

## First set of robustness tests conducted for OMP2018

A. Ross-Gillespie and D.S. Butterworth<sup>1</sup>

Email: mlland028@myuct.ac.za

### Summary

Results are presented for a total of ten robustness tests. The first five robustness tests investigate alternative assumptions regarding future surveys and commercial catchability. Further robustness tests involve testing alternative mortality-at-age vectors and sensitivity to data arising from different species splitting algorithms. The final robustness tests explored in this document decrease past or future carrying capacity, simulating a situation where expected recruitment fails. For all robustness tests, simulations suggest that OMP2018 is able to provide adequate management advice for the situations simulated. For robustness tests simulating predictable situations (e.g. if it were known that surveys would be discontinued in the future), the OMP needs to be re-tuned to avoid adverse impacts on TACs and/or resource status. The Appendix reports the results of three additional robustness tests concerning future surveys requested by the Panel during the course of the workshop, none of which indicated that OMP2018 would not be able to provide adequate management advice.

### Introduction

This document provides a full set of results for the robustness tests conducted for OMP2018 prior to the acceptance of the OMP, which imposes a cap of 160 000t and increases the previous OMP tuning parameters  $b_s$  (the parameters which determine the extent to which TAC responds to changes in survey and CPUE indices) by 5%<sup>2</sup>. Furthermore, OMP2018 fixes the 2019 and 2020 TAC at the 2018 TAC increased by 10% (i.e. an increase from 133 119t to 146 431t).

Results for a total of ten robustness tests are included in this document. Given time constraints, these robustness tests were conducted for the Reference Case (RC) Operating Model (OM) only – this is the RS02 OM that assumes a Ricker stock-recruitment relationship with 1958 as the central year of the species shift in the catch (see MARAM/IWS/2018/Hake/P3).

The first five robustness tests relate to assumptions regarding future surveys. Note that the RC model assumes that two surveys (West Coast (WC) and South Coast (SC)) will be conducted each year by the research vessel.

#### RT1: No future surveys (NS)

For missing surveys, the OMP formula for the TAC calculation is adjusted so that the survey indices are removed from the calculation (see MARAM/IWS/2018/Hake/BG4).

#### RT2: All future surveys are conducted by industry vessels (AI)

The standard method for generating future survey indices is:

$$I_y^{surv} = q^{surv} B_y^{surv} e^{\epsilon_y} \quad (1)$$

where  $I_y^{surv}$  is the survey index generated for year  $y$ ,  $q^{surv}$  is the survey catchability coefficient estimated when fitting the Operating Model (OM),  $B_y^{surv}$  is the survey biomass for year  $y$  and  $e^{\epsilon_y}$  is an error term.

Let  $q_R$  be the catchability coefficient of the research vessel (i.e. the *Africana*) and let  $q_I$  be the catchability coefficient of the industry vessel. Let  $R$  be the ratio of the catchability coefficient of the industry to that of the research vessel, i.e.

<sup>1</sup> Marine Resource Assessment and Management Group, Department of Mathematics and Applied Mathematics, University of Cape Town, Rondebosch 7701.

<sup>2</sup> Note that when the  $b$  parameters for *M. paradoxus* and for *M. capensis* are modified, this is always such as maintains their original ratio. Sometimes therefore, these  $b$  parameters are referenced in the singular rather than the plural.

$R=q_I/q_R$ . MARAM/IWS/2018/Hake/BG5 reports that this ratio  $R$  is distributed with a median of 0.80 and standard error of 0.18 (the computations to come assume this distribution to be log-normal).<sup>3</sup>

When generating future survey indices under the assumption that an industry vessel will conduct the surveys, a further error term  $e^{\eta_y}$  is added to equation (1) with  $\eta_y \sim N(\ln 0.80, 0.22^2)$  so that  $e^{\eta_y}$  is log-normally distributed with median of 0.80 and standard error of 0.18.<sup>4</sup> The whole equation is further scaled by a factor of 1/0.80 to account for the likely lower catchability of the industry vessels (this factor is included because this bias would be known, so that in application the result from a future survey would be divided by this factor). Equation 1 consequently becomes:

$$I_y^{surv} = q^{surv} B_y^{surv} e^{\epsilon_y} e^{\eta_y} / 0.80 \tag{2}^5$$

**RT3: Two surveys (WC and SC) take place every second year (ES)**

As for no future surveys, the survey indices are removed from the TAC calculation in years for which no survey is to take place.

**RT4: RC assumptions (i.e. all future surveys continue as normal with the research vessel) but with an undetected increase in CPUE catchability of 2% p.a. (RC+UCI).**

For an undetected increase in commercial catchability, a steady increase of 2% p.a. in the CPUE  $q$  value was assumed, as this was considered a likely extreme case scenario.

**RT5: No future surveys with an undetected increase in catchability (NS+UCI).**

Robustness test (1) is combined with an undetected increase in commercial catchability, as for Robustness test RT4.

The remaining five robustness tests conducted prior to the recommendation of OMP-2018 are described below.

**RT6: The original natural mortality-at-age vectors (OM)**

The RC OM is reconditioned using the mortality-at-age vectors from the Rademeyer 2017 model for the two hake species instead of the ones estimated by the predation model. These mortality-at-age vectors are fixed at 0.75 for lower ages and 0.375 for higher ages.

**RT7: Use of the original algorithm 2013 species splitting results for the CPUE and catch data (A13).**

The RC OM is reconditioned using the data from the original Algorithm 2013 species splitting algorithm, which is the same species splitting algorithm used to generate the CPUE and catch data for the 2017 assessment model. MARAM/IWS/2018/Hake/BG6 provides more details on the species splitting algorithms, with the table below summarising the key differences and similarities between the three models that have been considered for the RC and robustness tests.

<b>Alg13 (robustness)</b>	<b>Model A6 (robustness)</b>	<b>Model A6b (RC)</b>
Research data only	Research + observer data	Research + observer data
Random trawlID effect excluded	Random trawlID effect included	Random trawlID effect excluded
Fish numbers	Fish mass (unweighted)	Fish mass (unweighted)
Includes fish size category and depth as covariates	Includes fish size class and geographical location (lat/long) as categorical variables and depth as a covariate	Includes fish size class and geographical location (lat/long) as categorical variables and depth as a covariate

**RT8: Use of the Model A6 species splitting results for the CPUE and catch data (A6).**

<sup>3</sup> When a single  $q$  value for all industry vessels was estimated in the RC OM, the resulting estimate was roughly 0.80, which is consistent with what has been estimated in MARAM/IWS/2018/Hake/BG5.

<sup>4</sup> If  $\ln R$  is normally distributed with mean  $\mu$  and variance  $\sigma^2$ , then the median of the distribution of  $R$  is given by  $\theta_R = e^\mu$  and the variance by  $\sigma_R = (e^{\sigma^2} - 1)(e^{2\mu + \sigma^2})$ . Since  $\theta_R = 0.80$  and  $\sigma_R = 0.18$ , it follows that  $\mu = \ln 0.80$  and  $\sigma = 0.22$ .

<sup>5</sup> Note that there was an error in this equation in FISHERIES/2018/OCT/SWG-DEM/71, which has been corrected here and the robustness test has been re-run.

The RC OM is reconditioned using the data from an alternative species splitting algorithm Model A6 (see table above and MARAM/IWS/2018/Hake/BG6 for more details).

**RT9: Decrease past carrying capacity (PK).**

Carrying capacity for both species decreases linearly by 30% between 1980 and 2000. Carrying capacity before 1980 is at the 1917 value, while carrying capacities after 2000 are at 70% of their 1917 values.

**RT10: Decrease future carrying capacity (FK)**

Carrying capacity for both species decreases linearly by 30% between 2018 and 2022. Carrying capacities after 2022 are at 70% of their 1917 values.

For robustness tests 9 and 10 that decrease the carrying capacity, a dynamic  $B_{MSY}$  has been calculated<sup>6</sup>.

The Appendix reports results for three additional robustness tests concerning future surveys that were conducted during the course of the workshop.

### Tuning of robustness tests

For robustness tests simulating situations that are predictable (i.e. it is known beforehand that they will occur) and for which management action could in theory be taken (e.g. a situation where no future surveys take place), the robustness tests have been re-tuned so that there is “equivalent risk” as the RC OM. This has been done as there is a trade-off between changes in catches and final abundance as the projection assumptions are changed, making comparisons difficult. For “equivalent risk” results, the value for the  $b_s$  control parameters for each relevant robustness test (i.e. the three robustness tests which simulate “known” situations, namely NS, AI and ES) have been modified so that the lower 5%-ile on the final (2042) *M. paradoxus* depletion is identical to that for the RC. Again, the  $b_s$  parameters have been modified so as to maintain the original *M. paradoxus* to *M. capensis* ratio.

### Results

The primary performance statistics are listed in Table 1a and b. Graphic illustrations of the performance statistics for the RC and 10 robustness tests are given in Figure 1a and b, while the corresponding plots of the median “trajectories”<sup>7</sup> and probability envelopes for  $B/B_{MSY}$ , catch, effort and CPUE are shown in Figure 2a and b.

### Discussion

Some brief discussion points are noted below.

#### Future surveys

- After re-tuning, results for the scenario of future surveys every second year are fairly similar to the RC projections. If surveys are undertaken by industry vessel, this may result in slightly higher TACs over the next 25 years, but not substantially so. The biggest impact is associated with no future surveys; however provided that the OMP is re-tuned should this scenario occur, the impact is mitigated and the OMP seems able to provide reliable management advice.
- Should an unexpected increase in the commercial catchability occur (both with and without future surveys) this will likely lead to higher overall TACs and consequently a somewhat more depleted *M. paradoxus* resource,

<sup>6</sup> Dynamic  $K$  is the trajectory that would be followed by a population in the absence of catches in circumstances where population dynamics parameters (in particular those of the stock-recruitment relationship) change over time. This differs from the trajectory for  $K$  itself as a result of demographically induced delay effects. Dynamic  $B_{MSY}$  is conventionally assumed to remain the same proportion of dynamic  $K$  during such changes. For the performance statistics in the tables though, true and dynamic  $K$  are the same, as in the years for which the statistics are reported, the transient effects associated with the differences between the two have died out, so that the extra computational complexity of computing dynamic  $B_{MSY}$  for all years could be bypassed. The  $B/B_{MSY}$  trajectories plotted in Figure 2b for the robustness tests changing  $K$  are, however, consequently marginally in error.

<sup>7</sup> The reason for placing trajectories in quotes here is that the means and medians here do not reflect actual achieved trajectories in some simulation, but rather joining the mean or median values from the predicted distributions for successive future years. Quotation marks will not be used for this word elsewhere in the document.

although the lower 5<sup>th</sup> percentile of the  $B_{2042}/B_{MSY}$  estimate remains above or close to one in all cases. It would seem that the undetected increase in catchability is not likely to have serious management implications.

- There is very little difference in the projections using the Model A6b and Model A6 species-splitting data. When comparing the Model A6b results with those from the original Algorithm 2013 species splitting algorithm, it is interesting to note that there is very little difference between the two models in the short-term projected TAC values for the next few years, despite the more optimistic estimate of the status of depletion for *M. paradoxus* under the new Model A6b species splitting algorithm. The 25-year TAC projections are actually more optimistic under the old species splitting algorithm (possibly because of the less depleted *M. capensis* and higher exploitable biomass estimated by this model). The simulations suggest that depletion levels for *M. paradoxus* are fairly similar after 25 years regardless of which species splitting algorithm has been assumed.
- If the original mortality-at-age vectors assumed for the 2017 model were to be closer to reality than those estimated by the predation model, the OMP still performs adequately, with slightly higher TACs predicted over the next 25 years.
- Decreasing past carrying capacity (simulating a situation where recruitment fails) results in a more optimistic perception of the *M. paradoxus* resource, although the lower bounds of the *M. capensis* depletion estimates are more pessimistic. However, the method for calculation of dynamic  $B_{MSY}$  could still be refined, in which case the  $B/B_{MSY}$  estimates might change. Projected catch in this case is fairly similar to the RC. Decreasing future carrying capacity has a more severe negative impact on projected TAC, with a projected increase in effort and decrease in CPUE in the short-term. However, long-term projections are more similar to the RC, with  $B/B_{MSY}$  estimates well above one for both species at the end of the projection period, suggesting that the rules of the OMP should be able to cope adequately with a situation where recruitment fails in the future.
- None of the robustness tests reported on in this document provide evidence that OMP2018 could not be used to sustainably manage the hake resource in the various situations simulated.

## Acknowledgements

Computations were performed using facilities provided by the University of Cape Town's ICTS High Performance Computing team: [hpc.uct.ac.za](http://hpc.uct.ac.za). AR-G acknowledges post-doctoral support from the Claude Leon Foundation.

**Table 1a:** The first row of the Table list the performance statistics for OMP2018 (cap of 160 00t and a 5% increase in  $b$ ), which are reported as medians of the equally weighted simulations from the nine RS models (i.e. medians of 9x100 values) with their 90% probability intervals, except for the second columns under  $TAC_{av}$ (25 yrs) and  $TAC_{av}$ (4yrs) where the means and standard deviations are also listed.

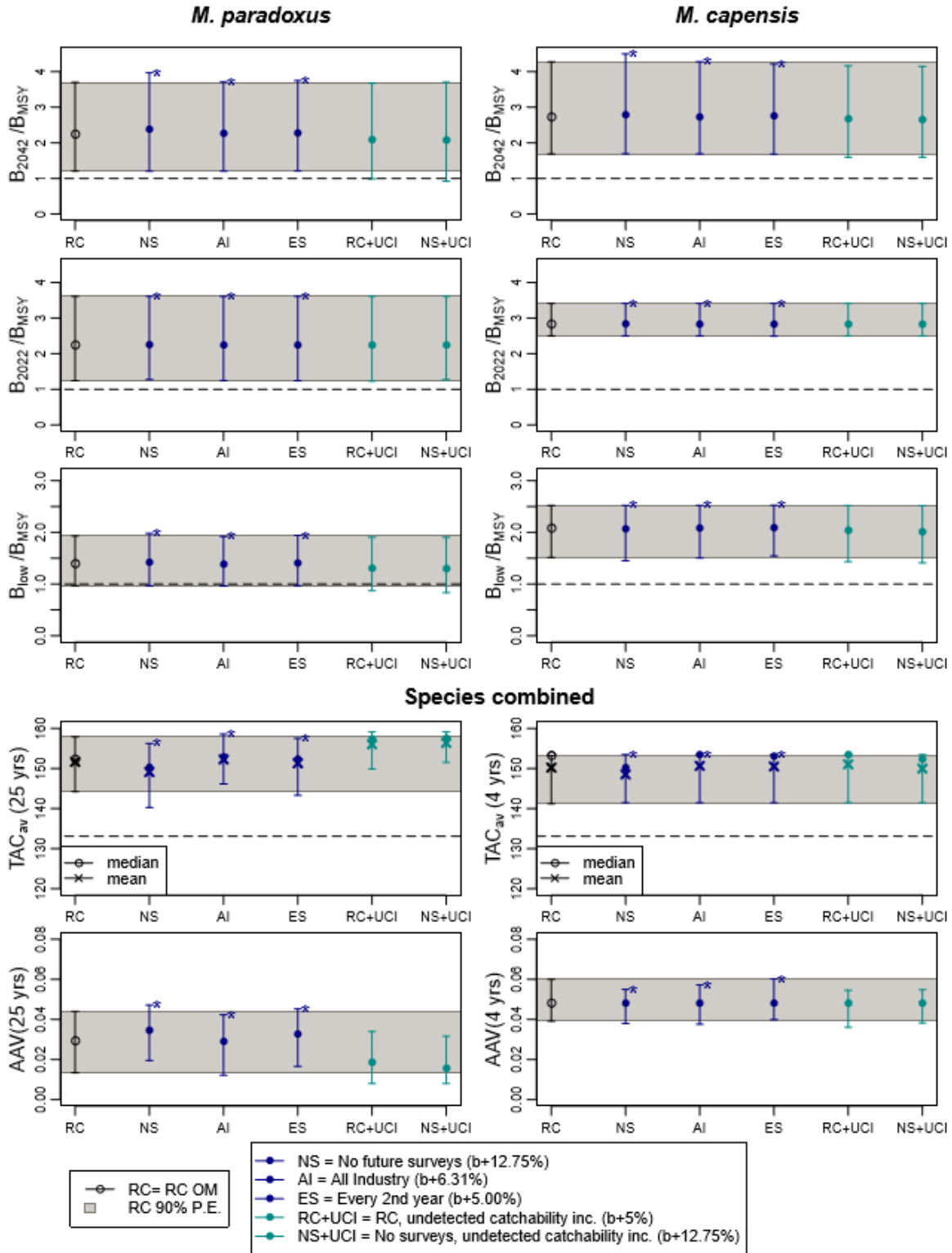
The second part of the table lists the performance statics for the RS02 Reference Case (RC) OM (the RS OM with Ricker stock recruitment and central year of catch shift 1958) and 10 robustness tests, where all tests have been conducted for RS02 only. Where robustness tests involve future occurrences for which it would likely be known beforehand whether they would occur (marked with a star in the table below), the robustness test  $b_s$  parameters were tuned so that the risk (defined as the lower 5<sup>th</sup> percentile of the  $B_{2042}/B_{MSY}$  value for *M. paradoxus*) is within half a percent of the RC results. The abbreviation “NS” stands for “no surveys” and “UCI” stands for “undetected catchability increase”.

The performance statistics here in Table 1a are  $B^{sp}/B_{MSY}$  for 2042 (final year of the projection period) and for 2022 (end of the life span of OMP2018) and  $B_{low}/B_{MSY}$  (the lowest value of this statistic in the projection period).

		<i>M. paradoxus</i>						<i>M. capensis</i>							
		Cap	b	$B_{2042}/B_{MSY}$		$B_{2022}/B_{MSY}$		$B_{low}/B_{MSY}$		$B_{2042}/B_{MSY}$		$B_{2022}/B_{MSY}$		$B_{low}/B_{MSY}$	
Equal weighting across the nine RS OMs															
Proposed OMP2018		160	+5%	2.67	(1.37,4.81)	2.10	(1.01,3.43)	1.47	(0.90,2.15)	3.01	(0.91,6.70)	2.94	(0.39,5.18)	2.40	(0.32,4.52)
Robustness test results for RS02 only															
0	RC	160	+5%	2.24	(1.21,3.70)	2.24	(1.24,3.61)	1.39	(0.97,1.93)	2.73	(1.69,4.27)	2.83	(2.50,3.41)	2.08	(1.51,2.52)
1	No surveys*		+12.75%	2.38	(1.20,3.96)	2.25	(1.27,3.61)	1.42	(0.97,1.98)	2.79	(1.69,4.50)	2.84	(2.50,3.41)	2.07	(1.45,2.52)
2	All industry*		+1.25%	2.27	(1.21,3.73)	2.24	(1.26,3.61)	1.40	(0.96,1.96)	2.72	(1.71,4.29)	2.83	(2.50,3.41)	2.07	(1.52,2.52)
3	Every second year*	160	+5%	2.28	(1.21,3.75)	2.25	(1.24,3.61)	1.41	(0.97,1.94)	2.75	(1.68,4.21)	2.83	(2.49,3.41)	2.09	(1.54,2.52)
4	RC + UCI		+5%	2.09	(0.99,3.67)	2.24	(1.23,3.61)	1.31	(0.87,1.90)	2.68	(1.60,4.16)	2.83	(2.50,3.41)	2.04	(1.43,2.52)
5	NS+ UCI		+12.75%	2.08	(0.92,3.70)	2.24	(1.27,3.61)	1.30	(0.83,1.90)	2.65	(1.60,4.15)	2.83	(2.50,3.41)	2.01	(1.41,2.52)
6	Original mortality		+5%	2.28	(0.98,4.12)	2.25	(1.24,3.66)	1.31	(0.81,1.76)	2.75	(1.72,4.65)	2.76	(2.25,3.48)	1.78	(1.20,2.34)
7	Alg13 sp. splitting		+5%	2.16	(1.19,3.60)	1.93	(1.06,3.05)	1.37	(0.89,1.74)	3.11	(2.00,4.39)	3.11	(2.70,3.71)	2.58	(1.94,2.80)
8	Model A6 sp. splitting	160	+5%	2.23	(1.26,3.70)	2.35	(1.37,3.70)	1.47	(1.03,1.93)	2.85	(1.75,4.40)	2.92	(2.57,3.47)	2.18	(1.58,2.61)
9	Decrease past K		+5%	3.65	(1.89,6.02)	3.67	(1.88,6.01)	2.14	(1.49,2.94)	2.54	(1.31,4.18)	2.55	(1.30,4.18)	1.49	(1.03,2.04)
10	Decrease Future K		+5%	2.77	(1.45,4.57)	2.85	(1.58,4.80)	1.48	(1.07,2.06)	1.78	(0.93,2.94)	1.83	(1.02,3.09)	0.95	(0.69,1.32)

**Table 1b:** Continuation of the performance statistics from Table 1a, listing the catch related performance statistics. The statistics reported here are  $C_{av}$  (the average catch over the projection period and over the next four years) both in median and mean terms, and AAV (the average inter-annual proportional change in catch over the projection period and over the next four years), in median terms.

	Species combined											
	Cav (25 yrs)				Cav (4 yrs)				AAV (25 years)		AAV (4 years)	
	Median	(90% P.I)	Mean	(sd)	Median	(90% P.I)	Mean	(sd)				
Equal weighting across the nine RS OMs												
Proposed OMP2018	154.08	(140.59,158.87)	152.33	(6.16)	148.13	(140.97,153.22)	147.50	(5.48)	0.025	(0.008,0.048)	0.048	(0.038,0.063)
Robustness test results for RS02 only												
0 RC	152.33	(144.23,157.93)	151.65	4.85	153.22	(141.21,153.22)	150.22	4.30	0.029	(0.013,0.044)	0.048	(0.039,0.060)
1 No surveys*	150.32	(140.27,156.26)	149.11	5.55	150.19	(141.47,153.48)	148.54	4.85	0.035	(0.019,0.047)	0.048	(0.038,0.055)
2 All industry*	152.25	(142.32,158.18)	151.17	5.06	153.04	(141.47,153.48)	150.00	4.53	0.031	(0.016,0.045)	0.048	(0.040,0.062)
3 Every 2nd year*	152.34	(143.35,157.50)	151.31	5.38	153.08	(141.47,153.48)	150.50	4.15	0.033	(0.017,0.045)	0.048	(0.040,0.060)
4 RC + UCI	157.35	(149.92,159.13)	156.02	3.60	153.48	(141.47,153.48)	151.05	3.92	0.019	(0.008,0.034)	0.048	(0.036,0.055)
5 NS+ UCI	157.57	(151.55,159.13)	156.37	3.57	152.33	(141.47,153.48)	149.98	4.41	0.016	(0.008,0.032)	0.048	(0.038,0.055)
6 Original mortality	152.89	(145.86,158.94)	152.48	5.04	153.48	(141.47,153.48)	151.05	3.92	0.027	(0.009,0.042)	0.048	(0.037,0.056)
7 Alg13 sp. Splitting	156.01	(147.85,159.13)	154.61	4.58	153.48	(141.47,153.48)	150.69	4.17	0.020	(0.008,0.039)	0.048	(0.039,0.058)
8 Model A6 sp. Splitting	151.57	(143.81,157.75)	151.11	4.71	153.48	(141.47,153.48)	150.88	4.06	0.032	(0.017,0.045)	0.048	(0.038,0.056)
9 Decrease past K	150.55	(142.51,157.38)	150.10	5.00	152.92	(141.47,153.48)	149.91	4.61	0.034	(0.019,0.046)	0.048	(0.040,0.060)
10 Decrease Future K	139.21	(128.56,151.24)	139.79	7.43	151.61	(141.47,153.48)	149.39	4.80	0.052	(0.035,0.076)	0.048	(0.039,0.060)



**Figure 1a:** Zeh plots of the performance statistics from Table 1, for the RC and first five robustness tests. These results are thus for the RS02 RC only and not an equal weighting across the RS. The statistics are  $B^{SP}/B_{MSY}$  for 2042 and 2022,  $B^{SP(low)}/B_{MSY}$  (the lowest value of this statistic in the projection period to 2042),  $TAC_{av}$  (the average catch over the projection period (25 years) and over the next four years) and  $AAV$  (the average inter-annual proportional change in catch over the projection period (25 years) and over the next four years). Medians and 90% probability intervals are shown. For  $TAC_{av}$ , the means are also shown by crosses. For each plot the 90% probability envelope for the RC is indicated by the grey shaded area for comparison purposes. Robustness tests that have been returned to match the RC risk level for *M. paradoxus* are marked by an asterisk.

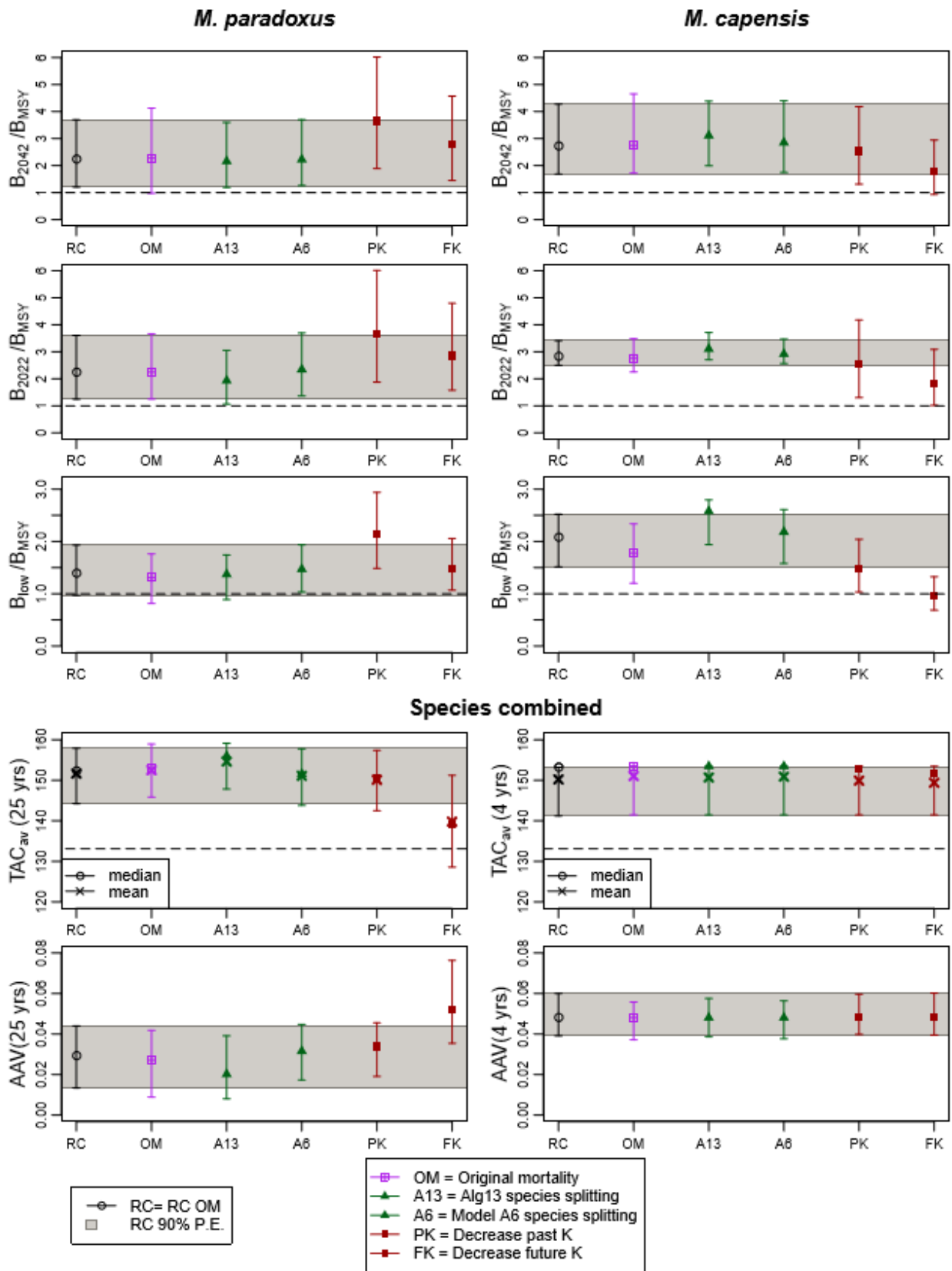
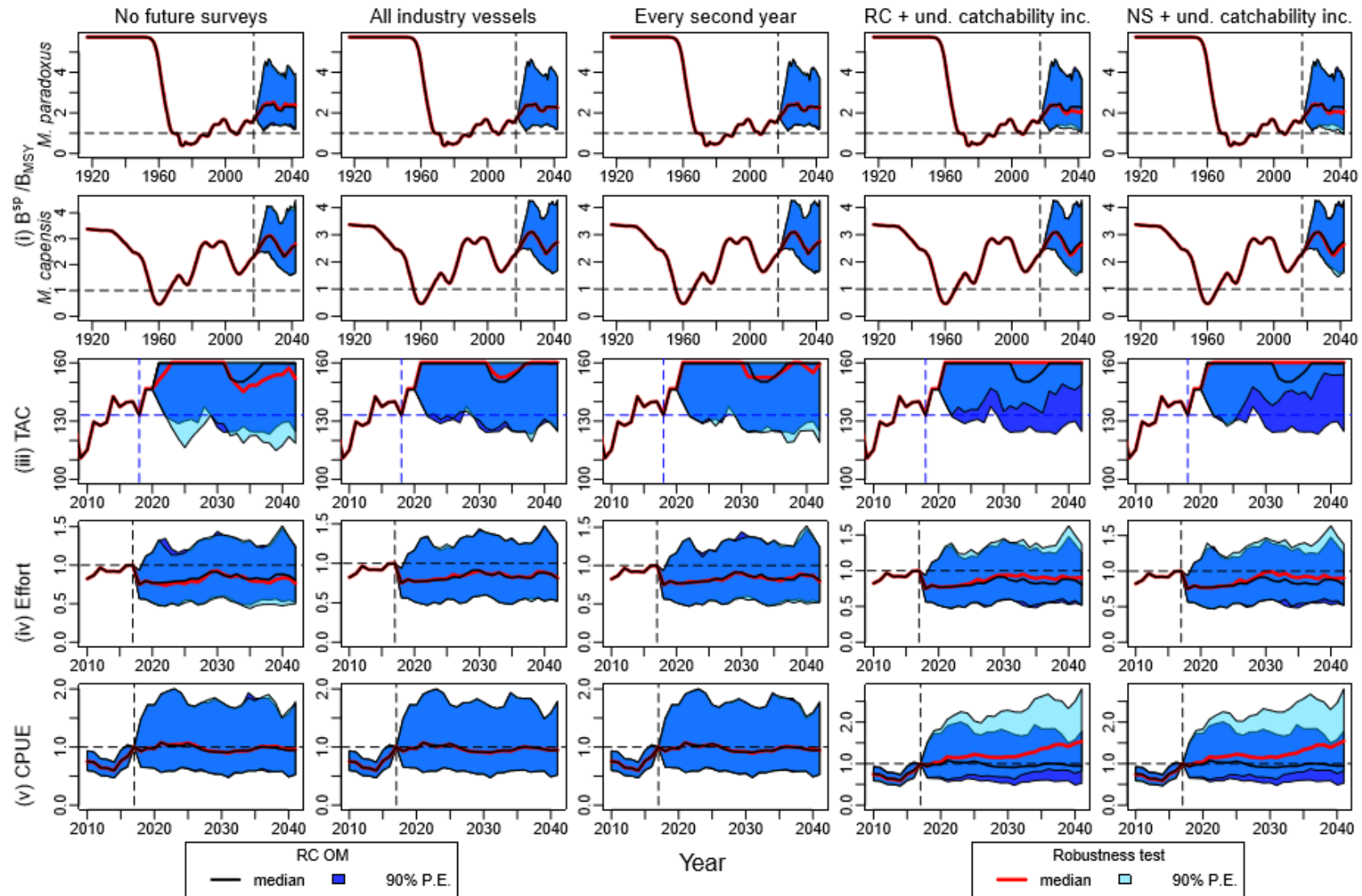


Figure 1b: Zeh plots of the performance statistics from Table 1, for the RC and second set of five robustness tests.



**Figure 2a:** Projected trajectories for  $B^{SP} / B_{MSY}$ , TAC, effort and CPUE are shown for the RC and first five robustness tests. The RS02 RC OM results are shown by the black lines (median trajectories) and dark blue shading (90% P.E.), while the robustness test results are indicated by the red lines and light blue shading. Areas of overlaps between the two P.E.'s are indicated by an intermediate blue.

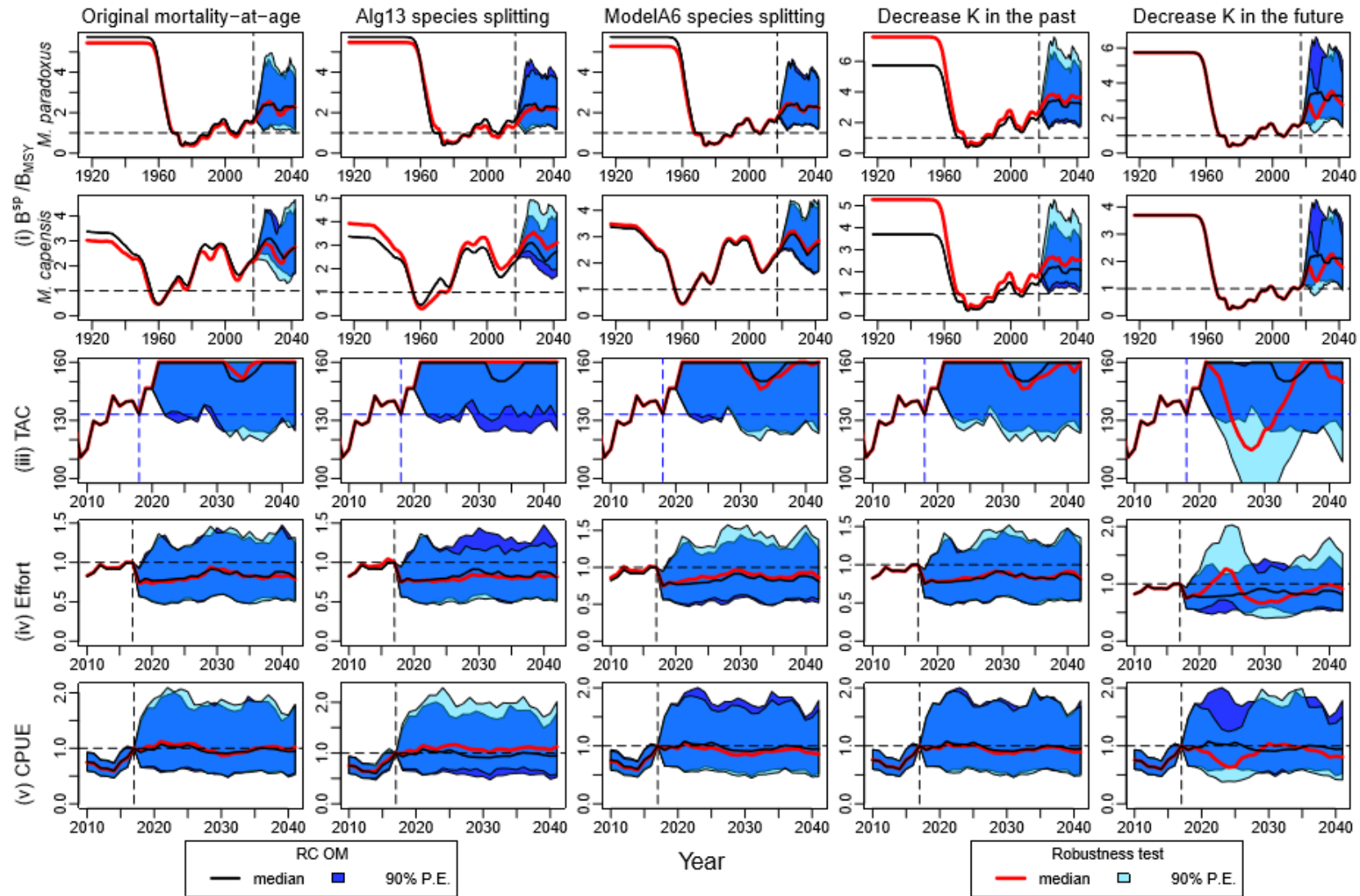


Figure 2b: Projected trajectories are shown for the RC and second set of five robustness tests.

## Appendix: Additional robustness tests concerning no future surveys

This Appendix reports results for three additional robustness tests requested by the Panel:

1. Decreasing future recruitment

Previously a robustness test was conducted where future carrying capacity was decreased by 30% over the course of the first five years projected, and then maintained at the reduced level for the remainder of the projection period. This reduction in  $K$  was intended as a proxy for recruitment failure. Here a robustness test is conducted where the recruitment values for 2018-2022 are halved, but return to normal levels for the remainder of the projection period.

2. Decreasing future recruitment plus no surveys

The robustness test above is combined with the no future surveys robustness test.

3. No future surveys plus undetected increase in catchability of 4%

This is a repeat of RT 5 of Hake/P6a, except that the commercial catchability coefficient is increased by 4% per annum instead of 2%.

Figure A1 plots the performance statistics and Figure A2 the trajectories for the three robustness tests.

Simulations suggest that while a recruitment failure over the next five years would likely result in an appreciably reduced TAC for almost 20 years, the OMP seems to perform adequately in preserving the hake resource, with the depletion levels after 25 years being very similar to the RC estimates.

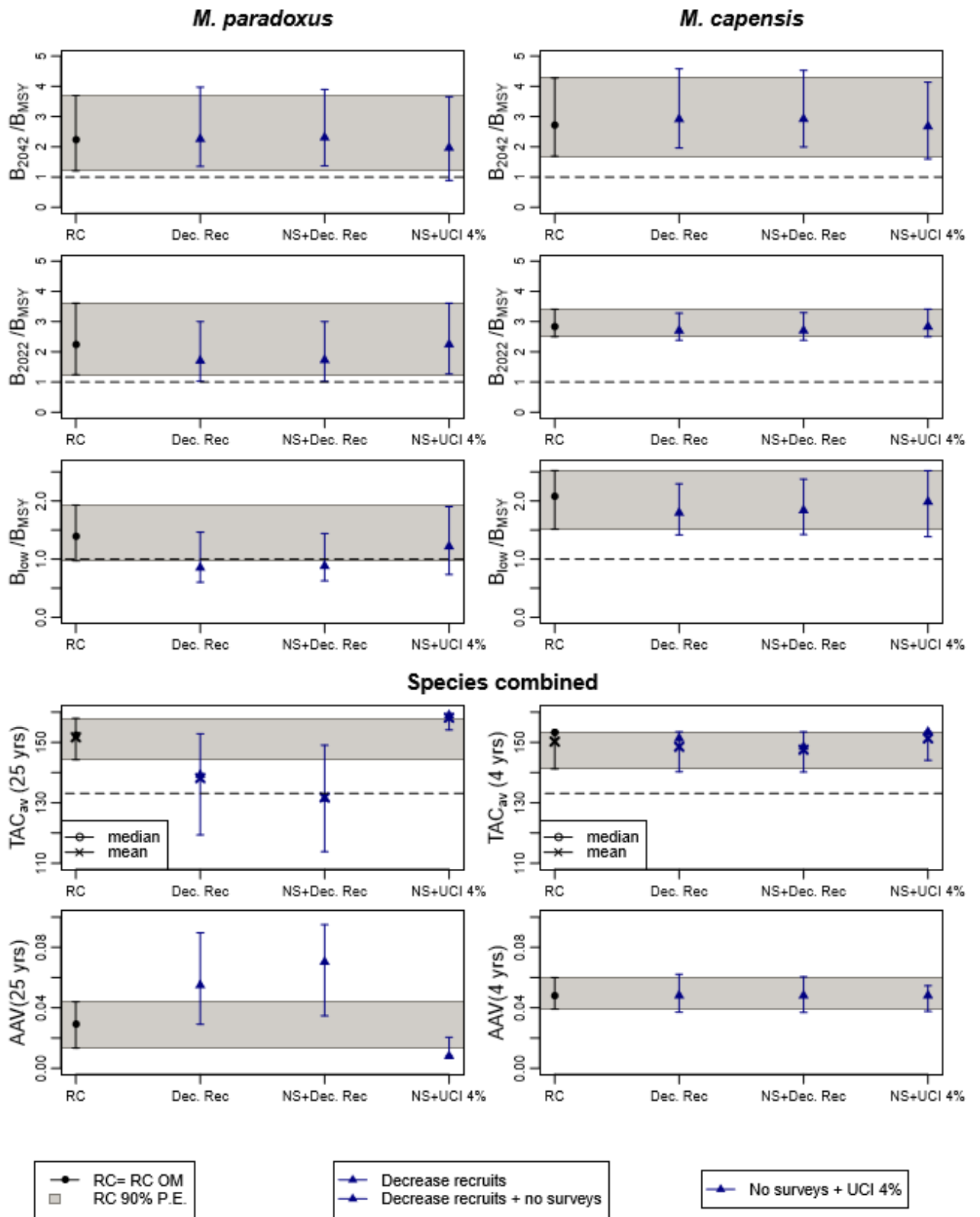
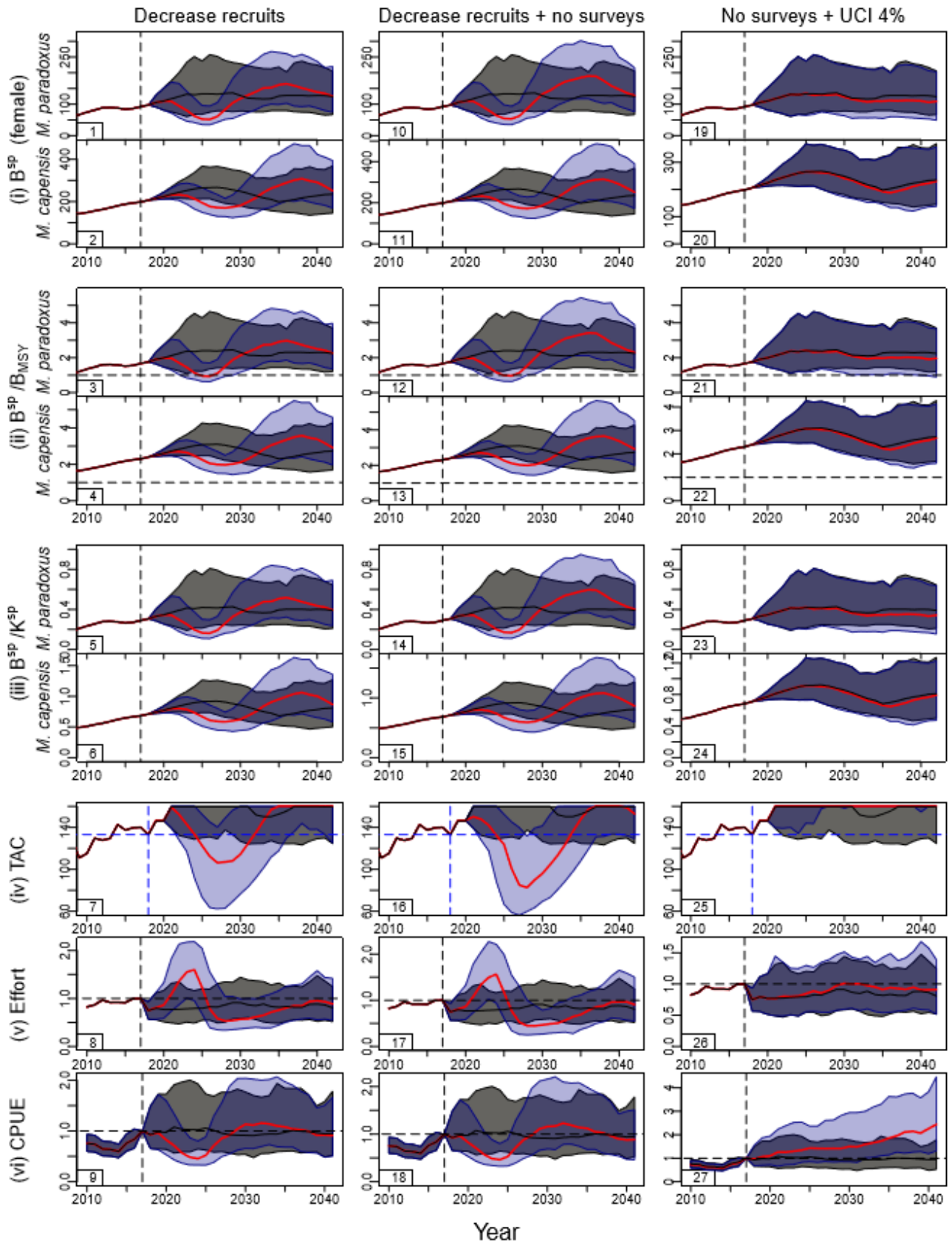


Figure A3: Zeh plots of the performance statistics for the RC and three additional robustness tests.



**Figure A4:** Trajectories for the RC OM (black lines for medians and grey shaded areas for 90% probability envelopes) are contrasted with the three additional robustness tests (red lines for medians and blue shaded area for 90% probability envelopes).