SCRS/2020/072 Collect. Vol. Sci. Pap. ICCAT, 77(2): 389-416 (2020)

WESTERN ATLANTIC BLUEFIN TUNA STOCK ASSESSMENT 1950-2018 USING STOCK SYNTHESIS: PART I. MODEL SPECIFICATION AND DATA

Y. Tsukahara[1], J. Walter[2], A. Kimoto[3] and M. Ortiz[3]

SUMMARY

This document describes a stock assessment model using Stock Synthesis (version 3.30) for the Western Atlantic population of Bluefin tuna. This document describes initial model set up, fleet definitions, selectivity and parameterizations. The model runs from 1950 to 2018 and was fit to length composition data, conditional length at age (otolith age-length pairs input as an age-length key), 13 indices and 13 fishing fleets. Growth was internally estimated in the model and natural mortality was scaled with a Lorenzen function. These input and model settings were only slightly changed from those used in 2017, commensurate with this being a strict update. Two models (early and late maturity) were used for advice in 2017 and the same are retained here. Some slight parameter changes were necessary and are documented below with relatively minor impact but improvements in model stability and fit. This paper represents the first in a series of three papers that will describe the full assessment process.

RÉSUMÉ

Ce document décrit un modèle d'évaluation des stocks utilisant Stock Synthesis (version 3.30) pour la population de thon rouge de l'Atlantique Ouest. Ce document décrit la configuration initiale du modèle, les définitions de flottille, la sélectivité et les paramétrages. Le modèle s'étend de 1950 à 2018 et s'est ajusté aux données de composition par taille, la taille conditionnelle par âge (paires d'otolithes âge-longueur saisies comme une clé âge-longueur), 13 indices et 13 flottilles de pêche. La croissance a été estimée en interne dans le modèle et la mortalité naturelle a été mise à l'échelle avec une fonction de Lorenzen. Ces configurations de données d'entrée et de modèle n'ont été que légèrement modifiées par rapport à celles utilisées en 2017, dans la mesure où il s'agit d'une mise à jour stricte. Deux modèles (maturité précoce et maturité tardive) ont été utilisés pour obtenir un avis en 2017 et les mêmes sont retenus ici. Quelques légères modifications des paramètres ont été nécessaires et sont documentées ci-dessous. Elles ont eu un impact relativement mineur mais ont amélioré la stabilité et l'ajustement du modèle. Ce document est le premier d'une série de trois documents qui décriront le processus d'évaluation complet.

RESUMEN

Este documento describe un modelo de evaluación de stock que utiliza Stock Synthesis (versión 3.30) para la población de atún rojo del Atlántico occidental. Este documento describe la configuración inicial del modelo, las definiciones de flota, la selectividad y las parametrizaciones. El modelo abarca desde 1950 hasta 2018 y se ajustó a los datos de composición por tallas, la talla por edad condicional (pares de otolitos edad-talla introducidos como clave de edad-talla), 13 índices y 13 flotas de pesca. El crecimiento se estimó internamente en el modelo y la mortalidad natural se escaló con una función Lorenzen. Estas configuraciones de entrada y del modelo sólo se modificaron ligeramente con respecto a los utilizados en 2017, en consonancia con el hecho de que se trata de una actualización estricta. Se utilizaron dos modelos (madurez temprana y tardía) para el asesoramiento en 2017 y se mantienen los mismos aquí. Fueron necesarios algunos ligeros cambios en los parámetros, que se documentan a continuación, con un impacto relativamente menor, pero con mejoras en la estabilidad y el ajuste

[1] NRIFSF. 5-7-1, Orido, Shimizu, Shizuoka, 424-8633. JAPAN. tsukahara_y@ affrc.go.jp
[2] NOAA Fisheries, Southeast Fisheries Center, 75 Virginia Beach Drive, Miami, Florida, 33149, UNITED STATES
[3] ICCAT Secretariat, C/ Corazón de María 8 – 6th floor, 28002 Madrid, Spain

del modelo. Este documento representa el primero de una serie de tres documentos que describirán el proceso de evaluación completo.

KEYWORDS

Stock assessment, bluefin tuna, Stock Synthesis

Introduction

Stock Synthesis (SS) is an integrated statistical catch-at-age model which is widely used for many stock assessments in the United States and throughout the world (Methot and Wetzel 2013 http:// https://vlab.ncep.noaa.gov/web/stock-synthesis). SS takes relatively unprocessed input data and incorporates many of the important processes (mortality, selectivity, growth, etc.) that operate in conjunction to produce observed catch, size and age composition and CPUE indices. Because many of these inputs are correlated, the concept behind SS is that they should be modeled together, which helps to ensure that uncertainties in the input data are properly accounted for in the assessment. SS is comprised of three subcomponents: 1) a population subcomponent that recreates an estimate of the numbers/biomass at age using estimates of natural mortality, growth, fecundity, etc.; 2) an observational sub-component that consists of observed (measured) quantities such as CPUE or proportion at length/age; and 3) a statistical sub-component that uses likelihoods to quantify the fit of the observations to the recreated population.

According to the terms of reference outlined in the 2019 SCRS report (Anon 2020, Appendix 5) the stock assessment in 2020 represents as strict of an update of the 2017 stock assessment as possible. The 2017 assessment terminal year of data was 2015 and for this assessment we update with data up until 2018. Model settings for biological and fisheries information are identical to the previous assessment with some minor changes and should remain the same unless there are obvious misfits between current settings and latest data. The data about catch, abundance indices, size composition and age composition data were updated to 2018. This paper describes the data, model set up and parameter settings. Two subsequent papers will describe model diagnostics and initial results and then final results and projections.

Model Specification

Overview

Overall the WBFT SS model uses size composition information, conditional age at length data (essentially an age-length key using the age-length pair data available for WBFT), 13 indices and landings going back to 1950 (Figure 1). Catch at age for the Japan longline, as derived from cohort slicing is input in the model but not used in fitting for the purposes of evaluating the predicted CAA from SS with the assumed CAA for the VPA.

Basic equations and technical specifications underlying Stock Synthesis can be found in Methot and Wetzel (2011). In these models we use both SS version 3.30.14 converted from version 3.24 used in 2017. Version 3.30 is latest software with some improvements and modifications from SS version 3.24P which was used for previous assessment in 2017. Theoretically, there is little difference between them, though version 3.30 has more advanced modeling of time dimensions, growth and improved projection capacity and is fully supported by the developers. Two base case models (early and late maturity) are constructed, similar to 2017.

The model assumed the Western Atlantic Bluefin tuna stock structure (West of 45º longitude) with no spatial structure otherwise. Fleet structure was designed to generally alias spatial/temporal structure with fleets were separated according to whether they occurred in the Gulf of Mexico or the Atlantic and when there was a clear separation in size structure due to either selectivity or availability.

The model starts in 1950 and runs to 2018 (**Figure 1**). Conditions were assumed to be near-virgin in 1950 with two fleets, USA_ TRAP and USA_CAN_HARPOON, assumed to have equilibrium catches equal to the average of 1950-1955, respectively, 434.5 and 310 t and initial Fs estimated for one of the fleets. An annual time step was assumed for the model with 13 fleets assumed to take catch out continuously over the year. Individual 13 indices were adjusted to account for the timing within the year when the index occurs.

Key settings and Input data

Biology

A single sex was assumed for the model and spawning biomass was assumed to be the summed mass of all mature fish. Fish are born at age 0 and the model uses a plus group age of 35. Maturity at age was modeled with two vectors representing either early or late spawning (**Figure 2**). Natural mortality was modeled with a Lorenzen function scaled according to the growth model with a reference M of 0.1 applied to a reference age of 20. The M of 0.1 corresponds to the Hoenig (1983) estimator of Z for a maximum age of 35. Growth was modeled with a Richards 3 parameter formulation and initially input as the Ailloud *et al* (2017) growth parameters but then all growth parameters, except for length at age 0.5 (43 cm) which was fixed, were freely estimated in the model (Linf, K, Richards parameter and the CV on young and old fish). Fecundity was modeled as proportional to weight (eggs=a*Wt^b) and the overall Western Atlantic length weight relationship was used to convert size to weight (1.52E 05* length^3.05305). Biological vectors input or initial value for estimation in SS (italics) are shown below:

Age	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	…	35
early spawning	0	0	0	0.25	0.5	1	1	1	1	1	1	1	1	1	1	…	1
late spawning	0	0	0	0	0	0	0	0.01	0.04	0.19	0.56	0.88	0.98	1	1	…	1
M (Lorenzen scaled)	*0.40*	*0.33*	*0.27*	*0.23*	*0.20*	*0.18*	*0.16*	*0.14*	*0.13*	*0.12*	*0.12*	*0.11*	*0.11*	*0.11*	*0.11*	…	0.10
growth (mid year size)	43	*58*	*75*	*93*	*113*	*133*	*152*	*170*	*186*	*200*	*212*	*222*	*231*	*238*	*243*	….	*266*

Stock-recruitment relationship.

A Beverton-Holt stock recruit relationship was assumed and spawning biomass was equal to the biomass of the mature population according to the two maturity vectors outlined in the biology section. Parameters of the stock recruitment relationship (steepness and R0) were freely estimated. The variance in interannual recruitment deviations (sigmaR) was estimated between a range of 0.2 to 2 using the Method and Taylor bias correction ramping to facilitate estimability.

Deviations from the stock-recruitment relationship were assumed to follow a lognormal distribution estimated on a logscale as N(0, sigmaR) variates with a min and max of -5 and 5, respectively. Zero recruitment deviations were assumed until the start of informative data on age structure, i.e. annual deviates were only estimated from 1961-2018. The lognormal bias correction ($-0.5\sigma^2$) for the mean of the stock recruit relationship was applied during the period 1961-2017 with a bias correction ramp applied prior to 1961 and after 2011 according to the Methot and Taylor (2011) recommended bias correction ramping. This bias correction ramping was updated for the 2020 models and recruitment deviations extended to 2017.

Fleet and index definitions

Fleet definition for catch and index fleet is exact same as previous assessment.

Overall the model consists of 13 fleets (**Table 1**):

1. JAPAN_LL 2. USA_CAN_PSFS 3. USA_CAN_PSFB 4. USA_TRAP 5. USA_CAN_HARPOON
6. USA_RRFB 7. USA_RRFS 8. OTHER_ATL_LL 9. CAN_HOOKLINE 10. GOM_LL_US_MEX
11. JLL_GOM 12. CAN_TRAP 13. CAN_GSL1

and 13 indices, though two (tagging and the oceanographic index were not used)
1. JAPAN_LL 2. US_RR_66_114 3. US_RR_115_144 4. US_RR_GT145 5. US_RR_GT177
6. US_RR_GT195 7. USPLL_GOM 8. JLL_GOM 9. CAN_NS 10. GOM larval
11. tagging (not used) 12. CAN_ACOUSTIC 13. oceanographic (not used)
14. JAPAN_LL2 15. USPLL_GOM_LL2

Total catch (task I)

The total catches were calculated by the Secretariat (**Table 2, Figure 3**) with some modifications as noted to the fleets, above. Catch in metric tons was used in the model for all fleets, and was assumed to be known essentially without error (standard error =0.05). Initial equilibrium catch was input for USA_trap and USA_CAN_Harpoon that had non-negligible catches in 1950. Initial F was estimated for these fleets but was assumed to be zero for all other fleets. To provide initial equilibrium catches for USA_TRAP and USA_CAN_HARPOON the average for 1950-1955 was input (434.5 and 310 t, respectively). The initial F for the USA_CAN_Harpoon fleet was fixed at 0.0014762, an estimate from previous model runs but necessary to fix as it often hit a minimum bound.

Catch per unit effort data

While retained in the data file the SS models exclude the Gulf of Mexico oceanographic index and the historic tagging index from likelihood component. All indices were input with a CV of 0.2 for each year (input as a log scale standard error in model). This decision was similar to the decisions made for the VPA and other models. CPUE indices were assumed to have a lognormal error structure. No time blocks on indices were modeled as indices that required splits were input as separate indices with unique catchabilities, while catchability for three indices, US_RR_GT177, i.e. CAN_GSLNS and CAN_ACOUSTIC, were linked with the Atlantic multidecadal oscillation (AMO) for July, August and September as an environmental factor (see SCRS 2020/. CPUE input data are shown here (**Figure 4**) but fits to CPUE data will be shown in the second paper that documents preliminary Results.

Conditional age at length inputs

Otolith age-length data was available from the same five labs that provided data in 2017, with substantial additional numbers of age-length pairs available (**Table 3**). Much of the data has gone through extensive re-evaluation and scrutiny of aging protocols (SCRS-2019-132) resulting in updates to several of the datasets used in 2017.

Consistent with the nature of this assessment as an update we include age data from 2016-2018 (terminal year of the model) and also to include the historical data from the years that it was originally used in the 2017 assessment. Data from University of Maryland Center for Environmental Sciences (UMCES) remained unchanged from the 2017 models.

The data was screened for outlier length-weight pairs by noting observations +/- 3 empirical standard deviations from the mean size at age. In many cases these were due to length conversions from different units and could be corrected in the original files. The remaining outliers that could not be confidently identified as being due to size conversion errors were removed from the age dataset (**Figure 5**).

Similar to the treatment of the data in 2017, when gear types were not recorded expert opinion was necessary to assign gear based on landing port and these remain the same fleets as in 2017. In the Panama City dataset, a number of small fish without gear were assumed to be USA_RRFS as the samples likely came from the Large Pelagics Biological Survey that generally surveys the US recreational fleet.

This process of updating the years of data from 2016-2018 and replacing the previously used samples with the revised age reads resulted in a similar dataset as in 2017 but with additional years of data (**Table 1**). It did result in removing a substantial amount of new ageing data from the time period 1973-83 from Canada DFO and from University of Maine from 2012-2015 but this would have been data not used in 2017. The total number of age-length pairs available were 9307 from years 1973-2018 with 6552 remaining following screening and following the strict update protocols.

Age-length data was assigned to 9 different fleets (**Figure 6**). Age information was input with an aging error vector assuming a CV of approximately 0.1 for most ages (SCRS/2014/038). In 2017 an aging bias vector derived from paired otolith-spine samples was used. However, a review of aging protocols (SCRS/2019/132) indicated that some of this bias may have been due to the previous assumption regarding the timing of opaque band formation. A revised adjustment criterion was proposed to convert the count of bands into ages and all historical reads (except the UMCES samples) were revised accordingly, obviating the need to input a bias in the aging vector. Hence only a vector of aging error was input to the update models.

age class	Age	standard error	CV
0	0.5	0.14	0.28
1	1.5	0.41	0.27
2	2.5	0.54	0.22
3	3.5	0.62	0.18
4	4.5	0.73	0.16
5	5.5	0.75	0.14
6	6.5	0.89	0.14
7	7.5	1.07	0.14
8	8.5	1.09	0.13
9	9.5	1.14	0.12
10	10.5	1.22	0.12
11	11.5	1.34	0.12
12	12.5	1.52	0.12
13	13.5	1.85	0.14
14	14.5	2.04	0.14
15	15.5	1.76	0.11
16	16.5	1.66	0.10
17	17.5	1.44	0.08
18	18.5	1.53	0.08
19	19.5	2.2	0.11
20	20.5	2.31	0.11
21	21.5	2.42	0.11
22	22.5	2.54	0.11
23	23.5	2.65	0.11
24	24.5	2.76	0.11
25	25.5	2.87	0.11
26	26.5	2.99	0.11
27	27.5	3.10	0.11
28	28.5	3.21	0.11
29	29.5	3.32	0.11
30	30.5	3.44	0.11
31	31.5	3.55	0.11
32	32.5	3.66	0.11
33	33.5	3.77	0.11
34	34.5	3.89	0.11

Ages were adjusted according to SCRS/2019/132. An additional adjustment to the ages for input to Stock Synthesis was to subtract one half of a year to the age to account for the assumed (within SS) January 1st birthdate so that SS correctly tracks cohorts. Age data was input as conditional age at length data (similar to an age-length key) where the main assumption is that the ages are randomly collected within a length bin e.g., within a 5 cm length bin all the samples a random sample. We also show histograms of the age composition by year for visual purposes (**Figure 7**). As the sampling is not representative of all fleets, across all sizes this greatly relaxes the assumption of random sampling across all size classes for a fishery.

Catch at age input

Similar to the 2017 model, catch at age was input for the Japan longline fleet which did not have conditional age at length data. Catch at age data was not fit in the likelihood component but was input for diagnostic purposes to evaluate the consistency of decisions used to construct the CAA with internal modeling of growth and selectivity in SS. Catch at age was only updated to 2015 as it was not necessary to include later years simply for diagnostic purposes.

Size frequency information

Development of the raw size frequency input to SS followed the same process as in 2017 (*SCRS/2017/166*). Some data cleaning was conducted (removing outliers due to extreme skewness, kurtosis, or extremely small or large sizes for particular fleets) but the size composition information was used in its most raw format as provided by individual CPCs (**Figures 8 & 9**). In the 2017 model the outliers were not removed from the dataset and can be seen in several of the residual plots. In 2020 these outliers were removed which cleans up many of the Pearson residual plots allowing for the central patterns to emerge. Data was input is straight fork length in centimeters and modeled with 5 cm length bins between 30 and 350 cm in the model.

Size frequencies for the remainder of the 12 fleets indicate relatively consistent size structure over time with the exception of several fleets with sparse data (**Figures 10, 11**). Length composition data is modeled assuming a multinomial distribution.

Selectivity

Selectivity was parameterized (**Table 1, Figures 12, 13**) as length-based for most fleets/surveys as either 6 parameter double normal which could take on either dome or asymptotic shape or as logistic on the basis of visual examination of the length composition data. There was time block selectivity for two fleet, Japan_LL and USA_RRFS. For the Japan LL time varying selectivity with deviation was assumed from 2011 to 2015 to be aligned with the target change of this fishery because of fishery management. The selectivity deviations were not expanded for 2016-2018. Several surveys had a special selectivity parameterization with the larval survey assumed to have selectivity of the GOM_LL_US_MEX index and fishery. The oceanographic index was excluded in the likelihood component but was retained to evaluate the potential fit and was modeled with a selectivity equal to exp(rec devs). For the US_RR_66_114 and US_RR_115_144 indices, selectivities were assumed to be double normal and fixed to reflect constant selectivity between the size classes that the index is designed to represent e.g. 66-114 cm and 115-144 cm, respectively. In several cases when the double normal selectivity showed either a steady increasing or decreasing limb these were modeled to allow for either a smooth increase or decrease to avoid sharp and unrealistic breaks. For four selectivity parameters in three fleets, i.e. ascending (P3) for USA_CAN_PSFS, top (P2) and descending (P4) for USA_TRAP and ascending (P3) for USA_CAN_PSFB, a symmetric beta prior was used with a mean of 0.5, 0.5, -4 and 1.2, respectively, and a standard deviation of 0.1 to avoid the model hitting bounds on this parameter.

Data weighting

Francis and Hilborn (2011) indicates that often in complex integrated models there is conflicting sources of information, stemming from fitting to either the length composition data, or abundance index data and often the numerically abundant length composition information dominates the likelihood. Length composition data was initially input with a sample size of 100 and conditional age at length data was input with the actual sample size. In most cases, the effective N was much higher than the input N indicating that that the effective sample should be reduced for most fleets. Input sample size for length and age data input was iteratively adjusted so that the harmonic mean effective N equaled the input N using variance adjustments (McAllister and Iannelli 1997). Input weights, as follow, generally substantially downweighted the length composition as well as the conditional length at age data. Age composition data input for the Japan_LL was not fit in the model likelihood and removed using the lambda emphasis factors. The iterative reweighting of the models was repeated for the 2020 update models but only for one maturity run and the same weights used for the other run.

No adjustment to index weighting was performed in the current iterations of the models.

Converting from SS 3.24 to 3.30.

In 2017 SS version 3.24 was used. The most recent version of SS (3.30) more precise control in modeling with many new features. The latest executable is desirable to use for stock assessment provided that it can give a comparable result. Therefore, the result by SS 3.24 and SS 3.30 are compared to examine the performance comparison. The conversion was conducted by ss_tran.exe developed by the SS developers before the data was updated. The conversion proceeded as follows:

Step	model	description	LL	diff	SSB1950
1	WBFT 12	base model, late maturity	5413.97		176229
2	WBFT12_pre conversion	Take off devs to JapanLL selectivity, change F method to 2	5436.63	22.66	176107
3	WBFT12_post conversion	convert to 3.30	5445.80	31.83	176443q
4	post conversion no Q priors	Remove priors on env Q	5430.69	16.72	176241
5	WBFT12_3.30 devs added	add devs to JapanLL selectivity, change F method to 3	5407.04	-6.93	176188

The converted model was then updated with the new data and reweighted using the McAllister and Iannelli (1997) composition weighting method. The process was not done for Run 13 (low age at maturity) due to its nearly identical diagnostic behavior as Run 12. SS 3.30 handles time varying deviations differently than in 3.24. In 3.24 the deviations in annual selectivity were penalized with a deviation vector. This vector was multiplied by the parameter estimate to and a prespecified input standard error of 0.2 was used. SS version 3.30 has much greater control over parameter deviations, even allowing them to be estimated parameters, if desired. Hence the configuration of the deviations in annual peak selectivity had to be reconfigured to reflect SS 3.30 input specifications. In 3.30 the deviation parameter is now multiplied by the standard error parameter, rather than deviations being penalized according to a specified standard error (the approach in v.3.24) necessitating a change in the value of the input standard error. To mimic the similar degree of change in 3.30, a range of standard errors were explored to obtain similar effects as in Run 12 with a value of 50 for the standard error of the parameter deviations giving similar performance, however it resulted in slight differences in the log-likelihoods for the deviations of about 3 log likelihood units, contributing to half of the difference between the 2017 model in 3.24 and 3.30. Overall, given that the WBFT model has time-varying selectivities that have received extensive updating in SS 3.30 the very small difference in log-likelihood and the nearly identical trends and absolute magnitude in SSB and recruitment (**Figure 14**) indicate quite close agreement for the transition to 3.30.

Changes to 2017 models

A series of minor changes to the models were as follows:

1. The time block on USA_RRFS as switch to between 1950-1992 and 1993-future to allow the change in selectivity to extend past the additional years of the model.
2. Bounds on many of the parameters were reduced to facilitate the way that SS 3.30 conducts jittering using the min and max bounds rather than as a percent deviation from the parameter starting value.
3. Change parameter standard error settings for the time varying selectivity deviations from 0.2 to 50. This preserves the same specification of the degree of allowable change but with the new method that 3.30 uses for deviations. See "Converting from SS 3.24 to 3.30" for details.
4. Removal of the aging bias vector under the assumption that the revised aging protocols corrected the previous aging bias of otoliths relative to spines.
5. Beta priors input for the 3 catchability parameters to keep the model from hitting min/max bounds.
6. Normal priors for several selectivity parameters to aid in model convergence

Model Diagnostics

Model convergence was assessed using several means.

1. The first diagnostic was whether the Hessian, (i.e., the matrix of second derivatives of the likelihood with respect to the parameters) inverts.
2. The second measure is the maximum gradient component which, ideally, should be low. The third diagnostic was a jitter analysis of parameter starting values to evaluate whether the model has converged to a global solution, rather than a local minimum. Starting values of all estimated parameters were

randomly perturbed according to a normal distribution defined where the pr(par min)=0.01 and pr(par max)=0.99). This is different than in SS 3.24 where the jitters were defined by a user-specified variation around the input parameter starting value. This necessitated changes to the input min/max values to accommodate this change.

3. Parameter coefficients of variation where the CV of the parameter estimate comes from the model estimated variance from the variance-covariance matrix
4. Likelihood profiles were completed for three key model parameters: steepness of the stock-recruit relationship (*h*) and the log of unexploited equilibrium recruitment (R_0) and sigma R. Likelihood profiles elucidate conflicting information among various data sources, determine asymmetry around the likelihood surface surrounding point estimates and evaluate the precision of parameter estimation.
5. Evaluation of fits to residuals for indices and length composition,
6. Retrospective analyses. Retrospective analyses are also standard diagnostic practice and were conducted on models 1-2 with 5 year retrospective peels.
7. Sensitivity to different indices (index jackknife evaluation)

Another model diagnostic is parametric bootstrapping. Uncertainty in parameter estimates and derived quantities can as well bias between the maximum likelihood estimates and estimates obtained by bootstrapping were investigated using a parametric bootstrap approach. Bootstrapping is a standard technique used to estimate confidence intervals for model parameters or other quantities of interest and was used in 2017 to generate the kobe matrix. There is a built-in option to create bootstrapped data-sets using SS. This feature performs a parametric bootstrap using the error assumptions and sample sizes from the input data to generate new observations about the fitted model expectations. The model was refit to approximately 100 bootstrapped data-sets and the distribution of the parameter estimates was used to represent the uncertainty in the parameters and derived quantities of interest.

Parameters Estimated

Overall 110 parameters are estimated in the model, consisting of 7 growth parameters 1 initial F parameter, 31 selectivity parameters, 6 catchability, 5 deviations, 3 stock recruitment parameters and 57 recruitment deviations. Several selectivity and catchability parameters were input with Bayesian priors to aid model stability.

Benchmark and fishing mortality calculations

For overall fishing mortality rate, an F0.1 proxy calculated from the yield per recruit curve was used in 2017 and will also be used here. Given the substantial changes in overall selectivity over time the F01 and benchmarks will be estimated on a year-specific basis according to the fle*et al*location in that year. Fishing mortality will be calculated as the average true (instantaneous) F over ages 10-20.

Uncertainty Quantification

In 2017 uncertainty in parameter estimates was quantified by computing asymptotic standard errors for each parameter. Asymptotic standard errors are calculated by inverting the Hessian matrix after the model fitting process.

For construction of the K2SM parametric bootstrapping using 500 bootstraps of each model was conducted. Comparisons of parametric bootstrapping with the multivariate lognormal approximation approach (Winker *et al*., 2019) indicate little benefit from the added time involved in bootstrapping with greatly increased times to produce the K2SM and the BFT WG may want to consider using the approximation approach used for yellowfin tuna in 2019.

Results

Results of diagnostic evaluations will be shown in presentations during the meeting, pending final approval of the data inputs and initial modeling decisions. They will be included in the second paper of this three-part series. Overall the model set up and basic data inputs reflect very minimal changes from the 2017 models. The primary change being additional years of data and the changes to the aging conventions for the age data and removal of the ageing bias vector.

Acknowledgements

We are grateful for the contributions of the Secretariat, the combined efforts of a vast number of colleagues who, through GBYP, have resurrected age-length information, catches and other historical data that facilitate creating an integrated model that uses as much data as possible for WBFT.

References

Ailloud, L. E., M. V. Lauretta, A. R. Hanke, W. J. Golet, R. J. Allman, M. Siskey, D. H. Secor, J. M. Hoenig. 2017. Improving growth estimates for Western Atlantic bluefin tuna using an integrated modeling approach. Fisheries Research 191: 17–24.

Busawon, D. S.; E. Rodriguez-Marin, *et al*. 2014. Evaluation of an Atlantic Bluefin Tuna Otolith Reference Collection. SCRS-2014-038.

Francis, R.C., and Hilborn, R. 2011. Data weighting in statistical fisheries stock assessment models. Canadian Journal of Fisheries and Aquatic Sciences **68**(6): 1124–1138. NRC Research Press

Mather, F.J., J. M. Mason A. C. Jones. 1995. Historical document: life history and fisheries of Atlantic Bluefin tuna. NOAA technical Memorandum NMFS-SEFSC 370.

McAllister M. K., Ianelli J. N. 1997. Bayesian stock assessment using catch-age data and the sampling-importance resampling algorithm, Canadian Journal of Fisheries and Aquatic Sciences, 1997, vol. 54 (pg. 284-300)

Methot, R.D. 2000. Technical description of the stock synthesis assessment program. U.S. Dept. Commer., NOAA Tech. Memo. NMFS-NWFSC-43, 46 p.

Methot, R.D. and Taylor, R.G. (2011) Adjusting for bias due to variability of estimated recruitments in fishery assessment models. Canadian Journal of Fisheries and Aquatic Sciences 68:1744-1760.

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.

Porch, C. and A. Hanke. 2017. Estimating the fraction of western Atlantic bluefin tuna that spawn by age from size frequency data collected on the Gulf of Mexico spawning grounds. SCRS/2017/164

Sakagawa, G. 1975. The Purse-Seine Fishery tor Bluefin Tuna in the Northwestern Atlantic Ocean. Marine Fisheries Review. 37(3):

Schirripa, MJ. F. Abascal, I. Andrushchenko, G. Diaz, J. Mejuto, M. Ortiz, M.N. Santos, J. Walter. 2016. A hypothesis of a redistribution of North Atlantic swordfish based on changing ocean conditions. Deep Sea Research Part II: Topical Studies in Oceanography. 130.

Winker, H., Kell, L.T., Fu, D, Sharma, R., Courtney D., Carvalho F., Schirripa M and Walter, J. 2019. A rapid approach to approximate Kobe posteriors from age-structured assessment models with applications to North Atlantic shortfin mako. ICCAT – SCRS/2019/093.

Table 1. Names and fishery definitions of the fleets used in the SS model

Num.	Fleet/Index	Selectivity (all length based except fleet 15)	Time block Selectivity	use	start	end
1	JAPAN_LL	Double Normal	Y (1950-2009)	Y	1957	2018
2	USA_CAN_PSFS	Double Normal	N	Y	1950	1984
3	USA_CAN_PSFB	Double Normal	N	Y	1950	2015
4	USA_TRAP	Double Normal	N	Y	1950*	1974
5	USA_CAN_HARPOON	Logistic	N	Y	1950*	2018
6	USA_RRFB	Double Normal	N	Y	1950	2018
7	USA_RRFS	Double Normal	Y (1950-1992)	Y	1950	2018
8	OTHER_ATL_LL	Logistic	N	Y	1952	2018
9	CAN_HOOKLINE	Logistic	N	Y	1950	2018
10	GOM_LL_US_MEX	Logistic	N	Y	1972	2018
11	JLL_GOM	Logistic	N	Y	1974	1981
12	CAN_TRAP	Logistic	N	Y	1950	2018
13	CAN_GSL1	Logistic	N	Y	1950	1987
14	IND1_JAPAN_LL	mirror JAPAN_LL	N	Y	1976	2018
15	IDX2_US_RR_66_114	Double normal	N	Y	1993	2018
16	IDX3_US_RR_115_144	Double normal	N	Y	1993	2018
17	IDX4_US_RR_LT145	mirror RRFS	N	Y	1980	1992
18	IDX5_US_RR_GT177	mirror RRFB	N	Y	1993	2018
19	IDX6_US_RR_GT195	mirror RRFB	N	Y	1983	1992
20	IDX7_USPLL_GOM	mirror GOM_LL	N	Y	1987	1992
21	IDX8_JLL_GOM	mirror JLL_GOM	N	Y	1974	1981
22	IDX9_CAN_NS	mimic CAN_HL	N	Y	1984	2018
23	IDX10_GOM larval	mimic GOM_LL	N	Y	1977	2018
24	IDX11_tagging	NA	N	N	1970	1981
25	IDX12_CAN_ACOUSTIC	mimic CAN_GSL1	N	Y	1994	2018
26	IDX13_oceanographic	exp(rec devs)	N	N	1993	2011
27	IND14_JAPAN_LL2	mirror JAPAN_LL	N	Y	2010	2018
28	IND15_USPLL_GOM_LL2	mirror GOM_LL	N	Y	1993	2018

*fishery starts with equilibrium catch

Table 2. Task 1 landings input for SS3.

Year	JAPAN_LL	USA_CAN_PSFS	USA_CAN_PSFB	USA_TRAP	USA_CAN_HARPOON	USA_RRFB	USA_RRFS	OTHER_ATL_LL	CAN_HOOKLINE	GOM_LL_US_MEX	JLL_GOM	CAN_TRAP	CAN_GSL1
equ. Cat.	0	0	0	434.5	310	0	0	0	0	0	0	0	0
1950	0	1	0	346	459	88	38	0	0	0	0	10	75
1951	0	85	15	491	263	155	1	0	0	0	0	27	86
1952	0	0	0	135	323	95	0	7	0	0	0	65	69
1953	0	0	0	766	197	86	5	1	0	0	0	0	29
1954	0	47	8	531	129	46	13	0	0	0	0	0	49
1955	0	0	0	377	135	14	4	5	0	0	0	0	9
1956	0	0	0	181	47	14	2	0	0	0	0	0	3
1957	30	0	0	404	58	19	15	16	0	0	0	0	4
1958	32	117	21	869	61	64	3	40	0	0	0	0	0
1959	200	664	117	302	125	58	7	83	0	0	0	79	14
1960	339	235	42	204	119	45	10	1	0	0	0	32	5
1961	373	768	135	79	78	43	24	0	0	0	0	79	41
1962	1219	3203	565	87	44	237	135	132	0	0	0	137	40
1963	6191	4905	866	74	22	668	427	367	0	0	0	229	90
1964	12044	4378	773	161	24	309	200	303	0	0	0	318	99
1965	9147	2831	500	166	55	590	385	318	0	0	0	81	94
1966	2471	855	151	134	46	2182	1439	604	0	0	0	87	111
1967	694	17700	312	139	53	196	114	2432	0	0	0	174	56
1968	272	584	103	25	61	282	175	1393	0	0	0	101	180
1969	116	1118	0	38	30	757	113	477	0	0	0	193	170
1970	66	3335	953	53	72	447	57	202	0	0	0	130	151
1971	1375	3166	603	47	166	949	123	15	0	0	0	59	88
1972	321	1549	462	29	160	1058	111	18	0	23	0	29	188
1973	1097	1387	269	13	86	546	31	30	0	29	0	144	239
1974	824	892	68	20	214	185	2361	41	0	39	81	256	409
1975	237	2009	311	0	233	461	122	49	0	24	1276	144	206
1976	790	1365	217	0	189	382	28	246	0	37	2112	172	342
1977	1033	1292	210	0	157	512	60	118	0	14	2625	372	302
1978	709	1117	113	0	158	645	51	80	0	28	2436	221	208
1979	1298	1012	369	0	143	647	95	101	0	22	2323	31	214
1980	1420	537	221	0	102	553	82	37	0	10	2516	47	259
1981	1759	516	394	0	109	460	73	37	0	90	2012	41	279
1982	292	101	136	0	86	367	92	68	0	14	0	68	436
1983	711	109	275	0	159	616	121	118	0	12	0	7	426
1984	696	57	344	0	115	558	119	73	0	75	0	3	261
1985	1092	0	377	0	166	610	139	50	0	98	0	20	122
1986	584	0	360	0	127	419	97	577	0	124	0	0	41
1987	960	0	367	0	122	565	161	136	0	142	0	17	33
1988	1109	0	383	0	151	471	129	197	275	173	0	14	0
1989	468	0	385	0	187	622	166	255	579	101	0	1	0
1990	550	0	384	0	129	501	476	151	432	156	0	2	0
1991	688	0	237	0	129	570	483	150	479	193	0	0	0

Table 2. Task I landings input for SS3, continued.

Year	JAPAN_LL	USA_CAN_PSFS	USA_CAN_PSFB	USA_TRAP	USA_CAN_HARPOON	USA_RRFB	USA_RRFS	OTHER_ATL_LL	CAN_HOOKLINE	GOM_LL_US_MEX	JLL_GOM	CAN_TRAP	CAN_GSL1
1992	512	0	300	0	105	441	116	261	433	127	0	1	0
1993	581	0	295	0	121	558	209	148	372	71	0	29	0
1994	427	0	301	0	102	642	93	139	274	56	0	79	0
1995	387	0	249	0	120	661	260	184	457	58	0	72	0
1996	436	0	245	0	128	529	355	221	453	55	0	90	0
1997	330	0	250	0	153	762	190	181	383	26	0	59	0
1998	691	0	249	0	169	640	169	170	475	26	0	68	0
1999	365	0	248	0	154	673	103	648	473	62	0	44	0
2000	492	0	275	0	202	637	50	516	514	72	0	16	0
2001	506	0	196	0	122	1006	249	179	481	30	0	16	0
2002	575	0	208	0	68	1008	519	320	547	45	0	28	0
2003	57	0	265	0	98	677	315	285	449	76	0	84	0
2004	470	0	32	0	48	389	329	195	470	160	0	32	0
2005	265	0	178	0	46	257	170	163	541	129	0	8	0
2006	376	0	4	0	50	218	158	236	664	102	0	3	0
2007	277	0	28	0	40	235	399	155	412	88	0	4	0
2008	492	0	0	0	54	307	352	154	499	119	0	23	0
2009	162	0	11	0	84	717	143	290	427	122	0	23	0
2010	353	0	0	0	66	573	111	280	364	70	0	39	0
2011	578	0	0	0	100	420	173	341	342	27	0	26	0
2012	289	0	2	0	83	421	149	260	381	153	0	17	0
2013	317	0	43	0	70	251	115	243	377	55	0	11	0
2014	302	0	42	0	79	379	100	242	371	92	0	20	0
2015	347	0	39	0	103	582	112	163	427	62	0	6	0
2016	345	0	0	0	78	723	145	180	354	66	0	10	0
2017	346	0	0	0	99	658	142	178	369	46	0	13	0
2018	407	0	0	0	74	767	114	186	387	88	0	3	0

*gray shaded years are a product of an interpolated decline from PSFS to PSFB over 1980-1984.

** blue shaded years are a product of splitting PSFS and RR FS and RR FB

*** very minor “other” task I allocated to similar or most abundant fishery (usually US RRFB)

Table 3. Table of otolith age-length pairs by sampling laboratory (DFO: Canada Department of Ocean and Fisheries, St Andrews Biological Station; PC: US NMFS Panama City Lab; UMaine: University of Maine; UMCES: University of Maryland Center for Environmental Sciences).

	WBFTagesAll.4.27 (full 2020 dataset, no exclusions or outliers removed)				WBFTagesWithSSgear (2017 dataset)				WBFTagesStrictUpdateRemoveEarly (Strict update dataset)			
year	DFO	PC	UMaine	UMCES	DFO	PC	UMaine	UMCES	DFO	PC	UMaine	UMCES
1973	1											
1974				2				2				2
1975	180			154				154				154
1976	342			68				68				68
1977	269			26				26				25
1978	315			97				97				96
1979	72											
1980	137											
1981	170											
1982	33											
1983	347											
1996				75				75				75
1997				34				34				33
1998				43				43				43
1999				21				21				21
2000				6				6				6
2002				54				54				54
2009		80				77	35			79		
2010	63	60	293		63	60	293		62	60	292	
2011	292	276	342	108	288	271	328	108	288	273	339	106
2012	288	237	147	143	289	235		143	284	235		142
2013	327	135	247	114	330	134		114	321	135		114
2014	298	207	290		297	205			296	206		
2015	254	169	144		245	164			254	169		

2016	338	274	293						338	272	287	
2017	512	243							499	240		
2018	439	248							437	247		
total	4677	1929	1756	945	1512	1146	656	945	2779	1916	918	939

not included in strict update dataset

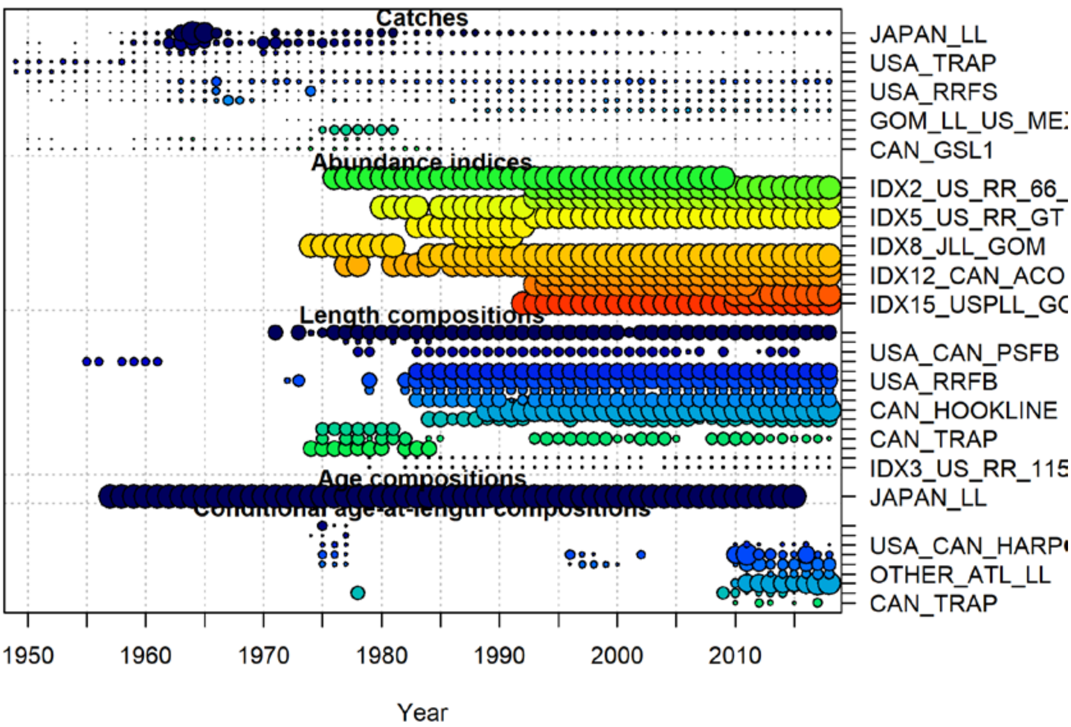

Figure 1. Time series of data inputs to the WBFT SS model.

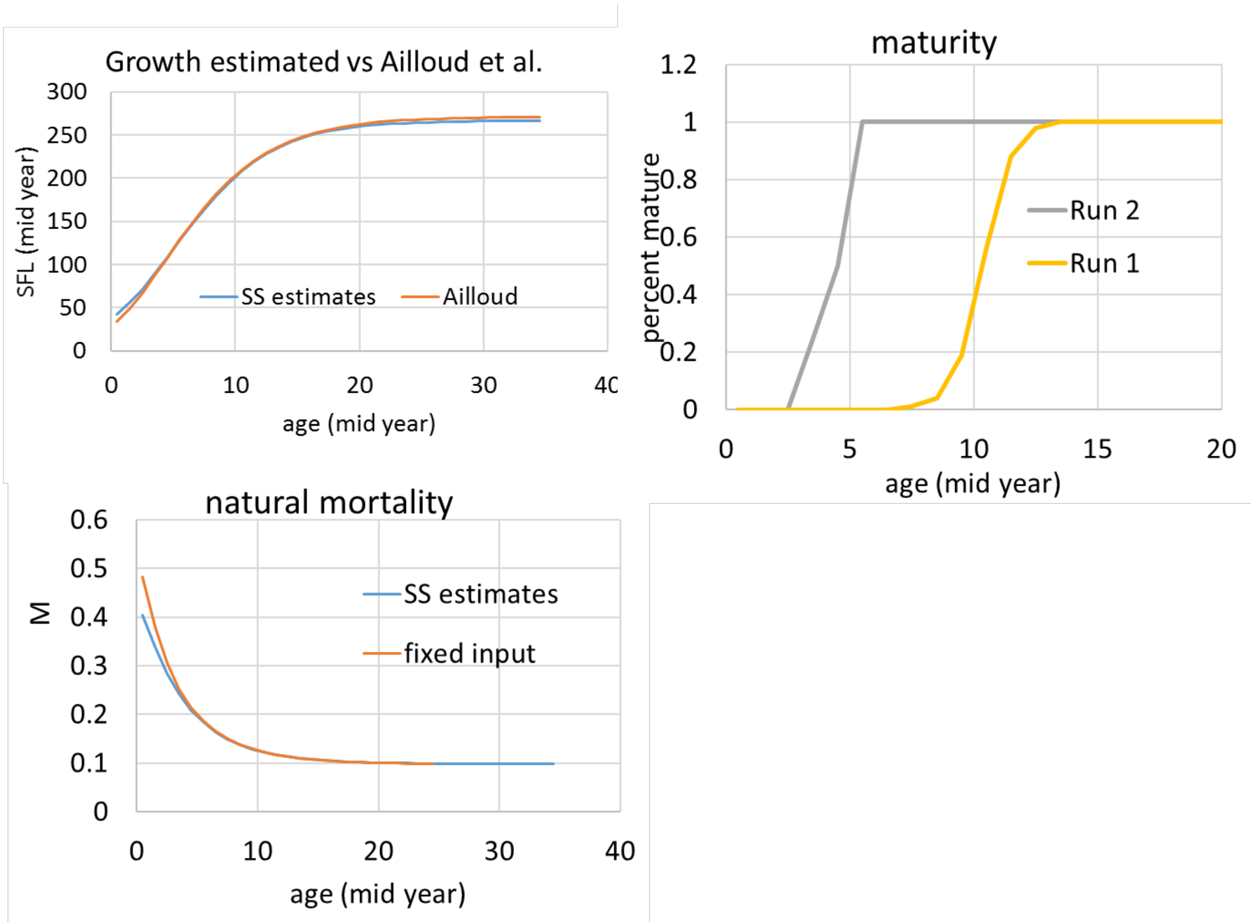

Figure 2. Estimated growth (from 2017 model) using a Richards function compared with Ailloud *et al.* (2017) growth estimate, maturity and mortality at age vector as scaled by SS using M=0.01 on age 20 and scaled by the growth curve.

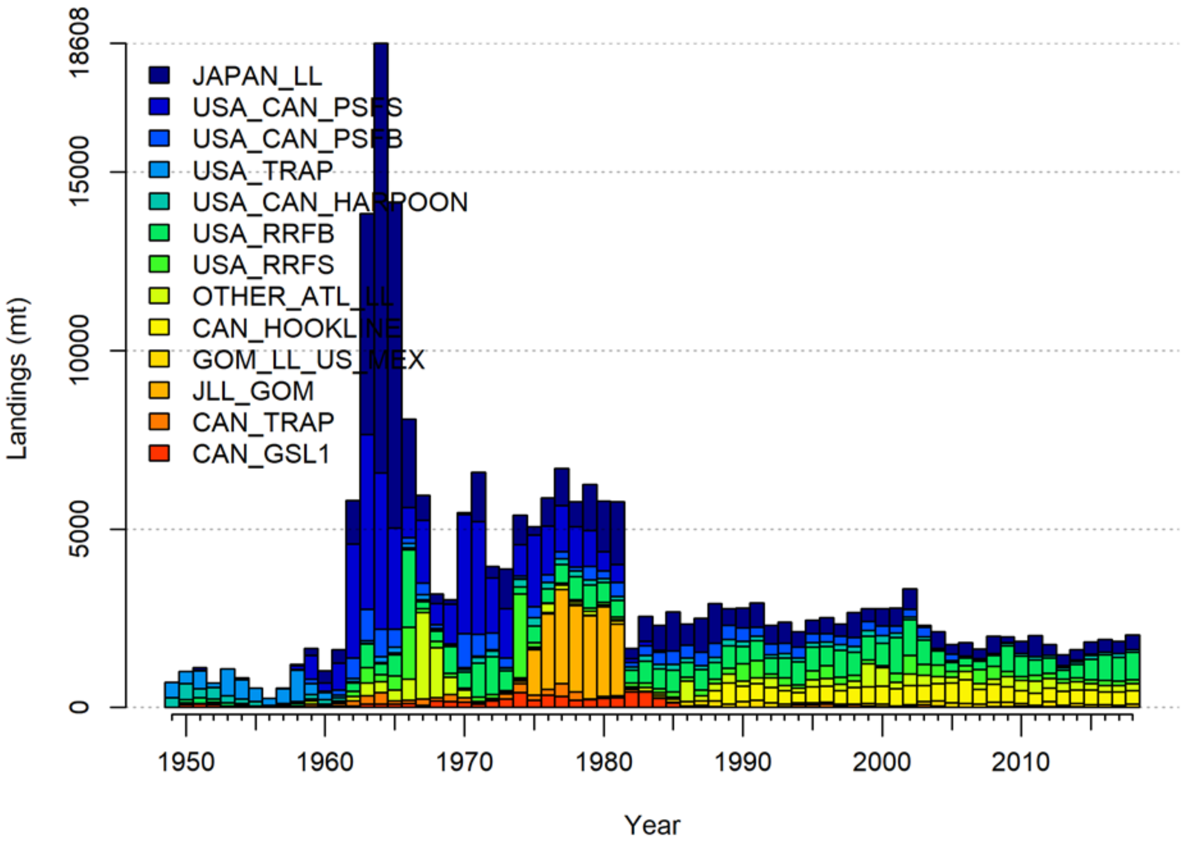

Figure 3. Task I catch by SS fleet.

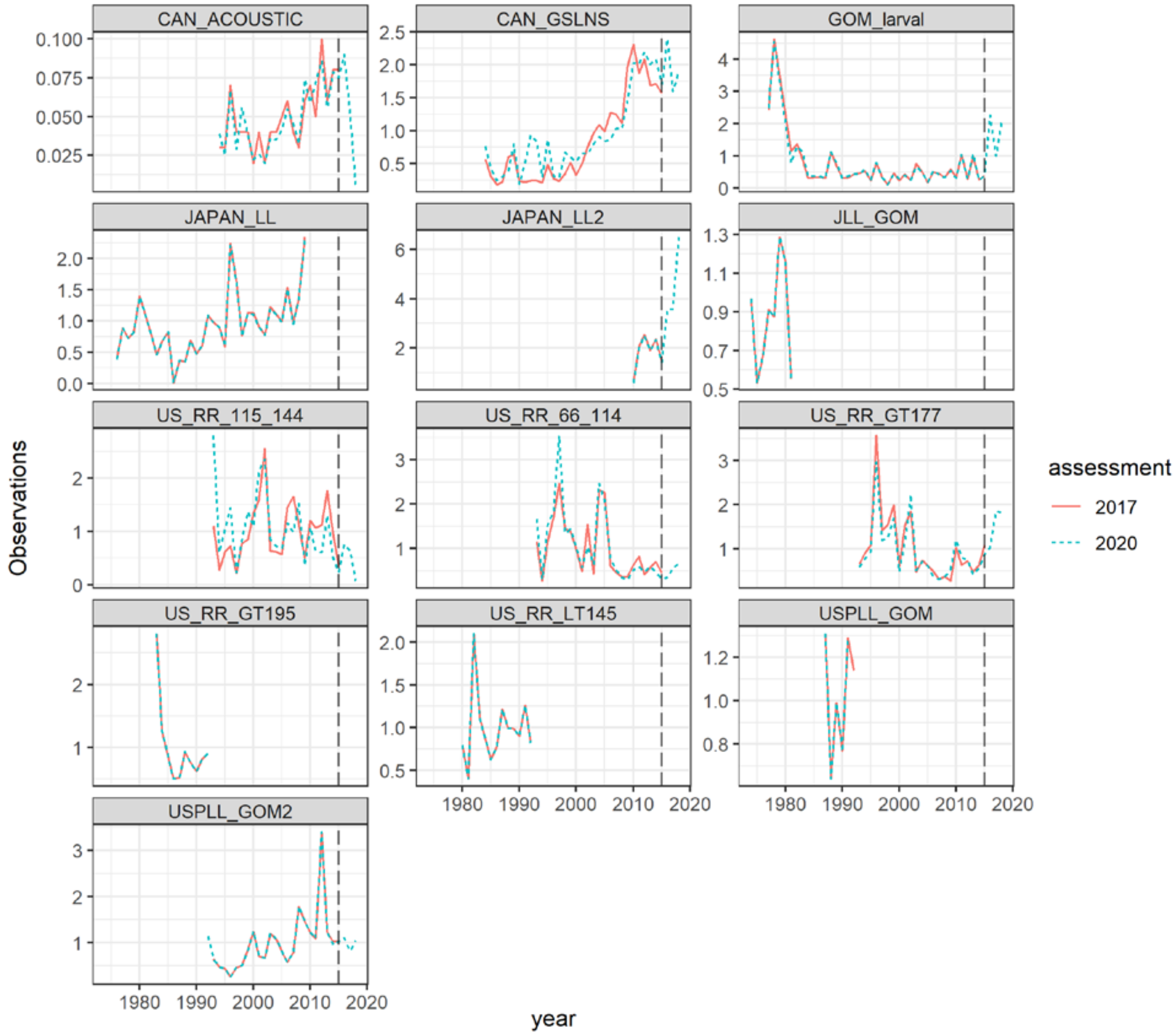

Figure 4. Indices used in SS3 assessment compared with indices used in 2017.

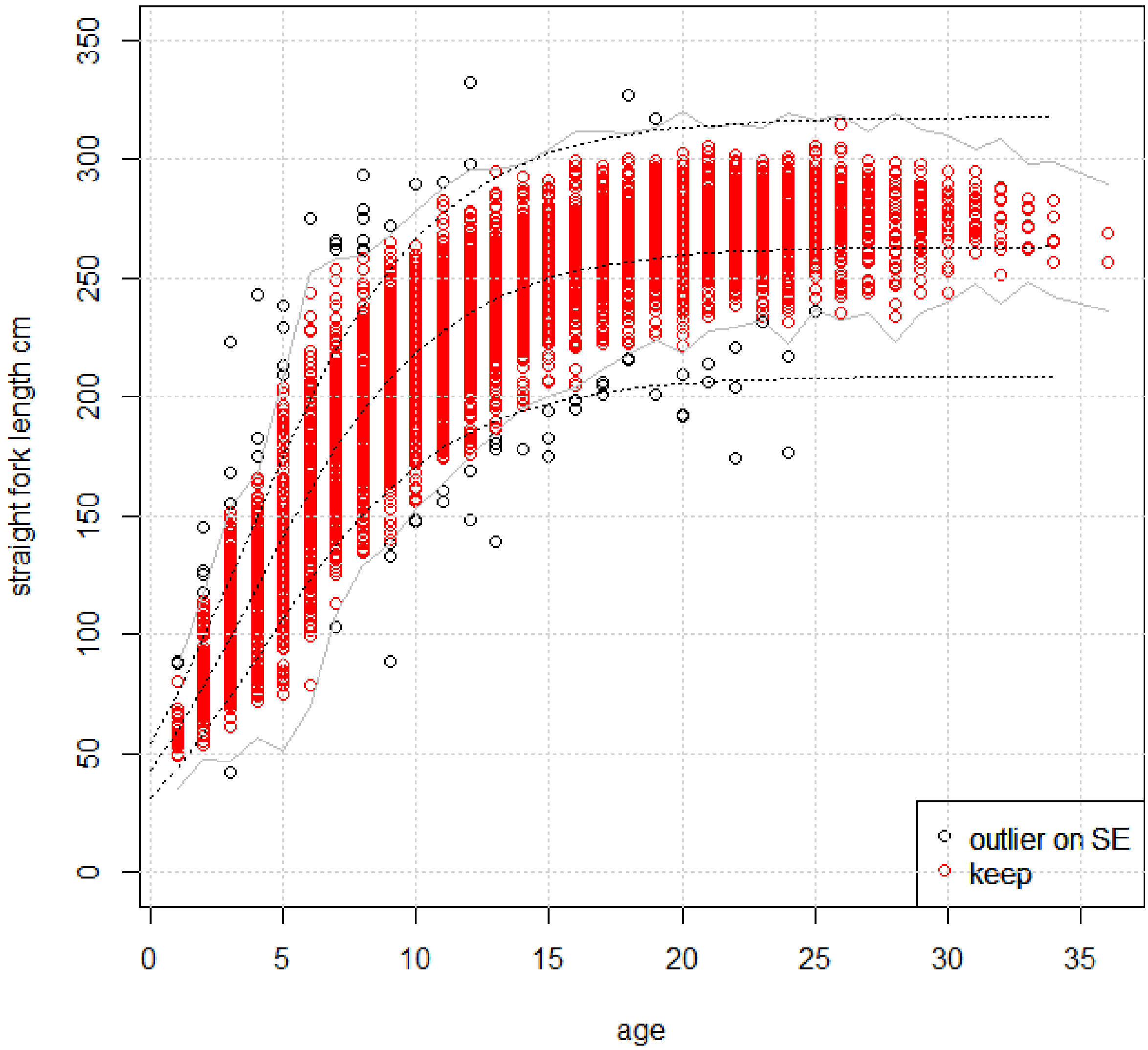

Figure 5. Straight fork length to age data. Black dots are observations that are +/- 3 standard deviations (gray lines) from the mean size at age. The dashed black lines are the mid year size at age as estimated by Stock Synthesis in 2017 +/- 3 standard deviations using the Richards growth function.

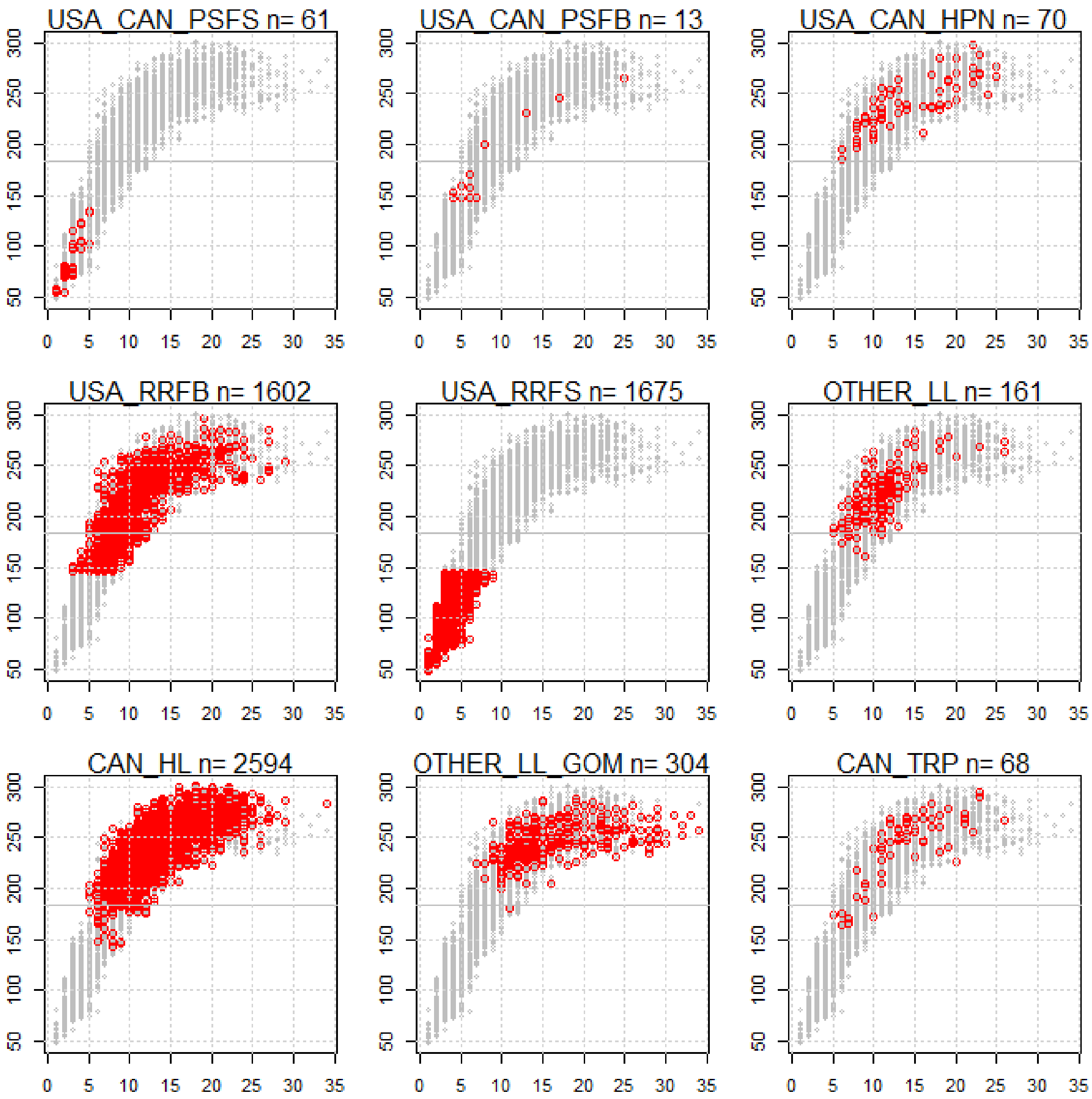

Figure 6. WBFT age-length data assigned (outliers exclude and only strict update data) to each fleet (red dots). Total age-length data are represented by the gray dots).

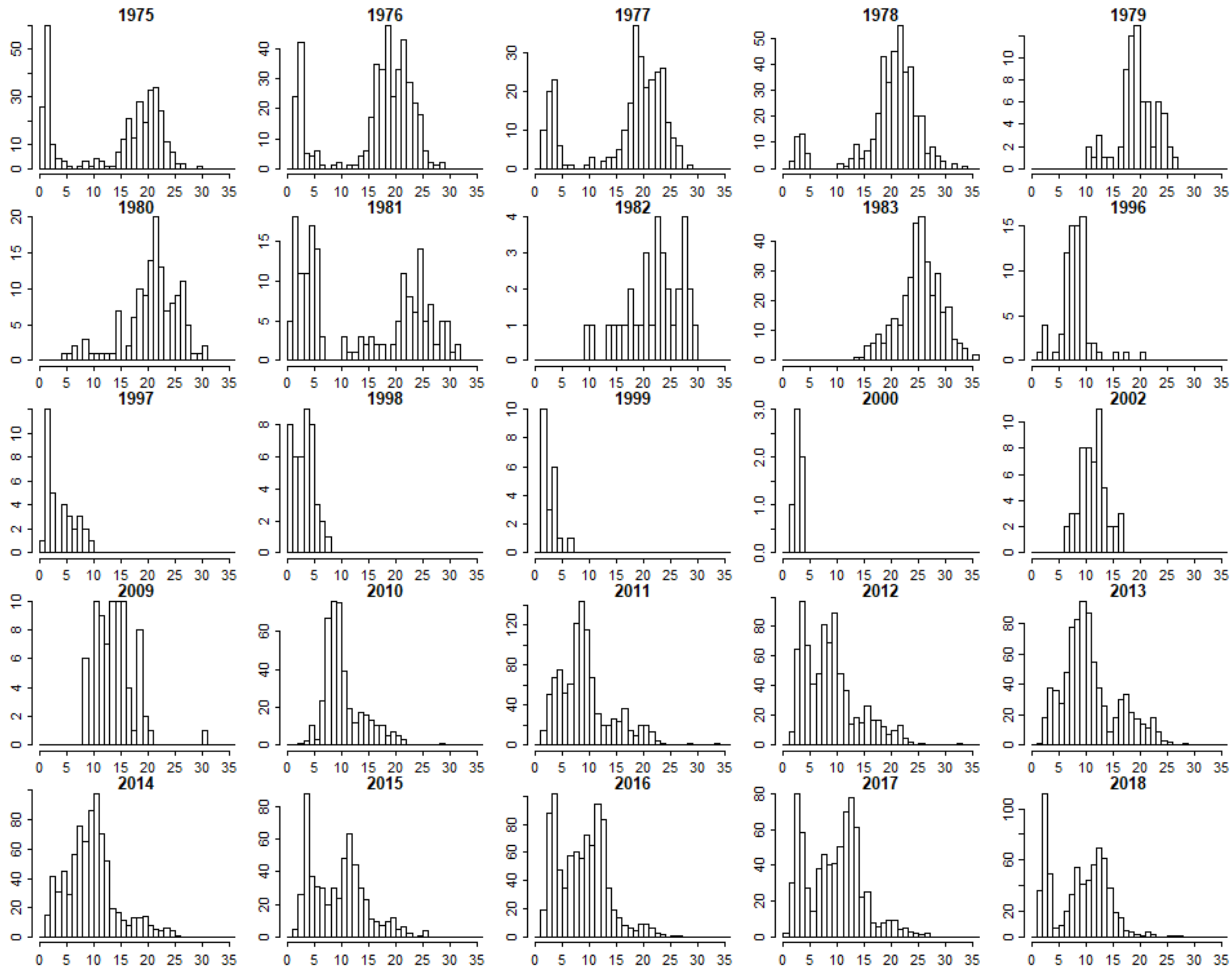

Figure 7. Histograms of age (Not just strict update dataset) data by year. Note that this is all gears and not necessarily representative of the all fleets and is not how the data are input to Stock Synthesis.

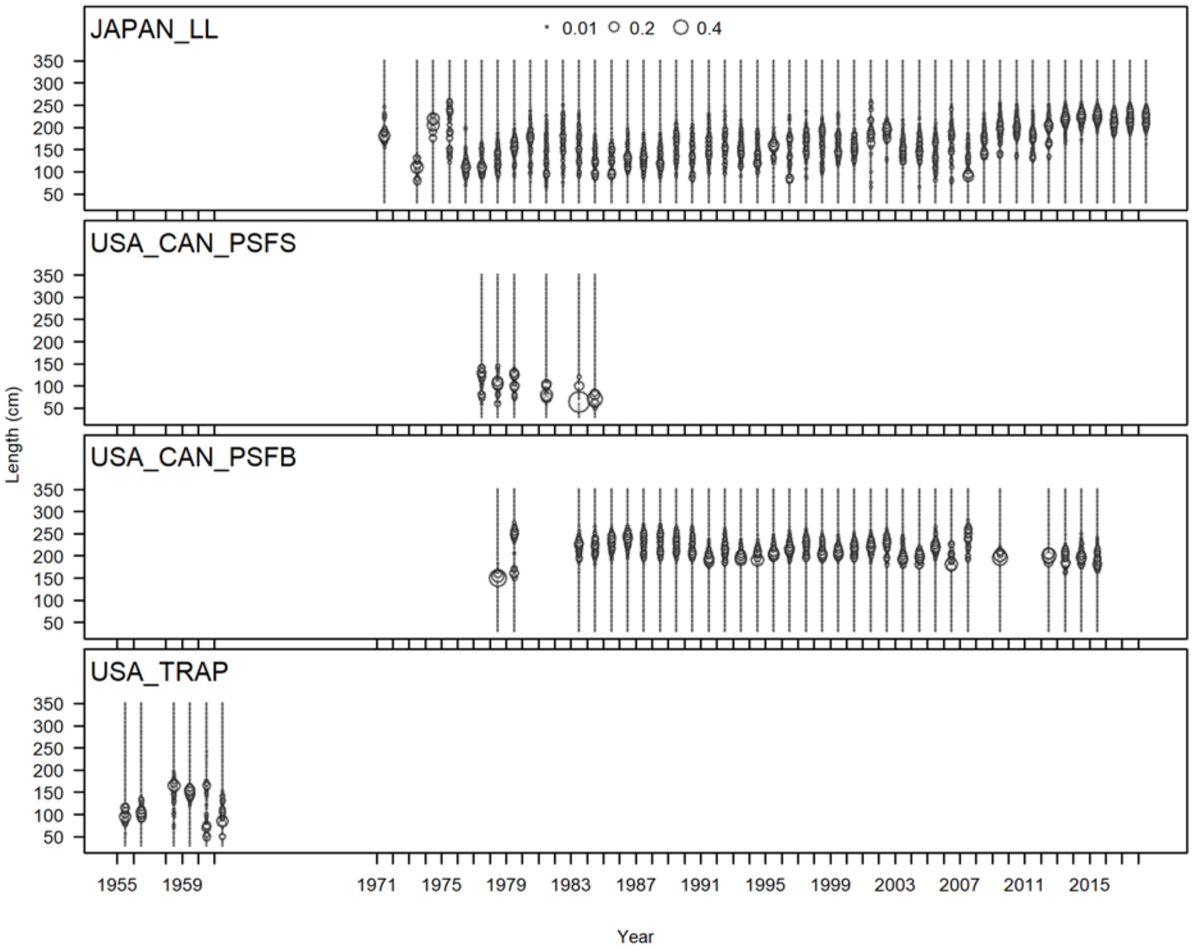

Figure 8. Size composition input for fleets 1-4.

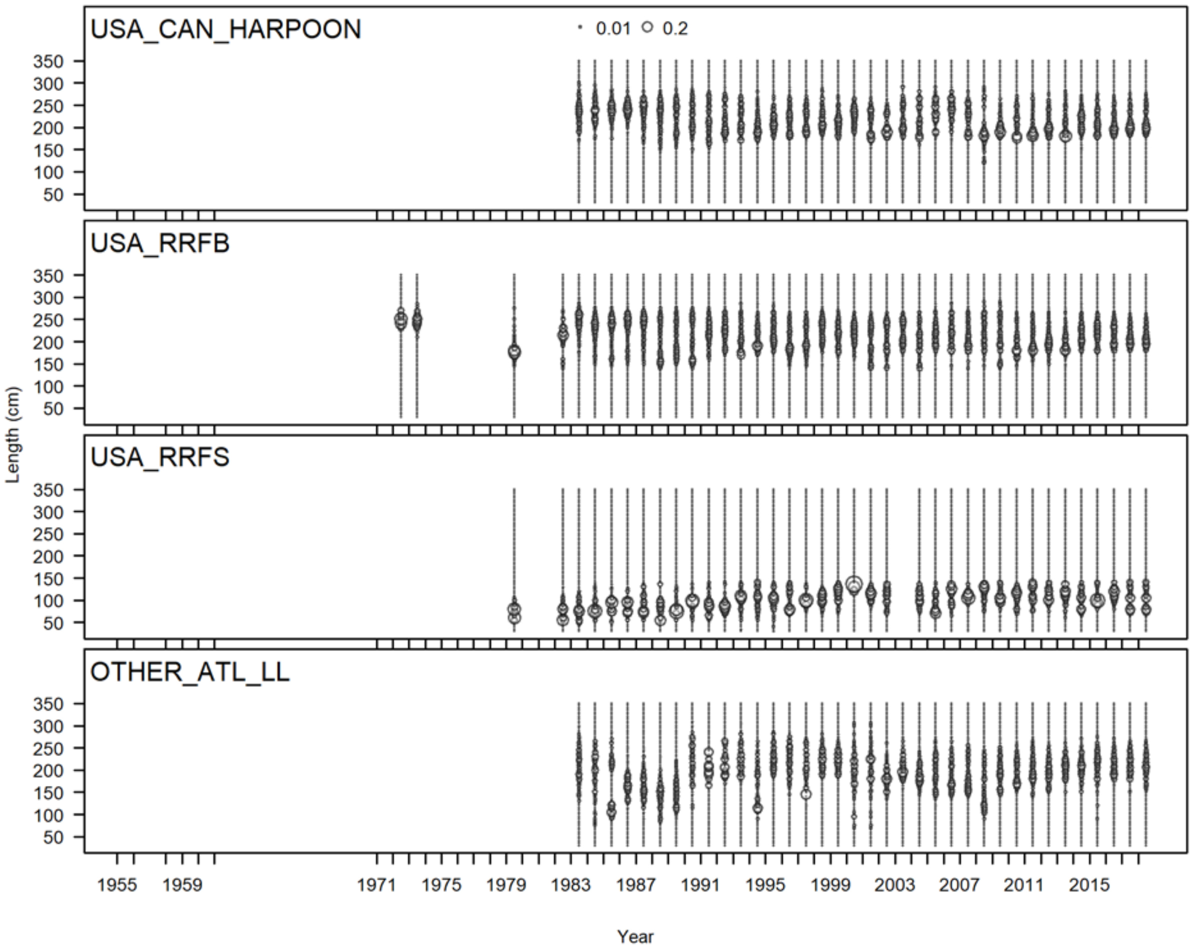

Figure 9. Size composition input for fleets 5-8.

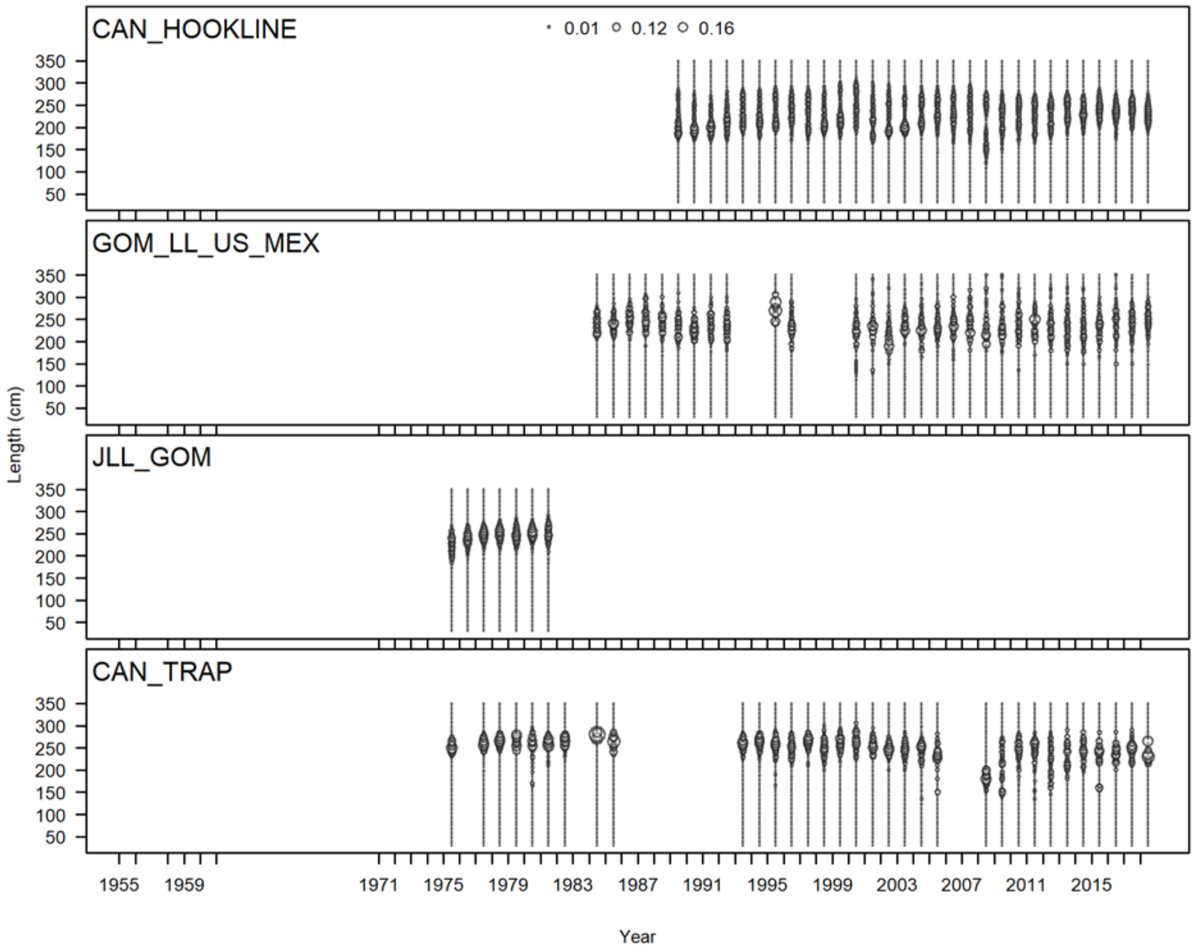

Figure 10. Size composition input for fleets 9-12.

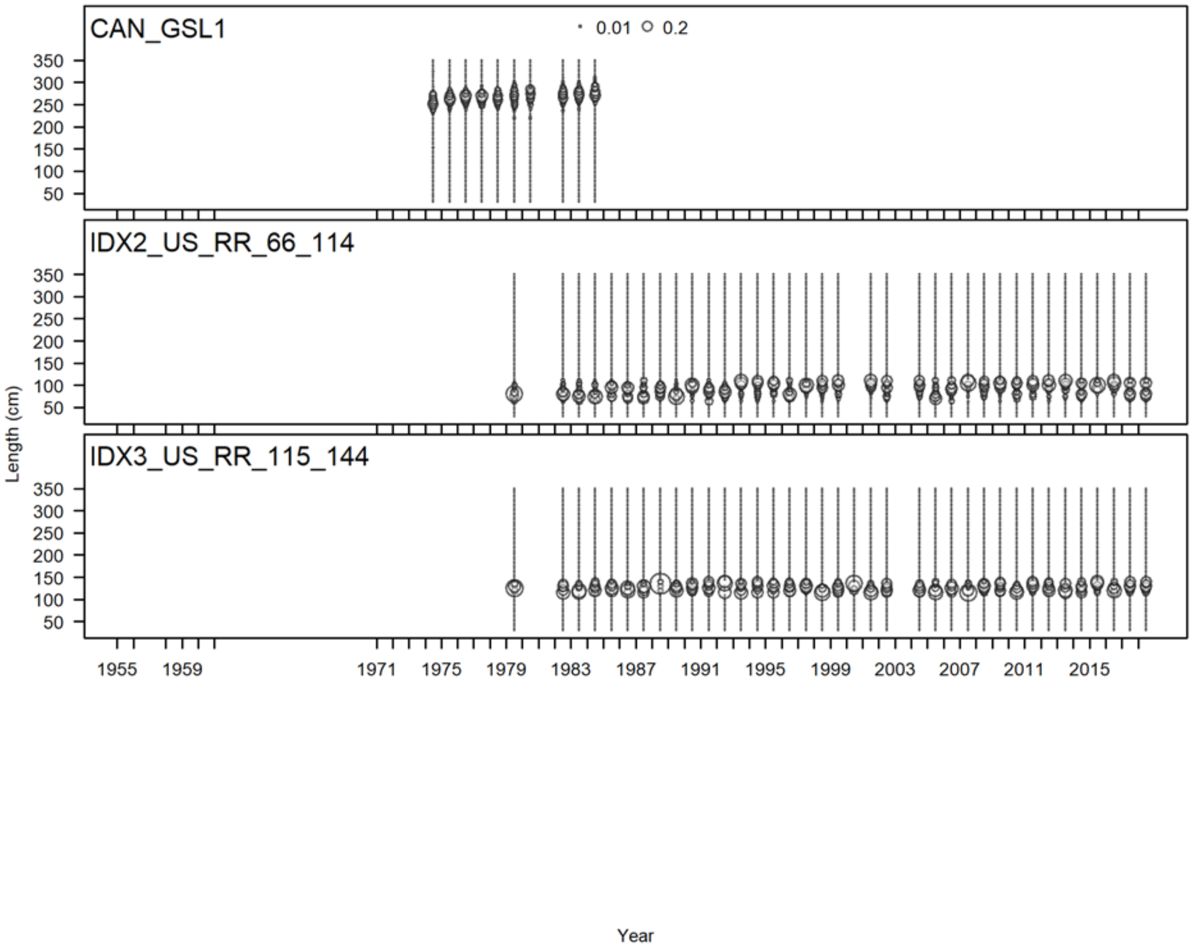

Figure 11. Size composition input for fleets 13, index 5, index 6.

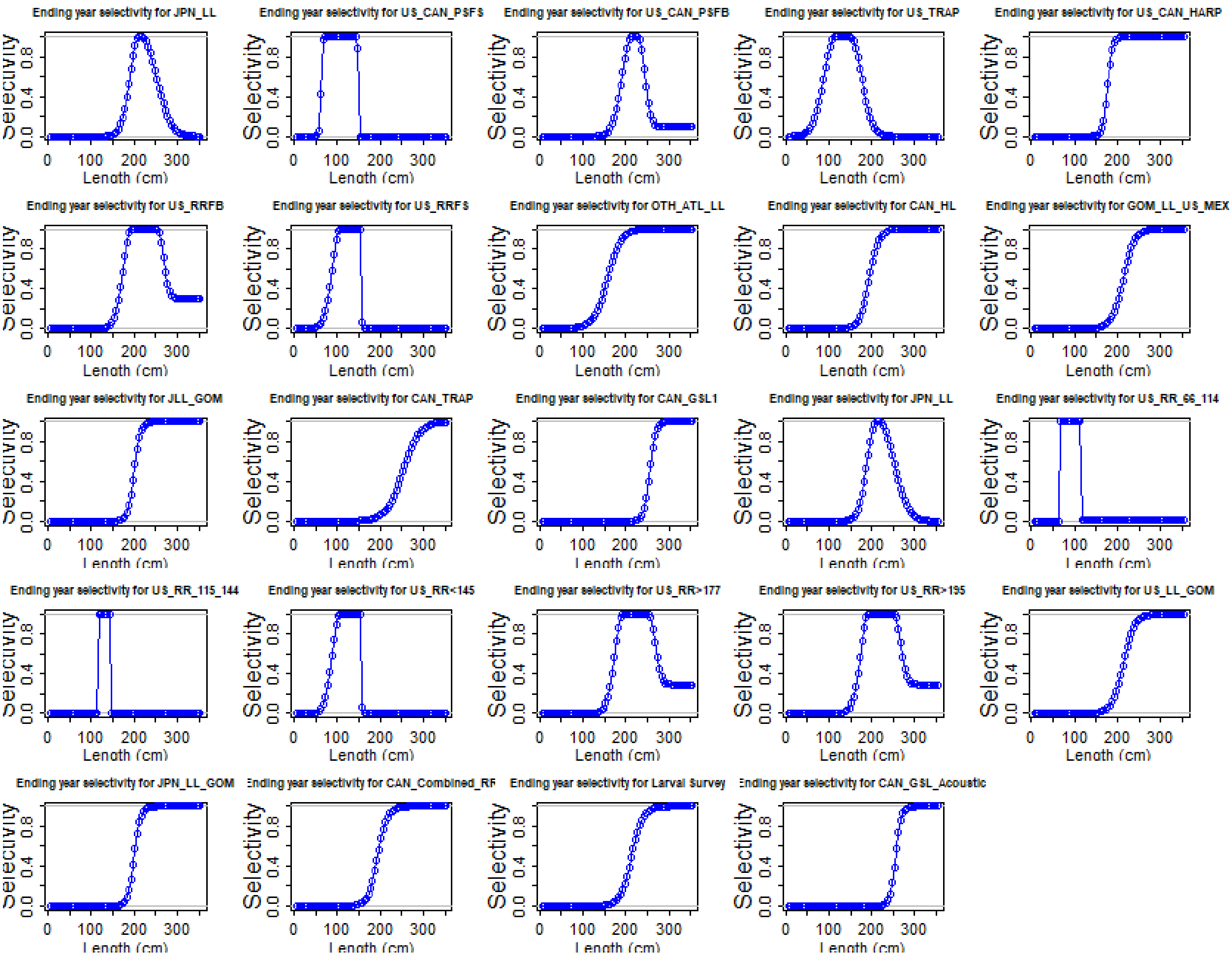

Figure 12. Estimated selectivities by fleet and survey.

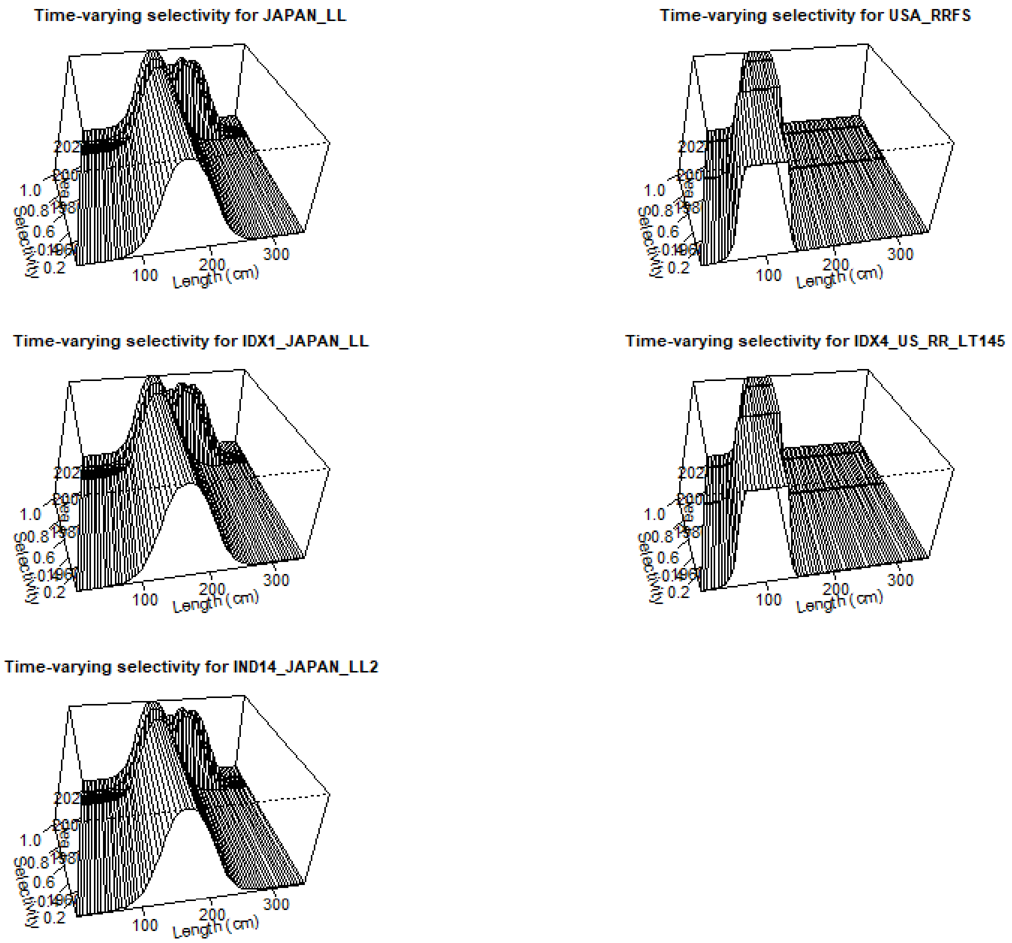

Figure 13. Estimated time varying selectivities.

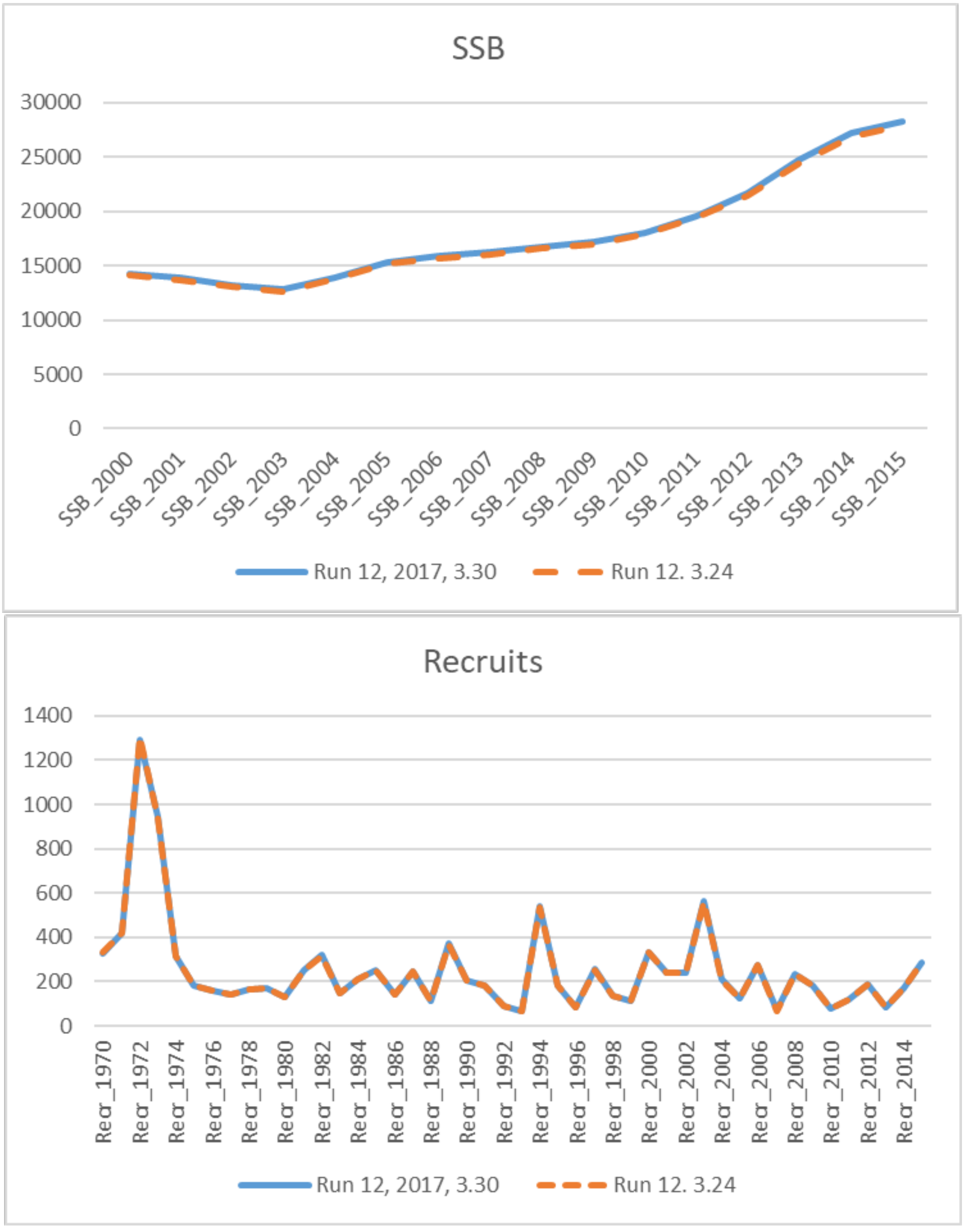

Figure 14. Conversion of Run 12 in 2017 from SS 3.24 to 3.30.

.