Second set of robustness tests conducted for OMP2018

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

Email: mlland028@myuct.ac.za

Summary

Results are provided for a further 24 robustness tests for OMP2018. All tests have been conducted for the RC OM, and most do not have an appreciable impact on the assessment and projection results. The tests with the biggest impact were increasing or decreasing the proportion of the total fishing mortality in the offshore catch on *M. paradoxus*, adjusting the weights of the ALK and CAL data in the negative log-likelihood, accounting for discards and adjusting the ageing assumptions. None of these robustness tests, however, suggest serious concerns for the conservation of the hake resource when managed under OMP-2018.

Introduction

All robustness tests have been conducted for the RC OM (RS02 with a Ricker stock-recruitment and a central shift in catches occurring in 1958). Table 1 provides a summary of the robustness tests, with more detailed descriptions provided in the main text below. These robustness tests originate from a variety of sources, including robustness tests conducted for previous OMPs and recommendations made at past International Stock Assessment Workshops. MARAM/IWS/2018/Hake/P5 provides more information on the origins of these tests.

Detailed descriptions of the robustness tests

The robustness tests reported on in this document are described below.

RT11: Decrease future survey CV's by a factor of 1/sqrt(2)

This robustness tests aims to simulate a situation where the sampling methods are improved in the future (e.g. through more intensive surveying), resulting in a smaller CV. Note that strictly speaking this robustness test should be retuned to match the risk of the RC OM, given that it simulates a "known" situation. However, as the results are so similar to those for the RC OM this was not deemed to be necessary.

RT12: Trend in F-proportions over time in the future

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 (a) increased and (b) 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 (a) 0.96 or (b) 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. Where proportions exceed one as a result of these changes, they are fixed at 0.9999 (as Fprop=1 results in divide by zero errors).

RT13: Restrict survey q's to be less than one

At IWS 2017, the panel noted that it was unexpected that several of the estimated q values should exceed one. As such, they recommended that the OM should be fitted restricting the survey q to be less than one and to examine which data are possibly driving these estimates to exceed 1.

RT14: Adjust survey estimates from industry vessels for month effect

As the surveys conducted by industry vessels were typically conducted a little later in the year, it has been suggested to adjust the abundance estimates from these surveys to take this into account. This was done using estimates for month effects from the GLM fits to commercial CPUE, which are listed in Table 2.

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

Surveys conducted on research vessels have taken place (on average) from 10 January – 16 February, and those conducted by industry vessels on average from 12 February – 17 March. While these dates do not correspond exactly to the full calendar months in Table 2, for this robustness test it was assumed that the month January corresponds to research vessel surveys and the month February to industry vessel surveys. These assumptions could clearly be refined, but given the relatively small differences in month effects, such further refinement is unlikely to have a substantial impact. Therefore, in order to take month effect into account, abundance estimates from surveys conducted by industry vessels were divided by a factor of 1.0299 (=1.0340/1.0039) for *M. capensis* and a factor of 1.0136 (=1.0089/0.9954) for *M. paradoxus.*

RT15: Introduce a 150 000t cap from 1995

During discussions on whether to increase the TAC cap imposed by the OMP (which was set at 150 000t for OMP2014 and has been increased to 160 000t for OMP2018), a concern was expressed that the last time the catches went above 150 000t (in the late 1990s and early 2000s), this was followed by a considerable decline in hake abundance. In order to try ascertain whether this decline in resource was more likely due to these high catches or to some other factor such as a recruitment failure, the OM was re-run introducing a 150 000t cap from 1995 (i.e. this is not a re-conditioning as the RC MLE estimates where retained, but with the catch reduced in years where it exceeded 150 000t. If the total catch in year *i* was *X*>150, then the catch for each fleet was reduced by 150/X). The 150 000t mark was exceeded six times in the period 1995-2004, and when the catch is capped at 150 000t during this period, this results in a reduction of 37 115t in removals compared to what actually occurred.

RT16: Alternative survey calibration factors to cater for Africana gear change

The calibration factors assumed in the RC OM are $\left(q^{new}/q^{old}\right)^{capensis} = 0.652$ for *M. capensis* and $\left(q^{new}/q^{old}\right)^{paradoxus} = 0.883$ for *M. paradoxus*. Values of (a) 0.90 and (b) 0.60 are tested for *M. capensis.*

RT17: q for CPUE drops by 20% from 1997 to 2002 as a result of shorter tows

The length of the tows was reduced in this period for quality reasons, such as reducing the proportion of crushed fish. This could conceivably reduce the catchability coefficient, as for long tows the mesh in the cod end tends to get blocked so that smaller fish cannot pass through that mesh (resulting in proportionately higher landed catches). Reducing the tow length means a proportionately longer period when the smaller fish escape and thus a lower catchability coefficient and reduced CPUE.

RT18: Increase weighting of recent CPUE and survey data

In order to test whether historical data have an undue influence on particularly the recent assessment results, the CPUE and survey data from 2011 onwards are upweighted by a factor of 5 (i.e. their log-likelihood contributions are multiplied by 5) so that recent results correspond more closely to recent data trends).

RT19: Down-weight data for industry vessel surveys by a factor of 10

Given uncertainty about calibration factors for the industry vessels, data arising from these surveys are down-weighted by a factor of 10, i.e. their log-likelihood contributions are multiplied by 0.1.

RT 20: Alternative log-likelihood weightings

The RC OM assumes a weighting for the age-length key data of 0.01 in the negative log-likelihood and the catch-at-length data are weighted by a factor of 0.10. These robustness tests decrease and increase these weights by a factor of 10.

RT21: Discards

As the RC OM ignores discards, this robustness test aims to take discards into account. With a similar approach to such robustness tests in previous years, discarding by offshore and inshore trawlers is modelled as an increase in the commercial selectivity of 0.2 for ages 1 and 2 for *M. capensis* and *M. paradoxus*; the loss of fish from longlines is also included by doubling fishing mortality from this fleet for ages 1 and 2 for both species. These increases in selectivity and *F* are taken into account immediately **before** the catches-at-age are computed in the model and deducted from the population; thus the discards effect only the estimates of total population size and not any of the fits to the data.

RT22: Biological ageing

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.

RT23: Gender-specific natural mortality

The natural mortality-at-age vectors for males are increased by 0.05 throughout, and those for females decreased by 0.05.

RT24: Increase M at larger ages

The RC OM uses the mortality-at-age vectors from the hake predation model (MARAM/IWS/2018/Hake/BG7), which assumes a basal mortality rate (the only contribution to the natural mortality for larger hake that are not subject to predation by other hake) of 0.20. In this robustness test, the natural mortality is increased to 0.3 at larger ages (from age 7 for *M. capensis* and age 10 for *M. paradoxus*), in a manner so as to keep the mortality-at-age vectors monotonic. These vectors are illustrated in Figure 1.

RT25: Asymptotically flat offshore and inshore trawl selectivities for M. capensis

The panel for IWS 2017 recommended that a robustness test be conducted where at least one fleet or survey having a near asymptotic selectivity pattern for *M. capensis*, to examine what is otherwise driving the dome-shaped selectivity estimated. The OM was run forcing the offshore and inshore trawl selectivities for *M. capensis* to be asymptotically flat.

RT26: M. capensis offshore trawl fleet selectivity is assumed to be equal to that for the inshore trawl fleet

The *M. capensis* offshore trawl fleet selectivity is assumed to be identical to the inshore trawl fleet selectivity (note that the BC assumes a fixed shift in selectivity towards older ages, as might be expected, but with a somewhat arbitrary choice for the associated magnitude). Note this OM tended to estimate an unrealistically high *M. capensis* pristine spawning biomass (over 2 000 000t). As there was not sufficient time for extensive troubleshooting, the OM was re-fit simply by fixing the Ricker *h* parameters at the RC estimates, which served to keep the K^{sp} estimate in check.

Robustness tests still outstanding

The list below provides information on robustness tests that have not yet been conducted to date. The label "IWS 2017" indicates a recommendation made by the panel for the 2017 International Stock Assessment Workshop while "OMP-2014" indicates that the robustness test in question was considered during the development of OMP-2014.

- Examine the consequences of fishery selectivity changing in the future (IWS 2017).
- Impose a lower bound on the von Bertalanffy κ value (IWS 2017).
- Include OMs that are fitted to the geostatistical estimates of biomass (to 500m) in the robustness tests (IWS 2017).
- Develop a robustness test that includes the data from the region 500-750m (IWS 2017)
- Alternative maturity-at-length vectors (OMP-2014)
- Maximum proportion of a cohort catchable in one year decreases from 90% to 70% (OMP-2014)

- No shrinkage of recent recruitment (OMP-2014)
- 40:60 male:female birth ratio instead of 50:50 (OMP-2014)
- Past stock-recruitment σ_R=0.25 (OMP-2014)
- MPA: possible effects on future CPUE (OMP-2014)
- Change efficiency in offshore trawl fleet in 1994/1995 (OMP-2014)

Additionally, the OM starting in 1978, which the panel for IWS 2017 recommended be made part of the RS, is not yet fitting at the same level of reliability as the other RS OMs and robustness tests, and consequently has not been included as a robustness test itself at this stage. However, the results that have been achieved to date have been included in the Appendix.

Results

The negative log-likelihood components for the RC OM and for the robustness tests which had to be reconditioned are listed in absolute terms Table 3a and relative to the RC OM in Table 3b. The primary performance statistics are illustrated in Figure 2a and b, while the corresponding plots of the median "trajectories" and probability envelopes for B^{sp}, B/B_{MSY}, B/K^{sp}, catch, effort and CPUE are shown in Figure 3a-h. Note that the colours chosen for the different robustness tests have no significance other than to help link the performance statistics results in Figure 2a and b with the trajectories in in Figure 3a-h.

Discussion

Of the 24 robustness tests explored here, many did not give results which differed substantially from those for the RC. Those for which there was a notable difference are discussed briefly below. None of the robustness tests explored here, however, suggest serious concerns for the conservation of the hake resource and the performance of OMP2018.

- Increasing the *F*-proportions (i.e. increasing the proportions of the total fishing mortality in the offshore catch relating to *M. paradoxus*) has the predictable effect of reducing the *M. paradoxus* and increasing the *M. capensis* resource abundance. The converse is true when the *F*-proportions are decreased. In both cases the median trajectories for B^{sp} remain well above B_{MSY}, although for the former (12a), the lower bounds of the 90% P.E. do go below B_{MSY}.
- Restricting the survey *q*s to be less than one has the effect of reducing the estimates of both species relative to B_{MSY}, although the *M. paradoxus* estimate of B^{sp} relative to K^{sp} is more optimistic. The negative log-likelihood for this OM is worse than the RC by 6.66 points, but these are fairly evenly spread between worse fits to the CPUE data, survey abundances and commercial catches-at-length, which are countered by a better fit to the survey catches-at-length. Thus, there is no one data source that stands out as a driving force behind the q>1 estimates.
- Adjusting the weight of the ALK and CAL data does have some impact on the appearance of the spawning biomass trajectories, although in median terms these remain well above B_{MSY}. Adjusting the weight of the CAL data seems to have the biggest impact, with trajectories for depletion for both species differing substantially as these data are up- or down-weighted.
- Accounting for past discards has the effect of increasing the estimate of pristine biomass and indicating a more heavily depleted current *M. paradoxus* resource. If in future selectivity is changed to prevent any need for discards, then the *M. paradoxus* resource is predicted to increase rapidly. If discards continue in future, the population relative to K^{sp} stays more depleted than for the RC, but the B/B_{MSY} trajectories remain fairly similar to those for the RC.
- Adjusting the ageing assumptions also has varying impacts on the appearance of the biomass trajectories, although not in an entirely predictable manner (i.e. RT22a with ageing out by one year is not an intermediate between RT22b (ageing halved) and the RC) – possibly there are subtle trade-offs when fitting to the ALK data and other data sources. Nonetheless, the B/B_{MSY} trajectories in the recent past and the future are very similar and remain above or close to one.
- The introduction of an artificial 150 000t cap from 1995 onwards (RT15) had minimal impact on the assessment results. The *M. capensis* estimate of depletion (B^{sp}/K^{sp}) at the 2007 low point did not change and for *M. paradoxus* it increased from 0.17 to 0.19.

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 main text. The column titled "OM re-run" indicates whether or not the robustness test in question required the OM to be re-conditioned.

No.	Description	OM re-run?
RT11	Decrease future survey CV's by a factor of 1/sqrt(2)	No
RT12	Trend in F-proportions for *M. paradoxus* over time in the future a. 2% p.a. increase for 10 years to a maximum of 0.9999), then constant b. 2% p.a. decrease for 10 years, then constant	No
RT13	Restrict survey q's to be less than one	Yes
RT14	Adjust survey estimates for industry vessels for month effect	Yes
RT15	Introduce a 150 000t cap from 1995	No[2]
RT16	Alternative survey calibration factors to cater for *Africana* gear change a. Increase cap factor to 0.90 b. Decrease cap factor to 0.60	Yes
RT17	*q* for CPUE drops by 20% from 1997 to 2002 as a result of shorter tows	Yes
RT18	Increase weighting of recent CPUE and survey data	Yes
RT19	Down-weight data for industry vessel surveys by a factor of 10, i.e. log-likelihood contributions multiplied by 0.1	Yes
RT20	Log-likelihood weighting (current ALK weighting is 0.01 and CAL weighting is 0.1) a. Weight ALK by 0.001 b. Weight ALK by 0.1 c. Weight CAL by 0.01 d. Weight CAL by 1.0	Yes
RT21	Discards a. Include discards in the past, but not the future b. Include discards in the past and future c. Past discards are halved, no discards in the future	Yes
RT22	Biological ageing a. Ageing of both species is out by one year b. Ageing of both species is halved	Yes
RT23	Gender-specific natural mortality- increase males +0.05, decrease female -0.05	Yes
RT24	Increase *M* at larger ages (currently 0.2)	Yes
RT25	Asymptotically flat offshore and inshore trawl selectivities for *M. capensis*	Yes
RT26	*M. capensis* offshore trawl fleet selectivity is assumed to be equal to that for the inshore trawl fleet	Yes

[2] Strictly speaking this is not a robustness test as it aims to answer the theoretical question of what would had happened if the catches had been less than what they really were, i.e. it does not reflect a plausible reality. It was, however, convenient to include the results in this document for record purposes.

Table 2: Estimates of the effect of month on abundance estimates, as provided by J. Glazer from GLM fits to commercial CPUE data. The estimates from the GLM (for ln CPUE) have been exponentiated, so that the CPUE is directly proportional to the values in the table.

Month	*M. capensis*	*M. paradoxus*
Jan	1.0039	0.9954
Feb	1.0340	1.0089
Mar	1.0521	1.0062
Apr	1.0721	1.0234
May	1.1598	1.0988
Jun	1.2152	1.1365
Jul	1.1673	1.0936
Aug	1.0704	1.0327
Sep	0.9774	0.9527
Oct	0.9166	0.8908
Nov	0.9327	0.9259
Dec	1.0000	1.0000

Table 3a: Negative log-likelihood components for the RC OM and the robustness tests for which the OM had to be reconditioned. Values in grey and with a star indicate that owing to the nature of the robustness test, these negative log-likelihood components are not comparable with the RC OM. The robustness test numbers are repeated in the last column for ease of comparison.

Run	GLM CPUE	ICSEAF	Surv. abun.	Comm. CAL	Survey CAL	Rec. resid.	ALKs	Total	Run
RC OM	−202.88	−37.72	−34.53	−1507.21	−1503.31	9.44	122.05	−3154.14	RC
RT13 (surv q<1)	−199.02	−34.89	−31.77	−1506.14	−1506.78	9.64	121.40	−3147.48	RT13
RT14 (ind. month effect)	−202.94	−37.72	−34.34	−1507.20	−1503.27	9.47	122.04	−3153.95	RT14
RT15 (150t cap from 1990)	−203.02	−38.01	−34.26	−1507.04	−1503.37	9.47	122.03	−3154.19	RT15
RT16a (Incr calibration)	−203.53	−37.33	−35.63	−1507.01	−1503.70	9.56	121.93	−3155.64	RT16a
RT16b (Decr calibration)	−202.78	−37.75	−33.79	−1507.24	−1503.03	9.42	122.06	−3153.09	RT16b
RT17 (CPUE q drop 97−02)	−194.19	−38.61	−36.19	−1505.39	−1501.85	9.21	122.84	−3144.15	RT17
RT18 (Upweight recent abund.)	−382.66*	−37.76	−1.11*	−1506.61	−1501.91	10.57	121.69	−3297.77*	RT18
RT19 (Downweight industry)	−202.14	−37.72	−43.16*	−1507.10	−1503.34	9.14	122.20	−3162.10*	RT19
RT20a (Weight ALK 0.001)	−206.11	−38.64	−41.71	−1527.09	−1513.53	12.47	27.04*	−3287.53*	RT20a
RT20b (Weight ALK 0.1)	−191.72	−37.67	−25.76	−1502.40	−1489.62	10.43	1116.28*	−2120.46*	RT20b
RT20c (Weight CAL 0.01)	−215.55	−37.09	−40.13	−140.90*	−144.23*	11.31	117.15	−449.38*	RT20c
RT20d (Weight CAL 1.0)	−153.55	−37.39	−26.58	−15177.39*	−15176.67*	32.80	148.97	−30389.70*	RT20d
RT21a&b (Past discards)	−198.61	−39.22	−31.63	−1487.39	−1504.78	10.13	121.70	−3129.77	RT21a&b
RT21c (Halve past discards)	−202.57	−38.38	−34.29	−1506.04	−1503.88	9.60	121.27	−3154.26	RT21c
RT22a (Ageing out by a year)	−200.26	−39.55	−35.13	−1500.44	−1498.41	13.08	138.44*	−3122.24*	RT22a
RT22b (Halve ageing)	−204.75	−37.35	−37.94	−1498.98	−1489.85	13.15	224.00*	−3031.16*	RT22b
RT23 (Gender−specific mort)	−198.51	−35.96	−34.11	−1505.71	−1497.65	8.36	122.50	−3141.01	RT23
RT24 (Increase large M)	−205.44	−37.61	−35.34	−1511.20	−1503.28	9.31	122.07	−3161.42	RT24
RT25 (asymp. mcap sel)	−195.59	−37.56	−34.53	−1417.00	−1508.81	10.52	127.39	−3055.56	RT25
RT26 (Off=Insh)	−216.12	−36.76	−36.27	−1503.41	−1499.77	9.28	125.19	−3157.84	RT26

Table 3b: The negative log-likelihood components from **Table 3**a are reported relative to the values of the RC OM. Values in grey indicate that the negative log-likelihood is worse than for the RC OM. In this table, values for negative log-likelihood components that are non-comparable with the RC have been excluded (marked with a -). For these cases, the values 'Total' column correspond to the differences in comparable likelihood components only (and are marked with a star).

Run	GLM CPUE	ICSEAF	Surv. abun.	Comm. CAL	Survey CAL	Rec. resid.	ALKs	Total	Run
RC OM	−202.88	−37.72	−34.53	−1507.21	−1503.31	9.44	122.05	−3154.14	RC
RT13 (surv q<1)	3.86	2.83	2.76	1.06	−3.47	0.20	−0.65	6.66	RT13
RT14 (ind. month effect)	−0.06	−0.00	0.19	0.00	0.04	0.03	−0.01	0.19	RT14
RT15 (150t cap from 1990)	−0.15	−0.29	0.26	0.17	−0.06	0.03	−0.02	−0.05	RT15
RT16a (Incr calibration)	−0.65	0.40	−1.10	0.19	−0.39	0.12	−0.12	−1.50	RT16a
RT16b (Decr calibration)	0.10	−0.02	0.74	−0.03	0.27	−0.02	0.01	1.05	RT16b
RT17 (CPUE q drop 97−02)	8.69	−0.89	−1.66	1.82	1.46	−0.23	0.79	9.99	RT17
RT18 (Upweight recent abund.)	−	−0.04	−	0.60	1.39	1.12	−0.36	2.72*	RT18
RT19 (Downweight industry)	0.74	0.00	−	0.10	−0.03	−0.30	0.15	0.67*	RT19
RT20a (Weight ALK 0.001)	−3.23	−0.92	−7.18	−19.88	−10.22	3.03	−	−38.40*	RT20a
RT20b (Weight ALK 0.1)	11.15	0.05	8.77	4.80	13.69	0.99	−	39.46*	RT20b
RT20c (Weight CAL 0.01)	−12.67	0.64	−5.60	−	−	1.87	−4.90	−20.66*	RT20c
RT20d (Weight CAL 1.0)	49.33	0.34	7.95	−	−	23.36	26.93	107.90*	RT20d
RT21a&b (Past discards)	4.27	−1.50	2.90	19.82	−1.47	0.69	−0.35	24.37	RT21a&b
RT21c (Halve past discards)	0.31	−0.65	0.24	1.16	−0.57	0.16	−0.78	−0.12	RT21c
RT22a (Ageing out by a year)	2.62	−1.82	−0.60	6.77	4.89	3.64	−	15.49*	RT22a
RT22b (Halve ageing)	−1.87	0.37	−3.42	8.22	13.45	3.71	−	20.47*	RT22b
RT23 (Gender−specific mort)	4.37	1.77	0.42	1.50	5.65	−1.08	0.46	13.13	RT23
RT24 (Increase large M)	−2.56	0.11	−0.81	−3.99	0.02	−0.13	0.02	−7.28	RT24
RT25 (asymp. mcap sel)	7.29	0.16	−0.00	90.20	−5.51	1.08	5.34	98.58	RT25
RT26 (Off=Insh)	−13.24	0.96	−1.74	3.79	3.54	−0.16	3.15	−3.70	RT26

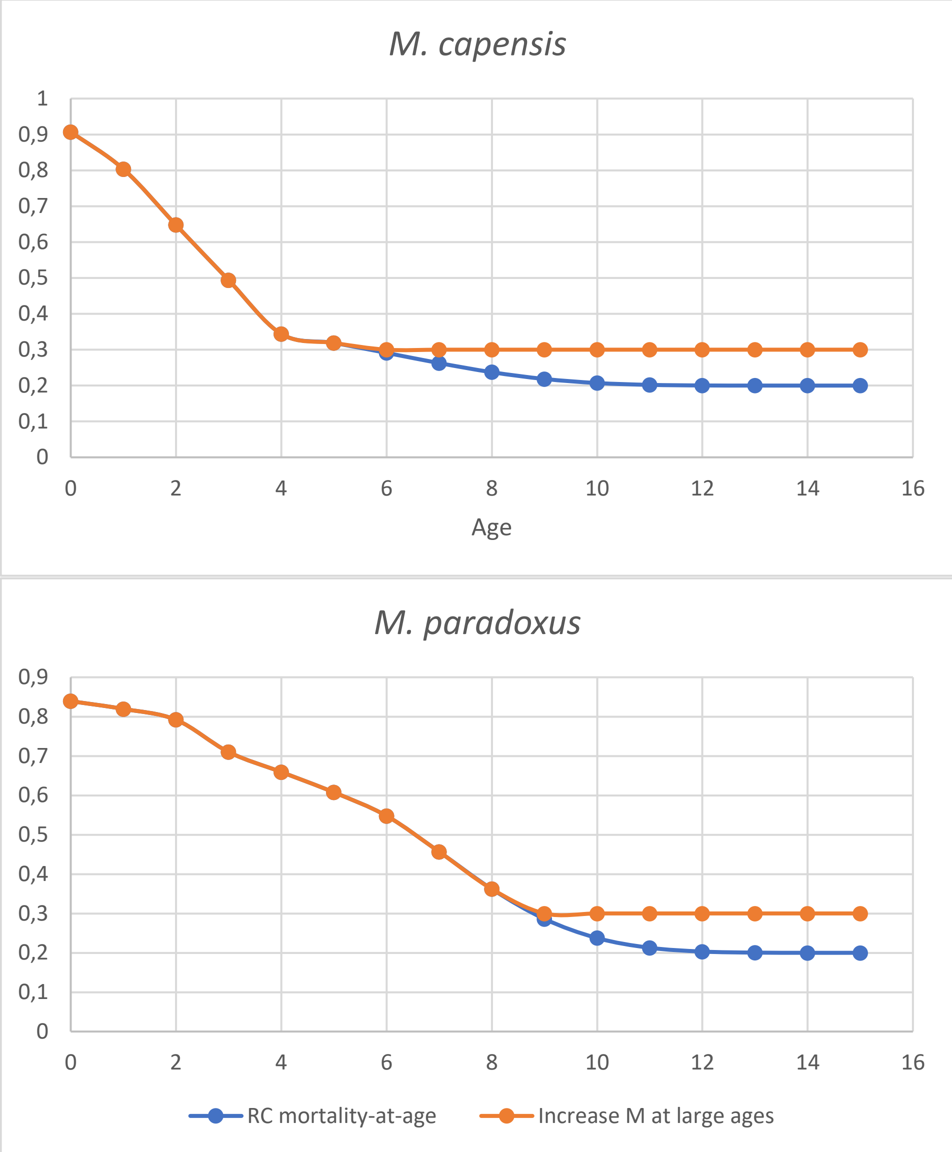

Figure 1: The natural mortality-at-age vectors assumed for the RC OM and for robustness test 19, where the mortality is increased at larger ages.

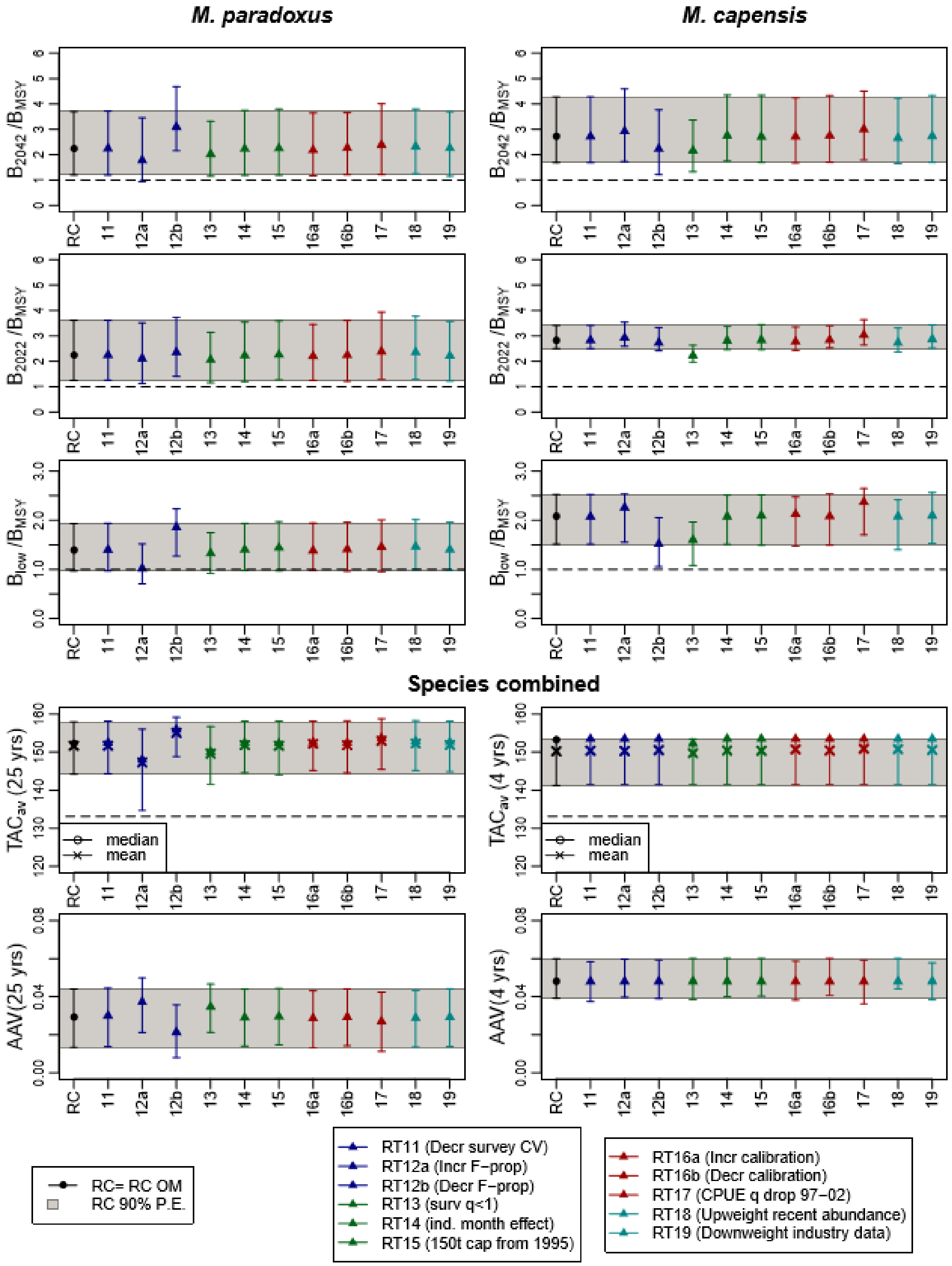

Figure 2a: Zeh plots of the performance statistics from Table 3, for the RC and **first 11 robustness tests**. The statistics are B^{sp}/B_{MSY} for 2042 and 2022, $B^{sp}(low)/B_{MSY}$ (the lowest value of this statistic in the projection period to 2042), TAC_{av} (the average catch over the projection 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. The colours have no significance other than to match to the colours in Figure 3a-h.

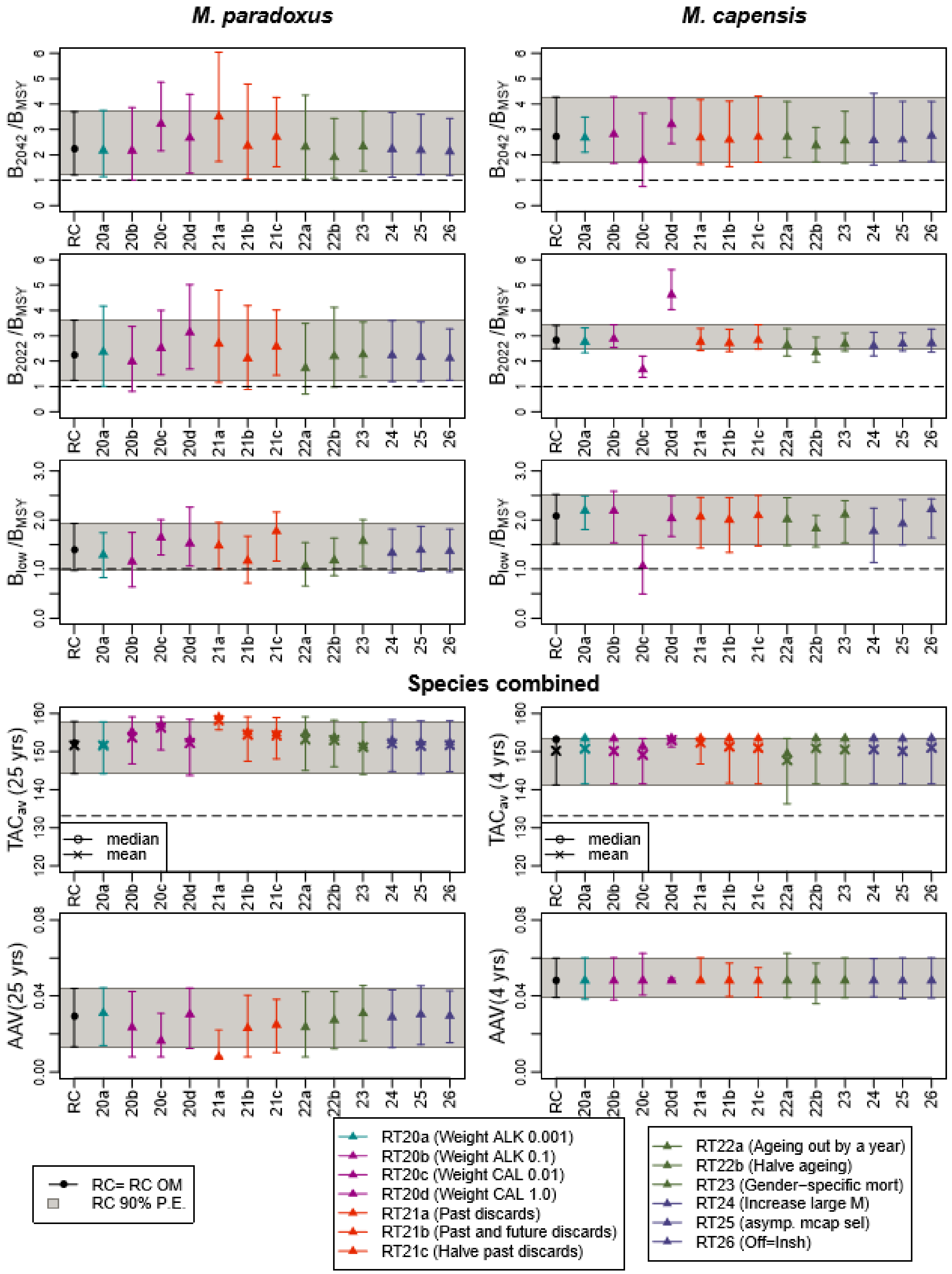

Figure 2b: Zeh plots of the performance statistics from Table 3, for the RC and **second set of 13 robustness tests**.

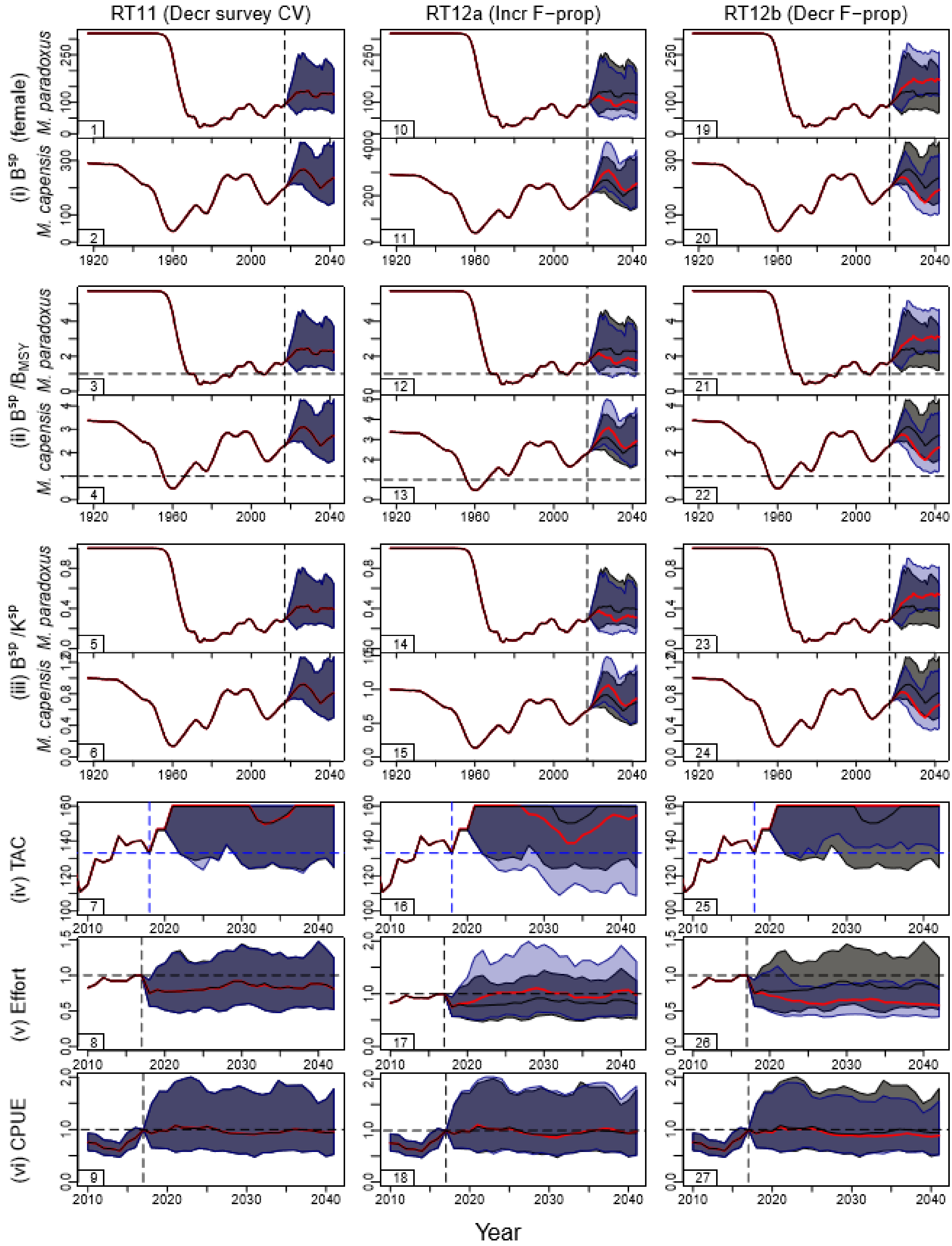

Figure 3a: Projected trajectories for female spawning biomass (in absolute terms, relative to B_{MSY} and relative to K^{sp}), TAC, effort and CPUE are shown for the RC and **robustness tests RT11, RT12a and RT12b**. The RS02 RC OM results are shown by the black lines (median trajectories) and grey shading (90% P.E.), while the robustness test results are indicated by the red lines and coloured shading.

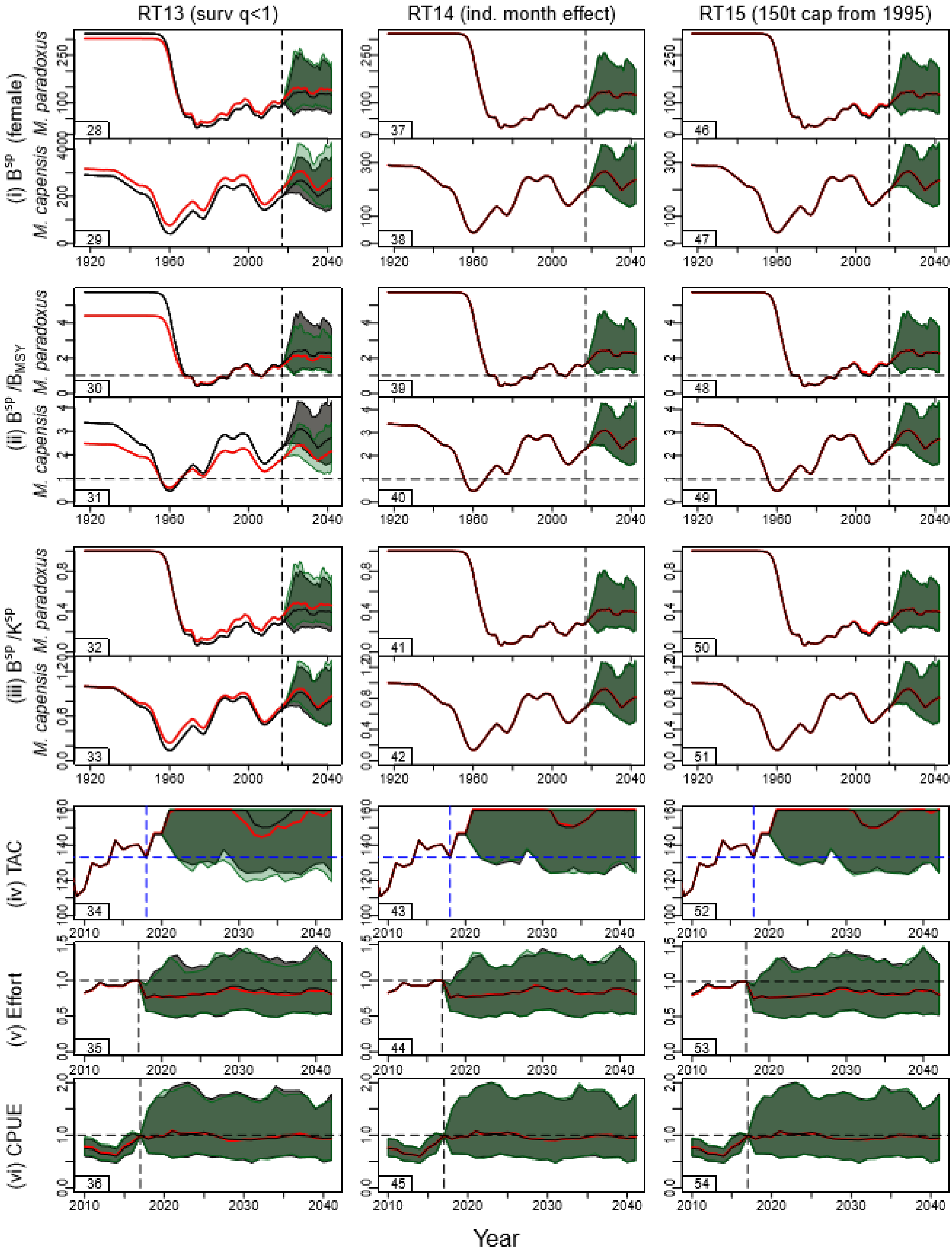

Figure 3b: Projected trajectories for **robustness tests RT13, RT14 and RT15.**

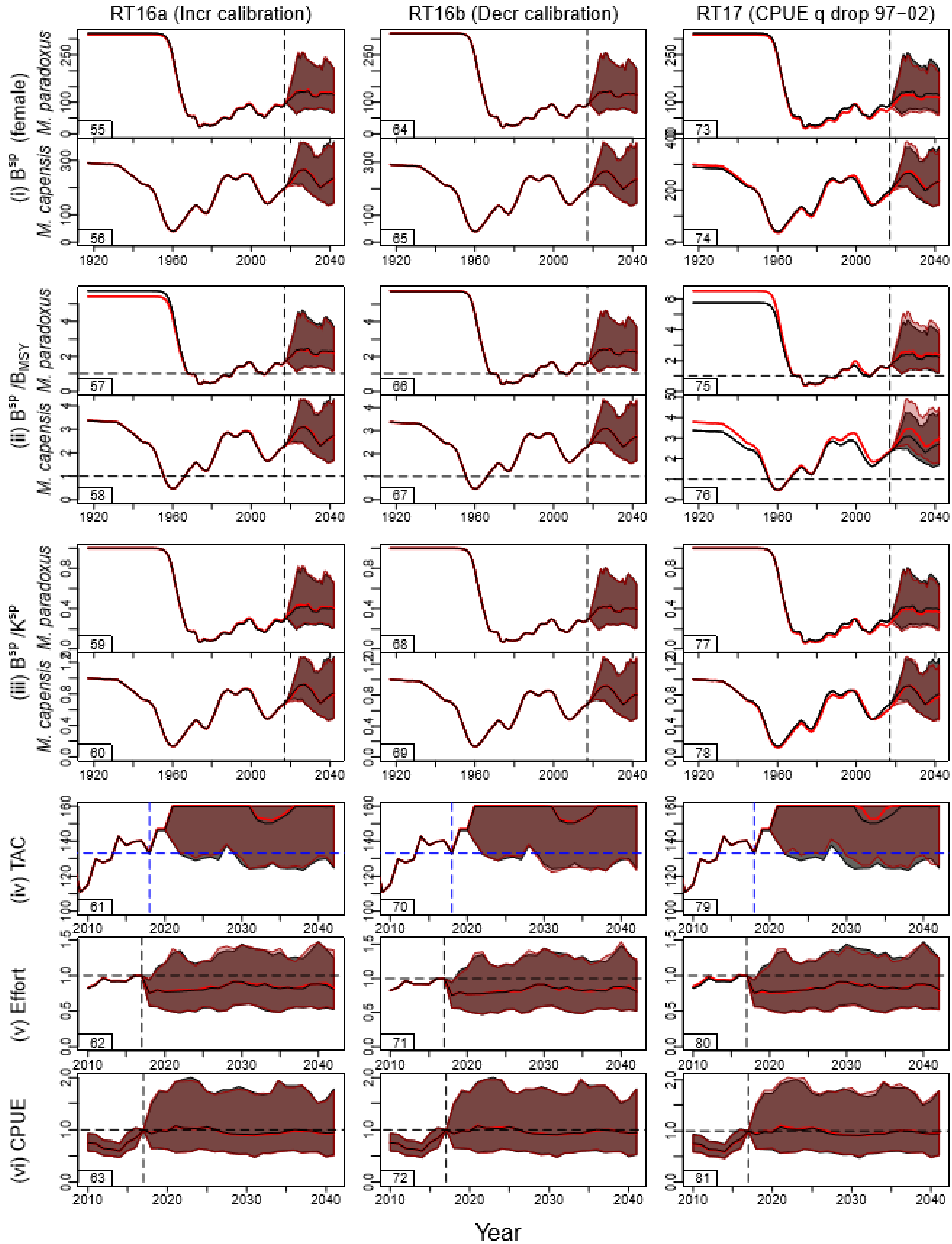

Figure 3c: Projected trajectories for **robustness tests RT16a, RT16b and RT17.**

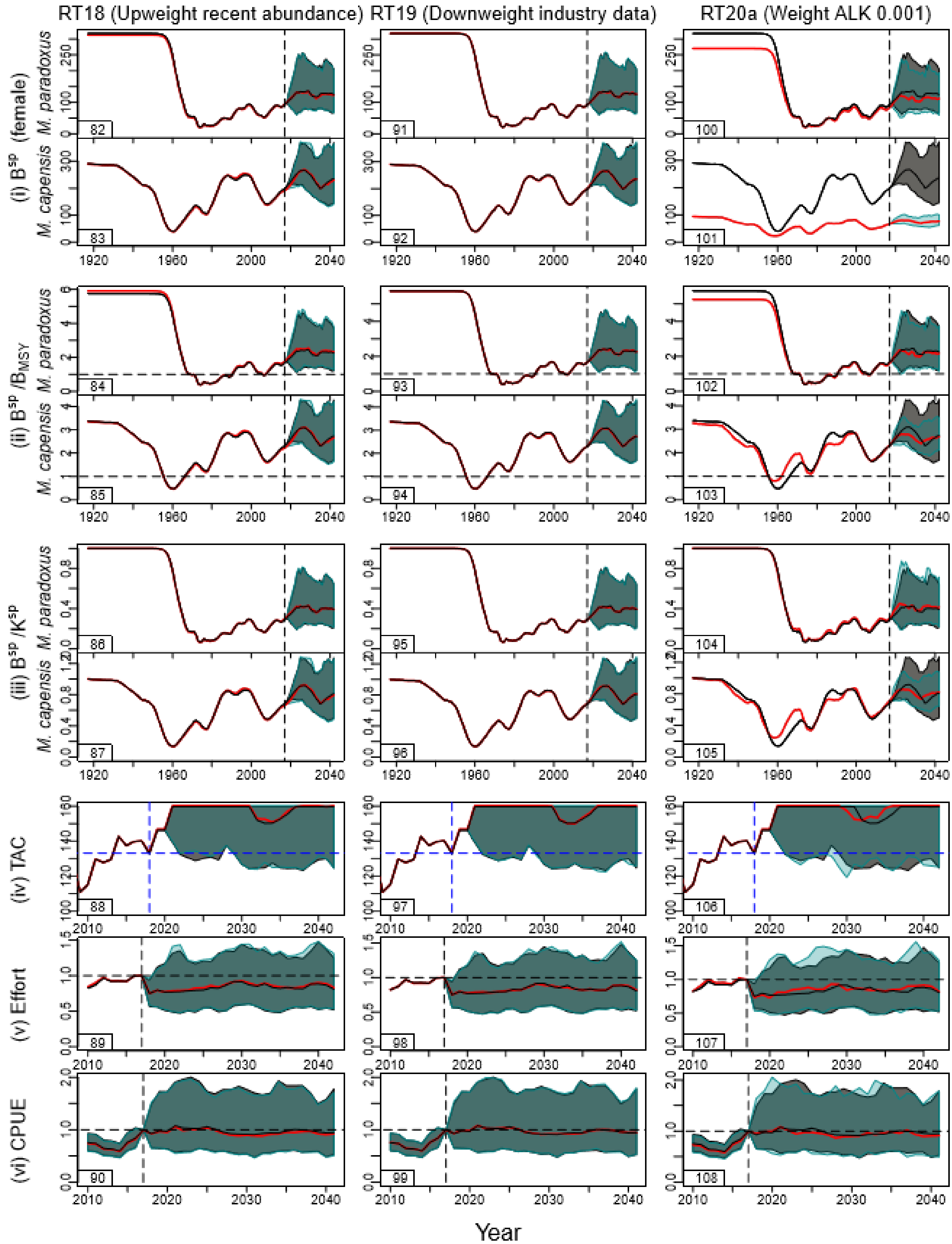

Figure 3d: Projected trajectories for **robustness tests RT18, RT19 and RT20a.**

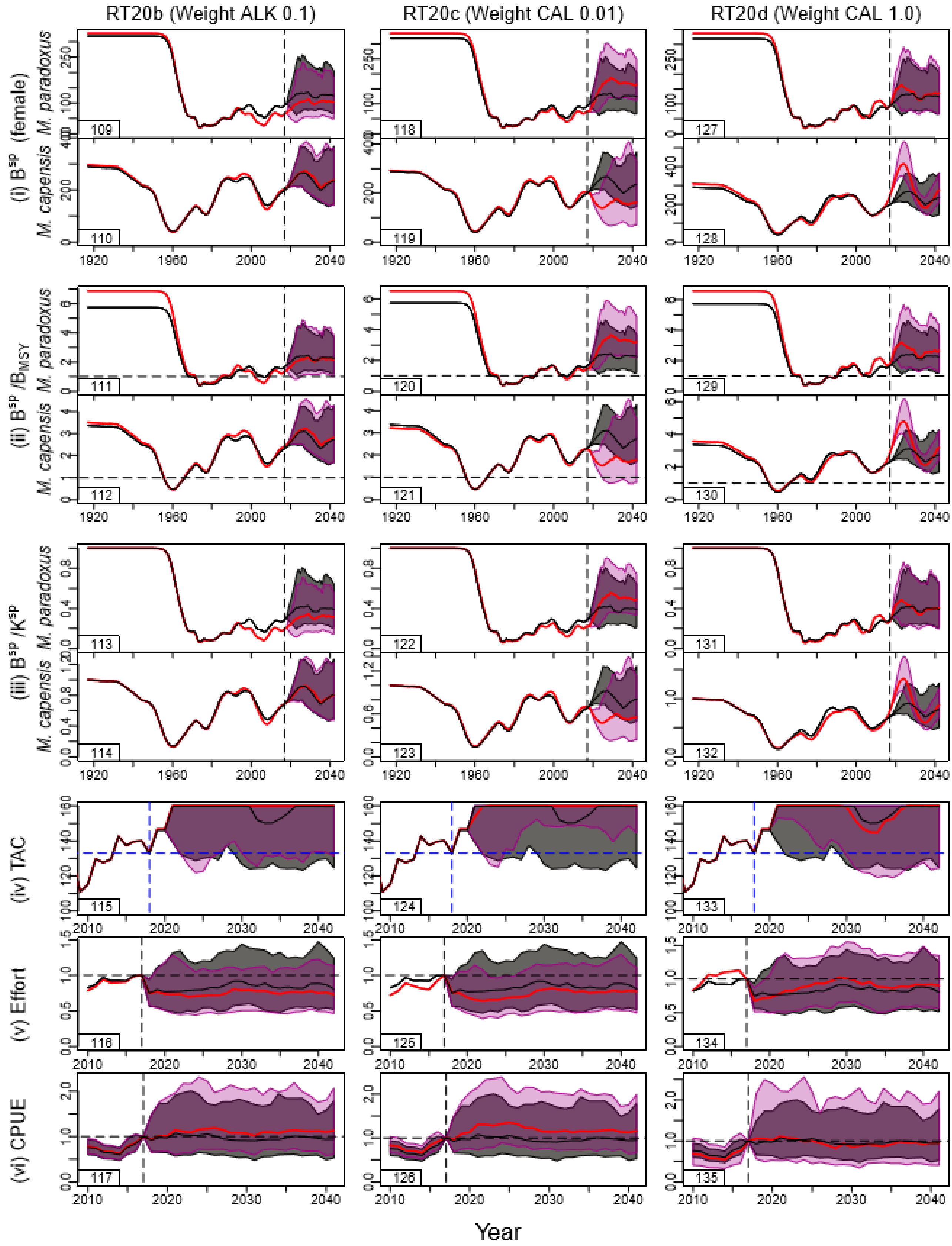

Figure 3e: Projected trajectories for **robustness tests RT20b, RT20c and RT20d.**

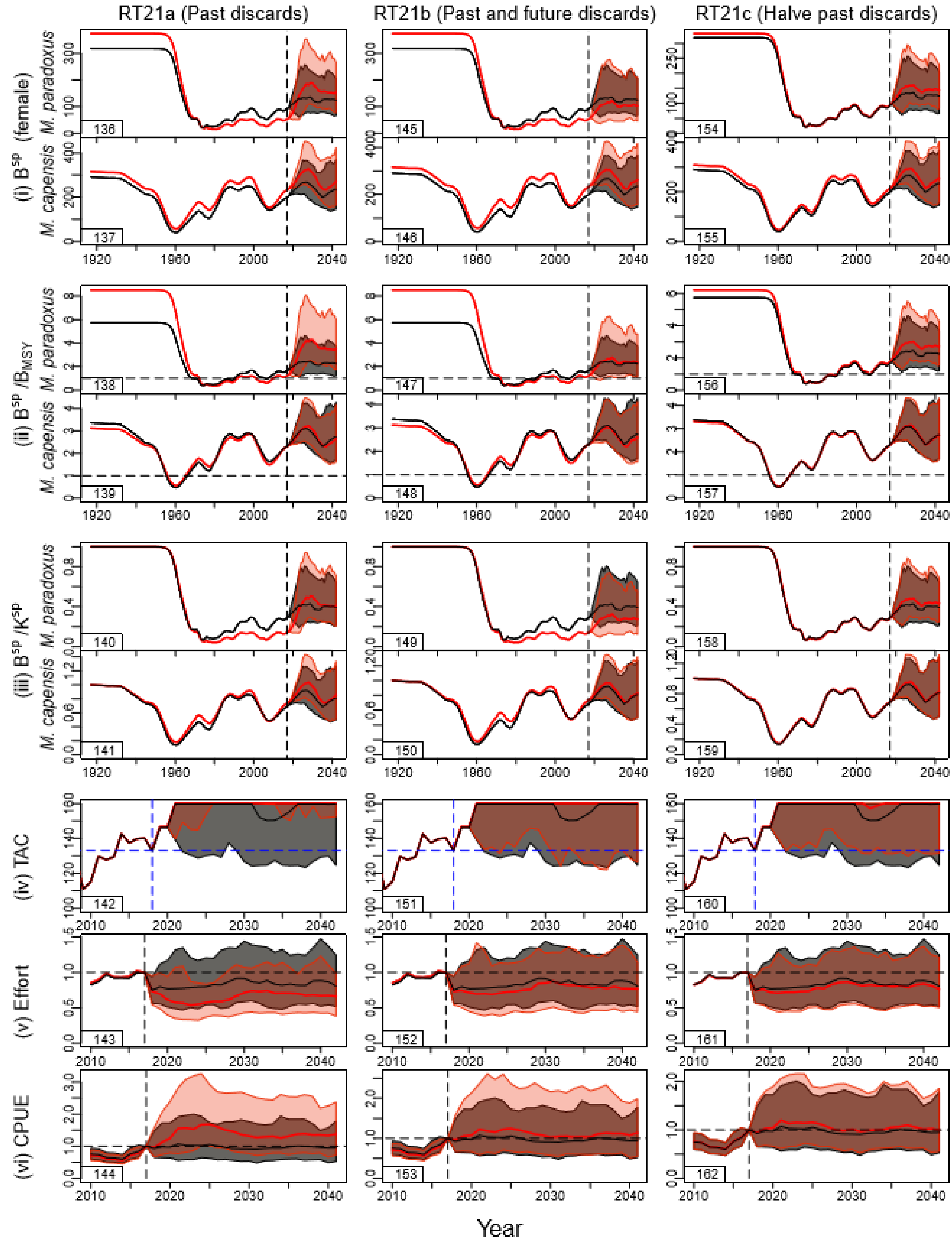

Figure 3f: Projected trajectories for **robustness tests RT21a, RT21b and RT21c.**

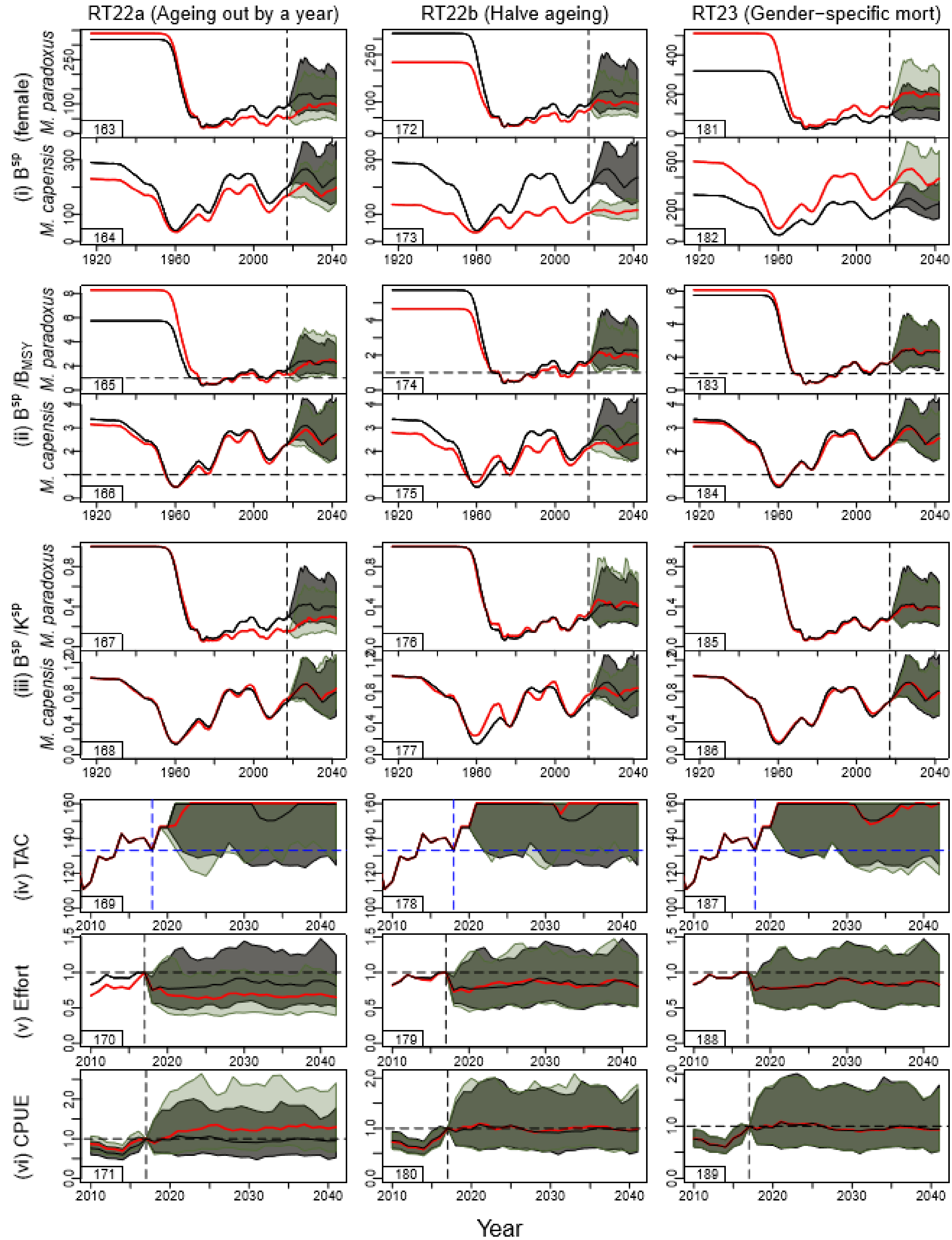

Figure 3g: Projected trajectories for **robustness tests RT22a, RT22b and RT23.**

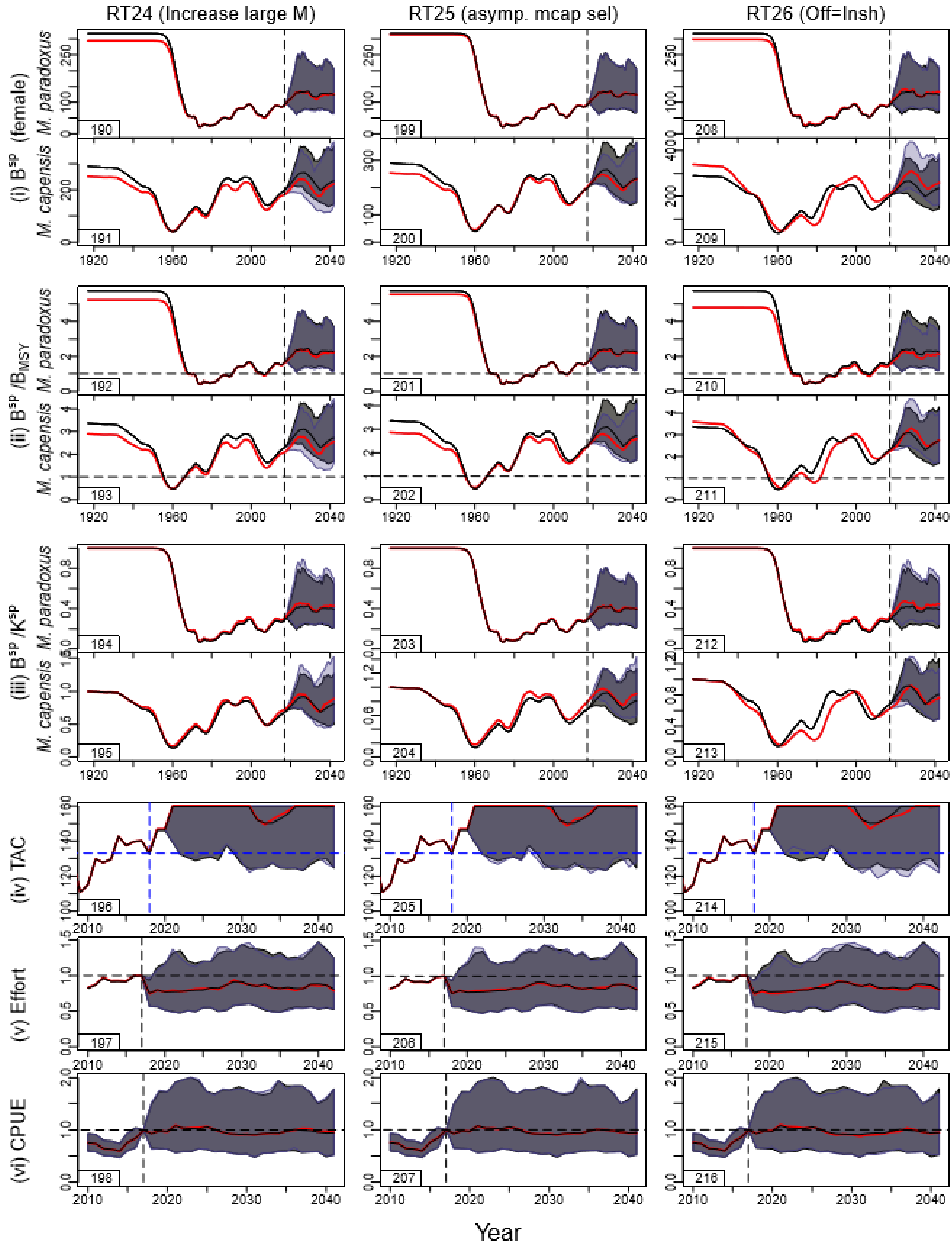

Figure 3h: Projected trajectories for **robustness tests RT24, RT25 and RT26.**

Appendix: Results for the models starting in 1978

This appendix gives the results of the models starting in 1978 alongside those of the other nine models in the RS as reported in the RS conditioning document MARAM/IWS/2018/Hake/P3.

Method

The main differences between the model starting in 1917 and in 1978 are the assumptions which have to be made related to the population structure in the first year (*y0*) of the model. Under pristine equilibrium conditions (i.e. 1917) the assumption is made that the number of fish in age cohort *a*+1 can be calculated as $N_{y0,a+1} = N_{y0,a}e^{-M_a}$. Since the population would not have been at equilibrium conditions in 1978, some different assumptions have to be made for that case.

The numbers in the first cohort of age zero are given by the stock recruitment relationship. For the 1978 runs presented in this document, the subsequent numbers-at-age are given by:

$$N_{1978,a+1} = N_{1978,a}e^{-(M_a+\varphi)} \tag{A1}$$

where M_a is the natural at age *a* and φ is an additional mortality component to account for fishing mortality in preceding years.

Furthermore, the proportion of spawning biomass in 1978 relative to pristine values in 1917 is estimated, i.e. by a parameter $\theta = B^{sp}_{1978}/K^{sp}$. The model can also be adjusted to allow $N_{1978,a+1}$ to be estimated for all ages up to some specific age $a*$, but this has not yet been attempted at this stage.

When first implementing this model, serious issues were experienced with instability in the minimisation. In order to produce some results, several parameters had to be fixed. These parameters are:

- The offshore selectivity parameters (West coast and South Coast) - nine parameters
- The inshore selectivity parameters - three parameters
- The scaling factor for female *M. paradoxus* selectivity (south coast surveys only) - two parameters
- The sigma parameters for the age-length matrix variance - 12 parameters
- The growth curve parameters - 12 parameters

These parameters were fixed at the (best fitting) RS02 "Reference Case" run values. Additionally, the stock-recruitment parameters (*h* and *γ*) appeared to be very poorly estimated, so these were also fixed at the RS02 values.

Results and discussion

Table A4a and Table A4b list key parameter estimates, while Table A5 gives the negative log-likelihood components. Figure A4a and b show the spawning biomass trajectories for *M. paradoxus* and *M. capensis* respectively, and Figure A5 shows the fits to the CPUE data.

Results for RS04 (Ricker model starting in 1978) and RS08a (Beverton-Holt with *h*=0.9, also starting in 1978) are relatively comparable with the other models in their respective groups. RS08b however (Beverton-Holt with *h*=0.7, starting in 1978) shows very different trends. Care should be taken when interpreting these results because of the estimation instability described above. Therefore, while these models certainly show potential for providing good fits to the data, they need to be developed further before they are of a similar reliability to the other RS models and robustness test models.

Table A4a: Key parameter estimates for the RS models (biomass units are thousand tons), including the runs starting in 1978. Cases where the current spawning biomass is below its MSY value are in bold font.

Model name	Central Year	Stock Recruit	*M. paradoxus* K^{sp}	B^{sp}_{MSY}	B^{sp}_{2017}	B^{tot}_{2017}	B^{sp}_{2017}/K^{sp}	$B^{sp}_{2017}/B^{sp}_{MSY}$	B^{sp}_{MSY}/K^{sp}	MSY	*M. capensis* K^{sp}	B^{sp}_{MSY}	B^{sp}_{2017}	B^{tot}_{2017}	B^{sp}_{2017}/K^{sp}	$B^{sp}_{2017}/B^{sp}_{MSY}$	B^{sp}_{MSY}/K^{sp}	MSY
(0) Oct 2017	1958	Ricker	515	115	127	245	0.25	1.11	0.22	137	196	63	141	334	0.72	2.23	0.32	81
(1) RS01	1952		340	53	88	196	0.26	1.65	0.16	144	412	96	294	647	0.71	3.06	0.23	112
(2) RS02	1958	Ricker	318	55	93	206	0.29	1.67	0.17	145	290	86	198	446	0.68	2.30	0.30	84
(3) RS03	1963		266	63	103	223	0.39	1.62	0.24	146	465	142	343	750	0.74	2.42	0.31	106
(4) RS04	Start `78		308	53	72	165	0.23	1.36	0.17	143	234	69	146	337	0.63	2.11	0.30	68
(5) RS05a	1952		520	50	77	181	0.15	1.53	0.10	141	418	84	35	104	0.08	**0.42**	0.20	53
(6) RS06a	1958	Beverton-Holt (h=0.9)	527	51	84	194	0.16	1.64	0.10	140	1213	215	877	1874	0.72	4.07	0.18	134
(7) RS07a	1963		540	51	95	219	0.18	1.85	0.10	142	1553	274	1180	2507	0.76	4.31	0.18	170
(8) RS08a	Start `78		624	81	81	184	0.13	1.00	0.13	142	408	76	137	316	0.34	1.80	0.19	47
(9) RS05b	1952		77	16	26	165	0.34	1.67	0.20	153	536	154	90	217	0.17	**0.58**	0.29	48
(10) RS06b	1958	Beverton-Holt (h=0.7)	82	17	30	177	0.36	1.78	0.20	154	1442	398	1045	2224	0.72	2.63	0.28	120
(11) RS07b	1963		88	18	36	216	0.41	2.07	0.20	165	746	217	59	152	0.08	**0.27**	0.29	69
(12) RS08b	Start `78		2004	489	98	224	0.05	**0.20**	0.24	315	501	142	112	264	0.22	**0.79**	0.28	44

Table A4b: Some further parameter estimates. Note that B^{sp}_{1978}/K^{sp} is a model output for the first three models in each group and an estimable parameter (θ) for the last in each group where the assessment starts in 1978.

			M. paradoxus					*M. capensis*				
Model name	Central Year	Stock Recruit	K^{sp}	h	γ	B^{sp}_{1978}/K^{sp}	φ	K^{sp}	h	γ	B^{sp}_{1978}/K^{sp}	φ
(0) Oct 2017	1958	Ricker	514.9826	1.264696	0.382267	0.111883	NA	196	1.342332	0.856512	0.342578	NA
(1) RS01	1952		340	1.50	0.34	0.07	NA	412	2.00	0.58	0.41	NA
(2) RS02	1958	Ricker	318	1.62	0.42	0.08	NA	290	2.00	0.85	0.37	NA
(3) RS03	1963		266	1.90	0.71	0.12	NA	465	1.60	0.79	0.42	NA
(4) RS04	Start `78		308	1.62	0.42	0.13	0.39	234	2.00	0.85	0.40	0.25
(5) RS05a	1952		520	0.90	NA	0.13	NA	418	0.90	NA	0.09	NA
(6) RS06a	1958	Beverton-Holt (h=0.9)	527	0.90	NA	0.18	NA	1213	0.90	NA	0.67	NA
(7) RS07a	1963		540	0.90	NA	0.27	NA	1553	0.90	NA	0.71	NA
(8) RS08a	Start `78		624	0.90	NA	0.07	0.37	408	0.90	NA	0.25	0.26
(9) RS05b	1952		77	0.70	NA	0.29	NA	536	0.70	NA	0.19	NA
(10) RS06b	1958	Beverton-Holt (h=0.7)	82	0.70	NA	0.35	NA	1442	0.70	NA	0.70	NA
(11) RS07b	1963		88	0.70	NA	0.44	NA	746	0.70	NA	0.08	NA
(12) RS08b	Start `78		2004	0.70	NA	0.02	0.47	501	0.70	NA	0.21	0.22

Table A5: Negative log-likelihood components are shown for the RS models. For the models starting in 1978, the historical CPUE data are not relevant as they end in 1978, neither are the Age-length keys as the growth curve parameters have been fixed for these models. Additionally, three years' of South Coast offshore catch-at-length data are ignored in the 1978 model, as this series starts in 1975 – hence the commercial catch-at-length values are also not comparable with the other RS models. These columns are marked with a star to indicate that the incomparability of the Start '78 runs, and additionally grey font and italics have been used to show values that are not comparable across the models. For the Oct 2017 model the incomparability is as a result of the old treatment of the catch-at-length data. The values in brackets indicate the difference between the comparable -lnL for a given run with the minimum across the RS.

Model name	Central Year	Stock Recruit	Total -lnL		historical CPUE*	GLM CPUE	Survey	Comm. CAL*	Comm. Sex-disagg CAL	Survey CAL	Survey sex-disagg CAL	Age-length Keys*	Rec. Resid.
(0) Oct 2017	1958	Ricker	*-5251.5*	(-)	-40.8	-191.4	-35.1	*-1330.6*	*-1110.6*	*-709.7*	*-1968.3*	124.5	10.4
(1) RS01	1952	Ricker	-3151.2	(2.9)	-37.5	-200.9	-34.4	-823.6	-682.0	-413.3	-1090.7	122.3	8.9
(2) RS02	1958		-3154.1	(0.0)	-37.7	-202.9	-34.5	-825.6	-681.6	-413.3	-1090.0	122.0	9.4
(3) RS03	1963		-3153.0	(1.2)	-36.9	-202.7	-34.4	-823.2	-682.1	-413.3	-1090.9	121.8	8.5
(4) RS04	Start `78		-	-	NA	-193.0	-36.1	*-804.2*	-682.0	-411.8	-1090.6	NA	9.5
(5) RS05a	1952	BH (h=0.9)	-3134.9	(19.3)	-40.2	-183.5	-33.4	-827.9	-681.3	-416.9	-1087.3	123.1	12.6
(6) RS06a	1958		-3122.8	(31.4)	-36.0	-172.0	-32.6	-821.3	-686.1	-416.8	-1094.4	124.2	12.2
(7) RS07a	1963		-3120.1	(34.0)	-37.0	-167.9	-32.0	-821.0	-685.5	-417.0	-1094.7	124.3	10.7
(8) RS08a	Start `78		-	-	NA	-195.4	-36.6	*-801.8*	-681.8	-411.8	-1090.4	NA	7.5
(9) RS05b	1952	BH (h=0.7)	-3122.2	(31.9)	-37.7	-185.6	-35.0	-826.5	-680.1	-417.2	-1092.5	139.2	13.1
(10) RS06b	1958		-3106.6	(47.6)	-35.5	-174.0	-33.4	-821.4	-681.6	-418.6	-1096.5	140.8	13.6
(11) RS07b	1963		-3117.8	(36.3)	-37.5	-181.4	-33.8	-830.8	-677.7	-418.3	-1091.4	141.2	11.3
(12) RS08b	Start `78		-	-	NA	-207.0	-37.0	*-799.8*	-680.9	-413.0	-1089.8	0.0	7.7

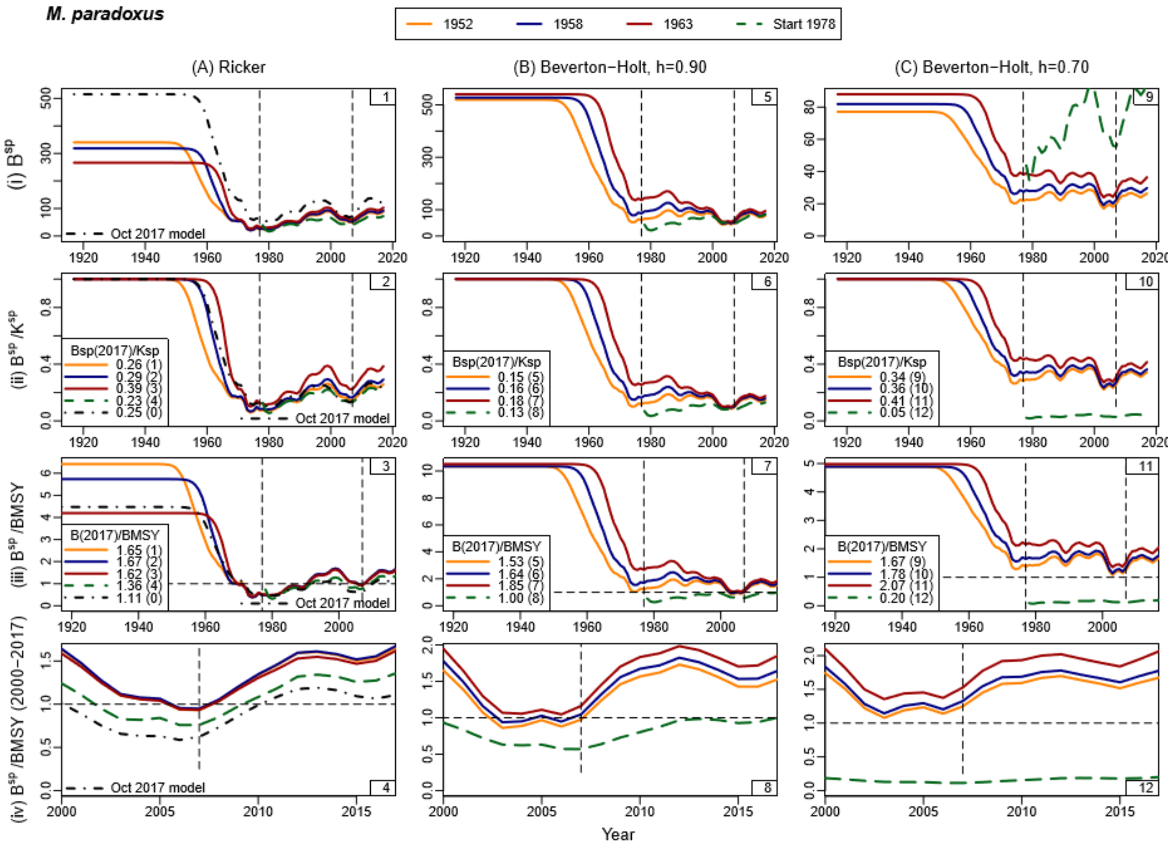

Figure A4a: Female spawning biomass trajectories are shown for ***M. paradoxus***. In the plots, yellow lines have been used for the models with the central year of shift occurring in 1952, blue lines for the 1958 models and red lines for the 1963 models. Dashed green lines have been used for the model starting in 1978. The Oct 2017 model has been included in the first column with the RS Ricker models (black dash-dot lines).

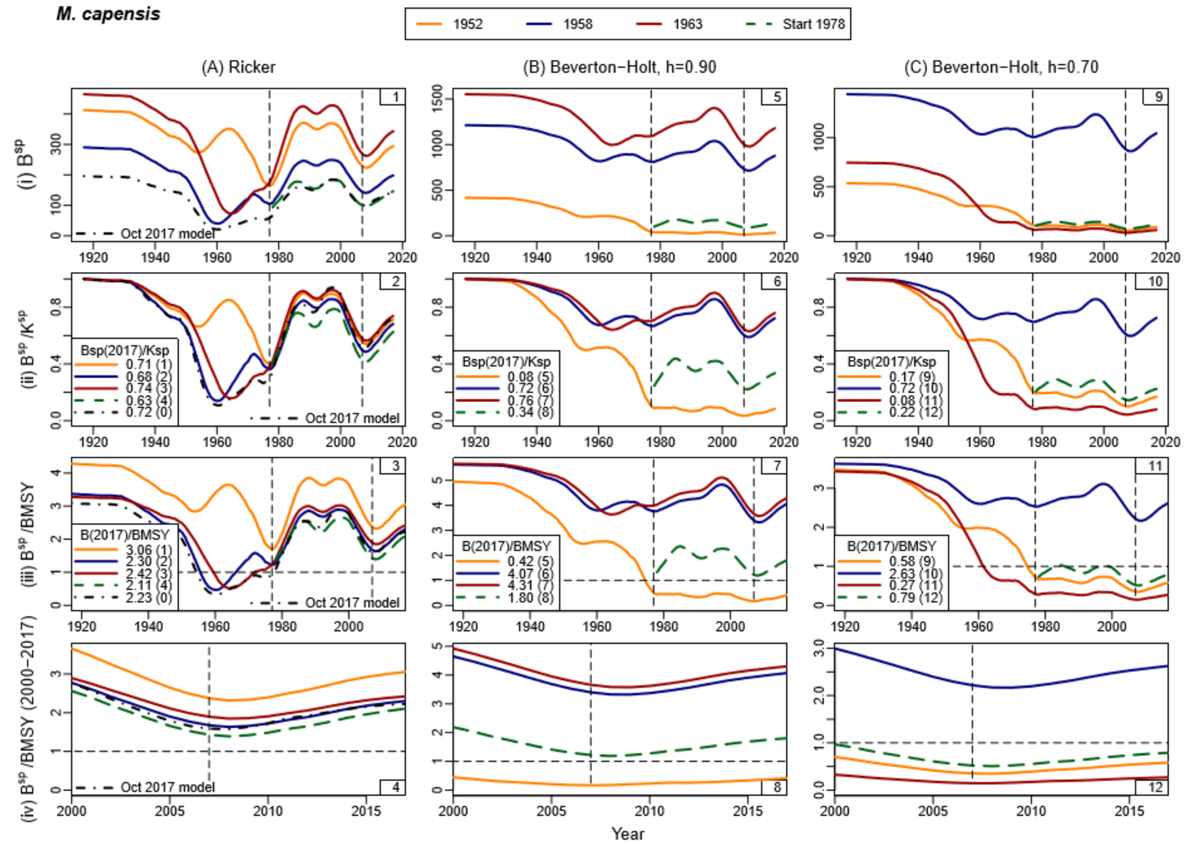

Figure A4b: Female spawning biomass trajectories are shown for ***M. capensis.***

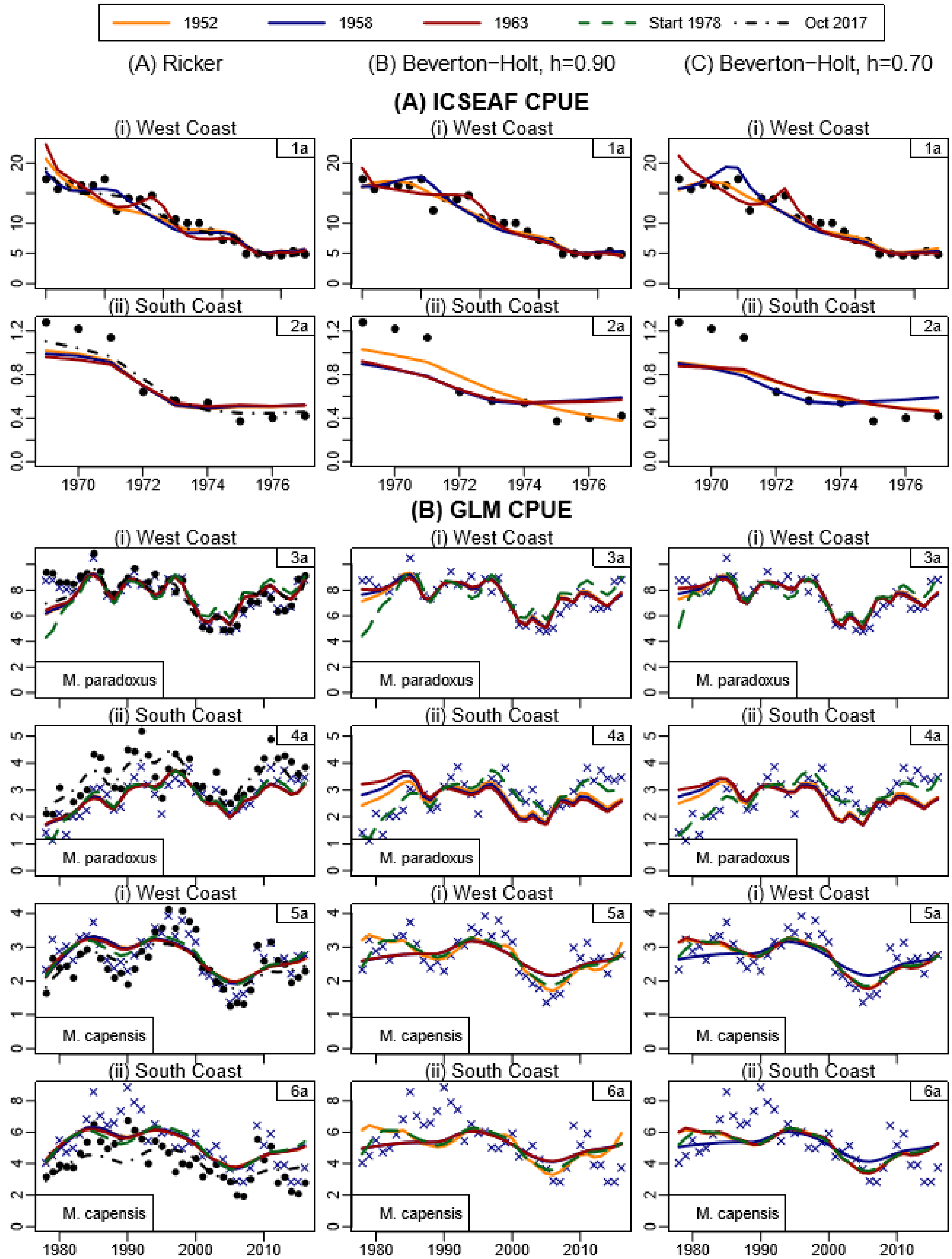

Figure A5: Fits to the ICSEAF and commercial CPUE data. All three columns show the new data for the GLM CPUE. The first column, which shows the Ricker models including the Oct 2017 model, additionally shows the old data with the filled circles.