Jack mackerel management strategy evaluation for SC14

Operating models, candidate management procedures, robustness tests, and matters for Scientific Committee consideration

2026-07-24

Jack Mackerel Management Strategy Evaluation: Synthesis for SC14

Status and purpose. This report synthesizes the June 2026 SCW17 MSE workshop and subsequent model and management-procedure development.

Executive summary

The jack mackerel MSE has progressed from a broad workshop design into a narrower comparison of candidate management procedures (CMPs) tested against a reference operating model (OM) and a structured set of robustness OMs. SCW17 established the principal design choices: simple and interpretable empirical rules, simulation of correlated index errors, dynamic reference points, explicit robustness testing, a broad performance-measure set, and development of an exceptional-circumstances protocol. Post-workshop work corrected and extended the implementation, tuned the leading rule families, and examined their performance under one-stock and two-stock hypotheses, recruitment stress, movement, selectivity, and alternative annual TAC-change limits.

A novel feature of this MSE is its treatment of the five abundance indicators as related observations rather than independent pieces of evidence. The OMs simulate their observation errors jointly using historical estimates of each series' uncertainty, temporal autocorrelation, and cross-correlation. This allows the evaluation to represent shared movements among indicators, including correlation that may arise when the population shifts spatially.

Among the dozens of MPs evaluated at the SCW17 workshop, the analysts concluded that the hockey-stick and power-ramp designs represented by these two CMPs, together with reasonable variants of those designs, were suitable to carry forward as proposals for consideration at SC14. This analytical conclusion identifies a focused set for Scientific Committee consideration; it is not a formal recommendation or adoption decision.

Two CMPs provide the central comparison:

- **HS+20** uses a hockey-stick harvest control rule (HCR), a 2,000-kt catch target, a 270-kt minimum, and a combined indicator from five standardized abundance indices.
- **PR+20** uses a power-ramp HCR, a 1,500-kt catch target, the same minimum and indicator, and a more responsive reduction in advice as the indicator declines.

These two CMPs anchor the central rule-shape comparison, but they were not the only procedures examined in detail. Six retuned annual-limit variants were also evaluated under the reference and robustness OMs: **HS-20**, **HSsym**, **HS-30**, **PR-20**, **PRsym**, and **PR-30**. They test reverse-asymmetric, symmetric, and larger-decrease constraints around the two central HCR families. The paired catch trajectories for HS+20 and HS-20 illustrate one such annual-limit contrast in Figure 1.

Both were tuned under the reference OM to an approximately 60% mean probability of dynamic Kobe-green status during 2041–2050. In the comparisons reviewed on 24 July, HS+20 generally provided higher short-term catch and lower interannual catch variability, whereas PR+20 responded more quickly to declining indices and showed comparatively stronger recovery under the recruitment-crash and cyclic-recruitment scenarios. Differences were often modest, depended on the OM and performance measure, and do not establish a preferred or "optimal" CMP. Reference-OM vulnerable-biomass performance across the evaluated variants is summarized in Figure 2.

The most consequential robustness finding came from the two-stock hypotheses. Without movement, the southern component passed the agreed benchmark while the northern component required an exploratory change of about 3.07 percentage points in simulated catch distribution from north to south to reach the same benchmark. This is a sensitivity result, not an allocation recommendation. With movement, no tested fixed catch distribution achieved the long-term benchmark for both components. Equal weighting of indices could also mask a declining southern component when northern indices increased. Separate annual tracking of northern and southern indices is therefore an important candidate element of exceptional-circumstances monitoring.

The Scientific Committee is asked to consider the adequacy of HS+20 and PR+20 as the central rule-shape comparison; the annual TAC-change variants to retain; the role of all five indices, including the only fishery-independent series; the interpretation of two-stock robustness; the use of final 500-replicate runs; and the work needed to complete banking-and-borrowing and exceptional-circumstances evaluation.

Purpose, scope, and decision status

The MSE is intended to compare complete management procedures in a closed-loop simulation before one is considered for implementation. Each CMP converts agreed observations into annual catch advice; each OM represents a plausible description of stock and fishery dynamics against which the CMP is tested. Performance measures describe biological outcomes, yield, stability, and fishery feasibility over specified periods.

This report melds the technical foundation assembled in the benchmark-preparation and MSE-planning materials, the design record from the SCW17 MSE Workshop report, and the post-workshop analysis summarized in the 24 July task-team meeting report and 25 July progress presentation. The preparation site preceded SCW17 and provided the detailed model specifications, diagnostics, working papers, and technical discussions needed to update the operating models used at SCW17 and in subsequent MSE work. Their different roles are summarized in Table 1.

Table 1: Source materials and their roles in this synthesis.

Source	Role in this synthesis
Benchmark preparation and MSE planning	Pre-SCW17 benchmark papers, OM specifications, diagnostics, index analyses, and preparatory meeting records that supported updates to the OMs used at SCW17 and beyond.
SCW17 workshop report, 15–19 June 2026	Agreed MSE design, OM and CMP principles, reference-point approach, performance framework, and planned robustness work.
Post-workshop `jmMSE` analysis	Implemented CMP definitions, tuning runs, reference and robustness simulation outputs, and reproducibility records.
Task-team meeting, 24 July 2026	Interpretation of results, wording cautions, issues for SC14, and agreed follow-up.
Progress presentation dated 25 July 2026	Detailed tuning sequence, OM contrasts, and graphical robustness diagnostics available at the time of review.
Draft weighted scorecard, 23 July 2026	Example of an editable multi-criteria comparison method; not an exhaustive candidate list, metric list, or decision.

The Scientific Committee provides scientific advice and the Commission weighs management trade-offs and makes management decisions. A shortlist in an MSE analysis is not an adoption decision.

Development from SCW17 to the current analysis

SCW17 continued work begun at SCW15 and the 2026 benchmark. The benchmark-preparation and MSE-planning site preserves the detailed model specifications, diagnostics, index working papers, and preparatory meeting discussions that preceded SCW17. That record provided the technical basis for updating the operating models taken into the workshop and used in subsequent MSE analyses.

SCW17 aimed to prepare an OM set, trial CMPs, performance statistics, robustness scenarios, and documentation for SC14. The workshop favored simple empirical rules because they are transparent, reproducible, and can be locked before use. Trial CMPs were used to learn about rule behavior and were not final recommendations.

Post-workshop work made the following material advances:

1. Index projection errors were extended to include iteration-specific variability, autocorrelation, and cross-correlation among index deviations.
2. Dynamic biological reference points and additional measures, including vulnerable-biomass and exploitation measures, were incorporated.
3. A straight-slope comparison rule was added.
4. The power-ramp implementation was corrected to remove discontinuities around its limit. The correction was required for consistent implementation but did not overturn the broad workshop conclusions.
5. A projected effort constraint was added to avoid implausible catches that would require extreme effort.
6. CMP tuning was performed as a nested screen of HCR family, catch target, minimum catch, index limit, index set, simulation replication, and robustness performance.
7. Alternative annual TAC-change limits were retuned instead of being compared using triggers tuned under a different constraint.

The 25 July presentation used the word "optimal" for some within-family runs. The 24 July meeting clarified the interpretation: those runs were useful within-family representatives, but the analysis does not establish an optimal procedure across management objectives.

Operating models and simulation methods

Reference and robustness set

The current grid centers on the one-stock benchmark model `h1_0.16` with steepness 0.65. Robustness OMs test major structural and parametric uncertainties rather than assigning probability to one forecast (Table 2).

Table 2: Reference and robustness operating-model set.

OM label	Main contrast	Analytical role
om11	One stock, model 0.16, steepness 0.65	Reference OM for current tuning

OM label	Main contrast	Analytical role
om11_1	Alternative selectivity	Younger/smaller fish exposed to greater fishing pressure
om11_2	One-year recruitment crash	Acute recruitment stress
om11_3	Seven-year cyclic recruitment reduction	Prolonged recruitment stress
om12	One stock, steepness 0.8	Productivity robustness
om13	One stock, SC13-era/model 1.14 configuration	Assessment/configuration robustness
om21	Two stocks, no movement	Stock-structure and catch-distribution robustness
om21_1	Two stocks with movement	Spatial redistribution and indicator-conflict robustness
om22	Two stocks, steepness 0.8	Two-stock productivity robustness
om23	Two stocks, model 1.14 configuration	Two-stock assessment/configuration robustness

The OM set (reference and robustness) is not a declaration that all states are equally plausible. Its purpose is to expose CMP sensitivity to assumptions that could matter to management outcomes. Any later OM weighting should be explicit and documented.

Recruitment, observation, and implementation uncertainty

Recruitment variability strongly influences stock trajectories and dynamic reference points. The core simulations therefore retain recruitment uncertainty, while discrete crash and cyclical-reduction cases act as stress tests. SCW17 treated ENSO and climate effects as robustness questions at this stage rather than embedding a single climate forecast in the core OM.

Each CMP receives simulated abundance indices rather than true stock status. Projected index deviations include autocorrelation and cross-correlation so that the observation model better represents the historical behavior of multiple surveys and CPUE series. Simulation diagnostics should show the distributions of index-error standard deviations, autocorrelation parameters, and correlations used.

Dynamic reference points

Environmental recruitment variability can make static reference points misleading. SCW17 strongly recommended calculating the ratio B_{MSY}/B_0 within each MCMC draw, multiplying it by the corresponding dynamic B_0, and evaluating spawning biomass relative to that dynamic B_{MSY}. The current tuning objective uses a dynamic Kobe-green measure that combines biomass and fishing-pressure status.

Reference-point choices remain consequential. In particular, alternative selectivity can change F_{MSY} and therefore F/F_{MSY} even when biomass and catch trajectories are similar. Interpretation should separate changes caused by reference-point definitions from changes in underlying CMP behavior.

Candidate management procedures

Common indicator and annual calculation

The central CMPs use five observed series: northern Chile acoustic biomass, Chilean CPUE, Peruvian artisanal and industrial indices, and offshore CPUE. Each series is expressed relative to its 2019–2023 mean. The standardized values are combined with equal weights within each year, and the CMP uses a three-year recent average.

> **!** Novel treatment of correlated indicators
>
> The combined indicator does not assume that five series provide five independent confirmations of stock change. In each OM iteration, projected index errors are generated jointly from the historical index deviations: series-specific log-scale standard deviations describe uncertainty, lag-1 parameters preserve temporal autocorrelation, and an iteration-specific correlation matrix preserves cross-correlation among indices. These relationships matter because common environmental effects, fleet behavior, and spatial redistribution of the population can cause several indices to rise or fall together.
> The resulting closed-loop test therefore exposes the CMP to realistic combinations of agreement, redundancy, and divergence among indicators. The novelty lies in evaluating the simple, transparent equal-weight composite under this estimated multivariate observation process; equal weighting itself is not covariance weighting and does not mathematically remove correlation.

Most data are assumed to arrive with a one-year lag, while offshore CPUE has a two-year lag. Under the tested timing, a calculation during 2026 for 2027 advice would use 2025 values for the acoustic, Chilean CPUE, and Peruvian indices, and the 2024 offshore CPUE value. Same-year 2026 observations do not enter that tested 2027 calculation unless the timing rule is explicitly changed and re-evaluated.

Naming convention and candidate variants

The working labels encode the HCR family and the most distinguishing annual TAC-change constraint. `HS` means hockey-stick and `PR` means power ramp. `+20` identifies the base 20% maximum increase, `-20` identifies the reverse asymmetry with a 20% maximum decrease, `sym` identifies symmetric 15% limits, and `-30` identifies a 30% maximum decrease. These labels should be used in figures and interpretive prose. The legacy MP identifiers remain in the table only to preserve the link to code, saved outputs, and earlier documents. The complete mapping is given in Table 3.

Central rule-shape comparison

All eight CMPs use the same five-index mean, 2019–2023 reference period, three-year averaging window, 270-kt minimum catch, 0.1 index limit, and 1,385-kt initial simulated catch. The six non-base variants were retuned under the reference OM after changing the annual constraint.

Table 3: Naming convention, legacy identifiers, and definitions of the eight candidate variants.

Working label	Legacy ID	HCR family	Target or breakpoint (kt)	Tuned trigger	Annual decrease/increase limits
HS+20	MP29	Hockey-stick	2,000	2.125	-15% / +20%
HS-20	MP43	Hockey-stick	2,000	3.0625	-20% / +15%
HSsym	MP45	Hockey-stick	2,000	3.0625	-15% / +15%
HS-30	MP47	Hockey-stick	2,000	3.0625	-30% / +20%
PR+20	MP32	Power ramp	1,500	1.421875	-15% / +20%
PR-20	MP44	Power ramp	1,500	1.9140625	-20% / +15%
PRsym	MP46	Power ramp	1,500	1.9375	-15% / +15%
PR-30	MP48	Power ramp	1,500	1.890625	-30% / +20%

The hockey-stick increases advice linearly between the limit and trigger and then reaches a plateau. The power ramp reduces catch more quickly when the indicator is below its trigger and can continue to increase above its stated breakpoint according to the implemented rule. These structural differences, not the labels alone, create the main responsiveness-versus-stability contrast.

The progress presentation also reported separate 500-replicate tuning checks for the central HS and PR configurations. Their triggers were close to the 100-replicate results, and the reported performance differences were small. They are replication checks, not new conceptual CMPs. Final candidate summaries should use the completed 500-replicate results when available.

Annual TAC-change variants

The base CMPs limit annual advice to a 15% decrease and a 20% increase. The task team identified a limited set of alternatives for retuning and comparison, summarized in Table 4.

Table 4: Annual TAC-change-limit variants by HCR family.

HCR family	Base	Reverse asymmetry	Symmetric	Larger decrease
Hockey-stick	HS+20: -15%/+20%	HS-20: -20%/+15%	HSsym: -15%/+15%	HS-30: -30%/+20%
Power ramp	PR+20: -15%/+20%	PR-20: -20%/+15%	PRsym: -15%/+15%	PR-30: -30%/+20%

The variants change both management responsiveness and the tuned trigger. They should therefore be compared as complete, retuned CMPs. The available 100-replicate reference and nine-OM robustness runs support exploratory comparison; they do not replace final higher-replication evaluation of the reduced set.

The following reference-OM plot shows 30 reproducibly selected simulation iterations for the two hockey-stick asymmetry cases over 2025–2050. Iteration IDs are paired exactly across panels: for example, iteration 6 would be shown for both MPs if selected, and every selected ID occurs once in each panel. Each iteration has its own line color, and the same iteration uses the same color in both panels. HS-20 is MP43 (`tun43`), with annual limits of -20%/+15%; HS+20 is MP29 (`tun29`), with limits of -15%/+20%. The vertically stacked panels share a common catch scale, as shown in Figure 1.

Choice of indices

A CPUE/fishery-index-only comparison was more responsive but produced slightly lower short-term catch and higher short-term catch variability. The task team retained all five indices for the current central comparison. The northern Chile acoustic survey is the only fishery-independent abundance index in the set and provides information that fishery-dependent series cannot fully replace. Conversely, equal weighting across regions can become problematic if northern and southern components diverge. These two findings support retaining the acoustic series while also monitoring regional index groups separately.

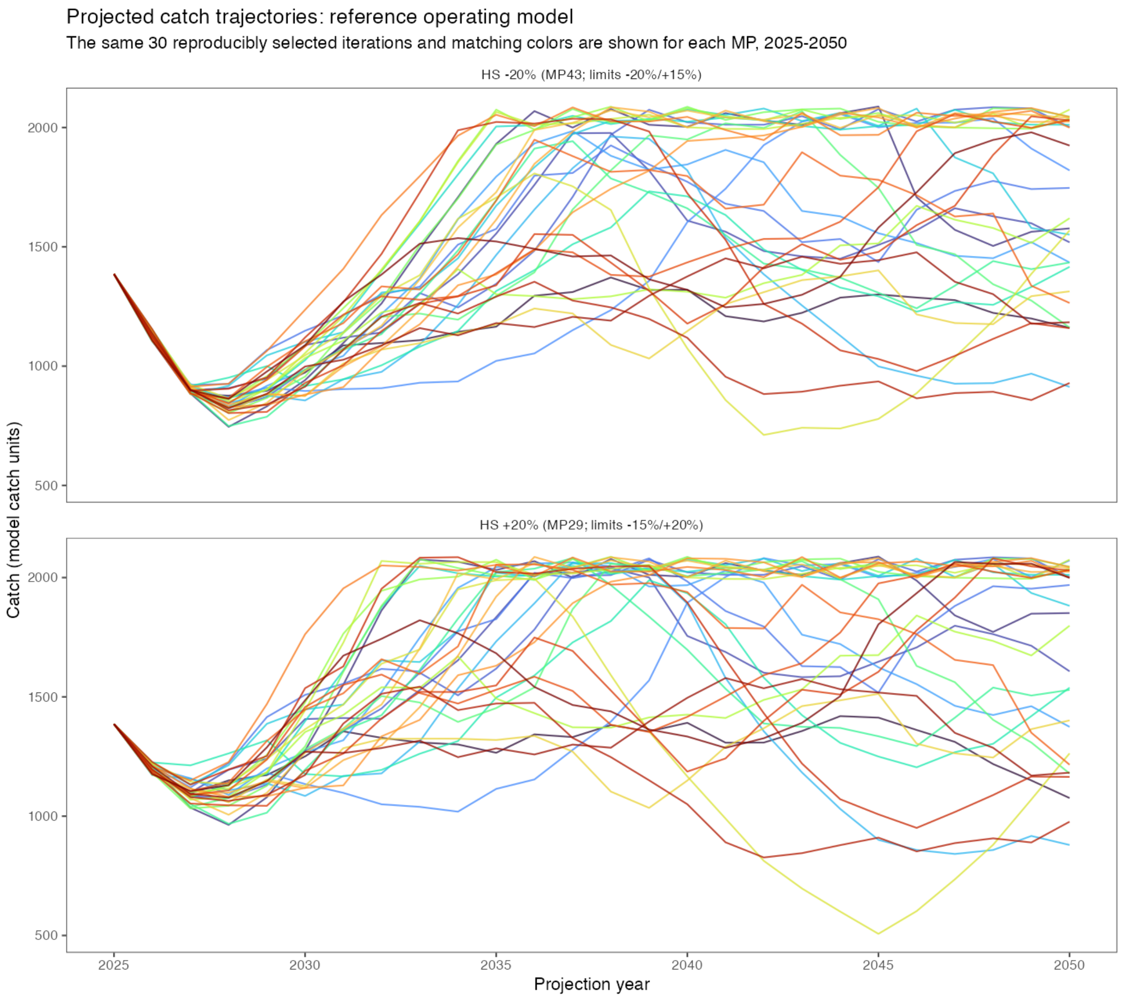

Figure 1: Projected catch trajectories for the same 30 simulation iteration IDs under both hockey-stick MPs in the reference operating model. Each iteration retains the same line color in both panels so matching trajectories can be traced through time and across MPs.

Performance evaluation and the draft scorecard

Performance framework

SCW17 agreed that CMPs should be evaluated across safety, yield, stability, and feasibility rather than by the tuning target alone. Tuning several CMPs to the same Kobe-green objective makes that objective a control for comparison, not a sufficient ranking statistic. The full performance framework can include:

- probabilities of dynamic Kobe-green, low biomass, overfishing, and stock collapse;
- spawning biomass, vulnerable biomass, and fishing pressure relative to reference points;
- short-, medium-, and long-term catch;
- interannual catch change and the frequency with which TAC-change limits bind;
- low-catch, minimum-catch, and fishery-shutdown-like outcomes;
- effort or catchability feasibility diagnostics; and
- robustness measures across individual OMs and stress tests.

A short headline table can aid Scientific Committee review, but the complete metric set and underlying distributions should remain available in the technical annex and interactive outputs.

The augmented candidate exports include vulnerable biomass relative to its 2025 level and relative to equilibrium vulnerable biomass at the fishing mortality that produces MSY. Vulnerable biomass was selected to provide a performance measure related to the fish available to the fishing fleet under the modeled selectivity, complementing spawning-biomass measures of biological stock condition. Reference-OM results are shown in Figure 2 and the corresponding robustness results in Figure 3.

Performance metrics by CMPs

The quilt in Figure 4 is an unweighted comparison of the eight annual-change variants listed in Table 3 under the reference OM. Annual values were averaged over 2041–2050 within each simulation iteration; cell labels are medians of those iteration means. Color is normalized independently within each metric across the eight variants.

Purple means lower relative preference and light lavender means higher relative preference (0 is darkest; 1 is lightest). Higher SB/SB_{MSY}, catch, VB/VB_{2025}, and VB/VB_{MSY} are treated as better; lower F/F_{MSY} and lower IACC are treated as better. Lower F/F_{MSY} represents lower fishing pressure, while catch is shown separately. The within-column color scale is not an absolute biological or management threshold and must not be interpreted as a rank or recommendation.

The underlying values, preferred directions, metric ranges, and normalized scores are saved in `candidate_quilt_reference_summary.csv`. An editable scorecard input layer is saved in

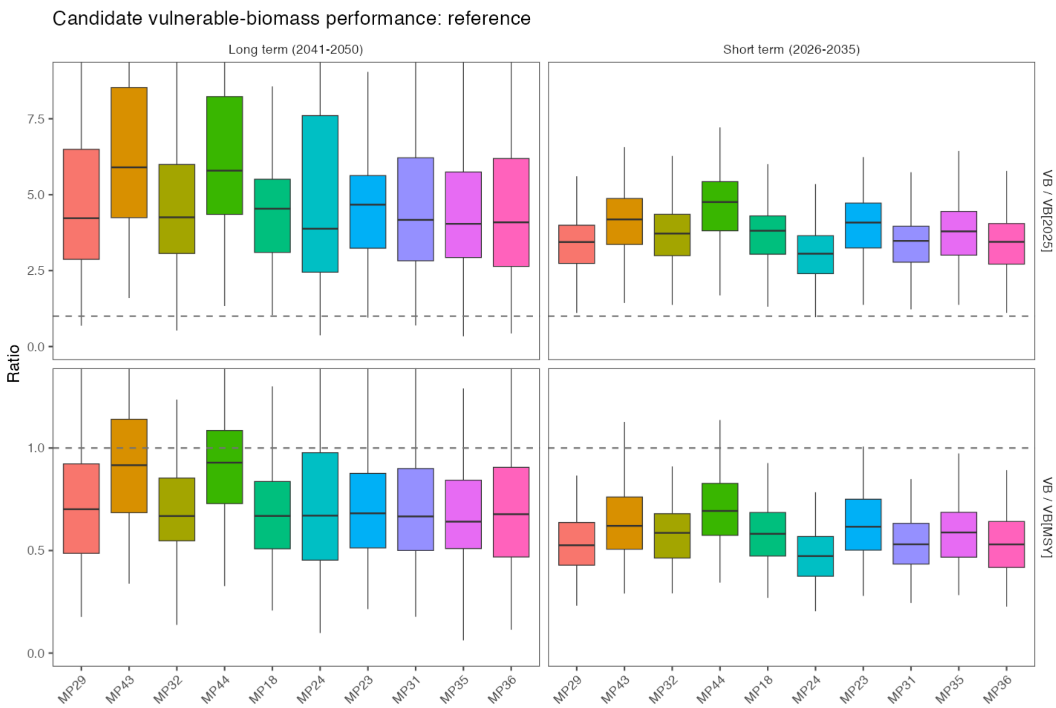

Figure 2: Vulnerable-biomass performance for the reference operating model.

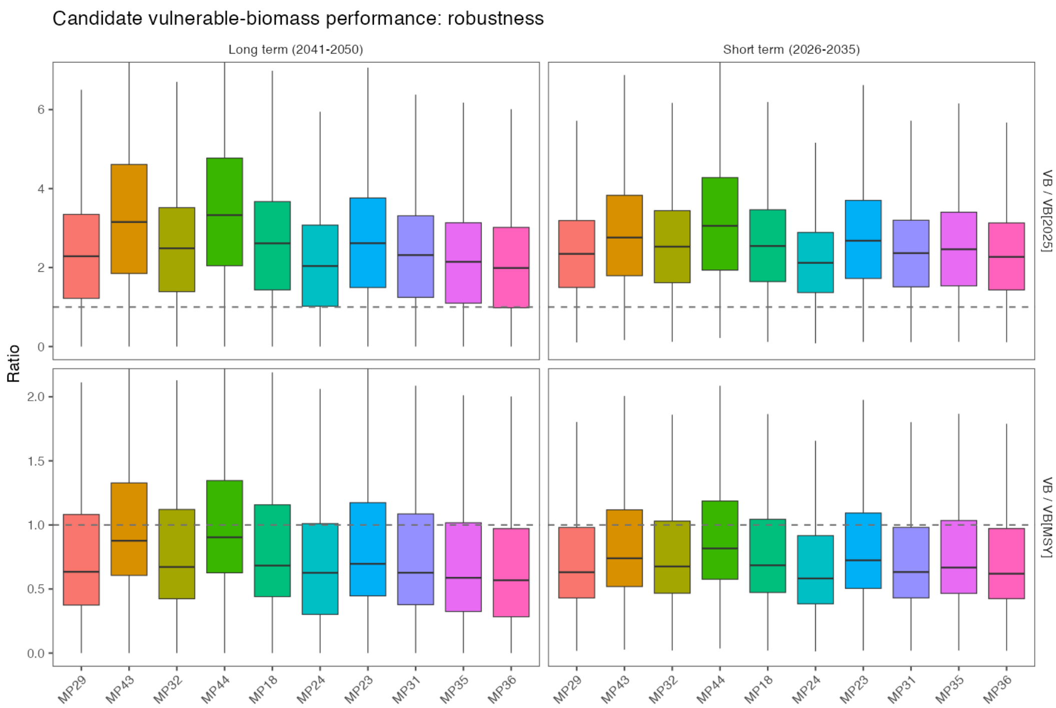

Figure 3: Vulnerable-biomass performance across the robustness operating models.

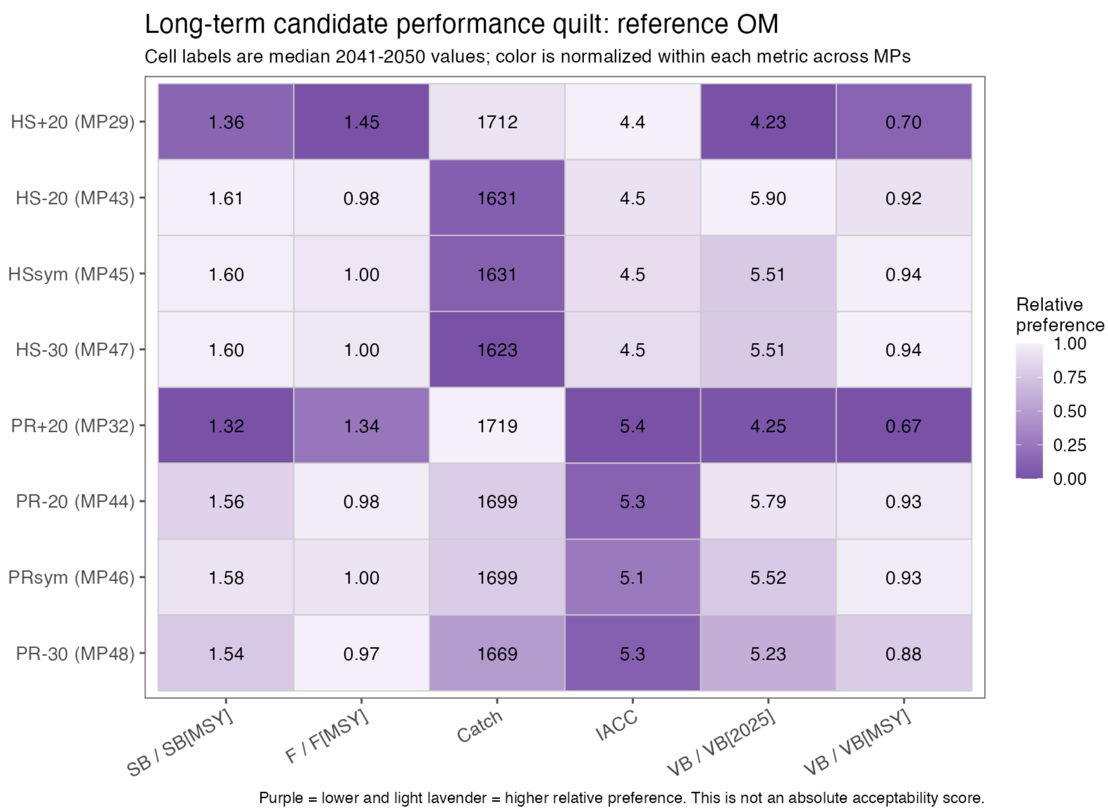

Figure 4: Reference-OM candidate performance quilt for the eight naming-convention variants. Cell labels are median 2041–2050 values. Purple indicates lower relative preference and light lavender indicates higher relative preference within each metric (0 is darkest; 1 is lightest).

`candidate_scorecard_input_reference.csv`. It converts quilt relative preference to a 0–100 benefit score, records the preferred direction and raw value, and supplies editable `include` and `weight` columns. The default weight is one for every metric, but no weighted total is calculated in this export. Changing weights or aggregating scores remains an explicit user decision under the method below.

Extensible weighted-scorecard method

The performance quilt in Figure 4 now supplies a reproducible worked scorecard for the eight candidate variants under the reference OM. Let x_{jm} be the value for CMP j and metric m, and let $x_m^{\min}$ and $x_m^{\max}$ be the minimum and maximum across the included CMPs. The directionally normalized preference score p_{jm} is

$$p_{jm} = \begin{cases} \dfrac{x_{jm} - x_m^{\min}}{x_m^{\max} - x_m^{\min}}, & \text{if higher values are preferred,} \\ \dfrac{x_m^{\max} - x_{jm}}{x_m^{\max} - x_m^{\min}}, & \text{if lower values are preferred.} \end{cases}$$

Thus $p_{jm} = 0$ is the least preferred observed value and $p_{jm} = 1$ is the most preferred observed value for that metric and comparison set. If all CMPs have the same value for a metric, they are tied at $p_{jm} = 1$. With inclusion indicator I_m and non-negative weight w_m, the aggregate score is

$$S_j = 100 \, \frac{\sum_m I_m w_m p_{jm}}{\sum_m I_m w_m}.$$

The contribution of metric m to the total is $C_{jm} = 100 I_m w_m p_{jm} / \sum_m I_m w_m$, so $S_j = \sum_m C_{jm}$. The R implementation reads the editable Figure 4 input, checks that every CMP has one value for every included metric, and writes both the totals and metric-level contributions.

The worked results in Table 5 include all six metrics from Figure 4. The equal-weight view assigns one-sixth to every metric. The balanced view assigns half the total weight to fishing performance—25% each to Catch and IACC—and half to stock condition—12.5% each to SB/SB_{MSY}, F/F_{MSY}, VB/VB_{2025}, and VB/VB_{MSY}. All inputs use the reference OM and the 2041–2050 summary window used in Figure 4.

Table 5: Equal-weight and balanced scorecard results for the eight performance-metric-quilt CMPs under the reference OM, 2041–2050. The balanced score assigns 25% each to Catch and IACC and 12.5% each to SB/SBMSY, F/FMSY, VB/VB[2025], and VB/VB[MSY].

Equal-weight rank	CMP	Equal-weight score	Balanced score	Balanced rank
1	HS-20 (MP43)	81.65	73.75	1
2	HSsym (MP45)	78.02	70.77	3
3	PRsym (MP46)	77.69	71.42	2
4	PR-20 (MP44)	76.70	68.79	4
5	HS-30 (MP47)	76.61	68.52	5
6	PR-30 (MP48)	62.01	53.66	6
7	HS+20 (MP29)	36.29	51.28	7
8	PR+20 (MP32)	20.93	28.20	8

The machine-readable totals are in `candidate_scorecard_result_reference.csv`, and the individual weighted contributions are in `candidate_scorecard_contributions_reference.csv`. These values are a reproducible calculation, but both orders remain **illustrative rather than management rankings**: neither weighting scheme is an agreed statement of management priorities, and the scores are relative to this candidate set.

The current Figure 4–based worked scorecard is an **illustrative framework**, not the final scorecard. It contains eight CMPs and six performance metrics. Please note:

- candidate rows can be added, removed, or replaced as the shortlist evolves;
- performance columns can be added as further safety, yield, stability, feasibility, and robustness metrics are validated;
- preferred direction, units, periods, transformation, and weight must be recorded for every metric; and
- normalization must be recalculated whenever the candidate set changes.

Min–max normalization is relative to the alternatives included. Adding or removing a CMP can therefore change every normalized score even when underlying MSE results do not change. Weights express management priorities and should be agreed and recorded rather than presented as scientific constants. A weighted score is a decision aid, not a recommendation, and should always be shown with the underlying metric values and sensitivity analyses.

The draft example also includes a “green margin” based on mean biomass and fishing-pressure ratios. That diagnostic is not the simulated probability of being in the Kobe-green quadrant and must not replace the full-distribution MSE probability metric.

Main performance and robustness findings

Reference comparison

The preliminary reference-OM scorecard in Table 5 provides the current eight-CMP comparison. HS-20 ranks first under both the equal-weight and balanced schemes. HSsym and PRsym form the next group: HSsym ranks second with equal weights, while PRsym ranks second when Catch and IACC jointly receive half of the total weight. PR-20 and HS-30 follow closely under both schemes, then PR-30. The two base variants, HS+20 and PR+20, receive lower composite scores because their relatively strong catch or stability outcomes are offset by weaker values for several stock-condition metrics.

This ordering is preliminary and relative to the CMPs, metrics, directions, 2041–2050 window, and weights included in the scorecard. It does not mean that one CMP dominates every biological, yield, stability, and responsiveness objective. The underlying values and trade-off figures should therefore remain the basis for interpreting why the order changes with management priorities.

Reference set trade-offs

Vulnerable biomass. The near-term reference comparison uses mean annual catch and mean VB/VB_{MSY} over 2026–2035 (Figure 5). Points are CMP medians and bars are marginal 5th–95th percentiles.

The long-term reference comparison uses mean annual catch and mean VB/VB_{2025} over 2041–2050 with the same aggregation (Figure 6).

Spawning biomass. The reference-set yield–spawning-biomass comparison uses mean catch during 2026–2030 and mean SB/SB_{MSY} during 2036–2040. Points show means among 100 replicates, and horizontal and vertical bars show the separate 10th–90th percentile ranges for the two measures.

Figure 7 shows the reference-OM pattern. The less restrictive base configurations lie toward higher near-term catch and lower later SB/SB_{MSY}, while variants allowing a larger annual decrease or smaller annual increase generally shift toward lower near-term catch and higher later biomass. The uncertainty ranges overlap substantially, so the plot should not be read as a deterministic ranking.

Figure 8 provides a complementary reference-OM view of biological status and catch stability. The horizontal axis is the replicate mean SB/SB_{MSY} during 2036–2040. The vertical axis is mean IACC during the same period, where annual IACC is $100|C_t/C_{t-1} - 1|$: the plotted period therefore covers the five annual changes from 2035–2036 through 2039–2040. Lower IACC indicates more stable catch. Points are means among 100 replicates; horizontal and vertical bars are the separate 10th–90th percentile ranges of the replicate-level period means.

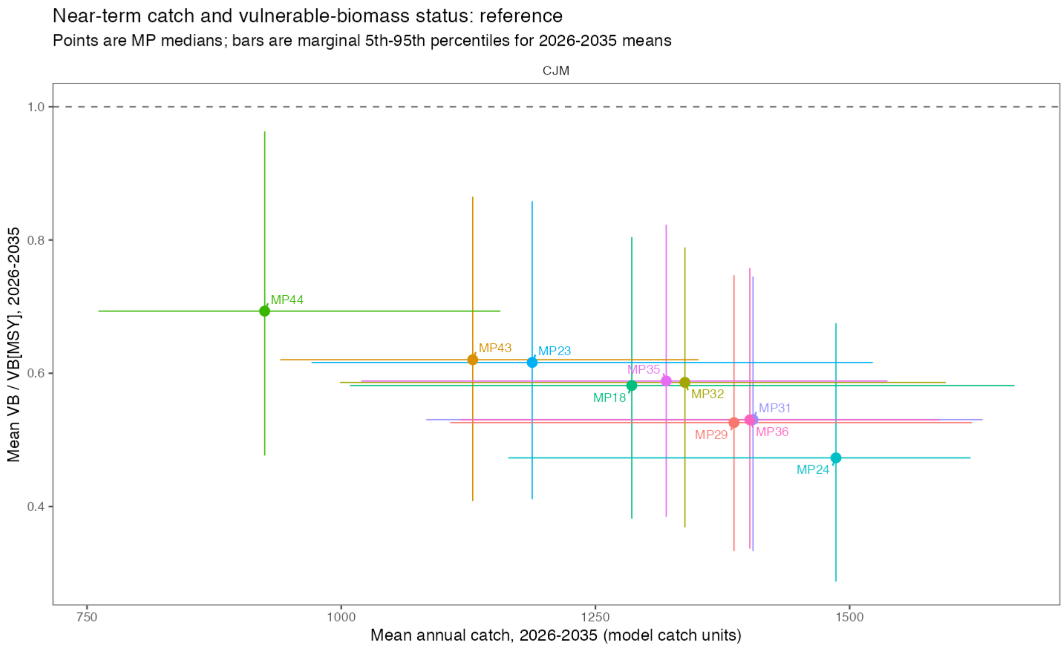

Figure 5: Near-term catch versus VB/VBMSY for the reference operating model, 2026–2035.

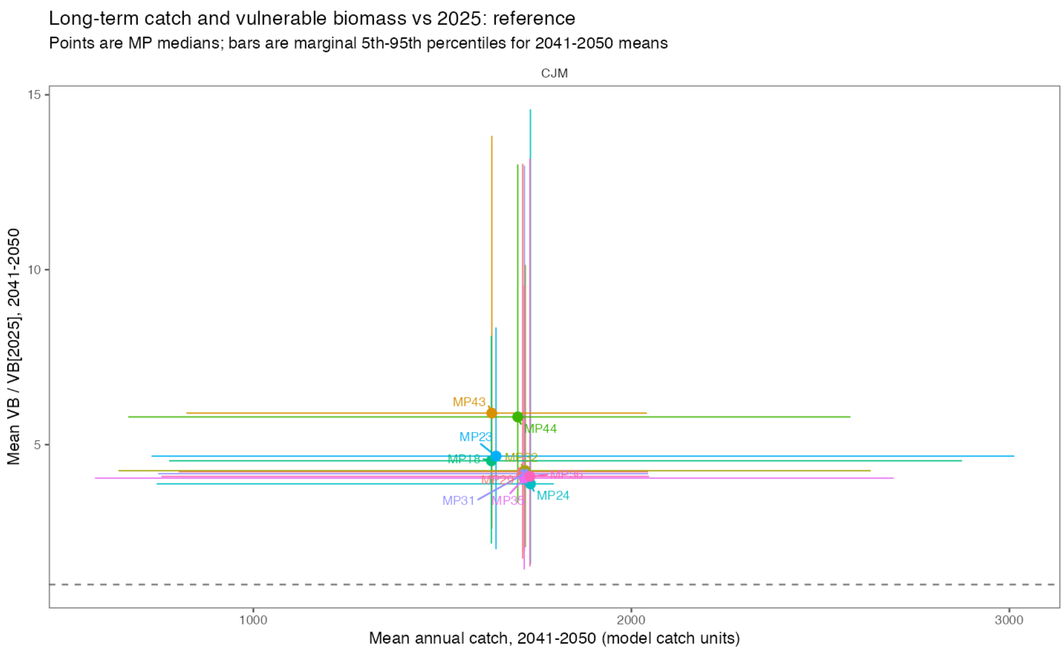

Figure 6: Long-term catch versus VB/VB2025 for the reference operating model, 2041–2050.

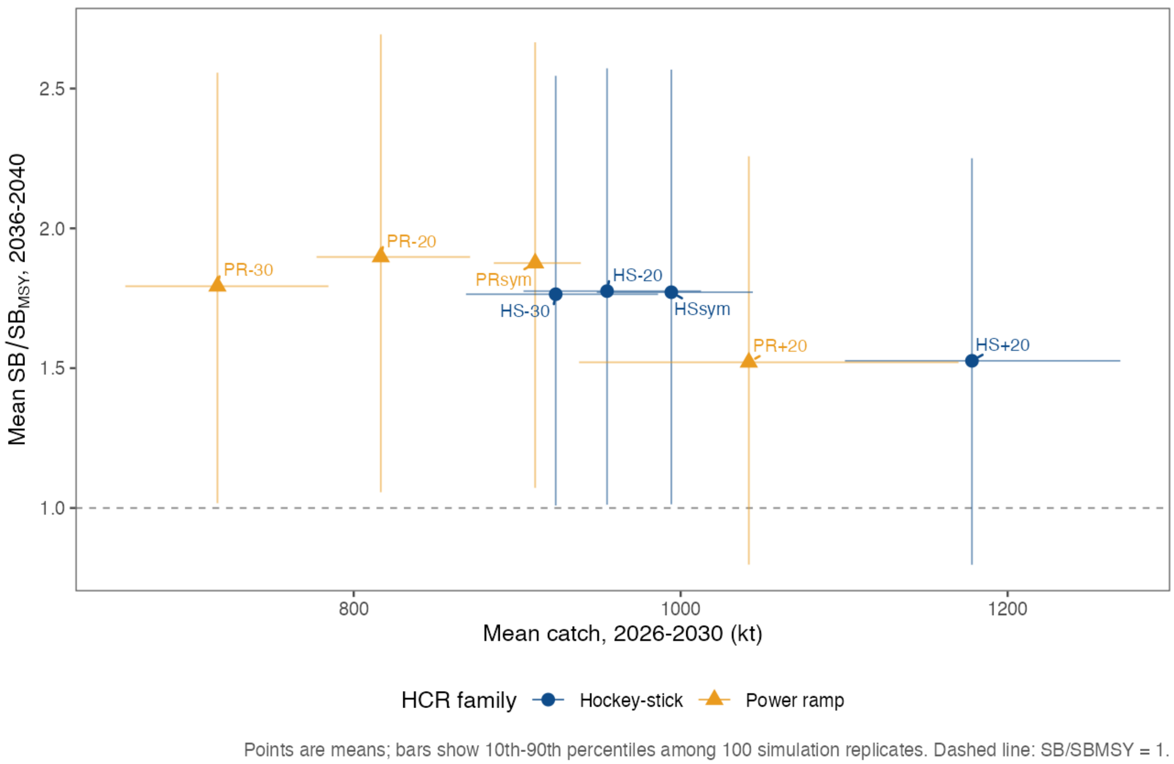

Figure 7: Preliminary reference-OM trade-off for the eight HS and PR configurations.

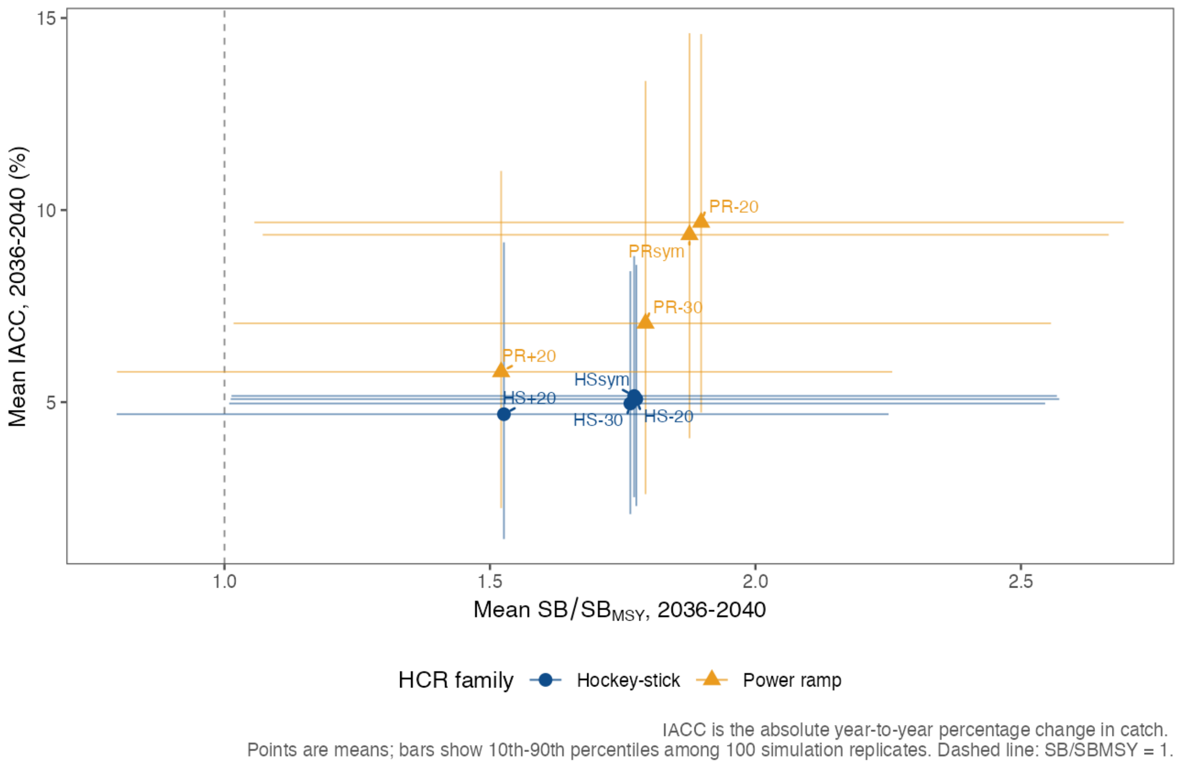

Figure 8: Preliminary reference-OM spawning-biomass and interannual-catch-variability trade-off for the eight HS and PR configurations.

As in Figure 7, overlap is substantial and the marginal bars do not show within-replicate covariance.

Trade-offs under robustness cases

Recruitment stress

The discrete recruitment crash and seven-year cyclic reduction lowered long-term catch and increased catch variability for both central CMPs. PR+20 responded more quickly and produced comparatively stronger long-term recovery in these tests, while HS+20 retained the relative stability advantages seen in the reference case. These are robustness contrasts, not forecasts that either recruitment path will occur.

Alternative selectivity

The alternative-selectivity OM exposed younger and smaller fish to greater fishing pressure. Similar catch levels could then occur with lower spawning biomass and poorer Kobe performance. HS+20 showed a somewhat better long-term trade-off than PR+20 in the presented results, but differences were modest and were affected by the changed F_{MSY} reference point.

Two-stock hypotheses and catch distribution

In the two-stock OM without movement, the southern component remained above the agreed benchmark while the northern component did not. An exploratory shift of approximately 3.07 percentage points in simulated catch distribution from north to south brought both components to the 60% long-term benchmark. This illustrates sensitivity to spatial exploitation; it is not an allocation or catch-distribution recommendation.

In the movement OM, biomass transferred from the southern to the northern component. Because the five indices were combined with equal weights, increasing northern indices could raise the overall indicator while southern spawning biomass declined. No tested time-invariant catch distribution resolved the long-term benchmark for both components. The test is deliberately stressful and does not establish that this movement process is occurring. It does establish a management-relevant warning signal: persistent divergence between northern and southern indices.

Vulnerable biomass. The near-term catch and VB/VB_{MSY} comparison across robustness OMs is shown in Figure 9. It uses the same 2026–2035 period, CMP medians, and marginal 5th–95th percentiles as the reference comparison.

The long-term robustness comparison in Figure 10 uses mean annual catch and mean VB/VB_{2025} over 2041–2050.

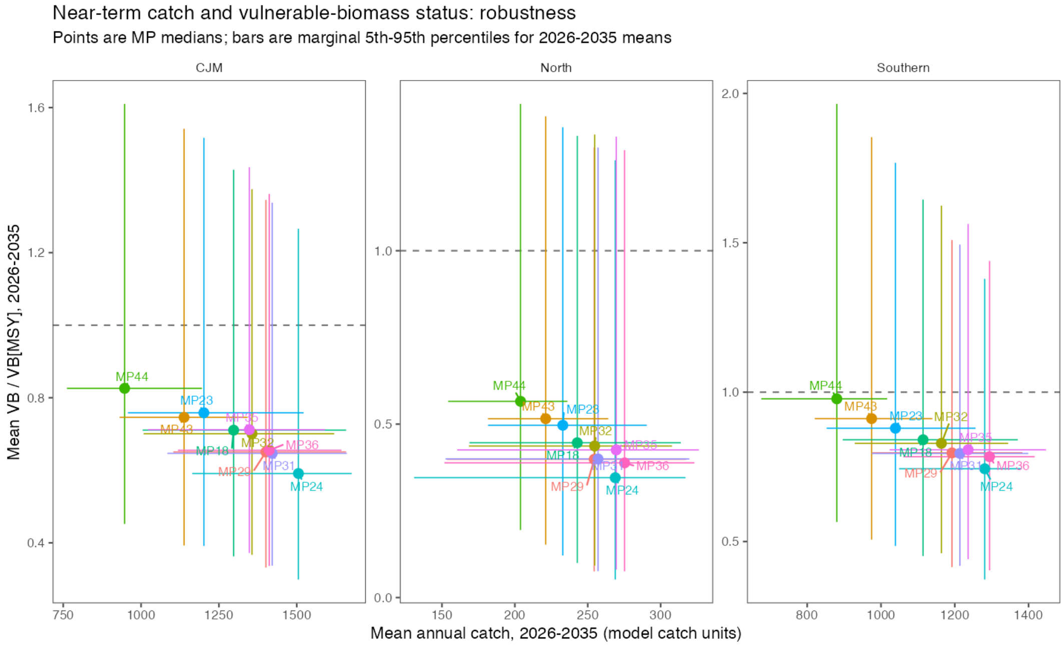

Figure 9: Near-term catch versus VB/VBMSY across robustness operating models, 2026–2035.

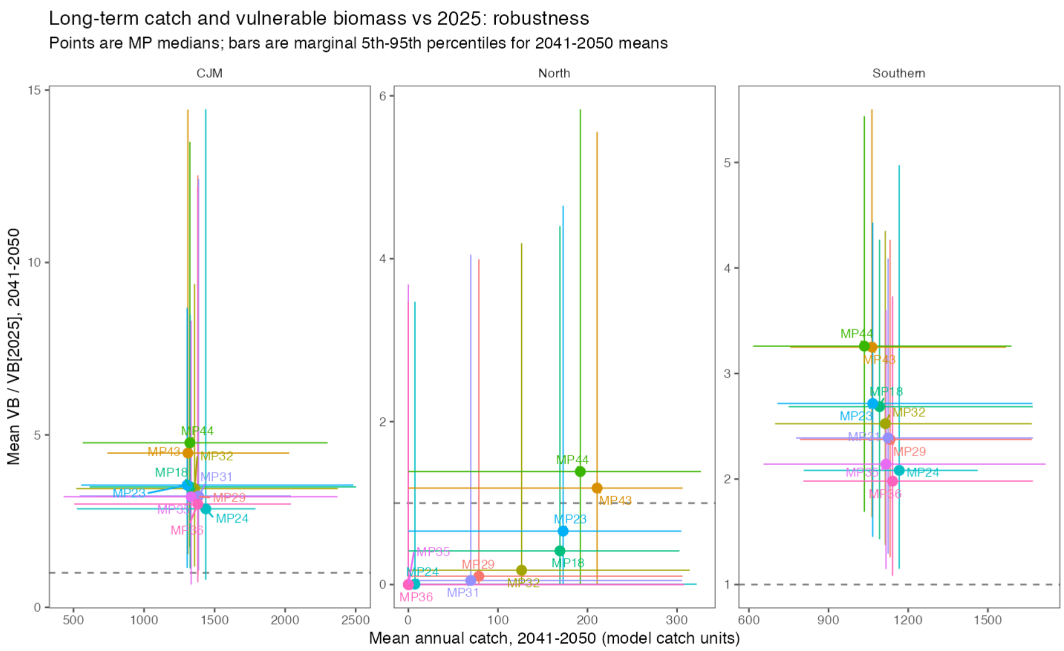

Figure 10: Long-term catch versus VB/VB2025 across robustness operating models, 2041–2050.

Spawning biomass. The saved results also support preliminary yield–spawning-biomass comparisons across the single-stock and two-stock robustness cases.

The same broad direction is visible across the single-stock robustness cases (Figure 11), but the location and separation of configurations changes by OM. Recruitment stress reduces later biomass and changes the apparent trade-off, while productivity and assessment-configuration alternatives move both axes.

For the two-stock cases (Figure 12), each row is a biological component and the horizontal axis is **component catch**, not total catch. The plots therefore diagnose stock-specific consequences and must not be interpreted as an allocation comparison. They reinforce the need to keep northern and southern outcomes separate.

The spawning-biomass figures in Sections 6.2 and 6.3 are preliminary for four reasons. First, they use the completed 100-replicate runs, not the planned final higher-replication comparison. Second, the x and y uncertainty bars are marginal percentiles and do not show the within-replicate covariance between plotted measures. Third, spawning biomass is shown as a standardized dynamic-reference-point ratio rather than in tonnes. Fourth, these views cover only selected yield, biological, and catch-stability periods; they do not include fishing pressure, Kobe-green probability, low-catch risk, feasibility, or the other performance measures needed for a decision. The plotted summary values are saved in `doc/data/sc14-preliminary-tradeoffs.csv` and can be regenerated with `R/plot_sc14_tradeoffs.R`.

Banking, borrowing, and implementation constraints

SCW17 treated banking and borrowing as a sensitivity to apply to the final CMP set rather than as feedback within the HCR. Ten-percent banking-and-borrowing runs were still in progress at the 24 July meeting. The technical annex should state whether the provision materially changes catch, stability, or biological performance and whether its effect differs among CMPs.

Effort feasibility is another implementation constraint. A TAC can be mathematically available but difficult to catch when fish distribution or catchability changes. The OM includes an effort cap to prevent unrealistic projected effort. Low realized catch, catch-rate proxies, and the frequency of binding effort constraints should be retained as supporting diagnostics.

Exceptional circumstances and annual monitoring

An exceptional-circumstances protocol should identify when new information is inconsistent with the conditions under which a selected CMP performed acceptably. SCW17 considered the ECP software framework as a proof of concept: robustness simulations can identify data patterns associated with poor CMP performance, and detection rules can then be evaluated for false-positive and false-negative errors.

The current analysis identifies at least two monitoring priorities:

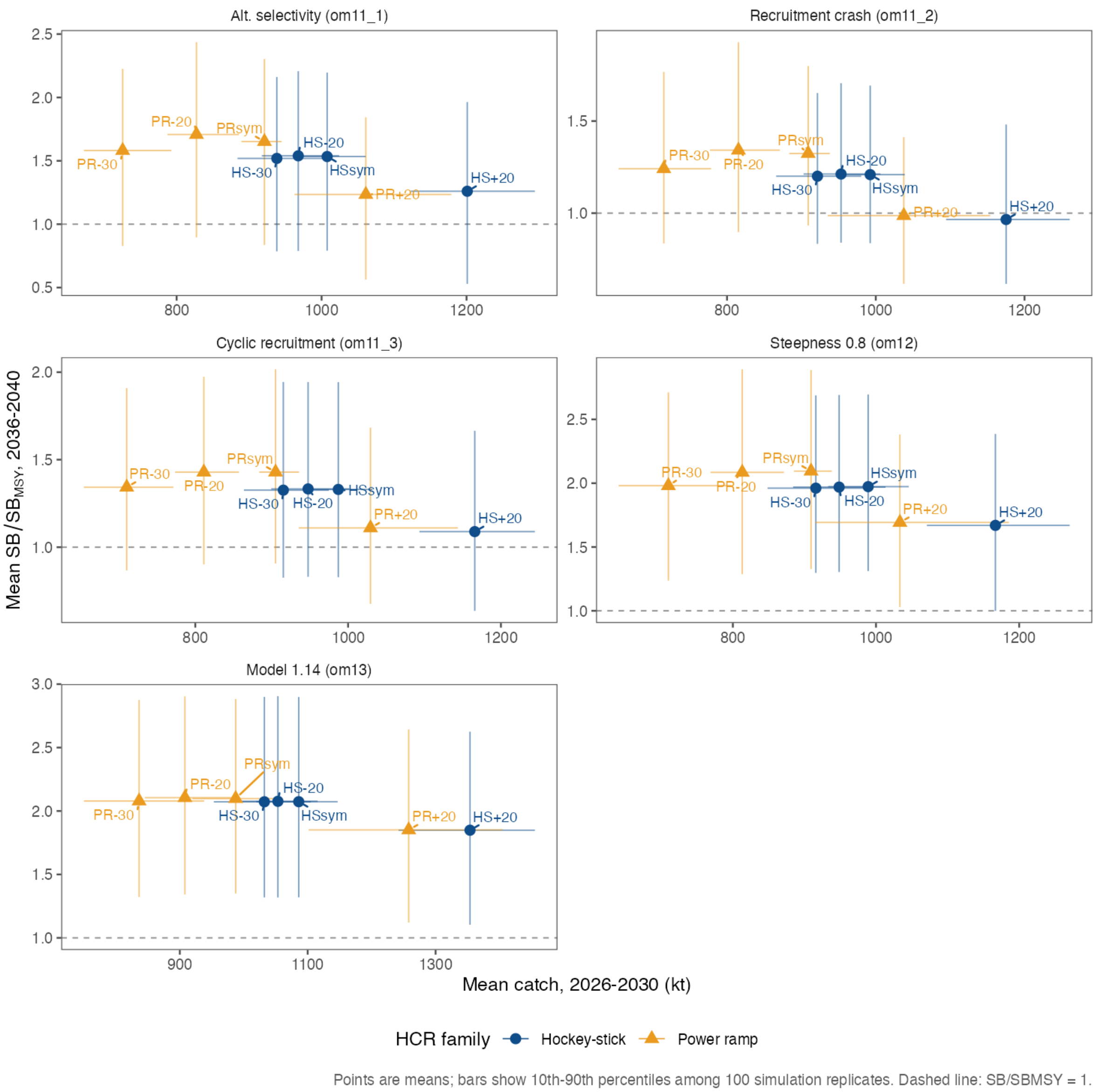

Figure 11: Preliminary trade-offs across the five single-stock robustness OMs.

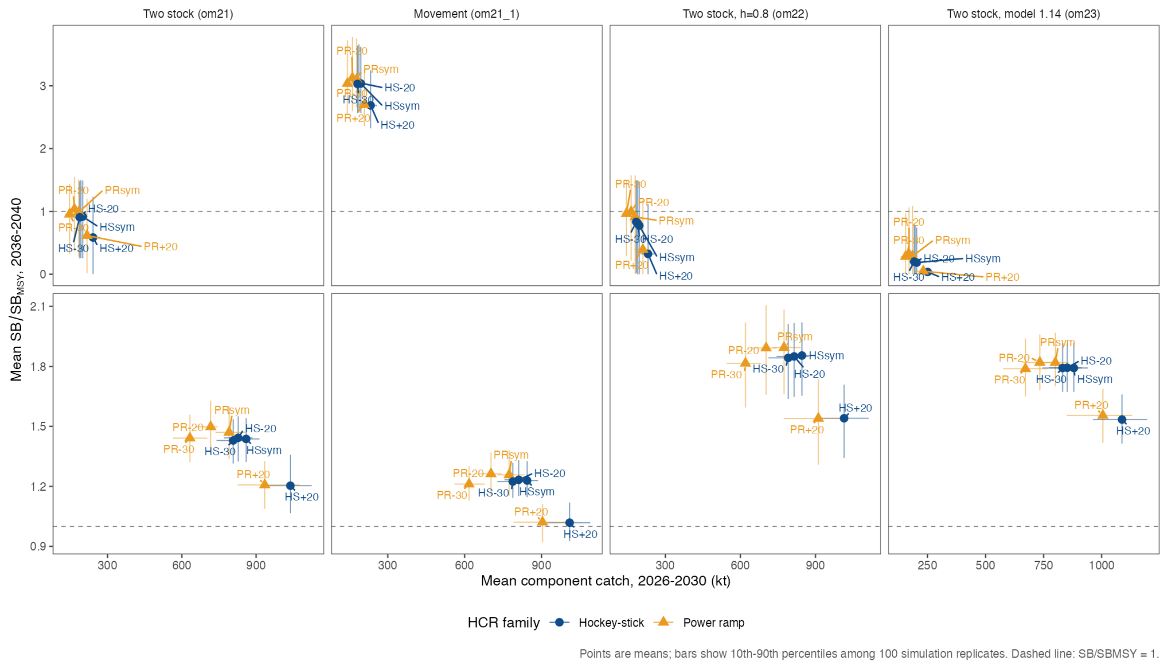

Figure 12: Preliminary component-level trade-offs across the four two-stock robustness OMs.

1. persistent divergence between northern and southern abundance indicators, particularly in relation to the two-stock-with-movement hypothesis; and
2. whether current survey, catch, and fishery information lies outside OM distributions or is associated with simulated CMP failure states.

Preliminary oral fishery updates reported at the 24 July meeting are not validated data submissions and do not by themselves establish exceptional circumstances. A protocol should specify the observations, data timing, triggering evidence, review authority, possible responses, and return-to-rule conditions before operational use.

Summary and conclusions

This MSE combines the operating-model development and management-procedure design undertaken before and during SCW17 with the subsequent implementation, tuning, robustness testing, and performance evaluation. The candidate procedures apply transparent hockey-stick and power-ramp HCRs to a combined indicator constructed from five standardized abundance series. A distinctive feature is that the OMs simulate uncertainty, temporal autocorrelation, and cross-correlation among those series, rather than treating multiple indicators as independent confirmations of stock change.

The analysis narrows the broader workshop set to two central rule families and reasonable annual TAC-change variants. Candidate performance is evaluated against a reference OM and a structured robustness set spanning productivity, recruitment stress, selectivity, stock structure, movement, and spatial catch distribution. The results demonstrate trade-offs among biological condition, catch, stability, and responsiveness rather than identifying a single CMP that dominates every objective. The performance quilt and worked scorecards provide transparent summaries, but the underlying distributions, metric definitions, weights, and robustness results remain necessary for interpretation. For example, MP43 and MP45 (HS-20 and HSsym) perform well under the reference OM and remain comparatively strong across the robustness set, although their relative performance still depends on the metric, weighting, and robustness scenario considered.

The available work supports taking a reduced subset of CMPs forward for Commission selection, accompanied by annual monitoring and a formal exceptional-circumstances process. Because this synthesis contains result-dependent statements, scores, and conclusions, **this section must be updated dynamically in future report versions whenever simulation results, candidate definitions, performance metrics, or agreed weights change.**

The Scientific Committee is asked to consider the following:

Noting

- that the approaches provide valid combinations of the available information for application in the tested HCRs;
- that the work identifies reasonable annual TAC-change variants informed by polling of Commission members;
- that the results satisfy the Commission requests and recorded management preferences;
- the use of all five indices and the complementary value of the fishery-independent acoustic series;
- that northern and southern indicators should continue to be tracked separately, as in the past, alongside the combined CMP indicator and monitored through the annual assessment process, including when evaluating whether exceptional circumstances may be occurring;
- that the available banking-and-borrowing results and implementation appear feasible within the levels investigated; and
- that the robustness tests provide plausible scenarios for testing the CMPs.

Endorsing

- the analytical work as satisfying the request from the Commission;
- adoption of a subset of the CMPs presented so that the Commission may select one among them to satisfy the trade-offs it may consider;
- development of a formal exceptional-circumstances protocol, including evaluation of detection performance and pre-agreed responses; and

- review and documentation of the full performance-metric set and any weighting process, while retaining the underlying outcomes rather than relying on a composite score alone.

Work required before final advice

Note

The remaining analytical and documentation tasks are summarized in Table 6.

Table 6: Work required before final scientific advice.

Work item	Why it matters	Current status in source material
Complete final 500-replicate tuning and reduced-set evaluation	Stabilizes final reported triggers and performance distributions	Requested; 100-versus-500 checks showed small differences
Finalize TAC-change comparison	Tests responsiveness and stability under management-relevant constraints	Retuned 100-replicate reference and robustness variants available
Complete banking-and-borrowing sensitivity	Tests an implementation provision outside the HCR	Pending at 24 July meeting
Add index-error parameter distributions	Makes observation uncertainty auditable	Requested for technical annex
Improve index and HCR diagnostics	Clarifies responsiveness and the acoustic survey's role	Requested for technical annex
Finalize metric definitions and periods	Prevents ambiguous or directionally incorrect scoring	Full framework still being assembled
Expand and sensitivity-test the scorecard	Allows changing CMP and metric sets without implying a fixed ranking	Draft framework available
Specify exceptional-circumstances protocol	Defines detection, authority, response, and return to the CMP	Proof-of-concept stage
Freeze reproducible inputs and outputs	Supports independent review and exact reruns	Run records exist; final archive still required

Source, terminology, and interpretation notes

This synthesis uses "CMP" for a procedure still under evaluation and "MP" when referring to the identifiers used in the analysis. "Catch distribution" is used for two-stock sensitivity tests; those tests are not allocation advice. "Dynamic Kobe green" refers to performance calculated using the dynamic reference-point implementation described above.

Numerical values from presentations and meeting materials should be checked against the frozen final simulation archive before submission. Where a presentation takeaway and the later meeting interpretation differed in tone, this report follows the meeting's more cautious interpretation. The meeting report itself was prepared from meeting notes, transcript material, and an automated meeting summary; preliminary oral information requires confirmation against formal submissions.

Sources

SPRFMO Jack Mackerel Working Group. 2026. Report of the Jack Mackerel Management Strategy Evaluation Workshop - SCW17, 15–19 June 2026, updated 7 July 2026.

SPRFMO Jack Mackerel MSE Task Team. 2026. *Meeting report: Candidate management procedures, robustness testing, and SC14 preparation*, meeting held 24 July 2026. Draft for member review.

Mosqueira, I., and B. Bergès. 2026. *Management Strategy Evaluation (MSE) for SPRFMO jack mackerel: Progress since SCW17.* Presentation dated 25 July 2026.

SPRFMO. 2026. Jack mackerel MSE 2026 project site. Source for candidate definitions, analysis components, performance outputs, supporting materials, and report artifacts used in the current MSE documentation.