IOTC Bigeye Tuna Management Procedure Evaluation Update June 2021[1]

Management Procedure Evaluation Status

- Management Procedure (MP) evaluation is being pursued in the strict sense (i.e. as in the International Whaling Commission and Commission for the Conservation of Southern Bluefin Tuna), in which the data to be input into the MP, the analysis, and the Harvest Control Rule (HCR) are all defined in advance and simulation-tested together.
- MPs have evolved incrementally since TCMP-03 (2019), including the addition of a new class of MP that conducts internal projections, in a manner that is analogous to how the Commission might interpret the Kobe 2 Strategy Matrix.
- A set of 6 MPs is presented here to illustrate typical performance. Simulations assumed that the first MP-derived TAC setting would be in 2023 and catches in the period 2020-2022 would be the average reported catch from 2017-2019.
- The bigeye reference set Operating Model (simulator) is being iteratively developed in line with IOTC technical working party requests (WPTT and WPM). Since no changes were requested for the Operating Model in 2020, the next stage of the development process requires a formal endorsement of the OM by the Scientific Committee, at which point the selection and implementation of an MP should become the priority. Independent review should be part of this process.
- Continued funding for scientific and technical development is being sought from the Australian Government for two years August 2021 – June 2023.
- The main feedback priorities for TCMP-04 are a review of the MP tuning objectives requested in 2019, and further guidance on secondary management objectives.

Bigeye MP Development Guidance from TCMP-03 (2019)

The tuning objective refers to a single key management objective that the MPs can achieve precisely (e.g. achieving $SB \geq SB_{MSY}$ with a 50% probability by 2024). Tuning objectives commonly relate to a desirable biomass (in terms of the risk of exceeding reference points and/or a rebuilding timeframe). This has a strong influence on the obtainable yield (because biomass risk and attainable catch are closely related). Tuning ensures that candidate MPs are identical with respect to a high priority objective, and makes it easier to select among MPs on the basis of performance with respect to secondary management objectives (e.g. yield and catch stability). Ideally the Commission will have narrowed down the tuning objectives to 1 or 2 before MP selection. This will allow MP developers to focus MP development within an appropriate range. TCMP-03 (2019) retained two of the tuning objectives requested by TCMP-02 (2018):

B2: Pr(Kobe green zone 2030:2034) = 0.6. The stock status is in the Kobe green quadrant over the period 2030-2034 exactly 60% of the time (averaged over all simulations).

[1] D. Kolody & P. Jumppanen, CSIRO, Australia (email: dale.kolody@csiro.au), with guidance from the IOTC Working Party on Methods MSE Task Force (project steering committee). Funding support from the Australian Department of Foreign Affairs and Trade (through FAO) and CSIRO (funding agencies do not necessarily endorse the results).

B3: Pr(Kobe green zone 2030:2034) = 0.7. The stock status is in the Kobe green quadrant over the period 2030-2034 exactly 70% of the time (averaged over all simulations).

TCMP-03 (2019) further recognized the desirability of other MP constraints:

- Total Allowable Catch (TAC) to be set every 3 years (and held constant between settings)
- A maximum of 15% change to the TAC (increase or decrease) relative to the previous TAC

Candidate Management Procedures (MPs)

Six candidate MPs were selected for presentation, one Harvest Control Rule (HCR) functional form from each of 3 MP classes, with control parameters adjusted to achieve the 2 tuning objectives and constraints above:

- **CPUE** – an "empirical" MP that seeks to stabilize the standardized longline CPUE at a target level, that ideally should correspond to a desirable stock size (Figure 1).
- **PT-HS** – a model-based MP, which fits a simple population model, then prescribes the TAC as a hockeystick-shaped function of the estimated stock depletion (Figure 2).
- **PT-Proj** – a model-based MP, which fits a simple population model, then uses internal projections to solve for the constant TAC that is most likely to attain a pre-defined biomass depletion target in a specific number of years (Figure 3).

MP evaluation results are labelled with the abbreviation above and a suffix denoting the tuning objective (e.g. "**PT-HS.B3**" is the model-based hockeystick MP that attains the B3 tuning objective). Results from a broader range of candidate MPs are included in Appendix 3.

Summary of Bigeye Candidate MP Performance with the Reference Set Operating Model

Candidate MP performance statistics are summarized and ranked in Table 1 for the 15 year time window (2023-2037). More detailed summary tables are included in Appendix 1 (including a range of time windows from 1-20 years). Figures 4-10 provide the standard IOTC graphical performance summaries, from which we note:

- The 2 tuning objectives requested by TCMP-03 are easily achievable, appear to be consistent with IOTC objectives, and result in similar MP performance.
- Over the period 2023-2037, all of these MPs suggest >95% chance of maintaining the spawning biomass above the limit reference point (Table 1).
- Over the period 2023-2037, all of these MPs suggest at least a 75% probability that the average catches will be higher than recent catches (Figure 4).
- All of the MPs demonstrate qualitatively similar time series behaviour, with catch tending to increase until 2035, before stabilizing or declining slightly. Biomass tends to increase until 2030, and then tends to decline steadily to the end of the projection period shown (2040).
- The time series behaviour of the model-based MPs is more stable than the CPUE-based MPs. The CPUE-based MPs suggest that biomass is likely to be below the target level by 2040, and still declining.
- The model-based MPs are near or above the biomass target in 2040, with the projection-based MPs showing the most stable behaviour approaching 2040 (Figure 8). The projection-based MPs also have the lowest catch variability.

Bigeye Candidate MP Performance with respect to Robustness Tests

Robustness test results are presented to the TCMP for the first time in 2021. Robustness tests usually describe the expected MP performance under circumstances that are considered to be less likely than the reference set Operating Model, but which are still plausible, and potentially troublesome. The 6 tuned MPs above were applied to 7 robustness test Operating Models. Performance is described in Appendix 2. Most of the robustness tests had an adverse outcome on MP performance, but none of the results suggested severe adverse impacts, and all MPs were similarly affected. e.g. As shown in Figure A2.1 – A2.4, if there are 8 quarters of below average recruitment in the next couple years, this will reduce the probability of the stock status remaining in the Kobe green zone, and will decrease the expected average catches. However, the probability of breaching the spawning biomass limit remains low, and average catches are still estimated to remain above current catches for all MPs.

Feedback Requests for the TCMP

Discussion of the following questions may help guide for the next iteration of MP development:

1) Do the current MP tuning objectives encompass the range of MP performance of potential interest to the Commission?

2) From the 6 MPs described (Table 1, Figures 4-10), which ones have the most desirable behaviour?

3) Would tuning objectives be easier to interpret and communicate if they were expressed in different units? e.g. As an alternative, the current tuning levels could be approximately re-expressed in terms of the 15 year averaged Kobe plot (Figure 6). i.e. If the tuning objective was redefined such that Pr(mean(B(2023:2037)/B(MSY)) > 1.20) = 0.5, all of the median points in Figure 6 would be located at exactly the same SB/SB_{MSY}.

"CPUE" class MPs

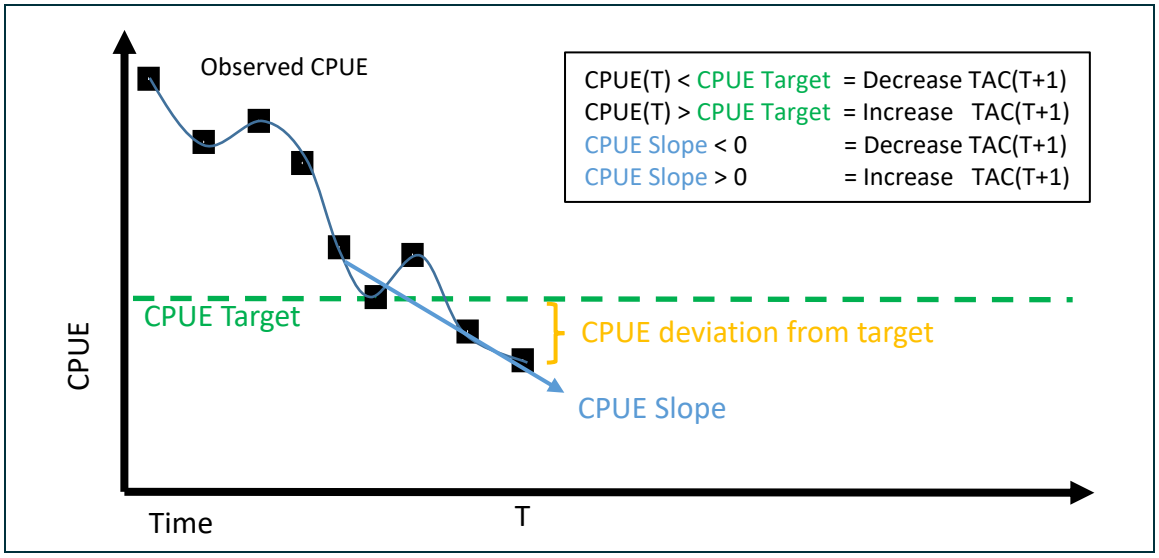

Figure 1. The CPUE class of MPs attempt to manage the fishery to achieve a target value of standardized longline CPUE (annual, regionally-averaged). The next TAC is increased relative to the current TAC if current CPUE is above the target CPUE and the CPUE trend is increasing. Conversely, the next TAC is decreased relative to the current TAC if current CPUE is below the target CPUE and the CPUE trend is decreasing. If the CPUE location relative to the target and CPUE slope are in opposite directions, the TAC change could be in either direction, depending on the magnitude of these indicators, and the associated control parameters. Control parameters include: 1) the number of years in the CPUE slope calculation, 2) responsiveness to CPUE target deviation, 3) responsiveness to CPUE slope and 4) the CPUE target (the tuning parameter in this case). The TAC change constraint will also affect MP behaviour.

"PT-HS" - Model-based MPs with a "hockeystick" type of Harvest Control Rule

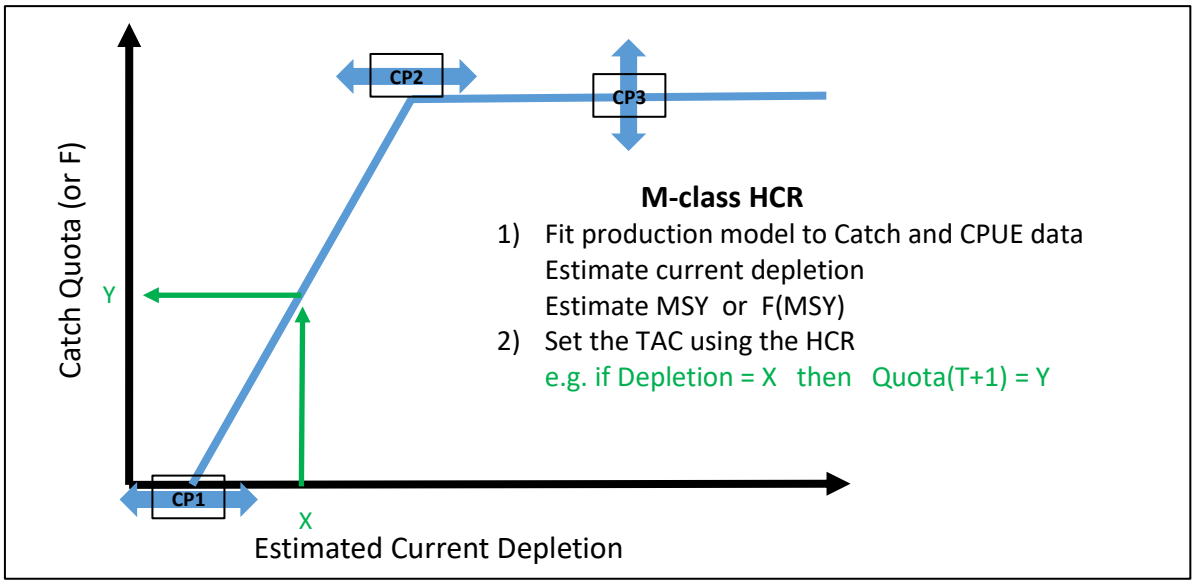

Figure 2. The PT-HS class of MPs involve two steps: 1) fit a surplus production model (Pella-Tomlinson Random Effects version), and 2) applying a hockeystick-shaped Harvest Control Rule (HCR) to the model estimates. The individual PT-HS MPs differ in terms of the Control Parameters (CP1-CP3) that define the shape of the HCR (and potentially the TAC change constraints). In the examples presented here, CP1 and CP2 were constant, while numerical optimization was used to find the value of CP3 that achieves the precise tuning objective.

"PT-Proj" - Model-based MPs with internal projections that aim for a specific future depletion target

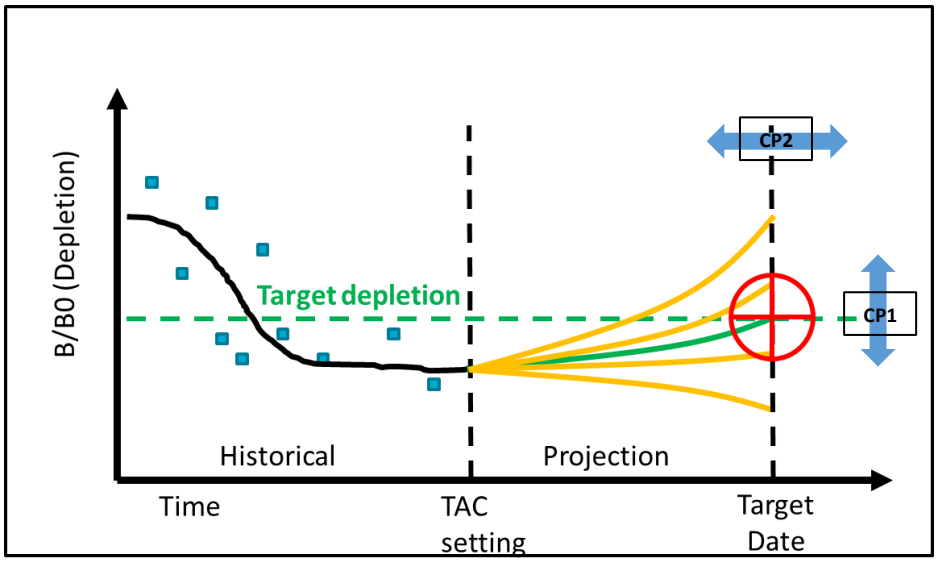

Figure 3. The PT-Proj class of MPs involve two steps: 1) fit a surplus production model (Pella-Tomlinson Random Effects version), and 2) use constant catch projections to estimate the TAC required to reach a target depletion level in a pre-defined number of years. The individual PT-HS class MPs differ in terms of the Control Parameters (CP1-CP2) that define the biomass target and how many years should be taken to reach the target (this latter may be constant, or an asymmetrical function of estimated depletion, e.g. aim to reach the target more quickly if the stock is more depleted). Numerical optimization was used to find the specific value of the biomass target (CP1) that achieves the precise tuning objective.

Table 1. Performance of candidate MPs with respect to key performance measures averaged over the 15 year period starting with the first MP-based TAC setting. Darker shading is better.

	15 year Performance Measure (2023-2037)				
Management Procedure	[1]SB/SB_{MSY}	[2]Prob(Green)	[3]Prob(*SB*>limit)	[4]Mean Catch	[5]Catch Variability
CPUE.B2	1.15 (0.77-1.48)	0.59	0.96	99.4 (73.0-119.8)	4.41
PT-HS.B2	1.20 (0.82-1.53)	0.67	0.98	94.7 (69.2-116.5)	4.32
PT-Proj.B2	1.13 (0.78-1.45)	0.60	0.97	96.4 (78.0-113.3)	4.07
CPUE.B3	1.19 (0.83-1.49)	0.67	0.98	94.1 (76.3-110.2)	3.97
PT-HS.B3	1.11 (0.77-1.42)	0.60	0.97	96.5 (77.0-112.5)	3.61
PT-Proj.B3	1.17 (0.83-1.47)	0.67	0.98	93.7 (74.6-109.7)	3.62

[1]***SB / SB_{MSY}*** = Median (10th-90th percentile) of the distribution of the mean ratio of Spawning Biomass (*SB*) relative to Spawning Biomass that would sustain *MSY* (SB_{MSY}).

[2]**Prob(Green)** = Proportion of results with stock status in the Kobe plot green quadrant across all years and simulations. (This is the mean of the distribution in Figure 4)

[3]**Prob(*SB*>limit)** = Proportion of results in which $SB > 0.5\ SB_{MSY}$ across all years and simulations. (This is the mean of the distribution in Figure 4)

[4]**Mean Catch** = Median (10th-90th percentile) of the distribution of the mean annual catch (1000 tonnes) across all years and simulations. (This is the mean of the distribution in Figure 4)

[5]**Catch Variability** (%AAV) = Mean of the absolute value of the change in catch between consecutive years, expressed as a percentage.

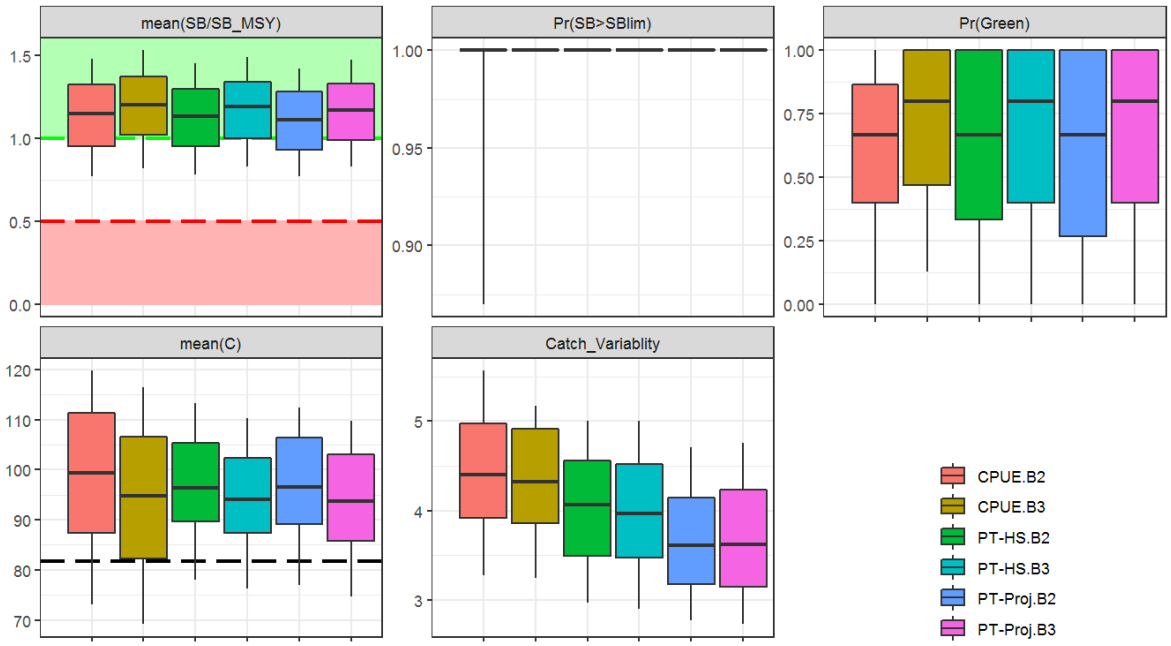

Figure 4. Boxplots comparing candidate MPs evaluated with the reference set Operating Model, with respect to key performance measures averaged over the 15 year period 2023 - 2037. Horizontal line is the median, boxes represent 25th - 75th percentiles, whiskers (thin vertical lines) represent the 10th-90th percentile range. Red and green horizontal reference lines represent the interim limit and target reference points for the mean SB/SB_{MSY} performance measure. The horizontal dashed black line is the mean reported catch 2017-2019.

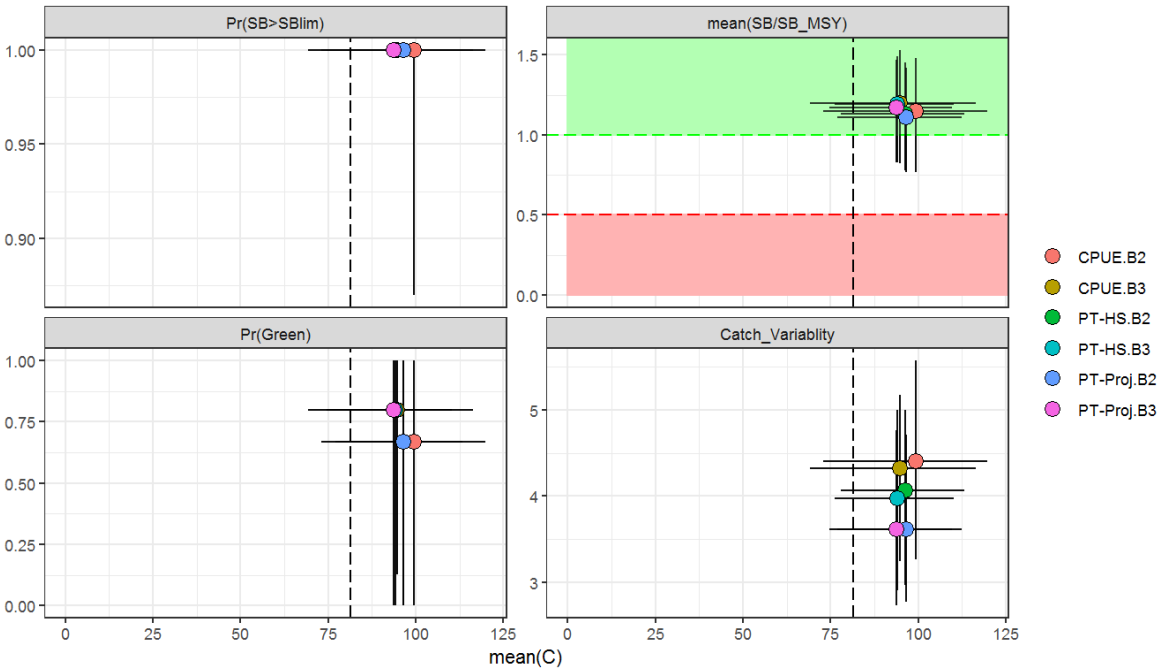

Figure 5. Trade-off plots comparing candidate MPs evaluated with the reference set Operating Model, with respect to catch on the X-axis, and 4 other key performance measures on the Y-axis, each averaged over the 15 year period 2023 - 2037. Circle is the median, lines represent 10th-90th percentiles. Red and green horizontal lines represent the interim limit and target reference points for the mean SB/SB_{MSY} performance measure. The dashed vertical black line is mean catch 2017-2019.

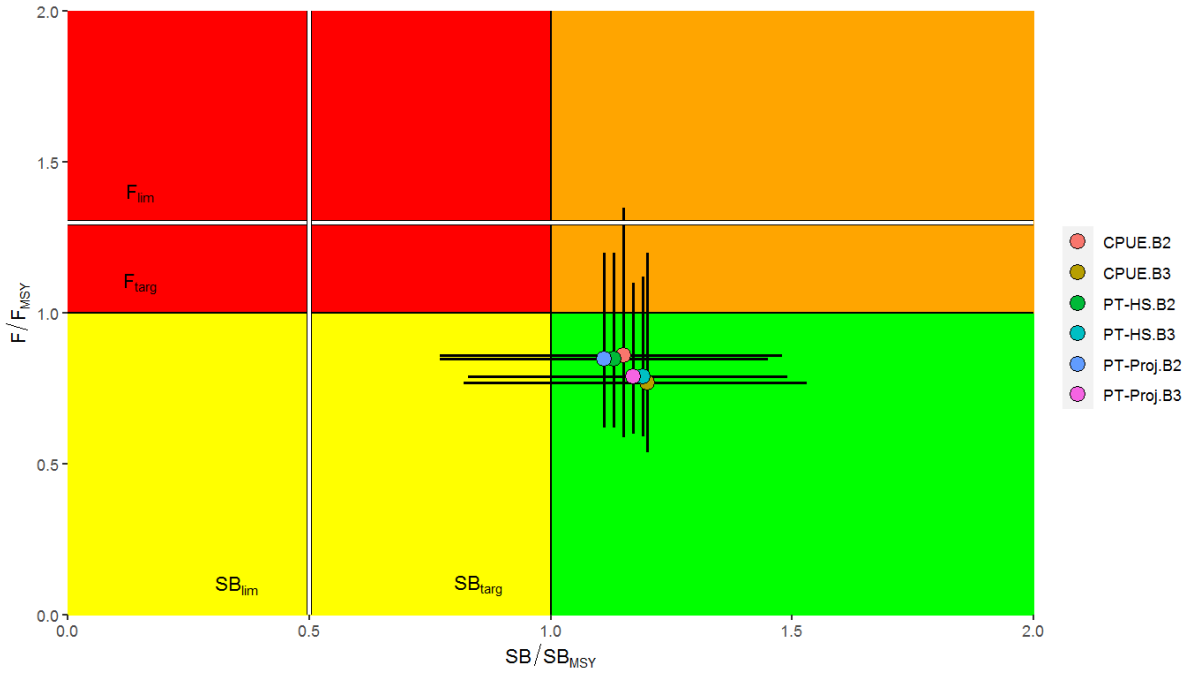

Figure 6. Kobe plot comparing candidate MPs evaluated with the reference set Operating Model, on the basis of the expected 15 year average (2023-2037) performance. Circle is the median, lines represent 10th-90th percentiles.

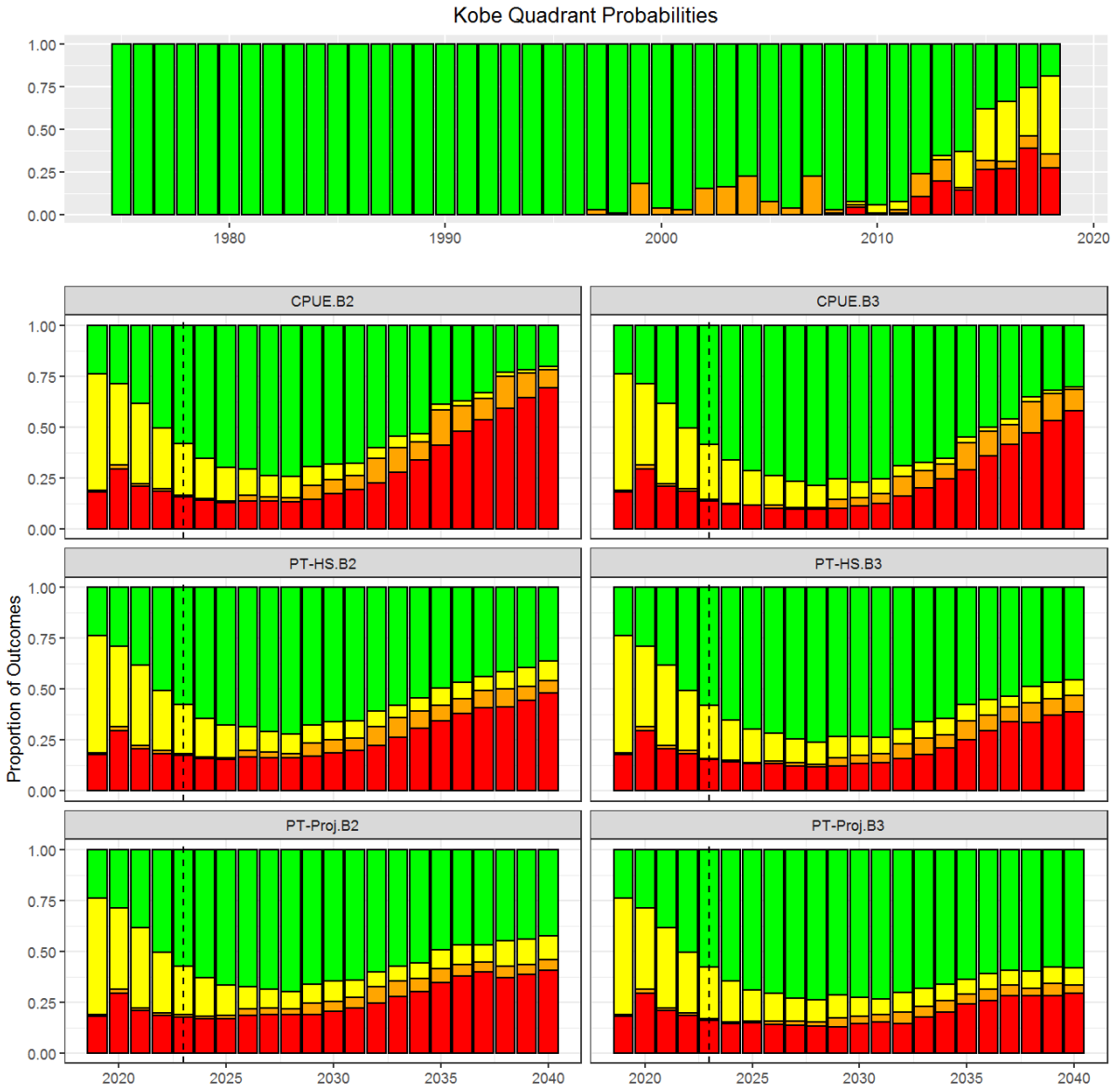

Figure 7. Proportion of simulations in each of the Kobe quadrants over time for each of the candidate MPs evaluated with the reference set Operating Model. Historical estimates are included in the top panel. The lower panels are projections, with the first MP application indicated by the broken vertical line (2023).

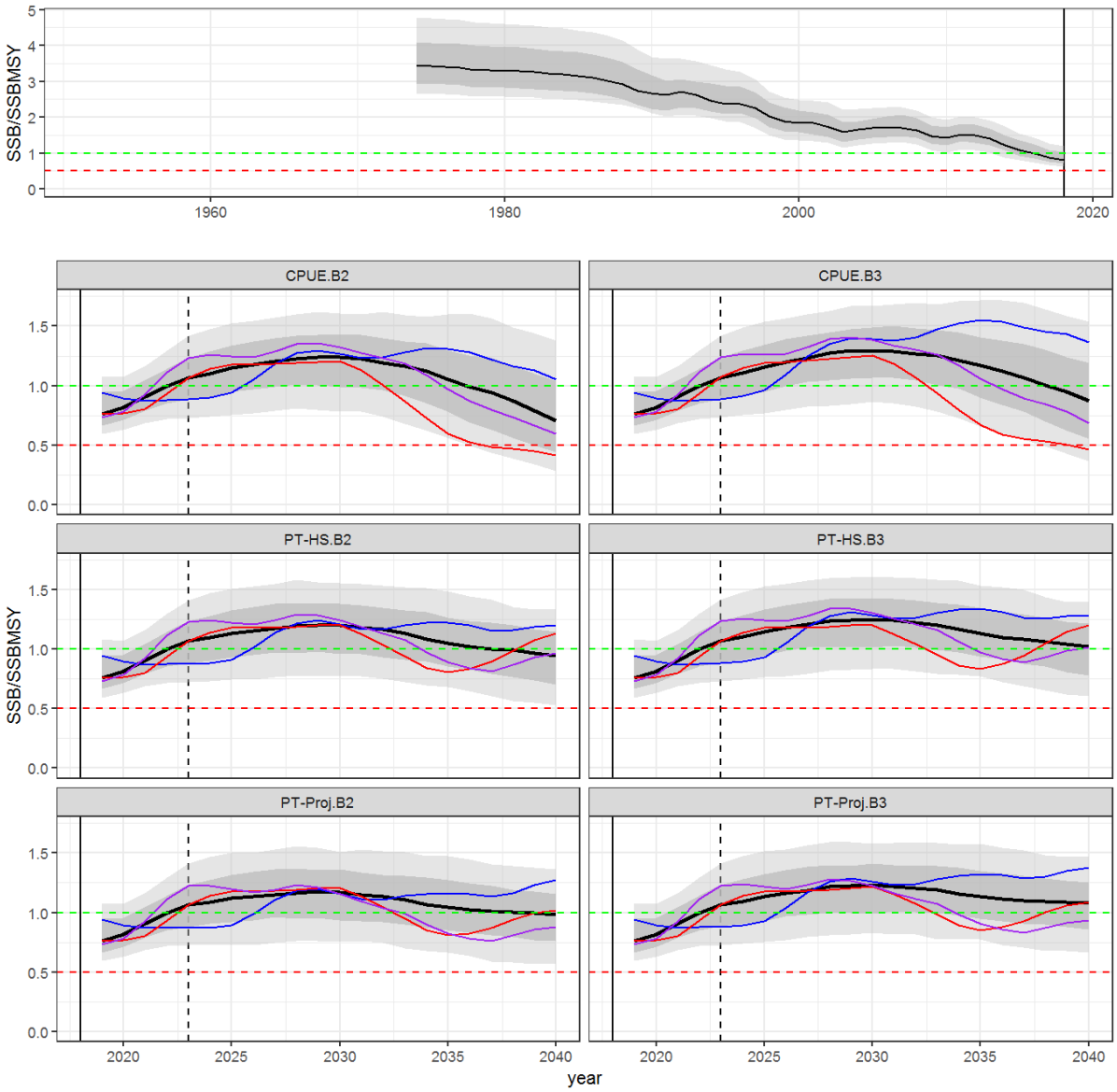

Figure 8. Time series of spawning stock size for the candidate MPs evaluated with the reference set Operating Model. The top panel represents the historical estimates from the reference case operating model, and lower plots represent the projection period. The solid vertical line represents the last year used in the historical conditioning. The broken vertical line represents the first year that the MP is applied. The median is represented by the bold black line, the dark shaded ribbon represents the 25th-75th percentiles, the light shaded ribbon represents the 10th-90th percentiles. Thick broken lines represent the interim target (green) and limit (red) reference points. The 3 thin coloured lines represent examples of individual realizations (the same OM scenarios across MPs and performance measures), to illustrate that individual variability greatly exceeds the median.

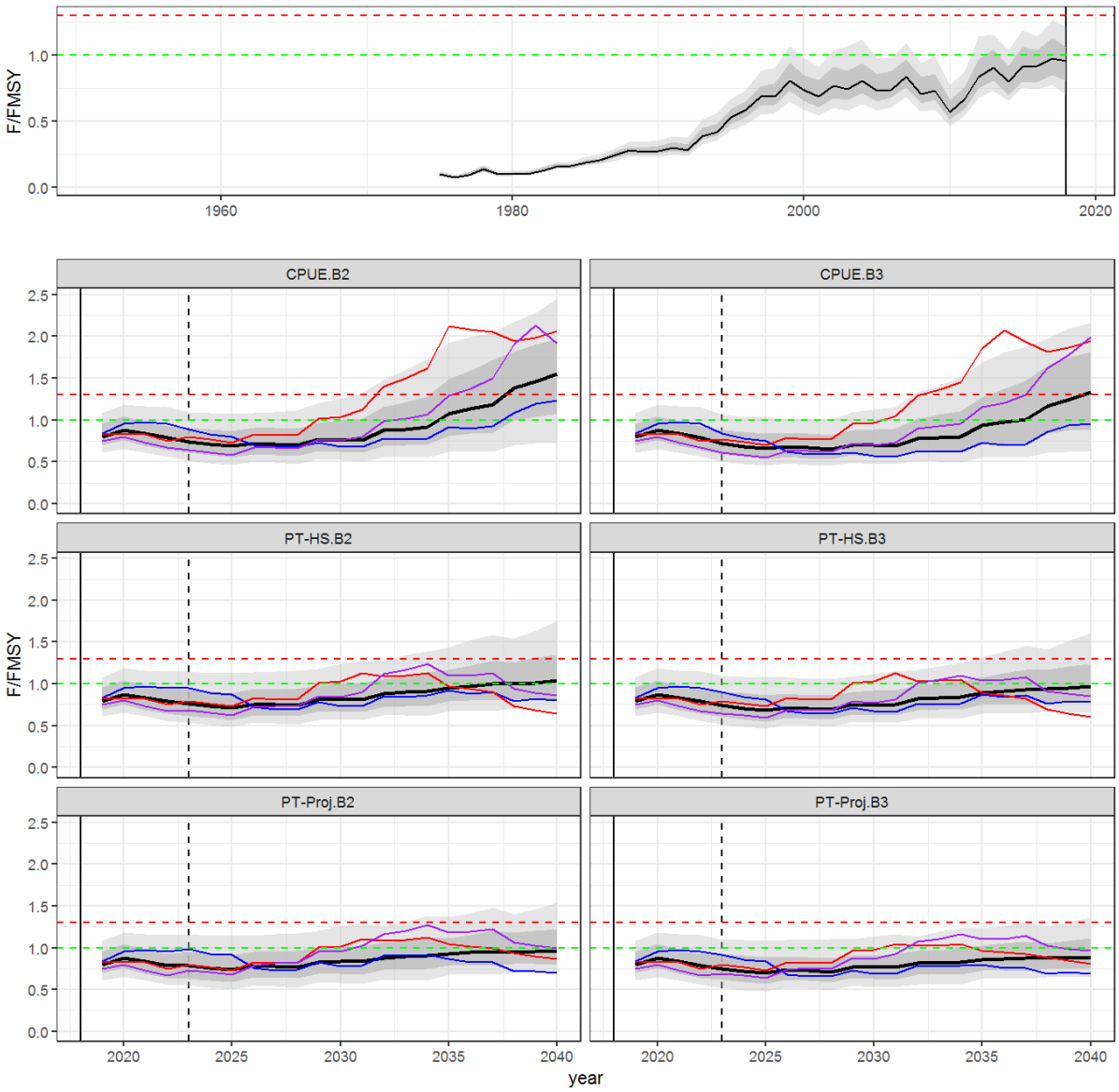

Figure 9. Time series of fishing intensity (relative to F_{MSY}) for the candidate MPs evaluated with the reference set Operating Model. The top panel represents the historical estimates from the reference case operating model, and lower plots represent the projection period. The solid vertical line represents the last year used in the historical conditioning. The broken vertical line represents the first year that the MP is applied. The median is represented by the bold black line, the dark shaded ribbon represents the 25th-75th percentiles, the light shaded ribbon represents the 10th-90th percentiles. Thick broken lines represent the interim target (green) and limit (red) reference points. The 3 thin coloured lines represent examples of individual realizations (the same OM scenarios across MPs and performance measures), to illustrate that individual variability greatly exceeds the median.

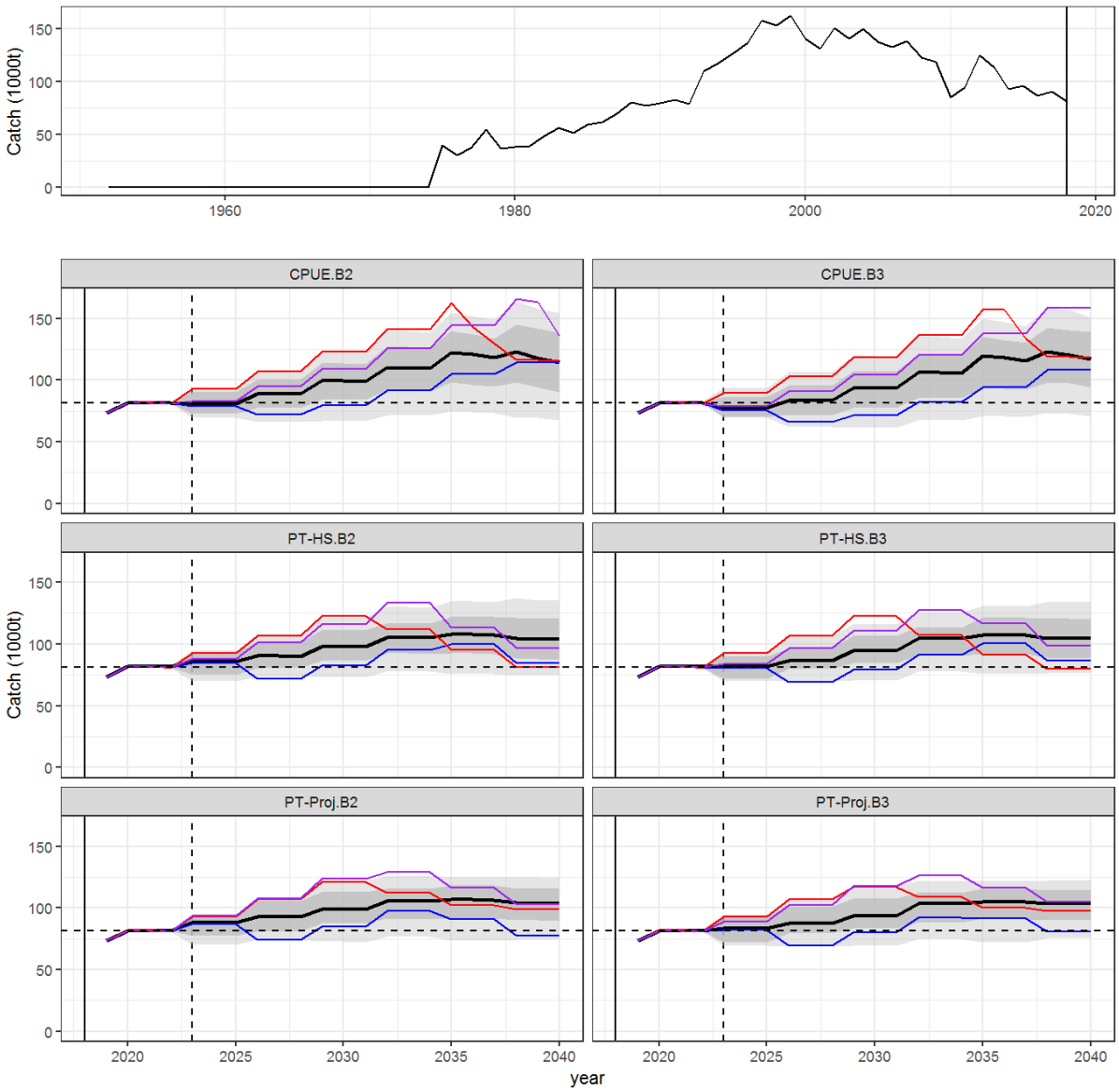

Figure 10. Time series of catch for the candidate MPs evaluated with the reference set Operating Model. The top panel represents the historical estimates from the reference case operating model, and lower plots represent the projection period. The solid vertical line represents the last year used in the historical conditioning. The broken vertical line represents the first year that the MP is applied. The median is represented by the bold black line, the dark shaded ribbon represents the 25th-75th percentiles, the light shaded ribbon represents the 10th-90th percentiles. The broken black horizontal line represents recent (mean 2017-2019) catch. The 3 thin coloured lines represent examples of individual realizations (the same OM scenarios across MPs and performance measures), to illustrate that individual variability greatly exceeds the median.

Appendix 1. Candidate Management Procedure summary performance tables for a range of time periods from 1-20 years (aggregated over regions and fisheries).

Table A1.1. Candidate MP performance for standard IOTC performance measures aggregated over a 10 year period.

	10 year Performance Measure (2023-2032)				
Management Procedure	[1]SB/SB_{MSY}	[2]Prob(Green)	[3]Prob(*SB*>limit)	[4]Mean Catch	[5]Catch Variability
CPUE.B2	1.19 (0.81-1.52)	0.68	0.98	91.5 (69.8-111.0)	5.00
PT-HS.B2	1.23 (0.85-1.55)	0.72	0.99	86.6 (66.6-108.1)	4.95
PT-Proj.B2	1.16 (0.79-1.50)	0.66	0.98	92.4 (76.8-108.1)	4.84
CPUE.B3	1.20 (0.81-1.52)	0.70	0.99	88.9 (74.3-105.2)	4.72
PT-HS.B3	1.13 (0.77-1.48)	0.64	0.98	93.5 (76.3-108.8)	4.59
PT-Proj.B3	1.18 (0.82-1.50)	0.69	0.99	89.5 (72.8-106.9)	4.64

Table A1.2. Candidate MP performance for standard IOTC performance measures aggregated over a 20 year period.

	20 year Performance Measure (2023-2042)				
Management Procedure	[1]SB/SB_{MSY}	[2]Prob(Green)	[3]Prob(*SB*>limit)	[4]Mean Catch	[5]Catch Variability
CPUE.B2	1.05 (0.69-1.41)	0.49	0.90	105.0 (77.1-124.2)	4.95
PT-HS.B2	1.12 (0.78-1.50)	0.57	0.93	103.3 (73.1-120.9)	4.83
PT-Proj.B2	1.08 (0.76-1.37)	0.54	0.96	99.3 (80.4-113.4)	4.40
CPUE.B3	1.14 (0.81-1.42)	0.61	0.97	97.5 (79.9-111.2)	4.22
PT-HS.B3	1.08 (0.77-1.37)	0.55	0.96	99.1 (79.0-112.9)	3.64
PT-Proj.B3	1.15 (0.83-1.45)	0.64	0.98	97.0 (77.0-110.3)	3.62

[1]***SB / SB_{MSY}*** = Median (10th-90th percentile) of the distribution of the mean ratio of Spawning Biomass (*SB*) relative to Spawning Biomass that would sustain *MSY* (SB_{MSY}).

[2]**Prob(Green)** = Proportion of results with stock status in the Kobe plot green quadrant across all years and simulations.

[3]**Prob(*SB*>limit)** = Proportion of results in which $SB > 0.5\ SB_{MSY}$ across all years and simulations.

[4]**Mean Catch** = Median (10th-90th percentile) of the distribution of the mean annual catch (1000 tonnes) across all years and simulations.

[5]**Catch Variability** (%AAV) = Mean of the absolute value of the change in catch between consecutive years, expressed as a percentage.

Table A1.3. Candidate MP performance for standard IOTC performance measures for the year 2023.

Status : maximise stock status		**1 year average**					
		CPUE.B2	**PT-HS.B2**	**PT-Proj.B2**	**CPUE.B3**	**PT-HS.B3**	**PT-Proj.B3**
Mean spawner biomass relative to pristine	SB/SB_0	0.46	0.46	0.46	0.46	0.46	0.46
Minimum spawner biomass relative to pristine	SB/SB_0	0.46	0.46	0.46	0.46	0.46	0.46
Mean spawner biomass relative to SBMSY	SB/SB_{MSY}	1.77	1.77	1.78	1.78	1.78	1.78
Mean fishing mortality relative to FMSY	F/F_{tar}	0.54	0.54	0.52	0.53	0.50	0.51
Mean fishing mortality relative to target	F/F_{MSY}	0.54	0.54	0.52	0.53	0.50	0.51
Probability of being in Kobe green quadrant	SB,F	0.83	0.83	0.83	0.83	0.84	0.83
Probability of being in Kobe red quadrant	SB,F	0.12	0.11	0.11	0.10	0.09	0.08
Safety : maximise the probability of remaining above low stock status (i.e. minimise risk)							
Probability of spawner biomass being above 20% of SB0	SB	0.99	0.99	0.99	0.99	0.99	0.99
Probability of spawner biomass being above BLim	SB	1.00	1.00	1.00	1.00	1.00	1.00
Yield : maximise catches across regions and gears							
Mean catch (1000 t)	C	103.21	102.91	102.36	102.20	101.07	100.79
Mean relative CPUE (aggregate)	C	0.83	0.84	0.79	0.80	0.77	0.78
Mean catch relative to MSY	C/MSY	0.85	0.85	0.85	0.85	0.85	0.85
Stability: maximise stability in catches to reduce commercial uncertainty							
Mean absolute proportional change in catch	C	15.00	15.00	15.00	15.00	15.00	15.00
% Catch coefficient of variation	C	NA	NA	NA	NA	NA	NA
Probability of shutdown	C	0.00	0.00	0.00	0.00	0.00	0.00

Table A1.4. Candidate MP performance for standard IOTC performance measures for the 5 year period 2023-2027.

Status : maximise stock status		**5 year average**					
		CPUE.B2	**PT-HS.B2**	**PT-Proj.B2**	**CPUE.B3**	**PT-HS.B3**	**PT-Proj.B3**
Mean spawner biomass relative to pristine	SB/SB_0	0.31	0.31	0.30	0.31	0.30	0.30
Minimum spawner biomass relative to pristine	SB/SB_0	0.28	0.28	0.28	0.28	0.28	0.28
Mean spawner biomass relative to SBMSY	SB/SB_{MSY}	1.14	1.16	1.13	1.15	1.12	1.13
Mean fishing mortality relative to FMSY	F/F_{tar}	0.72	0.68	0.76	0.71	0.78	0.73
Mean fishing mortality relative to target	F/F_{MSY}	0.72	0.68	0.76	0.71	0.78	0.73
Probability of being in Kobe green quadrant	SB,F	0.67	0.69	0.66	0.68	0.64	0.67
Probability of being in Kobe red quadrant	SB,F	0.14	0.12	0.17	0.14	0.18	0.15
Safety : maximise the probability of remaining above low stock status (i.e. minimise risk)							
Probability of spawner biomass being above 20% of SB0	SB	0.88	0.89	0.87	0.88	0.87	0.88
Probability of spawner biomass being above BLim	SB	0.99	0.99	0.99	0.99	0.99	0.99
Yield : maximise catches across regions and gears							
Mean catch (1000 t)	C	83.74	79.37	87.28	83.39	89.64	85.53
Mean relative CPUE (aggregate)	C	0.71	0.68	0.74	0.71	0.76	0.73
Mean catch relative to MSY	C/MSY	1.03	1.05	1.02	1.03	1.01	1.02
Stability: maximise stability in catches to reduce commercial uncertainty							
Mean absolute proportional change in catch	C	4.40	4.46	4.63	4.43	4.91	4.92
% Catch coefficient of variation	C	0.08	0.08	0.08	0.08	0.08	0.08
Probability of shutdown	C	0.00	0.00	0.00	0.00	0.00	0.00

Table A1.5. Candidate MP performance for standard IOTC performance measures for the 10 year period 2023-2032.

Status : maximise stock status		**10 year average**					
		CPUE.B2	**PT-HS.B2**	**PT-Proj.B2**	**CPUE.B3**	**PT-HS.B3**	**PT-Proj.B3**
Mean spawner biomass relative to pristine	SB/SB_0	0.31	0.33	0.31	0.32	0.31	0.32
Minimum spawner biomass relative to pristine	SB/SB_0	0.26	0.27	0.27	0.27	0.26	0.27
Mean spawner biomass relative to SBMSY	SB/SB_{MSY}	1.19	1.23	1.16	1.20	1.13	1.18
Mean fishing mortality relative to FMSY	F/F_{tar}	0.75	0.70	0.78	0.74	0.80	0.75
Mean fishing mortality relative to target	F/F_{MSY}	0.75	0.70	0.78	0.74	0.80	0.75
Probability of being in Kobe green quadrant	SB,F	0.68	0.72	0.66	0.70	0.64	0.69
Probability of being in Kobe red quadrant	SB,F	0.16	0.12	0.18	0.14	0.20	0.15
Safety : maximise the probability of remaining above low stock status (i.e. minimise risk)							
Probability of spawner biomass being above 20% of SB0	SB	0.89	0.91	0.88	0.90	0.87	0.89
Probability of spawner biomass being above BLim	SB	0.98	0.99	0.98	0.99	0.98	0.99
Yield : maximise catches across regions and gears							
Mean catch (1000 t)	C	91.53	86.55	92.36	88.94	93.51	89.47
Mean relative CPUE (aggregate)	C	0.77	0.73	0.78	0.76	0.79	0.76
Mean catch relative to MSY	C/MSY	1.08	1.12	1.05	1.08	1.04	1.07
Stability: maximise stability in catches to reduce commercial uncertainty							
Mean absolute proportional change in catch	C	5.00	4.95	4.84	4.72	4.59	4.64
% Catch coefficient of variation	C	0.14	0.13	0.11	0.11	0.11	0.10
Probability of shutdown	C	0.00	0.00	0.00	0.00	0.00	0.00

Table A1.6. Candidate MP performance for standard IOTC performance measures for the 15 year period 2023-2037.

Status : maximise stock status		**15 year average**					
		CPUE.B2	**PT-HS.B2**	**PT-Proj.B2**	**CPUE.B3**	**PT-HS.B3**	**PT-Proj.B3**
Mean spawner biomass relative to pristine	SB/SB_0	0.30	0.32	0.30	0.32	0.30	0.31
Minimum spawner biomass relative to pristine	SB/SB_0	0.21	0.23	0.24	0.25	0.24	0.25
Mean spawner biomass relative to SBMSY	SB/SB_{MSY}	1.15	1.20	1.13	1.19	1.11	1.17
Mean fishing mortality relative to FMSY	F/F_{tar}	0.86	0.77	0.85	0.79	0.85	0.79
Mean fishing mortality relative to target	F/F_{MSY}	0.86	0.77	0.85	0.79	0.85	0.79
Probability of being in Kobe green quadrant	SB,F	0.59	0.67	0.60	0.67	0.60	0.67
Probability of being in Kobe red quadrant	SB,F	0.24	0.18	0.23	0.18	0.24	0.17
Safety : maximise the probability of remaining above low stock status (i.e. minimise risk)							
Probability of spawner biomass being above 20% of SB0	SB	0.85	0.89	0.87	0.89	0.86	0.89
Probability of spawner biomass being above BLim	SB	0.96	0.98	0.97	0.98	0.97	0.98
Yield : maximise catches across regions and gears							
Mean catch (1000 t)	C	99.43	94.69	96.42	94.12	96.53	93.72
Mean relative CPUE (aggregate)	C	0.84	0.80	0.83	0.80	0.82	0.80
Mean catch relative to MSY	C/MSY	1.05	1.10	1.04	1.08	1.04	1.08
Stability: maximise stability in catches to reduce commercial uncertainty							
Mean absolute proportional change in catch	C	4.41	4.32	4.07	3.97	3.61	3.62
% Catch coefficient of variation	C	0.18	0.18	0.14	0.14	0.11	0.11
Probability of shutdown	C	0.00	0.00	0.00	0.00	0.00	0.00

Table A1.7. Candidate MP performance for standard IOTC performance measures for the 20 year period 2023-2042.

Status : maximise stock status		**20 year average**					
		CPUE.B2	**PT-HS.B2**	**PT-Proj.B2**	**CPUE.B3**	**PT-HS.B3**	**PT-Proj.B3**
Mean spawner biomass relative to pristine	SB/SB_0	0.28	0.31	0.29	0.31	0.29	0.31
Minimum spawner biomass relative to pristine	SB/SB_0	0.15	0.18	0.21	0.22	0.22	0.23
Mean spawner biomass relative to SBMSY	SB/SB_{MSY}	1.05	1.12	1.08	1.14	1.08	1.15
Mean fishing mortality relative to FMSY	F/F_{tar}	1.05	0.94	0.89	0.84	0.89	0.83
Mean fishing mortality relative to target	F/F_{MSY}	1.05	0.94	0.89	0.84	0.89	0.83
Probability of being in Kobe green quadrant	SB,F	0.49	0.57	0.54	0.61	0.55	0.64
Probability of being in Kobe red quadrant	SB,F	0.36	0.28	0.29	0.23	0.28	0.21
Safety : maximise the probability of remaining above low stock status (i.e. minimise risk)							
Probability of spawner biomass being above 20% of SB0	SB	0.76	0.82	0.83	0.87	0.84	0.89
Probability of spawner biomass being above BLim	SB	0.90	0.93	0.96	0.97	0.96	0.98
Yield : maximise catches across regions and gears							
Mean catch (1000 t)	C	105.01	103.29	99.26	97.52	99.08	96.95
Mean relative CPUE (aggregate)	C	0.90	0.88	0.84	0.83	0.84	0.81
Mean catch relative to MSY	C/MSY	0.98	1.04	1.02	1.06	1.02	1.08
Stability: maximise stability in catches to reduce commercial uncertainty							
Mean absolute proportional change in catch	C	4.95	4.83	4.40	4.22	3.64	3.62
% Catch coefficient of variation	C	0.21	0.21	0.15	0.15	0.12	0.12
Probability of shutdown	C	0.00	0.00	0.00	0.00	0.00	0.00

Appendix 2. Candidate Management Procedure summary performance graphics for a range of robustness tests.

The process of developing, tuning and selecting an MP emphasizes performance as tested with a reference set Operating Model (OM), which include the most important and "relatively high" probability uncertainties in the model structure, parameters, and data. In contrast, robustness test OMs are typically considered to be less likely, and often focus on particularly challenging circumstances with potentially negative consequences to be avoided. Robustness tests are conducted using the MPs that were tuned to the reference set Operating Model. They can provide confidence that an MP can deal with a broader range of situations, and may be useful for discriminating among candidate MPs that perform similarly with respect to the reference set.

In this appendix, we report against 7 Robustness tests, 6 requested by the IOTC Working Parties on Methods and/or Tropical Tunas, and a final one related to falling recent catches:

1) What happens if there are 8 quarters of poor recruitment (55% of expected + usual stochastic error, as defined for yellowfin based on historical assessments)?
2) What happens if the (annualized, spatially-aggregated) longline CPUE observation error CV is increased to 30% (auto-correlation = 0.5) in projections?
3) What happens if there is a consistent 10% future over-catch (reported accurately), for all fleets (starting in 2023, the first year of MP-based TAC setting)?
4) What happens if there is a consistent 10% future over-catch (unreported), for all fleets (starting in 2023, the first year of MP-based TAC setting)?
5) What happens if there is a consistent 10% future over-catch (exactly half of which is reported), for all fleets (starting in 2023, the first year of MP-based TAC setting)?
6) What happens if the longline CPUE catchability trend is 2% per year in the projections (but the same as the reference set OM historically)?
7) What happens if there is a 5 year period (2020-2024) in which industry drops catch by 20% from recent levels (or under-catches the TAC by 20%) before the quotas become restrictive (i.e. economic or logistical restrictions to the fishery suddenly disappear)?

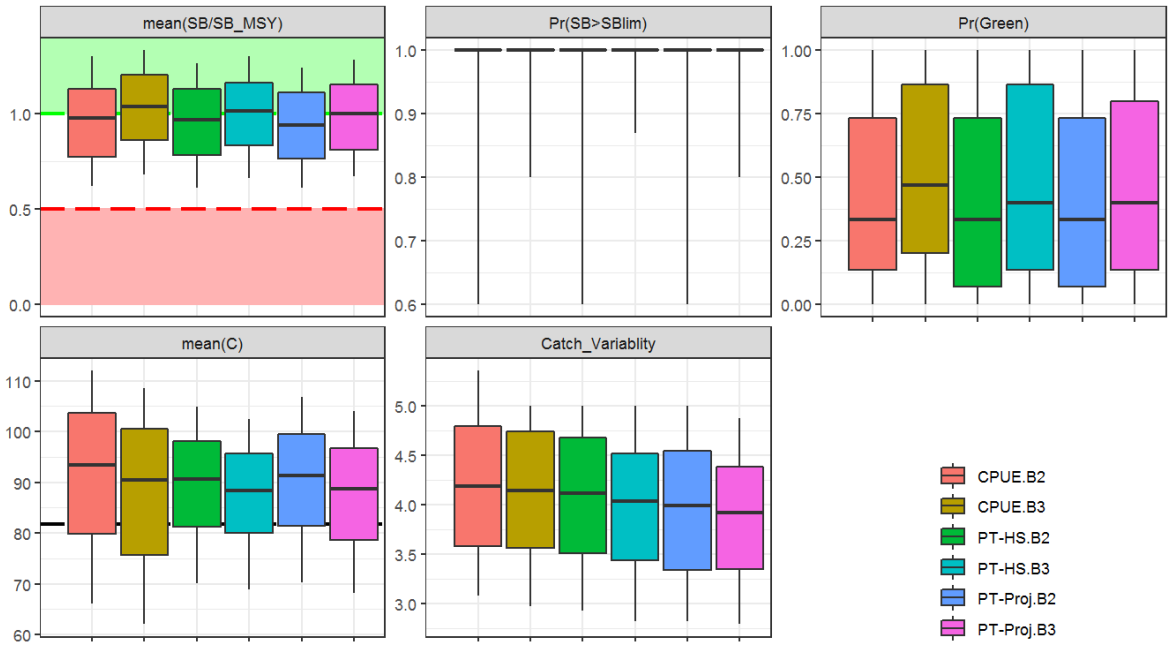

Figure A2.1. Boxplots comparing candidate MPs evaluated with the robustness test 1 (recruitment shock), illustrating key performance measures averaged over the period 2023 - 2037. Horizontal line is the median, boxes represent 25th - 75th percentiles, whiskers (thin vertical lines) represent the 10th-90th percentile range. Red and green horizontal reference lines represent the interim limit and target reference points for the mean SB/SB$_{MSY}$ performance measure. The horizontal dashed black line is the mean reported catch 2017-2019.

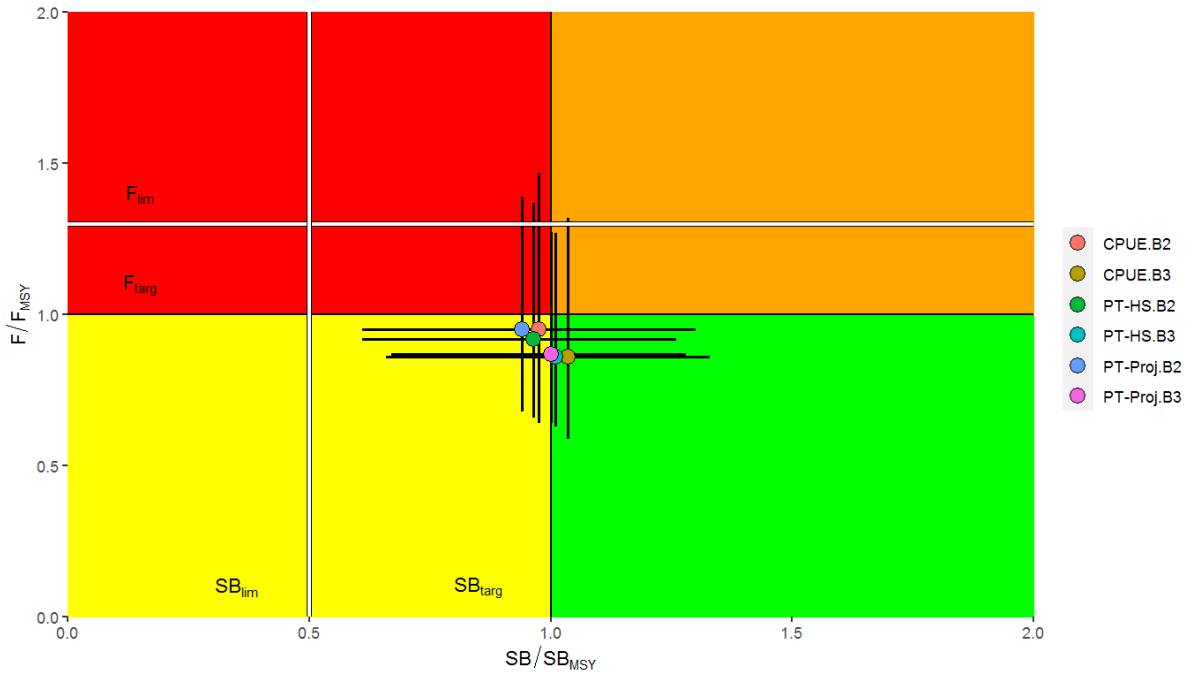

Figure A2.1. Kobe plot comparing candidate MPs evaluated with the robustness test 1 (recruitment shock), on the basis of the expected 15 year average (2023-2037) performance. Circle is the median, lines represent 10th-90th percentiles.

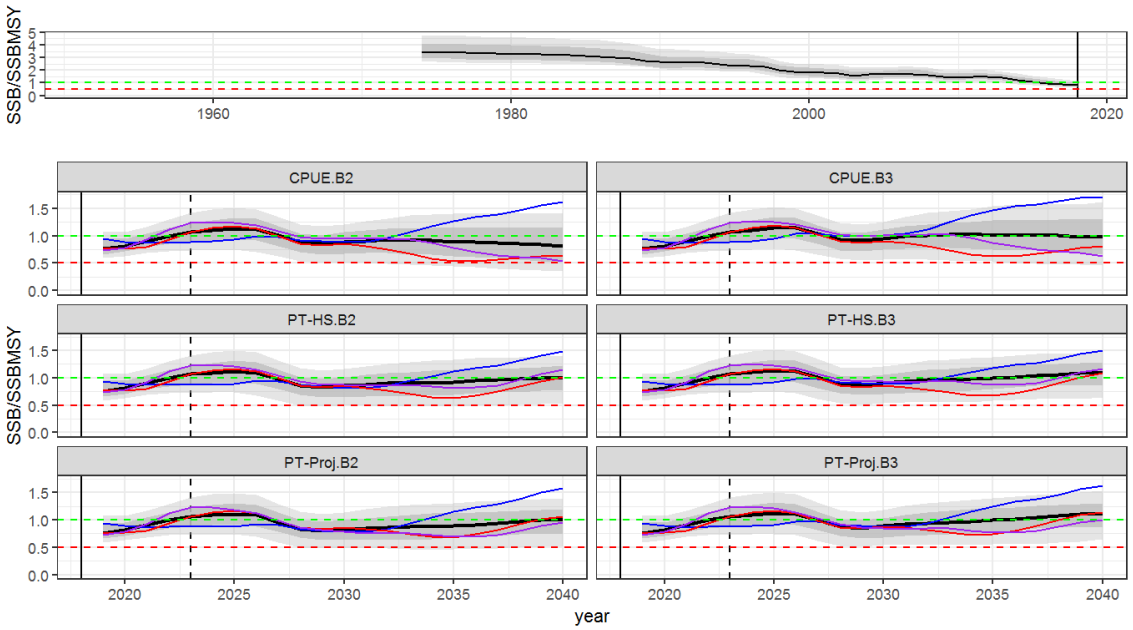

Figure A2.3. Time series of spawning stock size for the candidate MPs with the robustness test 1 (recruitment shock). The top panel represents the historical estimates from the reference case operating model, and lower plots represent the projection period. The solid vertical line represents the last year used in the historical conditioning. The broken vertical line represents the first year that the MP is applied. The median is represented by the bold black line, the dark shaded ribbon represents the 25th-75th percentiles, the light shaded ribbon represents the 10th-90th percentiles. Thick broken lines represent the interim target (green) and limit (red) reference points. The 3 thin coloured lines represent examples of individual realizations (the same OM scenarios across MPs and performance measures), to illustrate that individual variability greatly exceeds the median.

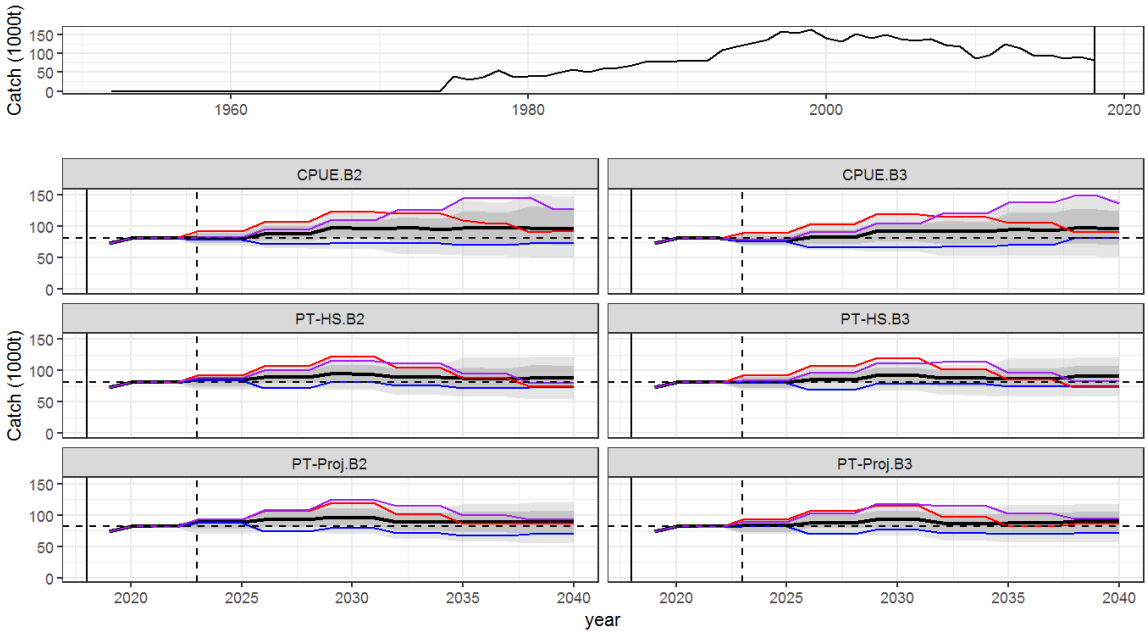

Figure A2.4. Time series of catch for the candidate MPs evaluated with the robustness test 1 (recruitment shock). The top panel represents the historical estimates from the reference case operating model, and lower plots represent the projection period. The solid vertical line represents the last year used in the historical conditioning. The broken vertical line represents the first year that the MP is applied. The median is represented by the bold black line, the dark shaded ribbon represents the 25th-75th percentiles, the light shaded ribbon represents the 10th-90th percentiles. The broken black horizontal line represents recent (2017) catch. The 3 thin coloured lines represent examples of individual realizations (the same OM scenarios across MPs and performance measures), to illustrate that individual variability greatly exceeds the median.

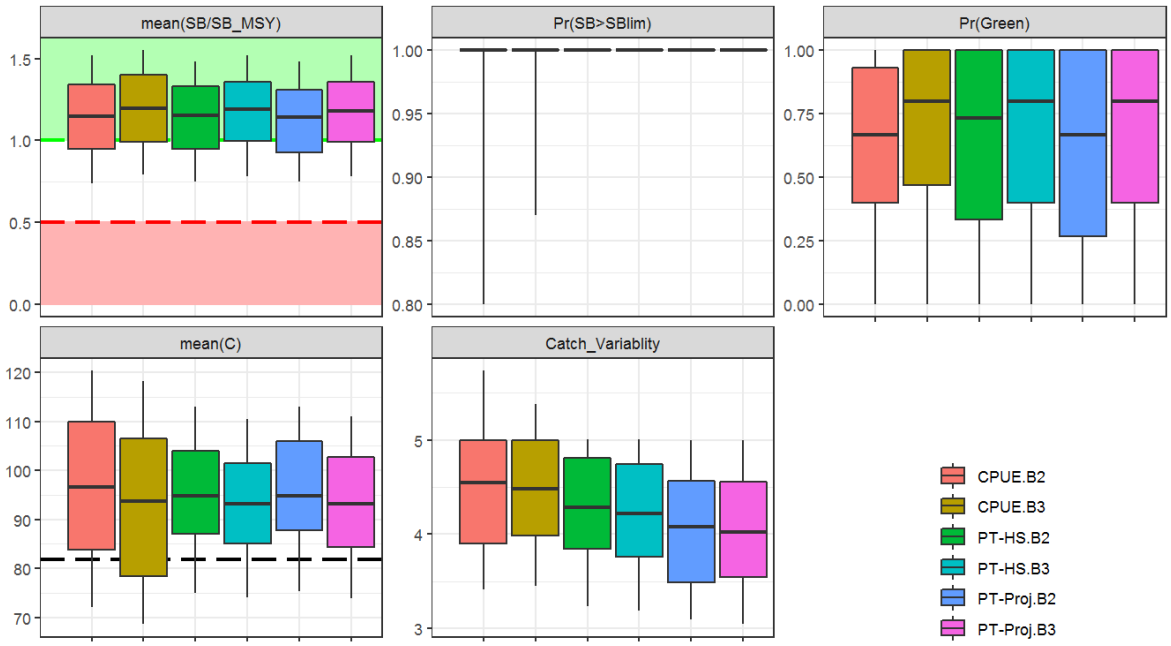

Figure A2.5. Boxplots comparing candidate MPs evaluated with the robustness test 2 (CPUE CV = 30%), illustrating key performance measures averaged over the period 2023 - 2037. Horizontal line is the median, boxes represent 25th - 75th percentiles, whiskers (thin vertical lines) represent the 10th-90th percentile range. Red and green horizontal reference lines represent the interim limit and target reference points for the mean SB/SB$_{MSY}$ performance measure. The horizontal dashed black line is the mean reported catch 2017-2019.

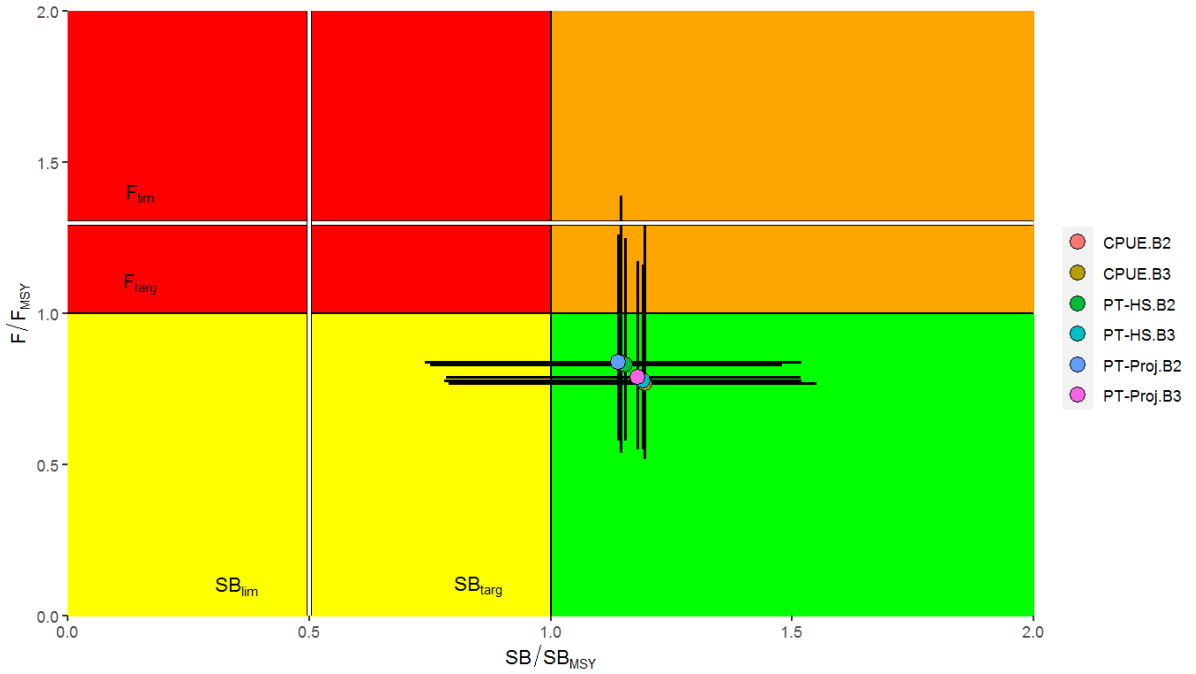

Figure A2.6. Kobe plot comparing candidate MPs evaluated with the robustness test 2 (CPUE CV = 30%), on the basis of the expected 15 year average (2023-2037) performance. Circle is the median, lines represent 10th-90th percentiles.

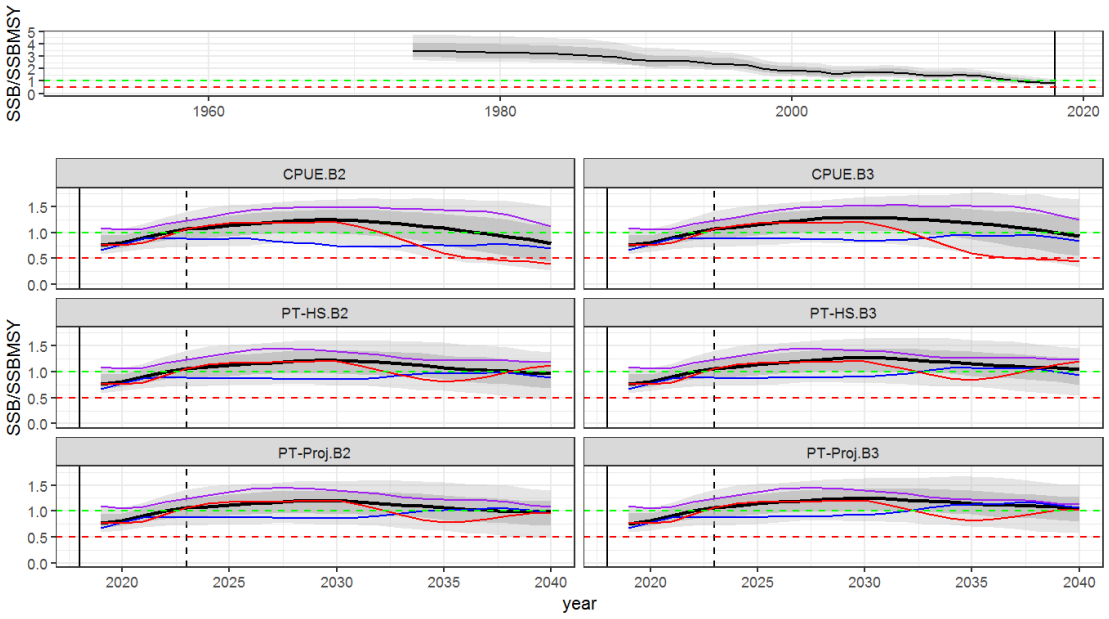

Figure A2.7. Time series of spawning stock size for the candidate MPs with the robustness test 2 (CPUE CV = 30%). The top panel represents the historical estimates from the reference case operating model, and lower plots represent the projection period. The solid vertical line represents the last year used in the historical conditioning. The broken vertical line represents the first year that the MP is applied. The median is represented by the bold black line, the dark shaded ribbon represents the 25th-75th percentiles, the light shaded ribbon represents the 10th-90th percentiles. Thick broken lines represent the interim target (green) and limit (red) reference points. The 3 thin coloured lines represent examples of individual realizations (the same OM scenarios across MPs and performance measures), to illustrate that individual variability greatly exceeds the median.

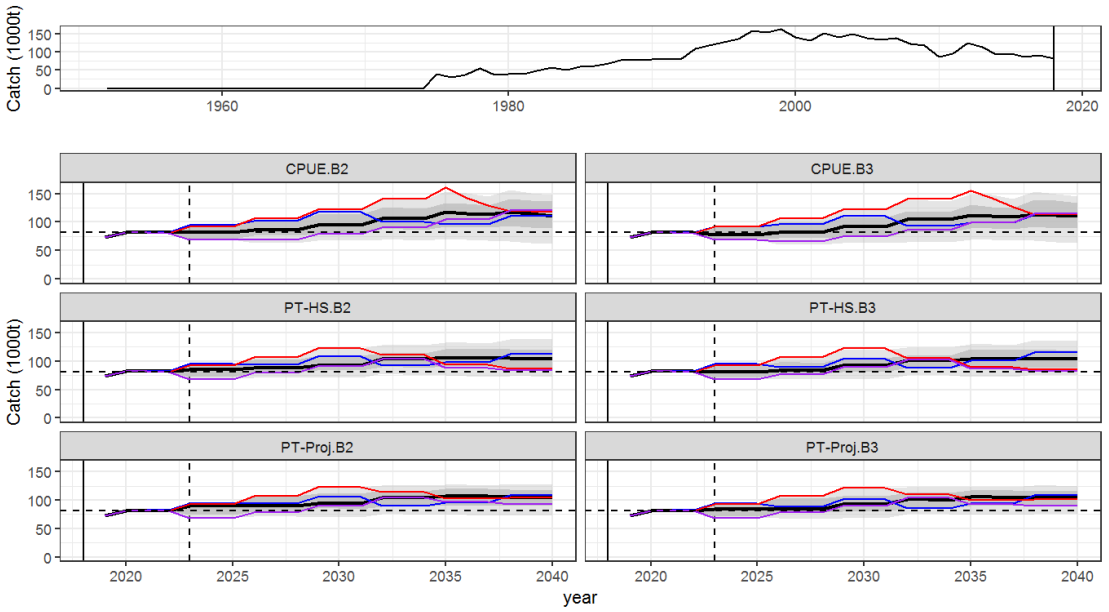

Figure A2.8. Time series of catch for the candidate MPs evaluated with the robustness test 2 (CPUE CV = 30%). The top panel represents the historical estimates from the reference case operating model, and lower plots represent the projection period. The solid vertical line represents the last year used in the historical conditioning. The broken vertical line represents the first year that the MP is applied. The median is represented by the bold black line, the dark shaded ribbon represents the 25th-75th percentiles, the light shaded ribbon represents the 10th-90th percentiles. The broken black horizontal line represents recent (2017) catch. The 3 thin coloured lines represent examples of individual realizations (the same OM scenarios across MPs and performance measures), to illustrate that individual variability greatly exceeds the median.

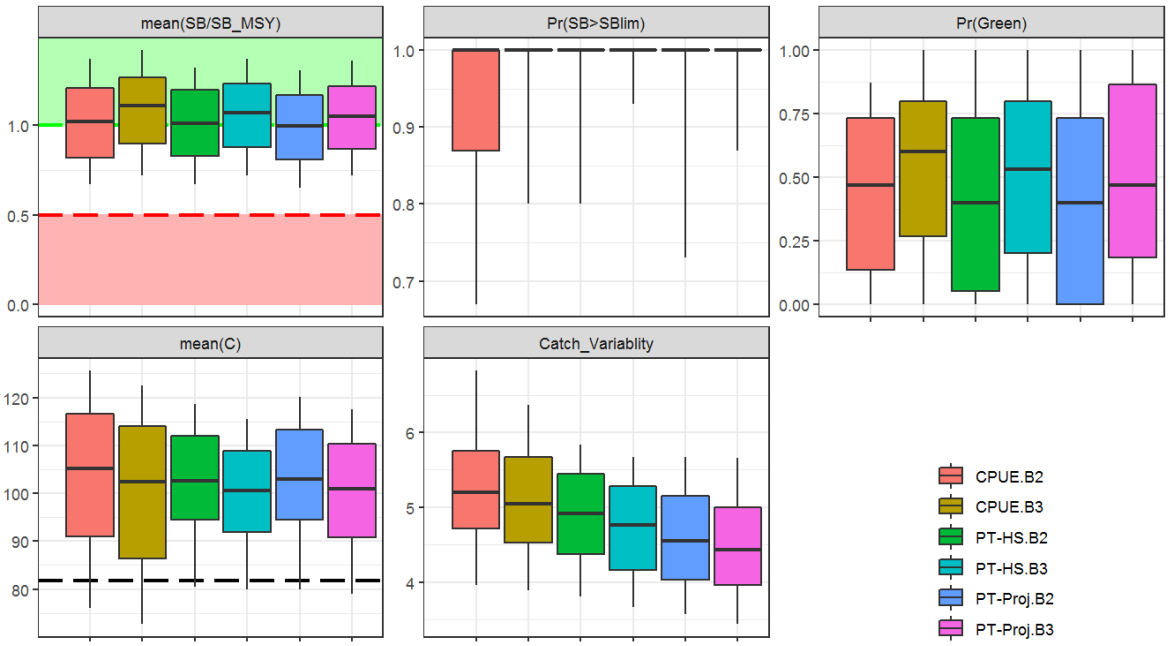

Figure A2.9. Boxplots comparing candidate MPs evaluated with the robustness test 3 (10% reported overcatch), illustrating key performance measures averaged over the period 2023 - 2037. Horizontal line is the median, boxes represent 25th - 75th percentiles, whiskers (thin vertical lines) represent the 10th-90th percentile range. Red and green horizontal reference lines represent the interim limit and target reference points for the mean SB/SB$_{MSY}$ performance measure. The horizontal dashed black line is the mean reported catch 2017-2019.

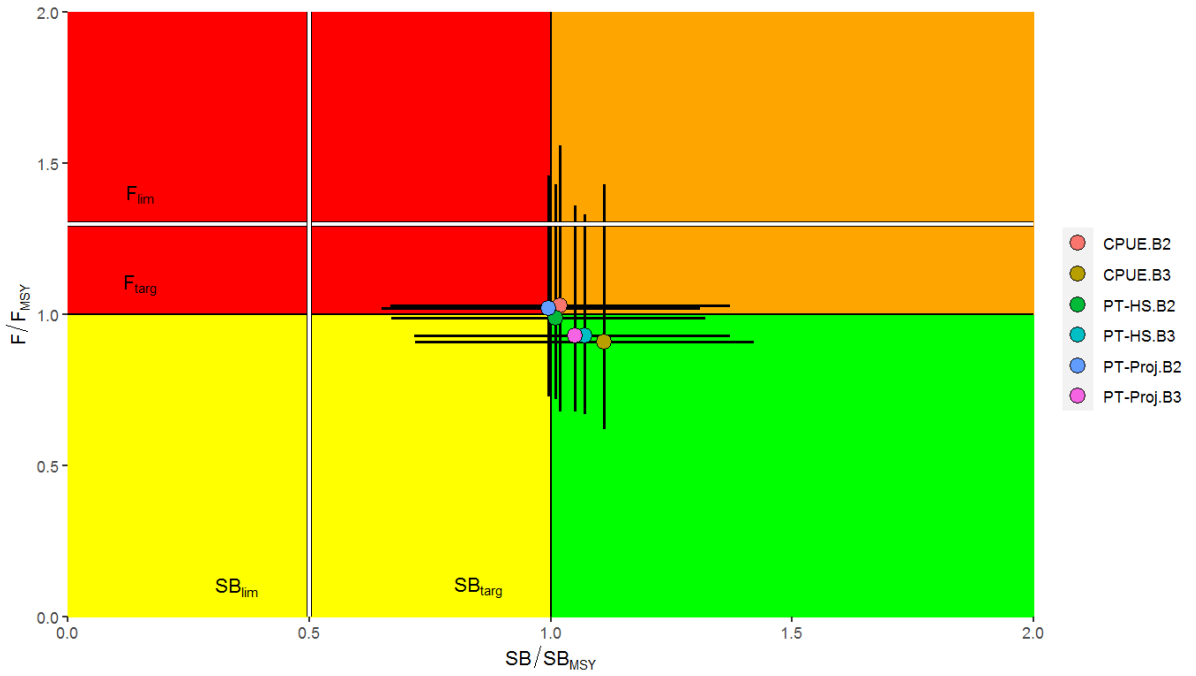

Figure A2.10. Kobe plot comparing candidate MPs evaluated with the robustness test 3 (10% reported overcatch), on the basis of the expected 15 year average (2023-2037) performance. Circle is the median, lines represent 10th-90th percentiles.

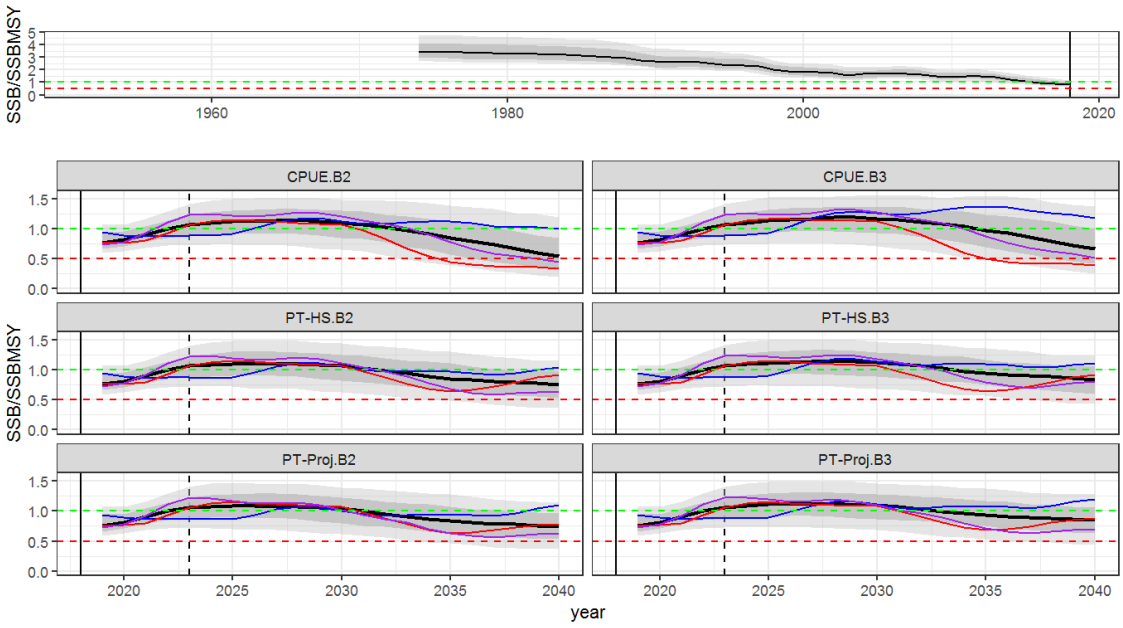

Figure A2.11. Time series of spawning stock size for the candidate MPs with the robustness test 3 (10% reported overcatch). The top panel represents the historical estimates from the reference case operating model, and lower plots represent the projection period. The solid vertical line represents the last year used in the historical conditioning. The broken vertical line represents the first year that the MP is applied. The median is represented by the bold black line, the dark shaded ribbon represents the 25th-75th percentiles, the light shaded ribbon represents the 10th-90th percentiles. Thick broken lines represent the interim target (green) and limit (red) reference points. The 3 thin coloured lines represent examples of individual realizations (the same OM scenarios across MPs and performance measures), to illustrate that individual variability greatly exceeds the median.

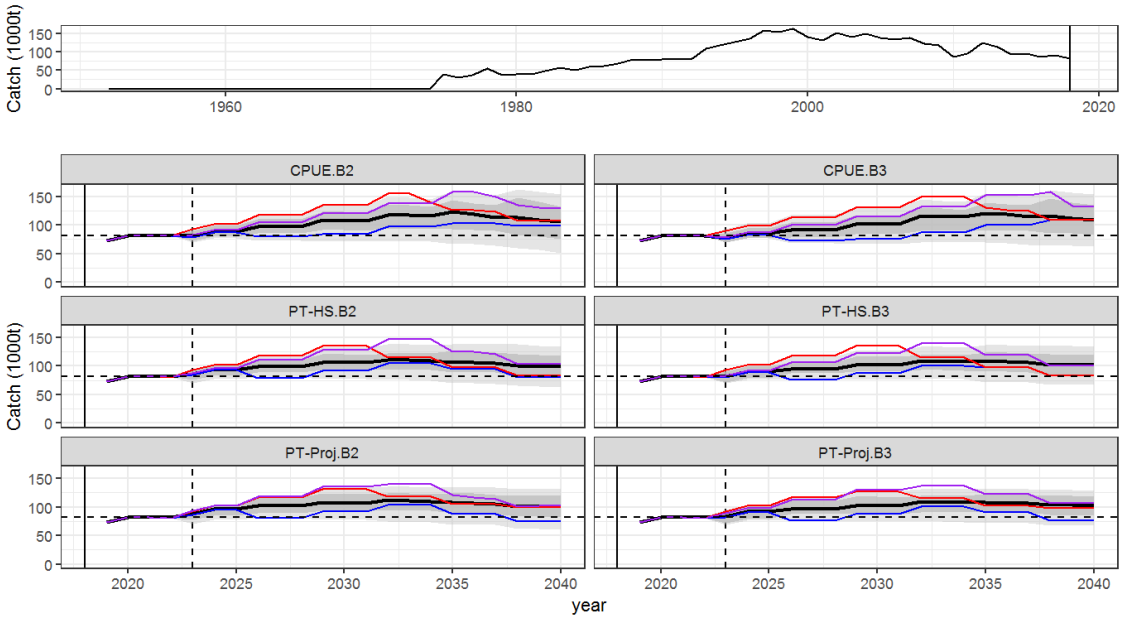

Figure A2.12. Time series of catch for the candidate MPs evaluated with the robustness test 3 (10% reported overcatch). The top panel represents the historical estimates from the reference case operating model, and lower plots represent the projection period. The solid vertical line represents the last year used in the historical conditioning. The broken vertical line represents the first year that the MP is applied. The median is represented by the bold black line, the dark shaded ribbon represents the 25th-75th percentiles, the light shaded ribbon represents the 10th-90th percentiles. The broken black horizontal line represents recent (2017) catch. The 3 thin coloured lines represent examples of individual realizations (the same OM scenarios across MPs and performance measures), to illustrate that individual variability greatly exceeds the median.

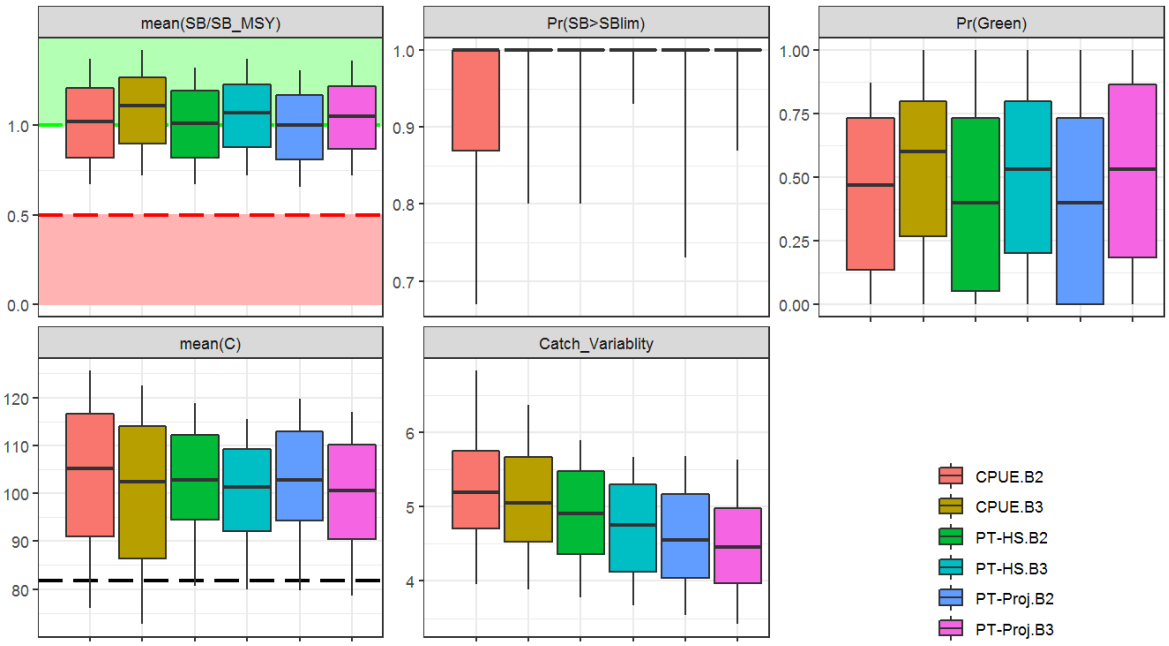

Figure A2.13. Boxplots comparing candidate MPs evaluated with the robustness test 4 (10% unreported overcatch), illustrating key performance measures averaged over the period 2023 - 2037. Horizontal line is the median, boxes represent 25th - 75th percentiles, whiskers (thin vertical lines) represent the 10th-90th percentile range. Red and green horizontal reference lines represent the interim limit and target reference points for the mean SB/SB$_{MSY}$ performance measure. The horizontal dashed black line is the mean reported catch 2017-2019.

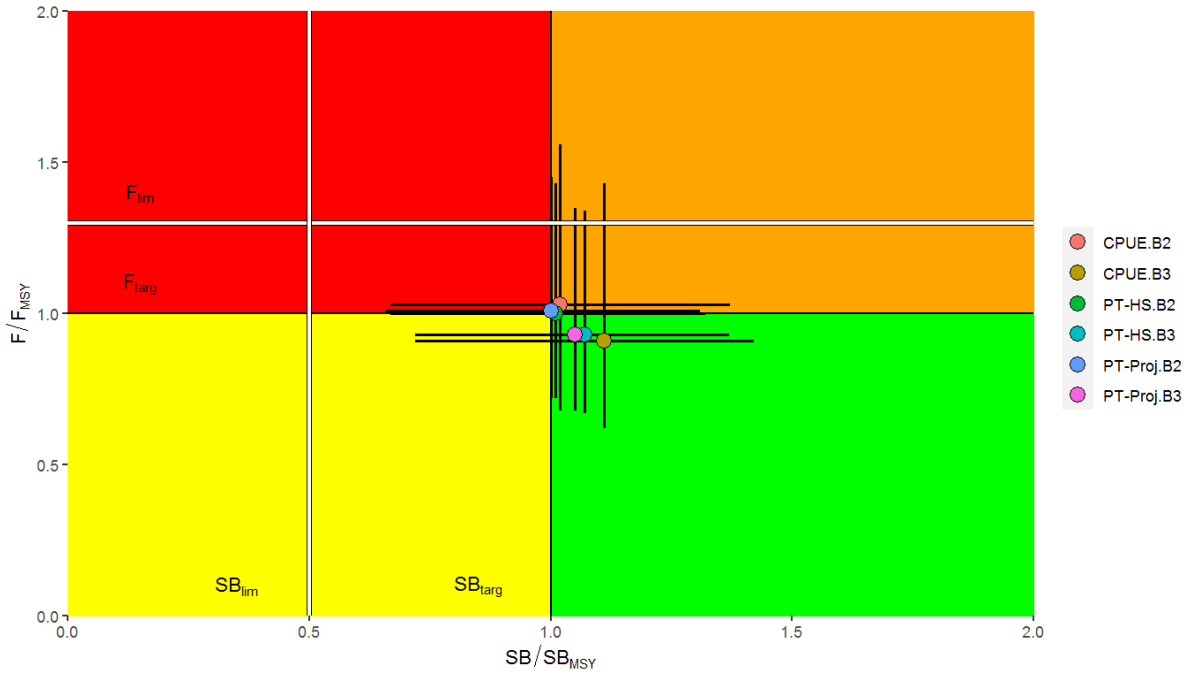

Figure A2.14. Kobe plot comparing candidate MPs evaluated with the robustness test 4 (10% unreported overcatch), on the basis of the expected 15 year average (2023-2037) performance. Circle is the median, lines represent 10th-90th percentiles.

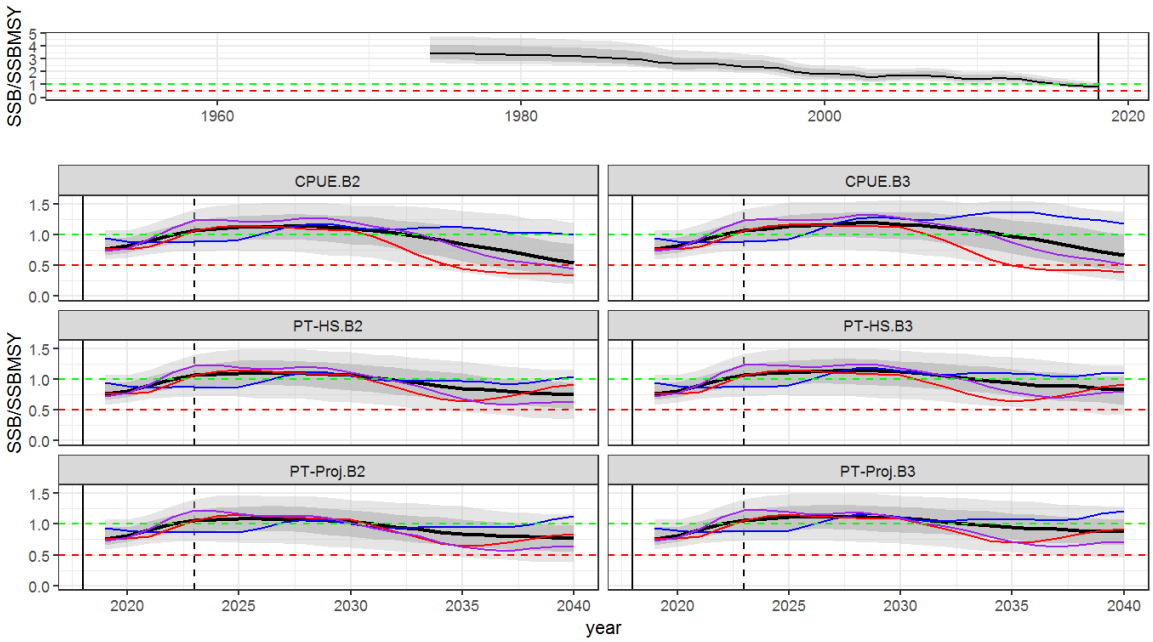

Figure A2.15. Time series of spawning stock size for the candidate MPs with the robustness test 4 (10% unreported overcatch). The top panel represents the historical estimates from the reference case operating model, and lower plots represent the projection period. The solid vertical line represents the last year used in the historical conditioning. The broken vertical line represents the first year that the MP is applied. The median is represented by the bold black line, the dark shaded ribbon represents the 25th-75th percentiles, the light shaded ribbon represents the 10th-90th percentiles. Thick broken lines represent the interim target (green) and limit (red) reference points. The 3 thin coloured lines represent examples of individual realizations (the same OM scenarios across MPs and performance measures), to illustrate that individual variability greatly exceeds the median.

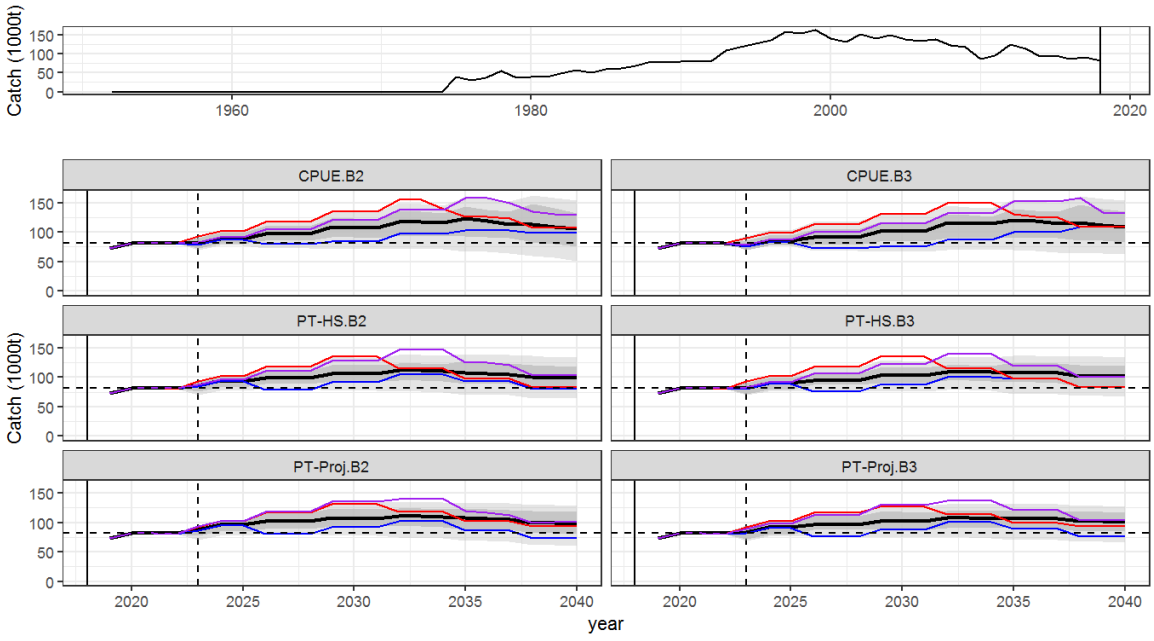

Figure A2.16. Time series of catch for the candidate MPs evaluated with the robustness test 4 (10% unreported overcatch). The top panel represents the historical estimates from the reference case operating model, and lower plots represent the projection period. The solid vertical line represents the last year used in the historical conditioning. The broken vertical line represents the first year that the MP is applied. The median is represented by the bold black line, the dark shaded ribbon represents the 25th-75th percentiles, the light shaded ribbon represents the 10th-90th percentiles. The broken black horizontal line represents recent (2017) catch. The 3 thin coloured lines represent examples of individual realizations (the same OM scenarios across MPs and performance measures), to illustrate that individual variability greatly exceeds the median.

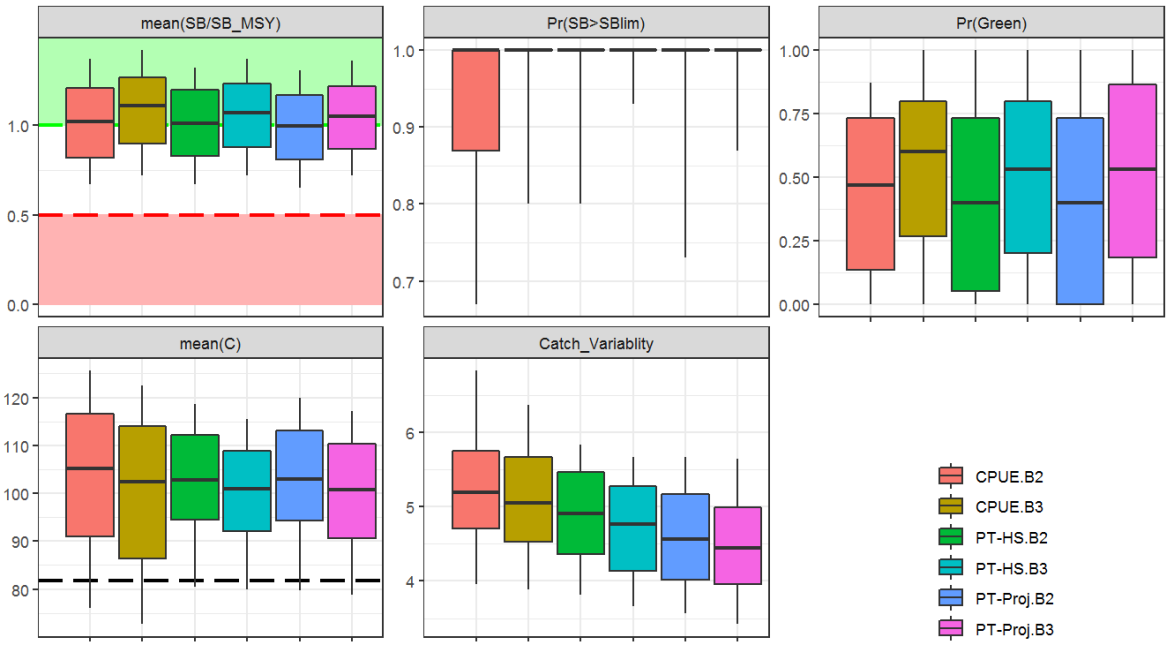

Figure A2.17. Boxplots comparing candidate MPs evaluated with the robustness test 5 (10% overcatch, half of which is reported), illustrating key performance measures averaged over the period 2023 - 2037. Horizontal line is the median, boxes represent 25th - 75th percentiles, whiskers (thin vertical lines) represent the 10th-90th percentile range. Red and green horizontal reference lines represent the interim limit and target reference points for the mean SB/SB_{MSY} performance measure. The horizontal dashed black line is the mean reported catch 2017-2019.

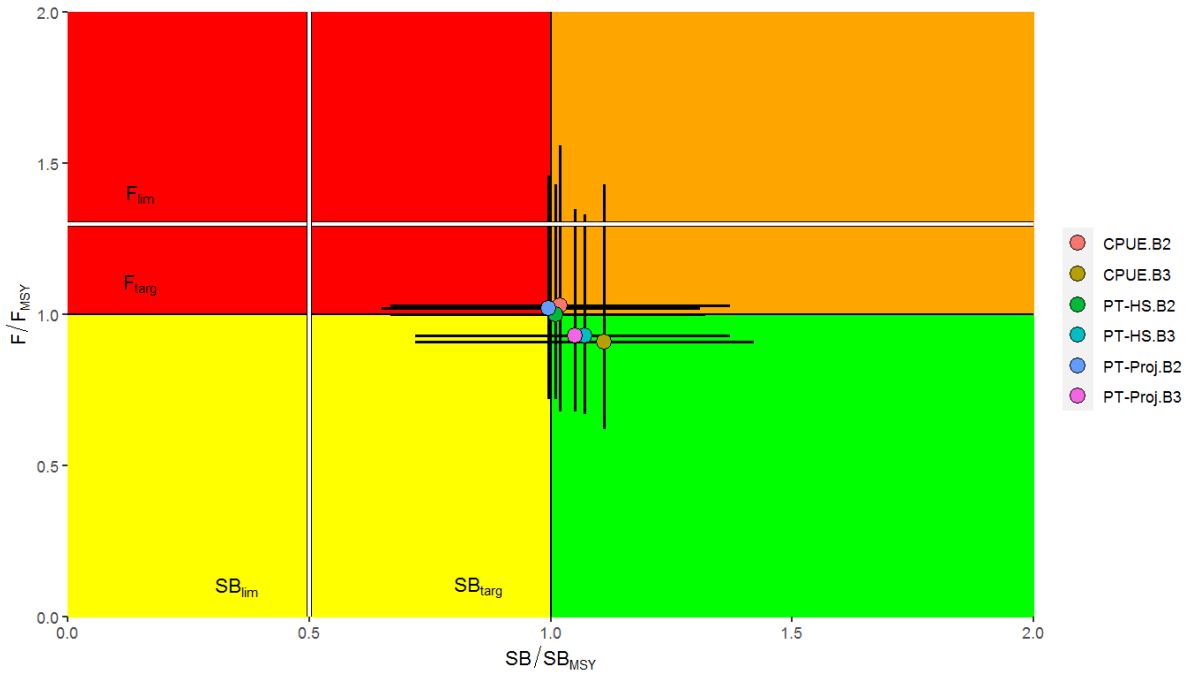

Figure A2.18. Kobe plot comparing candidate MPs evaluated with the robustness test 5 (10% overcatch, half of which is reported), on the basis of the expected 15 year average (2023-2037) performance. Circle is the median, lines represent 10th-90th percentiles.

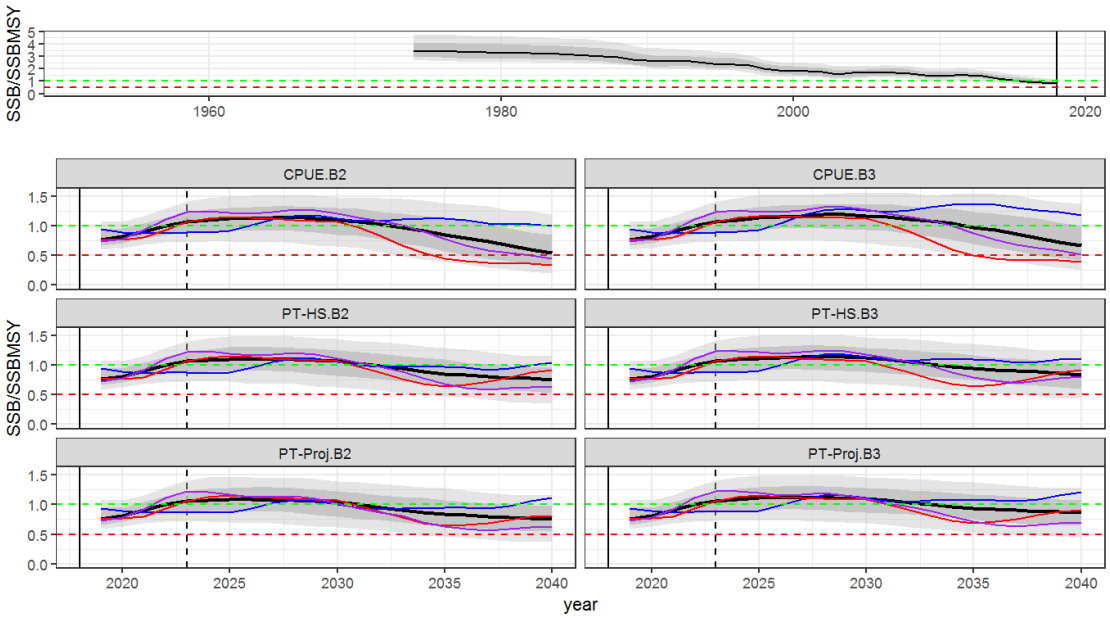

Figure A2.19. Time series of spawning stock size for the candidate MPs with the robustness test 5 (10% overcatch, half of which is reported). The top panel represents the historical estimates from the reference case operating model, and lower plots represent the projection period. The solid vertical line represents the last year used in the historical conditioning. The broken vertical line represents the first year that the MP is applied. The median is represented by the bold black line, the dark shaded ribbon represents the 25th-75th percentiles, the light shaded ribbon represents the 10th-90th percentiles. Thick broken lines represent the interim target (green) and limit (red) reference points. The 3 thin coloured lines represent examples of individual realizations (the same OM scenarios across MPs and performance measures), to illustrate that individual variability greatly exceeds the median.

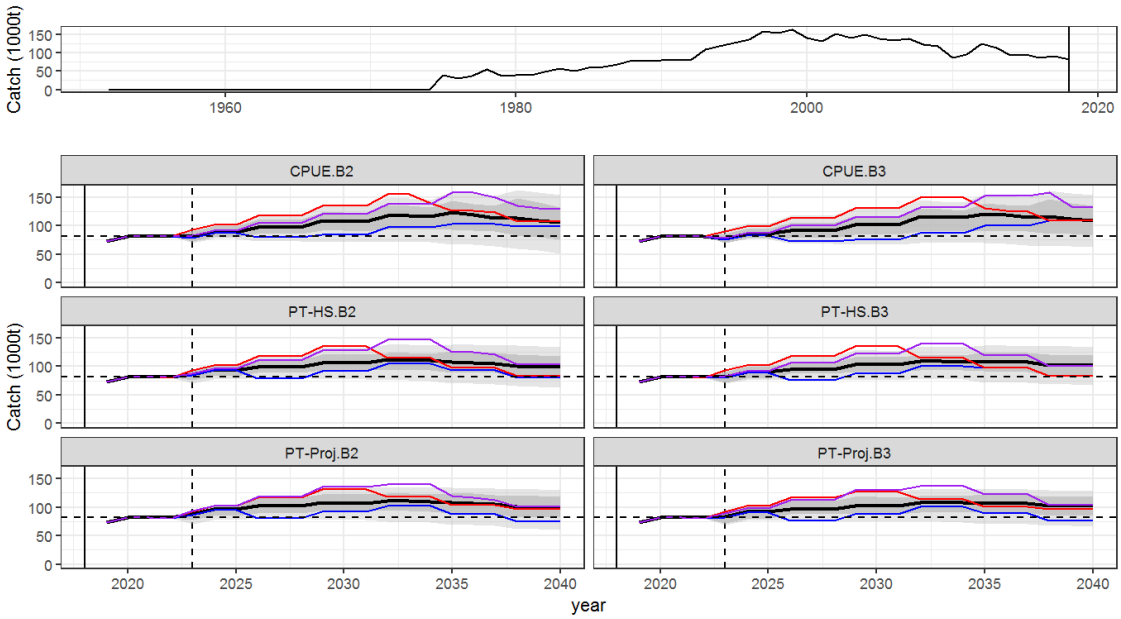

Figure A2.20. Time series of catch for the candidate MPs evaluated with the robustness test 5 (10% overcatch, half of which is reported). The top panel represents the historical estimates from the reference case operating model, and lower plots represent the projection period. The solid vertical line represents the last year used in the historical conditioning. The broken vertical line represents the first year that the MP is applied. The median is represented by the bold black line, the dark shaded ribbon represents the 25th-75th percentiles, the light shaded ribbon represents the 10th-90th percentiles. The broken black horizontal line represents recent (2017) catch. The 3 thin coloured lines represent examples of individual realizations (the same OM scenarios across MPs and performance measures), to illustrate that individual variability greatly exceeds the median.

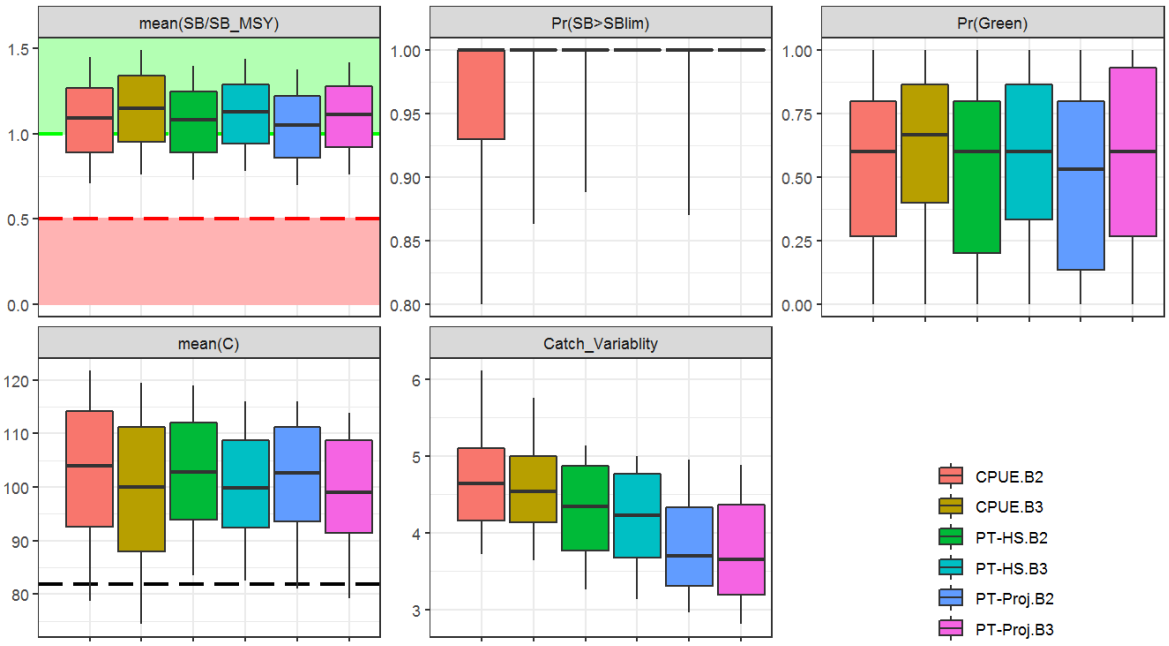

Figure A2.21. Boxplots comparing candidate MPs evaluated with the robustness test 6 (2% per year longline increasing catchability trend), illustrating key performance measures averaged over the period 2023 - 2037. Horizontal line is the median, boxes represent 25th - 75th percentiles, whiskers (thin vertical lines) represent the 10th-90th percentile range. Red and green horizontal reference lines represent the interim limit and target reference points for the mean SB/SB$_{MSY}$ performance measure. The horizontal dashed black line is the mean reported catch 2017-2019.

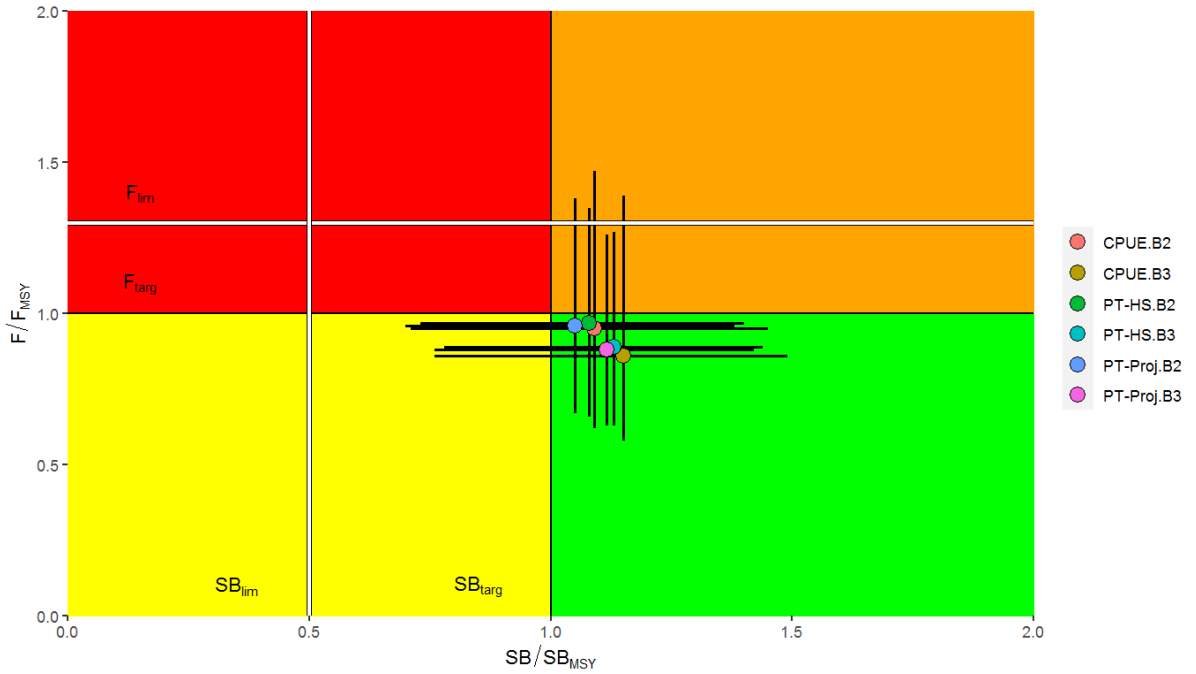

Figure A2.22. Kobe plot comparing candidate MPs evaluated with the robustness test 6 (2% per year longline increasing catchability trend), on the basis of the expected 15 year average (2023-2037) performance. Circle is the median, lines represent 10th-90th percentiles.

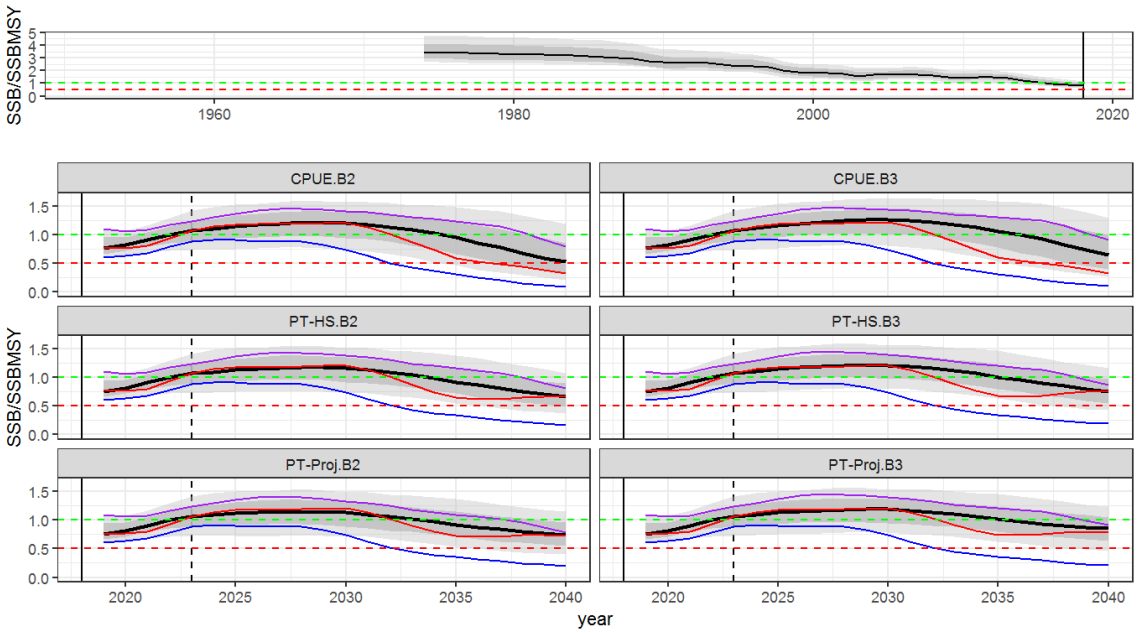

Figure A2.23. Time series of spawning stock size for the candidate MPs with the robustness test 6 (2% per year longline increasing catchability trend). The top panel represents the historical estimates from the reference case operating model, and lower plots represent the projection period. The solid vertical line represents the last year used in the historical conditioning. The broken vertical line represents the first year that the MP is applied. The median is represented by the bold black line, the dark shaded ribbon represents the 25th-75th percentiles, the light shaded ribbon represents the 10th-90th percentiles. Thick broken lines represent the interim target (green) and limit (red) reference points. The 3 thin coloured lines represent examples of individual realizations (the same OM scenarios across MPs and performance measures), to illustrate that individual variability greatly exceeds the median.

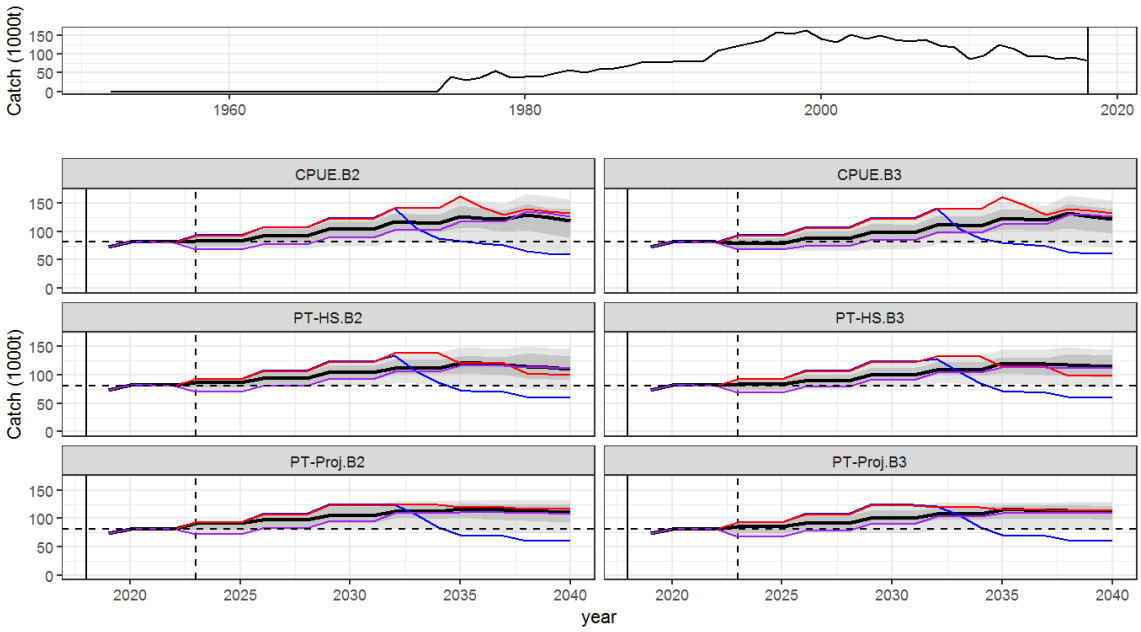

Figure A2.24. Time series of catch for the candidate MPs evaluated with the robustness test 6 (2% per year longline increasing catchability trend). The top panel represents the historical estimates from the reference case operating model, and lower plots represent the projection period. The solid vertical line represents the last year used in the historical conditioning. The broken vertical line represents the first year that the MP is applied. The median is represented by the bold black line, the dark shaded ribbon represents the 25th-75th percentiles, the light shaded ribbon represents the 10th-90th percentiles. The broken black horizontal line represents recent (2017) catch. The 3 thin coloured lines represent examples of individual realizations (the same OM scenarios across MPs and performance measures), to illustrate that individual variability greatly exceeds the median.

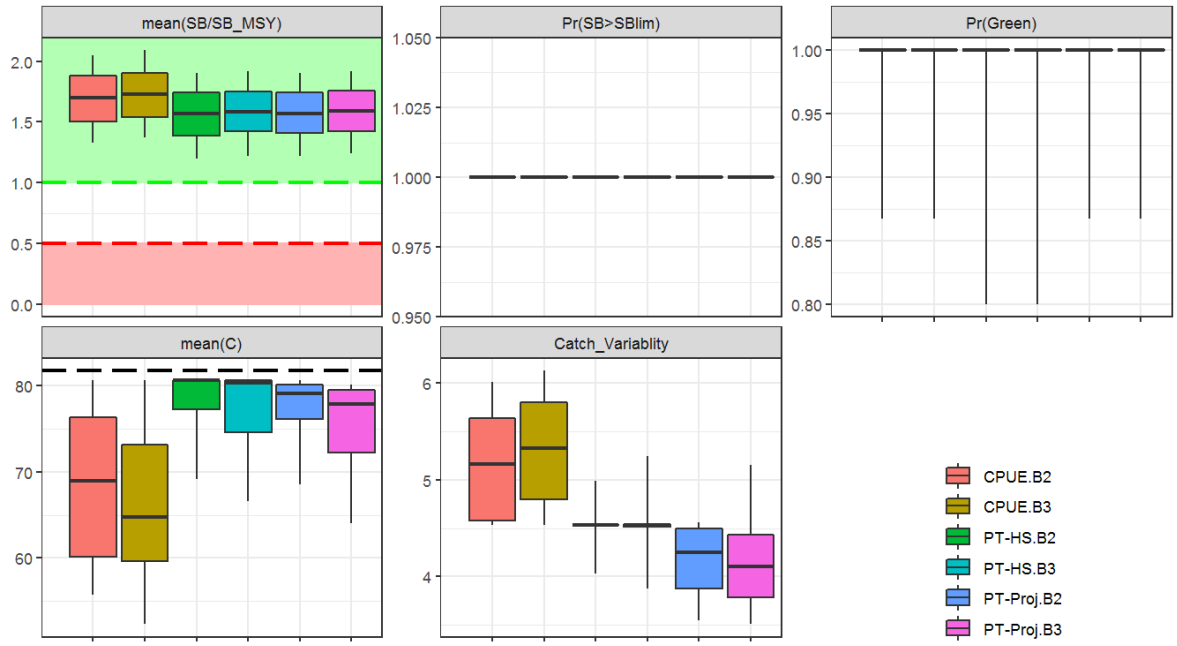

Figure A2.25. Boxplots comparing candidate MPs evaluated with the robustness test 7 (5 years of under-catch), illustrating key performance measures averaged over the period 2023 - 2037. Horizontal line is the median, boxes represent 25th - 75th percentiles, whiskers (thin vertical lines) represent the 10th-90th percentile range. Red and green horizontal reference lines represent the interim limit and target reference points for the mean SB/SB_{MSY} performance measure. The horizontal dashed black line is the mean reported catch 2017-2019.

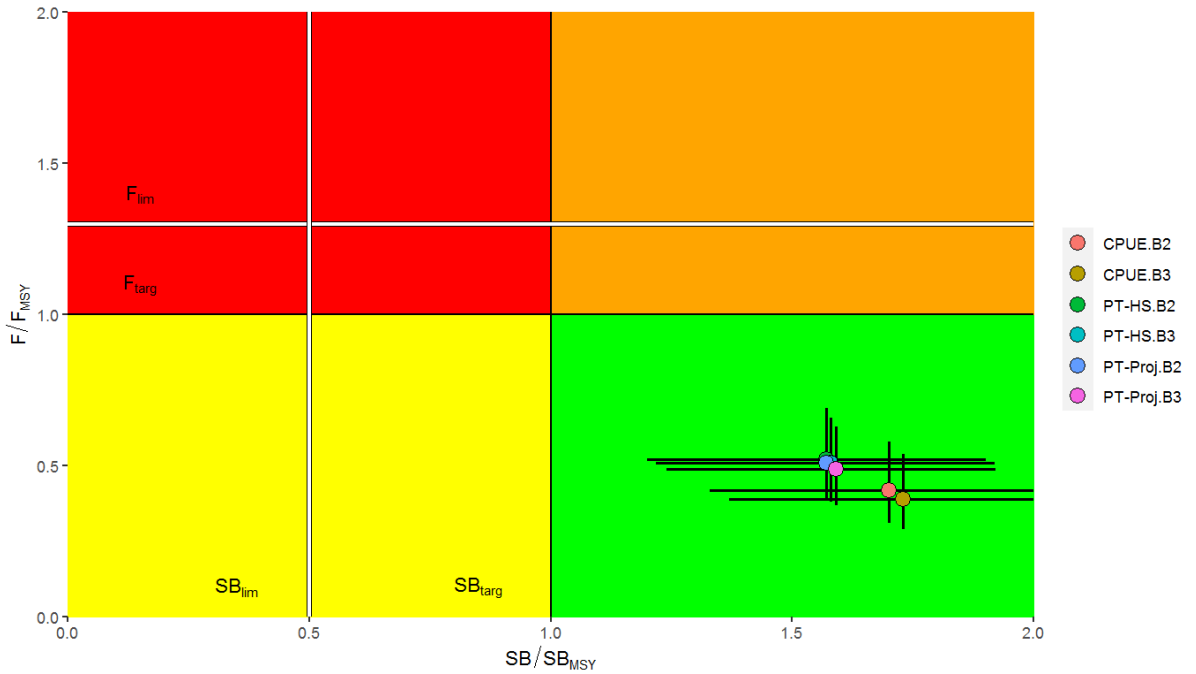

Figure A2.26. Kobe plot comparing candidate MPs evaluated with the robustness test 7 (5 years of under-catch), on the basis of the expected 15 year average (2023-2037) performance. Circle is the median, lines represent 10th-90th percentiles.

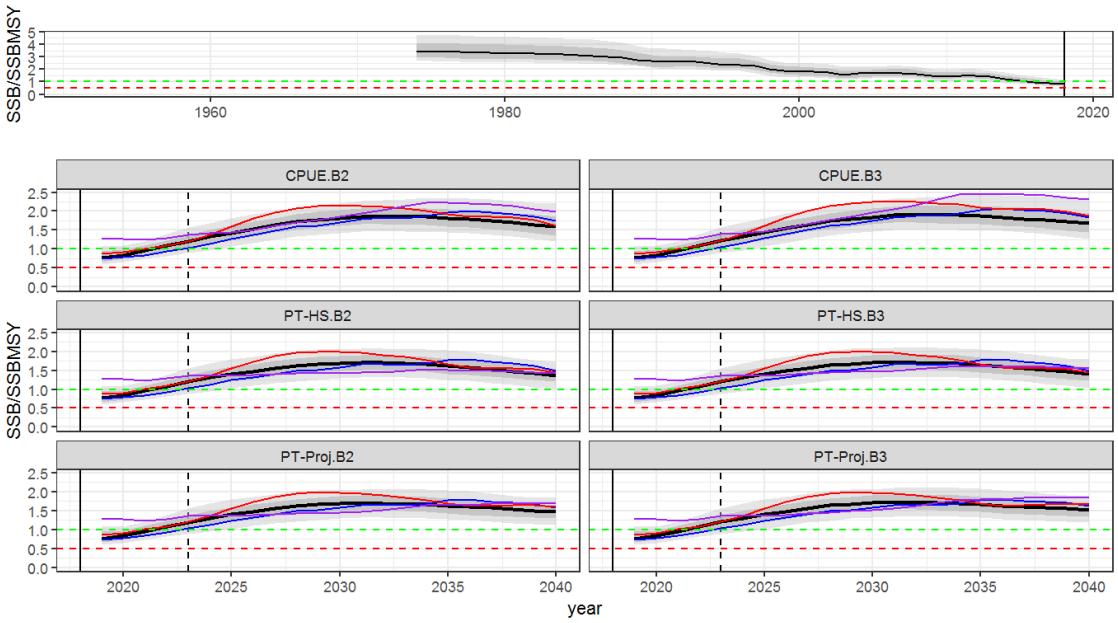

Figure A2.27. Time series of spawning stock size for the candidate MPs with the robustness test 7 (initial 5 years of under-catch). The top panel represents the historical estimates from the reference case operating model, and lower plots represent the projection period. The solid vertical line represents the last year used in the historical conditioning. The broken vertical line represents the first year that the MP is applied. The median is represented by the bold black line, the dark shaded ribbon represents the 25th-75th percentiles, the light shaded ribbon represents the 10th-90th percentiles. Thick broken lines represent the interim target (green) and limit (red) reference points. The 3 thin coloured lines represent examples of individual realizations (the same OM scenarios across MPs and performance measures), to illustrate that individual variability greatly exceeds the median.

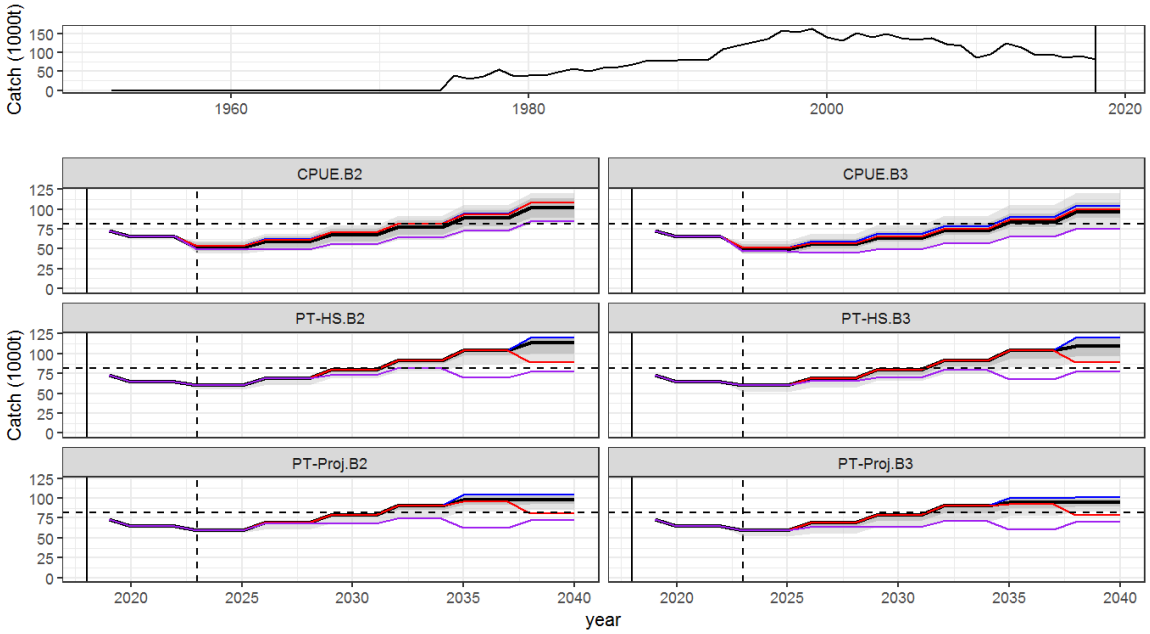

Figure A2.28. Time series of catch for the candidate MPs evaluated with the robustness test 7 (initial 5 years of under-catch). The top panel represents the historical estimates from the reference case operating model, and lower plots represent the projection period. The solid vertical line represents the last year used in the historical conditioning. The broken vertical line represents the first year that the MP is applied. The median is represented by the bold black line, the dark shaded ribbon represents the 25th-75th percentiles, the light shaded ribbon represents the 10th-90th percentiles. The broken black horizontal line represents recent (2017) catch. The 3 thin coloured lines represent examples of individual realizations (the same OM scenarios across MPs and performance measures), to illustrate that individual variability greatly exceeds the median.

Appendix 3. Additional MP Evaluation results[2]

The 6 MPs described in the main text were selected from the 32 MPs (16 functional forms X 2 tuning objectives) presented in this appendix. The selected MPs were chosen to represent the 3 classes of MPs discussed in the main text, where the long form name below is used in the github source code:

CPUE = IT5.t15g2020
PT-HS = PT41FM.t15.tmb
PT-Proj = PTBoB0Targ.t15.pr25

The 6 MPs selected for the main text were from the 32 MPs shown (Figures A.3.2 to A.3.10) are among the best in each class in terms of minimizing the risk of violating biomass limits. Time series behaviour of the MP in the period 2035-2040 (after the tuning period) might provide the most interesting performance contrast. However, the more important point is that all MPs are quite similar, and the tuning objective set by the TCMP is likely to provide the most important performance driver over the medium term period in which it might be expected that an MP would be active without a review.

[2] Note that there is a minor difference between the Operating Models used for Appendix 3 and the main text, related to the assumed bridging catches between the last year of data used in Operating Model conditioning (2018) and the first year of MP-based TAC setting (2023). In appendix 3, the bridging catches for each year 2019-2022 were assumed to be the mean of the reported 2017-2019 catches. The main text adopted the reported 2019 catch for 2019, and the 2017-2019 mean for 2020-2022.

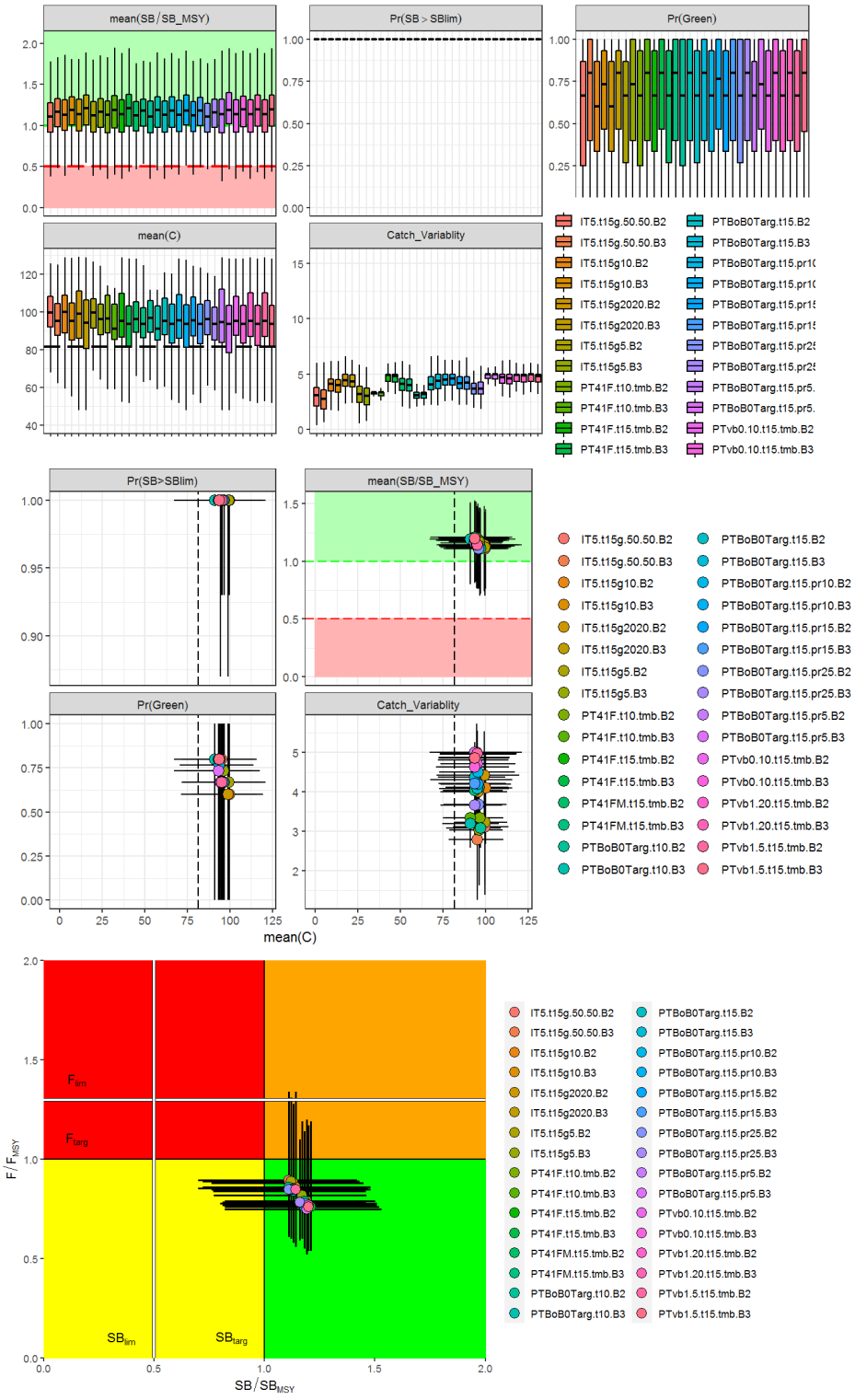

Figure A3.1. Fifteen year (2023-2037) time-aggregated MP evaluation summary plots for 32 MPs tuned MPs (16MPs X 2 Tuning objectives).

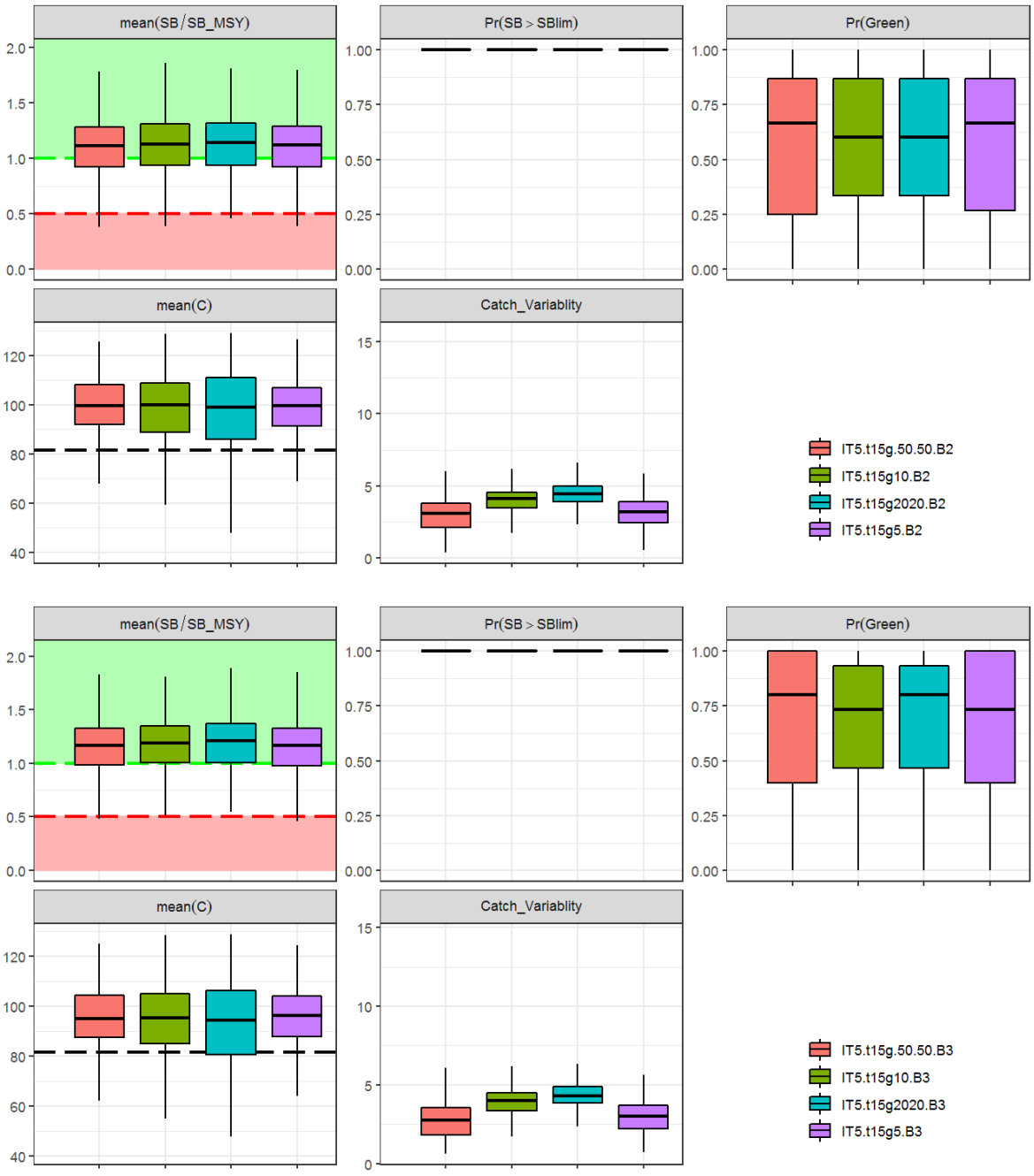

Figure A3.2. CPUE-based MP 15 year (2023-2037) summary plots (top is tuning B2, bottom B3)

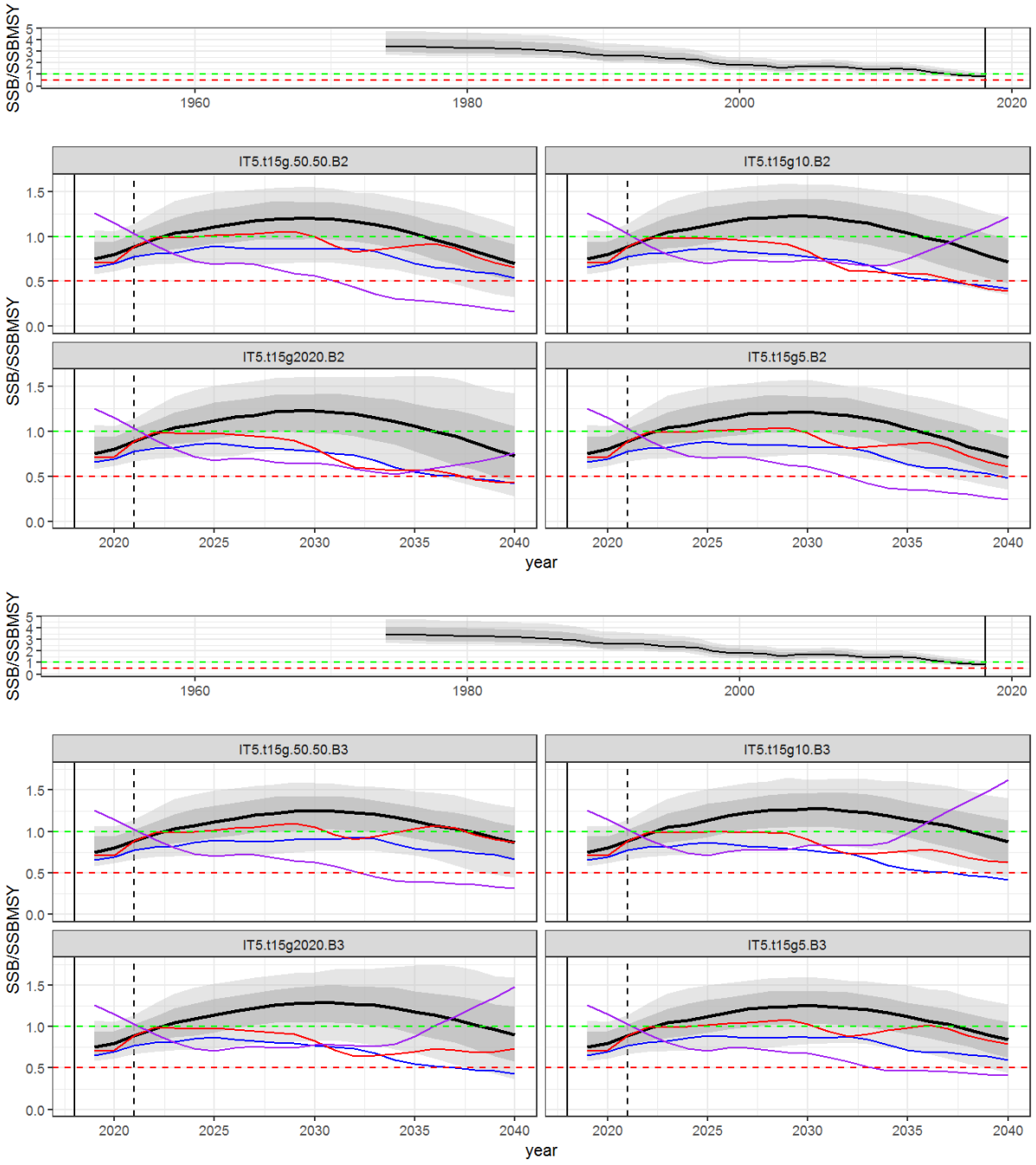

Figure A3.3. CPUE-based MP SSB time series (top is tuning B2, bottom B3)

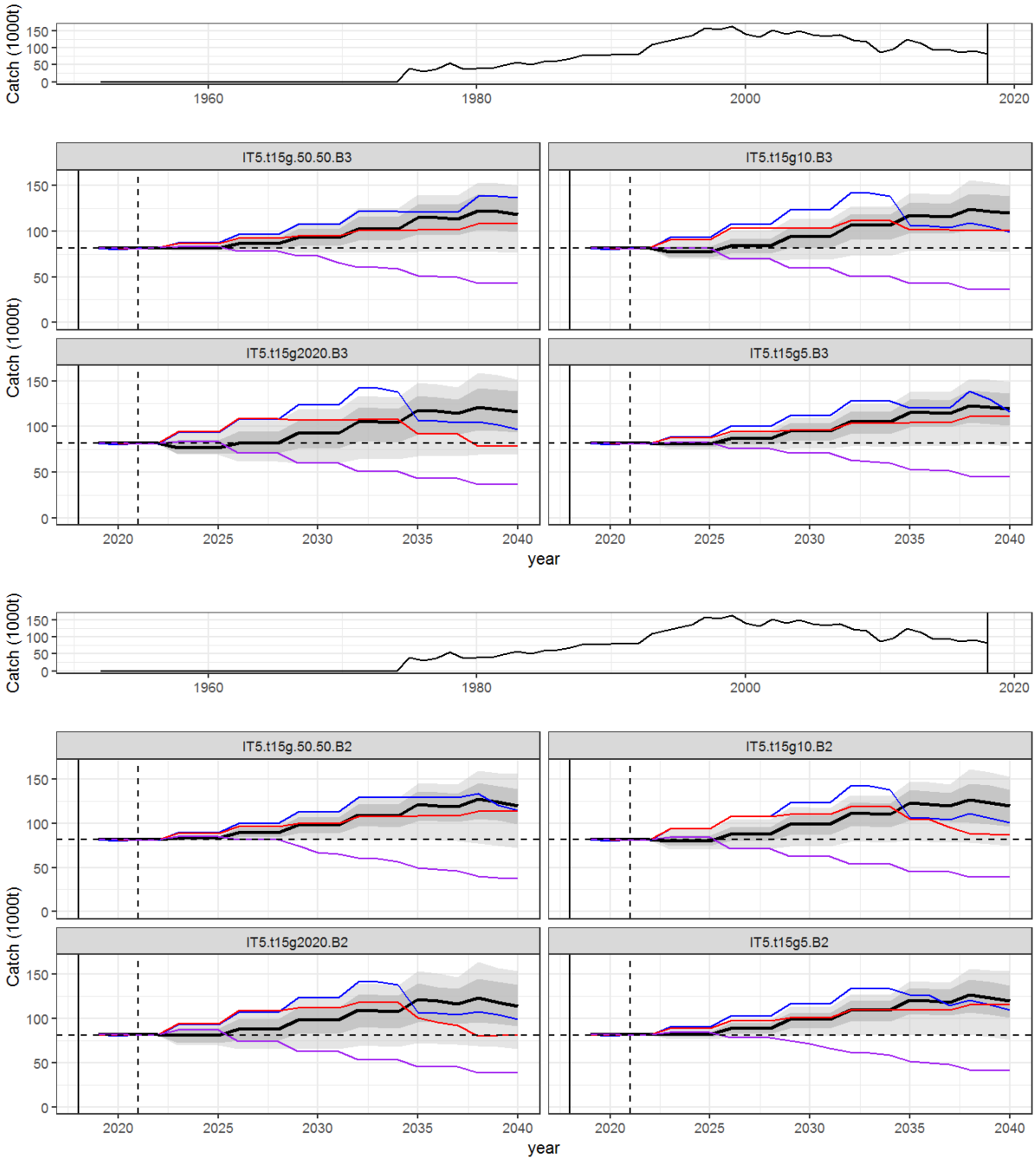

Figure A3.4. CPUE-based MP Catch time series (top is tuning B2, bottom B3)

Figure A3.5. PT-HS Pella-Tomlinson + Hockeystick MP 15 year (2023-2037) summary plots (top is tuning B2, bottom B3)

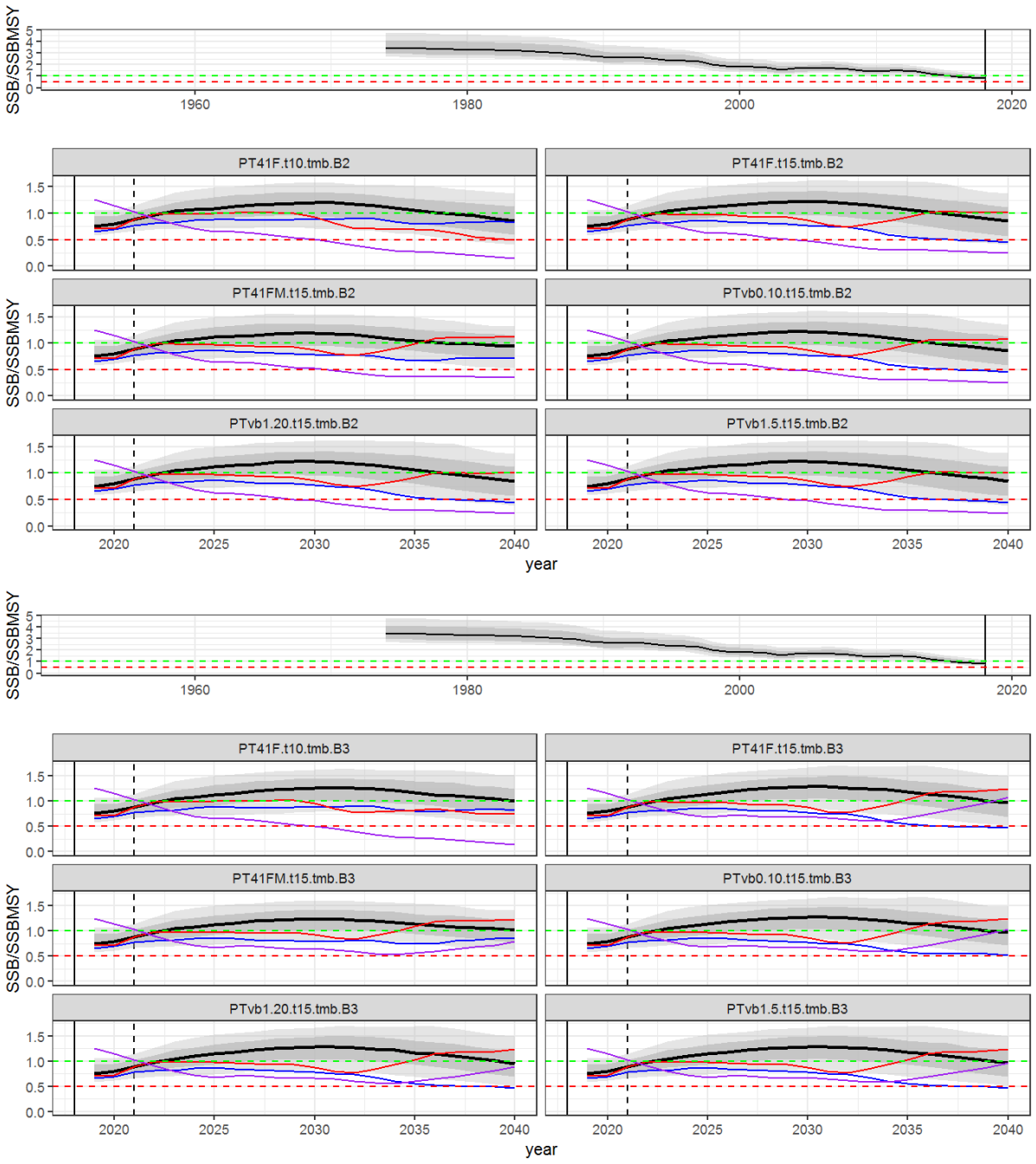

Figure A3.6. PT-HS Pella-Tomlinson + Hockeystick MP SSB time series (top is tuning B2, bottom B3)

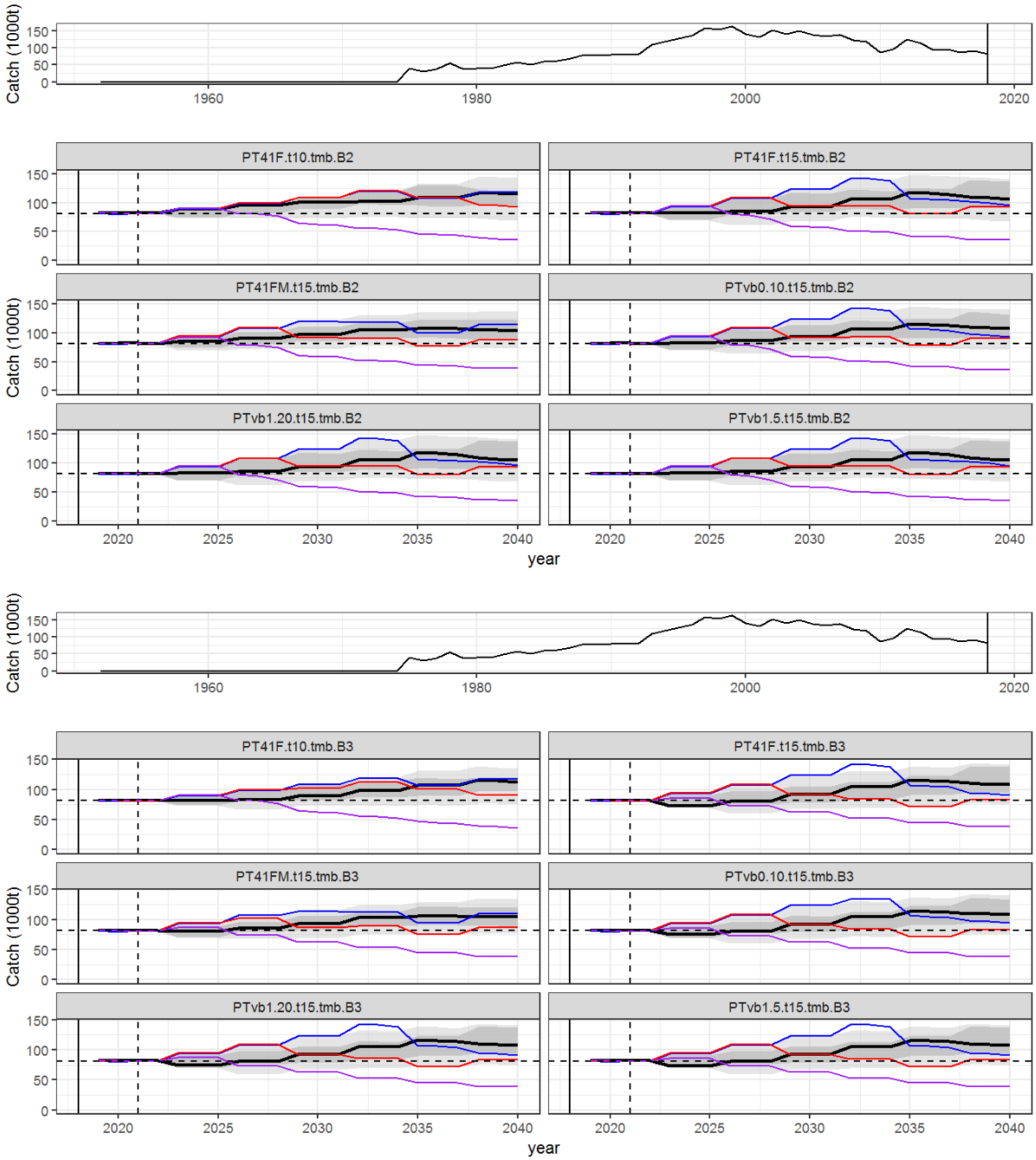

Figure A3.7. PT-HS Pella-Tomlinson + Hockeystick MP Catch time series (top is tuning B2, bottom B3)

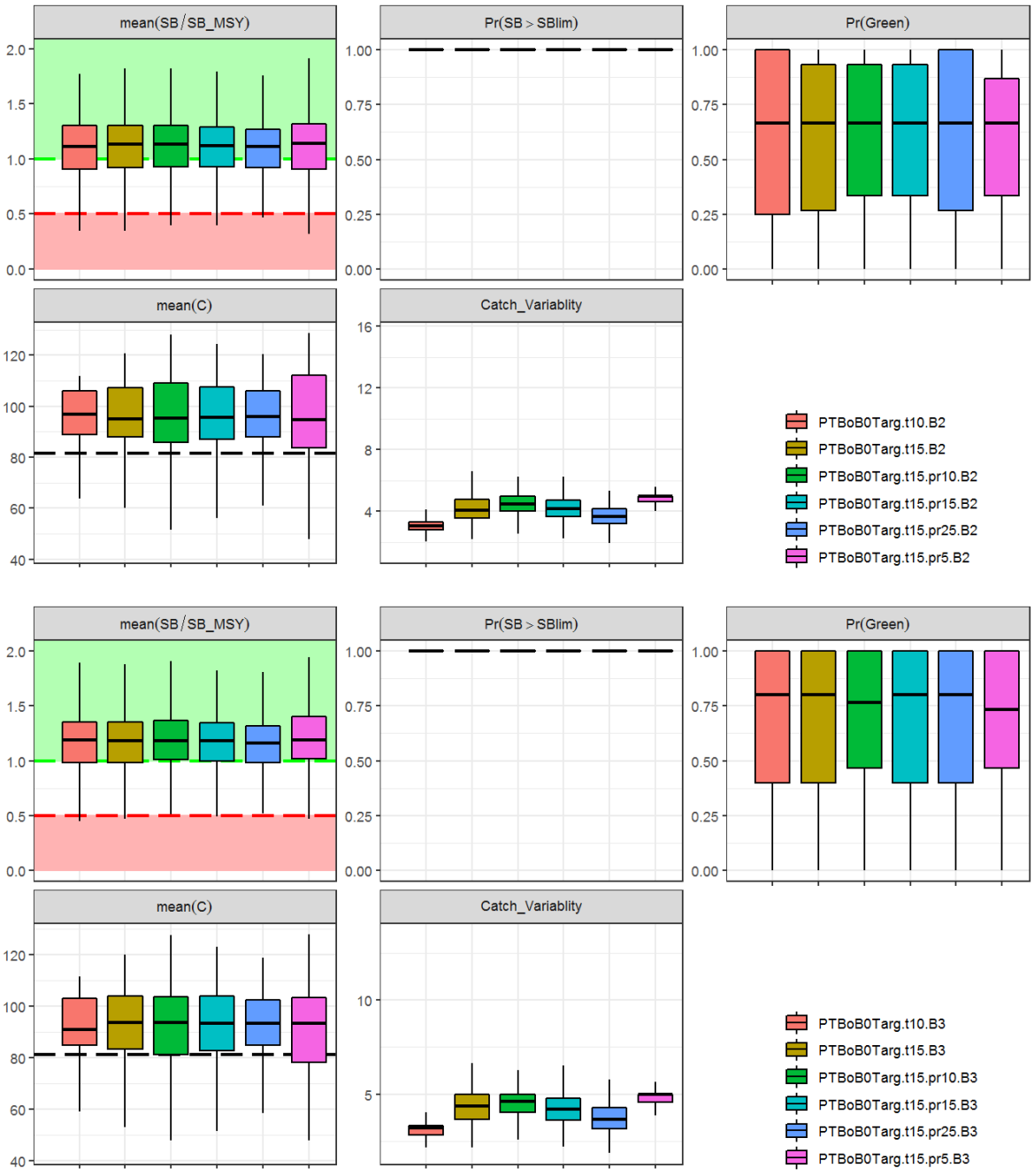

Figure A3.8. PT-Proj Pella-Tomlinson + internal projection MP 15 year (2023-2037) summary plots (top is tuning B2, bottom B3)

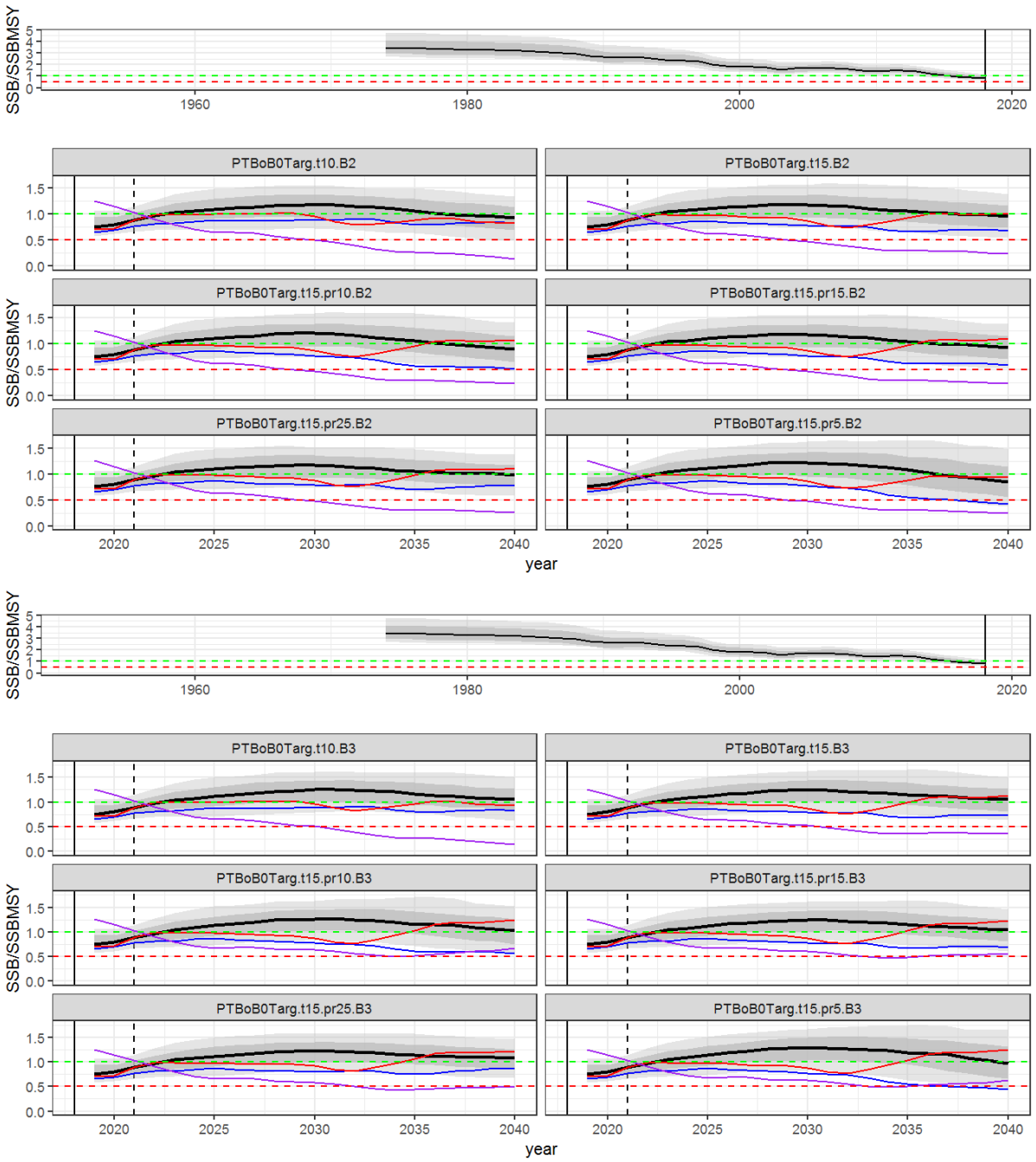

Figure A3.9. PT-Proj Pella-Tomlinson + internal projection MP SSB time series (top is tuning B2, bottom B3)

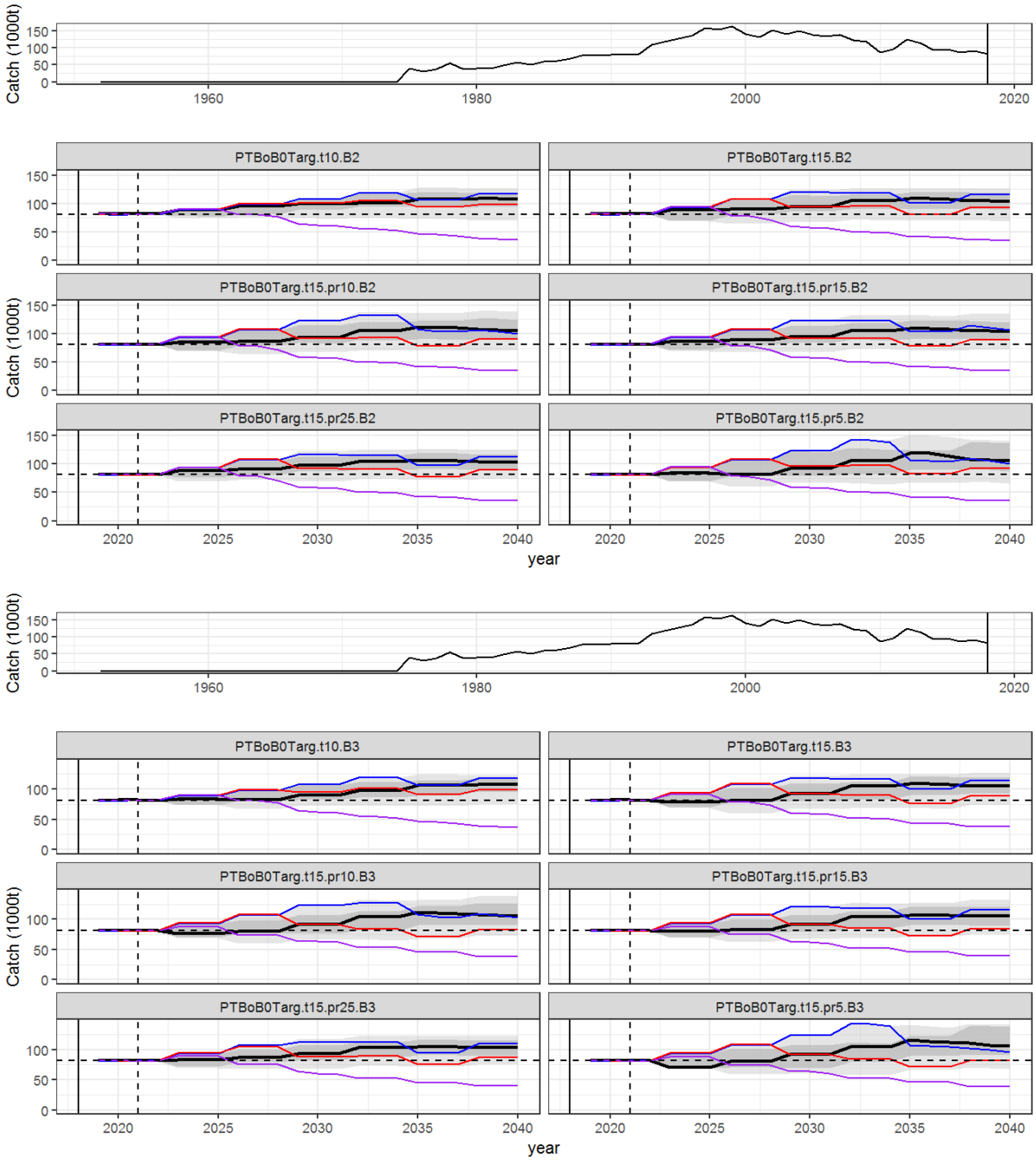

Figure A3.10. PT-Proj Pella-Tomlinson + internal projection MP Catch time series (top is tuning B2, bottom B3)