**Terms of Reference for an ad hoc contract in support of the EWG on stocks assessments and to improve the assessment of the stock of hake in GSAs 1, 5, 6 and 7.**

1. Background

This ad hoc contract is meant to complete the stock assessment runs for hake in GSA 1-5-6-7 that we identified in STECF EWG 25-06 and further discussed during PLEN 25- 02.

2.Terms of Reference to STECF

ToRs:

1) Create the FLStock and FLIndex for the assessment, using the data held by the JRC, submitted through the EU Mediterranean and Black Sea data call.

2) Convert length-based input data for stock assessment (commercial catches and MEDITS abundance indices) into age-based data using the stochastic slicing method commonly applied by the EWG. The slicing should be performed quarterly, applying the combined growth model as defined in EWG2506.

a. Output: FLStock and FLIndex objects ready for use in the statistical catch-at-age model implemented by a4a.

3) Improve the a4a model proposed in EWG2506 (run 3.2 with fmodel tensor to increase flexibility in capturing changes in the selection pattern) and provide associated diagnostics. In particular, the impact of different levels of smoothness in the fmodel should be evaluated by running a grid of options with varying k values. Assess whether any other submodels beyond the fmodel require modifications.

a. Output: updated submodels, fitted object, diagnostics and grid results.

4) Run the model selected based on the outcomes of item 2 using the age-based data produced in item 1.

a. Output: fitted model and diagnostics.

**ToR 1.- Create the FLStock and FLIndex for the assessment, using the data held by the JRC, submitted through the EU Mediterranean and Black Sea data call.**

The FLStock and FLIndex objects were produced with the most up-to-date data from the official data call available in the STECF EWG 25 09 workspace, and using the code stored in the STECF EWG 25 06 workspace (combination of the code from the scenarios “2025\_BaseCase\_Update\_priority\_2” and “2025\_MEDBALAR\_priority 3”). To fulfil this ToR, the deterministic slicing was implemented with the growth parameters agreed during the STECF EWG 25 06 (Linf=110; K=0.178, T0=-0.0005).

The code developed to merge the BALAR and MEDITS data was not included in the STECF EWG 25 06 workspace. For this reason, the code necessary to merge and clean up the BALAR and MEDITS data had to be developed again as part of this ad-hoc contract. The results obtained were slightly different to the values obtained during the STECF EWG 25 06. For this reason, it was decided to use the csv files with the ta, tb and tc data for BALAR and MEDITS prepared during the EWG 25 06. The reasons for this difference will need to be analyzed during the upcoming EWG 25 09.

The resulting FLIndex object, the FLStock object by country and GSA area, as well as the FLStock object with all countries and GSAs combined, are stored in the ToR1 folder.

**ToR 2.- Convert length-based input data for stock assessment (commercial catches and MEDITS abundance indices) into age-based data using the stochastic slicing method commonly applied by the EWG. The slicing should be performed quarterly, applying the combined growth model as defined in EWG2506.**

The expected outputs for this ToR 2 are the FLStock and FLIndex objects ready for use in the statistical catch-at-age model implemented by a4a.

This ToR required three separate and consecutive steps:

1. Producing length distributions of landings, discards and total catch by quarter.
2. Implement the stochastic slicing by quarter using the sex combined growth model.
3. Move individuals at age 0 in quarters 1 and 2 to the age 1.

The code developed during the EWG 25 06 to conduct the above mentioned three steps was not developed in a format that could be integrated with the code available to handle the data and perform the stock assessment for hake GSA\_1567, and only a small part of the code produced during the EWG 25 06 could be recycled. For this reason, the necessary code to produce the length distributions by quarter, and implementing the stochastic slicing had to be developed as part of the work of this ad-hoc contract. The resulting modelling scenario was called “2025\_MEDBALAR\_priority\_4\_stoc\_slic\_quarter”. The folder containing this scenario, and the resulting FLStock objects are included in the ToR 2 folder.

The catch at age in numbers obtained when applying the stochastic slicing by quarter was compared to the catch at age produced by the deterministic slicing. Important differences were found. To rule out that the problem might be related with the estimation of length distribution by quarter, the stochastic slicing was applied to length distributions estimated by year, and compared to the estimated total annual catch at age with deterministic slicing also implemented to length distributions by year. Differences between both approaches was still very important. Within the ToR 2 folder, the folders containing the comparisons are included. Here, as an example, the ages where differences are more important are presented (see figures 1 and 2).

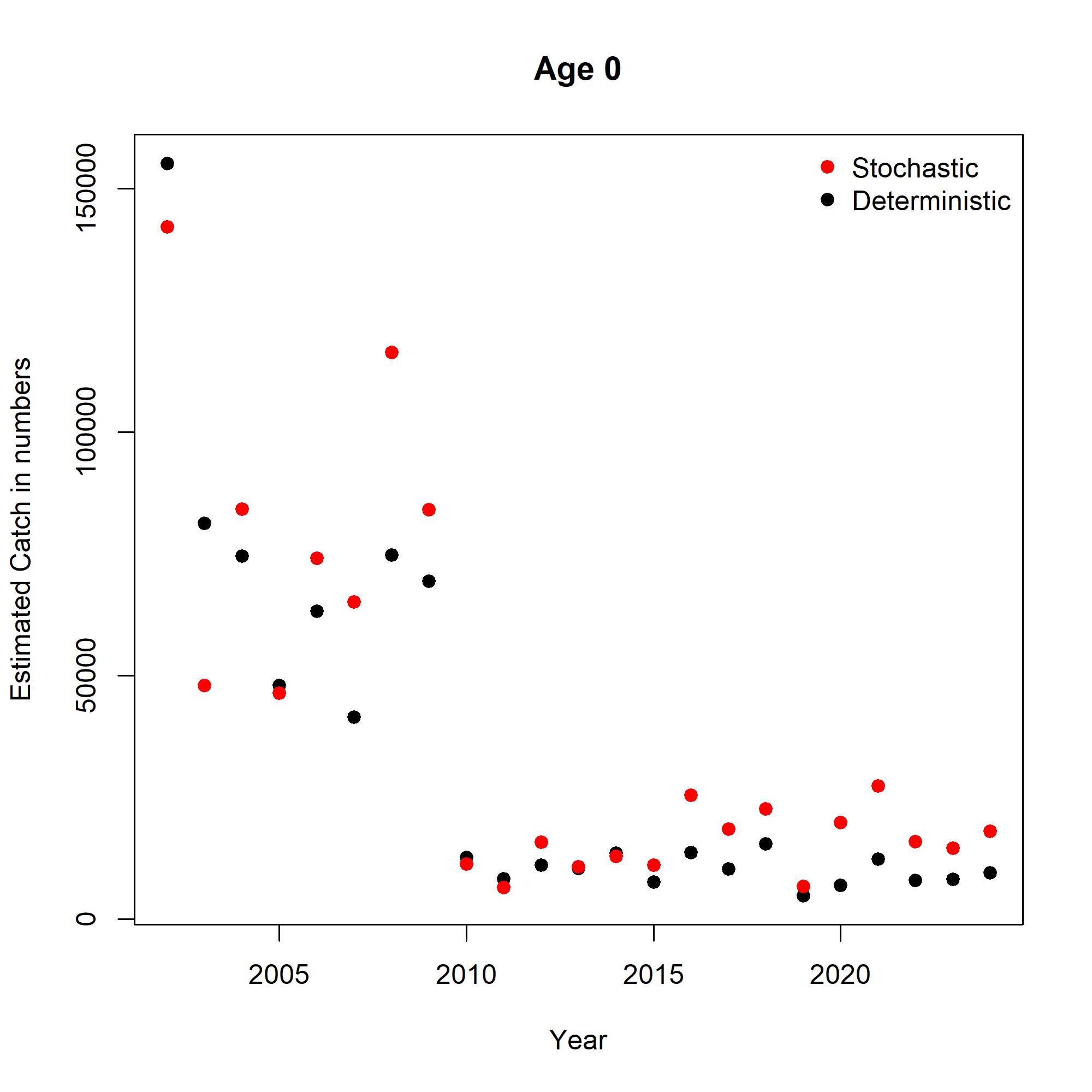


Figure 1.- Estimated total annual catch at age 0 in numbers by seasonal stochastic slicing by season and deterministic slicing by year.

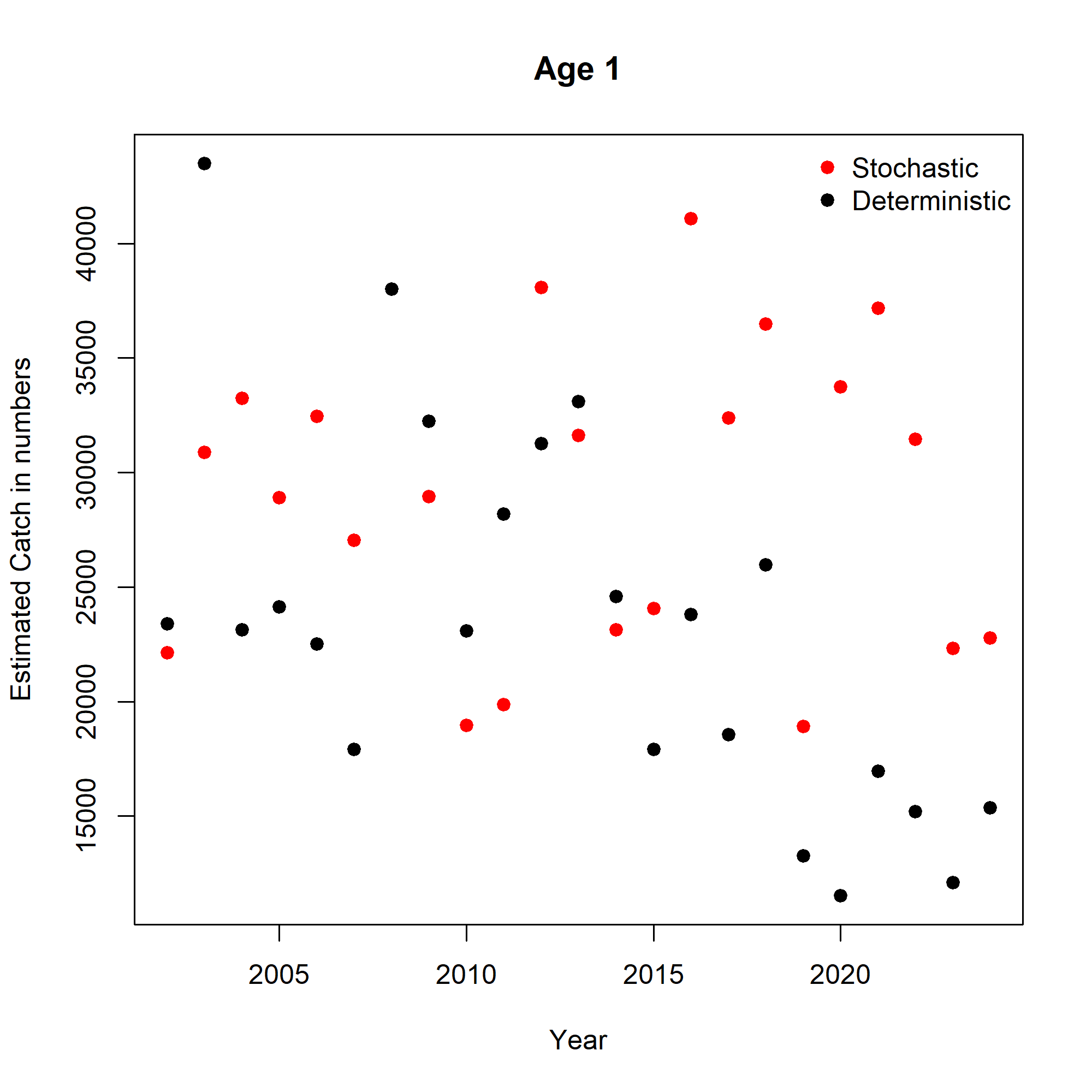
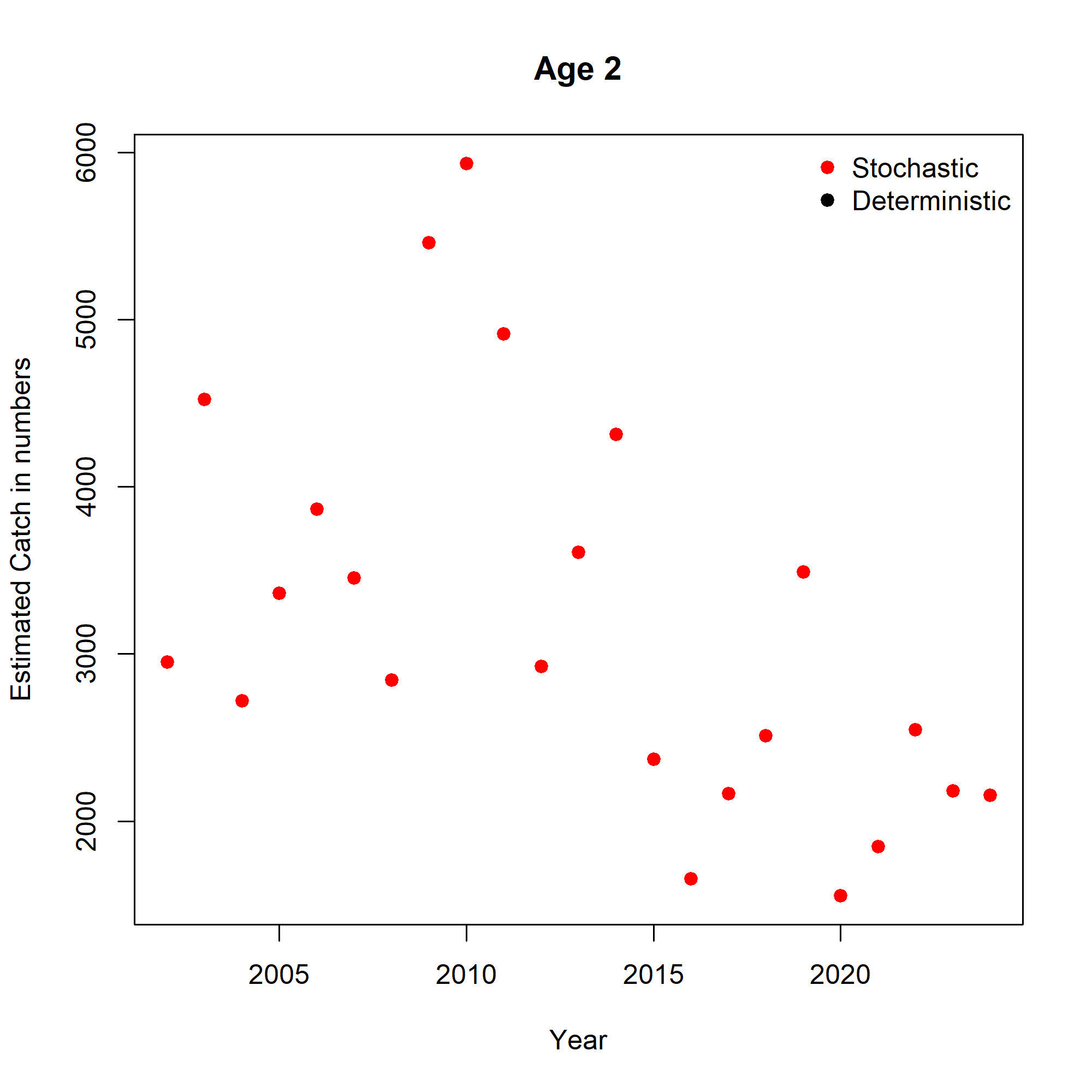
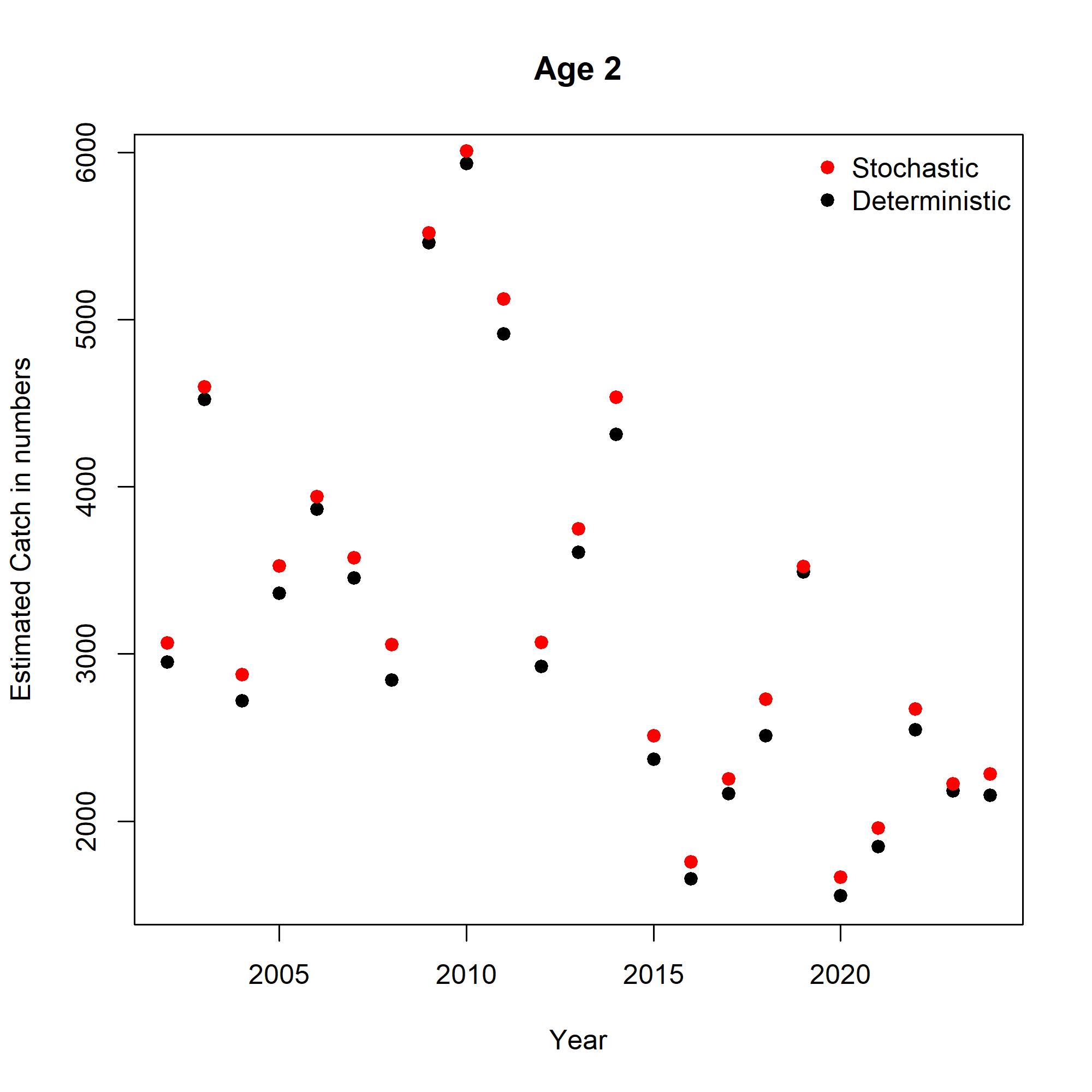


Figure 2.- Estimated total annual catch at age 1 in numbers by seasonal stochastic slicing by season and deterministic slicing by year.

It was finally found out that part of the problem was related with the number of iterations indicated in the mvrnorm() function used to produce the distribution of the growth parameters Linf, K and T0. If 250 iterations were set in the function mvrnorm, the mean value obtained for Linf, K and T0 was significantly different than the values from Melon-Duval selected during the EWG 25 06 and used in the deterministic slicing (Linf=110, K=0.178, T0=-0.0005). When the number of iterations was set at a very high number (100000 iterations or higher), the random distributions produced had mean values that were very closed to the deterministic growth parameters. When the stochastic slicing was run with 100000 iterations, the differences in the estimated catch at age obtained from stochastic and deterministic approaches were small when the stochastic slicing was applied to length distributions by year (see figure 3), although important differences at age 3 were still found. In case of stochastic slicing by quarter, the same approach was repeated (100000 iterations) to the length distributions estimated by quarter. In this case, although differences in annual estimates of catch at age were smaller, they were still important (see the figures in the folder “5.- determ\_versus\_stoch\_slicing\_my\_quarter” within the folder “ToR\_2/Comparisons”.

*Figure 3.- Difference between annual total catch at age 2 estimated by stochastic or deterministic slicing, in both cases applied to length distributions by year. In the left panel, the stochastic slicing is run with 100000 iterations to produce the distributions of the Linf, K and T0 parameters, while in the right panel 1000 iterations were run.*

In addition to the effect of the number of iterations when producing the growth parameters distributions, differences were found in the annual total catch in numbers by length estimated by quarter or by year in France-GSA07 in regular length distribution, and in Spain-GSA07 filled-gaps length distribution. The problem with the France-GSA07 data is due to the discards data. In the discards data there are 2283 entries for which the season column has a -1, which end up being removed. Regarding the Spain-GSA07 filled gaps data, this seems to be necessary to reconstruct the LFDs in 2020, since no biological sampling was available that year for any metier/gear. It can be expected that when filling up the gaps using the information from other metiers and areas different estimates of LFs by quarter are produced compared to year.

Hence, in summary the issues are:

1.- There are differences when estimating length distributions by quarter or by year for FRANCE-GSA07 and filled-gaps data for SPAIN-GSA07

2.- The stochastic slicing requires of very high number of iterations, and still, some differences should be expected.

At this stage, and for the work that has to be done as part of ToR 3, as a compromise solution it was decided to continue by using the deterministic slicing by year. In addition, it was not possible carrying out the correction of transferring individuals of age 0 in quarters 1 and 2 to age 1, given that the length distributions are estimated by year, and not by quarter.

ToR 3.- Improve the a4a model proposed in EWG2506

The most advanced scenario run during the EWG 25 06 was the scenario 3.4, for which a tensor spline was introduced in the fishing mortality model (fmodel) to increase flexibility in capturing changes in the selection pattern. In this ToR it is expected that the impact of different levels of smoothness in the fmodel are evaluated by running a grid of options with varying k values. The expected output are the updated submodels, fitted object, diagnostics and grid results.

The fmodel implemented in the stock assessment working group in 2024 (STECF EWG 24 10) was:

fmodel <- ~s(age, k=4) + s(year, k=8) + s(year, k=8, by=as.numeric(age==0)) + s(year, k=8, by=as.numeric(age==4))

Which includes a spline with 4 knots to model changes of F by age, a spline with 8 knots to model changes of F over the years, a specific spline for F at age 0 with 8 knots, and another spline specific for age 4 with another 8 knots. The intention when including a spline for age 0 was accounting for the change in selectivity derived from the legislation that force to use a square mesh of 40 mm in the cod end. The spline at age 4 was intended to account for a sharp decline in the survey index for ages 4-5.

During the STECF EWG 25 06, the time series was extended 5 more years, starting in 2002. The lengthening of the time series allowed some extra contrast and some more years of data previous to some important changes in the distribution of catches by fleet, with a decline of longliners contribution to total catch since 2010, and an increase of catch by gillnetters. The fmodel selected during the EWG 25 06 was:

fmodel <- ~s(age,k = 4) +s(year, k = 11) + te(age,year, k=c(3,6))

Which includes an spline by age (4 knots) and by year (11 knots), but also a tensor spline for age and year with 3 and 6 knots respectively. A tensor spline is a mathematical product that help model the individual smoothness of each variable (in this case age and year) while also capturing their joint interaction. The independent variables in the case of tensors are modeled with different numbers of basic functions allowing different amount of smoothness in each dimension. The smoothness, as it is the case in the splines, is defined by the number of knots (parameter k).

In this ToR the first goal is to continue exploring different combinations of smoothness (number of knots) for each of these three components in the model. Once one fmodel (or a selection of fmodels) has been selected, the stock-recruitment model (srmodel) and the survey catchability model (qmodel) will also be explored.

3.1.- Range of smoothness tested in the fmodel

A wide range of combinations of smoothness in the age spline (from 2 to 4 knots) , year spline (from 6 to 11 knots) and age-year tensor spline (combinations of 2-4 knots for age and 6-11 knots for year) were tested. In total, 220 models were run exploring different combinations of smoothness in the fmodel. The folder with the code and the results of those explorations is included in the ToR\_3 folder, with the name of “7.- 2025\_MEDBALAR\_priority 4 detslic\_fmodel\_explor”. The figures and tables with the diagnostics and model estimates for each of the 220 models are included within this folder, in the subfolder named as “output\_model\_explor”. The diagnostic explored for each model were:

* Observed versus estimated catch at age in numbers, and total catch
* Standardized residuals of catch at age in numbers, and total catch
* Observed versus estimated survey index at age in numbers
* Standardized residuals of survey index at age in numbers
* Degree of normal distribution of residuals (qqplots)
* Estimated fishing mortality (F) at age over years
* Estimated fishing selectivity at age over years
* Estimated survey catchability (Q) at age (constant over years)
* Retrospective pattern over the last 5 years (including Mohn´s Rho, maximum and minimum retro)
* Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC)

The table “select\_criteria\_all\_models.csv” contains the AIC, BIC, Mohn´s rho, maximum and minimum value in the retrospective pattern for the estimated Spawning Stock Biomass (SSB), Fishing mortality (Fbar) and Recruitment for each of the 220 models run. Moreover, this table also contains the equations used in each model run for the fmodel, qmodel and srmodel submodels.

The first selection criteria implemented was the Mohn´s rho. Those models with a Mohn´s rho higher than 0.10 or lower than -0.10 were discarded. This restriction resulted in a reduction of the number of models from 220 to 47 models. The next selection criteria was discarding those models where the maximum or minimum value in the retro pattern was above or below the 50% (that is 0.50 or -0.5 ratio between the value of Fbar, SSB or recruitment in the last year in the retro compared to that year in the complete assessment). This restriction resulted in a shorter list of 6 models. If we were less restrictive with the limit to the Mohn´s rho value, accepting those models with a Mohn´s rho smaller than 0.15, the number of models selected were 72. However, when applying the limit of 0.5 (or 50%) in the retro in any year, the list of models selected was exactly the same as before. The list of models with their respective fmodel, qmodel and srmodel are presented in table 1.

*Table 1.- Models selected when the limit of Mohn´s rho <0.10 and maximum retro value of 0.5 (or 50%) in any year is applied.*

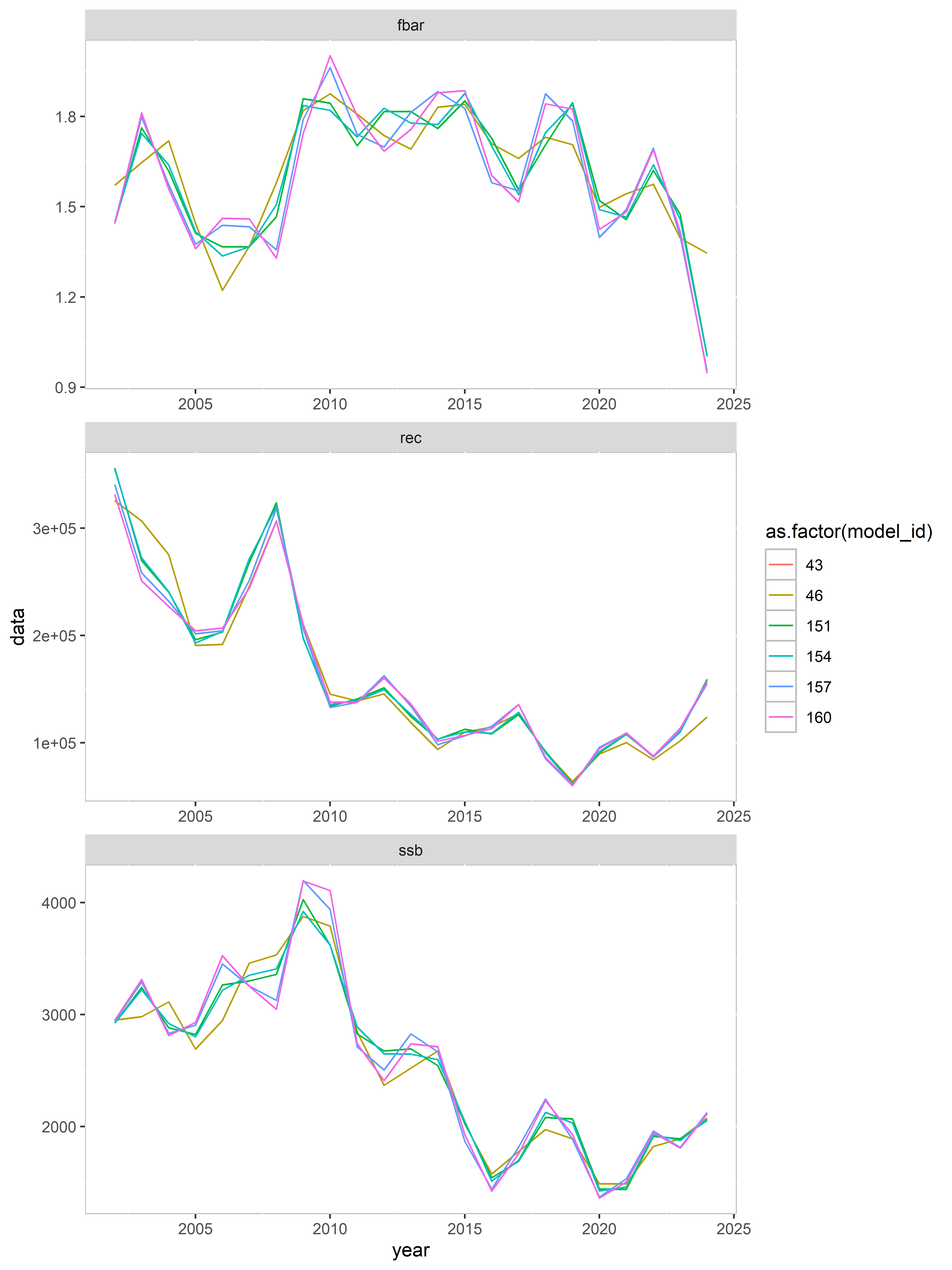
|  |  |  |  |
| --- | --- | --- | --- |
| model\_id | fmod\_eq | qmod\_eq | srmod\_eq |
| 157 | ~s(age, k = 4) + s(year, k = 8) + te(age, year, k = c(3, 10)) | ~I(1/(1 + exp(-age))) | ~factor(year) |
| 160 | ~s(age, k = 4) + s(year, k = 9) + te(age, year, k = c(3, 10)) | ~I(1/(1 + exp(-age))) | ~factor(year) |
| 151 | ~s(age, k = 4) + s(year, k = 6) + te(age, year, k = c(3, 10)) | ~I(1/(1 + exp(-age))) | ~factor(year) |
| 154 | ~s(age, k = 4) + s(year, k = 7) + te(age, year, k = c(3, 10)) | ~I(1/(1 + exp(-age))) | ~factor(year) |
| 46 | ~s(age, k = 4) + s(year, k = 7) + te(age, year, k = c(3, 7)) | ~I(1/(1 + exp(-age))) | ~factor(year) |
| 43 | ~s(age, k = 4) + s(year, k = 6) + te(age, year, k = c(3, 7)) | ~I(1/(1 + exp(-age))) | ~factor(year) |

The model estimates (figure 4) showed two clear groups, models 43 and 46 in one side, and models 151, 154, 157 and 160 on the other side. Hence, the difference in the perspective of the F is due to the different number of knots affecting the year in the tensor spline. Despite the differences, except for the last year in the assessment, all models showed very similar trends and absolute values.

Out of the 6 selected models, the one with the lowest AIC was the model number 157, whose fmodel was set as:

fmodel = ~s(age, k = 4) + s(year, k = 8) + te(age, year, k = c(3, 10))

Given that all models produced similar results, the model number 157 was selected as the candidate model given its lower AIC score (see table 2). Next, the main diagnostics and model estimates are presented within the ToR 4 section.



*Figure 4.- SSB, recruitment and Fbar for the 6 selected models (see table 1).*

*Table 2.- AIC, BIC and Mohn´s rho for Fbar, SSB and Rec for the six models selected.*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| model\_id | aic | bic | mohn\_rho\_Fbar | mohn\_rho\_SSB | mohn\_rho\_Rec |
| 157 | 294.3 | 540.2 | -0.071 | 0.077 | -0.087 |
| 160 | 294.4 | 543.8 | -0.085 | 0.07 | -0.077 |
| 151 | 303.2 | 542.1 | -0.091 | 0.059 | -0.078 |
| 154 | 304.4 | 546.8 | -0.087 | 0.058 | -0.075 |
| 46 | 311.5 | 522.3 | -0.049 | 0.075 | -0.01 |
| 43 | 312.5 | 519.8 | -0.061 | 0.073 | -0.004 |

ToR 4.- Run the model selected

The expected outputs for this ToR 4 are the fitted model and diagnostics. The folder with the model, the necessary code to reproduce the entire process from the input data to the model output are included within the folder “8.- 2025\_MEDBALAR\_priority 4 detslic\_selmod\_assess”

Input data:

The input data shows a decrease in total catch in tons since 2010 (figure 5), which was the results of the decline in commercial catches at all ages (figure 6). However, the catch at age in numbers seems to indicate that the decline in catch at age has occurred later in time the younger the age (figure 6). The survey index (figure 7) also indicates a decline in abundance index at age over the historic period. However, the reduction is not as pronounced as in the commercial catch at age.

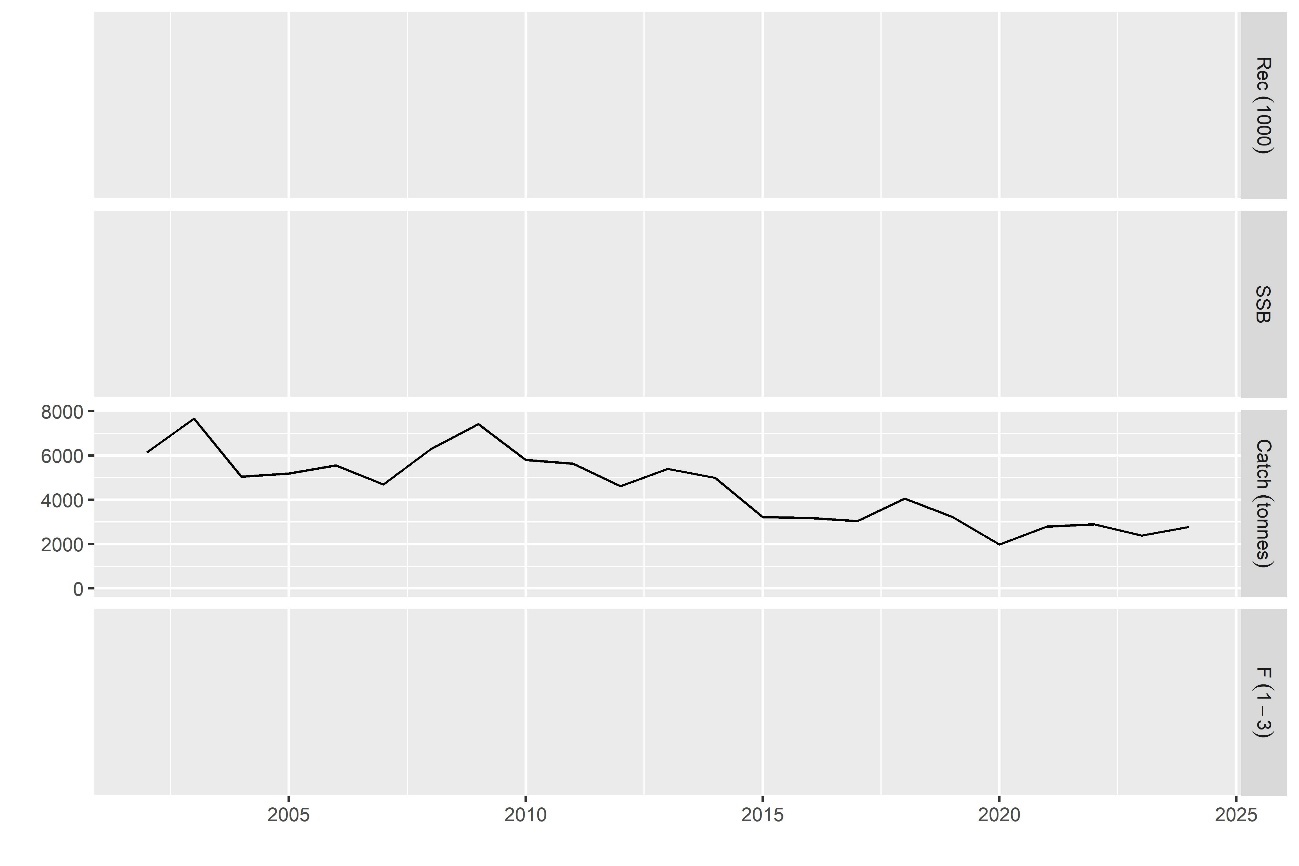


Figure 5.- Total annual catch in tons over the historic period (2002-2024)

The average weight at age shows increasing trends at ages 0 and 1, while a decreasing trend in observed at age 2 (figure 8).

The cohort consistency plots indicate a relatively high degree of linear correlation (pearson coefficient equal or higher than 0.5 at all ages) between consecutive ages in the commercial catches (figure 9). In the survey index, correlation was high between ages 1 and 2, and 2 and 3, but low between ages 0 and 1, and 3 and 4plus (figure 10).

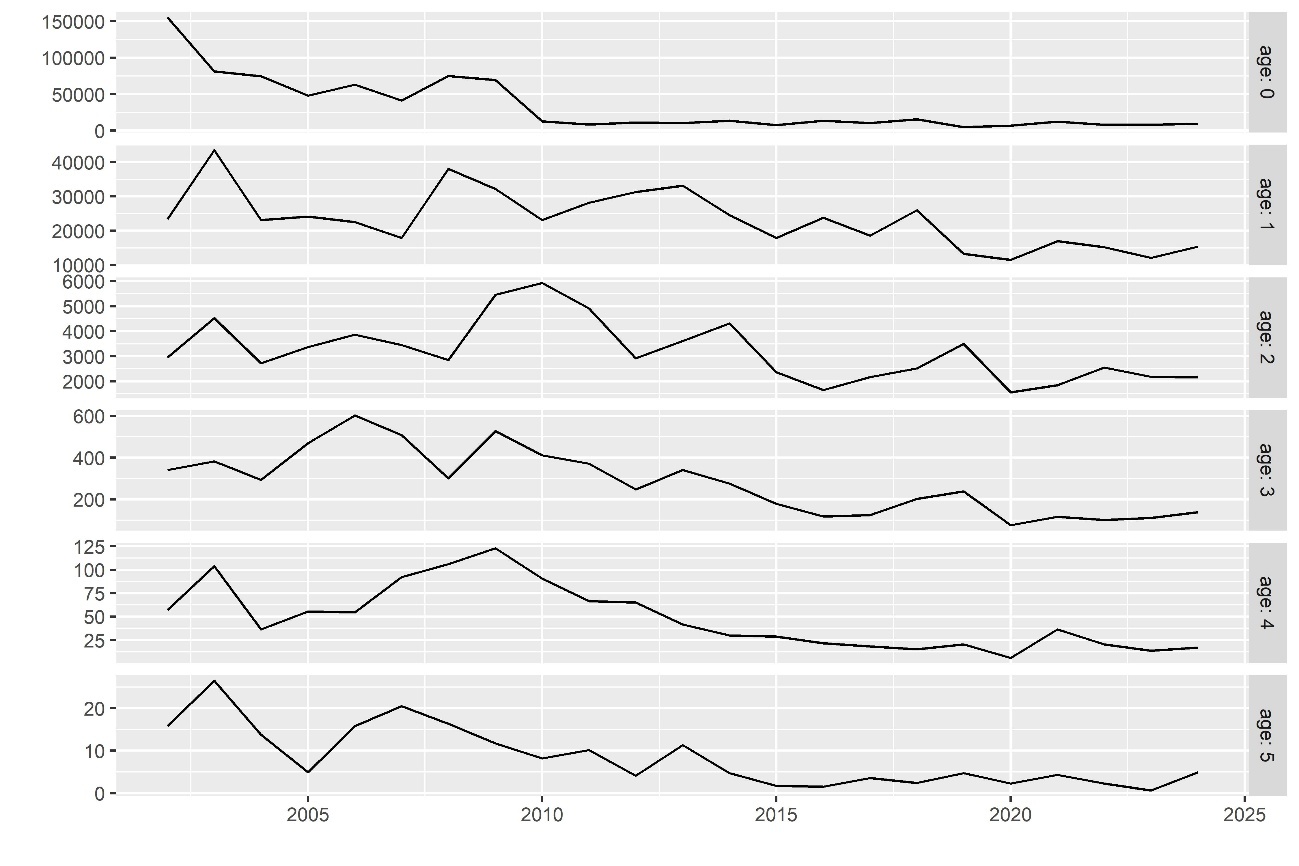


Figure 6.- Total catch in numbers by age over the historic period (2002-2024)

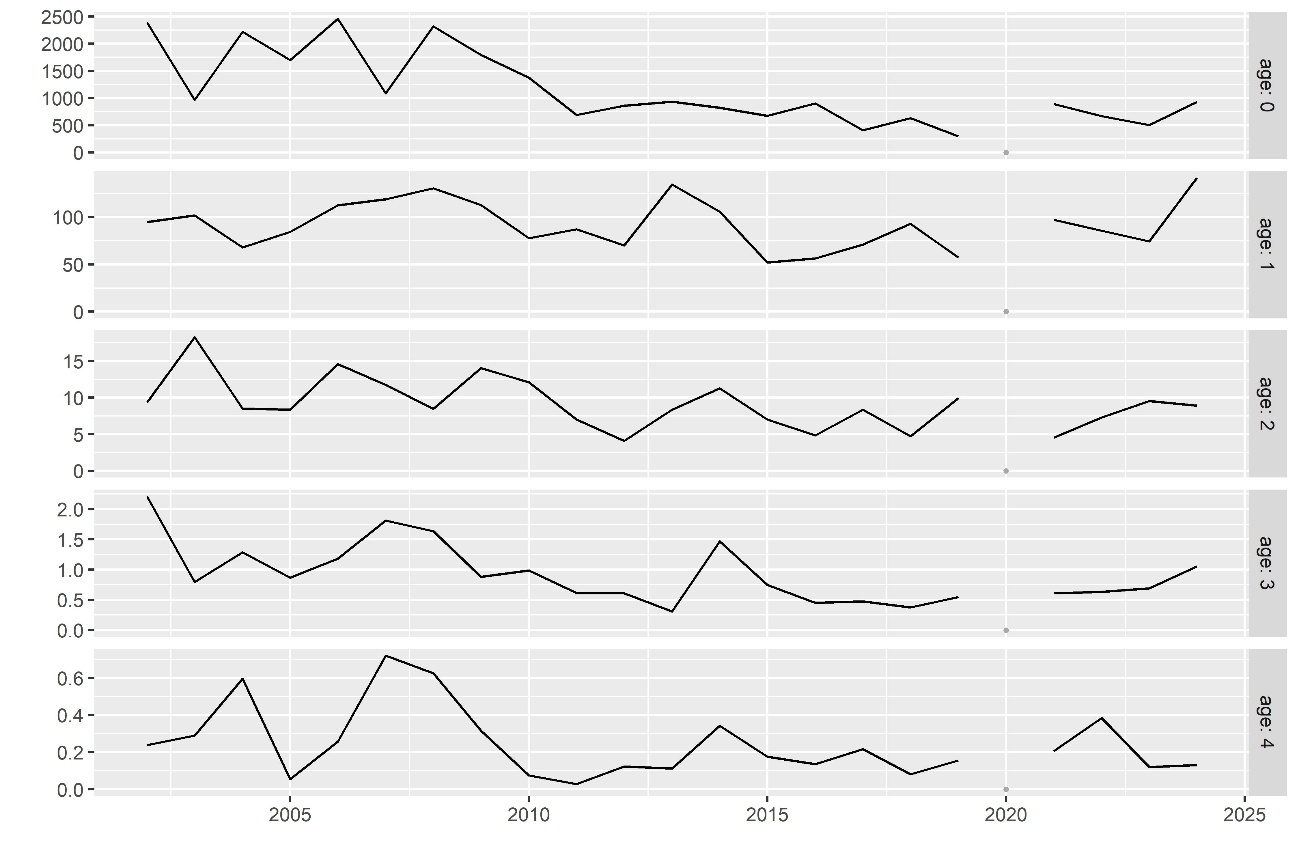


Figure 7.- Abundance survey index by age over the historic period (2002-2024). In 2020 there is no survey index due to COVID pandemic.

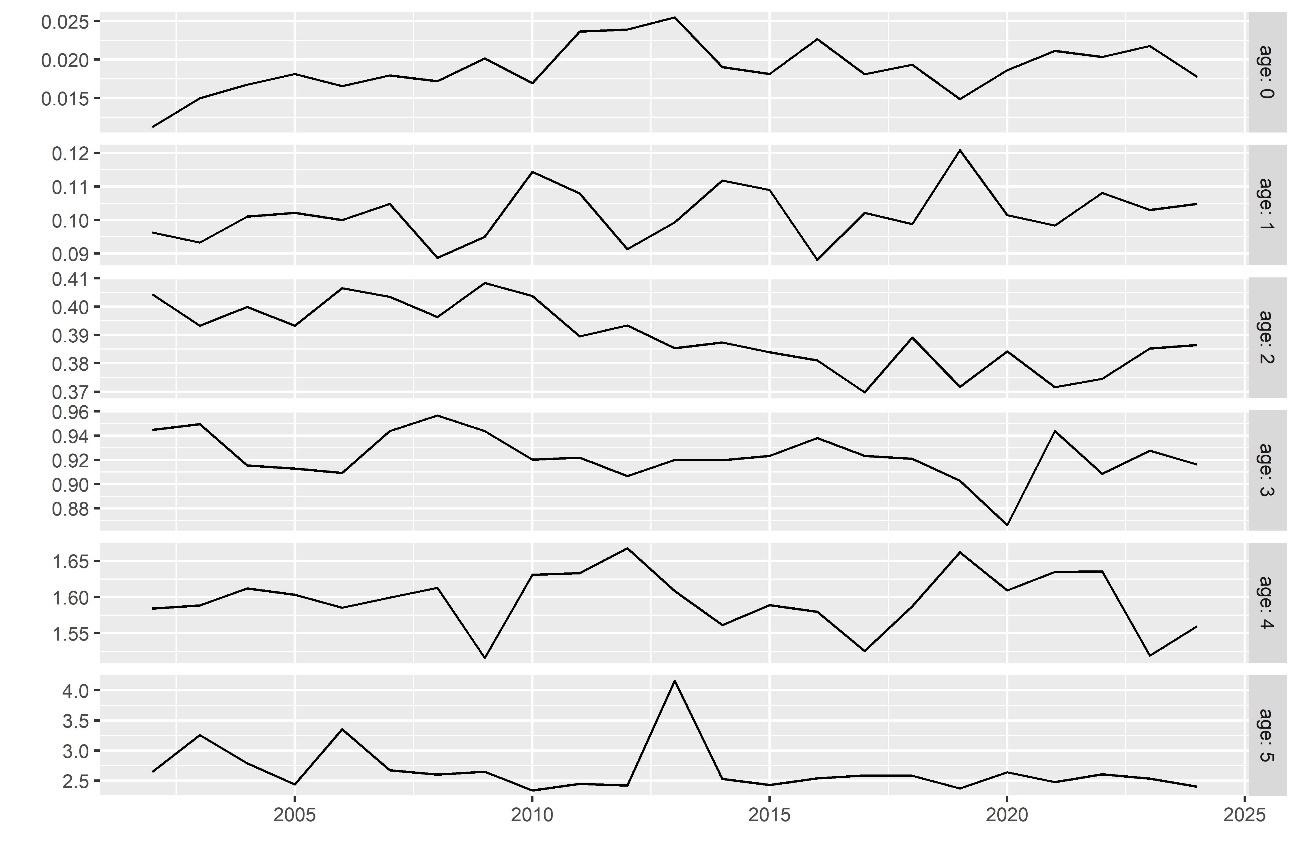


Figure 8.- Mean weight at age in kg over the historic period (2002-2024).



Figure 9.- Cohort consistency plots for commercial catch in numbers between consecutive ages.



Figure 10.- Cohort consistency plots for abundance survey index between consecutive ages.

Diagnostics:

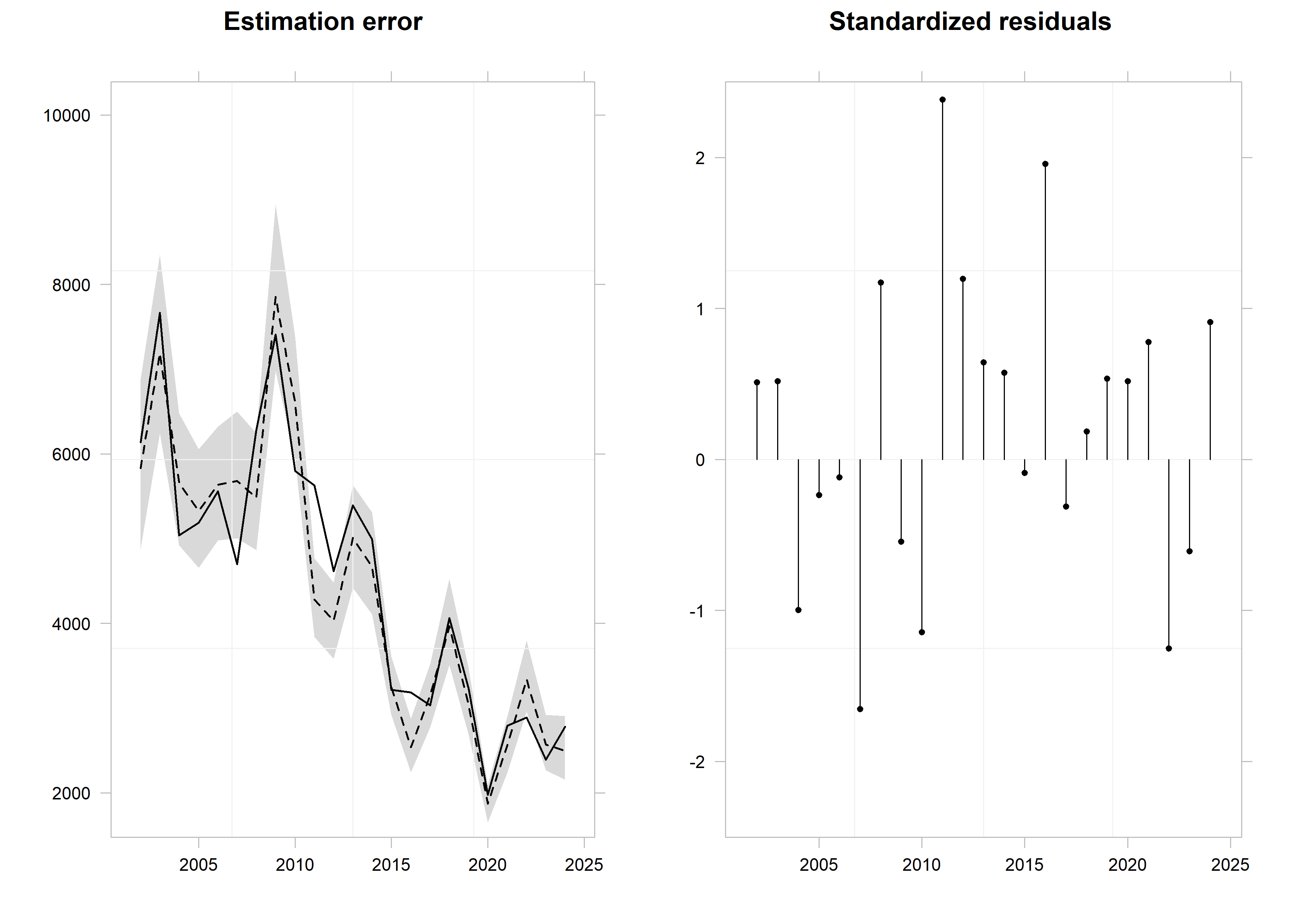
In table 3, the main parameters concerning the model fit are presented. There were no problems during the model fit, and the model converged. As explained in ToR 3 section, the selected model (model id number 157), obtained the best diagnostics in terms of Mohn´s rho and AIC.

Regarding the model estimates compared to the observed data, the model showed a good fit to the observed data (figures 11 to 15). The total commercial catches were tracked with high degree of accuracy by the model estimates, without a clear trend in the standardized residuals (figure 11). The log residuals of commercial catches and abundance survey index at age (figures 12 to 15), showed that commercial catch at age was tracked closely over time, with not a clear trend in the residuals. However, the survey index at ages 0, 1 and 2, shows some trend over time.

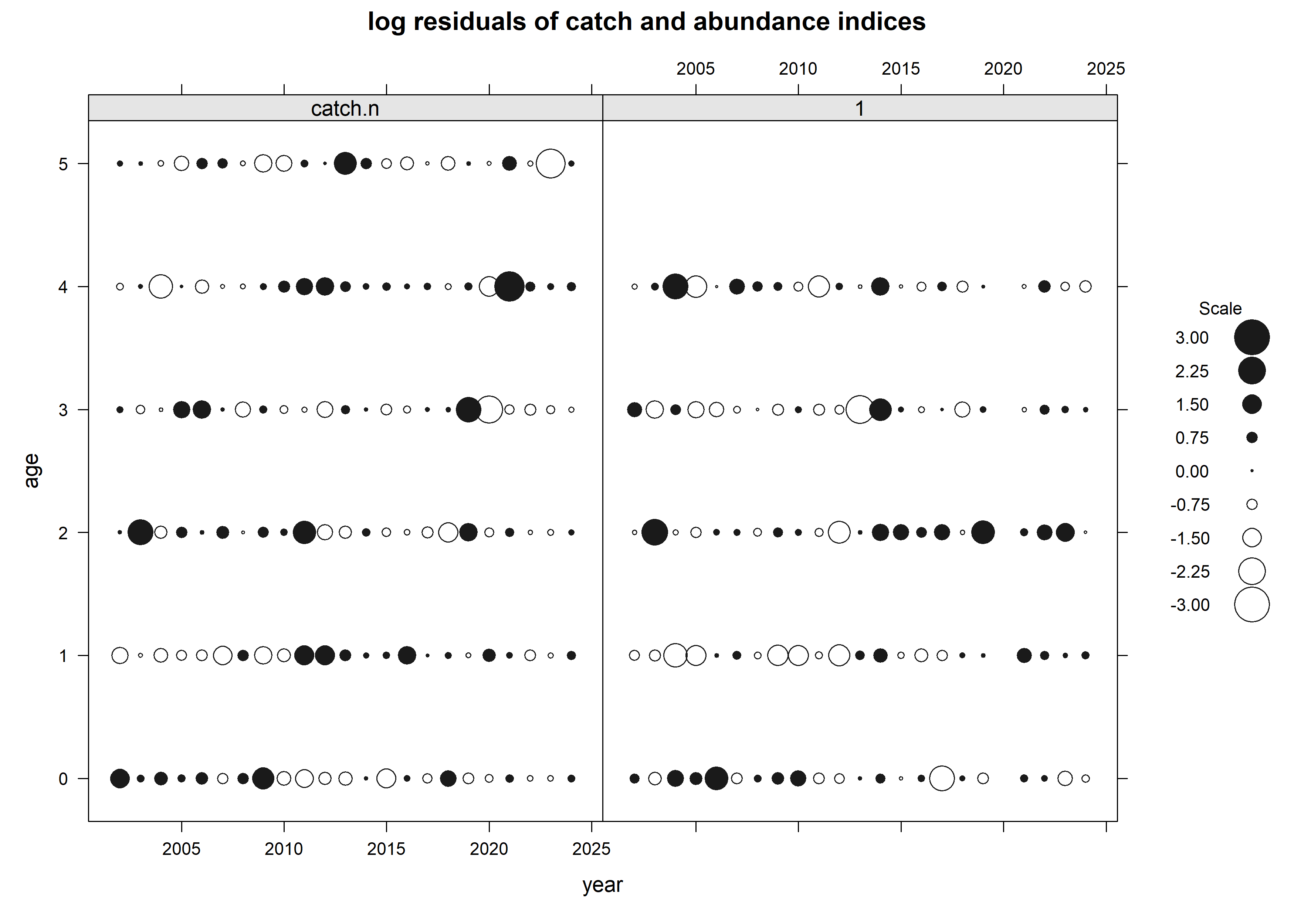
The retrospective pattern indicates a tendency to underestimate the Fbar, since 2020, with a positive correction when a new year of data is available (Figure 16). Despite of this observation, the retrospective pattern produced low Mohn´s rho values.

Table 3.- Summary of the selected model fit.

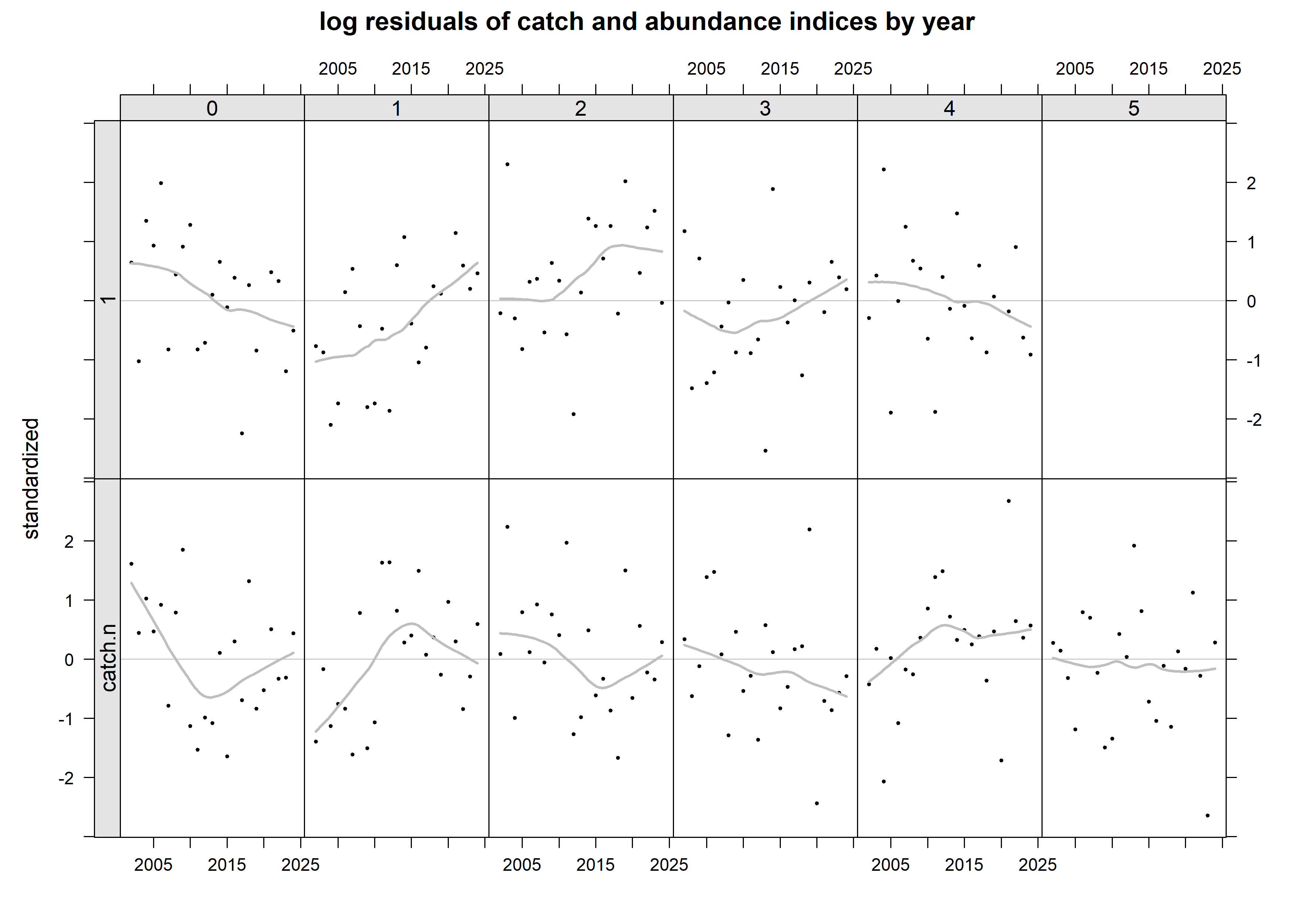
|  |  |
| --- | --- |
| **Parameter** | **Value** |
| nopar | 70 |
| nlogl | 77.1479531 |
| maxgrad | 3.03E-06 |
| nobs | 248 |
| gcv | 0.25391929 |
| convergence | 0 |
| accrate | NA |
| nlogl\_comp1 | 25.4619 |
| nlogl\_comp2 | 51.686 |



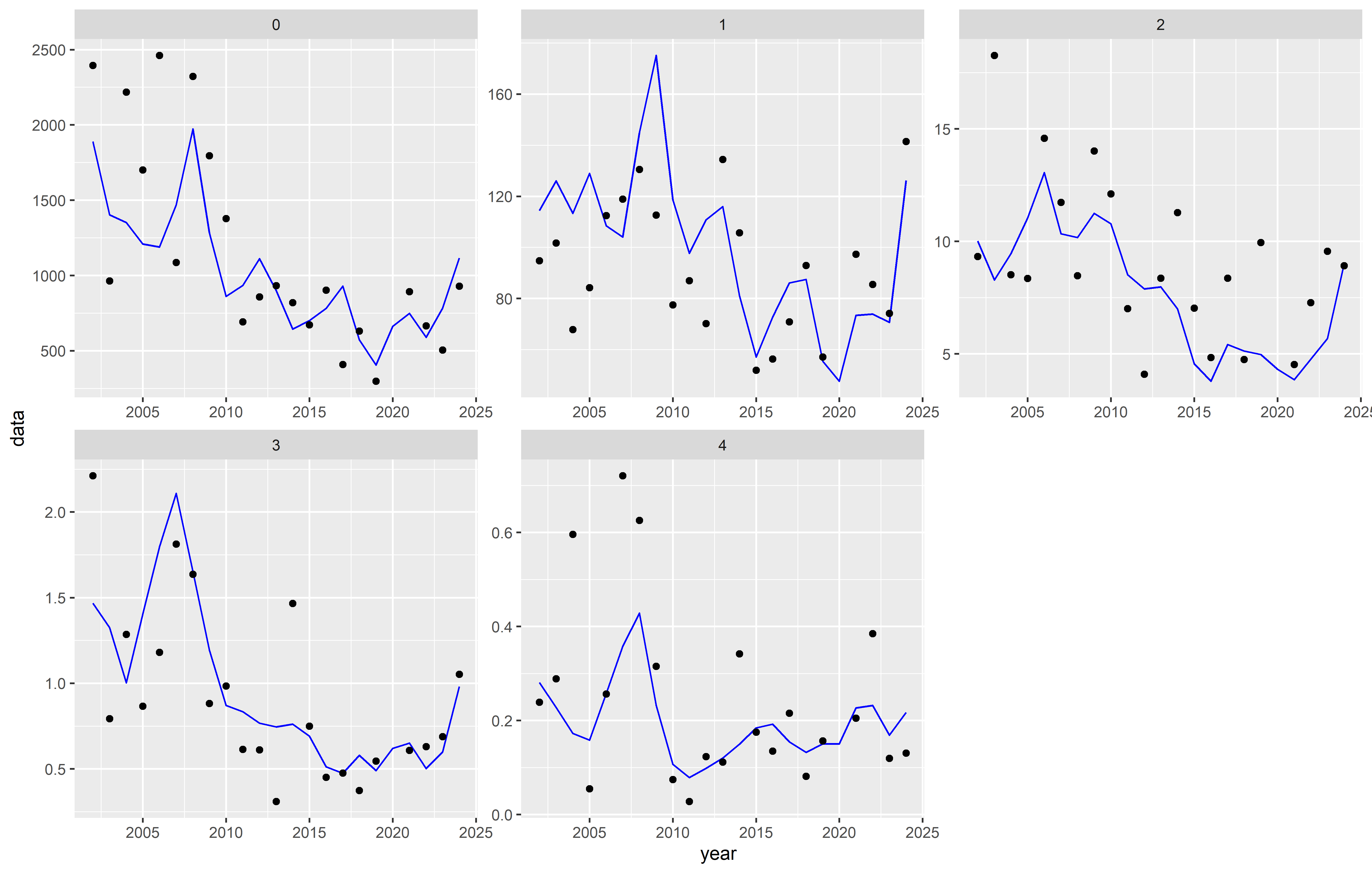
*Figure 11.- Estimated (dashed line) versus observed (solid line) total catch (in tons) over the historic period (left panel), and standardized residuals (right panel).*



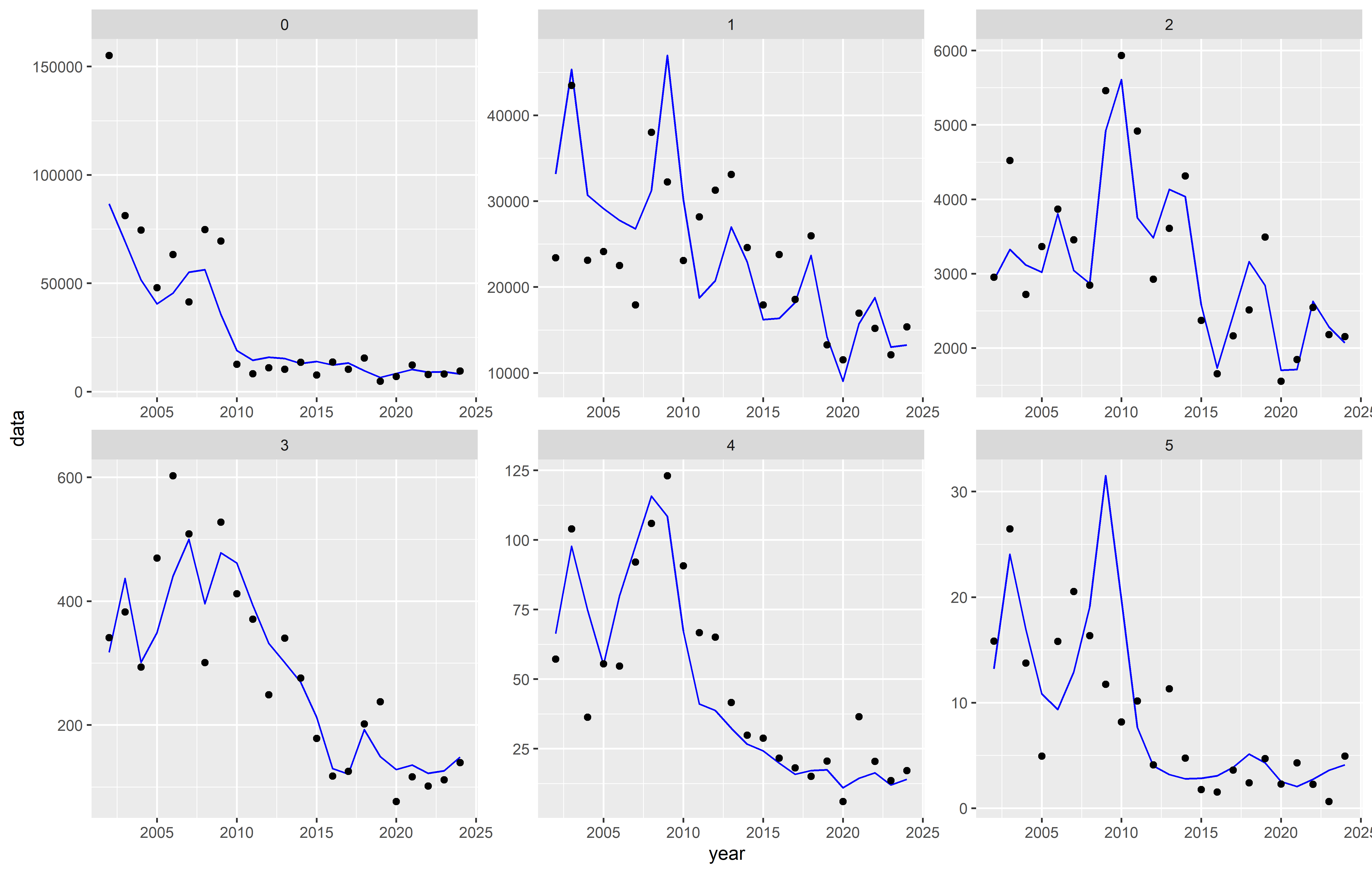
*Figure 12.- Estimated logarithm of the residuals for commercial catch at age (left panel) and abundance survey index by age (right panel).*



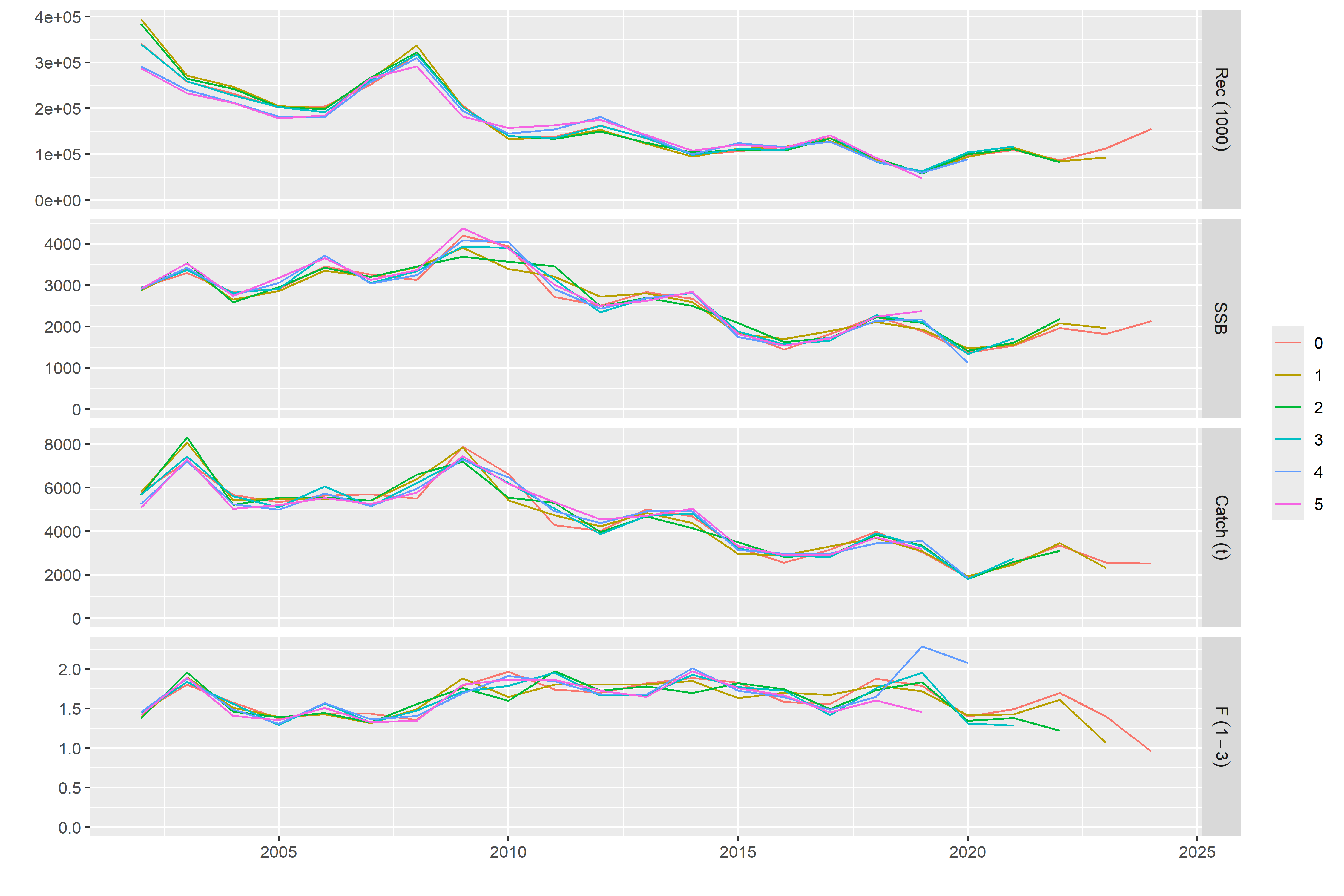
*Figure 13.- Estimated logarithm of the residuals over time for commercial catch and survey index by age. A smother is added to facilitate the identification of trends over time.*



*Figure 14.- Estimated (blue line) versus observed (black dots) abundance survey index by age over the historic period.*



*Figure 15.- Estimated (blue line) versus observed (black dots) commercial catch in numbers (x1000) by age over the historic period.*



*Figura 16.- Retrospective pattern. The Mohn´s rho value obtained for Fbar, SSB and recruitment was, respectively, -0.0714 0.0766 and -0.087.*

Model estimates:

The estimated catchability at age followed a negative exponential curve, since it was assumed like that when setting the qmodel. Since it is assumed constant over time, no changes in the year dimension are estimated (figure 17).

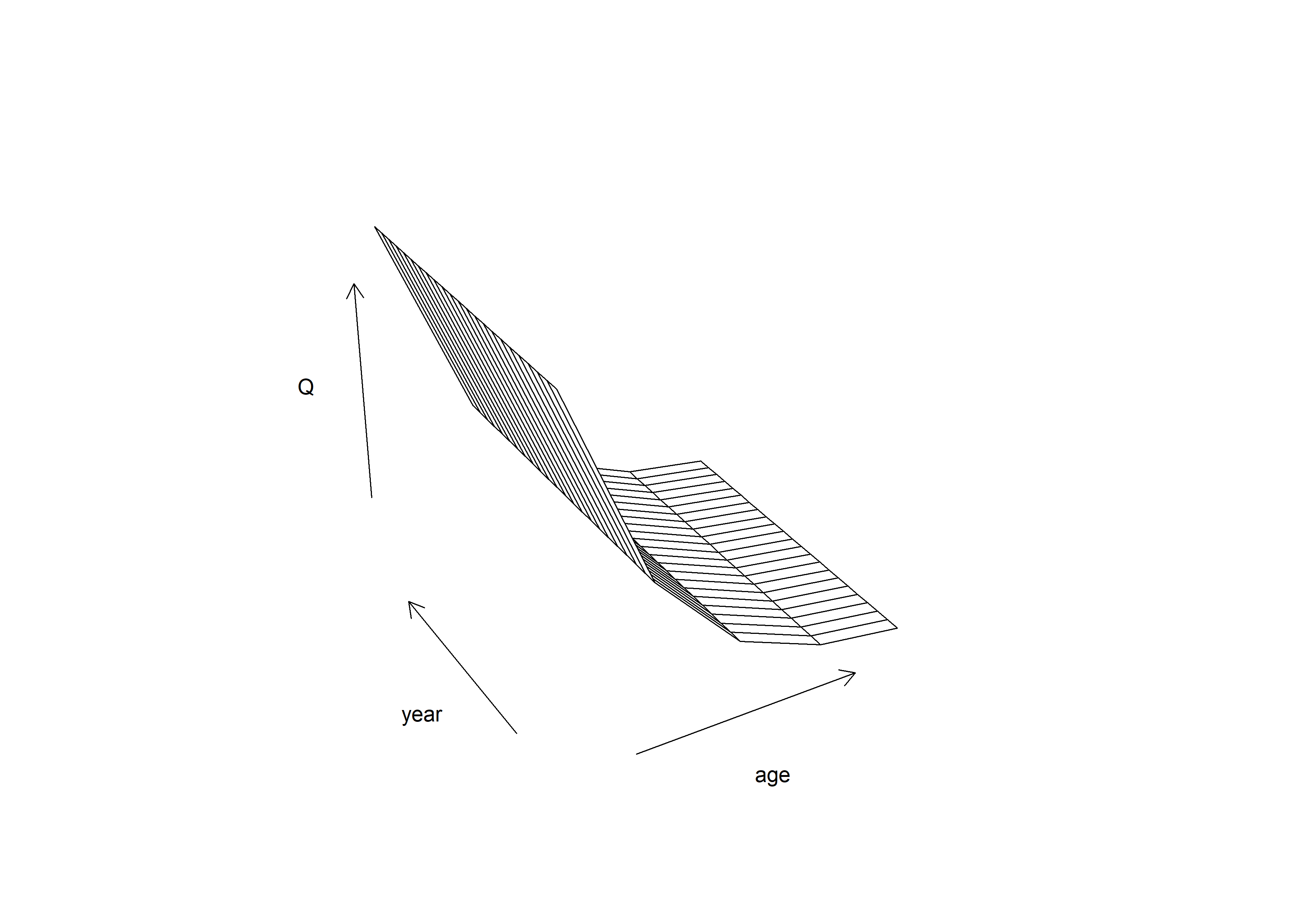
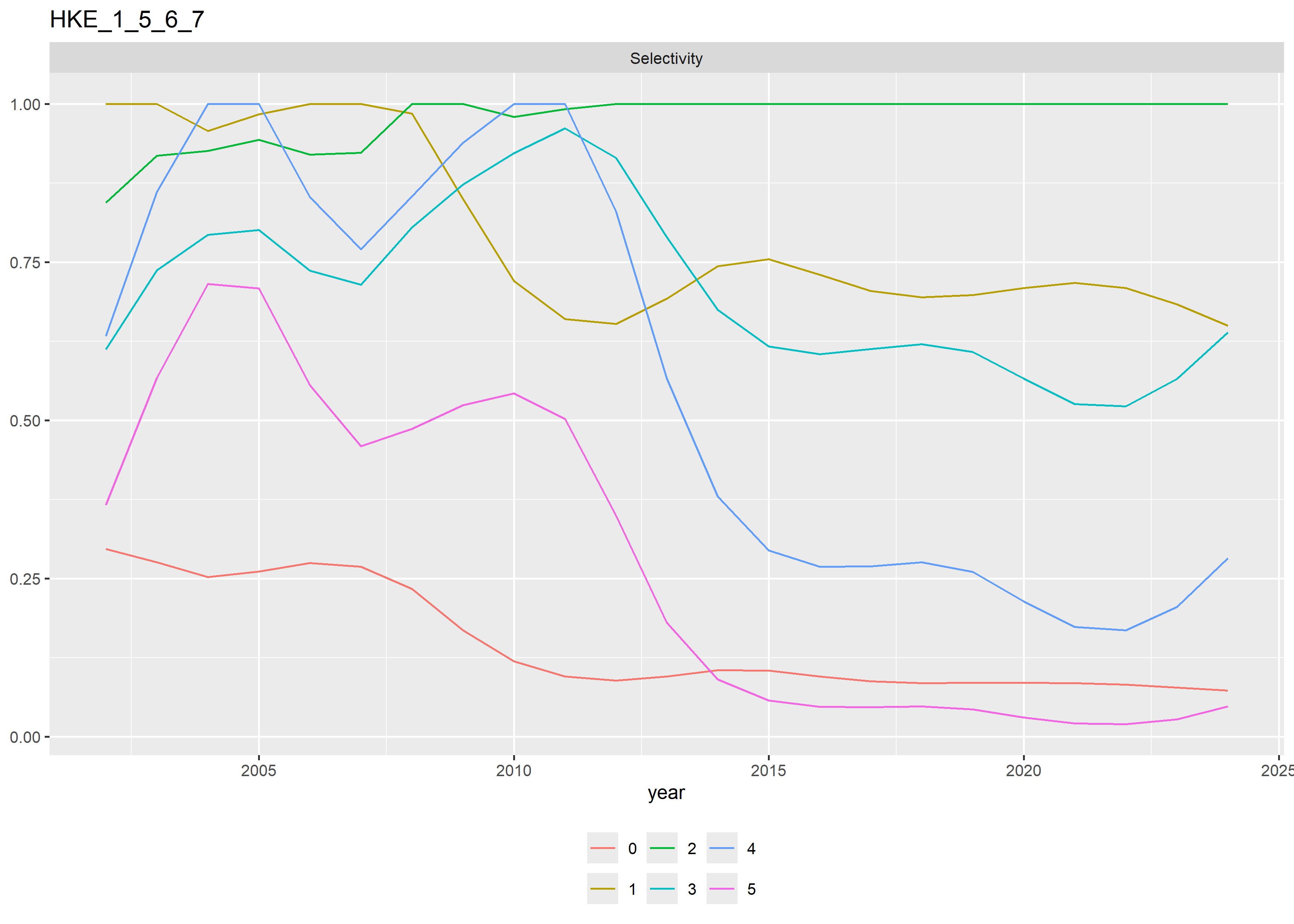


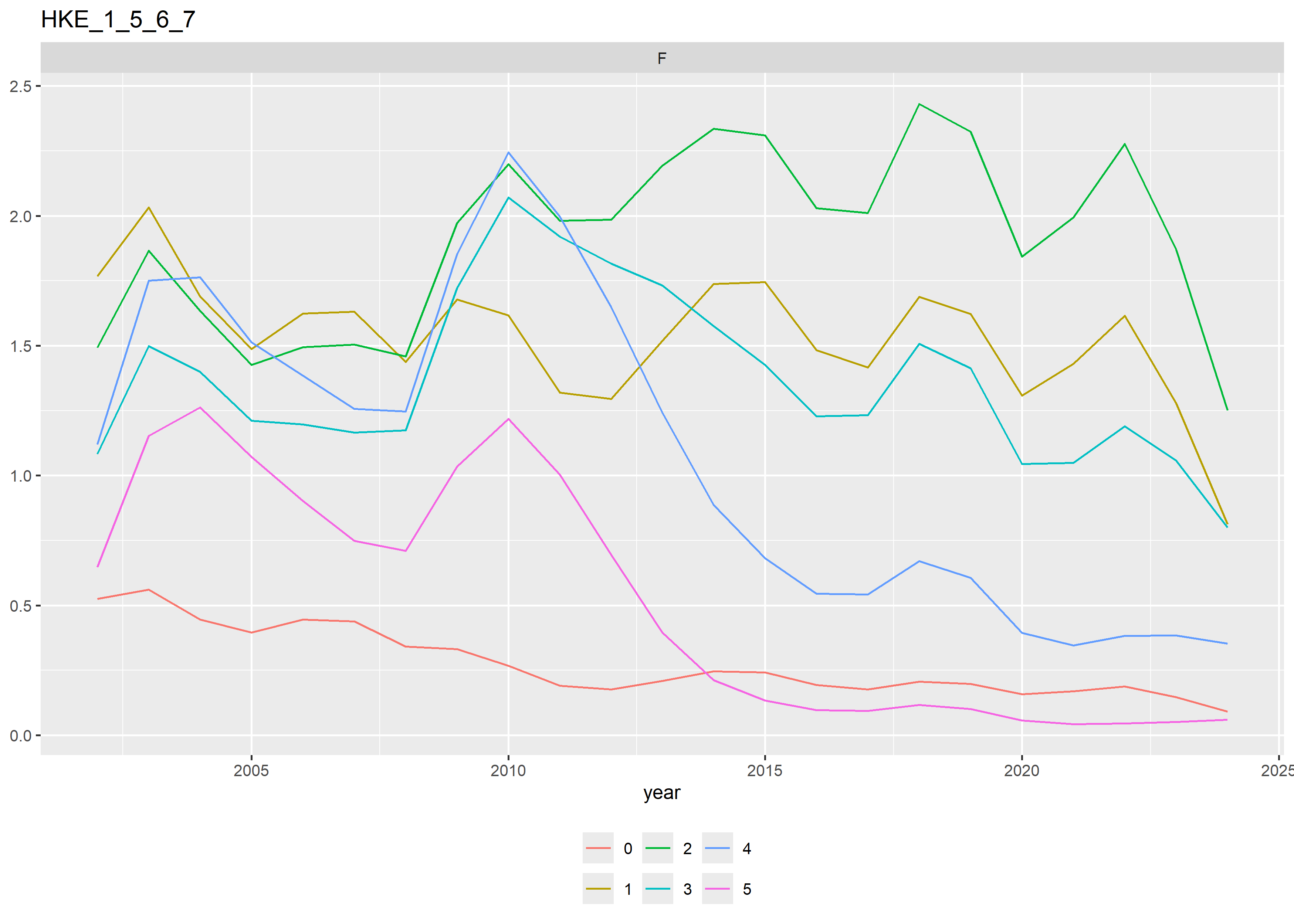
Figure 17.- Catchability curve, showing catchability at age over the historic period.

The estimated selectivity at age, and the estimated F at age over the historic period (figures 18 and 19 respectively), showed very similar trends. The pattern of selectivity and F at ages 3, 4 and 5 showed a declining trend starting around 2010, passing from a period of high to a stage of low selectivity/F. At the age 0, the estimated decline in selectivity started some years earlier, aprox. 2007. While selectivity at age 1 also experienced a decline in 2008, the F curve at this age didn´t show any clear trend or regime shift over the historic period. At the age 2, selectivity has always being very high, while the F at this age has been oscillating. Since 2022, the model estimates that F has started a declining trend at ages 0 to 3.

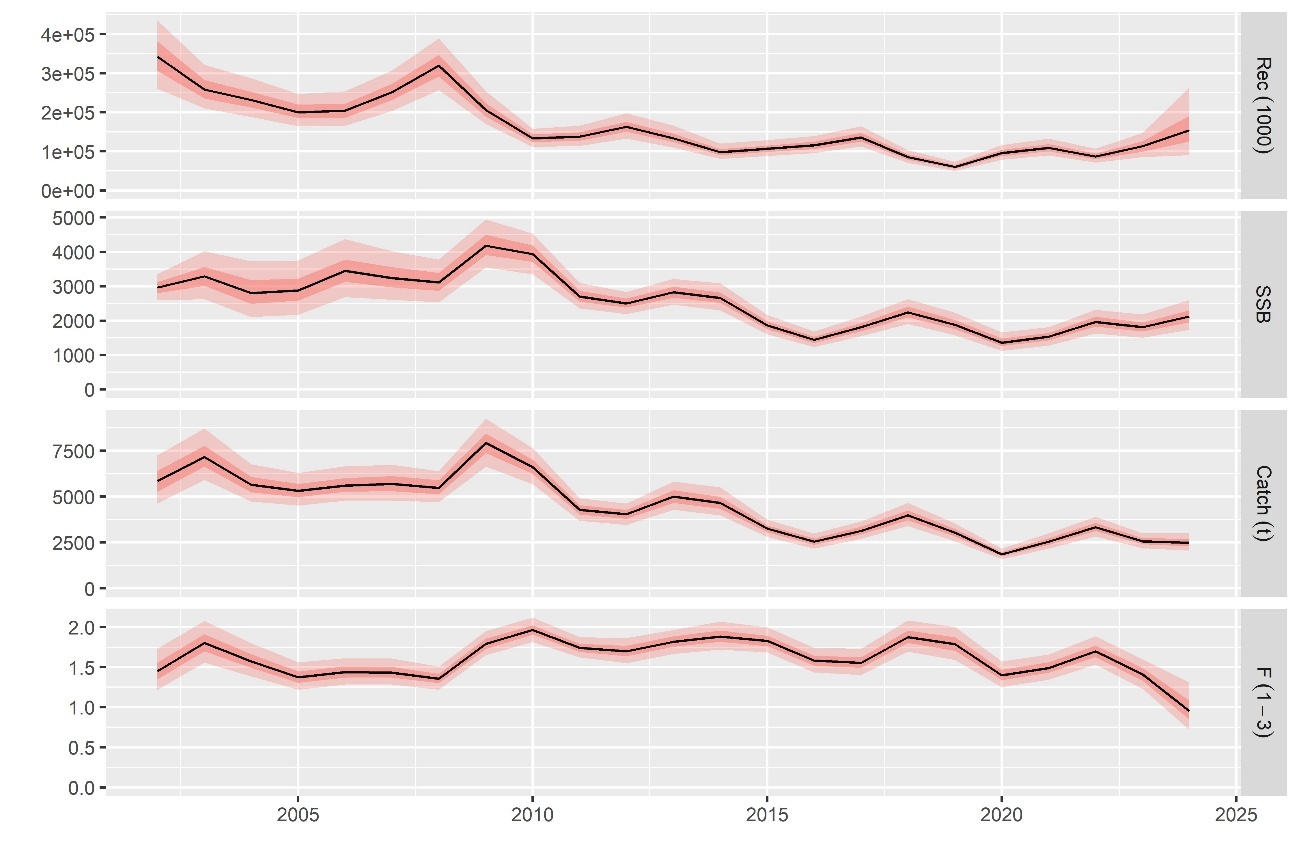


*Figure 18.- Estimated selectivity at age over the historic period*

The estimated catch (in tons), SSB (in tons), Recruitment (in x106) and Fbar are presented in figure 20 and table 4. The model estimates that since 2019 there is a mild increasing trend in recruitment and SSB, and a decreasing trend in Fbar, specially sin 2022. The Fbar estimated in 2024 is the lowest of in the historic period.



*Figure 19.- Estimated fishing mortality (F) at age over the historic period*



*Figure 20.- Estimated catch (in tons), SSB (in tons), Recruitment (in x103) and Fbar.*

*Table 4.- Estimated catch (in tons), SSB (in tons), Recruitment (in x103) and Fbar.*

|  |  |  |  |
| --- | --- | --- | --- |
| Catch | SSB | Recruitment | Fbar |
| 5786 | 2943 | 340685 | 1.448 |
| 7221 | 3291 | 258352 | 1.799 |
| 5656 | 2832 | 232050 | 1.574 |
| 5332 | 2905 | 201608 | 1.375 |
| 5634 | 3453 | 204280 | 1.438 |
| 5687 | 3255 | 251349 | 1.433 |
| 5491 | 3126 | 318356 | 1.356 |
| 7890 | 4198 | 206075 | 1.791 |
| 6618 | 3941 | 133030 | 1.962 |
| 4273 | 2712 | 137738 | 1.74 |
| 4019 | 2504 | 162458 | 1.699 |
| 5003 | 2828 | 134069 | 1.814 |
| 4671 | 2667 | 98234 | 1.883 |
| 3251 | 1868 | 106525 | 1.827 |
| 2541 | 1442 | 115476 | 1.58 |
| 3146 | 1816 | 135703 | 1.553 |
| 3974 | 2246 | 85267 | 1.875 |
| 3039 | 1886 | 60163 | 1.786 |
| 1864 | 1370 | 95803 | 1.398 |
| 2553 | 1532 | 108982 | 1.49 |
| 3341 | 1961 | 86734 | 1.694 |
| 2565 | 1813 | 112325 | 1.402 |
| 2502 | 2125 | 155033 | 0.954 |