Working Document to WGHANSA 2022 (23-27 May 2022)

Exploratory assessment of anchovy 27.9a-west using a surplus production model.

Alexandra A. Silva[1], Laura Wise[1], Fernando Ramos[2], Margarita Rincon[2], Susana Garrido[1], Andres Uriarte[3], Tobias Mildenberger[4]

[1] IPMA, Lisbon, Portugal

[2] *Centro Oceanográfico de Cádiz, Instituto Español de Oceanografía (IEO-CSIC), Cádiz, Spain.*

[3] AZTI, Pasaia, Spain

[4] National Institute of Aquatic Resources, Technical University of Denmark, Lyngby, Denmark

Abstract

The aim of this WD was to explore surplus production models to assess the western component of the anchovy 27.9a stock. Models were fitted to catch per quarter or semester (1991 – 2021) and to one biomass index, the spring acoustic survey (1999 – 2021), or two biomass indices, the acoustic survey and the autumn groundfish survey (1991 – 2018) using SPiCT. Various assumptions regarding the shape of the production curve, the initial biomass depletion and the intrinsic growth rate of the population were combined such that models varied from nearly unconstrained (more complex) to increasingly constrained (less complex). Bi-annual catch data and two survey indices lead to a higher number of convergent models. Several models passed all ICES criteria to accept a SPiCT assessment, except for a higher level of uncertainty in F/F_{MSY} than the agreed one for long-lived stocks. A model assuming a Schaefer production curve, a prior on *r* from a meta-analysis and, an initial depletion rate of 80%, showed better retrospective analysis, survey hindcast cross-validation and convergence performance than other candidate models. The results indicated that F/F_{MSY} was below 1 across most of the period, B/B_{MSY} fluctuated well below 1 until 2010 and above 1 since 2016. The present results may be considered for further work in a benchmark workshop.

1. Introduction

The anchovy 27.9a stock spans the ICES Division 9a corresponding to the region between Cape Finisterre and the Strait of Gibraltar in the Gulf of Cadiz. Anchovy distributed off the western coast of the Iberian Peninsula, from Cape Finisterre to Cape Saint Vincent is the west component of the stock. The southern component ranges from Cape Saint Vincent to the Strait of Gibraltar, the southern waters of the Iberian Peninsula. ICES provides separate catch advice annually for each of the stock components using a common basis: the rule “one-over-two” constrained by an uncertainty cap of +/- 80% of the former catch advice (ICES 2018, 2020).

In the case of the southern component, the rule uses an SSB indicator estimated in a Gadget assessment model, using length-age based catches and, length-age based abundance indices from two acoustic surveys, ECOCADIZ and PELAGO. For the western component, the rule uses an indicator obtained by adding the biomasses estimated in the acoustic surveys PELAGO and PELACUS which together cover the area. The western component is data-poor. The limited data available before the 2000s is related to a near absence of the species in the area. Monitoring such small catches and very low abundance was practically impossible. Monitoring of the western component population started in the late 1990s as a “by-product” of acoustic surveys targeting sardines while catches started to be sampled systematically in the late 2010s (ICES, 2018).

The use of estimates from a stock assessment model may have advantages over the direct use of survey estimates in terms of catch advice. Models, as they integrate several sources of data and may take both observation and process error into account, become more robust to specific situations of bias or noise in the case of a single indicator, such as a research survey. The fact that anchovy is a short-lived species precludes the application of assessment and reference points methods developed by ICES for medium- and long-lived data-limited species, as they are often based on equilibrium assumptions (approximately constant recruitment over time) (ICES, 2018). This fact promoted the search for alternative methods, work that has been developed within the scope of the ICES WKDLSSLS. In 2021, the WKDLSSLS concluded that short-lived stocks that have sufficiently long time series (catch data and total biomass indicators) can be assessed with surplus production models (SPMs, also called biomass dynamic models) (ICES, 2021a), provided the data have enough contrast. Scientific advice can be formulated based on F_{MSY} and rules for achieving MSY should include biomass limits and uncertainty buffers (Mildenberger *et al*. 2021). The F_{MSY} rule will be most successful if applied to an assessment including an indicator of population biomass immediately before the management period and which includes most age classes of the exploitable population.

During the WKDLSSLS workshops, SPMs were applied to various short-lived stocks using SPiCT (SPiCT, Stochastic Surplus Production Model in Continuous Time; (Mildenberger *et al*. 2021; Pedersen & Berg, 2017), namely to the west and south components of the anchovy in division 9a, sprat on the west coast of Scotland and sardines in sub-area 7 (Celtic Sea). In the case of anchovy from the 9a south component, SPiCT showed a good performance and results comparable to those of the analytical model in use. Classical surplus production models were generally not applied to assess short/medium lived stocks, due to the high variability. The appearance of SPMs that allow observation and process error, such as SPiCT, increased the chances of good results with short-lived species (Zhou *et al*. 2009).

In this WD, we explored SPMs to assess the anchovy 9a-west stock component using SPiCT. Various combinations of catch data and survey indices and various model configurations were explored.

2. Material and methods

Data

- catch biomass, t, per quarter or semester from the beginning of the first quarter of 1991 to the end of the second quarter of 2021
- total biomass, t, in the spring acoustic surveys PELACUS+PELAGO 1999 – 2021 (gaps in 2000, 2004 and 2012) (Massé *et al.*, 2017; Doray *et al.*, 2021)
- mean biomass and corresponding standard deviation (SD), kg h^{-1}, in groundfish surveys October/December 1991 – 2018 (autumn, with gaps in 1994 and 2012). The computation of indices followed the methodology provided by Cochran (1977) for stratified random sampling and the survey methodology is described in ICES (2017).

Survey indices were corrected to reflect the exploitable biomass, assumed to correspond to the biomass of individuals > 10 cm total length, the minimum length present in the commercial catches. For both survey series, the differences between the corrected and uncorrected data were minor (see Figure 2.1 for the acoustic survey; in the groundfish survey, there were differences in 1997 and from 2014 to 2016, all below 3% except for 2015 where they were 23%).

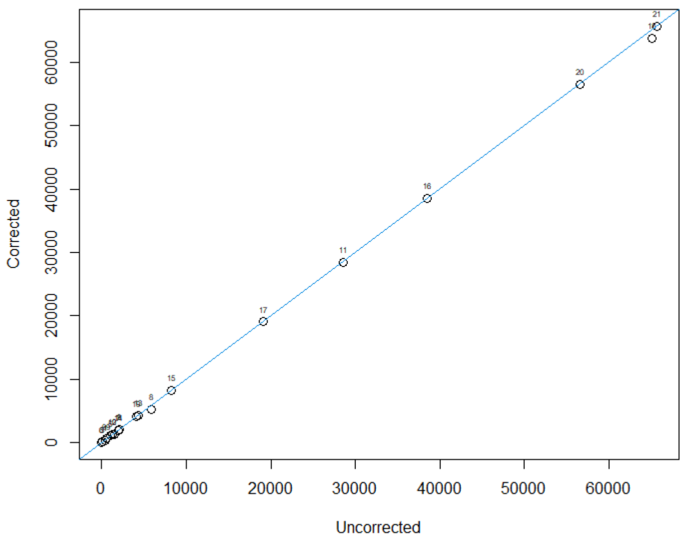

Figure 2.1 - Anchovy 9a-west: the relationship between uncorrected (all length classes) and corrected (biomass of length classes 10+) acoustic biomass.

Models were fitted to catch per quarter or semester and to one abundance index, the acoustic survey, or to both indices, with various assumptions regarding the shape of the production curve, the initial biomass depletion and the intrinsic growth rate of the population (see below). SPiCT fits surplus production models which incorporate dynamics in both biomass and fisheries and observation error of both catches and biomass indices. SPiCT uses a re-parametrization of the Pella and Tomlinson (1969) equation:

$dB_t = r/(n-1) * B_t*(1-(B_t/K)^{n-1}) - F_tB_t$

where B_t is the exploitable population biomass, F_t is the instantaneous fishing mortality rate, r is the intrinsic growth rate of the population, K is the carrying capacity and n is a unit-less parameter determining the shape of the production curve. The fraction B_1/K, where B_1 is the biomass in the first year of the assessment ($1-B_1/K$ is termed the initial depletion rate), is often difficult to estimate from the data. Data available on historical catches may be used to set priors for this parameter.

All models start in the middle of the calendar year (July 1st), following the ICES advice calendar for this stock. Assessment years go from 1 July of year *y* to 30 June of year *y+1*.

The time of catch (timeC) and survey (timeI) observations in the model is shown in Table 2.1.

Table 2.1 – Anchovy 9.a-west: Time of catch and survey observations. The forecast period is shown in bold.

Year	Time of catch observations				Time of survey observations	
	Quarterly data		Biannual data		Acoustic survey (spring)	Groundfish survey (autumn)
1991	1	1990.50	1	1990.5	1990.75	1991.25
1991	2	1990.75				
1991	3	1991.00	2	1991.0		
1991	4	1991.25				
1992	1	1991.50	1	1991.5	1991.75	1992.25
1992	2	1991.75				
1992	3	1992.00	2	1992.0		
1992	4	1992.25				
…	…	…	…	…	…	…
2021	1	2020.50	1	2020.5	2020.75	
2021	2	2020.75				
2021	**3**	**2021.00**	**2**	**2021.0**		
2021	**4**	**2021.25**				
2022	**1**	**2021.50**	**1**	**2021.5**		
2022	**2**	**2022.00**				

Coefficients of variation (CV) of groundfish indices were used as weighting factors of the data points to reflect differences in observation error. Acoustic surveys were given equal weight (=1) over time since estimates of observation error were not available. For better numerical stability all indices and weighing factors were scaled to have a mean = 1.

Priors for n, B_1/K, and r were combined such that models varied from nearly unconstrained (more complex) to increasingly constrained (less complex) (Table 2.2; Figure 2.2). The n.Thorson and r.Thorson priors were derived from n and r parameters for Clupeiforms and *Engraulis encrasicolus*, respectively, obtained in meta-analyses (Thorson *et al.* 2012; Thorson, 2020). Default priors (lognormal, mean = ln(1), SD = 2) were applied to the ratios of process error of fishing mortality/biomass to observation error in catches/abundance indices.

Table 2.2 – Anchovy 9.a-west: Prior means and standard deviations for n, B1/k and r parameters. In all cases prior probability distributions are lognormal. SD of n.Thorson and r.Thorson priors calculated as sqrt[mean(r)^2 /predictive error(r)^2].

	Prior		
Parameter	Name	exp(Mean)	Standard deviation
n	Default	2.00	2.00
	Schaefer	2.00	1.00E-03
	Fox	1.00	1.00E-03
	n.Thorson	0.60	0.57
B_1/K	20	0.20	0.50
	50	0.50	0.50
	80	0.80	0.80
r	r.Thorson	1.98	0.28

Figure 2.2 – Anchovy 9.a-west: Diagram of all possible prior combinations.

n		B_1/K		r
n.none		BKnone		r.none
Default	**X**	BK20	**X**	r.Thorson
Schaefer		BK50		
Fox		BK80		
n.Thorson				

To find one or a few final models, standard criteria of convergence, goodness-of-fit and consistency were checked according to ICES guidelines for the acceptance of a SPiCT assessment (2021b) and to recommendations of Pedersen and Berg (2017) and Carvalho *et al.* (2021). The following checklist was applied:

1) Convergence: successful completion of the fit, finite and reasonable confidence intervals; all absolute values of parameter correlations below 0.95; low sensitivity to initial values;
2) Goodness-of-fit: residuals normal (Shapiro-Wilk test; q-q plot), unbiased (t-test comparing the mean to zero; scatterplot of standardized residuals) and independent (Ljung and Box (1978) test on four lags; empirical auto-correlation plot);
3) Consistency: 5-year retrospective Mohn's Rho of B/BMSY and F/FMSY between -0.22 and 0.30; consistent retrospective trajectories across the historical period;
4) Prediction skill: mean absolute scaled error (MASE) of each abundance index of 7-year hindcast cross-validation, below 1 and, as low as possible.

The checklist was applied sequentially, apart from the sensitivity test to initial values, a time-consuming procedure, which was therefore applied only to models that passed the checklist. The sensitivity test consists of perturbing initial parameter values by random proportions between -2 and +2 and re-fitting the model. The recommended number of trials is 30. A vector of the distance between the estimates of the main model parameters of each trial and those of the base model is provided. The closest the distances are to zero the better although quantitative thresholds to accept a model have not been defined yet. Here, we calculated the 50^{th} and the 90^{th} percentiles of the distance vector and the proportion of vectors which failed to converge as ad-hoc indices to compare models.

3. Results

3.1. Overview of anchovy catches and abundance

The historical series of anchovy catches in Portugal from 1943 to 2020 showed fluctuations around a mean of 722 t (SD = 1075 t) apart from a period of consistently higher catches since 2017 (mean ± SD = 6944 ± 1705 t) and a single high value in 1943 (7476 t) (Figure 3.1). Although there were no data from Spanish catches far back in time, assuming Portuguese catches made the bulk of the catches of the stock component as seen in recent years, there were no signs of overexploitation of the resource at the beginning of the assessment period (1991). In the late 2010s the abundance "took off" reaching unprecedented levels in recent years (Figure 3.2). The index of abundance in the autumn groundfish survey in a given year *y* is significantly positively correlated with the index of abundance in the acoustic survey the following year, *y+1* (r = 0.91, p<0.001). Both indices presented marked fluctuations since 2015; the groundfish survey showed an increase to very high abundance in 2017, which resulted mainly from a single haul with a catch of 600 Kg of adult anchovy; abundance drops markedly the following autumn (2018). A similar, although less dramatic, variation was observed in the acoustic survey from 2017 to 2018.

Total catches showed a strong seasonal component, being the highest in the 3rd quarter of the year and decreasing from the 3rd quarter of a year to the second quarter of the next year (Figure 3.3). On average, 36% of the catches were obtained in the first semester.

Length frequency distributions (LFDs) of catches and surveys by semester available from the period 2015 – 2020 indicate that surveys observe smaller sized anchovy than caught in the fisheries (Figure 3.4). The difference is more pronounced in the second semester, with the autumn groundfish survey showing large proportions of individuals around 11 cm (possibly recruits) in some years.

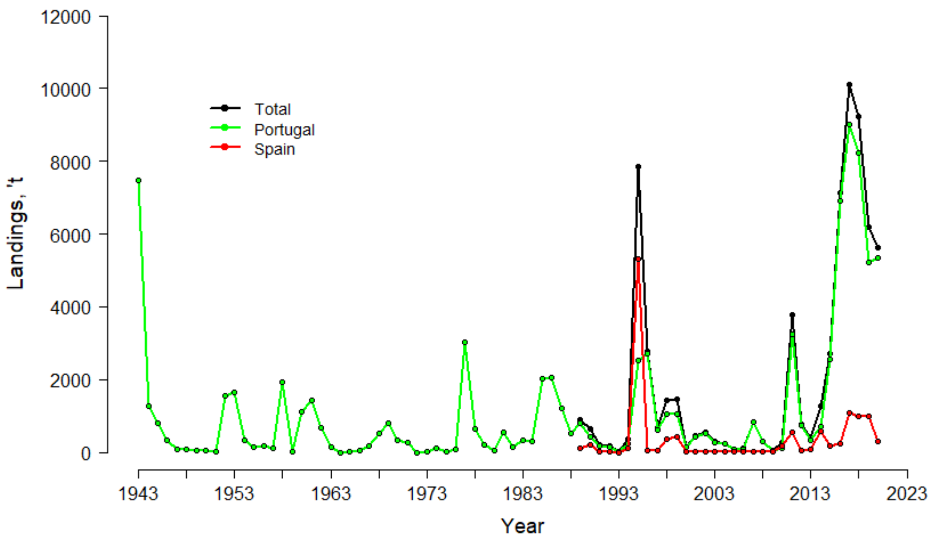

Figure 3.1 – Anchovy 9a-west: annual catch 1943 – 2021 by country and in total.

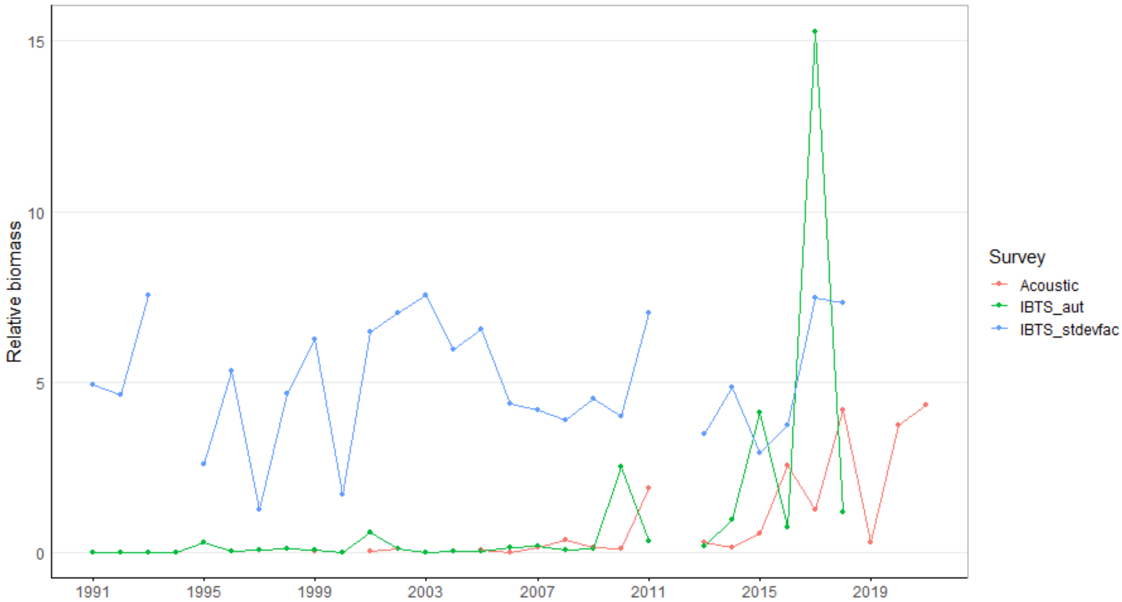

Figure 3.2 – Anchovy 9a-west: Index of the abundance of the spring acoustic survey 1999 – 2021 and index of abundance and coefficient of variation of the Portuguese autumn groundfish survey 1991 – 2018. Each survey and CV observation was divided by the mean of the corresponding series, therefore each series has a mean = 1 (the CV series was multiplied by 5 to improve the readability of the figure).

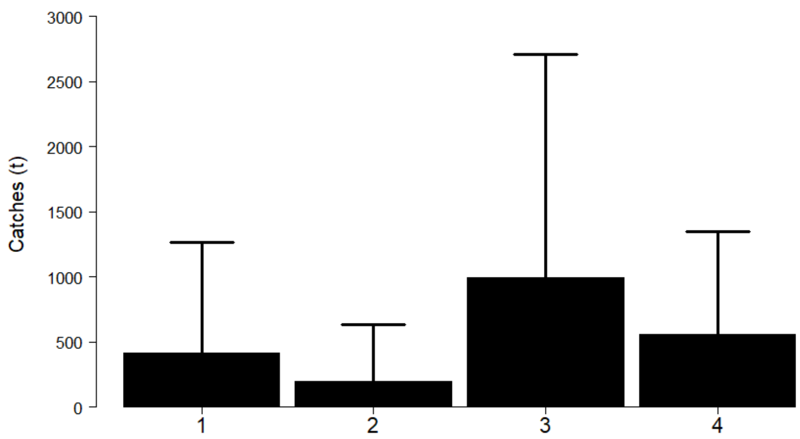

Figura 3.3 - Anchovy 9a-west: Mean catches by quarter in the period 1991 – 2021. Bars represent 1 standard deviation.

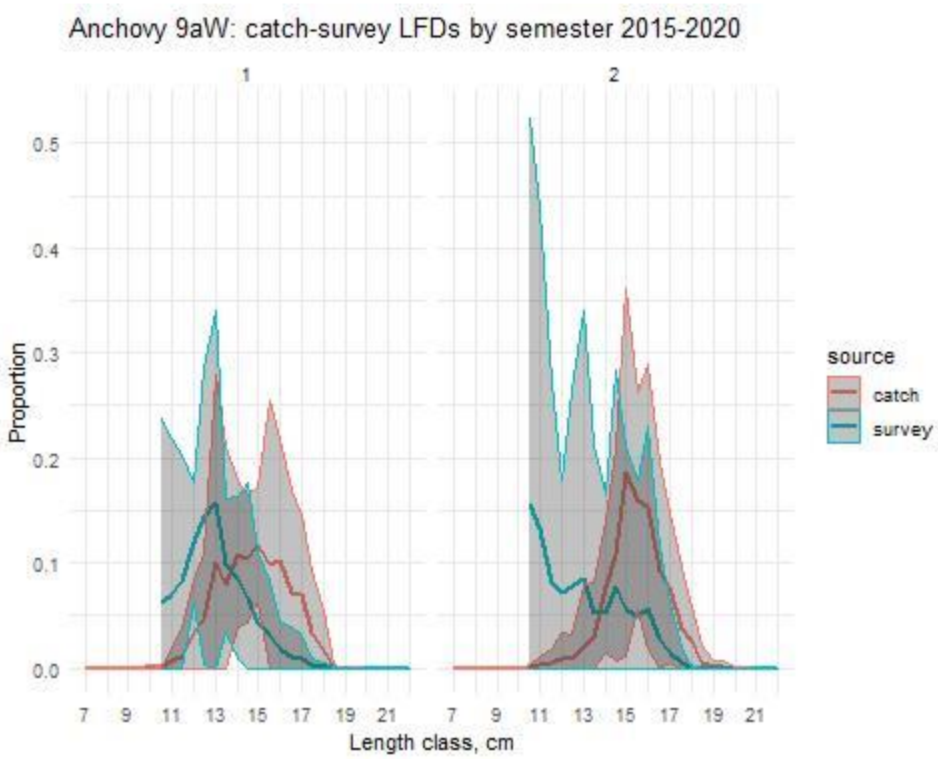

Figure 3.4 – Anchovy 9a-west: Mean proportion of individuals per ½ cm length class in the catches and the surveys by semester in the period 2015 – 2020.

3.2. Model diagnostics and results

Table 3.1 presents a summary of the main model diagnostics, parameters and derived quantities with corresponding estimates of uncertainty, for models which converged, had random, unbiased and independent residuals and, showed a minimum of four converged retrospective runs with consistent trajectories over time. Out of the initial 160 models, fourteen were retained on this step, 10 based on bi-annual catch data of which 4 used the acoustic survey and 6 used both the acoustic and the groundfish survey. All 14 models had, at least, one parameter with a prior from Table 2.2.

None of the models complied with the ICES guideline about the magnitude of B/B_{MSY} and F/F_{MSY} confidence intervals. Considering that larger uncertainty is expected for small pelagic fish due to their highly variable dynamics, this criterion was relaxed to admit models which estimated B/B_{MSY} and F/F_{MSY} confidence intervals spanning 2 orders of magnitude of the point estimates (T. Mildenberger, personal communication).

Models 11 to 14 showed the best performance in hindcast cross-validation of survey indices and overall good resistance to jittering of initial parameters (Table 3.1). Except for model 11, all showed MSY-K correlations above + 0.95; on the other hand, model 11 showed high sensitivity to the perturbation of initial values in a few trials. The four models had a similar performance regarding the checklist criteria and comparable point and uncertainty estimates of parameters (Table 3.1). The second retrospective trajectory, corresponding to the run with the 2019 acoustic survey and the 2018 groundfish survey as the last survey data points indicated considerably higher F/F_{MSY} and lower B/B_{MSY} in 2019 (with some backward effect) than the remaining retrospective runs (Figure 3.1). These surveys showed a 90% drop in biomass from the previous year's surveys and were followed by an increase of biomass of more than 1000%.

While any of the four models could be considered for further analysis, model 12, assuming a Schaefer production curve (n = 2), a Thorson prior on r (lognormal, mean=0.68, SD=0.30) and a lognormal prior on B1/K with mean = 0.20 (CV=0.50), corresponding to an initial depletion of 80%, seemed to have a slightly better retrospective, hindcast and convergence performance than the other 3 models (Table 3.1; Figures 3.1 and 3.2). Residuals complied with the assumptions of normality, no bias and independence (Figure 3.3). The retrospective pattern of the period 2016 – 2021 was positive for both B/B_{MSY} and F/F_{MSY} and, according to Mohn's Rho, substantially stronger for the latter while still below the threshold for short-lived species of 0.30. MASE scores were <1 for both surveys indicating the model had a superior prediction skill than the naïve baseline forecast (MASE=0.5 means twice as accurate as of the naïve forecast, i.e.; assuming the same abundance next year; Carvalho *et al.* 2021). The groundfish survey appears to have a better prediction skill than the acoustic survey; however, it is unclear if the fewer number of years used to calculate the MASE of the groundfish survey, 5 instead of 7 years, may have affected the result and prevented a fair comparison. Posterior distributions indicated that there is not much information on the data to estimate the intrinsic growth rate (Figure 3.4). Estimates of alpha ratios indicated that biomass process error was around double the observation error for both surveys (Table 3.1). On the other hand, the fishing mortality process error was about half the catch observation error. The estimate of B_{1991}/K (mean=0.11, CV=0.52) pointed to a depleted stock at the beginning of the assessment period.

Historical variations of B/B_{MSY} and F/F_{MSY} are shown in Figure 3.5. Point estimates of F/F_{MSY} were below 1 across most of the period. However, the huge confidence interval until the mid-2000s prevents any conclusion about the state of the stock. B/B_{MSY} fluctuated well below 1 until 2010. Since 2016, the stock has fluctuated slightly above B_{MSY}. On the 30 of June 2021, the end of the assessment period, the relative fishing mortality was estimated to be 0.06 and the relative total exploitable stock biomass was estimated to be 1.15, suggesting that the stock was healthy.

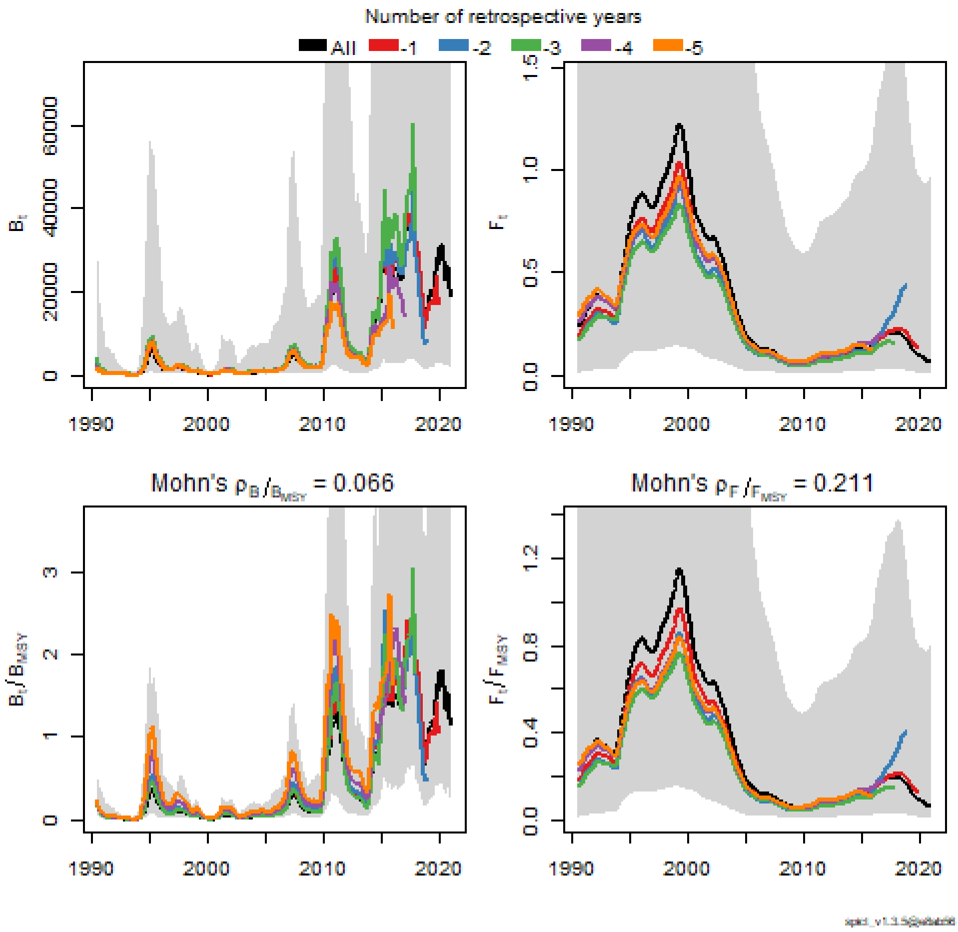

Figure 3.1 – Anchovy 9a-west: Retrospective error of B_{MSY} and F_{MSY} (top panel) and B/B_{MSY} and F/F_{MSY} (bottom panel) of model 12.

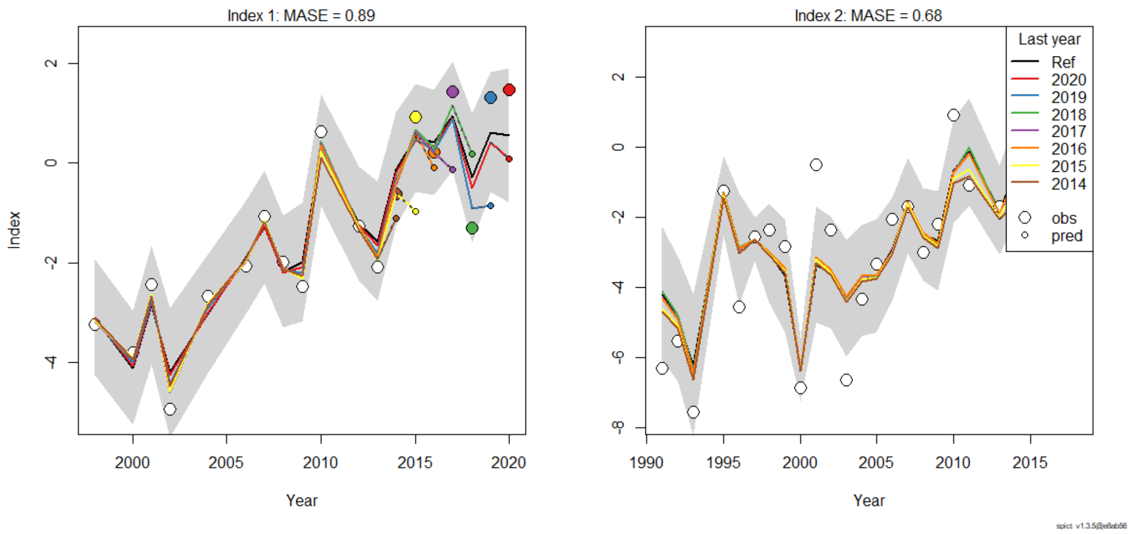

Figure 3.2 – Anchovy 9a-west: Hind-cast cross-validation results for the acoustic (left) and groundfish survey indices (right). The reference result corresponds to the result of Model 12; seven and five hindcast runs were carried out for the acoustic and the groundfish surveys, respectively (the last groundfish survey was in 2018).

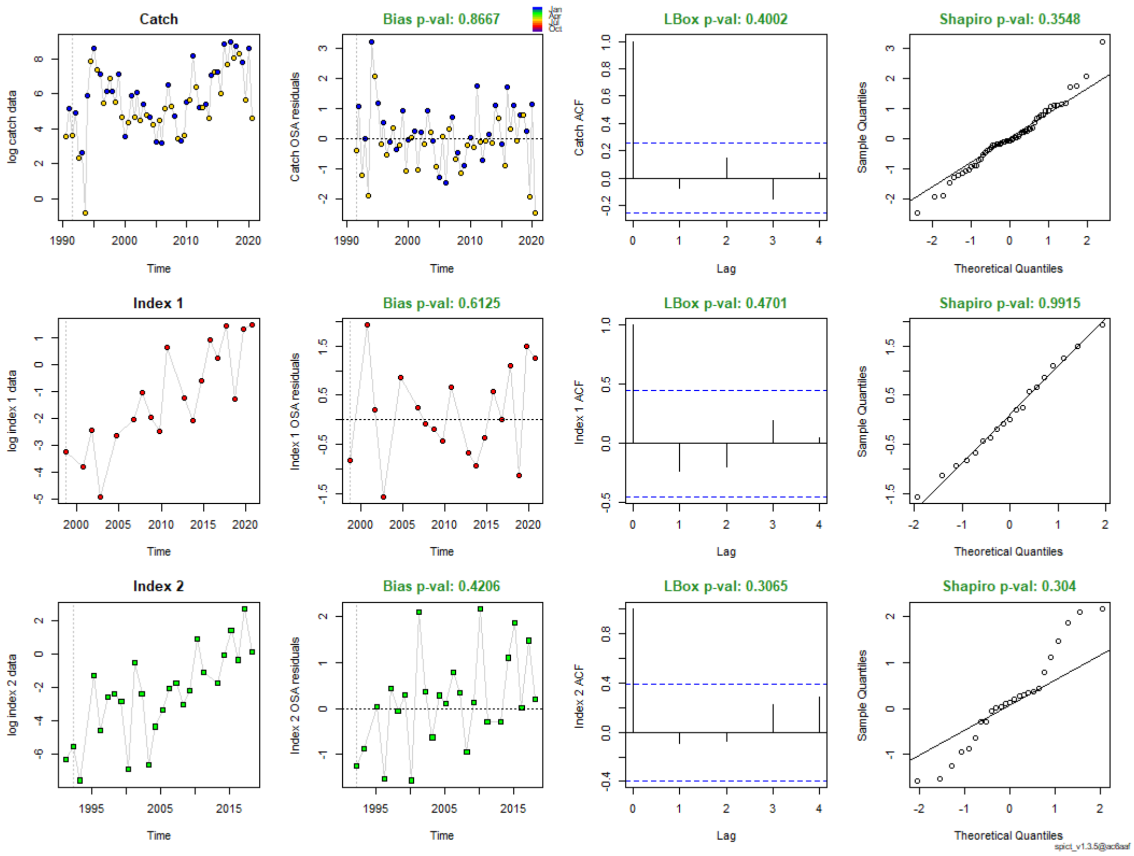

Figure 3.3 – Anchovy 9a-west: Plots of catch and survey residuals of Model 12.

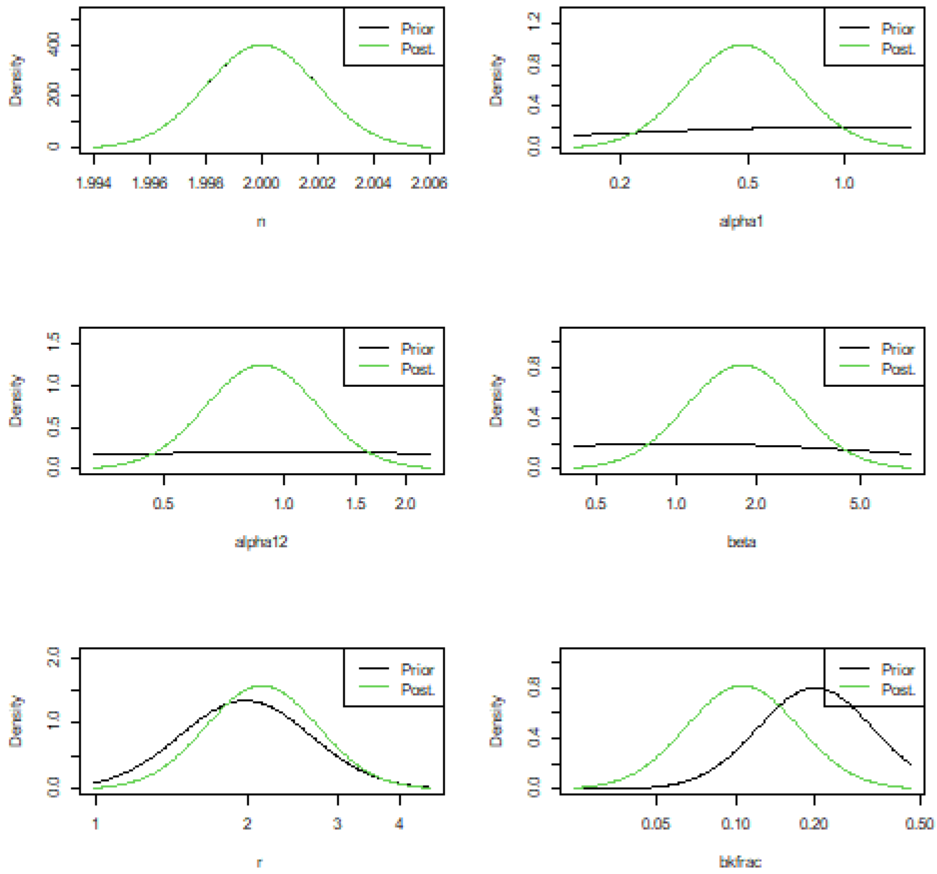

Figure 3.4 - Anchovy 9a-west: Prior and posterior distributions of n, alphas, beta, r and bk fraction of Model 12.

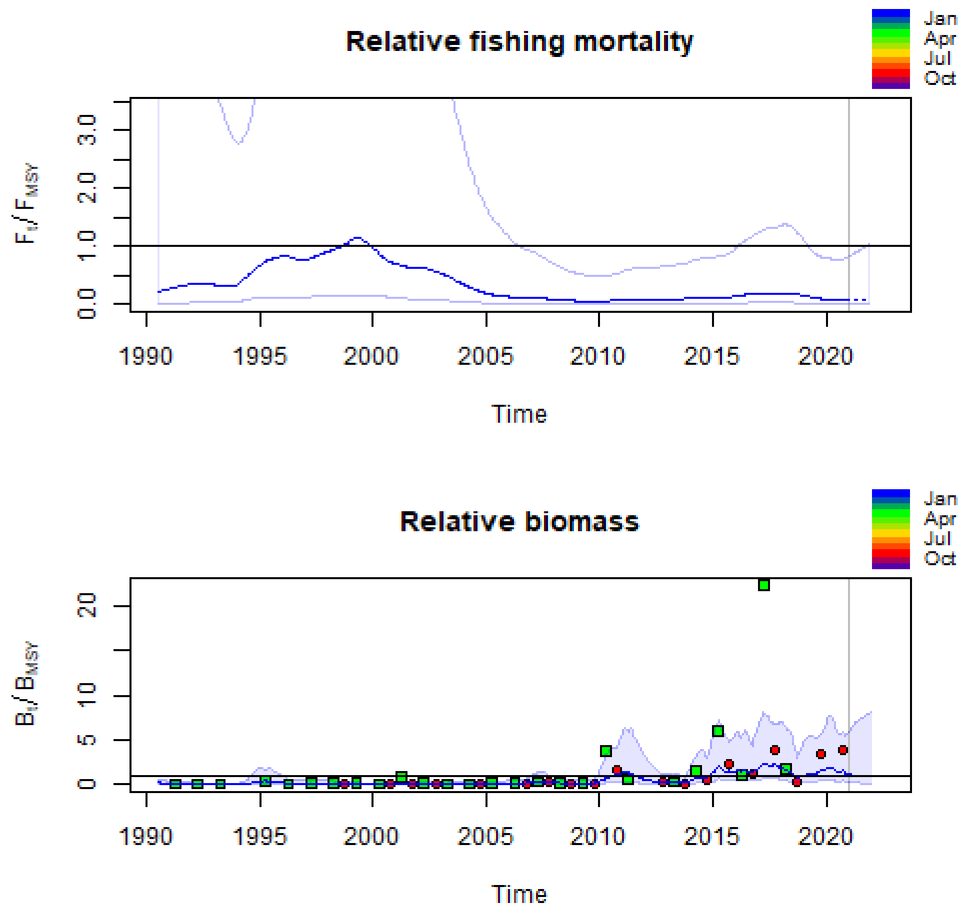

Figure 3.5 – Anchovy 9a-west: Historical F/F_{MSY} and B/B_{MSY} trajectories over the period 1991 – 2021. 95% CIs of relative biomass and fishing mortality are shown using shaded blue regions. The end of the data range is shown using a vertical grey line. Data are shown using points coloured by season.

3.3. Sensitivity of Model 12 to potentially biased survey data points

Three sensitivity tests of model 12 to down-weighting the following survey data points were carried out:

1) 2019 acoustic survey
2) 2019 acoustic survey and 2018 groundfish survey
3) 2017 groundfish survey

In all cases, the standard deviation of the data point was increased by a factor of 3, meaning an increase from 1 to 3 in the case of acoustic surveys and an increase from 1.44, to 4.32 in the case of 2018 groundfish survey.

Compared to Model 12, both models 1) and 2) showed a small decrease in the CV of B/B_{MSY} and the MASE of the acoustic survey (both around 8%) (Table 3.2). Changes in the CVs of the remaining parameters were negligible. The divergence of the second peel of the retrospective analysis decreased substantially in both runs compared with model 12 (Figures 3.1 and 3.6). Graphically, the fit of the model to the biomass in the two most recent years 2 improved (Figure 3.7). However, the Mohn's Rho of B/B_{MSY} and F/F_{MSY} increased 124% and 9% in comparison to model 12, respectively, something that was contrary to the expectation given the graphical pattern

Regarding test 3), down-weighting the 2017 groundfish data point decreased substantially the Mohn's Rho of F/F_{MSY} (67%) at the cost of cancelling the predictive power of the survey (MASE=1.1) (Table 3.2).

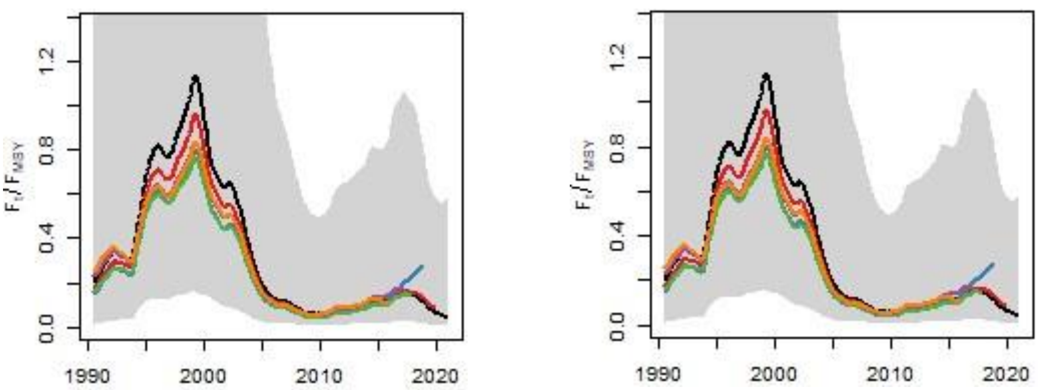

Figure 3.6 – Anchovy 9a-west: F/FMSY retrospective runs of models with down-weighted 2019 acoustic survey (left) and both the latter and the 2018 groundfish survey (right).

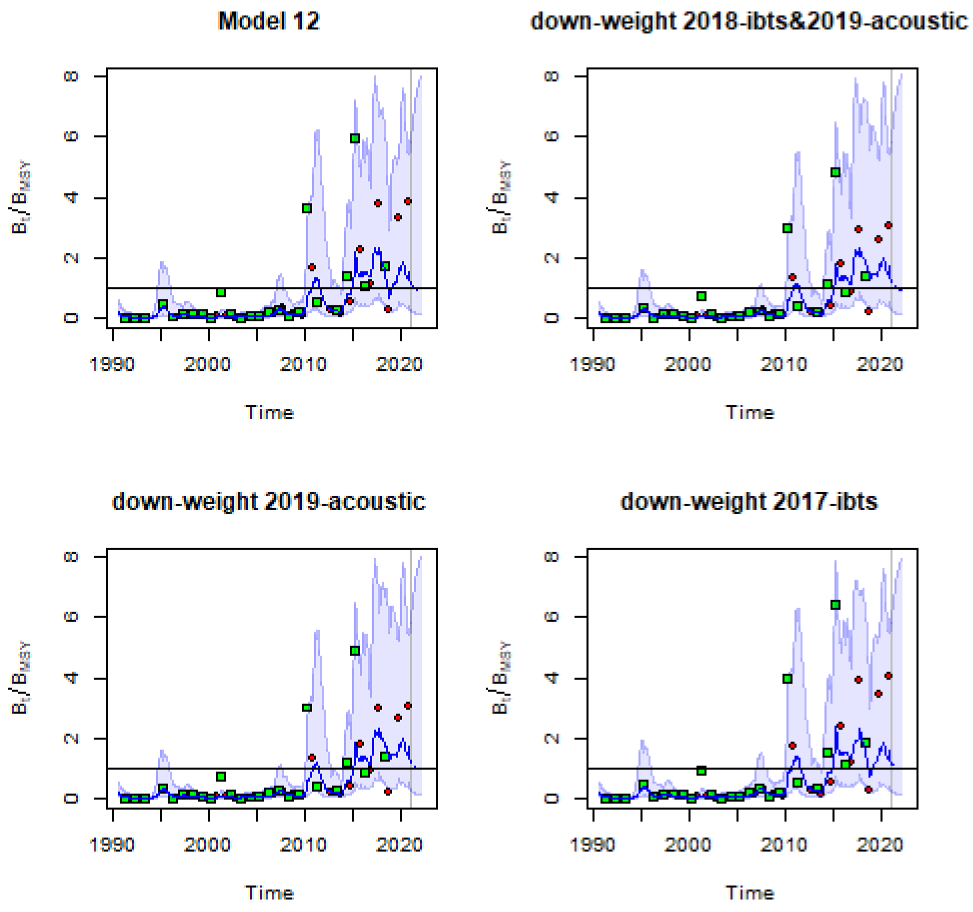

Figure 3.7 – Anchovy 9a-west: Plots of B/BMSY estimates of Model 12 and the models with down-weighted 2018 groundfish and 2019 acoustic surveys observations, down-weighted 2019 acoustic survey observation and down-weighted 2017 groundfish survey observation. The y-axis is truncated at 8 therefore the 2017 groundfish survey observation is not visible.

3.4. Sensitivity of Model 12 to the default prior assumptions on alpha and beta

Three additional sensitivity tests were:

4) Excluding the beta default prior
5) Excluding the alfa default prior
6) Excluding beta and alfa default priors

All tests had small effects on the CVs of parameters and derived quantities (Table 3.2). The main improvement when estimating alpha and beta parameters without priors was a decrease of 25-30% on F/F_{MSY} Mohn's Rho. Although at the same time the Mohn's Rho of B/B_{MSY} increased about 10%, the values were still well below the limits. Therefore, the free estimation of both alpha and beta parameters might be an option to consider in the final model.

4. Discussion

The following bullet points summarise the discussion in the group plenary:

- Surveys may not always represent the exploitable biomass, as they observe larger proportions of small individuals in years of good recruitment; small individuals may also be under-represented in the catches if there is slipping in those years; the fact that surveys are point observations in time may contribute to the differences observed in the LFDs; it may be sufficient that surveys cover the general length range caught in the fisheries; in future work, it may be worth to test the influence of a larger cut-off length (e.g. >13 cm) or corrected LFDs following Pedersen *et al.* (2017);
- The index of biomass of the autumn groundfish survey appears to be an acceptable index of abundance of anchovy since it showed a significantly positive correlation with that of the spring acoustic survey in the following year; both indices should continue to be explored for assessment purposes;
- PELACUS estimates are available since 2007 therefore this is the first year in the ICES assessment. In the present WD, the acoustic survey index starts in 1999, the first year with PELAGO survey estimates. Total abundance was assumed to be equal to the PELAGO estimates from 1999 to 2005 since abundance estimates of PELACUS in that period was assumed to be zero. The group considered this assumption to be acceptable since it was based on statements that PELACUS surveys were carried out although the estimation of anchovy was not possible due to its low abundance:
  "*Spanish acoustic surveys aimed at sardine have been conducted in Sub-division IXa North and Division VIIIc since 1983. Results from these surveys for the Sub-division IXa North have shown the scarce presence or even the absence of anchovy in this area (Carrera et al., 1999; Carrera, 1999, 2001). This situation still continues in the most recent years (surveys in the 2003-2007 period, see Porteiro et al., 2005; WD Iglesias et al., 2007)*". ICES, 2007, page 598).

- The possibility of using PELACUS and PELAGO separately was not considered an option because PELACUS coverage is not representative of the stock since it is just a small part of the western component distribution area.

- The decision to down-weight the 2018/2019 survey data points should be discussed with survey experts;
- The group noted that in the best model (model 12), F/F_{MSY} was estimated to be near the lowest historical harvest rate level calculated in the ICES assessment and well below the average of the historical series (Figure 4.1); the wide confidence intervals, namely in the past, may partly result from some very high harvest rates in combination with gaps in the acoustic survey series;
- Finally, the seasonal F parameter was inadvertently fixed equal to 1 in the bi-annual models but should have been estimated; the correction of this issue (running: `inp$phases$logphi=1)` was found just before the meeting, there was no time to re-run the models. It is noted that correcting this issue may introduce changes to the results presented so far.
- The WG considered that the present approach may be considered for further work in a benchmark workshop.

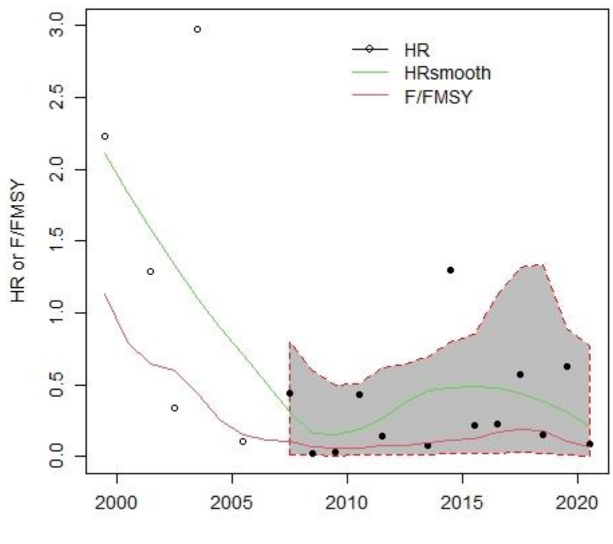

Figure 4.1 - Anchovy 9a-west: Estimates of F/F_{MSY} from the best SPiCT model (model 12) and harvest rates (ratio between annual catch and the PELAGO+PELACUS biomass) used in the ICES assessment (dots). The white dots show harvest rates before 2007 which are not used in the advice.

References

Carvalho, F., Winker, H., Courtney, D., Kapur, M., Kell, L., Cardinale, M., Schirripa, M. *et al*. 2021. A cookbook for using model diagnostics in integrated stock assessments. Fisheries Research, 240: 105959 .

Doray, M., Boyra, G., and van der Kooij, J. (Eds.). 2021. ICES Survey Protocols – Manual for acoustic surveys coordinated under ICES Working Group on Acoustic and Egg Surveys for Small Pelagic Fish (WGACEGG). 1st Edition. ICES Techniques in Marine Environmental Sciences Vol. 64. 100 pp. https://doi. org/10.17895/ices.pub.746

ICES. 2018. Report of the Benchmark Workshop on Pelagic Stocks (WKPELA 2018), 12–16 February 2018, ICES HQ, Copenhagen, Denmark. ICES CM 2018/ACOM:32. 313 pp.

ICES. 2020. Working Group on Southern Horse Mackerel, Anchovy and Sardine (WGHANSA). ICES Scientific Reports. 2:41. 655 pp. http://doi.org/10.17895/ices.pub.5977.

ICES. 2021a. Workshop on Data-Limited Stocks of Short-Lived Species (WKDLSSLS3). ICES Scientific Reports. 3:86. 60 pp. https://doi.org/10.17895/ices.pub.8145.

ICES 2021b. Benchmark Workshop on the development of MSY advice for category 3 stocks using Surplus Production Model in Continuous Time; SPiCT (WKMSYSPiCT) (tech. rep.). ICES Scientific Reports. 3:20. 316 pp. https://doi.org/10.17895/ices.pub.7919

Kell, L. T., Sharma, R., Kitakado, T., Winker, H., Mosqueira, I., Cardinale, M., & Fu, D. (2021). Validation of stock assessment methods: Is it me or my model talking? ICES Journal of Marine Science, https://doi. org/10.1093/icesjms/fsab104

Massé, J., Uriarte, A., Angélico, M. M., and Carrera, P. (Eds.) 2018. Pelagic survey series for sardine and anchovy in ICES subareas 8 and 9 – Towards an ecosystem approach. ICES Cooperative Research Report No. 332. 268 pp. https://doi.org/10.17895/ices.pub.4599.

Mildenberger, T. K., Berg, C. W., Kokkalis, A., Hordyk, A. R., Wetzel, C., Jacobsen, N. S., ... & Nielsen, J. R. (2022). Implementing the precautionary approach into fisheries management: Biomass reference points and uncertainty buffers. Fish and Fisheries, 23(1), 73-92.

Pedersen, M. W., & Berg, C. W. (2017). A stochastic surplus production model in continuous time. Fish and Fisheries, 18, 226–243.https://doi.org/10.1111/faf.12174

Zhou, A. E. Punt, R. Deng, C. M. Dichmont, Y. Ye and J. Bishop, “Modified Hierarchical Bayesian Biomass Dynamics Models for Assessment of Short-Lived Invertebrates: A Comparison for Tropical Tiger Prawns,” Marine & Freshwater Research, Vol. 60, No. 12, 2009, pp. 1298-1308. doi:10.1071/MF09022

Table 3.1 - Anchovy 9a-west: Summary of data, assumptions, diagnostics and results of the models fitted to anchovy data which converged, had random, unbiased and independent residuals and, showed a minimum of four converged retrospective runs with consistent trajectories over time (14 out of 40 initial models).

						Model parameters and coefficient of variation																		Absolute reference points			Relative reference points		Coefficient of variation			Mohn's Rho			Sensitivity to initial values			Hindcast crossvalidation (MASE)			
Model number	Catch aggregation	Indices of abundance	n prior	r prior	B1/K prior	n	n.cv	bkfrac	bkfrac.cv	r	r.cv	K	K.cv	$alpha_{ac}$.est	$alpha_{ac}$.cv	$alpha_{IBTS}$.est	$alpha_{IBTS}$.cv	beta	beta.cv	q_{ac}.est	q_{ac}.cv	q_{IBTS}.est	q.cv	BMSY	FMSY	MSY	B/BMSY	F/FMSY	B/BMSY	F/FMSY	Abs(param correlations) > 0.95	B/BMSY	F/FMSY	Comments on retro	P50 dist.	P90 dist.	Prop. fail	Acoustic survey (5 runs)	Groundfish survey (3 runs)	Number not converged (order)	other comments
1	quarter	1-acoustic	Schaefer	None	BK50	2	0	0.42	0.56	1.26	0.71	45837	1.20	0.79	0.55			1.17	0.36	4.4E-05	1.35			NaN	0.37	NaN	1.52	0.02	0.82	2.35		0.06	0.13	retro-2 contrasts with the others; retrospective -5 not converged	0.47	193286.45	0.19	0.96	-	3 (5, 6, 7)	
2	quarter	1-acoustic and 2-groundfish	Thorson	None	BK20	1.05	0.26	0.13	0.54	0.68	0.88	24611	1.08	0.49	0.90	0.47	0.44	1.81	0.53	7.2E-05	4.8E-05	1.08	0.99	NaN	0.62	NaN	1.88	0.04	1.58	1.91		-0.08	0.26	retrospective -5 not converged	0.12	0.49	0.16	1.00	0.47	2 (6, 7)	Odd shape of the production curve
3	quarter	1-acoustic and 2-groundfish	Thorson	None	BK50	0.92	0.39	0.33	0.54	0.52	1.15	20525	0.98	0.53	1.10	0.53	0.52	1.48	0.49	7.4E-05	4.3E-05	0.83	0.78	7240	0.57	4113	2.59	0.04	2.01	1.98		-0.11	0.09		0.15	11327.14	0.26	1.13	0.68		
4	quarter	1-acoustic and 2-groundfish	Thorson	None	BK80	0.87	0.36	0.53	0.53	0.48	0.99	18658	0.98	0.54	1.18	0.52	0.46	1.39	0.4	7.4E-05	4.2E-05	0.84	0.78	6401	0.56	3555	2.98	0.03	1.98	2.04		-0.16	0.15		0.13	11469.48	0.23	1.16	0.72		
5	semester	1-acoustic	Schaefer	None	BK50	2	0	0.42	0.55	1.24	0.85	50828	1.35	0.79	0.63			1.07	0.46	3.9E-05	1.50			NaN	0.36	NaN	1.4	0.04	0.86	3.11		0.11	0.18		0.13	72941.05	0.06	1.07	-	1 (7)	
6	semester	1-acoustic	Schaefer	Thorson	BK50	2	0	0.42	0.55	1.79	0.26	40789	1.19	0.61	0.41			1.07	0.47	5.5E-05	1.19			NaN	0.74	NaN	1.28	0.04	0.89	3.01	MSY-K	0.18	0.24	retro-2 contrasts with the others	0.08	0.31	0.16	1.05	-	1 (7)	
7	semester	1-acoustic	Schaefer	None	BK80	2	0	0.68	0.55	1.24	0.73	47786	1.23	0.81	0.59			1.00	0.44	4.1E-05	1.37			NaN	0.37	NaN	1.44	0.04	0.83	2.94		0.12	0.17		0.29	136430.71	0.00	0.91	-	2 (6, 7)	
8	semester	1-acoustic	Schaefer	Thorson	BK80	2	0	0.69	0.55	1.78	0.26	38238	1.15	0.62	0.42			1.02	0.46	5.8E-05	1.14			NaN	0.73	NaN	1.31	0.04	0.87	2.86	MSY-K	0.19	0.19		0.14	0.44	0.16	1.07	-	1 (7)	
9	semester	1-acoustic and 2-groundfish	Schaefer	None	None	2.0	0.0	0.00	2.20	2.53	0.36	67843	1.82	0.43	0.67	0.48	0.36	1.70	0.49	4.7E-05	3.8E-05	1.56	1.41	NaN	2.64	NaN	0.96	0.03	1.03	2.16	MSY-K	0.08	0.12	Trajectories diverge to the past	0.25	16322.83	0.06	0.74	0.66		
10	semester	1-acoustic and 2-groundfish	Schaefer	Thorson	None	2.0	0.0	0.00	2.20	2.18	0.24	86715	1.99	0.47	0.73	0.44	0.33	1.72	0.5	3.7E-05	3.0E-05	1.65	1.51	NaN	1.37	NaN	0.97	0.03	1.03	2.39	MSY-K; q_{ac}-q_{IBTS}	0.08	-0.11		0.45	1.06	0.10	0.75	0.67		
11	semester	1-acoustic and 2-groundfish	Schaefer	None	BK20	2.0	0.0	0.10	0.53	2.58	0.48	25387	1.56	0.44	0.78	0.48	0.40	1.86	0.53	9.2E-05	5.7E-05	1.62	1.66	NaN	2.94	NaN	1.11	0.07	0.96	2.16		0.05	0.24		2.72	108623.70	0.13	0.89	0.66		
12	semester	1-acoustic and 2-groundfish	Schaefer	Thorson	BK20	2.0	0.0	0.11	0.52	2.13	0.26	34166	1.31	0.48	0.88	0.42	0.33	1.78	0.52	6.5E-05	4.0E-05	1.26	1.25	NaN	1.25	NaN	1.15	0.06	0.95	2.18	MSY-K	0.07	0.21	retro-2 contrasts with the others	0.09	0.35	0.03	0.89	0.68		
13	semester	1-acoustic and 2-groundfish	Schaefer	Thorson	BK50	2.0	0.0	0.26	0.54	2.09	0.26	37817	1.34	0.47	0.93	0.43	0.34	1.63	0.51	5.7E-05	3.2E-05	1.30	1.29	NaN	1.16	NaN	1.18	0.05	0.96	2.33	MSY-K	0.07	0.25		0.12	0.49	0.03	0.93	0.71		
14	semester	1-acoustic and 2-groundfish	Schaefer	Thorson	BK80	2.0	0.0	0.42	0.55	2.07	0.26	38702	1.33	0.48	0.98	0.43	0.34	1.53	0.5	5.4E-05	2.9E-05	1.29	1.28	NaN	1.14	NaN	1.20	0.05	0.96	2.36	MSY-K	0.07	0.27		0.12	0.53	0.10	0.94	0.73		

Table 3.2 – Anchovy 9a-west: Percentage of change of point estimates and coefficients of variation between Model 12 and each of the sensitivity test models.

Sensitivity test	Model	BBMSY	BBMSY.CV	FFMSY	FFMSY.CV	B.MoR	F.MoR	bkfrac	bkfrac.cv	r	r.cv	K	K.cv	$alpha_{ac}$.est	$alpha_{ac}$.cv	$alpha_{IBTS}$.est	$alpha_{IBTS}$.cv	beta	beta.cv	q_{ac}.est	q_{ac}.cv	q_{IBTS}.est	q.cv	P50 dist.	P90 dist.	Prop. fail	MASE Acoustic survey (5 runs)	MASE Groundfish survey (3 runs)
13	Down-weight acoustic 2019	9.6	-8.4	-33.3	-1.8	123.9	19.3	-9.1	0.0	-1.4	0.0	22.9	0.8	-25.0	-1.1	11.9	-3.0	1.1	-1.9	3.4	-1.0	-0.8	-0.6	-52.9	1.4	400.0	-7.7	0.000
14	Down-weight groundfish 2018 & acoustic 2019	8.7	-7.4	-33.3	-1.4	127.0	28.4	-9.1	0.0	-0.9	0.0	24.7	2.3	-27.1	-2.3	11.9	0.0	1.7	-1.9	2.6	-0.7	-0.1	0.0	-52.9	-1.4	300.0	-7.7	0.000
15	Down-weight groundfish 2017	0.9	-1.1	0.0	-1.8	118.6	-67.1	-9.1	0.0	2.3	0.0	-8.5	-1.5	0.0	-8.0	0.0	6.1	0.0	0.0	4.9	1.5	-0.7	-0.2	17.6	71.6	0.0	0.3	1.1
16	No beta prior	0.0	0.0	0.0	0.5	-5.7	-4.3	0.0	0.0	0.5	0.0	-1.6	0.8	0.0	0.0	0.0	0.0	3.4	3.8	1.7	1.7	1.3	1.4	100.0	80.3	200.0	-0.1	-0.2
17	No alpha prior	0.9	-1.1	0.0	0.9	10.0	-25.7	0.0	0.0	0.9	0.0	1.6	0.8	-4.2	-1.1	7.1	0.0	0.0	0.0	-0.8	-0.5	0.9	0.9	41.2	84.1	300.0	-0.8	-0.2
18	No alpha&beta priors	0.9	0.0	0.0	1.4	11.4	-30.5	0.0	0.0	0.9	0.0	0.0	2.3	-4.2	-2.3	7.1	0.0	3.9	3.8	0.9	1.2	2.3	2.3	105.9	71.0	200.0	-0.9	-0.3