A Preliminary Base-case model in SS3.30 for the 2023 North Pacific Swordfish Assessment

Michelle L Sculley

4/1/23

# A preliminary base-case model in SS3.30 for the 2023 North Pacific Swordfish Assessment

Michelle L. Sculley,\* Yi-Jay Chang,+ Hirotaka Ijima^

\*NOAA NMFS Pacific Islands Fisheries Science Center, Honolulu, HI

michelle.sculley@noaa.gov

+National Taiwan University, Taipei, Taiwan

^Fisheries Research Agency, Yokohama, Japan

# Abstract

A preliminary base-case model in Stock Synthesis 3.30 for North Pacific (NP) swordfish (*Xiphias gladius*) is described for consideration as the 2023 base-case model. The base-case model covers the Western and Central North Pacific north of the Equator and the Eastern Pacific Ocean north of 10°N from 1975 to 2021. It includes data from three International Scientific Committee for the Conservation of Tuna and Tuna-like Species (ISC) countries and other countries in aggregate from the Western and Central Pacific Fisheries Commission (WCPFC) and the Inter-American Tropical Tuna Commission (IATTC). This paper describes the data available for inclusion and the proposed base-case model. The model converges and appears to fit the data well. Initial diagnostics do not indicate major problems. Preliminary results suggest the North Pacific swordfish stock is being not fished above FMSY and spawning stock biomass is above SSBMSY.

# 5Introduction

The International Scientific Committee for the Conservation of Tuna and Tuna-like Species (ISC) Billfish Working Group (BILLWG) has proposed to run a benchmark assessment on North Pacific swordfish (*Xiphias gladius,* SWO). Data were compiled from the International Scientific Committee for North Pacific Tuna and Tuna-like Species (ISC) member countries and other Western and Central Pacific Fisheries Commission (WCPFC) and Inter-American Tropical Tuna Commission (IATTC) countries. Countries were asked to contribute catch, CPUE, and size-frequency data. It was decided to run the assessment using a sex-specific model in Stock Synthesis version 3.30.20 (Methot and Wetzel, 2013) with fleets-as-areas spearating the Eastern Pacific Ocean (EPO) region from the Western and Central North Pacific (WCNPO) region, as defined in the December 2022 BILLWG meeting. The available data and the preliminary model results and diagnostics for a base-case model will be presented in this document for consideration at the ISC BILLWG SWO stock assessment meeting.

# Methods

## Spatial Temporal Structure

Data were compiled by region assuming a two region model of the North Pacific Ocean with boundaries based upon those detailed in Ichinokawa and Brodziak (2008) with the modification that the Eastern Pacific Ocean (EPO) region ends at 10°N ([Figure 1](#fig-1)). Countries were asked to contribute catch, CPUE, and length frequency data partitioned by these two regions so that the North Pacific could be modeled with two implicit areas using fleets as areas. The working group agreed to start model in 1975.

|  |
| --- |
| Figure 1: Western and Central North Pacific Ocean and North Eastern Pacific Ocean swordfish stock boundaries for the 2023 North Pacific swordfish assessment. Spatial structure is treated implicitly using fleets as areas. |

## Definition of Fisheries

Data are available for fourteen different fleets in the WCNPO: 14 catch time series, with some fleets split into early and late periods and 8 CPUE indices of which one is a recruitment index. The fleet names and numbers are detailed in Table 1. The data available for each fleet is in [Figure 2](#fig-2). The acronyms in the fleet names are defined as follows: WCNPO is Western and Central North Pacific Ocean; EPO is Eastern Pacific Ocean; OSDWLL is offshore distant water longline; OSDWCOLL is offshore distant water and coastal longline; early is the early time period; late is the late time period, Area1 and 2 are the Japanese fishery areas in the WCNPO as defined in Ijima 2022; OSDF is offshore driftnet gear; CODF is coastal driftnet gear, JPN\_WCNPO\_Other is Japanese small-scale coastal longline vessels which are not under obligation to submit logbook data, bait, and net fishing gear; DWLL is distant water longline gear, TWN\_WCNPO\_Other is Taiwanese offshore longline, coastal longline, gillnet, harpoon and other gears; LL is longline gear; shallow is the Hawaii shallow-set sector; deep is the Hawaii deep-set sector; GN is gillnet gear; US\_WCNPO\_Other is harpoon and other gears; WCPFC is other WCPFC and IATTC longline gear in the WCNPO; IATTC is longline gear in the EPO north of 10ºN.

|  |
| --- |
| Figure 2: Catch, CPUE index, and size composition data included in the 2023 NP swordfish stock assessment. The size of the bubble indicates the relative number of observations available. |

## Catch

Some of the time series of catch were divided into early and late periods to coincide with divisions of the CPUE indices (Table 1, [Figure 3](#fig-3)). Three ISC countries contributed catch time series: Japan, Taiwan, and the US. In addition, catch from countries reporting to the WCPFC and IATTC were obtained from each RFMO, respectively. The CV for catch was set to 0.05 for all fleets. Catch for fleets with only annual data were divided equally into each quarter.

|  |
| --- |
| Figure 3: Annual catch of NP swordfish by country or commission and area. |

## Relative Abundance Indices

Each of the eight CPUE time series were assigned to a quarter based upon the recommendations of the country providing the index and are assumed to represent the quarter in which the highest catches take place for each fishery. Japanese longline fleets (S1-4) were all assigned to quarter 1; Taiwanese longline fleet (S5) was assigned to quarter 3; US longline deep-set (S7) was assigned to quarter 2, and US longline shallow-set (S8 and S9) were assigned to quarter 1. US longline deep-set fleet S6 was included as an index of recruitment because the fishery catches large numbers of young-of-the-year fish (Fleet type 31, Bohaboy and Sculley, 2023). The CPUE indices were assumed to be linearly proportional to biomass where catchability (*q*) was assumed to be constant and occur in the first month of the quarter assigned.

The CVs for each CPUE index were assumed to be equal to their respective calculated SEs on the log scale. The minimum CV was scaled to a minimum of 0.20 or the root-mean-square error (RMSE) (i.e., square root of the residual variance) of what we would expect the assessment model to fit the CPUE index best by adding a constant to each CV value. This was calculated as the square root of the residual variance of a loess smoother fit to each index (Francis 2011, Lee et al., 2014).

where *Y*t is the observed CPUE in year *t* on the log scale, is the predicted CPUE in year *t* from the smoother fit to the data on the log scale, and *N* is the number of CPUE observations. RMSE values for each index are listed in Table 2. If the input SE was greater than these values, it was left unchanged.

## Length Composition

Length composition data were available for six WCNPO fleets and two EPO fleets ([Figure 4](#fig-4)). Length composition data were available in quarterly time steps. Quarters with fewer than 15 total samples were removed from the time series due to limited sample size, as agreed upon by the modeling sub-group. In addition, the length composition data for F5 were excluded as they were sparse. Data were fit using a multinomial error structure. Length composition data were weighted based upon the Francis (2011) method. Input effective sample size was estimated as the total number of fish measured in each year-quarter divided by 10, with any year-quarter over fifty set to fifty to reduce the influence of large sampling events. Then, weighting was attempted based upon the T.A1.8 equation (Francis 2011) as calculated by the model using r4ss, an R package for plotting SS results (R version 3.4.0, R Core Team, 2017, r4ss version 1.28.0, Taylor et al., 2017). Length composition data were only down-weighted if the model suggested it, otherwise the fleets were not adjusted. The only fleet to be down-weighted using the Francis methods was F9, the US deep-set LL data.

|  |
| --- |
|  |

|  |
| --- |
|  |

Figure 4: Length Composition data available in 5cm size bins for the 2023 NP swordfish stock assessment.

## Initial Base-case Model Description

The assessment was conducted with Stock Synthesis (SS) version 3.30.20.00 released 09/30/2022 using Otter Research ADMB 13.0 (Methot and Wetzel 2013). The model was set up as a single area model with two sexes and four seasons (quarters) with fleets as areas. Spawning was assumed to occur in May (month 5) while recruitment was assumed to occur in July (month 7). Age at recruitment was calculated based upon the model estimated average selectivity at age based upon the quarterly selectivity at length. The maximum age of swordfish was set to 15 years. Sex specific biological parameters were used, with sex- and age-specific natural mortality (Table 3) as agreed upon in the BILLWG data preparatory meeting (ISC Billfish WG 2023). In addition, the CV of the growth curve was set to 0.1 for small males and females and 0.15 for large males and females, and the sex ratio at birth was assumed to be 1:1. The model used a Beverton-Holt spawner-recruit relationship with steepness (h) fixed at 0.9 and sigmaR (σr) initially fixed at 0.6 and rescaled to 0.42 based upon the model suggestion.

Twenty-seven fleets were included in the model: 19 catch fleets and 8 survey fleets. Initial equilibrium catch was estimated for 1974.

Main recruitment deviations were estimated from 1985-2021, as this is the period during which there are size composition data. The recruitment deviations were bias-adjusted based upon the estimates from Methot and Taylor (2011) provided from the model results. No bias adjustment was applied to recruitment deviations from 1965-1967. 1967-1995 was the “ramp-up” period where the bias adjustment of σr was 0 at the beginning of the period and increased linearly to the maximum bias adjustment 0.94 in 1995. Full bias adjustment was from 1995-2021. The early period of recruitment deviations represents a data-poor period where there is little information to drive recruitment. The main recruitment period represents a data-rich period where there is enough data to drive the bias-adjustment of the recruitments. The ramp up period allows for a gradual ramp up of the bias-adjustment between the data-poor and data-rich periods.

The population model and the fishery length data had 51 five cm length bins from 10-260+ cm. The population had 16 annual ages from age 0 to 15+. There were no age data. Fishery length data were used to estimate selectivity patterns which controlled the size distribution of the fishery removals. All fleets with length data were estimated as six parameter double normal (dome-shaped) selectivity patterns. Survey selectivity patterns mirrored their respective catch fleets (Table 4). Estimating parameters six of the double normal selectivity pattern for F9 resulted in an improved fit to the size composition data for that fleet.

Model estimated time series of total biomass (B in metric tons, mt = 1000 kg), age 1+ total biomass (B1+ mt), female spawning biomass (SSB mt) and recruitment (R in 1000s of fish) were tabulated on an annual basis. Annual exploitation rate (F) was calculated as Catch/B1+. Stock status indicators were calculated based upon MSY-based reference points as proxies, given that the WCPFC has not set biological or other reference points for swordfish.

## Convergence Criteria and Diagnostics

The model was assumed to have converged if the standard error of the estimated parameters could be derived from the inverse of the negative hessian matrix. Various convergence diagnostics were also evaluated. Excessive CVs (>50%) on estimated parameters would suggest uncertainty in the parameter estimates or model structure. A gradient of >0.001 would suggest poorly fit parameter estimates. The correlation matrix was also evaluated to identify highly correlated (>95%) and non-informative (<0.01) parameters. Parameter estimates hitting bounds of the prior was also indicative of poor model fit.

Several diagnostics were run to evaluate the fit of the model to the data. An Age-Structure Population Model (APSM) was used to evaluate the influence of the length composition data on the population trends (Carvalho *et al.*, 2017). The ASPM was also used to explore how each CPUE index informed the population trends by running one-off ASPMs for each index. Profiling the likelihood on R0, where the R0 is fixed at a range of values around the maximum likelihood estimate and then the likelihood is estimated, was used to identify influential data components (Lee *et al.,* 2014). A runs test was used to evaluate randomness in the residuals of the CPUE data (Carvalho *et al.,* 2021). Residual plots and plots of the observed vs expected data were examined to evaluate goodness-of-fit. Finally, a retrospective analysis and hindecast cross-validation were used to evaluate the predictive ability of the model (Carvalho *et al.,* 2021).

# Results

## Model fit

The base-case model ran in about 10 minutes, estimated 89 parameters, and had a total likelihood of 1338.75 The inverse Hessian was positive definite, which allowed for the estimation of parameter standard deviations and suggests that the model converged, and the maximum gradient component was less than 0.001. None of the parameter estimates hit a bound, two selectivity parameters had a correlation above 0.95 and no parameters had correlations below 0.01. All twenty early recruitment deviations (1965-1984) and 30 of the 37 main recruitment deviations had CVs > 50%. Five of the 34 selectivity parameters had CVs >50%, most of which were for the estimated width of the peak (i.e. a very small number).

Fits to the abundance indices were relatively good ([Figure 5](#fig-5) - [Figure 12](#fig-12)). The expected CPUE trend for S5 TWN LL was relatively flat and did not fit the data well, and the observed CPUE in the last 3-5 years for S4 JPN LL area 1 declined but the expected CPUE increased. Three indices failed the runs test, S5, S6, and S8 (Figure 11). S6 is the US deep set LL fleet which is a recruitment index and not fitted as a survey index, so its fit can be ignored. S8 is the US shallow-set LL fleet in the late time period, and indicates a long term decline. Additional investigation into the CPUE indices indicate that there does appear to be some conflict between the trends for each fleet, which could be driving some of the mis-fit within the model.

|  |  |
| --- | --- |
|  |  |

Figure 5: Fit to the S1 Japanese early area 1 LL CPUE index. Left is the input CPUE with CV and the model fit CPUE (blue line). Right is the annual residuals of that fit.

|  |  |
| --- | --- |
|  |  |

Figure 6: Fit to the S2 Japanese late area 1 LL CPUE index. Left is the input CPUE with CV and the model fit CPUE (blue line). Right is the annual residuals of that fit.

|  |  |
| --- | --- |
|  |  |

Figure 7: Fit to the S3 Japanese early area 2 LL CPUE index. Left is the input CPUE with CV and the model fit CPUE (blue line). Right is the annual residuals of that fit.

|  |  |
| --- | --- |
|  |  |

Figure 8: Fit to the S4 Japanese late area 2 LL CPUE index. Left is the input CPUE with CV and the model fit CPUE (blue line). Right is the annual residuals of that fit.

|  |  |
| --- | --- |
|  |  |

Figure 9: Fit to the S5 Chinese Taipei late LL CPUE index. Left is the input CPUE with CV and the model fit CPUE (blue line). Right is the annual residuals of that fit.

|  |  |
| --- | --- |
|  |  |

Figure 10: Fit to the S6 US deep-set LL CPUE recruitment index. Left is the input CPUE with CV and the model fit CPUE (blue line). Right is the annual residuals of that fit.

|  |  |
| --- | --- |
|  |  |

Figure 11: Fit to the S7 US shallow-set LL early CPUE index. Left is the input CPUE with CV and the model fit CPUE (blue line). Right is the annual residuals of that fit.

|  |  |
| --- | --- |
|  |  |

Figure 12: Fit to the S8 US shallow-set LL early CPUE index. Left is the input CPUE with CV and the model fit CPUE (blue line). Right is the annual residuals of that fit.

|  |
| --- |
| Figure 13: Results from a runs test for each CPUE index. Red indicates the index failed the test (residuals are not random), green indicates the index passed the test. |

Estimated selectivity for each fleet are in . Fits to the length composition data were also relatively good ([Figure 17](#fig-17) - [Figure 20](#fig-20)), although there are still problems fitting the US deep-set longline data (F9).  The fit to the US size data is challenging because size distribution changes substantially seasonally, with a sharp peak of small fish entering the fishery in quarters 3 and 4 which are not observed in such large numbers in quarters 1 and 2. Attempts to separate the fleet into a quarterly fleets to estimate selectivity are ongoing but not yet successful. In addition, F1 JPN LL area 1 late and F4 IATTC both fail the runs test ([Figure 21](#fig-21)).

|  |
| --- |
| Figure 14: Time-varying selectivity estimated for F01 Japan LL area 1 late. |

|  |
| --- |
| Figure 15: Time-varying selectivity estimated for F02 Chinese Taipei LL late. |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  | | --- | | F03\_US\_WCNPO\_LL\_shallow\_late | |  | |  | | --- | | F04\_IATTC | |  | |  | | --- | | F06\_JPN\_WCNPO\_OSDWLL\_early\_Area1 | |  |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| |  | | --- | | F07\_JPN\_WCNPO\_CODF | |  | |  | | --- | | F08\_US\_WCNPO\_LL\_shallow\_early | |  | |  | | --- | | F09\_US\_WCNPO\_LL\_deep | |

Figure 16: Selectivity estimates for each of the 6 fleets without time-varying parameters.

|  |  |
| --- | --- |
|  |  |

|  |  |
| --- | --- |
|  |  |

Figure 17: Fits to the annual mean length (left panels) and quarterly residuals (right panels) for Japan LL Area 1 late (top) and Chinese Taipei LL late (bottom) length composition data. The blue line indicates the estimated mean length, open dots indicate input mean length with black bars indicating the distribution of the length data with the added variance. Open circles indicate negative residuals and closed circles indicate positive residuals.

|  |  |
| --- | --- |
|  |  |

|  |  |
| --- | --- |
|  |  |

Figure 18: Fits to the annual mean length (left panels) and quarterly residuals (right panels) for US shallow-set LL late (top) and IATTC EPO (bottom) length composition data. The blue line indicates the estimated mean length, open dots indicate input mean length with black bars indicating the distribution of the length data with the added variance. Open circles indicate negative residuals and closed circles indicate positive residuals.

|  |  |
| --- | --- |
|  |  |

|  |  |
| --- | --- |
|  |  |

Figure 19: Fits to the annual mean length (left panels) and quarterly residuals (right panels) for Japan LL Area 1 early (top) and Japan coastal driftnet (bottom) length composition data. The blue line indicates the estimated mean length, open dots indicate input mean length with black bars indicating the distribution of the length data with the added variance. Open circles indicate negative residuals and closed circles indicate positive residuals.

|  |  |
| --- | --- |
|  |  |

|  |  |
| --- | --- |
|  |  |

Figure 20: Fits to the annual mean length (left panels) and quarterly residuals (right panels) for US shallow-set LL early (top) and US deep-set LL (bottom) length composition data. The blue line indicates the estimated mean length, open dots indicate input mean length with black bars indicating the distribution of the length data with the added variance. Open circles indicate negative residuals and closed circles indicate positive residuals.

Warning: combining data from multiple seasons

|  |
| --- |
| Figure 21: Results from a runs test for each size composition time series. Red indicates the time series failed the test (residuals are not random), green indicates the time series passed the test. |

Model estimates of age 1+ biomass show a general increase in biomass from through the entire assessment time horizon, with some periodic dips in abundance ([Figure 22](#fig-22)). Initial spawning stock biomass was estimated to be approximately 4,200 mt and virgin SSB was around 25,000 mt ([Figure 23](#fig-23)). Annual fishing mortality is reported as the average for fish ages 1-10 ([Figure 24](#fig-24)). Fishing mortality was above MSY for all except 6 years and excepting 2015, has been below FMSY since 2014. Recruitment deviations suggested three periods of recruitment: high recruitment from 1975 to 1993, a period with little data and large variability around the estimates, average recruitment from 1994 to 2003 where recruitment varied around equilibrium recruitment, and low recruitment from 2004 to 2020. The log of the deviations were generally between 0.6 and -0.6 ([Figure 25](#fig-25)). Current depletion, as estimated as the age 1+ biomass in 2020 compared to the virgin age 1+ biomass was estimated to be 0.09.

|  |
| --- |
| Figure 22: Estimated biomass (mt) of NP swordfish 1+ from the base-case model. |

|  |
| --- |
| Figure 23: Estimated NP swordfish Spawning Stock Biomass (SSB) from the with 95% confidence intervals. SSBMSY is indicated by the dashed green line. |

|  |
| --- |
| Figure 24: Estimated annual fishing mortality (average ages 1-10) the with 95% confidence intervals. FMSY is indicated by the dashed green line. |

|  |
| --- |
| Figure 25: Estimated annual recruitment (thouseands of age-0 fish) with 95% confidence intervals. |

## Diagnostics

Profiling on R0 showed that the recruitment estimates were highly influential in the model results, but there was relative agreement between the CPUE indices and the length composition data on the lower bound of R0 ([Figure 26](#fig-26) - @fig28). The US data (CPUE and length comp) drive the model dynamics suggesting an ln(R0) below 6.0, and Chinese Taipei data and Japanese size composition data suggesting an ln(R0) around 7. Japanese CPUE data suggest an ln(R0) around 6.1` (Tables 4-5).

|  |
| --- |
| Figure 26: Likelihood profile over R0 for the base-case model: total likelihood (black circles), recruitment (blue triangles), length composition data (light blue crosses), and survey/CPUE indices (yellow x). |

|  |
| --- |
| Figure 27: Likelihood profile over R0 by CPUE index for the base-case model. |

|  |
| --- |
| Figure 28: Likelihood profile over R0 for each length composition time series for the base-case model. |

Results from the ASPM model showed the same population trend as the full model during 1975-1993, which is the time period without any size composition data. After 1994, the ASPM biomass increases drastically and deviates completely from the base-case model (Figures 29 and 30). Further investigation running an ASPM with a single early and late index at a time indicates a similar pattern for all CPUE indices (Figure 31). The most likely explanation for this is that the population dynamics prior to 1994 are being driven by the two CPUE indices available (Figure 32). After 1994 catch decreases and CPUE for all fleets flattens out which would indicate that a stock is recovering. However, the size composition data indicate that the majority of the catch is juvenile fish. Continued removals of individuals before they have a chance to reproduce would continue to cause the stock to decline below MSY levels. This indicates that the size composition data are an integral component of the model, without which we would not have a full picture of the fishing effect on the stock.

|  |
| --- |
| Figure 29: Spawning stock biomass trend for the ASPM model run (dashed line, triangles) and the base-case model (solid line, circles). Grey shading indicates 95% confidence intervals for each model. |

The retrospective analysis indicates that a significant retrospective pattern exists for both biomass and fishing mortality (Mohn’s rho = 0.2 and -0.14, respectively, Figure 33). Generally, biomass is overestimated and fishing mortality is underestimated. Results of the hindcast with cross-validation indicate that of the four CPUE indices at the end of the assessment horizon, only Chinese Taipei had reasonable predictive ability (MASE = 0.9), with all other fleets MASE > 1 (Figure 34). Comparing the predictive ability of the size composition data, two fleets had very good predictive ability (MASE <0.5, F2 and F18), five had good predictive ability (MASE <1 and >0.5, F1, F4, F5, F6, F16) and one had poor predictive ability (MASE <1, F14, Figure 35). F14 only had one datapoint in the analysis which likely explains why it had poor predictive ability.

length(dirvec) as input to SSgetoutput:6

reading output fromC://users//michelle.sculley//documents//2023 SWO ASSESS//ModelDev//Current Best/Diagnostics/Retros/Retro\_full/retrospectives/retro0/Report.sso

added element 'replist1' to list

reading output fromC://users//michelle.sculley//documents//2023 SWO ASSESS//ModelDev//Current Best/Diagnostics/Retros/Retro\_full/retrospectives/retro-1/Report.sso

added element 'replist2' to list

reading output fromC://users//michelle.sculley//documents//2023 SWO ASSESS//ModelDev//Current Best/Diagnostics/Retros/Retro\_full/retrospectives/retro-2/Report.sso

added element 'replist3' to list

reading output fromC://users//michelle.sculley//documents//2023 SWO ASSESS//ModelDev//Current Best/Diagnostics/Retros/Retro\_full/retrospectives/retro-3/Report.sso

added element 'replist4' to list

reading output fromC://users//michelle.sculley//documents//2023 SWO ASSESS//ModelDev//Current Best/Diagnostics/Retros/Retro\_full/retrospectives/retro-4/Report.sso

added element 'replist5' to list

reading output fromC://users//michelle.sculley//documents//2023 SWO ASSESS//ModelDev//Current Best/Diagnostics/Retros/Retro\_full/retrospectives/retro-5/Report.sso

added element 'replist6' to list

Summarizing 6 models:

imodel=1/6

N active pars = 94

imodel=2/6

N active pars = 94

imodel=3/6

N active pars = 94

imodel=4/6

N active pars = 94

imodel=5/6

N active pars = 94

imodel=6/6

N active pars = 94

Summary finished. To avoid printing details above, use 'verbose = FALSE'.

Warning in min(x): no non-missing arguments to min; returning Inf

Warning in max(x): no non-missing arguments to max; returning -Inf

Warning in min(x): no non-missing arguments to min; returning Inf

Warning in max(x): no non-missing arguments to max; returning -Inf

Warning in min(x): no non-missing arguments to min; returning Inf

Warning in max(x): no non-missing arguments to max; returning -Inf

Index Season MASE MAE.PR MAE.base  
1 S1\_JPN\_WCNPO\_OSDWLL\_early\_Area1 1 NA NA NA  
2 S2\_JPN\_WCNPO\_OSDWCOLL\_late\_Area1 1 2.2260081 0.4270844 0.1918611  
3 S3\_JPN\_WCNPO\_OSDWLL\_early\_Area2 1 NA NA NA  
4 S4\_JPN\_WCNPO\_OSDWLL\_late\_Area2 1 0.8542094 0.1507977 0.1765348  
5 S5\_TWN\_WCNPO\_DWLL\_late 3 1.8995232 0.4318851 0.2273650  
6 S6\_US\_WCNPO\_LL\_deep 2 1.7588687 0.2122219 0.1206582  
7 S7\_US\_WCNPO\_LL\_shallow\_early 1 NA NA NA  
8 S8\_US\_WCNPO\_LL\_shallow\_late 1 1.7200811 0.1981002 0.1151691  
 MASE.adj n.eval  
1 NA 0  
2 2.2260081 5  
3 NA 0  
4 0.8542094 5  
5 1.8995232 5  
6 1.7588687 5  
7 NA 0  
8 1.7200811 5

|  |
| --- |
| Figure 30: Hind casting cross-validation (HCxval) results for Japanese longline late area 1 (top right), Japanese LL late area 2 (top left), Chinese Taipei deep water longline late (center right), US Hawaii deep-set longline (center left) CPUE, and US Hawaii deep-set longline (bottom left) fits, showing observed (large points with dashed line), fitted (solid lines), and one-year-ahead forecast values (small terminal points) in the old growth model. The observations used for cross-validation are highlighted as color-coded solid circles with associated 95% confidence intervals (light-grey shading). The model reference year refers to the endpoint of each one-year-ahead forecast and the corresponding observation. The mean absolute scaled error (MASE) score associated with each CPUE time series is denoted in each panel. |

Warning: combining data from multiple seasons  
Warning: combining data from multiple seasons  
Warning: combining data from multiple seasons  
Warning: combining data from multiple seasons  
Warning: combining data from multiple seasons  
Warning: combining data from multiple seasons

Warning in qlnorm(0.975, meanlog = log(y), sdlog = indexSEvec): NaNs produced

Warning in qlnorm(0.025, meanlog = log(y), sdlog = indexSEvec): NaNs produced

Warning in qlnorm(0.975, meanlog = log(y), sdlog = indexSEvec): NaNs produced

Warning in qlnorm(0.025, meanlog = log(y), sdlog = indexSEvec): NaNs produced

Warning in min(x): no non-missing arguments to min; returning Inf

Warning in max(x): no non-missing arguments to max; returning -Inf

Warning in min(x): no non-missing arguments to min; returning Inf

Warning in max(x): no non-missing arguments to max; returning -Inf

Warning in qlnorm(0.975, meanlog = log(y), sdlog = indexSEvec): NaNs produced

Warning in qlnorm(0.025, meanlog = log(y), sdlog = indexSEvec): NaNs produced

Index Season MASE MAE.PR MAE.base  
1 F1\_JPN\_WCNPO\_OSDWCOLL\_late\_Area1 1 1.5857805 0.03010630 0.01898516  
2 F2\_TWN\_WCNPO\_DWLL\_late 1 0.8512948 0.04041926 0.04747975  
3 F3\_US\_WCNPO\_LL\_shallow\_late 2 0.4687310 0.03793153 0.08092387  
4 F4\_IATTC 2 NA NA NA  
5 F6\_JPN\_WCNPO\_OSDWLL\_early\_Area1 1 NA NA NA  
6 F7\_JPN\_WCNPO\_CODF 2 1.5758234 0.01912729 0.01213796  
7 F8\_US\_WCNPO\_LL\_shallow\_early 1 NA NA NA  
8 F9\_US\_WCNPO\_LL\_deep 3 0.6216355 0.08359865 0.13448177  
 MASE.adj n.eval  
1 0.3010630 5  
2 0.4041926 5  
3 0.3793153 4  
4 NA 0  
5 NA 0  
6 0.1912729 1  
7 NA 0  
8 0.6216355 5

|  |
| --- |
| Figure 31: Hind casting cross-validation (HCxval) results for size composition mean lengths, showing observed (large points with dashed line), fitted (solid lines), and one-year-ahead forecast values (small terminal points) in the old growth model. The observations used for cross-validation are highlighted as color-coded solid circles with associated 95% confidence intervals (light-grey shading). The model reference year refers to the endpoint of each one-year-ahead forecast and the corresponding observation. The mean absolute scaled error (MASE) score associated with each size composition time series is denoted in each panel. |

The likelihood profile indicated that the US data componenets (CPUE and size composition data) are majority contributors to the likelihood. Attempts to downweight the US LL size composition data results in a model that fails to converge (due to a Hessian that is not positive definite). However, the ASPM models indicate that the US CPUE data suggest a similar trend as the other CPUE indices. To further explore the effect of the US data, models were run removing all US data except catch and fitting the parameters. While the models including the Japanese size data also failed to converge, the results for all the models indicated an estimated ln(Ro) between 6.2 and 6.4, and SSBzero between 12500 and 15500 metric tons (Table 6). This suggests that even though the US data is a significant contributor to the likelihood, the estimated population size would be similar without the data.