Analysis of causes of retrospective pattern in Southern hake stock assessment with GADGET

Santiago Cerviño López
Instituto Español de Oceanografía (C.O. de Vigo)
e-mail: santiago.cervino@ieo.es

Summary
The strong retrospective pattern preclude the use of Souther hake stock assessment for fishery advice. An analysis of plausible causes developed in ICES WKFORBIAS (Woodshole. USA. 9-17 Nov 2019) was performed. Results suggest that catch underestimation after 2010 is the most probably cause. Alternative explanations such as an increase of natural mortality or migration out of the area could also produce similar retrospective pattern. Although there are not evidences supporting the latest further work is needed to evaluate to what extent all the three can contribute the retrospective pattern.

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 been increased since them and particularly in recent years, moving from figures around 0.2 in latest years and raised to around 0.4 last assessment year. Next table shows the Mohn's Rho values for SSB and F in the last 3 assessment years:

Name	SSB	F
Final Run 2017	-0.28	0.23
Final Run 2018	-0.3	0.24
Final Run 2019	-0.45	0.31

The working plan for WKFORBIAS will mainly focus on ToR *b) "Categorize the potential causes for bias as identified in ICES stock assessments with respect to factors like stock longevity, quality of input data (catch and survey), model assumptions and environmental changes"* and consists on exploring and testing alternative model configurations and their impact on hake retrospective bias, 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

1. Likelihood data out
2. Biological realistic alternatives to growth and M / Selectivity realistic alternatives / Catchability realistic alternatives
3. Over-catch scenarios.
4. Residual analysis

1. Likelihood data out scenarios

Two different approaches have been applied: first leaving apart 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 apart 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.

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 2019 ICES model.

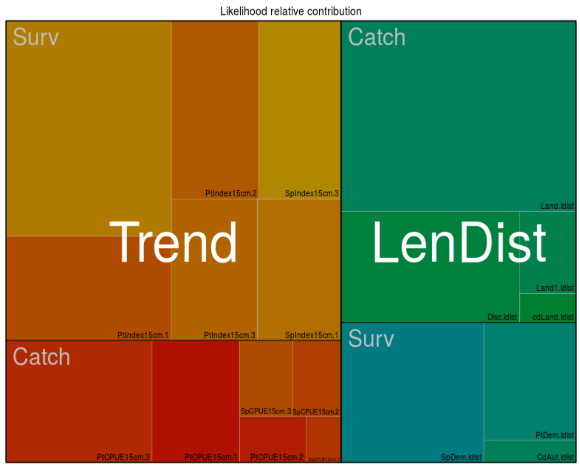

Fig. 2. Relative contribution of the 19 likelihoods to the global fit. These are grouped in catch vs. survey dependent and length distribution vs time trends. Survey trends and catch trends (CPUEs) are split in Portuguese and Spanish, and each one in 3 length groups. Catch length distribution includes 4 different likelihoods for discards and landings in different periods; and survey length distribution includes 3 different surveys.

The contribution of different likelihoods to the overall likelihood value is dominated by the Survey trends, and inside these, the main contributor is Group 2 of the Spanish survey (20 to 35 cm). This was not the case when the model was developed. At this time the 4 main groups had similar contribution. Since the weights of these components have not changed, the cause of this change has to be the quality of the fit that has got worse in recent years.

Next table shows the impact on retrospective pattern (measured as Mohn's Rho Index) of survey trends likelihood components.

Type	SSB	F
No Survey Tr	-0.18	0.17
No Survey Tr 5-20	-0.46	0.3
No Survey Tr 20-35	-0.18	0.17
No Survey Tr 35-50	-1.1	0.5

Type refers to the likelihood component that is given apart in each run. 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. A more extensive analysis (not presented here) leaving apart other likelihood components shows no impact on the Monh's Rho index.

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 the only thread to follow.

2. Biological, selectivity and catchability scenarios.
Based on the expert knowledge 12 alternative scenarios were developed for M, growth, selectivity and survey catchability.

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 benchmarck. Alternative models consider lower M (0.3): M=0.3 but with higher M for young ages and also with higer M for older ages. Changes in growth were also considered

Selection scenarios changing fishery process. Current selection for recent years include separate landings (logistic from 1994 to now) and and discards (asimetric normal from 1994 to now). Other "fleets" are separated in the past. Three scenarios were explored with alternative selections (dome shaped) for fleets in recent years (after 2005).

Current catchability models are linear in log scale. However there are reasons to think that some dende-dependent process can be in act. For instance, in the periods of large abundance (2005-2010) the density increases outside the survey area. Two different scenarios with alternative dense-dependent catchability configuration were also explored.

None of these provide a clear improvement of the retrospective pattern.

3. Overcatch scenarios

Current catches used in the Southern hake stock assessment model are not the official ones but scientifically estimated. The estimation system changed after 2010 and there are reasons to think that it could be still underestimated. However the amount are uncertain. To test the impact of catch overestimation on the retrospective pattern (Monh's Rho index) some scenarios were run with overshooting after 2010 of 10%, 20%, 30% and 40%. the results are presented in the following table.

Catch increase	SSB	F	Likelihood
0% (WGBIE 19)	-0.45	0.31	1242
10%	-0.34	0.25	1229
20%	-0.23	0.19	1195
30%	-0.15	0.13	1182
40%	-0.08	0.08	1173

Results are quite consistent since for the 3 indicators; when catch increases after 2010 the Monh's Rho index for SSB and F approaches to zero. Furthermore, the quality of the fit (likelihood value) is also reduced. Results suggest that catch underestimation after 2010 can explain the observed retrospective pattern.

4. Residual analysis of survey trends

Two possible explanations for the retro have been identified: Contradictory signals in surveys and Catch underestimation. And both can be linked?

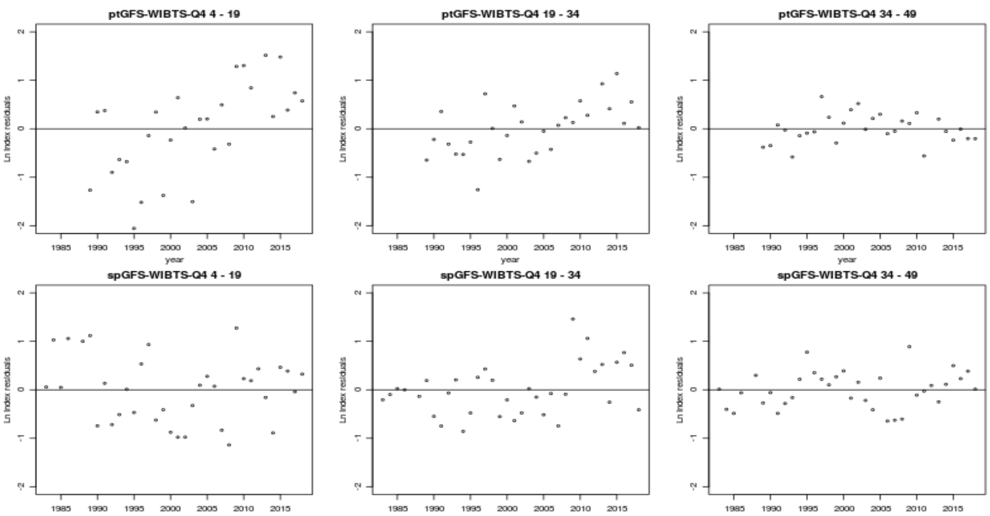

Fig. 2. Log scale model residuals for Pt Survey (upper pannels) and Sp Survey (lower pannels) at length classes 4-19 cm, 19-34 cm and 34-49 cm (from left to right).

Model residuals shows a clear positive pattern for both surveys (Sp and Pt) at length groups 4-19 cm and 19-34 cm after 2010. This is less clear for Sp Survey (4-19 cm). Largest length groups (34-49 cm) do not show this pattern. The model is not able to follow the estimated survey abundance by these two surveys after 2010. There is not reason to think that have been changes in surveys catchability.

Discussion

Contribution to ToR 2 *"Categorize the potential causes for bias as identified in ICES stock assessments with respect to factors like stock longevity, quality of input data (catch and survey), model assumptions and environmental changes"*. Methodology implemented to find out the causes of Southern hake retrospective pattern follows a simple approach changing models settings (data or process) and quantify the impact on Monh's Rho index. The settings explored can be classify in 3 different groups: 1) eliminate individual (or groups) likelihood data to identify whether one of them (or a group) is causing the pattern; 2) change critical parameters (M, growth) or process (selectivity function) and 3) explore catch misreporting. None of them alone provides a clear solution but the combination of them, specifically the elimination of survey likellihood data with misreporting scenarios and survey residual analysis put the focus on misreporting as a feasible cause of retrospective pattern.

Linking the residuals of survey trends fit with overshooting scenarios reveals that something strange happened after 2010. There is not any evidence supporting changes in the population at this time. However there has been many changes in the fisheries management and fishermen behavior. Individual quotas were implemented in an important part of the catches and fishermen complains about the insufficient quotas. Furthermore, we missed the fisherman support we used to have to validate the catches used in the model. The catch data we are using now are the best information available although could be underestimated. The analysis performed here support this.

Alternative explanations for the misreporting scenarios can be mimic with other processes. From a simulation point of view an increase of natural mortality or migrations out of the stock area after 2010 can also produce similar results explaining the retrospective pattern. However, although there is not any evidence supporting these behavior in this period, further work is needed to check to what extent these processes could contribute to the retrospective pattern.

An additional problem of this analysis is the convergence problem that makes that some scenarios finalize on non realistic results. Convergence problems are signal of complex likelihood space with alternative local solutions. This can be caused by confounding information from different likelihood data or model misspecification. Could this kind of undefined likelihood maximize the retrospective pattern? If this is the case it should be expected than a model with less retrospective patter also have less convergence problems.

Further work is needed to check the potential contribution of different plausible causes of retrospective pattern. Catch (landings and discards) should be reviewed and fishermen support in this review would be quite valuable. Current catch data used in the model are scientific estimations and the process to estimate them should also be reviewed. Alternative hypothesis work should focus on changes in predator abundance (dolphins and hake) and stomach contents to check whether a change in M after 2010 is feasible. Migration out of Southern stock is more difficult to validate without specific tagging experiences although some indirect evaluation could be done based on recent papers estimating genetic migration and also looking at abundance in adjacent stocks.