Updated assessment of the South African sardine resource using data from 1984-2020

C.L. de Moor*

Correspondence email: carryn.demoor@uct.ac.za

A quantitative assessment of the South African sardine resource has been updated to include data from 1984 to 2020. This two mixing-component hypothesis assumes no stock recruitment relationship during conditioning. The west component abundance remains at a low level, but the spawner biomass is estimated to be higher than that estimated a year ago. These results have shown that the model is sensitive to the inclusion of parasite prevalence-at-length data and further investigation into the most appropriate manner with which these data should be included in the model is warranted. Further sensitivity testing and the consideration of alternative hypotheses is required before a 'baseline' assessment can be selected.

Keywords: assessment, population dynamics, parasite prevalence-at-length, sardine

Introduction

This document presents results for a 'simple updated' assessment of South African sardine, using data from 1984-2020. The assessment assumes the sardine population consists of two mixing 'components', with a west component distributed west of Cape Agulhas and a south component distributed south-east of Cape Agulhas. This hypothesis assumes there is movement of fish from the west component into the south component and a small contribution of south component spawning to west component recruitment. The former is modelled by an annually-varying proportion of west component sardine moving to form part of the south component while the latter is modelled by assuming a time-invariant 8% of south component spawner biomass contributes to west component effective spawner biomass.

Data and Population Dynamics Model

The population dynamics model for the South African sardine resource is detailed in Appendix A (extended from de Moor (2020a). The data used in this assessment are listed in de Moor *et al*. (2021). The prevalence-at-length data collected from annual acoustic surveys were originally included in the assessment model to help improve the confidence with which movement (in particular) was estimated (de Moor *et al*. 2017, Figure 1). The inclusion of a time series of parasite prevalence-at-length from commercial catches (aligned with the four annual model quarters) as well as parasite intensity-at-length data into the likelihood of the model has been proposed many times, but not yet undertaken due to time constraints. Given low samples of large fish on the south coast in the November 2019 and 2020 acoustic surveys, van der Lingen (2021) proposed that samples from commercial catches collected between September and November of these years be added to the samples from the acoustic survey off the south coast in 2019 and from the west and south coasts in 2020. Further work is still ongoing to update van der Lingen (2021) such that the comparisons of historical prevalence-at-length from survey and commercial data are undertaken using samples from the same area and time period. In the meantime, the assessment is run both with and without the prevalence-at-length data from survey and/or commercial samples in 2019 and 2020 to consider the impact of including commercial data for two years only.

Results and Discussion

Table 1 lists the changes in model fits to the data and estimates of final year biomass depending on which parasite prevalence-at-length data are used to condition the model. Including the prevalence-at-length data sampled from acoustic surveys in the model results in some improvement in the estimation of the annual proportion moving in the past decade, at the expense of fitting other

* MARAM (Marine Resource Assessment and Management Group), Department of Mathematics and Applied Mathematics, University of Cape Town, Rondebosch, 7701, South Africa.

data (Figures 2a,c, Table 1). In some years, the confidence intervals are decreased substantially. The confidence intervals in the estimate of the proportion moving in 2013 are mutually exclusive between including or excluding the prevalence data, which merits further investigation (this was not observed by de Moor *et al*. 2017, see Figure 1). The inclusion of the 2019 prevalence-at-length data only results in a substantial change in estimated movement in 2019 (Figure 2a), with a substantial increase/decrease in $B^S_{w,2019}$ / $B^S_{s,2019}$, with a smaller increase/decrease in $B^S_{w,2020}$ / $B^S_{s,2020}$. The inclusion of 2020 prevalence-at-length data results in a substantial increase in the estimated biomass in November 2020, with little change to the November likelihood (Table 1). The Hessian-based confidence interval of the estimate of the proportion moving in November 2020 when the additional commercial samples are included, only just includes the best estimate achieved without the 2020 commercial samples (Figures 2b,c).

While narrowing of the confidence intervals of some estimated parameters with little deterioration in the fit to the acoustic survey data was expected *a priori* by including the prevalence-at-length data in the model (de Moor *et al*. 2017), these results indicate there may be some conflict between the data used to condition this model and this should be explored further.

The model is able to fit the survey estimates of abundance well in most years (Figures 3 and 4). The years for which the model predicted value exceeded the 95% confidence interval of the survey estimate typically corresponded to years for which there is conflicting information between the November biomass and May/June recruit surveys. The relatively high model predicted recruitment on the south coast in May/June 2018 to 2020 are now primarily informed by length frequency data given the absence of survey estimates of recruitment in these recent years. There is a deterioration in the model's ability to fit these survey data in the most recent three years when prevalence-at-length data are included in the likelihood.

The model fit to the survey length frequencies (Figures 6, 7) and commercial length frequencies (Figures 9 and 10) is relatively good, with little noticeable change to that achieved by de Moor (2020a), and with similar survey and commercial selectivities to that estimated by de Moor (2020a) (Figures 5 and 8). The fit to survey and commercial length frequencies improves substantially when the prevalence-at-length data are excluded from the likelihood (Figures 7, 10).

Figure 11 shows the growth curves which are modelled to vary by cohort. For model iv) the average $t_{0,j,y}$ for the von Bertalanffy growth curves for the south component is 0.25 (close to November) with a maximum difference in $t_{0,j,y}$ between cohorts of approximately 8 months. Even with an adjustment to the growth curve for the west component, the average $t_{0,j,y}$ is approximately 7 months before November, with a maximum difference in $t_{0,j,y}$ between cohorts of about 7.5 months. These ranges are similar to that of de Moor (2020a), and as previously suggested further alternative adjustments to the growth curve should be explored. However, for model i) which excludes the parasite data from the likelihood, the the average $t_{0,j,y}$ for the von Bertalanffy growth curves for the west component is 0.18 (approx. 2 months after 1 November) with a maximum difference in $t_{0,j,y}$ between cohorts of approximately 6 months, while the average $t_{0,j,y}$ for the south component corresponds with 1 October. These ranges correspond closer to that expected a priori, i.e. a birth date close to 1 November for the west component and earlier (due to winter spawning) for the south component. Figure 12 shows the length-at-age distributions for one example year.

The model estimated annual proportion of west component sardine infected by the "tetracotyle" type digenean endoparasite is shown in Figure 13. The model fit to the parasite prevalence data between 2010 and 2016 is relatively unchanged from that of de Moor (2020a), but is updated in 2018 to 2020 with the addition of new data (Figure 14a). The change in prevalence-at-length in

2019 and 2020, with the inclusion of samples from commercial catches between September and November can be seen by comparing Figures 14a and 14b.

The model estimated November recruitment is plotted against spawner biomass and effective spawner biomass in Figure 15, indicating the low west component [effective] spawner biomass and recruitment in recent years. The west component effective spawner biomass includes 8% of the south component spawner biomass. The proportion of this effective west component spawner biomass that consists of south component spawner biomass is estimated to be high after the turn of the century and again in recent years, reaching 57% in 2019 (Figure 16). This reflects the low current west component spawner biomass.

The updated and extended time series of data has some correction to the de Moor (2020a) estimate of November biomass in 2019 (Figures 1 and 17) and west component recruitment in May/June 2019 (November 2018) (Figures 2 and 16). However, this correction is not as extensive as that noted a year ago (de Moor 2020a). This model option iv) estimates the effective west and south component spawner biomasses to be 39 000t and 168 000t in November 2020, respectively, while model option i) estimates the effective west and south component spawner biomasses to be 29 000t and 186 000t.

Figure 18 shows the historical harvest proportion on the sardine, which has often been substantially higher on the west component than on the south component.

In Summary

This document has presented an updated assessment of South African sardine. For many parameters and model outputs, only results at the joint posterior mode have been provided thus far, and readers are reminded that there is estimation error about these values.

The assessment model i) estimates the west component biomass in November 2020 to be about 79 000t - 18% of the historical (1984-2019) average. The west component spawner biomass is estimated to be about 13 000t (24% of the historical average) and the effective spawner biomass is estimated to be 29 000t (42% of the historical average). This is a lower biomass, but higher spawner biomass than estimated last year (de Moor 2020a). The corresponding values from assessment model iv) are 187 500t (44% of historical biomass), 24 000t (43% of historical spawner biomass) and 39 000t (53% of historical effective spawner biomass) which are all higher than that estimated for November 2019 by de Moor (2020a).

The assessment model iv) estimates the south component biomass in November 2020 to be about 724 000t which is above the historical average and the south component spawner biomass and effective spawner biomass are estimated to be at 99% of the historical average at 202 000t and 186 000t, respectively. The assessment model iv) estimates the south component biomass in November 2020 to be about 483 500t which is 90% of the historical average and the south component spawner biomass and effective spawner biomass are estimated to be at 88% of the historical average at 182 000t and 167 500t, respectively.

It is expected that a range of sensitivity testing as well as alternative hypothesis will be considered prior to a baseline(s) being selected for the Operating Models that will be used to simulation test OMP-22. However, this initial updated assessment is presented now to be used to provide 'ad hoc' management advice for sardine (in the same manner as that followed the previous two years), given that the OMP is not being used following the declaration of Exceptional Circumstances for sardine.

References

de Moor CL. 2020a. South African sardine assessment posterior distributions and sensitivity tests. DEFF: Branch Fisheries Document FISHERIES/2020/DEC/SWG-PEL/138.

de Moor CL. 2020b. Final baseline assessment of the South African sardine resource using data from 1984-2018. DEA: Branch Fisheries Document FISHERIES/2020/MAR/SWG-PEL/28.

de Moor CL and Butterworth DS. 2016. Assessment of the South African sardine resource using data from 1984-2015: Results at the joint posterior mode for the two mixing-stock hypothesis. DAFF Branch Fisheries Document: FISHERIES/2016/JUL/SWG-PEL/22REV2.

de Moor CL, Butterworth DS and van der Lingen CD. 2017. The quantitative use of parasite data in multistock modelling of South African sardine (Sardinops sagax). *Can. J. Fish. Aquat. Sci*. 74:1895-1903.

de Moor CL, Merkle D, Coetzee J and van der Lingen C. 2021. The data used in the 2021 sardine assessment. DEFF: Branch Fisheries Document FISHEREIS/2021/APR/SWG-PEL/22.

OLSPS 2020. Updated length weight regression parameters for sardine. DEA: Branch Fisheries Document FISHERIES/2020/MAR/SWG-PEL/22REV.

van der Lingen CD. 2021. Comparing prevalence-at-length plots for western and southern sardine from Pelagic Biomass Survey and commercial catch samples, 2010-2018. DEFF: Branch Fisheries Document FISHERIES/2021/MAR/SWG-PEL/15.

van der Lingen CD, Fréon P, Fairweather TP, van der Westhuizen JJ. 2006. Density-dependent changes in reproductive parameters and condition of southern Benguela sardine *Sardinops sagax*. African Journal of Marine Science 28(3&4): 625-636.

Table 1. The contributions to the objective function from likelihood and prior components, together with the associated estimated survey bias parameters, $k_{j,N}^{S}$ and $k_{j,r}^{S}$, model predicted biomass in 2018-2020, $B_{j,y}^{S}$, and model predicted spawner biomass in 2018-2020, $B_{j,y}^{sp,S}$. The Hessian-based CV is included in parentheses for $B_{w,2020}^{S}$.

Option	Prevalence Included In Likelihood																	
	2010-2018 from surveys	2019 from survey	2020 from survey	2019 from commercial	2020 from commercial	Obj fn	$-lnL$	$-lnL^{Nov}$	$-lnL^{rec}$	$-lnL^{com\ propl}$	$-lnL^{sur\ propl}$	$-lnL^{prev}$	$ln(k_{ac}^{S})$	$move_{y,1}$	η_{y}^{t}	$\bar{l}_{1,y}$	$k_{j,N}^{S}$	$k_{w,r}^{S}$
i)	x	x	x	x	x	-705.30	-771.83	56.32	38.56	-458.39	-408.32	3533.94	-1.401	-32.74	-20.47	121.04	0.75	0.48
ii)	Y	x	x	x	x	978.91	904.59	62.54	41.18	-450.88	-398.61	1650.37	-1.33	-32.06	-13.44	120.81	0.77	0.58
iii)	Y	Y	x	x	x	1125.42	1051.58	63.00	40.49	-449.38	-393.88	1791.34	-1.28	-31.80	-14.22	120.76	0.77	0.62
iv)	Y	Y	Y	x	x	1211.03	1136.88	63.08	40.62	-449.34	-394.47	1876.99	-1.25	-31.78	-13.87	120.76	0.78	0.62
v)	Y	Y	x	Y	x	1126.73	1052.16	63.06	40.41	-449.20	-394.45	1792.33	-1.29	-31.80	-13.47	120.74	0.77	0.63
vi)	Y	Y	Y	Y	x	1213.01	1138.61	63.10	40.74	-429.34	-395.78	1879.89	-1.23	-32.09	-13.93	120.75	0.78	0.63
vii)	Y	Y	Y	Y	Y	1363.32	1288.04	63.85	40.38	-449.08	-393.35	2026.24	-1.22	-31.72	-12.88	120.74	0.78	0.63
	2010-2018 from surveys	2019 from survey	2020 from survey	2019 from commercial	2020 from commercial	$B_{w,2018}^{S}$	$B_{w,2019}^{S}$	$B_{w,2020}^{S}$	$B_{w,2018}^{sp,S}$	$B_{w,2019}^{sp,S}$	$B_{w,2020}^{sp,S}$	$B_{s,2018}^{S}$	$B_{s,2019}^{S}$	$B_{s,2020}^{S}$	$B_{s,2018}^{sp,S}$	$B_{s,2019}^{sp,S}$	$B_{s,2020}^{sp,S}$	
i)	x	x	x	x	x	35.6	45.3	79.1 (0.92)	8.5	6.8	9.4	308.8	450.2	723.8	132.9	135.9	213.6	
ii)	Y	x	x	x	x	51.5	64.2	73.3 (0.87)	6.5	1.1	5.3	254.2	434.6	715.9	132.9	123.0	285.7	
iii)	Y	Y	x	x	x	36.6	149.7	92.8*	6.6	6.8	14.1	230.5	279.1	590.1	121.8	106.3	209.1	
iv)	Y	Y	Y	x	x	31.5	113.1	187.6 (0.56)	7.6	6.6	24.1	237.7	271.6	483.5	122.9	110.4	182.2	
v)	Y	Y	x	Y	x	35.7	155.2	90.8 (0.86)	6.6	6.3	12.9	248.5	280.5	588.1	122.6	116.4	202.5	
vi)	Y	Y	Y	Y	x	29.9	119.10	192.2 (0.56)	7.7	6.0	23.8	252.7	277.4	485.9	123.4	119.7	179.5	
vii)	Y	Y	Y	Y	Y	28.5	122.6	215.9 (0.46)	8.5	5.3	23.2	250.7	258.7	434.4	121.3	117.9	166.1	

*Non positive definite Hessian

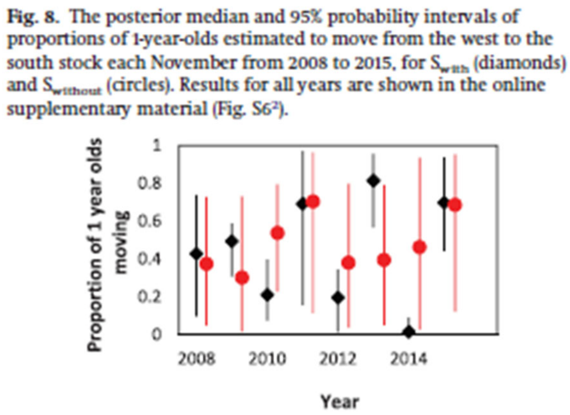

Figure 1. Figure 8 of de Moor et al. (2017) showing the improvement in the ability for the model to estimate annual proportions moving given the inclusion of parasite prevalence-at-length data in the likelihood.

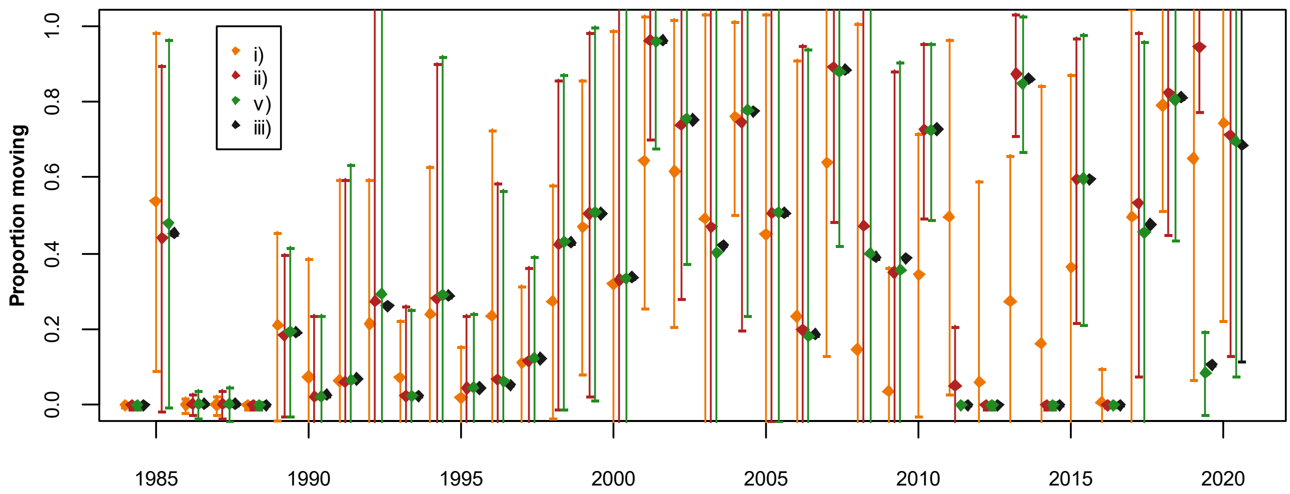

Figure 2a. The model estimated annual proportion (with 95% Hessian-based confidence interval for options i), ii), v)) of west component 1-year olds that move to the south component for model scenarios which exclude prevalence-at-length data in 2020.

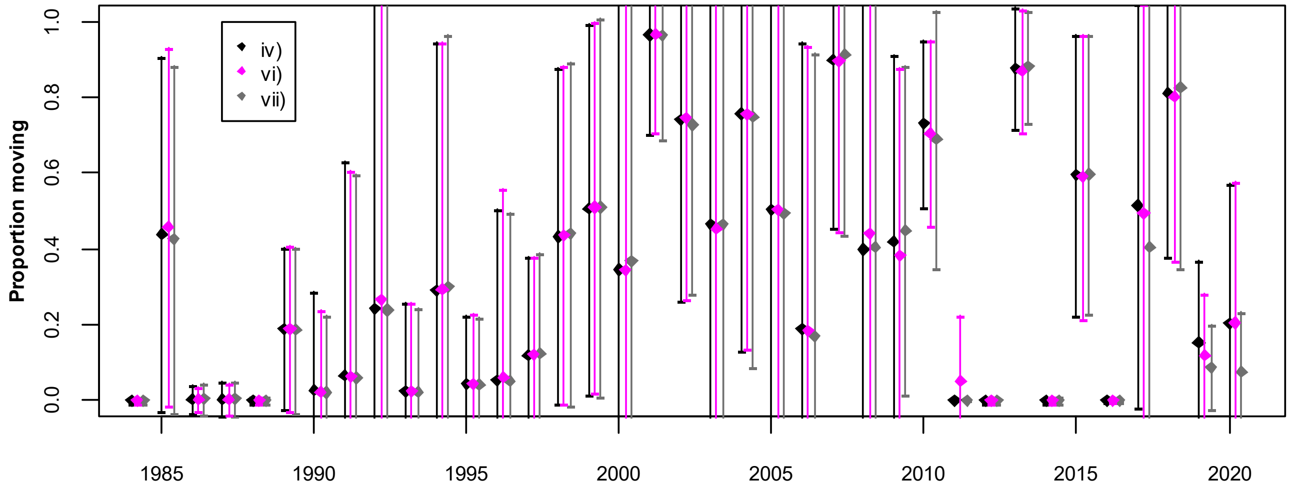

Figure 2b. The model estimated annual proportion (with 95% Hessian-based confidence interval) of west component 1-year olds that move to the south component for model scenarios which include prevalence-at-length data in 2020.

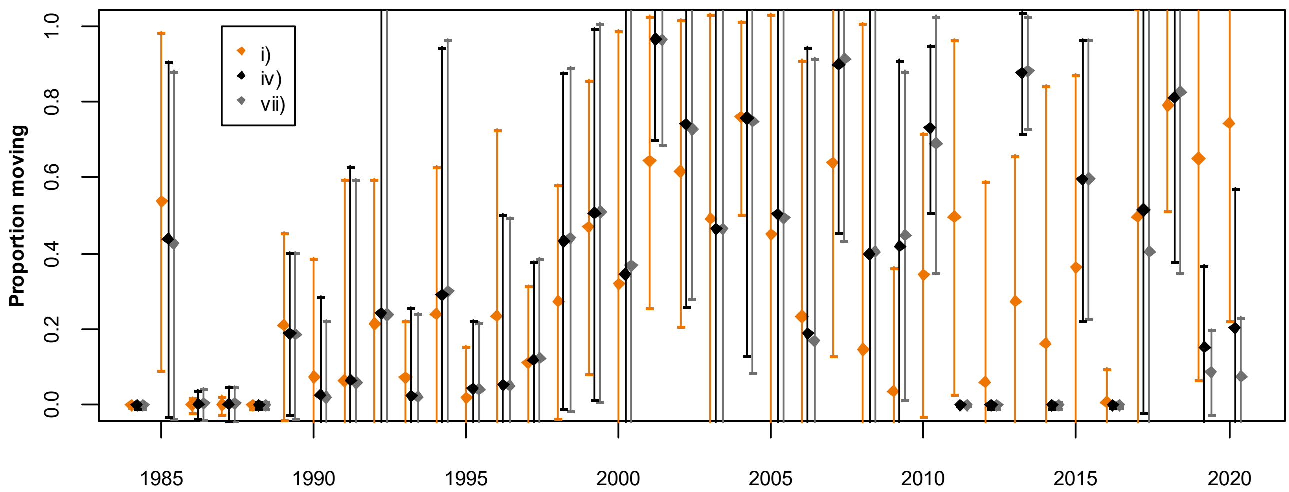

Figure 2c. The model estimated annual proportion (with 95% Hessian-based confidence interval) of west component 1-year olds that move to the south component for model scenarios which i) exclude all prevalence-at-length data, iv) include all prevalence-at-length data from samples collected from acoustic surveys and vii) additionally include prevalence-at-length data from samples collected from commercial catches in 2019 (south only) and 2020.

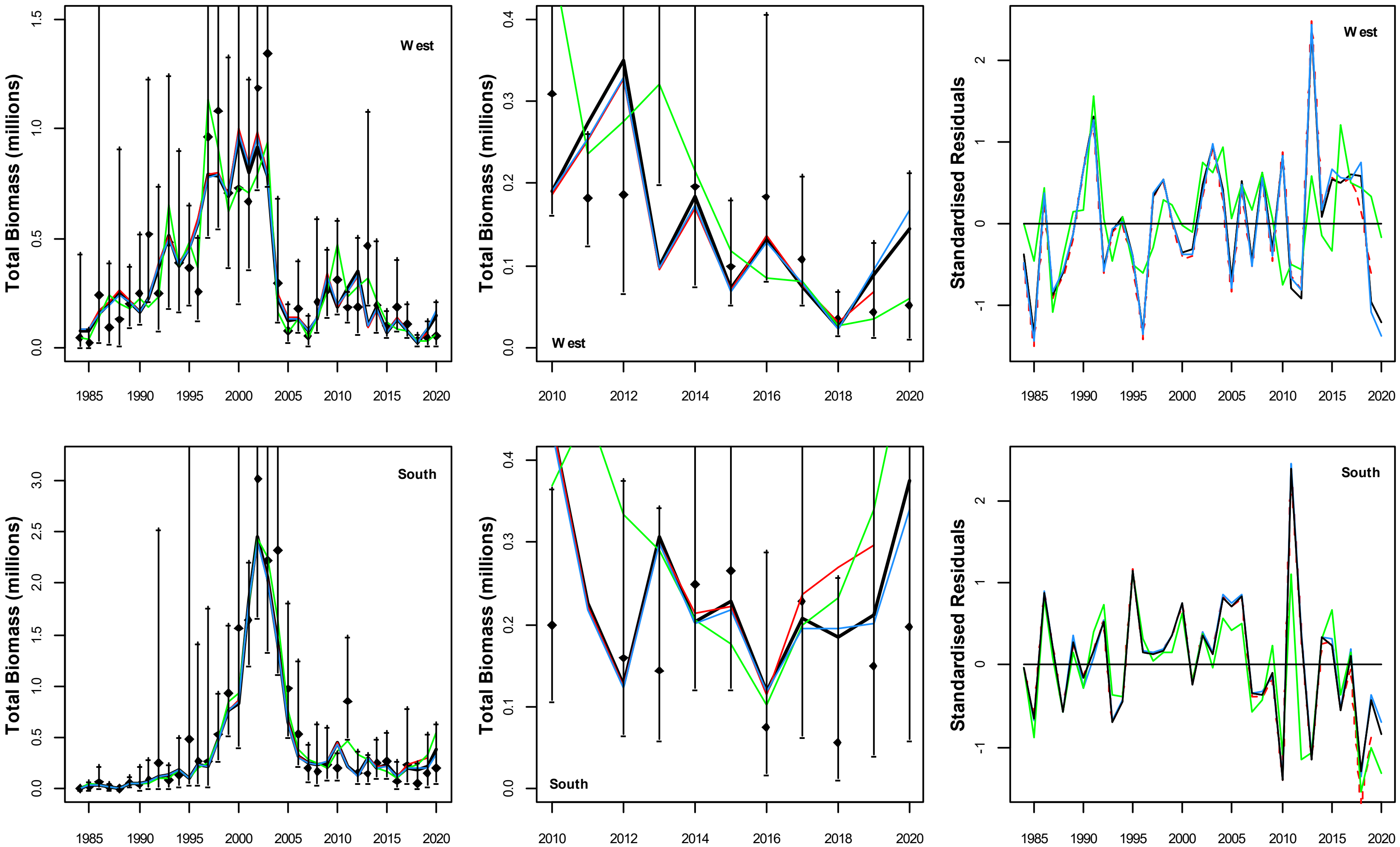

Figure 3. Acoustic survey estimated and model predicted November sardine total biomass from 1984 to 2020 for model option iv) (black) compared to i) (green) and vii) (blue). The observed indices are shown with 95% confidence intervals. The centre plot shows only the most recent 11 years of the left hand plot. The standardised residuals (i.e. the residual divided by the corresponding standard deviation, including additional variance where appropriate) from the fits are given in the right hand plots. The red lines indicate the November biomass predicted by the baseline model of de Moor (2020a) and associated residuals.

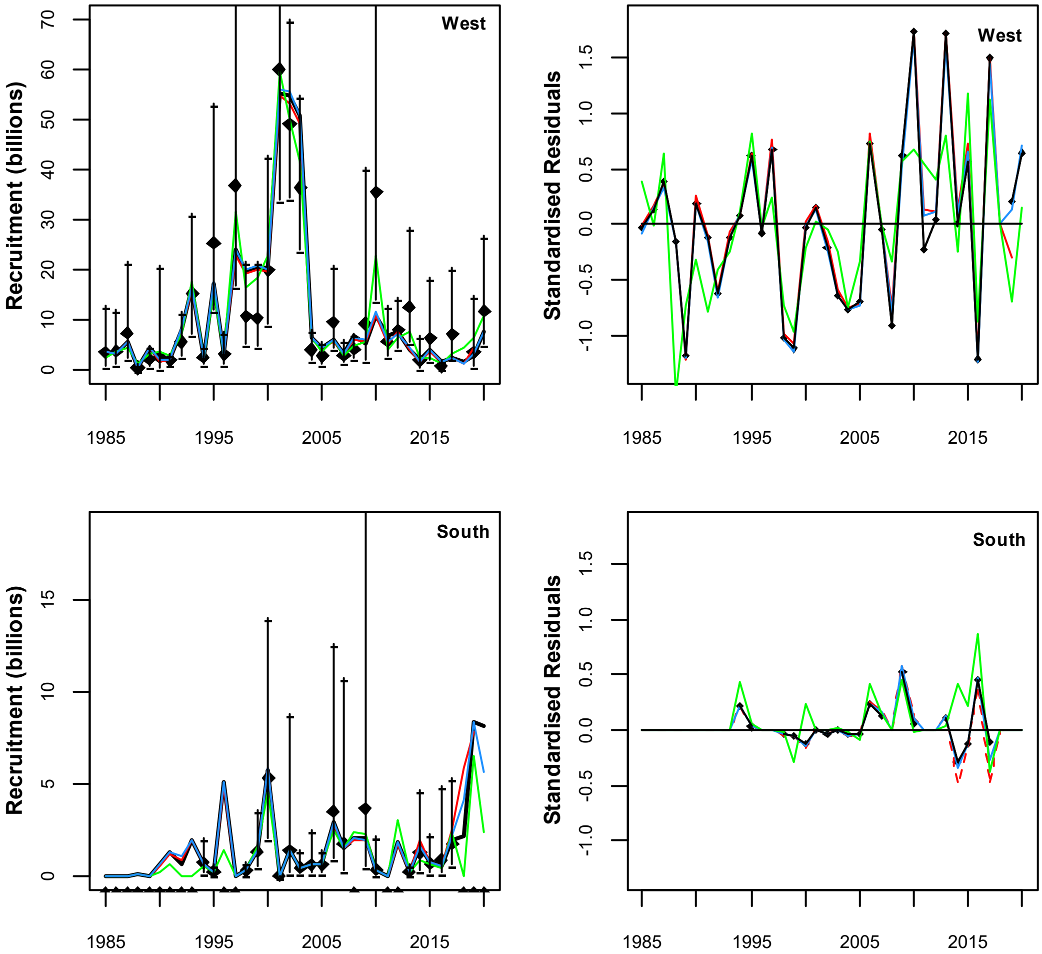

Figure 4. Acoustic survey estimated and model predicted sardine recruitment numbers from May/June 1985 to 2020 for model option iv) (black) compared to i) (green) and vii) (blue). There was no survey observation in 2018; the model predicted value corresponds to the recruitment predicted at 8th June 2018 which is the average start date of the survey from 2016, 2017 and 2019 surveys. The survey indices are shown with 95% confidence intervals. The standardised residuals from the fit are given in the right hand plots. The red lines indicate the May recruitment predicted by baseline model of de Moor (2020a) and associated residuals.

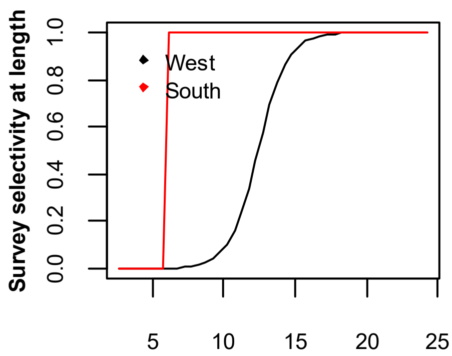

Figure 5. The model estimated November survey selectivity at length for model option iv).

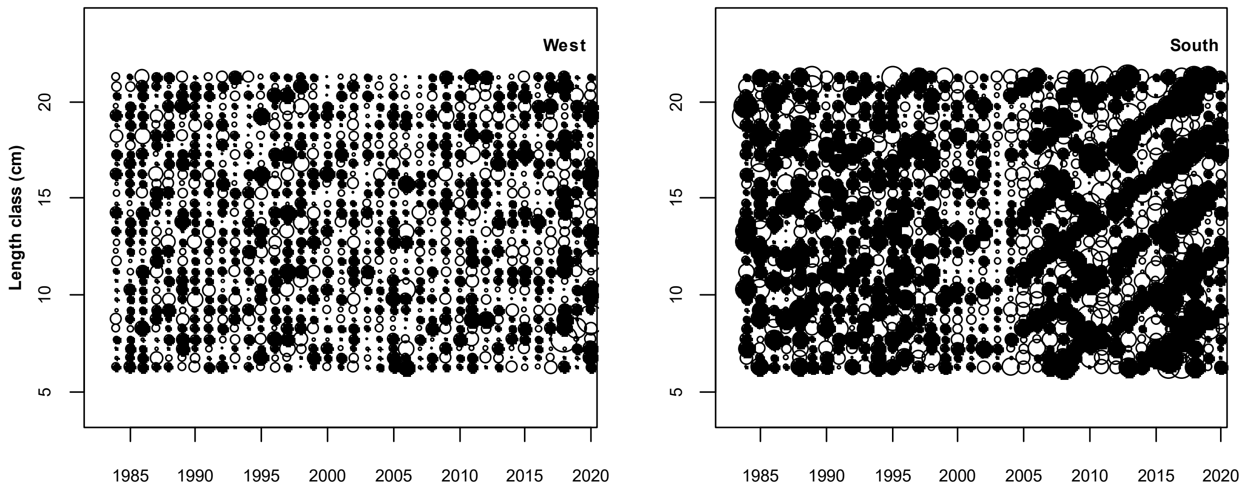

Figure 6. Residuals from the fit of the model predicted proportions-at-length in the November survey to the hydroacoustic survey estimated proportions for model option iv).

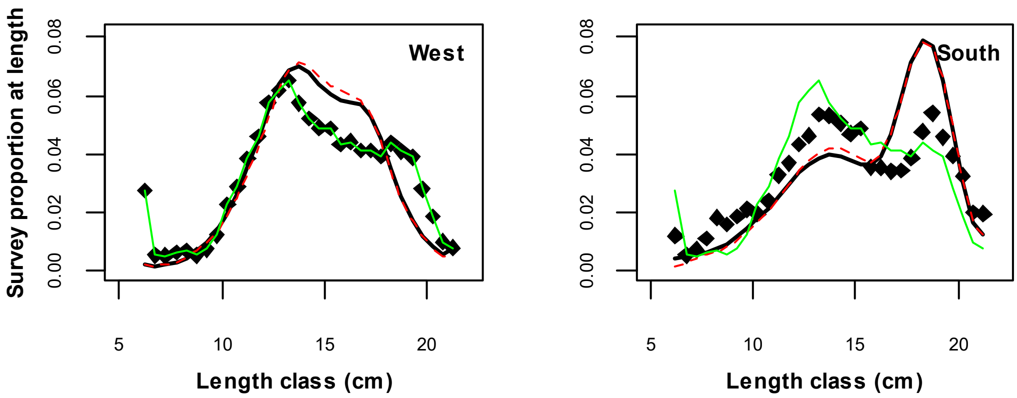

Figure 7. Average (over all years) model predicted and observed proportion-at-length in the November survey for model option iv) (black) compared to i) (green), with the red lines indicating that predicted by baseline model of de Moor (2020a).

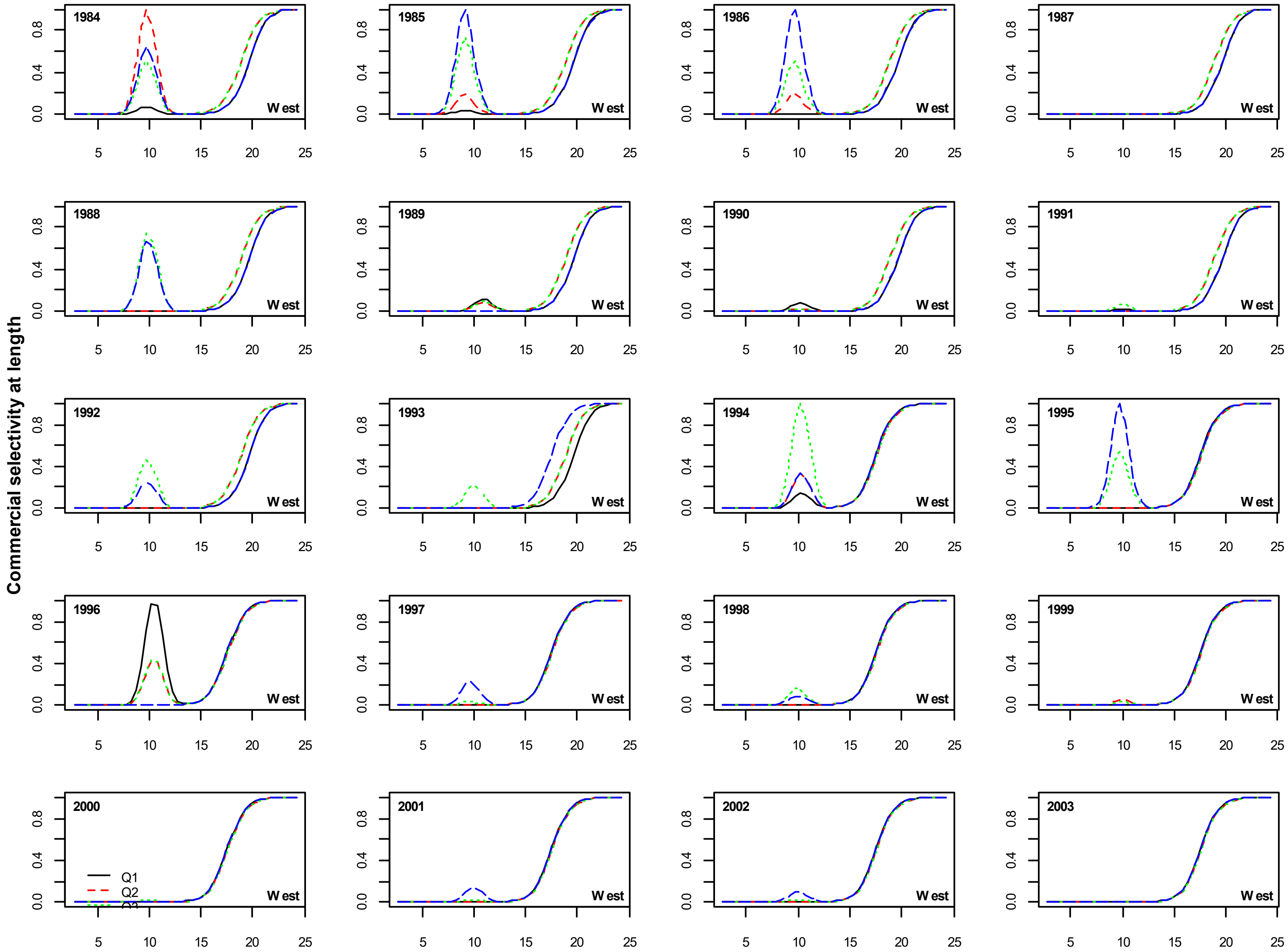

Figure 8. The model estimated commercial selectivity at length for model option iv).

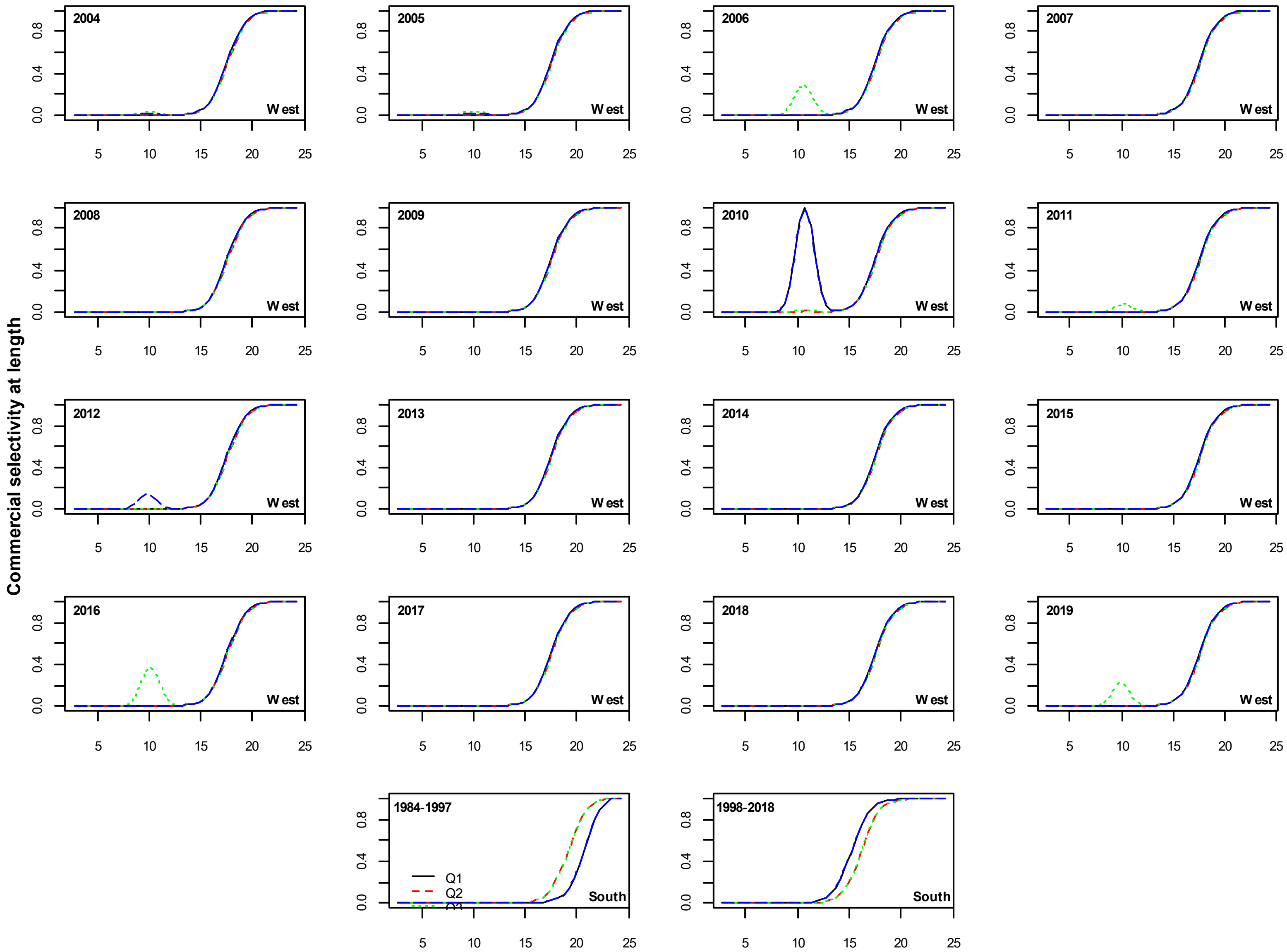

Figure 8 (continued).

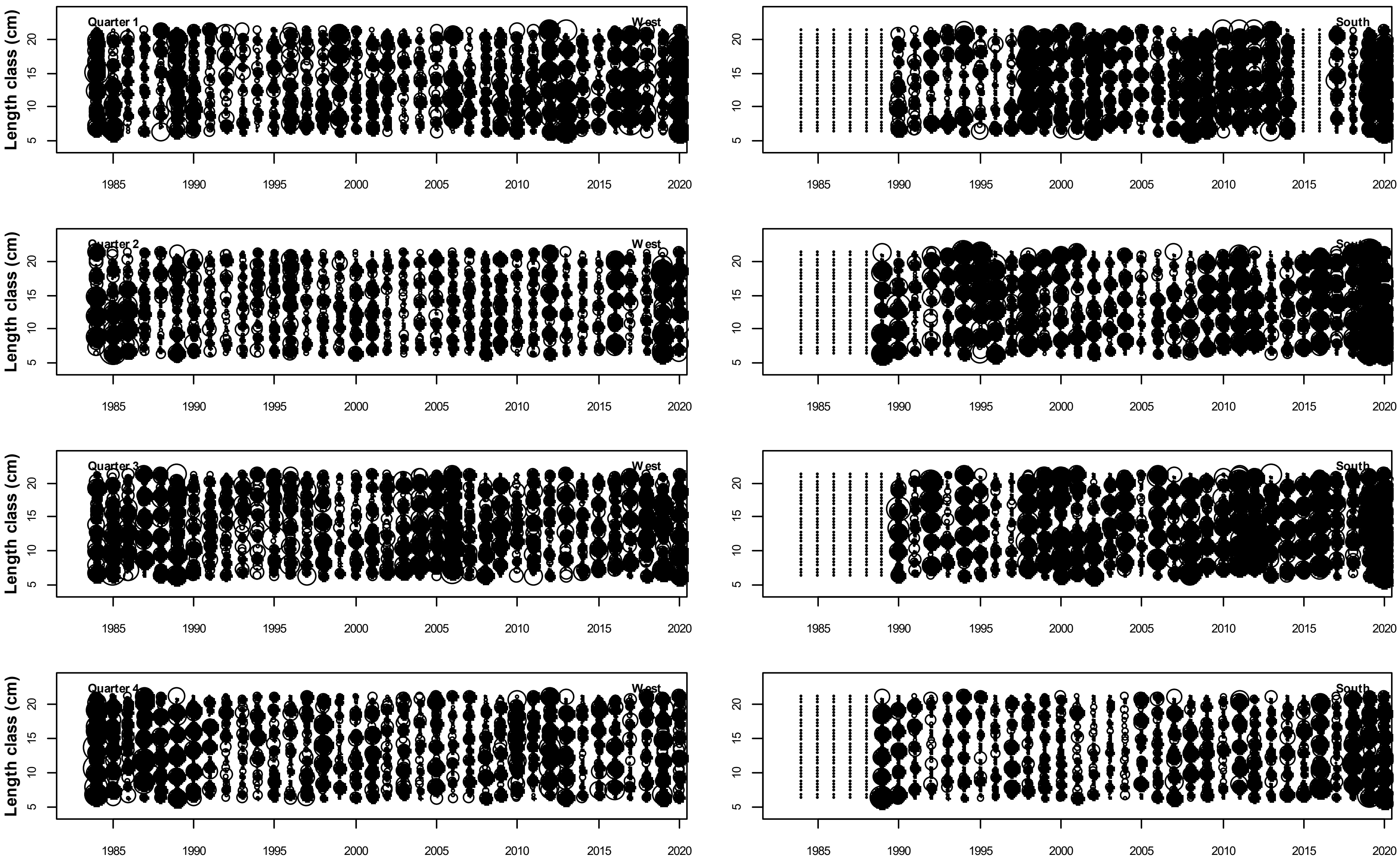

Figure 9. Residuals from the fit of the model predicted proportions-at-length in the quarterly commercial catch to the observed proportions for model option iv).

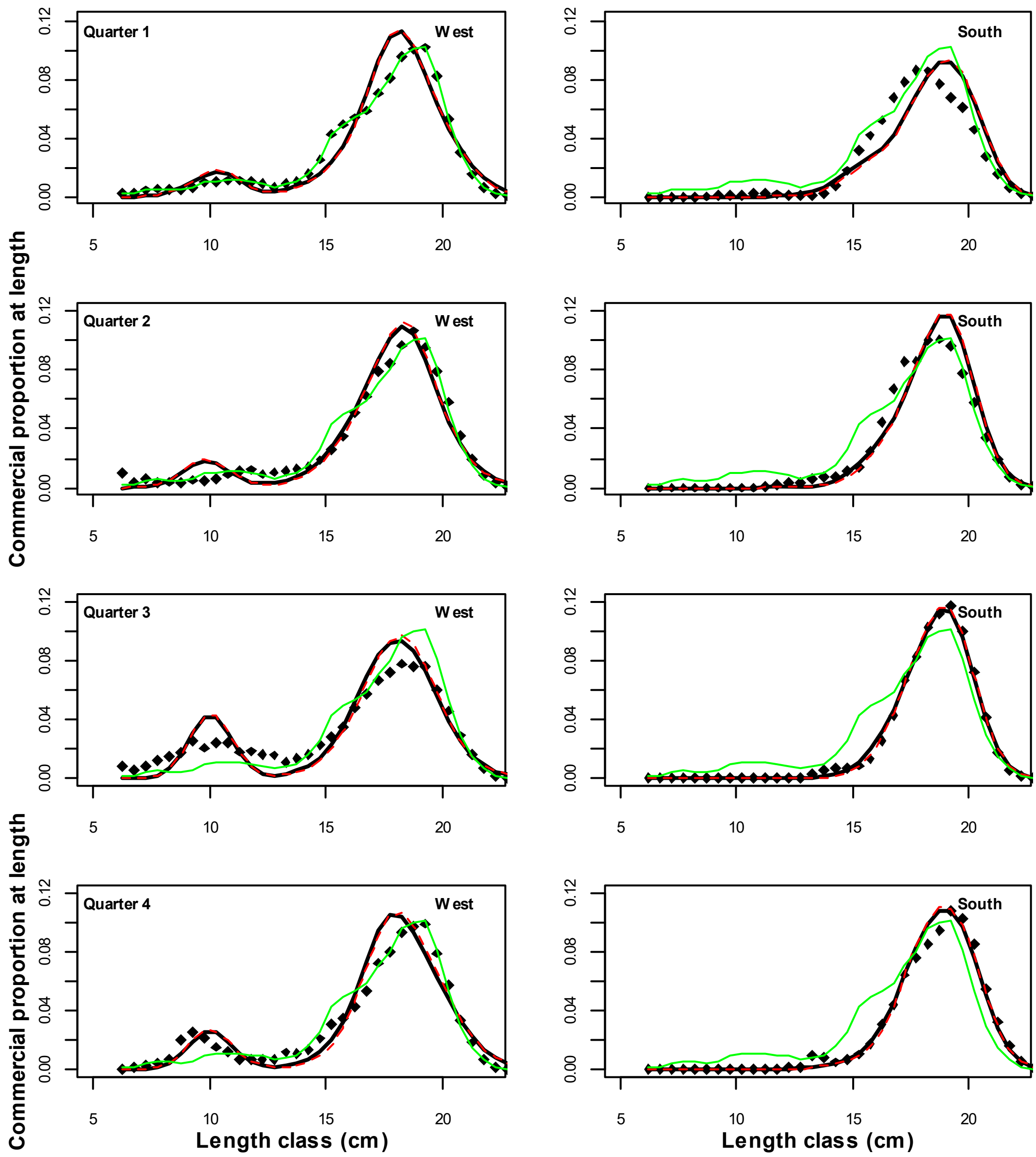

Figure 10. Average (over all years) quarterly model predicted and observed proportion-at-length in the commercial catch for model option iv) (black) compared to i) (green), with the red lines indicating that predicted by baseline model of de Moor (2020a).

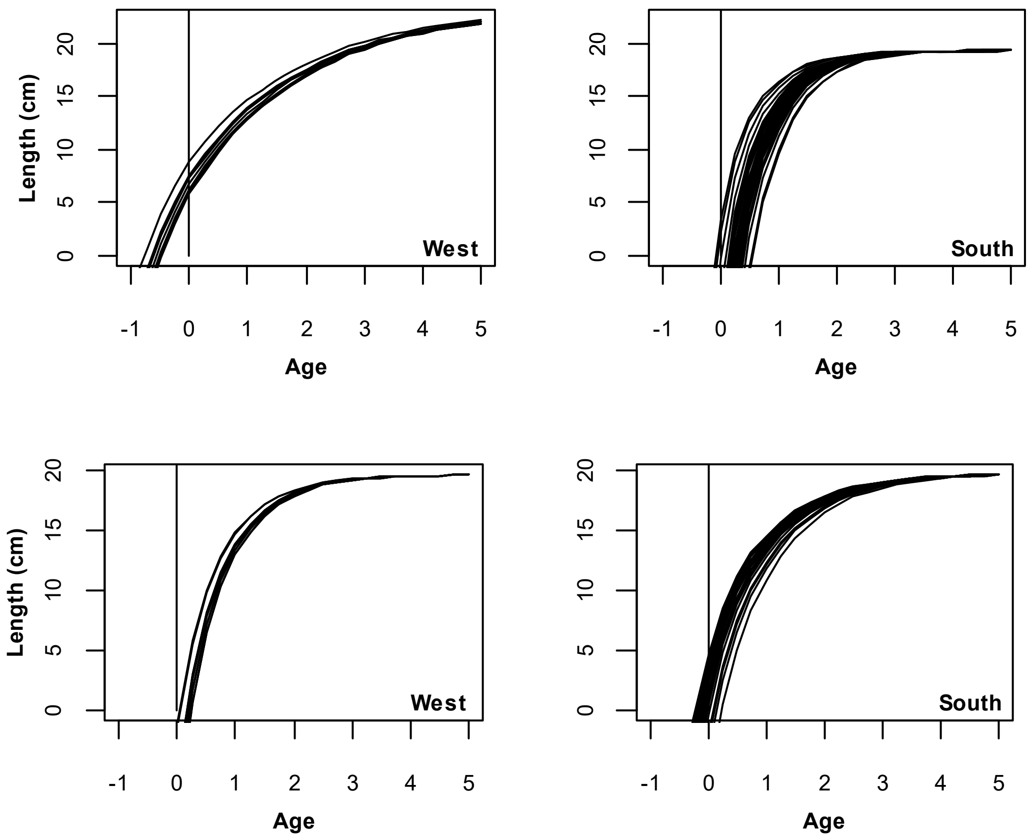

Figure 11. The von Bertalanffy growth curves (by cohort) estimated for model option iv) (above) and model option i) (below) by allowing for auto-correlated residuals for the variation about the age at which length is zero.

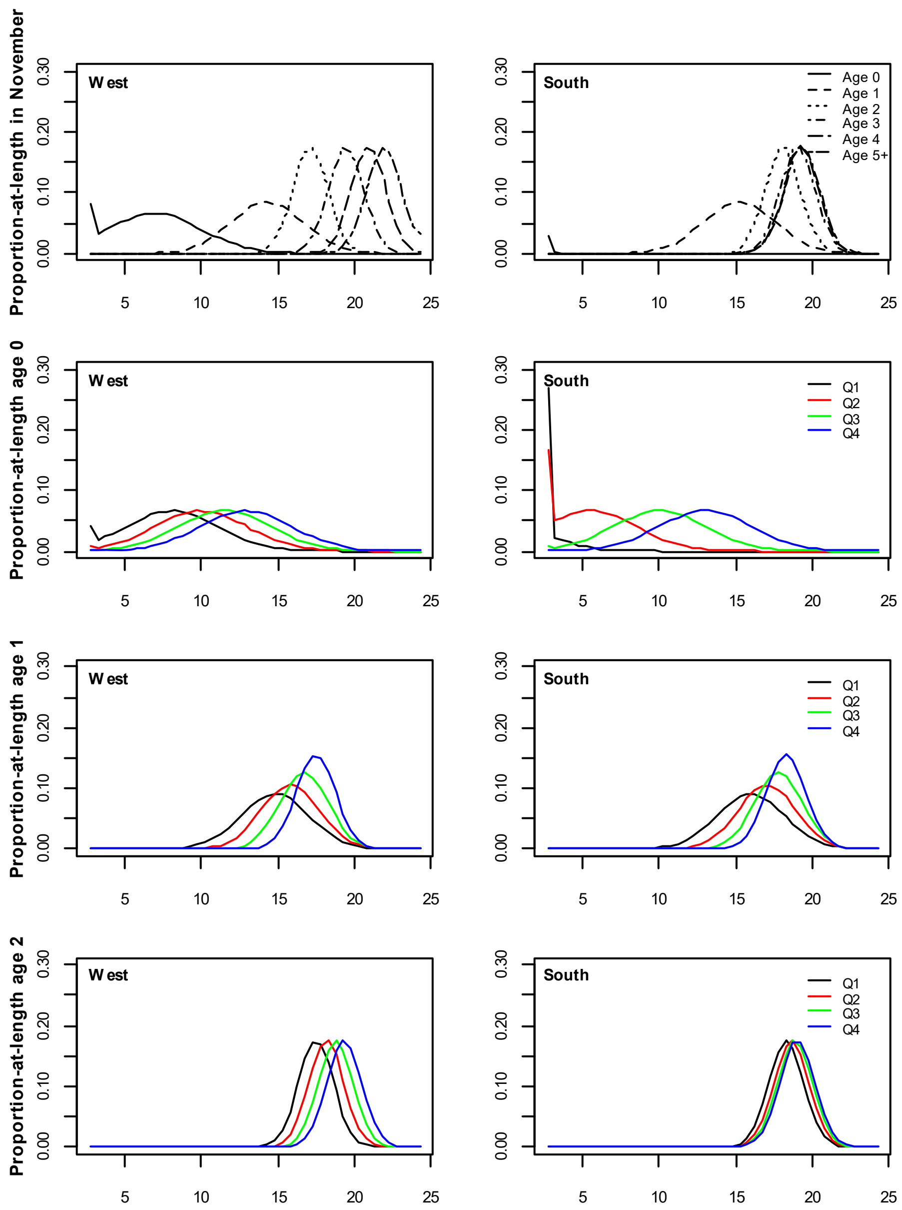

Figure 12. The model option iv) estimated distributions of proportions-at-length for each age **in 2010**, given at the time of the biomass survey (1 November, top row), and middle of each quarter of the year (corresponding to the times commercial catch is modelled to be taken) for age 0, 1 and 2 (subsequent rows).

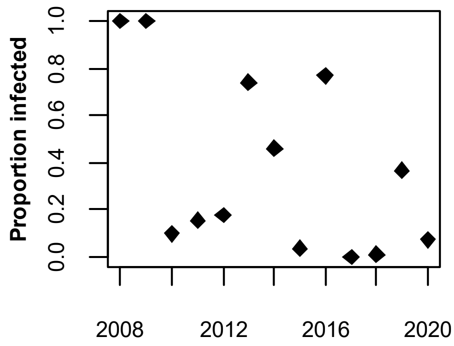

Figure 13. The model option iv) estimated proportion of west component sardine infected with the parasite between 2008 and 2020. (Annual infection rate is arbitrarily assumed to be 0 prior to 2008.)

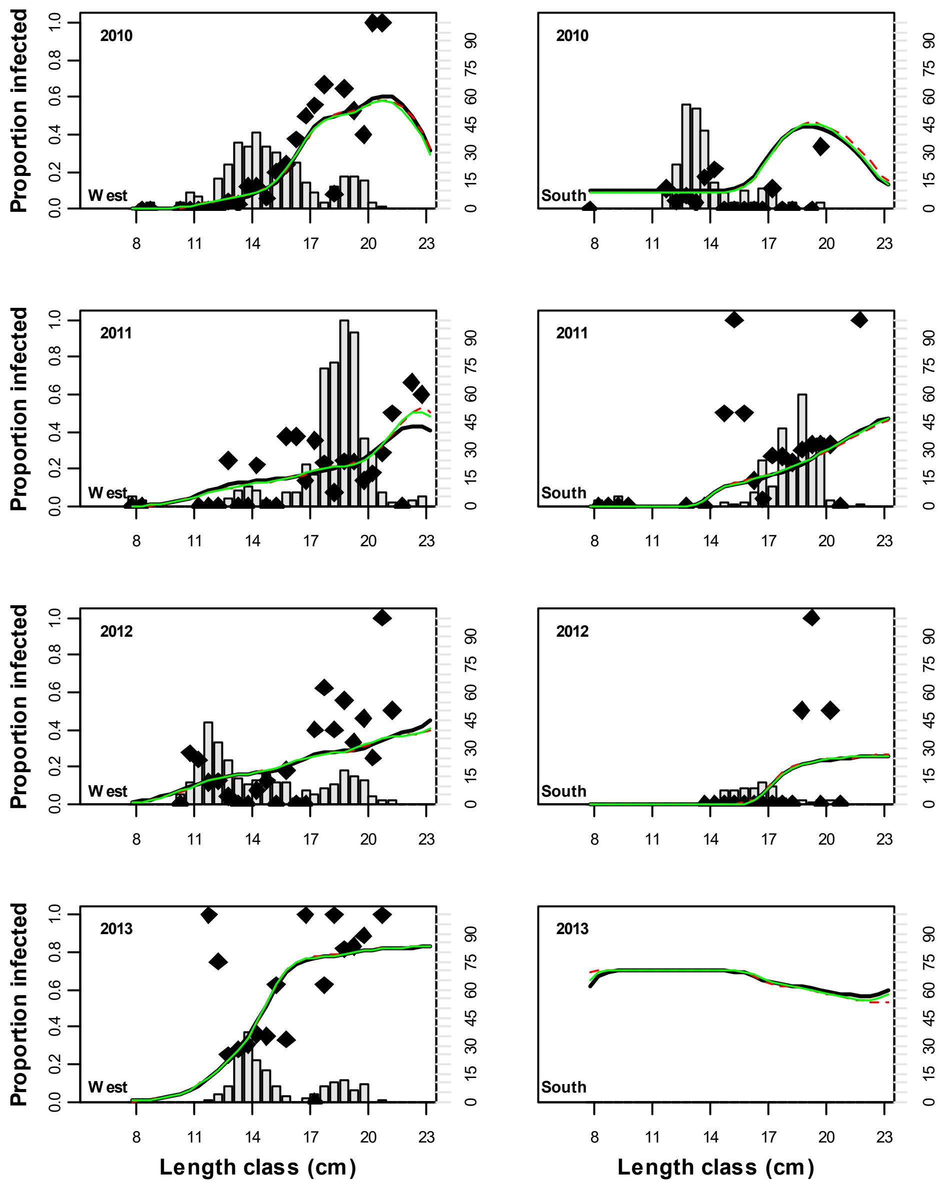

Figure 14a. The model estimated proportions-at-length of west and south stock sardine infected with the parasite (i.e. parasite prevalence-by-length) between 2010 and 2020, together with the observed proportions-at-length. Results are shown for model option iv) (black) compared to i) (green), with the red lines indicating the proportions predicted by

baseline model of de Moor (2020a). The sample size for each length class is given by the grey bars, plotted against the right vertical axis.

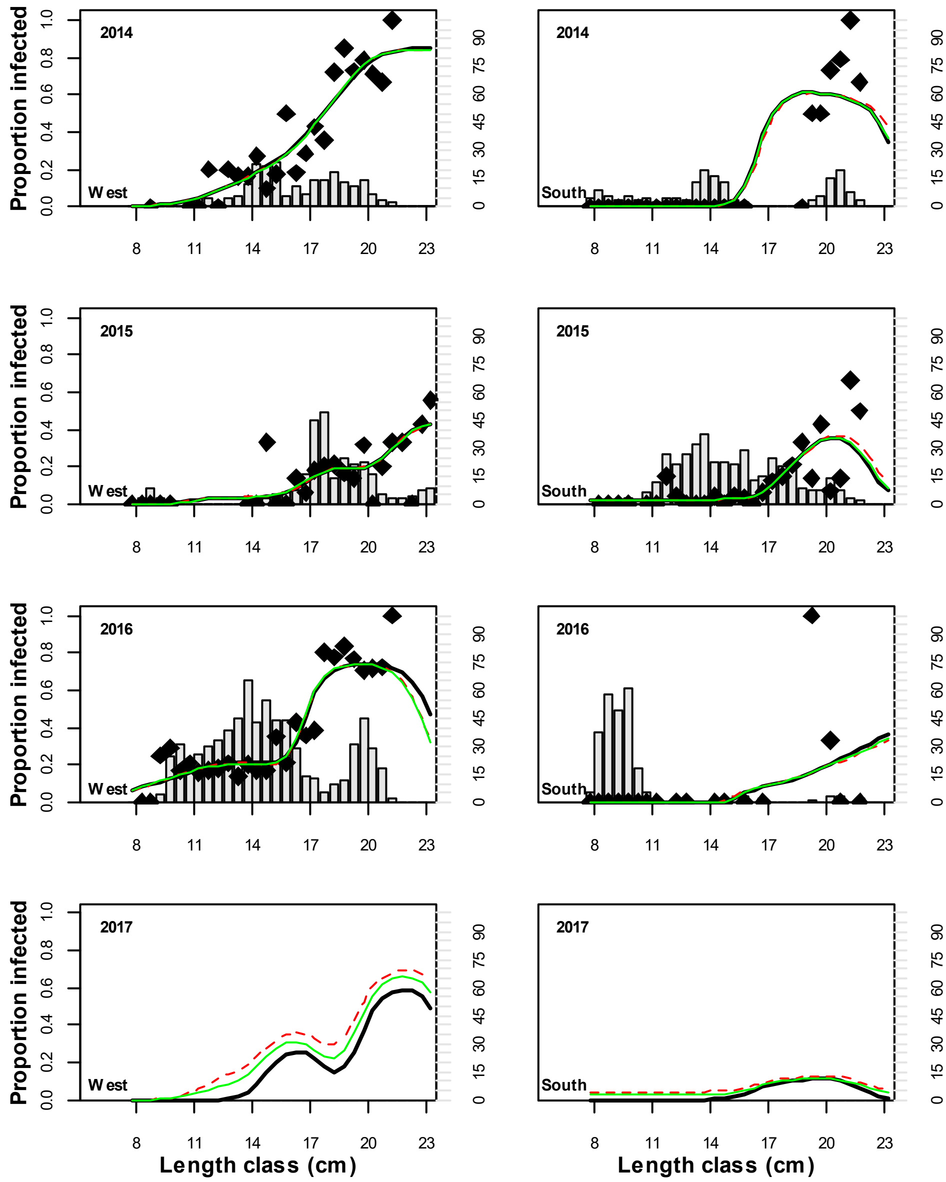

Figure 14a (continued).

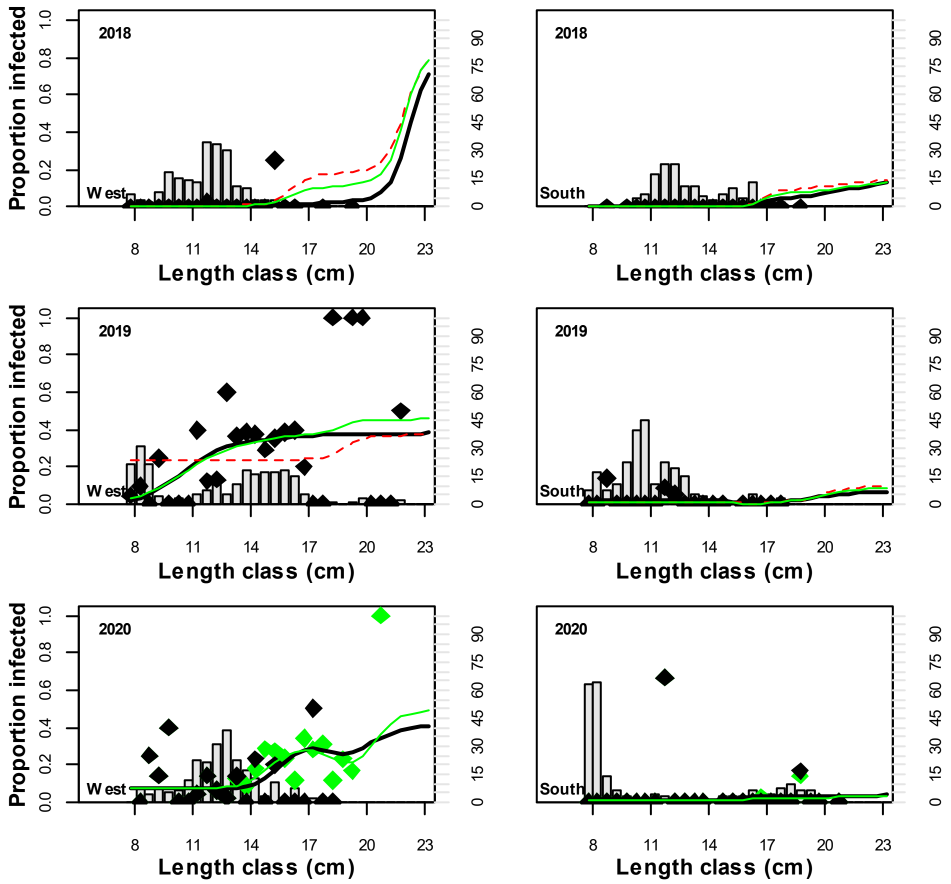

Figure 14a (continued).

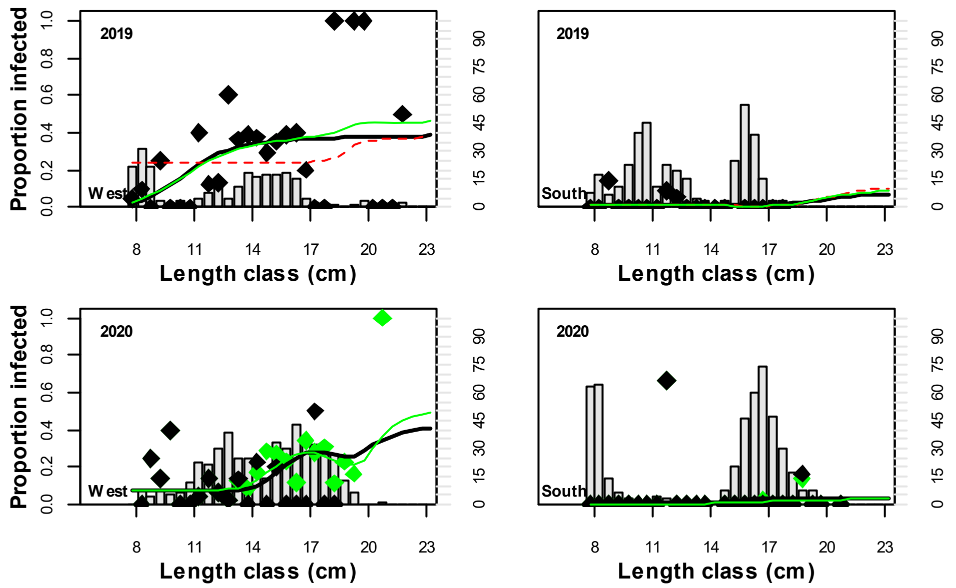

Figure 14b. As for Figure 15a for 2019 and 2020, but including the commercial samples in the observed proportions-at-length on the west (2020) and south (2019 and 2020) coasts.

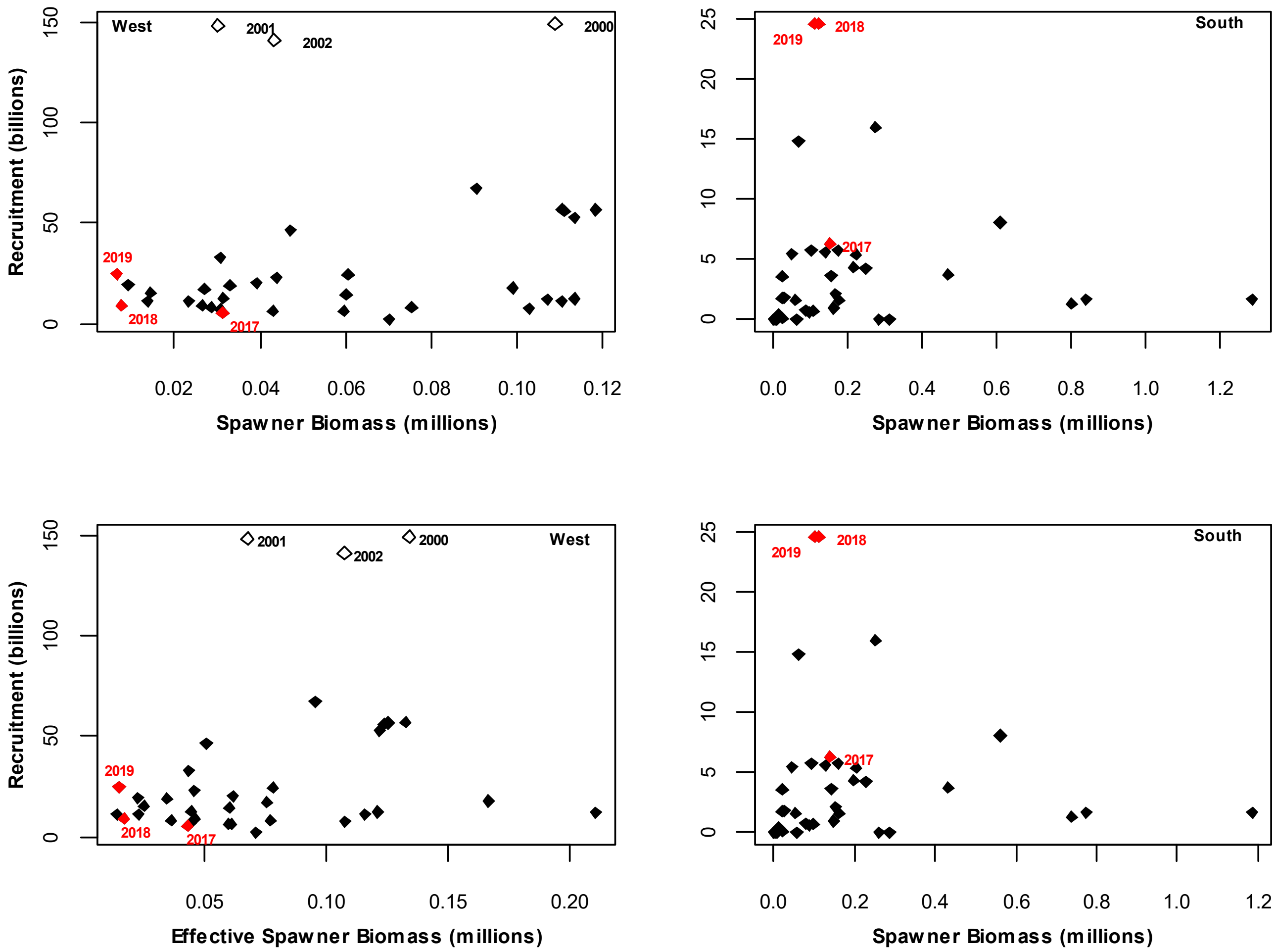

Figure 15. Model option iv) predicted sardine recruitment (in November) plotted against spawner biomass (top panels) and effective spawner biomass (lower panels) from November 1984 to November 2019. The open diamonds indicate the years of peak west component recruitment.

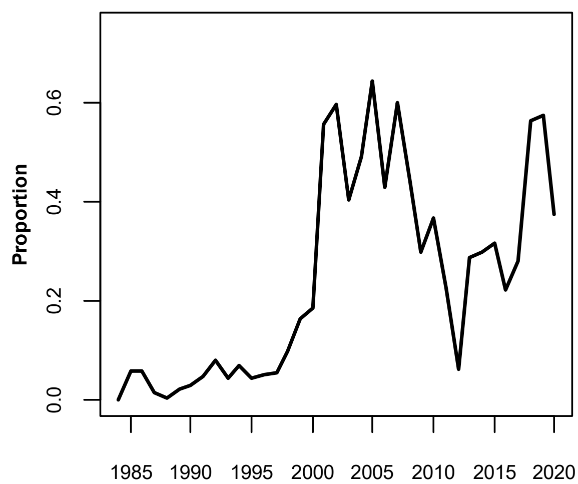

Figure 16. The proportion of west component effective spawner biomass (defined as west component spawner biomass combined with 8% of south component spawner biomass) that consists of south component spawner biomass (i.e. $SSB^{S}_{j=s,y}/SSB^{eff,S}_{j=W,y}$).

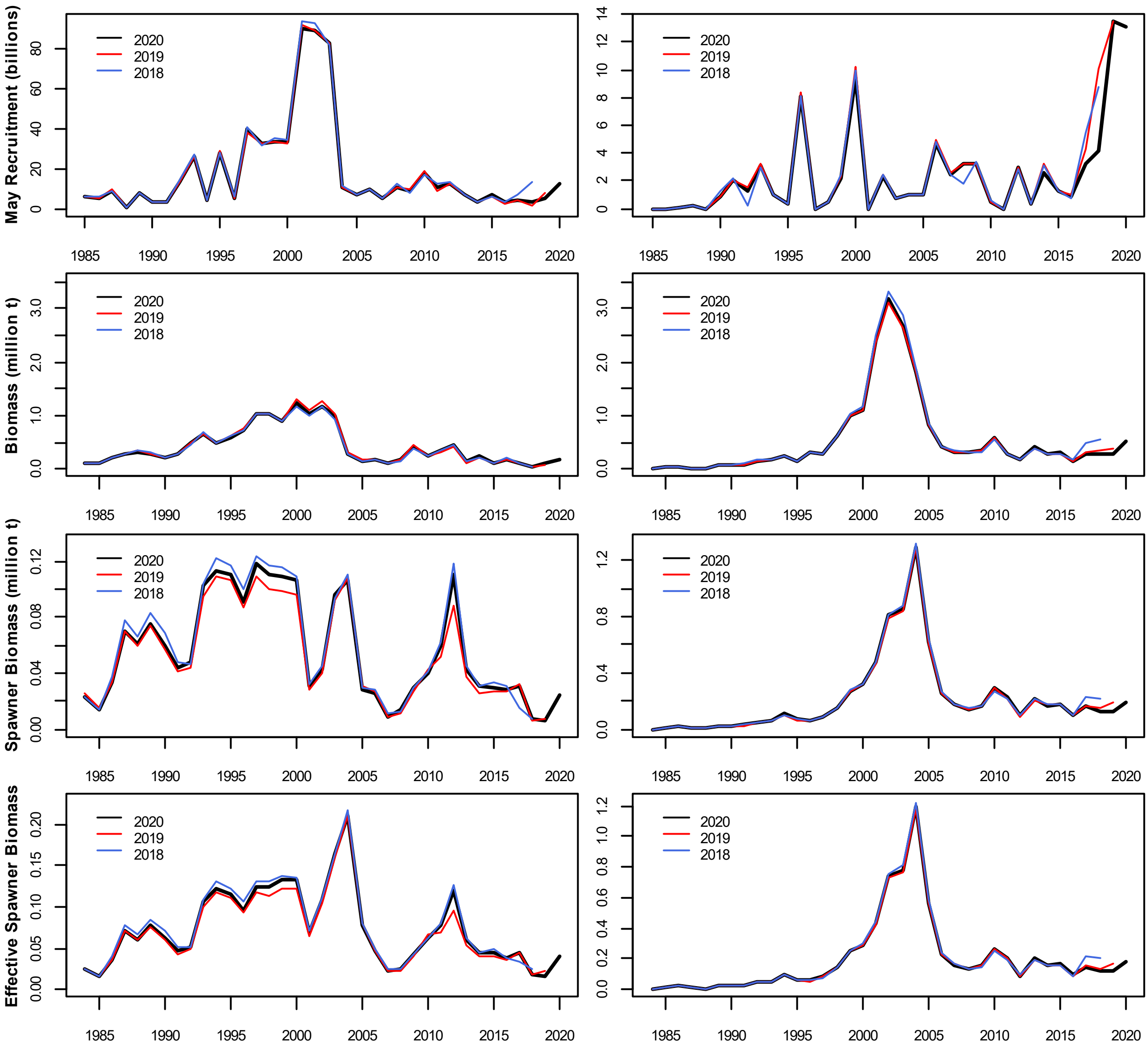

Figure 17. Model predicted May recruitment (without correction for survey bias), total biomass (without correction for survey bias), spawner biomass and effective spawner biomass and for the west and south components from model option iv) ("2020") compared to that of de Moor (2020a) ("2019") and de Moor (2020b) ("2018").

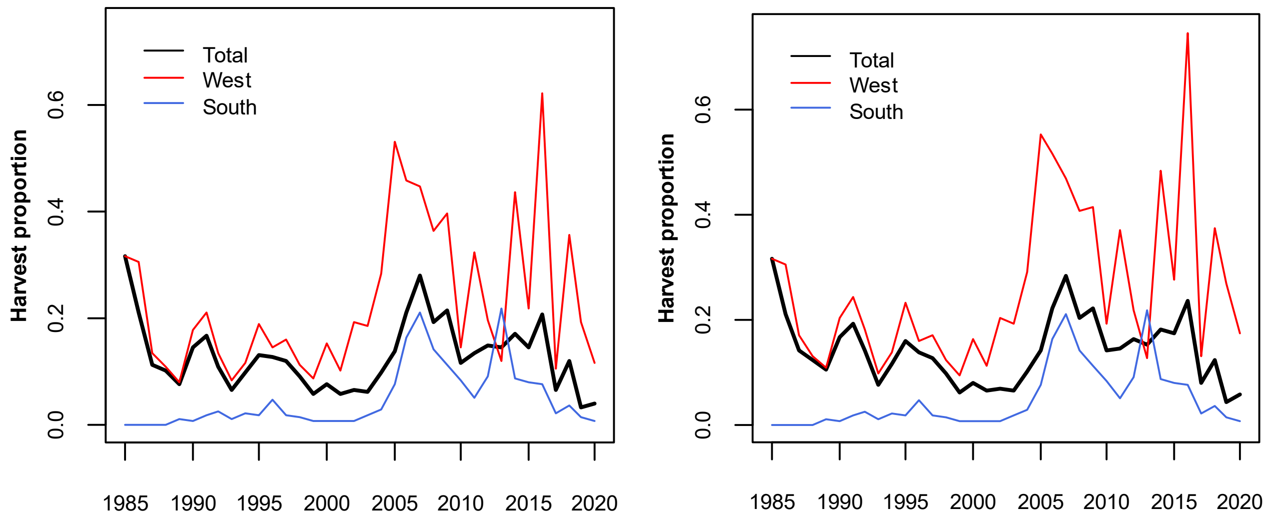

Figure 18. The exploitation rate (simply calculated as the observed annual (Nov-Oct) catch tonnage as a proportion of the model predicted total biomass). The left plot excludes small sardine bycatch with anchovy while the right plot includes these landings.

Appendix A: Bayesian operating model for the South African sardine resource

This assessment provides the generalised operating model for the South African sardine resource (used for this baseline two mixing-component hypothesis as well as a single stock hypothesis[1]). The assessment is run from November $y_1 = 1984$ to November $y_n = 2020$, with the following subscript notation:

- quarters $q = 1$ denoting November $y-1$ to January y, $q = 2$ denoting February to April y, $q = 3$ denoting May to July y and $q = 4$ denoting August to October y;
- ages $a = 0$ to a plus group of $a = 5^+$;
- lengths from a minus group of $l = 2.5^- cm$ to a plus group of $l = 24^+ cm$;
- components $j = W$ or $j = S$ denote the west and south components, respectively, where only the west component equations are used in the single component hypothesis;
- infection $p = NI$ or $p = I$ denote the sardine uninfected and infected with the digenean 'tetracotyle-type' metacercarian endoparasite, respectively.

All parameters are defined in Tables A1 and A2.

Population Dynamics

Numbers-at-age at 1 November before movement or infection

$$N^{S*}_{j,p,y,a} = \left(\left(\left(\left(N^{S}_{j,p,y-1,a-1} e^{-M^S_{y,a-1}/8} - C^S_{j,p,y,1,a-1}\right) e^{-M^S_{y,a-1}/4} - C^S_{j,p,y,2,a-1}\right) e^{-M^S_{y,a-1}/4} - C^S_{j,p,y,3,a-1}\right) e^{-M^S_{y,a-1}/4} - C^S_{j,p,y,4,a-1}\right) e^{-M^S_{y,a-1}/8}$$

$$p = I, NI, y_1 \le y \le y_n, 1 \le a \le 4$$

$$N^{S*}_{j,p,y,5^+} = \left(\left(\left(\left(N^{S}_{j,p,y-1,4} e^{-M^S_{y,4}/8} - C^S_{j,p,y,1,4}\right) e^{-M^S_{y,4}/4} - C^S_{j,p,y,2,4}\right) e^{-M^S_{y,4}/4} - C^S_{j,p,y,3,4}\right) e^{-M^S_{y,4}/4} - C^S_{j,p,y,4,4}\right) e^{-M^S_{y,4}/8} +$$
$$\left(\left(\left(\left(N^{S}_{j,p,y-1,5^+} e^{-M^S_{y,5^+}/8} - C^S_{j,p,y,1,5^+}\right) e^{-M^S_{y,5^+}/4} - C^S_{j,p,y,2,5^+}\right) e^{-M^S_{y,5^+}/4} - C^S_{j,p,y,3,5^+}\right) e^{-M^S_{y,5^+}/4} - C^S_{j,p,y,4,5^+}\right) e^{-M^S_{y,5^+}/8}$$

$$p = I, NI, y_1 \le y \le y_n \qquad \text{(A1)}$$

Infection of west component sardine in the two mixing-component hypothesis; in the single component hypothesis $I_y = 0$ as the parasite data have no influence so that they are not included in the likelihood

$$\mathrm{N}^{\mathrm{S**}}_{\mathrm{W,NI,y,a}} = (1 - \mathrm{I_y})\mathrm{N}^{\mathrm{S*}}_{\mathrm{W,NI,y,a}} \qquad y_1 \le y \le y_n, 1 \le a \le 5^+$$
$$\mathrm{N}^{\mathrm{S**}}_{\mathrm{W,I,y,a}} = \mathrm{N}^{\mathrm{S*}}_{\mathrm{W,I,y,a}} + \mathrm{I_y}\mathrm{N}^{\mathrm{S*}}_{\mathrm{W,NI,y,a}} \qquad y_1 \le y \le y_n, 1 \le a \le 5^+$$
$$\mathrm{N}^{\mathrm{S**}}_{\mathrm{S,p,y,a}} = \mathrm{N}^{\mathrm{S*}}_{\mathrm{S,p,y,a}} \qquad p = I, NI, y_1 \le y \le y_n, 1 \le a \le 5^+ \qquad \text{(A2)}$$

Movement of west component ($j = W$) sardine to the south component ($j = S$) in the two mixing-component hypothesis; in the single component hypothesis $move_{y,a} = 0$

$$\mathrm{N}^{\mathrm{S}}_{\mathrm{W,p,y,a}} = (1 - \mathrm{move_{y,a}})\mathrm{N}^{\mathrm{S**}}_{\mathrm{W,p,y,a}} \qquad p = I, NI, y_1 \le y \le y_n, 1 \le a \le 5^+$$

[1] For the single stock hypothesis, both abundance indices and proportion-at-length data are combined for the full area and parasite prevalence-by-length is excluded from the likelihood.

$\mathrm{N^{S}_{S,p,y,a}} = \mathrm{N^{S**}_{S,p,y,a}} + \mathrm{move_{y,a}} \mathrm{N^{S**}_{W,p,y,a}}$ $\quad p = I, NI, y_1 \le y \le y_n,\ 1 \le a \le 5^+$ (A3)

Numbers-at-age mid-way through each quarter (for use in catch equations)

$N^S_{j,p,y,1,a} = N^S_{j,p,y-1,a} e^{-M^S_{y,a}/8}$ $\quad p = I, NI, y_1 \le y \le y_n,\ 0 \le a \le 5^+$

$N^S_{j,p,y,q,a} = \left(N^S_{j,p,y,q-1,a} - C^S_{j,p,y,q-1}\right) e^{-M^S_{y,a}/4}$ $\quad p = I, NI, y_1 \le y \le y_n,\ 2 \le q \le 4, 0 \le a \le 5^+$ (A4)

Numbers-at-length at 1 November (after infection and movement)

The model estimated numbers-at-length range from a 2.5cm minus group to a 24cm plus group, denoted 2.5^- and 24^+, respectively, in the remaining text.

$N^S_{j,p,y,l} = \sum_{a=0}^{5^+} A^{sur}_{j,y,a,l} N^S_{j,p,y,a}$ $\quad p = I, NI, y_1 \le y \le y_n,\ 2.5^- cm \le l \le 24^+ cm$ (A5)

The model predicted numbers-at-length of ages 1+ only are given by:

$N^{S,1+}_{j,p,y,l} = \sum_{a=1}^{5^+} A^{sur}_{j,y,a,l} N^S_{j,p,y,a}$ $\quad p = I, NI, y_1 \le y \le y_n,\ 2.5^- cm \le l \le 24^+ cm$ (A6)

The proportion of sardine of age a in component j that fall in length group l at 1 November, $\mathrm{A^{sur}_{j,y,a,l}}$, is calculated under the assumption that length-at-age is normally distributed about a von Bertalanffy growth curve, with a modification to the somatic growth rate at low ages:

$A^{sur}_{j,y,a,l} \sim N\left(L_{j,y,a}, \vartheta^2_a\right)$ [2]

where $L_{j,y,a} = \begin{cases} L^{small}_{j,y-a,\infty}\left(1 - e^{-1.5\kappa_j\left(a - t^{small}_{0,j,y-a}\right)}\right) & a \le 0.5 \\ L_{j,\infty}\left(1 - e^{-\kappa_j\left(a - t_{0,j,y-a}\right)}\right) & a \ge 0.5 \end{cases}$ $\quad y_1 \le y \le y_n,\ 2.5^- cm \le l \le 24^+ cm$ (A7)[3]

with $t_{0,j,y} = t_{0,j} + \varepsilon^t_y$ (A8)[4]

And $\varepsilon^t_y = \begin{cases} \eta^t_y & y = y_1 \\ \rho^t \varepsilon^t_{y-1} + \sqrt{1 - (\rho^t)^2}\,\eta^t_y & y_1 < y \le y_n \end{cases}$

Natural mortality

Natural mortality is modelled to vary annually in an autocorrelated manner around a median as follows (although the baseline assumes no such correlation – see Table A.1):

$M^S_{y,a=0} = \bar{M}^S_{ju} e^{\varepsilon^{ju}_y}$ with $\varepsilon^{ju}_{1984} = \eta^{ju}_{1984}$ and $\varepsilon^{ju}_y = \rho \varepsilon^{ju}_{y-1} + \sqrt{1-\rho^2}\,\eta^{ju}_y$, $y_1 \le y \le y_n$ (A9)

$M^S_{y,a=1+} = \bar{M}^S_{ad} e^{\varepsilon^{ad}_y}$ with $\varepsilon^{ad}_{1984} = \eta^{ad}_{1984}$ and $\varepsilon^{ad}_y = \rho \varepsilon^{ad}_{y-1} + \sqrt{1-\rho^2}\,\eta^{ad}_y$, $y_1 \le y \le y_n$ (A10)

Spawning biomass and biomass associated with the November survey

$SSB^S_{j,y} = \sum_p \sum_{l=2.5^-}^{24^+} f^S_{j,y,l} N^{S,1+}_{j,p,y,l} w^S_{j,l}$ $\quad y_1 \le y \le y_n$ (A11)

$SSB^{eff,S}_{j=W,y} = \xi_W SSB^S_{W,y} + (1 - \xi_S) SSB^S_{S,y}$ $\quad y_1 \le y \le y_n$

[2] Given the allowance for early/late recruitment in varying $t_{0,y}$ estimates annually, there may be some proportion of this distribution below a length of zero (due to late recruitment). In these cases, this proportion is removed from the proportion-at-length of the minus length class.

[3] The proportion is calculated as the area under the curve between the lower limit and upper limit of length class l. The lower and upper tails are included in the proportions calculated for the minus and plus groups, respectively.

[4] Additive error allows for early or late recruitment. While the timing of recruitment may vary between stocks due to differing environmental conditions on the west and south coasts, the same autocorrelation parameters are assumed here for simplicity reasons.

$$SSB_{j=S,y}^{eff,S} = (1-\xi_W)SSB_{W,y}^S + \xi_S SSB_{S,y}^S \qquad y_1 \le y \le y_n \qquad \text{(A12)}$$

$$B_{j,y}^S = k_{j,N}^S \sum_p \sum_{l=2.5^-}^{24^+} N_{j,p,y,l}^S w_{j,l}^S \qquad y_1 \le y \le y_n \qquad \text{(A13)}$$ [5][6]

Commercial selectivity

$$S_{j,y,q,l} = \begin{cases} 0 & l \le 5.5cm \\ \chi_{j,y,q} exp\left\{-\frac{(l+0.25-\bar{l}_{1,y})^2}{(\sigma_1^{sel})^2}\right\} + \frac{1}{1+exp\{-(l+0.25-\bar{l}_{2,j,y,q})/(\sigma_2^{sel})^2\}} & 6cm \le l \le l_{max} = 23cm \\ S_{j,y,q,lmax} & l > l_{max} \end{cases}$$ [7]

$$y_1 \le y \le y_n,\ 1 \le q \le 4 \qquad \text{(A14)}$$

$$S_{j,y,q,a} = \sum_{l=2.5^-}^{24^+} A_{j,y,q,a,l}^{com} S_{j,y,q,l} \qquad y_1 \le y \le y_n,\ 1 \le q \le 4,\ 0 \le a \le 5^+ \qquad \text{(A15)}$$

$$A_{j,y,q,a,l}^{com} \sim N\left(L_{j,y,q,a}^{com}, \left[\left(1-\frac{(2q-1)}{8}\right)\vartheta_a + \frac{(2q-1)}{8}\vartheta_{a+1}\right]^2\right)$$

where $L_{j=1,y,q,a}^{com} = \begin{cases} L_{j,y-a,\infty}^{small}\left(1-e^{-1.5\kappa_j\left(a+(2q-1)/8-t_{0,j,y-a}^{small}\right)}\right) & a \le 0.5 \\ L_{j,\infty}\left(1-e^{-\kappa_j\left(a+(2q-1)/8-t_{0,j,y-a}\right)}\right) & a \ge 0.5 \end{cases}$

and $L_{j=2,y,q,a}^{com} = L_{j,\infty}\left(1-e^{-\kappa_j\left(a+(2q-1)/8-t_{0,j,y-a}\right)}\right)$

$$y_1 \le y \le y_n,\ 1 \le q \le 4, 0 \le a \le 5^+,\ 2.5^- cm \le l \le 24^+ cm \qquad \text{(A16)}$$ [8]

Bycatch in the anchovy directed fishery

$$C_{j,p,y,q,a}^{bycatch} = \begin{cases} N_{j,p,y,q,a}^S F_{j,y,q,a}^{By} & 0 \le a \le 1 \\ 0 & 2 \le a \le 5^+ \end{cases} \qquad p = I, NI, y_1 \le y \le y_n,\ 1 \le q \le 4 \qquad \text{(A17)}$$ [9]

Catch in the directed sardine and round herring bycatch fisheries

$$C_{j,p,y,q,a}^{dir} = \left(N_{j,p,y,q,a}^S - C_{j,p,y,q,a}^{bycatch}\right) S_{j,y,q,a} F_{j,y,q} \qquad p = I, NI, y_1 \le y \le y_n,\ 1 \le q \le 4, 0 \le a \le 5^+ \qquad \text{(A18)}$$

Total catch

$$C_{j,p,y,q,a}^S = C_{j,p,y,q,a}^{bycatch} + C_{j,p,y,q,a}^{dir} \qquad p = I, NI, y_1 \le y \le y_n,\ 1 \le q \le 4, 0 \le a \le 5^+ \qquad \text{(A19)}$$

Fished proportion of the available biomass from the sardine bycatch with the anchovy directed fishery

$$F_{j,y,q=1,a=0}^{By} = \frac{\sum_{m=11}^{12} \sum_{l<lcut_{y-1,m}} C_{j,y-1,m,l}^{RLF,fleet=3} + \sum_{l<lcut_{y,m}} C_{j,y,1,l}^{RLF,fleet=3}}{\sum_p N_{j,p,y,q=1,a=0}^S}$$

---

[5] The biomass in y_n excludes age 0 fish, although the contribution of age 0 fish to the total biomass should be minor.

[6] A time invariant weight-at-length is used in this equation. Previous assessments adjusted the November weight-at-length annually, informed by the average weight of sardine sampled during the survey, to account for the differing condition factor of sardine at the time of the survey. However, recent discussions have clarified that the hydro-acoustic survey estimate of total biomass depends on the size of the fish swim bladder which depends (through the time invariant target strength relationship) on fish length only but not on the condition (skinniness/fattiness) of the fish at the time of the survey. A time-invariant weight-at-length therefore provides most appropriate basis to estimate biomass from the population model to correspond to the time series of biomasses from the survey (which is independent of sardine condition factor).

[7] The $l + 0.25$ denotes the middle of length class l. This function is renormalized to a maximum of 1.

[8] The proportion is calculated as the area under the curve between the lower limit and upper limit of length class l. The lower and upper tails are included in the proportions calculated for the minus and plus groups, respectively.

[9] "Selectivity" is incorporated in $F_{j,y,q,a}^{By}$, as the sardine bycaught is typically independent of sardine abundance, but rather correlated with anchovy recruitment which varies from year to year.

$$F^{By}_{j,y,q=1,a=1} = \frac{\sum_{m=11}^{12} \sum_{l \geq lcut_{y-1,m}} C^{RLF,fleet=3}_{j,y-1,m,l} + \sum_{l \geq lcut_{y,m}} C^{RLF,fleet=3}_{j,y,1,l}}{\sum_p N^S_{j,p,y,q=4,a=1}}$$

$$F^{By}_{j,y,q=2,a=0} = \frac{\sum_{m=2}^{4} \sum_{l<lcut_{y,m}} C^{RLF,fleet=3}_{j,y,m,l}}{\sum_p N^S_{j,p,y,q=2,a=0}} \qquad F^{By}_{j,y,q=2,a=1} = \frac{\sum_{m=2}^{4} \sum_{l \geq lcut_{y,m}} C^{RLF,fleet=3}_{j,y,m,l}}{\sum_p N^S_{j,p,y,q=2,a=1}}$$

$$F^{By}_{j,y,q=3,a=0} = \frac{\sum_{m=5}^{7} \sum_{l<lcut_{y,m}} C^{RLF,fleet=3}_{j,y,m,l}}{\sum_p N^S_{j,p,y,q=3,a=0}} \qquad F^{By}_{j,y,q=3,a=1} = \frac{\sum_{m=5}^{7} \sum_{l \geq lcut_{y,m}} C^{RLF,fleet=3}_{j,y,m,l}}{\sum_p N^S_{j,p,y,q=3,a=1}}$$

$$F^{By}_{j,y,q=4,a=0} = \frac{\sum_{m=8}^{10} \sum_{l<lcut_{y,m}} C^{RLF,fleet=3}_{j,y,m,l}}{\sum_p N^S_{j,p,y,q=4,a=0}} \qquad F^{By}_{j,y,q=4,a=1} = \frac{\sum_{m=8}^{10} \sum_{l \geq lcut_{y,m}} C^{RLF,fleet=3}_{j,y,m,l}}{\sum_p N^S_{j,p,y,q=4,a=1}} \qquad \text{(A20)}$$

A penalty is imposed within the model to ensure that $F^{By}_{j,y,q,a} < 0.95$.

Fished proportion of the available biomass from the directed sardine catch and sardine bycatch with round herring fishery

$$F_{j,y,q=1} = \frac{\sum_{fleet=1}^{2} \sum_{m=11}^{12} \sum_{l \geq 6cm} C^{RLF,fleet}_{j,y-1,m,l} + \sum_{fleet=1}^{2} \sum_{l \geq 6cm} C^{RLF,fleet}_{j,y,1,l}}{\sum_p \sum_{a=0}^{5+} \left(N^S_{j,p,y,1,a} - C^{bycatch}_{j,y,1,a}\right) S_{j,y,1,a}}$$

$$F_{j,y,q=2} = \frac{\sum_{fleet=1}^{2} \sum_{m=2}^{4} \sum_{l \geq 6cm} C^{RLF,fleet}_{j,y,m,l}}{\sum_p \sum_{a=0}^{5+} \left(N^S_{j,p,y,2,a} - C^{bycatch}_{j,y,2,a}\right) S_{j,y,2,a}}$$

$$F_{j,y,q=3} = \frac{\sum_{fleet=1}^{2} \sum_{m=5}^{7} \sum_{l \geq 6cm} C^{RLF,fleet}_{j,y,m,l}}{\sum_p \sum_{a=0}^{5+} \left(N^S_{j,p,y,3,a} - C^{bycatch}_{j,y,3,a}\right) S_{j,y,3,a}}$$

$$F_{j,y,q=4} = \frac{\sum_{fleet=1}^{2} \sum_{m=8}^{10} \sum_{l \geq 6cm} C^{RLF,fleet}_{j,y,m,l}}{\sum_p \sum_{a=0}^{5+} \left(N^S_{j,p,y,4,a} - C^{bycatch}_{j,y,4,a}\right) S_{j,y,4,a}} \qquad \text{(A21)}$$

A penalty is imposed within the model to ensure that $S_{j,y,a,l} F_{j,y,q} < 0.95$. Fish <6cm were seldom[10] caught and were thus not used in fitting this model. Commercial selectivity-at-length is fixed to zero for length classes <6cm (equation A12).

Number of recruits associated with the recruit survey

$$N^S_{j,y,r} = k^S_{j,r} \left(\left(N^S_{j,NI,y,2,0} - C^S_{j,NI,y,2,0}\right) e^{-(1/8+0.5t^S_y/12) M^S_{y,0}} - \tilde{C}^S_{j,y,0bs} \right) e^{-0.5 t^S_y \times M^S_{y,0}/12} \quad 1985 \leq y \leq y_n \qquad \text{(A22)}$$

Multiplicative survey bias

$$k^S_{j,N} = k^S_{ac} \qquad \text{(A23)}$$

$$k^S_{j=W,r} = k^S_{cov} \times k^S_{ac} \qquad \text{(A24)}$$

$k^S_{j=S,r} = k^S_{covS} \times k^S_{cov} \times k^S_{ac}$ (for the two mixing-component hypothesis only) (A25)

Survey trawl selectivity

$$S^{survey}_{j,l} = \begin{cases} 0 & l = 2.5^- cm \\ \left[1 + exp\left\{-\left(l + 0.25 - S_{50,j}\right)/\delta_j\right\}\right]^{-1} & 3cm \leq l \leq 24^+ cm \end{cases} \qquad y_1 \leq y \leq y_n \qquad \text{(A26)}$$

[10] Less than 6% of the quarters west of Cape Agulhas, less than 2% of the quarters south-east of Cape Agulhas and less than 4% of the quarters for the whole coast.

Proportion-at-length associated with the November survey

$$p^S_{j,y,l} = \begin{cases} \frac{\sum_p \sum_{l\le 6cm} N^S_{j,p,y,l} S^{survey}_{j,l}}{\sum_p \sum_{l=2.5^-}^{24^+} N^S_{j,p,y,l} S^{survey}_{j,l}} & l = 6^- cm \\ \frac{\sum_p N^S_{j,p,y,l} S^{survey}_{j,l}}{\sum_p \sum_{l=2.5^-}^{24^+} N^S_{j,p,y,l} S^{survey}_{j,l}} & 6.5cm \le l \le 20.5cm \\ \frac{\sum_p \sum_{l=21}^{23.5} N^S_{j,p,y,l} S^{survey}_{j,l}}{\sum_p \sum_{l=2.5^-}^{24^+} N^S_{j,p,y,l} S^{survey}_{j,l}} & l = 21 - 23.5cm \\ \frac{\sum_p N^S_{j,p,y,l} S^{survey}_{j,24^+}}{\sum_p \sum_{l=2.5^-}^{24^+} N^S_{j,p,y,l} S^{survey}_{j,l}} & l = 24^+ cm \end{cases} \quad {}^{11} \qquad y_1 \le y \le y_n \qquad \text{(A27)}$$

Proportion-at-length of fish infected with the parasite in November

$$P^S_{j,y,l} = \frac{N^S_{j,I,y,l}}{\sum_p N^S_{j,p,y,l}} \qquad y_1 \le y \le y_n,\ 7.5cm \le l \le 23cm \qquad \text{(A28)}$$

Catch-at-length from the directed and round herring bycatch fisheries

$$C^{dir}_{j,p,y,q,l} = \sum_{a=0}^{5^+} \left(N^S_{j,p,y,q,a} - C^{bycatch}_{j,p,y,q,a}\right) A^{com}_{j,q,a,l} S_{j,y,q,l} F_{j,y,q} \quad {}^{12}$$

$$p = I, NI, y_1 \le y \le y_n,\ 1 \le q \le 4,\ 2.5^- cm \le l \le 24^+ cm \quad \text{(A29)}$$

Proportion-at-length associated with the directed catch and round herring bycatch

$$p^{coml,S}_{j,y,q,l} = \begin{cases} \frac{\sum_p C^{dir}_{j,p,y,q,l}}{\sum_p \sum_{l=6}^{24^+} C^{dir}_{j,p,y,q,l}} & 6cm \le l \le 22.5cm \\ \frac{\sum_p \sum_{l=23}^{24^+} C^{dir}_{j,p,y,q,l}}{\sum_p \sum_{l=6}^{24^+} C^{dir}_{j,p,y,q,l}} & l = 23^+ cm \end{cases} \quad {}^{13} \qquad y_1 \le y \le y_n,\ 1 \le q \le 4 \qquad \text{(A30)}$$

<u>Fitting the Model to Observed Data (Likelihood)</u>

$$-lnL = -lnL^{Nov} - lnL^{rec} - lnL^{sur\,propl} - lnL^{com\,propl} - lnL^{prev} \qquad \text{(A31)}$$

where

$$-lnL^{Nov} = 0.5 \sum_j \sum_{y=y_1}^{y_n} \left\{ \frac{5^5 \left(\frac{\left|ln\left(\hat{B}^S_{j,y}\right) - ln\left(B^S_{j,y}\right)\right|}{\sqrt{\left(\sigma^S_{j,y,Nov}\right)^2 + \left(\phi^S_{ac}\right)^2 + \left(\lambda^S_{j,N}\right)^2}} \right)^5}{5^5 + \left(\frac{\left|ln\left(\hat{B}^S_{j,y}\right) - ln\left(B^S_{j,y}\right)\right|}{\sqrt{\left(\sigma^S_{j,y,Nov}\right)^2 + \left(\phi^S_{ac}\right)^2 + \left(\lambda^S_{j,N}\right)^2}} \right)^5} \right\}^{2/5} + ln\left[2\pi \left(\left(\sigma^S_{j,y,Nov}\right)^2 + (\phi^S_{ac})^2 + \left(\lambda^S_{j,N}\right)^2 \right)\right] \qquad \text{(A32)}$$

[11] The inclusion of model predicted proportion-at-length 24⁺cm is deliberate to take into account the zero samples of 24⁺cm sardine in the survey.

[12] Note the model predicted commercial catch of lengths <6cm is zero, from a zero commercial selectivity in equation A.13. This is consistent with the range of length classes in the observed commercial proportions-at-lengths.

[13] Note the model predicted commercial catch of lengths <6cm is zero, from a zero commercial selectivity in equation A.13. This is consistent with the range of length classes in the observed commercial proportions-at-lengths.

$$-lnL^{rec} = 0.5\sum_j \sum_{y=y_2}^{y_n} \left\{ \frac{5^5\left(\frac{\left|ln\left(\hat{N}^S_{j,y,r}\right)-ln\left(N^S_{j,y,r}\right)\right|}{\sqrt{\left(\sigma^S_{j,y,rec}\right)^2+\left(\phi^S_{ac}\right)^2+\left(\lambda^S_{j,r}\right)^2}}\right)^5}{5^5+\left(\frac{\left|ln\left(\hat{N}^S_{j,y,r}\right)-ln\left(N^S_{j,y,r}\right)\right|}{\sqrt{\left(\sigma^S_{j,y,rec}\right)^2+\left(\phi^S_{ac}\right)^2+\left(\lambda^S_{j,r}\right)^2}}\right)^5} \right\}^{2/5} + ln\left[2\pi\left(\left(\sigma^S_{j,y,rec}\right)^2+\left(\phi^S_{ac}\right)^2+\left(\lambda^S_{j,r}\right)^2\right)\right] \quad \text{(A33)}$$

$$-lnL^{sur\,propl} = w^{sur}_{propl}\sum_j \sum_{y=y_1}^{y_n} \left\{ \sum_{l=6}^{21^+} \left\{ \frac{\left(\sqrt{\hat{p}^S_{j,y,l}}-\sqrt{p^S_{j,y,l}}\right)^2}{2\left(\sigma^S_{j,sur}\right)^2} + ln\left(\sigma^S_{j,sur}\right) \right\} + \frac{\left(0-\sqrt{p^S_{j,y,24^+}}\right)^2}{2\left(\sigma^S_{j,sur}\right)^2} + ln\left(\sigma^S_{j,sur}\right) \right\} \quad \text{[14] (A34)}$$

$$-lnL^{com\,propl} = w^{com}_{propl}\sum_j \sum_{y=y_1}^{y_n} \sum_{q=1}^{4} \sum_{l=6}^{23^+} \left\{ \frac{\left(\sqrt{\hat{p}^{S,coml}_{j,y,q,l}}-\sqrt{p^{S,coml}_{j,y,q,l}}\right)^2}{2\left(\sigma^S_{j,com}\right)^2} + ln\left(\sigma^S_{j,com}\right) \right\} \quad \text{(A35)}$$

$$-lnL^{prev} = \sum_j \sum_{y=2010}^{2018} \sum_{l=7.5cm}^{23cm} -n^{prev}_{j,y,l}\, ln\left(P^S_{j,y,l}\right) - \left(N^{prev}_{j,y,l} - n^{prev}_{j,y,l}\right) ln\left(1-P^S_{j,y,l}\right) \quad \text{(A36)}$$

A "robustified likelihood" is used for the contributions from the hydro-acoustic surveys to ensure no undue influence from any extreme (outlying) values for residuals. The functional form chosen to robustify makes negligible difference for standardised residuals of magnitude three or less, but essentially treats large standardised residuals as if they do not exceed five in magnitude.

[14] The 21^+ group in this equation consists of the length classes 21cm, 21.5cm, 22cm, 22.5cm, 23cm and 23.5cm.

Table A1. Assessment model parameters and variables with associated fixed values or prior distributions and, for derived variables, associated equation numbers. As the majority of prior distributions are uninformative, notes are provided only for informative priors and/or bounds.

	Parameter / Variable	Description	Units / Scale	Fixed Value / Prior Distribution	Equation	Notes
Annual numbers and biomass	$N_{j,p,y,a}^{S}$	Model predicted numbers-at-age a at the beginning of November in year y of component j that are uninfected ($p = NI$) or infected ($p = I$) with the endoparasite	Billions	$ln(N_{j,NI,y,0}^{S})/10 \sim U(-10,3.2)$ $N_{j,I,y,0}^{S} = 0$	A1 - A3	
	$N_{j,p,1983,a}^{S}$	Initial numbers-at-age a in component j	Billions	$N_{j,NI,1983,a=1}^{S} \sim U(0,50)$ $N_{j,NI,1983,a}^{S} = 0,\ 2 \leq a \leq 5^{+}$ $N_{j,I,1983,a}^{S} = 0,\ 0 \leq a \leq 5^{+}$		
	$N_{j,p,y,q,a}^{S}$	Model predicted numbers-at-age a mid-way through quarter q of year y of component j that are uninfected ($p = NI$) or infected ($p = I$) with the endoparasite	Billions		A4	
	I_y	Proportion of uninfected west component sardine that are infected with the endoparasite in year y (two mixing-component hypothesis only)		$= 0,\ y_1 \leq y \leq 2007$ $\sim U(0,1),\ 2008 \leq y \leq y_n$		
	$move_{y,a}$	Proportion of west component sardine of age a which move to the south component at the beginning of November of year y (two mixing-component hypothesis only)	-	$move_{y,1} \sim Beta(1.05,1.05)$ $move_{y,2+} = \phi move_{y,1}$ $\phi \sim U(0,1)$		
	$SSB_{j,y}^{S}$	Model predicted spawning biomass of component j at the beginning of November in year y	Thousand tons		A11	
	$SSB_{j,y}^{eff,S}$	Model predicted effective spawning biomass of component j at the beginning of November in year y	Thousand tons		A12	
	$B_{j,y}^{S}$	Model predicted total biomass of component j at the beginning of November in year y, associated with the November survey	Thousand tons		A13	
	ξ_j	Proportion of j-component spawner biomass that contributes to the effective spawning biomass on the same coast		$\xi_W = 1$ $\xi_S = 0.92$		Alternative values considered in robustness tests
	$w_{j,l}^{S}$	Mean mass of sardine of component j in length class l	Grams	$5.6876 \times 10^{-6} \times l^{3.140026}$		OLSPS (2020)

Table A1 (Continued).

	Parameter / Variable	Description	Units / Scale	Fixed Value / Prior Distribution	Equation	Notes
Annual numbers and biomass	$f_{j,y,l}^S$	Proportion of component j sardine that are mature in length class l in year y	-	$\left[1+e^{-(l-17.2)/1.17}\right]^{-1}$ $1984 \leq y \leq 1987$ $\left[1+e^{-(l-18.6)/1.26}\right]^{-1}$ $1988 \leq y \leq 1995$ $\left[1+e^{-(l-19.4)/1.40}\right]^{-1}$ $1996 \leq y \leq 2003$ $\left[1+e^{-(l-17.4)/0.95}\right]^{-1}$ $2004 \leq y \leq 2018$		Refit from data used by van der Lingen *et al*. (2006) using midpoints of length classes. Assuming maturity post-2003 reflects that of 1965-1975 as maturity is hypothesized to be density dependent (van der Lingen *et al*. 2006) and both these periods correspond to low biomass following a peak in abundance
	$N_{j,y,r}^S$	Model predicted number of juveniles of component j at the time of the recruit survey in year y	Billions		A23	
Natural mortality	$M_{y,a}^S$	Rate of natural mortality of age a in year y	Year^{-1}	$M_{y,0}^S = 1.0$ $M_{y,1+}^S = 1.0$	A9 and A10	Selected based on maximized joint posterior, and subject to a compelling reason to modify from previous assessment
	$\bar{M}_{ju}^S$	Median juvenile rate of natural mortality	Year^{-1}	1.0		
	$\bar{M}_{ad}^S$	Median rate of natural mortality for 1+ sardine	Year^{-1}	0.8		
	ε_y^{ju}	Annual residuals about juvenile natural mortality rate	-		A9	
	ε_y^{ad}	Annual residuals about natural mortality rate for 1+ sardine	-		A10	
	η_y^{ju}	Normally distributed error in calculating ε_y^{ju}	-	$N(0,\sigma_j^2)$		
	η_y^{ad}	Normally distributed error in calculating ε_y^{ad}	-	$N(0,\sigma_{ad}^2)$		
	σ_j	Standard deviation in the annual residuals about juvenile natural mortality	-	0		See robustness tests
	σ_{ad}	Standard deviation in the annual residuals about natural mortality for ages 1+	-	0		See robustness tests
	ρ	Annual autocorrelation coefficient	-	0		See robustness tests

Table A1 (Continued).

	Parameter / Variable	Description	Units / Scale	Fixed Value / Prior Distribution	Equation	Notes
Proportions-at-length and growth curve	$N^{S}_{j,p,y,l}$	Model predicted numbers-at-length l at the beginning of November in year y of component j that are uninfected ($p = NI$) or infected ($p = I$) with the endoparasite	Billions		A5	
	$p^{S}_{j,y,l}$	Model predicted proportion-at-length l of component j associated with the November survey in year y	-		A27	
	$A^{sur}_{j,y,a,l}$	Proportion of age a of component j sardine that falls in the length group l in November of year y	-		A7	
	κ_j	Somatic growth rate parameter for component j	Year^{-1}	$U(0,3)$		
	$L_{j,\infty}$	Maximum length (in expectation) of component j	Cm	$L_{j,\infty} = \frac{L_{j,1}e^{-2\kappa_j} - L_{j,3}}{e^{-2\kappa_j} - 1}$ where $L_{j,a=1} \sim U(5,25)$ $L_{j,a=3} - L_{j,a=1} \sim U(5,25)$		
	$L^{small}_{j,\infty,y}$	Maximum length (in expectation) of component $j = 1$ for the growth curve below $a \acute{=} 0.5$ years in year y	Cm	$= \frac{L_{j,\infty}\left(1 - e^{-\kappa_j(\acute{a} - t_{0,j,y})}\right)}{\left(1 - e^{-1.5\kappa_j\left(\acute{a} - t^{small}_{0,j,y}\right)}\right)}$		
	$t_{0,j,y}$	Age at which the length (in expectation) is zero for component j in year y	Year		A8	
	$t^{small}_{0,j,y}$	Age at which the length (in expectation) is zero for component $j = 1$ in year y for the growth curve below $a = 0.5$ years	Year	$= \frac{1}{1.5\kappa_j} ln\left(\frac{e^{1.5\kappa_j\acute{a} - \kappa_j(\acute{a} - t_{0,j,y})}}{1.5 + (1 - 1.5)e^{-\kappa_j(\acute{a} - t_{0,j,y})}}\right)$		
	$t_{0,j}$	Average age at which the length (in expectation) is zero	Year	$\frac{1}{\kappa_j} ln\left\{\frac{e^{\kappa_j}\left(L_{j,1} - L_{j,3}\right)}{L_{j,1}e^{-2\kappa_j} - L_{j,3}}\right\}$		
	ε^{t}_{y}	Annual autocorrelated residuals about the age at which the length is zero			A8	
	η^{t}_{y}	Annual uncorrelated residuals about the age at which the length is zero		$N(0, 0.2^2)$		
	ρ^{t}	Autocorrelation coefficient in these residuals		$U(-1,1)$		
	ϑ_a	Standard deviation of the distribution about the mean length for age a	-	$U(0,3)$, $a = 0,1,2^{+}$		Upper bound precludes unrealistically large lengths for young fish
	$p^{com,S}_{j,y,q,l}$	Model predicted proportion-at-length l of component j in the directed catch and round herring bycatch during quarter q of year y	-		A30	
	$A^{com}_{j,y,q,a,l}$	Proportion of age a of component j sardine that falls in the length group l mid-way through quarter q of year y	-		A16	
	$P^{S}_{j,y,l}$	Model predicted proportion-at-length l of component j that are infected with the endoparasite, at the time of the November survey in year y			A28	

Table A1 (Continued).

	Parameter / Variable	Description	Units / Scale	Fixed Value / Prior Distribution	Equation	Notes
Selectivity	$S_{j,l}^{survey}$	Survey selectivity-at-length l in the November survey for component j	-		A26	Some smaller fish escape through the trawl net
	$S_{50,j}$	Length at which survey selectivity is 50% for component j	Cm	$U(2.5{,}20)$		
	δ_j	Inverse of slope of survey selectivity-at-length ogive when selectivity is 50% for component j	-	$U(0.05{,}50)$		
	$S_{j,y,q,l}$	Commercial selectivity-at-length l during quarter q of year y of component j	-		A14	
	$S_{j,y,q,a}$	Commercial selectivity-at-age a during quarter q of year y of component j	-		A15	
	$\chi_{j,y,q}$	Height of the Gaussian component for component j relative to the height of the logistic component in quarter q of year y	-	$U(0,1)$ for $j=1$ = 0 for $j=2$		No bycatch modelled for south component
	$\bar{l}_{1,y}$	Mean of the Gaussian distribution for in year y	mm	$N(100, 10^2)$		
	$\bar{l}_{2,j,y,q}$	Length at 50% selectivity in the logistic component for component j in quarter q of year y	mm	$\bar{l}_{2,j,y,1} - \bar{l}_{1,2000} \sim U(0{,}150)$ $\bar{l}_{2,j,y,2} - \bar{l}_{1,2000} \sim U(0{,}150)$ $\bar{l}_{2,j,y,3} = \bar{l}_{2,j,y,2}$ $\bar{l}_{2,j,y-1,4} = \bar{l}_{2,j,y12}$		Estimated for two time periods per component: 1984-1993, 1994-2018 (west) and 1984-1997, 1998-2018 (south)
	$(\sigma_1^{sel})^2$	Variance parameter of the Gaussian distribution	mm	$U(20{,}150)$		
	$(\sigma_2^{sel})^2$	Variance parameter of the logistic distribution	mm	$U(0{,}100)$		
Multiplicative bias	$k_{j,N}^S$	Multiplicative bias associated with the November survey of component j	-		A23	
	$k_{j,r}^S$	Multiplicative bias associated with the recruit survey of component j	-		A24 – A25	
	k_{ac}^S	Multiplicative bias associated with the hydro-acoustic survey	-	$ln(k_{ac}^S) \sim N(-0.311, 0.094^2)$		Appendix B of de Moor and Butterworth (2016)
	k_{cov}^S	Multiplicative bias associated with the coverage of the recruits during the recruit survey in comparison to the coverage of the biomass during the November survey	-	Uniform prior on logit transpose of k_{cov}^S, such that $0.3 \le k_{cov}^S \le 1$		Lower bound selected in discussions with scientists on these surveys and their field experience
	k_{covS}^S	Multiplicative bias associated with the coverage of the south component recruits in comparison to the west component recruits during the recruit survey		$U(0{,}1)$		

Table A1 (Continued).

	Parameter / Variable	Description	Units / Scale	Fixed Value / Prior Distribution	Equation	Notes
Catch	$C^{S}_{j,p,y,q,a}$	Model predicted number of age a fish of component j caught during quarter q of year y that are uninfected ($p = NI$) or infected ($p = I$) with the endoparasite	Billions		A19	
	$lcut_{y,m}$	Cut off length for recruits in month m of year y	Cm	de Moor *et al*. 2020		Differ by month and year as informed by the recruit surveys
	$C^{bycatch}_{j,p,y,q,a}$	Number of age a fish of component j bycaught in the anchovy-directed fishery in quarter q of year y that are uninfected ($p = NI$) or infected ($p = I$) with the endoparasite	Billions		A17	
	$C^{dir}_{j,p,y,q,a}$	Number of age a fish of component j caught in the sardine-directed and round herring bycatch fisheries in quarter q of year y that are uninfected ($p = NI$) or infected ($p = I$) with the endoparasite	Billions		A18	
	$C^{dir}_{j,p,y,q,l}$	Number of length l fish of component j caught in the sardine-directed and round herring bycatch fisheries in quarter q of year y	Billions		A29	
	$F^{By}_{j,y,q,a}$	Fished proportion in quarter q of year y for age class a of component j, of bycatch in the anchovy-directed fishery	-		A20	
	$F_{j,y,q}$	Fished proportion in quarter q of year y for a fully selected age class a of component j, by the directed and round herring bycatch fisheries	-		A21	
Likelihood	$-lnL^{Nov}$	Contribution to the negative log likelihood from the model fit to the November survey biomass data	-		A32	
	$-lnL^{rec}$	Contribution to the negative log likelihood from the model fit to the recruit survey data	-		A33	
	$-lnL^{surpropl}$	Contribution to the negative log likelihood from the model fit to the November survey proportion-at-length data	-		A34	
	$-lnL^{compropl}$	Contribution to the negative log likelihood from the model fit to the quarterly commercial proportion-at-length data	-		A35	
	$-lnL^{surprev}$	Contribution to the negative log likelihood from the model fit to the November parasite prevalence-at-length data	-		A36	
	ϕ^{S}_{ac}	CV associated with factors which cause bias in the acoustic survey estimates and which vary inter-annually rather than remain fixed over time	-	=0.227		Appendix B of de Moor and Butterworth (2016)
	$\left(\lambda^{S}_{j,N/r}\right)^2$	Additional variance (over and above $\left(\sigma^{S}_{j,y,Nov/rec}\right)^2$ and $(\phi^{S}_{ac})^2$) associated with the November/recruit surveys of component j	-	$U(0,10)$		

Table A1 (Continued).

Parameter / Variable	Description	Units / Scale	Fixed Value / Prior Distribution	Equation	Notes
w_{propl}^{sur}	Weighting applied to the remaining survey proportion-at-length data	-	$= 0.5 \times 0.167$		To allow for autocorrelation[15]
$\sigma_{j,sur}^{S}$	Standard deviation associated with the survey proportion-at-length data of component *j*	-		$\sqrt{\frac{\sum_{y=y_1}^{y_n}\sum_{l=6^-}^{21^+}\left(\sqrt{\hat{p}_{j,y,l}^{S}}-\sqrt{p_{j,y,l}^{S}}\right)^2}{\sum_{y=y_1}^{y_n}\sum_{l=6^-}^{21^+}-1}}$[16]	Closed form solution
w_{propl}^{com}	Weighting applied to the commercial proportion-at-length data		$= 0.5 \times 0.04$		To allow for autocorrelation[17]
$\sigma_{j,com}^{S}$	Standard deviation associated with the commercial proportion-at-length data of stock *j*	-	$\sqrt{\frac{\sum_{y=y_1}^{y_n}\sum_{q=1}^{4}\sum_{l=6}^{23^+}\left(\sqrt{\hat{p}_{j=1,y,q,l}^{comlS}}-\sqrt{\cdots}\right.}{\sum_{y=y_1}^{y_n}\sum_{q=1}^{4}\sum_{l=6}^{23^+}1}}$ $\sqrt{\frac{\sum_{y=y_1}^{y_n}\sum_{q=1}^{4}\sum_{l=13}^{23^+}\left(\sqrt{\hat{p}_{j=2,y,q,l}^{comlS}}-\sqrt{\cdots}\right.}{\sum_{y=y_1}^{y_n}\sum_{q=1}^{4}\sum_{l=13}^{23^+}1}}$	Closed form solution[18]	$\sigma_{j,com}^{S}$

[15] Based upon data being available ~6 times more frequently than annual age data which contain maximum information content on this.
[16] The 21^+ group in this equation consists of the length classes 21cm, 21.5cm, 22cm, 22.5cm, 23cm and 23.5cm.
[17] Based upon data being available ~4x6 times more frequently than annual age data which contain maximum information content on this.
[18] A shorter range of lengths is used for the south component given the near absence of data outside this range, resulting in small/zero residuals, which would negatively bias this estimate.

Table A2. Assessment model data, detailed in de Moor *et al*. (2021)[19].

Quantity	Description	Units / Scale
t_y^S	Time lapsed between 1 May and the start of the recruit survey in year y	Months
$\tilde{C}_{j,y,0bs}^S$	Number of juveniles of component j caught between 1 May and the day before the start of the recruit survey in year y	Billions
$C_{j,y,m,l}^{RLF,fleet}$	Number of fish in length class l landed by $fleet$ in month m of year y of component j. $fleet = 1$ denotes the sardine directed fishery, $fleet = 2$ denotes the sardine bycatch with round herring (1984-2011) or ≥14cm sardine bycatch (2012-19) and $fleet = 3$ denotes the juvenile sardine bycatch with anchovy (1984-2011) or <14cm sardine bycatch (2012-19)	Billions
$\hat{B}_{j,y}^S$	Acoustic survey estimate of biomass of component j from the November survey in year y	Thousand tons
$\sigma_{j,y,Nov}^S$	Survey sampling CV associated with $\hat{B}_{j,y}^S$ that reflects survey inter-transect variance	-
$\hat{N}_{j,y,r}^S$	Acoustic survey estimate of recruitment of component j from the recruit survey in year y	Billions
$\sigma_{j,y,rec}^S$	Survey sampling CV associated with $\hat{N}_{j,y,r}^S$ that reflects survey inter-transect variance	-
$\hat{p}_{j,y,l}^S$	Observed proportion (by number) of component j in length group l in the November survey of year y	-
$\hat{p}_{j,y,q,l}^{S,com}$	Observed proportion (by number) of the directed catch and round herring bycatch of fish of component j and length group l during quarter q of year y	-
$n_{j,y,l}^{prev}$	Number of sardine of component j in length class l sampled from the November survey in year y that were tested and found to be infected with the endoparasite	Numbers
$N_{j,y,l}^{prev}$	Number of sardine of component j in length class l sampled from the November survey in year y that were tested for infection with the endoparasite	Numbers

[19] Note that the expected mass by length class and month, used to calculate $C_{j,y,m,l}^{RLF,fleet}$ and $\tilde{C}_{j,y,0bs}^S$, is given by $EM_{y,l,m} = 0.0000090193 \times l_{mid}^{3.066305} \times N_{y,l,m}$ for the west component and $EM_{y,l,m} = 0.00023041 \times l_{mid}^{2.463739} \times N_{y,l,m}$ for the south component.