NORTHERN COMMITTEE
Twenty-First Regular Session

14 – 15 July 2025
Toyama, Japan (Hybrid)

ISC25/ANNEX 08
Report of the Pacific Bluefin Tuna Management Strategy Evaluation

WCPFC-NC21-2025/WP-04
(IATTC-NC-JWG10-2025/WP-02)

Pacific Bluefin Tuna Working Group
ISC[1]

[1] International Scientific Committee for Tuna and Tuna-like Species in the North Pacific Ocean

ANNEX 08

25th Meeting of the
International Scientific Committee for Tuna
and Tuna-Like Species in the North Pacific Ocean
Busan, Republic of Korea
June 15-20, 2025

REPORT OF THE PACIFIC BLUEFIN TUNA MANAGEMENT STRATEGY EVALUATION[1]

June 2025

[1] Prepared for the 25th Meeting of the International Scientific committee on Tuna and Tuna-like Species in the North Pacific Ocean (ISC) held June 15-20, 2025, in Busan, Republic of Korea. Document should not be cited without permission of the authors.

Left Blank for Printin

Table of Contents

1. Executive Summary

History and Goal of PBF Management Strategy Evaluation

Pacific bluefin tuna (PBF) is a highly migratory species whose range covers the entire North Pacific and which sustains economically important fisheries in Chinese Taipei, Japan, Korea, Mexico and the United States. Due to its broad range, the stock is managed internationally by two Regional Fisheries Management Organizations (RFMOs), the Western and Central Pacific Fisheries Commission (WCPFC) and the Inter American Tropical Tuna Commission (IATTC). The WCPFC-NC and IATTC PBF Joint Working Group (JWG) was started to coordinate PBF management between these two RFMOs. Fishing records date back to the 1800s, and the stock has experienced high fishing pressure, with spawning stock biomass (SSB) falling to 2% of the unfished SSB ($2\%SSB_{F=0}$) in 2009 and 2010. Following the decline of the stock, management measures were put in place by the RFMOs to rebuild the stock to a first rebuilding target of $6.3\%SSB_{F=0}$, and then a second rebuilding target of $20\%SSB_{F=0}$. These management measures were successful, with SSB surpassing the second rebuilding target in 2021.

Now that the stock has rebuilt to the second rebuilding target of $20\%SSB_{F=0}$, the RFMOs have tasked the ISC PBF working group (WG) with developing a management strategy evaluation (MSE) to inform the development of a long-term management procedure (MP) for PBF. MSE is a process that evaluates the tradeoffs and performance of candidate MPs under a range of uncertainties using computer simulations. Testing MPs in an MSE allows for the ruling out of those MPs that do not perform adequately in computer simulations as we would not expect them to perform well in the real world. It also enables managers to identify specific management objectives and quantitative metrics with which to evaluate performance and lays bare the tradeoffs between them.

The RFMOs finalized the candidate harvest control rules (HCRs) to be tested and agreed on the management objectives and performance metrics with which to evaluate their performance in 2023 and requested the ISC PBF WG to finalize the MSE in 2025. In February 2025, after being presented with a set of preliminary results by the ISC, the RFMOs further reduced the HCRs to be tested in the MSE to a final set. This PBF MSE examined the performance of 16 candidate management procedures, relative to the set of management objectives and performance metrics agreed-upon by the RFMOs given uncertainties, using a closed loop computer simulation. The closed loop simulation recreates the real-world management process, from data collection, assessment of stock status, and management procedure implementation (Fig. ES1).

An MP establishes management actions (here, the setting of a total allowable catch, TAC) with the aim of achieving the stated management objectives. It specifies (1) what harvest control rule (HCR) will be applied, (2) how stock status estimates will be calculated (here, via a stock assessment), and (3) how data will be monitored. The MPs in this MSE only differ in terms of the HCRs and associated control points used. As in the real world, estimates of the condition of the PBF stock relative to control points are calculated via a simulated stock assessment, referred to as the estimation model (EM). For this MSE, the EM is an age-structured production model with estimated recruitment deviates (ASPM-R+). The + indicates that size frequency data from the Taiwanese and Japanese longline fleets were included and their selectivities were estimated. It is a simplified version of the 2024 PBF stock assessment model. The virtual stock is monitored by collecting data on catch and size composition as would occur in the real world. Data on catch, size

composition, and the index of abundance are generated, with observation errors, from operating models (OMs), which are mathematical representations of the possible true dynamics of the stock and fisheries (Fig. ES1). These observations are then fed into the simulated stock assessment (i.e., the EM). As in the real world, the results from the simulated assessment are then used to inform the management of the PBF fisheries, based on the candidate HCR being tested (Fig. ES1). The resulting management action (i.e., TAC) then impacts the simulated fleets and the PBF stock (Fig. ES1). At the end of the 23-year long simulation, output from the OMs is used to compute performance metrics to assess the performance relative to the set of management objectives of each of the candidate HCRs.

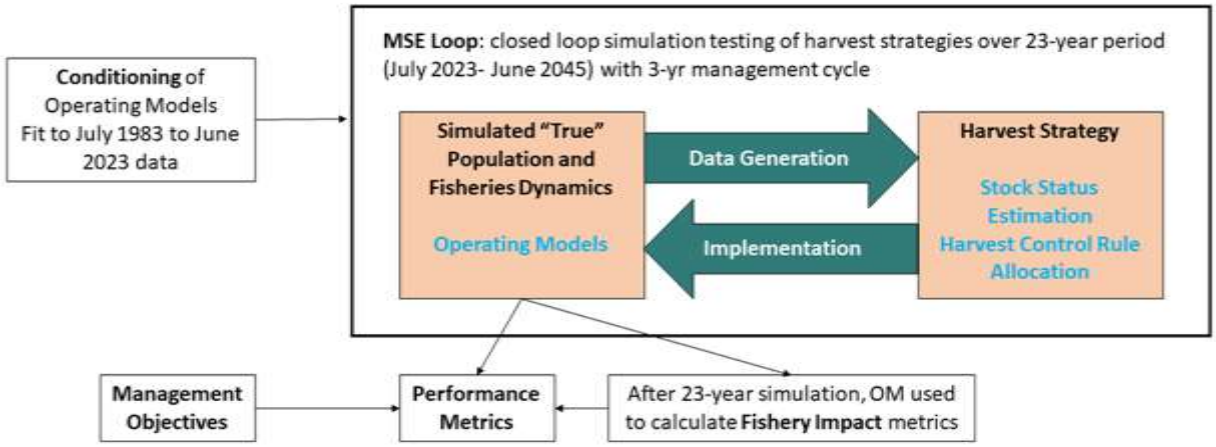

Figure ES1. Overview of the PBF MSE closed-loop simulation framework showing the MSE feedback loop where data are sampled with error from the operating models and fed into the management procedure, which includes a simulated assessment, which determines stock status and informs the harvest control rule (HCR). The HCR then determines a management action (i.e., TAC) which then affects the dynamics of the "true" population in the operating models.

Management Objectives and Performance Indicators

The management objectives and associated performance indicators for this MSE were agreed upon by the RFMOs following two PBF MSE workshops and additional discussions at two JWG meetings. These are outlined in Table ES1. Performance indicators were used to quantitatively evaluate the performance of the HCRs tested relative to the management objectives.

Harvest Control Rules

The HCRs and reference points considered in this MSE (Table ES2) were put forward by the JWG. The HCRs specify a management action based on SSB estimates in relation to biomass-based control points. More specifically, the HCRs identify, given stock status, a desired fishing mortality (F) on the stock, calculated as 1-SPR, where SPR is the spawning potential ratio, the ratio of the cumulative spawning biomass that an average recruit is expected to produce over its lifetime when the stock is fished at the current fishing level to the cumulative spawning biomass that could be produced by an average recruit over its lifetime if the stock was unfished (Fig. ES2).

Within the MSE simulation, a TAC is then set using the desired F and the current biomass from the EM. The TAC is then kept constant for three years until the next assessment. In addition, the first expected TAC to be applied in 2026 is calculated based on the EM but outside the MSE simulation loop. To do so, the EM was updated with catches and an updated index of abundance

for fishing year 2023 (i.e., up to June 2024), the latest year for which data are available. The potential TACs are listed in Table ES4.

Table ES1. List of management objectives and performance indicators put forward by the JWG and used in the PBF management strategy evaluation. SSB refers to spawning stock biomass, LRP to limit reference point, and F to fishing mortality, measured as 1-SPR where SPR is the spawning potential ratio, the ratio of the cumulative spawning biomass that an average recruit is expected to produce over its lifetime when the stock is fished at the current fishing level to the cumulative spawning biomass that could be produced by an average recruit over its lifetime if the stock was unfished. F_{TARGET} is the target reference point based on fishing mortality.

Category	Management Objective	Performance Indicator
Safety	• There should be a less than 20%* probability of the stock falling below the LRP	• Probability that SSB< LRP in any given year of the evaluation period
Status	• To maintain fishing mortality at or below F_{TARGET} with at least 50% probability	• Probability that $F \leq F_{TARGET}$ in any given year of the evaluation period • Probability that SSB is below the equivalent biomass depletion levels associated with the candidates for F_{TARGET}
Stability	• To limit changes in overall catch limits between management periods to no more than 25%, unless the ISC has assessed that the stock is below the LRP	• Percent change upwards in catches between management periods excluding periods when SSB<LRP • Percent change downwards in catches between management periods excluding periods when SSB<LRP
Yield	• Maintain an equitable balance in proportional fishery impact between the WCPO and EPO	• Median fishery impact (in %) on SSB in the terminal year of the evaluation period by fishery and by WCPO fisheries and EPO fisheries
	• To maximize yield over the medium (5-10 years) and long (10-30 years) terms, as well as average annual yield from the fishery	• Expected annual yield over 5-10 years of the evaluation period, by fishery • Expected annual yield over 10-30 years of the evaluation period, by fishery • Expected annual yield in any given year of the evaluation period, by fishery
	• To increase average annual catch in all fisheries across WCPO and EPO	

*The acceptable levels of risk may vary depending on the LRP selected, but should be no greater than 20%.

Table ES2. List of harvest control rules (HCRs) tested in the PBF MSE. The target reference point (F_{TARGET}) is an indicator of fishing mortality based on SPR. SPR is the spawning potential ratio. An F_{TARGET} of $F_{SPR40\%}$ is associated with a fishing mortality that would leave 40% of the SSB per recruit compared to the unfished state. An F_{TARGET} of $F_{SPR20\%}$ implies a higher fishing mortality (i.e., 1-SPR of 0.8) and would result in a SSB per recruit of 20% of the unfished SPR. The threshold (ThRP) and limit reference points (LRP) are SSB-based and refer to the specified percentage of equilibrium unfished SSB ($SSB_{F=0}$). The minimum F (F_{min}) refers to the fraction of the F_{TARGET} that the fishing intensity is set to when SSB is below the LRP, except for HCRs 4 and 12, which specify a specific fishing mortality. Note that for HCRs 5 and 13, when the ThRP is breached, the HCR switches from constant fishing mortality at the F_{TARGET} to a constant TAC set at the catch limits defined in CMM2021-02 (WCPFC 2021) and C-21-05 (IATTC 2021). While HCRs 5, 6, 7, 13, 14, and 15 do not use LRPs as control points, an LRP of median SSB from 1952-2014 (6.3% $SSB_{F=0}$) has been specified by the JWG to compute performance metrics. HCRs 9 to 16 are identical to HCRs 1 to 8, except for the allocation of fishing pressure between the Western Central Pacific Ocean (WCPO) fleet segment and the Eastern Pacific Ocean (EPO) fleet segment. HCRs 1 to 8 were tuned to reach a fishery impact ratio between the WCPO and EPO of 80% to 20% (80:20), while HCRs 9 to 16 were tuned to reach a WCPO:EPO fishery impact ratio of 70:30.

HCR number	F_{TARGET}	Control Point 1 (ThRP)	Control Point 2 (LRP)	Number of Control Points	F_{min}	WCPO:EPO Impact Ratio
1	FSPR30%	20%$SSB_{F=0}$	15%$SSB_{F=0}$	2	10% F_{TARGET}	80:20
2	FSPR30%	25%$SSB_{F=0}$	15%$SSB_{F=0}$	2	10% F_{TARGET}	80:20
3	FSPR40%	25%$SSB_{F=0}$	20%$SSB_{F=0}$	2	10% F_{TARGET}	80:20
4	FSPR30%	20%$SSB_{F=0}$	10%$SSB_{F=0}$	2	FSPR70%	80:20
5	FSPR25%	20%$SSB_{F=0}$	NA	1	NA	80:20
6	FSPR20%	20%$SSB_{F=0}$	NA	1	NA	80:20
7	FSPR25%	15%$SSB_{F=0}$	NA	1	NA	80:20
8	FSPR30%	20%$SSB_{F=0}$	7.7%$SSB_{F=0}$	2	5% F_{TARGET}	80:20
9	FSPR30%	20%$SSB_{F=0}$	15%$SSB_{F=0}$	2	10% F_{TARGET}	70:30
10	FSPR30%	25%$SSB_{F=0}$	15%$SSB_{F=0}$	2	10% F_{TARGET}	70:30
11	FSPR40%	25%$SSB_{F=0}$	20%$SSB_{F=0}$	2	10% F_{TARGET}	70:30
12	FSPR30%	20%$SSB_{F=0}$	10%$SSB_{F=0}$	2	FSPR70%	70:30
13	FSPR25%	20%$SSB_{F=0}$	NA	1	NA	70:30
14	FSPR20%	20%$SSB_{F=0}$	NA	1	NA	70:30
15	FSPR25%	15%$SSB_{F=0}$	NA	1	NA	70:30
16	FSPR30%	20%$SSB_{F=0}$	7.7%$SSB_{F=0}$	2	5% F_{TARGET}	70:30

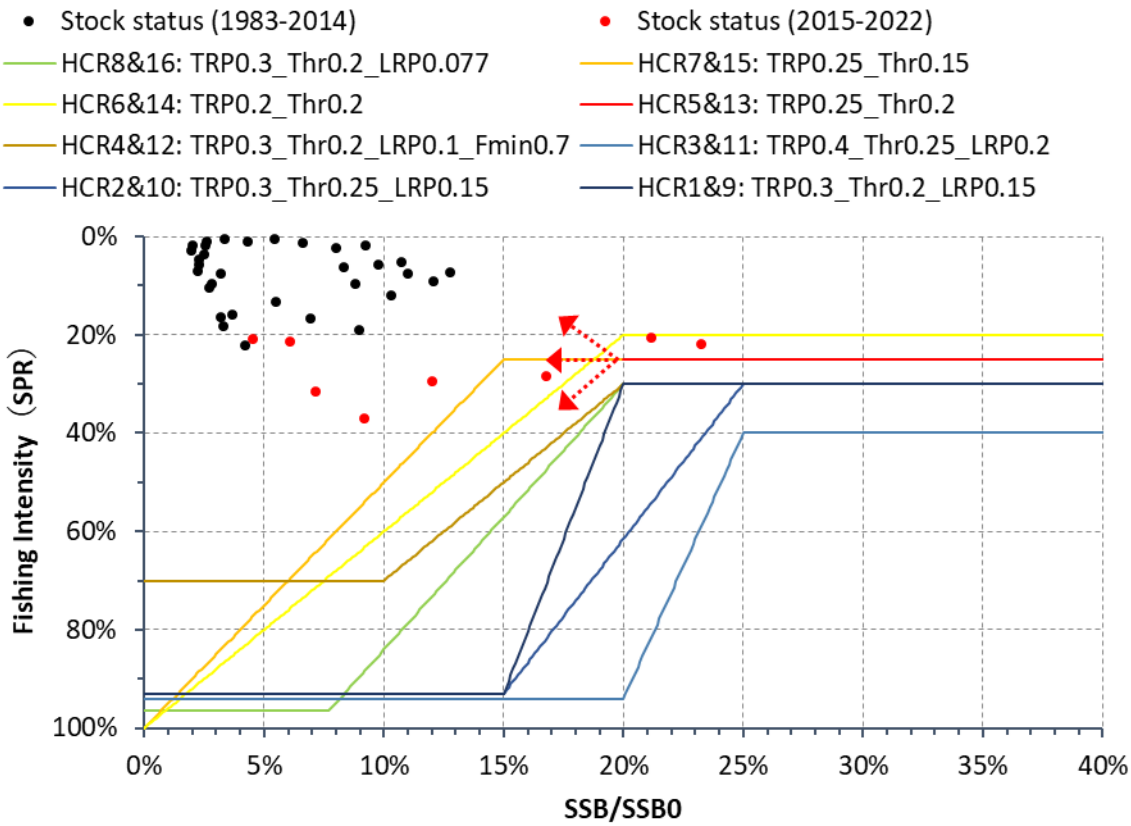

Figure ES2. Candidate HCR evaluated in the PBF MSE. Fishing intensity is an indicator of fishing mortality based on SPR. SPR is the spawning potential ratio that would result from the current year's pattern and intensity of fishing mortality relative to the unfished stock. $SSB/SSB_{F=0}$ is SSB relative to the equilibrium unfished SSB ($SSB_{F=0}$). The points are annual estimates of SPR and relative SSB from the latest PBF stock assessment (ISC 2024). Red dots represent the years when stricter catch limits were in place to rebuild the stock. For HCR 5 (red line), a constant catch management, which was similar to the one applied in 2015-2022, is used if the SSB breaches a control point set at $20\%SSB_{F=0}$. Resulting illustrative fishing intensities for a constant catch are shown as dashed arrows. Note that HCRs 9 to 16 are not represented as they are identical in shape to HCRs 1 to 8.

These HCRs define the management action to be taken (i.e., F) given the estimated ratios of SSB to biomass-based control points from the simulated stock assessments. All the HCRs considered in this MSE have a target state based on fishing mortality (F_{TARGET}). This is the target reference point (TRP) and the state that management wants to achieve. Some HCRs have two control points, with the first being labeled the threshold reference point (ThRP) and the second being labeled the limit reference point (LRP). Having two control points generally helps avoid reaching low biomass levels, where severe management action is taken, and rebuild the stock back to a target state faster. Figure ES2 outlines, for each HCR, the allowed F based on the status of estimated SSB relative to $SSB_{F=0}$. For all HCRs, if SSB is above the first control point, F is managed to be at the F_{TARGET} (Fig. ES2). If SSB falls below the first control point, the allowed F is reduced, except for HCR 5 and 13, in proportion to the estimated relative SSB down to a minimum level at the second control point for HCRs with two control points or down to 0 for those with one control point, to allow biomass to increase back to the target (Fig. ES2). For HCRs 5 and

13, a constant catch management, which was similar to the one applied in 2015-2022, is applied if the SSB breaches its first control point. Historically, the stock has been under intense fishing pressure, and F as estimated by the latest stock assessment has never been at a 40%SPR level, even when the stricter management measures were in place (Fig. ES2).

It is important to note that the LRPs and TRPs in the HCRs serve both as control points of management actions and as measuring sticks to evaluate performance. However, control points can differ from the LRPs and TRPs. LRPs and TRPs, in principle, can also simply play the role of reference points to evaluate the performance of HCRs. In these cases, the level of the LRPs and TRPs would only be used as measuring sticks without affecting the management actions under the HCRs.

Uncertainties considered

MSE recreates the real-world management process to ensure that management procedures will work even in the presence of errors in the observations, assessment, and implementation. The PBF MSE framework therefore adds realistic error to the data used in the simulated stock assessments (i.e., the EMs). As in the real world, the MSE framework also runs the EM every three years and estimates stock status with this data to ensure that estimation error is considered. The MSE also simulates a realistic lag between the availability of data used in the assessment and the implementation of management actions. For instance, the first EM in the MSE uses data up to fishing year 2023 (i.e., up to June 2024) to set a TAC that is applied starting in calendar year 2026. TACs are provided in three categories of fleets; WPO large fish, WPO small fish, and EPO, based on the recent (2015-2022) selectivity. Since the fleets may catch more than assigned by the TAC due to discards, the MSE also includes an implementation error by adding 1.2% higher catch than set by the HCR to EPO recreational fleets, 5% higher to the WCPO fleets except for the Japanese troll for penning fleet, which is set at 100% higher to account for potentially high discards.

In addition to uncertainty related to the management process, the MSE also considers uncertainty stemming from our limited understanding of the true population or fisheries dynamics. This was addressed by developing 20 different OMs, each representing an equally plausible "true" version of the system. In developing the potential OMs, the ISC PBF WG reviewed potential sources of uncertainty for the PBF stock and identified natural mortality, growth, and the steepness parameter as the most influential sources of uncertainty. The PBF WG then diagnosed plausible ranges for these parameters and developed population dynamics models using the resulting parameter combinations. Models that passed a series of quantitative diagnostic tests to ensure they were plausible and could reasonably replicate past PBF observations were selected as a reference set and given equal weight. Models that demonstrated unsatisfactory diagnostics were discarded. The OM reference set spans a wide range of stock statuses (Fig. ES3). All results and performance metrics are calculated across this entire reference set.

In addition to the reference set, the PBF WG also developed three robustness tests. These are less likely than the reference set and so should not be given the same weight, but are still considered plausible. They are a way to test HCR behavior under extreme conditions detrimental to stock productivity. These robustness tests were: 1) a doubling of discards; 2) an effort creep for the Taiwanese longline fleet on which the main index of abundance is based; and 3) about a 40% 10-year long drop in recruitment, starting from 2042. These robustness OMs were constructed by modifying OM1, which has the same settings as the 2024 base-case assessment model. Results for the robustness set are presented separately. Finally, as PBF recruitment can vary greatly between years due to unknown environmental factors, even when SSB remains stable, the MSE also

considered process uncertainty in recruitment. This was done by, for each OM, sampling recruitment deviations from a normal distribution with a mean of 0 and standard deviation σ_R=0.6 in log space.

For each HCR-OM combination, 100 iterations with different random trajectories in recruitment were run. Less than 1% of all the simulated assessments had estimation issues and had an extremely high estimation error (> 1000% absolute relative error) or produced unrealistically low estimated SSB (less than 1 fish) that were not seen in the OMs and were not caused by the HCRs. These unrealistically low estimated SSBs appeared to be caused by unrealistic estimation error due to non-convergence. While this only happened for EMs in some assessment years, iterations, and OMs, to ensure the HCRs were exposed to the same recruitment trends, we discarded the iterations associated with this estimation issue for all OMs and HCRs, leaving a total of 81 iterations per OM/HCR combination with which to compute performance metrics. Removing these iterations was considered reasonable given that it did not greatly affect the performance metrics (see details in main text).

Table ES3. List of the 20 operating models (OMs) in the reference set representing different productivity scenarios and their parameter specifications. The models were considered equally plausible and given equal weight in the calculation of performance metrics. M_{2+} refers to natural mortality for age 2 and older fish, L_2 refers to the length at age 3, and h refers to steepness. OM 1 has the same parameter specifications as the current base case stock assessment for Pacific bluefin tuna.

OM #	M_{2+}	L_2	h	OM #	M_{2+}	L_2	h
1	0.25	118.57	0.999	12	0.25	118.57	0.99
2	0.25	118	0.91	13	0.25	119	0.99
3	0.193	118.57	0.97	14	0.25	118	0.97
4	0.193	118	0.999	15	0.25	119	0.97
5	0.193	118	0.99	16	0.25	118	0.95
6	0.193	118.57	0.99	17	0.25	118.57	0.95
7	0.193	119	0.99	18	0.25	119	0.95
9	0.25	118	0.999	19	0.25	118	0.93
10	0.25	119	0.999	20	0.25	118.57	0.93
11	0.25	118	0.99	21	0.25	119	0.93

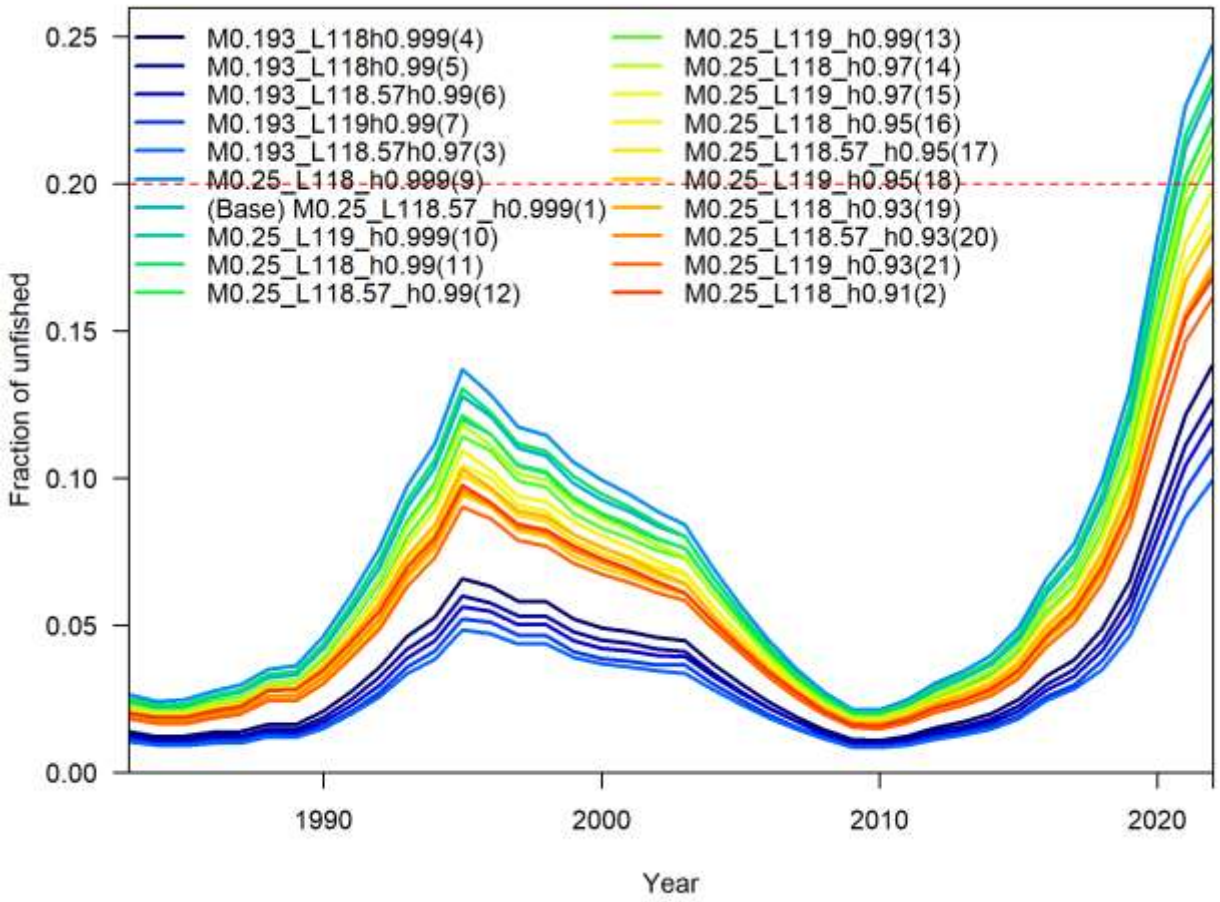

Figure ES3. Historical trajectory of the relative spawning stock biomass estimated from each of the 20 operating models (OMs) in the reference set representing different productivity scenarios and their parameter specifications. The dashed line indicates the rebuilding target at $20\%SSB_{F=0}$. The models were considered equally plausible and given equal weight in the calculation of performance metrics. M refers to natural mortality for age 2 and older fish, L refers to the length at age 3, and h refers to steepness. The OM number in parentheses refers to Table ES3. OM 1 has the same parameter specifications as the current base case stock assessment for Pacific bluefin tuna.

Results

The results of the MSE analysis can be summarized in eight main points:

1. *All HCRs were able to maintain a low probability (<20%) of the stock breaching their respective LRP and the IATTC's interim reference point for tropical tunas of $7.7\%SSB_{F=0}$. In addition, all HCRs except for HCRs 6 and 14 were also able to maintain a low probability (<20%) of breaching the second rebuilding target of $20\%SSB_{F=0}$. Under all HCRs, median SSB increased from initial conditions to levels above their respective targets (Fig. ES4).*

Even when considering the range of uncertainties in stock productivity, recruitment variability, observation, estimation, and implementation, all HCRs met the safety objective and had a less than a 20% probability of SSB being below their respective LRP and a less than 10% probability of breaching the IATTC's interim reference point for tropical tunas (Figs. ES5 and ES6, Table ES4). Furthermore, all HCRs except 6 and 14, had a less than 20% probability of SSB being below the second rebuilding target of 20%$SSB_{F=0}$ (Fig. ES7, Table ES4). Also, under all HCRs, median SSB increased from initial conditions to levels above their respective targets (Fig. ES4).

The PBF WG has no specific recommendation for an LRP with which to test safety performance, especially given that the PBF stock has recovered from a very low level of SSB (2% of $SSB_{F=0}$).

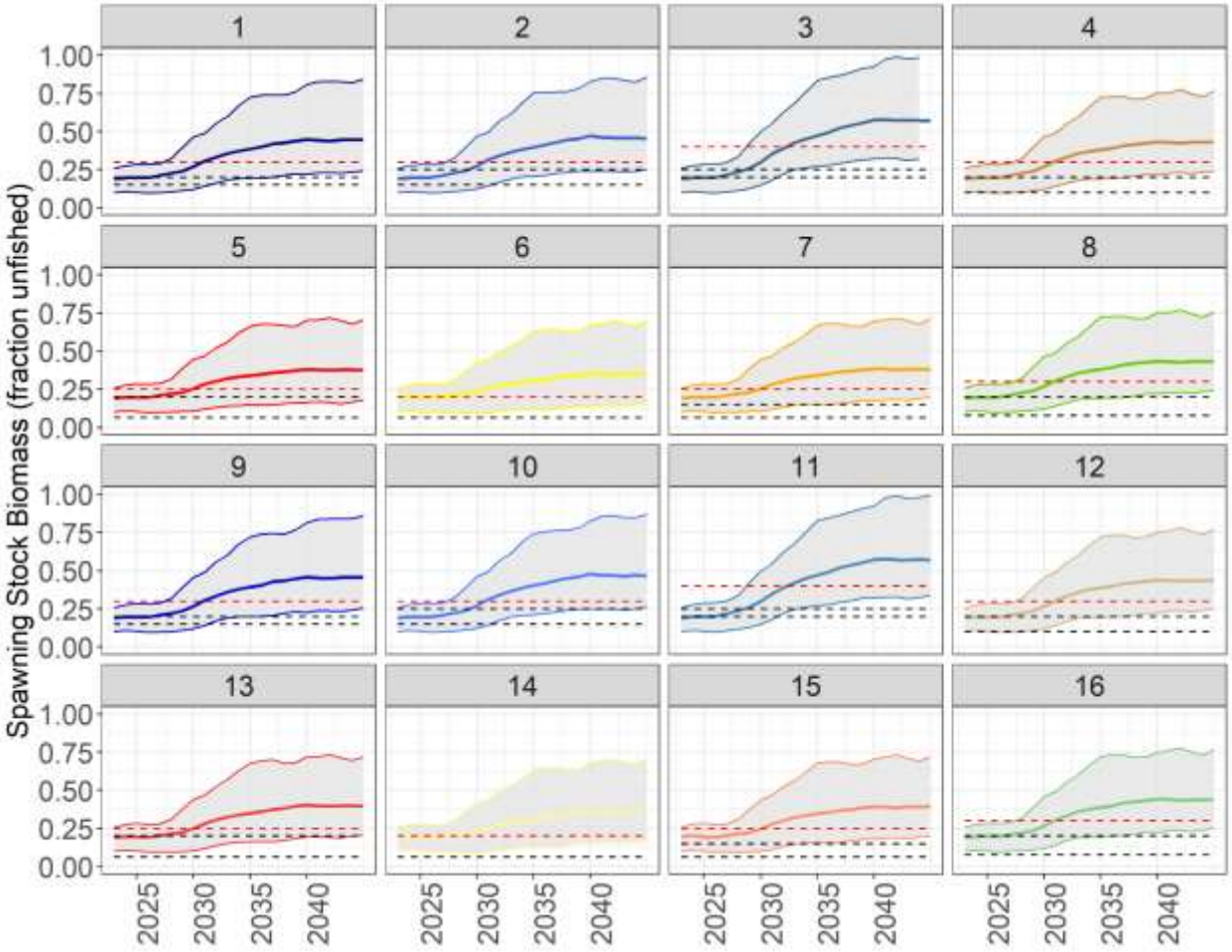

Figure ES4. Trends in median relative spawning stock biomass (SSB/unfished SSB, thick solid color lines) from the operating models under all iterations and reference scenarios by harvest control rule (HCR). The grey shading represents trends in the 5th to 95th quantile range. The lowest black dotted line represents the lowest control point for each HCR, and the highest black dotted line represents the highest. The dashed red line represents the SSB associated with the respective F_{TARGET}. Note that HCRs 5, 6, 7, 13, 14, and 15 do not have a second control point, so the lowest dashed line marks the LRP specified by the JWG to assess performance.

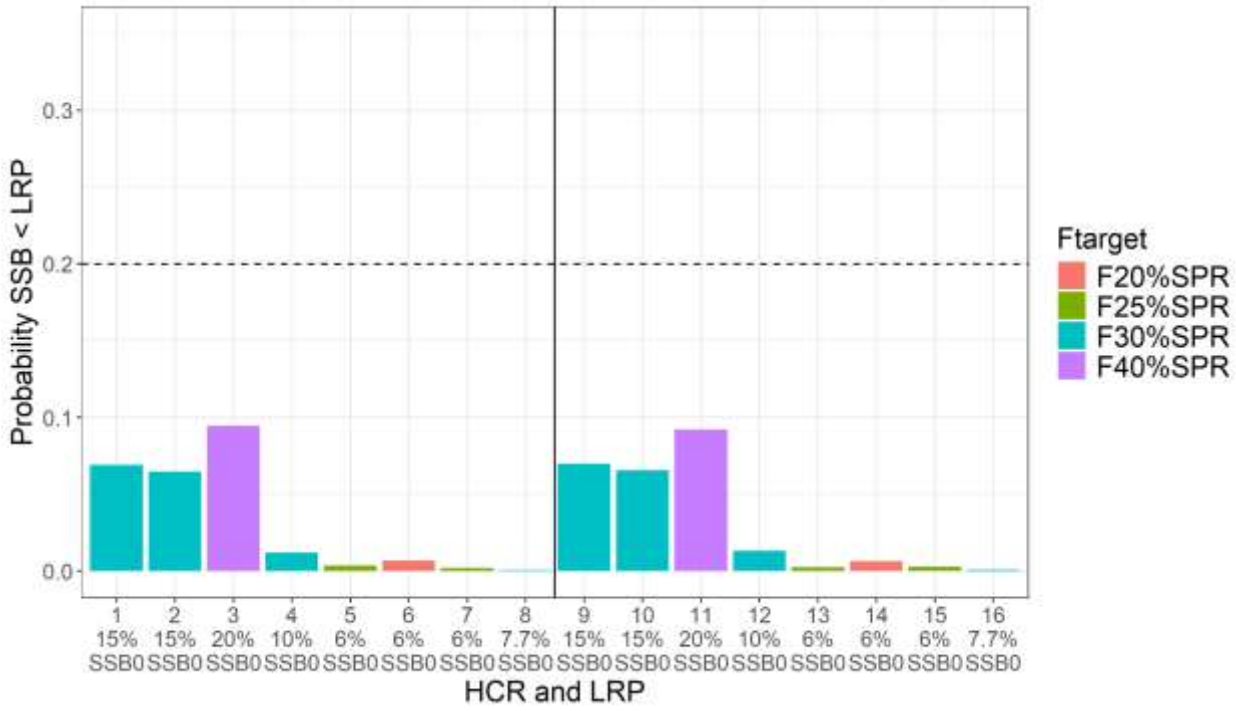

Figure ES5. Probability, for each harvest control rule (HCR), of spawning stock biomass (SSB) being below the limit reference point (LRP) specified by each HCR across all reference scenarios, iterations, and simulation years. Colors represent the F_{TARGET} reference point associated with each HCR. The x-axis shows both the HCR number and the LRP relative biomass level associated with each HCR. The vertical solid line separates HCRs 9 to 16, which are tuned to an EPO:WCPO impact ratio of 30:70 but are otherwise the same as HCRs 1 to 8. EPO stands for Eastern Pacific Ocean and WCPO for Western Central Pacific Ocean.

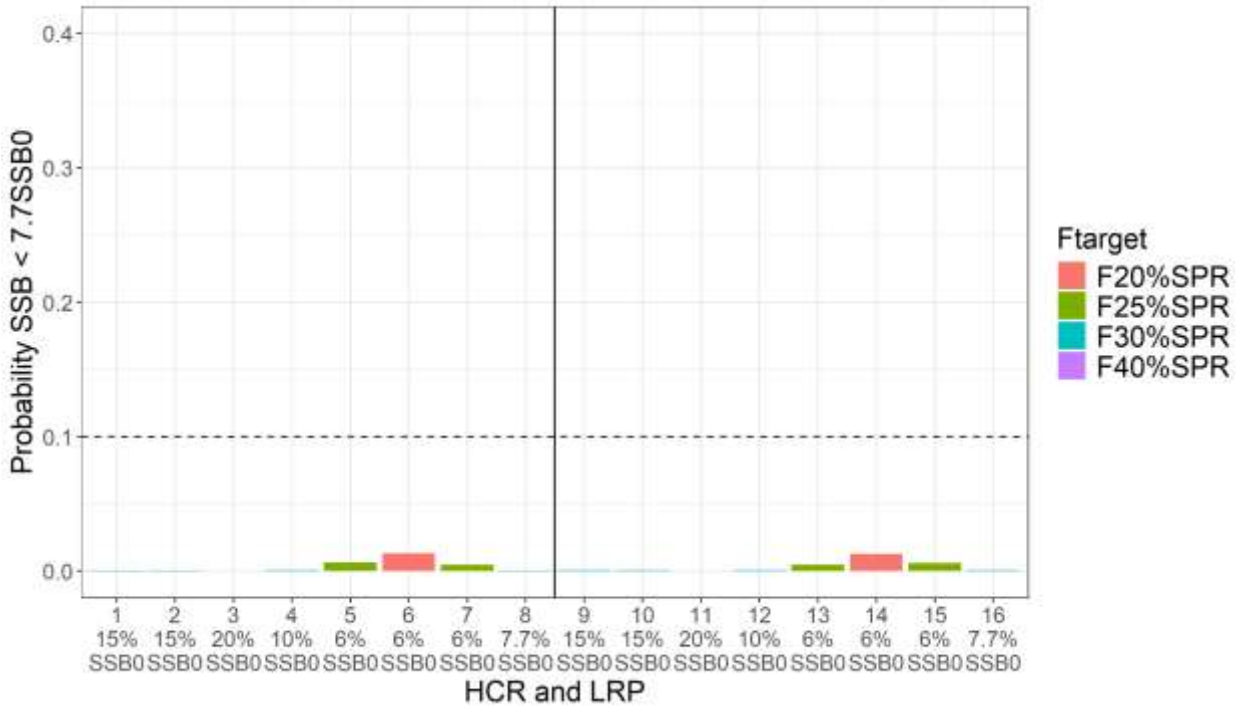

Figure ES6. Probability, for each harvest control rule (HCR), of spawning stock biomass (SSB) being less than 7.7%$SSB_{F=0}$ across all reference scenarios, iterations, and simulation years. The x-axis shows both the HCR number and the LRP relative biomass level associated with each HCR. Colors represent the F_{TARGET} reference point associated with each HCR. The horizontal dotted line represents a 10% probability. The vertical solid line separates HCRs 9 to 16, which are tuned to an EPO:WCPO impact ratio of 30:70 but are otherwise the same as HCRs 1 to 8. EPO stands for Eastern Pacific Ocean and WCPO for Western Central Pacific Ocean.

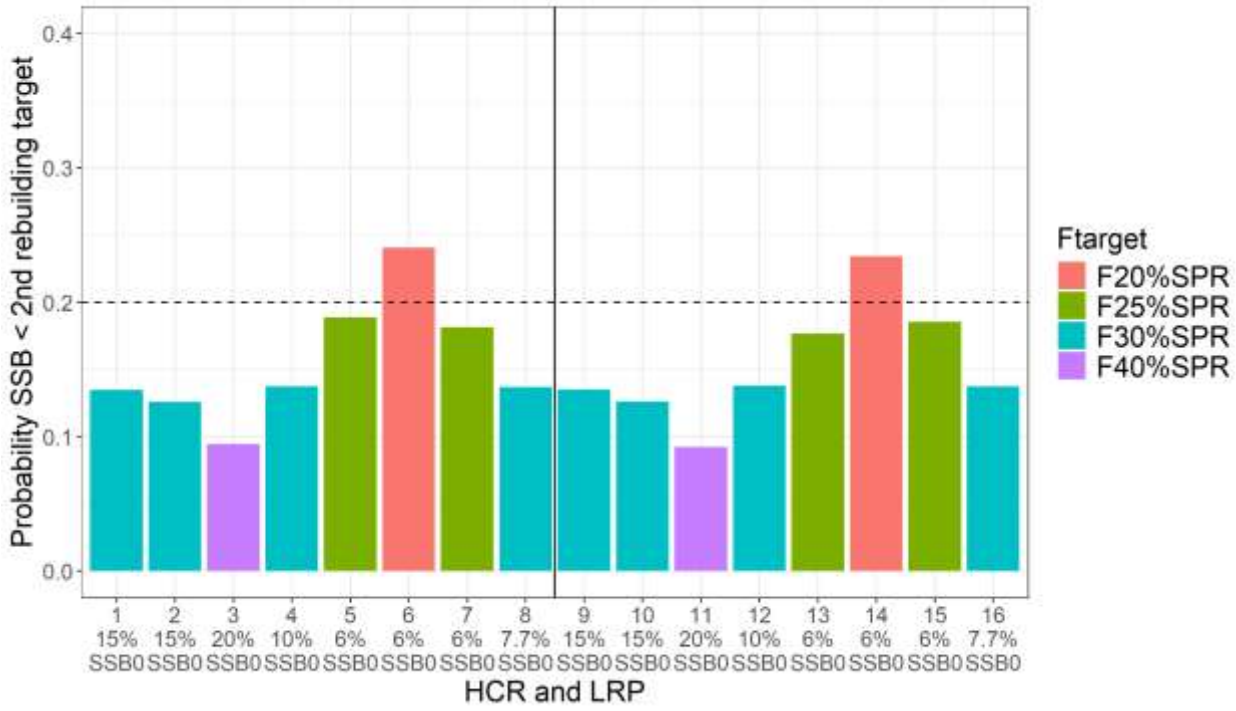

Figure ES7. Probability, for each harvest control rule (HCR), of spawning stock biomass (SSB) being less than 20%$SSB_{F=0}$ across all reference scenarios, iterations, and simulation years. The x-axis shows both the HCR number and the LRP relative biomass level associated with each HCR. Colors represent the F_{TARGET} reference point associated with each HCR. The horizontal dotted line represents a 20% probability. The vertical solid line separates HCRs 9 to 16, which are tuned to an EPO:WCPO impact ratio of 30:70 but are otherwise the same as HCRs 1 to 8. EPO stands for Eastern Pacific Ocean and WCPO for Western Central Pacific Ocean.

2. *There was a tradeoff between the safety metrics (e.g., probability of being at or above the second rebuilding target of 20%$SSB_{F=0}$) and yield metrics (e.g., median annual catch in mt). Those HCRs that had the highest probability of SSB being at or above the second rebuilding target had the lowest yield metrics and vice-versa.*

Due to their higher F_{TARGET}, HCRs 3 and 11 maintained a higher SSB and had the highest probability of SSB being at or above the second rebuilding target of 20%$SSB_{F=0}$, but this came at the cost of lower yields (Fig. ES8), with these HCRs having the lowest total catch, as well as the lowest fleet segment specific (i.e., WCPO large, WCPO small, and EPO) TACs (Figs. ES9, ES10, ES11, and ES12, Table ES4). HCRs with the same F_{TARGET} perform similarly for safety and yield metrics.

Given tradeoffs between the different performance indicators, the choice of a preferred HCR is dependent on the priorities of the respective managers and stakeholders regarding the different management objectives and their level of risk aversion.

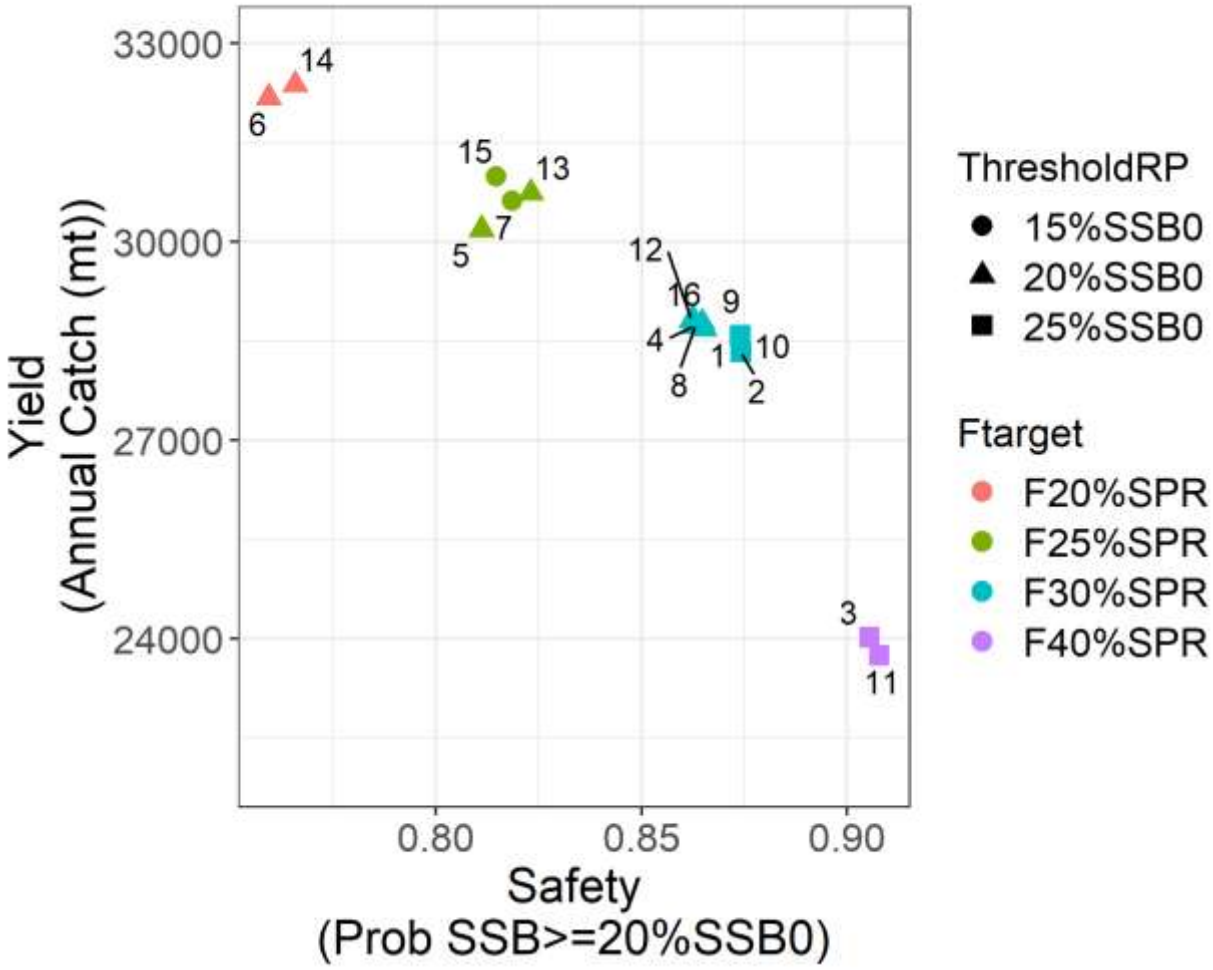

Figure ES8. Median annual total catch versus the probability of spawning stock biomass (SSB) being at or above the second rebuilding target of $20\%SSB_{F=0}$. Note that to ensure that for both measures a higher value is better, here we reversed the second performance metric shown in Fig. ES5 to be the probability of $SSB \geq 20\%SSB_{F=0}$ instead of the probability of $SSB < 20\%SSB_{F=0}$. Each HCR is labeled and colored according to their F_{TARGET}. Each symbol represents a different ThresholdRP, which is the first control point for each HCR and stands for Threshold Reference Point.

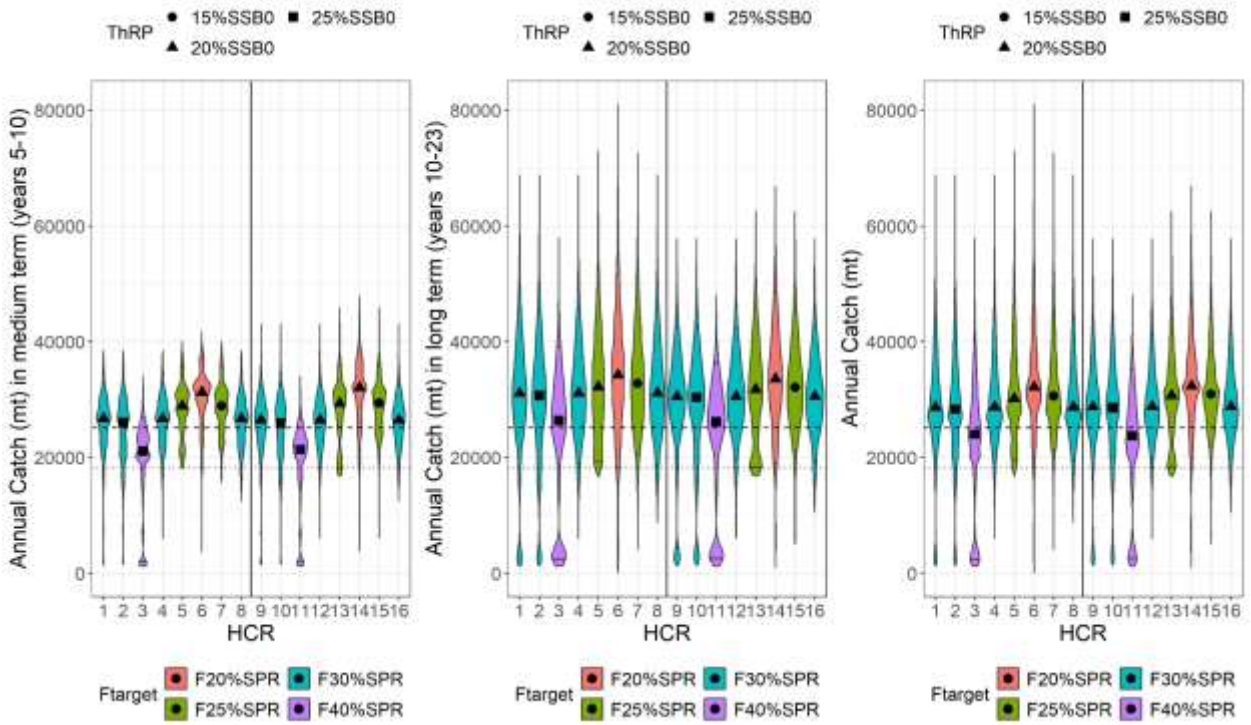

Figure ES9. Violin plots showing the probability density of total annual catch (including discards and the EPO recreational fleet) for each harvest control rule (HCR) across all iterations, reference scenarios, and simulation years in the medium term (first panel), long term (second panel), and all years (third panel). The medium term shows the annual catch distribution over years 5 to 10 of the simulation, while the long term shows the distribution over years 10 to 23 of the simulation. Colors represent the F_{TARGET} reference point associated with each HCR. The marker inside each violin plot is the median value for the medium term, long term, or annual catch, and horizontal solid lines within each violin represent the 5th to 95th quantile range. The shape of each marker represents the ThresholdRP (ThRP), which is the first control point for each HCR and stands for Threshold Reference Point. The dotted line identifies the total catch limit set by the WCPFC's CMM 23-02 plus IATTC's Resolution C-21-05, effective in 2024, plus EPO recreational catches for the calendar year 2023. The dashed line identifies the total catch limit set by the WCPFC's CMM 24-01 plus IATTC's Resolution C-24-02, effective in 2025, plus EPO recreational catches for the calendar year 2023. For the IATTC's resolution, catch limits were based on half of the biennial TAC. The vertical solid line separates HCRs 9 to 16, which are tuned to an EPO:WCPO impact ratio of 30:70 but are otherwise the same as HCRs 1 to 8. EPO stands for Eastern Pacific Ocean and WCPO for Western Central Pacific Ocean.

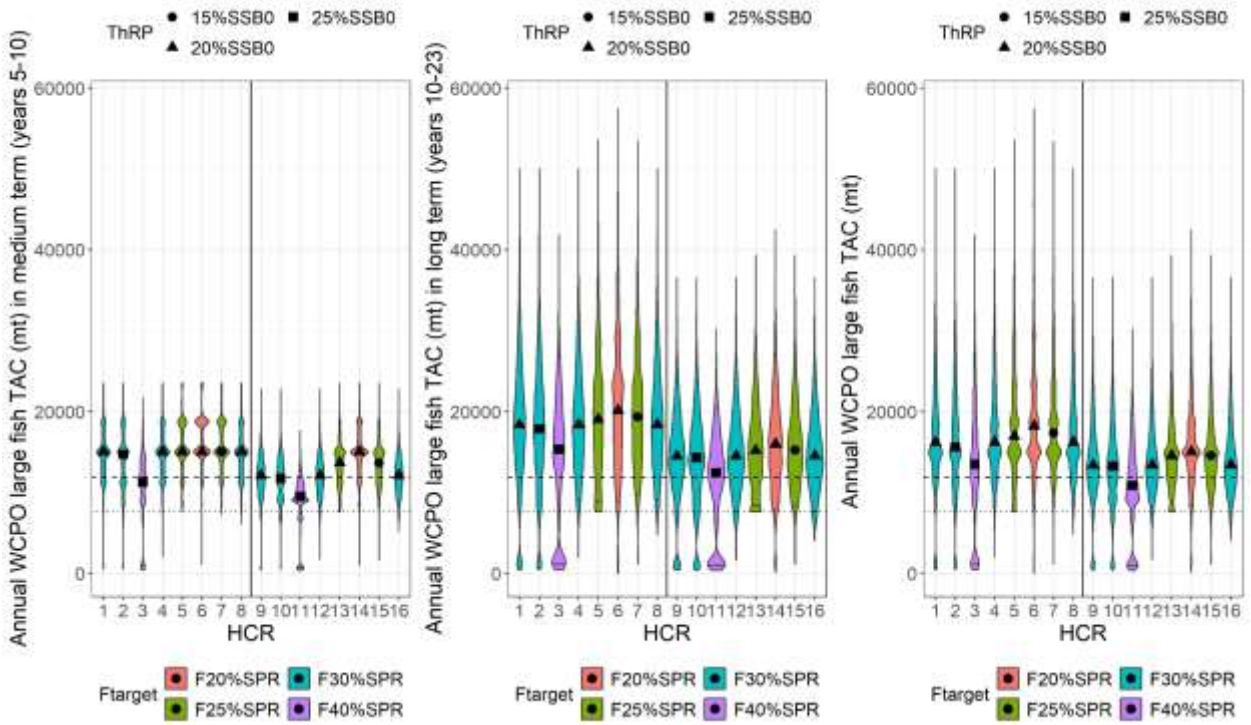

Figure ES10. Violin plots showing the probability density of the TAC for the Western Central Pacific Ocean (WCPO) large fish fleets for each harvest control rule (HCR) across all iterations, reference scenarios, and simulation years in the medium term (first panel), long term (second panel), and annually (third panel). The medium term shows the annual catch distribution over years 5 to 10 of the simulation, while the long term shows the distribution over years 10 to 23 of the simulation. Colors represent the F_{TARGET} reference point associated with each HCR. The marker inside each violin plot is the median value for the medium term, long term, or annual TAC, and horizontal solid lines within each violin represent the 5th to 95th quantile range. The shape of each marker represents the ThresholdRP (ThRP), which is the first control point for each HCR and stands for Threshold Reference Point. The dotted line identifies the catch limit for large fish set by the WCPFC's CMM 23-02, effective in 2024. The dashed line identifies the catch limit for large fish set by the WCPFC's CMM 24-01, effective in 2025. The vertical solid line separates HCRs 9 to 16, which are tuned to an EPO:WCPO impact ratio of 30:70 but are otherwise the same as HCRs 1 to 8. EPO stands for Eastern Pacific Ocean.

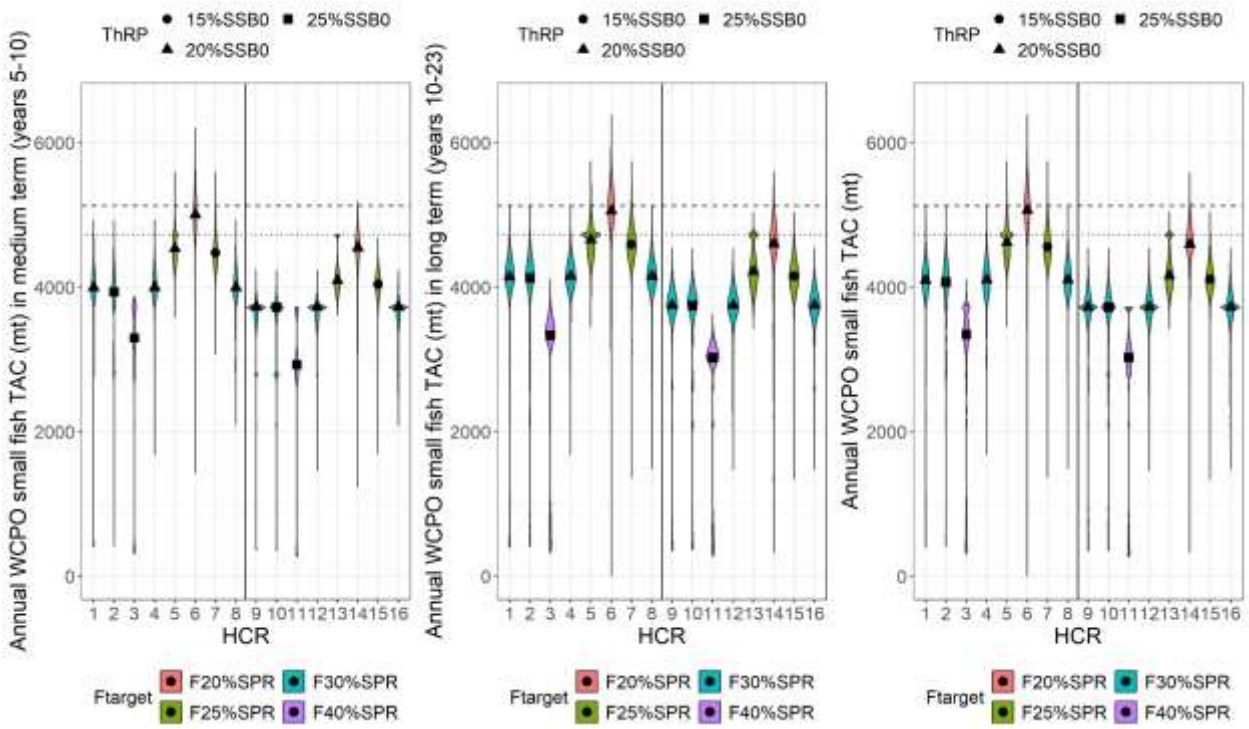

Figure ES11. Violin plots showing the probability density of the TAC for the Western Central Pacific Ocean (WCPO) small fish fleets for each harvest control rule (HCR) across all iterations, reference scenarios, and simulation years in the medium term (first panel), long term (second panel), and annually (third panel). The medium term shows the annual catch distribution over years 5 to 10 of the simulation, while the long term shows the distribution over years 10 to 23 of the simulation. Colors represent the F_{TARGET} reference point associated with each HCR. The marker inside each violin plot is the median value for the medium term, long term, or annual TAC, and horizontal solid lines within each violin represent the 5th to 95th quantile range. The shape of each marker represents the ThresholdRP (ThRP), which is the first control point for each HCR and stands for Threshold Reference Point. The dotted line identifies the catch limit for small fish set by the WCPFC's CMM 23-02, effective in 2024. The dashed line identifies the catch limit for small fish set by the WCPFC's CMM 24-01, effective in 2025. The vertical solid line separates HCRs 9 to 16, which are tuned to an EPO:WCPO impact ratio of 30:70 but are otherwise the same as HCRs 1 to 8. EPO stands for Eastern Pacific Ocean.

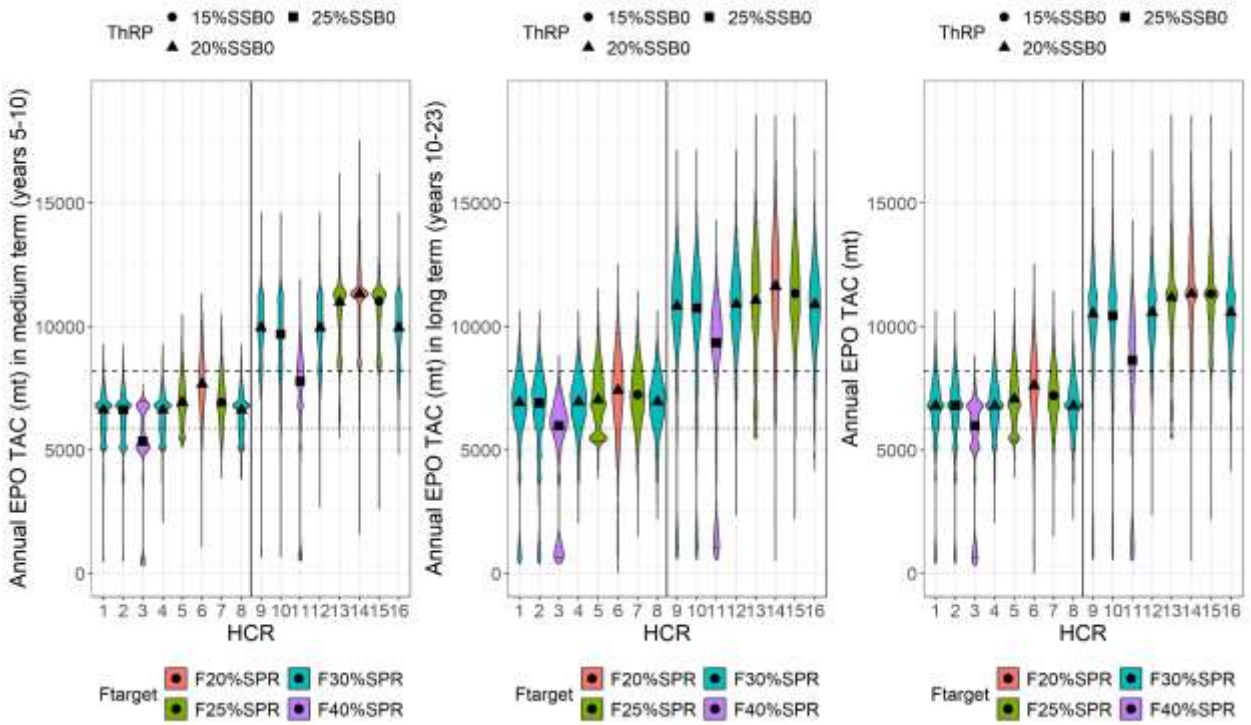

Figure ES12. Violin plots showing the probability density of the TAC for the Eastern Pacific Ocean (EPO) fleets for each harvest control rule (HCR) across all iterations, reference scenarios, and simulation years in the medium term (first panel), long term (second panel), and annually (third panel). The medium term shows the annual catch distribution over years 5 to 10 of the simulation, while the long term shows the distribution over years 10 to 23 of the simulation. Colors represent the F_{TARGET} reference point associated with each HCR. The marker inside each violin plot is the median value for the medium term, long term, or annual TAC and horizontal solid lines within each violin represent the 5th to 95th quantile range. The shape of each marker represents the ThresholdRP (ThRP), which is the first control point for each HCR and stands for Threshold Reference Point. The dotted line identifies the catch limit for the EPO set by IATTC's Resolution C-21-05, effective in 2024, plus EPO recreational catches for the calendar year 2023. The dashed line identifies the catch limit set by IATTC's Resolution C-24-02, effective in 2025, plus EPO recreational catches for the calendar year 2023. Catch limits were based on the half of the biennial TAC. Note that in the MSE, the EPO TAC includes recreational fleets. The vertical solid line separates HCRs 9 to 16, which are tuned to an EPO:WCPO impact ratio of 30:70, but are otherwise the same as HCRs 1 to 8. WCPO stands for Western Central Pacific Ocean.

3. *Catch in the medium and long term for all HCRs is expected to be higher than the current catch limit, except for HCRs 3 and 11 in the medium term. However, the expected TAC trends differ among fleets, with only the WCPO large fish fleet and the EPO fleet under a 70:30 impact ratio increasing above current catch limits.*

Median catches of all HCRs, except for HCRs 3 and 11 in the medium term and across all years, reached higher levels than the current catch limit (Fig. ES9). All HCRs had a long term catch higher than the current catch limit (Fig. ES9). Across all HCRs, the increase in catch was due to increases in the WCPO large fish TAC, although the EPO TAC can be increased under a 70:30 impact ratio (Figs. ES10, ES11, and ES12). The WCPO large fish TAC was always higher than the current catch limits for all HCRs, except for HCRs 3 and 11 in the medium term and for HCR 11 across all years (Fig. ES10). The WCPO small fish TAC was always smaller than the current catch limits for all HCRs (Fig. ES11). The EPO TAC was larger only for HCRs 9 to 16, which had a higher EPO fisheries impact, with HCR 11 in the medium term being an exception (Fig. ES12). In the MSE, allocation of catch across the different fleet segments is set by the relative allocation of fishing mortality across fleets, which is set to the 2015-2022 baseline agreed upon by the JWG. These patterns are also affected by the fact that as the population biomass grows throughout the simulation, more biomass accumulates in older age classes, while average numbers of recruits and juveniles targeted by the WCPO small fish fleet segment and EPO may remain more stable. Furthermore, the TAC is dependent on estimates of numbers at age from the terminal year, which for young age classes are uncertain due to the lack of a recruitment or juvenile index. Thus, the estimation model tends to always estimate current recruitment to the average of the stock-recruitment function, leading to relatively low and stable small fish TACs.

4. *HCRs 1, 2, 3, 9, 10, and 11 had more instances of drastic (>25%) declines in catches due to severe management intervention resulting from breaching their respective LRP more often than other HCRs.*

HCRs 1, 2, 3, 9, 10, and 11 have longer lower tails in the annual catch violin plots in Fig. ES9, implying more instances of very low catch values. This is a result of more instances of severe management intervention due to their higher LRPs, which are breached more often than other HCRs. Indeed, worm plots of total TAC show that these HCRs have more instances where TAC declines dramatically (Fig. ES13) and these HCRs have the lowest 5th quantiles of TAC (Figs. ES9 and ES14).

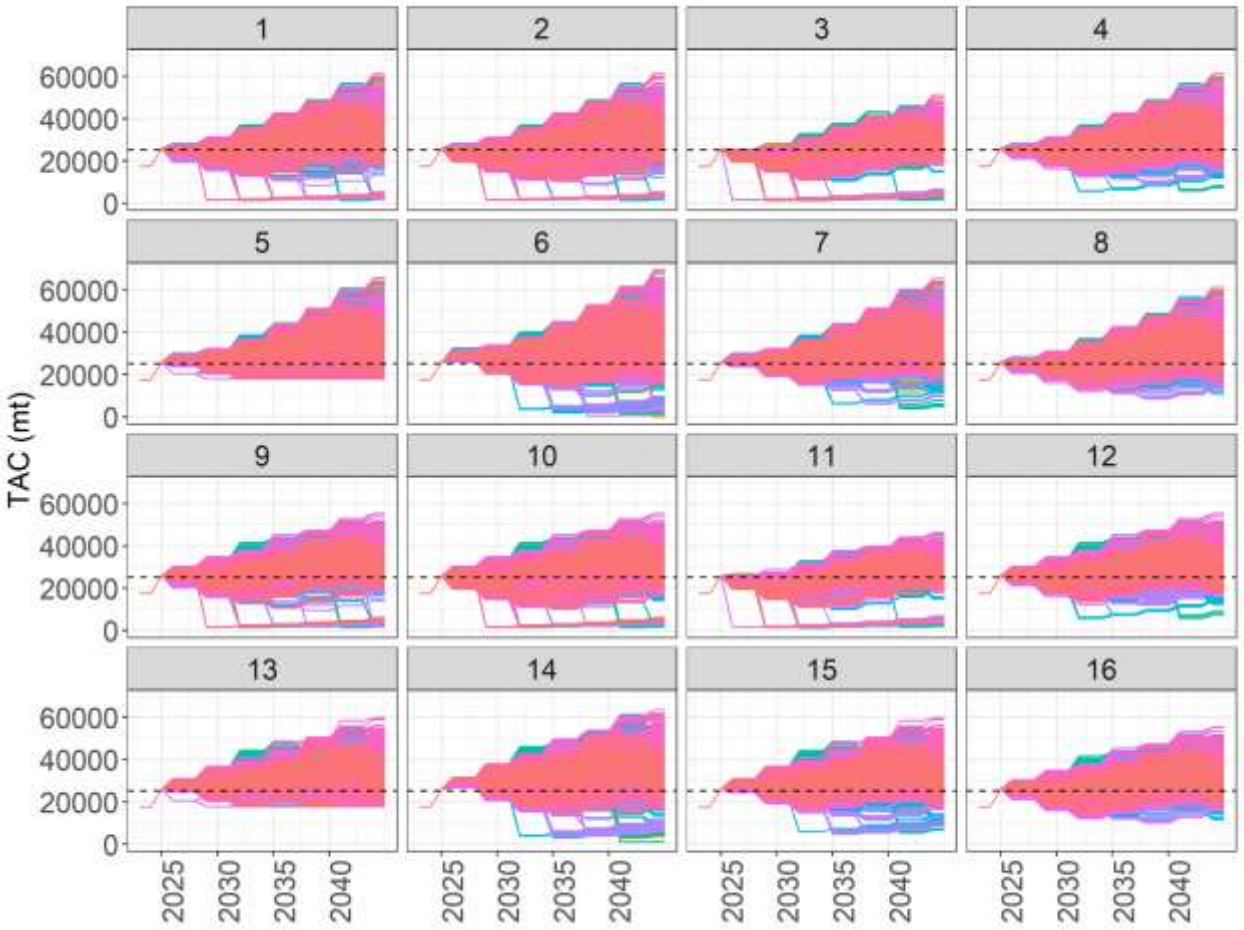

Figure ES13. Worm plots of the total allowable catch (TAC) set by each harvest control rule (HCR) for individual runs across all reference scenarios. Each panel presents the results for the labeled HCR. Trajectories represent separate iterations differing in simulated random recruitment deviates. The dashed line represents the current catch limit set by the WCPFC's CMM 24-01 and IATTC's Resolution C-24-02, plus EPO recreational catches for the calendar year 2023. For the IATTC's resolution, catch limits were based on half of the biennial TAC.

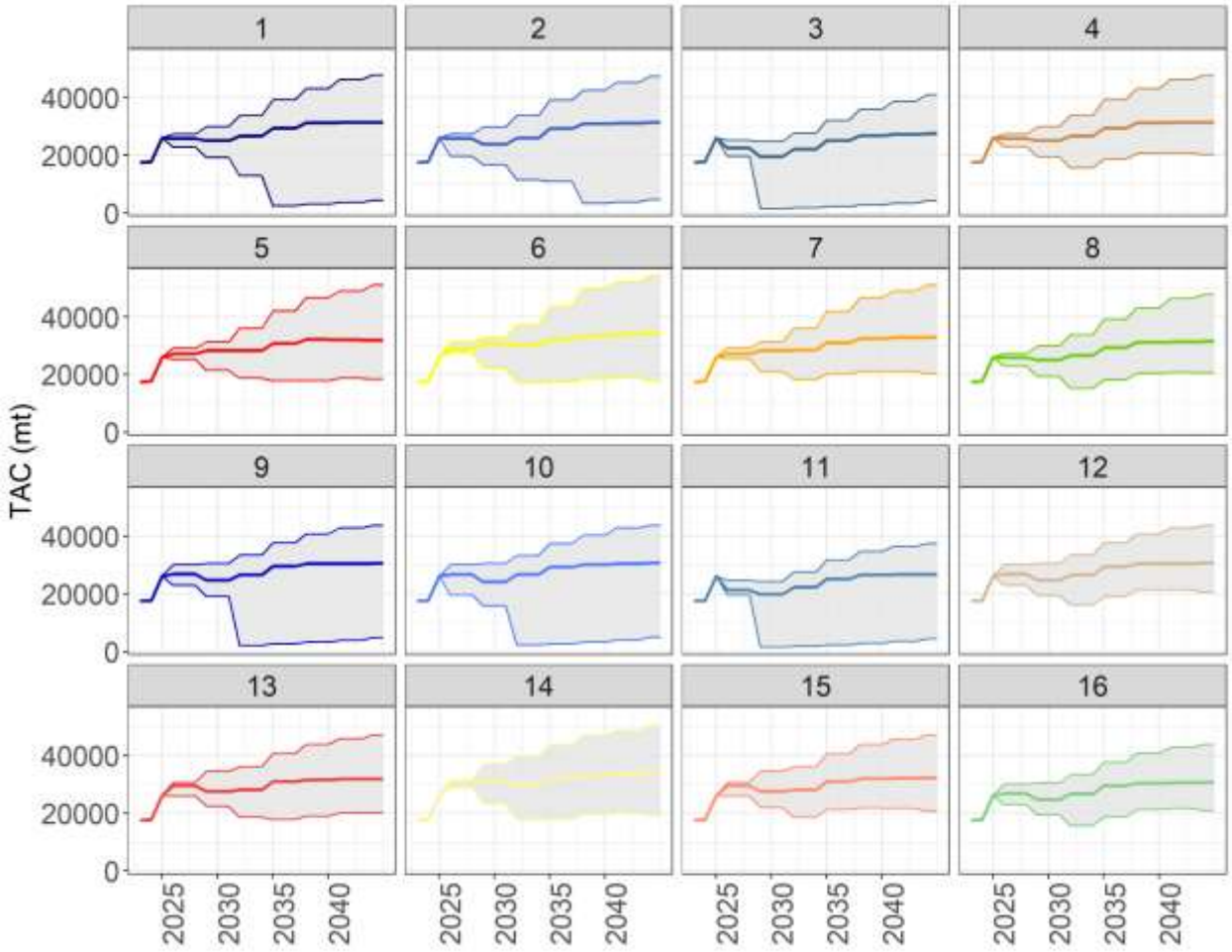

Figure ES14. Trends in median total allowable catch (TAC) set by each harvest control rule (HCR) under all iterations and reference scenarios. The grey shading represents trends in the 5th to 95th quantiles of TAC.

5. HCRs with a first control point (i.e., ThRP) closer to the target SSB (SSB associated with their F_{TARGET}) had lower catch stability.

HCRs 2, 5, 6, 10, 13, and 14 have a first control point that is closer to the target SSB than other HCRs (Table ES2). This leads to more frequent large reductions in F and lower stability (Figs. ES15, ES16, Table ES4). HCRs 3 and 11 have the largest differences between their first control point and the SSB associated with their F_{TARGET} and have the highest catch stability when SSB is at or above the LRP (Figs. ES15, ES16, Table ES4). Nonetheless, due to the built-in 25% limit on TAC change in each HCR, all HCRs met the stability objective.

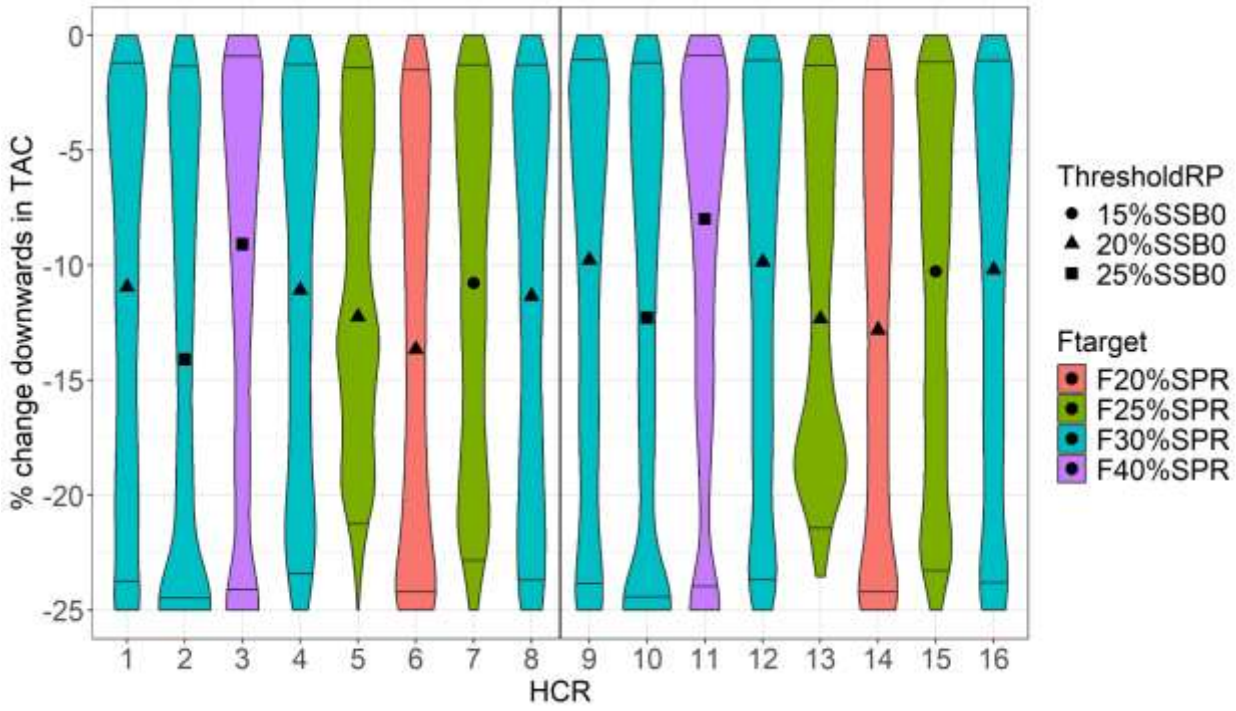

Figure ES15. Violin plots showing the probability density of downward changes in TAC between management periods when SSB≥LRP for each harvest control rule (HCR) across all iterations, reference scenarios, and simulation years. Each HCR is colored according to their F_{TARGET}. The marker inside each violin plot is the median downward change in TAC, and horizontal solid lines within each violin represent the 5th to 95th quantile range. Each symbol represents a different ThresholdRP, which is the first control point for each HCR and stands for Threshold Reference Point. The vertical solid line separates HCRs 9 to 16, which are tuned to an EPO:WCPO impact ratio of 30:70, but are otherwise the same as HCRs 1 to 8. EPO stands for Eastern Pacific Ocean and WCPO for Western Central Pacific Ocean.

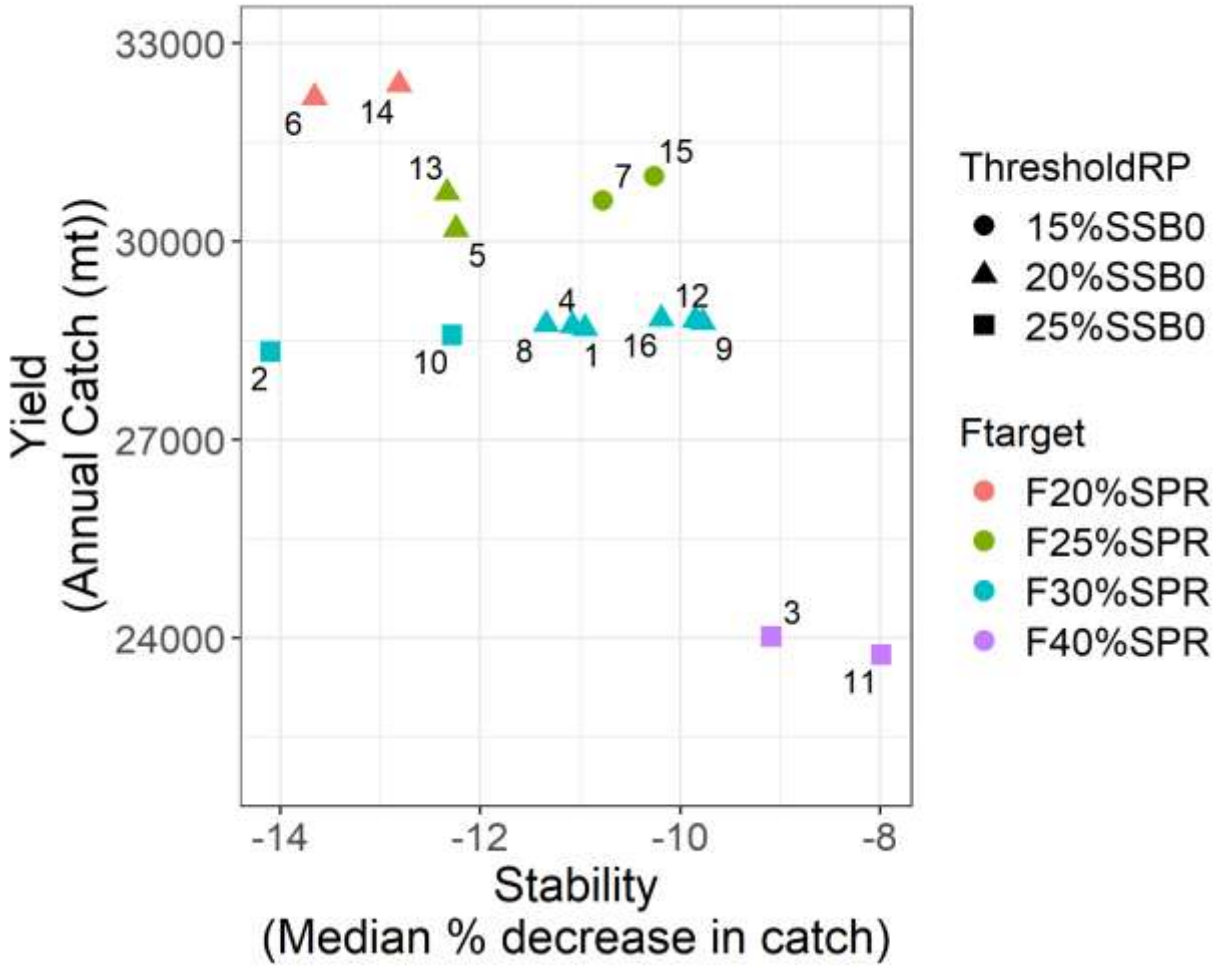

Figure ES16. Median annual total catch versus the median decrease in catch between management periods. Each HCR is labeled colored according to their F_{TARGET}. Each symbol represents a different ThresholdRP, which is the first control point for each HCR and stands for Threshold Reference Point.

6. All HCRs met the status objective of maintaining fishing mortality at or below the F_{TARGET} with at least 50% probability.

Despite uncertainties in stock productivity, recruitment variability, observation, estimation, and implementation, all HCRs met the status objective and maintained fishing mortality at or below the F_{TARGET} with at least 50% probability (Fig. ES17, Table ES4). For all HCRs, this probability was higher than 50% because the EM estimated fishing mortality as being lower than in the OMs, leading to a median F that was lower than the F_{TARGET} for all HCRs. The probability was highest for HCRs 1, 2, 3, 9, 10, and 11 because they had a higher LRP, resulting in drastic management interventions occurring more often. Once F fell to these low levels, it was slow to

increase due to the 25% limit in TAC changes between management periods, even if biomass rebuilt quickly, leading to median F being lower.

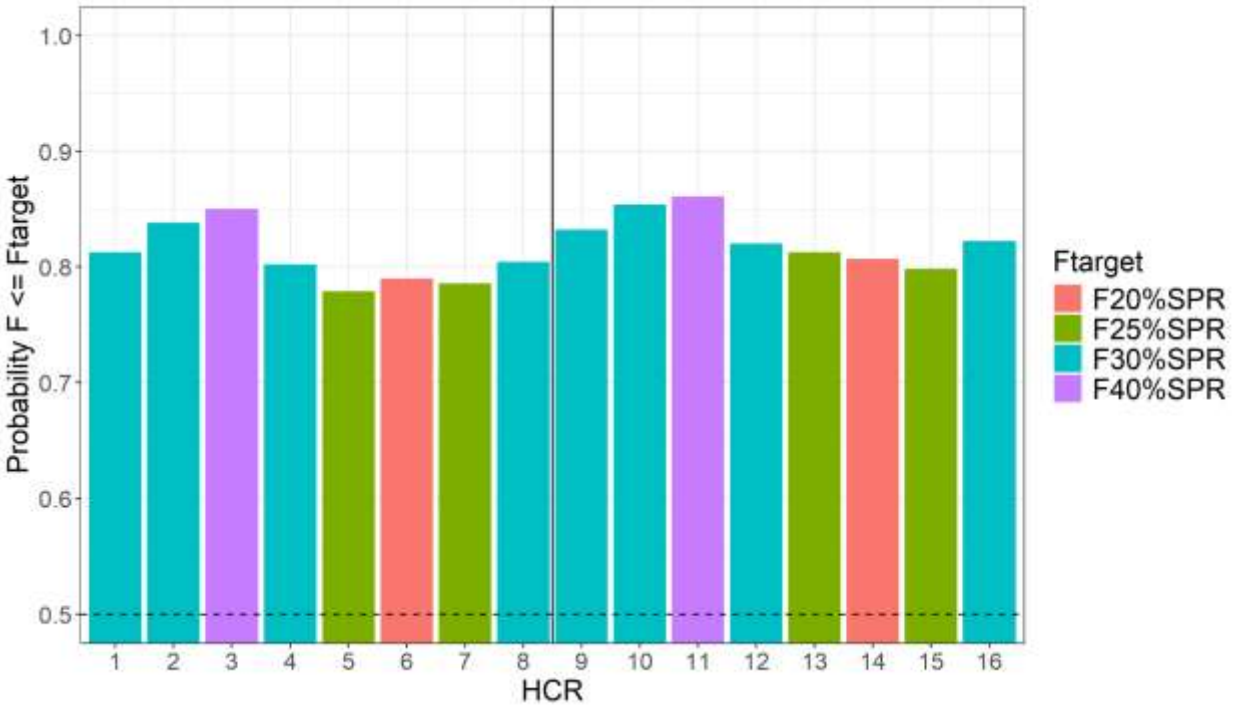

Figure ES17. Plot of the first status performance metric, the probability, for each harvest control rule (HCR), of fishing mortality (F, 1-SPR) being less or equal to the F_{TARGET} across all reference scenarios, iterations, and simulation years. Each HCR is colored according to their F_{TARGET}. The horizontal dotted line represents a 50% probability. The vertical solid line separates HCRs 9 to 16, which are tuned to an EPO:WCPO impact ratio of 30:70 but are otherwise the same as HCRs 1 to 8. EPO stands for Eastern Pacific Ocean and WCPO for Western Central Pacific Ocean.

7. *The different fisheries impact ratios only affected yield metrics but other performance metrics remained almost unchanged.*

HCRs 1 to 8 maintained the current WCPO:EPO fisheries impact ratio (about 80:20), while HCRs 9 to 16 were tuned to meet a 70:30 ratio. We would then expect higher yields for EPO fleets and lower yields for WCPO fleets under HCRs 9 to 16 (Figs. ES7, ES8, and ES9). All other metrics remained quite similar (Table ES4). Other performance metrics remained almost unchanged as shown in various tables and figures.

8. *Under robustness tests, all HCRs were robust to discard and effort-creep uncertainty, but performance deteriorated under extreme drops (40%) in recruitment over a 10-year period.*

Under robustness tests, where HCRs faced more unlikely but still possible situations, the performance naturally deteriorated as they were placed in more extreme conditions. Nonetheless, all HCRs were fairly robust to the "doubling of the discards" scenario and the "effort-creep"

scenario. However, although the degree was different among HCRs, all HCRs had difficulty in dealing with the "recruitment drop" scenario. This is expected because the MPs only respond to the assessed terminal SSB. Since PBF fully mature at 5 years of age and the abundance trend was informed only by the longline CPUE index, which informs the relative biomass of age 7 and older, it takes several years for the EM to detect a decline in SSB from the recruitment drop and for the MPs to initiate a significant reduction in catches. In the meantime, small fish catches remain an important component of the fishing mortality. Once the EM eventually detected the decrease in SSB, F was curtailed, and median SSB ultimately rebuilt to target levels for all HCRs. It is therefore important to carefully monitor the recruitment and also SSB through regular assessments to detect in a timely manner if a chronic decline in recruitment has occurred and to consider appropriate exceptional circumstances provisions to swiftly deal with such a situation. For more details, see the main body of the report.

Key Limitations

- Fleet selectivity was assumed to be constant at the current average of 2015-2022 levels throughout the simulation. If fleet operations and targeting behavior change in the future so that the size composition of catch of specific fleets differs widely from what was simulated, results from this analysis may no longer be applicable.
- The operating models were conditioned on data from 1983 onwards, thus the management procedures tested here are robust to uncertainty in productivity that was bounded by those historical observations. If future population dynamics strongly diverge from the past, results from this analysis may no longer be applicable.

Table ES4. Performance indicators for each harvest control rule (HCR) across all iterations, evaluation years, and operating models. SSB refers to spawning stock biomass, LRP to limit reference point, $SSB_{F=0}$ refers to unfished spawning stock biomass, F refers to fishing mortality measured as 1-SPR where SPR is spawning potential ratio, TAC refers to total allowable catch, WCPO refers to Western Central Pacific Ocean and EPO refers to Eastern Pacific Ocean. Note that to ensure that for all indicators a higher value is better, here we reversed the performance metrics showed in Figures ES5 and ES7 to be the probability of SSB≥LRP and of SSB≥20%$SSB_{F=0}$. The % change upwards in TAC (% change TAC +) was set to negative so that high values (smaller -) are better. The % change downwards does not include years when SSB is below LRP as provided by the management objective. The value including years when SSB is below LRP is provided in the main body of the report. The 2026 TAC[2] is the total TAC and the TAC for each fleet segment that could be applied in 2026 if each of the HCR would be adopted. It is calculated based on biomass status estimated by EM. Color shadings reflect the range of each column. Highest levels have dark green, lowest light yellow, and different shades of green to yellow are in between. As there is no optimal impact, the EPO impact column does not have a color.

Performance Indicators

Reference Set

	Prob SSB => LRP	Prob SSB => 20%SSBo	Prob F <= Ftarget	Prob SSB => SSBtarget	% change TAC +	% change TAC -	EPO Impact	Median annual catch	Median years 5-10 annual catch	Median years 11-23 annual catch	Median WCPO large fish annual TAC	Median WCPO small fish annual TAC	Median EPO annual TAC	2026 TAC	2026 TAC WCPO large fish	2026 TAC WCPO small fish	2026 TAC EPO
1	93	87	81	62	-14	-11	23	28685	26744	31094	16174	4093	6794	25868	14836	4512	6520
2	94	87	84	64	-15	-14	23	28330	26054	30691	15618	4069	6794	25868	14836	4512	6520
3	91	91	85	56	-17	-9	24	24026	21135	26361	13472	3346	5971	24366	14836	3844	5686
4	99	86	80	61	-13	-11	23	28722	26745	31124	16221	4102	6794	25868	14836	4512	6520
5	100	81	78	66	-13	-12	22	30183	28894	32227	16965	4617	7054	27485	14836	5161	7488
6	99	76	79	76	-14	-14	22	32174	31286	34249	18243	5063	7609	29437	14836	5939	8662
7	100	82	79	67	-13	-11	23	30616	28940	32814	17330	4557	7192	27485	14836	5161	7488
8	100	86	80	61	-14	-11	23	28741	26746	31127	16222	4101	6794	25868	14836	4512	6520
9	93	86	83	63	-13	-10	32	28773	26503	30537	13378	3722	10528	27942	14073	4392	9476
10	93	87	85	65	-16	-12	32	28582	25973	30368	13242	3722	10433	27942	14073	4392	9476
11	91	91	86	56	-16	-8	32	23748	21378	26147	10877	3023	8632	23653	10724	3844	9085
12	99	86	82	62	-12	-10	33	28812	26505	30572	13414	3722	10568	27942	14073	4392	9476
13	100	82	81	68	-13	-12	30	30735	29380	31768	14567	4160	11175	29323	14836	5010	9476
14	99	77	81	77	-15	-13	31	32369	32077	33617	15040	4592	11323	30061	14836	5749	9476
15	100	81	80	67	-12	-10	32	30988	29413	32137	14567	4108	11323	29323	14836	5010	9476
16	100	86	82	62	-12	-10	33	28826	26507	30582	13413	3722	10565	27942	14073	4392	9476

[2] Bycatch allowance described in paragraph 4 of WCPFC CMM2024-01, notably 200t by NZ and 40t by Australia, is not considered part of TAC in this MSE thus not reflected in the calculation of 2026 TAC

2. Introduction

Management strategy evaluation (MSE) is a process whereby the performance of a set of management procedures relative to management objectives put forward by managers and stakeholders is assessed under a range of uncertainties using a closed-loop computer simulation (Punt et al. 2016). To assess the impact of each management procedure on the stock and fisheries, an MSE simulates all aspects of a management procedure, from data collection, to stock status estimation, to management implementation. Their simulated performance relative to the pre-agreed upon management objectives then helps managers and stakeholders select between them.

Pacific Bluefin tuna (*Thunnus orientalis*, hereafter PBF) is considered a single North Pacific-wide stock and is managed by two Regional Fisheries Management Organizations (RFMOs), the Western and Central Pacific Fisheries Commission (WCPFC) and the Inter American Tropical Tuna Commission (IATTC). PBF is currently assessed by the Pacific Bluefin Working Group (PBF WG or WG) of the International Scientific Committee for Tuna and Tuna-like Species in the North Pacific Ocean (ISC). Previous assessments showed that PBF had been subject to intense fishing pressure, and is mainly harvested by Chinese Taipei, Japan, Korea, Mexico, and the United States (U.S.). Records of landings go back to the early 1800's in Japan and to the 1900's in the U.S. (ISC 2024a). Catches have fluctuated widely historically, and were lowest in 1990 (ISC 2024a). Spawning stock biomass (SSB) relative to equilibrium unfished SSB ($SSB_{F=0}$) was at its lowest in 2009 and 2010, at 2% $SSB_{F=0}$ (ISC 2024a). That's when management measures were first introduced by the WCPFC to maintain effort below 2002-2004 levels to help rebuild the stock (WCPFC 2009), followed by similar measures in the IATTC area in 2012 (IATTC 2012). In 2013 and 2014 when SSB remained below 4% $SSB_{F=0}$ (ISC 2024a), additional measures to halve the catch of juvenile fish (<30 kg) were introduced (WCPFC 2013, WCPFC 2014, IATTC 2014). In 2017, an interim "rebuilding period" management procedure specifying a first rebuilding target of 6.3%$SSB_{F=0}$ (the median relative SSB from 1952-2014) and a second rebuilding target of 20%$SSB_{F=0}$ was adopted (WCPFC 2017). It specified that catch limits be applied to help reach the specific rebuilding targets over a pre-defined period based on stock status projections conducted by the ISC (WCPFC 2017). The more stringent management measures were successful in rebuilding the stock, which reached the second rebuilding target in 2021 (ISC 2024a). After rebuilding, some catch increases were adopted based on stock projections results (WCPFC 2023a, WCPFC 2024, IATTC 2024), according to an interim management procedure (WCPFC 2023b, IATTC 2023a), but no long-term management procedure or reference points have been adopted for PBF.

To help inform development of a long-term management procedure for PBF, the WCPFC NC and IATTC requested, via the WCPFC NC and IATTC PBF Joint Working Group (JWG), that the ISC PBF WG develop an MSE, to be presented to the JWG in 2025 (JWG 2022). In preparation for the MSE, the ISC organized two workshops to promote understanding of MSE, its elements, and roles of managers and stakeholders (ISC 2018a, 2019). In 2019, the JWG proposed an initial

set of more than 100 candidate harvest control rules (HCRs) to be tested in the PBF MSE (WCPFC NC 2019). In 2023, after a preliminary evaluation of these HCRs by the ISC (Tommasi et al. 2023a), the JWG put forward a set of 12 HCRs to be tested (WCPFC NC 2023a), and also agreed on the management objectives and performance metrics with which to evaluate performance of the HCRs in the MSE (WCPFC NC 2023b). After being presented with a set of preliminary results at an intersessional meeting in February 2025, the JWG finalized the set of HCRs to be tested in the MSE.

This report provides a description of the PBF MSE framework developed by the ISC PBF WG, and of results of the PBF MSE, including an assessment of performance of the 16 finalized HCRs with respect to the PBF management objectives specified by the JWG (NC 2023b). The MSE evaluated all the candidate HCRs and associated reference points proposed by the JWG. The work represents the final set of MSE analyses in support of development of a long-term management procedure for PBF. Section 3 contains background information on the biology, fisheries, and management of PBF, as well as management objectives and performance indicators, reference points, and candidate harvest control rules, and uncertainties considered in this set of MSE simulations. Section 4 provides a detailed overview of the MSE framework, including operating and estimation models, while Section 5 is focused on the MSE results.

3. Background

3.1 Biology

Because of its high commercial value, large body size, strong swimming ability, and physiological traits, various biological studies have been carried out on PBF for a long time. That information was implicitly or explicitly incorporated into the current stock assessment model, which is the basis of the operating models and estimation model in this MSE.

3.1.1 Stock Structure and Distribution

Bluefin tunas in the Pacific and Atlantic Oceans were once considered a single species (*Thunnus thynnus*) with two subspecies (*Thunnus thynnus orientalis* and *Thunnus thynnus thynnus*, respectively), but are now recognized as distinct species (*Thunnus orientalis* and *Thunnus thynnus* for Pacific and Atlantic bluefin tunas, respectively) based on genetic and morphometric studies (Collette 1999). This taxonomic distinction is adopted by pertinent tuna RFMOs, the Food and Agriculture Organization of the United Nations (FAO), and ISC.

PBF are mainly distributed in subtropical and temperate latitudes between 20ºN and 50°N, although they are occasionally encountered in tropical waters and in the southern hemisphere (Fujioka et al. 2015). There are several spawning grounds of PBF in the western North Pacific Ocean (WPO) (Ohshimo et al. 2018, Tanaka et al. 2020), and two

of them have been considered major spawning grounds: 1) waters between the Ryukyu Islands in Japan and the eastern coast of Taiwan islands; and 2) the southern portion of the Sea of Japan (Schaefer 2001). Conversely, no evidence of PBF reproduction has been observed in the eastern Pacific Ocean (EPO) despite there being enough older PBF for reproduction (Dewar et al. 2022). A study on the natal origin of adult PBF caught either in the waters around the Ryukyu Islands or in the Sea of Japan indicated that they originated from both of these spawning grounds (Uematsu et al. 2018). Similarly, elemental analysis of otoliths from adult PBFs caught in the waters around Taiwan indicated that they also originated from both known spawning grounds (Rooker et al. 2021). Additionally, age-1 PBFs caught in the EPO have been traced back to both known spawning grounds using trace elements in their otoliths (Wells et al. 2020). These findings support the notion of a single stock for PBF, as there is no significant difference in the natal origin between the two known spawning grounds. Genetics and tagging studies (e.g., Bayliff 1994, Tseng and Smith 2012) further support the assumption of a single stock for PBF. A review conducted by Nakatsuka (2019) concluded that there is no evidence exclusively suggesting the existence of multiple stocks after examining available genetic and reproductive information, otolith and vertebrae data, and fishery data. As a result, a single highly migrating stock is assumed in the PBF assessment within the ISC and is acknowledged by the RFMOs (WCPFC and IATTC).

Despite substantial inter-annual variations in movement in terms of the numbers of migrants, timing of migration, and migration routes, the movements of PBFs are among the most extensively documented among highly migratory species. Mature adults in the WPO typically migrate northward to feeding grounds following spawning (Fujioka et al. 2024), although a small proportion of fish may move southward or eastward (Itoh 2006). Fish aged 0-1 that have hatched in the waters surrounding the Ryukyu Islands and eastern Taiwan main Island migrate northward with the Kuroshio Current during the summer as they grow, while age-0 fish that have hatched in the Sea of Japan migrate along the coastlines of Japan and Korea (Inagake et al. 2001, Itoh et al. 2003). Depending on oceanic conditions, an undetermined portion of immature fish aged 1-3 in the WPO makes a seasonal clockwise eastward migration across the North Pacific Ocean (stable isotope in muscle tissues: Tawa et al. 2017, Madigan et al. 2017), spending several years as juveniles in the EPO before returning to the WPO (Inagake et al. 2001). The drivers behind this trans-Pacific migration have been hypothesized to be limitations in food sources in the WPO and favorable oceanographic conditions (Polovina 1996). While PBFs are in the EPO, juveniles make seasonal north-south migrations along the west coast of North America (Kitagawa et al. 2007, Boustany et al. 2010). In spring, they are found in the waters off the southern coast of Baja California, and as summer approaches and waters warm, they move northwest into the southern California Bight. By fall, PBF distribute in the waters off central and northern California. After spending 3-4 years in the EPO, PBF migrate westward, presumably for spawning,

as no spawning grounds have been observed outside of the WPO. This westward migration typically occurs from December to March as they begin their migration along the coast of California (Boustany et al. 2010). The considerable seasonal (Fujioka et al. 2021) and interannual variations in trans-Pacific movement make it challenging to quantify migration rates accurately.

3.1.2 Reproduction

PBF are known as iteroparous spawners, meaning they spawn multiple times throughout their life. Spawning events are confined to specific areas and seasons: from April to July in the northwestern Pacific Ocean, and from July to August in the Sea of Japan. These conclusions are drawn from histological studies on PBF gonads (Yonemori 1989, Ashida et al. 2015, Okochi et al. 2016, Ashida et al. 2021, Ashida et al. 2022) and the distribution of PBF larvae (Yabe et al. 1966). Some histological studies showed that approximately 80% of fish caught in the Sea of Japan from June to August and weighing around 30 kg (corresponding to 3 years old or age 2.75 in the assessment model) were mature (Tanaka 2006, Okochi et al. 2016). Nearly all fish caught in the waters surrounding the Ryukyu Islands and eastern Taiwan Island were larger than 60 kg (> 150 cm fork length (FL)) (Chen et al. 2006, Ashida et al. 2015). These fish were at least 5 years old (age 4.75 in the model) and all were mature. Although large PBFs have also been observed in the EPO, particularly in recent years in Southern California, Dewar et al. (2022) reported no evidence of PBF reproduction in the EPO based on histological examinations of ovaries and ichthyoplankton data. Although there is a clear difference in age/size of PBF between the spawning areas, it was also confirmed that the spawners from different natal origins appeared in both major spawning areas. These data indicate that PBF segregate on each spawning ground by age/size and exhibit age/size-based migration, at least during the spawning season.

3.1.3 Growth

Age determination of PBF has been established through various methods such as vertebral ring counts (Aikawa and Kato 1938), scale ring counts (Yukinawa and Yabuta 1967), tag-recapture studies (Bayliff et al. 1991), and otolith observations (daily increments: Foreman 1996; annual rings: Shimose et al. 2008, 2009, Shimose and Takeuchi 2012). A standardized technique for age determination of PBF based on otolith samples was developed at the Pacific Bluefin and North Pacific Albacore Tuna Age Determination Workshop in 2014 (Shimose and Ishihara 2015) held by the ISC. This workshop initiated the large-scale age determination of annuli rings of otolith samples for PBF collected from troll, purse seine, set-net, handline, and longliners between 1992 and 2014. The work also examined the daily increments of otolith samples caught by the troll and set-net fisheries on the west coast of Japan between 2011 and 2014. In addition to analyzing the number of opaque zones in otoliths, post-bomb radiocarbon

dating was used to validate age estimation, and the results were consistent with otolith thin sections (Ishihara et al. 2017). Fukuda et al. (2015), further contributed by estimating growth curves based on the analysis of annuli data from 1,782 fish (70.5-271 cm in fork length [FL], corresponding to 1-28 years old) and daily increment data from 228 fish (18.6-60.1 cm in FL, corresponding to 51-453 days old after hatching). The examination of variances in length at age was also conducted using a dataset consisting of over 7,000 paired age-length samples collected between 1992 and 2022 (Tsukahara et al. 2024). These paired length at age samples showed a gradual decrease in CV of length across ages from 15% to 7% for ages 2-7, stabilizing at 5-6% for those aged 8 and older. It was hypothesized that the CV of length at ages 0 and 1 would be higher despite the absence of estimates for these ages (referenced as Figure 2-3 in the ISC 2024 stock assessment report).

In 2023, Ishihara et al. analyzed the same dataset using different sampling methods to estimate growth parameters and revealed that the growth rate and asymptotic length were robust and estimated similarly to Fukuda et al. (2015) regardless of sampling methods. The estimated possible estimation error of the growth parameters in this study was incorporated as one of the sources of the uncertainty in the OMs of the MSE (see Sections 4.1.1.2.1.2 and 4.1.1.3.1). The growth curve assumed in the OM and EM are generally consistent with previous studies (Shimose et al. 2009, Shimose and Takeuchi 2012, Shimose and Ishihara 2015, Fukuda et al. 2015); fish grow rapidly up to age 5 (approximately 160 cm FL), after which growth slows down. By age 12, the fish reach 226 cm FL on average, corresponding to 90% of the maximum FL for PBF. Fish larger than 250 cm FL are primarily older than age 20, indicating that the potential lifespan of this species is at least 20 years. Fish larger than 300 cm FL are rarely found in commercial catches.

The otolith samples were also integrated into the 2022 assessment model as conditional age-at-length (CAAL) data to estimate CVs of length at age. This integration was imperative as the model's expectations necessitated an understanding of the age structure of the population, with CAAL data predominantly used within population dynamics models (Piner et al. 2016, Lee et al. 2017). A quasi-age-structured production model with recruitment variation (ASPM-R) was produced, wherein the recruitment deviations and length-based selectivity were specified at MLE from the previous assessment model (Tsukahara et al. 2024). The CV estimate at age-0 was approximately 28%, and gradually decreased to around 4% by age 3. The large CV at age-0 was potentially attributed to the variation in size among age-0 fish originating from the two spawning grounds with distinct main spawning periods, although the assessment model assumed that age-0 fish originated from a singular spawning ground for the sake of simplicity.

The length-weight relationship of PBF was based on the von Bertalanffy growth curve used in the stock assessment (ISC 2024a, Section 4.1.1.2.1.4).

3.1.4 Natural Mortality

Natural mortality coefficients (M) are one of the most difficult parameters to be reliably estimated in the stock assessment model, based on simulation studies (Lee et al. 2011, Lee et al. 2012). M for the 2024 assessment was assumed to be age-specific: high at a young age, decreasing as fish mature, and stabilizing afterwards. Natural mortality for age-0 fish was derived from conventional tagging studies conducted on PBF (Takeuchi and Takahashi 2006, Iwata et al. 2012a, Iwata et al. 2014). In the absence of direct estimates beyond age 0, natural mortality for age-1 fish was estimated based on length-adjusted M values derived from conventional tagging studies conducted on southern bluefin tuna (Polacheck et al. 1997, ISC 2009). This adjustment accounted for the differences in the life-history between PBF and southern bluefin tuna. A constant M for mature fish was then derived from the median value obtained through a suite of empirical and life-history based methods to represent age 2 and older fish (Aires-da-Silva et al. 2008, ISC 2009). In 2023, additional estimations of the M of age 2 and older fish were performed through a suite of empirical and life history parameter methods to comprehend the possible range of this parameter (Lee and Tommasi 2023). The choice of age-based M schedule in the OM is shown in section 4.1.1.3.

3.2 Fisheries

The main fisheries from each fishing nation and the RFMOs' management measures are summarized in this section, and the fleet structures and associated data used in the OMs are summarized in section 4.1. Although PBF catch records are sparse prior to 1952, there are some PBF landing records dating back to 1804 from coastal Japan and the early 1900s for U.S. fisheries operating in the EPO. PBF catch estimates were high from 1929 to 1940, with a peak catch of approximately 47,635 t (36,217 t in the WPO and 11,418 t in the EPO) in 1935 but sharply declined during World War II. PBF catches increased significantly after 1949 as Japanese fishing activities expanded across the North Pacific Ocean (Muto et al. 2008). By 1952, most fishing nations had adopted a more consistent catch reporting process. From 1952 to 2022, annual catches of PBF by ISC member countries exhibited wide fluctuations (Figure 2A). Among these nations, PBF are mainly harvested by five countries, with Japan having the largest catches, followed by Mexico, the U.S., Chinese Taipei, and Korea. Although catches in tropical waters and in the southern hemisphere have historically been small and sporadic, there was a notable increase in the southern hemisphere catch in 2020, reaching around 50 t (WCPFC 2023a). During this period, reported catches peaked at 40,383 t in 1956 and 34,612 t in 1981, reaching the low of 8,653 t in 1990, followed by an increase to over 30,000 t in 2000 and 2004 before declining to about 12,000 t in 2017.

Although PBF are caught by various fishing gears, most of the catch is from purse seine fisheries (Figure 2B). In Japan, major active PBF fisheries include longlines, purse seines, trolling, and set-nets. Most PBF fisheries in Japan operate inside its Exclusive Economic Zone

(EEZ). The distant-water longline fisheries also catch PBF, but their catch is small compared to other active fisheries. Overall, total annual catches by Japanese fisheries have fluctuated between a maximum of 34,000 t in 1956 and a minimum of 6,000 t in 1990 (calendar year). For more details of Japanese fisheries taking PBF see Yamada (2007).

In the U.S., two major active PBF fisheries (purse seine and recreational/sport fisheries) catch PBF off the west coast of North America. Initially, the U.S. purse seine fishery harvested a large amount of PBF for canning in the waters off Baja California until Mexico established its EEZ in 1976, leading to the exclusion of U.S purse seine vessels. Subsequently, after 1983, the U.S. purse seine fishery opportunistically caught PBFs (Aires-da-Silva et al. 2007). Currently, the majority of U.S. PBF catch is from recreational fisheries operating in U.S. and Mexican waters (Heberer and Lee 2019).

The Mexican purse seine fishery experienced rapid development after Mexico established its EEZ and is now the most important large pelagic fishery in Mexico. This fishery is closely monitored through an at-sea observer program with 100% coverage, captains' logbooks and Vessel Monitoring Systems (VMS), and recently, stereoscopic cameras (Dreyfus and Aires-da-Silva 2015, Dreyfus 2018). While purse seine sets target yellowfin tuna (*Thunnus albacares*, the dominant species in the catch) in tropical waters, PBFs are caught near Baja California for farming. The Mexican PBF catch history recorded three large annual catches (above 7,000 t) in the years 2004, 2006, and 2010.

In Korea, PBF are primarily caught by the offshore large purse seine fishery (OLPS), although there have been reports of small amounts of catches from the coastal fisheries in recent years. The catch of the OLPS fishery was below 500 t until the mid-1990s, peaked at 2,601 t in 2003, and since then has fluctuated between 600 t and 1,900 t. In 2018, the catch of the OLPS fishery was 523 t. The main fishing ground of the OLPS fishery is off Jeju Island, with the vessels occasionally operating in the Yellow Sea and the East Sea (Yoon et al. 2014, Lee et al. 2018).

The amount of PBF caught by the Taiwanese fisheries (including small-scale longline, purse seine, large-scale pelagic driftnet, set net, offshore and coastal gillnet, and bottom longline fisheries) was small (<300 t) between the 1960s and the early 1980s. After 1984, the total landings gradually increased to over 300 t, mostly due to the small-scale longline vessels (<100 gross registered tonnage (GRT)) targeting spawners for the sashimi market from April to June. The highest observed catch was 3,000 t in 1999, followed by a rapid decline to less than 1,000 t in 2008 and a subsequent drop to about 200 t in 2012. The catch then slightly increased to around 500 t in 2018 and showed a significant increase to more than 1,800 t in 2022.

3.3 Management

The trend in catch is associated with the RFMOs' management efforts as well as the stock condition. In 2010, the WCPFC introduced the first conservation and management measure (CMM) to regulate the fishing effort towards PBF, followed by catch limit management for

small PBF (<30 kg in body weight) within its convention area (WCPFC CMMs 2009-07 and 2010-04). The catch limit was further reduced in 2014 (WCPFC CMM 2013-09) and 2015 (WCPFC CMM 2014-04) to ensure that the catches of small PBF remained below 50% of the 2002-2004 average level, and the catches of large PBF (>30 kg in body weight) remained below the 2002-2004 average level. In the IATTC area, conservation and management measures were introduced in 2012 (IATTC resolution C-12-09) to regulate the catches for all size ranges of PBF within its convention area. Additional reductions in catch limits were established in 2015 to ensure that total commercial catches remained below 6,200 t. In 2021, both the WCPFC and IATTC adopted CMMs for PBF to be implemented during 2022-2024, allowing for an increase in the large PBF catch limits. In the current (2025 calendar year) measures (WCPFC CMM 2024-01 and IATTC Resolution C-24-02), the catch limits were further increased to ensure that total catches in the WCPFC and IATTC convention areas remained below 16,994 and 6292.5 t (half of biennial IATTC limit for commercial catches during 2025 and 2026), respectively.

3.4 Management Objectives and Performance Indicators

The overarching objectives for PBF management specified in the current management procedure are to support thriving Pacific bluefin tuna fisheries across the Pacific Ocean, maintain or restore the stock at levels capable of producing maximum sustainable yield, maintain an equitable balance of fishing privileges among CCMs and to find an equitable balance between the fisheries in the western and central Pacific Ocean (WCPO) and those in the EPO (WCPFC 2023b, IATTC 2023a). However, more specific candidate management objectives were identified and agreed upon by the JWG to evaluate the performance of the different candidate harvest control rules in the MSE (WCPFC NC 2023b). The management objectives are outlined in Table ES1 and relate to four categories: safety, status, stability, and yield. To quantitatively evaluate the performance of the HCRs tested relative to the management objectives, the JWG represented these management objectives into quantitative performance indicators, also outlined in Table ES1.

The safety objective relates to conservation of the stock and maintenance of its reproductive potential and states that there should be a less than 20% probability of the stock falling below the limit reference point (LRP). As there is no agreed-upon LRP for PBF, the JWG provided different LRPs, associated with different HCRs (Table ES1), with which to evaluate this performance metric. For some HCRs, the LRP also acts as a control point for a management action (see Sections 3.5 and 3.6). The safety performance indicator is the probability that SSB< LRP in any given year of the evaluation period, calculated as the number of times for each HCR that SSB was below the LRP across all the simulation years, reference set of operating models (OMs), and iterations. As the LRP is specific to each of the HCRs being considered, the PBFWG also developed two additional safety performance indicators with a reference point common to all HCRs, the probability that SSB< $20\%SSB_{F=0}$, the second rebuilding target for this species, or that SSB< $7.7\%SSB_{F=0}$, the IATTC's interim LRP for

tropical tunas, in any given year of the evaluation period. The evaluation period over which performance metrics are calculated is from year 2026, when the first simulated TAC is applied, through the end of the simulation in the year 2045.

The status objective relates to fishing mortality being around the target level, or the level decision-makers want to achieve. It is "To maintain fishing mortality at or below F_{TARGET} with at least 50% probability". As for the LRP, there is no fishing mortality target reference point (F_{TARGET}) that has been adopted for PBF, and the JWG proposed a set of potential F_{TARGET}s, each associated with a particular HCR (Table ES1). The status performance indicator is the probability that that $F \leq F_{TARGET}$ in any given year of the evaluation period, calculated as the number of times for each HCR that F was at or below the F_{TARGET} across all the simulation years, reference set of operating models (OMs), and iterations. A second status performance metric was also developed by the JWG, and is the probability that SSB is below the equivalent biomass depletion levels associated with the candidates for F_{TARGET}, calculated as the number of times for each HCR that SSB was below the equivalent biomass depletion levels associated with the candidates for F_{TARGET} across all the simulation years, reference set of OMs, and iterations.

The stability objective focuses on maintaining low variability in catches between management periods to promote consistency for the fishing industry. It is "to limit changes in overall catch limits between management periods to no more than 25%, unless the ISC has assessed that the stock is below the LRP". There are two stability performance indicators, one looking at increases in catch limits, one looking at decreases. They are the percent change upwards or downwards in catches between management periods excluding periods when SSB<LRP. Note that catch limits are set every three years and kept constant in between management periods, so this metric compares the catch limit to the catch limit set three years prior.

There are three yield objectives. The first is to "maintain an equitable balance in proportional fishery impact between the WCPO and EPO". Its performance metric is the median fishery impact (in %) on SSB in the terminal year of the evaluation period by WCPO and EPO fisheries. Fishery impact examines the effect of a particular fishery group (e.g. by gear or region) on SSB. It is computed by simulating what the SSB would have been in the absence of catches from that fishery group and depends not only on the amount of catch of that fishery group but also on the size composition of that catch. For instance, catching juvenile fish would have a larger impact on SSB than catching the same amount of mature fish as those fish are removed before they reach their full growth potential or reproduce (Wang et al. 2009). Proportional fishery impact is the fishery impact of a particular group relative to the impact of all the fisheries combined and has become a quantity routinely computed and presented to managers in the PBF stock assessment (ISC 2024). Here we calculate the proportional fishery impact (in %) between the EPO and WCPO fleet segments using the algorithm developed by Wang et al. (2009), customized for PBF in ISC (2013), and modified to be used with the output of the PBF MSE OMs by Tommasi et al. (2023b).

The second yield management objective is "to maximize yield over the medium (5-10 years) and long (10-30 years) terms, as well as average annual yield from the fishery". The associated performance indicators are the expected annual catch over years 5-10, years 10-30, or in any given year of the evaluation period, by fishery. We present these performance indicators for the total catch including discards as well as the total allowable catch (TAC) for each fleet segment, which are WCPO small fish, WCPO large fish, and EPO. We use the same performance indicators to evaluate the last yield objective, "to increase average annual catch in all fisheries across WCPO and EPO" by summing the total annual catch from all three fleet segments over years 5-10, years 10-30, or in any given year of the evaluation period.

3.5 Reference Points

Reference points are the benchmarks to which current stock level and/or fishing intensity are compared to when assessing the status of the stock. When used in an HCR, reference points can also correspond to control points that, when breached, trigger a management action. In this MSE, three types of reference points are used; and they often also act as control points. They are target reference points (TRPs), LRPs, and threshold reference points (ThRPs). It is important to note that the LRPs and TRPs in the HCRs serve both as control points of management actions and as measuring sticks to evaluate performance (i.e., reference points). However, control points can differ from the LRPs and TRPs. LRPs and TRPs, in principle, can also simply play the role of reference points to evaluate the performance of HCRs. In these cases, the level of the LRPs and TRPs would only be used as measuring sticks without affecting the management actions under the HCRs.

A TRP refers to a desired state that the management wants to achieve. The TRPs for all the HCRs evaluated in this MSE are based on fishing mortality (F) and referred to as F_{TARGET}. Fishing mortality is defined as 1-SPR, where SPR is the spawning potential ratio, or the SSB per recruit relative to the unfished population. The F_{TARGET}s are labeled as FSPRx%, where x refers to an SPR value. For instance, FSPR40% represents an F that leads to a SSB per recruit that fluctuates around 40% of the unfished SSB per recruit (i.e., removing about 60% of the SSB or a fishing intensity of 0.6). In contrast, a TRP of FSPR20% leads to a SSB per recruit that is around 20% of unfished SSB per recruit (i.e., a fishing intensity of 0.8). Therefore, an F_{TARGET} of FSPR20% implies a higher fishing mortality and fishing harder than FSPR40%, and the average level of SSB desired is lower. The candidate HCRs put forward by the JWG for testing in the MSE are associated with one of four F_{TARGET}s: FSPR40%, FSPR30%, FSPR25%, and FSPR20%.

Note that no F_{TARGET} reference point has been adopted for PBF. In the IATTC, when MSY-based reference points can be reliably estimated from stock assessment models, MSY-based interim TRPs are used for management (IATTC 2023b). By contrast, the Scientific Committee (SC) of the WCPFC recommends using MSY-based reference points as limits, when steepness parameter of the stock-recruitment relationship can be reliably estimated (WCPFC SC 2011). However, PBF is a Level 2 stock as there is uncertainty about steepness, and MSY-based

reference points cannot be reliably estimated. In those instances the WCPFC recommends using proxies based on FSPRx% for fishing mortality reference points and based on a fraction of unfished SSB for biomass reference points (WCPFC SC 2011). For those tuna or billfish species for which there is not enough information (e.g. about the stock-recruitment relationship) to reliably estimate $B_{msy.}$ IATTC also recommends an MSY-proxy (IATTC 2023b). Proposed interim TRP proxies are whichever is higher between 30% of the dynamic unfished SSB or the SSB_{msy} under current relative age specific fishing mortality when the spawner-recruitment relationship follows the Beverton-Holt function with an assumed steepness (h) of 0.75 (Maunder et al. 2023a).

In contrast to TRPs, LRPs define levels of biomass or fishing mortality that should be avoided with high probability to prevent recruitment overfishing and stock collapse. The WCPFC agreed that the risk of breaching the LRP should not be higher than 20% (WCPFC 2016), and its SC recommended using MSY-based reference points or their proxies described above as LRPs (SC 2011). For the IATTC, a risk of 10% of breaching the interim LRP has been adopted, where the LRP is the SSB that produces 50% of the virgin recruitment (R0) when the spawner-recruitment relationship follows the Beverton-Holt function with an assumed steepness of 0.75 (IATTC 2023b). The SSB at this LRP is equal to $7.7\%SSB_{F=0}$ and it is independent of the stock for which it is applied (IATTC 2023b). National standard guidelines in the U.S. suggest a default minimum stock size threshold of $(1\text{-}M)*B_{msy}$ or $0.5B_{msy}$, whichever is greater, where M is the natural mortality (Restrepo et al. 1998).

For this MSE, the JWG put forward five potential LRPs, which often also act as control points triggering management action as specified in the HCRs. Four candidate LRPs are defined as fractions of $SBB_{F=0}$: $20\%SSB_{F=0}$, $15\%SSB_{F=0}$, $10\%SSB_{F=0}$, and $7.7\%SSB_{F=0}$ (which is the IATTC interim LRP described above). The last candidate LRP is the median SSB from 1952-2014, which was estimated by the 2022 stock assessment model to be 40,725 t (ISC 2022) and corresponding to $6.3\%SSB_{F=0}$. This SSB level was also the WCPFC initial rebuilding target for PBF (WCPFC 2017).

The ThRP is solely used as a control point to trigger management action in conjunction with an HCR to bring the stock back to the target level faster, without breaching the LRP. Three candidate ThRP based on a fraction of $SSB_{F=0}$ have been proposed: $25\%SSB_{F=0}$, $20\%SSB_{F=0}$, and $15\%SSB_{F=0}$. While no reference points have been adopted for PBF, there is an interim harvest control rule where $20\%SSB_{F=0}$ is used as a ThRP that triggers management action when breached (WCPFC 2023b, IATTC 2023a).

3.6 Candidate Harvest Control Rules

Candidate HCRs to test in the MSE were put forward by the JWG and specify a management action depending on the status of the stock. The PBF MSE is model-based, meaning that the inputs to the HCR (i.e. current SSB and control points) are derived from a stock assessment, which is similar to the current PBF management system. The HCR specifies a F based on the condition of the PBF population as estimated by the stock assessment (i.e.

the estimation model, EM, in the MSE) relative to the HCR control points. The F is then translated into a Total Allowable Catch (TAC) for implementation. As agreed by the JWG, a three year management cycle is implemented whereby a stock assessment is conducted every three years and the TAC is kept constant between management periods. The MSE also simulates a realistic lag between the end of the assessment period and the implementation of the resulting TAC. For example, the current 2025 PBF catch limits are based on an assessment with data up to the end of fishing year 2022 (June 2023). Thus, the first EM in the MSE ends in the fishing year 2023 to set a TAC starting in the calendar year 2026 (Fig. 3A).

Figure ES2 depicts, for each of the eight HCRs under consideration, the management actions (i.e. changes in F) associated with a specific estimate of relative SSB (SSB as a fraction of $SSB_{F=0}$). Note that since HCRs 9 to 16 are the same as HCRs 1 to 8 except for the allocation of F between the EPO and WCPO, only HCRs 1 to 8 are depicted in ES2 and described below. Table ES2 lists the specific control points that trigger a change in management action when breached for each HCR. HCRs 1 to 4, and 8 have two control points (Fig. ES2, Table ES2). If relative SSB is at or above the first control point (ThRP), F is set to the F_{TARGET}. If relative SSB is below the ThRP, but above the second control point (here corresponding to the LRP), F is reduced proportionally to the SSB to levels below the F_{TARGET} to avoid breaching the LRP, and to bring the SSB back above the ThRP to target levels faster than if the F would have been kept constant at the F_{TARGET}. If SSB falls below the LRP, the F is maintained constant at a low level until SSB is rebuilt above the LRP. This minimum F (F_{min}) is a fraction of the F_{TARGET} except for HCR 4 which specifies a fishing mortality of FSPR70% (Table ES2).

By contrast, HCRs 6 and 7 have 1 control point, corresponding to the ThRP (Table ES2, Figure ES2). When the control point is breached, the F is reduced proportionally to the SSB to levels below the F_{TARGET}, eventually decreasing to 0 when the biomass is 0 (Fig. ES2). There is no additional management action when the LRP is breached, so these HCRs do not have an F_{min}.

HCR 5 also uses only 1 control point at the ThRP. When this is breached the HCR switches from using a constant F set to the F_{TARGET} to a constant catch set to the CMM2021-02 (WCPFC 2021) and C-21-05 (IATTC 2021) limits (Table ES2, Fig. ES2). Note that the resulting F relative to the F_{TARGET} once HCR 5 switches to the constant catch control is hard to define a priori as it will depend on the age structure of the population given the relative impact of different fleets under the catch limit, hence the illustrative dotted arrows in Fig. ES2.

The JWG also specified that HCRs be tested with a limit that constrains changes in TAC, between consecutive management periods, to no more than 25%, unless SSB falls below the LRP. The allocations were also tuned to reach the WCPO:EPO fishery impact ratio of 70:30 or the status quo ratio (about 80:20) (WCPFC NC 2023a). Thus, 16 total HCRs have been evaluated within this MSE, with the 8 HCRs in Fig. ES2 tuned to two different impact ratios.

3.7 Uncertainties Considered in MSE Process

PBF management advice has been based on stock assessments grounded in a single, best-case set of assumptions, with some uncertainties considered in the future projection (ISC, 2024a). For example, in projections since the 2016 stock assessment (ISC, 2016; 2018b, 2020, 2022, 2024a), observation and process (resampled from the recruitment estimates) uncertainties have been consistently incorporated, with one projection scenario further including parameter uncertainty. Due to the limited understanding of the true underlying system and challenges in effective implementation, uncertainties can dominate the errors in fishery management. Thus, despite limited uncertainties used in past PBF management advice, it is recommended that future long-term management advice be based on a fully specified set of rules (i.e., a management procedure) tested through MSE simulations across diverse uncertainty scenarios applied to PBF.

These uncertainties typically include errors in data and observation systems, model uncertainty, parameter uncertainty, process uncertainty, and implementation uncertainty, as outlined by Punt et al. (2016). Errors in data and observation systems arise from sampling and resources monitoring, particularly in collecting catch data, size compositions, and/or abundance indices (e.g., standard deviation for each observation in CPUE, effective sample size given to the compositions). In this MSE, we account for observation error by generating data with error from the OMs, which are mathematical representations of the possible true dynamics of the stock and fisheries (Fig. ES1). For more details on data generation see Section 4.2.1.

Model and parameter uncertainty stem from a lack of knowledge about the true population dynamics. The former relates to assumptions regarding the functional forms used to represent the biological process (e.g., whether the stock recruitment relationship follows Beverton-Holt or Ricker, or whether fishery selectivity is asymptotic or dome shaped, etc.). The latter relates to the uncertainty in fixed parameter values in the population dynamics models (e.g., steepness, M, growth parameters). To address these uncertainties, 20 different OMs, each representing an equally plausible "true" version of the system, were developed and used in the MSE (see Section 4.1.1.3 for more details).

Process uncertainty arises from seemingly unpredictable natural variability in population processes affecting abundance, such as future recruitment. The MSE accounts for this uncertainty by, for each OM, running 100 unique iterations differing in their recruitment deviations. The random recruitment deviations were sampled from a normal distribution with mean 0 and standard deviation of 0.6 in log space. The same deviation was assumed in the 2024 PBF stock assessment (ISC 2024a). The recruitment deviations are kept the same across HCRs for the same iteration to ensure HCRs encounter the same random recruitment trajectories.

Implementation uncertainty occurs when management actions are not perfectly executed. For example, actual catch levels may exceed TACs, or discarding and unreported landings may exist but remain undocumented. The MSE considers this uncertainty by adding an

implementation error to the TAC resulting from each of the candidate HCRs (see Section 4.2.4 for more details).

4. MSE Framework Description

4.1 Operating Models (OMs)

The OMs are computational representations of the real world, designed to simulate plausible population dynamics by conditioning on stock assessment models and incorporating key sources of uncertainty pertaining to the species of interest.

4.1.1 Conditioning process

The conditioning process is an important step in developing OMs, as it aligns the simulated system with the best available data. This includes fitting the OMs to historical data derived from the latest stock assessment to ensure that simulated population dynamics are consistent with observed trends in the data. Since OMs can differ from the current stock assessment due to divergent assumptions about biological parameters, the latest stock assessment model is used as a starting point for conditioning. During the conditioning, parameters related to productivity (e.g., steepness, natural mortality, growth parameters) are calibrated to the alternative assumptions and all other parameters (e.g., global scale, equilibrium fishing mortality, and selectivity) are estimated.

4.1.1.1 Data used for conditioning

The historical data used in the OMs are based on the 2024 stock assessment (ISC 2024a). Three main types of data were collected, compiled, and used in the assessment: catch and unaccounted mortality data, abundance indices, and size composition data.

First, catch statistics for fleets catching PBF were compiled quarterly for 26 fleets from 1983 to 2022 (Table 3A). The total annual catch fluctuated widely, with a historical maximum of 33,975 t in 2000 and a minimum of 8,585 t in 1990. Annual catches averaged about 14,000 t over the last decade (2013-2022). Most catches were attributed to purse seine fisheries. Unaccounted mortality includes fishery-induced deaths not reflected in landing data, such as discard mortalities. Estimates were provided for Japanese, Korean, and U.S. fisheries, while Mexico and Chinese Taipei reported no discards. Unaccounted mortalities for Fleet 24 (those from Japanese fisheries and Korean purse seine fisheries) were set to 5% of the WCPO total catches except Fleet 14 (Japanese troll for penning), unaccounted mortalities for Fleet 25 (those from Japanese fisheries for penning) were set to 100% of the total catches for Fleet 14, and unaccounted mortalities for the EPO recreational fleet were set to 1.2% of the total EPO recreational catches.

Second, abundance indices were derived from fishery-specific catch and effort data, standardized using statistical methods. Three longline CPUE series (two from Japanese fleets and one from Taiwanese fleets) served as adult abundance indices, while a Japanese troll index was used as a recruitment index for 1983-2010. Input coefficients of variation (CVs) were set at 0.2 for all indices and years, unless higher values were estimated. Logbook data from 1983-1992 (offshore/distant water) and 1993-2019 (coastal operations) were used for Japanese longline indices, with adjustments to exclude small-sized fish, in order to maintain consistency after 2017. The Japanese longline indices were discontinued after 2020 due to changes in management measures resulting in substantial declines in fishery activities during the main PBF fishing season. The Taiwanese longline index was derived from operations in their southern fishing ground (2002-2022). The Japanese troll index has been proven to be an informative indicator of recruitment since the 1980s, but it was discontinued after 2017 due to changes in management measures. Furthermore, this index post-2010 was identified as the cause of the negative retrospective pattern in the assessment. The substantial increase in catch for juvenile PBF farming after 2010, coupled with the implementation of mandatory licensing for troll vessels starting after 2010, may have compromised the representativeness of the troll index after 2010.

Third, quarterly size composition data (length or weight) were compiled for 1983-2022. Size composition data for certain fleets were excluded due to concerns about sampling quality or representativeness. In summary, size data from Japanese longline, Taiwanese longline, Japanese purse seine, Korean purse seine, Mexico purse seine, US purse seine, US recreational, Japanese setnet, troll, and pole-and line fisheries were used. Input sample sizes for the size composition data were sourced from various criteria for each fleet. Depending on the corresponding fisheries and available data, the input sample size includes “Number of fish measured”, “Number of landing wells sampled”, “Number of the total month of wells sampled by port”, and “Number of haul wells sampled”.

The only difference in data between the 2024 stock assessment and the 20 OMs is that OMs converted the catch in number for some fleets to catch in weight to streamline the MSE feedback process at each simulated assessment time step.

4.1.1.2 Base Case Operating Model

The base case OM (OM1) was derived from the assumptions and structure of the 2024 stock assessment model, but with a simplified selectivity configuration (see details below).

4.1.1.2.1 Biological and Demographic Assumptions

4.1.1.2.1.1 Sex specificity

The base case OM assumes no sexual dimorphism because previous studies showed that the sex ratios of females to males were not statistically different from 1:1 (Chen et al. 2006, Shimose and Takeuchi 2012). Although males typically attain larger sizes than females after reaching sexual maturity (Maguire and Hurlbut 1984, Shimose et al. 2009, Shimose and Takeuchi 2012), available age-length data by sex were unbalanced and sample sizes were relatively small. Furthermore, the lack of sex information in fishery data makes the use of sexual dimorphism in the base case OM impractical, leading to the assumption of a single-sex population.

4.1.1.2.1.2 Growth

A time-invariant sex-combined length-at-age relationship was externally estimated from over 2,000 paired age-length otolith samples (Fukuda et al. 2015, Ishihara et al. 2023) and used in the base case OM. This relationship was re-parameterized to fit the von Bertalanffy growth equation using Stock Synthesis and adjusted for the assumed birth date (1st of July, i.e., the first day of the fishing year),

$$L_2 = L_{inf} + (L_1 - L_{inf})e^{-K(A_2 - A_1)}$$

where L_1 and L_2 are the lengths (cm) associated with ages (years) near the first (A_1) and second (A_2) ages, L_{inf} is the asymptotic average length-at-age (Francis 1988), and K is the growth coefficient (y^{-1}). The growth parameters K, L_1, and L_2 were fixed, with K at 0.188 y^{-1}, and L_1 and L_2 at 19.05 cm and 118.57 cm for age 0 and age 3, respectively, based on the length-at-age relationship by Fukuda et al. (2015). L_{inf} was re-parameterized as:

$$L_{inf} = L_1 + \frac{L_2 - L_1}{1 - e^{-K(A_2 - A_1)}}$$

L_{inf} was then calculated as 249.917 cm. The process errors, modeled as the CVs, were functions of the mean length at age, $CV = f(length_{age})$. Based on the estimated variances from the conditional age-at-length data (Tsukahara et al. 2024), the CVs were fixed at 0.278 and 0.0401 for ages 0 and 3, respectively. Linear interpolation between 0-3 was used to generate the process error for intervening ages, and ages 3 and older were assumed to be the same as age 3.

4.1.1.2.1.3 Age modeled

The maximum observed age in the paired age-length samples were around 28 years old, but sample sizes substantially decrease after 20 years old (Fukuda et al. 2015). The mean length-at-age for fish older than 20 years old remained stable. Therefore, age 20 was treated as an accumulator for all older ages (dynamics are simplified in the

accumulator age), where approximately 0.15% of an unfished cohort remained, based on the natural mortality schedule. Ages from 0 to 20 were modeled.

4.1.1.2.1.4 Weight at length

A sex-combined weight-length relationship was used to convert fork length (L) in cm to weight (W_L) in kg (Kai 2007). The relationship is:

$$W_L = 1.7117 * 10^{-5} * L^{3.0382}$$

where W_L is the weight at length L. This weight-length relationship was assumed to be time-invariant and fixed.

4.1.1.2.1.5 Natural Mortality

Natural mortality (M) was assumed to be time-invariant and age-specific and was fixed in the base case OM. Age-specific M estimates for PBF were derived from a meta-analysis of different estimators based on empirical and life history methods to represent juvenile and adult fish (Aires-da-Silva et al. 2008). The M of age 0 fish was estimated from a tagging study, as discussed in detail previously. Age-specific estimates of M were fixed in the model: 1.6 y^{-1} for age 0, 0.386 y^{-1} for age 1, and 0.25y^{-1} for age 2 and older fish.

4.1.1.2.1.6 Recruitment and reproduction

PBF spawn throughout spring and summer (April-August) in various areas of the western Pacific Ocean, as inferred from egg and larvae collections and examination of female gonads. In the base case OM model, spawning was assumed to commence at the beginning of April (fishing month 10). Based on Tanaka (2006), age-specific estimates of the proportion of mature fish were fixed: 0.2 at age 3, 0.5 at age 4, and 1.0 at age 5 and older fish as of April 1st. PBF ages 0-2 fish were assumed to be immature. Recruitment is assumed to occur in fishing month 1.

A standard Beverton and Holt stock-recruitment relationship (SR) was used in this assessment. The expected recruitment for year (R_y) is a function of spawning biomass (SSB_{y-1}), an estimated unfished equilibrium spawning biomass (SSB_0), a specified steepness parameter (h), and an estimated unfished recruitment (R_0).

$$R_y = \frac{4hR_0SSB_{y-1}}{SSB_0(1-h) + SSB_{y-1}(5h-1)}\ e^{-0.5b_y\sigma_R^2+\hat{R}_y}$$

$$\hat{R}_y \sim Normal(0, \sigma_R^2)$$

Annual recruitment deviations from the SR relationship ($\hat{R}_y$) were estimated from 1982 to 2021 and assumed to follow a normal distribution with a specified standard deviation (σ_R) in natural log space (Methot and Taylor 2011, Methot and Wetzel 2013). This σ_R penalizes recruitment deviations from the spawner-recruitment curve. The central tendency, penalizing the log (recruitment) deviations for deviating from zero, was assumed to sum to zero over the estimated period. Estimation of recruitment is

known to be difficult in the penalized likelihood estimation framework (Maunder and Deriso 2003), so a tuning σ_R approach was used to match the standard deviation of the estimated recruitment deviations. Several repeated model runs were conducted to numerically estimate a value of σ_R based on Methot and Taylor (2011), resulting in a σ_R set to be 0.6 in the model, which was about the variability of deviates if it was estimated by the model. This σ_R value allows the model to be less sensitive to our assumptions about the steepness.

A log-bias adjustment pattern fraction (b) was applied to the 1982-2019 recruitment estimates to assure unbiased estimation of mean recruitment. Because this b was internally calculated in the model, a two-step procedure was used to apply the estimation of b based on Methot and Taylor (2011). The first model run estimated recruitment deviations and variability around these values without adjusting the bias accurately. The b was also calculated in the first model run based on the estimated recruitment deviations and σ_R, which was 0.9336. The base case OM applied this b value obtained from the first run. The closer b is to the max value of 1, the more informative the data are about recruitment deviations, and vice versa, because b is in log space.

The steepness of the stock-recruitment relationship (h) was defined as the fraction of recruitment when the spawning stock biomass is 20% of $SSB_{F=0}$, relative to R_0. Previous studies have indicated that h tends to be poorly estimated due to the lack of information in the data about this parameter (Magnusson and Hilborn 2007, Conn et al. 2010, Lee et al. 2012). Lee et al. (2012) concluded that steepness could be estimable within the stock assessment models when models were correctly specified for relatively low productivity stocks with good contrast in spawning stock biomass. However, the internally estimated h may be imprecise and biased for PBF as it is a highly productive species. Independent estimates of steepness that incorporated biological and ecological characteristics of the species (Iwata 2012, Iwata et al. 2012b) reported that the mean of h was around 0.999, close to the asymptotic value of 1.0. Therefore, steepness was fixed at 0.999 in the base case OM. It was noted that these estimates were highly uncertain due to the lack of information on PBF's early life history stages.

4.1.1.2.2 Model structure

The model assumed a single well-mixed stock for PBF, which is supported by tagging and genetic studies (Section 3.1.1).

4.1.1.2.2.1 Initial conditions

When populations are exploited prior to the onset of data collection, models must make assumptions about what occurred before the start of the dynamic period. These models often make an assumption of the population being in an equilibrium state in this pre-dynamic period, which can result in a population in the initial year being either at an unfished equilibrium, in equilibrium with an estimated mortality rate influenced by

data on historical equilibrium catch, or exhibiting estimable age-specific deviations from equilibrium. Two approaches describe extreme alternatives for dealing with the influence of equilibrium assumptions on the estimated dynamics.

The first approach is to start the dynamic model as far back in time as necessary to assume that there was no fishing prior to the dynamic period. Usually, this entails creating a series of hypothetical catches that extend backward in time and diminish in magnitude with temporal distance from the present. The other approach is to estimate (where possible) parameters defining initial conditions.

Because of the significance (in both time and magnitude) of the historical catch prior to 1983, the base case OM used the second method (estimate) to develop non-equilibrium initial conditions that estimated: 1) R_1 offset, 2) initial fishing mortality rates, and 3) early recruitment deviations. The R_1 offset was estimated to reflect the initial equilibrium recruitment relative to R_0, which had been estimated in the models. The equilibrium fishing mortality rate (Feq) was estimated because the initial equilibrium involved not only natural mortality but also fishing mortality. The estimation of Feq can be based on the equilibrium catch, which is the catch taken from a stock for which removals and natural mortality are balanced by stable recruitment and growth. Although the model did not fit equilibrium catch (having no influence on the total likelihood function for deviating from assumed equilibrium catch), Feq was freely estimated for the Japanese set-net fleet for seasons 1-3 because it represented a fleet that mainly took small fish.

Nine-year recruitment deviations prior to the start of the dynamic period were estimated to adjust the equilibrium initial age composition before starting the dynamic to be a non-equilibrium initial age composition. The model first applied the R_1 offset and initial Feq level to an equilibrium age composition to obtain the preliminary numbers-at-age. Then it applied the recruitment deviations for the specified number of younger ages (information came from the size compositions for early years in the assessment) to these numbers-at-age. Since the number of estimated ages in the initial age composition is less than the maximum age, the older ages retained their equilibrium levels. Because the older ages in the initial age compositions will have less information, the bias adjustment was set to be zero.

4.1.1.2.2.2 Selectivity

Selectivity for the fishery fleets

Selectivity is the observation model process that links composition data to underlying population dynamics. For non-spatial models, this process combines the contact selectivity of the gear and population availability to the gear. The former is defined as the probability that the gear catches a fish of a given size/age, and the latter is the probability that a fish of a given size/age is spatially available to the gear. In the case of PBF, variable trans-Pacific movement rates of juvenile fish cause temporal

variability in the availability component of selectivity for those fisheries catching migratory juveniles. Therefore, in addition to estimating length-based gear selectivity, time-varying age-based selectivity was estimated to approximate the time-varying age-based movement rate. The use of time-varying selection results in better fits to the composition data compared to the time-invariant selection model, which had adverse consequences on fits to other prioritized data (ISC 2014, ISC 2016).

We also used a combination of model processes (time-varying length- and age-based selectivity) and data weightings to ensure goodness of fit to size composition for the fleets that caught high numbers of fish (ISC 2024a). In general, fleets with large catches of migratory ages, good quality of size composition data, and no CPUE index were modeled with time-varying selection (Lee et al. 2015). Fleets taking mostly age-0 fish or adults were treated as time-invariant. Fleets with small catches or poor size composition data were either aggregated with similar fleets or given low weights. Details are given below.

Fishery-specific selectivity was estimated by fitting length and weight composition data for each fleet except for fleets whose selectivity patterns were assumed to be the same as other fleets based on the similarity of the size of fish caught. The weight and length composition data for some fleets were not used to estimate its selectivity due to poor quality of sampling, limited observations, or/and unclear sampling scheme. The size composition data for the discard fleets were not available, but it was assumed that their selectivity pattern was similar to that of the retained catch.

Fleets with an associated CPUE index were modeled as time-invariant length-based selection patterns to account for the gear selectivity. Due to the nature of their size compositions (non-migratory ages caught by these fleets, either age-0 fish or spawners, resulting in a single well-behaved mode), functional forms of logistic or double normal curves were used for these fleets. The choice of asymptotic (logistic curves) or dome-shaped (double normal curves) selection pattern was based on the assumption that at least one of these fleets sampled from the entire population above a specific size (asymptotic selectivity pattern), which helps to stabilize population estimation. This assumption was evaluated in a previous study, which indicated that the Taiwanese longline fleet from the south fishing ground consistently produced the best fitting model, when an asymptotic selection was used for that fleet (Piner 2012). This assumption along with the observed sizes and life history parameters set an upper bound to population size.

Fleets without an associated CPUE index were categorized into fleets taking fish of non-migratory ages (mainly age-0 fish or spawners) and fleets taking fish of migratory ages (mainly ages 1-6). Selectivity for non-CPUE fleets taking fish of non-migratory ages was modeled as time-invariant length-based selection patterns to account for the gear contact and time-invariant age-based availability patterns to account for the additional ages available to the fleets. Due to the nature of their size compositions with

a single well-behaved mode, functional forms of double normal curves were estimated. As for non-CPUE fleets taking fish of migratory ages, both length- and age-based selectivity patterns were estimated (Lee et al. 2015). Selection is then a product of the age- and length-based selection patterns. In general, the pattern for the length-based selection was time-invariant asymptotic or dome-shaped, while the age-based selection estimated separate parameters for each age and was time-varying for migratory ages. Because of the large number of parameters involved, fleets without significant catch did not include the time-varying age-based component. Additionally, three EPO fleets were modeled with time-varying length-based selection due to the possible difference in growth between EPO and WPO. Detailed selectivity specification for each fleet can be found in the 2024 stock assessment report (ISC 2024a). While the OMs, like the 2024 stock assessment, had time varying selectivity for some fleets as described above, in the forward MSE simulation, selectivity was kept constant to 2015-2022 average values, which was agreed to by the JWG.

Selectivity for the abundance index

Selectivity for each relative abundance index was assumed to be time-invariant and the same as the fishery from which each respective index was derived.

4.1.1.2.2.3 Catchability

Catchability (q) was estimated assuming that each index of abundance was proportional to the vulnerable biomass/numbers with a scaling factor of q, which was assumed to be constant over time. Vulnerable biomass/numbers depended on the fleet-specific selection pattern and underlying population numbers-at-age.

4.1.1.2.2.4 Data Observation Models

The statistical model estimates the best-fit model parameters by minimizing a negative log-likelihood value that consists of likelihoods for data and prior information components. The likelihood components consisted of catch, CPUE indices, size compositions, and a recruitment penalty. The observed total catch data assumed a lognormal error distribution. An unacceptably poor fit to catch was defined as models that did not remove >99% of the total observed catch from any fishery. Fishery CPUE and recruitment deviations were fit assuming a lognormal error structure. Size composition data assumed a multinomial error structure.

4.1.1.2.2.5 Data Weighting

Three types of weighting were applied in the base case OM: (1) length compositions, (2) catch data, and (3) CPUE abundance data. Catch data were weighted with S.E.=0.1 (log space) for all fleets, which is relatively precise for catches, except for unaccounted mortality fleets (S.E.=0.3). CPUE observations were generally weighted

with CV=0.2 unless the standardization model indicated higher uncertainty. In that case, the larger CV estimated from the standardization was used. The weights given to fleet-specific quarterly composition data via effective sample size were based on an ad-hoc method. Generally, sample sizes were low (<15 effective sample sizes) based on the number of well-measured samplings from the number of hauls or daily/monthly landings except for the longline fleets. For longline fleets, because only the numbers of fish measured are available (the numbers of trips or landings measured were not available), the sample size was scaled relative to the average sample size and standard deviation of the sample size of all other fisheries based on the number of fish sampled.

4.1.1.3 Model Structure of alternative Operating Models in reference set

The ISC PBFWG identified productivity parameters as the most influential and uncertain factors among the examined uncertainties, which include model uncertainty and errors in data and observation systems. These productivity parameters include length at age 3 (L_2), natural mortality for age 2 and older (M_2^+), and the steepness of the stock-recruitment relationship (h). The uncertainty grid encompassed combinations of values for length at age 3 (L_2), steepness (h), and natural mortality rate for ages 2 and older (M_2^+). This uncertainty grid was used to develop a series of candidate models for the reference OM grid. Diagnostics of these candidate models were then used to select the final OM grid. The process used to finalize the OM grid is summarized below, but for more details refer to Lee et al. (2023), Lee and Tommasi (2023), and Lee and Tommasi (2024).

4.1.1.3.1 Determining the range of productivity parameters

A suite of empirical estimators for M_2^+ was used to investigate the range of M for mature fish. These estimators were based on the maximum age, the von Bertalanffy growth function, and the age at maturity. In a recent review by Maunder et al. (2023b), the methods for estimating M were examined, and they recommended focusing on the maximum observed age (*tmax*) as it provides a more direct relationship with M. Among the estimators based on *tmax*, the formula M=5.4/*tmax* was suggested by Then et al. in 2015, Hamel and Cope in 2022, and Maunder et al. (2023b) (equation T3.2.1). This equation is derived from models that assess the probability of a fish surviving to a particular age under a specific level of total mortality. Based on historical age data (Fukuda et al. 2015), the maximum observed age is 28, corresponding to an M value of 0.193 year^{-1}. The base case assessment assumed an M_2^+ value of 0.25 year^{-1} (ISC 2024a), where *tmax* is about 22 years old according to the formula. We considered these values as bounding the potential uncertainty in M_2^+ and explored here alternative model structures with an M_2^+ specified at 0.193 or 0.25 year^{-1}.

The length-at-age data from otoliths, collected by Japanese and Taiwanese scientists between 1992 and 2014 (Fukuda et al. 2015), were used to explore the range of length-at-age 3 in the first quarter (L_2 = 3.0 years old). Ishihara et al. (2023) further bootstrapped these length-at-age data using different sampling methods and data points, revealing that the median of estimated length-at-age 3 ranged from 118.57 to 118.82. The 95% confidence interval for the estimated L_2 was within ±2 cm from the median. In the base case stock assessment model, L_2 was specified at 118.57 cm fork length, with the CV of L_2 at 4.4%. After synthesizing the model-based estimates and bootstrapped analyses, we consequently selected a range of potential L_2 as spanning from 118 to 119 cm. In the following analyses we tested alternative models with L_2 values of 118, 118.57, or 119 cm.

There is less information available to guide the choice of a range for parameter h than for M_2^+ or L_2 due to the lack of early life history data. Independent estimates of steepness that incorporate biological and ecological characteristics of the stock (Iwata 2012; Iwata et al. 2012b) reported that the mean of h was around 0.999. We explored a broad range of h values, ranging from 0.8 to 1.

4.1.1.3.2 Diagnostics of candidate models in the reference OM grid

The candidate models within the OM grid associated with the uncertainty in identified productivity parameters and their plausible values were evaluated. Fishery data were integrated into each model in the OM grid, and a suite of diagnostic tools including jitter analyses, goodness-of-fit assessments, likelihood profiling on R_0, retrospective analyses, and ASPM-R was applied to identify and eliminate underperforming models. The PBF WG considered that models with poor diagnostics indicated poorly conditioned models, which were likely unrepresentative of the real world, and thus the data generated by these models would be unsuitable for use in the MSE.

4.1.1.3.2.1 Convergence and stability

To evaluate convergence towards a global minimum, we conducted 25 jitter analyses for each candidate model within the OM grid. This process involved randomly perturbing the initial values of all parameters by 10% and subsequently re-running the model. The primary objective of these jittering analyses was to ensure that none of the randomly generated starting values of parameters led to a solution with a lower total negative log-likelihood (NLLs) compared to the reference model. The final reference model must have the lowest total NLL and a positive-definite Hessian matrix. These analyses served as a quality control procedure to confirm that the model was not converging towards a local minimum.

When M_2^+ was 0.25, the percentage of jitter runs resulting in a positive-definite Hessian matrix generally increased with higher steepness values, regardless of L_2 (Lee and Tommasi 2024). However, when M_2^+ was 0.193, the percentage of jitter runs resulting in a positive-definite Hessian matrix was low when steepness values were between 0.95 and 0.97. Any combination of parameter values within the OM grid with 0% of runs resulting in a positive-definite Hessian matrix was not considered in the subsequent diagnostics and was given a score of zero for this diagnostic.

4.1.1.3.2.2 Goodness-of-fit

We used total NLLs to guide our assessment of the goodness-of-fit for both data components (abundance indices and size composition). We utilized the NLL values from the 2024 stock assessment as the basis to determine how well each candidate model in the OM grid fit each data component. A statistically significant worse fit relative to the base model (Section 4.1.1.2) was defined as an increase in NLLs exceeding 1.92 units.

The NLL values for the index data components suggest that most candidate models in the OM grid performed similarly to or better than the base model (Lee and Tommasi 2024). The NLLs for the size compositions indicate that more models in the OM grid achieved a fit similar to or better than the base model as L_2 decreased. For all data compositions, the NLLs diagnostic concluded that more models in the OM grid achieved a fit similar to or better than the base model as L_2 decreased. The index and size composition components provided inconsistent grid profiles; therefore, goodness-of-fit was not considered in the final OM selection process.

4.1.1.3.2.3 R_0 likelihood profile

The R_0 likelihood profile served as a tool for assessing which data sources provided information on the global population scale parameter and for pinpointing regions where conflicts arose among these sources (Lee et al. 2014). The profile involved running a series of models, where the $\ln(R_0)$ parameter was fixed (not estimated) at a range of values both above and below the estimate derived within the model. This process quantifies the extent of loss of fit for each data component resulting from changing the population scale. Data components rich in information on population scale will exhibit substantial degradation in fit when the population scale deviates from the best estimate.

Following the completion of all profile runs, the degradation in fit was computed by subtracting the overall and component's minimum NLL (or best fit) across all profile runs from the overall and component's NLL from each specific profile run, respectively. We calculated the 95% confidence interval for the changes in NLL around R_0^{MLE} (R_0 at

the minimal total likelihood estimates), corresponding to half of the chi-squared values for p=0.95 with 1 degree of freedom. Ultimately, if R_0^c for the data component at the minimal likelihood estimates falls outside the 95% confidence interval for R_0^{MLE}, it indicates a conflict with the overall model. Conversely, if R_0^c for the data component at the minimal likelihood estimates falls inside the 95% confidence interval for R_0^{MLE}, the data component aligns with the overall model population scale. This entire process was iterated for each candidate model in the OM grid.

The R_0 profile results indicated that only the size components provided consistent estimates of the global population scale ($\ln(R_0)$) for the base grid, with R_0 at the minimal likelihood estimates for the size data component falling within the 95% confidence interval for R_0^{MLE} (Lee and Tommasi 2024). This consistency, as seen in the base model, was also observed in most of the other models in the OM grid. Any candidate model in the OM grid lacking the same consistency as in the base model was given a score of zero for this diagnostic.

4.1.1.3.2.4 Retrospective analyses

A retrospective analysis was used to examine consistency of model output once recent data were systematically removed from each of the candidate models in the OM grid. The underlying assumption is that estimates of historical abundance using all data are more accurate than estimates from retrospective models that ignore recent data. Therefore, this analysis reveals potential biases within model estimates. A 7-year retrospective analysis was conducted across all models in the OM grid by sequentially removing one year of data from the end of the time series. Subsequently, the Mohn's rho statistic (Hurtado-Ferro et al. 2014) was calculated to quantify the severity of retrospective patterns. A greater absolute Mohn's rho indicates a consistently obvious pattern of change in the retrospective models.

The Mohn's ρ value for spawning stock biomass from the base grid was 0.01 (Lee and Tommasi 2024). Other grids exhibited similar or smaller Mohn's ρ values compared to the base grid. When M_2^+ was 0.25, the retrospective pattern increased as h decreased, accompanied by a larger absolute Mohn's ρ. However, when M_2^+ was 0.193, the retrospective pattern decreased as h decreased with a smaller absolute Mohn's ρ. Any candidate model in the OM grid with an absolute Mohn's ρ value larger than 0.1 was given a score of zero for this diagnostic.

4.1.1.3.2.5 Age-structured production model with recruitment (ASPM-R)

The age-structured production model diagnostic (ASPM; Maunder and Piner 2015) served as a diagnostic tool to evaluate the current state of the production function and to identify potential misspecifications in the system dynamics (Carvalho et al. 2017).

To account for cohort growth, we modified the ASPM, introducing the ASPM-R model, which allows for recruitment deviations to be specified at previously estimated values in addition to selectivities.

Initially, each candidate model in the OM grid was fitted to catch, size compositions, and abundance indices (adult and recruitment indices) as in the assessment model, but with alternative productivity assumptions. Subsequently, an ASPM-R model was developed and ran, incorporating recruitment deviations and selectivities specified at the estimates from each model in the OM grid. Therefore, each candidate model in the OM grid had an associated ASPM-R version of that model. The ASPM-R models estimated scaling parameters ($\ln(R_0)$ and R_1) and the initial fishing mortality rates, fitting to catch and adult abundance indices.

Differences in model fits (i.e., NLLs) between the ASPM-R base model and each candidate model in the OM grid was used to determine whether the two models performed similarly. If the difference was more than 1.92 units, the PBF WG considered that specific candidate model in the OM grid to have significantly degraded performance.

The NLLs from the ASPM-R models generally deteriorated when h was smaller than the base value, regardless of M_2^+ or L_2 values (Lee and Tommasi 2024). The selected range of h expanded when either M_2^+ or L_2 was larger. In the case of M_2^+=0.25, the selected h values ranged from 0.99 to 0.999 when L_2 was 118.57, while the selected h values expanded from 0.97 to 0.999 when L_2 was 119. Any candidate model in the OM grid displaying a statistically significant degradation in the NLLs of its associated ASPM-R model, thus indicating a poorly estimated production relationship, was given a score of zero for this diagnostic.

4.1.1.3.2.6 Ensemble diagnostic results

The selections are based on the sum of the scores from convergence, R_0 profile, retrospective, and ASPM-R analyses for each candidate model in the reference OM grid (Lee and Tommasi 2024). The scores range from 0 to 4, with the highest score indicating successful passage of all four diagnostics. The scores reveal conflicting information across retrospective analyses, R_0 profile, and ASPM-R. Specifically, ASPM-R favored higher values for M_2^+ and h, while R_0 profile leaned towards lower values for h. In summary, only grids that passed three or more diagnostics were selected in the reference set of the OM, and those were equally weighted to evaluate the MPs in the MSE process. Their productivity parameters are outlined in Table ES3 and the uncertainty range in relative SSB for the selected OMs is shown in Figure ES3.

4.1.1.4 Robustness set of Operating Models

Three robustness scenarios were run to test HCR behavior under extreme conditions detrimental to stock productivity. These scenarios are less likely than the reference set, but are still considered plausible. They should not be given the same weight as the reference set, and thus results are presented separately. These robustness scenarios, which are described in more detail below, were: 1) about a 40% 10-year long drop in recruitment; 2) an effort creep for the Taiwanese longline fleet on which the main index of abundance is based; and 3) a doubling of discards. For all the robustness scenarios, OMs were constructed by modifying OM1, which has the same settings as the 2024 base-case assessment model. As in the reference set, for each robustness OM, 100 iterations were run to account for process error.

4.1.1.4.1 Recruitment drop

The recruitment drop simulations had a later end year (2066) and hence lasted longer than all runs in the reference set, which ended in 2045. The drop in recruitment started in 2042, once the median SSB for each HCR had increased to target levels and stabilized from the lower initial levels in 2023. To simulate the drop in recruitment, in years 2042 to 2051 of the forward simulation, recruitment deviations were sampled at random with replacement from a set of 11 predefined low recruitment deviations from the historical period. These low recruitment deviations were taken from years 1980 to 1988 and 1991 and 1992 of the 2022 PBF assessment (ISC 2022). Although recruitment years classes in 1980-1989 (recruitment deviations in 1979-1988) were defined as the low recruitment period used to select recruitment for projections in the 2016 PBF assessment (ISC 2016), the PBF WG did not include the year 1989 recruitment deviation in the low recruitment resampling period for this robustness test because the 1989 recruitment deviation was positive. In addition, the recruitment deviations for 1991 and 1992 were included in the low recruitment resampling period because those were the lowest estimated recruitment deviations. The low recruitment period in this simulation was 10 years long because the historical period of low recruitment lasted approximately 10 years (ISC 2016). In the other (i.e., non-low recruitment) years of the simulation, recruitment deviations were sampled in the same way as the reference set, from a normal distribution with mean 0 and standard deviation of 0.6 in log space. This method resulted in about a 40% drop in recruitment for the period of 2042 to 2051 for all the iterations in this robustness test (Fig. 4A). Note that in the reference set, performance metrics were calculated over the 20-year period from 2026 (when the first simulated TAC from the candidate HCRs is applied) to 2045. Although the performance metrics in this recruitment drop robustness test were calculated also from a 20-year period, the time period was instead from 2047 through 2066 to cover the period of low SSB resulting from the drop in recruitment.

4.1.1.4.2 Catchability Change

A fishery-dependent CPUE index is a common data source for stock assessments, and it is critically important in cases when a survey abundance index is not available, as is the case of the PBF assessment. One of the issues with using a fishery-dependent CPUE index is the possible change in catchability (q) over time due to the technological

and operational changes in fisheries (i.e. effort creep; Palomares and Pauly 2019, Kleiven et al. 2022). Although there has been an internal consistency among the adult CPUE indices, recruitment index, and catch time series in the PBF stock assessment, simulations with an unaccounted catchability change (2% increase/year) were conducted to test the robustness of the candidate HCRs to this potential effort creep.

OM 1 was tuned (conditioned) with the Taiwanese longline CPUE index (TLL index). In this robustness scenario, this CPUE index was assigned a hypothetical catchability increase of 2% per year (downward adjustment of index values) since 2002 under OM1 (Fukuda 2024). To simulate an "overlooked effort creep", the MSE simulation loop was modified to adjust the TLL index downward by 2% per year since 2002 in the OM, while the q in the EM remained constant (overestimating an increasing trend of CPUE after 2010).

4.1.1.4.3 Increase in Discards

The PBF stock assessment has taken into account possible unobserved mortality related to the possible post-release mortality or unreported catch as a part of the removals from the stock. This unobserved mortality was expressed as three fleets: 1) fleet 24 which is the unaccounted mortality in the WPO in the biomass unit (weight); 2) fleet 25, which is the unaccounted mortality in the WPO in number of fish; and 3) fleet 26, which is the unaccounted mortality in the EPO in number of fish (Table 3A; ISC 2024a).

Although the current assessment performed well in terms of the consistency between the total removals and the trend of the stock, the unobserved mortality could become more influential in the future because of its "unknown" nature (Nishikawa et al. 2024). Thus, the WG decided on using a robustness scenario that assumed an unaccounted mortality that was twice that assumed in the assessment and described in Section 4.2.4 (ISC 2024b). This assumption does not require changes in the conditioning of the OM but was modeled as an additional unknown removal during the MSE "future" simulation process. While the OM experienced a doubling of discards, the EM was run with the same assumption as the current assumption in the stock assessment. This robustness test was run using OM1.

4.2 "Future" Process

Once all the selected OMs were "conditioned" and the PBF WG deemed that the OMs could reasonably reconstruct past patterns in PBF observations, they were projected forward in time in a closed loop simulation (Fig. ES1). In the closed loop simulation, there was feedback between the PBF population and the management actions. For each candidate HCR, each of the 20 OMs was projected forward in time from 2023 to 2045, for 100 different iterations to account for process uncertainty in recruitment (Section 3.7). At each time step of the 23-year simulation, each OM simulated the "true" population dynamics of PBF stock and

the fisheries operating on it, given the TAC set by a candidate HCR. Catch, CPUE, and size composition data with error were sampled from the OM every three years (based on a 3-year stock assessment frequency, Section 4.2.1) and, as in the real world, input with error into a stock assessment model (i.e., the EM, Section 4.2.2) (Fig. ES1). This ensures management procedure performance is evaluated given realistic errors in observation and estimation.

Collection of data and stock status estimation are all components of the management procedure, which also includes the HCR and any specifications about how catches may be allocated to different fleet segments (Fig. ES1). In the PBF MSE, the only component of the management procedure that changes among the different candidate options is the HCR. Once data with error was input into the EM (i.e., the simulated stock assessment model), as in the real world, the EM estimated the "current" population levels and fishing mortality as well as reference points. Estimates of stock status and reference points were then input into a management module, which consisted of a HCR with specific reference points (Table ES2). The management module (Section 4.2.3) sets the fleet-segment specific TACs and these catches were then input into the OM with some implementation error (Section 4.2.4). The population dynamics of the stock were then simulated for three time-steps (i.e., three years), using the OM with the fleet-segment specific TACs and implementation errors, until the next simulated stock assessment with the EM, which closes the simulation loop. This simulation loop continued until the end of the 23-year simulation. The different components of the forward closed loop simulations are described in more detail in the sections below.

The first simulated assessment in the MSE ran in the fishing year 2023 (July 2023 to June 2024 in calendar year) and set the TAC for calendar year 2026. Thus, to run the OM, catches prior to 2026 needed to be specified as the first TAC fed back into the OM occurs in 2026. Member countries provided catch data for July to December 2023 for the MSE. Catches for 2024 and 2025 were based on the agreed upon catch limits (IATTC 2021, IATTC 2024, WCPFC 2023a, WCPFC 2024). These catch limits are by country and, for WCPO countries, also size category (smaller or larger than 30 kg). However, fleets in the MSE OM are at a finer level than country and size category to account for variable selectivity at age by season and region of operation. Thus, the country and size category catch limit were split between the fleets for a specific country and size category using the same catch ratios as in the last year of the 2024 stock assessment. Note that catch limits for the US and Mexico were combined as, like the stock assessment, the OM has a combined EPO commercial fleet. Since recreational catches in the EPO were not part of the catch limit, EPO recreational catches for 2024 and 2025 were set to be the same as in the last year of the 2024 stock assessment. Discards are also not included in the catch limits, so discards for Fleet 24 (discards from Japanese fisheries and Korean purse seine fisheries) were set to 5% of the WCPO total catches except Fleet 14 (Japanese troll for penning), discards for Fleet 25 (discards from Japanese fisheries for penning) were set to 100% of the total catches for Fleet 14, and discards for the EPO recreational fleet were set to 1.2% of the total EPO recreational catches.

The code to run the PBF MSE framework is available at a github web site[3]. It is written in R, and, since the OM and EM are based on the Stock Synthesis (SS3) software, it allows SS3 to be run directly from R. The repository contains all the directories, files, and functions needed to run the MSE framework for OM1 (or scenario 1), management procedure 1, and all the harvest control rules.

The PBF MSE directory structure has the following format:

PBF_MSE/management procedure/hcr/OM/iteration/time step

Before running the main "*PBF_MSE_prll_all_hcrs_sam24_ncmm.R*" code, the repository needs to be cloned from GitHub to ensure that all the required directories and files are on the user's computer. In this code, the user specifies the management procedure and scenario (OM) being run. It also calls one of the main PBF_MSE functions, which are dependent on the HCR being run. HCRs 1 to 4 and 8 use "*PBF_MSE_hs1_for_sam24_ncmm.R*", while HCR 5 uses "*PBF_MSE_hs1_hcr8_for_sam24_ncmm.R*", and HCRs 6 and 7 use "*PBF_MSE_hs2_910_for_sam24_ncmm.R*". Note that while the LRP and ThRP are set in the call to these functions, the F_{TARGET} needs to be specified in the SS3 forecast file in the HCR folder as the algorithm requires SS3 to determine in the benchmark section the F multiplier associated with the F_{TARGET}. Thus, there is an SS3 forecast file under each HCR directory. The scenario refers to the OM being used. The code uses the scenario to select which OM files in the Condition folder to use for the MSE forward simulation. The iteration reflects which random recruitment deviation time series is being run and the time step refers to the estimation time step (i.e., when an assessment would be run). The MSE simulation is run for 23 years and an assessment is run every three years. In OM1, the generation time for PBF is ~9 years, so the simulation time horizon corresponds to about 3 PBF generations. Note that the code is quite computationally expensive to run. To reduce storage and run times, it should first be tested with perfect estimation by setting sa=0 in the functions described above.

4.2.1 Data Generation

Catch, CPUE, and size composition data is generated using the SS3 bootstrap data generation routine (Methot and Wetzel 2013). First, the new catch data, given the TAC and implementation error, is added to the OM data files and dummy data is put in for the CPUE index and the size composition data. The code also automatically specifies the starter file to generate a parametric bootstrap file. The data generation routine then creates a new data set of random observations, which includes observation errors using the same variance properties (standard error of fleet specific catch, standard error of the CPUE indices, and effective sample size of the size composition data) and error structure (lognormal for catch and CPUE, multinomial for the size composition data) assumed during the conditioning phase. The new data with observation error are then input into the EM.

[3] https://github.com/detommas/PBF_MSE

For the forward simulation, catch data was assigned a CV of 0.1. The index of abundance was from the Taiwanese Longline and assigned a CV of 0.2 as for the conditioning period. The effective sample size for the size composition data was set to the average of the conditioning period as specified in Table 4A . The MSE data generation workflow also implements the bootstrap bias corrections outlined in Lee et al. (2021).

4.2.2 Estimation Model

The EM used in the PBF MSE process is an age-structured production model with estimated recruitment deviates (ASPM-R+). This ASPM-R+ is based on an implementation in Stock Synthesis (SS3; Methot and Wetzel 2013) which has been widely used as a diagnostic tool to evaluate model fitting in stock assessments (Maunder and Piner 2015). The use of ASPM-R+ as an EM can substantially reduce computation time in the MSE process (approximately -58% timesaving) without degrading estimation performance, compared to using the full SS3 model as an EM (Takahashi et al. 2024). In the ASPM-R+, all the selectivity parameters are fixed at those estimated from the full SS3 model (OM1, i.e., similar to the base case of the 2024 stock assessment model) except for fleets 1 (JPN_LL) and 3 (TWN_LLSouth) and only the size composition data for those fleets are included. Recruitment deviates added to the ASPM-R+ are also estimated from the full SS3 OM1.

Thus, the ASPM-R+ EM has basically the same modeling structure as the base-case OM (i.e., OM1). In MSE simulation, unlike the OM, which is run with an assumed, pre-specified set of parameters, the EM, mimicking a stock assessment in the real world, estimates parameters given the (simulated) data. Discrepancies between OM and EM output are therefore driven by observation and estimation errors. Estimates of terminal year SSB, terminal year numbers at age, and unfished SSB from the EM are input into the HCR. The EM is also used to compute the F multiplier that will achieve, given the biology, selectivity, and relative fishing mortality between fleets, the SPR-based F_{TARGET} specified by the HCR (See Section 4.2.3).

4.2.3 Management Module

The management module uses outputs from the EM and a candidate HCR to derive a TAC by fleet segment (EPO, WCPO small fish, WCPO large fish). The first step in this process is to use the EM to find the multiplier of the current fishing mortality that would achieve the F_{TARGET} set by the HCR given a specified relative fishing mortality across fleets, selectivity and biology. This is done using the SS3 forecast algorithm. More specifically, the EM has an associated forecast file, where the F_{TARGET} of the HCR under consideration is specified. The forecast file also specifies how the fishing mortality is allocated between fleets. This is defined as the relative apical fishing mortality for each fleet and season ($RelF_{t,f,s}$ where t = a year or multi- year period, f =

fleet and *s* = season). The $RelF_{t,f,s} = F'_{t,f,s}/(\text{sum of } F'_{t,f,s} \text{ for all fleets and seasons})$ where F' is the apical (i.e. fully selected) fishing mortality for a specific season, fleet, and year or averaged over a multi-year period. The forecast file also specifies over which period to average the age selectivity of each fleet and the biological parameters. The JWG agreed to set the selectivity and the relative F used in calculation of the F multiplier to the average of 2015-2022 for HCRs 1 to 8 (JWG 2024). Since the EM has constant biological parameters, the period used for biology does not influence the outcome. The algorithm finds the multiplier of the current fishing mortality that would achieve the F_{TARGET} ($Fmult_{TARGET}$) given the specified relative F, selectivity and biology. The management module then uses the HCR to assess if the $Fmult_{TARGET}$ should be changed depending on stock status, i.e., the level of terminal year SSB relative to the control points (Table 4B). The F multiplier, age selectivity per fleet, and the relative fishing mortality across fleets are then used to find the fishing mortality at age by fleet and season that would meet the fishing mortality set by the HCR. To do so, the F multiplier is first multiplied by the $RelF_{f,s}$ specified in the forecast file to find the apical F for each fleet and season ($F'_{f,s}$).

$$F'_{f,s} = Fmult * RelF_{f,s}$$

These apical Fs are then multiplied by the age selectivity per fleet and season ($Sel_{a,f,s}$) to obtain the fishing mortality at age per fleet and season ($F_{a,f,s}$).

$$F_{a,f,s} = F'_{f,s} * Sel_{a,f,s}$$

To then obtain a TAC from the fishing mortality set by the HCR, terminal year numbers at age, natural mortality, and weight at age from the EM are used. The starting numbers at age for season 1 ($N_{a,1}$) are calculated from the EM terminal year numbers at age in season 1 ($Nty_{a,1}$) and the terminal year total mortality at age (Zty_a, natural plus the fishing mortality) according to:

For ages 1 to A-1 where A is the oldest age, $N_{a,1} = Nty_{a-1,1} * \exp(-Zty_{a-1})$

For age A, $N_{A,1} = Nty_{A-1,1} * \exp(-Zty_{A-1}) + Nty_{A,1} * \exp(-Zty_A)$

For age 0, $N_{0,1}$ = Recruits from terminal year season 4

The catch at age by fleet at age for season 1 ($C_{a,f,1}$), is calculated as:

$$C_{a,f,1} = F_{a,f,1}/Z_a * N_{a,1} * (1-\exp(-0.25*Z_a))$$

Where $F_{a,f,1}$ is the fishing mortality at age for season 1 that would meet that fishing mortality set by the HCR calculated above. This is then multiplied by the weight at age and summed across ages to obtain the TAC by fleet for season 1. Note that we multiply Z_a by 0.25 as this is the season duration since there are four seasons in a year in the PBF OM. The numbers at age for seasons 2 to 4 ($N_{a,s}$) are calculated from the mortality and numbers at age in the previous seasons:

$$N_{a,s} = N_{a,s-1} * \exp(-0.25 * Z_a)$$

In each season, the catch at age and TAC is calculated as shown above for season 1. The total TAC for the EPO is found by summing the TAC for all the EPO fleets, including recreational fleets, for all seasons. The WCPO small fish TAC is found by summing the TAC for the WCPO small fish fleets, and ages 0 to 2 of the WCPO mixed fleets, across all seasons. The WCPO large fish TAC is found by summing the TAC for the WCPO large fish fleets, and for ages 3 and older of the WCPO mixed fleets, across all seasons. This corresponds to the TAC set by the fishing mortality specified in the HCR. However, unless biomass is below the LRP, for each HCR, there is a 25% limit on the change in TAC between management periods. Thus, when biomass is above the LRP, for each fleet segment, if the HCR-derived TAC is smaller or greater than the previous TAC by more than 25%, the TAC is simply set to a 25% increase or decrease from the previous TAC and the HCR-derived TAC is not used.

4.2.4 Implementation Error

In an MSE, the implementation error accounts for errors in reporting, problems with compliance, operational implementation of the management measure, unforeseen changes in fisher behavior, or discards that lead to the actual catch being different from the TAC. For PBF, the recent, more stringent management measures and growing PBF population, have increased the potential for release of PBF of an undesirable size (ISC 2024a). In the PBF MSE, therefore, implementation error reflects the presence of discards, with the catch fed into the OM being more than the TAC specified by the HCR due to discards. The extra discard catch is implemented by using three additional discard fleets and the number of discards was the same as what was assumed in the latest stock assessment. Fleet 24 (discards from Japanese fisheries and Korean purse seine fisheries) discards were set to 5% of the WCPO total catch except Fleet 14 (Japanese troll for penning), discards for Fleet 25 (discards from Japanese fisheries for penning) were set to 100% of the Fleet 14 catch, and discards for the EPO recreational fleet were set to 1.2% of the total EPO recreational catches.

4.2.5 Fishery Impact Ratio Tuning

The JWG specified that HCRs be evaluated in the MSE with allocations tuned to reach a WCPO:EPO fishery impact ratio of 70:30 or 80:20 (WCPFC NC 2023a). A fishery impact ratio or proportional fishery impact is the fishery impact of a particular group (e.g. EPO) relative to the impact of all the fisheries combined and has become a quantity routinely computed and presented to managers in the PBF stock assessment (ISC 2024a). Its value depends on the relative exploitation pattern across fleets (i.e. allocation) and the age selectivity of each fleet. Use of the current (2015-2022) allocation baseline (relative F) and age selectivity in the EM and management module algorithm (Section 4.2.3) leads to an EPO:WCPO impact ratio close to 80:20 in OM 1 for HCRs 1 to 8 (Fig. 4B). Therefore, there was no need to tune the relative F in the EM to achieve the desired 80:20 fishery impact ratio.

To reach the 70:30 fishery impact ratio, the ISC PBF WG developed a method to determine by how much the relative F of the EPO in the EM and management module algorithm (Section 4.2.3) should be increased over the 2015-2022 baseline to meet a pre-determined impact ratio between the EPO and WCPO (see Tommasi and Lee 2024). While Tommasi and Lee (2024) ran the method with no estimation error, in this MSE, the method was used with the EM to find the final tuning factor. Briefly, 100 MSE simulations were run for HCR 1 under OM 1 with progressively different EPO/WCPO relative Fs specified in the EM forecast file to assess the effect of changes in relative F on the proportional EPO/WCPO fishery impact metric. Only one HCR was used because Tommasi et al. (2023a) showed that the fishery impact ratio was relatively consistent across different HCRs as it was dependent on the specified relative F. We also only used OM 1, rather than finding a tuning factor for each OM, to consider uncertainty in the estimation of the tuning factor. OM 1 is based on the current assessment and is the most similar OM to the EM, and we assumed that other versions of reality (i.e. OMs) would not be known during operational application of the management procedure.

For each of the different relative F levels, 100 runs were carried out, and the EPO proportional fishery impact was computed and plotted against the relative F level (Fig. 4C). In R version 4.1.3, we fit a polynomial regression with a quadratic term to the relative F and proportional fishery impact output, and used the estimated relationship to compute what increase in EPO relative F compared to the 2015-2022 baseline would be needed to be to obtain a WCPO:EPO proportional fishery impact ratio of 70:30. We found that the relative F of the EPO fleets needed to be increased by 6.5% and that of the WCPO fleets decreased accordingly. The forecast file with this new relative F was then used in the EM during the MSE simulation to evaluate the HCRs with a WCPO:EPO fishery impact ratio of 70:30.

5. Results

The simulations resulted in a total of 32,000 runs (20 OMs x 16 HCRs x 100 iterations) for the reference set. Each run had 8 simulated stock assessments (i.e., EMs), for a total of 256,000 EMs. When reviewing the output, it was noticed that 224 EMs produced unrealistically low (<1e-60) terminal SSB estimates not seen in the OM (compare Figs 5A and 5B). In addition, 20 EMs had absolute relative errors in SSB, calculated as $100*|(SSB_{OM}-SSB_{EM})|/SSB_{OM}$, greater than 1000%. These unrealistic values pointed to estimation issues in the EMs. These EMs were further investigated and all of the 244 EMs had extremely large gradients (> than 100), indicating non-convergence. The low SSB values were associated with extremely high estimated Fs, and later low Fs as the low SSB triggered HCRs to implement drastic management actions (Fig. 5C). The management action then resulted in high SSB values later in the simulation (Figs. 5A and 5B).

When performing assessments in the real world, an initial assessment may fail to converge but assessment scientists will almost certainly diagnose and fix the issue by making small adjustments to the initial values of the parameters. This cannot occur in a simulation with limited resources. Since the unrealistic SSB values were caused by model non-convergence, and such models would not have been used in the real world to make management decisions, it was decided to remove the problematic EMs. While these estimation issues only rarely occurred (less than 1% of the EMs in some years, OMs, HCRs, and iterations), the PBF WG decided to remove the 19 iterations associated with the problematic EMs for all OMs and HCRs. This was done to ensure performance metrics were computed for HCRs exposed to the same recruitment trends. Removing the iterations associated with the problematic EMs resulted in the removal of the unrealistically low estimated SSB values as well as the high SSB later in the simulation (compare Figs. 5A and 5D). Nevertheless, the performance metrics and the ranking of HCRs for each performance metric were not greatly affected (compare Table ES4 to Table 5F).

All results below are therefore presented only for the 81 iterations without any associations with the problematic EMs. Performance metrics for the reference set are calculated across 20 OMs and 81 iterations and 20 evaluation years (2026 to 2045 as the first simulated TAC was applied in 2026). The same iterations were also used to evaluate results from the robustness scenarios. While not discussed in the text, Table 5F shows the performance metrics computed using all the 100 iterations, which were highly similar to the performance metrics with 81 iterations.

Before looking at results for each performance metrics, it is instructive to look at relative SSB and F time series plots from the EM and OM to check if a management action was triggered as expected and what the impact was on the population and fisheries. Worm plots (e.g. Fig. 5D) display each of the time series that were simulated for each iteration and scenario. Each line represents a possible simulated reality that may occur. Areas where more lines overlap represent more likely realties than others. In viewing these plots note that in the MSE, the F and TAC resulting from the HCRs depend on what the EM (i.e. the simulated stock assessment) estimates the stock status to be, which is different from the true stock status simulated by the OM because the EM has estimation error. We are interested in assessing if HCR performance still meets objectives despite this estimation error.

Initial relative SSB in all iterations and scenarios is estimated by the EM to be above the LRP of all HCRs except for HCRs 3 and 11, which have the highest LRP at $20\%SSB_{F=0}$ (Fig. 5D). Relative SSB is also estimated for most scenarios and iterations to be above the ThRP of all HCRs except 2, 3, 10, and 11, which have the highest ThRP of $25\%SSB_{F=0}$ (Fig. 5D). Thus, for most HCRs, we would not expect drastic management actions to be triggered for calendar year 2026,

when the first TAC is applied. The first management action is applied in calendar year 2026 given conditions in the initial fishing year 2023 because of the lag between estimation of stock status and management action. Indeed, we see a decline in TAC for most worm plots lines around 2026 for HCRs 3 and HCR 11. For these two HCRs, some iterations and scenarios even showed a drastic management action (steep drop in TAC) occurring during this first management period because the LRP was breached (Fig. ES13). This early decline in TAC for both HCR 3 and 11 is also reflected in the median and 5^{th}-95^{th} quantile plot of TAC, with median TAC declining already in 2026 (Fig. ES14). For most other HCRs, median TAC remains relatively similar or declines only slightly compared to 2025 levels in the medium term (Fig. ES14). However, in the long term, all HCRs show a higher median TAC as compared to 2025 levels (Fig. ES14) due to the increase in estimated biomass under HCR management (Fig. 5D). We note that HCRs 1, 2, 3, 9, 10, and 11, which have the highest LRPs, have more iterations with drastic drops in TAC because the LRP had been breached (Fig. ES13). In these instances, once the TAC falls drastically due to the LRP being breached, it can only increase slowly from these low levels due to the 25% limit on TAC change between periods (Fig. ES13), even if biomass recovers above the LRP and ThRP, which happens rapidly under the low F.

The resulting TAC is a combination of the F set by the HCR as well as the estimated biomass. Initially, TAC patterns follow trends in the estimated F from the EM. As for the TAC, only HCRs 3 and 11 show F worm plots with large declines in 2026, and drastic management actions for some iterations and scenarios occurring during the first management period (Fig. 5E). Also similar to the patterns in TAC, HCRs 1, 2, 3, 9, 10, and 11, which have the highest LRPs, have more iterations with drastic drops in F because the LRP had been breached (Fig. 5E). Like median TAC, median estimated F declines in the medium term relative to the 2025 level for all HCRs. However, unlike the higher long-term median TACs, median estimated F remains constant in the long-term for all HCRs (Fig. 5F). While for most HCRs, estimated F in the long-term remains constant around the $F_{TARGET,}$ HCRs 1, 2, 3, 9, 10, and 11 have a lower F than target, due to more management intervention and the 25% limit on TAC change between periods, which limits potential increases in TAC and makes management more conservative (Fig. 5F).

Looking at trends in estimated relative SSB and F from the EM helps us understand patterns in the resulting TAC, but performance metrics are based on output from the OM. This is because we want to assess impacts of the management measures on the simulated "true" population. Actual F from the OM is lower than estimated F. Indeed, unlike in the EM, median F from the OM is already declining in 2026 for most HCRs, not just HCRs 3 and 11, albeit it declines faster for these two (Fig. 5G). Furthermore, median long term F remains below the F_{TARGET} for all HCRs in the OM (Fig. 5G). This is due to estimation error whereby the EM estimates F to be higher than it actually is (compare Figures 5Fand 5G). In terms of patterns in relative SSB from the OM, there is a lag between reductions in F and resulting increases in relative SSB, with OM median relative SSB only starting to increase around 2027 (Fig. ES4). In the long term, for all HCRs, OM median relative SSB increases above target levels because long term median F is below target (Fig. 5G and 5H). There is also estimation error in the relative SSB, with the EM initially overestimating biomass (compare Fig. 5D and 5H) because the EM has a parametrization similar to the most productive OM, OM 1. For example, while the EM estimates initial relative SSB to be above the LRP for all iterations and scenarios except those of HCRs 3 and 11 (Fig.5D), relative SSB from

the OM in some iterations is actually also lower than the LRP of HCRs 1, 2, 9, and 10, and at the LRP of HCRs 4, 8, 12, and 16 (Fig. 5H). Despite this estimation error, all HCRs are able to rebuild relative SSB above the LRP for all iterations and scenarios in the long term (Fig. 5H). The next sections will outline results for each performance metric under management objective category, based on OM output.

5.1 Safety

Performance of the different HCRs with respect to the safety management objective, *there should be a less than 20% probability of the stock falling below the LRP*, was measured by computing the probability of SSB<LRP in any given year of the evaluation period (Table ES1). Three performance metrics were calculated for this management objective, each with a different LRP. First, we used the LRP that the JWG specified for each HCR and that in some cases was used as a control point (Table ES2). Second, we used the second rebuilding target of $20\%SSB_{F=0}$ as a consistent LRP for all HCRs, notwithstanding what lower biomass control point each HCR used. Third, we used the IATTC's interim LRP of $7.7\%SSB_{F=0}$ as another consistent LRP across HCRs. The two latter calculations were not requested by the JWG but were nonetheless provided to ease comparison.

All HCRs are able to maintain the probability of relative SSB falling below their own LRP to less than 20% (Fig. ES5). HCRs 1, 2, 3, 9, 10, and 11, which have the highest LRPs at 15% or $20\%SSB_{F=0}$ (Table ES2), perform poorer. This is due to their performance at the start of the simulation. Initial SSB is the same across HCRs, and is already lower than the LRPs for these HCRs for some iterations and scenarios (Fig. 5H), leading to higher probability that SSB<LRP in the initial years of the simulation (Fig. 5I). However, management action lowers this probability quickly and by year 2029 the probability of relative SSB being lower than $15\%SSB_{F=0}$ is already less than 15% for HCRs 1, 2, 9, and 10 and by 2030 also the probability of relative SSB being lower than $20\%SSB_{F=0}$ is less than 20% for HCRs 3 and 11 (Fig. 5I). This safety metric can also be interpreted as the probability of drastic management action, as when SSB<LRP the TAC can change more than 25%.

Changes in performance between HCRs for the first safety performance metric are dependent on both the values of SSB as well as the LRP. However, for the second and third safety performance metrics, differences between HCRs are only dependent on SSB levels because consistent reference levels of $20\%SSB_{F=0}$ or $7.7\%SSB_{F=0}$ are used. Violin plots of relative SSB show that HCRs 3 and 11, which have the lowest F_{TARGET} at FSPR40%, have the highest relative SSB, and HCRs 6 and 14 with the highest F_{TARGET} at FSPR20% have the lowest relative SSB (Fig. 5J). This is expected as an F that leaves more SPR also leaves more biomass. The violin plot also shows that more of the distribution of HCRs 6 and 14 is below $20\%SSB_{F=0}$ (Fig. 5J), thus we expect these HCRs to have lower performance with respect to the second safety performance metric. Indeed, performance of the second metric is related to the F_{TARGET} levels. HCRs with lower F_{TARGET} do better and HCRs with a higher F_{TARGET} do poorer (Fig. ES7). Among HCRs with the same F_{TARGET} of FSPR30%, HCRs 2 and 10

perform slightly better (Fig. ES7). These HCRs have a higher ThRP (Table ES2) and the associated earlier reduction in F is associated with a slightly higher median relative SSB (Fig. 5J). While HCRs with higher F_{TARGET} have a lower probability of SSB being below $20\%SSB_{F=0}$, the probability is less than 20% for all HCRs except HCRs 6 and 14 (Fig. ES7). Poorer performance of HCRs 6 and 14 is due to the slower biomass increases under these HCRs with less conservative control points, which leads to a higher probability of SSB being below $20\%SSB_{F=0}$, as compared to HCRs with more conservative control points (Fig. 5K). Patterns for the third safety performance metric, the probability of $SSB<7.7\%SSB_{F=0}$ are similar to the second, with HCRs with a higher F_{TARGET} performing worst (Fig. ES6). However, all HCRs have a very low probability (<10%) of SSB being below $7.7\%SSB_{F=0}$ (Fig. ES6).

The PBF WG has no specific recommendation for an LRP with which to test safety performance, especially given that the PBF stock has recovered from a very low level of SSB (2% of $SSB_{F=0}$).

5.2 Status

Performance of the different HCRs with respect to the status management objective, *to maintain fishing mortality at or below F_{TARGET} with at least 50% probability*, was measured by two performance indicators. The first computed the probability that $F \leq F_{TARGET}$ in any given year of the evaluation period. The second calculated the probability that SSB is below the equivalent biomass depletion levels associated with the candidates for F_{TARGET} (Table ES1).

Under ideal management, with no estimation or implementation error, yearly stock assessments, and no lag between status estimation and application of management action, we would expect F in the long run to stabilize around the F_{TARGET}, and thus the probability of $F \leq F_{TARGET}$ to be around 50%. We saw however, that, due to estimation error and the 25% limit on TAC increases between management periods, all HCRs tend to be more conservative and result in a median F below the F_{TARGET} in the long term (Fig. 5G). This is also evident in the violin plot of F (Fig. 5L). HCRs 1, 2, 4, 8, 9, 10, 12, and 16 have a median SPR of around 40% instead of 30%, while HCRs 3 and 11 of 60% rather than 40%, HCRs 5, 7, 13, and 15 of 35% instead of 25%, and HCRs 6 and 14 of 30% instead of 20%. Thus, we would expect the first status performance metric to be met by all HCRs. Indeed, the probability that $F \leq F_{TARGET}$ was at least 50% for all HCRs (Fig. ES17).

The violin plot also shows that HCRs 3 and 11, and to a lesser extent, HCRs 1, 2, 9, and 10, have fatter lower tails (Fig. 5L). This is due to their higher LRP, and thus higher number of iterations requiring drastic management action (i.e. steep decrease in F as LRP has been breached) (Fig. 5E). Once F falls to these low levels, it is slow to increase due to the 25% limit in TAC changes between management periods, leading to fatter lower tails in the violin plots (Fig. 5L). These HCRs, given their fatter lower tails, have then the highest probability of $F \leq F_{TARGET}$ (Fig. ES17). HCRs 6 and 14 also show a more variable F distribution (longer

lower tails, Fig. 5L) as the SSB associated with the F_{TARGET} of F20%SPR is close to the ThRP of 20%$SSB_{F=0}$, requiring frequent management action.

Median SSB levels also tend to be above target (Fig. 5M). Median relative SSB for HCRs 3 and 11 is just above 40%$SSB_{F=0}$, while SSB for HCRs 1, 2, 4, 8, 9, 10, 12, and 16 is at about 35%$SSB_{F=0}$, SSB for HCRs 5, 7, 13, and 15 is at 30%$SSB_{F=0}$, and SSB for HCRs 6 and 14 is around 27%$SSB_{F=0}$. Thus, the probability of SSB being below the equivalent biomass depletion levels associated with the candidates for F_{TARGET} is less than 50% for all HCRs (Fig. 5N).

5.3 Stability

Performance of the different HCRs with respect to the stability management objective, *to limit changes in overall catch limits between management periods to no more than 25%, unless the ISC has assessed that the stock is below the LRP*, was measured by the percent change upwards or downwards in catches between management periods excluding periods when SSB<LRP (Table ES1). Since the JWG specified that each HCR be assessed with a limit of 25% on changes in TAC between management periods, unless the LRP has been breached, we would not expect the TAC to vary by more than 25% when SSB is ≥ LRP. Indeed, the maximum percent change upward or downward in TAC was 25% (Fig. ES15 and 5O). Within this range, the most conservative HCRs 3 and 11 have the highest stability in terms of decreases in TAC. These HCRs have the largest difference between the biomass associated with their F_{TARGET} and their ThRP ($SSBF_{TARGET}$ - ThRP=15%$SSB_{F=0}$). By contrast, HCRs 2, 5, 6, 10, 13 and 14, which have a smaller difference between the target and ThRP ($SSBF_{TARGET}$ - ThRP= 0% or 5% of $SSB_{F=0}$), have low stability (Fig. ES15). When stability is calculated including years when SSB<LRP, downward changes in TAC can at times be greater than 50% due to drastic management action (Fig. 5P). HCRs 2, 6, 10, and 14 still retain the lowest median downward changes in TAC (Table 5A). However, HCRs 3 and 11 are no longer the best performing (Fig. 5P, Table 5A). This is a result of more drastic management intervention for these HCRs at the start of the simulation, as shown by their higher probability of being below their LRP at the start of the simulation (Fig. 5I).

With regards to increases in TAC, HCRs 3, 11, 2 and 10 have the highest changes (Fig. 5O). These HCRs have the highest ThRP of 25%$SSB_{F=0}$, which implies earlier drastic management action, a faster rebuilding back to the target, and an associated larger increase in SSB, leading to a larger increase in TAC relative to other HCRs. A violin plot of the whole distribution of changes in TAC demonstrates that median change in TAC is positive (Fig. 5Q), highlighting that all HCRs have more increases than decreases in TAC.

5.4 Yield

Unlike the other categories, yield has multiple management objectives (Table ES1). The first, *maintain an equitable balance in proportional fishery impact between the WCPO and*

EPO was measured by computing the median fishery impact on SSB (in %) in the terminal year of the evaluation period by fishery and by WCPO fisheries and EPO fisheries. Fishery impact is dependent on both the total catch of the fleet and the size composition of that catch as fish of different sizes and hence maturities contribute differently to SSB. Therefore, this performance metric is affected by both the selectivity and relative F (i.e. allocation of fishing mortality) of each fleet. For HCRs 1-8, the relative F and selectivity were set to the 2015-2022 average, as agreed by the JWG, while HCRs 9-16 used a relative F that would lead to a 70:30 impact ratio (see Section 4.2.5 for details). It is important to note that the impact ratios are not expected to be exactly at 80:20 or 70:30 after uncertainty in productivity is considered and the impact ratio computed across all reference set OMs. This is because the impact of using the 2015-2022 baseline leads to a median WCPO:EPO impact ratio of about 80:20 under OM 1 for HCRs 1 to 8, and the 70:30 impact ratio runs were tuned using the EM under OM 1 for HCRs 9 to 16. Median fishery impact was similar across HCRs with the same relative F (Table ES4, Fig. 5R). Median EPO fishery impact ranged from 22 to 24 for HCRs 1 to 8, while it was between 30 and 33 for HCRs 9 to 16 (Table ES4, Fig. 5R).

The second and third yield management objectives, *to maximize yield over the medium (5-10 years) and long (10-30 years) terms, as well as average annual yield from the fishery* and *to increase average annual catch in all fisheries across WCPO and EPO*, were measured by the same performance metrics. These were the median annual catch in any given year of the simulation period, over the medium term (years 5-10 of the simulation period), and over the long term (years 10-23 of the simulation period). The metrics were computed across all fisheries as well as for the WCPO large fish fleet segment, the WCPO small fish fleet segment, and the EPO fleet segment. HCRs with a higher F_{TARGET} showed a higher yield performance (Fig. ES9). Catch was highest for HCRs 6 and 14 as they had the highest F_{TARGET} of 20%SPR, while HCRs 3 and 11 had the lowest as they had the lowest F_{TARGET} of 40%SPR (Fig. ES9). HCRs 3 and 11, and to a lesser extent 1, 2, 9, and 10 had a fatter lower tail in their catch distribution (Fig. ES9), similar to the fishing mortality violin plot (Fig. 5L). This was due to their higher LRPs, higher number of iterations requiring drastic management action, and the 25% limit in TAC changes between management periods, which slowed increases to higher TAC levels once catch falls to these low levels. Annual catch also had a longer tail for HCRs 6 and 14 (Fig. ES9) as the common ThRP of these HCRs is the same as the biomass associated with the F_{TARGET}, leading to frequent management intervention. Of the HCRs with a F_{TARGET} of FSPR30%, HCRs 4, 8, 12, and 16 have lower variability in annual catch than HCRs 1, 2, 9, and 10 as a result of their lower control points, but have comparable median catch (Fig. ES9). Of the HCRs with a F_{TARGET} of FSPR25%, HCRs 5 and 13 have a shorter lower tail because when the ThRP is breached, the TAC is set to the catch limits defined in CMM2021-02 (WCPFC 2021) and C-21-05 (IATTC 2021) and so cannot fall as low as for HCRs 7 and 15. All HCRs had a median annual catch higher than the current catch limits in the long term. However, median annual catch in the medium term or across all years was only higher than the 2024 catch limits for HCRs 3 and 11 (Fig. ES9).

As expected, there were no large differences in total catch between HCRs with a 80:20 or 70:30 WCPO:EPO impact ratio. However, HCRs 9 to 16, with higher EPO impacts, had lower TACs for WCPO large and small fish, and higher EPO TACs than HCRs 1 to 8 (Figs. ES10, ES11, ES12).

Within the same impact ratio, the TACs for specific fleet segments followed the same pattern as total catch, being highest for those HCRs with the highest F_{TARGET} (Figs. ES10, ES11, ES12). As for total catch, all HCRs had a higher median annual WCPO large fish TAC in the long term than the current catch limits (Fig. ES10). Annual WCPO large fish TAC was higher than current catch limits for all HCRs except 3 and 11 in the medium term, and except for HCR 11 across all years (Fig. ES10). By contrast, median annual WCPO small fish TAC was less than the current catch limits for all HCRs (Fig. ES11). Median annual TAC for the EPO was also less than the current catch limits for all HCRs with a 2015-2022 baseline EPO impact ratio, but was higher for all HCRs when the EPO impact was increased to 30%, except for HCR 11 in the medium term (Fig. ES12).

5.5 Tradeoffs between Performance Metrics

There were no best performing HCRs for all management objectives as tradeoffs between performance metrics were evident (Table ES4). HCRs with a lower F_{TARGET} performed best in terms of the second safety performance metric, but at the cost of reduced yield (Fig. ES8). HCRs with F_{TARGET}s of FSPR25% or FSPR30% were able to maintain the probability of SSB being above $20\%SSB_{F=0}$ to more than 80% while maintaining higher catches than HCRs with a F40%SPR F_{TARGET} (Fig. ES8). While HCRs with a FSPR40% F_{TARGET} performed best in terms of stability and poorest in terms of yield, there was no clear tradeoff between yield and stability due to the high variability in performance among HCRs with F_{TARGET}s of FSPR25% or FSPR30% (Fig. ES16 and 5S). For the same F_{TARGET}, HCRs with a lower ThRP generally did better in terms of stability because of lower management intervention, but maintained similar performance in terms of yield and safety (Fig. ES16 and 5S). All HCRs performed well in terms of status, with the probability of $F \leq F_{TARGET}$ being always greater than 50%. However, HCR 11, which had the lowest yield, had the highest probability of $F \leq F_{TARGET}$ (Fig. 5T). Some of the HCRs performing best in terms of yield, performed worst in terms of status. However, there was high variability in the status performance among HCRs with F_{TARGET}s of FSPR25% or FSPR30% and hence no clear tradeoff between status and yield (Fig. 5T). There was also no tradeoff between status, stability, and safety performance metrics (Fig. 5U and 5V). HCRs with a FSPR40% F_{TARGET} performed well in terms of both status, stability, and safety (Fig. 5U and 5V).

5.6 Potential First TAC

In the PBF MSE, an age-structured production model with estimated recruitment deviates (ASPM-R+) was used as an EM (see section 4.2.2). The stock synthesis software,

which has been used in the stock assessment of PBF, can alter its model structure from the fully integrated model to a simple age structured production model, and the ASPM-R+ in this MSE was developed by modifying the 2024 PBF stock assessment base-case (Takahashi et al., 2024).

Here we calculate, outside of the MSE framework, what the potential annual TAC for calendar years 2026 to 2028 would be for each of the candidates MP, if they were to be adopted. To do so, the 2024 stock assessment model, which terminated at the 2022 fishing year, was updated for a year up to the 2023 fishing year. Quarterly catch by fleet and the Taiwanese longline CPUE abundance index were updated, and all parameters were re-estimated. Then, the updated model was diagnosed to check for critical model misspecification, such as falling into a local minimum or an obvious misfit to the abundance index. After confirming that the updated model passed the diagnostics, the updated model was modified to the ASPM-R+ (i.e., the EM). Figure 5W shows the estimated $SSB/SSB_{F=0}$ by the ASPM-R+. The terminal relative SSB was estimated to be $28.8\%SSB_{F=0}$.

This EM was used to find the age structure (numbers at age) and $SSB/SSB_{F=0}$ at the terminal year (2023), as well as the multiplier of the current fishing mortality that would achieve the F_{TARGET} given the assumed selectivity (F2015-2022). As was done when testing the MPs in the MSE, the relative fishing intensity across the fleets was set to the 2015-2022 base-line for HCRs 1-8, while it was tuned to achieve the 70:30 fishery impact ratio between the WCPO and EPO fisheries groups for HCRs 9-16 (see section 4.2.5). Once the EM calculated those parameters, the F multipliers were further adjusted to control the fishing intensity if required by each HCR based on stock status (see section 4.2.3). The potential annual TAC for 2026-2028 was obtained from the fishing mortality set by the HCR, terminal age structure (numbers at age), natural mortality, and weight at age assumed.

The WCPO small fish TAC was found by summing the calculated catch for the WCPO small fish fleets and the calculated catch for ages 2 and younger of the WCPO mixed fleets, across all seasons. The WCPO large fish TAC was found by summing the calculated catch for the WCPO large fish fleets, and the calculated catch for ages 3 and older of the WCPO mixed fleets, across all seasons. The EPO TAC was found by summing the calculated catch for the EPO commercial fisheries and EPO recreational fisheries, across all seasons. This corresponds to the TAC set by the fishing mortality specified in the HCR. However, for each fleet group (i.e., WCPO small fish, WCPO large fish, and EPO) if the calculated TAC was higher or lower than the TAC in 2025 by more than 25%, instead of using the HCR TAC, the 2025 TAC was simply increased (or decreased) by 25%[4]. If the biomass is below the LRPs, the 25% limitation would not be applied, but it was not the case at this time. The estimated potential TACs are shown in Table ES4. Note that catches by the unaccounted mortality fleets (e.g. fleets 24-26) are not included in the TAC calculated above.

[4] Bycatch allowance described in paragraph 4 of WCPFC CMM2024-01, notably 200t by NZ and 40t by Australia, is not considered part of TAC in this MSE thus not reflected in the calculation of 2026 TAC

5.7 Robustness scenarios

All HCRs were fairly robust to a doubling of discards not detected in the EM. Median SSB was lower under higher discards (Fig. 5Z1) and safety performance was poorer (Table 5B). However, all HCRs maintained a less than 20% probability of being below their own LRP or below the IATTC's interim LRP of 7.7%$SSB_{F=0}$ (Fig. 5Z2). Similar to the results from the reference set, the probability of breaching 20%$SSB_{F=0}$ was higher than 20% for HCRs 6 and 14 under the discard scenario (Fig. 5Z2). Due to the higher discards, median F was higher for all HCRs (Fig. 5Z3). Nonetheless, all HCRs maintained a probability of F being at or below F_{TARGET} higher than 50% (Fig.5Z4). Furthermore, all HCRs except 3 and 11 had a less than 50% probability of SSB being below SSB levels associated with the F_{TARGET} (Fig. 5Z4). For HCRs 3 and 11, the probability of SSB being below SSB levels associated with the F_{TARGET} was 54% and 53%, respectively. The ranking of HCRs in terms of stability performance were also similar to that of the reference set (compare Tables 5B and 5C), with HCRs 2, 5, 6, 10, 13, and 14 performing worst in terms of downward changes in TAC (Fig. 5Z5). In terms of yield, patterns were again similar to the reference set, with total catches increasing only slightly relative to the reference set, even with the increase in F, due to the decrease in SSB (Fig. 5Z6 and Tables 5B and 5C).

In the effort creep for the Taiwan longline fleet robustness scenarios, where the effort creep starts in 2002 and is undetected in the EM, the OM median SSB was lower than target and F higher than F_{TARGET} at the start of the simulation (Figs. 5Z7 and 5Z8). Nevertheless, management intervention was able to bring median SSB at or above target levels and median F close to target levels for all HCRs in the long term (Figs. 5Z7 and 5Z8). The overall probability of SSB breaching the IATTC's interim LRP of 7.7%$SSB_{F=0}$ remained below 10% for all HCRs (Fig. 5Z9). The probability of breaching their own LRP remained below 20% for all HCRs except HCRs 3 and 11, which have the highest LRP of 20%$SSB_{F=0}$ (Fig. 5Z10). However, the probability of breaching the second rebuilding target of 20%$SSB_{F=0}$ was higher than 20% for all HCRs (Fig. 5Z11). This was due to the high initial probabilities of breaching 20%$SSB_{F=0}$, but probabilities fall below 20% after 6 or 8 years for HCRs with an F_{TARGET} of SPR40% or SPR30% (Fig. 5Z12). For HCRs with the highest F_{TARGET} of 20% or 25%, the probability of breaching 20%$SSB_{F=0}$ remained higher than 20% throughout the simulation under the effort creep scenario.

With regards to status under the effort creep scenario, the probability of F being at or below the F_{TARGET} was at or above 50% only for HCRs 2, 3, 10, and 11 (Fig. 5Z13) due to their higher first control point and earlier management intervention. The probability of F being at or below the F_{TARGET} is really low (12% or lower) at the start of simulation due to the effort creep but all HCRs were able to bring the probability up to 50% following management intervention (Fig. 5Z14). This happened faster for some HCRs depending on their control points (Fig. 5Z14). Furthermore, all HCRs have a higher than 50% probability of SSB being less than the target SSB (Fig. 5Z15). However, that probability fell to 50% or below for all HCRs by year 9 or 10 following management intervention (Fig. 5Z16).

Under the effort creep scenario, there was less catch stability than in the reference set due to increased drastic management intervention (compare Figs. 5Z5 and 5Z17). The decrease in stability was particularly apparent for HCRs 2, 3, 10, and 11, which had the most conservative control points (Fig. 5Z17). Increased management intervention and lower biomass affected overall catch, which was lower under the effort creep scenario (compare table 5C to table 5D). Median annual catch for all HCRs, except those with the highest $F_{TARGETs}$ of FSPR20% or FSPR25%, were below current catch limits in the medium term. However, median annual catch for all HCRs increased above current catch limits in the long term, except for HCRs 3 and 11 (Fig. 5Z18).

Although the degree was different among HCRs, all HCRs had difficulty in dealing with the 10-year recruitment drop scenario. Median SSB declined to levels close to or below the LRP for all HCRs (Fig. 5Z19). Indeed, all HCRs had a higher than 20% probability of breaching their respective LRPs except for HCRs 8, 12, 15, and 16 (Fig. 5Z20). These HCRs have some of the lowest LRPs of 6.3, 7.7, or 10%$SSB_{F=0}$, but F_{TARGET}s of at least FSPR25% and do not switch, like HCRs 5 and 13, to a constant catch management once the ThRP is breached. All HCRs had a higher than 20% probability of breaching 20%$SSB_{F=0}$ (Fig. 5Z21). Furthermore, under this robustness test, all HCRs tuned to meet a 30:70 EPO:WCPO impact ratio (HCRs 9 to 16) did better in safety than their 20:80 HCRs counterpart (Fig. 5Z20) because of the lower fishing mortality on age 0 fish.

Given their different $F_{TARGETS}$ and thus different starting relative SSB levels, median SSB for some HCRs did not fall as low as others (Fig. 5Z19). For instance, HCRs 1 to 3 and 9 to 11 were able to maintain a higher median SSB and had a lower than 10% probability of breaching the interim IATTC LRP of 7.7%$SSB_{F=0}$ (Fig. 5Z22). HCRs 1 to 4, 8, 9 to 12, and 16 had a lower than 10% probability of breaching the first rebuilding target of 6.3%$SSB_{F=0}$ (Fig. 5Z23). These probabilities were computed for 2047 to 2066, 4 years after the start of the recruitment drop through the end of the simulation. Looking at the safety metrics for each year of the simulation showed that HCRs with a higher F_{TARGET} of F25%SPR or F20%SPR had probabilities of SSB being below 7.7%$SSB_{F=0}$ for more years than other HCRs (Fig. 5Z24). This was because they maintained SSB at a lower target level and thus reached lower SSB levels faster once the recruitment drop started in 2042 (Fig. 5Z19). Furthermore, more conservative control points led to faster rebuilding. For example, among the HCRs with a FSPR30% F_{TARGET}, the HCRs with a higher 2nd control point at 20%$SSB_{F=0}$ (i.e., HCRs 1, 2, 9, and 10) reduced the probability of breaching 7.7%$SSB_{F=0}$ to levels below 10% faster than HCRs with a lower 2nd control point at 10 or 7.7%$SSB_{F=0}$ (HCRs 4, 8, 12, and 16) (Fig. 5Z24). Among HCRs with a FSPR25% $F_{TARGET,}$ those rules switching to a pre-specified constant TAC level (i.e., HCRs 5 and 13) showed the slowest decline in the probability of breaching 7.7%$SSB_{F=0}$ to levels below 10% (Fig. 5Z24) and had the slowest rates of SSB increase (Fig. 5Z19).

Even under the recruitment drop scenario, the probability of F remaining at or below the F_{TARGET} was at least 50% for all HCRs, except 5 and 13 (Fig. 5Z25). However, median F

started increasing soon after the start of the low recruitment period and eventually breached target levels for all HCRs (Fig. 5Z26). Since management actions for all HCRs were based on SSB levels, it was only when SSB had been estimated by the EM to have breached the SSB-based control points that management actions began to reduce F. In the meantime, as there are fisheries catching young fish, F increased and median F breached target levels. Median relative SSB started decreasing around 2048, 6 years after the recruitment drop began (Fig. 5Z19). However, the EMs were able to detect the declining SSB and median F started to decrease after control points were breached, quickly falling below target levels (Fig. 5Z25). At the end of this robustness simulation, median F remained well below target levels for all HCRs (Fig. 5Z26) due to the 25% limit on TAC changes between management periods. Some HCRs, such as 1, 2, 9, and 10, which have a F_{TARGET} of FSPR30% and a second control point of 15%$SSB_{F=0,}$ breached their control points earlier and showed earlier drastic declines in F, leading to a lower 5th quantile of F earlier in the simulation than other HCRs (Fig. 5Z25). HCRs 6 and 14, which have a FSPR20% F_{TARGET} and a single control point at 20% also showed an earlier drop in the 5th quantile of F.

Under the recruitment drop scenario, patterns in TAC reflected those of F and SSB. Median TAC started declining with SSB around 2048 and then more steeply following management intervention (Fig. 5Z27). Like for F, HCRs 1, 2, 6, 9, 10, and 14 had more variability in TAC closer to the start of the recruitment drop. Once TAC dropped to low levels, it only increased slowly due to the 25% limit on TAC changes between management periods (Fig. 5Z27), leading to a sharp increase in SSB above target levels (Fig. 5Z19). Indeed, the median upward changes in TAC was 25% for all HCRs except 5 and 13 (Table 5E), which use a catch limit rule when under the ThRP and for which the TAC never fell to low levels (Fig. 5Z27). Note, however, that while TAC was set to this minimum catch level, sometimes there was not enough biomass to meet this TAC, so actual catch did fall to lower levels for some iterations under these HCRs due to the low biomass (Fig. 5Z28).

Under the recruitment drop scenario, median annual catch in the medium term (2048 to 2053) was above the current catch limit for all HCRs, but fell below it in the long term (2053 to 2066) following management intervention (Fig. 5Z29). In the reference set, yield performance was strongly related to F_{TARGET} levels, with HCRs with a higher F_{TARGET} doing best and those with the lowest F_{TARGET} doing poorest. By contrast, in the recruitment drop scenario, HCRs 5 and 13 with a F_{TARGET} of FSPR25% and HCRs 4, 8, 12, and 16, with an F_{TARGET} of FSPR30% performed best in terms of median annual catch (Fig. 5Z29). Unlike the reference set, there was also a large difference in performance in terms of annual catch among HCRs with the same $F_{TARGET.}$ HCRs 1, 2, 9, and 10 had the lowest yield while HCRs 4, 8, 12, and 16, with the same F_{TARGET} of FSPR30%, had the highest (Fig. 5Z29). Moreover, in the robustness scenario, the tradeoff between yield and safety was not apparent. HCRs 3 and 11, with the lowest F_{TARGET}, had higher yield and higher safety than HCRs 6 and 14, which had the highest F_{TARGET} (Fig. 5Z30). HCRs 5 and 13 had a high yield due to switching to constant catch control when biomass declined below their 20%$SSB_{F=0}$ control point, but

this came at the cost of safety (Fig. 5Z30). These HCRs also had the best stability (i.e., the smallest downward change in TAC) (Table 5E, Fig. 5Z31), but had the slowest rate of rebuilding above SSB target levels (Fig. 5Z19). HCRs 4, 8, 12, and 16 had similar yield to HCRs 5 and 13, but had better safety (Fig. 5Z30).

6. Key Limitations and Exceptional Circumstances

The following key limitations were identified:

- Fleet selectivity was assumed to be constant at the current average of 2015-2022 levels throughout the simulation. If fleet operations and targeting behavior change in the future so that the size composition of catch of specific fleets differs widely from what was simulated, results from this analysis may no longer be applicable.
- The operating models were conditioned on data from 1983 onwards, thus the management procedures tested here are robust to uncertainty in productivity that was bound by those historical observations. If future population dynamics strongly diverge from the past, results from this analysis may no longer be applicable.

If one of the MPs presented here were to be adopted for PBF, meta rules may be put in place to assess if exceptional circumstances exist. Exceptional circumstances are situations outside the range for which robustness of the MPs was evaluated and for which a different management action than specified by the adopted MP may be taken. As exemplified by Preece et al. (2023), in addition to the key limitations above, development of such meta rules could consider changes in input data sources to the EM, major changes in our understanding of PBF biology, substantial changes in fishery operations, unusual trends in fishery and stock indicators (e.g., a sustained drop in a recruitment index), and if, once a TAC is placed, total removals differ significantly (i.e., more than what was specified by the implementation error) from what is recommended by the MP. From the results of the "recruitment drop" robustness test, it became apparent that no candidate MP was robust to a chronic drop in recruitment due to a lack of a recruitment index in the MP. Therefore, the monitoring of recruitment and inclusion of recruitment drop as part of exceptional circumstances determination is critical.

7. References

Aikawa, H. and Kato, M. 1938. Age determination of fish. Bulletin of the Japanese Society of Scientific Fisheries. 1938.

Aires-da-Silva, A., Hinton, M. G. and Dreyfus, M. 2007. Standardized catch rates for PBF caught by United States and Mexican-flag purse seine fisheries in the Eastern Pacific Ocean (1960-2006). Working paper submitted to the ISC PBF Working Group Meeting, 11-18 December 2007, Shimizu, Japan. ISC/07/PBFWG-3/01.

Aires-da-Silva, A., Maunder, M., Deriso, R., Piner, K. and Lee, H.-H. 2008. An evaluation of the natural mortality schedule assumed in the PBF 2008 stock assessment and proposed changes. Working paper submitted to the ISC PBF Working Group Meeting, 10-17 December 2008, Ishigaki, Japan. ISC/08/PBF-2/04.

Ashida, H., Suzuki, N., Tanabe, T., Suzuki, N. and Aonuma, Y. 2015. Reproductive condition, batch fecundity, and spawning fraction of large Pacific bluefin tuna *Thunnus orientalis* landed at Ishigaki Island, Okinawa, Japan. Environmental Biology of Fishes 98(4): 1173-1183.

Ashida, H., Okochi, Y., Oshimo, S., Sato, T., Ishihara, Y., Watanabe, S., Fujioka, K., Furukawa, S., Kuwahara, T., Hiraoka, Y. and Tanaka, Y. 2021. Differences in the 76 reproductive traits of Pacific bluefin tuna *Thunnus orientalis* among three fishing grounds in the Sea of Japan. Marine Ecology Progress Series. 662: 125-138.

Ashida, H., Shimose, T., Okochi, Y., Tanaka, Y. and Tanaka, S. 2022. Reproductive dynamics of Pacific bluefin tuna (*Thunnus orientalis*) off the Nansei Islands, southern Japan. Fisheries Research 249: 106256. https://doi.org/10.1016/j.fishres.2022.106256.

Bayliff, W. H. 1994. A review of the biology and fisheries for northern bluefin tuna, *Thunnus thynnus*, in the Pacific Ocean. FAO Fish. Tech. Pap. 336/2: 244-295.

Bayliff, W. H. 1991. Status of northern bluefin tuna in the Pacific Ocean. In: Deriso, R.B. and Bayliff, W.H. (eds.), World meeting on stock assessment of bluefin tunas: strengths and weakness. Special Rep. 7, Inter-Amer. Trop. Tuna Comm., La Jolla, CA. pp. 29-88.

Boustany, A. M., Matteson, R., Castleton, M., Farwell, C. and Block, B. A. 2010. Movements of pacific bluefin tuna (*Thunnus orientalis*) in the Eastern North Pacific revealed with archival tags. Progress in Oceanography 86 (1-2), 94-104.

Carvalho, F., Punt, A.E., Chang, Y.-J., Maunder, M. N. and Piner, K. R. 2017. Can diagnostic tests help identify model misspecification in integrated stock assessments? Fisheries Research 192: 28-40.

Chen, K., Crone, P. and Hsu, C.-C. 2006. Reproductive biology of female Pacific bluefin tuna *Thunnus orientalis* from south-western North Pacific Ocean. Fisheries Science 72, 985–994.

Collette, B. B. 1999. Mackerels, molecules, and morphology. In Séret, B. and Sire, J.-Y. (eds.), Proceedings of 5th IndoPacific Fish Conference, Nouméa, New Caledonia, 1997. Société Française d'Ichthyologie, Paris, France. pp. 149-164.

Conn, P. B., Williams, E. H. and Shertzer, K. W. 2010. When can we reliably estimate the productivity of fish stocks? Can. J. Fish. Aquat. Sci. 67:1-13.

Dewar, H., Snodgrass, O. E., Muhling, B. A. and Schaefer, K. M. 2022. Recent and historical data show no evidence of Pacific bluefin tuna reproduction in the southern California Current system. PLoS ONE 17(5): e0269069.

Foreman, T. 1996. Estimates of age and growth, and an assessment of ageing techniques, for northern bluefin tuna, *Thunnus thynnus*, in the Pacific Ocean. InterAm. Trop. Tuna Comm., Bull., 21(2): 75-105.

Francis, R.I.C.C., 1988. Maximum likelihood estimation of growth and growth variability from tagging data. N.Z. J. Mar. Freshwater Res. 22:42-51.

Fujioka, K., Masujima, M., Boustany, A. M., and Kitagawa, T. 2015. Horizontal movements of Pacific bluefin tuna. In Kitagawa, T. and Kimura, S. (eds.), Biology and ecology of bluefin tuna. CRC Press, Boca Raton London, New York. pp. 101-122.

Fujioka, K., K. Sasagawa, Kuwahara, T., Estess, E. E., Takahara, Y., Komeyama, K., Kitagawa, T., Farwell, C. J., Furukawa, S., Kinoshita, J., Fukuda, H., Kato, M., Aoki, A., Abe, O., Ohshimo, S. and Suzuki, N. 2021. Habitat use of adult Pacific bluefin tuna *Thunnus orientalis* during the spawning season in the Sea of Japan: evidence for a trade-off between thermal preference and reproductive activity. Mar. Ecol. Prog. Ser. 668: 1-20. https://doi.org/10.3354/meps13746.

Fujioka, K., Asai, S., Tsukahara Y., Fukuda, H. and Nakatsuka, S. 2024. Recruitment abundance index of Pacific bluefin tuna based on real-time troll monitoring survey data using Vector Autoregressive Spatio-Temporal (VAST) model analysis. Working paper submitted to the ISC PBF Working Group Meeting, 29 February – 8 March, Kaohsiung, Chinese- Taipei, ISC/24/PBFWG-1/04.

Fukuda, H. 2024. Consideration about a possible unseen change in catchability in the standardized CPUE for the robustness test of the PBF MSE. Working paper submitted to the ISC PBFWG meeting, 10-13 December, webinar. ISC/24/PBFWG-2/06.

Fukuda, H., Yamasaki, I., Takeuchi, Y., Kitakado, T., Shimose, T., Ishihara, T., Ota, T., Watai, M., Lu, H.-B. and Shiao, J.-C. 2015. Estimates of growth function from length-at-age data based on otolith annual rings and daily rings for Pacific Bluefin tuna. Working paper submitted to the ISC PBF Working Group Meeting, 18-25 November, Kaohsiung, Chinese-Taipei. ISC/15/PBFWG-2/11.

Hamel, O. S., Cope, J. M., 2022. Development and considerations for application of a longevity-based prior for the natural mortality rate. Fish. Res. 256:106477. https://doi.org/10.1016/j.fishres.2022.106477.

Heberer, L. N. and Lee, H.-H. 2019. Updated size composition data from the San Diego Commercial Passenger Fishing Vessel (CPFV) recreational fishery for Fleet 15: Eastern Pacific Ocean Sport Fisheries, 2014-2019. Working paper submitted to the ISC PBF Working Group Meeting, 18-23 November, La Jolla, USA. ISC/19/PBFWG-2/06.

Hurtado-Ferro, F., Szuwalski, C. S., Valero, J. L., Anderson, S. C., Cunningham, C. J., Johnson, K. F., Licandeo, R., McGilliard, C. R., Monnahan, C. C., Muradian, M. L., Ono, K., Vert-Pre, K. A., Whitten, A. R., Punt, A. E.. 2014. Looking in the rear-view mirror: bias and retrospective patterns in integrated, age-structured stock assessment models. Ices J. Mar. Sci., 72(1): 99-110.

IATTC. 2012. Conservation and Management Measure for Bluefin Tuna in the Eastern Pacific Ocean. 83rd Meeting of the Inter-American Tropical Tuna Commission. Resolution C-12-09. Available at https://www.iattc.org/getattachment/f36addc0-9cca-48de-a480-5b299e4a1734/C-12-09_Conservation-of-bluefin-tuna.pdf

IATTC. 2014. Measures for the Conservation and Management of Pacific Bluefin Tuna in the Eastern Pacific Ocean, 2015-2016. 87th Meeting of the Inter-American Tropical Tuna Commission. Resolution C-14-06. Available at https://www.iattc.org/getattachment/09ec6d12-6815-4c4e-9135-8f8dc64f83b0/C-14-06_Conservation-of-bluefin-2015-2016.pdf

IATTC. 2021. Measures for the Conservation and Management of Pacific Bluefin Tuna in the Eastern Pacific Ocean. 98th Meeting of the Inter-American Tropical Tuna Commission. Resolution C-21-05. Available at https://www.iattc.org/GetAttachment/b425762e-aba3-4727-ac13-5c9eadd175ac/C-21-05-Active_Bluefin-tuna.pdf

IATTC. 2023a. Amendment to Resolution C-21-01 on a Long-term Management Framework for the Conservation and Management of Pacific Bluefin Tuna in the Eastern Pacific Ocean. 101st Meeting of the Inter-American Tropical Tuna Commission. Resolution C-23-01. Available at https://www.iattc.org/GetAttachment/700f386a-4a72-433c-91b9-08395c948969/C-23-01_Bluefin-tuna.pdf

IATTC. 2023b. Amendment to Resolution C-16-02 on Harvest Control Rules for Tropical Tunas (Yellowfin, Bigeye, and Skipjack). 101st Meeting of the Inter-American Tropical Tuna Commission. Resolution C-23-06. Available at https://www.iattc.org/GetAttachment/cbda923b-b77c-4f4d-a44a-3cdbe3b5fbc6/C-23-06_Harvest-Control-Rules-amends-and-replaces-C-16-02.pdf

IATTC. 2024. Measures for the Conservation and Management of Pacific Bluefin Tuna in the Eastern Pacific Ocean. 102nd Meeting of the Inter-American Tropical Tuna Commission. Resolution C-24-02. Available at https://www.iattc.org/GetAttachment/b02f2675-e880-40a0-bc9b-dabda92adaad/C-24-02_Bluefin-tuna.pdf

Inagake, D., Yamada, H., Segawa, K., Okazaki, M., Nitta, A. and Itoh., T. 2001. Migration of young bluefin tuna, *Thunnus orientalis* Temminck et Schlegel, through archival tagging experiments and its relation with oceanographic condition in the western North Pacific. Bull. Natl. Res. Inst. Far Seas Fish. 38: 53-81.

ISC. 2009. Report of the PBF Working Group Workshop, 10-17 December 2008, Ishigaki, Japan. Annex 4. Report of the Ninth Meeting of the International Scientific Committee for Tuna and Tuna-like Species in the North Pacific Ocean, Plenary Session. 15-20 July 2009, Kaohsiung, Chinese Taipei. ISC/09/ANNEX/04. Available at https://isc.fra.go.jp/pdf/ISC09/Annex_4_ISC9_PBFWG_Dec08.pdf

ISC. 2013. Fishery Impact Plots: Evaluating Fishery Impacts on the Current Stock Status. Document submitted to 13th Meeting of the International Scientific Committee for Tuna and Tuna-like Species in the North Pacific Ocean. ISC/13/ANNEX/14. Available at https://isc.fra.go.jp/pdf/ISC13/Annex14_PB_final_version_0.pdf

ISC. 2014. Stock assessment of Bluefin tuna in the Pacific Ocean in 2014. ISC/14/ANNEX/04. Available at https://isc.fra.go.jp/pdf/2014_Intercessional/Annex4_Pacific_Bluefin_Assmt_Report_2014-June1-Final-Posting.pdf

ISC. 2016. 2016 Pacific Bluefin Tuna Stock Assessment. Document submitted to 16th Meeting of the International Scientific Committee for Tuna and Tuna-like Species in the North Pacific Ocean. ISC/14/ANNEX/09. Available at https://isc.fra.go.jp/pdf/ISC16/ISC16_Annex_09_2016_Pacific_Bluefin_Tuna_Stock_Assessment.pdf

ISC. 2018a. Summary Report of the Pacific Bluefin Tuna Management Strategy Evaluation Workshop. Document submitted to 18th Meeting of the International Scientific Committee for Tuna and Tuna-like Species in the North Pacific Ocean. ISC/18/ANNEX/08. Available at

https://www.isc.fra.go.jp/pdf/ISC18/ISC_18_ANNEX_08_Summary_Report_of_Pacific_Bluefin_MSE_Workshop_(May_2018)_FINAL.pdf

ISC. 2018b. Stock assessment of Pacific Bluefin Tuna in the Pacific Ocean in 2018. ISC/18/ANNEX/14. Available at https://isc.fra.go.jp/pdf/ISC18/ISC_18_ANNEX_14_Pacific_Bluefin_Tuna_Stock_Assessment_2018_FINAL.pdf

ISC. 2019. Report of the Second Pacific Bluefin Tuna Management Strategy Evaluation Workshop. Document submitted to 19th Meeting of the International Scientific Committee for Tuna and Tuna-like Species in the North Pacific Ocean. ISC/19/ANNEX/10. Available at https://isc.fra.go.jp/pdf/ISC19/ISC19_ANNEX10_Report_of_the_second_PBF_MSE_Workshop.pdf

ISC. 2020. Stock assessment of Pacific Bluefin Tuna in the Pacific Ocean in 2020. ISC/20/ANNEX/11. Available at https://isc.fra.go.jp/pdf/ISC20/ISC20_ANNEX11_Stock_Assessment_Report_for_Pacific_Bluefin_Tuna.pdf

ISC. 2022. Stock Assessment of Pacific Bluefin Tuna in the Pacific Ocean in 2022. Document submitted to 22nd Meeting of the International Scientific Committee for Tuna and Tuna-like Species in the North Pacific Ocean. ISC/22/ANNEX/13. Available at https://isc.fra.go.jp/pdf/ISC22/ISC22_ANNEX13_Stock_Assessment_for_Pacific_Bluefin_Tuna.pdf

ISC. 2024a. Stock Assessment of Pacific Bluefin Tuna in the Pacific Ocean in 2024. Document submitted to 24th Meeting of the International Scientific Committee for Tuna and Tuna-like Species in the North Pacific Ocean. ISC/24/ANNEX/13. Available at https://isc.fra.go.jp/pdf/ISC24/ISC24_ANNEX13-Pacific_Bluefin_Tuna_Stock_Assessment_in_2024-FINAL.pdf

ISC. 2024b. Report of the Pacific Bluefin Tuna Working Group Intersessional Meeting. Document submitted to 25th Meeting of the International Scientific Committee for Tuna and Tuna-like Species in the North Pacific Ocean. ISC/25/ANNEX/XX

Ishihara, T., Abe, O., Shimose, T., Takeuchi, Y. and Aires-da-Silva, A. 2017. Use of post-bomb radiocarbon dating to validate estimated ages of Pacific bluefin tuna *Thunnus orientalis*, of the North Pacific Ocean. Fisheries Research 189: 35-41.

Ishihara, T., Tanaka, H., Fukuda, H., Tawa, A., Ashida, H., Tanaka, Y. 2023. Estimation of confidence intervals for the von Bertalanffy growth function parameters using the bootstrap method with a data set of direct age estimates from otoliths. Working paper submitted to

the ISC PBF Working Group Meeting, 21-24 March 2023, Shinagawa, Japan. ISC/23/PBFWG-1/10: 8p. Available at: https://isc.fra.go.jp/pdf/PBF/ISC23_PBF_1/ISC23_PBF_1_10.pdf

Itoh, T. 2006. Sizes of adult bluefin tuna *Thunnus orientalis* in different areas of the western Pacific Ocean. Fish. Sci. 72: 53–62.

Itoh, T., Tsuji, S. and Nitta, A. 2003. Migration patterns of young PBF (*Thunnus orientalis*) determined with archival tags. Fish. Bull. 101: 514-534.

Iwata, S. 2012. Estimate the frequency distribution of steepness for PBF. Working paper submitted to the ISC PBF Working Group Meeting, 10-17 November 2012, Honolulu, Hawaii, USA. ISC/12/PBFWG-3/01: 13p. Available at: http://isc.fra.go.jp/pdf/PBF/ISC12_PBF_3/ISC12_PBFWG-3_01_Iwata.pdf

Iwata, S., Fujioka, K., Fukuda, H. and Takeuchi, Y. 2012a. Reconsideration of natural mortality of age -0 PBF and its variability relative to fish size. Working paper submitted to the ISC PBF Working Group Meeting, 31 January-7 February 2012, La Jolla, California, USA. ISC/12/PBFWG-1/13.

Iwata, S., Fukuda, H., Abe, O., and Takeuchi, Y. 2012b. Estimation of steepness of PBFT -By using biological feature. Working paper submitted to the ISC PBF Working Group Meeting, 31 January-7 February 2012, La Jolla, California, USA. ISC/12/PBFWG-1/15: 9p. Available at: http://isc.fra.go.jp/pdf/PBF/ISC12_PBF_1/ISC12-1PBFWG15_Iwata.pdf

Iwata, S., Kitakado, T., Fujioka, K., Fukuda, H. and Takeuchi, Y. 2014. Revision of the natural mortality estimates for age-0 fish of the Pacific bluefin tuna using 1996-2012 mark-recapture data in Tosa Bay. Working paper submitted to the ISC PBF Working Group Meeting, 17-22 February 2014, La Jolla, USA. ISC/14/PBFWG-1/08.

JWG. 2022. Chairs' Summary of the 7th Joint IATTC and WCPFC-NC Working Group Meeting on the Management of Pacific Bluefin Tuna. JWG07-2022-00. Available at https://meetings.wcpfc.int/node/16046

JWG. 2024. Chairs' Summary of the 9th Joint IATTC and WCPFC-NC Working Group Meeting on the Management of Pacific Bluefin Tuna. JWG09-2024-00. Available at https://meetings.wcpfc.int/node/22881

Kai, M. 2007. Weight-length relationship of North Western PBF. Working paper submitted to the ISC PBF Working Group Meeting, 11-18 December 2007, Shimizu, Japan. ISC/07/PBFWG-3/07

Kitagawa, T., Boustany, A. M, Farwell, C. J., Williams, T. D., Castleton, M.R. and Block, B.A. 2007. Horizontal and vertical movements of juvenile bluefin tuna (*Thunnus orientalis*) in

relation to seasons and oceanographic conditions in the eastern Pacific Ocean. Fisheries Oceanography 16(5): 409-421.

Kleiven, A. R., Espeland, S. H., Stiansen, S., Ono, K., Zimmermann, F. and Olsen, E. M. 2022. Technological creep masks continued decline in a lobster (*Homarus gammarus*) fishery over a century. Sci. Rep. 12: 3318. https://doi.org/10.1038/s41598-022-07293-2

Lee, H.-H., Maunder, M. N., Piner, K. R. and Methot, R. D. 2011. Estimating natural mortality within a fisheries stock assessment model: An evaluation using simulation analysis based on twelve stock assessments. Fish. Res.109(1): 89-94.

Lee, H.-H., Maunder, M. N., Piner, K. R. and Methot, R. D. 2012. Can steepness of the stock-recruitment relationship be estimated in fishery stock assessment models? Fish. Res.125-126: 254-261.

Lee, H.H., Piner, K. R., Methot, R. D. Jr. and Maunder, M. N. 2014. Use of likelihood profiling over a global scaling parameter to structure the population dynamics model: an example using blue marlin in the Pacific Ocean. Fish.Res.158:138-146.

Lee, H.-H., Piner, K. R., Maunder, M. N., and Methot, R. D. 2015. Simulation of methods of dealing with age-based movement in PBF stock assessment. Working paper submitted to the ISC PBF Working Group Meeting, 18-25 November, Kaohsiung, Chinese-Taipei. ISC/15/PBFWG-2/12

Lee, H.H., Piner, K. R., Maunder, M. N., Taylor I. G. and Methot, R. D. 2017. Evaluation of alternative modelling approaches to account for spatial effects due to age-based movement. Can. J. Fish. Aquat. Sci. 74(11): 1832-1844

Lee, H., Fukuda, H., Tsukahara, Y., Piner, K., Maunder, M., and Methot, R. 2021. The devil is in the details: Investigating sources of bootstrapped bias in the Pacific bluefin tuna assessment and the associated impact on the future projections. Working paper submitted to the ISC PBF Working Group Meeting, 20-27 April, Webinar, ISC/21/PBFWG-1/07. Available at: https://isc.fra.go.jp/pdf/PBF/ISC21_PBF_1/ISC_21_PBFWG_1_07_H_Lee.pdf

Lee, H.-H., Tommasi, D., Hiromu, F., Piner, K.R. 2023. Evaluating productivity parameter uncertainty using the age-structured production model diagnostic with recruitment. Working paper submitted to the ISC PBF Working Group Meeting, 21-24 March, Shinagawa Japan. ISC/23/PBFWG-1/09

Lee, H.-H., Tommasi, D. 2023. Evaluating the uncertainty grid: Applying diagnostic tools. Working paper submitted to the ISC PBF Working Group Meeting, 27 November- 1 December, Webinar. ISC/23/PBFWG-2/12.

Lee, H.-H., Tommasi, D. 2024. Evaluating the uncertainty grid using the 2024 stock assessment: Applying diagnostic tools. Working paper submitted to the ISC PBF Working Group Meeting, 29 February- 8 March, Kaohsiung, Chinese-Taipei. ISC/24/PBFWG-1/15.

Lee, S. I., Kim, D. N. and Lee, M. K. 2018. CPUE standardization of Pacific bluefin tuna caught by Korean offshore large purse seine fishery (2003-2016). Working paper submitted to the ISC PBF Working Group Meeting, 5-12 March, La Jolla USA. ISC/18/PBFWG-1/04

Madigan, D. J., Boustany, A. and Collette, B. B. 2017. East not least for Pacific bluefin tuna. Science 357:356–357.

Magnusson, A. and Hilborn, R. 2007. What makes fisheries data informative? Fish and Fisheries. 8: 337-358.

Maguire, J. J. and Hurlbut, T. R.. 1984. Bluefin tuna sex proportion at length in the Canadian samples 1974-1983. Col. Vol. Sci. Pap. ICCAT, 20(2): 341-346.

Maunder, M. N. and Deriso, R. B. 2003. Estimation of recruitment in catch-at-age models. Can. J. Fish. Aquat. Sci., 60: 1204-1216.

Maunder, M. N. and Piner, K. R. 2015. Contemporary fisheries stock assessment: many issues still remain. ICES J. Mar. Sci. 72, https://doi.org/10.1093/icesjms/fsu015

Maunder, M. N., Aires-da-Silva, A., Minte-Vero, C., Valero, J. 2023a. Interim limit and target reference points for tuna, billfishe, and other highly productive fishes in the Eastern Pacific Ocean. 14th Meeting of the Scientific Advisory Committee of the Inter-American Tropical Tuna Commission, Document SAC-14 INF-O. Available at https://www.iattc.org/GetAttachment/6408e9a2-2597-4e79-ab98-68b7fd126e02/SAC-14-INF-O_Proposed-interim-target-and-limit-reference-points.pdf

Maunder, M. N. and Piner, K. R. 2015. Contemporary fisheries stock assessment: many issues still remain. ICES J. Mar. Sci. 72, https://doi.org/10.1093/icesjms/fsu015

Maunder, M. N., Hamel, O. S., Lee, H., Piner, K. R., Cope, J. M., Punt, A. E., Ianelli, J. N., Castillo-Jordan, C., Kapur, M. S., Methot, R. D., 2023b. A review of estimation methods for natural mortality and their performance in the context of fishery stock assessment. Fisheries Research. 257: 106489. https://doi.org/10.1016/j.fishres.2022.106489

Methot Jr., R. D. and Taylor, I. G. 2011. Adjusting for bias due to variability of estimated recruitments in fishery assessment models. Can. J. Fish. Aquat. Sci. 68, 1744-1760.

Methot, R. D. and Wetzel, C. R. 2013. Stock synthesis: A biological and statistical framework for fish stock assessment and fishery management. Fish. Res. 142: 86–99. Elsevier B.V. doi:10.1016/j.fishres.2012.10.012.

Muto, F., Takeuchi, Y. and Yokawa, K. 2008. Annual Catches By Gears of Pacific Bluefin Tuna before 1952 In Japan and Adjacent Areas. Working paper submitted to the ISC PBF Working Group Meeting, 28 May - 4 June 2008, Shimizu, Japan. ISC/08/PBF-01/04.

Nakatsuka, S. 2019. Stock Structure of Pacific Bluefin Tuna (*Thunnus orientalis*) for Management Purposes—A Review of Available Information, Reviews in Fisheries Science & Aquaculture, 28(2), 170–181. https://doi.org/10.1080/23308249.2019.1686455

Nishikawa, K. and Fukuda H. 2024. Sensitivity analysis with different unseen mortality assumptions based on the 2024 stock assessment model. Working paper submitted to the ISC PBF Working Group Meeting, 10-13 December 2024, Webinar. ISC/24/PBFWG-2/05

Ohshimo, S., Sato, T., Okochi, Y., Tanaka, S., Ishihara, T., Ashida, H. and Suzuki, N. 2018. Evidence of spawning among Pacific bluefin tuna, *Thunnus orientalis*, in the Kuroshio and Kuroshio-Oyashio transition area. Aquat. Living Resour., 31: 33. https://doi.org/10.1051/alr/2018022

Okochi, Y., Abe, O., Tanaka, S., Ishihara, Y. and Shimizu, A. 2016. Reproductive biology of female Pacific bluefin tuna, *Thunnus orientalis*, in the Sea of Japan. Fisheries Research 174 30-39.

Palomares, M. L. D., and Pauly, D. 2019. On the creeping increase of vessels' fishing power. Ecology and Society 24(3):31(7 pages). https://doi.org/10.5751/ES-11136-240331

Piner, K. 2012. Selection of an asymptotic selectivity pattern. Working paper submitted to the ISC PBF Working Group Meeting, 30 May-6 June 2012, Shimizu, Japan. ISC/12/PBFWG-2/15: 6p.

Piner, K., Lee, H.-H. and Maunder, M. N. 2016. Evaluation of using random-at-length observations and an equilibrium approximation of the population age structure in fitting the von Bertalanffy growth function. Fisheries Research 180:128-137.

Polovina, J. J. 1996. Decadal variation in the Trans-Pacific migration of northern Bluefin tuna (*Thunnus thynnus*) coherent with climate-induced change in prey abundance. Fish. Oceanogr. 5:114-119.

Polacheck, T., Hearn, W. S., Miller, C., Whitelaw, W. and Stanley, C. 1997. Updated estimates of mortality rates for juvenile SBT from multi-year tagging of cohorts. CCSBT/SC/9707/26.

Preece, A. L, and Davies, C. R. 2023. Evaluation of exceptional circumstances – SBT 2023. CSIRO, Australia. CCSBT-ESC/2308/31. Available at https://www.ccsbt.org/system/files/2023-08/ESC28_31_AU_2023ExcepCircs.pdf

Punt, A. E., Butterworth, D.S., de Moor, C. L., De Oliveira, J. A. A. and Haddon, M. 2016. Management strategy evaluation: best practices. Fish and Fisheries 17, 303–334.

Restrepo, V. R., Thompson, G. G., Mace, P. M., Gabriel, W. L., Low, L. L., MacCall, A. D., Methot, R. D., Powers, J. E., Taylor, B. L., Wade, P. R., and Witzig, J. F. 1998. Technical Guidance on the Use of Precautionary Approaches to Implementing National Standard 1 of the Magnuson-Stevens Fishery Conservation and Management Act, NOAA Technical Memorandum NMFS–F/SPO–31. Available at https://www.st.nmfs.noaa.gov/Assets/stock/documents/Tech-Guidelines.pdf

Rooker, J.R., Wells, R. J. D., Block, B. A., Lui, H., Baumann, H., Chiang, W.-C., Zapp Sluis, M., Miller, N. R., Mohan, J. A., Ohshimo, S., Tanaka, Y., Dance, M. A., Dewar, H., Snodgrass, O. E. and Shiao, J.-C. 2021. Natal origin and age-specific egress of Pacific bluefin tuna from coastal nurseries revealed with geochemical markers. Scientific Reports 11: 14216.

Schaefer, K. 2001. Reproductive biology of tunas. in B. A. Block and E. D. Stevens. (eds.), Tuna - Physiology, Ecology, and Evolution. Academic Press, San Diego. pp. 225-270.

Shimose, T. and Ishihara, T. 2015. A manual for age determination of Pacific Bluefin tuna *Thunnus orientalis*. Bull. Fish. Res. Agen. No. 40, 1-11.

Shimose, T. and Takeuchi, Y. 2012. Updated sex-specific growth parameters for PBF *Thunnus orientalis*. Working paper submitted to the ISC PBF Working Group Meeting, 31 January-7 February 2012, La Jolla, California, USA. ISC/12/PBFWG 1/12.Shimose, T., Tanabe, T., Kai, M., Muto, F., Yamasaki, I., Abe, M., Chen, K.S. and Hsu, C.-C. 2008. Age and growth of Pacific bluefin tuna, *Thunnus orientalis*, validated by the sectioned otolith ring counts. ISC/08/PBF/01/08. 19pp.

Shimose, T., Tanabe, T., Chen, K.-S. and Hsu, C.-C. 2009. Age determination and growth of PBF, *Thunnus orientalis*, off Japan and Taiwan. Fish. Res. 100:134-139.

Takahashi, N., Tsukahara, Y., Tommasi, D., and Fukuda, H. 2024. Final considerations of the use of SS3 ASPM-R as an estimation model in PBF MSE. ISC/24/PBFWG-2/03 Working document submitted to the ISC Pacific Bluefin Tuna Working Group Workshop, 10 – 13 December 2024, online.

Takeuchi, Y. and Takahashi, M. 2006. Estimation of natural morality of age-0 PBF from conventional tagging data. Working paper submitted to the ISC PBF Working Group Meeting, 31 January-7 February 2006, Shimizu, Japan. ISC/06/PBFWG/07.

Tanaka, S. 2006. Maturation of bluefin tuna in the Sea of Japan. Working paper submitted to the ISC PBF Working Group Meeting, 16-20 January 2006, Shimizu, Shizuoka, Japan. ISC/06/PBFWG/09.Then, A. Y., Hoenig, J. M., Hall, N. G., Hewitt, D. A., 2015.

Evaluating the predictive performance of empirical estimators of natural mortality rate using information on over 200 fish species. ICES J. Mar. Sci. 72, 82–92.

Takahashi, N., Tsukahara, Y., Tommasi, D., and Fukuda, H.. 2024. Final considerations of the use of SS3 ASPM-R as an estimation model in PBF MSE. ISC/24/PBFWG-2/03 Working document submitted to the ISC Pacific Bluefin Tuna Working Group Workshop, 10 – 13 December 2024, online.

Takeuchi, Y. and Takahashi, M. 2006. Estimation of natural morality of age-0 PBF from conventional tagging data. Working paper submitted to the ISC PBF Working Group Meeting, 31 January-7 February 2006, Shimizu, Japan. ISC/06/PBFWG/07.

Tanaka, Y., Tawa, A., Ishihara, T., Sawai, E., Nakae, M., Masujima M. and Kodama, T. 2020. Occurrence of Pacific bluefin tuna *Thunnus orientalis* larvae off the Pacific coast of Tohoku area, northeastern Japan: possibility of the discovery of the third spawning ground. Fisheries Oceanography 29: 46–51.

Tawa, A., Ishihara, T., Uematsu, Y., Ono, T. and Ohshimo, S. 2017. Evidence of westward transoceanic migration of Pacific bluefin tuna in the Sea of Japan based on stable isotope analysis. Mar Biol 164, 94. https://doi.org/10.1007/s00227-017-3127-8

Then, A. Y., Hoenig, J. M., Hall, N. G., Hewitt, D. A., 2015. Evaluating the predictive performance of empirical estimators of natural mortality rate using information on over 200 fish species. ICES J. Mar. Sci. 72, 82–92.

Tommasi, D., and Lee, H. 2024. Revised Method to Tune the Relative Fishing Mortality in the Pacific Bluefin tuna MSE for the Requested Proportional Fishery Impact. Working paper submitted to the ISC PBF Working Group Meeting, 29 February- 8 March, Kaohsiung, Chinese-Taipei. ISC/24/PBFWG-01/10. Available at https://isc.fra.go.jp/pdf/PBF/ISC24_PBF_1/2024_ISC_PBFWG-1_10.pdf

Tommasi, D., Lee, H., Piner. K. 2023a. Performance of Candidate Model-based Harvest Control Rules for Pacific Bluefin Tuna. Working paper submitted to the ISC PBF Working Group Meeting, 21-24 March 2023 Shinagawa, Japan. ISC/23/PBFWG-01/14. Available at https://isc.fra.go.jp/pdf/PBF/ISC23_PBF_1/ISC23_PBF_1_14.pdf

Tommasi, D., Lee, H., Piner. K. 2023b. Calculation of Fishery Impact Performance metric for the Pacific Bluefin Management Strategy Evaluation. Working paper submitted to the ISC PBF Working Group Meeting, 21-24 March 2023 Shinagawa, Japan. ISC/23/PBFWG-01/12. Available at https://isc.fra.go.jp/pdf/PBF/ISC23_PBF_1/ISC23_PBF_1_12.pdf

Tommasi, D., and Lee, H. 2024. Revised Method to Tune the Relative Fishing Mortality in the Pacific Bluefin tuna MSE for the Requested Proportional Fishery Impact. Working paper

submitted to the ISC PBF Working Group Meeting, 29 February- 8 March, Kaohsiung, Chinese-Taipei. ISC/24/PBFWG-01/10. Available at https://isc.fra.go.jp/pdf/PBF/ISC24_PBF_1/2024_ISC_PBFWG-1_10.pdf

Tseng, M.-C. and Smith, P.J . 2012. Lack of genetic differentiation observed in Pacific bluefin tuna (*Thunnus orientalis*) from Taiwanese and New Zealand waters using mitochondrial and nuclear DNA markers. Marine and Freshwater Research 63(3) 198-209.

Tsukahara Y., Ishihara, T., Chang, S.-K., Shiao, J.-C., Lee, H.-H. and Fukuda, H. 2024. Re-evaluation of coefficient of variance (CV) in growth curve using the latest otolith data. Working paper submitted to the ISC PBF Working Group Meeting, 29 February- 8 March, Kaohsiung, Chinese-Taipei. ISC/24/PBFWG-1/01

Uematsu, Y., Ishihara, T., Hiraoka, Y., Shimose, T. and Ohshimo, S. 2018. Natal origin identification of Pacific bluefin tuna (*Thunnus orientalis*) by vertebral first annulus. Fish. Res. 199: 26–31.

Wang, S., Maunder, M. N., Aires-da-Silva, A., and Bayliff, W. H. 2009. Evaluating fishery impacts: Application to bigeye tuna (*Thunnus obesus*) in the eastern Pacific Ocean. Fisheries Research: 99, 106-111.

WCPFC. 2009. Conservation and Management Measure for Pacific Bluefin Tuna. Sixth Regular Session of the Western and Central Pacific Fisheries Commission. CMM 2009-07. Available at https://cmm.wcpfc.int/measure/cmm-2009-07

WCPFC. 2013. Conservation and Management Measure for Pacific Bluefin Tuna. Tenth Regular Session of the Western and Central Pacific Fisheries Commission. CMM 2013-09. Available at https://cmm.wcpfc.int/measure/cmm-2013-09

WCPFC. 2014. Conservation and Management Measure for Pacific Bluefin Tuna. Eleventh Regular Session of the Western and Central Pacific Fisheries Commission. CMM 2014-04. Available at https://cmm.wcpfc.int/measure/cmm-2014-04

WCPFC. 2016. Summary Report. Thirteenth Regular Session of the Western and Central Pacific Fisheries Commission. WCPFC-2016-SC12. Available at https://meetings.wcpfc.int/node/10084

WCPFC. 2017.Harvest Strategy for Pacific Bluefin Tuna Fisheries. Fourteenth Regular Session of the Western and Central Pacific Fisheries Commission. HS 2017-02. Available at https://www.wcpfc.int/doc/hs-2017-02/harvest-strategy-pacific-bluefin-tuna-fisheries

WCPFC. 2021. Conservation and Management Measure for Pacific Bluefin Tuna. Eighteenth Regular Session of the Western and Central Pacific Fisheries Commission. CMM 2021-02. Available at https://cmm.wcpfc.int/measure/cmm-2021-02

WCPFC. 2023a. Conservation and Management Measure for Pacific Bluefin Tuna. Twentieth Regular Session of the Western and Central Pacific Fisheries Commission. CMM 2023-02. Available at https://cmm.wcpfc.int/measure/cmm-2023-02

WCPFC. 2023b. Harvest Strategy for Pacific Bluefin Tuna Fisheries. Western and Central Pacific Fisheries Commission. HS 2023-02. Available at https://www.wcpfc.int/doc/hs-2023-02/harvest-strategy-pacific-bluefin-tuna

WCPFC. 2024. Conservation and Management Measure for Pacific Bluefin Tuna . Twenty-first Regular Session of the Western and Central Pacific Fisheries Commission. CMM 2024-01. Available at https://cmm.wcpfc.int/measure/cmm-2024-01

WCPFC NC. 2019. Candidate Reference Points and Harvest Control Rules for Pacific Bluefin Tuna. Fifteenth Regular Session of the Northern Committee of the Western and Central Pacific Fisheries Commission. WCPFC16-2019-NC15, Attachment G. Available at https://meetings.wcpfc.int/node/11422

WCPFC NC. 2023a. Candidate Harvest Control Rules and Reference Points to Evaluate in the Management Strategy Evaluation. Summary Report of Nineteenth Regular Session of the Northern Committee of the Western and Central Pacific Fisheries Commission. WCPFC20-2023-NC19, Attachment G. Available at https://meetings.wcpfc.int/node/19726

WCPFC NC. 2023b. Candidate Operational Management Objectives and Performance Indicators for Pacific Bluefin Tuna. Summary Report of Nineteenth Regular Session of the Northern Committee of the Western and Central Pacific Fisheries Commission. WCPFC20-2023-NC19, Attachment F. Available at https://meetings.wcpfc.int/node/19726

WCPFC SC. 2011. Summary Report of the WCPFC Scientific Committee Seventh Regular Session. Available at https://meetings.wcpfc.int/meetings/sc07

Wells R. J. D., Mohan J. A., Dewar H., Rooker J. R., Tanaka Y. and Snodgrass O. E. 2020. Natal origin of Pacific bluefin tuna from the California Current Large Marine Ecosystem. Biol Lett. 16: 20190878. http://dx.doi.org/10.1098/rsbl.2019.0878.

Yabe, H., Ueyanagi, S. and Watanabe, H. 1966. Studies on the early life history of bluefin tuna *Thunnus thynnus* and on the larvae of the southern bluefin tuna *T. maccoyii*. Report of Nankai Region Fisheries Research Laboratory, 23, 95-129.

Yamada, H. 2007. Reviews of Japanese fisheries and catch statistics on the PBF. Working paper submitted to the ISC PBF Working Group Meeting, 16-23 April 2007, Shimizu, Japan. ISC/07/PBFWG-1/01.

Yukinawa, M. and Yabuta, Y. 1967. Age and growth of bluefin tuna, T*hunnus thynnus* (Linnaeus), in the North Pacific Ocean. Nankai Regional Fish. Res. Lab.

Yonemori, T. 1989. To increase the stock level of the highly migrated pelagic fish. In Marine ranching (Agriculture, Forestry and Fisheries Research Council secretariat, eds.), 8-59. Koseisha-Koseikaku, Tokyo, Japan [In Japanese].

Yoon, S. C., Kim, Z. G., Lee, S. I. and Lee, D. W. 2014. Preliminary analysis of catch at size for Pacific bluefin tuna (*Thunnus orientalis*) caught by Korean offshore large purse seiner. Working paper submitted to the ISC PBF Working Group Meeting, 17-22 February 2014, La Jolla, California, USA. ISC/14/PBFWG-1/12: 13p.

8. Glossary

- **Depletion -** can be defined as spawning biomass depletion or total biomass depletion. It shows what fraction of unfished biomass (spawning or total) the current biomass is. It is calculated as the ratio of the current to unfished biomass (spawning or total).
- **Estimation Model (EM)** – An analytical model that takes data generated with error by the operating model (e.g. catch, abundance index) and produces an estimate of stock status. This often mirrors a stock assessment model.
- **Fishing intensity** – a fishing mortality indicator based on SPR. SPR is the SSB per recruit that would result from the current year's pattern and intensity of fishing mortality relative to the unfished stock. A fishing intensity of F30%SPR would result in 30% of the SSB per recruit relative to the unfished state. This is equivalent to a fishing instensity (1-SPR) of 70%.
- **Harvest control rule (HCR)** - Pre-agreed upon set of rules that specify a management action (e.g. setting the total allowable catch or location/timing of closures) based on a comparison of the status of the system to specific reference points.
- **Management procedure (or management strategy)** - a framework for deciding which fisheries management actions (such as setting a TAC) will achieve stated management objectives. It specifies (1) what harvest control rule will be applied, (2) how stock status estimates will be calculated (e.g. via a stock assessment), and (3) how catch or effort will be monitored.
- **Limit reference point (LRP)** – A benchmark current stock status is compared to and that should not be exceeded with a high probability. It can be biomass-based (e.g. $SSB_{LIMIT)}$ or fishing intensity-based (e.g. F_{LIMIT}).
- **Management Objectives** – High-level goals of a management plan (e.g. prevent overfishing or promote profitability of the fishery).
- **Management Strategy Evaluation (MSE)** – a simulation-based analysis to evaluate trade-offs achieved by alternative harvest (or management) strategies and to assess the consequences of uncertainty in achieving management objectives
- **Operating Model (OM)** – Mathematical representation of plausible versions of the true dynamics of the system under consideration. These are conditioned on historical data. Generally, multiple OMs are required to represent the range of uncertainty in different factors. OMs can range in complexity (e.g. from single species to ecosystems models) depending on the management objectives and management strategies being evaluated.
- **Performance metrics** - Quantitative indicators that are used to evaluate each HCR and serve as a quantitative representation of the management objectives.
- **Spawning potential ratio (SPR)** – the ratio of spawning stock biomass per recruit under fishing to spawning stock biomass per recruit under unfished conditions.
- **SSB** – spawning stock biomass.

- **$SSB_{F=0}$** – equilibrium unfished spawning stock biomass.
- **Target reference point (TRP)** - A benchmark which a current stock levels is compared to. It represents a desired state that management intends to achieve. It can be biomass-based (e.g. SSB_{TARGET}) or fishing mortality-based (e.g. F_{TARGET}).
- **Threshold reference point** – A benchmark current stock status is compared to. Its value is between that of a target and limit reference point. It represents a control point below which a management action is undertaken to bring the stock back to a target state.

9. Tables

Table 3A. Definition of fleets in the stock assessment (ISC 2024a) and MSE of Pacific bluefin tuna *Thunnus orientalis*.

Fleet #	Fleet description	Abundance index
1	Japanese longline fisheries (JPN_LL) for all seasons for 1983-1992, and for season 4 for 1993-2016	S1 (1993-2019), S2 (1983-1992)
2	Japanese longline fisheries (JPN_LL) for seasons 1-3 for 1993-2016 and all seasons for 2017-2022	
3	Taiwanese longline fishery (TWN_LL) in southern fishing ground for 1983-2022	S5 (2002-2022)
4	Taiwanese longline fishery (TW_LL) in northern fishing ground for 2000-2022	
5	Japanese tuna purse seine fishery off the Pacific coast of Japan (JPN_TPS_ PO) for 1983-2022	
6	Japanese tuna purse seine fishery in the Sea of Japan (JPN_TPS_SOJ) for 1983-2022	
7	Japanese tuna purse seine fishery in the Sea of Japan for farming (JPN_TPS_SOJ Farming) for 2016-2022	
8	Japanese small pelagic fish purse seine fishery in the East China Sea (JPN_SPPS) for seasons 1, 3, and 4 for 1987-2022	
9	Japanese small pelagic fish purse seine fishery in the East China Sea (JPN_SPPS) for season 2 for 1988-2022	
10	Japanese small pelagic fish purse seine fishery in the East China Sea for farming (JPN_SPPS Farming) for 2014-2022	
11	Korean offshore large scale purse seine fishery (KOR_LPPS) for 1983-2022	
12	Japanese troll fishery (JPN_Troll) for seasons 2-4 for 1983-2022	S3 (1983-2010)
13	Japanese troll fishery (JPN_Troll) for season 1 for 1983-2022	
14	Japanese troll fishery for farming (JPN_Troll Farming) for season 1 for 1998-2022	

15	Japanese pole and line fishery (JPN_PL) for 1983-2022	
16	Japanese set-net fisheries (JPN_Setnet) for seasons 1-3 for 1983-2022	
17	Japanese set-net fisheries (JPN_Setnet) for season 4 for 1983-2022	
18	Japanese set-net fisheries in Hokkaido and Aomori (JPN_Setnet (HK_AM)) for 1983-2022	
19	Japanese other fisheries (JPN_Others), mainly small-scale fisheries in the Tsugaru Strait for season 2 for 1983-2022	
20	Eastern Pacific Ocean commercial purse seine fishery (U.S. dominant) (EPO_COMM(-2001)) for 1983-2001	
21	Eastern Pacific Ocean commercial purse seine fishery (Mexico dominant) (EPO_COMM(2002-)) for 2001-2022	
22	Eastern Pacific Ocean sports fishery (EPO_SP(2014-)) for 2014-2022	
23	Eastern Pacific Ocean sports fishery (EPO_SP(-2013)) for 1983-2013	
24	Unaccounted mortality fisheries (in weight) in WPO (WPO_Disc_Weight) for 2017-2022	
25	Unaccounted mortality fisheries (in number) in WPO (WPO_Disc_Num) for 1998-2022	
26	Unaccounted mortality fisheries (in number) in EPO (EPO_Disc_Num) for 1999-2022	

Table 4A. Fleet-specific effective sample sizes used to generate data with error from the operating model for input into the estimation model. Note that for some fleets, size composition data is available for more than one season.

Fleet #	Season	Effective Sample Size
1	11.5	10
2	2.5	9
2	11.5	14
3	8.5	14
4	2.5	11
5	11.5	10
6	5.5	12
6	8.5	7
6	11.5	4
12	11.5	11
14	8.5	9
15	5.5	6
15	11.5	6
17	11.5	3
18	5.5	10
20	11.5	16
21	8.5	3
21	11.5	10

Table 4B. Details of how the multiplier of the current fishing mortality that would achieve the F_{TARGET} ($Fmult_{TARGET}$) is changed in the MSE management module depending on the stock status as specified by each candidate harvest control rule. $SSB_{current}$ is the terminal year spawning stock biomass relative to unfished. ThRP is the threshold reference point, LRP is the limit reference point. $Fmult_{min}$ is the fraction of the $Fmult_{target}$ as defined by F_{min} in Table y.

Stock Status	HCR	F multiplier
$SSB_{current} \geq$ ThRP	All	$Fmult = Fmult_{TARGET}$
LRP < $SSB_{current}$ < ThRP	1, 2, 3, 4, 8	$Fmult = (Fmult_{TARGET} - Fmult_{min})*(SSB_{current} - LRP)/(ThRP-LRP) + Fmult_{min}$
$SSB_{current}$ < ThRP	6, 7	$Fmult = (Fmult_{TARGET} /ThRP)*SSB_{current}$
	5	No Fmult, managed by constant catch set by CMM
$SSB_{current} \leq$ LRP	1, 2, 3, 4, 8	$Fmult = Fmult_{min}$

Table 5A. Stability performance indicator calculated as the median % downward change in total allowable catch (TAC) between management periods with and without years when SSB is below LRP for each harvest control rule (HCR) across all iterations, evaluation years, and reference set operating models. (% change TAC +)

HCR	Median % downward change in TAC when SSB≥LRP	Median % downward change in TAC all years
1	-11	-12
2	-14	-15
3	-9	-12
4	-11	-11
5	-12	-12
6	-14	-14
7	-11	-11
8	-11	-11
9	-10	-11
10	-12	-13
11	-8	-10
12	-10	-10
13	-12	-12
14	-13	-13
15	-10	-10
16	-10	-10

Table 5B. Performance indicators for each harvest control rule (HCR) across all iterations, and evaluation years for operating model 1 under the doubling of discards robustness test. SSB refers to spawning stock biomass, LRP to limit reference point, $SSB_{F=0}$ refers to unfished spawning stock biomass, F refers to fishing mortality measured as 1-SPR where SPR is spawning potential ratio, TAC refers to total allowable catch, WCPO refers to Western Central Pacific Ocean and EPO refers to Eastern Pacific Ocean. Note that to ensure that for all indicators a higher value is better, here we reversed the performance metrics shown in Figures ES5 and ES7 to be the probability of SSB≥LRP and of SSB≥20%$SSB_{F=0}$. The % change upwards in TAC (% change TAC +) was set to negative so that high values (smaller -) are better. The % change downwards does not include years when SSB is below LRP as provided by the management objective. Color shadings reflect the range of each column. Highest levels have dark green, lowest light yellow, and different shades of green to yellow are in between. As there is no optimal impact, the EPO impact column does not have a color.

Performance Indicators

Double Discards Robustness Scenario

	Prob SSB => LRP	Prob SSB => 20%SSBo	Prob F <= Ftarget	Prob SSB => SSBtarget	% change TAC +	% change TAC -	EPO Impact	Median annual catch	Median years 5-10 annual catch	Median years 11-23 annual catch	Median WCPO large fish annual TAC	Median WCPO small fish annual TAC	Median EPO annual TAC
1	97	92	59	56	-12	-9	19	30160	28792	31571	16979	4028	6754
2	97	93	61	58	-13	-13	19	29986	28142	31552	16714	4021	6772
3	98	98	59	46	-14	-6	19	26216	23428	28023	14503	3299	5915
4	99	92	58	55	-11	-9	19	30180	28792	31571	16987	4028	6754
5	99	83	54	69	-11	-13	18	31544	31424	32311	17624	4516	6872
6	97	75	59	75	-11	-15	18	33068	33326	33258	18532	4946	7256
7	99	84	57	70	-11	-11	18	31765	31429	32632	17779	4474	6976
8	100	92	58	55	-11	-9	19	30180	28792	31571	16987	4027	6754
9	98	93	63	57	-10	-7	29	29386	28125	30573	13294	3722	10424
10	98	94	64	59	-13	-10	29	29354	27804	30455	13290	3722	10335
11	99	99	63	47	-11	-6	30	25418	22996	27202	11412	3017	8919
12	100	93	63	57	-10	-8	29	29395	28125	30573	13294	3722	10433
13	100	87	62	72	-11	-10	28	31107	30864	31156	14166	4120	10910
14	98	77	63	77	-14	-14	28	33126	33504	32085	15040	4553	11323
15	99	86	61	72	-11	-9	29	31247	30909	31601	14180	4093	11067
16	100	93	63	57	-10	-8	29	29386	28125	30573	13294	3722	10425

Table 5C. Performance indicators for each harvest control rule (HCR) across all iterations, and evaluation years for operating model 1 in the reference set. SSB refers to spawning stock biomass, LRP to limit reference point, $SSB_{F=0}$ refers to unfished spawning stock biomass, F refers to fishing mortality measured as 1-SPR where SPR is spawning potential ratio, TAC refers to total allowable catch, WCPO refers to Western Central Pacific Ocean and EPO refers to Eastern Pacific Ocean. Note that to ensure that for all indicators a higher value is better, here we reversed the performance metrics shown in Figures ES5 and ES7 to be the probability of SSB≥LRP and of SSB≥20%$SSB_{F=0}$. The % change upwards in TAC (% change TAC +) was set to negative so that high values (smaller -) are better. The % change downwards does not include years when SSB is below LRP as provided by the management objective. Color shadings reflect the range of each column. Highest levels have dark green, lowest light yellow, and different shades of green to yellow are in between. As there is no optimal impact, the EPO impact column does not have a color.

Performance Indicators

Reference Set OM1

	Prob SSB => LRP	Prob SSB => 20%SSBo	Prob F <= Ftarget	Prob SSB => SSBtarget	% change TAC +	% change TAC -	EPO Impact	Median annual catch	Median years 5-10 annual catch	Median years 11-23 annual catch	Median WCPO large fish annual TAC	Median WCPO small fish annual TAC	Median EPO annual TAC
1	98	95	69	62	-12	-8	19	29439	27760	31227	17389	4033	6794
2	98	95	70	64	-13	-12	19	29414	27170	31075	17347	4030	6794
3	99	99	69	51	-14	-6	20	25591	22579	27585	14885	3311	5970
4	100	95	69	61	-12	-8	19	29446	27760	31227	17393	4032	6794
5	100	89	66	77	-11	-12	19	30843	30293	32162	18181	4517	6990
6	99	81	67	81	-11	-13	19	32149	32105	33148	18799	4977	7375
7	100	90	68	77	-11	-10	19	30951	30326	32432	18217	4488	7055
8	100	95	69	61	-12	-9	19	29439	27760	31203	17389	4034	6794
9	99	96	73	62	-9	-7	30	28859	27319	30168	13582	3722	10486
10	99	96	74	64	-13	-9	30	28791	27134	30097	13547	3722	10392
11	99	99	73	51	-11	-6	31	24861	22309	26726	11632	3023	8998
12	100	96	73	62	-9	-7	30	28864	27319	30201	13582	3722	10489
13	100	91	71	78	-11	-9	29	30625	29912	31420	14464	4125	11096
14	99	83	71	83	-13	-12	29	32236	32421	32112	15040	4569	11323
15	100	90	71	78	-10	-8	30	30715	29937	31572	14463	4105	11168
16	100	96	73	62	-9	-7	30	28864	27319	30201	13582	3722	10490

Table 5D. Performance indicators for each harvest control rule (HCR) across all iterations, and evaluation years for operating model 1 in the effort creep robustness scenario. SSB refers to spawning stock biomass, LRP to limit reference point, $SSB_{F=0}$ refers to unfished spawning stock biomass, F refers to fishing mortality measured as 1-SPR where SPR is spawning potential ratio, TAC refers to total allowable catch, WCPO refers to Western Central Pacific Ocean and EPO refers to Eastern Pacific Ocean. Note that to ensure that for all indicators a higher value is better, here we reversed the performance metrics shown in Figures ES5 and ES7 to be the probability of SSB≥LRP and of SSB≥20%$SSB_{F=0}$. The % change upwards in TAC (% change TAC +) was set to negative so that high values (smaller -) are better. The % change downwards does not include years when SSB is below LRP as provided by the management objective. Color shadings reflect the range of each column. Highest levels have dark green, lowest light yellow, and different shades of green to yellow are in between. As there is no optimal impact, the EPO impact column does not have a color.

Performance Indicators

Effort Creep Robustness Scenario

	Prob SSB => LRP	Prob SSB => 20%SSBo	Prob F <= Ftarget	Prob SSB => SSBtarget	% change TAC +	% change TAC -	EPO Impact	Median annual catch	Median years 5-10 annual catch	Median years 11-23 annual catch	Median WCPO large fish annual TAC	Median WCPO small fish annual TAC	Median EPO annual TAC
1	82	67	45	42	-17	-15	23	26330	23884	27691	14440	4066	6793
2	84	70	50	47	-21	-22	24	25569	22504	26748	13970	3984	6506
3	78	78	53	45	-20	-21	24	21226	19679	22490	11280	3377	5936
4	96	66	42	40	-16	-14	23	26748	23884	28325	14793	4092	6794
5	96	53	38	40	-15	-13	23	27542	26576	27518	15040	4616	6841
6	94	48	48	48	-19	-17	23	29236	29411	27966	15040	4959	7170
7	97	56	42	42	-16	-13	23	28087	26617	28932	15040	4539	7104
8	99	66	42	41	-16	-15	23	26756	23884	28377	14760	4095	6794
9	81	66	48	42	-18	-16	34	26482	23988	28229	11803	3659	10108
10	84	69	55	48	-23	-21	35	25406	22997	26859	11460	3619	9628
11	77	77	52	41	-21	-19	34	22085	19126	24363	9735	2984	8187
12	95	65	43	40	-16	-16	35	27343	23988	29020	12078	3678	10261
13	96	56	44	42	-16	-17	26	27091	26070	26799	12909	4095	9952
14	94	48	49	48	-20	-17	33	29505	28872	28716	13592	4418	10561
15	96	55	42	41	-16	-15	34	28809	26262	29536	12941	4029	10825
16	98	65	43	40	-16	-17	34	27250	23988	29094	12088	3684	10270

Table 5E. Performance indicators for each harvest control rule (HCR) across all iterations, and evaluation years for operating model 1 in the recruitment drop robustness scenario. SSB refers to spawning stock biomass, LRP to limit reference point, $SSB_{F=0}$ refers to unfished spawning stock biomass, F refers to fishing mortality measured as 1-SPR where SPR is spawning potential ratio, TAC refers to total allowable catch, WCPO refers to Western Central Pacific Ocean and EPO refers to Eastern Pacific Ocean. Note that to ensure that for all indicators a higher value is better, here we reversed the performance metrics shown in Figures ES5 and ES7 to be the probability of SSB≥LRP and of SSB≥20%$SSB_{F=0}$. The % change upwards in TAC (% change TAC +) was set to negative so that high values (smaller -) are better. The % change downwards does not include years when SSB is below LRP as provided by the management objective. Color shadings reflect the range of each column. Highest levels have dark green, lowest light yellow, and different shades of green to yellow are in between. As there is no optimal impact, the EPO impact column does not have a color.

Performance Indicators

Recruitment Drop Robustness Scenario

	Prob SSB => LRP	Prob SSB => 20%SSBo	Prob F <= Ftarget	Prob SSB => SSBtarget	% change TAC +	% change TAC -	EPO Impact	Median annual catch	Median years 5-10 annual catch	Median years 11-23 annual catch	Median WCPO large fish annual TAC	Median WCPO small fish annual TAC	Median EPO annual TAC
1	71	62	62	48	-25	-20	20	4751	29344	2975	1991	988	1396
2	74	64	63	49	-25	-22	20	6857	29109	3379	3517	1705	2433
3	74	74	54	38	-25	-18	19	18220	27895	3256	11777	2590	4268
4	79	56	51	40	-25	-21	21	19378	29716	11506	12185	3175	4705
5	64	30	32	23	-16	-16	25	19699	28606	18486	9511	4725	5473
6	75	49	60	49	-25	-24	23	9564	28893	3467	6944	3211	4190
7	79	50	51	43	-25	-22	22	17006	30301	6963	11513	3447	4738
8	86	56	50	40	-25	-22	20	19743	29829	14212	12532	3126	4669
9	73	62	60	48	-25	-18	29	6182	27971	3180	2180	1294	3032
10	75	64	62	49	-25	-23	30	7679	27971	3462	5966	2115	5551
11	75	75	55	39	-25	-17	29	17310	26365	3280	8919	2450	6774
12	82	57	52	40	-25	-21	31	19184	28201	12942	9149	2894	7312
13	78	39	38	30	-15	-19	24	19519	27993	18486	9511	4725	6658
14	78	51	60	51	-25	-23	31	12798	28242	5459	5525	3026	6809
15	82	51	51	43	-25	-22	30	17900	28881	8585	8499	3112	7230
16	89	57	51	41	-25	-23	29	19522	28404	14817	9406	2846	7157

Table 5F. Performance indicators for each harvest control rule (HCR) across all evaluation years, operating models in the reference set, and 100 iterations, including those associated with non-converged EMs. SSB refers to spawning stock biomass, LRP to limit reference point, $SSB_{F=0}$ refers to unfished spawning stock biomass, F refers to fishing mortality measured as 1-SPR where SPR is spawning potential ratio, TAC refers to total allowable catch, WCPO refers to Western Central Pacific Ocean and EPO refers to Eastern Pacific Ocean. Note that to ensure that for all indicators a higher value is better, here we reversed the performance metrics shown in Figures ES5 and ES7 to be the probability of SSB≥LRP and of SSB≥20%$SSB_{F=0}$. The % change upwards in TAC (% change TAC +) was set to negative so that high values (smaller -) are better. The % change downwards does not include years when SSB is below LRP as provided by the management objective. The value including years when SSB is below LRP is provided in the main body of the report. The 2026 TAC is the total TAC and the TAC for each fleet segment that could be applied in 2026 if each of the HCR would be adopted. It is calculated based on biomass status estimated by EM. Color shadings reflect the range of each column. Highest levels have dark green, lowest light yellow, and different shades of green to yellow are in between. As there is no optimal impact, the EPO impact column does not have a color.

Performance Indicators

Reference Set - 100 iterations

	Prob SSB => LRP	Prob SSB => 20%SSBo	Prob F <= Ftarget	Prob SSB => SSBtarget	% change TAC +	% change TAC -	EPO Impact	Median annual catch	Median years 5-10 annual catch	Median years 11-23 annual catch	Median WCPO large fish annual TAC	Median WCPO small fish annual TAC	Median EPO annual TAC	2026 TAC	2026 TAC WCPO large fish	2026 TAC WCPO small fish	2026 TAC EPO
1	92	85	81	61	-15	-11	23	28299	26687	30548	15793	4092	6794	25868	14836	4512	6520
2	93	86	84	63	-16	-15	23	27924	25932	30205	15207	4063	6794	25868	14836	4512	6520
3	90	90	85	55	-18	-9	24	23746	21011	25910	13218	3344	5961	24366	14836	3844	5686
4	98	85	80	60	-14	-12	23	28344	26691	30592	15844	4103	6794	25868	14836	4512	6520
5	99	79	77	65	-13	-12	22	29791	28774	31537	16523	4627	7020	27485	14836	5161	7488
6	99	74	79	74	-14	-14	22	31822	31160	33426	17607	5060	7568	29437	14836	5939	8662
7	100	80	78	66	-13	-11	23	30250	28820	32173	16905	4560	7168	27485	14836	5161	7488
8	100	85	80	60	-14	-12	23	28360	26690	30598	15868	4102	6794	25868	14836	4512	6520
9	92	85	83	62	-14	-10	33	28456	26429	30070	13177	3722	10469	27942	14073	4392	9476
10	93	86	85	64	-17	-13	33	28284	25878	29915	13042	3722	10348	27942	14073	4392	9476
11	90	90	86	55	-17	-8	32	23462	21294	25815	10685	3016	8556	23653	10724	3844	9085
12	98	85	82	60	-13	-10	33	28528	26432	30133	13218	3722	10520	27942	14073	4392	9476
13	99	80	81	66	-14	-13	30	30460	29272	31238	14345	4163	11103	29323	14836	5010	9476
14	99	75	81	75	-15	-14	31	32223	32045	33027	15040	4588	11323	30061	14836	5749	9476
15	99	80	80	65	-12	-11	32	30691	29293	31671	14351	4103	11295	29323	14836	5010	9476
16	100	85	82	61	-13	-11	33	28530	26433	30148	13220	3722	10517	27942	14073	4392	9476

10. Figures

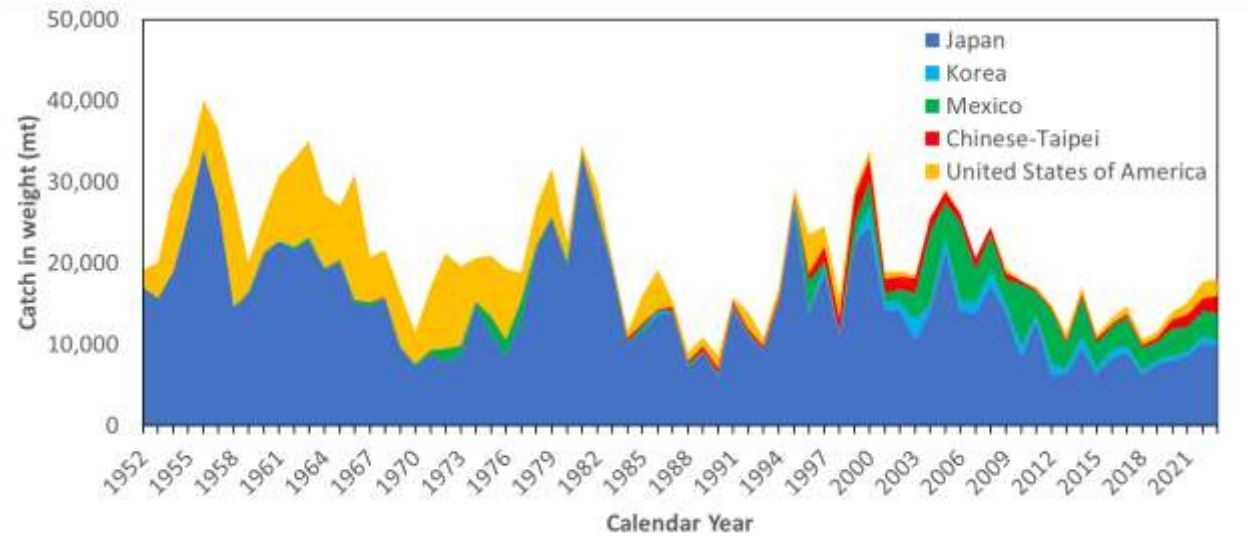

Figure 2A. Annual catch (tons) of Pacific bluefin tuna (*Thunnus orientalis*) by ISC member countries from 1952 through 2023 (calendar year) based on ISC official statistics.

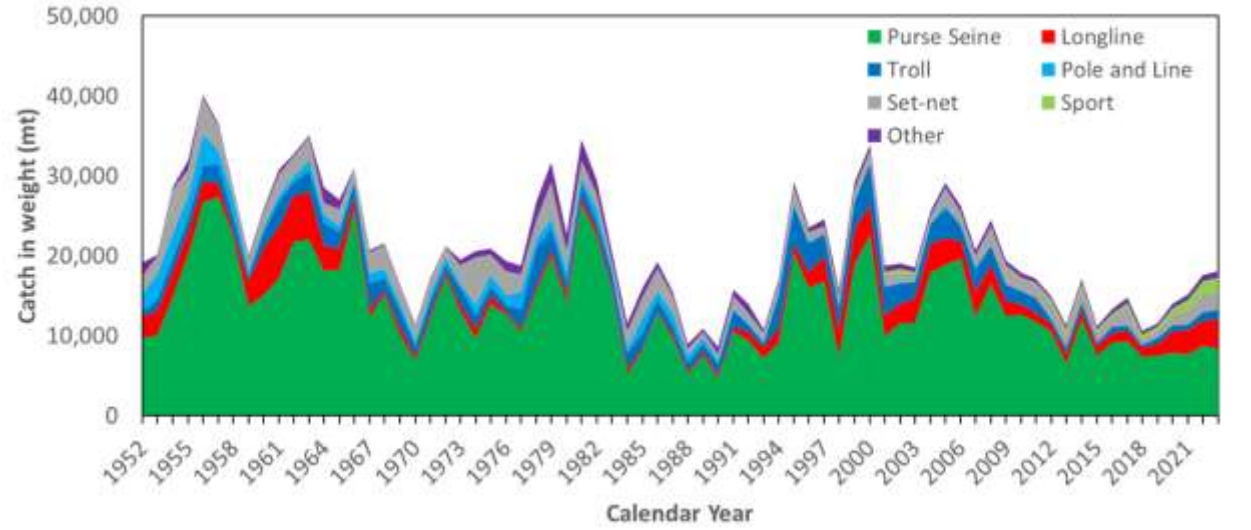

Figure 2B. Annual catch (tons) of Pacific bluefin tuna (*Thunnus orientalis*) by gear type by ISC member countries from 1952 through 2023 (calendar year) based on ISC official statistics.

Figure 3A. Overview of the simulated management cycle for the first five years of the simulation. As the current Pacific bluefin tuna assessment, the simulated stock assessment or EM (estimation model) uses a fishing year (FY) from July to June. For example, fishing year 2023 is from July 2023 to June 2024. The MSE simulation starts in 2023 and the first EM is conducted with data up to FY 2023, and sets an annual TAC from 2026 to 2028. The next EM happens three years later, uses data up to FY 2026, and sets a TAC for 2029 to 2031. In the years before the first TAC comes into effect, 2023 to 2025, observed catch or existing CMMs are used to set the catches.

Figure 4A. Worm plots of recruits (Number of individuals) for individual runs from the operating model of the recruitment drop robustness scenario. Each panel presents the results for the labeled HCR. Trajectories represent separate iterations differing in simulated random recruitment deviates.

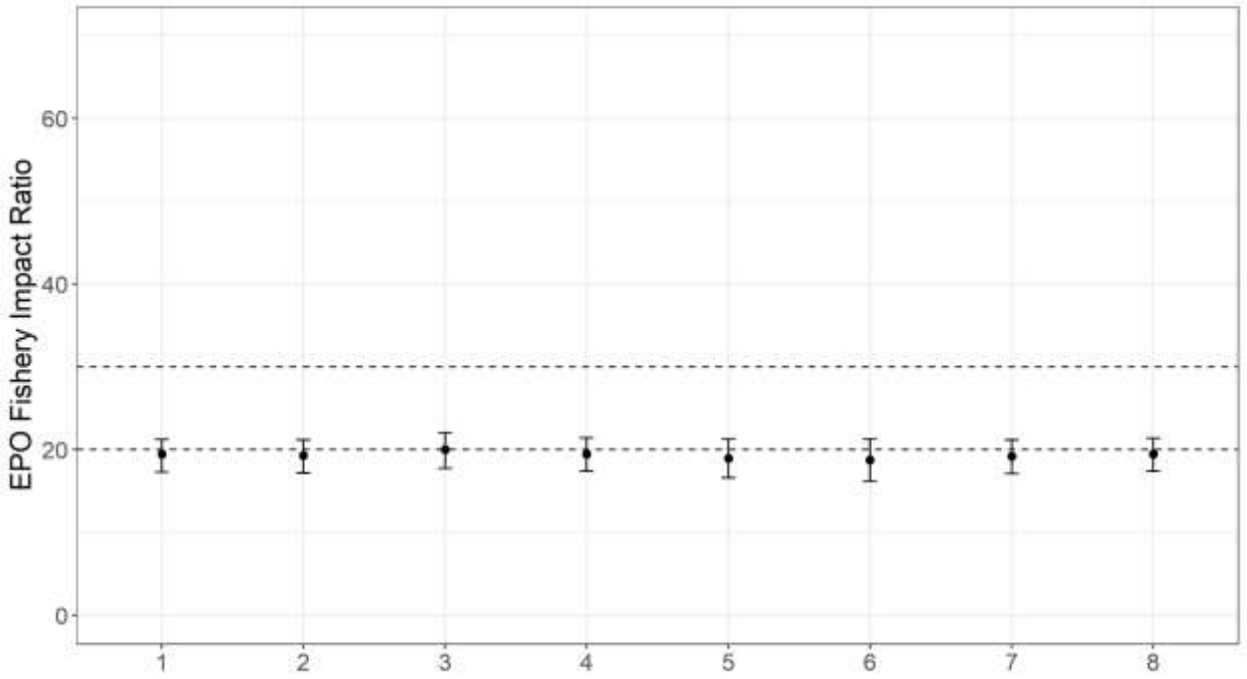

Figure 4B. Median (dot) and 5th to 95th quantile range (error bars) of fishery impact on spawning stock biomass for the Eastern Pacific Ocean (EPO) fleet segment in the terminal year of the MSE simulation across all iterations for operating model 1 for HCRs 1 to 8. The simulation was run using a 2015-2022 baseline relative fishing mortality between fleets. The impact of the Western and Central Pacific Ocean fleet segment is 100 – the EPO impact.

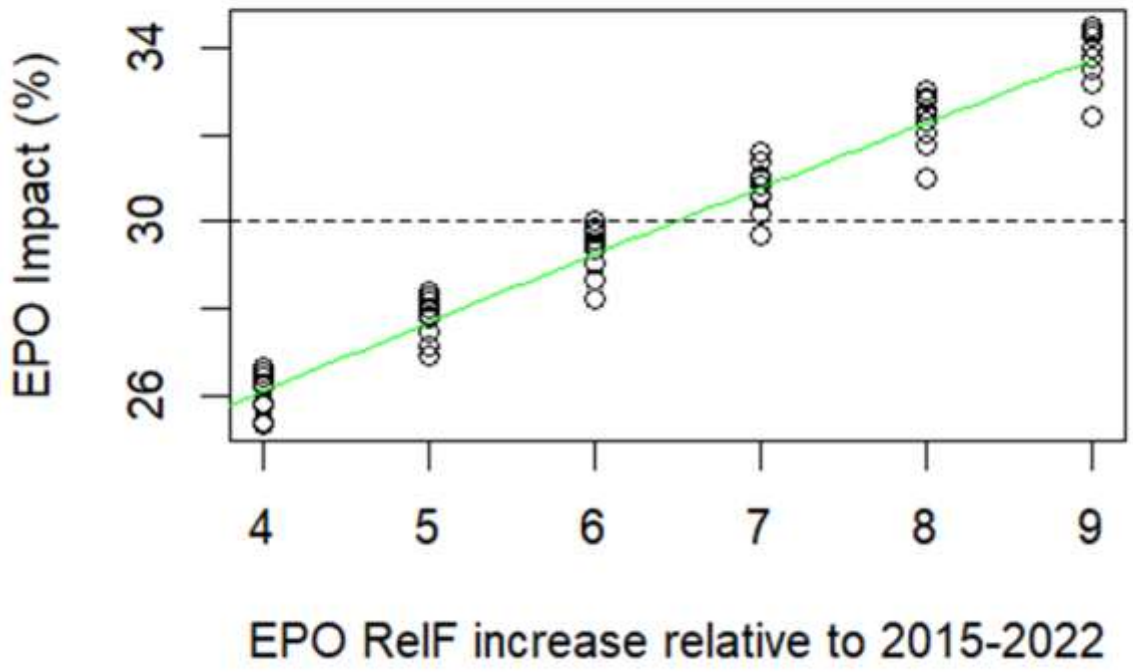

Figure 4C. EPO relative F increase from 2015-2022 average levels and associated proportional EPO fishery impact for the 100 simulations at each different relative F level (black circles). The green line is the best fit polynomial regression and the dotted line shows the desired 30% EPO proportional fishery impact. The intersection of the green and dotted line is the relative F increase relative to 2015-2022 average levels required to reach a 30% EPO proportional fishery impact.

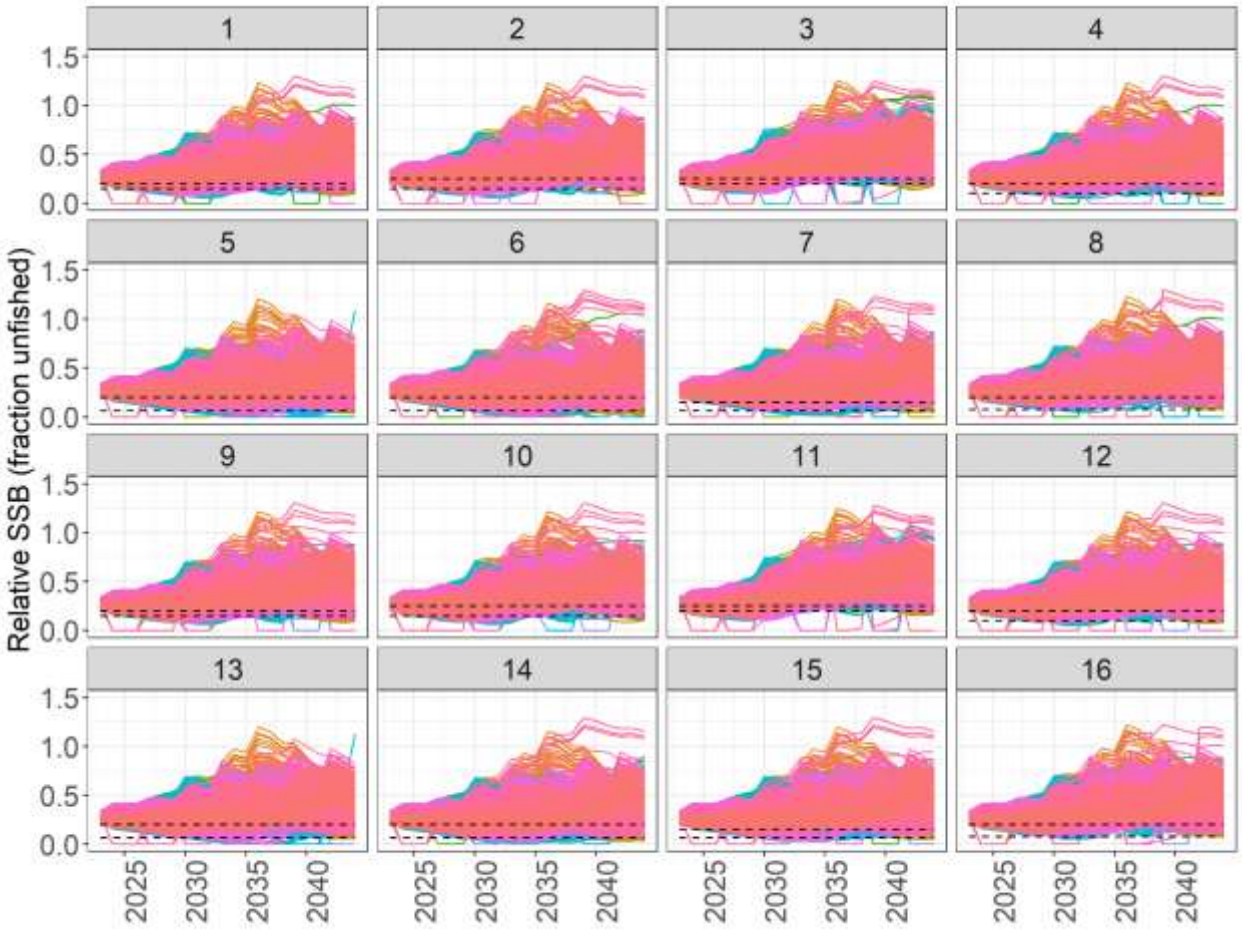

Figure 5A. Worm plots of estimated relative spawning stock biomass ($SSB/SSB_{F=0}$) from the estimation model (EM, i.e. simulated stock assessment) for individual runs and all reference scenarios. Each panel presents the results for the labeled HCR. Trajectories represent the 100 iterations differing in simulated random recruitment deviates for each reference scenario. Each estimated SSB trajectory was compiled using the estimated terminal SSB from the first EM in 2023 and the estimated SSB for the last three years of each of the other 7 EMs.

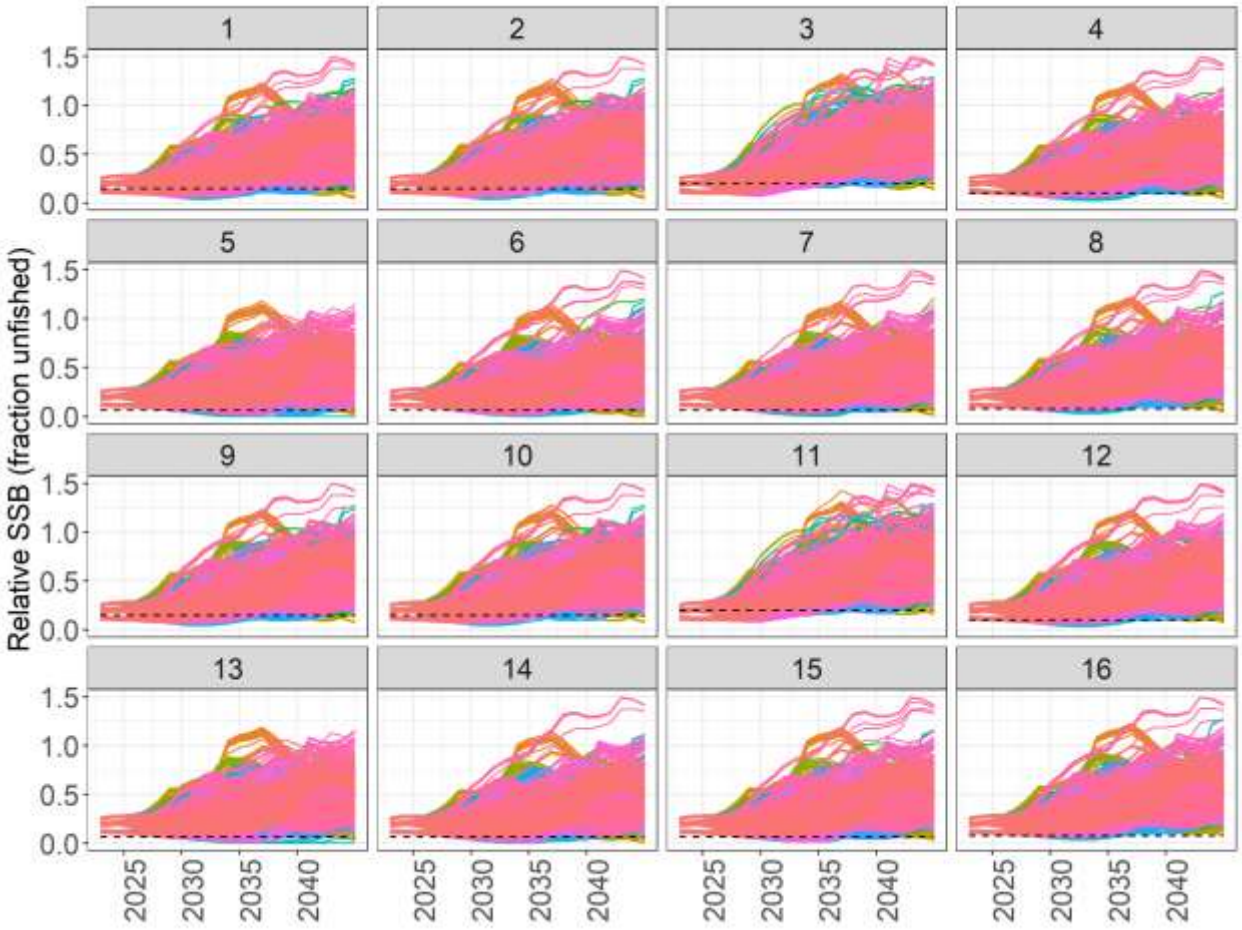

Figure 5B. Worm plots of estimated relative spawning stock biomass ($SSB/SSB_{F=0}$) from the operating model (OM) for individual runs for all reference scenarios. Each panel presents the results for the labeled HCR. Trajectories represent the 100 iterations differing in simulated random recruitment deviates for each reference scenario.

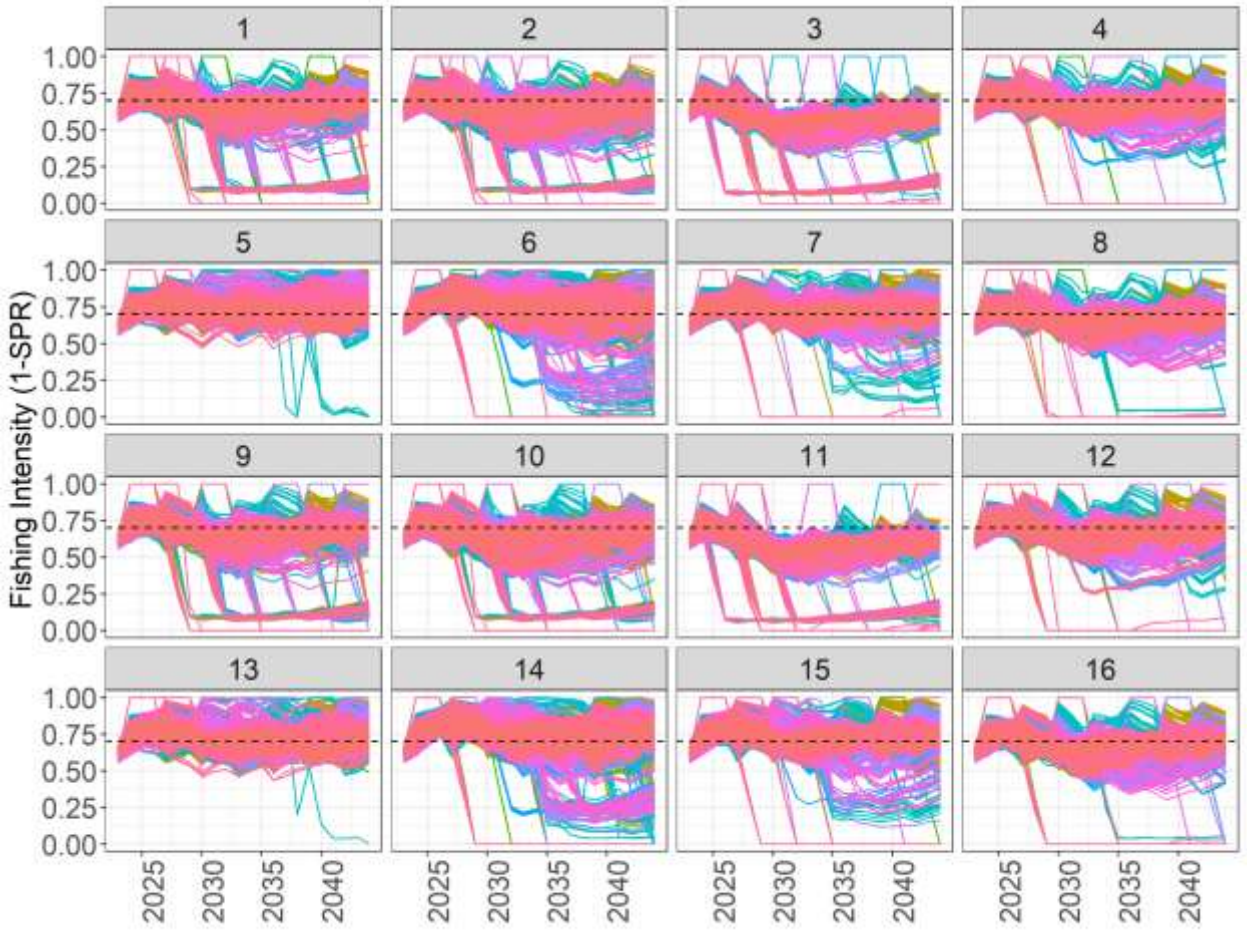

Figure 5C. Worm plots of estimated fishing intensity (F, 1-SPR) from the estimation model (EM, i.e. simulated stock assessment) for individual runs and all reference scenarios. Each panel presents the results for the labeled HCR. Trajectories represent the 100 iterations differing in simulated random recruitment deviates for each reference scenario. The dotted line represents an F of FSPR30%, 0.70. Each estimated F trajectory was compiled using the estimated terminal F from the first EM in 2023 and the estimated SSB for the last three years of each of the other 7 EMs.

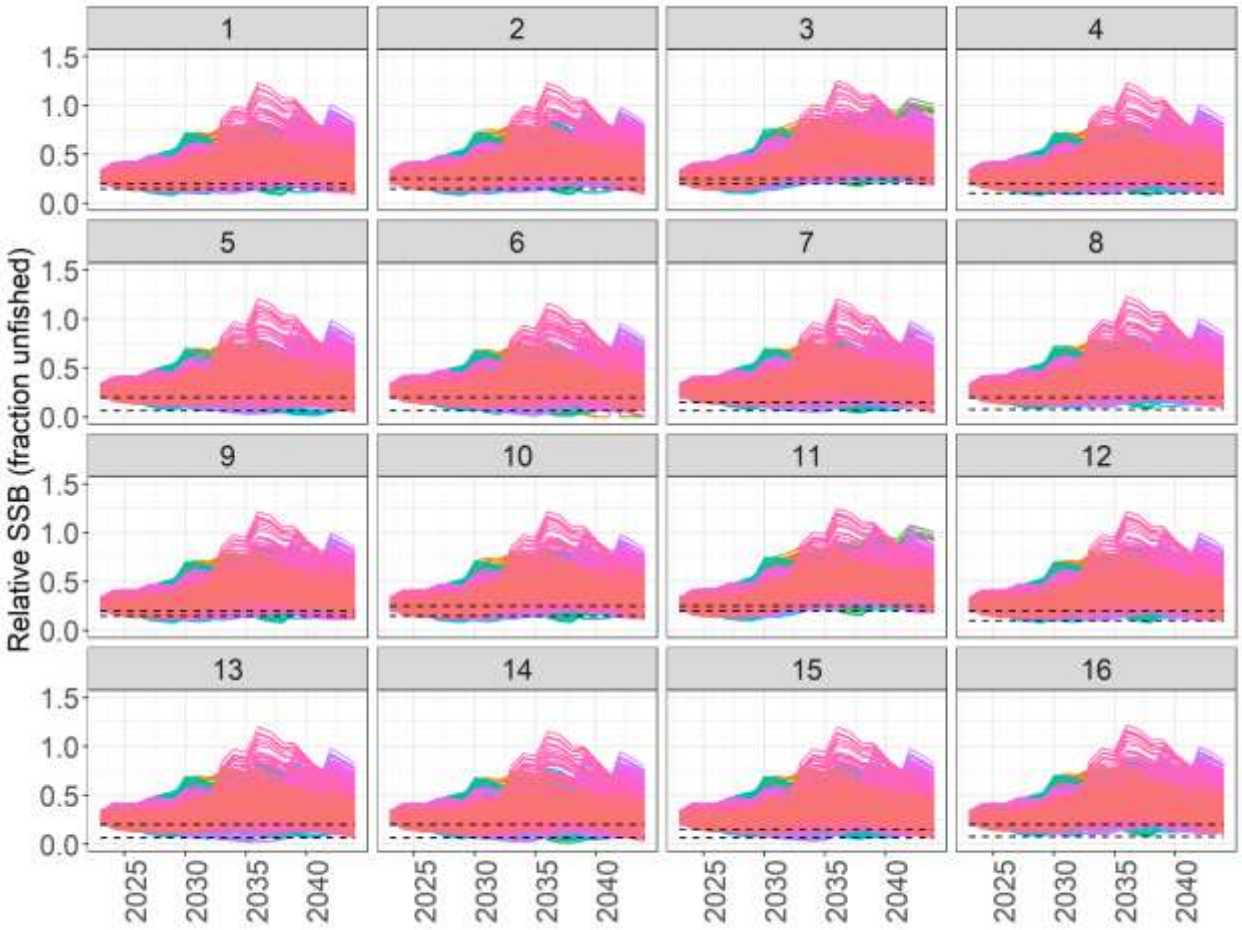

Figure 5D. Worm plots of estimated relative spawning stock biomass ($SSB/SSB_{F=0}$) from the estimation model (EM, i.e. simulated stock assessment) for individual runs for each harvest control rule (HCR) and all reference scenarios. Each panel presents the results for the labeled HCR. Trajectories represent separate iterations differing in simulated random recruitment deviates for each OM. Each estimated SSB trajectory is compiled using the estimated terminal SSB from the first EM in 2023 and the estimated SSB for the last three years of each of the other 7 EMs. The lowest dotted line represents the limit reference point (LRP) for each HCR and the highest the threshold reference point (ThRP). Note that HCRs 5, 6, 7, 13, 14, and 15 don't use the LRP as a control point. Trajectories for only the 81 iterations without estimation issues are presented.

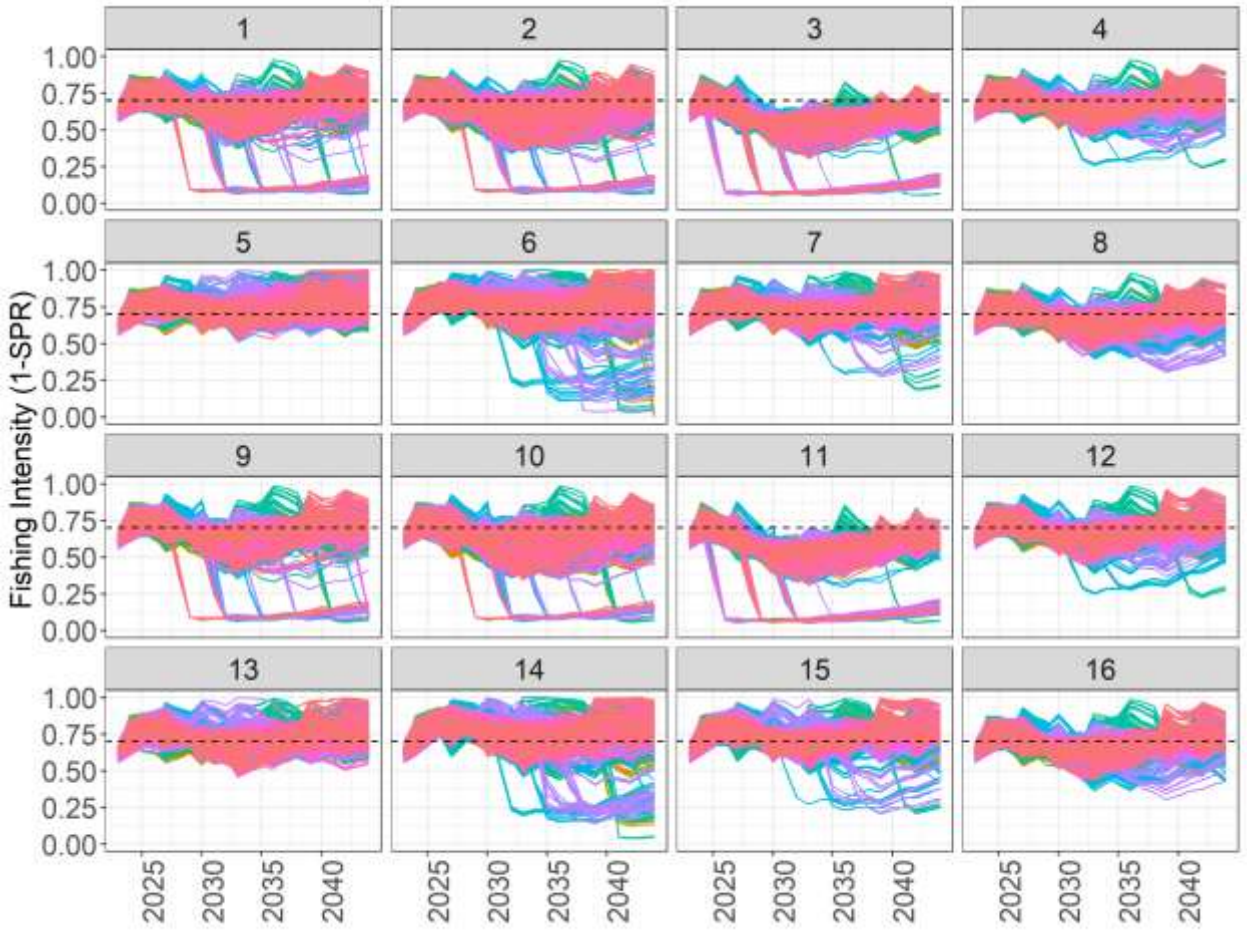

Figure 5E. Worm plots of estimated fishing intensity (F, 1-SPR) from the estimation model (EM, i.e. simulated stock assessment) for individual runs for each harvest control rule (HCR) and all reference scenarios. Each panel presents the results for the labeled HCR. Trajectories represent separate iterations differing in simulated random recruitment deviates. The dotted line represents an F of FSPR30%, 0.70. Trajectories for only the 81 iterations without estimation issues are presented.

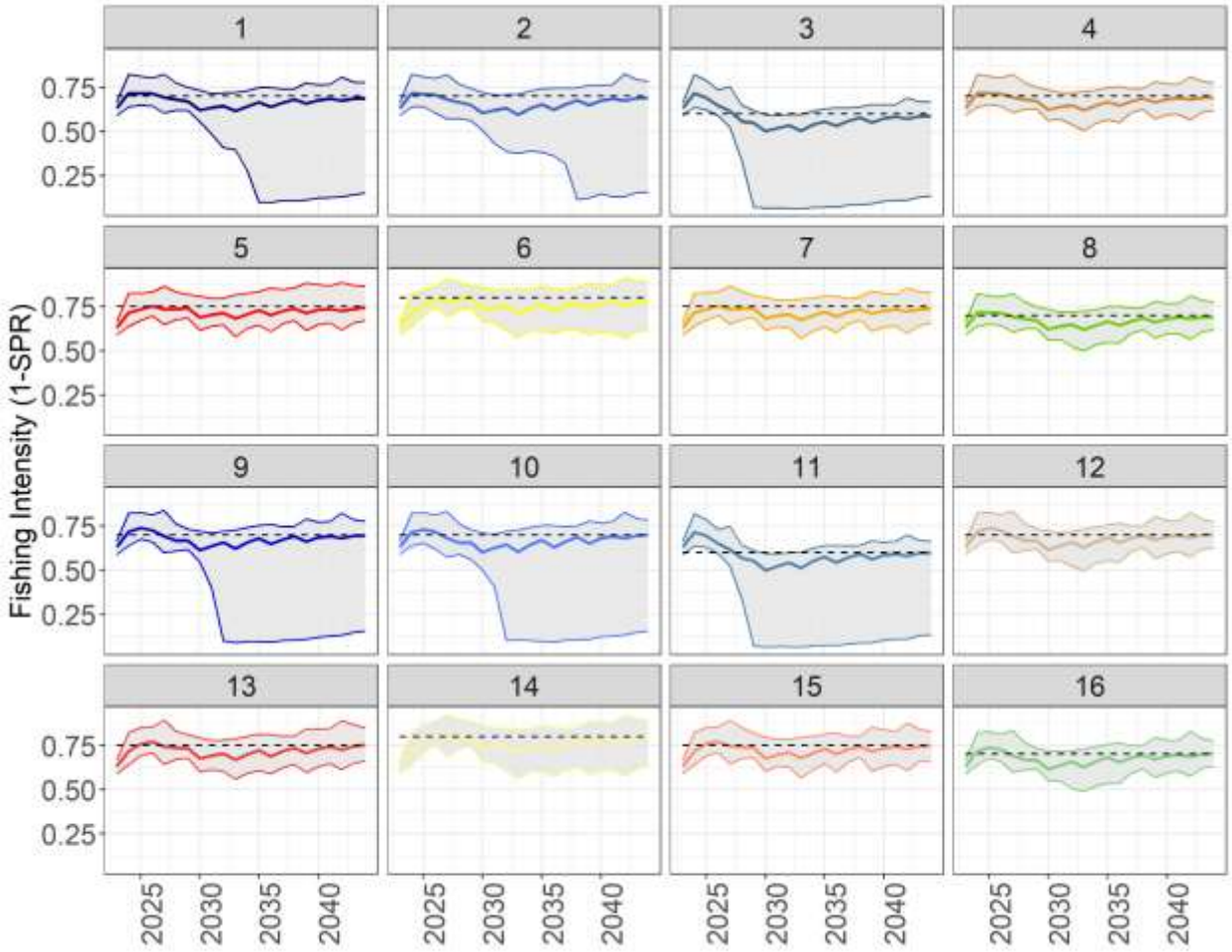

Figure 5F. Trends in median estimated fishing intensity (F, 1-SPR, solid color line) from the estimation model (EM, i.e. simulated stock assessment) under all iterations and reference scenarios by harvest control rule (HCR). The grey shading represents trends in the 5th to 95th quantiles of F. The dotted line represents the F_{TARGET} associated with each HCR.

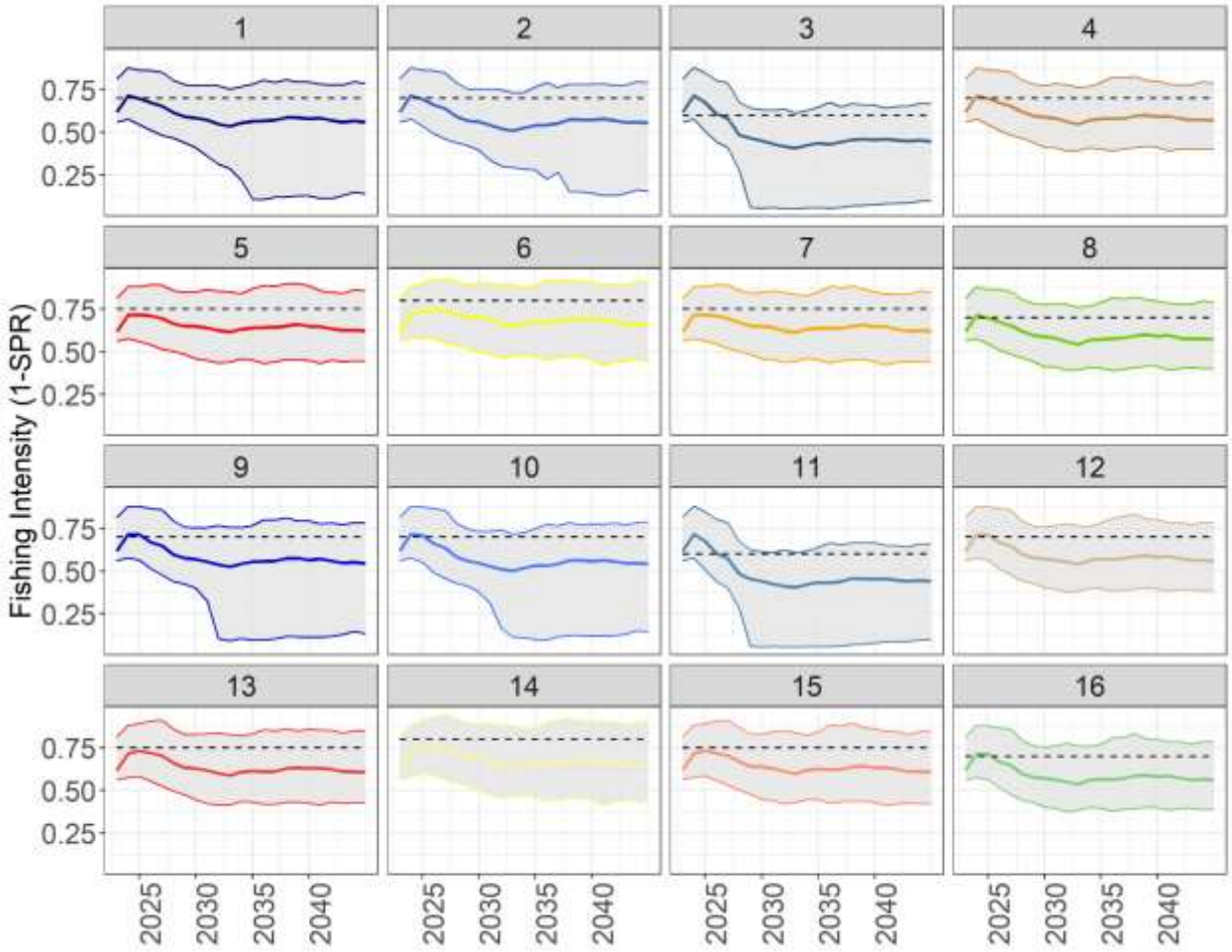

Figure 5G. Trends in median fishing intensity (F, 1-SPR, solid color line) from the operating model under all iterations and reference scenarios by harvest control rule (HCR). The grey shading represents trends in the 5th to 95th quantiles of F. The dotted line represents the F_{TARGET} associated with each HCR.

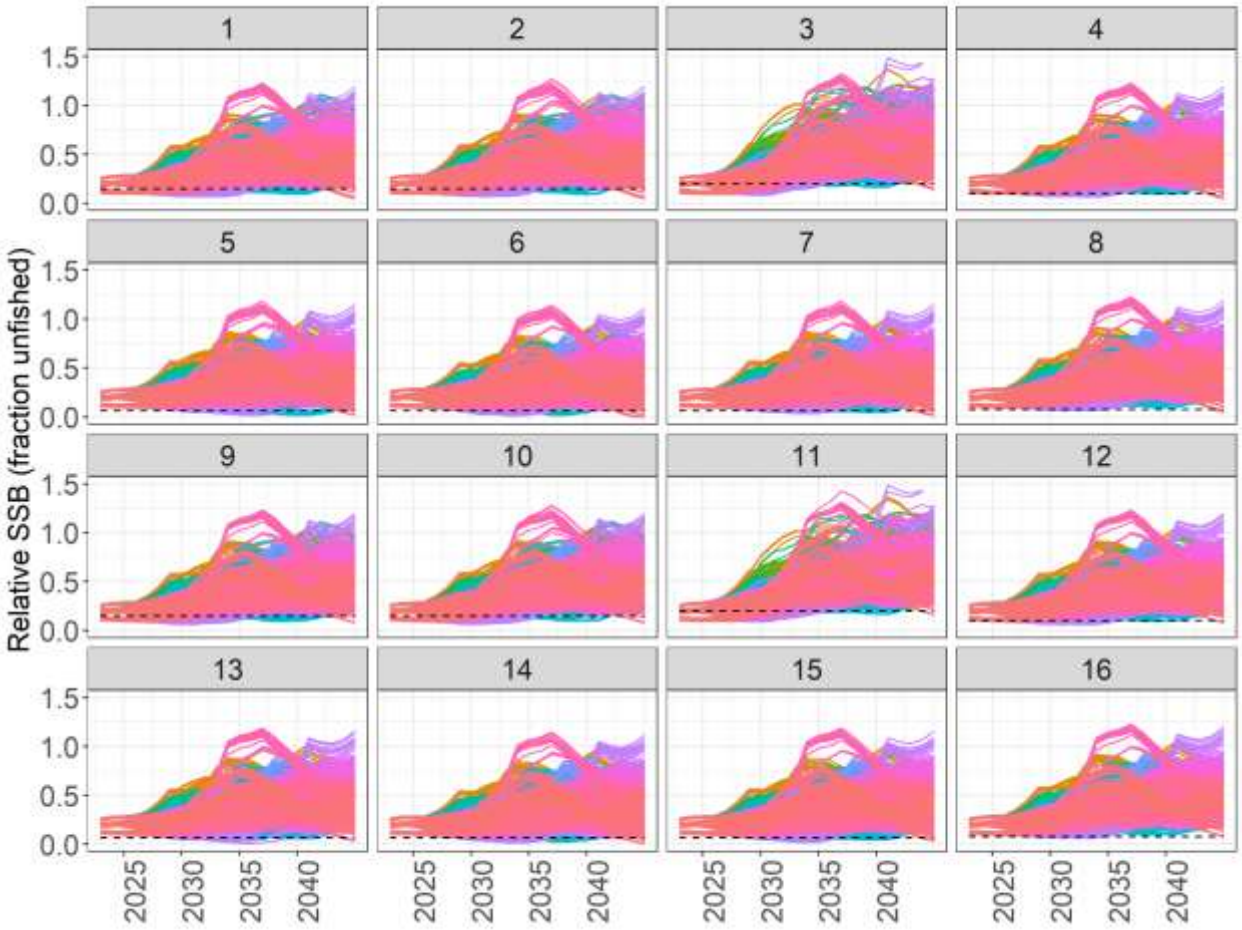

Figure 5H. Worm plots of relative spawning stock biomass ($SSB/SSB_{F=0}$) from the operating model for individual runs for each harvest control rule (HCR) and all reference scenarios. Each panel presents the results for the labeled HCR. Trajectories represent separate iterations differing in simulated random recruitment deviates. The lowest dotted line represents the limit reference point (LRP) for each HCR and the highest the threshold reference point (ThRP). Note that HCRs 5, 6, 7, 13, 14, and 15 don't use the LRP as a control point.

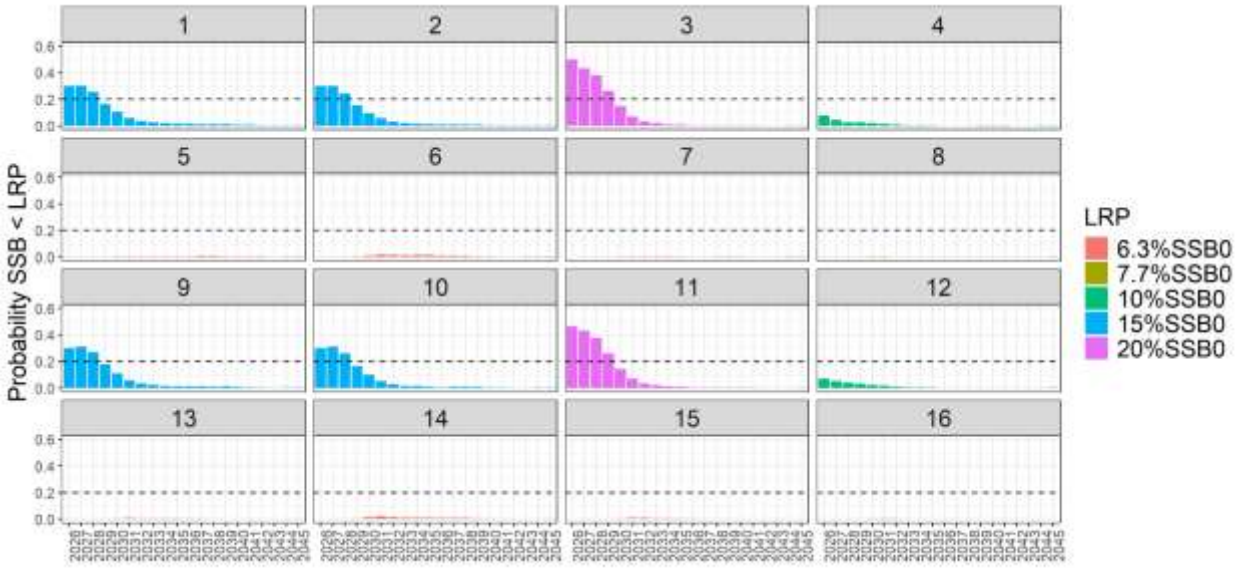

Figure 5I. Plot of the first safety performance metric, the probability, for each harvest control rule (HCR), of spawning stock biomass (SSB) being less than the limit reference point (LRP) specified by each harvest control rule (HCR) across all reference scenarios, and iterations for each simulation year. Colors represent the limit reference point (LRP) associated with each HCR. Each panel presents the results for the labeled HCR.

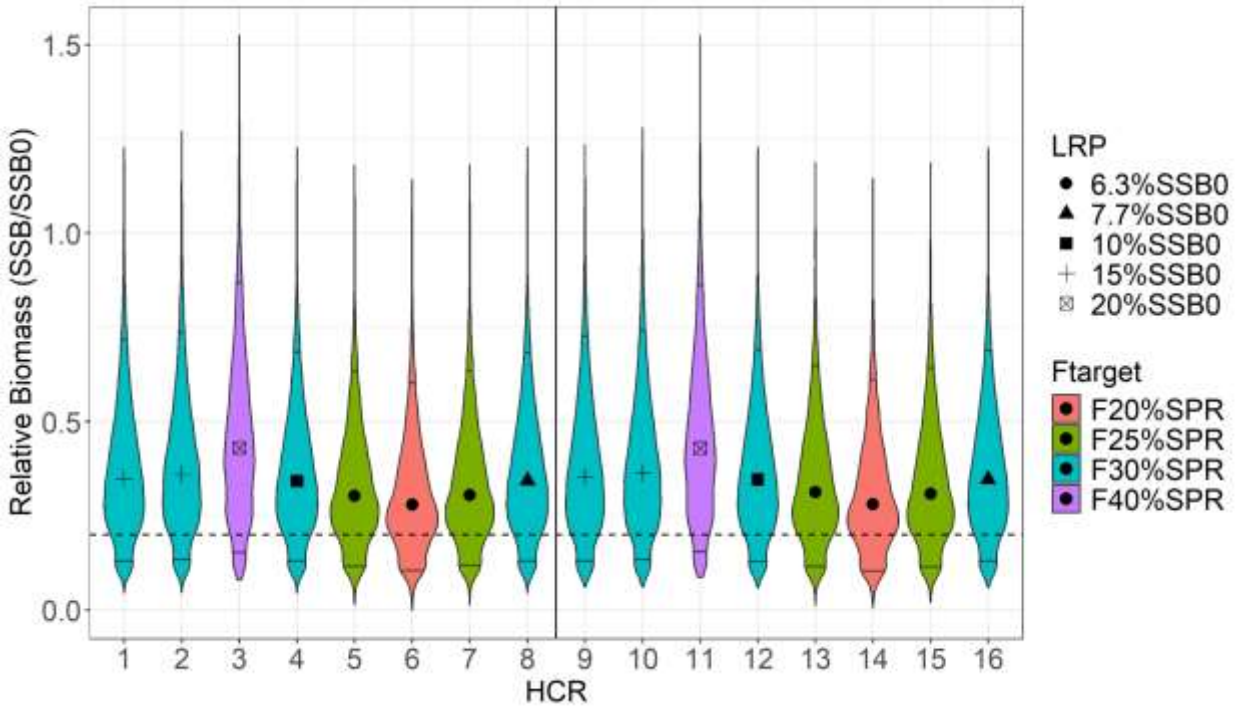

Figure 5J. Violin plot showing the probability density of relative spawning stock biomass ($SSB/SSB_{F=0}$) from the operating model for each harvest control rule (HCR) across all iterations, reference scenarios, and simulation years. The marker inside each violin plot is the median relative SSB and horizontal solid lines within each violin represent the 5th to 95th quantile range. The shape of each marker represents the limit reference point (LRP). Colors represent the F_{TARGET} reference point associated with each HCR. The horizontal dotted line represents the second rebuilding target of $20\%SSB_{F=0}$. The vertical solid line separates HCRs 9 to 16, which are tuned to an EPO:WCPO impact ratio of 30:70, but are otherwise the same as HCRs 1 to 8. EPO stands for Eastern Pacific Ocean and WCPO for Western Central Pacific Ocean.

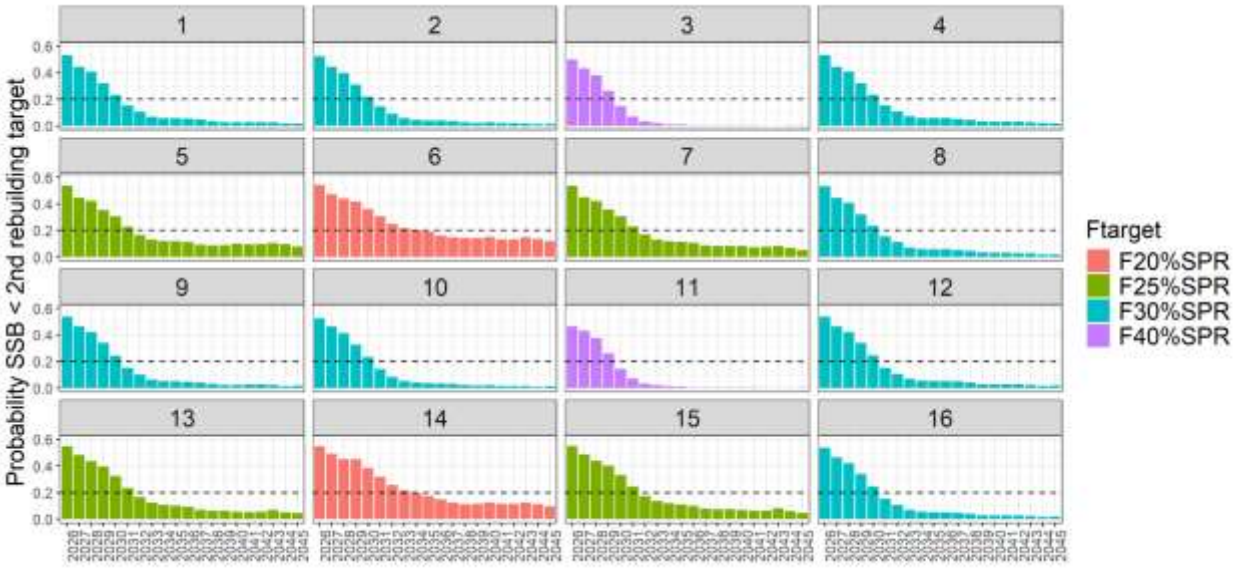

Figure 5K. Plot of the second safety performance metric, the probability, for each harvest control rule (HCR), of spawning stock biomass (SSB) being less than 20%$SSB_{F=0}$ across all reference scenarios, and iterations for each simulation year. Each panel presents the results for the labeled HCR. Colors represent the F_{TARGET} reference point associated with each HCR.

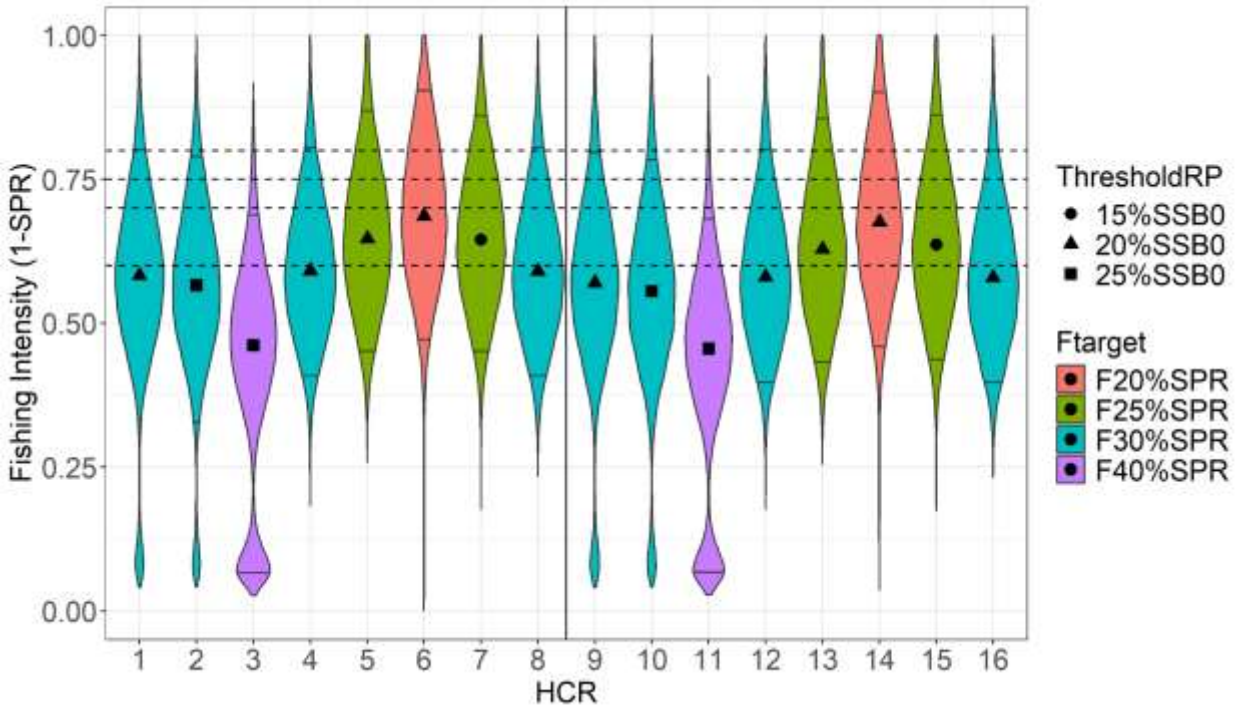

Figure 5L. Violin plot showing the probability density of fishing mortality (F, 1-SPR) from the operating model for each harvest control rule (HCR) across all iterations, reference scenarios, and simulation years. The marker inside each violin plot is the median F and horizontal solid lines within each violin represent the 5th to 95th quantile range. The shape of each marker represents the threshold reference point (ThresholdRP). Colors represent the F_{TARGET} reference point associated with each HCR. The horizontal dotted lines represent the different F_{TARGET}s in 1-SPR associated with different HCRs. The vertical solid line separates HCRs 9 to 16, which are tuned to an EPO:WCPO impact ratio of 30:70, but are otherwise the same as HCRs 1 to 8. EPO stands for Eastern Pacific Ocean and WCPO for Western Central Pacific Ocean.

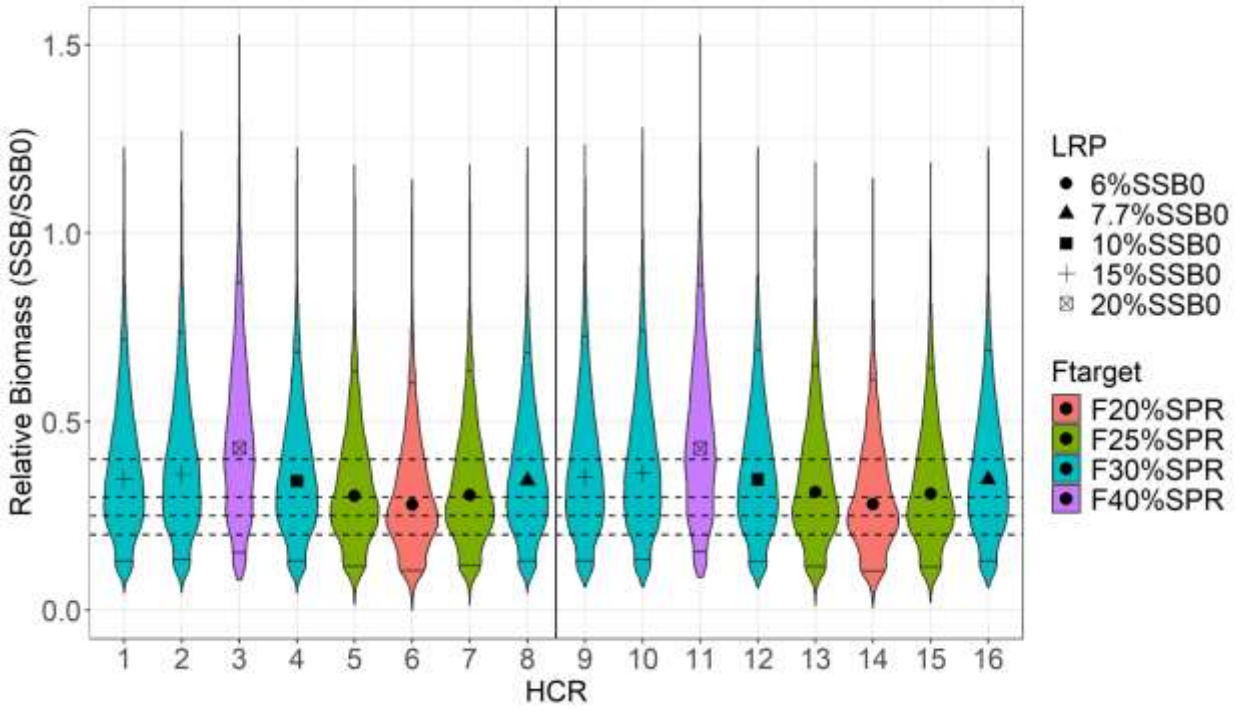

Figure 5M. Violin plot showing the probability density of relative spawning stock biomass ($SSB/SSB_{F=0}$) from the operating model for each harvest control rule (HCR) across all iterations, reference scenarios, and simulation years. The marker inside each violin plot is the median relative SSB and horizontal solid lines within each violin represent the 5th to 95th quantile range. The shape of each marker represents the limit reference point (LRP). Colors represent the F_{TARGET} reference point associated with each HCR. The horizontal dotted lines represent the relative SSB associated with each different F_{TARGET}. The vertical solid line separates HCRs 9 to 16, which are tuned to an EPO:WCPO impact ratio of 30:70, but are otherwise the same as HCRs 1 to 8. EPO stands for Eastern Pacific Ocean and WCPO for Western Central Pacific Ocean.

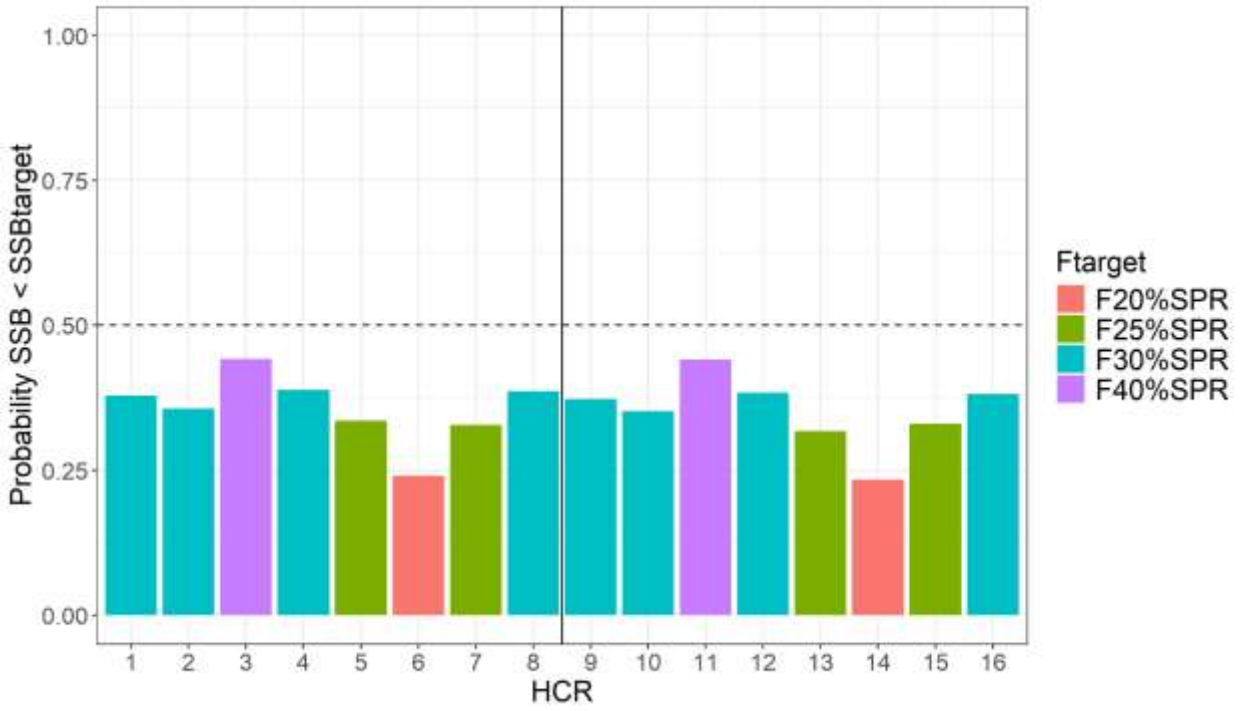

Figure 5N. Plot of the second status performance metric, the probability, for each harvest control rule (HCR), of spawning stock biomass (SSB) being below the equivalent biomass depletion level associated with the candidates for F_{TARGET} across all reference scenarios, iterations, and simulation years. The horizontal dotted line represents a 50% probability. Colors represent the F_{TARGET} reference point associated with each HCR. The vertical solid line separates HCRs 9 to 16, which are tuned to an EPO:WCPO impact ratio of 30:70, but are otherwise the same as HCRs 1 to 8. EPO stands for Eastern Pacific Ocean and WCPO for Western Central Pacific Ocean.

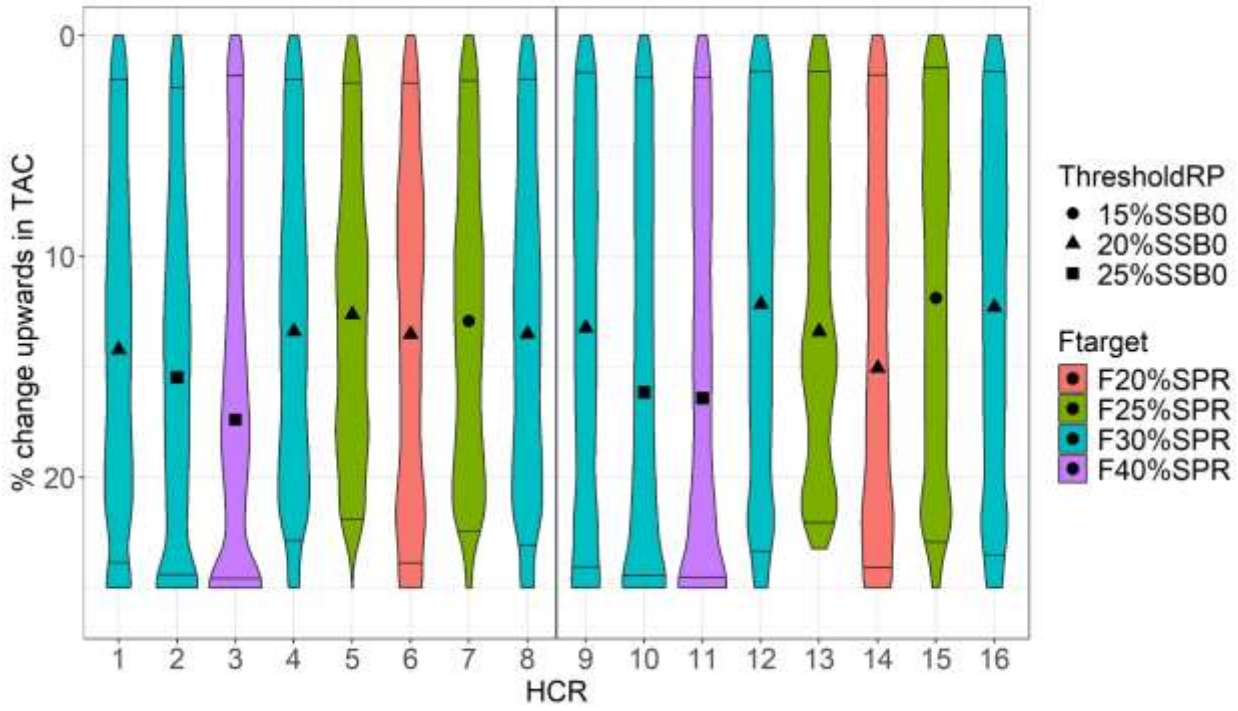

Figure 5O. Violin plot showing the probability density of upward changes in TAC between management periods when SSB≥LRP for each harvest control rule (HCR) across all iterations, reference scenarios, and simulation years. The marker inside each violin plot is the median upward change in TAC and horizontal solid lines within each violin represent the 5th to 95th quantile range. The shape of each marker represents the threshold reference point (ThresholdRP). Colors represent the F_{TARGET} reference point associated with each HCR. The vertical solid line separates HCRs 9 to 16, which are tuned to an EPO:WCPO impact ratio of 30:70, but are otherwise the same as HCRs 1 to 8. EPO stands for Eastern Pacific Ocean and WCPO for Western Central Pacific Ocean.

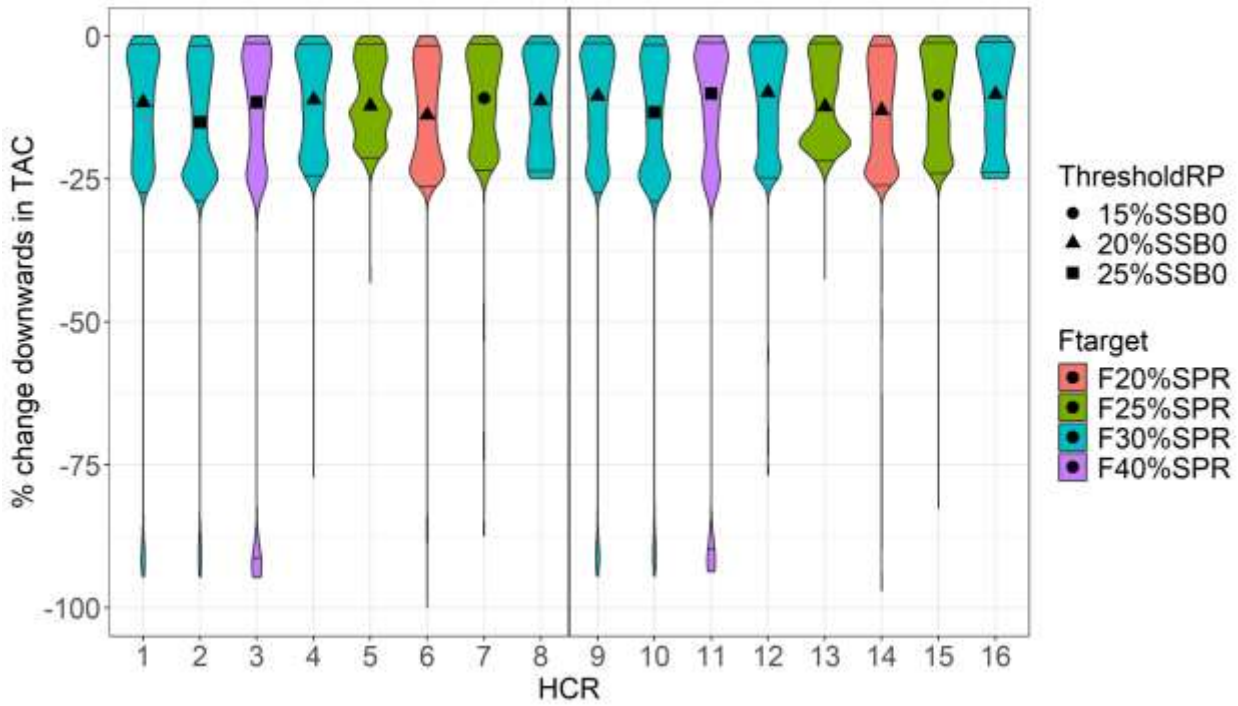

Figure 5P. Violin plot showing the probability density of downward changes in TAC between management periods for all simulation years for each harvest control rule (HCR) across all iterations and reference scenarios. The marker inside each violin plot is the median downward change in TAC and horizontal solid lines within each violin represent the 5th to 95th quantile range. The shape of each marker represents the threshold reference point (ThresholdRP). Colors represent the F_{TARGET} reference point associated with each HCR. The vertical solid line separates HCRs 9 to 16, which are tuned to an EPO:WCPO impact ratio of 30:70, but are otherwise the same as HCRs 1 to 8. EPO stands for Eastern Pacific Ocean and WCPO for Western Central Pacific Ocean.

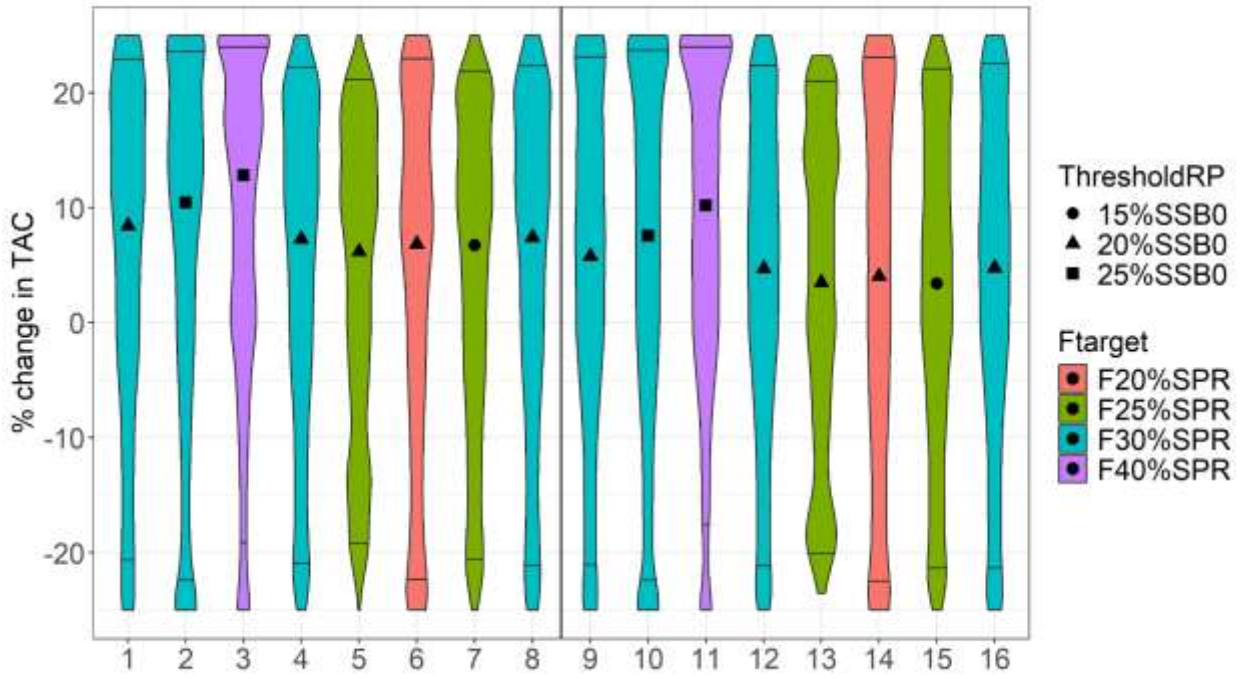

Figure 5Q. Violin plot showing the probability density of changes in TAC between management periods when SSB≥LRP for each harvest control rule (HCR) across all iterations, reference scenarios, and simulation years. The marker inside each violin plot is the median change in TAC and horizontal solid lines within each violin represent the 5th to 95th quantile range. The shape of each marker represents the threshold reference point (ThresholdRP). Colors represent the F_{TARGET} reference point associated with each HCR. The vertical solid line separates HCRs 9 to 16, which are tuned to an EPO:WCPO impact ratio of 30:70, but are otherwise the same as HCRs 1 to 8. EPO stands for Eastern Pacific Ocean and WCPO for Western Central Pacific Ocean.

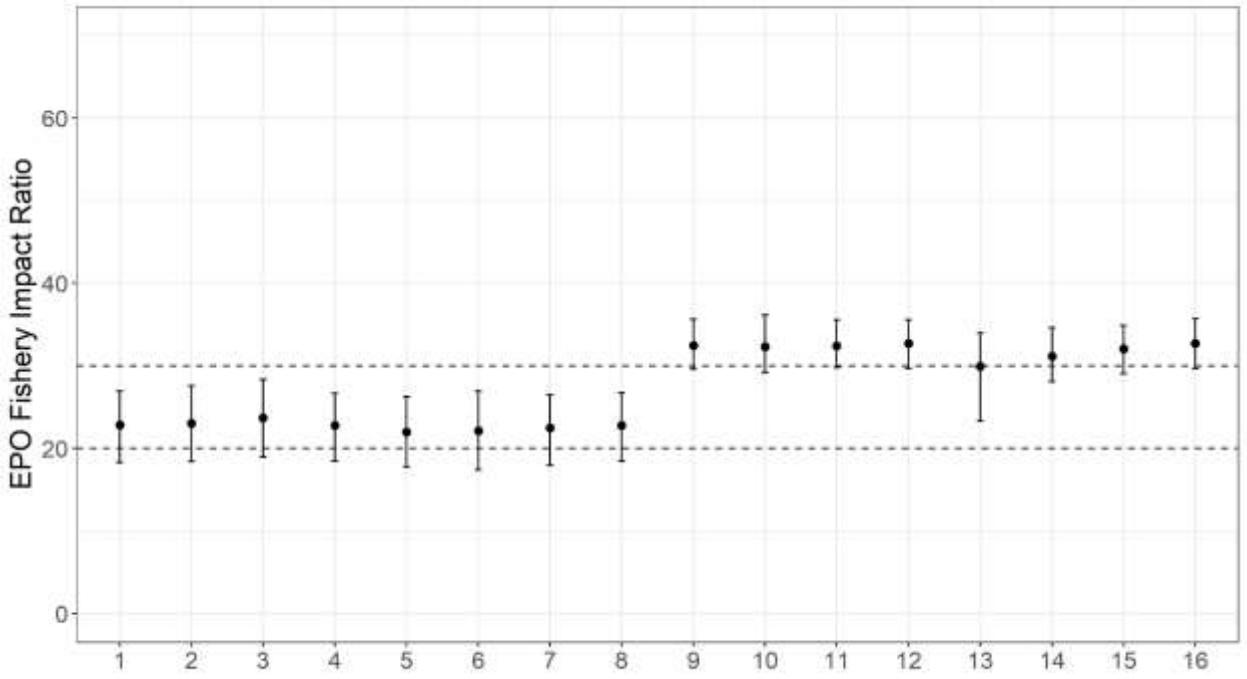

Figure 5R. Median (dot) and 5th to 95th quantile range (error bars) of fishery impact on spawning stock biomass for the Eastern Pacific Ocean (EPO) fleet segment in the terminal year of the MSE simulation across all iterations and reference OMs for each HCR. The impact of the Western and Central Pacific Ocean fleet segment is 100 – the EPO impact.

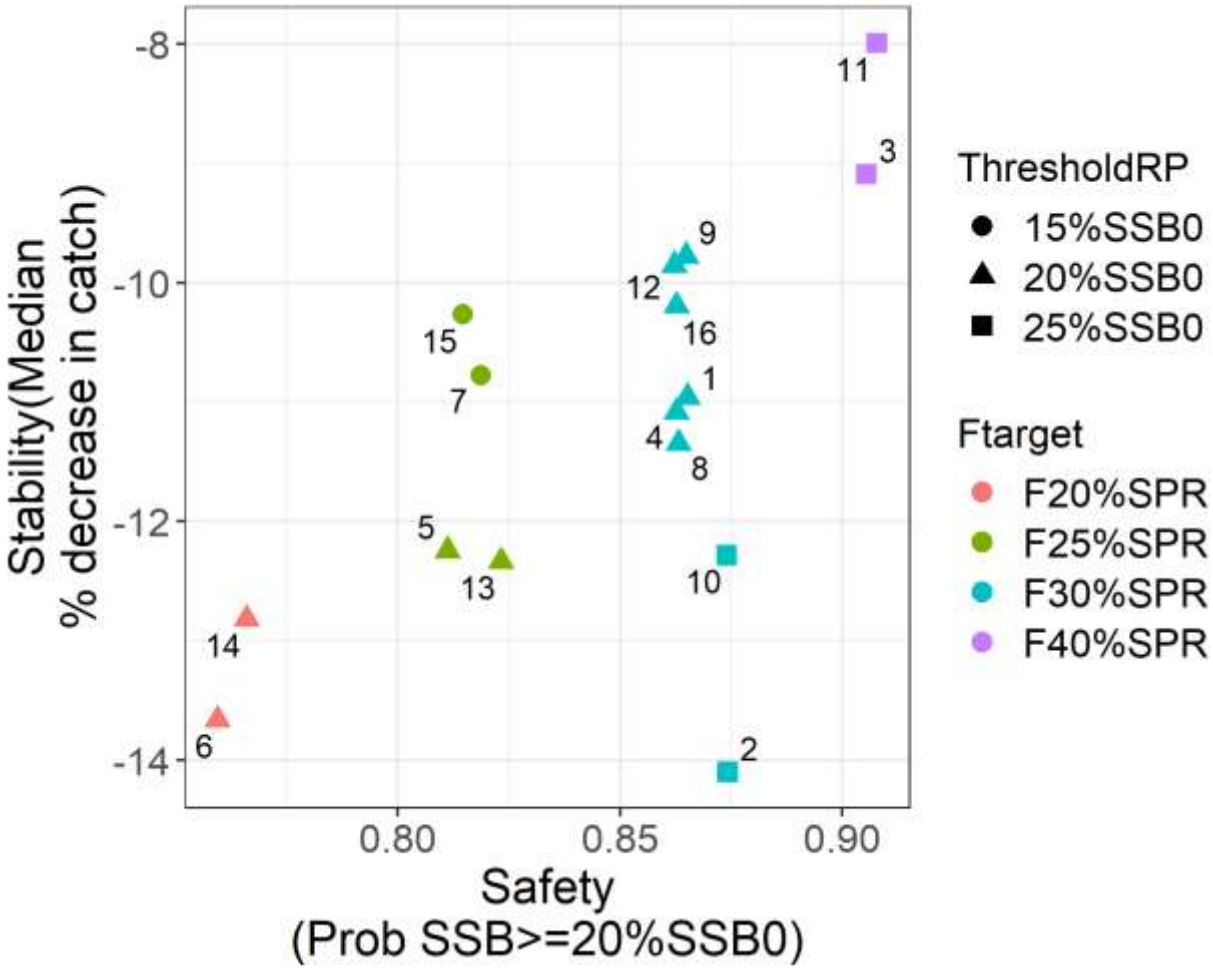

Figure 5S. Median decrease in catch between management periods versus the probability of spawning stock biomass (SSB) being at or above the second rebuilding target of $20\%SSB_{F=0}$. Note that to ensure that for both measures a higher value is better, here we reversed the second performance metric shown in Fig. ES3 to be the probability of $SSB \geq 20\%SSB_{F=0}$ instead of the probability of $SSB < 20\%SSB_{F=0}$. Each HCR is labeled and represented by a symbol colored according to their F_{TARGET}. The symbol shape represents the ThresholdRP is the first control point for each HCR and stands for threshold reference point.

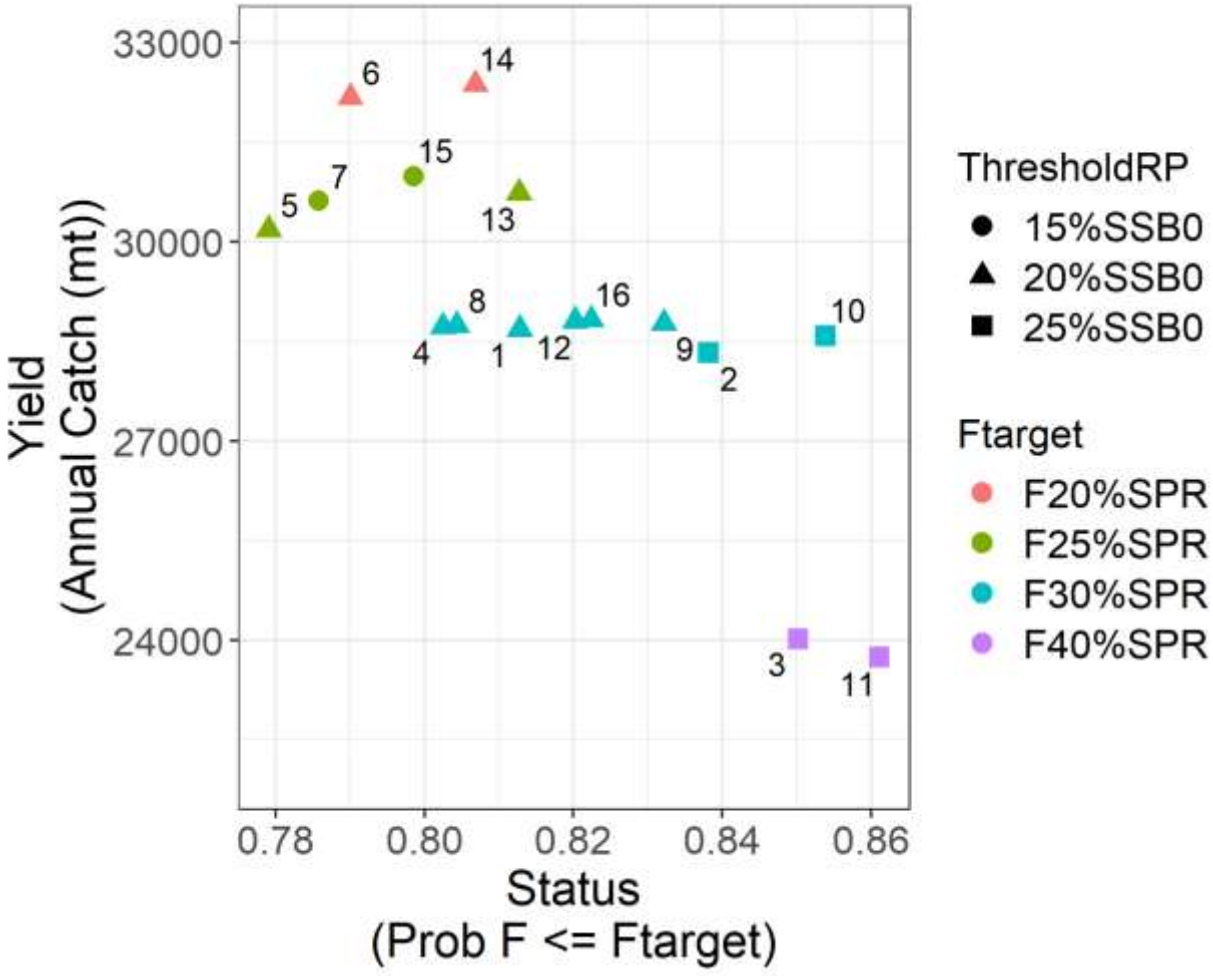

Figure 5T. Median annual total catch versus the probability of fishing morality (F) being at or below the F_{TARGET}. Each HCR is labeled and represented by a symbol colored according to their F_{TARGET}. The symbol shape represents the ThresholdRP is the first control point for each HCR and stands for threshold reference point.

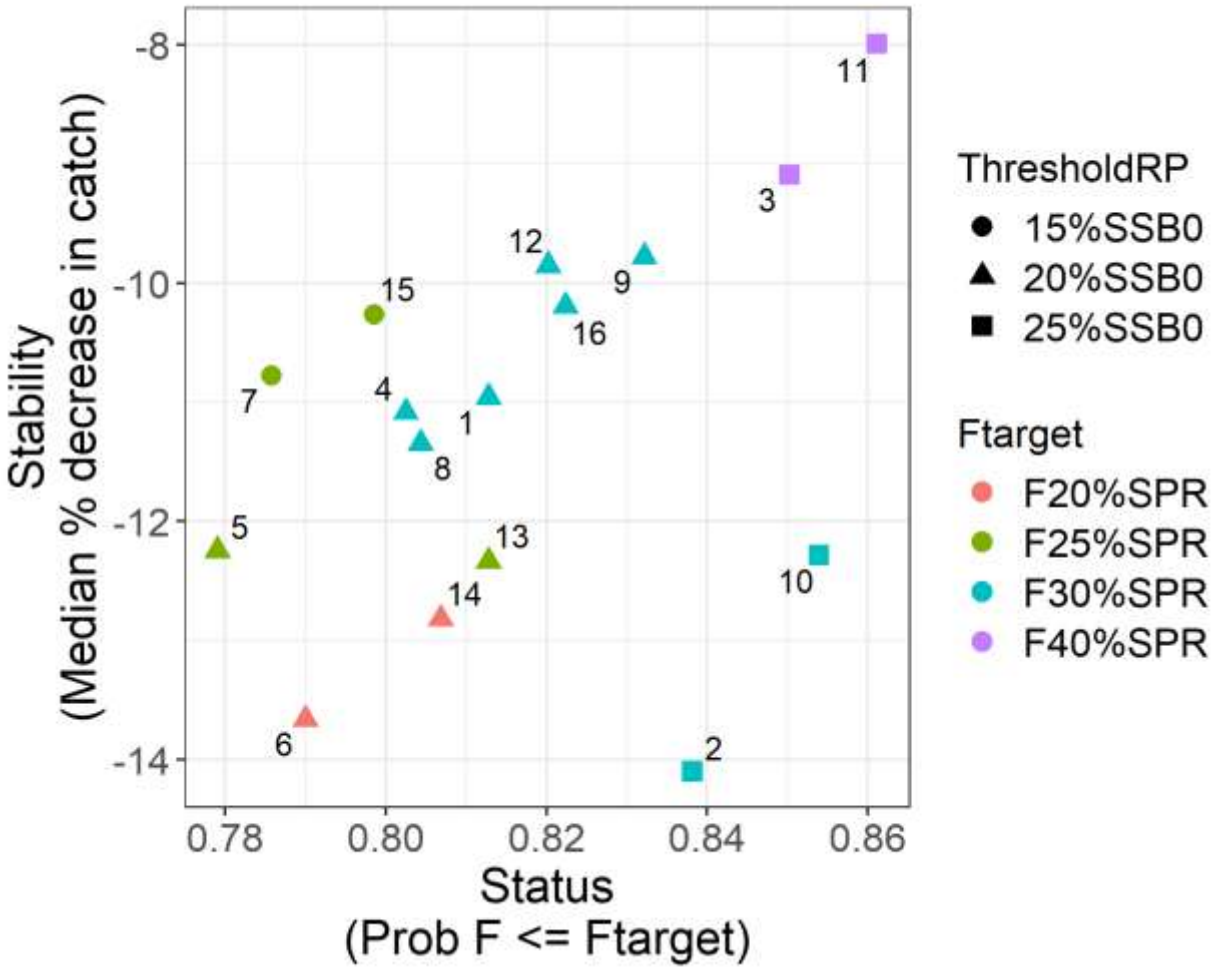

Figure 5U. Median decrease in catch between management periods versus the probability of fishing morality (F) being at or below the F_{TARGET}. Each HCR is labeled and represented by a symbol colored according to their F_{TARGET}. The symbol shape represents the ThresholdRP is the first control point for each HCR and stands for threshold reference point.

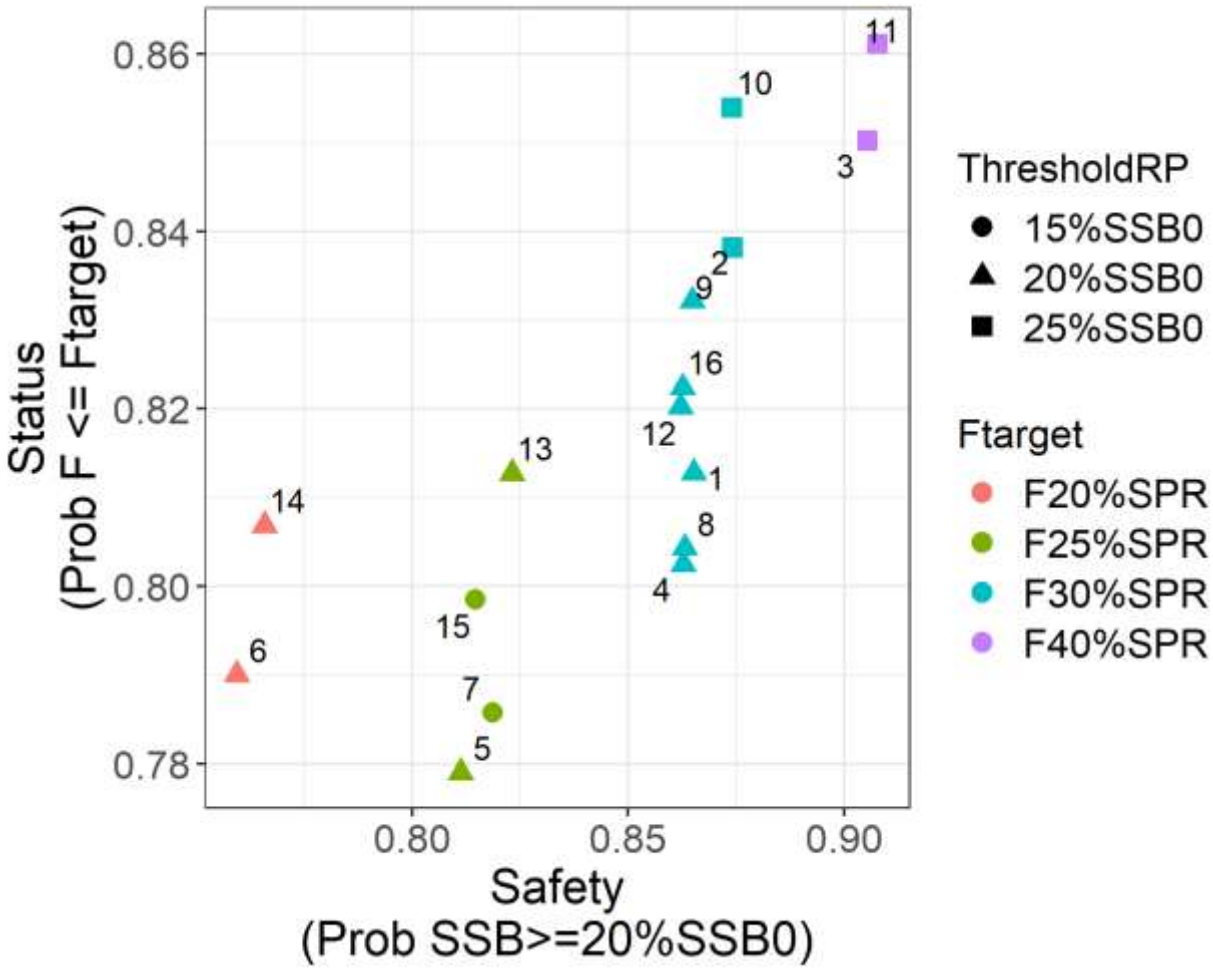

Figure 5V. Probability of spawning stock biomass (SSB) being at or above the second rebuilding target of $20\%SSB_{F=0}$ versus the probability of fishing morality (F) being at or below the F_{TARGET}. Note that to ensure that for both measures a higher value is better, here we reversed the second performance metric shown in Fig. ES3 to be the probability of $SSB \geq 20\%SSB_{F=0}$ instead of the probability of $SSB < 20\%SSB_{F=0}$. Each HCR is labeled and represented by a symbol colored according to their F_{TARGET}. The symbol shape represents the ThresholdRP is the first control point for each HCR and stands for threshold reference point.

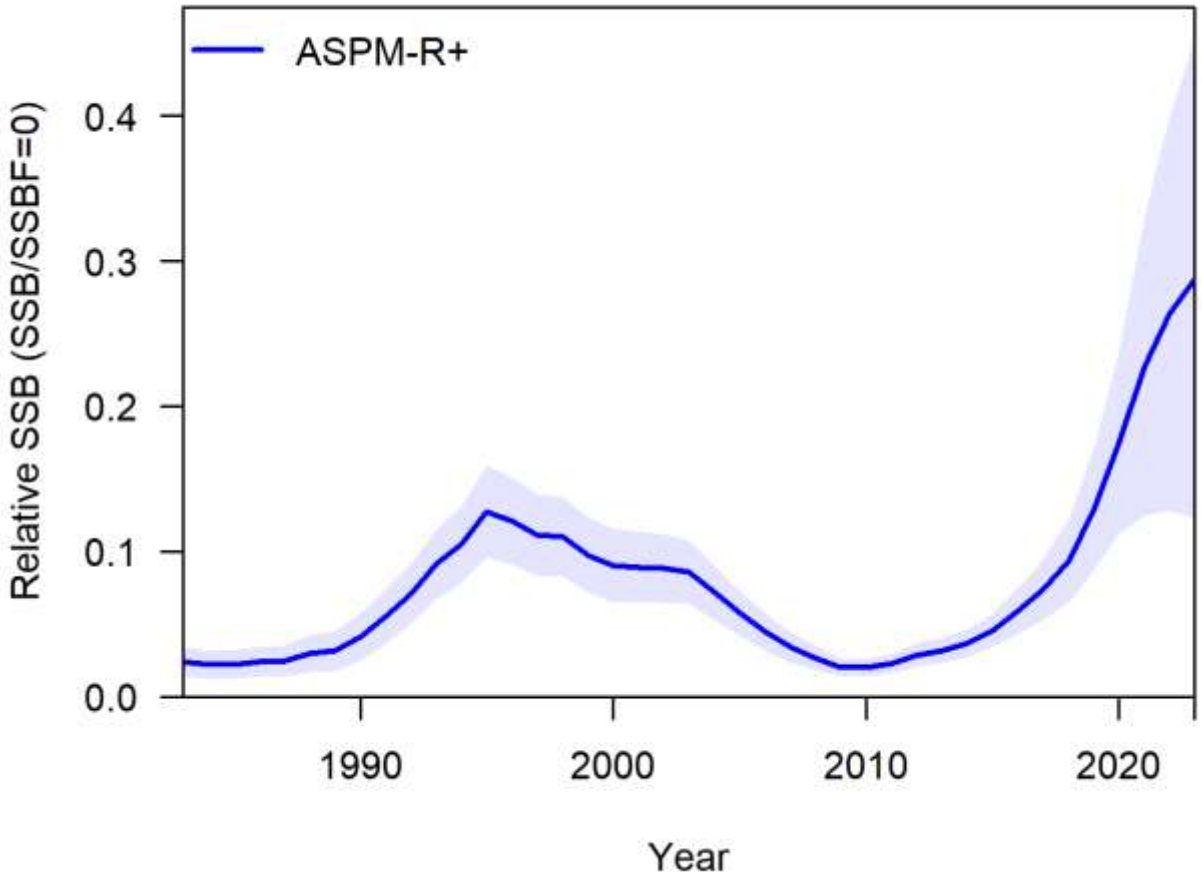

Figure 5W. Trajectory of relative spawning stock biomass ($SSB/SSB_{F=0}$) of Pacific bluefin tuna *Thunnus orientalis* (1983-2023) estimated from the ASPM-R+ model used to calculate the potential TAC for 2026-2028. The solid line is the point estimate, and shaded area delineates the 90% confidence interval from the normal approximation of the Hessian matrix.

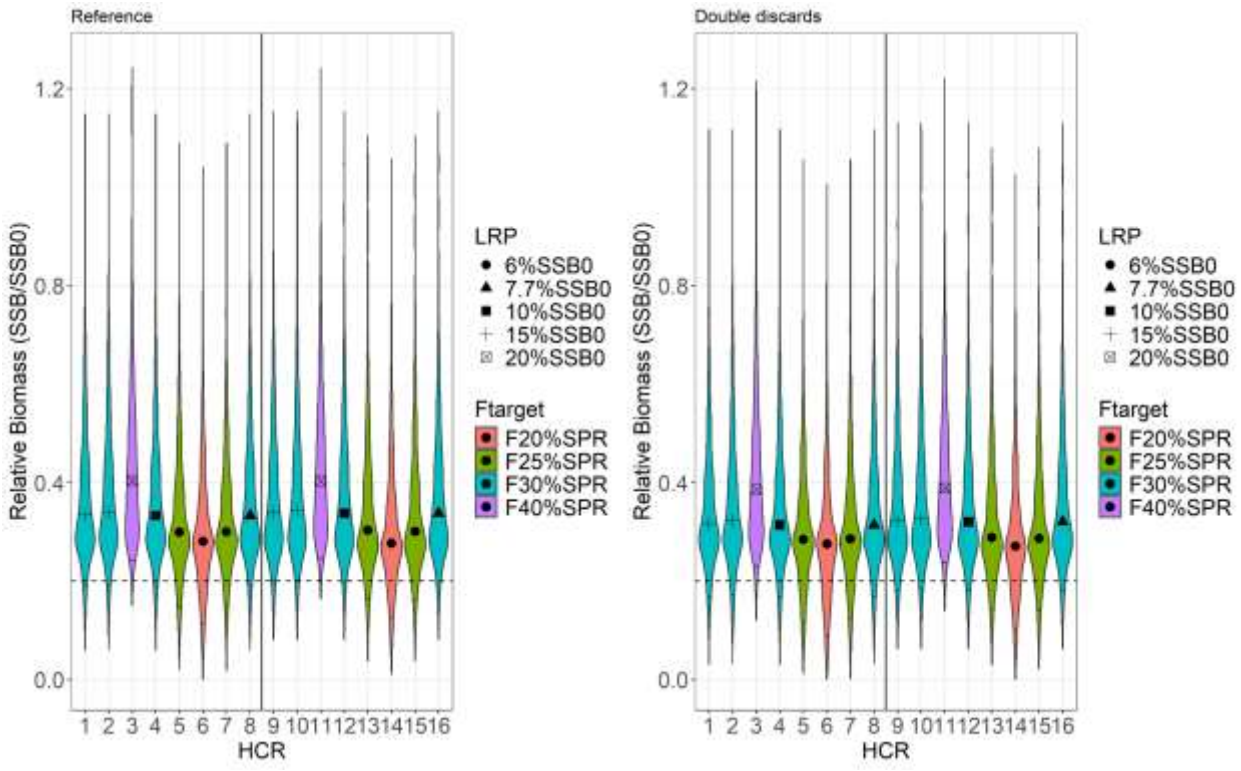

Figure 5Z1. Violin plot showing the probability density of relative spawning stock biomass (SSB/ $SSB_{F=0}$) from operating model 1 for each harvest control rule (HCR) across all iterations and simulation years for the reference run and the doubling of discards robustness test. The marker inside each violin plot is the median relative SSB and horizontal solid lines within each violin represent the 5th to 95th quantile range. The shape of each marker represents the limit reference point (LRP). Colors represent the F_{TARGET} reference point associated with each HCR. The horizontal dotted line represents the second rebuilding target of $20\%SSB_{F=0}$. The vertical solid line separates HCRs 9 to 16, which are tuned to an EPO:WCPO impact ratio of 30:70, but are otherwise the same as HCRs 1 to 8. EPO stands for Eastern Pacific Ocean and WCPO for Western Central Pacific Ocean.

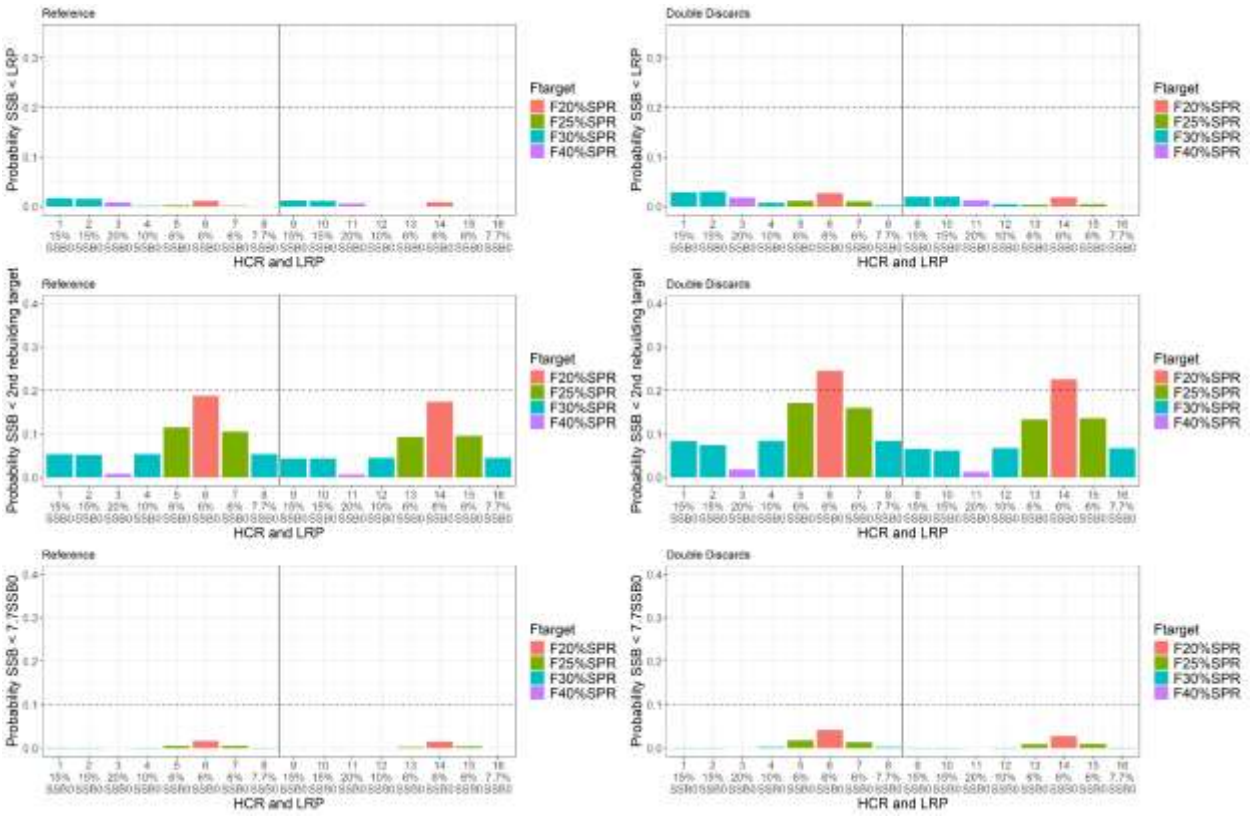

Figure 5Z2. Plot of the safety performance metrics, the probability, for each harvest control rule (HCR), of spawning stock biomass (SSB) being less than the limit reference point (LRP) specified by each harvest control rule (HCR) or 20%$SSB_{F=0}$ or 7.7%$SSB_{F=0}$ for operating model 1 across all iterations and simulation years for the reference run and the doubling of discards robustness test. Colors represent the F_{TARGET} reference point associated with each HCR. The vertical solid line separates HCRs 9 to 16, which are tuned to an EPO:WCPO impact ratio of 30:70, but are otherwise the same as HCRs 1 to 8. EPO stands for Eastern Pacific Ocean and WCPO for Western Central Pacific Ocean.

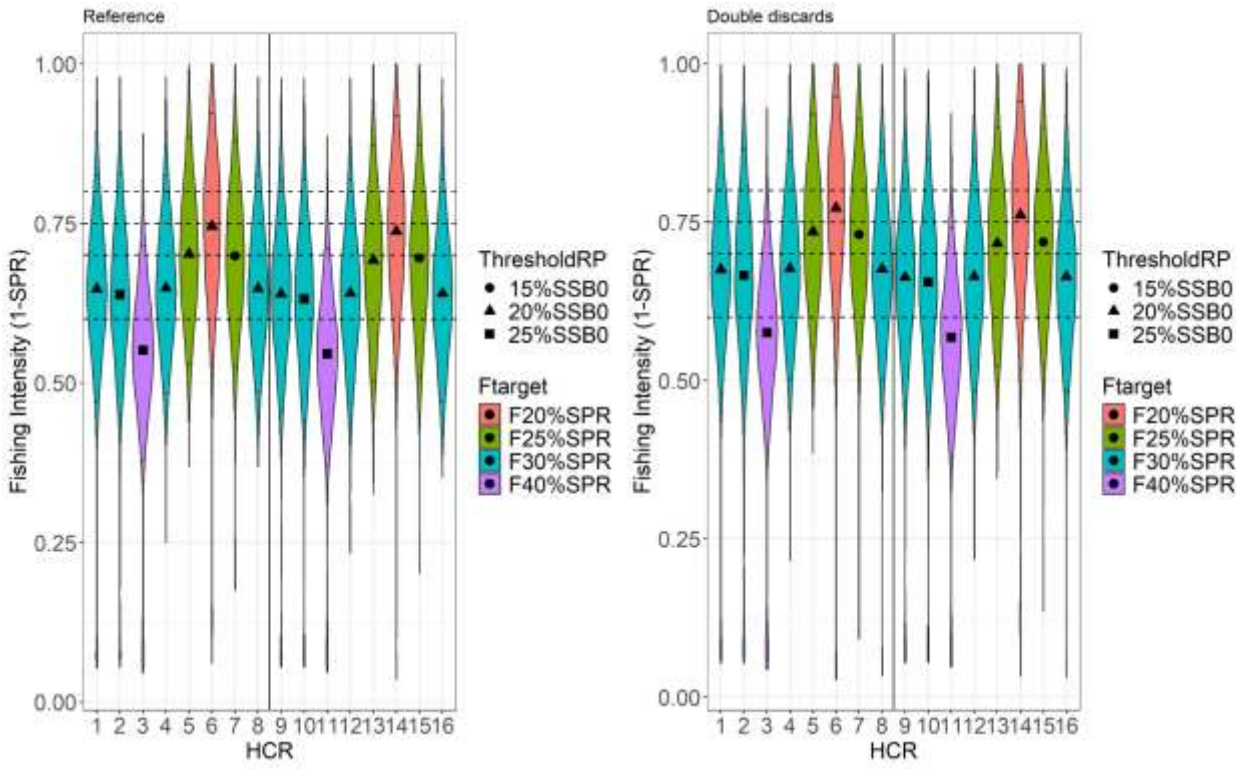

Figure 5Z3. Violin plot showing the probability density of fishing mortality (F, 1-SPR) from operating model 1 for each harvest control rule (HCR) across all iterations and simulation years for the reference run and the doubling of discards robustness test. The marker inside each violin plot is the median F and horizontal solid lines within each violin represent the 5th to 95th quantile range. The shape of each marker represents the threshold reference point (ThresholdRP). Colors represent the F_{TARGET} reference point associated with each HCR. The horizontal dotted lines represent the different F_{TARGET}s in 1-SPR associated with different HCRs. The vertical solid line separates HCRs 9 to 16, which are tuned to an EPO:WCPO impact ratio of 30:70, but are otherwise the same as HCRs 1 to 8. EPO stands for Eastern Pacific Ocean and WCPO for Western Central Pacific Ocean.

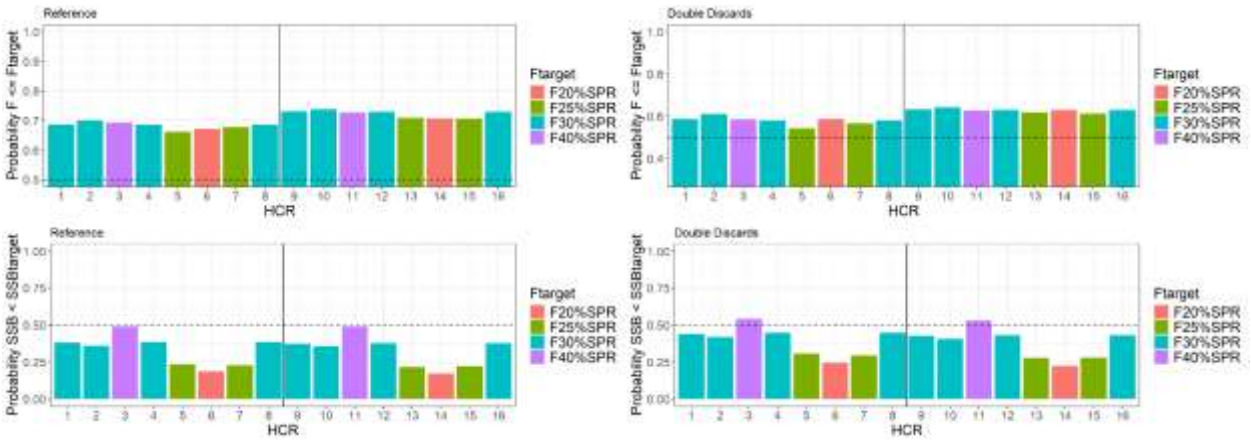

Figure 5Z4. Plot of the status performance metrics, the probability of fishing mortality (F, 1-SPR) being less or equal to the F_{TARGET} and the probability of spawning stock biomass (SSB) being below the equivalent biomass depletion level associated with the candidates for F_{TARGET} for operating model 1 across all iterations and simulation years for the reference run and the doubling of discards robustness test. The horizontal dotted line represents a 50% probability. Colors represent the F_{TARGET} reference point associated with each HCR. The vertical solid line separates HCRs 9 to 16, which are tuned to an EPO:WCPO impact ratio of 30:70, but are otherwise the same as HCRs 1 to 8. EPO stands for Eastern Pacific Ocean and WCPO for Western Central Pacific Ocean.

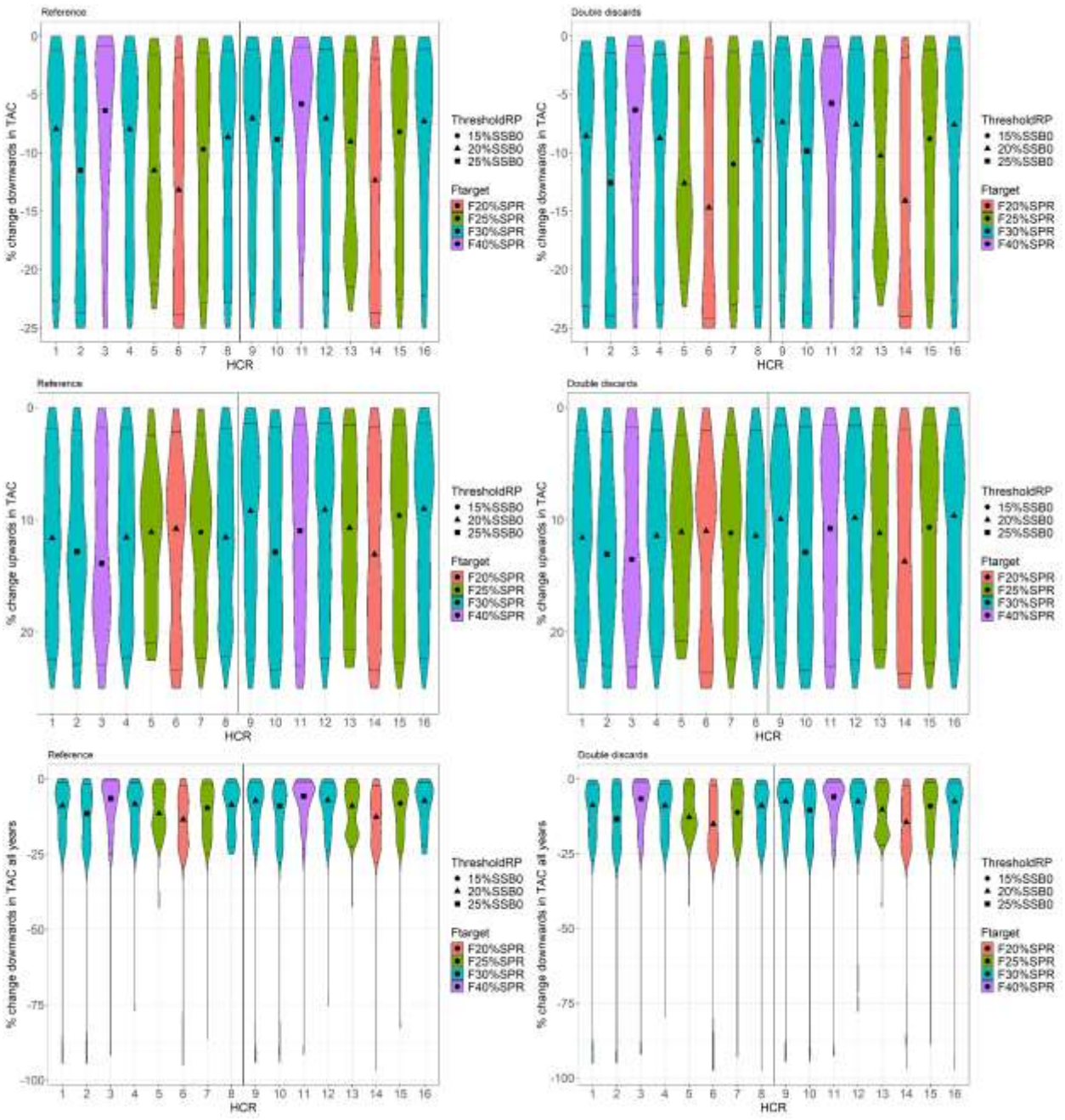

Figure 5Z5. Violin plot showing the probability density of downward changes in TAC (top panel) and of upwards changes in TAC (middle panel) between management period when SSB≥LRP, and of downward changes in TAC across all management periods (bottom panel) for operating model 1 for each harvest control rule (HCR) across all iterations and simulation years for the reference run and the doubling of discards robustness test. The marker inside each violin plot is the median downward change in TAC and horizontal solid lines within each violin represent the 5th to 95th quantile range. The shape of each marker represents the threshold reference point (ThresholdRP). Colors represent the F_{TARGET} reference point associated with each HCR. The vertical solid line separates HCRs 9 to 16, which are tuned to an EPO:WCPO impact ratio of 30:70, but are otherwise

the same as HCRs 1 to 8. EPO stands for Eastern Pacific Ocean and WCPO for Western Central Pacific Ocean.

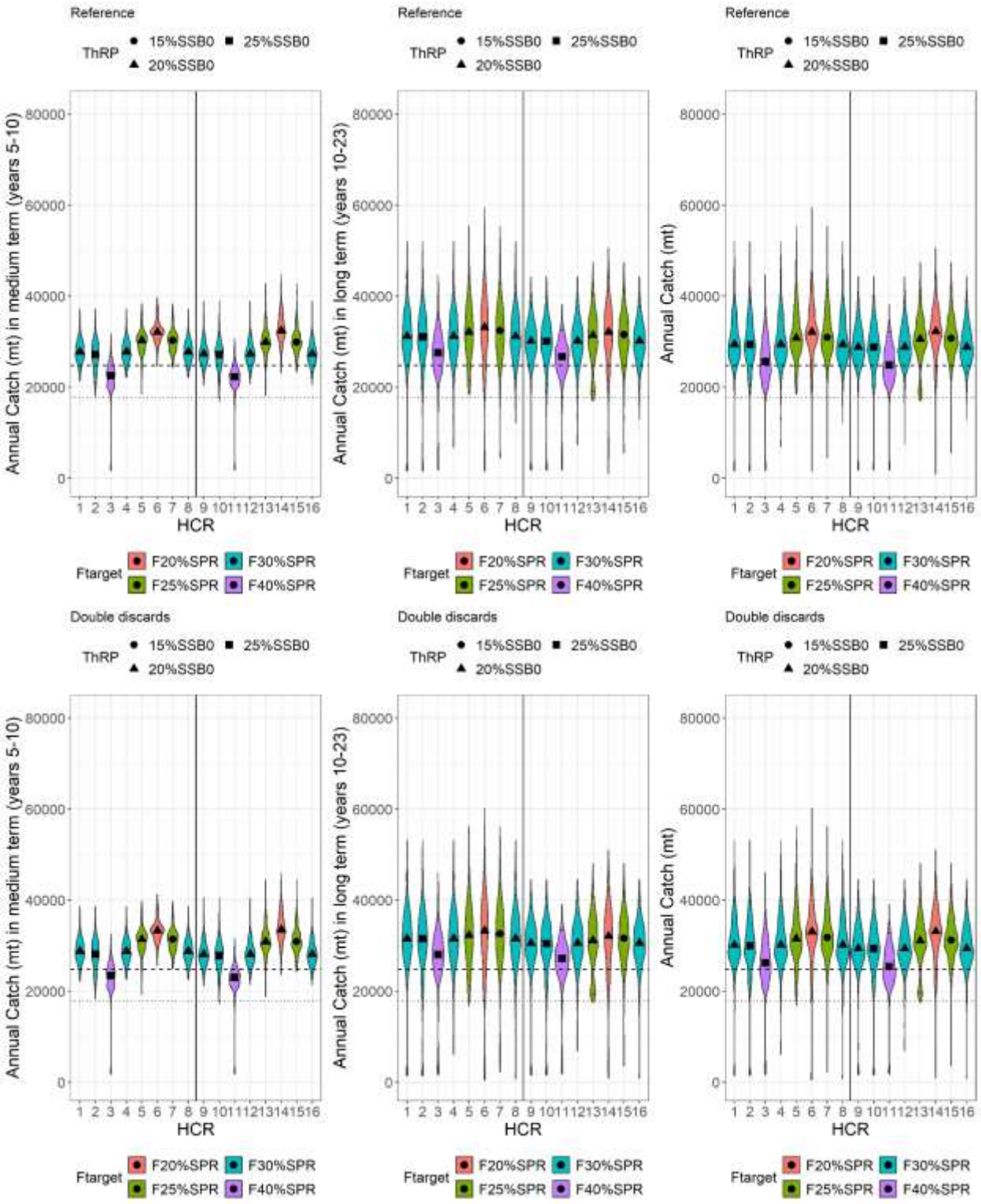

Figure 5Z6. Violin plot showing the probability density of total annual catch (including discards and the EPO recreational fleet) in the medium term (years 5 to 10, first column), long term (years 10 to 23, second column), and across all years (third column) for operating model 1 for each

harvest control rule (HCR) across all iterations for the reference run and the doubling of discards robustness test. The marker inside each violin plot is the median medium term, long term, or annual catch and horizontal solid lines within each violin represent the 5th to 95th quantile range. The ThresholdRP is the first control point for each HCR and stands for threshold reference point. Colors represent the F_{TARGET} reference point associated with each HCR. The dotted line identifies the total catch limit set by the WCPFC's CMM 23-02 plus IATTC's Resolution C-21-05, effective in 2024, plus EPO recreational catches for calendar year 2023. The dashed line identifies the total catch limit set by the WCPFC's CMM 24-01 plus IATTC's Resolution C-24-02, effective in 2025, plus EPO recreational catches for calendar year 2023. For the IATTC's resolution, catch limits were based on half of the biennial TAC. The vertical solid line separates HCRs 9 to 16, which are tuned to an EPO:WCPO impact ratio of 30:70, but are otherwise the same as HCRs 1 to 8. EPO stands for Eastern Pacific Ocean and WCPO for Western Central Pacific Ocean.

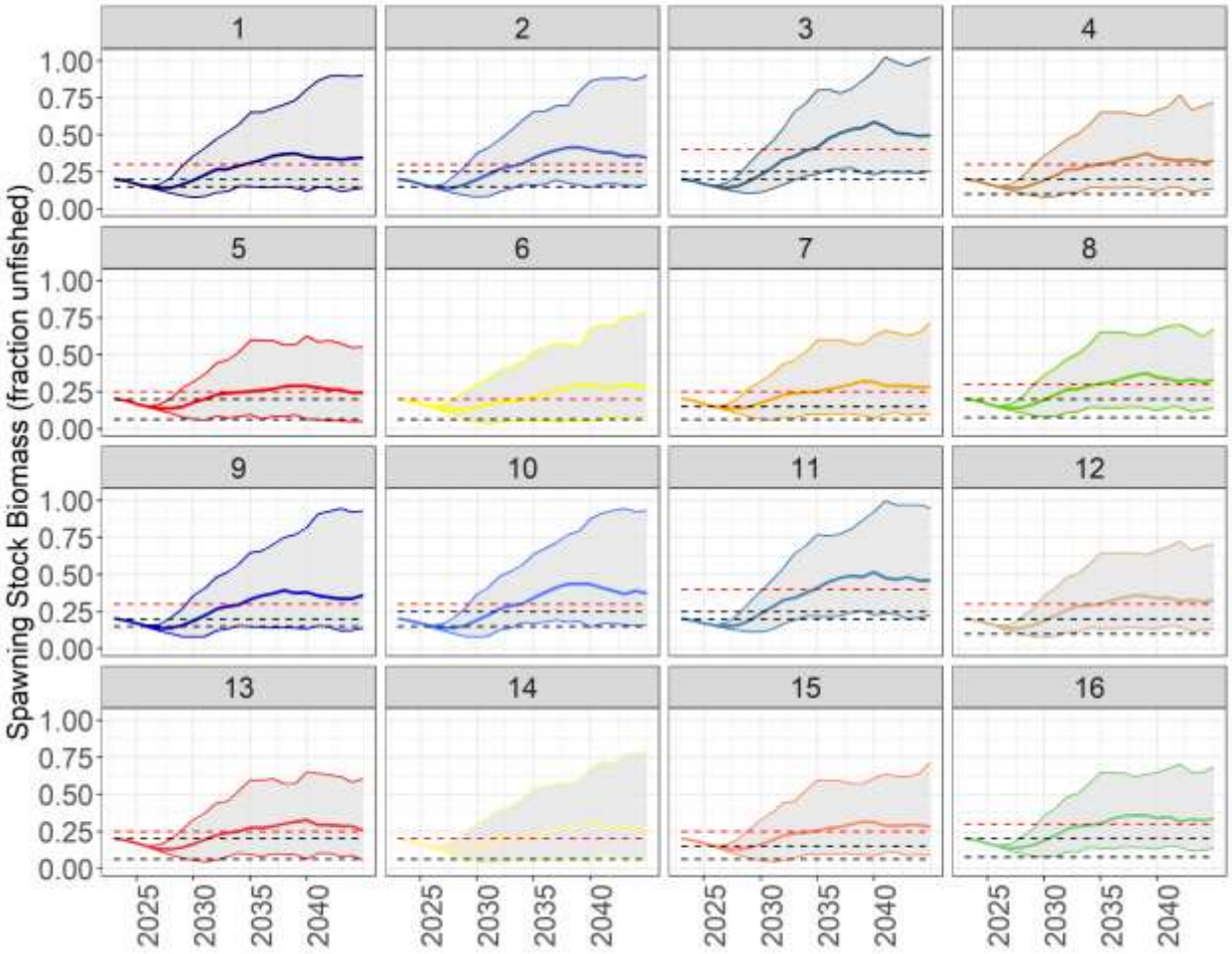

Figure 5Z7. Trends in median relative spawning stock biomass (SSB/unfished SSB, thick solid color line) from operating model 1 under all iterations for the effort creep robustness scenario by harvest control rule (HCR). The grey shading represents trends in the 5th to 95th quantiles range. The lowest black dotted line represents the limit reference point (LRP) for each HCR and the highest the threshold reference point (ThRP). The dotted red line represents the SSB target. Note that HCRs 5, 6, 7, 13, 14, and 15 don't use the LRP as a control point.

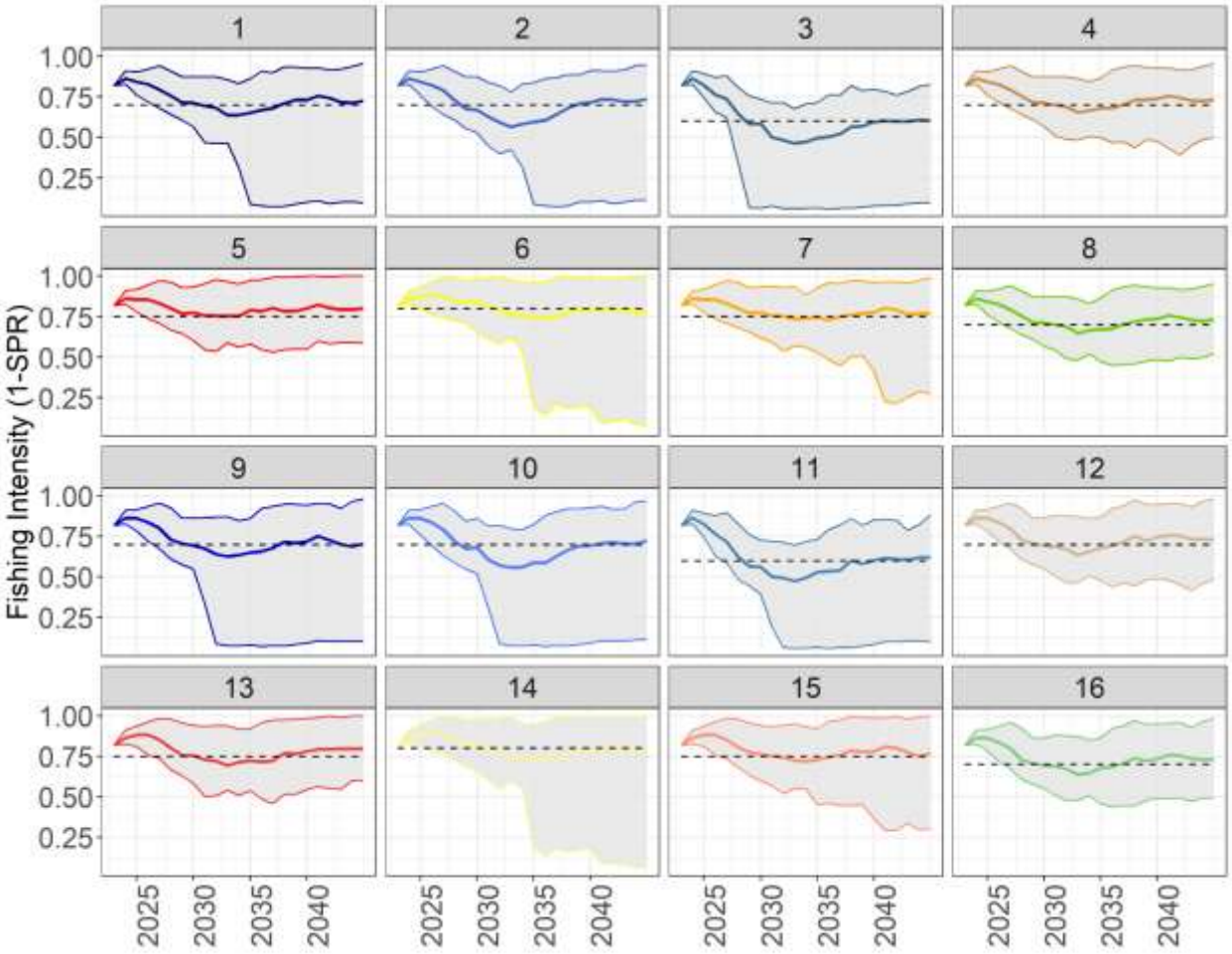

Figure 5Z8. Trends in median fishing intensity (F, 1-SPR, solid color line) from operating model 1 under all iterations for the effort creep robustness scenario by harvest control rule (HCR). The grey shading represents trends in the 5th to 95th quantiles of F. The dotted line represents the F_{TARGET} associated with each HCR.

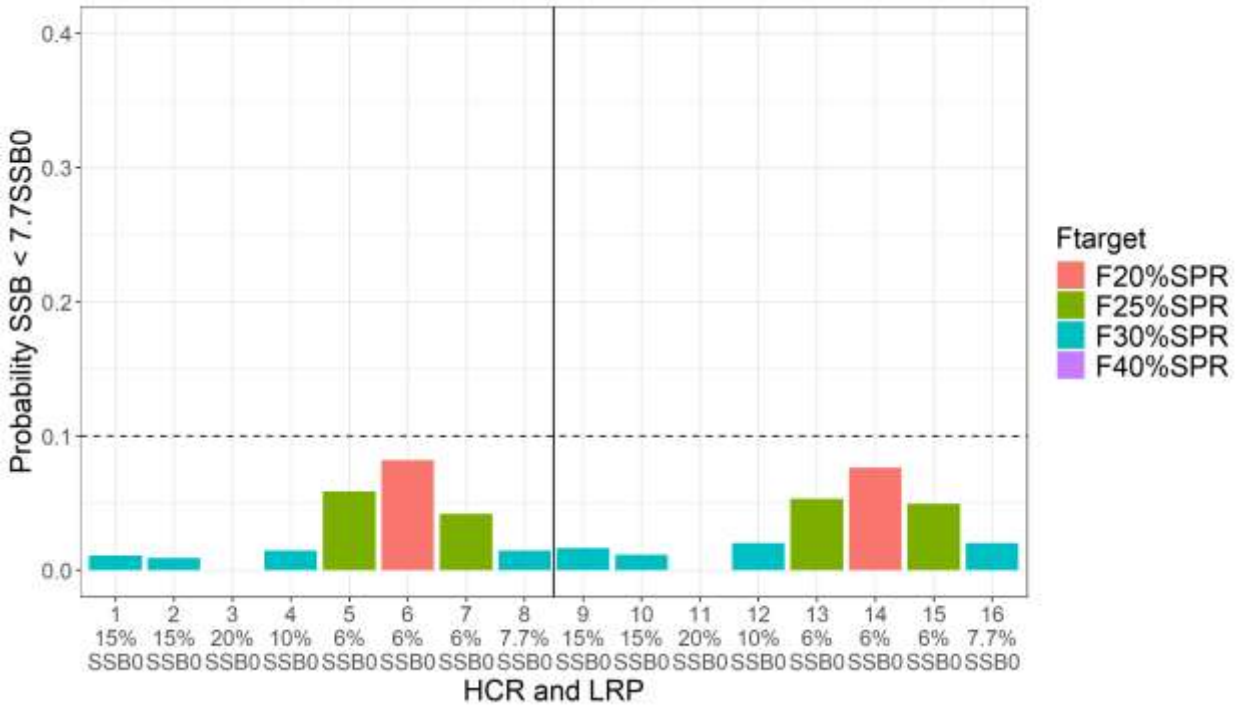

Figure 5Z9. Probability, for each harvest control rule (HCR), of spawning stock biomass (SSB) being less than 7.7%$SSB_{F=0}$ across all iterations and simulation years for the effort creep robustness scenario. The x axis specifies both the HCR number and the limit reference point (LRP) associated with each HCR. The horizontal dotted line represents a 10% probability. The vertical solid line separates HCRs 9 to 16, which are tuned to an EPO:WCPO impact ratio of 30:70, but are otherwise the same as HCRs 1 to 8. EPO stands for Eastern Pacific Ocean and WCPO for Western Central Pacific Ocean.

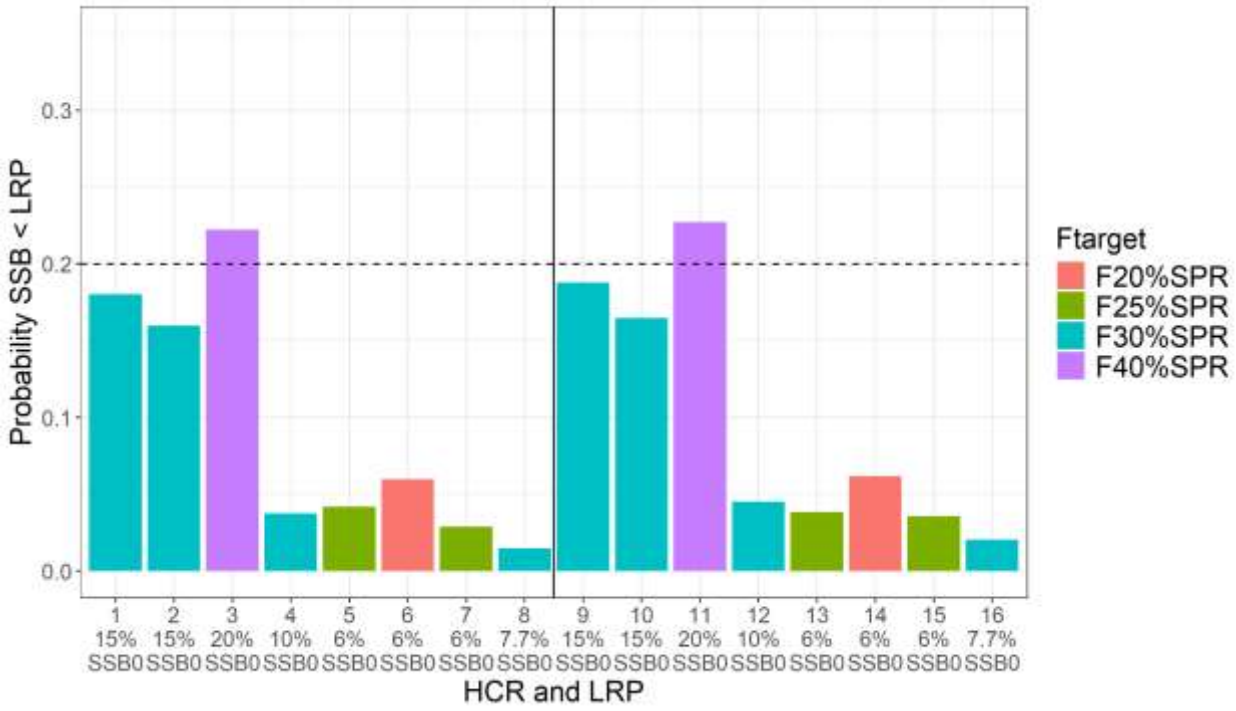

Figure 5Z10. Probability, for each harvest control rule (HCR), of spawning stock biomass (SSB) being less than their own limit reference point (LRP) across all iterations and simulation years for the effort creep robustness scenario. The x axis specifies both the HCR number and the LRP relative biomass level associated with each HCR. The horizontal dotted line represents a 10% probability. The vertical solid line separates HCRs 9 to 16, which are tuned to an EPO:WCPO impact ratio of 30:70, but are otherwise the same as HCRs 1 to 8. EPO stands for Eastern Pacific Ocean and WCPO for Western Central Pacific Ocean. Colors represent the F_{TARGET} reference point associated with each HCR.

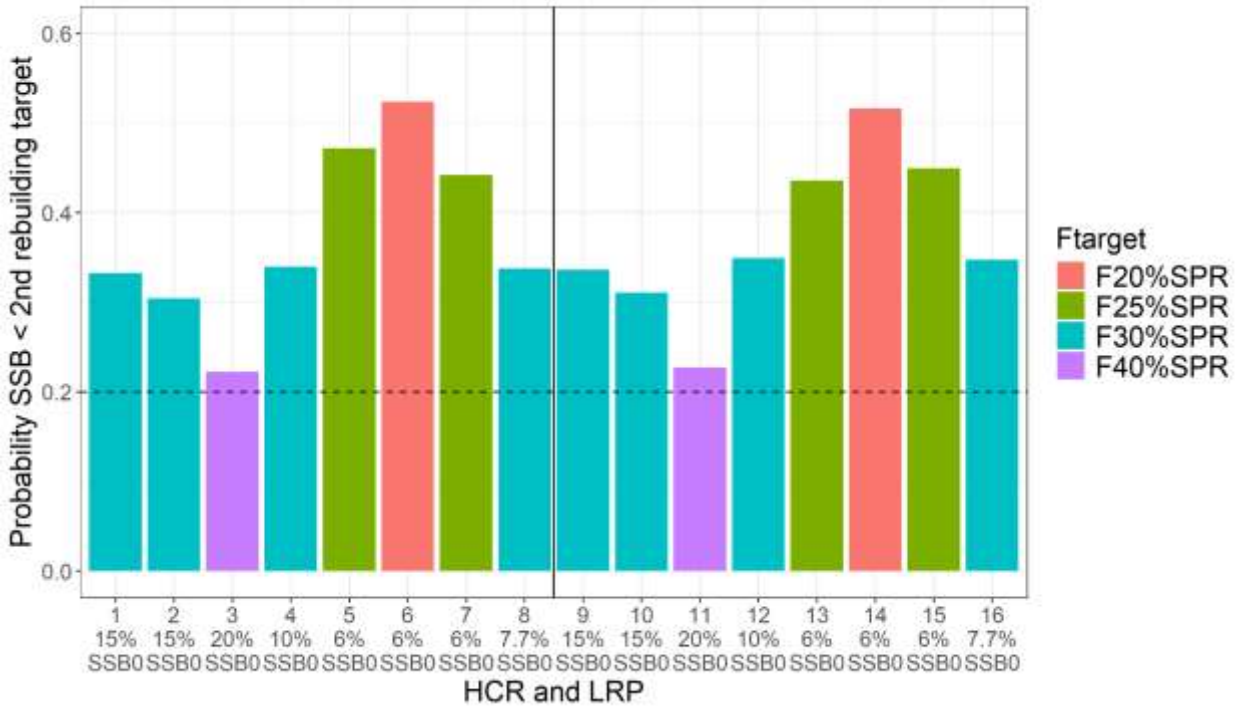

Figure 5Z11. Probability, for each harvest control rule (HCR), of spawning stock biomass (SSB) being less than 20%$SSB_{F=0}$ across all iterations and simulation years for the effort creep robustness scenario. The x axis specifies both the HCR number and the limit reference point (LRP) associated with each HCR. The horizontal dotted line represents a 20% probability. The vertical solid line separates HCRs 9 to 16, which are tuned to an EPO:WCPO impact ratio of 30:70, but are otherwise the same as HCRs 1 to 8. EPO stands for Eastern Pacific Ocean and WCPO for Western Central Pacific Ocean. Colors represent the F_{TARGET} reference point associated with each HCR.

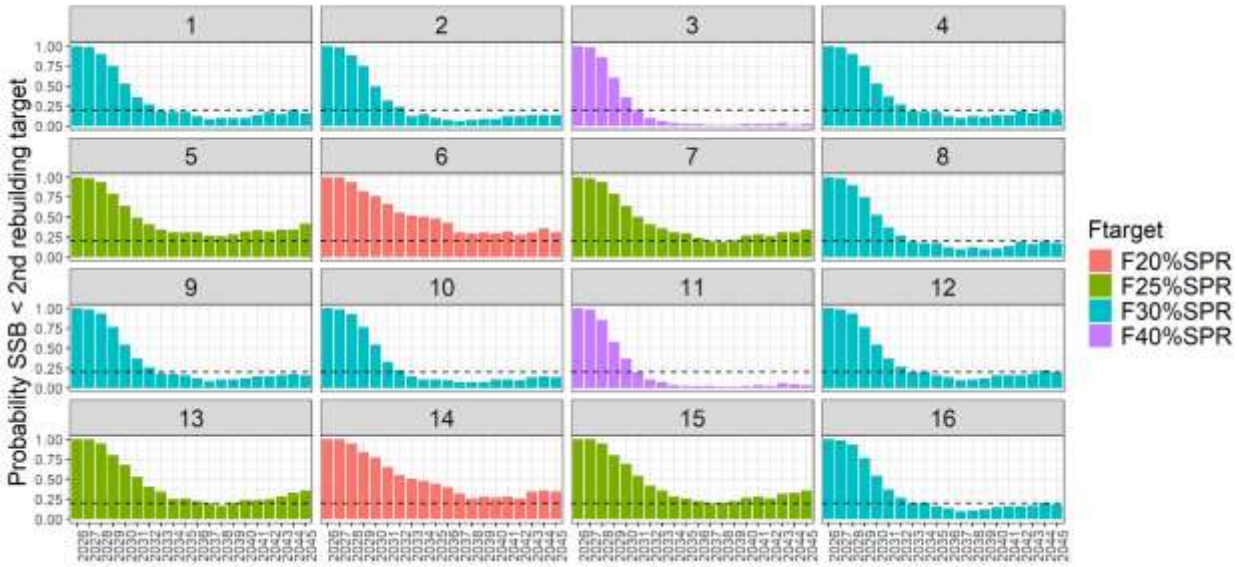

Figure 5Z12. Probability, for each harvest control rule (HCR), of spawning stock biomass (SSB) being less than 20%$SSB_{F=0}$ for the effort creep robustness scenario across all iterations for each simulation year. Each panel presents the results for the labeled HCR. Colors represent the F_{TARGET} reference point associated with each HCR.

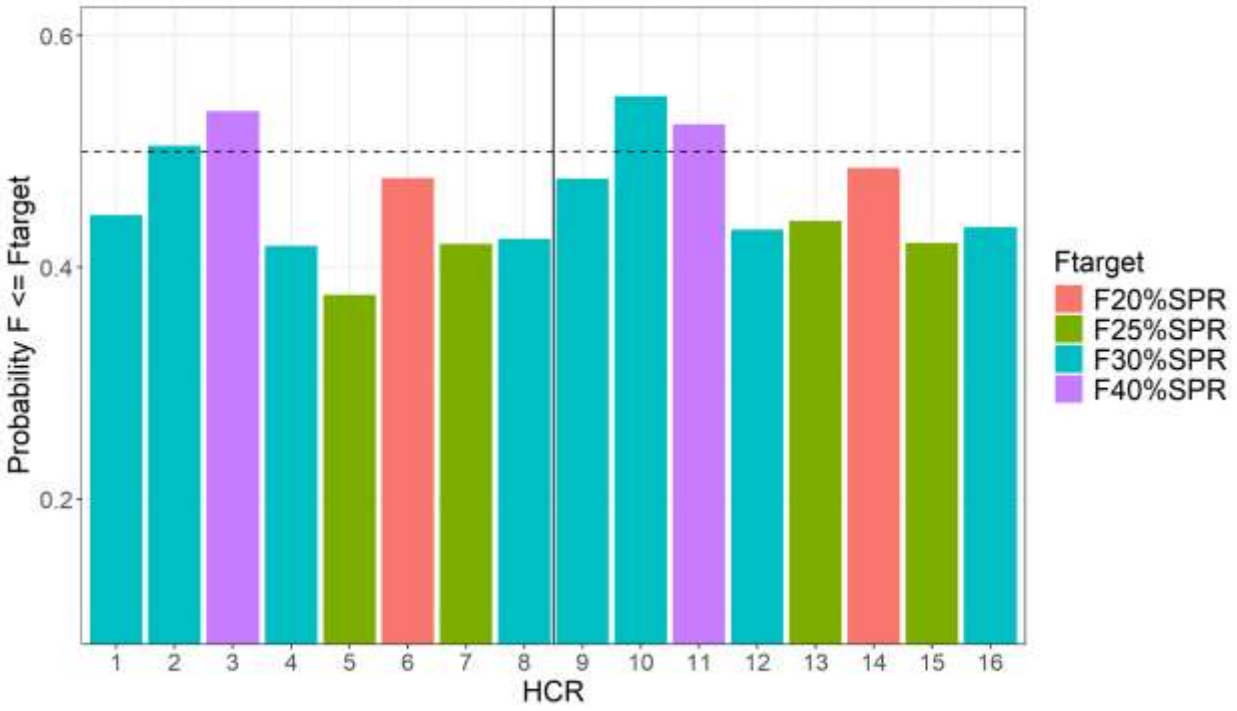

Figure 5Z13. Probability, for each harvest control rule (HCR), of fishing mortality (F, 1-SPR) being less or equal to the F_{TARGET} across all iterations and simulation years for the effort creep robustness scenario. The horizontal dotted line represents a 50% probability. The vertical solid line separates HCRs 9 to 16, which are tuned to an EPO:WCPO impact ratio of 30:70, but are otherwise the same as HCRs 1 to 8. EPO stands for Eastern Pacific Ocean and WCPO for Western Central Pacific Ocean. Colors represent the F_{TARGET} reference point associated with each HCR.

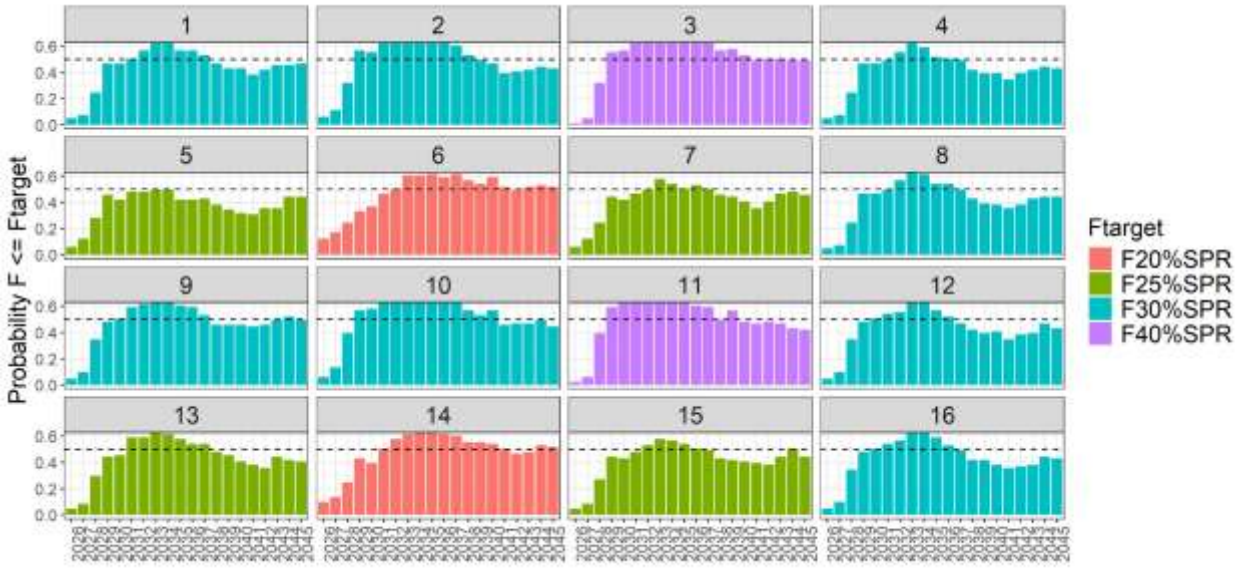

Figure 5Z14. Probability, for each harvest control rule (HCR), of fishing mortality (F, 1-SPR) being less or equal to the F_{TARGET} for the effort creep robustness scenario across all iterations for each simulation year. Each panel presents the results for the labeled HCR. Colors represent the F_{TARGET} reference point associated with each HCR.

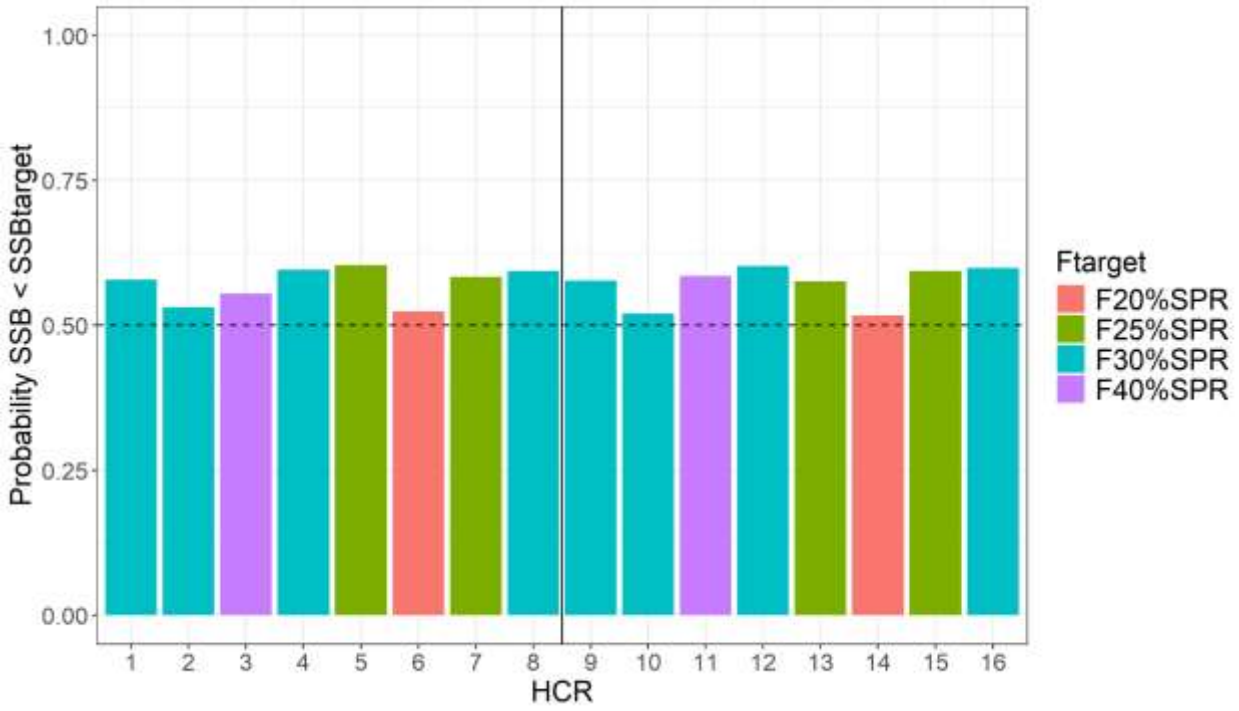

Figure 5Z15. Probability, for each harvest control rule (HCR), of spawning stock biomass (SSB) being below the equivalent biomass depletion level associated with the candidates for F_{TARGET} across all iterations and simulation years for the effort creep robustness scenario. The x axis shows both the HCR number and the LRP relative biomass level associated with each HCR. The horizontal dotted line represents a 20% probability. The vertical solid line separates HCRs 9 to 16, which are tuned to an EPO:WCPO impact ratio of 30:70, but are otherwise the same as HCRs 1 to 8. EPO stands for Eastern Pacific Ocean and WCPO for Western Central Pacific Ocean. Colors represent the F_{TARGET} reference point associated with each HCR.

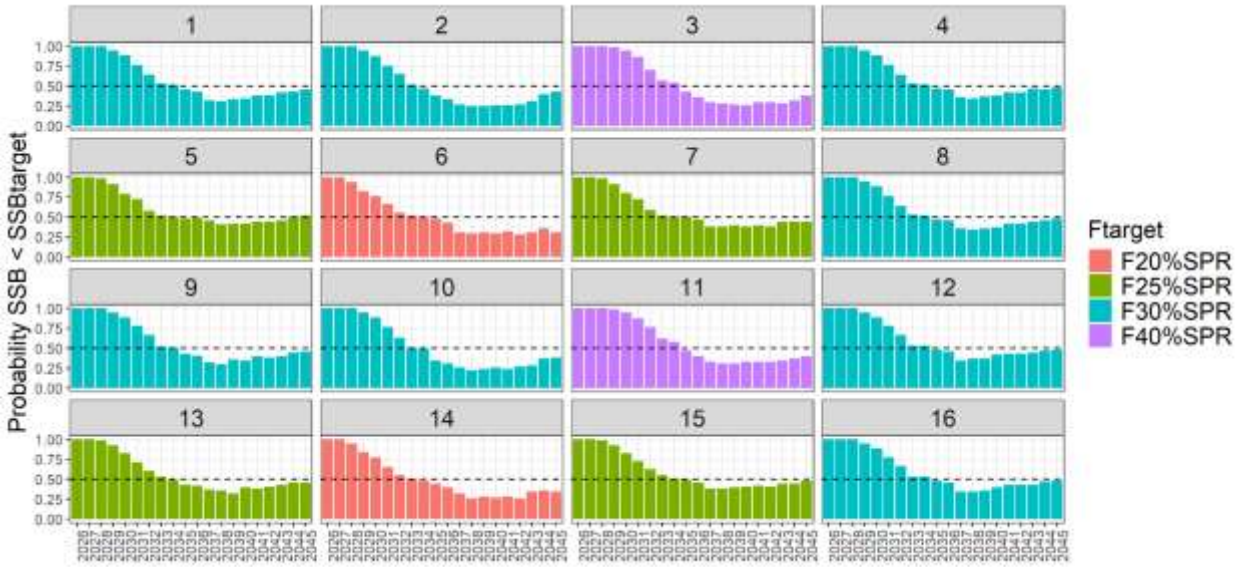

Figure 5Z16. Probability, for each harvest control rule (HCR), of spawning stock biomass (SSB) being below the equivalent biomass depletion level associated with the candidates for F_{TARGET} for the effort creep robustness scenario across all iterations for each simulation year. Each panel presents the results for the labeled HCR. Colors represent the F_{TARGET} reference point associated with each HCR.

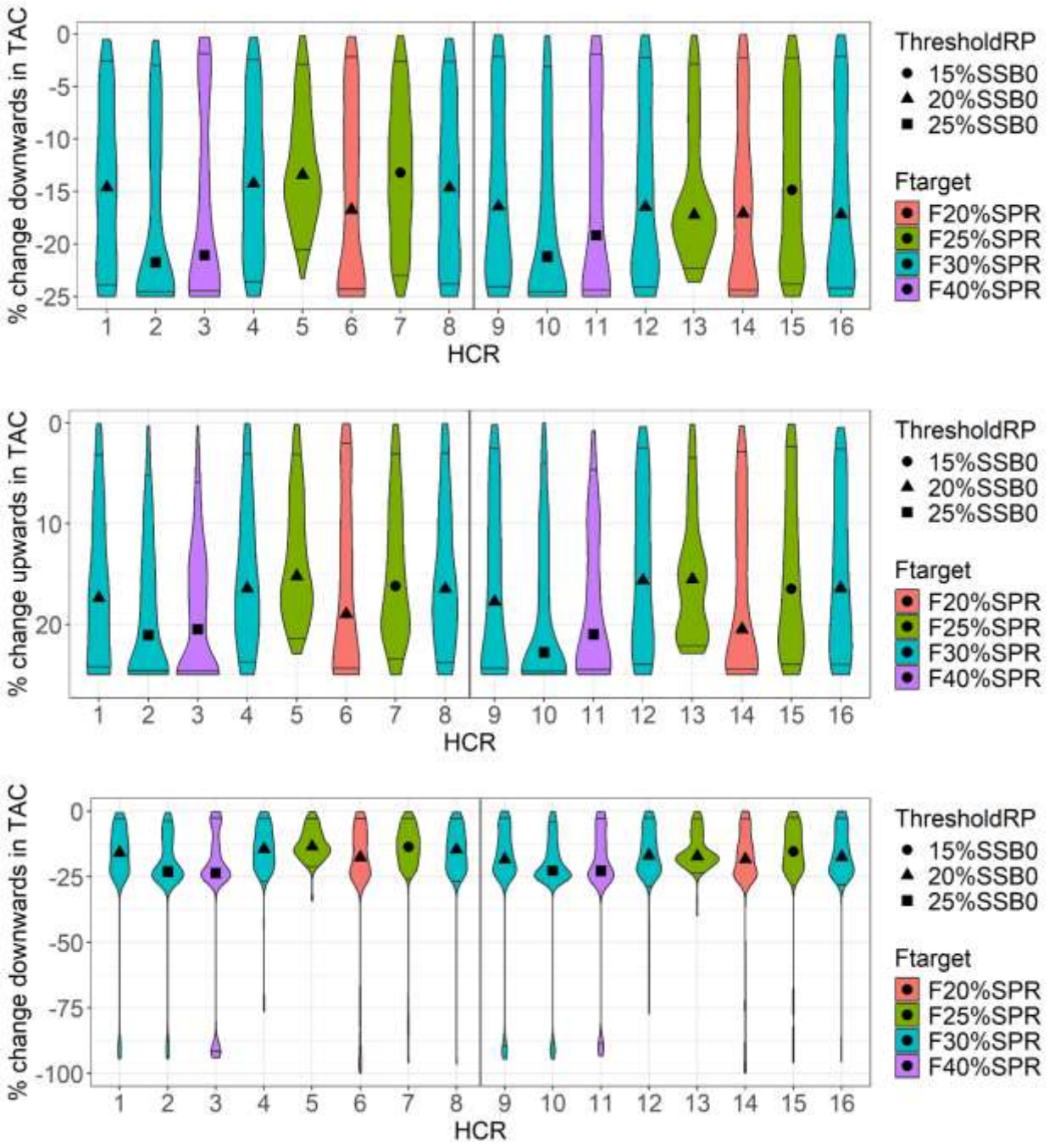

Figure 5Z17. Violin plot showing the probability density of downward changes in TAC (top panel) and of upwards changes in TAC (middle panel) between management period when SSB≥LRP, and of downward changes in TAC across all management periods (bottom panel) for each harvest control rule (HCR) across all iterations and simulation years for the effort creep robustness scenario. The marker inside each violin plot is the median downward change in TAC and horizontal solid lines within each violin represent the 5th to 95th quantile range. The shape of each marker represents the threshold reference point (ThresholdRP). Colors represent the F_{TARGET}

reference point associated with each HCR. The vertical solid line separates HCRs 9 to 16, which are tuned to an EPO:WCPO impact ratio of 30:70, but are otherwise the same as HCRs 1 to 8. EPO stands for Eastern Pacific Ocean and WCPO for Western Central Pacific Ocean.

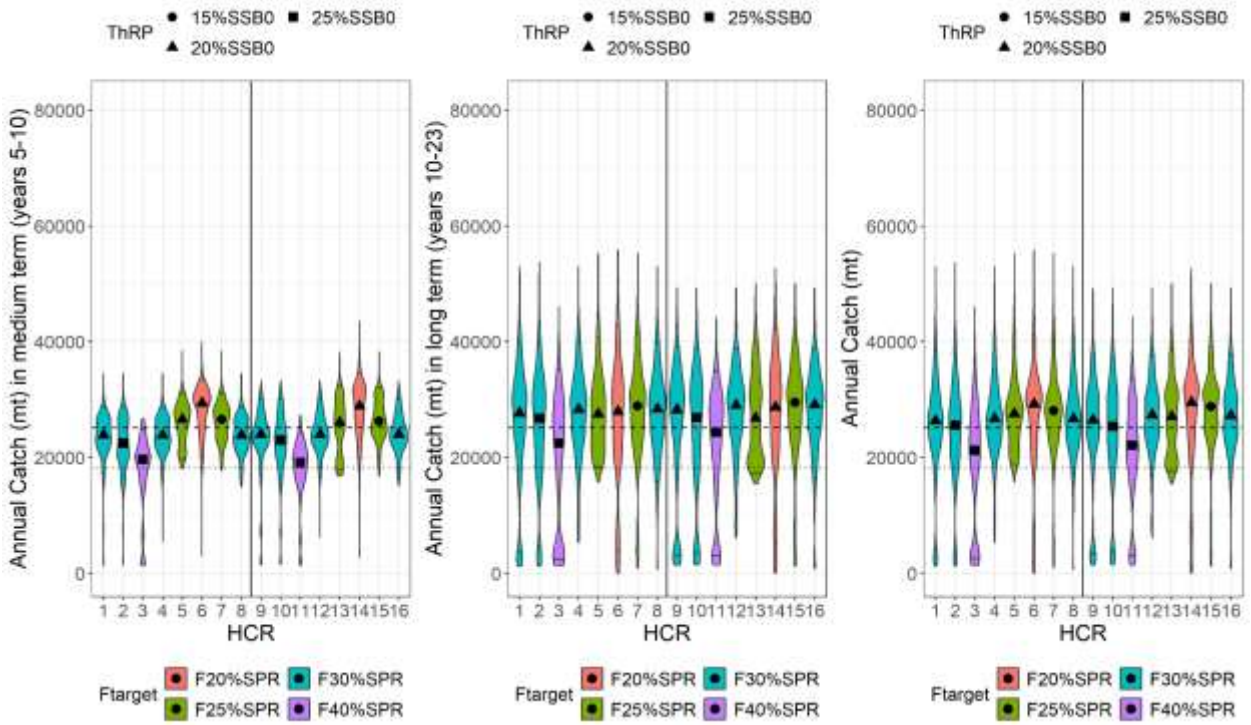

Figure 5Z18. Violin plot showing the probability density of total annual catch (including discards and the EPO recreational fleet) in the medium term (years 5 to 10, first panel), long term (years 10 to 23, second panel), and across all years (third panel) for each harvest control rule (HCR) across all iterations for the effort creep robustness scenario. The marker inside each violin plot is the median medium term, long term, or annual catch and horizontal solid lines within each violin represent the 5th to 95th quantile range. The ThRP is the first control point for each HCR and stands for threshold reference point. Colors represent the F_{TARGET} reference point associated with each HCR. The dotted line identifies the total catch limit set by the WCPFC's CMM 23-02 plus IATTC's Resolution C-21-05, effective in 2024, plus EPO recreational catches for calendar year 2023. The dashed line identifies the total catch limit set by the WCPFC's CMM 24-01 plus IATTC's Resolution C-24-02, effective in 2025, plus EPO recreational catches for calendar year 2023. For the IATTC's resolution, catch limits were based on half of the biennial TAC. The vertical solid line separates HCRs 9 to 16, which are tuned to an EPO:WCPO impact ratio of 30:70, but are otherwise the same as HCRs 1 to 8. EPO stands for Eastern Pacific Ocean and WCPO for Western Central Pacific Ocean.

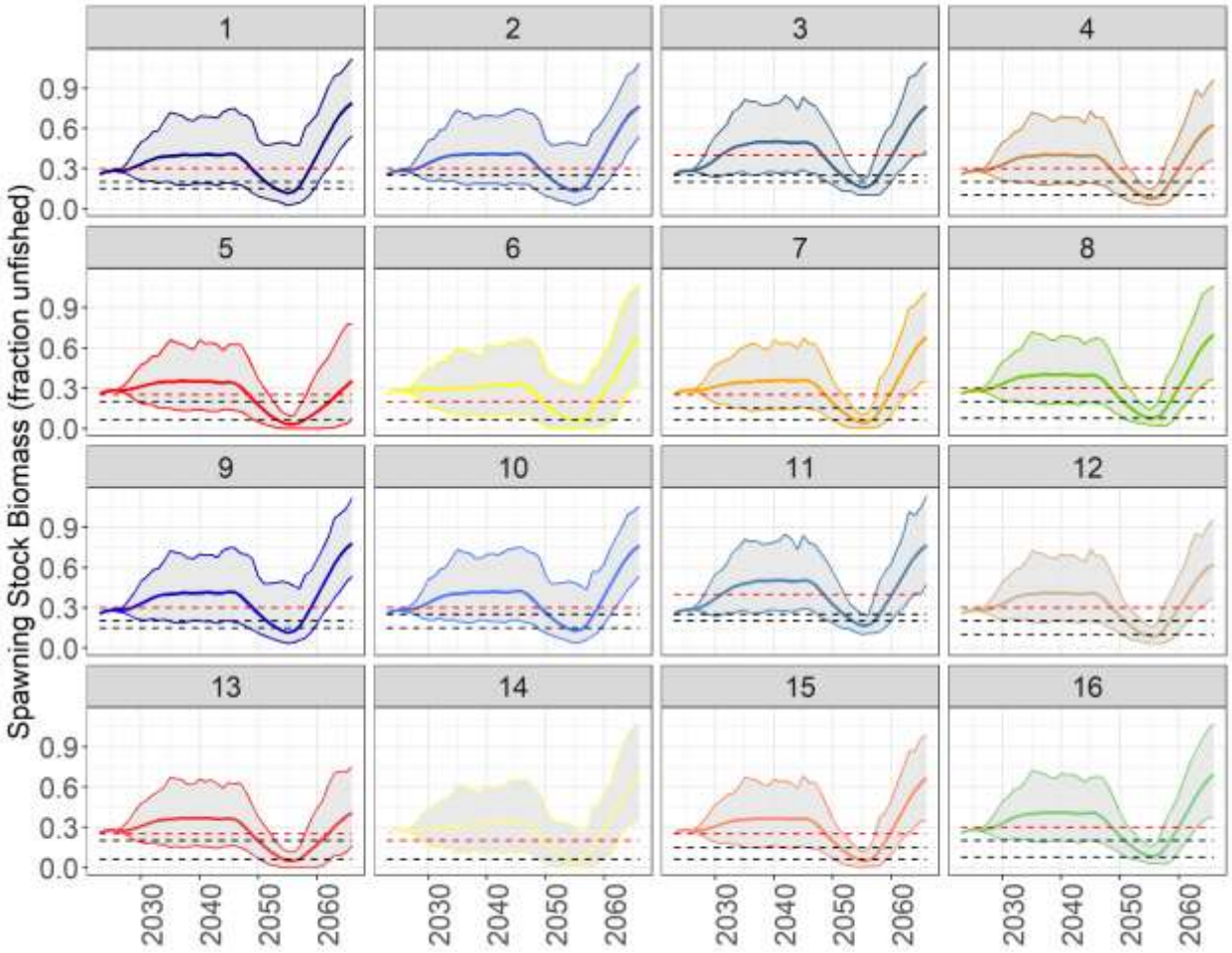

Figure 5Z19. Trends in median relative spawning stock biomass (SSB/unfished SSB, thick solid color line) from operating model 1 under all iterations for the recruitment drop robustness scenario by harvest control rule (HCR). The grey shading represents trends in the 5th to 95th quantiles range. The lowest black dotted line represents the limit reference point (LRP) for each HCR and the highest the threshold reference point (ThRP). The dotted red line represents the SSB target. Note that HCRs 5, 6, 7, 13, 14, and 15 do not use the LRP as a control point.

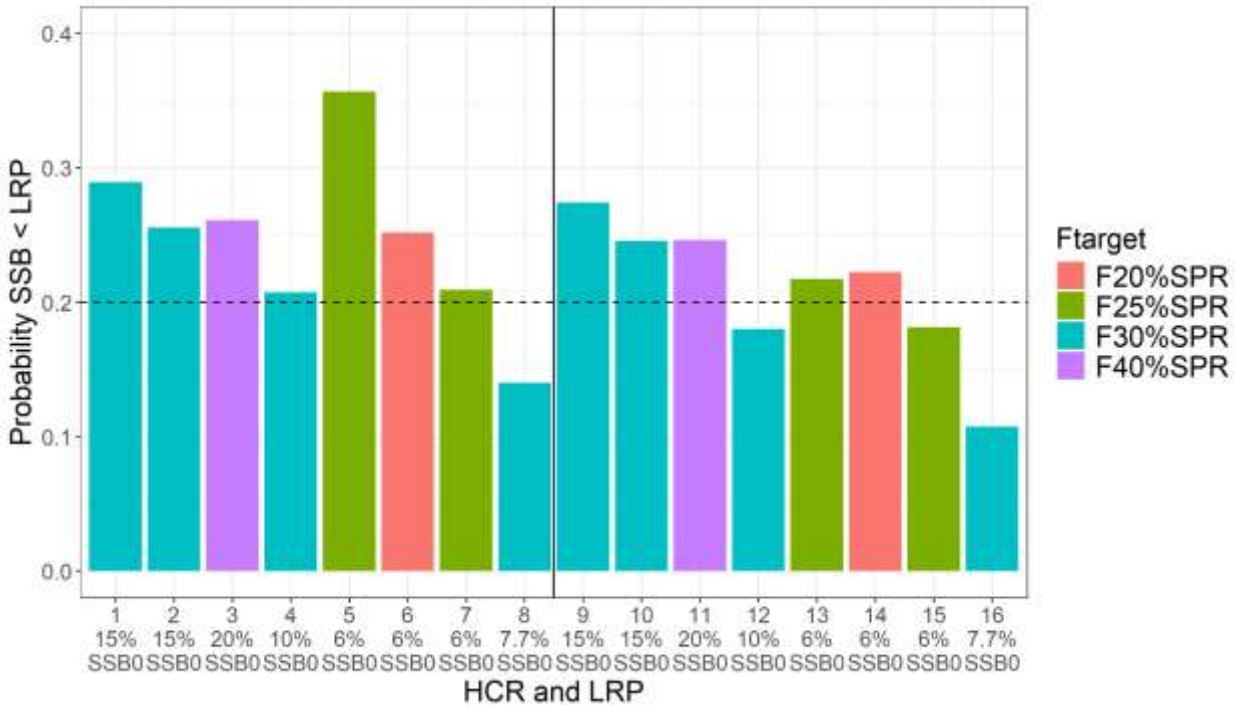

Figure 5Z20. Probability, for each harvest control rule (HCR), of spawning stock biomass (SSB) being less than their own limit reference point (LRP) across all iterations and simulation years for the recruitment drop robustness scenario. The x axis specifies both the HCR number and the LRP relative biomass level associated with each HCR. The horizontal dotted line represents a 20% probability. The vertical solid line separates HCRs 9 to 16, which are tuned to an EPO:WCPO impact ratio of 30:70, but are otherwise the same as HCRs 1 to 8. EPO stands for Eastern Pacific Ocean and WCPO for Western Central Pacific Ocean. Colors represent the F_{TARGET} reference point associated with each HCR.

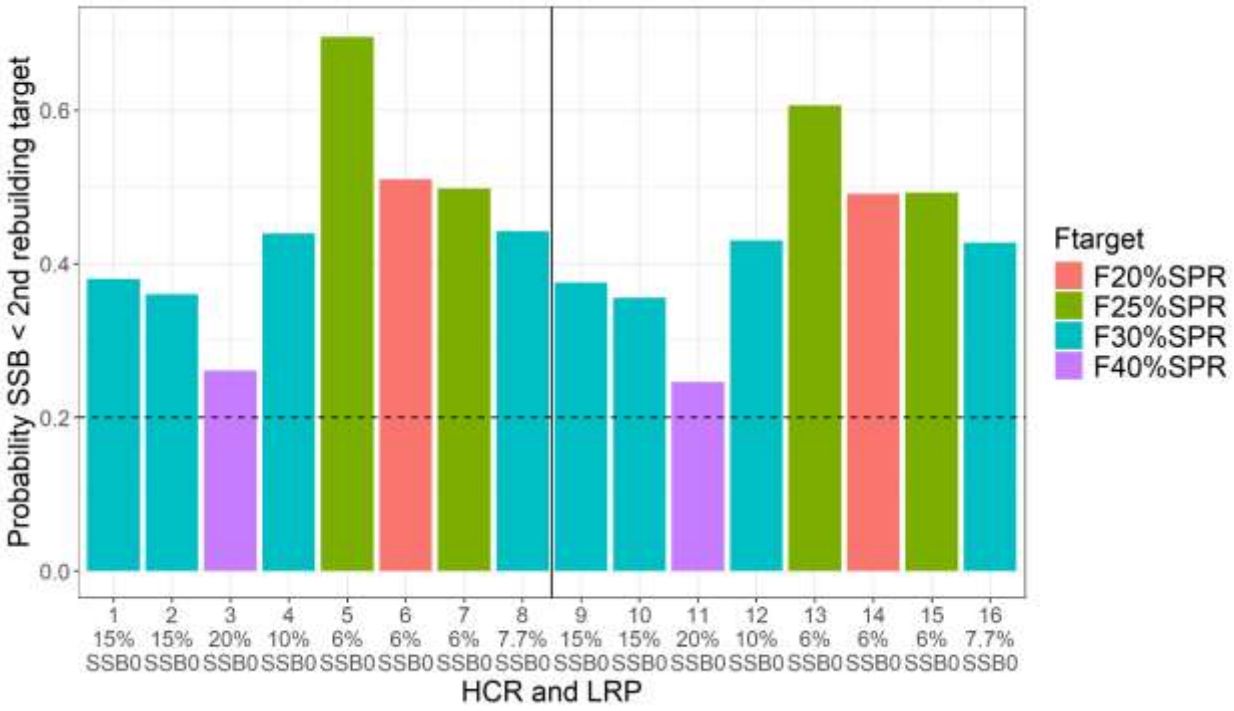

Figure5Z21. Probability, for each harvest control rule (HCR), of spawning stock biomass (SSB) being less than 20%$SSB_{F=0}$ across all iterations and simulation years for the recruitment drop robustness scenario. The x axis specifies both the HCR number and the limit reference point (LRP) associated with each HCR. The horizontal dotted line represents a 20% probability. The vertical solid line separates HCRs 9 to 16, which are tuned to an EPO:WCPO impact ratio of 30:70, but are otherwise the same as HCRs 1 to 8. EPO stands for Eastern Pacific Ocean and WCPO for Western Central Pacific Ocean. Colors represent the F_{TARGET} reference point associated with each HCR.

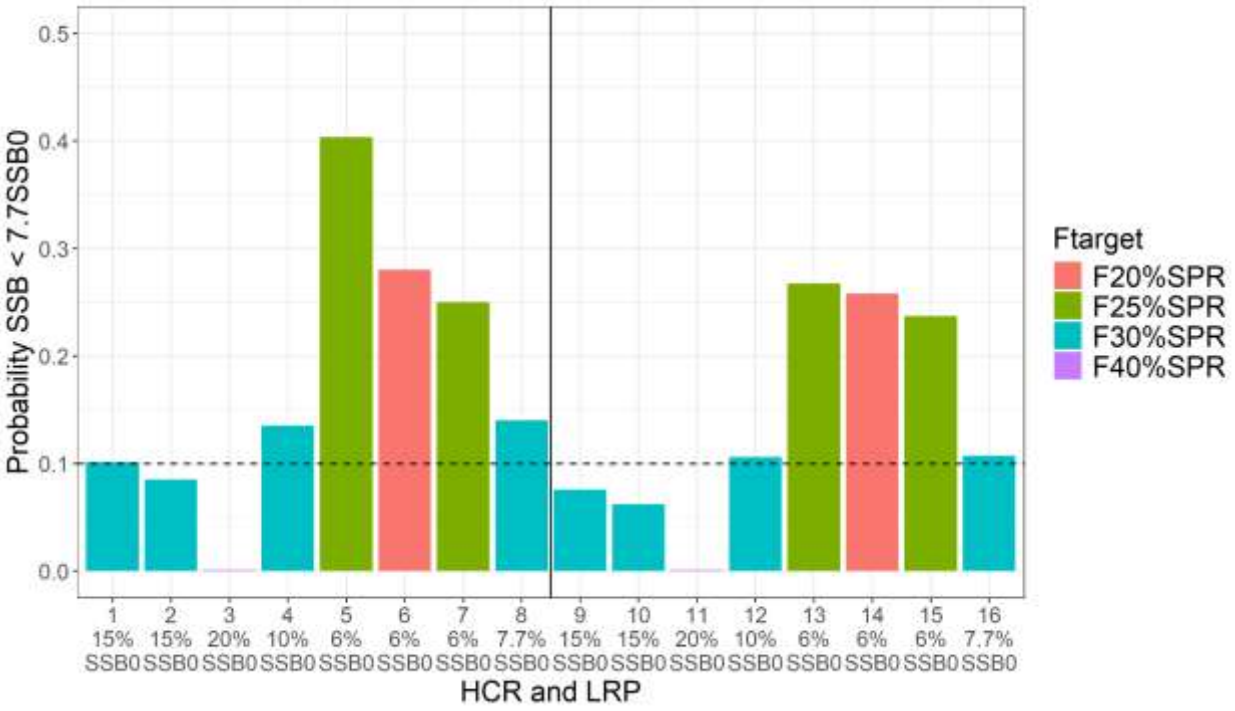

Figure 5Z22. Probability, for each harvest control rule (HCR), of spawning stock biomass (SSB) being less than 7.7%$SSB_{F=0}$ across all iterations and simulation years for the recruitment drop robustness scenario. The x axis specifies both the HCR number and the limit reference point (LRP) associated with each HCR. The horizontal dotted line represents a 10% probability. The vertical solid line separates HCRs 9 to 16, which are tuned to an EPO:WCPO impact ratio of 30:70, but are otherwise the same as HCRs 1 to 8. EPO stands for Eastern Pacific Ocean and WCPO for Western Central Pacific Ocean. Colors represent the F_{TARGET} reference point associated with each HCR.

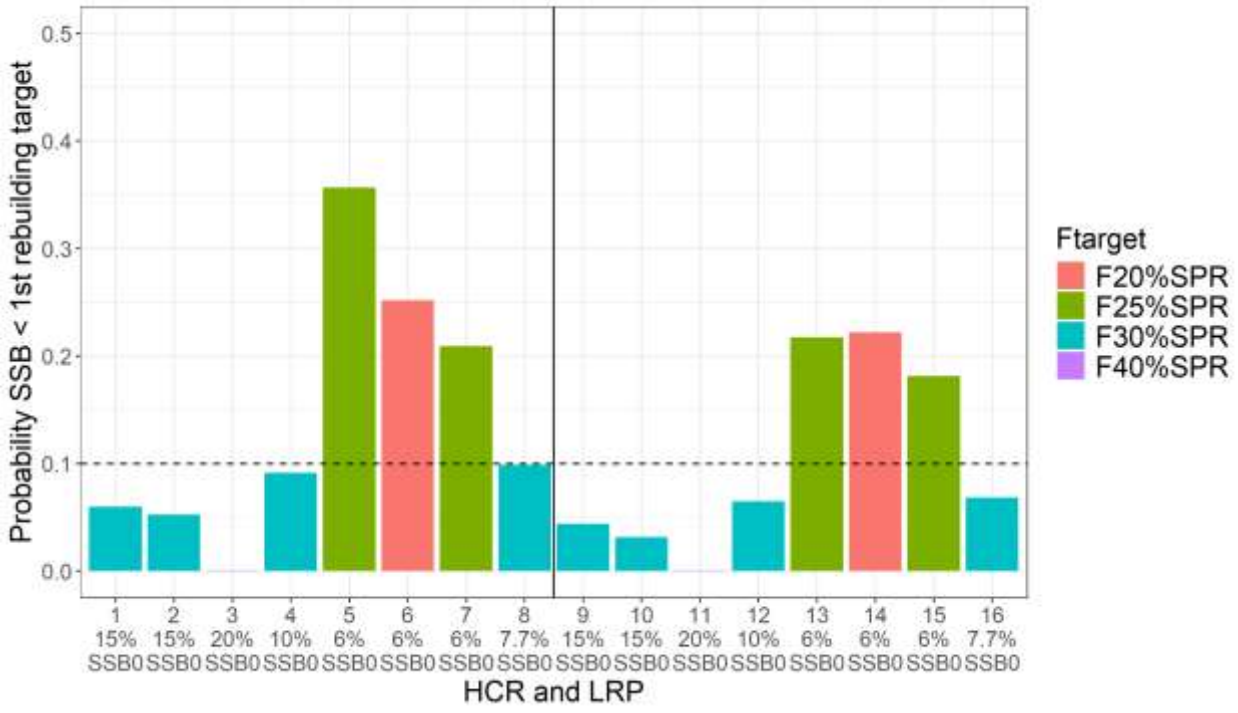

Figure 5Z23. Probability, for each harvest control rule (HCR), of spawning stock biomass (SSB) being less than 6.3%$SSB_{F=0}$ across all iterations and simulation years for the recruitment drop robustness scenario. The x axis specifies both the HCR number and the limit reference point (LRP) associated with each HCR. The horizontal dotted line represents a 10% probability. The vertical solid line separates HCRs 9 to 16, which are tuned to an EPO:WCPO impact ratio of 30:70, but are otherwise the same as HCRs 1 to 8. EPO stands for Eastern Pacific Ocean and WCPO for Western Central Pacific Ocean. Colors represent the F_{TARGET} reference point associated with each HCR.

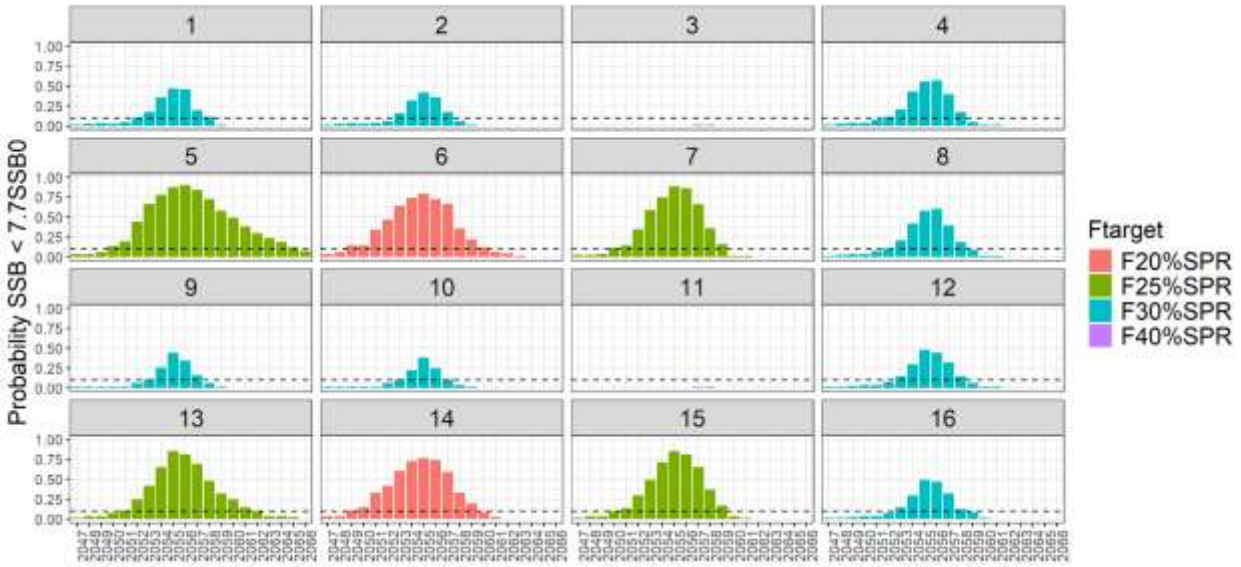

Figure 5Z24. Probability, for each harvest control rule (HCR), of spawning stock biomass (SSB) being less than 7.7%$SSB_{F=0}$ across all iterations for each simulation year for the recruitment drop robustness scenario. The x axis specifies both the HCR number and the limit reference point (LRP) associated with each HCR. The horizontal dotted line represents a 10% probability. The vertical solid line separates HCRs 9 to 16, which are tuned to an EPO:WCPO impact ratio of 30:70, but are otherwise the same as HCRs 1 to 8. EPO stands for Eastern Pacific Ocean and WCPO for Western Central Pacific Ocean. Colors represent the F_{TARGET} reference point associated with each HCR.

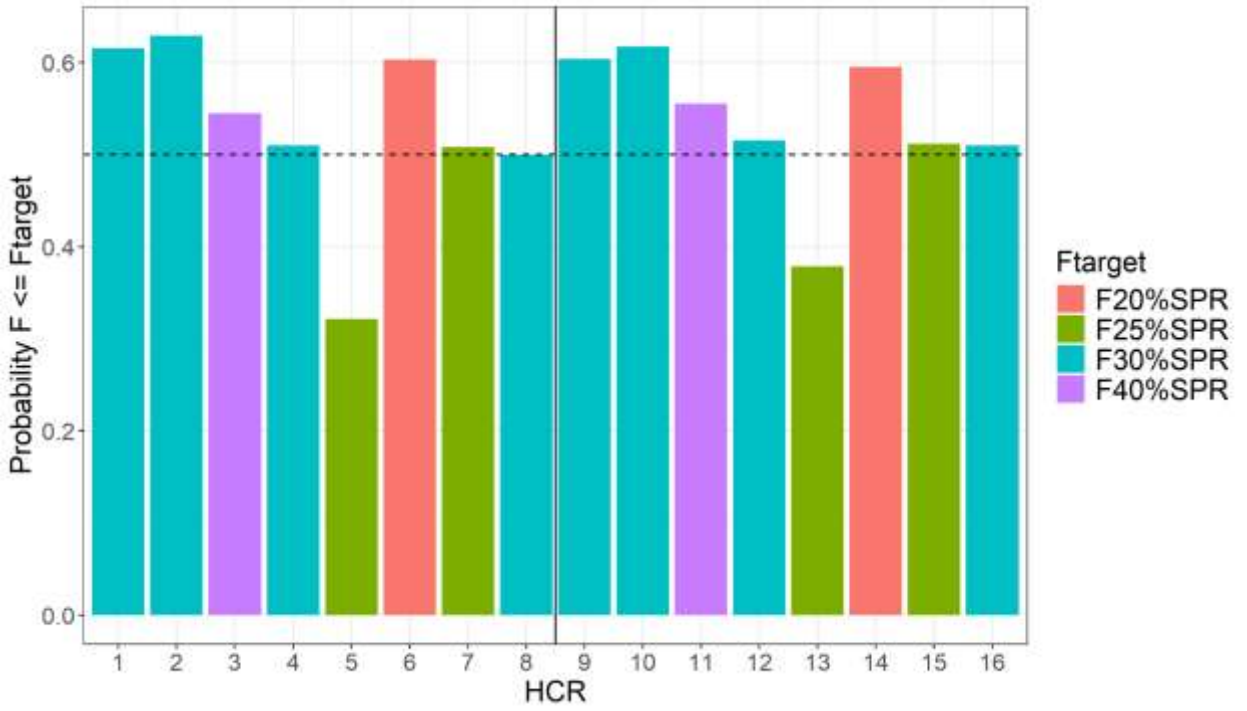

Figure 5Z25. Probability, for each harvest control rule (HCR), of fishing mortality (F, 1-SPR) being less or equal to the F_{TARGET} across all iterations and simulation years for the recruitment drop robustness scenario. The horizontal dotted line represents a 50% probability. The vertical solid line separates HCRs 9 to 16, which are tuned to an EPO:WCPO impact ratio of 30:70, but are otherwise the same as HCRs 1 to 8. EPO stands for Eastern Pacific Ocean and WCPO for Western Central Pacific Ocean. Colors represent the F_{TARGET} reference point associated with each HCR.

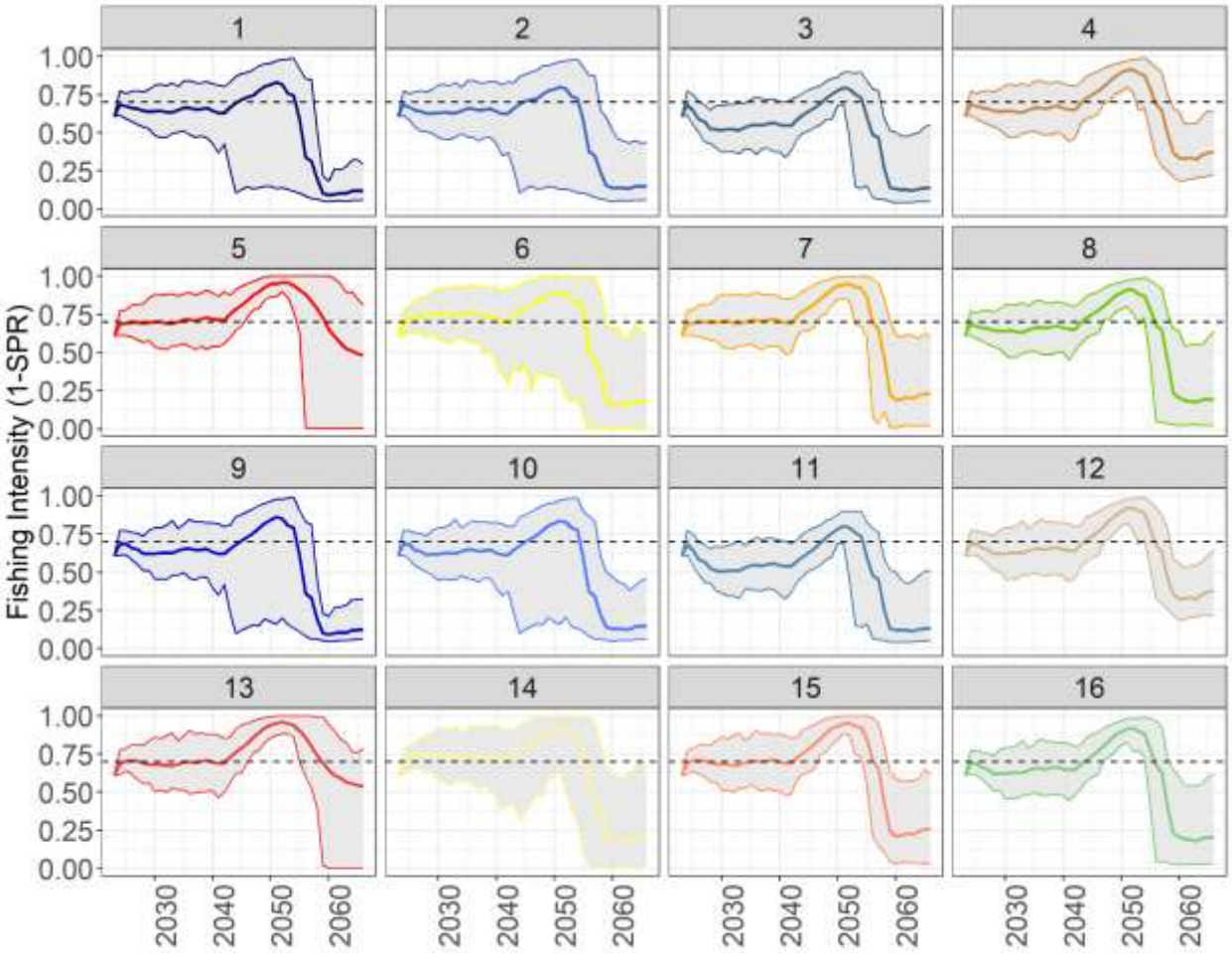

Figure 5Z26. Trends in median fishing intensity (F, 1-SPR, solid color line) from operating model 1 under all iterations for the recruitment drop robustness scenario by harvest control rule (HCR). The grey shading represents trends in the 5th to 95th quantiles of F. The dotted line represents the F_{TARGET} associated with each HCR.

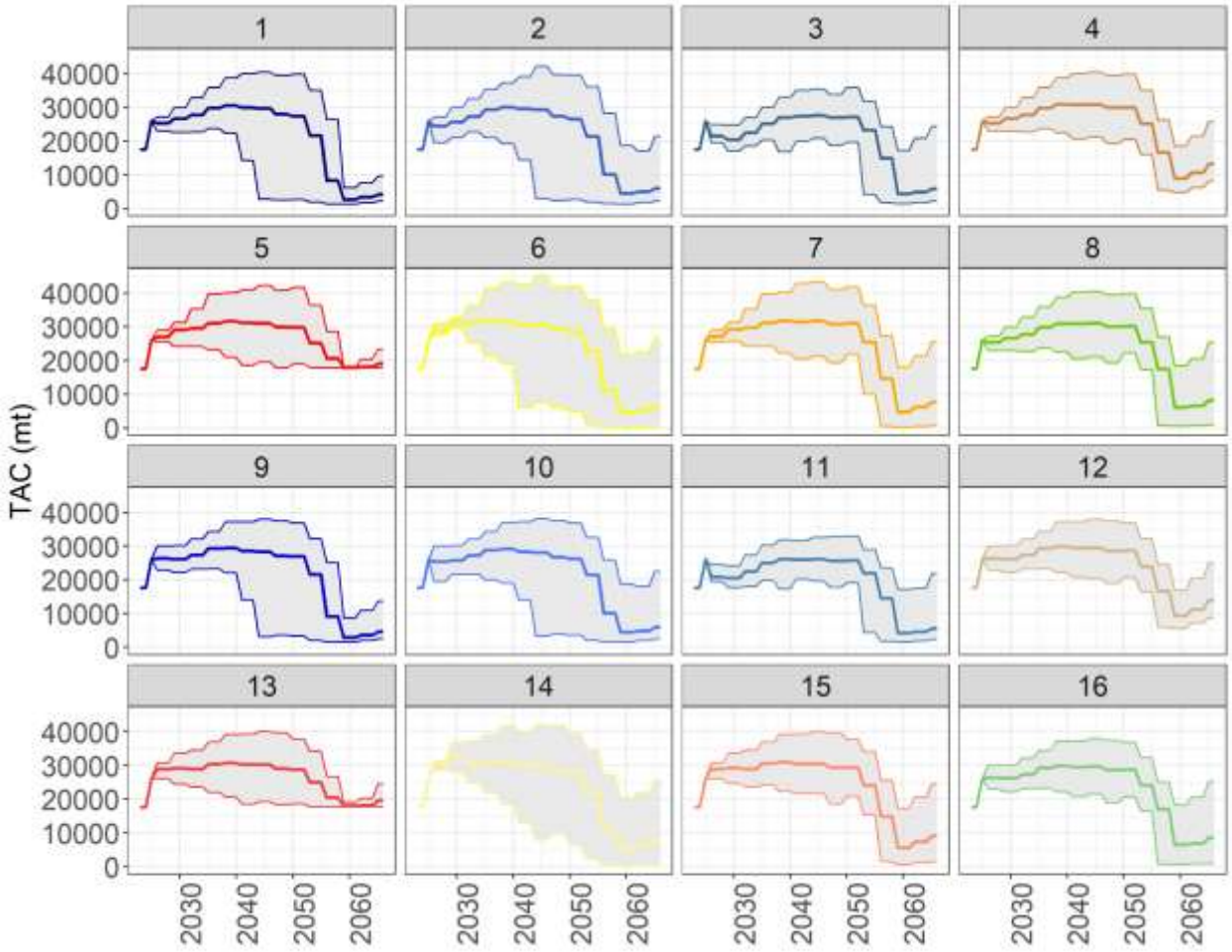

Figure 5Z27. Trends in median total allowable catch (TAC) set by each harvest control rule (HCR) from operating model 1 under all iterations for the recruitment drop robustness scenario by harvest control rule (HCR). The grey shading represents trends in the 5th to 95th quantiles of TAC.

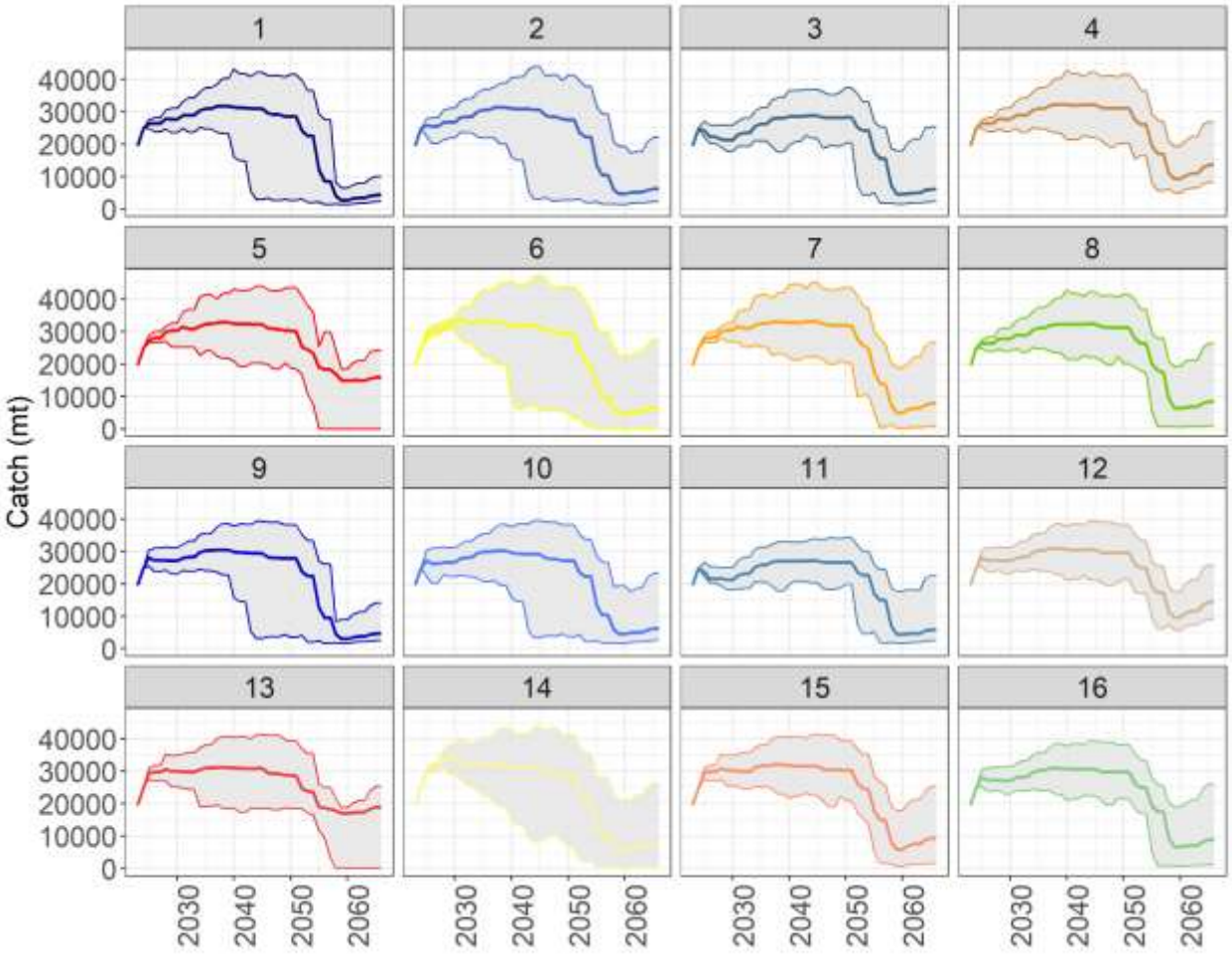

Figure 5Z28. Trends in median catch from operating model 1 under all iterations for the recruitment drop robustness scenario by harvest control rule (HCR). The grey shading represents trends in the 5th to 95th quantiles of catch.

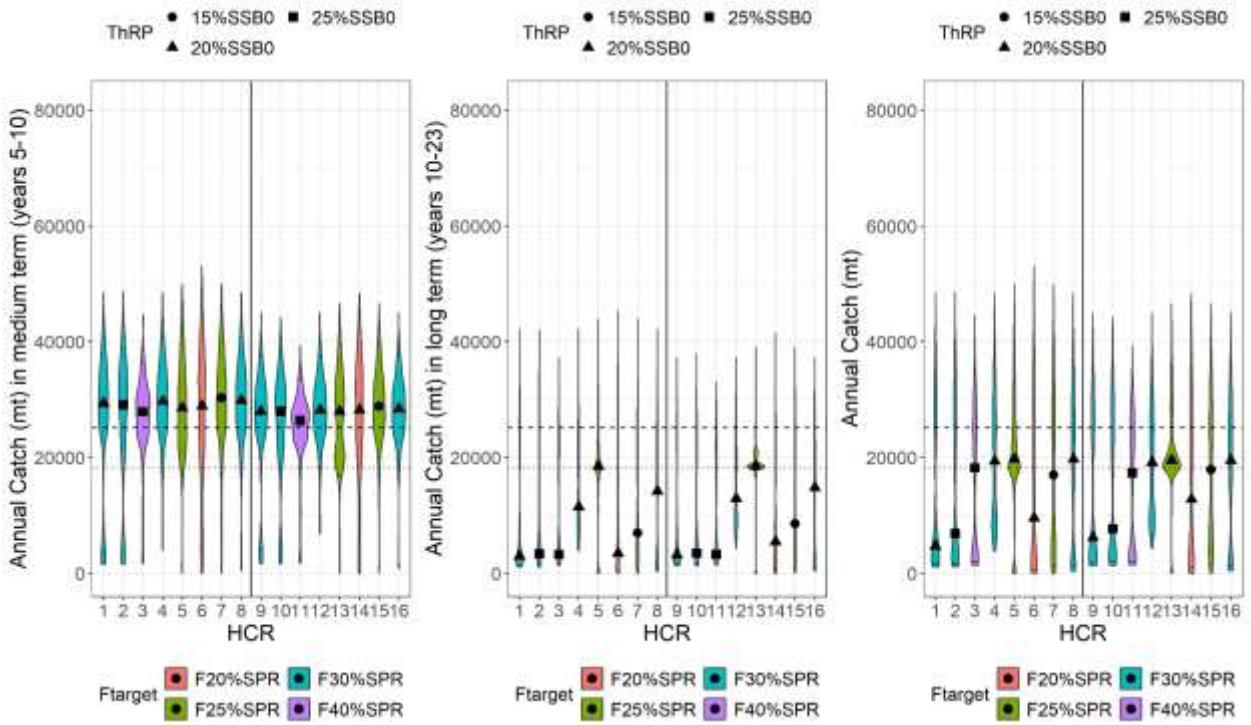

Figure 5Z29. Violin plot showing the probability density of total annual catch (including discards and the EPO recreational fleet) in the medium term (years 5 to 10 since 2044, first panel), long term (years 10 to 23 since 2044, second panel), and across all years (third panel) for each harvest control rule (HCR) across all iterations for the effort creep robustness scenario. The marker inside each violin plots is the median medium term, long term, or annual catch and horizontal solid lines within each violin represent the 5th to 95th quantile range. The ThRP is the first control point for each HCR and stands for threshold reference point. Colors represent the F_{TARGET} reference point associated with each HCR. The dotted line identifies the total catch limit set by the WCPFC's CMM 23-02 plus IATTC's Resolution C-21-05, effective in 2024, plus EPO recreational catches for calendar year 2023. The dashed line identifies the total catch limit set by the WCPFC's CMM 24-01 plus IATTC's Resolution C-24-02, effective in 2025, plus EPO recreational catches for calendar year 2023. For the IATTC's resolution, catch limits were based on half of the biennial TAC. The vertical solid line separates HCRs 9 to 16, which are tuned to an EPO:WCPO impact ratio of 30:70, but are otherwise the same as HCRs 1 to 8. EPO stands for Eastern Pacific Ocean and WCPO for Western Central Pacific Ocean.

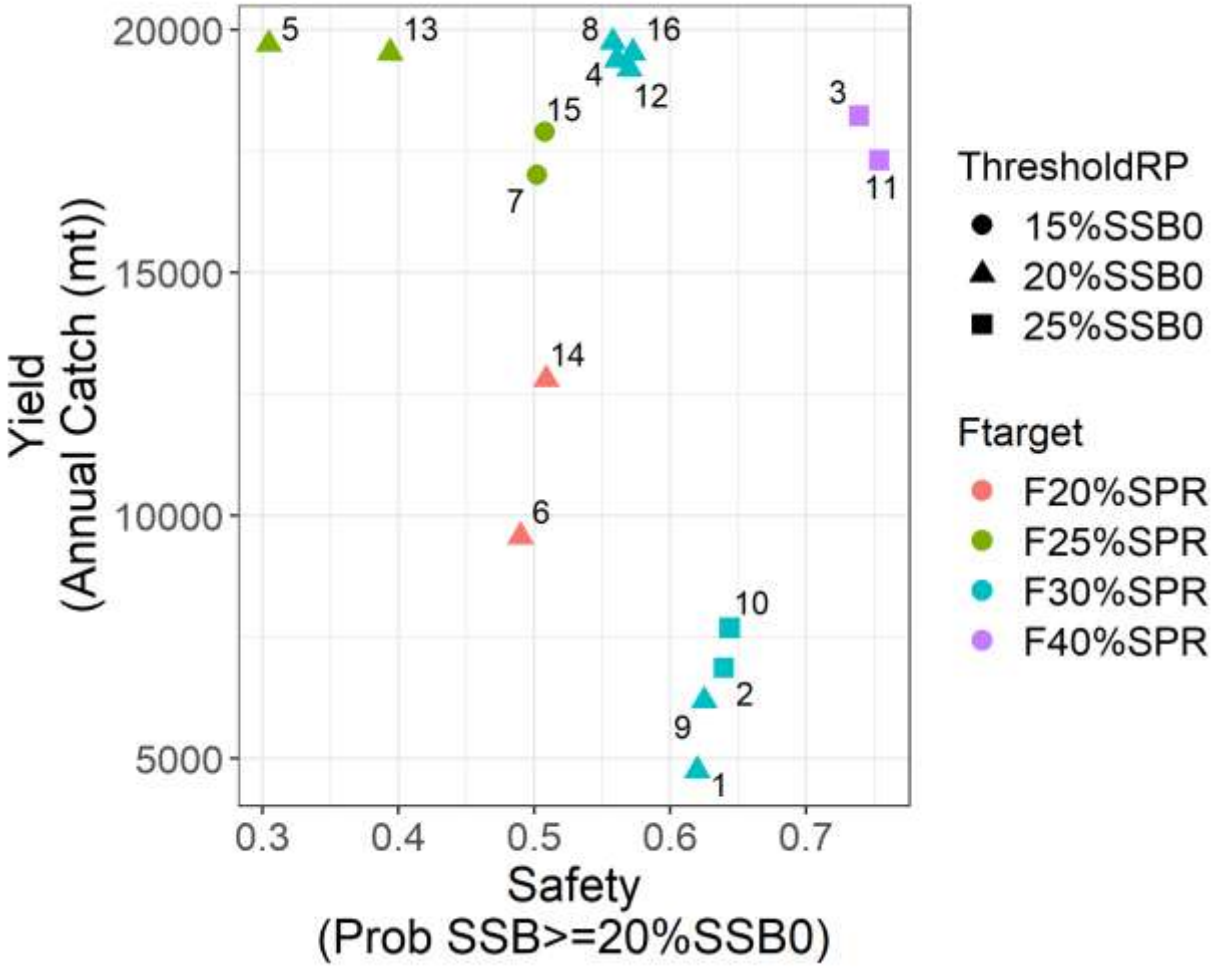

Figure 5Z30. Median annual total catch versus the probability of spawning stock biomass (SSB) being at or above the second rebuilding target of $20\%SSB_{F=0}$ under the recruitment drop scenario. Note that to ensure that for both measures a higher value is better, here we reversed the second performance metric shown in Fig. ES3 to be the probability of $SSB \geq 20\%SSB_{F=0}$ instead of the probability of $SSB < 20\%SSB_{F=0}$. Each HCR is labeled and represented by a symbol colored according to their F_{TARGET}. The ThresholdRP is the first control point for each HCR and stands for Threshold reference point.

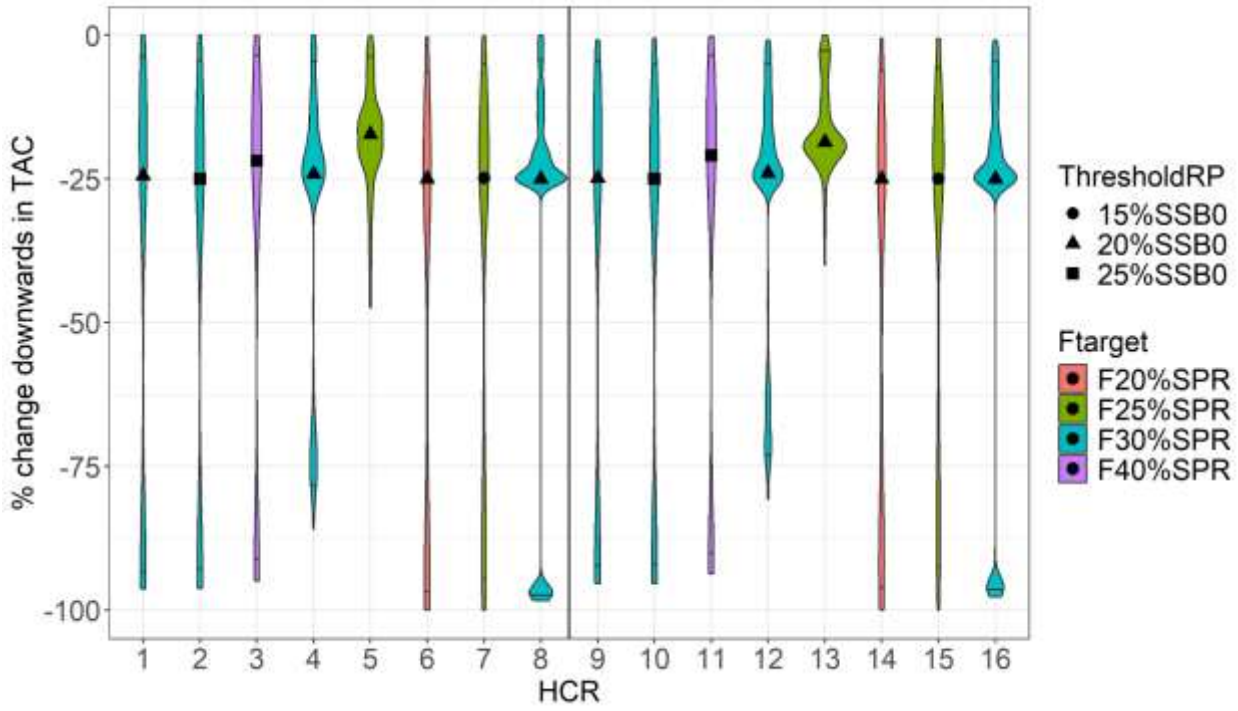

Figure 5Z31. Violin plot showing the probability density of downward changes in TAC across all management periods (bottom panel) for operating model 1 for each harvest control rule (HCR) across all iterations and simulation years for the recruitment drop robustness scenario. The marker inside each violin plot is the median downward change in TAC and horizontal solid lines within each violin represent the 5th to 95th quantile range. The shape of each marker represents the threshold reference point (ThresholdRP). Colors represent the F_{TARGET} reference point associated with each HCR. The vertical solid line separates HCRs 9 to 16, which are tuned to an EPO:WCPO impact ratio of 30:70, but are otherwise the same as HCRs 1 to 8. EPO stands for Eastern Pacific Ocean and WCPO for Western Central Pacific Ocean.