comparisons between the observed data and model output.

Even when each modeling component within the overall model (ROMS, NPZD, FEAST, FAMINE, MSE) provided plans that included statistical techniques to measure variance and accuracy, the number and diversity of variables in each model made it impossible to provide the desired level of accuracy for each output from the integrated model. For example, even if it is possible to explain 50% or more of the variance of the data used in a particular model, the cascading effect of such variability or lack of accuracy on processes outside that model may be greater. For example, initial sea temperature estimates in the ROMS model, considered to be within acceptable ranges in an oceanographic context, drove the bioenergetics of lower and upper trophic levels towards and beyond their upper tolerance limits. Moreover, it moved the location and extent of the cold pool – a key environmental factor known to impact the dynamics of groundfish stocks (NPFMC, 2012) – thus changing critical temporal and spatial ecosystem dynamics.

4.4 What alternative models are plausible competitors whose performance should be tested against the model being developed?

All models should be recognized as simplifications of the system under consideration. The EMC recognized the need for multiple alternative models so that the predictive skill of the proposed model could be evaluated relative to alternative (generally less complex) models, and because it is not uncommon for the predictions from ecosystem models to be very sensitive to their structure.

The EMC envisioned complementing and competing models: in particular, correlative models to be developed as part of the Bering Sea Project (Mueter et al., 2011; Siddon et al., 2011, 2013a, b; Heintz et al., 2013), and existing models such as MSM (Jurado-Molina et al., 2005) and the Ecopath model for the eastern Bering Sea (Aydin et al., 2007), as well as currently used single-species stock assessments. Also developed were a multi-species biomass dynamics model for walleye pollock, Pacific cod, arrowtooth flounder (the three main species in FEAST), and small mouthed flatfish (not in FEAST) (Uchimaya et al., this issue), and a statistical model linking recruitment of walleye pollock to variability in late summer sea surface temperatures and to the biomass of major predators (Mueter et al. 2011).

4.5 How will the achieved predictive power of the model be compared against the performance of plausible alternatives, and how will this guide subsequent choices about model form and parameterization?

The quality of fishery models is generally assessed in terms of hindcast skill, i.e. the ability to replicate the data used for model calibration, and this is clearly a minimum requirement for any ecosystem (or other) model. Considerable effort has been dedicated to developing metrics for evaluating hindcast skill for stock assessment models, including residual analysis and Bayesian methods for posterior predictive checks. However, the EMC expected model performance (and model refinement) to be based on forecast as well as hindcast skill.

Given the expected performance of FEAST's forecast skill, several attributes, including those linked to the stock assessment models, required calibration. The predictions, which could be compared among models, included spatial aspects such as species distribution by age, as well as key regional and length-specific trophic interactions (e.g., Buckley et al., this issue).

The ability to review the performance of forecasts based on the FAMINE and MSE components of the integrated model was limited given lack of sufficient computational resources. However, forecast skill could have been evaluated by running the calibrated end-to-end model to a year other than the most recent year and projecting forward. Unfortunately, time constraints of the overall project, given the available computational resources, precluded this.

4.6 What data are available to drive, calibrate, and test the model?

This question recognized that data are used in multiple ways in ecosystem models. The EMC envisaged that some sources of data would be included in the model as "facts". However, data in this context also include values for parameters that are pre-specified based on auxiliary information. For example, when applying models such as Ecosim, diet is frequently assumed to be known. All models, ecosystem or otherwise, include parameters that are not known from auxiliary information but which must be estimated from the monitoring data. The model fitting process should ideally involve minimizing some form of objective function involving discrepancies between the observed data and model predictions. However, it is computationally infeasible to fit large complex ecosystem models such as FEAST or Atlantis (Fulton et al., 2011a) to monitoring data, so the model calibration process is more heuristic than formal. The EMC considered model validation a key step in the modeling process and expected that some of the available data would be kept away from the modelers to allow an independent test of model skill. Use of this form of cross-validation is common in some modeling fields, but is relatively uncommon with fisheries modeling where, given the general lack of data, all of the available information is used for model calibration.

The primary sources of data for FEAST were the historical databases kept by the Alaska Fisheries Science Center (NOAA) for fish age, length, weight, distribution, feeding habits and fishery catches. Data for the models of the lower trophic levels and the ROMS model were based on past data, as well as from moorings and process studies that were part of the Bering Sea Project. The FAMINE model was driven using data on fishing effort and ice cover, whereas the MSE model used information generated by FEAST. However, no current amount of field work could provide the data needed to estimate all parameters and validate all levels of the end-to-end model. In hindsight, the availability and consolidation of such data proved to be a bottleneck for model development, particularly for the NPZD model and the process studies.

4.7 How will the existing data be used to quantify model fit and predictive power?

Evaluating model fit (hindcast skill) is a key element of single-species stock assessment, and extensive terms of reference have been developed to detect violations of the ability to replicate data (e.g., PFMC, 2012). How to evaluate hindcast skill, however, is not as developed for multi-species models (see, however, Gaichas et al., 2010, 2011), and particularly not for models that produce spatial outputs, owing to spatial autocorrelation in the data available for evaluating model skill. Simple metrics (e.g., all species remain in the system) have been used to evaluate model fit and hindcast skill for ecosystem models, but these metrics are not nearly as sophisticated as those used for single-species stock assessments.

Evaluating predictive power involves similar issues to evaluating hindcast skill, but with additional complexity: assumptions made when making future predictions need to be specified and evaluated carefully. A variety of approaches were used to validate the components of the end-to-end model. For example, the climate models used for the forecast were selected based on performance in the Bering Sea, mainly their ability to capture ice cover and the Pacific Decadal Oscillation (Wang et al., 2010).

Validation of physical characteristics (correlations between observed and model estimates) such as ice cover and temperature was carried out by Danielson et al. (2011) for the 60-layer ROMS North East Pacific 5 model. The smaller grid used for the Bering 10K ROMS-NPZD and Bering 10K ROMS-NPZD-FEAST-FAMINE model has a reduced vertical resolution from 60 to 10 levels. Hermann et al. (2013) conducted both correlation and principal component analyses using available time series for physical data, such as temperatures at mooring 2 (M2), ice extent and salinity; multivariate analysis was performed using data from the Bering Sea. Herman et al. (2013) also used temperature, salinity and total chlorophyll from the Alaska Fisheries Science Center’s annual Bering-Aleutian Salmon International Survey (BASIS) research cruises in a multivariate analysis. Gibson and Spitz

(2011) conducted a sensitivity analysis of the NPZD portion of the end-to-end integrated model. Assessments of fish movement and distribution patterns (I. Ortiz, UW, unpublished results), biophysical processes (Ortiz et al., this issue) and fish bioenergetics (K. Aydin, NOAA, unpublished results) were also conducted.

For FEAST, historical data from 1982 to 2007 were used to estimate parameters related to the fish bioenergetics (length-weight relationships and length-energy density) and the relationship found between recruitment and fall condition of age-0 pollock was used to assess model performance. Refinements of these processes were made based on the field studies. For spatial aspects, historical data were used to construct initial conditions for fish in all years from 1971 to 2010. This allowed testing of single individual years. However, since only the first year uses initial conditions derived from data, for multiyear runs, subsequent years could be validated using the remaining historical data.

Ideally, a more holistic validation of the entire end-to-end model could have been achieved had there been both cold and warm years during the field seasons encompassed by the Bering Sea Project. Contrast in environmental conditions during the fieldwork years was originally envisaged in the proposals that led to the Bering Sea Project. However, all field years were cold, thus precluding this approach to model validation.

In general, FEAST succeeded in capturing the general growth, movement and distribution of fish, and was sensitive to cold and warm years. However, the model failed to predict recruitment and survival of age-zero fish satisfactorily for multi-year historical runs in which small age-structure errors could accumulate over the run, and the numbers of age-1 pollock had to be nudged to their stock assessment estimated numbers at the end of each model-year.

4.8 What pertinent future data are anticipated to become available within the time frame of the project and how will these future data be used to quantify model fit and predictive power?

The FETE involved model development, data collection occurring in parallel, and this question was developed to ensure that fieldwork and modelling were integrated. Obtaining data for the lower trophic levels for cold and warm years was not feasible due to the lack of warm years during the field program (Stabeno et al., 2012). Several data sets that became available during the program were integrated into the modeling efforts (either for parameter estimation or to assess model performance), namely improved spatial distribution of age-0 and age-1 pollock, zooplankton surveys, acoustic estimates of euphausiids, winter distribution of the pollock spawning stock, seasonal energy density of juvenile pollock, consumption of small, medium and large copepods by fish, and a series of data from the lower-trophic-level component. Several of these data sets, e.g. pollock bioenergetics, acoustic estimates of euphausiid biomass, and additional oceanographic data, are now regularly updated and have become part of standard surveys due to their usefulness for supporting analyses. Other data gaps have led to new analyses (such as zooplankton seasonal and spatial patterns) and pilot projects (winter zooplankton sampling).

4.9 How has it been determined that the proposed quantity and quality of data can be expected to be sufficient for the intended use in tuning and testing the model?

This question attempted to integrate the remainder of the questions, and hence provide an overall basis for evaluating the design of the modeling. Unfortunately, this question won’t be fully addressed until the end-to-end model has been applied more extensively.

5 Discussion: Best practices and future directions

The approach for developing end-to-end models for management purposes outlined by Marasco et al. (2007) is comprehensive, and, when combined with the questions developed by the EMC, should have led to a process in the FETE where a set of models was selected that were relevant to the system at hand, could be calibrated to existing data and tested through comparison with independent data sources, and were useful for evaluating management strategies in an ecosystem context. Practice, however, often differs from theory, and hence here we summarize our experience and distill what we consider best practices to facilitate subsequent efforts and end-to-end modeling in general.

5.1 Be realistic about what can be accomplished within a given timeline

It is important to be realistic about the constraints due to the size and complexity of a model before work starts on its development and parameterization. In the case of the Bering Sea Project, the complexity of the FETE effort only became fully apparent as the project proceeded. For example, coupling the individual models was a major undertaking, which, although recognized as a key task when the overall project was designed, and a goal that was achieved, was an ongoing constraint on the speed of model development. As such, a **significant amount of effort should be spent early on fully scoping out the model needs, especially in terms of integration**. Most modelers are generally well aware of their individual needs and are somewhat realistic about what can be done. Developing end-to-end models for actual ecosystems and management, however, is a much younger endeavor, resulting in a tendency to underestimate challenges and project outcomes on the basis of potential rather than reality.

5.2 Larger-scale software projects need logistical support on a par with fieldwork

Care should be taken when a project's scale exceeds that of an individual or a small team and encompasses multiple institutions. While technology scales, large-scale software development, as an activity, does not (Brooks, 1995). Scientists used to working as individuals, on individual pieces of code, need to expect time devoted to logistics of working with large computers at multiple institutions, transferring files, and keeping source code synced. When coupling models from different disciplines and modeling teams, code is often written independently and then synchronized. **Software and hardware management and familiarity with the structure and parameters of all components of the model are critical for achieving a working end-to-end model**.

5.3 Clear separation of scientific versus logistics oversight

Rose et al. (2010) note that the challenge of interdisciplinary research is "as much of a people challenge as a technical one". In the case of the integrated modeling work, the first few years were coordinated through the EMC. Their role was to guide and facilitate, but not to make final decisions. The questions designed by the EMC included both scientific concerns (comparing outputs to data) and logistical concerns (time frame of data). However, the EMC functioned almost entirely as a scientific review body during the initial stages of the actual work on the project. Logistics were initially to be handled by the modelers collectively; while a lead modeler was appointed, it was primarily in a communication/coordination role rather than as a firm project leader.

As the project developed and many modelers focused on their own timelines and model developments, it became clear **that a modeling facilitator was needed to help maintain a unified standard and expectation across projects in terms of cross-collaboration, facilitation, product delivery, priorities and overall model management.** Such an independent, but informed, coordinator was appointed during the latter part of the project and helped to keep the overall outcome in mind whenever individual goals and timelines were in conflict. A third model of how an independent group can facilitate and oversee a modeling project is provided by the Gulf of Alaska Integrated Ecosystem Research Program (GOAIERP). This is a much smaller project than the Bering Sea Project with a markedly smaller modeling component. In particular, there is no attempt to develop an end-to-end model for the Gulf of Alaska at present, so the logistics involved in the modelling are markedly less. In this case, an individual was contracted by NPRB on an as–needed basis to provide guidance to the modeling group.

5.4 Open and frequent communication with field biologists

In addition to being the source of most of the data for validation, field biologists provide expert advice and direction when confronted with modeling decisions for which there are apparently equally suitable options or no data. Close communication with groups of field biologists also facilitates consensus building, improved understanding of model structure and ultimately, and acceptance of the model. In the FETE, much effort was put towards facilitating frequent conversations between modelers and field teams, and the latter consequently had a clear expectation that ongoing data collection would ‘feed into’ the modeling. This might have been a realistic expectation if it were a simple issue of adding data to a data file and running the model. However, adding data can lead to changes in the model structure because the model structure is, by definition, tailored to the data. There is also a lag time between data collection, analyses and pattern/process identification. While it is obviously desirable to allow data collection efforts to feed into model development and parameterization, the process should not be considered routine, fast, easy or not disruptive to the overall modeling process**. Addressing the issue of how to integrate new data into the modeling process needs to be addressed early in the project design, and the logistic constraints need to be recognized**. For example, new data could be used for validation purposes in the final year of a project if sufficient data are collected to parameterize the model in the first place. This issue was identified at the start of the project, but the extent of the task was not totally understood at the time. The possibility of the results of a major piece of fieldwork calling for a major change to model structure was not recognized at the time the project was designed, but rather later during development.

5.5 Adequacy and availability of data for model validation/testing

Ideally, the existing data and the temporal and spatial coverage of the key variables in the models should match. In the FETE, many of the oceanographic and lower trophic level data available to validate the model came from point data, e.g., moorings, which provide reliable time-series but poor geographic coverage, or from oceanographic stations, spread over a large area but with no associated long-term time-series. Eventually, an effort was made to use other sources of data (such as, for example, temperatures collected during annual fishery surveys) appropriate for model validation. In addition, a series of data sources were combined to define regions of similar bio-physical characteristics that could be used for model comparison rather than relying on point sources (Ortiz et al., 2012). The existing data should also be compiled and made available in advance. For both the oceanographic and the lower trophic level modeling efforts, data and validation came late in the process, too late for the benefits of improved parameters to be included in the simulations coupling fish dynamics. **Future attempts at end-to-end modeling should involve a group to identify all potential data sources, a designated entity in charge of compiling, formatting, and disseminating such datasets, and the creation of the framework by which to conduct model validation.**

5.6 Most work is sequential and iterative as opposed to simultaneous and independent (non iterative)

All models have to be integrated and re-validated as a whole. The size of this task is highly dependent on overall model structure and level of coupling/linkage between the different model components. This is not a one-time occurrence and demands longer timelines, as response time depends on each party's time availability and priorities, in addition to the actual difficulty of the problem itself. Therefore, **even when one of the components of an end-to-end model is considered finalized, time should be allotted to support further implementation and testing of subsequent coupled versions of the integrated model**.

In the FETE, this issue proved particularly challenging for the use of MSE, as forward projections could not commence until the remainder of the Bering 10K ROMS-NPZD-FEAST-FAMINE model had been developed and validated. Having an MSE component from the start of the program meant that management quantities to be extracted from the model (e.g. spawning stock biomass for fish stocks) were built into the model design from the start, rather than in an *ad hoc* manner afterwards. However, the first viable (hindcast) version of the fully-coupled model was finalized only after six years, so the "top-of-the-food chain" portions of the project (MSEs and Economics) ended up being much more limited in scope than intended. We propose two alternatives to address this problem:

(1) Start projects of this type in multiple phases. In particular, phase 1 would involve developing the ecosystem component model that will operate together as a system model while phase 2 would involve refining the system model and also conducting the MSE. Phase 1 would involve steps such as a stakeholder workshop to identify the management strategies to evaluate and also the specification of the data that are needed to apply to selected management strategies. These steps are needed so that the biological component of the system model is structured to generate the data needed as the basis for the MSE.

(2) Conduct the MSE as part of the FETE, but also develop a "simple" system model as a component of the project so that some MSE results can be obtained. It is likely that some management strategies will fail to achieve the management objectives using a simple model. It would be expected that management strategies that 'fail' for simple system models will also 'fail' for more sophisticated and realistic system models.

It should be noted that there is a cost associated with developing ecosystem models to evaluate management options beyond that required to increase ecosystem understanding. For example, the management strategies to be evaluated required data on the age structure of fishery and survey catches. The original design of the FEAST model involved modeling population length- but not age-structure; including population age-structure in FEAST increased the number of variables for pollock, Pacific cod and arrowtooth flounder from approximately 180 to 1386 and reduced the number of length bins from 20 to 14. The management strategy evaluations also required fisheries by sector (catcher vs catcher/processor vessels) in addition to by gear and species, thus doubling the number of modelled fisheries. Moreover, the need to manage according to total catch quotas also required the model to be stopped at regular intervals during the simulation to keep track of total catches and effort allocation, which added additional complexity to the overall project.

5.7 Mismatch of required performance levels and performance measures between single discipline approaches and multidisciplinary ones

When development of the fish model in BSIERP started, there was an incomplete understanding of the state of development of the oceanographic model. Later, it was noted that the oceanographic model predictions of temperature were biased by approximately 2°C. This bias was considered acceptable within an oceanographic context, but unacceptable for the bioenergetics in the fish model, and for the consequences of temperature on fish distribution. **Particular emphasis should be placed on differences in required scales of results between models**. For example, a 1-dimensional version of the coupled ROMS-NPZD was developed early on in the modeling for calibration to a specific data source (the M2 mooring). It was initially thought and planned that the 1-D model would be sufficient to quickly test and calibrate the fish model while it was under development. However, the combination of M2 being a poor location for fish due to productivity, and the importance of horizontal movement for calibrating fish growth, meant that the testbed had to await a 3D model, thus slowing down achievement of planned milestones.

Models are always a mix of mechanistic and statistical aspects. FEAST is a primarily mechanistic model with as few embedded phenomenological correlations amongst variables as possible. This pertains to (but not exclusively) the EMC's questions regarding data availability and usage. Some data were used to set up the mechanics, some data were used to test model performance (e.g. the spatial distribution of fish species by age and length), and some were used as a given process part of the system. It is important to distinguish between using data as "facts", and the steps or mechanics of growth and data used to evaluate performance of a synergistic property. How much a model is “steered” towards the mechanistic vs. the phenomenological gradient is a constant choice, and while some guidelines and principles are general and applicable to all ecosystem modeling, some are specific as they depend on the nature of the project. **Decision making should be consistent with both the mechanistic and the phenomenological gradient throughout the entire project. Individual component performance metrics should be in line with the overall purposes of the model and not with a discipline-specific need or standard. Alternatively, if there are multiple purposes, there needs to a clear process for prioritizing those purposes.**

The mismatch in levels of performance between single discipline and multidisciplinary work often requires a recalibration of the various components once coupled so general patterns can be captured. Further model refinement improves timing, magnitude and other attributes and decreases the need to compensate the mismatch between models.

5.8 Lack of familiarity with model limitations pertaining to other disciplines

There is a learning curve when working with multidisciplinary models that can only be gained by experience and joint collaboration. While all the modelers involved had experience developing models within their field of expertise, most were unaware or unfamiliar with computing languages, common practices, model structure, model restrictions and expectations from the other disciplines. This resulted in serious implications for model design. For example, the fish modelers assumed that time savings could occur through coarser time steps (which couldn’t be done due to physical constraints), while the physicists assumed that the fish could be modeled with fewer state variables covering length and ages of fish (which couldn’t be done due to biological and MSE constraints). A consequence of this was much longer run times and hence increased difficulties with model development and calibration. In addition, the funded proposal was modified through discussions with the funding bodies, other researchers on the project and the EMC. Consequently, the workplan for the modelling was modified during the project development process instead of during the proposal development phase. **Clear, transparent communications between all components needs to occur during proposal development and early phases of the program to avoid misunderstandings and to dispel wrong assumptions. Moreover, the relationship between realism and run times needs to be recognized during the project design stage.**

5.9 Coherence of final products from different funding agencies

Different components of the project were completed at different times, and the early finishers were thus initially disengaged from the synthesis. Eventually, the issue was addressed by several synthesis projects being funded. The mis-match in the funding of synthesis efforts reinforces the **importance of including adequate time for synthesis as well as for time for modelers to deal with requests from, and interaction with, other modelers and field biologists from all components involved in the integrated program**. A program needs to start with a synthesis of the kinds of data that will be needed to address the central questions driving the program, as well as a synthesis at the end. This wrap-up synthesis requires that many if not most of the basic papers from the program are in press so that they are available to the synthesis teams. Pushing the final synthesis too early means that much of the material derived from the field and modeling program will not be available for the synthesis.

6 Conclusions: program legacy

Looking at each individual project separately, the Bering Sea Program's modeling effort, or FETE, was extremely successful by most scientific funding standards. The oceanographic model, the NPZD model, and the fish growth/movement model, can be seen as separate 3-year modelling projects; compared to a traditional sequential approach (completing work bottom-up from physics to fish), the overall program condensed 9 years of research into 6 years. Advances were made in physical modelling of the region (Danielson et al., 2011, 2012; Hermann et al., 2013), measuring uncertainties in NPZD models (Gibson and Spitz, 2011), and quantifying seasonal versus interannual environmental effects on the growth, feeding rates, and survival of fish (K. Aydin, NOAA, unpublished results), effects of prey availability and temperature on fish distribution (I. Ortiz, UW, unpublished results), and year-round biophysical processes and their effect on fish and fisheries (Ortiz et al., this issue).

The structure of the overall Bering Sea Project, including in-depth principal investigator meetings and structured workshops between modelers and observationalists, facilitated strong connections for specific components. This is reflected by the large number of observationalist and modeler partnerships that developed during the project. Modelers have brought key results from ROMS, NPZD, and/or FEAST (such as predicted euphausiid densities) to the ongoing NSF synthesis project, fueling modelling and data analysis well beyond the scope of the original program (e.g., Sigler et al., this issue).

The project has also had ramifications in the ongoing monitoring of the Bering Sea. The Alaska Fisheries Science Center is continuing the development of the FETE and is currently using it to target specific model parameter uncertainties for extended research during ongoing monitoring activities. This new, integrated activity should significantly operationalize the FETE, both as model and field components, to provide EBFM advice on an ongoing basis. Combined, these factors have the potential of creating an institutional structure that will link modeling and field work more tightly into the future. Additionally, the program has brought fisheries modeling into the developing field of high-performance computing and high-performance data applications.

The MSE project included an initial workshop with attendance from a broad range of stakeholders and decision makers, and included the development of potential management scenarios. The end results are visible in the North Pacific Fishery Management Council's current research priorities, which include the development of management strategy evaluations and continued production of whole-ecosystem models for integrated ecosystem assessment.

Every model, just like every field measurement, is in some sense "wrong"; a model, however complex, is a simplification of reality. The researcher's challenge is to consider modeling like field research, as an ongoing, iterative process, producing new questions as well as answers. The models, as proposed, included a brief to change the very way that field research interacted with models. In that, they were highly successful; the legacy that this project left is visible today in the ongoing collaborations between researchers of the Bering Sea, stakeholders, agencies, management bodies, and the public.

Ultimately, the question that needs to be answered is whether it will ever be feasible to construct a FETE that follows all of the steps outlined by Marasco et al. (2007), and fully addresses the questions developed by the EMC. We believe that the Bering 10K ROMS-NPZD-FEAST-FAMINE model has already increased understanding about the Bering Sea ecosystem and its fisheries, even if it could not follow all of the steps nor fully address all of the questions. Nevertheless, the guidance provided through the work of the EMC, along with the experience gained through this project, suggests that a FETE will enhance the development and use of end-to-end models to increase understanding of ecosystems and provide useful information for both management and research prioritization.

The lessons learned during the development of the FETE are applicable to future model development work in the North Pacific but also in regions where similar endeavors are being undertaken such as the Benguela (Travers-Trolet et al., 2014) and the California (e.g. Fulton et al., 2011a; Kaplan et al., 2012) current systems. These lessons are particularly relevant when considering the development of permanent operational programs for EBFM, such as the Integrated Ecosystem Assessment program of NOAA (Levin et al., 2009), where it is envisioned that ecosystem models, if coupled with ongoing feedback from field researchers and management, may form an organizing principle for a core EBFM team to provide ecosystem-based management and research advice in an ongoing fashion.

7 Acknowledgements

This paper is derived from the authors' presentations at the Daniel Goodman Memorial Symposium (20-21 March 2014), and we thank the organizers for inviting us to be part of that important occasion. The authors would like to acknowledge the central contribution of the EMC to the Bering Sea Project, and in particular the contributions of Dan Goodman, former Science Panel member of the NPRB, whose idea it was to form an EMC, and who chaired that effort during its existence. Without the framework Dan and the committee provided, this modeling effort would have been fragmented and less tuned to delivering products useful not only for scientific understanding, but also for direct application to management objectives and decision making. AEP and IO were partially supported by the North Pacific Research Board. GLH was partially supported by National Science Foundation grant number 1107250. FKW was fully supported by NPRB in his former role as Science Director and modeling manager for NPRB. Mike Sigler, Martin Dorn (AFSC), Chris Harvey (NWFSC), three anonymous reviewers and the guest editor, Tom Van Pelt, are thanked for their comments on earlier versions of this paper. This work was partially funded by the Joint Institute for the Study of the Atmosphere and Ocean (JISAO) under NOAA Cooperative agreement No. NA10OAR4320148, Contribution No. 2424. This paper is BEST-BSIERP Bering Sea Project publication number 157 and NPRB publication number 535.

8 References

Aydin, K., Gaichas, S. Ortiz, I, Kinzey, D. Friday, N., 2007. A Comparison of the Bering Sea, Gulf of Alaska, and Aleutian Islands Large Marine Ecosystems Through Food Web Modeling. NOAA Tech Memo NMFS-AFSC 178. 298 pp.

Bering Ecosystem Study (BEST) Science Plan. 2004. Fairbanks, AK: Arctic Research Consortium of the U.S. ix+82 pp.

Brooks, F., 1995. The Mythical Man-Month. Addison-Wesley, New York.

Buckley, T.W., Ortiz, I., Kotwicki, S., Aydin, K., This issue. Summer diet composition of walleye pollock in the eastern Bering Sea, 1987-2011, and predator-prey relationships with copepods and euphausiids. Deep-Sea Res. II 00, 00–00.

Butterworth, D.S., 2007. Why a management procedure approach? Some positives and negatives. ICES J. Mar. Sci. 64, 613–617.

Curchitser, E.N., Haidvogel, D.B., Hermann, A.J., Dobbins, E.L., Powell, T.M., Kaplan, A., 2005. Multi-scale modeling of the North Pacific Ocean: Assessment and analysis of simulated basin-scale variability (1996–2003). J. Geophys. Res. 110, C11021, doi:10.1029/2005JC002902.

Danielson, S., Curchitser, E., Hedstrom, K., Weingartner, T., Stabeno, P.J., 2011. On ocean and sea ice modes of variability in the Bering Sea. J. Geophys. Res. 116, C12034, http://dx.doi.org/10.1029/2011JC007389.

Danielson S., Hedstrom, K., Aagaard, K., Weingartner, T., Curchitser, E., 2012. Wind-induced reorganization of the Bering shelf circulation. Geophys. Res. Lett., 39, L08601, http://dx.doi.org/10.1029/2012GL051231.

Dichmont, C.M., Deng, A., Punt, A.E., Ellis, N., Venables, W.N., Kompas, T., Ye, Y., Zhou, S., Bishop, J., 2008. Beyond biological performance measures in Management Strategy Evaluation: Bringing in economics and the effects of trawling on the benthos. Fish. Res. 94, 238–250.

Dichmont, C.M., Ellis, N., Bustamante, R.H., Deng, R., Rickell, S., Pascual, R., Lozano-Montes, H., Griffiths, S., 2013. Evaluating marine spatial closures with conflicting fisheries and conservation objectives. J. Appl. Ecol. 50, 1060–1070.

FAO Fisheries Department. 2003. The ecosystem approach to fisheries. FAO Technical Guidelines for Responsible Fisheries. No. 4, Suppl. 2. Rome, FAO.

Francis, R.C., Hixon, M.A., Clarke, M.E., Murawski, S.A., Ralston, S., 2007. Ten commandments for ecosystem-based fisheries sciences. Fisheries 35, 217–233.

Fulton, E.A., Link, J.S., Kaplan, I.C., Savina-Rolland, M., Johnson, P., Armsworth, C., Home, P., Gorton, R., Gamble, R.J., Smith, A.D.M., Smith, D.C., 2011a. Lessons in modeling and management of marine ecosystems: the Atlantis experience. Fish and Fish. 12, 171–188.

Fulton E.A., Smith, A.D.M., Smith, D.C., 2007. Alternative management strategies for Southeastern Australian Commonwealth Fisheries: Stage 2: Quantitative Management Strategy Evaluation. Report to the Australian Fisheries Management Authority and the Fisheries Research and Development Corporation. CSIRO Marine and Atmospheric Research.

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

Gaichas, S.K., Aydin, K.Y., Francis, R.C., 2010. Using food web model results to inform stock assessment estimates of mortality and production for ecosystem-based fisheries management. Can. J. Fish. Aquatic Sci. 67, 1490–1506.

Gaichas, S.K., Aydin, K.Y., Francis, R.C., 2011. What drives dynamics in the Gulf of Alaska? Integrating hypotheses of species, fishing, and climate relationships using ecosystem modelling. Can. J. Fish. Aquat. Sci. 68, 1553–1578.

Gibson, G.A., Spitz., Y.H., 2011. Impacts of biological parameterization, initial conditions, and environmental forcing on parameter sensitivity and uncertainty in a marine ecosystem model for the Bering Sea. J. Mar. Sys. 88, 214–231

Goodman, D., Mangel, M., Parkes, G., Quinn, T., Restrepo, V., Smith, T., Stokes, K., 2002. Scientific review of the harvest strategy currently used in the BSAI and GOA groundfish fishery management plans. North Pacific Fishery Management Council. Anchorage, AK.

Heintz, R.A., Siddon, E.C., Farley Jr., E.V., Napp, J.M., 2013. Correlation between recruitment and fall condition of age-0 pollock (*Theragra chalcogramma*) from the eastern Bering Sea under varying climate conditions. Deep-Sea Res. II 94, 150–156.

Hermann, A.J., Gibson, G.A., Bond, N.A., Curchitser, E.N., Hedstrom, K., Cheng, W., Wang, M., Stabeno, P.J., Eisner, L., Ciecel, K.D., 2013. A multivariate analysis of observed and modeled biophysical variability on the Bering Sea shelf: Multidecadal hindcasts (1970-2009) and forecasts (2010-2040). Deep Sea Res. II 94, 121–139.

Holsman, K.K., Ianelli, J., Aydin, K., Punt, A.E., Moffitt, E.A., This issue. A comparison of fisheries biological reference points estimated from temperature-specific multi-species and single-species climate-enhanced stock assessment models. Deep Sea Res. II 00, 00–00.

Jurado-Molina, J., Livingston, P., Ianelli, J., 2005. Incorporating predation interactions in a statistical catch-at-age model for a predator-prey system in the eastern Bering Sea. Can. J. Fish. Aquat. Sci.62, 1865–1873.

Kaplan, I.C., Horne, P.J., Levin, P.S., 2012. Screening California Current fishery management scenarios using the Atlantis end-to-end ecosystem model. Prog. Ocean. 102, 5-18.

Legutke, S., Voss, R., 1999. The Hamburg atmosphere-ocean coupled model ECHO-G. Technical Report~18, German Climate Computer Center (DKRZ).

Levin, P.S., Fogarty, M., Murawski, S.A., Fluharty, D., 2009. Integrated Ecosystem Assessments: Developing the scientific basis for ecosystem-based management of the ocean. PLoS Biology 7(1): e1000014. doi:10.1371/journal.pbio.1000014

Marasco, R.J., Goodman, D., Grimes, C.B., Lawson, P.W., Punt, A.E., Quinn II, T.J., 2007. Ecosystem-based fisheries management: some practical suggestions. Can. J. Fish. Aquat. Sci. 64, 928–939.

McLeod, K.L., Lubchenco, J., Palumbi, S.R., Rosenburg, A.A., 2005. Scientific consensus statement on marine ecosystem-based management. Signed by 221 academic scientists and policy experts with relevant expertise. Communication Partnership for Science and the Sea (COMPASS) [online]. Available from http://compassonline.org/?q=EBM.

Moffitt, E., Punt, A.E., Holsman, K., Aydin, K.Y., Ianelli, J.N., Ortiz, I., This issue. Moving towards Ecosystem Based Fisheries Management: Options for parameterizing multi-specie sharvest control rules. Deep Sea Res. II 00, 00–00.

Mueter, F.J., Bond, N.A., Ianelli, J.N., Hollowed, A.B., 2011. Expected declines in recruitment of walleye pollock (*Theragra chalcogramma*) in the eastern Bering Sea under future climate change. ICES J. Mar. Sci. 68, 1284–1296.

Murawski, S.A., Matlock, G.C., editors. 2006. Ecosystem Science Capabilities Required to Support NOAA's Mission in the Year 2020. U.S. Dep. Commerce, NOAA Tech. Memo. NMFS-F/SPO-74.

North Pacific Fishery Management Council (NPFMC). 2012. Stock Assessment and Fishery Evaluation Report for the groundfish resources of the Bering Sea / Aleutian Islands region. North Pacific Fishery Management Council 605 West 4th Ave., Suite 306 Anchorage, AK 99501. 1297pp.

Ortiz, I., Wiese, F.K., Grieg, A., 2012. Marine Regions Boundary Data for the Bering Sea Shelf and Slope. UCAR/NCAR-Earth Observing Laboratory/Computing, Data, and Software Facility. Dataset. http://dx.doi.org/10.5065/D6DF6P6C.

Ortiz, I., Aydin, K., Hermann, A.J., Gibson, G., This issue. Climate to fisheries: a vertically integrated model for the eastern Bering Sea. Deep Sea Res. II 00, 00–00.

Pacific Fishery Management Council (PFMC). 2012. Terms of Reference for the Groundfish and Coastal Pelagic Species Stock Assessment and Review Process for 2013-2014. Pacific Fishery Management Council, 7700 NE Ambassador Place, Portland, OR 97220, USA.

Pikitch, E.K., Santora, C., Babcock, E.A., Bakun, A., Bonfil, R., Conover, D.O., Dayton, P., Doukakis, P., Fluharty, P., Heneman, B., Houde, E.D., Link, J., Livingston, P.A., Mangel, M., McAllister, M.K., Pope, J., Sainsbury, K.J., 2004. Ecosystem-based fishery management. Science 305, 346–347.

Punt, A.E., A’mar, T., Bond, N.A., Butterworth, D.S., de Moor, C.L., Oliveira, J.A.A., Haltuch, M.A., Hollowed, A.B., Szuwalski, C., 2014a. Fisheries management under climate and environmental uncertainty: Control rules and performance simulation. ICES J. Mar. Sci. 71, 2208-2220.

Punt, A.E., Butterworth, D.S., de Moor, C.L., De Oliveira, J.A.A. and M. Haddon. 2014b Management Strategy Evaluation: Best practices. Fish and Fisheries http://doi/101111/faf.12104.

Rose, K.A., Allen, J.I., Artioli, Y., Barange, M., Blackford, J., Carlotti, F., Cropp, R., Daewell, U., Edwards, K., Flynn, K., Hill, S,L., HilleRisLambers, R., Huse, G., Mackinson, S., Megrey, B., Moll, A,. Rivkin, R., Salihoglu, B., Schrum, C., Shannon, L., Shin, Y., Smith, S.L., Smith, C., Solidoro, C., St. John, M., Zhou, M., 2010. End-to-End modeling for the analysis of marine ecosystems: challenges, issues and next steps. Mar. Coast. Fish. 2, 115–130.

Sainsbury, K.J., Punt, A.E., Smith, A.D.M., 2000. Design of operational management strategies for achieving fishery ecosystem objectives. ICES J. Mar. Sci. 57, 731–741.

Siddon, E.C., Duffy-Anderson, J.T., Mueter, F.J., 2011. Community-level response of fish larvae to environmental variability in the southeastern Bering Sea. Mar. Ecol. Prog. Ser. 426, 225–239.

Siddon, E.C., Heintz, R.A., Mueter, F.J., 2013a. Conceptual model of energy allocation in walleye pollock (*Theragra chalcogramma*) from age-0 to age-1 in the southeastern Bering Sea. Deep-Sea Res. II 94, 140–149.

Siddon, E.C., Kristiansen, T., Mueter, F.J., Holsman, K.K., Heintz, R.A., Farley, E.V., 2013b. Spatial match-mismatch between Juvenile fish and prey provides a mechanism for recruitment variability across contrasting climate conditions in the eastern Bering Sea. PLOS ONE 8 (12), e84526.

Sigler, M.F., Heintz, R.A., Hunt, G.L. Jr., Lomas, M.W., Napp, J.M., Stabeno, P.J., This issue. A Mid-trophic View of Subarctic Productivity: Lipid Storage, Location Matters and Historical Context. Deep-Sea Res. II. 00, 00–00.

Sissenwine, M.P., Murawski, S.A., 2004. Moving beyond “intelligent tinkering”: advancing an ecosystem approach to fisheries. Mar. Ecol. Progr. Ser. 274, 291–295.

Smith, A.D.M., 1994. Management strategy evaluation – the light on the hill. *In* Population dynamics for fisheries management, pp. 249–253. Ed. by D. A. Hancock. Australian Society for Fish Biology, Perth.

Smith, A.D.M., Sainsbury, K.J., Stevens, R.A., 1999. Implementing effective fisheries management systems – management strategy evaluation and the Australian partnership approach. ICES J. Mar. Sci. 56, 967–979.

Stabeno, P.J., Kachel, N.B., Moore, S.E., Napp, J.M., Sigler, M., Yamaguchi, A., Zerbini, A.N., 2012. Comparison of warm and cold years on the southeastern Bering Sea shelf and some implications for the ecosystem. Deep Sea Res. II, 65-70, 31-45. doi:http://dx.doi.org/10.1016/j.dsr2.2012.02.020

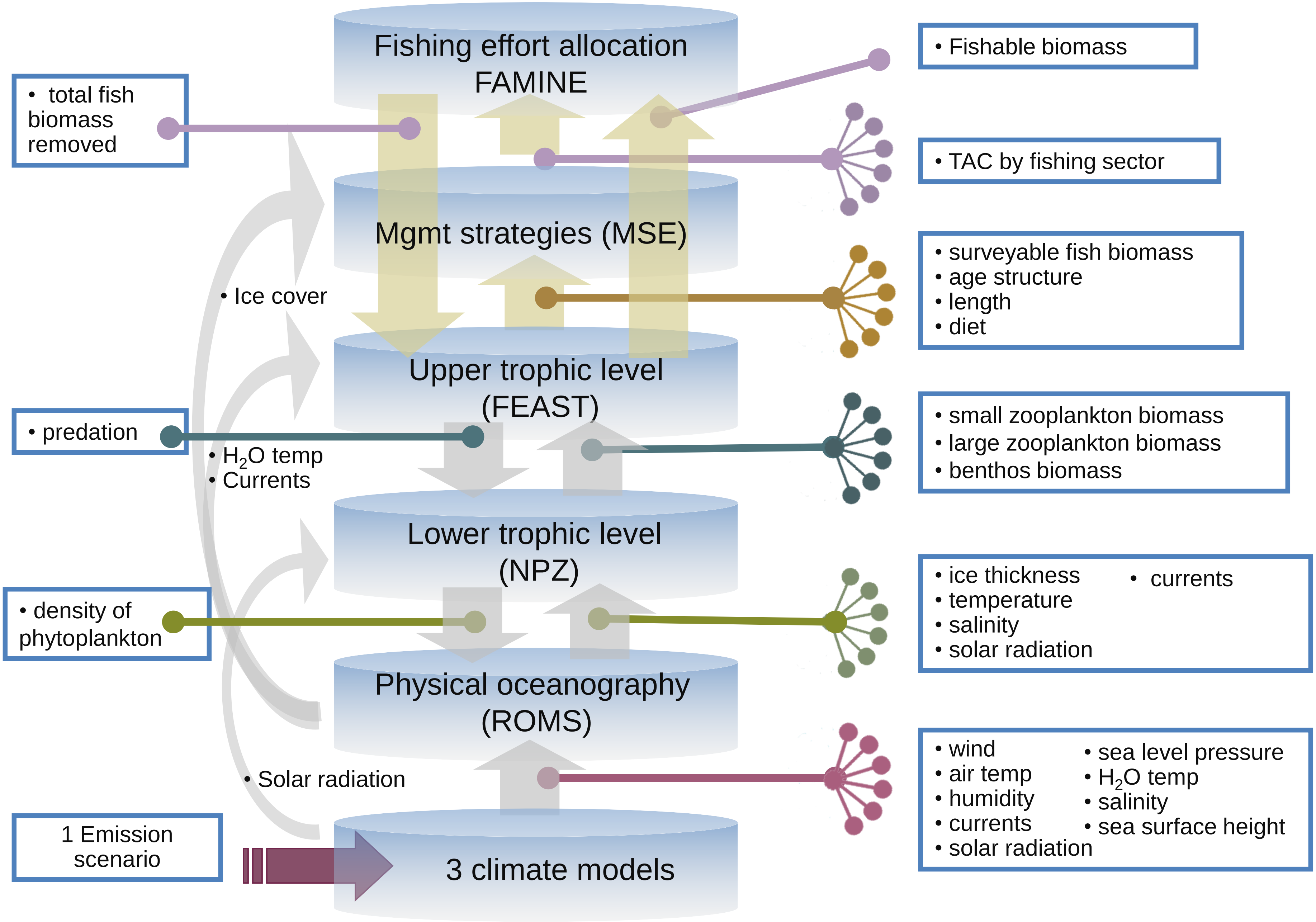
Travers-Trolet, M., Shin, Y-J., Field, J.G., 2014. An end-to-end coupled model ROMS-N2P2Z2D2-OSMOSE of the southern Benguela foodweb: parameterisation, calibration and pattern-oriente validation. African J. Mar. Sci. 36, 11-29.

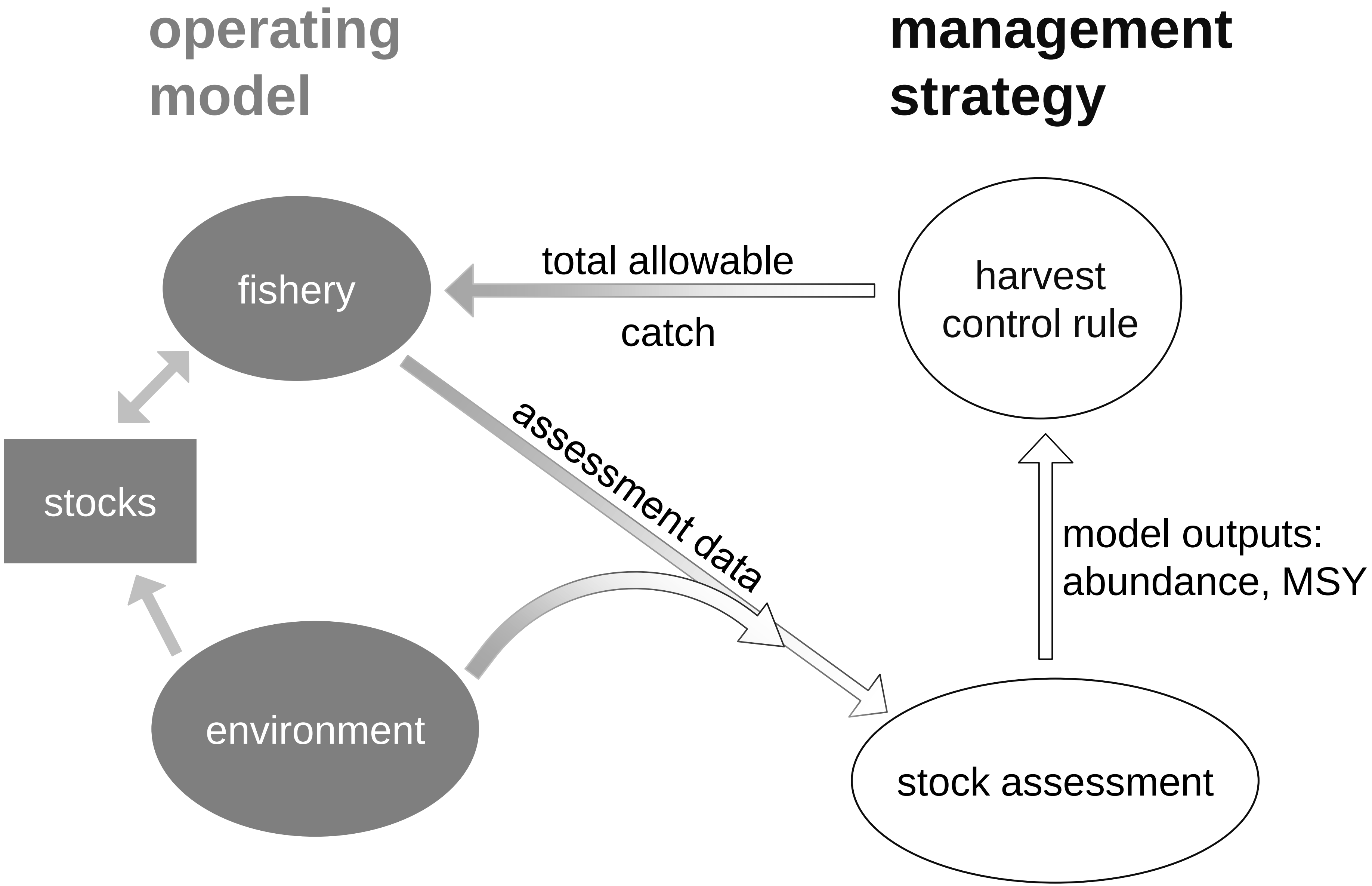
Uchimaya, T., Kruse, G.H., Mueter, F.J., This issue. A multispecies biomass dynamics model for investigating predator-prey interactions in the Bering Sea groundfish community. Deep Sea Res. II 00, 00–00.

Wang, M., Overland, J.E., Bond, N.A, 2010. Climate projections for selected large marine ecosystems. J. Mar.e Sys. 79, 258–266.

Wiese, F.K., Wiseman, Jr., W.J., Van Pelt, T.I., 2012. Bering Sea linkages. Deep Sea Res. II. 65-70, 2–5.

Witherell, D., Pautzke, C.P., Fluharty, D., 2000. An ecosystem-based approach for Alaska groundfish fisheries. ICES J. Mar. Sci. 57, 771–777.


Fishing effort allocation FAMINE
Mgmt strategies (MSE)
Upper trophic level (FEAST)
Lower trophic level (NPZ)
Physical oceanography (ROMS)
3 climate models
• total fish biomass removed
• predation
• density of phytoplankton
1 Emission scenario
• Ice cover
• H_2O temp
• Currents
• Solar radiation
• Fishable biomass
• TAC by fishing sector
• surveyable fish biomass
• age structure
• length
• diet
• small zooplankton biomass
• large zooplankton biomass
• benthos biomass
• ice thickness
• temperature
• salinity
• solar radiation
• currents
• wind
• air temp
• humidity
• currents
• solar radiation
• sea level pressure
• H_2O temp
• salinity
• sea surface height



operating model

management strategy

fishery

stocks

environment

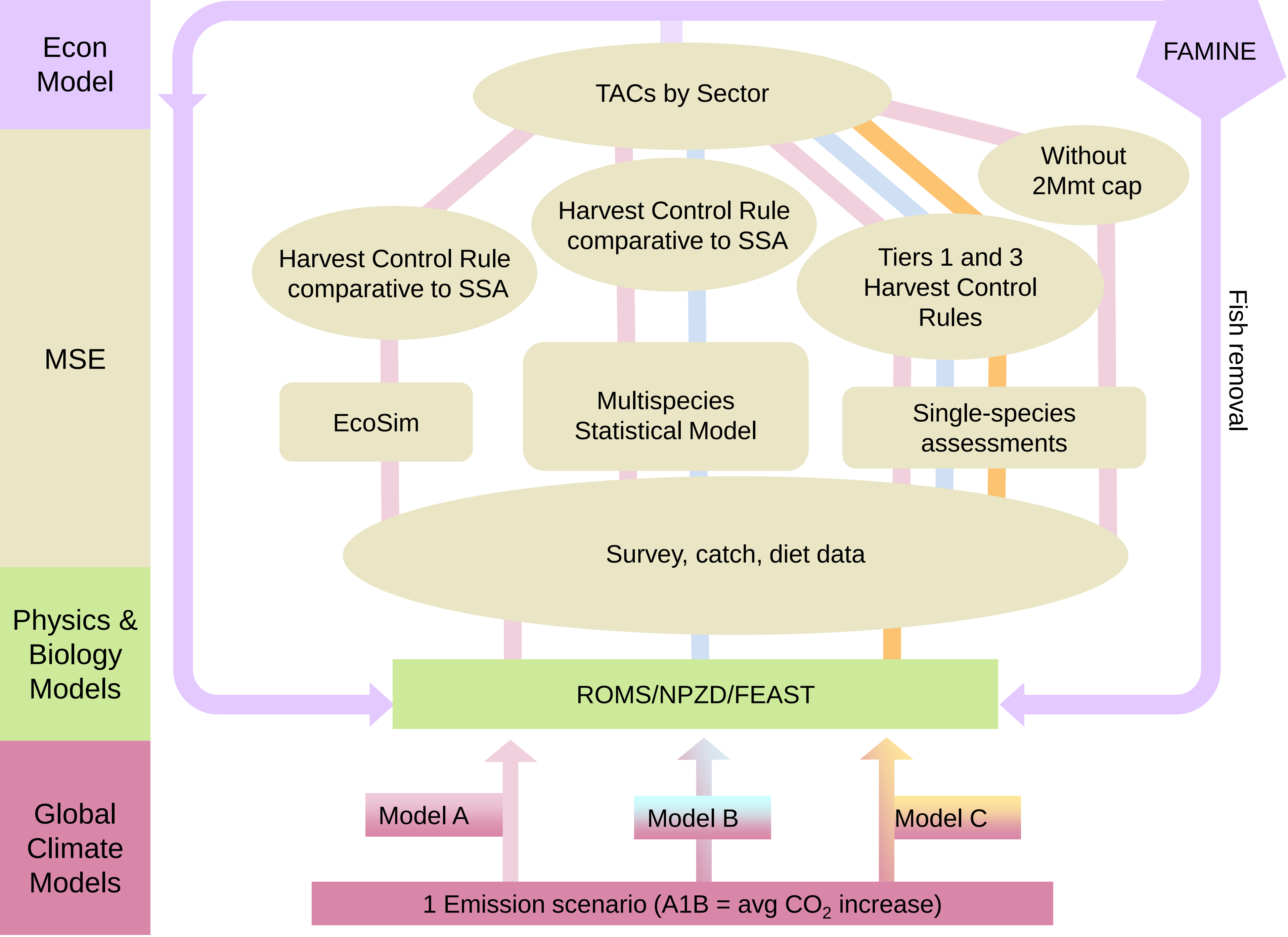
harvest control rule

stock assessment

total allowable catch

assessment data

model outputs: abundance, MSY


Econ Model
MSE
Physics & Biology Models
Global Climate Models
FAMINE
TACs by Sector
Without 2Mmt cap
Harvest Control Rule comparative to SSA
Harvest Control Rule comparative to SSA
Tiers 1 and 3 Harvest Control Rules
Fish removal
EcoSim
Multispecies Statistical Model
Single-species assessments
Survey, catch, diet data
ROMS/NPZD/FEAST
Model A
Model B
Model C
1 Emission scenario (A1B = avg CO_2 increase)