Fisheries and Oceans Canada | Pêches et Océans Canada
Ecosystems and Oceans Science | Sciences des écosystèmes et des océans

Canadian Science Advisory Secretariat (CSAS)

Research Document 2021/007

Pacific Region

A management procedure framework for groundfish in British Columbia

Sean C. Anderson[1], Robyn E. Forrest[1], Quang C. Huynh[2], Elise A. Keppel[1]

[1]Pacific Biological Station
Fisheries and Oceans Canada, 3190 Hammond Bay Road
Nanaimo, British Columbia, V9T 6N7, Canada

[2]Institute for the Oceans and Fisheries
University of British Columbia AERL, 2202 Main Mall
Vancouver, British Columbia, V6T 1Z4, Canada

December 2021

Foreword

This series documents the scientific basis for the evaluation of aquatic resources and ecosystems in Canada. As such, it addresses the issues of the day in the time frames required and the documents it contains are not intended as definitive statements on the subjects addressed but rather as progress reports on ongoing investigations.

Published by:

Fisheries and Oceans Canada
Canadian Science Advisory Secretariat
200 Kent Street
Ottawa ON K1A 0E6

http://www.dfo-mpo.gc.ca/csas-sccs/
csas-sccs@dfo-mpo.gc.ca

ISSN 1919-5044
ISBN 978-0-660-38696-6 Cat No. Fs70-5/2021-007E-PDF

Correct citation for this publication:

Anderson, S.C., Forrest, R.E., Huynh, Q.C., Keppel, E.A. 2021. A management procedure framework for groundfish in British Columbia. DFO Can. Sci. Advis. Sec. Res. Doc. 2021/007. vi + 139 p.

Aussi disponible en français :

Anderson, S.C., Forrest, R.E., Huynh, Q.C., Keppel, E.A. 2021. Un Cadre des procédures de gestion pour le poisson de fond en Colombie-Britannique. Secr. can. de consult. sci. du MPO. Doc. de rech. 2021/007. vi + 150 p.

TABLE OF CONTENTS

ABSTRACT

The Pacific Region Groundfish Integrated Fisheries Management Plan lists approximately 80 species-area fish stocks for which annual total allowable catches are required, most of which are applied as individual transferable quotas within the British Columbia (BC) integrated groundfish fishery. The majority of fish stocks encountered by the integrated groundfish fishery are considered data-limited, where data-limited stocks are defined as those with insufficient data to reliably estimate stock status or estimate abundance or productivity with conventional stock assessment methods such as statistical catch-at-age models. In recent decades, Fisheries and Oceans Canada (DFO) groundfish stock assessments have focused on data-rich stocks, resulting in a subset of stocks with full stock assessments, while many stocks with less informative data remain unassessed.

The DFO Sustainable Fisheries Framework, legislated via the Fish Stocks provisions in the *Fisheries Act*, requires that fish stocks be managed at sustainable levels—specifically at biomass levels above the Limit Reference Point (LRP). For data-limited stocks, data are often insufficient to adequately account for uncertainty in the assessment of stock status relative to biological reference points in traditional stock assessments. Instead of focusing on the explicit knowledge of current stock status, we propose a management-oriented approach that emphasizes selecting management procedures (MPs) that have a high likelihood of maintaining fish stocks above implicitly known reference points across multiple plausible states of nature, regardless of the quality and quantity of available data.

Worldwide there has been a movement towards MP (or management strategy evaluation) approaches to providing science advice on fish stocks via closed-loop simulation. Closed-loop simulation differs from conventional stock assessment because it simulates feedback between the implementation of MPs and a simulated system representing the fish stock and its environment, described by one or more operating models (OMs). This document presents a methodology for developing appropriate OMs, testing suites of MPs, and identifying MPs that best meet the objectives of fisheries management and stakeholders. We outline six best-practice steps for MP approaches: (1) defining the decision context, (2) setting objectives and performance metrics, (3) specifying OMs, (4) selecting candidate MPs, (5) conducting closed-loop simulations, and (6) presenting results to evaluate trade-offs. We then describe our proposed approach (the "MP Framework") and how it aims to accomplish each of these best-practice steps. Included in our framework are provisional conservation and fishery objectives and performance metrics based on Sustainable Fisheries Framework policies, a provisional library of data-limited MPs that are appropriate for BC groundfish stocks, and provisional visualizations to help decision-makers evaluate performance of MPs and trade-offs amongst MPs.

We undertake a case study of the Rex Sole (*Glyptocephalus zachirus*) stock in the West Coast Vancouver Island groundfish management area (Area 3CD) to demonstrate an application of the MP Framework. The case study develops six reference-set OMs and two robustness-set OMs. The case study reveals a set of survey-index-based MPs, constant catch, and surplus-production-based MPs that achieve > 0.9 probability (9 times out of 10 chance) of maintaining biomass above the LRP in the long term (35–50-years in the future) while maintaining a > 0.8 probability (4 times out of 5 chance) of maintaining catches at or above recent (5-year) average levels in the near future (1–10 years) in the reference-set OMs. We also present performance metrics related to the long-term probability of biomass remaining above the Upper Stock Reference, the long-term probability of fishing below F_{MSY} (fishing mortality at maximum sustainable yield), the long-term probability of maintaining catches above recent aver-

age levels, and the probability of catch variability remaining below historical levels. Four of the MPs achieved only slightly lower performance metrics in the robustness OMs, compared to the reference-set OMs, while other MPs were more sensitive to these OM robustness scenarios.

We highlight issues regarding reference points, MP tuning, assessment frequency and triggers, the inclusion of environmental effects, assessing the value of information, and use of this framework as part of stock rebuilding plans. Throughout, our framework emphasizes transparency and reproducibility and to that end we develop an associated package for the statistical software R that facilitates applications of the framework. Overall, we intend this framework to improve the capacity for Pacific DFO Science to provide evidence-based catch advice for more groundfish stocks—regardless of data limitations—in a standardized and transparent manner consistent with the DFO Sustainable Fisheries Framework, the Fish Stocks provisions in the *Fisheries Act*, and international best practices.

1 INTRODUCTION

1.1 BACKGROUND

Fishery-independent surveys conducted by Fisheries and Oceans Canada (DFO) encounter over 200 groundfish species in the Pacific region. Of these, approximately 100 are regularly caught in the British Columbia (BC) integrated groundfish fishery. The Pacific Region Groundfish Integrated Fisheries Management Plan (IFMP)(DFO 2017a) lists approximately 80 species-area fish stocks for which annual total allowable catches (TACs) are required, most of which are applied as individual transferable quotas (ITQs) within the BC integrated groundfish fishery.

In recent decades, advice on TACs for Pacific groundfish species has been developed using statistical stock assessment models that rely on indices of abundance, age-composition, commercial catch, and biological data that enable estimation of key population metrics such as stock status, relative depletion, recruitment, growth, fishery selectivity, and natural and fishing mortality. These stock assessment models are data and resource-intensive, with only a few assessments produced each year with available data and Science sector resources.

The majority of fish stocks encountered by the integrated groundfish fishery are considered data-limited, where data-limited stocks are defined as those with insufficient data to: (1) reliably estimate stock status; or (2) estimate abundance or productivity with conventional stock assessment methods such as statistical catch-at-age models (Dowling et al. 2015a, 2015b). Many of these stocks lack current assessment advice.

1.2 MOTIVATION

The Canadian Sustainable Fisheries Framework (SFF) lays the foundation for the precautionary approach (PA) to fisheries management in Canada (DFO 2006, 2009). The Precautionary Approach Framework (DFO 2009) relies on the definition of biological reference points (BRPs), which define biomass targets as well as low biomass thresholds to be avoided with high probability. The approach requires that fishing mortality be adjusted in relation to two levels of stock status—an Upper Stock Reference (USR) and a Limit Reference Point (LRP) (Figure 1). The LRP and USR delineate three stock status zones ("healthy," "cautious," and "critical"). Of particular importance is the LRP, which is defined as the spawning biomass below which serious harm may occur to the stock. Other key elements of the Precautionary Approach Framework include a harvest control rule (HCR), which determines allowable removal rates in each of the three stock status zones (e.g., Figure 1), and the requirement to take into account risk and uncertainty when developing BRPs and determining stock status in relation to BRPs (DFO 2006, 2009).

On June 21, 2019, major amendments to Canada's Fisheries Act received Royal Assent, thus passing them into Canadian law. Among many other amendments, provisions in the new *Fisheries Act* require that fish stocks be managed at sustainable levels, specifically at biomass levels above the LRP. Hereafter referred to as the "Fish Stocks provisions," the provisions legislate elements of Canada's Sustainable Fisheries Framework and Precautionary Approach Framework, and state:

> **Measures to maintain fish stocks** 6.1(1) In the management of fisheries, the Minister shall implement measures to maintain major fish stocks at or above the level necessary to promote the sustainability of the stock, taking into account the biology of the fish and the environmental conditions affecting the stock.

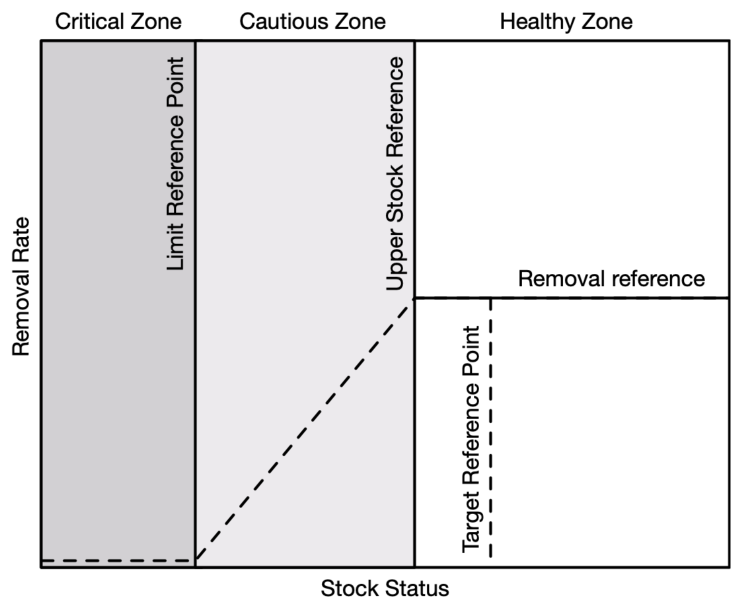

Figure 1. Illustration of DFO's Precautionary Approach Framework. Based on DFO (2009).

> **Limit reference point** 6.1 (2) If the Minister is of the opinion that it is not feasible or appropriate, for cultural reasons or because of adverse socio-economic impacts, to implement the measures referred to in subsection (1), the Minister shall set a limit reference point and implement measures to maintain the fish stock above that point, taking into account the biology of the fish and the environmental conditions affecting the stock.
>
> **Plan to rebuild** 6.2(1) If a major fish stock has declined to or below its limit reference point, the Minister shall develop a plan to rebuild the stock above that point in the affected area, taking into account the biology of the fish and the environmental conditions affecting the stock, and implement it within the period provided for in the plan.

For data-limited stocks, data are generally insufficient to adequately account for uncertainty in development of BRPs and assessment of stock status relative to BRPs. To comply with the legal requirements identified in the Fish Stocks provisions for data-limited species, it is therefore necessary to develop defensible methods and a framework for setting catch limits that promote sustainability and maintain stocks above their LRP, recognizing that, in many cases, the LRP and stock status cannot be reliably estimated.

The objective of the framework proposed in this document is to preserve the *intent* of the Precautionary Approach Framework, legislated in the Fish Stocks provisions, and provide decision-makers with a quantified characterization of risk with respect to the likelihood of maintaining stocks above the LRP, regardless of the quality and quantity of available data.

1.3 TIERED APPROACHES

Many other nations and governing bodies have established frameworks for assessing data-limited fish stocks—often through "tiered" approaches (e.g., United States (Newman et al. 2015), ICES (ICES 2012), Australia (Smith et al. 2009)). A tiered approach places stocks into ranked categories (tiers) based on the types and quality of available data. Specific data-limited methods for setting catch limits are identified for each tier. Arbitrary buffers may also be added to recommended catch limits in each tier (e.g., reduce recommended catch limit by 20%), in an attempt to account for the additional uncertainty inherent in more data-limited tiers.

In 2016, a CSAS workshop was held in the Pacific region to review international tiered approaches, as a preliminary evaluation of the suitability of a tiered approach for the BC integrated groundfish fishery (DFO 2016). Instead of recommending a discrete tiered approach based on data-availability, the workshop recommended considering data-richness on a continuous scale and simulation-testing multiple alternative management procedures (MPs) on a stock-by-stock basis, using a management-oriented, or management strategy evaluation (MSE), approach to identify data-limited MPs that would best meet policy and fishery objectives, explicitly accounting for risk.

1.4 MANAGEMENT-ORIENTED APPROACHES

Worldwide, there has been a movement towards management-oriented approaches to stock assessment and fisheries management (e.g., de La Mare 1998; Butterworth and Punt 1999; Plagányi et al. 2007; Rademeyer et al. 2007; Berkson and Thorson 2015; Geromont and Butterworth 2015a; Carruthers et al. 2016; Punt et al. 2016). Whereas traditional stock assessment methods focus explicitly on estimation of biological parameters and stock status, a management-oriented approach focuses on identifying MPs that perform best with respect to policy and fishery objectives when implemented in a “closed-loop” simulation environment (Figure 2).

Closed-loop simulation is distinguished from conventional stock assessment because it simulates feedback between implementation of the MPs and the simulated system representing the fish stock and its environment, described by one or more operating models (OMs). This approach not only takes into account the effect of the MPs on the system but also the future data collected from the system and its use in the MPs (Punt et al. 2016). Data-limited MPs may be represented by a simple empirical rule, such as an adjustment to catch based on a change in an index of abundance or may be based on more complex methods incorporating multiple data sources and statistical models.

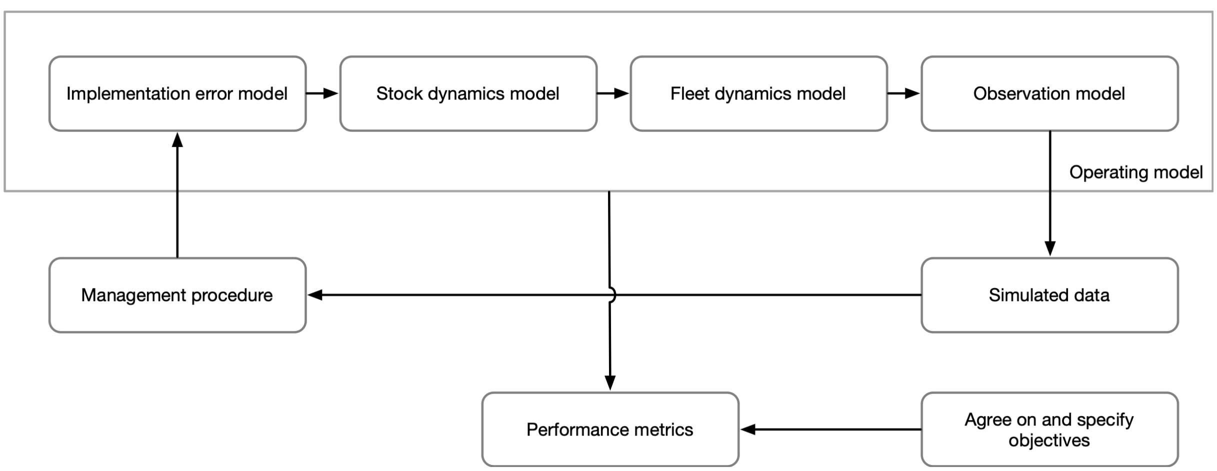

Figure 2. Illustration of the fisheries closed-loop simulation process following Punt et al. (2016). The management procedure may be based on a simple data rule (e.g., decrease the allowable catch x% if the survey index decreases y%) or it might be an estimation model combined with a harvest control rule.

The scientific literature reports a large number of data-limited MPs. When tested in closed-loop simulation environments, data-limited MPs have shown variable performance across conditions such as species life-history and levels of biomass depletion (e.g., Carruthers et al. 2014, 2016). Closed-loop simulation-testing of data-limited MPs allows for evaluation of relative performance of MPs across a range of uncertainties in, for example, underlying fish biology, observation error, estimation error, and implementation error (e.g., Kell et al. 2006; Carruthers et al. 2016). Importantly, a management-oriented approach employing closed-loop simulation enables ranking of MPs with respect to meeting objectives and enables evaluation of trade-offs among competing objectives (e.g., policy-based conservation objectives and economic objectives).

Since 2017, a partnership agreement between the University of British Columbia (UBC) and DFO (DFO 2017b) has supported development of two open-source software packages for MSE, implemented in R (R Core Team 2019): the Data Limited Methods toolkit (DLMtool) (Carruthers and Hordyk 2018a; Carruthers and Hordyk 2018b) and the Management Strategy Evaluation toolkit (MSEtool) (Huynh et al. 2019). After several years of development, these packages provide some of the fastest, most flexible and extensible software for conducting MSE for fisheries, ranging from data-poor to data-rich, enabling rapid assessment of multiple MPs according to customizable conservation and fisheries objectives, and evaluation of key trade-offs.

1.5 OBJECTIVES OF THIS DOCUMENT

This document aims to develop a management procedure framework ("MP Framework") for selecting methods for provision of catch advice for Pacific groundfish stocks that do not have sufficient data to estimate the parameters needed for full statistical stock assessments. Therefore, the framework will provide advice for stocks where biological reference points and stock status cannot be reliably estimated. In doing so, it aims to improve the capacity of DFO to meet the requirements of Canada's Sustainable Fisheries Framework and the Fish Stocks provisions of the *Fisheries Act* for data-limited species. The MP Framework will apply closed-loop simulation to formalize the process of testing and selecting MPs for data-limited groundfish stocks, which will support the provision of scientific advice to fisheries managers. We emphasize that biological reference points and stock status, which are integral to the Sustainable Fisheries Framework and Fish Stocks provisions, are also integral to the MP Framework, even if they are not explicitly stated in the catch advice. Instead reference points and stock status are incorporated into the OMs (see Appendix A) and management procedures are selected based on explicitly defined probabilities of maintaining OM stock status above reference points. We choose the terminology MP Framework to clearly specify the focus on selection of best-performing MPs, rather than focusing on the available data.

This document presents a methodology for developing appropriate OMs, testing suites of MPs, and identifying MPs that best meet the objectives of fisheries management, First Nations and stakeholders. This work builds on recent work by Anderson et al. (2019), which is a comprehensive, reproducible data synopsis for nearly all Pacific groundfish species encountered by DFO surveys and commercial fisheries in BC. Although the proposed framework is software agnostic, in the current implementation, data and empirical analyses from the groundfish synopsis packages (Anderson et al. 2020c; Keppel et al. 2020) are used to develop OMs using the DLMtool (Carruthers and Hordyk 2018a) and MSEtool packages (Huynh et al. 2019), implemented in the statistical software R (R Core Team 2019) (now collectively named openMSE). The OMs are then used in simulation-testing a range of data-limited MPs against performance metrics. While this document focuses on data-limited MPs, the main features of the framework could be extended to more data-rich species to evaluate performance of data-moderate or data-rich MPs (e.g., different configurations of statistical catch-age models coupled with harvest control rules).

An R package written by the authors for this framework, ggmse (Anderson et al. 2020b) (formerly gfdlm), houses many tools and custom visualizations to aid in the process. Steps of the process are described more fully in the following sections of the document.

1.6 EXPECTED BENEFITS

There are a number of expected benefits of this project. Compared to the current status quo, this framework will allow DFO to:

- provide evidence-based catch advice for more stocks, including stocks for which advice was not previously available;
- develop a standardized and transparent approach across stocks;
- test performance of data-limited MPs for providing catch advice with respect to meeting sustainability and fishery objectives; and
- help build an understanding of the most important data needs and research priorities for reducing uncertainty in stock assessment advice.

Critically, the framework will develop methods to improve capacity for DFO Science to provide sustainable, data-limited catch advice, consistent with Canada's Precautionary Approach Framework as legislated under the Fish Stocks provisions, therefore enabling data-limited stocks to be brought into compliance with the Fish Stocks provisions of the *Fisheries Act*.

Extensions to the framework could be used to inform:

- rebuilding plans and recovery potential assessments (e.g., Haggarty et al. 2021);
- methods for accounting for environmental conditions in management decisions, as required under the Fish Stocks Provisions;
- assessing the value of collecting more information; and
- evaluating the performance of data-moderate and data-rich MPs.

See Section 5 for discussion of these.

We emphasize that while the framework does not explicitly report biological reference points and stock status, which often cannot be reliably estimated for data-limited stocks, the framework is designed to evaluate MPs with respect to the probability of maintaining stocks above reference points over a defined range of uncertainties. Reference points are therefore an integral component of the framework.

2 BEST PRACTICES FOR MANAGEMENT PROCEDURE APPROACHES

Punt et al. (2016) reviewed best practices for MSE and identified five key steps in the process (Steps 2–6 below). In large part, the DLMtool software has been designed to allow practitioners to follow these steps [Figure 3; Carruthers and Hordyk (2018a)]. We also identify a critical first step (Step 1 below): defining the decision context (Gregory et al. 2012; Cox and Benson 2016). In most practical applications the steps in the MSE process will be iterative. For example, objectives or performance metrics may be refined after experience is gained with their performance in simulations or the real world (de La Mare 1998; Plagányi et al. 2007; Cox and Kronlund 2008; Punt et al. 2016). OMs and MPs may also be refined or revised in light of new information or changes to fishing operations or the ecosystem (e.g., Plagányi et al. 2007; Pestal et al. 2008).

In this section we provide an overview of the six best practice steps. We describe how each of the six steps is implemented in the proposed MP Framework in Section 3.

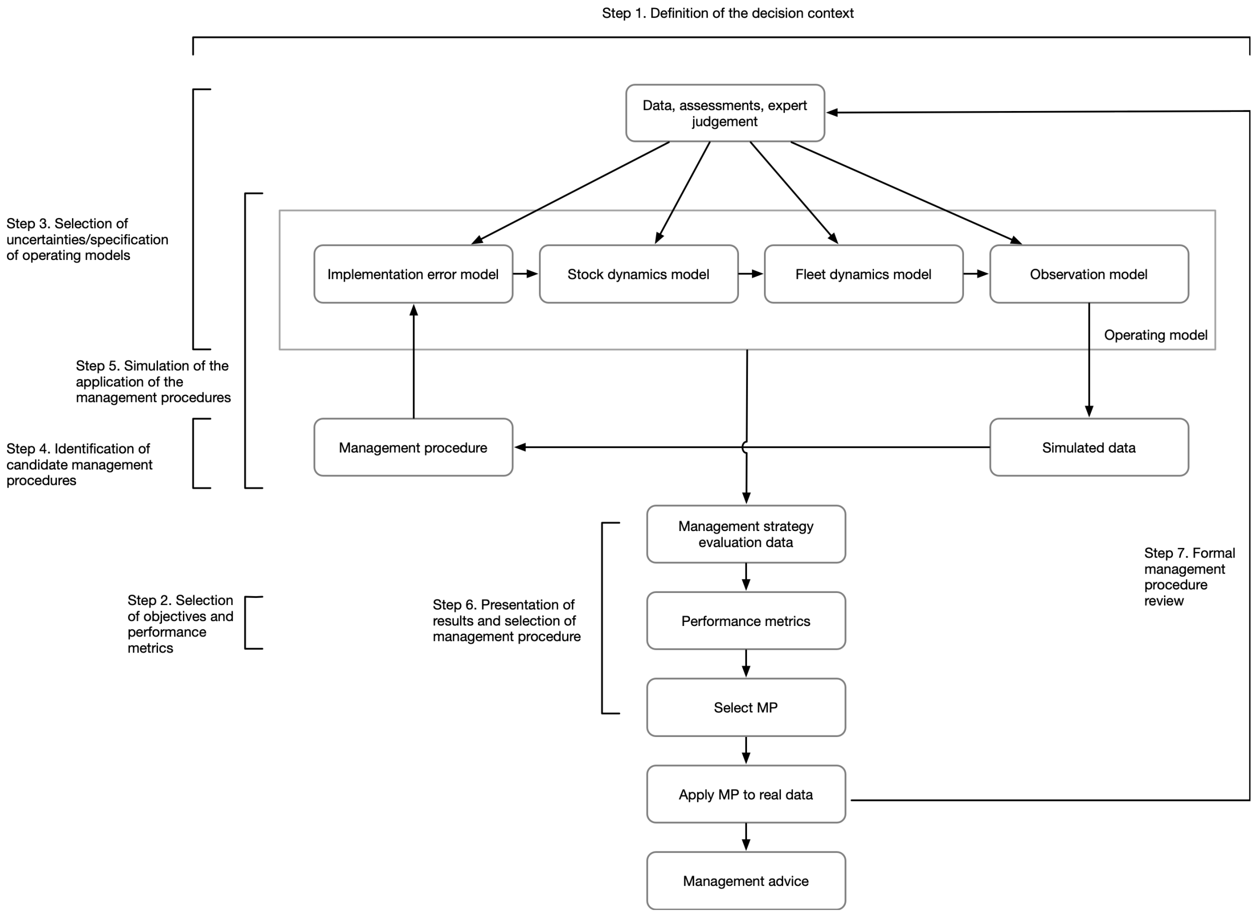

Figure 3. The steps of the MSE process following Punt et al. (2016) as implemented in DLMtool. Adapted from Carruthers and Hordyk (2018a). This figure expands on Figure 2.

2.1 STEP 1: DEFINE THE DECISION CONTEXT

Key questions to guide defining the decision context for MSE include:

- What is the exact decision to be made?
- What is the time frame for making the decision?
- How often will the decision be evaluated and updated? E.g., will a TAC decision be in place for one year or several years?
- What are the boundaries on the project and decision?
- What are the legislative and policy requirements?
- What are specific roles and responsibilities of parties involved? Parties include Science, Management, First Nations, industry, academia, and/or non-governmental organizations (NGOs).
- How will the final decision be made? For example, it may be necessary to rank or weight objectives if there are large trade-offs with respect to performance for different objectives.

- How will the process be governed? For example, how will acceptability of trade-offs be determined? How will meetings be facilitated? How will consultation be managed? Failures in governance of the decision-making process can lead to a less successful decisions in terms of acceptance and compliance (Smith et al. 1999; Armitage et al. 2019).

Definition of the decision context is the role of managers, stakeholders, First Nations, and other key interested parties. Engagement of resource users at all stages of the MSE cycle from the outset is critical, as it increases the likelihood that the process will be considered credible, objectives will reflect real objectives, and MPs will be successfully implemented as planned (Smith et al. 1999; Punt et al. 2016; Armitage et al. 2019).

2.2 STEP 2: SELECTION OF OBJECTIVES AND PERFORMANCE METRICS

Clear management and fishery objectives and the performance metrics that measure them must be identified. Objectives may initially be high level and "strategic" (e.g., achieve sustainable fisheries, maintain economic prosperity, maintain cultural access) but these must be converted into operational "tactical" objectives that can be expressed as quantitative performance metrics (de La Mare 1998; Hilborn 2007; Punt et al. 2016). Fully quantified objectives include a metric, the desired probability of success, and a time frame to achieve the objective (e.g., probability of maintaining the stock above the LRP is greater than 0.95 [19 times out of 20], in each and every year of a 50-year period).

Since the properties of the underlying system represented by the OM are known exactly, a wide range of biological and economic metrics can be calculated from the OM (Carruthers and Hordyk 2018a). However, having too many performance metrics can make the final decision process complex. Performance metrics should be chosen so they can be understood by decision-makers and participants, and to facilitate a tractable decision-making environment (Punt et al. 2016).

Objectives should be developed with the participation of managers, stakeholders, First Nations, and other interested parties (e.g., Smith et al. 1999; Plagányi et al. 2007; Cox and Kronlund 2008; Mapstone et al. 2008). Hilborn (2007) identified four categories of fishery objectives: biological, economic, social and political, noting that many conflicts in fisheries in fact arise from conflicting objectives. Within each of these categories, different resource users and interest groups will place value on different components, leading to inevitable trade-offs. For example, while most resource users may place higher value on increased biological production, specific resource user groups, such as aboriginal users, may prefer more stable access to yields and increased participation (Plagányi et al. 2013) or have spatial objectives (Okamoto et al. 2020). Recreational users may prefer lower yields and larger trophy fish (Hilborn 2007). Achieving agreement on the list of objectives and performance metrics, especially with multiple user groups is critical, can be time-consuming and should be iterative, as participants build familiarity with each other and the process. Incorporating social and cultural objectives may be particularly important and is lacking in many decision making processes (Stephenson et al. 2017; Benson and Stephenson 2018; Okamoto et al. 2020; but see, for example, Plagányi et al. 2013; DFO 2019, 2020). During Step 1 (Section 2.1), attention should be given to governance of the process to ensure that there is meaningful participation by different groups and that participants can agree on a final set of objectives to adequately represent their interests and characterize the main trade-offs (Stephenson et al. 2017; Okamoto et al. 2020). An overview of possible approaches is provided in Benson and Stephenson (2018).

2.3 STEP 3: SELECTION OF UNCERTAINTIES/SPECIFICATION OF OPERATING MODELS

Uncertainties inherent in the underlying system are represented in the OM. Uncertainty in the OMs may be related to: the biology of the stock (e.g., growth, natural mortality, recruitment, migration); the dynamics of the fleet (e.g., targeting behaviour, selectivity of the fishing gear); the observation process (e.g., bias or imprecision in survey data or age/length composition data); and/or the implementation process (e.g., exceeding catch limits) (Carruthers and Hordyk 2018a).

Some of this uncertainty (e.g., range of values of natural mortality or other parameters) may be captured within a single OM by expressing distributions for these parameters. However, it is unlikely that the full range of uncertainties thought to influence the system can be captured in a single OM. Therefore, best practice recommends dividing MSE trials into a "reference set," using a core set of OMs that include the most important uncertainties (e.g., depletion of the stock or range of natural mortality values), and a "robustness set," representing other plausible OM formulations that represent alternative structural hypotheses (Rademeyer et al. 2007). These authors recommend that the reference set of OMs include the most important uncertainties, which are both highly plausible and have major impacts on results. While there is no established formal means for selecting OMs for the reference set, Punt et al. (2016) suggest that best practice start from a common set of factors which commonly have a large impact on MSE performance due to uncertainties. They provide a list of factors which commonly have a large impact on MSE performance due to uncertainty (their Table 3) and suggests that, at a minimum, MSE processes should consider: parameter uncertainty (related to productivity and stock size); process uncertainty; and observation error (see also Cooke 1999). Implementation uncertainty is another important source of uncertainty, for example uncertainty in actual catches relative to TACs. Where available, stock assessments can be consulted for the major sources of uncertainty. For example, the initial reference set may be based on the set of sensitivity analyses presented in the stock assessment (possibly with additional scenarios), while the robustness set may include a broader, more exploratory set of uncertainties [e.g., Atlantic Bluefin Tuna (*Thunnus thynnus*), T. Carruthers, personal communication, May 12 2020].

Interactions among uncertainties may be considered by evaluating all combinations of uncertainty factors (e.g., Rademeyer and Butterworth 2006a). This may be presented in grid format (e.g., Carruthers et al. 2020). However, this may not always be computationally possible, and it is more common to select 'base' levels for each factor and then develop alternative OMs which vary one (or more than one) factor in turn (Punt et al. 2016). An iterative approach may be required, where sensitivity tests are run to determine which combinations produce the largest differences in results (e.g., Rademeyer and Butterworth 2006b; Carruthers et al. 2020).

Once an agreed-upon reference set of OMs has been determined, a wider range of OMs (the robustness set) should be developed to capture a wider range of uncertainties that may be less plausible but should nonetheless be explored (Rademeyer et al. 2007). These may include effects related to environmental change (e.g., time-varying mortality, climate-driven recruitment, predator-prey relationships); structural representation of population dynamics (e.g., form of the stock-recruit relationship); or fleet dynamics (e.g., selectivity). Punt et al. (2016) also note that, in some cases, where the data used to parameterize the OM are in conflict (e.g., two indices of abundance are in conflict), the best practice may be to develop alternative OMs based on the different data sources. Other uncertainties in past reliability or future availability of data may also be captured in the robustness set (Rademeyer et al. 2007).

Reference and robustness sets may be selected through an iterative process examining the impact of uncertainties on the MSE performance. For example, Rademeyer and Butterworth (2006b) evaluated 28 preliminary robustness tests for a South African hake case study but discontinued tests that produced results very similar to the reference set trials. OMs may be weighted, where weightings may be based on qualitative plausibility criteria (Butterworth et al. 1996), or may be quantitative based on model-selection criteria based on fits to data (e.g., AIC [Akaike information criterion]). However, Punt et al. (2016) urged caution in using model-selection criteria to weight OMs unless there was very high confidence in the reliability of the likelihood function, which is unlikely to be the case in data-limited situations. Placing less plausible OMs into the reference set may somewhat alleviate the need to consider weighting of OMs.

Ideally, OMs should be conditioned on real data to ensure they can reproduce historical observations (e.g., Cox and Kronlund 2008; Forrest et al. 2018). In data-limited cases without reliable historical observations, this may not be possible. In these cases, best practice would be to develop a set of OMs that differ in terms of major uncertainties, especially related to stock productivity and current depletion level.

Development of OMs is principally the responsibility of Science, although input from stakeholders, First Nations and other parties is desirable, especially with respect to identifying key uncertainties and ensuring plausibility of the OMs.

2.4 STEP 4: IDENTIFICATION OF CANDIDATE MANAGEMENT PROCEDURES

The scientific literature now reports many MPs for data-limited fisheries, more than 80 of which have been integrated into the DLMtool software (Carruthers et al. 2016; Carruthers and Hordyk 2018a). Management procedures for fisheries managed by catch limits are generally either model-based, where data are integrated into a stock assessment model and outputs are used to calculate catch limits, or empirical, where data are used in an algorithm to directly determine the catch limit (e.g., adjustment of catch based on change in index of abundance) (Punt et al. 2016). Empirical MPs can make use of a variety of data types including catch, population indices, fish lengths, and fish ages.

Empirical MPs take data sampled from the system, such as a survey index, apply an algorithm, and make a catch recommendation. An example is the “Iratio” MP (ICES 2012; Jardim et al. 2015), where the mean survey index value from the last two years is divided by the mean survey index value three to five years before present. This provides a ratio indicating whether the survey has increased or decreased, which is then multiplied by the previous year’s catch to generate a new catch recommendation. If the survey index has been trending up, then the catch recommendation will increase, and vice versa. Model-based MPs fit a statistical population model (e.g., surplus production model) to observed data to estimate biological reference points and stock biomass. These are then incorporated into a harvest control rule (e.g., Figure 1) to determine the catch limit for the following year.

Given the large number of MP options available, a screening step is desirable. For example, MPs that do not return a catch limit (e.g., spatial closures or effort-based MPs) can be immediately screened out if management requires a catch limit. Also, unless the decision context involves considering the value of collecting new information, it is important to test MPs for which information or data are available (Punt et al. 2016). For example, MPs that rely on age-composition data or an estimate of current depletion may not be feasible for many data-limited BC groundfish stocks. It is also important to consider the period of time for which MPs will be left in place.

Unless MPs will be updated annually, it will be important to design MPs that are left in place for several years and test their performance (e.g. Haggarty et al. 2021). While it is important to work with a manageable set of MPs, it is also important not to screen too aggressively, to make sure good candidate MPs are not screened out early on.

In general, identification of available MPs is the role of Science. Managers, stakeholders and First Nations should be involved in identifying desirable MPs (e.g., Cox and Kronlund 2008; Plagányi et al. 2013) and provide input on feasibility of implementing some MPs and their likely success in terms of acceptance and compliance (Armitage et al. 2019).

2.5 STEP 5: SIMULATION OF THE APPLICATION OF THE MANAGEMENT PROCEDURES

Once the OM and MPs are fully specified, the closed-loop simulation replicates can be run, following the process illustrated in Figure 3. Critically, the simulations include feedback between the OM and the MP, where the OM generates data at each time step, which is used to apply the MP, which generates a catch recommendation, which is removed from the OM, which generates the next time step of data, and so forth until the projection period is complete.

Typically, a large number of replicate simulations are run for each OM-MP combination. Replicates may differ in terms of OM process error, observation errors and random draws from ranges of OM parameters, meaning that each replicate provides a different set of simulated data to the MPs. The number of replicates should be selected to ensure that performance metrics can be calculated with adequate precision (Punt et al. 2016), which can be indicated by MPs being consistently ranked in the same order regardless of additional replicates (Carruthers and Hordyk 2018a). The MSE should output enough information to calculate performance metrics for the MPs, and also to evaluate the behaviour and performance of the MSE itself (e.g., whether all trials converged, ranges of OM parameter values, and trajectories of key OM variables such as biomass and catch).

There may be a need to reduce the number of candidate MPs to a manageable set. Analysts can screen out MPs that do not meet a basic set of requirements for a broad range of stocks (e.g., MPs that result in a high probability of stocks being below the LRP). Such a procedure of screening out poorly performing MPs has been termed “satisficing” (Miller and Shelton 2010), where MPs must meet a minimum-defined standard to be accepted. Satisficing criteria may be used at the screening stage and can also be used at the final MP selection stage to help streamline the decision-making process. Satisficing criteria may be less strict at the preliminary screening stage, to ensure that potentially successful MPs are not screened out of the process too early.

Running the simulations is the role of Science. Feedback from managers, stakeholders and First Nations should be sought throughout the process, to enable iterative refinement of the models and outputs (e.g., Cox and Kronlund 2008).

2.6 STEP 6: PRESENTATION OF RESULTS AND SELECTION OF MANAGEMENT PROCEDURE

Selection of an MP involves addressing trade-offs (e.g., between conservation and economic performance metrics), and therefore is the purview of managers, stakeholders, First Nations, and interested parties (Punt et al. 2016). Ultimately, selection of an MP may be a subjective process, depending on the magnitude of trade-offs. It may be necessary to rank performance metrics in order of priority before the process starts. The role of Science in this step is to ensure that results are clearly presented to decision-makers. Ideally this should include presentation of graphical outputs that enable clear comparison of MPs with respect to performance metrics and trade-offs (Punt 2017).

Two basic approaches may be used in selecting the final MP: satisficing and trade-off evaluation, where satisficing involves setting minimum performance standards (described in Section 2.5) and trade-off evaluation involves decision-makers and stakeholders finding a balance among competing performance metrics (Punt 2017). We already described an early satisficing step to screen out poorly performing MPs. A second satisficing step may be used towards the end of the process, to further screen out MPs that do not meet a minimum standard and to simplify the decision-making environment (Miller and Shelton 2010). After this, MP selection may move to a trading-off stage with a final reduced set of MPs. An iterative process may also be required, where MPs and/or OMs are refined following examination of results (e.g., Cox and Kronlund 2008). In cases where there is a reference and robustness set of OMs, OMs can be weighted on the basis of plausibility, although this may require a qualitative, expert-driven approach and may not be straightforward (Punt et al. 2016).

Carruthers and Hordyk (2018a) also discuss a final step (Step 7 in Figure 3), which is formal review of the selected MP once it has been implemented with real data. Formal review includes evaluation of whether the MP is performing as expected. For example, this could be done by comparing whether real relative abundance indices follow similar trajectories to those predicted by the OMs under the selected MP. In this document, we do not demonstrate this formal review, but recognize that ongoing review of the performance of MPs following their application is a critical component of MSE, where OMs and MPs may be continuously refined as new data become available (Cox and Kronlund 2008; Carruthers and Hordyk 2018b).

Selection of the final MP should ideally result from careful specification of the objectives and performance metrics. In cases where a trade-off remains, or multiple MPs achieve sufficient performance, it is the role of managers—with input from stakeholders and First Nations and advice from Science—to select the final MP.

3 METHODS: THE PROPOSED FRAMEWORK

We present the steps of a proposed framework for selecting MPs based on trade-offs in performance amongst MPs. Sections 3.1 to 3.6 describe the methods of the framework, as it would be applied in the provision of catch advice, organized according to the six best-practice steps described in Section 2. We note that the elements of the framework as laid out here, represent primarily the role of Science. When this framework is applied in provision of management advice, decision-makers, stakeholders, and other interested parties (e.g., First Nations, Nongovernmental Organizations [NGOs], and academics) should be engaged throughout the process, particularly in defining the decision context, setting objectives and performance metrics, and selecting MPs (e.g., Cox and Kronlund 2008).

3.1 STEP 1: DEFINE THE DECISION CONTEXT

For quota-managed groundfish species in BC, the decision to be made is which MP to use to determine catch limits for the period until the next available catch advice. The time frame for making the decision should be stated in the request for science advice. The boundaries on the project should be decided by a “technical committee,” convened for each assessment, and typically comprised of representatives from DFO Science, Fisheries Management, First Nations, commercial and recreational fishing representatives, NGOs and other interested parties, as required. Examples of project elements to be scoped include key uncertainties to be included and excluded in the OMs, data to be included and excluded, and explicit trade-offs to be considered. These will be discussed in more detail in the following sections.

The final decision on which MP to use to determine catch limits should be made based on a consensus by the Regional Peer Review committee, after review of the scientific content of the advice (including the structure and content of the OMs), and consideration of the relative performance of the MPs with respect to meeting stated objectives and trade-offs among performance metrics. The Regional Peer Review committee will typically be comprised of the technical committee plus a much broader range of interested parties representing DFO Science, Fisheries Management, First Nations, commercial and recreational fishing representatives, NGOs and other interested parties.

The simulation framework tests the performance of specific MPs and ultimately a single catch limit from the final selected MP. The framework does not test posthoc adjustments to the catch limit recommended by an MP. This is in contrast to the decision tables presented in most Pacific Region groundfish stock assessments, where a range of possible catch limits with forecast of future stock status under each catch limit is provided for decision-making. Regardless of the framework used for advice, we note that it remains the purview of the Fisheries Minister to make posthoc adjustments to catch limits, based on cultural, social, or economic considerations, in accordance with the *Fisheries Act* (Sections 2 and 6).

3.2 STEP 2: SELECTION OF OBJECTIVES AND PERFORMANCE METRICS

Here we describe a set of provisional objectives and associated performance metrics, refined after discussions with our technical advisory group. In applications of the framework, objectives and performance metrics should be refined on a stock-by-stock basis, with advice from Fisheries Management, First Nations, commercial and recreational fishing representatives, NGOs and other affected parties. Other objectives and performance metrics could be added (e.g., cultural objectives), decided on a stock-by-stock basis. The time-frame of interest for setting objectives and calculating performance metrics may also be determined on a stock-by-stock basis, as trade-offs change over time (Cox and Kronlund 2008).

Key provisional conservation objectives are guided by the Precautionary Approach Framework (DFO 2006, 2009), elements of which are incorporated into the Fish Stocks provisions of the *Fisheries Act* (see Section 1.2). Additional objectives related to fisheries catch and variability in annual fisheries catch are based on precedents in other DFO Pacific Region analyses (e.g., Cox and Kronlund 2008; Forrest et al. 2018; Cox et al. 2019).

We propose the following provisional tactical conservation and fisheries objectives:

1. Maintain stock status above the LRP in the long term with an agreed upon probability.
2. Maintain stock status above the upper stock reference (USR) in the long term with an agreed upon probability.
3. Maintain a fishing exploitation rate below the rate at maximum sustainable yield with an agreed upon probability.
4. Given the above conservation objectives are achieved, maximize short- and long-term fisheries catch.

5. Given the above conservation objectives are achieved, minimize variability in fisheries catch from year to year.

Objective 1 is implicit in Section 6.1 (1) of the *Fisheries Act* and explicit in Section 6.1(2) (see Section 1.2 of this document). Objectives 2 and 3 are interpretations of Section 6.1 (1) of the *Fisheries Act*, where it is stated that "the Minister shall implement measures to maintain major fish stocks at or above the level necessary to promote the sustainability of the stock." The term sustainability has many definitions as they pertain to fisheries (Hilborn et al. 2015; Marentette and Kronlund 2020). DFO Science is currently evaluating the language of the Fish Stocks provisions to clarify definitions of sustainability with respect to managing Canadian fisheries. Here, we assume that the objective of maintaining the stock in the Healthy Zone (Figure 1) with non-zero probability is consistent with the language of maintaining stocks at or above levels to promote sustainability. We use provisional values of LRP = 0.4 B_{MSY} and USR = 0.8 B_{MSY}, suggested under the Precautionary Approach Framework (DFO 2009) (we herein use B to refer to spawning biomass). See Marentette et al. (2021) for a discussion of the evolving thinking on the role of these reference points in Canadian policy and legislation.

The specific probabilities assigned to successfully achieving each objective likely need consideration on a stock-by-stock basis. For Objective 1, international best practice suggests the probability of maintaining stocks above the LRP should be 90–95% (Sainsbury 2008; McIlgorm 2013; ICES 2018; Marentette and Kronlund 2020), while the probability of reaching a target biomass (e.g., the threshold to the Healthy Zone or some pre-defined target above the USR) can be lower at around 50% (McIlgorm 2013).

We propose the following provisional performance metrics, where MSY refers to maximum sustainable yield, B_{MSY} refers to equilibrium spawning biomass at MSY, and F_{MSY} refers to the fishing mortality that produces MSY over the long term:

1. LT LRP: Probability $B > 0.4\ B_{\mathrm{MSY}}$ (over long-term year range)
2. LT USR: Probability $B > 0.8\ B_{\mathrm{MSY}}$ (over long-term year range)
3. FMSY: Probability $F < F_{\mathrm{MSY}}$ (over entire projection)
4. STC: Probability catch > reference catch (over years 1–10)
5. LTC: Probability catch > reference catch (over long-term year range)
6. AADC: Probability AADC (average absolute interannual difference in catch) < historical AADC (over entire projection)

All of the above performance metrics are expressed as probabilities of being above or below some criteria, where the probability is calculated as the average number of times the criteria is met across replicates and years, e.g.,

$$\mathrm{LT\ LRP} = \frac{\sum_{n=1}^{n_{\mathrm{rep}}} \sum_{t=t_1}^{t_2} B_t > 0.4 B_{\mathrm{MSY}}}{t_2 - t_1 + 1},$$

where t_1 and t_2 are the first and last years over which the metric is calculated, and n_{rep} is the number of replicate simulations. The other performance metrics are calculated in the same way.

In the above list of performance metrics, LT LRP and LT USR are conservation metrics measuring Objectives 1 and 2 over the long term. FMSY is a conservation performance metric measuring Objective 3 over the whole projection period. LTC and STC are economic metrics, representing Objective 4, measured in the short and long-term, respectively. AADC is an economic metric, representing Objective 5, measured over the whole projection period (see description below).

We suggest averaging the long-term performance metrics over a short window (e.g., 5–15 years) before the final year. The projection period and associated long-term year range should be defined on a stock-by-stock basis recognizing that shorter-lived stocks could use a shorter projection period, while longer-lived stocks such as rockfishes (*Sebastes* spp.) may require a longer projection period (DFO 2009, 2013). However, barring some other specified requirement for a definition of long-term (e.g., an associated Committee on the Status of Endangered Wildlife in Canada [COSEWIC] process), we suggest considering “long-term” as the minimum of 1.5–2 generation times of the species (DFO 2009, 2013) or 50 years, whichever is longer. We selected 50 years as a time-frame that should lead to relatively stable behaviour of MPs for shorter-lived stocks such as flatfishes (e.g., Rex Sole case study in Appendix E; Forrest et al. (2018)). Cox and Kronlund (2008) used a projection period of 40 years for Sablefish (*Anopoploma fimbria*) and calculated average long-term performance metrics over years 21–40. Cox et al. (2019) used a projection period of 20 years for Pacific Herring (*Clupea pallasii*) and calculated average long-term performance metrics over the whole projection period. For a longer-lived stock, the first application of this framework for decision-making was an evaluation of rebuilding strategies for Inside Yelloweye Rockfish (*S. ruberrimus*) (Haggarty et al. 2021). These authors evaluated long-term performance metrics over 1.5 generation times (in the 56th year). They also calculated probability of extinction metrics, required for COSEWIC evaluation, after 100 years. Preliminary simulation-testing may be needed to evaluate stability in projections to ensure that a long enough time period has been chosen.

In most studies, conservation metrics are only presented over the long-term, reflecting legislated or policy-driven sustainability objectives. However, in some cases, shorter-term conservation metrics may also be valuable, for example to ensure that long-term rebuilding plans for depleted stocks don’t come at a cost of further depleting the stock in the short term (e.g., Haggarty et al. 2021).

We suggest the “short-term” year range, which in the examples above only applies to the economic performance metric STC, should reflect some time period that is of near-term interest to current participants in the fishery. Our suggestion of 1–10 years represents a starting point and could be modified for specific fisheries.

The catch objectives STC and LTC are provisionally defined in terms of some reference catch that is deemed necessary or desirable for economic reasons to maintain the fishery in question or to maintain the multispecies BC groundfish fishery as a whole. This reference level could be obtained through consultation with stakeholders or by identifying the average or minimum catch in recent years (e.g., minimum catch in the last five years).

The performance metric AADC represents average absolute interannual difference in catch. We chose this representation of variability to be easy to understand and simple to compare historical values. It is calculated by subtracting each year’s catch from the previous year’s catch and taking the mean of the absolute value of these numbers, i.e.,

$$\mathrm{AADC} = \frac{\sum_{t=t_1}^{t_2} |C_t - C_{t-1}|}{t_2 - t_1 + 1},$$

where t_1 and t_2 are the first and last years over which the metric is calculated. In other words, AADC represents the expected deviation in catch from year to year. Alternatives such as standard deviations in log space or coefficients of variation have a more abstract meaning and have been found to be challenging for some stakeholders to interpret (Punt 2017). We suggest reporting AADC relative to the AADC in some historical period that stakeholders identify as having acceptable annual variation in catch. We note that variability in catch is likely to be more important for target species than for incidentally caught species. However, large fluctuations in incidentally caught species may represent a problem in multispecies fisheries, where sudden increases in abundance of low-quota species co-occurring with target species may limit the ability of the fishery to realise TACs of target species.

Performance metrics in the MP Framework are easily customizable. Performance metrics for applications of the framework may, for example, reflect a broader range of objectives such as those associated with rebuilding plans or COSEWIC criteria (e.g., Haggarty et al. 2021).

In cases where performance metrics are calculated over a range of years, care needs to be taken to clearly report how summary statistics are calculated. Provisionally, we suggest calculating performance statistics across replicates and the entire time window as defined for the performance metric. For example, we calculated the F_{MSY} performance metric across all replicates and years simultaneously. Alternative calculations would include calculating the performance statistics for a specific year of interest, calculating the proportion of years in which the performance metric was met (Australian Government 2018), ensuring that the performance metric threshold is met in each and every year (ICES 2016), or each and every replicate. Best practices for these calculations vary across jurisdictions (Marentette and Kronlund 2020) and different methods may imply different risk tolerances. For example, requiring that a threshold is met in each and every year (ICES 2016) is a stricter requirement than requiring a threshold to be met on average over replicates and years. Therefore, the method for calculating the performance metrics may need to be determined iteratively, in consultation with managers, stakeholders and First Nations.

3.3 STEP 3: SELECTION OF UNCERTAINTIES/SPECIFICATION OF OPERATING MODELS

DLMtool OMs are organized into four main components representing a real fished system:

1. population dynamics of the fish stock (e.g., growth, recruitment, mortality);
2. fishery dynamics (e.g., selectivity, spatial targeting);
3. observation processes (e.g., bias and precision in survey indices); and
4. management implementation (e.g., percentage overages of catch limits).

Parameters under the four components are entered into “slots” [terminology referring to a feature of the “S4” object-oriented programming system in R; R Core Team (2019)], described in detail in our Appendix A and in Appendix B of Carruthers and Hordyk (2018a).

DLMtool allows the incorporation of uncertainty in most OM parameters through optional specification of a distribution. See Appendix B of Carruthers and Hordyk (2018a) for a full list of parameters for which a distribution can be specified. To isolate the effects of specific sources of uncertainty on performance of MPs, we recommend development of alternative OMs that change the value (or distribution) of one or more parameters and/or data sources of interest (Punt et al. 2016). In general, we recommend developing more than one OM, dividing OMs into a reference set of core OMs representing the most important plausible model uncertainties, and a robustness set for testing sensitivity to a broader range of structural uncertainties (Rademeyer et al. 2007). We do not attempt to weight OMs in this framework. For BC groundfish stocks, reference-set uncertainties will likely be based on key uncertainties identified in typical groundfish stock assessments, namely natural mortality (*M*), steepness of the stock-recruit relationship (*h*), and initial depletion (i.e., depletion from an unfished state at the beginning of the projection period). For stocks with available stock assessments, these should be consulted to identify major sources of uncertainty.

Candidate uncertainties to include in OMs in the robustness set may include:

- Changes in predation rate (e.g., seal predation).
- Changes in availability of prey.
- The effectiveness of or changes to closed areas such as Rockfish Conservation Areas (RCAs).
- Alternative representations of survey and commercial fleet size selectivity.
- Alternative catch histories (e.g., for species such as rockfishes, which were historically reported under generic species names).
- Implementation error (actual catches are above or below the TAC).

In BC commercial groundfish fisheries, 100% of trawl vessels have carried observers since 1996 (Turris 2000) and 100% of line vessels have carried electronic monitoring (EM) systems since 2006 (Stanley et al. 2015). Therefore, recent catch data are considered to be reliable and can be expected to remain so in the future, barring any major changes or interruptions to the observer programs. For some species that are important recreationally or to First Nations, there may be other uncertainties associated with catch data. We suggest that implementation error scenarios concerning catch reporting should be developed in collaboration with fishery managers, including First Nations managers, and the commercial or recreational industry, as appropriate to the stock.

In some cases, the above uncertainties may be included in the reference set. In general an iterative approach may be required to reduce the robustness down to a set of OMs that provide contrast to results obtained with reference-set OMs (Rademeyer et al. 2007). We discuss recommendations for treatment and presentation of results from the reference and robustness sets in Section 3.6. To ensure transparency and reproducibility, we recommend that the full specification of OM parameters be clearly documented in appendices attached to the working paper and the code to accomplish the simulations be version controlled and archived on publication.

We note that DLMtool OMs include a large number of parameters that can vary through time, or which can be set to be deliberately biased. To simplify the OM and focus on what are likely the most important axes of uncertainty, we suggest fixing most parameters to be time invariant and unbiased (Appendix C). Exceptions would occur when one or more of these parameters represent axes of uncertainty for specific stocks, or when certain time-varying parameters are already accepted components of the stock assessment.

Best practice recommends conditioning OMs with observed data so they can reproduce historical observations (e.g., indices of abundance, age-composition data). DLMtool's companion software package, MSEtool (Huynh et al. 2019), includes an efficient implementation of a stock reduction analysis (SRA) (Kimura and Tagart 1982; Walters et al. 2006) to help with this process (Appendix B). An SRA is effectively a catch-at-age model that estimates the combinations of unfished recruitment, depletion, fishing effort, recruitment deviations, and selectivity that would be consistent with the observed data given assumptions about other parameters (e.g., growth, natural mortality). The SRA is run n times, with the number of replicates n matching that for the closed-loop simulations. The SRA draws from parameter ranges specified in the OM (e.g., Appendix F), estimates key parameters and updates those parameters in the OM (see Appendix A for details). SRA replicates that do not converge (with convergence defined as a positive-definite covariance matrix) are discarded. The SRA implementation in MSEtool can be conditioned on catch or effort time series (Appendix B). For most applications for BC groundfish stocks, we suggest conditioning on catch, as historical trajectories of catch tend to be more reliable than time series of effort, especially given uncertainties in how to best represent and interpret effort in multispecies fisheries. Further details on the SRA OM-conditioning model are provided in Appendix B.

For some data-limited stocks, indices of abundance may be considered less reliable due to sampling difficulties or rarity. In these cases, as long as there is a time series of catch data, we still recommend using the SRA model for conditioning, recognizing that a much broader range of uncertainties will need to be considered in the set of OMs, including large uncertainty in stock size, productivity and current depletion level. Where there is no index of abundance, we recommend developing a wide range of OMs conditioned on available catch data, which differ in terms of major uncertainties, especially related to stock productivity and current depletion level.

We recognize that, for some data-limited stocks, there may be some efficiency gains in developing more generic OMs that capture the major biological, fleet, and observation characteristics of a set of similar stocks. However, we recommend focusing first on species for which customized OMs can be developed, which are conditioned on observed data.

3.4 STEP 4: IDENTIFICATION OF CANDIDATE MANAGEMENT PROCEDURES

We screened all MPs available in DLMtool as of November 2019 along with recent MPs used in BC groundfish reports to consider their appropriateness for the framework. This represents a fairly comprehensive set of data-limited MPs available in the primary literature or agency reports to date. Here, we describe the types of MPs available and the process by which we identified a provisional set of MPs, then we explain how some MPs were tailored to the BC groundfish needs. We describe provisional candidate MPs in detail in Appendix D.

DLMtool includes MPs that make many different kinds of management recommendations. These recommendations include adjustments to total allowable catch (TAC), effort, or spatial allocation of catch or effort. For the framework, we focus on MPs that make TAC recommendations because BC groundfish are managed in general by quotas. Therefore, all MPs considered here take some subset of the data generated by an OM and provide a recommended catch for the subsequent year.

We focus on two main types of MPs: empirical and model-based MPs. In choosing from the available empirical and model-based MPs, we excluded MPs based on a number of requirements that would rarely be met for our stocks. We excluded MPs that required knowledge of absolute abundance since there are unlikely to be cases where we have such knowledge in a data-limited case. We excluded MPs that required recent age composition data because we intend this framework to be applied to stocks for which recent age-composition data are not available. We excluded MPs that required knowledge of depletion and steepness of the stock-recruit relationship since these are likely to be major axes of uncertainty for the stocks to which this framework will be applied. While it is necessary to explore these axes of uncertainty within the OM, implementing an MP on real data when that MP requires knowledge of depletion and steepness would require additional assumptions. In a few cases, we excluded MPs that we found difficult to communicate (e.g., MPs based on the value of the index relative to the time series mean and standard error, Jardim et al. 2015). Such MPs did not perform appreciably differently from included MPs and we felt their exclusion would not lead to a loss in overall performance of the framework.

A library of provisional MPs included in this framework is described in Appendix D.

3.4.1 Empirical MPs

Commercial catch data are available for all BC groundfish stocks with relative certainty since 1996 for BC trawl fisheries and 2008 for BC hook-and-line fisheries. Fisheries-independent trawl and longline surveys have been conducted systematically since the early 2000s for BC groundfish and the population indices derived from these data likely represent some of the most informative data for many data-limited groundfish stocks in BC. Fish lengths are collected on both surveys and commercial fishing trips for many species. However, length-based MPs often require strong assumptions and require that the simulated length-composition data are sufficiently "messy" to reflect the real-world length-composition data, which often have large and inconsistent variances among years and length bins. Simulating realistic length-composition data is challenging and we have not sufficiently investigated best practices for simulating length-compositions within the DLMtool software. Furthermore, length-based empirical MPs may be less responsive to changes in indices of abundance and perform more poorly than other data-limited MPs (Sagarese et al. 2018). Reliable and abundant age-composition data are generally not available for the data-limited species for which this framework is designed. For the above reasons, we propose MPs that make use of only catch and population index data as provisional candidate MPs.

We can divide MPs that make use of catch and/or population index data into four categories: constant catch, index ratio, index slope, and index target MPs:

1. Constant-catch MPs set the recommended catch to some fixed level, typically based on recent or historical catches. Importantly, constant-catch MPs do not incorporate feedback between the management system and the population—they make the same catch recommendation regardless of trends in the population index. Nonetheless, they represent simple MPs that in many cases represent the status quo or slight modifications of the status quo.
2. Index-ratio MPs base their catch recommendation on a ratio of a population index in one time period compared to another time period—generally a recent period (e.g., last year) compared to a short period before that (e.g. 2–3 years ago).
3. Index-slope MPs fit a regression of population index data compared to time and make a catch recommendation based on the slope of the regression. They are closely related to index-ratio MPs.
4. Index-target MPs compare recent population index values to the value of the index at a fixed, agreed-upon historical time period to make a catch recommendation that aims to maintain the population index at the fixed historical value. In this regard, index-target MPs differ subtly but importantly from the index-ratio and index-slope MPs, which compare recent index values to a moving window of index values as time progresses.

We tailored many of the available empirical MPs to suit BC groundfish stocks (Appendix D). For example, most of the available survey data for BC groundfish are collected biennially in any one spatial region. Therefore, we modified many of the index-based MPs to reflect this reality. Typically, this involved adding variants of MPs that considered longer time windows when calculating averages or slopes to account for the fact that there was only half the available data compared to an annual survey. In other cases, we added alternative versions of MPs that encompassed a wider range of control variables. For example, the Islope MPs (Geromont and Butterworth 2015a) as originally described and implemented in DLMtool implicitly set the catch recommendation in the first year to 60–80% of the mean catch from the recent five years (assuming a neutral survey index). Since we do not a priori expect BC groundfish stocks to be overfished, we adjusted the relevant control parameters in our provisional MPs to explore a wider range of initial catch recommendations.

3.4.2 Model-based MPs

In addition to the empirical MPs, we suggest a surplus production (SP) model be considered amongst candidate MPs. We provisionally include the SP model coded in MSEtool (Huynh et al. 2019) and based on Fletcher (1978) (Appendix D Section D.4.1) paired with a number of possible harvest control rules (HCRs). We suggest consideration of both Schaefer (1954) and Fox (1970) production models since it is not clear, until simulation-tested, which will generate better performance statistics for a given stock. We suggest a weakly informative prior probability distribution be set on the intrinsic rate of population increase r, possibly following recommendations such as those in McAllister et al. (2001). Alternative prior probability distributions could be considered tuning parameters in alternative MPs. The SP model estimates must be paired with an HCR to form a complete MP. We provisionally suggest a number of HCRs in Appendix D Section D.4.

3.4.3 Dealing with multiple survey indices within MPs

The vast majority of published data-limited MPs are based on single indices of abundance. This presents a challenge for BC groundfish stocks, since the trawl and hook-and-line fisheries-independent surveys that cover our coast do so on a biennial basis, alternating amongst areas. We suggest the following three possible solutions:

1. Build and test OMs for areas associated with a single index. If these areas are considered simultaneously in a single application of the MP Framework then we suggest comparing performance of the MPs across all areas and, if possible, choosing an MP that performs reasonably well across all areas. This could be accomplished, for example, via a minimax-style solution, where an MP is selected that performs the least poorly across all areas. If MP performance across areas differs substantially, then different MPs may be needed for different areas. A potential problem with this approach is when stocks are larger than the surveyed area and the information captured in a single index does not represent the entire stock.
2. Develop a single index by "stitching" multiple survey indices together, likely with the application of geostatistical spatiotemporal modeling. This is an active area of research (e.g., Shelton et al. 2014; Thorson et al. 2015; Anderson et al. 2019; Anderson and Ward 2019) and will likely become more common within the fisheries literature and in stock assessments. While MPs could be developed that average or in some other way combine multiple survey indices, geostatistical modelling is likely to offer a more coherent way of combining survey data from multiple survey protocols or spatial areas.
3. Develop MPs that incorporate multiple survey indices (e.g., Cox et al. 2019). Many existing data-limited MPs, such as those described in this document (Appendix D), could be modified to incorporate information from multiple indices. For example, separate indices could be

fed through the MP algorithms independently and the most biologically precautionary TAC recommendation could be made.

3.4.4 Reference MPs

In addition to the candidate MPs, it is important to include reference MPs. Provisionally, we suggest the following reference MPs:

- No fishing
- Fishing at F_{MSY}
- Fishing at 0.75 F_{MSY}
- Maintaining the current TAC

The purpose of reference MPs is not to explore viable management strategies but to bound the range of expected or possible performance and contextualize whether differences between performance statistics among MPs are meaningful (Punt et al. 2016). For example, the "no fishing" reference MP provides information on maximum possible stock levels and the maximum possible rate of rebuilding under a rebuilding scenario. The "fishing at 0.75 F_{MSY}" MP illustrates performance under an omniscient manager with perfect information. The MP that maintains the current TAC is included because it illustrates what is likely the near-term default had the framework not been implemented and illustrates the long-term performance expectations given current exploitation levels.

3.4.5 Including new MPs

The candidate MPs proposed here are a provisional library from which to build. Data-limited MP development is a rich area of research, likely still in its infancy, and beyond minor adjustments to existing MPs, MP development is not the focus of this document. More MPs may be developed as part of the application of this framework and will certainly be developed elsewhere in the literature. MPs used when applying this framework may also be "tuned" to perform well for specific stocks by adjusting algorithm parameters or priors. The framework presented here has been designed to accommodate new MPs and we expect the library of candidate MPs to grow over time, with the framework providing a means to rigorously test new MPs through the closed-loop simulation process.

3.5 STEP 5: SIMULATION OF THE APPLICATION OF THE MANAGEMENT PROCEDURES

Once the objectives, performance metrics, conditioned OMs and MPs are fully specified, a closed-loop simulation framework (Figures 2 and 3) can be applied to test relative performance of the MPs with respect to meeting the stated objectives.

We recommend beginning with a satisficing step, where trial simulations are run to screen out MPs that do not meet a basic set of performance criteria (Miller and Shelton 2010). For example, in our illustrative Rex Sole case study (Appendix E), we screened out MPs that did not achieve a long-term 90% probability of keeping the stock above the LRP (LT LRP $>$ 0.9) and a short-term 80% probability of catch being above recent average catch (STC $>$ 0.8). We recommend an iterative approach to selecting the satisficing criteria on a stock-by-stock basis. The purpose is to simplify the decision-making process by focusing on a manageable number of MPs. If possible, criteria should not be so strict as to filter out MPs with broadly acceptable performance or to filter out almost all MPs. Similarly, criteria should be strict enough to filter out consistently under-performing MPs. We recommend selection of satisficing criteria be done iteratively with the technical committee.

DLMtool is designed to follow standard MSE operating procedure (Figure 3). For each MP, the OM is used to simulate the various data streams required by the MP at each time step, then the population biomass is projected forward under the prescribed MP at each subsequent time step until the projection period is complete. Performance is then evaluated through calculation of performance statistics in the OM. DLMtool makes use of the C++ programming language and parallel processing, making the simulations computationally efficient (Carruthers and Hordyk 2018a).

For each OM-MP combination, multiple replicate projections are run to account for observation and process errors in the data streams. This is achieved by adding stochastic noise to the data (e.g., indices of abundance) before passing them to the MP. Coefficients of variation in the data should be consistent with those in historical observations (Rademeyer et al. 2007). We suggest selecting a sufficient number of replicates so that the rank order of MPs across the performance metrics remains consistent regardless of additional replicates (Carruthers and Hordyk 2019). This can be checked by plotting the cumulative performance statistics against the number of replicates, to check that the rank order of MPs with respect to performance statistics will not change with the addition of more replicates (e.g., Figure 4). In our experience with BC groundfish, the number of required replicates is likely to be at least 100.

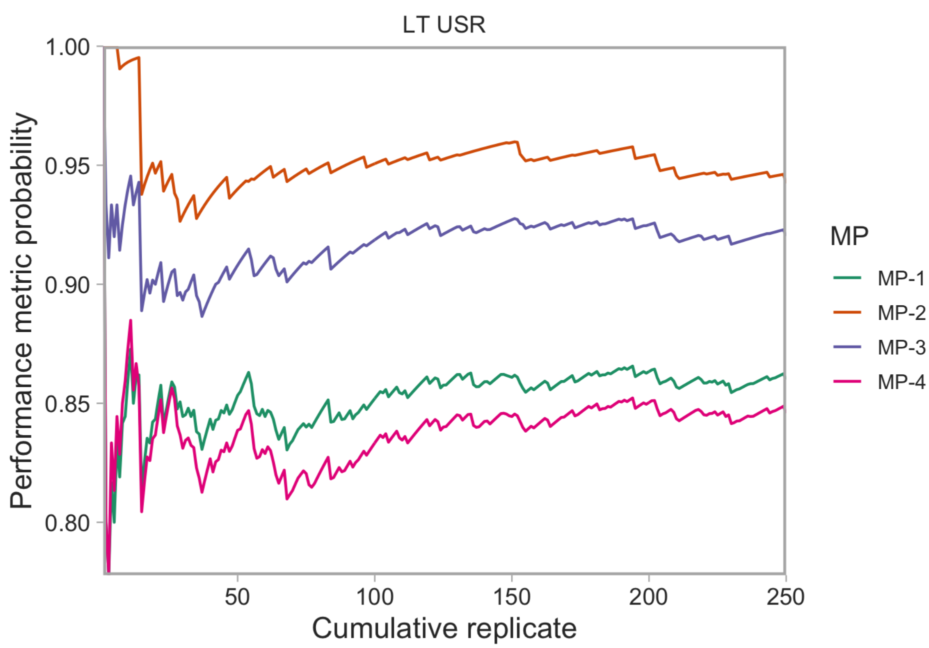

Figure 4. This visualization illustrates the degree to which the closed-loop simulation replicates have reached convergence in the rank order of MP performance. The y-axis represents the proportion of cumulative replicates that have achieved the performance metric. The cumulative replicates are shown along the x-axis. Each panel represents a different performance metric and the colours illustrate the various MPs. A wiggly line indicates that a given MP is still varying in its performance metric value as more replicates are added. Lines that are crossing indicate changes to the rank order of MP performance as more replicates are added. A plot is consistent with convergence when the lines are remaining roughly horizontal and parallel and are not crossing each other towards the right of the panel.

3.6 STEP 6: PRESENTATION OF RESULTS AND SELECTION OF MANAGEMENT PROCEDURE

3.6.1 Reference and robustness sets

In Step 3 (Section 3.3), we recommended dividing OM scenarios into a reference set, which includes a range of plausible alternative OM scenarios that impact results, and a robustness set, which includes a wider range of alternative OM scenarios, which may be less supported by available data, but still potentially impact results (Rademeyer et al. 2007). These authors recommended reducing the set of OM robustness scenarios down to those which yield the most contrasting results from the reference set. Therefore, before the final MP is selected, the role of the robustness set is to provide a check that the final MP performs well across a more diverse set of OM scenarios. Poor performance of an MP under one of these OM scenarios may influence a decision-maker to select another MP that performs well under both OM reference and robustness scenarios.

We recommend presenting performance metrics from the reference and robustness sets separately. For most visualizations, we recommend that the reference set performance metrics are averaged across all OM reference-set scenarios. An exception is the table of performance metrics (e.g., Figure 5), which is presented in two ways: (i) minimum value of the performance metric across all OM reference-set scenarios; and (ii) average value of the performance metric across all OM reference-set scenarios. The first is a “worst-case scenario” approach, whereas the second case integrates across the whole reference set. We recommend presenting performance metrics from the individual OM robustness-set scenarios separately. This facilitates visualization of performance of MPs that perform well in the reference set under specific more diverse assumptions (Rademeyer et al. 2007).

3.6.2 Visual presentation of performance metrics

Here we focus on developing a set of provisional visualizations that facilitate comparison of performance metrics across MPs and evaluation of trade-offs amongst them. We have developed an R package ggmse (Anderson et al. 2020b) (Appendix G) for generating these visualizations. We expect that some or all of the visualizations will be refined over time as users gain familiarity and specific needs arise.

First, we suggest a graphical representation of a probability table (Figure 5) to visualize performance metric results. This visualization lends itself to a large number of MPs and so works well for displaying results for *all* MPs—not necessarily just satisficed MPs. By shading the cells according to their underlying performance metric value, the visualization draws the eye to similarities and differences across MPs. We suggest sorting the rows by performance metrics of particular interest, shading the MP name differently for reference MPs, and highlighting any satisficing criteria by outlining cells that pass satisficing thresholds. The default colour shading is colour-blind proof, prints accurately in grey scale, and is visually linear in its gradient (Garnier 2018).

Second, we suggest a visualization that lends itself to summarizing performance for a small set of satisficed MPs across OM reference-set scenarios (Figure 6). This figure highlights average performance across OM scenarios (dots) along with the range of performance values seen across OM scenarios (thin lines). This figure also illustrates the range of performance after dropping the OM scenarios with the highest and lowest of each performance metric values (thicker lines).

We suggest two possible visualizations for highlighting performance-metric trade-offs (Figures 7 and 8). The first is a bivariate dot plot, which lends itself to comparing any two performance metrics. The second is a "radar" plot (also known as a "kite" or "spider" plot). Radar plots lend themselves to comparing across many performance metrics, but become difficult to interpret as the number of performance metrics grows (e.g., above six) and one needs to be careful when interpreting these plots because the arrangement of the "spokes" can affect perception of the resulting shapes and humans have been shown to be slower and less accurate at interpreting radial displays of data compared to Cartesian displays (Diehl et al. 2010; Feldman 2013; Albo et al. 2016).

To understand the processes leading to the performance metrics, we recommend that the results from the application of this framework should include visualizations of historical and projected B/B_{MSY}, F/F_{MSY}, and catch. We suggest two versions: one that lends itself to careful inspection of individual OM scenarios and an understanding of individual replicate behaviour (Figure 9) and another that lends itself to comparing the time series across OM scenarios (Figure 10). Inspection of these time series can provide a diagnostic check that the parameter space is being adequately sampled among replicates. For example, a bifurcated time series could indicate that a boundary condition is being reached in some replicates and would indicate that one or more parameters may need to be sampled from a broader distribution. Inspection of these time series can also improve understanding of the performance of MPs, and may lead to addition of new performance metrics if the existing set fails to capture some important behaviour that becomes apparent in the time series. For example, if some MPs are creating variable TAC recommendations in the initial implementation years this may suggest the need to specify a short-term TAC variability objective and performance metric. Alternatively, if some MPs are resulting in short-term declines in B/B_{MSY} that eventually recover by the long-term window, this may suggest the need for a short-term conservation objective and performance metric or the calculation of performance metric probabilities via the minimum probability in any year (e.g., ICES 2016). These plots may also indicate MPs that are likely to cause extinction of the stock, even if this behaviour is masked by long term average performance metrics. For example, like many OMs, the DLMtool OM prevents negative biomass in the event that TACs are greater than the stock biomass. Therefore MPs which do not consider feedback from the index of abundance, such as constant catch MPs, may sometimes generate TACs greater than the available biomass, which would be buffered against by the OM. Inspection of Figure 9 would indicate this behaviour.

Visualizing the simulated population-index time series is particularly important when using MPs that rely on the population index (e.g., Figure 11). First, for simple MPs that closely resemble recent management, future simulations of the population index can be used as a type of posterior predictive check (Gelman et al. 2014) to ask whether the future simulations look like they are plausible given historical observations of the same population index. Second, the range of future projected index values can be used as a trigger for re-evaluating an OM. For example, if a future observed survey index deviates from some quantile of simulated survey index values, the observed state of nature can be considered unlikely given the assumed OMs (sometimes called "exceptional circumstances" (Butterworth 2008)). We note that the cumulative probability of the observed index exceeding some quantile of simulated values increases with time, so such a rule will generate more than the tail probability of false re-evaluation triggers. For example, even if the OM perfectly captured reality, there is approximately a 0.4 probability (2 in 5 chance) that the observed index will at some point exceed the 95% quantile of simulated index values after 10 years. Closed-loop simulation itself can be used to establish appropriate triggers with acceptable statistical power (Carruthers and Hordyk 2018b).

Finally, we suggest two visualizations that illustrate the trade-off between F/F_{MSY} and B/B_{MSY} across replicates for the various MPs (Figure 12 and 13). One is a standard Kobe plot showing F/F_{MSY} vs. B/B_{MSY} for the final year of the projection (Figure 12). This visualization highlights the parameter space with the highest probability density via quantile kernel-density-estimate contour lines. The other visualization shows the trajectory of F/F_{MSY} vs. B/B_{MSY} through time (Figure 13). The final year of Figure 13 is an alternative representation of Figure 12. Figure 13 also presents another diagnostic check of the behaviour of MPs over time, for example by indicating how often the stock falls within the critical zone. While some MPs may meet long-term conservation metrics, they may cause conservation problems in the shorter term, possibly indicating the need to consider aditional short-term conservation metrics.

	LT LRP	LT USR	FMSY	STC	LTC	AADC
MP-1	>0.99	>0.99	>0.99	0.05	0.34	0.58
MP-5	>0.99	>0.99	>0.99	<0.01	0.12	0.83
MP-3	>0.99	>0.99	>0.99	<0.01	0.04	0.90
MP-4	>0.99	>0.99	>0.99	<0.01	0.01	0.98
MP-2	>0.99	>0.99	>0.99	<0.01	<0.01	0.99
MP-ref	>0.99	>0.99	>0.99	<0.01	<0.01	>0.99
MP-13	>0.99	>0.99	>0.99	<0.01	<0.01	>0.99
MP-12	>0.99	>0.99	>0.99	<0.01	<0.01	>0.99
MP-11	>0.99	>0.99	>0.99	<0.01	<0.01	>0.99
MP-10	>0.99	>0.99	>0.99	<0.01	<0.01	>0.99
MP-9	>0.99	>0.99	>0.99	<0.01	<0.01	>0.99
MP-8	>0.99	>0.99	>0.99	<0.01	<0.01	>0.99
MP-7	>0.99	>0.99	>0.99	<0.01	<0.01	>0.99
MP-6	>0.99	>0.99	>0.99	<0.01	<0.01	>0.99
MP-17	>0.99	0.98	>0.99	0.09	0.45	0.86
MP-18	>0.99	0.96	0.99	0.94	0.68	0.60
MP-15	0.99	0.94	0.98	>0.99	0.82	0.85
MP-14	0.99	0.92	0.97	0.97	0.71	0.50
MP-19	0.98	0.86	0.91	0.96	0.71	0.38
MP-16	0.98	0.85	0.90	0.97	0.74	0.41

Figure 5. This probability table illustrates performance metric values across a number of MPs. See Section 3.2 for definitions of the various performance metrics (columns). The MPs are ordered by decreasing performance metric values from top to bottom starting with the left-most performance metric and using columns from left to right to break any ties. The colour shading reflects the underlying numbers and is included to make the differences in the values more readily apparent. Outlined cells represent MPs that met a given satisficing criterion. MP names shaded grey represent reference MPs.

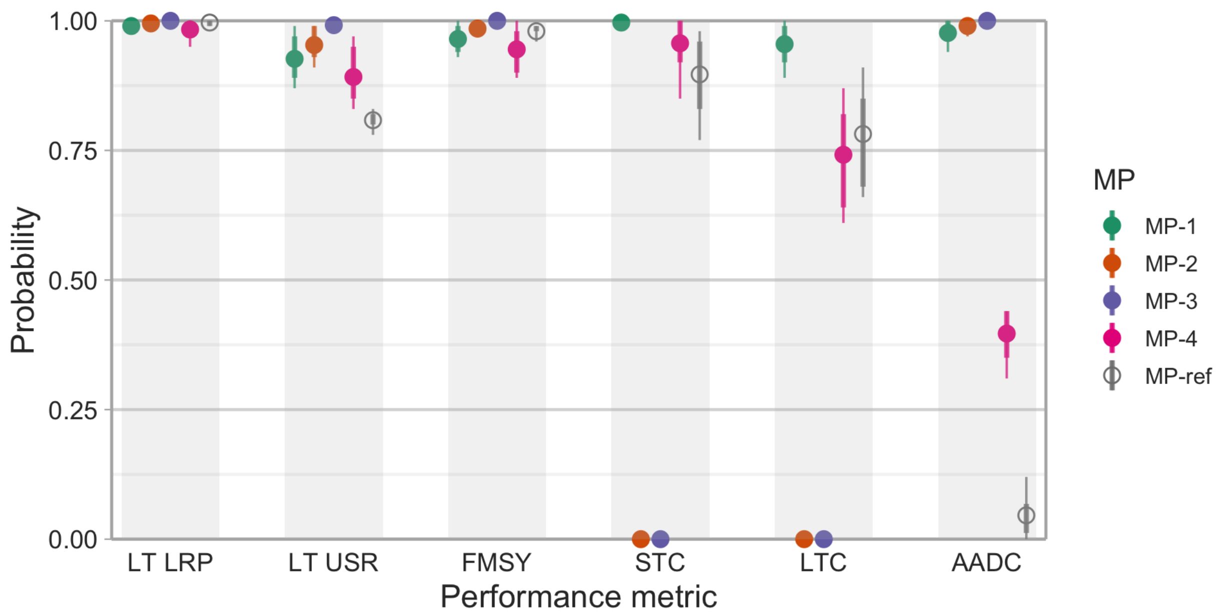

Figure 6. This visualization summarizes performance of a small number of MPs (e.g., satisficed MPs) across OM scenarios. Dots represent mean performance across OM scenarios. Thin lines represent the range of performance across OM scenarios. Thicker lines represent the range of performance across OM scenarios after dropping the highest and lowest OM scenario within each performance metric. This visualization can also be used without the line segments to represent performance for individual OM scenarios (e.g., OM robustness scenarios). Reference MPs are indicated by open circles (True). Non-reference MPs are indicated by closed circles (False).

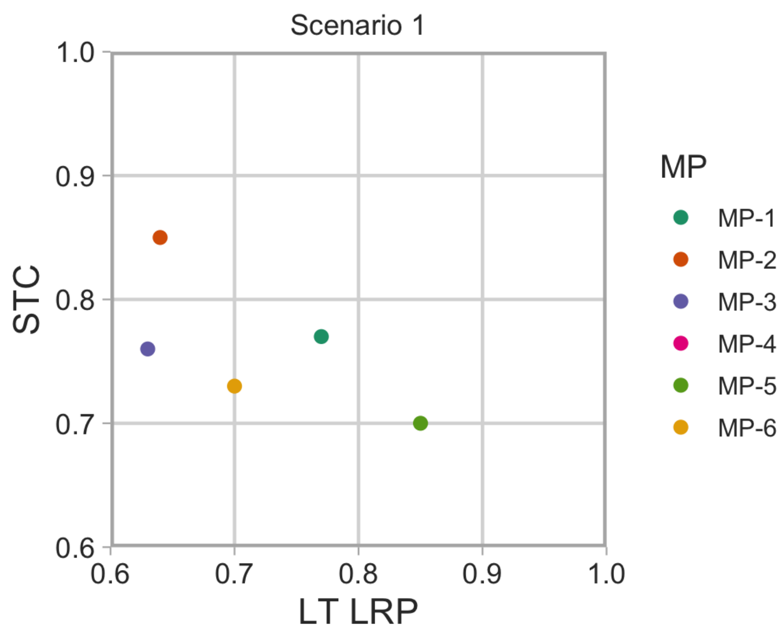

Figure 7. This visualization demonstrates bivariate trade-offs between two performance metrics for a given OM scenario or average across OM scenarios. The bottom and left axis correspond to two performance metrics and individual dots represent MPs.

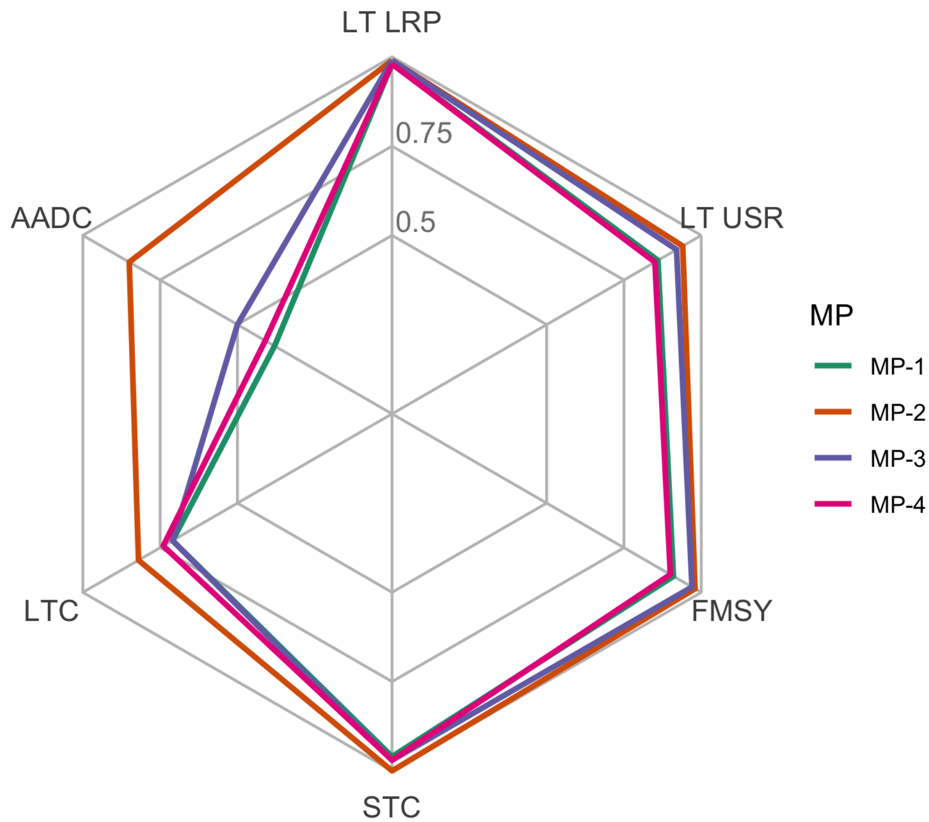

Figure 8. This radar plot illustrates performance metric trade-offs amongst a set of MPs. Each spoke represents a performance metric (Section 3.2). Each line represents an MP. The position of each line on each spoke represents the probability of that MP achieving that performance metric. Therefore, lines closer to the outside correspond to MPs that are determined to have a higher probability of achieving that performance metric. Care should be taken when interpreting these plots because the arrangement of the "spokes" can affect perception of the resulting shapes (Diehl et al. 2010; Feldman 2013; Albo et al. 2016).

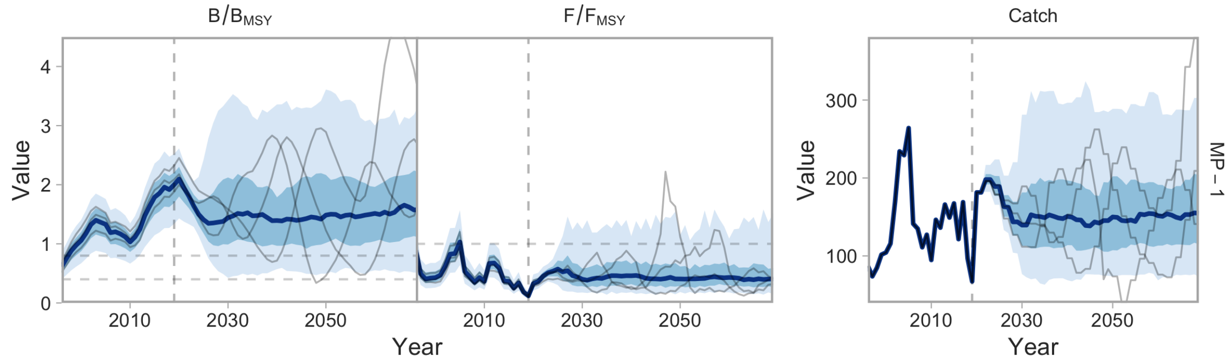

Figure 9. This visualization illustrates historical and projected B/B_{MSY} (where B indicates spawning biomass), F/F_{MSY}, and catch for various MPs for a single OM. Only one MP is shown here. In applications of the framework, all satisficed MPs would be shown as separate panels. Dark lines indicate the median value and the darker and lighter shaded ribbons indicate the 50% and 90% quantiles. Thin gray lines represent illustrative simulation replicates. The vertical dashed line indicates the last year of the historical period. The horizontal dashed lines indicate B/B_{MSY} = 0.8 and 0.4 and F/F_{MSY} = 1. Note that the simulations are mean-unbiased and so the median B/B_{MSY} and F/F_{MSY} are not expected to lie perfectly on the 1 line even if fishing perfectly at F_{MSY}.

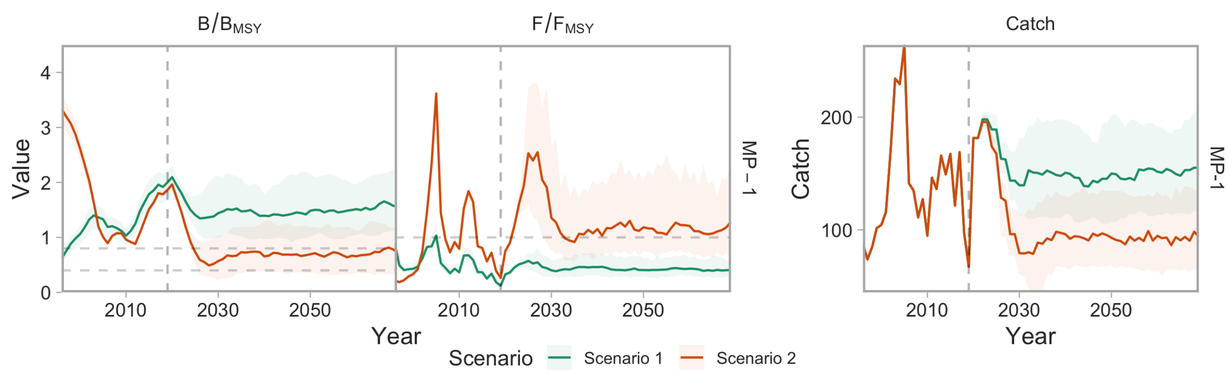

Figure 10. This visualization highlights sensitivity of the historical conditioned and projected time series across OM scenarios for two hypothetical OM scenarios. In applications of the framework with more than two OM scenarios, more lines would be shown. Only one MP is shown here. In applications of the framework, all satisficed MPs would be shown as separate panels. The solid lines correspond to median values and the shaded ribbons correspond to 50% quantiles to indicate variability across replicates. The vertical dashed line indicates the last year of the historical period. Whereas Figure 9 helps compare across MPs within a given OM scenario, and highlights example replicate performance, this visualization helps compare time series within a given MP across OM scenarios. This visualization lends itself to multiple rows where each row represents a different MP. Note that sensitivity of the time series to OM scenarios does not necessarily correspond to sensitivity in terms of rank order of MPs, which is ultimately the important result in a management-oriented approach.

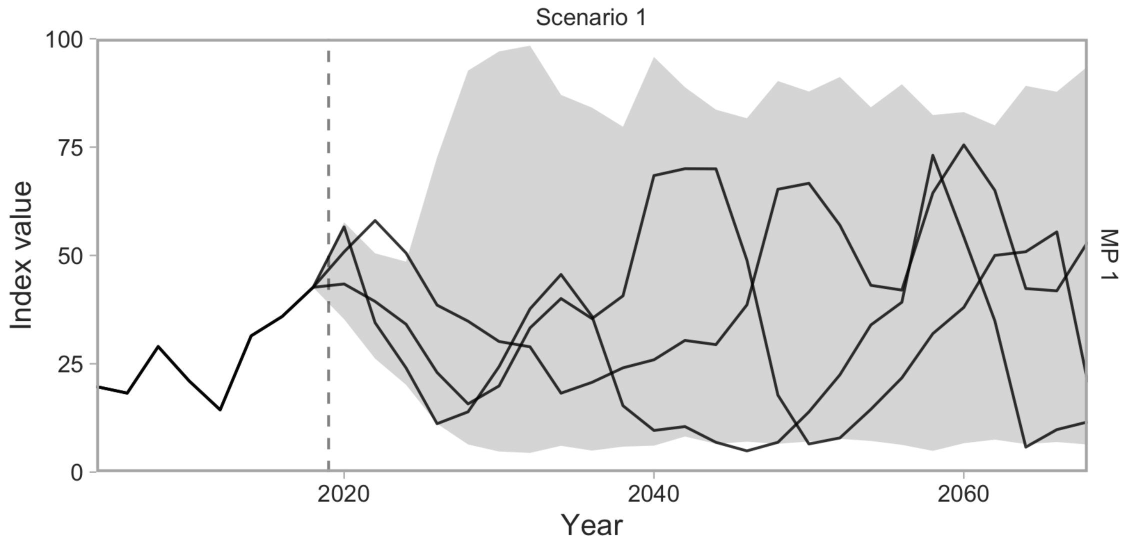

Figure 11. This visualization illustrates the historical and projected relative population index values. The black lines represent example replicates. The shaded grey region represents the 95% quantile of the simulated index values. The vertical dashed line represents the last historical year of the observed data (final year of the catch-time series, see Appendix A). In the historical period, the index represents the observed index (usually a survey index). In cases such as the one illustrated here, the survey is performed every second year. For index-based MPs, understanding the expected envelope of possible future survey trends is important for evaluating whether the system continues to be consistent with the OM assumptions in the future. For simple MPs, this visualization is also helpful for assessing whether the simulation is generating future index projections that are consistent with the observed historical trends.

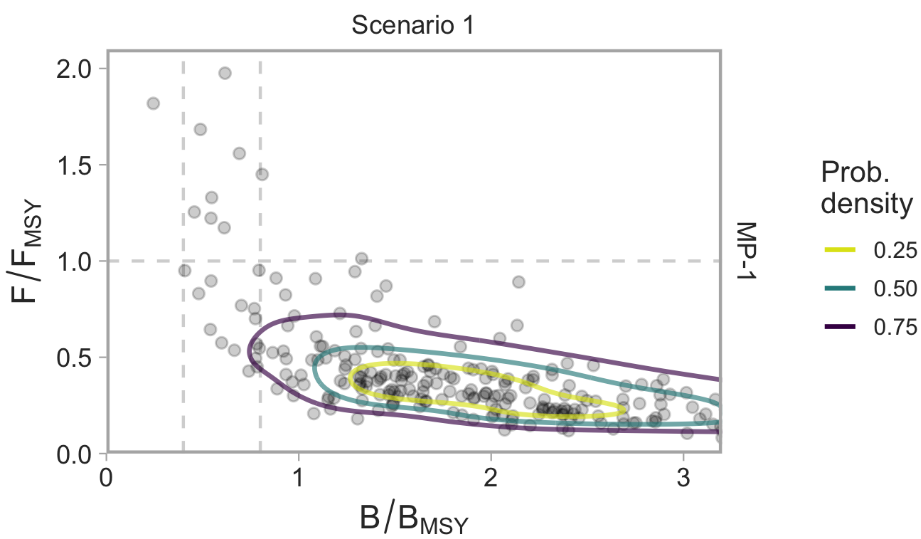

Figure 12. Kobe (B/B_{MSY} vs. F/F_{MSY}) plots for the final year of the projections across all replicates. Dots represent individual replicates. The contour lines indicate two-dimensional kernel-density-smoothed quantiles, calculated in log space. For example, the 0.50 contour lines encompass approximately 50% of the replicates. The vertical dashed lines show B/B_{MSY} = 0.4 (left) and 0.8 (right). The horizontal dashed line shows F/F_{MSY} = 1. Replicates with values beyond the outer axis limits are shown on the axis limit with an open circle (e.g., bottom right corner).

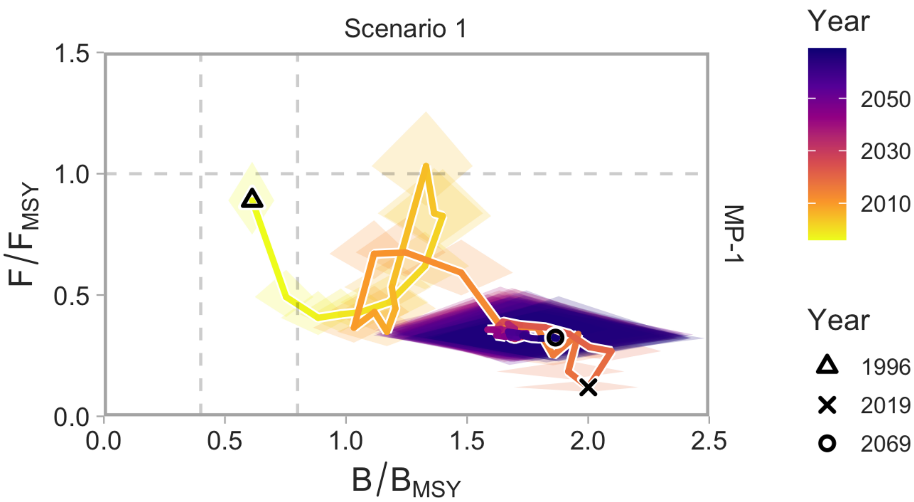

Figure 13. This visualization illustrates the time trajectory of F/F_{MSY} vs. B/B_{MSY} values summarized across replicates for a single MP. The solid line corresponds to the median value. The vertical dashed lines show B/B_{MSY} = 0.4 (left) and 0.8 (right). The horizontal dashed line shows F/F_{MSY} = 1. Each diamond represents the 50% quantile of B/B_{MSY} (horizontal) and F/F_{MSY} (vertical). There is one diamond per year of the historical and projection period. The lines and diamonds change colour over time and specific points in time are illustrated with symbols (first year, last historical year, and last projected year). For reference, the data underlying the last time slice in this figure are the same as Figure 12.

4 REX SOLE CASE STUDY SUMMARY

We undertook a case study of the Rex Sole (*Glyptocephalus zachirus*) stock in the West Coast Vancouver Island (WCVI) groundfish management area (Area 3CD) to demonstrate an application of the MP Framework. This stock was selected because it lacks recent ageing data, lacks current assessment advice, and has a contrasting life-history with that of Strait of Georgia Yelloweye Rockfish (*Sebastes ruberrimus*), a fish stock for which this framework has been applied concurrently (Haggarty et al. 2021). We briefly summarize the Rex Sole case study here. Full details are available in Appendix E and Appendix F. The case-study is presented to illustrate the framework and its outputs but *is not intended for the provision of catch advice at this time*.

Rex Sole are a flatfish caught primarily in the bottom-trawl fishery in BC. They can live as long as 29 years, grow as large as 58 cm, and weigh up to 1.44 kg. Otoliths have been collected regularly on BC surveys, but these otoliths have not been aged. Rex Sole are caught regularly in the DFO synoptic bottom trawl surveys and relative biomass indexes calculated from these surveys have a relatively low level of observation error (CV [coefficient of variation] = 0.07–0.12) compared to many other groundfish stocks caught in the same survey. Rex Sole in WCVI do not have an existing stock assessment or TAC.

Decision context: The decision to be made was which MP to use to determine catch limits for the period until the next available catch advice.

Objectives and performance metrics: We defined objectives and performance metrics following the provisional suggestions in Section 3.2 that are guided by the Precautionary Approach Framework. Since 1.5–2 Rex Sole generations would be less than 50 years, we chose to run our projections for 50 years (Section 3.2).

We defined the performance metrics as:

1. LT LRP: Probability $B > 0.4\ B_{\mathrm{MSY}}$ (years 35–50)
2. LT USR: Probability $B > 0.8\ B_{\mathrm{MSY}}$ (years 35–50)
3. FMSY: $\mathrm{P}(F < F_{\mathrm{MSY}})$ (years 1–50)
4. STC: Probability catch > reference catch (years 1–10)
5. LTC: Probability catch > reference catch (years 35–50)
6. AADC: Probability AADC (average absolute interannual difference in catch) < historical AADC (years 1–50)

where reference catch was defined as the average catch from the last five years.

Operating model specification: We established six reference-set OMs encompassing uncertainty in depletion of the stock prior to 1996, natural mortality (M), stock-recruit steepness (h), fishery and survey length selectivity, and the inclusion or exclusion of commercial CPUE effort data (Appendix E Section E.4.1). We further established two robustness-set OMs encompassing less plausible but possible additional sources of structural uncertainty: (1) an OM scenario that assumes the stock was lightly fished before 1996 and excludes the CPUE data; and (2) an OM scenario that assesses robustness to future increases in natural mortality (possibly due to climate change or shifts in predator abundance) (Appendix E Section E.4.2).

We documented the initial parameterization of the OMs (Appendix F) and the conditioning of the OMs using stochastic stock reduction analysis (SRA; Appendix B). The conditioning estimated biomass depletion in the last historical year, the magnitude of unfished recruitment, fishing mortality at age by year, and historical recruitment deviations. We retained only parameter combinations that had plausible fits to the observed data in the SRA (Appendix E Section E.4.3).

Candidate management procedure specification: We worked with the full set of provisional candidate MPs described in Appendix D. We modified the MPs to only observe even years of the survey index to reflect the biennial nature of the WCVI synoptic bottom trawl survey.

Closed-loop simulation application: We applied satisficing criteria based on the LT LRP and STC performance metrics (minimum LT LRP > 0.9, minimum STC > 0.8). We discarded some MPs with near-identical performance to render a manageable set of seven MPs for further consideration. These MPs included index-target MPs, surplus production models paired with a harvest control rule, and two constant-catch MPs.

Presentation of results: We presented a number of visualizations featuring the OM reference-set scenarios to aid evaluating trade-offs amongst the seven satisficed MPs for final decision-making. We averaged results across OM reference-set scenarios, to integrate across the various uncertainties represented by these OMs. We presented probability results in terms of the average (Figures E.13) and minimum performance (Figures E.14), where the latter accounts for the worst-case scenario among OM reference-set scenarios. We suggest that Figures E.13 and E.14 would be the primary source of information used by decision-makers to reach a final MP-selection decision. To aid this decision, we included a number of other visualizations to help decision-makers understand the underlying dynamics and trade-offs of the various OMs across all satisficed MPs.

Finally, we assessed performance of the satisficed MPs under the OM robustness scenarios. We present results from the two OM robustness scenarios separately to allow decision-makers to see performance of MPs under these less plausible OM scenarios. Poor performance of an MP under one of these OM scenarios may influence a decision-maker to select another MP that performs well under both OM reference and robustness scenarios. Four Itarget MPs (Appendix D Section D.3.2) were relatively robust to the OM scenario where the stock was lightly fished prior to 1996 and to increasing natural mortality in the future. Conversely, the constant catch and surplus production MPs had substantially lower probabilities of achieving the LT LRP objective under the lightly fished OM robustness scenario.

We provide a full set of results and interpretation of results in Appendix E.

5 DISCUSSION

In this document we have presented a framework for implementing an MP approach to providing Science advice for groundfish in British Columbia. We outlined an approach to developing appropriate OMs, testing suites of MPs, and identifying MPs that best meet the objectives of fisheries management and stakeholders. Ultimately, it is our aim that the MP Framework improves the capacity of DFO to meet the requirements of Canada's Sustainable Fisheries Framework and the Fish Stocks provisions of the *Fisheries Act* for data-limited species. Although the MP Framework does not emphasize explicit knowledge of fish-stock reference points, it emphasizes selecting management approaches that have a high likelihood of maintaining fish stocks above implicitly known reference points across multiple plausible states of nature. In the following discussion we highlight issues regarding reference points, MP tuning, assessment frequency and triggers, the inclusion of environmental effects, assessing the value of information, and use of this framework as part of stock rebuilding plans.

5.1 IMPLICIT VS. EXPLICIT KNOWLEDGE OF LIMIT REFERENCE POINTS

This MP Framework, and all MSE processes, differ from conventional stock assessments in the way science advice is delivered. In most BC groundfish stock assessments (e.g., Yamanaka et al. 2011; Starr and Haigh 2017; Forrest et al. 2020), catch advice is presented in the form of decision tables, where probabilities of breaching reference points (e.g., probability of the stock falling below the LRP) are presented over a range of possible future TAC levels. Uncertainty can be incorporated into the process in two main ways: (1) within a single model, through treating model parameters (e.g., M, R_0, h, process, and observation error terms) as random variables; and/or (2) developing alternative models to test sensitivity to model assumptions. In the latter case, results from some of these sensitivity models may be averaged to produce a model-averaged decision table (e.g., Forrest et al. 2020), integrating uncertainties across multiple models. This approach depends on explicit reporting of reference points and estimation of stock status. Following the production of a decision table, it is then the job of the decision-makers to select a future TAC based on the probabilities presented in the decision table and their consideration of other factors such as economic needs of the fishery combined with their risk tolerance. In this process, consideration of risk (i.e., probability of breaching reference points and resulting impacts) occurs at the final step of the decision-making process and may not always be transparent or be related to agreed-upon objectives.

MP frameworks differ from conventional assessments in two key ways: (1) reference points and stock status are not explicitly reported (or at least not emphasized); and (2) objectives related to the probability of breaching reference points must be agreed upon at the beginning of the process, i.e., at Step 2 of the best practices (Section 2.2). Reference points and stock status are therefore still an integral component of the framework—they are calculated in the OMs and are built into the performance metrics. Critically, agreement on acceptable risk (e.g., acceptable probabilities of breaching reference points) must be reached at the beginning of the process so that performance metrics and satisficing criteria can be established. The final decision point in this process is the MP that delivers a TAC that meets objectives, while ideally also achieving acceptable trade-offs among other objectives such as catch or variability in catch. An advantage of MP frameworks is that all objectives must be transparently stated and are "baked in" to the final catch advice.

We note that, for many stocks, especially data-limited stocks, it is not possible to reliably estimate biological reference points or estimate stock status. MP frameworks such as this one may be especially important for these stocks The Sustainable Fisheries Framework and the Fish Stocks provisions of the *Fisheries Act* require that fish stocks be maintained at sustainable levels, and particularly above the LRP (Section 1.2). This framework implicitly preserves the intent of these policies, despite the fact that reference points and stock status are not explicitly provided. The MP Framework therefore increases capacity for provision of Sustainable Fisheries Framework- and Fish Stocks provisions-compliant catch advice for data-limited stocks. We recommend that products such as DFO's sustainability survey be flexible to accommodate status reports from MP-approach processes, which may use alternative wording such as: "Under the current management procedure, the stock has a less than [X] probability ([Y] times out of [N] chance) of being below the LRP averaged over a [Z]-year time frame." Alternatively, if a minimum-performance-by-year calculation (ICES 2016) is performed: "... less than [X] probability ([Y] times out of [N] chance) of being below the LRP in each and every year over a [Z]-year time frame."

5.2 TUNING MPS

Many of the MPs in this framework are characterized by one or more parameters that control how the TAC should change in response to changes in the survey index (Appendix D). For example, the Itarget MPs have four "tuning" parameters, w, x, λ, and δ, which control the rate and scale of TAC adjustments in response to changes in the index (Appendix D Section D.3.2). In our Rex Sole case study, we tested six versions of the Itarget MPs over a fairly coarse scale of the tuning parameters. In this particular case, several of the MPs performed well with respect to satisficing criteria so we did not explore more combinations. However, in some applications of the framework it may be desirable to iteratively tune MPs on a fine scale to achieve desired performance outcomes (see Sagarese et al. 2018 for discussion on tuning MPs).

There is a trade-off between testing a larger set of generic MPs across a coarse array of tuning parameters and homing in on better-performing MPs via the satisficing step vs. focusing effort on a few MPs that are highly "tuned" to achieve desired outcomes. In cases where generic MPs perform poorly, the latter approach may be preferred. This latter approach may also be preferred in more mature processes with strong stakeholder engagement, where MPs can be tuned iteratively to meet a more refined set of objectives (e.g., Cox and Kronlund 2008). Ultimately, the decision about whether to evaluate generic or finely tuned MPs will be made on a stock-by-stock basis. The process may start with more generic MPs and graduate to more finely tuned MPs as experience is gained with the performance of specific MPs.

5.3 REASSESSMENT FREQUENCY AND TRIGGERS

In general, the purpose of an MP framework is to identify and select a robust MP that can be left in place for an agreed amount of time. We do not recommend a specific interval between assessments in this framework and suggest this should be done on a stock-by-stock basis. We suggest that the MP Framework itself can be used to test appropriate re-assessment intervals for individual fish stocks (e.g., Huynh et al. 2020). Interim checks between assessments are also recommended to ensure the selected MP is performing as expected.

In addition to the best practice steps described in Section 2, Carruthers and Hordyk (2018a) describe a final evaluation step, where performance of the selected MP is formally reviewed once it has been implemented. Departures from an MP's predicted performance have been termed "exceptional circumstances," where the observed system dynamics fall outside the range of OM scenarios specified in the OM(s), over which the MPs were demonstrated to be robust (Butterworth 2008). Exceptional circumstances can be caused either by misspecification of the original OM(s) or can be due to unforeseen changes in the future system dynamics that were not captured in the original OM(s) (e.g., changes in natural mortality, growth, recruitment, or fishing dynamics). Evidence for exceptional circumstances, occurring within the recommended assessment interval, would trigger a review of the OM(s) and MP, possibly resulting in a new OM, or an adjustment to the selected MP (Carruthers and Hordyk 2018b).

In established MSE processes (e.g., Cox and Kronlund 2008), informal evaluation of the performance of MPs may be done at regular intervals as the MP is applied and new data are gathered (e.g., survey and commercial CPUE information). Carruthers and Hordyk (2018b) list several examples of MSEs where formal protocols for detecting exceptional circumstances have been established. In general, formal protocols have included monitoring the biomass index, catch, and sometimes other data-types such as age-composition data, and comparing observations to the OM predictions. Examples of triggers for re-evaluation include observed data falling outside some confidence interval of the OM-predicted data (e.g., 90% or 95%). Carruthers and Hordyk

(2018b) recommend testing the statistical power of formal protocols to detect exceptional circumstances. This may be especially important for data-limited species, where statistical power may be low due to large uncertainty in OM dynamics. For example, if the confidence interval of a predicted index is extremely large due to OM uncertainties, the likelihood of future observed indices falling outside its range may be low. It may therefore be necessary to use more rigorous test statistics, possibly based on multiple sources of data (e.g., examples provided in Carruthers and Hordyk 2018b).

We recommend regular evaluation of the performance of MPs recommended by this framework but recognize that a formal protocol has not yet been established. We therefore also recommend further analyses to evaluate protocols for detecting exceptional circumstances as a matter of priority.

5.4 INCLUDING ENVIRONMENTAL EFFECTS

The Fish Stocks provisions of the *Fisheries Act* state that fisheries management decisions shall take “into account the biology of the fish and the environmental conditions affecting the stock” (Section 1.2). Changing environmental conditions can affect fish stocks in many ways, including impacting natural mortality rates, growth rates and condition, recruitment success, and spatial distribution, which may impact fishery catchability or selectivity. Robust methods for including environmental considerations into single-species catch advice are not well-established in fisheries decision-making environments (but see Haltuch and Punt 2011; Crone et al. 2019; Haltuch et al. 2019). This is because of large uncertainties associated with observing complex marine environments and understanding mechanisms linking fish population dynamics with environmental change (Myers et al. 1995). Unless mechanisms are well-understood (e.g., Swain and Benoît 2015), incorporating environmental variables into stock assessments does not necessarily improve advice (Punt et al. 2014).

Closed loop simulation-testing is an important means of evaluating how MPs perform in the presence of uncertainty and a changing environment (e.g., Haltuch et al. 2009; Haltuch and Punt 2011; Punt 2011). In this approach, hypotheses about how environmental variables impact population dynamics are incorporated into the OMs, with the performance of alternative MPs (which may or may not include environmental variables) evaluated in the same way as presented in this framework. Multiple OMs can be developed that represent multiple hypotheses for environmental effects. OMs could explicitly include environmental variables (e.g., linking ocean temperature with recruitment deviations, or linking predator abundance with *M*), or could be mechanism-free (e.g., allowing growth parameters or M to vary through time with no specific driver). The latter, mechanism-free approaches are already straightforward to incorporate into DLMtool operating models, via growth or mortality parameters (see Section A.6). The Rex Sole case study included one OM scenario in the robustness set where M increased linearly through time. In this case, no specific mechanism for changing M was provided.

Environmental changes may also increase the frequency of unexpected events such as very large or very small recruitment events. These are sometimes referred to as “black-swan” events (e.g., Taleb 2007; Anderson et al. 2017; Anderson and Ward 2019) or “spasmodic recruitment” (Caddy and Gulland 1983) and can be incorporated into operating models through the addition of alternative distributions for recruitment anomalies to allow for occasional very large or very small events. An upcoming application of the MP Framework for the Inside stock of Yelloweye Rockfish will include an OM scenario that incorporates occasional sudden large recruitment events (Haggarty et al. 2021). We recommend using the tools developed for this MP Framework to further explore simulation-testing approaches to identify MPs that are robust to changing environmental conditions.

5.5 ASSESSING THE VALUE OF INFORMATION

The MP Framework provides the tools to evaluate the "value of information" (VoI) on the performance of MPs. Value-of-information analyses assess whether performance can be improved with the addition of resources invested in data-collection. Conversely, such analyses can evaluate whether similar performance could be achieved if fewer data were available. These are essentially sensitivity analyses to the addition or removal of data sources or changes to data quantity or quality.

Typical data types that could be included in value-of-information analyses in applications of the MP Framework for BC groundfish include survey-index and age-composition data. In terms of survey-index data, this framework could be used to test the sensitivity of MP performance to changes in the precision of population indexes. For example, applications of the framework could quantify the expected change in the probability of maintaining a stock above its LRP in the long term given reductions or expansions in random-stratified survey programs.

The MP Framework could also be used to evaluate whether MP performance could be improved by ageing otoliths for a given stock. This would be achieved by using the OM to simulate age-composition data, possibly with different effective sample sizes, and testing whether MPs that rely on age-composition data (e.g., a full age-structured model) outperform simpler, data-limited MPs such as the ones tested in this document. DLMtool's companion package MSEtool contains a large number of data-rich MPs (age-structured models) that could be used for this purpose (Huynh et al. 2019). This type of analysis could be used to answer questions about whether data-rich MPs would better meet management objectives for a given fish stock and, if so, how much ageing data would be required for a full assessment (e.g., Sagarese et al. 2018). The effects of ageing imprecision and bias on assessment performance could also be evaluated. We note that simulating age- and length-composition data with sufficiently realistic "noise" can be challenging. Evaluating the performance of MPs with composition data that represent the underlying true composition more closely than real data may overestimate the performance of MPs that rely on composition data.

We suggest that expanding the framework to include VoI considerations may also require development of new objectives and performance metrics associated with the costs of expanding research programs, collecting more data, or benefits such as marginal improvements in accuracy of advice and improvements in catch accessibility.

5.6 REBUILDING PLANS

The Precautionary Approach policy and the Fish Stocks provisions of the *Fisheries Act* (Section 1.2) require that management measures be put in place if fish stocks are assessed to be in the critical zone (i.e., below the LRP). Rebuilding plans require that management measures should rebuild stocks out of the critical zone within a specified time frame (e.g., one or two generation times), with a specified probability. The MP Framework can be easily modified to test performance of alternative MPs with respect to meeting rebuilding objectives for stocks requiring rebuilding plans. The only real modifications needed are to the objectives and performance metrics, which are already customizable within this framework.

A challenge for data-limited species is how to determine whether a stock is in the critical zone in the first place, required in order to trigger a rebuilding plan. In light of new legislative requirements for rebuilding under the Fish Stocks provisions, DFO Science is currently developing more detailed guidance on considerations for the design of rebuilding strategies (Kronlund et al. 2021). Drawing on precedents by Health Canada (Weight of Evidence Working Group 2018), Kronlund et al. (2021) suggest a "weight of evidence" approach to determine whether rebuilding is required. A weight of evidence approach may include consideration of combined contributions of individual studies (totality of evidence), and expert judgement-assigned weights for each line of evidence, where a line of evidence may consist of one or more studies. We suggest that if conditioned OMs place a high probability of a fish stock being in the critical zone across a range of plausible OM assumptions, this could contribute to the lines of evidence used to trigger a rebuilding plan.

This framework is currently being applied to evaluate alternative data-limited MPs for the rebuilding plan for the Inside stock of Yelloweye Rockfish in British Columbia (Haggarty et al. 2021).

5.7 DATA-MODERATE AND DATA-RICH STOCKS

This framework has outlined the steps needed to evaluate the performance of data-limited MPs, with a focus on BC groundfish species. Data-richness occurs along a gradient and it may be desirable to apply the framework to evaluate the performance of data-moderate or data-rich MPs. The principles of this framework are the same along the data-richness continuum. The MSEtool R package (Huynh et al. 2019) contains data-moderate (e.g., delay-difference models) and data-rich (e.g., statistical catch-age models) MPs that could easily be incorporated into this framework with some modification. However, we note again that the framework is software agnostic. Other software tools have been used in a similar fashion for data-rich species in BC (e.g., Sablefish: Cox and Kronlund 2008; Pacific Herring: Cox et al. 2019).

5.8 LIMITATIONS OF THE FRAMEWORK

As in all MSE processes, results of this framework will depend on the degree to which uncertainties within the real system are captured within the OMs. For this reason we recommended developing multiple OMs to capture the key, most plausible hypotheses about the system in the reference set, and a wider range of uncertainties in the robustness set. However, it is inevitable that some uncertainties will not be considered, either because they are unknown or because including them would create unworkable complexity in the modelling and decision-making environment. For example, species for which spatial considerations are important (e.g., highly migratory species, or species with strong environmental drivers of productivity). Some considerations will inevitably be considered outside the scope of the process due to limitations in available data, time, or expertise. Therefore it is important to evaluate the performance of selected MPs once they are implemented, either through informal or formal means (Section 5.3).

Our OMs were conditioned on observed data, using the SRA model in MSEtool. Outputs of this model rely on the quality of the available data, and also on the assumed distributions of its input parameters. In particular, assumptions about selectivity will be a key source of uncertainty for species with little or no age composition data. Therefore, selectivity should be treated as an axis of uncertainty in most applications. Furthermore, as the SRA assumes almost no observation error in historical catch data, uncertainty may be underestimated in its outputs. For trawl-caught BC groundfish, catch estimates are considered reliable since the 1996 introduction of 100% at-sea observer coverage. For line-caught species, catch data are considered reliable since

2006, when 100% electronic monitoring was introduced. For some species, especially those with low monetary value, catch data prior to these years may be more uncertain and it may be necessary to include alternative scenarios to account for uncertainty in catch (e.g. Haggarty et al. 2021).

Finally, the success of this framework and any MSE process will depend on adequate engagement by fishery managers, First Nations, and stakeholders. For some species, different user groups may have diverse objectives, creating large tradeoffs and a complex decision-making environment. In these cases, careful governance of the process and attention to Steps 1 and 2 (Section 2) will be especially critical.

6 ACKNOWLEDGEMENTS

We are grateful to the members of our technical advisory group Rob Kronlund, Adam Keizer, Chris Grandin, Ashleen Benson, Dana Haggarty, Rowan Haigh, Roger Kanno, Paul Starr, Rob Tadey, and Greg Workman for helpful discussions and comments on earlier drafts of this document.

We thank Chris Grandin, who developed a number of visualization and report-generation functions within the ggmse R package that greatly improved this document and will be invaluable in applications of the framework.

We thank Tom Carruthers and Adrian Hordyk for their work on developing DLMtool and MSEtool. This project would have been a much bigger undertaking without these tools. Tom and Adrian's helpful guidance and willingness to add many new features on request have greatly improved this framework.

We thank John Holmes for supporting our work on this framework within the Stock Assessment and Research Division at the Pacific Biological Station.

We thank Jaclyn Cleary and Ann-Marie Huang for helpful advice on other MSE processes in BC.

We thank DFO's Ocean and Freshwater Science Contribution Program for supporting the development of DLMtool and MSEtool.

Finally, we thank Carrie Holt and Meaghan Bryan for insightful reviews that greatly improved this document.

7 REFERENCES

Abookire, A.A. 2006. Reproductive biology, spawning season, and growth of female rex sole (*Glyptocephalus zachirus*) in the Gulf of Alaska. Fishery Bulletin 104(3): 350–359.

Albo, Y., Lanir, J., Bak, P., and Rafaeli, S. 2016. Off the radar: Comparative evaluation of radial visualization solutions for composite indicators. IEEE Trans. Vis. Comput. Graph. 22(1): 569–578.

Anderson, S.C., Branch, T.A., Cooper, A.B., and Dulvy, N.K. 2017. Black-swan events in animal populations. Proc. Natl. Acad. Sci. USA 114(12): 3252–3257.

Anderson, S.C., Grandin, C., Edwards, A.M., Grinnell, M.H., Ricard, D., and Haigh, R. 2020a. csasdown: Reproducible CSAS reports with bookdown. R package version 0.0.8.

Anderson, S.C., Grandin, C., and Forrest, R.E. 2020b. ggmse: Tools for working with openMSE (including DLMtool, MSEtool, and SAMtool). R package version 0.0.1.9000.

Anderson, S.C., Keppel, E.A., and Edwards, A.M. 2019. A reproducible data synopsis for over 100 species of British Columbia groundfish. DFO Can. Sci. Advis. Sec. Res. Doc. 2019/041. vii + 321 p.

Anderson, S.C., Keppel, E.A., Edwards, A.M., and Grandin, C. 2020c. gfplot: Data extraction and plotting for DFO PBS groundfish stocks. R package version 0.1.4.

Anderson, S.C., and Ward, E.J. 2019. Black swans in space: Modelling spatiotemporal processes with extremes. Ecology 100(1): e02403.

Armitage, D.R., Okamoto, D.K., Silver, J.J., Francis, T.B., Levin, P.S., Punt, A.E., Davies, I.P., Cleary, J.S., Dressel, S.C., Jones, R.R., Kitka, H., Lee, L.C., MacCall, A.D., McIsaac, J.A., Poe, M.R., Reifenstuhl, S., Shelton, A.O., Schmidt, J.O., Thornton, T.F., Voss, R., and Woodruff, J. 2019. Integrating Governance and Quantitative Evaluation of Resource Management Strategies to Improve Social and Ecological Outcomes. BioScience 69(7): 523–532.

Australian Government, D. of A. and W.R. 2018. Guidelines for the Implementation of the Commonwealth Fisheries Harvest Strategy Policy. Canberra.

Bache, S.M., and Wickham, H. 2014. magrittr: A forward-pipe operator for R. R package version 1.5.

Benson, A.J., and Stephenson, R.L. 2018. Options for integrating ecological, economic, and social objectives in evaluation and management of fisheries. Fish Fish. 19(1): 40–56.

Berger, A.M., Edwards, A.M., Grandin, C.J., and Johnson, K.F. 2019. Status of the Pacific Hake (Whiting) stock in U.S. And Canadian waters in 2019. Prepared by the Joint Technical Committee of the U.S. And Canada Pacific Hake/Whiting Agreement. National Marine Fisheries Service; Fisheries and Oceans Canada.

Berkson, J., and Thorson, J.T. 2015. The determination of data-poor catch limits in the United States: Is there a better way? ICES J. Mar. Sci. 72(1): 237–242.

Beverton, R. J. H., and Holt, S.J. 1957. On the dynamics of exploited fish populations. Fisheries Investigations, 19, 1-533.

Botsford, L.W. 1981. Optimal fishery policy for size-specific, density-dependent population models. J. Math. Biol. 12(3): 265–293.

Butterworth, D., Punt, A., and Smith, A. 1996. On plausible hypotheses and their weighting, with implications for selection between variants of the Revised Management Procedure. Rep.-Int. Whal. Comm. 46: 637–642.

Butterworth, D.S. 2008. Some lessons from implementing management procedures. Edited by K. Tsukamoto, T. Kawamura, T. Takeuchi, T.D. Beard, Jr., And M.J. Kaiser. *In* Fisheries for Global Welfare and Environment, 5th World Fisheries Congress 2008. TERRAPUB, Toyko. pp. 381–397.

Butterworth, D.S., and Punt, A.E. 1999. Experiences in the evaluation and implementation of management procedures. ICES J. Mar. Sci. 56(6): 985–998.

Caddy, J.F., and Gulland, J.A. 1983. Historical patterns of fish stocks. Mar. Policy 7(4): 267–278.

Carruthers, T., and Hordyk, A. 2019. Data-limited methods toolkit (DLMtool 5.4.0). User guide.

Carruthers, T.R., and Hordyk, A. 2018a. The data-limited methods toolkit (DLMtool): An R package for informing management of data-limited populations. Methods Ecol. Evol. 9: 2388–2395.

Carruthers, T.R., and Hordyk, A.R. 2018b. Using management strategy evaluation to establish indicators of changing fisheries. Can. J. Fish. Aquat. Sci.: 1–16.

Carruthers, T.R., Hordyk, A.R., Huynh, Q.C., Singh, R., and Barrett, T.J. 2020. A Framework for Conditioning Operating Models for the Southwest Nova Scotia/Bay of Fundy Spawning Component of 4VWX Herring.

Carruthers, T.R., Kell, L.T., Butterworth, D.D.S., Maunder, M.N., Geromont, H.F., Walters, C., McAllister, M.K., Hillary, R., Levontin, P., Kitakado, T., and Davies, C.R. 2016. Performance review of simple management procedures. ICES J. Mar. Sci. J. Cons. 73(2): 464–482.

Carruthers, T.R., Punt, A.E., Walters, C.J., MacCall, A., McAllister, M.K., Dick, E.J., and Cope, J. 2014. Evaluating methods for setting catch limits in data-limited fisheries. Fish. Res. 153: 48–68.

Cooke, J.G. 1999. Improvement of fishery-management advice through simulation testing of harvest algorithms. ICES J. Mar. Sci. 56(6): 797–810.

Cope, J., Dick, E.J., MacCall, A., Monk, M., Soper, B., and Wetzel, C. 2015. Data-moderate stock assessments for brown, China, copper, sharpchin, stripetail, and yellowtail rockfishes and English and rex soles in 2013. Pacific Fishery Management Council, 7700 NE Ambassador Pl, Suite 101, Portland, OR 97220.

Cox, S.P., and Benson, A.J. 2016. Roadmap to more sustainable Pacific herring fisheries in Canada: A step-by-step guide to the management strategy evaluation approach.

Cox, S.P., Benson, A.J., Cleary, J.S., and Taylor, N.G. 2019. Candidate limit reference points as a basis for choosing among alternative harvest control rules for Pacific Herring (*Clupea pallasii*) in British Columbia. DFO Can. Sci. Advis. Sec. Res. Doc. 2019/050. viii + 47 p.

Cox, S.P., Doherty, B., Benson, A.J., Johnson, S.D., and Haggarty, D. 2020. Evaluation of potential rebuilding strategies for Outside Yelloweye Rockfish in British Columbia. DFO Can. Sci. Advis. Sec. Res. Doc. 2020/069. viii + 135 p.

Cox, S.P., and Kronlund, A.R. 2008. Practical stakeholder-driven harvest policies for groundfish fisheries in British Columbia, Canada. Fish. Res. 94(3): 224–237.

Cox, S.P., Kronlund, A.R., and Benson, A.J. 2013. The roles of biological reference points and operational control points in management procedures for the Sablefish (*Anoplopoma Fimbria*) fishery in British Columbia, Canada. Environ. Conserv. 40(04): 318–328.

Crone, P.R., Maunder, M.N., Lee, H., and Piner, K.R. 2019. Good practices for including environmental data to inform spawner-recruit dynamics in integrated stock assessments: Small pelagic species case study. Fish. Res. 217: 122–132.

de La Mare, W.K. 1998. Tidier fisheries management requires a new MOP (management-oriented paradigm). Reviews in Fish Biology and Fisheries 8(3): 349–356.

DFO. 2006. A Harvest Strategy Compliant with the Precautionary Approach. DFO Can. Sci. Advis. Sec. Sci. Advis. Rep. 2006/023.

DFO. 2009. A Fishery Decision-Making Framework Incorporating the Precautionary Approach.

DFO. 2013. Guidance for the development of rebuilding plans under the Precautionary Approach framework: Growing stocks out of the critical zone.

DFO. 2016. Proceedings of the Pacific regional peer review on A Review of International Best Practices to Assigning Species to Tiers for the Purposes of Stock Assessment Based on Data Availability and Richness; May 30-31, 2016. DFO Can. Sci. Advis. Proceed. Ser. 2016/051.

DFO. 2017a. Pacific Region Integrated Fishery Management Plan. Groundfish. Effective February 21, 2017.

DFO. 2017b. DLMtool phase II: Developing a management strategy evaluation package for advancing the science and management of data-limited and at-risk Canadian fish stocks. PAC2016.12.

DFO. 2019. Evaluation of Management Procedures for Pacific Herring (*Clupea pallasii*) in the Strait of Georgia and the West Coast of Vancouver Island Management Areas of British Columbia. DFO Can. Sci. Advis. Sec. Sci. Advis. Rep. 2019/001.

DFO. 2020. Evaluation of Management Procedures for Pacific Herring (*Clupea pallasii*) in Haida Gwaii, Prince Rupert District, and the Central Coast Management Areas of British Columbia. DFO Can. Sci. Advis. Sec. Sci. Resp. 2020/003.

Diehl, S., Beck, F., and Burch, M. 2010. Uncovering strengths and weaknesses of radial visualizations — an empirical approach. IEEE Trans. Vis. Comput. Graph. 16(6): 935–942.

Dowling, N.A., Dichmont, C.M., Haddon, M., Smith, D.C., Smith, A.D.M., and Sainsbury, K. 2015a. Empirical harvest strategies for data-poor fisheries: A review of the literature. Fish. Res. 171: 141–153.

Dowling, N.A., Dichmont, C.M., Haddon, M., Smith, D.C., Smith, A.D.M., and Sainsbury, K. 2015b. Guidelines for developing formal harvest strategies for data-poor species and fisheries. Fish. Res. 171: 130–140.

Fargo, J.J. 2012. Biomass estimates and biological reference points for flatfish species resident to the coast of British Columbia, Canada. DFO Can. Tech. Rep. Fish. Aquat. Sci. 2995: 88 p.

Feldman, R. 2013. Filled radar charts should not be used to compare social indicators. Soc. Indic. Res. 111(3): 709–712.

Fletcher, R.I. 1978. On the restructuring of the Pella-Tomlinson system. Fish. Bull. 76: 515–521.

Forrest, R.E., Anderson, S.C., Grandin, C.J., and Starr, P.J. 2020. Assessment of Pacific Cod (*Gadus macrocephalus*) for Hecate Strait and Queen Charlotte Sound (Area 5ABCD), and West Coast Vancouver Island (Area 3CD) in 2018. DFO Can. Sci. Advis. Sec. Res. Doc. 2020/070. v + 215 p.

Forrest, R.E., Holt, K.R., and Kronlund, A.R. 2018. Performance of alternative harvest control rules for two Pacific groundfish stocks with uncertain natural mortality: Bias, robustness and trade-offs. Fish. Res. 206: 259–286.

Fox, W.W. 1970. An exponential surplus-yield model for optimizing exploited fish populations. Transactions of the American Fisheries Society 99(1): 80–88.

Garnier, S. 2018. viridis: Default color maps from 'matplotlib'. R package version 0.5.1.

Gelman, A., Carlin, J.B., Stern, H.S., Dunson, D.B., Vehtari, A., and Rubin, D.B. 2014. Bayesian Data Analysis. *In* Third edition. Chapman & Hall, Boca Raton, FL.

Geromont, H.F., and Butterworth, D.S. 2015a. Complex assessments or simple management procedures for efficient fisheries management: A comparative study. ICES J. Mar. Sci. 72(1): 262–274.

Geromont, H.F., and Butterworth, D.S. 2015b. Generic management procedures for data-poor fisheries: Forecasting with few data. ICES J. Mar. Sci. 72(1): 251–261.

Goodrich, B., Gabry, J., Ali, I., and Brilleman, S. 2018. Rstanarm: Bayesian applied regression modeling via Stan. R package version 2.17.4.

Grandin, C.J., and Forrest, R.E. 2017. Arrowtooth Flounder (*Atheresthes stomias*) stock assessment for the west coast of British Columbia. DFO Can. Sci. Advis. Sec. Res. Doc. 2017/25. v + 87 p.

Gregory, R., Failing, L., Harstone, M., Long, G., and McDaniels, T.L. (*Editors*). 2012. Structured decision making: A practical guide to environmental management choices. Wiley-Blackwell, Oxford.

Haggarty, D.R., Huynh, Q.C., Anderson, S.C., Bresch, M., Keppel, E.A., and Forrest, R.E. 2021. Evaluation of potential rebuilding strategies for Inside Yelloweye Rockfish (*Sebastes ruberrimus*) in British Columbia. DFO Can. Sci. Advis. Sec. Res. Doc.: In press.

Haltuch, M.A., Brooks, E.N., Brodziak, J., Devine, J.A., Johnson, K.F., Klibansky, N., Nash, R.D.M., Payne, M.R., Shertzer, K.W., Subbey, S., and Wells, B.K. 2019. Unraveling the recruitment problem: A review of environmentally-informed forecasting and management strategy evaluation. Fish. Res. 217: 198–216.

Haltuch, M.A., and Punt, A.E. 2011. The promises and pitfalls of including decadal-scale climate forcing of recruitment in groundfish stock assessment. Can. J. Fish. Aquat. Sci. 68(5): 912–926.

Haltuch, M.A., Punt, A.E., and Dorn, M.W. 2009. Evaluating the estimation of fishery management reference points in a variable environment. Fish. Res. 100(1): 42–56.

Hilborn, R. 2007. Defining success in fisheries and conflicts in objectives. Marine Policy 31(2): 153–158.

Hilborn, R., Fulton, E.A., Green, B.S., Hartmann, K., Tracey, S.R., and Watson, R.A. 2015. When is a fishery sustainable? Can. J. Fish. Aquat. Sci. 72(9): 1433–1441.

Huynh, Q.C., Hordyk, A.R., and Carruthers, T. 2019. MSEtool: Management strategy evaluation toolkit. R package version 1.4.3.

Huynh, Q.C., Hordyk, A.R., Forrest, R.E., Porch, C.E., Anderson, S.C., and Carruthers, T.R. 2020. The interim management procedure approach for assessed stocks: Responsive management advice and lower assessment frequency. Fish Fish. In press.

ICES. 2012. ICES implementation of advice for data-limited stocks in 2012 in its 2012 advice. ICES CM 2012/ACOM 68.

ICES. 2016. ICES criteria for defining multi-annual plans as precautionary. ICES Advice Technical Guidelines, Section 12.4.10. *In* ICES Advice 2016, Book 12.

ICES. 2018. ICES Advice Basis.

Jardim, E., Azevedo, M., and Brites, N.M. 2015. Harvest control rules for data limited stocks using length-based reference points and survey biomass indices. Fish. Res. 171: 12–19.

Kell, L.T., De Oliveira, J.A.A., Punt, A.E., McAllister, M.K., and Kuikka, S. 2006. Chapter 15 Operational management procedures: An introduction to the use of evaluation frameworks. Developments in Aquaculture and Fisheries Science 36: 379–407.

Keppel, E.A., Anderson, S.C., Edwards, A.M., and Grandin, C. 2020. gfdata: Data extraction for DFO PBS groundfish stocks. R package version 0.0.0.9000.

Kimura, D.K., and Tagart, J.V. 1982. Stock Reduction Analysis, another solution to the catch equations. Can. J. Fish. Aquat. Sci. 39(11): 1467–1472.

Kristensen, K., Nielsen, A., Berg, C.W., Skaug, H., and Bell, B.M. 2016. TMB: Automatic differentiation and Laplace approximation. J. Stat. Softw. 70(5): 1–21.

Kronlund, A.R., Marentette, J.R., Olmstead, M., Shaw, J., and Beauchamp, B. 2021. Considerations for the design of rebuilding strategies for Canadian fish stocks. DFO Can. Sci. Advis. Sec. Res. Doc.: In press.

Mace, P., M., and Doonan, I.J. 1988. A generalized bioeconomic simulation model for fish population dynamics. N.Z. Fish. Assess. Res. Doc. 88/4.

Mapstone, B.D., Little, L.R., Punt, A.E., Davies, C.R., Smith, A.D.M., Pantus, F., McDonald, A.D., Williams, A.J., and Jones, A. 2008. Management strategy evaluation for line fishing in the Great Barrier Reef: Balancing conservation and multi-sector fishery objectives. Fish. Res. 94(3): 315–329.

Marentette, J.R., and Kronlund, A.R. 2020. A Cross-Jurisdictional Review of International Fisheries Policies, Standards and Guidelines: Considerations for a Canadian Science Sector Approach. DFO Can. Tech. Rep. Fish. Aquat. Sci. 3342: xii + 169 p. Ottawa.

Marentette, J.R., Kronlund, A.R., and Cogliati, K.M. 2021. Specification of Precautionary Approach Reference Points and Harvest Control Rules in Domestically Managed and Assessed Major Fish Stocks In Canada. DFO Can. Sci. Advis. Sec. Res. Doc.: In press.

Maunder, M.N. 2012. Evaluating the stock-recruitment relationship and management reference points: Application to summer flounder (*Paralichthys* dentatus) in the U.S. Mid-Atlantic. Fish. Res. 125-126: 20–26.

McAllister, M.K., Pikitch, E.K., and Babcock, E.A. 2001. Using demographic methods to construct Bayesian priors for the intrinsic rate of increase in the Schaefer model and implications for stock rebuilding. Can. J. Fish. Aquat. Sci. 58(9): 1871–1890.

McGilliard, C.R., and Palsson, W. 2017. Assessment of the rex sole stock in the Gulf of Alaska.

McIlgorm, A. 2013. Literature study and review of international best practice in fisheries harvest strategy policy approaches. A report to the Department of Agriculture, Fisheries and Forestry (DAFF) Canberra, by the Australian National Centre for Ocean Resources and Security (ANCORS), University of Wollongong. Final Report, March 26, 2013.

Miller, D.C.M., and Shelton, P.A. 2010. "Satisficing" and trade-offs: Evaluating rebuilding strategies for Greenland halibut off the east coast of Canada. ICES J. Mar. Sci. 67(9): 1896–1902.

Munk, K.M. 2001. Maximum ages of groundfishes in waters off Alaska and British Columbia and considerations of age determination. Alaska Fishery Research Bulletin 8(1).

Myers, R.A., Bowen, K.G., and Barrowman, N.J. 1999. The maximum reproductive rate of fish at low population sizes. Can. J. Fish. Aquat. Sci. 56: 2404–2419.

Myers, R., Barrowman, N., and Thompson, K. 1995. Synchrony of recruitment across the North Atlantic: An update.(or, "now you see it, now you don't!"). ICES J. Mar. Sci. 52(1): 103–110.

Newman, D., Berkson, J., and Suatoni, L. 2015. Current methods for setting catch limits for data-limited fish stocks in the United States. Fish. Res. 164: 86–93.

Okamoto, D.K., Poe, M.R., Francis, T.B., Punt, A.E., Levin, P.S., Shelton, A.O., Armitage, D.R., Cleary, J.S., Dressell, S.C., Jones, R., Kitka, H., Lee, L.C., MacCall, A.D., McIsaac, J.A., Reifenstuhl, S., Silver, J.J., Schmidt, J.O., Thornton, T.F., Voss, R., and Woodruff, J. 2020. Attending to spatial social–ecological sensitivities to improve trade-off analysis in natural resource management. Fish Fish. 21(1): 1–12.

Pestal, G., Ryall, P., and Cass, A. 2008. Collaborative Development of Escapement Strategies for Fraser River Sockeye: Summary Report 2003 - 2008.

Plagányi, É.E., Putten, I. van, Hutton, T., Deng, R.A., Dennis, D., Pascoe, S., Skewes, T., and Campbell, R.A. 2013. Integrating indigenous livelihood and lifestyle objectives in managing a natural resource. Proc. Natl. Acad. Sci. USA 110(9): 3639–3644.

Plagányi, É.E., Rademeyer, R.A., Butterworth, D.S., Cunningham, C.L., and Johnston, S.J. 2007. Making management procedures operational—innovations implemented in South Africa. ICES J. Mar. Sci. 64(4): 626–632.

Porch, C.E., Turner, S.C., and Schirripa, M.J. 2004. The commercial landings of red snapper in the Gulf of Mexico from 1872 to 1962. SEDAR, North Charleston, South Carolina.

Punt, A.E. 2011. The impact of climate change on the performance of rebuilding strategies for overfished groundfish species of the U.S. West coast. Fish. Res. 109(2-3): 320–329.

Punt, A.E. 2017. Strategic management decision-making in a complex world: Quantifying, understanding, and using trade-offs. ICES J. Mar. Sci. 74.

Punt, A.E., A'mar, T., Bond, N.A., Butterworth, D.S., Moor, C.L. de, De Oliveira, J.A.A., Haltuch, M.A., Hollowed, A.B., and Szuwalski, C. 2014. Fisheries management under climate and environmental uncertainty: Control rules and performance simulation. ICES J. Mar. Sci. 71(8): 2208–2220.

Punt, A.E., Butterworth, D.S., de Moor, C.L., De Oliveira, J.A.A., and Haddon, M. 2016. Management strategy evaluation: Best practices. Fish Fish. 17(2): 303–334.

R Core Team. 2019. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Rademeyer, R.A., and Butterworth, D.S. 2006a. Summary of the most recent South African hake assessments. BCLME Joint Hake Research Planning workshop, Cape Town, South Africa.

Rademeyer, R.A., and Butterworth, D.S. 2006b. Management procedure robustness tests for the South African *Merluccius paradoxus* and *M. capensis* resources under a constant catch scenario.

Rademeyer, R.A., Plagányi, É.E., and Butterworth, D.S. 2007. Tips and tricks in designing management procedures. ICES J. Mar. Sci. 64(4): 618–625.

Ricker, W.E. 1954. Stock and recruitment. J Fish. Res. Board Can. 11(5): 559–623.

Sagarese, S.R., Harford, W.J., Walter, J.F., Bryan, M.D., Isely, J.J., Smith, M.W., Goethel, D.R., Rios, A.B., Cass-Calay, S.L., Porch, C.E., Carruthers, T.R., and Cummings, N.J. 2018. Lessons learned from data-limited evaluations of data-rich reef fish species in the Gulf of Mexico: Implications for providing fisheries management advice for data-poor stocks. Canadian Journal of Fisheries and Aquatic Sciences: 1–16.

Sainsbury, K. 2008. Best Practice Reference Points for Australian Fisheries. Report to the Australian Fisheries Management Authority.

Schaefer, M.B. 1954. Some aspects of the dynamics of populations important to the management of the commercial marine fisheries. Bulletin of the Inter-American Tropical Tuna Commission 1(2): 27–56.

Shelton, A.O., Thorson, J.T., Ward, E.J., and Feist, B.E. 2014. Spatial semiparametric models improve estimates of species abundance and distribution. Can. J. Fish. Aquat. Sci. 71(11): 1655–1666.

Smith, A.D.M., Sainsbury, K.J., and Stevens, R.A. 1999. Implementing effective fisheries-management systems – management strategy evaluation and the Australian partnership approach. ICES J. Mar. Sci. 56(6): 967–979.

Smith, D., Punt, A., Dowling, N., Smith, A., Tuck, G., and Knuckey, I. 2009. Reconciling Approaches to the Assessment and Management of Data-Poor Species and Fisheries with Australia's Harvest Strategy Policy. Marine and Coastal Fisheries: Dynamics, Management, and Ecosystem Science: 244–254.

Stan Development Team. 2020. RStan: The R interface to Stan. R package version 2.19.3.

Stanley, R.D., Karim, T., Koolman, J., and McElderry, H. 2015. Design and implementation of electronic monitoring in the British Columbia groundfish hook and line fishery: A retrospective view of the ingredients of success. ICES Journal of Marine Science 72(4): 1230–1236.

Starr, P.J., and Haigh, R. 2017. Stock assessment of the coastwide population of Shortspine Thornyhead (*Sebastolobus alascanus*) in 2015 off the British Columbia coast. DFO Can. Sci. Advis. Sec. Res. Doc. 2017/015. ix + 174 p.

Stephenson, R.L., Benson, A.J., Brooks, K., Charles, A., Degnbol, P., Dichmont, C.M., Kraan, M., Pascoe, S., Paul, S.D., Rindorf, A., and Wiber, M. 2017. Practical steps toward integrating economic, social and institutional elements in fisheries policy and management. ICES J. Mar. Sci. 74(7): 1981–1989.

Swain, D.P., and Benoît, H.P. 2015. Extreme increases in natural mortality prevent recovery of collapsed fish populations in a Northwest Atlantic ecosystem. Mar. Ecol. Prog. Ser. 519: 165–182.

Taleb, N.N. 2007. The black swan: The impact of the highly improbable. Random House Inc., New York.

Thorson, J.T., Jensen, O.P., and Zipkin, E.F. 2014. How variable is recruitment for exploited marine fishes? A hierarchical model for testing life history theory. Can. J. Fish. Aquat. Sci. 71(7): 973–983.

Thorson, J.T., Shelton, A.O., Ward, E.J., and Skaug, H.J. 2015. Geostatistical delta-generalized linear mixed models improve precision for estimated abundance indices for West Coast groundfishes. ICES J. Mar. Sci. 72(5): 1297–1310.

Turris, B.R. 2000. A comparison of British Columbia's ITQ fisheries for groundfish trawl and sablefish: Similar results from programmes with differing objectives, designs and processes. Food; Agriculture Organization (FAO). pp. 254–261.

Walters, C.J., Martell, S.J.D., and Korman, J. 2006. A stochastic approach to stock reduction analysis. Can. J. Fish. Aquat. Sci. 63(1): 212–223.

Weight of Evidence Working Group. 2018. Weight of evidence: General principles and current applications at Health Canada. Prepared for the Task Force on Scientific Risk Assessment by the Weight of Evidence Working Group. Health Canada.

Wickham, H. 2016. ggplot2: Elegant graphics for data analysis. Springer-Verlag New York.

Wilderbuer, T.K., and Nichol, D.G. 2015. Assessment of the other flatfish stock complex in the Bering Sea and Aleutian Islands. *In* Stock assessment and fishery evaluation report for the groundfish resources of the Bering Sea/Aleutian Islands Regions. North Pacific Fishery Management Council, 605 W 4 th Ave, Suite 306, Anchorage AK 99501. pp. 1047–1064.

Yamanaka, K.L., McAllister, M.K., Olesiuk, P.F., Etienne, M.-P., Obdradovich, S.G., and Haigh, R. 2011. Stock assessment for the inside population of Yelloweye Rockfish (*Sebastes ruberrimus*) in British Columbia, Canada for 2010. DFO Can. Sci. Advis. Sec. Res. Doc. 2011/129. xiv + 131.

APPENDIX A. DLMTOOL OPERATING MODEL

The population dynamics operating model (OM), as implemented in DLMtool, is an age-structured model described below. The OM is flexible, with options to allow a number of parameters (e.g., natural mortality, growth, selectivity) to vary through time (see Section A.6). Multiple sub-areas can also be described, and outcome uncertainty can be represented through the addition of implementation uncertainty and bias parameters. Full documentation of the DLMtool OM, including the range of possible parameterizations, is described in the appendices of Carruthers and Hordyk (2018a), adapted here for the MP Framework. Here, we adapt the documentation in Carruthers and Hordyk (2018a), following their notation conventions where possible for consistency. We present the simplest version of the OM, without time-varying parameters or multiple sub-areas. This simple OM structure should be suitable as a base OM for most stocks intended to be assessed within the framework. More complex structural assumptions, including time-varying parameters, could be developed as OM scenarios when justified for individual stocks.

Following the lead of Carruthers and Hordyk (2018a), we denote parameters that are user-defined and sampled stochastically across replicates with a tilde (e.g. $\widetilde{M}$). For example, $\widetilde{M}_i \sim f(\theta)$ represents the i^{th} simulation draw of parameter M being sampled from the distribution function f and parameters θ. In most cases, the stochastic parameters could be sampled from a uniform distribution or specified directly after sampling from any distribution or model.

There are two distinct time-periods in the simulation framework: (1) the historical period, which includes the years from the first year of the catch time-series t_1 to the final year of the catch time series t_c (where "c" represents the "current" year); and (2) the projection period, which covers the period from the first year after t_c to the final projection year t_N. The historical period is conditioned to historical observations using an age-structured stock-reduction analysis [SRA; Kimura and Tagart (1982); Walters et al. (2006)], described in Appendix B. The closed-loop simulation, with application of the MPs and calculation of performance metrics, begins in the first year of the projection period (year $t_c + 1$).

OM development in the MP Framework follows three steps:

1. Set parameter values and ranges in the OM—OM equations are provided in this appendix; main parameters are defined in Table A.1; default parameter settings are provided in Appendix C. Parameter values may be drawn from biological and fishery data, local stock assessments, the scientific literature, meta-analyses from similar species, stock assessments from other regions, expert judgement, or other sources where relevant. Examples of user-defined parameter settings drawn from a range of sources are provided in Appendix F, and in Haggarty et al. (2021).
2. Pass the OM parameters to the SRA model, which conditions the OM by fitting an age-structured SRA to historical observed catches and indices of abundance (with options to fit to age- and/or length-composition data if available). This results in conditioned estimates of model parameters and estimates of historical biomass and historical fishing mortality (in years t_1 to t_c), which are consistent with historical observations. Equations for the SRA model are provided in Appendix B.
3. Pass the conditioned parameter values back to the OM (now the "conditioned" OM) for use in the closed-loop simulation projections, starting in year $t_c + 1$.

The SRA updates one or more of the following OM parameters, depending on data availability (see Appendix B for details):

- B_{t_c}/B_0 (or "D"; depletion in the last historical year t_c)

- R_0 (unfished recruitment)
- θ_{AC} (or "AC"; first-order autocorrelation of recruitment deviations)
- $\varepsilon_{\mathrm{R},y}$ for years t_1 to t_c (annual recruitment deviations for the historical period)
- $F_{a,y}$ (fishing mortality at age by year for the historical period)
- Selectivity parameters (age-based or length-based).

These parameters are defined in context in the following sections. Non-updated parameters are used as specified in the OM.

Note that covariance among SRA parameter estimates is preserved when they are passed back to the OM. In general, covariance among other specified parameters is not accounted for since parameters are specified by single values or distributions. However, it is possible to include covariance among parameters if this information is available. For example, distributions of growth and maturity parameters for the Rex Sole case study are drawn from their posterior probability distributions (Appendix F, Section F).

A.1 POPULATION DYNAMICS MODEL

A.1.1 GROWTH AND MATURITY

The OM assumes that growth follows a von Bertalanffy model:

$$L_a = \widetilde{L}_\infty \left(1 - \exp(-\widetilde{\kappa}(a - \widetilde{t}_0)\right)), \tag{A.1}$$

where L_a represents length at age a, $\widetilde{\kappa}$ represents the growth rate, $\widetilde{L}_\infty$ represents the maximum length and $\widetilde{t}_0$ represents the theoretical age where length is zero.

Weight-at-age (w_a) is assumed to be related to length-at-age (L_a) by:

$$w_a = \widetilde{\alpha}_W \; {L_a}^{\widetilde{\beta}_W}, \tag{A.2}$$

where α_W and β_W are species-specific length-weight coefficients. Maturity is specified according to length, using a logistic model with parameters describing length-at-50% maturity ($\widetilde{\theta}_{l50}$) and the length increment from 50% to 95% maturity ($\widetilde{\theta}_{l50-95}$). Maturity (m_l) at length (l) is then calculated as:

$$m_l = \frac{1}{1 + e^{-\ln 19\left(\frac{l - \widetilde{\theta}_{l50}}{\widetilde{\theta}_{l50-95}}\right)}}, \tag{A.3}$$

where maturity (m_a) at age is calculated from the length-at-age:

$$m_a = \frac{1}{1 + e^{-\ln 19\left(\frac{a - \theta_{a50}}{\theta_{a95} - \theta_{a50}}\right)}}, \tag{A.4}$$

and θ_{a50} and θ_{a95} represent the age at 50% and 95% maturity given by the inverse von-Bertalanffy growth curve:

$$\theta_{a50} = \frac{-\ln\left(1 - \widetilde{\theta}_{l50}/\widetilde{L}_\infty\right)}{\widetilde{\kappa}} + \widetilde{t}_0, \tag{A.5}$$

$$\theta_{a95} = \frac{-\ln\left(1 - \left(\widetilde{\theta}_{l50}/\widetilde{L}_\infty + \widetilde{\theta}_{l50\_95}/\widetilde{L}_\infty\right)\right)}{\widetilde{\kappa}} + \widetilde{t}_0. \tag{A.6}$$

A.1.2 POPULATION DYNAMICS

Leading parameters of the population dynamics model, defined as fixed values or ranges by the user, are: unfished recruitment R_0, the steepness parameter of the stock-recruit relationship (h) (Mace and Doonan 1988), and natural mortality (M). Recruitment and numbers-at-age are initialized in the first historical year using the SRA model (Appendix B, Equations B.3 to B.6). During the projection period (years $y = t_c + 1$ to $y = t_N$), the number of fish recruited to the first age group at the beginning of each year $N_{y,a=1}$ is calculated as a function of the previous year's spawning biomass B_{y-1}. The framework assumes a Beverton-Holt stock-recruit relationship (Beverton and Holt 1957) with annual lognormal recruitment deviations $\varepsilon_{R,y}$, but analysts can alternatively choose a Ricker function (Ricker 1954):

$$N_{y,a=1} = R_y = \begin{cases} \dfrac{\alpha^{\mathrm{B}} B_{y-1}}{1 + \beta^{\mathrm{B}} B_{y-1}} \varepsilon_{R,y} & \text{Beverton-Holt} \\ \alpha^{\mathrm{R}} B_{y-1} \exp(-\beta^{\mathrm{R}} B_{y-1}) \varepsilon_{R,y} & \text{Ricker,} \end{cases} \tag{A.7}$$

where α^{B} and β^{B} are the parameters of the Beverton-Holt stock-recruit relationship, and α^{R} and β^{R} are the parameters of the Ricker stock-recruit relationship, derived from user-defined parameters steepness ($\widetilde{h}$) and $\widetilde{R}_0$, where $\alpha^{\mathrm{B}} = \frac{4\tilde{h}}{(1-\tilde{h})\phi_0}$, $\beta^{\mathrm{B}} = \frac{5\tilde{h}-1}{(1-\tilde{h})B_0}$ and $\alpha^{\mathrm{R}} = \frac{(5\tilde{h})^{1.25}}{\phi_0}$, $\beta^{\mathrm{R}} = \frac{\log(5\tilde{h})}{B_0}$. The parameter ϕ_0 is unfished spawners-per-recruit (see Equation A.41, with $F^e = 0$) and B_0 is unfished spawning biomass, calculated as

$$B_0 = \sum_{a=1}^{A} m_a \; w_a \; \widetilde{R}_0 e^{-\widetilde{M}(a-1)}, \tag{A.8}$$

where A is the user-defined maximum age and $\widetilde{R}_0 e^{\widetilde{M}(a-1)}$ gives unfished numbers-at-age. Spawning biomass B_y in a given year is calculated as

$$B_y = \sum_{a=1}^{A} w_a m_a N_{y,a}, \tag{A.9}$$

where $N_{y,a}$ is annual numbers-at-age, given in Equation A.13.

Annual log recruitment deviations ($\log \varepsilon_{R,y}$) are generated by first sampling from a normal distribution with standard deviation $\widetilde{\sigma}_R$:

$$\log \delta_{R,y} \sim \mathrm{Normal}\left(\frac{-0.5\widetilde{\sigma}_R^2(1 - \widetilde{\theta}_{\mathrm{AC}})}{\sqrt{1 - \widetilde{\theta}_{\mathrm{AC}}^2}}, \widetilde{\sigma}_R \right), \tag{A.10}$$

where the mean term includes a mean-bias correction for autocorrelated lognormal variables and the term $\widetilde{\theta}_{\mathrm{AC}}$ represents first-order autocorrelation. We use the notation Normal(mean, standard deviation) rather than Normal(mean, variance) here and throughout. Temporal autocorrelation $\widetilde{\theta}_{\mathrm{AC}}$ is added to these initial draws as:

$$\log \varepsilon_{R,y} = \widetilde{\theta}_{\mathrm{AC}} \log \varepsilon_{R,y-1} + \delta_{R,y} \sqrt{\left(1 - \widetilde{\theta}_{\mathrm{AC}}^2\right)}, \tag{A.11}$$

$$\varepsilon_{R,y} = \exp\left(\log \varepsilon_{R,y}\right). \tag{A.12}$$

The number of fish N in each year $y > t_c$ is then calculated from the numbers in the previous year and age class, subject to the instantaneous total mortality rate (Z) at age (a):

$$N_{y,a} = \begin{cases} R_y & a = 1 \\ N_{y-1,a-1}\, e^{-Z_{y-1,a-1}} & 1 < a < A \\ N_{y-1,a-1}\, e^{-Z_{y-1,a-1}} + N_{y-1,a}\, e^{-Z_{y-1,a}} & a = A, \end{cases} \tag{A.13}$$

where A is the maximum age class and the annual total mortality rate $Z_{y,a}$ is given by

$$Z_{y,a} = \widetilde{M} + v_a F_{y,a}, \tag{A.14}$$

where $F_{y,a}$ is fishing mortality-at-age during year y, defined in Equation A.24), and v_a is selectivity-at-age, defined in Equation A.16.

Note that Equation A.13 treats the maximum age class A as a plus group (i.e., the age class A includes all fish of age A and older). This is not the default in DLMtool, which usually makes the assumption that A is large enough to include all age classes (i.e., $N_{y,a} = N_{y-1,a-1} e^{-Z_{y-1,a-1}}$ is used for all age classes including $a = A$). However, when the model is conditioned using the SRA (Appendix B), the SRA defaults to treating A as a plus group (Equation B.9) and, to ensure consistency between the historical and projection periods, the SRA forces DLMtool to also treat A as a plus group for the closed-loop simulations in the projection period. Since we recommend always conditioning OMs for BC groundfish using the SRA, we include the plus group in Equation A.13.

A.2 FLEET DYNAMICS

During the projection period, the rate of fishing mortality-at-age ($F_{y,a}$) is calculated from the TAC prescribed by the MP ($\mathrm{TAC}_{\mathrm{MP},y}$). To distribute catches over ages it is first necessary to calculate the distribution of vulnerable biomass ($B^V_{y,a}$) across ages:

$$B^V_{y,a} = N_{y,a} w_a v_a \dot{R}_a \tag{A.15}$$

where $\dot{R}_a$ represents retention rate at age and the selectivity-at-age v_a (or vulnerability-at-age) is calculated according to whether users define selectivity to be asymptotic (logistic) or dome-shaped, using the following length-based function:

$$v_a = \begin{cases} 2^{-\frac{(\widetilde{L}_a - \widetilde{L}^{\mathrm{FS}})^2}{\widetilde{\sigma}^2_{\mathrm{asc}}}} & \text{if } L_a < \widetilde{L}^{\mathrm{FS}}, \\ 1 & \text{if logistic and } L_a \geq \widetilde{L}^{\mathrm{FS}}, \\ 2^{-\frac{(\widetilde{L}_a - \widetilde{L}^{\mathrm{FS}})^2}{\widetilde{\sigma}^2_{\mathrm{des}}}} & \text{if dome-shaped and } L_a \geq \widetilde{L}^{\mathrm{FS}} \end{cases} \tag{A.16}$$

where $\widetilde{L}^{\mathrm{FS}}$ represents the user-defined length-at-full-selectivity. The standard deviation of the ascending limb ($\widetilde{\sigma}^2_{\mathrm{asc}}$) is given by

$$\sigma_{\mathrm{asc}} = \frac{\widetilde{L}^5 - \widetilde{L}^{\mathrm{FS}}}{\sqrt{-\log_2 0.05}} \tag{A.17}$$

where $\widetilde{L}^5$ is the user-defined length-at-5%-selectivity.

The standard deviation of the descending limb is given by

$$\sigma_{\text{des}} = \frac{\widetilde{L}_\infty - \widetilde{L}^{\text{FS}}}{\sqrt{-\log_2 \widetilde{V}_{L_\infty}}}, \tag{A.18}$$

where $\widetilde{V}_{L_\infty}$ is the user-defined selectivity-at-maximum-length. Setting this parameter to 1 necessarily fixes selectivity to be logistic (Equation A.18 is undefined when $\widetilde{V}_{L_\infty} = 1$).

DLMtool currently models all catch as coming from a single fleet. However, if the OM is conditioned using the SRA model (Appendix B), the SRA can accommodate multiple fleets, and selectivity is fleet-specific. Fleet-specific selectivity in the SRA is calculated using Equation B.1 with fleet-specific user-defined settings for $\widetilde{L}^5$, $\widetilde{L}^{FS}$ and $\widetilde{V}_{L_\infty}$, or is estimated if age- or length-composition data are provided. In this case, selectivity in the DLMtool OM in the projection period is replaced in Equation A.14 with the SRA-conditioned estimate of fishing mortality-at-age in the final year (t_c) of the historical period ($\Sigma_f v_{a,f} F_{t_c,f}$), normalized by dividing by apical F (maximum F across ages) in that year. This provides fractions of F-at-age derived from the relative selectivity-at-age across fleets (i.e., catch-weighted selectivity-at-age). The closed-loop simulation projections therefore assume that the relative selectivities across fleets remains constant in the projection period.

Similarly, if the OM is conditioned using the SRA model, analysts can also specify (or estimate) selectivity parameters for the individual indices of abundance. In this case, the SRA passes all of the indices back to DLMtool, preserving the estimated or user-defined selectivities-at-age for each index. It is the coding of specific MPs that determines which index is used in the MPs during the projection period.

The realized projected catches $C_{y,a}$ are the TAC recommendations across ages (with possible implementation error). Projected catches may account for retention rate ($\dot{R}_{y,a}$) and post-release discard mortality rate $\widetilde{\theta}_{M\text{disc}}$ in the presence of discarding. Implementation uncertainty (I_{TAC}, Equation A.37) may also be accounted for (e.g., in the presence of consistent under-utilization of TACs):

$$C_{y,a} = \frac{B^V_{y,a}}{\sum_a^A B^V_{y,a}} \text{TAC}_{\text{MP},y}\ I_{\text{TAC},y}\ \frac{\dot{R}_{y,a} + \left(1 - \dot{R}_{y,a}\right)\widetilde{\theta}_{\text{Mdisc}}}{\dot{R}_{y,a}} \tag{A.19}$$

where (I_{TAC}) is user-defined (according to other parameters; Equation A.37), and the overall retention rate-at-age $\dot{R}_{y,a}$, is a combination of an age-specific retention $r_{y,a}$ with a maximum value of 1 (Equation A.21), and a constant rate of discarding $\widetilde{\gamma}$:

$$\dot{R}_{y,a} = r_{y,a}\ (1 - \widetilde{\gamma}). \tag{A.20}$$

where age-specific retention $r_{y,a}$, is modelled using the same form of double-normal curve as selectivity:

$$r_{y,a} = \begin{cases} 2^{-\frac{(L_{y,a} - \widetilde{L}_{\text{rmax}})^2}{\sigma^2_{\text{rasc}}}} & L_{y,a} \le \widetilde{L}_{\text{rmax}}, \\ 2^{-\frac{(L_{y,a} - \widetilde{L}_{\text{rmax}})^2}{\sigma^2_{\text{rdesc}}}} & L_{y,a} > \widetilde{L}_{\text{rmax}}, \end{cases} \tag{A.21}$$

where $\widetilde{L}_{\text{rmax}}$ is the length at maximum retention. The standard deviation parameter of the ascending limb is given by the length at 5% retention $\widetilde{L}_{r5}$:

$$\sigma_{\text{rasc}} = \frac{\widetilde{L}_{r5} - \widetilde{L}_{\text{rmax}}}{\sqrt{-\log_2 0.05}} \tag{A.22}$$

while the standard deviation of the descending limb is given by retention $\widetilde{r}_{L_\infty}$ of fish of length $\widetilde{L}_\infty$:

$$\sigma_{\text{rdesc}} = \frac{\widetilde{L}_\infty - \widetilde{L}_{\text{rmax}}}{\sqrt{-\log_2 \widetilde{r}_{L_\infty}}}. \tag{A.23}$$

The retention curve approaches an asymptotic curve as $\widetilde{r}_{L_\infty} \to 1$.

Fishing mortality rates are then calculated from the realized catches $C_{y,a}$ subject to the constraint that they do not exceed F_{max}:

$$F_{y,a} = \min\left(-\ln\left(1 - \frac{C_{y,a}}{N_{y,a} w_a}\right), \widetilde{F}_{\text{max}}\right). \tag{A.24}$$

F_{max} can be adjusted but is set to 3 by default.

A.3 OBSERVATION DYNAMICS

The observation dynamics emulate the collection of data for use in the MPs. Two fundamental types of data are simulated by the OM: (1) time series data (e.g., annual catches from 1970–2017); and (2) catch composition data (e.g., length or age samples).

A.3.1 TIME SERIES DATA

Time series data are simulated with various types of error that would be expected from real-life sampling of fisheries data (e.g., lognormal observation error on indices of abundance). The standard index of abundance in DLMtool is calculated by adding observation error and bias to annual total biomass (B_y^T) through a term $\omega_{I,y}$ that includes bias and imprecision in the index observations:

$$I_y^{\text{obs}} = \omega_{I,y} \frac{{B_y^T}^{\widetilde{\beta}}}{\frac{1}{t_c}\sum_{i=1}^{t_c} {B_i^T}^{\widetilde{\beta}}} \tag{A.25}$$

where B_y^T is total biomass given by the sum over ages of the weight-at-age w_a, and numbers at age $N_{y,a}$:

$$B_y^T = \sum_{a=1}^{A} w_a N_{y,a} \tag{A.26}$$

and where $\widetilde{\beta}$ is a hyperstability/hyperdepletion parameter. When $\widetilde{\beta}$ is 1 the index is linearly related to spawning biomass B_y. When $\widetilde{\beta}$ is greater than 1 the index is hyperdeplete and moves faster than true spawning biomass changes. When $\widetilde{\beta}$ is lower than 1 the index is hyperstable and moves slower than true spawning biomass changes. When the observed index is calculated, it is normalized to have a mean value of 1 over all years.

The term $\omega_{I,y}$ represents imprecision in observations via $\widetilde{\sigma}_I$:

$$\omega_{I,y} = \exp\left(\varepsilon_{I,y} - \frac{\widetilde{\sigma}_I^2}{2}\right), \tag{A.27}$$

where the lognormal error term, ε, is drawn from a normal distribution whose standard deviation σ_I can be sampled at random in each simulation:

$$\varepsilon_{I,y} \sim \text{Normal}(0, \widetilde{\sigma}_I). \tag{A.28}$$

The DLMtool OM can also be set up to reflect one or more real observed indices of biomass or abundance in the historical period and apply the specified survey selectivity-at-age to the projected period. We recommend this approach for the MP Framework assuming that the OM has been conditioned via an SRA model (Appendix B). In this case, the index values in years t_1 to t_c are specified as data and the index values I (without observation error) in year y (for $y > t_c$) and for survey s are calculated as:

$$I_{y,s} = q_s \sum_{a=1}^{A} v_{a,s} N_{y,a} w_{y,a} \tag{A.29}$$

for a biomass-based index and:

$$I_{y,s} = q_s \sum_{a=1}^{A} v_{a,s} N_{y,a} \tag{A.30}$$

for an abundance-based index. The symbol q_s represents catchability as estimated via the SRA (Equation B.19 or B.20). The selectivity-at-age ($v_{a,s}$) is modelled using the same logistic form as for the commercial fleet (Equations A.16–A.18) but with its own shape as specified by the user or estimated via the SRA (Equations B.1).

The OM then scales and adds observation error to each index as:

$$I^{\text{obs}}_{y,s} = \omega_{I,y,s} \frac{{I_{y,s}}^{\widetilde{\beta}}}{\frac{1}{t_c}\sum_{i=1}^{t_c} I_{i,s}}, \tag{A.31}$$

with $\omega_{I,y,s}$ calculated as:

$$\omega_{I,y,s} = \exp\left(\varepsilon_{I,y,s} - \frac{\widetilde{\sigma}_I^2}{2}\right), \tag{A.32}$$

where the lognormal error term ε, is drawn from a normal distribution whose standard deviation $\widetilde{\sigma_I}$ is sampled at random in each simulation:

$$\varepsilon_{I,y,s} \sim \text{Normal}\left(0, \widetilde{\sigma}_I\right). \tag{A.33}$$

A.3.2 CATCH COMPOSITION DATA

Two types of catch composition observations are simulated, catches-by-age-class-by-year (CAA) and catches-by-length-class-by-year (CAL). Although we do not propose any provisional MPs that use CAA or CAL, future applications may explore the value of information of these types of data, e.g., related to aging fish otoliths. Therefore, we describe the observation model component of CAA. We do not describe the observation model of CAL at this time.

The observation model for CAA uses a simple multinomial distribution to generate observed catches at age for the projection period $C^{\text{obs}}_{y,a}$, accounting for the user-defined effective sample size ($\widetilde{\text{ESS}}_{\text{CAA}}$, the number of independent observations) and the average annual number of samples (n_{CAA}, number of individuals aged). For example, $\widetilde{\text{ESS}}_{\text{CAA}}$ independent catch samples-at-age (e.g., 20 per year) are sampled in proportion p_a to the catches-at-age predicted by the model ($C_{y,a}$, Equation A.19):

$$\dot{C}^{\text{obs}}_{y,a} \sim \text{Multinomial}\left(\widetilde{\widetilde{\text{ESS}}}_{\text{CAA}}, p_a = \ C_{y,a}\right). \tag{A.34}$$

For each year, the proportion of samples at age is then inflated to match the total sample size $\widetilde{n}_{\mathrm{CAA}}$ and rounded to the nearest integer (nint):

$$C^{\mathrm{obs}}_{y,a} = \mathrm{nint}\left(\frac{\dot{C}^{\mathrm{obs}}_{y,a}\,\widetilde{n}_{\mathrm{CAA}}}{\widetilde{\mathrm{ESS}}_{\mathrm{CAA}}}\right). \tag{A.35}$$

Due to rounding, this model generates frequency-at-age data that are approximately equal to the average annual sample size:

$$\sum_{a}^{A} C^{\mathrm{obs}}_{y,a} \approx \widetilde{n}_{\mathrm{CAA}}. \tag{A.36}$$

A.4 IMPLEMENTATION DYNAMICS

This framework only considers MPs that provide TAC advice for quota-managed groundfish fisheries. Given that BC groundfish fisheries are subject to 100% at-sea and dockside observer coverage, we make the assumption that under-reporting of catch will be negligible in projection years. However, for non-target species, there is a possibility that TACs will not be fully used. The implementation uncertainty in TACs is applied in Equation A.19.

The TAC implementation uncertainty term (I_{TAC}) is the product of a user-defined fraction of the TAC taken $\widetilde{b}_{\mathrm{TAC}}$ and a degree of inter-annual variability controlled by $\widetilde{\sigma}_{\mathrm{TAC}}$.

$$I_{\mathrm{TAC},y} = \widetilde{b}_{\mathrm{TAC}} \cdot \exp\left(\varepsilon_{\mathrm{TAC},y} - \frac{\widetilde{\sigma}^2_{\mathrm{TAC}}}{2}\right). \tag{A.37}$$

For example, $b_{\mathrm{TAC}} = 0.7$ is equivalent to 30% catch underages. The error term $\varepsilon_{\mathrm{TAC},y}$ is drawn from a normal distribution whose standard deviation $\widetilde{\sigma}_{\mathrm{TAC}}$ is user-defined and can be sampled at random in each simulation:

$$\varepsilon_{TAC,y} \sim \mathrm{Normal}\left(0, \widetilde{\sigma}_{\mathrm{TAC}}\right). \tag{A.38}$$

A.5 CALCULATION OF MSY-BASED REFERENCE POINTS

Biological reference points (BRPs) in the MP Framework are currently based on the provisional reference points suggested in Canada's Precautionary Approach Framework (DFO 2006, 2009), where the limit reference point (LRP) is defined as the OM value of $0.4B_{\mathrm{MSY}}$ and the upper stock reference (USR) is defined as the OM value of $0.8B_{\mathrm{MSY}}$. B_{MSY} is defined as the equilibrium spawning biomass B^e that would occur if the stock were fished at the constant rate of fishing mortality F^e that produces maximum sustainable yield MSY.

Calculation of BRPs is done using estimated parameters from the conditioning stage, described in Appendix B. In the absence of annual variability in life history and/or selectivity parameters, F_{MSY} is calculated by numerically solving for the value of equilibrium F^e that maximizes the equilibrium yield Y^e

$$Y^e = F^e R^e \phi^b, \tag{A.39}$$

where R^e is equilibrium recruitment (Equation A.40) and ϕ^b is equilibrium vulnerable biomass-per-recruit, defined in Equation A.44 below.

In this framework, we assume a Beverton-Holt stock-recruit relationship (Beverton and Holt 1957), although DLMtool also allows users to select the Ricker form (Ricker 1954):

$$R^e = \begin{cases} \dfrac{\alpha^{\mathrm{B}}\phi - 1}{\beta^{\mathrm{B}}\phi} & \text{Beverton-Holt} \\ \dfrac{\log(\alpha^{\mathrm{R}}\phi)}{\beta^{\mathrm{R}}\phi} & \text{Ricker,} \end{cases} \tag{A.40}$$

where α^{B} and β^{B} are the parameters of the Beverton-Holt stock recruit relationship, and α^{R} and β^{R} are the parameters of the Ricker stock recruit relationship, defined above (see Equation B.10.

Following the approach of Botsford (1981), equilibrium spawning biomass-per-recruit is calculated

$$\phi = \sum_{a=1}^{A} \iota_a w_a m_a, \tag{A.41}$$

where ι_a is equilibrium survivorship-at-age:

$$\iota_a = \begin{cases} 1, & a = 1 \\ \iota_{a-1} e^{-Z^e_{a-1}}, & 1 < a < A \\ \dfrac{\iota_{a-1} e^{-Z^e_{a-1}}}{1 - e^{-Z^e_a}}, & a = A, \end{cases} \tag{A.42}$$

and Z^e_a is equilibrium total mortality-at-age:

$$Z^e_a = \widetilde{M} + F^e_a v_a. \tag{A.43}$$

Finally, using the same approach, ϕ^b is calculated:

$$\phi^b = \sum_{a=1}^{A} \iota_a w_a v_a. \tag{A.44}$$

After numerically solving Equation A.39 for F_{MSY}, B_{MSY} is calculated:

$$B_{\mathrm{MSY}} = \phi \mathrm{MSY} R^e, \tag{A.45}$$

with $F^e = F_{\mathrm{MSY}}$ in Equation A.42.

Note that, as for Equation A.13, the maximum age class A is treated as a plus group in Equation A.42. While this is not the default setting in DLMtool, it becomes the default when the SRA is used to condition the OM (see explanation in Section A.1).

A.6 TIME-VARYING PARAMETERS

The OM as described above assumes that model parameters representing growth, selectivity, and fishery catchability are time invariant. DLMtool allows users to set a number of key parameters to vary through time, according to user-specified slope and variation parameters. For example, if the growth parameters $\widetilde{L}_\infty$ and $\widetilde{\kappa}$ are set to be time-varying (Carruthers and Hordyk 2018a, their Appendix C, Equations 7 and 8), then the von Bertalanffy growth curve will vary through time. This will affect other model variables that depend on length, including weight-at-age (Equation A.2), maturity-at-length and age (Equations A.3 and A.4), and selectivity-at-age (Equation A.16).

Similarly, M can be set to be time-varying, either as a function of: (1) a user-defined percentage annual increase or (2) weight-at-age relative to maximum weight, or both (Carruthers and Hordyk 2018a, their Appendix C, Equation 4). Fishery catchability q can also be set to vary through time during the projection period (Carruthers and Hordyk 2018a, their Appendix D, Equation 42). Note that setting growth and mortality parameters to be time-varying in the DLMtool OM will result in them being carried through to the SRA model.

In the presence of time-varying parameters, B_0 is calculated as

$$B_0 = \frac{\sum_{y=1}^{\theta_{a50}} B^y_{\mathrm{unfished}}}{\theta_{a50}}, \tag{A.46}$$

which is the average unfished spawning biomass over the first θ_{a50} (age-at-maturity) years of the historical period (rounded up to the nearest integer) and B^y_{unfished} is calculated B_0 (Equation A.8). Other alternative interpretations of B_0 could also be programmed (e.g., average B_0 across the whole historical period, or at the end of the historical period), depending on the desired benchmark or target.

Time-varying growth, natural mortality and/or selectivity will result in time-varying MSY-based reference points, because DLMtool calculates MSY and related metrics (e.g., B_{MSY}) at each time-step in the model, in which case annual performance metrics will be based on annual estimates of reference points. Analysts should exercise caution with time-varying B_{MSY} and B_0. Declining reference points can imply improved stock status relative to lower reference points, possibly leading to a “ratcheting” effect as both stock and reference points decline. Alternatively, it is possible to base performance metrics on an average B_{MSY} from the historical period or on some other desired benchmark (e.g., the estimated average biomass from a productive historical period DFO 2009; Forrest et al. 2018).

It is recommended that appropriate benchmarks or targets in the presence of time-varying productivity be resolved in collaboration with fisheries management, stakeholders, and First Nations, as performance metrics based on reference points such as B_{MSY} and B_0 will be impacted by the methods used to calculate them. Such decisions should be documented to ensure transparency.

How to treat reference points when parameters, particularly productivity parameters such as $\widetilde{M}$ or $\widetilde{\kappa}$, are time-varying is an active area of research (e.g., Haltuch et al. 2009; Punt et al. 2014; Haltuch et al. 2019). We suggest that the MP Framework is a suitable environment for further simulation-testing alternative approaches for time-varying reference points and suggest this is an important research area for future applications of the framework.

A.7 DLMTOOL SLOTS

Table A.1. DLMtool operating model (OM) 'slots', parameter symbols, and brief descriptions. In the 'OM elements' column, 'Obs.' and 'Imp.' refer to the observation and implementation components, respectively. 'Slots' is the technical name for the parameter elements within the DLMtool R package. All parameters in this table (except n_a, α_W, and β_W) can be defined as stochastic with draws from a distribution for each simulation replicate.

OM element	Slot	Parameter	Description
Stock	`maxage`	n_a	The maximum age of individuals that is simulated.
Stock	`R0`	R_0	The magnitude of unfished recruitment.
Stock	`M`	M	Natural mortality rate.
Stock	`h`	h	Steepness of the stock-recruit relationship.
Stock	`Perr`	σ_R	SD of recruitment deviations in log space.
Stock	`AC`	θ_{AC}	Autocorrelation in recruitment deviations.
Stock	`Linf`	L_∞	von Bertalanffy maximum expected length.
Stock	`K`	κ	von Bertalanffy growth parameter.
Stock	`t0`	t_0	von Bertalanffy theoretical age at length zero.
Stock	`L50`	θ_{l50}	Length at 50% maturity.
Stock	`L50_95`	θ_{l50-95}	Length increment from 50% to 95% maturity.
Stock	`a`	α_W	Length-weight parameter.
Stock	`b`	β_W	Length-weight parameter.
Fleet	`L5`	L^5	Shortest length corresponding to 5% vulnerability.
Fleet	`LFS`	L^{FS}	Shortest length that is fully vulnerable to fishing.
Fleet	`Vmaxlen`	V_{L_∞}	The vulnerability of fish at maximum expected length.
Fleet	`LR5`	L_{r5}	Shortest length fish corresponding to 5% retention.
Fleet	`LFR`	L_{rmax}	Shortest length fish that are fully retained.
Fleet	`Rmaxlen`	r_{L_∞}	The retention of fish at L_∞.
Obs.	`Iobs`	σ_I	Observation error SD of the relative abundance index in log space.
Obs.	`Cobs`	σ_C	Observation error SD of the catch in log space.
Obs.	`Cbiascv`	b_C	CV controlling the sampling of bias in catch observations.
Obs.	`beta`	β	Hyperstability/hyperdepletion parameter.
Obs.	`CAA_ESS`	$\mathrm{ESS}_{\mathrm{CAA}}$	Effective sample size of multinomial catch-at-age observation error model.
Obs.	`CAA_nsamp`	n_{CAA}	Number of catch-at-age observations per time step.
Imp.	`TACFrac`	b_{TAC}	Mean fraction of TAC taken.
Imp.	`TACSD`	σ_{bTAC}	SD in lognormal fraction of TAC taken.

APPENDIX B. STOCK REDUCTION ANALYSIS FOR CONDITIONING OPERATING MODELS

B.1 A BACKGROUND ON STOCK REDUCTION ANALYSIS

The operating model (OM) defined in Appendix A is conditioned by fitting an age-structured stock-reduction analysis (SRA) (Kimura and Tagart 1982; Walters et al. 2006) to historical observed catches and indices of abundance (and to age- and length-composition data if available). This step produces estimates of parameters that are conditioned to produce time-series trajectories consistent with historical observations. This is especially important in data-limited or data-moderate situations, where the lack of an accepted assessment makes it difficult to parameterize historical depletion and fishing mortality F.

In the literature, the term "stock reduction analysis" has been used to describe a model in which the predicted total catch matches the observed catch. Kimura and Tagart (1982) presented SRA as an alternative to a virtual production analysis (VPA) or surplus production models. The SRA approach is more applicable to data-limited species than, for example, a VPA, which requires annual catch-at-age data. The SRA uses an age-structured modeling approach that incorporates natural mortality and recruitment information for reconstructing the stock history, rather than taking a pooled biomass-dynamics approach (e.g., surplus production model). Another advantage of the SRA approach for BC groundfish fisheries is that it accommodates conditioning on multiple indices of abundance (e.g., accommodating indices from multiple fishery-independent surveys and/or commercial CPUE series).

For data-limited species, where it may only be possible to provide ranges of values for key parameters, the SRA approach tries to reduce the range of plausible parameter values to be consistent with historical observations, via its likelihood functions. In other words, it helps address the broader question: What combinations of historical fishing mortality and recruitment could have generated the observed data (Walters et al. 2006)?

We note that an SRA does not necessarily require age-composition data, as long as credible estimates of maturity-at-length or age and vulnerability-at-length or age can be provided. However, if age- or length-composition data are available, even for one or a few years, these can be used to inform vulnerability-at-age in the model.

B.2 STOCK REDUCTION ANALYSIS IN THE CONTEXT OF DLMTOOL

The SRA described here can be fit using TMB (Kristensen et al. 2016) with the R package MSEtool via the function `SRA_scope()`. The function takes a DLMtool OM object (Appendix A) and historical observed data (observed catches, indices of abundance and, if available, composition data), fits the SRA to the observed data, and returns a list with an updated OM and outputs from the SRA.

The approach can be stochastic (with Monte Carlo sampling) if the operating model is specified with ranges on key parameters. For example, the steepness parameter of the stock-recruit relationship (h) is usually highly uncertain. The initial OM can specify a range of values for h, for example: $h \sim \mathrm{Uniform}(0.6, 0.9)$. In this case, with $i = 1$ to n (e.g., $n = 250$) replicates in the closed-loop simulation, the SRA function will sample n steepness values from this distribution and fit the SRA model n times. The SRA model reconstruction from the i^{th} fit will be conditioned on the i^{th} sampled value of h. The sampled values of h (as well as all other input parameters to the SRA) are then saved alongside the estimated parameters to form the conditioned OM.

B.3 THE STOCK REDUCTION ANALYSIS MODEL

The SRA model can be conditioned on catch or fishing effort. For BC groundfish species, we recommend conditioning the model on catch. Catch data since the introduction of 100% at-sea and dockside observer coverage in 1996 are known with very little error. Effort data in multi-species fisheries are more difficult to interpret. If the model is conditioned on catch, then the SRA will generate predicted catches that match the observed catches. A full time series of the conditioning variable (i.e., catch) is required.

If the catch time series is long enough, the historical period can be assumed to begin at unfished conditions ($B_{t=1} = B_0$, where $B_{t=1}$ is spawning biomass in the first historical year and B_0 represents equilibrium stock size under unfished conditions). However, for some BC groundfish species, catch records may be less reliable prior to the introduction of comprehensive catch monitoring (e.g., in 1996 for the BC groundfish trawl fleet). In these cases, the SRA model is set up such that a value for equilibrium catch in each fleet f prior to the first year of the catch time series is assumed (C_f^{eq}). We note that C_f^{eq} need not be the true catch prior to the first year—factors such as recruitment, predator mortality, or any number of catch trends could have caused the estimated initial depletion—but C_f^{eq} represents the corresponding catch in equilibrium. The SRA will then use this value to estimate the initial depletion ($\frac{B_{t=1}}{B_0}$) in the first year of the historical period. Therefore, C_f^{eq} is a convenient means of initializing the model at different levels of depletion.

Initial depletion is difficult to estimate with precision without other sources of information (e.g., a long index of abundance, or age-composition data). We therefore recommend treating C_f^{eq} as a major axis of uncertainty if this approach is taken. An alternative approach is to use catch-reconstruction methods (e.g., Porch et al. 2004; Starr and Haigh 2017) to reconstruct the catch time series back to the time when the stock was considered unfished. While this approach avoids the uncertainty associated with estimating depletion based on C_f^{eq}, it may introduce other sources of uncertainty associated with the reconstruction approach.

In addition to the conditioning time series, additional data types can be used, which do not need to be available in every year of the historical time series:

- Time series of indices of abundance (either as surveyed biomass or fishery-dependent CPUE)
- Age-composition data
- Length-composition data
- Mean length (in commercial or survey data)

Multiple surveys and fleets can be accommodated by the SRA. Including one of these data types in addition to the conditioning catch time series is generally needed to inform depletion estimates. Even in cases where availability of these data is sparse over time, they can still be informative for estimates of depletion. For example, an age-composition sample from a single recent year that shows a very truncated age structure can imply a heavily depleted stock. Age-composition data from one or a few years can also be informative about selectivity, which in turn informs estimates of stock size.

B.3.1 PARAMETERS OF THE SRA MODEL

The required pre-specified OM parameters needed for SRA scoping are as follows (DLMtool slot names are provided in Table A.1):

- Growth parameters L_∞, κ and a_0.
- Length-weight conversion factors α_W and β_W.
- Maturity parameters θ_{l50} and θ_{l50-95}.
- Natural mortality M.
- Steepness h.

If growth, maturity, or natural mortality are set to be time-varying in the historical period, then the SRA will implement time-varying life history in the estimation model as well.

- The stock-recruit relationship type (Beverton-Holt or Ricker)
- Unfished recruitment (R_0), which is used as the starting value for estimation if the model is conditioned on catch.
- Selectivity parameters L_5, L_{LFS} and V_{L_∞}.

If there are no age- or length-composition data, then selectivity in the model is fixed to these values. Otherwise, the ascending limb of selectivity is estimated with age or length composition data. If the selectivity is assumed to be dome-shaped, then the descending limb can either be fixed or estimated in the SRA. See Equation B.1.

- The recruitment deviation first-order autocorrelation (θ_{AC}) is estimated post-hoc from the recruitment deviation estimates.
- The standard deviation of recruitment deviations σ_R is taken as input. Historical recruitment deviations are updated by the SRA model (Equation B.10), while recruitment deviations in the projection period are sampled with autocorrelation (Equation A.10).

If initial depletion (Equation B.7) is estimated, then the annual recruitment deviations in the first year are adjusted in order to produce the estimated abundance-at-age in the first year of the SRA model (Equation B.10).

If multiple fleets are used for conditioning, then selectivity-at-age ($v_{a,f}$, Equation B.1) will be updated based on the relative fishing mortality among fleets. The default assumption in the projection period of the closed-loop simulation is that the selectivity and relative fishing mortality among fleets are identical to those in the last historical year t_c. See Section A.2.

B.4 DESCRIPTION OF THE SRA MODEL

B.4.1 VULNERABILITY-AT-AGE AND MORTALITY

Fleet-specific selectivity-at-age (or vulnerability-at-age) ($v_{a,f}$) is length-based and modelled in the same way as in the DLMtool OM (Appendix A), with the only difference being that multiple fleets (f) can be accommodated. For fleet f with asymptotic selectivity, a two-parameter logistic function is used, with parameters defining the length-of-5%-selectivity (L^5_f) and the length-of-full-selectivity L^{FS}_f. For dome-shaped selectivity, a third parameter, the selectivity at L_∞, $v_f^{L_\infty}$ is also used. Length-based selectivity is converted to age-based selectivity as:

$$v_{a,f} = \begin{cases} 2^{-[(L_a - L_f^{\mathrm{FS}})/(\sigma_f^{\mathrm{asc}})]^2} & \text{if } L_a < L_f^{\mathrm{FS}} \\ 1 & \text{if logistic and } L_a \geq L_f^{\mathrm{FS}} \\ 2^{-[(L_a - L_f^{\mathrm{FS}})/(\sigma_f^{\mathrm{des}})]^2} & \text{if dome-shaped and } L_a \geq L_f^{\mathrm{FS}}, \end{cases} \tag{B.1}$$

where $L_{y,a}$ is the mean length-at-age, and $\sigma_f^{\mathrm{asc}} = (L_f^5 - L_f^{\mathrm{FS}})/\sqrt{-\log_2(0.05)}$ and $\sigma_f^{\mathrm{des}} = (L_\infty - L_f^{\mathrm{FS}})/\sqrt{-\log_2(v^{L_\infty})}$ control the shape of the ascending and descending limbs, respectively, of the selectivity function. In this parameterization, length-based selectivity is constant over time. The corresponding age-based selectivity is constant over time if growth is not time-varying.

See Appendix A, Section A.2 for descriptions of handling of multiple fleets in the DLMtool OM during the projection period.

Total mortality Z in year y and for age a is given by

$$Z_{y,a} = M + \Sigma_f v_{a,f} F_{y,f}, \tag{B.2}$$

where $F_{y,f}$ is fishing mortality in year y and fleet f, and M is natural mortality, assumed here to be constant.

B.4.2 INITIAL POPULATION DISTRIBUTION

Numbers-at-age in the first year of the model $y = 1$ are assumed to be in an equilibrium state:

$$N_{1,a} = \begin{cases} R^{\mathrm{eq}} & a = 1 \\ N_{1,a-1}\exp(-Z_a^{\mathrm{eq}}) & a = 2, \ldots, A-1 \\ \dfrac{N_{1,a-1}\exp(-Z_a^{\mathrm{eq}})}{1 - \exp(-Z_A^{\mathrm{eq}})} & a = A, \end{cases} \tag{B.3}$$

where the R^{eq} is the equilibrium recruitment (Equation B.6) and Z_a^{eq} is the equilibrium total mortality rate:

$$Z_a^{\mathrm{eq}} = M + \Sigma_f v_{a,f} F_f^{\mathrm{eq}}. \tag{B.4}$$

If the stock is assumed to be unfished in the first year of the historical period ($y = t_1$), unfished conditions are modelled by setting $F_f^{\mathrm{eq}} = 0$. In practical terms, this is done by the user setting the equilibrium catch for each fleet to zero ($C_f^{\mathrm{eq}} = 0$).

If the stock was not unfished in year t_1, the population is assumed to be in an equilibrium state with catch equal to a user-defined equilibrium catch (C_f^{eq}), defined as a fraction of the observed catch in year t_1. In this case the SRA estimates predicted equilibrium catch ($C_f^{\mathrm{eq,pred}}$), which is fit to C_f^{eq}, with very low standard deviation (0.01) in the likelihood function (Equation B.27). Equilibrium catch is predicted by the Baranov equation, summed across fleets ($f = 1 : nf$) and ages ($a = 1 : A$):

$$C_f^{\mathrm{eq,pred}} = \sum_{a=1}^{A} \frac{v_{a,f} F_f^{\mathrm{eq}}}{Z_a^{\mathrm{eq}}} (1 - \exp^{-Z^{\mathrm{eq}}}) N_{1,a} w_a, \tag{B.5}$$

where F_f^{eq} are estimated parameters.

Once Z_a^{eq} is obtained, then the equilibrium recruitment is calculated, using either a Beverton-Holt (Beverton and Holt 1957) or Ricker (Ricker 1954) stock-recruit relationship:

$$R^{\mathrm{eq}} = \begin{cases} \dfrac{\alpha^{\mathrm{B}}\phi - 1}{\beta^{\mathrm{B}}\phi} & \text{Beverton-Holt} \\ \dfrac{\log(\alpha^{\mathrm{R}}\phi)}{\beta^{\mathrm{R}}\phi} & \text{Ricker,} \end{cases} \tag{B.6}$$

where ϕ is equilibrium spawners-per-recruit (see Equation A.41), α^{B} and β^{B} are the parameters of the Beverton-Holt stock recruit relationship, and α^{R} and β^{R} are the parameters of the Ricker stock recruit relationship, derived from user-defined parameters steepness (h) and R_0, where $\alpha^{\mathrm{B}} = \frac{4h}{(1-h)\phi_0}$, $\beta^{\mathrm{B}} = \frac{5h-1}{(1-h)B_0}$ and $\alpha^{\mathrm{R}} = \frac{(5h)^{1.25}}{\phi_0}$, $\beta^{\mathrm{R}} = \frac{\log(5h)}{B_0}$, and where ϕ_0 and B_0 are unfished spawners-per-recruit and unfished spawning biomass, respectively, and B_0 is derived from $R_0\phi_0$.

Initial spawning depletion is then given by:

$$\frac{B_1}{B_0}, \tag{B.7}$$

where

$$B_1 = \sum_a^A w_a m_a N_{1,a}. \tag{B.8}$$

It should be apparent from Equations B.5, B.6 and A.41 that in data-limited cases, the parameters R_0 and F_f^{eq} are highly confounded since both inform population size through R^{eq}. In such cases, it is strongly recommended to treat C^{eq} as a major axis of uncertainty as it will be a strong determinant of initial depletion and stock size. We also note that the parameters (q_s) scaling survey observations to vulnerable biomass (Equation B.19) will also be confounded with R_0 and F_f^{eq}.

B.4.3 DYNAMICS EQUATIONS

After setting the equilibrium population age distribution in the first year of the model, the population abundance $N_{y,a}$ in subsequent years is

$$N_{y,a} = \begin{cases} R_y & a = 1 \\ N_{y-1,a-1}\, e^{-Z_{y-1,a-1}} & a = 2, \ldots, A-1, \\ N_{y-1,a-1}\, e^{-Z_{y-1,a-1}} + N_{y-1,a}\, e^{-Z_{y-1,a}} & a = A, \end{cases} \tag{B.9}$$

where R_y is the annual recruitment (Equation B.10) and A is the maximum age, treated as a plus-group.

Annual recruitment R_y is modelled as:

$$R_y = \begin{cases} \dfrac{\alpha^{\mathrm{B}} B_{y-1}}{1 + \beta^{\mathrm{B}} B_{y-1}} \exp(\varepsilon_y - 0.5\tau^2) & \text{Beverton-Holt} \\ \alpha^{\mathrm{R}} B_{y-1} \exp(-\beta^{\mathrm{R}} B_{y-1}) \exp(\varepsilon_y - 0.5\tau^2) & \text{Ricker,} \end{cases} \tag{B.10}$$

where ε_y are recruitment deviates (here in log space vs. in natural space in Appendix A) and τ is the standard deviation of the deviates, and B_y is annual spawning biomass, given by:

$$B_y = \sum_a^A w_a m_a N_{y,a}, \tag{B.11}$$

where $m_{y,a}$ and $w_{y,a}$ are the maturity-at-age and weight-at-age, respectively (Equations A.1 and A.2).

B.4.4 CATCH-AT-AGE

If the model is fit to age-composition data, the catch (in numbers) C^N at age for fleet f is

$$C^N_{y,a,f} = \frac{v_{a,f}F_{y,f}}{Z_{y,a}}N_{y,a}(1-\exp(-Z_{y,a}). \tag{B.12}$$

$F_{y,f}$ can be estimated as parameters (or solved iteratively to match the observed catch).

B.4.5 CATCH-AT-LENGTH

If the model is fit to length-composition data, the catch-at-length is calculated assuming a normally distributed length-at-age $P(\ell, a)$, where

$$C^N_{y,\ell,f} = \sum_a C^N_{y,a,f}P(\ell|a), \tag{B.13}$$

and

$$P(\ell|a) = \begin{cases} \phi^L(L'_{\ell+1}) & \ell = 1 \\ \phi^L(L'_{\ell+1}) - \phi(L'_\ell) & \ell = 2,\ldots,L-1, \\ 1-\phi^L(L'_\ell) & \ell = L, \end{cases} \tag{B.14}$$

where L'_ℓ is the length at the lower boundary of length bin ℓ and $\phi^L(L'_\ell)$ is the cumulative distribution function of a normal variable with mean $\tilde{L}_{y,a}$ (the expected mean length at age a) and standard deviation $\tilde{L}_{y,a} \times CV^L$, where CV^L is the coefficient of variation in mean length at age.

The fitted catch in weight $C_{y,f}$ is

$$C_{y,f} = \sum_a C^N_{y,a,f}w_{y,a}. \tag{B.15}$$

The mean length of the catch $\bar{L}_{y,f}$ is

$$\bar{L}_{y,f} = \frac{\sum_\ell L_\ell C^N_{y,\ell,f}}{\sum_\ell C^N_{y,\ell,f}}, \tag{B.16}$$

where L_ℓ is the midpoint of the length bin ℓ.

The proportion of the catch-at-age is

$$p_{y,a,f} = \frac{C^N_{y,a,f}}{\sum_a C^N_{y,a,f}}. \tag{B.17}$$

The proportion of the catch-at-length is

$$p_{y,\ell,f} = \frac{C^N_{y,\ell,f}}{\sum_\ell C^N_{y,\ell,f}}. \tag{B.18}$$

B.5 SURVEY

If the s^{th} survey is biomass-based, then the survey value $I_{y,s}$ is calculated as

$$I_{y,s} = q_s \sum_a v_{a,s} N_{y,a} w_{y,a}, \tag{B.19}$$

where q is the scaling coefficient and s indexes each survey.

If the survey is abundance-based, then

$$I_{y,s} = q_s \sum_a v_{a,s} N_{y,a}. \tag{B.20}$$

B.6 LIKELIHOODS

If the model is conditioned on catch, and fishing mortality rates are estimated parameters, then the log-likelihood component Λ_1 of the catch is

$$\Lambda_1 = \sum_f \left[\lambda_f^C \sum_y \left(-\log(0.01) - \frac{[\log(C_{y,f}^{\text{obs}}) - \log(C_{y,f}^{\text{pred}})]^2}{2 \times 0.01^2} \right) \right], \tag{B.21}$$

where obs and pred indicate observed and predicted quantities, respectively, and λ are likelihood weights. With a very small standard deviation for the catch likelihood (0.01) relative to the variance in other likelihood components, the predicted catch should match the observed catch almost perfectly.

The log-likelihood component Λ_2 of survey data is

$$\Lambda_2 = \sum_s \left[\lambda_s^I \sum_y \left(-\log(\sigma_{y,s}) - \frac{[\log(I_{y,s}^{\text{obs}}) - \log(I_{y,s}^{\text{pred}})]^2}{2\sigma_{y,s}^2} \right) \right]. \tag{B.22}$$

The log-likelihood component Λ_3 of catch-at-age data is

$$\Lambda_3 = \sum_f \lambda_f^A \left[\sum_y O_{y,f}^A \sum_a p_{y,a,f}^{\text{obs}} \log(p_{y,a,f}^{\text{pred}}) \right], \tag{B.23}$$

where O^A is the annual sample sizes for the age compositions.

The log-likelihood component Λ_4 of catch-at-length data is

$$\Lambda_4 = \sum_f \lambda_f^L \left[\sum_y O_{y,f}^L \sum_\ell p_{y,\ell,f}^{\text{obs}} \log(p_{y,\ell,f}^{\text{pred}}) \right], \tag{B.24}$$

where O^L is the annual sample sizes for the length compositions.

The log-likelihood component Λ_5 of observed mean lengths in the catch is

$$\Lambda_5 = \sum_f \lambda_f^{\bar{L}} \left[\sum_y \left(-\log(\omega_f) - \frac{[\bar{L}_{y,f}^{\text{obs}} - \bar{L}_{y,f}^{\text{pred}}]^2}{2\omega_f^2} \right) \right], \tag{B.25}$$

where ω_f is the standard deviation of mean lengths.

The log-likelihood component Λ_6 of annual estimated recruitment deviates ε_y in log space is

$$\Lambda_6 = \Sigma_y \left(-\log(\tau) - \frac{\varepsilon_y^2}{2\tau^2} \right), \tag{B.26}$$

where τ is the standard deviation of recruitment deviates.

The log-likelihood component Λ_7 of the equilibrium catch is

$$\Lambda_7 = \sum_f \lambda_f^C \left(-\log(0.01) - \frac{[\log(C_f^{\text{eq,obs}}) - \log(C_f^{\text{eq,pred}})]^2}{2 \times 0.01^2} \right). \tag{B.27}$$

The total log-likelihood LL to be maximized is

$$\mathrm{LL} = \sum_{i=1}^{7} \Lambda_i. \tag{B.28}$$

APPENDIX C. DEFAULT SLOTS

The following are OM values or “slots” that we suggest by default be “turned off” or set to default values for the sake of simplicity. On a case-by-case basis, an analyst might choose to set some of these slots to other values or include other values as part of a robustness set. However, unless specified otherwise, the following slots can be assumed to be specified as follows.

C.1 DEFAULT STOCK SLOTS

C.1.1 M2

(Optional) Natural mortality rate at age. Vector of length maxage. Positive real number.

By default, we will not specify natural mortality rate by age.

```
stock_default@M2
#> numeric(0)
```

C.1.2 MEXP

Exponent of the Lorenzen function assuming an inverse relationship between M and weight. Uniform distribution lower and upper bounds. Real numbers <= 0.

By default, we will not use this functionality.

```
stock_default@Mexp <- c(NA_real_, NA_real_)
```

C.1.3 MSD

Inter-annual variability in natural mortality rate expressed as a coefficient of variation. Uniform distribution lower and upper bounds. Non-negative real numbers.

By default, we will assume that natural mortality is not time-varying, although this might be an important slot to explore as part of a robustness set for some stocks.

```
stock_default@Msd
#> numeric(0)
```

C.1.4 PERIOD

(Optional) Period for cyclical recruitment pattern in years. Uniform distribution lower and upper bounds. Non-negative real numbers

By default, we will not assume a cyclical recruitment pattern.

```
stock_default@Period
#> numeric(0)
```

C.1.5 AMPLITUDE

(Optional) Amplitude in deviation from long-term average recruitment during recruitment cycle (e.g., a range from 0 to 1 means recruitment decreases or increases by up to 100% each cycle). Uniform distribution lower and upper bounds. 0 < Amplitude < 1.

By default, we will not assume a cyclical recruitment pattern.

```r
stock_default@Amplitude
#> numeric(0)
```

C.1.6 KSD

Inter-annual variability in growth parameter k expressed as coefficient of variation. Uniform distribution lower and upper bounds. Non-negative real numbers.

By default, we will assume that growth is not time-varying.

```r
stock_default@Ksd
#> numeric(0)
```

C.1.7 LINFSD

Inter-annual variability in maximum length expressed as a coefficient of variation. Uniform distribution lower and upper bounds. Non-negative real numbers.

By default, we will assume that growth is not time-varying.

```r
stock_default@Linfsd
#> numeric(0)
```

C.1.8 SIZE_AREA_1

The size of area 1 relative to the total area (area 1 + area 2). Set to 0.5 to approximate single area model. Uniform distribution lower and upper bounds. Positive real numbers.

We will set this to 0.5 to mimic a single area model.

```r
stock_default@Size_area_1 <- c(0.5, 0.5)
```

C.1.9 FRAC_AREA_1

The fraction of the unfished biomass in area 1 relative to the total area (area 1 + area 2). Uniform distribution lower and upper bounds. Positive real numbers.

We will set this to 0.5 to mimic a single area model.

```r
stock_default@Frac_area_1 <- c(0.5, 0.5)
```

C.1.10 PROB_STAYING

The probability of individuals in area 1 remaining in area 1 over the course of one year. Uniform distribution lower and upper bounds. Positive fraction.

We will set this to 0.5 to mimic a single area model.

```r
stock_default@Prob_staying <- c(0.5, 0.5)
```

C.1.11 SRREL

Type of stock-recruit relationship. Single integer value, switch (1) Beverton-Holt (2) Ricker.

We will use a Beverton-Holt stock-recruit relationship as has been used in most BC groundfish assessments and is assumed in the estimates of steepness defined in this operating model.

```r
stock_default@SRrel <- 1L
```

C.1.12 FDISC

Fraction of discarded fish that die. Uniform distribution lower and upper bounds.

As described below, provisionally, we suggest setting up the OM such that all discards are considered part of the catch. Since the discard rate will be set to zero, this parameter will have no effect. For specific OM scenarios it could be explored.

```r
stock_default@Fdisc <- c(1, 1)
```

C.2 DEFAULT FLEET SLOTS

C.2.1 EFFYEARS

Years representing join-points (vertices) of time-varying effort. Vector. Non-negative real numbers.

By default, we will populate this from the SRA model.

```r
fleet_default@EffYears
#> numeric(0)
```

C.2.2 EFFLOWER

Lower bound on relative effort corresponding to EffYears. Vector. Non-negative real numbers.

By default, we will populate this from the SRA model.

```r
fleet_default@EffLower
#> numeric(0)
```

C.2.3 EFFUPPER

Upper bound on relative effort corresponding to EffYears. Vector. Non-negative real numbers.

By default, we will populate this from the SRA model.

```r
fleet_default@EffUpper
#> numeric(0)
```

C.2.4 ESD

Additional inter-annual variability in fishing mortality rate. Uniform distribution lower and upper bounds. Non-negative real numbers.

We will assume that fishing mortality is not time-varying by default.

```r
fleet_default@Esd <- c(0, 0)
```

C.2.5 QINC

Average percentage change in fishing efficiency (applicable only to forward projection and input controls). Uniform distribution lower and upper bounds. Non-negative real numbers.

We will assume that fishing efficiency is not time-varying by default.

```
fleet_default@qinc <- c(0, 0)
```

C.2.6 QCV

Inter-annual variability in fishing efficiency (applicable only to forward projection and input controls). Uniform distribution lower and upper bounds. Non-negative real numbers.

We will assume that fishing efficiency is not time-varying by default.

```
fleet_default@qcv <- c(0, 0)
```

C.2.7 SELYEARS

(Optional) Years representing join-points (vertices) at which historical selectivity pattern changes. Vector. Positive real numbers.

We will not use this functionality by default.

```
fleet_default@SelYears
#> numeric(0)
```

C.2.8 ABSSELYEARS

(Optional) Calendar years corresponding with SelYears (e.g., 1951, rather than 1), used for plotting only. Vector (of same length as SelYears). Positive real numbers.

We will not use this functionality by default.

```
fleet_default@AbsSelYears
#> numeric(0)
```

C.2.9 L5LOWER

(Optional) Lower bound of L5. Vector. Non-negative real numbers.

We will not use this functionality by default.

```
fleet_default@L5Lower
#> numeric(0)
```

C.2.10 L5UPPER

(Optional) Upper bound of L5. Vector. Non-negative real numbers.

We will not use this functionality by default.

```
fleet_default@L5Upper
#> numeric(0)
```

C.2.11 LFSLOWER

(Optional) Lower bound of LFS. Vector. Non-negative real numbers.

We will not use this functionality by default.

```
fleet_default@LFSLower
#> numeric(0)
```

C.2.12 LFSUPPER

(Optional) Upper bound of LFS. Vector. Non-negative real numbers.

We will not use this functionality by default.

```
fleet_default@LFSUpper
#> numeric(0)
```

C.2.13 VMAXLOWER

(Optional) Lower bound of Vmaxlen. Vector. Fraction.

We will not use this functionality by default.

```
fleet_default@VmaxLower
#> numeric(0)
```

C.2.14 VMAXUPPER

(Optional) Upper bound of Vmaxlen. Vector. Fraction.

We will not use this functionality by default.

```
fleet_default@VmaxUpper
#> numeric(0)
```

C.2.15 MPA

(Optional) Matrix specifying spatial closures for historical years.

We will not use this functionality by default.

```
fleet_default@MPA
#> <0 x 0 matrix>
```

C.2.16 DR

Discard rate: the fraction of caught fish that are discarded. Uniform distribution lower and upper bounds. Fraction.

Since all the discards are included in the catch, we will set this to 0.

C.2.17 LR5

Shortest length corresponding to 5% retention. Uniform distribution lower and upper bounds.

This should have no impact on the results because we are assuming 100% discard mortality and reporting of all catch by default.

```
fleet_default@LR5
#> numeric(0)
```

C.2.18 LFR

Shortest length that is fully retained. Uniform distribution lower and upper bounds.

This should have no impact on the results because we are assuming 100% discard mortality and reporting of all catch by default.

```
fleet_default@LFR
#> numeric(0)
```

C.2.19 RMAXLEN

The retention of fish at `Linf`. Uniform distribution lower and upper bounds.

This should have no impact on the results because we are assuming 100% discard mortality and reporting of all catch by default.

```
fleet_default@Rmaxlen <- c(1, 1)
```

C.2.20 VMAXLEN

The vulnerability of fish at `Linf`. Uniform distribution lower and upper bounds. Fraction.

We will set this to 1 by default. Setting this to a value lower than 1 would be assuming dome-shaped selectivity.

```
fleet_default@Vmaxlen <- c(1, 1)
```

C.3 DEFAULT OBSERVATION SLOTS

C.3.1 CREFBIASCV

Log-normal CV for sampling persistent bias in MSY. Uniform distribution lower and upper bounds.

The provisional MPs do not make use of this slot.

```
obs_default@Crefbiascv
#> numeric(0)
```

C.4 DEFAULT IMPLEMENTATION SLOTS

C.4.1 TACFRAC

Mean fraction of Total Allowable Catch (TAC) taken. Uniform distribution lower and upper bounds. Positive real number.

We will not use this functionality by default.

```
imp_default@TACFrac
#> [1] 1 1
```

C.4.2 TACSD

Log-normal coefficient of variation in the fraction of TAC taken. Uniform distribution lower and upper bounds. Non-negative real numbers.

We will not use this functionality by default.

```
imp_default@TACSD
#> [1] 0 0
```

C.4.3 SIZELIMFRAC

The real minimum size that is retained expressed as a fraction of the size. Uniform distribution lower and upper bounds. Positive real number.

We will not use this functionality by default.

```
imp_default@SizeLimFrac
#> [1] 1 1
```

C.4.4 SIZELIMSD

Log-normal coefficient of variation controlling mismatch between a minimum size limit and the real minimum size retained. Uniform distribution lower and upper bounds. Non-negative real numbers.

We will not use this functionality by default.

```
imp_default@SizeLimSD
#> [1] 0 0
```

C.4.5 CAA_NSAMP

Number of catch-at-age observations per time step. Uniform distribution lower and upper bounds.

This slot will not be relevant in the main operating model since none of the proposed MPs make use of age data. It could be re-enabled if evaluating the value of information inherent in ageing fish and using MPs that make use of them.

```
obs_default@CAA_nsamp
#> numeric(0)
```

C.4.6 CAA_ESS

Effective sample size (independent age draws) of the multinomial catch-at-age observation error model. Uniform distribution lower and upper bounds.

This slot will not be relevant in operating models unless MPs make use of age-composition data.

```
obs_default@CAA_ESS
#> numeric(0)
```

C.4.7 CAL_NSAMP

Number of catch-at-length observations per time step. Uniform distribution lower and upper bounds.

This slot will not be relevant in operating models unless MPs make use of length-composition data.

```r
obs_default@CAL_nsamp
#> numeric(0)
```

C.4.8 CAL_ESS

Effective sample size (independent length draws) of the multinomial catch-at-length observation error model. Uniform distribution lower and upper bounds.

This slot will not be relevant in operating models unless MPs make use of length-composition data.

```r
obs_default@CAL_ESS
#> numeric(0)
```

APPENDIX D. DATA-LIMITED MANAGEMENT PROCEDURES

Here we include an overview of MPs that may be used as part of this framework. The MPs are a combination of those from the literature and ones that have been used in recent assessments for other BC groundfish. We focus on "output-controlled" MPs—MPs that recommend a TAC—for consistency with the management approach for BC groundfish. The list is not exhaustive and will likely grow or be adapted over time.

We note that the notation between MPs is not always consistent (e.g., different symbols may be used for the slope of a relative abundance index), but we have aimed to maintain consistency with the DLMtool documentation or the primary source publications wherever possible.

D.1 CONSTANT-CATCH MANAGEMENT PROCEDURES

MPs in this category set a constant catch, often based on some average historical catch. This type of MP is static and therefore does not incorporate feedback between subsequent stock status and MP recommendations. Nonetheless, these MPs are simple, may represent only a minor modification to the status quo, and in some circumstances, may satisfy performance metrics.

D.1.1 CC_HIST20: AVERAGE HISTORICAL CATCH

This is a simple management procedure in which the mean historical catch is calculated and used to set a constant TAC (with "CC" indicating constant catch). The TAC in year y is calculated as:

$$\mathrm{TAC}_y = \frac{\sum_{i=1}^{n} C_i}{n}, \tag{D.1}$$

where n is the number of historical years, and C_i is the catch in historical year i. For our purposes in BC, we suggest using average catch from the last 20 years, which encompasses a time period after the implementation of trawl ITQs in 1996. We denote this MP "CC_hist20":

$$\mathrm{TAC}_y = \frac{\sum_{i=n-20+1}^{n} C_i}{20}. \tag{D.2}$$

For specific stocks, analysts might consider other starting years that are relevant to that stock.

D.1.2 CC: CONSTANT CATCH

For these MPs, TAC is calculated as some fraction of the average historical catch over the last 5 years (Geromont and Butterworth 2015a):

$$\mathrm{TAC}_y = \alpha \frac{\sum_{i=n-5+1}^{n} C_i}{5}, \tag{D.3}$$

where α is some value greater than zero. The TAC is constant for all future projections. We suggest the following provisional "CC" MPs, which differ in their value of α:

- CC1.2: $\alpha = 1.2$
- CC1.1: $\alpha = 1.1$
- CC1.0: $\alpha = 1.0$
- CC0.9: $\alpha = 0.9$

- CC0.8: $\alpha = 0.8$
- CC0.7: $\alpha = 0.7$
- CC0.6: $\alpha = 0.6$

D.2 INDEX-SLOPE AND INDEX-RATIO MANAGEMENT PROCEDURES

Management procedures in this category make a TAC recommendation based on the change in a relative abundance index over time. The term "slope" is used since many MPs in this category fit a linear regression to the relative abundance index (usually in log space) and make a recommendation that generally decreases the TAC if the slope is negative and increases the TAC if the slope is positive. Some of the MPs are based on ratios of relative abundance in certain years.

D.2.1 IRATIO: MEAN INDEX RATIO

This MP adjusts the TAC by a ratio, α, with the numerator being the mean index in the most recent two years of the time series and the denominator being the mean index in the three years prior to those in the numerator. This MP is based on Method 3.2 used by ICES for data-limited stocks (ICES 2012; Jardim et al. 2015). The TAC is calculated as:

$$\mathrm{TAC}_y = \alpha C_{y-1}, \tag{D.4}$$

$$\alpha = \frac{I_{y-1} + I_{y-2}}{2} \Bigg/ \frac{I_{y-3} + I_{y-4} + I_{y-5}}{3}, \tag{D.5}$$

where C_{y-1} is the catch from the previous year and α is the ratio of the mean index in the most recent 2 years of the time series and the mean index in 3–5 years before current time. Due to the biennial nature of most BC surveys for any one area, we propose an alternative version of this MP, "Iratio2," which calculates the ratio with the last 2 years of available survey observations in the numerator and the last 3 years of available survey observations in the denominator.

D.2.2 GB_SLOPE: GEROMONT AND BUTTERWORTH INDEX SLOPE

This MP adjusts TAC based on previous catch and the trend in a relative abundance index to aim for relatively stable catch rates (Geromont and Butterworth 2015a) (Figure D.1). The TAC is calculated as:

$$\mathrm{TAC}_y = C_{y-1}(1 + \lambda \beta_I), \tag{D.6}$$

where C_{y-1} is catch from the previous year, β_I is the slope of a linear regression of the ln abundance index over the last n years (default of $n = 5$), and λ is a control parameter between 0 and 1 that adjusts how quickly TAC is adjusted based on the slope of the index. The TAC is subject to the following conditions that limit the rate at which the TAC can be adjusted up or down:

- if next TAC > 1.2 last catch, then TAC = 1.2 $\times$ last catch
- if next TAC < 0.8 last catch, then TAC = 0.8 $\times$ last catch.

The default λ value is 1 in DLMtool.

Here we propose adding a version with a lower value of λ ($\lambda = 0.66$), which is therefore less responsive to changes in the relative abundance index. Furthermore, we propose extending the linear regression to encompass 6 or 8 years, due to the biennial nature of the synoptic trawl surveys, thereby encompassing a constant number of years with data. We denote these MPs as "GB_slope_6_1," "GB_slope_6_0.66," "GB_slope_8_1," and "GB_slope_8_0.66," where the numbers indicate the number of years and λ value, respectively. The number of years and the λ parameter could be explored as tuning parameters for specific stocks.

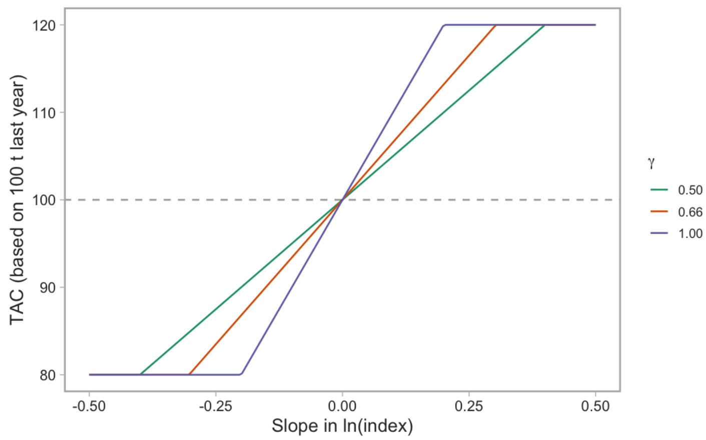

Figure D.1. Illustration of GB_slope across different values of λ, the gain/smoothing parameter that controls how sensitive the TAC recommendations are to changes in the relative abundance index.

D.2.3 ISLOPE: INDEX SLOPE TRACKING

These MPs incrementally adjust the TAC in an attempt to maintain a constant relative abundance index (Figure D.2). The MPs are similar to "GB_slope" with the addition of a parameter that determines the TAC in the first projection year and different choices of the λ parameter. The TAC is calculated as:

$$\mathrm{TAC}_y = \mathrm{TAC}^*(1 + \lambda\beta_I), \tag{D.7}$$

where, in the first projection year, TAC^* is $(1 - x)$ multiplied by the mean catch from the last 5 historical years. In subsequent years, TAC^* is the TAC from the previous year. Again, λ is a gain or smoothing parameter, and βI is the slope of the ln abundance index over the past n years (we have set $n = 6$).

There are four variants of this procedure as described in Geromont and Butterworth (2015b):

- Islope1: The least biologically precautionary, with $\lambda = 0.4$ and $x = 0.2$
- Islope2: Increasingly biologically precautionary, with $\lambda = 0.4$ and $x = 0.3$
- Islope3: Increasingly biologically precautionary, with $\lambda = 0.4$ and $x = 0.4$
- Islope4: The most biologically precautionary, with $\lambda = 0.2$ and $x = 0.4$

Because of the x values, all 4 versions start TAC at 60–80% of recent average catch. For our BC groundfish fisheries, we do not have an a priori expectation that stocks are currently being overfished. We therefore propose the following set that start TAC at 80% or 100% of average catch over the last 5 historical years:

- Islope0.4_100: $\lambda = 0.4$ and $x = 0$
- Islope0.4_80: $\lambda = 0.4$ and $x = 0.2$
- Islope0.2_100: $\lambda = 0.2$ and $x = 0$

- Islope0.2_80: $\lambda = 0.2$ and $x = 0.2$

We have additionally modified this MP to add a maximum proportional increase in TAC from one year to the next of 1.2.

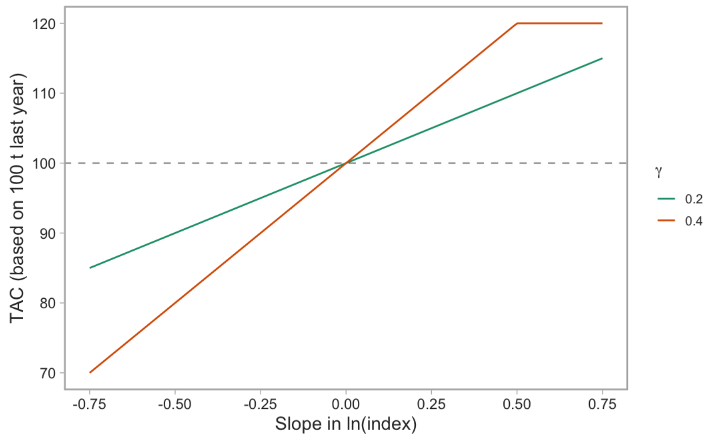

Figure D.2. Illustration of the Islope MPs across 2 values of λ. The x parameter only affects the TAC in the initial projection year and is therefore not shown.

D.2.4 IDX: INDEX-BASED MP FROM COX ET AL. (2020)

This MP was used in the rebuilding plan for outside Yelloweye Rockfish in BC (Cox et al. 2020) (Figure D.3). The MP assigns TAC based on:

$$\mathrm{TAC}_y = \begin{cases} \mathrm{TAC}_{\mathrm{Floor}}, & \text{if } \Delta I_y \leq \delta_{\mathrm{min}} \\ (1 + \Delta I_y)\mathrm{TAC}_{y-1}, & \text{if } \delta_{\mathrm{min}} < \Delta I_y \leq \delta_{\mathrm{max}} \\ (1 + \delta_{\mathrm{max}})\mathrm{TAC}_{y-1}, & \text{if } \Delta I_y > \delta_{\mathrm{max}}, \end{cases} \tag{D.8}$$

where δ_{min} is the most negative drop allowed in the relative biomass index before the fishery is closed that year (by default assuming $\mathrm{TAC}_{\mathrm{Floor}}$ is 20% of the average catch from the last 5 historical years) and ΔI_y is the current index value divided by the last observed index value minus 1. We set $\delta_{\mathrm{min}} = -0.5$ as in (Cox et al. 2020), but this could be tuned for individual stocks. The maximum increase in TAC is capped at $\delta_{\mathrm{max}} = 0.25$ by default.

This MP can be additionally smoothed:

$$\mathrm{TAC}_y = \lambda \cdot \mathrm{TAC}_y + (1 - \lambda)\mathrm{TAC}_{y-1}, \tag{D.9}$$

where λ controls the degree of smoothing and can range between 0 and 1. Cox et al. (2020) used $\lambda = 0.5$. We define these MPs for DLMtool as “IDX” ($\delta_{\mathrm{min}} = -0.5$, $\delta_{\mathrm{max}} = 0.25$) and “IDX_smooth” (same as IDX but with $\lambda = 0.5$ to split the difference between the potentially proposed TAC and the one previously recommended).

$\mathrm{TAC}_{\mathrm{Floor}}$ would ideally be set on a stock-specific basis to a reasonable value required to maintain other groundfish fisheries. Provisionally, in the absence of a stock-specific floor value, “IDX” will set the floor to 20% of the average catch from the last 5 historical years.

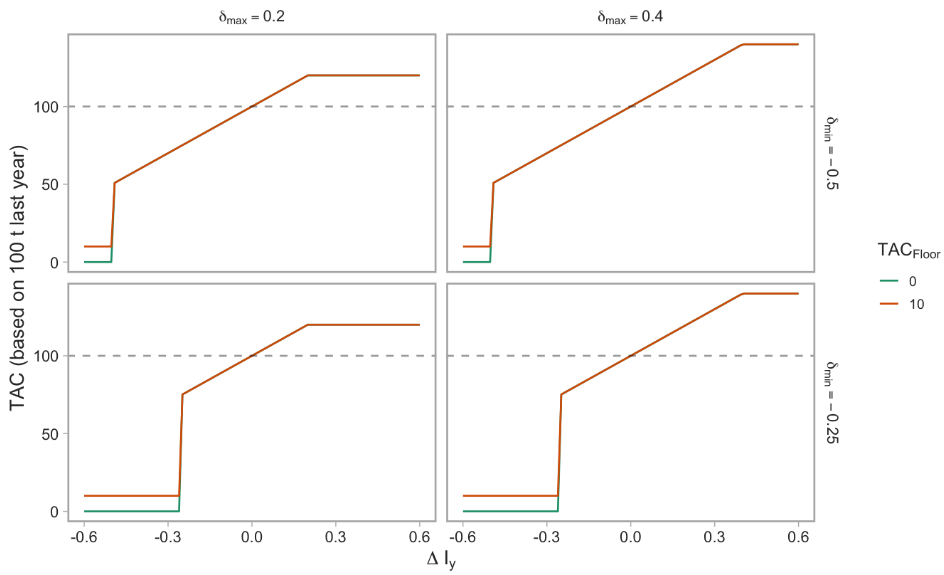

Figure D.3. Illustration of the IDX MPs across two values of δ_{min} and δ_{max}, also illustrating two values of TAC_{Floor} .

D.3 INDEX-TARGET MANAGEMENT PROCEDURES

MPs in this category aim to maintain a relative abundance index at some reference level. Typically, this reference level is set based on an assumption of what the index would be if the stock was at B_{MSY}. Since it would be challenging to apply this to real data in a manner consistent with this assumption, we instead propose modifying this MP to set the target index level to a historical index level, which would be chosen on a stock-by-stock basis. Provisionally, we set the reference level to the mean index value from the 10 years prior to the projection period. Analysts, in consultation with managers and stakeholders, may choose to adjust this reference period or choose some alternative reference index value. The performance of choices of reference index can be evaluated in the closed-loop simulation.

D.3.1 IT: ITERATIVE INDEX TARGET

These are index target MPs where the TAC is modified according to current index levels (the mean index over the last 5 years) relative to a target level. Traditionally the target level is set to the index value at B_{MSY}, subject to observation error. As noted above, we modified this MP to set the target level as the average index value from the 10 year period prior to the projection period.

The TAC is calculated as:

$$\mathrm{TAC}_y = C_{y-1} I_\delta, \tag{D.10}$$

where C_{y-1} is the catch from the previous year and I_δ is the ratio of the mean index over the past five years to the reference index level.

There are two variants of this procedure:

- IT5 where the maximum annual changes to the TAC are 5%
- IT10 where the maximum annual changes to the TAC are 10%

We denote our historical variants “IT5_hist” and “IT10_hist.”

D.3.2 ITARGET: INCREMENTAL INDEX TARGET

The “Itarget” MPs (Geromont and Butterworth 2015a, 2015b) incrementally adjust the TAC based on reference catch and abundance index values (Figure D.4).

If $I_{\mathrm{recent}} \geq I_0$ the TAC is calculated as:

$$\mathrm{TAC}_{y+1} = \mathrm{TAC}^* \left[w + (1-w) \left(\frac{I_{\mathrm{recent}} - I_0}{I_{\mathrm{target}} - I_0} \right) \right], \tag{D.11}$$

otherwise:

$$\mathrm{TAC}_{y+1} = w \cdot \mathrm{TAC}^* \left(\frac{I_{\mathrm{recent}}}{I_0} \right)^2, \tag{D.12}$$

These calculations depend on:

- I_{recent}, the average index over the most recent five years;
- I_{ave}, the average index over the 10 years prior to the projection period;
- C_{ave}, the average historical catch over the last five years of the historical period;
- λ, the fraction of I_{ave} below which future TACs are reduced quadratically (Figure D.4);
- δ, the fraction of I_{ave} defining the target index value;
- x, the proportional difference between the future catch and C_{ave}; and
- w, a smoothing parameter that defines the “steepness” of the adjustment slope.

then:

- $I_0 = \lambda I_{\mathrm{ave}}$;
- $I_{\mathrm{target}} = \delta I_{\mathrm{ave}}$; and
- $\mathrm{TAC}^* = x C_{\mathrm{ave}}$, the catch target.

Geromont and Butterworth (2015a) and Geromont and Butterworth (2015b) propose a number of configurations of these MPs. We propose starting with the following provisional versions and investigating tuning across a more complete range of parameter values if any of the provisional versions perform well.

We recommend the following provisional base values: $\lambda = 0.2$, $\delta = 1$, $w = 0.5$, $x = 1$. This represents:

- the TAC adjustment decreasing quadratically at 20% of the average index over the last 10 historical years ($\lambda = 0.2$);
- a target index of the average index value over the last 10 years ($\delta = 1$);
- a moderate steepness of the adjustment slope ($w = 0.5$); and
- a target catch equal to the average historical catch over the last five years of the historical period ($x = 1$).

We then recommend varying each of the parameters while holding the others at their default values (Figure D.5):

- Itarget_base: $\lambda = 0.2$, $\delta = 1$, $w = 0.5$, $x = 1$
- Itarget_w0.8: $\lambda = 0.2$, $\delta = 1$, $w = 0.8$, $x = 1$

- Itarget_x0.2: $\lambda = 0.2$, $\delta = 1$, $w = 0.5$, $x = 1.2$
- Itarget_x0.8: $\lambda = 0.2$, $\delta = 1$, $w = 0.5$, $x = .8$
- Itarget_d1.2: $\lambda = 0.2$, $\delta = 1.2$, $w = 0.5$, $x = 1$
- Itarget_d0.8: $\lambda = 0.2$, $\delta = 0.8$, $w = 0.5$, $x = 1$

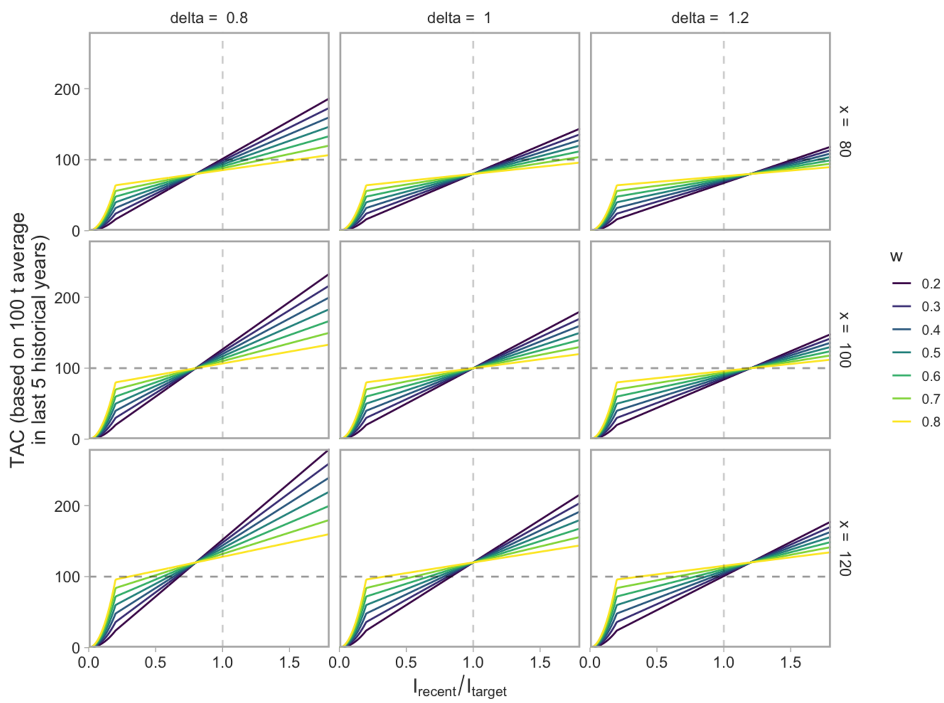

Figure D.4. Illustration of possible Itarget MPs across a range of tuning parameters based on $I_{ave} = 1$, showing the effect of varying w. In all cases, $I_0 = \lambda I_{ave} = 0.2$, below which the TAC is reduced quadratically. Note that delta refers to δ.

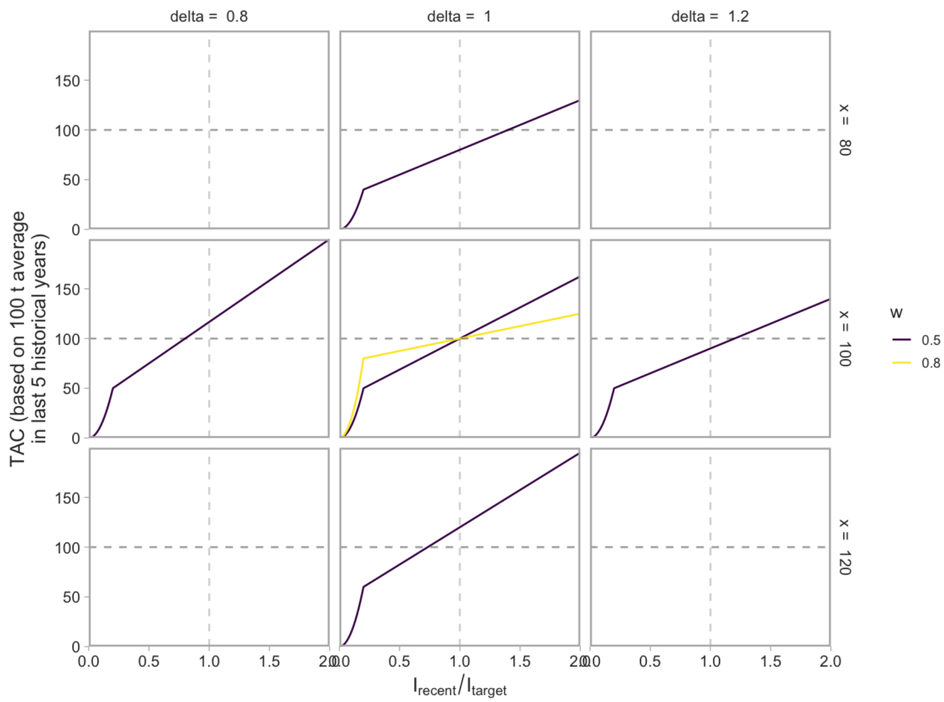

Figure D.5. Illustration of provisional Itarget MPs based on $I_{ave} = 1$, with recommended provisional combinations of parameters. In all cases, $I_0 = \lambda I_{ave} = 0.2$, below which the TAC is reduced quadratically. Note that delta refers to δ.

D.3.3 ITM: INDEX TARGET BASED ON NATURAL MORTALITY RATE

"ITM" is an index-target MP where the TAC is modified according to current index levels with the window defining "current" based on the assumed natural mortality M (Figs D.6).

The MP is defined as:

$$\mathrm{TAC}_y = \mathrm{TAC}_{y-1}\delta I, \tag{D.13}$$

where δI is the ratio of the mean index over $4\,(1/M)^{1/4}$ years to the reference index. The maximum fractional change in TAC is determined by x, defined as $x = \max\left((5 + 25M)/100,\ 0.2\right)$. As in the other reference index MPs, we use an historical period of the index to set the reference index level—provisionally the last 10 years before the projection period.

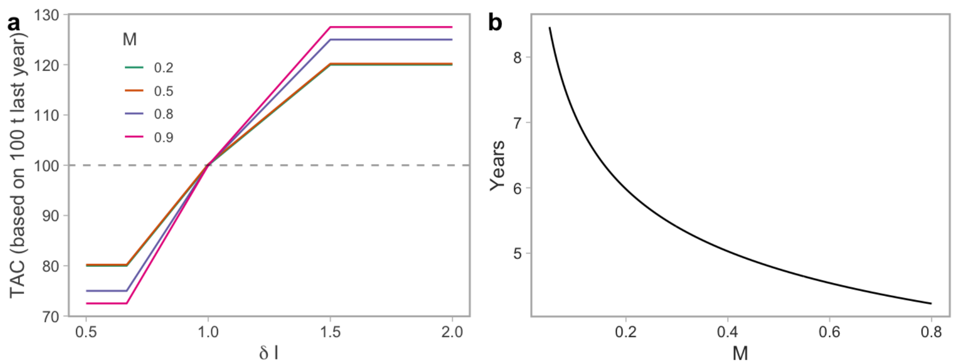

Figure D.6. Illustration of the ITM MP. (a) Illustration of how TAC recommendation is based on natural mortality (M) and δI, the proportional change in the recent relative abundance index compared to a reference level. Note that the effect of M only appears for large values of M—0.2 and 0.5 are overlapping in this figure. (b) Illustration of how M relates to the number of years over which to calculate the mean recent relative abundance index.

D.4 MODEL-BASED MANAGEMENT PROCEDURES

D.4.1 SURPLUS MODEL PRODUCTION

In addition to the empirical MPs described above, we propose including a surplus production model paired with a harvest control rule as a model-based MP. Here, we use the surplus production model implemented in MSEtool (Huynh et al. 2019) and TMB (Kristensen et al. 2016) and based on Fletcher (1978).

Biomass B in year y is calculated as

$$B_y = B_{y-1} + P_{y-1} - C_{y-1}, \tag{D.14}$$

where C_y is the observed catch and P_y is the surplus production given by

$$P_y = \gamma \times \mathrm{MSY} \times \left(\frac{B_y}{K} - \left[\frac{B_y}{K}\right]^n\right), \tag{D.15}$$

where K is the carrying capacity, MSY is the estimated maximum sustainable yield, and n is the parameter that controls the shape of the production curve, and γ is defined as

$$\gamma = \frac{1}{n-1} n^{n/(n-1)}. \tag{D.16}$$

The production parameter n is typically fixed, producing the Schaefer model (Schaefer 1954) with a symmetric productive curve ($B_{\mathrm{MSY}}/K = 0.5$) when $n = 2$. The Fox model is the limiting case of the Fletcher parameterization as $n \to 1$, where

$$K = e \times B_{MSY}, \tag{D.17}$$

$$r = F_{\mathrm{MSY}}, \tag{D.18}$$

and

$$P_y = -e \times \mathrm{MSY} \times \frac{B_y}{K} \times \log\left(\frac{B_y}{K}\right). \tag{D.19}$$

By conditioning the model on observed catch, the predicted index $\hat{I}_y$ is

$$\hat{I}_y = \hat{q}\hat{B}_y \tag{D.20}$$

and the harvest rate is

$$\hat{F}_y = \frac{C_y}{\hat{B}_y}. \tag{D.21}$$

The likelihood of the observed index I_y, assuming a lognormal distribution, is

$$\log(I_y) \sim \mathrm{Normal}(\log[\hat{I}_y], \sigma), \tag{D.22}$$

where σ represents the standard deviation.

We pair these surplus production models with the following harvest control rules (Figure D.7):

1. HCR-4010: Above 40% of estimated B/B_0 (biomass divided by unfished equilibrium spawning biomass), $F_y = \hat{F}_{\mathrm{MSY}}$; at 10% of estimated B/B_0, $F_y = 0$; between 10% and 40%, interpolate the adjustment factor linearly. This is a commonly applied HCR in the fisheries literature and on the US West Coast (e.g., Berger et al. 2019).
2. HCR-8040: Above 80% of estimated B/B_{MSY}, $F_y = \hat{F}_{\mathrm{MSY}}$; at 40% of estimated B/B_{MSY}, $F_y = 0$; between 40% and 80%, interpolate the adjustment factor linearly. Note that this reference point is based on B_{MSY} whereas HCR-4010 is based on B_0. This HCR creates operational control points that mimic the provisional biological upper stock reference and limit reference points from DFO's Sustainable Fisheries Framework (Figure 1), where operational control points define the thresholds of management action (i.e., reducing fishing mortality). We note, however, that operational control points do not necessarily need to match the biological reference points (BRPs) to be consistent with the Sustainable Fisheries Framework. For example, a model may generate biased estimates of B/B_{MSY} and be better paired with operational control points that differ from the BRPs to obtain performance that meet sustainability objectives defined by the BRPs (e.g., Cox et al. 2013).
3. HCR-6040: A slightly less biologically conservative HCR than HCR-8040 (Cox et al. 2013). This HCR does not begin ramping down the TAC from MSY until $B/B_{\mathrm{MSY}} < 0.6$.

We denote these management procedures: "SP8040" (Schaefer surplus production model, 8040 HCR), "SP8040_Fox" (Fox surplus production model, 8040 HCR), etc. We only include the Schaefer surplus production model with the 4010 rule, since the choice of production function should not affect a depletion-based harvest control rule.

We also include the following "meta" HCR rules (e.g., Cox et al. 2020), that could be modified as needed for specific stocks:

- The maximum TAC increase in any one year is 20%.
- The maximum decrease in TAC in any one year is 50%.
- The minimum TAC is 10% of the catch in the last historical year.
- The TAC is not increased unless the MP recommends increasing it by at least 5%.

We include these (modifiable) rules in an MP defined as `SP_gf()` in the gfdlm R package (Anderson et al. 2020b).

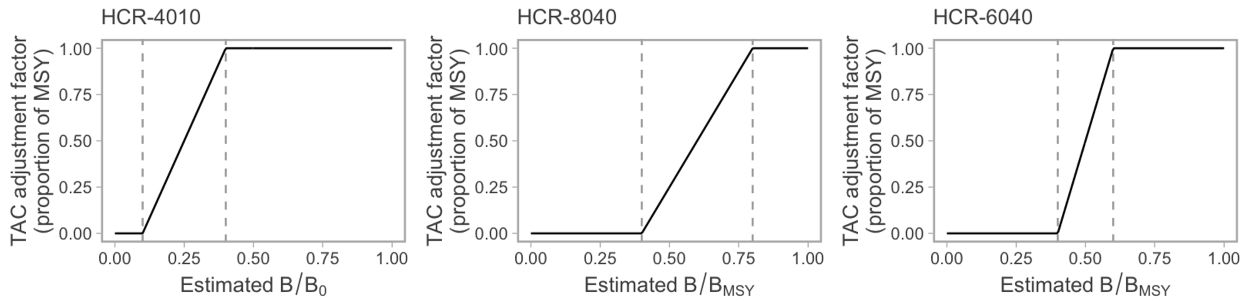

Figure D.7. Proposed provisional harvest control rules to associate with model-based MPs.

APPENDIX E. WEST COAST VANCOUVER ISLAND REX SOLE CASE STUDY

E.1 BACKGROUND

Rex Sole (*Glyptocephalus zachirus*) occur from central Baja California to the western Bering Sea, and are generally distributed throughout coastal BC. They are sometimes referred to as "Witch Sole." Rex Sole are primarily caught in the groundfish bottom-trawl fishery in BC, which implemented 100% at-sea observer coverage and 100% dockside monitoring in 1996.

Rex Sole are caught in BC at depths between 20 m and greater than 1000 m, with most captured between 75 m and 450 m. Rex Sole as large as 58 cm and weighing up to 1.44 kg have been caught in BC waters. The oldest aged individual was a 34 cm female aged 15 years; however, there is very little available ageing data for this species, with the only data coming from the Hecate Strait Multispecies Assemblage Bottom Trawl survey in 1998. Otoliths have been collected on BC surveys from 2000–present although they have not been aged. Rex Sole have been aged up to 29 years (female) in the Gulf of Alaska (Abookire 2006).

Rex Sole are commonly caught in the West Coast Vancouver Island (WCVI) synoptic groundfish bottom trawl survey (Figure E.1 and E.2). While they are caught in the commercial mixed-species bottom trawl fishery (Figures E.3 and E.4), there is no directed fishery or assigned quota for them. They are commonly caught with Pacific Cod (*Gadus macrocephalus*), Dover Sole (*Microstomus pacificus*), English Sole (*Parophrys vetulus*) and Arrowtooth Flounder (*Atheresthes stomias*). Rex Sole have never been formally assessed in BC. However, biomass of Rex Sole in Hecate Strait was estimated in 2012 using research survey data collected between 1984 and 2003 (Fargo 2012). The WCVI synoptic bottom trawl survey has shown a moderate decline in biomass density from ~2008 to 2012 and an increase from 2014 to 2018 (Figures E.1 and E.2). Standardized commercial trawl CPUE (catch per unit effort) increased in the late 1990s, showing a moderate decline from ~2004 to 2010, and a moderate increase until about 2016 (Figure E.2). There are many samples of length, weight, otoliths, and maturity of Rex Sole from the WCVI synoptic survey, but no aged fish (Figure E.5). There are limited biological samples of any kind from the commercial fleet (Figure E.5).

In the US, Rex Sole are caught in the bottom trawl fishery and are considered to form three separate stocks: the Gulf of Alaska stock (McGilliard and Palsson 2017), the Bering Sea/Aleutian Islands stock (Wilderbuer and Nichol 2015), and the West Coast stock from Washington to California (Cope et al. 2015).

We demonstrate the MP Framework here, through a case study of the Rex Sole stock in the WCVI groundfish management area (Area 3CD). This detailed case study is illustrative and is not intended for catch advice at this time. We frame our case study according to the steps outlined in Section 3.

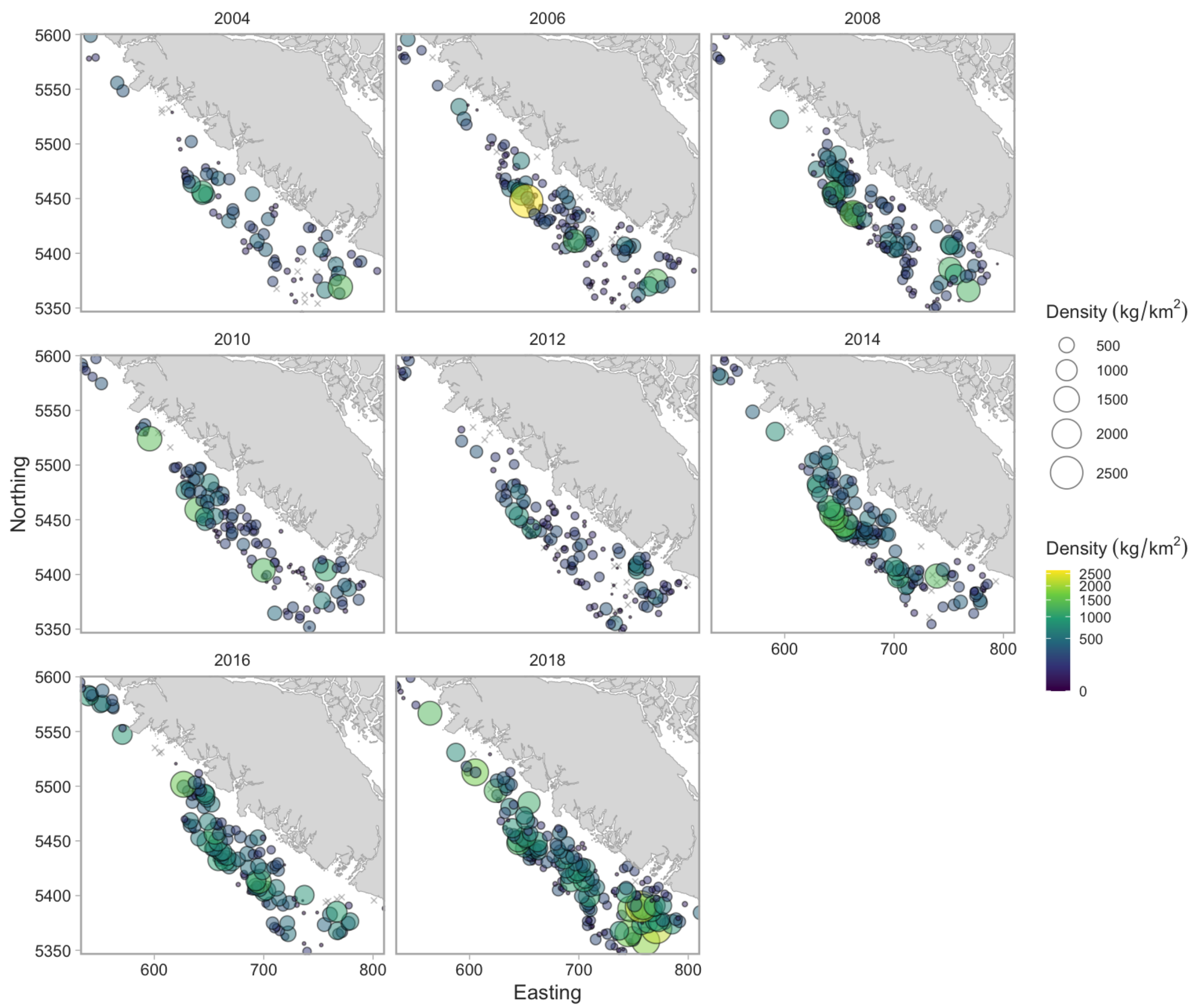

Figure E.1. Individual survey tows for the West Coast Vancouver Island synoptic bottom trawl survey. Light gray crosses indicate survey sets that did not catch Rex Sole. The area and colour of the circles are proportional to the density of Rex Sole for that survey set. Eastings and Northings are for UTM zone 9.

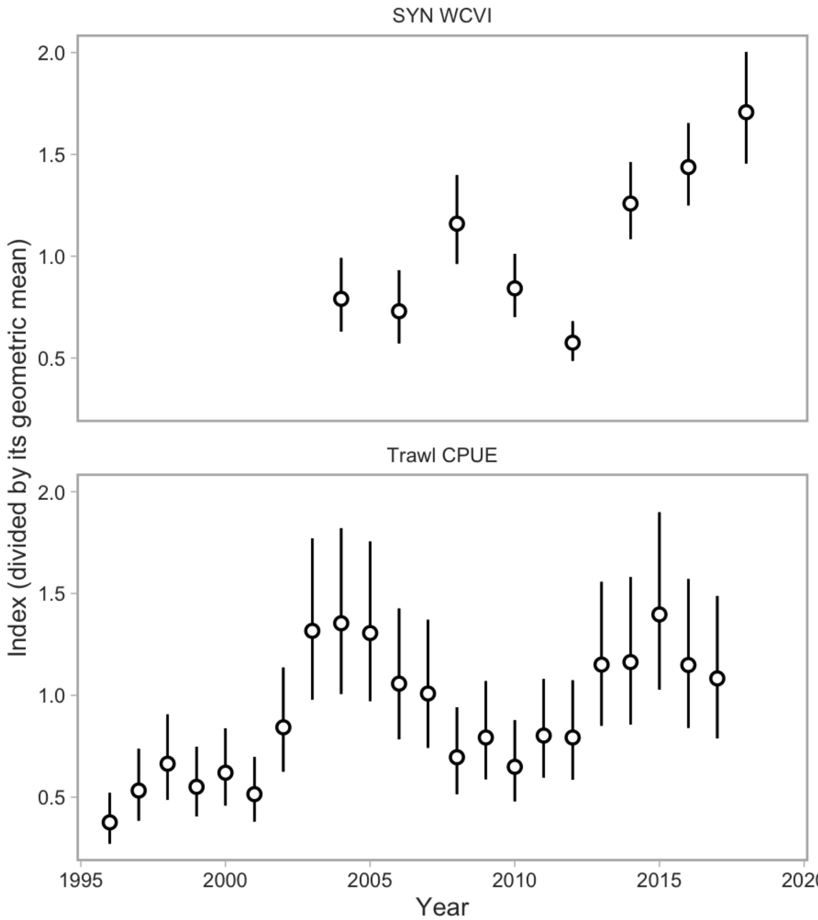

Figure E.2. West Coast Vancouver Island synoptic bottom trawl survey and standardized commercial bottom trawl CPUE for Rex Sole standardized according to the methods in Anderson et al. (2019). Each index is divided by its geometric mean for years 2004–2019 for visualization purposes. Dots represent mean estimates and line segments represent 95% confidence intervals. The index in the top panel is derived from the data shown in Figure E.1 with confidence intervals generated via a stratified bootstrap procedure (for details, see, for example, Anderson et al. 2019).

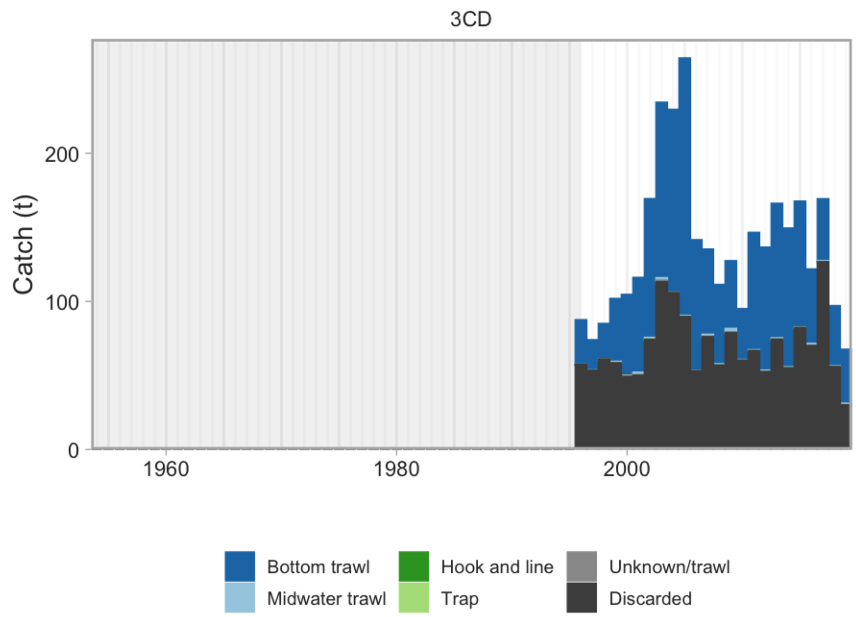

Figure E.3. Commercial catch for Rex Sole. Note that limited catch in the early 1990s is considered unreliable and is not shown. The shaded region prior to 1996 indicates the time period before the at-sea observer program was implemented for bottom and midwater trawl fleets.

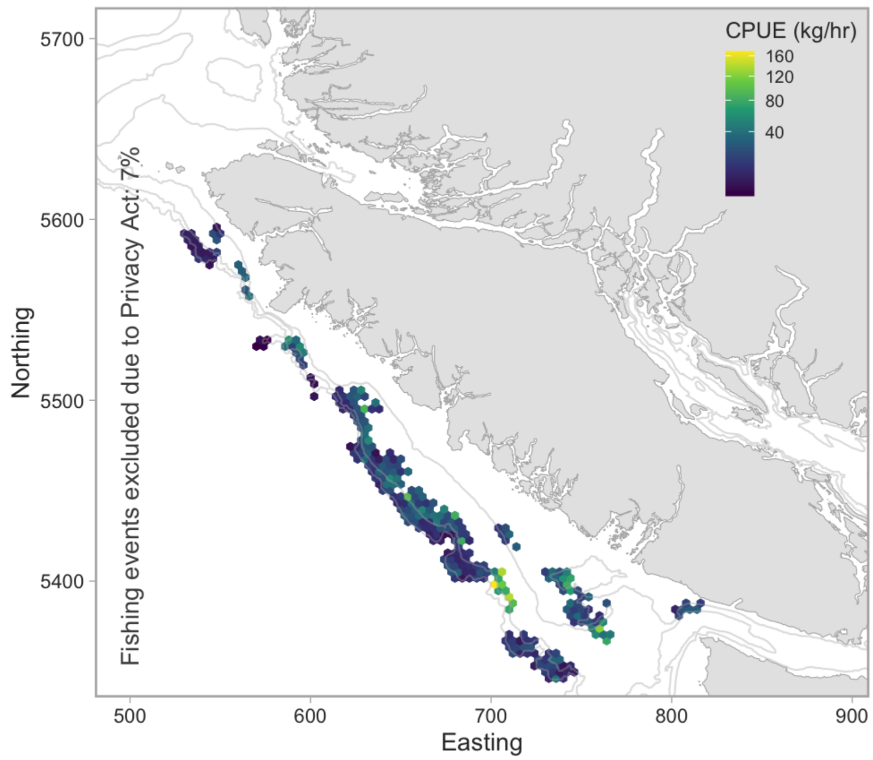

Figure E.4. Spatial commercial trawl CPUE for Rex Sole for tows that caught any Rex Sole from 2013–2019. Cells are 4 km wide and are only shown in cases where there are at least three unique vessels in a given cell to meet federal privacy requirements. CPUE is calculated as the weight of catch (landings plus discards) divided by hours fished for all positive tows from the groundfish trawl sector. Trawl data are shown from 2013 onwards after the trawl footprint was frozen.

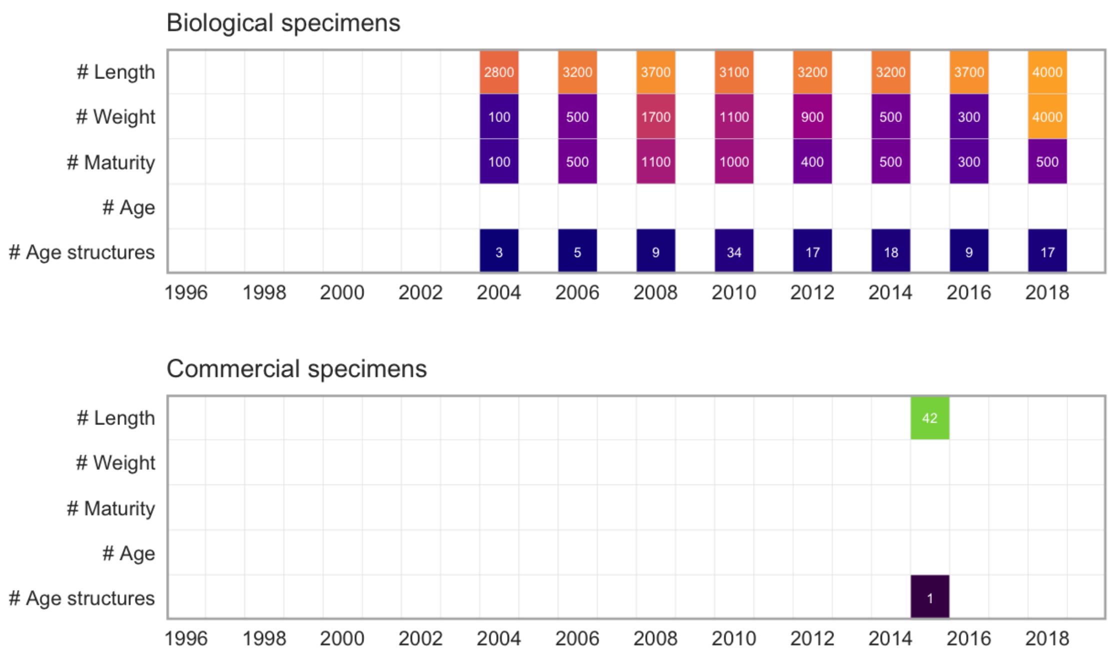

Figure E.5. Specimen availability for Rex Sole. Shown are the number of available fish specimens that have had their length measured, have been weighed, had their maturity assessed, had their age assessed, and for which ageing structures are available for ageing. Blank panels indicate year-measurement combinations without any data. Shading of these cells reflects the relative number of specimens available with the actual number of specimens indicated in the cells to the nearest round number.

E.2 STEP 1: DEFINE THE DECISION CONTEXT

The decision to be made is which MP to use to determine catch limits for the period until the next available catch advice. The time-frame for making the decision would be stated in the Request for Science Advice. In applications of the MP Framework, additional aspects of the decision context would be established through a technical committee. See Section 3.1.

E.3 STEP 2: SELECTION OF OBJECTIVES AND PERFORMANCE METRICS

We defined the objectives for Rex Sole based on the provisional objectives presented in Section 3.2. These objectives are guided by the Precautionary Approach Framework, elements of which are incorporated into the Fish Stocks provisions of the *Fisheries Act*.

1. Maintain stock status above the LRP in the long term with high probability.
2. Maintain stock status above the USR in the long term with moderately high probability.
3. Maintain a fishing exploitation rate below the rate at maximum sustainable yield with some probability.
4. Given the above conservation objectives are achieved, maximize short- and long-term fisheries catch.
5. Given the above conservation objectives are achieved, minimize variability in fisheries catch from year to year.

Our estimated length-at-50% maturity (25.8 cm), coupled with our estimated growth parameters, imply an age-at-50% maturity of approximately 4–5 years. Since 1.5 to 2 generations would be less than 50 years, we chose to base our long-term performance metrics on a 50-year projection.

The performance metrics measuring the objectives were:

1. LT LRP: Probability $B > 0.4\ B_{\mathrm{MSY}}$ (years 35–50).
2. LT USR: Probability $B > 0.8\ B_{\mathrm{MSY}}$ (years 35–50).
3. FMSY: $\mathrm{P}(F < F_{\mathrm{MSY}})$ (years 1–50).
4. STC: Probability catch > reference catch (years 1–10).
5. LTC: Probability catch > reference catch (years 35–50).
6. AADC: Probability AADC (average absolute interannual difference in catch) < historical AADC (years 1–50).

Reference catch was defined as the average catch from the last five years. We measured the performance metrics as probabilities integrated over years and replicates simultaneously.

E.4 STEP 3: SELECTION OF UNCERTAINTIES AND SPECIFICATION OF OPERATING MODELS

We considered the major axes of uncertainty for Rex Sole to be:

1. The magnitude of the commercial catch prior to 1996 (influencing initial depletion in 1996)
2. The influence of the commercial CPUE index
3. The value of the steepness parameter (h) of the stock-recruit relationship

4. The value of natural mortality (*M*)
5. The age selectivity pattern of the commercial fleet and survey gear

We divided OM scenarios into a reference set and a robustness set (Section 2.3). We defined six OM reference-set scenarios that differed from each other in these respects and two OM robustness-set scenarios (Table E.1), with parameter values provided in Appendix F. We describe each of the OM scenarios below.

Table E.1. Rex Sole OM scenarios. "C^{eq}" refers to "catch equilibrium" in 1995 and is used to parameterize the level of biomass depletion at the beginning of the recorded catch history. Unless otherwise specified, "C^{eq}" was set at 200% in all scenarios.

OM scenario name	Set type
C^{eq} 200%	Reference
C^{eq} 250%	Reference
Higher M	Reference
Higher steepness	Reference
Lower selectivity	Reference
No CPUE C^{eq} 250%	Reference
No CPUE C^{eq} 50%	Robustness
M increasing	Robustness

E.4.1 REFERENCE SET

E.4.1.1 Magnitude of pre-1996 commercial catch and importance of commercial CPUE

Prior to 1996, reported catches for WCVI Rex Sole were much lower than post-1995 and there were no reported discards. We have removed these catches from Figure E.3 because these data are considered unreliable due to a lack of reporting requirements and at-sea observer coverage prior to 1996. Assumptions about pre-1996 catch levels affect the SRA model's estimates of the depletion level in 1996 (Appendix B). The SRA model parameter C^{eq}, describes the equilibrium catch in weight prior to the first year of the model (1996 for Rex Sole), which is used to estimate equilibrium fishing mortality F^{eq} prior to the first year of the model, which in turn is used to calculate total mortality Z^{eq} needed to calculate recruits and numbers in the first year of the model (Equation B.5). We note that the use of F^{eq} and C^{eq} make the assumption that the population biomass and fishery were in equilibrium prior to 1996. While this is likely an unrealistic assumption, it is a convenient means of initializing a population model in the presence of fishing. For this reason, we treated the value of C^{eq} as a major axis of uncertainty.

We attempted to fit C^{eq} values ranging from 0 to 2.5 times the 1996 catch, which imply that equilibrium catch in 1995 was 0% to 250% of the 1996 catch, respectively. The SRA model estimated implausibly high values of fishing mortality ($F_{a,y} > 3$) for values of $C^{\mathrm{eq}} < 200\%$ of 1996 catch, suggesting that there may have been considerable unrecorded, possibly discarded, catch prior to 1996, also implying that the stock was unlikely to have been close to B_0 in 1995. Therefore, our reference set includes only two levels of C^{eq} that we thought rendered plausible OM conditioning. We include an additional OM scenario with the 1995 stock closer to B_0 in the robustness set (Section E.4.2).

We believe the assumption of unrecorded catch prior to 1996 is consistent with the data and knowledge of the multispecies groundfish fishery for three reasons. First, the standardized commercial CPUE index shows an increasing trend from 1996 to 2003 before the SYN WCVI survey began (Figure E.2). Given that Rex Sole were not a targeted species, it seems unlikely that the fishery began catching Rex Sole coincident with the introduction of 100% at-sea observer coverage and dockside monitoring. Increasing CPUE is also consistent with (but not necessarily indicative of) a stock rebuilding from a depleted state. While there are many caveats associated with commercial CPUE time series, two factors suggest this trend is plausible: (a) commercial CPUE has somewhat mirrored the broad patterns seen in the SYN WCVI survey index since the survey began (Figure E.2); and (b) Rex Sole have not been a targeted species by the commercial trawl fleet. The absence of targeting removes some of the reasons why CPUE trends may not reflect abundance trends. Second, the SYN WCVI survey has approximately doubled its relative index value for Rex Sole since the early 2000s, which would be consistent with the stock starting in a somewhat depleted state in 1996. Third, other species that are commonly caught with Rex Sole in catch records since 1996 were reported to have been caught in substantial quantities before 1996 (Anderson et al. 2019). It therefore seems reasonable to assume that there was considerable unrecorded catch prior to 1996.

We considered three OM reference-set scenarios and one OM robustness-set scenario that varied in the assumed value of C^{eq} (Table E.1). For the reference set:

- The OM scenario "C^{eq} 200%" assumes that equilibrium catch prior to 1996 was twice that in 1996.
- The OM scenario "C^{eq} 250%" assumes that equilibrium catch prior to 1996 was 2.5 times that in 1996.
- The OM scenario "No CPUE C^{eq} 250%" excludes the commercial CPUE index from the SRA model (Figure E.2). When the commercial CPUE data were excluded, values of $C^{\mathrm{eq}} < 2.5$ rendered implausible SRA model fits ($F_{a,y} > 3$) without other changes to the model such as shifting length selectivity.

For all other OM scenarios C^{eq} was set at 200%. We did explore some OM scenarios with $C^{\mathrm{eq}} > 250\%$ (not shown). These OM scenarios did not provide greater contrast in terms of the ranking of MPs and we chose to cap C^{eq} at 250% in this case study.

E.4.1.2 Steepness of the stock-recruit relationship (*h*)

We considered one alternative OM scenario that differed in terms of the value of *h* (Table E.1). Unless otherwise stated, all other OM scenarios drew from a baseline beta distribution $h \sim \mathrm{beta}(13.4, 2.4)$, which results in a mean and standard deviation of 0.85 and 0.09 (see Section F.1.4).

- The OM scenario "Higher steepness" uses the value $h = 0.95$.

E.4.1.3 Natural mortality (*M*)

We considered one alternative OM scenario that differed in terms of the value of *M* (Table E.1). Unless otherwise stated, all other OM scenarios drew from a baseline uniform distribution *M* = 0.17–0.25 y^{-1}, which starts at the value of 0.17 y^{-1} used for the Gulf of Alaska stock assessment of Rex Sole (McGilliard and Palsson (2017), see Section F.1.3) and explores a slightly higher range since *M* is likely to be somewhat greater at a lower latitude.

- The OM scenario "Higher M" considers that *M* could be latitude-dependent and may be substantially higher for stocks in warmer BC waters (*M* = 0.3 y^{-1}).

E.4.1.4 Selectivity

We considered one alternative OM scenario that differed in terms of the length- and age-selectivity of the commercial fleet and the WCVI synoptic survey (Table E.1). Unless otherwise stated, all other OM scenarios drew from a length selectivity curve that approximately matched the maturity ogive ($L^5 = 22$, $L^{\mathrm{FS}} = 32$).

- The OM scenario "Lower selectivity" considers that the selectivity curve could be shifted to the left resulting in both the survey and the commercial fleet capturing younger fish ($L^5 = 17$, $L^{\mathrm{FS}} = 28$).

We considered the above six OM scenarios to represent the most plausible uncertainty OM scenarios for Rex Sole and grouped them into the "reference set" of OMs (see Section 2.3).

E.4.2 ROBUSTNESS SET

To illustrate application of the framework in the presence of additional sources of structural uncertainty, we included two OM robustness scenarios (Table E.1):

- The "No CPUE C^{eq} 50%" OM scenario assumes that catch at equilibrium in year 1995 was 50% of the catch in 1996 ($C^{\mathrm{eq}} = 0.5$). In other words, this OM scenario assumes that the stock started in a much less depleted state than the other OM scenarios. In order to fit the SRA model with plausible $F_{a,y}$ values, we also had to omit the commercial CPUE and use the selectivity curve described in the "Lower selectivity" OM scenario.
- The "M increasing" OM scenario assumes that *M* was fixed at 0.2 y^{-1} in the historical period and increases monotonically from 0.2 to 0.4 y^{-1} over the projection period. This could represent increasing natural mortality due to a process such as climate change or shifts in predator abundance. We include this OM scenario simply to illustrate including a time-varying parameter in the OM projection period that reflects an underlying environmental process, and to assess robustness of MPs to such a scenario.

Results from the Robustness Set will be presented separately in Section E.7 and will be used to show whether performance of MPs is significantly affected by these other sources of uncertainty.

We include a detailed explanation of the initial OM specification in Appendix F.

E.4.3 OM CONDITIONING

After specifying most of the OM parameters (Appendix F), we conditioned the OMs using the SRA model (with 250 replicates) described in Appendix B. We conditioned the models using: (1) commercial catch data from 1996–2019; (2) the synoptic WCVI (SYN WCVI) trawl relative biomass index (a depth-stratified random survey); and (3) commercial bottom trawl CPUE. We down-weighted the commercial CPUE index by inflating the standard errors 1.5 times since we have greater confidence in the SYN WCVI data reflecting underlying biomass trends. We specified a range of C^{eq} values as described above to reflect uncertainty in initial stock depletion.

The SRA was able to fit to the WCVI synoptic trawl survey relative biomass index and commercial bottom trawl CPUE reasonably well for all OM scenarios (Figure E.6). The SYN WCVI survey was better fit than the commercial CPUE, likely due to the smaller standard errors (Figure E.6). The SRA models fit the catch data almost perfectly by design, via setting the standard deviations of the observation error to a value of 0.01 (Equation B.21).

We used the SRA to populate the following parameters in the conditioned OMs:

- B_{t_c}/B_0 (or "D"; depletion in the last historical year)
- R_0 (unfished recruitment)
- θ_{AC} (or "AC"; first-order autocorrelation of recruitment deviations)
- $F_{a,y}$ (fishing mortality at age by year)
- $\varepsilon_{\mathrm{R},y}$ for years t_1 to t_c (historical recruitment deviations)

The OM reference and robustness scenarios rendered a range of estimated parameter values (Figures E.7, E.10, E.11). In fitting the SRA across replicate draws from the initial OMs, we discarded any models that did not converge (with convergence defined as a positive-definite covariance matrix) or had $F_{a,y} > 3$. The majority of replicates were retained (Table E.2). However, by discarding replicates from non-converged models, the input parameters carried through to the conditioned OMs varied slightly (Figure E.8). The discarded parameter values represent implausible parameter combinations according to the SRA model. The implied spawning stock biomass depletion trajectories during the historical period from the eight OMs were similar with the exception of the "No CPUE C^{eq} 50%" OM scenario (Figure E.9).

Table E.2. Fraction of replicates retained after conditioning with the SRA model.

OM Scenario	Fraction retained
1 - C^{eq} 200%	1.00
2 - C^{eq} 250%	1.00
3 - Higher M	1.00
4 - Higher steepness	1.00
5 - Lower selectivity	0.99
6 - No CPUE C^{eq} 250%	1.00
7 - No CPUE C^{eq} 50%	0.92
8 - M increasing	1.00

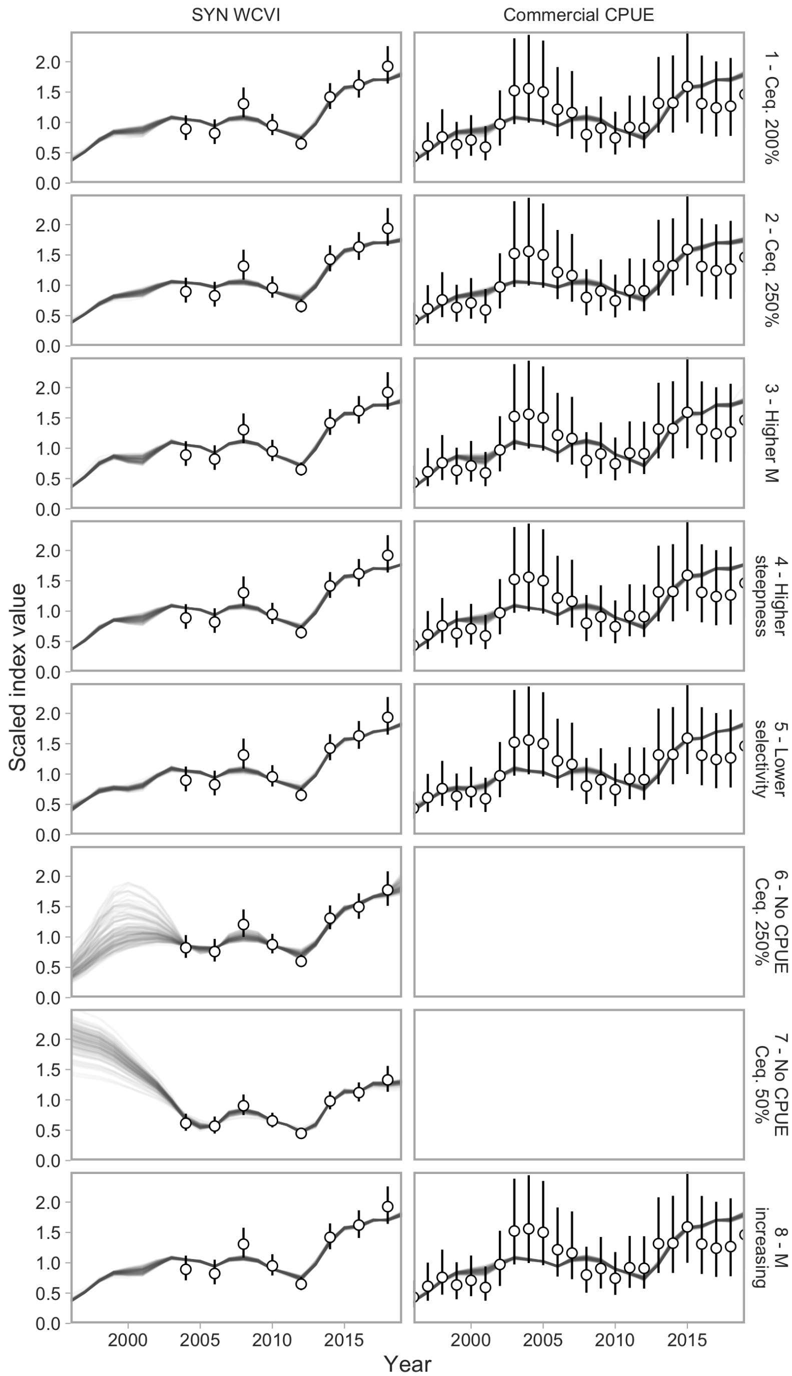

Figure E.6. SRA model fits to the SYN WCVI and commercial bottom trawl CPUE relative biomass indices. Panels from top to bottom represent OM scenarios. Thin lines represent individual SRA model fits across stochastic draws from the various OM parameters. Dots represent index mean and line segments represent 2 times the standard errors as entered into the SRA models. In the case of the commercial CPUE, these standard errors have been inflated 1.5 times from the fitted standard errors to down-weight the commercial CPUE index relative to the survey index.

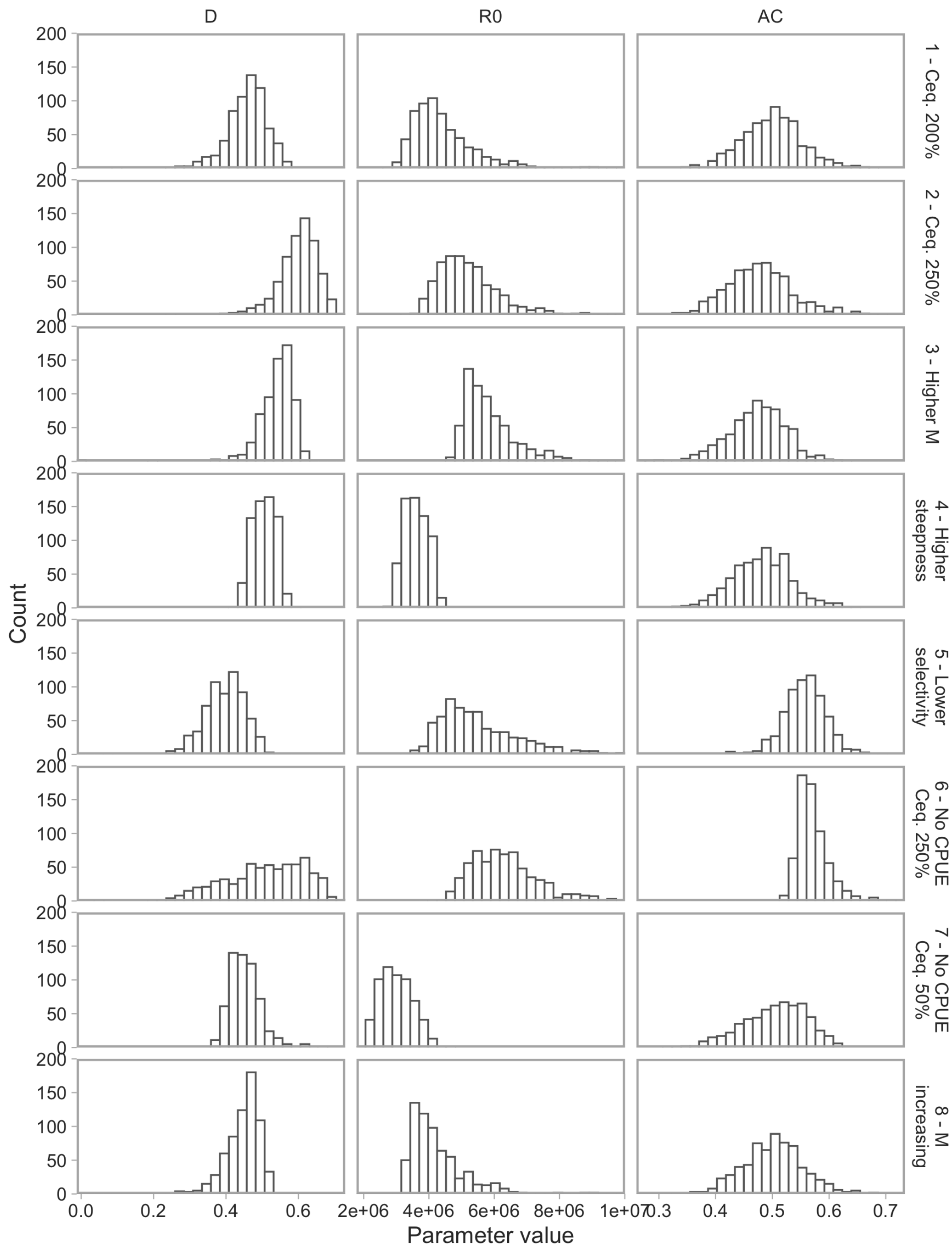

Figure E.7. Histograms of parameters estimated by the SRA. Fishing mortality at age, the historical depletion trajectory, and historical recruitment deviations are also derived from the SRA model. Historical depletion trajectories are shown in Figure E.9, apical fishing mortality by year are shown in Figure E.10, and historical recruitment deviations are shown in Figure E.11. D refers to depletion. AC refers to θ_{AC}.

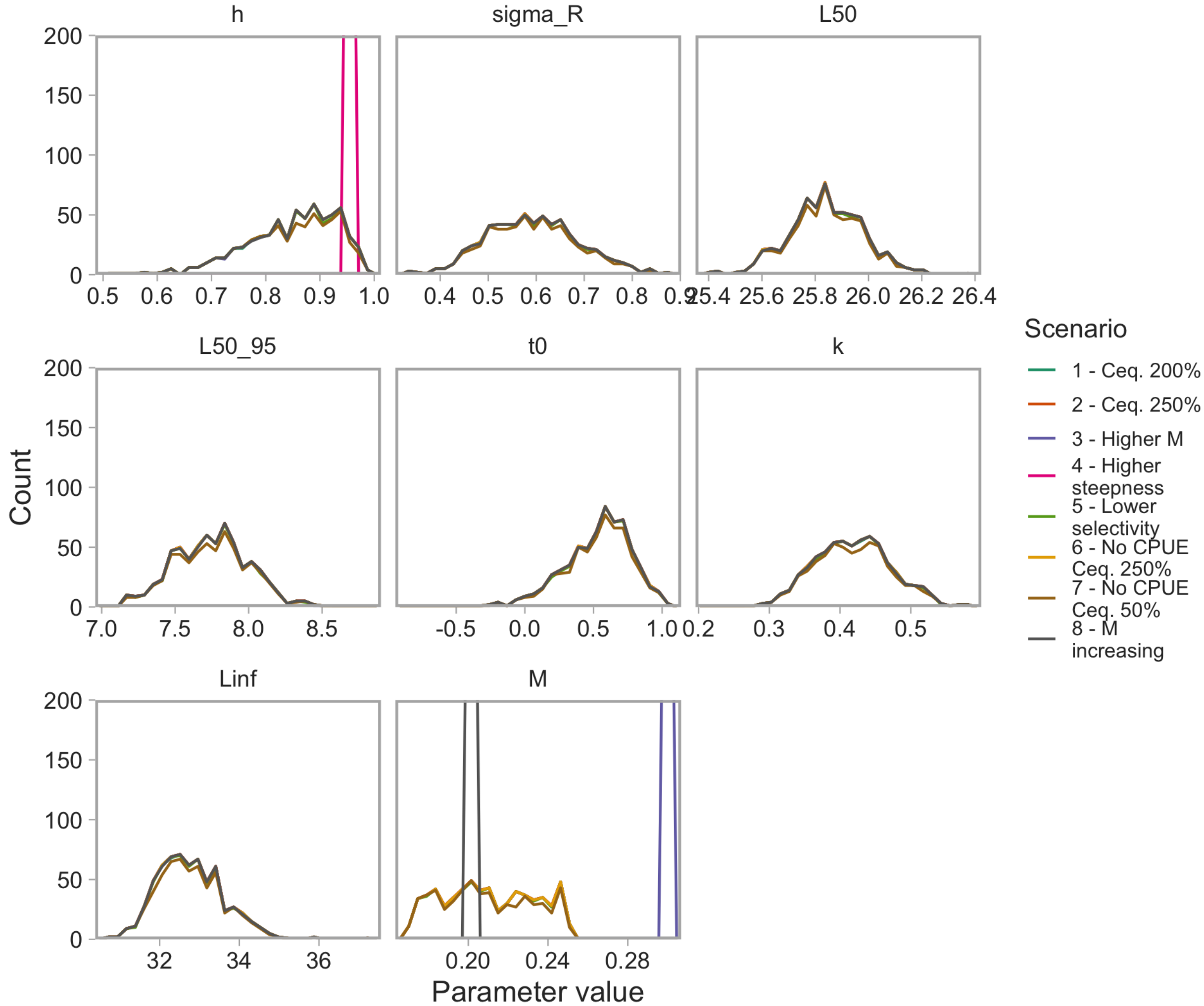

Figure E.8. Frequency polygons for parameters input to the SRA. Parameters were sampled stochastically as input into the SRA (see Appendix F for ranges). Any samples associated with non-converged SRA models were discarded. Since most SRA models converged, these distributions are similar across OM scenarios unless they were specified differently in the OM. The large spikes represent fixed parameter values in the OM. sigma_R refers to σ_R.

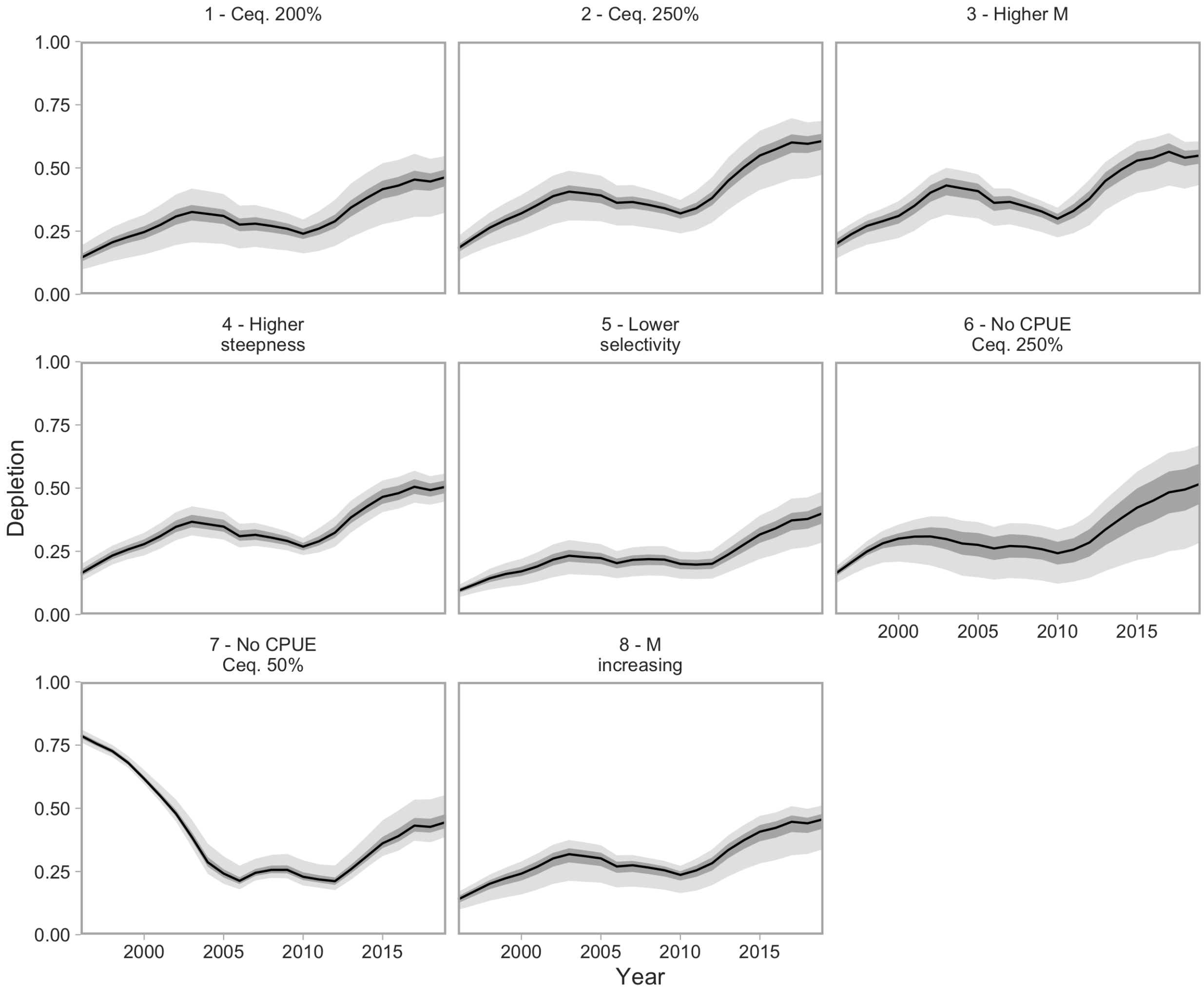

Figure E.9. Spawning stock biomass (B) depletion trajectories for reference and robustness set OMs. Depletion is represented as a fraction of B_0 (spawning stock biomass at unfished equilibrium). Note that the "M increasing" OM scenario only represents M increasing in the future projections—not the historical time period illustrated here—and uses a fixed value of M in the historical period. Lines represent medians, and dark and light-grey shading represent 50% and 95% quantiles across replicates.

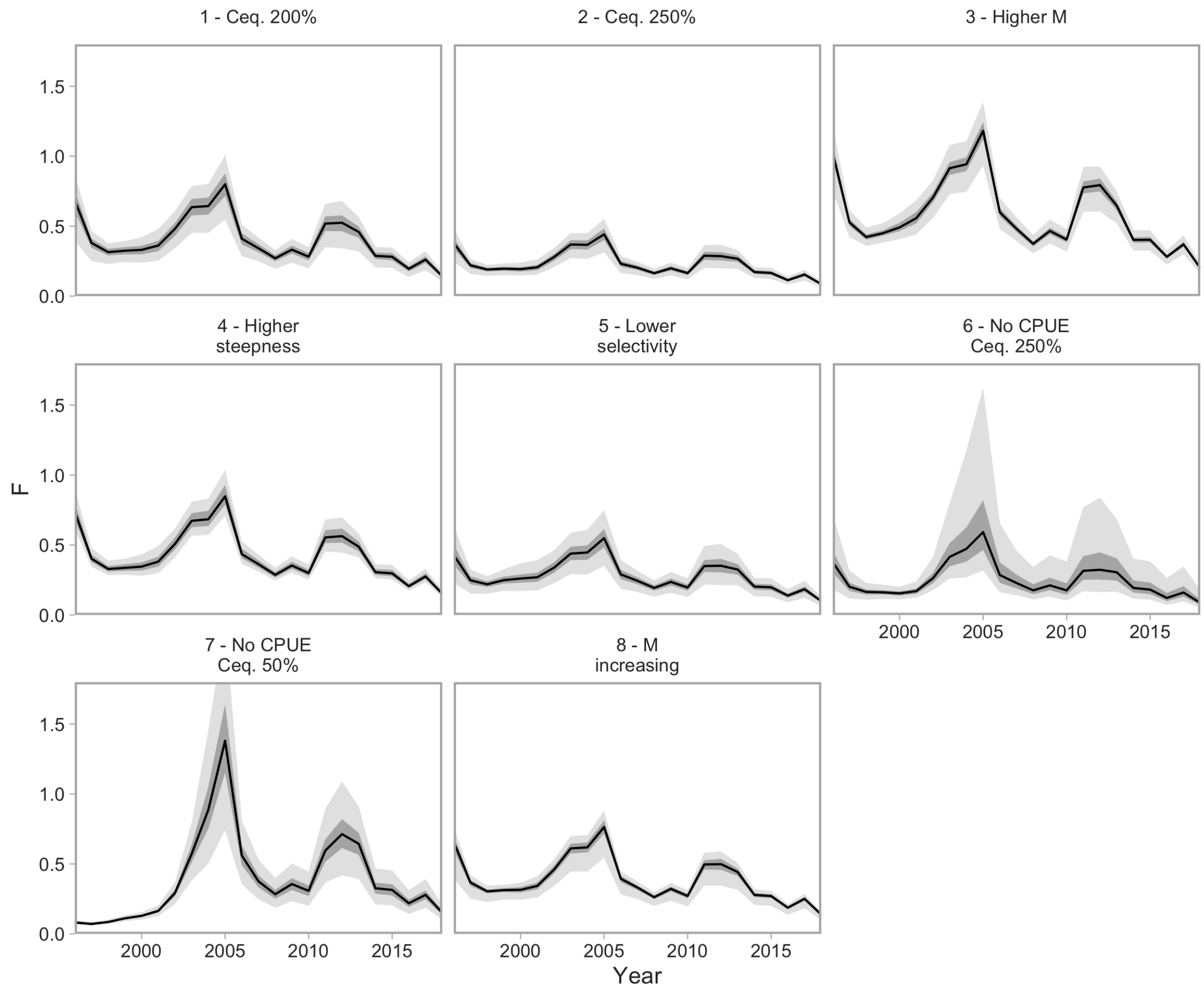

Figure E.10. Apical fishing mortality (F_y) trajectories for reference and robustness set OMs. Apical fishing mortality is the maximum $F_{a,f}$ experienced by fish of any age in a given year. Lines represent medians, and dark and light grey shading represent 50% and 95% quantiles across replicates.

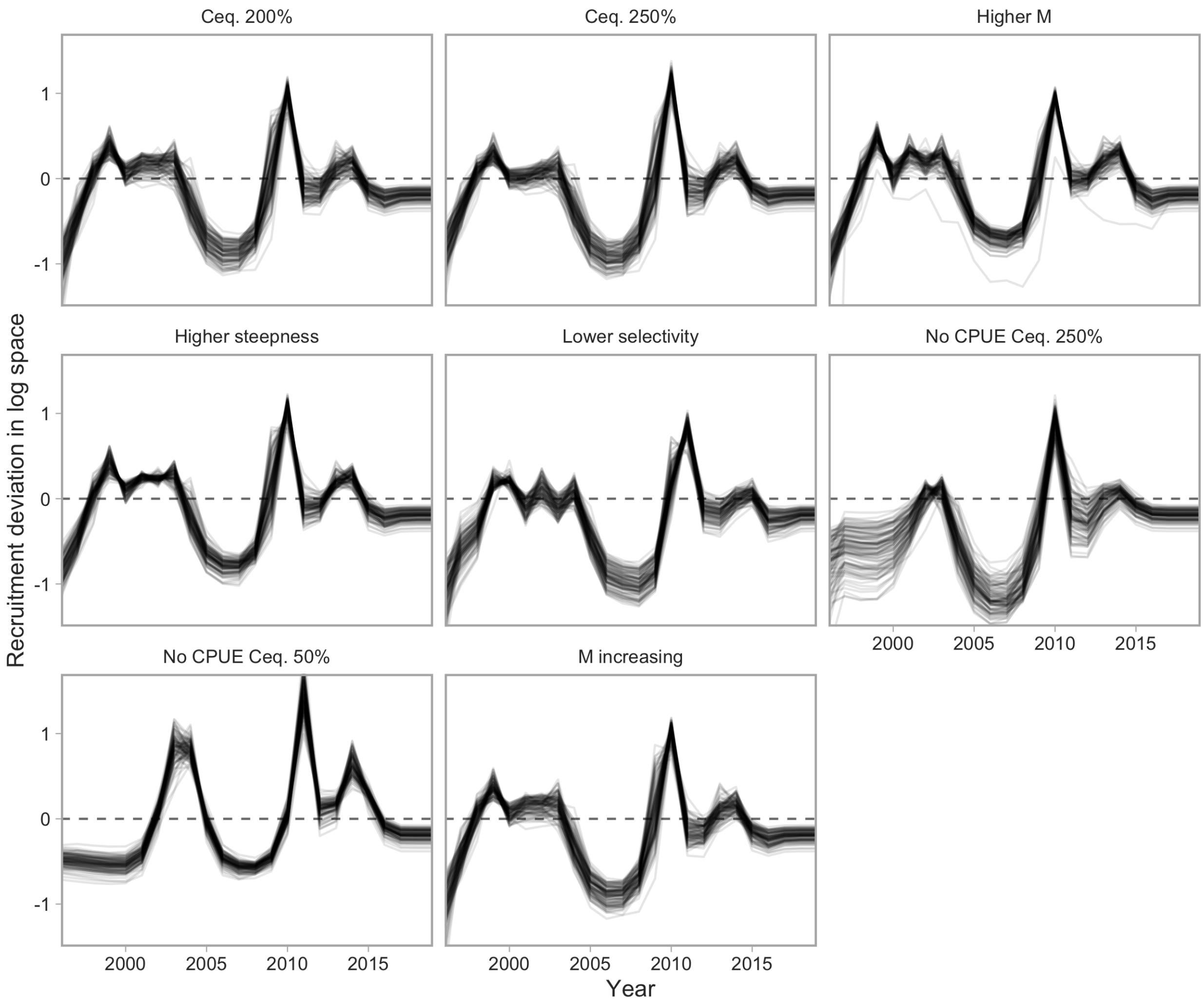

Figure E.11. Historical recruitment deviations estimated by the SRA model (in log space). Lines represent 100 random samples from the replicates.

E.5 STEP 4: IDENTIFICATION OF CANDIDATE MANAGEMENT PROCEDURES

We started with the full set of provisional candidate MPs described in Appendix D (Table E.3). Early explorations with the SP models suggested little effect of the three alternative harvest control rules (HCRs) described in Section D.4.1 (presumably because the projected biomass spent little time on the HCR “ramp”). There was also very little difference in performance between the Fox and Schaefer production models. Therefore, for the SP models we focused mostly on tuning the prior on r, the intrinsic population growth rate. We evaluated three priors on r: $\mathrm{Normal}(0.4, 0.1)$, $\mathrm{Normal}(0.5, 0.1)$, and $\mathrm{Normal}(0.6, 0.1)$. With the Schaefer surplus production function, these correspond to $F_{\mathrm{MSY}} = 0.2$, $F_{\mathrm{MSY}} = 0.25$, and $F_{\mathrm{MSY}} = 0.3$, which approximately match the values of *M* explored in the OMs. We included the provisional “meta” HCR rules described in Appendix D Section D.4.1 and given that the stock does not have an existing TAC, we set the initial TAC for the purpose of the “meta” rules to the average catch from the last five years of the historical period. In applications of the framework, additional tuning of SP models and index-based MPs could be explored if needed.

Table E.3. Candidate MPs. The ‘.’ in front of some MPs denotes versions of the MPs that only observe the survey index every second year to mimic our biennial WCVI synoptic survey. .SP6040_0.4 refers to an SP model with a prior on r of 0.4, etc.

Management procedure	MP type
CC_hist20	Constant catch
CC1.2	Constant catch
CC1.1	Constant catch
CC1.0	Constant catch
CC0.9	Constant catch
CC0.8	Constant catch
CC0.7	Constant catch
CC0.6	Constant catch
.Iratio2	Index ratio
.GB_slope6_0.66	Index slope
.GB_slope6_1	Index slope
.GB_slope8_0.66	Index slope
.GB_slope8_1	Index slope
.Islope0.2_80	Index slope
.Islope0.2_100	Index slope
.Islope0.4_80	Index slope
.Islope0.4_100	Index slope
.IDX	Index ratio
.IDX_smooth	Index ratio
.IT10_hist	Index target
.IT5_hist	Index target
.Itarget_base	Index target
.Itarget_w0.8	Index target
.Itarget_x0.2	Index target
.Itarget_x0.8	Index target
.Itarget_d1.2	Index target
.Itarget_d0.8	Index target

Management procedure	MP type
.ITM_hist	Index target
.SP6040_0.4	Surplus production
.SP6040_0.5	Surplus production
.SP6040_0.6	Surplus production
.SP8040_0.6	Surplus production
.SP4010_0.6	Surplus production
.SP6040_0.6_fox	Surplus production
NFref	Reference
FMSYref	Reference
FMSYref75	Reference

E.6 STEP 5: SIMULATION OF THE APPLICATION OF THE MANAGEMENT PROCEDURES

We ran the closed-loop simulations across 250 stochastic replicates using DLMtool version 5.4.2, R version 4.0.4, and the simulation random seed set to `1`, with the OMs and MPs described above. We assessed convergence of the closed-loop simulation by plotting the cumulative performance metrics as replicates were added (Figure E.12). We deemed 250 replicates sufficient since the rank order of MPs was consistent for this number of replicates (Figure E.12).

To determine which MPs would be carried forward as satisficed MPs (Miller and Shelton 2010), we began by assessing average and minimum performance across all candidate MPs for the reference set of OMs (Figures E.13 and E.14). To obtain a manageable number of MPs for further consideration, we set satisficing thresholds of LT LRP $>$ 0.9 (9 times out of 10) and STC $>$ 0.8 (4 times out of 5). We could apply these criteria to either the average (Figures E.13) or minimum performance (Figures E.14) and such a decision is likely best made by a technical working group or decision-makers coupled with stakeholders in an applied setting. In this case study, it made little difference whether we chose the average or minimum performance metrics to apply the satisficing criteria. For the purposes of this case study, we chose to apply the satisficing criteria to the minimum performance metrics. This implies that the satisficed MPs pass these criteria even in the worst-case OM reference-set scenario. Applying these criteria resulted in eleven remaining MPs (.Itarget_base, .Itarget_d0.8, .Itarget_w0.8, .Itarget_x0.2, .SP4010_0.6, .SP6040_0.6, .SP6040_0.6_fox, .SP8040_0.6, CC_hist20, CC1.1, and CC1.2).

We additionally decided to omit four MPs from further consideration. First, we chose to retain only one satisficed SP model MP (.SP8040_0.6) since the SP model MPs that differed only in production-function shape (.SP6040_0.6_fox) or HCR (.SP6040_0.6 and .SP4010_0.6) had near identical performance (Figure E.14). Second, we removed CC_hist20 since it was outperformed by the other constant catch MPs—LT LRP and STC were both exceeded by all other constant-catch MPs (Figure E.14). This left seven MPs (.Itarget_base, .Itarget_d0.8, .Itarget_w0.8, .Itarget_x0.2, .SP8040_0.6, CC1.1, and CC1.2).

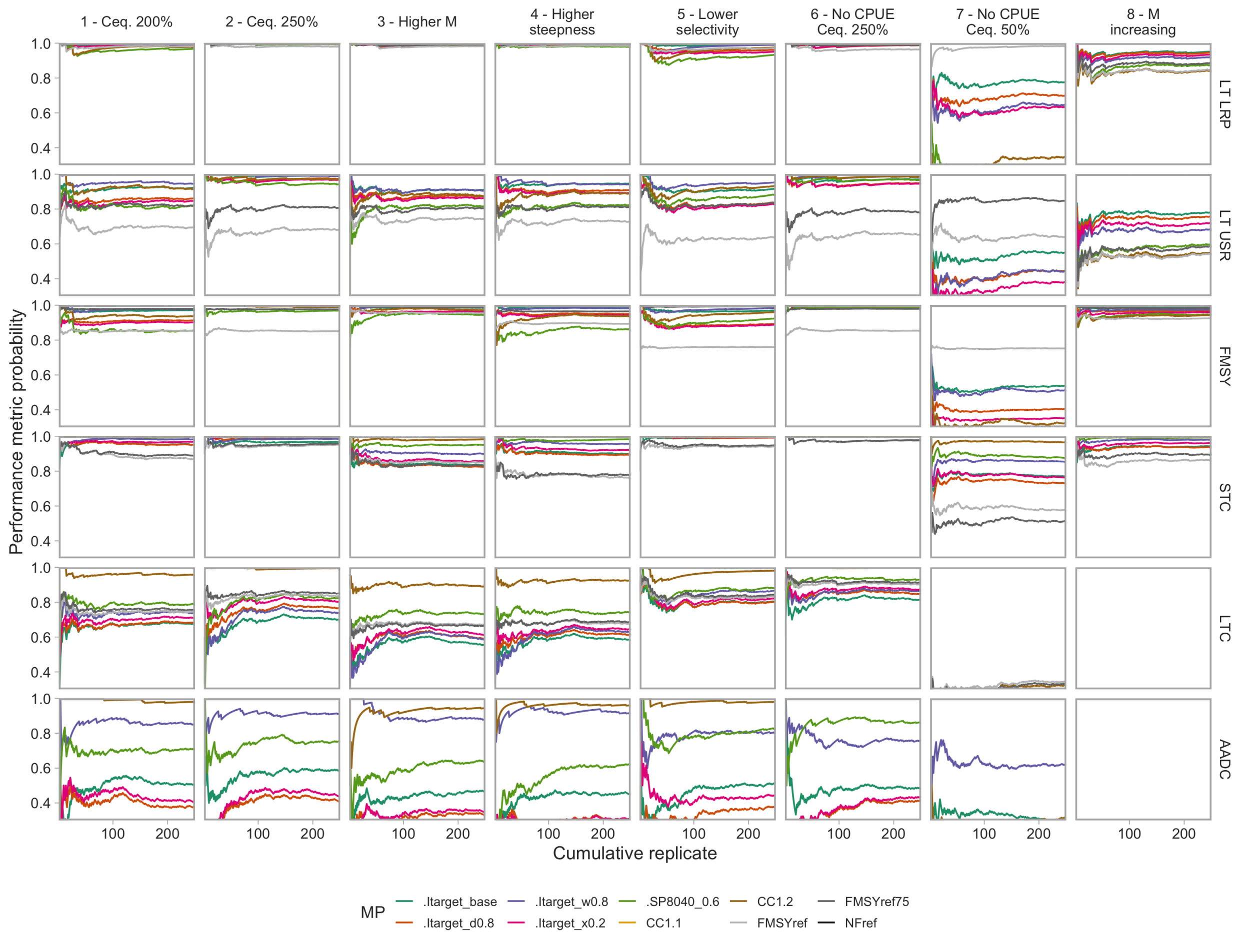

Figure E.12. Assessing convergence of the Rex Sole closed-loop simulations on consistent rank order of MPs within performance metrics. Colours represent individual satisficed and reference MPs. Lines that do not cross by the final replicates indicate that rank order among replicates had converged. Although not shown, we also checked that the satisficing rules had converged (i.e., the selection of satisficed MPs did not change with additional replicates). Note that in the "M increasing" OM scenario the LTC and AADC metrics do not appear because they are below the lower y-axis limit (or perfectly at 1 in the case of AADC and NFref).

	LT LRP	LT USR	FMSY	STC	LTC	AADC
.GB_slope8_1	>0.99	>0.99	>0.99	<0.01	0.07	0.84
.GB_slope6_1	>0.99	>0.99	>0.99	<0.01	0.02	0.92
.GB_slope8_0.66	>0.99	>0.99	>0.99	<0.01	0.01	0.97
.GB_slope6_0.66	>0.99	>0.99	>0.99	<0.01	<0.01	0.99
.Islope0.4_80	>0.99	>0.99	>0.99	<0.01	<0.01	>0.99
.Islope0.4_100	>0.99	>0.99	>0.99	<0.01	<0.01	>0.99
.Islope0.2_80	>0.99	>0.99	>0.99	<0.01	<0.01	>0.99
.Islope0.2_100	>0.99	>0.99	>0.99	<0.01	<0.01	>0.99
NFref	>0.99	>0.99	>0.99	<0.01	<0.01	>0.99
CC0.6	>0.99	>0.99	>0.99	<0.01	<0.01	>0.99
.ITM_hist	>0.99	>0.99	>0.99	<0.01	<0.01	>0.99
.IT5_hist	>0.99	>0.99	>0.99	<0.01	<0.01	>0.99
.IT10_hist	>0.99	>0.99	>0.99	<0.01	<0.01	>0.99
.IDX_smooth	>0.99	>0.99	>0.99	<0.01	<0.01	>0.99
.IDX	>0.99	>0.99	>0.99	<0.01	<0.01	>0.99
.Iratio2	>0.99	>0.99	>0.99	0.03	0.28	0.61
CC0.7	>0.99	>0.99	>0.99	<0.01	<0.01	>0.99
CC0.8	>0.99	0.99	>0.99	<0.01	<0.01	>0.99
.Itarget_x0.8	>0.99	0.99	>0.99	0.03	0.34	0.86
.Itarget_d1.2	>0.99	0.96	0.99	0.88	0.63	0.59
CC0.9	>0.99	0.98	>0.99	<0.01	<0.01	>0.99
FMSYref75	>0.99	0.81	0.98	0.90	0.78	0.05
CC1.0	>0.99	0.95	0.98	<0.01	<0.01	0.99
.Itarget_w0.8	>0.99	0.95	0.99	0.97	0.74	0.85
.Itarget_base	>0.99	0.94	0.98	0.94	0.69	0.50
CC1.2	0.99	0.93	0.96	>0.99	0.96	0.98
CC1.1	0.99	0.93	0.96	>0.99	0.96	0.98
.Itarget_d0.8	0.98	0.90	0.95	0.94	0.72	0.37
.Itarget_x0.2	0.98	0.89	0.94	0.96	0.74	0.40
CC_hist20	0.98	0.90	0.94	0.99	0.94	0.96
.SP8040_0.6	0.98	0.87	0.92	0.99	0.81	0.73
.SP4010_0.6	0.98	0.87	0.92	0.99	0.82	0.73
FMSYref	0.98	0.69	0.86	0.89	0.77	0.02
.SP6040_0.6_fox	0.97	0.86	0.92	0.99	0.82	0.73
.SP6040_0.6	0.97	0.86	0.92	0.99	0.82	0.73
.SP6040_0.4	0.94	0.82	0.85	0.98	0.68	0.50
.SP6040_0.5	0.94	0.81	0.88	0.98	0.69	0.55

Figure E.13. Average performance of all candidate MPs across the reference set of OMs. MPs are ordered by decreasing performance metric values from top to bottom starting with the left-most performance metric (LT LRP) and using columns from left to right to break any ties. The colour shading reflects the probabilities. Outlined cells represent MPs that met a particular performance metric's satisficing criteria. Using this set of criteria, MPs would be "satisficed" if cells in both *"LT LRP" and "STC" were outlined. Light grey MPs indicate reference MPs.*

	LT LRP	LT USR	FMSY	STC	LTC	AADC
.GB_slope8_1	>0.99	>0.99	>0.99	<0.01	0.03	0.75
.GB_slope6_1	>0.99	>0.99	>0.99	<0.01	<0.01	0.85
NFref	>0.99	>0.99	>0.99	<0.01	<0.01	>0.99
CC0.6	>0.99	>0.99	>0.99	<0.01	<0.01	>0.99
.ITM_hist	>0.99	>0.99	>0.99	<0.01	<0.01	>0.99
.IT5_hist	>0.99	>0.99	>0.99	<0.01	<0.01	>0.99
.IT10_hist	>0.99	>0.99	>0.99	<0.01	<0.01	>0.99
.IDX_smooth	>0.99	>0.99	>0.99	<0.01	<0.01	>0.99
.IDX	>0.99	>0.99	>0.99	<0.01	<0.01	>0.99
.Islope0.4_80	>0.99	>0.99	>0.99	<0.01	<0.01	0.99
.Islope0.4_100	>0.99	>0.99	>0.99	<0.01	<0.01	0.99
.Islope0.2_80	>0.99	>0.99	>0.99	<0.01	<0.01	0.99
.Islope0.2_100	>0.99	>0.99	>0.99	<0.01	<0.01	0.99
.GB_slope6_0.66	>0.99	>0.99	>0.99	<0.01	<0.01	0.97
.GB_slope8_0.66	>0.99	>0.99	>0.99	<0.01	<0.01	0.94
.Iratio2	>0.99	0.99	>0.99	0.01	0.18	0.50
CC0.7	>0.99	0.99	>0.99	<0.01	<0.01	>0.99
CC0.8	>0.99	0.98	>0.99	<0.01	<0.01	>0.99
.Itarget_x0.8	>0.99	0.97	>0.99	0.01	0.24	0.78
.Itarget_d1.2	>0.99	0.93	0.99	0.76	0.50	0.55
CC0.9	0.99	0.94	0.99	<0.01	<0.01	0.99
.Itarget_base	0.99	0.91	0.97	0.83	0.55	0.45
CC1.0	0.99	0.91	0.97	<0.01	<0.01	0.97
.Itarget_w0.8	0.99	0.90	0.98	0.90	0.58	0.76
FMSYref75	0.99	0.78	0.96	0.77	0.66	<0.01
CC1.2	0.98	0.87	0.93	0.98	0.89	0.94
CC1.1	0.98	0.87	0.93	0.98	0.89	0.94
.Itarget_d0.8	0.96	0.84	0.89	0.82	0.59	0.30
FMSYref	0.96	0.64	0.76	0.76	0.67	<0.01
CC_hist20	0.95	0.84	0.90	0.97	0.85	0.90
.Itarget_x0.2	0.95	0.83	0.89	0.85	0.61	0.31
.SP8040_0.6	0.93	0.81	0.85	0.95	0.73	0.62
.SP4010_0.6	0.93	0.81	0.85	0.95	0.73	0.62
.SP6040_0.6_fox	0.93	0.80	0.85	0.95	0.73	0.62
.SP6040_0.6	0.93	0.80	0.85	0.95	0.73	0.62
.SP6040_0.4	0.84	0.72	0.74	0.92	0.57	0.35
.SP6040_0.5	0.81	0.70	0.79	0.94	0.57	0.47

*Figure E.14. Minimum performance of all candidate MPs across the reference set of OMs. This figure is the same as Figure E.13 but shows the **minimum** performance metric across the OMs in the reference set for the purposes of applying satisficing rules. In other words, this figure illustrates the worst performance of each MP across the reference set of OMs.*

E.7 STEP 6: PRESENTATION OF RESULTS AND SELECTION OF MANAGEMENT PROCEDURE

E.7.1 REFERENCE SET RESULTS

Performance within the reference set of OMs varied across MPs and performance metrics (Figures E.13, E.14, E.15, and E.16). There was variability in performance among the individual OM reference-set scenarios with the most variability in both the long-term and entire projection period performance metrics: LT USR, FMSY, LTC, and AADC (Figure E.15). Between the two performance metrics used in the initial satisficing step, *average* LT LRP varied between 0.98 and 1.00, and STC varied between 0.94 and 1.00 across MPs within the reference set (Figure E.13 and E.15). *Minimum* LT LRP varied between 0.93 and 0.99, and STC varied between 0.82 and 0.98 across MPs within the reference set (Figure E.14 and E.15).

A dot-and-line plot of the performance metrics aggregated across OM scenarios helps compare performance across satisficed MPs (Figure E.16). Long-term catch (LTC) was roughly comparable between a fishing strategy of fishing at F/F_{MSY} and the satisficed Itarget or .SP8040_0.6 MPs (Figure E.16). Short-term catch (STC) of all satisficed MPs exceeded these reference MPs (Figure E.16). Fixed-catch MPs of 1.1- or 1.2-times average catch from the last five years rendered similar performance to other satisficed MPs across most performance metrics within the OM reference-set scenarios, with somewhat higher LTC and AADC metrics (Figure E.16). These fixed-catch MPs do not incorporate feedback between the management and biological systems.

Trade-offs among performance metrics for most of the satisficed MPs were relatively minor within the OM reference-set scenarios (Figures E.17 and E.18). Within the "C^{eq} 200%" OM scenario, there was a slight trade-off between STC and LT LRP (Figure E.17). The time series trajectories of the projected survey index (Figure E.19), B/B_{MSY}, F/F_{MSY}, and catch (Figures E.20–E.25) further demonstrate performance across the various MPs and reference-set OMs. Kobe plots demonstrate either the final B/B_{MSY} vs. F/F_{MSY} status among replicates (Figure E.26) or the trajectory of these stock status values through time (Figure E.27).

1 - Ceq. 200%

	LT LRP	LT USR	FMSY	STC	LTC	AADC
.ltarget_w0.8	0.99	0.94	0.98	0.98	0.73	0.85
.ltarget_base	0.99	0.92	0.97	0.95	0.67	0.50
CC1.1	0.98	0.91	0.94	>0.99	0.95	0.98
CC1.2	0.98	0.91	0.94	>0.99	0.95	0.98
.ltarget_d0.8	0.98	0.86	0.91	0.95	0.68	0.38
.ltarget_x0.2	0.98	0.85	0.90	0.97	0.71	0.41
.SP8040_0.6	0.97	0.82	0.85	0.99	0.79	0.70

2 - Ceq. 250%

	LT LRP	LT USR	FMSY	STC	LTC	AADC
.ltarget_w0.8	>0.99	0.99	>0.99	0.99	0.74	0.91
.ltarget_base	>0.99	0.99	>0.99	0.97	0.70	0.59
CC1.1	>0.99	0.97	0.99	>0.99	0.99	>0.99
CC1.2	>0.99	0.97	0.99	>0.99	0.99	>0.99
.ltarget_d0.8	>0.99	0.97	>0.99	0.99	0.77	0.41
.ltarget_x0.2	>0.99	0.97	>0.99	>0.99	0.80	0.44
.SP8040_0.6	0.99	0.94	0.97	>0.99	0.82	0.75

3 - Higher M

	LT LRP	LT USR	FMSY	STC	LTC	AADC
.ltarget_w0.8	0.99	0.90	0.99	0.90	0.58	0.88
.ltarget_base	0.99	0.91	0.99	0.83	0.55	0.46
CC1.1	0.99	0.87	0.97	0.98	0.89	0.94
CC1.2	0.99	0.87	0.97	0.98	0.89	0.94
.ltarget_d0.8	0.99	0.87	0.96	0.82	0.59	0.33
.ltarget_x0.2	0.99	0.86	0.96	0.85	0.61	0.35
.SP8040_0.6	0.99	0.81	0.94	0.95	0.73	0.63

4 - Higher steepness

	LT LRP	LT USR	FMSY	STC	LTC	AADC
.ltarget_w0.8	>0.99	0.94	0.98	0.96	0.63	0.92
.ltarget_base	>0.99	0.95	0.98	0.90	0.58	0.45
CC1.1	0.99	0.89	0.93	>0.99	0.92	0.96
CC1.2	0.99	0.89	0.93	>0.99	0.92	0.96
.ltarget_d0.8	0.99	0.91	0.95	0.89	0.61	0.30
.ltarget_x0.2	0.99	0.89	0.94	0.92	0.64	0.31
.SP8040_0.6	0.98	0.81	0.86	0.98	0.74	0.62

5 - Lower selectivity

	LT LRP	LT USR	FMSY	STC	LTC	AADC
.ltarget_w0.8	0.99	0.95	0.98	>0.99	0.87	0.80
.ltarget_base	0.99	0.92	0.97	0.99	0.80	0.51
CC1.1	0.98	0.93	0.96	>0.99	0.98	0.98
CC1.2	0.98	0.93	0.96	>0.99	0.98	0.98
.ltarget_d0.8	0.96	0.84	0.89	0.99	0.80	0.38
.ltarget_x0.2	0.95	0.83	0.89	>0.99	0.82	0.44
.SP8040_0.6	0.93	0.88	0.92	>0.99	0.88	0.83

6 - No CPUE Ceq. 250%

	LT LRP	LT USR	FMSY	STC	LTC	AADC
.ltarget_w0.8	>0.99	0.99	>0.99	>0.99	0.86	0.76
.ltarget_base	>0.99	0.97	0.99	>0.99	0.81	0.48
CC1.1	>0.99	0.99	>0.99	>0.99	>0.99	>0.99
CC1.2	>0.99	0.99	>0.99	>0.99	>0.99	>0.99
.ltarget_d0.8	0.99	0.95	0.98	>0.99	0.85	0.41
.ltarget_x0.2	0.99	0.95	0.98	>0.99	0.87	0.43
.SP8040_0.6	0.99	0.96	0.99	>0.99	0.93	0.86

Figure E.15. Performance of satisficed MPs for the reference-set OMs. MPs are ordered by decreasing performance metric values from the averaged reference set (Figure E.13). These are the same data underlying Figure E.13 and Figure E.14 but shown for individual OM reference-set scenarios and only for satisficed MPs.

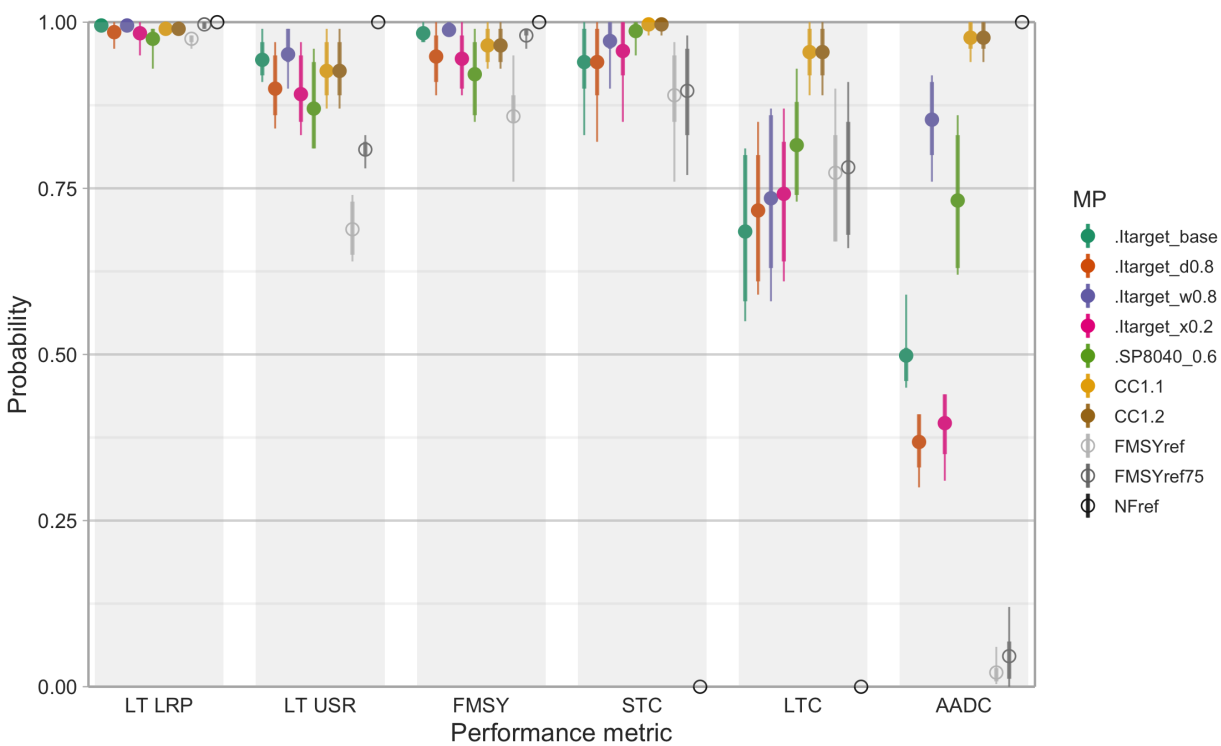

Figure E.16. Dot-and-line plot of performance metrics averaged across OM scenarios. Dots represent average performance metric values and thin lines represent the range of values across OM scenarios. Thick lines represent the range of values across OM scenarios after dropping the high and low values. Reference MPs are indicated by open circles (True). Non-reference MPs are indicated by closed circles (False).

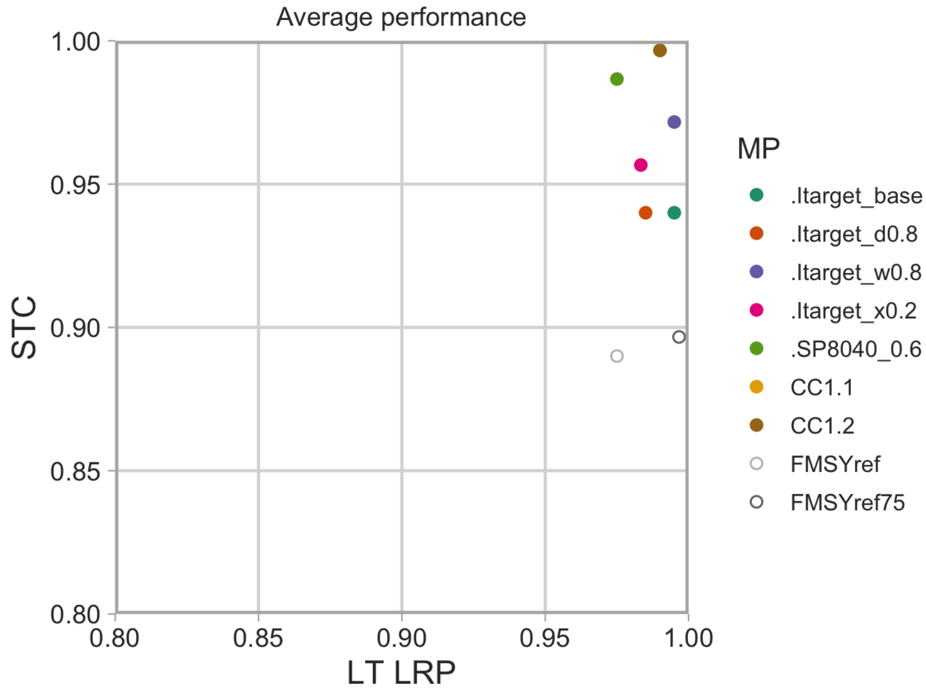

Figure E.17. Trade-off between LT LRP and STC average performance metrics across the OM reference-set scenarios. Reference MPs are indicated by open circles (True). Non-reference MPs are indicated by closed circles (False).

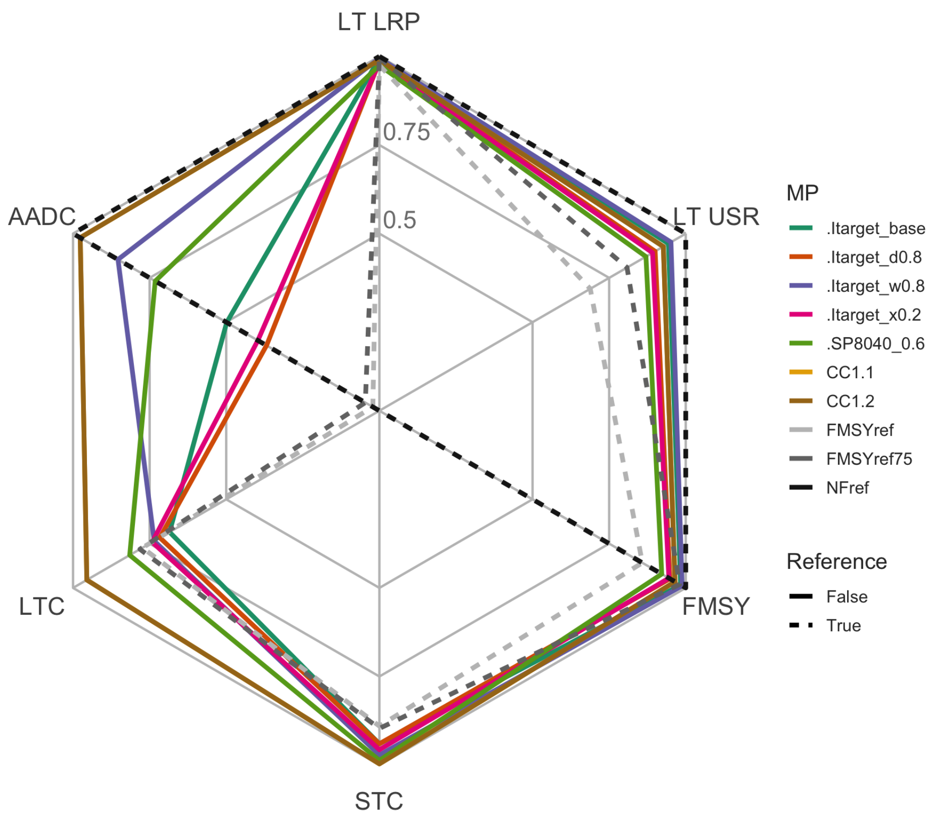

Figure E.18. Radar-plot representation of average performance metric trade-offs for the OM reference-set scenarios. The outside of the hexagon represents a performance metric probability of 1.0; in the middle represents a value of 0. Dashed lines represent reference MPs.

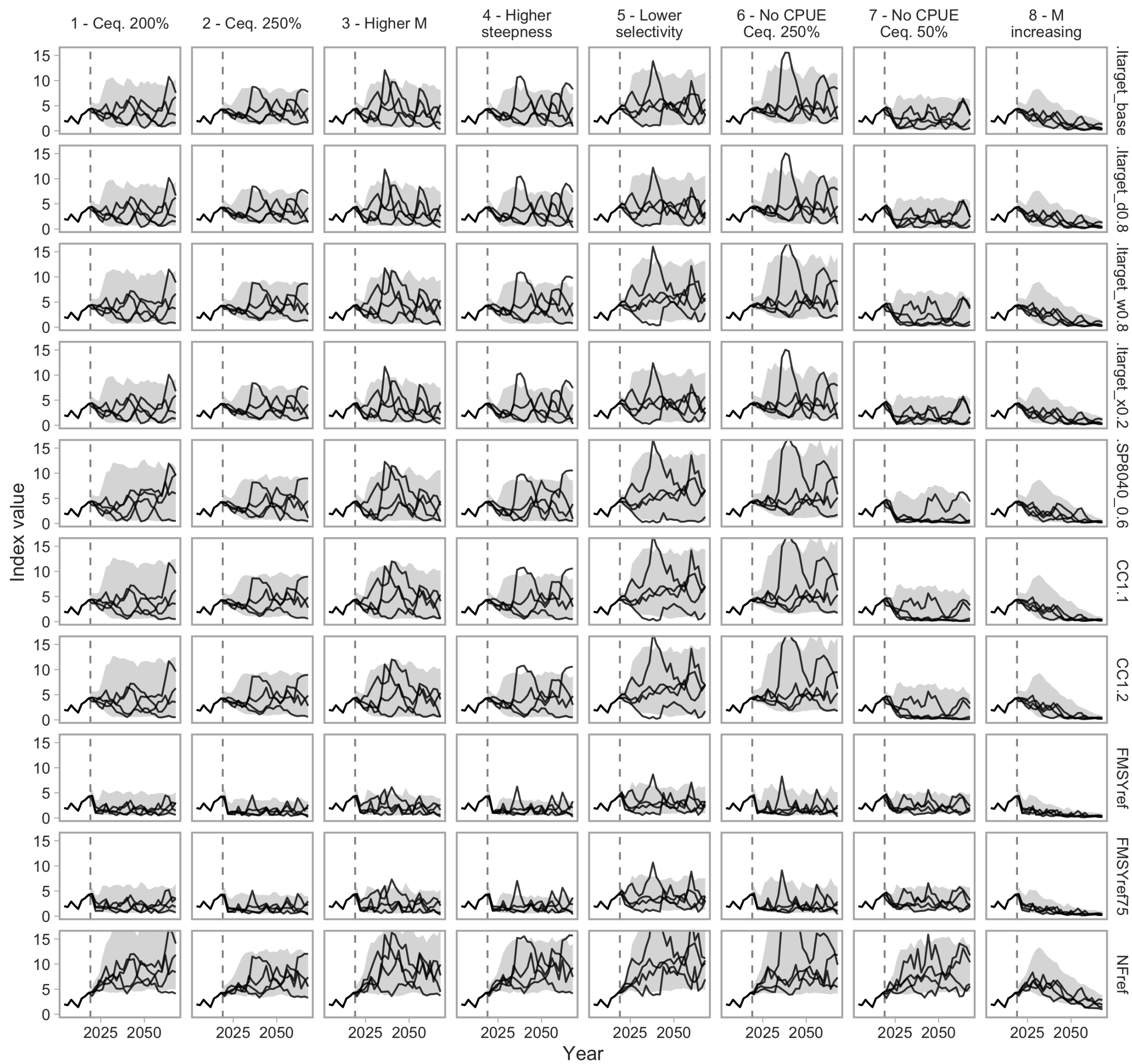

Figure E.19. Historical and projected SYN WCVI relative biomass index values. Vertical dashed line represents 2019. The shaded region represents the 95% quantile and individual lines represent four sample replicates.

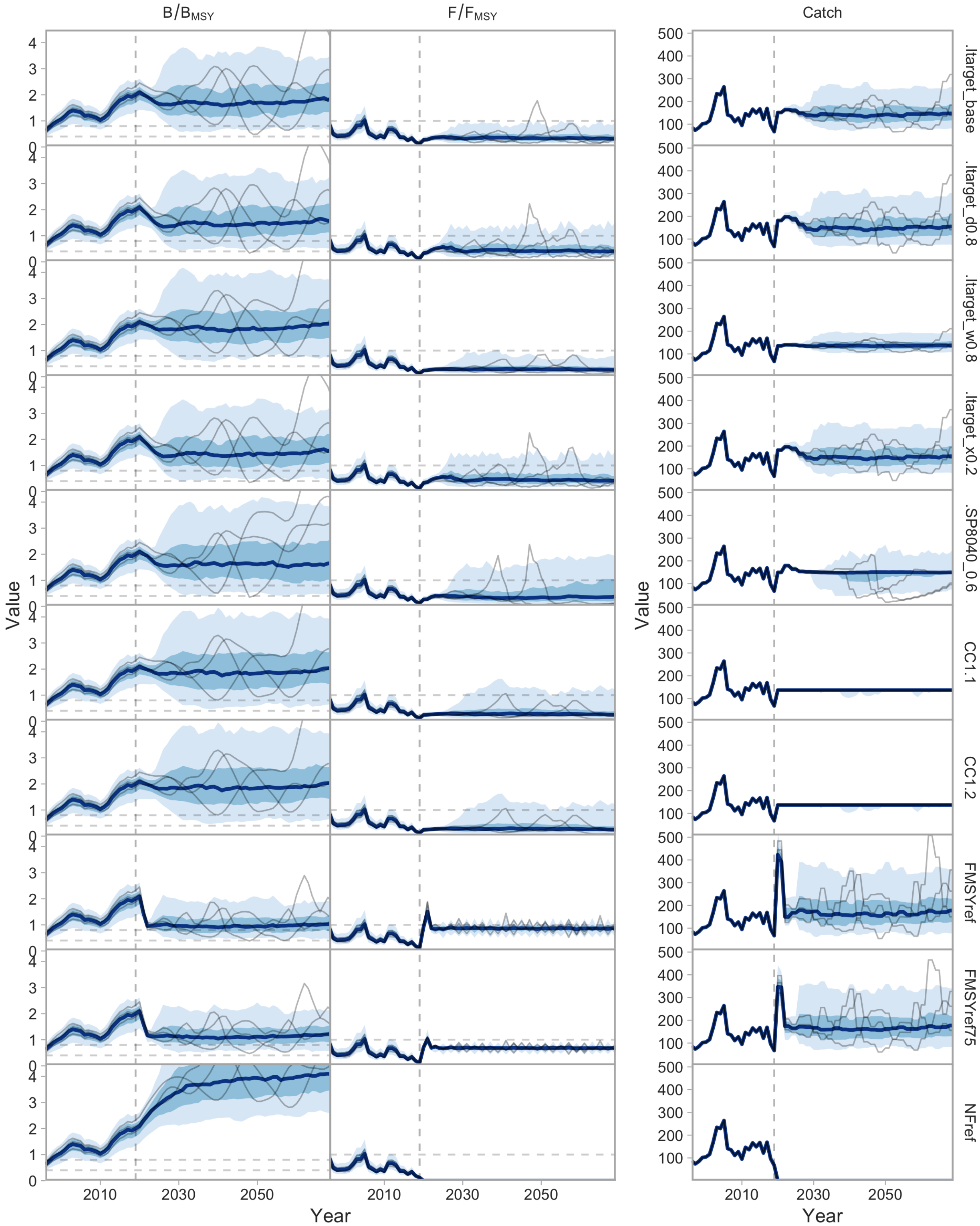

Figure E.20. B/B_{MSY}, F/F_{MSY}, and catch from the historical and projected time periods for the "C^{eq} 200%" OM. Dark line indicates the median value and the darker and lighter shaded ribbons indicate the 50% and 90% quantiles. Thin gray lines represent illustrative simulation replicates. The vertical dashed line indicates the last year of the historical period. The horizontal dashed lines indicate B/B_{MSY} = 0.8 and 0.4 and F/F_{MSY} = 1.

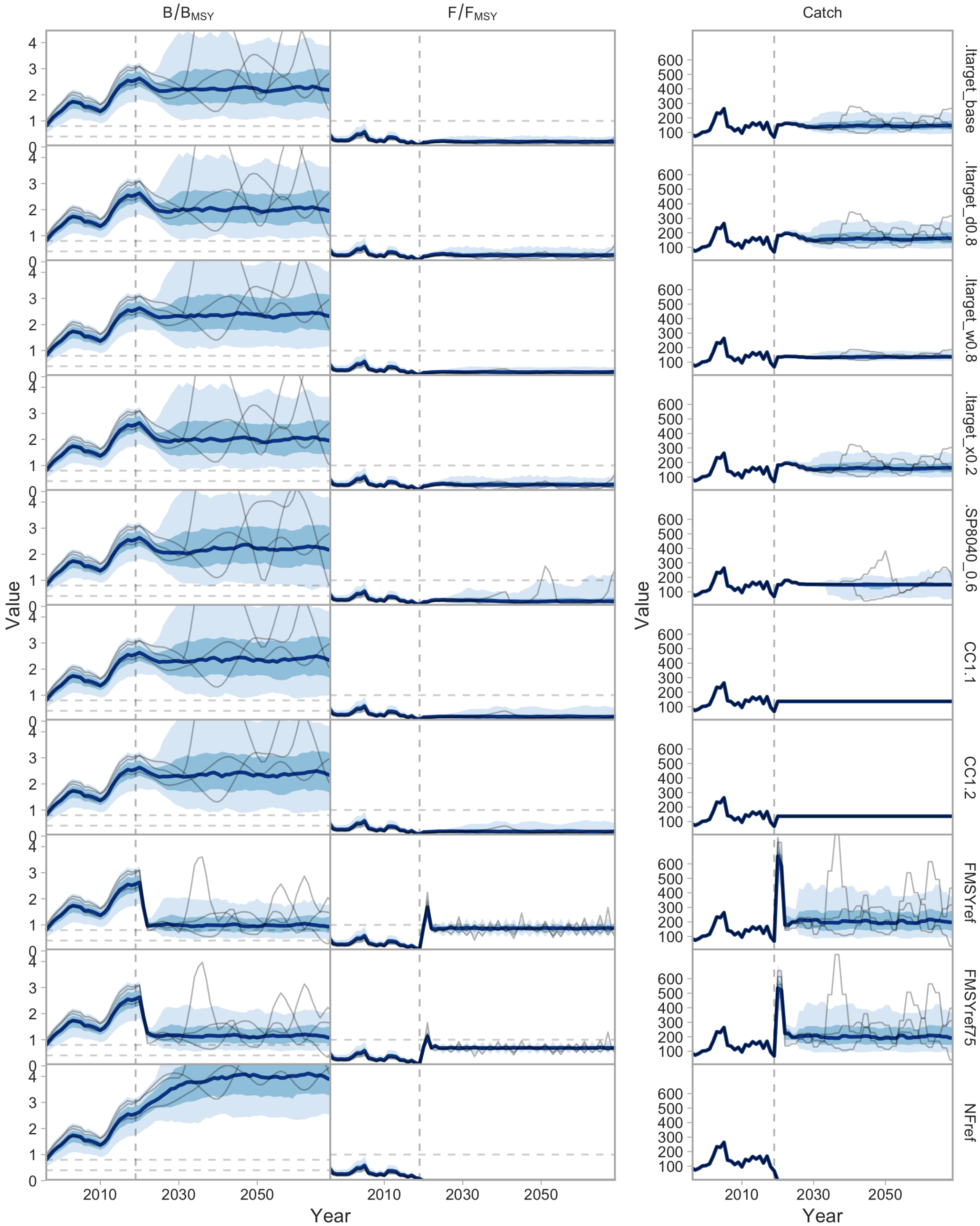

Figure E.21. Same as Figure E.20 but for the OM "C^{eq} 250%."

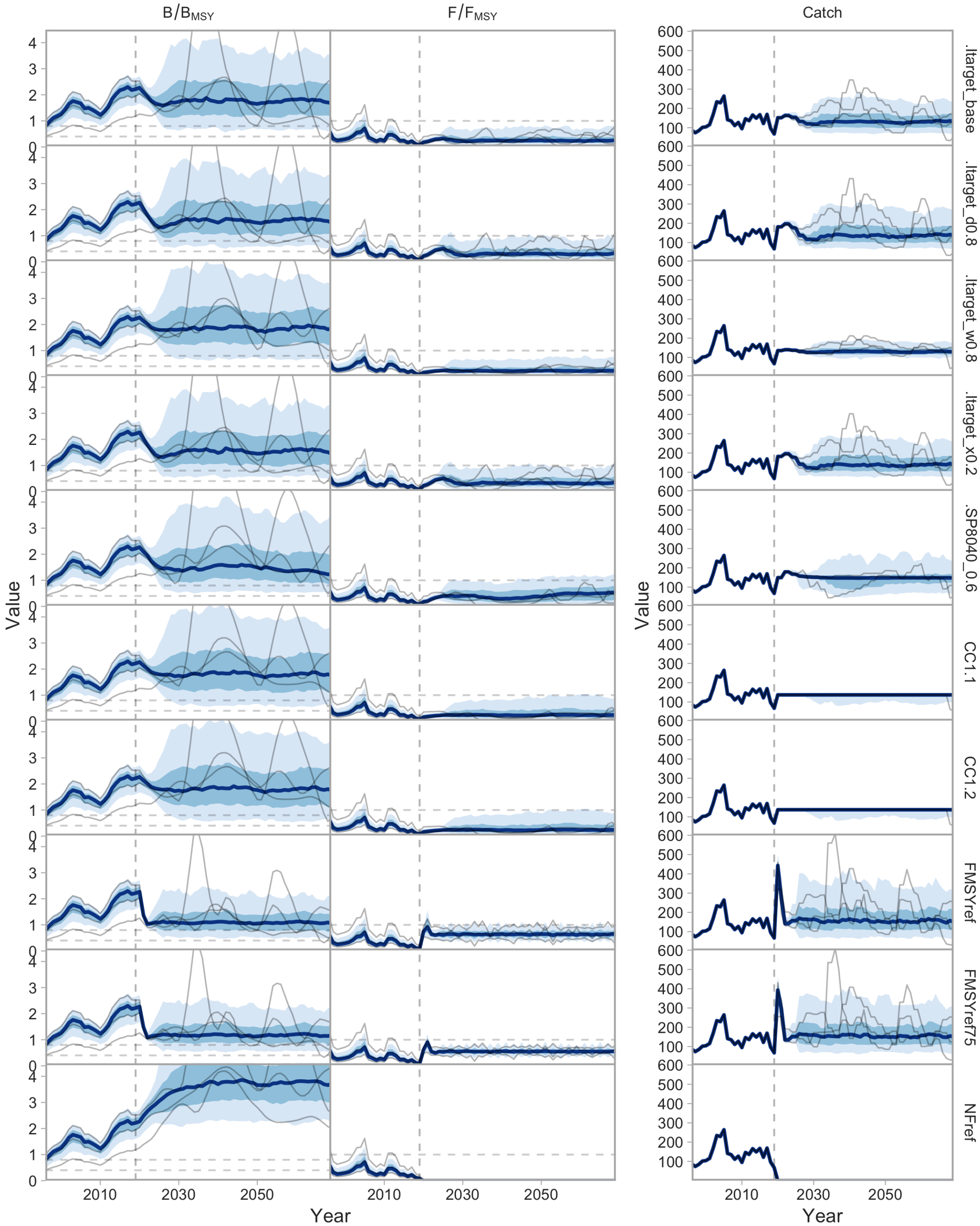

Figure E.22. Same as Figure E.20 but for the OM "Higher M."

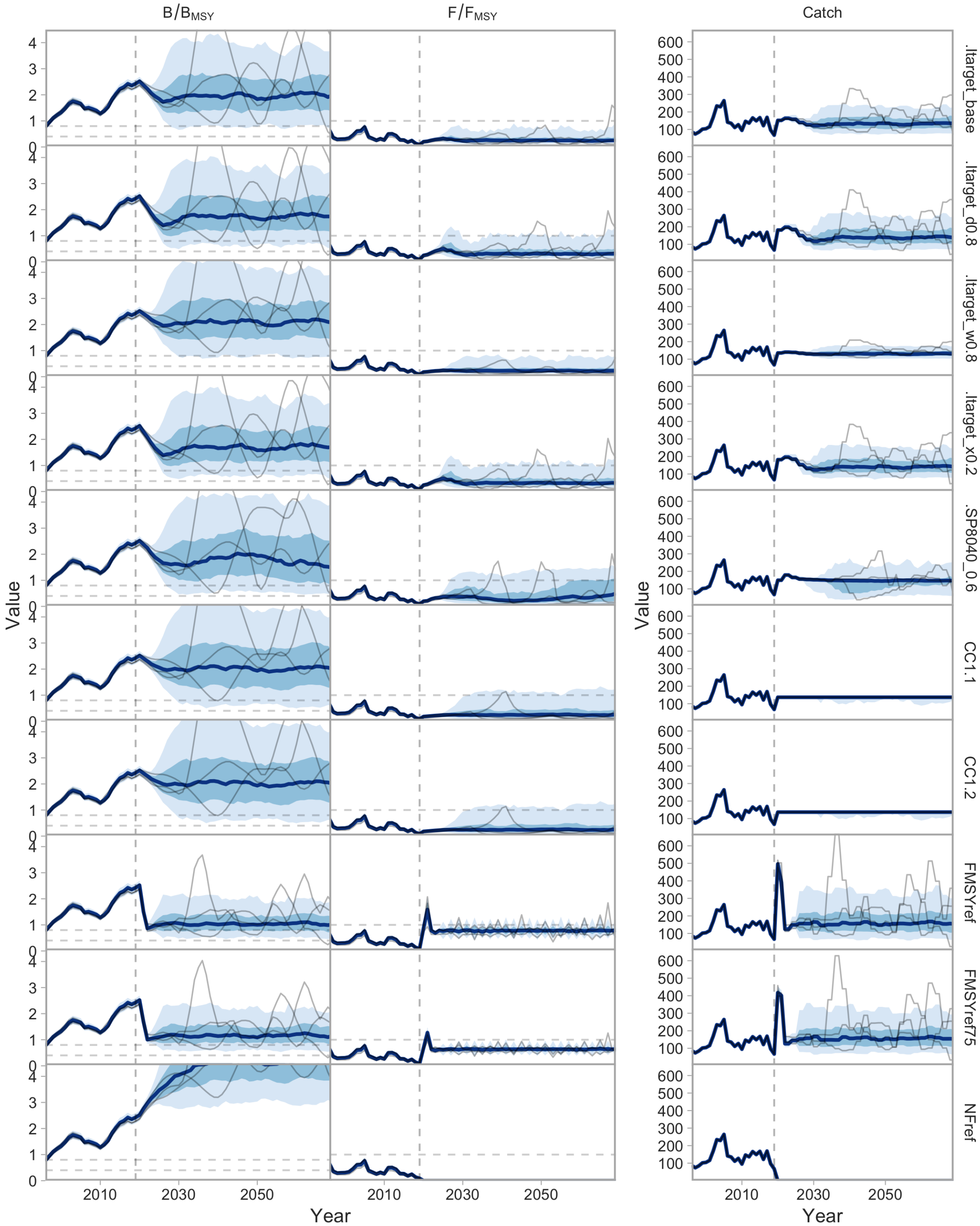

Figure E.23. Same as Figure E.20 but for the OM "Higher steepness."

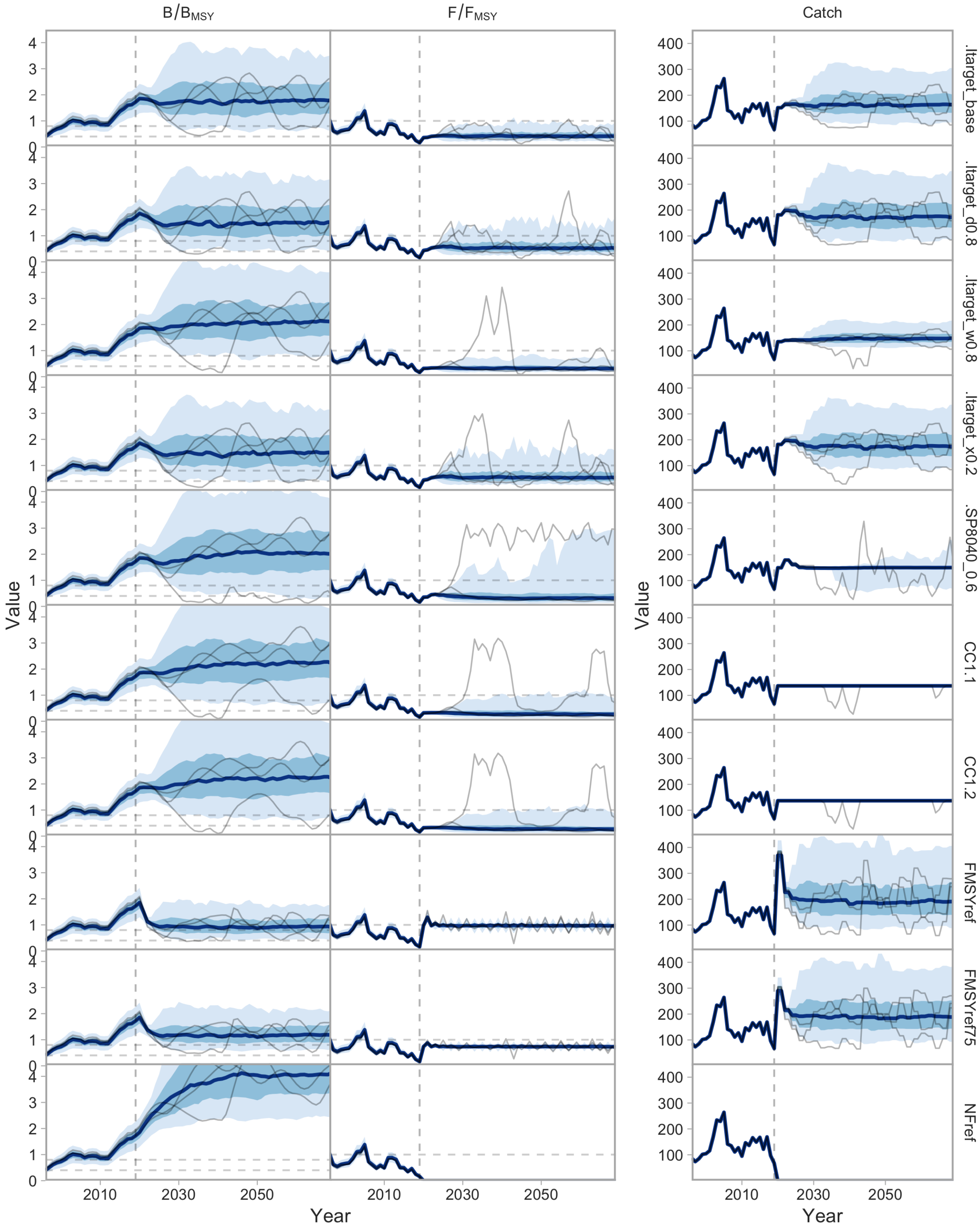

Figure E.24. Same as Figure E.20 but for the OM "Lower selectivity."

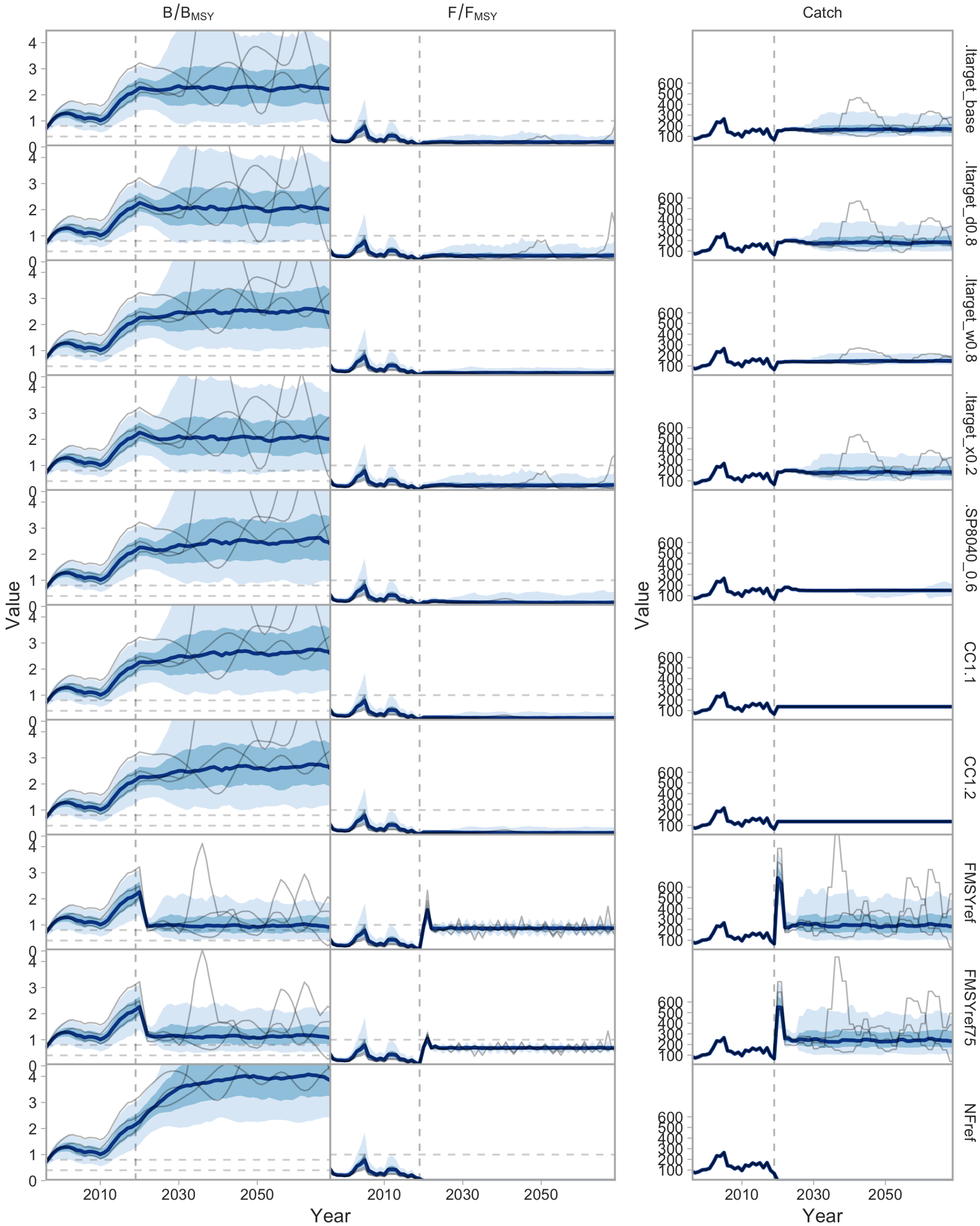

Figure E.25. Same as Figure E.20 but for the OM "No CPUE C^{eq} 250%."

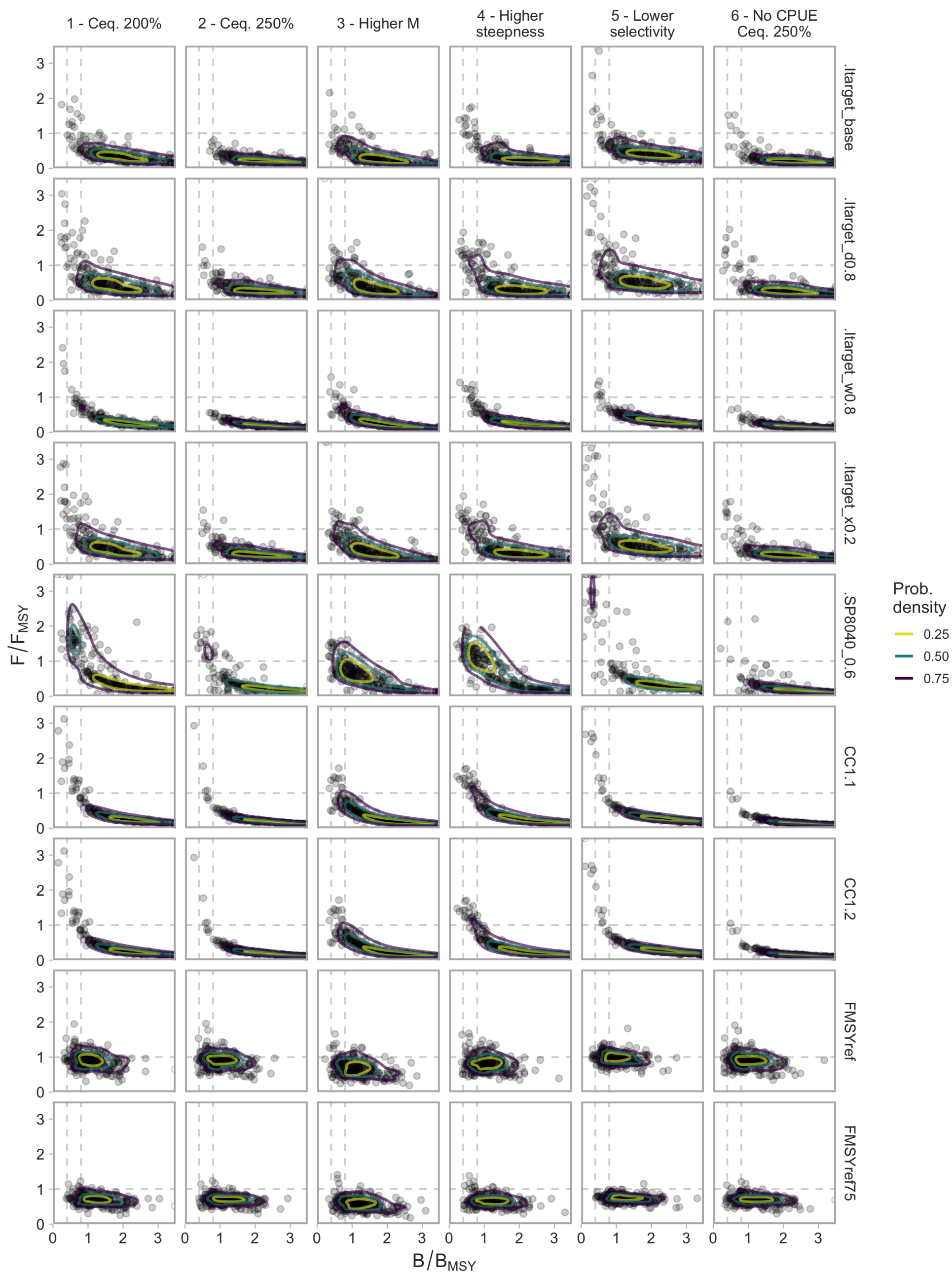

Figure E.26. B/B_{MSY} and F/F_{MSY} values from the final year of the projections across all replicates. Dots represent individual replicates. Vertical dashed lines show B/B_{MSY} = 0.4 (left) and 0.8 (right). Horizontal dashed line shows F/F_{MSY} = 1. Contour lines indicate two-dimensional kernel-density-smoothed quantiles at 0.25, 0.50, and 0.75 levels, calculated in log space.

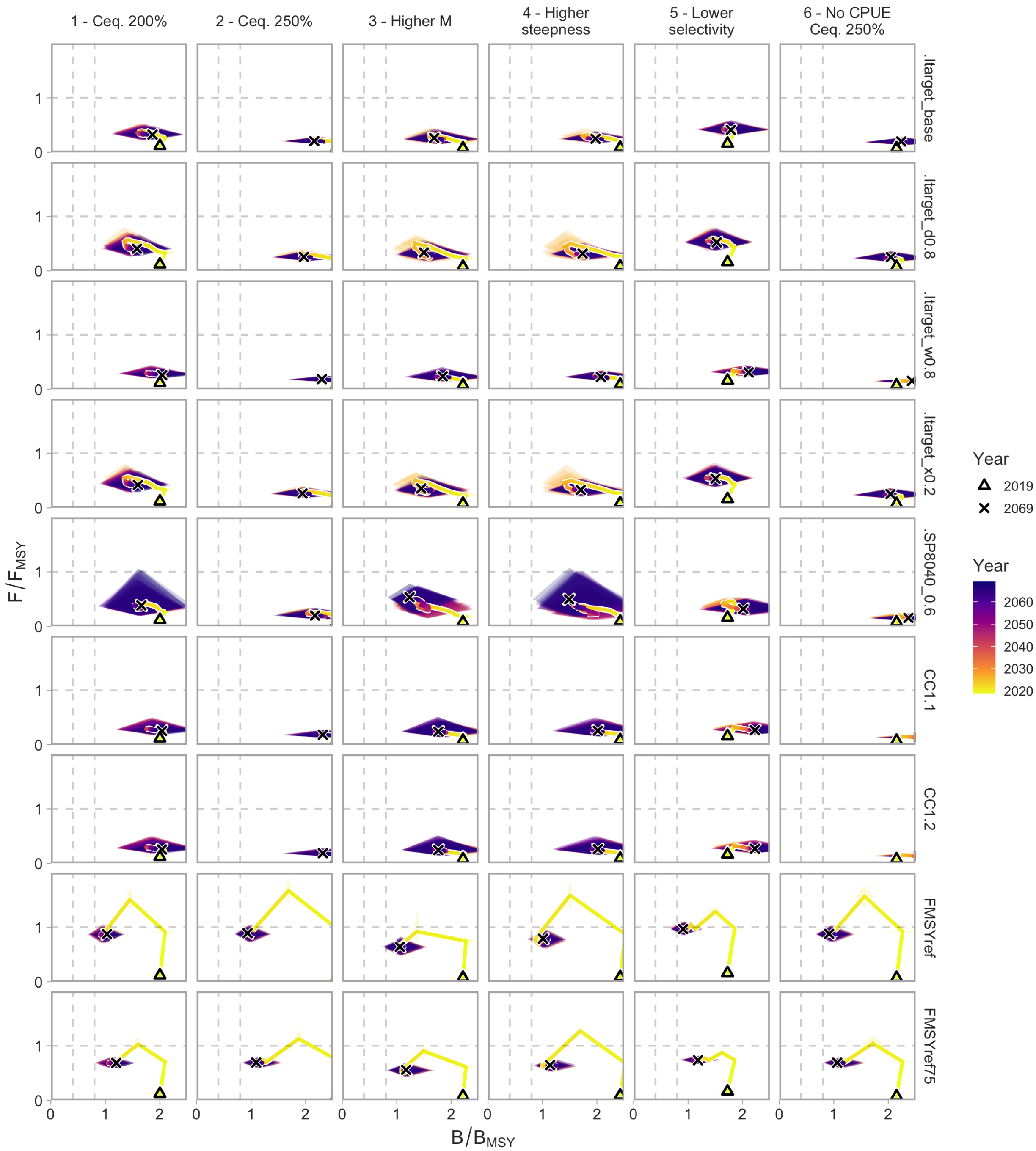

Figure E.27. Trajectory of B/B_{MSY} and F/F_{MSY} values summarized across replicates. The solid line corresponds to the median value. Vertical dashed lines show B/B_{MSY} = 0.4 (left) and 0.8 (right). Horizontal dashed line shows F/F_{MSY} = 1. Each diamond represents the 50% quantile of B/B_{MSY} (horizontal) and F/F_{MSY} (vertical).

E.7.2 ROBUSTNESS SET RESULTS

The "No CPUE C^{eq} 50%" OM robustness scenario represents markedly different biomass and fishing mortality historical trajectories to any of the OM reference-set scenarios. This OM scenario estimates the stock to be smaller (smaller R_0 values in Figure E.7) with high F_y (apical F by year) values around 2005 (Figure E.10). In this OM scenario, the Itarget MPs are able to achieve LT LRP > 0.60 (3 times out of 5), while the constant catch and surplus production model achieved LT LRP < 0.35 (7 times out of 20) (Figures E.28 and E.29). Other performance metrics show less contrast among the MPs, although LT USR also remains higher for the Itarget vs. constant-catch or surplus-production MPs (Figure E.28 and E.29).

This OM robustness scenario demonstrates stronger trade-offs among performance metrics than the OM reference-set scenarios (Figures E.30 and E.31). In particular, there is an obvious trade-off between STC and LT LRP (Figures E.30). The constant catch MPs achieved the highest STC values but at the expense of lower LT LRP values. The Itarget MPs struck more of a balance between LT LRP and STC (Figures E.30). The projection figures show the values underlying the performance metrics (Figure E.32). Within this scenario, over time, the .SP8040_0.6, CC1.1, and CC1.2 MPs all had a tendency to overfish the stock (Figure E.32).

For the "M increasing" OM scenario, the satisficed MPs were fairly robust to linear increases in natural mortality in the projection period, with the exception that all satisficed MPs had a considerably lower probability of achieving the long-term catch (LTC) metric (Figures E.28–E.31). Notably, the Itarget satisficed MPs maintained LT LRP > 0.9 despite increasing *M*, whereas the fixed-catch and surplus-production MPs all resulted in LT LRP < 0.9. The Itarget MPs were able to ramp down catch as *M* increased (Figure E.33) to maintain most replicates above the LRP.

By plotting the projections from all OM scenarios (Figures E.20–E.25 and E.32–E.33) on the same panels, we can examine the sensitivity of the projections to the various OM scenario conditions (Figure E.34). The historical trajectories were most sensitive to the assumptions of the "No CPUE C^{eq} 50%" OM scenario compared to the other OM scenarios. The projected trajectories were relatively similar across all OM reference-set scenarios, while differing for the "No CPUE C^{eq} 50%" and "M increasing" OM robustness scenarios. Note that sensitivity of the projections does not necessarily correspond to sensitivity of the rank order of MPs (Figure E.28), which is ultimately the important characteristic for decision making.

	7 - No CPUE Ceq. 50%						8 - M increasing					
	LT LRP	LT USR	FMSY	STC	LTC	AADC	LT LRP	LT USR	FMSY	STC	LTC	AADC
.Itarget_base	0.77	0.55	0.54	0.77	0.29	0.30	0.95	0.78	0.99	0.94	0.02	0.01
.Itarget_d0.8	0.70	0.45	0.40	0.73	0.29	0.17	0.94	0.75	0.97	0.94	0.04	<0.01
.Itarget_w0.8	0.64	0.44	0.51	0.85	0.26	0.62	0.92	0.68	0.98	0.98	0.02	0.16
.Itarget_x0.2	0.63	0.38	0.35	0.76	0.28	0.17	0.93	0.72	0.96	0.96	0.04	<0.01
CC1.1	0.34	0.21	0.32	0.97	0.32	0.32	0.84	0.54	0.94	>0.99	0.15	0.09
CC1.2	0.34	0.21	0.32	0.97	0.32	0.32	0.84	0.54	0.94	>0.99	0.15	0.09
.SP8040_0.6	0.29	0.12	0.19	0.88	0.23	0.17	0.87	0.59	0.94	0.99	0.12	0.07

Figure E.28. Performance of satisficed MPs for the robustness set OMs.

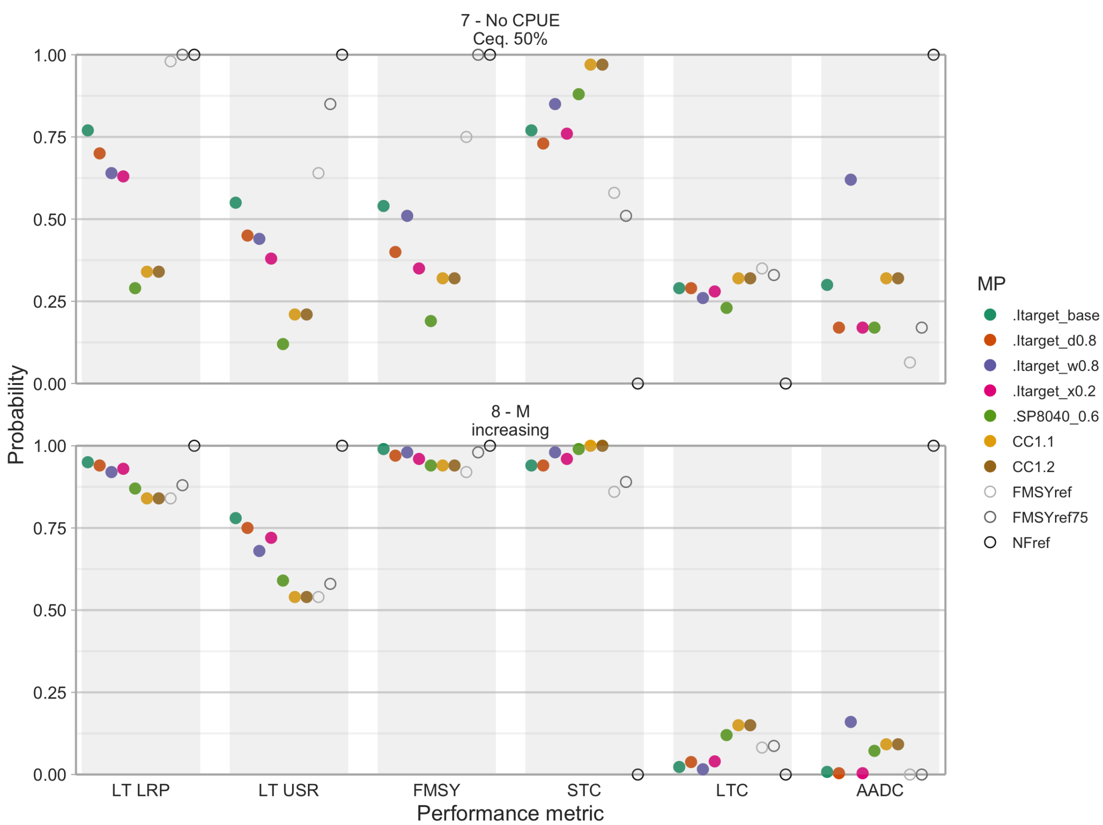

Figure E.29. Performance of satisficed MPs for the robustness set OMs.

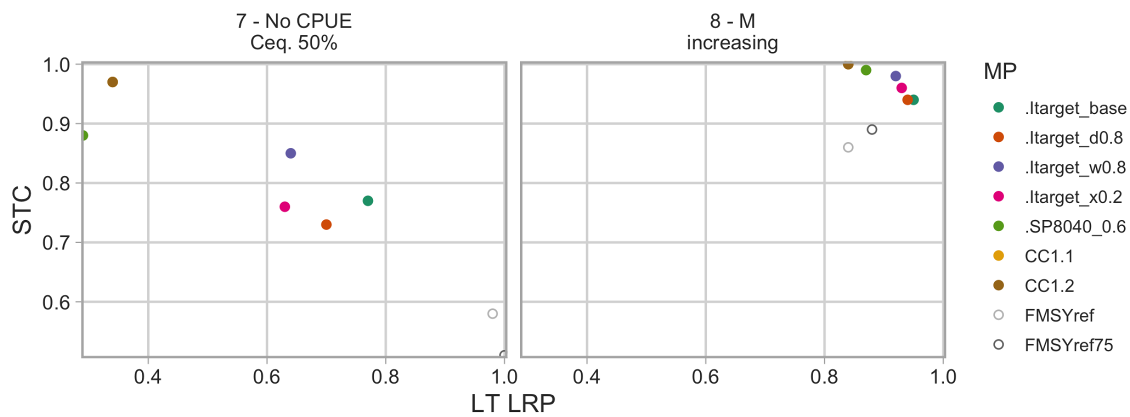

Figure E.30. Trade-off plot between LT LRP and STC performance metric values for satisficed MPs for the robustness set OMs.

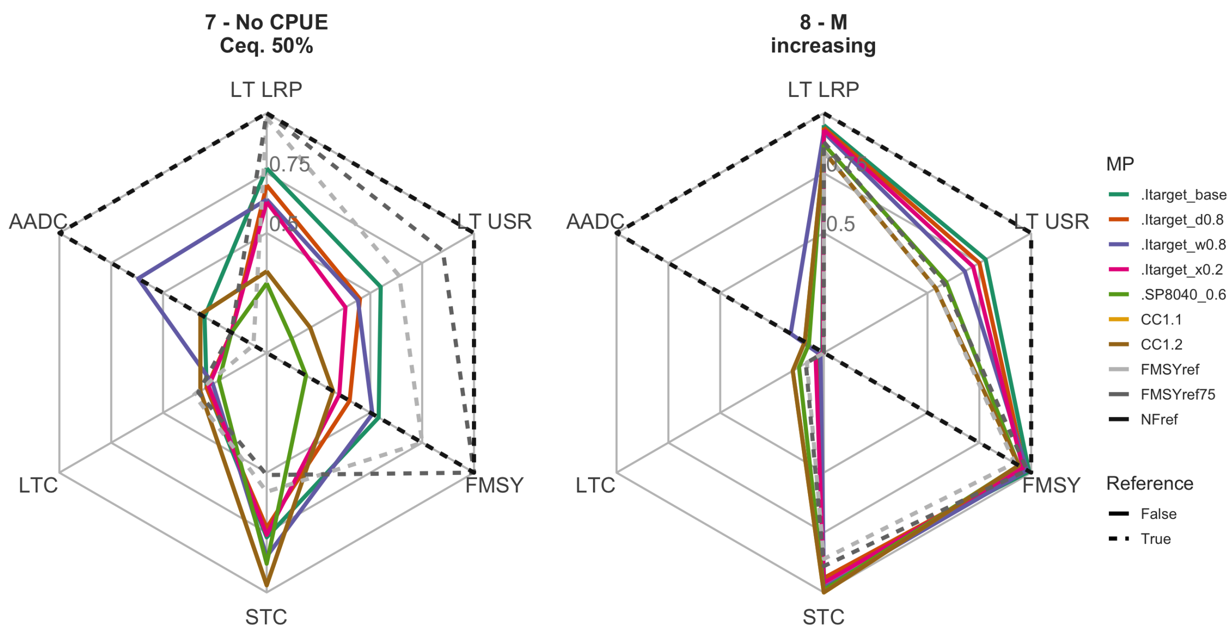

Figure E.31. Radar-plot representation of the trade-off in performance metric values for satisficed MPs for the robustness set OMs.

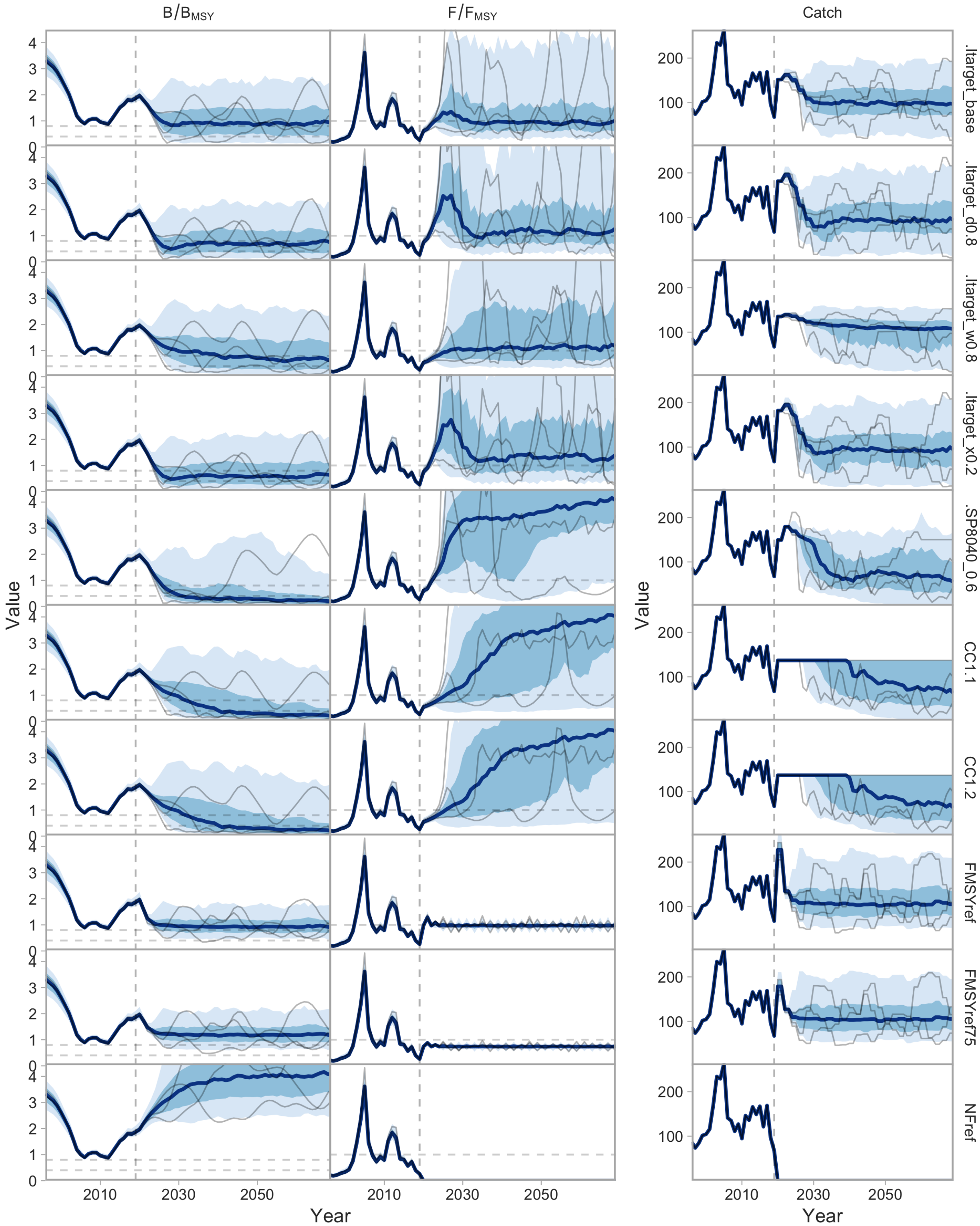

Figure E.32. Same as Figure E.20 but for the robustness OM "No CPUE C^{eq} 50%."

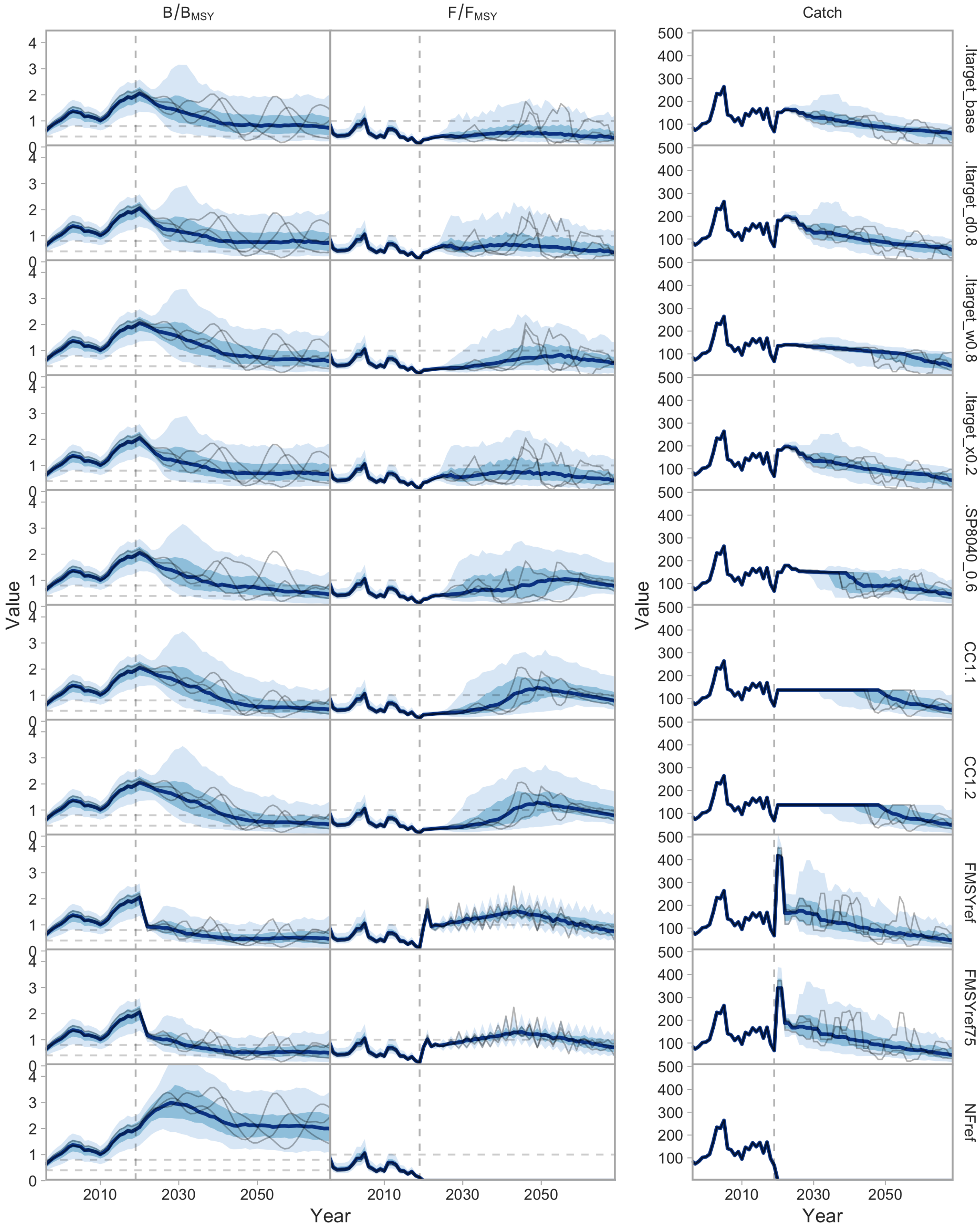

Figure E.33. Same as Figure E.20 but for the robustness OM "Increasing M."

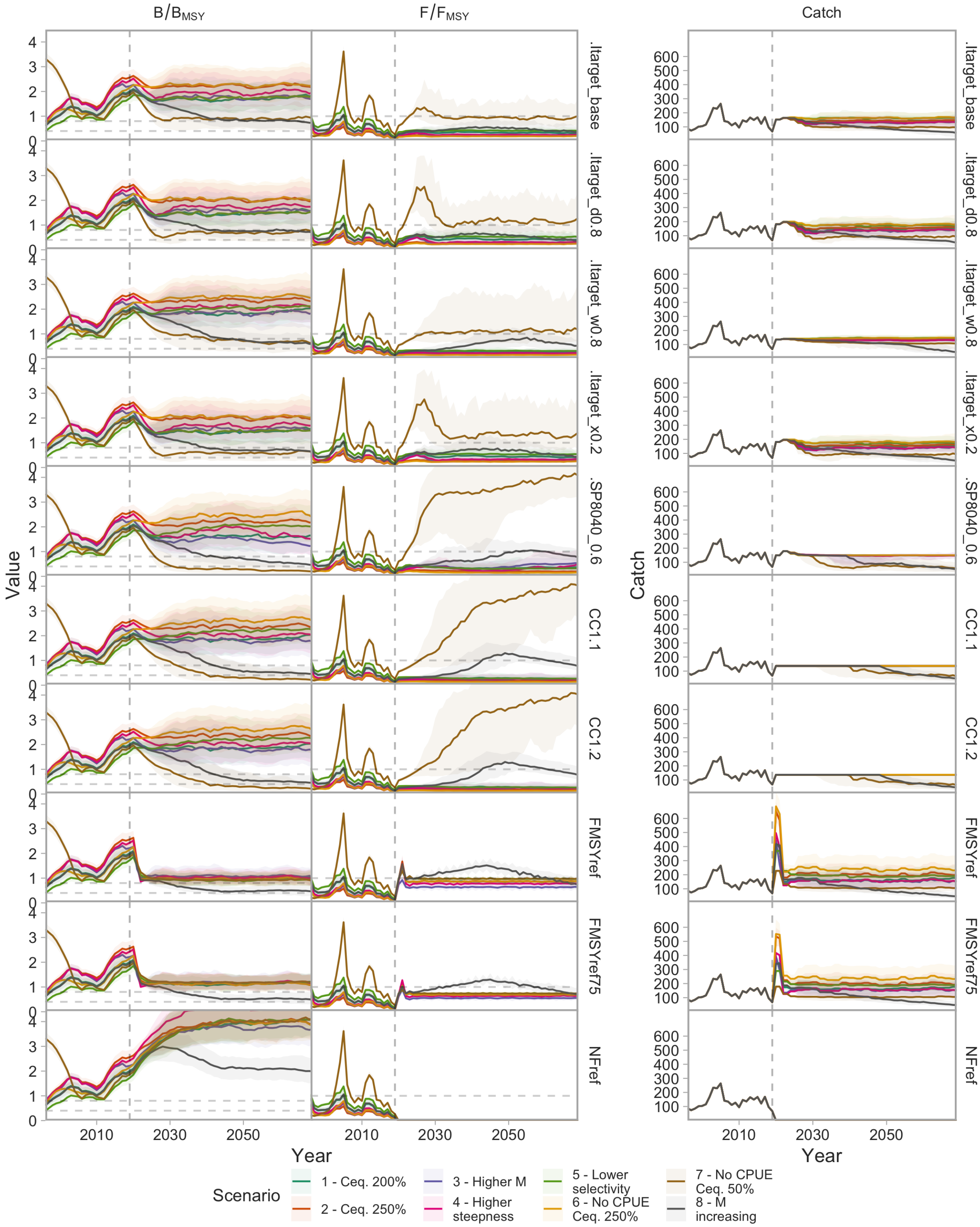

Figure E.34. B/B$_{MSY}$, F/F$_{MSY}$, and catch from the historical and projected time periods. Colours represent the reference- and OM robustness-set scenarios. Lines represent medians and shaded regions represent 50% quantiles. Satisficed and reference MPs are shown from top to bottom. Vertical dashed lines represent 2019.

APPENDIX F. REX SOLE OPERATING MODEL DEFINITION

This section describes the specification of the Rex Sole operating model. Some OM scenarios in the reference and robustness sets assign alternative options for some of the parameters. We note this below where it occurs.

F.1 STOCK SLOT DESCRIPTIONS

We begin by setting the random seed for sampling from distributions below:

```
set.seed(282943)
```

F.1.1 COMMON_NAME

Common name of the species.

```
stock_rex@Common_Name <- "Rex Sole"
```

F.1.2 SPECIES

Scientific name of the species.

```
stock_rex@Species <- "Glyptocephalus zachirus"
```

F.1.3 MAXAGE

The maximum age of individuals that is simulated. Positive integer.

The maximum recorded age for a Rex Sole in BC waters is 15 in the Hecate Strait Multispecies Assemblage survey in 1998. We do not have ages in our database for 3CD to inform the maximum age for this region. Munk (2001) note the maximum age observed for Rex Sole in the Gulf of Alaska is 27 years. This maximum age becomes the plus group when the SRA is run.

```
stock_rex@maxage <- 27
```

F.1.4 M

Natural mortality rate.

McGilliard and Palsson (2017) used 0.17 y^{-1} for M but they also note that growth rates and size can vary with latitude, which can affect M. M is likely slightly higher in BC at lower latitudes. Here, we add a range of possible M values starting with a lower bound of that from McGilliard and Palsson (2017). The maximum age observed in our database of 15 implies a higher value of M, although this maximum age is likely an underestimate. We include an alternative OM scenario with a higher M value (M = 0.3 y^{-1}).

```
stock_rex@M <- c(0.17, 0.25)
```

F.1.5 H

Steepness of the stock-recruit relationship.

Maunder (2012) cite Myers et al. (1999) who found a median value of steepness of 0.80 for flatfish and note that, because of the bias in the value estimated for North Atlantic Cod in that paper, the value might be closer to 0.94. Grandin and Forrest (2017) estimated steepness for Arrowtooth Flounder, but found their posterior was highly influenced by their prior (they estimated a median of 0.88 with a 95% confidence interval of 0.69–0.98, which nearly matches the prior). Here we will use the same prior used in Grandin and Forrest (2017) (Figure F.1) but we also include an alternative OM scenario with steepness fixed to a relatively high value of 0.95.

```
cpars$h <- rbeta(nsim, 13.4, 2.4)
```

Note that if the samples of h were lower, they would ideally be constrained to $[0.2, 1)$ via a transformation. For example: `0.8 * rbeta(nsim, 13.4, 2.4) + 0.2`.

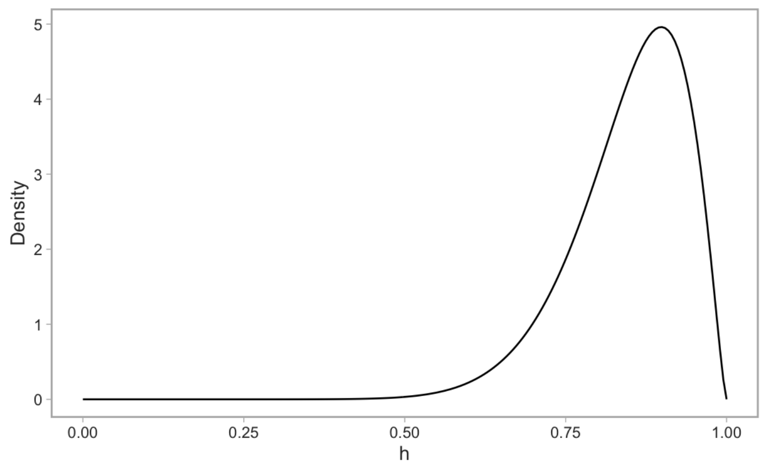

Figure F.1. Density of h prior.

F.1.6 PERR

Process error, the CV of lognormal recruitment deviations.

Thorson et al. (2014) found the mean standard deviation of log recruitment deviations for Pleuronectiformes to be about 0.64. McGilliard and Palsson (2017) used 0.6 for Gulf of Alaska stocks. We will use a normal distribution centered around this value of 0.6 with a bit of variation represented by an SD of 0.1:

```
cpars$Perr <- rnorm(nsim, 0.6, 0.1)
```

F.1.7 AC

Autocorrelation in recruitment deviations.

This parameter will be replaced by the stock reduction analysis (SRA) model based on a post-hoc calculation performed on the estimated historical recruitment deviations.

```
cpars$AC <- c(0, 1)
```

F.1.8 LINF

Maximum length.

Although we do not have any ages in our database for WCVI, there exist a limited number of ages from the Hecate Strait Multispecies Assemblage Survey and a von Bertalanffy model fits reasonably well (Figure F.2). We will extract samples from the model posterior (Figure F.3) and save them as custom parameters.

We fit the model with Stan (Stan Development Team 2020) using the model implemented in the gfplot R package (Anderson et al. 2020c) according to the equation:

$$L_i \sim \text{Log-normal}\left(\log(l_\infty(1 - \exp(-k(A_i - t_0)))), \sigma\right), \tag{F.1}$$

where L_i and A_i refer to the length and age of fish i, σ refers to the scale parameter of the log-normal distribution, and all other parameters refer to the von Bertalanffy equation. We used weakly informative priors: $k \sim \text{Normal}(0, 2)$, $L_{\text{inf}} \sim \text{Normal}(0, \varphi)$, $t_0 \sim \text{Normal}(0, 20)$, and $\sigma \sim \text{Student-t}(3, 0, 2)$, where φ refers to the 99% quantile of the observed lengths (36 cm). We sampled 8000 iterations across four chains from the posterior and checked that the chains were consistent with convergence via the Rhat (potential scale reduction factor) and ESS (effective sample size) metrics along with trace plots of the samples (Gelman et al. 2014).

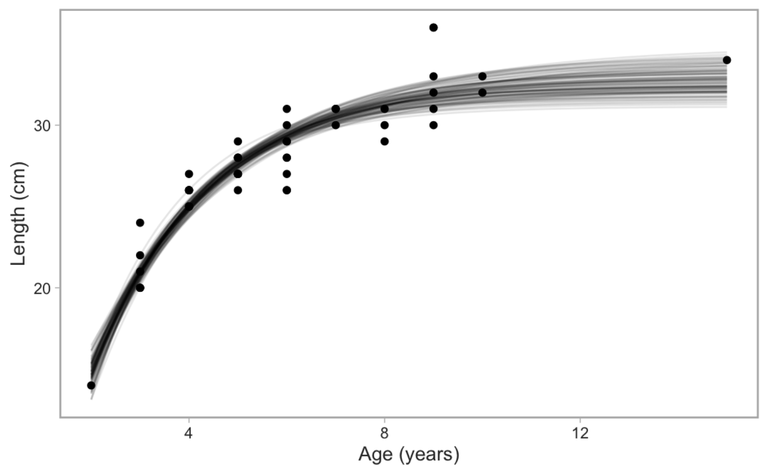

Figure F.2. von Bertalanffy model fit to lengths and ages from the Hecate Strait Multispecies Assemblage Survey. Dots represent individual fish and lines represent 100 samples from the posterior.

```r
msa <- gfdata::get_survey_samples("rex sole", ssid = 2)
vb_model <- gfplot::fit_vb(msa, sex = "all", method = "mcmc", iter = 8000)
vb_post <- rstan::extract(vb_model$model)
set.seed(3829)
i <- sample(seq_along(vb_post$k), nsim)
```

```r
cpars$Linf <- as.numeric(vb_post$linf[i])
```

F.1.9 K

von Bertalanffy growth parameter k.

```r
cpars$K <- as.numeric(vb_post$k[i])
```

F.1.10 T0

von Bertalanffy theoretical age at length zero.

The t_0 parameter estimated in our model is slightly positive, although there are very few fish under the age of 4 in this data set to inform the model.

```
cpars$t0 <- as.numeric(vb_post$t0[i])
```

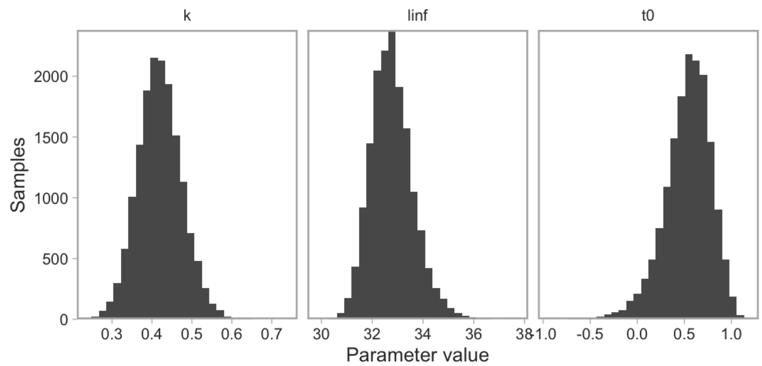

Figure F.3. von Bertalanffy parameter posteriors.

F.1.11 L50

Length at 50% maturity.

We have sufficient maturity and length data to fit a logistic maturity model. McGilliard and Palsson (2017) noted that length at maturity varied by latitude, but that age at maturity was fairly consistent among stocks.

Although we visualize the relationship for male and female fish separately (Figure F.4), we fit a Bayesian logistic GLM (generalized linear model) with rstanarm (Goodrich et al. 2018) to sample parameter values from a model that estimates maturity for an average male/female fish from our samples (Figure F.6). We used priors of Normal(0, 50) for the intercept and Normal(0, 10) for the effect of length and male/female. We sampled 2000 iterations for each of four chains from the posterior and check that the chains were consistent with convergence using the same approach as described above for the von Bertalanffy model.

```
m_mat <- gfplot::fit_mat_ogive(drex$survey_samples, type = "length")
d <- m_mat$data
d$sex <- d$female - mean(d$female)
d$length <- d$age_or_length
mat_bayes <- rstanarm::stan_glm(mature ~ length * sex,
  data = d, family = binomial(link = "logit"),
  cores = parallel::detectCores(), iter = 2000,
  prior_intercept = rstanarm::normal(0, scale = 50, autoscale = FALSE),
  prior = rstanarm::normal(0, scale = 10, autoscale = FALSE),
  seed = 8983
)
post_mat <- as.data.frame(mat_bayes)
logit_perc <- function(a, b, perc) -(log((1 / perc) - 1) + a) / b
L50 <- logit_perc(a = post_mat$`(Intercept)`, b = post_mat$length, perc = 0.5)
i <- sample(seq_along(post_mat$length), nsim)
```

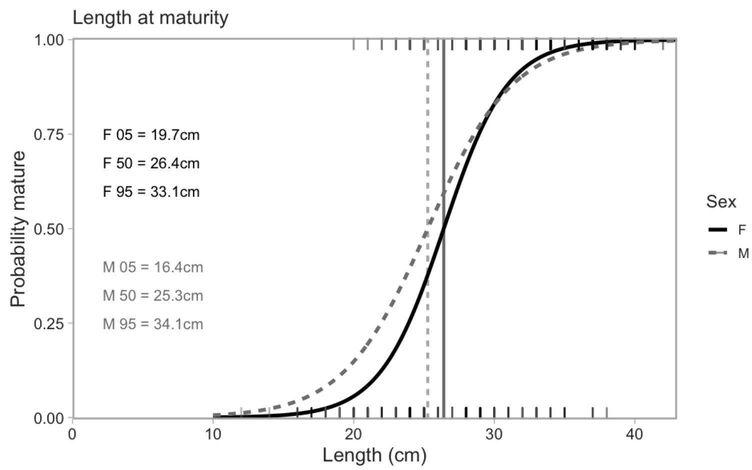

Figure F.4. Predicted maturity by length for males and females separately. "Rug lines" along the top and bottom indicate individual sampled fish with some overlap. Curved lines represent predictions from a logistic regression. Vertical line represents the length at 50% maturity.

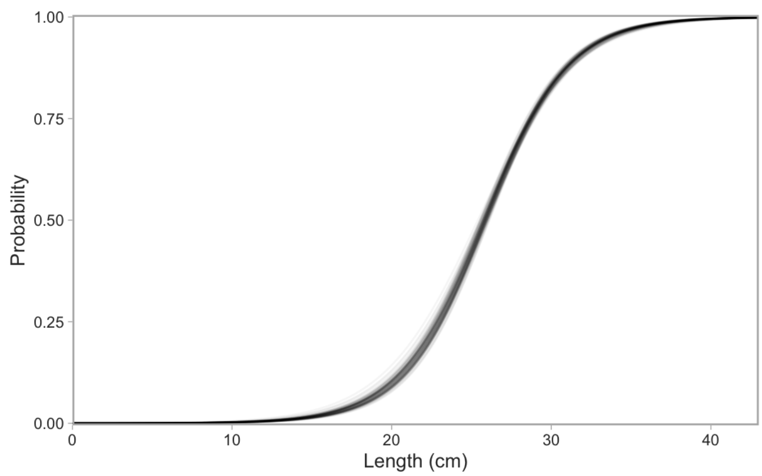

Figure F.5. Predicted maturity by length from the Bayesian model for a fish of average male/female sex. Individual thin translucent lines represent 100 draws from the posterior.

```
cpars$L50 <- L50[i]
```

Our estimated average length at 50% maturity (25.8 cm) is between the length at 50% maturity reported for Rex Sole off Oregon (24cm) and the Gulf of Alaska (35cm) (Abookire 2006).

F.1.12 L50_95

Length increment from 50% to 95% maturity.

We can calculate this as a derived parameter from our logistic maturity model:

```
L95 <- logit_perc(a = post_mat$`(Intercept)`, b = post_mat$length, perc = 0.95)
```

```
L50_95 <- L95 - L50
cpars$L50_95 <- L50_95[i]
```

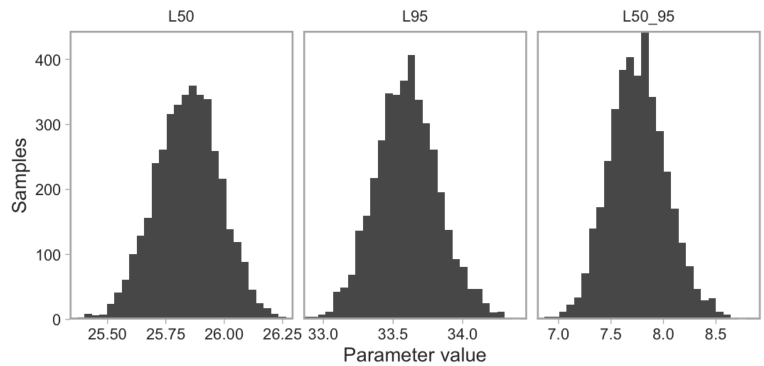

Figure F.6. Posterior samples of length at maturity parameters. L50 represents length at 50% maturity, L90 represents length at 95% maturity, and L50_95 represents the length difference between L50 and L95.

F.1.13 D

Current level of stock depletion $SSB_{current}/SSB_{unfished}$.

This will be populated via the SRA conditioning model and so we will leave it blank here.

```
stock_rex@D
#> numeric(0)
```

F.1.14 A

Length-weight parameter alpha. Positive real number.

We have sufficient length and weight data to fit a model to our data for both sexes combined:

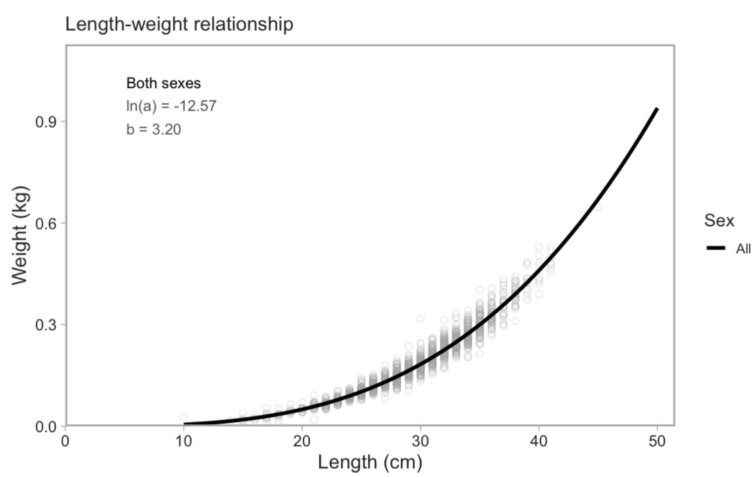

Figure F.7. Length-weight relationship for Rex Sole sampled from the WCVI Synoptic Survey with both sexes combined. The model is fit as in Anderson et al. (2019) (regression of log(weight) vs. log(length)) using the R package TMB (Kristensen et al. 2016) and a Student-t observation model with degrees of freedom of 3 to downweight any outliers.

```
mlw <- gfplot::fit_length_weight(drex$survey_samples, sex = "all")
gfplot::plot_length_weight(object_all = mlw, col = c("All" = "black"))
```

```
stock_rex@a <- exp(mlw$pars[["log_a"]])
round(log(stock_rex@a), 2)
#> [1] -12.57
round(stock_rex@a, 7)
#> [1] 3.5e-06
```

F.1.15 B

Length-weight parameter beta. Positive real number.

```
stock_rex@b <- mlw$pars[["b"]]
round(stock_rex@b, 2)
#> [1] 3.2
```

F.2 FLEET SLOT DESCRIPTIONS

F.2.1 NYEARS

The number of years for the historical spool-up simulation.

We will set the historical run-up to start in 1996.

```
catch_yrs <- unique(drex$catch$year)
c(catch_yrs[1], catch_yrs[length(catch_yrs)])
#> [1] 1996 2019
fleet_rex@nyears <- length(catch_yrs)
fleet_rex@nyears
#> [1] 24
```

F.2.2 L5

Shortest length corresponding to 5% vulnerability.

There are only 42 fish with measured lengths in the WCVI commercial data, which is insufficient to estimate selectivity in the SRA. We will make the assumption that commercial and survey selectivity approximately match the maturity ogive (but are fully selected by L_∞), similar to the assumption made in a delay-difference model that maturity and vulnerability to the fishery occur at the same age. For this we use $L^5 = 22$ and $L^{\mathrm{FS}} = 32$. We will include an alternative assumption that the selectivity curve is shifted to the left making younger fish vulnerable to the fishery ($L^5 = 17$ and $L^{\mathrm{FS}} = 28$). These curves will be specified in the SRA.

```
fleet_rex@L5 <- c(22, 22)
```

F.2.3 LFS

Shortest length that is fully vulnerable to fishing.

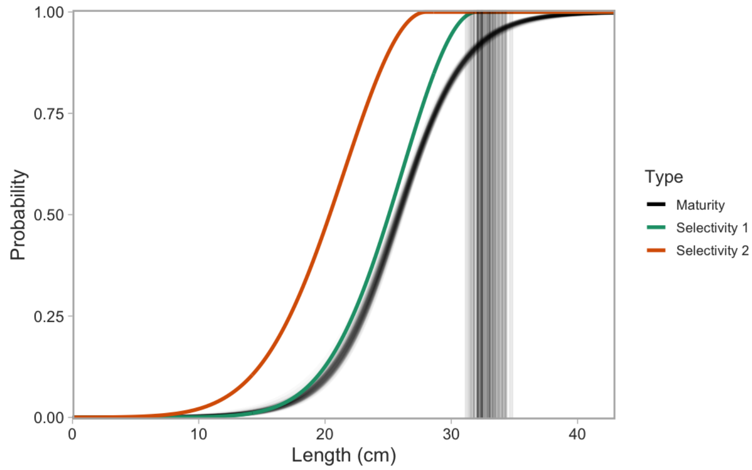

Figure F.8. Maturity-at-length and 2 selectivity curves used in the operating models. Maturity ogive represents 100 draws from the posterior. Vertical lines represent samples from L_∞.

```
fleet_rex@LFS <- c(34, 34)
```

F.2.4 CURRENTYR

The current calendar year (final year) of the historical simulations.

```
fleet_rex@CurrentYr <- 2019
```

F.3 OBS SLOT DESCRIPTIONS

F.3.1 COBS

Log-normal catch observation error expressed as a CV. Uniform distribution lower and upper bounds.

Since 1996, observation error on catch should be negligible.

```
obs_rex@Cobs <- c(0, 0)
```

F.3.2 CBIASCV

Log-normal CV controlling the sampling of bias in catch observations for each simulation. Uniform distribution lower and upper bounds.

We will keep this at 0 because observation error on catch should be very small with 100% on board and dockside monitoring.

```
obs_rex@Cbiascv <- c(0, 0)
```

F.3.3 IOBS

Observation error in the relative abundance indices expressed as a CV. Uniform distribution lower and upper bounds.

We will sample with replacement from the CVs of the observed WCVI synoptic trawl survey index values for Rex Sole.

```
set.seed(2943)
cpars$Iobs <- sample(drex$survey_index$re, size = nsim, replace = TRUE)
```

```
round(mean(cpars$Iobs), 2)
#> [1] 0.09
round(sort(unique(cpars$Iobs)), 2)
#> [1] 0.07 0.07 0.08 0.08 0.09 0.09 0.11 0.12
```

F.3.4 BETA

A parameter controlling hyperstability/hyperdepletion where values below 1 lead to hyperstability (an index that decreases more slowly than true abundance) and values above 1 lead to hyperdepletion (an index that decreases more rapidly than true abundance). Uniform distribution lower and upper bounds.

We will set this to 1 to imply no hyperstability or hyperdepletion of the synoptic trawl survey population index.

```
obs_rex@beta <- c(1, 1)
```

F.4 IMP SLOT DESCRIPTIONS

F.4.1 TACFRAC

Mean fraction of TAC taken. Uniform distribution lower and upper bounds.

We will assume that there is no persistent bias from replicate to replicate of the fraction of TAC taken.

```
imp_rex@TACFrac <- c(1.0, 1.0)
```

F.4.2 TACSD

Log-normal CV in the fraction of TAC taken. Uniform distribution lower and upper bounds.

We will assume there is no variation in the fraction of TAC taken from year to year.

```
imp_rex@TACSD <- c(0, 0)
```

APPENDIX G. R PACKAGE TO SUPPORT THE MP FRAMEWORK

We developed the R package ggmse (originally named 'gfdlm') to provide a set of tools that facilitate applications of this framework. The package includes our provisional library of MPs, a number of plotting functions for summarizing simulation output and model fits, and a variety of other utilities for manipulating DLMtool MPs and performance metrics, and creating reports based on DLMtool simulations (such as Appendix F in this document).

G.1 PLOTTING FUNCTIONS

All plotting functions are set up to take either:

1. a named list of DLMtool simulations from `DLMtool::runMSE()`,
2. a named list of SRA model fits from `MSEtool::SRA_scope()`, or
3. a named list of performance metric data frames from `ggmse::get_probs()`.

In all cases, the list elements should contain different OM scenarios from the reference and/or robustness sets and the names should reflect the OM scenario names.

Note that the `pkg::f()` syntax refers to a function `f()` within the package `pkg` without loading the package via the `library()` function. We have written code in this appendix using this syntax so that it is clear where various functions originate.

The following are the main plotting functions:

- `plot_index_fits()` takes a named list of SRA model fits and creates a plot of historical population index data and model fits. For example, Figure E.6.
- `plot_convergence()` takes a named list of DLMtool simulations and creates a plot to assess replicate convergence. For example, Figure E.12.
- `plot_tigure()` takes a named list of performance metric data frames and creates a shaded probability table of performance metrics. For example, Figure E.13.
- `plot_dots()` takes a named list of performance metric data frames and creates a dot or dot-and-line plot of performance metrics. For example, Figure E.16.
- `plot_tradeoff()` takes a named list of performance metric data frames and creates a trade-off plot of two selected performance metrics. For example, Figure E.17.
- `plot_radar()` takes a named list of performance metric data frames and creates a radar plot. For example, Figure E.18.
- `plot_index()` takes a named list of DLMtool simulations and creates a plot showing historical and projected population index values. For example, Figure E.19.
- `plot_main_projections()` takes a named list of DLMtool simulations and creates a time-series plot of historical and projected B/B_{MSY}, F/F_{MSY}, and catch. For example, Figure E.20.
- `plot_kobe()` takes a named list of DLMtool simulations and creates a Kobe plot (F/F_{MSY} vs. B/B_{MSY}) for the last projected year. For example, Figure E.26.
- `plot_worms()` takes a named list of DLMtool simulations and creates a Kobe plot through time. For example, Figure E.27.
- `plot_scenario_projections()` takes a named list of DLMtool simulations and creates a time-series plot of historical and projected B/B_{MSY}, F/F_{MSY}, and catch with all OM scenarios overlaid on the same plots using different colours. For example, Figure E.34.

- `plot_factory()` is "the one [plotting function] to rule them all." This function takes a named list of DLMtool simulations along with a number of other arguments specifying characteristics like satisficed MPs and performance metrics and generates common versions of all the figures described above. The output is a named list of ggplot2 (Wickham 2016) elements that can be viewed, modified, and/or saved for inclusion in a report. For example, this function was used to create all the main figures in the Rex Sole case study in this document (Figures E.12–E.34).
- `pm_factory()` is a function factory that generates a performance metric function. For example, the following creates a performance metric `LT LRP` that is based on spawning biomass, a reference level of 0.4 of B_{MSY}, and calculated over years 36–50 in the projection period.

```
`LT LRP` <- ggmse::pm_factory("SBMSY", ref = 0.4, yrs = c(36, 50))
```

G.2 MP MANIPULATION FUNCTIONS

`use_AddInd()` and `reduce_survey()` are function factories that modify MPs to use the DLMtool "additional index" data and discard selected survey years, respectively. The additional index data (as used in this framework) reflects the observed index in the historical period and an index that reflects the survey or fleet selectivity pattern in the projection period. For example, the following creates a new MP named `.Itarget_base` by taking the MP `Itarget_base` and (1) modifying it to use the first "additional index" and (2) omitting every even year of the survey to emulate a biennial survey:

```
oddify <- function(x) seq(2, x, by = 2)
.Itarget_base <- ggmse::Itarget_base %>%
  ggmse::use_AddInd() %>%
  ggmse::reduce_survey(index = oddify)
```

Note that `%>%` refers to the "pipe" function from the magrittr R package (Bache and Wickham 2014). It can be read as "and then." Technically, it places the output from the function before the `%>%` into the first argument of the function after the `%>%`.

G.3 REPORTING UTILITIES

`create_rmd()` creates or updates a template for specifying a DLMtool OM. For example, it was used to create Appendix F.

APPENDIX H. COMPUTATIONAL ENVIRONMENT

This version of the document was generated on 2021-08-13 11:19:41 with R version 4.1.0 (2021-05-18) (R Core Team 2019) and R package versions:

	Package	Version	Date
bookdown	bookdown	0.22	2021-04-22
cowplot	cowplot	1.1.1	2020-12-30
csasdown	csasdown	0.0.10.9000	2021-07-08
DLMtool	DLMtool	5.4.3	2021-08-13
dplyr	dplyr	1.0.7	2021-06-18
gfdata	gfdata	0.0.0.9000	2021-07-05
gfdlm	gfdlm	0.0.1.9001	2021-08-13
gfplot	gfplot	0.1.4	2021-07-06
ggplot2	ggplot2	3.3.5	2021-06-25
glmmTMB	glmmTMB	1.1.2	2021-07-09
kableExtra	kableExtra	1.3.4	2021-02-20
knitr	knitr	1.33	2021-04-24
MSEtool	MSEtool	1.6.0	2020-05-05
purrr	purrr	0.3.4	2020-04-17
rmarkdown	rmarkdown	2.9	2021-06-15
rstan	rstan	2.21.2	2020-07-27
rstanarm	rstanarm	2.21.1	2020-07-20
TMB	TMB	1.7.20	2021-04-08

The source code for this document is available at:
https://github.com/pbs-assess/gfmp/tree/f524374.

This document was compiled with the R package csasdown (Anderson et al. 2020a).