Gadget for anchovy 9a South: Model description and results to provide catch advice and reference points (WGHANSA-1 2023)

Margarita María Rincón[a,*], Fernando Ramos[a], Jorge Tornero[a], Susana Garrido[b], Bjarki Elvarsson[c], Jamie Lentin[c]

[a] *Instituto Español de Oceanografía (IEO-CSIC), Centro Oceanográfico de Cádiz, Puerto pesquero, Muelle de Levante s/n, Apdo. 2609, 11006 Cádiz, Spain*
[b] *Instituto Portugues do Mar e da Atmosfera-IPMA, Av. Brasília, 6, 1449-006 Lisboa, Portugal*
[c] *Marine and Freshwater Research Institute, Fornubúdum 5 220 Hafnarfjördur, Iceland*

1. Background

The model specifications presented below correspond to those benchmarked in WKPELA 2018. The main difference is that results are presented now for the end of the second quarter of each year instead of be presented at the end of the fourth quarter. This responds to practical modifications in the definition of the assessment year, now it goes from July 1st to June 30th of the next year. Specific model assumptions for this year are presented in section 2.2 and 3, as well as estimated parameters after optimization in Table 2.

2. Model Description

Gadget is an age-length-structured model that integrates different sources of information in order to produce a diagnose of the stock dynamics. It works making forward simulations and minimizing an objective (negative log-likelihood) function that measures the difference between the model and data, the discrepancy is presented as a likelihood score for each time period and model component.

The general Gadget model description and all the options available can be found in Gadget manual (Begley, 2004) and some specific examples can be found in Taylor et al. (2007), Elvarsson et al. (2014) and WKICEMSE assessment for Ling (Elvarsson, 2017). The latest was used as a guide for this document.

The Gadget model implementation consists in three parts, a simulation of biological dynamics of the population (simulation model), a fitting of the model to observed data using a weighted log-likelihood function (observation model) and the optimization of the parameters using different iterative algorithms.

A list of the symbols used and estimated parameters is presented in Table 2 and a graph with the Gadget model structure presented in the last benchmark (WKPELA 2018) is available at Gadget structure graph.

*Corresponding author
Email address: margarita.rincon@ieo.csic.es (Margarita María Rincón)

2.1. Simulation model

The model consists of one stock component of anchovy (*Engraulis encrasicolus*) in the ICES subdivision, 9.a South-Atlantic Iberian waters, Gulf of Cádiz. Gadget works by keeping track of the number of individuals, $N_{a,l,y,t}$, at age $a = 0, \ldots, 3$, at length $l = 3, 3.5, 4, 4.5, \ldots, 22$, at year $y = 1989, \ldots, 2023$, and each year divided into quarters $t = 1, \ldots, 4$.. The last time step of a year involves increasing the age by one year, except for the last age group, which its age remains unchanged and the age group next to is added to it, like a 'plus group' including all ages from the oldest age onwards (Taylor et al., 2007).

Growth

The growth function is a simplified version of the Von Bertalanffy growth equation, defined in Begley (2004) as the LengthVBSimple Growth Function (*lengthvbsimple*). Length increase for each length group of the stock is given by the equation below:

$$\Delta l = (l_\infty - l)(1 - e^{k\Delta t}), \tag{1}$$

where Δt is the length of the timestep, $l_\infty = 19\ cm$ (fixed) is the terminal length and k is the growth rate parameter.

The corresponding increase in weight (in Kg) of the stock is given by:

$$\Delta w = a((l + \Delta l)^b - l^b), \tag{2}$$

with $a = 3.128958e^{-6}$ and $b = 3.277667619$ set as fixed and extracted from all the samples available in third and fourth quarters from 2003 to 2017. The growth functions described above calculate the mean growth for the stock within the model. In a second step the growth is translated into a beta-binomial distribution of actual growths around that mean with parameters β and n. The first is fitted by the model as described in Taylor et al. (2007) and the second represents the number of length classes that an individual is allowed to grow in a quarter and it is fixed and equal to 5.

Initial abundance and recruitment

Stock population in numbers at the starting point of the simulation is defined as:

$$N_{a,l,1,1} = 10000\nu_a q_{a,l}, \quad a = 0, \ldots, 3, l = 3, \ldots, 20$$

Where ν_a is an age factor to be calculated by the model and $q_{a,l}$ is the proportion at lengthgroup l that is determined by a normal density with a specified mean length and standard deviation for each age group. Mean length at age (μ_a) and its standard deviation (σ_a) were extracted from all the data available from 1989 to 2018 including three surveys that are not included in the model: ARSA, ECOCADIZ-RECLUTAS and SAR survey (See table 2). The mean weight at age for this initial population is calculated by multiplying a reference weight corresponding to the length by a relative condition factor assumed as 1. This reference weight at length was calculated using the formula $w = al^b$, with a and b as defined before. In Gadget files this was specified as a normal condition distribution (*Normalcondfile*).

@

Similarly to the process of calculate the initial abundance described above, the recruitment specifies how the stock will be renewed. Recruits enter to the age 0 population at quarters 2, 3, 4 (because of the Gadget order of calculations for each time step this is equivalent to have recruitment one quarter later, i.e. in quarters 3,4 and 1 of the next year) of all years, respectively, as follows:

$$N_{0,l,y,t} = p_{l,t} R_{y,t}, \quad t = 2,3,4, l = 3,\ldots,15,$$

where $R_{y,t}$ represents recruitment at year y and quarter t, and $p_{l,t}$ the proportion in lengthgroup l that is recruited at quarter t which is sampled from a normal density with mean (μ) and standard deviation (σ_t) calculated by the model. The mean weight for these recruits is calculated by multiplying the reference weight corresponding to the length by a relative condition factor assumed as 1. Reference weight at age was the same used to calculate the initial population mean weight at age explained above. In Gadget files this was specified also as a normal condition distribution (*Normalcondfile*).

Fleet operations

In the model the fleets act as predators. There are three fleets inside the model: two for surveys (ECOCADIZ acoustic survey and PELAGO acoustic survey) and one for commercial landings including all fleets: Spanish purse-seine, trawlers, Portuguese purse-seine, and others. The main fleet is Spanish purse-seine representing more than a 90 % of all the catches from 2001 to 2016 and more than a 80 % from 1989 to 2000. It is also the only fleet with a lenght distribution available, then we decide to include all commercial reported data in the same fleet which is mostly the Spanish purse-seine.

Surveys fleets are assumed to remove 1 Kg in each of the quarters when the surveys take place while the commercial fleet is assumed to remove the reported number of individuals each quarter. This total amount of biomass (for the surveys) or numbers (for the commercial fleet) landed is then split between the length groups according to the equations 3 and 4 respectively, as follows:

$$C_{l,y,t} = \frac{E_{y,t} S_{l,T} N_{l,y,t} W_l}{\sum_l S_{l,T} N_{l,y,t} W_l}, \tag{3}$$

and

$$C_{l,y,t} = \frac{E_{y,t} S_{l,T} N_{l,y,t}}{\sum_l S_{l,T} N_{l,y,t}}, \tag{4}$$

where $E_{y,t}$ represents biomass landed (in Kg) at year y and quarter t in equation 3 and numbers landed in equation 4, W_l corresponds to weight at length and $S_{l,T}$ represents the suitability function that determines the proportion of prey of length l that the fleet is willing to consume during period $T, T = 1,2,3$ where $T = 1$ corresponds to the period 1989-2000, $T = 2$ to 2001-**2022** and $T = 3$ to 1989-**2022**.

For this model the suitability function chosen for the fleet and surveys is specified in Gadget manual as an ExponentialL50 function (*expsuitfuncl50*), and it is defined as follows:

$$S_{l,T} = \frac{1}{1 + e^{\alpha_T (l - l_{50,T})}} \tag{5}$$

where $l_{50,T}$ is the length of the prey with a 50% probability of predation during period T and α_T a parameter related to the shape of the function, both parameters are estimated from the data within the Gadget model. The whole model time period (1989-**2022**) has been splited into two different periods for suitability parameters of the commercial fleet because of changes in size regulation for the fishery around 1995 that become effective around 2001.

2.2. Observation model

Data are assimilated by Gadget using a weighted log-likelihood function. The model uses as likelihood components two biomass survey indices: ECOCADIZ acoustic survey and PELAGO acoustic survey; age - length keys from the commercial fleet (Spanish purse-seine), PELAGO survey and the ECOCADIZ survey; and length distributions for the commercial fleet, PELAGO and ECOCADIZ surveys (see Table 2.2 for a detailed description of the likelihood data used in the model).

Biomass Survey indices

The survey indices are defined as the total biomass of fish caught in a survey. The survey index is compared to the modelled abundance using a log linear regression with slope equal to 1 (*fixedslopeloglinearfit*), as follows:

$$\ell = \sum_t (\log(I_{y,t}) - (\alpha + \log(N_{y,t}))^2 \tag{6}$$

where $I_{y,t}$ is the observed survey index at year y and quarter t and $N_{y,t}$ is the corresponding population biomass calculated within the model. Note that the intercept of the log-linear regression, $\alpha = \log(q)$, with q as the catchability of the fleet (i.e $I_{y,t} = qN_{y,t}$).

Catch distribution

Age-length distributions are compared using l lengthgroup at age a and time-step y, t for both, commercial and survey fleets with a sum of squares likelihood function (*sumofsquares*):

$$\ell = \sum_y \sum_t \sum_l (P_{a,l,y,t} - \pi_{a,l,y,t})^2 \tag{7}$$

where $P_{a,l,t,y}$ is the proportion of the data sample for that time/age/length combination, while $\pi_{a,l,t,y}$ is the proportion of the model sample for the same combination, as follows:

$$P_{a,l,t,y} = \frac{O_{a,l,y,t}}{\sum_a \sum_l O_{a,l,y,t}} \tag{8}$$

and

$$\pi_{a,l,t,y} = \frac{N_{a,l,y,t}}{\sum_a \sum_l N_{a,l,y,t}}, \tag{9}$$

where $O_{a,l,y,t}$ corresponds to observed data.

When only length or age distribution is available. It is compared using equation 7 described above but considering all ages or all lengths, respectively.

Understocking

If the total consumption of fish by all the predators (fleets in this case) amounts to more than the biomass of prey available, then the model runs into ”understocking”. In this case, the consumption by the predators is adjusted so that no more than 95% of the available prey biomass is consumed, and a penalty, given by the equation 10 below, is applied to the likelihood score obtained from the simulation (Stefansson 2005, sec 4.1.)

$$\ell = \sum_t U_t^2 \tag{10}$$

where U_t is the understocking that has occurred in the model for that timestep.

Penalties

The BoundLikelihood likelihood component is used to give a penalty weight to parameters that have moved beyond the bounds in the optimisation process. This component does specify the penalty that is to be applied when these bounds are exceeded.

$$\ell_i = \begin{cases} lw_i(val_i - lb_i)^2 & \text{if } val_i < lb_i \\ uw_i(val_i - ub_i)^2 & \text{if } val_i > ub_i \\ 0 & otherwise \end{cases}$$

Where $lw_i = 10000$ and $uw_i = 10000$ are the weights applied when the parameter exceeds the lower and upper bounds, respectively, val_i is the value of the parameter and, lb_i and ub_i are the lower and upper bounds defined for the parameter.

2.3. Order of calculations

The order of calculations is as follows:

1. **Printing**: model output at the beginning of the time-step
2. **Consumption**: by the fleets
3. **Natural mortality**
4. **Growth**
5. **Recruitment**: new individuals enter to the population
6. **Likelihood comparison**: Comparison of estimated and observed data, a likelihood score is calculated

7. **Printing**: model output at the end of the time-step
8. **Ageing**: if this is the end of year the age is increased

Because of this order of calculations the time step of indexes, age-length keys and length distributions of the surveys are defined in Gadget a quarter before.

2.4. Implementation, weighting procedure

Input data (Likelihood files) were prepared for Gadget format using the *mfdb* R package (Lentin, 2014), running and weighting procedures were implemented in R with the gadget.iterative function from *Rgadget* package. This function follows the approach presented in Taylor et al. (2007) and in the appendix of Elvarsson et al. (2014) based on the iterative reweighting scheme of Stefánsson (1998) and Stefansson (2003), which is summarized as follows:

Let $\mathbf{w_r}$ be a vector of length L with the weights of the likelihood components (excluding understocking and penalties) for the run r, and $SS_{i,r}, i = 1, \ldots, L$, the likelihood score of component i after run r. First, a Gadget optimization run is performed to get a likelihood score ($SS_{i,1}$) for each likelihood component assuming that all components have a weight equal to one, i.e., $\mathbf{w_1} = (1, 1, \ldots, 1)$. Then, a separated optimization run for each of the components (L optimization runs) is performed using the following weight vectors:

$$\mathbf{w_{i+1}} = (1/SS_{1,1}, \ldots, (1/SS_{i,1}) * 10000, 1/SS_{i+1,1}, \ldots, 1/SS_{L,1}), i = 1, \ldots, L.$$

Resulting likelihood scores $SS_{i,i+1}$ are then used to calculate the residual variance, $\hat{\sigma}_i^2 = SS_{i,i+1}/df^*$ for each component, that is used to define the final weight vector as

$$\mathbf{w} = (1/\hat{\sigma}_1^2, \ldots, 1/\hat{\sigma}_L^2).$$

Where degrees of freedom df^* are approximated by the number of non-zero data points in the observed data for each component. Finally, the total objective function is the sum of all likelihoods components multiplied by their respective weights according to the vector $\mathbf{w}$.

In order to assign weights to the individual likelihood components (See table 2.2) in the procedure described above, all the survey indices were grouped together.

2.5. Initial parameters and optimization

Initial parameter values with their boundaries and settings for the optimising algorithms can be found in initial values for parameters file and optimization file. The optimization algorithms converged in individual and weighted runs.

3. Remarkable Model Assumptions (in bold the terms associated to the more recent assumptions)

- Due to lack of information of length distributions and Age-length keys for commercial catches in the first and second quarter of 2020, for this year assessment, the length distribution of those quarters in year 2020 was approximated using the joint distribution of 2018 and 2019. For the Age-length key the one for the PELAGO 2020 survey was used.

- Due to discrepancies on mean lenght and weight at age in PELAGO survey for 2023 a crossvalidation for age composition was required. This crossvalidation reveals some missestimations in the otolith reading suggesting that more analysis is needed to agreed on the definitive age composition. For this reason age distribution of PELAGO survey in 2023 was removed from the model.

- Due to technical problems there are no data available for ECOCADIZ survey in 2021 and 2022.

- The model was implemented quarterly from 1989 to the second quarter of **2023**.

- All commercial fleets where grouped into only one from 1989 to **2023** second quarter: The Spanish purse-seine. The Spanish purse-seine which represents more than a 90 % of all the catches from 2001 to 2016 and more than a 80 % from 1989 to 2000. It is also the only fleet with a lenght distribution available. For the first two quarters of year **2023**, provisional catches estimations of Spanish (until May 18th) purse-seine fleet were used and catches for June were estimated as the **39%** of January to May catches based on historical records from 2009 to **2022**. There were not any catches for Portuguese purse-seine in these two quarters.

- It was decided to include also discards (available from 2014 onwards) in WGHANSA-1-2020. This decision was taken because they were already accounted for some years in the previous assessments to 2020 but we did not notice about that. Since then we include discards in catches data.

- The parameters for weight-length relationship equation ($w = al^b$,) were assumed fixed and defined as $a = 3.128958e^{-6}$ and $b = 3.277667619$. Those values were calculated from all the samples available in third and fourth quarters from 2003 to 2017.

- Natural mortality at age was also considered fixed with $M_0 = 2.21$ and $M_1, M_2, M_3 = 1.3$,.

- There was a minimum landing size restriction from 1995, that were only effective until 2001. As a consequence it was neccesary to define different suitability parameters for two different periods. One from 1989 to 2000, and the other from 2001 to **2022**.

- Age 0 individuals were removed for **all** the data input corresponding to ECOCADIZ survey. It was noticed that age 0 was not removed from the length distribution in the assessments prior to 2021.

- It was noticed that the length distribution for year 2020 in ECOCADIZ survey was not included in the model used for 2021 assessment. We include that missing information in the model described in this document.

- Recruits enter to the age 0 population at quarters 2, 3 and 4 (because of the Gadget order of calculations for each time step this is equivalent to have recruitment one quarter later, i.e. in quarters 3,4 and 1 of the next year) of all years except the last year, because at the end of June there are no recruits (zero age individuals). Then, biomass and abundance estimates at the end of the second quarter need to be corrected removing age 0 individuals.

4. Natural mortality selection

Natural mortality selection is justified by the following arguments:

- Natural mortality was preferred to be selected from classical indirect formulations based on life history parameters. For it we used the R package *FSA* to obtain empirical estimates of natural mortality.

- For the estimation of the natural mortality rate, the Von Bertalanffy growth parameters and the maximum age that the species can live were used. Growth parameters of the Von Bertalanffy function were taken from Bellido et al. (2000) ($l_\infty = 18.95, k = 0.89, t_0 = -0.02$), and for the maximum observed age, we explored a range from age 3 to 5, but finally age 4 was considered adequate. A total of 13 estimators were produced using the R package *FSA* and the a value of $M = 1.3$ was undertaken (midway between the median and the mean of the available estimates for Agemax=4).

- Currently is generally accepted that Natural mortality may decrease with age, as far as it presumed to be particularly greater at the juvenile phase. It was agreed to adopt for the adult ages of anchovy (ages 1 to 4) the constant natural mortality estimated before (1.3), but for the juveniles (age 0) a greater one in proportion to the ratio of natural mortality at ages 0 and 1 (M_0/M_1) resulting from the application of the Gislason et al. (2010) method for modelling natural mortality as a function of the growth parameters. For it we used four vectors of length-at-age: derived from the Von Bertalanffy growth function in Bellido et al. (2000) for ages 1-5, from the ECOCADIZ-RECLUTAS survey for ages 0-3, the average of the length-at-age in the catches from 1987 to 2016 and the average of the length-at-age in the catches from 2007 to 2016. There was no major basis to select one or the other, we directly choosed the pattern shown by the ECOCADIZ-RECLUTAS data just because it seemed to be smoothest one (particularly for age 1 onwards as presumed here). The ratio M_0/M_1 is 2.722670/ 1.595922 = 1.7. Therefore $M_0 = 1.3 * 1.7 = 2.21$.

- In summary for anchovy 9a South, the adopted natural mortality by ages are $M_0 = 2.21, M_1 = 1.3$ and $M_2^+ = 1.3$ (similar at any older age).

5. Fit to data

A summary of likelihood scores is presented in Figure 1 while a comparison of estimated versus observed data is summarized in the following Figures:

Length distributions

- Figure 2: Length distribution of the commercial fleet.

- Figure 3: Length distribution of the ECOCADIZ acoustic survey.

- Figure 4: Length distribution of the PELAGO acoustic survey.
- Figure 5: Summary of residuals for length distributions.

Age distributions

- Figure 6: Age distribution of the commercial fleet.
- Figure 7: Age distribution of the ECOCADIZ acoustic survey.
- Figure 8: Age distribution of the PELAGO acoustic survey.
- Figure 9: Summary of residuals for age distributions.

Biomass survey indices fit

- Figure 10: Summary of biomass survey indices fit.

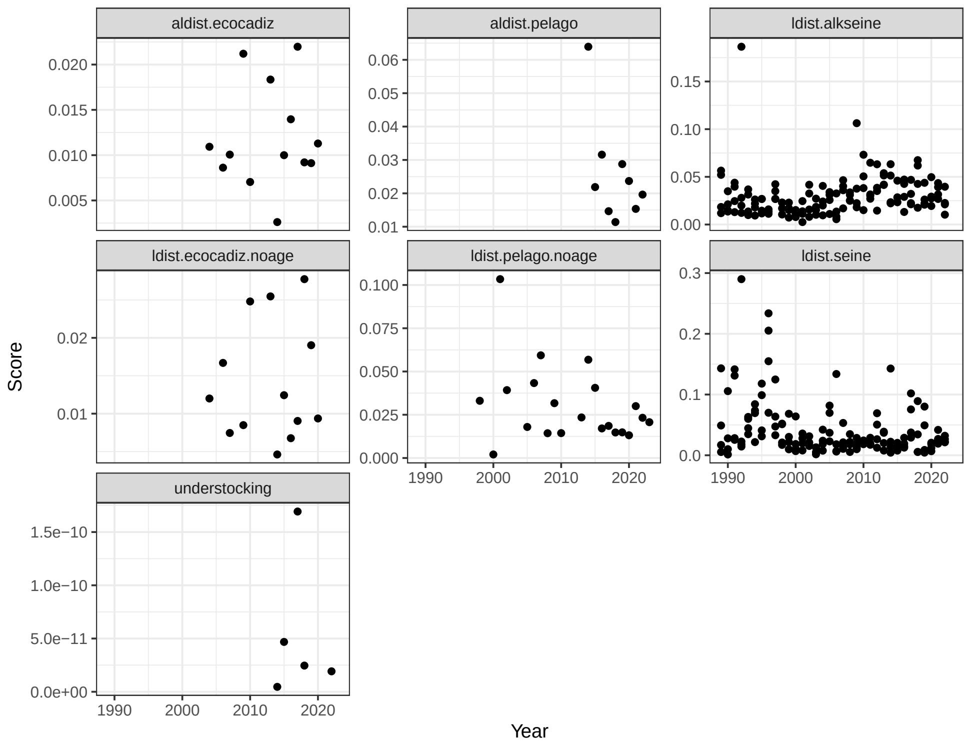

Figure 1: Likelihood scores for age-length key of ECOCADIZ survey, PELAGO survey and commercial landings (Upper panel) and length distribution of ECOCADIZ survey, PELAGO survey and landings. Dots represent the score for each quarter.

Index	
a	Age, $a = 0,\ldots,3$
l	Length, $l = 3,3.5,4,4.5,\ldots,22$
y	Years, $y = 1989,\ldots,2023$
t	Quartely timestep, $t = 1,\ldots,4$
T	$T = 1$ for period 1989-2000, $T = 2$ for period 2001-2022
Parameters	
Fixed	
a	Parameter of weight-length relationship $w = al^b$, $a = 3.128958 \times 10^{-6}$
b	Parameter of weight-length relationship $w = al^b$, $b = 3.277667619$
μ_a	Initial population mean length at age
	$\mu_0 = 9.99, \mu_1 = 12.1, \mu_2 = 15.2, \mu_3 = 16.1$
σ_a	Initial population standard deviation for length at age
	$\sigma_0 = 0.836, \sigma_1 = 0.5, \sigma_2 = 1, \sigma_3 = 1.2$
M_a	Natural mortality, $M_0 = 2.21, M_1 = 1.3, M_2 = 1.3, M_3 = 1.3$
n	Maximum number of length classes that an individual is supposed to grow $n = 5$
Estimated	
l_∞	Asympthotic length, l_∞=29.1744
k	Annual growth rate, $k = 0.0831751$
β	Beta-binomial parameter, $\beta = 5000$
ν_a	Age factor, $\nu_0 = 120000, \nu_1 = 118000,$
	$\nu_2 = 0.0601, \nu_3 = 1.25e-07$
μ	Recruitment mean length, $\mu = 9.86671$
σ_t	Recruitment length standard deviation by quarter, $\sigma_2 = 2.98305$, $\sigma_3 = 1.67904$, $\sigma_4 = 4$
$l_{50,T}$	Length with a 50% probability of predation during period T,
	$l_{50,1}^{seine} = 10.6$, $l_{50,2}^{seine} = 10.9$, $l_{50,3}^{ECO} = 12.8$, $l_{50,3}^{PEL} = 14.3$
α_T	Shape of function, $\alpha_1^{seine} = 0.407$, $\alpha_2^{seine} = 0.865$, $\alpha_3^{ECO} = 1.41$, $\alpha_3^{PEL} = 0.459$
Observed Data	
$E_{y,t}$	Number or biomass landed at year y and quarter t
W_l	Weight at length
$I_{y,t}$	Observed survey index at year y and quarter t
$P_{a,l,y,t}$	Proportion of the data sample over all ages and lengths for timestep/age/length combination
$O_{a,l,y,t}$	Observed data sample for time/age/length combination
$x_{a,y,t}$	Sample mean weight from the data for the timestep/age combination
Others	
Δl	Length increase
Δw	Weight increase
Δt	Length of timestep
$N_{a,l,y,t}$	Number of individuals of age a, length l in the stock at year and quarter y and t, respectively.
$q_{a,l}$	Proportion in lengthgroup l for each age group
$R_{y,t}$	Recruitment at year y and quarter t
$p_{l,t}$	Proportion in lengthgroup l that is recruited at quarter t
$C_{l,y,t}$	Total amount in biomass landed by surveys and in number caught by commercial fleet (discards 2014-2019)
$S_{l,T}$	Proportion of prey of length l that the fleet/predator is willing to consume during period T
$\pi_{a,l,y,t}$	Proportion of the model sample over all ages and lengths for that timestep/age/length combination
$\mu_{a,y,t}$	Mean length at age for the timestep/age combination
U_t	Understocking for timestep t
lw_i and uw_i	Weights applied when the parameter exceeds the lower or upper bound
lb_i and ub_i	Lower and upper bound defined for the parameter
val_i	Value of the parameter

Table 1: List of Symbols used in model specification and parameter estimates after optimization

Data source	type	Timespan	Likelihood function
Commercial catches	Length distribution	All quarters, 1989-**2022**	See eq. 7
(discards from 2014 onwards)	Age-length key	All quarters, 1989-**2022**	See eq. 7
ECOCADIZ acoustic survey	Biomass survey indexes	Second quarter 2004, 2006	see eq. 6
		third quarter 2007, 2009, 2010, 2013-2020	
	Length distribution	Second quarter 2004, 2006	see eq. 7
		third quarter 2007, 2009, 2010, 2013-2020	
	Age-length key	Second quarter 2004, 2006	see eq. 7
		third quarter 2007, 2009, 2010, 2013-2020	
PELAGO acoustic survey	Biomass survey indexes	First quarter 1999, 2001-2003	see eq. 6
		second quarter 2005-2010 and 2013-2023	
	length distribution	First quarter 1999, 2001-2003	see eq. 7
		second quarter 2005-2010, 2013-2023	
	Age-length key	second quarter 2014-2023	see eq. 7

Table 2: Overview of the likelihood data used in the model. Important remark: Due to lack of information of length distributions and Age-length keys for commercial catches in the first and second quarter of 2020, the length distribution was approximated using the joint distribution of 2018 and 2019 and the Age-length key used was the one for the PELAGO 2020 survey.

Figure 2: Comparison between observed and estimated catches length distribution. Black lines represent estimated data while gray lines represent observed data

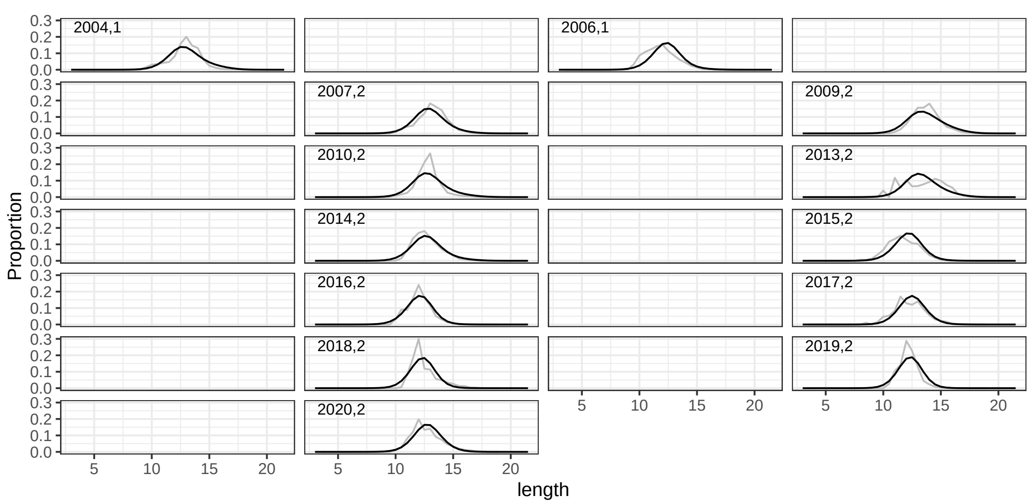

Figure 3: Comparison between observed and estimated catches length distribution for ECOCADIZ survey. Black lines represent estimated data while gray lines represent observed data

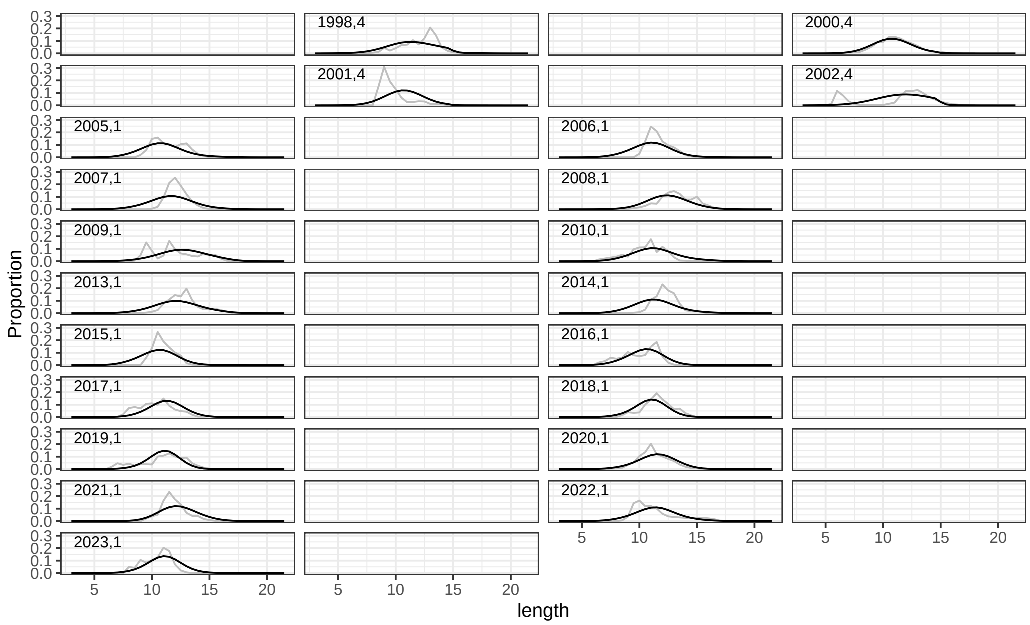

Figure 4: Comparison between observed and estimated catches length distribution for PELAGO survey. Black lines represent estimated data while gray lines represent observed data

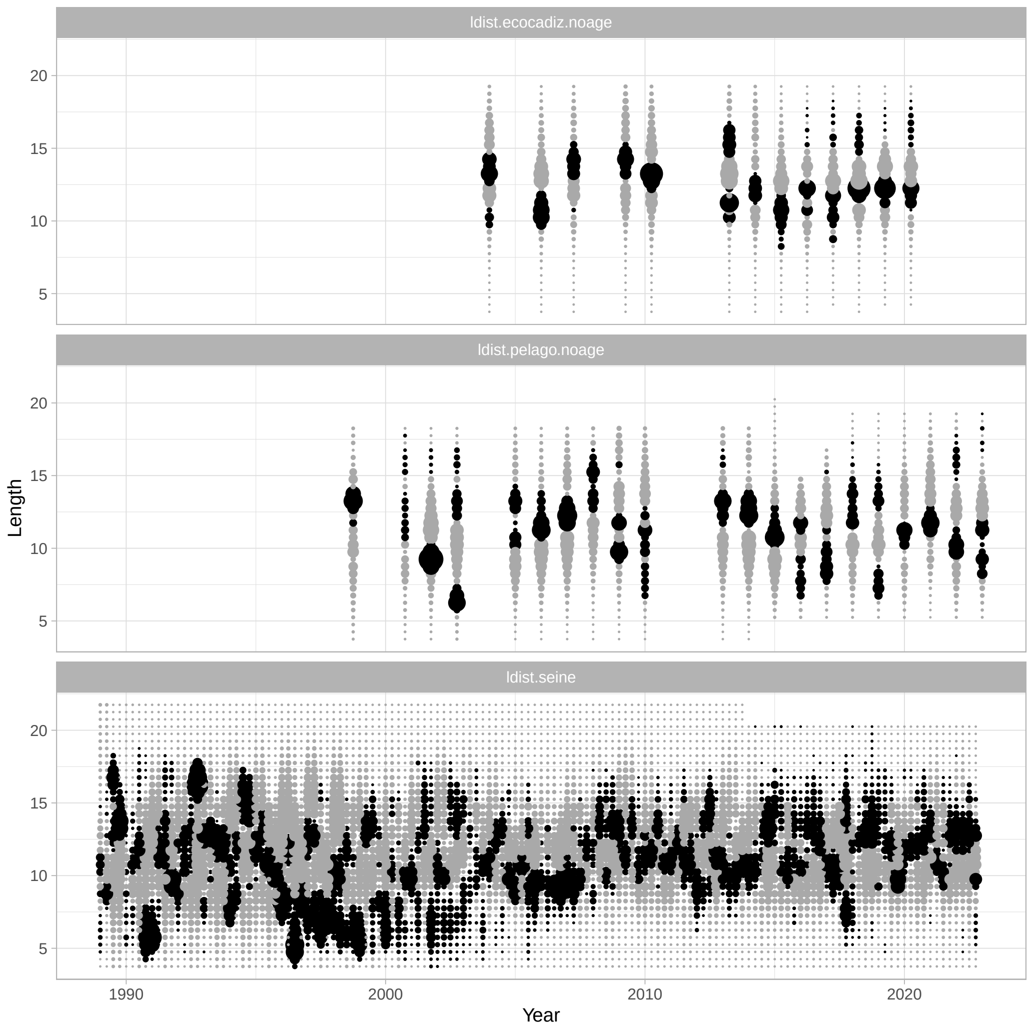

Figure 5: Standardised residual plots for the fitted length distribution from the ECOCADIZ survey, PELAGO survey and commercial landings. Black points denote a model underestimate and gray points an overestimated. The size of the points denote the scale of the standardised residual.

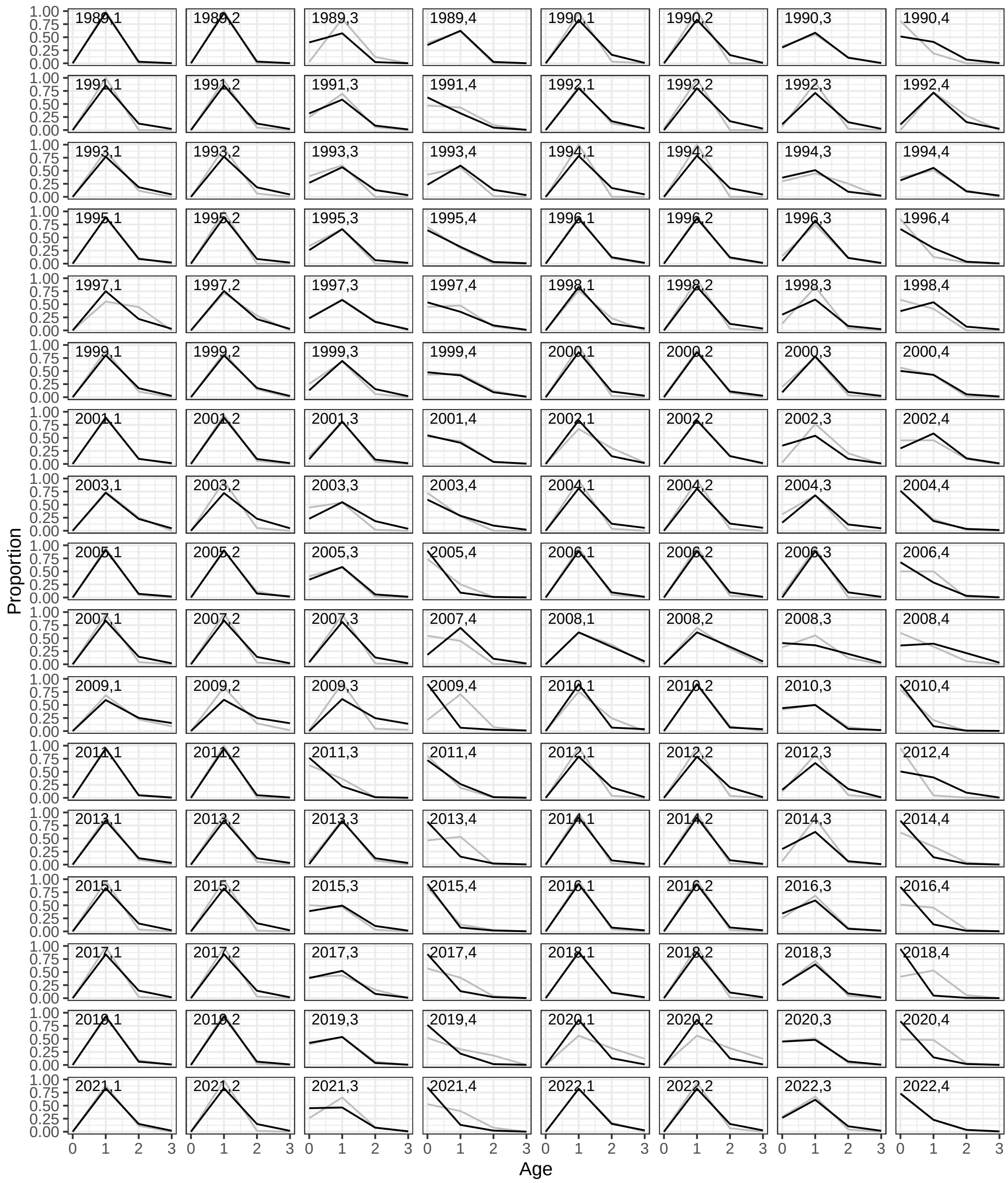

Figure 6: Comparison between observed and estimated catches age distribution. Black lines represent estimated data while gray lines represent observed data.

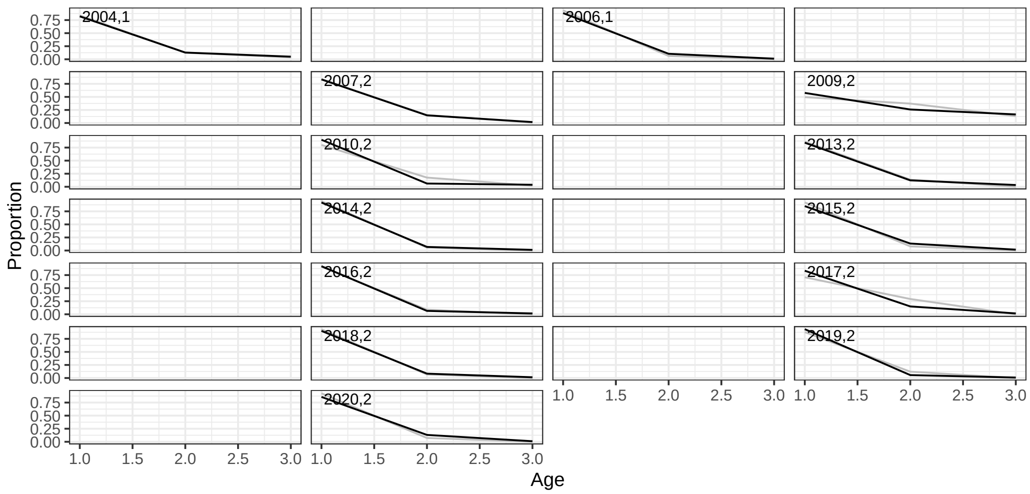

Figure 7: Comparison between observed and estimated ECOCADIZ survey age distribution. Black lines represent estimated data while gray lines represent observed data.

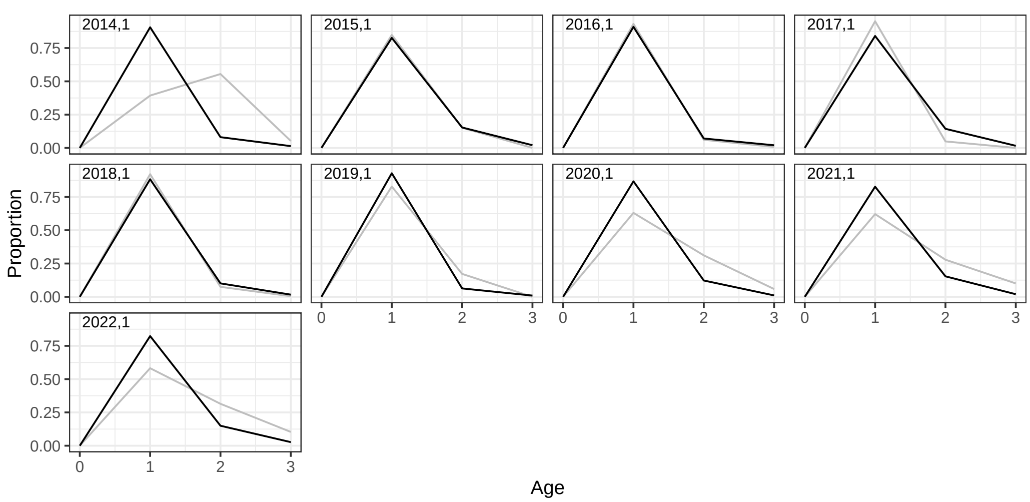

Figure 8: Comparison between observed and estimated PELAGO survey age distribution. Black lines represent estimated data while gray lines represent observed data.

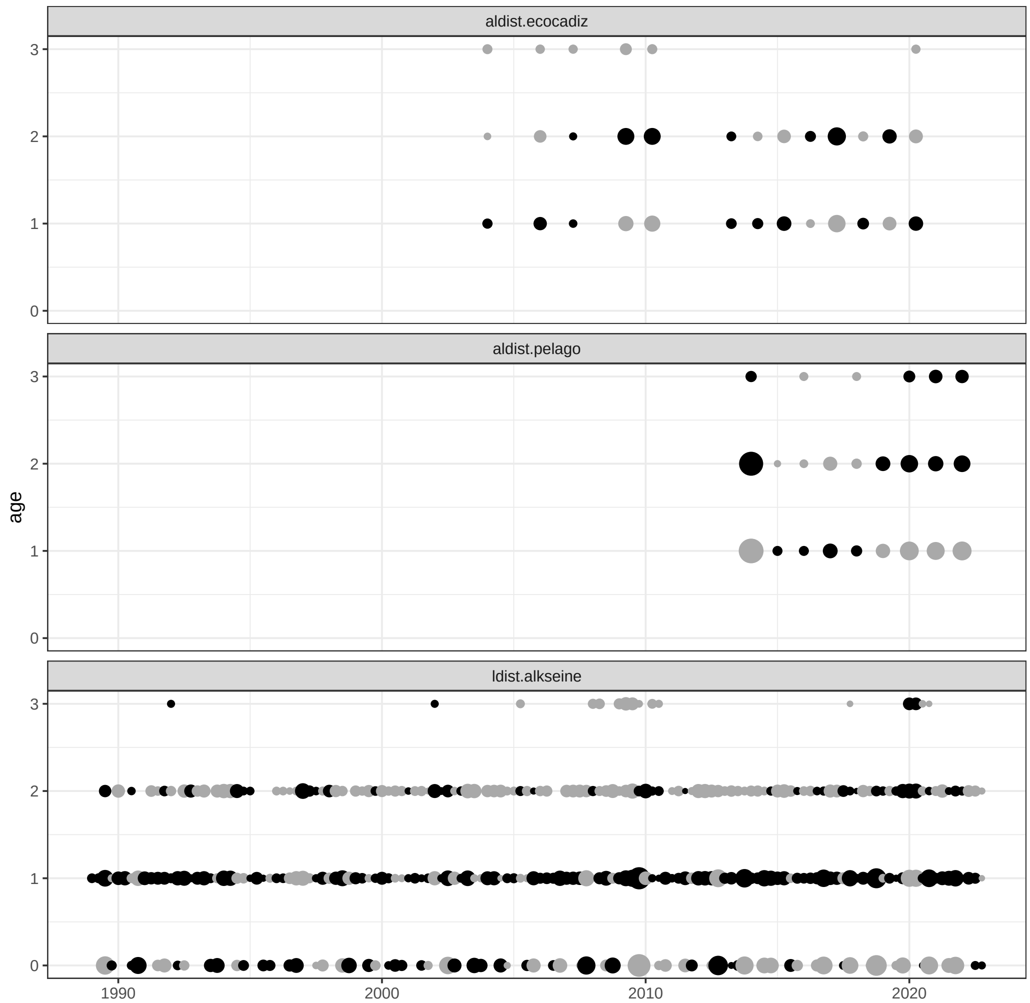

Figure 9: Standardised residual plots for the fitted age distribution from the ECOCADIZ survey, PELAGO survey and commercial fleet. Black points denote a model underestimate and gray points an overestimated. The size of the points denote the scale of the standardised residual.

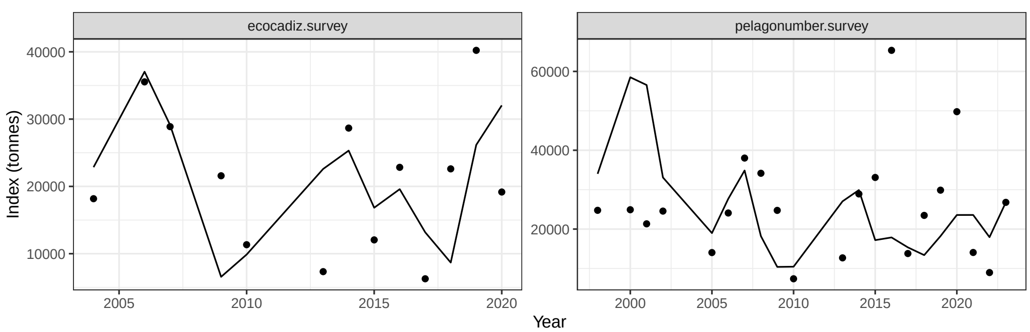

Figure 10: Comparison between observed and estimated survey indices. Black points represent observed data while black line represent estimated data

6. Model estimates

Parameter estimates after optimization are presented in Table 2. Detailed model outputs are available in Results folder on TAF repository, where each file corresponds to the following description:

- sidat: Model fit to the surveyindices
- suitability: Model estimated fleet suitability
- stock.recruitment: Model estimated recruitment
- res.by.year: Results by year
- catchdist.fleets: Data compared with model output for the length and age-length distributions
- stock.full: Modeled abundance and mean weight by year,step, length and stock
- stock.std: Modeled abundance, mean weight, number by age consumed by the fleet, stock and year
- stock.prey: Consumption of the fleet by length, year and step
- fleet.info: Information on catches, harvest rate and harvestable biomass by fleet, year and step
- params: parameter values used for the fit

6.1. Catchability

Figure 11 shows the catchability estimated by the model for the different surveys indices

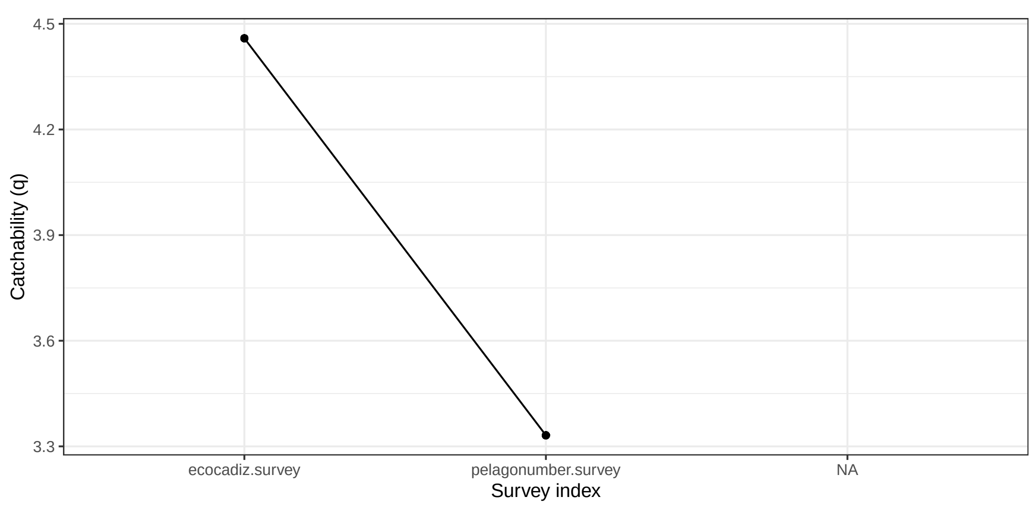

Figure 11: Estimated catchability parameters for the different survey indices

6.2. Estimated age composition

Figure 12 shows the estimated age composition of the population.

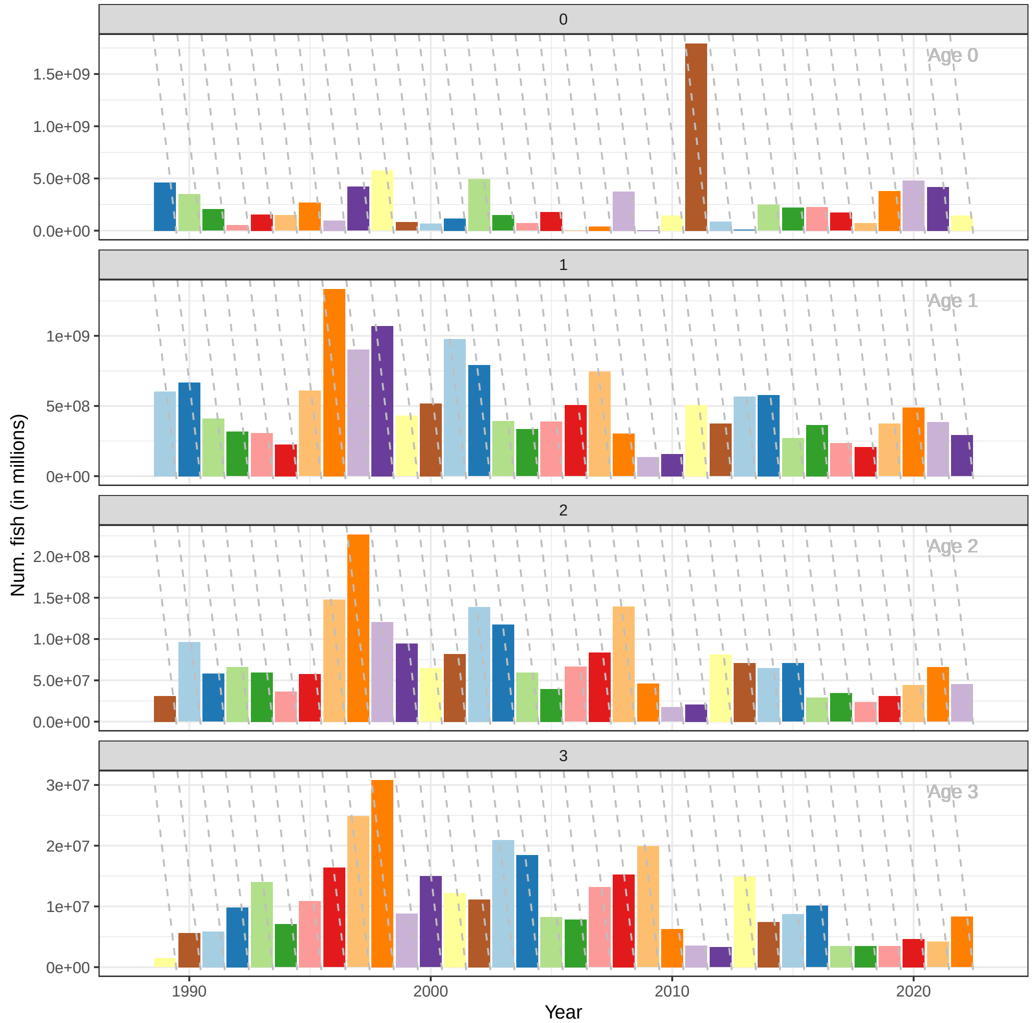

Figure 12: Estimated age composition of the population at the end of the second quarter for each year

6.3. Suitability

Figure 13 shows the fleet suitability functions estimated by the model for the commercial fleet and different surveys

6.4. Abundance, recruitment and Fishing mortality

Figure 14 presents model annual estimates for biomass, abundance (removing age 0 individuals to be accurate with the time of the assessment, see section 3 above for a detailed explanation), recruitment, fishing mortality and catches **at the end of the second quarter of each year**. Figure 15 shows annual estimates for biomass of individuals of age 1+ at the end of the second quarter of each year. Due to some inconsistencies in the maturity ogives not noticed during WKPELA 2018, we assume that all individuals with age 1 or higher (B_1+), are mature i.e. these abundance estimates result equivalent to spawning stock biomass estimates.

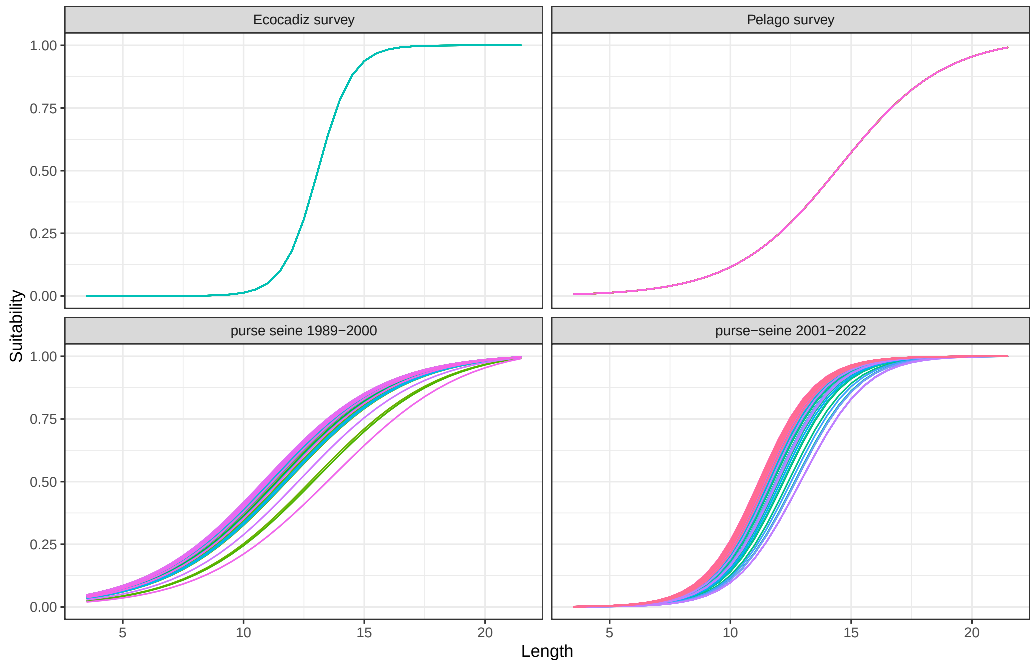

Figure 13: Estimated fleet suitability functions for the commercial fleet and different surveys.

6.5. Comparison with the last two years estimated time series

A comparison with the last two years estimated time series is presented in Figure 16. The pink line represents the current year estimated time series (the one estimated by the model described in this document), the green line, the estimated in 2022 and the blue line, the estimated in 2021. Trends are consistent even considering that each year the model updates the last values with the new information available.

7. Reference points

The methodology applied was the same decided in WKPELA 2018 (page 286 of WKPELA 2018 report (ICES, 2018)) following ICES guidelines for calculation of reference points for category 1 and 2 stocks and the report of the workshop to review the ICES advisory framework for short lived species ICES WKMSYREF5 2017 (ICES, 2017).

According to the above ICES guidelines and the S-R plot characteristics (Figure 17), this stock component can be classified as a “stock type 5” (i.e. stocks showing no evidence of impaired recruitment or with no clear relation between stock and recruitment (no apparent $S - R$ signal)). According to this classification, B_{lim} estimation is possible according to the standard method and it is assumed to be equal to $Bloss$ ($B_{lim} = B_{loss}$). For **2023** the value of B_{loss} for the 9a South anchovy corresponds to the estimated SSB in **2010** (1226.13 t), hence B_{lim} is set at 1226.13 t and the relative B_{lim} (divided by the mean value of B_1+) results equal to 0.286. Note that due to some inconsistencies in the maturity ogives used in WKPELA2018, age 1+ individuals (B_1+) are assumed as mature i.e. B_1+ class is equivalent to Stock Spawning Biomass (SSB) (see subsection 6.4 above).

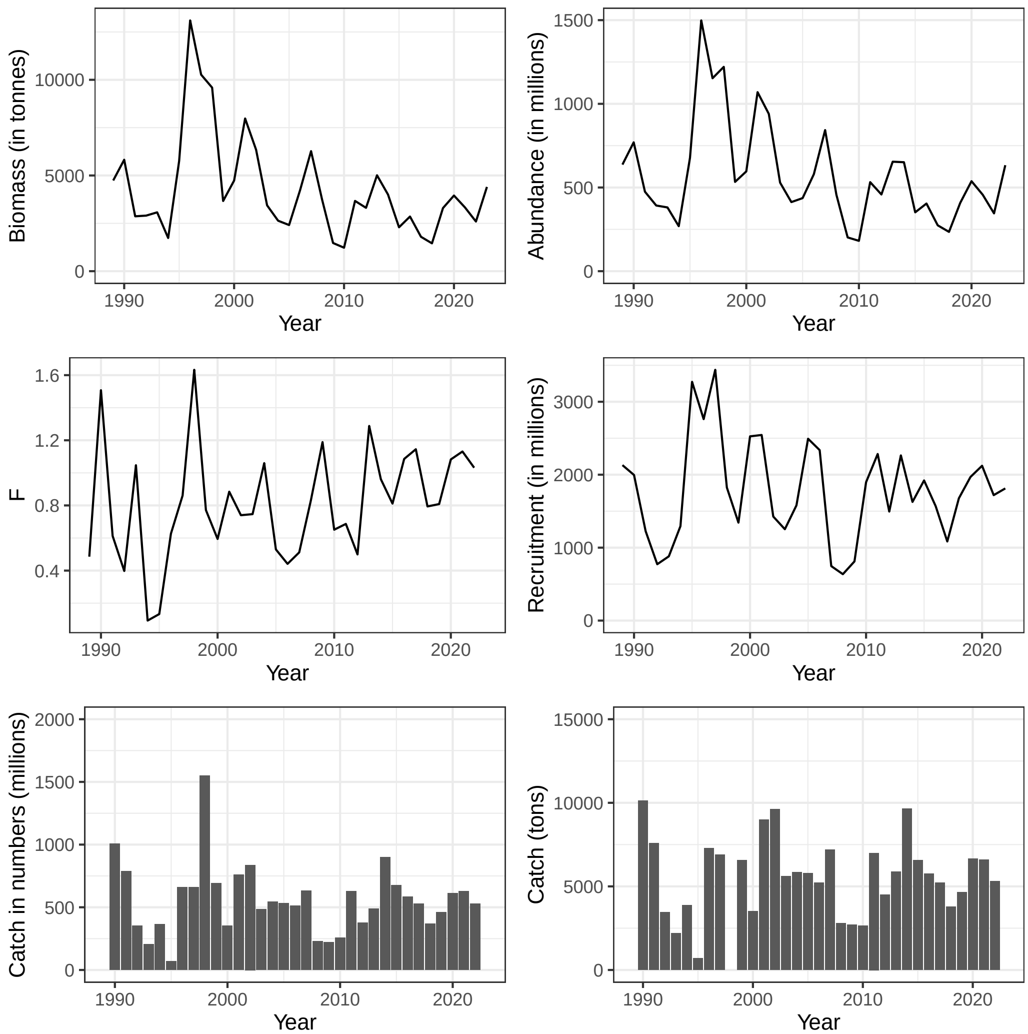

Figure 14: Annual catches time series (in numbers and biomass) compared with annual model estimates for abundance of individuals with more than one year of age(in numbers and biomass) recruitment and fishing mortality. Measures were summarized at the end of June each year, assuming that a year starts in July and ends in June of the next year. Recruitment was calculated including all the recruits of the previous year according to calendar year

ICES recommends to calculate B_{pa} as follows:

$$B_{pa} = e^{(1.645\sigma)} B_{lim},$$

where σ is the estimated standard deviation of $ln(SSB)$ in the last year of the assessment, accounting for the uncertainty in SSB for the terminal year. If σ is unknown and for short living species, as it is in our case, it can be assumed that $\sigma = 0.30$ (see page 34 of ICES WKMSYREF5 2017 report (ICES, 2017)), then $B_{pa} = e^{(1.645\sigma)} B_{lim} = 1.64 B_{lim}$. According to this B_{pa} is set at 2010.8532 t.

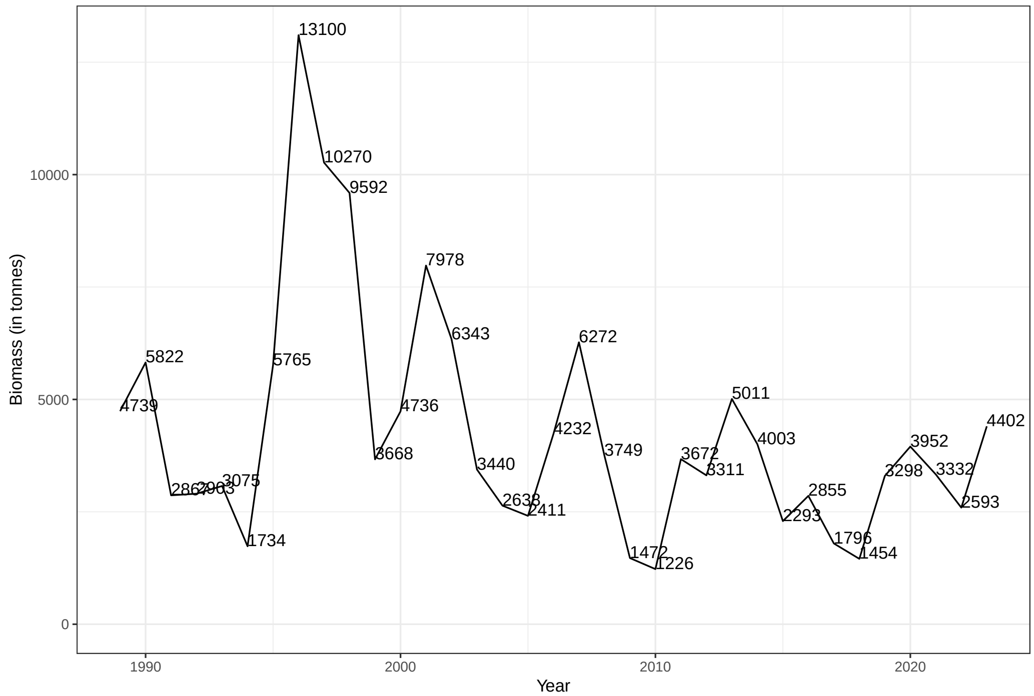

Figure 15: Estimated biomass time series at the end of quarter two (Age 0 removed to be consistent with recruitment at the end of the second quarter of the year). Note that under the assumption that all individuals in $B1+$ class are mature, this biomass is equivalent to SSB

8. Catch advice for July 2023 to June 2024

8.1. One over two harvest control rule

The ratio between the last year biomass estimate and the mean of the two previous years is:

$$\frac{B_y}{\overline{B_{y-1} + B_{y-2}}} = \frac{4402}{(3332 + 2593)/2} = 1.486$$

for B representing the estimated abundance by the model as shown in Figure 15. According to the report of WKLIFEVX (ICES,2021), if this ratio is above 1.8, the advice would be equal to the latest advice mutiplied by 1.8, if not, the latest advice would be multiplied by this ratio. In case the estimated abundance is below a biomass trigger, which in this case is B_{lim}, it is also multiplied by a biomass safe guard as follows:

$$C_{y+1} = \hat{C}_y * \min\left(1.8, \frac{B_y}{(B_{y-1} + B_{y-2})/2}\right)$$

where $\hat{C}_y$ is the value of adviced catches in the previous year. Then the adviced catches (in tonnes) for the next year (July 2023 to June 2024) would be:

$$C_{y+1} = 1694 * 1.49 = 2517.3.$$

This procedure modification has been implemented since this year and it is not specified in the Stock annex.

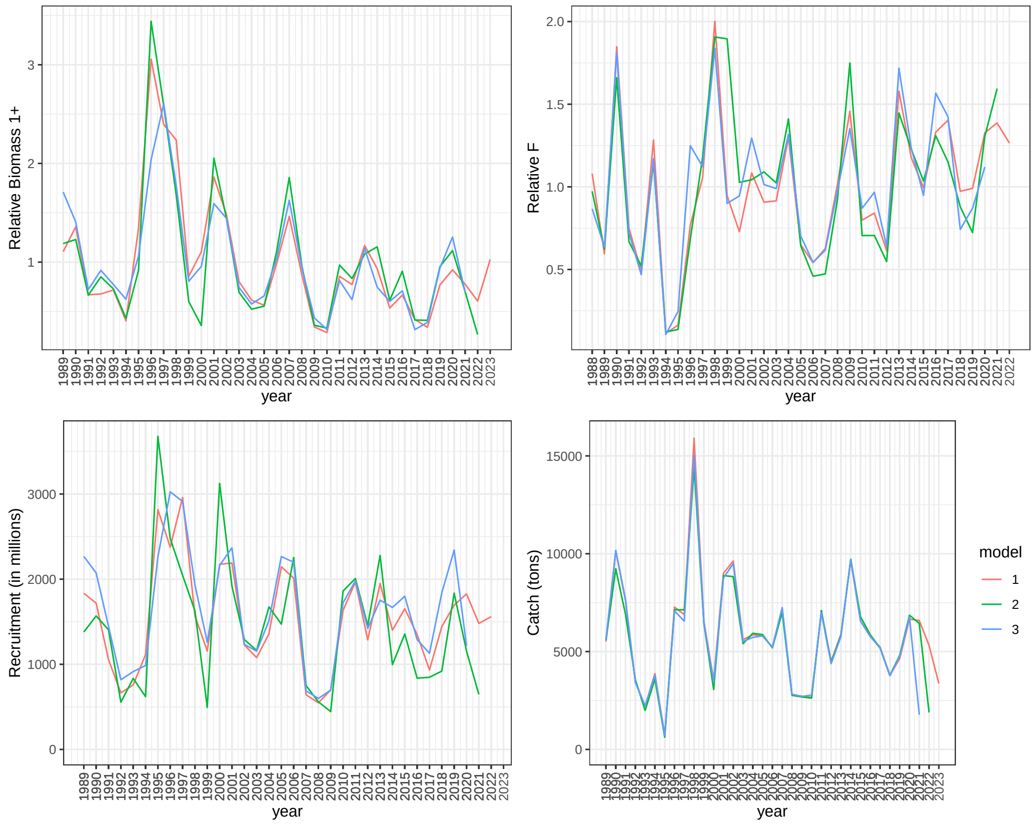

Figure 16: Comparison of estimates from different model implementations. Blue line corresponds to the current year estimated time series (the one estimated by the model described in this document), the green line, to the estimated in 2022 and the blue line, to the estimated in 2021. Measures were summarized at the end of June each year, assuming that a year starts in July and ends in June of the next year.

8.2. Constant harvest rate rule

According to this rule, adviced catches (in tonnes) for the next year (July 2023 to June 2024) would be the product of the last year biomass estimate and a constant harvest rate. In this case a rate of 0.5 was considered like the most suitable rate for this stock, as follows:

$$C_{y+1} = B_y * 0.5 = 4402 * 0.5 = 2201.$$

This procedure modification has been implemented since this year and it is not specified in the Stock annex.

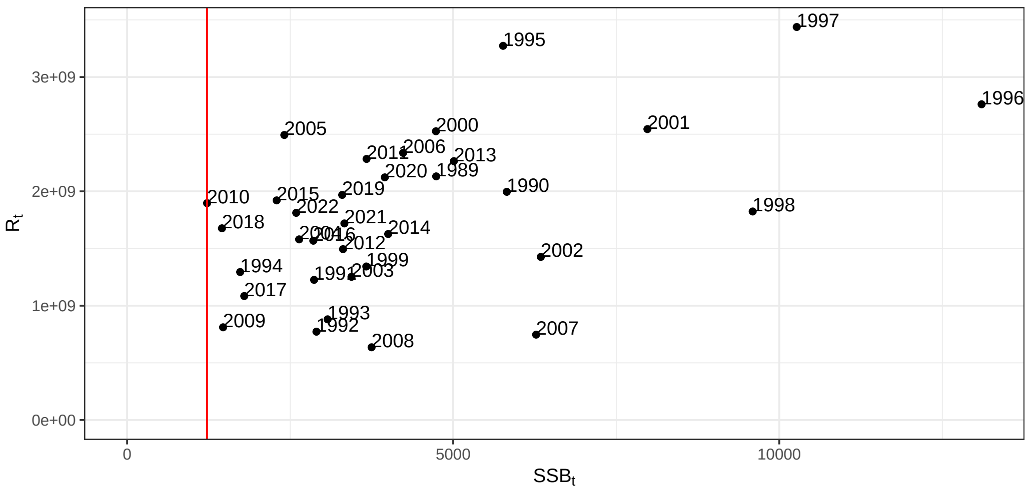

Figure 17: Estimated Stock Spawning biomass (SSB_t) vs. Recruitment (R_t), SSB_t corresponds to the Stock Spawning Biomass at the end of quarter 2 of year t, while R_t corresponds to the sum of the recruitment at the beginning of quarters 3,4 and 1 of years t and $t+1$, respectively.

9. Acknowledgements

We thank Jamie Lentin from Shuttlethread for the automatization of data input, Bjarki Elvarsson for having an open repository with very useful Gadget data processing routines and his valuable help, and to the members of WGHANSA group for their guidance and support.

We gratefully thank CESGA (Galician Supercomputing Center) for computational time at the FTII Supercomputer and technical assistance.

10. References

Begley, J., 2004. Gadget User Guide. URL: `http://www.hafro.is/gadget/files/userguide.pdf`.

Bellido, J.M., Pierce, G.J., Romero, J.L., Millan, M., 2000. Use of frequency analysis methods to estimate growth of anchovy (*Engraulis encrasicolus L.* 1758) in the gulf of cádiz (SW spain). Fisheries Research 48, 107–115.

Elvarsson, B., Taylor, L., Trenkel, V., Kupca, V., Stefansson, G., 2014. A bootstrap method for estimating bias and variance in statistical fisheries modelling frameworks using highly disparate datasets. African Journal of Marine Science 36, 99–110. URL: `http://www.tandfonline.com/doi/abs/10.2989/1814232X.2014.897253`, doi:`10.2989/1814232X.2014.897253`.

Gislason, H., Daan, N., Rice, J.C., Pope, J.G., 2010. Size, growth, temperature and the natural mortality of marine fish. Fish and Fisheries 11, 149–158.

Lentin, J., 2014. mfdb: MareFrame DB Querying Library. R package version 3.2-0.

Stefansson, G., 2003. Issues in Multispecies Models. Natural Resource Modeling 16, 415–437. URL: http://onlinelibrary.wiley.com/doi/10.1111/j.1939-7445.2003.tb00121.x/abstract, doi:10.1111/j.1939-7445.2003.tb00121.x.

Stefánsson, G., 1998. Comparing different information sources in a multispecies context. Fishery stock assessment models. Alaska Sea Grant College Program. AK-SG-98-01 , 741–758URL: http://mdgs.un.org/unsd/envaccounting/ceea/archive/Fish/Iceland.PDF.

Taylor, L., Begley, J., Kupca, V., Stefansson, G., 2007. A simple implementation of the statistical modelling framework Gadget for cod in Icelandic waters. African Journal of Marine Science 29, 223–245. URL: http://www.tandfonline.com/doi/abs/10.2989/AJMS.2007.29.2.7.190, doi:10.2989/AJMS.2007.29.2.7.190.