AN UPDATED ROADMAP FOR MSE DEVELOPMENT

T. R. Carruthers[1]

SUMMARY

An updated roadmap for the development of management strategy evaluation frameworks is presented that includes important features missing from the previous version (Carruthers, 2024). Those features include establishing the intended duration for MP usage, how often the MP will provide updated advice, identifying lags between data collection and MP usage, the need for independent code review and the weighting of operating models. The aim of the roadmap is to provide the participants of an MSE with a concise path to the adoption of an MP in which processes, products, and roles are clearly defined. The roadmap is intended to be comprehensive and inclusive of new MSE processes where, for example, managers are not yet familiar with MSE terminology, concepts, and procedures and may not yet have explicit performance objectives.

RÉSUMÉ

Une feuille de route actualisée pour le développement de cadres d'évaluation de la stratégie de gestion est présentée, incluant d'importantes caractéristiques manquantes dans la version précédente (Carruthers, 2024). Ces caractéristiques incluent l'établissement de la durée prévue de l'utilisation de la MP, la fréquence à laquelle la MP fournira un avis actualisé, l'identification des décalages entre la collecte des données et l'utilisation de la MP, la nécessité d'une révision indépendante du code et la pondération des modèles opérationnels. L'objectif de cette feuille de route est de fournir aux participants d'une MSE un chemin concis vers l'adoption d'une MP dans laquelle les processus, les produits et les rôles sont clairement définis. La feuille de route se veut exhaustive et inclura les nouveaux processus de la MSE dans le cadre desquels, par exemple, les gestionnaires ne sont pas encore familiarisés avec la terminologie, les concepts et les procédures de la MSE et n'ont peut-être pas encore d'objectifs de performance explicites.

RESUMEN

Se presenta una hoja de ruta actualizada para el desarrollo de marcos de evaluación de estrategias de ordenación que incluye características importantes que faltaban en la versión anterior (Carruthers, 2024). Estas características incluyen el establecimiento de la duración prevista para el uso del MP, la frecuencia con la que el MP proporcionará asesoramiento actualizado, la identificación de desfases entre la recopilación de datos y el uso del MP, la necesidad de una revisión independiente del código y la ponderación de los modelos operativos. El objetivo de la hoja de ruta es facilitar a los participantes una MSE con un camino conciso para la adopción de un MP en el que los procesos, productos y roles estén definidos de manera clara. La hoja de ruta pretende ser exhaustiva y estar destinada a los nuevos procesos de MSE en los que, por ejemplo, los gestores aún no están familiarizados con la terminología, los conceptos y los procedimientos de la MSE y puede que aún no tengan objetivos de desempeño explícitos.

---

[1] Blue Matter Science Ltd. 2150 Bridgman Ave, N. Vancouver. V7P 2T9

1. Introduction

The development of management strategy evaluation (MSE) frameworks varies strongly among stocks. Some frameworks such as that for Atlantic bluefin tuna, focus on stocks of high economic importance that are of general interest to a wide range of stakeholders, involve many nations, require the collaboration of numerous scientists, involve many data sources, a large number of possible system uncertainties and must characterize relatively complex fishery and population dynamics (Carruthers 2020). These intensive multi-year MSE processes strongly contrast with regional MSEs such as B.C. groundfish for which a management procedure can be adopted in a matter of months (Haggarty *et al*. 2022a).

Although the development of MSE frameworks can vary widely in their scope and demands, all share various fundamental components (Punt *et al*. 2016). All require the design of multiple candidate management procedures (CMPs), consider multiple performance metrics and evaluate robustness of CMPs against multiple operating models. Most identify primary (reference set) and secondary (robustness set) uncertainties. MSE processes necessarily require the input of participants diverse in their expertise, interest and experience including managers, policy makers, stakeholders, scientists and technical resources involved in coding and computation. In new MSE processes, it is important to lay the groundwork by clearly identifying the problem statement and providing introductory workshops and materials to ensure all participants have a shared understanding of concepts and terminology. Generally, all MSE processes require refinement of operating models, management procedures and tools for communicating results to stakeholders and managers. There is also a necessary order to the steps of MSE framework development in the scoping of uncertainties, the gathering and processing of data, the specification of operating models and the refinement of CMPs. It follows that there is a common set of MSE phases, tasks and processes.

Here, a MSE roadmap is proposed in which MSE development is organized by phase, task and process. The aim is to provide MSE participants with a concise path from status quo management to the adoption of an MP in which processes, products, and roles are clearly defined. In doing so, the intention is to maintain discipline and ensure efficient progress, and efficient use of participants time while avoiding unnecessary back tracking by imposing data guillotines. The roadmap is intended to be comprehensive and inclusive of new MSE processes where, for example, there may be a need for introductory workshops on MSE concepts and terminology, scientists may need to see working straw-dog MSE frameworks to maximise the benefit of their feedback, and managers may need an opportunity to refine management objectives as realistic MP performance is revealed. It follows that in management contexts where participants are familiar with MSE, the roadmap can be simplified and where appropriate, processes may be dropped if they are considered to be unnecessary.

Here an updated roadmap for the development of management strategy evaluation frameworks is presented that includes important features missing from the previous version (SCRS/2024/103).

2. Methods

Based on experience in the development of MSE frameworks for California state fisheries, BC groundfish (Haggarty *et al*. 2022a;b), Chilean pelagic stocks, western Atlantic skipjack tuna (Huynh *et al*. 2020), Atlantic bluefin tuna (Carruthers 2020), North Atlantic swordfish (Hordyk *et al*. 2021) and Bay of Fundy herring (Carruthers *et al*. 2022), a comprehensive set of processes in the development of MSE frameworks were identified and then organized in a sequence of tasks nested in phases. These were represented in a schematic where the order of processes follows the reading of text - across then down. The roadmap was designed to fit on a single page or be split into two for the single slide of a landscape presentation. All components of the roadmap are described in an accompanying descriptive table.

The updated roadmap and supporting documentation include the following changes:

- establishing the intended duration for MP usage (**Table 1**, Foundation phase – Dialogue meeting),
- how often the MP with provide updated advice (**Table 1**, Foundation phase – Dialogue meeting),
- establishing the lag between data collection and use of the MP to calculate management advice (**Table 1**, Foundation phase - Data prep. for OM conditioning – Review meta data)
- independent code review (**Figure 1** & **Table 1**, Initial phase – Technical milestone 1) and
- weighting of operating models (**Figure 1** and **Table 1**, Revision phase – Revise / simplify / weight OMs).

3. Results

The updated roadmap is presented in **Figure 1**. The description of roadmap phases, tasks and processes is included in **Table 1**. A glossary of terms in available in **Table 2**.

References

Carruthers T.R. 2020. Reference Set Operating models for Atlantic Bluefin Tuna assuming priors for Area-Specific Scale and western Stock Mixing. SCRS/2020/018. Retrieved from https://www.iccat.int/Documents/CVSP/CV077_2020/n_2/CV077020078.pdf [August 2022]

Carruthers T.R., Hordyk A.R., Huynh Q.C., Singh R., and Barrett T.J. 2022. A Framework for Conditioning Operating Models for the Southwest Nova Scotia/Bay of Fundy Spawning Component of 4VWX Herring. Can. Sci. Advis. Sec. Res. Doc. 2022/nnn. https://www.dfo-mpo.gc.ca/csas-sccs/Schedule-Horraire/2022/02_16-17-eng.html

Haggarty D.R., Huynh Q.C., Forrest R.E., Anderson S.C., Bresch M.J., and Keppell E.A. 2022a. Evaluation of potential rebuilding strategies for Inside Yelloweye Rockfish (*Sebastes ruberrimus*) in British Columbia. DFO Canadian Science Advisory Secretariat Research Document

Haggarty D.R., Siegel M.R., Litt MA., Huynh Q. 2022b. Quillback Rockfish Fishery and Conservation Objectives Workshop Summary Report. https://waves-vagues.dfo-mpo.gc.ca/library-bibliotheque/41059116.pdf

Hordyk A., Schirripa M., Rosa D. 2021. Updates to the operating models uncertainty grid for the North Atlantic swordfish MSE. SCRS/2021/099. https://iccat.int/Documents/CVSP/CV078_2021/n_7/CV078070240.pdf

Huynh Q.C., Carruthers T., Mourato B., Sant'Ana R., Cardoso L.G., Travassos P., and Hazin, F. 2020. Demonstration of a MSE framework for western Skipjack tuna, including operating model conditioning. SCRS/2020/140. Collect. Vol. Sci. Pap. ICCAT 77:121-144

Punt A.E., Butterworth D.S., de Moor C.L., De Oliveira J.A.A., and Haddon M. 2016. Management strategy evaluation: Best practices. Fish Fish. 17(2): 303–334. doi:10.1111/faf.12104

Table 1. A description of the phases, tasks and processes of the preliminary roadmap (**Figure 1**).

PHASE			Description
Task			
		Process	
PREREQUISITE PHASE			**Identifies the background to the problem, introduces MSE concepts and established membership to various working groups.**
	Identify Purpose		**The reason(s) why MSE is being pursued.**
	A1i	Problem Statement	What problem is being addressed in the application of MSE? For example: difficulty in developing a conventional stock assessment that can pass peer-review, requirement for simulation tested MPs, a need to demonstrate robustness to climate scenarios or conflicting views on how data should be used to inform management.
	Establish concepts and terminology		**Develop a common understanding of what MSE is, how it functions, the problems it may address and the associated terminology such that potential participants can understand their potential roles and interests in the process.**
	A2i	Introductory workshop(s)	Presentations by MSE experts intended for general audience including managers, stakeholders and scientists.
	A2ii	Online materials	Online documentation, presentations, explanatory videos and links in support of MSE introductory workshops.
	Identify membership		**Assign participants to at least one group such that processes can be organized by specialization to increase efficiency.**
	A3i	Working group	The wider MSE working group that meets regularly to discuss general aspects of MSE framework development which includes technical members, those involved in oversight and communications but does not include managers and stakeholders.
	A3ii	Technical	Participants with technical skills relating to data collection, processing, mathematical modelling, statistics, MP development or MSE coding. The technical group may spend dedicated time informally discussing quantitative approaches (e.g., solving mathematical equations, testing code, developing appropriate models or statistics).
	A3iii	Oversight / facilitation	Member of this group coordinate resources, set deliverable dates, and chair meetings. They are responsible keeping the MSE process on track.
	A3iv	Communications	The communications team focuses on developing processes, tools and materials that explain MSE concepts, methodologies, results to participants outside of the working group, particularly stakeholders and resource managers.
	A3v	Managers / stakeholders	The manager / stakeholder group includes participants that aim to use the outputs of the MSE for decision making or planning but do not have a direct role in the day-to-day development of the MSE framework.

FOUNDATION PHASE		Establishes the context for the MSE including the broad management objectives, key system uncertainties for OMs and investigation of available data and information sources to support operating model specification.
Qualitative objectives		**For example, maintain a stable productive stock biomass and fishery yields without substantial risk from overfishing.**
B1i	Review legal / comparable stocks	What legal frameworks are applicable? Objectives that have been established in similar fisheries that may be applicable. What objectives are laid out in relevant management guidelines?
B1ii	Dialogue meeting	Engage with managers and stakeholder to identify broad categories of objectives such as short-term yield, long-term yield, catch stability, probability of overfishing etc. Performance of management procedures can be considered one of the three principal design axes of MSE (operating model uncertainties and management procedures being the other two). At this stage it is desirable to establish the desired duration for MP usage (typically 5 – 8 years) and the update interval (how often the MP will be used to provide new advice, typically 1-2 years).
Outline system uncertainties		**What uncertainties in the fishery system should a management procedure be robust to?**
B2i	Establish primary uncertainties	These are typically those of particular relevance to the stock in question (for example a conflict among data inputs, or historical uncertainty over productivity regime) in addition to the typical important uncertainties in stock assessments that tend to strongly impact estimates of stock status, productivity and exploitation level such as stock resilience and natural mortality. Primary uncertainties may often be used to determine the reference set of operating models that are the primary basis for the comparative testing of CMPs
B2ii	Identify secondary uncertainties	Secondary uncertainties could include those with weaker empirical support, or scenarios for future fishery or population conditions that may occur but are not predictable. In some settings secondary uncertainties could include the hypotheses of subject area experts or stakeholders. Secondary uncertainties are often used to specify the robustness set of operating models that is used as an additional basis for discriminating among CMPs that perform similarly for the reference set of OMs.
B2iii	First cut TSD	The trial specifications document serves as a central reference for the MSE framework and comprehensively documents its specification to ensure reproducibility. The TSD includes descriptions of data, management objectives, OM specification, OM equations, OM model fitting, performance metrics, MP tuning, the simulation of projected data and exceptional circumstances documents.
Data prep. for OM conditioning		**Raw data that may inform the structure and parameters of the operating models are processed and organized to determine what aspects of operating models can be informed empirically.**
B3i	Review meta data	A meta data summary describes the various data types that are available, the time period over which they are available and their relative quantity / quality. An important aspect of the meta data review is identifying the future lags between data collection and MP implementation (i.e., when an MP is used in year y, for calculating a TAC in y +1, index data are only available up to and including y – 2).

	B3ii	Present data	It is important that the working group is provided with a comprehensive description of data sources so that the empirical basis for operating model development and calculation of advice by MPs is well understood.
	B3iii	Accept Initial data set for OM conditioning	Once organized and made available to the technical team for specification / conditioning of OMs, the first OM data guillotine is passed.
	First OM data guillotine		**Data guillotines are intended to maintain discipline in the development of operating models. If data are continually updated, it is not clear when the process of operating model specification and conditioning is complete for the next step.**
INITIAL PHASE			**This phase completes the first working MSE framework, finalizing the operating models and providing example MP projections.**
	Propose Initial OM sets		**Operating models represent plausible states of nature for fishery and population dynamics and in the context of MSE constitute 'what if' scenarios for robustness testing of CMPs. In this regard operating model projections are not forecasts but rather stress tests analogous to the simulated flight conditions used to train and evaluate prospective pilots.**
	C1i	Reference Case	The reference case operating model is a single model that is used a basis for demonstration and learning among working group members. Its specification, conditioning and properties are well understood and it can be therefore used as a suitable basis for examining alternative OM assumptions, parameterizations, data weighting, and other sensitivity analyses.
	C1ii	Reference Set	The reference set of operating models include primary sources of uncertainty and are the principal basis for the comparative evaluation of CMPs.
	C1iii	Robustness Set	The robustness set of operating models consists of secondary uncertainties that may be used to further discriminate among CMPs that perform similarly in the reference set. The robustness set provides a powerful basis for developing MPs that are robust to future scenarios and hypotheses that may have weaker empirical support but are nonetheless important to managers and stakeholders such as climate change or productivity.
	Technical milestone 1		**The production of a fully functional MSE software package.**
	C2i	Condition initial reference set	The operating models of the reference set are fitted to data to characterize fishery and population dynamics for the purposes of testing CMPs in closed-loop simulation in projections
	C2ii	Develop projection model	The MSE framework must be coded to project simulated conditions including the generation of future simulated data (observation error model), the calculation of advice by CMPs, the implementation of that advice in the fishery system (implementation model) and the subsequent exploitation of the stock.
	C2iii	Develop reference MP	A simple reference MP (for example maintaining current exploitation rate, or catch as a constant proportion of a relative abundance index) allows for the demonstration of the completed MSE software package.

C2iv	Code review		Before MSE results can be interpreted safely, it is important to conduct an independent code review whereby an external expert checks that the code for the operating models, candidate MPs and performance metrics is error free and matches the trial specifications document.
Presentation of initial MSE results			**Working group members are provided with the results of operating model fitting, operating fishing and population dynamics and an example projection. This critical step provides the first opportunity for the working group to appreciate and comment on the empirical support for the simulated dynamics, reference points and tangible results in the form of projected population outcomes and simulated data.**
C3i	OM fit		A description of the empirical plausibility of operating models based on the consistency between operating model predictions and observed data.
C3ii	Projection of reference MP		Demonstration of the interaction of a management procedure within the closed-loop simulation of the MSE software package.
C3iii	Plausible outcomes		The numerical outcomes of management procedure projections that often includes quantities such as simulated biomass, abundance, yield, exploitation rate and simulated data. This process allows working group members an early appreciation of the types of outcomes that may be realistic given the operating models (e.g., the approximate trade-off between yield and conservation outcomes).
C3iv	Dialogue meeting		This dialogue meeting provides the first report on MSE progress that includes arguably the most important methodological step (technical milestone 1 and a demonstration of the first management procedure projections). This meeting provides managers with their first appreciation of plausible projection outcomes that can be used to further discuss quantitative management performance metrics.
Finalize OMs			**Reference case, reference set and robustness set operating models are finalized and will only be revised in light of formal review that may be brought forward by exceptional circumstances protocols.**
C4i	Accept final dataset for OM conditioning		Input data used for operating model conditioning are finalized.
C4ii	Recondition OMs		Final versions of operating models are fitted to data.
Second OM data guillotine			**Operating models are now finalized and will not be revised by any newly available data.**
REVISION PHASE			**The MSE framework is revised to include bespoke MPs and reduce or weight the set of operating models if appropriate.**
Identify possible MP archetypes			**Depending on input data, availability of methods/expertise, and the management levers (e.g., catch limits, effort limits, size limits) various discrete classes of management procedure may be identified (e.g. empirical index target - TAC, model-based - TAC, model-based - TAC / size limit)**

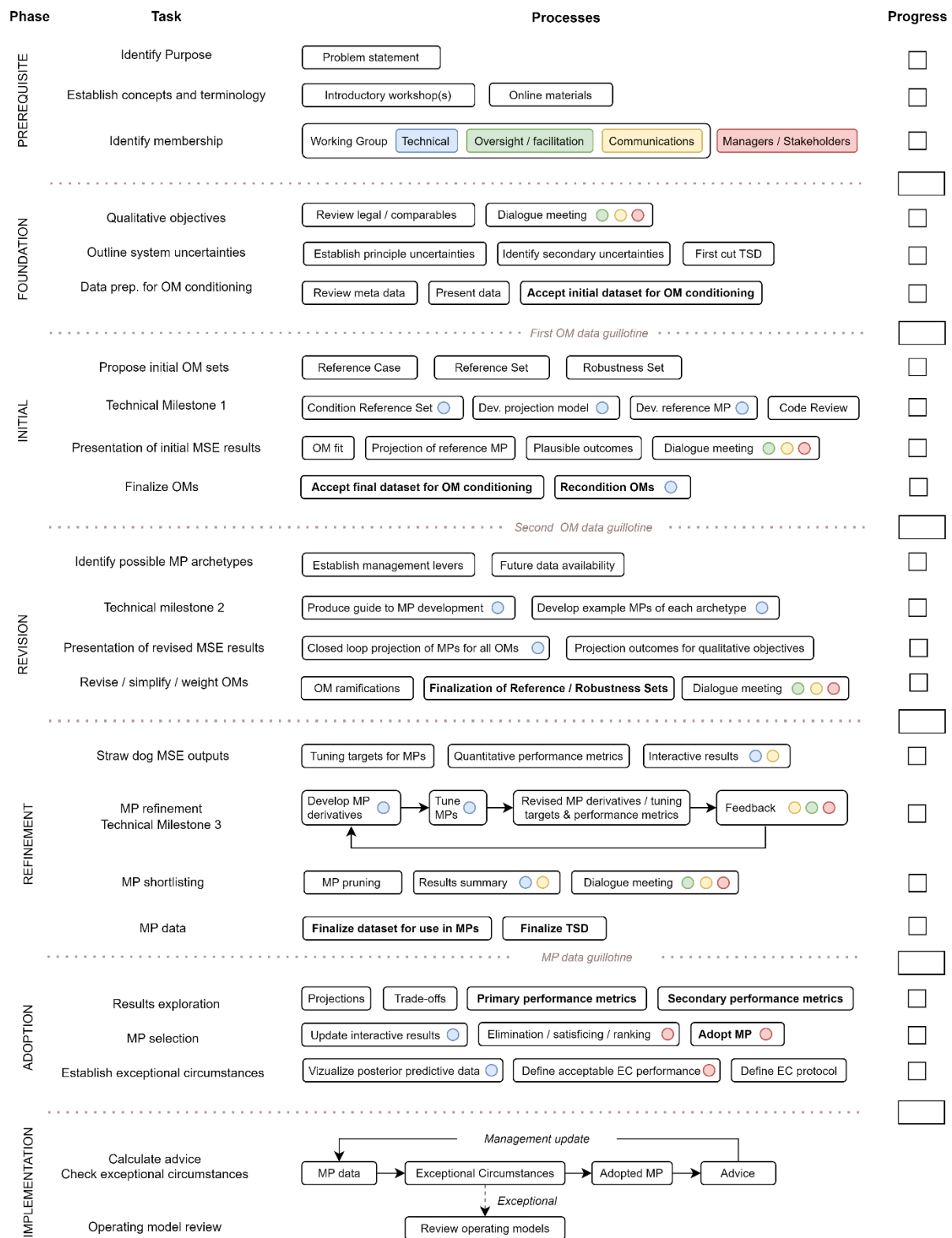|			
	D1i	Establish management levers	Candidate management procedures are limited to those types of management advice that are appropriate / possible (catch limits, spatio-temporal closures, gear restrictions, effort limits etc.).
	D2ii	Future data availability	Candidate management procedures are limited to use only those data that are currently available and thought to be available in the future (e.g., catches, relative abundance indices, survey indices, age composition data).
	Technical milestone 2		**A documented MSE software package accessible to CMP developers that may span a range of technical ability.**
	D3i	Produce guide to MP development	A concise guide to MP development and coding that allows for participation in CMP development, including CMP testing, refinement and tuning.
	D3ii	Develop example MPs for of each archetype	Codify examples of the various archetypes identified above to provide an example / template for CMP developers.
	Presentation of revised MSE results		**The working group are provided with updated projection results following operating model finalization in the previous phase.**
	D4i	Closed-loop projection of MPs for all OMs	Conduct an initial projection of all CMPs against all operating models and present results to the working group to outline the plausible range of performance outcomes, for informing CMP refinement and tuning targets.
	D4ii	Projection outcomes for qualitative objectives	Present a range of CMP outcomes for various interpretations of the qualitative objectives outlines above. For example, for the qualitative objective 'maximize yield' this could include quantitative metrics such as short-term mean yield over the first 10 projected years and long-term mean projected yield over the final 10 years of the projection.
	Revise / simplify / weight OMs		**Operating models may be revised or weighted based on how influential operating models are in determining performance outcomes and their relative credibility.**
	D5i	OM ramifications	An evaluation of how the relative importance of operating models in determining management outcomes. It may not be desirable to include a large number of operating models that provide similar projections or a similar test of MP robustness.
	D5ii	Finalization of reference/ robustness sets	The final set of operating models may be simplified or weighted according to their credibility / how consequential they are in evaluating the robustness of MPs. Weighting can be achieved using expert judgement (for example polling) or empirically using statistical methods that assign credibility to models. Weighting of operating models does not affect the MSE calculations, only the calculation of performance metrics and tuning of MPs.
	D5iii	Dialogue meeting	Managers are presented with MP projection results for their qualitative management objectives across various MP archetypes for the final set of operating models. These results communicate a realistic range of performance outcomes, presented by quantitate metrics and the first evaluation of likely trade-offs among objectives such as yield, yield variability and conservation. Managers should now have an appreciation of the types of outputs that are typical to the presentation of final MSE results (e.g., MP decision tables, bar plots, worm plots, trade-off plots).

REFINEMENT PHASE			Quantitative performance metrics are defined and CMPs are refined to achieve particular performance outcomes / performance tuning levels
	Straw dog MSE outputs		Following feedback from the working group and managers, methods for comparing the performance of CMPs are proposed.
	E1i	Tuning targets for MPs	A quantitative interpretation of a qualitative management objective that is used to standardize CMP performance for one axis of the principal performance trade-off (yield, conservation, yield variability) such that their performance can be evaluated on a more level playing field. Tuning can be useful in revealing which MP design is most effective (e.g., provides higher yields for the same conservation performance).
	E1ii	Quantitative performance metrics	One or more quantitative interpretations of each qualitative performance objective are proposed. For example: maximize yield - mean yield over all projection years; be biologically precautionary - probability of not overfishing and not overfished; stable catches - mean % absolute change in yield among projection years.
	E1iii	Interactive results	An approach is proposed for sharing results and allowing working group members, managers and stakeholders to investigate MSE results (e.g., a shiny app, R package, interactive html document).
	MP refinement (technical milestone 3)		An iterative process of MP revision following feedback from the working group
	E2i	Develop MP derivative	A revised version of the previous CMP in response to feedback from the group (e.g., including an approach to reduce catch variability, with a maximum catch limit imposed, that uses a weighted input data etc). These derivatives could include various levels of maximum catch, or maximum change in catch limits between years. I.e., the CMPs 'USA' and 'JPN' might have 'USA_V10' / 'JPN_V10' and 'USA_V20' / 'JPN_V20' in which catch recommendations are allowed to vary by 10% and 20% among years, respectively.
	E2ii	Tune MPs	Tune each CMP to one or more tuning targets (e.g., mean biomass relative to BMSY = 1 after 30 projected years).
	E2iii	Revised MP derivatives / tuning targets & performance metrics	The performance of new tuned CMP derivatives is calculated and presented for feedback from the working group.
	E2iv	Feedback	The working group evaluates CMP performance and proposes changes to CMP design, alternative derivatives and/or new tuning targets.
	CMP shortlisting		The set of candidate management procedures is reduced to simply results presentation while not affecting the range of performance outcomes / trade-offs represented by the full set.
	E3i	MP pruning	CMPs may be removed that, for example, do not meet performance requirements (e.g., a pre specified or legally required conservation objective), are outperformed by another CMP in all performance metrics ('dominated') or provide comparable results to another CMP but rely on larger assumptions, a greater range of less reliable data or are more complex in their methodology.

	E3ii	Results summary	Using the quantitative metrics identified above, results are presented to the working group in various formats including working papers, presentations and the approach for providing interactive results.
	E3iii	Dialogue meeting	Managers and stakeholders are presented with a terse set of MSE results for refined management procedures which now span the final range of performance outcomes and trade-offs.
	MP data		**The data used by the MP are finalized for the calculation of management advice in the first year of adoption**
	E4i	Finalize dataset for use in MPs	Any data streams that are used by the various CMPs are made available for use in calculation of management advice.
	E4ii	Finalize TSD	The MSE framework is now fully documented including available data, operating model equations, performance metrics and tuning targets.
ADOPTION PHASE			**A management procedure is selected, and adopted for provision of management advice with exceptional circumstances protocols.**
	Results exploration		**Final MSE results are presented for the short-list of CMPs**
	F1i	Projections	Projections of quantities of management interest are provided (e.g., yield, biomass, exploitation rate) that inform outcomes on a probabilistic and individual simulation basis (such that managers can visualise the outcomes of a particular projection).
	F1ii	Trade-offs	Trade-offs plots may be presented in, for example, box plots, spider diagrams or 2D scatter plots which communicate the relative performance costs/benefits among CMPs across the various performance axes.
	F1iii	Primary performance metrics	Based on feedback from managers, a subset of the performance metrics are identified that are used as the primary basis for comparing CMPs (as few as possible, preferably less than 5). Metrics may be removed if they are strongly colinear with other metrics (e.g., probability of overfishing and biomass trajectory).
	F1iv	Secondary performance metrics	A set of metrics that may be further used to discriminate among CMPs or are of particular importance to specific stakeholders or managers.
	MP selection		**An MP is selected from the set of CMPs**
	F2i	Update interactive results	The final results are presented in the interactive app.
	F2ii	Elimination / satisficing / ranking	Managers engage in a process of CMP elimination
	F2iii	Adopt MP	A single MP is selected for the provision of management advice

	Establish exceptional circumstances		**A system of detecting departures from operating model conditions is identified that may bring forward a review of operating models and the MP.**
	F3i	Visualize posterior predictive data	Simulated posterior predicted data are identified (usually for those data types used by the MP) that can be used to evaluate whether new observations (collected when the MP is in use), can be used to identify whether there is a departure from the simulated conditions of the operating models.
	F3ii	Define acceptable EC performance	The requirements of the EC protocols which may be quantitative (e.g., power to detect simulations where the MP leads low stock levels) or qualitative such as substantial catch overages or failure to collect data required by the adopted MP.
	F3iii	Define EC protocol	A full description of the protocol for invoking exceptional circumstances. This may or may not include prescribed actions such as a review of other data sources.
IMPLEMENTATION PHASE			
	Calculate advice / check exceptional circumstances		**The adopted MP is used for the provision of management advice.**
	G1i	MP data	The MP input data are processed for use in the EC protocol and calculation of advice by the MP
	G1ii	Exceptional Circumstances	EC protocols are checked using the new MP input data.
	G1iii	Adopted MP	Advice is calculated and presented to managers
	G1iv	Advice	Management advice is implemented
	Operating model review		**A possible outcome of triggering EC protocols**
	G2i	Review operating models	On triggering EC protocols, a formal review of the operating models may be necessary and following that, a review of the adopted MP.

Table 2. Glossary of terms.

Term	Description
MSE	Management Strategy Evaluation: a participatory process to establish management procedures (harvest strategies) that are robust to uncertainties in fishery and population dynamics.
OM	Operating Model: a mathematical description of fishery and population dynamics codified in a simulation framework for the robustness testing of candidate management procedures.
MP	Management Procedure (harvest strategy): a algorithm that calculates management advice from data (real or simulated).
CMP	Candidate Management Procedure. One of multiple possible management procedures that is to be comparatively evaluated by MSE.
MSE framework	The process, membership, meetings, documents, software package, management objectives and exceptional circumstances protocols that support the adoption of a management procedures.
Closed-loop simulation	The engine at the heart of MSE simulations: a codified representation of fishery and population dynamics (operating model) linked to an observation error model (data generation) a candidate management procedure, an implementation model (controls adherence to management advice) which accounts for feedback between the fishery system, data, recommendations and management actions to quantify management performance.
TSD	Trial Specifications Document: a description of the methodology of the MSE framework that ensures reproducibility including all decisions, background information and equations.
Reference Case	A single operating model familiar to the working group that can be used for didactive purposes such as exploring ideas, demonstrating concepts / sensitivities.
Reference Set	A set of operating models, sometimes represented by an orthogonal grid of operating models that represent the core uncertainties that CMPs should be robust to: the primary basis for the evaluation of CMPs.
Robustness Set	A secondary set of operating models used to further distinguish between CMPs that otherwise perform similarly for the reference set of OMs. These may include hypotheses that have a relatively weak empirical basis or uncertain future conditions for projections.
Data guillotine	A date after which new data will not be accepted for use in operating model or management procedure development.
OM conditioning	The process of fitting operating models to observed data statistically (similar to fitting of stock assessment models).
EC Protocols	Exceptional Circumstances protocols: an empirical check that observed data are consistent with those data expected to be observed when the MP is in use (a basis for detecting departures in systems dynamics away from the operating models for which the MP was demonstrated to be robust).

Phase	Task	Processes	Progress

PREREQUISITE

- Identify Purpose — Problem statement
- Establish concepts and terminology — Introductory workshop(s) · Online materials
- Identify membership — Working Group · Technical · Oversight / facilitation · Communications · Managers / Stakeholders

FOUNDATION

- Qualitative objectives — Review legal / comparables · Dialogue meeting
- Outline system uncertainties — Establish principle uncertainties · Identify secondary uncertainties · First cut TSD
- Data prep. for OM conditioning — Review meta data · Present data · **Accept initial dataset for OM conditioning**

First OM data guillotine

INITIAL

- Propose initial OM sets — Reference Case · Reference Set · Robustness Set
- Technical Milestone 1 — Condition Reference Set · Dev. projection model · Dev. reference MP · Code Review
- Presentation of initial MSE results — OM fit · Projection of reference MP · Plausible outcomes · Dialogue meeting
- Finalize OMs — **Accept final dataset for OM conditioning** · **Recondition OMs**

Second OM data guillotine

REVISION

- Identify possible MP archetypes — Establish management levers · Future data availability
- Technical milestone 2 — Produce guide to MP development · Develop example MPs of each archetype
- Presentation of revised MSE results — Closed loop projection of MPs for all OMs · Projection outcomes for qualitative objectives
- Revise / simplify / weight OMs — OM ramifications · **Finalization of Reference / Robustness Sets** · Dialogue meeting

REFINEMENT

- Straw dog MSE outputs — Tuning targets for MPs · Quantitative performance metrics · Interactive results
- MP refinement / Technical Milestone 3 — Develop MP derivatives → Tune MPs → Revised MP derivatives / tuning targets & performance metrics → Feedback
- MP shortlisting — MP pruning · Results summary · Dialogue meeting
- MP data — **Finalize dataset for use in MPs** · **Finalize TSD**

MP data guillotine

ADOPTION

- Results exploration — Projections · Trade-offs · **Primary performance metrics** · **Secondary performance metrics**
- MP selection — Update interactive results · Elimination / satisficing / ranking · **Adopt MP**
- Establish exceptional circumstances — Vizualize posterior predictive data · Define acceptable EC performance · Define EC protocol

IMPLEMENTATION

- Calculate advice / Check exceptional circumstances — *Management update* : MP data → Exceptional Circumstances → Adopted MP → Advice
- Operating model review — *Exceptional* : Review operating models

Figure 1. Updated MSE roadmap that includes processes for new applications where stakeholders and managers are not familiar to the concepts and terminology, and may not have yet established performance objectives. Unless specified by arrows, the process runs to the right and then downwards. Note that unless a specific group (colour) is assigned to a process (just a white box), all members of the working group are invited to participate.