Alaska Sablefish Model Update

Daniel Goethel, Dana Hanselman, Chris Lunsford, Cara Rodgveller,
Ben Williams, Katy Echave, Jane Sullivan, and Pete Hulson

Alaska Fisheries Science Center, Auke Bay Laboratories

September 2021

Executive Summary

Rapid changes in the sablefish resource as well as the associated fisheries have proven difficult to address in the sablefish stock assessment model currently used for management advice. Over the last few years the model has demonstrated increasing retrospective patterns and extensive retroactive downgrading of recent year classes. Updated data and analysis of growth, weight, and maturity indicate that values for these biological parameters used in the assessment should be revised. Similarly, degrading fits to important data sources (e.g., longline survey abundance indices) suggest that model parametrization and data weighting merit refinement. After exploring a variety of model updates and new parametrizations, the results of the current work suggest that five important changes would improve the 2021 sablefish assessment: 1) weight and growth for the recent time period (1996 – present) should be updated to reflect the full extent of available data; 2) maturity should be updated with an age- and length-based general additive model (GAM) that accounts for skipped spawning using histological information and incorporates data from recent maturity studies; 3) the catchability priors are no longer needed; 4) a recent time block for estimation of fishery and survey selectivity and fishery CPUE catchability will allow the model to better fit recent data and reduce internal model tension due to slightly incongruous trends in indices and compositional data; and 5) using data reweighting approaches (e.g., the Francis method) can help improve fits to abundance indices, limit retrospective patterns, and reduce retroactive downgrades in recruitment estimates and associated ABCs. Additionally, a brief update on the availability of data inputs for the 2021 SAFE is provided in Appendix A.1. Of particular importance, due to financial constraints, it appears unlikely that the fishery CPUE index will be updated in 2021.

Model *21.10_Proposed* is suggested as the best model for the provision of sablefish management advice, given that it incorporates each of the five proposed model changes. Model *21.10_Proposed* provides better fits to the longline survey RPN and fishery CPUE indices, albeit at the cost of degraded fits to the fishery age composition data. Compared to the current SAFE model (*16.5_Cont*), the resulting population trajectory from *21.10_Proposed* demonstrates less drastic reductions in SSB during the mid-2010s with more subtle rebuilding since 2017, primarily due to greatly reduced estimates of recent year class strength. Projected ABCs from the *21.10_Proposed* model are significantly lower than the *16.5_Cont* model and appear to be less volatile. Additionally, retrospective patterns and associated retroactive downgrades in recruitment year class strength have been reduced. Although the proposed model (*21.10_Proposed*) is not without flaws and requires further refinements to better reflect the dynamics of the sablefish resource and fishery, we believe it provides important tangible improvements over the current SAFE model (*16.5_Cont*).

Introduction

The Alaskan sablefish (*Anoplopoma fimbria*) resource has undergone rapid changes in population dynamics over the last decade as multiple, nearly consecutive and extremely large year classes have entered the population (Goethel et al., 2020). Although the mechanistic drivers of these large recruitment

events remain unclear (Shotwell et al., 2020), the resource complexion is now dominated by young, small, and primarily immature fish. Consequently, abundance and biomass has rebounded quickly from the lowest points on record in the mid-2010s to near historically high levels in recent years (Goethel et al., 2020). However, due to the partial maturity of these recent cohorts, spawning stock biomass (SSB), which forms the basis of the North Pacific Fisheries Management Council's (NPFMC) B_{40} harvest control rule (HCR), has yet to demonstrate as rapid of a recovery (although SSB has increased from the all-time low in 2018). Potentially associated with the influx of multiple unprecedented year classes, there have been apparent changes in condition (Shotwell et al., 2020) and potential impacts on growth and maturity (Echave, 2021; Williams and Rodgveller, 2021). However, many of the biological parameters input into the sablefish stock assessment model have not been updated in over a decade. For instance, growth and weight were last analyzed and updated for the 2008 assessment (Hanselman et al., 2007; Echave, 2021), while the maturity curve used in the 2020 SAFE (and all previous assessments) was developed by Sasaki (1985) using data from the late 1970s and early 1980s (Williams and Rodgveller, 2021). Given the large amount of data on length, weight, and maturity collected on the annual sablefish longline survey, similar samples collected by at-sea observers, and targeted sampling of maturity status during winter spawning months using histological data (Rodgveller et al., 2016, 2018), there is now ample information to explore updates to sablefish growth and maturity curves for the stock assessment.

Concomitant with changes in the resource, there have been rapid shifts within the directed fixed gear individual fishing quota (IFQ) and non-target trawl sectors. Associated with the extreme recruitment events, both sectors have been inundated with catch of small, comparatively low value sablefish (Goethel et al., 2020). Increasing abundance of juvenile sablefish in the eastern Bering Sea has led to increases in sablefish bycatch in the pelagic trawl fisheries in that region (Goethel et al., 2020). Similarly, the rapid increase in catch of small sablefish by the directed fixed gear sector (i.e., including both longline and pot gear types), has led to exploration of regulations to allow the release of small sablefish within the fixed gear sector (NPFMC, 2021). At the same time as the increase in catch of small sablefish, there has been a swift transition from longline to pot gear within the fixed gear sector in the Gulf of Alaska (i.e., with over 50% of the total IFQ landings of sablefish coming from pot gear in 2020; Goethel et al., 2020). In particular, the development of collapsible 'slinky' pots has allowed smaller vessels that were unable to utilize rigid pots to explore the use of pot gear; slinky pots are also less expensive than traditional pots making them more enticing for a wider array of sablefish IFQ stakeholders. The increase in pot gear is likely due to a combination of the wider utility of the slinky pots along with the increases in sperm whale depredation in the Gulf of Alaska on longline gear (Hanselman et al. 2019), which pot gear essentially eliminates. Additionally, the ability to incorporate escape rings into pot gears can help reduce the number of small sablefish landed and potentially increase the overall value of the landed catch (i.e., given that the IFQ fishery operates under a mandatory 100% retention regulation and small fish have lower value per pound). The increase in landings of small sablefish began in 2016 (associated with recruitment at age-2 of the large 2014 year class) and the shift towards pot gear began in 2017 when it was legalized in the GOA region (Goethel et al., 2020). Under such management and fleet changes, Wilberg et al. (2009) suggest that incorporating time-variation in fishery CPUE catchability represents best practice for stock assessment models. Thus, given the rapid change in gear composition in the fixed gear fishery, there is a need to explore whether the selectivity and catchability (i.e., associated with the fishery CPUE index) of the aggregated fixed gear fishery modeled in the assessment (i.e., combining all longline and pot gear into a single fleet) has altered in recent years.

Around the same time, the sablefish longline survey began observing large numbers of young fish. For instance, the survey age and length composition has been dominated by fish from the 2014, 2016, and 2017 year classes for the last five years. Due to the influx of young, small fish, the resultant longline survey abundance index has increased 2.5 fold since 2015, which is the year with the lowest index value on record. In 2020 the longline survey abundance index again increased by 30% from the 2019 value (Goethel et al., 2020). Although the increasing abundance indices are being driven by extreme recruitment events, there appears to be an increase in catch of small fish in deeper waters where they have

historically been rare. The mechanism driving the increases in catch of small fish in deeper water survey stations remain unknown, but it could be due to density-dependent effects (i.e., 'spillover' out of preferred juvenile habitat) or changes in water temperature. For instance, warmer shallow water may force young sablefish into deeper, cooler water at earlier ages. Such changes in apparent availability of small fish would influence survey selectivity. Given these changes in resource distribution and fishery composition, there is impetus to explore alternate model parametrizations for both selectivity and catchability (i.e., the addition of a new fishery and survey selectivity and fishery CPUE catchability time blocks all starting in 2016) to ensure that internal scaling and the effective age and length based selectivity being estimated by the model are appropriate.

Because catchability coefficients directly scale observed abundance or biomass indices to the actual total estimated population size, how these coefficients are parametrized within a stock assessment model can have important implications for determination of stock status and sustainable harvest levels (Wilberg et al., 2009). Currently, the sablefish assessment utilizes prior distributions for all catchability coefficients to ensure common scaling across indices as well as allowing longline indices to be temporally linked (i.e., maintain commonality in scaling across the assessment time frame where the longline survey transitioned from being run by Japanese scientists to a cooperative Japanese-U.S. survey, then eventually becoming run solely by the AFSC; Hanselman et al., 2007). Prior to the development of the catchability prior distributions, the sablefish model was sex-aggregated and had fewer abundance indices, thus fewer parameters. The domestic longline survey catchability was estimated freely, and the cooperative survey catchability had a fixed offset based on Kimura and Zenger (1988). The development of prior distributions was enacted to allow uncertainty in the link between abundance indices. But, it also served to stabilize parameter estimates, because, at that time, there was a more limited time series of data for many of the inputs to the assessment, particularly as the parametrization was moved to a two-sex model. However, stock assessment best practices generally suggest treating catchability parameters as free parameters to ensure adequate internal scaling within the assessment. Additionally, over the last few years, internal reviewers of the sablefish SAFE have requested explorations of model sensitivity to removal of the catchability priors.

Concurrent with exploring alternate parametrizations of selectivity and catchability, it is often advised that data weighting assumptions be refined to ensure that no single data input has undue influence on the model results and that the information content from abundance indices is adequately utilized (Francis, 2011, 2017). Given that selectivity and catchability can have a strong influence on internal scaling of the assessment model, if the parametrization of these values is altered it is important to ensure that the data are still being fit appropriately. Since 2016, the sablefish assessment has assumed fixed data weights based on advice during the 2016 CIE review. Recommendations from the CIE panel suggested that the longline survey index was being fit too precisely and the resultant proposed fixed data weights aimed to ensure that the age and length compositional data were more closely fit. Unfortunately, these weights were fixed prior to the influx of small fish and subsequent changes in the resource and fishery. In recent years, the assessment model has begun to demonstrate increasing retrospective patterns, primarily associated with uncertainty in the estimates of large recent year class strength. For instance, the estimate of the 2014 and 2016 year classes have been subsequently downgraded as new data have been incorporated into the model, with reductions to the 2014 year class exceeding 60% between first being estimated by the 2017 SAFE model and the current 2020 SAFE model estimate. Concurrently, the assessment model has demonstrated a propensity to predict longline survey abundance index values that are much larger than observed (e.g., overprediction by as much as 30% in recent years), which has led to potential overestimation of recruitment levels (i.e., as indicated by the retrospective patterns). The combination of priors on survey and fishery catchability as well as fixed data weights (i.e., with no use of data reweighting methods) could be potential sources for the observed retrospective patterns and degraded fit to recent longline survey data. Refining the model parametrization in combination with data reweighting methods (e.g., Francis reweighting) to better fit survey abundance indices may help reduce retrospective patterns and should be explored further.

Given the rapid changes in the resource and fishery and concomitant increases in retrospective patterns in the assessment, exploring a variety of potential model changes to the Alaskan sablefish assessment has been a high priority in recent years. Approaches for updating the biology (i.e., growth, weight, and maturity), model parametrization (i.e., addition of a recent fishery and survey selectivity time block along with alternate approaches to estimating catchability), and data weighting are described and the results compared, particularly in reference to the 2020 SAFE model. Also, an update on data availability for the 2021 SAFE is provided in Appendix A.1, focusing primarily on limitations associated with updates to the fishery CPUE index. The final proposed model being recommended based on this work for the 2021 sablefish SAFE makes important strides towards better representing sablefish biology, while reducing retrospective patterns and improving model stability.

Methods

Stock assessment model updates and explorations were grouped into three categories: 1) biological inputs; 2) model parametrization; and 3) data reweighting (see Table 1 for a list of model scenarios). Each model update was implemented individually to demonstrate the impact of each change as a one off alteration to the current 2020 Sablefish SAFE model (termed the Continuity model, *16.5_Cont*). A stepwise model building process was then implemented within each of the biological and model parametrization categories using a semi-factorial design (i.e., most, but not all, combinations of model changes were tested in a step-wise fashion, though not all model building steps are presented). For simplicity of presentation, we focus on the results of each update that is being recommended for inclusion in the final 2021 SAFE assessment. Although a variety of alternate model changes were tested, those deemed inappropriate, unrealistic, or otherwise unfit for operational assessment purposes (e.g., due to poor model performance, including poor fits to the data, unrealistic outputs, or stability issues) are not presented. Results within each category of model building are presented, then the final proposed model with data reweighting applied is compared to the continuity model (i.e., 2020 final SAFE model, *16.5_Cont*). Finally, results of important model diagnostics (i.e., data fits, residual patterns, retrospective analysis, and Markov Chain Monte Carlo) are analyzed and compared between the proposed model (*21.10_Proposed*) and the continuity model (*16.5_Cont*).

Continuity Model (16.5_Cont)

The 2020 sablefish SAFE final accepted assessment model (termed the Continuity model, *16.5_Cont*) is used as the basis of one off changes and baseline comparisons. The Continuity model is outlined in Goethel et al. (2020) and is implemented here exactly as it was for the provision of management advice in 2020. Model building towards the final proposed model for the 2021 sablefish SAFE (*21.10_Proposed*) is first undertaken within each group before the 'best' or most appropriate changes are combined into the final model. Each model scenario explored and discussed in this document is outlined in Table 1.

It is worth noting that all models subsequent to *16.5_Cont* contain a single alteration to the trawl fishery selectivity parametrization to improve model stability and better match the assumptions inherent in the selectivity parametrizations of the other fishery and survey fleets. Essentially, the parameter determining the shape of the gamma selectivity function for the trawl fishery was altered to be shared between males and females, which matches how the fixed gear fishery and longline survey logistic selectivity shape parameters are treated. The change in parametrization leads to subsequent models having one less selectivity parameter to estimate compared to the Continuity model (*16.5_Cont*), but impacts on model results were negligible and not discussed further.

Biological Inputs

Three potential updates to biology were considered, including updating the length-, weight-, and maturity-at-age. However, because growth and weight are intertwined processes, changes to these inputs are considered as a single update.

Growth and Weight

Growth and weight were last updated in the sablefish assessment in 2008 with data through 2004 (Hanselman et al., 2007). Data from two time periods (1981 – 1993 and 1996 -2004) were utilized to define and model two growth regimes (pre- and post-1995; Figure 1) where the time series breaks were determined primarily by changes in sampling design for sablefish data collected on the longline survey and used to estimate growth (Echave, 2021). Conversely, weight was not collected on the longline survey prior to 1996, so a single weight-at-age curve has been utilized for the entire assessment model time series using data collected from 1996 – 2004 (Figure 2).

Based on updated data through 2019 and the results of a cluster analysis, Echave (2021) recommended that both the growth curve and weight-at-age be updated. Additionally, results suggested that a new time block be added (i.e., pre- and post-2004) to account for apparent changes in growth over the last decade. However, at the moment there are no explicitly known biological or environmental mechanisms that might be driving growth changes since 2004. As such, the added complexity of an additional growth block was not considered in the current analysis. Thus, for updates to weight and growth, we utilized the results from Echave (2021) utilizing the single time block model and all available data through 2019 (i.e., with no additional time blocks in either process). The historic growth curve (pre-1995) remains unchanged, while the recent (post-1995) growth curve was updated with the new data (Figure 1). Similarly, the weight curve was updated with new data through 2019 and applied for all years in the assessment model (Figure 2). As noted, weight and growth were updated simultaneously and treated as a single model update (*21.1_Wt+Grt*; Table 1). The updated weight and growth parameters were maintained in subsequent models including *21.5_Upd_Bio_AL-Mat* and *21.10_Proposed*.

Maturity

Maturity in the sablefish assessment has always utilized a consistent age-based maturity curve developed by Sasaki (1985; see Figure 3), which was based on macroscopic maturity classifications and lengths collected in the summer during the late 1970s and early 1980s. There are a variety of potential issues with using these historic maturity estimates, especially considering potential recent changes in maturity and the documentation of skipped spawning in sablefish (Rodgveller et al., 2016, 2018; Goethel et al., 2020; Williams and Rodgveller, 2021). Maturity data collected using histological (as opposed to macroscopic) methods provide a more accurate determination of sablefish maturity, including skipped spawning. Moreover, utilization of General Additive Models (GAM) can better account for skipped spawning that cannot be adequately addressed using the more commonly applied General Linear Models (GLMs; Trippel and Harvey, 1992; Williams and Rodgveller, 2021). Additionally, given that maturity is typically dependent on a mixture of both age and length processes, models that account for both the length and age of mature fish are likely to better reflect true population maturity rates. For sablefish, Williams and Rodgveller (2021) demonstrate that an age-length GAM based on histological samples of sablefish and accounting for skipped spawning is likely to provide the most reliable estimate of maturity-at-age. Although data on skipped spawning is limited to three directed studies and the rate is variable, simulations demonstrate that ignoring skipped spawning when it is present is likely to cause increased bias compared to incorporating skipped spawning in maturity estimates but getting the average population rate of skipped spawning slightly incorrect (Williams and Rodgveller, 2021).

Given the recommendations of Williams and Rodgveller (2021), three maturity curves based on analysis of the histological data were explored. First, maturity was updated using an age-based general linear model (GLM) that ignored skipped spawning information, but utilized the recent histological data

(*21.2_Mat_Age_GLM_No_SS*). Although not a strict update of the maturity curve, this approach was deemed the most consistent with the methods of Sasaki (1985), but utilizing the more reliable and recently collected histological maturity information. However, given that skipped spawning has been observed for sablefish, use of the GLM maturity model was not recommended for further use (Williams and Rodgveller, 2021). Next, an age-based GAM maturity model, which includes skipped spawning information, was implemented using the histological data (*21.3_Mat_Age_GAM*). Finally, the recommended age-length based GAM maturity model, which also accounts for skipped spawning and uses the histological data, was utilized (*21.4_Mat_AL_GAM*). Because the latter model is partially based on length, changes to the growth curve cause changes to maturity-at-age. Thus, even though the maturity parameters are constant through time, the resultant maturity-at-age will change based on growth regimes in the assessment model. The input maturity-at-age based on the age-length GAM, therefore, differs before and after the growth time block in 1995, but also differs due to different underlying growth parameters in the Continuity model (*16.5_Cont*) and subsequent models that utilize the updated growth curves (note that maturity is input to the model, so the impact of changes in growth on maturity are calculated externally and input into the assessment). Based on recommendations by Williams and Rodgveller (2021) and the results of the current work, we utilize the age-length GAM maturity model for subsequent model building (i.e., *21.5_Upd_Bio_AL-Mat* and *21.10_Proposed*), but note that the input maturity for these models differs from that used in *21.4_Mat_AL_GAM* due to changes in the underlying growth parameters. Each of the maturity curves utilized in the various models are provided in Figure 3.

Update All Biology

The final model building scenario in the 'Biology Update' category implemented the combination of updated growth, weight, and maturity (*21.5_Upd_Bio_AL-Mat*). As noted, the model utilized the age-length GAM maturity model, but based on the updated growth parameters.

Model Parametrization

Increasing retrospective patterns over the last few years have provided impetus to explore alternate model parametrizations to better fit the observed data and address changes in fishery and resource dynamics. During the 2020 SAFE a wide variety of sensitivity runs were explored (Goethel et al., 2020). Of these, adding a recent selectivity time block to address apparent changes in targeting and availability of young sablefish in the fixed gear fishery and longline survey demonstrated the most promise, while also being the most defensible based on direct observation and knowledge of sablefish biology and harvesting. Similarly, removal of catchability priors has been consistently highlighted as a relatively straightforward potential model change that could improve scaling and model performance. Thus, both of these changes to model parametrization were explored further. It is worth noting that natural mortality has been consistently noted as needing further exploration and potential parametrization refinement within the sablefish model. Although recent analysis led to improvements in the estimation of natural mortality using priors (Hanselman et al., 2018) and alternate age- and time-varying parametrizations were explored in-depth for the 2020 SAFE, added complexity to the natural mortality formulation has often led to increased model instability along with seemingly unrealistic model outcomes (Goethel et al., 2020). Natural mortality will continue to be explored in the future, particularly in association with the goal of developing a tag-integrated assessment for sablefish, but no new formulations were explored or will be put forward for inclusion in the 2021 SAFE model.

Removal of Catchability Priors

As a direct scalar between the indices of abundance or biomass and the estimated population size, adequate parametrization of catchability coefficients is crucial within assessment models. As noted, the 2020 SAFE model (*16.5_Cont*) assumed priors on all catchability parameters to maintain consistent scaling across surveys and aid in model stability. However, the use of priors was implemented in 2007 and has not been addressed since that time, despite over a decade of additional data. To determine the

impact of using catchability priors and to explore whether these parameters can be freely estimated, model *21.6_No_q_Prior* treated all catchability coefficients as freely estimated parameters (Table 1). Although the number of estimated parameters did not explicitly change, the six catchability parameters were moved from constrained parameters to freely estimated. For subsequent model building scenarios, including *21.8_No_q_Add_Sel+q_Block* and *21.10_Proposed* models, all catchability coefficients were maintained as freely estimated parameters.

Addition of a Recent (Post-2016) Selectivity and Fishery CPUE Catchability Time Block

Sensitivity runs during the 2020 SAFE demonstrated that adding a selectivity time block in 2016 for both the fixed gear fishery and the longline survey improved fits to the longline survey relative population numbers (RPN) index, fishery and survey compositional data, and fishery CPUE index (Goethel et al., 2020). Additionally, given the rapid changes in the fixed gear fishery and data inputs associated with the fishery CPUE index, there is rationale to incorporate an associated recent time block for estimation of the fishery catchability coefficient. Adding a recent time block for fishery selectivity and associated catchability essentially assumes that fishery dynamics have changed, likely due to a combination of alterations in targeting behavior (i.e., to avoid large recent year classes of small, low-value sablefish) or distribution of gear types (i.e., an increasing shift towards pot gear and away from longline gear). A similar time block for survey selectivity implicitly assumes that availability to the survey gear has changed (i.e., young fish have moved into survey areas in recent years, mainly in deeper waters where they have not typically been sampled in the past). The 2020 SAFE model (*16.5_Cont*) had trouble rectifying the rapidly increasing survey index, the influx of large numbers of small and young fish in both the survey and fishery compositional data, and the relatively stagnant fishery CPUE index. The added flexibility provided by adding a post-2016 selectivity time block along with an associated fishery CPUE catchability time block will likely allow the model to better rectify conflicting signals within the various data sources and potentially account for processes that cannot be explicitly modeled (e.g., changes in targeting on low-value small fish, increasing use of pot gear, and potential redistribution of small fish into areas not previously inhabited). To address these potential changes, *Model 21.7_Add_Sel+q_Block* estimated new fishery and survey selectivity parameters (i.e., $a_{50\%}$ for males and females; four parameters) along with a new fishery CPUE catchability (*q*) coefficient (i.e., one parameter) resulting in a total of five additional parameters to be estimated (Table 1). The added time block and additional estimated parameters were maintained for the *21.8_No_q_Add_Sel+q_Block* and *21.10_Proposed* models.

Update All Model Parametrization

The final model building scenario in the 'Model Parametrization' category implemented the combination of removing catchability priors and allowing for a recent (post-2016) time block for fishery and survey selectivity and fishery CPUE catchability (*21.8_No_q_Add_Sel+q_Block*).

Data Weighting

Ensuring that a model adequately fits the available data is a prerequisite for developing a robust stock assessment. When fitting both abundance and compositional data in a model, data conflicts are common and determining appropriate statistical weights for each data source can be difficult. It is now considered best practices to perform reweighting procedures (e.g., Francis or McAllister-Ianelli) to ensure the model is 'right-weighted' and no single data source is dominating the negative log-likelihood and resulting model outputs (Francis, 2011, 2017). Additionally, it is suggested that reweighting procedures should be undertaken as the final step in the model development process to ensure consistent data weights that match the final assumptions and modeled processes (Maunder et al., 2017). Although a variety of reweighting approaches exist, the Francis method has been explored for other North Pacific species (e.g., GOA pollock and blackspotted/rougheye rockfish) and has been demonstrated to provide generally robust weights. Additionally, it can account for correlations among ages or length bins in the compositional data

by iteratively adjusting the data weights such that model mean age or length reflects the mean age and lengths observed in the compositional data.

The 2020 SAFE model (*16.5_Cont*) used fixed input data weights based on recommendations from the 2016 CIE review and these weights have not been altered since that review. The re-weighting that occurred at that time was based on targeting a standard deviation of normalized residuals (SDNR) approximately equal to one for each of the age and length (i.e., when no ages were available from a given fleet) compositional data sources. Exploratory analysis during the 2020 SAFE suggested that these fixed weights could be one potential source for increasing retrospective trends. Thus, we implement Francis reweighting with the continuity model (*21.9_Cont_Francis*) to determine whether the reweighting appears to lead to better data fits or alternate interpretations of the dynamics. Similarly, the final proposed model for 2021 (*21.10_Proposed*) utilizes Francis reweighting, as well (see 2.5 Final Proposed Model).

The methods applied for data reweighting follow Francis (2011) where the abundance index weights were fixed based on the input observed variance of each index and the compositional data weights were iteratively adjusted using a two stage approach. In Stage 1, the model was run with starting input compositional data (i.e., all sources of age and length composition data fit in the model) weights (exploratory runs demonstrated that final weights were insensitive to initial weights). Then, the compositional data weights were adjusted following Method TA1.8 and weighting assumption T3.4 of Francis (2011, Appendix Table A1; i.e., using the assumption of a multinomial distribution and accounting for correlations among ages or length bins). In Stage 2, the model was then rerun with the new weights. The weights were iteratively adjusted until the difference between the current weights and the revised weights were minimal (i.e., the weights converge; for sablefish this usually took less than 10 iterations).

Final Proposed Model

The final model being proposed for the 2021 sablefish SAFE (*21.10_Proposed*) combines the results of each model building stage (i.e., the final 'Biology Update', *21.5_Upd_Bio_AL-Mat*, and 'Model Parametrization', *21.8_No_q_Add_Sel+q_Block*, models), then Francis reweighting was applied. The final proposed model (*21.10_Proposed*) is analyzed in depth, particularly in comparison to parameter estimates and data fits of the 2020 accepted SAFE model (*16.5_Cont*).

Model Performance Criteria

A variety of performance criteria were utilized to determine model stability, adequacy, and robustness, which were compared across models. Model convergence was a minimum requirement to be considered further and this was gauged by having a maximum gradient component < 0.001 and a positive-definite Hessian matrix. A critical component of determining model performance was the fit to the data, particularly the tradeoff between age composition data from the fixed gear fishery and longline survey compared with the fit to the longline survey RPN and fishery CPUE indices. Similarly, residual patterns were explored visually to determine if any major patterns with time, age, or length were present. Although the negative log-likelihood (*nLL*) was utilized to gauge data fits, these were not necessarily directly comparable (e.g., due to changes in data weights and penalty terms).

For comparing the 2020 SAFE Continuity (*16.5_Cont*) and 2021 Proposed (*21.10_Proposed*) models, a full suite of diagnostic analyses were undertaken, including time series of model outputs, data fits, retrospective analysis, and Markov Chain Monte Carlo (MCMC). For the retrospective analysis, ten year data peels were utilized and Mohn's rho was calculated for terminal year SSB across all peels. Given the recent selectivity and catchability time block in the *21.10_Proposed* model, difficulties arise when performing retrospective peels before 2018. Essentially, very little data exists to estimate these additional parameters for 2016 and 2017 model peels. Although we still present the results of the retrospective analysis for all years, care should be taken when analyzing peels before and after 2018 given that these are fundamentally different models (we removed the recent selectivity and catchability time block for all

peels prior to 2018). For MCMC runs, the posterior distributions were computed based on one million MCMC simulations. The chain was thinned to 5,000 parameter draws to remove serial correlation between successive draws and a burn-in of 10% was removed from the beginning of the chain. MCMC runs were utilized to provide 95% credible intervals around estimates of SSB, biomass, and recruitment.

Results

All of the models explored converged with adequate maximum gradient values, indicating that most models were stable and parameter correlation was not a major issue (Table 2). Additionally, all models resulted in generally similar negative log-likelihood values (though these necessarily differed dramatically for models utilizing reweighting approaches), parameter estimates, and population time series trends, which lends further credence that a global minima in the likelihood surface is being achieved (Tables 2-3, Figures 4, 5, and 7).

Biology Updates

In terms of growth and associated weight, adding the full complement of data generally led to fish growing a bit slower yet reaching a larger maximum size in the recent (1996 – 2019) period (Figures 1-2). Updating maturity led to more complicated dynamics, given the range of maturity models (i.e., age-based and age-length based) explored. Using an age-based GLM and ignoring skipped spawning (model *21.2_Mat_Age_GLM_No_SS*) led to increased maturity at younger ages compared to Sasaki (1985; used in the *16.5_Cont* model), but slightly reduced maturity at ages five through twelve. Updating the maturity-at-age using an age-based GAM and the available histological data (model *21.3_Mat_Age_GAM*), including information on skipped spawning, led to decreases in the maturity at all ages, but particularly young and intermediate ages (i.e., ages two through five and nine through seventeen; Figure 3). The impacts on younger ages is due to the use of a more flexible GAM (as opposed to a GLM) that better reflects true maturity when skipped spawning occurs at intermediate ages, while the decreases at intermediate ages is directly associated with increased skipped spawning observed for these age classes (Williams and Rodgveller, 2021). The age-length model (models *21.4_Mat_AL_GAM* and *21.5_Upd_Bio_AL-Mat*) provided similar age-based maturity estimates to the values from Sasaki (1985; used in the *16.5_Cont* model) for the early time period (pre-1996), but with similar decreases at intermediate ages (i.e., associated with skipped spawning as demonstrated in the updated age-based model; Figure 3). In the recent time block (post-1996), maturity at younger ages was much lower for both growth models (i.e., based on the growth models used in the 2020 SAFE and the updated growth model), but then exceeded maturity from Sasaki (1985) for ages five through eight when the old growth curve was utilized (i.e., model *21.4_Mat_AL_GAM*). Conversely, using the age-length maturity model in combination with the updated growth curves (*21.4_Upd_Bio-AL-Mat*) led to maturity-at-age values that were consistently lower than Sasaki (1985) for all ages during the recent time block (post-1996).

Generally, updating the biological parameters did not alter model performance, parameter estimates, population trajectories, or fits to the data in any significant manner (Tables 2-3, Figure 4). As expected, the primary impact was to rescale SSB and associated reference points, which directly influenced stock status and harvest recommendations (Table 3). In particular, updates to growth and weight (model *21.1_Wt+Grt*) led to higher SSB, but similar increases in reference points given the new larger maximum size and weight; thus, the ABC decreased, because stock status was slightly lower than for model *16.5_Cont* (Table 3).

Utilizing the age-based maturity GLM without skipped spawning information (model *21.2_Mat_age_GLM_No_SS*) moderately reduced SSB, but did not greatly alter SSB trajectories (Figure 4). However, the biomass-based reference point only decreased slightly compared to a relatively strong reduction in terminal year SSB, which resulted in a lower stock status compared to the Continuity model (*16.5_Cont*) and a subsequent reduction in ABC (Table 3). Updating the age-based maturity curve (model

21.3_Mat_Age_GAM) had the largest impact due to the decreases in maturity-at-age, which led to strong scaling changes, including reductions in SSB albeit associated with similar reductions in the biomass reference point (Table 3, Figure 4). Utilizing the age-length maturity curve (model *21.4_Mat_AL_GAM*) led to an SSB trajectory about midway between the continuity (*16.5_Cont*) and updated age-based GAM maturity model (*21.3_Mat_Age_GAM*) and very similar to the trajectory of the model using the age-based GLM without skipped spawning information (*21.2_Mat_Age_GLM_No_SS*; Figure 4). But, the biomass-based reference point actually increased slightly compared to model *16.5_Cont* (Table 3), likely due to increased maturity values for the most abundant ages (i.e., ages five through eight) in the recent time block (post-1996).

The final 'Biology Update' model (*21.5_Upd_Bio_AL-Mat*), which incorporated the updated weight and growth curves along with the new age-length based maturity curve, closely matched the continuity model dynamics (*16.5_Cont*) for much of the time series, but estimated slower rebuilding in SSB over the last few years (Figure 4). The lack of rebuilding in SSB is due to the updated age-length maturity curve indicating that maturity of young and intermediate aged fish is much lower than assumed in the Sasaki (1985) maturity curve utilized in the continuity model (*16.5_Cont*). Therefore, because much of the population increase in recent years has been due to large 2014, 2016, and 2017 year classes, the *21.5_Upd_Bio_AL-Mat* model implies that these year classes are not as mature as previously assumed and SSB has not recovered as quickly. Conversely, primarily due to the changes in weight and growth, the associated biomass-based reference points have increased (Table 3). The dichotomous change in terminal SSB and biological reference points compared to the continuity model (*16.5_Cont*) leads to a significant decrease in the projected 2021 ABC (40 kt in the *21.5_Upd_Bio_AL-Mat* model compared to 52 kt in the *16.5_Cont* model).

Model Parametrization Updates

Similar to the biology updates, there were no major changes in general population trajectories, but magnitude and scale differed for the 'Model Parametrization' updates, especially in terms of the estimated strength of recent recruitment events (Tables 2-3, Figure 5). Similarly, fits to the data, particularly the longline fishery CPUE index, demonstrated some important deviations across model scenarios (Figure 6). Again, the primary impact across model scenarios was rescaling of SSB, associated reference points, and subsequent harvest recommendations (Table 3).

The main effect of allowing the catchability parameters to be freely estimated (model *21.5_No_q_Prior*) was minor variation in the catchability estimates (Table 4), which led to a rescaling of the SSB time series and slight reductions in recruitment estimates compared to the continuity model (*16.5_Cont*; Figure 5). Although the reference points did not change to any great extent, the terminal year SSB was slightly lower than in the Continuity model (*16.5_Cont*) resulting in a reduction in the ABC (Table 3). There was no appreciable change in model fits to the data, especially the abundance indices (Figure 6).

Adding a time block to the longline survey and longline fishery selectivity and fishery CPUE catchability (*21.6_Add_Sel+q_Block*) resulted in strong improvements in the fit to the fishery CPUE index (Figure 6). Concomitantly, the addition of the recent selectivity and catchability time block strongly reduced recent year class strength (Figure 5). The SSB time series was rescaled to a similar level as model *21.6_No_q_Prior* with a slightly decreased terminal year SSB estimate, but the biomass-based reference point underwent similar reductions and stock status only decreased slightly compared to the Continuity model (*16.5_Cont*; Table 3). However, the reduction in recent recruitment had a strong impact on projected biomass and rebuilding rates, which led to large reductions in the ABC (~35 kt), because the high projected ABC in future years associated with the Continuity model (*16.5_Cont*; Figure 5) are due to exceptionally high (and uncertain) recent year class estimates. The main factors driving the results of model *21.6_Add_Sel+q_Block* were estimated increases in selectivity of young fish (e.g., ages two through four; see Figure 14 for an example of changes in selectivity for the 'recent' time block from the *21.10_Proposed* model) and associated decreases in fishery catchability after 2016 (see Table 4 for *q*

estimates from the *21.10_Proposed* model). By reducing fishery catchability, the model was able to better rectify fishery CPUE, which underwent a strong reduction in 2016 and has yet to recover, with the longline survey index that has increased dramatically over the last five years (Figure 6). However, the increased selectivity estimates on younger fish forces the model to downgrade recruitment estimates. Allowing an increase in survey and fishery selectivity since 2016 allowed the model to interpret the increasing proportion of small, young fish in the composition data as a mixture of a change in availability, as well as, large year classes.

The final 'Model Parametrization Update' model (*21.8_No_q_Add_Sel+q_Block*), which removed the catchability priors and added the recent fishery and survey selectivity and fishery CPUE catchability time blocks, underwent a similar rescaling of the overall SSB as the *21.6_Add_Sel+q_Block* model (Figure 5). But, further reductions in recent recruitment estimates compared to previous models led the terminal SSB estimate and resulting stock status to be considerably more pessimistic compared to the continuity model (*16.5_Cont*; Table 3, Figure 5). The resulting downgrades in each of the 2014, 2016, and 2017 year classes was on the order of 20-50% compared to the Continuity model (*16.5_Cont*; Figure 5). Fits to the data generally followed the trends of model *21.6_Add_Sel+q_Block* with improved fit to the fishery CPUE data set compared to previous models. Most importantly, model *21.8_No_q_Add_Sel+q_Block* estimated a considerably lower terminal year SSB (74 kt) with only a slight decrease in the biomass-based reference point compared to model *21.6_Add_Sel+q_Block*, which resulted in an ABC of 29 kt. Once again, the reduction in ABC compared to the Continuity model (*16.5_Cont*) was strongly influenced by the large comparative reductions in recent recruitment estimates.

Data weighting comparisons

When Francis reweighting was applied to the Continuity model (*16.5_Cont*), the resulting model (*21.9_Cont_Francis*) estimated smaller recent recruitment events (particularly for the 2017 year class; Figure 7) and demonstrated much better fits to the longline survey RPN index (Figure 8), as well as the trawl survey biomass index (not provided). Although no strong scaling changes occurred in terms of SSB, the *21.9_Cont_Francis* model was more optimistic in terms of population trajectory in recent years with SSB not declining as rapidly and rebuilding quicker than the Continuity model (*16.5_Cont*; Figure 7). Additionally, the terminal SSB was higher and the biomass-based reference point was considerably lower than in the Continuity model (*16.5_Cont*; Table 3). However, because recruitment estimates for the 2017 year class were much smaller, the projected ABC decreased slightly from the *16.5_Cont* model.

As is expected from the Francis method, many of the final weights given to the compositional data were lower than the fixed weights used in the Continuity model (*16.5_Cont*; Table 5). The implicit downweighting of the compositional data allowed the model to better fit the index data (Figures 8-9), which is extremely important in terms of model interpretation of recent year class strength. Recent year classes are notoriously difficult for integrated models to accurately estimate (i.e., due to only a handful of data observations of these events) and the rapid changes in apparent resource productivity as observed by the influx of young, small fish in the survey and fishery age and length composition data has led to large uncertainty in the Continuity model (*16.5_Cont*) estimates of recent recruitment events. Although these year classes were estimated to be historically large, they have been undergoing large retroactive downgrades as more data on the strength of these recruitment events have become available (Goethel et al., 2020). The higher emphasis given to compositional data forced the *16.5_Cont* model to closely fit the rapid shift in the composition data since 2016, which led to the unprecedented estimates of year class strength at the cost of greatly overestimating all of the abundance and biomass indices (e.g., by more than 30% in the case of the longline survey RPN index). The converse is true after reweighting in the *21.9_Cont_Francis* model, where the survey indices are now better fit and recruitment estimates are slightly decreased.

However, the resultant data weights provide greater emphasis to fixed gear fishery length compositions over all other compositional data sources, including the associated fishery age compositions. These

weights are surprising given the relatively large number of sablefish otoliths sampled each year to determine age compositions (i.e., more than 1000 samples are taken from both the fishery and longline survey). Conversely, the longline survey age composition data is given more weight than the associated length compositions, but still lower relative weight than the fishery length composition data. It is unclear what underlying factor is driving the relative weights developed during the reweighting analysis, but there is likely model tension due to simultaneously fitting the length, age, and abundance index data sources. Additionally, uncertainty associated with assigning ages for young fish associated with the large recent recruitment events (i.e., ageing imprecision leading to a 'smearing effect' across large, consecutive year classes; Beamish and McFarlane, 1995) might be causing model difficulty rectifying age and length composition interpretations of year class strength for the 2014, 2016, and 2017 year classes. The decreased emphasis of fishery age composition data does lead to degraded fit to these data after reweighting (as discussed for the *21.10_Proposed* model; Figure 20).

It is important to note that the resulting data weights explicitly counter the recommendations of the 2016 CIE, which suggested that the longline survey index was being too closely fit at the expense of the compositional data (hence the recommendation to increase the weights of the composition data). However, these recommendations were developed before the large 2014 and subsequent 2016 and 2017 year classes began to be observed in the data and did not account for the resultant extreme overestimation of the survey index. Although not presented, retrospective patterns were considerably reduced for the *21.9_Cont_Francis* model compared to the *16.5_Cont* model, where the reweighted model demonstrated higher stability and fewer model scaling issues when data were removed in subsequent peels.

Final proposed model

The final proposed model for the 2021 SAFE incorporates the improvements noted in each set of model building exercises, including updating weight, growth, and maturity (i.e., using the age-length maturity model), removing catchability priors, allowing for a recent fishery CPUE catchability along with fishery and survey selectivity time block, then using the Francis method to reweight the final model. In general, model *21.10_Proposed* melds the mixture of trends and changes from each of the *21.5_Upd_Bio_AL-Mat*, *21.8_No_q_Add_Sel+q_Block*, and *21.9_Cont_Francis* models. The resulting population trend and scale in terms of SSB is very similar to the *16.5_Cont* model, but with a more pessimistic trend during the mid-1990s and early 2000s followed by a more optimistic trend (i.e., flat instead of declining) over the last five to ten years similar to the *21.9_Cont_Francis* model (Figure 7). Conversely, recent recruitment estimates are severely decreased compared to the *16.5_Cont* model and, notably, the 2017 year class estimate is even much lower than estimated in the *21.8_No_q_Add_Sel+q_Block* model (Figure 7). The reductions in recruitment are largely driven by the increased fishery and survey selectivity on younger ages in the recent (post-2016) time block (as discussed for model *21.6_Add_Sel+q_Block*; Figure 14) along with the reductions in relative weight given to the compositional data due to the application of Francis reweighting (as discussed for model *21.9_Cont_Francis*; Table 5, Figure 9). Compared to the continuity model (*16.5_Cont*), the new proposed model (*21.10_Proposed*) provides significantly improved fits to both the fishery CPUE and the longline survey RPN indices (Figure 8). Of particular interest, the fit to the RPN index improves starting in the mid-2010s and continues through the terminal year (Figures 8, 17-18). Whereas the *16.5_Cont* (and all other models developed) predict much lower RPNs than observed from 2014 to 2017 leading to strong declines in SSB during this period, the *21.10_Proposed* model better matches the timing of rebuilding observed in the survey RPN index resulting in a flatter population trajectory over this period (Figures 7-8). Similarly, by not overpredicting the value of the survey RPNs in the last three years, the *21.10_Proposed* model estimates more reasonable recruitment values that better reflect the observed data on abundance, which is then reflected by a more subtle rebuilding over the last three years (Figures 7-8). The terminal SSB estimate in the *21.10_Proposed* model is substantially lower than the *16.5_Cont* model, but the latter has an associated larger biomass-based reference point (likely due to the larger recent recruitment estimates and overall productivity), which results in almost identical stock status between the models (Table 3). However,

because recent recruitment estimates are greatly reduced in model *21.10_Proposed* compared to *16.5_Cont*, the resulting projected SSB does not increase as rapidly nor reach as high a magnitude; therefore, there are large reductions in future ABCs compared to *16.5_Cont* (i.e., an ABC of 27 kt in 2021 for the *21.10_Proposed* model; Table 3, Figure 12).

The final proposed model (*21.10_Proposed*) demonstrates limited retrospective patterns with a Mohn's rho of 8% compared with a value of 17% for the Continuity model (*16.5_Cont*; Figure 10). Although the retrospective pattern is reduced compared to model *16.5_Cont*, the results must be carefully interpreted. Because of the 2016 time block for selectivity and fishery CPUE catchability, models before and after the 2018 peel are not necessarily directly comparable (i.e., for peels before 2018 there is no estimation of new catchability and selectivity parameters for the post-2016 time block). However, models with consistent parametrizations (i.e., 2020, 2019, and 2018 peels; the black, purple, and pink lines in Figure 10) are nearly identical with very minor scaling differences compared to model peels prior to 2017. Perhaps more importantly, the issue of retroactive downgrading in recent recruitment estimates (i.e., for the 2014 and 2016 year classes), which has been an emergent problem since 2018 for the Continuity model (*16.5_Cont*), has been essentially eliminated with the new proposed model (*21.10_Proposed*; Figure 11). Again, it is worth noting the model parametrization difference between the 2017 and 2018 peels, which is clearly visible in the sudden decrease in the 2014 year class estimate (i.e., the original large value is based on the 2017 peel that does not include a recent time block for catchability and selectivity, which the subsequent 2018 peel does include; Figure 11). Consistent estimates of recruitment from one year to the next, as observed with the *21.10_Proposed* model, helps to prevent overly optimistic projected ABC values (Table 6), while also reducing the probability of future overfishing (because there is less probability that projected ABCs will be set too high due to overestimated recruitment). For instance, the impact of the potentially overoptimistic 2014 year class estimates (and subsequent year classes) in the *16.5_Cont* model are clearly observed in the resulting rapid increases in projected ABCs for this model in the retrospective analysis starting with the 2017 peel (i.e., 2018 projected ABC; Table 6). Because the *21.10_Proposed* model has much more modest and stable recruitment estimates, the projected ABCs increase more subtly, though by approximately 7 – 8 kt in each of the last two years. Again, it is worth noting the change in model formulation within the retrospective analysis, which is clearly observed in the sudden decrease in ABC from 2018 to 2019 (i.e., from the 2017 and 2018 retrospective peels, respectively). Overall, it appears that implemented ABCs and realized catch have generally fallen within sustainable thresholds given the projected ABCs from the *21.10_Proposed* model (although there is a potential that landings in 2019 were slightly above desired levels), despite potential overly optimistic projections from the *16.5_Cont* model (Table 6).

Based on the results of the MCMC runs for the *16.5_Cont* and *21.10_Proposed* models, levels of uncertainty appear to be similar for both SSB and biomass (Figures 12-13). Estimates of selectivity and fishing mortality generally agree among models, but with higher selectivity at younger ages in recent years for the new proposed model (*21.10_Proposed*; Figure 14). Fishing mortality was generally similar across the two models with slightly lower values throughout much of the 2010s and a slower reduction in the last few years for the *21.10_Proposed* model (Figures 15). Patterns in recruitment generally match, but, as discussed, the new proposed model (*21.10_Proposed*) estimates that recent year class sizes were much lower than predicted by the Continuity model (*16.5_Cont*; Figure 16). Fit to the observed indices is much improved in the new proposed model (*21.10_Proposed*), particularly the fishery CPUE, longline survey RPN, and trawl survey biomass indices (Figures 17-18). The fits to the age and length composition data in the new proposed model (*21.10_Proposed*) are generally good and reflect similar patterns observed in the Continuity model (*16.5_Cont*; Figures 19-25). But, degradation in the fit to the fixed gear fishery age composition data was observed, particularly due to overestimation of age-2 fish and underestimation of age three to seven fish (Figure 20).

Discussion

Rapid changes in the sablefish resource as well as the associated fisheries have proven difficult to address in the sablefish stock assessment model currently used for management advice, which has led to increasing retrospective patterns and extensive retroactive downgrading of recent year class strength (Goethel et al., 2020). Updated data and analysis of growth, weight, and maturity suggest that values for these biological parameters used in the assessment should be refined (Echave, 2021; Williams and Rodgveller, 2021). Similarly, degrading fits to important data sources (e.g., longline survey abundance indices) suggest that model parametrization and data weighting merit careful consideration. After exploring a variety of model updates and new parametrizations, the results of the current work indicate that five important changes should be considered for the 2021 sablefish assessment: 1) weight and growth for the recent time period (1996 – present) should be updated to reflect the full extent of available data; 2) maturity should be updated with an age-length GAM that accounts for skipped spawning using recent histological data; 3) the catchability priors are no longer needed; 4) a recent time block for estimation of fishery and survey selectivity and fishery CPUE catchability will allow the model to better fit recent data and reduce internal model tension due to slightly incongruous trends in indices and compositional data; and 5) using data reweighting approaches (e.g., the Francis method) can help improve fits to abundance indices, limit retrospective patterns, and reduce retroactive downgrades in recruitment estimates and associated ABCs. It is also worth noting that difficulties updating the fishery CPUE index for the 2021 SAFE (Appendix A.1) may have implications for the final 2021 assessment model, given that there is unlikely to be a 2020 CPUE index data point.

Therefore, model *21.10_Proposed* is suggested as the best model for the provision of management advice for the 2021 assessment year, given that it incorporates each of these model changes. The main impacts of these changes are that maximum weight and growth have increased, but the rate of growth is slightly lower for younger ages (Figures 1-2). Concomitantly, maturity is slightly lower than previously assumed, especially for young and intermediate ages in the recent time period (1996 – Present; Figure 3). The addition of a recent (post-2016) selectivity and fishery CPUE catchability time block suggests that catchability has slightly decreased in the fixed gear fishery (Table 4), but that the selectivity in both the fixed gear fishery and longline survey has increased for younger ages (Figure 14). Finally, recent recruitment appears to be high and well above average levels, but not as extreme as predicted from the Continuity model (*16.5_Cont*); in fact, the recent recruitment trend appears to be similar to the pattern observed in the late 1970s and early 1980s when the largest historic year class (i.e., the 1977 year class) was observed (Figure 16). Model *21.10_Proposed* provides much better fits to the longline survey RPN and fishery CPUE indices, which results in less drastic reductions in SSB during the mid-2010s with more subtle rebuilding since 2017 (Figure 13). However, this does come at the expense of some degraded fits to fishery age composition data (Figure 20). Consequently, the sablefish resource is not projected to rebuild to as high a level as previously thought (i.e., based on the Continuity model, *16.5_Cont*), though it is still expected to recover to well above the biomass-based reference points (Figure 12). As a result, projected ABCs from the *21.10_Proposed* model are significantly lower than the *16.5_Cont* model (Tables 3 and 6). Fortunately, the projected ABCs based on the retrospective analysis from model *21.10_Proposed* appear to align well with implemented ABCs and resultant realized catch, despite the projected ABCs from the Continuity model (*16.5_Cont*) suggesting much higher sustainable catch levels (Table 6). Moreover, projected ABCs appear to be less volatile using the proposed model (*21.10_Proposed*; Table 6), while retrospective patterns and associated retroactive downgrades in recruitment year class strength have been greatly diminished.

Although the proposed model (*21.10_Proposed*) does not (and can never) perfectly describe the dynamics of the sablefish resource and fishery, we believe it provides important tangible improvements over the current SAFE model (*16.5_Cont*). Many updates are consistent with first principles (i.e., biological updates) or statistical and assessment modeling best practices (i.e., freely estimating catchability parameters and using data reweighting approaches), while others appear appropriate given existing

hypotheses regarding sablefish dynamics (e.g., apparent increases in availability and selectivity, which may be due to density-dependent spillover from optimal juvenile habitat or warming water temperatures due to recent marine heatwaves that could be forcing juveniles into deeper, colder slope waters at earlier ages). Testing these hypotheses would require future process studies on fish behavior and gear selectivity. However, it is important to remember that the sablefish assessment assumes a single panmictic population across all management regions in Alaska, while the associated fishery dynamics are assumed to be homogeneous across the same domain. Thus, it can be difficult to relate observed or hypothesized changes at a regional scale to model changes at the Alaska-wide scale. Although it is hypothesized that recent density-dependent or environmental effects might have increased availability of small sablefish to the longline survey, thereby increasing recent selectivity of younger ages, such direct mechanistic explanations are not necessarily required to rationalize changes in model parametrization (i.e., allowing time-variation in catchability or selectivity parameters) that improve model performance and fits to observed data (Wilberg et al., 2009). In the future, continued improvements to the sablefish assessment model will be undertaken, including continued exploration of age- and time-varying natural mortality, alternate parametrizations of fishery selectivity (e.g., incorporating dome-shaped fishery selectivity or time-varying non-parametric approaches), better incorporation of pot gear dynamics into the assessment model, continued refinement to data weighting schemes, and the potential incorporation of the extensive tagging data available for sablefish. The sablefish team continually strives to refine and improve the sablefish assessment model and we envision that the proposed model updates will provide an important step towards continued sustainable management of the resource.

References

Beamish, R.J., and McFarlane, G.A. 1995. A discussion of the importance of aging errors, and an application to walleye pollock: the world's largest fishery. *In*: Recent Developments in Fish Otolith Research (D.H. Secor, J.M.. Dean, and S.E. Campana, *Eds*.). University of South Carolina Press: Columbia, SC. p. 545-565.

Echave, K. 2021. Updated growth analysis for Alaska sablefish. Report for the Joint Plan Teams in September 2021. Anchorage, AK: North Pacific Fishery Management Council.

Francis, R.I.C. 2011. Data weighting in statistical fisheries stock assessment models. Canadian Journal of Fisheries and Aquatic Sciences. 68: 1124-1138. doi:10.1139/F2011-025

Francis, R.I.C. 2017. Revisiting data weighting in fisheries stock assessment models. Fisheries Research. 192: 5-15. Doi: 10.1016/j.fishres.2016.06.006

Goethel, D.R., Hanselman, D.H., Rodgveller, C.J., Fenske, K.H., Shotwell, S.K., Echave, K.B., Malecha, P.W., Siwicke, K.A., and Lunsford, C.R. 2020. Assessment of the sablefish stock in Alaska. *In* "Stock Assessment and Fishery Evaluation Report for the Groundfish Resources of the GOA and BS/AI." Anchorage, AK: North Pacific Fishery Management Council.

Hanselman, D. H., C. Lunsford, J. Fujioka, and C. Rodgveller. 2007. Alaskan Sablefish. In Stock assessment and fishery evaluation report for the groundfish resources of the GOA and BS/AI as projected for 2008. North Pacific Fishery Management Council, 605 W 4th Ave, Suite 306 Anchorage, AK 99501.

Hanselman, D.H., C.J. Rodgveller, K.H. Fenske, S.K. Shotwell, K.B. Echave, P.W. Malecha, and C.R. Lunsford. 2018. Assessment of the sablefish stock in Alaska. In Stock assessment and fishery evaluation report for the groundfish resources of the GOA and BS/AI. North Pacific Fishery Management Council, 605 W 4th Ave, Suite 306 Anchorage, AK 99501. 216p.

Hanselman, D.H., C.J. Rodgveller, K.H. Fenske, S.K. Shotwell, K.B. Echave, P.W. Malecha, and C.R. Lunsford. 2019. Assessment of the sablefish stock in Alaska. In Stock assessment and fishery evaluation report for the groundfish resources of the GOA and BS/AI. North Pacific Fishery Management Council, 605 W 4th Ave, Suite 306 Anchorage, AK 99501. 263p.

Kimura, D. K., and H. H. Zenger. 1997. Standardizing sablefish (*Anoplopoma fimbria*) longline survey abundance indices by modeling the log-ratio of paired comparative fishing CPUEs. ICES J. Mar. Sci. 54:48-59.

Maunder, M.N., Crone, P.R., Punt, A.E., Valero, J.L., and Semmens, B.X. 2017. Data conflict and weighting, likelihood functions and process error. Fisheries Research. 192: 1-4.

NPFMC (North Pacific Fisheries Management Council). 2021. IFQ Sablefish Release Allowance. Anchorage, AK: North Pacific Fishery Management Council. https://meetings.npfmc.org/CommentReview/DownloadFile?p=0a70f8a0-25c2-4d5e-9044-f289c7f4e40d.pdf&fileName=C3%20IFQ%20Sablefish%20Release%20Allowance%20Analysis.pdf

Rodgveller, C. J. 2018. A comparison of methods for classifying sablefish maturity and skip spawning outside the spawning season. Mar. Coast. Fish. 10(6):563-576.

Rodgveller, C. J., J. W. Stark, K. B. Echave, and P.-J. Hulson. 2016. Age at maturity, skipped spawning, and fecundity of female sablefish (Anoplopoma fimbria) during the spawning season. Fish. Bull. 114:89-102.

Sasaki, T. 1985. Studies on the sablefish resources in the North Pacific Ocean. Bulletin 22, (1-108), Far Seas Fishery Laboratory. Shimizu, 424, Japan.

Shotwell, K., Goethel, D.R., Deary, A., Echave, K., Fenske, K., Fissel, B., Hanselman, D., Lunsford, C., Siwicke, K., and Sullivan, J. 2020. Ecosystem and socioeconomic profile of the sablefish stock in Alaska. *In* "Stock Assessment and Fishery Evaluation Report for the Groundfish Resources of the GOA and BS/AI." Anchorage, AK: North Pacific Fishery Management Council.

Trippel, E.A., and Harvey, H.H., 1991. Comparison of methods used to estimate age and length of fishes at sexual maturity using populations of white sucker (*Catostomus commersoni*). Canadian Journal of Fisheries and Aquatic Sciences 48, 1446–1459.

Williams, B., and Rodgveller, C. 2021. Updated maturity-at-age for female sablefish. Report for the Joint Plan Teams in September 2021. Anchorage, AK: North Pacific Fishery Management Council.

Wilberg, M.J., Thorson, J.T., Linton, B.C., and Berkson, J. 2009. Incorporating time-varying catchability into population dynamic stock assessment models. Reviews in Fisheries Science. 18(1): 7-24.

Tables

Table 1. Description of model runs with associated abbreviations.

Model Group	Scenario Name	Abbreviation	Description
Continuity	16.5. Continuity	*16.5_Cont*	The 2020 SAFE final model, which utilizes priors on catchability, fixed data weights, and no recent time blocks in the fishery or survey for catchability and selectivity parameter estimation.
Update Biology	21.1. Update Weight and Growth	*21.1_Wt+Grt*	The continuity model with updated weight and growth parameters based on the full complement of longline survey data from 1996-2019 (as described in Echave, 2021).
	21.2. Update Age-Based Maturity No Skipped Spawning	*21.2_Mat_Age_GLM_No_SS*	The continuity model with maturity updated using the age-based general linear model (GLM) and ignoring information on skipped spawning (i.e., strict update of maturity based on histological data only) from Williams and Rodgveller (2021).
	21.3. Update Age-Based Maturity	*21.3_Mat_Age_GAM*	The continuity model with maturity updated using the age-based general additive model (GAM) from Williams and Rodgveller (2021).
	21.4. Update Age-Length Maturity	*21.4_Mat_AL_GAM*	The continuity model with maturity updated using the age-length based general additive model (GAM) from Williams and Rodgveller (2021).
	21.5. Update Weight, Growth, and Age-Length Maturity	*21.5_Upd_Bio_AL-Mat*	The continuity model with weight and growth updated based on Echave (2021) and maturity updated using the age-length based general additive model (GAM) from Williams and Rodgveller (2021).
Model Parametrization	21.6. Remove Catchability Priors	*21.6_No_q_Prior*	The continuity model with all priors on catchability coefficients removed.
	21.7. Incorporate a Recent (post-2016) Time Block for Fishery and Survey Selectivity and Fishery CPUE Catchability Estimation	*21.7_Add_Sel+q_Block*	The continuity model with a recent time block (2016 - present) added to the longline fishery and longline survey for the estimation of selectivity parameters along with an associated fishery CPUE catchability parameter.
	21.8. Remove Catchability Priors and Add 2016 Selectivity and Fishery CPUE Catchability Time Block	*21.8_No_q_Add_Sel+q_Block*	The continuity model with all priors on catchability coefficients removed and a recent time block (2016 - present) added to the longline fishery and longline survey for the estimation of selectivity parameters and fishery CPUE catchability.
Data Weighting	21.9. Continuity with Francis Reweighting	*21.9_Cont_Francis*	The continuity model with data weights updated using the Francis (2011, 2016) reweighting method.
	21.10. Proposed Model	***21.10_Proposed***	The final proposed model where weight and growth are updated based on Echave (2021), maturity is updated using the age-length based general additive model (GAM) from Williams and Rodgveller (2021), catchability priors are removed, a recent time block (2016 - present) is added to the longline fishery and longline survey for the estimation selectivity parameters and fishery CPUE catchability, and data weights are updated using the Francis (2011, 2016) reweighting method.

Table 2. The maximum gradient component (*Max Grad*), total negative log-likelihood (*nLL*), and number of parameters *(# Pars)* for each model run. Note that all models aside from the Continuity (*16.5_Cont*) and Francis reweighted Continuity (*21.9_Cont_Francis*) models include a minor update to the trawl fishery selectivity parameterization, which reduced the number of estimated parameters by one compared to the Continuity model (see Section 2.1).

Model	Converged?	Max Grad	nLL	# Pars
16.5_Cont	TRUE	0.000301006496384589	1888.1	240
21.1_Wt+Grt	TRUE	6.28388663045004e-05	1845.4	239
21.2_Mat_Age_GLM_No_SS	TRUE	8.94369828047811e-05	1889.7	239
21.3_Mat_Age_GAM	TRUE	6.91647880822961e-05	1889.7	239
21.4_Mat_AL_GAM	TRUE	8.63499946136482e-05	1889.7	239
21.5_Upd_Bio_AL-Mat	TRUE	5.75877264453965e-05	1845.4	239
21.6_No_q_Prior	TRUE	0.000174866710556766	1907.33	239
21.7_Add_Sel+q_Block	TRUE	0.000268053382307611	1840.25	244
21.8_No_q_Add_Sel+q_Block	TRUE	9.24286507245969e-05	1841.4	244
21.9_Cont_Francis	TRUE	2.6246219892387e-05	724.48	240
21.10_Proposed	TRUE	0.00151063289005471	776.66	244

Table 3. Estimated terminal year (2020) parameters (i.e., fishing mortality, *F*, and spawning stock biomass, *SSB*), associated biological reference points and stock status determinations relative to a target SSB representing 40% (*SSB_40*) depletion from unfished SSB (*SSB$_0$*), and resultant Acceptable Biological Catch (*ABC*) based on the NPFMC B$_{40}$ HCR. For models with time-varying biology or selectivity, reference points and associated calculations utilize the most recent time block of values.

Model	2020 SSB (kt)	SSB_40 (kt)	2020 SSB/SSB_40	2020 F	F_40	2020 F/F_40	F_ABC	2021 ABC (kt)
16.5_Cont	94.43	126.84	0.74	0.05	0.1	0.5	0.1	52.41
21.1_Wt+Grt	99.1	135.16	0.73	0.04	0.09	0.44	0.09	44.88
21.2_Mat_Age_GLM_No_SS	87.17	124.22	0.7	0.05	0.09	0.56	0.09	44.23
21.3_Mat_Age_GAM	79.99	117.98	0.68	0.05	0.09	0.56	0.09	44.94
21.4_Mat_AL_GAM	90.72	127.17	0.71	0.05	0.09	0.56	0.09	48.35
21.5_Upd_Bio_AL-Mat	85.31	130.76	0.65	0.04	0.08	0.5	0.08	39.75
21.6_No_q_Prior	88.86	126.44	0.7	0.05	0.09	0.56	0.09	47.32
21.7_Add_Sel+q_Block	80.81	117.4	0.69	0.05	0.09	0.56	0.08	34.61
21.8_No_q_Add_Sel+q_Block	74.05	115.28	0.64	0.06	0.09	0.67	0.08	29.1
21.9_Cont_Francis	101.42	112.57	0.9	0.05	0.11	0.45	0.11	51.25
21.10_Proposed	85	114.19	0.74	0.06	0.08	0.75	0.08	27.09

Table 4. Comparison of catchability coefficient (*q*) estimates for each abundance or biomass index for the continuity (*16.5_Cont*) model, the model with catchability priors removed (*21.6_No_q_Prior*), and the new proposed model (*21.10_Proposed*). Note that the continuity (*16.5_Cont*) model uses priors on catchability parameters whereas the proposed model (*21.10_Proposed*) does not. Similarly, the proposed model (*21.10_Proposed*) has one additional catchability coefficient for the recent (post-2016) time block for fishery CPUE.

		Coop LL Survey	LL Survey	LL Survey Post-2016	Trawl Survey	LL Fishery CPUE Pre-1995 (Derby)	LL Fishery CPUE Post-1995 (IFQ)	LL Fishery CPUE Post-2016 (IFQ Recent)	JPN LL Fishery CPUE
	16.5_Continuity	5.96	7.96	Not Estimated	1.33	3.98	5.93	Not Estimated	6.55
Model	21.6_No_q_Prior	6.22	8.35	Not Estimated	1.39	4.20	6.26	Not Estimated	6.38
	21.10_Proposed	5.36	7.73	Not Estimated	1.07	3.83	6.81	3.45	8.02

(Column header group: **Index**)

Table 5. Comparison of input data weight for the continuity (*16.5_Cont*) model utilizing fixed data weights, the Francis reweighted continuity model (*21.9_Cont_Francis*), and the new proposed model (*21.10_Proposed*) that also utilizes Francis reweighting. Note that catch and index data weights are held constant throughout the Francis reweighting procedure. Additionally, indices have yearly input standard errors, while compositional data have yearly input effective samples sizes; neither of which are altered during the reweighting procedure.

Data Source	Model		
	16.5_Cont	*21.9_Cont_Francis*	*21.10_Proposed*
Fixed Gear Catch	50.000	50.000	50.000
Trawl Catch	50.000	50.000	50.000
Longline Survey RPN	0.448	0.448	0.448
Coop Survey RPN	0.448	0.448	0.448
Fixed Gear Fishery CPUE	0.448	0.448	0.448
Japan Longline Fishery CPUE	0.448	0.448	0.448
Trawl Survey RPW	0.448	0.448	0.448
Fixed Gear Age Composition	7.800	0.817	0.710
Longline Survey Age Composition	7.950	2.297	3.904
Coop Longline Survey Age Composition	1.000	1.123	1.167
Fixed Gear Fishery Length Composition Males	1.000	3.948	5.915
Fixed Gear Fishery Length Composition Females	1.000	4.423	6.223
Trawl Fishery Size Composition Males	4.100	0.324	0.327
Trawl Fishery Size Composition Females	4.100	0.523	0.396
Longline Survey Size Composition Males	1.000	0.904	1.772
Longline Survey Size Composition Females	1.000	0.986	1.885
Coop Survey Size Composition Males	1.000	1.229	1.182
Coop Survey Size Composition Females	1.000	1.923	1.960
Trawl Survey Size Composition Males	7.250	0.954	0.738
Trawl Survey Size Composition Females	7.250	1.274	0.719

Table 6. Comparison of observed catch, enacted ABC, and model projected ABC from retrospective runs of the Continuity model *(16.5_Cont)* and the new proposed model (*21.10_Proposed*). Projected ABCs are from retrospective peels representing a terminal data year equal to the year column minus one. Note that the projected ABCs may differ from the SAFE recommended ABCs as the reported values are based on the current retrospective runs where data inputs may differ from that used in the final SAFE models. All values are in metric tons (mt). The 2021 reported catch is the value as of September 1, 2021 as reported on AKFIN.

			Model	
Year	Catch (mt)	ABC (mt)	*16.5_Cont*	*21.10_Proposed*
2011	12,978	16,040	14,600	12,750
2012	13,869	17,240	14,400	13,464
2013	13,645	16,230	14,000	13,122
2014	11,588	13,722	12,100	12,042
2015	10,973	13,657	12,700	12,989
2016	10,257	11,795	11,300	11,476
2017	12,270	13,083	11,900	12,241
2018	14,341	14,957	25,700	16,829
2019	16,624	15,068	27,300	12,755
2020	19,006	22,009	43,600	19,914
2021	*13,112*	29,588	52,400	27,086

Figures
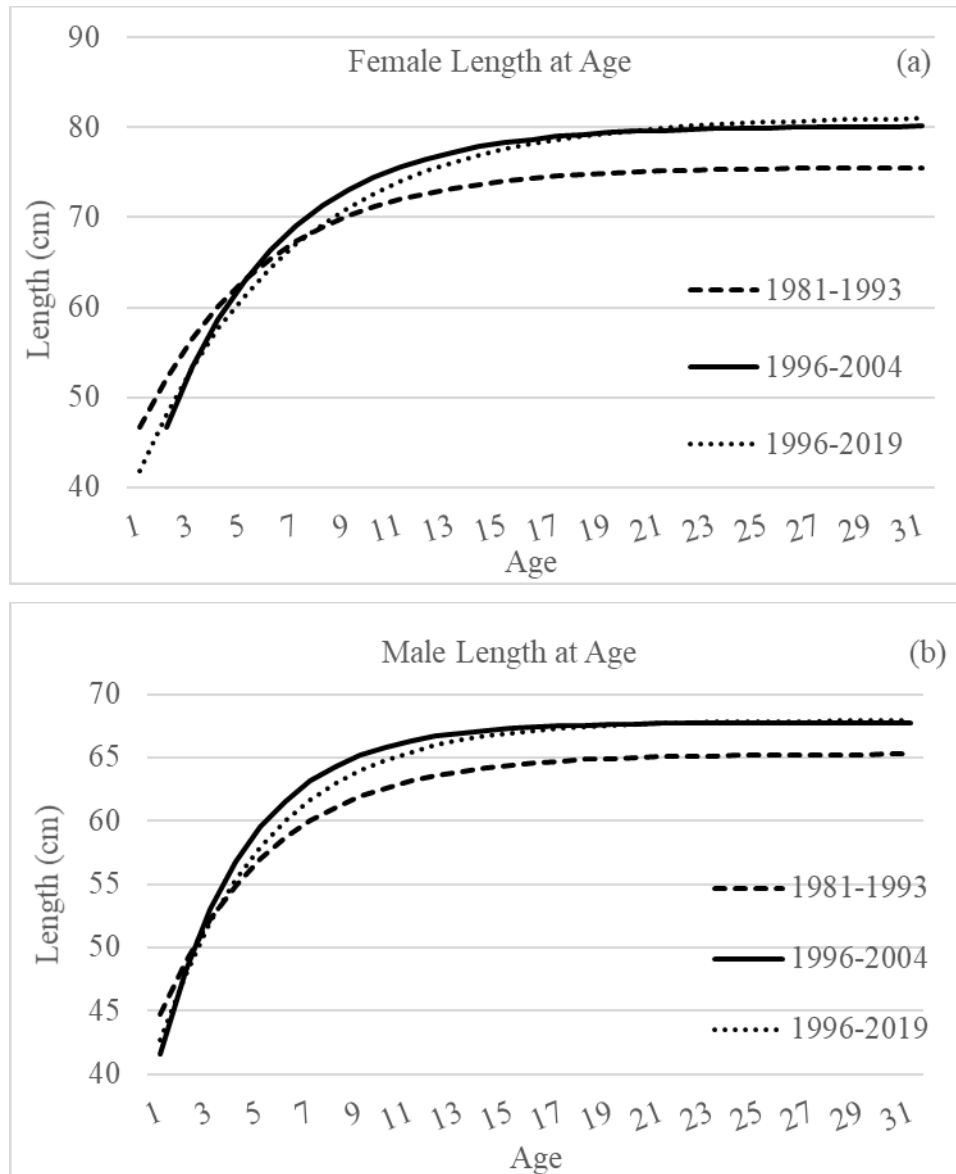
Figure 1. Growth curves used in the continuity (*16.5_Cont*) model (1981 -1993 and 1996 – 2004) and the new proposed model (*21.10_Proposed*; 1981-1993 and 1996 – 2019), as described in Echave (2021). The top panel illustrates the growth curve for females and the bottom panels shows the growth curve for males.
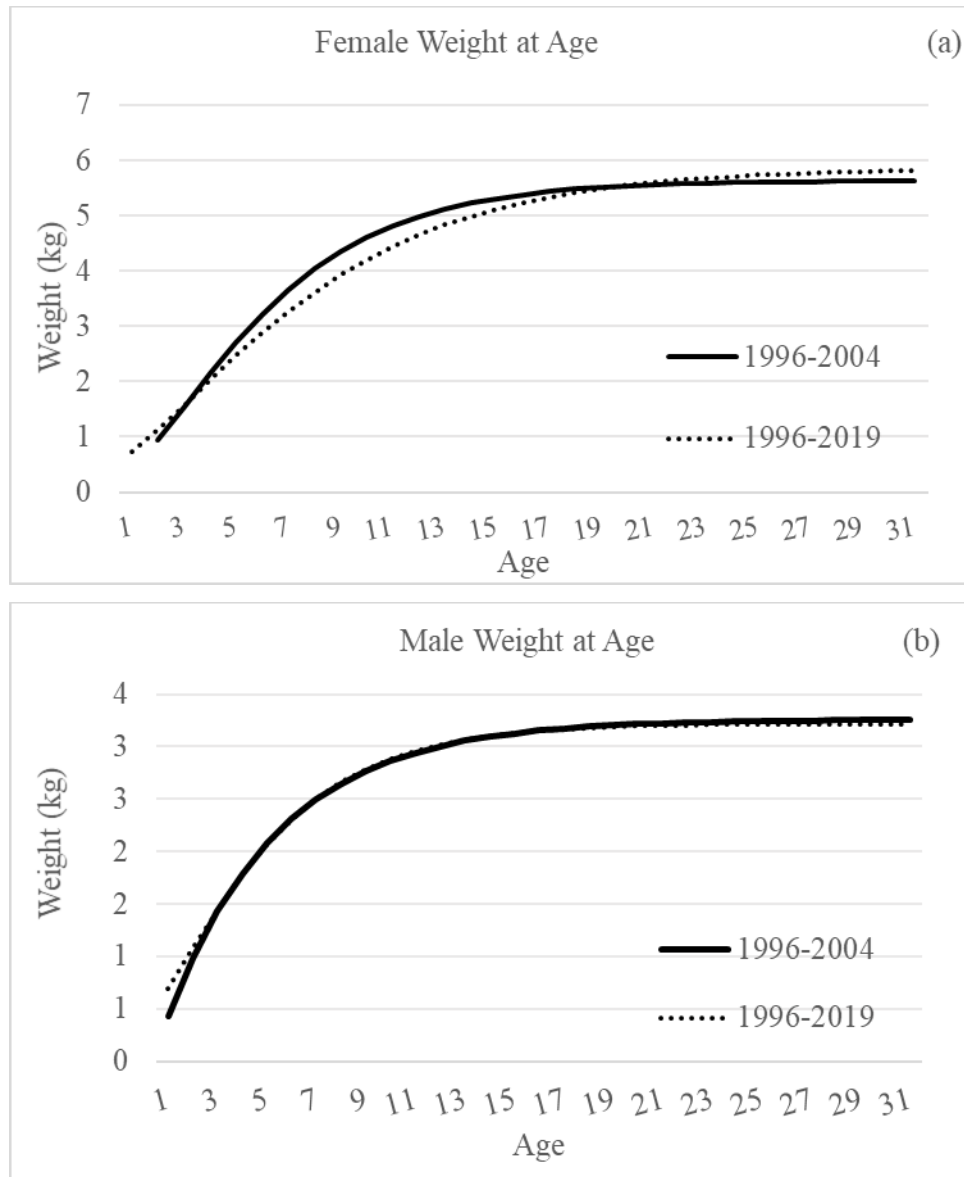
Figure 2. Weight-at-age used in the continuity (*16.5_Cont*) model (1996 – 2004) and the new proposed model (*21.10_Proposed*; 1996 – 2019), as described in Echave (2021). The top panel illustrates the weight-at-age curve for females and the bottom panels shows the weight-at-age curve for males.
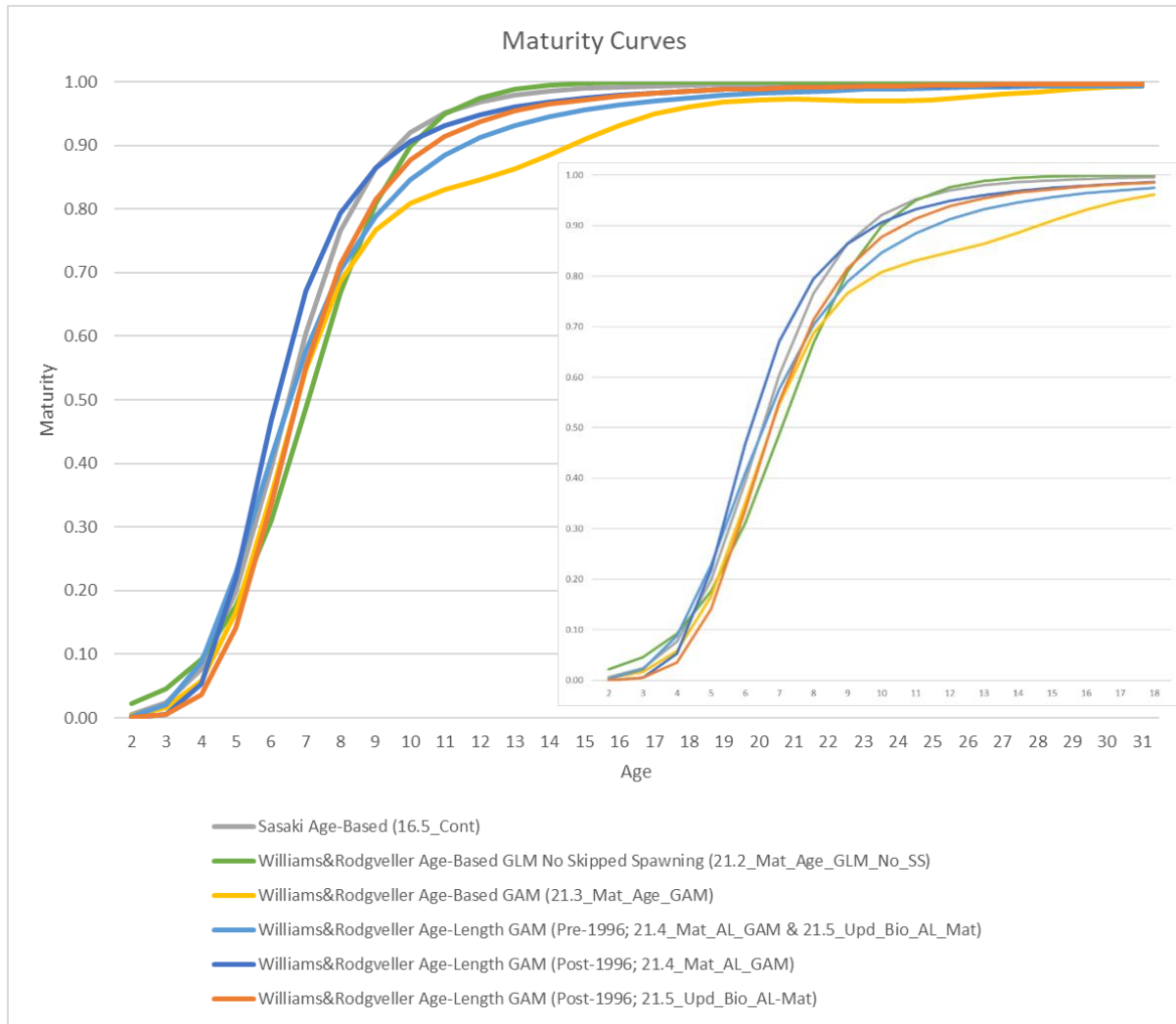
Figure 3. Current and updated maturity-at-age curves as described in Williams and Rodgveller (2021). The grey line illustrates the Continuity (*16.5_Cont*) model using the results of Sasaki (1985), while the green line is the updated age-based maturity curve based on histological data, a generalized linear model (GLM), and not including skipped spawning information (used in the *21.2_Mat_Age_GLM_No_SS* model) from Williams and Rodgveller (2021). The yellow line is the updated age-based maturity using a General Additive Model (GAM) and including skipped spawning information (used in the *21.3_Mat_Age_GAM* model) from Williams and Rodgveller (2021). The blue and orange lines show the updated age-length based maturity curve using a GAM from Williams and Rodgveller (2021). The light blue line illustrates the pre-1996 maturity-at-age, which is used by all age-length based maturity models (given that growth prior to 1996 is constant across all model implementations). The dark blue line is the post-1996 maturity-at-age used in model *21.4_Mat_AL_GAM* and based on the growth parameters used in the Continuity model (*16.5_Cont*). The orange line is the post-1996 maturity-at-age which used in the *21.5_Upd_Bio_AL-Mat* model, which uses the updated growth parameters. The light blue and orange age-length maturity curves are used in the new proposed model (*21.10_Proposed*). The change in maturity over time in the age-length GAM is due to the different growth stanzas (i.e., the updated growth curves developed by Echave, 2021, and illustrated in Figure 1), despite constant maturity parameters being assumed for all model years (i.e., the age-length maturity model calculates new maturity values when growth changes). Inset provides zoomed in view of most dynamic ages.
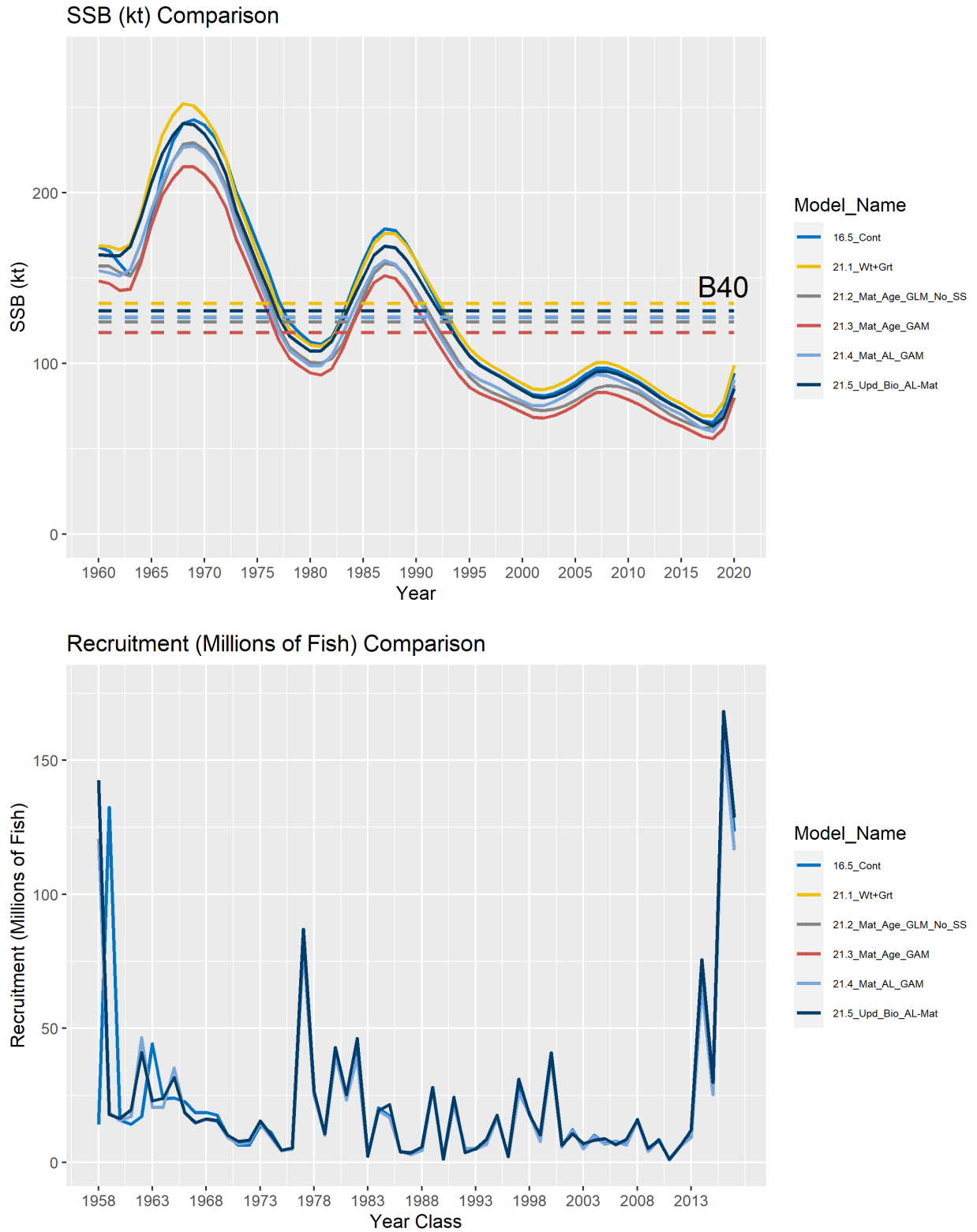
Figure 4. Model comparisons for spawning stock biomass (top panel) and recruitment (bottom panel) within the 'Biology Update' grouping.

SSB (kt) Comparison

Recruitment (Millions of Fish) Comparison
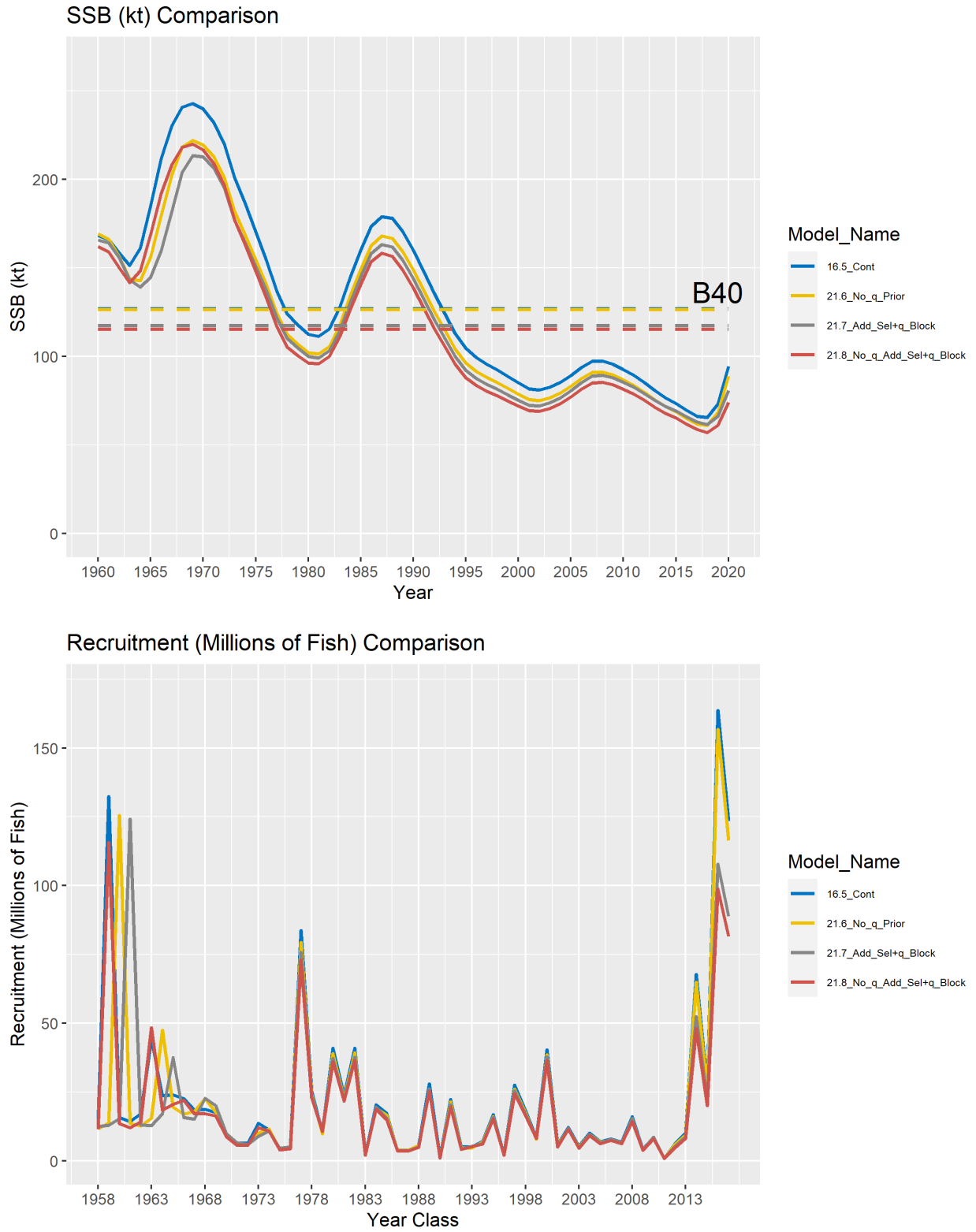
Figure 5. Model comparisons for spawning stock biomass (top panel) and recruitment (bottom panel) within the 'Model Parametrization Update' grouping.
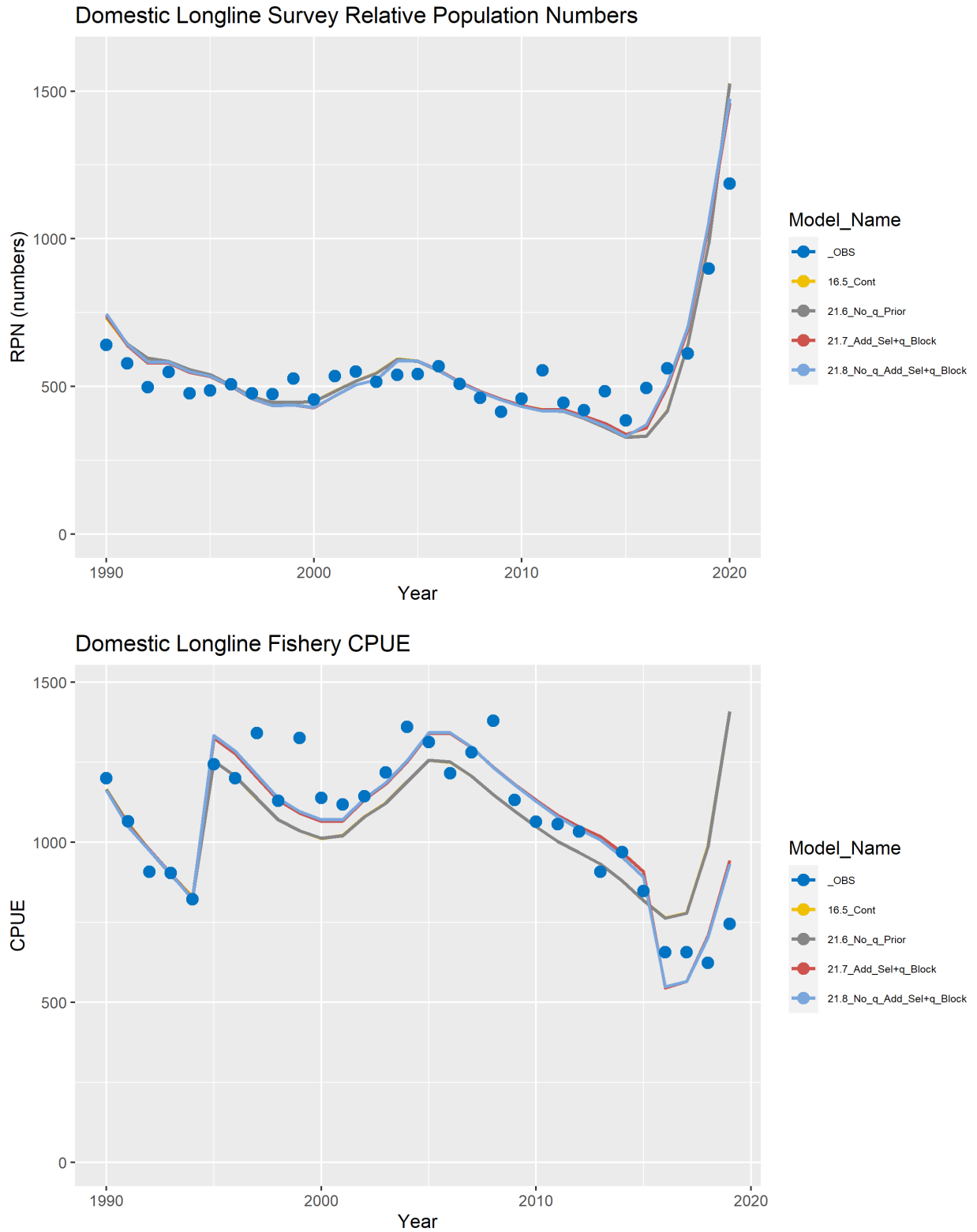
Figure 6. Model comparisons demonstrating fit to the domestic longline survey relative population numbers (RPN) index (top panel) and domestic longline fishery catch-per-unit effort (CPUE) index (bottom panel) within the 'Model Parametrization Update' grouping.
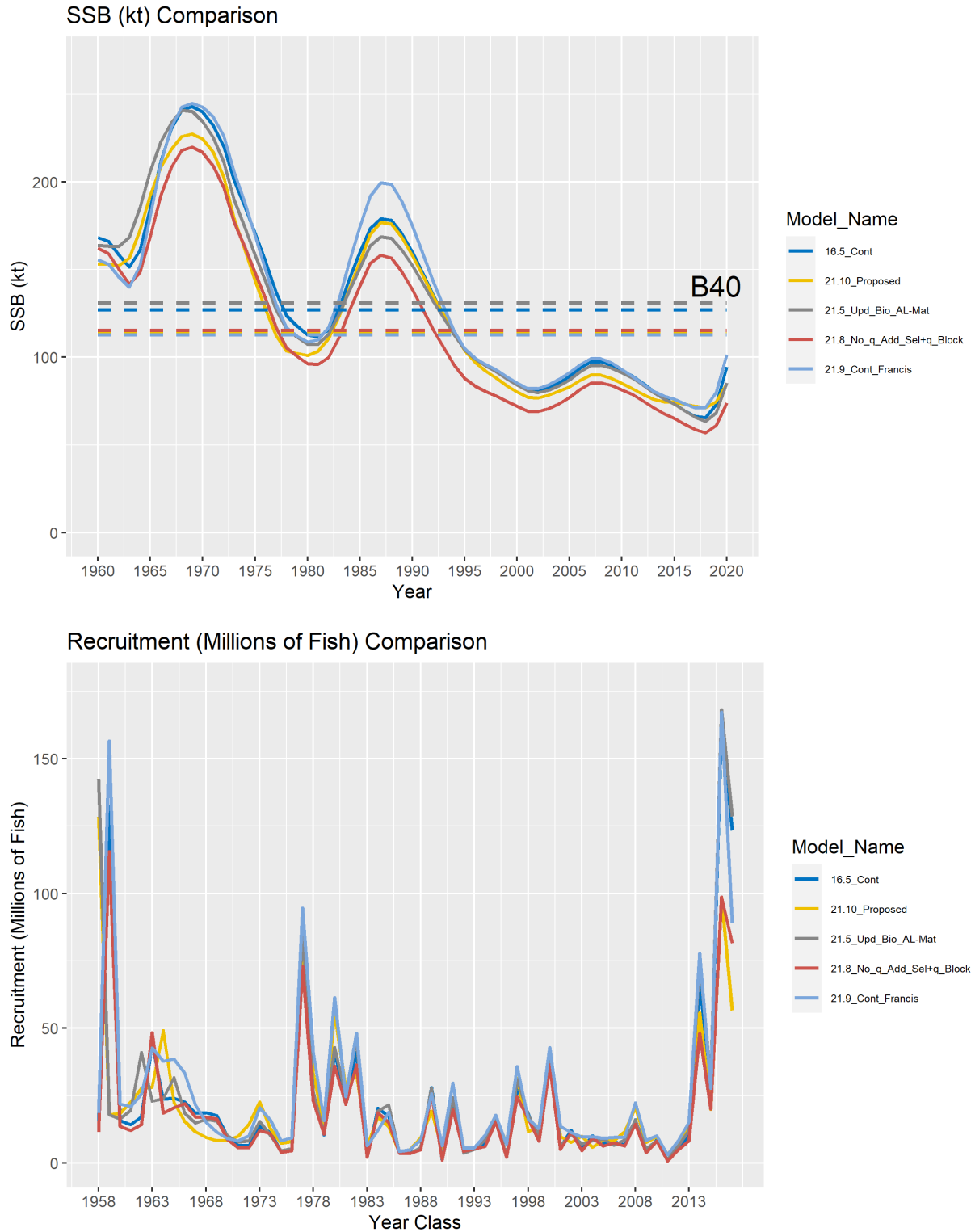
Figure 7. Model comparisons for spawning stock biomass (top panel) and recruitment (bottom panel) within the 'Data Weighting' grouping, including the new proposed model (*21.10_Proposed*), the full 'Biology Update' model (*21.5_Upd_Bio_AL-Mat*), and the 'Model Parametrization Update' model (*21.8_No_q_Add_Sel+q_Block*).
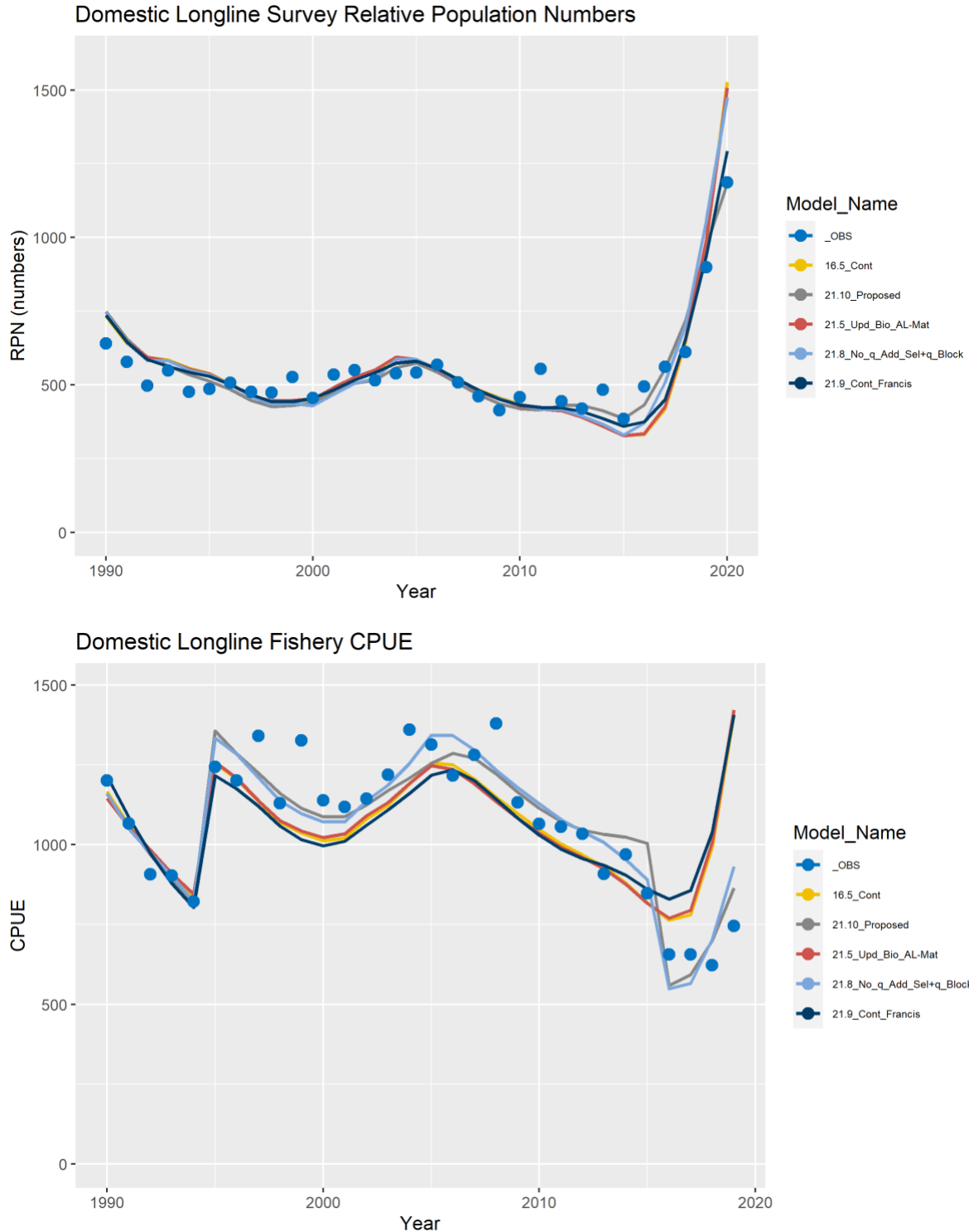
Figure 8. Model comparisons demonstrating fit to the domestic longline survey relative population numbers (RPN) index (top panel) and domestic longline fishery catch-per-unit effort (CPUE) index (bottom panel) within the 'Data Weighting' grouping, including the new proposed model (*21.10_Proposed*), the full 'Biology Update' model (*21.5_Upd_Bio_AL-Mat*), and the 'Model Parametrization Update' model (*21.8_No_q_Add_Sel+q_Block*).
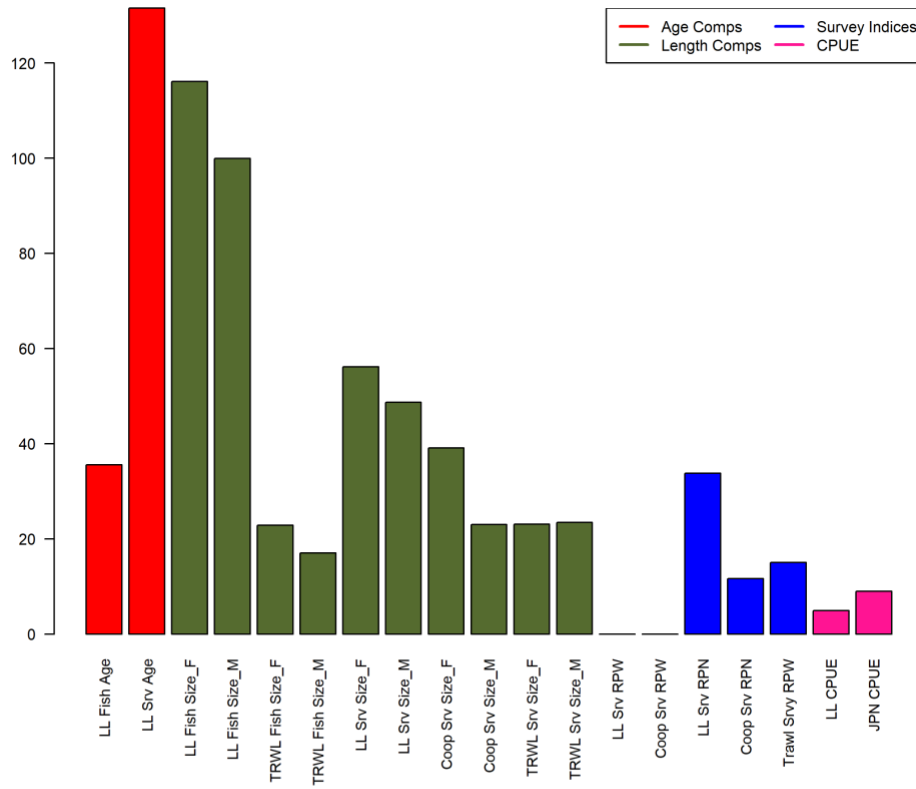
Figure 9. Component contributions to the total negative log-likelihood for each data source fit in the model. Results for the Continuity model (*16.5_Cont*) are in the top left panel, the Francis Reweighted Continuity model (*21.9_Cont_Francis*) are in the top right panel, and the new proposed model (*21.10_Proposed*) are in the bottom panel. Note differences in y-axis scale across panels.
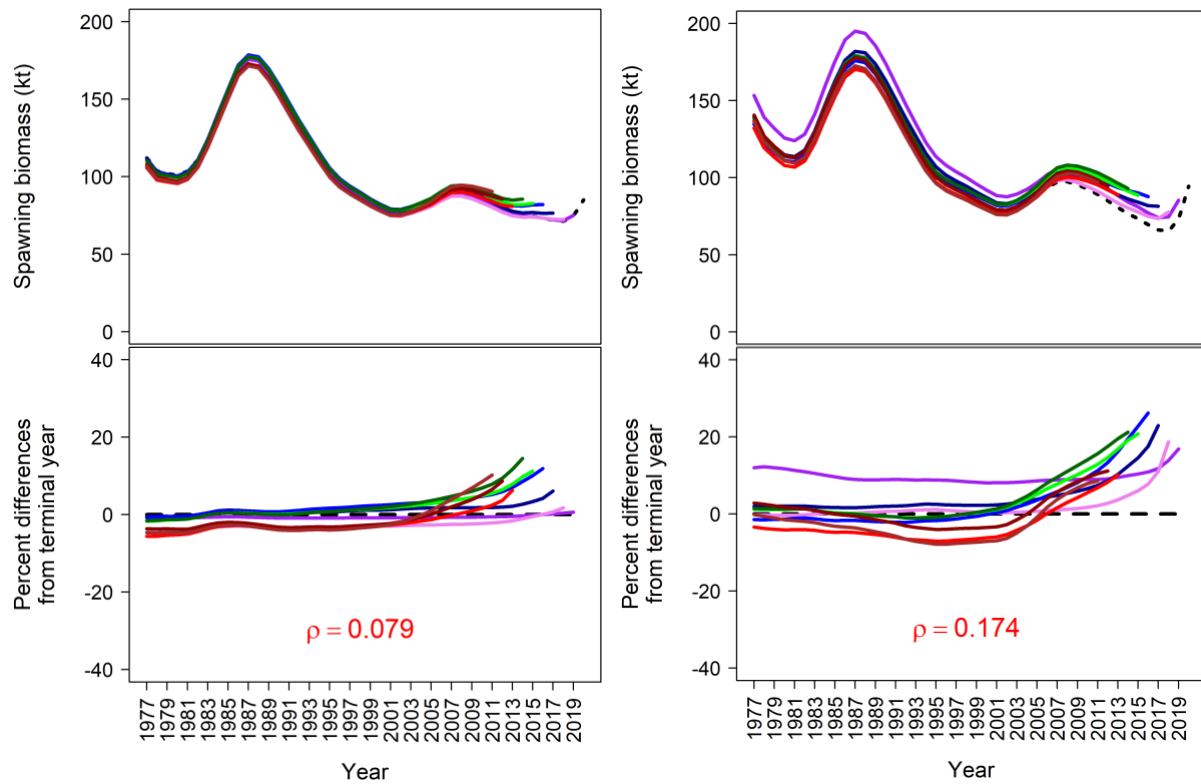
Figure 10. Results of a retrospective analysis for spawning stock biomass for the new proposed model (*21.10_Proposed*; left panel) and the Continuity model (*16.5_Cont*; right panel). Mohn's rho (*ρ*) is provided in red. Note that the proposed model (*21.10_Proposed*) retrospective analysis has a model change starting with the 2018 retrospective year (i.e., estimation of new longline survey and fishery catchability and selectivity parameters for the post-2016 time block). Thus, comparison of models for retrospective years before and after 2018 (i.e., starting with the dark blue line) is problematic.
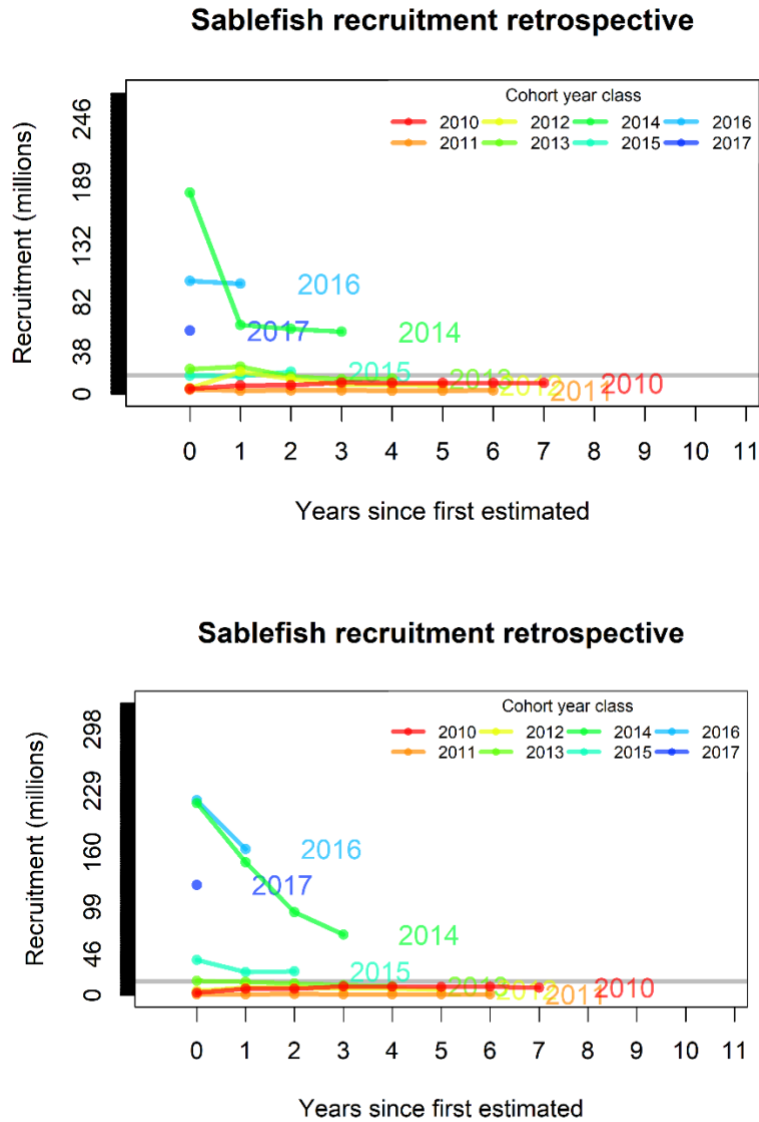
Figure 11. Squid plot demonstrating the refinement of age-2 recruitment estimates as new data years are added to the model based on the results of a retrospective analysis for the new proposed model (*21.10_Proposed*; top panel) and the Continuity model (*16.5_Cont*; bottom panel). Note that the proposed model (*21.10_Proposed*) retrospective analysis has a model change starting with the 2018 retrospective year (i.e., estimation of new longline survey and fishery catchability and selectivity parameters for the post-2016 time block). Thus, comparison of models for retrospective years before and after 2018 is problematic. The transition between model parametrizations is clearly visible in the estimation of the 2014 year class (i.e., the first estimate is based on the 2017 model without a separate catchability and selectivity time blocks for the recent, post-2016, period, while subsequent estimates are based on models with a recent time block).
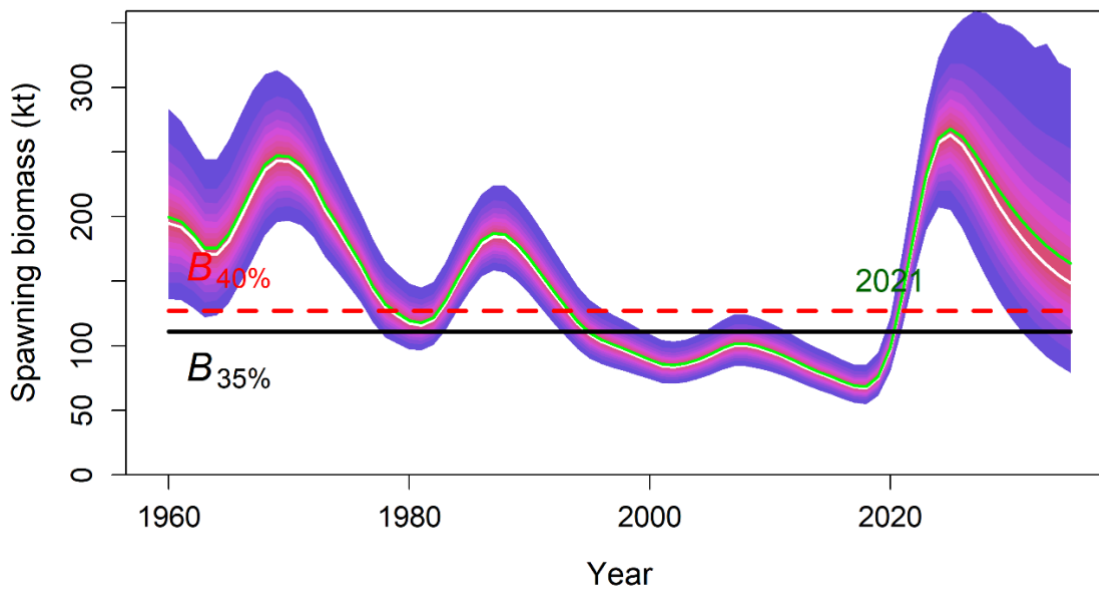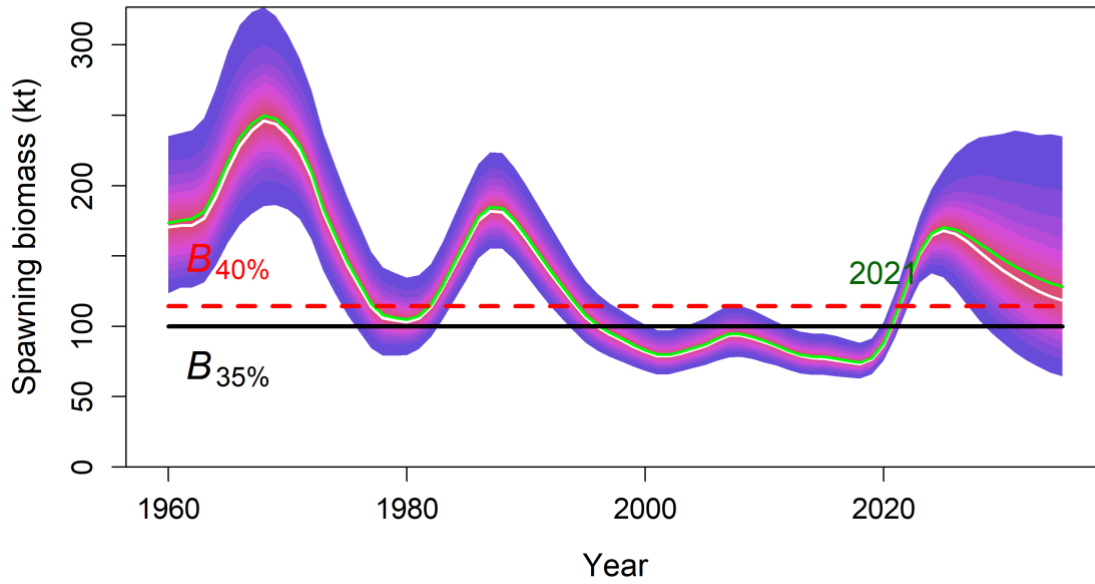
Figure 12. Results of a Markov Chain Monte Carlo (MCMC) analysis demonstrating estimates of female spawning biomass (kilotons) and their uncertainty from MCMC runs. The results for the new proposed model (*21.10_Proposed*) are in the top panel and those for the Continuity model (*16.5_Cont*) are in the bottom panel. White line is the median and green line is the mean, while shaded fill is 5% increments of the posterior probability distribution of spawning biomass based on MCMC simulations. Width of shaded area is the 95% credibility interval.
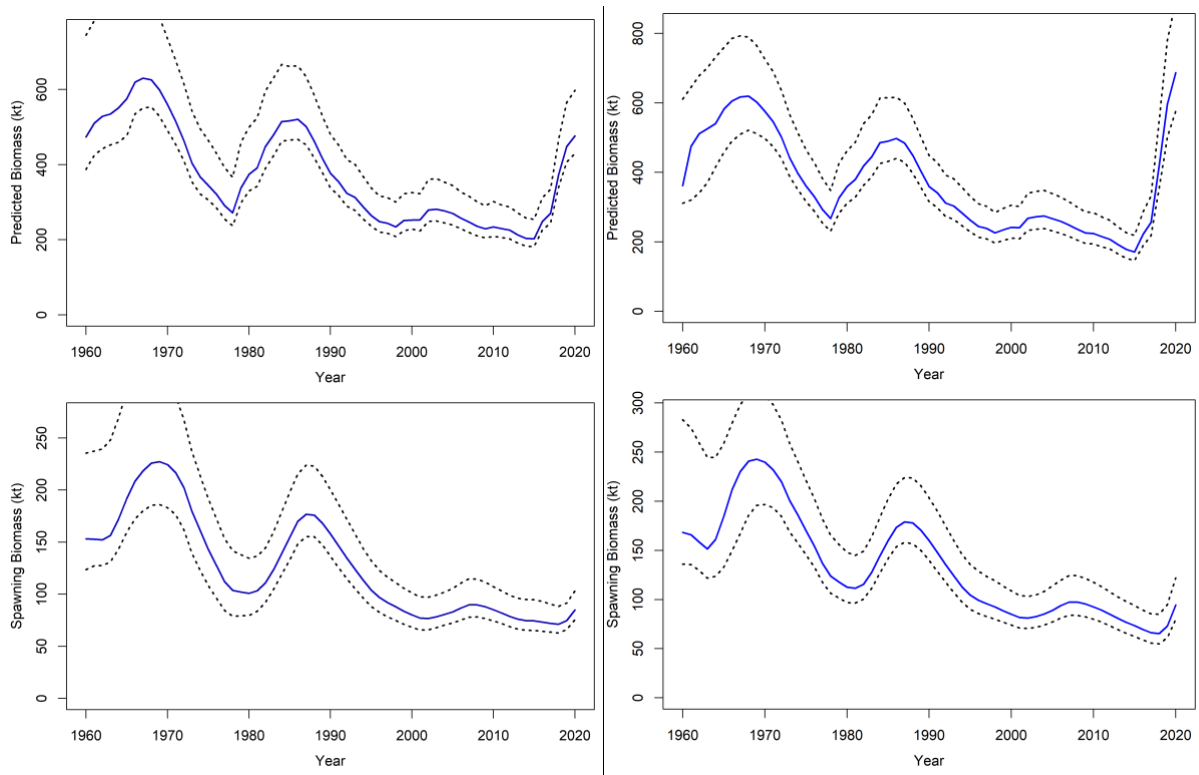
Figure 13. Estimated sablefish total biomass (top panel) and spawning biomass (bottom panel) with 95% MCMC credible intervals. The results for the new proposed model (*21.10_Proposed*) are in the left panel and those for the Continuity model (*16.5_Cont*) are in the right panel. Values are in kilotons.
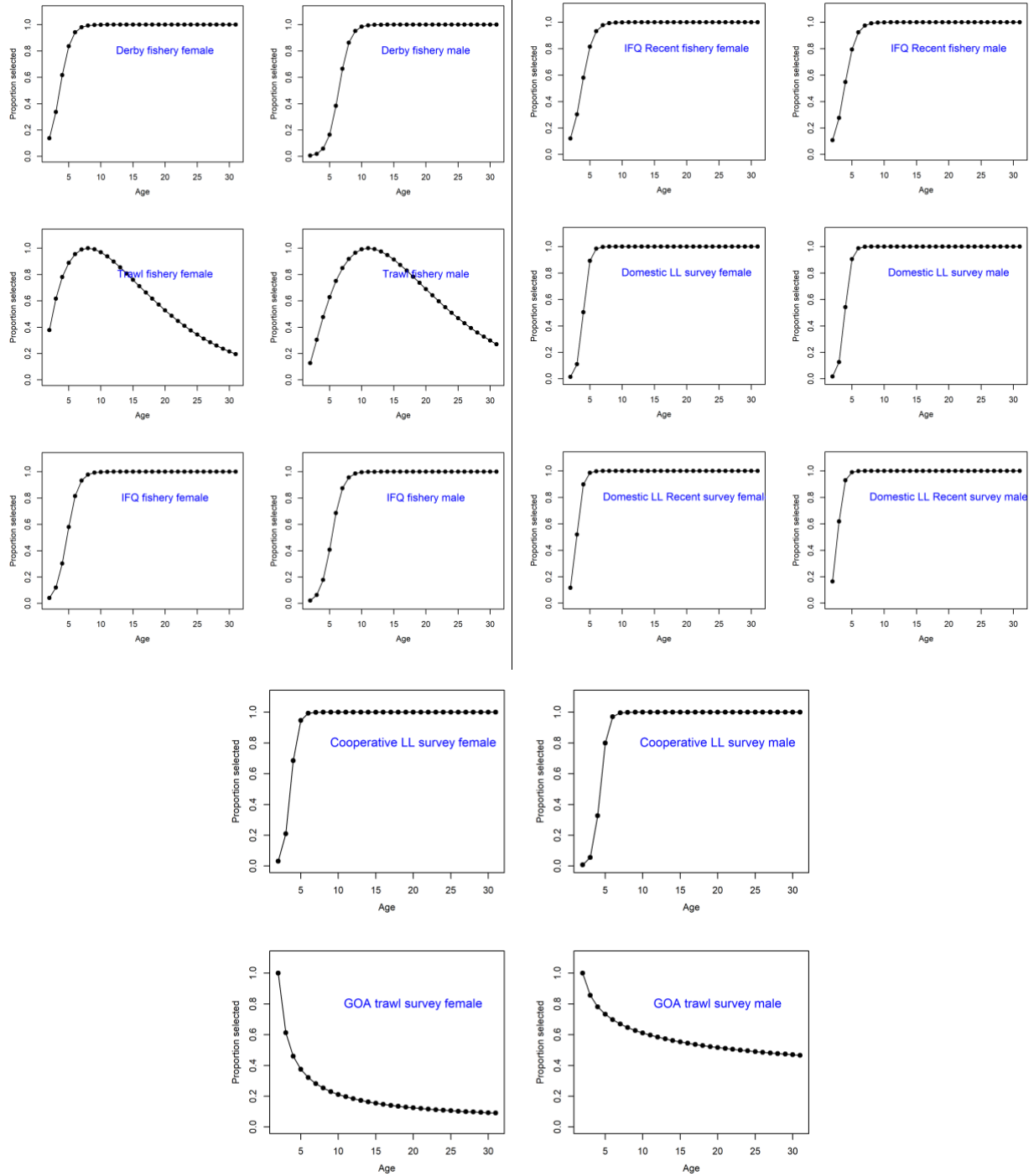
Figure 14. Estimated fishery and survey selectivity for the new proposed model (*21.10_Proposed*). The derby longline fishery occurred until 1994, then the fishery switched to an IFQ system in 1995. The 'Recent' in the longline fishery and survey selectivity names represents the recent time block implemented in 2016.
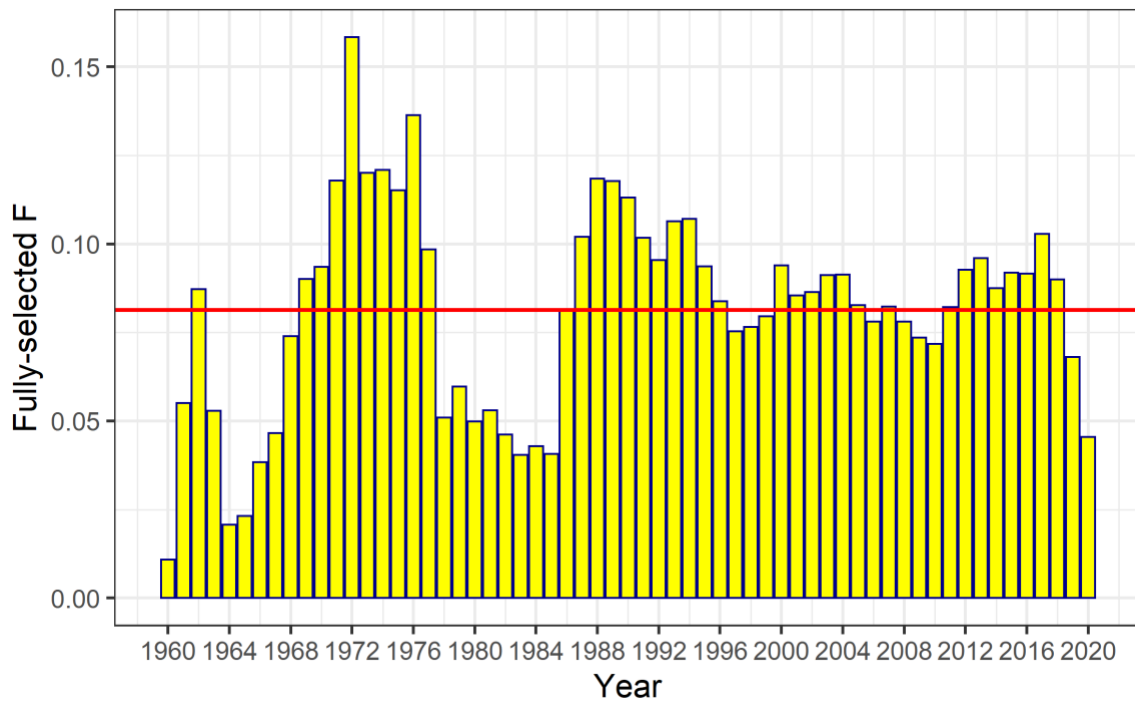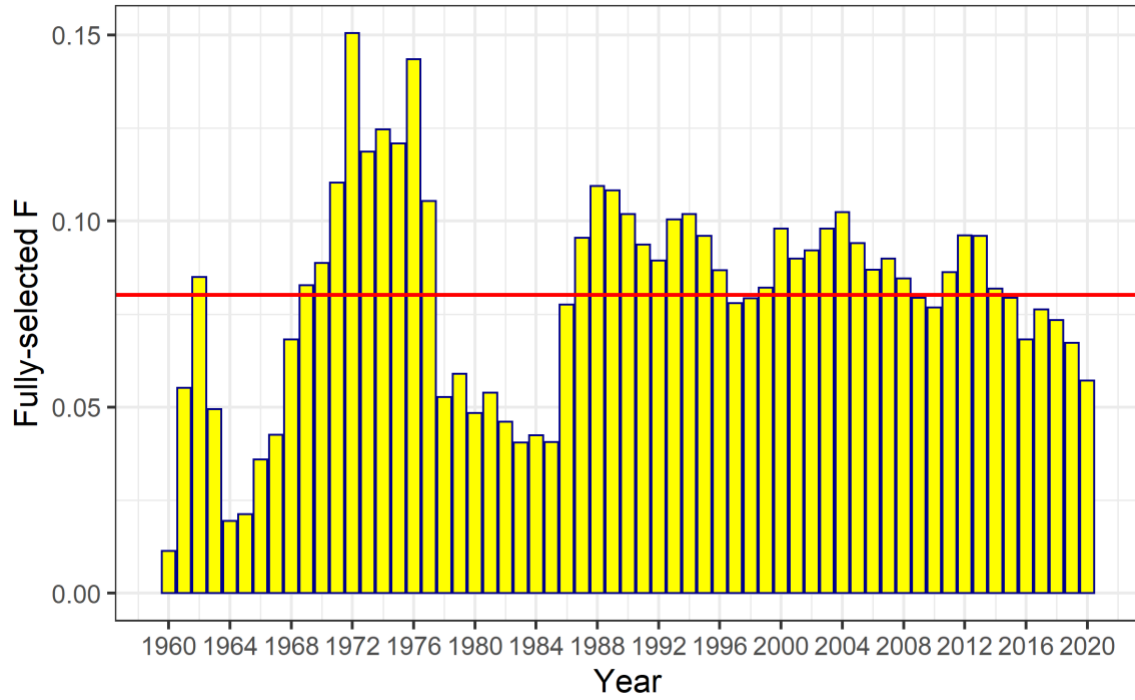
Figure 15. Time series of combined fully-selected fishing mortality for fixed and trawl gear for sablefish. The results for the new proposed model (*21.10_Proposed*) are in the top panel and those for the Continuity model (*16.5_Cont*) are in the bottom panel. Red line is the mean fishing mortality for the entire time series.
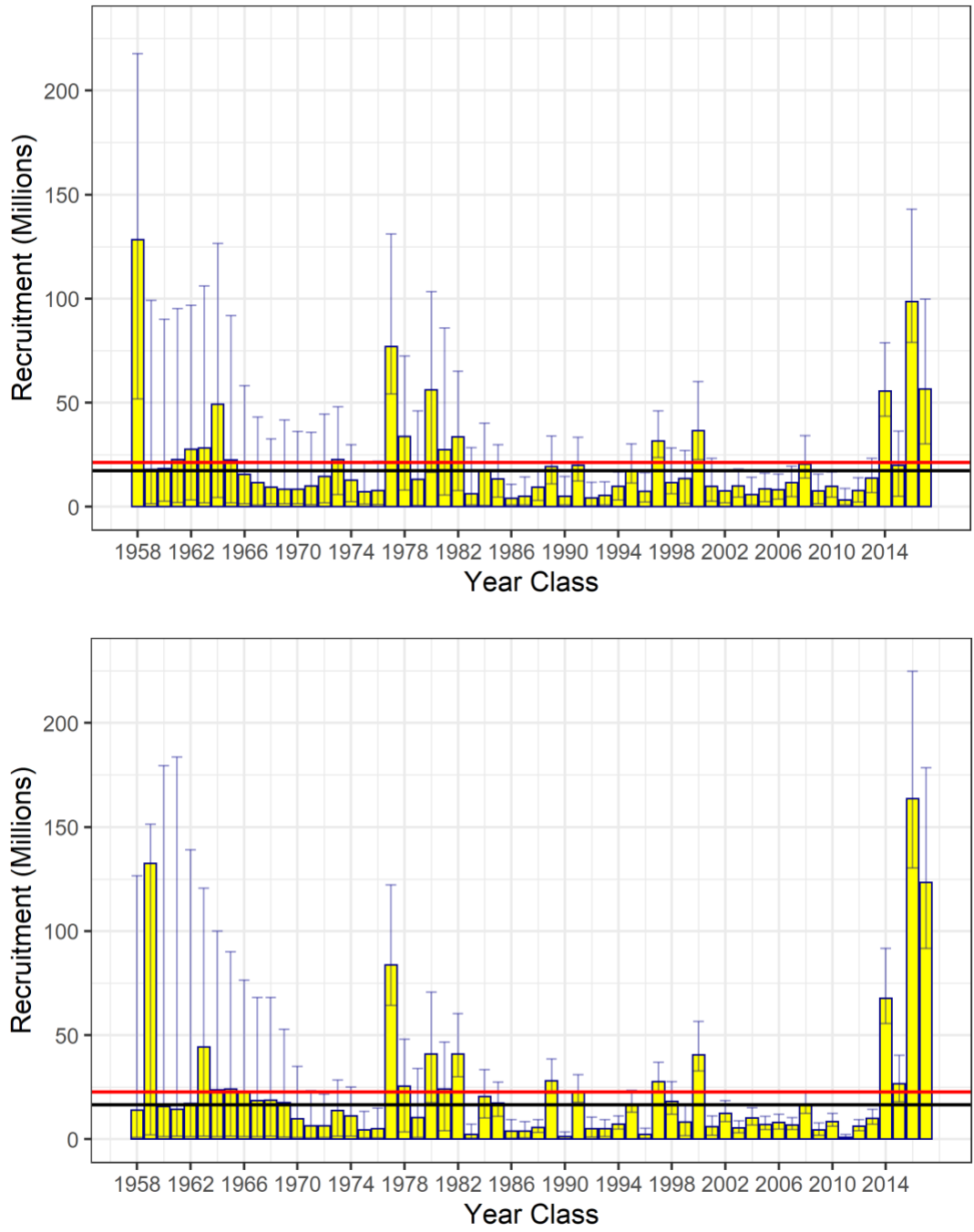
Figure 16. Estimated recruitment of age-2 sablefish (millions of fish) with 95% credible intervals from MCMC by year class (recruitment year minus two). The results for the new proposed model (*21.10_Proposed*) are in the top panel and those for the Continuity model (*16.5_Cont*) are in the bottom panel. Red line is overall mean, while black line is mean for recruitments from year classes between 1977 and 2017. Credible intervals are based on MCMC posteriors. The estimate for the 2018 year class (terminal year 2020 recruitment event) is omitted, because it is fixed to the estimated mean recruitment value (μ_r) with no deviation parameter estimated.
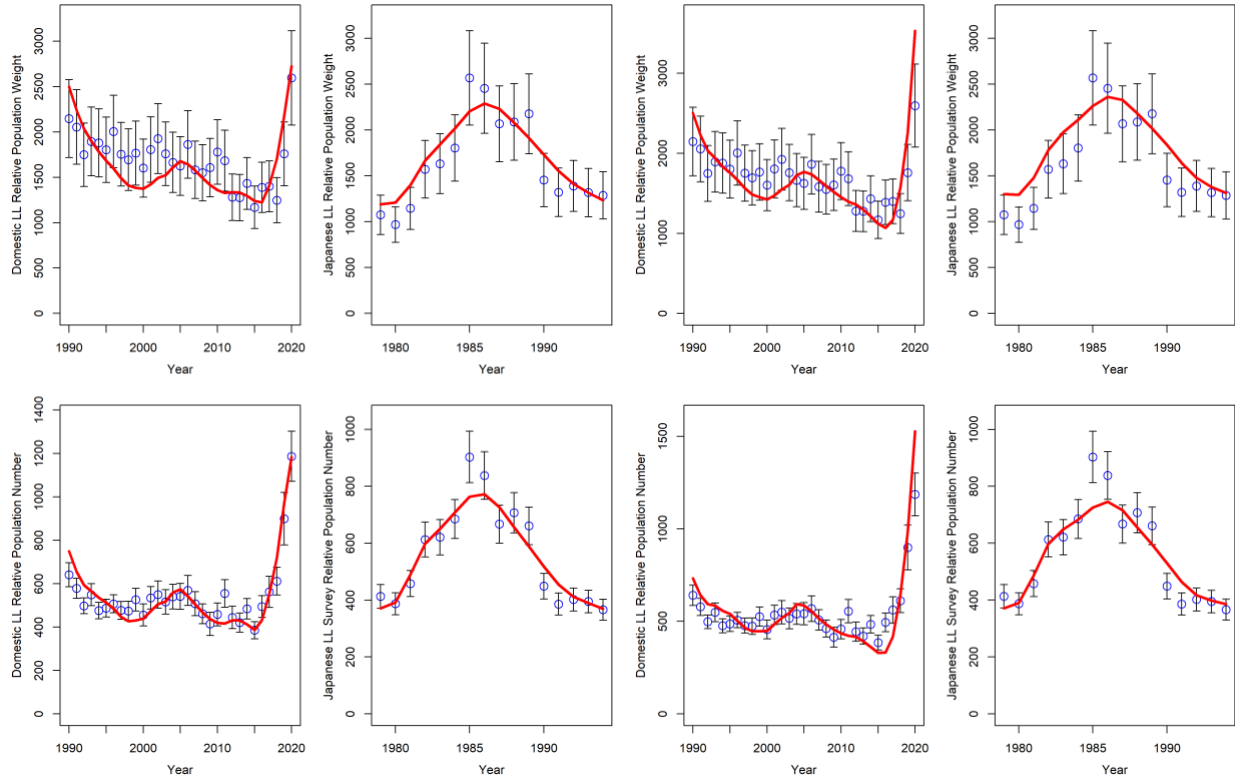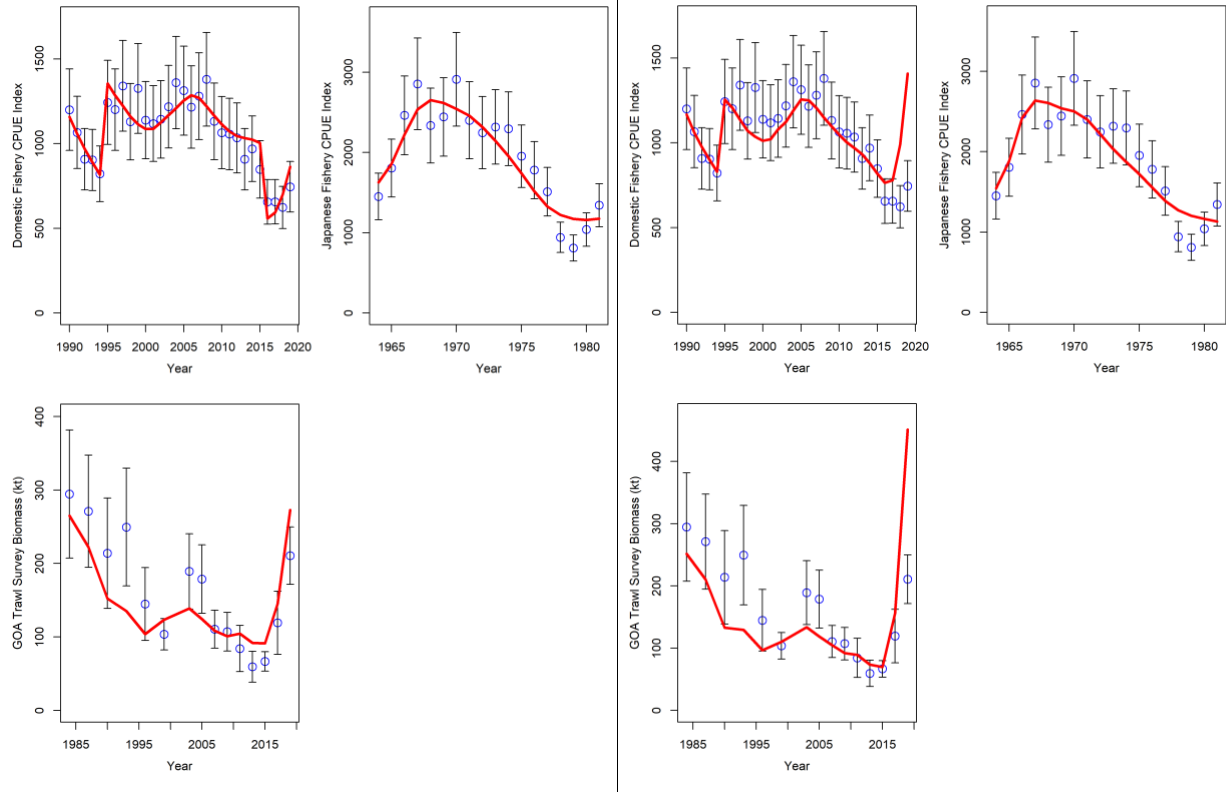
Figure 17. Fits of the new proposed model (*21.10_Proposed*; left two columns) and the Continuity model (*16.5_Cont*; right two columns) to abundance indices. Observed and predicted sablefish relative population weight and numbers for 1990 - 2020 for U.S. longline survey and for 1979 - 1994 for U.S.-Japan cooperative survey. Points are observed estimates with approximate 95% confidence intervals. Solid red line is the model predicted values. The relative population weights are not fit in the models, but are presented for comparison.

Figure 18 Fits of the new proposed model (*21.10_Proposed*; left two columns) and the Continuity model (*16.5_Cont*; right two columns) to abundance indices. Fishery CPUE indices are on top two panels. GOA trawl survey is on the bottom left panel. Points are observed values with approximate 95% confidence intervals, while solid red lines are model predictions.
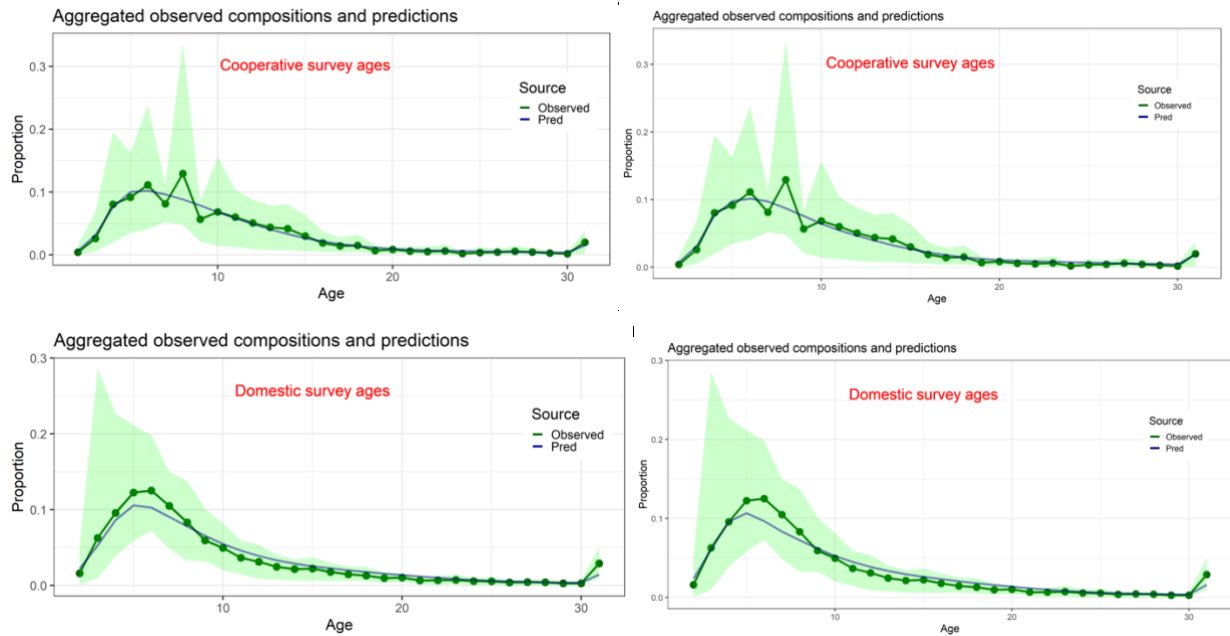
Figure 19. Fits of the new proposed model (*21.10_Proposed*; left panel) and the Continuity model (*16.5_Cont*; right panel) to aggregated compositional data. Mean observed (green line) cooperative (top panel) and domestic (bottom panel) longline survey age compositions aggregated across years along with the average fit of the model (blue line). The green bands are the 90% empirical confidence intervals.
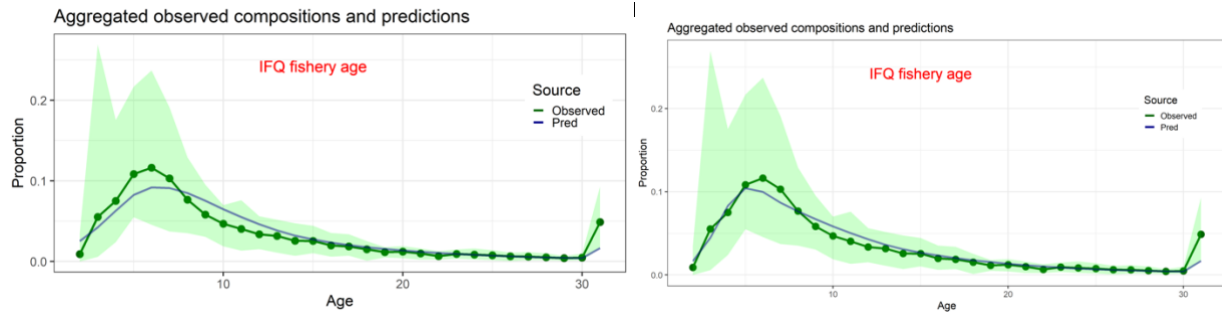
Figure 20 Fits of the new proposed model (*21.10_Proposed*; left panel) and the Continuity model (*16.5_Cont*; right panel) to aggregated compositional data. Mean observed (green line) domestic fixed gear fishery age compositions aggregated across years along with the average fit of the model (blue line). The green bands are the 90% empirical confidence intervals. Note the perceptibly worse fits of the *21.10_Proposed* model for ages two through five.
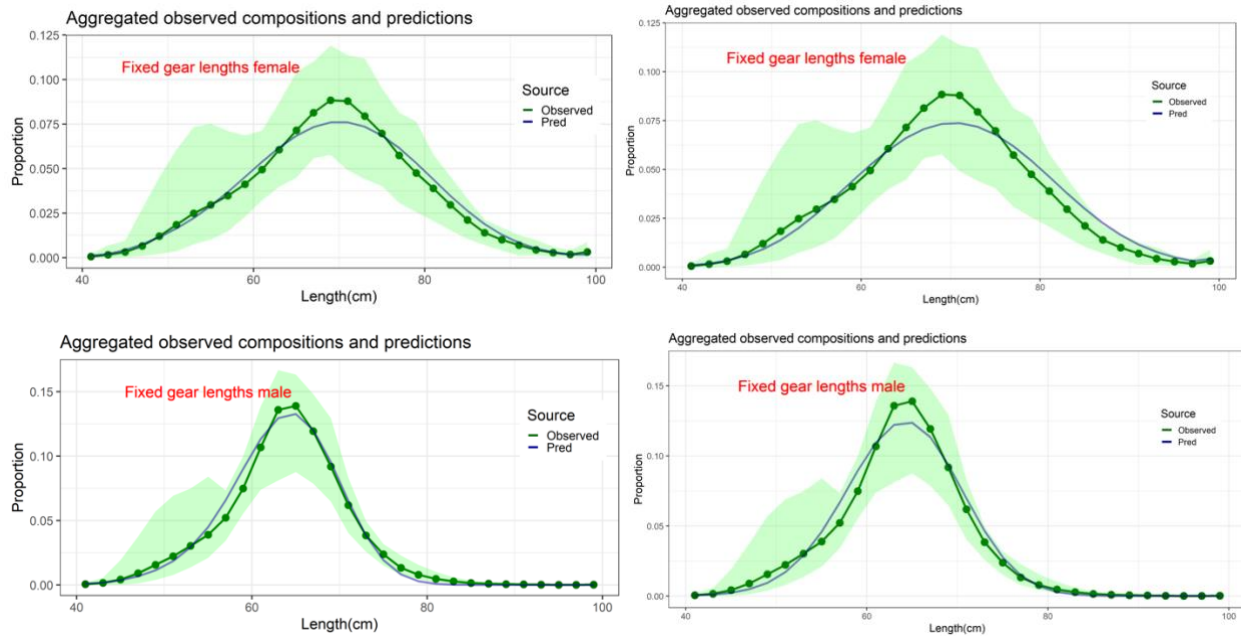
Figure 21. Fits of the new proposed model (*21.10_Proposed*; left panel) and the Continuity model (*16.5_Cont*; right panel) to aggregated compositional data. Mean observed (green line) domestic fixed gear fishery length compositions aggregated across years along with the average fit of the model (blue line). The green bands are the 90% empirical confidence intervals. Fit to female length compositions are provided in the top panel and fit to male length compositions are provided in the bottom panel.
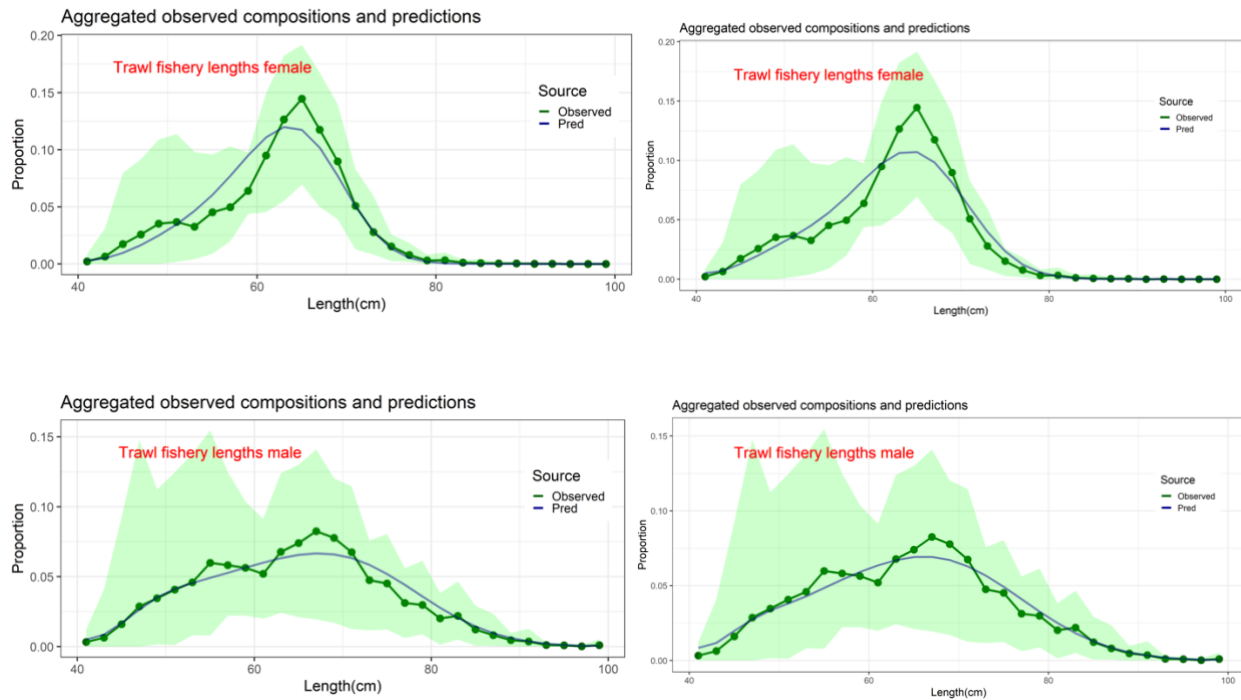
Figure 22. Fits of the new proposed model (*21.10_Proposed*; left panel) and the Continuity model (*16.5_Cont*; right panel) to aggregated compositional data. Mean observed (green line) domestic trawl fishery length compositions aggregated across years along with the average fit of the model (blue line). The green bands are the 90% empirical confidence intervals. Fit to female length compositions are provided in the top panel and fit to male length compositions are provided in the bottom panel.
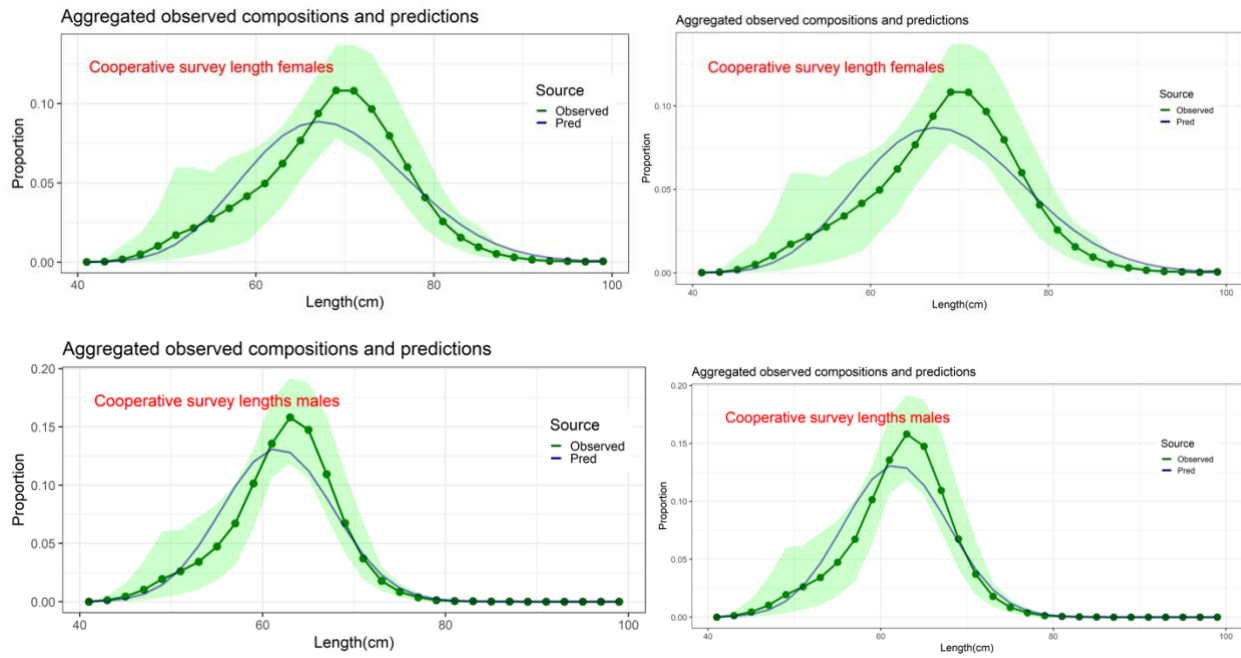
Figure 23. Fits of the new proposed model (*21.10_Proposed*; left panel) and the Continuity model (*16.5_Cont*; right panel) to aggregated compositional data. Mean observed (green line) cooperative longline survey length compositions aggregated across years along with the average fit of the model (blue line). The green bands are the 90% empirical confidence intervals. Fit to female length compositions are provided in the top panel and fit to male length compositions are provided in the bottom panel.
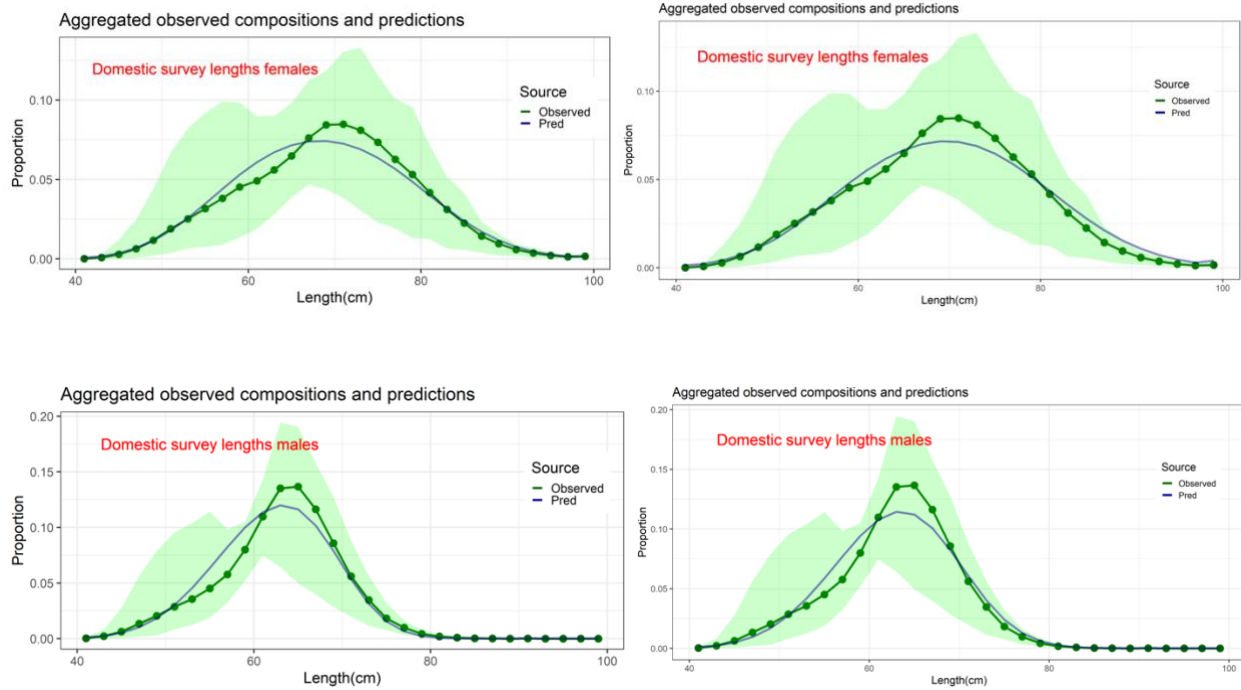
Figure 24. Fits of the new proposed model (*21.10_Proposed*; left panel) and the Continuity model (*16.5_Cont*; right panel) to aggregated compositional data. Mean observed (green line) domestic longline survey length compositions aggregated across years along with the average fit of the model (blue line). The green bands are the 90% empirical confidence intervals. Fit to female length compositions are provided in the top panel and fit to male length compositions are provided in the bottom panel.

Figure 25. Fits of the new proposed model (*21.10_Proposed*; left panel) and the Continuity model (*16.5_Cont*; right panel) to aggregated compositional data. Mean observed (green line) Gulf of Alaska trawl survey length compositions aggregated across years along with the average fit of the model (blue line). The green bands are the 90% empirical confidence intervals. Fit to female length compositions are provided in the top panel and fit to male length compositions are provided in the bottom panel.
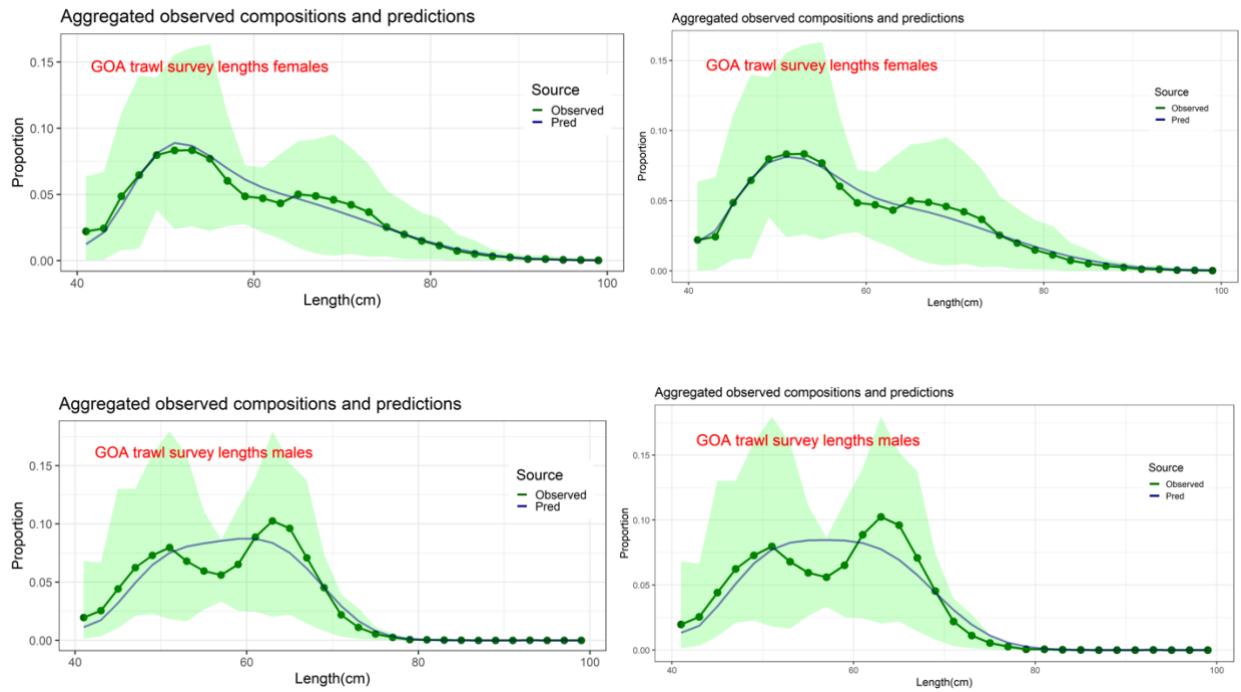
Appendix A.1: 2021 SAFE Data Updates

For the November stock assessment, we plan to include updated data from both survey and fishery data sources. The 2020 AFSC longline survey age compositions and the 2021 longline survey relative abundance and length data are expected to be available. The AFSC 2021 Gulf of Alaska bottom trawl survey biomass estimates and length compositions will also be available.

For fishery data, the final 2020 catch and projected 2021 catch will be included along with estimates of killer and sperm whale depredation and forecasted catch for future years based on the NPFMC B_{40} HCR. The availability of 2020 fishery catch rate data for inclusion in the fishery catch rate (i.e., CPUE) index are uncertain at this time. The 2020 observer data is available with the caveat that we cannot compute catch rates for use in the index with the data currently available from the Electronic Monitoring (EM) Program. Therefore, only trips with human observers have sufficient data at this time for input to the fishery catch rate index. However, the sample sizes from this subset of data are not sufficient alone to estimate catch rates from all sablefish management areas (Table A.1; see Goethel et al., 2020 for more information on data inputs for the CPUE index). The number of sets observed in 2020 were much lower than in previous years. These low sample sizes were likely due to: 1) an increase in pot fishing and EM compared to trips using human observers and hook-and-line gear (i.e., the catch rate index is based only on hook and line gear at this time); 2) the observer deployment plan; and 3) the COVID-19 pandemic and the overall lower number of human observed trips for all fisheries in 2020.

Additionally, we do not anticipate having 2020 data available from the sablefish volunteer logbook data in time for this year's assessment. This is the primary data source for the fishery catch rate dataset. This data is provided through a grant with the International Pacific Halibut Commission (IPHC), whose staff collect the logbooks dockside and keypunch the data before providing it to the sablefish authors. The funds available to the IPHC for this year's grant were insufficient to provide all specified services. As a result, keypunching of the 2020 logbooks have not been completed at this time due to a lack of funds and it is uncertain when this data will become available. As a result, we do not feel that sufficient sample sizes exist from solely human observed sablefish trips to compute a 2020 fishery catch index for use in the 2021 assessment model. Thus, without the ability to include EM sets or the availability of IPHC volunteer logbook sets, we recommend not including a 2020 data point in the fishery catch rate index for this year's assessment model.

Table A.1. Number of hook and line fixed gear sets covered by at-sea human observers from 2016 to 2020. NA represents years where data cannot be reported due to confidentiality concerns.

Year	AI	BS	WG	CG	WY	EY/SE
2016	184	0	251	732	140	228
2017	NA	14	81	389	86	229
2018	NA	NA	108	339	138	188
2019	NA	18	148	344	214	217
2020	0	10	13	90	68	109