Western and Central Pacific Fisheries Commission

SCIENTIFIC COMMITTEE
TWENTY-FIRST REGULAR SESSION

Tonga
13-21 August 2025

Summary report from the NOAA-SPC assessment model meeting on SWPO striped marlin

WCPFC-SC21-2025/SA-IP-XX
21 April 2025

N. Ducharme-Barth[1], C. Castillo-Jordán[2], M. Sculley[1], R. Ahrens[1], F. Carvalho[1]

[1]NOAA Fisheries, Pacific Islands Fisheries Science Center, 1845 Wasp Boulevard, Bldg. 176, Honolulu HI 96818
[2]Oceanic Fisheries Programme of the Pacific Community

Table of contents

1 Executive summary

This report presents outcomes from a joint Pacific Community (SPC) and NOAA Fisheries stock assessment modeling meeting held in Honolulu, Hawai'i from 20-24 January 2025. The meeting focused on addressing technical recommendations made by WCPFC SC20 concerning the 2024 Southwest Pacific Ocean (SWPO) striped marlin stock assessment.

Key issues identified by SC20 included: (1) poor fits to both size composition data and the relative abundance index, particularly in the last decade; (2) conflicts between the relative abundance index and size composition data; (3) inconsistencies in fits to length and weight composition data from the same fishery; (4) selectivity curve definition issues; and (5) inability to estimate model initial conditions.

The primary achievement of the meeting was transitioning the assessment model from MULTIFAN-CL to Stock Synthesis, which provided considerable advantages in resolving several key technical issues. This transition allowed for modeling selectivity as a function of length while maintaining annual population dynamics, which better aligns with the reproductive biology of striped marlin.

Through sequential model development and testing, several improvements were achieved:

1. **Improved initial conditions estimation**: Returning to a 1952 start year with an unfished state better addressed population initialization concerns.

2. **Enhanced selectivity modeling**: Length-based selectivity provides greater resolution.

3. **Refined data inclusion**: Selective exclusion of problematic size composition data improved model fit to remaining data components.

4. **Improved growth modeling**: Internal estimation of growth parameters using conditional age-at-length data helped address potential limitations from the externally derived growth curve.

5. **Accounting for historical catch uncertainty**: Implementation of catch uncertainty for early high-catch periods reduced sensitivity to potentially unreliable historical catch records.

Several tricky data and modeling challenges remain. These include: (1) spatiotemporal heterogeneity in fleet coverage limiting the reliability of abundance indices; (2) potential non-representativeness in mixed-fleet composition data; (3) limitations in the age data from opportunistic sampling; and (4) data weighting complexities arising from evident conflicts between data sources.

Recommendations for the final 2025 assessment include:

- Continue development in the Stock Synthesis framework
- Begin the model from 1952 assuming unfished initial conditions
- Develop multiple alternative abundance indices to evaluate sensitivity to different data sources
- Exclude problematic size composition data
- Consider ensemble modeling approaches to better characterize uncertainty
- Explore two-sex model development

These developments help address SC20's technical concerns and provide a framework for finalizing the 2025 SWPO striped marlin assessment.

2   Introduction

This report describes the outcomes and recommendations from a joint Pacific Community (SPC) and NOAA Fisheries stock assessment modeling meeting held in Honolulu, Hawai'i from 20 - 24 January 2025. Meeting activities focused on addressing technical recommendations made by SC20 (WCPFC, 2024) on the 2024 Southwest Pacific Ocean striped marlin stock assessment (Castillo-Jordan et al., 2024). For context, the 2024 Southwest Pacific Ocean striped marlin stock assessment (Castillo-Jordan et al., 2024) model, `diagnostic-2024` , is an age-structured population dynamics model implemented in MULTIFAN-CL (Fournier et al., 1998) that assumes a single, well-mixed population and approximates spatial structure using a fleets-as-areas approach. Major changes from the 2019 the Southwest Pacific Ocean striped marlin stock assessment (Ducharme-Barth et al., 2019) model, `diagnostic-2019` , include moving from a catch-errors to a catch-conditioned treatment of fishing mortality, and updating key biological assumptions (e.g., growth).

The SC20 identified the following technical issues needing further attention and requested that these issues be addressed prior to presenting the 2025 Southwest Pacific Ocean striped marlin stock assessment to SC21:

- lack of fit to either the size composition data or the relative abundance index (particularly in the last decade),
- conflict between the relative abundance index and the size composition data,
- inconsistencies in the fit to length and weight composition data from the same fishery,
- how selectivity curves are defined and,
- lack of ability to estimate model initial conditions.

It is in this context that the United States offered to host SPC staff for a modeling meeting and to provide technical assistance in the redevelopment of the Southwest Pacific Ocean striped marlin stock assessment (WCPFC, 2024). The meeting was held in person at the University of Hawai'i - Manoa campus with one participant from SPC and four participants from the NOAA Fisheries Pacific Islands Fisheries Science Center (PIFSC). The attendees are listed in the Appendix.

3   Model explorations

3.1   Pre-meeting

Following SC20 and prior to the modeling meeting, SPC continued development on the MULTIFAN-CL Southwest Pacific Ocean striped marlin stock assessment model to try and improve the fit to the size composition data. These additional investigations focused on how selectivity was defined in the model. In the 2024 Southwest Pacific Ocean striped marlin stock assessment (Castillo-Jordan et al., 2024), cubic-splines were used to define flexible selectivity curves as a function of age noting that implementation in MULTIFAN-CL restricts placement of nodes to age class values. Since the 2024 assessment assumed 10 annual age classes this placed a severe constraint on the shape of the selectivity curve (Figure 1), particularly at the youngest ages when growth was assumed to be quite rapid. A new model `mfcl-1979-20p3` (Figure 2) was reparametrized in MULTIFAN-CL using quarterly age in order to provide greater resolution in age-specific selectivity at the youngest sizes (Figure 3). The new quarterly age model also included a separate index for each of the four quarters which shared a stationary catchability, and selectivity curve. Additionally, time-varying observation error for each index was not activated and instead assumed a $\sigma = 0.17$.

While this new model `mfcl-1979-20p3` had the desired effect of improving the resolution of selectivity and fits to size composition data, it introduced a new problem into the population dynamics. Southwest Pacific Ocean striped marlin are not known to spawn year-round like tropical tunas, but rather have a restricted spawning season in the late austral spring and early austral summer where they form multiple discrete spawning aggregations (Kopf et al., 2012). However, in MULTIFAN-CL it is difficult to restrict recruitment in a quarterly model to occur in a single quarter of the calendar year. An orthogonal polynomial parametrization for recruitment was applied to restrict recruitment to the first quarter of the calendar year (following spawning in the early austral summer), however the model still introduced recruits into the model in all quarters of the year (Figure 4).

3.2 January 2025 meeting

Model investigations during the January 2025 stock assessment modeling meeting used the `mfcl-1979-20p3` MULTIFAN-CL model provided by SPC as a starting point. A table of key model convergence and fit metrics is provided in Table 1.

Model configuration files, output, and analytic code are available on GitHub[1] : https://n-ducharmebarth-noaa.github.io/2025-swpo-mls-meeting/.

3.2.1 Transition to Stock Synthesis

Discussions in the lead up to the January 2025 meeting identified that switching the modeling framework from MULTIFAN-CL to Stock Synthesis (Methot & Wetzel, 2013) could help resolve some of the issues identified by SC20 that the `mfcl-1979-20p3` model sought to address without compromising key population dynamics assumptions (e.g., annual population dynamics and recruitment). One of the key benefits to Stock Synthesis is that it can model selectivity as a function of length and/or age. This improves the resolution of the selectivity curve and fit to the size data while maintaining annual population dynamics.

An initial Stock Synthesis model `01-mls-base-1979` was configured to approximate the available MULTIFAN-CL `mfcl-1979-20p3` model. `01-mls-base-1979` implements a single-sex Stock Synthesis model spanning 1979-2023 with *annual* age-structure and *quarterly* population and fishing dynamics. An early recruitment period from 1969-1978 was used to set-up the initial population age structure, and the initial level of fishing mortality (`initF` [2]) was fixed to match the `mfcl-1979-20p3`. Fishery definitions and selectivity groupings were maintained with selectivity defined as a function of length assuming a double-normal functional form. All length and weight[3] composition were reformatted using the MULTIFAN-CL bin structure. To initialize the model, the input effective sample sizes for the weight and length composition matched the sum of the input observations from the `mfcl-1979-20p3` for each quarterly fishing instance. Rather than use the four quarterly indices, a single quarterly index corresponding to quarter 4 was used in the Stock Synthesis model. Quarter 4 was selected as this was the one most likely to represent changes in the spawning

---

[1]This repository will be made public following the 2025 SPC Pre-Assessment Workshop, noting that these represent preliminary model explorations and should not be used as the basis for management advice.

[2]The `mfcl-1979-20p3` model used total mortality, Z to set initial population depletion and conditions where $Z = 1.7 \times M$, M being natural mortality. In Stock Synthesis initial population conditions are set using initial fishing mortality F or `initF`. Initially `initF` was incorrectly set in `01-mls-base-1979` as $1.7 \times M$. `intF` was subsequently corrected to be $0.7 \times M$ in `25-mls-base-1979-corrected`.

[3]Weight bins $> 139kg$ were aggregated from $2kg$ to $4kg$ bins as in Stock Synthesis they can not be smaller than the length bins.

component of the population given the temporal reproductive dynamics described by Kopf et al. (2012). The time-varying observation error specified, but not used, in the `mfcl-1979-20p3` model were applied in this case.

Biological parameters were converted from the `mfcl-1979-20p3` model. Growth followed a von Bertalanffy function characterized by explicit parameter values: $L_1 = 88.38$ cm at age 0.25 years, $L_2 = 210.96$ cm at age 10 years, and $k = 0.84$, with CV structure following a length-at-age pattern. Reproduction dynamics incorporated a Beverton-Holt stock-recruitment relationship with steepness fixed at 0.8 and recruitment variance (σ_R) set to 0.2, while the maturity ogive was logistically parameterized with length at 50% maturity at 181.21 cm and a slope parameter of -0.20.

Initial results from the `01-mls-base-1979` (or `25-mls-base-1979-corrected`) models did not show particularly close agreement with the `mfcl-1979-20p3` temporal dynamics or spawning biomass scale, however estimates of population numbers were more comparable (Figure 5). However, this model was not expected to show good agreement as it assumed naïve sample sizes for the size composition data and the difference in treatment of observation error for the index. Most importantly, the model ran and could form the basis for further refinements to model structure in order to improve the fit to various data components.

3.2.2 Early refinements

Following the successful run of the `01-mls-base-1979` model, a pair of follow-on Stock Synthesis models were developed in a stepwise manner to address model fit and statistical treatment of data inputs. These modifications resulted in the `02-mls-ss3-1979` and `03-chg-selex-1979` model configurations.

The `02-mls-ss3-1979` modifications focused on improving model performance while preserving the fundamental assumptions of the baseline model. Selectivity parameterization was modified through three key adjustments: (1) converting fixed end logit parameters to estimated parameters for fleets using double-normal selectivity; (2) fixing top logit parameters enforce dome-shaped selection where empirically supported; and (3) replacing the New Zealand recreational fishery's double-normal selectivity with a simpler logistic function. These changes provided greater flexibility in fitting length composition data while maintaining biologically plausible selection patterns. Statistical treatment of composition data was refined by increasing the minimum tail compression value to 1e-02 and implementing fleet-specific effective sample size adjustments (down-weighting factors of 0.03-0.10) to match the level of down-weighting seen in the `mfcl-1979-20p3` model. Additionally, the CPUE index error structure was augmented with an observation error component (0.191) derived from loess smoother residual analysis. These modifications collectively addressed potential over-dispersion in composition data and explicitly accounted for potential un-modeled observation error in the standardized index.

Following the implementation of the `02-mls-ss3-1979` model, further adjustments focusing specifically on selectivity were explored in the `03-chg-selex-1979` model. For Fishery 2 (Japanese longline in sub-region 2), Fishery 9 (Australian recreational), Fishery 12 (aggregated longline in sub-region 2), and the standardized abundance index (Fleet 15), the top logit parameters were changed from fixed values to estimated parameters. Concurrently, for these same fleets, the end logit parameters were reverted from estimated parameters to fixed values (-999) to maintain parameter identifiability. This adjustment effectively allowed the peak of the selectivity curve to be estimated while allowing its right-hand tail behavior to decline gradually.

Modifiying selectivity modeling and size composition data weighting in the `02-mls-ss3-1979` model improved the fit to the index (Figure 6) and resulted in somewhat more reasonable selectivity curves. However, issues remained, noticeably for Fleets 2 (Japanese LL; sub-region 2) and 15 (Index). Further tweaking the selectivity in model `03-chg-selex-1979` degraded the fit to the index slightly but resolved some of the existing selectivity issues and generally improved fit to length (Figure 8) and weight (Figure 9) data.

3.2.3 Reverting model start year to 1952

One of the issues raised at SC20 was the lack of ability of the `diagnostic-2024` model to estimate initial conditions in 1979, rather than begin the model at an unfished equilibrium in 1952 as was done in the `diagnostic-2019` model. Initial conditions in 1979 in the `diagnostic-2024` model were tuned (assuming initial total mortality Z to be $1.7 \times M$, natural mortality) to match the level of depletion (~ 0.35) seen in 1979 from the `diagnostic-2019` model. Rationale given for changing the start year to 1979 in the `diagnostic-2024` model was to include vessel random effect in the standardized index, and to avoid the issue of implausibly high estimated recruitments in the early years that arose after switching to the updated growth curve.

Following the switch to Stock Synthesis, a new model beginning in 1952 `04-start-1952` was developed from the `03-chg-selex-1979` model to address the initial conditions issue raised with the `diagnostic-2024` model, and to see if the alternative selectivity formulation as length-based alleviated the issue of implausibly high estimated recruitments in the early years. The extended model incorporates historical catch data from 1952-1978 extracted from the `diagnostic-2019` model. The only size composition data that was included was the New Zealand recreational (Fleet 10) weight composition data. Given that this fishery assumed a logistic selectivity shape, observed declines in average weight could provide information about adult abundance trend and total mortality Z during the years prior to the start of the standardized index. Treatment of the initial condition (e.g., no longer including steepness effects in the initial equilibrium calculations and removal of `initF` parameter) was also changed to accommodate starting at an unfished condition. Lastly, the recruitment dynamics were reconfigured to accommodate the longer time series. Specifically, the recruitment deviation framework was expanded to span 1952-2021 (recruitment in the last year of the model was not estimated), with early recruitment deviations beginning in 1942 and bias adjustment parameters recalibrated for the extended time series.

Switching back to a 1952 start from an unfished condition in the `04-start-1952` model resulted in dynamics that were much similar to the `diagnostic-2019` model (Figure 10), especially in terms of more recent estimates of spawning biomass scale and depletion. The change in how recruitment is modeled between Stock Synthesis and MULTIFAN-CL with the use of the early recruitment period and recruitment bias-adjustment ramp eliminates the large spike in recruitment seen in the `diagnostic-2019` model and overall recruitment variability in the `04-start-1952` is moderated relative to the two MULTIFAN-CL models. Changing to the `04-start-1952` model did degrade fit to the index in later years (Figure 11).

3.2.4 Excluding size composition data

The next phase of developments considered the effect of removing size composition data from the model that were believed to be unsuitable for use in the model for either data quality, or sampling concerns. Prior to the meeting start, the SPC spent considerable time reviewing the size composition data for each fishery and provided guidance (P. Hamer, *pers. comm.*) on which data should be excluded. Based on this information,

two additional models `05-exclude-bad-comp` and `06-exclude-more-comp` were sequentially developed from `04-start-1952` to remove un-representative data, minimize data conflicts, and improve the fit to the remaining data (index and size composition). Note that these exclusions were largely exploratory and does not necessarily reflect final decisions on the inclusion/exclusion of data for a proposed 2025 stock assessment model. The cumulative changes leading to the `06-exclude-more-comp` are described below.

Length composition changes

- Fleet 1 (JP LL; sub-region 1): Eliminated all length composition due to small sample sizes and temporal patchiness.
- Fleet 2 (JP LL; sub-region 2): Eliminated all length compositions due to conflict with corresponding weight data and temporal patchiness.
- Fleet 6 & 7 (AU LL): Excluded all length compositions due to internal conflict with corresponding weight composition data and temporal patchiness relative to more comprehensive weight data.
- Fleet 8 (NZ LL): Removed all length compositions based on inadequate sample sizes and temporal patchiness.
- Fleet 11 (Aggregate LL; sub-region 1): Targeted removal of 2006-2014 length composition data to exclude TWDW data which may have data quality issues.
- Fleet 12 (Aggregate LL; sub-region 2): Eliminated all length compositions. Mixed-flag fishery compositions are spatiotemporally patchy and may fail to representatively capture aggregate fishery dynamics despite potentially having adequate sampling of individual fleet components.
- Fleet 13 (Aggregate LL; sub-region 3): Same as for Fleet 12.
- Fleet 14 (Aggregate LL; sub-region 4): Removed pre-2006 data that demonstrated anomalously small size observations.
- Fleet 15 (Index): Complete removal of length composition due to conflict with weight composition data and evidence of less comprehensive temporal sampling compared to the weight data.

Weight composition changes

- Fleet 1 (JP longline; sub-region 1): Removed all weight composition data due to small sample sizes and temporal inconsistencies
- Fleet 2 (JP longline; sub-region 2): Specifically removed 1979 weight composition data due to apparent sampling anomalies
- Fleet 10 (NZ recreational): Excluded post-1988 weight composition due to management-induced changes in size selectivity (encouraging non-retention of small individuals) that may bias post-1988 samples

Selectivity changes

The selectivity grouping for Fleet 12 was shared with Fleet 2, given the removal of all composition data for Fleet 12 (see Table 2 for updated fishery definitions).

Fits to the index (Figure 12) as well as the remaining composition data (length: Figure 13 and weight: Figure 14) were predictably improved when un-representative and conflicting composition data were removed.

3.2.5 Catch uncertainty

One of the concerns raised during development of the `diagnostic-2024` model was the influence of large catch event by the Fleet 2 (Japanese LL; sub-region 2) in 1954. This was also a concern in the `diagnostic-2019` model, and sensitivity to the starting year was explored in that previous assessment to exclude that large catch observation (Ducharme-Barth et al., 2019). A different approach was taken in the current explorations, and the `07-catch-uncertainty` was developed to account for potential uncertainty in the historical catch.

The `07-catch-uncertainty` model built on the `06-exclude-more-comp` model and bifurcated Fleet 2 (Japanese LL; sub-region 2) into two temporal components. This created a distinct fleet (F15_LL.JP.2_early) for pre-1979 data while maintaining the original fleet for more recent observations. Selectivity was shared between the two fleets. However, moving to Stock Synthesis `F_Method = 4` allowed some fleets (Fleets 1 - 14) to stay defined as catch-conditioned, where fishing mortality is calculated to fit the catch exactly, and have Fleet 15 estimate the fishing mortality (F) needed to fit to the catch with observation error. This structural modification also enabled differential parameterization of catch uncertainty, with higher standard errors (SE=0.2) assigned to quarters with higher catch observations (quarters 3-4) and lower uncertainty (SE=0.05) to quarters with lower catch observations (quarters 1-2). Additionally, the catch-conditioned F values for Fleet 15 were used as initial values in the estimation of F. The index fishery stayed the same but became Fleet 16. Parameterizing the model in this way allowed for potential uncertainty in historical catch to be propagated in model estimates in a computationally efficient manner.

Adding in catch uncertainty (Figure 15) did not greatly change time series estimates of key stock assessment outputs (Figure 16) though it did appear to improve fits to the index (Figure 17).

3.2.6 Treatment of growth

Using the `07-catch-uncertainty` as a baseline, a number of additional models were developed to explore alternative ways of dealing with growth. Growth was a major source of discussion as some of the issues identified in development of the `diagnostic-2024` model became more pronounced after switching to a new growth curve provided by Farley et al. (2021), with the new curve showing exceptionally rapid growth through the first year of life. This was flagged as a concern given the opportunistic nature of sampling for the otoliths from longline landed individuals which has reduced selectivity at smaller sizes and would only capture the most rapid growing individuals in the youngest age classes. In an effort to improve model fits and alleviate any potential mis-specification in the growth curve, two classes of models were developed.

Models `08-relax-growth` and `09-relax-growth-v2` attempted to improve the fit by defining variability in growth using CV as a function of length at age, rather than as a standard deviation, and setting the CV for yound and old fish to be 0.15. In the `08-relax-growth` model L_1 was set to 59.9 cm and assumed an age at L_1 of 0, while in the `09-relax-growth-v2` model L_1 was set to 88.3822 cm and assumed an age at L_1 of 0.25. These models ended up describing functionally identical growth curves, though it was initially unclear from the Stock Synthesis documentation that this would be the case.

Models `10-CAAL`, `11-CAAL-no-ageerr`, `12-CAAL-old-growth-SD`, and `13-CAAL-noAgeerr-OGsd` all attempted to improve fit by estimating growth internally using the aging data from the Farley et al. (2021) study as conditional age-at-length data. Data were associated with the appropriate fleet definition based on sampling location and flag of vessel sampled from. Though the aging data recorded decimal age, age was

entered in the model as the nearest model age less than or equal to the decimal age. Given that model ages were annual, this meant integer age values. For all models L_1, L_2 and k were freely estimated.

Models `10-CAAL` and `11-CAAL-no-ageerr` both defined variability in growth using CV as a function of length at age in the same way as models `08-relax-growth` and `09-relax-growth-v2`. Model `10-CAAL` included an aging error matrix with standard deviations of aging error standard deviation ranging from $\sigma_{Age} = 0.35$ at age 0 to $\sigma_{Age} = 2$ at age 10. This was done to account for both actual aging error and for converting from decimal age to integer age to match the model structure.

Models `12-CAAL-old-growth-SD` and `13-CAAL-noAgeerr-OGsd` both defined variability in growth using standard deviation as a function of length at age in the same way as model `07-catch-uncertainty`, where the standard deviations at the youngest and oldest model ages were ~ 10. Model `12-CAAL-old-growth-SD` included the same aging error matrix as `10-CAAL` while model `13-CAAL-noAgeerr-OGsd` did not.

Estimation of growth internally using conditionaly age-at-length data and size composition slowed growth (lower k) and increased L_2 (Figure 18), which had a small effect on recent time-series estimates (Figure 19). It also improved fits to the index (Figure 20), length composition (Figure 21) and weight composition (Figure 22) overall though some fisheries noticeably deteriorated (e.g., the `S01_INDEX` weight composition data). We note that although the variability in length at age between the formulations using CV vs. standard deviation were approximately the same, the models showed greater stability when variability in growth was defined using a standard deviation as a function of length at age.

3.2.7 Additional investigations

The `12-CAAL-old-growth-SD` model emerged from this process as a reasonable candidate from which a numer of additional investigations were attempted. These are briefly described below.

- `14-CAAL-target-earlyF`: Built off `13-CAAL-noAgeerr-OGsd` to implement a targeted approach to historical catch uncertainty by specifically isolating Japanese longline catches from 1954-1978 with values greater than 1 into a separate fleet (Fleet 15) with elevated uncertainty (0.2), while maintaining lower uncertainty (0.01) for all other fleet 2 catches. This refinement built directly upon `13-CAAL-noAgeerr-OGsd` but applied a more focused treatment of uncertainty to historical catch. This had negligible impacts on estimated quantities.
- `15-CAAL-qtrAge`: Attempted to input the conditional age-at-length data as quarterly ages but this was unsuccessful.
- `16-CAAL-rm-spike`: This model modified `12-CAAL-old-growth-SD` to remove the index observation in 1998 to see if attempting to fit this high point was driving the misfit to the index seen in 2000 - 2002. Assessment results with that observation removed showed little meaningful difference with the `12-CAAL-old-growth-SD` model results indicating that it is not likely to be the source of the index misfit.
- `17-CAAL-noAU-ASPM`: This model modified `12-CAAL-old-growth-SD` to keep the selectivity values for Fleet 6 (Australian LL; sub-region 2) fixed at the values estimated from model `12-CAAL-old-growth-SD` and removed all of the Fleet 6 weight composition data from the likelihood to see if the weight composition data from this fishery was driving the misfit to the index seen in 2000 - 2002. Fit to the index was substantially improved without the inclusion of this data indicating that conflict with this data source is likely the cause of the misfit.

10

- `18-CAAL-cUnc-cv40` : Built directly off of `12-CAAL-old-growth-SD` to evaluate assessment sensitivity to increased historical catch uncertainty by doubling the standard error values for Fleet 15 (early Japanese longline) from 0.2 to 0.4 (seasons 3-4) and from 0.05 to 0.1 (seasons 1-2). This targeted modification tested the sensitivity of model results to assumptions on historical variability in catch. This had a slight impact to biomass estimates in the early model period showing some sensitivity to this assumption.
- `19-CAAL-1979` : This model attempted to merge `03-chg-selex-1979` with `12-CAAL-old-growth-SD` by begining the model in 1979 with fixed `initF` , excluding all of the non-representative size composition data, and including the estimation of growth using conditional age-at-length data with an aging error matrix. This model ran but produced biomass estimates that were almost double `03-chg-selex-1979` .
- `20-CAAL-1979-estInitF` : Structurally identical to `19-CAAL-1979` except that `initF` is estimated. This model was unsuccesful as biomass estimates became unrealistically large.
- `21-CAAL-NZrecwtQtr` : Built off `12-CAAL-old-growth-SD` and re-allocated the New Zealand recreational fishery weight composition data from annual observations in the first quarter of the calendar year to the quarterly periods in which they were observed. This is a more appropriate treatment of the data, but did not result in percievable differences in model estimates.
- `22-1979-estInitF-v2` : Built off `03-chg-selex-1979` and successfully estimated `initF` . Time series estimates of recruitment, depletion, and spawning biomass appeared similar to the `2024-diagnostic` model.
- `23-CAAL-2sex` : This model implemented a 2-sex version of the `12-CAAL-old-growth-SD` model and estimated sex-specific growth as a proof of concept model. Other biological parameters were assumed to be equivalent between sexes for the sake of this example model. The model ran and produced similar estimates to the single-sex model however with a higher than desired maximum gradient indicating that estimating sex-specific growth internal to the model may be ambitious.
- `24-CAAL-Richards` : This model implemented a Richards growth model version of the `12-CAAL-old-growth-SD` model. Growth was very slightly more rapid with the estimation of a fourth growth parameter but stock status estimates were largely unchanged.

The key take-aways from these investigations were:

- the misfit to the index seen in 2000 - 2002 is driven by conflict with the Fleet 6 (Australian LL; sub-region 2) weight frequency data which suggests an increasing trend in mean size over the period 1998 - 2002, rather than a decline as seen in the index (Figure 23),
- a two sex model is technically feasible though estimation of sex-specific growth parameters may not be (Figure 24 & Figure 25),
- growth estimates are slightly sensitive to the functional form used, either von Bertalanffy or Richards (Figure 26),
- re-allocating New Zealand recreational fishing weight composition data to quarters is more correct but doesn't impact results (Figure 27) and,
- model estimates are slightly sensitive to the choice of uncertainty given to the early period catch (Figure 28 & Figure 29).

Single-sex model development could progress using either `12-CAAL-old-growth-SD` or `21-CAAL-NZrecwtQtr` as a starting point. Two-sex model development could progress using `23-CAAL-2sex` as a starting point noting that may need to be modified to match the quarterly treatment of New Zealand recreational fishery

weight composition data used in `21-CAAL-NZrecwtQtr`.

4 Challenges

This modeling meeting offered the opportunity to reflect on and identify several significant technical and analytical challenges that exist for the SWPO striped marlin stock assessment. These challenges span data availability, sampling representativeness, and model configuration concerns that collectively affect the ability to accurately estimate population status and dynamics.

4.1 Abundance index construction challenges

A fundamental constraint in developing a reliable abundance index is the absence of any single fleet that samples consistently across all spatiotemporal strata within the model domain. Historical fishing operations exhibit significant heterogeneity in spatial coverage, with distinct patterns by flag (Figure 30):

- Japanese longline vessels initially operated throughout regions 1-3 but have progressively retreated, maintaining operations primarily in region 1 and high-seas pockets of regions 2 and 3.
- Taiwanese longline vessels operate predominantly in regions 1 and 4.
- Australian longline vessels operate exclusively within their national EEZ in regions 2 and 3.

This spatial heterogeneity creates a situation where the abundance information provided by any standardized index derives from different spatial strata during different time periods. Without sufficient spatiotemporal overlap between fleets to properly estimate contrasts, systematic biases could be introduced into the abundance trends. This is particularly problematic if localized population dynamics exist, as the varying spatial coverage could potentially detect different segments of the population over time rather than providing a consistent measure of overall abundance.

Furthermore, this spatial heterogeneity creates potential conflicts between the index and the most informative size composition data, which in recent periods comes primarily from Australian fleets operating in regions 2 and 3. These data-type conflicts could manifest as competing signals in the likelihood function, creating challenges for appropriate model fitting and parameter estimation.

Lastly, the historical period lacks informative covariates, particularly vessel identifiers, prior to 1979 for Japanese vessels that were the primary fleet operating in the region. Without these covariates, there is reduced ability to account for vessel-specific catchability effects in the abundance standardization process.

4.2 Size composition data limitations

The size composition data present several technical challenges that affect their integration into the assessment model. Mixed-fishery composition data are particularly problematic due to their spatiotemporally patchy nature (P. Hamer, *pers. comm.*). Even when individual fleets are representatively sampled, these samples may not accurately characterize the aggregate fishery dynamics, especially when the proportion of catch by different fleets changes over time.

For a fast-growing species like striped marlin, variability in composition data can be artificially inflated relative to the true population variability, particularly for younger individuals. This inflation could occur when samples are combined across broad spatial areas within quarterly time steps if underlying migratory

patterns exist, which is believed to be the case for this species. Such artificial inflation of variability can lead to model misspecification and consequent poor fits to composition data, potentially biasing key model parameters including selectivity, recruitment, and fishing mortality.

4.3   Age data constraints

While the aging data available for SWPO striped marlin represents the best available information for the region and species, it comes with important limitations that affect growth parameter estimation. The sample size is relatively small and derives from opportunistic collection rather than a systematic sampling design, raising questions about how representative these data are of the overall population.

A particular concern is the potential for sampling bias in younger fish, where samples are likely to represent only the fastest-growing individuals within each age class due to size-selective sampling methods. This sampling bias can lead to significant overestimation of growth parameters, particularly L_1 and k. Such biases in growth estimation have cascading effects throughout the assessment, potentially affecting estimates of natural mortality, maturity-at-age, selectivity, and ultimately stock status and reference points. Estimating growth parameters internally using the conditional age-at-length data does not fully resolve this issue.

4.4   Data-weighting considerations

The presence of evident conflicts between different data sources necessitates careful consideration of data-weighting approaches. Appropriate data-weighting is crucial for balancing the model likelihood, but automated approaches (such as the Francis method) may not produce the desired results with sparse and potentially un-representative data.

For example, in the current model configuration, the weight composition data from Fleet 6 (Australian longline in region 2) dominates the likelihood. Paradoxically, application of the Francis weighting method would suggest giving this data source even more influence, potentially exacerbating rather than resolving data conflicts. This highlights the need for careful, considered approaches to data-weighting that balance statistical principles with expert judgment about data quality and representativeness.

In summary, these technical challenges collectively necessitate careful attention to uncertainty characterization in the assessment results. The model explorations described in this report represent initial steps toward addressing these challenges through alternative model configurations and data treatments, and a series of additional recommendations are made.

5   Recommendations

Based on the exploratory work completed during the modeling meeting, a series of technical recommendations were developed to guide the refinement of the SWPO striped marlin stock assessment prior to presentation at SC21. These recommendations address the primary concerns raised by SC20 while incorporating insights gained through the transition to Stock Synthesis and subsequent model explorations.

5.1 Modeling framework

The transition from MULTIFAN-CL to Stock Synthesis demonstrated clear advantages in addressing several key technical issues identified by SC20. Therefore, further development of the SWPO striped marlin assessment model is recommended to continue in the Stock Synthesis framework. This platform provides greater flexibility in defining length-based selectivity while maintaining annual population dynamics, which helps resolve the conflict between model structural assumptions and observed biological patterns of striped marlin reproduction.

5.2 Model configuration

5.2.1 Historical period and initial conditions

The exploratory models demonstrated that extending the model start year to 1952 (assuming unfished conditions) improved the ability to estimate initial population age-structure, addressing one of the key concerns from SC20. The inclusion of New Zealand recreational weight composition data, particularly from the earlier period, provided useful information on adult abundance trends during the historical period. To properly account for population dynamics throughout the assessment period, it is recommended to:

- Begin the model from 1952 assuming unfished initial conditions
- Estimate early recruitment deviations to appropriately configure the initial age structure
- Utilize available historical catch records to inform fishing mortality during the period 1952-1978

5.2.2 Growth parameterization

Growth remains a challenging component of the assessment. If model convergence allows, internal estimation of growth parameters using the conditional age-at-length data is recommended rather than fixing these values based on external information. Specific recommendations include:

- Test sensitivity to conditional age-at-length effective sample size weighting schemes
- Explore alternative ageing error assumptions
- If growth estimation proves problematic, consider a sensitivity analysis with fixed growth parameters

Additionally, the development of a two-sex model should be explored as a potential sensitivity to better capture sexually dimorphic growth patterns in striped marlin, though this should be considered a lower priority if it complicates model convergence.

5.3 Data components

5.3.1 Abundance indices

The abundance index is a critical component of the assessment, as fit to the index is prioritized. In order to properly integrate over potential uncertainties in the index, multiple alternative indices should be developed to evaluate sensitivity to different data sources and spatial configurations. A brief justification for each of these alternative indices is also provided.

- Japanese longline data only (all regions)
  - Historically, Japanese longline fishing had the largest spatial coverage (Figure 30; particularly in sub-regions 1-3) with a high concentration of effort in sub-region 2 during the austral spring and summer which coincides with known perceived striped marlin spawning aggregations (Kopf et al.,

2012). Modeling a single flag may alleviate issues that could arise when combining data across flags (e.g., catchability and selectivity differences).

- Combined Japanese/Taiwanese/Australian data (all regions)
  - As previously mentioned, though the Japanese longling fishery historically had broad spatial coverage no fishery consistently samples across the entire assessment domain or stock spatial distribution. Combining data from the three major fleets will improve spatiotemporal coverage at the risk of violating shared catchability and selectivity assumptions.
- Australian data only (utilizing the existing CSIRO index)
  - Having an index and the majority of the weight composition data come from the same fishery may reduce data conflicts within the model. However, if stock dynamics are not well mixed across the model domain this may not representatively index the entire stock.
- Japanese data only (restricted to sub-region 2)
  - A more focused version of the first option to index the spawning component of the population.
- Combined Japanese/Australian data (restricted to sub-regions 2 and 3)
  - A more focused version of the second option to index the spawning component of the population, while augmenting the spatiotemporal coverage of samples.

When combining data from multiple flags, flag identity should be incorporated as a covariate (fixed effect) in the standardization model. To better understand the sources of variance in the standardized index, particularly the 1998 spike in CPUE, it is recommended to plot nominal CPUE and all predicted quantities (standardized CPUE, encounter rate, positive catch rate) by space and time.

The integration of multiple models fitting to different individual indices in an ensemble approach should be considered to better characterize uncertainty in abundance trends. This approach would help address the concern about index-composition data conflicts by explicitly acknowledging potential inconsistencies between data sources.

5.3.2 Size composition data

Selective exclusion of problematic size composition data predictably improved model fit to the remaining data throughout the exploratory phase. The following guidelines should be applied when deciding which composition data to retain:

- Exclude data from mixed-flag fisheries where sampling is spatiotemporally patchy and there is a severe mismatch between countries providing composition samples and countries taking catch
- Remove Taiwanese length data not measured by observers due to potential quality concerns
- Exclude data that shows internal conflict between length and weight compositions from the same fishery (particularly evident in Australian and Japanese data). Only fit to one data type per fishery or selectivity grouping.
- Remove New Zealand recreational weight data following the 1988 policy change encouraging voluntary release of smaller individuals, as this could create a selectivity shift that is difficult to model.

5.3.3 Data weighting

The appropriate statistical treatment of composition data remains challenging for this assessment. The following approaches should be considered:

- Francis weighting may only be reliable for large sample sizes, in any case other assessments (e.g., ISC) only down-weight data based on Francis.
- The Dirichlet-multinomial approach may be more appropriate for fisheries with small sample sizes
- If neither approach provides satisfactory results, setting effective sample sizes based on the number of fishing trips rather than the number of individual fish sampled may provide a more realistic weighting scheme

5.4 Selectivity parameterization

The length-based selectivity approach in Stock Synthesis provides greater resolution than the age-based approach in MULTIFAN-CL, but careful consideration of parameterization is still required:

- For mixed-flag fisheries, define a selectivity that approximately removes fish at the appropriate size by either:
  - Mirroring the selectivity to a representative single-flag fishery operating in the same area
  - Selecting composition data from a single flag considered representative of the aggregated fleet
- Avoid combining data from different flags or sampling programs unless it can be demonstrated that they share similar selectivity patterns
- This recommendation presents a particular challenge for developing an appropriate selectivity curve for the abundance index when it represents a mixed-flag fishery

5.5 Further exploration and sensitivities

Several additional model configurations should be explored to fully characterize uncertainty in the assessment and evaluate the robustness of results:

5.5.1 Two-sex model

Development of a two-sex model structure would provide a more biologically realistic representation of striped marlin population dynamics, particularly given evidence of sexually dimorphic growth patterns. This approach would require: - Further refinement of sex-specific biological assumptions - Careful evaluation of sex-specific selectivity patterns - If two-sex growth estimation remains problematic, implementation with fixed sex-specific growth parameters should be considered

5.5.2 1979 model start year

While the 1952 start year is recommended for the base model, a comparative model beginning in 1979 should be developed to evaluate sensitivity to historical assumptions. This model should: - Attempt to estimate initial fishing mortality (`initF`) rather than fixing it based on external information - Provide a direct comparison with previous assessment configurations - Evaluate whether concerns about implausibly high early recruitments persist with the revised growth and selectivity parameterization

5.5.3 Growth parameterization sensitivities

Additional targeted sensitivities around growth estimation should be conducted: - Testing alternative weighting schemes for conditional age-at-length data - Exploring different ageing error matrices

These explorations will provide information about structural uncertainty in the assessment and help identify parameters or assumptions that exert significant influence on model outcomes.

6 Task Prioritization

Meeting participants developed the following suggested tasks to be prioritized for completion prior to the SPC Pre-Assessment Workshop:

1. Develop alternative relative abundance indices
   a. JP/TW/AU (all region data)
   b. JP only (only sub-region 2 data)
   c. JP/AU (only sub-region 2 & 3)
   d. JP only (all region data)
   e. AU (CSIRO index already exists)
2. Finalize updated size composition data

7 Declaration of Generative AI use

A generative artificial intelligence (AI) assistant, Anthropic Claude Sonnet 3.7, was used to parse meeting notes, and model code in the preparation of an initial draft of this report.

8 References

Castillo-Jordan, C., Day, J., Teears, T., Davies, N., Hampton, J., McKechnie, S., Magnusson, A., Peatman, T., Vidal, T., Williams, P., & Hamer, P. (2024). *Stock Assessment of Striped Marlin in the Southwest Pacific Ocean: 2024* (WCPFC-SC20-2024/SA-WP-03 (Ver. 2.01)). https://meetings.wcpfc.int/node/23120

Ducharme-Barth, N. D., Pilling, G., & Hampton, J. (2019). *Stock assessment of SW Pacific striped marlin in the WCPO* (WCPFC-SC15-2019/SC15-SA-WP-07). https://meetings.wcpfc.int/node/11232

Farley, J., Eveson, P., Krusic-Golub, K., & Kopf, K. (2021). *Southwest Pacific Striped Marlin Population Biology (Project 99)* (WCPFC-SC-17/SA-IP-11). https://meetings.wcpfc.int/node/12569

Fournier, D. A., Hampton, J., & Sibert, J. R. (1998). MULTIFAN-CL: A length-based, age-structured model for fisheries stock assessment, with application to South Pacific albacore, *Thunnus alalunga. Canadian Journal of Fisheries and Aquatic Sciences*, *55*(9), 2105–2116. https://doi.org/10.1139/f98-100

Kopf, R. K., Davie, P. S., Bromhead, D. B., & Young, J. W. (2012). Reproductive biology and spatiotemporal patterns of spawning in striped marlin *Kajikia audax. Journal of Fish Biology*, *81*(6), 1834–1858. https://doi.org/10.1111/j.1095-8649.2012.03394.x

Methot, R. D., & Wetzel, C. R. (2013). Stock synthesis: A biological and statistical framework for fish stock assessment and fishery management. *Fisheries Research, 142*, 86–99. https://doi.org/10.1016/j.fishres.2012.10.012

WCPFC. (2024). *SC20 Summary Report* (WCPFC21-2024-SC20). https://meetings.wcpfc.int/node/23997

9 Tables

Table 1: Summary of Southwest Pacific Ocean (SWPO) striped marlin assessment models. Columns correspond to model name (`Name`), start year (`Start`), number of estimated parameters (`N par`), whether the Hessian is positive-definite (`PDH`), the maximum gradient component (`MGC`), total negative log-likelihood (`Total`), the negative log-likelihood (`NLL`) for catch (`Catch`), the NLL for the index (`Index`), NLL for the length composition (`Length`), NLL for the weight composition (`Weight`), and NLL for the conditional age-at-length data (`Age`).

Name	Start	N par	PDH	MGC	Total	Catch	Index	Length	Weight	Age
01-mls-base-1979	1979	95	NA	0.00000e+00	34,595.2	0	131.205	14,032.1	19,930.4	NA
02-mls-ss3-1979	1979	94	NA	0.00000e+00	1,743.87	0.0001	-36.2442	536.606	1,113.85	NA
03-chg-selex-1979	1979	94	NA	0.00000e+00	1,417.15	0.0001	-33.4889	545.55	883.516	NA
04-start-1952	1952	121	NA	0.00000e+00	1,451.6	0.0043	-16.132	553.227	914.775	NA
05-exclude-bad-comp	1952	117	NA	0.00000e+00	988.205	0.0151	-17.2356	159.074	848.181	NA
06-exclude-more-comp	1952	117	NA	0.00000e+00	945.137	0.0122	-16.9754	127.86	836.134	NA
07-catch-uncertainty	1952	221	NA	0.00000e+00	938.389	0.8668	-28.032	127.047	842.638	NA
08-relax-growth	1952	221	NA	0.00000e+00	964.952	1.6964	-28.1126	113.733	885.742	NA
09-relax-growth-v2	1952	221	NA	0.00000e+00	967.708	1.7503	-28.4741	110.351	896.612	NA
10-CAAL	1952	224	NA	0.00000e+00	828.8	0.2057	-26.468	111.031	752.246	0.1534
11-CAAL-no-ageerr	1952	224	NA	0.00000e+00	831.152	0.5371	-25.5101	111.237	751.504	0.1975
12-CAAL-old-growth-SD	1952	224	NA	0.00000e+00	837.225	0.8048	-28.2912	107.915	755.857	0.1088
13-CAAL-noAgeerr-OGsd	1952	224	NA	0.00000e+00	837.716	0.8136	-28.3488	108.387	755.85	0.1437
14-CAAL-target-earlyF	1952	180	NA	0.00000e+00	837.668	0.8032	-28.3208	108.327	755.876	0.1443
16-CAAL-rm-spike	1952	224	NA	0.00000e+00	837.789	0.7971	-27.6842	107.972	755.946	0.1089
17-CAAL-noAU-ASPM	1952	220	NA	0.00000e+00	490.749	0.5313	-32.6067	107.487	424.924	0.1121
18-CAAL-cUnc-cv40	1952	224	NA	0.00000e+00	835.58	1.3891	-28.5947	107.699	754.594	0.1079
19-CAAL-1979	1979	93	NA	0.00000e+00	826.022	0	-20.7767	108.469	723.495	0.1123
20-CAAL-1979-estInitF	1979	97	NA	0.00000e+00	825.738	0	-20.9143	108.573	724.549	0.1147
21-CAAL-NZrecwtQtr	1952	224	NA	0.00000e+00	881.178	0.764	-28.1903	107.961	799.763	0.1089
22-1979-estInitF-v2	1979	98	NA	0.00000e+00	1,395.75	0.0006	-29.5129	554.753	846.235	NA
23-CAAL-2sex	1952	227	NA	0.00000e+00	350.06	0.326	-27.3106	104.981	275.717	0.1373
24-CAAL-Richards	1952	225	NA	0.00000e+00	880.964	0.7082	-27.9429	108.326	799.451	0.1096
25-mls-base-1979-corrected	1979	95	NA	0.00000e+00	33,891.9	0	86.4522	14,047.2	19,427.6	NA
26-CAAL-maxAge-15	1952	224	TRUE	4.38160e-03	836.63	0.7412	-27.8885	107.721	755.751	0.1154
27-CAAL-NZdblnorm	1952	226	TRUE	8.74690e-03	838.607	0.7447	-28.1002	108.664	756.843	0.1095
28-CAAL-noNZrec-ASPM	1952	222	TRUE	8.06018e-02	778.173	0.0497	-26.6697	106.312	701.581	0.1141
29-CAAL-ASPM-all	1952	187	FALSE	3.53616e-01	-60.605	0	-31.5233	0	0	0

30-CAAL-cUnc-cv2	1952	180	TRUE	6.37000e-05	833.995	1.0476	-29.0122	107.711	753.491	0.1068
31-CAAL-revASPM-01	1952	187	FALSE	3.53616e-01	-60.605	0	-31.5233	0	0	0
32-CAAL-revASPM-02	1952	187	TRUE	3.23000e-05	34.9355	0.073	-31.902	0	86.6999	0
33-CAAL-revASPM-03	1952	187	TRUE	8.10000e-06	-6.655	0.0972	-31.6097	14.028	33.8509	0
34-CAAL-revASPM-04	1952	187	FALSE	3.53616e-01	-60.605	0	-31.5233	0	0	0
35-CAAL-revASPM-05	1952	187	TRUE	2.61480e-01	-52.0766	0.0001	-31.5975	8.1764	0	0
36-CAAL-revASPM-06	1952	187	TRUE	5.36453e-02	297.05	0.1361	-27.8752	0	335.92	0
37-CAAL-revASPM-07	1952	187	TRUE	9.47000e-05	140.495	0.1077	-30.6363	0	188.943	0
38-CAAL-revASPM-08	1952	187	FALSE	3.53616e-01	-60.605	0	-31.5233	0	0	0
39-CAAL-revASPM-09	1952	187	TRUE	2.07000e-05	2.5168	0.0751	-31.5817	0	56.6072	0.0349
40-CAAL-revASPM-10	1952	187	TRUE	3.51000e-05	24.003	0.6255	-32.2372	21.7775	54.521	0.0498
41-CAAL-revASPM-11	1952	187	TRUE	2.78000e-05	-17.6667	0.0675	-30.9362	37.0394	0	0
42-CAAL-revASPM-12	1952	187	TRUE	2.46000e-05	-55.1441	0.0359	-30.7017	0	0	0.0266
43-CAAL-revASPM-13	1952	187	FALSE	3.53616e-01	-60.605	0	-31.5233	0	0	0
44-CAAL-revASPM-14	1952	187	TRUE	6.04800e-04	-28.1381	0.0118	-31.7188	28.6211	0	0
45-CAAL-revASPM-15	1952	187	FALSE	3.53616e-01	-60.605	0	-31.5233	0	0	0
46-CAAL-revASPM-16	1952	187	TRUE	2.47253e-01	-50.0407	0.001	-31.3906	0	9.09	0
47-CAAL-oneFleet-02	1952	191	TRUE	3.18000e-05	34.3446	0.046	-31.7004	0	86.2049	0
48-CAAL-oneFleet-03	1952	191	FALSE	3.46820e-02	-10.8216	0	-31.5814	13.9575	35.6651	0
49-CAAL-oneFleet-06	1952	191	FALSE	1.20274e-02	295.832	0.0394	-26.3424	0	335.535	0
50-CAAL-oneFleet-07	1952	191	TRUE	1.72800e-04	140.056	0.0389	-28.4726	0	188.151	0
51-CAAL-oneFleet-09	1952	191	TRUE	1.87700e-04	2.1876	0.0387	-31.2043	0	56.475	0.0355
52-CAAL-oneFleet-10	1952	189	TRUE	4.82200e-04	23.9808	0.5543	-32.2116	21.8754	54.4897	0.0502
53-CAAL-oneFleet-11	1952	187	TRUE	2.78000e-05	-17.6667	0.0675	-30.9362	37.0394	0	0
54-CAAL-exdblnorm	1952	217	FALSE	1.34253e+00	3,009.61	0	-13.3498	369.282	2,386.13	0.1093
55-CAAL-NZdblnorm-01	1952	224	TRUE	2.94000e-04	838.972	0.7885	-28.2607	108.725	756.998	0.1094
56-CAAL-NZdblnorm-25	1952	224	TRUE	4.07800e-04	839.549	0.7844	-28.2786	109.259	757.124	0.1096
57-CAAL-NZdblnorm-5	1952	224	TRUE	3.50500e-04	838.582	0.8175	-28.3413	108.474	756.674	0.1088
58-CAAL-NZdblnorm-75	1952	224	TRUE	2.78410e-03	838.665	0.8285	-28.3635	108.576	756.566	0.1087
59-CAAL-NZdblnorm-99	1952	224	TRUE	1.79000e-05	837.733	0.8658	-28.9994	107.488	757.041	0.1069
60-CAAL-NZdblnormV2-01	1952	219	TRUE	2.31000e-04	844.962	0.5344	-26.8388	111.05	760.821	0.1145
61-CAAL-NZdblnormV2-25	1952	219	TRUE	6.24599e-02	874.904	0.5326	-26.1835	141.969	758.363	0.1134
62-CAAL-NZdblnormV2-5	1952	219	TRUE	3.46400e-04	847.81	0.807	-27.8932	109.926	762.183	0.2186
63-CAAL-NZdblnormV2-75	1952	219	TRUE	2.42000e-05	855.298	0.8998	-30.5679	109.152	772.8	0.2194

64-CAAL-NZdblnormV2-99	1952	219	FALSE	1.55905e+04	1,021.12	0.0052	-33.055	136.515	925.633	0.1017

Table 2: Fleet definitions for model `06-exclude-more-comp`.

Fleet Number	Fleet Description	Selectivity grouping
1	Japan LL; sub-region 1	1
2	Japan LL; sub-region 2	2
3	Japan LL; sub-region 3	3
4	Japan LL; sub-region 4	4
5	Taiwan LL; sub-region 4	4
6	Australia LL; sub-region 2	5
7	Australia LL; sub-region 3	6
8	New Zealand LL; sub-region 3	6
9	Australia Recreational; sub-region 3	7
10	New Zealand Recreational; sub-region 3	8
11	Aggregate LL; sub-region 1	1
12	Aggregate LL; sub-region 2	2
13	Aggregate LL; sub-region 3	3
14	Aggregate LL; sub-region 4	4
15	Index (Aggregate LL; sub-regions 1-4)	9

10   Figures
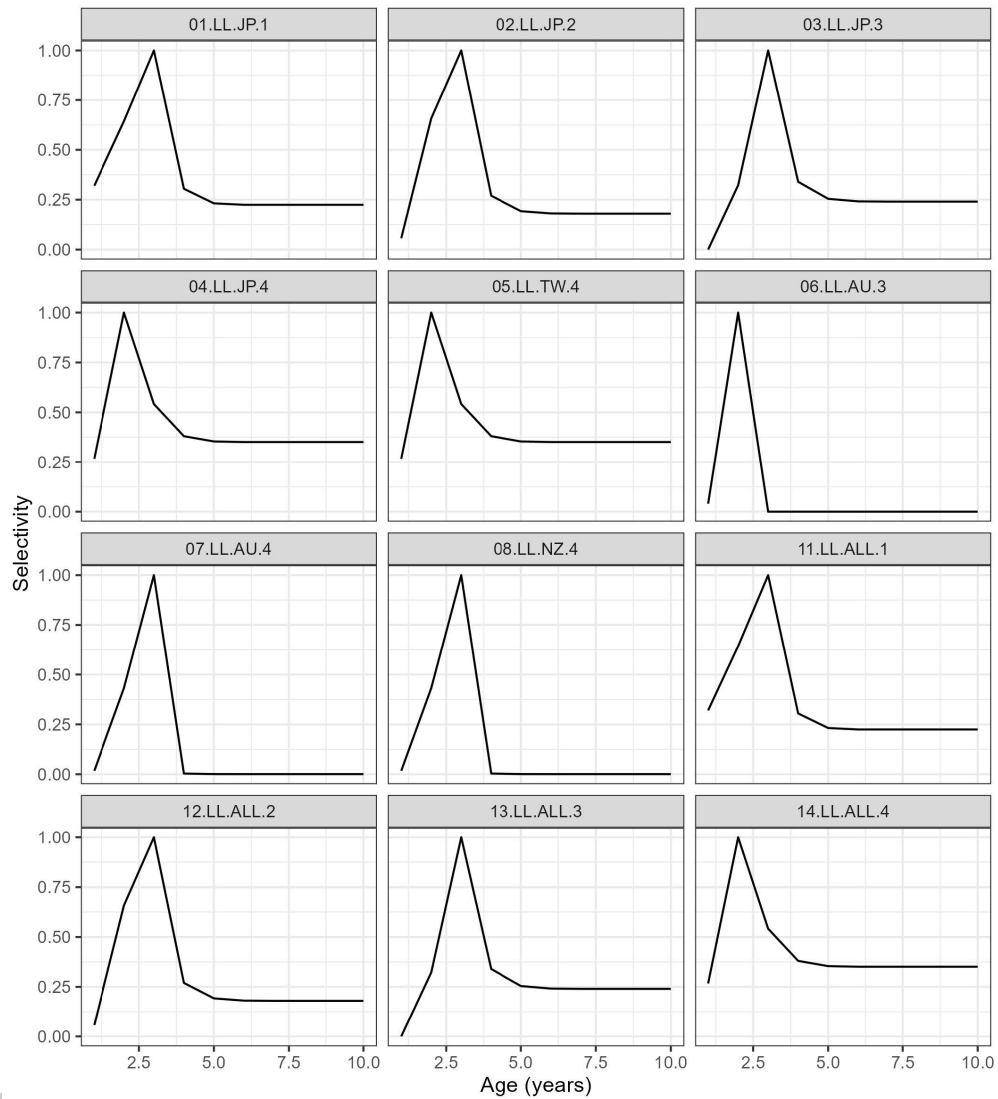
Figure 1: Age-specific fishery selectivity curves defined for longline fisheries in the `diagnostic-2024` model.
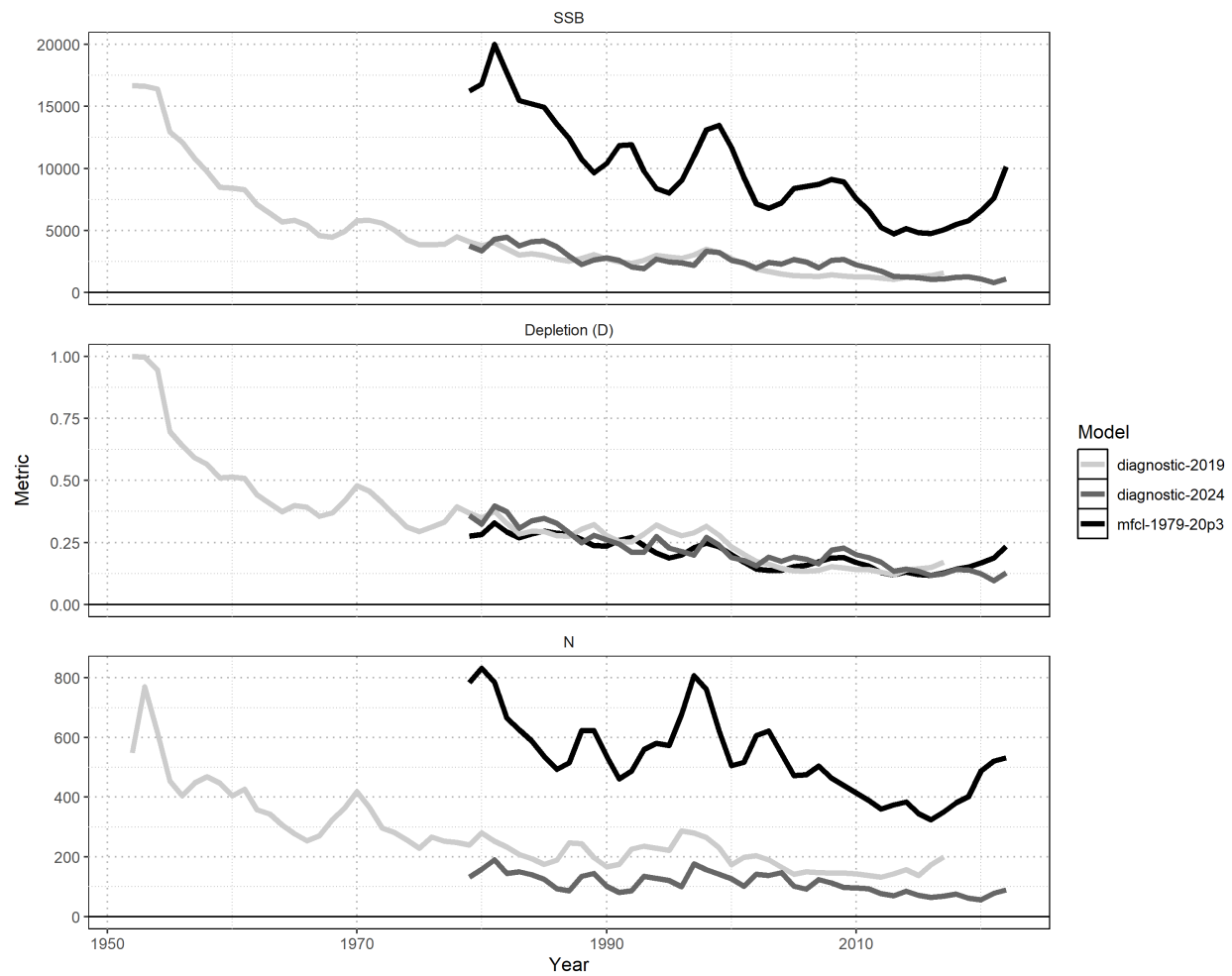
Figure 2: Time series comparison of Spawning Stock Biomass (SSB), Depletion (D), and total population numbers (N) across different models: `diagnostic-2019`, `diagnostic-2024`, `mfcl-1979-20p3`.
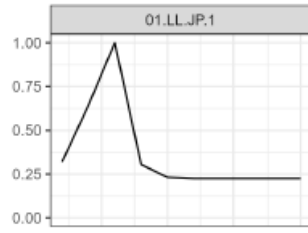
(a)
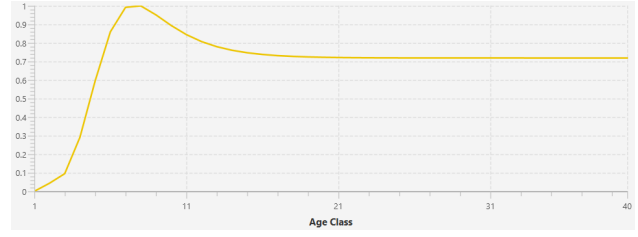
(b)

Figure 3: Age-specific selectivity for longline fishery 1 in a model with 10 annual ages `diagnostic-2024` (a) and updated age-specific selectivity in a model with 40 quarterly ages `mfcl-1979-20p3` .
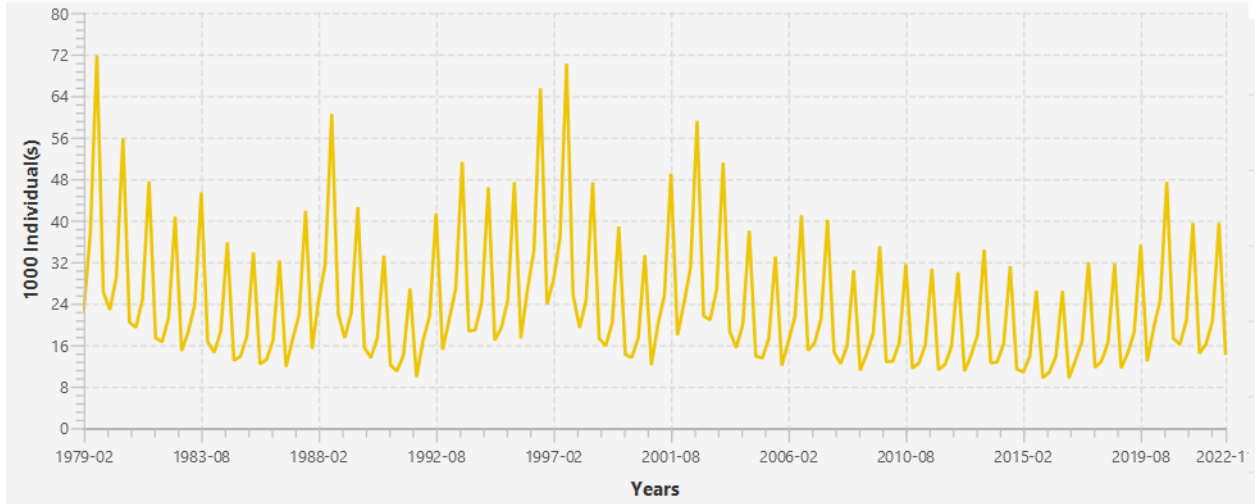
Figure 4: Recruitment estimated in the `mfcl-1979-20p3` model.
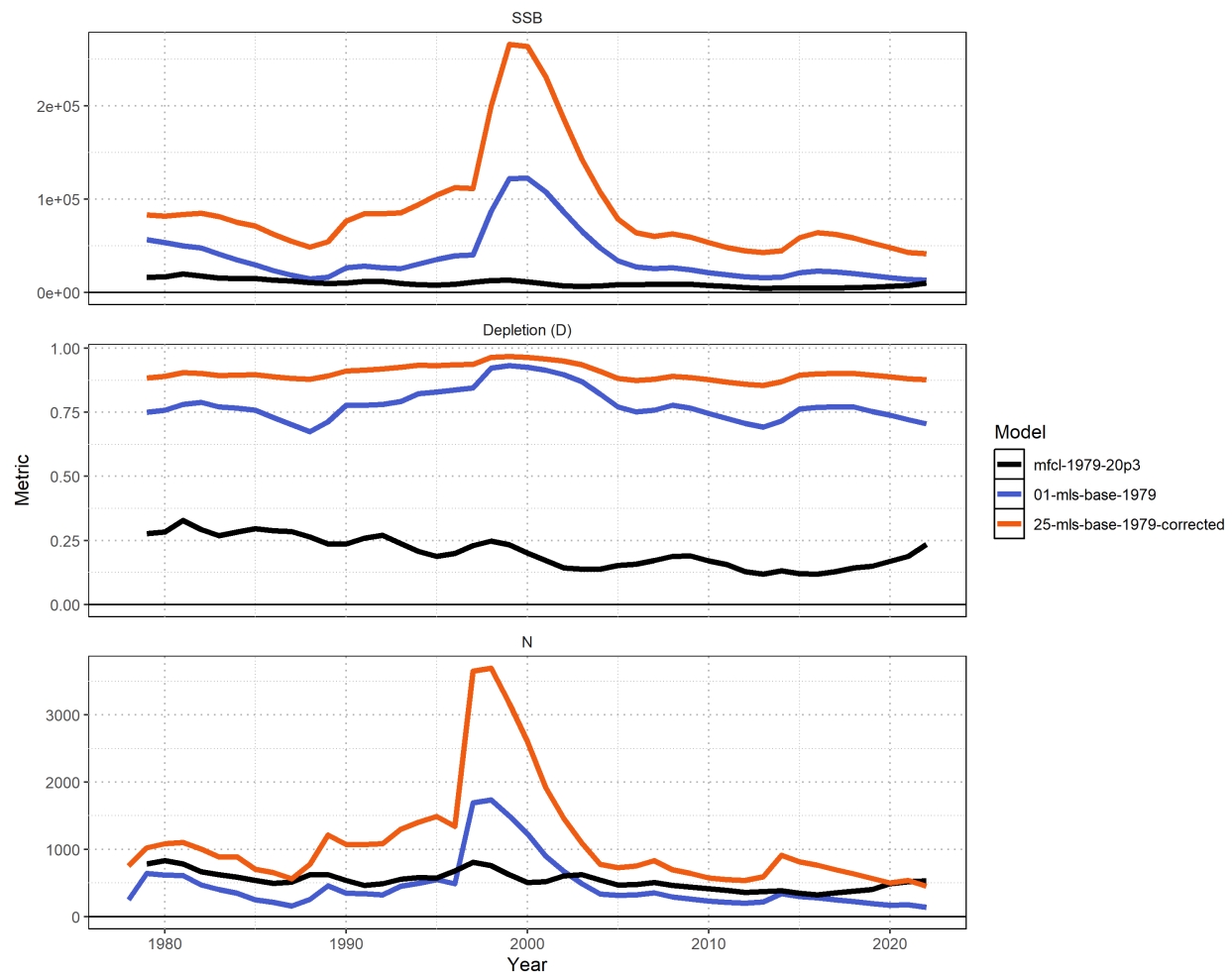
Figure 5: Time series comparison of Spawning Stock Biomass (SSB), Depletion (D), and to-
tal population numbers (N) across different models: `mfcl-1979-20p3`, `01-mls-base-1979`, and
`25-mls-base-1979-corrected`.
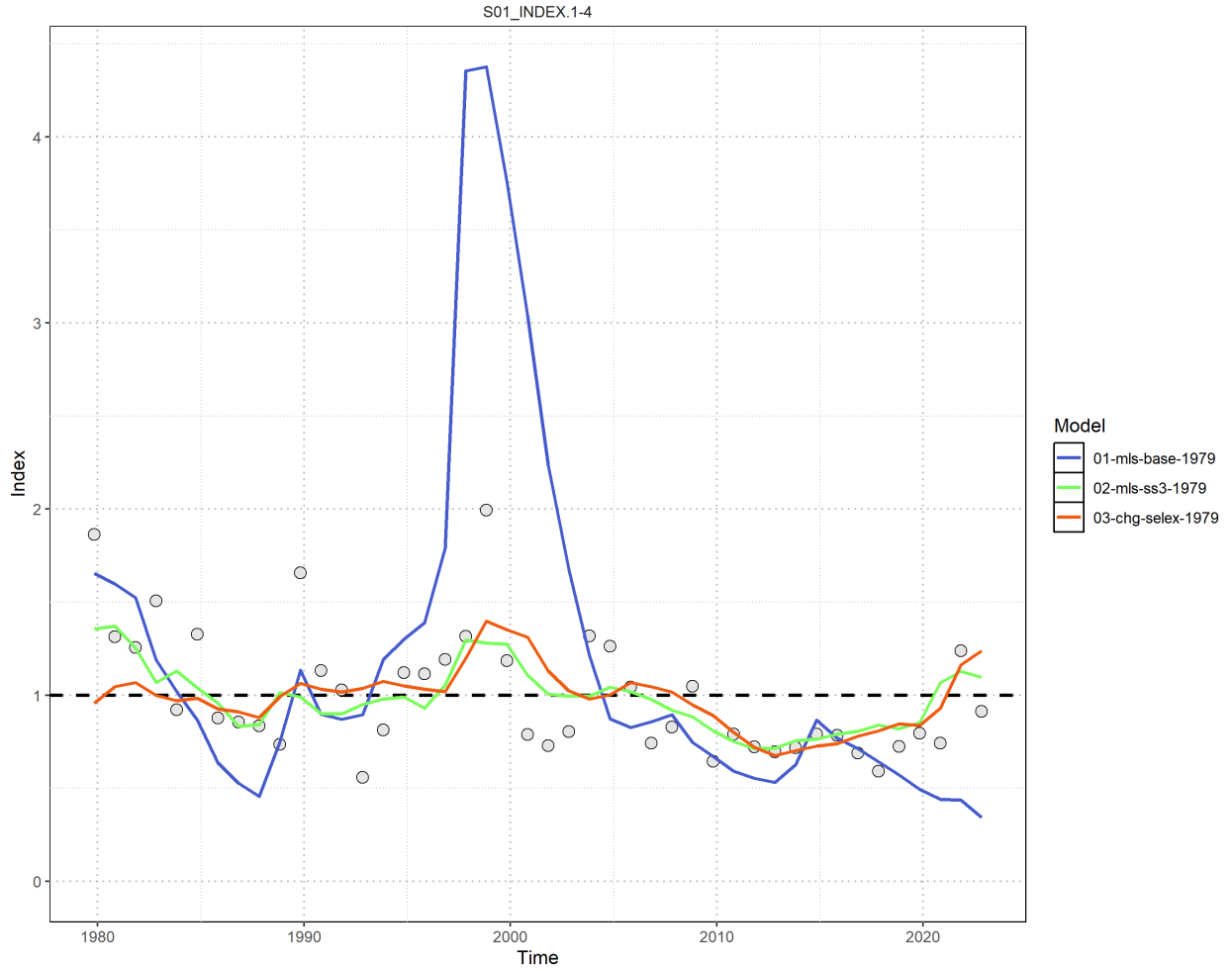
Figure 6: Fit to standardized CPUE index time series comparison showing observed values (circles) fitted by three different model configurations (lines): `01-mls-base-1979` (blue), `02-mls-ss3-1979` (green), and `03-chg-selex-1979` (orange).
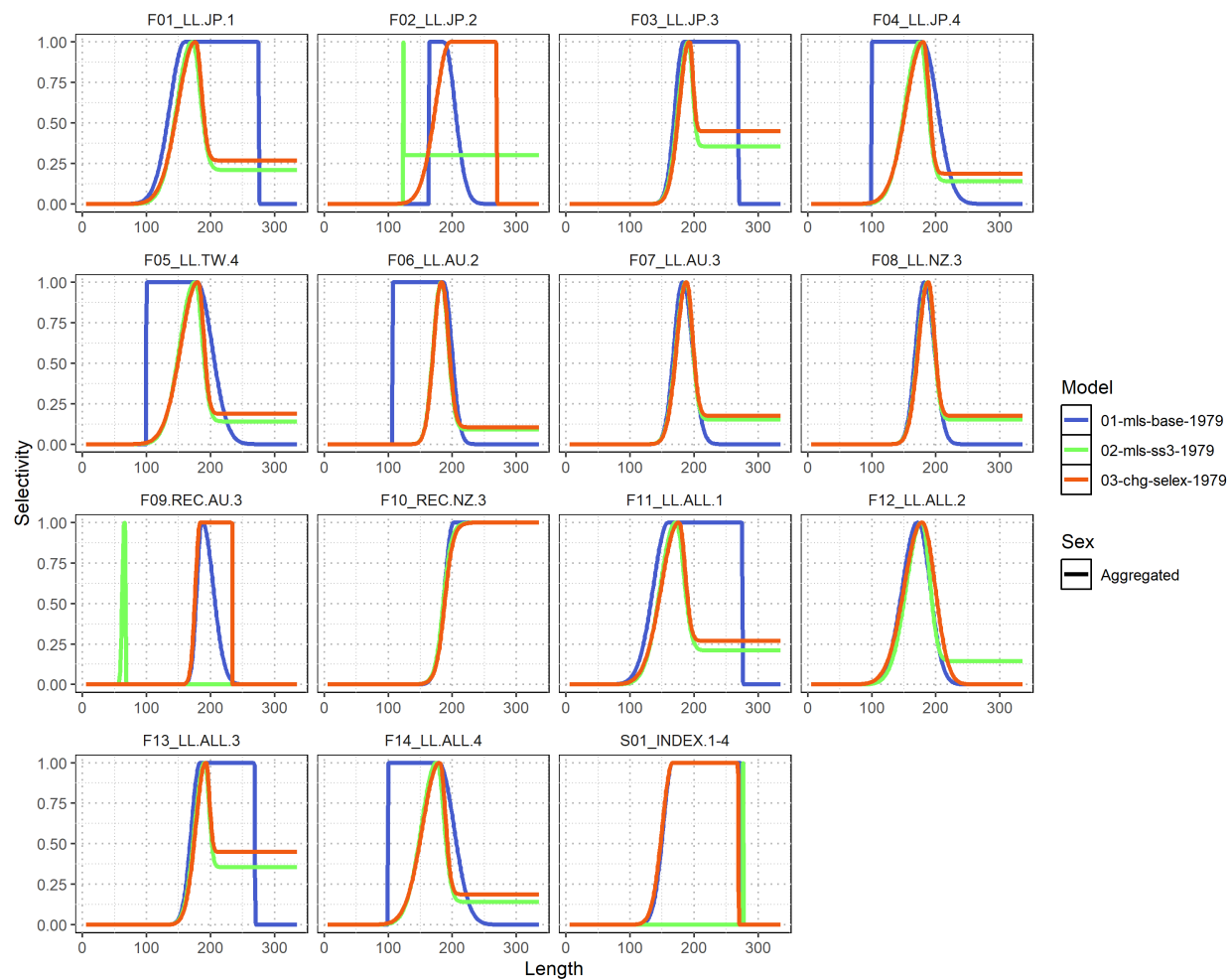
Figure 7: Selectivity curves by length across fisheries comparing three model configurations: `01-mls-base-1979` (blue), `02-mls-ss3-1979` (green), and `03-chg-selex-1979` (orange).
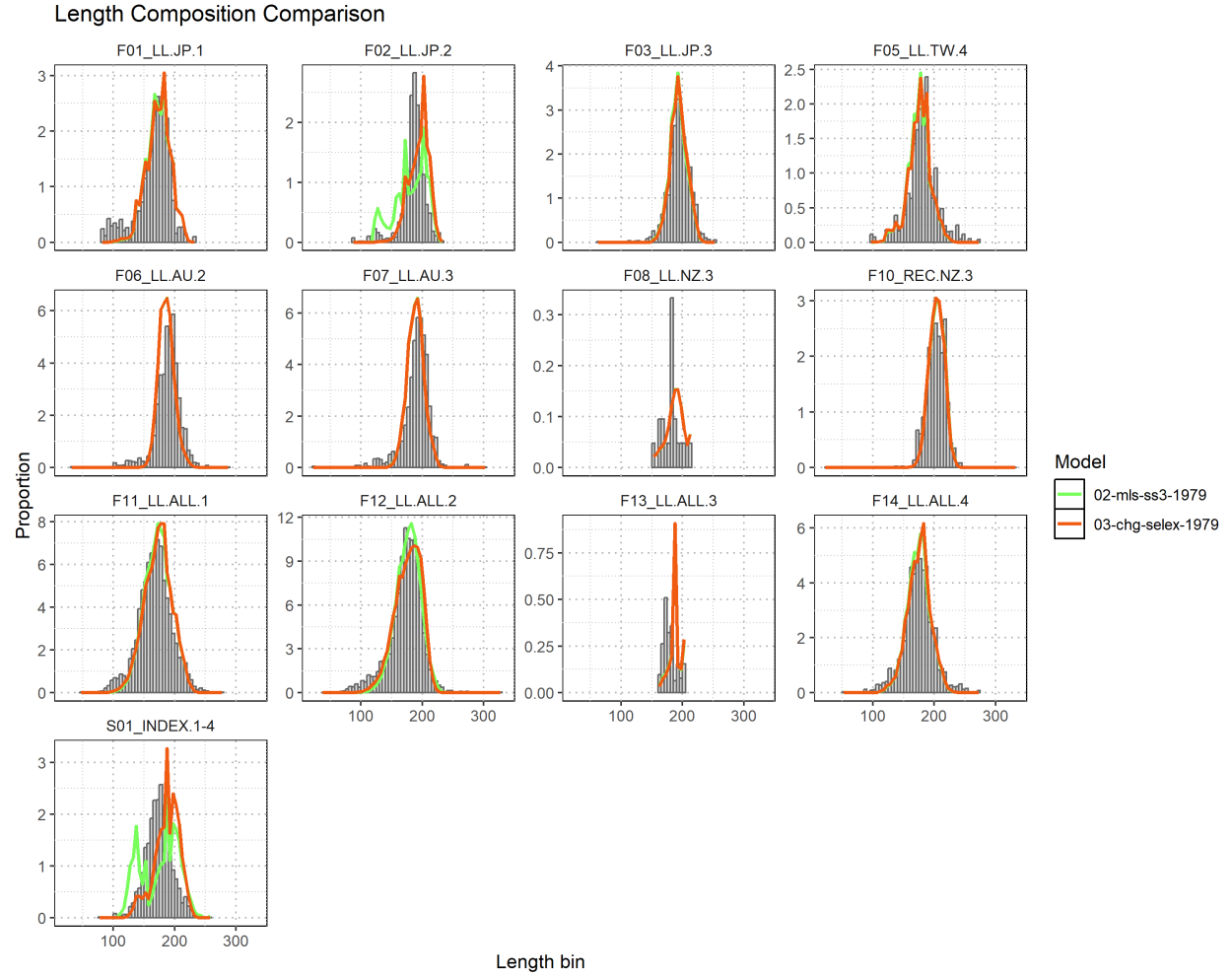
Figure 8: Fit to length composition showing observed and expected proportions by length bin. The histograms represent observed data while colored lines show model fits from two configurations: `02-mls-ss3-1979` (green) and `03-chg-selex-1979` (orange).
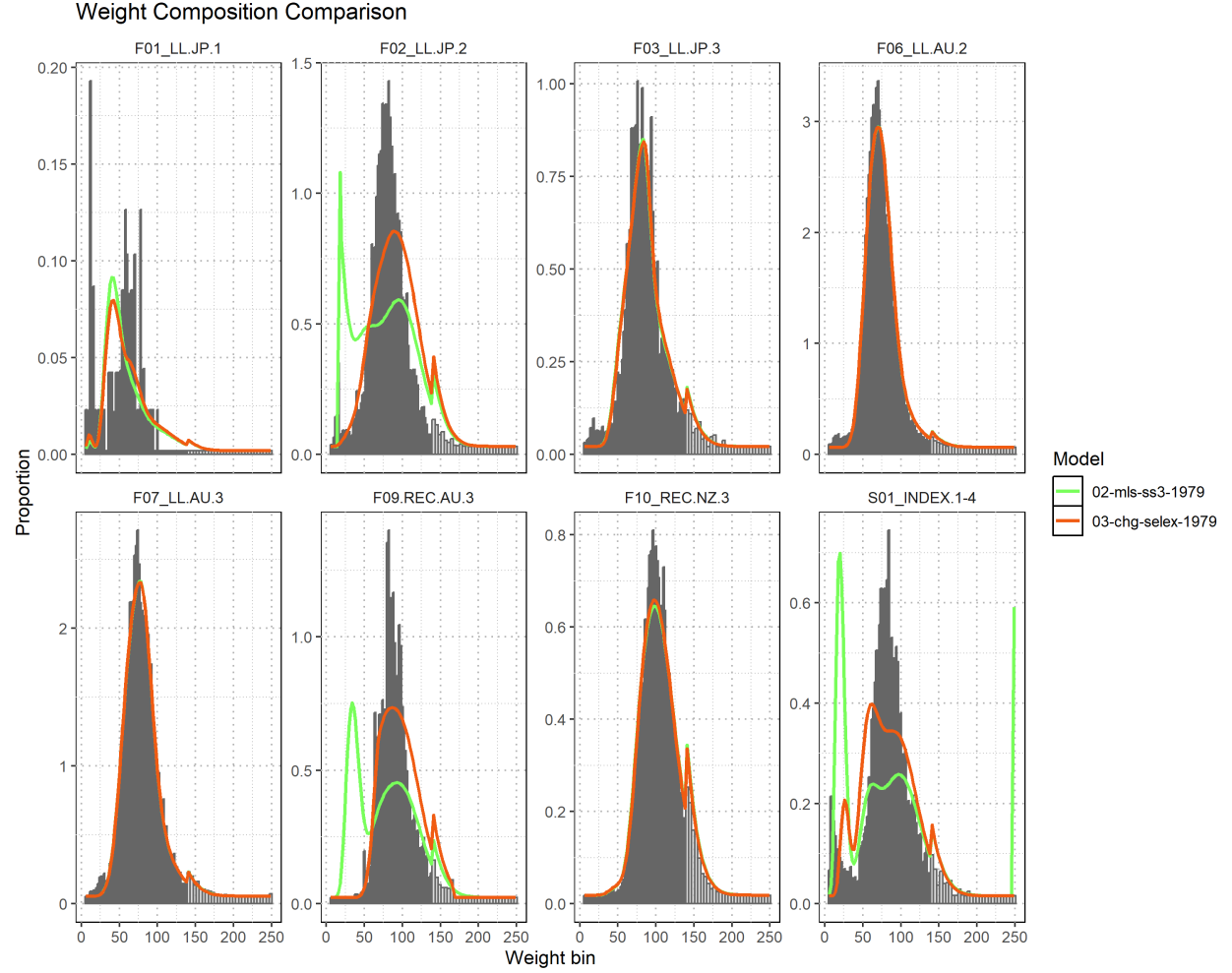
Figure 9: Fit to weight composition showing observed and expected proportions by weight bin. The histograms represent observed data while colored lines show model fits from two configurations: `02-mls-ss3-1979` (green) and `03-chg-selex-1979` (orange). Note that the discontinuity seen at ~ 140 is due to the change in bin width from 2 to 4 kg at that point.
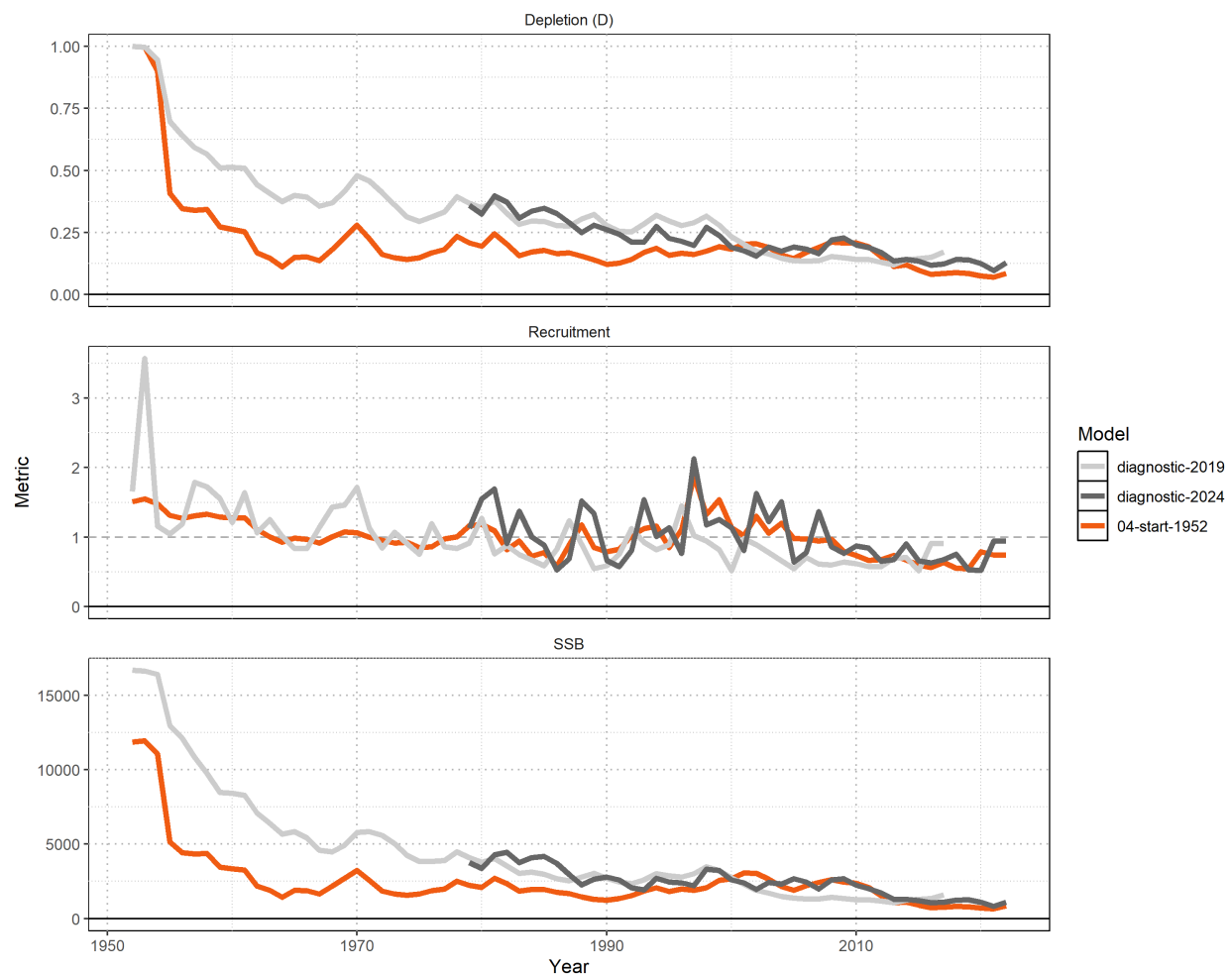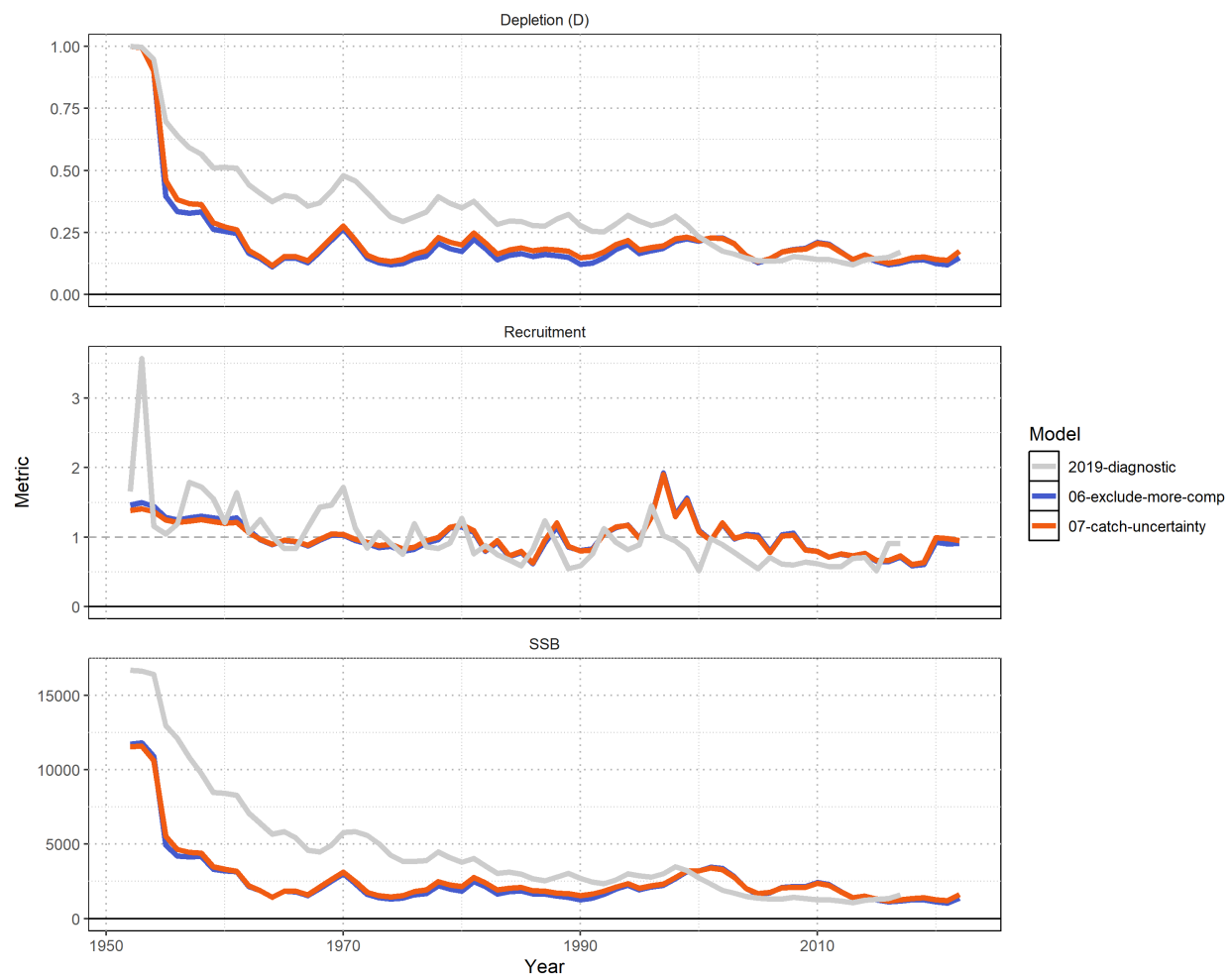
Figure 10: Time series comparison of key population metrics (Depletion (D), Recruitment, and Spawning Stock Biomass (SSB)) across different models: `diagnostic-2019` (light gray), `diagnostic-2024` (dark gray), and `04-start-1952` (orange).
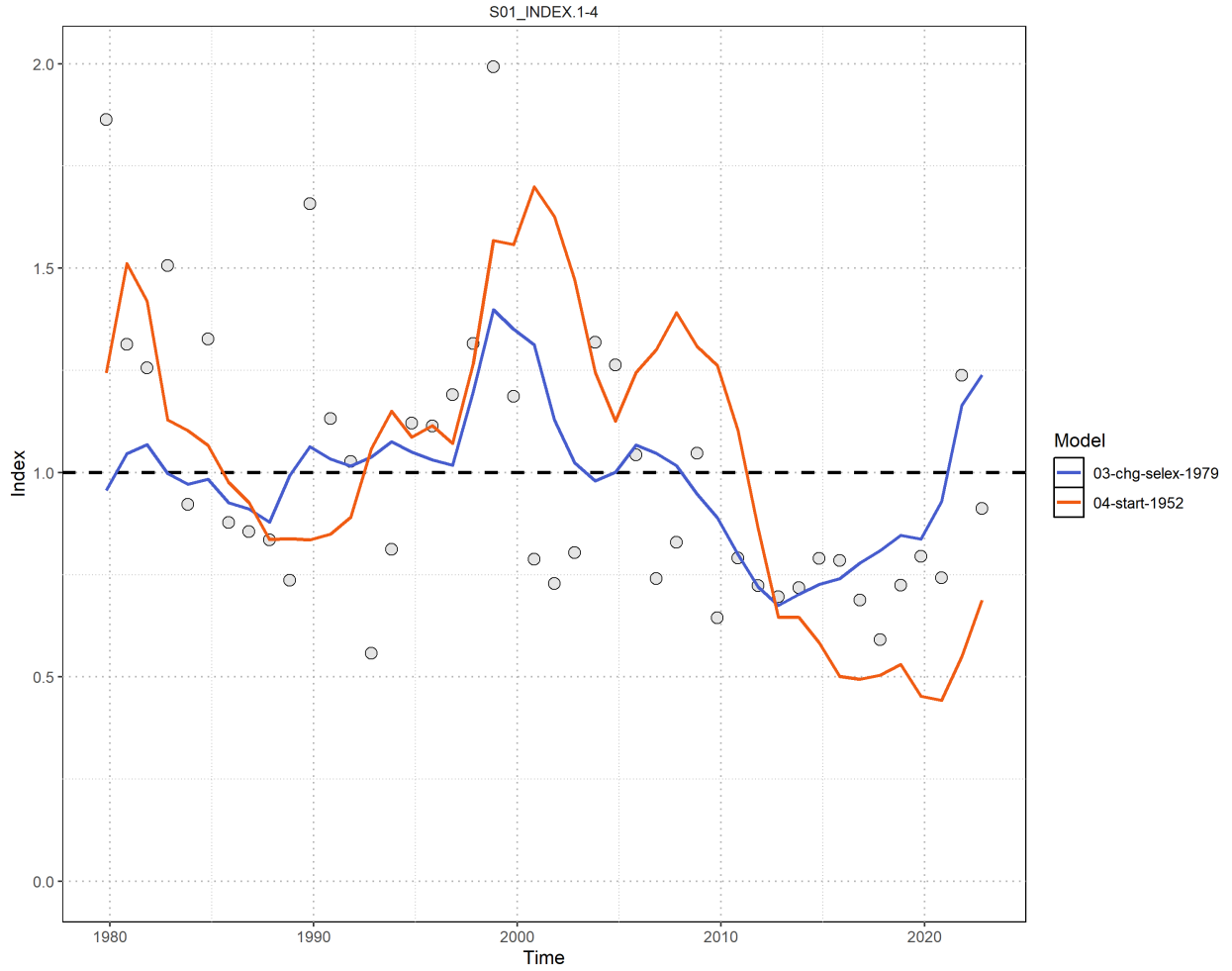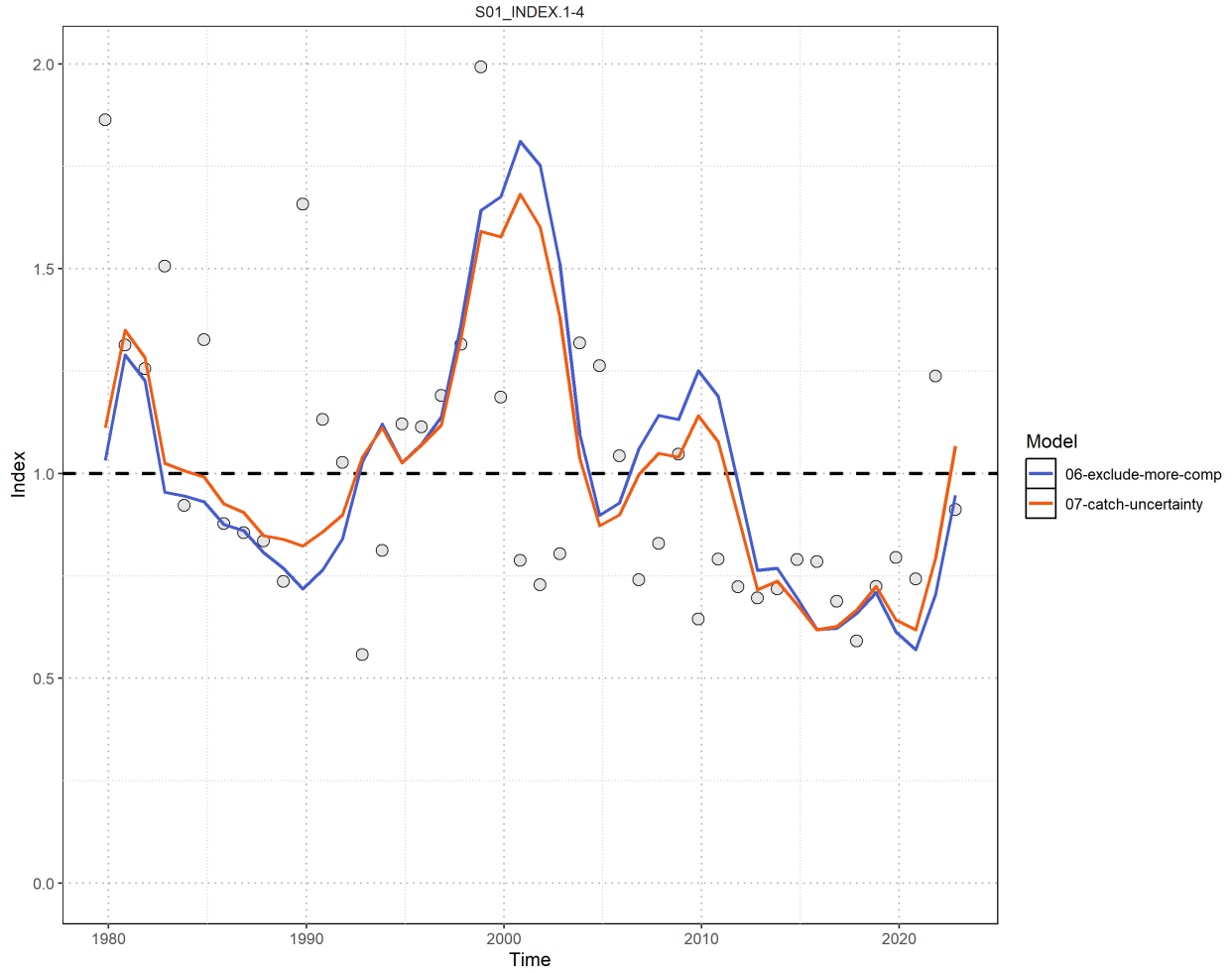
Figure 11: Fit to standardized CPUE index time series comparison showing observed values (circles) fitted by two different model configurations: `03-chg-selex-1979` (blue) and `04-start-1952` (orange).
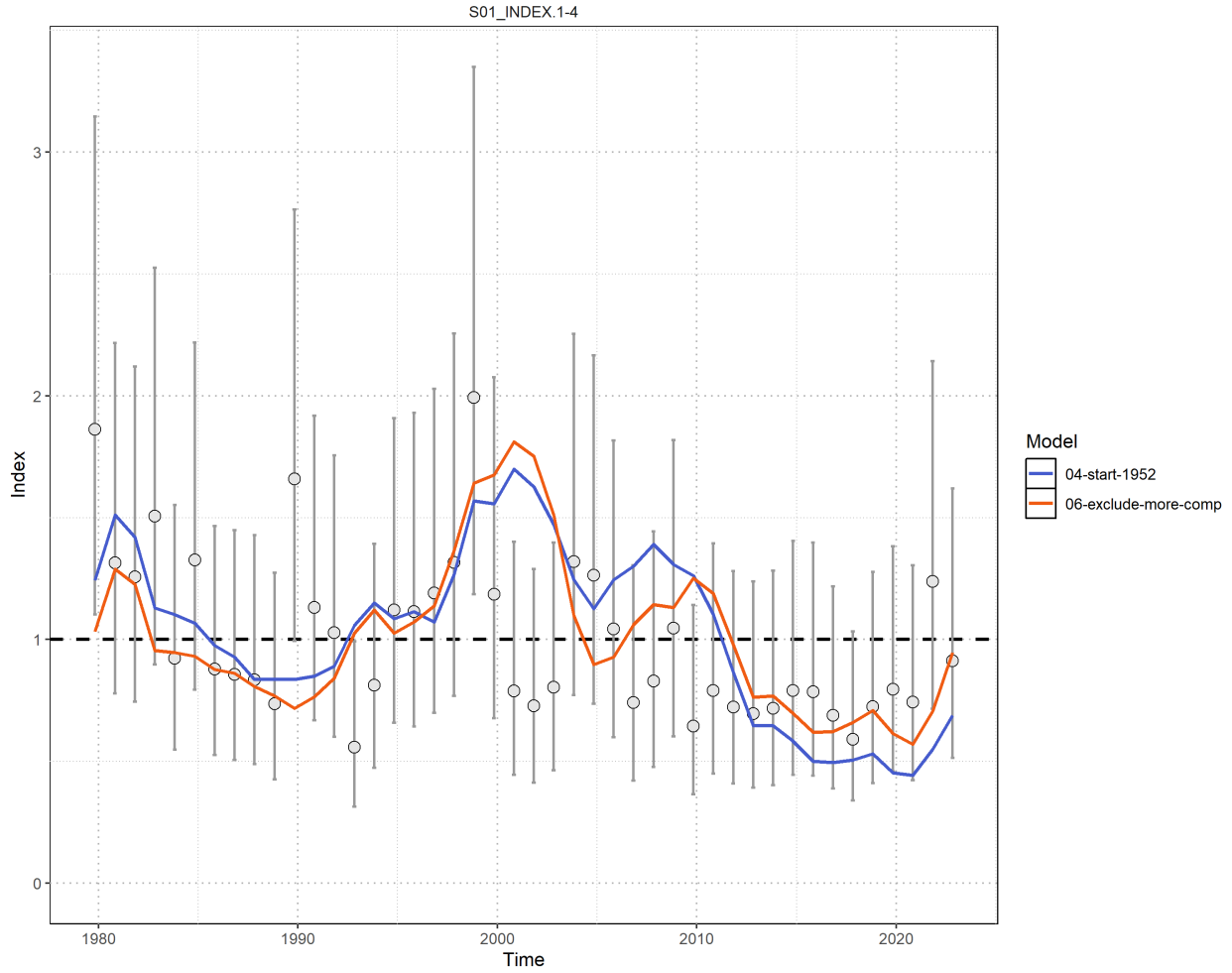
Figure 12: Fit to standardized CPUE index time series comparison showing observed values (circles) fitted by two different model configurations: `04-start-1952` (blue) and `06-exclude-more-comp` (orange).
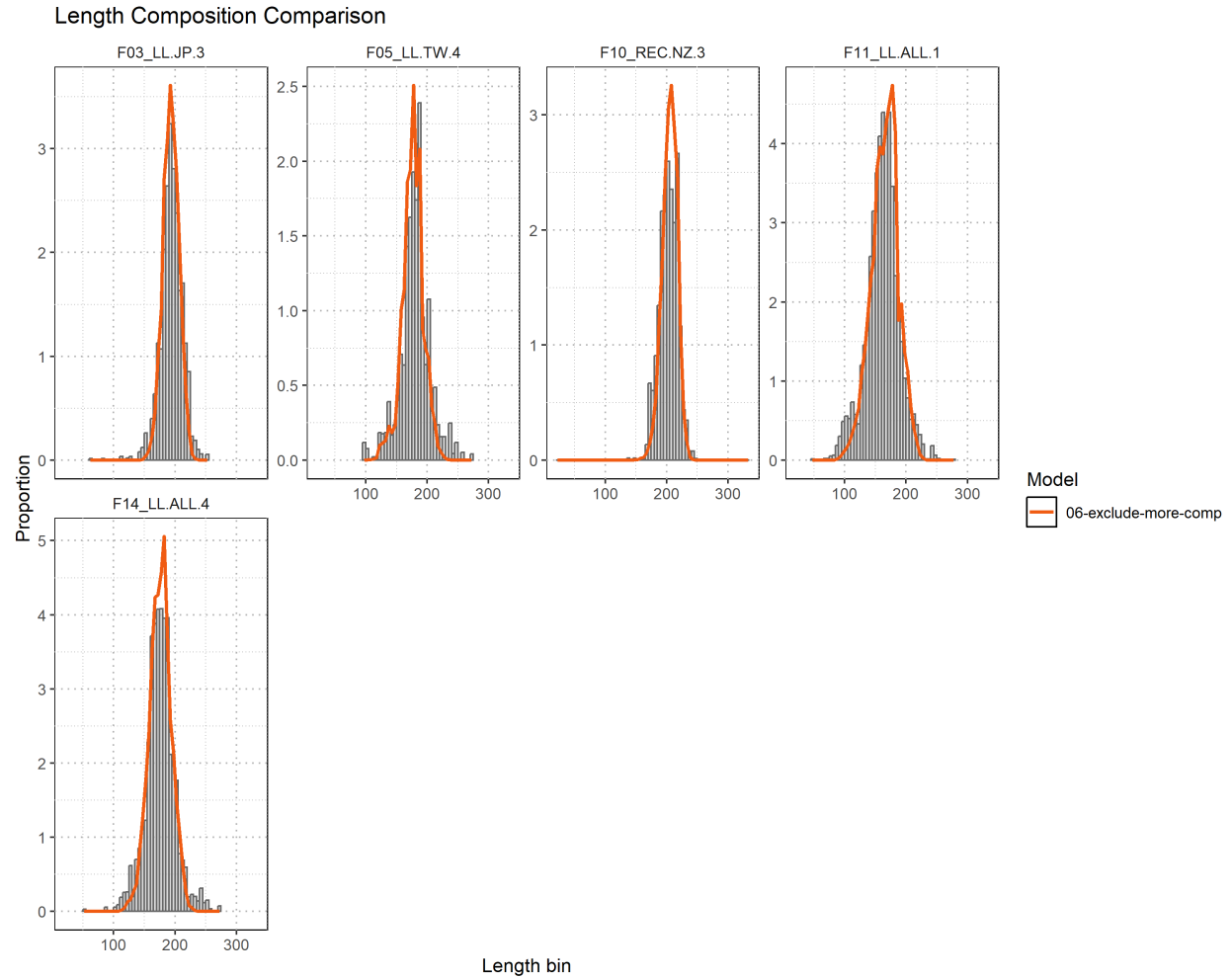
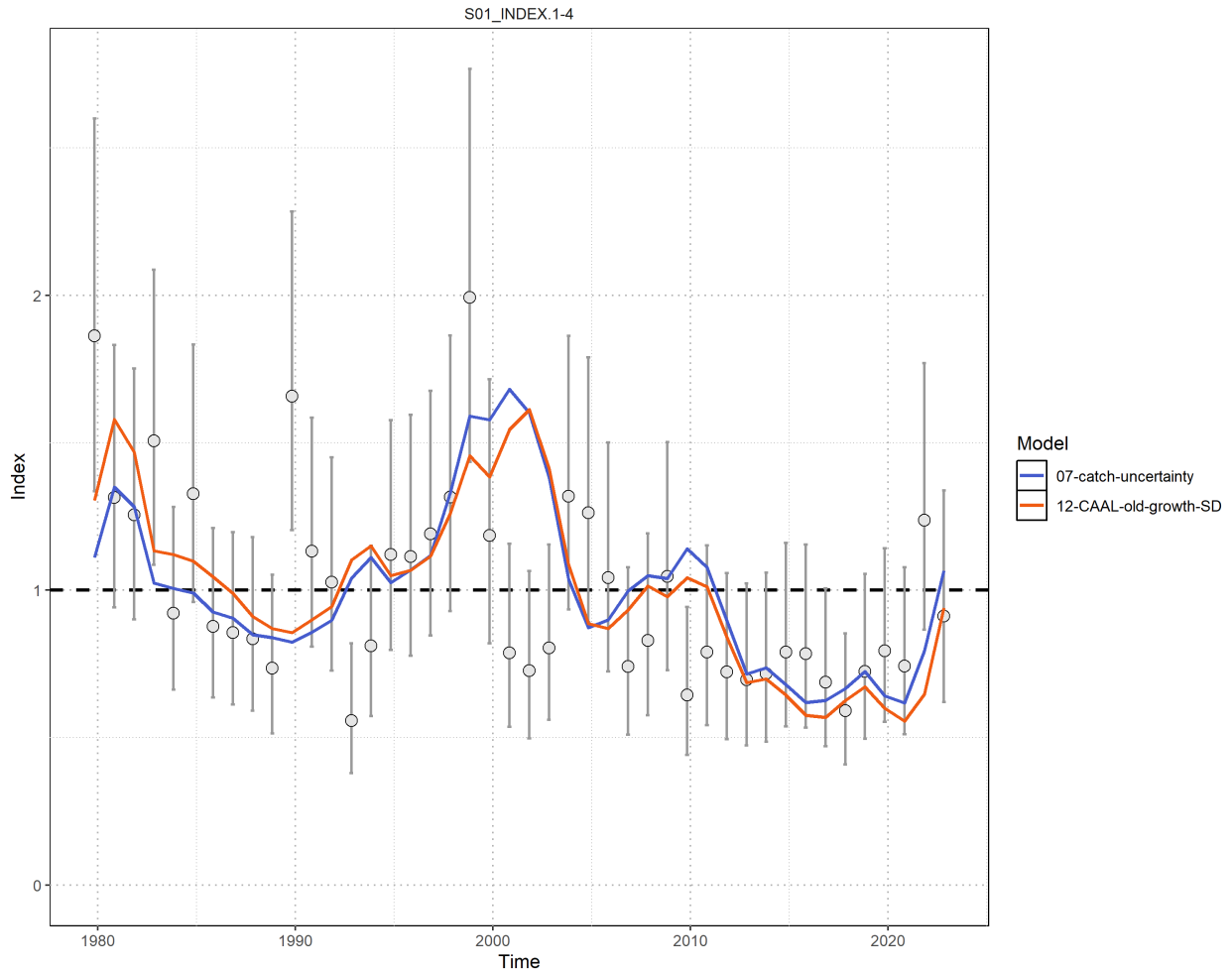Figure 13: Fit to length composition showing observed and expected proportions by length bin. The histograms represent observed data while colored lines show model fits from `06-exclude-more-comp` (orange).
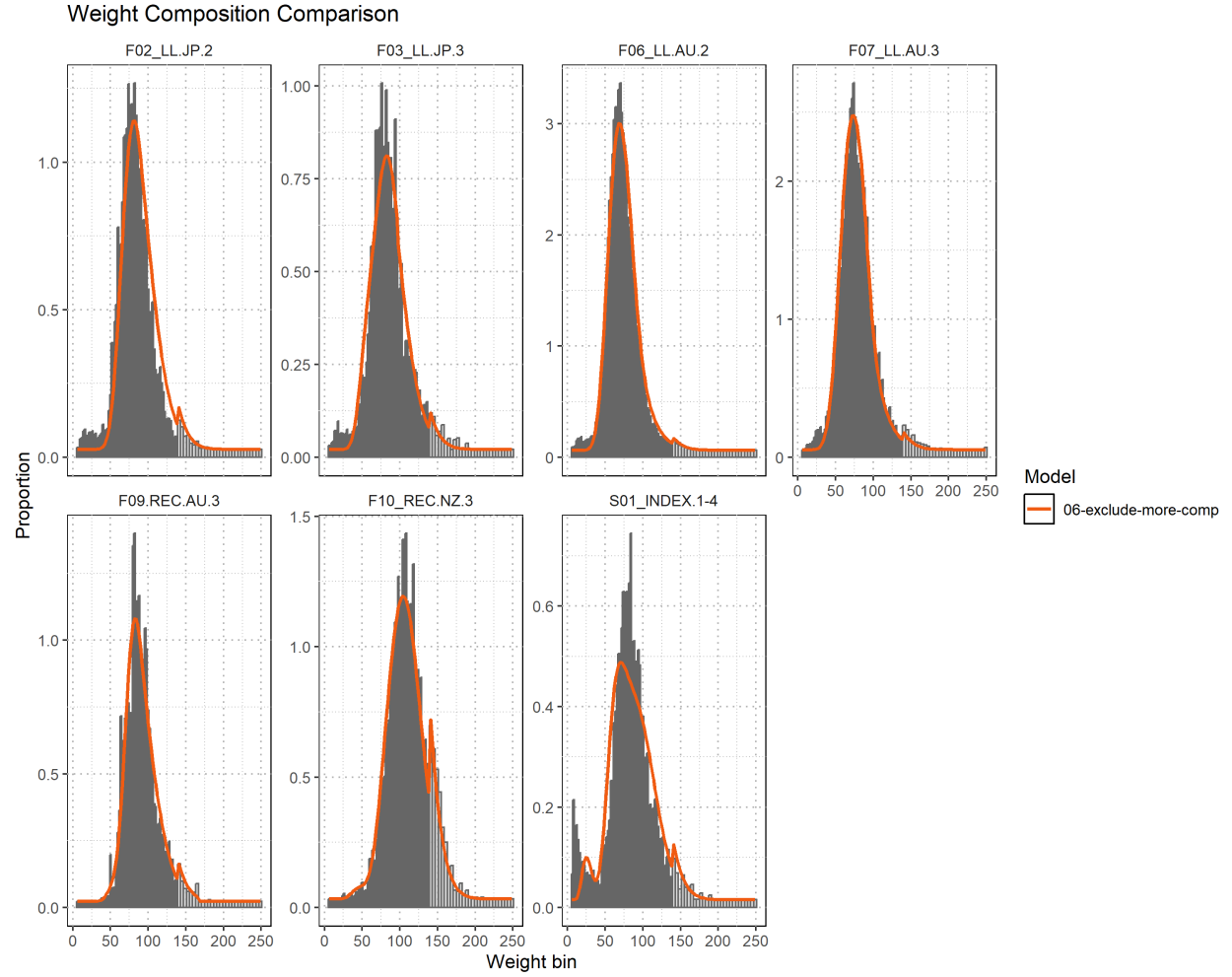
Figure 14: Fit to weight composition showing observed and expected proportions by weight bin. The histograms represent observed data while colored lines show model fits from `06-exclude-more-comp` (orange). Note that the discontinuity seen at ∼ 140 is due to the change in bin width from 2 to 4 kg at that point.
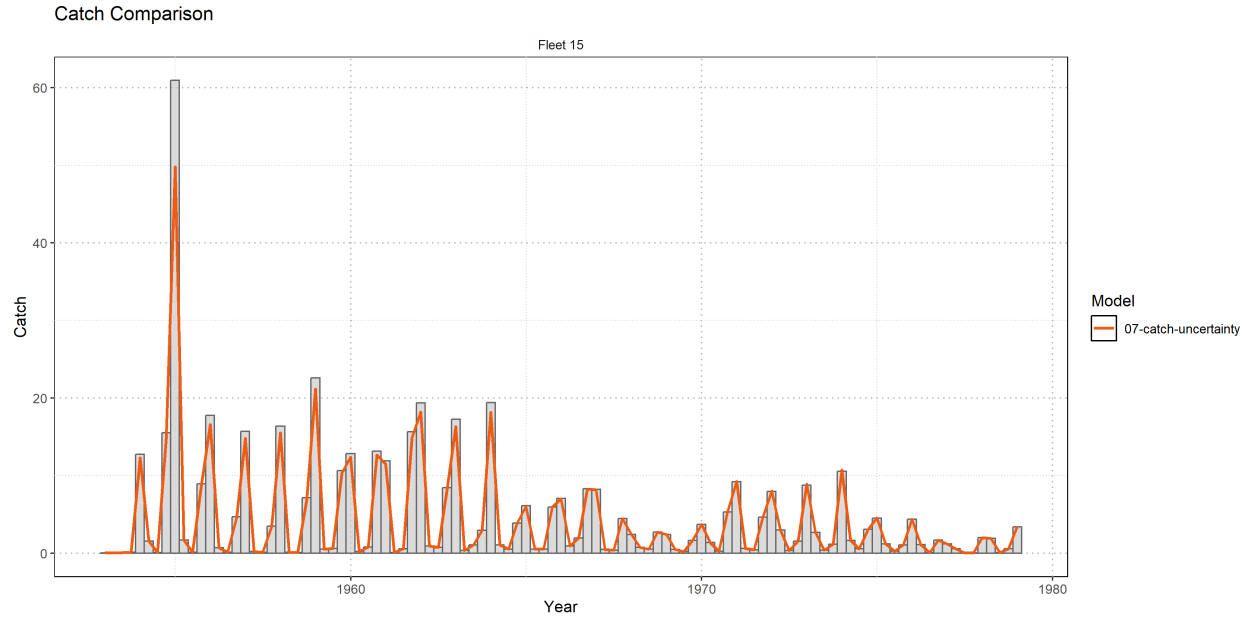
Catch Comparison

Figure 15: Fit to the catch for Fleet 15 (Early Japan LL; sub-region 2). The histograms represent observed data while colored lines show model fits from `07-catch-uncertainty` (orange).

Figure 16: Time series comparison of key population metrics (Depletion (D), Recruitment, and Spawning Stock Biomass (SSB)) across different models: `diagnostic-2019` (light gray), `06-exclude-more-comp` (blue), and `07-catch-uncertainty` (orange).

Figure 17: Fit to standardized CPUE index time series comparison showing observed values (circles) fitted by two different model configurations: `06-exclude-more-comp` (blue) and `07-catch-uncertainty` (orange).
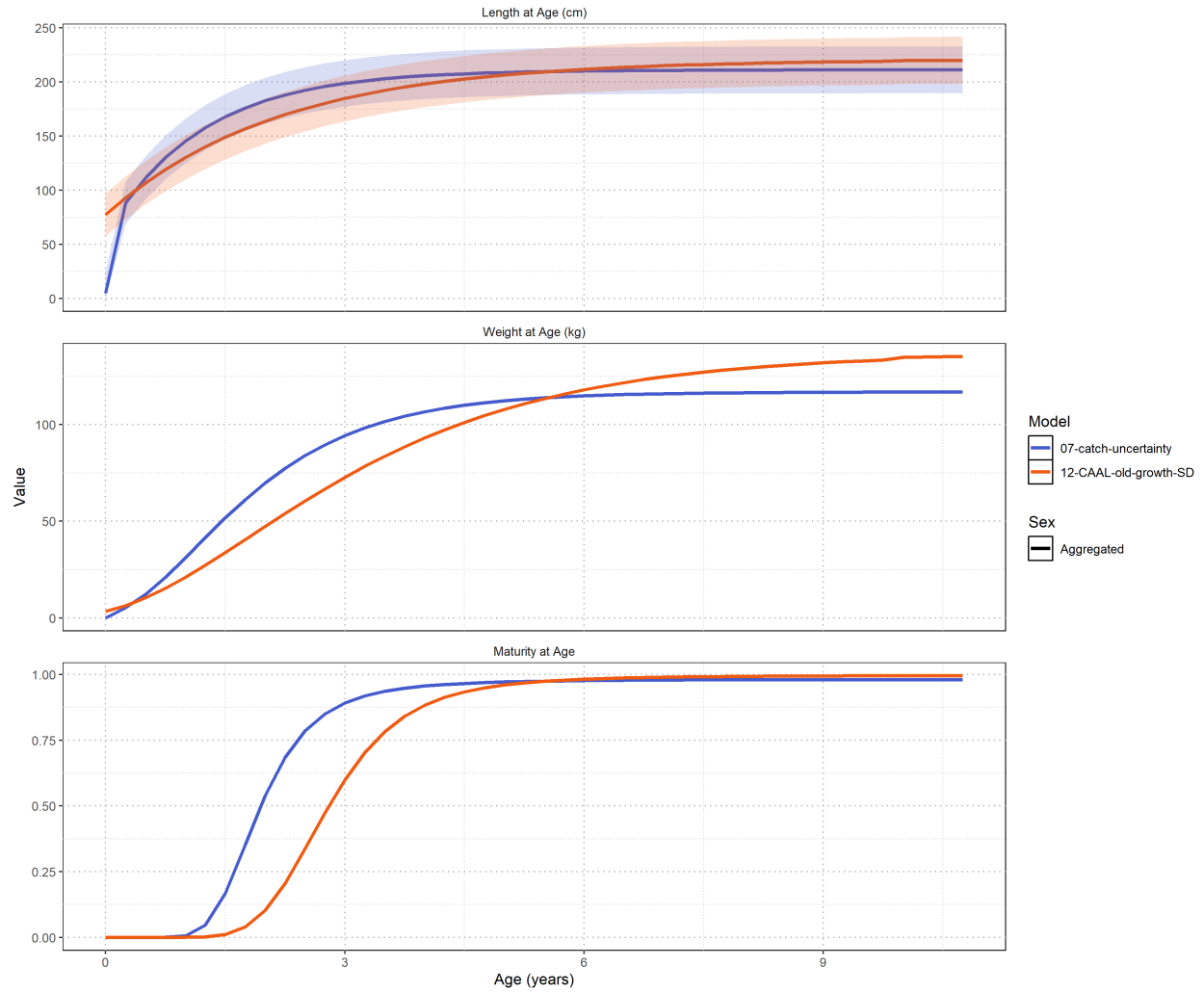
Figure 18: Length-at-age, weight-at-age, and maturity-at-age from two different models: `07-catch-uncertainty` (blue), and `12-CAAL-old-growth-SD` (orange). Variability in the length-at-age is shown in the top panel. Also weight-at-age and maturity-at-age are converted from length internal to the model using the probability of length-at-age.
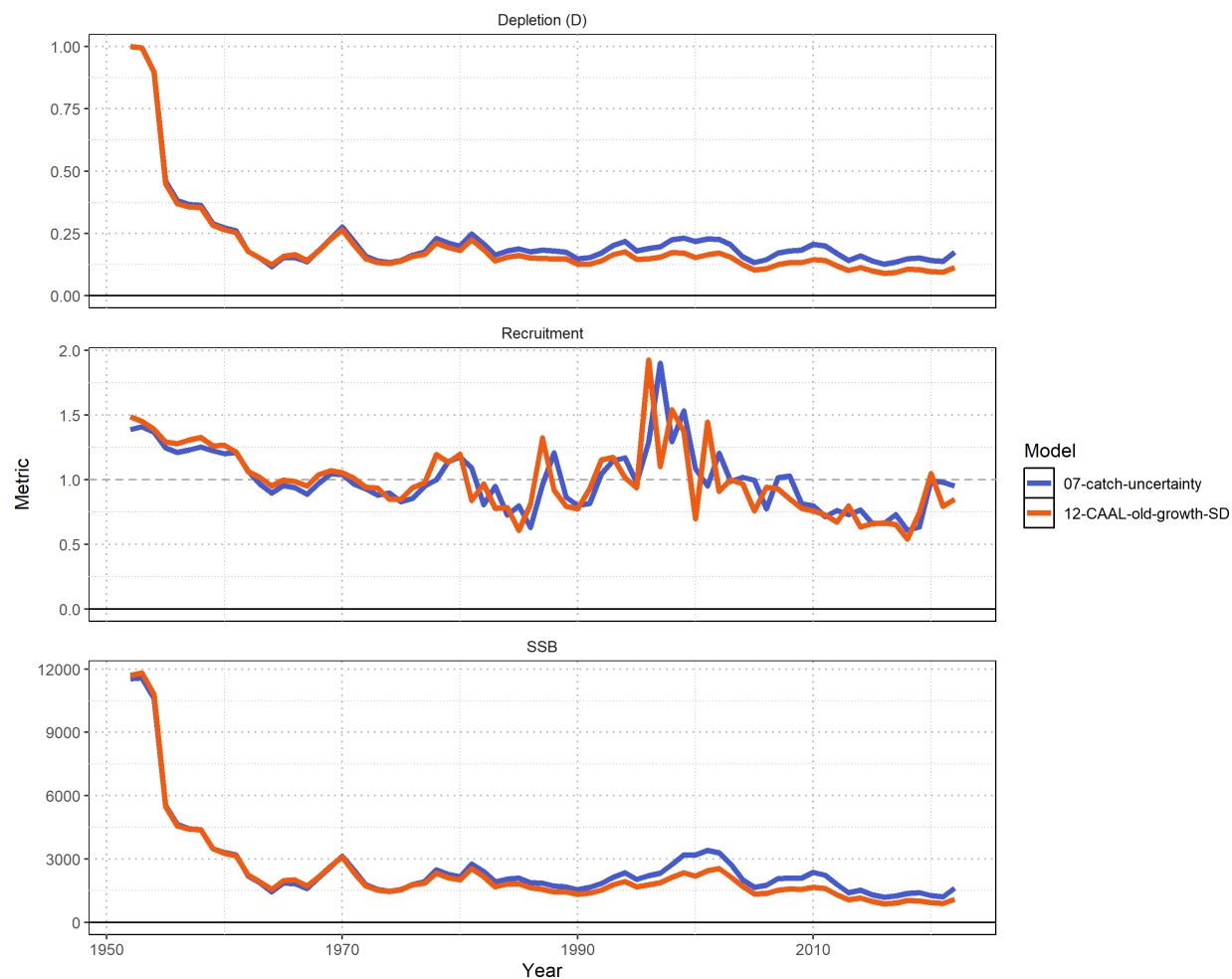
Figure 19: Time series comparison of key population metrics (Depletion (D), Recruitment, and Spawning Stock Biomass (SSB)) across different models: `07-catch-uncertainty` (blue), and `12-CAAL-old-growth-SD` (orange).
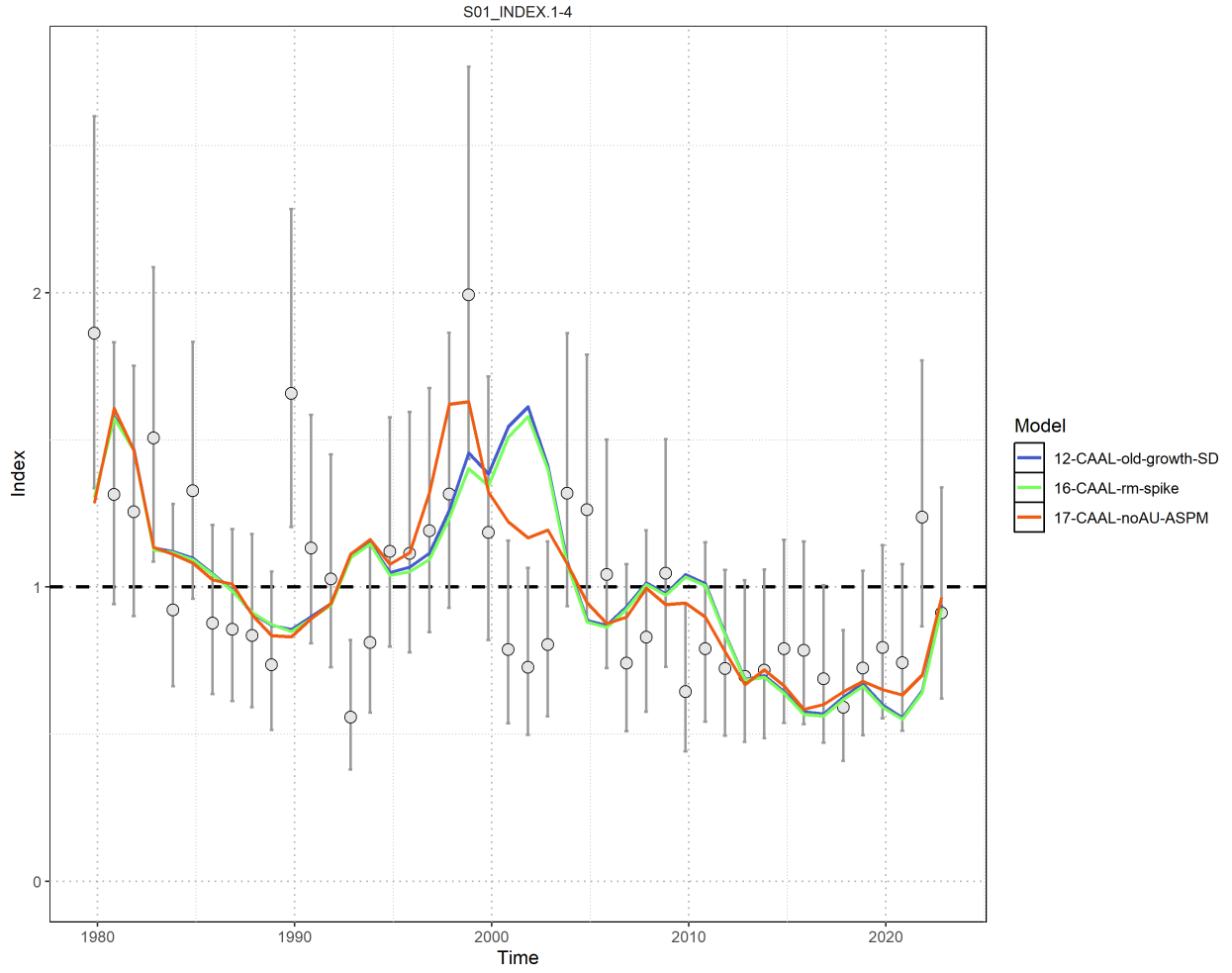
Figure 20: Fit to standardized CPUE index time series comparison showing observed values (circles) fitted by two different model configurations: `07-catch-uncertainty` (blue) and `12-CAAL-old-growth-SD` (orange).
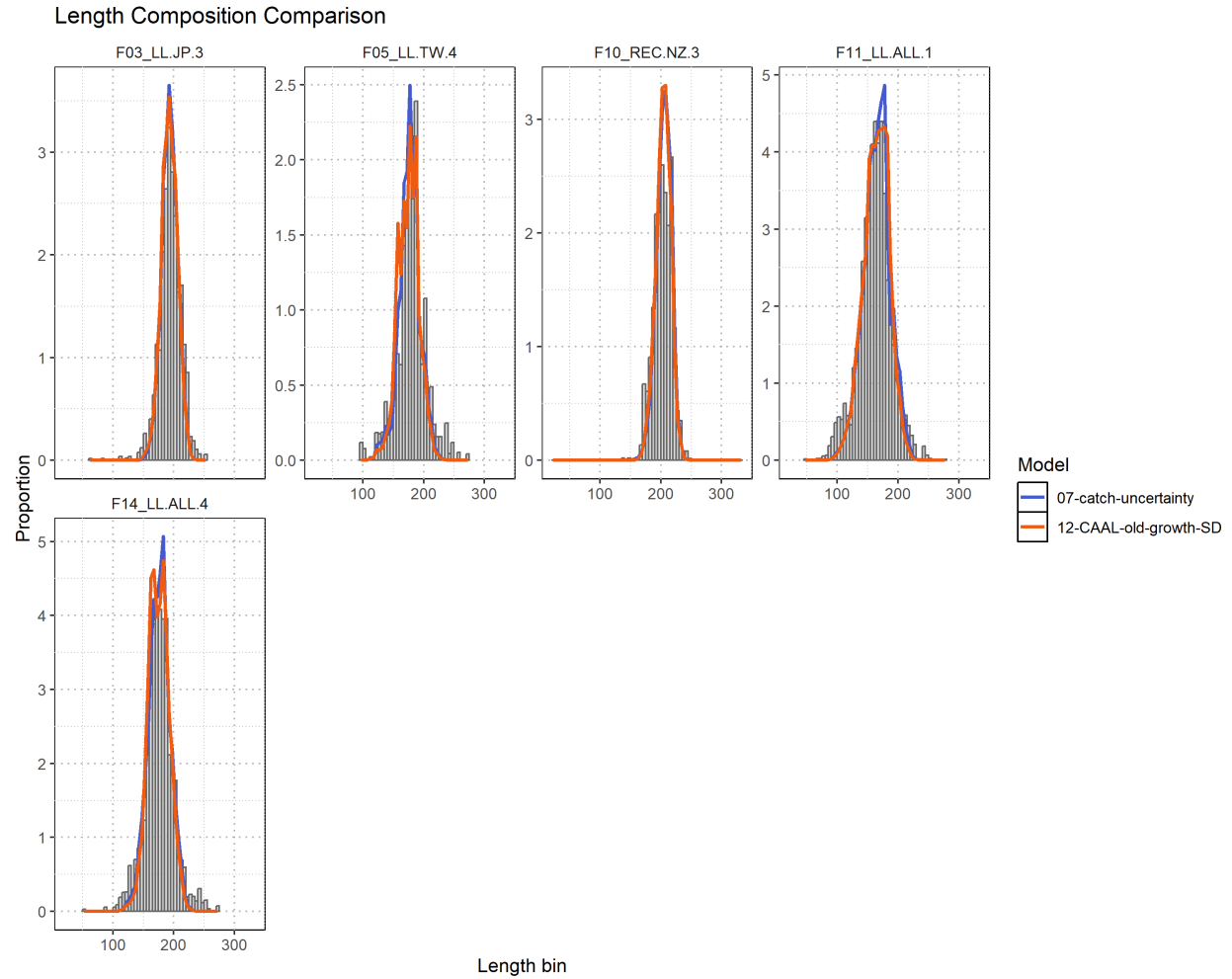
Figure 21: Fit to length composition showing observed and expected proportions by length bin. The histograms represent observed data while colored lines show model fits from two configurations: `07-catch-uncertainty` (blue) and `12-CAAL-old-growth-SD` (orange).
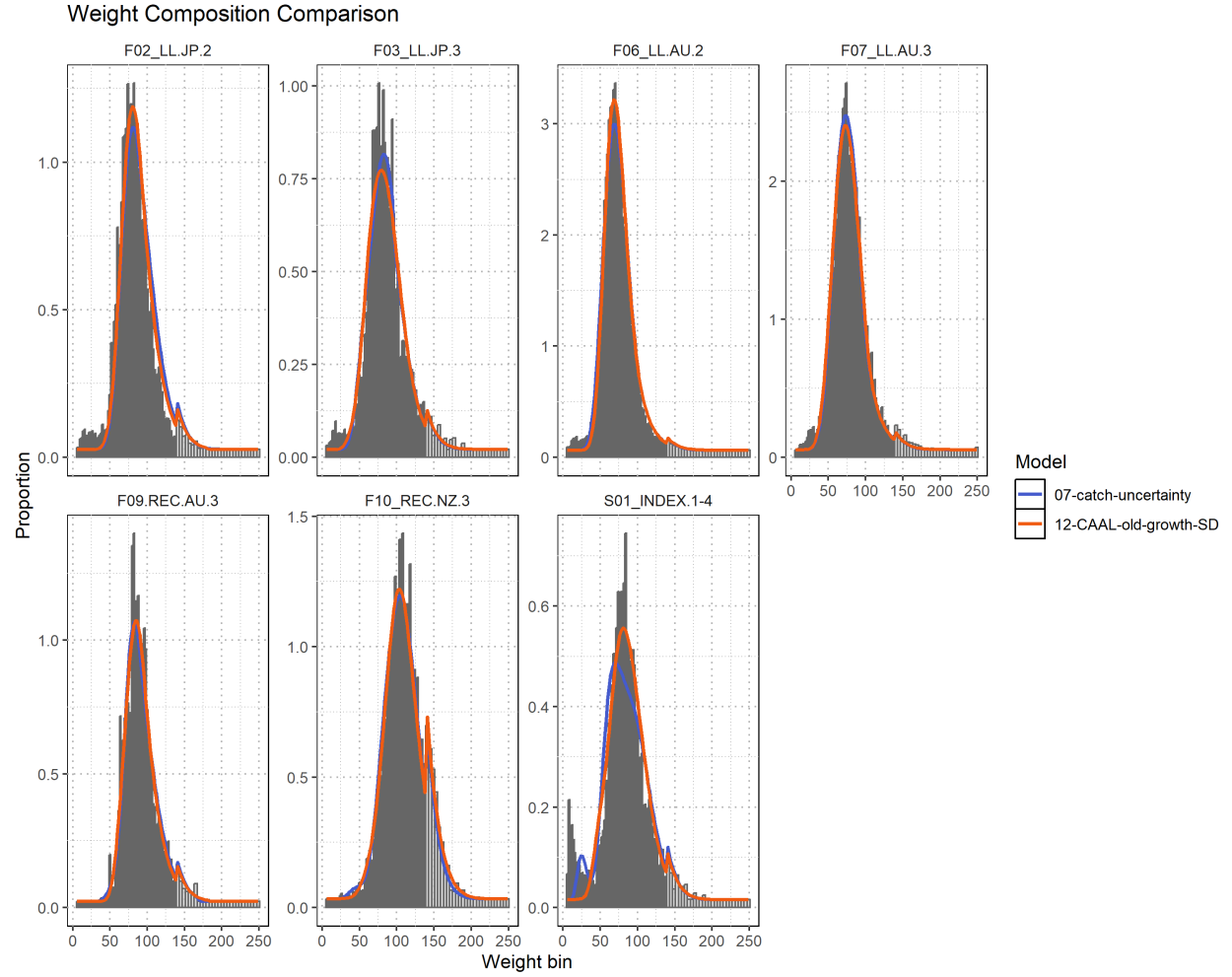
Figure 22: Fit to weight composition showing observed and expected proportions by weight bin. The histograms represent observed data while colored lines show model fits from two configurations: `07-catch-uncertainty` (blue) and `12-CAAL-old-growth-SD` (orange). Note that the discontinuity seen at ~ 140 is due to the change in bin width from 2 to 4 kg at that point.

Figure 23: Fit to standardized CPUE index time series comparison showing observed values (circles) fitted by three different model configurations: `12-CAAL-old-growth-SD` (blue), `16-CAAL-rm-spike` (green), and `17-CAAL-noAU-ASPM` (orange).
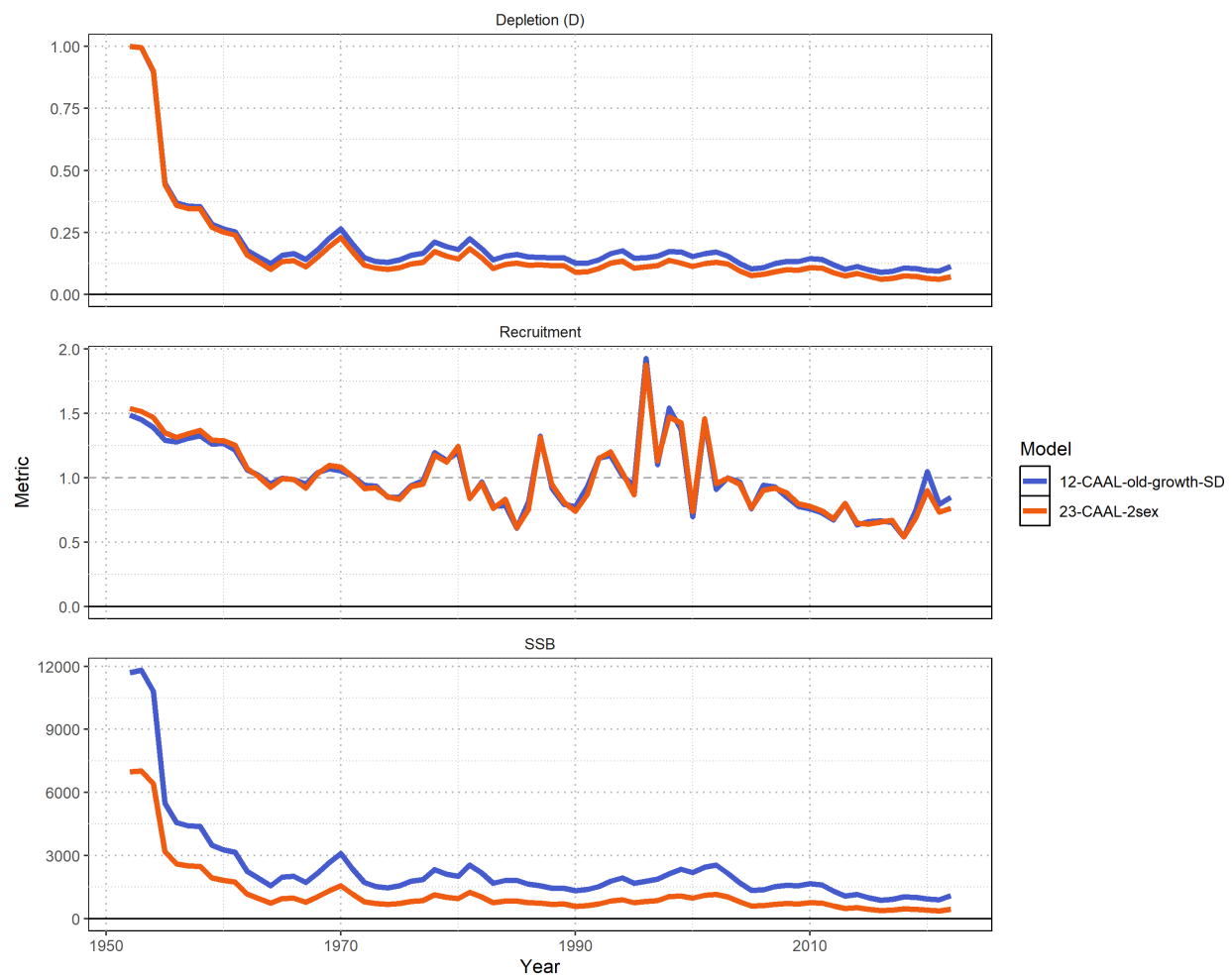
Figure 24: Time series comparison of key population metrics (Depletion (D), Recruitment, and Spawning Stock Biomass (SSB)) across different models: `12-CAAL-old-growth-SD` (blue), and `23-CAAL-2sex` (orange).
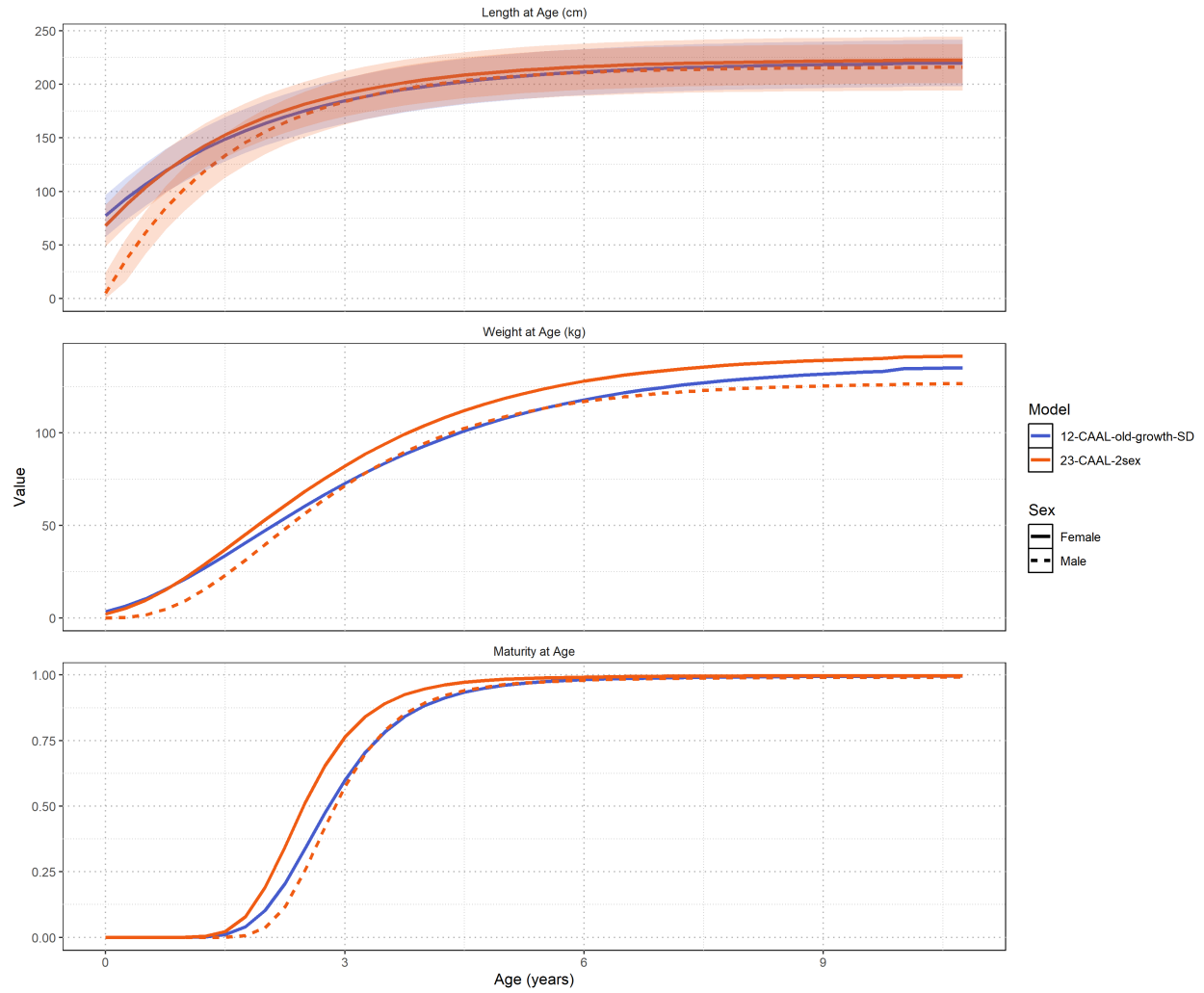
Figure 25: Length-at-age, weight-at-age, and maturity-at-age from two different models: `12-CAAL-old-growth-SD` (blue), and `23-CAAL-2sex` (orange). Variability in the length-at-age is shown in the top panel. Also weight-at-age and maturity-at-age are converted from length internal to the model using the probability of length-at-age. Sex is given by line type.
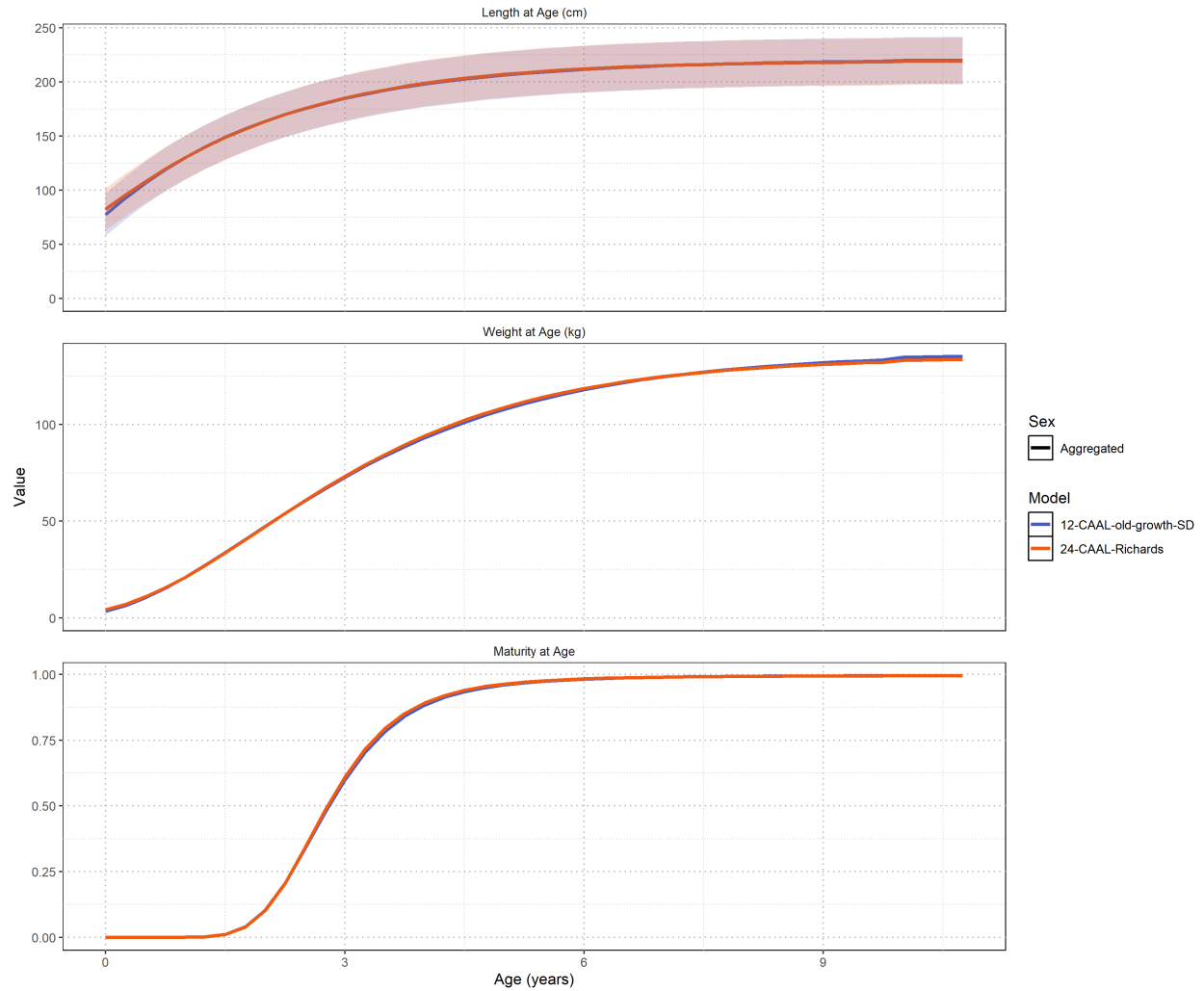
Figure 26: Length-at-age, weight-at-age, and maturity-at-age from two different models: `12-CAAL-old-growth-SD` (blue), and `24-CAAL-Richards` (orange). Variability in the length-at-age is shown in the top panel. Also weight-at-age and maturity-at-age are converted from length internal to the model using the probability of length-at-age.

Figure 27: Time series comparison of key population metrics (Depletion (D), Recruitment, and Spawning Stock Biomass (SSB)) across different models: `12-CAAL-old-growth-SD` (blue), and `21-CAAL-NZrecwtQtr` (orange).
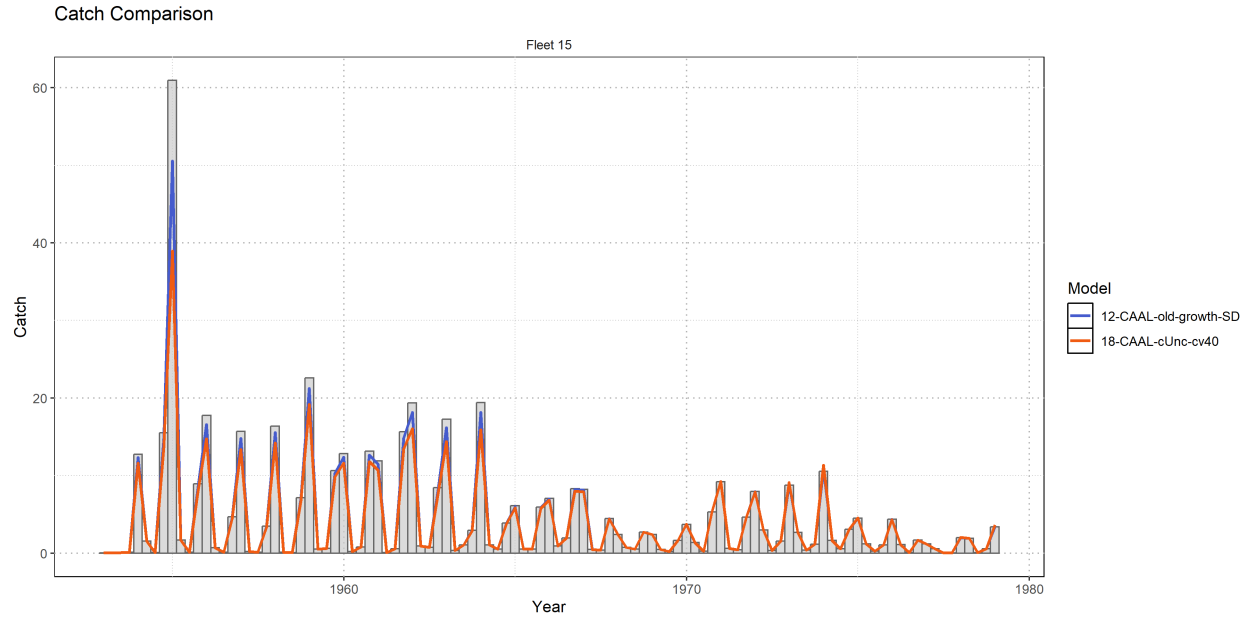
Catch Comparison

Figure 28: Fit to the catch for Fleet 15 (Early Japan LL; sub-region 2). The histograms represent observed data while colored lines show model fits from `12-CAAL-old-growth-SD` (blue), and `18-CAAL-cUnc-cv40` (orange).
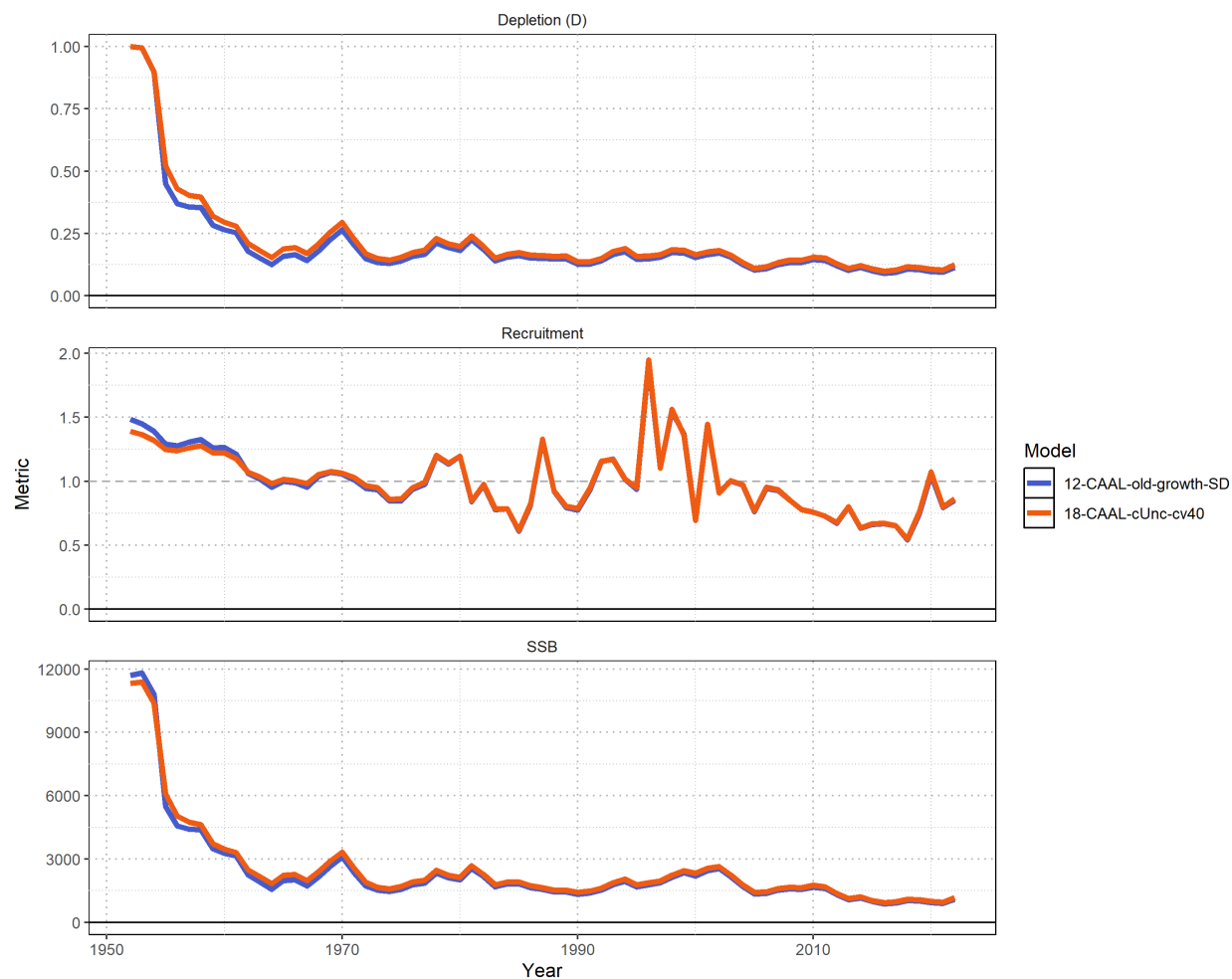
Figure 29: Time series comparison of key population metrics (Depletion (D), Recruitment, and Spawning Stock Biomass (SSB)) across different models: `12-CAAL-old-growth-SD` (blue), and `18-CAAL-cUnc-cv40` (orange).
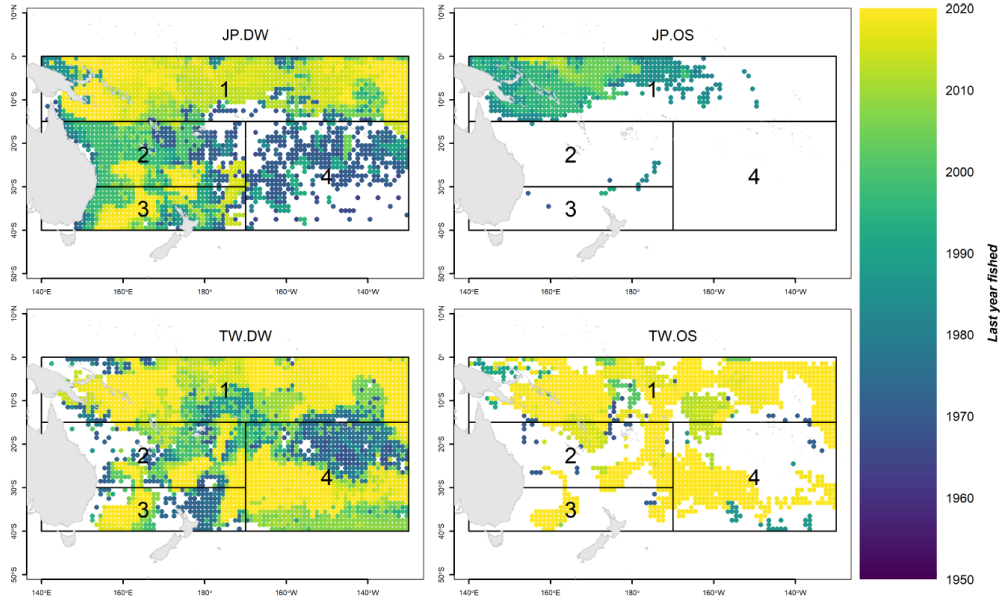
Figure 30: Reprinted from WCPFC-SC-15/SA-IP-07 Figure 16. The spatiotemporal distributions of longline fishing effort for the distant-water (DW) and offshore (OS) components of the Japanese (JP) and Chinese Taipei (TW) fleets. Lighter colors indicate the $1° \times 1°$ grid cells have been fished more recently while darker colors indicate that the grid cell was last fished earlier in the model period.

Appendix

Table 3: Meeting participants list.

Name	Affiliation
Claudio Castillo-Jordán	SPC
Rob Ahrens	PIFSC
Felipe Carvalho	PIFSC
Nicholas Ducharme-Barth	PIFSC
Michelle Sculley	PIFSC