

Design of management procedures for black grouper fisheries: an exploratory Management Strategy Evaluation (MSE)



FINAL

April 09, 2026

Vania Henriquez and William Harford

NATURE ANALYTICS

551 Lakeshore Rd E, Suite 105, Mississauga, Ontario, L5G 0A8, Canada

www.natureanalytics.ca

Report prepared for
Florida Fish and Wildlife Conservation Commission
Fish and Wildlife Research Institute

Executive summary

Management of black grouper (*Mycteroperca bonaci*) fisheries within in the U.S. Southeastern Region (South Atlantic and Gulf of America) is made complicated by concerns about information quality, which has previously led the Florida Fish and Wildlife Conservation Commission (FWC) to suspend a stock assessment (SEDAR 48, 2017). Accordingly, the FWC is engaged in an on-going process of exploring alternative decision-making approaches. Here, a process was initiated to progress the design and evaluation of multi-indicator management procedures (MPs) to confront issues of data quality and availability and robustness to ecological uncertainties. Performance of MPs was quantified using management strategy evaluation (MSE) that relied upon a simulated fishery system with characteristics representative of the black grouper fishery. Notably, MP performance was measured against a target reference point of 30% SPR, which serves as an MSY proxy for black grouper, as prescribed by the Gulf Council and South Atlantic Council. At least three multi-indicator MPs were demonstrated to have the potential to produce stability in stock biomass while subtly guiding longer-term outcomes towards the Council prescribed management target. It should be noted that this analysis represents a desktop MSE designed to evaluate the robustness of candidate management procedures under a range of operating model scenarios, rather than to provide immediate stock status determinations or prescriptive management advice. However, with continued work, an MP approach could lead to an interim policy for decision-making during which improved data collection is considered and conventional stock assessment becomes viable.

Table of contents

Background	3
Methods	5
Overview	5
Management procedures	6
Operating models	17
Reference OMs	18
Robustness OM	27
Performance metrics & evaluation.....	28
Results	30
Reference OMs	30
Robustness OM	38
Discussion	40
Synthesis	41
How to refine the operating models?.....	43
How to refine the management procedures?.....	44
Conclusions.....	46
References	47
Tables	51
Figures	66
Appendix: All MP figures.....	90

Background

Management of black grouper (*Mycteroperca bonaci*) fisheries within in the U.S. Southeastern Region (South Atlantic and Gulf of America) is made complicated by concerns about information quality. Information quality has been called into question, especially as it relates to the accuracy and precision of landings, which has hindered stock status determination (SEDAR 48, 2017). For example, in 2017, the Florida Fish and Wildlife Conservation Commission (FWC) initiated an age- and length-structured stock assessment using Stock Synthesis (Methot and Wetzel 2013), which was subsequently suspended due to high uncertainty in landings history. Despite concerns about landings history, the availability of various data streams is an invitation to consider alternative pathways for provision of management advice. Accordingly, the FWC is engaged in an on-going process of exploring alternative decision-making approaches.

One such alternative approach is through the design and evaluation of management procedures (MPs). An MP consists of three components: data collection, data analysis, and a pre-agreed decision rule or harvest control rule (HCR). These components are intended to work cohesively as a decision-making framework (Bentley and Stokes 2009). As an initial step in considering alternative MPs for black grouper fisheries management, the FWC completed an information synthesis that explored the viability and limitations of the MP approach (Henriquez et al. 2024). This information synthesis for black grouper fisheries management resulted in (i) delineation of key uncertainties and ranking of apparent quality of available data streams, (ii) identification of types and styles of MPs that aligned with higher-quality data streams, and (iii) demonstration of MPs that have been applied to similar fisheries (Henriquez et al. 2024). In the case of black grouper fisheries, multiple data streams are available yet cannot be statistically integrated within

a conventional stock assessment, necessitating exploration of alternative decision-making frameworks (Harford et al. 2021; Henriquez et al. 2024).

In a new phase of work that is reported herein, FWC has requested a desktop-based design and exploration of multi-indicator MPs that could be applied to a fishery system sharing similar characteristics to the Florida black grouper and its fisheries (Harford et al. 2021; Walter III et al. 2023). An examination of MPs was carried out using Management Strategy Evaluation (MSE) as progress towards a strategic understanding of how MPs for black grouper fisheries could cope with existing data limitations. *Strategic* MSE can provide coarse policy guidance, whereas *tactical* MSE typically requires more in-depth investigation that involves engagement between stakeholders, decision-makers, and analysts (Walter III et al. 2023). An MSE was conducted by completing the following tasks. First, a structured approach was taken in designing a suite of multi-indicator MPs that relied upon the previously completed information synthesis. Second, a numerical simulation of a fishery system was created, known as an operating model (OM), which was representative of black grouper population dynamics and its fisheries, and processes for generating imperfect observation from available data streams. Finally, numerical simulations were used to evaluate the anticipated performance and trade-offs among the suite of MPs using quantifiable performance metrics reflecting fishery management objectives relevant to the State of Florida.

Methods

Overview

At time of report preparation, the authors developed OM s with reliance principally on SEDAR reports and related working papers, with limited access to raw datasets (e.g., SEDAR 19, 2010; SEDAR 48, 2017). The OM s were specified to include two fisheries that were representative of commercial and recreational sectors, and with a focus on scenario-driven construction of plausible recent-past (2000 to 2025) trends in fishing effort and population size. In specifying OM s, data limitations typically necessitate scenario-driven development, whereas in data-rich cases, where a stock assessment is available, OM s are based more closely on the information produced by the stock assessment. OM s were specified to explicitly reflect a focal set of uncertainties, against which the performance and robustness of MP s was evaluated. OM s were classified into reference OM s and a robustness OM (Rademeyer et al. 2007). The reference OM s represented the most plausible scenarios, reflecting recent past (2001 to 2025) trends in fishing mortality and population size (see section: *Reference OM s*). A robustness OM included a scenario that was less likely but still plausible (see section: *Robustness OM*) (Punt et al. 2016). Given reliance on contemporary information, MP design emphasized reliance on near-past observations rather than those farther afield, figuratively speaking, because more recent data collection programs generally appear to be more reliable (Henriquez et al. 2024). Forward-projection of MP s was simulated based on MP implementation in 2026 for a duration of 50 years.

Management procedures

Design of MPs, with emphasis on multi-indicator MPs, was carried out according to the organizational approach outlined in Harford et al. (2021), and consisted of the following four steps: (i) identify indicators, (ii) select reference points and operational control points (OCPs), (iii) understand key uncertainties, and (iv) assemble a multi-indicator framework. Indicators were derived from observations of the fishery system as proxies for more direct variables of interest. In practice, this meant that degree of information limitations was used to differentiate the structures of proposed MPs. For example, more recent data collection generally appears to be more reliable, and thus, MP design emphasized reliance on more contemporary observations.

Step one: identify indicators

Getting organized around MP design involves screening the strengths and shortcomings of indicators and resolving the ways information can be assimilated (not unlike conventional stock assessment). Henriquez et al. (2024) identified data streams relevant to black grouper fishery management and provided a qualitative assessment of data reliability (Table 1). Here, a more constrained suite of data streams was used in MP design (Table 2). Life history parameters such as growth, length-weight, and maturity were deemed as acceptable for use in designing MPs. Two indices of abundance were used in developing indicators: Florida State Reef Fish Survey (SRFS), a fishery-dependent survey of recreational angling; and, the Reef Fish Visual Census (RVC), a fishery-independent survey occurring in approximate spatial alignment with the center of distribution of black grouper. In Henriquez et al. (2024), MRIP was identified and potentially reliable; however, following an update of the OM with more recent SRFS data, it appears that the SRFS abundance index has relatively low precision and likely requires some additional scrutiny

regarding its inclusion in delineating useful indicators. However, for completeness it is included in this analysis as a replacement for MRIP. Three sources of length composition data were used in developing indicators: SRFS, RVC, and those from commercial longline catches.

Using these available data streams, indicators were specified by considering (i) linkages between indicator attributes and condition of the fishery system, and (ii) reliability and expected information content obtained from monitoring programs (e.g., Magnusson and Hilborn 2007; Harford and Babcock 2016), principally based on estimates of data precision from previous SEDAR reports (SEDAR 19, 2010; SEDAR 48, 2017) (Table 2; see details therein). For example, the RVC has the potential to inform relative cohort strength entering, or soon to enter, the fisheries. Indicators derived from fishery-dependent length compositions can be used as proxies for changes in demography relative to useful targets (ICES 2015; Medeiros-Leal et al. 2023). For the given set of chosen available data streams, a variety of indicators were specified with the dual aims of maintaining stability in black grouper spawning biomass as well as long-term achievement of a proxy for maximum sustainable yield (MSY). Indicators that were derived from indices of abundance, used a rolling-window approach where present conditions are contrasted against near-past conditions with the aim of maintaining biomass stability through TAC adjustments. An equivalent approach was implemented for length composition data, using a rolling-window of annual mean length estimates (Table 2). Additionally, a length-based indicator was derived by combining fishery-dependent length compositions from recreational fishery and commercial fishery to produce mean length in the combined catch. This mean length indicator was contrasted against a length-based target that served as a proxy for MSY to guide biomass levels towards achievement of MSY (Jardim et al. 2015).

Step two: selecting reference points and operational control points

Within the context of U.S. fisheries management, National Standard 1 – Optimum Yield (NS1), reference points collectively refer to: Status Determination Criteria (SDC), maximum sustainable yield (MSY), OY, ABC and ACL (National Standard 1 - Optimum Yield 2022). SDC means the measurable and objective factors including maximum fishing mortality threshold (MFMT), OFL, and minimum stock size threshold (MSST), and their proxies. Many of the above stated SDC require estimation within conventional stock assessment or other demographics models, both of which usually rely on reliable catch histories (MacCall 2009; Free et al. 2023). However, NS1 offers alternative pathways to ABC calculation without the need to calculate conventional SDC (National Standard 1 - Optimum Yield 2022). As an example, the South Atlantic Fisheries Management Council specifies ABC control rules based on severity of scientific uncertainty, with some tiers used to calculate ABC directly without OFL calculation (Free et al. 2023, their supplementary information). Likewise, the New England Fishery Management Council manages fisheries for Thorny Skate (*Amblyraja radiata*) without MSY, OFL and OY calculation, instead using an index of relative abundance with a decision-rule that establishes ABC (NEFMC 2022).

Design of the suite of MPs emphasized the use of model-free or empirical indicators (i.e., derived more-or-less directly from raw data) to inform output-control HCRs, which in turn are used to adjust annual Total Allowable Catch (TAC). While the authors do not explicitly examine ABC control rules, this desktop MSE is intended as a step towards the possibility of ABC calculation without explicit calculation of most SDC, as per NS1. Thus, indicators and their use within HCRs are presented to adjust catch limits (in the generic sense of the term), with

suitability for application within State and Federal fisheries management policy requiring further discussion and clarification.

While reference points are commonly used to judge the status of the exploited population relative to management objectives, these differ from operational control points (OCPs; Cox et al. 2013). OCPs are quantities that are used in decision-making to invoke adjustments to TACs. For example, an indicator may be regarded as an imprecise approximation for a biological reference point, which is linked to an OCP that produces a high probability achieving that reference point despite errors in interpretation of an indicator (Cox et al. 2013). Selection of OCPs should proceed in a way that leads to the fishery toward achieving management objectives (PFMC 2020), which is evaluated in this report using MSE. Below, calculation of indicators is explained, including necessary OCPs and related parameters. All indicators follow a similar rationale for informing annual adjustments to a TAC, in that, when an indicator quantity does not exceed a defined threshold, it does not invoke TAC changes, promoting catch stability, but when an indicator quantity surpasses the threshold, it causes a proportionate response relative to distance from the threshold (Fig. 1).

Stability indicators using indices of abundance - Two approaches were applied to calculating indicators based on indices of abundance, each aimed at detecting changes in relative abundance and adjusting the TAC to maintain biomass stability and avoid further stock decline. Both approaches rely on annual index quantities within a ‘moving window’ over the most recent 5 years of data. Five years was selected because the SRFS was modified in 2020 to include state-wide implementation (previously excluding the Florida Keys), thus, it was assumed that

contemporary observations from 2021 to 2025 were the most reliable information on which to base MPs with implementation beginning in 2026. For simplicity, the most recent five years were used for all abundance indices.

The *slope-based* approach applied a linear regression to an index of abundance over the five-year period (Geromont and Butterworth 2014). A slope estimate produced from the *slope-based* approach was standardized (*std. slope*) by dividing it by the mean index value over the same period, making it scale-independent, where a value of zero indicates no change in the index of abundance. The TAC adjustment resulting from the *slope-based* calculation is:

$$Slope_Adj[I] = \begin{cases} 1, & \text{abs}(std.slope[I]) < \theta_{slope} \\ 1 + \gamma_{slope}std.slope[I], & \text{otherwise} \end{cases} \quad (1)$$

where I is the indicator data source, θ_{slope} is a responsiveness threshold, set to 0.05 to avoid frequent TAC adjustments, and where γ_{slope} controls the magnitude of TAC adjustment and was set to 1.0 (Fig. 1A). The *sequential ratio* approach calculates year-to-year changes in an index of abundance over the most recent 5 years of data (Apostolaki and Hillary 2009). The sequential ratio was calculated:

$$SeqRatio[I] = \sum_{i=1}^4 \omega_i \log\left(\frac{I_{i+1}}{I_i}\right), \quad (2)$$

where ω_i are weights that can be informed by index precision or recency of observations, and for which equal weights (1/4 each) were applied here. The TAC adjustment for the *sequential ratio* is then calculated:

$$SeqRatio_Adj[I] = \begin{cases} 1, & \text{abs}(SeqRatio[I]) < \theta_{SeqRatio} \\ 1 + \gamma_{SeqRatio}SeqRatio[I], & \text{otherwise} \end{cases} \quad (3)$$

where $\theta_{SeqRatio}$ is a responsiveness threshold, set to 0.05 to avoid frequent TAC adjustments, and where γ controls the magnitude of TAC adjustment and was set to 1.0 (Fig. 1A).

Stability indicators using mean length – Like the indicators based on relative abundance, stability seeking mean length indicators were aimed at maintaining biomass stability and avoiding further stock decline. Stability indicators using mean length rely on a ‘moving window’ over the most recent five years of data, where five recent years were selected for the same reasons outlined for the stability indicators using indices of abundance. The *average length ratio* indicator is calculated as the current year's mean length relative to the average mean length over the previous five years. The TAC adjustment from the *average length ratio* was calculated as:

$$LenRatio_Adj[I] = \begin{cases} 1 - PenMult[I] \times \left(1 - \frac{LenRatio[I]}{(1 - \theta_{LenRatio})}\right), & LenRatio[I] < (1 - \theta_{LenRatio}) \\ 1 + RewardMult[I] \times \left(1 + \frac{LenRatio[I]}{1 + \theta_{LenRatio}}\right), & LenRatio[I] > (1 + \theta_{LenRatio}) \\ 1, & otherwise \end{cases} \quad (4)$$

Penalty and reward multipliers, $PenMult[I]$ and $RewardMult[I]$, respectively, enable disproportionate adjustments in TAC, given that changes in mean length may not be proportional to changes in stock biomass. For all length-based stability indicators these parameters are set at 1.0. These multipliers are also known as ‘gain’ parameters and enable TAC adjustments to be more aggressively applied when the indicator is below the target, for example (Fig. 1B; Geromont and Butterworth 2015; Anderson 2021). The responsiveness threshold, $\theta_{LenRatio}$, was set to 0.05 to avoid frequent TAC adjustments. Specification of $\theta_{LenRatio}$ was determined through simulating the magnitude of observation error that would be expected for a fishery system in equilibrium. That is, the OM was used to simulate the observation process for mean length where the only source of variation arose from observation error (not changes in stock size), which in this case produced a coefficient of variation of 0.05.

Target seeking indicator using mean length – Using mean length in the catch, an indicator was specified to maintain size composition in proximity to MSY by responding to changes in size structure dynamics, fishing mortality and recruitment patterns. A length-based target, as a proxy for achievement of MSY, was calculated following the methodology outlined in Jardim et al. (2015). In their Appendix A, Jardim et al. (2015) describe a generalized version of the expected mean length in the catch, as originally derived by Beverton and Holt (1957):

$$Target[I] = \frac{K/M + L_c(\gamma_{Ltarget} + 1)}{K/M + \gamma_{Ltarget} + 1}, \quad (5)$$

where K is von Bertalanffy growth coefficient, M is natural mortality, L_c is length at full vulnerability to fishing, and $\gamma_{Ltarget} = F/M$ is the ratio of fishing mortality to natural mortality that produces the *Target* length. In a multi-fleet fishery, the interpretation of $\gamma_{Ltarget}$ is more complex. The value of *Target* to be applied against mean length from the combined recreational and commercial fisheries (see Table 2) was calculated using information obtained through simulation, using a modified version of the OM. In this case, the OM was used to solve for equilibrium level of spawning potential ratio (SPR) of 30% under constant life history ($M = 0.195 \text{ yr}^{-1}$, $K = 0.134 \text{ yr}^{-1}$, $L_{inf} = 136.7 \text{ cm}$) and constant yield allocation to each fleet (recreational = 63%, commercial = 37%, $L_c = 60.96 \text{ cm}$ or 24 inches) but where steepness of the stock-recruitment relationship was drawn from a Beta distribution (see section: **Operating models**). The SPR target of 30% was selected because both the Gulf Council and South Atlantic Council use 30% SPR as an MSY proxy for black grouper. Under these OM conditions, the median was *Target* = 77.260 cm, with a CV of 0.004. Solving the above equation from Jardim et al. (2015) using $M = 0.195 \text{ yr}^{-1}$, $K = 0.134 \text{ yr}^{-1}$, produces $\gamma_{Ltarget} = 1.501$. The same process was applied using 50% SPR, as an alternative MSY proxy (see Harford et al. 2019) and as a more precautionary OCP to aid in avoiding depletion below the Council's 30% SPR target,

produced median of $Target = 78.675$ cm and $\gamma_{Ltarget} = 1.246$. Having established $Target$, the $Ltarget$ indicator is:

$$Ltarget30[I] = \frac{\text{Current year's mean length [I]}}{Target\ 30\%[I]}, \quad (6)$$

and,

$$Ltarget50[I] = \frac{\text{Current year's mean length [I]}}{Target\ 50\%[I]}. \quad (7)$$

The TAC adjustment from the $Ltarget$ was calculated as:

$$LTarget30_Adj[I] = \begin{cases} 1 - PenMult[I] \times \left(1 - \frac{Ltarget30[I]}{(1-\theta_{Ltarget})}\right), & Ltarget30[I] < (1 - \theta_{Ltarget}) \\ 1 + RewardMult[I] \times \left(1 + \frac{Ltarget30[I]}{1+\theta_{Ltarget}}\right), & Ltarget30[I] > (1 + \theta_{Ltarget}) \\ 1, & otherwise \end{cases} \quad (8)$$

and,

$$LTarget50_Adj[I] = \begin{cases} 1 - PenMult[I] \times \left(1 - \frac{Ltarget50[I]}{(1-\theta_{Ltarget})}\right), & Ltarget50[I] < (1 - \theta_{Ltarget}) \\ 1 + RewardMult[I] \times \left(1 + \frac{Ltarget50[I]}{1+\theta_{Ltarget}}\right), & Ltarget50[I] > (1 + \theta_{Ltarget}) \\ 1, & otherwise \end{cases} \quad (9)$$

As in ***Stability indicators using mean length***, the penalty and reward multipliers are used; however, a brief tuning exercise was applied that resulted in penalty and reward multipliers of 2.0 and 0.5, respectively, to aid in avoiding cyclical and ratcheting effects sometimes observed with length-based indicators (Klaer et al. 2012; Hordyk et al. 2015; Bellquist et al. 2022). The responsiveness threshold, $\theta_{Ltarget}$, was set to 0.05 to avoid frequent TAC adjustments and follows the same rationale as described for the ***Stability indicators using mean length***.

Step three: understanding key uncertainties

Priority uncertainties associated with MPs were the apparent precision and bias of data streams from which indicators were derived. Data precision was addressed in two ways. First, best-available levels of data precision were obtained and used to simulate imperfect information from observation sub-model routines (see section: *Reference OMs*). Second, the robustness OM consisted of an alternative used to examine the effects of data reliability on MP performance (see section: *Robustness OM*). While this desktop MSE relied solely on statistical precision as a measure of indicator reliability, it is worth noting that Rice and Rochet (2005) provide a more complete set of nine screening criteria for data reliability, including ease of measurement and interpretation, and emphasize broader inclusivity of stakeholder and technical viewpoints within the indicators screening process.

Step four: assembling a multi-indicator framework

Using the set of indicators described in *Step two*, suites of single-indicator MPs and multi-indicator MPs were specified (Table 3). Single-indicator MPs were examined to better understand expected performance outcomes of each indicator, prior to construction of more complex multi-indicator MPs. For clarity, all MPs followed the same algorithm for producing TAC adjustments, described in three steps below.

First, the previous TAC was obtained. Calculation of the previous TAC was $\sum_{i=y-1}^{y-x} TAC_i/x$, which avoids abrupt transitions from one year to the next. We used the formula $x =$

$\left\lceil -\frac{\log(0.05)}{\bar{M}} - rec\ age \right\rceil$, where *rec age* (set to 4.0 years) indicates the approximate age of recruitment to the fishery at the minimum size limit of 24 inches, and x approximates cohort exposure time to recent-past catch rates (Apostolaki and Hillary 2009; Klaer et al. 2012). This

gives $\underline{x} = 11$ years of cohort exposure to the fisheries. In cases where the previous $\underline{x} = 11$ years included years prior to 2026, the imperfectly observed catches (in biomass via an observation model sub-routine) were used in replacement of previous TACs.

Second, an HCR was used to calculate a preliminary TAC:

$$TAC_{Preliminary} = TAC_{Previous} \times X_{Adj}[I], \quad (10)$$

where $X_{Adj}[I]$ is a placeholder for any of the indicators used in single-indicator MPs. Multi-indicator MPs follow a multiplicative approach, for example:

$$TAC_{Preliminary} = TAC_{Previous} \times X1_{Adj}[I] \times X2_{Adj}[I]. \quad (11)$$

Third, a TAC stability clause was imposed. All MPs tested included the following stability constraints, aimed at avoiding unusually large TAC adjustments. This stability constraint was implemented such that annual TAC change could not exceed +/- 10% relative to the previous TAC. Finally, allocation to each fleet was assumed to remain unchanged during MP projections. The recent actual allocation from 2024 for the South Atlantic region, where the recreational fleet is allocated 63.12% and the commercial fleet is allocated 36.88% of the total TAC was used for all MPs.

Single-indicator MPs – These MPs were intended to be informative about expected performance outcomes of each indicator, prior to developing multi-indicator MPs (Table 3). In total, nine single-indicator MPs were evaluated, utilizing stability indicators using indices of abundance, stability indicators using mean length, and a target-seeking indicator using mean length.

Multi-indicator MPs – Following an evaluation of single-indicator MPs, an additional set of MPs was specified to examine whether combining indicators could address limitations or improve MP performance, or both (Table 4). Multi-indicator MPs measured different aspects of stock condition and were utilized strategically to produce a composite profile the state of resource, leading to TAC adjustment. For example, abundance indices track biomass trends and promote stability, while size indicators detect changes in fishing mortality, recruitment, and population structure. As a second example, in several of the multi-indicator MPs, fishery-independent survey representative of the RVC (see Table 2) was used. In these cases, the RVC appears to provide the best-available indication of cohort strength entering or soon to enter the fishery. When using the RVC, both the index of abundance and mean length were used as indicators (Table 4). Thus, when an increase in relative abundance is detected, along with a decrease in mean length, a larger cohort may be entering the fishery, but in this case counter directional adjustments to the TAC are imposed. This results in a ‘wait and see’ approach that delays increasing harvest of a potentially large cohort, allowing the cohort to grow. This strategy attempts to maximize the potential for the cohort to contribute to the fishery, as described by Walters and Martell (2004; their chapter 3). As a third example, the mean length within the commercial fishery (longline fishery; see Table 2) is included in several multi-indicator MPs because this fishery targets mature adults, and thus, it is worthwhile to monitor changes in this component of the stock. As a technical note, index-based indicators relied on the *Sequential ratio* methodology because the availability of annual weighting parameters offers the potential for further refinements, perhaps related to index precision or recency of observation, that is not afforded by the *Slope* methodology. In addition, in multi-indicator MPs labelled 13, 14, 15, the SRFS abundance index was excluded given its low precision, enabling focus on a subset of

higher precision data sources (Table 4). In all, six multi-indicator MPs were evaluated, with the intention of providing a more robust stock condition assessment than any single indicator alone (Table 4).

Operating models

The fishery system, or Operating Model (OM), is a numerical model of the population dynamics of the fish stock, its fisheries, and monitoring programs for fishery-dependent or fishery-independent data, or both. In specifying an OM, historical trends in fish stock dynamics and fishing mortality or effort, catch, and other regulations serve as the precursor conditions from which forward-projections of an MP are made. For black grouper, the severity of data limitations required a data-limited approach for OM specification. This approach emphasized ‘tuning’ alternative scenarios about recent past trends, rather than OM conditioning based on a wealth of information that might otherwise be available from conventional stock assessment.

The OM(s) along with other components that comprise MSE were implemented within the R package FishSimGTG (Harford 2024). FishSimGTG was used to conduct numerical modeling of fish populations, including management strategy evaluation (MSE), and is an age-structured model that represents multiple concurrent cohorts using the growth-type-group methodology (Walters and Martell 2004). A detailed description of FishSimGTG can be found at <https://natureanalytics-ca.github.io/fishSimDoc/> and source code is publicly available on GitHub at <https://github.com/natureanalytics-ca/fishSimGTG>.

Reference OMs

Population dynamics

Population dynamics of black grouper were modeled as a single stock occurring within a single area corresponding to the U.S. Southeastern Region (South Atlantic and Gulf of America). Temporal dynamics occurred according to an annual time step, where growth occurred first, followed by reproduction, and then by total mortality (i.e., natural mortality plus fishing mortality). Somatic growth followed a von Bertalanffy function, maturity was an asymptotic function of length, and proportional transition from female to male was specified as a function of length, as black grouper is a protogynous hermaphrodite. Numbers-at-age were modeled as a single sex, and hermaphroditism was addressed in the calculation of reproductive output-at-age. Recruitment of age-0 fish was calculated according to the Beverton-Holt stock-recruitment function, parameterized using steepness, and where reproductive output (weight-per-female) was calculated as a power function of length (Table 5).

Most life history parameters were specified as point estimates (Table 5). However, uncertainties in productivity parameters (natural mortality, steepness, inter-annual recruitment variation) were specified using statistical sampling distributions. Natural mortality and inter-annual recruitment variation were specified as ranges, with minimums and maximums corresponding to the bounds of corresponding uniform distributions. During each iteration (simulation), draws from the corresponding uniform distributions were made (Table 5). The Hamel and Cope (2022) longevity-based estimator produced a point-estimate of natural mortality of 0.16 year^{-1} , based on the maximum reported age of 33 years (SEDAR 19, 2010). A range for natural mortality was obtained using the 50% confidence interval from Hamel and Cope (2022), as the 95% confidence

interval contained unreasonably high values. This specified range for natural mortality encompassed other previously considered point estimates (O’Hop and Beaver 2009; SEDAR 19, 2010). Inter-annual recruitment deviations are calculated on a log-scale with standard deviation specified as a uniform distribution between 0.2 and 0.6. A beta distribution for steepness was informed by the meta-analysis of Shertzer and Conn (2012), who describe steepness using a beta probability distribution with mode of 0.84 (approximate 95% centered distribution: 0.31 to 0.97; approximate interquartile range: 0.74 to 0.86), and by Thorson (2020) with a mean of 0.68 (95% centered distribution: 0.38 to 0.97; interquartile range: 0.58 to 0.78). The beta distribution for steepness most closely mimicked that of Shertzer and Conn (2012) but was bound between > 0.2 and 1.0. This approach produced a mode of 0.83 and centered 50% of the distribution between 0.74 and 0.86.

Fishery characteristics

Two fisheries were specified, which for simplicity are termed ‘recreational fishery’ (fleet 1) and ‘commercial fishery’ (fleet 2) and correspond approximately to fisheries targeting sub-adult/early adult and adult components of the stock, respectively (Table 6). Landings from recreational and commercial sectors tend to occur at different depths, with recreational fisheries occurring in shallower waters < 60 m and commercial fisheries occurring in deeper waters typically beyond 40 m (O’Hop and Beaver 2009). Length compositions from recreational landings and shallow-water fishery-independent collections tend to contain mainly sub-adults and fewer individuals above the length at 50% maturity (O’Hop and Beaver 2009). Conversely, length compositions from commercial landings contain mainly adults and fewer individuals below the length at 50% maturity (O’Hop and Beaver 2009). The predominance of subadults in shallower depths is also

apparent from a fishery-independent Reef Visual Census (RVC) conducted in depths of up to 33 m, and which observes few individuals above the length at maturity (Ault and Smith 2009). There is, however, a paucity of information about coarse-scale movement and corresponding movement rates at-length or at-age that would be necessary to explicitly represent migration in simulation of stock distribution. Therefore, to remedy this issue, selectivity of each fishery implicitly reflects the combined availability and vulnerability to the gear.

Representation of the recreational fishery was principally based on SRFS, with related information about selectivity of the recreational fleet obtained from the NOAA Marine Recreational Information Program (MRIP) (Table 6; Fig. 2). Recreational selectivity was parameterized as dome-shaped, following the SEDAR 19 (2010) approach for recreational fleets that operate in shallow waters and infrequently encounter older fish. Length-based parameters were derived by converting the SEDAR 19 (2010) age-based selectivity (from 1999–2008) to length-based selectivity. Selectivity peaked at age 2, which yielded a peak selectivity length of 38.3 cm. The dome-shaped pattern reflects that recreational fleets typically fish in shallower waters than commercial vessels. Consequently, selectivity declines for older ages, as black grouper older than 4 years (>67 cm) increasingly migrate to depths greater than 40 m—beyond the typical operating range of most recreational vessels. In the OM, recreational selectivity was tuned using an exponential logistic selectivity function (dome-shaped) to mimic the MRIP selectivity curves estimated in SEDAR 19 (2010). Like commercial fishing, retention in the recreational fishery was specified to be knife-edge at the minimum legal-size limit of 24 inches (60.96 cm). Discard mortality for the recreational fishery, which operates at shallower depths than the commercial fishery (range 1–50 m) had recommended mortality range of 9–26% from

SEDAR 48 (2017) based on discards from Florida for-hire fishery (2005-2013) using a tag-recapture model developed for gag grouper, and set the mid-point of 17.5% in the OM.

Representation of the commercial fishery was based on longline selectivity using a logistic function, obtained from SEDAR 19 (2010), which applied the age-structured model ASAP2 (Fig. 3). A logistic function is appropriate for the longline fishery because this fleet primarily encounters older black grouper, unlike shallower-operating fleets that rarely encounter older individuals (Table 6). Age-based selectivity from SEDAR 19 (2010) was converted to length-based using von Bertalanffy growth parameters. Based on this conversion, the age at 50% selectivity (i.e., $A_{50} = 7$ years) corresponded to an L_{50} of 86.3 cm, and the age at 95% selectivity (i.e., $A_{95} = 12$ years) corresponded to an L_{95} of 110.8 cm. These selectivity parameters are supported by SEDAR 48 (2017; their Fig. 2.8), which shows an increase in the occurrence of black grouper between 70 and 80 cm at longline operating depths (36–85 m, mean 57 m) (Table 6). Retention in the commercial fishery (i.e., the probability that a fish is selected by the fishing gear and is kept rather than discarded) was specified to be knife-edge at the minimum legal-size limit of 24 inches (60.96 cm). As the OMs were initiated in the year 2000 (see: *Initial conditions and historical dynamics*), historical regulatory changes in size limits were not included in the characterization of the commercial fishery. Discard mortality rates were assumed to be depth-dependent following SEDAR 48 (2017; their table 2.5) and based on gag grouper (*Mycteroperca microlepis*) as a proxy species. The commercial longline fishery operates within a typical depth range of 36 to 85 m, with recommended discard mortality of 25-50%. The midpoint of 37.5% discard mortality assumed in the OM.

Fishery-dependent observation models

An important subcomponent of an OM is the observation model(s). The observation model(s) simulate imperfect resource monitoring data with similar error structure to what would be available in practice. The observation model generates monitoring datasets, which are then fed into MPs.

Length composition, but not an index of biomass, was simulated for the commercial fishery (Table 7). At the time of report preparation, CPUE was unavailable for the commercial fishery, and thus was not included in accordance with the project objective of coping with existing data limitations. Simulated sampling of length composition, including precision of imperfect information was based approximately on sampling characteristics of the NOAA Trip Interview Program (TIP). Biological samples, including lengths, are collected at docks by port agents along the southeastern US coast, where commercial catches are landed. While number of length samples varies by year and gear type, precision of the length composition was based on sample sizes from the longline fishery, where the average number of black grouper measured per year was 176 between 1986 and 2015 (SEDAR 48, 2017). Simulated observation of length composition from the commercial longline fishery was generated using a multinomial sampling distribution with sample size value of 100 for each historical and forward-projection year, with bin width 1 cm. Imperfect observation of catch in number and biomass was based on an assumed CV of 0.1 (Table 7). While catch history is not utilized by any MP, catch estimates are required for estimation of weighted-average mean length across fleets and for previous TAC calculations.

An index of abundance and observation of length composition were both simulated for the recreational fishery. Samples from the recreational fishery were generated based on characteristics of SRFS; a program which collects fishery-dependent data through intercepts with recreational anglers. Prior to the SRFS, a related survey known as MRIP modified its methodology in 1991 and enhanced training for field samplers to help avoid confusion between species identification of black grouper and gag grouper. In 2020, SRFS received funding for state-wide implementation. An annual index of abundance was simulated by applying observation error to the ‘true simulated’ vulnerable abundance for the recreational fishery using lognormal error structure with bias correction, with specified range for Coefficient of Variation (CV) sampled for each iteration (Table 7). The CV range was approximated based on analysis of catch estimates from SRFS that were provided by FWRI. The survey was assigned a catchability coefficient (q) based on SEDAR 19 estimates (their Table 3.3.4.7). The catchability was sampled from a uniform distribution bounded between $2.62\text{e-}06$ and $4.07\text{e-}06$, representing the 95% confidence interval around the point estimate of $3.27\text{e-}06$. The hyperstability/hyperdepletion parameter, b , was fixed at 1.0. Simulated sampling of recreational length composition was generated using a multinomial sampling distribution with sample size value of 100 for each historical and forward-projection year, with bin width 1 cm. Imperfect observation of catch in number and biomass was based on a specified CV (Table 7). While catch history is not utilized by any MP, catch estimates are required for estimation of weighted-average mean length across fleets and for previous TAC calculations.

Fishery-independent observation models

An index of abundance as well as length composition were simulated for a shallow water survey with similar characteristics to the NMFS – University of Miami Reef Visual Census (RVC) survey monitoring program (Brandt et al. 2009; Smith et al. 2011; 2017). Selectivity of the simulated reef survey selectivity was parameterized as dome-shaped because the survey operates at shallow reef depths (1–33 m maximum) and likely does not sample the full availability of black grouper lengths due to ontogenetic depth migration and operational constraints of shallow reef sampling (Table 6; Fig. 4). This results in a descending limb of the selectivity curve as larger fish move beyond the survey's depth limit. The peak selectivity at ~ 24.2 cm was derived directly from the estimated selectivity at age reported in SEDAR 19 (2010; their Fig. 3.3.5.5). The shallow water survey was tuned to the RVC survey selectivity parameters to generate a descending limb of the selectivity function structure similar to that reported in SEDAR 19 (2010).

An annual index of abundance for the shallow water survey was simulated by applying observation error to the ‘true simulated’ abundance selected by this survey using lognormal error structure with bias correction, CV uniformly sampled from 0.086 to 0.208 for each iteration (Table 7). The CV range was derived directly from SEDAR 48 (2017) estimates of RVC survey precision, reflecting uncertainty in diver-based visual census methods where detection probability varies with environmental conditions, fish behavior, and observer experience. The survey was assigned a catchability coefficient (q) based on SEDAR 19 estimates (Table 3.3.4.7). The catchability was sampled from a uniform distribution bounded between 1.55e-06 and 2.32e-06, representing the 95% confidence interval around the point estimate of 1.89e-06. The

hyperstability/hyperdepletion parameter, b , was bounded between 0.8 and 1.2 (uniform distribution). Simulated sampling of reef survey length composition was generated using a multinomial sampling distribution with sample size value of 150 for each historical and forward-projection year, with bin width 1 cm.

Initial conditions and historical dynamics

In fishSimGTG, simulations begin with an initial equilibrium age-structure based on a specified level of stock depletion (i.e., spawning stock biomass relative to unfished spawning stock biomass). The year 2000 was specified as the initial equilibrium year, as cohort size through the 2000's were relatively stable, as was the relative fishing mortality of the recreational fishery; the larger of the two fisheries (SEDAR 19, 2010). Likewise, regulatory changes around the same time resulted in consistency across U.S. jurisdictions for the minimum size limit for black grouper (with the exception of federal waters of the Gulf of America, 24 inch size limit introduced in 2016) (SEDAR 19, 2010; SEDAR 48, 2017). Uncertainty in the initial state of the resource was specified as a uniform distribution on SPR between 0.4 and 0.6 to reflect a fully exploited black grouper stock (Huntsman et al. 1994; SEDAR 19, 2010). Like the productivity parameters, each iteration (simulation run) was initialized with a draw from the corresponding uniform distribution of initial SPR. When multiple fleets are included in an OM, information is required about how to distribute total fishing mortality among different fleets. To do this, target catch proportions were specified, which represent the percentage of the total catch allocated to each fleet. To initiate the model, it was assumed that allocation for the initial year was consistent with more recent allocation from 2024 for the South Atlantic region, where the recreational fleet

is allocated 63.12% and the commercial fleet is allocated 36.88% of the total TAC (by comparison, Gulf of America: 76% commercial, 24% recreational).

To create a reference set of OMs, two scenarios about historical trends in fishing mortality were developed, based on relative trends produced in SEDAR 19 (2010; their Table 3.3.4.4):

OM1 - Stable fishing mortality: The commercial and recreational fisheries were specified to have reasonably stable fishing mortality rates between 2000 and 2025. Inter-annual variation in fishing mortality was specified (based on information from MRIP fishing mortality estimates) as a log-scale standard deviation of 0.25, while a similar estimate of 0.13 was obtained from commercial longline fishing mortality estimates. In both cases, inter-annual variability was estimated from fishing mortality estimates from 2000 to 2008 obtained from SEDAR 19 (2010). A uniform distribution for inter-annual variation in fishing mortality was specified between 0.13 and 0.25 for both fisheries.

OM2 – Increasing fishing mortality: In the second reference OM, increasing fishing mortality in both fisheries was assumed, not necessarily intended to reflect actual fishing patterns, but instead an important strategic evaluation of the response of MPs to an upward trend in fishing mortality. In this OM, both fisheries were assumed to have increasing fishing mortality at a rate of 0.5% per year, producing approximately a 12.5% increase in total fishing effort over 50 year MP projections. A uniform distribution for inter-annual variation in fishing mortality was specified between 0.13 and 0.25 for both fisheries.

Robustness OM

A robustness OM was specified to evaluate an alternative plausible scenario that deviated from the reference OM's assumptions. The robustness trials are used to determine whether the management procedure will behave as intended in less likely scenarios, though still plausible (Punt et al. 2016). In general, it is common to evaluate different scenarios about data availability, changes in productivity or carrying capacity, or changes in the fishery characteristics, such as selectivity, catchability, etc. (Geromont and Butterworth 1998; Rademeyer et al. 2007)

OM2 - R1- Index (recreational & survey) hyperstability: The recreational fishery and the RVC survey appear to occur near the center of the distribution of black grouper, and therefore could be subject to hyperstability. The SEDAR 48 report (SEDAR 2017) notes that the RVC index is "primarily a sub-adult index" sampling shallow waters (1-33m), with only ~3% of observed fish at or above maturity length. The survey is conducted in the center of the distribution of the black grouper region; it does not cover the entire distribution of the stock and has a limited depth coverage (compared to adult habitat extending to 100m). Limited depth coverage is also apparent for recreational fishery. These features create potential for hyperstability, where the index might remain high even when the overall population declines. The hyperstability parameter ($b < 1$) for the recreational SRFS index and the fishery independent RVC survey was drawn from uniform distributions: Uniform (0.7, 0.9) for the RVC survey and Uniform (0.6, 0.8) for the SRFS index. These values are based on the meta-analysis by Harley et al. (2001), who examined a large number of CPUE-abundance series across multiple demersal species (e.g., cod, flatfish, and gadiformes) and gear types, finding mean values of 0.64–0.75 for the hyperstability parameter, with most estimates ranging between 0.6 and 0.9. Although that study did not include reef-

associated species or recreational fisheries, it remains the most comprehensive empirical study of hyperstability in the literature and was used here as a proxy for black grouper in the absence of species-specific estimates. A stronger degree of hyperstability was assumed for the recreational index, if recreational fishers actively target known productive locations, effectively tracking areas of highest fish density even when abundance declines.

Performance metrics & evaluation

Performance metrics are statistics that summarize the results of the simulation, and these metrics are used to evaluate how well a specific MP achieves some or all predefined policy goals for the resource (Rademeyer et al. 2007). Within the scope of conducting a strategic MSE, predefined policy goals were considered as those stated in regional fisheries management plans. In the South Atlantic region, the black grouper fishery is managed by the South Atlantic Fishery Management Council under the [Snapper-Grouper Fishery Management Plan](#) (FMP) implemented in 1983.

There are four main goals, each of which aligns with the National Standards to ensure sustainable and responsible fishery management as mandated by the Magnuson-Stevens Fishery Conservation and Management Act (MSFCMA): Goal 1 (Science): Management decisions for the snapper grouper fishery rely on robust science, using both qualitative and quantitative data analyzed promptly and transparently to build stakeholder confidence; Goal 2 (Management): Implement strategies for the snapper grouper fishery that focus on rebuilding and conserving fish stocks, considering regional differences in the fishery and fishing communities' social and economic needs; Goal 3 (Communication): Use interactive outreach strategies to encourage continuous engagement and support two-way communication between managers and

stakeholders of the snapper grouper fishery, enhancing understanding of the science and management involved. Goal 4 (Governance): Adopt a transparent, balanced, expeditious decision-making process that includes flexible but transparent protocols and strategies for managing the snapper grouper fishery.

In the Gulf of America, black grouper are managed under the [Reef Fish Fishery Management Plan](#) implemented in 1984 (GMFMC 1984). The FMP guides the management of reef fish resources under the authority of the Gulf Council, with the aims of: (1) rebuilding declining reef fish stocks, (2) establish a reporting system for monitoring, (3) conserve and enhance reef fish habitats, and (4) minimize conflicts among resource user groups.

MP performance was quantified using the following metrics:

$P(SB_{year} / SB_{final_historical} > 1.0)$ - Probability that spawning biomass in each projection year exceeds the final historical year (baseline). It is calculated across all iterations, where each iteration represents a different realization of uncertainty. By comparing projection performance to the final historical year, we assess whether the management procedure can maintain or improve stock biomass relative to the most recent conditions. This probability will be calculated in the short term (1-10 year projection period), medium term (11-20 year projection period), and long term (>20 years projection period).

Mean catch over projection period - This performance metric quantifies the probability that total catch (aggregated across all areas and fleets) in each projection year exceeds the final historical year catch. It is calculated across iterations to account for uncertainty. This performance metric is

a direct measure of commercial and recreational fishery value that allows for economic planning (processing capacity, markets). It will be calculated for the short, medium, and long term.

P(SPR>30%) - Both the Gulf Council and South Atlantic Council use 30% SPR as an MSY proxy for black grouper. This metric is computed across iterations each year. The SPR reference point represents the reproductive capacity of the stock relative to unfished conditions, with values above 30% generally indicating adequate to maintain stock productivity. By evaluating the probability of maintaining $SPR > 0.3$, we assess whether the management procedure can sustain sufficient spawning biomass to support long-term stock resilience. This performance metric was calculated in the short, medium, and long term.

Results

Reference OMs

Operating model 1 (OM1)

Under OM1, most of the stability-seeking MPs (1–6) maintained spawning biomass relatively close to its final historical level (i.e., 2025) throughout the 50-year projection, indicating that spawning biomass remained near that level on average (Fig. 5; Fig. A1). The inner ribbon (50% CI) stayed relatively tight around the median for all six MPs, while the outer ribbon (95% CI) expanded progressively over time, reflecting increasing uncertainty across iterations in the longer term (Fig. A1). There was similarity between sequential ratio (MPs 1-3) and slope-based MPs (4-6) in spawning biomass trends, suggesting both approaches maintain spawning biomass at similar levels under OM1 conditions. All the stability MPs maintained SPR well above the 0.30

threshold (dotted reference line), given that OM1 simulated a flat historical SPR trend between 0.40 and 0.60 (Fig. 5; Fig. A2). Overall, MPs 1–6 showed stable SPR trajectories around 0.5, with 50% CI bands generally ranging from approximately 0.45 to 0.55, while the 95% CI expanded to roughly 0.35-0.75 by the end of the projection.

Catches for both fleets under stability MPs 1-6 remained relatively stable, showing a median relative catch near 1.0 with moderate uncertainty. The 50% CI remained tight for both fisheries. In comparison, the 95% CI expanded over time, with the lower bound approaching zero in some iterations by the end of the projection, as a subset of iterations experienced substantial long-term catch declines. Overall, we found little difference between sequential ratio (MPs 1-3) and slope-based (MPs 4-6) (Fig. 5; Fig. A3). Like the other stability MPs, MP9 (L_mean_ratio), which uses a length- indicator rather than an index-based one, performed similarly to the index-based stability MPs (Fig. 5; Fig. A1 – A3).

A more detailed comparison of stability MPs was done by selecting three of the most representative MPs: MP3, MP6, and MP9 (Fig. 6). All three MPs exhibited remarkably similar spawning biomass trajectories. Median relative SB remained near 1.0 throughout the projection for the three MPs; however, MP9 (L_mean_ratio) showed a slightly wider 95% ribbon in the upper bound, suggesting more variability among iterations when using length information alone. SPR trajectories were stable around 0.5 for all three MPs, well above the 0.30 target. The minimal differences among MPs suggest that both abundance-based approaches (MP3, MP6) and the length-based approach (MP9) may result in similar stability and corresponding fishing mortality rates. For the same set of three MPs (MP3, MP6, and MP9), Fleet 1 (recreational) and

Fleet 2 (commercial) maintained catches close to baseline levels by the end of the projection period. The three MPs showed nearly identical catch trajectories by the end of the projection, and all converged to similar long-term catch levels, with the lower 90% bound declining toward zero in some iterations across all three MPs. Overall, slope vs. sequential ratio (MP3 vs. MP6) produced similar outcomes despite different mathematical approaches to detecting abundance trends. Additionally, the length-based MP (MP9) performed comparably to abundance-based MPs. This suggests that temporal trends in mean length provided a signal to guide TAC adjustments, at least under the stable conditions simulated in OM1 (Fig. 6).

The two target-seeking length-based MPs (target-seeking MPs; MP7 and MP8) performed differently, consistent with the differences between their respective reference lengths ($LF_{30} = 77.26$ cm, and $LF_{50} = 78.67$ cm). The long-term SPR remained above 30% (or 0.3) for both MPs. However, MP7 systematically moves the stock towards 30% SPR, very slowly over many years, resulting in lower SPR relative to the final historical year and correspondingly higher (Fig. 5, $L_mean_LF_M$ at LF_{30} , pink line). In OM1, noting initial SPR ranged between 0.4 and 0.5, the slight downward trend in biomass towards MP7's intended SPR target of 0.3 is evident in Figure 7.

Three of the multi-indicator MPs (MP10, MP11, MP12) performed similarly across metrics (Fig. 8) with median spawning biomass relative to the last historical year of 1.0 and SPR stable around 0.50 throughout the projection. MP10 (sequential ratio CPUE + Survey + LF_{30}) and MP11 (sequential ratio CPUE + Survey + LF_{50}) exhibited almost indistinguishable trajectories (Fig. 7). The inclusion of LF_{30} indicator does produce a downward trend towards its intended SPR target

of 0.3, but this trend is impeded by the other stability-seeking indicators included in this MP. In contrast, MP12 which does not contain a target-seeking indicator (i.e., MP 12 consists of sequential ratio CPUE + Survey + Lmean ratio) showed a slightly wider upper 95% SB ribbon in the long term, suggesting that the L_mean_ratio indicator produces marginally more variable biomass outcomes, which is consistent with what was observed for MP9 among the single-indicator stability MPs. Catch patterns were comparable for all three MPs. Recreational fleet catch remained near 1.0 at the median, but the lower 95% bound declined over the long term. The commercial fleet catch was more stable with tighter ribbons.

The next three multi-indicator MPs (MP13, MP14, MP15) (Fig. 9) differ from MPs 10-12 (Fig. 8) in that they use the RVC as an index of abundance, excluding the recreational abundance index, and incorporate both RVC and commercial fleet-specific length indicators. For OM1, the outcomes of MP14, and MP15 were like MP10 - MP12. Median SB remained near 1.0, and SPR patterns were similar across all three, with a median of around 0.50. However, the inclusion of LF30 indicator within MP 13 produced a downward trend towards its intended SPR target of 0.3, outperforming all other multi-indicator MPs in this regard.

Probabilistic outcomes for OM1

Spawning Biomass Performance (Fig. 10; Table 8): The stability MPs performed well; they maintained the probability of keeping biomass above the historical year with probabilities of 40–48% in the longer term. This performance aligns with the design of these MPs, which prioritize maintaining stable conditions relative to recent historical years. In that context, a 40–50% probability represents roughly 50/50 odds of being above or below the baseline. Comparing

target-seeking MPs with multi-indicator MPs, slight differences emerged by the end of the projection period. In the long term, MP7 (LF30) had the lowest $P(SB > SB_{hist})$ at 0.439, as this MP was intended to move SPR lower, towards its 30% SPR target. Multi-indicator MPs with survey and length components (MP13-MP15) showed slightly higher probabilities (0.450-0.479), indicating that using multiple information sources offers marginally better biomass maintenance, impeded the effect of LF30 and providing guardrails through inclusion of stability-seeking indicators. However, these differences are minor compared to the overall uncertainty.

SPR Performance (Fig. 11; Table 8): Most MPs maintained ~85% probability of keeping SPR above 30% throughout the projection period. MP15 (seqRatio_survey_Lrvc_Lcom) showed the highest long-term SPR probability (0.922), while MP7 (LF30) showed among the lowest (0.857), suggesting the less conservative LF30 target allows more fishing, which slightly reduces the probability of maintaining SPR above 0.30.

Catch Performance (Fig. 12; Table 8): The probability of exceeding baseline catch differed most among MPs. In the short term, values clustered around 0.49-0.52, near the baseline. MP7 (LF30) and MP8 (LF50) had the highest short-term probabilities (0.512 and 0.517), consistent with the length-based target approach allowing slightly higher catch initially. In the long term, all MPs declined to 0.43-0.51, with MP7 maintaining the highest probability (0.512). The stability MPs (1-6) and L_mean_ratio (MP9) generally had lower catch probabilities than the multi-indicator and MPs, except for multi-indicator MP15, which had the lowest probability in the long term.

Trade-offs under OM1

Short-term (1-10 years): Most MPs were clustered tightly, with median slightly below 1 and median catch near 1, indicating similar performance, as all MPs started with comparable initial TAC levels based on recent historical catches (Fig. 13). The exceptions were MP7 and MP8, positioned in the left upper portion, achieving the highest median catch but with slightly lower SB. MP15 (seqRatio_survey_Lrvc_Lcom) was positioned lowest in the catch axis, despite comparable spawning biomass levels, while MP4 showed the lowest catch among the stability MPs (Fig. 13).

Medium-term (11-20 years): In the medium term, MPs began to spread out. MP7 (LF30) and MP8 (LF50) remained in the upper portion with the highest median catch (~1.01) and SB near 0.98, with the best combined performance in the medium term. The multi-indicator MP10 achieved high catch (~1.01) but at the cost of lower SB (~0.97). The stability MPs (MP1-6) and MP12 clustered in the lower-left area with lower catch and SB (Fig. 13).

Long-term (>20 years): All MPs fell below 1.0 for both catch (0.97 - 1.00) and SB (0.94 - 0.98), indicating a general decline in both metrics over the long term. MP7 (LF30) maintained the highest median catch (~1.00) but with SB around 0.96, representing a trade-off between catch and biomass. MP 8 and MP13 maintained the best overall balance with both high catch (~0.99) and high SB (~0.98). On the opposite end, MP15 and MP6 showed among the lowest catch values, while MP4 was also positioned in the lower portion. The general pattern suggests that MPs incorporating length-based indicators (MP7, MP8, MP13) tended to sustain higher catch in

the long term, while stability MPs (MP1-6) clustered lower (Fig. 13).

Operating model 2 (OM2)

OM2 simulated a more complex management context than OM, in that fishing effort was increased by 0.5% annually during the historical period, resulting in downward trends in biomass between 2000 and 2025 (Figs. 14 -18). All MPs showed absolute differences in performance relative to OM1, but the relative ranking among MPs remained largely consistent (Figs. 14 - 17). The most visible difference from OM1 was a clear downward trend in median spawning biomass relative to the final year of the historical period (2025) across all MPs, declining to approximately 0.60 - 0.80 by the end of the projection period compared to 0.90 - 0.95 under OM1. The 90% confidence intervals also showed lower bounds approaching zero in some iterations by the end of the projection. The comparison panel (Fig. 18) confirms that all MPs followed a similar declining trajectory, with MP7 (LF30, pink line) declining more steeply than the cluster (as intended), reaching approximately 0.60 by the end of the projection.

Median SPR (Figs. 14 - 17) declined gradually from ~0.50 to approximately 0.40 - 0.44 by the end of the projection period but remained above the 0.30 reference target for MPs 3, 6, 9, 7, 8, and 10-15. The lower 90% bound dropped below 0.30 in some iterations for several MPs, particularly MP7 and MP8. The comparison panel (Fig. 18) shows that all MPs exhibit SPR that declines steadily, rather than remaining stable as in OM1.

The increasing effort scenario was clearly reflected in fishing mortality trajectories (Fig. 18). Commercial and recreational F increased to approximately 1.2 - 1.4 relative to historical levels

for most MPs. MP7 again showed the highest F , consistent with its less conservative reference length intended to guide the stock towards 30% SPR, which allows more fishing pressure. Despite increasing F , median catch for both fleets declined slightly over the projection, reflecting that higher effort could not compensate for declining abundance (Fig. 18).

Probabilistic outcomes for OM2

The probability table (Table 9) and Figs. 19 - 21 confirmed the declining trends observed in the projection plots. $P(SB > SB_hist)$ dropped from approximately 0.40 - 0.42 in the short term to 0.35–0.39 in the long term, reflecting the sustained biomass decline under increasing effort. $P(Catch > Catch_hist)$ followed a similar pattern, declining from ~0.47- 0.52 in the short term to 0.37-0.45 in the long term, with MP7 (LF30) again showing the highest catch probabilities but at the cost of lower SB probabilities. $P(SPR > 0.3)$ shows short-term values around 0.97 that decline to 0.75-0.83 in the long term. It is worth noting that MP13 and MP15 maintained the highest long-term $P(SPR > 0.3)$ between 0.82–0.83, suggesting the use of multi-indicators performed the highest at restricting further biomass declines, relative to the other MPs. This is further illustrated by comparison with target-seeking MP7 which showed the lowest $P(SPR > 0.3)$ (0.76) because this MP was intended to accelerate biomass towards its 30% SPR target. Overall, under the pessimistic conditions of OM2, all MPs produced $P(SPR > 0.3)$ above 0.75.

Trade-offs under OM2

The short-term trade-off pattern (Fig. 22) showed most MPs clustered with SB slightly below 1.0 (0.96 - 1.00) and catch near 1.0, with MP12 positioned lower in catch than the rest. However, by the medium term, all MPs shifted to lower SB (0.90 - 0.95), and most fell below 1.0 for catch,

with MP7 (LF30) separating as the top performer, maintaining the highest catch (~ 1.007) and still high SB. MP15 showed the weakest performance in terms of catch. In the long term, all MPs declined substantially (SB: 0.80 - 0.90, Catch: 0.93 - 0.98). MP7 (LF30) maintained the highest catch (~ 0.99) but with the lowest SB (~ 0.88). The multi-indicator MPs (MP10, MP13) maintained a relatively good balance between catch and biomass.

Robustness OM

OM2-R1 tested the effect of hyperstability in both the fishery-independent survey (RVC, $b \sim U(0.7, 0.9)$) and the recreational CPUE (SRFS, $b \sim U(0.6, 0.8)$) under the same increasing fishing effort scenario as OM2. This robustness test was designed to illustrate that indices may remain artificially high while abundance declines, potentially leading MPs to receive overly optimistic signals about the stock size.

Median trends for OM2 – R1

The overall median patterns (Fig. 23) closely resembled OM2, with all MPs showing declining median SB, declining SPR, increasing F, and declining catch. MP7 (LF30, pink line) again separated from the cluster with the most different trajectory. It showed a steeper SB decline and higher F. Two notable differences emerged relative to OM2. First, MP9 (L_mean_ratio) diverged more strongly in SB, declining to approximately 0.72 by the end of the projection compared to the main cluster. Second, the overall spread among MPs was slightly wider than under OM2, suggesting hyperstability amplifies differences in MP responsiveness.

Trade-offs under OM2 - R1

The short-term patterns were similar to OM2, with MPs clustered near 1.0 for both metrics. MP9 (L_mean_ratio) was positioned in the upper left portion of the plot, with the highest short-term catch and the lowest spawning biomass. In the medium term, MP9 again had the highest catch, while MP8 maintained a high catch with higher SB. In the long term, all MPs declined below 1.0 for both metrics (SB: 0.80 - 0.92, Catch: 0.93 - 1.00), with MP8 (LF50) maintaining the best balance between catch and spawning biomass, while MP1, 4, 12, and 15 showed the weakest performance in terms of catch levels while maintaining spawning biomass similar to the main cluster of MPs (Fig. 24).

Probabilistic outcomes for OM2 – R1

Comparing the probability tables for the six MPs (MP1 -6) that rely solely on survey and/or CPUE indices showed that, in the short-term, values were virtually identical between OM2 and OM2-R1 across all metrics. $P(\text{SB} > \text{SB_hist})$, $P(\text{Catch} > \text{Catch_hist})$, and $P(\text{SPR} > 0.3)$ (Table 10). However, $P(\text{Catch} > \text{Catch_hist})$ showed a consistent pattern: MPs using CPUE showed lower catch probabilities under OM2-R1. MP4 (islope_cpue) dropped from 0.507 to 0.462, MP3 (seqRatio_cpue_survey) from 0.500 to 0.472, and MP6 (islope_cpue_survey) from 0.504 to 0.483. This is consistent with the hypothesis of CPUE having stronger hyperstability ($b = 0.6\text{--}0.8$) than the survey ($b = 0.7\text{--}0.9$).

In the medium term, the pattern continued. MP4 (islope_cpue) showed the largest drop in $P(\text{Catch} > \text{Catch_hist})$, from 0.494 to 0.454, while MP6 (islope_cpue_survey) dropped from

0.489 to 0.471. Survey-only MPs (MP2, MP5) showed smaller or negligible changes, consistent with the lower hyperstability applied to the survey index. $P(\text{SPR} > 0.3)$ showed slight declines for some MPs (e.g., MP2 from 0.928 to 0.917, MP6 from 0.916 to 0.909) but remained above 0.90.

In the long term, differences between OM2 and OM2-R1 became less consistent across MPs, with $P(\text{SPR} > 0.3)$ showing mixed results, some MPs slightly higher, others slightly lower. This may suggest that over 50 years, the cumulative effects of increasing fishing effort had a greater impact than hyperstability signals. In summary, the most detectable impact of hyperstability was on CPUE-dependent MPs in the short- and medium-term, particularly on catch performance. MPs relying on the survey (less hyperstability) were less affected. However, all differences were modest in terms of probability changes. In addition the $\pm 10\%$ stability clause likely limited the extent to which biased signals from indices could translate into substantially different TAC recommendations.

Discussion

Design and evaluation of multi-indicator MPs was used to illustrate the potential for management responsiveness in the absence of stock assessment. This desk MSE provides a structured approach to evaluate how different MPs perform across plausible uncertainty scenarios (e.g., uncertainty in steepness, depletion, recruitment variability, and historical fishing mortality trends). By simulating each OM against each MP, outcomes in the face of the uncertainty were quantified using performance metrics.

Design of MPs and related OM specifications were integrated within an MSE to examine strategic policy goals, rather than tactical decisions (Punt et al. 2016). Strategic MSE relies on predefined management goals and related generic performance metrics, whereas tactical MSE require a more in-depth process of ongoing engagement between stakeholders and analysts (Walter III et al. 2023). Strategic MSE can provide coarse policy guidance based on representative yet simplified simulations of a fishery system. Conversely, tactical MSE is likely to necessitate a more subtle and detailed representation of the statistical properties of various elements of the fishery system, built upon analysis of actual raw data streams.

Synthesis

The performance of the MPs revealed a clear hierarchy, with the multi-indicator MPs 13, 14, and 15 (Table 4; MP 13: seqRatio_survey_Lrvc_Lcom_LF_M(LF30); MP 14: seqRatio_survey_Lrvc_Lcom_LF_M(LF50); MP 15: seqRatio_survey_Lrvc_Lcom) emerging as potentially beneficial MPs that are worth further consideration and refinement. The combination of stability-seeking and target-seeking MPs, while sometimes working slightly antagonistically, also beneficially function as guardrails to TAC adjustments where data precision is low. Under OM1, stable biomass conditions, the inclusion of the target-seeking LF30 indicator within MP 13 illustrates movement of the stock towards its intended SPR target of 0.3, outperforming all other multi-indicator MPs in this regard. For a long-lived species, the stability indicators included in MP 13 slow this process, avoiding over-shooting targets due to time-lags between changes in abundance and corresponding length composition, as well as the cyclical and ratcheting effects

sometimes observed with length-based indicators (Klaer et al. 2012; Hordyk et al. 2015; Bellquist et al. 2022).

Within MPs 13, 14, and 15, the inclusion of ‘primarily’ sub-adults from the RVC length composition as well as length composition from adults retained in the commercial catch are worth elevating. While our contrasts of MPs did not demonstrate disparate performance of MPs containing Lrv and Lcom indicators (see Table 4), moving forward, it is likely worth considering further refinement of how these indicators influence TAC adjustment. The concept of tracking cohort strength, weak or strong, and responding accordingly, is well established in the fisheries literature and has implications for long-lived species where a few cohorts can dominate the catch over several years (Walters and Martell 2004; their chapter 3). For example, each of these three MPs were aimed at detecting a larger cohort may be entering the fishery, producing a ‘wait and see’ response that delays increasing harvest of a potentially large cohort, allowing the cohort to grow. Likewise, mean length within the commercial fishery (longline fishery; see Tables 2 & 4) that tracks demographic attributes of adults is likely worthwhile monitoring. While mean length within the commercial fishery cannot be used within analytical length-based stock assessments (i.e., LB-SPR) because removals occurring in the recreational fishery would render assessments of commercial lengths as inaccurate, relative changes in mean length could reflect changes in commercial fishing pressure.

The comparison between OM1 (stable historical fishing mortality) and OM2 (increasing historical fishing mortality) demonstrates the value of testing MPs under contrasting assumptions about stock history. The finding that most MPs differed in performance under OM2 reveals that

the scale and constraints on TAC adjustments imposed by design of these MPs may require additional refinement, and perhaps exploration of other complementary management actions. Similar results related to the degree of responsiveness required under increasing fishing effort have been illustrated elsewhere (Harford et al. 2016). Within U.S. fishery management policy, the circumstance of rebuilding also needs to be considered in context of rebuilding requirements under NS1, which may require a different set of decision-making criteria (i.e., recovery trajectories obtained within a set timeframe).

How to refine the operating models?

For black grouper, refinement of OM_s could focus on several priorities. First, improvement of OM realism could focus on deeper assessment of the statistical properties of empirical data for catch, indices, and length composition. Data quality and quantity is likely to determine viable MP_s, thus additional allocation of effort to refining observation sub-models is suggested. Second, life history parameters for growth, maturation, and sex change were constant in OM_s, and likewise, these quantities were used for Target-seeking indicators (i.e., *L_{target30}* and *L_{target50}*). This issue of MP_s having ‘perfect knowledge’ of some aspects of black grouper life history should be examined using OM_s where these parameters are inconsistent (or mismatched) with those used in MP_s to examine for robustness to life history uncertainty. Third, the joint management of the black grouper stock by the Gulf Council and the South Atlantic Council, along with catch allocation with these regions could necessitate a two-area model. This additional model complexity could also enable higher resolution in observation sub-models, if geographic differences persist in sampling methodologies. Related to allocation, there is also a

need to introduce implementation error (i.e., TAC are not precisely removed from the stock) to identify MPs that are robust to implementation challenges.

How to refine the management procedures?

The sequential approach of testing single-indicator MPs, followed by multi-indicator MPs, was beneficial from a few perspectives. From an educational perspective, the approach builds understanding of MP design and performance expectations, rather than proceeding directly to more complex multi-indicator MPs. From a resource monitoring perspective, the approach provides a pathway for management to prioritize data streams, potentially focusing on refinement of data quality and quantity. For example, stability-based MPs can provide a baseline for MP complexity, and could be modified as additional data collection improves, including the addition of other indicators, achieving a multi-indicator approach. Finally, from an adaptability perspective, this approach provides an early iteration of a roadmap towards MP refinement, where MP design evolves with growing knowledge rather than remaining static.

The product of the MSE presented herein is a technical analysis that quantifies performance of MPs in terms of expected achievement of management objectives. Once an understanding of MP performance is revealed to managers and stakeholders, refinement of the larger set of MPs is often required, including removal of those strategies which perform poorly. The reader is directed to two approaches for narrowing and selection of an MP(s). First, consider *satisficing*: selecting an MP or narrowing to a sub-set of MPs can involve identifying those MPs that meet some minimum performance criteria across a set of standards (Miller and Shelton 2010). A related approach involves deciding whether an MP poses sufficiently low risk to the resource,

including irreparable damage, across a broad set of circumstances (i.e., OM conditions) (Darcy and Matlock 1999). Second, consider *ranked order changes of MPs*: in this case, MPs are ranked based on performance. Various sensitivity analyses or robustness tests are used to understand whether the ranked order of MPs would change under different conditions. This process is a component of decision-analysis, which is used to evaluate whether ranking of options is robust to uncertainty, with better options holding-up despite uncertainty (Peterman 2004; Więckowski and Sałabun 2023).

Narrowing of MP options is sometimes followed by fine-tuning of the remaining strategies, which can require additional MSE. Where fine-tuning of MPs is of interest, the following guidance is provided. First, *try to avoid generalized MPs*: development of data-limited MPs has been rapid, with several examples of attempts to generalize MP approaches (Geromont and Butterworth 2014). However, more recent arguments have been made that bespoke MPs or non-generic MPs, perhaps built upon the shoulders of generic foundations, are likely to lead to better solutions (Sagarese et al. 2018; Dowling et al. 2019; Kell et al. 2022). Thus, it is worth considering more nuanced exploration of operational control points (OCPs), for example, the weight parameters on *Sequential ratio* indicators, perhaps especially for MPs 13, 14 and 15. Further, if a more realistic OM is developed, it is worth ‘tuning’ OCPs against this new level of realism. Second, *explore model-based MPs*: exploration of other indicators as well as some forms of model-based assessment could be considered (e.g., Stock Synthesis Data-limited tool; <https://github.com/shcaba/SS-DL-tool>). Third, *refine performance metrics*: generic performance metrics could be expanded or replaced with those most relevant to stakeholders and decision-makers (Punt 2017). Related, refinement of MPs should consider how MPs can be implemented

in practice to best meet legal mandates and reference points as set out in NS1. For example, consider transforming the generic TAC setting approaches presented here to those that align more directly with existing tier-based control rules reflecting degree of severity in scientific uncertainty (i.e., South Atlantic Fisheries Management Council; Free et al. 2023). Finally, *focus on process*: the narrowing and refinement of MPs illustrates that MSE becomes a constituent tool within a larger process of manager and stakeholder engagement that is both transparent and pragmatic and can expectantly lead to support for implementing a viable management strategy.

Conclusions

The black grouper fishery lacks a conventional stock assessment, leaving decision-makers without the typically available tools to guide harvest decisions. Here, a process was initiated for progress towards design and evaluation of multi-indicator MPs to confront issues of data quality and availability, robustness to ecological uncertainties, and exploration of alternative data streams to inform aspects of stock status within a cohesive decision-making framework. At least three multi-indicator MPs were demonstrated with potential to produce stability in stock biomass while subtly guiding longer-term outcomes towards a Council prescribed sustainability target. Two next steps are proposed, consisting of (i) in-depth refinement of the OM(s) to better reflect specific characteristics of the black grouper fishery system through analysis of empirical data and (ii) engaging in refinement of MPs through directed collaborative opportunities. With continued work, an MP approach could lead to an interim policy for decision-making while improved data collection is considered and until conventional stock assessment becomes viable.

References

- Anderson, S. 2021. A management procedure framework for groundfish in British Columbia. Canadian Science Advisory Secretariat (CSAS).
- Apostolaki, P., and Hillary, R. 2009. Harvest control rules in the context of fishery-independent management of fish stocks. *Aquat. Living Resour.* **22**(2): 217–224. doi:10.1051/alr/2009022.
- Ault, J.S., and Smith, S.G. 2009. Ault-Smith Notes on Reef-fish Visual Census (RVC) Population Statistics Estimation for Black Grouper (*Mycteroperca bonaci*) and Red Grouper (*Epinephelus mori*) in the Florida Keys and Dry Tortugas Regions. DRAFT. S19DW10.
- Bellquist, L., Harford, W.J., Hurd, F., Jackson, A., Prince, J.D., Freiwald, J., Neumann, A., Likins, J., and Wilson, J.R. 2022. Use of management strategy evaluation to understand the value of citizen science in managing an iconic California recreational fishery. *Estuar. Coast. Shelf Sci.* **278**: 108112. doi:10.1016/j.ecss.2022.108112.
- Bentley, N., and Stokes, K. 2009. Contrasting Paradigms for Fisheries Management Decision Making: How Well Do They Serve Data-Poor Fisheries? *Mar. Coast. Fish.* **1**(1): 391–401. doi:10.1577/C08-044.1.
- Beverton, R.J.H., and Holt, S.J. 1957. On the dynamics of exploited fish populations. Chapman and Hall, London UK.
- Brandt, M.E., Zurcher, N., Acosta, A., Ault, J.S., Bohnsack, J.A., Feeley, M.W., Harper, D.E., Hunt, J.H., Kellison, T., McClellan, D.B., Patterson, M.E., and Smith, S.G. 2009. A cooperative multi-agency reef fish monitoring protocol for the Florida Keys coral reef ecosystem. National Park Service, United States Department of the Interior. Natural Resource Report NPS/SFCN/NRR - 2009/150.
- 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**(4): 318–328. Cambridge University Press. doi:10.1017/S0376892913000271.
- Darcy, G.H., and Matlock, G.C. 1999. Application of the precautionary approach in the national standard guidelines for conservation and management of fisheries in the United States. *ICES J Mar Sci* **56**(6): 853–859. doi:10.1006/jmsc.1999.0533.
- Dowling, N.A., Smith, A.D.M., Smith, D.C., Parma, A.M., Dichmont, C.M., Sainsbury, K., Wilson, J.R., Dougherty, D.T., and Cope, J.M. 2019. Generic solutions for data-limited fishery assessments are not so simple. *Fish Fish.* **20**(1): 174–188. doi:10.1111/faf.12329.
- Free, C.M., Mangin, T., Wiedenmann, J., Smith, C., McVeigh, H., and Gaines, S.D. 2023. Harvest control rules used in US federal fisheries management and implications for climate resilience. *Fish Fish.* **24**(2): 248–262. doi:10.1111/faf.12724.
- Geromont, H.F., and Butterworth, D.S. 1998. Management procedure robustness trials for west coast hake. Marine and Coastal Management South Africa Document WG/06/98/D:H:36. 13 pp.
- Geromont, H.F., and Butterworth, D.S. 2014. Generic management procedures for data-poor fisheries: forecasting with few data. *ICES J Mar Sci* **72**(1): 251–261. doi:10.1093/icesjms/fst232.

- Geromont, H.F., and Butterworth, D.S. 2015. Complex assessments or simple management procedures for efficient fisheries management: a comparative study. *ICES J Mar Sci* **72**(1): 262–274. doi:10.1093/icesjms/fsu017.
- Gulf of Mexico Fishery Management Council (GMFMC). 1984. Gulf Reef Fish Fishery Management Plan. <https://www.fisheries.noaa.gov/management-plan/gulf-reef-fish-fishery-management-plan>
- Hamel, O.S., and Cope, J.M. 2022. Development and considerations for application of a longevity-based prior for the natural mortality rate. *Fish. Res.* **256**: 106477. doi:10.1016/j.fishres.2022.106477.
- Harford, W.J. 2024. FishSimGTG: an $\{R\}$ package for simulation of fish population dynamics and evaluation of management strategies. <https://github.com/natureanalytics-ca/fishSimGTG>.
- Harford, W.J., Amoroso, R., Bell, R.J., Caillaux, M., Cope, J.M., Dougherty, D., Dowling, N.A., Hurd, F., Lomonico, S., Nowlis, J., Ovando, D., Parma, A.M., Prince, J.D., and Wilson, J.R. 2021. Multi-Indicator Harvest Strategies for Data-Limited Fisheries: A Practitioner Guide to Learning and Design. *Front. Mar. Sci.* **8**. Frontiers. doi:10.3389/fmars.2021.757877.
- Harford, W.J., and Babcock, E.A. 2016. Aligning monitoring design with fishery decision-making: examples of management strategy evaluation for reef-associated fisheries. *Aquat. Living Resour.* **29**(2): 205. doi:10.1051/alr/2016018.
- Harford, W.J., Gedamke, T., Babcock, E.A., Carcamo, R., McDonald, G., and Wilson, J.R. 2016. Management strategy evaluation of a multi-indicator adaptive framework for data-limited fisheries management. *Bull Mar Sci* **92**(4): 423–445. doi:10.5343/bms.2016.1025.
- Harford, W.J., Sagarese, S.R., and Karnauskas, M. 2019. Coping with information gaps in stock productivity for rebuilding and achieving maximum sustainable yield for grouper–snapper fisheries. *Fish Fish.* **20**(2): 303–321. doi:10.1111/faf.12344.
- Harley, S.J., Myers, R.A., and Dunn, A. 2001. Is catch-per-unit-effort proportional to abundance? *Can. J. Fish. Aquat. Sci.* **58**(9): 1760–1772. NRC Research Press. doi:10.1139/f01-112.
- Henriquez, V., Moore, P., and Harford, W.J. 2024. Design of alternative management procedures (MPs) for black grouper fisheries: An exploration of MP viability and data limitations. Report prepared for: State of Florida Florida Fish and Wildlife Conservation Commission.
- Hordyk, A.R., Loneragan, N.R., and Prince, J.D. 2015. An evaluation of an iterative harvest strategy for data-poor fisheries using the length-based spawning potential ratio assessment methodology. *Fish Res* **171**: 20–32. doi:10.1016/j.fishres.2014.12.018.
- Huntsman, G.R., Potts, J., and Mays, R.W. 1994. A preliminary assessment of the populations of seven species of grouper (Serranidae, Epinephelinae) in the western Atlantic ocean from Cape Hatteras, North Carolina to the Dry Tortugas, Florida. *Proceedings. 43rd Gulf and Caribbean Fisheries Institute* 193-213.
- ICES. 2015. Report of the fifth Workshop on the Development of Quantitative Assessment Methodologies based on Life-history traits, exploitation characteristics, and other relevant parameters for data-limited stocks (WKLIFE V). report, ICES Expert Group reports (until 2018). doi:10.17895/ices.pub.19283927.v1.

- 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. doi:10.1016/j.fishres.2014.11.013.
- Kell, L.T., Minto, C., and Gerritsen, H.D. 2022. Evaluation of the skill of length-based indicators to identify stock status and trends. *ICES J. Mar. Sci.* **79**(4): 1202–1216. doi:10.1093/icesjms/fsac043.
- Klaer, N.L., Wayte, S.E., and Fay, G. 2012. An evaluation of the performance of a harvest strategy that uses an average-length-based assessment method. *Fish. Res.* **134–136**: 42–51. doi:10.1016/j.fishres.2012.08.010.
- MacCall, A.D. 2009. Depletion-corrected average catch: a simple formula for estimating sustainable yields in data-poor situations. *ICES J. Mar. Sci. J. Cons.* **66**(10): 2267–2271. doi:10.1093/icesjms/fsp209.
- Magnusson, A., and Hilborn, R. 2007. What makes fisheries data informative? *Fish Fish.* **8**: 337–358.
- Medeiros-Leal, W., Santos, R., Peixoto, U.I., Casal-Ribeiro, M., Novoa-Pabon, A., Sigler, M.F., and Pinho, M. 2023. Performance of length-based assessment in predicting small-scale multispecies fishery sustainability. *Rev. Fish Biol. Fish.* **33**(3): 819–852. doi:10.1007/s11160-023-09764-9.
- Methot, R.D., and Wetzel, C.R. 2013. Stock synthesis: A biological and statistical framework for fish stock assessment and fishery management. *Fish. Res.* **142**: 86–99. doi:10.1016/j.fishres.2012.10.012.
- 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. doi:10.1093/icesjms/fsq083.
- National Standard 1 - Optimum Yield. 2022. National Standard 1—Optimum Yield. 50 CFR § 600.310. <https://www.govinfo.gov/app/details/CFR-2022-title50-vol12/CFR-2022-title50-vol12-sec600-310>.
- NEFMC. 2022. Northeast skate complex fishery management plan 2022. New England Fishery Management Council (NEFMC).
- O’Hop, J., and Beaver, R. 2009. Age, growth, and maturity of black grouper (*Mycteroperca bonaci*) - Crabtree and Bullock (1998) revisited. SouthEast Data, Assessment, and Review (SEDAR) 19. Atlantic States Marine Fisheries Commission. SEDAR19-DW-09.
- Peterman, R.M. 2004. Possible solutions to some challenges facing fisheries scientists and managers. *ICES J Mar Sci* **61**: 1331–1343.
- PFMC. 2020. Pacific coast groundfish fishery management plan: for the California, Oregon, and Washington groundfish fishery. Pacific Fishery Management Council (PFMC), Portland, Oregon, USA.
- Punt, A.E. 2017. Strategic management decision-making in a complex world: quantifying, understanding, and using trade-offs. *ICES J. Mar. Sci.* **74**. doi:10.1093/icesjms/fsv193.
- 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**: 303–334. doi:10.1111/faf.12104.
- 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. doi:10.1093/icesjms/fsm050.

- Rice, J.C., and Rochet, M.-J. 2005. A framework for selecting a suite of indicators for fisheries management. *ICES J. Mar. Sci.* **62**(3): 516–527. Oxford Academic.
doi:10.1016/j.icesjms.2005.01.003.
- Sagarese, S.R., Rios, A.B., Cass-Calay, S.L., Cummings, N.J., Bryan, M.D., Stevens, M.H., Harford, W.J., McCarthy, K.J., and Matter, V.M. 2018. Working Towards a Framework for Stock Evaluations in Data-Limited Fisheries. *North Am. J. Fish. Manag.* **38**(3): 507–537. doi:10.1002/nafm.10047.
- SEDAR. 2010. Stock assessment report: Gulf of Mexico and South Atlantic Black Grouper. South Atlantic Fishery Management Council. SEDAR 19.
- SEDAR. 2017. SEDAR 48 - Southeastern U.S. black grouper data workshop report.
- Shertzer, K.W., and Conn, P.B. 2012. Spawner-Recruit Relationships of Demersal Marine Fishes: Prior Distribution of Steepness. *Bull Mar Sci* **88**(1): 39–50.
doi:10.5343/bms.2011.1019.
- Smith, S.G., Ault, J.S., Bohnsack, J.A., Harper, D.E., Luo, J., and McClellan, D.B. 2011. Multispecies survey design for assessing reef-fish stocks, spatially-explicit management performance, and ecosystem condition. *Fish. Res.* **109**: 25–41.
- Thorson, J.T. 2020. Predicting recruitment density dependence and intrinsic growth rate for all fishes worldwide using a data-integrated life-history model. *Fish Fish.* **21**(2): 237–251.
doi:10.1111/faf.12427.
- Walter III, J.F., Peterson, C.D., Marshall, K., Deroba, J.J., Gaichas, S., Williams, B.C., Stohs, S., Tommasi, D., and Ahrens, R. 2023. When to conduct, and when not to conduct, management strategy evaluations. *ICES J. Mar. Sci.* **80**(4): 719–727.
doi:10.1093/icesjms/fsad031.
- Walters, C.J., and Martell, S.J.D. 2004. *Fisheries Ecology and Management*. Princeton University Press, USA.
- Więckowski, J., and Sałabun, W. 2023. Sensitivity analysis approaches in multi-criteria decision analysis: A systematic review. *Appl. Soft Comput.* **148**: 110915.
doi:10.1016/j.asoc.2023.110915.

Tables

Table 1. Reproduced from Henriquez et al. (2024). Categorization and reliability of the available data. The green circle represents good or acceptable quality; the yellow circle represents medium quality, and the red circle represents bad quality.

Data	Description	Data category	Data reliability
A50%	Age at 50% maturity	Life history	●
L50%	Length at 50% maturity	Life history	●
M	Natural mortality	Life history	●
Linf	Asymptotic length in VBGF	Life history	●
K	Growth parameter in VBGF	Life history	●
t0	Age at zero length in VBGF	Life history	●
<i>a</i>	Parameter <i>a</i> in weight–length relationship	Life history	●
<i>b</i>	Parameter <i>b</i> in weight–length relationship	Life history	●
A50% transition	Age at 50% of transition	Life history	●
L50% transition	Length at 50% of transition	Life history	●
Ct	Commercial landings	Catch	●
Crt	Recreational landings	Catch	●
CAL	Commercial length samples	Catch composition	●
Reef Fish Visual Census (RVC)	Fishery independent survey (sub-adults)	Population indices	●
MRIP (South Atlantic)	Fishery dependent survey	Population indices	●
Southeast Regional Headboat Survey (SRHS)	Fishery dependent survey	Population indices	●

Table 2. Indicators specified from chosen data streams. FD is fishery-dependent and FI is fishery-independent.

Observation model	Source	Information type	Linkage to fishery	Indicators
Recreational fishery (FD)	SRFS	Annual abundance index Common length: ~61 to 100 cm ~24 to 39 inches	Primarily tracks female sub-adult to adult abundance.	1. Stability: index slope of recent directional trend. 2. Stability: sequential ratio of recent directional trend.
Recreational fishery (FD)	SRFS	Annual mean length in the catch. Common length: ~61 to 100 cm ~24 to 39 inches	Tracks history of recreational fishing pressure and cohort size fluctuations.	1. Stability: average ratio of recent directional trend.
Commercial fishery (FD)	Longline	Annual mean length in the catch. Common length: > ~86 cm > 34 inches	Tracks history of commercial fishing pressure and cohort size fluctuations.	1. Stability: average ratio of recent directional trend.
Catch(N)-weighted recreational & commercial (FD)	SRFS& Longline	Annual mean length in the catch.	Tracks history of fishing pressure and cohort size fluctuations.	1. Stability: average ratio of recent directional trend. 2. Target seeking: relative to proxy MSY target
Survey (FI)	RVC	Annual abundance index Common length: ~25 to 100 cm ~10 to 39 inches	Primarily tracks female sub-adult to adult abundance.	1. Stability: index slope of recent directional trend. 2. Stability: sequential ratio of recent directional trend.
Survey (FI)	RVC	Annual mean length Common length: ~25 to 100 cm ~10 to 39 inches	Tracks history of fishing pressure and cohort size fluctuations.	1. Stability: average ratio of recent directional trend.

Table 3. Single-indicator management procedures. HCR is harvest control rule, *I* is the indicator data source.

Name	Type	Indicator data source	HCR
MP1: seqRatio_cpue	Stability	I = Recreational fishery SRFS, annual abundance index	$TAC\ Preliminary = TAC\ Previous \times SeqRatio_Adj[I]$
MP2: seqRatio_survey	Stability	I = Survey (FI) RVC, annual abundance index	$TAC\ Preliminary = TAC\ Previous \times SeqRatio_Adj[I]$
MP3: seqRatio_cpue_survey	Stability	I1 = Recreational fishery SRFS, annual abundance index I2 = Survey (FI) RVC, annual abundance index.	$TAC\ Preliminary$ $= TAC\ Previous$ $\times mean(SeqRatio_Adj[I1], SeqRatio_Adj[I2])$ (weighted average of two indices using equal weights)
MP 4: islope_cpue	Stability	I = Recreational fishery SRFS, annual abundance index	$TAC\ Preliminary = TAC\ Previous \times Slope_Adj[I]$
MP 5: islope_survey	Stability	I = Survey (FI) RVC, annual abundance index	$TAC\ Preliminary = TAC\ Previous \times Slope_Adj[I]$
MP 6: islope_cpue_survey	Stability	I1 = Recreational fishery SRFS, annual abundance index I2 = Survey (FI) RVC, annual abundance index.	$TAC\ Preliminary = TAC\ Previous \times mean(Slope_Adj[I1], Slope_Adj[I2])$ (weighted average of two indices using equal weights)
MP 7: L_mean_LF_M(LF30)	Target	I = Catch(N)-weighted recreational & commercial (FD) SRFS& Longline, annual mean length in the catch.	$TAC\ Preliminary = TAC\ Previous \times LTarget30_Adj[I]$
MP 8: L_mean_LF_M(LF50)	Target	I = Catch(N)-weighted recreational & commercial (FD) SRFS& Longline, annual mean length in the catch.	$TAC\ Preliminary = TAC\ Previous \times LTarget50_Adj[I]$
MP 9: L_mean_ratio	Stability	I = Catch(N)-weighted recreational & commercial (FD) SRFS& Longline, annual mean length in the catch.	$TAC\ Preliminary = TAC\ Previous \times LenRatio_Adj[I]$

Table 4. Multi-indicator management procedures. HCR is harvest control rule, *I* is the indicator data source.

Name	Type	Indicator data source	HCR
MP 10: seqRatio_cpue_survey_LF_M(LF30)	Stability & Target	I1 = Recreational fishery SRFS, annual abundance index I2 = Survey (FI) RVC, annual abundance index. I3 = Catch(N)-weighted recreational & commercial (FD) SRFS& Longline, annual mean length in the catch.	<i>TAC Preliminary</i> $= TAC Previous$ $\times mean(SeqRatio_Adj[I1], SeqRatio_Adj[I2])$ $\times LTarget30_Adj[I3]$
MP 11: seqRatio_cpue_survey_LF_M(LF50)	Stability & Target	I1 = Recreational fishery SRFS, annual abundance index I2 = Survey (FI) RVC, annual abundance index. I3 = Catch(N)-weighted recreational & commercial (FD) SRFS& Longline, annual mean length in the catch.	<i>TAC Preliminary</i> $= TAC Previous$ $\times mean(SeqRatio_Adj[I1], SeqRatio_Adj[I2])$ $\times LTarget50_Adj[I3]$
MP 12: seqRatio_cpue_survey_L_mean_ratio	Stability	I1 = Recreational fishery SRFS, annual abundance index I2 = Survey (FI) RVC, annual abundance index. I3 = Catch(N)-weighted recreational & commercial (FD) SRFS& Longline,	<i>TAC Preliminary</i> $= TAC Previous$ $\times mean(SeqRatio_Adj[I1], SeqRatio_Adj[I2])$ $\times LenRatio_Adj[I3]$

		annual mean length in the catch.	
MP 13: seqRatio_survey_Lrvc_Lcom_LF_M(LF30)	Stability & Target	I1 = Survey (FI) RVC, annual abundance index. I2 = Survey (FI) RVC, annual mean length. I3 = Commercial fishery (FD), longline, annual mean length in the catch. I4 = Catch(N)-weighted recreational & commercial (FD) SRFS& Longline, annual mean length in the catch.	<i>TAC Preliminary</i> $= TAC_{Previous} \times SeqRatio_{Adj}[I1]$ $\times LenRatio_{Adj}[I2]$ $\times LenRatio_{Adj}[I3]$ $\times LTarget30_{Adj}[I4]$
MP 14: seqRatio_survey_Lrvc_Lcom_LF_M(LF50)	Stability & Target	I1 = Survey (FI) RVC, annual abundance index. I2 = Survey (FI) RVC, annual mean length. I3 = Commercial fishery (FD), longline, annual mean length in the catch. I4 = Catch(N)-weighted recreational & commercial (FD) SRFS& Longline, annual mean length in the catch.	<i>TAC Preliminary</i> $= TAC_{Previous} \times SeqRatio_{Adj}[I1]$ $\times LenRatio_{Adj}[I2]$ $\times LenRatio_{Adj}[I3]$ $\times LTarget50_{Adj}[I4]$

MP 15: seqRatio_survey_Lrvc_Lcom	Stability	I1 = Survey (FI) RVC, annual abundance index. I2 = Survey (FI) RVC, annual mean length. I3 = Commercial fishery (FD), longline, annual mean length in the catch.	<i>TAC Preliminary</i> $= TAC_{Previous} \times SeqRatio_{Adj}[I1]$ $\times LenRatio_{Adj}[I2]$ $\times LenRatio_{Adj}[I3]$
-------------------------------------	-----------	---	--

Table 5. Life history parameters used in operating models. OM is operating model.

Parameter	Quantity used in OM		Notes and citation
	Point estimate	Range	
Growth, maturity, sex transition			
Length type			TL
Length units			cm
Weight units			kg
Loo von Bertalanffy	136.66		Recommended by SEDAR 48 (SEDAR 2017)
K von Bertalanffy (year ⁻¹)	0.1338		
t0 von Bertalanffy	-0.4568		
Length at 50% maturity	85.56		Recommended by SEDAR 48 (SEDAR 2017)
Length-increment to 95% maturity	11.41		
Length-weight alpha	1.023602e-05		Recommended by SEDAR 48 (SEDAR 2017)
Length-weight beta	3.0722		
Length at 50% male	120.8		Recommended by SEDAR 48 (SEDAR 2017)
Length increment to 95% male	18.40		
CV on Linf across growth-type-groups	0.1		Consistent with variation in length-at-age in SEDAR 48 (SEDAR 2017)
Productivity parameters			
Natural mortality (year ⁻¹)		(0.13, 0.20)	Uniform (min, max) 95% confidence interval from (Hamel and Cope 2022)
Log-normal recruitment deviations		(0.2, 0.6)	Uniform (min, max)
Steepness		(2.6, 1.42)	Beta(shape,scale)Similar to (Shertzer and Conn 2012)

Table 6. Parameters for fishery characteristics and fishery-independent survey.

Object	Parameters	Notes and citation
Commercial fishery		
Length type	TL	
Length units	cm	
Length at 50% selectivity	86.27	Logistic selectivity function (SEDAR 19, 2010)
Length-increment to 95% selectivity	24.57	Logistic selectivity function (SEDAR 19, 2010)
Knife-edge retention	60.96	Current minimum size limit
Discard mortality	37.5%	Recommended by SEDAR 48 (SEDAR 2017)
Recreational fishery		
Length type	TL	
Length units	cm	
Length at the peak of selectivity	38.3	Tuned parameter
Knife-edge retention	60.96	Current minimum size limit
Discard mortality	17.5%	Recommended by SEDAR 48 (SEDAR 2017)
Fishery-independent survey (RVC)		
Length type	TL	
Length units	cm	
Length at the peak of selectivity	24.2	Dome-shaped using exponential logistic SEDAR 19 (2010)

Table 7. Parameters for fishery-dependent and fishery-independent observation model sub-routines used in reference operating models.

Object	Parameters	Notes and citation
Recreational index		abundance
Catchability coefficient (q)	(2.62e-06, 4.07e-06)	Uniform (min, max) SEDAR 19 (2010)
Coefficient of Variation (CV)	(0.36, 0.95)	Uniform (min, max) Assumed CV on CPUE $\geq$ CV on catches
Recreational catch		
Coefficient of Variation (CV)	0.42	Abundance; Analysis provided by FWRI
Coefficient of Variation (CV)	0.32	Biomass; Analysis provided by FWRI
Recreational length composition		TL, cm
Multinomial sample size length composition	100	Highly uncertain, 100 used as initial guess
Commercial length composition		TL, cm
Multinomial sample size length composition	100	Typically > 100, quantity used as initial guess
Commercial catch		
Coefficient of Variation (CV)	0.1	
Coefficient of Variation (CV)	0.1	
Fishery-independent index (RVC)		abundance
Catchability coefficient (q)	(1.56e-06, 2.41e-06)	SEDAR 19 (2010)
Coefficient of Variation (CV)	(0.086, 0.208)	SEDAR 48 (2017)
Fishery-independent length composition (RVC)		TL, cm
Multinomial sample size length composition	150	Unknown at time of reporting, 150 initial guess

Table 8. Average performance of MPs for OM1 across short-term (years 1–10), medium-term (years 11–20), and long-term (years 21–50) projection periods. Performance metrics include the probability that spawning biomass exceeds the final historical year level ($P(SB > SB_{hist})$), the probability that catch exceeds the final historical year level ($P(Catch > Catch_{hist})$), and the probability that the spawning potential ratio exceeds the 0.3 reference point ($P(SPR > 0.3)$). Values represent the mean probability across projection years within each period. Cell colors indicate performance level using a color scale applied independently to each column using conditional formatting, where green corresponds to the maximum value in each column (high probability), yellow to the midpoint (moderate probability), and red to the minimum (low probability).

MP	Projected year	$P(SB > SB_{hist})$	$P(Catch > Catch_{hist})$	$P(SPR > 0.3)$
MP1: seqRatio_cpue	(1-10)	0.476	0.500	0.990
MP2: seqRatio_survey	(1-10)	0.469	0.498	0.990
MP3: seqRatio_cpue_survey	(1-10)	0.468	0.495	0.994
MP4: islope_cpue	(1-10)	0.469	0.488	0.991
MP5: islope_survey	(1-10)	0.469	0.494	0.988
MP6: islope_cpue_survey	(1-10)	0.477	0.494	0.992
MP7: L_mean_LF_M(LF30)	(1-10)	0.472	0.512	0.994
MP8: L_mean_LF_M(LF50)	(1-10)	0.472	0.517	0.990
MP9: L_mean_ratio	(1-10)	0.464	0.500	0.991
MP10: seqRatio_cpue_survey_LF_M(LF30)	(1-10)	0.470	0.507	0.990
MP11: seqRatio_cpue_survey_LF_M(LF50)	(1-10)	0.468	0.499	0.992
MP12: seqRatio_cpue_survey_L_mean_ratio	(1-10)	0.476	0.509	0.992
MP13: seqRatio_survey_Lrvc_Lcom_LF_M(LF30)	(1-10)	0.470	0.493	0.994
MP14: seqRatio_survey_Lrvc_Lcom_LF_M(LF50)	(1-10)	0.469	0.491	0.994
MP15: seqRatio_survey_Lrvc_Lcom	(1-10)	0.480	0.476	0.995
MP1: seqRatio_cpue	(11-20)	0.476	0.479	0.957
MP2: seqRatio_survey	(11-20)	0.442	0.490	0.968
MP3: seqRatio_cpue_survey	(11-20)	0.457	0.477	0.966
MP4: islope_cpue	(11-20)	0.487	0.477	0.970
MP5: islope_survey	(11-20)	0.474	0.480	0.953
MP6: islope_cpue_survey	(11-20)	0.478	0.480	0.953
MP7: L_mean_LF_M(LF30)	(11-20)	0.457	0.529	0.961
MP8: L_mean_LF_M(LF50)	(11-20)	0.473	0.535	0.959
MP9: L_mean_ratio	(11-20)	0.463	0.494	0.957
MP10: seqRatio_cpue_survey_LF_M(LF30)	(11-20)	0.460	0.518	0.966
MP11: seqRatio_cpue_survey_LF_M(LF50)	(11-20)	0.455	0.498	0.965
MP12: seqRatio_cpue_survey_L_mean_ratio	(11-20)	0.467	0.479	0.969
MP13: seqRatio_survey_Lrvc_Lcom_LF_M(LF30)	(11-20)	0.457	0.498	0.973
MP14: seqRatio_survey_Lrvc_Lcom_LF_M(LF50)	(11-20)	0.475	0.484	0.974
MP15: seqRatio_survey_Lrvc_Lcom	(11-20)	0.482	0.484	0.979

Table 8. Continued.

MP	Projected year	P(SB > SB hist)	P(Catch > Catch hist)	P(SPR > 0.3)
MP1: seqRatio_cpue	(21-50)	0.472	0.442	0.875
MP2: seqRatio_survey	(21-50)	0.437	0.433	0.898
MP3: seqRatio_cpue_survey	(21-50)	0.450	0.445	0.896
MP4: islope_cpue	(21-50)	0.472	0.438	0.898
MP5: islope_survey	(21-50)	0.472	0.439	0.885
MP6: islope_cpue_survey	(21-50)	0.481	0.439	0.882
MP7: L_mean_LF_M(LF30)	(21-50)	0.439	0.512	0.857
MP8: L_mean_LF_M(LF50)	(21-50)	0.469	0.483	0.874
MP9: L_mean_ratio	(21-50)	0.453	0.439	0.873
MP10: seqRatio_cpue_survey_LF_M(LF30)	(21-50)	0.447	0.482	0.891
MP11: seqRatio_cpue_survey_LF_M(LF50)	(21-50)	0.466	0.462	0.874
MP12: seqRatio_cpue_survey_L_mean_ratio	(21-50)	0.462	0.456	0.896
MP13: seqRatio_survey_Lrvc_Lcom_LF_M(LF30)	(21-50)	0.450	0.476	0.908
MP14: seqRatio_survey_Lrvc_Lcom_LF_M(LF50)	(21-50)	0.466	0.454	0.914
MP15: seqRatio_survey_Lrvc_Lcom	(21-50)	0.479	0.429	0.922

Table 9. Average performance of MPs for OM2 across short-term (years 1–10), medium-term (years 11–20), and long-term (years 21–50) projection periods. Performance metrics include the probability that spawning biomass exceeds the final historical year level ($P(SB > SB_{hist})$), the probability that catch exceeds the final historical year level ($P(Catch > Catch_{hist})$), and the probability that the spawning potential ratio exceeds the 0.3 reference point ($P(SPR > 0.3)$). Values represent the mean probability across projection years within each period. Cell colors indicate performance level using a color scale applied independently to each column using conditional formatting, where green corresponds to the maximum value in each column (high probability), yellow to the midpoint (moderate probability), and red to the minimum (low probability).

MP	Projected year	$P(SB > SB_{hist})$	$P(Catch > Catch_{hist})$	$P(SPR > 0.3)$
MP1: seqRatio_cpue	(1-10)	0.425	0.474	0.976
MP2: seqRatio_survey	(1-10)	0.424	0.485	0.979
MP3: seqRatio_cpue_survey	(1-10)	0.413	0.500	0.974
MP4: islope_cpue	(1-10)	0.417	0.507	0.971
MP5: islope_survey	(1-10)	0.421	0.484	0.977
MP6: islope_cpue_survey	(1-10)	0.403	0.504	0.978
MP7: L_mean_LF_M(LF30)	(1-10)	0.408	0.518	0.972
MP8: L_mean_LF_M(LF50)	(1-10)	0.421	0.492	0.973
MP9: L_mean_ratio	(1-10)	0.420	0.484	0.977
MP10: seqRatio_cpue_survey_LF_M(LF30)	(1-10)	0.401	0.507	0.976
MP11: seqRatio_cpue_survey_LF_M(LF50)	(1-10)	0.411	0.501	0.975
MP12: seqRatio_cpue_survey_L_mean_ratio	(1-10)	0.418	0.477	0.978
MP13: seqRatio_survey_Lrvc_Lcom_LF_M(LF30)	(1-10)	0.415	0.488	0.981
MP14: seqRatio_survey_Lrvc_Lcom_LF_M(LF50)	(1-10)	0.414	0.470	0.975
MP15: seqRatio_survey_Lrvc_Lcom	(1-10)	0.415	0.485	0.976
MP1: seqRatio_cpue	(11-20)	0.405	0.453	0.928
MP2: seqRatio_survey	(11-20)	0.394	0.460	0.928
MP3: seqRatio_cpue_survey	(11-20)	0.382	0.475	0.921
MP4: islope_cpue	(11-20)	0.373	0.494	0.918
MP5: islope_survey	(11-20)	0.385	0.462	0.917
MP6: islope_cpue_survey	(11-20)	0.364	0.489	0.916
MP7: L_mean_LF_M(LF30)	(11-20)	0.386	0.522	0.905
MP8: L_mean_LF_M(LF50)	(11-20)	0.395	0.486	0.907
MP9: L_mean_ratio	(11-20)	0.404	0.457	0.908
MP10: seqRatio_cpue_survey_LF_M(LF30)	(11-20)	0.385	0.502	0.921
MP11: seqRatio_cpue_survey_LF_M(LF50)	(11-20)	0.378	0.482	0.918
MP12: seqRatio_cpue_survey_L_mean_ratio	(11-20)	0.398	0.454	0.936
MP13: seqRatio_survey_Lrvc_Lcom_LF_M(LF30)	(11-20)	0.373	0.471	0.934
MP14: seqRatio_survey_Lrvc_Lcom_LF_M(LF50)	(11-20)	0.382	0.445	0.930
MP15: seqRatio_survey_Lrvc_Lcom	(11-20)	0.389	0.439	0.927

Table 9. Continued.

MP	Projected year	P(SB > SB_hist)	P(Catch > Catch_hist)	P(SPR > 0.3)
MP1: seqRatio_cpue	(21-50)	0.389	0.382	0.800
MP2: seqRatio_survey	(21-50)	0.359	0.388	0.805
MP3: seqRatio_cpue_survey	(21-50)	0.357	0.412	0.792
MP4: islope_cpue	(21-50)	0.361	0.387	0.790
MP5: islope_survey	(21-50)	0.380	0.376	0.789
MP6: islope_cpue_survey	(21-50)	0.352	0.400	0.774
MP7: L_mean_LF_M(LF30)	(21-50)	0.348	0.451	0.757
MP8: L_mean_LF_M(LF50)	(21-50)	0.360	0.410	0.773
MP9: L_mean_ratio	(21-50)	0.370	0.377	0.765
MP10: seqRatio_cpue_survey_LF_M(LF30)	(21-50)	0.354	0.418	0.787
MP11: seqRatio_cpue_survey_LF_M(LF50)	(21-50)	0.353	0.388	0.792
MP12: seqRatio_cpue_survey_L_mean_ratio	(21-50)	0.392	0.376	0.804
MP13: seqRatio_survey_Lrvc_Lcom_LF_M(LF30)	(21-50)	0.357	0.414	0.822
MP14: seqRatio_survey_Lrvc_Lcom_LF_M(LF50)	(21-50)	0.371	0.377	0.818
MP15: seqRatio_survey_Lrvc_Lcom	(21-50)	0.373	0.383	0.833

Table 10. Average performance of MPs for *OM2-R1- Index (recreational & survey)* *hyperstability* across short-term (years 1–10), medium-term (years 11–20), and long-term (years 21–50) projection periods. Performance metrics include the probability that spawning biomass exceeds the final historical year level ($P(SB > SB_{hist})$), the probability that catch exceeds the final historical year level ($P(Catch > Catch_{hist})$), and the probability that the spawning potential ratio exceeds the 0.3 reference point ($P(SPR > 0.3)$). Values represent the mean probability across projection years within each period. Cell colors indicate performance level using a color scale applied independently to each column using conditional formatting, where green corresponds to the maximum value in each column (high probability), yellow to the midpoint (moderate probability), and red to the minimum (low probability).

MP	Projected year	P(SB > SB _{hist})	P(Catch > Catch _{hist})	P(SPR > 0.3)
MP1: seqRatio_cpue	(1-10)	0.419	0.483	0.972
MP2: seqRatio_survey	(1-10)	0.423	0.470	0.976
MP3: seqRatio_cpue_survey	(1-10)	0.419	0.472	0.979
MP4: islope_cpue	(1-10)	0.424	0.462	0.974
MP5: islope_survey	(1-10)	0.413	0.498	0.975
MP6: islope_cpue_survey	(1-10)	0.412	0.483	0.977
MP7: L_mean_LF_M(LF30)	(1-10)	0.406	0.496	0.975
MP8: L_mean_LF_M(LF50)	(1-10)	0.415	0.485	0.977
MP9: L_mean_ratio	(1-10)	0.407	0.534	0.971
MP10: seqRatio_cpue_survey_LF_M(LF30)	(1-10)	0.423	0.468	0.979
MP11: seqRatio_cpue_survey_LF_M(LF50)	(1-10)	0.400	0.500	0.974
MP12: seqRatio_cpue_survey_L_mean_ratio	(1-10)	0.424	0.479	0.976
MP13: seqRatio_survey_Lrvc_Lcom_LF_M(LF30)	(1-10)	0.409	0.492	0.977
MP14: seqRatio_survey_Lrvc_Lcom_LF_M(LF50)	(1-10)	0.420	0.466	0.976
MP15: seqRatio_survey_Lrvc_Lcom	(1-10)	0.407	0.479	0.978
MP1: seqRatio_cpue	(11-20)	0.396	0.451	0.910
MP2: seqRatio_survey	(11-20)	0.406	0.470	0.917
MP3: seqRatio_cpue_survey	(11-20)	0.378	0.460	0.928
MP4: islope_cpue	(11-20)	0.401	0.454	0.905
MP5: islope_survey	(11-20)	0.376	0.464	0.925
MP6: islope_cpue_survey	(11-20)	0.391	0.471	0.909
MP7: L_mean_LF_M(LF30)	(11-20)	0.391	0.493	0.913
MP8: L_mean_LF_M(LF50)	(11-20)	0.415	0.482	0.921
MP9: L_mean_ratio	(11-20)	0.371	0.533	0.897
MP10: seqRatio_cpue_survey_LF_M(LF30)	(11-20)	0.384	0.471	0.925
MP11: seqRatio_cpue_survey_LF_M(LF50)	(11-20)	0.376	0.482	0.912
MP12: seqRatio_cpue_survey_L_mean_ratio	(11-20)	0.401	0.449	0.912
MP13: seqRatio_survey_Lrvc_Lcom_LF_M(LF30)	(11-20)	0.382	0.467	0.912
MP14: seqRatio_survey_Lrvc_Lcom_LF_M(LF50)	(11-20)	0.384	0.466	0.931
MP15: seqRatio_survey_Lrvc_Lcom	(11-20)	0.393	0.463	0.928

Table 10. Continued.

MP	Projected year	P(SB > SB_hist)	P(Catch > Catch_hist)	P(SPR > 0.3)
MP1: seqRatio_cpue	(21-50)	0.376	0.376	0.796
MP2: seqRatio_survey	(21-50)	0.368	0.366	0.779
MP3: seqRatio_cpue_survey	(21-50)	0.355	0.368	0.785
MP4: islope_cpue	(21-50)	0.388	0.366	0.790
MP5: islope_survey	(21-50)	0.367	0.371	0.803
MP6: islope_cpue_survey	(21-50)	0.386	0.370	0.783
MP7: L_mean_LF_M(LF30)	(21-50)	0.350	0.416	0.748
MP8: L_mean_LF_M(LF50)	(21-50)	0.383	0.412	0.799
MP9: L_mean_ratio	(21-50)	0.360	0.411	0.748
MP10: seqRatio_cpue_survey_LF_M(LF30)	(21-50)	0.360	0.386	0.790
MP11: seqRatio_cpue_survey_LF_M(LF50)	(21-50)	0.345	0.386	0.773
MP12: seqRatio_cpue_survey_L_mean_ratio	(21-50)	0.376	0.377	0.791
MP13: seqRatio_survey_Lrvc_Lcom_LF_M(LF30)	(21-50)	0.352	0.381	0.775
MP14: seqRatio_survey_Lrvc_Lcom_LF_M(LF50)	(21-50)	0.386	0.391	0.793
MP15: seqRatio_survey_Lrvc_Lcom	(21-50)	0.372	0.384	0.795

Figures

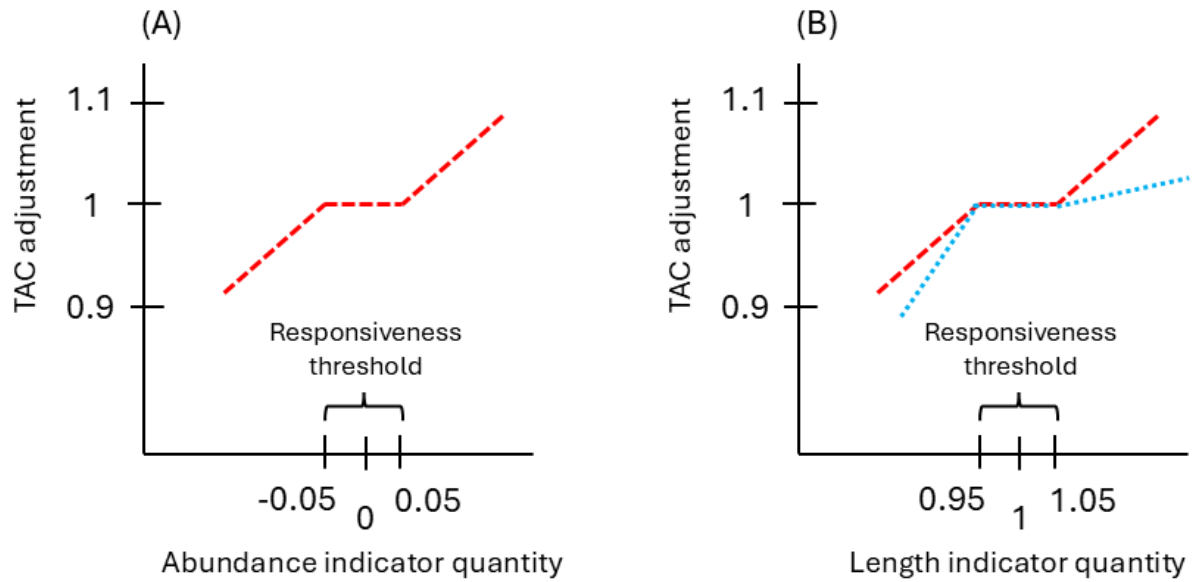


Figure 1. Demonstration of TAC adjustment process when using (A) index-based indicators, and (B) length-based indicators. Responsiveness threshold defines the range of indicator quantities over which TAC adjustments are avoided (i.e., TAC adjustment multiplier equals 1). In (B), solid line and dashed line show two variants of penalty and reward multipliers, producing different degrees of responsiveness.

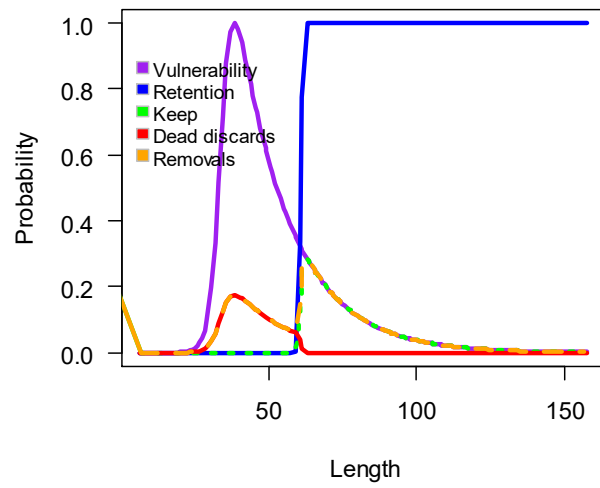


Figure 2. Recreational fishery characteristics.

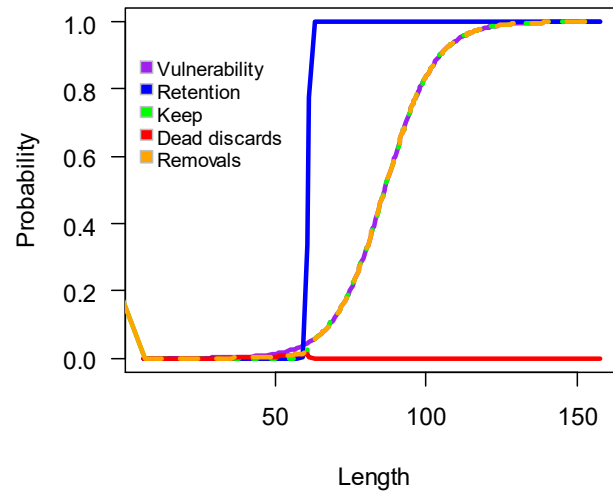


Figure 3. Commercial fishery characteristics.

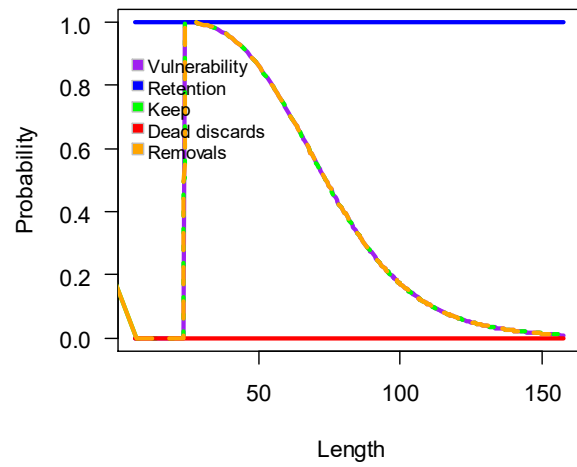


Figure 4. Survey retention characteristics.

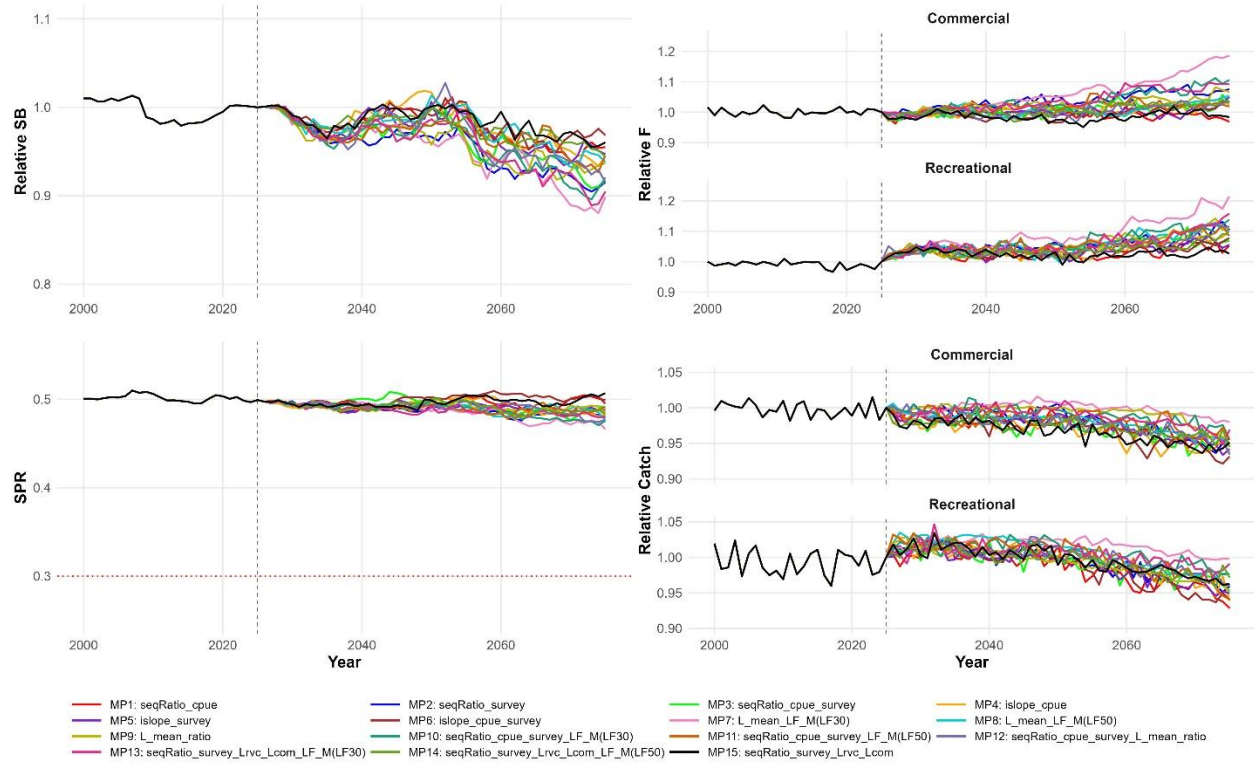


Figure 5. Median trajectories of spawning biomass, SPR, fishing mortality, and catch under all management procedures for OM1. Metrics are relative to the final historical year (except for SPR). The vertical dashed line marks the start of projections; the horizontal dotted line shows $SPR = 0.30$. Lines represent the median of MPs 1-15 across 300 simulation iterations.

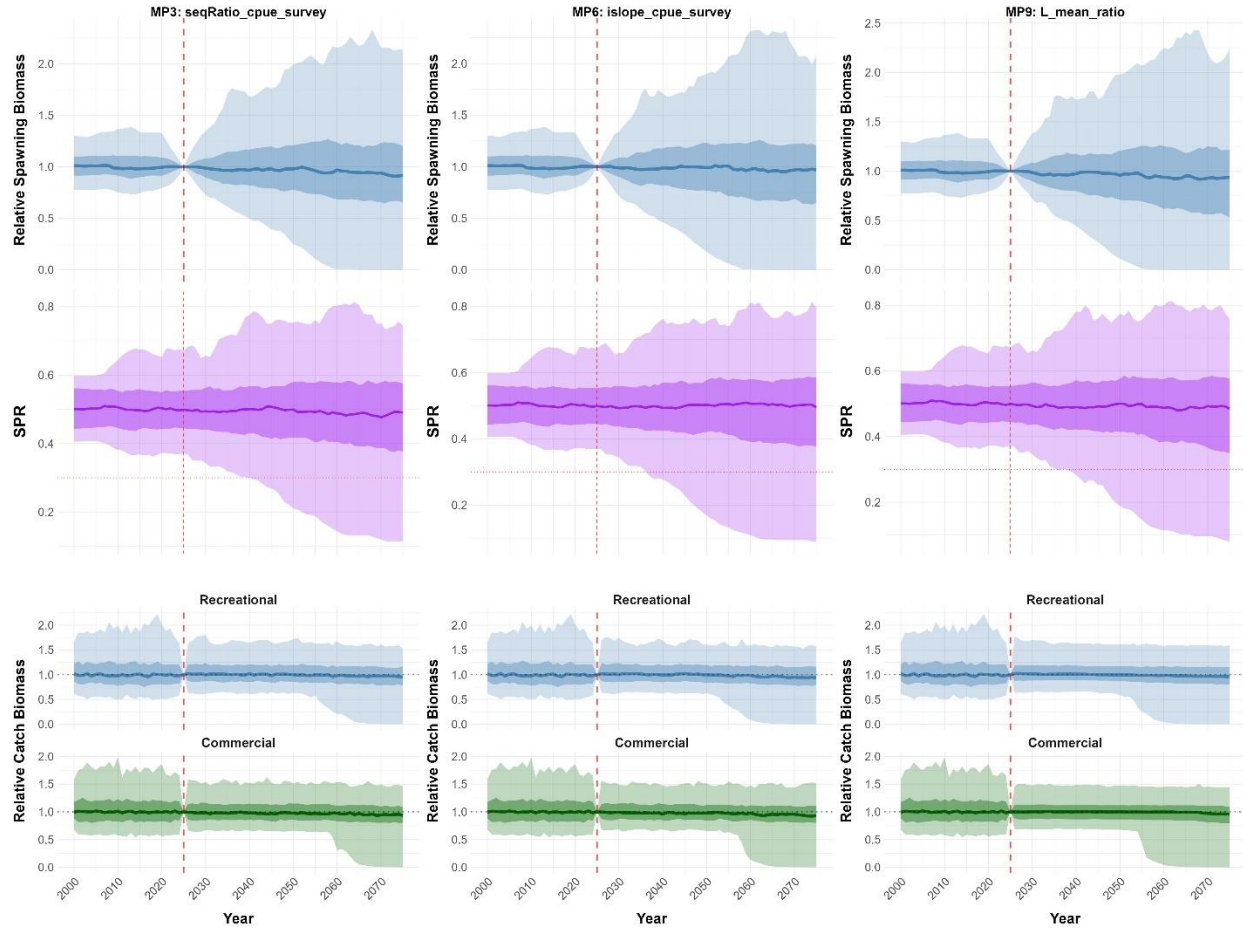


Figure 6. Comparison of three stability-based management procedures under OM 1. Panels show relative spawning biomass (top row), SPR (middle row), and relative catch biomass by fleet (bottom rows) for MP3 (seqRatio_cpue_survey), MP6 (islope_cpue_survey), and MP9 (L_mean_ratio). Solid lines represent median trajectories; dark shaded bands show the 50% confidence interval (25th–75th percentiles); light shaded bands show the 95% confidence interval (2.5th–97.5th percentiles). All relative values are normalized to the final historical year (except SPR). Vertical dashed lines indicate the start of projections; horizontal dotted lines in SPR panels show SPR 30%. Results are based on 300 simulation iterations.

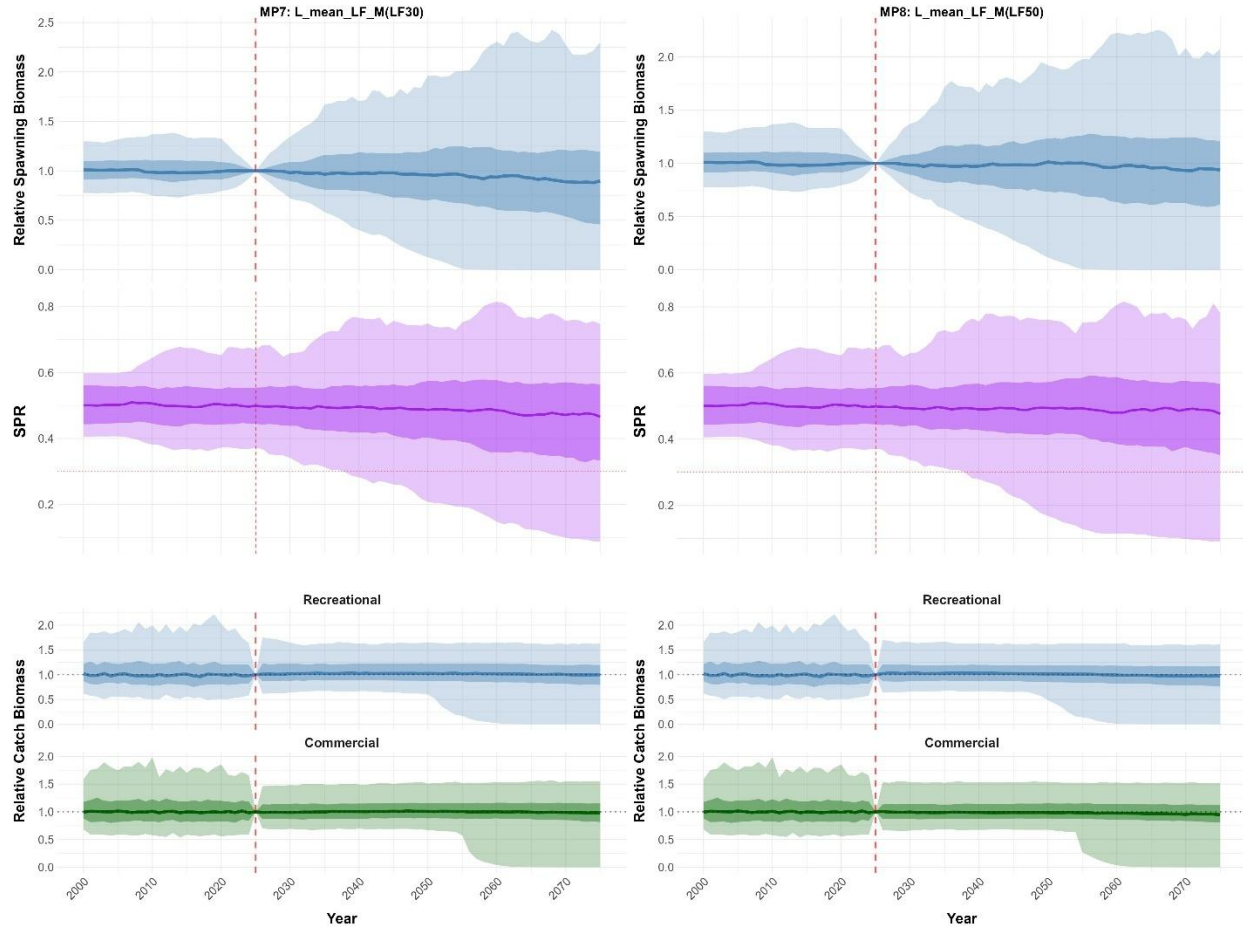


Figure 7. Target-seeking MPs under OM1. Relative SB, SPR, and catch by fleet for MP7: L_mean_LF_M(LF30) and MP8: L_mean_LF_M(LF50). Medians (solid lines), dark shaded bands show the 50% confidence interval (25th–75th percentiles); light shaded bands show the 95% confidence interval (2.5th–97.5th percentiles). Dotted lines: SPR = 0.30 threshold. Results are based on 300 simulation iterations.

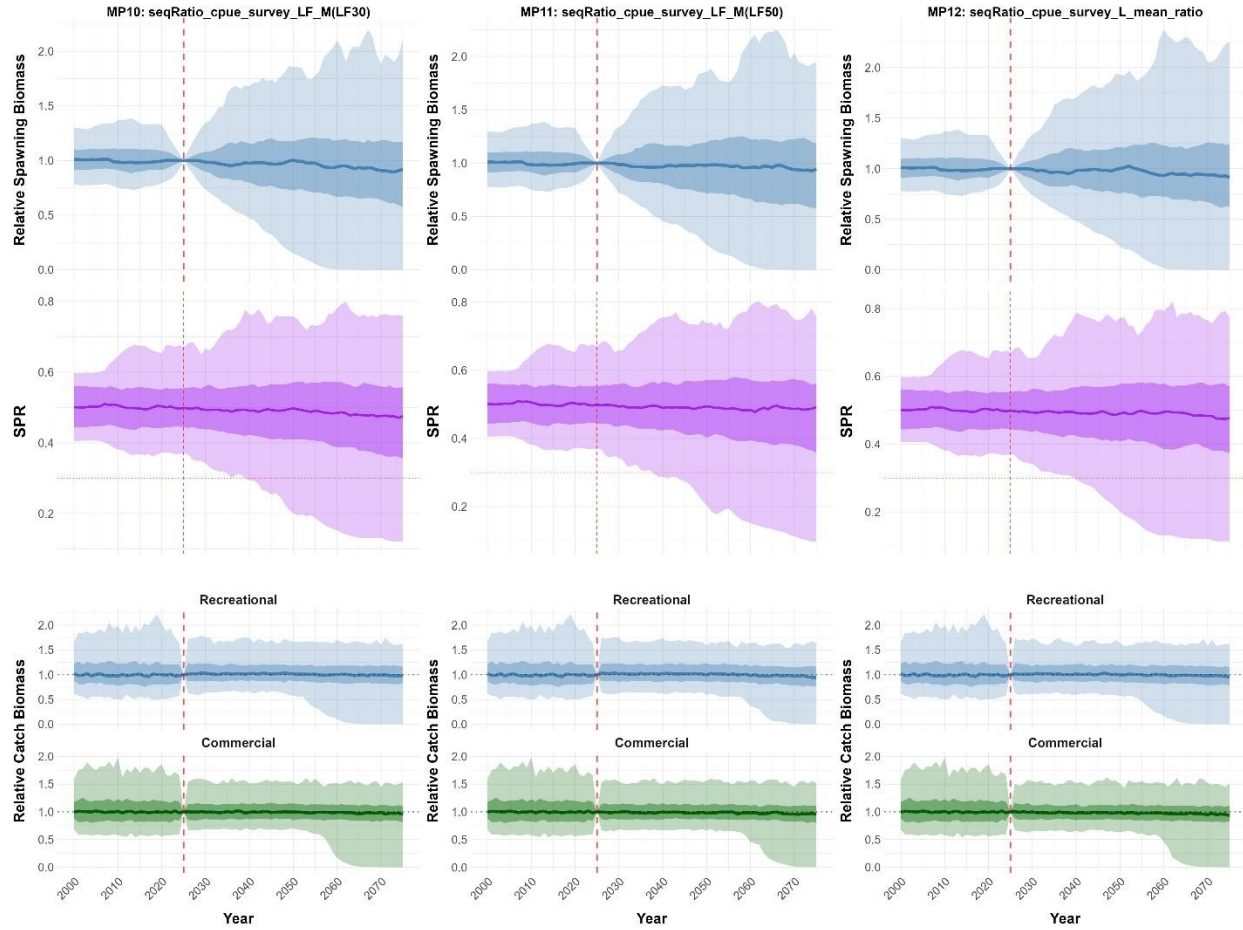


Figure 8. Multi-indicator MPs under OM1. Relative SB, SPR, and catch by fleet for MP10: seqRatio_cpue_survey_LF_M(LF30), MP11: seqRatio_cpue_survey_LF_M(LF50) and MP12: seqRatio_cpue_survey_L_mean_ratio. Medians (solid lines), dark shaded bands show the 50% confidence interval (25th–75th percentiles); light shaded bands show the 95% confidence interval (2.5th–97.5th percentiles). Dotted lines: SPR = 0.30 threshold. Results are based on 300 simulation iterations.

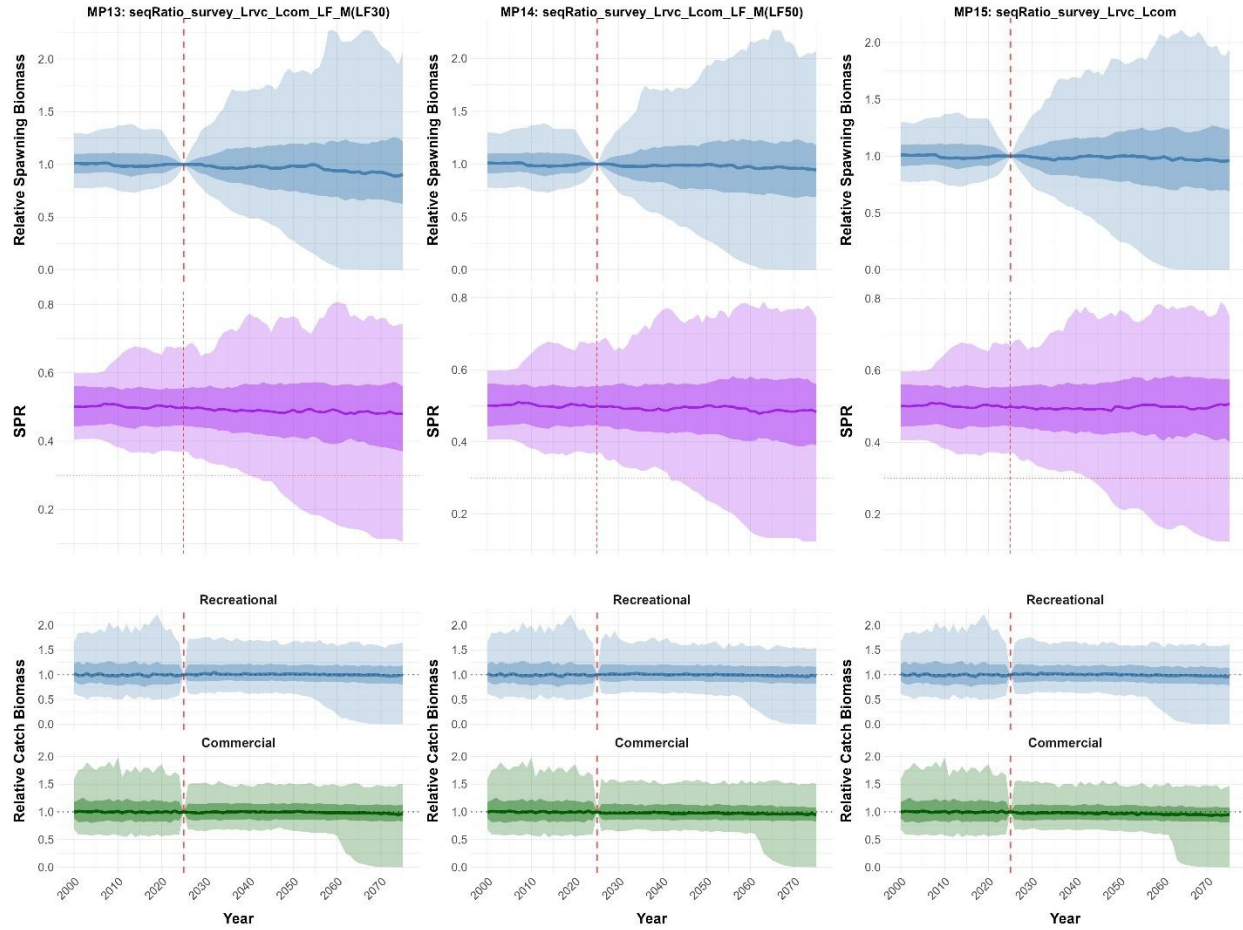


Figure 9. Multi-indicator MPs under OM1. Relative SB, SPR, and catch by fleet for MP13: seqRatio_survey_Lrvc_Lcom_LF_M(LF30), MP14: seqRatio_survey_Lrvc_Lcom_LF_M(LF50) and MP15: seqRatio_survey_Lrvc_Lcom. Medians (solid lines), dark shaded bands show the 50% confidence interval (25th–75th percentiles); light shaded bands show the 95% confidence interval (2.5th–97.5th percentiles). Dotted lines: SPR = 0.30 threshold. Results are based on 300 simulation iterations.

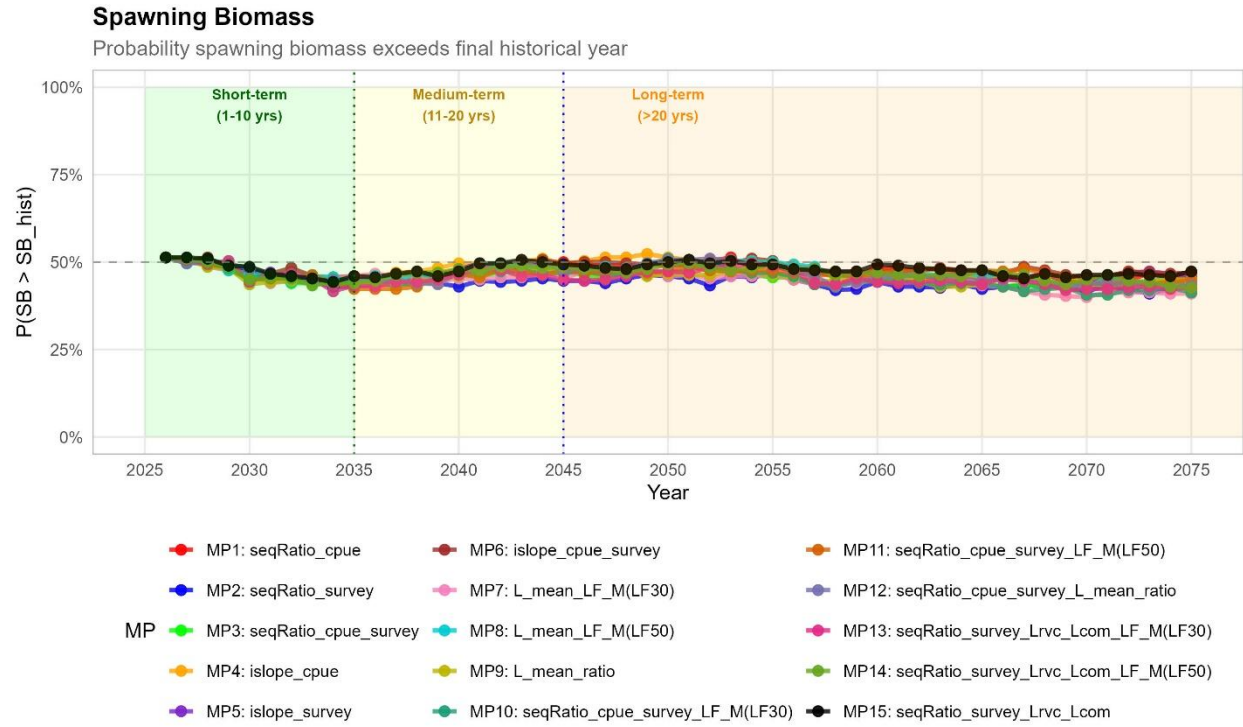


Figure 10. Performance metrics for spawning biomass (Probability that spawning biomass exceeds final historical year) across all management procedures under OM 1. Panels show probability of achieving performance targets over short-term (1-10 yrs, green), medium-term (11-20 yrs, yellow), and long-term (>20 yrs, orange) projection periods.

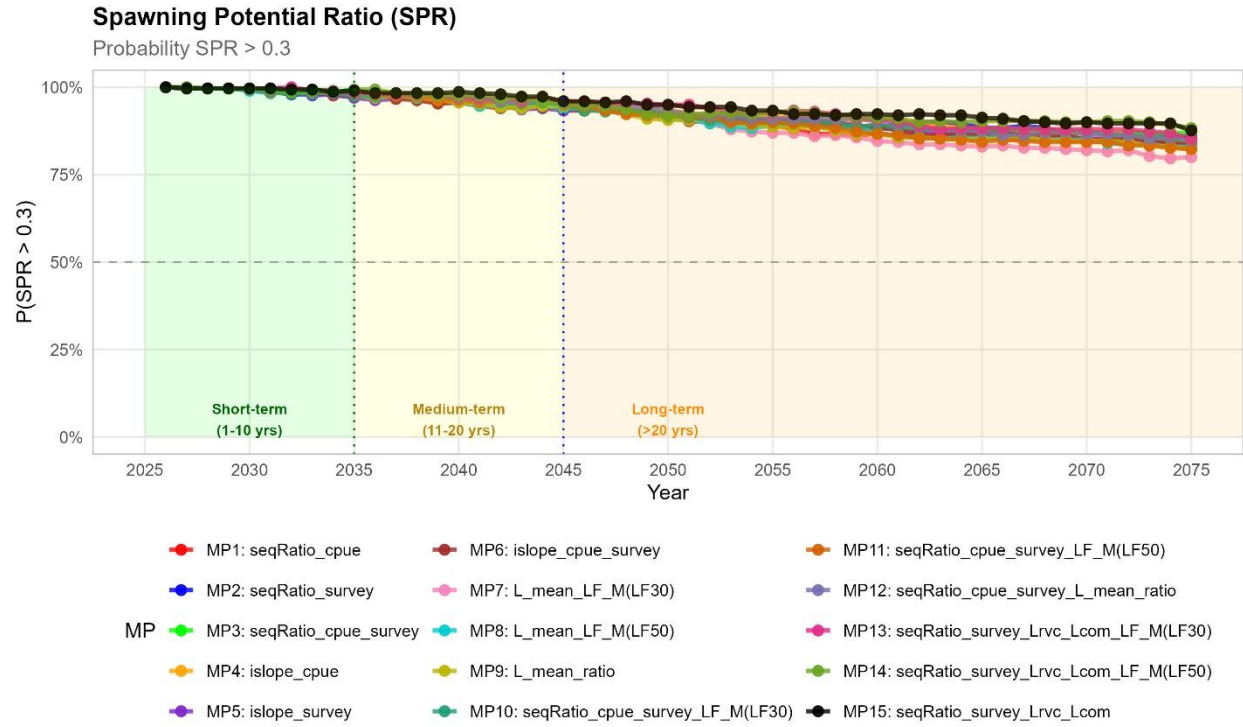


Figure 11. Performance metrics for SPR (probability SPR > 0.3) across all management procedures under OM 1. Panels show probability of achieving performance targets over short-term (1-10 yrs, green), medium-term (11-20 yrs, yellow), and long-term (>20 yrs, orange) projection periods.

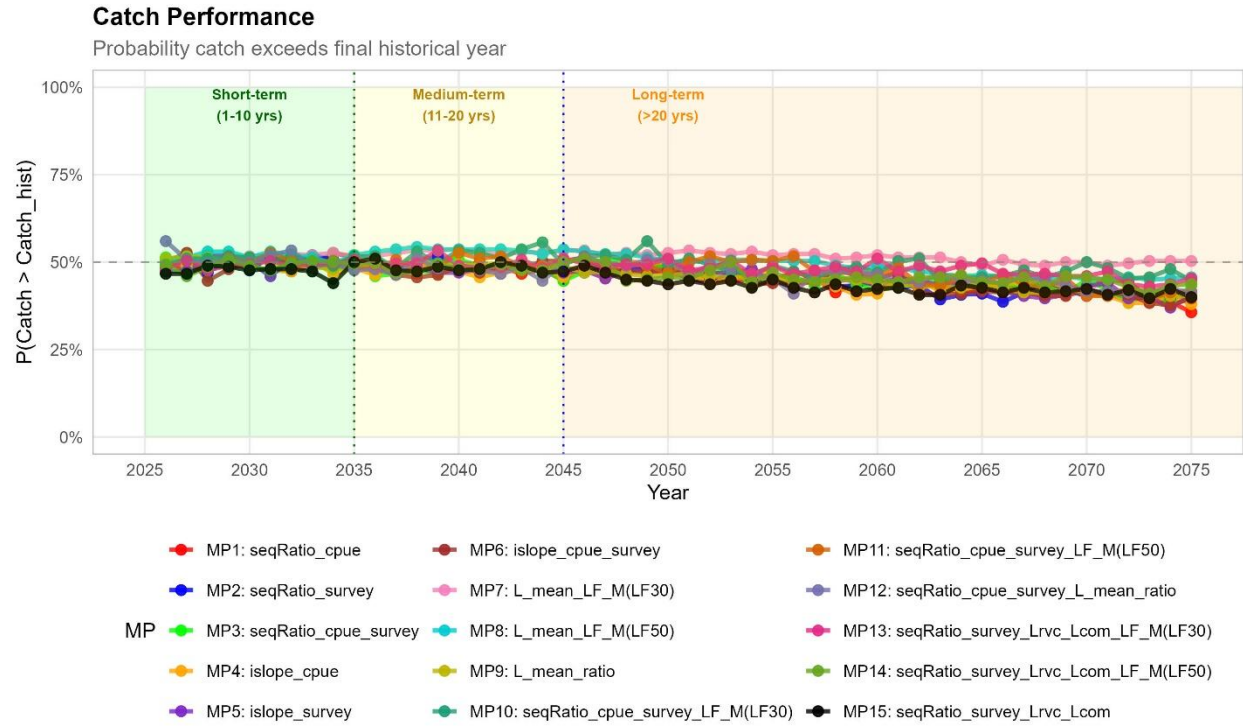


Figure 12. Performance metrics for catch performance (probability catch exceeds final historical year) across all management procedures under OM 1. Panels show probability of achieving performance targets over short-term (1-10 yrs, green), medium-term (11-20 yrs, yellow), and long-term (>20 yrs, orange) projection periods. Horizontal dashed line shows 50% probability threshold.

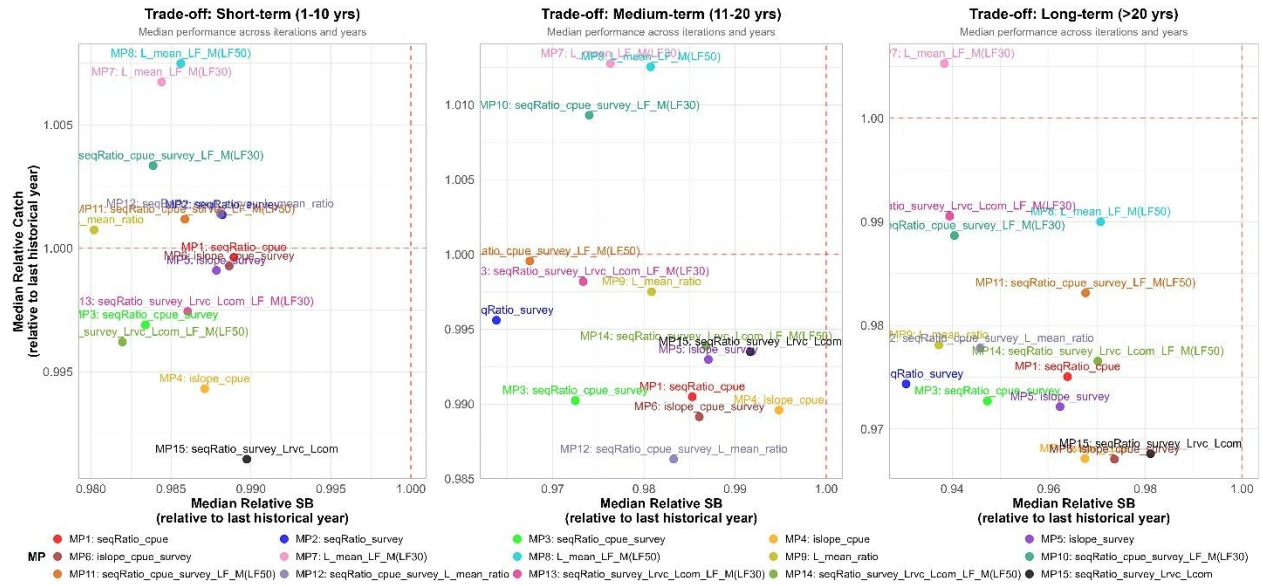


Figure 13. Trade-off between median relative catch and median relative spawning biomass across all management procedures under OM 1 for the short, medium, and long term. Each color represents an individual MP.

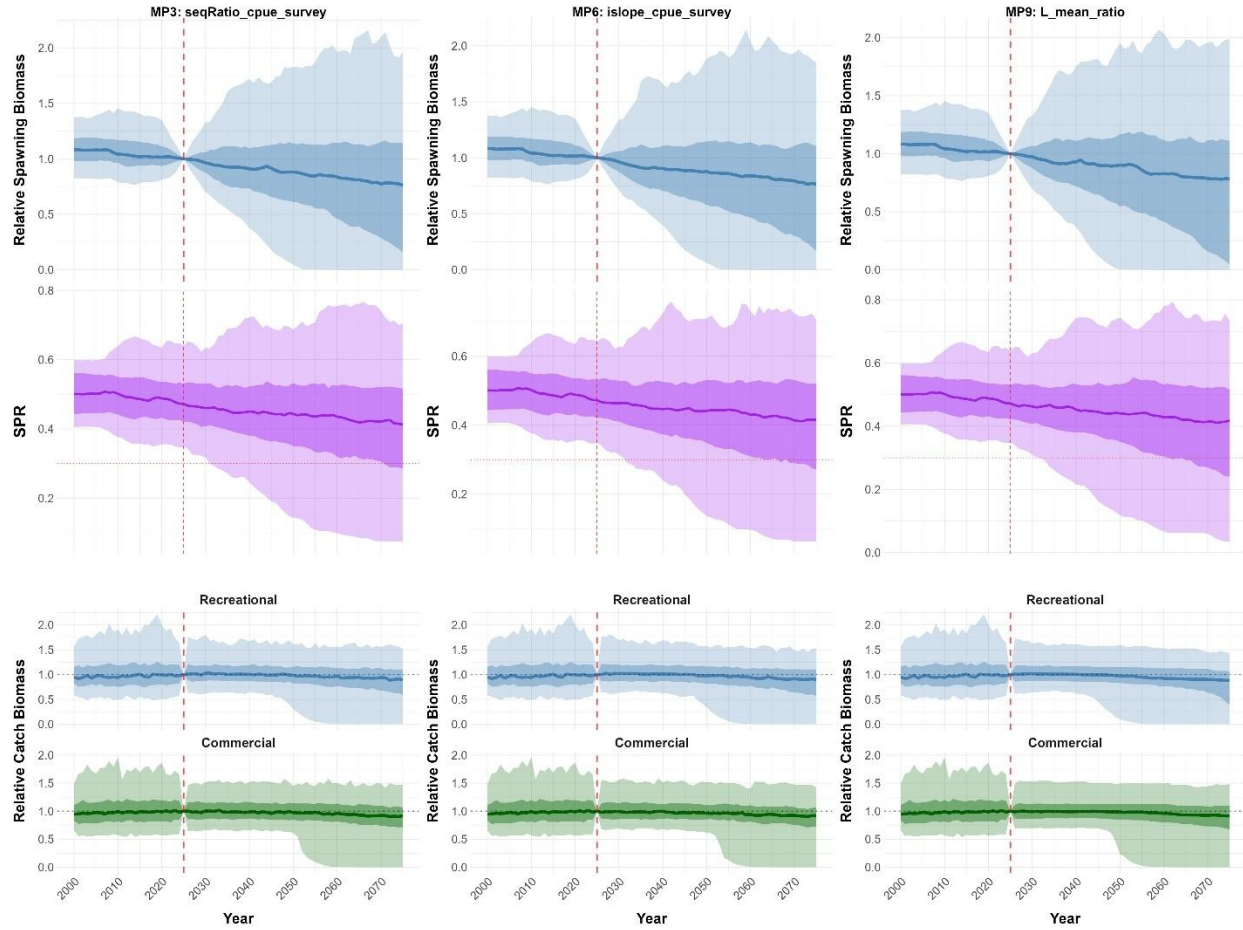


Figure 14. Comparison of three stability-based management procedures under OM 2. Panels show relative spawning biomass (top row), SPR (middle row), and relative catch biomass by fleet (bottom rows) for MP3 (seqRatio_cpue_survey), MP6 (islope_cpue_survey), and MP9 (L_mean_ratio). Solid lines represent median trajectories; dark shaded bands show the 50% confidence interval (25th–75th percentiles); light shaded bands show the 95% confidence interval (2.5th–97.5th percentiles). All relative values are normalized to the final historical year (except SPR). Vertical dashed lines indicate the start of projections; horizontal dotted lines in SPR panels show SPR 30%. Results are based on 300 simulation iterations.

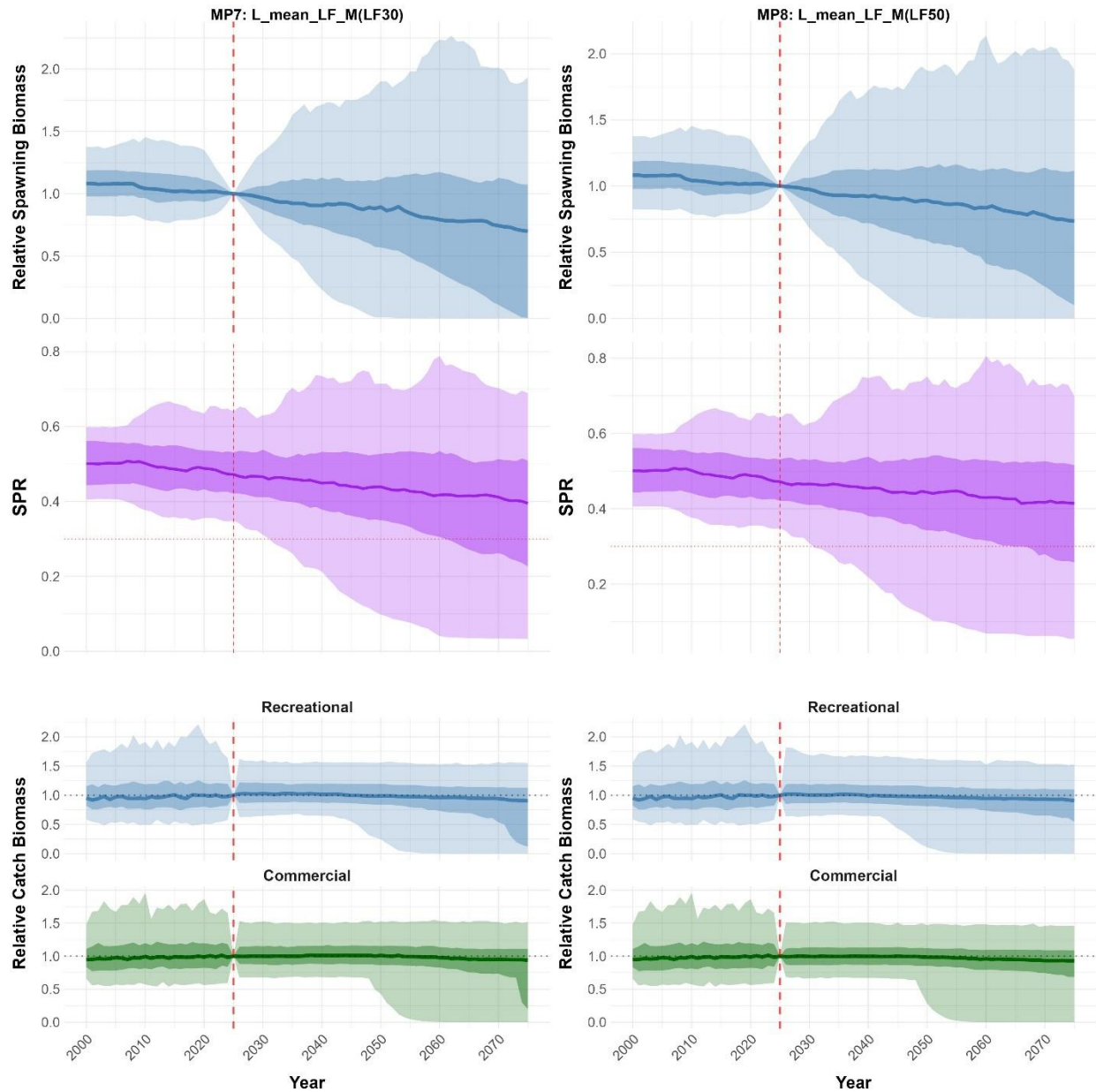


Figure 15. Target-seeking MPs under OM2. Relative SB, SPR, and catch by fleet for MP7: MP7: L_mean_LF_M(LF30) and MP8: MP7: L_mean_LF_M(LF50). Medians (solid lines), dark shaded bands show the 50% confidence interval (25th–75th percentiles); light shaded bands show the 95% confidence interval (2.5th–97.5th percentiles). Dotted lines: SPR = 0.30 threshold. Results are based on 300 simulation iterations.

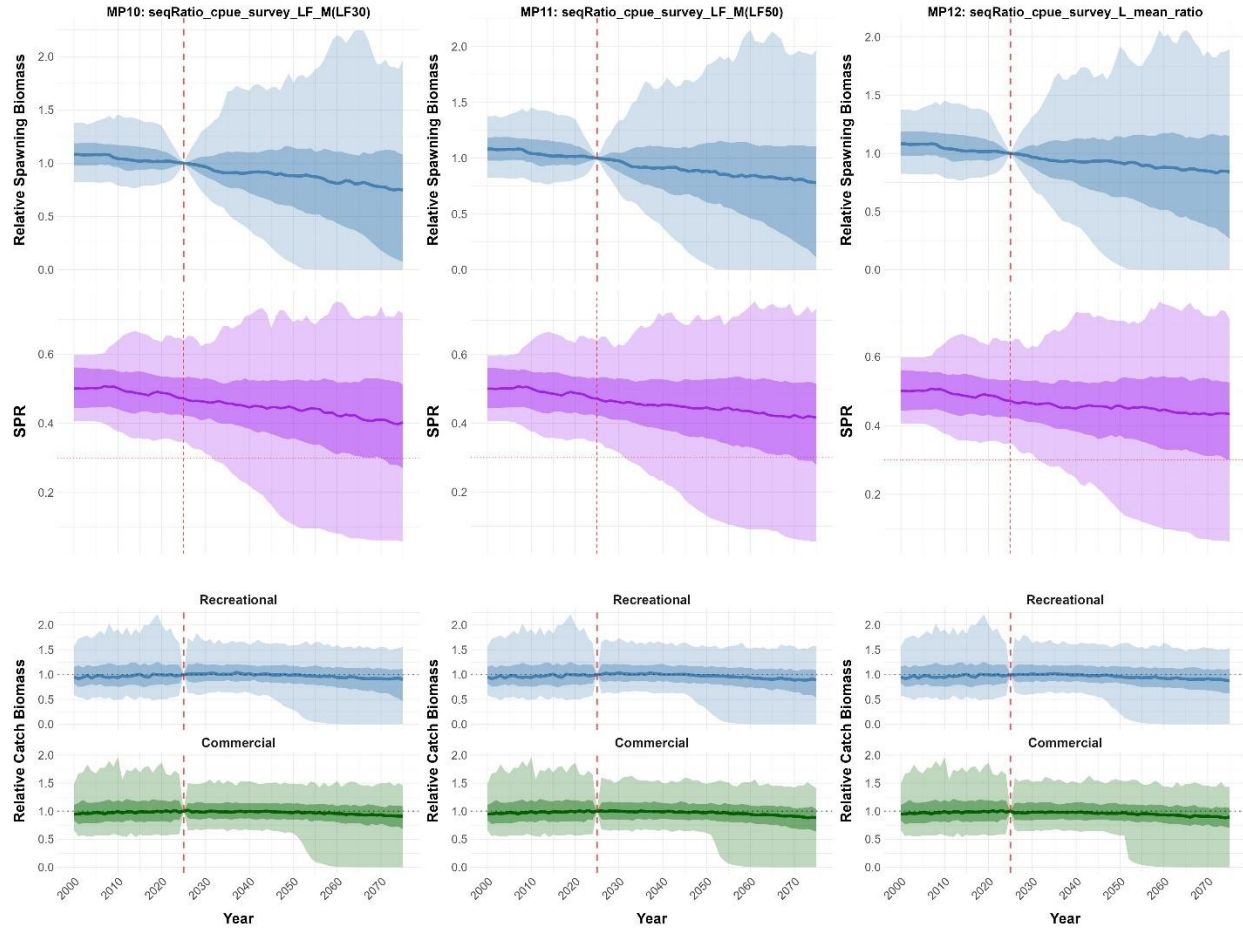


Figure 16. Multi-indicator MPs under OM2. Relative SB, SPR, and catch by fleet for MP10: seqRatio_cpue_survey_LF_M(LF30), MP11: seqRatio_cpue_survey_LF_M(LF50) and MP12: seqRatio_cpue_survey_L_mean_ratio. Medians (solid lines), dark shaded bands show the 50% confidence interval (25th–75th percentiles); light shaded bands show the 95% confidence interval (2.5th–97.5th percentiles). Dotted lines: SPR = 0.30 threshold. Results are based on 300 simulation iterations.

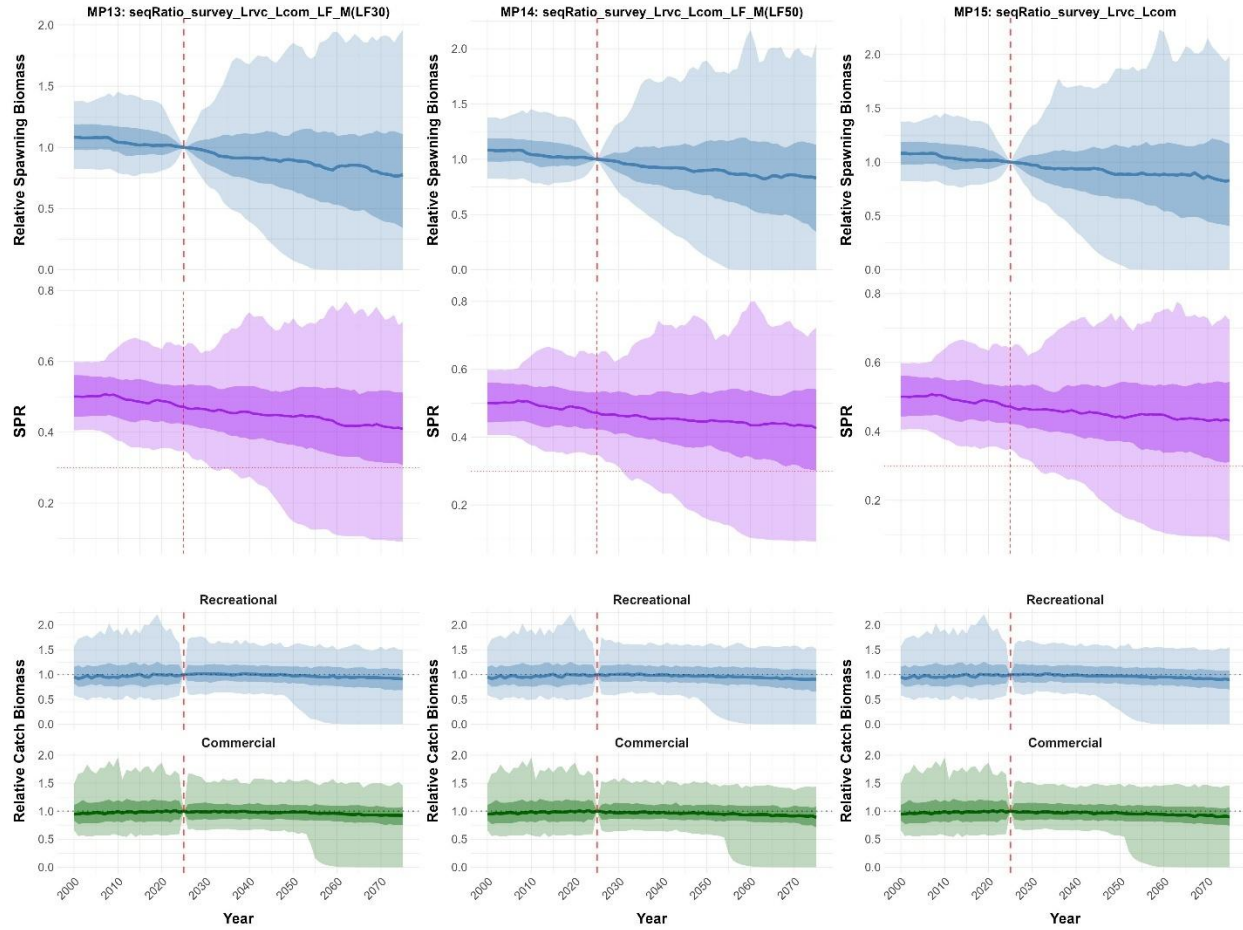


Figure 17. Multi-indicator MPs under OM2. Relative SB, SPR, and catch by fleet for MP13: seqRatio_survey_Lrvc_Lcom_LF_M(LF30), MP14: seqRatio_survey_Lrvc_Lcom_LF_M(LF50) and MP15: seqRatio_survey_Lrvc_Lcom. Medians (solid lines), dark shaded bands show the 50% confidence interval (25th–75th percentiles); light shaded bands show the 95% confidence interval (2.5th–97.5th percentiles). Dotted lines: SPR = 0.30 threshold. Results are based on 300 simulation iterations.

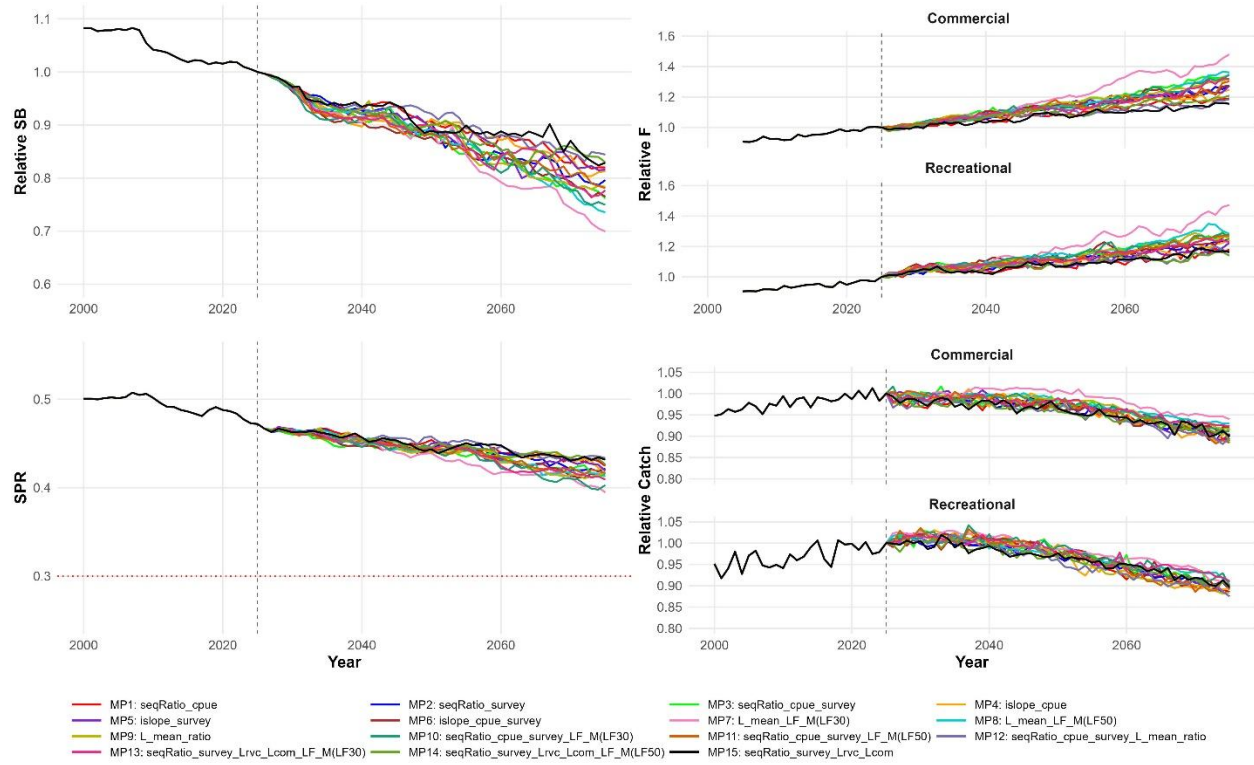


Figure 18. Median trajectories of spawning biomass, SPR, fishing mortality, and catch under all management procedures for OM2. Metrics are relative to the final historical year (except for SPR). The vertical dashed line marks the start of projections; the horizontal dotted line shows $SPR = 0.30$. Lines represent the median of MPs 1-15 across 300 simulation iterations.

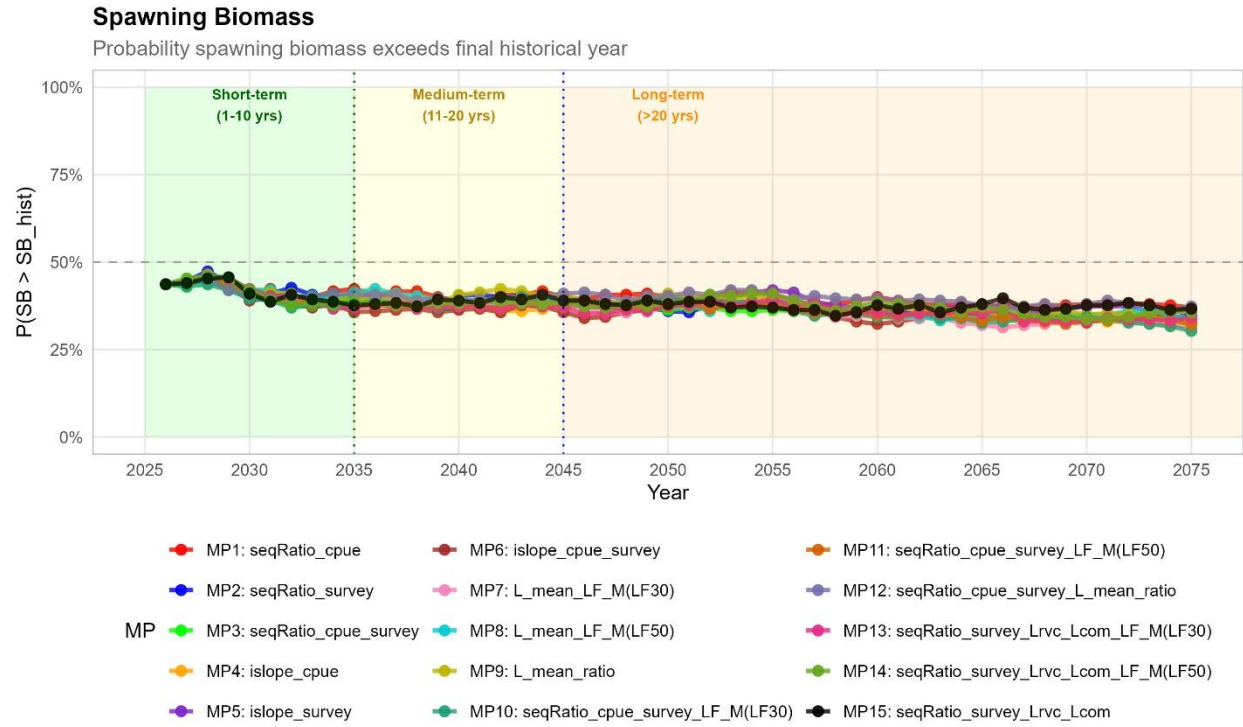


Figure 19. Performance metrics for spawning biomass (Probability that spawning biomass exceeds final historical year) across all management procedures under OM2. Panels show probability of achieving performance targets over short-term (1-10 yrs, green), medium-term (11-20 yrs, yellow), and long-term (>20 yrs, orange) projection periods.

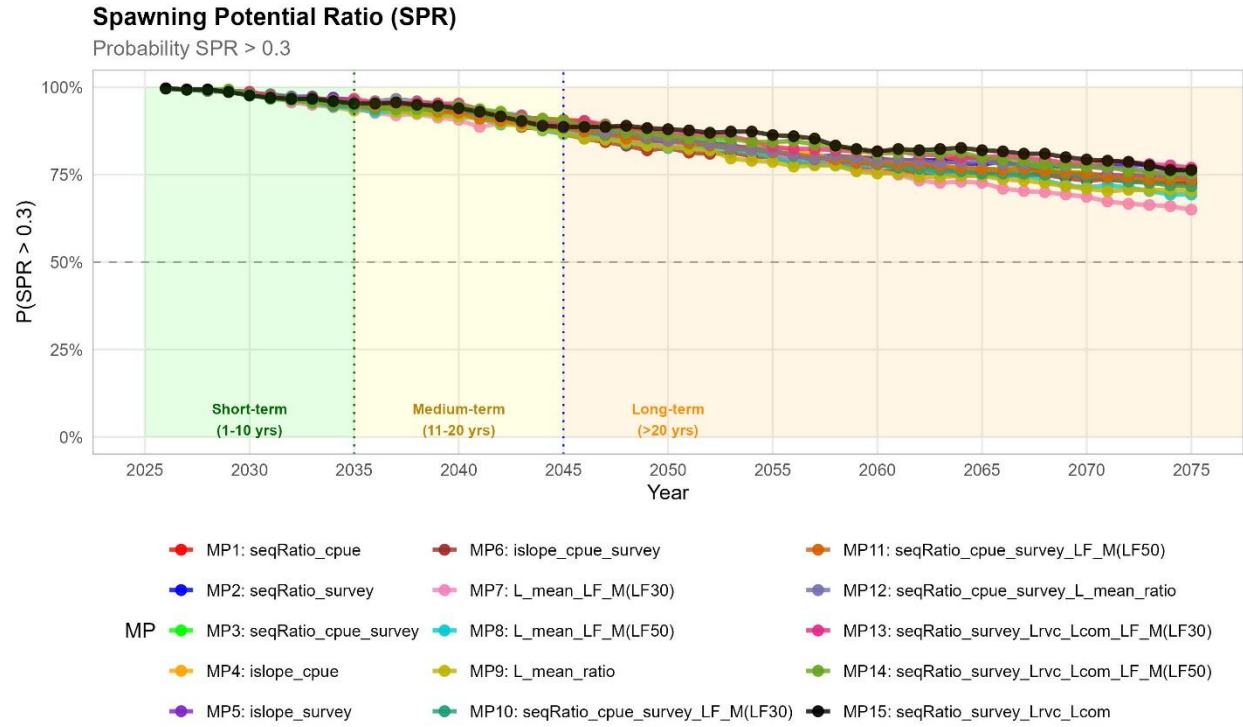


Figure 20. Performance metrics for SPR (probability SPR > 0.3) across all management procedures under OM2. Panels show probability of achieving performance targets over short-term (1-10 yrs, green), medium-term (11-20 yrs, yellow), and long-term (>20 yrs, orange) projection periods.

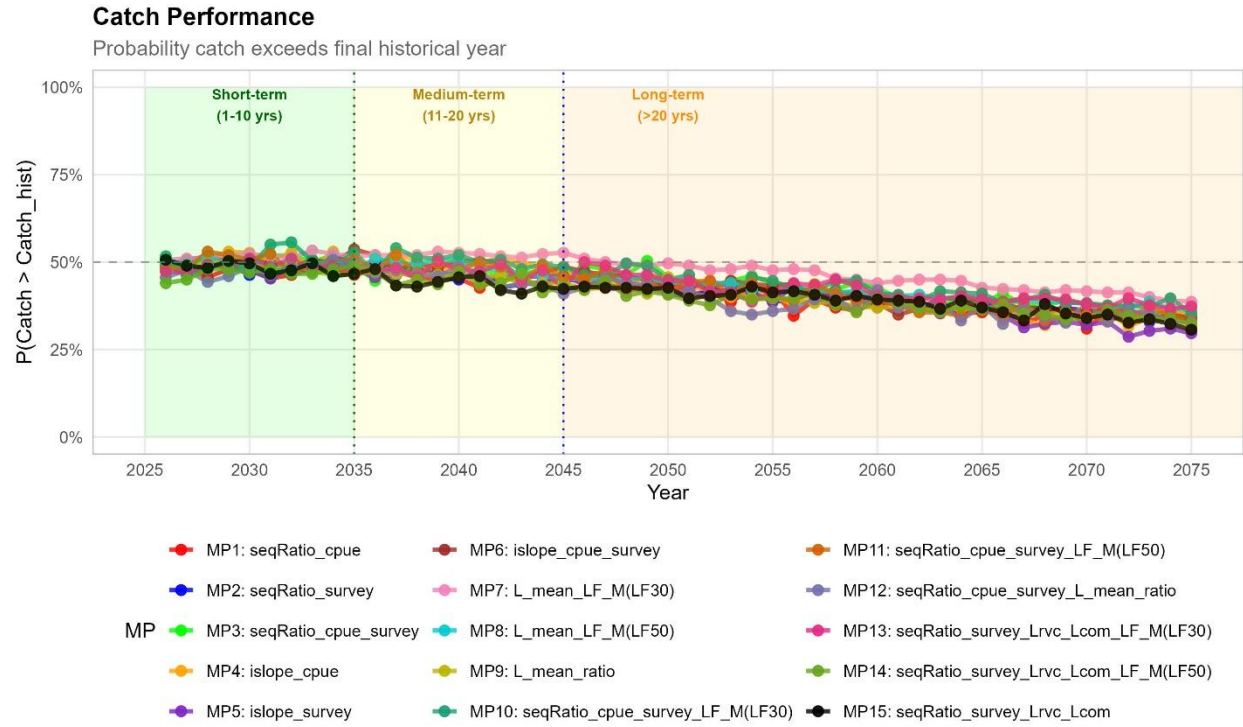


Figure 21. Performance metrics for catch performance (probability catch exceeds final historical year) across all management procedures under OM2. Panels show probability of achieving performance targets over short-term (1-10 yrs, green), medium-term (11-20 yrs, yellow), and long-term (>20 yrs, orange) projection periods. Horizontal dashed line shows 50% probability threshold.

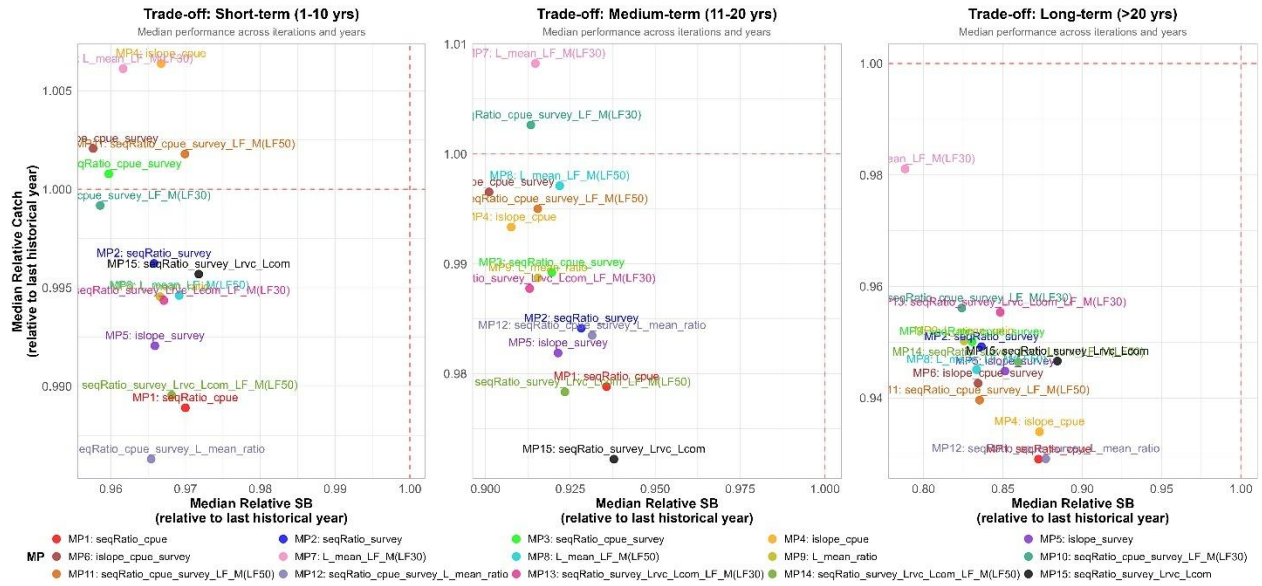


Figure 22. Trade-off between median relative catch and median relative spawning biomass across all management procedures under OM2 for the short, medium, and long term. Each color represents an individual MP.

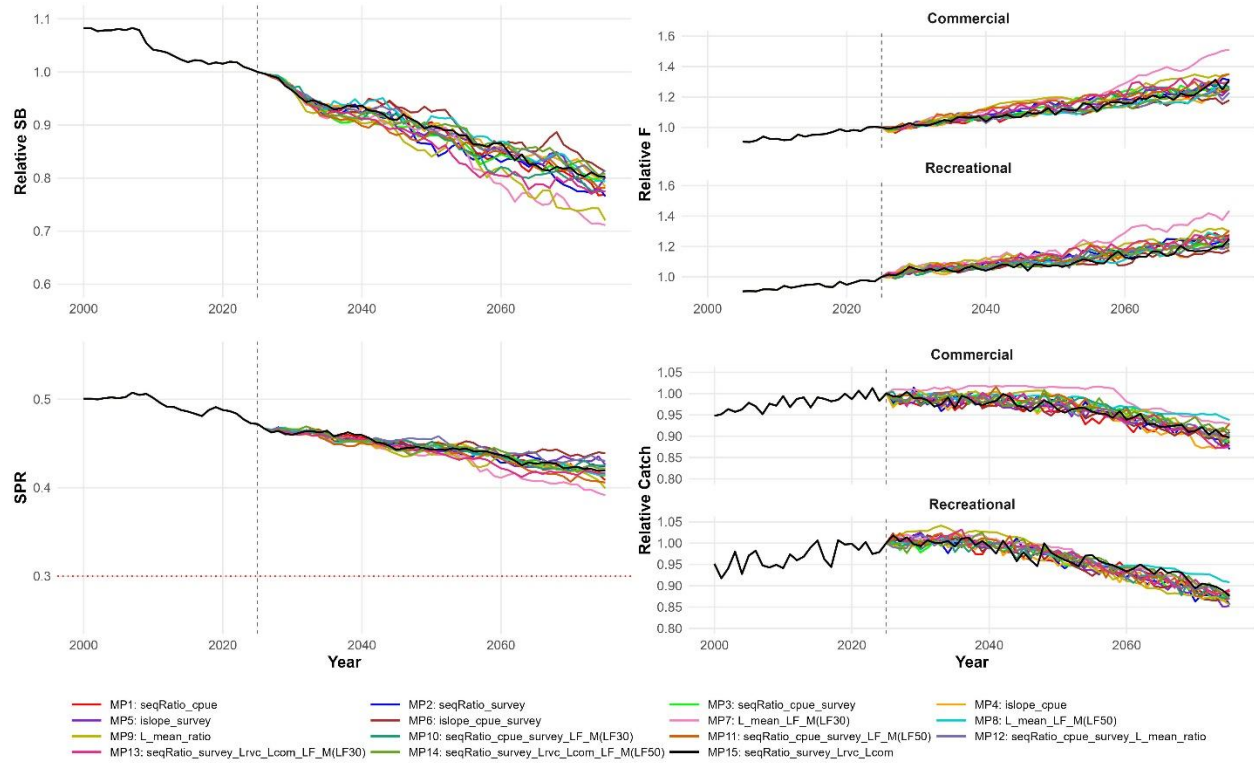


Figure 23. Median trajectories of spawning biomass, SPR, fishing mortality, and catch under all management procedures for *OM2-R1-Index (recreational & survey) hyperstability*. Metrics are relative to the final historical year (except for SPR). The vertical dashed line marks the start of projections; the horizontal dotted line shows SPR = 0.30. Lines represent the median of MPs 1-15 across 300 simulation iterations.

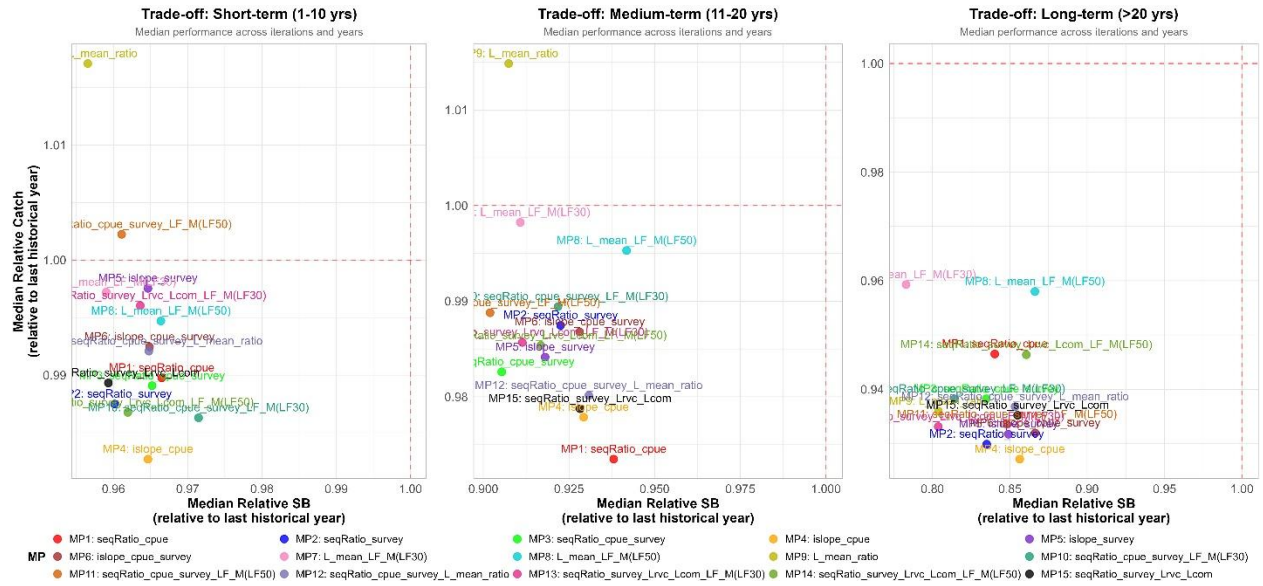


Figure 24. Trade-off between median relative catch and median relative spawning biomass across all management procedures under *OM2-R1- Index (recreational & survey) hyperstability* for the short, medium, and long term. Each color represents an individual MP.

Appendix: All MP figures

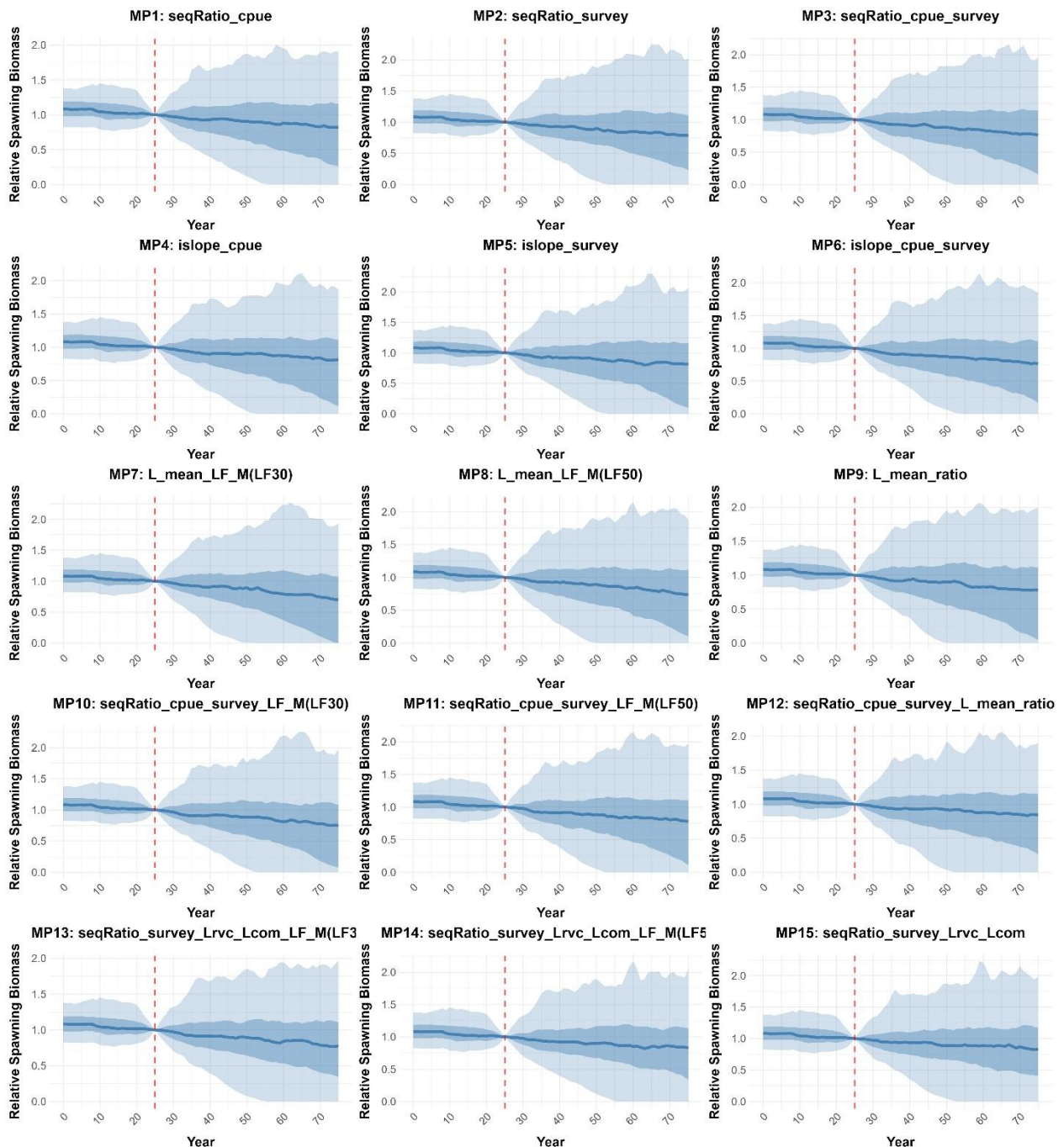


Figure A1. Historical and projected relative spawning biomass for all 15 management procedures under OM1. Each panel shows the median (solid line), 50% confidence interval (dark shading), and 95% confidence interval (light shading) across iterations. Values are standardized by the last historical year. The dashed vertical line marks the start of the projection period.

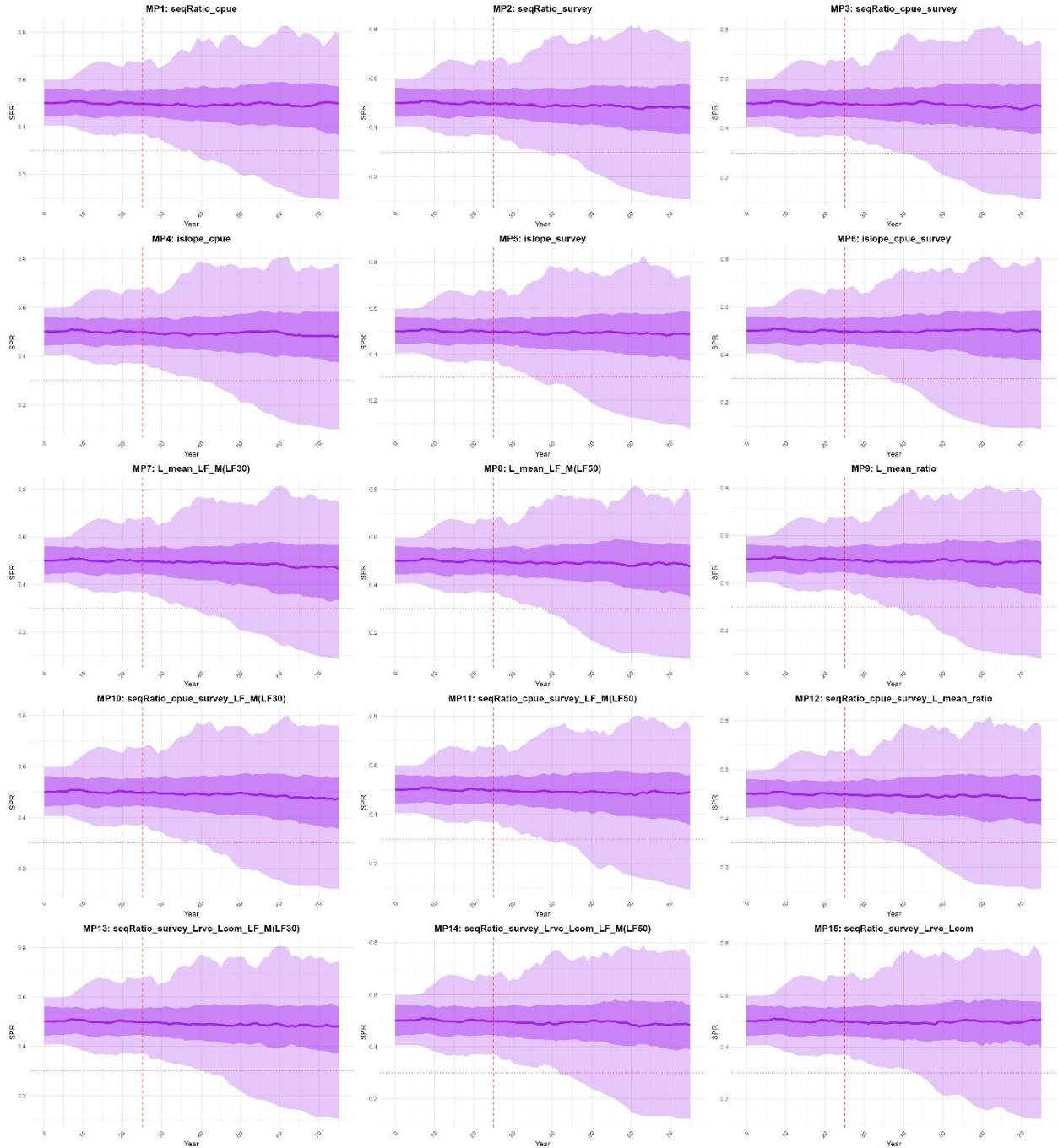


Figure A2. Historical and projected spawning potential ratio (SPR) for all 15 management procedures under OM1. Each panel shows the median (solid line), 50% confidence interval (dark shading), and 95% confidence interval (light shading) across iterations. The horizontal dotted line indicates the SPR = 0.30 reference point. The dashed vertical line marks the start of the projection period.

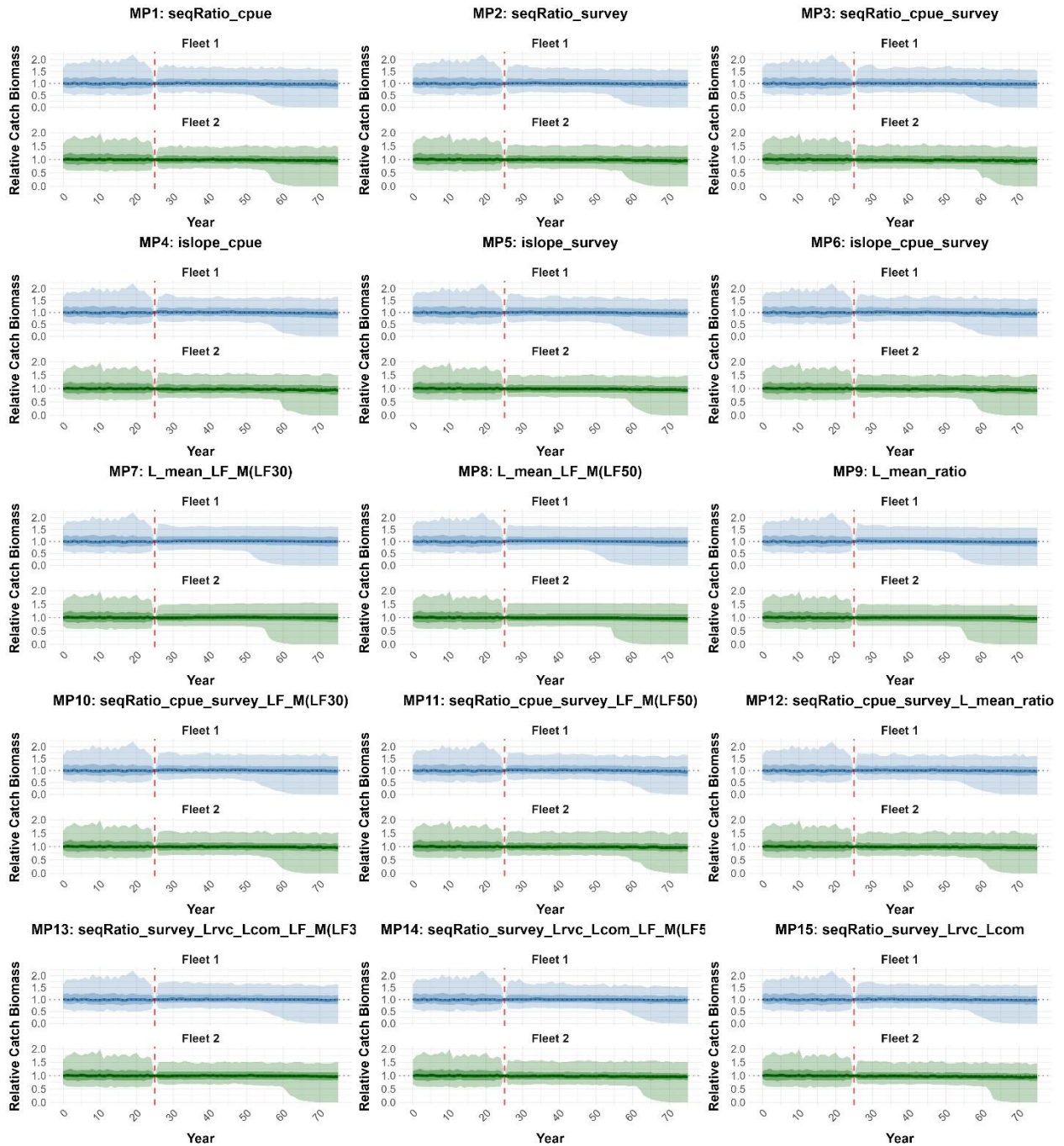


Figure A3. Historical and projected relative catch biomass by fleet (recreational, top; commercial, bottom) for all 15 management procedures under OM1. Each panel shows the median (solid line), 50% confidence interval (dark shading), and 95% confidence interval (light shading) across iterations. Values are standardized by the last historical year. The dashed vertical line marks the start of the projection period (2025).