Southern Hake Retrospective Analysis (WD-5. ICES WGBIE 2020)

Santiago Cerviño (IEO), santiago.cervino@ieo.es
Hugo Mendes (IPMA), hmendes@ipma.pt

5 May 2020

Contents

Introduction

Southern hake stock assessment model is a length-based model developed in GADGET (https://github.com/Hafro/gadget) and the latest assessment report is available in www.ices.dk (WGBIE 2019 Report). Gadget is a forwards age-length projection model. Quarterly from 1982 to 2018 with 0 to 15+ ages and 1 to 130 length classess.

The retrospective pattern was not an issue when this model was first implemented in 2010. However the magnitude of this retrospective pattern evaluated as Monh's Rho indices have increased in recent years moving from figures around 0.2 in recent years and raised to around 0.4 last assessment year (see Tab 1 and Fig 1).

Table 1. Rho figures in last 4 years including updated 2020 run

Name	Rho SSB	Rho F
Final Run 2017	-0.28	0.23
Final Run 2018	-0.3	0.24
Final Run 2019	-0.45	0.31
Updated Run 2020	-0.56	0.35

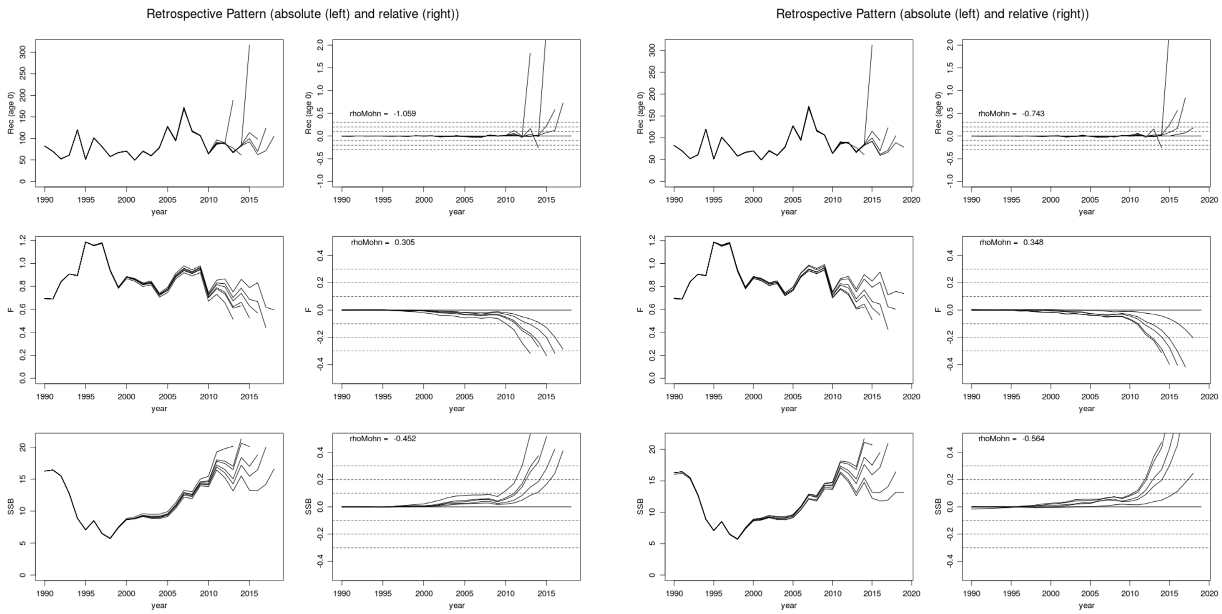


Figure 1. ICES WGBIE 2019 hake retrospective plot with Mohn's Rho figures

An analysis of plausible causes was first developed in ICES WKFORBIAS (Woodshole. USA. 9-17 Nov 2019). The approach consisted on exploring and testing alternative model configurations and their impact on hake retrospective pattern, quantified as a Mohn's rho index. Furthermore, this preliminary analysis can complement the others analysis and checks (convergence, likelihoods, residuals, etc) that can help to explain the sources of the retrospective pattern for this stock. Results suggested that catch underestimation after 2010 is the most probably cause. Although alternative explanations such as an increase of natural mortality or migration out of the area could also produce similar retrospective pattern.

The work initiated in WKFORBIAS was continued with additional scenarios and tests. A total of 53 scenarios were performed and the Rho index was calculated to any of them spending more than 1630 computation hours in the FinisTerrae-II equipment belonging to CESGA (Centro de Supercomputación de Galicia).

Scenarios analysis

Preliminary analysis include scenarios type "one data type out". First blocks of likelihood data types, afterwards running the modes without each likelihood data. Afterwards some scenarios addresing population dynamics uncertain parameters (growth and M), selectivity and catchability were also performed, and finally scenarios were overcatch was simulated and also scenarios were the first years of the time series were removed. In summary, the

- Biological realistic alternatives to growth and M.
- Selectivity realistic alternatives
- Catchability realistic alternatives
- Overcatch scenarios.
- Cut first years

Likelihood data out

Two different approaches have been applyed: first leaving appart big groups of data with similar characteristics (time trends data, length distribution data, fisheries dependent data or survey data) and second, depending on the results, leaving appart small groups or individual likelihoods data each time. The aim of this is to identify whether a specific data type is driving the retro pattern.

Leaving appart big groups of likelihood data

The hake GADGET model uses likelihood data that can be categorized in two different ways. Depending on the sampling origin, the data can be **catch-dependent**, i.e. quarterly length distribution data for landings and discards (with some gaps) and yearly LPUEs for two different fleets (with some gaps two) or **survey-dependent**; there are 3 different demersal surveys performed in 3 different areas covering the whole stock distribution providing hake data on yearly length distribution and abundance indices. On the other side, depending on the model dimension (time and size) we have two groups of likelihood data: **length distribution** data from catch and surveys and **time trends** from surveys and LPUEs. All these data are contributing to 19 likelihood functions, each one with an external weight. The total likelihood is the sum of these 19 product of weight * likelihood. Next plot shows the relative contribution of each likelihoods component to the total likelihood in the Southern hake ICES model.

Table 2. Gadget model Likelihood components

Description	period	area	Likelihood component
Landings -Length distribution	1994–lastYear	Iberia	Land1.ldist
Landings - Length distribution	1982–1993	Iberia	Land.ldist
Landings – Cadiz Length distr.	1994–lastYear	Gulf of Cadiz	cdLand.ldist
Spanish GFS- Length distribution	1982–lastYear	North Spain	SpDem.ldist
Port. GFS - Length distribution	1989–lastYear	Portugal	PtDem.ldist
Cadiz GFS- Length distribution	1990–lastYear	Gulf of Cadiz	CdAut.ldist
Discards - Length distribution	1994,98,99,2004–lastYear	Iberia	Disc.ldist
Sp GFS Abund: 4-19cm	1982–lastYear	North Spain	SpIndex15cm.1
Sp GFS Abund: 20-35cm	1982–lastYear	North Spain	SpIndex15cm.2
Sp GFS Abund: 36-51cm	1982–lastYear	North Spain	SpIndex15cm.3
Pt. GFS: 4-19 cm	1989–lastYear	Portugal	PtIndex15cm.1
Pt. GFS: 20-35cm	1989–lastYear	Portugal	PtIndex15cm.2
Pt. GFS: 36-51cm	1989–lastYear	Portugal	PtIndex15cm.3
Sp LPUE: 25-39 cm	1994–2012	North Spain	Spcpue15cm.1
Sp LPUE: 40-54cm	1994–2012	North Spain	Spcpue15cm.2
Sp, LPUE: 55-70cm	1994–2012	North Spain	Spcpue15cm.3
Pt. Stand. LPUE:25-39cm	1989–lastYear	Portugal	Ptcpue15cm.1
Pt. Stand. LPUE:40-54cm	1989–lastYear	Portugal	Ptcpue15cm.2
Pt. Stand. LPUE:55-70cm	1989–lastYear	Portugal	Ptcpue15cm.3

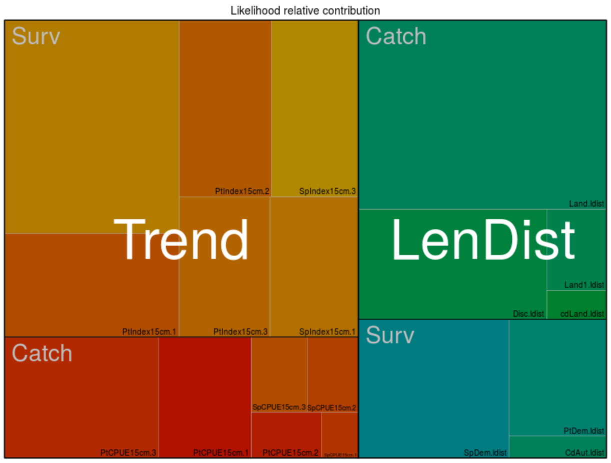


Fig 2. Relative contribution of the 19 likelihoods to the global fit. These are grouped in catch vs. survey dependent and length distribution vs time trends.

A preliminary retrospective analysis are performed deleting each one of these group of data. These are the 4 model names:

- **NoCatch**: without catch length distribution and LPUES data.
- **NoSurvey**: without survey length distribution and abundance indices data.
- **NoLength**: without length distribution from catch and surveys.
- **NoTrends**: without abundance indices (LPUEs and survey indices)

Summary table for

Description	Rho SSB	Rho F	Rho Rec	Comment
WGBIE19 Final Run	-.45	.31	-1.1	
No catch data	0.56	-1.5	0.2	NO CONVERGENCE. Parallel retro
No Survey data	**-0.18**	**0.17**	0.5	**F and SSB Reduction <0.2!**
No Length Dist data	0.33	-0.8	-0.2	NO CONVERGENCE. Parallel retro
No Time Trends data	**0.35**	**-1.3**	0.68	**OPOSITE PATTERN!!!**. Rare retro in period 2000-2005

There are some convergence problems. **However, the elimination of survey data (length distribution and time trends) reduce it below 0.2**. To follow this clue, some additional scenarios were run to identify the specific origin of this Rho value. To carry on with this, some additional scenarios were developed. In this case, to avoid convergence problems caused by leaving a group of relevant likelihoods together, the likelihoods were removed one by one or in small groups. This new analysis focused the problem in survey scenarios too.

Leaving appart Survey likelihood datae

The NoSurvey scenario was the one suggested that this data can be the cause of the observed retro. To continue this clue some additional runs were performed. NOSurvey likelihood data includes length distribution data and time trends data for two surveys split in 3 15 cm length groups from 5 to 50 cm

Description	Rho SSB	Rho F	Rho Rec	Comment
No Survey Ld	-0.49	0.32	-1.1	
No Survey Tr	-0.18	0.17	0.35	

Description	Rho SSB	Rho F	Rho Rec	Comment
No Survey Tr 5-20	-0.46	0.3	-1.8	
No Survey Tr 20 35	-0.18	0.17	-0.47	**HERE IT IS!!!!**
No Survey Tr 35-50	-1.1	0.5	-0.75	

When the Survey trend is given apart, the SSB Rho is reduced from -0.45 to -0.18 and F from 0.31 to 0.17. The same results are getting when we leave apart the length group of 20-35 cm. This preliminary analysis focuses the problem on the survey trend (lengths 20 to 35) data. However surveys are the more confident source of information because the well controlled sampling process. Furthermore, the length clash between 20 and 35 cm are well represented in both surveys. It is difficult to think that this data source is the cause of the problem, thought it is clearly affecting the pattern. Additional runs can be performed to indentify whether the PtSurvey or the SpSurvey data are causing this problem showed that is the combination of both. When we leave apart only one the Rho is quite less reduced.

Biological (growth and natural mortality)

Rationale: Current model included M=0.4 for all ages; Linf=130; k and **beta** (dispersion parameter) are model estimated. Biological studies show that there are alternatives to this "best model" approach decided in last benchmark.

The following scenarios were explored - **M 0.3.** M = 0.3 for all ages. Initially K was estimated but the K and M are higly correlated and k was bounded at 0.1. The models was re estimated keeping k=0.17, that was the value estimated with M=0.4 model. - **M 0.3 Ages 0-1.** The same as previous but Mage0=1 and Mage1=0.6 (based on predation) - **M 0.3 Ages 0-1 M9plus.** The same as previous but M increases for ages older than 9 based on senescence. - **Linf100K17M28.** Median figures from an hierarchical bayesian analysis based on hake (12 spp data) life history invariants. - **beta10DeltaL10.** Beta is currently model estimated driven to quite low figures. Correction of growth dispersion to reduce the model estimated dispersion.

Description	Rho SSB	Rho F	Rho Rec	Comment
M 0.3	-0.29	0.24	-0.81	Slight reduction of Retro
M 0.3 Ages 0-1	-0.32	0.26	-0.88	No clear retro pattern. Convergence?
M 0.3 Ages 0-1 M9plus	-0.32	0.23	-0.91	Slight retro reduction. Increased before 2000.
Linf100K17M28	-0.29	0.27	-0.91	Slight retro reduction
beta10DeltaL10	-0.47	0.32	-0.65	Similar than base model

In general, alternative "realistic" biological parameters **improve only slightly the retrospective pattern**. However an M around 0.3 could contribute, at least partially, to the solve the problem.

Selection alternatives (changing fishery process)

Rationale: Current selection for recent years include separate landings (logistic from 1994 to now) and and discards (asymetric normal from 1994 to now). Other "fleets" are separated in the past. The scenarios explored are related with alternative selections for fleets in recent years using dome shaped instead of logistic.

- **Sel Change 2005**. Current landing "fleet" is split in two periods: 1994-04 and 2005-now.
- **Dome shaped (estim end)**. Current landing "fleets" with logistic selections are changed to dome shaped selection.

- **Dome shape (cte end)**. Current landing "fleets" with logistic selections are changed to dome shaped selection fixing the parameter that defines the right part function.

Description	Rho SSB	Rho F	Rho Rec	Comment
Sel Change 2005	-0.47	0.23	-1.1	Retro converge in 2005 but get wider before
Dome shape (estim end)	0.9	-4	0.35	Weird but interesting. Convergence problems. But a thread to follow.
Dome shape (cte end)	-0.48	0.29	-1	Wider retro in past. Pivot point around 2010

The first and third scenarios do not change substantially the base case. However the second one, that do not converge in some peels can help to future tests.

Catchability process

Rationale: Current catchability models are linear in log scale. However there are reasons to think that some dense-dependent process can be in act. For instance, in the periods of large abundance (2005-2010) the density increases outside the survey area. The scenarios explored are the following:

- **surveys 2 params**. The two surveys are now modeled with dense-dependent catchability
- **Pt CPUE 2 params**. The two CPUEs are now modeled with dense-dependent catchability

Description	Rho SSB	Rho F	Rho Rec	Comment
surveys 2 params	-0.84	0.42	-0.53	Worst than Base. Slight improve in recruitment.
Pt CPUE 2 params	-0.62	0.37	-0.98	Worst than Base

Overcatch scenarios

Rationale: Current catches used in the Southern hake stock assessment model are not the official ones but estimated. The sampled vessels catch and effort are used to raise each metier catch to total effort. The reason to set 2010 as the year to increase the catch are twofold: (1) estimation system changed after 2010, now is not dependent on the fishing sector collaboration; there were important changes in the regulation (e.g. share of quotas by vessel) and increased of inspection which resulted in a weaker collaboration; on the other side (2) there are more diagnostics showing that something happened after 2010, such as a increased of the retrospective pattern or some survey residuals that start to raise. To test the impact of catch overestimation on the retrospective pattern some scenarios were run with overshoting after 2010 of 10%, 20%, 30% and 40%. the results are presented in the following table.

Description	Rho SSB	Rho F	Rho Rec	Likelihood
0%	-.45	.31	-1.1	1242
10%	-0.34	0.25	-0.9	1229
20%	-0.23	0.19	-0.82	1195
30%	-0.15	0.13	-0.77	1182
40%	-0.08	0.08	-0.67	1173

Results are quite consistent since all the indicators used show the same continuous behavior: the Rho's SSB decreases with increased overcatch, the same for Rho F and Rhos recruitment and also for the likelihood, i.e. an increase of catches after 2010 gives a good Rho value with a better model fit.

Which likelihood figures are contributing more to this likelihood decrease? The following table summarise how change each likelihood component.

Lik	0%	10%	20%	30%	40%	50%
Len dist	100%	101%	100%	100%	100%	100%
SP Survey	100%	97%	95%	92%	90%	91%
PT Survey	100%	96%	92%	89%	87%	83%
SP CPUE	100%	97%	95%	93%	89%	93%
PT CPUE	100%	93%	97%	97%	96%	97%
Total	100%	96%	95%	93%	91%	90%

In general we see that the inclusion of additional catch does not affect the quality of the fit of the length distribution but those of the survey trends. This mean that adding catches after 2010 helps to explain better the whole likelihood data, although specifically the trends data for all, surveys and CPUEs.

Time series cut

Rationale: The initial time series (1982-93) of Spanish catches and length distribution in the Northern area were estimated in the 90s. There were not a records of catches by stock area since the fleet was allowed to catch in both areas. The bigger vessels were catching more frequently in the North, even though there were some missing that required stock assignment. IEO and AZTI scientist reviewed that data and made an assignment by stock. This work was critical in the 90s since this information was required to do both assessments. However the time series is now larger and an assessment can be developed without this information. In fact WKFORBIAS recommended to this kind of exercised when information is less reliable.
A GADGET run from 1994 to 2019 was performed. Retro plot and Rho figures were calculated in two different ways. First using the last assessment year estimated parameters for all the peels and second, using different starting figures to each peel. Results are presented in next figure.

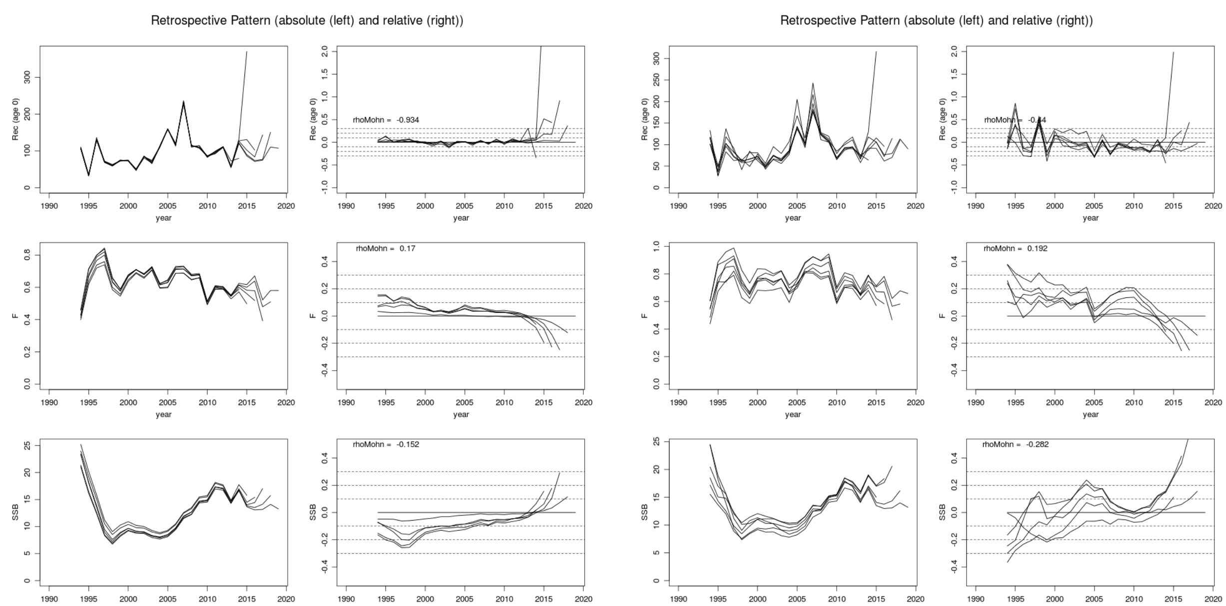
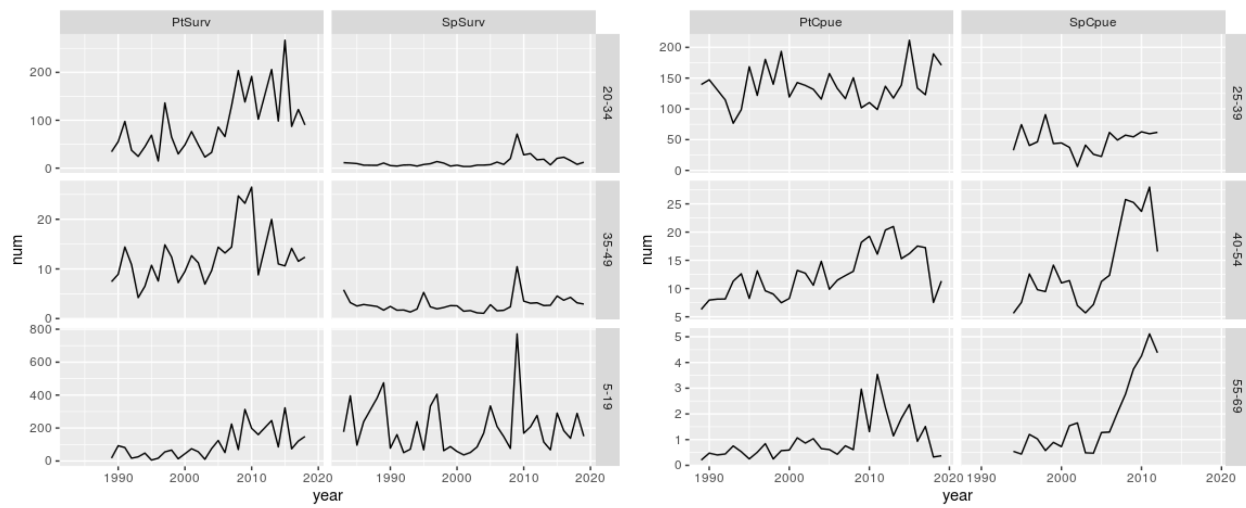


Figure shows the retrospective plot with peels estimated starting with same parameters (Left) and peels with different parameters (Right). In general we can see that the Retrospective pattern improves. However there is an issue with the convergence, that gets worse. Next table shows the Rho figures for these two runs and final (2019) and updated (2020) runs.

Name	Rho SSB	Rho F
Run 94 =param	-0.15	0.17
Run 94 != param	-0.26	0.19
Final Run 2019	-0.45	0.31
Updated Run 2020	-0.56	0.35

Discussion

Two possible explanations for the retro have been identified: Contradictory signals in abundance index trends and Catch underestimation. And both could be linked. Next plot shows the raw abundance trend data used to calibrate the GADGET model. These are updated to 2019.



Contradictory signals source inside the same index or among indices. Examples in the same index are SPSurv, with recruits (length 4-19 cm) without a clear trend, mainly noise, although upcoming classes (lengths 20-55) showing an increase after 2005. Something similar happens with PtCPUE, with length group (25-49 cm) increasing after 2010 and although upcoming classes (lengths 50-79 cm) show a decrease after 2010. Conflict signals among different indices can be seen, for instance, between PtSurvey (20-34), increasing after 2005, and PtCPUE (25-39), without this increase; also between SpSurvey (4-19), with no clear trend, and PtSurv (4-19) increasing after 2005; also the strong decrease of abundance after 2010 in PtCPUE (40-54 and 55-69) that is not seen in any other index.

These data conflicts affect the model fit as can be seen in the 2019 model residuals. Furthermore, when adding new data each year, the new model fit can give more credibility to a different index affecting the population abundance through the index catchability and its trends. If this change in index dominance is consistent in time it can produce a retrospective pattern. What it is observed in the change in population biomasas (SSB) estimated by GADGET is that the population increase in 2005-10 is not reproduced in the recent GADGET models, i.e. the index showing this change are having less contribution in recent models.

An additional source of mortality after 2010 produces a reduction of the retrospective pattern but also improves the quality of most of the indices fit, reducing also to total likelihood. However, although the scenario simulated implements an increase in catches, a similar result could be obtained with an increase of natural mortality or even with an increase of migrations out of the stock. Although an increase on catches can happen there is not alternative information to run the model. Natural mortality depends on their predators that are mainly hake (cannibalism) and common dolphins. The available information stomach contents although partial, it shows an increase of hake in hake diet after 2004, when the increase of hake abundance was first observed. Information on migrations out of the stock are also scarce. Tagging in Spain have not got success in recoveries and tagging n France, although successful in recoveries do not provide information on migrations out of Iberian peninsula. However there are some non-direct information that can help. These are the genetic studies showing that both stocks are quite connected and also the recent raise in abundance in the North stock, although none of these process implies necesarially migration from South to North.

A common pattern in all (most) the runs performed is that the retro get worse in last two years. Some additional data are supporting a decrease in abundance in last 2 years!!

Among the scenarios developed we found there were other things that help to reduce the retrospective pattern. First the reduction of natural mortality from 0.4 to 0.3 (with different M structures), although in any case Rho figures were reduced bellow rule of thumb (0.2). It also worked the cut of the time series starting in 1994 instead of the usual 1982. In this case there is a clear issue with the convergence that requires additional work to solve it. The reason for this can be that now the conflicting calibration indices are more in play. Furthermore, such a change would require a re estimation of the weights for all likelihoods, thick that was not tested.

Conclusions and further work

The main goal of this work was to explore the causes of the strong retrospective pattern with the aim of produce an alternative model to give advice to 2021 catches. Although this target was not achived, we better understan now of the potential causes producing the retrospective patter. HOwever there are still some options that can be tested in a short time.

- The option to increase mortality: fishing, natural (or even migration) do not seem reliable given the lack of external information to support any plausible alternative.
- Whether convergence problems are caused by conflicting signals inside calibration indices, it can be explored some alternatives such as changing CPUEs-at-legth with biomass CPUEs.
- Short time series combined (or not) with a lower M (= 0.3) could work in terms of retrospective pattern. However it requires a re-estimation of likelihood weights and probably dealing with convergence problems. This makes difficult to estimate the time needed to do it.

However it depends on the time available in the WGBIE but also whether we are able to go for a inter-becnhmark with this uncertain options.