 Collect. Vol. Sci. Pap. ICCAT, 79(2): 715-748 (2022)

# CURRENT STATUS OF THE NORTHERN SWORDFISH (*XIPHIAS GLADIUS*) STOCK IN THE ATLANTIC OCEAN 2022: POST-DECISIONAL STOCK ASSESSEMENT MODEL

Michael Schirripa[1]

*SUMMARY*

*This document describes the post-decisional stock assessment model configuration, diagnostics and results for the 2022 fully integrated assessment model for North Atlantic swordfish (*Xiphias gladius*). The CPUE indices used exhibited conflicts between themselves. Likewise, there was conflict between the trends suggested by the CPUE indices in general and those of the length compositions. These conflicts contributed to the overall uncertainty in the assessment results. An attempt was made to estimate the total discards of the fishery based on all available observation data. A suite of diagnostics were performed on the assessment model that further highlighted the conflicting data trends and the need for fixing, or providing informative priors, on several parameters. The stock was found not to be overfished and overfishing not to be occurring. Evaluation of the effectiveness of the current minimum size regulation was difficult to ascertain due to the period of time that has passed since the inception as well as the lack of observations of the amount and characteristics of discards.*

*RÉSUMÉ*

*Le présent document décrit la configuration post-décisionnelle du modèle d'évaluation du stock, les diagnostics et les résultats du modèle d'évaluation pleinement intégré de 2022 pour l'espadon de l'Atlantique Nord (*Xiphias gladius*). Les indices de CPUE utilisés présentaient des conflits entre eux. De même, il y avait un conflit entre les tendances suggérées par les indices de CPUE en général et celles des compositions par taille. Ces conflits ont contribué à l'incertitude générale des résultats de l'évaluation. Un essai a été fait pour estimer les rejets totaux de la pêcherie sur la base de toutes les données d'observation disponibles. Une série de diagnostics a été effectuée sur le modèle d'évaluation qui a mis en évidence les tendances contradictoires des données et la nécessité de corriger, ou de fournir des distributions a priori informatives, sur plusieurs paramètres. On a constaté que le stock n'était pas surexploité et qu'il n'était pas victime de surpêche. Il a été difficile d'évaluer l'efficacité de la réglementation actuelle sur la taille minimale en raison du temps écoulé depuis son entrée en vigueur et du manque d'observations sur la quantité et les caractéristiques des rejets.*

*RESUMEN*

*Este documento describe la configuración, los diagnósticos y los resultados del modelo de evaluación postdecisional para el modelo de evaluación plenamente integrado de 2022 del pez espada (Xiphias gladius) del Atlántico norte. Los índices de CPUE utilizados mostraron conflictos entre sí. Asimismo, hubo conflicto entre las tendencias sugeridas por los índices de CPUE en general y las de las composiciones por tallas. Estos conflictos contribuyeron a la incertidumbre general de los resultados de la evaluación. Se intentó estimar los descartes totales de la pesquería a partir de todos los datos de observación disponibles. Se realizó una serie de diagnósticos sobre el modelo de evaluación que pusieron de manifiesto las tendencias contradictorias de los datos y la necesidad de fijar, o de proporcionar distribuciones previas informativas sobre varios parámetros. Se determinó que el stock no estaba sobrepescado y que no se estaba produciendo sobrepesca. La evaluación de la eficacia de la actual regulación sobre tallas mínimas fue difícil de determinar debido al período de tiempo transcurrido desde su inicio, así como a la falta de observaciones sobre la cantidad y las características de los descartes.*

[1] NOAA Fisheries, Southeast Fisheries Center, Sustainable Fisheries Division, 75 Virginia Beach Drive, Miami, FL, 33149-1099, USA. Michael.Schirripa@noaa.gov

*KEYWORDS*

*Stock Synthesis, Swordfish, stock assessment, diagnostics, population dynamics*

## 1. Introduction

The Atlantic swordfish (*Xiphias gladius*) stock assessment meeting was held June 20-20, 2022. One of the modeling platforms employed for the northern stock was the fully integrated Stock Syntheses (SS3) (Methot and Wetzel 2013). This document describes the post-decisional stock assessment model configuration and diagnostics. Diagnostic tests used are those described in Carvalho et al. 2021.

## 2. Methods

This assessment was carried out using Stock Synthesis (SS3, V3.30.19.01;_safe;_compile_date:_Apr 15 2022). Schirripa and Hordyk (2020) describe the migration of the 2017 assessment from SS3 V3.24 to the latest version of SS3. Assumptions and configurations from the 2017 stock assessment were maintained as closely as possible while still maintaining model stability. The final base case model was labeled Base_v5.

The model configured with one area, two sexes, ten fishing fleets and seven surveys. All fleets were longline fleets except for the harpoon fishery. All observational data used to fit the model are shown in **Figure 1**.

***Landings.*** Landing were provided by the ICCAT Secretariat Task 1 data (**Figure 2**). Reported dead discards were not included in the landings as they are accounted in the "discards" section of the observational data.

***Regulatory dead discards.*** At the time of this assessment, complete estimates of the amount of live and dead discards were not available. Inspection of the freely available ICCAT Secretariat Task 1 data reveals that either only a few flags report dead discards, or, only a few generate dead discards. Length composition data suggests that several fleet report catching sub-legal size fish yet there is no accompanying report of dead discards. In an attempt to work around this lack of data, total discards were estimated within the model for the period of minimum size regulation 1993-2020 with information derived from each fleet's landings, length compositions that included discarded fish and explicitly modeling the fleets length-based selectivity and the minimum legal size. "Dummy" values for discards were input as data, however the lambda on these values was set to zero. In this way, the "dummy" values had no influence on the model fit, but the model still created expected values (i.e. estimated discards). Discards before the establishment of the minimum legal size were considered zero, which may not be the case. Until more information can be obtained on pre-regulation discards and lengths, this assumption is necessary.

Dead at-haulback from on-board observer data were available from several CPCs and used to characterize the percentage of undersized fish that were discarded dead. This approach assumes that all undersized fish dead at-haulback were discarded to comply with the minimum size regulation. Information on the percentage of undersized fish were available either directly from observer data (US and Canada) or indirectly from published literature on the topic (Coelho et al. 2019, Pan et al. 2022).

***Indices of abundance.*** The indices of abundance used were made available during the 2022 Data Prep meeting (Anonymous 2022). Fleets where catch was characterized as "retained only" were associated with that fishery and fleets that counted catch as "retained and discarded" were used as surveys.

The Spanish longline age-aggregated (SPN_1) and Combined CPUE (Combined_CPUE) were included in the configuration for convince of testing alternative hypotheses, but neither was used in the final model configuration.

***Lengths.*** Lengths of retained fish (1950-2020) and discarded fish (1993-2020) were provided by the ICCAT Secretariat. Compiling these lengths required the combining of port sampled lengths (retained catch) with onboard observer data (discarded lengths).

***Environmental data***. Estimates of the Atlantic Multidecadal Oscillation (AMO) were obtained from the NOAA website (https://psl.noaa.gov/data/correlation/amon.us.data) (AMO unsmoothed, detrended from the Kaplan SST V2. Calculated at NOAA PSL1; http://www.psl.noaa.gov/data/timeseries/AMO/). As in the 2017 assessment, the AMO was used to modulate the catchability of the following fleets/indices: Canada, Morocco, Age_1, Age_2 and Age_4.

***Data weighting.*** Weighting on all indices of abundance were scaled according to their reported variance while assuming a minimum of CV = 0.30. Length composition data was weighted according to Francis, R.I.C.C. (2011).

***Biology.*** Natural mortality for both male and female were fixed at 0.20 per age with a maximum age of 25 years. Maturity-at-age was taken from Sharma and Arocha (2017). Fecundity was made a function of body weight. Growth parameters were fixed at values developed during the 2017 ICCAT Swordfish Data Preparatory Meeting (Anon., in press) (**Figure 3**).

***Stock-recruitment relation***. A Beverton-Holt stock-recruitment relation was used with virgin recruitment (R0) freely estimated and steepness attempted to be estimated with a prior of 0.70 assuming a full beta distribution and a standard deviation of 0.06.

***Selectivity***. Selectivity was length-based with age 0 fish and older being fully selected. The double-normal function was used for those fisheries whose selectivity was estimated. The Japan_late fleet was mirrored to Japan_early; Portugal was mirrored to Spain; Chinese Taipei_late was mirrored to Chinese Taipei_early; the harpoon fleet used a logistic function; the "Other" fleet was mirrored to the US longline fishery. The harpoon and Canadian longline fisheries were assumed asymptotic while all other fisheries were allowed to estimate a dome-shaped selectivity.

***Retention.*** A retention function was applied to all fisheries (except harpoon) starting in 1993 to accommodate the minimum legal size regulation (see ICCAT Recommendation 1990-2). This regulation has very detailed exceptions that were not able to be modeled (e.g. allowance for a percent of fish < 125 cm based at the trip "landed" level). To avoid conflict with this exception, all retention functions assumed a knife-edge 119 cm minimum legal size.

***Discard mortality***. Direct observations of dead at-haulback mortality were used as proxy for discard mortality. Annual observations of dead at-haulback were available for the US and Canada via at-sea observer data. Annual averages for Chinese-Taipei were available from Pan et al. (2022). All other fisheries were assigned a value of 0.88 based on Coelho and Muñoz-Lechuga (2019). Note that using dead at-haulback as a proxy fails to consider the mortality of fish that were released alive but later died. Thus, these values likely under estimate the total discarding mortality.

## 3. Results

**Fit to CPUEs.** The SS3 fit to the indices of abundance are shown in **Figure 4 and 5.** In general, the age-specific indices were fit better than those that represented the entire age structure. The addition of the AMO as a modulator on annual catchablility to the CPUE's of Canada, Morocco, Age_1, Age_2 and Age_4 led to significantly better fits, judged by the parameter value being significantly different from zero (**Table 1**).

The trends in abundance as indicated by the CPUE's were not always consistent with one another. This lead to an estimate of biomass that was relatively flat over the past two decades. This result was one of the bases of uncertainty of the entire assessment. It could not be determined whether the conflicting signals were a result of many of the CPUEs being area specific, or perhaps some factor within the individual CPUE standardization process.

**Fit to length compositions.** The SS3 fit to the gear specific length compositions for all years is shown in **Figure 6.** Fit the length compositions aggregated across years fit very well. However, there were signs of year-specific residual patterns in some fleets. Efforts to alleviate these patterns lead to over-parameterization that eventually would have rendered the model less stable had they been maintained. It was determined that the better option was to keep the number the number of parameters at a minimum to maintain the stability of the model.

**Selectivity.** The SS3 estimated gear specific selectivities are shown in **Figure 7.** Four fleets (Canada, Chinese-Taipei early and late and harpoon) were estimated to be asymptotic while all others were dome-shaped.

**Discards.** The estimated overall regulatory discards for 1993-2020 are shown in **Figure 8.** Total estimated discards peaked (approximately 1900 mt) at the inception of the minimum size regulation in the mid to late 1990 and have shown a decreasing trend since. Estimates of recent discards average approximately 750 mt.

**Stock-recruitment.** The estimated stock-recruitment function and resulting recruitment deviations are shown in **Figure 9**. Noteworthy is the fact that the highest and lowest levels of recruitment were estimated for the same general level of SSB, suggesting at least some density-independent influence on recruit survival. Annual recruitment deviations did not show any sort of linear trend which suggests a random dispersal of points around the stock-recruitment function. Even so, six of the last eight years suggest are negative and will influence near future projections. Noteworthy is the inverted U-shape of the deviations, a pattern that is somewhat common in ICCAT assessments. This reoccurring pattern and possible causes may be deserving of further investigation.

***Model Diagnostics***

Several standard diagnostic techniques were investigated and/or applied to the three models to evaluate the stability of the parameter estimates and the model in general.

1. Joint residual and Run Test on CPUEs
2. Profiling *R0* (virgin recruitment) and *h* (steepness of S/R function)
3. ASPM diagnostic
4. Retrospective analysis
5. Hindcast cross-validation
6. Jitter analysis

***Joint residual and Run Test on CPUEs.*** To evaluate the overall model fit of the relative abundance indices and composition data, the joint-index residual plot was applied to the residuals from the fits to indices and mean length for multiple time series simultaneously. The code for this diagnostic plot was adapted from the JABBA R package (github.com/jabbamodel/JABBA) to Stock Synthesis output files and implemented as the plotting function ss3diags::SSplotJABBAres(), which provides the option to specify the type of data input.

Overall, the SMA joint-index residual plot indicated a good fit with the RMSE around 26% for the CPUE data (**Figure 10,** top) and 5.5% for the mean length data (**Figure 10**, bottom). Overall, the joint-index residual plot indicated a good fit to the CPUE data with RMSE around 30% without the use of the AMO. Use of the AMO environmental data to modulate catchability reduced the RMSE to 26 percent. The joint-length residual plot indicated a good fit to the length data with RMSE of 5.4 percent.

The runs test (Carvalho et al. 2017) is used to test for randomness of residuals of the model fit the CPUE and mean length time series (**Figure 11**). Failure to pass run tests is seen as an indication that process error is not being adequately characterized and that further model development is warranted. There was sufficient evidence ($p>=0.05$) to reject the hypothesis of randomly distributed residuals for 11 of the 13 indices of abundance when the environmental data was not used. When the AMO data was used, however, that rejection was reduced to only 8 of the 13 indices.

***Profiles on R0 and steepness.*** Profile analysis for R0 by data type is shown in **Figure 12.** Examination of profiling by individual data sources suggests that the length data is contributing the most to the estimation of R0 and the index data the next most (**Figure 13**). Profiles were also examined by individual length composition data. All three models resulted in minimums between the minimum and maximum of the range centered on R0 = 5.3. The longline lengths showed a clear minimum at R0=5.3 however the Sport lengths showed a minimum at the smallest value examined. For Model_2 there were two minimums for the Sport lengths, indicating a poor signal. This is not unexpected given the small sample size for that data. Model_3 showed the relatively best profile with three minimums at values similar to those shown by the R0 profiles.

***ASPM diagnostic.*** I used the following workflow to compute the ASPM diagnostic: (1) run the integrated Stock Synthesis model, (2) fix the selectivity parameters at the maximum likelihood estimates (MLEs), (3) turn off estimation of all parameters except R0 and the parameters representing the initial conditions (e.g., recruitment offset for the first time step of the model and initial fishing mortality parameters), (4) set the recruitment (and the initial age structure) deviates to zero (adjusting the bias-correction factor appropriately, see Methot and Taylor, 2011), and (5) fit the model to the indices of abundance only. Additionally, the model is run as for the ASPM but with recruitment deviates estimated (ASPMdev). Trends in relative SSB were then compared between the fully integrated stock assessment model, ASPM, and ASPMdev.

All three models estimated similar trends for SSB. However, the SSB from ASPM was higher than SSB from the fully integrated model and ASPMdev (**Figure 14**). The differences between the CPUE fits from the fully integrated model and the ASPM are explained by the estimated recruitment deviations in the fully integrated model.

The recruitment deviations allow for variability in age-0 recruitment and can be interpreted as the process error necessary to fit the observed trends in the CPUE data (**Figure 14**). By applying the ASPM diagnostic, it was possible to conclude that the variability in recruitment must be taken into account to estimate both the trends in CPUE 1 and the absolute scaling of SSB. The ASPMdev indicated that the CPUE data and the catches contained information on temporal variability on recruitment.

***Retrospective analysis.*** This analysis was carried out with 5-year peel. A 'rule of thumb', proposed by Hurtado-Ferro et al. (2015), suggests values of ρM that fall outside (-0.15 to 0.20) for SSB for longer-lived species, or outside (-0.22 to 0.30) for shorter-lived species indicates an undesirable retrospective pattern. In addition, the direction of the retrospective bias has implications for characterizing risk associated with management advice. A positive ρM for SSB is of particular concern because it implies a systemic overestimation of biomass, which would lead to over-optimistic quota advice if not taken into consideration (Hurtado-Ferro et al., 2015).

Results from this diagnostic show that a slight retrospective pattern is evident, however not significant (**Figure 15**). The estimates of SSB and F/Fmsy for each peal fall well within the 95% confidence interval of the reference model and the Mohn's rhos were small (rho for SSB = -0.02, rho for F/Fmsy = 0.03). These values are well within the rule proposed by Hurtado-Ferro et al. (2015).

***Hindcast cross validation.*** Implementing the HCxval diagnostic in Stock Synthesis required using the outputs produced for the retrospective routine generated by the R package r4ss. The forecasts are based on the settings in 'forecast.ss', which are also evoked when conducting future projections with the same model, only that the observed catches are used for the retrospective forecasts. A desirable feature of Stock Synthesis is that the software also computes the expected values of the observational data (e.g., abundance indices, length- or age-composition data) based on the forward-projections of the numbers- and catch-at-age matrices. Therefore, there are no additional computationally intensive tasks needed if HCxval is conducted in conjunction with retrospective analysis.

Of the ten indices of abundance used, six had MASE scores > 1 (failed) and four < 1 (passed), indicating that six had a superior prediction skill than the naïve baseline forecast of a random draw (**Figure 16**). The most accurate predictions were observed for the Age_3 index with a prediction residual MASE of 0.67 and the least accurate for the Portuguese longline CPUE with a prediction residual MASE of 3.08. The hindcast cross validation diagnostic again highlighted the lack of consistent signal emanating from all the CPUE indices considered.

The summary of the results of the above diagnostics are given in **Table 2**.

***Jitter analysis.*** A commonly used jitter value of 0.1 was used to test model stability. Results of the jitter analysis suggest realtively good model stability (**Figure 17**). Stability of the model configuration was challenged by the attempted estimation of the CPUE added variance parameters. Presumably due to the conflicting signal contained in the CPUE series, attempting to assign added variance to each on simultaneously lead to equally likely model fits. This inability is itself a useful model diagnostic highlighting the issues of conflicting CPUE times series. Ultimately, these added variance parameters were not estimated.

***Jack-knife analysis.*** To examine model sensitivity to each of the CPUE's each CPUE index was removed one at a time and the trends in SSB were compared (**Figure 18, top**). The removal of any one individual CPUE time series had virtually no effect on either the trend or the absolute value of SSB. This is likely due to the lack of a consistent signal emanating from the CPUEs as a whole. With thirteen indices, many of which are inconsistent with the others, the elimination of any one index will not lead to a significant change in the estimated trend in SSB. However, when all CPUE's were removed and only the information in the length compositions considered, a different absolute value of SSB was estimated as well a different trend. While the CPUEs in isolation suggest an increasing stock size over the past two decade, the length compositions, in isolation, suggest a decreasing trend over the most recent years.

To examine model sensitivity to each of the length compositions each fleet specific composition was removed one at a time and the trends in SSB were compared (**Figure 18, bottom**). The most influential length composition was the US fleet. This is likely due to this is the only fleet in which the lengths of discards were separated from the lengths of retained fish; the lengths from all other fleets were categorized as "combined". The model would very likely be improved by a greater distinction between the lengths of fish that were discarded and those that were retained.

***Sensitivity to assumed steepness assumption.*** Three values of steepness (0.88, 0.80 and 0.70) were considered and compared in terms of SSB, recruitment, $F/F_{MSY}$ and $SSB/SSB_{MSY}$. As the value of steepness decreased estimates of absolute SSB and $SSB/SSB_{MSY}$ increased, while the estimate of $F/F_{MSY}$ decreased (**Figure 19**). Trends in these derived quantities remained relatively consistent over the values of steepness examined.

A summary of the derived management quantities is given in Table 3.

**Effects on minimum size regulation**

One aspect of the Swordfish fishery not included in the 2017 assessment model is that of the minimum legal size limit adopted in 1991 (ICCAT Recommendation 1990-2). This regulation gave each CPC the choice of adopting a minimum size limit of 125 cm LJFL with an 11% allowance for undersized fish, or 119 cm LJFL with no allowance for undersize fish.

> *In order to protect small swordfish, the Contracting Parties take the necessary measures to prohibit the taking and landing of swordfish in the entire Atlantic Ocean weighing less than 25 kg live weight (125 cm lower jaw fork length); however, the Contracting Parties may grant tolerances to boats which have incidentally captured small fish, with the condition that this incidental catch shall not exceed 15 percent of the number of fish per landing of the total swordfish catch of said boats.*

The effectiveness of a minimum size regulation is directly proportional to the degree of discard mortality of the released fish is likely to incur. The greater the discard mortality the less effective the minimum size regulation is, to the extent that it can actually make the fishery less productive. As the most recent swordfish assessment model was chosen as the operating model (OM) for the swordfish Management Strategy Evaluation (MSE) effort these two tasks provided an opportunity to migrate the current swordfish assessment model, Stock Synthesis version 3.24 (SS3) to the latest version of Stock Synthesis 3.30 (SS).

A minimum legal size regulation can affect the fishery in at least two different ways. In the first way, the fishery can maintain its pre-regulatory fishing practices and simply discard those fish that do not meet the minimum size limit. In this case, the selectivity of undersized fish would be assumed not to differ between pre- and post-regulatory periods. This assumption would be modeled by estimating the selectivity parameters for one continuous period for the entire time series of length compositions and introducing a retention curve corresponding to the years of the regulation. In this case undersized fish would be assumed to be encountered at the same rate pre- and post-regulation and thus subjected to some assumed value of discard mortality.

Another way the fishery may respond to a minimum size regulation is for the fishery to change its pre-regulatory practices to avoid undersized fish. This could include changes such as, but not limited to, location, season, depth of gear, bait type, or to stop fishing entirely. In this case, the selectivity of undersized fish would be assumed to change between pre- and post-regulatory periods. This assumption would be modeled by estimating the selectivity separately for the pre- and post-regulatory period (i.e. time varying). Perhaps the most likely response.
The effect of the minimum size regulation was examined taking both of the above possibilities into consideration.

The percent of undersized fish observed dead at-haulback by year from the US (observer data), Canada (observer data), and averaged over years for Chinese (Pan et al. 2022) and Portugal (Coelho and Muñoz-Lechuga 2019) are shown in **Figure 20**. On average, the percent of undersized fish observed to be dead at-haulback were 73.6% from the US; 63.3% from Canada; 83.6% from the Chinese fleet, and 88% from the Portugal fleet. It is apparent from these observations and studies that there is a high mortality rate for undersized fish. However, what is not fully understood (and more important) is the ***encounter rate*** of undersized fish and how this rate might have changed over time. While we can get some information from an examination of the total estimated discards from the SS3 model (**Figure 8**), which shows a gradual decline in discards, we cannot be certain if this decline is due to a decrease in the encounter rate or a decrease in recruitment, or both.

Examination of the length compositions aggregated across pre- and post-regulation periods may also provide some indication of trends in encounter rate of undersized fish (**Figure 21**). Examination of these set of length compositions suggest that some fisheries (the US) have decreased their encounter rate while others (Morocco) may have increased their rate. However, given the long period that has elapsed since the inception of the regulation 1992 makes any indications suggested by the length compositions subject to a variety of other factors that have taken place in these fleets and the sampling of lengths over the past thirty years. In this regard, caution should be exercised before drawing any conclusions.

***Equilibrium, yield-per-recruit***. While ICCAT does not use yield-per-recruit (YPR) as a management benchmark for Atlantic swordfish, a useful metric and does not rely on many of the assumptions associated with a population level model. This YPR analysis makes the explicit assumption that the encounter rate of undersized fish does not change between the minimum sizes considered. In other words, selectivity of the fishery remains constant. The Ricker (1971) formulation of yield-per-recruit (YPR) was used to develop matrices of YPR by minimum size and discard mortality. A single, catch weighted overall fishery selectivity was used to represent all north Atlantic swordfish fisheries. The biological parameters used where the same as those used in the assessment. Discard mortalities, which includes dead at-haulback mortality, of 0, 25, 45 and 88% were examined.

Under the assumption that all fish encountered (i.e. hooked) survive, YPR is maximized at minimum size of approximately 160 cm and a high (> 0.80) fishing mortality (**Figure 22, upper left**). However, if the assumed discard mortality (which includes dead at-haulback) to be just 25% then the minimum size at maximum YPR drops to approximately 120 cm but YPR is decreased by approximately 37.5% at fishing mortality of approximately F = 0.30 (**Figure 22, upper right**). If the assumed discard mortality (which includes dead at-haulback) is assumed to be 45% then the minimum size at maximum YPR drops to < 100 cm and YPR is decreased by approximately 50% at fishing mortality of approximately F = 0.25 (**Figure 22, lower left**). If the assumed discard mortality (which includes dead at-haulback) to be 88% then the minimum size at maximum YPR drops to approximately 75 cm and YPR is decreased by approximately 50% from that for 0% discard mortality at fishing mortality of approximately F = 0.20 (**Figure 22, lower right**).

The spawning stock biomass-per-recruit (SSBR) was examined for the same levels of discard mortality. The SSBR estimates the expected lifetime reproductive potential of an average recruit (P), which is an important correlate of population growth potential. The ratio of the fished to unfished magnitude of P is the spawning potential ratio (SPR) and is a measure of the impact of fishing on the potential productivity of a stock (Goodyear 1993).

Under the assumption that all fish encountered (i.e. hooked) survive, SPR can be incrementally maximized by incrementally increasing the minimum size at fishing mortalities above approximately F = 0.40 (**Figure 23, upper left**). However, if the discard mortality (which includes dead at-haulback) is assumed to be just 25% the SPR isopleths become nearly vertical in nature, indicating that increases in minimum size do very little to increase SPR under any fishing mortality value (**Figure 22, upper right**). If the assumed discard mortality (which includes dead at-haulback) is assumed to be 45% the isopleths become even more vertical wiith virtually no gains in SPR with increases in minimum size (**Figure 22, lower left**). If the assumed discard mortality (which includes dead at-haulback) to be 88% then the minimum size at maximum the isopleths indicate that no gain in SPR is to be had be any increase in minimum size and that SPR is strictly a function of effort (**Figure 22, lower right**). In such a case, management measures that would decrease the encounter rate with small fish (e.g. closed areas/seasons) likely have a higher probability of increasing the SPR over time than do minimum sizes. However, if a minimum size regulation motivates fleets to decrease their effort on small fish the fishery can still realize the conservation effects, if with a high encounter mortality rate.

## 4. Discussion

Model diagnostics suggest that there are conflicts between the signals in population trend amongst the various CPUE times series as well as between the CPUEs in general and the length compositions. These conflicts could be due to length sampling error/bias, inadequate CPUE standardization (i.e. observational uncertainties) and/or population model misspecification (i.e. structural uncertainties). These uncertainties could be considered when formulating the management strategy evaluation grid of uncertainty, which uses the same SS3 model and configuration.

Further biological studies are recommended to challenge the hypothesis on one unit stock for the north Atlantic. The CPUE's provided to the assessment effort have the potential of being area-specific and may not reflect the overall trend of the population. The sampling of lengths and the subsequent weight assigned these lengths is another area that could benefit from further research. Given the sensitivity of the model to the length compositional data greater effort should be made to separate out the lengths of those fish that are discarded from those that are retained, if possible. An evaluation of the current minimum size regulation can only be carried out by having data on the discarding rate brought about by the regulation.

**Acknowledgements**

The author wishes to acknowledge the efforts of the ICCAT Secretariat (especially A. Kimoto, M. Ortiz and C. Palma) for their support in the development of this assessment.

**References**

Arocha, F., C. Moreno, L. Beerkircher, D. W. Lee and L. Marcano. 2003. Update on the growth estimates for the swordfish, Xiphias gladius, in the Northwestern Atlantic. Collect. Vol. Sci. Pap. ICCAT, 55(4): 1416-1429.

Carvalho, F., Winker, H., Courtney, D., Kapur, M., Kell, L., Cardinale, M., Schirripa, M., Kitakado, T., Yemane, D., Piner, K.R. and Maunder, M.N., 2021. A cookbook for using model diagnostics in integrated stock assessments. Fisheries Research, 240, p.105959.

Coelho, R. and Muñoz-Lechuga, R., 2019. Hooking mortality of swordfish in pelagic longlines: comments on the efficiency of minimum retention sizes. Reviews in Fish Biology and Fisheries, 29(2), pp.453-463.

Diaz, G.A., 2020. The effect of circle hooks vs J hooks on the at-haulback survival in the US Atlantic pelagic longline fleet. Vol. Sci. Pap. ICCAT, 77(4), pp.127-136.

FrancisR.I.C. Chris. Data weighting in statistical fisheries stock assessment models. Canadian Journal of Fisheries and Aquatic Sciences. 68(6): 1124-1138. https://doi.org/10.1139/f2011-025

Goodyear, C.P., 1993. Spawning stock biomass per recruit in fisheries management: foundation and current use. Canadian Special Publication of Fisheries and Aquatic Sciences, pp.67-82.

Hurtado-Ferro, F., Szuwalski, C.S., Valero, J.L., Anderson, S.C., Cunningham, C.J., Johnson, K.F., Licandeo, R., McGilliard, C.R., Monnahan, C.C., Muradian, M.L., Ono, K., Vert-Pre, K.A., Whitten, A.R., Punt, A.E., 2015. Looking in the rear-view mirror: bias and retrospective patterns in integrated, age-structured stock assessment models. Ices J. Mar. Sci. 72, 99–110. https://doi.org/10.1093/icesjms/fsu198.

Methot, R.D., Wetzel, C.R., 2013. Stock synthesis: a biological and statistical framework for fish stock assessment and fishery management. Fish. Res. 142, 86–99. https://doi.org/10.1016/j.fishres.2012.10.012.

Pan, B., Zhu, J., Lin, Q., Geng, Z., Wu, F. and Zhang, Y., 2022. Study on the catch, bycatch and discard of Chinese pelagic longline fisheries in the Atlantic Ocean. Aquaculture and Fisheries.

Schirripa, M. and Hordyk, A., 2020. Migrating the North Atlantic swordfish stock assessment model to an updated version of Stock Synthesis with analysis of the current minimum size regulation. Collect. Vol. Sci. Pap. ICCAT, 77(3), pp.654-668.

Sharma, R. and F. Arocha. 2017. Resiliency for north Atlantic swordfish using life history parameters. Collect. Vol. Sci. Pap. ICCAT, 74(3): 1306-1321.

Waters, J.R. and Huntsman, G.R., 1986. Incorporating mortality from catch and release into yield-per-recruit analyses of minimum-size limits. North American Journal of Fisheries Management, 6(4), pp.463-471.

**Table 1.** Estimated parameters, values and standard deviations of the North Atlantic swordfish SS3 assessment model, Base_v5.

| Parameter | Value | Phase | Min | Max | Init | Status | Parm_StDev | Gradient | Pr_type | Prior | Pr_SD |
|---|---|---|---|---|---|---|---|---|---|---|---|
| SR_LN(R0) | 6.78 | 1 | 6 | 8 | 6.78 | OK | 0.0134 | -1.46E-03 | No_prior | NA | NA |
| LnQ_base_CAN_3(3) | -7.05 | 1 | -10 | -5 | -7.05 | OK | 0.0807 | -3.34E-03 | No_prior | NA | NA |
| LnQ_base_JPN_ERLY_4(4) | -7.24 | 1 | -10 | -5 | -7.24 | OK | 0.0911 | -8.02E-04 | No_prior | NA | NA |
| LnQ_base_JPN_LATE_5(5) | -6.37 | 1 | -10 | -2 | -6.37 | OK | 0.1055 | 6.10E-03 | No_prior | NA | NA |
| LnQ_base_CHT_EARLY_7(7) | -7.67 | 1 | -10 | -5 | -7.67 | OK | 0.0612 | 9.74E-04 | No_prior | NA | NA |
| LnQ_base_CHT_LATE_8(8) | -6.82 | 1 | -10 | -5 | -6.82 | OK | 0.0921 | 3.82E-04 | No_prior | NA | NA |
| LnQ_base_MOR_9(9) | -10.77 | 1 | -15 | -7 | -10.77 | OK | 0.1314 | -2.23E-04 | No_prior | NA | NA |
| LnQ_base_US_Survey_12(12) | -6.54 | 1 | -10 | -5 | -6.53 | OK | 0.0728 | 4.41E-03 | No_prior | NA | NA |
| LnQ_base_PORT_Survey_13(13) | -11.02 | 1 | -15 | -5 | -11.02 | OK | 0.0749 | -1.48E-03 | No_prior | NA | NA |
| LnQ_base_Age-1(14) | -6.47 | 1 | -10 | -3 | -6.47 | OK | 0.0412 | -1.68E-04 | No_prior | NA | NA |
| LnQ_base_Age-2(15) | -6.02 | 1 | -10 | -3 | -6.02 | OK | 0.0428 | -6.72E-04 | No_prior | NA | NA |
| LnQ_base_Age-3(16) | -5.47 | 1 | -10 | -3 | -5.47 | OK | 0.0447 | -1.24E-03 | No_prior | NA | NA |
| LnQ_base_Age-4(17) | -5.05 | 1 | -7 | -3 | -5.05 | OK | 0.0500 | -7.13E-03 | No_prior | NA | NA |
| LnQ_base_Age-5+(18) | -6.07 | 1 | -10 | -3 | -6.06 | OK | 0.0644 | -1.80E-03 | No_prior | NA | NA |
| LnQ_base_CAN_3(3)_ENV_mult | -0.16 | 7 | -1 | 1 | -0.16 | OK | 0.0316 | -2.42E-04 | No_prior | NA | NA |
| LnQ_base_MOR_9(9)_ENV_mult | -0.13 | 7 | -1 | 1 | -0.13 | OK | 0.0376 | -2.63E-04 | No_prior | NA | NA |
| LnQ_base_Age-1(14)_ENV_mult | -0.24 | 7 | -1 | 1 | -0.24 | OK | 0.0257 | 3.17E-03 | No_prior | NA | NA |
| LnQ_base_Age-2(15)_ENV_mult | -0.19 | 7 | -1 | 1 | -0.19 | OK | 0.0264 | 1.96E-04 | No_prior | NA | NA |
| LnQ_base_Age-4(17)_ENV_mult | 0.12 | 7 | -1 | 1 | 0.12 | OK | 0.0317 | 1.57E-03 | No_prior | NA | NA |
| Size_DblN_top_logit_SPN_1(1) | -0.98 | 3 | -10 | 0.5 | -0.98 | OK | 0.2036 | 1.83E-03 | No_prior | NA | NA |
| Size_DblN_descend_se_SPN_1(1) | 8.58 | 3 | 0 | 15 | 8.58 | OK | 0.3771 | 1.05E-03 | No_prior | NA | NA |
| Size_DblN_end_logit_SPN_1(1) | -2.56 | 6 | -10 | 5 | -2.56 | OK | 0.7624 | 5.04E-04 | Normal | -2.60 | 0.78 |
| Size_DblN_descend_se_US_2(2) | 6.58 | 3 | 2 | 12 | 6.58 | OK | 0.1092 | -5.45E-04 | No_prior | NA | NA |
| Size_DblN_end_logit_US_2(2) | -1.29 | 6 | -8 | 0 | -1.29 | OK | 0.1236 | 3.20E-03 | Normal | -0.45 | 3.58 |
| Size_DblN_top_logit_JPN_ERLY_4(4) | -6.11 | 4 | -15 | -0.5 | -6.11 | OK | 1.1939 | -5.00E-05 | Normal | -6.10 | 1.20 |
| Size_DblN_descend_se_JPN_ERLY_4(4) | 4.30 | 3 | -10 | 15 | 4.30 | OK | 0.9210 | 7.11E-04 | Normal | 4.21 | 0.90 |
| Size_DblN_end_logit_JPN_ERLY_4(4) | 0.65 | 6 | -5 | 5 | 0.65 | OK | 0.1766 | -3.46E-04 | Normal | 0.64 | 0.19 |
| Size_DblN_top_logit_CHT_EARLY_7(7) | -6.08 | 4 | -10 | -3 | -6.08 | OK | 0.6101 | -7.76E-05 | Normal | -6.08 | 0.61 |
| Size_DblN_descend_se_CHT_EARLY_7(7) | -1.00 | 3 | -5 | 5 | -1.00 | OK | 0.2000 | 8.63E-05 | Normal | -1.00 | 0.20 |
| Size_DblN_end_logit_CHT_EARLY_7(7) | 8.47 | 6 | -5 | 15 | 8.47 | OK | 0.8504 | 5.56E-04 | Normal | 8.47 | 0.85 |
| Size_DblN_peak_MOR_9(9) | 125.41 | 3 | 80 | 180 | 125.41 | OK | 4.0826 | 1.15E-04 | No_prior | NA | NA |
| Size_DblN_top_logit_MOR_9(9) | -11.44 | 3 | -15 | -5 | -11.44 | OK | 2.2886 | 8.93E-05 | Normal | -11.44 | 2.29 |
| Size_DblN_ascend_se_MOR_9(9) | 5.98 | 3 | 0 | 15 | 5.98 | OK | 0.4623 | 1.01E-03 | No_prior | NA | NA |
| Size_DblN_descend_se_MOR_9(9) | 6.35 | 3 | 0 | 15 | 6.35 | OK | 0.5420 | -1.88E-03 | No_prior | NA | NA |
| Size_DblN_end_logit_MOR_9(9) | -1.23 | 5 | -5 | 5 | -1.23 | OK | 0.3752 | 1.14E-03 | No_prior | NA | NA |
| Size_inflection_HRPN_10(10) | 185.33 | 3 | 120 | 280 | 185.34 | OK | 1.9976 | -3.70E-03 | No_prior | NA | NA |
| Size_95%width_HRPN_10(10) | 29.91 | 3 | 0 | 135 | 29.91 | OK | 1.9581 | 4.05E-03 | No_prior | NA | NA |
| Size_DblN_peak_SPN_1(1)_BLK1repl_1950 | 88.20 | 3 | 50 | 120 | 88.20 | OK | 2.8039 | 1.68E-03 | No_prior | NA | NA |
| Size_DblN_peak_SPN_1(1)_BLK1repl_1993 | 84.17 | 3 | 50 | 120 | 84.15 | OK | 1.6004 | 8.94E-04 | No_prior | NA | NA |
| Size_DblN_ascend_se_SPN_1(1)_BLK1repl_1950 | 4.18 | 3 | -5 | 7 | 4.18 | OK | 0.6191 | -5.81E-04 | No_prior | NA | NA |
| Size_DblN_ascend_se_SPN_1(1)_BLK1repl_1993 | 3.74 | 3 | -5 | 7 | 3.74 | OK | 0.3945 | 2.01E-03 | No_prior | NA | NA |
| Size_DblN_peak_US_2(2)_BLK1repl_1950 | 82.96 | 3 | 50 | 170 | 82.96 | OK | 2.3534 | 7.51E-03 | No_prior | NA | NA |
| Size_DblN_peak_US_2(2)_BLK1repl_1993 | 88.41 | 3 | 50 | 170 | 88.38 | OK | 0.8115 | -9.21E-04 | No_prior | NA | NA |
| Size_DblN_top_logit_US_2(2)_BLK1repl_1950 | -0.32 | 3 | -15 | 5 | -0.32 | OK | 0.0829 | -1.69E-03 | Normal | -0.41 | 5.13 |
| Size_DblN_top_logit_US_2(2)_BLK1repl_1993 | -7.08 | 3 | -15 | 5 | -7.08 | OK | 1.2288 | 6.94E-04 | Normal | -6.19 | 1.69 |
| Size_DblN_ascend_se_US_2(2)_BLK1repl_1950 | 4.45 | 3 | -5 | 10 | 4.45 | OK | 0.3783 | -1.12E-03 | No_prior | NA | NA |
| Size_DblN_ascend_se_US_2(2)_BLK1repl_1993 | 5.81 | 3 | -5 | 10 | 5.81 | OK | 0.0953 | -2.40E-03 | No_prior | NA | NA |
| Retain_L_infl_US_2(2)_BLK2repl_2013 | 115.81 | 6 | 50 | 250 | 115.81 | OK | 0.2420 | 2.86E-03 | No_prior | NA | NA |
| Size_DblN_peak_CAN_3(3)_BLK1repl_1993 | 176.02 | 3 | 100 | 200 | 176.03 | OK | 5.1454 | -2.38E-04 | No_prior | NA | NA |
| Size_DblN_ascend_se_CAN_3(3)_BLK1repl_1950 | 10.72 | 3 | 0 | 15 | 10.72 | OK | 0.2798 | -2.36E-03 | No_prior | NA | NA |
| Size_DblN_ascend_se_CAN_3(3)_BLK1repl_1993 | 7.55 | 3 | 0 | 15 | 7.55 | OK | 0.1839 | 6.00E-04 | No_prior | NA | NA |
| Size_DblN_peak_JPN_ERLY_4(4)_BLK1repl_1950 | 202.68 | 3 | 55 | 317.5 | 202.68 | OK | 7.4707 | 1.27E-03 | No_prior | NA | NA |
| Size_DblN_peak_JPN_ERLY_4(4)_BLK1repl_1993 | 205.20 | 3 | 55 | 315 | 205.20 | OK | 5.5107 | 3.34E-04 | No_prior | NA | NA |
| Size_DblN_ascend_se_JPN_ERLY_4(4)_BLK1repl_1950 | 8.75 | 3 | 1 | 15 | 8.75 | OK | 0.2539 | -2.37E-03 | Normal | 12.86 | 4.80 |
| Size_DblN_ascend_se_JPN_ERLY_4(4)_BLK1repl_1993 | 8.04 | 3 | -5 | 15 | 8.04 | OK | 0.1426 | -2.07E-03 | No_prior | NA | NA |
| Size_DblN_peak_CHT_EARLY_7(7)_BLK1repl_1950 | 136.05 | 3 | 100 | 200 | 136.30 | OK | 3.4872 | -2.50E-03 | Normal | 136.00 | 27.00 |
| Size_DblN_peak_CHT_EARLY_7(7)_BLK1repl_1993 | 166.94 | 3 | 100 | 250 | 166.92 | OK | 5.5026 | -4.06E-04 | Normal | 167.00 | 33.00 |
| Size_DblN_ascend_se_CHT_EARLY_7(7)_BLK1repl_1950 | 4.74 | 3 | 0 | 15 | 4.78 | OK | 0.5097 | -1.36E-03 | Normal | 4.78 | 0.95 |
| Size_DblN_ascend_se_CHT_EARLY_7(7)_BLK1repl_1993 | 6.76 | 3 | 0 | 15 | 6.75 | OK | 0.2905 | -3.37E-04 | Normal | 6.76 | 1.35 |

**Table 2.** Summary statistics runs tests, retrospective analysis, and hindcast cross-validation (HCxval) model diagnostics applied to the North Atlantic swordfish Stock Synthesis Base_v5 model.

| Diagnostic | Index | type | Stastitic | Value | Result |
|---|---|---|---|---|---|
| Runs Test | CAN_3 | cpue | p-value | -0.71 | Failed |
| Runs Test | JPN_ERLY_4 | cpue | p-value | -0.38 | Failed |
| Runs Test | JPN_LATE_5 | cpue | p-value | -0.45 | Passed |
| Runs Test | CHT_EARLY_7 | cpue | p-value | -0.43 | Failed |
| Runs Test | CHT_LATE_8 | cpue | p-value | -0.71 | Failed |
| Runs Test | MOR_9 | cpue | p-value | -0.35 | Passed |
| Runs Test | US_Survey_12 | cpue | p-value | -0.39 | Failed |
| Runs Test | PORT_Survey_13 | cpue | p-value | -0.38 | Passed |
| Runs Test | Age-1 | cpue | p-value | -0.79 | Passed |
| Runs Test | Age-2 | cpue | p-value | -0.41 | Failed |
| Runs Test | Age-3 | cpue | p-value | -0.44 | Failed |
| Runs Test | Age-4 | cpue | p-value | -0.46 | Passed |
| Runs Test | Age-5+ | cpue | p-value | -0.38 | Failed |
| Runs Test | SPN_1 | len | p-value | 0.00 | Failed |
| Runs Test | US_2 | len | p-value | 0.00 | Failed |
| Runs Test | CAN_3 | len | p-value | 0.00 | Failed |
| Runs Test | JPN_ERLY_4 | len | p-value | 0.02 | Failed |
| Runs Test | JPN_LATE_5 | len | p-value | 0.32 | Passed |
| Runs Test | PORT_6 | len | p-value | 0.00 | Failed |
| Runs Test | CHT_EARLY_7 | len | p-value | 0.00 | Failed |
| Runs Test | CHT_LATE_8 | len | p-value | 0.00 | Failed |
| Runs Test | MOR_9 | len | p-value | 0.00 | Failed |
| Runs Test | HRPN_10 | len | p-value | 0.01 | Failed |
| Retrospective analysis | | SSB | Mohn's rho | -0.02 | Passed |
| Retrospective analysis | | Bratio | Mohn's rho | 0.00 | Passed |
| Retrospective analysis | | Fratio | Mohn's rho | 0.03 | Passed |
| Hcxval | CAN_3 | CPUE | MASE | 1.30 | Fail |
| Hcxval | JPN_LATE_5 | CPUE | MASE | 0.78 | Pass |
| Hcxval | CHT_LATE_8 | CPUE | MASE | 0.75 | Pass |
| Hcxval | MOR_9 | CPUE | MASE | 1.68 | Fail |
| Hcxval | US_Survey_12 | CPUE | MASE | 1.71 | Fail |
| Hcxval | PORT_Survey_13 | CPUE | MASE | 3.08 | Fail |
| Hcxval | Age-2 | CPUE | MASE | 0.88 | Pass |
| Hcxval | Age-3 | CPUE | MASE | 0.67 | Pass |
| Hcxval | Age-4 | CPUE | MASE | 1.30 | Fail |
| Hcxval | Age-5+ | CPUE | MASE | 1.71 | Fail |

**Table 3**. Derived quantities from for the North Atlantic swordfish SS3 model Base_v5.

| Derived Quantity | h = 0.70 | h = 0.80 | h = 0.88 |
|---|---|---|---|
| F_2020 | 0.559 | 0.679 | 0.781 |
| Bratio_2020 | 2.215 | 1.430 | 1.115 |
| SSB_unfished | 151,883 | 133,541 | 120,468 |
| Totbio_unfished | 335,057 | 294,595 | 265,755 |
| SmryBio_unfished | 335,057 | 294,595 | 265,755 |
| Recr_unfished | 1,112 | 977 | 882 |
| SSB_F01 | 0.125 | 0.170 | 0.198 |
| annF_F01 | 0.166 | 0.165 | 0.163 |
| Dead_Catch_F01 | 10,236 | 12,224 | 12,837 |
| SSB_SPR | 28,384 | 30,127 | 30,029 |
| annF_SPR | 0.136 | 0.136 | 0.137 |
| Dead_Catch_SPR | 11,609 | 12,515 | 12,638 |
| SSB_MSY | 18,707 | 22,464 | 23,666 |
| SPR_MSY | 0.217 | 0.220 | 0.223 |
| annF_MSY | 0.167 | 0.165 | 0.164 |
| Dead_Catch_MSY | 10,178 | 12,206 | 12,838 |
| Ret_Catch_MSY | 9,480 | 11,423 | 12,062 |
| B_MSY/SSB_unfished | 0.123 | 0.168 | 0.196 |

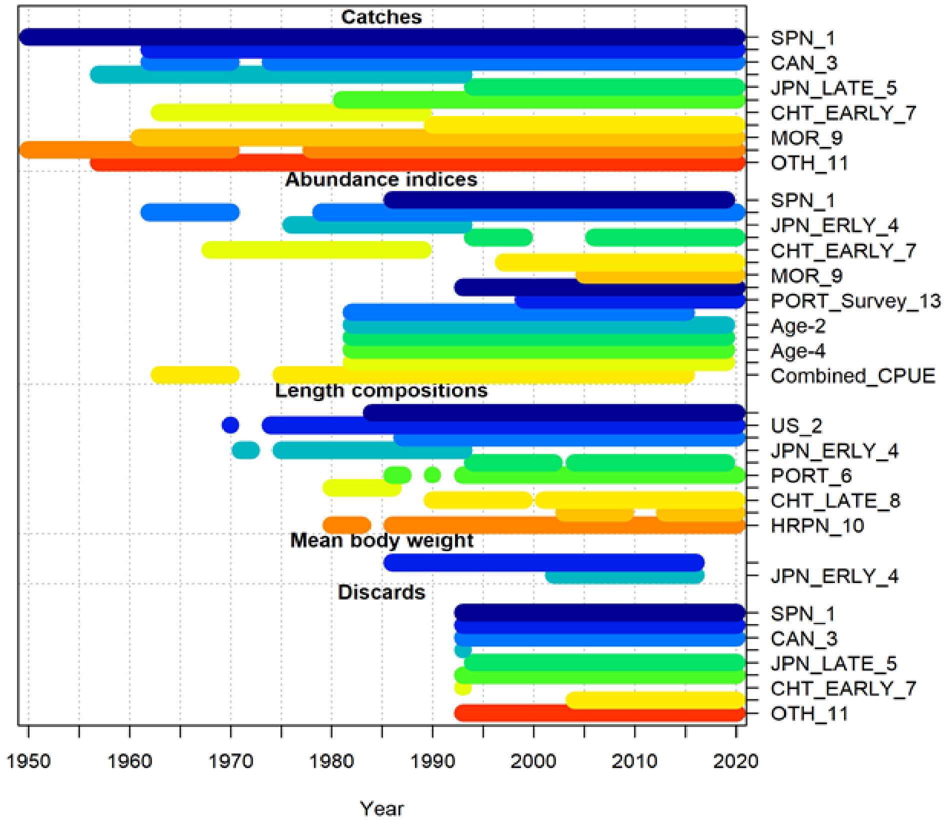



**Figure 1.** Observational data sources by year used in the Stock Synthesis models for North Atlantic swordfish.

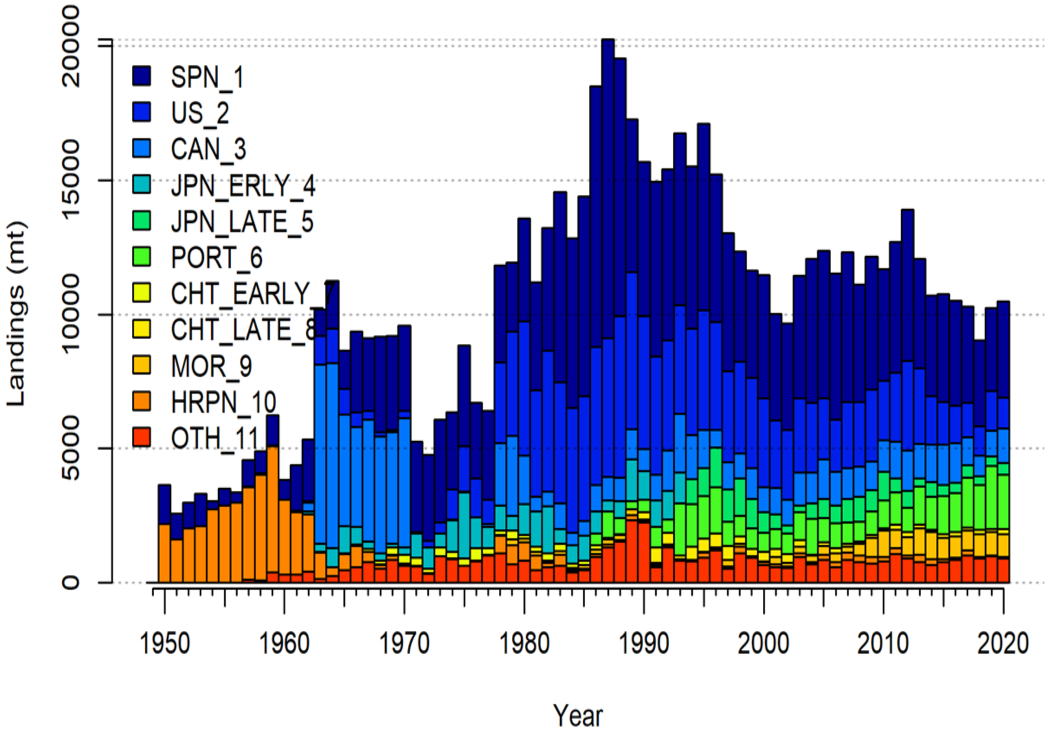



**Figure 2.** Total landings of North Atlantic swordfish year and gear, 1950-2020.

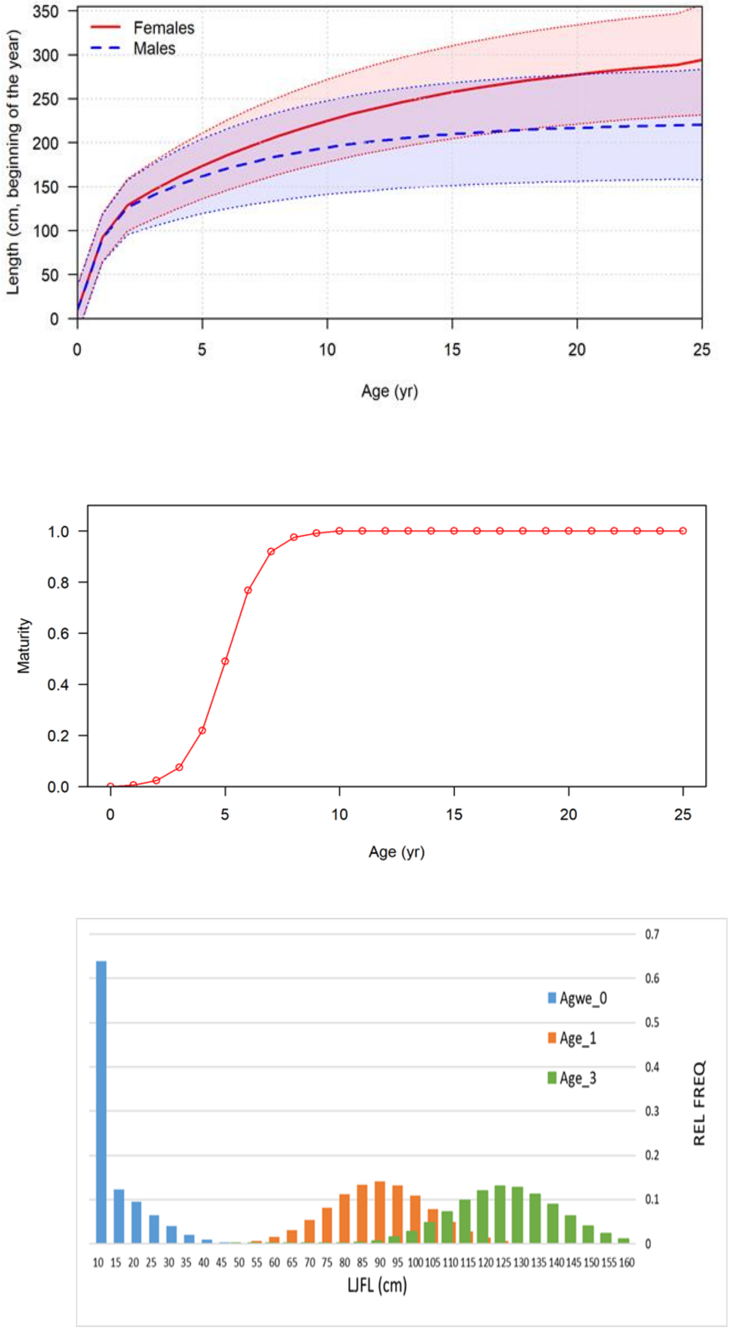



**Figure 3.** Growth (top), maturity (center) and size-at-age for ages 0, 1 and 2 for North Atlantic swordfish.

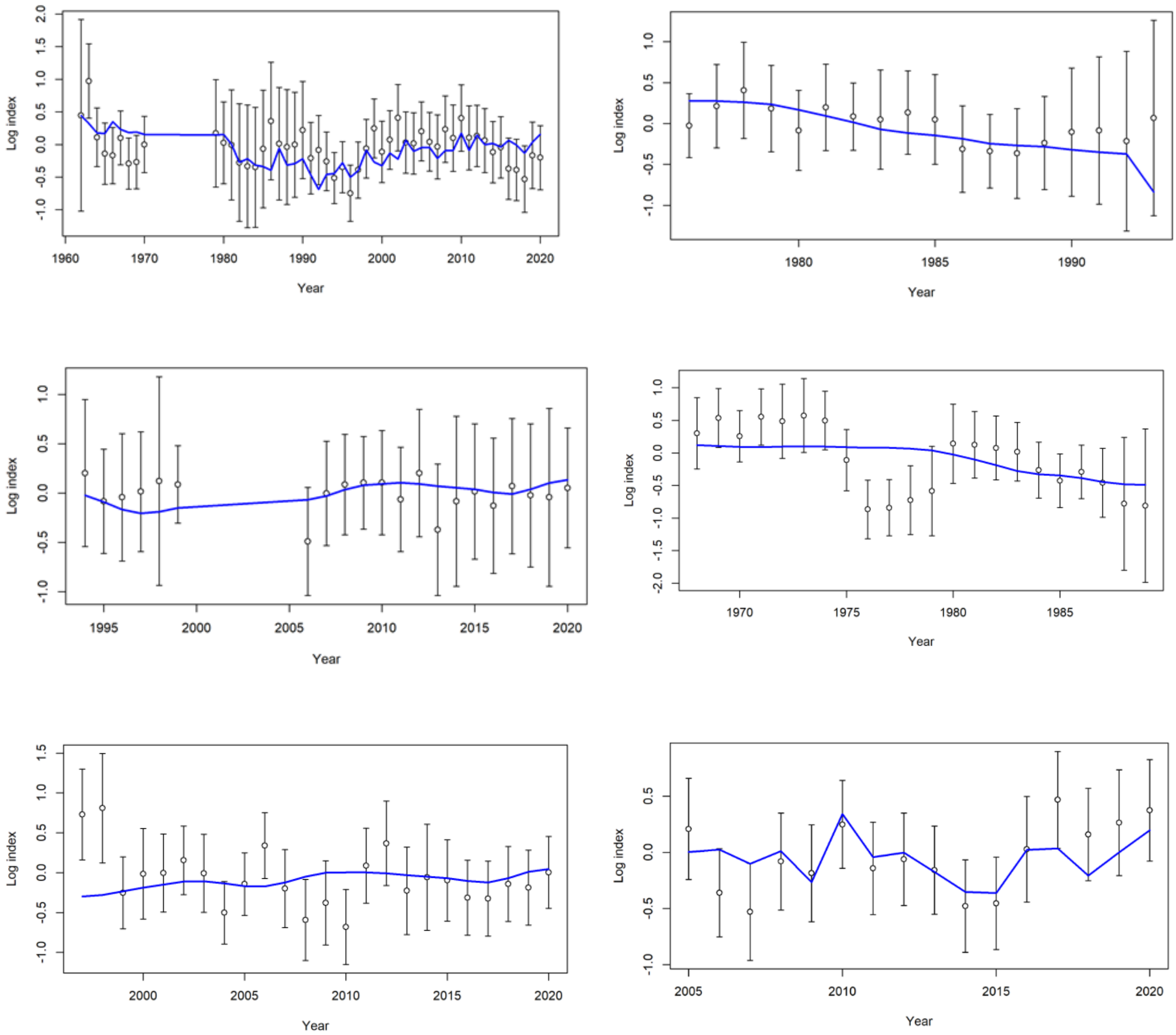



**Figure 4**. SS fit to North Atlantic swordfish indices of abundance. From left to right Canada, Japan early, Japan late, Chinese-Taipei early, Chinese-Taipei late and Morocco).

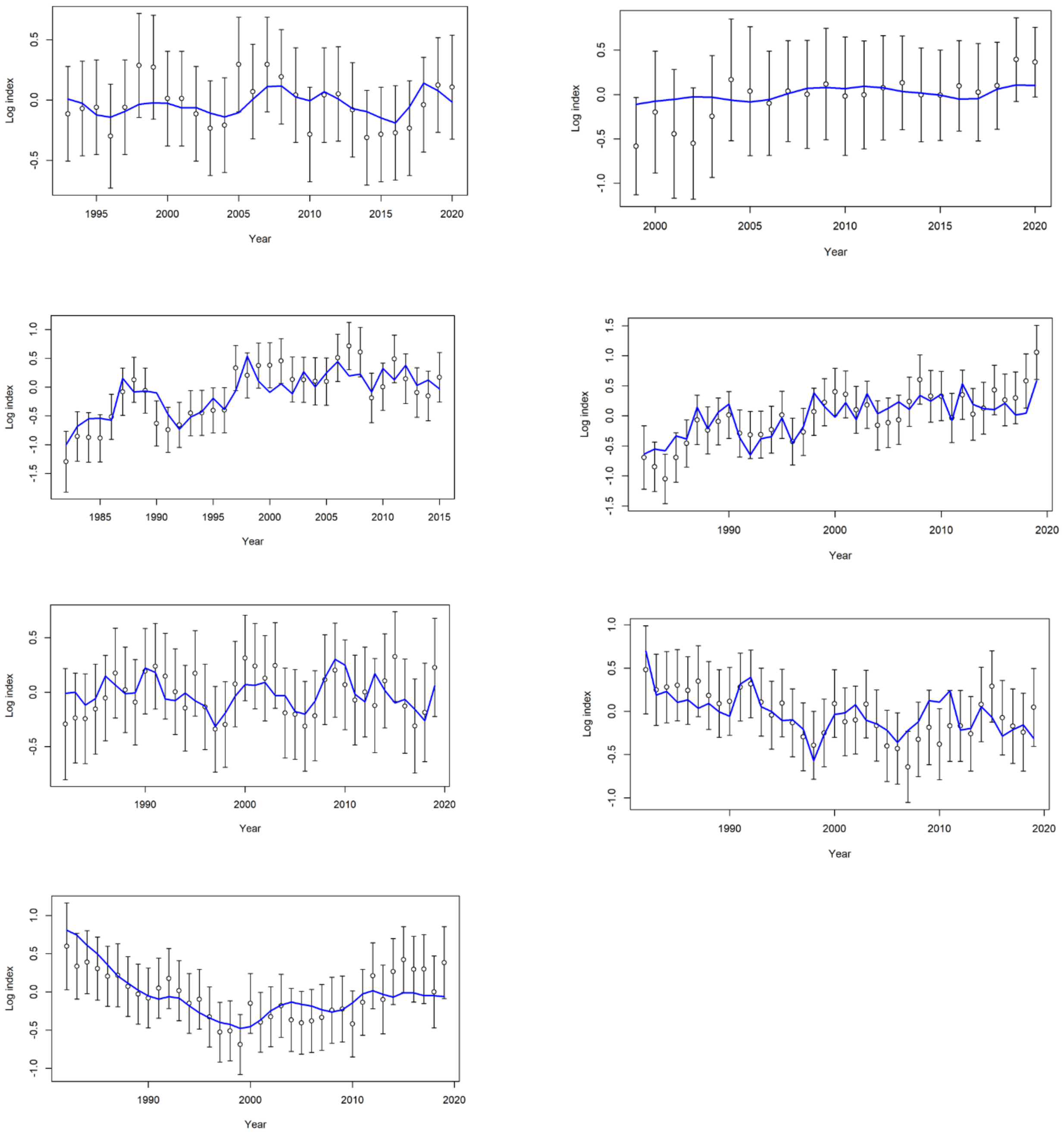



**Figure 5.** Fit to North Atlantic swordfish indices of abundance for the SS3 model. From left to right US, Portugal, age 1, 2, 3, 4 and 5+.

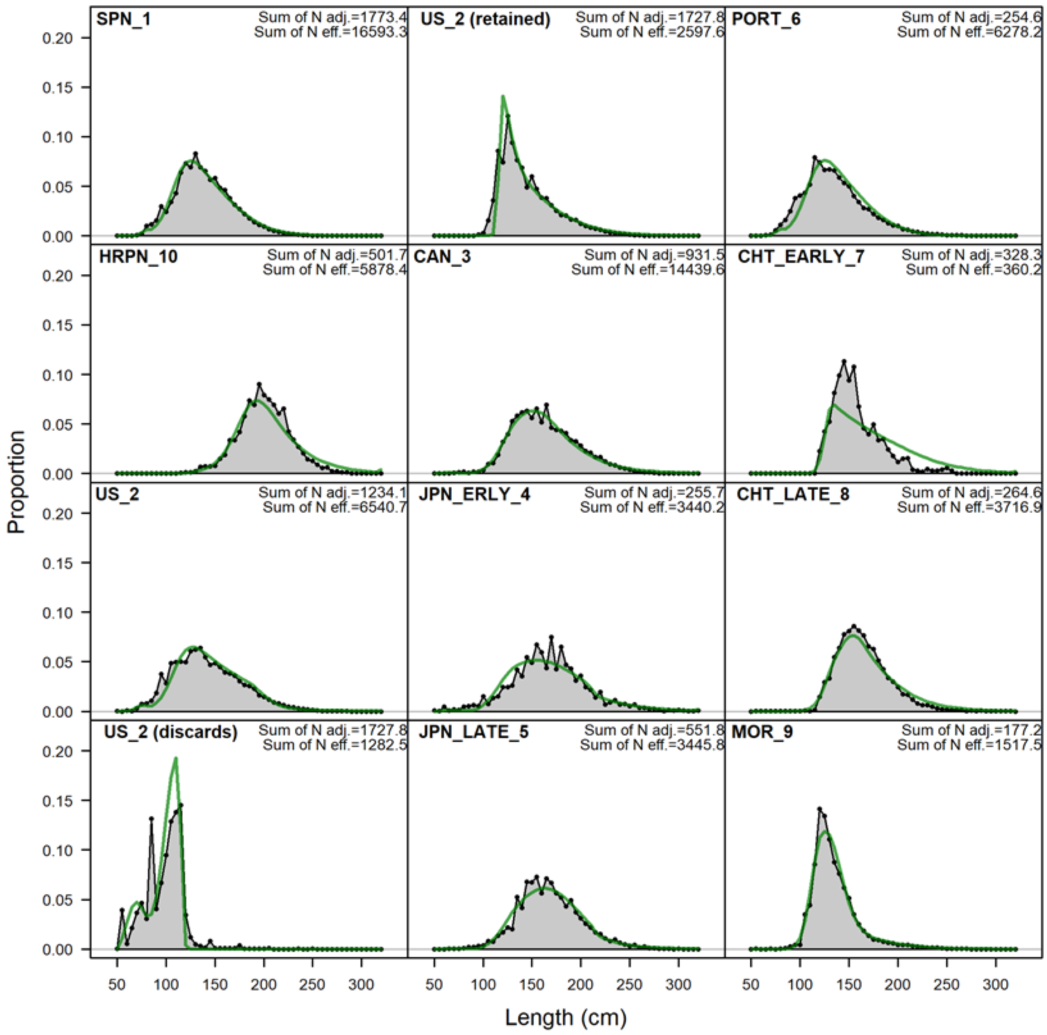



**Figure 6.** Length frequencies for North Atlantic swordfish and SS3 model fit by fishery across all years.

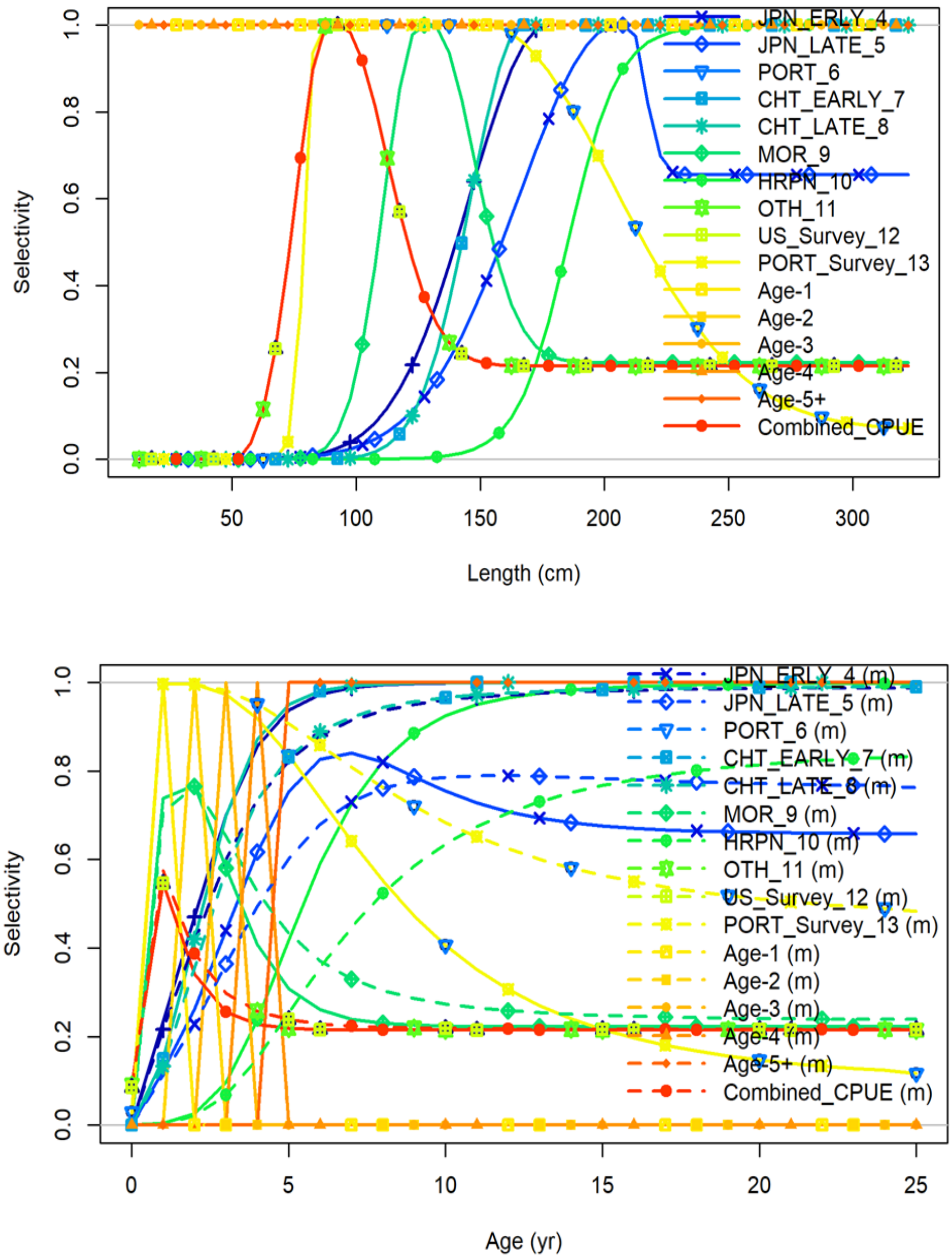



**Figure 7.** Estimated ending year (top) length and (bottom) age-length selectivities by fleet for North Atlantic swordfish from the SS3 model.

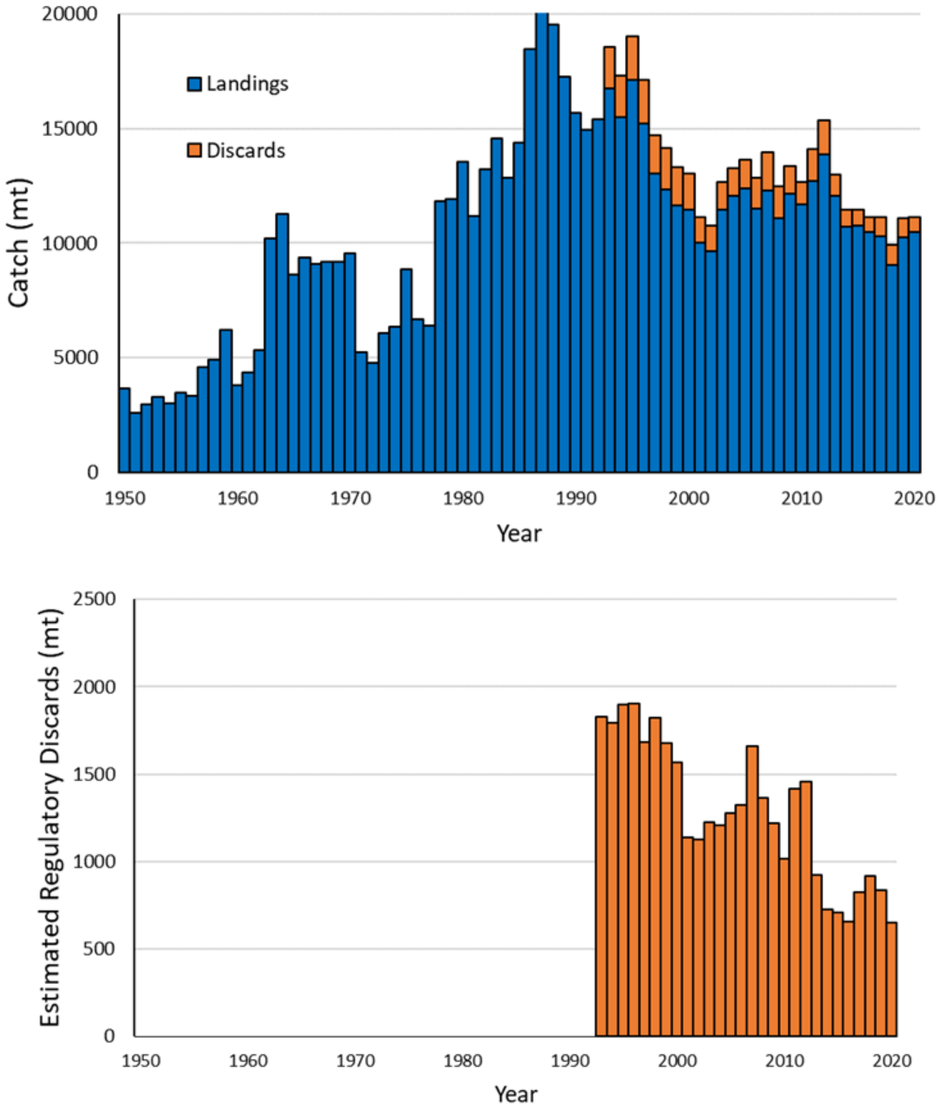
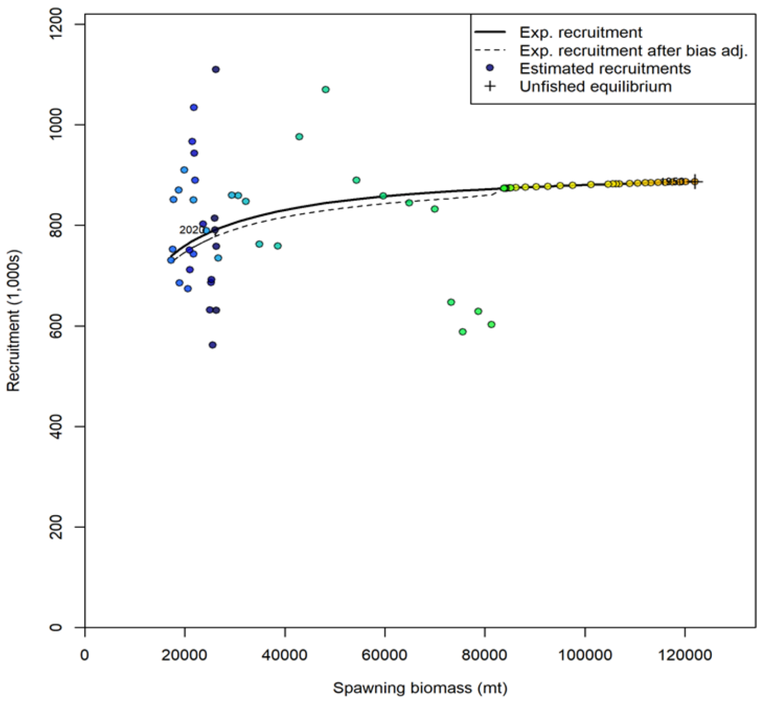



**Figure 8**. Reported (top) landings and (bottom) SS3 estimated discards of North Atlantic swordfish, 1950-2020.



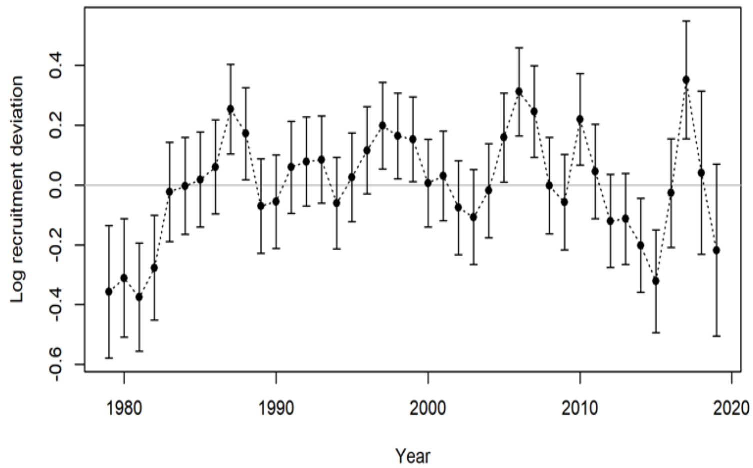




**Figure 9.** Estimated (top) stock-recruitment function and (bottom) recruitment deviations from SS3 model for North Atlantic swordfish.

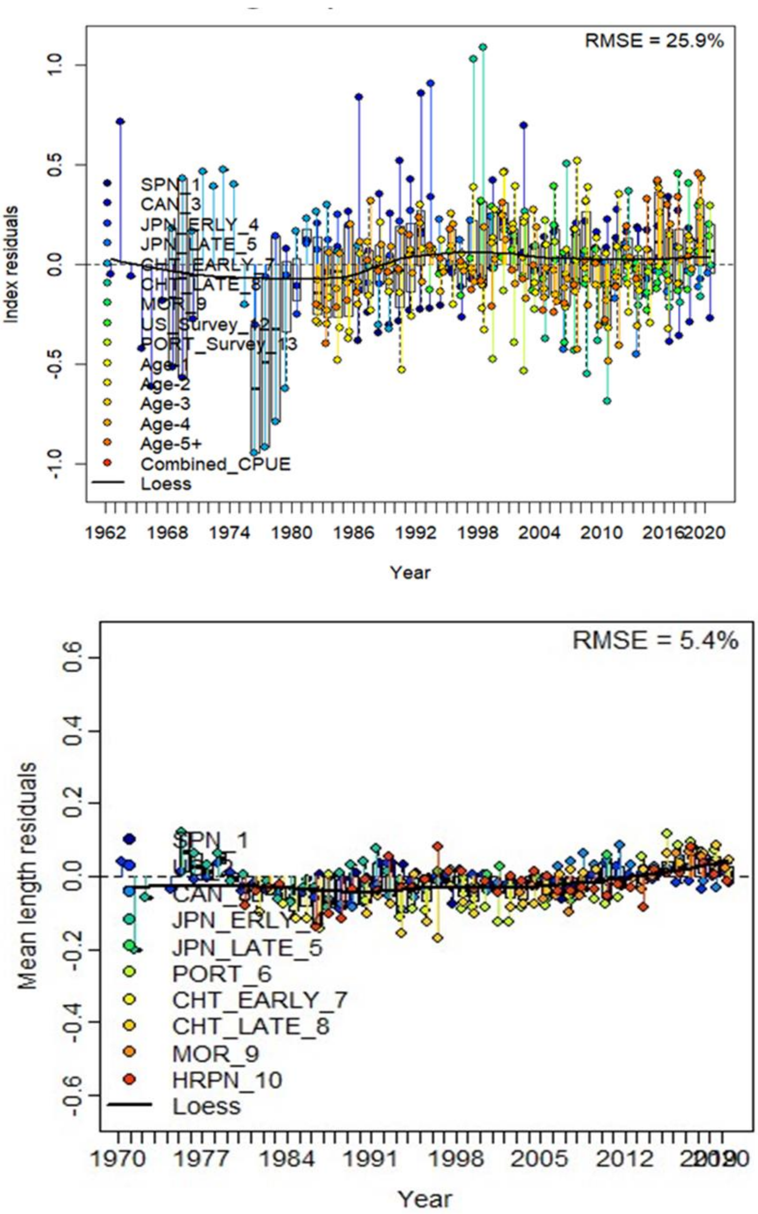



**Figure 10.** Joint residual plots for (top) multiple CPUE fits from different fishing fleets, (bottom) (c) annual mean length estimates for multiple fishing fleets from the North Atlantic swordfish.

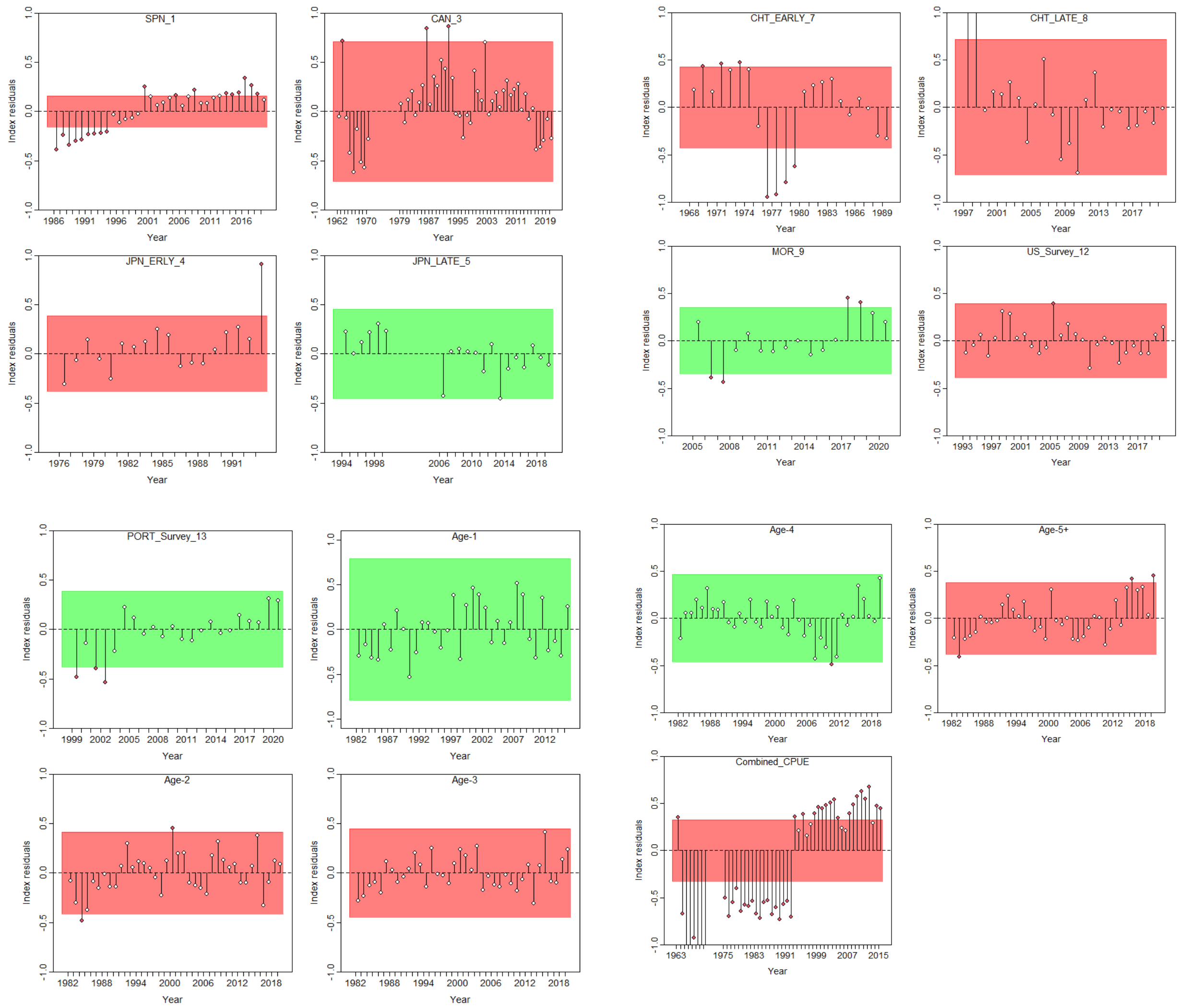



**Figure 11.** Runs test results illustrated for the indices of abundance for the SS model for North Atlantic swordfish. SPN_1 and Combined_CPUE were not used in the SS3 model.

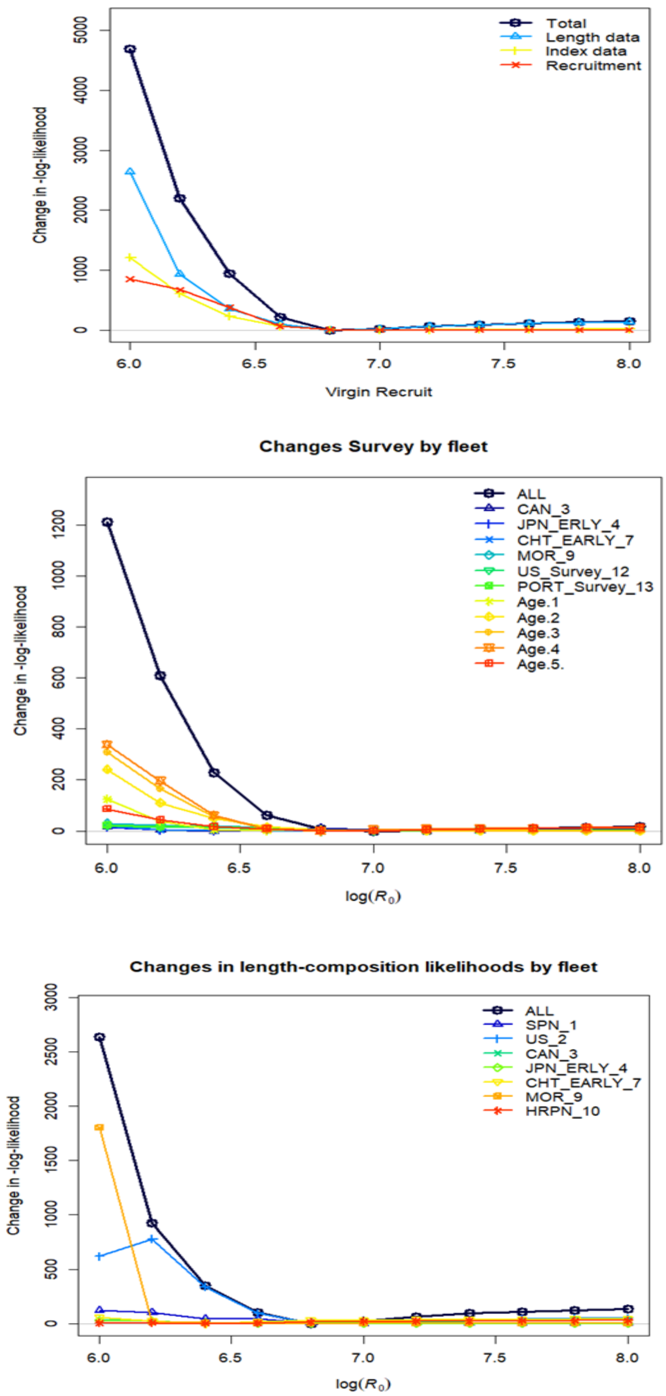



**Figure 12.** Log-likelihood profiles for R0 for the various data components included in the North Atlantic swordfish Stock Synthesis models, showing the contribution of (top) all data likelihood components, (center) among abundance indices and (bottom) among length composition data.

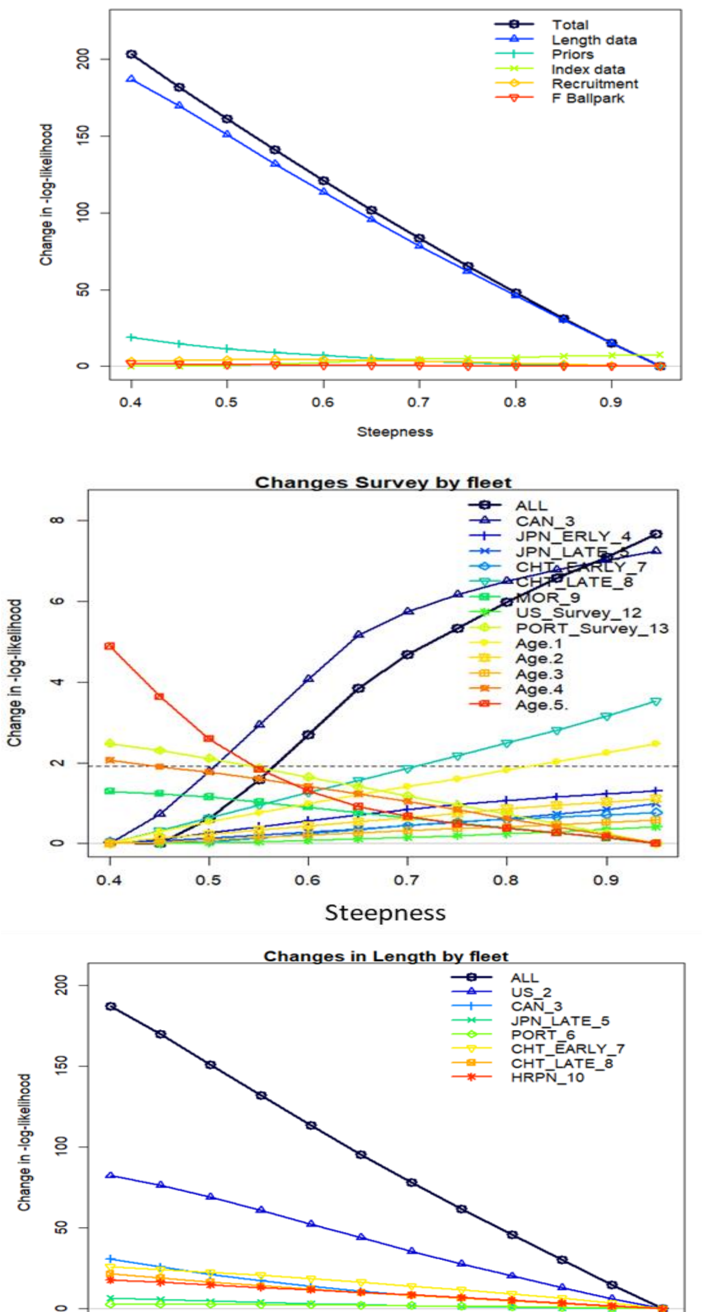



**Figure 13.** Log-likelihood profiles for steepness for the various data components included in the North Atlantic swordfish Stock Synthesis models, showing the contribution of (top) all data likelihood components, (center) among abundance indices and (bottom) among length.

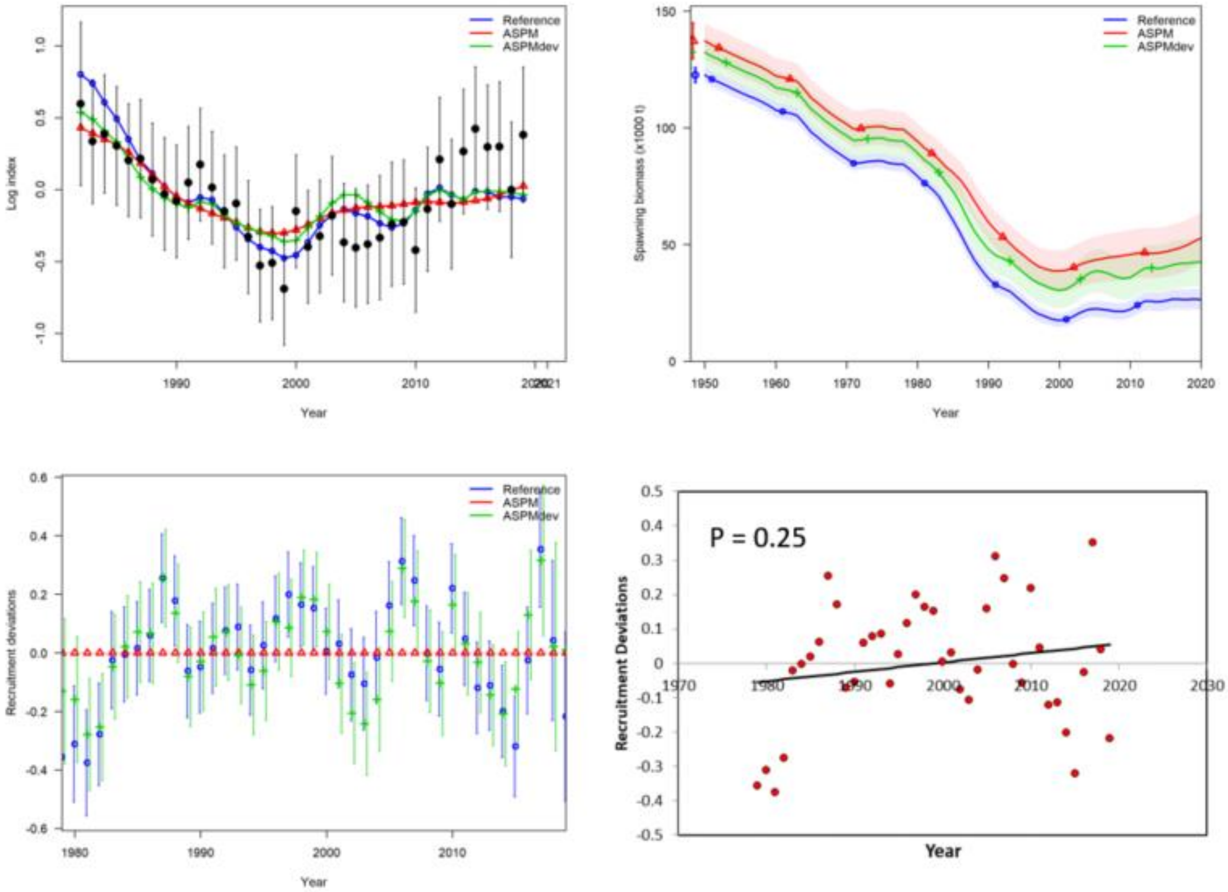



**Figure 14.** Comparison between the fully integrated base-case and the deterministic Age-Structured-Production Model (ASPM) results for North Atlantic swordfish showing observed and predicted values for (top, left) the age_4 CPUE index (top, right) spawning stock biomass trajectories ,and (bottom, left) recruitment deviation estimates. Also (bottom, right) regression of recruitment deviations vs. time showing non-significant relation.

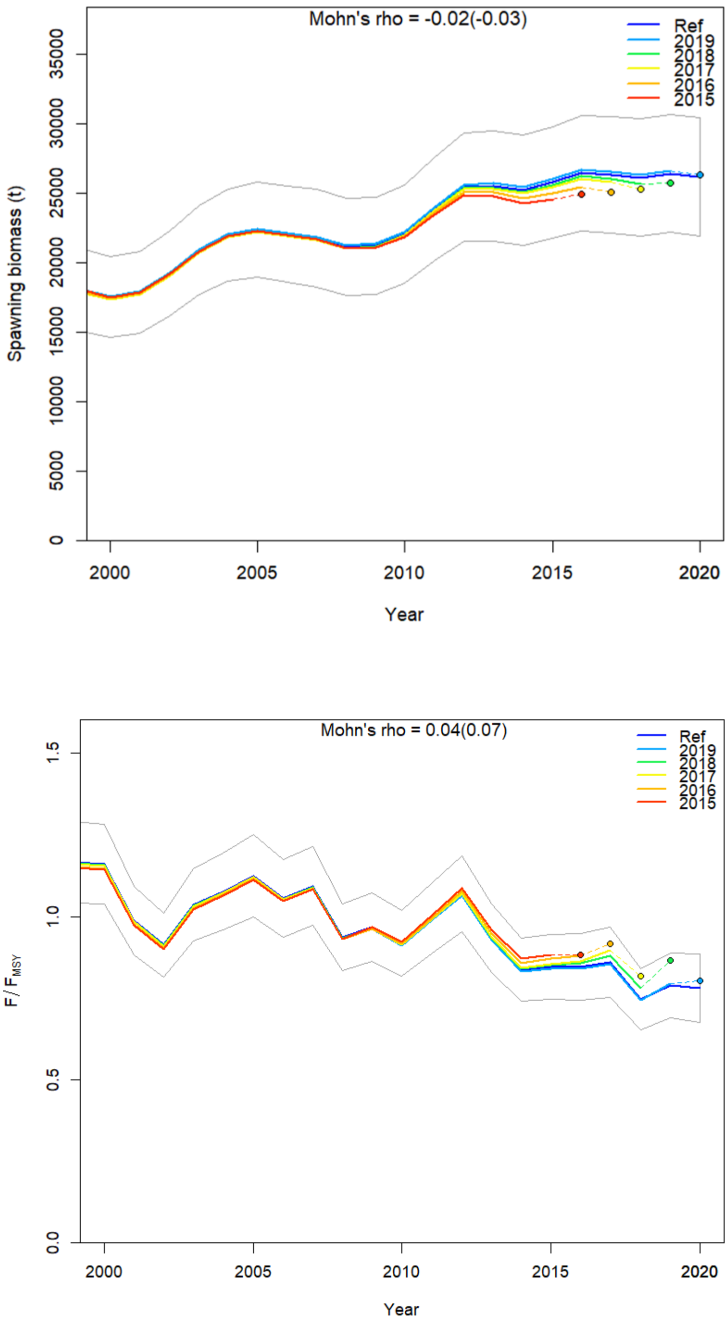



**Figure 15.** Retrospective analysis of spawning stock biomass (SSB) estimates for North Atlantic swordfish SS3 models conducted by re-fitting the reference model (Ref) after removing five years of observations, one year at a time sequentially. Mohn's rho statistic and the corresponding 'hindcast rho' values (in brackets) are printed at the top of the panels. One-year-ahead projections denoted by color-coded dashed lines with terminal points are shown for each model. Grey shaded areas are the 95 % confidence intervals from the reference model.

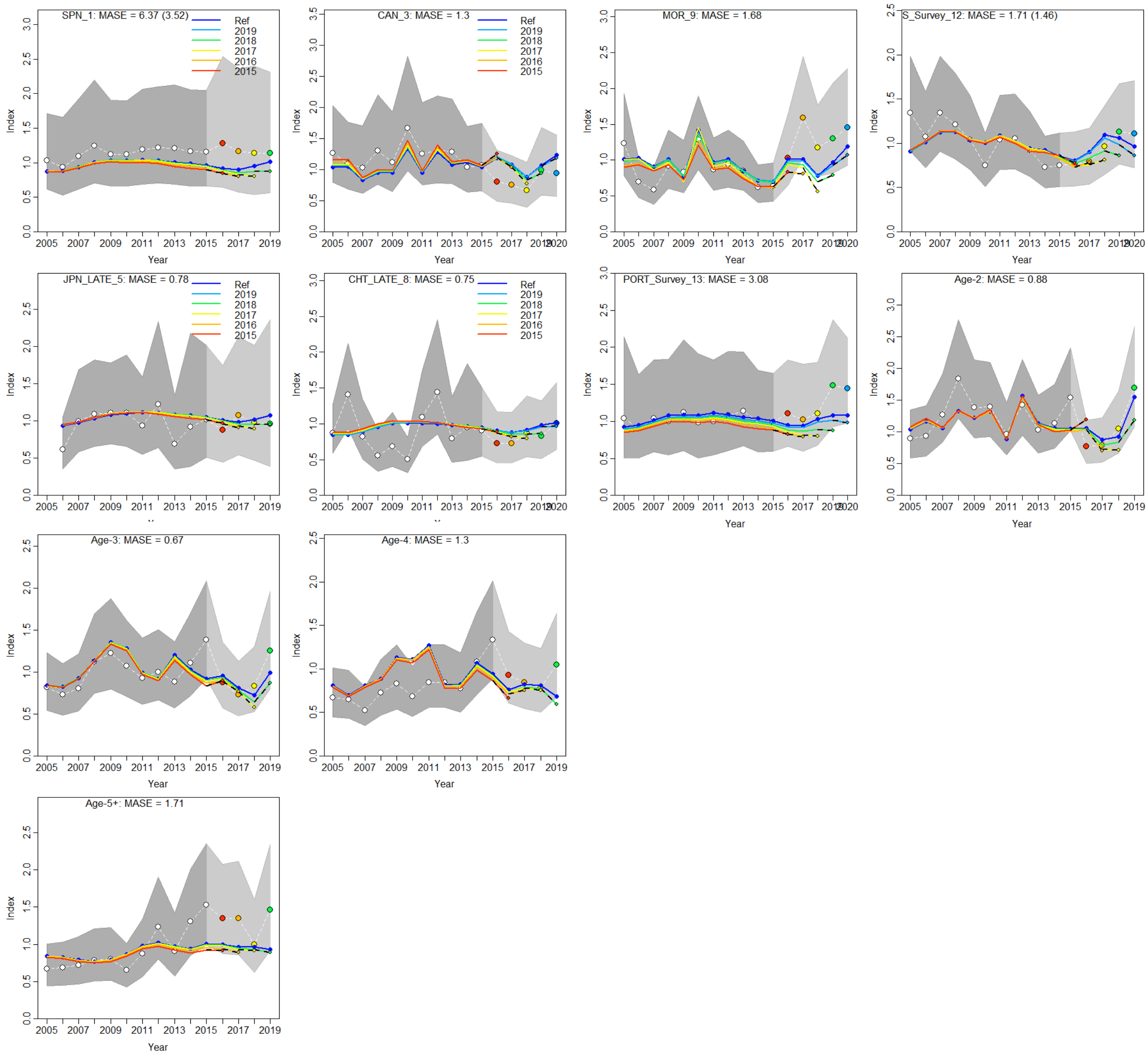



**Figure 16.** Hindcasting cross-validation (HCxval) results for three catch-per-unit-effort (CPUE) fits from the North Atlantic swordfish SS3 model, showing observed (large points connected with dashed line), fitted (solid lines) and one-year ahead forecast values (small terminal points). HCxval was performed using one reference model (Ref) and five hindcast model runs (solid lines) relative to the expected catch-per-unit effort (CPUE). The observations used for cross validation are highlighted as color-coded solid circles with associated 95 % confidence intervals (light-gray shading). The model reference year refers to the endpoints of each one-year-ahead forecast and the corresponding observation (i.e., year of peel + 1). The mean absolute scaled error (MASE) score associated with each CPUE time series is denoted in each panel. The SPN_1 CPUE was not used in the final model.

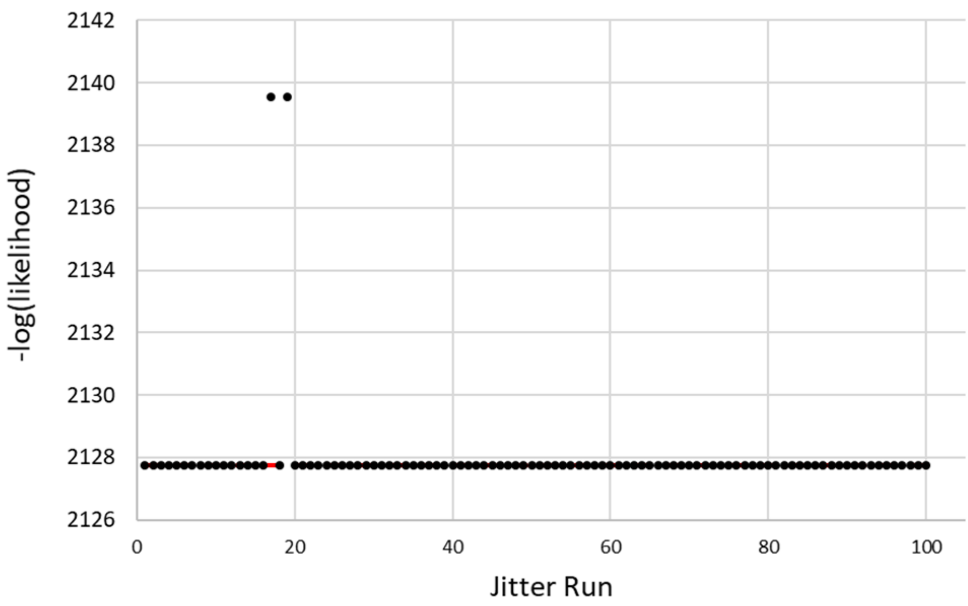



**Figure 17.** The jitter diagnostic for global convergence conducted for the North Atlantic swordfish SS3 models. On top panels (a) – (b), solid black circles represent the total likelihood obtained from jittered model runs. The red horizontal dashed line represents the total likelihood value from the base-case model.

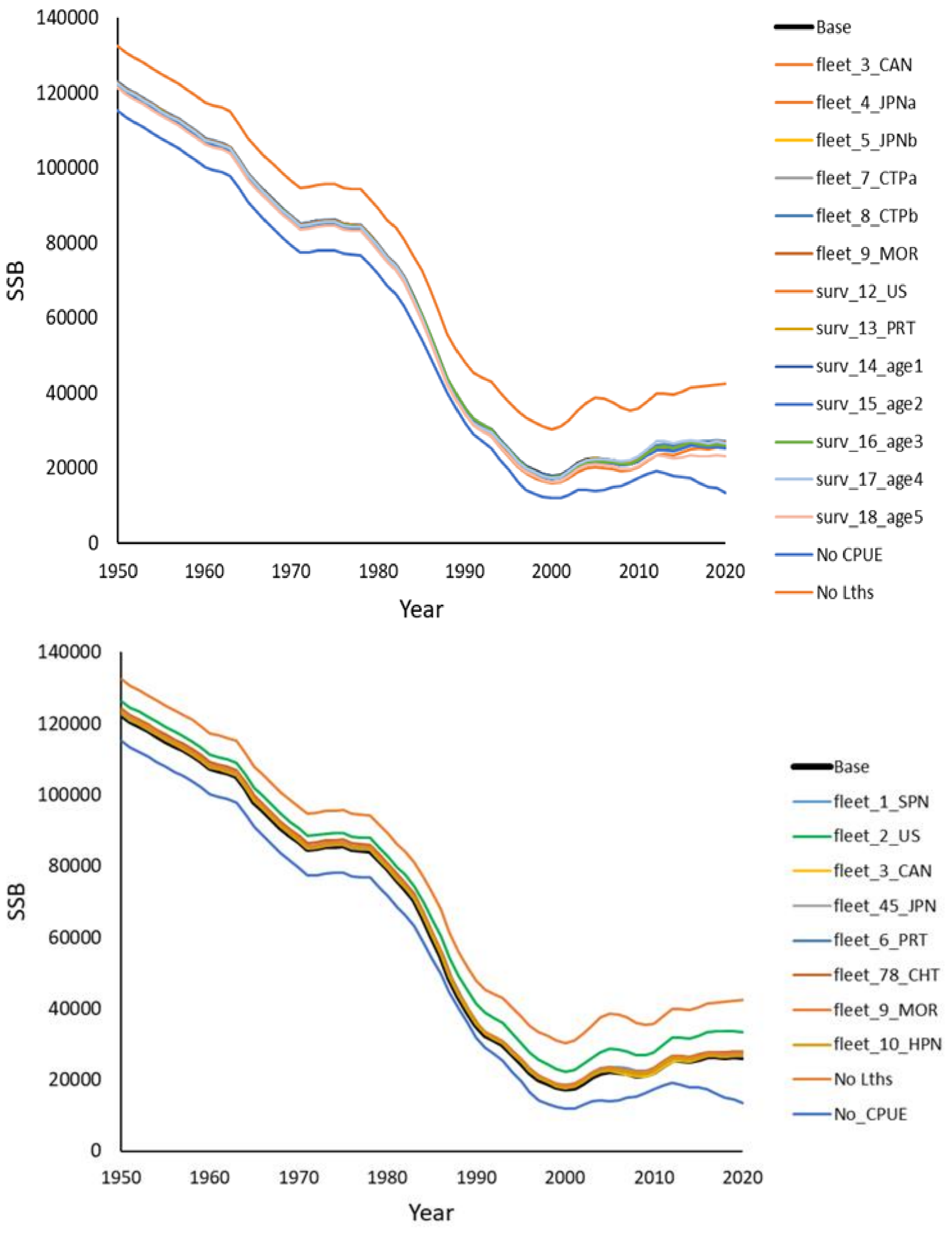



**Figure 18.** Jack-knife analysis (i.e., removal of one fleet at a time) of (top) CPUE and (bottom) length observational data for North Atlantic swordfish using the SS3 model.

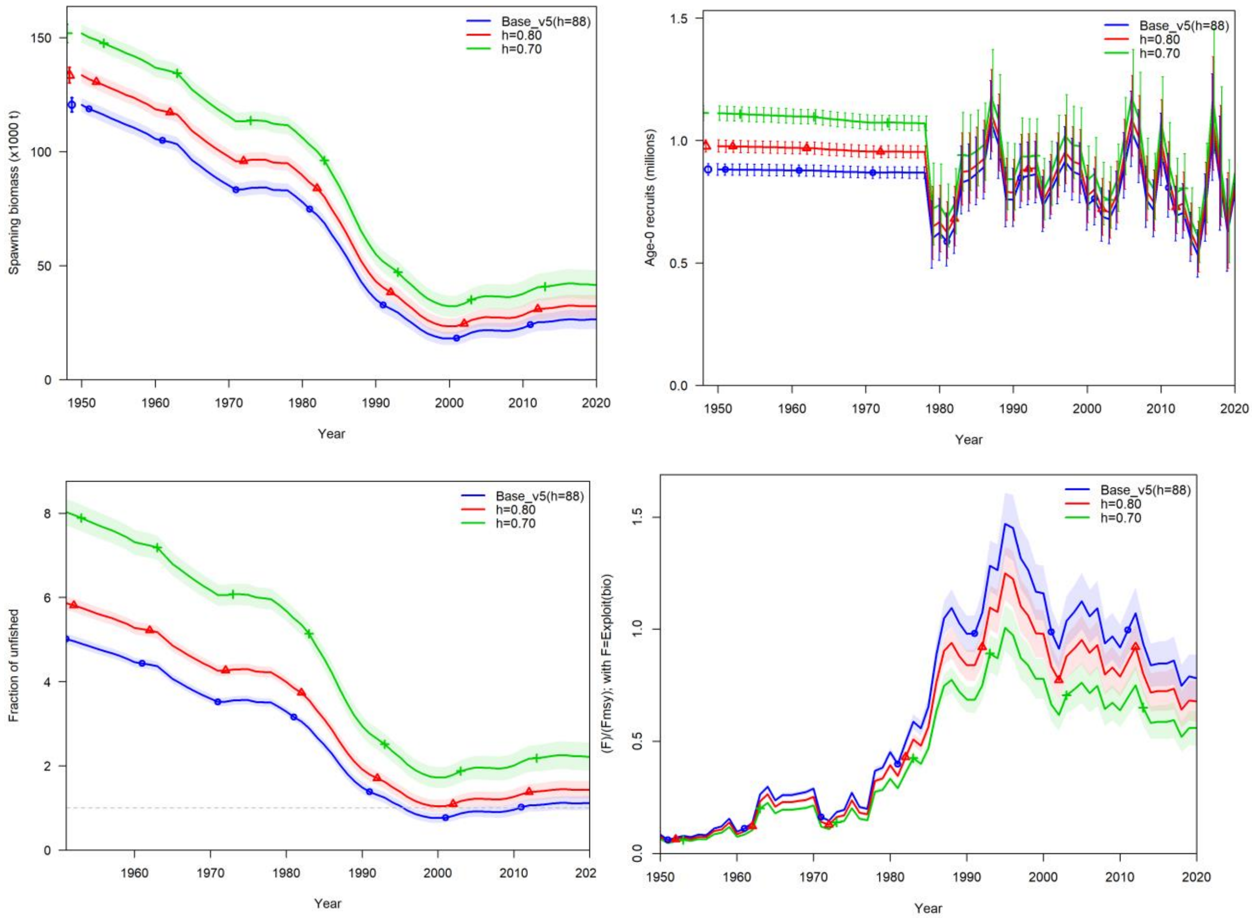



**Figure 19.** Model uncertainty as expressed via different assumptions for steepness (h =0.88, 0.8 and 0.7) for North Atlantic swordfish SS3 model (upper, left) spawning stock biomass; (upper, right) recruitment; (lower, left SSB/SSBmsy); (lower, right) F/Fmsy. Model Base_v5 used a h = 0.88.

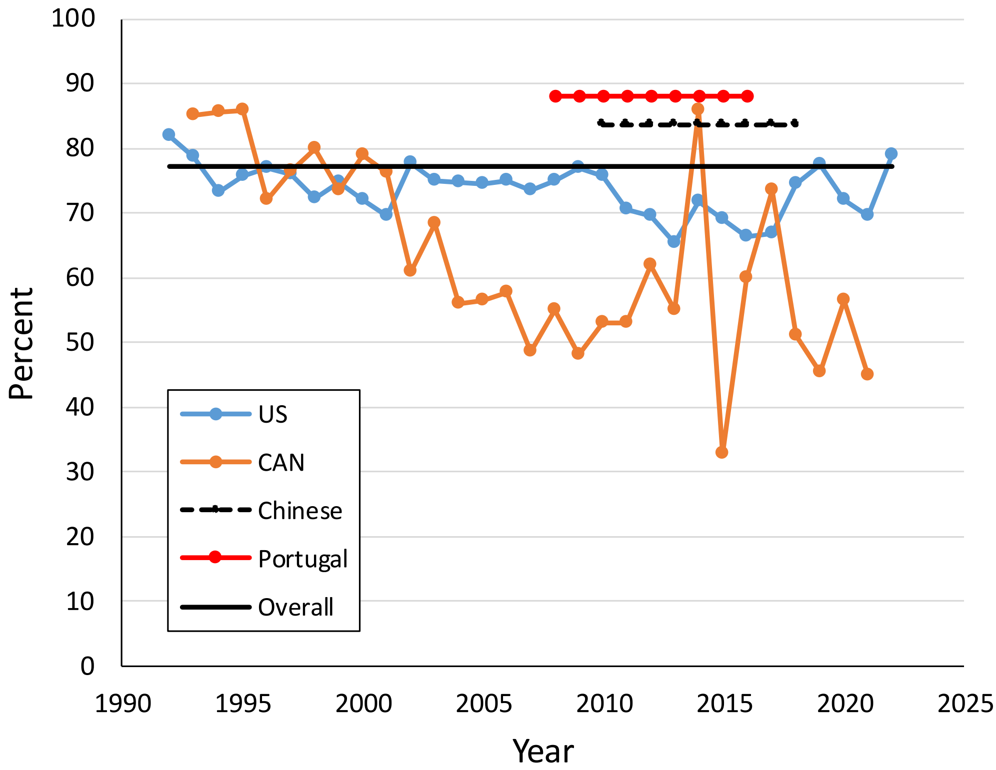



**Figure 20**. Percent of under sized North Atlantic swordfish dead at-haulback from observer data by year for from the US, Canada, and averaged over years for Chinese and Portugal. Black solid line is overall average.

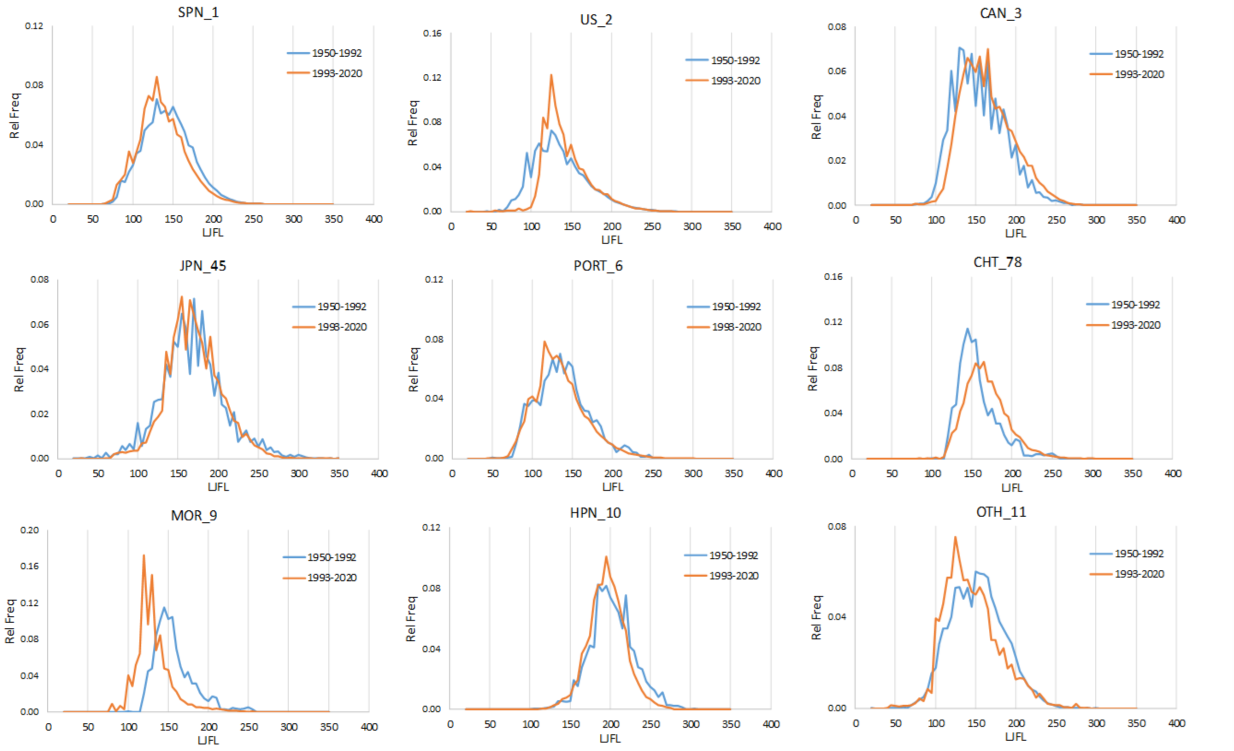



**Figure 21.** Length frequency distributions for North Atlantic swordfish by fishery combined over years pre- (1950-1992) and post- (1993-2020) minimum size regulation. The fisheries of US, Canada, Chinese-Taipei include lengths of discarded fish while the others do.

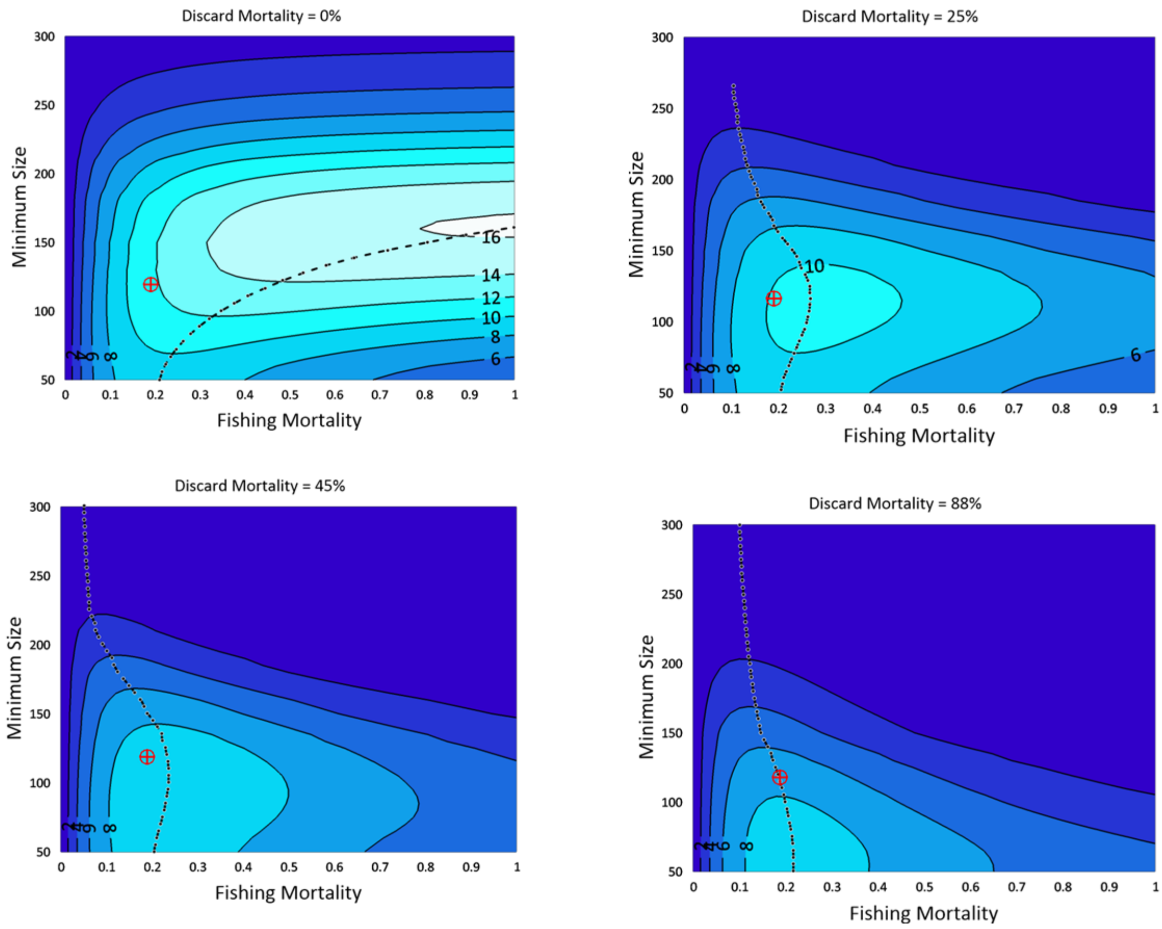



**Figure 22**. Yield-per-recruit isopleths as a function of fishing mortality and minimum size at capture for four different discard mortalities (0, 25, 45 and 88%). Dashed line is maximum yield-per-recruit; red cross is current estimate of fishing mortality and minimum size.

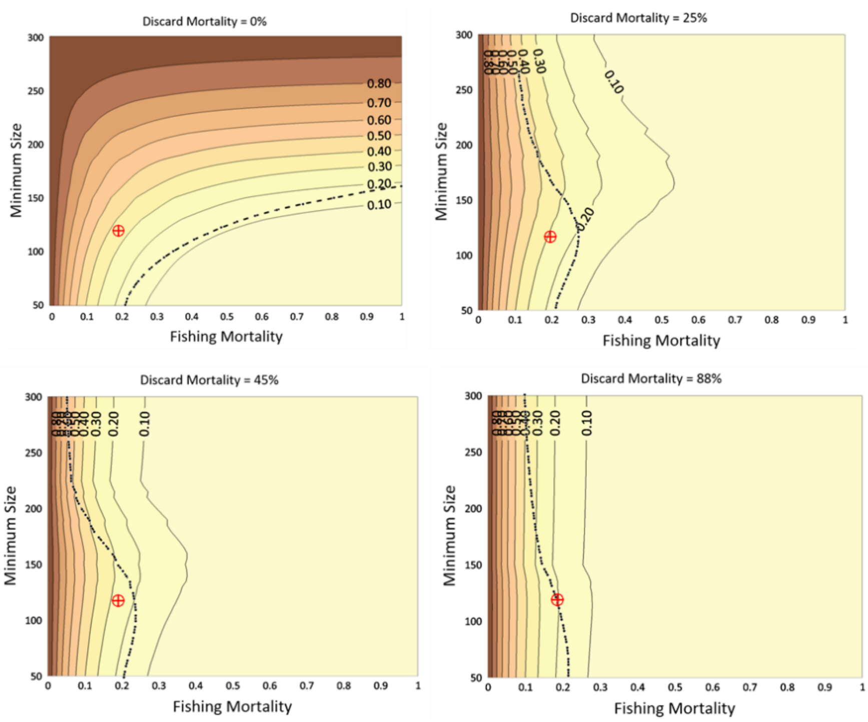



**Figure 23**.Spawning potential ratio (SPR) isopleths as a function of fishing mortality and minimum size at capture for four different discard mortalities (0, 25, 45 and 88%). Dashed line is maximum yield-per-recruit; red cross is current estimate of fishing mortality and minimum size.

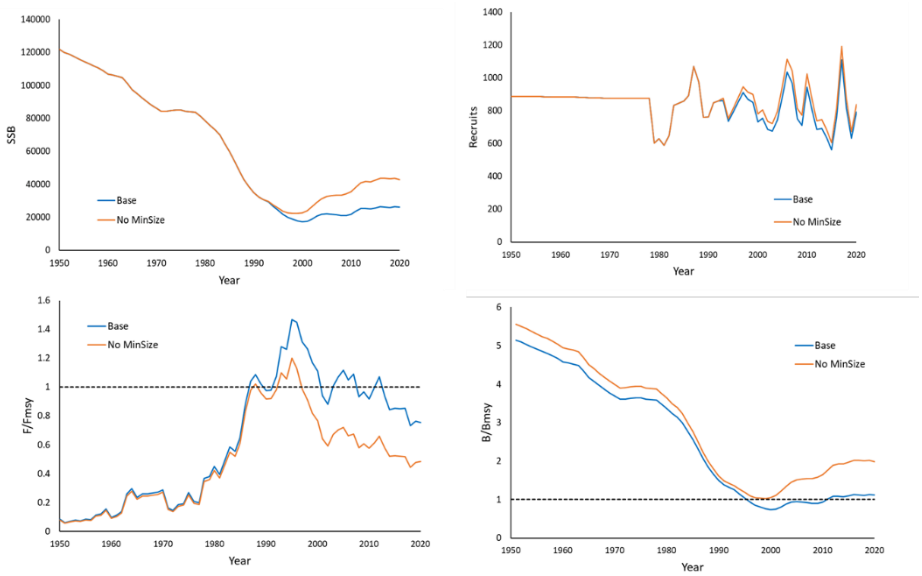



**Figure 24.** Comparison of status of North Atlantic swordfish (upper left) spawning stock biomass, (upper, right) recruitment, (lower, left) fishing mortality relative to $F_{MSY}$, and (lower, right) spawning stock biomass relative to $SSB_{MSY}$ for the SS3 Base-h88 model and a projection of a hypothetical model where no minimum size was invoked 1993. Projections make the critical assumption that the fishery would have continued to operate exactly the same with and without the use of the recommendation.