Report of the MSE workshop, July 14-18, 2025

2025-07-25 07:02

Executive summary

The Jack Mackerel Management Strategy Evaluation (MSE) Technical Workshop brought together scientists, technical experts, and external reviewers to review recent progress and refine the MSE framework being developed under SPRFMO. The primary goal of the workshop was to ensure that the modeling framework and management procedures (MPs) are scientifically sound, technically transparent, and aligned with management priorities.

Key Outcomes and Advancements

1. **MSE Framework Consolidation**
   Participants reviewed the jmMSE software package, confirming that it provides a robust and flexible platform for conducting MSEs. The package includes a reference set of operating models conditioned to historical data using MCMC, an efficient MP tuning algorithm, and tools for visualizing and comparing results.

2. **Robustness Testing**
   The workshop clarified the role and scope of robustness tests. These tests are intended to explore how CMPs perform under a range of plausible yet uncertain scenarios rather than represent definitive alternative models. Scenarios reflecting changes in recruitment, spatial availability, environmental regime shifts (e.g., El Niño), and stock structure were reviewed and refined for implementation.

3. **Indicator-Driven MPs and HCR logic**
   Empirical MPs based on one or more indicators were evaluated, with focus on two formulations:

   - TAC as a product of a target and a multiplier from an index.
   - TAC adjusted incrementally from the previous year based on index signals.

   We noted that the high current stock status (well within the "green" zone) tended to increase catch levels when tuning to achieve a desired P(Green). This can result in declining stock trends later in the projection period, even when the short-term performance criteria are met.

4. **Recommendations and Refinements**
   The group recommended additional diagnostics and refinements, including:

- Adding plots of how index trajectories relate to TACs.
- Including new performance metrics that reflect stock status and trends in the final projection years.
- Ensuring consistent treatment of selectivity, weights-at-age, and catch splits in both projections and reference point calculations.
- Exploring robustness scenarios that account for variability in fleet selectivity and biological assumptions, particularly where CPUE is used as an input.

5. **Documentation and Transparency**
   The group emphasized the importance of transparency in documenting model assumptions, data sources, and MP structure. The group agreed on priorities for improving documentation and sharing annotated examples of MP behavior.

6. **Next Steps and Implementation**
   The next phase of work will focus on finalizing the candidate MPs, running the robustness tests, and summarizing trade-offs across key performance indicators. In discussions we also identified future reporting needs, including summary tables and figures for managers, and exploration of reference points and evaluation criteria beyond the green zone probability.

Introduction

Management Strategy Evaluation (MSE) has emerged as a critical tool for fisheries management, especially in contexts where data are limited or uncertainty is high. Foundational software frameworks like FLR were developed to facilitate reproducible, cross-disciplinary evaluation of management strategies through simulation and decision analysis (Laurence T. Kell and Scott 2007). Building on this foundation, recent advances have expanded FLR's capacity for data-rich and data-limited systems alike, improving accessibility and integration with other tools (Hillary et al. 2023). Complementing these developments, a structured framework for evaluating methods and risk in data-limited fisheries has been proposed, providing practical guidance on the application of MSE in real-world settings (Carruthers et al. 2023).

The SCW15 Jack Mackerel MSE Technical Workshop was convened in response to the Scientific Committee's request for progress on developing and evaluating management procedures (MPs) for jack mackerel under the SPRFMO framework. The meeting was held over five days (14–18 July 2025) and hosted in a hybrid format, with active participation from in-person attendees in Seattle and remote collaborators from SPRFMO Member States and invited experts. This event followed on previous technical work, including the SCW14 benchmark, and focused on finalizing the reference set of operating models (OMs), implementing robustness tests, and refining MP candidates using the jmMSE software package.

Throughout the week, participants engaged in live coding sessions, software validation, model tuning, and scenario refinement. The agenda was intentionally flexible, allowing the group to respond dynamically to technical challenges—such as issues with index generation, selectivity artifacts, and catch variability under different MP formulations. The workshop emphasized transparency, reproducibility, and documentation, with clear objectives to improve the utility and credibility of the MSE outputs ahead of Scientific Committee and Commission review.

By way of review, we provide a general outline for the workflow for defining and evaluating Management Procedures (MPs). We divided the process into three main stages (Figure 1).

The workshop participants recognized that all of the pieces for this exercise were available and implemented. However, the group struggled with getting the candidate MPs defined relative to available indices ("stage 1" in the diagram).

The appendices provide the participants, the agenda and the daily activities in the workshop minutes. The review from the external experts is summarized in an appendix as well.
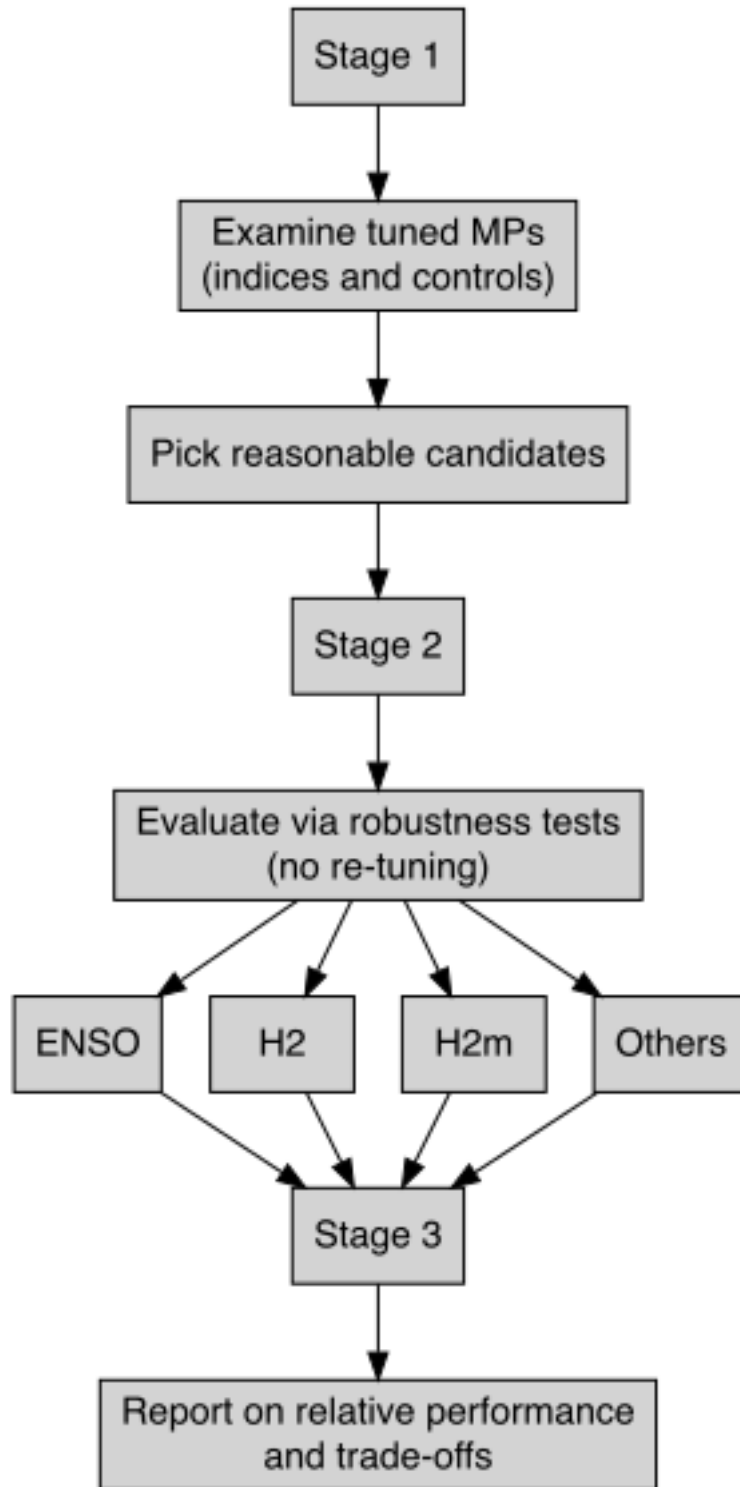
Figure 1: Workflow for evaluating and selecting candidate management procedures (MPs).

The following sections detail the discussion on how best to incorporate some environmental effects for projecting from the operating model. Specifically, we considered how to account for the effects of El Niño on recruitment and catchability/availability.

Simulated El Niño Effects in the Operating Model

To incorporate climate-driven variability into the Operating Model (OM), we define a scenario simulating El Niño–like events every five years beginning in 2030. These events affect primarily recruitment and distribution options. The analysts also proposed biological and fishery processes for El Niño conditions. The group discussed these and noted that they could be evaluated in the next round of stock assessment benchmark and MSE work.

The table below summarizes the proposed effects of simulated El Niño conditions on the Operating Model (OM), categorizing them by their expected direction, biological or fishery-based justification, and evaluation priority. The table is divided into two sections: effects that are prioritized for immediate evaluation and those deferred for further study.

The first items highlights two key El Niño-driven effects: A 30% increase in recruitment with a one-year lag, linked to ENSO-related early life stage survival (Figure 2). Shifts in catchability, with coastal regions experiencing increased availability and offshore regions seeing declines, reflecting observed onshore movement of fish during warm anomalies. This is also high-priority, with a focus on quantifying impacts on fishery removals.

The deferred effects were discussed and included the potential for reduced weight-at-age (potentially due to prey scarcity), earlier maturity (a stress response observed in small pelagics), and increased natural mortality (from predation or environmental stress). These are flagged for future study, pending historical data checks or further evidence. The table succinctly organizes hypotheses while clarifying immediate next steps for the OM framework.

The following tables summarizes the key effects, their expected directions, and justification.

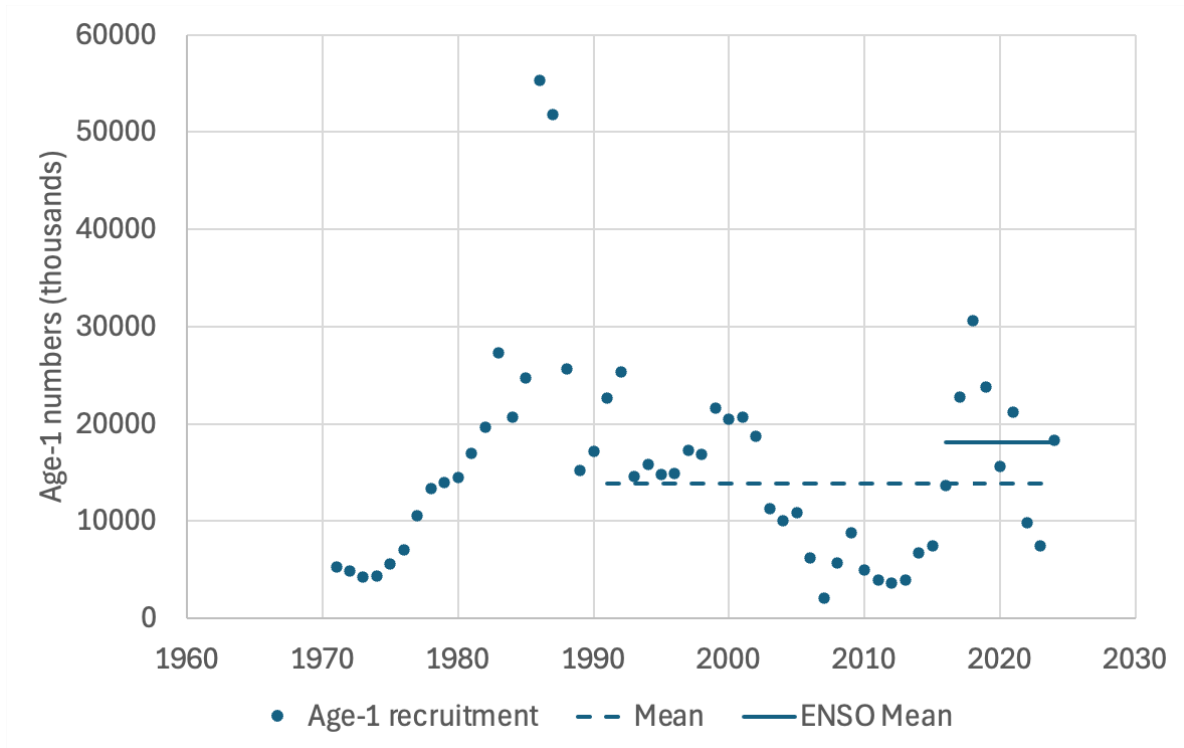| Effect	Direction	Justification
Recruitment ↑20%	↑ 1-year lag	ENSO-linked early life stage effects on recruitment
Regional availability	Coast catchability ↑ and offshore ↓	Onshore shift during warm anomalies

Discussed but deferred for further study

Figure 2: Recruitment estimates and mean values (horizontal lines) used to estimate the impact of ENSO effects.

Effect	Direction	Justification	Notes
Weight-at-age	↓Productivity	Lower prey density and observed condition declines	Check historical WAA anomalies
Age-1 maturity	Earlier maturity	Stress response seen in small pelagics	similar impact on Recruitment
Coastal selectivity	Age 1–2 sel ↑	Spatial contraction, availability change	Confirm from CPUE by age?
M ↑30%/20%	↓ Survival	Stress-induced mortality, predation	

Estimating Relative Availability from Catch Proportions

Taking an assumption that over a recent period that changes in a smoothed proportion of catch by coastal and offshore areas roughly relates to the effective catchability q (which includes both true catchability and availability) of the fishery, i.e.:

$$\text{Catch}_{\text{fleet}} \propto q_{\text{fleet}}$$

Thus, observed catch proportions can serve as a proxy for relative availability.

Table 3: **5-year moving averages of the proportion of the catch occurring in the "coastal" areas compared to the offshore fleet.**

Year Range	Coastal (%)	Offshore (%)
2004–2008	81	19
2005–2009	78	22
2006–2010	74	26
2007–2011	76	24
2008–2012	78	22
2009–2013	82	18
2010–2014	84	16
2011–2015	86	14
2012–2016	86	14
2013–2017	85	15
2014–2018	86	14
2015–2019	87	13
2016–2020	91	9
2017–2021	93	7

Year Range	Coastal (%)	Offshore (%)
2018–2022	94	6
2019–2023	94	6
2020–2024	94	6
Mean	**85**	**15**

Range of Change in Estimated Availability

Given this pattern we can assume a relative catchability due to an environmental effect. We note that the

- **Coastal effective catchability** increased from a low of **74%** (2006–2010) to a high of **94%** (2018–2024), a **+20 percentage point** change.
- **Offshore effective catchability** declined from **26%** to **6%**.

As a sensitivity, we could propose that the effective availability to the offshore fleet drops from 15% of the biomass (the mean) to gradually to 6% during El Niño periods (a 60% decline in q). This would apply to the data generated for offshore CPUE index in the simulations. For the coastal zones, the effect of El Niño would correspond to an 11% increase in the availability of fish relative to the mean (85%). These changes would apply to the Chilean SC CPUE index and the Peruvian CPUE index data generation. This is one proposal among many that could be imagined. For example a slightly more conservative range could be based on the 10th and 90th percentiles of estimated effective catchability (from proportional catches):

```
Coastal:
•    10th percentile: 77%
•    90th percentile: 94%
Offshore:
•    10th percentile: 6%
•    90th percentile: 23%
```

These shifts may provide some scope for showing the impact of changes in the relative abundance indicators in index values. These reflect patterns over the past two decades, possibly due to environmental changes.

Review of the OM specifications

The workshop reviewed the current specifications of the Operating Models (OMs). In particular, the assumptions for the reference point calculations were discusssed and contrasted with the 2024 assessment results and reports (South Pacific Regional Fisheries Management Organisation (2024)). Due to the terminal (2024) estimates of fisheries selectivities, the assessment report had anomalously high values for F_{MSY}

While the 2024 stock assessment produced high estimates of potential catch under the third tier of the harvest control rule—exceeding 4,900 kt based on F_{MSY}—this result was considered unrealistic due to likely upward bias in F_{MSY} estimates caused by strong selection on older fish. As a result, the Scientific Committee recommended constraining the 2025 TAC to be at or below 1,428 kt, representing only a 15% increase from 2024 levels and aligned with the Commission's guidance. In developing the Operating Model (OM), reference points such as F_{MSY} were instead based on longer-term averages to avoid the influence of short-term variability or cohort effects, ensuring more stable and precautionary management advice consistent with the MSE framework (Figure 3).

Several issues were identified with the current stock assessment that warrant further attention ahead of the next benchmark. Key among them are uncertainties in catch-at-age data stemming from differences in age determination methods across laboratories, as well as assumptions about mean body weight at age. The Scientific Committee emphasized the need for standardizing CPUE indices and improving data collection protocols, particularly regarding fleet-specific efficiency changes. Sensitivities to early age composition data—especially from the pre-1990 period—remain unresolved, with residual patterns noted for the North Chilean fleet. In addition, assumptions underlying selectivity and recruitment regimes were highlighted as critical sources of uncertainty, with substantial influence on reference points and management advice. Finally, the Committee underscored the importance of continued evaluation of single-stock versus two-stock model structures using simulation and MSE tools.
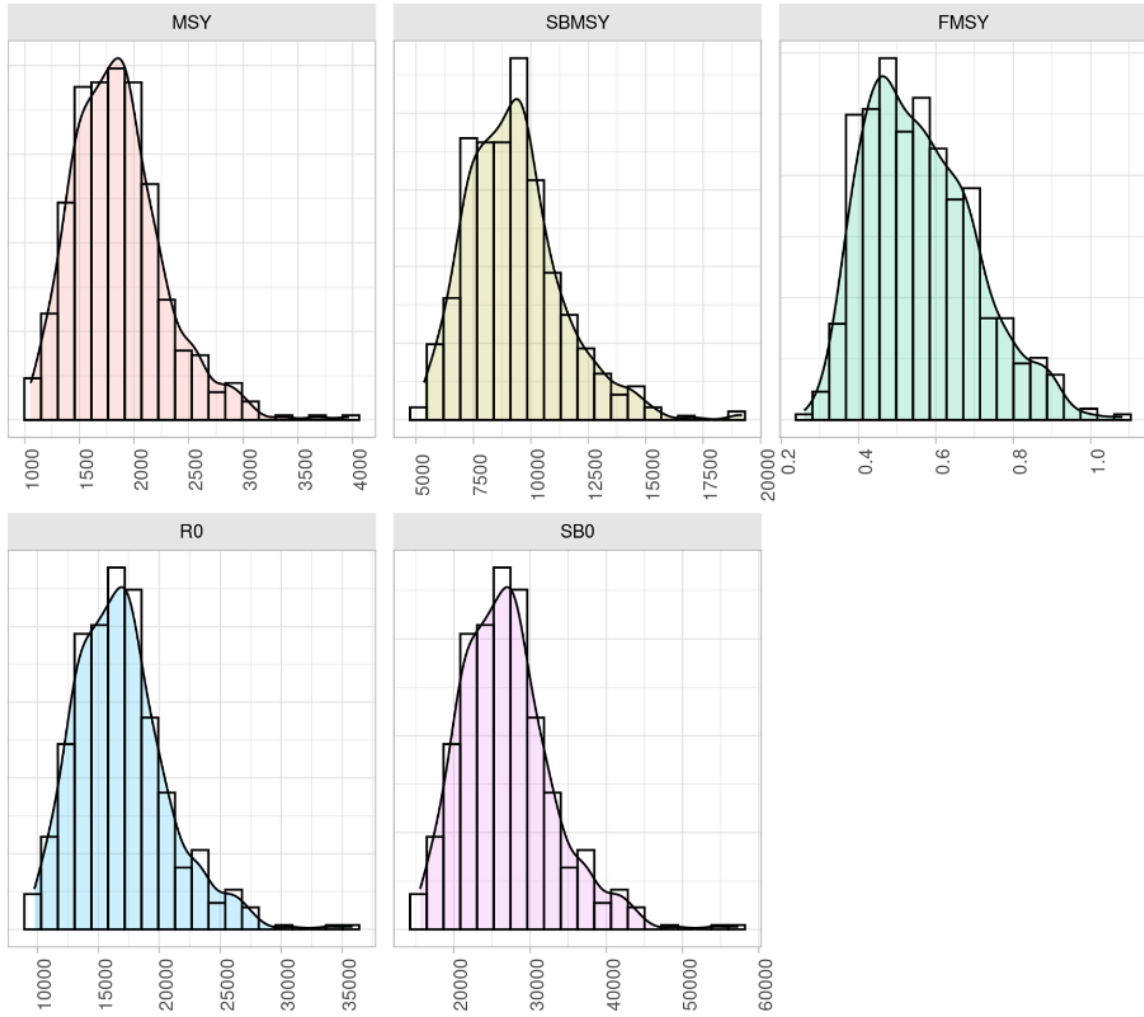
Figure 3: Distribution of reference points from the operating model accepted by the workshop.

Summary of Workshop Outcomes

The SCW15 workshop provided a venue for progressing the Jack Mackerel MSE work, resolving some technical issues, and evaluating multiple MP configurations. A key outcome was the identification of problems in how MPs interacted with the OMs. This raised the need for either resolution of MP specification issues prior to the SC and/or holding a separate follow-up technical meeting, ideally in person. This would focus discussion on narrowing MP options. Depending on this direction, it may mean that such a meeting would have to occur after February 2026 and the Commission meeting.

Participants were encouraged to document their activities during the workshop, including the methods explored and tuning targets used. Work tasked identified included:

- **Jim** evaluated 9 MPs (including `bufferdelta2`, `cpuescore2`, test acoustic, and combinations of CPUE indices with different `delta_TAC` values), all tuned to achieve 60% green status.

- **Jose/Chile** apply shortcut tuning methods to reach similar green zone targets.

Software and Technical Recommendations

- Continue using **FLR** as the main MSE engine unless there is a dedicated effort to migrate to **openMSE** or another platform.

- Improve naming conventions in code to reduce ambiguity. For example:

  - Functions like `cpuescore2.ind` and `cpuescore3.ind` are not intuitive.

  - The `target` argument is overloaded: it may refer to either a TAC or depletion target depending on context.

MSE Development Timeline and Deliverables

The group noted that MSE funding (in the form of providing support from external developers) may be available but would be contingent on:

- Coordination with the current analyst (Iago).

- Collaboration with the technical team.

- Openness to using **openMSE**.

- Clear timelines and deliverables.

Proposed deliverables and deadlines include:

- **Reference OMs** (no multistock): End of July

- **Robustness OMs** (no multistock): End of August

- **Shortcut calibration** to the JJM assessment: End of August

- **Range of shortcut MPs** run for all reference OMs: End of July

Planned products:

- Technical documentation and reports:

    - Draft Technical Summary Document (TSD) by end of July.

    - Technical working papers and presentations for:

        * Shortcut calibration to JJM

        * Reference set OM results for MP archetypes

        * Robustness OM results

        * MP performance summaries

    - Slick MSE results summary.

    - TSD v1 by end of September.

Recommendations

For SC:

- Adopt the current proposal structure with flexibility for future adjustment.

- Recommend a shortlist of MP options to simplify the selection process at the Commission level.

- Consider a placeholder method for calculating the 2026 TAC if MSE work is not yet finalized.

For Members:

- Commit to a shared MSE software base (FLR or openMSE).
- Engage in pre-SC online meetings to broaden participation in MSE discussions.

For Analyst (Iago):

- Prioritize enhancements discussed during the workshop:
  - Code clarity and naming conventions
  - Logical parameter usage across MPs
  - Refinement of FLR-to-dataframe functions
- Identify successor strategy after contract ends in 2025.

Summary Table: MP Methods and Configuration

Year	Method	Metric	Tuning Parameter	Other Parameters	Score Index	Comments
2024+	`buffer.hcr`	depletion	target	bufflow, buffup, limit	`cpuescore3.ind`	Original pkg function
2024+	`bufferdelta.hcr`	depletion	width	sloperatio	`cpuescore3.ind`	Modified; not compatible with z-score metrics
2024+	`bufferdelta2.hcr`	zscore	width	sloperatio	`cpuescore2.ind`	New; not compatible with depletion
2024+	`buffer2.hcr`	zscore	target	width (affects buffer)	`cpuescore2.ind`	Original; adjusted for zscore (limit = -2 SD)

References

Carruthers, Thomas R., Quang C. Huynh, Adrian R. Hordyk, David Newman, Anthony D. M. Smith, Keith J. Sainsbury, Kevin Stokes, et al. 2023. "Method Evaluation and Risk Assessment: A Framework for Evaluating Management Strategies for Data-Limited Fisheries." *Fish and Fisheries* 24 (6): 1335–50. https://doi.org/10.1111/faf.12726.

Hillary, Richard M., José M. Castro, James T. Thorson, Sean C. Anderson, and Laurence T. Kell. 2023. "The FLR Software Framework for Building Management Strategy Evaluation Systems: Recent Advances and Application to Data-Rich and Data-Limited Fisheries." *Fisheries Research* 263: 106585. https://doi.org/10.1016/j.fishres.2023.106585.

Laurence T. Kell, Paul Grosjean, Iago Mosqueira, and R. D. Scott. 2007. "FLR: An Open-Source Framework for the Evaluation and Development of Management Strategies." *ICES Journal of Marine Science* 64 (4): 640–46. https://doi.org/10.1093/icesjms/fsm012.

South Pacific Regional Fisheries Management Organisation. 2024. "SC12 Report Annex 07: Jack Mackerel Technical Annex." SPRFMO; https://sprfmo.int/assets/Meetings/02-SC/12th-SC-2024/SC12-Report-Annex-07-JM-Technical-Annex.pdf.

Appendix A, workshop participants

Figure 4: Participants at the SCW15 technical workshop. From left to right; back row: Iago Mosqueira, Aquiles Sepulveda, Jose Zenteno, Josymar Torrejó; front row: Ricardo Oliveros-Ramos, Grant Adams, Ana Parma, Jim Ianelli, Ignacio Paya.

Appendix B, Agenda

1. **Welcome and Introduction** (5 minutes)
   Objective: Set the stage for the focused discussion on Jack Mackerel MSE progress and future directions.
   Note: Briefly introduced the MSE as a tool for sustainable fisheries management, emphasizing its importance given the historical fluctuations in jack mackerel stock and exploitation levels.

2. **Current Status of Jack Mackerel MSE Work** (10 minutes)
   Objective: Provide a concise update on progress, including advancements in Operating Models (OMs) and MP testing.
   Note: MSE development has progressed rapidly. Updates include new data inputs, refinements from the SCW14 benchmark, and expanded uncertainty axes. MPs are under active testing.

3. **Review of Candidate MPs and Tuning Results** (30 minutes)
   Objective: Discuss MP structure, indicator choices, and implications of tuning to achieve target performance criteria.
   Note: Focused on empirical MPs using CPUE and acoustic indices. Tuning challenges under high biomass conditions were highlighted.

4. **Robustness Scenarios and Specification Refinement** (30 minutes)
   Objective: Finalize the list of scenarios for robustness testing.
   Note: Scenarios included El Niño-like variability, availability shifts, and alternative stock structures. Their role as comparative stress tests was reaffirmed.

5. **Evaluation Metrics and Visualization Tools** (20 minutes)
   Objective: Review tools for comparing MP performance.
   Note: Emphasis on visual summaries, including Kobe plots, probability tiles, and trade-off diagrams.

6. **Feedback from External Experts** (15 minutes)
   Objective: Integrate external review findings and technical suggestions.
   Note: Dr. Parma's input emphasized realistic assumptions, consistent reference points, and long-term performance evaluation.

7. **Wrap-up and Next Steps** (10 minutes)
   Objective: Identify action items and prepare for upcoming reporting deadlines.

Note: Plans were set for refining MPs, running full simulations, and summarizing results in a format accessible to decision-makers.

Appendix C, minutes

Day 1 summary

Participants and Setup

- The workshop included both in-person and online participation, with representatives from Peru, Chile, Argentina, Ecuador, and the Netherlands, among others.

- The agenda was described as ambitious, with a focus on hands-on technical work, including software installation and repository access.

Technical Infrastructure & Workflow

- Two main GitHub repositories are central: FLjjm (for building FLR objects and running JJM inside the MP) and jjmMSE (the main development site for the MSE work).

- The workflow follows the DAF (Data, Analysis, Framework) system, with clear steps for data preparation, OEM (Operating Model) conditioning, and performance analysis.

- Emphasis was placed on forking repositories and using branches for collaborative work, with a preference for merging at the end of the week.

- Docker was suggested as a potential solution for ensuring consistent environments across operating systems and participants, though not yet implemented.

Modeling and Simulation

- The group is working with both single-stock and two-stock hypotheses, including a two-stock model with future-applied connectivity based on movement matrices.

- Robustness scenarios are being explored, including cyclic environmental changes and their impacts on productivity.

- The group discussed the use of ".q" files for efficient storage and handling of large simulation outputs.

Key Scientific Discussions

- The calculation and interpretation of MSY (Maximum Sustainable Yield) and F_{MSY} were debated, with concerns about the realism of current methods in JJM.

- It was agreed that the 10-year average of MSY reference points would be used for performance evaluation, to avoid short-term volatility.

- Selectivity patterns and their impact on projections were a major topic. The group considered transitioning from terminal year selectivity to long-term averages over a five-year period to avoid unrealistic jumps in catch projections.

- There was consensus that the main focus should be on long-term performance of management procedures, but short- and mid-term results are also important for managers.

Environmental and Biological Scenarios

- The workshop addressed the need to model environmental variability, particularly El Niño events, and their impact on stock productivity, weight-at-age, and selectivity.

- Literature-informed scenarios were presented, with parameters for changes in mortality, recruitment, and spatial distribution.

- Regional differences (e.g., between far north and south stocks) and their implications for catchability and biological responses were discussed.

Action Items and Next Steps

- Participants will review and potentially refine the environmental scenarios, with a focus on realism and literature support.

- Further work is planned on selectivity transitions and the technical implementation of gradual changes.

- The group will continue to test and validate the workflow, with an emphasis on reproducibility and collaborative code development.

Day 2

Opening

- The workshop began with participant introductions, including new attendees such as Robert Robinson.

- Jim Ianelli welcomed participants, provided a recap of the previous day, and referenced a summary posted in the Teams channel for review.

Review of Previous Work and Agenda

- The agenda was described as ambitious, with a focus on technical aspects of Management Strategy Evaluation (MSE).

- Discussion included the effects of El Niño on recruitment and catch distribution, referencing an unsent summary email and a report on the topic.

Technical Discussions

a. Effects of El Niño and Distribution Shifts

- Jim presented an analysis of catch distribution changes between coastal and offshore fleets, proposing a 15% average offshore drop (60% decline in Q for offshore fleet) and an 11% increase for coastal zones.

- Participants debated the appropriateness of using long-term versus short-term averages and the need for smoothing changes rather than step changes.

- It was agreed that the effect should be applied to both availability and catchability in simulations, with implications for the acoustic survey indices.

b. Selectivity and Projection Periods

- The group discussed the transition from recent selectivity estimates to long-term means, with concerns about artifacts from using 10-year averages.

- Consensus moved toward using a representative period (2000–2010) for selectivity, avoiding recent years with anomalously high selection for older fish.

- The need for a smooth transition in selectivity assumptions for projections was emphasized.

c. Recruitment Regimes and Projections

- Analysis of recruitment means for projections considered the El Niño effect, with a proposed 23–30% bump up in recruitment for recent years.

- Debate ensued on which years to include for calculating means, with a focus on data consistency from 1991 onward.

- The group discussed the potential for regime shifts and the implications for robustness testing in MSE.

d. Model Implementation and Coding Practices

- Demonstrations were given on the use of R and package structures for running MSE simulations, including best practices for project setup and function sourcing.

- The importance of consistent selectivity and Q parameter normalization across projections and indices was highlighted.

Decisions and Action Items

- Adopt 2000–2010 as the reference period for selectivity in projections, with a smooth transition from current conditions.

- Apply a 23–30% recruitment increase for projections reflecting recent El Niño effects, with final years to be confirmed.

- Ensure normalization of selectivity and Q parameters is consistent across all indices and projections.

- Continue refining the codebase, documenting changes, and sharing updates among the technical team.

Other Business and Closing

- Participants shared experiences with data handling, model setup, and coding challenges.

- The workshop included informal discussions and technical clarifications.

- The session concluded with plans to continue reviewing model outputs and performance indicators, and to reconvene as needed for further technical work.

Day 3

Review of Technical Issues and Model Adjustments

- Discussion focused on the technical aspects of Management Strategy Evaluation (MSE) for jack mackerel.

- Participants examined the function and configuration of sliding buffers, control rules, and the impact of selectivity changes.

- The group reviewed how indices (e.g., CPUE, acoustics) are incorporated, including the use of averages over multiple years and weighting schemes for recent data.

- There was debate on the variance and correlation structure in observation models, and how these affect simulation results.

Timing and Implementation of Management Procedures (MPs)

- The workflow and timing for implementing MPs were clarified:

  - Data from 2026 would be used for advice in 2027.

  - The Scientific Committee (SC) would run the MP, with the Commission making final decisions.

  - Discussion on the need for preliminary versus finalized data and the implications for observation error.

- The importance of using the most recent data versus the stability of multi-year averages was highlighted.

Treatment of Effort Creep and Index Standardization

- Participants agreed that simulating effort creep in future projections is not necessary for the base case, but could be a robustness test.

- The importance of correcting historical indices for effort creep was emphasized, while future indices are assumed to be unbiased.

- There was consensus that the standardized or corrected index should be used for both simulation and real-world application.

Selection and Tuning of Management Procedures

- Multiple MPs were tested, each tuned to a 60% probability of meeting the Kobe green zone.

- The group compared different buffer widths and TAC change limits, analyzing their effects on catch variability and performance metrics.

- Discussion included the need for clear documentation of specifications and the importance of presenting a set of MPs with distinct trade-offs to the Commission, rather than recommending a single option.

Performance Metrics and Projections

- The team reviewed performance across near-term, medium-term, and long-term projections.

- Boxplots and other visualizations were used to compare MPs, focusing on catch variability, probability of stock being in the green zone, and interannual TAC changes.

- Concerns were raised about downward trends in some simulations and the need for additional performance statistics (e.g., probability of stock crash).

Next Steps and Action Items

- Agreement to refine the operating model (OM) and finalize Annex documentation on selectivity and other specifications.

- Plan to continue testing and tuning MPs, with further analysis of performance metrics.

- The SC will present a set of MPs to the Commission, along with the status quo as a fallback.

- Lunch arrangements and informal discussions concluded the session.

Day 4

Key Activities and Discussions

- Model Runs and Debugging:

  - Overnight and morning runs were conducted to examine the acoustics index using legacy targets and buffers.

- A significant focus was on debugging the generation of CPUE v3 for 2024 and 2025, investigating unexpected jumps in predicted values.

- Alternative normalization methods for CPUE were tested, following recommendations to improve stability.

- Analysis of Index Jumps:

  - The team identified that the jump in the index from 2024 to 2025 was primarily driven by increases in vulnerable biomass and changes in mean weight at age, rather than selectivity changes alone.

  - Weighted age calculations and their implications for projections were reviewed in detail, including the use of three-year means and the impact of preliminary data from 2024.

- Code Review and Live Demonstration:

  - Live coding sessions were held to demonstrate how to adjust the operating model (OM) to exclude problematic years (e.g., dropping 2024 and extending from 2023).

  - Smoothing techniques for selectivity and weights at age were discussed and implemented to reduce artificial jumps in projections.

- Uncertainty and Robustness:

  - The group discussed the treatment of process error, residual variability, and auto-correlation in indices.

  - Empirical approaches to setting CVs for indices were compared with default values, and the impact on future projections was considered.

- Action Items and Next Steps:

  - Further testing of the OM with adjusted years and smoothing is to be completed, with outputs to be pushed under new filenames to avoid disrupting ongoing work.

  - Continued analysis of the causes of high catches in test runs and further parameter tuning were assigned.

  - The group agreed to revisit and possibly refine the approach to handling preliminary data and smoothing in both indices and weights at age.

Notable Outcomes

- Consensus that both selectivity and mean weight at age contribute to observed index jumps, with smoothing and exclusion of problematic years being viable mitigation strategies.

- Agreement to document and communicate technical progress to the broader group, while maintaining a focus on robust, transparent modeling practices.

Day 5

Model Runs and Technical Issues

- The group reviewed progress on running various management procedures (MPs), focusing on the acoustic and CPUE indices. Issues with FL libraries and model reproducibility were discussed, with fixes applied to ensure models ran as expected.

- Robustness tests were conducted, particularly comparing the performance of different indices (acoustic, CPUE3, CPUE6, 3.6). The need to clarify how "shortcut" procedures treat stocks was debated, especially regarding biomass tracking and catch splits.

- The group noted that tuning parameters (e.g., width, slope ratio) and their impact on model performance remain a challenge, especially when standardizing across indices.

Interpretation and Presentation of Results

- There was significant discussion about interpreting outputs, particularly when catch trends did not align with biomass trends. Concerns were raised about the credibility of certain indices and the need for clearer communication of model behavior.

- The importance of visualizing trade-offs and the response of TAC to indices was emphasized, with ongoing efforts to develop summary figures for inclusion in reports.

Workflows, Code, and Collaboration

- Participants shared experiences with the codebase, noting progress in understanding and modifying functions, but also highlighting the need for further documentation and standardization.

- The group agreed on the value of reproducibility and transparency, with suggestions to document daily progress and maintain clear records for future reference.

Planning and Next Steps

- The group recognized that while technical progress was made, the process is ongoing. There was consensus on the need for another technical workshop (ideally in person) to continue development and evaluation of MPs.

- The timeline for delivering a report to the Scientific Committee (SC) and Commission was discussed, with acknowledgment that final recommendations are not yet possible. Instead, the report will focus on documenting progress, challenges, and a proposed work plan for the coming year.

- Concerns about funding, continuity (especially regarding software and contracts), and the need for member commitment to ongoing tool development were raised.

Recommendations and Reflections

- The group agreed to recommend continued development and evaluation of MPs, with an emphasis on transparency, reproducibility, and clear communication to managers.

- It was noted that, if a new MP is not ready for 2026, the existing method (Annex K) will remain in use.

- There was broad recognition of the complexity and time required for this process, and appreciation for the collaborative progress made during the workshop.

Appendix D. Summary comments from external experts

Dr. Ana Parma, along with Qi Lee reviewed the jmMSE framework and its application to evaluating candidate management procedures (CMPs) for jack mackerel. Their comments provide both a validation of the current approach and targeted suggestions for improvement.

General Observations on the jmMSE Tool

The jmMSE package offers a comprehensive and flexible platform for conducting Management Strategy Evaluation (MSE). It includes:

- A reference set of Operating Models conditioned to historical data using MCMC.

- A suite of visualization tools and performance metrics for comparing CMPs.

- An efficient tuning algorithm that adjusts user-selected MP parameters to achieve a target outcome, such as a probability of being in the green zone.

A variety of robustness tests were developed to address key uncertainties identified for jack mackerel, such as recruitment variability, fleet-specific availability, and stock structure assumptions (one vs. two stocks). Many of these assumptions relate to potential impacts from El Niño-like events. These robustness scenarios were refined during the workshop, and their role was clarified as stress tests—designed not to predict specific mechanisms, but to evaluate the relative performance of CMPs under plausible alternative futures.

Management Procedures Reviewed

Two primary classes of empirical MPs were evaluated:

1. **Target-based MPs:** TAC is calculated as a function of a fixed target catch multiplied by an index-driven adjustment factor.

2. **Incremental MPs:** TAC is adjusted from the previous period based on indicator trends, resulting in smoother changes over time.

During tuning (e.g., to achieve P(Green) = 0.6), both approaches frequently led to increased TACs and eventual stock declines—especially under a high current stock status. This points to the need for additional metrics that reflect long-term sustainability, not just near-term status probabilities.

Technical Recommendations

Some actionable recommendations were articulated:

- **Indicator Visualization:** Add plots showing the time series of indicators used to drive each HCR.

- **Projection-End Metrics:** Include statistics that summarize stock status and trends at the end of the projection period. This could include P(Green) in the final year or a new trend-based metric.

- **Weight-at-Age Specification:** Enable projections to use mean weights-at-age over a recent period (as done for selectivity), rather than fixing weights from the start year.

- **Consistency of Reference Points:** Ensure that FMSY and SSBMSY reference points used for performance metrics are consistent with the selectivity, weights-at-age, and fleet composition used in projections. This is critical since MPs are tuned relative to these reference points.

- **Observation Error Realism:** Consider adding robustness scenarios with variable selectivity and weight-at-age during projections to better represent realistic observation error in indices—particularly for commercial CPUE.

Appendix E. Notes on the Jack Mackerel MSE Framework

This document summarizes the structure and behavior of key objects used in the **jmMSE** Management Strategy Evaluation (MSE) framework for SPRFMO Jack Mackerel. It documents the modeling components (`h1`, `om`, `perf`), performance calculations, and the use of `getSlick()` and `FLslick()` to generate evaluation plots.

Defining Management Procedures using `mpCtrl` and `mseCtrl`

In the `mse` package, Management Procedures (MPs) are constructed as modular sequences of functional components that simulate how a fishery would be managed under alternative strategies. These strategies are defined using `mpCtrl`, which organizes component modules defined via `mseCtrl`.

This modular design allows you to:

- Select estimation methods for stock status (`est`)
- Define harvest control rules (`hcr`, `phcr`)
- Simulate implementation systems (`isys`)
- Include optional technical measures (`tm`)

Structure of `mpCtrl`

The `mpCtrl()` constructor takes a named list of components. Each component must be an `mseCtrl` object that defines the function to use (`method`) and its input parameters (`args`).

```
ctrl <- mpCtrl(list(
  est = mseCtrl(method = shortcut.sa, args = list(
    metric = "depletion",
    devs = metdevs,
    B0 = refpts(om)$SB0
  )),
```

```
  hcr = mseCtrl(method = buffer.hcr, args = list(
    target  = 1000,
    bufflow = 0.30,
    buffupp = 0.50,
    lim     = 0.10,
    min     = 0,
    metric  = "depletion"
  )),
  isys = mseCtrl(method = split.is, args = list(
    split = catch_props(om)$last5
  ))
))
```

Explanation of Components

Component	Description
est	The estimator module determines how the stock status is assessed during each management cycle.
hcr	The harvest control rule uses the estimated status to recommend a total allowable catch or effort level.
phcr	(Optional) Parametrization helper for complex or adaptive control rules.
isys	Simulates how the recommended catch is implemented across fleets, often using historical proportions or allocation logic.
tm	(Optional) Includes non-catch-based rules, such as minimum size limits or gear restrictions.

Common `method` and `args` Options

Component	Example `method`	Common `args`
est	perfect.sa, shortcut.sa	metric, devs, B0, years
hcr	buffer.hcr, hockeystick.hcr, trend.hcr	target, lim, bufflow, buffupp, trigger, metric

Component	Example `method`	Common `args`
`isys`	`split.is`, `fixed.is`	`split`, `noise`, `bias`
`tm`	user-defined	size or spatial constraints

Example: Building an `mseCtrl` for a Harvest Control Rule

Each `mseCtrl` specifies a method function and its arguments.

```
ctl <- mseCtrl(
  method = buffer.hcr,
  args = list(
    target  = 1000,
    bufflow = 0.30,
    buffupp = 0.50,
    lim     = 0.10,
    min     = 0,
    metric  = "depletion"
  )
)
```

Accessor Functions

You can programmatically access or modify components of an `mpCtrl` object:

```
# Access or change the HCR method
method(ctrl@hcr)

# Update the target biomass value
args(ctrl@hcr)$target <- 1200
```

Alternatively, use accessor functions:

```
hcr(ctrl) <- mseCtrl(method = new.hcr, args = list(...))
args(ctrl, "hcr")$target <- 1100
```

Summary

This framework enables:

- Rapid prototyping of management strategies

- Transparent comparisons across MPs

- Flexible integration with simulated operating models (OMs)

By separating each component of an MP into a function-object pair, the `mse` package supports reproducible, configurable, and extensible MSE design workflows.

To analyze the behavior of bufferdelta.hcr() over the range of index values used as input (i.e., the stock status metric like "depletion" or "zscore"), the key output to examine is the harvest control multiplier hcrm—which determines how much the TAC is adjusted relative to the previous TAC. This multiplier is a piecewise function of the index value at the data year.

Harvest Control Rule (HCR) with Buffer Delta

This HCR formulation uses a smoothed transition based on a buffer zone around a biomass or metric `target`. The response scalar $h(m)$, applied to the previous catch, is defined based on the relative metric value m (e.g., standardized index or depletion level), and follows a piecewise logic:

Let:

- m: observed metric (e.g., index value)
- t: target level
- w: buffer width
- $l = t - 2w$: limit threshold
- $b_{\text{low}} = t - w$: buffer lower bound
- $b_{\text{upp}} = t + w$: buffer upper bound
- r: slope ratio

Then the Harvest Control Rule (HCR) response multiplier $h(m)$ is:

$$h(m) = \begin{cases} \frac{1}{2}\left(\frac{m}{l}\right)^2, & \text{if } m \leq l \\ \frac{1}{2}\left(1 + \frac{m-l}{b_{\text{low}}-l}\right), & \text{if } l < m < b_{\text{low}} \\ 1, & \text{if } b_{\text{low}} \leq m < b_{\text{upp}} \\ 1 + r \cdot \frac{1}{2(b_{\text{low}}-l)}(m - b_{\text{upp}}), & \text{if } m \geq b_{\text{upp}} \end{cases}$$

The resulting Total Allowable Catch (TAC) is calculated as:

$$\text{TAC}_{\text{new}} = \text{TAC}_{\text{previous}} \cdot h(m)$$

Optional constraints on TAC changes can be applied:

$$\text{TAC}_{\text{new}} = \min(\text{TAC}_{\text{new}}, \text{TAC}_{\text{previous}} \cdot d_{\text{upp}})$$

$$\text{TAC}_{\text{new}} = \max(\text{TAC}_{\text{new}}, \text{TAC}_{\text{previous}} \cdot d_{\text{low}})$$

Behavior Summary

Let's denote:

m: index metric (e.g., depletion) at the data year

target: central value, e.g., 0.5

width: buffer width, e.g., 1.0

Then:

bufflow = target − width

buffupp = target + width

lim = target − 2 · width

HCR behavior across index values

Index value m	Description	Multiplier h(m)	TAC behavior
$m \leq \text{lim}$	Very low index	Quadratic increase from $0 \rightarrow 0.5$	Strong reduction
$\text{lim} < m < \text{bufflow}$	Between limit and lower buffer	Linear rise from 0.5 to 1	Moderate reduction
$\text{bufflow} \leq m < \text{buffupp}$	Within buffer zone	Flat at 1	No change
$m \geq \text{buffupp}$	Above upper buffer	Linear increase starting at 1 (slope = sloperatio)	Moderate increase

Example with Default Parameters

If you use the defaults:

target = 0.5

width = 1

sloperatio = 0.2

Then:

- ```
  bufflow = -0.5, buffupp = 1.5, lim = -1.5
  ```

- ```
  The flat zone is from -0.5 to 1.5 (note: with depletion metric, this wide range makes se
  ```

If using depletion as the metric, you'd typically want:

target = 0.4

width = 0.1

sloperatio = 0.2

$\rightarrow$ lim = 0.2, bufflow = 0.3, buffupp = 0.5 So the response curve looks like:

```
    |
 ↑  |          /
 h  |         /
 c  |------●-------●------
 r  |    /            \
 m  |   /              \
    |__|_____|____
        lim  bufflow buffupp   → depletion (m)
```

Suggestion: Plot Response Curve

Here's R code to visualize the multiplier hcrm across a range of input metric values:

```r
plot_hcrm <- function(target = 0.4, width = 0.1, sloperatio = 0.2, metric_range = seq(0, 1, 
  bufflow <- target - width
  buffupp <- target + width
  lim <- target - 2 * width

  hcrm <- ifelse(metric_range <= lim,
    ((metric_range / lim)^2) / 2,
```

36

```r
    ifelse(metric_range < bufflow,
      0.5 * (1 + (metric_range - lim) / (bufflow - lim)),
      ifelse(metric_range < buffupp,
        1,
        1 + sloperatio * 1 / (2 * (bufflow - lim)) * (metric_range - buffupp)
      )
    )
  )

  plot(metric_range, hcrm,
    type = "l", col = "blue", lwd = 2,
    xlab = "Metric (e.g., depletion)", ylab = "Harvest multiplier (hcrm)",
    main = "Response of TAC to Index Metric"
  )
  abline(h = 1, col = "gray", lty = 2)
  abline(v = c(lim, bufflow, buffupp), col = "red", lty = 3)
}
plot_hcrm(target = 0.4, width = 0.1, sloperatio = 0.2)
```

Response of TAC to Index Metric

Metric (e.g., depletion)
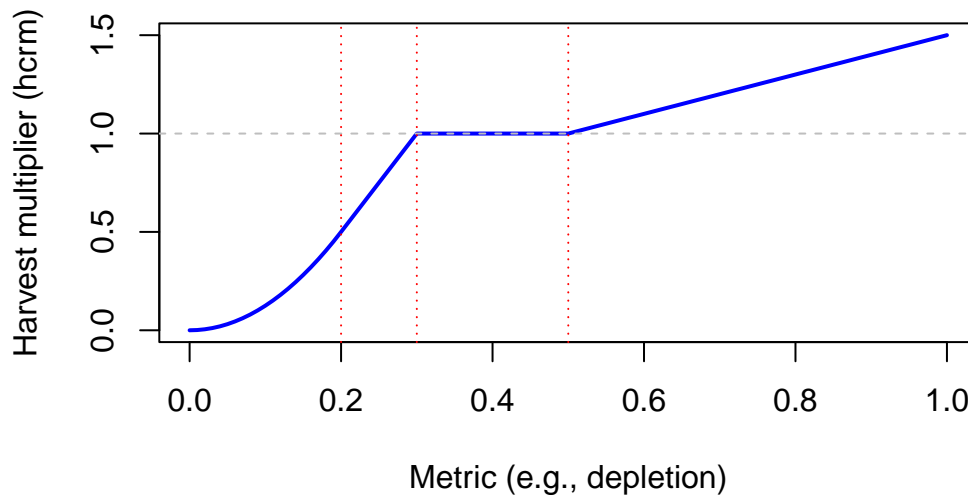
This shows the piecewise nature of the multiplier and can be tailored to any input metric (depletion, zscore, etc.).

Overview of cpuescore

In the jmMSE framework, different CPUE scoring functions are used to inform harvest control rules (HCRs). These functions standardize or compare CPUE time series across simulations and reference periods. The three primary scoring methods are:

- `cpuescore.z`
- `cpuescore.mean`
- `cpuescore.level`

---

1. Z-score Standardization: `cpuescore.z`

This method standardizes the CPUE values by subtracting the mean and dividing by the standard deviation across simulations:

$$\text{score}_{i,t} = \frac{\text{CPUE}_{i,t} - \mu_t}{\sigma_t}$$

Where:

μ_t = mean CPUE in year t across simulations

σ_t = standard deviation in year t across simulations

This produces a unitless, centered score:

```
score <- (met[, ac(dy[i])] %-% yearMeans(ref)) %/% sqrt(yearVars(ref))
```

Useful when you want to assess relative anomalies in CPUE from expected trends.

2. Mean Ratio: cpuescore.mean

This method compares the mean CPUE in recent years (dy) to a reference mean CPUE:

$$\text{score}_i = \frac{\bar{\text{CPUE}}_{dy,i}}{\bar{\text{CPUE}}_{ref,i}}$$

This is a relative index level and is not standardized:

```
score <- yearMeans(met[, ac(dy[i])]) %/% yearMeans(ref)
```

Often used when absolute differences in mean CPUE should affect TAC decisions.

3. Raw Index: cpuescore.level

This method passes through the CPUE values with no transformation:

$$\text{score}_{i,t} = \text{CPUE}_{i,t}$$

```
score <- met[, ac(dy[i])]
```

Used when raw or smoothed CPUE indices are deemed directly interpretable and comparable.

Example: Simulated Scores

The following example shows how the three `cpuescore` methods behave using simulated CPUE data across 100 simulations and 10 years.

```
library(dplyr)
library(tidyr)
library(ggplot2)

# Simulate CPUE data
set.seed(42)
n_sim <- 100
n_year <- 10
ind_sim <- matrix(rlnorm(n_sim * n_year, meanlog = log(1), sdlog = 0.2), nrow = n_sim)
```

```r
# Reference: all years; "recent" period: years 9-10
ref_mean <- rowMeans(ind_sim)
ref_sd <- apply(ind_sim, 1, sd)

z_scores <- (ind_sim[, 10] - ref_mean) / ref_sd
mean_scores <- rowMeans(ind_sim[, 9:10]) / ref_mean
raw_scores <- ind_sim[, 10]

# Reshape for plotting
score_df <- tibble(
  sim = 1:n_sim,
  z = z_scores,
  mean = mean_scores,
  level = raw_scores
) %>% pivot_longer(cols = -sim, names_to = "score_type", values_to = "score")

ggplot(score_df, aes(x = sim, y = score)) +
  geom_line() +
  # facet_wrap(~score_type, scales = "free_y") +
  facet_wrap(~score_type) +
  labs(
    title = "Simulated CPUE Score Types",
    x = "Simulation",
    y = "Score Value"
  ) +
  theme_minimal()
```
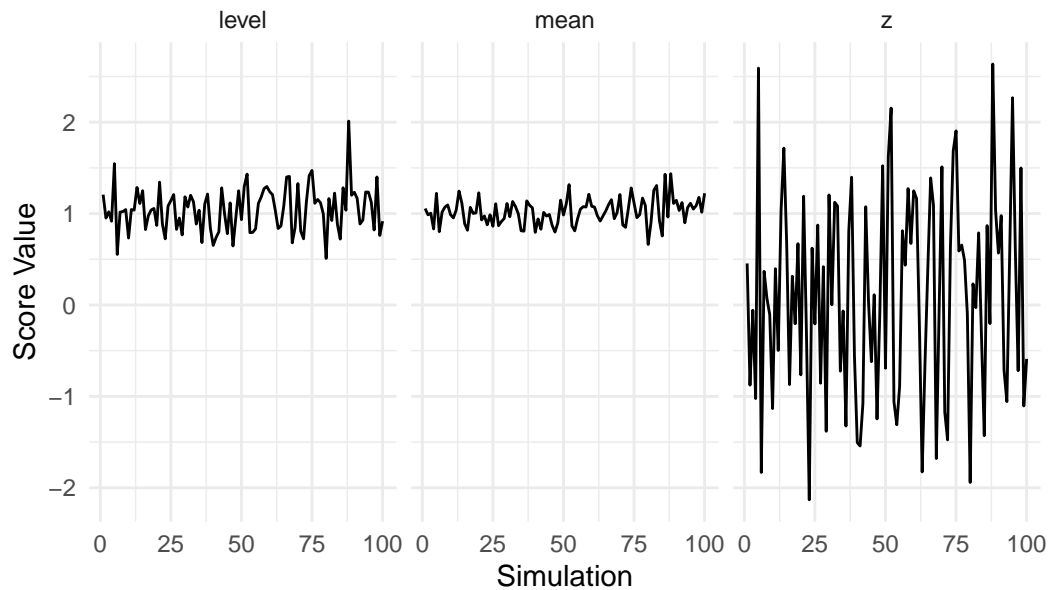
Simulated CPUE Score Types

Summary

The indices were refactored from the `cpuescore` functions to avoid confusion with surveys and can be summarized as follows:

Function	Description	Normalized	Use case
`indscore.z`	Standardize vs mean/sd		Compare anomalies across sims
`indscore.mean`	Mean ratio of dy vs ref		Index levels matter
`indscore.level`	Raw CPUE values used as-is		When index is well-calibrated

```
indscore.z <- function(
    stk, idx, index = 1, dlag = rep(args$data_lag, length(index)),
    refyrs = NULL, args, tracking) {
  dlag <- setNames(dlag, nm = names(idx)[index])
  ay <- args$ay
  dy <- ay - dlag
```

```r
  res <- as.list(setNames(nm = names(idx)[index]))

  for (i in names(res)) {
    met <- window(idx[[i]], end = dy[i]) # removed biomass()

    ref <- if (!is.null(refyrs)) met[, ac(refyrs)] else met

    res[[i]] <- (met[, ac(dy[i])] %-% yearMeans(ref)) %/% sqrt(yearVars(ref))

    dimnames(res[[i]])$year <- max(dy)

    track(tracking, paste0("score.mean.", i), ac(ay)) <- yearMeans(ref)
    track(tracking, paste0("score.sd.", i), ac(ay)) <- sqrt(yearVars(ref))
    track(tracking, paste0("score.ind.", i), ac(ay)) <- res[[i]]
  }

  return(list(stk = stk, ind = res, tracking = tracking, cpue = met))
}
bufferdelta.hcr <- function(stk, ind, target = 0.5, width = 1,
                            sloperatio = 0.2, dupp = NULL, dlow = NULL,
                            args, tracking) {
  # setup
  ay <- args$ay
  iy <- args$iy
  frq <- args$frq
  man_lag <- args$management_lag
  cys <- seq(ay + man_lag, ay + man_lag + frq - 1)

  # compute score
  score <- ind[[1]] # assuming single FLQuant named 'zscore', 'mean_ratio', or 'level'

  # define slope
  bufflow <- target - width
  buffupp <- target + width
  lim <- target - 2 * width
  slope <- sloperatio / width

  # compute h(score)
  h <- ifelse(score < lim, 0,
    ifelse(score < bufflow,
      slope * (score - lim),
      ifelse(score > buffupp, 1 + sloperatio * (score - buffupp), 1)
```

```
    )
  )

  tac_prev <- catch(stk)[, ac(iy - 1)]
  tac_new <- tac_prev * h

  # apply TAC limits if given
  if (!is.null(dupp)) tac_new <- pmin(tac_new, tac_prev * dupp)
  if (!is.null(dlow)) tac_new <- pmax(tac_new, tac_prev * dlow)

  catch(stk)[, ac(cys)] <- tac_new
  return(list(stk = stk, ind = ind, tracking = tracking))
}

args <- list(ay = 2025, iy = 2026, frq = 1, data_lag = 1, management_lag = 1)
idx <- list("cpue" = FLQuant(rlnorm(30), dimnames = list(year = 1996:2025)))

stk <- FLStock(catch = FLQuant(1000, dimnames = list(year = 2025:2030)))

result <- indscore.z(stk, idx, args = args, tracking = FLQuant(0))
output <- bufferdelta.hcr(result$stk, result$ind, args = args, tracking = result$tracking)

catch(output$stk)

args <- list(ay = 2025, iy = 2026, frq = 1, data_lag = 1, management_lag = 1)
idx <- list("cpue" = FLQuant(rlnorm(30), dimnames = list(year = 1996:2025)))

result <- indscore.z(stk, idx, args = args, tracking = FLQuant(0))
output <- bufferdelta.hcr(result$stk, result$ind, args = args, tracking = result$tracking)
```

Environment Objects

Object Name	Type / Purpose
h1	A list containing the full OM, OEM, and IEM for hypothesis H1 (qs file)
om	Iterated subset of the Operating Model from h1
oem, iem	Observation and implementation error models; extracted from h1

Object Name	Type / Purpose
omperf	Performance metrics of OM alone, usually C, F, SB for conditioning years
perf	Combined data frame of MP simulation performance results
getSlick	Function that merges MP/OM results and constructs a Slick summary object
FLslick	Constructor function that builds and returns a Slick object for plotting
sli	The returned Slick object for visualization (Kobe, Quilt, Spider, etc.)
ctrl	A list of control parameters for MPs (e.g., estimation methods, tuning devs)
condition	*Not found in current project files; possibly a misidentified object*

Helper Functions to Plot Results

```
plot_slick_quilt(sli, stat = "longterm C")
plot_slick_spider(sli, om_idx = 1, mp_idx = 2)
plot_slick_tradeoff(sli, stat = "PSBMSY")
```

Defining Management Procedures using `mpCtrl` and `mseCtrl`

In the `mse` package, Management Procedures (MPs) are modularly constructed using the `mpCtrl` class, which bundles together component controls (`mseCtrl`) for estimation, decision-making, and implementation.

Structure of `mpCtrl`

Components and Options

Component	Description	Example `method`	Typical `args`
est	Estimator for stock status	perfect.sa, shortcut.sa	metric, B0, devs, years

Component	Description	Example `method`	Typical `args`
`hcr`	Harvest Control Rule	`buffer.hcr`, `hockeystick.hcr`, `trend.hcr`	`target`, `lim`, `trigger`, `metric`, `buffer`
`phcr`	Parametrization of HCR (optional)	`parametric.hcr`, etc.	Custom parameters
`isys`	Implementation system (e.g., allocation strategy)	`split.is`	`split = catch_props(...)`, `noise`, `bias`
`tm`	Technical measures (e.g., size limits, gear rules)	–	Optional, rarely used in CJM MSE setup

Summary of `mseCtrl`

An `mseCtrl` object specifies **how a module runs**:

- `method`: a function to apply (e.g., `shortcut.sa`)
- `args`: list of arguments passed to the method

Example

```
mseCtrl(
  method = buffer.hcr,
  args = list(
    target  = 1000,
    bufflow = 0.30,
    buffupp = 0.50,
    lim     = 0.10,
    min     = 0,
    metric  = "depletion"
  )
)
```

Accessors and Replacements

You can modify or retrieve components using:

```
method(ctrl@hcr)
args(ctrl@hcr)$target <- 1200
```

Or, using provided accessors:

```
hcr(ctrl) <- mseCtrl(method = new.hcr, args = list(...))
args(ctrl, "hcr")$target <- 1100
```

This modular framework allows for flexible and transparent testing of MPs within the MSE simulation, including full customization of estimation, control rules, and implementation behavior. This notebook supports the JM MSE development process and is intended for use during scenario comparison, workshop reporting, and trade-off evaluation.

Slick Object Structure

The `Slick` object is the core summary container created by the `getSlick()` and `FLslick()` functions. It contains performance data used for visualization and evaluation across multiple Management Procedures (MPs) and Operating Models (OMs).

Slot	Contents
@Boxplot	MP × OM × performance indicators (boxplots)
@Kobe	SB/SBMSY vs F/FMSY over kobeyrs
@Quilt	Heatmap of average performance
@Spider	Scaled performance for visual trade-offs
@Timeseries	Time series of F, C, SB
@Tradeoff	Mean trade-off indicators (post-OM years)
@MPs, @OMs	Metadata: MP and OM definitions and labels

Creating and Visualizing a Slick Object

```
# Load OM and compute baseline performance
h1 <- qread("data/h1_1.07.qs")
om <- iter(h1$om, seq(100))
omperf <- performance(om, years = 1970:2023, statistics = statistics[c("C", "F", "SB")])
```

```
perf <- readPerformance("demo/performance.dat.gz")

# Combine with MP results (perf), filtering to "tune" runs
sli <- getSlick(perf[grep("tune", run)], omperf, kobeyrs = 2034:2042)
```

Optional: Shorten MP Labels for Plotting

```
shorten_mp <- function(mpnames) {
  gsub("h1_1.07_", "", mpnames) |>
    gsub("cpue2", "C2", .) |>
    gsub("cpue36", "C36", .) |>
    gsub("cpue3", "C3", .) |>
    gsub("scbad", "SCB", .) |>
    gsub("scgood", "SCG", .) |>
    gsub("scmedium", "SCM", .) |>
    gsub("buffer", "BUF", .) |>
    gsub("hcr_target_", "HCR_T", .) |>
    gsub("tune", "T", .)
}
perf[, mp := paste0("H1_", shorten_mp(mp))]
```