TOWARDS PROVIDING SCIENTIFC ADVICE FOR INDIAN OCEAN YELLOWFIN IN 2020

Agurtzane Urtizberea[1], Massimiliano Cardinale[2], Henning Winker[3], Richard Methot[4], Dan Fu[5], Toshi Kitakado[6], Carmen Fernández[7], Gorka Merino[8]

SUMMARY

In 2018 the advice of yellowfin tuna in the Indian Ocean (YFT) was based on a grid of 24 models, where all models were based on the age and length structured integrated assessment model Stock Synthesis (SS). However, due to several issues in the data inputs and model assumptions, the Science Committee of IOTC (SC) recommended a workplan to improve the YFT assessment. Therefore, in this document, based on the comments of the WPTT21, two different processes were conducted: i) some of the basic assumptions on the assessment model were analyzed in details and ii) a new procedure on how to select the models to be included in the final grid used for the advice is presented.

The current model treats seasons as continuous years, and this complicates the settings of the model as well as the interpretation of the results. Therefore, with the aim of simplifying the model but at the same time improve the understanding of the modeling part of the key processes in the dynamic of the stock such as movement and recruitment, we transform the non-seasonal model into an annual model with seasons. The models were compared using diagnostics where the fits to the data, the prediction skills and the retrospective pattern are used to evaluate the performance of each model. The results are promising but still more works need to be done with the annual model before using it for assessment.

The other process analyzed in this study is the selection of the models to be included in the final grid used for advice. In WPTT21 the group discussed which were the main axis of uncertainties in the model assumptions and proposed a grid of models that could cover that uncertainty. In this study, based on that original grid we present different hypothesis that encapsulate the main axis of uncertainties of the assessment models, and present a new procedure leading to the selection of the models to be included the final grid used for providing the advice.

KEYWORDS: yellowfin tuna, stock assessment, stock synthesis, diagnostics, grid.

1 AZTI, aurtizberea@azti.es;
2 SLU, massimiliano.cardinale@slu.se;
3 JRC, henning.winker@gmail.com
4 NOAA, richard.methot@noaa.gov
5 IOTC, Dan.fu@fao.orgda;
6 TUMSAT, kitakado@kaiyodai.ac.jp
7 IEO, carmen.fernandez@ieo.es
8 AZTI, gmerino@azti.es;

Contents

Introduction

The advice proposed in 2018 was given based on a reference grid of 24 models (Fu et al. 2018, Fu et al. 2018b) configured using Stock Synthesis (Methot and Wetzel 2013). However, there were several uncertainties around the stock assessment data inputs and assumptions, and therefore a workplan to address these issues was recommended by the SC of IOTC. This lead in 2019 to a new stock assessment model, although, new management advice could not be provided due to the complexity of the work, lack of agreement on key model aspects and time constraints during the meeting. For those reasons, the stock status was determined on the basis of the 2018 assessment integrated across a grid of 24 different model configurations aimed to encapsulate the main axis of uncertainties of the assessment.

In this study, with the intention of improving the assessment of yellowfin tuna in the Indian Ocean (YFT) two different components of the process that constitute the advice were analysed: i) the basic structure of the assessment models was analysed in details with the aim of improving the model and ii) a new procedure on how to select the models to be included in the final grid used for the advice is presented. One of the main complexities of the YFT assessment models included in the grid of 2018 is that all models treat seasons as continuous years. These settings make the model rather complex and the outputs difficult to interpret. Therefore, with the aim of simplifying the model, a preliminary analysis was done by converting the original model into an annual model with seasons. The performance of annual model was analysed based on diagnostics using the library ss3diag (Winker et al. 2020).

In the second part of this study, we present the key uncertainties on several aspects and assumptions of the model agreed at WPTT21. We represent the uncertainties through different model configurations and evaluate the performance of each of the models in terms of diagnostics using the library ss3diag (Winker et al. 2020).

Diagnostics

Below is the diagnostics used to evaluate the performance of the models in terms of fit to the data, prediction skills, and retrospective pattern. These analyses were performed with the library ss3diag (Winker et al. 2020) developed in R and are listed below:

-The runs test was estimated to evaluate whether residuals of the CPUEs and length frequency distributions were normally distributed or/and had time trends (Winker et al., 2018). A non-random pattern of residuals may indicate that some heteroscedasticity is present, or there is some leftover serial correlation (serial correlation in sampling/observation error or model misspecification). If the runs test indicates that the residuals are not larger than 1 then that means that the fit of the CPUE index for example is good. Runs test provides a significance level so that “pass” and “fail” of each residual time series can be statistically evaluated.

- Retrospective analysis was done to evaluate the reliability of parameter and reference point estimates and to reveal systematic bias in the model estimation. It involves fitting a stock assessment model to the full dataset. The same model is then fitted to truncated datasets where the data for the most recent years are sequentially removed. The retrospective analysis was conducted for the last 5 years of the assessment time horizon to evaluate whether there were any strong changes in model results. Mohn’s rho of retrospective pattern and forecast was estimated. The retrospective pattern are sensitive to the life history parameters and Hurtado-Ferro et al. (2014) proposed that for long-lived species values of Mohn's rho index higher than 0.20 or lower than -0.15 (upper and lower bounds of the 90% simulation intervals for the flatfish base case) should be cause for concern and taken as indicators of retrospective patterns.

- Hindcasting analysis was done following a similar analysis as in Kell et al. 2016 to evaluate model prediction skill of the CPUE. When conducting hindcasting, a model is fitted to the first part of a time series and then projected over the period omitted in the original fit. Prediction skill can then be evaluated by comparing the predictions from the projection with the observations using for example the MASE indicator (Hyndman and Athanasopoulos, 2013). The CPUE performs well if the MASE value of the hindcasting is lower than the value of 1 when predicting the index one year ahead.

Preliminary model transformation with season within years

A base case was chosen from the grid of the assessment of 2018 called M1 and modified to have seasons within years. However, two different approach were followed in the transformation of the model with two different model configurations as a result. One model called M2, which has very similar settings and follows the same assumptions as M1 while a second model called M3, which is the simplified version of M2. Table 1 shows the differences between the three models. The models were compared following the diagnostics described above. The results are very promising; the two models show similar pattern in terms of diagnostics and although M1 has the best score, the three models have very similar performances.

Methods:

First a base case was chosen with the same assumption in terms of growth and natural mortality as the previous assessments: the estimated growth by Fonteneau (2008) and the natural mortality estimated for IOTC YFT (Langley 2015). The steepness in the base case was assumed 0.8 and tagging data were not downweighted. This model was part of the grid in the assessment of 2018 (Fu et al. 2018) and based on this model, a second model was developed (M2) by transforming the base case into a model with seasons within years. M2 model has similar settings and assumptions to the base case and below are listed the changes done in the input files for this transformation:

Data:

-The year and seasons of catch, CPUE, length composition, and tagging data were modified.

-The age of tagging data was modified (unit year).

-The environmental data were removed (no need for the estimation of movement between seasons).

Control:

-Growth parameters were modified; *K* was estimated for each age (unit year).

-Natural mortality is assumed the average within each age (unit year).

-All selectivity parameters were modified from being based on age to base on length.

- In the M1 model recruitment happens every season (quarter) in area 1 and 4 with recruitment deviates (σ_R) defined quarterly while in the annual model M2, the annual recruitment is distributed between areas and also between seasons.

-The age of the movements was modified to the unit of year. In the M1 model the seasonal and temporal movement is characterized with oceanographic indices while in the M2 model seasonal movement is estimated within year without the need of any environmental variables.

- F ball park in the base case was assumed 0.1 in 220 (in 2001 -season 1) and in this model was modified to 0.4 (the sum of the year).

On the other hand, another model was developed (M3) with the aim to analyse every single component of the original model and in the cases where a clear explanation was missing in previous reports, then the following principles were used to change the model settings: Aim for model simplification, use SS manual and/or r4ss suggestions and performing of additional analysis. Thus, M3 model was not meant to mimic the original model but it was an attempt to produce a moderately different model configuration (based on the principles listed above) to evaluate its goodness of fit. Below is the list of the settings modified in the M3 model in comparison to the base case and table 1 shows the main differences between the three models.

- Recruitment distribution method was modified from 2 to 3. This is the optimum option for the settings of this model and only each settle entity will get a portion of the total recruitment coming from each spawning.

- All the parameters were estimated without priors, due to the lack of knowledge to choose the values of the priors.

-The deviance of recruitment between quarters (σ_R) in the base case was 0.6 but in M3 was changed to 0.3. Based on the suggestions of Dale et al. (2019) 0.6 corresponds to an annual deviance on recruitments of 0.3 (the mean of the four events).

- F ball park option deactivated due to the lack of information on this setting.

-The 2 time blocks defined for gillnets region 1a (fishery 1) and one for handlines in region 1 a (fishery 2) are not considered in model M3.

-Bias correction ramp (Methot and Taylor, 2011) in the base case was not activated but in the M3 Model the bias correction ramp was activated following the suggestions of r4ss (Table 2).

Results and Discussion:

The M2 model is quite similar to the M1 model, but due to the change on the structure of the model, some of the dynamic of the stock mainly in recruitment and movement are not simulated in the same way. The M1 model recruitment happens every season (quarter) in area 1 and 4 in a continuous way with recruitment deviates (σ_R) defined quarterly. But in the annual model, the annual recruitment is distributed between areas and also between seasons, so with a total of 8 partitions for each annual recruitment. In the case of movement in the M1 model the seasonal and temporal movement is characterized with oceanographic indices while in the annual model seasonal movement is estimated within year without the need of any environmental variables. So even if there are these differences between the three models, the general score is very similar but the best model is M1 with little difference compared to M3 (Table 2, Figure 2,3 and 4). The three models have the retrospective and forecast pattern within the acceptable range for long live species. The model that better fits the CPUE is the M3 model while the model that better predict the CPUE are M1 and M2.

Procedure on how to select the models to be included in the final grid for advice

Below we described the steps we followed to select the models to be included in the final grid for advice. The final grid will include only models that show good performance in terms of model diagnostic.

1-The different hypothesis are translated into different model configurations.

2-The models are run and checked for convergence issues.

3- The diagnostics of all the models were performed with the ss3diag library: the fits of the CPUEs and length composition of each fishery are tested with runs test, the predicting power of each CPUE for each season was analysed using hindcasting and retrospective pattern is measured with Mohn's rho value for SSB and predicting power for SSB is checked with forecast Mohn's rho value.

4-The performance of each model in term of diagnostics is measured as the percentage of pass of the tests.

5-The performance in terms of diagnostic of each model is summarized as a weighted mean of the success of the model in terms of diagnostics where the three components; runs test, hindcasting and retrospective pattern have the same weight.

6-The ranking of the models is done based on the score on the diagnostics of each model.

Hypotheses

Below we describe the hypotheses considered in a total of 48 scenarios, defined into 3 different levels (Figure 5). Between brackets at the end of the sentence is the short name of each type of scenario.

- Level 1: Is the stock fully mixed between the 4 areas?
  - Yes: one area model (1A)
  - No: 2 stocks (East & West without mixing between them) (2A) (More details in Urtizberea et al. (2019)).
  - Not completely mixed, and based on tagging data analysis 4 area are defined the western-tropical, western-temperate, eastern-tropical and western tropical (4A) (Figure 6)

- Level 2A (only 4area model): 2 hypotheses about tagging data, based on if tagging data are mixed or not completely:
  - Full weight on tagging data (lbda01)
  - Downweighted by 0.1 (lbda1)
- Level 2B: Combination of 2 hypotheses of growth and 2 hypothesis of mortality (Figure 7) :
  - Growth estimated by Fonteneau (Fonteneau 2008) (GF)
  - Growth estimated by Dortel; model 3 (Dortel et al. 2015) (GD)
  - Natural Mortality assumed in the assessments of YFT IOTC until 2018 (MB)
  - Natural Mortality scaled based on Atlantic YFT assuming Ma=2 =0.35 (ML)
- Level 3: 3 different hypotheses of steepness:0.7,0.8,0.9 (h)

Convergency and likelihood

The settings of the models are similar to the settings on the assessment model 2018, but the models were transformed to the last version of ss v3.30 and some changes were done to improve the model (Table 4) and to get that the 48 models do not have important boundary problems. Table 4 shows that all the scenarios converge and table 5 show the total likelihood, and the likelihood of each component. The lowest likelihoods within the models with tagging data is when tagging data are down weighted or not considered; the lowest LL is achieved in the models where growth is modelled with the GD and ML natural mortality. In the case of steepness, the results are not clear cut, and the three steepness can have the lowest LL depending on the number of areas in the model or if tagging data are considered.

Diagnostics and discussion

Table 7 shows the results of the diagnostics and in table 8 the weighted mean of the performance of each model in terms of diagnostics is presented. The models with 1 area have the lowest score in terms of diagnostics and therefore, in order to understand better the performance of the other variables the one area model was not consider in the rest of the plots (Figure 8 and Figure 9).

The models with LM, GF, steepness of 0.8 and tagging data downweighted are the models with the best diagnostics in general (Figure 8). From a total of 48 models, 18 models have a score between 76% and 70%. Between the best 18 models all the options of growth, natural mortality, steepness and tagging data are listed but the most common are the ML natural mortality (13 models), tagging data downweighted (9 models), the growth GF (10) and steepness of 0.8 and 0.7 (7 models each).

Figure 9 show some patterns in diagnostics:

-GD has a better performance with LM

- When the tagging data are considered then the models with ML and the GF perform better

-When the tagging data are not considered then the models with GD give better diagnostics.

A threshold of 70% on the model performance was choose to select models to be included in the final grid used for advice. This implies that a total of 18 models were selected and 30 models were excluded.

References

Carvalho, F., Punt, A.E., Chang, Y.-J., Maunder, M.N., and Piner, K.R. 2017. Can diagnostic tests help identify model misspecification in integrated stock assessments? Fisheries Research 192: 28-40.

Dortel E., Sardenne F., Bousquet N., Rivote E., Million J., Le Croizierc G., Chassot G. (2015). An integrated Bayesian modeling approach for the growth of IndianOcean yellowfin tuna. Fish. Res. (2014), Fisheries Research 163:69–84.

Fonteneau, A. 2008. A working proposal for a Yellowfin growth curve to be used during the 2008 yellowfin stock assessment. IOTC-2008-WPTT-4.

Fu D., Langley A. , Merino G. Urtizberea A. (2018). Preliminary Indian Ocean Yellowfin tuna stock assessment 1950-2017 (Stock Synthesis). IOTC-2018-WPTT20-33.

Fu D., Langley A. , Merino G. Urtizberea (2018b). Indian ocean yellowfin tuna ss3 model projections (2018b). IOTC-2018-WPTT20-33. IOTC-2018-SC21-16

Hyndman, R.J. and Athanasopoulos, G. 2013. Forecasting: principles and practice, an online text book. Retrieved September 16, 2012, from http://otexts.com/fpp/ .

Hurtado-Ferro, F., Szuwalski, C. S., Valero, J. L., Anderson, S. C., Cunningham, C. J., Johnson, K. F., & Ono, K. (2014). Looking in the rear-view mirror: bias and retrospective patterns in integrated, age-structured stock assessment models. ICES Journal of Marine Science, 72(1), 99-110.

Kell, L.T., Kimoto, A. and Kitakado, T., 2016. Evaluation of the prediction skill of stock assessment using hindcasting. Fisheries research, 183, pp.119-127.

Langley, A. 2015. Stock assessment of yellowfin tuna in the Indian Ocean using stock synthesis. IOTC–2015–WPTT17–30

Maunder, M.N., and Piner, K.R. 2014. Contemporary fisheries stock assessment: many issues still remain. ICES Journal of Marine Science 72(1): 7-18.

Methot, R.D. and Wetzel C.R. 2013. Stock synthesis: A biological and statistical framework for fish stock assessment and fishery management, Fisheries Research 142: 86-99.

Mohn, R. 1999. The retrospective problem in sequential population analysis: An investigation using cod fishery and simulated data. ICES Journal of Marine Science: Journal du Conseil 56(4): 473-488.

Urtizberea A, Fu D., Merino G., Methot R., Cardinale M., Winker H., Walter J., Murua H.. preliminary assessment of indian ocean yellowfin tuna 1950-2018 (Stock Synthesis, v3.30). IOTC-2018-WPTT21-50.

Thorson, J. T. In press. Predicting recruitment density dependence and intrinsic growth rate for all fishes worldwide using a data-integrated life-history model. Fish and Fisheries.

Thorson, J. T., S. B. Munch, J. M. Cope, and J. Gao. 2017. Predicting life history parameters for all fishes worldwide. Ecological Applications. 27(8): 2262–2276.

Winker, H., Carvalho, F., and Kapur, M., 2018. JABBA: Just Another Bayesian Biomass Assessment. Fisheries Research, Volume 204, August 2018, Pages 275-288.

Winker H., Carvalho F., Cardinale M. and Kell L. (2020). ss3diags: What the Package Does (One Line, Title Case). R package version 1.0.2.

Tables:

Table 1: The table show the variables analyzed under M3 model as sensitivity analysis and compared with the settings in the base case (M1) and M2 model. The variables that contain "x" in the column have that variable activated in the model settings.

Model	Description	seasons within years	F ballpark	Advanced option	time blocks	prior	sigmaR
M1	Base case	-	x		x	x	0.6
M2 (similar base case)	Similar to the base case	x	x		x	x	0.3
M3	Simple versión of M2	x	-	x	-		0.3

Table 2: The settings of the bias correction ramp in the M3 model.

1 # (0/1) to read 13 advanced options

0 #_recdev_early_start (0=none; neg value makes relative to recdev_start)

4 #_recdev_early_phase

0 #_forecast_recruitment phase (incl. late recr) (0 value resets to maxphase+1)

1 #_lambda for Fcast_recr_like occurring before endyr+1

1886 #_last_yr_nobias_adj_in_MPD; begin of ramp

1984 #_first_yr_fullbias_adj_in_MPD; begin of plateau

2016 #_last_yr_fullbias_adj_in_MPD

2016 #_end_yr_for_ramp_in_MPD (can be in forecast to shape ramp, but SS sets bias_adj to 0.0 for fcast yrs)

0.92 #_max_bias_adj_in_MPD (-1 to override ramp and set biasadj=1.0 for all estimated recdevs)

0 #_period of cycles in recruitment (N parms read below)

-5 #min rec_dev

5 #max rec_dev

0 #_read_recdevs

Table 3: The summary of the diagnostics for each of the model: The number of CPUE and fishery length composition that passed the run test, the number of season where the MASE of hindcasting of the CPUE in each region was lower or equal to 1, the mohn's rho value of the retrospective pattern and forecast. Table b) shows the success of the model in terms of percentage for each of the variable and the weighted mean.

a)

Model	Run	CPUE pass	NFishery LC pass	hind. Area1	hind. Area2	hind. Area3	hind. Area4	MohnR retro	MohnR forecast
M1	R14-Grid	2 & 4	7	1	2	3	2	-0.05	-0.04
M2	Annual	1	9	0	1	4	2	-0.09	-0.13
M3	Annual Simple	1,2, 3 & 4	9	0	1	1	2	-0.09	-0.12

b)

Model	Run	%CPUE pass	%NFishery LC pass	%hind. Area1	%hind. Area2	%hind. Area3	%hind. Area4	MohnR retro	MohnR forecast	W.Mean
M1	R14-Grid	50	27.27	25	50	75	50	100	100	63.63
M2	Annual	25	36.36	0	25	100	50	100	100	58.14
M3	Annual Simple	100	36.36	0	0	25	50	100	100	62.31

Table 4: The modifications done to the reference case of 2018 (Fu et al. 2018) and applied to all the models in the grid.

1. Convert 2018 reference model from SS 3.24 to 3.30 (changes in boundaries movement rates, and survey catchability becomes a parameter.	done, get some differences, probably due to some local minimum in movement rates.
2. The length composition bins modified from 2 to 4 cm	done
3. Add a constant of 0.01 to baitboat and handlines due to the patchy distribution	Residual pattern is improved.
4. Changes in purse seiners selectivity from based on age to based on length.	A little increase in likelihood.
6. Update catches (data 2019)	Done and the catches of European Purse seiners were revised in the WPTT21 meeting)
7. Update length compositions (data 2019)	done
8. Change the longline joint index. 4 area model: Scaled index with estimates of 2019 for each area. 2 area model: in the model region 1, the scaled index (2019) region 1b +region 2 (CPUE regions definition) and in the model region 2 (2019), the scaled index region 3 + region 4. 1 area model: the scaled index (2019) region 1b +region 2 (CPUE regions definition) +region 3 + region 4.	done
9. Remove all the length composition from 272 (2015-2018).	The same approach as in the assessment of 2018, because results very sensitive to the new length compositions.
10. Wider boundaries on the movement desviations	Some runs were touching boundaries of fleet 3 parameters R4/R8/R12/R20/R27/R28/R36 low boundary fleet 3 Some runs still touch the boundaries of the variability of movement, but they are very wide -50,50 so not big impact in the results.
11. Wider boundaries in fleet 3	R23 problems to converge and R20 ln(R0) parameter high gradients. So we kept both of them keeping the previous boundaries of fleet3
12. Fballpark	In the v3.24 of SS the reduction of lambda of Fballpark was done automatically, and in the v3.30 is done manually. Maxlambdaphase is set as 4 and the reducction of lambda to 0.01 in the last phase.

Table 5: The convergency table with the details of each of the scenarios.

Model	Run	LL	grad	hessian	ssb0	time	nparam	AIC
2A_lbda0_MB_GF_h08	R1	2988.34	0.00098718	yes	3316130	3 h, 6 min, 51 s.	403	790
2A_lbda0_ML_GF_h08	R2	3040.96	0.00085926	yes	4857900	3 h, 16 min, 23 s.	403	789.96
2A_lbda0_MB_GD_h08	R3	2933.14	0.00098541	yes	3394940	3 h, 2 min, 45 s.	403	790.03
2A_lbda0_ML_GD_h08	R4	2910.82	0.00070845	yes	5104590	2 h, 59 min, 25 s.	403	790.05
2A_lbda0_MB_GF_h07	R5	2972.13	0.00083675	yes	3574860	3 h, 16 min, 50 s.	403	790.01
2A_lbda0_ML_GF_h07	R6	3050.53	0.00060296	yes	5327640	3 h, 14 min, 2 s.	403	789.95
2A_lbda0_MB_GD_h07	R7	2948.31	0.00083867	yes	3675610	3 h, 10 min, 21 s.	403	790.02
2A_lbda0_ML_GD_h07	R8	2911.19	0.00037086	yes	5620170	3 h, 15 min, 29 s.	403	790.05
2A_lbda0_MB_GF_h09	R9	2988.44	0.00088404	yes	3136870	3 h, 2 min, 1 s.	403	789.99
2A_lbda0_ML_GF_h09	R10	3040.76	0.00093308	yes	4521090	3 h, 4 min, 15 s.	403	789.96
2A_lbda0_MB_GD_h09	R11	2954.24	0.00099559	yes	3221840	3 h, 14 min, 22 s.	403	790.02
2A_lbda0_ML_GD_h09	R12	2910.81	0.00091887	yes	4732950	3 h, 9 min, 42 s.	403	790.05
4A_lbda01_MB_GF_h08	R13	3611.8	0.00086119	yes	3231800	4 h, 11 min, 48 s.	452	887.62
4A_lbda1_MB_GF_h08	R14	8796.33	0.00098378	yes	2899270	4 h, 11 min, 48 s.	452	885.84
4A_lbda01_ML_GF_h08	R15	3666.7	0.00099127	yes	4877700	4 h, 3 min, 53 s.	452	887.59
4A_lbda1_ML_GF_h08	R16	8839.61	0.0007711	yes	4307390	4 h, 11 min, 59 s.	452	885.83
4A_lbda01_MB_GD_h08	R17	3574.26	0.00074881	yes	3258590	4 h, 12 min, 52 s.	452	887.64
4A_lbda1_MB_GD_h08	R18	8895.71	0.00040334	yes	3075190	4 h, 10 min, 47 s.	452	885.81
4A_lbda01_ML_GD_h08	R19	3550.51	0.00084514	yes	4659830	4 h, 3 min, 47 s.	452	887.65
4A_lbda1_ML_GD_h08	R20	8779.24	0.00068105	yes	4375530	4 h, 14 min, 56 s.	452	885.84
4A_lbda01_MB_GF_h07	R21	3604.73	0.00057287	yes	3470830	4 h, 3 min, 31 s.	452	887.62
4A_lbda1_MB_GF_h07	R22	8837.74	0.00075749	yes	3139580	4 h, 0 min, 8 s.	452	885.83
4A_lbda01_ML_GF_h07	R23	3677.49	0.0009444	yes	5401700	4 h, 17 min, 42 s.	452	887.58
4A_lbda1_ML_GF_h07	R24	8856.21	0.00025532	yes	4655820	4 h, 15 min, 11 s.	452	885.82
4A_lbda01_MB_GD_h07	R25	3576.04	0.00098698	yes	3546760	4 h, 7 min, 58 s.	452	887.64
4A_lbda1_MB_GD_h07	R26	8845.37	0.00078845	yes	3191410	4 h, 9 min, 38 s.	452	885.82

4A_lbda01_ML_GD_h07	R27	3536.12	0.00073878	yes	5226270	4 h, 7 min, 50 s.	452	887.66
4A_lbda1_ML_GD_h07	R28	8774.13	0.00056267	yes	4690730	4 h, 3 min, 28 s.	452	885.84
4A_lbda01_MB_GF_h09	R29	3600.51	0.00091144	yes	3080210	4 h, 21 min, 52 s.	452	887.62
4A_lbda1_MB_GF_h09	R30	8798.95	0.00078438	yes	2823980	4 h, 7 min, 29 s.	452	885.84
4A_lbda01_ML_GF_h09	R31	3672.33	0.0007101	yes	4595930	4 h, 4 min, 37 s.	452	887.58
4A_lbda1_ML_GF_h09	R32	8841.42	0.00060289	yes	4132220	4 h, 21 min, 26 s.	452	885.83
4A_lbda01_MB_GD_h09	R33	3589.85	0.0009343	yes	2863790	4 h, 0 min, 47 s.	452	887.63
4A_lbda1_MB_GD_h09	R34	8842.54	0.00075477	yes	2899880	4 h, 0 min, 1 s.	452	885.83
4A_lbda01_ML_GD_h09	R35	3537.45	0.0009345	yes	4545130	4 h, 11 min, 13 s.	452	887.66
4A_lbda1_ML_GD_h09	R36	8780.47	0.00075422	yes	4047860	4 h, 13 min, 58 s.	452	885.84
1A_lbda0_MB_GF_h08	R37	3035.97	0.00083218	yes	2964440	2 h, 42 min, 1 s.	403	789.96
1A_lbda0_ML_GF_h08	R38	3111.39	0.00095869	yes	3552790	2 h, 42 min, 37 s.	403	789.91
1A_lbda0_MB_GD_h08	R39	3111.39	0.00095869	yes	3552790	2 h, 36 min, 28 s.	403	789.91
1A_lbda0_ML_GD_h08	R40	3050.68	0.00078955	yes	5110860	2 h, 37 min, 6 s.	403	789.95
1A_lbda0_MB_GF_h07	R41	3119.85	0.00081793	yes	3573460	2 h, 44 min, 0 s.	403	789.91
1A_lbda0_ML_GF_h07	R42	3168.79	0.00646232	yes	5130070	2 h, 43 min, 42 s.	403	789.88
1A_lbda0_MB_GD_h07	R43	3097.13	0.00077667	yes	3824410	2 h, 39 min, 1 s.	403	789.92
1A_lbda0_ML_GD_h07	R44	3063.77	0.00094958	yes	5615900	2 h, 51 min, 57 s.	403	789.95
1A_lbda0_MB_GF_h09	R45	3144.7	0.00029017	yes	3214160	2 h, 38 min, 2 s.	403	789.89
1A_lbda0_ML_GF_h09	R46	3167.47	0.00084535	yes	4413610	2 h, 37 min, 48 s.	403	789.88
1A_lbda0_MB_GD_h09	R47	3114.49	0.00089779	yes	3376890	2 h, 41 min, 54 s.	403	789.91
1A_lbda0_ML_GD_h09	R48	3062.68	0.00094014	yes	4783630	2 h, 37 min, 25 s.	403	789.95

Table 6: The total likelihood and the likelihood of some of the components.

Model	Run	TOTAL	Survey	Length_comp	Tag_comp	Tag_negbin	Recruitment	Parm_priors	Parm_softbounds
2A_lbda0_MB_GF_h08	1	2988.34	-332.136	3347.46	0	0	-58.7357	13.6484	0.00615071
2A_lbda0_ML_GF_h08	2	3040.96	-328.205	3378.09	0	0	-51.4964	25.0095	0.00650245
2A_lbda0_MB_GD_h08	3	2933.14	-347.905	3282.03	0	0	-51.8276	31.8728	0.00628175
2A_lbda0_ML_GD_h08	4	2910.82	-339.383	3241.79	0	0	-46.8132	34.6242	0.0156307
2A_lbda0_MB_GF_h07	5	2972.13	-333.127	3324.67	0	0	-58.8008	22.877	0.00569367
2A_lbda0_ML_GF_h07	6	3050.53	-328.376	3388.26	0	0	-49.8693	22.8799	0.0067681
2A_lbda0_MB_GD_h07	7	2948.31	-347.175	3305.68	0	0	-50.7044	20.8607	0.00654632
2A_lbda0_ML_GD_h07	8	2911.19	-339.425	3241.11	0	0	-45.6119	34.8039	0.0156407
2A_lbda0_MB_GF_h09	9	2988.44	-331.067	3346.25	0	0	-59.2226	14.2288	0.0061185
2A_lbda0_ML_GF_h09	10	3040.76	-327.945	3379.28	0	0	-52.497	24.2112	0.00664491
2A_lbda0_MB_GD_h09	11	2954.24	-344.474	3307.07	0	0	-52.0207	24.0338	0.00638941
2A_lbda0_ML_GD_h09	12	2910.81	-340.306	3242.17	0	0	-47.9464	37.164	0.0156325
4A_lbda01_MB_GF_h08	13	3611.8	-338.33	3372.32	415.813	173.011	-57.2317	20.157	0.00613544
4A_lbda1_MB_GF_h08	14	8796.33	-327.092	3400.08	4021.88	1693.86	-53.7416	36.2993	0.00585092
4A_lbda01_ML_GF_h08	15	3666.7	-332.866	3390.64	417.764	171.67	-47.786	41.1413	0.00664488
4A_lbda1_ML_GF_h08	16	8839.61	-323.883	3424.89	4040.17	1659.81	-41.0456	54.7428	0.00604263
4A_lbda01_MB_GD_h08	17	3574.26	-348.073	3291.77	430.269	176.707	-50.8271	47.8546	0.00626106
4A_lbda1_MB_GD_h08	18	8895.71	-309.133	3412.38	4047.38	1720.54	-48.6345	48.8003	0.00637915
4A_lbda01_ML_GD_h08	19	3550.51	-346.145	3258.55	426.591	176.258	-39.1052	47.0067	0.0154998
4A_lbda1_ML_GD_h08	20	8779.24	-342.643	3324.18	4051.27	1687.37	-35.1099	65.1522	0.0151687
4A_lbda01_MB_GF_h07	21	3604.73	-342.005	3365.56	415.465	173.605	-56.4777	22.9755	0.00575061
4A_lbda1_MB_GF_h07	22	8837.74	-327.142	3417.3	4033.89	1698.11	-52.0529	39.6557	0.00645793
4A_lbda01_ML_GF_h07	23	3677.49	-320.984	3379.58	418.495	173.529	-44.7771	46.5878	0.00729905
4A_lbda1_ML_GF_h07	24	8856.21	-319.651	3427.79	4037.32	1668.37	-34.4579	52.6362	0.00659044

4A_lbda01_MB_GD_h07	25	3576.04	-354.627	3298.77	427.663	177.659	-49.2078	49.8948	0.00612867
4A_lbda1_MB_GD_h07	26	8845.37	-319.012	3361.38	4042.24	1717.67	-46.5799	65.8361	0.00627699
4A_lbda01_ML_GD_h07	27	3536.12	-378.937	3274.43	429.694	175.566	-35.8595	48.381	0.00764857
4A_lbda1_ML_GD_h07	28	8774.13	-344.751	3305.23	4045.84	1694.08	-30.0632	76.0152	0.0155536
4A_lbda01_MB_GF_h09	29	3600.51	-340.278	3361.53	416.105	172.957	-59.6375	28.0541	0.00553098
4A_lbda1_MB_GF_h09	30	8798.95	-328.883	3399.92	4025.77	1694.82	-55.6701	37.632	0.00549934
4A_lbda01_ML_GF_h09	31	3672.33	-332.721	3401.98	418.133	171.259	-50.6394	36.1804	0.00667948
4A_lbda1_ML_GF_h09	32	8841.42	-330.5	3435.94	4038.32	1659.6	-41.4857	54.4467	0.00657416
4A_lbda01_MB_GD_h09	33	3589.85	-384.987	3348.65	430.204	176.48	-48.9214	42.4279	0.00692482
4A_lbda1_MB_GD_h09	34	8842.54	-324.946	3340.33	4047.42	1718.61	-48.9181	81.9574	0.0069824
4A_lbda01_ML_GD_h09	35	3537.45	-348.453	3244.92	426.434	174.908	-42.1735	54.1006	0.0153837
4A_lbda1_ML_GD_h09	36	8780.47	-341.607	3333.72	4048.36	1682.79	-35.4241	65.6085	0.0153945
1A_lbda0_MB_GF_h08	37	3035.97	-178.717	3260.1	0	0	-58.2679	12.8378	0.00616512
1A_lbda0_ML_GF_h08	38	3111.39	-205.478	3344.22	0	0	-48.6803	21.3163	0.00663383
1A_lbda0_MB_GD_h08	39	3111.39	-205.478	3344.22	0	0	-48.6803	21.3163	0.00663383
1A_lbda0_ML_GD_h08	40	3050.68	-202.407	3264.24	0	0	-48.6596	37.4901	0.0151949
1A_lbda0_MB_GF_h07	41	3119.85	-199.995	3363.69	0	0	-57.084	13.2263	0.00623729
1A_lbda0_ML_GF_h07	42	3168.79	-198.176	3393.74	0	0	-51.0172	24.2352	0.00626446
1A_lbda0_MB_GD_h07	43	3097.13	-206.524	3329.87	0	0	-48.9996	22.7697	0.00604133
1A_lbda0_ML_GD_h07	44	3063.77	-202.646	3283.42	0	0	-47.6921	30.6713	0.0153805
1A_lbda0_MB_GF_h09	45	3144.7	-198.281	3385.57	0	0	-57.3394	14.7292	0.00687155
1A_lbda0_ML_GF_h09	46	3167.47	-198.003	3396.94	0	0	-53.5667	22.0864	0.00650414
1A_lbda0_MB_GD_h09	47	3114.49	-204.54	3347.62	0	0	-48.4381	19.8381	0.00677042
1A_lbda0_ML_GD_h09	48	3062.68	-202.007	3283.52	0	0	-49.1338	30.2766	0.0153655

Table 7: Diagnostics on each model the CPUE that pass the Runs Test (CPUE pass), the number of fishery length composition that pass the Runs Test (NFishery LC pass), the number of season that pass the hindcasting by area (hind. Area), Mohn's Rho value of the retrospective patter (MohnR retro), Mohn's Rho value of the forecast pattern (MohnR forecast).

Model	Run	CPUE pass	NFishery LC pass	hind. Area1	hind. Area2	hind. Area3	hind. Area4	MohnR retro	MohnR forecast
2A_lbda0_MB_GF_h08	1	1 & 2	9	1			2	-0.09	-0.12
2A_lbda0_ML_GF_h08	2	1 & 2	10	1			1	-0.09	-0.17
2A_lbda0_MB_GD_h08	3	1 & 2	9	1			1	-0.06	-0.07
2A_lbda0_ML_GD_h08	4	1 & 2	11	1			3	-0.08	-0.12
2A_lbda0_MB_GF_h07	5	1 & 2	10	1			2	-0.09	-0.12
2A_lbda0_ML_GF_h07	6	1 & 2	10	1			1	-0.1	-0.19
2A_lbda0_MB_GD_h07	7	1 & 2	11	1			1	-0.08	-0.1
2A_lbda0_ML_GD_h07	8	1 & 2	11	1			1	-0.08	-0.13
2A_lbda0_MB_GF_h09	9	1 & 2	9	0			2	-0.08	-0.1
2A_lbda0_ML_GF_h09	10	1 & 2	10	1			2	-0.09	-0.16
2A_lbda0_MB_GD_h09	11	1 & 2	9	1			0	-0.08	-0.11
2A_lbda0_ML_GD_h09	12	1 & 2	11	1			3	-0.07	-0.12
4A_lbda01_MB_GF_h08	13	1,2,3 & 4	6	1	2	3	2	-0.07	-0.03
4A_lbda1_MB_GF_h08	14	2 & 4	7	1	2	3	2	-0.05	-0.04
4A_lbda01_ML_GF_h08	15	1,2 & 3	9	0	3	4	3	-0.07	-0.08
4A_lbda1_ML_GF_h08	16	1,2,3 & 4	8	1	2	4	2	-0.05	-0.07
4A_lbda01_MB_GD_h08	17	1,2,3 & 4	8	0	2	4	1	-0.07	-0.01
4A_lbda1_MB_GD_h08	18	1& 4	9	0	1	4	1	-0.11	-0.05
4A_lbda01_ML_GD_h08	19	1,2 & 4	12	0	3	4	1	0	0.03
4A_lbda1_ML_GD_h08	20	1,2,3 & 4	10	0	1	4	1	-0.02	0
4A_lbda01_MB_GF_h07	21	1,2,3 & 4	6	1	3	3	3	-0.09	-0.07
4A_lbda1_MB_GF_h07	22	1,2 & 4	8	1	2	3	3	-0.09	-0.1
4A_lbda01_ML_GF_h07	23	1,2 & 4	8	0	2	4	3	-0.1	-0.11
4A_lbda1_ML_GF_h07	24	1,2,3 & 4	10	1	2	4	1	-0.06	-0.08
4A_lbda01_MB_GD_h07	25	1,2,3 & 4	9	0	1	2	0	-	-

4A_lbda1_MB_GD_h07	26	1	9	1	1	2	1	-0.04	0
4A_lbda01_ML_GD_h07	27	1,2 & 4	12	0	2	4	2	-0.04	-0.02
4A_lbda1_ML_GD_h07	28	1,2,3 & 4	10	0	1	4	1	-0.08	-0.05
4A_lbda01_MB_GF_h09	29	2,3 &4	6	1	3	3	2	-0.06	0
4A_lbda1_MB_GF_h09	30	2& 4	7	1	2	3	2	-0.08	-0.07
4A_lbda01_ML_GF_h09	31	1,2,3 &4	8	1	2	4	2	-0.1	-0.12
4A_lbda1_ML_GF_h09	32	1,2,3 & 4	8	0	1	4	2	-0.07	-0.08
4A_lbda01_MB_GD_h09	33	1,2, 3 & 4	12	0	1	0	0	0.04	0.16
4A_lbda1_MB_GD_h09	34	1,3 & 4	9	0	1	1	1	-0.11	0
4A_lbda01_ML_GD_h09	35	1,2,3 & 4	10	0	1	4	2	-0.08	-0.06
4A_lbda1_ML_GD_h09	36	1 & 4	11	0	1	4	1	-0.06	-0.03
1A_lbda0_MB_GF_h08	37	1	11	1				-0.1	-0.15
1A_lbda0_ML_GF_h08	38	1	11	0				-0.1	-0.14
1A_lbda0_MB_GD_h08	39	1	11	0				-0.1	-0.14
1A_lbda0_ML_GD_h08	40	1	10	1				-0.09	-0.18
1A_lbda0_MB_GF_h07	41	1	10	1				-0.13	-0.19
1A_lbda0_ML_GF_h07	42	0	10	0				-0.11	-0.21
1A_lbda0_MB_GD_h07	43	1	11	1				-0.1	-0.15
1A_lbda0_ML_GD_h07	44	1	11	1				-0.1	-0.19
1A_lbda0_MB_GF_h09	45	1	10	1				-0.11	-0.17
1A_lbda0_ML_GF_h09	46	1	10	0				-0.12	-0.22
1A_lbda0_MB_GD_h09	47	1	11	0				-0.09	-0.11
1A_lbda0_ML_GD_h09	48	1	11	1				-0.09	-0.19

Table 8: Diagnostics on each model % of CPUE that pass the Runs Test (%NCPUE pass), the percentage of fishery length composition that pass the Runs Test (% NFishery LC), the percentage of season that pass the hindcasting by area (hind. Area), Mohn's Rho value of the retrospective patter (MohnR retro), Mohn's Rho value of the forecast pattern (MohnR Forecast).

Model	Run	%NCPUE	%NFishery LC	%hind.	%hind.	%hind.	%hind.	MohnR	MohnR	Mean
4A_lbda01_MB_GF_h07	21	100	27.27	25	75	75	75	100	100	75.38
2A_lbda0_ML_GD_h08	4	100	50	25	0	0	75	100	100	75
2A_lbda0_ML_GD_h09	12	100	50	25	0	0	75	100	100	75
4A_lbda1_ML_GF_h08	16	100	36.36	25	50	100	50	100	100	74.81
4A_lbda01_ML_GF_h09	31	100	36.36	25	50	100	50	100	100	74.81
4A_lbda1_ML_GF_h07	24	100	45.45	25	50	100	25	100	100	74.24
4A_lbda01_ML_GF_h08	15	75	40.91	0	75	100	75	100	100	73.49
4A_lbda01_ML_GD_h09	35	100	45.45	0	25	100	50	100	100	72.16
4A_lbda01_ML_GD_h08	19	75	54.55	0	75	100	25	100	100	71.59
4A_lbda01_ML_GD_h07	27	75	54.55	0	50	100	50	100	100	71.59
4A_lbda01_MB_GF_h08	13	100	27.27	25	50	75	50	100	100	71.21
4A_lbda01_MB_GD_h08	17	100	36.36	0	50	100	25	100	100	70.64
4A_lbda1_MB_GF_h07	22	75	36.36	25	50	75	75	100	100	70.64
4A_lbda01_ML_GF_h07	23	75	36.36	0	50	100	75	100	100	70.64
4A_lbda1_ML_GF_h09	32	100	36.36	0	25	100	50	100	100	70.64
2A_lbda0_MB_GF_h07	5	100	45.45	25	0	0	50	100	100	70.08
4A_lbda1_ML_GD_h08	20	100	45.45	0	25	100	25	100	100	70.08
4A_lbda1_ML_GD_h07	28	100	45.45	0	25	100	25	100	100	70.08
2A_lbda0_MB_GF_h08	1	100	40.91	25	0	0	50	100	100	69.32
4A_lbda01_MB_GF_h09	29	75	27.27	25	75	75	50	100	100	69.13
2A_lbda0_MB_GD_h07	7	100	50	25	0	0	25	100	100	66.67
2A_lbda0_ML_GD_h07	8	100	50	25	0	0	25	100	100	66.67
1A_lbda0_MB_GF_h08	37	100	50	25	0	0	0	100	100	66.67
1A_lbda0_MB_GD_h07	43	100	50	25	0	0	0	100	100	66.67

2A_lbda0_MB_GD_h08	3	100	40.91	25	0	0	25	100	100	65.15
2A_lbda0_MB_GF_h09	9	100	40.91	0	0	0	50	100	100	65.15
4A_lbda1_MB_GF_h08	14	50	31.82	25	50	75	50	100	100	63.64
4A_lbda1_MB_GF_h09	30	50	31.82	25	50	75	50	100	100	63.64
4A_lbda1_ML_GD_h09	36	50	50	0	25	100	25	100	100	62.5
4A_lbda01_MB_GD_h09	33	100	54.55	0	25	0	0	100	100	61.18
2A_lbda0_MB_GD_h09	11	100	40.91	25	0	0	0	100	100	60.99
4A_lbda1_MB_GD_h08	18	50	40.91	0	25	100	25	100	100	60.99
4A_lbda1_MB_GD_h09	34	75	40.91	0	25	25	25	100	100	58.9
1A_lbda0_ML_GF_h08	38	100	50	0	0	0	0	100	100	58.33
1A_lbda0_MB_GD_h08	39	100	50	0	0	0	0	100	100	58.33
1A_lbda0_MB_GD_h09	47	100	50	0	0	0	0	100	100	58.33
4A_lbda1_MB_GD_h07	26	25	40.91	25	25	50	25	100	100	54.74
2A_lbda0_ML_GF_h09	10	100	45.45	25	0	0	50	100	0	53.41
1A_lbda0_ML_GD_h07	44	100	50	25	0	0	0	100	0	50
1A_lbda0_ML_GD_h09	48	100	50	25	0	0	0	100	0	50
2A_lbda0_ML_GF_h08	2	100	45.45	25	0	0	25	100	0	49.24
2A_lbda0_ML_GF_h07	6	100	45.45	25	0	0	25	100	0	49.24
1A_lbda0_ML_GD_h08	40	100	45.45	25	0	0	0	100	0	49.24
1A_lbda0_MB_GF_h07	41	100	45.45	25	0	0	0	100	0	49.24
1A_lbda0_MB_GF_h09	45	100	45.45	25	0	0	0	100	0	49.24
1A_lbda0_ML_GF_h09	46	100	45.45	0	0	0	0	100	0	40.91
4A_lbda01_MB_GD_h07	25	100	40.91	0	25	50	0	0	0	29.74
1A_lbda0_ML_GF_h07	42	0	45.45	0	0	0	0	100	0	24.24

Figures:

Figure 1: Runs test analysis of each CPUE of the M3 model.

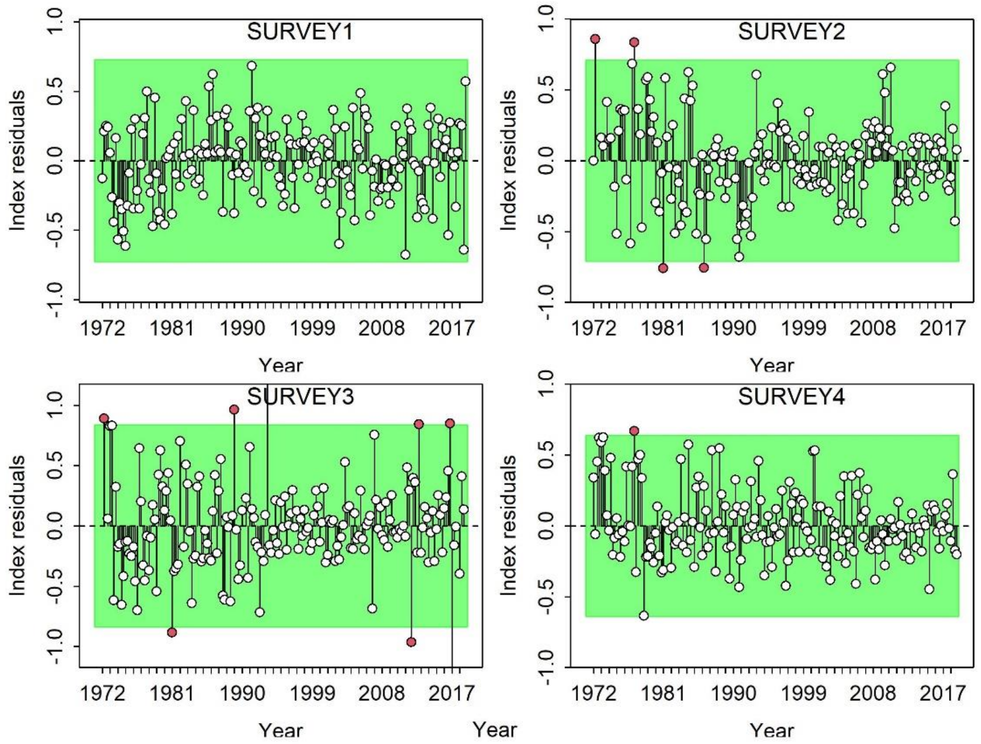


Figure 2: Runs test analysis of the residuals of the length composition of each fishery in the M3 model. In the table A1, is the definition of each fishery.

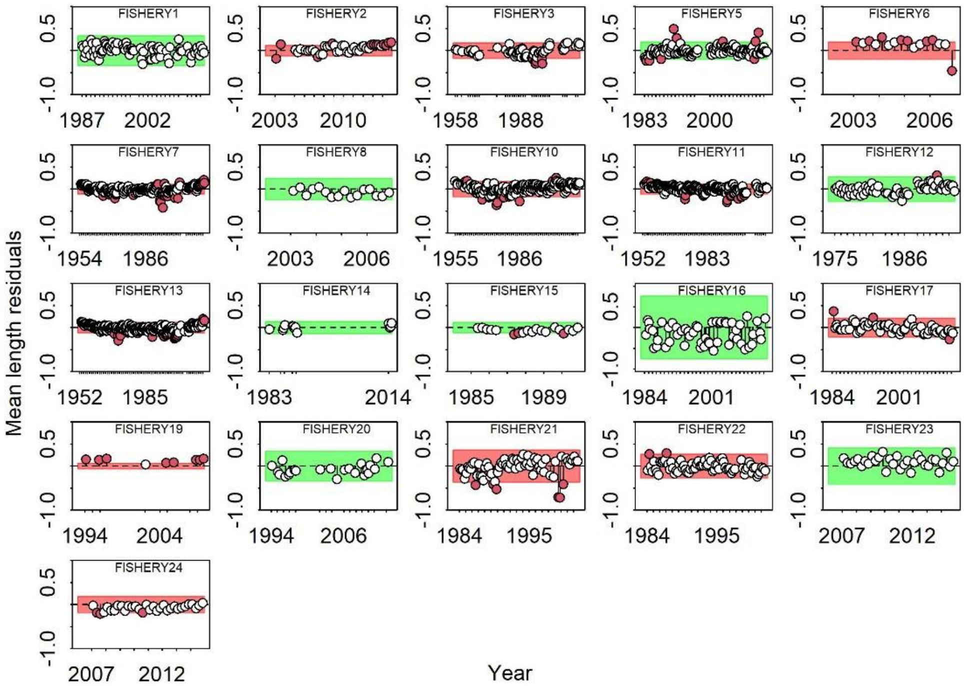
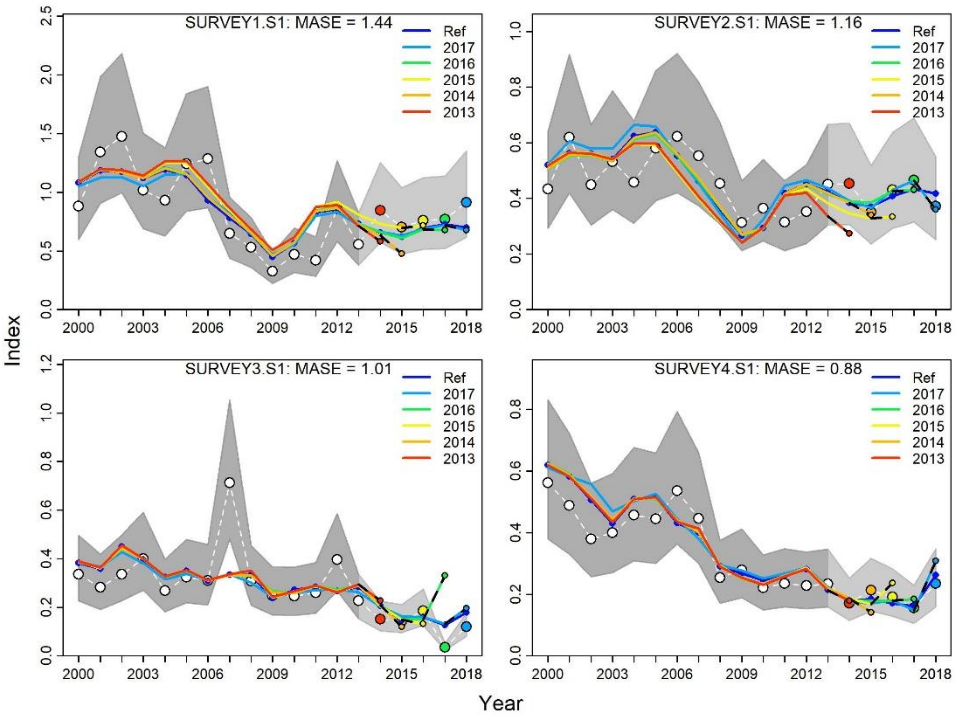
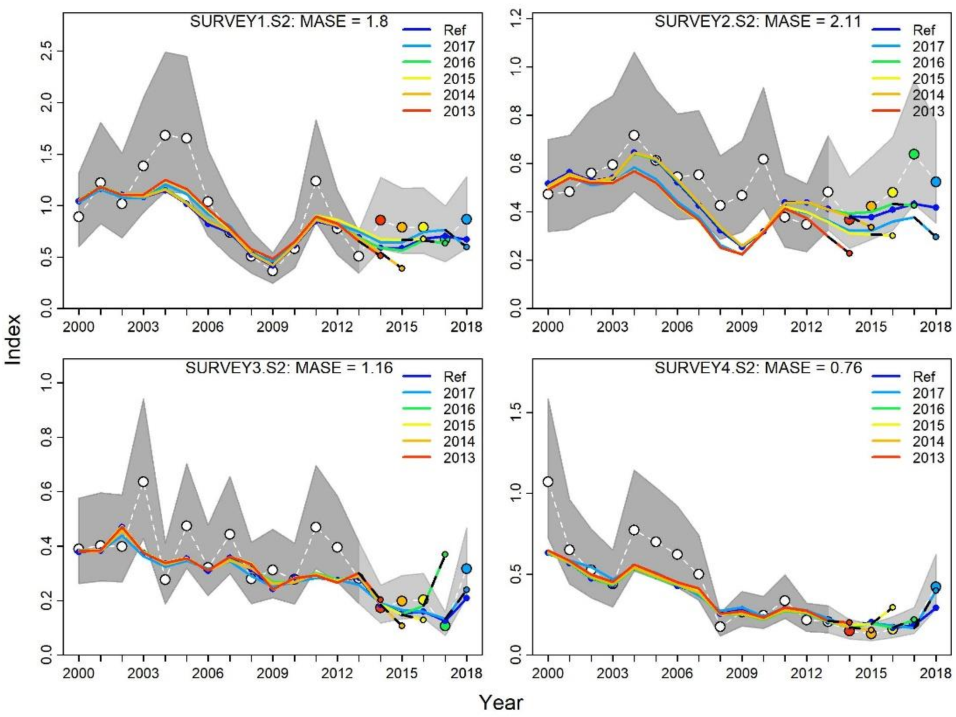


Figure 3: a) Hindcasting analysis of each CPUE in season 1, (b) in season 2, (c) in season 3 and (d) season 4 of M3 model.

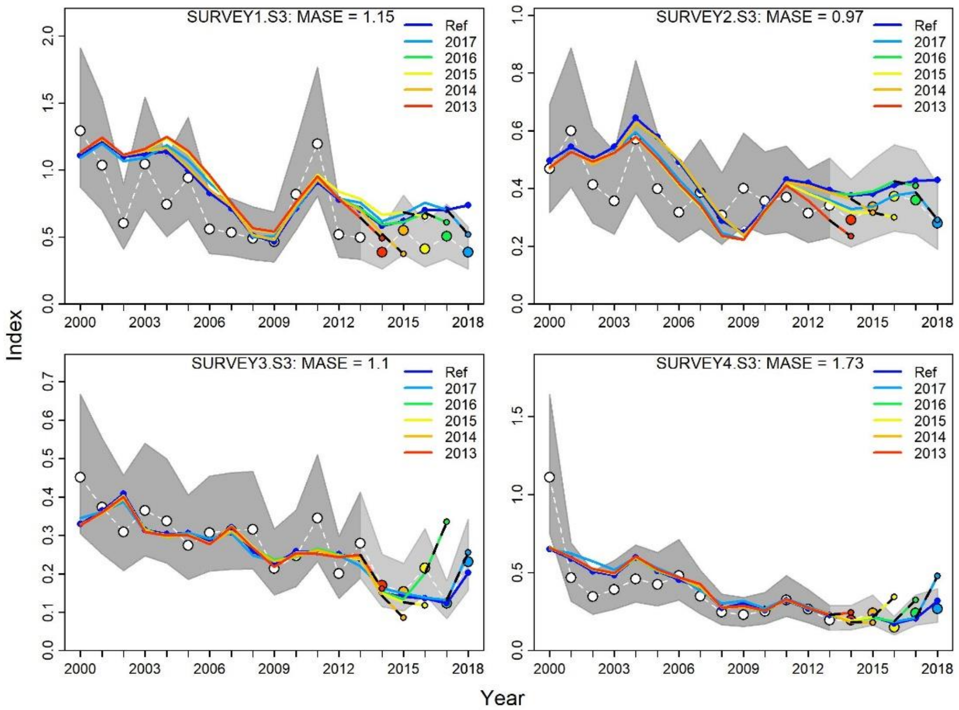
a)



b)



c)


SURVEY1.S3: MASE = 1.15
SURVEY2.S3: MASE = 0.97
SURVEY3.S3: MASE = 1.1
SURVEY4.S3: MASE = 1.73
Ref
2017
2016
2015
2014
2013
Index
Year

d)

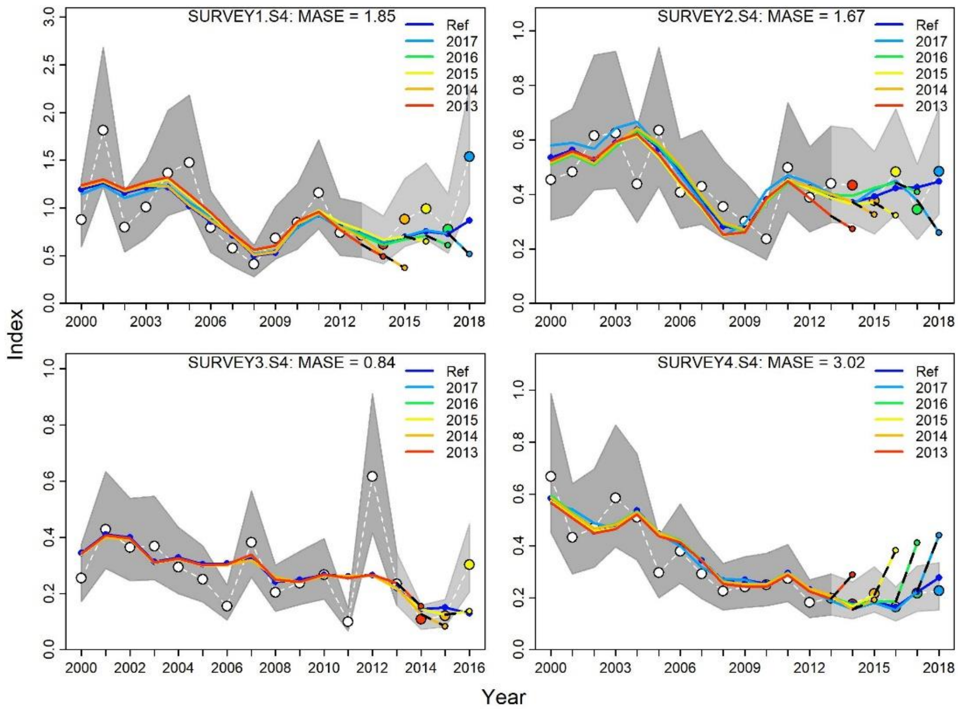

SURVEY1.S4: MASE = 1.85
SURVEY2.S4: MASE = 1.67
SURVEY3.S4: MASE = 0.84
SURVEY4.S4: MASE = 3.02
Ref
2017
2016
2015
2014
2013
Index
Year

Figure 4: The retrospective analysis of M3 model and the Mohn's rho value. The discontinuous line is the forecast of each retro and the value between brackets is the Mohn's rho of the forecast.

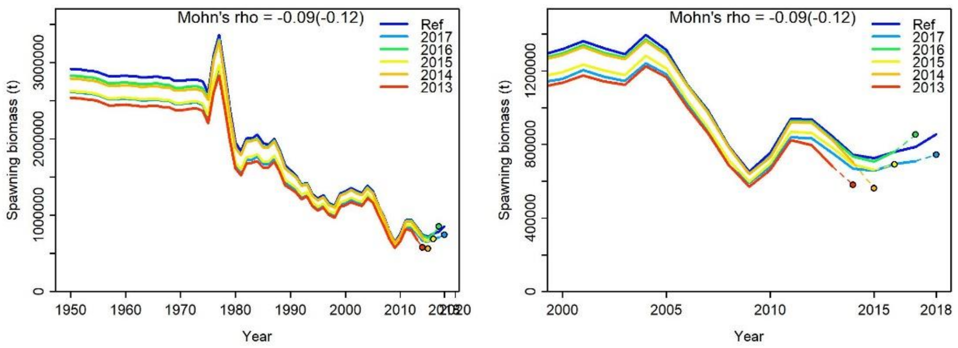


Figure 5: The different hypothesis analyzed in the grid.

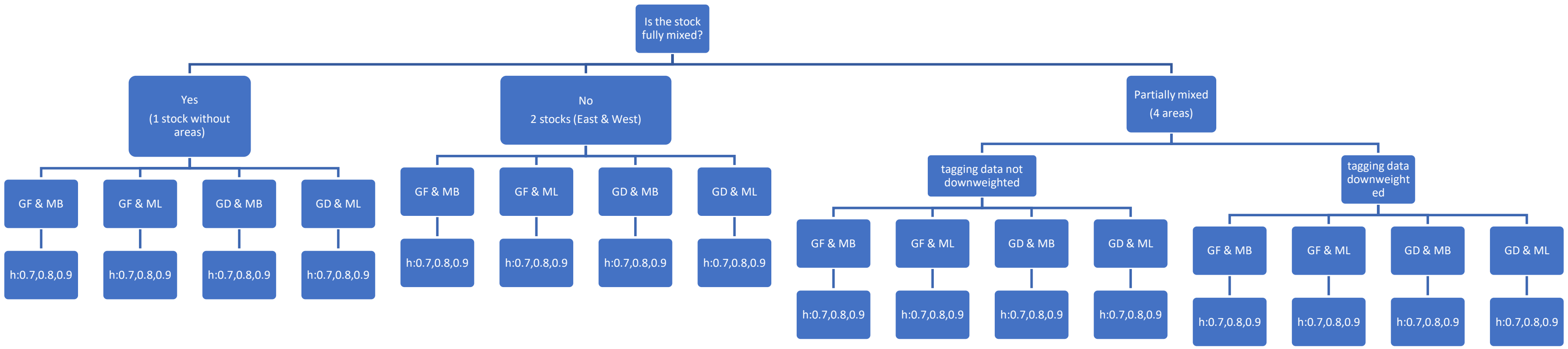


Figure 6: Spatial stratification of the Indian Ocean for the 4-area assessment model. The black arrows represent the configuration of the movement parameterization of the base assessment model.

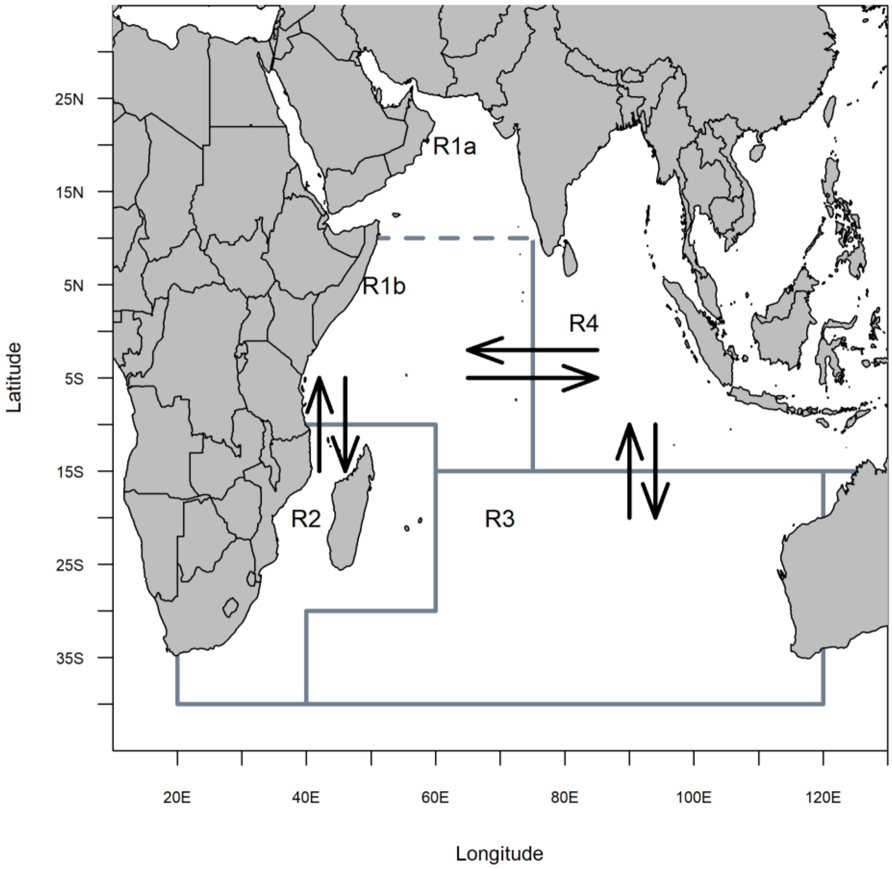
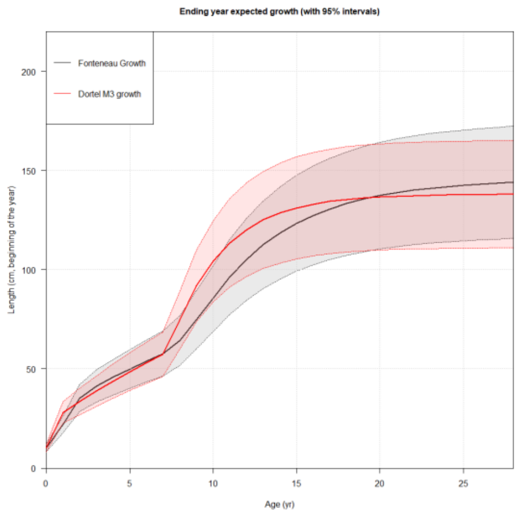


Figure 7: Left) Comparison of the length at age assuming the growth model estimated by Fonteneau (2018) and by Dortel (2015) in the model 3. Right) Comparison of the natural mortality at age used in the assessment of 2018 (MB) and the natural mortality based on the YFT of the Atlantic (ML).



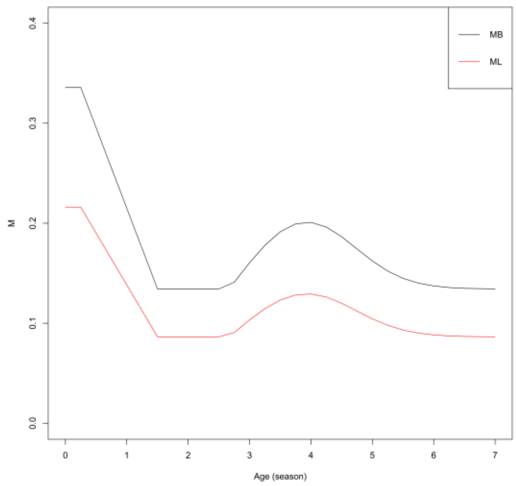


Figure 8: The weighted mean of the success on the three components of diagnostics by area, by natural mortality, growth, steepness and tagging data. The model with only one area is only considered in the first plot of the weighted mean by area.

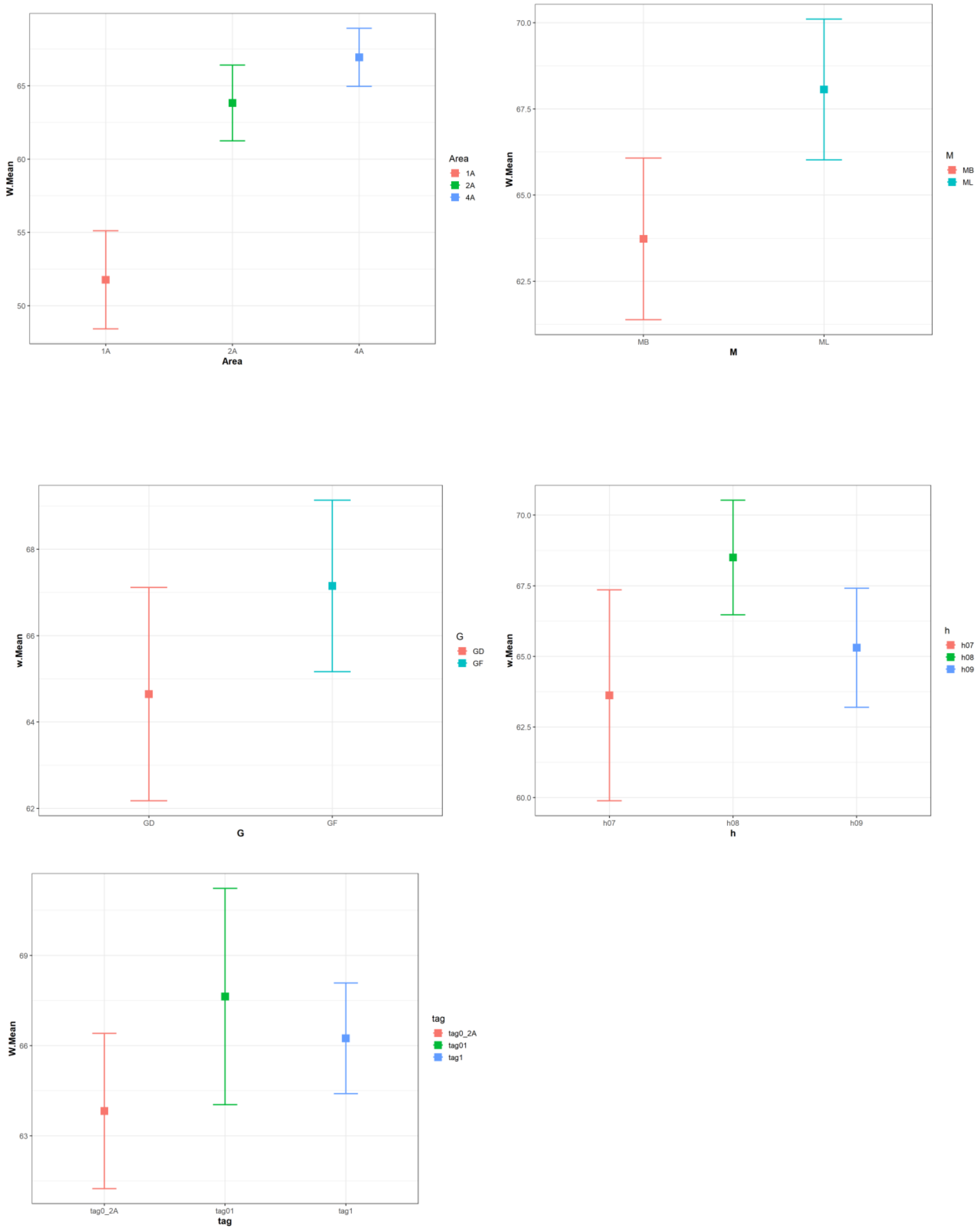
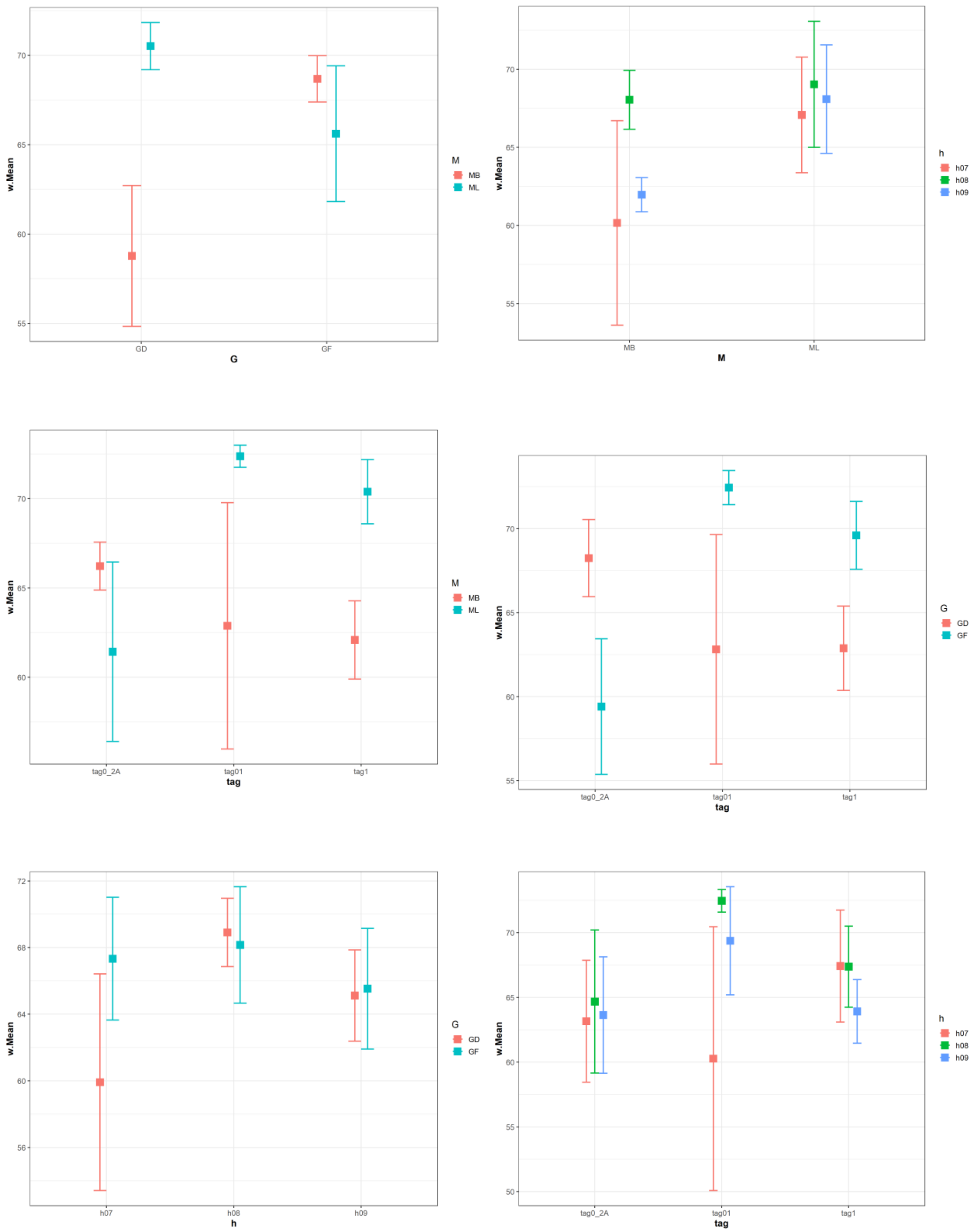


Figure 9: The weighted mean of the success on the three components of diagnostics by combining the different variables considered in the uncertainty grid. The model with only one area is only considered in the first plot of the weighted mean by area.



ANNEX

Table A1: The characteristics of the fleets defined in the model, the fleet number within ss3, the area, the gear, and the preference size composition, and the function of selectivity assumed within the model.

fleet	area	gear	years	40-60	75-100	100-150	150+	comment	selex
5	1b	bait		xx	x	x			DN5
1	1a	gill		x	xx	x			DN1
2	1a	hand			xx	xx	x		DN2
3	1a	LL			x	xx			LL3
7	1b	LL			x	xx	x		LL3
4	1a	oth		x	xx				DN4
22	1b	PS-log	<2003	xx		x			CS8
8	1b	PS-log	2003-6	xx		--			CS8
24	1b	PS-log	>2006	xx					CS8
21	1b	PS-school	<2003	x		xx			CS6
6	1b	PS-school	2003-6		x	xx			CS6
23	1b	PS-school	>2006	x		xx			CS6
9	1b	troll						no data	DN9
10	2	LL			x	xx	x		LL3
16	2	PS-school		xx		xx			CS6
17	2	PS-log		xx		x			CS8
18	2	troll						no data	DN9
11	3	LL			x	xx	--		LL3
12	4	gill		xx	x	--			DN12
13	4	LL			x	xx			LL3
14	4	oth		xx					DN14
15	4	troll		xx					DN9
19	4	PS-school				xx			CS6
20	4	PS-log		xx		--			CS8
25	4	LL-fresh				xx			LL25