Results for a selection of robustness tests re-run for the updated 2019 Hake Reference Set

A. Ross-Gillespie and D.S. Butterworth[1]

Email: mlland028@myuct.ac.za

Summary

Results for a total of 14 robustness tests re-run for the updated 2019 Reference Case Operating Model (OM) are contrasted with results for the same robustness tests conducted for the 2018 Reference Case OM. Three of the robustness tests were also run for the 2019 Beverton-Holt central-year OMs (which estimate a heavily depleted *M. capensis* population) to check whether performance of OMP2018 is satisfactory even under those circumstances. Results for this selection of robustness tests indicate that the performance of OMP2018 remains adequate.

Introduction

This document provides results for a selection of the robustness tests that were originally conducted for OMP2018, but are re-run here for the updated 2019 Reference Set[2]. A total of 14 robustness tests have been re-run, where these tests were chosen based on three broad criteria:

(1) greatest management concerns (e.g. tests concerning uncertainty about future surveys),

(2) previous results indicated most concern for the conservation of the hake resource and/or stability of the TAC projections (e.g. tests decreasing the past or future carrying capacity and decreasing future recruits), and

(3) perceived biological importance (e.g. uncertainty about ageing information on the hake).

In most cases tests were conducted for the Reference Case (RC) Operating Model (OM) only[3], especially tests requiring a reconditioning of the OM or tuning of the *b* parameter[4], both of which would be very time-consuming exercises to conduct for the whole RS. A total of three robustness tests which indicated the most concern for the conservation of the populations (in particular for *M. capensis* given the depleted status estimated for this stock by the Beverton-Holt OMs) were conducted for the three central year OMs (Ricker 1958, Beverton-Holt 1958 for h=0.90 and 0.70) to check for adequate performance of OMP2018 even under the more pessimistic estimation of the *M. capensis* status under these revised Beverton-Holt OMs.

Table 1 provides a summary of the robustness tests which have been re-run for the updated RS. Note that in addition to these, RT13 (which restricts the survey *q's* to be less than one) has also been conducted for the updated RC and the two central-year Beverton-Holt OMs, and these results can be viewed in Appendix D of FISHERIES/2019/MAR/SWG-DEM/03.

---

[1] Marine Resource Assessment and Management Group, Department of Mathematics and Applied Mathematics, University of Cape Town, Rondebosch 7701.

[2] See MARAM/IWS/DEC2018/Hake/P6a and b for specifications of and results for all the robustness tests conducted for OMP2018 and FISHESRIES/2019/MAR/SWG-DEM/03 for details on the updated RS.

[3] The RC is the RS02 Ricker model with central year 1958, and is the RS OM with the best negative log-likelihood.

[4] For robustness tests simulating situations that are predictable (i.e. it is known beforehand that they will occur) and for which responsive management action could in theory be taken beforehand (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, which makes comparisons difficult. For "equivalent risk" results, the value for the b_s control parameters for each relevant robustness test have been modified so that the lower 5%-ile on the final (2042) *M. paradoxus* depletion is identical to that for the RC. The b_s parameters have been modified so as to maintain the original *M. paradoxus* to *M. capensis* ratio.

Results

Zeh plots of the performance statistics are shown in Figure 1a and b, while plots of trajectories of biomass, TAC, effort and CPUE are shown in Figure 2a and b. Both Figures 1 and 2 compare the results for the 2018 RC and robustness tests conducted for the 2018 RC with the updated 2019 RC and robustness tests conducted for the 2019 RC. Three of the robustness tests that showed greatest concern for the *M. capensis* resource and/or future TAC stability were also re-run for the two central-year Beverton-Holt OMs to check for adequate performance of OMP2018 even under the more pessimistic estimation of the current status of the *M. capensis* resource under the Beverton-Holt OMs. Trajectories for an equal weighting across the three central-year OMs are shown in Figure 3. Note that the plots in Figure 3 contrast results for the 2019 RC and its robustness tests, and do not include results for the 2018 RC. Table 2a and b list the negative log-likelihood components for the 2018 and 2019 RC OMs, and the robustness tests that required a reconditioning of the OM.

Discussion

On the whole, results for robustness tests conducted for the updated 2019 RC are very similar to those obtained in 2018 prior to this update. Differences in the range of the probability intervals in the Zeh plots are presumably most likely due to the small differences in the Reference Case results themselves, which are carried over into the robustness test results. The two robustness tests that test ageing assumptions do however provide certain somewhat different results to those previously (see Figure 2c). It is possible that some of the robustness tests conducted for the 2018 RC corresponded to a local minimum in the fitting process – the negative log-likelihood components are much better for the 2019 robustness tests and fairly comparable with the Reference Case components (see Table 2a and b) – and that this is the main cause of these differences.

There is no evidence from these results that the performance of OMP2018 is not adequate (or at least not equivalently adequate to as assessed in 2018). In all cases, by the end of the projection period, the population trajectories for both species show either upward trends or oscillations about well-recovered levels relative to B_{MSY}, even if the situations simulated result in initial declines (e.g. the robustness tests simulating recruitment failures in the immediate future, RT27, RT28). This indicates that OMP2018 should be able to cope in making adequate TAC adjustments should such situations occur.

Acknowledgements

Computations were performed using facilities provided by the University of Cape Town's ICTS High Performance Computing team: *hpc.uct.ac.za*.

Table 1: Summary of the robustness tests for which results are given in this document. More detailed descriptions of the robustness tests are provided in the Appendix. The column titled "OM re-run" indicates whether or not the robustness test in question required the OM to be re-conditioned. The column titled "Tuning required" indicates whether the robustness test simulates a situation that is likely to be known and for which the OMP *b* parameter would need to be to be (re-)tuned. The last column shows whether the robustness test in question has been conducted for the Reference Case (RC) OM only, or for all three central year OMs.

No.	Description	OM re-run?	Tuning required?	Results available for:
RT1	No future surveys **(NS)**	No	Yes	RC
RT2	All future surveys are conducted by industry vessels **(AI)**	No	Yes	RC
RT3	Two surveys (WC and SC) take place every second year **(ES)**	No	Yes	RC
RT4	RC assumptions (i.e. all future surveys continue as normal with the research vessel) but with an undetected increase in commercial catchability of 2% p.a. **(RC+UCI 2%)**	No	No	RC
RT5	No future surveys with an undetected increase in commercial catchability of a. 2% p.a. **(NS+UCI 2%)** b. 4% p.a. **(NS+UCI 4%)**	No	Use tuning values from RT1	RC
RT6	The original natural mortality-at-age vectors **(OM)**	Yes	No	RC
RT9	Decrease past carrying capacity **(PK)**	Yes	No	RC
RT10	Decrease future carrying capacity **(FK)**	No	No	RC, and B-H central-year OMs
RT12	Trend in F-proportions for *M. paradoxus* over time in the future - 2% p.a. decrease for 10 years, then constant **(DF)**	No	No	RC, and B-H central-year OMs
RT22	Biological ageing a. Ageing of both species is out by one year **(ALK-1)** b. Ageing of both species is halved **(ALK/2)**	Yes	No	RC
RT27	Decreasing future recruitment **(DR)**	No	No	RC, and B-H central-year OMs
RT28	Decreasing future recruitment and no future surveys **(DR+NS)**	No	Use tuning values from RT1	RC

Table 2a: Negative log-likelihood components for the **2018 RS02 RC OM** and the robustness tests that required a reconditioning of the OM. Table (i) lists the values in absolute term and Table (ii) relative to the RC values.

(i) Negative log−likelihood components for the 2018 RC and robustness tests requiring a reconditioning.

Run	GLM CPUE	ICSEAF CPUE	Survey abun.	Comm. CAL	Survey CAL	Recruit. resid.	ALKs	Penalties	Total (w/o pen.)
2018 RS02 RC	**−202.88**	**−37.72**	**−34.53**	**−1507.21**	**−1503.31**	**9.44**	**122.05**	**0.31**	**−3154.45**
Original mortality (OM)	−207.87	−39.69	−36.93	−1514.71	−1504.10	9.00	121.01	0.30	−3173.56
Decrease past K (PK)	−201.73	−38.26	−34.05	−1507.95	−1503.71	7.34	122.09	0.28	−3156.53
Ageing off by one (ALK−1)	−200.26	−39.55	−35.13	−1500.44	−1498.41	13.08	138.44	0.20	−3122.44
Ageing halved (ALK/2)	−204.75	−37.35	−37.94	−1498.98	−1489.85	13.15	224.00	0.91	−3032.07

(ii) Negative log−likelihood components relative to the first row (grey highlighting indicates a worse negative log−likelihood).

Run	GLM CPUE	ICSEAF CPUE	Survey abun.	Comm. CAL	Survey CAL	Recruit. resid.	ALKs	Penalties	Total (w/o pen.)
2018 RS02 RC	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**
Original mortality (OM)	−4.99	−1.97	−2.40	−7.51	−0.79	−0.44	−1.04	−0.01	−19.11
Decrease past K (PK)	1.15	−0.54	0.48	−0.74	−0.41	−2.10	0.05	−0.04	−2.07
Ageing off by one (ALK−1)	2.62	−1.82	−0.60	6.77	4.89	3.64	16.39	−0.11	32.01
Ageing halved (ALK/2)	−1.87	0.37	−3.42	8.22	13.45	3.71	101.95	0.60	122.38

Table 2b Negative log-likelihood components for the **2019 RS02 RC OM** and the robustness tests that required a reconditioning of the OM. Table (i) lists the values in absolute term and Table (ii) relative to the RC values.

(i) Negative log−likelihood components for the 2019 RC and robustness tests requiring a reconditioning.

Run	GLM CPUE	ICSEAF CPUE	Survey abun.	Comm. CAL	Survey CAL	Recruit. resid.	ALKs	Penalties	Total (w/o pen.)
2019 RS02 RC	**−203.88**	**−37.74**	**−34.69**	**−1508.45**	**−1502.68**	**9.37**	**123.55**	**0.06**	**−3154.52**
Original mortality (OM)	−208.48	−39.53	−36.71	−1516.85	−1501.95	8.58	121.47	0.02	−3173.46
Decrease past K (PK)	−203.04	−38.37	−34.26	−1509.18	−1503.13	7.28	123.47	0.05	−3157.21
Ageing off by one (ALK−1)	−201.59	−37.08	−36.31	−1507.29	−1509.40	10.87	127.12	0.11	−3153.67
Ageing halved (ALK/2)	−210.26	−40.01	−39.60	−1513.76	−1493.45	12.60	68.39	0.04	−3216.09

(ii) Negative log−likelihood components relative to the first row (grey highlighting indicates a worse negative log−likelihood).

Run	GLM CPUE	ICSEAF CPUE	Survey abun.	Comm. CAL	Survey CAL	Recruit. resid.	ALKs	Penalties	Total (w/o pen.)
2019 RS02 RC	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**
Original mortality (OM)	−4.60	−1.79	−2.02	−8.40	0.73	−0.78	−2.08	−0.04	−18.94
Decrease past K (PK)	0.84	−0.63	0.43	−0.72	−0.45	−2.09	−0.07	−0.01	−2.69
Ageing off by one (ALK−1)	2.29	0.65	−1.62	1.17	−6.72	1.50	3.57	0.05	0.85
Ageing halved (ALK/2)	−6.38	−2.28	−4.91	−5.30	9.24	3.23	−55.16	−0.02	−61.57

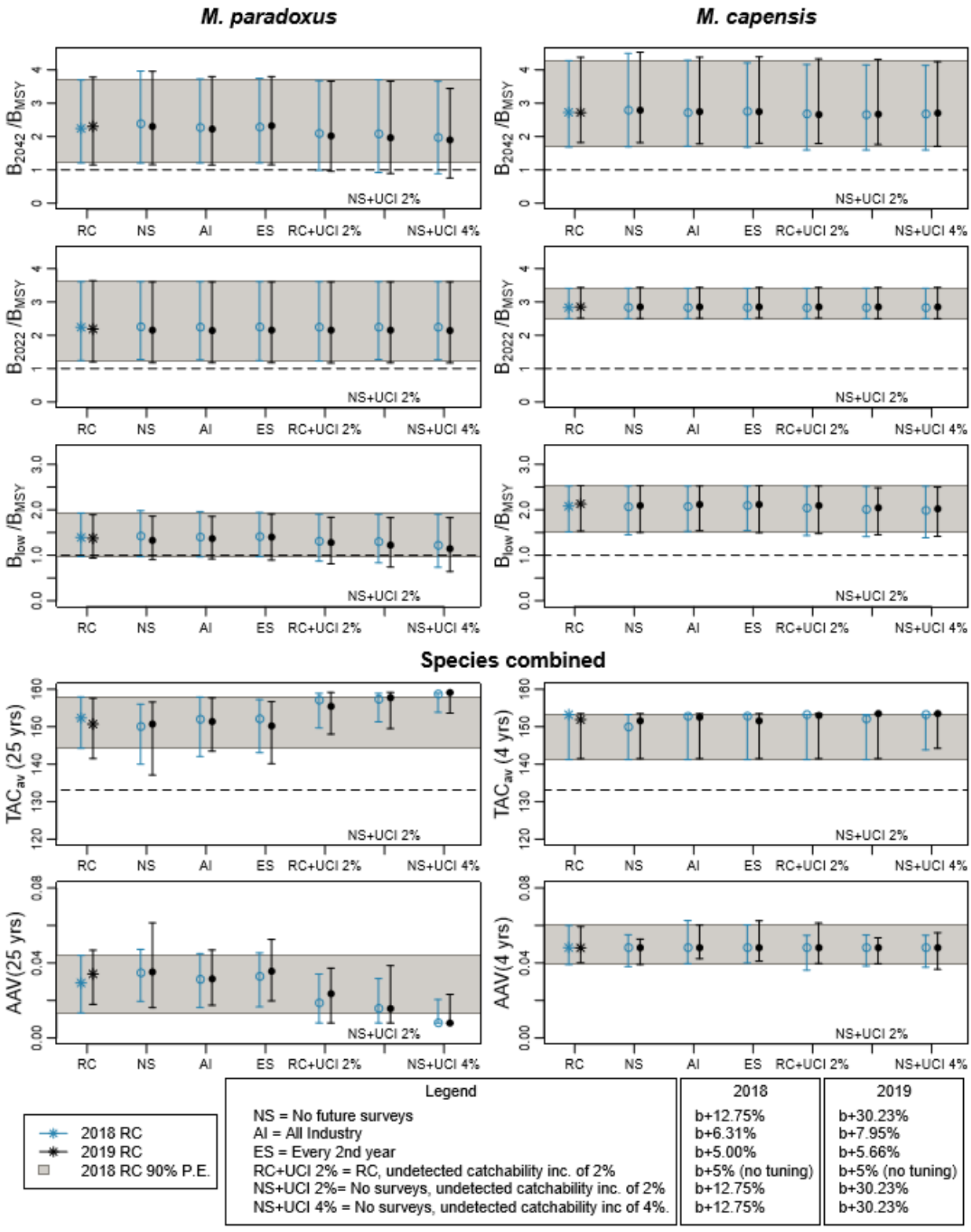

Figure 1a: Zeh plots of the key performance statistics**, for the RC and first six robustness tests**. In each case, results are shown for the 2018 (blue) and 2019 (black) RC and robustness tests conducted for each. These results are for the RS02 RC only. The statistics are B^{sp}/B_{MSY} for 2042 and 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 each plot, the 90% probability interval for the 2018 RC is indicated by the grey shaded area for comparison purposes. All of the robustness tests except RC+UCI2% have been retuned to match the RC risk level for *M. paradoxus* (see footnote 4 on page 1), and the retuned *b* parameters for the 2018 and 2019 OMs are indicated in the legend.

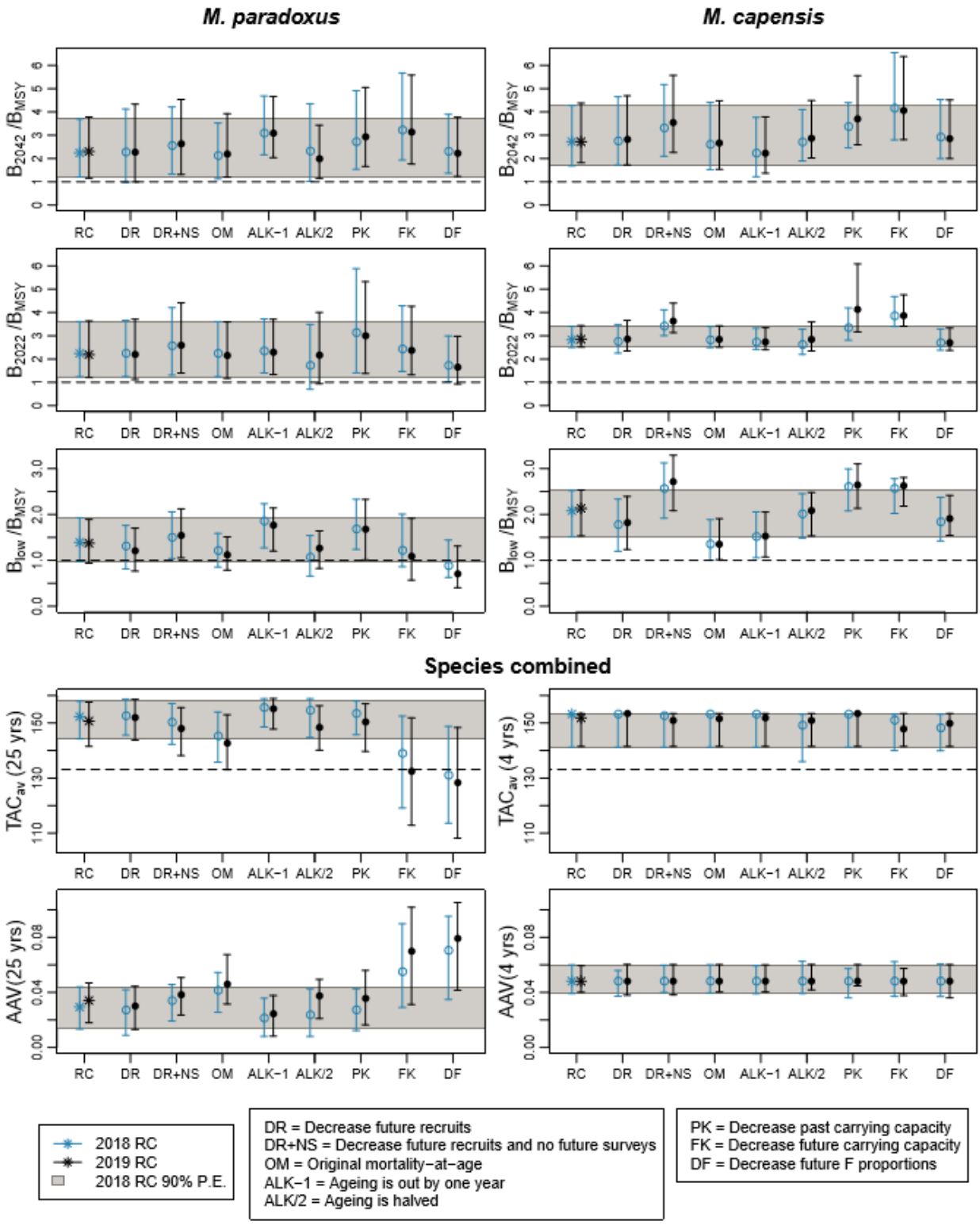

Figure 1b: Zeh plots of the performance statistics from **Error! Reference source not found., for the RC and remaining eight robustness tests**. Once again, these robustness test results are for the RC only, and results are shown comparing the 2018 and 2019 RC OMs and robustness tests conducted for each. No retuning (see footnote 4 on page 1) was required for any of the robustness tests included in this Figure, although the RT1 *b* tuning parameter values were used for the DR+NS robustness test.

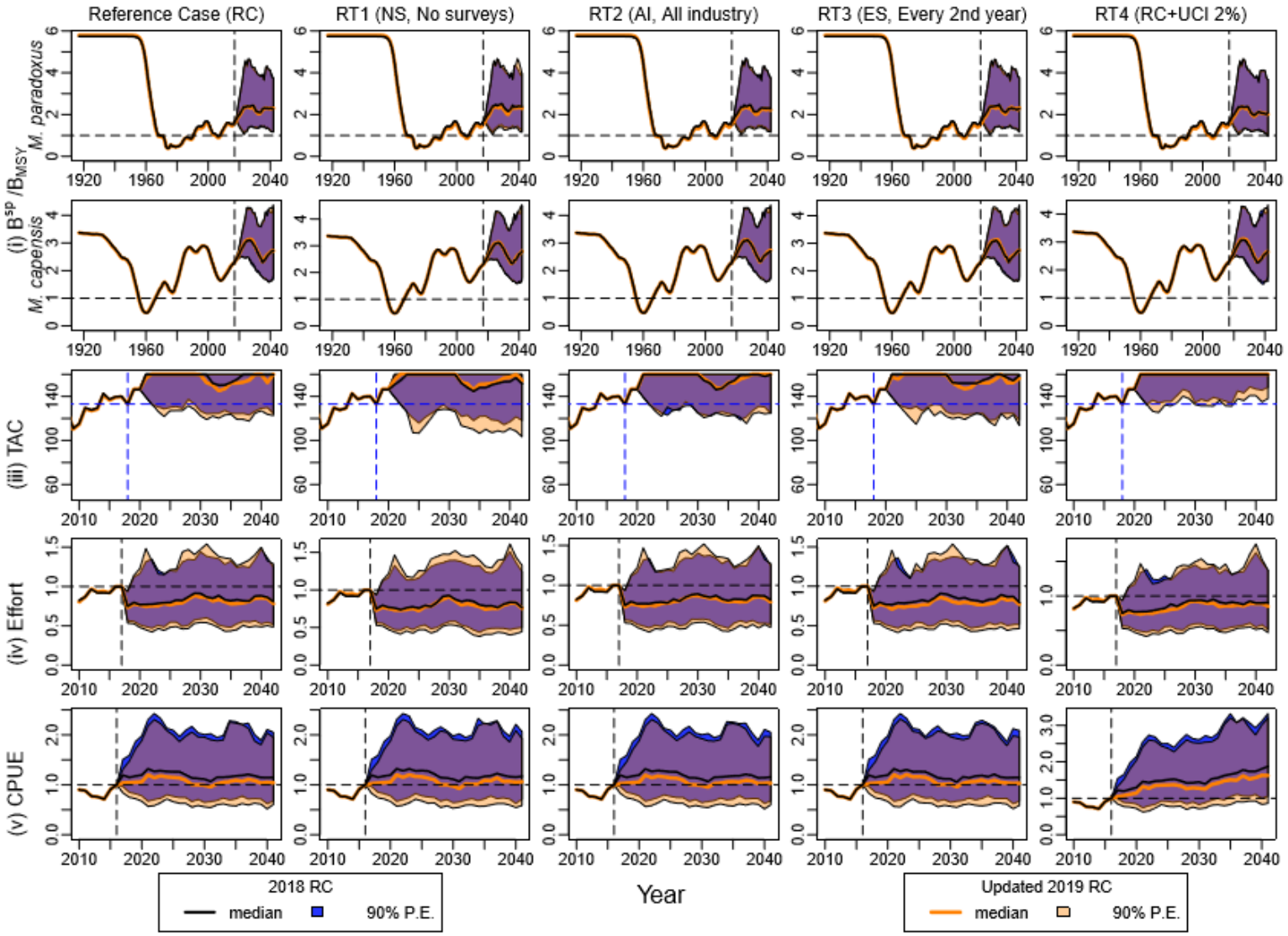

Figure 2a: Projected trajectories for B^{sp}/B_{MSY}, TAC, effort and CPUE are shown for the **RC and first four robustness tests**. In each plot, trajectories for the 2018 RC or robustness test conducted on the 2018 RC (black median lines with blue 90% probability envelopes) are contrasted with the 2019 updated RC or robustness test conducted on the 2019 RC (orange median lines with yellow 90% P.E.s). Areas of overlap between the two P.E.'s are indicated by an intermediate purple.

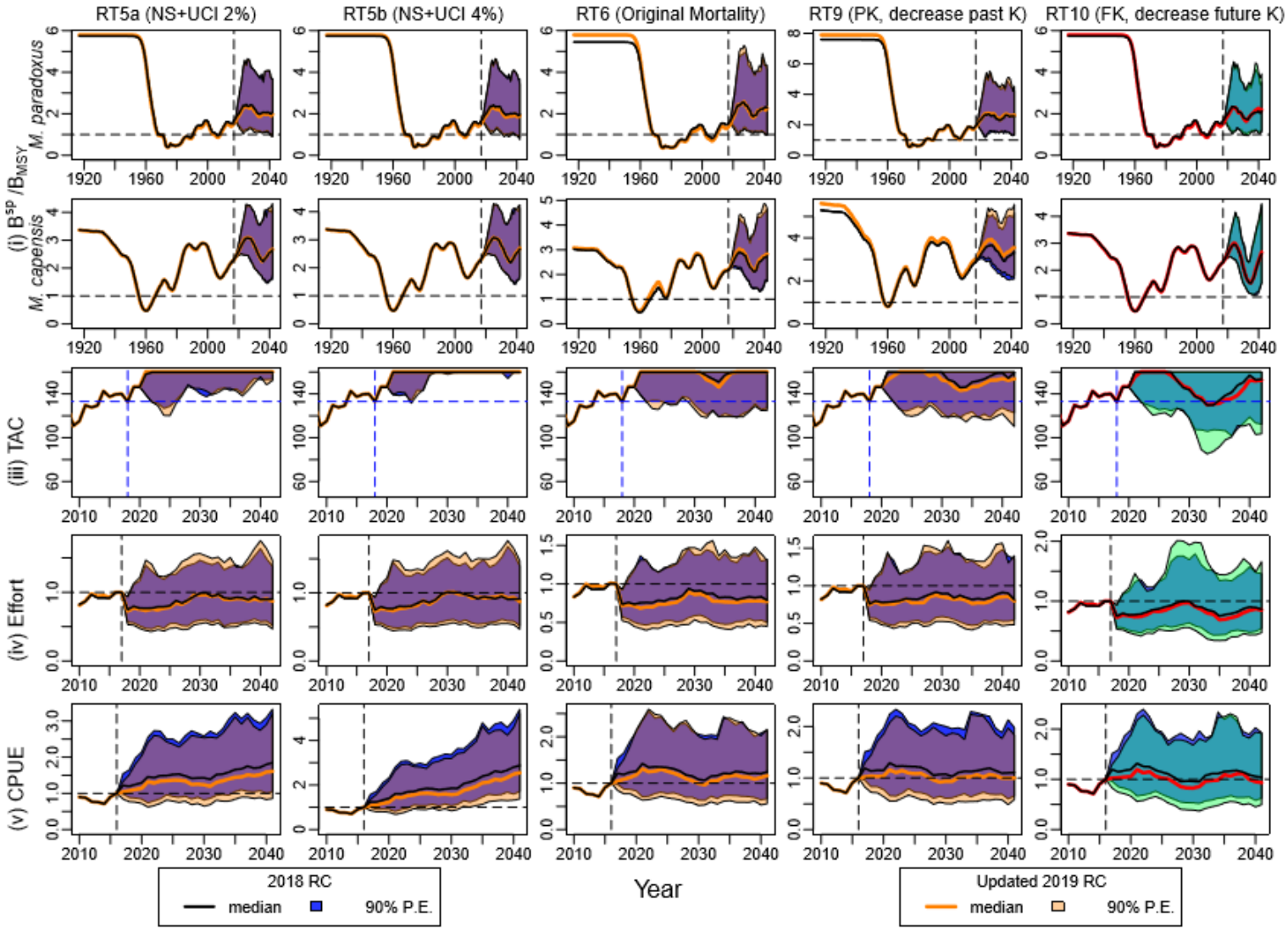

Figure 2b: Projected trajectories are shown for the **next five robustness tests**. Columns with blue and green shading indicate that the robustness test has been repeated for the Beverton-Holt central-year OMs, and results for an equal weighting across the three central-year OMs can be seen in Figure 3. Results shown in this Figure, however, are for the RC OM only.

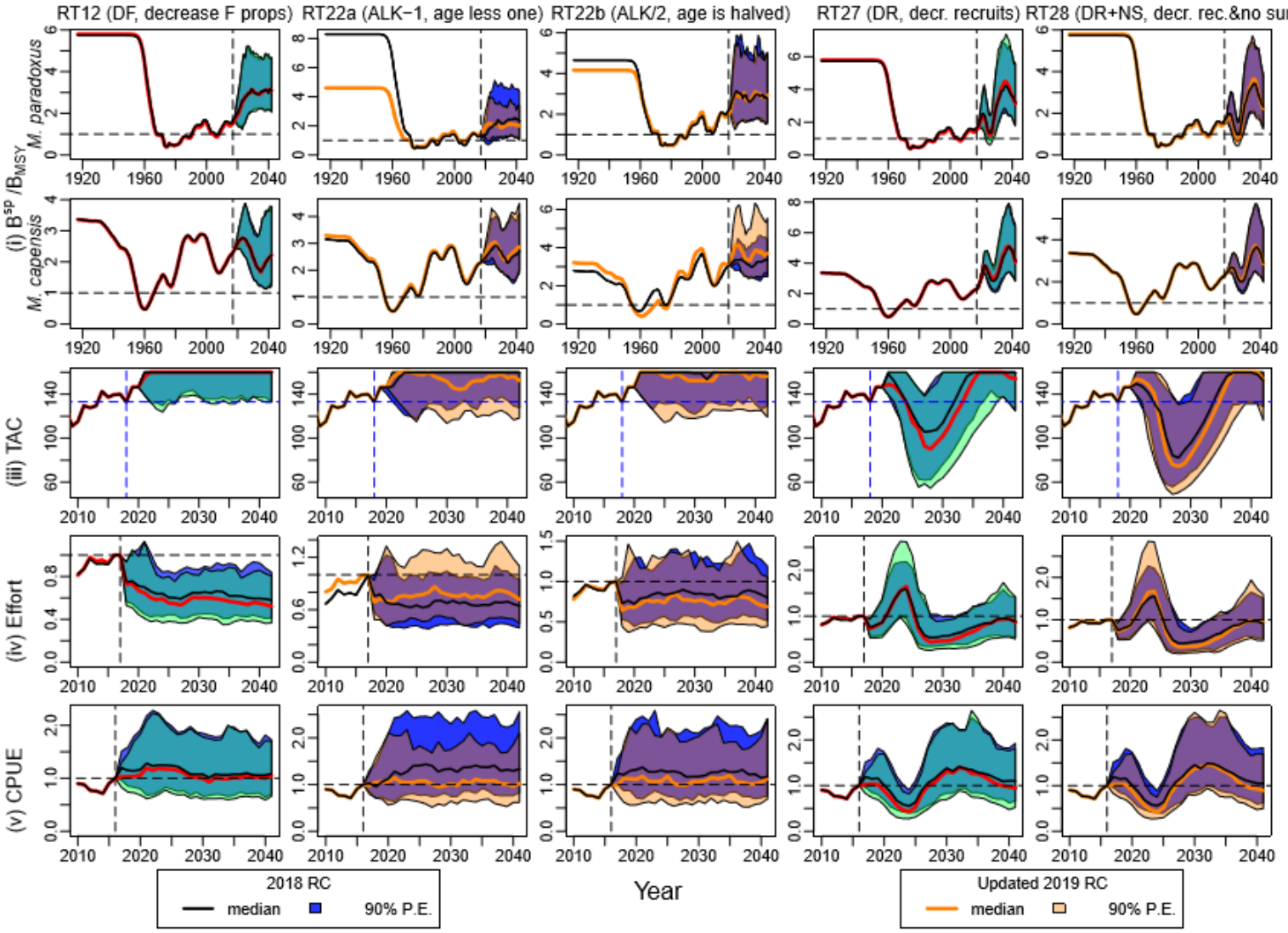

Figure 2c: Projected trajectories are shown for the **last five robustness tests**. Columns with blue and green shading indicate that the robustness test has been repeated for the Beverton-Holt central-year OMs, and results for an equal weighting across the three central-year OMs can be seen in Figure 3. Results shown in this Figure, however, are for the RC OM only.

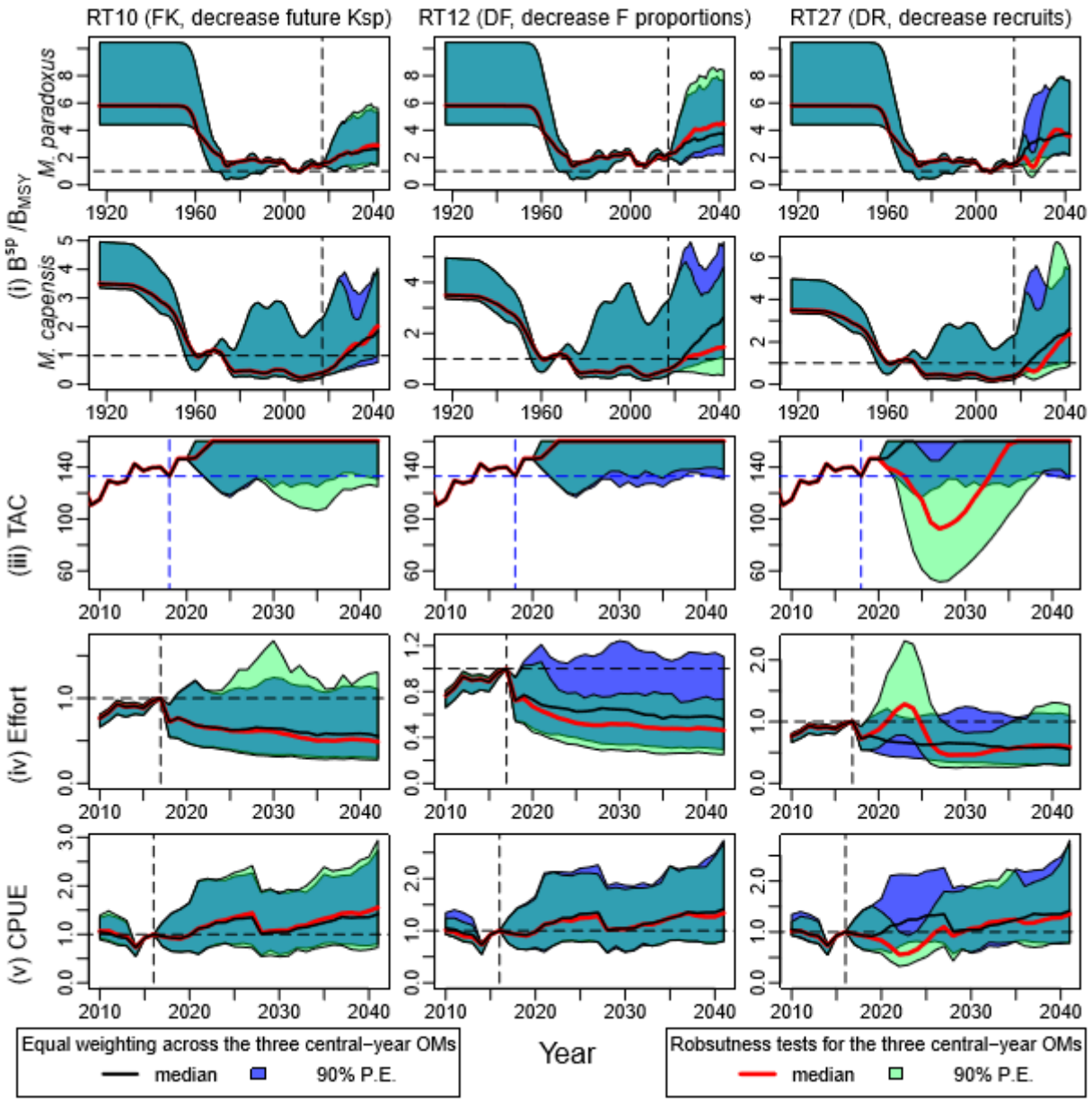

Figure 3: Projected trajectories are shown for an **equal weighting across the three central-year OMs** for a selection of three robustness tests that indicated greatest concern with respect to conservation of the hake resource and/or TAC stability. Median and 90% P.E. trajectories for an equal weighting across the three central-year OMs are shown by the black lines and blue shaded area, while an equal weighting for the robustness tests results for the three OMs are shown by the red lines and green shaded areas. **Note that this figure compares the 2019 RC with robustness tests conducted on the 2019 RC, i.e. no results for the 2018 RC have been included here.**

Appendix: Detailed descriptions of the robustness tests re-run for the updated RS

The descriptions of the robustness tests provided in this Appendix are based on those from MARAM/IWS/DEC2018/Hake/P6a and b, repeated here for ease of reference.

*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 for details of the OMP formula).

*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^{ϵ_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. $R=q_I/q_R$. 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). [5]

When generating future survey indices under the assumption that an industry vessel will conduct the surveys, a further error term e^{η_y} is added to equation (1) with $\eta_y \sim N(ln0.80, 0.22^2)$ so that e^{η_y} is log-normally distributed with median of 0.80 and standard error of 0.18.[6] 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:

$$I_y^{surv} = q^{surv} B_y^{surv} e^{\epsilon_y} e^{\eta_y} / 0.80 \quad (2)$$

*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. Two values are tested: a 2% increase p.a. and a 4% increase per annum, with the latter being an additional robustness tests requested by the Panel for the 2018 International Stock Assessment Review Workshop.

*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.

[5] 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.

[6] If ln*R* is normally distributed with mean μ and variance σ^2, then the median of the distribution of *R* is given by $\theta_R = e^{\mu}$ and the variance by $\sigma_R = \left(e^{\sigma^2} - 1\right)\left(e^{2\mu+\sigma^2}\right)$. Since $\theta_R = 0.80$ and $\sigma_R = 0.18$, it follows that $\mu = ln0.80$ and $\sigma = 0.22$.

*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[7].

*RT12: Trend in F-proportions over time in the future **(DF)***

The offshore trawl fishery F-proportions (i.e. the proportion of the total fishing mortality in the offshore catch on *M. paradoxus*), generated according to Appendix A and B of MARAM/IWS/2018/Hake/P4) are decreased by 2% p.a. (where this 2% is in absolute terms, i.e. if a proportion was 0.94 in 2018, it will be 0.92 in 2019) for a period of 10 years, after which the proportions are constant. Thus, effectively only the proportions generated for 2018 are used, as the subsequent proportions are derived from these initial values.

*RT22: Biological ageing **(ALK-1** and **ALK/2)***

The RC OM is fit to age-length keys, which consist of pairs of age-length readings, allowing parameters for the von Bertalanffy growth curve to be estimated. Therefore, to test the assumptions that the age readings are not correct and that either (a) ageing of both species is out by one year, or (b) the ageing of both species is halved, the age column of the age-length keys simply needs to be decreased by one or divided by two. Adjusting the age-length-key error matrices is a little more complicated. These are the matrices consisting of error distributions for each age, for each species and from five different readers. For robustness tests 22(a), the columns and rows of the matrices are shifted by one, leaving the column and row corresponding to the plus group 15+ blank. These values have been interpolated by borrowing from data in adjoining cells. Each row is then normalised so that it once again sums to one. For test 22(b) where the age is halved, the data corresponding to ages 0 and 1 are combined to form the row for the new age 1, ages 2 and 3 combined for the new age 2, and so on. This results in an error matrix up to age 7 only, and the code has correspondingly been adjusted to take these errors into account up to age 7 only.

Note that strictly speaking the predation model (which fixes the von Bertalanffy parameters at the estimates from 2017 RC model) should have been re-conditioned with the von Bertalanffy estimates from these robustness tests to obtain updated mortality-at-age vectors that reflect the ageing assumptions made in the robustness tests. There was however insufficient time to attempt this.

*RT27: Decreasing future recruitment **(DR)***

Robustness Test RT10 (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) was intended as a proxy for recruitment failure. The Panel for the 2018 International Stock Assessment Workshop requested a robustness test where the recruitment values for 2018-2022 are halved, but return to normal levels for the remainder of the projection period.

*RT28: Decreasing future recruitment plus no surveys **(DR+NS)***

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

[7] 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.