Received: 28 July 2022 | Revised: 28 October 2022 | Accepted: 23 November 2022

DOI: 10.1111/faf.12726

ORIGINAL ARTICLE

FISH and FISHERIES WILEY

Method evaluation and risk assessment: A framework for evaluating management strategies for data-limited fisheries

Thomas R. Carruthers[1,2] | Quang C. Huynh[2] | Adrian R. Hordyk[2] | David Newman[3] | Anthony D. M. Smith[4] | Keith J. Sainsbury[5] | Kevin Stokes[6] | Alexander Morison[7] | David Agnew[8] | Ana Parma[9] | Ignacio Sobrino[10] | Catherine Longo[11]

[1]Institute for the Oceans and Fisheries, AERL, University of British Columbia, Vancouver, British Columbia, Canada

[2]Blue Matter Science Ltd, Vancouver, British Columbia, Canada

[3]Natural Resources Defense Council, New York, New York, USA

[4]Centre for Marine Socioecology, University of Tasmania, Hobart, Tasmania, Australia

[5]Centre for Marine Socioecology, University of Tasmania, Battery Point, Tasmania, Australia

[6]Stokes.net.nz Ltd., Wellington, New Zealand

[7]Morison Aquatic Sciences, Oakbank, South Australia, Australia

[8]Convention for the Conservation of Antarctic Living Marine Resources, Hobart, Tasmania, Australia

[9]Centro Científico Tecnológico del Consejo Nacional de Investigaciones Científicas y Técnicas, Puerto Madryn, Argentina

[10]Instituto Español de Oceanografía, Madrid, Spain

[11]Marine Stewardship Council, London, UK

Correspondence
Thomas R. Carruthers, Institute for the Oceans and Fisheries, AERL, 2202 Main Mall, University of British, Columbia, Vancouver, BC V6T 1Z4, Canada.
Email: t.carruthers@oceans.ubc.ca

Funding information
David and Lucile Packard Foundation; Marine Stewardship Council; Natural Resources Defense Council; United Nations Food and Agricultural Organization

Abstract

Fisheries managers are in need of quantitative tools to inform decisions regarding selection of robust management practices, prioritising research gaps and stocks to focus on, particularly where there are limited resources or data. To support these decisions, the use of Management Strategy Evaluation (MSE), that is, closed loop simulation-testing of management procedures, is widely regarded as best practice. However, applying MSE is time- and computationally intensive, and requires highly skilled expertise and processes for stakeholder input and peer review. For data- and capacity-limited fisheries, MSE may be particularly challenging to implement. Yet, these are the contexts where it is most critical to test assumptions, evaluate the implications of all sources of uncertainty and identify the most informative data sources. To facilitate wider use of MSE, the Method Evaluation and Risk Assessment (MERA) framework was developed as an accessible online interface, with quick processing time, focused on generic data-limited management procedures, but allowing progression to tailored and more data-rich methods. The framework links a quantitative questionnaire and data input standard to a flexible operating model with optional customisation via command line access to the back-end open-source R libraries. Here, we illustrate a case study application of MERA for the bocinegro (*Pagrus pagrus*, Sparidae) fishery in the Gulf of Cadiz, where in conjunction with fishery stakeholders, a custom management procedure was developed and tested and key research gaps and data collection priorities were identified. We discuss implications for wider use of MSE in various contexts, including eco-certification and fishery improvement projects.

KEYWORDS
capacity-limited, eco-certification, fishery improvement, management procedure, management strategy evaluation, simulation testing

1 | INTRODUCTION

Fisheries management systems typically involve an ongoing cycle of harvest, data collection, data processing, resource assessment, provision of management recommendations and enforcement (Walters & Martell, 2004). Within such a system, managers of fisheries must make a number of critical decisions. For example, these may include the choice of stock assessment approach and how assessment outputs are to be interpreted in the provision of management advice (i.e. a harvest control rule). Managers are often also expected to guide their science programme in the direction most in need of research, apportioning monetary budgets among various species and programmes such as data collection, scientific research, resource assessment and enforcement. Given that management generally involves the use of public money in the stewardship of public resources, there is broad consensus that decision making should be guided by evidence-based approaches, and that for many of these decisions it is best practice to use Management Strategy Evaluation (MSE) (Goethel et al., 2019; Punt et al., 2016). However, use of MSE is still mostly limited to a small minority of high value, data-rich stocks within well-resourced management areas of developed economies, with only a few exceptions (Goethel et al., 2019; Punt et al., 2016).

As many as 90% of the world's fisheries lack sufficient time-series data (e.g. annual catches, relative abundance indices) or capacity to apply conventional stock assessments, and derive management advice from such models (Costello et al., 2012). A large fraction of data-limited fisheries are small-scale fisheries in the developing world. Our experience working in such fisheries in Mexico, Indonesia, East Africa and South America, that most are essentially unmanaged with fishing pressure constrained only by prevailing socioeconomic factors. Where management measures are implemented for such fisheries, these are generally input controls such as maximum fishing effort (e.g. days at sea), seasonal closures, spatial closures and size limits (Mees, 2007; Pilling et al., 2008; Ruddle, 1996).

In an effort to expand the evaluation and management of the world's stocks, a variety of 'data-limited' methods have been developed. Several of these methods focus on estimating stock status using limited data streams such as catch time series and life-history traits (e.g. Free et al., 2020), through mechanistic models and empirically derived functions or proxies for depletion (Hordyk et al., 2015; Ovando et al., 2021). These methods, however, rely on assumptions that are particularly difficult to verify in data-limited contexts, with risk of biased or uninformative results (e.g. Free et al., 2020; Ovando et al., 2021; Rosenberg et al., 2018). Even when required assumptions are verified, their use is limited to gaining a snapshot of the status of a stock prior to the implementation of a management approach. This is because they rely on an assumed relationship between observed catch, or observed size structure, and stock population biomass that no longer holds true if catches or effort are being regulated (Free et al., 2020). In particularly data-poor cases, estimates of biological risk have been obtained using semi-qualitative approaches such as Productivity Susceptibility Analysis (PSA: Hobday et al., 2011). For example, PSA may be used to audit fisheries to the Marine Stewardship Council (MSC) Fisheries Standard (Martin et al., 2012; MSC, 2018) and in Seafood Watch fisheries ratings (2016). There are, however, limitations to reproducibility and validation of these expert knowledge-based methods (Hordyk & Carruthers, 2018), and as is the case for other methods described above, they cannot be used to evaluate management responsiveness. MSC, for example, requires that a PSA is used as part of a Risk-Based Framework where it is combined with available data (e.g. CPUE trends, in a Consequence Analysis (CA)) and its use is highly precautionary and restricted to specific circumstances (MSC, 2018).

These approaches may help assess the status quo, in some cases, but are not useful to guide management decisions or monitor their effectiveness. On the other hand, it is possible to monitor and manage fisheries sustainably through data-limited harvest strategies (e.g. Carruthers et al., 2014; Carruthers & Hordyk, 2018; MSC, 2018) that, rather than relying on estimates of population status, use indicators or proxies to inform specific control rules. These 'data-limited' indicators may be as simple as, progressive dynamic adjustments of catch or effort limits, for example, until a particular size-based indicator target is reached (e.g. Prince & Hordyk, 2019) or according to relative abundance index levels (Carruthers et al., 2016; Geromont & Butterworth, 2015; Hoshino et al., 2020; Jardim et al., 2015), or trigger action based on a system of interlinked indicators with respective thresholds (Dowling et al., 2015; Harford et al., 2021).

When adopting such data-limited management procedures (MPs, rules for calculating management advice from data, often referred to as 'harvest strategies') it becomes particularly crucial to conduct formalised performance testing, such as MSE, as, by their very nature, these MPs rely on fewer data.

There are two options for evaluating competing management strategies (including data collection, stock assessment modelling, harvest control rules and enforcement): experimentation and simulation testing using theoretical models. While models may provide unreliable predictions if they fail to adequately capture critical system dynamics (e.g. changing environmental conditions, unobserved exploitation, variable data quality), it is often the best option because the experimental approach is not practically feasible in most fisheries. Even if experimentation was feasible, the statistical power to detect system changes over relevant time horizons may be expected to be low (Legg & Nagy, 2006). Experimentation may also be expensive after accounting for the costs of additional data collection and lost fishery yields.

The theoretical testing approach relies on the development of systems dynamics models ('operating models') that encompass a range of plausible states of nature, to evaluate the expected performance of alternative management strategies. While operating models can be used to inform a wide range of management questions (e.g. the expected benefit of more stringent enforcement or alternative data streams), previous applications in fisheries have generally focused on MSE and the comparative evaluation of MPs (Butterworth & Punt, 1999; Punt et al., 2016). While there have been criticisms of the use of operating models (Rochet & Rice, 2009, in reference to MSE), the potential advantages of the approach have made it an ongoing priority for developed fishery management systems at various scales, for example, California state fisheries (CDFW, 2018), federal fisheries in the United States (NOAA, 2019) and Canada (Anderson et al., 2021; Kronlund et al., 2013) and for high seas tuna stocks (Anon, 2018). Common impediments to more widespread development and adoption of the operating model approach to fisheries management include relatively high costs (compared with a one-off assessment of the population) and the unavailability of suitably qualified analysts and data to inform various scenarios for system dynamics. Several tools have been proposed to assist with the implementation of data-limited fisheries methods. These are either focused on providing proxies to gain a snapshot of stock status, but do not test their performance in informing management, or provide qualitative guidance on potential options based on structured processes involving stakeholders and expert opinion (e.g. FishPath, 2022), but require separate software to then analytically test their assumptions and performance.

The demanding costs and expertise of the operating model approach are particularly difficult to overcome in many data-limited fisheries, that is, based in developing economies or operated by small-scale producers, which lack the technical and financial investment needed to support complex analytical evaluations and large participatory processes. In these cases, we argue that more efficient ways of organising key information about the fishery through a simpler data input approach, more effectively communicating technical concepts across different types of experts (e.g. between observers in the field and modellers) as well as other stakeholders through interactive data visualisation, can make the process more cost-effective and accessible, encouraging awareness and buy-in for this type of rigorous testing.

Here, we describe lessons learned from application of the MSE framework MERA for a range of data-limited and data moderate fisheries (Anon, 2022; Loneragan et al., 2021; MERA, 2022). The primary objective of MERA (2022), initially commissioned by the MSC, is to provide an accessible, open-source tool to support the adoption of robust data-limited MPs, by incentivising MSE testing and guiding the user through stepping stones towards improved data and more advanced methods. MERA is designed to assist in rapidly synthesising and documenting available information of a fishery in a working operating model that can be used to inform strategic decisions at various management levels including data collection, species prioritisation, MP selection and enforcement.

A secondary objective is to ensure flexibility and extensibility by linking MERA to flexible models and statistical libraries to allow for bespoke operating models, MPs and MSE when required, as well as tailored diagnostics. This ensures that MERA may be used through time, as the knowledge of the fishery and available data improve. MERA is intended to be applicable to a wide range of fisheries, data availability, biological, ecological and exploitation characteristics. Additionally, it can be used to present information and results to diverse user groups including regional fishery management organisations, seafood ecolabelling auditors and international development agencies such as the United Nations Food and Agricultural Organization.

Here we illustrate an applied case study of developing MSE for the data-limited bocinegro (*Pagrus pagrus*, Sparidae) fishery in the Gulf of Cadiz (Spain) using MERA, and discuss lessons learned on what can help overcome the barriers to wider use of MSE, particularly in data-limited fisheries.

2 | MATERIALS AND METHODS

2.1 | Overview

MERA has two inputs: a mandatory quantitative questionnaire and an optional standardised fishery data input table (Figure 1). The questionnaire contains 30 questions, 19 related to the fishery dynamics, that is, biological as well as harvesting characteristics and history, 7 questions about the management system, and a further 4 regarding the types and quality of data that are available (see Appendix A and Appendix B for details on the mapping of questions to operating model dynamics. See the Appendices of Carruthers & Hordyk, 2018 for all operating model equations).

After completing the questionnaire users select one of three modes:

1. Management Planning—determining a suitable management procedure,
2. Management Performance—evaluating the current management procedure,
3. Status Determination—calculating stock status using available data.

The operating model is specified by sampling parameters for fishery and population dynamics. The historical dynamics of the

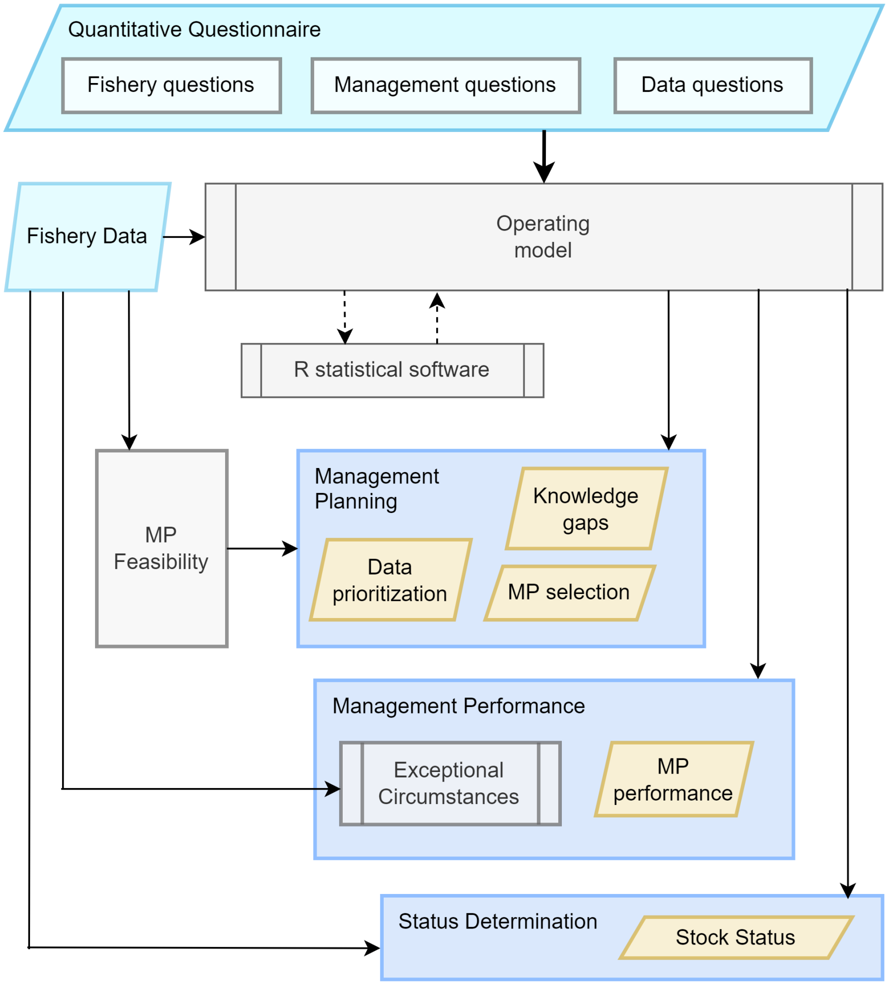

FIGURE 1 MERA application structure and modes of use. Dashed lines represent optional functionality.

fishery can be reconstructed using only the sketched effort and specified depletion in the quantitative questionnaire. However, if time series data are available such as annual catches, relative abundance indices or size composition data, the operating model may be conditioned (statistically fitted) on those data. Figure 2 illustrates these two pathways to specifying the operating model. For more detail on model conditioning and data requirements, we refer the reader to the MERA user guide (Carruthers, 2022). The status determination mode always requires additional data to estimate stock status. Any operating model can be structured to include important aspects of system dynamics that are not included in the quantitative questionnaire (such as spatial structure, fleet structure or sex-specific growth). In such cases, the operating model derived by the MERA questionnaire may be exported into the R statistical environment, modified and re-imported. MERA operating models are generated by the openMSE R packages (Hordyk et al., 2022): DLMtool (Carruthers & Hordyk, 2018) and MSEtool, and SAMtool (Huynh et al., 2022). Where required, these packages contain functions to implement complex dynamics such as ontogenetic habitat shifts (e.g. age-based movement), fine-scale movement dynamics, time-varying movement, temporal changes in growth and natural mortality rate, shifting fleet size selectivity and economic constraints.

2.2 | Quantitative questionnaire

The quantitative questionnaire is designed to include prevalent fishery dynamics that are critical in determining the relative performance among classes of management procedures (see Tables A1–A3 for the various questions, answers and corresponding operating model parameter ranges). For example, longevity determines temporal variability in biomass and the responsiveness of the stock to changes in fishing dynamics which in turn affects whether MPs need to be closely linked to recent data. Furthermore, management objectives may require MPs that are biologically conservative and are likely to rebuild the population or those that are either less responsive to stock status or recommend higher exploitation rates. Similarly, fishery selectivity, size at maturity, discard rates and post-release survival, all determine the relative success of MPs that regulate the size of capture (e.g. a minimum size limit).

For every question, all answers are initially selected by default, incurring the widest range of uncertainty in dynamics. If the user wishes to reduce the number of selected answers for a question, they are obliged to provide a written justification for the commensurate narrowing of uncertainty (e.g. citing values from other sources, Figure 3). This, in combination with assigning broad

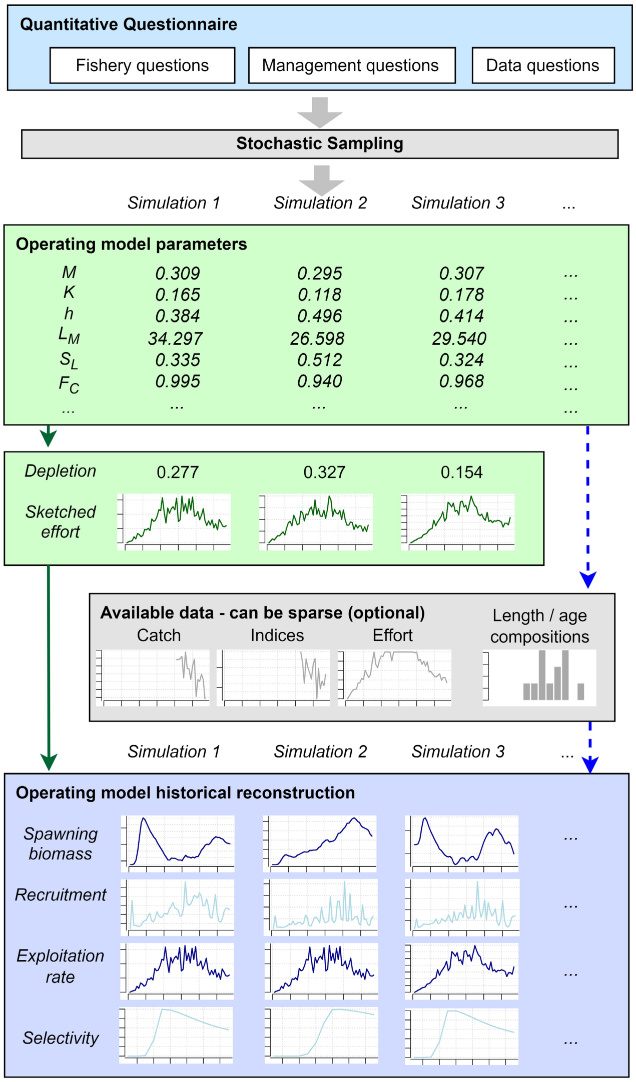

FIGURE 2 Specification of MERA operating models. Operating model parameters are stochastically sampled from the quantitative questionnaire to provide multiple simulations of system dynamics. These include population parameters such as natural mortality rate (M), von-Bertanffy growth rate (K), steepness of the stock-recruitment relationship (h), length at 50% maturity (L_M) and fisheries parameters such as selectivity of the largest fish (S_L) and observation model parameters such as bias (a fraction) of catches that are reported (F_C). The operating model can be specified using only the quantitative questionnaire, including sampled values for stock depletion and the sketched historical effort (data-limited pathway, green solid arrows on the left) or alternatively, each simulation can be conditioned on available time series data (blue dashed arrows on the right).

ranges to each selectable answer, means that MERA is predisposed towards uncertain system dynamics, and hence a more precautionary evaluation of biological risk (Risk Assessment), a more stringent test of MP robustness (Management Planning, Management Performance) and greater uncertainty in estimated stock status (Status Determination).

2.3 | Constructing operating models from the quantitative questionnaire

By default, operating models are constructed by sampling parameters from a truncated normal distribution defined by the parameter ranges of answers in the questionnaire. The exceptions are the parameters controlling somatic growth rate, which are imputed from the answers for natural mortality rate and maturity using meta-analysis of life-history parameters (joint distributions of natural mortality rate, length at 50% maturity and von Bertalanffy growth parameter κ) (Thorson et al., 2017). Time series of historical exploitation pattern and age-dependent variables such as maturity, growth and fishing gear selectivity are also sampled. For each simulation (each parameter/time series sample), numerical optimisation solves for the magnitude of historical exploitation (i.e. fishing mortality) that matches the sampled level of stock depletion (see Carruthers and Hordyk (2018) for more details).

2.4 | Importing data

Data can be imported to MERA either in a standard object class ('Data', see Hordyk et al., 2022) or from a suitably formatted spreadsheet or text file. These data can be incomplete and patchy, including biological parameters controlling mortality and growth, and time series of fisheries catches, indices of abundance (total biomass, vulnerable biomass, spawning biomass) and size composition of catches. Once imported, these real fishery data can be used to statistically fit operating models, estimate the status of the stock (Status Determination mode) and to identify which MPs are feasible (Management Planning mode). Given a completed MERA questionnaire that includes a sketch of historical fishing effort, an operating model can be conditioned using only a single observation of catch or two observations of a relative abundance index. By importing data, various aspects of the operating model are now estimated (overriding aspects of the questionnaire), ensuring that simulated dynamics such as the scale of the fishery and population, the historical pattern of exploitation and the degree of stock depletion, are consistent with empirical observations. Additionally, uncertainties in the data are also then reflected in the operating model.

3 | CASE STUDY

Here, we demonstrate the use of some of the MERA modes in the case study of the bocinegro fishery of the Gulf of Cadiz, Spain. The case study was developed at a workshop convened by the Marine Stewardship Council and the University of British Columbia in Cadiz, Spain in October 2018, including participants from the Instituto Español de Oceanografía that were familiar with the local system and providing scientific advice. The workshop participants specified and documented a MERA questionnaire and discussed the history

FIGURE 3 MERA user interface and quantitative questionnaire. In this screen shot, the user has specified the second of 19 fishery questions, regarding species longevity, and has provided a link to justify their selection (bocinegro, *Pagrus pagrus*, in the Gulf of Cadiz).

and current status of the fishery, data availability and current management. For more detailed information about this case study, the MERA files are available (Anon, 2022) from the online library of MERA case studies (MERA, 2022).

While the commercial fishery for bocinegro began in the 1960s, catch data and a catch-per-unit-effort based relative abundance index are only available since 2003 (Figure 4). The highest catch, in 2006, was just 67 tonnes. As such, bocinegro provides a typical

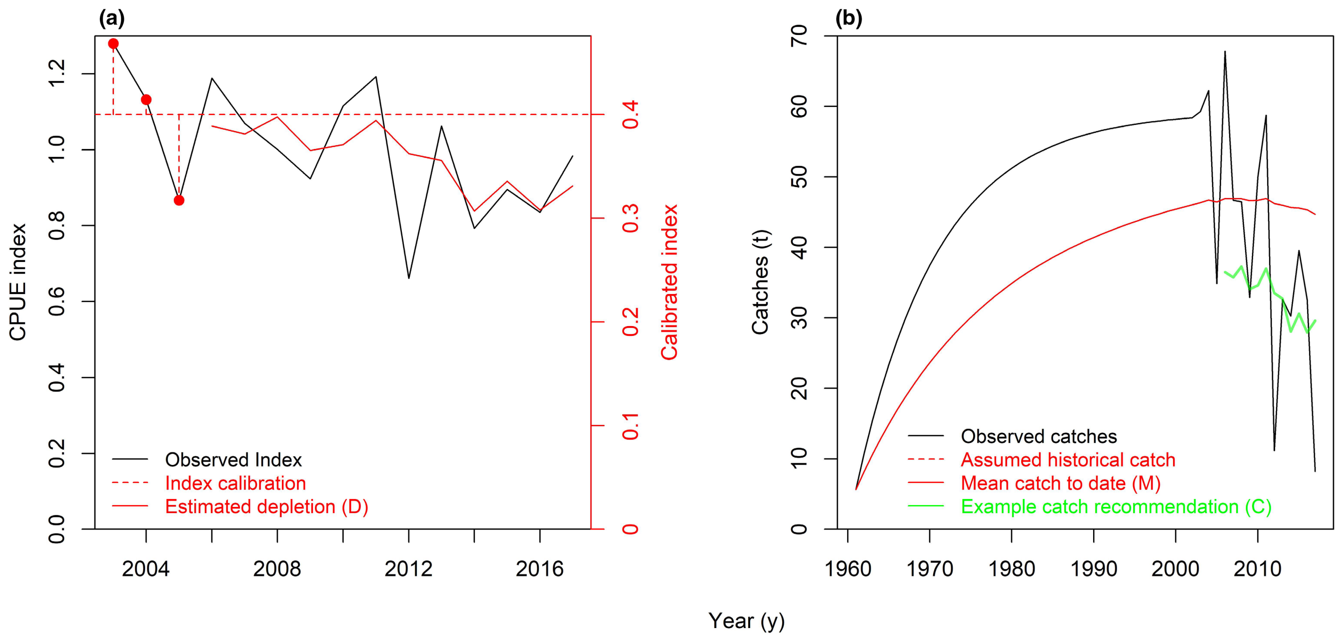

FIGURE 4 Observed (black) CPUE relative abundance index and catch data for bocinegro. Red annotations describe the assumptions of the custom MP. Stock depletion was assumed proportional to the index and calibrated such that the mean value of the index from 2003 to 2005 (red points) corresponded with a depletion level of 0.4 (this assumed level of depletion is a control parameter in the MP). The red solid line in panel a is estimated stock depletion *D*: The mean calibrated index over the most recent 3 years. In panel b, increases in historical catches were assumed to decelerate to an asymptote (red dashed line) equal to the 2003 value from which mean historical values could be calculated (red solid line). For illustrative purposes, example catch recommendations (green solid line) are included that are calculated by the custom MP for previous years from 2006 to 2017 (the catches that would have been recommended by the MP given the data to date).

example of a small-scale, data-limited fishery (for further details on the bocinegro MERA questionnaire see Anon, 2018).

To test a realistic management option, a custom MP was developed to provide catch advice from the data that are currently available. MERA includes more than 100 pre-coded MPs ranging from data-poor (e.g. management prescriptions such as size limits) to data-rich (e.g. stock assessments fitted to indices of abundance and size composition data). These MPs can provide management advice in the form of, catch limits, fishing effort limits, size limits and spatial closures, or combinations thereof. However, in the case of bocinegro, only catch limits were considered viable by the participants since other management approaches could not be reliably enforced. Initial investigation of generic MP types identified index ratio approaches (setting catch advice to a fixed ratio of the abundance index data) as the best performing approach given the systems dynamics of bocinegro. However, such approaches were not feasible since they require a complete catch history that is not available for bocinegro. Additionally, they also provided relatively uncertain biological performance outcomes over the medium term (e.g. a wide range of stock depletion over a 10 year projection). A custom MP was developed with the broad objective of providing comparable yields to the current fishery but sufficient responsiveness to obtain acceptable long-term conservation performance. The MP imputed catches for the missing years in the real historical data and also included a control parameter for assumed stock depletion over recent years that allowed for a simple control rule to be used to maintain a narrower projected range of stock biomass. The custom MP assumed that relative abundance indices were calibrated such that the mean index value from 2003 to 2005 corresponded with a stock depletion of 40% (an MP control parameter) (Figure 4a) and that prior to the first observation in the year 2000, historical catches had increased rapidly from zero in 1959 and then decelerated to an asymptote equal to the catch in 2000 (Figure 4b). The value of the depletion control parameter was selected as an intermediate value in the range of depletion estimates for bocinegro over the years 2003–2005. In the bocinegro case study, this control parameter remained unchanged at the value initially proposed. However, consistent with MP refinement in other settings, the value of the control parameter could have been adjusted to better achieve management performance objectives, for example, reducing short-term mean yields in favour of reduced risk of depleted biomass levels over the long term. Based on this control parameter, an estimate of stock depletion *D*, was obtained as the mean calibrated index over the most recent 3 years. For year *y*, the custom MP sets catch advice *C*, according to mean historical catches to date *M*, and estimated depletion where: $C_y = 2M_{y-1}D_{y-1}$. This was an index-based adaptation of the 'MC_adj' MP tested by Harford and Carruthers (2017). It should be noted that when developing such an MP, the focus is evaluating the management performance of the MP as revealed by close-loop testing rather than the theoretical validity of the functional form and assumptions of the MP algorithm.

3.1 | Management planning mode

Management Planning mode aims to provide direct guidance to fishery managers on prioritising data collection (see 'value of information analysis' below), critical knowledge gaps (see 'cost of current uncertainties analysis', below), the system of fishery management (e.g. catch control, size limits, spatial closures), the relative importance of enforcing management measures and ultimately the selection of an MP that is likely to meet their management objectives. If real fishery data are uploaded, the feasibility of applying the MPs to the data is automatically evaluated and can be used to filter the results (i.e. MPs are flagged that require more data types or more complete data than are available).

At the heart of the Management Planning mode is a closed-loop simulation that compares the performance of numerous candidate MPs, including more than 100 data-limited and data-moderate MPs from the data-limited MSE R package DLMtool (Carruthers & Hordyk, 2018) and 20 data-rich (stock assessment-based) MPs from the data-rich MSE R package SAMtool (Huynh et al., 2022). Using these packages, users can rapidly develop custom MPs in R and import these into MERA. The Management Planning mode provides access to a wide range of information about what occurred during MSE testing, including projected biomass, catches and fishing effort. These can be summarised in time-series projection plots, and performance metrics can be compared among MPs in performance trade-off plots (Figure 5).

The Management Planning results for bocinegro show a characteristic negative trade-off among MPs between biomass objectives and yield objectives (Figure 5). The shape of the trade-off was convex, providing some MPs with a desirable compromise among yield and biomass outcomes (e.g. the Custom MP or constant current fishing effort 'Cur. Eff'). An idealised statistical catch-at-age stock assessment (fishing at FMSY estimated from perfect information) obtained higher probabilities of obtaining more than 50% MSY yields at the cost of a lower probability of remaining above the biomass reference level of 40% of unfished biomass levels. Given the data available and that management can only enforce a TAC (Total Allowable Catch) limit, the Custom MP was the only feasible management option. The Custom MP was a derivative of MC_adj (Harford & Carruthers, 2017) and provided comparable performance to MC_adj that relies on a reliable estimate of stock depletion. The Custom MP performed similarly to status quo fishing effort (Cur. Eff) but with

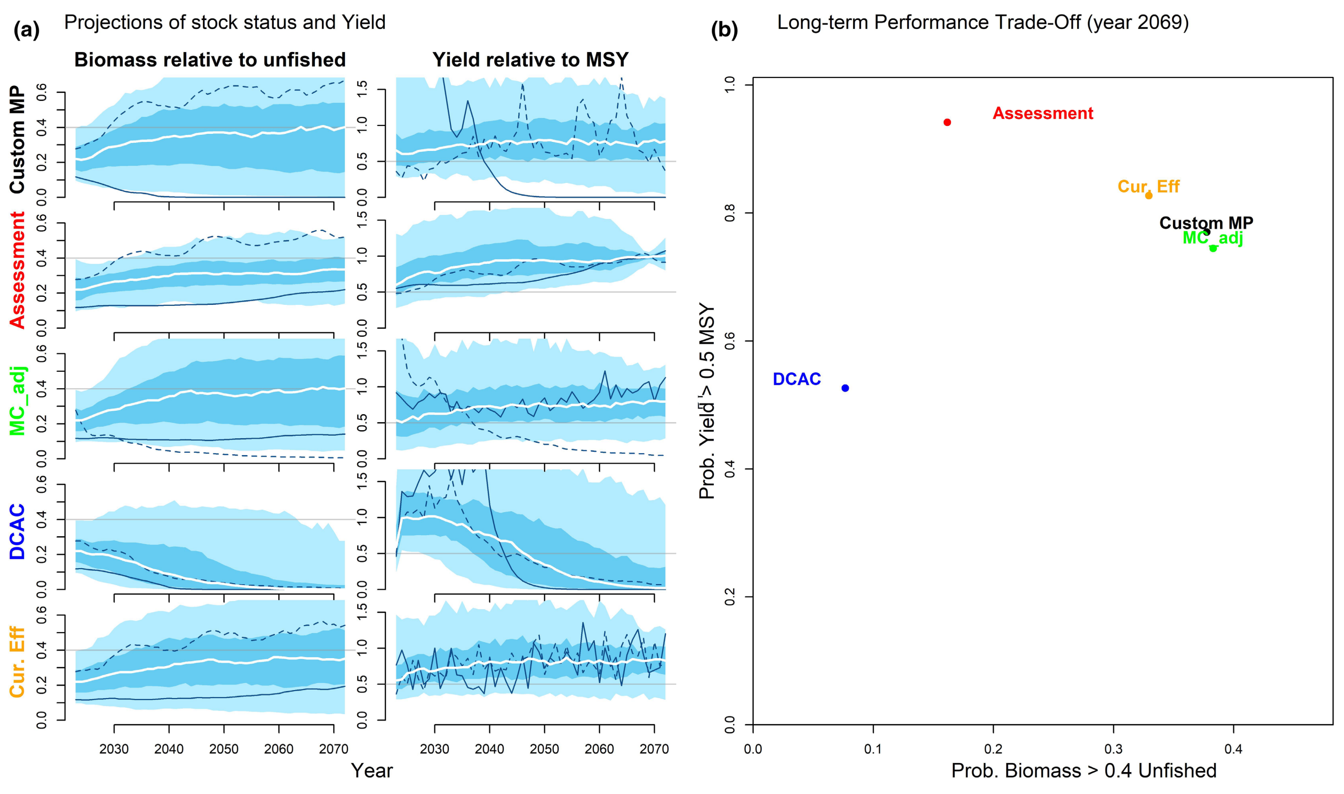

FIGURE 5 Projected fishery biomass and yield for five MPs and the corresponding performance trade-off in the long-term probability of exceeding reference levels (biomass above 40% of unfished, yield above 50% of MSY). Over a 50-year period, the performance of the Custom MP was compared with a data-rich stock assessment (Assessment—a perfect-information, statistical catch-at-age model fishing at F_{MSY}), a depletion-based harvest control rule (MC_adj—from which the Custom MP was derived, see above), Depletion-Corrected Average Catch (DCAC, MacCall, 2009) and status-quo current fishing effort (Cur. Eff). The dark blue and light blue shaded areas of the projection plots (a) are the 50% and 90% probability intervals (200 simulations). The white line is the median value over all simulations. The solid and dashed dark blue lines are two example simulations. A reference level of 40% unfished biomass was intended as a proxy of MSY biomass (grey horizontal line). To calculate yield performance an arbitrary level of 50% of MSY was used. The performance trade-off plot (b) reveals a typical negative relationship between long-term biomass and yield outcomes among all MPs except DCAC.

slightly lower yield performance and better biological performance (Figure 5). The average catch approach DCAC (MacCall, 2009) provided anomalous performance and tended to crash the stock providing worse yield and biomass performance than the other MPs.

For any MP that is tested in closed-loop simulation, performance can be compared against the simulated quality of each data type, on a simulation-by-simulation basis. For example, if users specify catch under-reporting between 30% and 10% (Data question 2, answer 2) these sampled values for catch reporting can be correlated with the projected yield for each simulation to determine the potential value of improving catch reporting. Similar to parameters controlling the quality of data, the parameters of the operating model system dynamics (e.g. natural mortality rate, somatic growth rate, catch overages) also can be related to quantities such as yield. This 'Cost of Current Uncertainties' analysis reveals which areas of current understanding are most in need of further investigation, including uncertainties in how well management advice is implemented (Figure 6).

In the bocinegro case study, the strongest determinant of future yields was the simulated range of stock resilience ('Steepness') that controls the extent to which expected recruitment is impacted by reductions in stock level (Fishery question 4, F4) followed by the simulated range of current stock depletion (Fishery question 3, F3), the interannual variability in recruitment (Fishery question 15, F15) and the implementation error in the TAC (the extent to which TAC implementation varies around TAC advice, Management question 3, M3) (Figure 6). Similarly, the value of information analysis (B) reveals that the yield performance of the Custom MP for bocinegro was similarly impacted by observation error in annual catches (Catch Err.), observation error in the relative abundance index (Index Err.) and the level of consistent biases in catch reporting (Catch bias).

3.2 | Management performance mode

It is considered best practice in MSE to compare real data that are collected while an MP is in use with those data predicted when the MP was adopted (Carruthers & Hordyk, 2019; Punt et al., 2016). If the predicted and observed data differ substantially this may indicate unexpected MP performance due to operating model misspecification (often called 'exceptional circumstances'), and therefore require a review of the operating models (and potentially the advice provided by the MP). MERA records projected catch and length data in addition to vulnerable, total biomass and spawning biomass indices. The user can then collect and upload to MERA any observed data of any of these types and compare these with those projected (Figure 7). This provides a necessary empirical check that the basis for MP adoption is still supported by observations, and additional motivation to continue and improve data collection while an MP is in use (Carruthers & Hordyk, 2018).

An MP has not yet been adopted for use in managing the bocinegro fishery. To demonstrate Management Performance mode, three simulated data sets were generated for the bocinegro case study, using the Custom MP. The 'consistent' data were generated from the same 'Base' operating model that provides the posterior predicted data. The 'less consistent' data were generated from the same Base operating model but with stock status and natural mortality rate that were 20% lower. The 'inconsistent' data set was generated from an operating model with stock status and natural mortality rate that were 40% lower. Figure 7 shows the ability to detect these operating model misspecifications using the exceptional circumstances protocols contained in MERA. In this case, an increasing frequency of 'outlier' observations can be seen in the 'less consistent' and 'inconsistent' data sets, demonstrating that the protocols would have correctly identified those cases of operating model misspecification.

3.3 | Status determination mode

The status of an exploited population is central to legal frameworks, seafood certification standards and a quantity of principal interest to various stakeholders including environmental NGOs. MERA uses an age-structured rapid conditioning model similar to a stock assessment (function 'RCM' of the SAMtool package, Huynh et al., 2022)

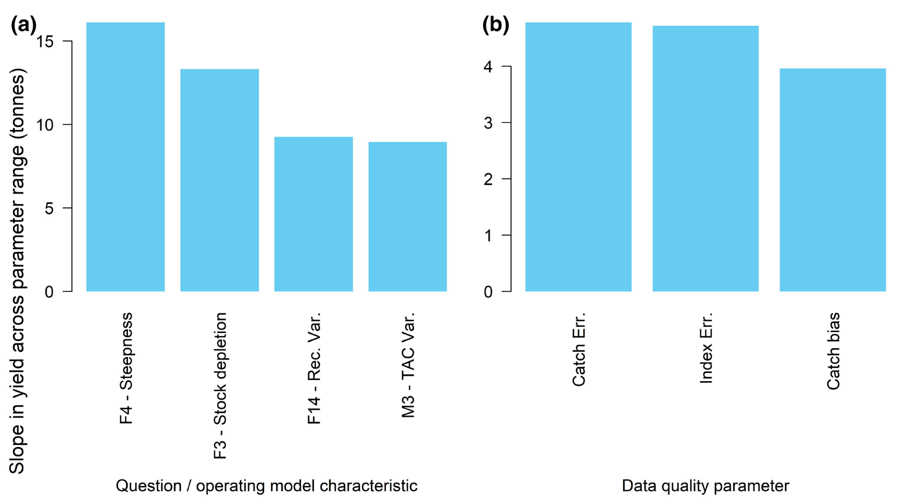

FIGURE 6 Cost of current uncertainties (a) and value of information (b) diagnostics for the custom MP in the management planning mode. The cost of current uncertainties analysis examines the value (expressed here as the slope in fishery yield with respect to each parameter) across the parameters corresponding with the answers provided for each question. For example, as illustrated in panel a, the largest difference in yield occurs across the specified range in stock resilience ('steepness', fishery dynamics question 4, F4).

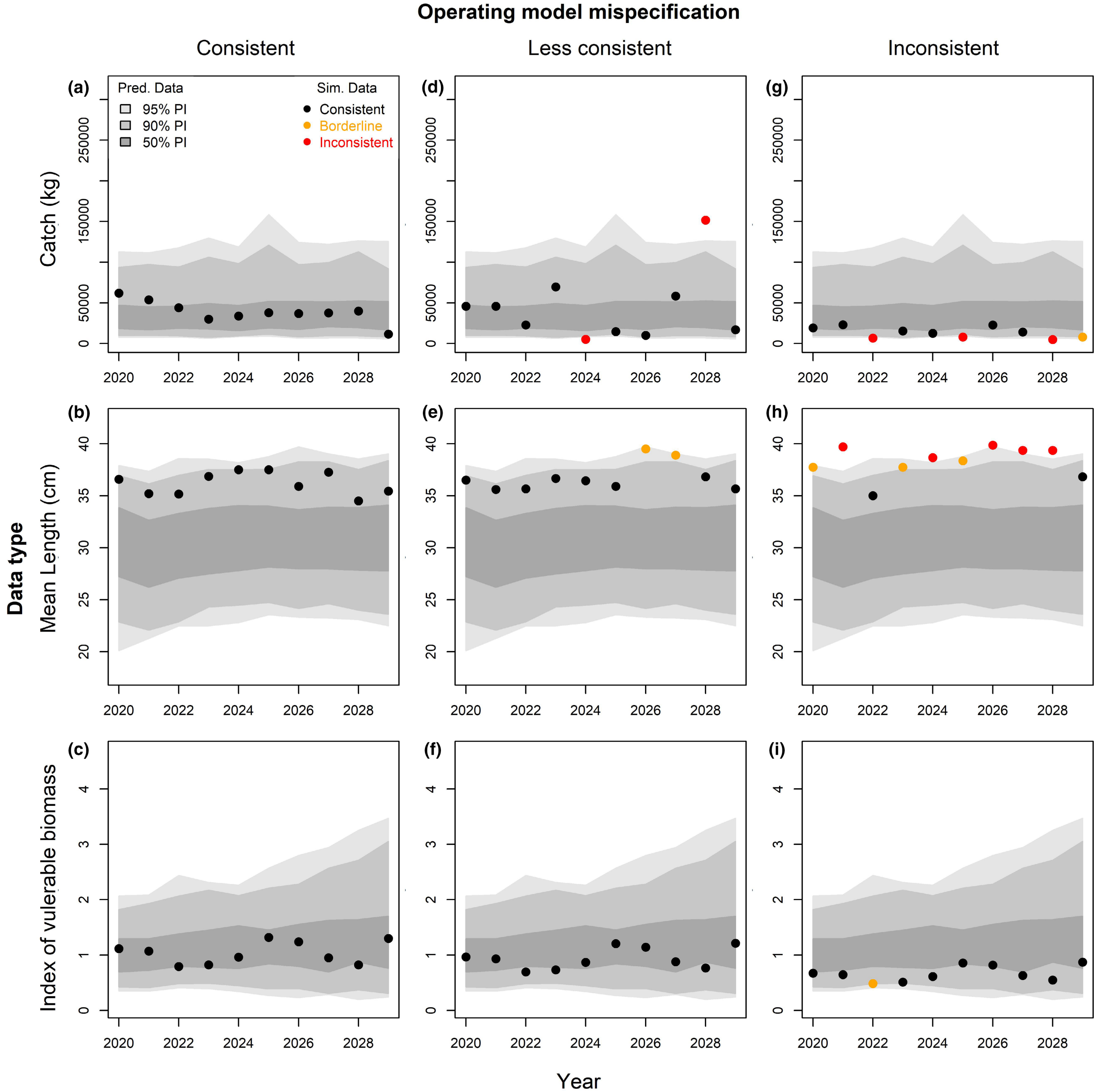

FIGURE 7 Exceptional circumstances in the management performance mode. Since an MP has not been in use in the case of bocinegro, simulated data were used to demonstrate exceptional circumstances protocols. These simulated future data originate from the bocinegro MP and an unmodified version of the bocinegro operating model (consistent, panels a–c), the same MP projected for an operating model that is more depleted and has lower natural mortality rate (less consistent, panels d–f) and lastly an operating model where the stock is even more depleted and has even lower natural mortality rate (inconsistent, panels g–i). In each row the posterior predicted data (shaded grey areas) of the base operating model is unchanging, only the simulated projected data (points) vary. Data that fall inside the 90th probability interval are coloured black, those between the 90th and 95th probability intervals are coloured orange, and those outside of the 95th probability interval are coloured red.

that uses the Template Model Builder R library to estimate stock status (i.e. 'stock depletion', spawning stock biomass relative to equilibrium unfished levels) based on more than 30 combinations of data types (e.g. time series of catches, effort, relative abundance, size and age composition). MERA automatically detects the models that can be applied for status determination when the user uploads a data set.

For the bocinegro case study, four status determination approaches were available that used the various combinations of data. The approach applied was 'C_E_I' which used all three of the available data sets: recent catches (C) to scale the stock, fishing effort (E) sketched by the user in MERA to derive the historical pattern of exploitation, and a recent fishery-dependent index of vulnerable biomass (I). Status Determination reconstructed the stock in which rapid declines from unfished conditions occurred in the first 20 years following gradual rebuilding and a recent decline over the last 15 years (Figure 8). The current spawning stock biomass is relatively uncertain and estimated to be between 15% and 45% of unfished levels (90% CI).

4 | DISCUSSION

MERA is designed to provide a simulation-testing framework that is more readily accessible than more traditional command line MSE software. A graphical interface with a simple questionnaire and guided steps is effective in bringing all the information around a fishery into a single, transparently documented, quantitative framework with efficient computation. The framework can be applied as a first approach to select MPs that can achieve specified performance objectives, prioritise data collection and establish indicators for detecting exceptional circumstances. This type of framework can guide practitioners new to MSE to understand its application, or, for more experienced users, it can be a means of quickly facilitating group discussions and input from other stakeholders. We argue that these uses are particularly valuable in data- and capacity-limited fisheries to promote robustness testing and careful consideration of the major sources of uncertainty.

MERA is intended to offer a quantitative solution to fisheries that can only use data-limited and data-moderate options, by identifying MPs that are expected to achieve management objectives without accurate and precise estimates of stock status. The three use cases demonstrate how MERA can be used to account for uncertainty in the fishery system in order to identify robust management procedures and quantify the value of alternative data collection and research programmes. With operating models and closed-loop simulation testing, MSE can account for biological, fishing and observation dynamics and provide quantitative MP performance outputs, for example, probabilistic estimates of biomass relative to reference levels, which are central to fishery legal frameworks and eco-certification standards.

The potential capabilities of the framework provided by MERA were explored in an application to the bocinegro fishery in the Gulf of Cadiz (see Loneragan et al., 2021 for MERA application to seven stocks in Indonesia). The bocinegro fishery is small and management advice is not actually provided. It was used here as an example of how MERA might be used. Feasible MPs were developed, research priorities were identified and stock status was quantified. To provide potential TAC advice for bocinegro, a new custom MP was designed that imputed missing catch data prior to 2003. When subjected to simulation testing in the Management Planning mode, the custom MP performed reasonably well and was identified as a feasible management option with an intermediate trade-off between biological and yield outcomes. Value of Information and Cost of Current Uncertainties analyses identified research and data priorities for bocinegro, including data-processing/collection protocols for improving the precision of the annual catch data and relative abundance indices. Given the scale of the bocinegro fishery, the maximum theoretical value of these improvements was small in absolute terms at around 5 tonnes per year but large in relative terms (10–15% of recent catches) for research programmes relating to data precision (Figure 6).

The patchy bocinegro data allowed for only four of the possible 30 status determination methods that use catch or relative abundance index data. The availability of data for bocinegro is typical of a data-limited fishery, where catch and index data are unavailable for the time period of the historical fishery. The operating model conditioning approach used in MERA can use historical fishing effort data or user-sketched effort to constrain the historical stock

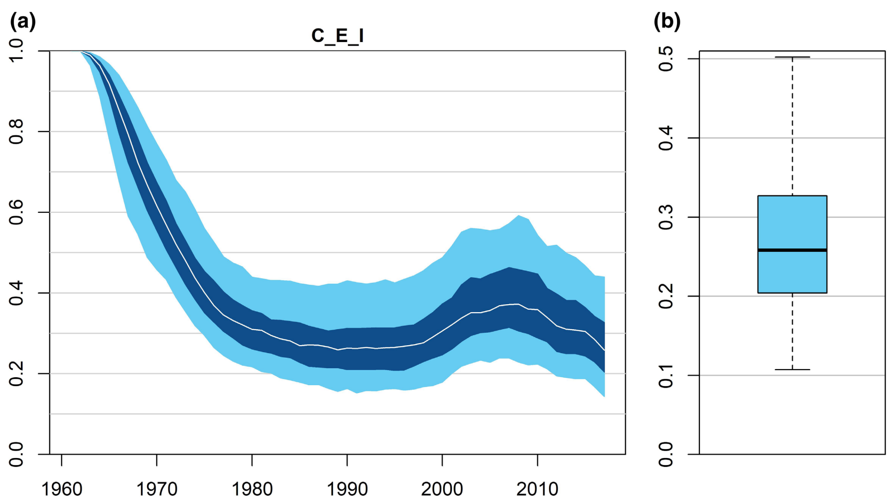

FIGURE 8 Spawning stock biomass relative to unfished (biomass in 1961) estimated by the rapid conditioning model using the combination of data 'C_E_I' that includes recent catches (C) sketched fishery effort (E), and recent relative abundance indices (I). Panel a is the estimated historical stock status where the white line denotes the median value, the dark blue box represents the 50th probability interval and the light blue region represents the 90% probability interval. Panel b is a boxplot where the 50% interval is denoted by the shaded box and the 95% probability interval is denoted by the whiskers.

reconstruction. When applied to bocinegro the conditioning approach led to relatively wide estimates of depletion in the range of 0.15–0.45 (90% CI) which provides a suitably challenging basis for the testing of candidate management procedures for a data-limited fishery.

There is a critical need for tools to inform the design and selection of robust fishery management systems in the face of imperfect information and high uncertainty. This can be considered a primary issue for global fisheries management since around 90% of stocks are data-limited and lacking sufficient data to conduct a conventional assessment (Costello et al., 2012). This fraction is still as high as 60% of stocks even in countries with developed fisheries management systems (e.g. USA, Newman et al., 2015). Without frameworks to demonstrate the potential value of alternative management options and data collection programmes, progress towards sustainability is stalled in many fisheries. System dynamics modelling offers a possible path to address these challenges since it can evaluate the implicit performance of a prescriptive management procedure (e.g. a size limit or spatial closure) for a data-limited fishery for which it may never be possible to defensibly derive an explicit estimate of stock status (Carruthers & Hordyk, 2018).

By allowing for rigorous evaluation of management options for data-limited fisheries, frameworks such as MERA may unlock new incentive systems for improving the science and management of these stocks. This dynamic approach can facilitate more rigorous and formalised documentation of qualitative and expert knowledge used to inform a management procedure. It can also promote more thorough explorations of the risks posed by a broad range of sources of uncertainty, including poor data quality or management enforcement gaps. More broadly, MP testing also emphasises implementation of robust and responsive management procedures over the identification of stock status as a means to achieve biological and fishery objectives (Hilborn, 2002). The bocinegro case study illustrates a first step of an iterative ongoing approach that cycles through status estimation, MP identification, MP adoption and monitoring of exceptional circumstances.

The versatile and powerful uses of MSE come with trade-offs associated with its analytical complexity, broad scope of information covered, and its iterative, multi-stakeholder nature.

Even with a more user-friendly front-end such as MERA, the construction and conditioning of operating models require dedicated expertise. The interpretation of outputs demands a good understanding of the analytical properties of the underlying operating model and of the management procedures being tested. In addition, a dedicated effort is needed to engage different experts to inform a realistic operating model, collate and prepare data from different sources—especially in capacity-limited contexts where data governance and understanding of data collection protocols may be lacking. Further, generic MPs such as those included by default in MERA will only be suitable in some contexts. In the bocinegro case-study, for example, the default MPs resulted in highly uncertain outcomes. A custom MP, coded through the command line version of the underlying packages (openMSE) and re-imported in MERA, resulted in better yield and biological performance, suggesting it made better use of the available data. Many fisheries are multi-gear, multi-fleet and, particularly in the case of small-scale operators, target mixed species. In such cases, command-line tailoring or additional modelling tools may be necessary to appropriately incorporate key variables (e.g. spatial, multi-species interactions), or to consider other key influential drivers (e.g. environmental, economic). MERA offers the advantage of being easily linked to R command line packages that cover the gamut of these additional analyses, but these further applications will have to rely on a well-trained MSE expert familiar with the underlying R packages (i.e. Hordyk et al., 2022).

For species that are strongly environmentally driven and short-lived, MERA is unlikely to produce precise projections that clearly distinguish between the relative performances of alternative management procedures. Since MERA assumes an annual model for population dynamics equations, it may not be appropriate for approximating the dynamics of species with multiple reproductive events during the year or those that change sex. This implies that MERA would not be suitable for evaluating pelagic fish stocks that live <3 years, and many invertebrates, without significant tailoring of the operating model (i.e. exporting the operating model from MERA, altering it using bespoke code and importing it back into MERA).

To follow best practice in the implementation of MSEs, multiple MERA analyses (operating models and closed-loop simulation tests) may be required so as to test different realistic alternative operating models, and explorations of the simulations delivering more extreme values (Punt et al., 2016) which may be easier to automate through command line coding rather than the MERA interface.

Lastly, a governance process is needed to ensure the inclusion of stakeholders in defining the management targets, assuring a transparent and rigorous analysis and peer-review process, facilitating the clear communication of outcomes to all interested parties, and endorsing and adopting final recommendations. These activities rely on highly skilled staff and meeting costs that are often underestimated when using data-limited methods. Trained experts and well-designed processes may not even be available if MSE is not an embedded practice in the management advisory system.

Despite the need for expert support and, in many cases, additional work tailoring the MERA code or performing analyses outside of the MERA environment, we propose that this tool can play an important facilitation and capacity building role in introducing MSE to data-limited fisheries' management systems. It can help gather all relevant information from different experts to build an appropriate operating model, through an interactive questionnaire; make the best use of available data for conditioning; create greater awareness of the pitfalls of data-limited methods, encouraging more careful scrutiny; support consistent documentation for all parameter selections; and incentivise stepping stones of improvement, by helping prioritise what data to collect and review the current procedure in place using auxiliary data.

Implementation of an appropriate procedure, coupled with monitoring of some response variables that are able to track progress of the fishery will enable data-limited fisheries that are in fishery improvement projects (FIPs) to demonstrate their progress to FIP providers and retailers, and ultimately can build the best practice and evidence that will be required to ensure they meet sustainability benchmarks and, perhaps, even gain certification to eco-certification standards such as MSC.

ACKNOWLEDGEMENTS

MERA was initially commissioned by the Marine Stewardship Council, and benefits from the ongoing support of the David & Lucile Packard Foundation, the Marine Stewardship Council, the Natural Resources Defense Council, the Walton Family Foundation and the United Nations Food and Agricultural Organization.

DATA AVAILABILITY STATEMENT

The MERA code and documentation are hosted at https://github.com/blue-matter/MERA/ and the application, case studies and user guide are available at https://www.merafish.org.

ORCID

Thomas R. Carruthers https://orcid.org/0000-0001-9173-9043

REFERENCES

Anderson, S. C., Forrest, R. E., Huynh, Q. C., & Keppel, E. A. (2021). A management procedure framework for groundfish in British Columbia. *DFO Canadian Science Advisory Secretariat (CSAS) Research Document*, *2021/007*, vi + 139.

Anon. (2018). Report of the 2018 joint tuna RFMO management strategy evaluation working group meeting, Seattle. 24 pp. http://www.tuna-org.org/Documents/tRFMO_MSE_2018_TEXT_final.pdf

Anon. (2022). *MERA questionnaire report for bocinegro, Pagrus pagrus in the Gulf of Cadiz*. Instituto Español de Oceanografía. https://www.merafish.org/mera_library/bocinegro/bocinegro/

Butterworth, D. S., & Punt, A. E. (1999). Experiences in the evaluation and implementation of management procedures. *ICES Journal of Marine Science*, *56*(6), 985–998. https://doi.org/10.1006/jmsc.1999.0532

Carruthers, T. R. (2022). MERA User Guide. https://blue-matter.github.io/openMSE/MERA-User-Guide.html

Carruthers, T. R., & Hordyk, A. R. (2018). The data-limited methods toolkit (DLMtool): An R package for informing management of data-limited populations. *Methods in Ecology and Evolution*, *12*(9), 2388–2395. https://doi.org/10.1111/2041-210X.13081

Carruthers, T. R., & Hordyk, A. R. (2019). Using management strategy evaluation to establish indicators of changing fisheries. *Canadian Journal of Fisheries and Aquatic Sciences*, *76*(9), 1653–1668. https://doi.org/10.1139/cjfas-2018-0223

Carruthers, T. R., Kell, L., Butterworth, D., Maunder, M., Geromont, H., Walters, C., McAllister, M. K., Hillary, R., Levontin, P., Kitakado, T., & Davies, C. (2016). Performance review of simple management procedures. *ICES Journal of Marine Science*, *73*(2), 464–482. https://doi.org/10.1093/icesjms/fsv212

Carruthers, T. R., Punt, A. E., Walters, C. J., MacCall, A., McAllister, M. K., Dick, E. J., & Cope, J. (2014). Evaluating methods for setting catch limits in data-limited fisheries. *Fisheries Research*, *153*, 48–68. https://doi.org/10.1016/j.fishres.2013.12.014

CDFW. (2018). *Master plan for fisheries: A guide for implementation of the marine life management act*. California Department of Fish & Wildlife. https://nrm.dfg.ca.gov/FileHandler.ashx?DocumentID=159222&inline

Costello, C., Ovando, D., Hilborn, R., Gains, S. D., Deschenes, O., & Lester, S. E. (2012). Status and solutions for the world's unassessed fisheries. *Science*, *338*, 517–520. https://doi.org/10.1126/science.1223389

Dowling, N. A., Dichmont, C. M., Haddon, M., Smith, D. C., Smith, A. D. M., & Sainsbury, K. (2015). Empirical harvest strategies for data-poor fisheries: A review of the literature. *Fisheries Research*, *171*, 141–153. https://doi.org/10.1016/j.fishres.2014.11.005

FishPath. 2022. *The nature conservancy*. https://fishpath.org/

Free, C. M., Jensen, O. P., Anderson, S. C., Gutierrez, N. L., Kleisner, K. M., Longo, C., Minto, C., Osio, G. C., & Walsh, J. C. (2020). Blood from a stone: Performance of catch-only methods in estimating stock biomass status. *Fisheries Research*, *223*, 105452. https://doi.org/10.1016/j.fishres.2019.105452

Geromont, H. F., & Butterworth, D. S. (2015). Generic management procedures for data-poor fisheries: Forecasting with few data. *ICES Journal of Marine Science*, *72*(1), 251–261. https://doi.org/10.1093/icesjms/fst232

Goethel, D. R., Lucey, S. M., Berger, A. M., Gaichas, S. K., Karp, M. A., Lynch, P. D., & Walter, J. F. (2019). Recent advances in management strategy evaluation: Introduction to the special issue "under pressure: Addressing fisheries challenges with management strategy evaluation". *Canadian Journal of Fisheries and Aquatic Sciences*, *76*(10), 1689–1696. https://doi.org/10.1139/cjfas-2019-0084

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., & Wilson, J. R. (2021). Multi-indicator harvest strategies for data-limited fisheries. A practitioners guide to learning and design. *Frontiers in Marine Science*, *8*, 757877. https://doi.org/10.3389/fmars.2021.757877

Harford, W. J., & Carruthers, T. R. (2017). Interim and long-term performance of static and adaptive management procedures. *Fisheries Research*, *190*, 84–94. https://doi.org/10.1016/j.fishres.2017.02.003

Hilborn, R. (2002). The dark side of reference points. *Bulletin of Marine Science*, *70*(2), 403–408.

Hobday, A. J., Smith, A. D. M., Stobutzki, I. C., Bulman, C., Daley, R., Dambacher, J. M., Deng, R. A., Dowdney, J., Fuller, M., Furlani, D., Griffiths, S. P., Johnson, D., Kenyon, R., Knuckey, I. A., Ling, S. D., Pitcher, R., Sainsbury, K. J., Sporcic, M., Smith, T., … Zhou, S. (2011). Ecological risk assessment for the effects of fishing. *Fisheries Research*, *108*(2–3), 372–384. https://doi.org/10.1016/j.fishres.2011.01.013

Hordyk, A., Ono, K., Valencia, S., Lonergran, N., & Prince, J. (2015). A novel length-based empirical estimation method of spawning potential ratio (SPR), and tests of its performance, for small-scale, data-poor fisheries. *ICES Journal of Marine Science*, *72*(1), 217–231. https://doi.org/10.1093/icesjms/fsu004

Hordyk, A. R., & Carruthers, T. R. (2018). A quantitative evaluation of a qualitative risk assessment framework: Examining the assumptions and predictions of the Productivity Susceptibility Analysis (PSA). *PLOS ONE*, *13*(6), e0198298. https://doi.org/10.1371/journal.pone.0198298

Hordyk, A.R., Carruthers, T.R., & Huynh, Q. (2022). openMSE: Open-source libraries for management strategy evaluation. https://openmse.com/

Hoshino, E., Hillary, R., Davies, C., Satria, F., Sadiya, L., Ernawati, T., & Proctor, C. (2020). Develoment of pilot empirical harvest strategies for tropical tuna in Indonesian archipelagic waters: Case studies of skipjack and yellowfin tuna. *Fisheries Research*, *227*, 105539. https://doi.org/10.1016/j.fishres.2020.105539

Huynh, Q., Carruthers, T. R., & Hordyk, A. R. (2022). *SAMtool: Stock assessment methods toolkit*. https://cran.r-project.org/web/packages/SAMtool/

Jardim, E., Azevedo, M., & Brites, N. M. (2015). Harvest control rules for data limited stocks using length-based reference points and survey biomass indices. *Fisheries Research*, *171*, 12–19. https://doi.org/10.1016/j.fishres.2014.11.013

Kronlund, A. R., Holt, K. R., Cleary, J. S., & Shelton, P. A. (2013). *Current approaches for the provision of scientific advice on the precautionary approach for Canadian fish stocks: Sections 8 – Management strategy evaluation*. Canadian Science Advisory Secretariat. Research Document - 2013/081. http://www.dfo-mpo.gc.ca/csas-sccs/publications/resdocs-docrech/2013/2013_081-eng.html

Legg, C. J., & Nagy, L. (2006). Why most conservation monitoring is, but need not be, a waste of time. *Journal of Environmental Management*, *78*(2), 194–199. https://doi.org/10.1016/j.jenvman.2005.04.016

Loneragan, N. R., Wiryawan, B., Hordyk, A. R., Halim, A., Proctor, C., Satria, F., & Yulianto, I. (2021). *Proceedings from workshops on management strategy evaluation of data-limited fisheries: Towards sustainability – Applying the method evaluation and risk assessment tool to seven Indonesian fisheries*. Murdoch University, Western Australia, and IPB University, Indonesia. https://researchrepository.murdoch.edu.au/id/eprint/60810/

MacCall, A. (2009). Depletion-corrected average catch: A simple formula for estimating yields in data-poor situations. *ICES Journal of Marine Science*, *66*, 2267–2271. https://doi.org/10.1093/icesjms/fsp209

Martin, S. M., Cambridge, T. A., Grieve, C., Nimmo, F. M., & Agnew, D. J. (2012). An evaluation of environmental changes within fisheries involved in the marine stewardship council certification scheme. *Reviews in Fisheries Science*, *20*(2), 61–69. https://doi.org/10.1080/10641262.2011.654287

Mees, C. C. (2007). Managing fishing effort in multispecies fisheries. In D. D. Hoggarth, S. Abeyasakera, R. Arthur, J. R. Beddington, R. W. Burn, A. S. Halls, G. P. Kirkwood, M. Mcallister, P. Medly, C. C. Mees, G. B. Parkes, G. M. Pilling, R. C. Wakeford, & R. L. Welcomme (Eds.), *Stock assessment for fishery management – A framework guide to the use of the fish stock assessment tools of the fisheries management science Programme* (Vol. *487*, pp. 189–194). FAO Fisheries Technical Paper.

MERA. (2022). *Method evaluation and risk assessment*. Version 5.4. www.merafish.org

MSC. (2018). MSC fisheries certification process version 2.01, marine stewardship council. https://www.msc.org/docs/default-source/default-document-library/for-business/program-documents/fisheries-program-documents/msc-fisheries-certification-process-v2.1.pdf?sfvrsn=5c8c80bc_24

Newman, D., Berkson, J., & Suatoni, L. (2015). Current methods for setting catch limits for data-limited fish stocks in the United States. *Fisheries Research*, *164*, 86–93. https://doi.org/10.1016/j.fishres.2014.10.018

NOAA. (2019). *Integrated ecosystem assessment*. U.S. National Oceanic and Atmospheric Administration. https://www.integratedecosystemassessment.noaa.gov/publications

Ovando, D., Hilborn, R., Monnahan, C., Rudd, M., Sharma, R., Thorson, J. T., Rousseau, Y., & Ye, Y. (2021). Improving estimates of the state of global fisheries depends on better data. *Fish and Fisheries*, *22*, 1377–1391. https://doi.org/10.1111/faf.12593

Pilling, G. M., Apostolaki, P., Failler, P., Floros, C., Large, P. A., Morales-Nin, B., Reglero, P., Stergiou, K. I., & Tsikliras, C. (2008). Assessment and management of data-poor fisheries. In A. Payne, J. Cotter, & T. Potter (Eds.), *Advances in fisheries science: 50 years on from Beverton and Holt* (pp. 280–305). Blackwell Publishing. https://doi.org/10.1002/9781444302653.ch12

Prince, J., & Hordyk, A. R. (2019). What to do when you have almost nothing: A simple quantitative prescription for managing extremely data-poor fisheries. *Fish Fish*, *20*, 224–238. https://doi.org/10.1111/faf.12335

Punt, A. E., Butterworth, D. S., de Moor, C. L., De Oliveira, J. A. A., & Haddon, M. (2016). Management strategy evaluation: Best practices. *Fish and Fisheries*, *17*, 303–334. https://doi.org/10.1111/faf.12104

Rochet, M.-J., & Rice, J. C. (2009). Simulation-based management strategy evaluation: Ignorance disguised as mathematics? *ICES Journal of Marine Science*, *66*, 754–762. https://doi.org/10.1093/icesjms/fsp023

Rosenberg, A. A., Kleisner, K. M., Afflerbach, J., Anderson, S. C., Dickey-Collas, M., Cooper, A. B., Fogarty, M. J., Fulton, E. A., Gutiérrez, N. L., Hyde, K. J., Jardim, E., Jensen, O. P., Kristiansen, T., Longo, C., Minte-Vera, C. V., Minto, C., Mosqueira, I., Osio, G. C., Ovando, D., … Ye, Y. (2018). Applying a new ensemble approach to estimating stock status of marine fisheries around the world. *Conservation Letters*, *11*, e12363. https://doi.org/10.1111/conl.12363

Ruddle, K. (1996). Traditional management of reef fishing. In N. V. C. Polunin & C. M. Roberts (Eds.), *Reef fisheries* (pp. 315–335). Chapman and Hall.

SeaFood Watch. (2016). *Seafood watch standard for fisheries*. Monterey Bay Aquarium Seafood Watch. https://www.seafoodwatch.org/recommendations/our-standards

Thorson, J. T., Munch, S. B., Cope, J. M., & Gao, J. (2017). Predicting life history parameters for all fishes worldwide. *Ecological Applications*, *27*(8), 2262–2276. https://doi.org/10.1002/eap.1606

Walters, C. J., & Martell, S. J. D. (2004). *Fisheries ecology and management* (p. 448). Princeton University Press.

How to cite this article: Carruthers, T. R., Huynh, Q. C., Hordyk, A. R., Newman, D., Smith, A. D. M., Sainsbury, K. J., Stokes, K., Morison, A., Agnew, D., Parma, A., Sobrino, I., & Longo, C. (2023). Method evaluation and risk assessment: A framework for evaluating management strategies for data-limited fisheries. *Fish and Fisheries*, *24*, 279–296. https://doi.org/10.1111/faf.12726

APPENDIX A

THE QUANTITATIVE QUESTIONNAIRE

TABLE A1 Quantitative fishery questions regarding population and exploitation dynamics

Fishery question		Answers (multiple choice)	Operating model parameter values
1. Fishery description	Provide an overview of the resource including references to supporting information	Name the species, geographical location, management authority and duration of management	–
2. Longevity	What is the maximum age (A) of the species?	Very short-lived Short-lived Moderate life span Moderately long-lived Long-lived Very long-lived	$5 < A < 7$ $7 < A < 10$ $10 < A < 20$ $20 < A < 40$ $40 < A < 80$ $80 < A < 160$
3. Stock depletion	What is the status of spawning stock biomass compared to 'unfished levels' (D)	Crashed Very depleted Depleted Moderately depleted Healthy Underexploited	$0.01 < D < 0.05$ $0.05 < D < 0.1$ $0.1 < D < 0.15$ $0.15 < D < 0.3$ $0.3 < D < 0.5$ $0.5 < D < 0.8$
4. Resilience	What fraction of unfished recruitment occurs at 20% of unfished spawning stock biomass (h)	Not resilient Low resilience Moderate resilience Resilient Very resilient	$0.25 < h < 0.3$ $0.3 < h < 0.5$ $0.5 < h < 0.7$ $0.7 < h < 0.9$ $0.9 < h < 0.99$
5. Historical effort pattern	How has fishing intensity varied historically (e.g. annual days of fishing)?	Stable Two-phase Boom-bust Gradual increases Stable, recent increases Stable, recent declines	Adjustable skew, magnitude of recent changes and time-series truncation
6. Inter-annual variability in historical effort	What is the magnitude of inter-annual changes in fishing effort (σ_E) among years?	Not variable Variable Highly variable	$10\% < \sigma_E < 20\%$ $20\% < \sigma_E < 50\%$ $50\% < \sigma_E < 100\%$
7. Historical fishing efficiency changes	What percentage change in fishing efficiency (Δ_h) can be expected over previous years	Strongly declining Declining Stable Increasing Strongly increasing	$-3\% < \Delta_h < -2\%$ $-2\% < \Delta_h < -1\%$ $-1\% < \Delta_h < 1\%$ $1\% < \Delta_h < 2\%$ $2\% < \Delta_h < 3\%$
8. Future fishing efficiency changes	What percentage change in fishing efficiency (Δ_f) can be expected over future years	Strongly declining Declining Stable Increasing Strongly increasing	$-3\% < \Delta_f < -2\%$ $-2\% < \Delta_f < -1\%$ $-1\% < \Delta_f < 1\%$ $1\% < \Delta_f < 2\%$ $2\% < \Delta_f < 3\%$
9. Length at maturity	At what fraction of asymptotic length (L_M) can 50% of fish be assumed to be sexually mature?	Very small Small Moderate Moderate to large Large	$0.4 < L_M < 0.5$ $0.5 < L_M < 0.6$ $0.6 < L_M < 0.7$ $0.7 < L_M < 0.8$ $0.8 < L_M < 0.9$
10. Selectivity of small fish	Relative to asymptotic length, at what size do fish first become 50% vulnerable to fishing (S)?	Very small Small Half asymptotic length Large Very large	$0.1 < S < 0.2$ $0.2 < S < 0.4$ $0.4 < S < 0.6$ $0.6 < S < 0.8$ $0.8 < S < 0.9$

(Continues)

TABLE A1 (Continued)

Fishery question		Answers (multiple choice)	Operating model parameter values
11. Selectivity of large fish	What is the selectivity of fish of asymptotic length (S_L)?	Asymptotic selectivity Declining selectivity Dome-shaped selectivity Strong dome-shape	$S_L = 1$ $0.75 < S_L < 1$ $0.25 < S_L < 0.75$ $0 < S_L < 0.25$
12. Discard rate	Of the fish that are caught, what fraction are discarded (F_D)?	Low Low – moderate Moderate Moderate – high High	$0 < F_D < 1\%$ $1\% < F_D < 10\%$ $10\% < F_D < 30\%$ $30\% < F_D < 50\%$ $50\% < F_D < 70\%$
13. Post-release mortality rate	Of the fish that are discarded, what fraction die due to capture (F_R)?	Low Low – moderate Moderate Moderate – high High Almost all die	$0 < F_R < 5\%$ $5\% < F_R < 25\%$ $25\% < F_R < 50\%$ $50\% < F_R < 75\%$ $75\% < F_R < 95\%$ $95\% < F_R < 100\%$
14. Recruitment variability	What is the magnitude of inter-annual changes in recruitment (σ_R)	Very low Low Moderate High Very high	$10\% < \sigma_R < 20\%$ $20\% < \sigma_R < 60\%$ $60\% < \sigma_R < 120\%$ $120\% < \sigma_R < 180\%$ $180\% < \sigma_R < 240\%$
15. Size of existing spatial closures	What percentage of the species habitat is included in existing marine spatial closures (r_h)?	None Small Small-moderate Moderate Large Very large Huge	$r_h = 0$ $0 < r_h < 5\%$ $5\% < r_h < 10\%$ $10\% < r_h < 20\%$ $20\% < r_h < 30\%$ $30\% < r_h < 40\%$ $40\% < r_h < 50\%$
16. Spatial mixing in/out of existing spatial closures	Among years, what fraction of fish leave the spatial closure and enter the fished area (P_h)?	Very low Low Moderate High Fully mixed	$0 < P_h < 1\%$ $1\% < P_h < 5\%$ $5\% < P_h < 10\%$ $10\% < P_h < 20\%$ $20\% < P_h < 50\%$
17. Size of future spatial closures	What percentage of the species habitat is included in proposed future marine spatial closures (r_f)?	None Small Small-moderate Moderate Large Very large Huge	$r_f = 0$ $0 < r_f < 5\%$ $5\% < r_f < 10\%$ $10\% < r_f < 20\%$ $20\% < r_f < 30\%$ $30\% < r_f < 40\%$ $40\% < r_f < 50\%$
18. Spatial mixing in/out of future spatial closures	Among years, what fraction of fish are expected to leave the spatial closure and enter the fished area (P_f)?	Very low Low Moderate High Fully mixed	$0 < P_f < 1\%$ $1\% < P_f < 5\%$ $5\% < P_f < 10\%$ $10\% < P_f < 20\%$ $20\% < P_f < 50\%$
19. Initial stock depletion	At the start of the historical time series, what was the stock level as a fraction of theoretical unfished stock size (D_1)	Very low Low Moderate High Asymptotic unfished	$0.1 < D_1 < 0.15$ $0.15 < D_1 < 0.3$ $0.3 < D_1 < 0.5$ $0.5 < D_1 < 1$ $D_1 = 1$

TABLE A2 Quantitative questions regarding the type of management system and its implementation

Management question		Answers (multiple choice)	Operating model parameter values
1. Type of fishery management that is possible	Can fishery exploitation be controlled by measure such as Total annual catches (TAC), Total annual effort (TAE).	TAC TAE Size limit Time-area closures	–
2. TAC offset	What fraction (F_C) of recommended catches are taken by the fishery	Large underages Underages Slight underages Taken exactly Slight overages Overages Large overages	$40\% < F_C < 70\%$ $70\% < F_C < 90\%$ $90\% < F_C < 100\%$ $95\% < F_C < 105\%$ $100\% < F_C < 110\%$ $110\% < F_C < 150\%$ $150\% < F_C < 200\%$
3. TAC implementation variability	Given the offset between catch recommendations and catches of the fishery what is the maximum annual deviation (d_C) from this offset?	Constant Not variable Low variability Variable High variable	$0 < d_C < 1\%$ $1\% < d_C < 5\%$ $5\% < d_C < 10\%$ $10\% < d_C < 20\%$ $20\% < d_C < 40\%$
4. TAE offset	What fraction (F_E) of recommended catches are taken by the fishery	Large underages Underages Slight underages Taken exactly Slight overages Overages Large overages	$40\% < F_E < 70\%$ $70\% < F_E < 90\%$ $90\% < F_E < 100\%$ $95\% < F_E < 105\%$ $100\% < F_E < 110\%$ $110\% < F_E < 150\%$ $150\% < F_E < 200\%$
5. TAE implementation variability	Given the offset between effort recommendations and effort of the fishery what is the maximum annual deviation (d_E) from this offset?	Constant Not variable Low variability Variable High variable	$0 < d_E < 1\%$ $1\% < d_E < 5\%$ $5\% < d_E < 10\%$ $10\% < d_E < 20\%$ $20\% < d_E < 40\%$
6. Size limit offset	What fraction of a recommended minimum size limit (F_S) is taken by the fishery	Much smaller Smaller Slightly smaller Taken exactly Slightly larger Larger Much larger	$40\% < F_S < 70\%$ $70\% < F_S < 90\%$ $90\% < F_S < 100\%$ $95\% < F_S < 105\%$ $100\% < F_S < 110\%$ $110\% < F_S < 150\%$ $150\% < F_S < 200\%$
7. Size limit implementation variability	Given the offset between recommended minimum size limits and the minimum size that is taken, what is the maximum deviation (d_S) from this offset?	Constant Not variable Low variability Variable High variable	$0 < d_S < 1\%$ $1\% < d_S < 5\%$ $5\% < d_S < 10\%$ $10\% < d_S < 20\%$ $20\% < d_S < 40\%$

TABLE A3 Data questions regarding the types of data that are available and the quality of these data

Data question		Answers (multiple choice)	Operating model parameter values
1. Types of data that are available	What data types are collected and processed for making management recommendations using management procedures?	Historical annual catches Recent annual catches Historical abundance index Recent abundance index Fishing effort Size composition data Age composition data Growth Absolute biomass survey	–
2. Catch reporting bias	What is the % difference between the catches reported and those taken (θ_C)	Strong under-reporting Under-reporting Slight under-reporting Reported accurately Slight over-reporting	$-50\% < \theta_C < -30\%$ $-30\% < \theta_C < -10\%$ $-10\% < \theta_C < 0$ $-5\% < \theta_C < 5\%$ $0 < \theta_C < 10\%$
3. Hyperstability in indices	How linear is the relationship between the index *I*, and the abundance *A*, where $I \propto A^{\beta}$	Strong hyperdepletion Hyperdepletion Proportional Hyperstability Strong hyperstability	$0 < \beta < 1\%$ $1\% < \beta < 5\%$ $5\% < \beta < 10\%$ $10\% < \beta < 20\%$ $20\% < \beta < 40\%$
4. Overall data quality	How extensive, accurate and precise are other aspects of the data collection?	Perfect Good (accurate/precise) Data moderate Data poor (inaccurate/ imprecise)	See Table B1 for a detailed description of these observation error models

APPENDIX B

OPERATING MODEL ASSUMPTIONS

TABLE B1 Observation model types. Ranges provided are uniform random variables, with one draw per simulation. For example, for a simulation, a value of 0.011 may be drawn for the lognormal standard deviation of catch observation error from the uniform random variable in the range of zero to 0.05. All annual observation errors for this simulation are then sampled from a lognormal distribution with this standard deviation. For further details on the observation error model used in MERA see Carruthers and Hordyk (2018)

		Overall data quality (answers to data question #4)			
Quantity	Type of error	Perfect	Good	Data moderate	Data poor
Error in annual catch observations	Lognormal SD	U[0, 0.05]	U[0.1, 0.2]	U[0.1, 0.3]	U[0.2, 0.6]
Error in annual relative abundance index observations	Lognormal SD	U[0, 0.05]	U[0.1, 0.25]	U[0.1, 0.4]	U[0.2, 0.6]
Number of annual catch at age samples	Multinomial sample size	U[500, 900]	U[50, 100]	U[25, 50]	U[10, 20]
Number of annual catch at length samples	Multinomial sample size	U[500, 900]	U[50, 100]	U[25, 50]	U[10, 20]
Biases in estimates of natural mortality rate	Lognormal SD	0.01	0.05	0.2	0.4
Error in annual estimates of current stock status	Lognormal SD	U[0, 0.05]	U[0.025, 0.1]	U[0.05, 0.1]	U[0.05, 0.2]
Biases in estimates of current stock status	Lognormal SD	0.01	0.2	0.5	0.75