Aleutian Islands Golden King Crab Stock Assessment Draft Models

Tyler Jackson
Alaska Department of Fish and Game, tyler.jackson@alaska.gov

September 2024

Response to Comments

SSC June 2024

Comment: "*The SSC requests the rationale for using the terminal year minus four year approach to define the reference period for future assessments*"

Response: This comment will be addressed during the final assessment.

. . . .

Comment: "*The SSC recommends that the CPT explore whether to conduct this final assessment on the same cycle as other crab assessments in September/October to better align the assessment with the annual cycle of catch mortality.* "

Response: ADF&G and the CPT do not think moving the timing of the final assessment to September would be suitable because the fishery opens in August.

. . . .

Comment: "*The SSC recommends prioritizing further consideration of data weighting, as the Francis re-weighting continues to be an issue in this assessment.*"

Response: This analysis explores data weighting, specifically with respect to size composition data.

. . . .

Comment: "*The SSC places a high priority on incorporating information from the cooperative survey into the assessment and supports the CPT recommendation to incorporate this survey as a separate fleet.*"

Response: See models 25.1 and 25.1b.

. . . .

Comment: "*Further examination of the retrospective pattern in terms of magnitude, direction and cause continues to be important.*"

Response:

Retrospective bias in MMB appears to arise from data conflict that also results in poor fit to post rationalization observer CPUE in the EAG, which has the same retrospective pattern. Data conflict is less apparent in the WAG. This document discusses model misspecfication, but more work is necessary.

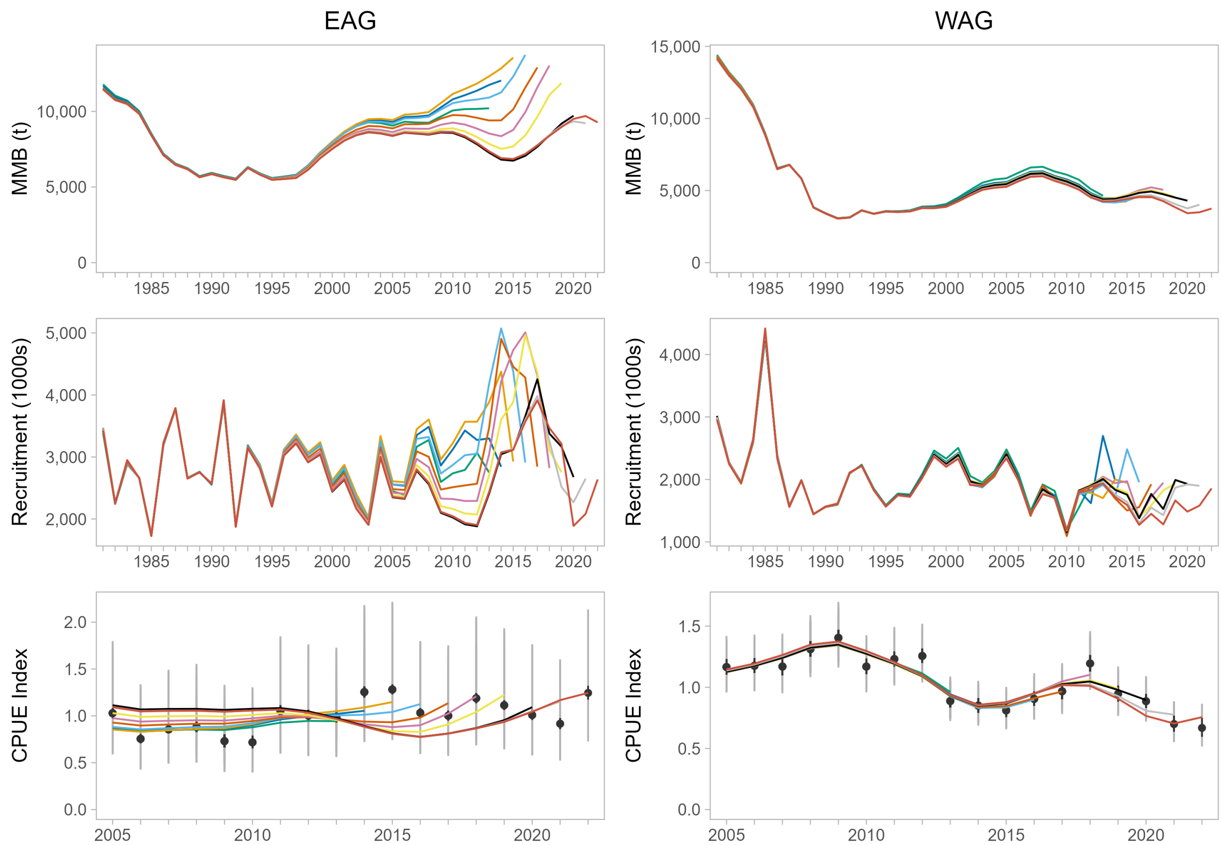

Figure 1: MMB trajectory, recruitment, and fits to observed CPUE index for EAG retrospective peels up to 10 years.

. . . .

Comment: "*The CPT suggested that next year's model should be 25.0. The SSC reminds the CPT and authors that new model year numbers are only applicable if there is a major structural model change.*"

Response: Noted.

. . . .

Comment: "*The current method of projecting the remaining landings for the current incomplete season seems overly complicated and the SSC recommends a more straightforward method for determining total catch be considered, such as basing it on the average fraction harvested to date.* "

Response: See response to this comment in the 2024 SAFE. If the directed fishery is incomplete at the time of the final assessment, total catch will be computed as usual, though assuming the TAC is achieved and CPUE remains the same as at the time data were pulled. Total effort will estimated by dividing the TAC in terms of the estimated number of crab by current CPUE. See Appendix A of Jackson (2024a) for details of total catch estimation.

. . . .

CPT May 2024

Comment: "*Use the standard convention for model numbering, i.e. the models for the May 2025 assessment, will be 25.xx and not 24.xx.*"

Response: This is done.

. . . .

Comment: "*Document why the 1993 bycatch and total catch size-composition data are not included in this and past assessments.*"

Response: Issues with 1993 observer data (not bycatch) have been resolved by model 23.1 (data). There was no observer coverage in the 1993 crab year for EAG directed fishing and for WAG, most pots were previously removed because they were of unknown size. Those pots are now included for size composition data only, which also affects total catch estimation.

. . . .

Comment: "*Explore reasons for the retrospective pattern for the EAG.*"

Response: See above. This work is ongoing.

. . . .

Comment: "*Consider models for the EAG and WAG that allow for the bias-correction in recruitment, especially given there is virtually no information in the data on the sizes of the recruitments before 1985.*"

Response: See model 23.1c.

. . . .

Comment: "*Include the EAG cooperative survey data (index and size-composition) as an additional fleet.*"

Response: See models 25.1 and 25.1b of this analysis.

. . . .

Comment: "*Fit models that assume that the size-composition data are Dirichlet-multinomial distributed instead of Francis weighting the size-composition data.*"

Response: See models 25.0c and 25.0d of this analysis.

. . . .

Comment: "*Explore the reasons for the implausible values for groundfish fishing mortality in some years for some of the retrospectives and some of the jitter runs.*"

Response: This was presented at the May 2024 plan team meeting. The model estimated a large recruitment pulse preceding 1996, which appeared to allow for better fit to 1996 observer index data. Because size composition data in the directed fishery and the base natural mortality rate could not support such a recruitment pulse, the model 'killed off' the extra crab in the bycatch fishery which does not have associated composition data and large CV. Running the model with a .pin file that specifies appropriate starting values for F deviations or increasing the penalty on F deviations in that fleet resolves the issue.

. . . .

Comment: "*Consider starting the model in a non-equilibrium state around 1981.*"

Response: Model 25.0 and its derivations addresses this.

. . . .

Comment: "*Revisit estimation of size-at-maturity given the addition of new data.*"

Response: Due to the short turn around from the May 2024 assessment to now, this was prioritized lower. It can be revisited during the next cycle.

. . . .

Comment: "*Continue exploration of CPUE standardization, including investigation of models with block:year interactions and using geostatistical methods.*"

Response: CPUE standardization will be revisited in the next cycle.

. . . .

Comment: "*Explore time-varying catchability (e.g. as blocks) rather than the use of additional variance to reconcile the trends in CPUE and those in abundance. Given the known difficulties estimating time-variation in catchability, this could be explored as part of a simulation study – with initial discussions at the January 2025 modeling workshop.*"

Response: It would be good to review simulation using GMACS during the next modelling workshop so this could be explored.

. . . .

SSC June 2023

Comment: "*Further examination of the retrospective pattern in terms of magnitude, direction and cause continues to be important.*"

Response: See above, this work is ongoing.

. . . .

CPT May 2023

Comment: "*Identify and eliminate the conflict between the model and the data giving rise to the retrospective patterns for EAG models. Revisit the analysis considering a model with time-varying catchability, but impose a penalty on the devs to allow the index data to inform the model.*"

Response: This work is ongoing. Training on simulating data using GMACS would be a useful tool in identifying data model misspecification.

. . . .

Comment: "*Calculate reference points using both combined-area and area-specific size-at-maturity values.*"

Response: Models that use area specific size at maturity will be explored in a future cycle.

. . . .

Comment: "*Continue work to obtain an index using the cooperative pot survey data for use in the EAG assessment model.*"

Response: Work on cooperative survey data is continually improving, see Appendix A of this document and model 25.1 and 25.1b.

. . . .

Comment: "*Explore models that provide better fits to EAG CPUE data.*"

Response: Resolving poor fit to EAG CPUE data is the primary goal of data weighting explorations in this cycle. Specifically, model 25.0b2 forces a better fit to those data.

. . . .

Model Explorations

CPT and SSC comments recommend a variety of modifications to the assessment model including exploration of data weighting, addition of the cooperative survey, a combined EAG and WAG model, improvements to observer CPUE standardization and evaluating area specific size at maturity. Given the short turn around time since the last final assessment (May 2024) not all of these issues could be explored in this cycle. Poor fit to index data and associated retrospective bias in MMB in the EAG has been an ongoing issue for several assessment cycles, so evaluating data weighting was given high priority. In addition, there has been a cooperative survey in the EAG during eight of the last nine years and there has been much recent effort to make those data available to the assessment process. Models explored in this document address prior comments on questionable data, initial conditions, data weighting, and the cooperative survey (See Figure 2 for flow chart). These include:

- **23.1**: base model, accepted for specifications in May 2024;
  - **23.1 (update)**: base model, implemented in GMACS version 2.20.16;
  - **23.1 (season)**: 23.1 (update), with reconfigured seasons to allow for output of N matrix on June 30th of the terminal year.
  - **23.1 (data)**: 23.1 (season), with updated total size composition and total catch time series (i.e., including additional pot types). **This model becomes the new version of the base model, 23.1**;
- **23.1c**: Model 23.1, with corrected recruitment bias correction from 1960 - 1980;
- **25.0**: Model 23.1, with non-equilibrium initial conditions, starting in 1981;
- **25.0a**: Model 25.0, with equal emphasis factors on all likelihood components;
- **25.0b**: Model 25.0a, with bootstrap estimated stage 1 effective sample sizes;
- **25.0c**: Model 25.0b, using the Dirichlet multinomial likelihood for size composition data;
- **25.0d**: Model 25.0a, using the Dirichlet multinomial likelihood for size composition data;
- **25.1 (EAG Only)**: Model 25.0b, with cooperative survey index and size composition data;
- **25.1b (EAG Only)**: Model 25.0c, with cooperative survey index and size composition data.

Updates to Base Model

The 2024 final assessment was conducted using GMACS version 2.01.M.10. This document uses the updated version 2.20.16. Small changes in the likelihood values of priors are expected, though all other model processes should be unchanged for this stock. Model 23.1 and all previous GMACS models for this assessment used the following annual structure consisting of six seasons:

1. Instantaneous: Start of year, output numbers at size;
2. July 1 to the mid-point of the directed fishery: Natural mortality;
3. Instantaneous: Removal of directed catch, bycatch, fit index and size composition;
4. Mid-point of the directed fishery to February 15: Natural mortality;
5. Instantaneous: Estimate MMB;
6. February 15 - June 30: Natural mortality, recruitment, growth.

Annual structure was reconfigured by swapping the duration of seasons five and six. Season five now runs from February 15 - June 30, with the same model process occurring, though with the addition of MMB estimation. MMB is estimated at the beginning of a season, so there is no changed to estimates. Season six is now instantaneous (June 30) and is used to output numbers at size so that the ADF&G harvest strategy can utilize estimated legal male abundance as close to the upcoming fishery as possible. Nothing happens during season 1, which has length zero and serves as a place holder. This is a small process change that has no effect on model results or reference point calculation (Table 1 - 3).

Observer Coverage During the 1993/94 Fishery

Previous assessment models removed observer data for the EAG in 1993, but retained those data for the WAG (Jackson 2024b; Siddeek et al., 2023). The CPT recommended investigating those data and better documenting the reason for use or non-use.

Prior to 1996, the longitudinal dividing line between the EAG and WAG was 171° W lon, and fishing seasons did not conform to the post-rationalization crab year. Jackson (2024a) re-computed time series data by applying the current management boundary (174° W lon) and crab year to these early time series data. The 1993/94 season in the EAG (east of 171° W lon) was open from September 1, 1993 to March 1, 1994, and there was no observer coverage in the EAG during that season. Observer data that are assigned to the 1993/94 EAG season were actually from the easternmost portion (174° W lon - 171° W lon) of the WAG fishery in 1992/93, which ran from November 1, 1992 to August 15, 1993 (i.e., July 1 - August 15, 1993 get assigned to the 1993/94 crab year). Previous assessments included these data as part of the 1993/94 EAG retained catch and retained size composition, but remove these data from the total catch and total size composition time series, which seems appropriate.

Observer size composition for the WAG in 1993/94 was also flagged as suspect, given the stark difference between those data and the surrounding seasons. Observer data for WAG that are assigned to the 1993/94 season come from both the 1992/93 and 1993/94 seasons, and size composition was recorded for all observer pots. Previous assessments removed various uncommon or non-target species gear types by recommendation from the fleet (M. Siddeek, ADF&G personal communication, 2023). Those include: Dungeness crab pots, pyramid pots, conical pots, hair crab pots, snail pots, cods, dome shaped pots, ADF&G research pots with stretch mesh instead of escape rings, and rectangular pots measuring 9'x 9', 8 1/2' x 8 1/2', 9 1/2' x 9 1/2', 8' x 9', 8' x 10', 9' x 10', 7' x 8', or with unknown dimension. In 1993, nearly all WAG observer (162 / 174) pots belonged to one of those categories, mostly unknown sized rectangular pots (160). Including all rectangular pots has little impact on the size composition in other years in either subdistrict, so the total size composition time series was revised to include all rectangular pots (Figure 3 and 4). Changes to total size composition data also have a minor impact on total catch estimates (Table 4). Pot types included in CPUE estimation will remain unchanged. These difference are evaluated in model 23.1 (data), for which there was very little difference in model results (Figures 3 - 5) or reference points (Table 3).

Recruitment Bias Correction

Model 23.1 suggests steadily decreasing MMB from 1960 - 1981 due to how previous versions of GMACS implemented bias correction of log recruitment deviations. Model 23.1 estimated initial recruitment as a model parameter (R_0) and recruitment by year is estimated as series of annual deviations. Log-deviations have a tendency to go to zero in lieu of data, though standard errors are large and decrease as the spin up period approaches the beginning of the data time series in 1981. GMACS version 2.01.M.10 added $e^{\frac{\sigma^2}{2}}$ to expected values of log-deviations, resulting in decreasing recruitment, and thus MMB. Model 23.1c implemented bias correction as $b_t e^{\frac{\sigma^2}{2}}$, in which b_t is a vector of 0 for years preceding data (1960 - 1980) and 1 for years with time series data (1981 - 2023).

Intitial conditions

Previous assessment models started in 1960 assuming the population was in an unfished state, with no fishing mortality until 1891 (Siddeek et al. 2023; Jackson 2024b). Model 25.0 starts in a non-equilibrium state in 1981 and initial numbers at size were estimated from parameters representing initial recruitment (R_{init}) and scaled deviates for each size class. The size bin at which crab first become legally retained, 136 - 140 mm CL, was used as the reference size class.

Data Weighting

Base model data weighting

Recent CPT and SSC comments reflect a need to reevaluate data weighting. The data weighting scheme for the base model was as follows:

- **Catch Data**: Retained catch *cv* was assumed to be 0.0316 for all years and the emphasis (λ) on the likelihood is $\lambda = 4$. Total catch *cv* was the graded number of observer sampled pots with non-zero catches ($m_{nz,i}$) in which the maximum weight (ω_i) is $\max[\omega_i] = 250$, scaled as

$$cv_i = \sqrt{e^{\frac{1}{2\omega_i}} - 1} \tag{1}$$

$$\omega_i = \frac{\max[\omega_i] m_{nz,i}}{\max[m_{nz,i}]} \tag{2}$$

Likelihood emphasis for total catch was $\lambda = 2$. Groundfish bycatch *cv* is assumed to be 1.3108 for all years and $\lambda = 1$.

- **Index Data**: Standard errors of year coefficients from CPUE standardization models were used to compute annual *cv* for each of the index series. Observer index *cv* ranged from 0.017 - 0.049 in the EAG and 0.019 - 0.059 in the WAG. Fish ticket *cv* was slightly greater, ranging from 0.044 - 0.178 in the EAG to 0.038 - 0.093 in the WAG. Since these estimates were overly precise due to the vast number of observer and fish ticket records, additional *cv* was estimated as a parameter for each series. Additional *cv* estimates were larger than observed *cv* for all indices, between 0.15 - 0.23 in the EAG, and 0.06 - 0.24 in the WAG (Table 10 and 11; Jackson 2024b). Likelihood emphasis for all index series was $\lambda = 1$.
- **Size Composition Data**: The base model assumed a multinomial error distribution for size composition data with effective sample size based on the number of vessel days and observer days for retained and total size composition data, respectively. Effective sample sizes were tuned using the Francis (2011) method for the multinomial. Based model Francis weights were $\lambda_{ret} = 0.209$ and $\lambda_{tot} = 0.432$ in the EAG and $\lambda_{ret} = 0.143$ and $\lambda_{tot} = 0.521$ in the WAG. Likelihood emphasis for both size composition series was $\lambda = 1$.
- **Tagging Data**: Sample size for growth data were the tag return sample sizes and likelihood emphasis was $\lambda = 1$.

Likelihood Emphasis

Likelihood emphasis factors (λ) provide a means of weighting the sum of likelihood components to influence the overall model fit. The accepted model (23.1) assigns the most weight to retained ($\lambda_{ret} = 4$) and total ($\lambda_{tot} = 2$) catch data. Model 25.0a evaluated resetting all likelihood emphasis factors equal to 1 (i.e., all data sources weighted equally).

Model 25.0b2 explores increasing emphasis on index data ($\lambda = 2$). This model is not an attempt to resolve the data conflict in the EAG, but to force a better fit to index data without a change in assumptions and examine the sensitivity in other model processes.

Multinomial Effective Sample Size

Model 25.0b evaluates Francis weighting of size composition data using sample sizes based on variability in the data. Stage 1 effective sample sizes N_{eff} were estimated using a bootstrap approach based on Stewart and Hamel (2014). The resampling design was two-staged, first within the primary sampling unit (delivery or observer pot) and second by individual crab. Resampling was done with replacement to observed sample sizes. Effective sample size for a given year was estimated as

$$N_{eff} = \frac{\sum_l P_l(1 - P_l)}{\sum_l (P_l - B_l)^2} \tag{3}$$

where P_l and B_l are the observed and bootstrap proportion of crab measured in length bin, l, respectively. Annual estimates of effective sample size were computed as the mean of 500 replicates for each year for the retained size composition and 100 replicates for each year for the total size composition. Resampling iterations were low due to computational needs. Estimated N_{eff} are in Table 5 and 6. Since estimated N_{eff} was quite large for many years, observed stage 1 sample sizes were reduced to

$$N'_{eff} = \min(2000, N_{eff}) \tag{4}$$

Dirichlet Multinomial

Model 25.0c assumes the Dirichlet multinomial error distribution for both series of size composition data (Thorson et al. 2017). Stage 1 effective sample sizes were the same as model 25.0b. Model 25.0d also assumes the Dirichlet multinomial for size composition data, but uses stage 1 effective sample sizes from model 25.0a. This was included to illustrate that the Dirichlet overdispersion parameter hits the upper bound (β_{DM} [UB] = 1,000) with smaller input sample sizes than used in model 25.0c, indicating that the estimated sample size was larger than the input sample size.

Cooperative Survey

The AIGKC cooperative pot survey was initiated in 2015 and has continued every year since, with the exception of 2020. The survey has been confined to the EAG in all years but 2018, when it was extended to the WAG. Appendix A details the survey design and analytical approach. Briefly, an index of total male abundance was estimated from a Tweedie GLMM that included fixed, parametric effects of soak time (days) and depth with string as a random effect nested within survey subarea. Survey size composition was estimated by the same method as observer data, and effective sampling size was estimated using the bootstrap approach described above.

Model assumptions relating to survey data include:

- The survey occurs during season 3 (same as the directed fishery);
- There is no extra CV estimated for the survey index;
- Catchability is estimated independent of the directed fishery;
- Selectivity is logistic and estimated independent of the directed fishery;
- There is no retention or handling mortality. Retention of legal crab is accounted for in the directed fishery catch. Thus, $\bar{F} = 0$ and $F_{35\%}$ is not applied to this fleet;
- Size composition data are fit using a multinomial likelihood with input sample sizes tuned using Francis weighting (model 25.1) or the Dirichlet multinomial (model 25.1b).
- Survey size composition data are independent of observer (i.e., total) size composition data. Observers are not deployed during survey trips.

Diagnostics

Retrospective Runs

Retrospective bias was evaluated by iteratively re-running a model and ‘peeling’ (i.e. removing) the terminal year for each iteration. Mohn’s ρ (Mohn 1999) was used to compare retrospective bias in MMB between models:

$$\text{Mohn's } \rho = \frac{1}{n}\sum_{y=1}^{n}\frac{|\text{MMB}_y - MMB|}{MMB} \tag{5}$$

where MMB_y is the terminal year mature male biomass for each peel, MMB is the mature male biomass for the full model, and n is the number of peels. Here the difference in MMB was computed using the absolute value to avoid erroneously low ρ estimates since bias can be both positive and negative. Retrospective runs were performed for models 23.1, 25.0b, 25.0c, and 25.1, peeling up to 10 years. Only 8 years were peeled for model 25.1, since the cooperative survey data only extend back to 2015.

Recruitment Profiles

Likelihood profiles of recruitment have been used as a diagnostic tool for identifyingdata conflict, specifically in relation to selectivity (Ichinokawa et al., 2014; Wang et al., 2014; Maunder and Piner 2015; Carvalho et al., 2017). Typically, relative likelihood profiles are compared to profiles of a model using data simulated without error. Differences between the realized profile and the simulated profile can be interpreted as misspecification, though this method may have a low detection rate (Carvalho et al., 2017). No data simulations were performed in this analysis (the author does request training in using data simulation features of GMACS). Here likelihood profiles were constructed to evaluate which data components were most informative for estimating R_0 (model 23.1c) and $\bar{R}$ (model 25.0b). These parameters were chosen due to their influence on abundance and MMB. Difference in likelihood profiles were also compared between EAG and WAG models, since those models were structured the same and use similar data.

Likelihood profiles were constructed by fixing R_0 or $\bar{R}$ at a range of values between the 95% confidence interval of the maximum likelihood estimate, while re-estimating all other parameters. The weight associated with the F dev penalty on groundfish fishery bycatch was increased to 0.5 to avoid erroneously high estimates for WAG models 23.1c and 25.0b, and EAG 25.0b. Likelihoods were scaled to relative values by subtracting the minimum likelihood for each component. Likelihood components evaluated were catch data, growth data (i.e., tagging data), index data, size composition data, and the recruitment penalty.

Results and Discussion

Initial Conditions

Changing the basis correction on recruitment before 1981 (model 23.1c) resulted in a flatter MMB trend in years preceding data, but did not totally eliminate the decreasing ramp in the EAG (Figure 6). Changes in recruitment bias correction had only marginal effect, if any, on other model processes. Starting the model in 1981 increased recruitment in the starting year in both subdistricts, and to a lesser extent from 2017 - 2023 in the EAG. The size distribution of recruits was the same among models. The 1981 recruitment estimate has little effect on other model processes. The only data pre-1985 are retained catch data. Retained size composition data starts in 1985 and total catch data begins in 1990. Differences in fits to size composition and index data were minimal, though present in several years in both subdistricts. There was an increase in MMB_{prj} in the EAG for model 25.0 (Table 8), owing to the small increase in recruitment and MMB towards the end of the time series. Despite there being a small change in recruitment, starting the model in 1981 is preferred as it slightly reduces the number of estimable parameters and eliminates the issue of decreasing MMB and recruitment before the data time series.

Data Weighting

Reducing the emphasis on retained and total catch data so that likelihood components are equally weighted had little impact on model results in either subdistrict. Bootstrap estimated N_{eff} for size composition data were highly variable and larger than the observed number of crab measured in many cases. This is possible when bootstrap proportion at size approximates observed size composition with high precision. Annual means of bootstrap N_{eff} were less than observed sample sizes in all years and followed the same trend, though the bootstrap distribution was highly variable (Table 5 and 6). Sample sizes were largest for the retained catch in the 1990s and have stabilized since the 2000s when sampling design restricted effort (Table Table 5 and 6).

Francis weighting with bootstrap estimated stage one N_{eff} (model 25.0b) had little impact on fits to catch and size composition data in both subdistricts. Francis weights were smaller than in model 25.0a and resulted in a similar stage two N_{eff} for retained size composition and an approximate time series average of model 25.0a for total size composition (most years had stage one N_{eff} capped at 2000) (Figure 11 and 12). Model 25.0b estimated slightly lower recruitment from 2020 - 2023, which is reflected in fit to index data and estimated MMB (Figure 17 and 18).

Model 25.0c estimated N_{eff} for both retained and total size composition data considerably higher than models that used Francis weighting (Figure **??**). Estimated N_{eff} for model 25.0d was the same as stage one sample sizes since overdispersion (β_{DM}) reached its upper bound for both size composition series. Increased weighting of size composition data by Dirichlet multinomial models resulted in larger estimates of added CV on post-rationalized observer CPUE data, thus down weighting those data (Table 7). Naturally, fits to index data in the EAG were worse for models 25.0c and 25.0d (Figure 8). This was also the case in the WAG to a lesser degree (Figure 9). Increasing emphasis on index data (model 25.0b2) resulted in much lower Francis weights than model 25.0b and near zero additional CV on post-rationalized index data for the EAG (Table 7; Figure 8) (i.e., the tuning process down weighted size composition and further increased weight on index data, fitting the data near exactly). Weights tuning was less sensitive to increasing index likelihood emphasis in the WAG, but had the same general effect.

Fits to size composition data were good and differed little among models (Figure 10). All models characterize mean size of the retained catch well in the first half of the time series, and then tended to under predict and then over predict mean size during stretches of years in both subdistricts (Figure 13 and 15). This pattern was least prevalent in model 25.0c and 25.0d, as those models weighted size composition data more. All models except 25.0b2 in the EAG characterized mean carapace length of the total size composition well (Figure 14 and 16).

Model 25.0c estimated lower recruitment from 2014 - 2023 in the EAG, but not the WAG. Forcing a better fit to post-rationalized observer CPUE (model 23.0b2) resulted in large, annual swings in recruitment in the EAG (Figure 17). Such large annual fluctuations are likely implausible. Increasing the recruitment smoothness penalty would temper recruitment swings (not shown in this document), though there is no objective bias for the value of that penalty for this stock. In general, WAG recruitment was less sensitive to data weight in the latter part of the time series that in the EAG. Model 25.0c estimated lower MMB from 2014 - 2023 in the EAG, following the same trend in recruitment (i.e., recruitment can occur at maturity). Model 25.0b2 estimated a different MMB trajectory than other models in the EAG, and inter-annual differences were somewhat erratic (Figure 18). All other EAG models had similar MMB estimates, with models 25.0 and 25.0a being slightly greater than models 25.0b and 25.0d (models with bootstrap N_{eff}) at the end of the time series. MMB was less sensitive to data weighting in the WAG (Figure 18).

Explorations in weighting size composition and index data did not resolve data conflicts in the EAG, but demonstrated that model results are indeed sensitive to weighting, which is indicative of mispecification (Maunder and Piner 2015; 2017). EAG model misspecification is likely within the fishing process as fitting post-rationalized observer CPUE data resulted in biologically implausible recruitment and MMB trends (Figure 17 and 18). Index and total size composition data used in this assessment are based on at-sea observations during the directed fishery. Since crab are counted and measured in all observer samples these data sources are generated from the same sampling process, and thus are subject to the same potential biases. It is unlikely that CPUE standardization is able to account for all variability arising from factors that influence catchability and the nature of fishing (targeting legal males, balancing logistics with maximizing CPUE, etc.) likely lends itself to variability in availability at size. Estimating additional observation error in index data and tuning size composition weight allows observation error to compensate for unmodelled process error, somewhat ignoring the index data. The CPT recommended evaluating time varying processes as part of a simulation study, which should be a priority during the next cycle. It seems less likely that data conflict is driven by natural mortality or growth, as both are based on EAG tagging data (Siddeek et al. 2016; Siddeek et al. 2022) and used in WAG models. WAG models contain less apparent data conflict and were not sensitive to data weighting.

Cooperative Survey

Models 25.1 and 25.1b added the EAG cooperative survey as an additional fleet. Appendix A details estimation of the cooperative survey index and size composition data. Fits to survey index roughly followed the average survey index from 2015 - 2019, and decreased in 2022 and 2023. Predicted survey index did not capture the large decrease in 2021 (Figure 19). Both models that included the cooperative survey fit post-rationalized observer CPUE similarly to model 25.0c (Figure 19). Estimates of additional CV were larger than models without survey data (Table 7).

Both models fit survey size composition data reasonably well, except for in 2018 - the only year in which survey size composition substantially deviates from the fishery total size composition (Figure 20; Appendix A Figure 14). Francis weighting (model 25.1) resulted in lower weights on size composition survey size composition data than the Dirichlet multinomial (model 25.1b), though predicted proportions at size were not very different.

Recruitment estimates tended to align closely with the parent model (25.0b or 25.0c) until the beginning of the survey time series, at which both models 25.1 and 25.1b undergo a steeper decline mirroring model 25.0c (Figure 21). Estimated MMB followed a similar trajectory, with both models estimating terminal year MMB considerably lower than the base model (Figure 22).

Previous efforts to incorporate the cooperative survey spliced survey CPUE into the observer time series (Siddeek et al. 2023) and included the survey as a separate fleet, but allowed for estimation of additional CV on index data (Jackson 2024). Models 25.1 and 25.1b are the most appropriate use of the survey data that have been presented thus far. Including the survey data does not appear to necessarily improve overall model performance (as mentioned above it actually results in down weighting and poorer fit to index data). Since MMB and management quantities do appear to be sensitive to the addition of survey data, there should be careful consideration of whether the survey design appropriately captures population trends. The current survey analysis post-stratified survey stations (i.e., strings) based on the three strata design implemented since 2022. Number of stations per stratum has varied annually, as has ability to haul strings set at all selected stations. Further vessels tend to not overlap in the survey footprint, so vessel effects would be masked by spatial variability, or vice versa. Due to logistic constraints, there were fewer stations sampled in 2024 compared to previous years, though the spatial extent was similar to 2023.

Retrospective Bias

All EAG models demonstrated the same pattern of retrospective bias in MMB that has been an issue for years (Figure 23). For models without the cooperative survey, Mohn's ρ ranged from 0.36 - 0.499. Retrospective bias was worse for model 25.0b than 23.1c, though both are problematic in peels further back than 2020. The retrospective pattern of model 25.1 should be interpreted cautiously, as the survey time series is not long. Its worth noting that the 2019 peel has less retrospective bias than models 23.1c and 25.0b. Retrospective bias is not a concern for models 23.1c or 25.0b in the WAG ($\rho = 0.111$ and 0.115, respectively). Model 23.1c had issues with large groundfish F in 1996, though that is a known problem which would be fixed by increase the penalty on F deviations (Figure 24).

Recruitment Profiles

Profiles of R_0 and $\bar{R}$ were fairly noisy. The recruitment penalty and size composition data appear to be most informative for both EAG models. In contrast, index data appear to be more informative for R_0 and $\bar{R}$ in the WAG (Figure 25). The general shape of likelihood profiles for R_0 and $\bar{R}$ are odd, and should be revisited as data conflict in these models continues to be explored.

Author Recommendation

Three models should be brought forward for the final May 2025 assessment: 23.1c, 25.0b, and 25.1. Model 23.1c should replace the 2024 accepted model in the 2025 final assessment. There is negligible difference between models after the start of the data time series in 1981, and removing bias correction pre-1981 is correct given the lack of data. Relative to 23.1c, model 25.0b starts the model in a non-equilibrium state, uses equal emphasis factors ($\lambda = 1$) on all likelihood components, and bootstrap resampling to estimate stage 1 effective sample sizes for size composition data. Reducing emphasis on retained and total catch had minimal impact on model fit and establishes a level platform for further exploring model misspecification and data weighting. Using bootstrapping to estimate stage 1 effective sample sizes establishes size composition weights as proportional to sampling variances more so than does the number of vessel or observer days. Model 25.1 is evaluated for the EAG only and uses the cooperative survey as an additional fleet. The cooperative survey time series has been developing since 2015 and this analysis is the best use of survey data presented, thus far.

However, it may be prudent to troubleshoot data conflicts in EAG models prior to accepting a final model that uses the cooperative survey.

Literature Cited

Carvalho F, AE Punt, YJ Chang, MN Maunder, KR Piner. 2017. Can diagnostic tests help identify model misspecification in integrated stock assessments? Fisheries Research 192: 28-40.

Ichinokawa M, H Okamura, Y Takeuchi. 2014. Data conflict caused by model mis-specification of selectivity in an integrated stock assessment model and its potential effects on stock status estimation. Fisheries Research 158: 147-157.

Jackson 2024a. Aleutian Islands golden king crab stock assessment draft models. North Pacific Fishery Management Council, Anchorage, AK. January 2024 eAgenda Document

Jackson, T. 2024b. Aleutian Islands golden king crab stock assessment 2024. North Pacific Fishery Management Council, Anchorage, AK.

Maunder MN, KR Piner. 2015. Contemporary fisheries stock assessment: many issues still remain. ICES Journal of Marine Science 72: 7 - 18.

Maunder MN, KR Piner. 2017. Dealing with data conflicts in statistical inference of population assessment models that integrate information from multiple diverse data sets. Fisheries Research 192: 16 - 27.

Siddeek MSM, J Zheng, AE Punt, V Vanek. 2016. Estimation of size–transition matrices with and without molt probability for Alaska golden king crab using tag–recapture data. Fisheries Research 180: 161 - 168.

Siddeek MSM, B Daly, V.Vanek, J Zheng, C Siddon. 2022. Length-based approaches to estimating natural mortality using tagging and fisheries data: The example of the eastern Aleutian Islands, Alaska golden king crab (*Lithodes aequispinus*). Fisheries Research 251: 106304.

Siddeek MSM, B Daly, T Jackson. 2023. Aleutian Islands golden king crab stock assessment. North Pacific Fishery Management Council, Anchorage, AK.

Thorson JT, KF Johnson, RD Methot, IG Taylor. 2017. Model-based estimates of effective sample size in stock assessment models using the Dirichlet-multinomial distribution. Fisheries Research 192: 84 - 93.

Wang SP, MN Maunder, KR Piner, A Aires-da-Silva, HH Lee. 2014. Evaluation of virgin recruitment profiling as a diagnostic for selectivity curve structure in integrated stock assessment models. Fisheries Research 158: 158-164.

Tables

Table 1: Comparison of likelihood components for EAG models 23.1 using GMACS version 2.01.M.10 and 2.20.16, the updated annual timing structure, and the updated observer size data.

Component	v2.01.M.10	v2.20.16	v2.20.16 + season	v2.20.16 + data
catch	-472.858	-472.858	-472.858	-473.410
index	-41.816	-41.816	-41.816	-41.795
size	841.851	841.851	841.851	840.779
recruitment	19.447	19.447	19.447	19.485
tagging	2,694.969	2,694.969	2,694.969	2,694.986
penalites	0.150	0.150	0.150	0.150
priors	25.724	26.793	26.793	26.793
total	3,067.467	3,068.536	3,068.536	3,066.988

Table 2: Comparison of likelihood components for WAG models 23.1 using GMACS version 2.01.M.10 and 2.20.16, the updated annual timing structure, and the updated observer size data.

Component	v2.01.M.10	v2.20.16	v2.20.16 + season	v2.20.16 + data
catch	-434.541	-434.541	-434.541	-430.302
index	-66.820	-66.820	-66.820	-66.435
size	769.625	769.625	769.625	780.883
recruitment	21.844	21.844	21.844	21.466
tagging	2,698.615	2,698.615	2,698.615	2,698.487
penalites	0.059	0.059	0.059	0.059
priors	25.724	26.793	26.793	26.793
total	3,014.507	3,015.575	3,015.575	3,030.951

Table 3: Comparison of reference point estimates for EAG and WAG models 23.1 using GMACS version 2.01.M.10 and 2.20.16, and the updated annual timing structure.

Area	Model	MMB (t)	$B_{35\%}$ (t)	$\frac{MMB}{B_{35\%}}$	$\bar{R}_{1987-2017}$	$F_{35\%}$	F_{OFL}	OFL (t)
EAG	23.1 v2.01.M.10	7,551	6,905	1.09	2,781	0.55	0.55	2,825
EAG	23.1 v2.20.16	7,551	6,905	1.09	2,781	0.55	0.55	2,825
EAG	23.1 season	7,551	6,905	1.09	2,781	0.55	0.55	2,825
EAG	23.1 data	7,547	6,905	1.09	2,781	0.55	0.55	2,823
WAG	23.1 v2.01.M.10	3,837	4,638	0.83	1,866	0.54	0.44	900
WAG	23.1 v2.20.16	3,837	4,638	0.83	1,866	0.54	0.44	900
WAG	23.1 season	3,837	4,638	0.83	1,866	0.54	0.44	900
WAG	23.1 data	3,767	4,498	0.84	1,808	0.54	0.44	899

Table 4: Timeseries of directed fishery total catch (t) computed using the status quo (SQ) and the updated size composition data.

	EAG		WAG	
Crab Year	SQ	Update	SQ	Update
1990	3,521	3,521	2,695	2,695
1991	3,943	4,017	1,731	1,705
1992	5,054	5,118	1,289	1,201
1993			1,978	1,887
1994	3,974	3,974	5,191	5,197
1995	4,658	4,658	3,171	3,171
1996	3,207	3,208	2,290	2,291
1997	2,900	2,897	1,855	1,856
1998	2,949	2,951	1,590	1,590
1999	2,541	2,541	2,079	2,079
2000	2,592	2,592	2,313	2,314
2001	2,154	2,154	2,176	2,176
2002	1,871	1,871	1,889	1,889
2003	1,855	1,854	1,782	1,782
2004	1,671	1,670	1,839	1,839
2005	1,620	1,620	1,646	1,646
2006	1,617	1,617	1,400	1,400
2007	1,755	1,755	1,593	1,593
2008	1,774	1,774	1,697	1,696
2009	1,793	1,793	1,682	1,682
2010	1,702	1,702	1,602	1,602
2011	1,801	1,801	1,540	1,542
2012	1,946	1,946	1,778	1,778
2013	1,853	1,854	1,880	1,880
2014	1,965	1,965	1,584	1,584
2015	2,206	2,206	1,522	1,522
2016	2,214	2,214	1,493	1,493
2017	2,332	2,334	1,420	1,420
2018	2,778	2,778	1,639	1,639
2019	3,039	3,039	1,614	1,614
2020	2,604	2,605	1,763	1,763
2021	2,386	2,388	1,567	1,567
2022	2,078	2,078	1,122	1,122
2023	2,304	2,304	1,130	1,130

Table 5: Time series of number of crab measured (N) in retained catch sampling and bootstrap estimated effective sample size (N_{eff}). Total measure only includes crab > 100 mm carapace length.

	EAG				WAG			
		Bootstrap N_{eff}				Bootstrap N_{eff}		
Year	N	Min	Mean	Max	N	Min	Mean	Max
1985	4,853	368	1,616	10,780	798	71	295	2,100
1986	3,416	323	1,244	4,878	1,161	82	447	3,144
1987	2,001	157	712	2,555	397	14	136	735
1988	23,298	1,170	5,135	27,150	58,128	1,532	9,880	45,485
1989	67,350	1,846	13,262	61,292	98,269	3,526	24,802	170,250
1990	26,461	1,887	7,800	43,289	47,367	1,245	9,493	55,240
1991	30,238	1,572	7,560	40,935	45,973	2,360	10,450	66,042
1992	29,583	1,483	7,946	90,286	36,129	2,040	9,510	50,055
1993	10,841	611	2,781	12,937	9,818	671	2,741	22,910
1994	16,892	702	4,676	21,577	16,926	941	5,244	48,794
1995	25,974	1,604	6,814	23,368	3,646	198	1,021	5,656
1996	7,829	594	2,436	16,782	22,510	1,297	6,732	65,248
1997	13,340	879	4,381	26,052	29,437	1,783	8,469	45,414
1998	8,114	693	2,587	13,749	25,304	1,257	7,050	38,357
1999	6,497	273	1,833	9,481	22,371	1,869	6,980	87,756
2000	11,256	662	3,240	23,923	17,880	1,391	5,831	124,639
2001	4,892	391	1,640	11,917	19,297	1,045	5,423	26,293
2002	5,729	416	1,715	11,092	17,527	1,181	5,036	27,146
2003	5,220	310	1,510	7,065	12,331	844	3,552	18,505
2004	3,794	238	1,184	5,121	12,948	803	3,430	15,463
2005	3,083	242	1,047	6,098	11,983	753	3,401	13,166
2006	2,195	171	717	3,905	12,618	892	4,043	41,736
2007	3,255	269	1,089	6,375	9,669	921	3,512	17,535
2008	2,645	181	769	3,585	10,526	742	3,526	27,339
2009	2,355	190	803	3,076	9,790	1,000	3,497	20,368
2010	2,353	195	817	4,831	9,818	1,103	3,590	17,716
2011	2,507	111	767	4,237	10,639	748	3,499	22,443
2012	2,926	195	943	6,069	6,542	512	1,989	7,599
2013	2,560	219	886	4,169	2,408	213	749	3,481
2014	2,175	233	898	6,778	2,929	208	964	5,497
2015	2,298	206	851	4,046	2,759	152	932	7,165
2016	2,697	240	1,005	3,927	2,552	208	778	3,535
2017	2,636	215	896	5,651	2,313	173	738	5,290
2018	2,717	227	968	6,965	2,076	184	705	4,806
2019	2,969	282	999	4,522	2,618	232	923	5,661
2020	2,914	281	1,104	7,268	2,831	188	892	4,021
2021	2,725	245	970	14,399	2,512	222	909	4,433
2022	2,138	180	703	2,506	1,813	117	643	3,771
2023	2,719	233	1,012	5,625	1,811	152	677	4,826

Table 6: Time series of number of crab measured (N) in observer size composition sampling and bootstrap estimated effective sample size (N_{eff}).

	EAG				WAG			
		Bootstrap N_{eff}				Bootstrap N_{eff}		
Year	N	Min	Mean	Max	N	Min	Mean	Max
1990	2,600	58	604	2,521	4,241	337	1,823	7,716
1991	5,807	627	2,353	6,979	12,227	574	4,222	14,342
1992	5,834	504	2,005	8,203	10,347	1,089	4,209	15,109
1993					2,451	238	980	3,590
1994	1,235	154	584	1,744	24,024	1,015	7,674	21,943
1995	96,369	7,019	37,208	181,699	72,709	5,766	24,343	90,109
1996	104,843	12,229	45,370	157,744	91,843	8,059	33,243	104,754
1997	77,173	9,800	30,922	97,069	54,607	4,247	19,650	96,531
1998	83,328	6,622	28,481	112,682	40,878	2,227	13,584	58,762
1999	73,728	7,260	27,495	121,430	61,180	7,224	23,695	77,031
2000	28,334	2,204	11,991	37,418	72,459	5,290	22,607	125,901
2001	35,606	4,270	15,282	87,783	63,710	3,600	21,779	97,374
2002	24,536	2,558	11,458	42,044	41,830	2,608	14,224	72,468
2003	22,859	1,956	10,510	32,126	38,569	3,266	13,858	40,488
2004	19,481	1,952	8,059	24,837	34,824	1,865	11,158	34,536
2005	12,451	1,916	5,227	18,173	24,111	1,589	9,005	55,666
2006	9,463	1,089	4,379	13,694	27,008	2,635	11,776	38,936
2007	12,448	1,613	5,361	12,878	26,643	2,699	9,252	44,526
2008	15,715	2,199	7,582	22,068	25,191	2,458	10,787	67,840
2009	13,972	1,547	5,317	13,524	30,106	1,715	8,776	26,582
2010	15,283	1,780	6,146	15,697	24,575	2,895	9,604	47,126
2011	18,994	2,170	7,368	23,272	26,054	1,735	9,260	40,707
2012	20,648	1,730	8,122	26,561	32,869	1,120	11,559	31,291
2013	23,800	2,219	10,249	25,101	29,736	1,860	11,203	35,729
2014	22,365	1,801	8,438	36,607	25,491	1,000	8,254	43,203
2015	30,904	2,849	10,523	40,661	27,855	2,651	12,064	70,118
2016	33,943	2,369	12,335	32,803	24,156	1,111	8,141	37,390
2017	34,151	3,074	13,905	47,659	20,084	2,130	6,858	20,581
2018	32,665	1,837	13,570	51,419	22,382	1,938	8,200	34,085
2019	33,199	3,576	13,598	54,796	21,562	1,432	7,264	28,097
2020	28,824	2,640	11,974	54,644	26,734	1,964	10,157	35,811
2021	20,196	1,185	8,262	41,869	18,776	1,762	7,401	24,433
2022	17,805	1,704	8,530	27,651	17,484	1,068	6,345	20,000
2023	19,893	1,356	7,487	26,356	14,649	1,445	5,446	21,973

Table 7: Estimated additional CV by index series and model in the EAG and WAG.

		23.1	23.1c	25.0	25.0a	25.0b	25.0b2	25.0c	25.0d	25.1	25.1b
EAG	Pre-Rat. Observer CPUE	0.26	0.26	0.26	0.24	0.22	0.17	0.21	0.24	0.22	0.21
EAG	Post-Rat. Observer CPUE	0.23	0.23	0.23	0.22	0.24	0.00	0.31	0.27	0.32	0.36
EAG	Early Fish Ticket CPUE	0.19	0.19	0.19	0.20	0.21	0.07	0.26	0.24	0.22	0.26
WAG	Pre-Rat. Observer CPUE	0.24	0.24	0.24	0.24	0.20	0.15	0.18	0.20		
WAG	Post-Rat. Observer CPUE	0.07	0.07	0.07	0.07	0.07	0.04	0.10	0.10		
WAG	Early Fish Ticket CPUE	0.17	0.16	0.17	0.17	0.15	0.09	0.20	0.20		

Table 8: Comparison of biological reference points for EAG models.

Model	MMB (t)	$B_{35\%}$ (t)	$\frac{MMB}{B_{35\%}}$	$\bar{R}_{1987-2020}$	$F_{35\%}$	F_{OFL}	OFL (t)
23.1	7,547	6,905	1.09	2,781	0.55	0.55	2,823
23.1c	7,539	6,904	1.09	2,781	0.55	0.55	2,822
25.0	8,058	6,939	1.16	2,789	0.54	0.54	2,973
25.0a	8,053	6,908	1.17	2,775	0.55	0.55	2,970
25.0b	7,464	6,846	1.09	2,743	0.55	0.55	2,755
25.0b2	6,324	6,439	0.98	2,573	0.51	0.50	2,096
25.0c	5,140	6,633	0.77	2,662	0.59	0.44	1,345
25.0d	7,311	6,846	1.07	2,752	0.58	0.58	2,710
25.1	4,754	6,547	0.73	2,619	0.55	0.38	1,036
25.1b	4,158	6,528	0.64	2,615	0.58	0.35	774

Model	MMB (mil lb)	$B_{35\%}$ (mil lb)	$\frac{MMB}{B_{35\%}}$	$\bar{R}_{1987-2020}$	$F_{35\%}$	F_{OFL}	OFL (mil lb)
23.1	16.64	15.22	1.09	2,781	0.55	0.55	6.22
23.1c	16.62	15.22	1.09	2,781	0.55	0.55	6.22
25.0	17.76	15.30	1.16	2,789	0.54	0.54	6.55
25.0a	17.75	15.23	1.17	2,775	0.55	0.55	6.55
25.0b	16.46	15.09	1.09	2,743	0.55	0.55	6.07
25.0b2	13.94	14.20	0.98	2,573	0.51	0.50	4.62
25.0c	11.33	14.62	0.77	2,662	0.59	0.44	2.97
25.0d	16.12	15.09	1.07	2,752	0.58	0.58	5.97
25.1	10.48	14.43	0.73	2,619	0.55	0.38	2.28
25.1b	9.17	14.39	0.64	2,615	0.58	0.35	1.71

Table 9: Comparison of biological reference points for WAG models.

Model	MMB (t)	$B_{35\%}$ (t)	$\frac{MMB}{B_{35\%}}$	$\bar{R}_{1987-2020}$	$F_{35\%}$	F_{OFL}	OFL (t)
23.1	3,767	4,498	0.84	1,808	0.54	0.44	899
23.1c	3,757	4,494	0.84	1,807	0.55	0.45	894
25.0	3,762	4,491	0.84	1,803	0.54	0.45	892
25.0a	3,762	4,491	0.84	1,803	0.54	0.45	892
25.0b	3,705	4,504	0.82	1,800	0.54	0.44	872
25.0b2	3,493	4,417	0.79	1,757	0.54	0.41	754
25.0c	3,590	4,540	0.79	1,795	0.54	0.42	784
25.0d	3,722	4,522	0.82	1,793	0.54	0.43	849

Model	MMB (mil lb)	$B_{35\%}$ (mil lb)	$\frac{MMB}{B_{35\%}}$	$\bar{R}_{1987-2020}$	$F_{35\%}$	F_{OFL}	OFL (mil lb)
23.1	8.31	9.92	0.84	1,808	0.54	0.44	1.98
23.1c	8.28	9.91	0.84	1,807	0.55	0.45	1.97
25.0	8.29	9.90	0.84	1,803	0.54	0.45	1.97
25.0a	8.29	9.90	0.84	1,803	0.54	0.45	1.97
25.0b	8.17	9.93	0.82	1,800	0.54	0.44	1.92
25.0b2	7.70	9.74	0.79	1,757	0.54	0.41	1.66
25.0c	7.91	10.01	0.79	1,795	0.54	0.42	1.73
25.0d	8.21	9.97	0.82	1,793	0.54	0.43	1.87

Figures

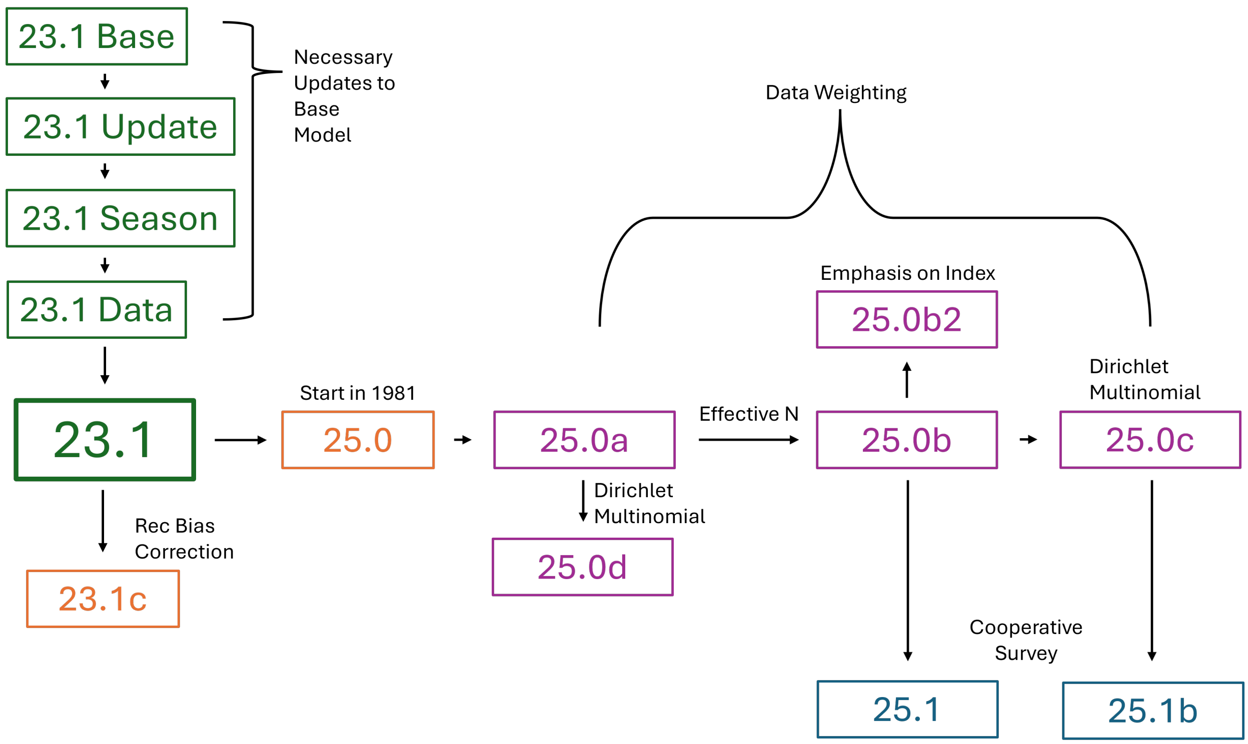

Figure 2: Flow chart of models evaluated in this analysis. Green indicates updates to the base model, orange evaluates initial conditions, purple explores data weighting, and blue adds the cooperative survey.

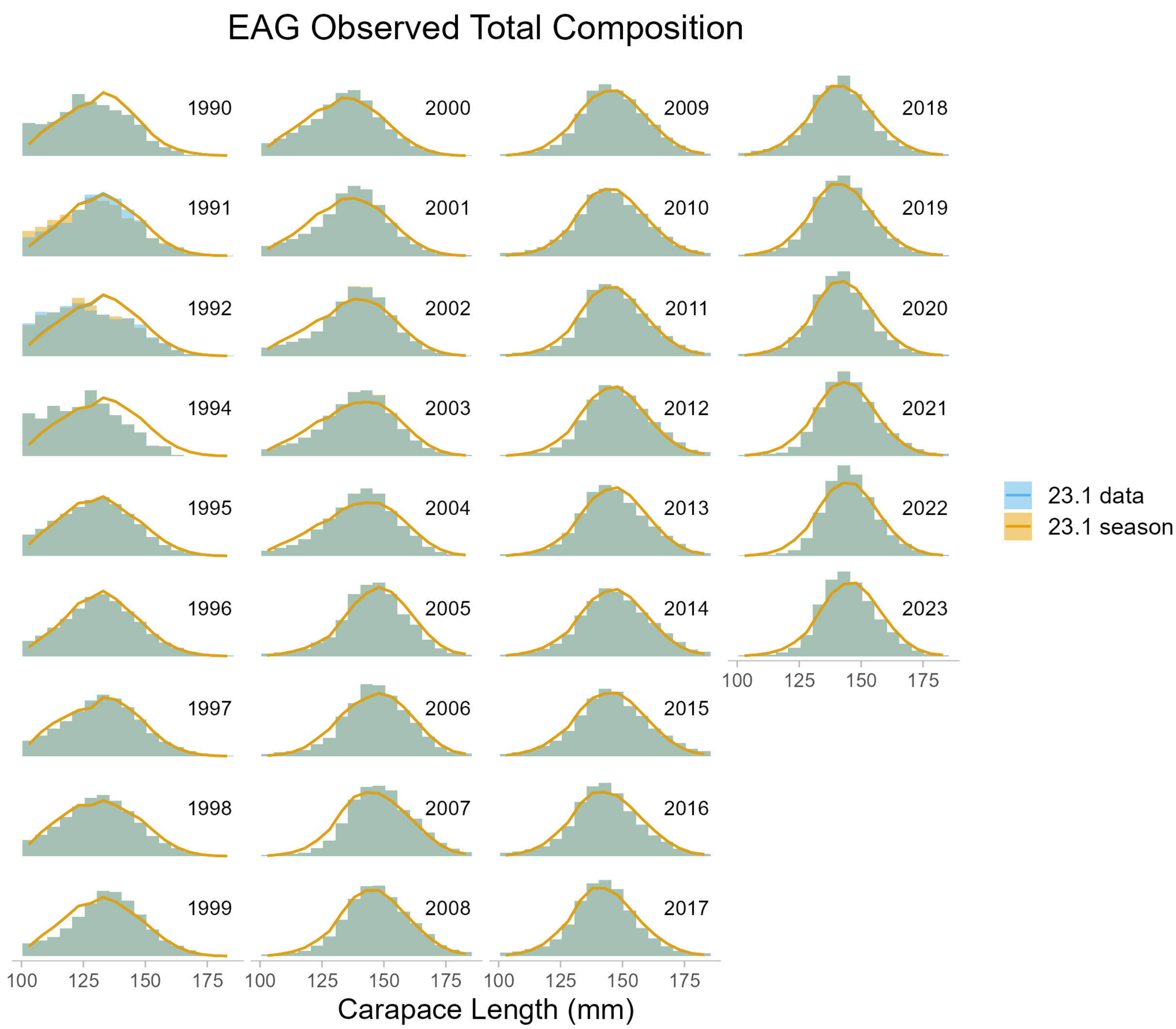

Figure 3: Fit to EAG total size composition data using the previous subset of gear types (Jackson 2024b; Siddeek et al. 2023) and the revised subset of gear types (explained above).

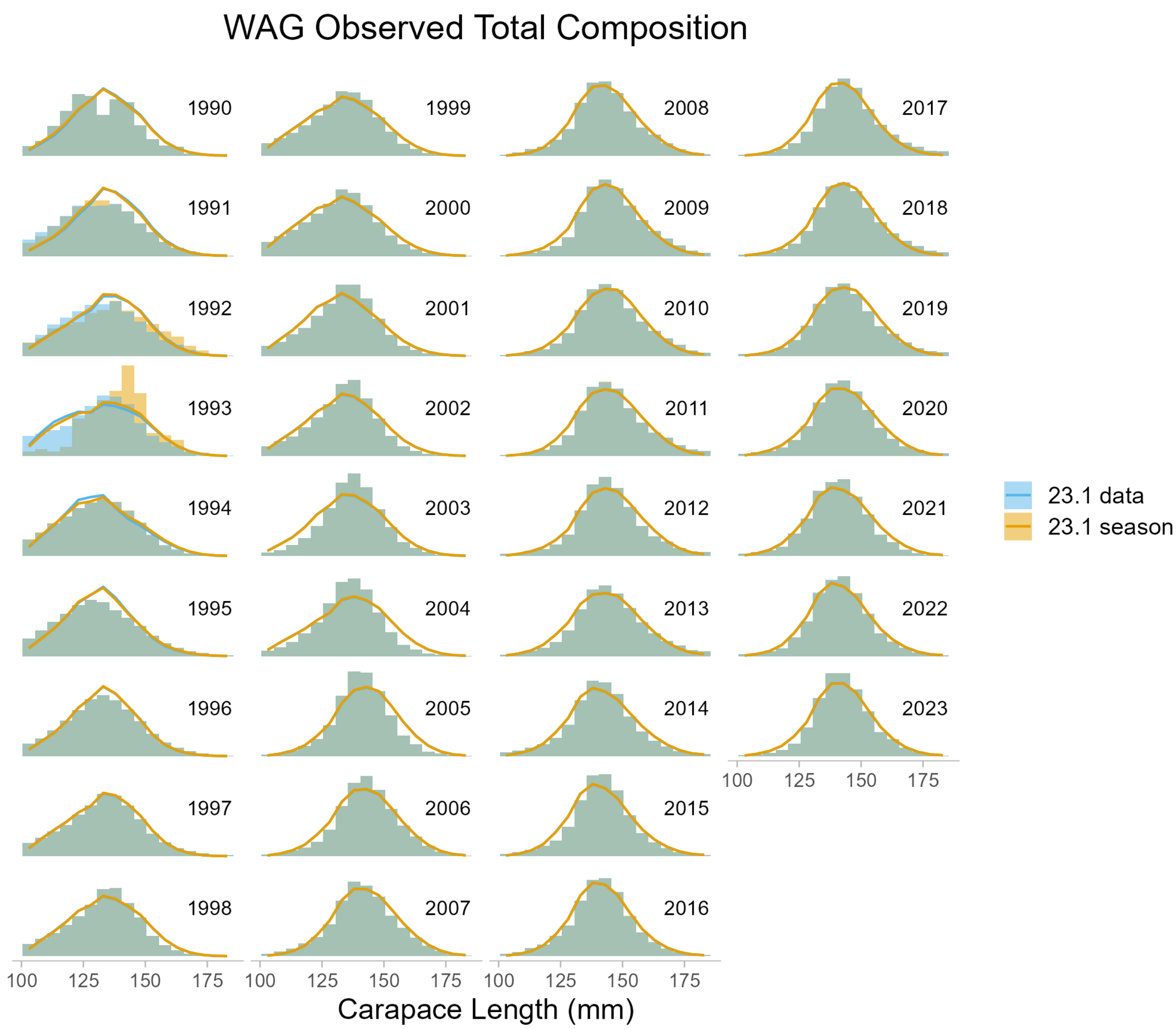

Figure 4: Fit to WAG total size composition data using the previous subset of gear types (Jackson 2024b; Siddeek et al. 2023) and the revised subset of gear types (explained above).

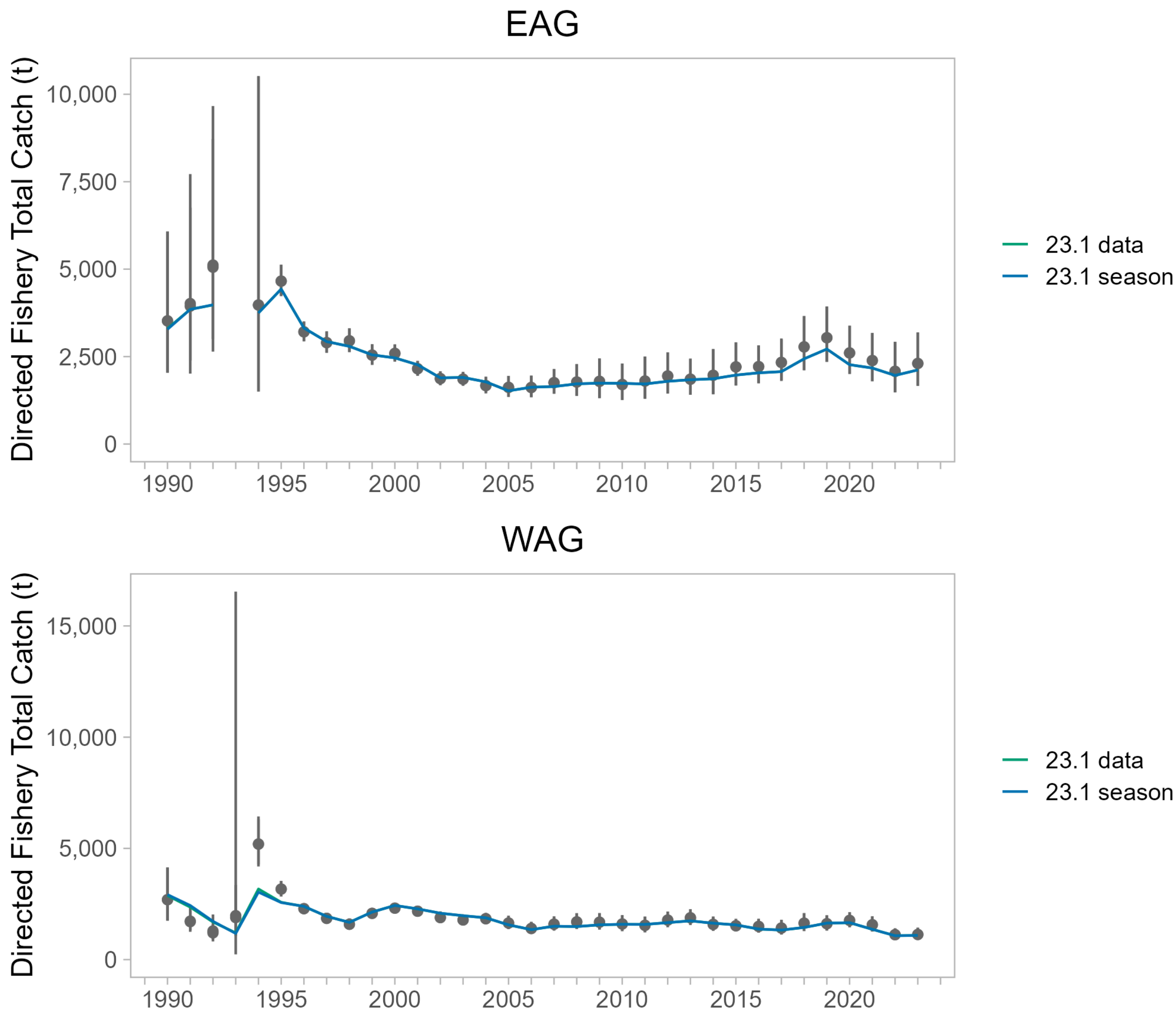

Figure 5: Fit to total catch data using the previous subset of gear types (Jackson 2024b; Siddeek et al. 2023) and the revised subset of gear types (explained above).

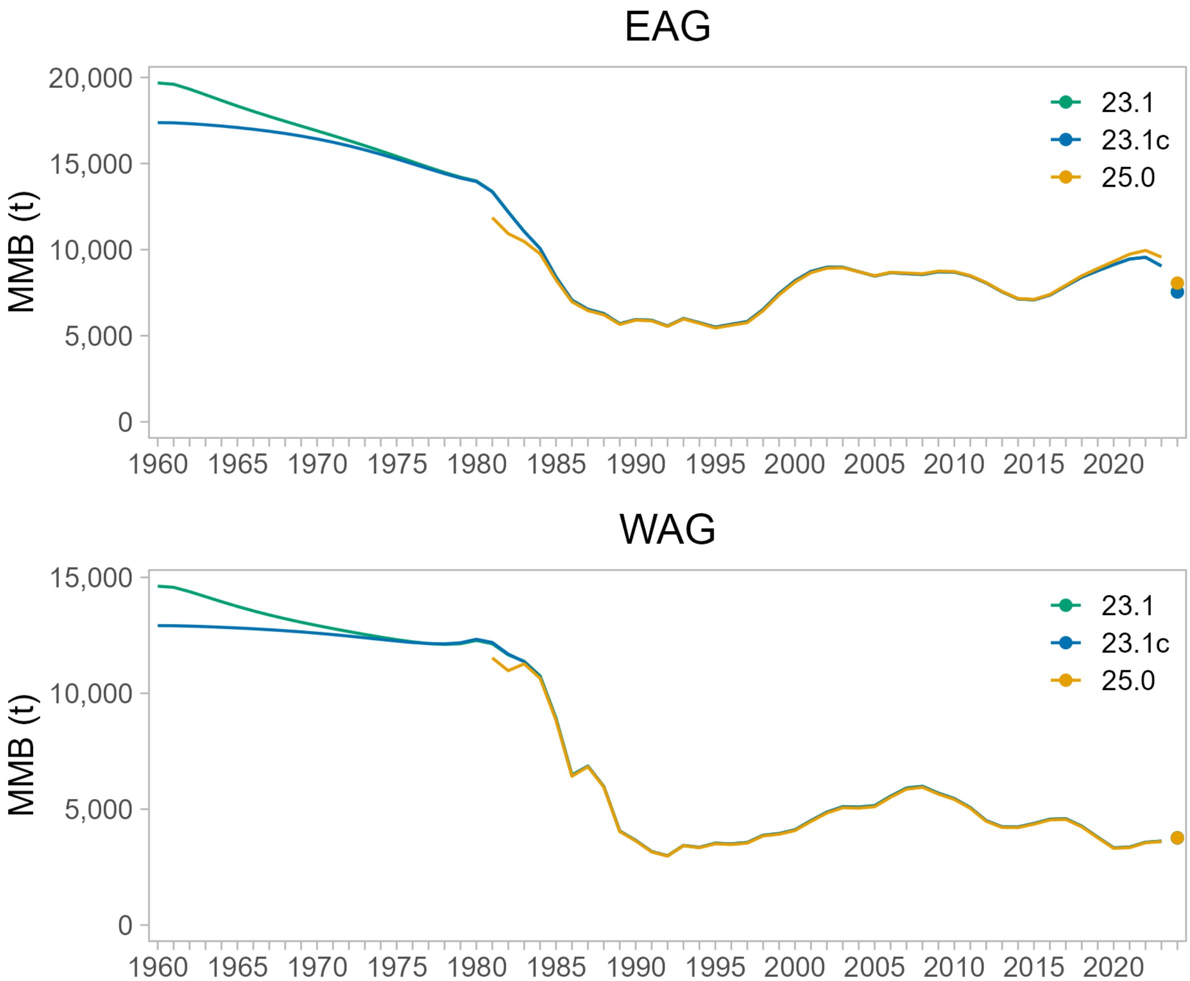

Figure 6: MMB trajectory for models 23.1, 23.1c, and 25.0. Model 23.1c modifies recruitment bias correction during 1960 - 2023, and model 25.0 starts the model in 1981 in non-equilibrium conditions.

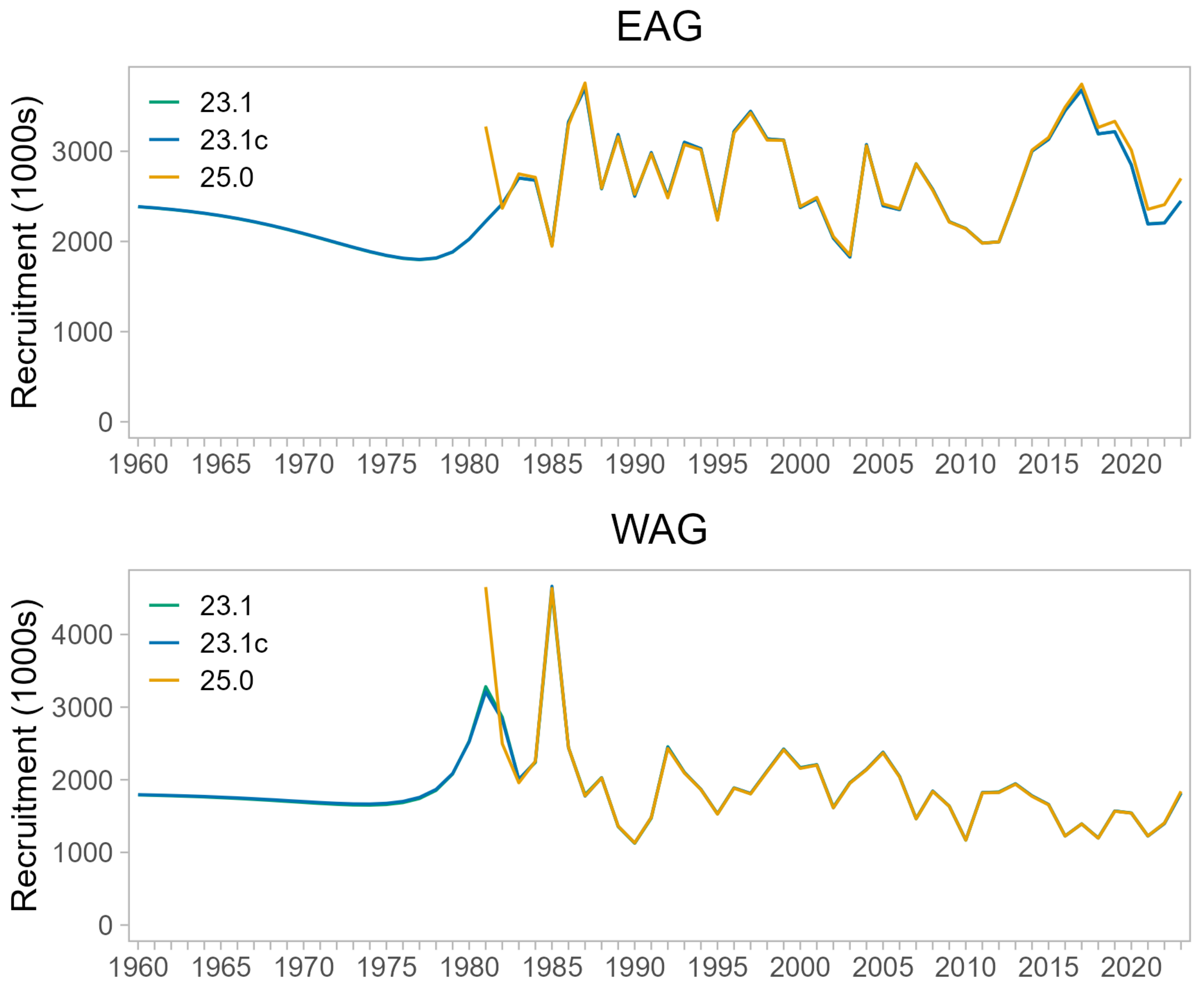

Figure 7: Recruitment for models 23.1, 23.1c, and 25.0. Model 23.1c modifies recruitment bias correction during 1960 - 2023, and model 25.0 starts the model in 1981 in non-equilibrium conditions.

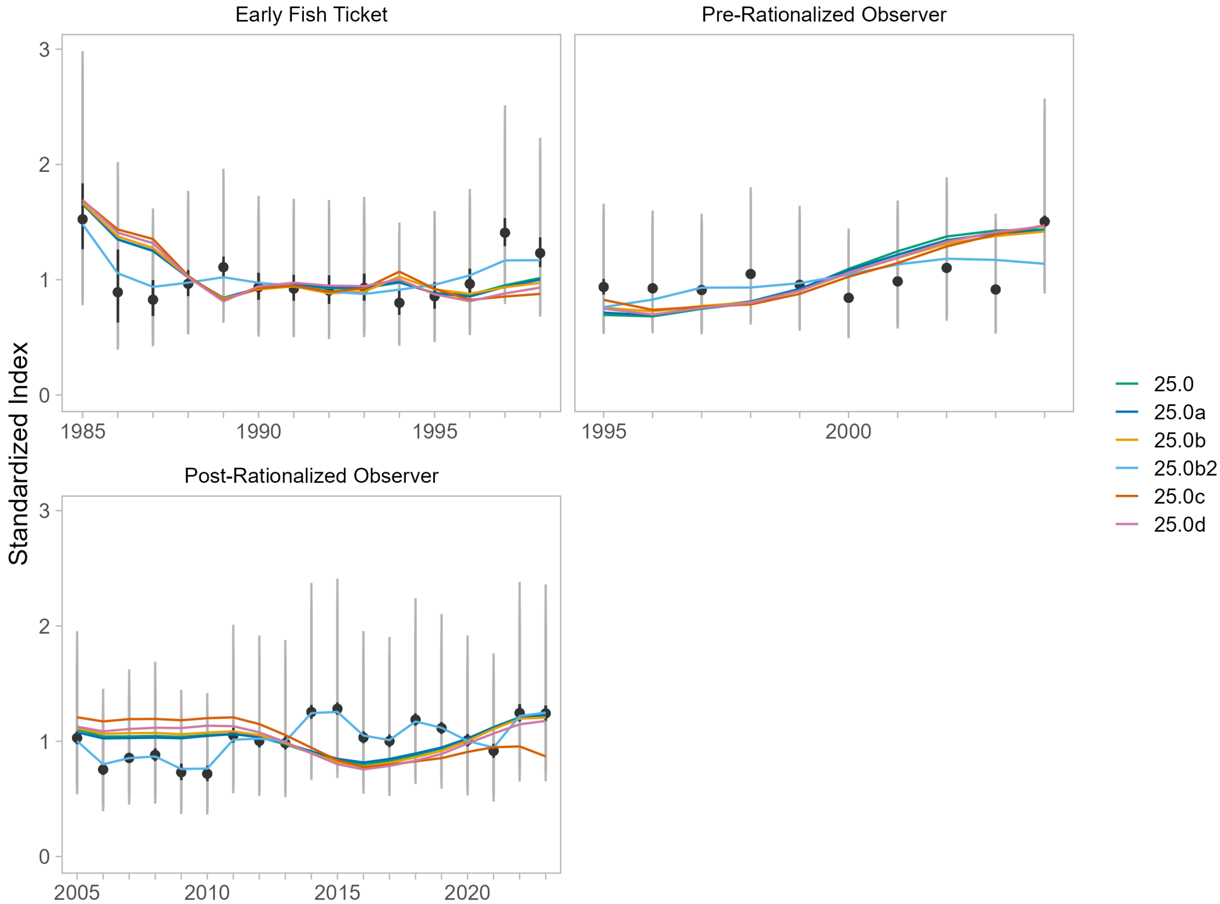

Figure 8: EAG fits to index data for models that explored data weighting. Models 25.0 - 25.0b tune size composition weights using Francis weighting and models 25.0c and 25.0d use the Dirichlet multinomial. Black bars indicate observed 95% confidence intervals, grey bars indicate estimated additional error for which model was greatest for each index. Estimated added CV by model is in Table 7.

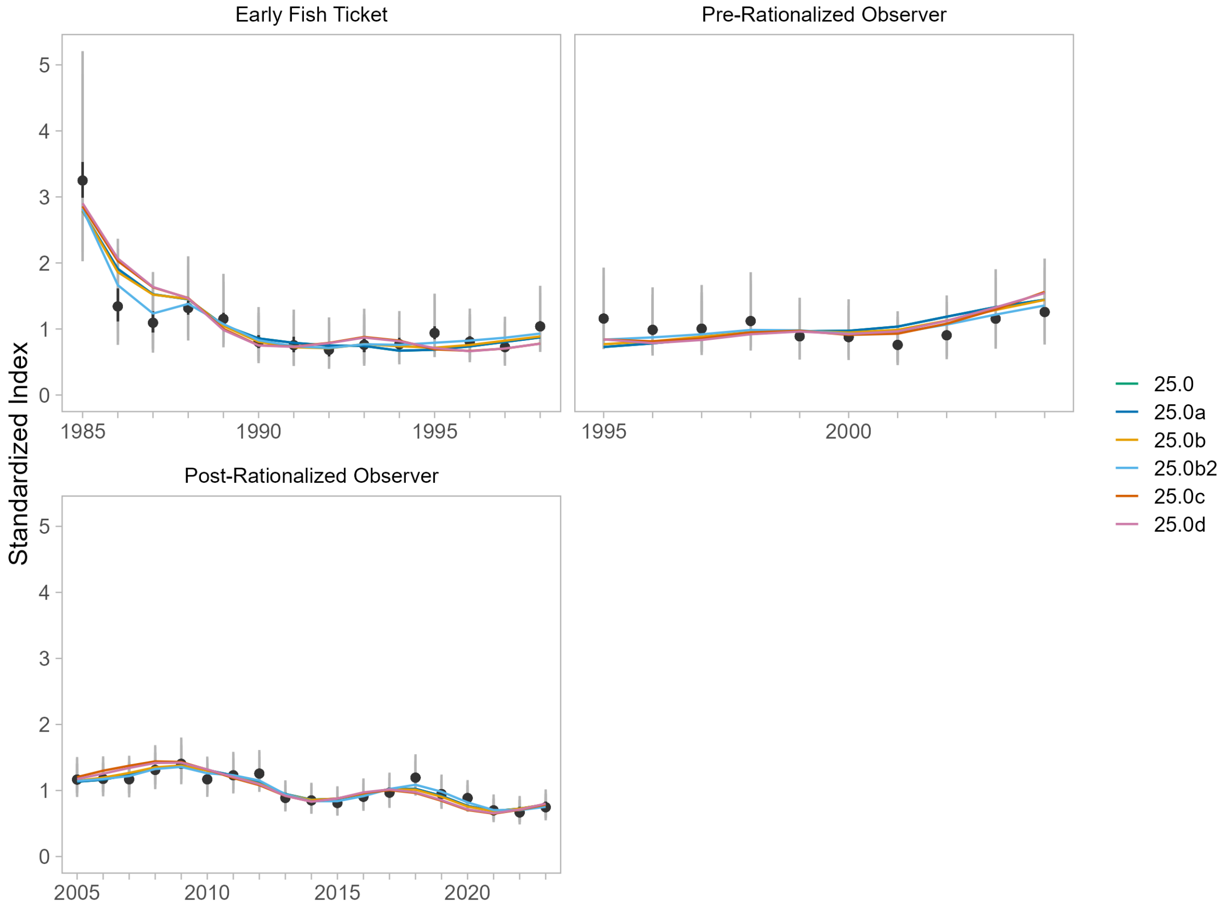

Figure 9: WAG fits to index data for models that explored data weighting. Models 25.0 - 25.0b tune size composition weights using Francis weighting and models 25.0c and 25.0d use the Dirichlet multinomial. Black bars indicate observed 95% confidence intervals, grey bars indicate estimated additional error for which model was greatest for each index. Estimated added CV by model is in Table 7.

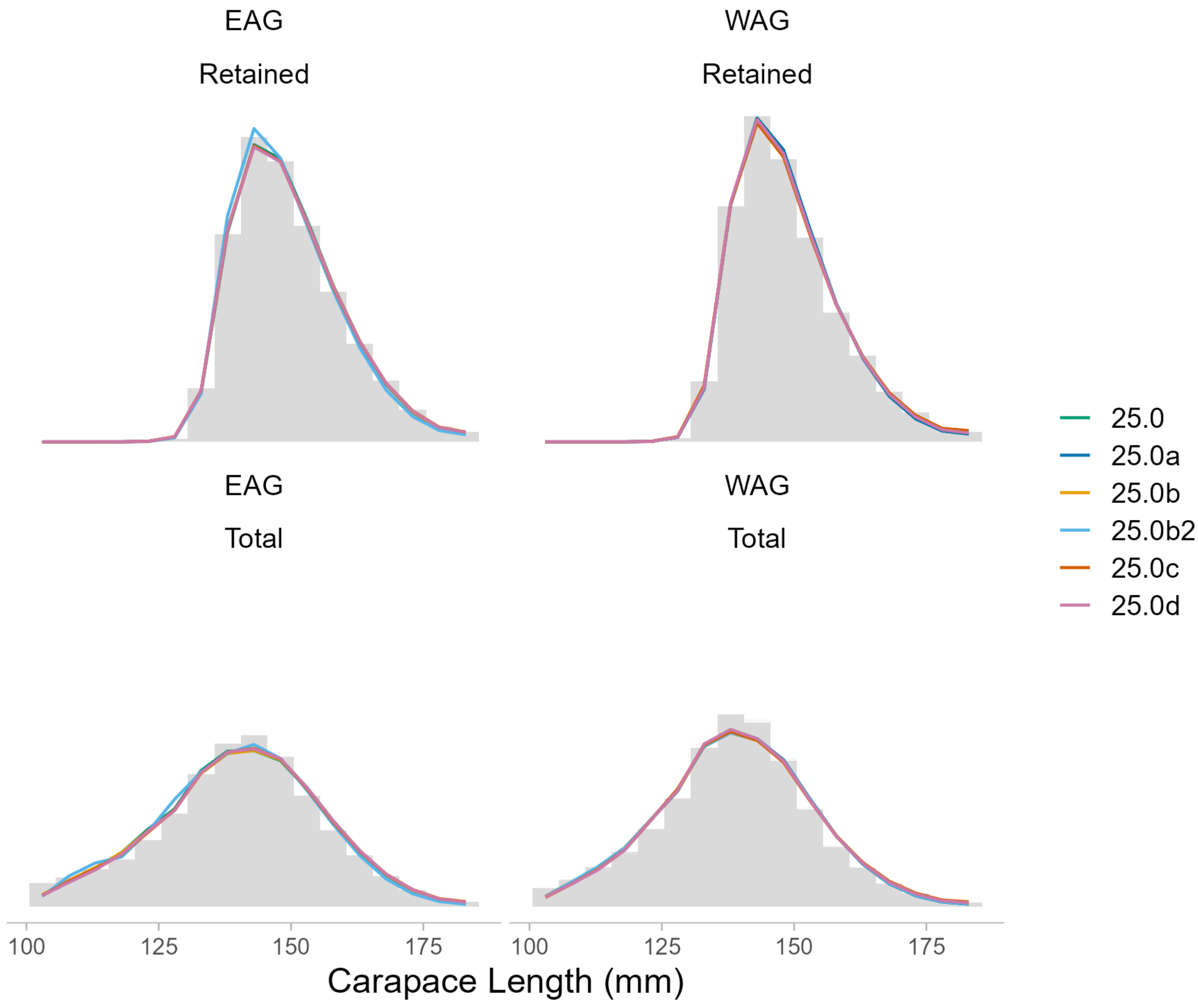

Figure 10: Fits to size composition data in aggregate for models that explored data weighting. Models 25.0 - 25.0b tune size composition weights using Francis weighting and models 25.0c and 25.0d use the Dirichlet multinomial.

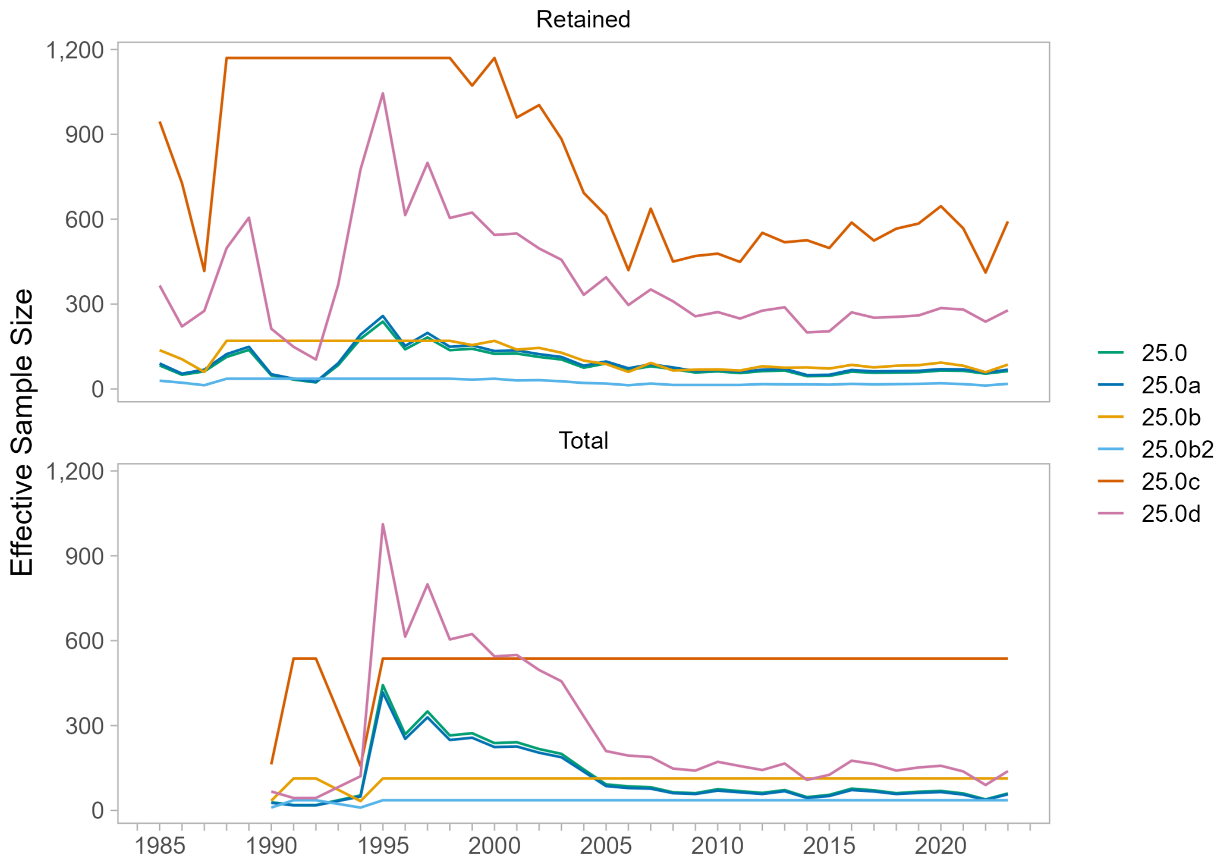

Figure 11: Stage two effective sample size for retained and total size composition data in the EAG. Models 25.0, 25.0a, 25.0b, and 25.0b2 use a multinomial likelihood with sample sizes tuned using Francis weighting and models 25.0c and 25.0d use a Dirichlet multinomial likelihood.

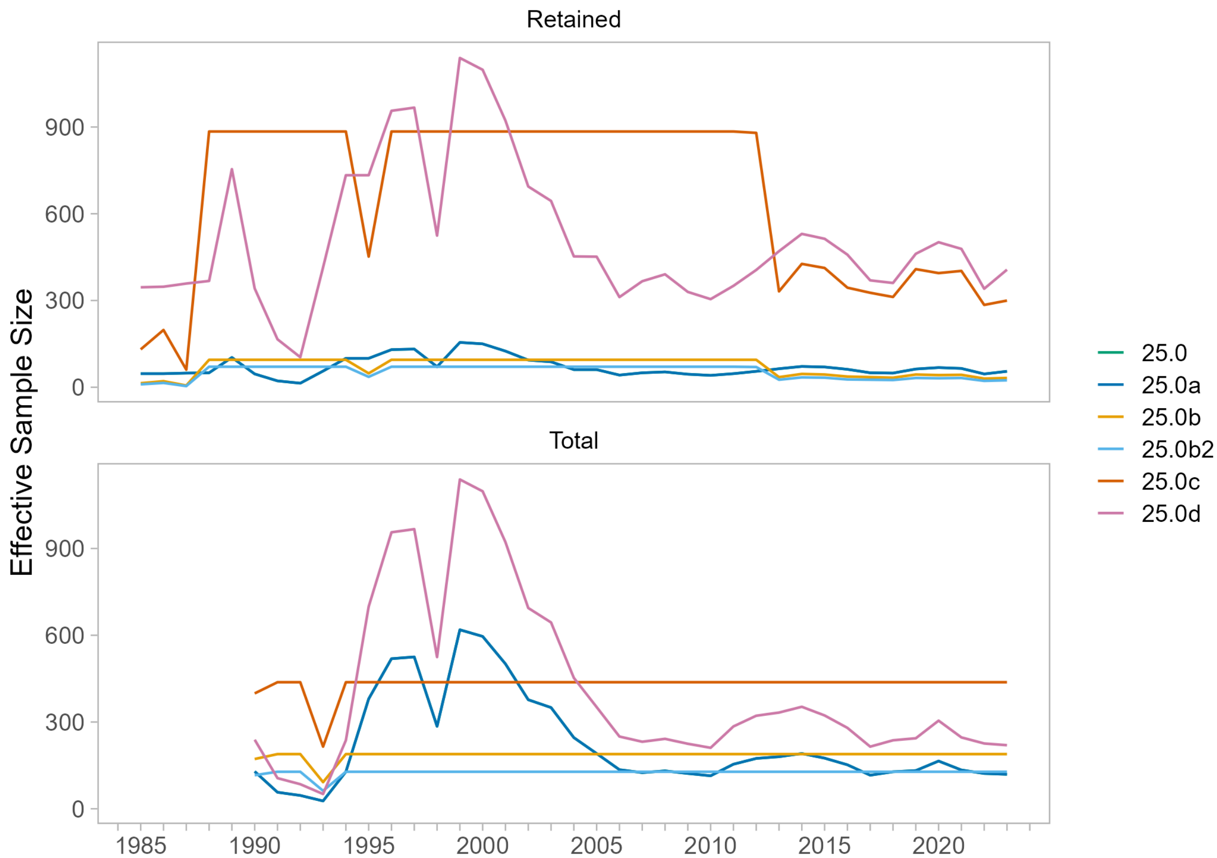

Figure 12: Stage two effective sample size for retained and total size composition data in the WAG. Models 25.0, 25.0a, 25.0b, and 25.0b2 use a multinomial likelihood with sample sizes tuned using Francis weighting and models 25.0c and 25.0d use a Dirichlet multinomial likelihood.

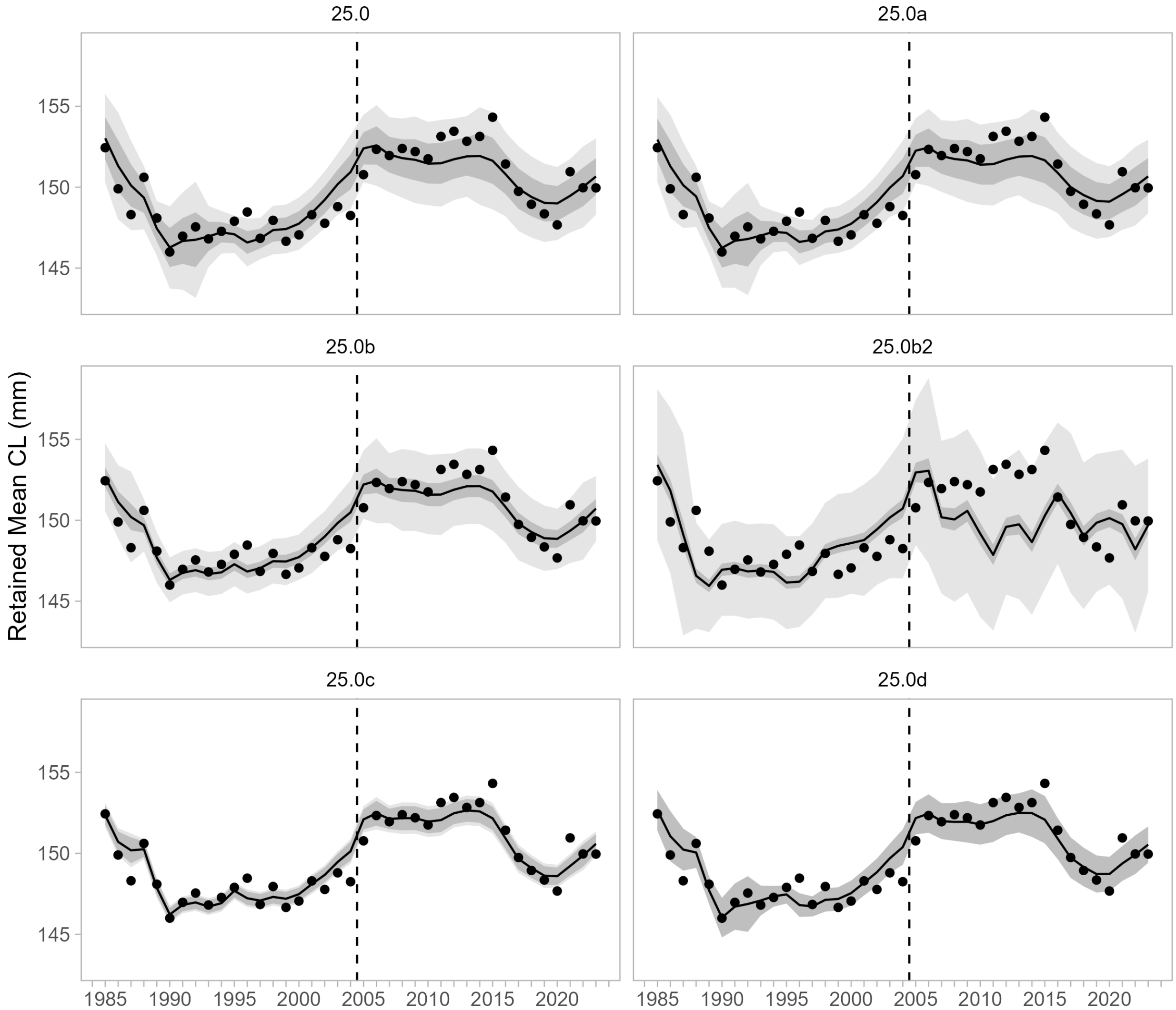

Figure 13: EAG observed (dots) and predicted (lines) mean carapace length (mm) of the retained size composiion. Grey bands indicate 95% confidence interval assuming stage one (darker) and stage 2 (lighter) effective sample sizes.

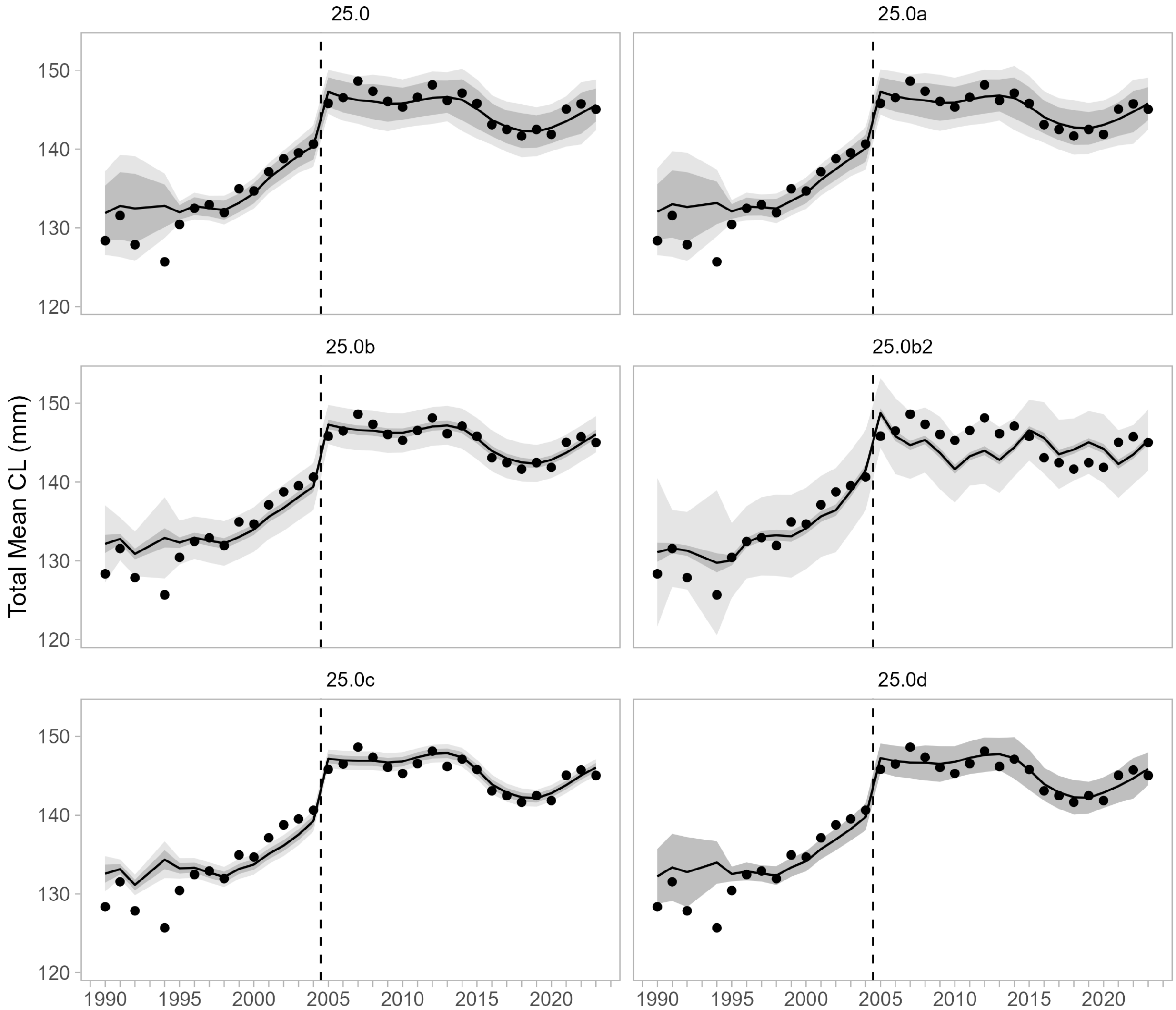

Figure 14: EAG observed (dots) and predicted (lines) mean carapace length (mm) of the total size composiion. Grey bands indicate 95% confidence interval assuming stage one (darker) and stage 2 (lighter) effective sample sizes.

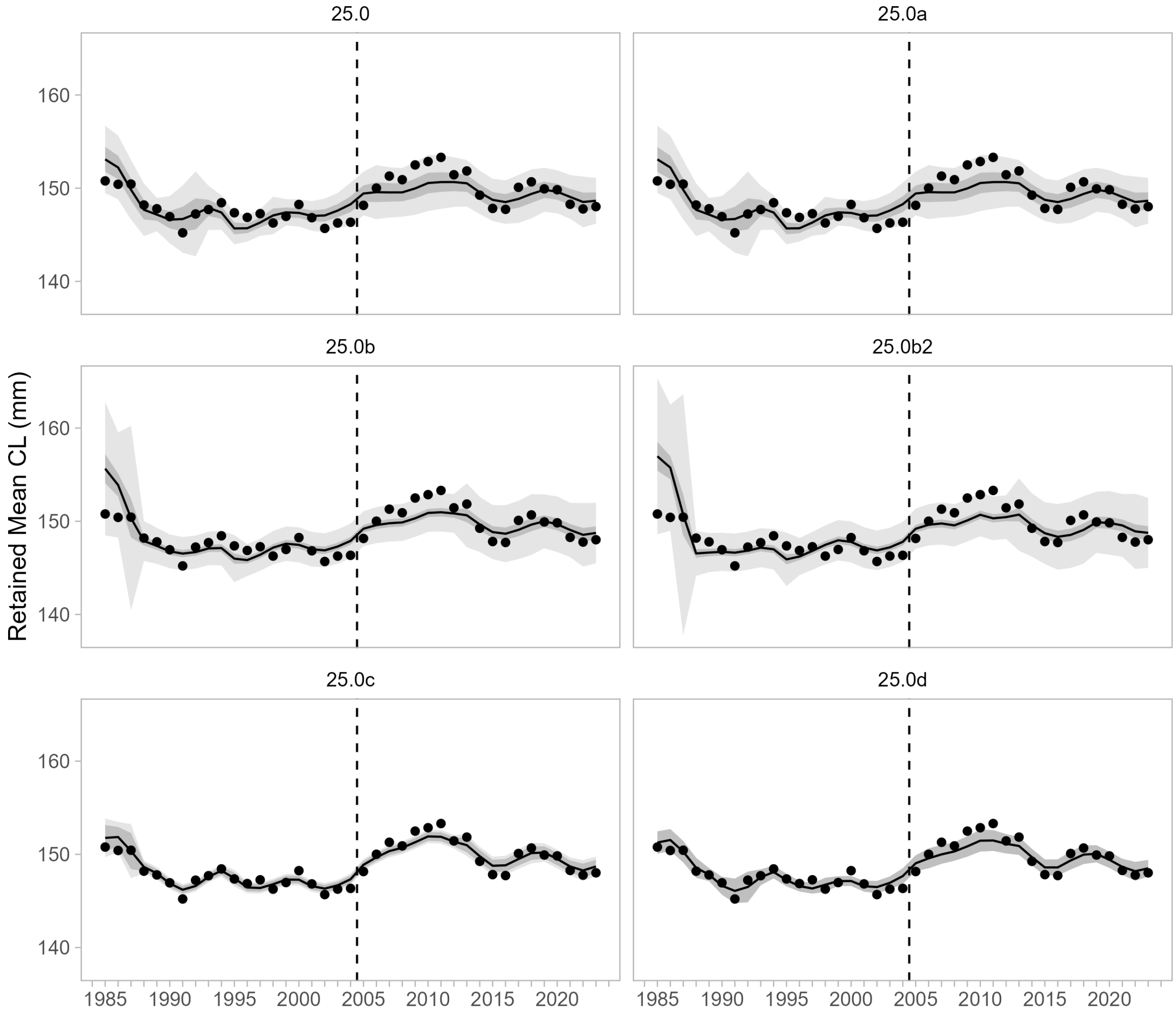

Figure 15: WAG observed (dots) and predicted (lines) mean carapace length (mm) of the retained size composiion. Grey bands indicate 95% confidence interval assuming stage one (darker) and stage 2 (lighter) effective sample sizes.

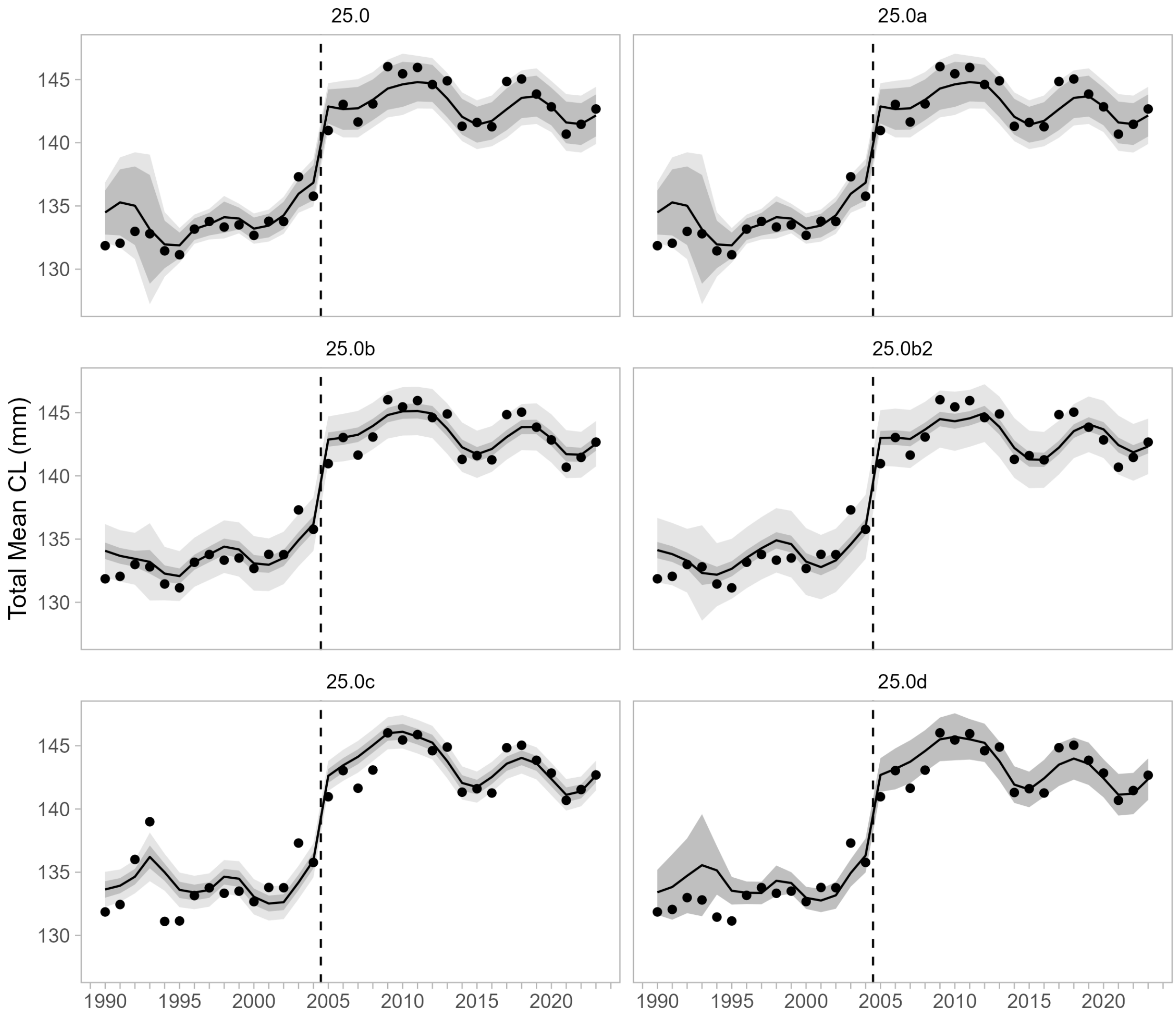

Figure 16: WAG observed (dots) and predicted (lines) mean carapace length (mm) of the total size composiion. Grey bands indicate 95% confidence interval assuming stage one (darker) and stage 2 (lighter) effective sample sizes.

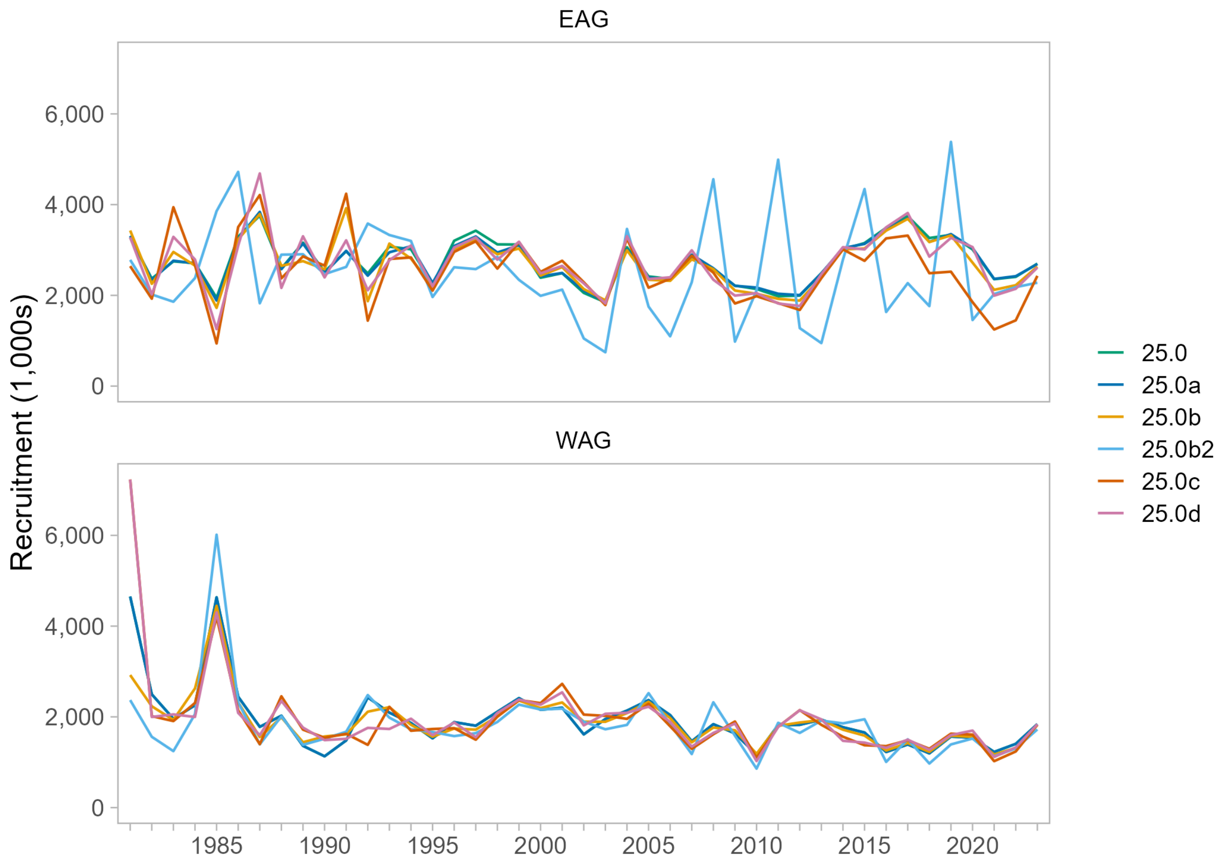

Figure 17: Recruitment trajectory for models that explored data weighting. Models 25.0 - 25.0b tune size composition weights using Francis weighting and models 25.0c and 25.0d use the Dirichlet multinomial.

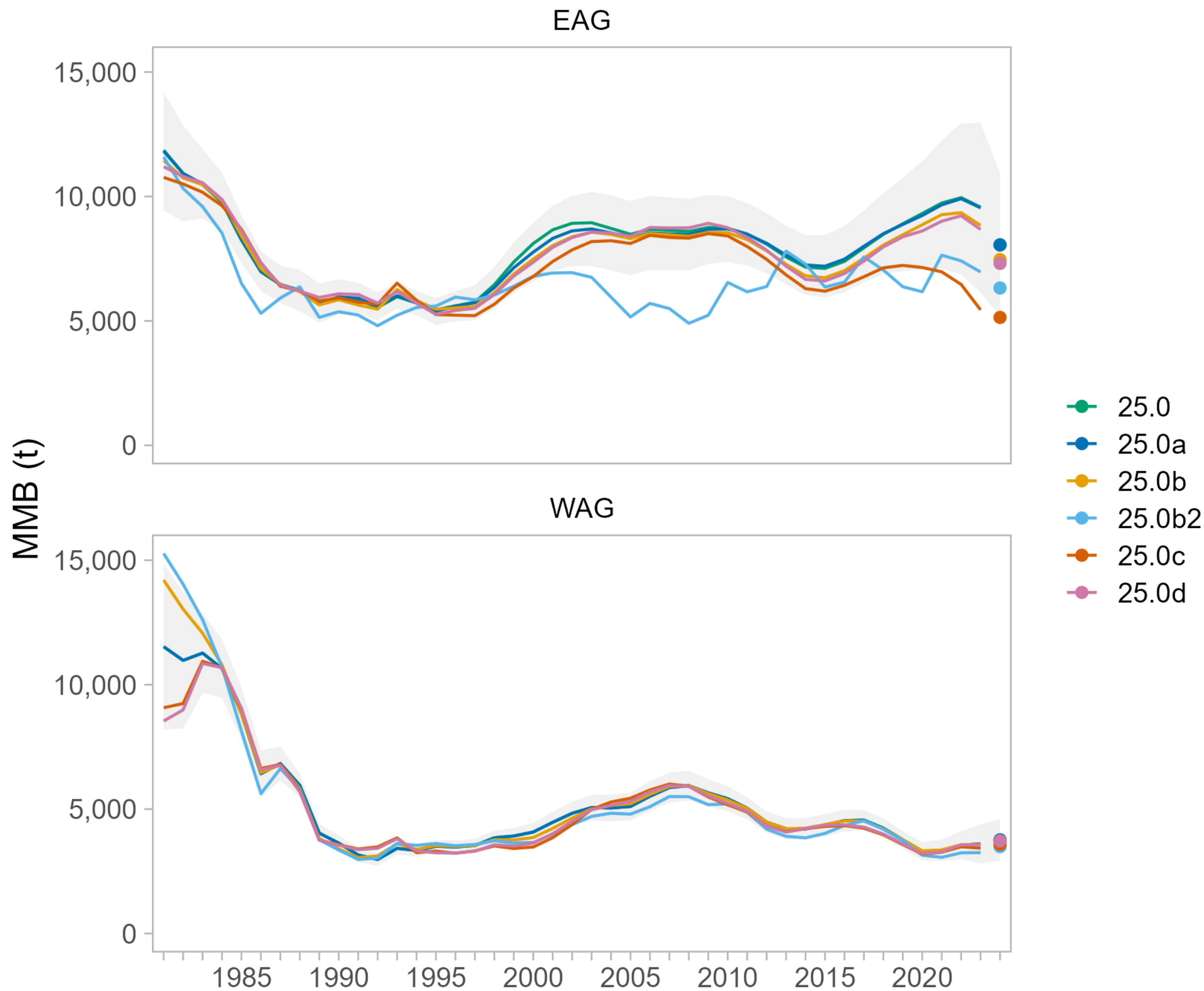

Figure 18: Mature male biomass trajectory for models that explored data weighting. Error bars represents a 95% confidence interval for model 25.0a. Models 25.0 - 25.0b tune size composition weights using Francis weighting and models 25.0c and 25.0d use the Dirichlet multinomial.

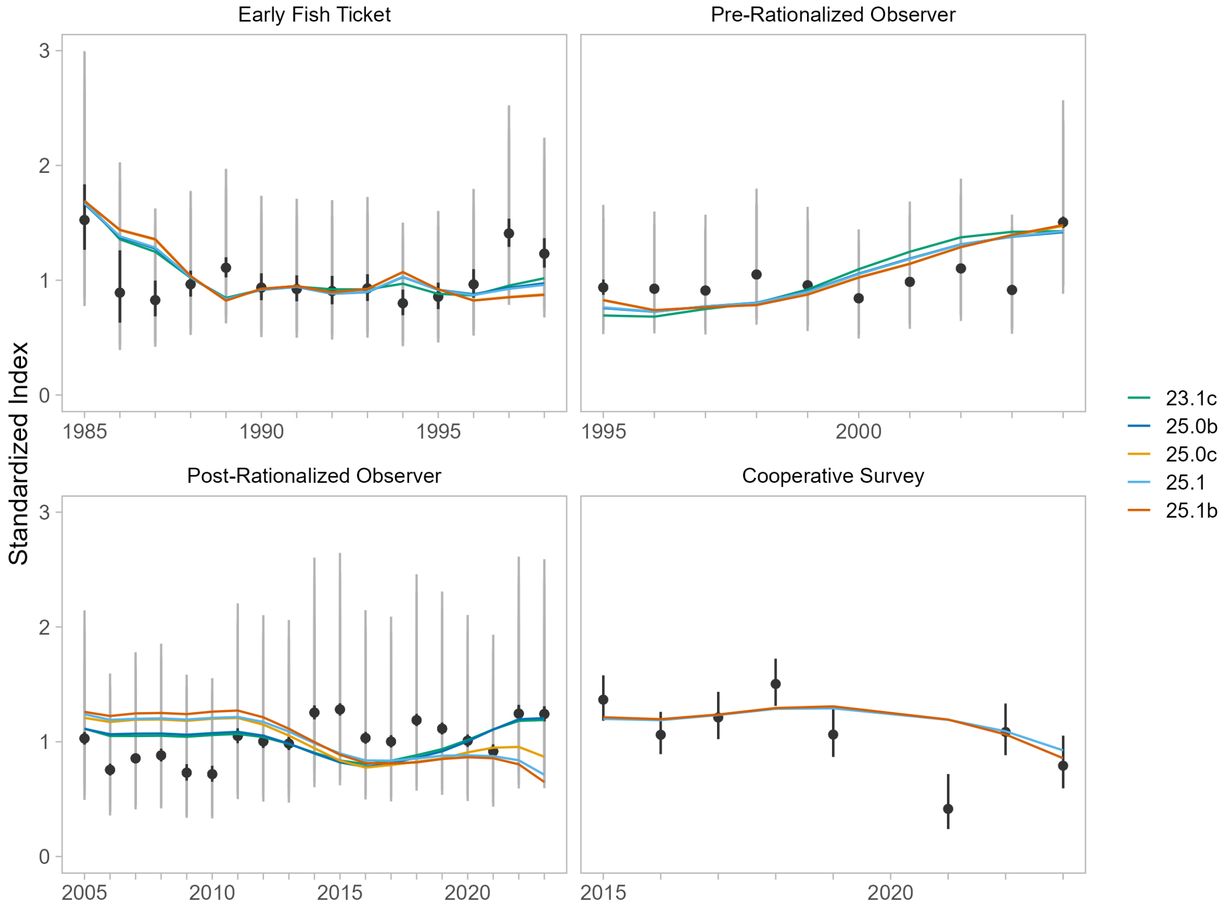

Figure 19: Fits to index data among EAG models. Black bars indicate observed 95% confidence intervals, grey bars indicate estimated additional error for which model was greatest for each index. Estimated added CV by model is in Table 7. Models 25.1 and 25.1b incorporate the cooperative survey.

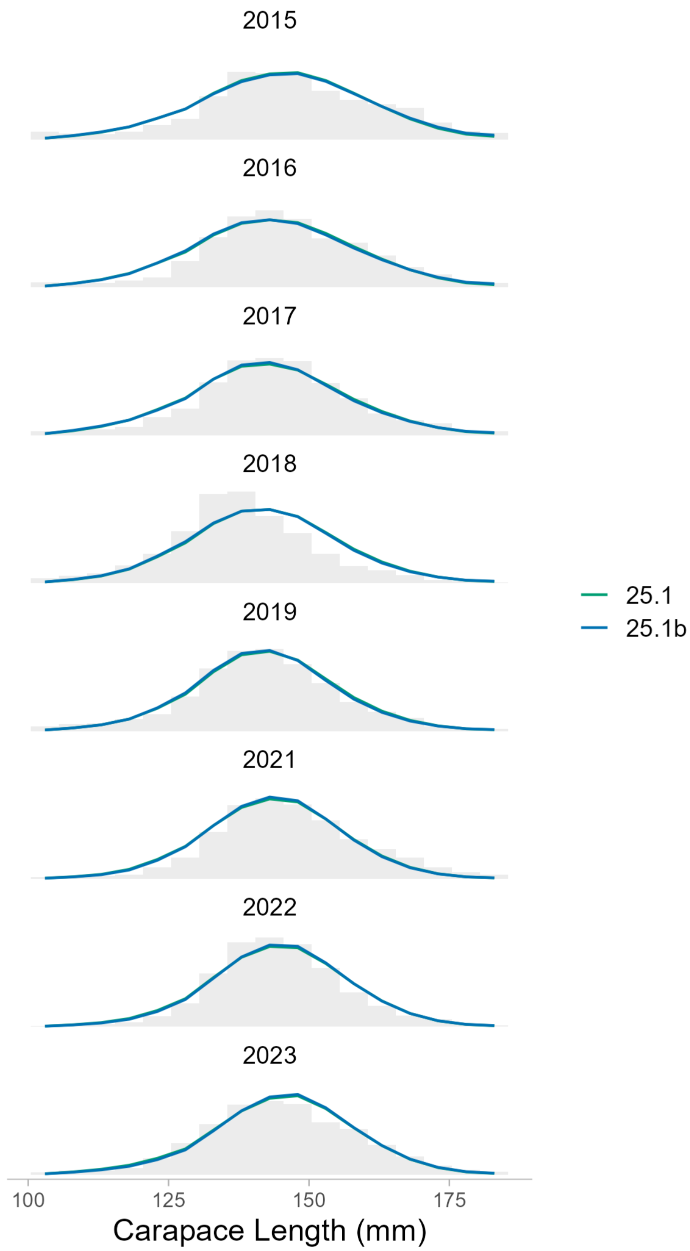

Figure 20: Fits to cooperative survey size composition data EAG models 25.1 and 25.1b.

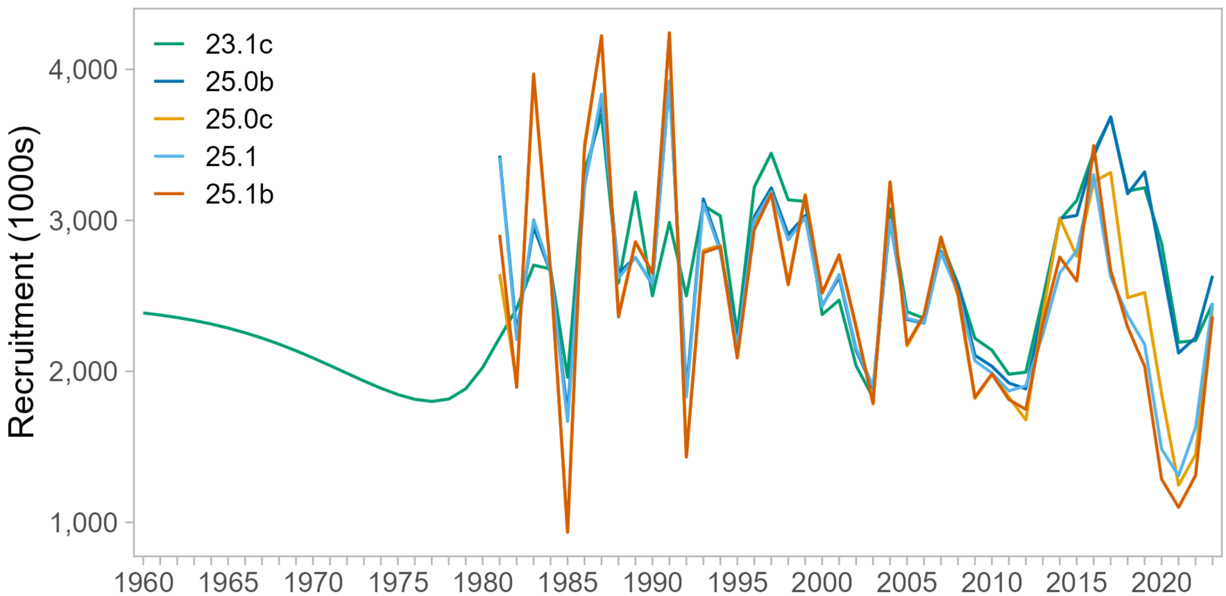

Figure 21: Recruitment estimates for EAG models. Models 25.1 and 25.1b included the cooperative survey as a fleet.

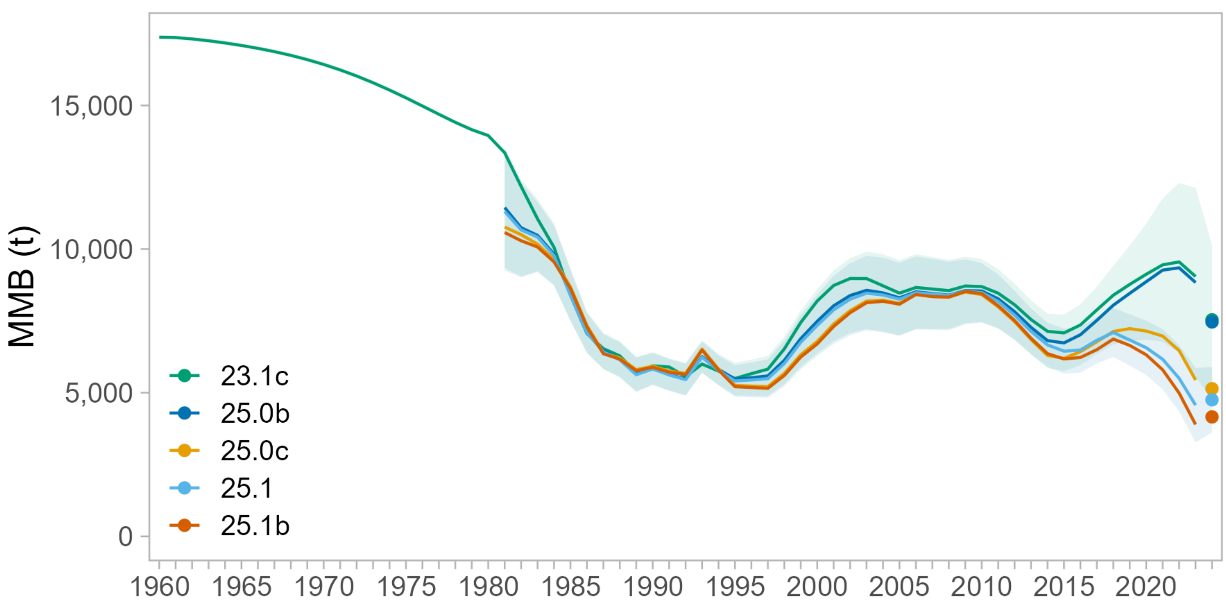

Figure 22: MMB trajectory for EAG models. Models 25.1 and 25.1b included the cooperative survey as a fleet. Shaded areas represent 95% confidence intervals for model 25.0b and 25.1.

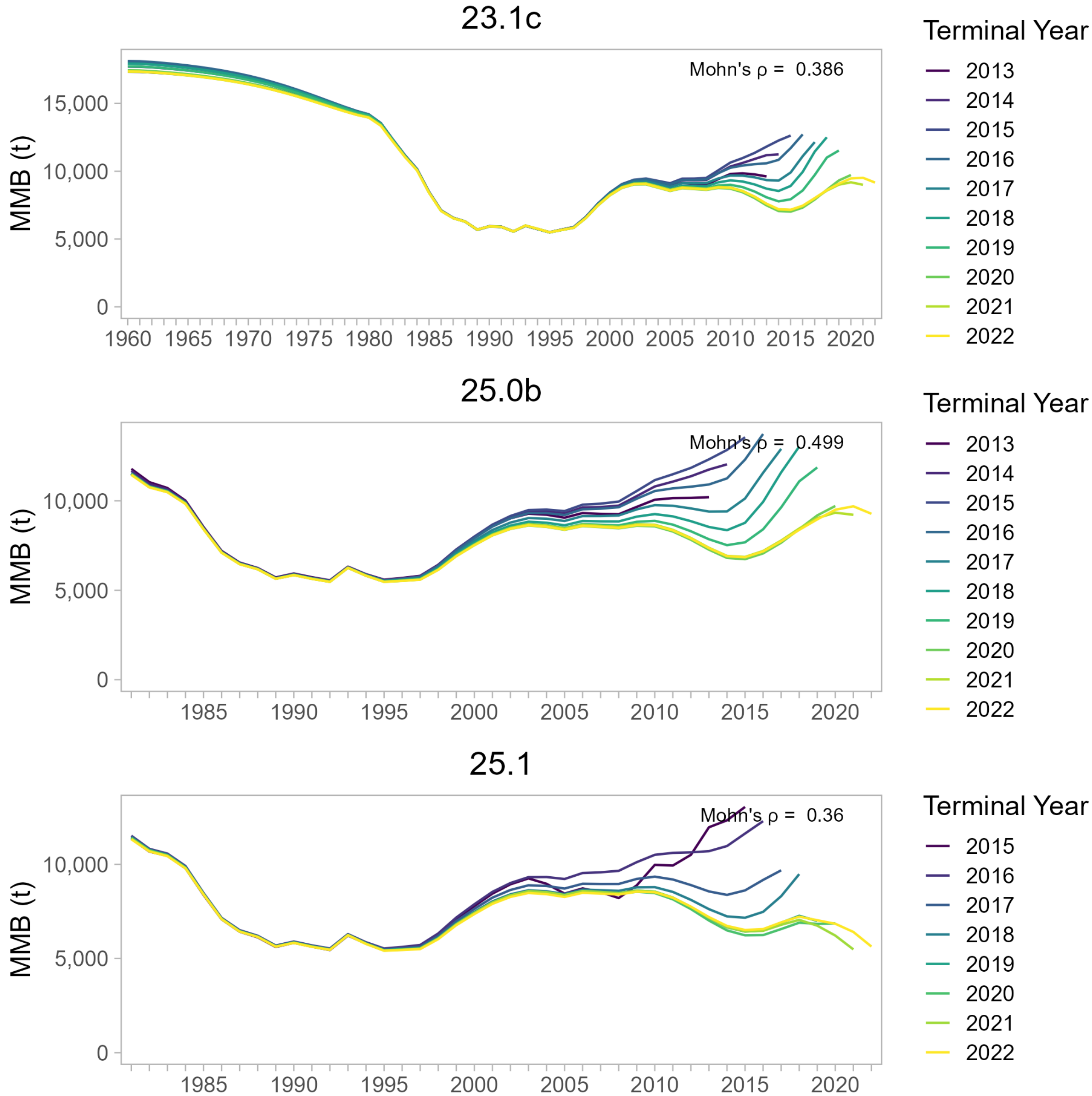

Figure 23: Retrospective patterns in mature male biomass (t) for models 23.1c, 25.0b, and 25.1 in the EAG.

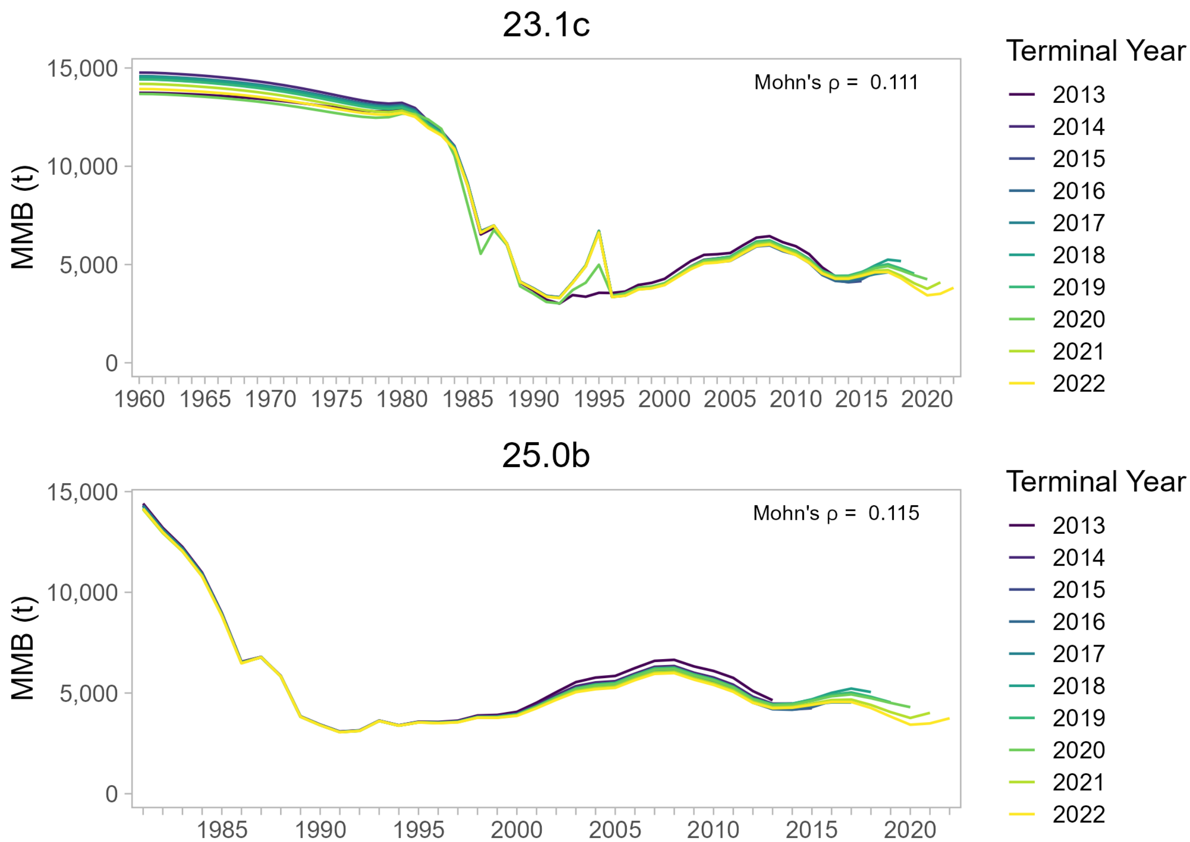

Figure 24: Retrospective patterns in mature male biomass (t) for models 23.1c and 25.0b in the WAG.

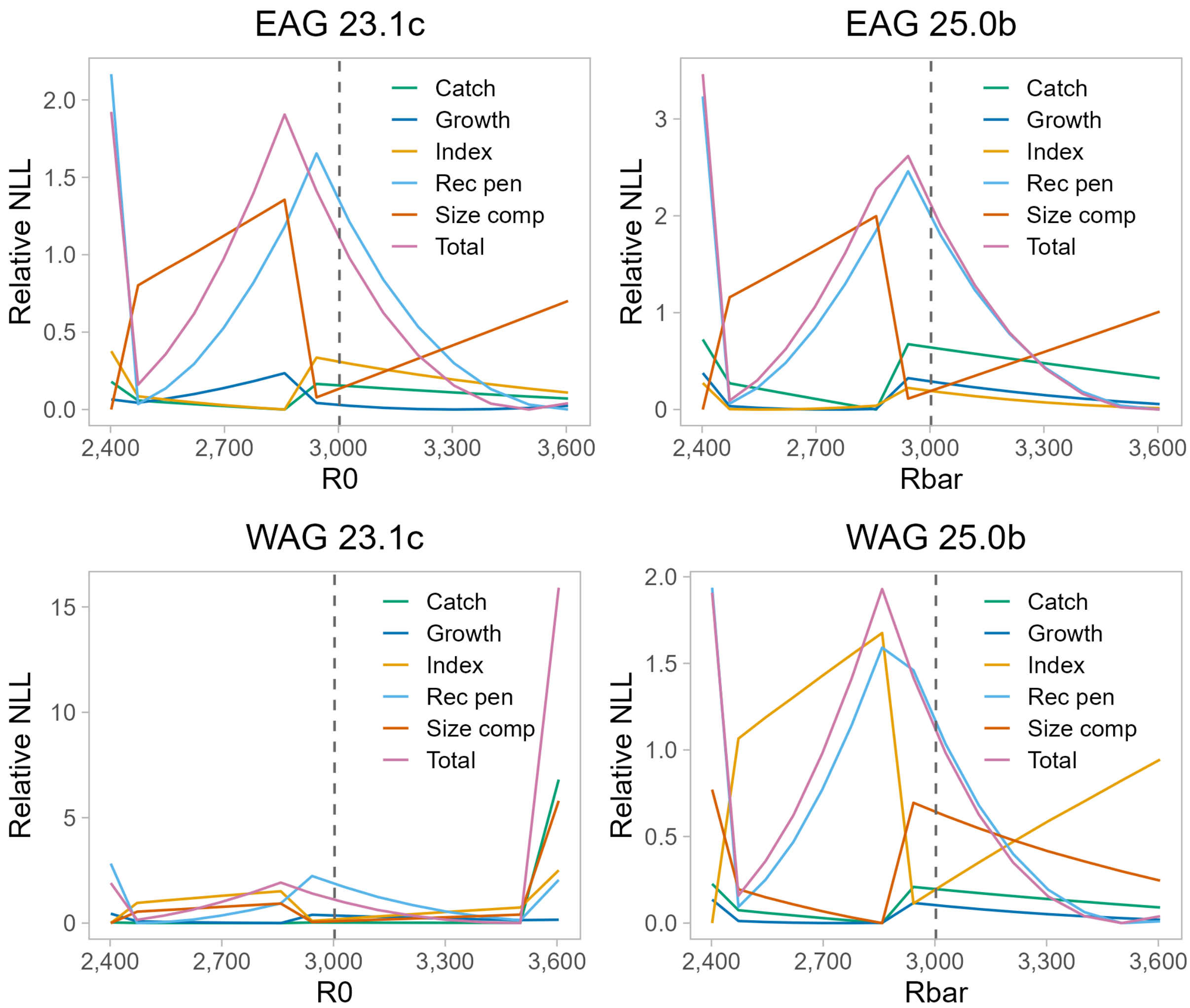

Figure 25: Likelihood profiles of R_0 and $\bar{R}$ for models 23.1c and 25.0b, respectively. The Grey dashed line indicated the maximum likelihood estimate.

Previously Addressed Comments

SSC Feb 2024

Comment: "*The SSC recommends that any new substantial standardization changes should be reviewed during the next cycle, not during specifications in May/June 2024*

Response: The only revisions to CPUE standardization between model explorations and the final assessment addressed poor model diagnostics, though this will be noted for the future.

....

CPT Jan 2024

Comment: "*The CPT recommends that the CPUE standardization be revised for the 2024 assessment by:*

- *exploring the use of a Tweedie instead of the negative binomial distribution;*
- *dropping the data for gear types 4 and 13 which have few observations;*
- *reporting DHARMa residuals and providing influence plots as additional diagnostics; and*
- *exploring the basic data used for the fish ticket CPUE index because the data on which the standardization is based for the current analyses include many zero observations – this may be because the extracted data may include trips for red king crab in the Aleutians. If the residual pattern for the fish ticket analysis (Fig. 44 of Appendix B) is not resolved, results should be presented in May 2024 for model runs that use and ignore the fish ticket CPUE index."*

Response: All of these recommendations were addressed in CPUE standardization except dropping gear types 4 and 13. This recommendation will be followed up in 2025 model explorations.

. . . .

Comment: "*Include measures of uncertainty (for at least one model configuration) in the plots for the estimates of recruitment and MMB*

Response: This has been addressed.

. . . .

Comment: "*Include a plot of the survey index overlaid on the observer CPUE index (EAG)*

Response: This plot will be included in documents that evaluate models containing survey data.

. . . .

Comment: "*Describe why the MMB for the EAG declines substantially before 1980 while this is not the case for the WAG*

Response: This is explained in section 4.g of Jackson (2024).

. . . .

Comment: "*Start the y-axis for the plots of recruitment and MMB at zero*

Response: This has been addressed.

. . . .

Comment: "*Include the number of parameters in likelihood tables*

Response: This has been addressed.

. . . .

Comment: "*Apply jittering to ensure that the reported parameters correspond to the global minimum of the objective function.*

Response: Jitter analysis was performed for the two author preferred model scenarios, model 23.1 and 23.1b.

. . . .

SSC June 2023

Comment: "*The SSC agrees with the CPT recommendation for a 25% buffer for this assessment and supports the resulting ABC. For the future, the SSC specifically requests that jitter and retrospective analyses be conducted for all final models that have the potential to be used for setting harvest specifications*"

Response: Retrospective analyses were performed here, and jitter analysis will be performed on the author preferred model in the final assessment.

. . . .

Comment: "*The SSC places a high priority on incorporating information from the cooperative survey into the assessment and supports the CPT recommendation that this be incorporated as a separate fleet.*"

Response: Model 23.2 explores the utility of the pot survey as an additional fleet.

. . . .

Comment: "*Further examination of the retrospective pattern in terms of magnitude, direction and cause continues to be important.*"

Response: More work will be done to address the retrospective pattern in the EAG during the next cycle.

. . . .

Comment: "*Revisit the choice to maintain the recruitment years at 1987 – 2017 rather than successively adding recent years to the time series, as is done for other crab stocks.*"

Response: See response to similar comment above.

. . . .

Comment: "*The CPT recommended removing the data on the smallest size bin for the total catch prior to 2005/2006. The SSC requests first plotting these data and the model fit and providing further consideration of why these data may or may not be representative of the fishery at that time.*"

Response: For clarification, the CPT recommended to removed data on crab below the smallest size bin (i.e. $\leq$ 100 mm) that were being included in the 101-105 mm bin.

. . . .

Comment: "*The current method of projecting the remaining landings for the current incomplete season seems overly complicated and the SSC recommends that a more straightforward method for determining total catch be considered, such as basing it on the average fraction harvested to date.*"

Response: In May 2024, total catch will be determined using the effort required to achieve the TAC at current CPUE on the date when data were pulled. See Appendix A for details of total catch estimation.

. . . .

Comment: "*Further analysis and discussion of the retrospective pattern is needed to justify the size of the buffer used.*"

Response: This will be noted during the final assessment in May 2024.

. . . .

CPT May 2023

Comment: "*Continue work to obtain an index using the cooperative pot survey data for use in the EAG assessment model.*"

Response: Model 23.2 explores the utility of the pot survey as an additional fleet. Now see model 25.1 and 25.1b.

. . . .

Comment: "*Plot observed vs. predicted values for fitted data to help diagnose misfits.*"

Response: It's unclear what model process this is referring to. When applicable, observations are always plotted with fitted data in this document.

. . . .

Comment: "*Add confidence intervals to plots of fits to catch data (i.e., retained catch, total catch) reflecting assumed data uncertainty.*"

Response: All plots of catch and index data now include confidence intervals.

. . . .

Comment: "*Perform retrospective analyses for all models that have the potential to serve as the basis for calculating reference points.*"

Response: Retrospective analyses were performed for all EAG and WAG models, and presented for 22.1e2, 23.0a, 23.1, 23.1b, 23.2, and AI 23.1b.

. . . .

Comment: "*The cooperative survey should be fit as an additional CPUE index, not substituted for existing indices as was done for models 22.1g and 22.1h.*"

Response: That is what has been explored here.

. . . .

Comment: "*Size-composition data should not include a "minus" group (i.e., crab smaller than the smallest size bin used in the model).*"

Response: This is rectified by model 23.1.

. . . .

Comment: "*The data used to determine the total catch size-compositions in the two areas should be re-examined to determine whether the abundances in the smallest size bin from 1990 to 2004 are correct.*"

Response: Appendix A recomputes size composition time series using data directly pulled from the observer database. Updated time series still appear to contain a disproportionate amount crab 101-105 mm CL, even without minus-sized crab (model 23.1). This is possibly do to escape mesh not being required until the 1997 season.

. . . .

Comment: "*Explore models that provide better fits to EAG CPUE data.*"

Response: More work in this area is needed during the next cycle.

. . . .

Comment: "*Use GAMs rather than GLMs to standardize the CPUE indices (e.g., use the R package "mgcv").*"

Response: All models derivative of 23.0a take this approach.

. . . .

Comment: "*Show both the original CV's and effective CV's (i.e., incorporating additional variance) when showing fits to the CPUE index time series.*"

Response: This has been done in all plots showing fits to CPUE index.

. . . .

Comment: "*In the SAFE document*

- *Add a note to explain that retained catch can exceed TAC in some years due to the cost recovery fishery associated with the cooperative survey.*
- *Drop Appendix D.*

- *Remove tier designation from area-specific management Table.*
- *Add explanation for extrapolation of total catch in final year"*

Response: All items will be addressed in the May 2024 SAFE document.

. . . .