# Performance of Candidate Model-based Harvest Control Rules for Pacific Bluefin Tuna

Desiree Tommasi[a,b], Huihua Lee[a], Kevin Piner[c]

a: NOAA Fisheries, Southwest Fisheries Science Center,
La Jolla, CA, USA
b: Institute of Marine Sciences, University of California Santa Cruz,
Santa Cruz, CA, USA
c: NOAA Fisheries (Retired),
Encampment, WY, USA

**March 2023**

Working document submitted to the ISC Pacific bluefin tuna Working Group, International Scientific Committee for Tuna and Tuna-Like Species in the North Pacific Ocean (ISC), from 21 to 24 March 2023, Shinagawa, Japan.

**Summary**

The Western and Central Pacific Fisheries Commission of the Northern Committee (WCPFC NC) and the Inter American Tropical Tuna Commission (IATTC) requested, via the Joint Working Group (JWG), that the ISC PBF working group develop a Management Strategy Evaluation (MSE) to help inform development of a long-term management strategy for PBF (JWG 2022). As part of the MSE process the JWG proposed a set of candidate reference points and associated model-based harvest control rules (HCRs) to be evaluated using MSE (WCPFC 2019, Annex F). The total number of candidate HCRs to be tested is 115. It would not be feasible to timely evaluate all these HCRs in an MSE simulation that includes an estimation model due to long run times. We therefore evaluate performance of these HCRs in a faster MSE simulation with no estimation error to guide selection of a subset of HCRs for evaluation in an MSE simulation with an estimation model.

**Introduction**

Management strategy evaluation (MSE) is a process whereby the performance of a set of harvest strategies relative to some management objectives and performance metrics of interest to stakeholders is assessed under a range of uncertainties using a computer simulation (Punt et al. 2016). The two Regional Fisheries Management Organizations (RFMOs) tasked with managing the Pacific Bluefin tuna (PBF) stock, namely the Western and Central Pacific Fisheries Commission of the Northern Committee (WCPFC NC) and the Inter American Tropical Tuna Commission (IATTC) requested, via the Joint Working Group, that the ISC PBF working group develop an MSE to help inform development of a long-term management strategy for PBF. According to the JWG workplan the MSE results would be presented at JWG10 in 2025 (JWG 2022 – Annex G). As part of the MSE process the JWG proposed a set of candidate reference points and associated model-based HCRs (i.e. where stock status is determined via an assessment model) to be evaluated using MSE (WCPFC 2019, Annex F). Considering the different potential combination of reference points, 115 potential model-based HCRs have been put forward (Table 1).

To capture the range of uncertainty in the system, an MSE simulation includes a set of operating models (OMs), which are mathematical representations of the true dynamics of the population and fisheries of interest. Having more than one OM allows an MSE to evaluate performance of harvest controls under different assumptions related to the biology of the stock, the fisheries, or the management system (Punt et al. 2016). An MSE also accounts for process uncertainty (e.g., in recruitment) by running many simulations with different recruitment trajectories. The OMs are used to simulate trends in the population under a range of different management procedures in a closed-loop forward simulation. Data are generated with error from the OM and, for a model-based management procedure (i.e. where stock status is determined via an assessment model), are input to the estimation model (EM, i.e. the stock assessment model). For PBF, running the stock assessment during each assessment period in the forward simulation is computationally expensive. To ensure timely completion of the PBF MSE, the number of HCRs to be evaluated in an MSE with an estimation model needs to be reduced. Here we evaluate performance of 90 of the 115 proposed HCRs using a closed-loop MSE simulation with no estimation model to help guide selection of which HCRs to test further considering estimation error.

**Table 1.** List of harvest control rules (HCRs) for harvest strategy 1a and 1b. The target reference point ($F_{target}$) is an indicator of fishing intensity based on SPR. SPR is the spawning stock biomass (SSB) per recruit that would result from the current year's pattern and intensity of fishing mortality relative to the unfished stock. An $F_{target}$ of FSPR40% is associated with a fishing intensity that would leave 40% of the SSB per recruit as compared to the unfished state. An $F_{target}$ of FSPR30% implies a higher fishing intensity (i.e., 1-SPR of 0.7) and would result in a SSB per recruit of 30% of the unfished SPR. The threshold and limit reference points are SSB-based and refer to the specified percentage of unfished SSB ($SSB_{F=0}$). The minimum F refers to the fraction of the $F_{target}$ that the fishing intensity is set to when SSB is below the limit reference point.

| HCR Type | HCR # | Limit Reference Point | Threshold Reference Point | Target Reference Point | Minimum F |
|---|---|---|---|---|---|
| 1a or 1b | 1 | 5%$SSB_{F=0}$ | 15%$SSB_{F=0}$ | FSPR15% | 5%$F_{target}$ |
| | 2 | 5%$SSB_{F=0}$ | 15%$SSB_{F=0}$ | FSPR20% | 5%$F_{target}$ |
| | 3 | 5%$SSB_{F=0}$ | 15%$SSB_{F=0}$ | FSPR30% | 5%$F_{target}$ |
| | 4 | 5%$SSB_{F=0}$ | 15%$SSB_{F=0}$ | FSPR40% | 5%$F_{target}$ |
| | 5 | 5%$SSB_{F=0}$ | 20%$SSB_{F=0}$ | FSPR20% | 5%$F_{target}$ |
| | 6 | 5%$SSB_{F=0}$ | 20%$SSB_{F=0}$ | FSPR30% | 5%$F_{target}$ |
| | 7 | 5%$SSB_{F=0}$ | 20%$SSB_{F=0}$ | FSPR40% | 5%$F_{target}$ |
| | 8 | 5%$SSB_{F=0}$ | 25%$SSB_{F=0}$ | FSPR30% | 5%$F_{target}$ |
| | 9 | 5%$SSB_{F=0}$ | 25%$SSB_{F=0}$ | FSPR40% | 5%$F_{target}$ |

| | | | | | |
|---|---|---|---|---|---|
| | 10 | 7.7%$SSB_{F=0}$ | 15%$SSB_{F=0}$ | FSPR15% | 5%$F_{target}$ |
| | 11 | 7.7%$SSB_{F=0}$ | 15%$SSB_{F=0}$ | FSPR20% | 5%$F_{target}$ |
| | 12 | 7.7%$SSB_{F=0}$ | 15%$SSB_{F=0}$ | FSPR30% | 5%$F_{target}$ |
| | 13 | 7.7%$SSB_{F=0}$ | 15%$SSB_{F=0}$ | FSPR40% | 5%$F_{target}$ |
| | 14 | 7.7%$SSB_{F=0}$ | 20%$SSB_{F=0}$ | FSPR20% | 5%$F_{target}$ |
| | 15 | 7.7%$SSB_{F=0}$ | 20%$SSB_{F=0}$ | FSPR30% | 5%$F_{target}$ |
| | 16 | 7.7%$SSB_{F=0}$ | 20%$SSB_{F=0}$ | FSPR40% | 5%$F_{target}$ |
| | 17 | 7.7%$SSB_{F=0}$ | 25%$SSB_{F=0}$ | FSPR30% | 5%$F_{target}$ |
| | 18 | 7.7%$SSB_{F=0}$ | 25%$SSB_{F=0}$ | FSPR40% | 5%$F_{target}$ |
| | 19 | 15%$SSB_{F=0}$ | 20%$SSB_{F=0}$ | FSPR20% | 5%$F_{target}$ |
| | 20 | 15%$SSB_{F=0}$ | 20%$SSB_{F=0}$ | FSPR30% | 5%$F_{target}$ |
| | 21 | 15%$SSB_{F=0}$ | 20%$SSB_{F=0}$ | FSPR40% | 5%$F_{target}$ |
| | 22 | 15%$SSB_{F=0}$ | 25%$SSB_{F=0}$ | FSPR30% | 5%$F_{target}$ |
| | 23 | 15%$SSB_{F=0}$ | 25%$SSB_{F=0}$ | FSPR40% | 5%$F_{target}$ |
| | 24 | 20%$SSB_{F=0}$ | 25%$SSB_{F=0}$ | FSPR30% | 5%$F_{target}$ |
| | 25 | 20%$SSB_{F=0}$ | 25%$SSB_{F=0}$ | FSPR40% | 5%$F_{target}$ |
| | 26 | 5%$SSB_{F=0}$ | 15%$SSB_{F=0}$ | FSPR15% | 10%$F_{target}$ |
| | 27 | 5%$SSB_{F=0}$ | 15%$SSB_{F=0}$ | FSPR20% | 10%$F_{target}$ |
| | 28 | 5%$SSB_{F=0}$ | 15%$SSB_{F=0}$ | FSPR30% | 10%$F_{target}$ |
| | 29 | 5%$SSB_{F=0}$ | 15%$SSB_{F=0}$ | FSPR40% | 10%$F_{target}$ |
| | 30 | 5%$SSB_{F=0}$ | 20%$SSB_{F=0}$ | FSPR20% | 10%$F_{target}$ |
| | 31 | 5%$SSB_{F=0}$ | 20%$SSB_{F=0}$ | FSPR30% | 10%$F_{target}$ |
| | 32 | 5%$SSB_{F=0}$ | 20%$SSB_{F=0}$ | FSPR40% | 10%$F_{target}$ |
| | 33 | 5%$SSB_{F=0}$ | 25%$SSB_{F=0}$ | FSPR30% | 10%$F_{target}$ |
| | 34 | 5%$SSB_{F=0}$ | 25%$SSB_{F=0}$ | FSPR40% | 10%$F_{target}$ |
| | 35 | 7.7%$SSB_{F=0}$ | 15%$SSB_{F=0}$ | FSPR15% | 10%$F_{target}$ |
| | 36 | 7.7%$SSB_{F=0}$ | 15%$SSB_{F=0}$ | FSPR20% | 10%$F_{target}$ |
| | 37 | 7.7%$SSB_{F=0}$ | 15%$SSB_{F=0}$ | FSPR30% | 10%$F_{target}$ |
| | 38 | 7.7%$SSB_{F=0}$ | 15%$SSB_{F=0}$ | FSPR40% | 10%$F_{target}$ |
| | 39 | 7.7%$SSB_{F=0}$ | 20%$SSB_{F=0}$ | FSPR20% | 10%$F_{target}$ |

| | | | | | |
|---|---|---|---|---|---|
| | 40 | 7.7%$SSB_{F=0}$ | 20%$SSB_{F=0}$ | FSPR30% | 10%$F_{target}$ |
| | 41 | 7.7%$SSB_{F=0}$ | 20%$SSB_{F=0}$ | FSPR40% | 10%$F_{target}$ |
| | 42 | 7.7%$SSB_{F=0}$ | 25%$SSB_{F=0}$ | FSPR30% | 10%$F_{target}$ |
| | 43 | 7.7%$SSB_{F=0}$ | 25%$SSB_{F=0}$ | FSPR40% | 10%$F_{target}$ |
| | 44 | 15%$SSB_{F=0}$ | 20%$SSB_{F=0}$ | FSPR20% | 10%$F_{target}$ |
| | 45 | 15%$SSB_{F=0}$ | 20%$SSB_{F=0}$ | FSPR30% | 10%$F_{target}$ |
| | 46 | 15%$SSB_{F=0}$ | 20%$SSB_{F=0}$ | FSPR40% | 10%$F_{target}$ |
| | 47 | 15%$SSB_{F=0}$ | 25%$SSB_{F=0}$ | FSPR30% | 10%$F_{target}$ |
| | 48 | 15%$SSB_{F=0}$ | 25%$SSB_{F=0}$ | FSPR40% | 10%$F_{target}$ |
| | 49 | 20%$SSB_{F=0}$ | 25%$SSB_{F=0}$ | FSPR30% | 10%$F_{target}$ |
| | 50 | 20%$SSB_{F=0}$ | 25%$SSB_{F=0}$ | FSPR40% | 10%$F_{target}$ |
| 2 | 1 | 5%$SSB_{F=0}$ | | FSPR10% | |
| | 2 | 5%$SSB_{F=0}$ | | FSPR15% | |
| | 3 | 5%$SSB_{F=0}$ | | FSPR20% | |
| | 4 | 5%$SSB_{F=0}$ | | FSPR30% | |
| | 5 | 5%$SSB_{F=0}$ | | FSPR40% | |
| | 6 | 7.7%$SSB_{F=0}$ | | FSPR10% | |
| | 7 | 7.7%$SSB_{F=0}$ | | FSPR15% | |
| | 8 | 7.7%$SSB_{F=0}$ | | FSPR20% | |
| | 9 | 7.7%$SSB_{F=0}$ | | FSPR30% | |
| | 10 | 7.7%$SSB_{F=0}$ | | FSPR40% | |
| | 11 | 15%$SSB_{F=0}$ | | FSPR20% | |
| | 12 | 15%$SSB_{F=0}$ | | FSPR30% | |
| | 13 | 15%$SSB_{F=0}$ | | FSPR40% | |
| | 14 | 20%$SSB_{F=0}$ | | FSPR30% | |
| | 15 | 20%$SSB_{F=0}$ | | FSPR40% | |

**Methods**

The JWG put forward three potential model-based HCRs: HCR1a, HCR1b, and HCR2 (WCPFC 2019, Fig. 1). The HCRs specify a fishing intensity (F, 1-SPR), where SPR is the spawning stock biomass (SSB) per recruit that would result from the current year's pattern and intensity of fishing mortality relative to the unfished stock, based on a comparison of the current spawning stock biomass (SSB) relative to biomass-based reference points. F is at the $F_{target}$ if biomass is above the biomass-based threshold reference (ThRP) for HCR1a and 1b, or above the limit reference point (LRP) for HCR2. HCRs start ramping down F at either the ThRP (HCR1a and 1b) or the LRP. HCR1a and 1b decline F down to the LRP to a minimum level of F ($F_{min}$) while under HCR2 F declines down to 0. The decline in F for HCR1b follows a sigmoidal curve, but it's linear for the other HCRs. The equations detailing, for each HCR, how F changes in relation to stock status are described in Table 2. We use the MSE framework presented in Tommasi and Lee 2022 and run the simulation with no assessment model error (i.e. no estimation model) to reduce run times. The Tommasi and Lee (2022) framework and associated code (available at https://github.com/detommas/PBF_MSE) was expanded to include R scripts to run any of these HCRs and also modified according to feedback from the November ISC PBF WG meeting to:

- Set a TAC every three rather than two years, following a simulated three-year assessment schedule,
- Run the feedback control MSE simulation for 24 rather than 30 years,
- For all fleets, keep selectivity constant at the 2017-2019 values for the forward simulation. These are the selectivity values also used in the benchmark calculations to compute the F multiplier required to keep fishing intensity at the target level.

We note that the code was also modified to use the Lee et al. 2021 bootstrap correction when generating data from the OM to input into the EM, but this new capability was not used in this analysis as the simulation was run assuming no estimation error.

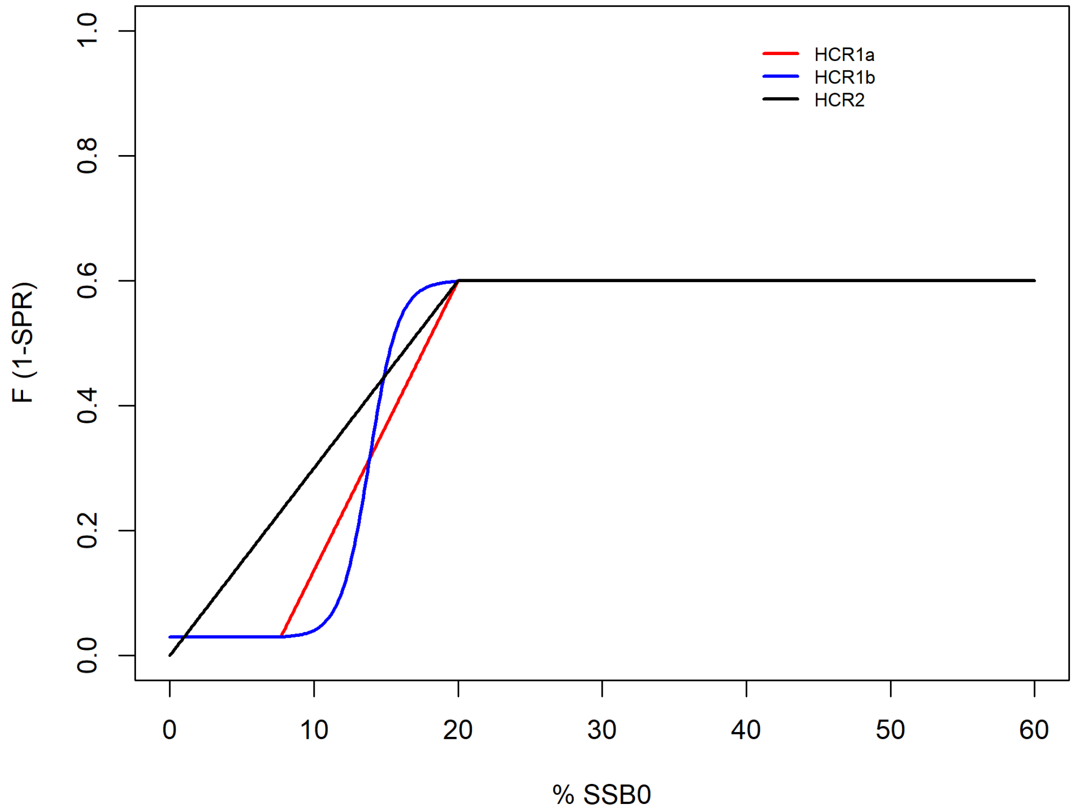


**Figure 1.** Example harvest control rules (HCRs) 1a, 1b, and 2 proposed by the JWG (WCPFC 2019). In this example, all HCRs have a target reference point of F40. This corresponds to a 1-SPR fishing intensity of 0.6 that would generate 40% of the unfished spawning potential. HCR1a and 1b have a threshold reference point of 20% of unfished SSB ($SSB_0$) and a limit reference point (LRP) of 7.7%$SSB_0$. HCR2 has an LRP of 20%$SSB_0$.

**Table 2.** Details of candidate harvest controls at specific current spawning stock biomass ($SSB_{current}$) relative to SSB reference points proposed for the PBF MSE. ThRP is the SSB based threshold reference point, LRP is the SSB-based limit reference point, and $F_{target}$ is the target reference point.

| Stock Status | HCR | Fishing Intensity (1-SPR) |
|---|---|---|
| $SSB_{current} \geq ThRP$ | 1a or 1b | $F = F_{target}$ |
| $LRP < SSB_{current} < ThRP$ | 1a | $F = (F_{target} - F_{min}) * (SSB_{current} - LRP)/(ThRP - LRP) + F_{min}$ |
| | 1b | $F = ((F_{target} /(1+exp(-(SSB_{current} - (((ThRP-LRP)/2)+LRP)))))+F_{min}) * (F_{target} /(F_{target} + F_{min}))$ |

| | | |
|---|---|---|
| $SSB_{current}$ < LRP | 1a or 1b | $F_{min}$ |
| $SSB_{current} \geq$ LRP | 2 | $F = F_{target}$ |
| $SSB_{current}$ < LRP | 2 | $F = (F_{target}/LRP)* SSB_{current}$ |

We test all 65 HCRs of type 1a and 2. However, for HCR1b we only test the 25 HCRs using the more aggressive $F_{min} = 10\%F_{target}$ (HCR1b 26-50, Table 1) since for HCRs 1a there was no large difference in performance between the two different candidate $F_{min}$ (see Results section).

As described in Tommasi and Lee (2022), the PBF MSE uses a modified version of the short 2022 Stock Synthesis (SS) PBF stock assessment model (Fukuda et al. 2022) as the base case operating model (OM). The OM has been conditioned using historical data and is run with no estimation using parameters set in the .par file during the forward simulation. Catches in the OM .dat file are updated every three years as set by the TAC determined by the HCR. Thus, in the 24-year simulations a TAC is set eight times. For each of the 90 HCRs, we run 100 different iterations to account for recruitment process uncertainty, for a total of 9,000 24-years runs.

Output from the OM, for each year of the MSE simulation, is used to calculate the performance metrics proposed at JWG07 and outlined in Table 3. Note that the fishery impact management objective is dependent on the relative exploitation pattern (relative fishing mortality across fleet), which is set to the 2017-2019 average by design (Tommasi and Lee 2022). Therefore, the fishery impact performance metrics do not vary widely by HCR (Tommasi and Lee 2023) and are not presented here. Since catch for the first three years of the simulation is set to the CMM catch limits and the HCR starts being applied in 2024 (Tommasi and Lee 2022), we calculate performance metrics using output from 2024 onwards.

**Table 3.** List of operational management objectives and performance metrics for Pacific Bluefin tuna generated during JWG07 and to be revised at JWG08 (JWG07 Annex E). SSB refers to female spawning stock biomass, LRP to limit reference point. F is the fishing intensity (1-SPR) and $F_{target}$ is the target reference point.

| Category | Operational Management Objective | Performance Metric |
|---|---|---|
| Safety | There should be a less than [5-20%] 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 [50-75]% probability | Probability that $F \leq F_{target}$ in any given year of the evaluation period |
| Stability | To limit changes in overall catch limits between management periods to no more than [15%] downwards [unless the ISC has assessed that there is a greater than 50%chance the stock is below the LRP] | Percent change upwards in catches between management periods excluding periods when SSB<LRP<br><br>Percent change downwards in catches between management periods excluding periods when SSB<LRP |
| Yield | [Maintain a proportional fishery impact between the WCPO and EPO [similar to the average proportional fishery impact from1971-1994] | Median fishery impact (in %) on SSB in any given year of the evaluation period by fishery and by WCPO fisheries and EPO fisheries<br><br>The probability that the proportional EPO fishery impact is at least the1971-1994 average in any given year |
| | To maximize yield over the medium (5-10 years) and long (10-30 years) terms, as well as average annual catch yield from the fishery. | Expected annual yield over years5-10 of the evaluation period, by fishery.<br><br>Expected annual yield over years10-30 of the evaluation period, by fishery.<br><br>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] | Expected annual yield in any given year of the evaluation period |

## Results

### *Safety*

There is no adopted LRP for PBF and different candidate HCRs (Table 1) have different candidate LRPs. We first compute the Safety performance metric relative to the LRP specified in each HCR. Most HCRs are able to maintain SSB at or below the LRP with a probability of less than 20% (Fig. 2). The exceptions are HCRs 6 and 11 for HCR type 2. These HCRs have an LRP that is relatively close to the SSB associated with the $F_{target}$ (Table 3), leading to a higher probability that the LRP be crossed under natural recruitment variability. For instance, HCR11 has an LRP of 15%$SSB_0$ and an $F_{target}$ of FSPR20%. By contrast, HCR19 for HCR type 1a and HCR44 for type 1a and 1b have the same LRP and $F_{target}$ as HCR11 type 2, but the presence of a ThRP reduces F before the LRP is reached, lowering the probability of it being breached as compared to the type 2 HCR (Fig. 2).

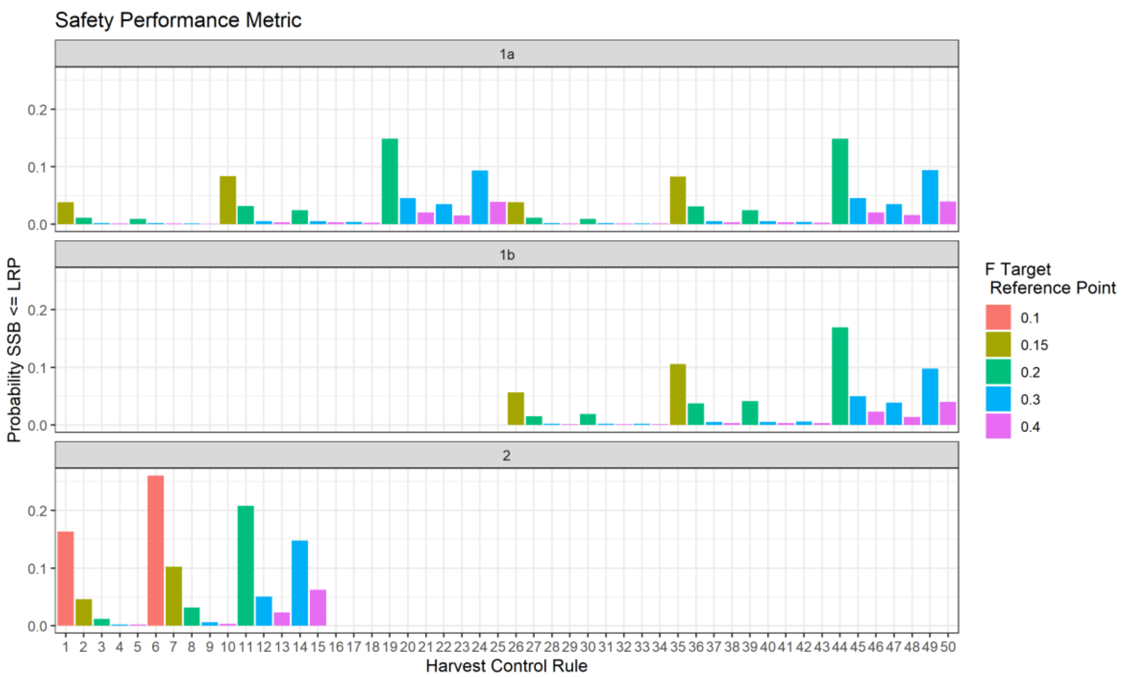


**Figure 2.** Plot of the safety performance metric, the probability in any given year of the simulation of spawning stock biomass (SSB) being below the limit reference point (LRP) as specified in each harvest control rule (HCR), calculated across 100 iterations. Panels show results for each HCR type: 1a, 1b, and 2.

Since different HCRs have different LRPs, the performance metric as computed above does not show performance across a common level of safety. We therefore compute the

Safety performance metric (Table 3) also relative to two potential safety thresholds of 20%$SSB_0$, the second rebuilding target for PBF, and 7.7%$SSB_0$, IATTC's interim LRP for tropical tunas. As expected, HCRs with an $F_{target}$ of FSPR30% or FSPR40% are the only ones with a less than 20% probability of SSB being at or below 20%$SSB_0$. This is because, on average, SSB will be at the SSB associated with the $F_{target}$. Indeed, HCRs with a higher $F_{target}$ have a higher median SSB across all iterations and simulation years (Fig. 4).

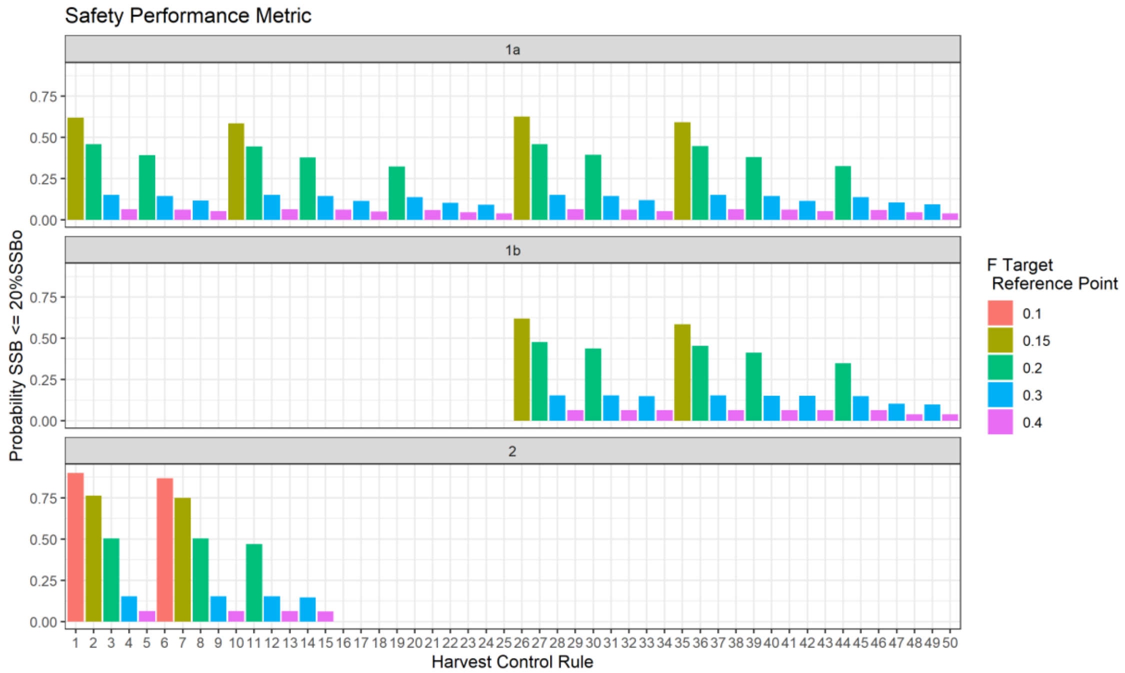


**Figure 3.** Plot of the safety performance metric computed as the probability in any given year of the simulation of spawning stock biomass (SSB) being below the 2nd rebuilding target of 20%$SSB_0$, calculated across 100 iterations. Panels show results for each HCR type: 1a, 1b, and 2.

All HCRs except HCR1 and HCR6 for type 2, the only ones with an $F_{target}$ of FSPR10%, have a probability of SSB being at or below a safety threshold of 7.7%$SSB_0$ higher than 20% (Fig. 5).

We note also that, notwithstanding the way the safety performance metric is presented, HCR type 1a and 1b perform similarly and, for HCR1a there is no large difference in performance when using a $F_{min}$ fraction of 5% or 10% given the same reference points.

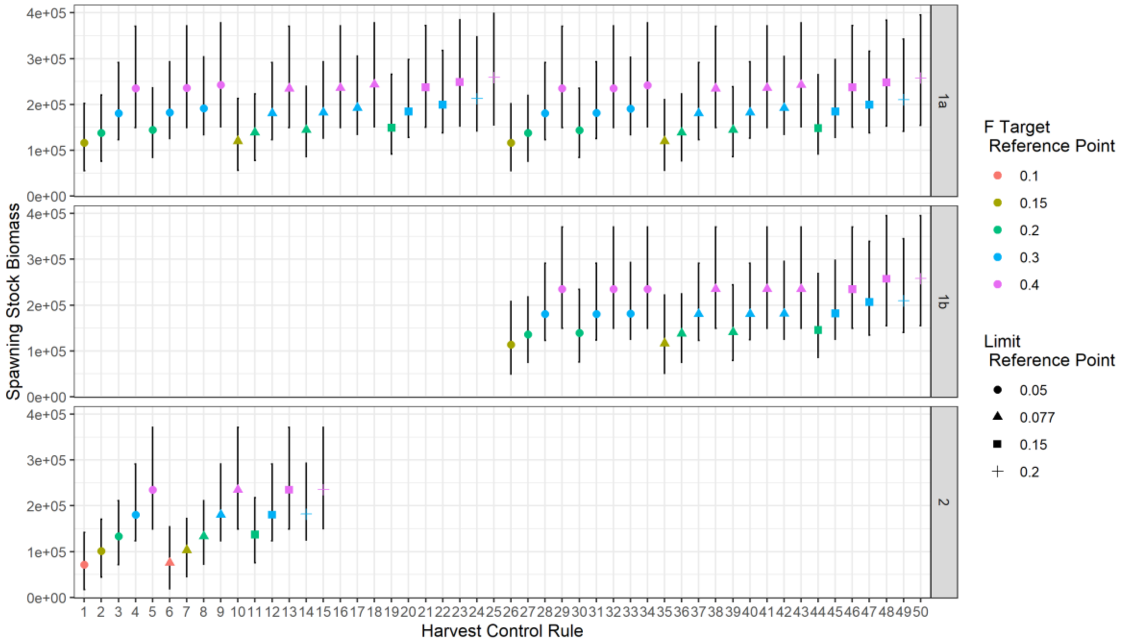


**Figure 4.** Median spawning stock biomass across 100 iterations and all years of the forward simulation for each HCR. The vertical bars represent the 5th to 95th quantile range.

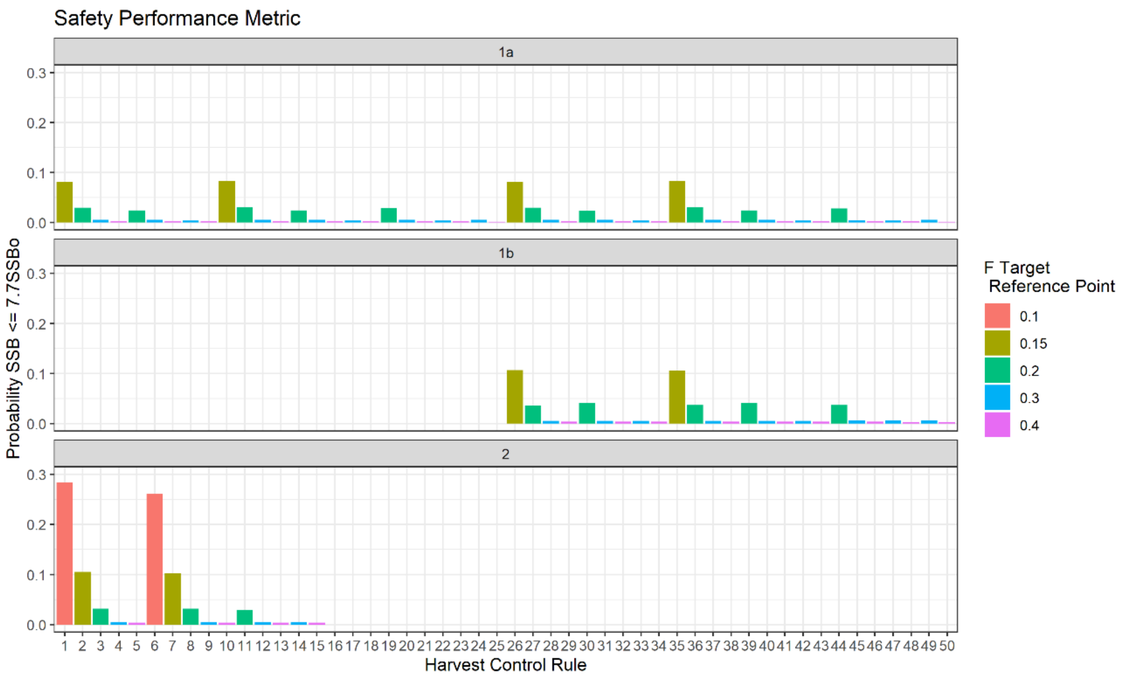


**Figure 5.** Plot of the safety performance metric computed as the probability in any given year of the simulation of spawning stock biomass (SSB) being 7.7%$SSB_0$, calculated across 100 iterations. Panels show results for each HCR type: 1a, 1b, and 2.

*Status*

The average level of fishing intensity (1-SPR) is set by the $F_{target}$ and thus HCRs with a higher $F_{target}$ have a higher median fishing intensity across all iterations and simulation years (Fig. 6). Figure 6 also shows that HCRs of type 1a and 1b with an $F_{target}$ of FSPR15% have a wider spread due to more management intervention. HCRs 19, 24, 25, 44, 49, and 50 also have a wider range as, for a given $F_{target}$ and ThRP, they have the steepest reduction in F down to the LRP. We note also that type 1b HCRs with $F_{target}$ FSPR15% or FSPR20% are more variable than corresponding type 1a HCRs. Type 2 HCRs have a gentler decline when the control point is breached and have thus lower variability in F than the other HCRs (Fig. 6).

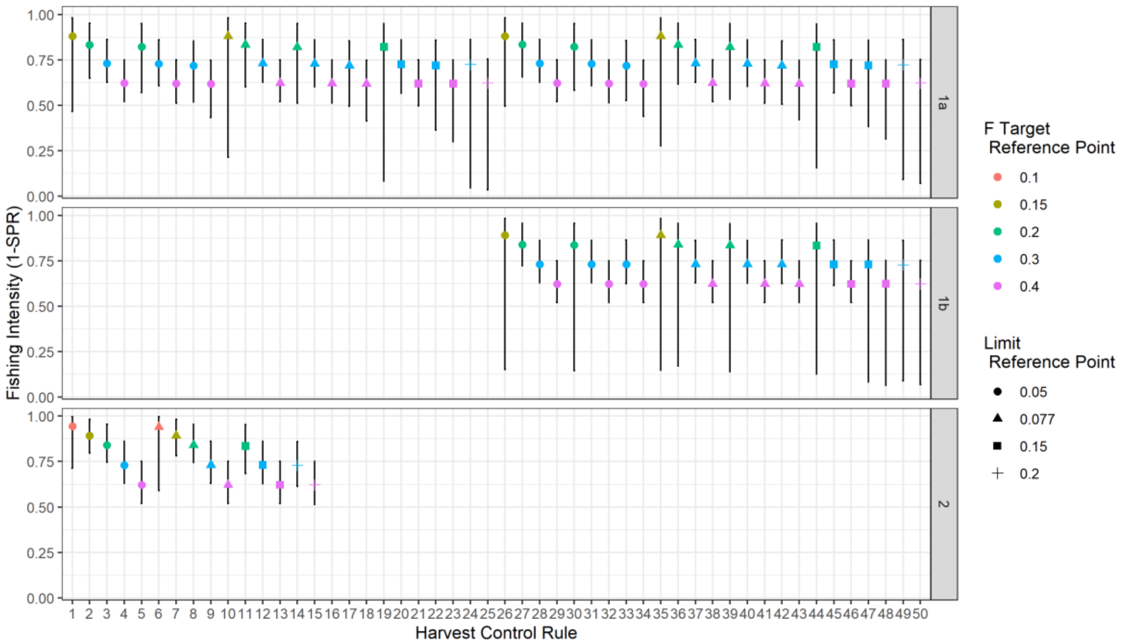


**Figure 6.** Median fishing intensity(1-SPR) across 100 iterations and all years of the forward simulation for each HCR. The vertical bars represent the $5^{th}$ to $95^{th}$ quantile range.

Without recruitment variability, we would expect the F to remain at the $F_{target}$ after an initial adjustment from the 2020 starting age structure. To make sure the MSE operating model responds as expected, we run HCR4 type 1a without recruitment deviations and see that the F stabilizes in the last 10 years of the simulation (Fig. 7) to an F of 0.61, close to the 0.6 expected with an $F_{target}$ of FSPR40%. We therefore compute the status performance metric over the last 10 years of the simulation. We chose this HCR for this check as it has a high $F_{target}$ and a low ThRP (15%$SSB_0$), which is never breached so F does not decline because of management intervention.

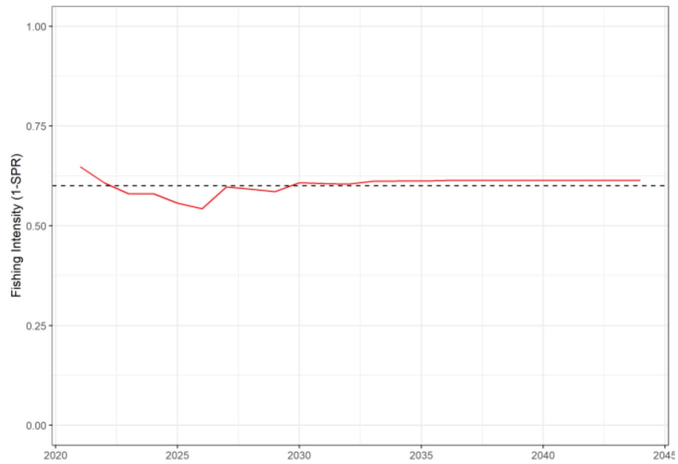


**Figure 7.** Time series of fishing intensity (1-SPR) with no recruitment variability for HCR4 of type 1a, which has an $F_{target}$ of FSPR40%. The dotted line represents the target fishing intensity of 0.6.

All HCRs have a probability of at least 50% of F being at or below the $F_{target}$ (Fig. 8). In general, HCRs with a lower reference point have a higher probability of F being at or below the $F_{target}$ because of the higher management intervention, which sets the F below the $F_{target}$. As for the Safety performance metric, HCR type 1a and 1b perform relatively similarly and, for HCR1a there is no large difference in performance when using a $F_{min}$ fraction of 5% or 10% given the same LRP, ThRP, and $F_{target}$ (Fig. 8).

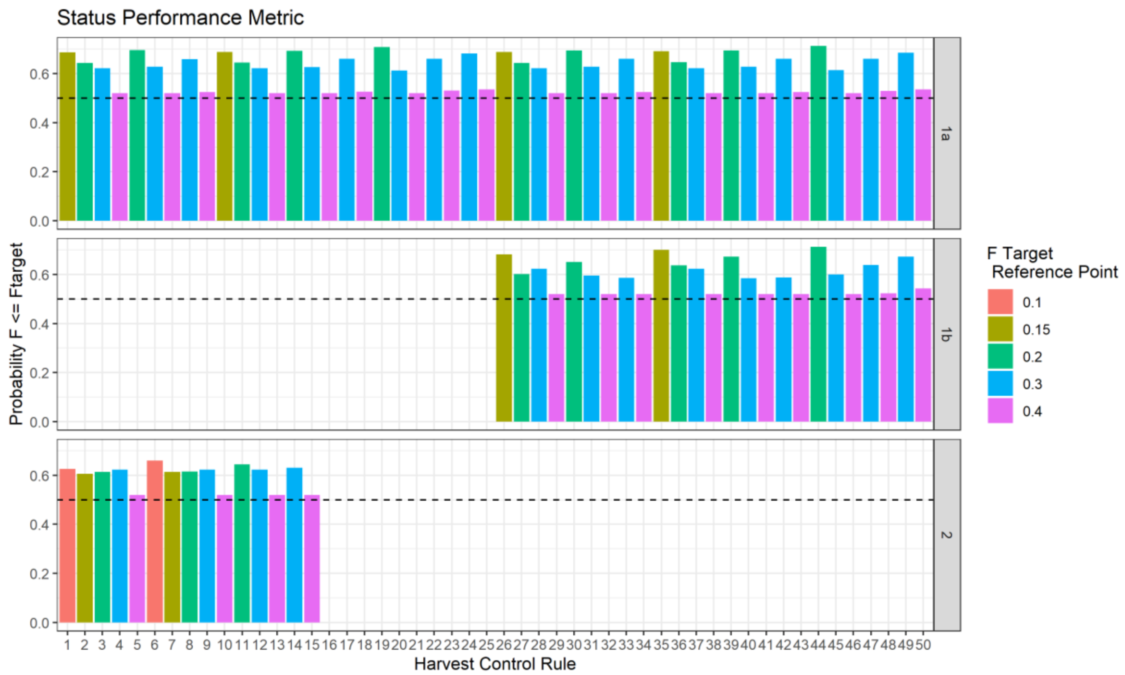

**Figure 8.** Plot of the status performance metric computed as the probability in any given year of the simulation of fishing intensity (F, 1-SPR) being at or below the $F_{target}$ specified in each HCR, calculated across 100 iterations. Panels show results for each HCR type: 1a, 1b, and 2.

*Stability*

Median % downward change in catch between the three-year management periods ranges between 9 to 27% for HCR1a, 9 to 24% for HCR1b, and 9 to 24% for HCR2 (Fig. 9). Median % downward change in catch is largest for HCRs with a lower $F_{target}$ and these HCRs also show a higher variability in % downward change in catch (Fig. 9). For the same reference points, HCRs of type 1b have a slightly higher median % downward change in catch as compared to HCRs of type 1a, but have a higher variability (Fig. 9). Type 2 HCRs have the lowest variability in % downward change in catch as the required change in F once the management control point is crossed is more gradual. Note that, even without considering instances when SSB<LRP, no HCR has a maximum % downward change in catch that is 15% or less. For instance, even HCR4, which maintains SSB above the ThRP and LRP, has a maximum % downward change in catch of 43% due to the feedback on biomass of maintaining a constant TAC for three years under natural recruitment variability.

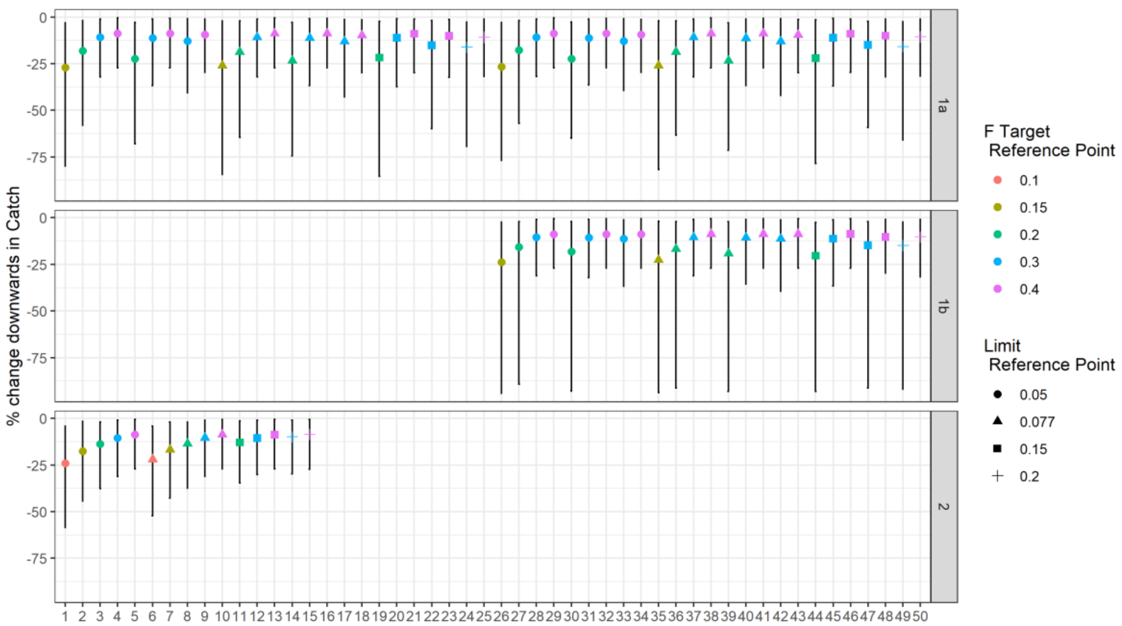


**Figure 9.** Median percent change downwards in annual catch between management periods, excluding periods when SSB<LRP, across 100 iterations and all years of the

forward simulation for each HCR. The vertical bars represent the 5th to 95th quantile range.

*Yield*

As for the spawning stock biomass, the largest differences in annual catch were associated with different $F_{target}$ reference points. Median annual catch over the entire simulation period was highest, but most variable, for HCRs with the highest $F_{target}$ (Fig. 10). Given the same ThRP and $F_{target}$, median catch was highest for HCRs with the highest LRP as biomass was maintained higher, but this was at the cost of more management intervention and higher catch variability (Fig. 9 Type 1a and 1b, e.g. compare HCR24 to 8, 17, 22). For the same reference points, catch was more variable for type 1b HCRs as compared to 1a. Type 2 HCRs, for the same LRP and $F_{target}$, had lower, but less variable, median catch (Fig. 9).

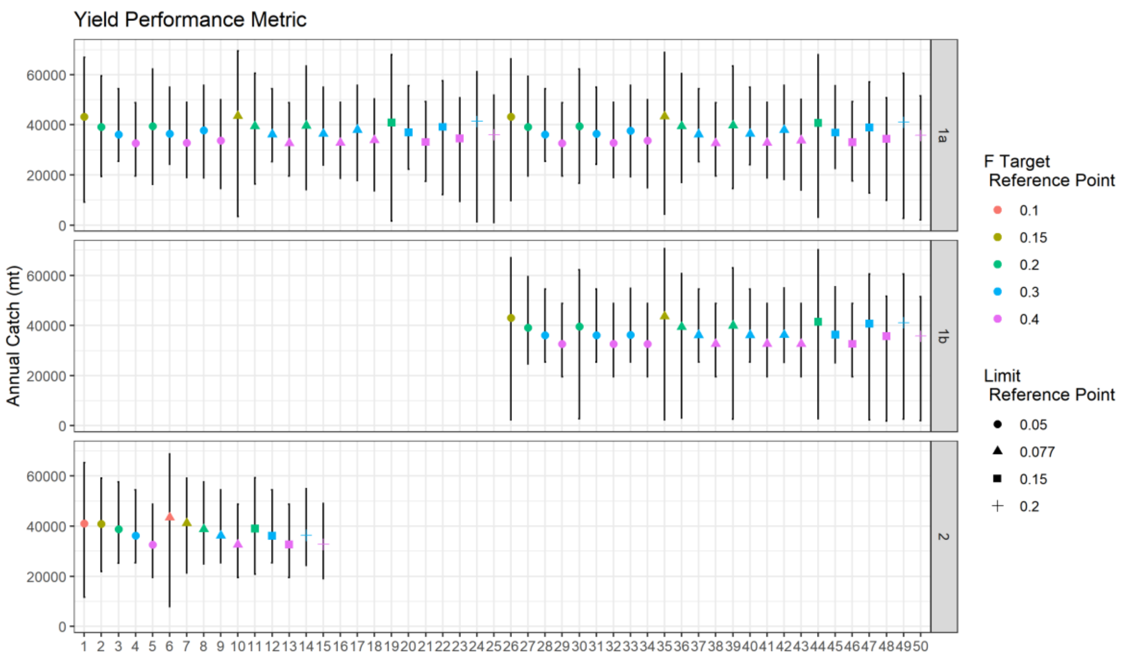


**Figure 10.** Median annual catch across 100 iterations and all years of the forward simulation for each HCR by HCR type. The vertical bars represent the 5th to 95th quantile range.

As for annual catch, medium term catch is also lowest for HCRs with the lowest $F_{target}$, (Fig. 11) but, for type 1a and 1b HCRs, HCRs with the lowest $F_{target}$ of F15 perform more poorly in terms of medium term catch than HCRs with an $F_{target}$ of F20 and comparably to HCRs with an $F_{target}$ of F30 (Fig. 11). Similarly, for HCR2, F10 HCRs perform more poorly than F15, F20, and F30 HCRs (Fig. 11). The difference in performance between

HCRs with different $F_{target}$ reference points of type 1a and 1b is even less stark, but HCRs with F15 and F20 $F_{target}$ have the most variable long-term catch (Fig. 12). This is because while annual catch is on average higher with a higher $F_{target}$, catch is also more variable, leading to comparable catch when averaged over the long term. For HCR2, long term catch is highest for HCRs with an $F_{target}$ of F15 or F20 (Fig. 12).

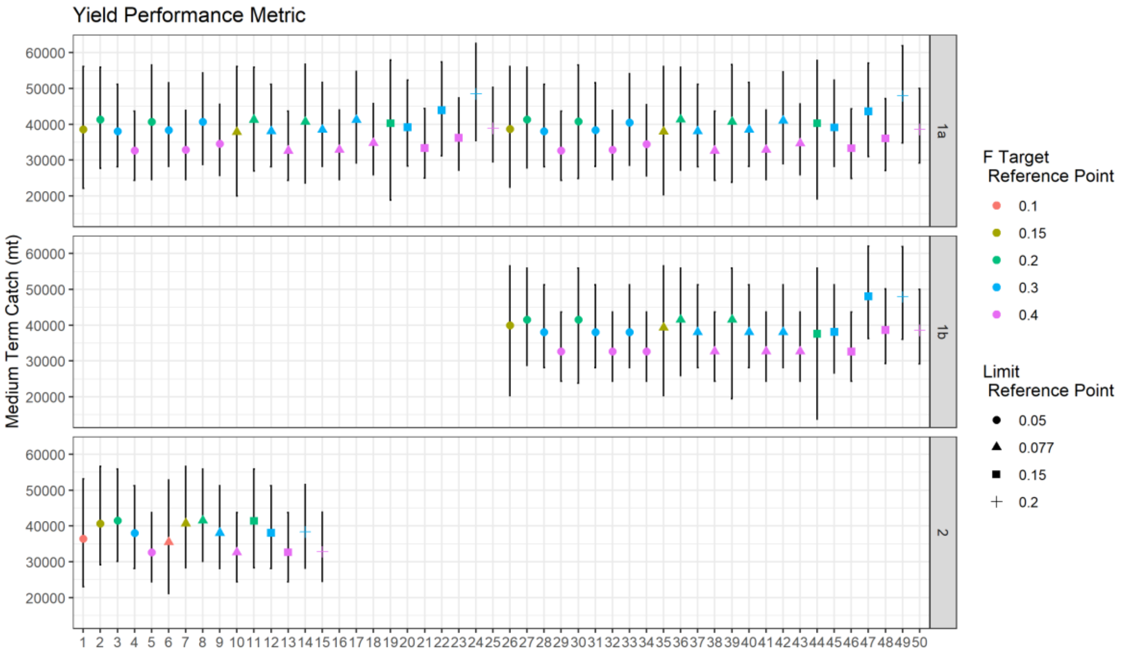


**Figure 11.** Median medium term catch across 100 iterations for each HCR by HCR type. The vertical bars represent the 5th to 95th quantile range. Medium term catch is the mean catch over years 5-10 of the forward simulation.

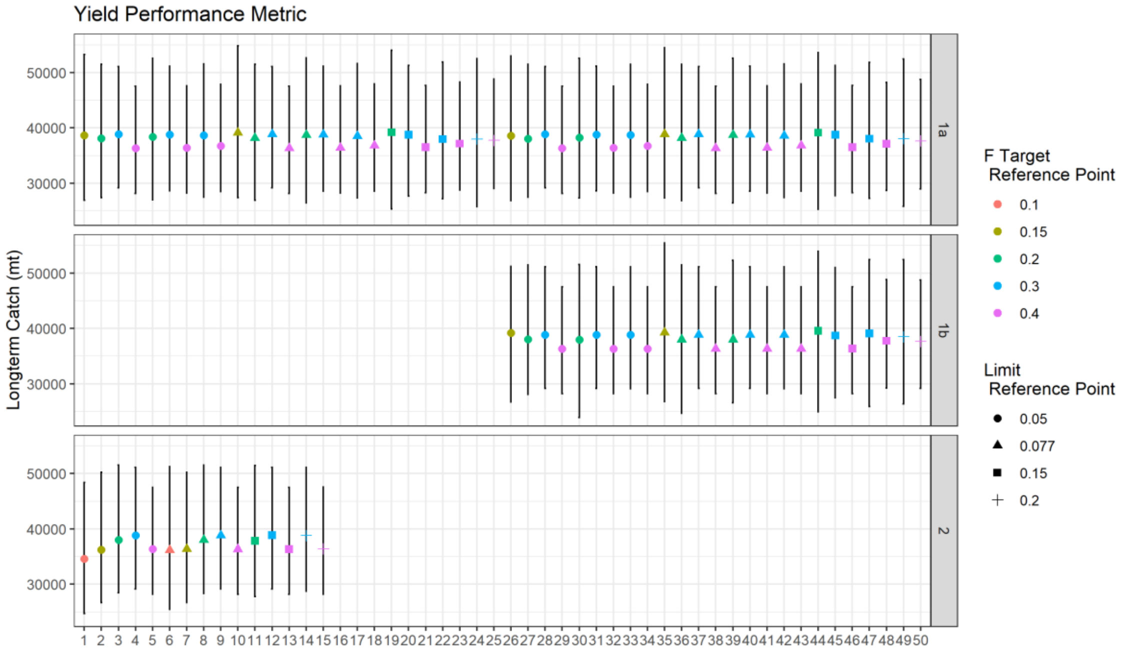


**Figure 12.** Median long term catch across 100 iterations for each HCR by HCR type. The vertical bars represent the 5th to 95th quantile range. Long term catch is the mean catch over years 10-24 of the forward simulation.

*Tradeoffs between management objectives*

Tradeoffs were evident between performance metrics, particularly between Yield and Safety and Yield and Stability. HCRs performing best in terms of the annual catch Yield metric, performed worst in terms of Stability (Fig. 13, panels a). However, when looking at patterns across HCRs sharing the same $F_{target}$ (i.e. same color symbols on Fig. 13, panel a), HCRs that performed better in terms of Yield also performed well in terms of Safety.

There was also a tradeoff between the annual catch Yield metric and Stability, with F15 rules performing best in Yield but worst in Stability and F40 rules performing worst in Yield but worst in Stability (Fig. 13, panel d). Furthermore, many HCRs, particularly with F20 or F30 $F_{targets}$, performed similarly in terms of Yield, but showed more drastic differences in terms of stability (Fig. 13, panel d).

While there was an overall positive relationship between Safety and Stability metrics (Fig. 13, panel b), within HCRs sharing the same $F_{target}$ (i.e. same color symbols on Fig. 13, panel b) the pattern was reversed and HCRs with the higher Safety performance performed worst in terms of Stability.

When considering the Yield long term performance metric, trade-offs with Safety and Stability where not as clear-cut as the relationship was non-linear (Fig. 13, panels c and f). HCRs performing best in terms of long-term Yield had an intermediate performance in terms of Safety or Stability (Fig. 13, panels c and f).

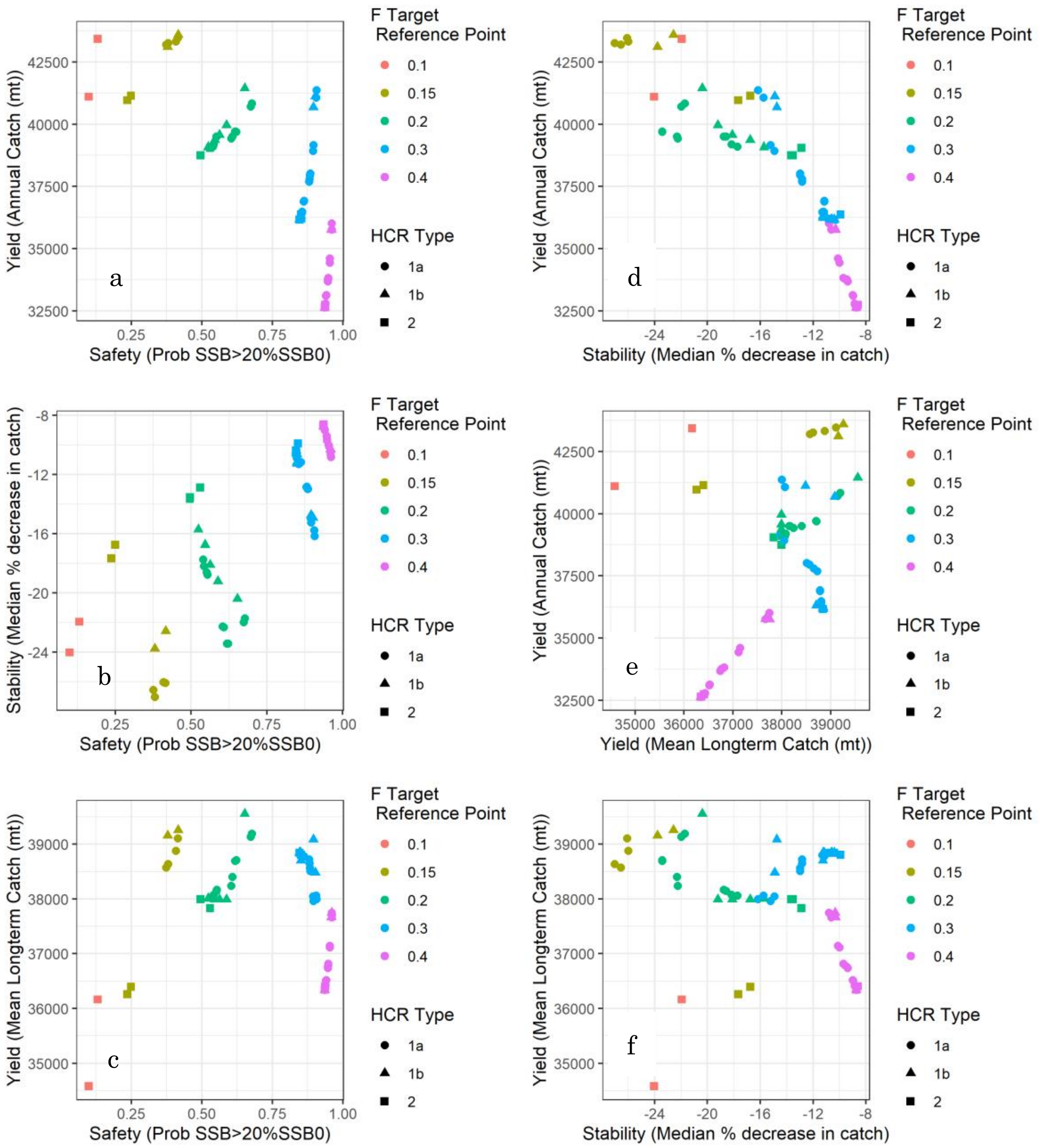


**Figure 13.** Scatter plots comparing performance across different metrics for each HCR. Safety is defined, for each HCR, as the probability of SSB being above the second PBF rebuilding target of 20%SSB0. Yield is measured as either the median annual catch across all iterations and simulation years or as the mean long term catch, which is the

median catch averaged over years 10-24 of the simulation across all iterations. Stability in the median % decrease in catch between management periods.

**Discussion**

We have expanded the capabilities of the PBF MSE framework to include the option of testing all the model-based HCRs put forward so far by the JWG. MSE simulations with only one base case operating model and no estimation error show that all HCRs proposed are able to meet the Status management objective and maintain F around the $F_{target}$ in the last 10 years of the simulation with at least 50% probability. The HCRs are also able to maintain a 20% or less probability of SSB falling below their own LRP, but when Safety is compared against a common level, such as the second rebuilding target of 20%$SSB_0$, then performance across HCRs varies. Furthermore, while some HCRs are able to maintain a median decrease in catch between management periods of 15% or less, all HCRs have a decrease in catch that is higher than 15% in at least one management period and iteration. Thus, if limiting changes in catch to within 15% in any management period is a desirable objective, HCRs with a built-in limit to changes in TAC should be tested.

There exists no single best-performing HCR as there are tradeoffs among management objectives. The $F_{target}$ reference point was the most important determinant of performance across multiple management objectives. HCRs with a higher $F_{target}$ perform best in terms of Safety (measured relative to 20%$SSB_0$) and Stability, but at the cost of lower annual catch. Long term Yield is maximized at intermediate biomass and thus HCRs with the lowest (F10) or highest (F40) $F_{target}$ perform best in terms of long term catch.

Overall, the value of $F_{min}$ did not have a large impact on performance. Furthermore, for the same reference points combination, HCRs of type 1a and 1b performed similarly in terms of Safety, Status, and Yield metrics, but type 1b HCRs had lower Stability than type 1a, particularly for F20 and F30 HCRs. For the same $F_{target}$, type 2 HCRs performed worst than type 1a or 1b in terms of both Yield and Safety, but best in terms of Stability.

HCRs not performing adequately under this best-case scenario are not expected to improve in performance under the final MSE simulation with multiple operating models, estimation, and implementation error. We hope the results here presented can be useful to the PBF WG and the JWG to screen candidate HCRs that underperform relative to the proposed management objectives or that perform similarly to help identify a narrower set of HCRs to be tested in the final phase of the MSE process.

**References**

Fukuda, H., Tsukahara, Y. and Nishikawa, K. 2022. Update of the PBF population dynamics model using short time series data (1983-) and the sensitivity runs for the robustness test. ISC/22/PBFWG-1/06.

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

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. 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

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.

Tommasi, D., Lee, H. 2022. Overview of the Pacific Bluefin tuna management strategy evaluation workflow. ISC/22/PBFWG-2/06.

Tommasi, D., Lee, H., Piner, K. 2023. Calculation of the fishery impact performance metric for the Pacific Bluefin Tuna Management Strategy Evaluation. ISC/23/PBFWG-1/xx.

WCPFC 2019. Commission for the Conservation and Management of Highly Migratory Fish Stocks in the Western and Central Pacific Ocean. Northern Committee Fifteenth Regular Session Annex F Candidate Reference Points and Harvest Control Rules for Pacific Bluefin Tuna. Available at https://meetings.wcpfc.int/meetings/nc15