May 2020 Report on Developments in the Tanner Crab Assessment

William T. Stockhausen
Alaska Fisheries Science Center
April 2020

1.0 Introduction

This report presents discusses developments in the Tanner crab assessment since the Fall 2019 assessment and provides a list of author-recommended alternative models to be considered for the Fall 2020 assessment. It is organized in the following sections:

- 2: Responses to SSC and CPT Comments
- 3: Analyses
- 4: Model development
- 5: Model scenarios
- 6: Model results
- 7: Recommended scenarios for Fall 2020

2.0 Responses to SSC and CPT Comments

1. Responses to the most recent two sets of SSC and CPT comments on assessments in general.

Oct 2019 SSC Meeting
SSC Comment: The SSC reminds authors to use the model numbering protocols that allows the SSC to understand the year in which a particular version of the model was first introduced.
Response: The requested numbering protocols have been implemented, with the 2019 assessment model "backdated" and referred here as 19.03 (where it was referred to 19F03 during the 2019 assessment).

SSC Comment: the SSC requests that the CPT consider developing a standard approach for projecting the upcoming year's biomass that does not include removing the entire OFL for stocks where recent mortality has been substantially below the OFL. This may appreciably change the projected biomass levels for stocks such as Tanner crab, where actual catch mortality has been less than 10% of the OFL .
Response: The CPT has not yet developed a standard approach for doing so.

SSC Comment: the SSC encouraged authors to work together to create a standard approach for creating priors on selectivity and catchability from these (BSFRF/NMFS side-by-side trawl) data for use in the respective assessments. A hierarchical comparison of all species pooled, separated species, and separated sexes may be helpful for understanding where statistically supported differences exist. Where sample sizes are modest (e.g., snow crab), bootstrapping, or a sample size-weighted estimate rather than a raw average may be useful for aggregating across years.
Response: Results from one preliminary attempt to develop priors on sex/size-specific catchability (q *x* selectivity) and availability for Tanner crab are presented in this report. An option to use such priors has also been added to the Tanner crab assessment model code, but was not utilized in this report due to time constraints.

September 2019 Crab Plan Team Meeting
No new general comments.

June 2019 SSC Meeting
SSC Comment: The SSC reminded authors to use the model numbering protocols that allow the SSC to understand the year in which a particular version of the model was first introduced.
Updated response (05/20): The requested numbering protocols have been implemented, with the 2019 assessment model "backdated" and referred here as 19.03 (where it was referred to 19F03 during the 2019 assessment).

Original response: The Tanner crab assessment has not fully implemented this suggestion. The 2018 assessment model was labeled 18AM17, which does not follow the guidelines. Here, that model is referred to as M19F00 ("00" designating the base model from which other scenarios proceed in the 2019 assessment, "F" denoting the "final" scenarios proposed in May). This also does not reflect the requested model numbering. However, the model numbering adopted herein should allow subsequent model numbering to follow the guidelines (so that the author's preferred model M19F03 would become 19.03 in the future).

May 2019 Crab Plan Team Meeting
No new general comments.

October 2018 SSC Meeting
SSC Comment: The SSC reminded authors to use the model numbering protocols that allow the SSC to understand the year in which a particular version of the model was first introduced.
Updated response (05/20): The requested numbering protocols have been implemented, with the 2019 assessment model "backdated" and referred here as 19.03 (where it was referred to 19F03 during the 2019 assessment).

Original response: Model numbering was consistent with this guideline for the model scenarios presented by the author to the CPT in September 2018. However, the CPT recommended a model based on the 2017 assessment which was labeled 18AM17 to designate the 2017 assessment model updated with 2018 data, which did not follow the guidelines.

SSC Comment: The SSC encourages authors (using VAST estimates of survey biomass) to consider whether or not the apparent reduction in uncertainty in survey biomass is appropriately accounted for with their models.
Updated response (05/20): Two model scenarios fitting VAST estimates of survey biomass were included in this report: one which fit the estimates without adjusting the variance estimates and one which estimated parameters describing "extra" uncertainty (i.e., re-inflating the uncertainty of the VAST estimates). While the model fit without estimating "extra" uncertainty was "worse" from a strictly likelihood perspective (larger z-scores) compared to that from the same model fit to the standard design-based estimates, the predicted values "fit" the VAST estimates better from a visual standpoint (i.e., on a scale unweighted by the uncertainty). Unfortunately, the attempt to compensate for the possible over-shrinkage of uncertainty in the VAST estimates by estimating parameters related to "extra" uncertainty failed because the model converged to with the parameters at their upper bounds (equivalent to "extra" CVs of 270%).

Original response: The Tanner crab assessment does not yet use VAST-based estimates of survey biomass.

September 2018 Crab Plan Team Meeting

No new general comments.

2. Responses to the most recent two sets of SSC and CPT comments specific to the assessment.

October 2019 SSC Meeting

SSC comment: The SSC requested that for the next assessment, models be reparameterized, simplified, or have parameter bounds adjusted such that no parameters remain at the bounds after estimation.
Response: TBD.

SSC comment: Use the standard model numbering approach.
Response: Done.

SSC comment: In next year's assessment, project biomass using a mortality level consistent with recent years, rather than the full OFL (see general CPT comments).
Response: TBD.

SSC comment: Provide a retrospective analysis for future assessments.
Response: Code has been implemented to allow doing so in a timely manner.

SSC comment: Add the 2018 BSFRF/NMFS side-by-side data for all future analyses of that time-series.
Response: BSFRF has not yet provided this data.

SSC comment: Report the values for natural mortality actually used for calculation of reference points in the appropriate table(s).
Response: This is an issue for the CPT, not the author (the relevant tables are in the Introduction to the SAFE and are updated by the CPT).

SSC comment: Provide additional information on data weighting. Specifically, identify standardized residuals appreciably greater than would be expected by chance (e.g., values of 4 and larger), report mean input and harmonic mean effective sample sizes by source for evaluation of model fit, and consider basing input sample sizes on the number of trips/hauls sampled rather than number of individual crab measured..
Response: TBD.

September 2019 CPT Meeting

The CPT suggested exploring appropriate values for catchability. For example, runs that fit to the BSFRF data and fix availability to empirical estimates to contrast the outcomes with runs in which availability is estimated could be informative for what is driving the small estimates of catchability in the author-preferred model.
Response: Empirical estimates of availability and selectivity are developed from BSFRF and NMFS side-by-side (SBS) selectivity study data for Tanner crab in Sections 3.3 and 3.4. These are used in Model Scenarios 20.07 (fixed empirical availability curves used to fit BSFRF SBS and NMFS EBS survey data and estimate NMFS EBS survey catchability) and 20.08 (fixed empirical selectivity curves used to fit NMFS EBS survey data), and compared with results from the 2019 assessment model.

The CPT suggested exploring the relationship between natural mortality, growth, and overestimates of large crab. For example, estimate growth outside the model to attempt to address the overestimates of large crab.
Response: Work is in progress to address this issue.

The CPT suggested exploring maturity states for growth increment data and make recommendations for directions for growth model development.
Response: Work is in progress to address this issue.

The CPT requested include the data to which the models are fit for the survey biomasses figures in the presentation.
Response: The data was dropped for clarity of comparison among model predictions of survey biomass. The data will be included in future plots of this sort.

The CPT requested that if 'catchability' is to be used for something similar to 'fully-selected fishing mortality', perhaps translate it to a 0-1 scale and distinguish it from survey catchability so that it is clear that there is mortality associated with it.
Response: The term "catchability" was used to describe the rate at which "fully-selected" crab are captured in a fishery. Because some discards are assumed to survive, this is not equivalent to "fully-selected fishing mortality" (if discard mortality were 0, there would be *no* mortality associated with capture in a bycatch fishery). Perhaps "capturability" would cause less confusion?

The CPT requested that the author explore ways to provide a retrospective analysis of the assessment model.
Response: A substantial effort was made to add the capability to perform a retrospective analysis to the assessment model. This has now been completed and retrospective analyses will be provided at the Fall 2020 CPT Meeting.

June 2019 SSC Meeting
The SSC endorsed the CPT suggestions from its May meeting.
Response: none.

The SSC requested an evaluation of all parameters estimated to be at or very near bounds, or substantially limited by priors (unless those priors can be logically defended).
Original response (9/19): Two tables of parameters estimated at or near their bounds are provided (Tables 18 and 19). These parameters are estimated at their bounds in all (or nearly all) of the scenarios examined here. The parameters include one related to peak retention in the directed fishery prior to 1997 (at its upper bound on the logit scale, implying full retention of large legal males) and two related to the probability of undergoing terminal molt (effectively 1 for males in the largest model size bin and 0 for females in the smallest model size bin). These could be fixed in future models (the latter two are in several scenarios here). Survey catchability parameters for the 1975-1981 time period were also estimated at their lower bound (0.5). This might not be unreasonable given the reduced areal coverage of these surveys relative to later surveys and the spatial limits of the Tanner crab stock. However, it would be worthwhile to explore the effect of reducing these bounds. The remaining parameters are related to selectivity functions describing the size-specific capture efficiency of the fisheries and surveys. Two at their lower bounds are probably inconsequential (pS2[10] and pS4[1]) and are related to the ascending and descending slopes of the dome-shaped selectivity describing male bycatch in the snow crab fishery prior to 1997. A double-normal is used to describe the dome shape, but an alternative function (e.g., a single normal) might have better estimation properties. The size at 50% selected was estimated at its upper bound (90 mm CW) for NMFS survey selectivity in the 1975-1981 time period pS1[1]). This results in an almost linear function, rather than asymptotic, across the size range. This result may reflect the changing interaction between the areas surveyed (availability) and the gear selectivity in this time period as the survey gradually extended from the southeast shelf and Bristol Bay where adult males were prevalent to the north and west where more immature males would be encountered, effectively "seeing" relatively more large males than small males. Two other survey-related selectivity parameters, describing the size difference between crab at 50% and 95% selected) were estimated at their upper bounds for the both males and females in the NMFS EBS trawl survey in the 1982-present time period (pS2[2] and pS2[4]). The selectivity functions are assumed to be logistic, with the other estimated parameter being the size at 95% selected. The practical consequence of this is that small crab (females in particular) are described as fairly well-selected (> 50% for females) relative to fully-selected (sex-specific) large crab. This result may reflect conflicts from between the model assumption of equal sex ratios for recruitment in

the 25-40 mm CW range, apparent equal abundances and spatial patterns for males and females at small sizes in the NMFS EBS survey, and assumed logistic selectivity. The selectivity parameter describing the size at 50% selected for males in the groundfish fisheries during 1987-1996 was estimated in all scenarios at its lower bound (40 mm CW), probably a consequence of fairly substantial catches of small crab in some years (e.g., 1993, Figure 12). Finally, three parameters at their upper bounds (pS1[23], pS1[24], and pS1[27]) are related to the size at 95% selected in the BBRKC fishery in the 1997-2004 (males) and 2005+ (males and females) time periods. The upper bounds (180 for males, 140 for females) were selected to reflect the largest possible sizes reasonably expected in the model, so the resulting selectivity functions are essentially positively-sloped linear functions with values fixed at 0.95 at the parameter bound because the other estimated logistic parameter estimates a large size at 50% selected (see selectivity curves in Figure 46).

May2019 Crab Plan Team Meeting

CPT comment: compare the estimated selectivity to the ratio of NMFS to BSFRF numbers at length. Is estimated and empirical catchability/availability/selectivity the same? Does the empirical selectivity look logistic?"

Original response (9/19): The model-estimated availability of Tanner crab to the survey gears in the side-by-side (SBS) study areas was compared to "empirical" estimates of availability using the ratio of numbers-at-size in the NMFS SBS datasets to those form the full NMFS EBS survey. The results are shown in Figure 53. While there are some similarities between the two sets, there are also substantial differences when conceptually they should be the same. Results for the empirical size-specific relative catchability (the ratio of NMFS to BSFRF estimated abundance at size) are shown in Figure 65, but are not compared directly to the estimated selectivity. The mean curves appear reasonably logistic, with approximate asymptotes of ~0.6 for males and ~0.4 for females. If the BSFRF surveys are regarded as providing estimates of absolute abundance (catchability=1 for all sizes), this would suggest fully-selected NMFS survey "q"'s are ~0.6 for males and ~0.4 for females--which are about 50% higher than the estimates (0.43 and 0.24, respectively) from the assessment model, but within the 95% confidence intervals for males (0.37-0.49) (but not females: 0.19-0.29).

CPT comment: show the fits to the BSFRF length composition data by year as well as in aggregate.

Original response (9/19): These fits are shown in Appendix B.

CPT comment: check the bounds of parameters when estimating the BSFRF data.

Original response (9/19): Fitting the BSFRF data results in no better, or worse, performance in terms of parameters hitting their bounds.

CPT comment: indicate whether or not Hessians were produced.

Original response (9/19): Hessians were produced for the "best" model runs for all scenarios and .std files were obtained.

CPT comment: Suggest rationale for chosen weighting for the second difference smoothing on the availability curve.

Original response (9/19): The rationale for the selected weighting is that it reflects a preference toward a smoothly-varying function, reflecting an assumption that crab of similar sizes would tend to be found together with no abrupt dichotomies (which would justify a smaller smoothing weight) in spatial distribution with size. However, this assumption has not been examined in detail.

CPT comment: Compare trends in largest crab to fishing pressure and area occupied by stock.

Original response (9/19): This is a good suggestion that, time permitting, will be addressed before the January 2020 CPT meeting.

CPT comment: Compare the maximum sizes seen in the fishery to the survey.
Original response (9/19): Another good suggestion that, time permitting, will be addressed before the January 2020 CPT meeting.

CPT comment: Consider blocking for estimation of growth and probability of maturing.
Original response (9/19): This has been on the "to do" list for a while now, but with relatively low priority. The problem is that the principal data which the model relies on for estimating both processes is, except for size compositions, only available (from a practical standpoint) since 2006 for male maturity ogives and since 2015 for (both sexes) molt increment data. The ability of the model to reliably estimate changes in these processes is thus somewhat doubtful.

CPT comment: Make incorporating chela height data in the assessment a priority because this might address changes in the probability of maturing over time
Original response (9/19): Chela height data, in the form of male maturity ogives based on collections of chela heights since 2006, is incorporated in several model scenarios examined here, including the author-preferred scenario.

CPT comment: Provide retrospective analysis and calculate Mohn's rho for MMB
Updated response (5/20):

Original response (9/19): Retrospective analyses for Tanner crab are complicated given the recent fishery closures and short time frames for molt increment and maturity ogive data. Time did not permit making retrospective analyses for the model scenarios considered herein. However, a retrospective analysis for the CPT-selected assessment model could be presented at the January 2020 CPT meeting.

3.0 Analyses

3.1 Carapace width-weight regressions

Weights for many crab sampled in the surveys and by observers in the fisheries are based on measured carapace width (CW), with CW converted to weight using previously-determined regressions of crab weight on CW from crab observed in the NMFS EBS bottom trawl surveys on which measurements of both types were taken (Lang et al., 2018). For female Tanner crab, separate CW-weight regressions are used to determine weights for immature and mature crab while a single regression is used for male crab. However, none of these regressions depends on shell condition. “New shell” crab are assumed to have undergone a molt within the year previous to capture, while “old shell” crab are assumed to have undergone their terminal molt more than a year previous to capture. Following a molt, a new shell crab puts on mass and “fills out” its larger shell over a period of several months. Given the typical timing (summer) of the NMFS EBS survey on which these regressions are based, this suggests that measured new shell crab probably underwent a molt within at most a couple of months prior to capture in the survey and are still in a period of “filling out” when captured in the survey. This, in turn, suggests that CW-weight regressions that do not take shell condition into account may over-estimate weights for new shell crab and under-estimate weights for old shell crab.

As suggested by Figure 3.1.1, the suggestion that the regressions may underestimate weights for old shell crab seems to be supported by applying the current CW-weight regressions for male Tanner crab to males classified by shell condition captured in the NMFS EBS bottom trawl survey on which both CW and weight measurements were taken. The discrepancy is much less apparent for female crab (Figure 3.1.2) because separate regressions are used for immature (always new shell) and mature females (mix of new and old shell). As such, it may be worthwhile to revisit the regressions for male Tanner crab.

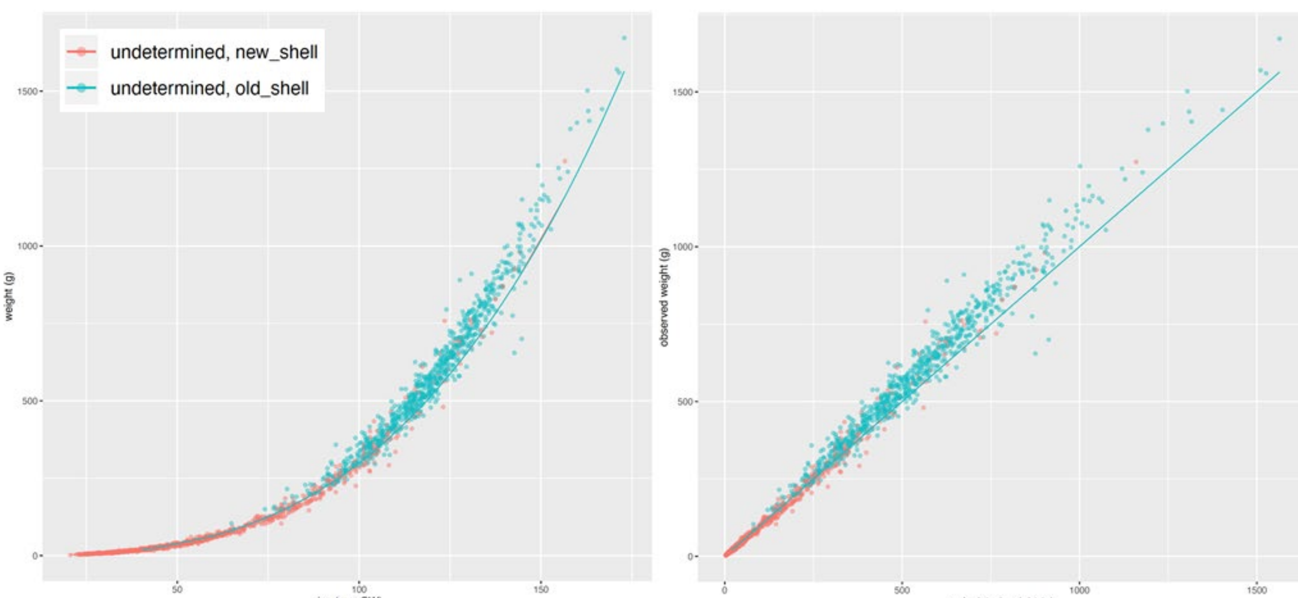

Figure 3.1.1. Observed (dots) and predicted (lines) weights for male Tanner crab from the NMFS EBS bottom trawl survey based on current CW-weight regressions. Left plot is weight vs. CW. Righthand plot is observed weight vs. calculated weight. Orange coloring denotes values for new shell crab, green denotes values for old shell crab.

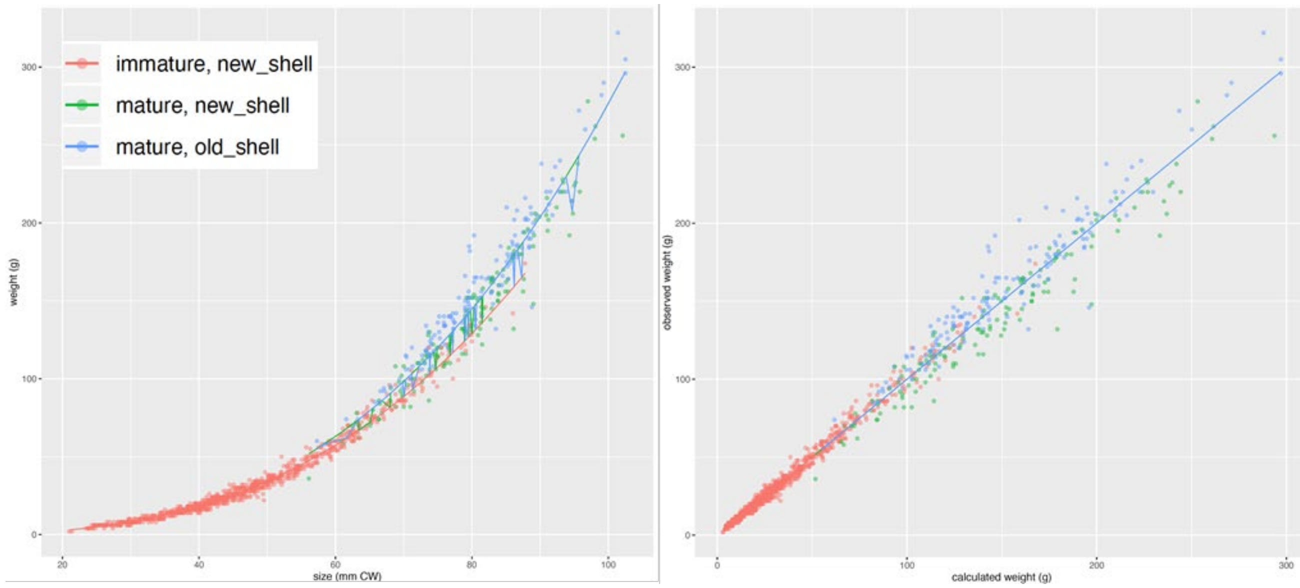

Figure 3.1.2. Observed (dots) and predicted (lines) weights for female Tanner crab from the NMFS EBS bottom trawl survey based on current CW-weight regressions. Left plot is weight vs. CW. Righthand plot is observed weight vs. calculated weight. Orange coloring denotes values for immature (new shell) crab, green denotes values for mature new shell crab, and blue denotes mature old shell crab.

3.2 Revised biomass and size composition estimates from BSFRF/NMFS side-by-side studies

The BSFRF and NMFS conducted "side-by-side" (SBS) selectivity studies for Tanner crab in 2013-2018, although BSFRF data for 2018 has not yet been made available to the author. These data were included as estimates of biomass and size compositions within the SBS study areas in the model fits for several scenarios considered during last year's Tanner crab assessment. The areas used in the SBS studies differed each year and were a small subset (~20%) of the full EBS NMFS survey area. For last year's assessment, estimates of total biomass and size compositions within the SBS areas were calculated ignoring the area stratification used to calculate total biomass and size compositions for the full EBS survey. While not wrong, per se, this approach was inconsistent with that used for the full EBS survey and could result in situations where size-specific availability of crab was greater than 1 (i.e., an estimate of size-specific abundance in the SBS area was greater than that for the full EBS survey). In order to avoid the possibility that availability in the SBS area was greater than 1, estimates of total biomass and size compositions were re-calculated for this report (and future use) using the same stratification as that for the full EBS survey, with catches at non-SBS stations in the full survey set to 0. This change had very small effects on the estimates of total biomass (Figure 3.2.1) and size composition (Figure 3.2.2) in the SBS areas but did increase the associated variances somewhat. It did, however, ensure that size-specific availability was always in the interval from 0 to 1.

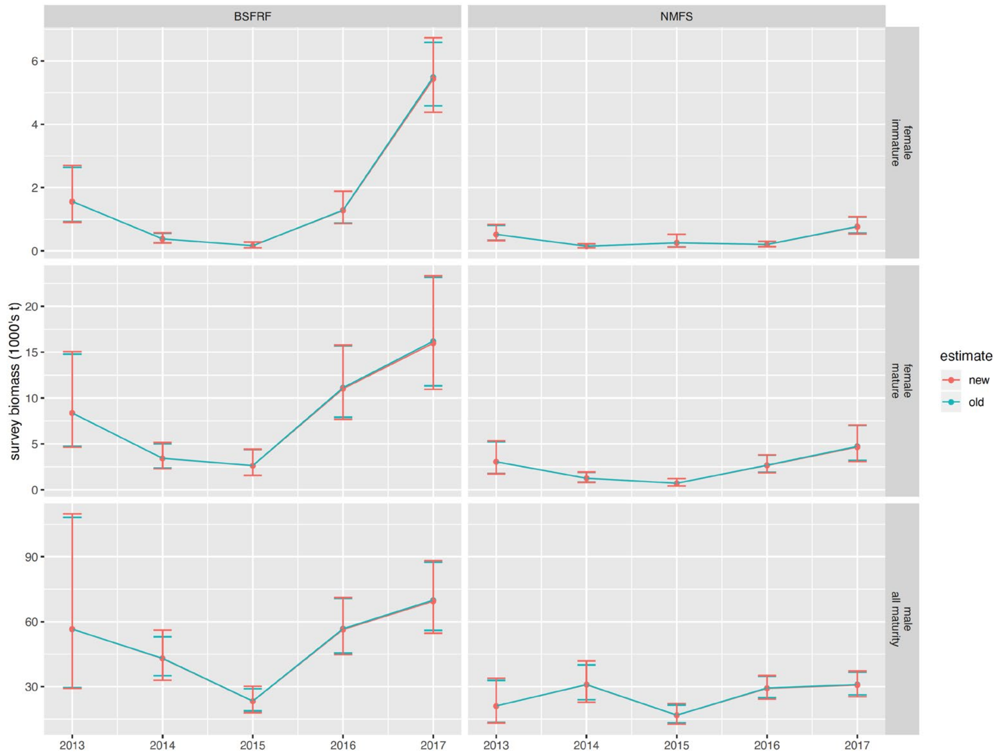

Figure 3.2.1. Comparison of estimates of total biomass in the SBS study areas using two different approaches (old: with no area stratification; new: using NMFS EBS survey area stratification). Note: the estimates from the two approaches are almost indistinguishable.

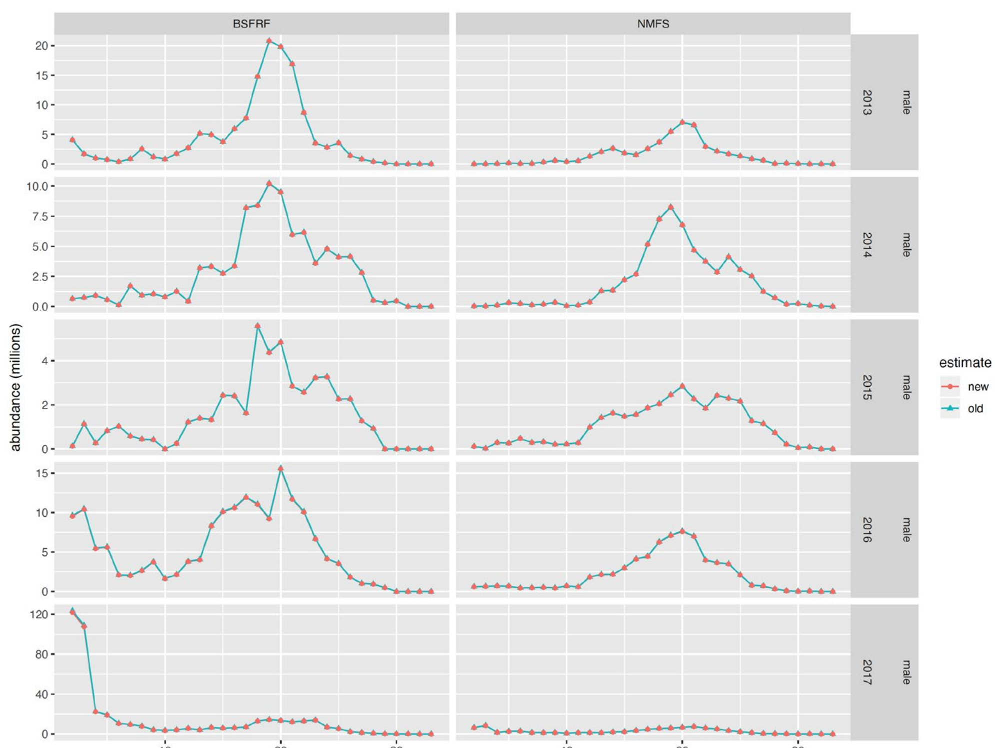

Figure 3.2.2. Comparison of estimates of male size compositions in the SBS study areas using two different approaches (old: with no area stratification; new: using NMFS EBS survey area stratification). Note: the estimates from the two approaches are almost indistinguishable.

3.3 Empirical estimates of survey availability

"Availability" refers to the fraction of individuals in a population that are available to be captured in a survey or by a fishing fleet (i.e., the fraction that could conceivably be caught). In the case of surveys, availability is considered to be 1 when the survey area encompasses the entire population/stock area and no individuals occupy refuge habitats that cannot be surveyed (e.g., untrawlable rocky bottom). When a survey area does not encompass the entire population/stock area, then availability will be less than 1. In the assessment model, size-specific survey abundance is related to size-specific population abundance by

$$N_z^{survey} = A_z^{survey} \cdot C_z^{survey} \cdot N_z^{population} \tag{1}$$

where z represents size, N represents size-specific abundance, A represents size-specific availability, and C represents size-specific survey catchability.

Availability in the NMFS EBS bottom trawl survey is considered to be 1 for all Tanner crab in the EBS stock. Because the SBS selectivity studies were conducted on smaller areas than the full survey, availability to the NMFS and BSFRF gear in these studies must be less than or equal to 1. The BSFRF gear is assumed to catch all crab within the footprint of a tow (i.e., it is non-selective and provides an estimate of absolute abundance), so the following relationships are assumed in the assessment model to hold:

$$N_z^{NMFS\ EBS} = 1 \cdot C_z^{NMFS} \cdot N_z^{EBS} \tag{2}$$

$$N_z^{NMFS\ SBS} = A_z^{SBS} \cdot C_z^{NMFS} \cdot N_z^{EBS} \tag{3}$$

$$N_z^{BSFRF\ SBS} = A_z^{SBS} \cdot 1 \cdot N_z^{EBS} \tag{4}$$

where we have substituted the assumptions $A_z^{EBS} \equiv 1$ and $C_z^{BSFRF} \equiv 1$ in the equations above.

In the scenario in last year's assessment that incorporated the SBS studies data, both the NMFS survey catchability (C_z^{NMFS}) and the size-specific availability for each SBS study area (A_z^{SBS}) were estimated in the model—the former assuming a logistic form for survey catchability and the latter using a non-parametric, "smoothed" approach that placed fewer constraints on the shapes of the size-specific availability curves. However, dividing Equation 3 by Equation 2 and rearranging yields:

$$A_z^{SBS} = \frac{N_z^{NMFS\ SBS}}{N_z^{NMFS\ EBS}} \tag{5}$$

which allows one to estimate availability outside the assessment model using the ratios of size compositions from the NMFS survey in the SBS and full EBS areas.

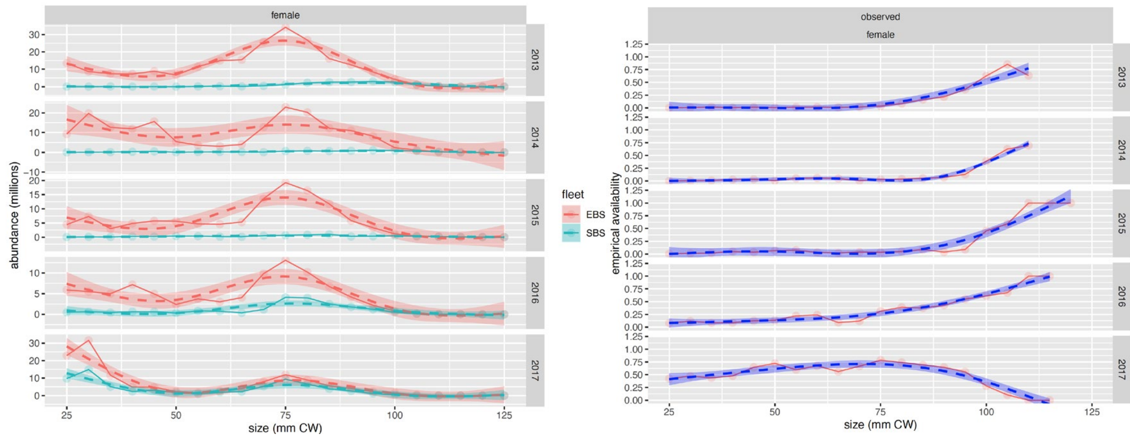

Figure 3.3.1. Lefthand plot: NMFS bottom trawl survey size compositions for female Tanner crab from the EBS (orange) and SBS (green) survey areas for 2013-2017. Solid lines and dots are “raw” estimates, dashed lines are smoothed fits using cubic splines. Righthand plot: empirical availability curves calculated using Equation 5. Red lines and dots: “raw” curves, dashed lines and fills: smoothed fits.

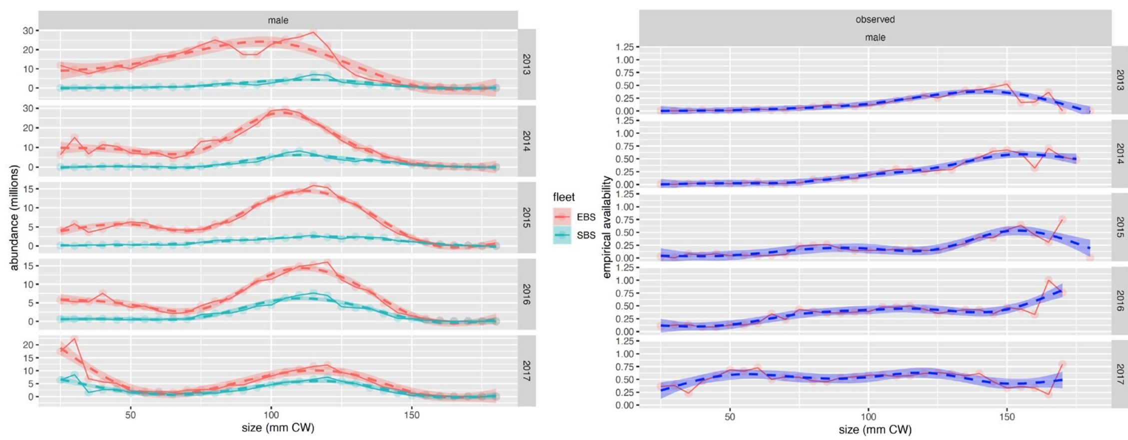

Figure 3.3.2. Lefthand plot: NMFS bottom trawl survey size compositions for male Tanner crab from the EBS (orange) and SBS (green) survey areas for 2013-2017. Solid lines and dots are “raw” estimates, dashed lines are smoothed fits using cubic splines. Righthand plot: empirical availability curves calculated using Equation 5. Red lines and dots: “raw” curves, dashed lines and fills: smoothed fits.

The SBS and EBS size compositions, and resulting “raw” empirical availability curves, from the 2013-2017 SBS selectivity studies are shown in Figures 3.3.1 and 3.3.2. Availability was generally small (< 0.25) for female Tanner crab smaller than 80 mm CW and increases with size in 2013-2015 when the study areas were in the inner and middle shelf domains near Bristol Bay (Figure 3.3.1). When the study area shifted west to the Pribilof Islands in 2017, availability increased for small females but decreased for large females. Availabilities for male Tanner crab showed similar patterns for small crab (< 100 mm), but availability tended to decrease with size for the largest males (> 150 mm CW), except in 2016. Clearly, though, the patterns are different between the sexes and on an annual basis.

The CPT and SSC suggested it might be possible to use these data to set priors on availability curves estimated in the assessment model for scenarios which fit the SBS data. In order to develop such priors, I undertook a bootstrapping analysis to explore the uncertainty in the empirical estimates. Bootstrapping was done as a hierarchical process for each SBS year. For each bootstrap replicate, n_y^{SBS} hauls were randomly selected with replacement from those in the SBS study area for that year, after which $n_y^{EBS} - n_y^{SBS}$ were then randomly selected with replacement from the EBS survey hauls conducted outside the SBS area for that year, where n_y^{SBS} was the number of hauls conducted in the SBS area in year y and n_y^{SBS}

was the number of hauls conducted in the full EBS survey area in year *y*. Following haul selection, n_h individuals were randomly selected with replacement from the individuals actually captured in that haul, with n_h representing the number of individuals captured. If a haul was selected multiple times, individuals from that haul were re-sampled each time. Sex-specific EBS and SBS size compositions were then calculated using 5 mm bins from 25 mm to 180 mm CW for each bootstrap realization using the standard, stratified EBS design-based approach, after which sex-specific "raw" availabilities were calculated for each 5 mm size bin using Equation 5. For each SBS year, bootstrap replicates were drawn 1000 times.

Figures 3.3.3 and 3.3.4 summarize the results from this bootstrapping analysis, showing the variability of the estimates within each 5 mm size bin for each year, as well as a smoothed fit across the size bins. Not surprisingly, the smoothed fits are similar to those shown in Figures 3.3.1 and 3.3.2, but the scale of variability within any size bin seems reasonably large compared to the fit. Although the final step remains to be taken, one possible approach to developing priors for availability curves estimated in the assessment model to describe the SBS data would be to use beta distributions with mean and variance parameters equal to those from the bootstrapping analysis.

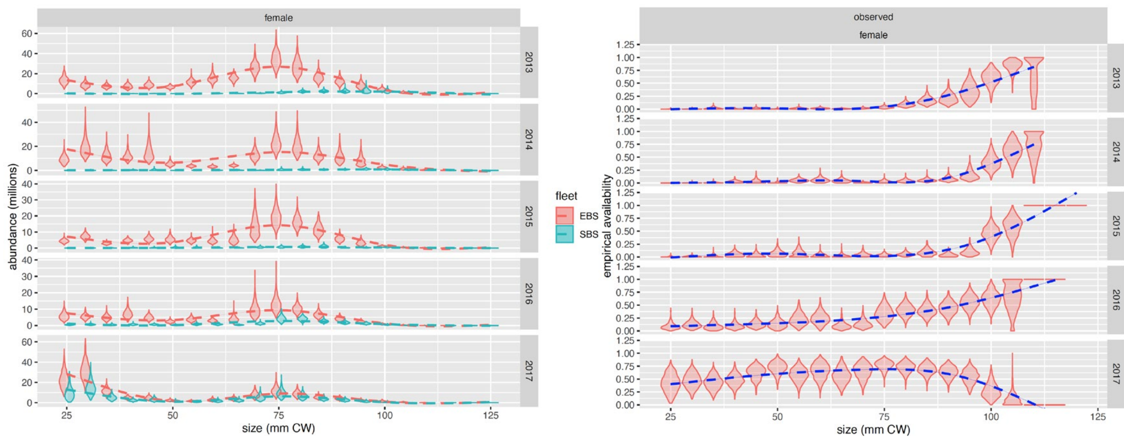

Figure 3.3.3. Lefthand plot: Bootstrapped NMFS bottom trawl survey size compositions for female Tanner crab from the EBS (orange) and SBS (green) survey areas for 2013-2017. Violin plots summarize "raw" estimates, dashed lines are smoothed fits using cubic splines. Righthand plot: bootstrapped empirical availability curves calculated using Equation 5. Red violin plots: "raw" estimates, dashed lines: smoothed fits.

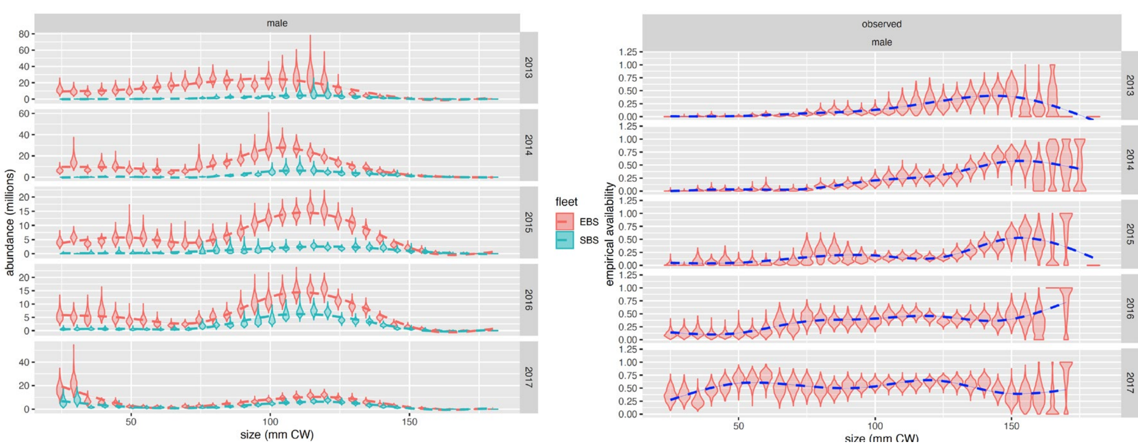

Figure 3.3.4. Lefthand plot: Bootstrapped NMFS bottom trawl survey size compositions for male Tanner crab from the EBS (orange) and SBS (green) survey areas for 2013-2017. Violin plots summarize "raw" estimates, dashed lines are smoothed fits using cubic splines. Righthand plot: bootstrapped empirical availability curves calculated using Equation 5. Red violin plots: "raw" estimates, dashed lines: smoothed fits.

3.4 Empirical estimates of survey catchability

Using Equations 2-4, it is also possible to estimate NMFS survey catchability (C_z^{NMFS}) directly from the SBS study data. Dividing Equation 3 by Equation 4 yields:

$$C_z^{NMFS} = \frac{N_z^{NMFS\,SBS}}{N_z^{BSFRF\,SBS}} \tag{6}$$

which provides an empirical estimate of NMFS survey catchability outside the assessment model.

The NMFS and BSFRF size compositions and resulting "raw" empirical catchability curves from the 2013-2017 SBS selectivity studies are shown in Figures 3.4.1 and 3.4.2.

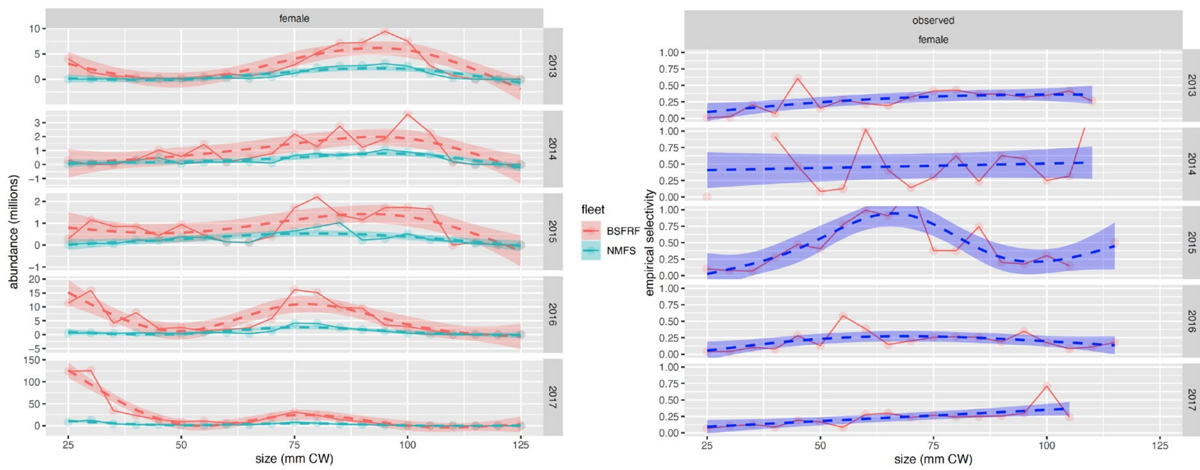

Figure 3.4.1. Lefthand plot: BSFRF (orange) and NMFS (green) size compositions for female Tanner crab from the SBS study areas for 2013-2017. Solid lines and dots are "raw" estimates, dashed lines are smoothed fits using cubic splines. Righthand plot: empirical catchability curves calculated using Equation 6. Red lines and dots: "raw" curves, dashed lines and fills: smoothed fits.

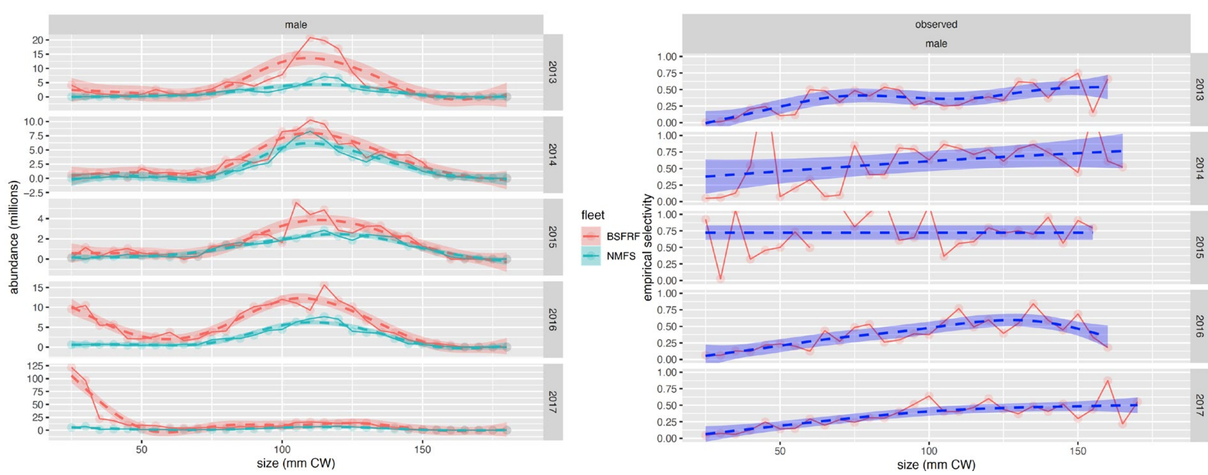

Figure 3.4.2. Lefthand plot: BSFRF (orange) and NMFS (green) size compositions for male Tanner crab from the SBS study areas for 2013-2017. Solid lines and dots are "raw" estimates, dashed lines are smoothed fits using cubic splines. Righthand plot: empirical catchability curves calculated using Equation 6. Red lines and dots: "raw" curves, dashed lines and fills: smoothed fits.

The empirical catchability curves estimated directly from the SBS data appear to imply a fair bit of interannual variability in size-specific catchability, as well as some indications that the shape of these curves is not really logistic (as assumed in the assessment model). However, the assessment model estimates a single sex-specific catchability curve for the entire 1982-2019 time period, so smooth, sex-specific GAM models were fit to the "raw" empirical catchability results using the R "mgcv" package with a log link to a normal error distribution model such that

$$\eta_z \sim N(s(z), \sigma_z^2), \text{ with } \ln\left(E_y[\hat{C}_{z,y}^{EBS}]\right) = \eta_z \tag{7}$$

where $\hat{C}_{z,y}^{EBS}$ is the "raw" empirical catchability estimate at size z in year y and $E_y[\cdot]$ is the expectation taken over y. The resulting mean curves and 80% confidence intervals are shown in Figure 3.4.3. Interestingly, these curves lend support to the use of logistic curves as a basis for describing NMFS survey size selectivity, with fully-selected catchability coefficients ("survey q's") of about 0.4 for females and 0.6 for males. Note that the mean curves could be used to fix NMFS survey catchability outside the assessment model, but that the mean and confidence intervals could also be used to define priors for catchability curves estimated within the assessment model. However, these priors might not reflect the full level of sampling variability inherent in the data.

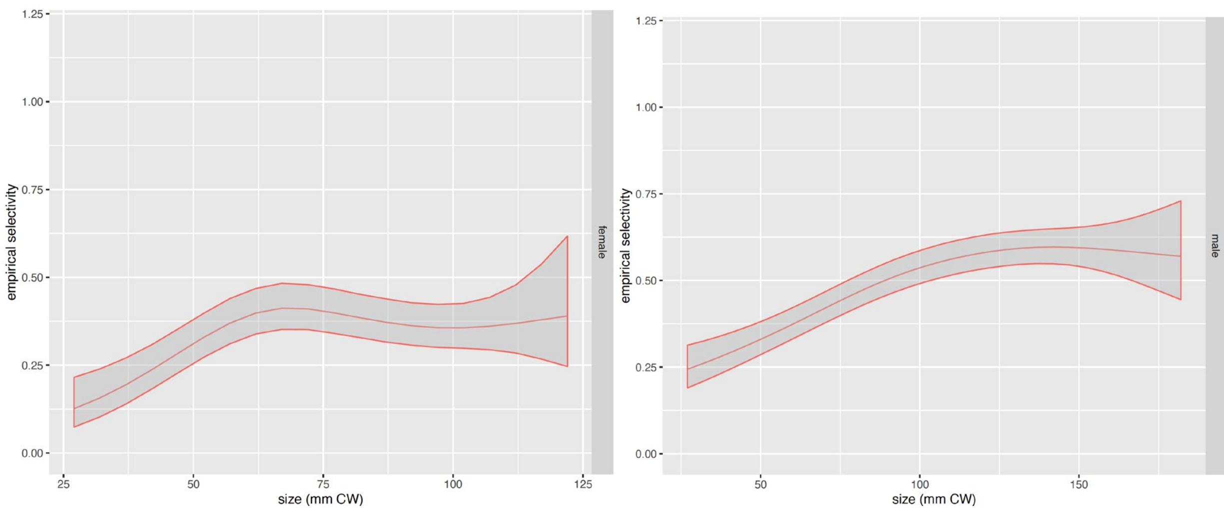

Figure 3.4.3. Estimated empirical catchability curves based on fitting the "raw" empirical catchability curves in Figures 3.4.1 (females) and 3.4.2 (males) using generalized additive models (GAMs; Wood 2017) assuming log link functions and normal distributions. The R package "mgcv" (Wood, 2011) was used to fit the GAMs. The center line in each plot represents the mean, the fills represent 80% confidence intervals. Note the difference in x-axis scales.

To address that concern, a bootstrapping analysis was undertaken to provide priors for NMFS survey catchability that more fully incorporated sampling variability into the analysis. To that end, the SBS data was resampled using a hierarchical approach similar in spirit to that discussed in the previous section. For each bootstrap replicate, n_y^{SBS} hauls were randomly selected with replacement from the paired SBS hauls in the SBS study area for year y. Following the paired haul selection, n_h^{BSFRF} individuals were randomly selected with replacement from the n_h^{BSFRF} individuals actually captured in the BSFRF haul, and n_h^{NMFS} individuals were randomly selected with replacement from the n_h^{NMFS} individuals actually captured in the associated NMFS haul. If paired hauls were selected multiple times, individuals from those hauls were re-sampled each time. Sex-specific BSFRF and NMFS size compositions were then calculated using 5 mm bins from 25 mm to 180 mm CW for each bootstrap realization using the standard, stratified EBS design-based approach, after which sex-specific "raw" catchabilities were calculated for each 5 mm size bin using Equation 6. Bootstrap replicates were drawn 1000 times for each SBS year.

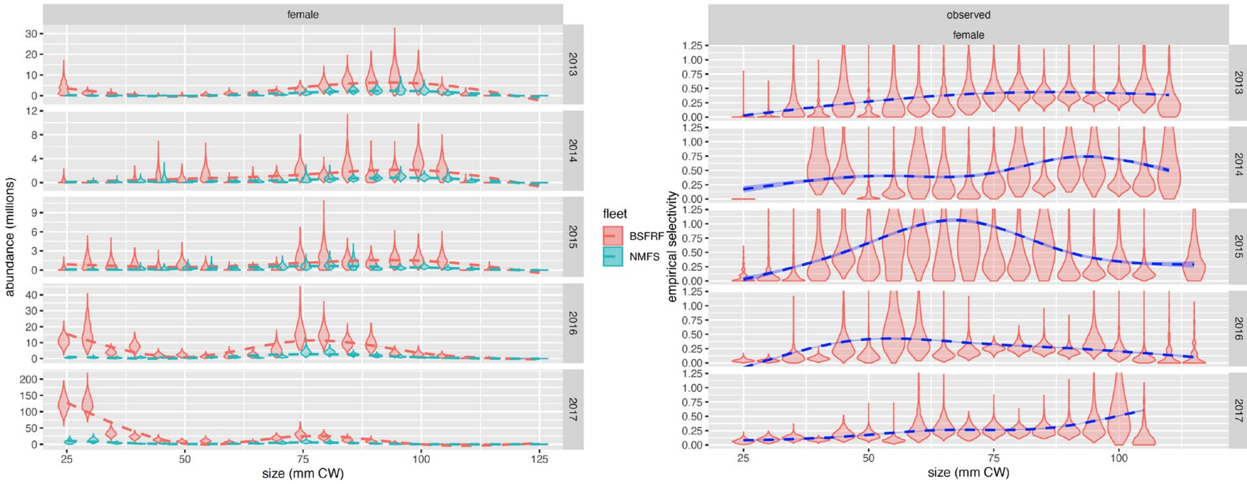

Figure 3.4.4. Lefthand plot: Bootstrapped BSFRF (orange) and NMFS (green) size compositions for female Tanner crab from the SBS survey areas for 2013-2017. Violin plots summarize "raw" estimates, dashed lines are smoothed fits using cubic splines. Righthand plot: bootstrapped empirical selectivity curves calculated using Equation 5. Red violin plots: "raw" estimates, dashed lines: smoothed fits.

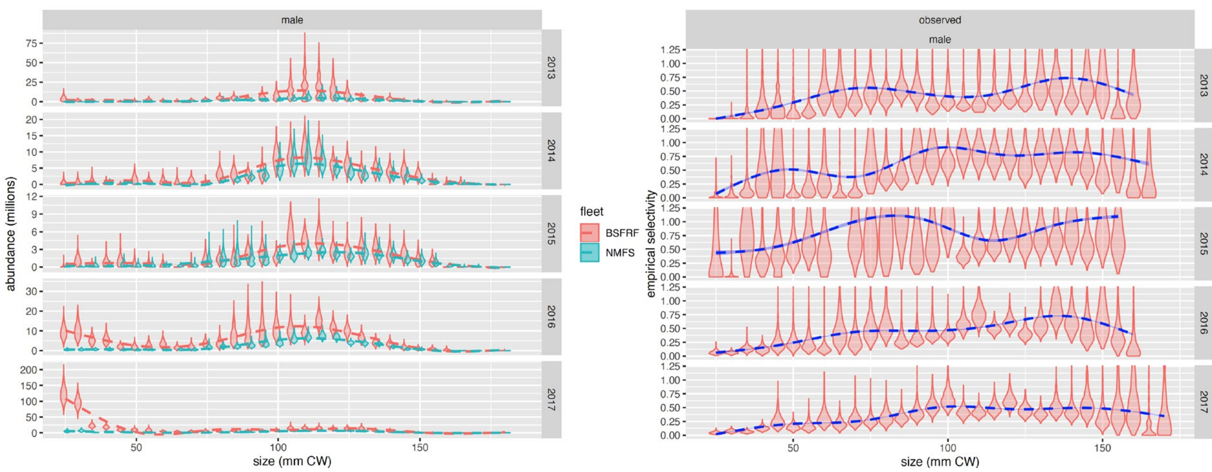

Figure 3.4.5. Lefthand plot: Bootstrapped BSFRF (orange) and NMFS (green) size compositions for female Tanner crab from the SBS survey areas for 2013-2017. Violin plots summarize "raw" estimates, dashed lines are smoothed fits using cubic splines. Righthand plot: bootstrapped empirical selectivity curves calculated using Equation 5. Red violin plots: "raw" estimates, dashed lines: smoothed fits.

Figures 3.4.4 and 3.4.5 summarize the results from this bootstrapping analysis, showing the variability of the estimates within each 5 mm size bin for each year, as well as a smoothed fit across the size bins. Not surprisingly, the smoothed fits are similar to those shown in Figures 3.4.1 and 3.4.2, but the scale of variability within any size bin (for any year) suggests that NMFS survey catchability for Tanner crab may not be well-estimated using this approach. As a consequence, any priors developed in the future based on this analysis may provide little constraint on catchability curves estimated within the assessment model.

3.5 VAST model-based estimates for survey biomass and abundance

Jon Richar (AFSC Kodiak) provided VAST (Vector Autoregressive Spatio-Temporal; Thorson, 2019) model-based estimates of Tanner crab survey abundance and biomass as alternatives to the survey design-based estimates that are currently fit in the assessment model. The time series are shown in Figures 3.5.1 for males and 3.5.2 for immature and mature females. The principal difference between the two approaches is the size of the confidence intervals associated with the estimates, with the VAST estimates having notably smaller confidence intervals in almost all years. In most cases, the VAST estimates agree with their design-based estimate counterparts. Notable exceptions include male biomass estimates for 1976-78, male abundance estimates for 1976, and mature female abundance in 1975. However, a number of estimates for mature female biomass are substantially different (i.e., outside the confidence intervals for the design-based estimates): 1980-1982, 1988-1992, 2006, and 2012-2015.

Figure 3.5.1 Comparison of VAST model-based estimates of NMFS male survey biomass and abundance with those from the standard design-based approach (labelled 19.03). The VAST estimates are labelled "20.04 VAST". Error bars denote 80% lognormal confidence intervals. Note: graphs in righthand column focus on 2004-2019.

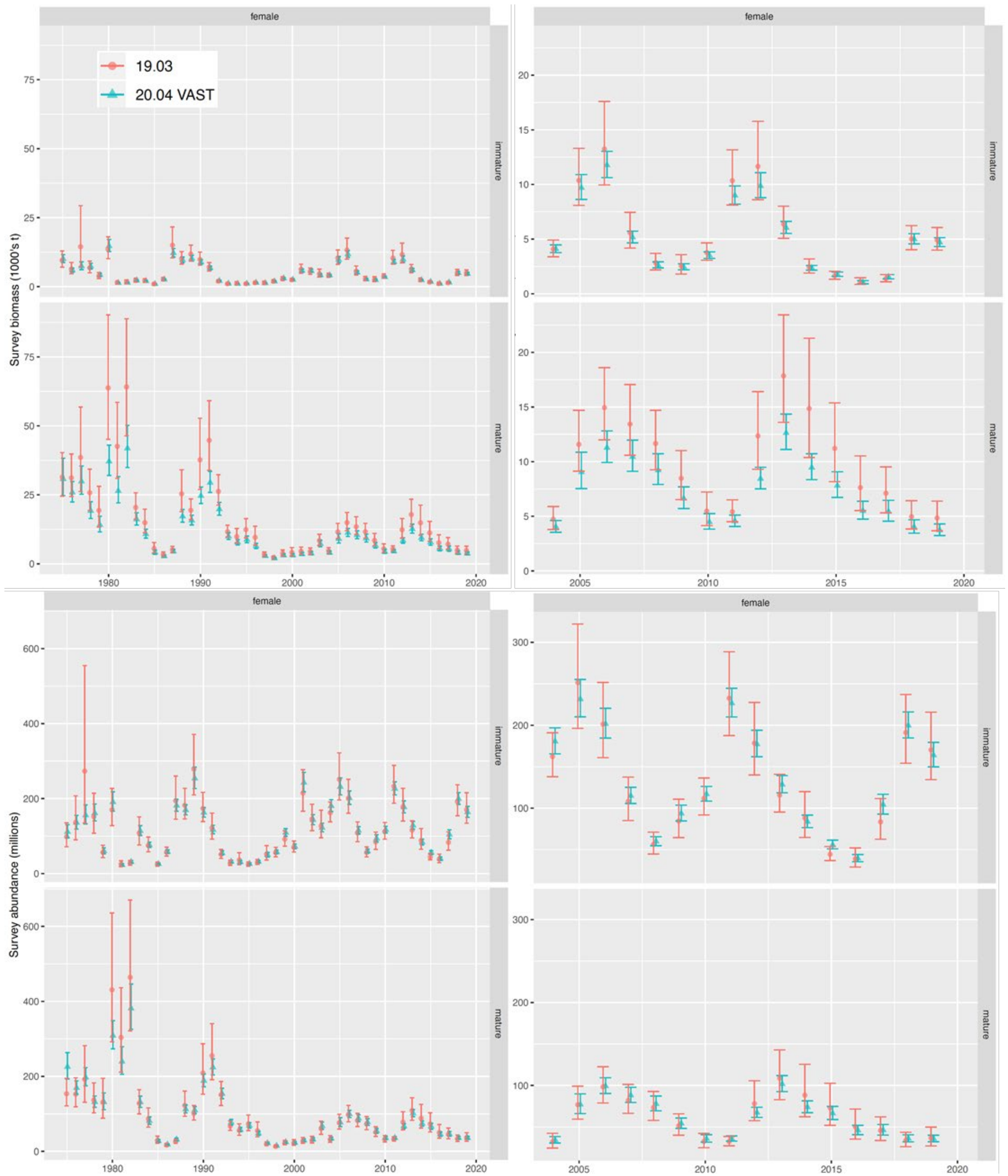

Figure 3.5.2. Comparison of VAST model-based estimates of NMFS survey female biomass and abundance with those from the standard design-based approach (labelled 19.03). The VAST estimates are labelled “20.04 VAST”. Error bars denote 80% lognormal confidence intervals. Note: graphs in righthand column focus on 2004-2019.

4.0 Model development since Sept 2019

Several new features have been added to TCSAM02, the Tanner crab assessment model code. These include:

1. the ability to programmatically specify a retrospective model run (i.e., running the model with a specified number of the most recent years of data and associated parameters excluded from the model fit and estimation)
2. an option to estimate selectivity/availability curves based on cubic splines
3. an option to apply selectivity (catchability) and/or availability curves estimated outside the model to survey or fishery data
4. an option to apply prior probabilities determined outside the model to selectivity (catchability) and/or availability curves estimated inside the model
5. an option to estimate "additional uncertainty" parameters associated with a survey

The code is available on GitHub. The current development branch is "202005CPT. Model scenarios incorporating features 2, 3, and 5 are evaluated as part of this report. As noted in Sections 3.3 and 3.4, prior probabilities have not yet been completely developed (although the process is underway) to apply to selectivity (catchability) and/or availability curves estimated inside the model, so scenarios using feature 4 above could not tested at this point.

5.0 Model scenarios

Eight new model scenarios were considered for this report, as outlined in the following table:

Scenario	Parameters	Progression	Description
19.03	343	--	Accepted model for the 2019 Tanner crab assessment (identified as M19F03 in Stockhausen, 2019.)
20.01 (RecZCs)	345	19.03 +	Recruitment size distribution estimated
20.02 (TruncSrv)	339	20.01 +	NMFS surveys 1982-2019 only
20.03 (CbSpls)	359	19.03 +	NMFS survey selectvity estimated using cubic splines
20.04 (VAST)	343	19.03 +	VAST estimates for NMFS surrvey abundance and biomass
20.05 (VAST+XU)	347	20.04 +	Additional survey uncertainty estimated
20.06 (SBS)	610	20.01 +	SBS NMFS and BSFRF biomass time series (revised) and size comp.s
20.07 (SBS+FACs)	345	20.01 +	SBS BSFRF biomass time series (revised) and size comp.s, with availability fixed from SBS studies
20.08 (SBS+FCCs)	339	20.01 +	sex/size-specific EBS NMFS survey catchability fixed using selectivity from SBS studies

The first three model scenarios are direct variations on the 2019 assessment model (Stockhausen, 2019), designated 19.03 here in keeping with the SSC's requested model numbering conventions (it was referred to as M19F03 in Stockhausen, 2019). In 19.03 (and previous models), the size composition for recruits entering model each year has been described using a gamma probability distribution with fixed shape and scale parameters which was applied in all model years. In contrast, Scenario 20.01 estimates the shape and scale parameters of the gamma distribution to better characterize the average size distribution of recruits. 19.03 (and all scenarios considered here other than 20.02) included NMFS survey biomass and size composition data for surveys from 1975 to 1981 in the model optimization; however, this data was collected using a different gear type and varying spatial coverage—which necessitates estimating different selectivity curves and catchability coefficients to describe the survey data from this time period. Excluding this data simplifies the model optimization by reducing the number of parameters and possibly removing conflicting data. The utility of this is explored in Scenario 20.02, which was otherwise identical to 20.01 but excluded fits to NMFS survey biomass and size composition data for surveys prior to 1982 in the model optimization. 19.03 also used logistic functions to describe NMFS survey selectivity in both time periods, but this may be overly restrictive in terms of the assumed shapes of the curves. Scenario 20.03 explores the utility of cubic splines to more flexibly describe NMFS survey selectivity in the 1975-1981 and 1982-2019 time periods.

Model-based estimates of survey biomass using VAST (Vector Autoregressive Spatio-Temporal; Thorson, 2019) typically have higher precision (smaller CVs) than standard designed-based estimates for the NMFS EBS survey, and VAST's ability to "fill in" missing data may be particularly helpful in more accurately characterizing survey biomass for Tanner crab in the 1975-1981 time period. Scenario 20.04 includes fits to VAST model-based estimates for biomass from the NMFS surveys (1975-2019) rather than the design-based estimates fit in 19.03 (and all other scenarios except 20.02 and 20.05). Note, however, that the size compositions fit in all scenarios are the same design-based estimates. The higher precision of the VAST estimates can make it problematic for an assessment model to fit them. One approach taken in several groundfish assessments that use VAST estimates in the model optimization is to estimate parameters reflecting "additional" uncertainty associated with these estimates, effectively downweighting these estimates relative to other data sources fit in a model. Scenario 20.05 (VAST+XU) extends 20.04 by estimating 4 new parameters reflecting additional ("extra") uncertainty associated with the VAST biomass estimates for males/females in two survey time periods (1975-1981 and 1982-2019). These parameters effectively inflate the variance associated with the VAST biomass estimates, making it "easier" for the model to fit the estimates

Scenario 20.06 is based on 20.01 (i.e., it fits the design-based estimates of NMFS survey biomass as in 19.03 as well as estimates the size distribution at recruitment), but it also fits BSFRF (Bering Sea Research Foundation) and NMFS survey estimates of biomass and size compositions from the side-by-side (SBS) selectivity studies conducted in 2013-2017. In addition to estimating parameters to describe the assumed logistic selectivity of the full (EBS) NMFS survey, as in 20.01, Scenario 20.06 estimates 265 additional parameters to describe size-specific annual availability of Tanner crab to the BSFRF and NMFS gear in the SBS study area in a "non-parametric" fashion: each annual (2013-2017) availability "function" is described by a sex-specific number of parameters (males: 32; females: 21) with constraints on the smoothness of the resulting curves applied as penalties in the likelihood (thus reducing the degrees of freedom/effective number of parameters). The BSFRF gear is assumed to be non-selective, with catchability equal to 1. Scenario 20.07 replaces the estimated non-parametric availability curves with fixed availability curves (FACs) determined empirically outside the assessment model using the NMFS EBS and SBS data (see Section 3.3 above). The use of FACs in this scenario eliminates the need to fit the NMFS SBS data in the assessment model. Finally, Scenario 20.08 essentially goes one step further than 20.07 and uses fixed catchability curves (FCCs) determined empirically outside the model from the NMF and BSFRF SBS data (see Section 3.4 above) to describe the sex/size-specific catchability (selectivity *x* fully-selected catchability) of the full (EBS) NMFS survey during the 1982-2019 time period. Use of the FCCs eliminates the need to fit either the BSFRF or NMFS SBS data in the model, because these data were used to determine the empirical curves outside the model.

6.0 Model results

6.1 Population-level quantities and processes for all model scenarios

All model scenarios were evaluated using 1000+ model runs and jittering of initial parameter values to increase confidence that the converged model exhibiting the smallest objective function value had truly arrived at the global, rather than a local, minimum. All scenarios achieved minima that exhibited small maximum gradients (Table 6.1) and whose hessian was invertible to obtain approximate estimates of parameter uncertainty. To some extent, though, all of the scenarios considered here (including 19.03, the base scenario) are somewhat disappointing in one fashion or another.

Table 6.1 Summary of model scenario results. *N* refers to the number of estimated parameters. Except for the group {19.03, 20.01, 20.03}, the objective function values for the converged models are not directly comparable between scenarios.

model scenario	N	objective function value	max gradient	OFL (1000's t)	Fofl	projected B (1000's t)	current B (1000's t)	Fmsy	Bmsy (1000's t)	MSY (1000's t)	unfished B (1000's t)	average recruitment (millions)
19.03	343	3228.46	1.35E-04	29.5	1.12	39.7	82.6	1.18	41.6	19.5	119.0	393.8
20.01 (RecZCs)	345	3202.35	5.38E-03	30.4	1.17	40.0	84.1	1.24	42.1	19.8	120.4	473.3
20.02 (TruncSrv)	339	3227.47	9.96E-05	23.7	0.89	36.7	71.3	0.90	37.2	16.5	106.4	336.5
20.03 (CbSpls)	359	2975.18	2.59E-04	79.1	2.31	75.6	185.3	2.54	82.2	40.1	234.8	988.2
20.04 (VAST)	343	4069.76	9.62E-04	16.0	0.95	25.1	50.2	1.12	29.2	15.1	83.3	295.7
20.05 (VAST+XU)	347	2783.07	3.68E-04	21.6	1.07	36.5	67.2	1.42	47.2	19.1	134.7	407.4
20.06 (SBS)	610	3367.16	9.02E-04	26.0	1.07	36.5	74.7	1.17	39.5	18.6	112.9	439.9
20.07 (SBS+FACs)	345	3506.61	1.47E-04	23.0	0.94	35.2	69.3	1.00	37.5	17.2	107.0	379.0
20.08 (SBS+FCCs)	339	3229.69	1.22E-04	18.3	0.82	30.5	57.6	0.87	32.0	14.0	91.4	266.2

Results from groups of closely-related scenarios are discussed in more detail in subsequent sections.

6.2 VAST Scenarios 20.04 and 20.05

Predicted survey biomass based on fits to the standard design-based estimates of NMFS survey biomass and the VAST estimates of survey biomass in Scenarios 19.03 and 20.04, respectively, follow relatively similar trends (Figure 6.2.1). The trends predicted in Scenario 20.05, which fitted the VAST estimates but also estimated parameters related to "extra" uncertainty in those estimates are, however, extremely different in scale from those of 19.03 and 20.04—even though the associated RMSEs (root mean square error) for mature males and mature females are smaller for Scenario 20.05 (both ~1) than for either 19.3 or 20.04 (~2.4 and 3.4, respectively), suggesting the 20.05 fits the VAST data quite well. These apparent contradictions are the consequence of two of the four estimated parameters related to extra uncertainty being estimated at their upper bounds in the converged 20.05 scenario (1.0 on the ln-scale, corresponding to extra CVs of 270%). This level of extra uncertainty is clearly too large, and so further consideration of results from Scenario 20.05 is dropped from this report. The reasons that this approach failed are currently under investigation.

Although comparison of the survey biomass likelihood components for 19.03 (191.6 likelihood units) and 20.05 (687.7 units) would indicate the model fit to the VAST data is much worse than that to the design-based data, the predicted trajectories in 20.05 appear to follow the temporal fluctuations in the VAST survey biomass estimates more closely than in those in 19.03 follow the design-based estimates (Figure 6.2.2). Both of these aspects are driven by the "shrinkage" exhibited in the VAST estimates toward the mean: the smaller VAST confidence intervals "encourage" the model to follow VAST estimates more closely (relative to other data) while inflating the "importance" of the mismatch between observed estimates and model predictions in the likelihood/objective function. Finding a balance to these two rather contradictory effects was the purpose in attempting to estimate "extra" uncertainty in 20.05, but this was obviously not successful (at least by the result achieved here). In general, 20.05 differs little from 19.03 in its fits to the aggregated survey size compositions (Figure 6.2.3), although it fits the size compositions in 1986 and 1987 much more poorly (at least 9 likelihood units difference in each year) for each sex compared to 19.03. It seems a bit strange, then, that the two scenarios would estimate similarly-shaped selectivity curves in the 1982-2019 period for both sexes and for females in the 1975-1981 time period, but differ rather substantially in the shapes of the males in the 1975-1981 time period (Figure 6.2.4).

Fits to retained and total catch biomass in the fisheries are almost identical for the two scenarios (Figures 6.2.5 and 6.2.6), as are estimated the selectivity curves (Figure 6.2.7).

Estimated natural mortality rates are slightly higher in 20.05 compared with 19.03 in the majority of the model period (mature female rates are lower during the "enhanced mortality" period 1980-1984; Figure 6.2.8). The estimated size-specific probability of terminal molt for females is identical for the two scenarios, while growth rates are slightly larger for males are slightly larger, and slightly smaller for females, in 20.05 compared with 19.03 (Figure 6.2.8).

The timing of estimated fluctuations in recruitment are almost identical for the two scenarios after 1988, with 19.03 exhibiting higher estimates when recruitment peaked (Figure 6.2.9). Prior to 1989, the trends are somewhat more different, with 19.03 exhibiting generally higher levels, except in the late 1960's and 1985. Fluctuations in population biomass are almost synchronous between the two scenarios, 19.03 exhibiting somewhat higher values.

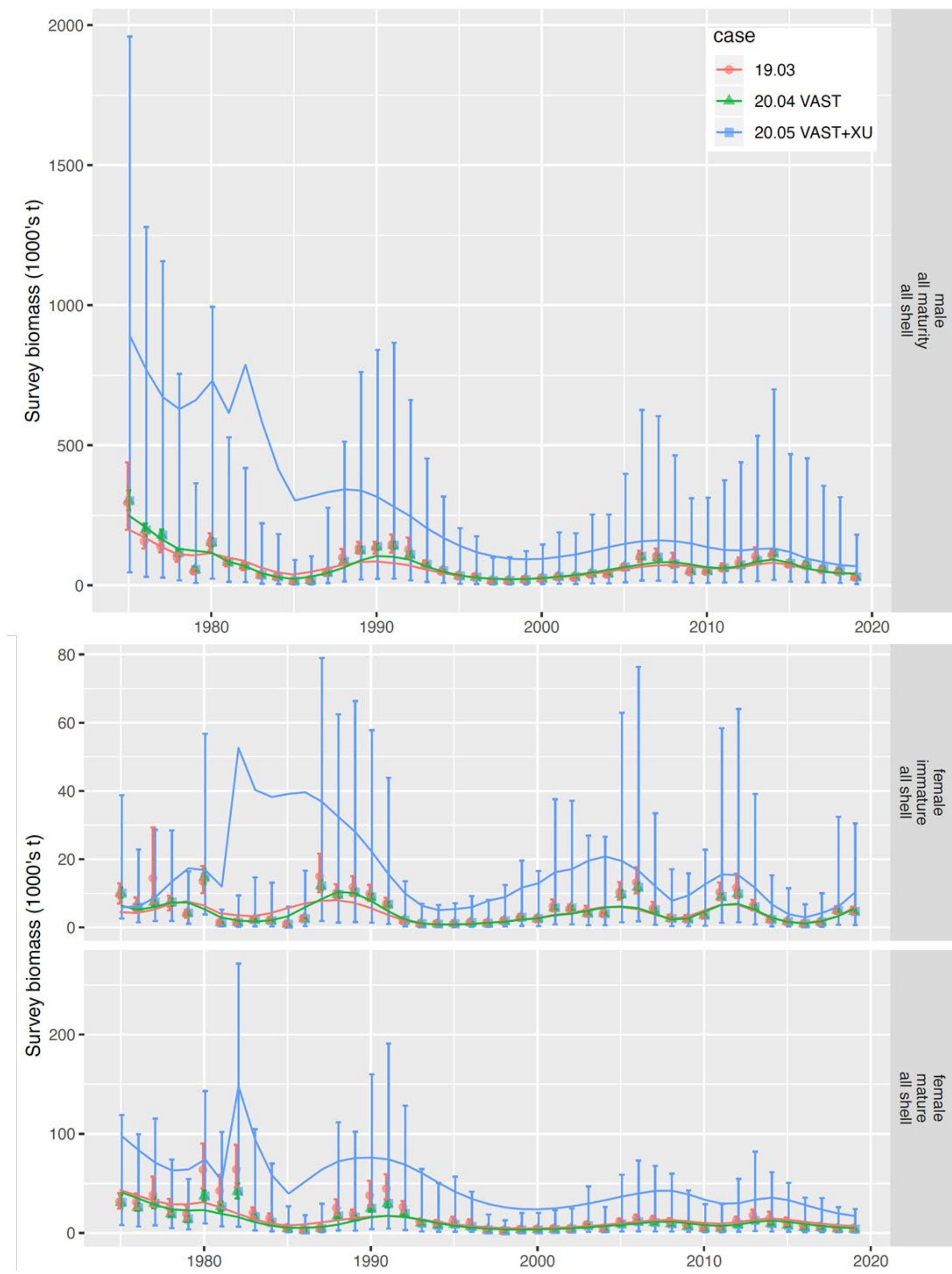

Figure 6.2.1. Comparison of model fits to design-based (19.03) and VAST-based (20.04, 20.05) survey biomass for Scenarios 19.03, 20.04, and 20.05. Confidence intervals are 80%; confidence intervals for 20.05 include estimated "extra" uncertainty.

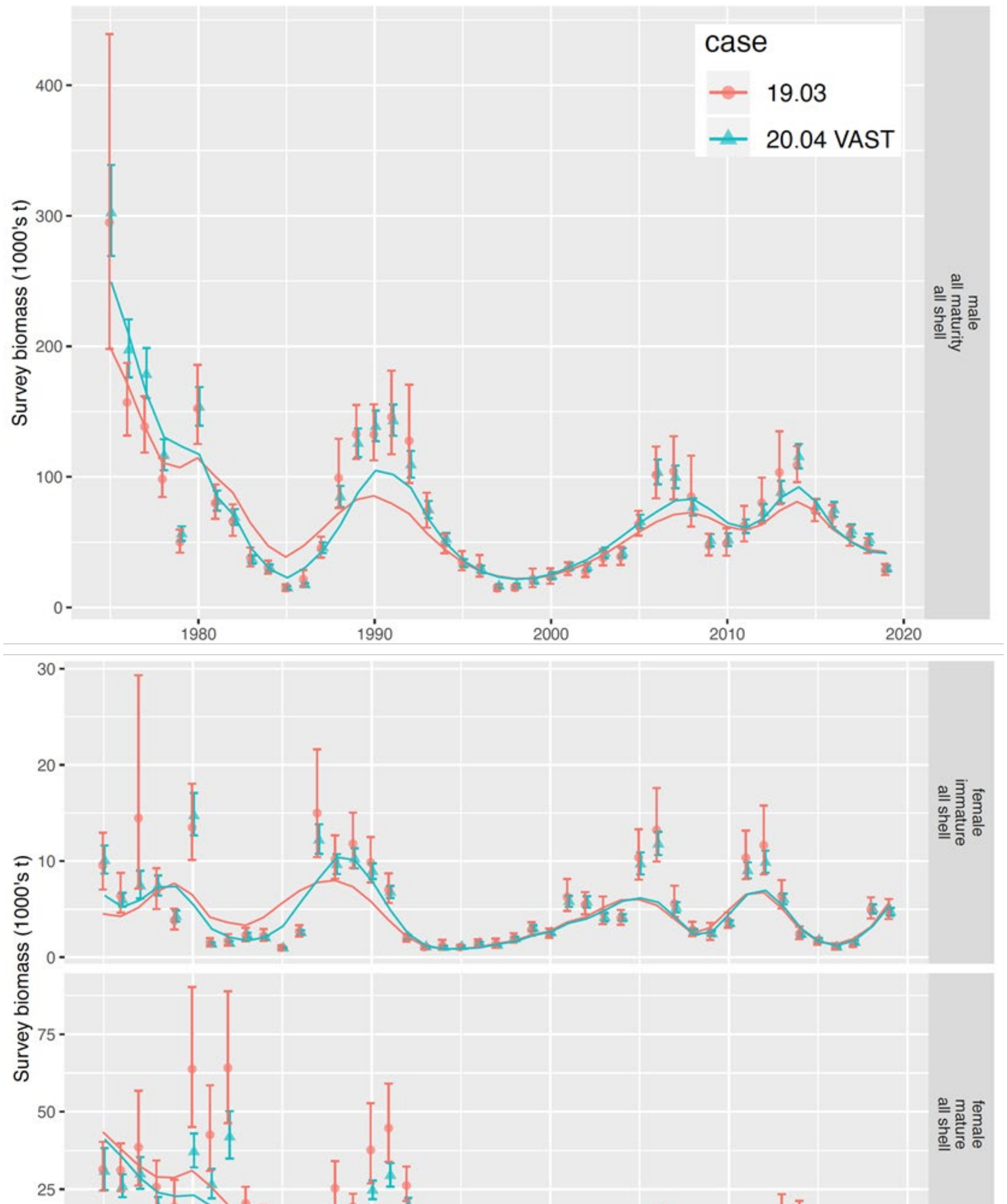

Figure 6.2.2. Comparison of model fits to design-based (19.03) and VAST-based (20.04) survey biomass for Scenarios 19.03 and 20.04.

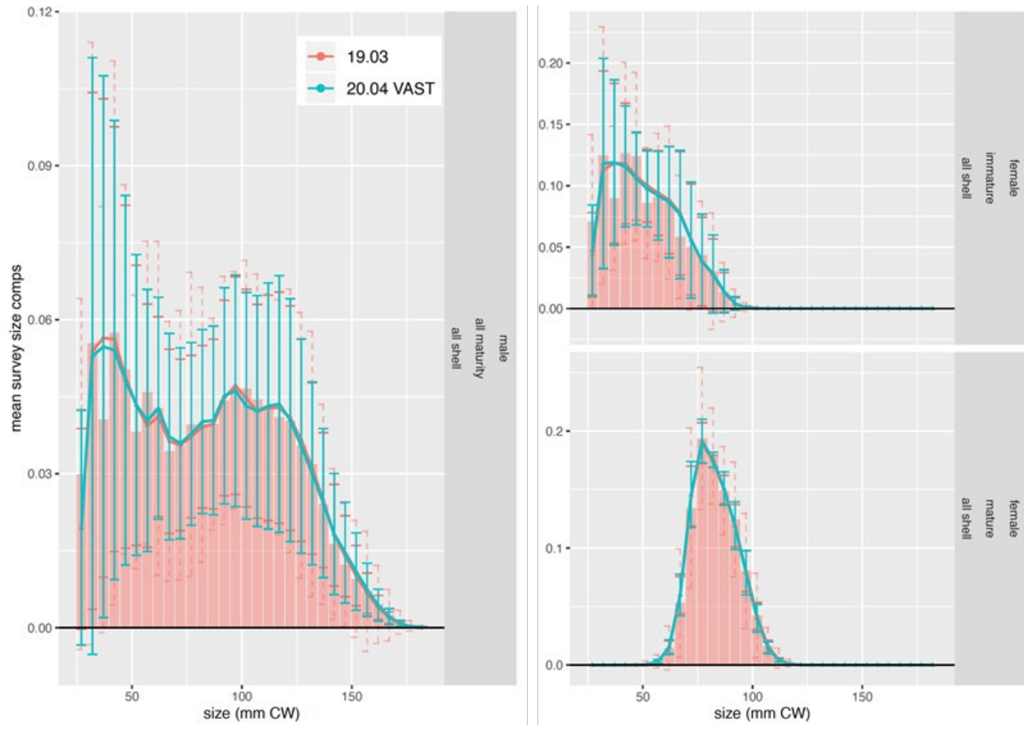

Figure 6.2.3. Comparison of averaged model fits to survey size compositions for Scenarios 19.03 and 20.04.

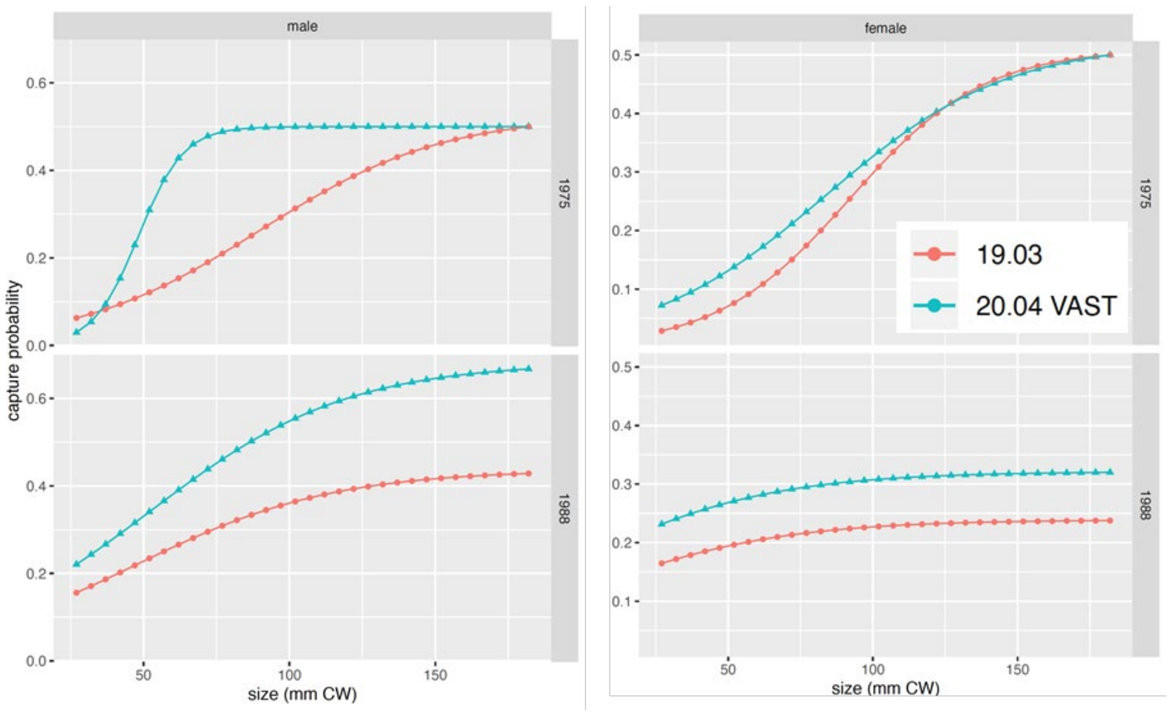

Figure 6.2.4. Comparison of NMFS survey catchability for Scenarios 19.03 and 20.04.

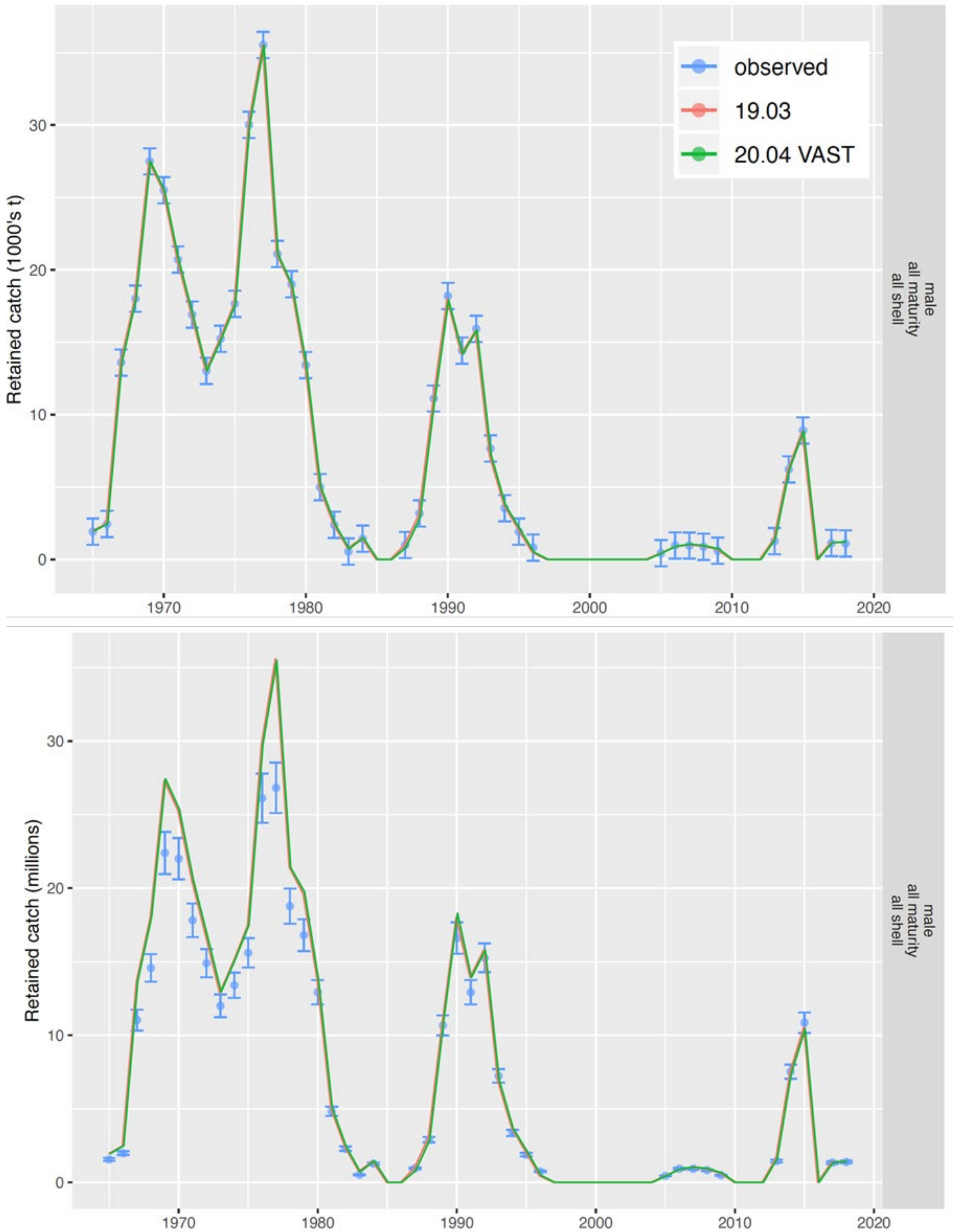

Figure 6.2.5. Comparison of model fits to retained catch biomass and abundance for Scenarios 19.03 and 20.04.

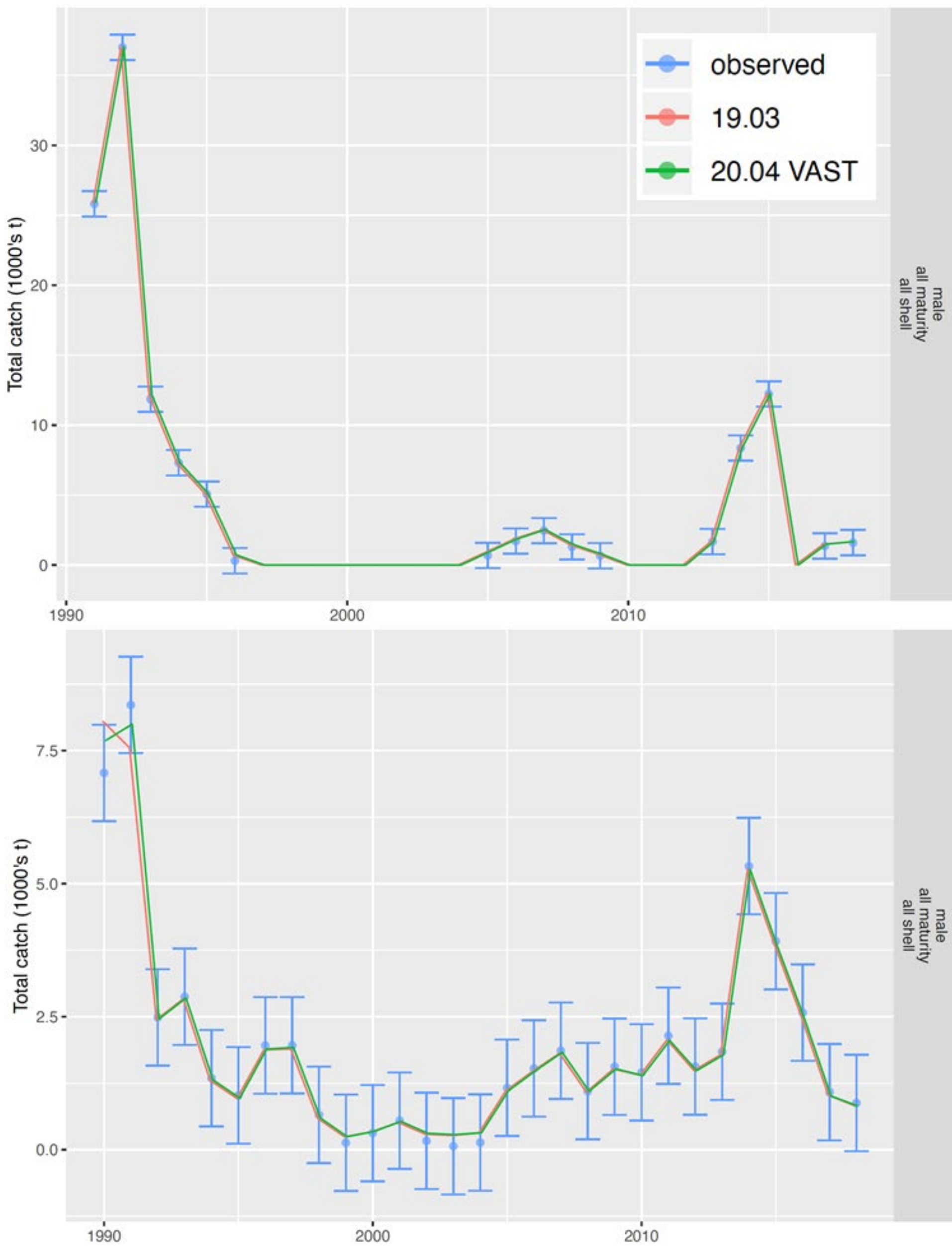

Figure 6.2.6. Comparison of model fits to total catch biomass in the directed (TCF) and snow crab (SCF) fisheries for Scenarios 19.03 and 20.04.

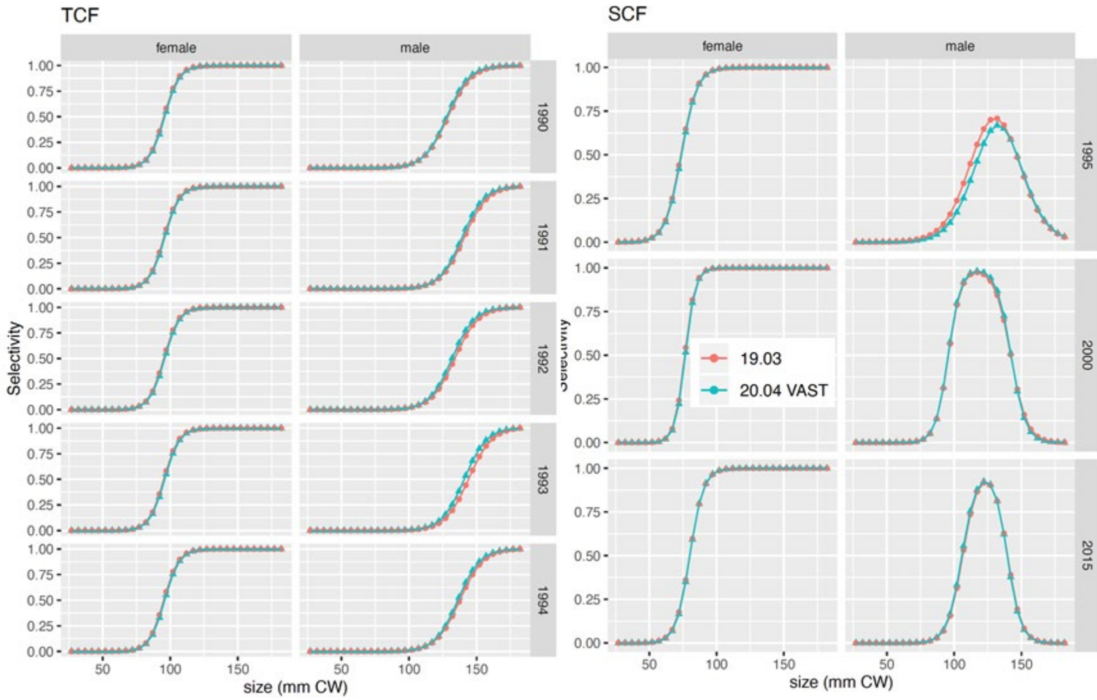

Figure 6.2.7. Comparison of estimated fishery total catch selectivity curves in several time periods for the directed Tanner crab fishery (TCF) and snow crab fishery (SCF).

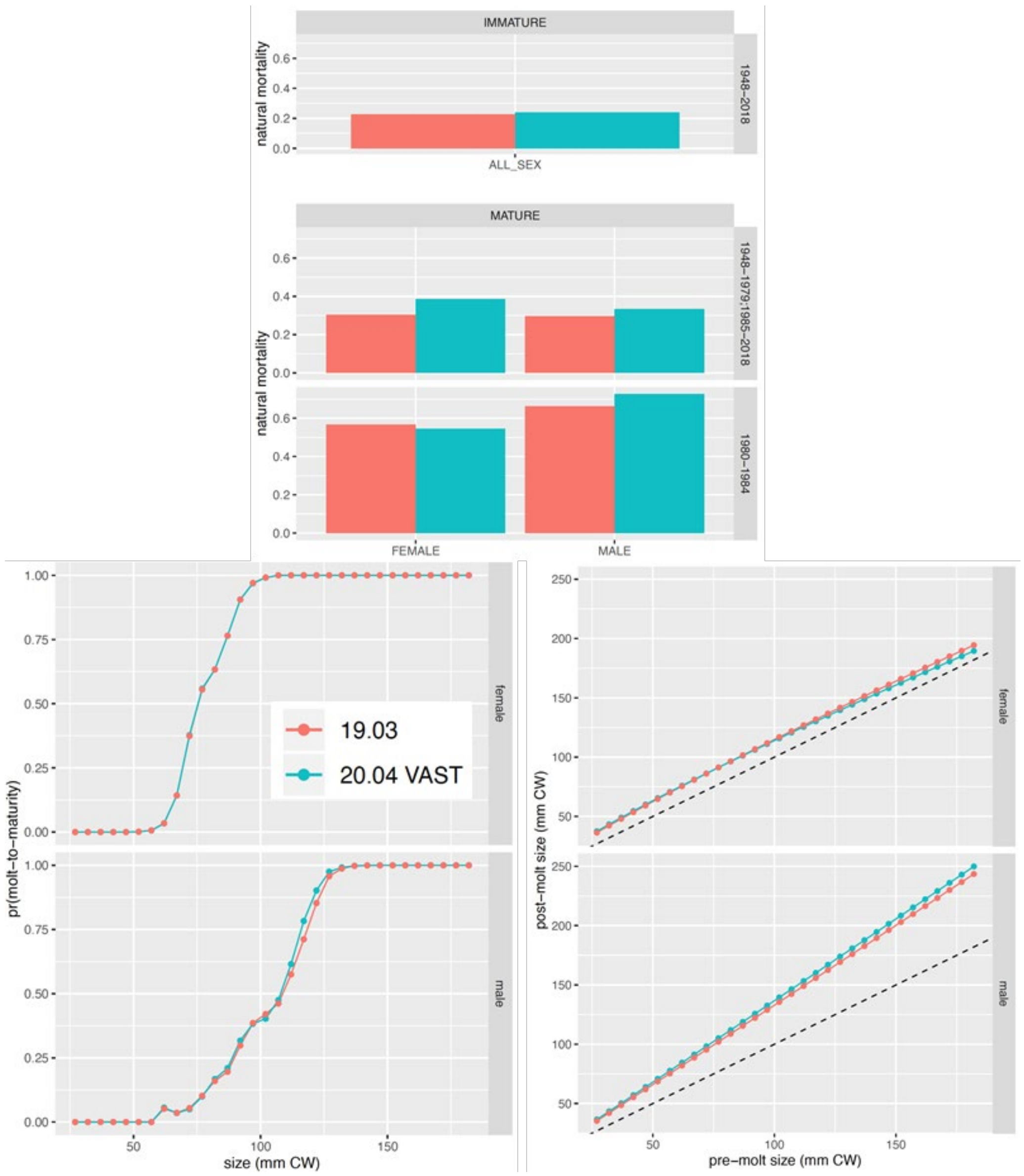

Figure 6.2.8. Comparison of estimated natural mortality rates (top), probabilities of terminal molt (lower left) and post-molt size (lower right) for Scenarios 19.03 and 20.04.

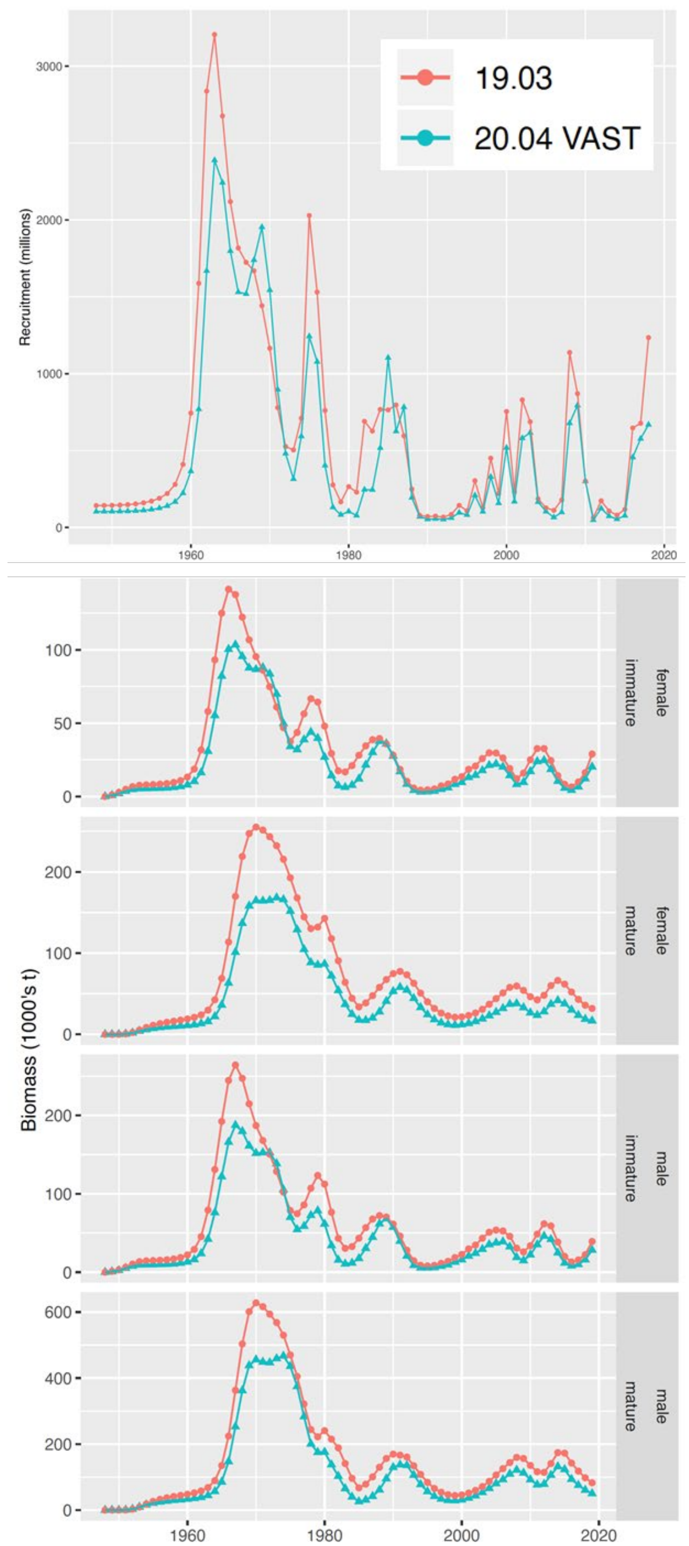

Figure 6.2.9. Comparison of recruitment and population biomass trends from Scenarios 19.03 and 20.04.

6.3 20.01 RecZCs, 20.02 TrncSrvs and 20.03 CbSpls

Scenarios 20.01, 20.02, and 20.03 consider the effects of estimating the parameters determining the size distribution at recruitment (all three scenarios), excluding the oldest survey data collected using different gear and areal coverage (20.02), and estimating survey selectivity curves not constrained to be logistic (20.03).

Fits to survey biomass are almost identical across all four scenarios after 1995 (Figure 6.3.1), and for 19.03 and 20.01 in particular across the entire time period. 20.02 generally estimates somewhat higher survey biomass during 1975-1981 than do the other scenarios, but 20.02 is not fitting the survey data in this time period. 20.03 fairly consistently has the lowest estimates for mature biomass across the entire time period (1975-2019). All four scenarios predict somewhat different aggregate survey size compositions below ~70 mm CW, while only 20.03 exhibits differences from the other scenarios at larger male crab sizes (Figure 6.3.2). 20.03 overpredicts the aggregate proportions of the largest males (>145 mm CW) in the survey.

The similarity in fits to both survey biomass and size compositions among all four scenarios is rather surprising given the substantially different survey catchability curves estimated in the four scenarios (Figure 6.3.3). 19.03 and 20.01 estimate quite different logistic catchability curves for females after 1981, with the one from 19.03 being almost constant with size, but the estimated curves for males and for females prior to 1982 are quite similar. The catchability curves estimated in 20.02 after 1982 tend to differ substantially from those in 20.01 only in scale (the curves shown for 20.02 prior to 1982 were not estimated).

Fits to retained and total catch biomass in the fisheries are almost identical for the four scenarios (Figures 6.3.4-5). The selectivity curves are also almost identical for scenarios 19.03, 20.01 and 20.02, as well, but they are right-shifted by ~5 mm CW for all but females in the snow crab fishery in Scenario 20.03 (Figure 6.3.6).

Recruitment size distributions were fixed (and equal) in Scenarios 19.03 and 20.03, while 20.01 and 20.02 estimated distributions that where higher proportions were slightly left-shifted to smaller sizes (Figure 6.3.7). Estimated natural mortality rates were fairly similar across all four scenarios for immature crab and for mature males and females in most of the model time period, but the estimated rate for mature males in the “enhanced mortality” period (1980-1984) was much higher in Scenario 20.03 (and lower in 20.02) than the other scenarios (Figure 6.3.7). The estimated size-specific probability of terminal molt for females was almost identical across the four scenarios, with slightly more variation for males (males in 20.02 were slightly more likely to remain immature at sizes < 115 mm CW but slightly more likely to undergo terminal molt at sizes greater than this, while the reverse was true for 20.03; Figure 6.3.7). Estimated mean growth rates were slightly smaller for males in 20.03 but otherwise similar for both sexes across all four scenarios (Figure 6.3.7).

The timing of estimated fluctuations in recruitment is almost identical for the four scenarios across the entire model time period—the scenarios differ primarily in scale, with mean recruitment substantially higher in Scenario 20.03 (Figure 6.2.9). Recruitment is slightly lower in 20.02 than the 19.03 and 20.01 prior to 1990, but afterward is essentially the same as these two scenarios. Estimated trends in population biomass basically show the same differences among scenarios as recruitment.

Of the three alternative scenarios examined in this section, 20.01 and 20.02 seem to be reasonable alternatives to 19.03. Estimating the two additional parameters governing the size distribution improves the model fit in 20.01 by 26 likelihood units over that in 19.03 (Table 6.1), although most of this change is due to a decrease in ~50 likelihood units in the priors on survey Q for females. The time frame for fitting to NMFS survey data in 20.03 (1982-2019) is consistent with that for snow crab and avoids issues with inconsistent gear types and survey coverages. On the other hand it does exclude an important part of the available time series for Tanner crab when abundance was apparently very high. However, the incorporation of VAST model-based estimates for survey abundance and biomass (Scenario 20.04), rather than the current (19.03) design-based estimates, presumably ameliorates some of the potential inconsistencies in the observed time series due to changes in survey coverage in the pre-1982 survey period, as well as some of the gear effects and thus reduces any advantages to moving forward with 20.02. Scenario 20.03 exhibited unlikely differences in population scale, although not the timing of fluctuations or trends, relative to the other scenarios due to the way the cubic splines used to estimate survey catchability were constrained at large crab sizes to asymptote to 1. This was done partly to allow interpretation of estimated Q's as "fully-selected" Q's, but the resulting curves do not seem reasonable and the implementation details need to be reconsidered. However, estimating less constrained parametric functions than the logistic functions used in 19.03 to describe survey catchability seems a worthwhile endeavor, so model scenarios like 20.03, although not recommended as a candidate for the Fall 2020 assessment, should remain areas for future development.

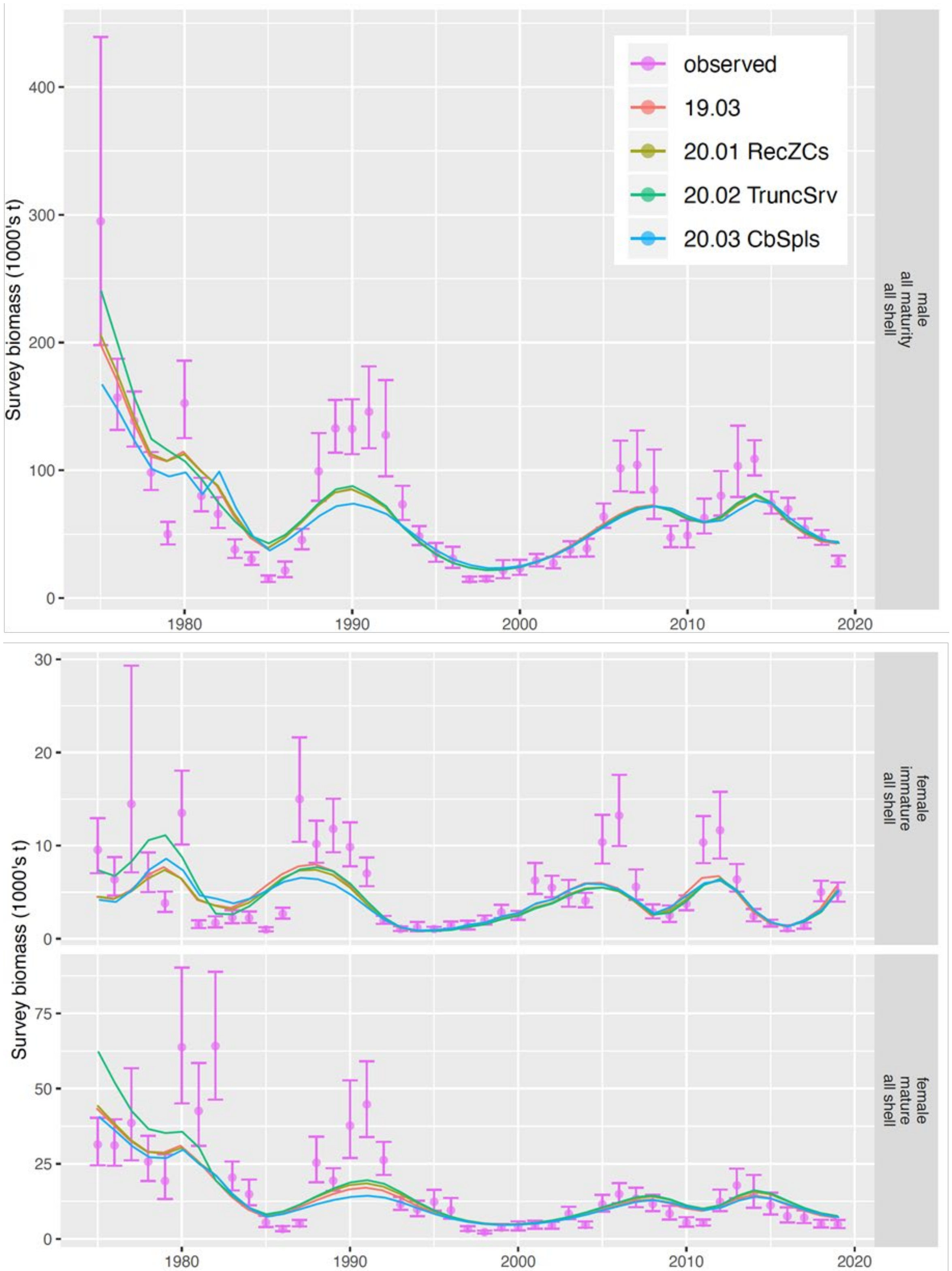

Figure 6.3.1. Comparison of NMFS survey biomass fits for Scenarios 19.03, 20.01, 20.02, and 20.03. Notes: the fits in 20.02 exclude the 1975-1981 data.

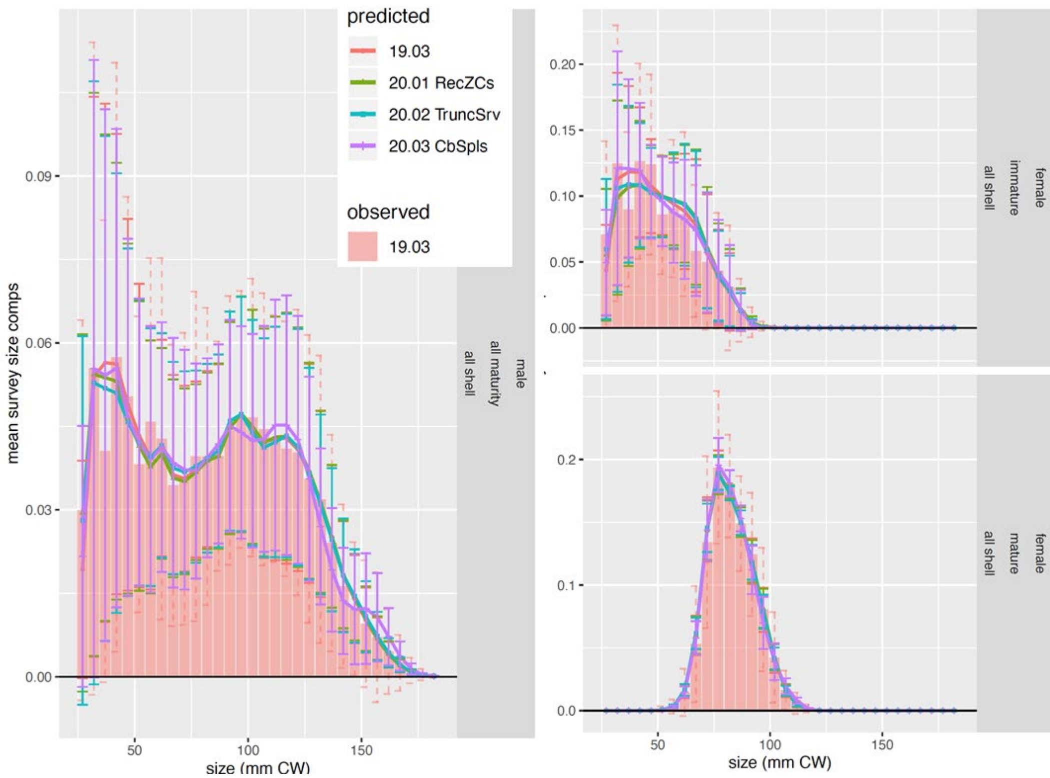

Figure 6.3.2. Comparison of mean predicted size compositions to observed compositions for Scenarios 19.03, 20.01, 20.02, and 20.03.

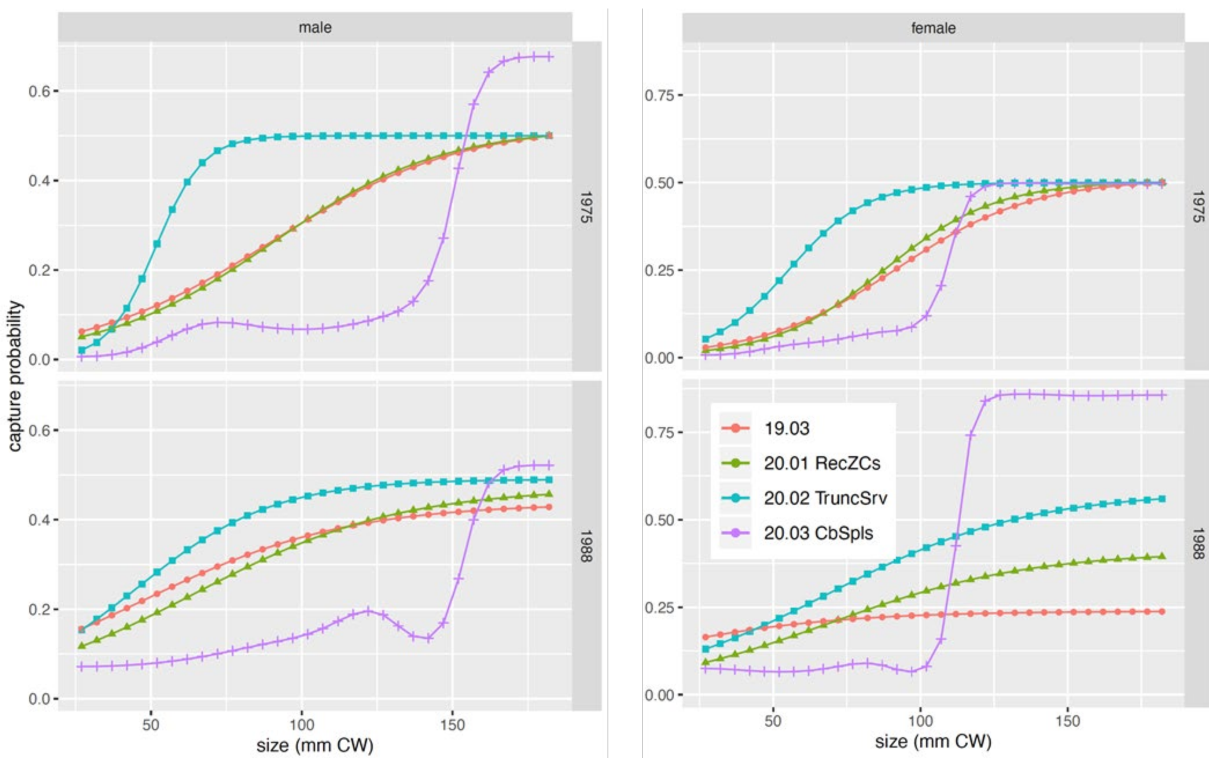

Figure 6.3.3. Comparison of estimated survey catchability curves for Scenarios 19.03, 20.01, 20.02, and 20.03. Notes: 1) the curves for Scenario 20.02 in 1975 were not estimated; 2)the cubic spline curves have been forced to asymptote to a constant (females > 125 mm CW; males > 170 mm CW).

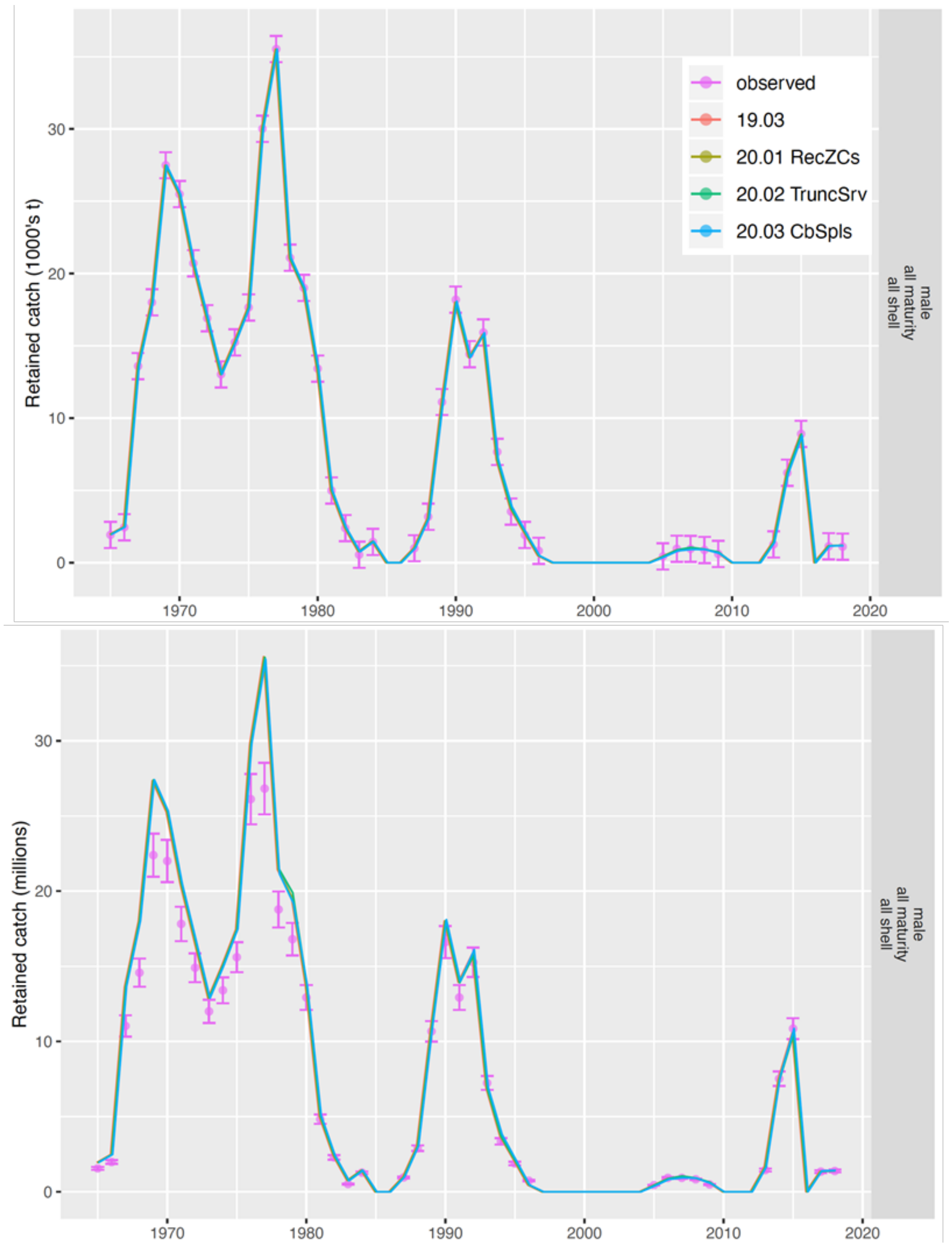

Figure 6.3.4. Comparison of fits to retained catch biomass and abundance in the directed fishery for Scenarios 19.03, 20.01, 20.02 and 20.03.

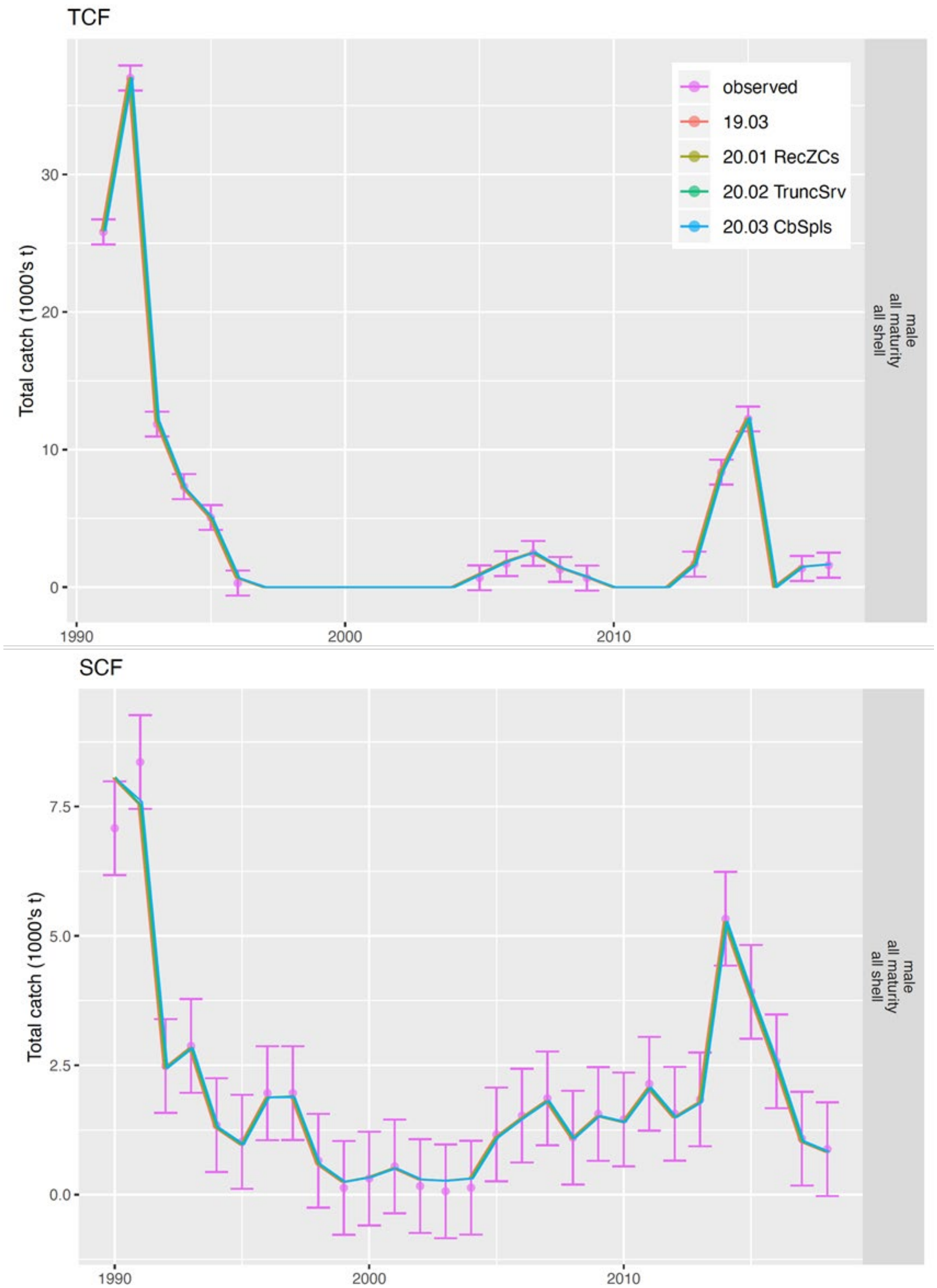

Figure 6.3.5. Comparison of fits to total catch biomass in the directed fishery (TCF) and the snow crab fishery (SCF) for Scenarios 19.03, 20.01, 20.02 and 20.03.

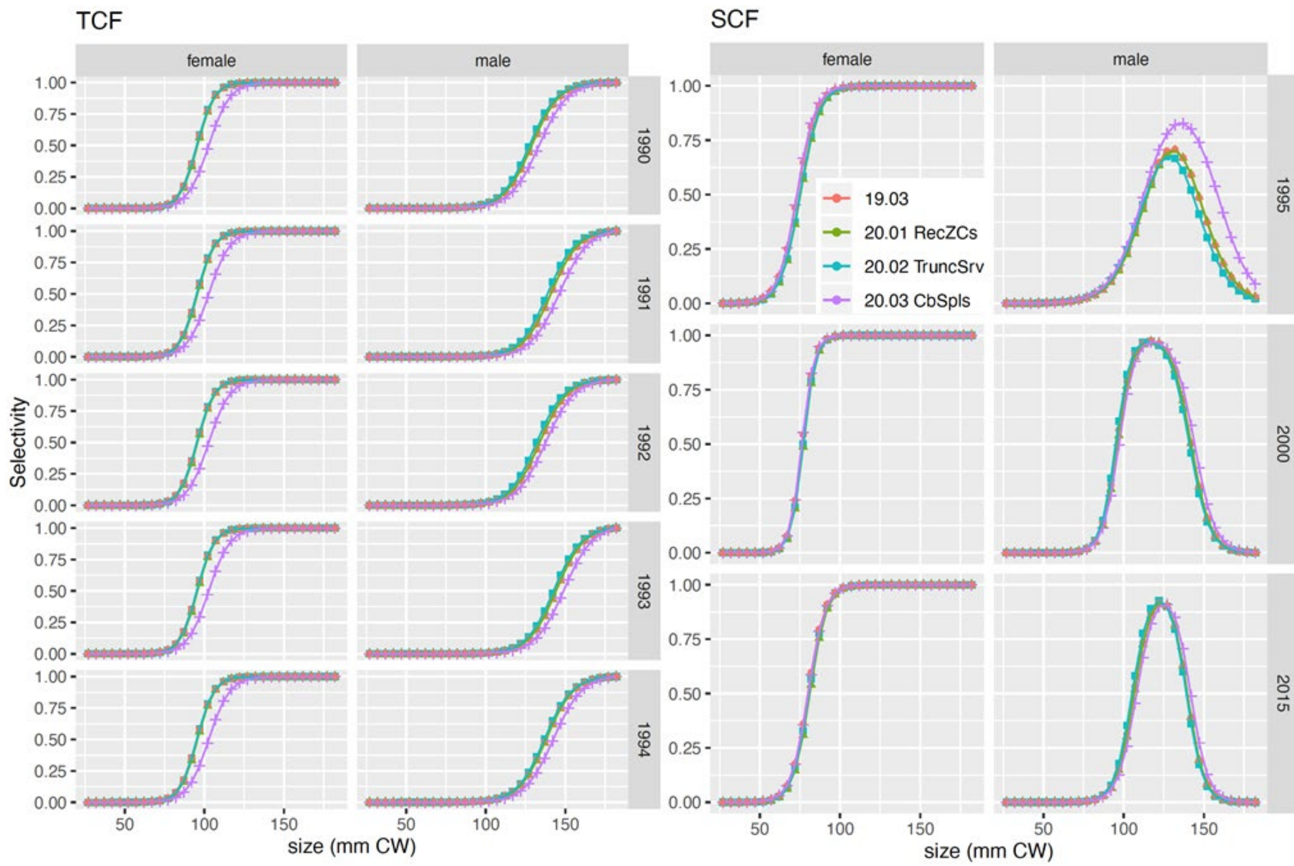

Figure 6.3.6. Comparison of estimated fishery selectivity curves in different time periods for the directed fishery (TCF) and the snow crab fishery (SCF) for Scenarios 19.03, 20.01, 20.02 and 20.03.

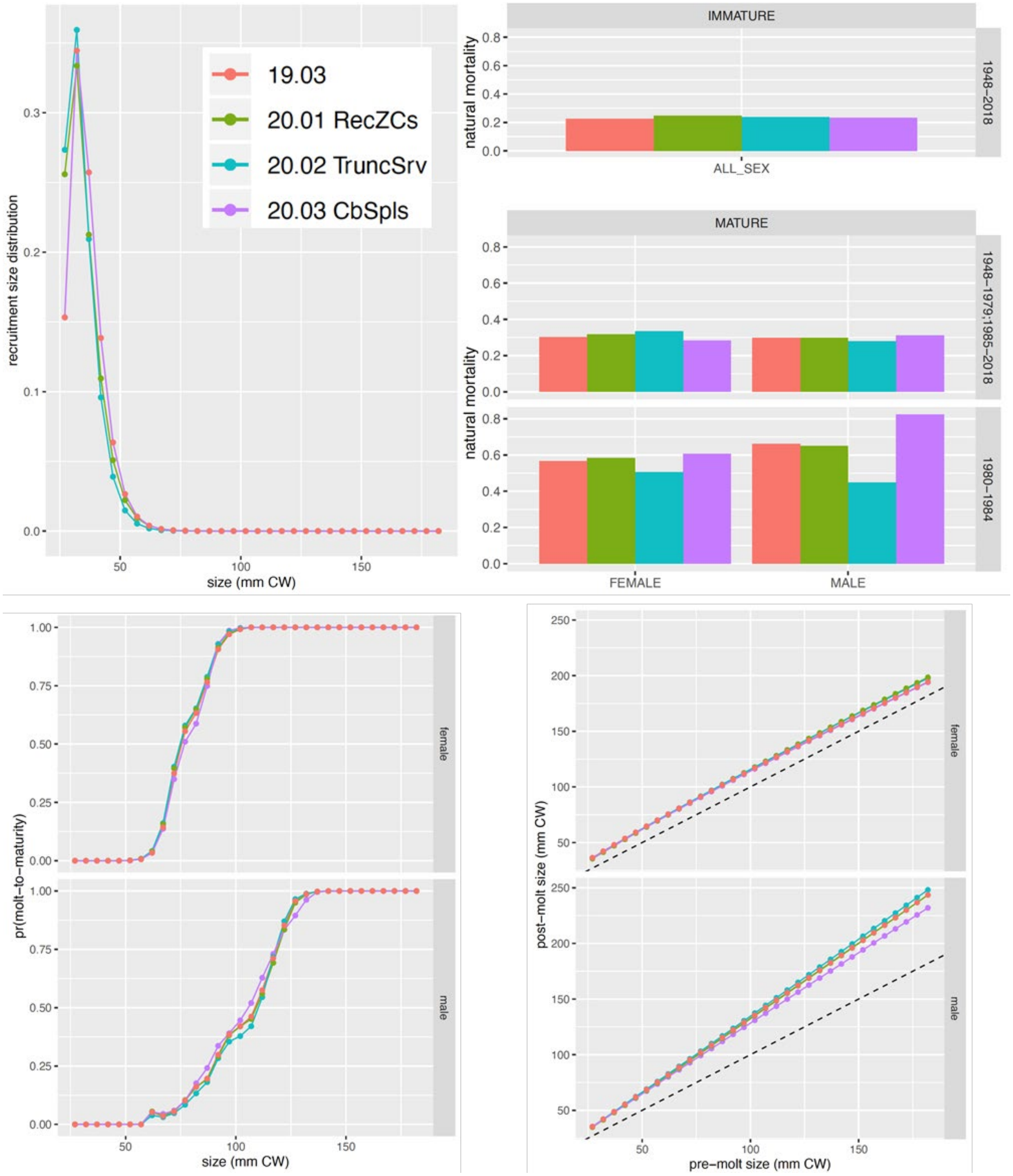

Figure 6.3.7. Comparison of estimated recruitment size distributions (upper left), natural mortality rates (upper right), probabilities of terminal molt (lower left) and mean post-molt size (lower right) for Scenarios 19.03, 20.01, 20.02, and 20.03.

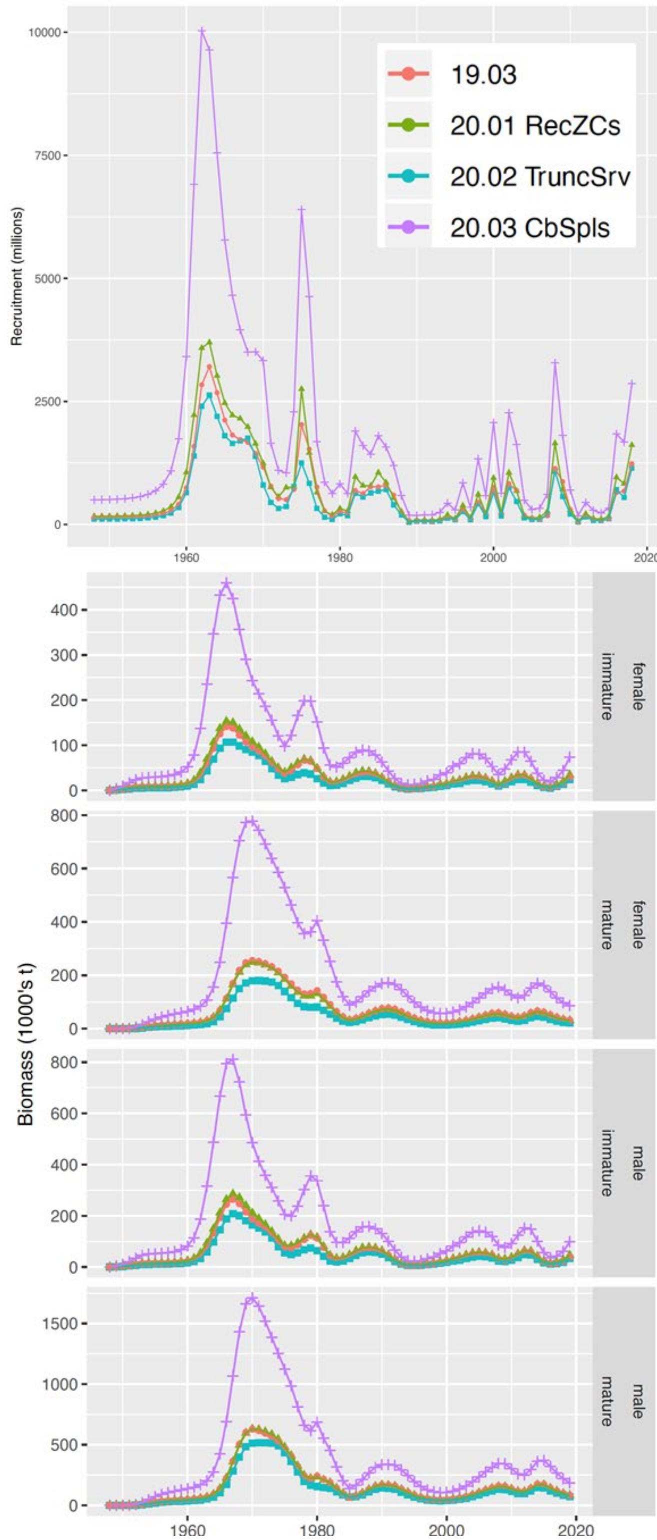

Figure 6.3.8. Comparison of estimated trends in recruitment and population biomass for Scenarios 19.03, 20.01, 20.02 and 20.03.

6.4 SBS Scenarios

Scenarios 20.06, 20.07, and 20.08 incorporate the BSFRF and NMFS survey data from the side-by-side (SBS) selectivity studies for Tanner crab conducted in 2013-2017 in one form or another. 20.06 estimates availability curves for each SBS year and fits survey biomass and size compositions from all the survey data. 20.07 fits the NMFS EBS survey biomass and size compositions, BSFRF survey biomass and size compositions, and uses availability curves determined outside the model from the EBS and SBS NMFS survey data (see Section 3.3). 20.08 fits the NMFS EBS survey biomass and size compositions from 1982-2019 using a catchability curve determined using the BSFRF and NMFS SBS data (see Section 3.4).

Fits to NMFS EBS survey biomass are quite similar across all four scenarios after 1995 (Figure 6.4.1a). 20.07 and 20.08 predict slightly higher biomass than 19.03 when biomass is high (pre-1980, 1988-1990, 2012-2014). 20.06 and 20.07 follow the estimates of biomass in both the NMFS and BSFRF SBS surveys reasonably well for females in both surveys (Figure 6.4.1b). For males, though, 20.07 follows the estimates much better than 20.06, with 20.06 exhibiting large discrepancies in 2016 from the observed values in both the NMFS and BSFRF SBS surveys. Fits to the mean NMFS EBS survey size compositions are similar across all four models for males and mature females, but 20.08 differs from the other scenarios in that it overestimates the relative abundance of immature females in the 50-75 mm CW size range (Figure 6.4.2a). 20.06 fits the mean size composition from the NMFS SBS data for males in the 100-125 mm CW range better than 20.07 does, but still under-predicts the observed values. 20.06 accurately predicts the proportions of large males relative to the observed values, whereas 20.07 over-predicts these proportions (Figure 6.4.2b). On the other hand, 20.07 accurately predicts the proportions of large females relative to the observed values, whereas 20.06 under-predicts these proportions. Interestingly, the same observations pertain to the fits to the BSFRF SBS mean size compositions (Figure 6.4.2c).

The sex-specific availability curves determined outside the model in Scenario 20.07 and estimated in 20.06 are somewhat similar to each other for 2013, are less similar in 2014, 2016 and 2017 (but follow similar large-scale trends), but differ rather substantially in 2015 with completely opposite trends with size (Figure 6.4.3a). In fact, it is rather unbelievable (and not supported by the full NMFS survey data) that availability for males in 2015 was unity up to ~125 mm CW, as estimated in Scenario 20.06. These comments carry over to the estimated catchability curves for the NMFS SBS survey data shown in Fig. 6.4.3b because these are, as assumed here, simply the availability curves multiplied by the catchability curves from the 1982+ time period for the full NMFS EBS survey (shown in 6.4.3c). The catchability curves for the full NMFS EBS survey as estimated in each scenario (Figure 6.4.3c) appear to represent a fairly mixed bag of results when one would hope to see more consistency. Estimated catchability (either from inside or outside the model) for males in the 1982+ time period has a fairly consistent shape across the scenarios, but differs in the fully-selected value, ranging from ~0.4 in Scenario 19.03 (estimated in the model) to 0.6 in Scenario 20.08 (estimated outside the model). For females, the 1982+ curve estimated outside the model (20.08) rises to full selection rapidly with size, while the curves from 20.06 and 20.07 rise more slowly—and that from 19.03 hardly rises at all.

The catchability curves for the 1975-1981 period prior to the NMFS EBS survey gear standardization in 1982 are estimated logistic curves in each scenario (Figure 6.4.3c). Estimated fully-selected catchability is 0.5 for both sexes and all scenarios here, with the estimated parameter value constrained by its lower imit. For males, the estimated size at 50% selection is much smaller (~50 mm CW) in Scenarios 20.07 and 20.08 than in Scenarios 19.03 and 20.06 (~80 mm CW). Based on previous assessments, the value for this parameter does not seem robust to changes in the model assumptions. In contrast, the size at 50% selection for females in this time period is more smoothly varying, possibly in conjunction with the slope of the catchability curve for females in the 1982+ time period.

Fits to retained and total catch biomass in the fisheries are almost identical for the four scenarios (Figures 6.4.4-5), as are the selectivity curves for the directed and snow crab fisheries (Figure 6.4.6).

Estimated recruitment size distributions in 20.06, 20.07 and 20.08 had higher proportions in the first two size bins and lower proportions in larger bins relative to the fixed size distribution in 19.03 (Figure 6.4.7, upper left). Estimated natural mortality rates were fairly similar across all four scenarios for immature crab and for mature males and females in most of the model time period, but the estimated rates for mature crab in the "enhanced mortality" period (1980-1984) were noticeably higher in Scenario 20.06 than the other scenarios (Figure 6.4.7, upper right). The estimated size-specific probability of terminal molt for was almost identical across the four scenarios for both sexes (Figure 6.4.7, lower left). Estimated mean growth rates were similar for both sexes across all four scenarios (Figure 6.4.7, lower right).

Once again, the timing of estimated fluctuations in recruitment was almost identical for the four scenarios across the entire model time period—the scenarios differ primarily in scale, with recruitment generally highest in Scenario 20.06, followed closely by 19.03, and lowest in Scenario 20.08 (Figure 6.4.8). Estimated trends in population biomass exhibit a different ranking among scenarios, with 19.03 larger than 20.06 for mature crab because estimated natural mortality rates are slightly higher for mature crab in 20.06.

Of the three alternative scenarios examined in this section, the most reasonable appears to be 20.07. This scenario fits the NMFS EBS and the BSFRF SBS survey data using empirical availability curves developed outside the model based on the NMFS SBS survey data. This scenario thus uses the BSFRF SBS data to inform fully-selected catchability ("Q") for the NMFS survey, but avoids the cost and model instability associated with estimating over 250 additional parameters in 20.06. The empirical availability curves estimated outside the model used in 20.07 are based on simple, direct calculations, whereas the availability curves estimated inside the model can be influenced indirectly by the values of many other parameters through the model optimization. Additionally, the "flat" availability curves for females in 2015 and 2016 (Figure 6.4.3b) estimated in Scenario 20.06 seem highly unlikely and are in substantial disagreement with the empirically-determined curves. The use of fixed *catchability* curves estimated outside the model in 20.08 is somewhat questionable because these are based on the assumption that the ratios of size composition data from the NMFS and BSFRF SBS data reflect catchability in the full EBS survey. This seems questionable given the fair amount of temporal variation exhibited in the empirical catchability curves developed from the SBS studies on a yearly basis (although this could be flipped around, as well, to question the use of time-invariant catchability for the NMFS survey since the gear change in 1982).

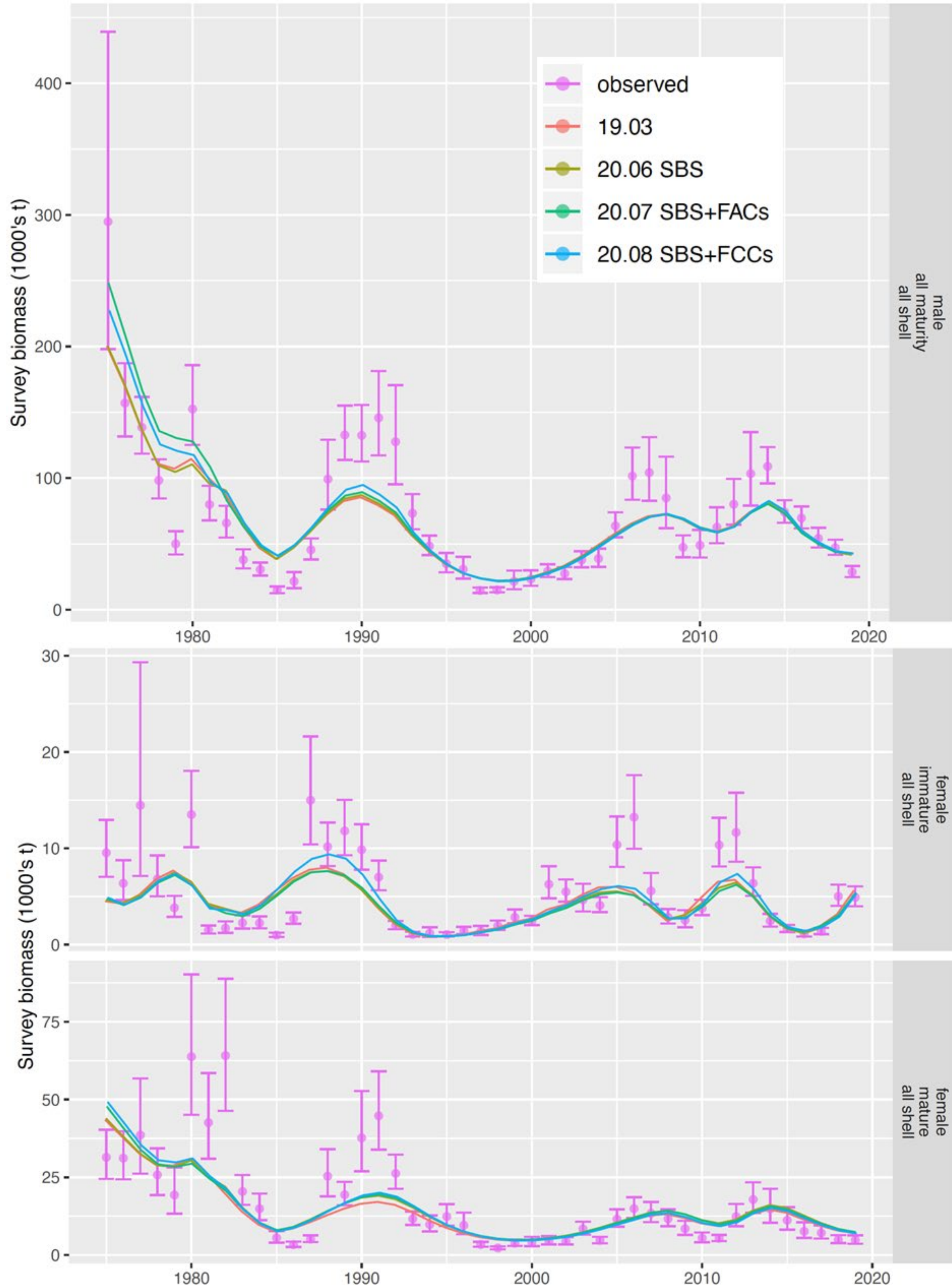

Figure 6.4.1a. Comparison of fit to NMFS EBS survey biomass for Scenarios 19.03, 20.06, 20.07, and 20.08.

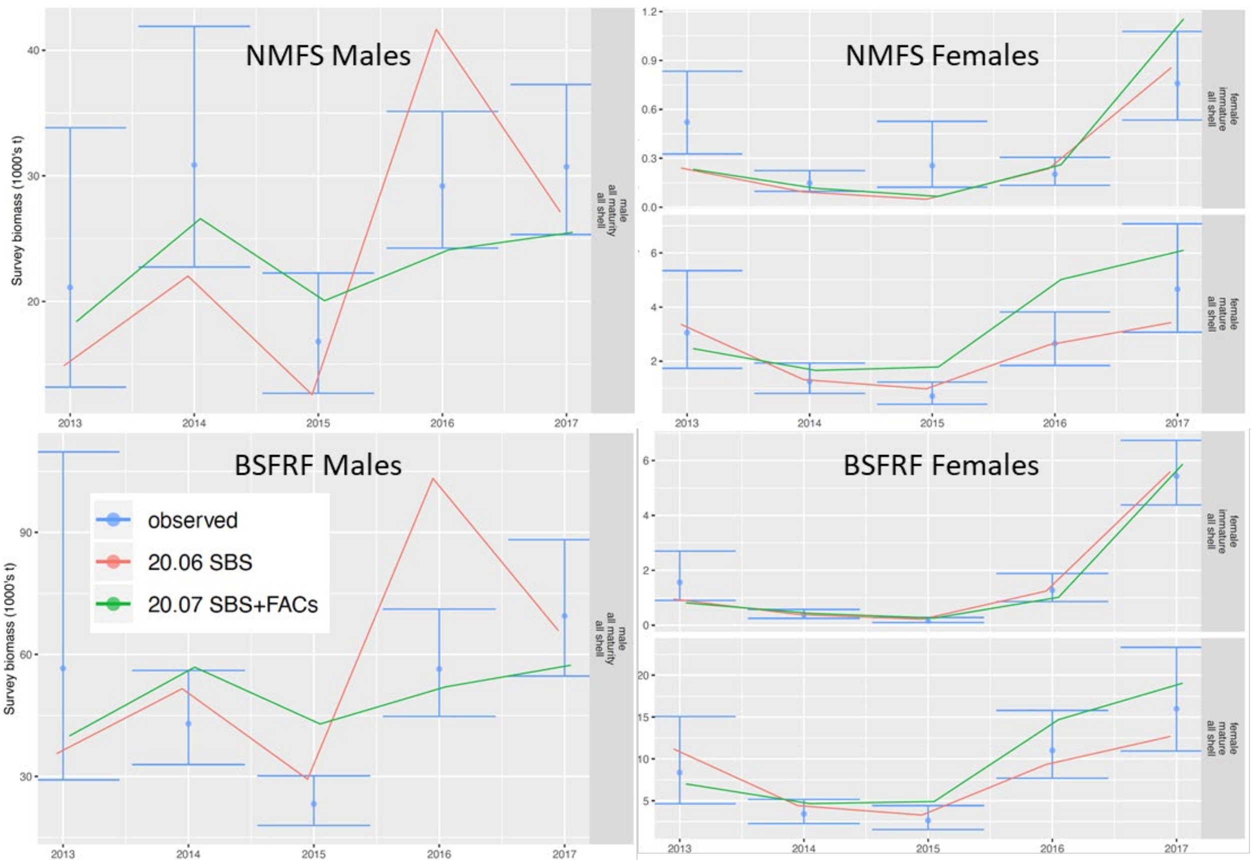

Figure 6.4.1b. Comparison of fit to survey biomass estimates from the SBS selectivity studies for Scenarios 19.03, 20.06, and 20.07 (20.08 did not fit SBS data).

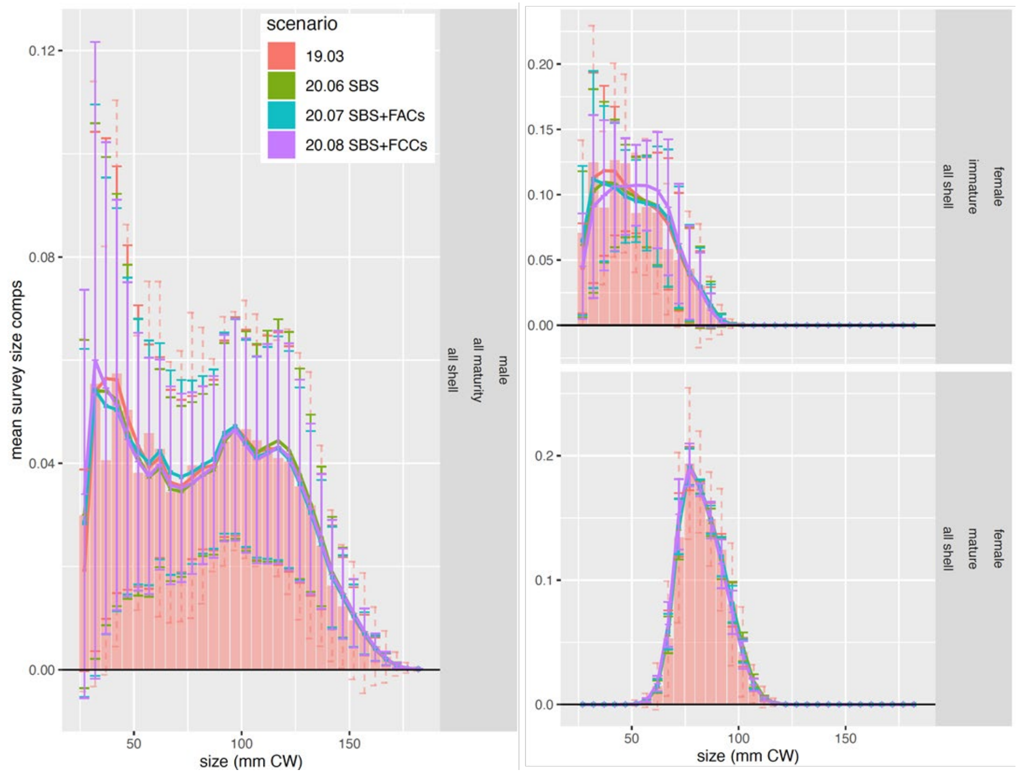

Figure 6.4.2a. Comparison of mean predicted size compositions to observed NMFS EBS survey compositions for Scenarios 19.03, 20.06, 20.07, and 20.08.

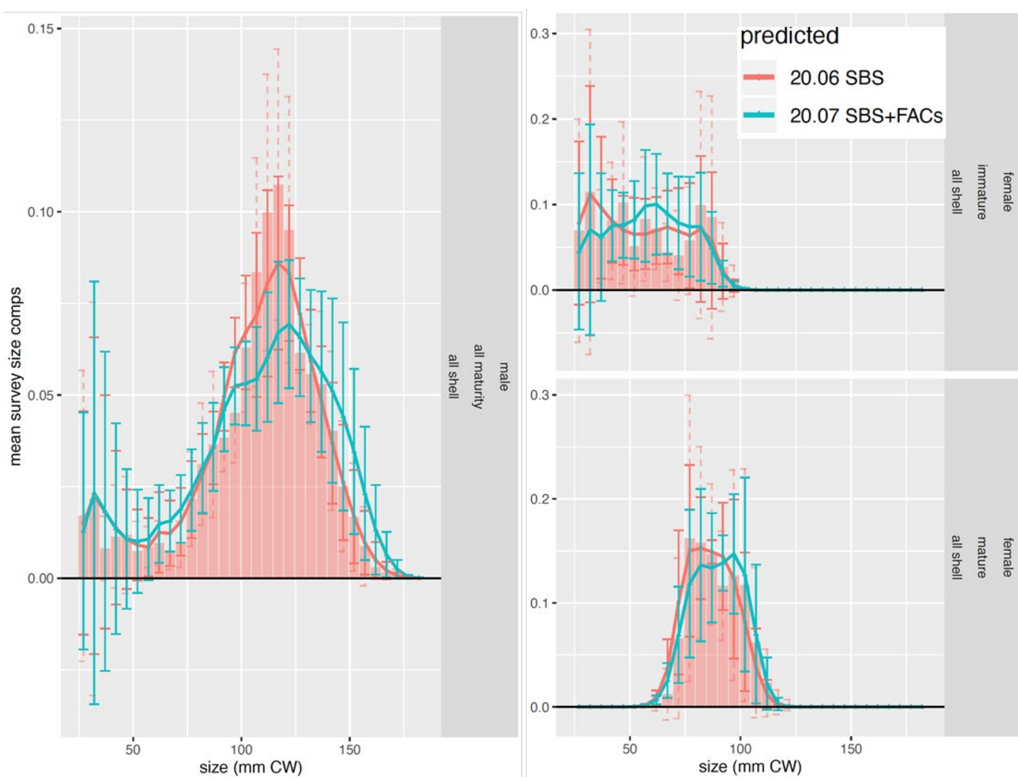

Figure 6.4.2b. Comparison of mean predicted size compositions to observed NMFS SBS survey compositions for Scenarios 19.03, 20.06, and 20.07 (20.08 did not fit SBS data).

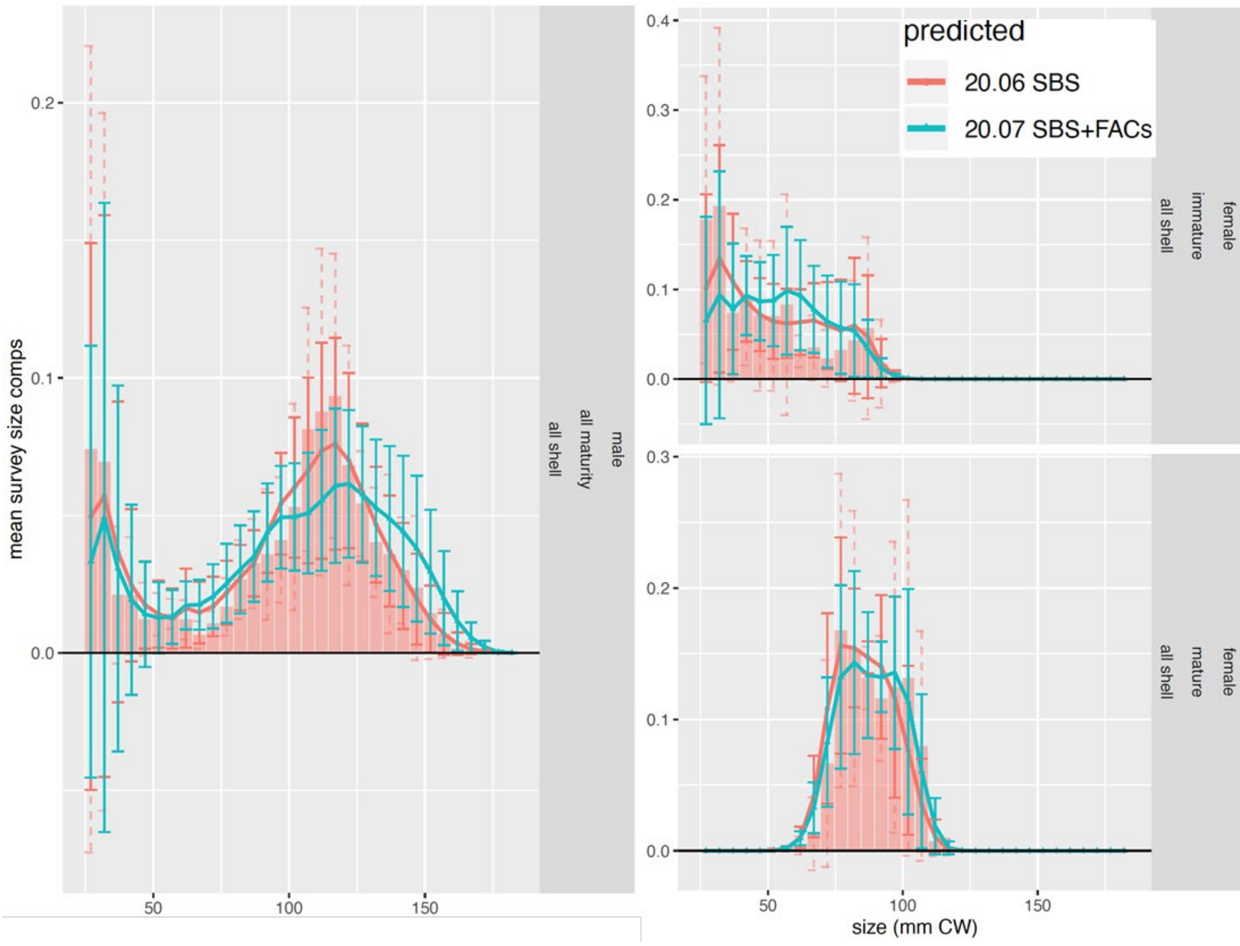

Figure 6.4.2c. Comparison of mean predicted size compositions to observed BSFRF SBS survey compositions for Scenarios 19.03, 20.06, and 20.07 (20.08 did not fit SBS data).

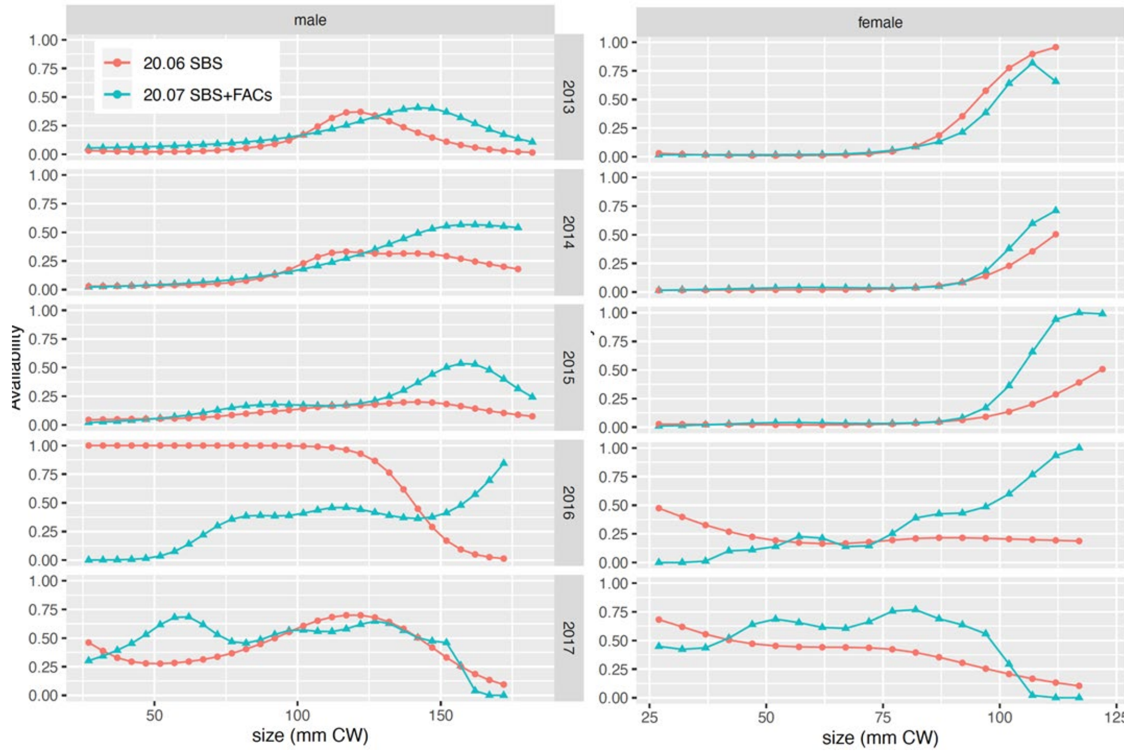

Figure 6.4.3a. Comparison of estimated survey *availability* curves for the SBS survey data from Scenarios 20.06 and 20.07. Notes: 1) availability curves are not used in Scenarios 19.03 and 20.08; 2) the curves for Scenario 20.07 were determined outside the model.

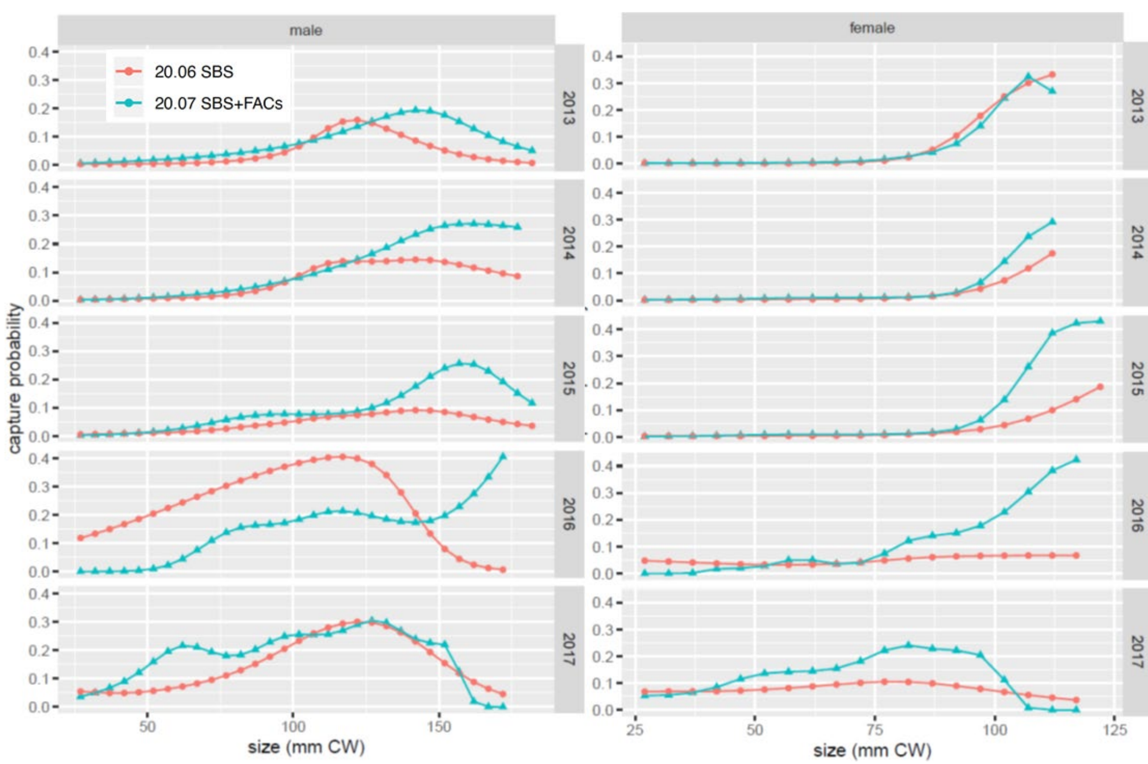

Figure 6.4.3b. Comparison of estimated survey *catchability* curves for the NMFS SBS survey data for Scenarios 20.06 and 20.07. Notes: 1) availability curves were not used in Scenarios 19.03 and 20.08; 2) the curves for Scenario 20.07 were determined outside the model.

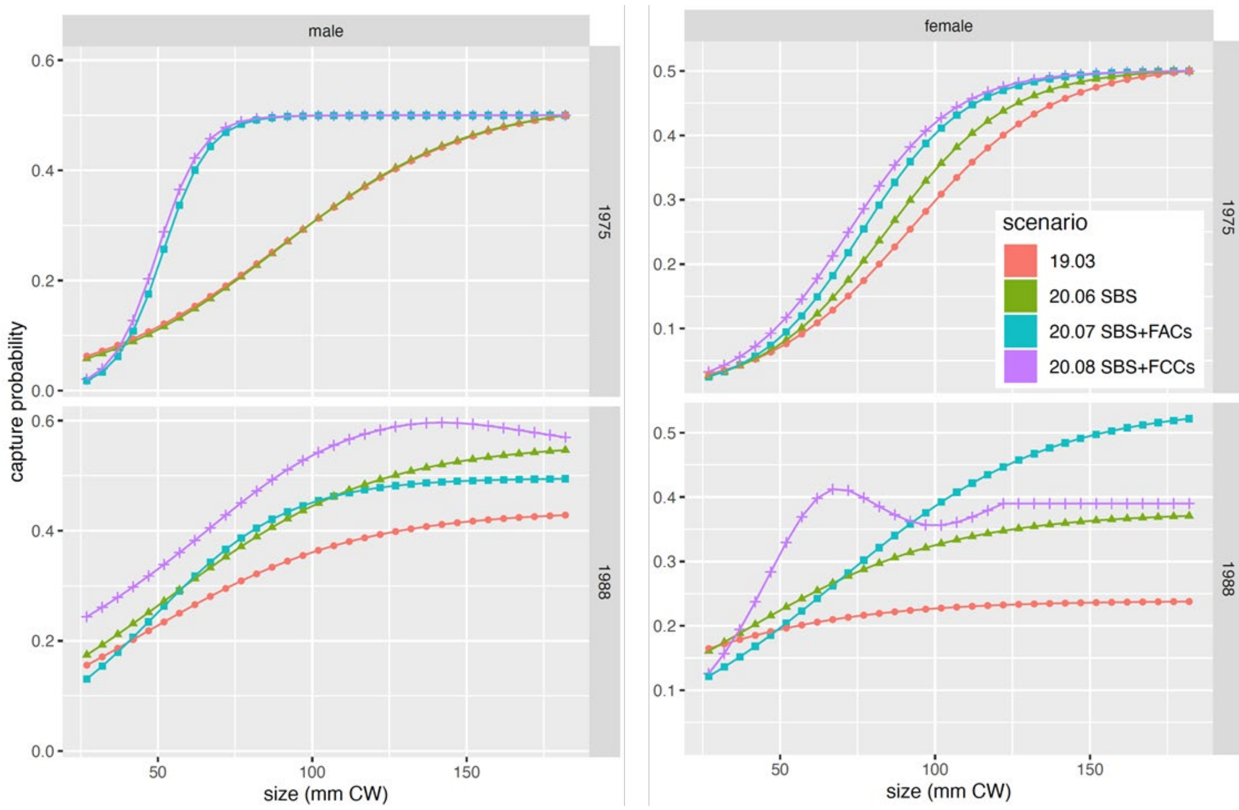

Figure 6.4.3c. Comparison of estimated survey *catchability* curves for the NMFS EBS survey from Scenarios 19.03, 20.06, 20.07, and 20.08. Note: the catchability curves for 20.08 in the 1982-2019 period were determined outside the model.

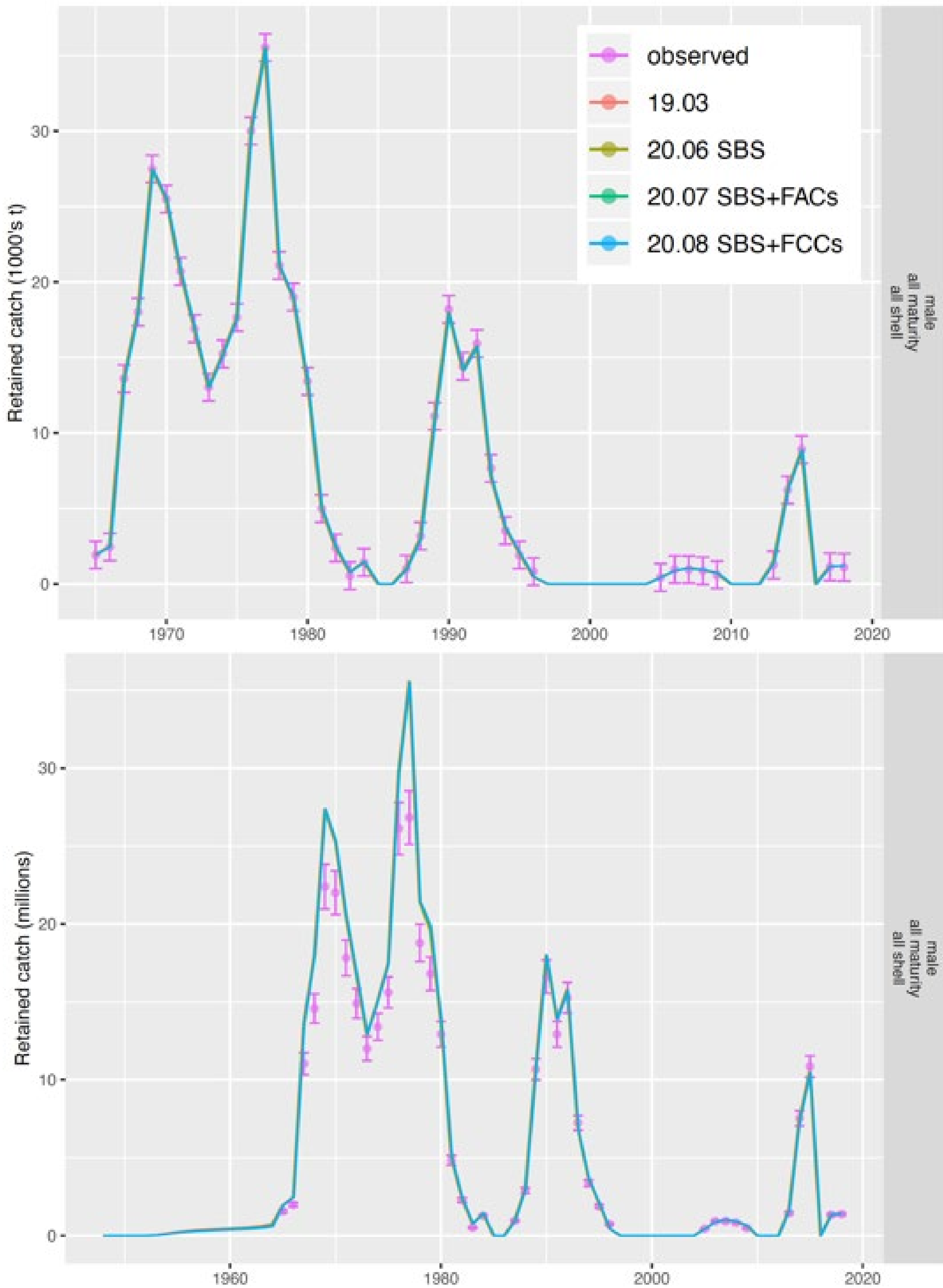

Figure 6.4.4. Comparison of fits to retained catch biomass and abundance in the directed fishery for Scenarios 19.03, 20.06, 20.07, and 20.08.

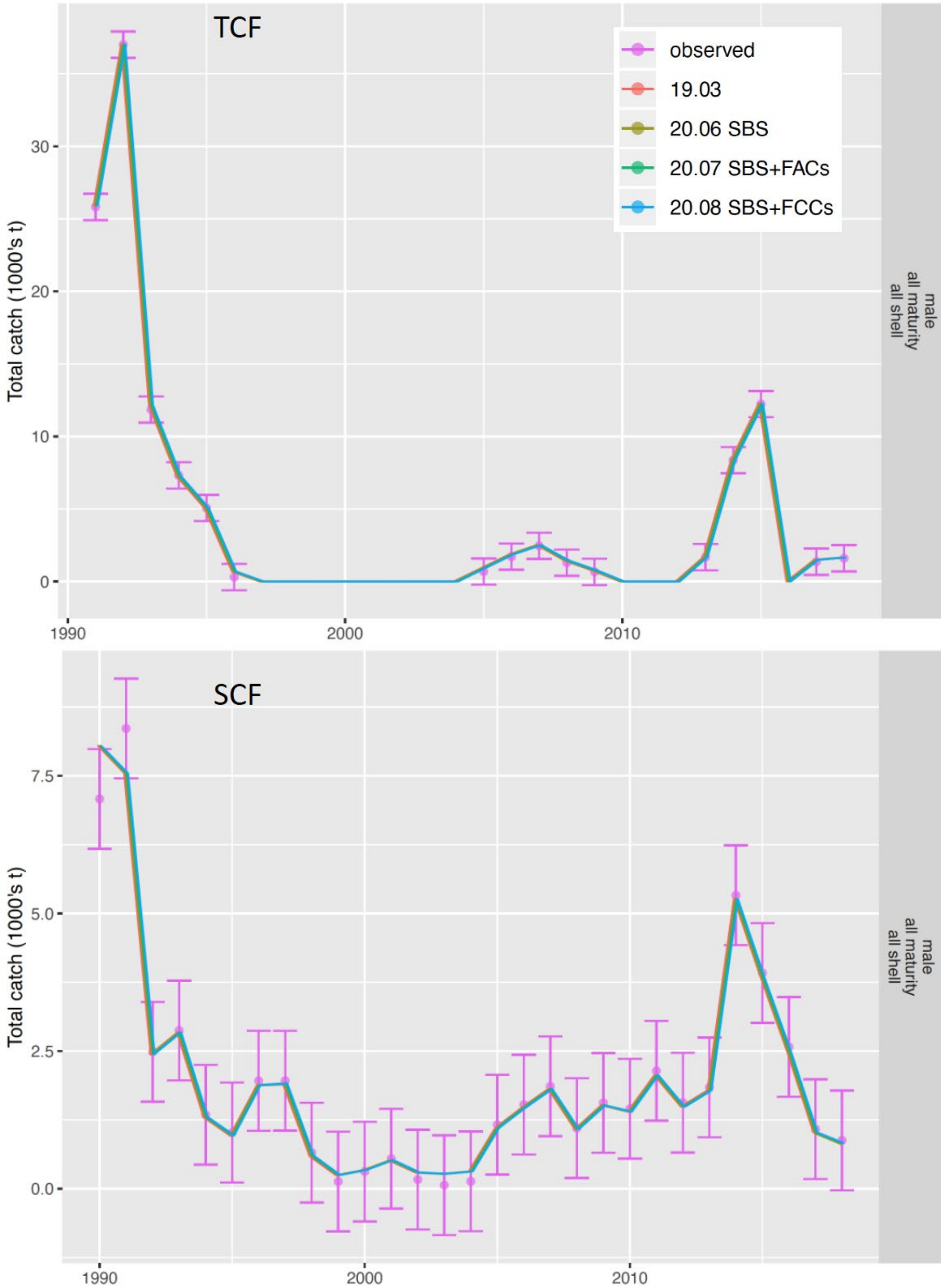

Figure 6.4.5. Comparison of fits to total catch biomass in the directed Tanner crab fishery (TCF) and the snow crab fishery (SCF) for Scenarios 19.03, 20.06, 20.07, and 20.08.

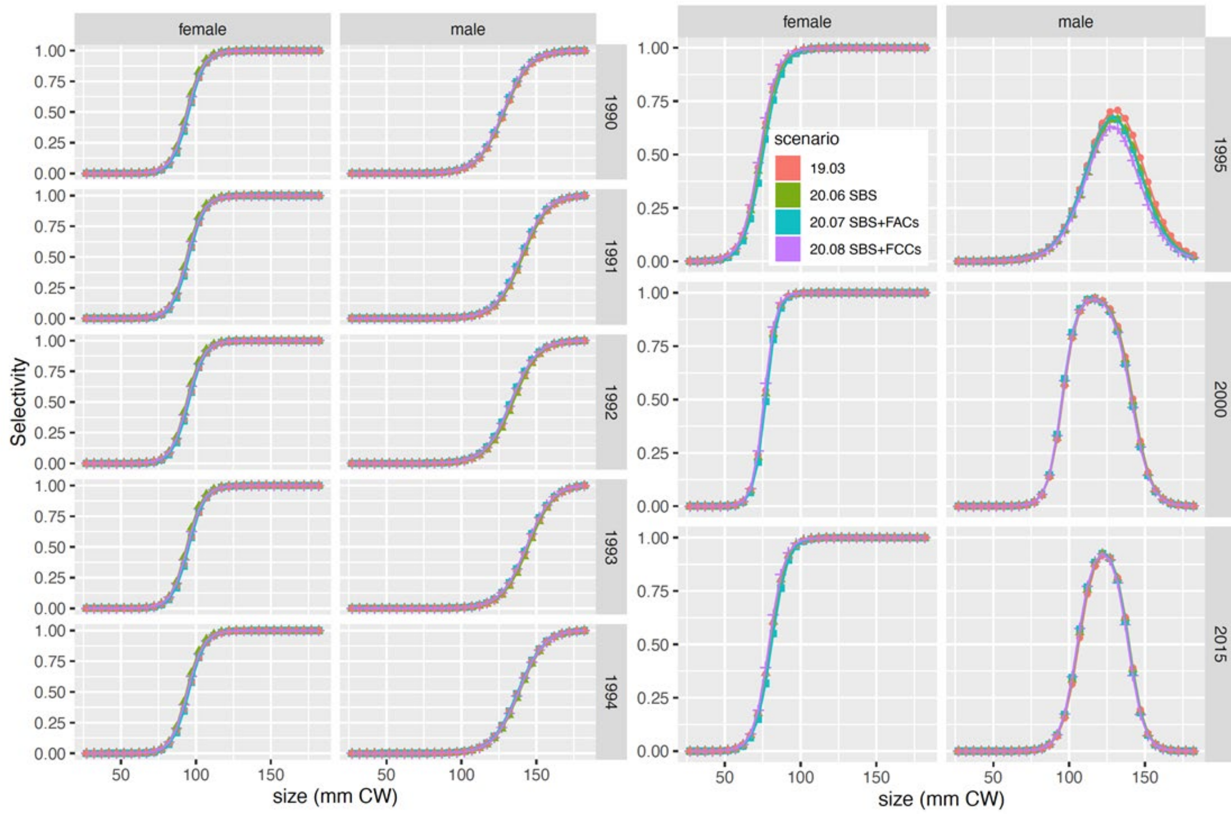

Figure 6.4.6. Comparison of estimated fishery selectivity curves in different time periods for the directed fishery (TCF) and the snow crab fishery (SCF) for Scenarios 19.03, 20.06, 20.07, and 20.08.

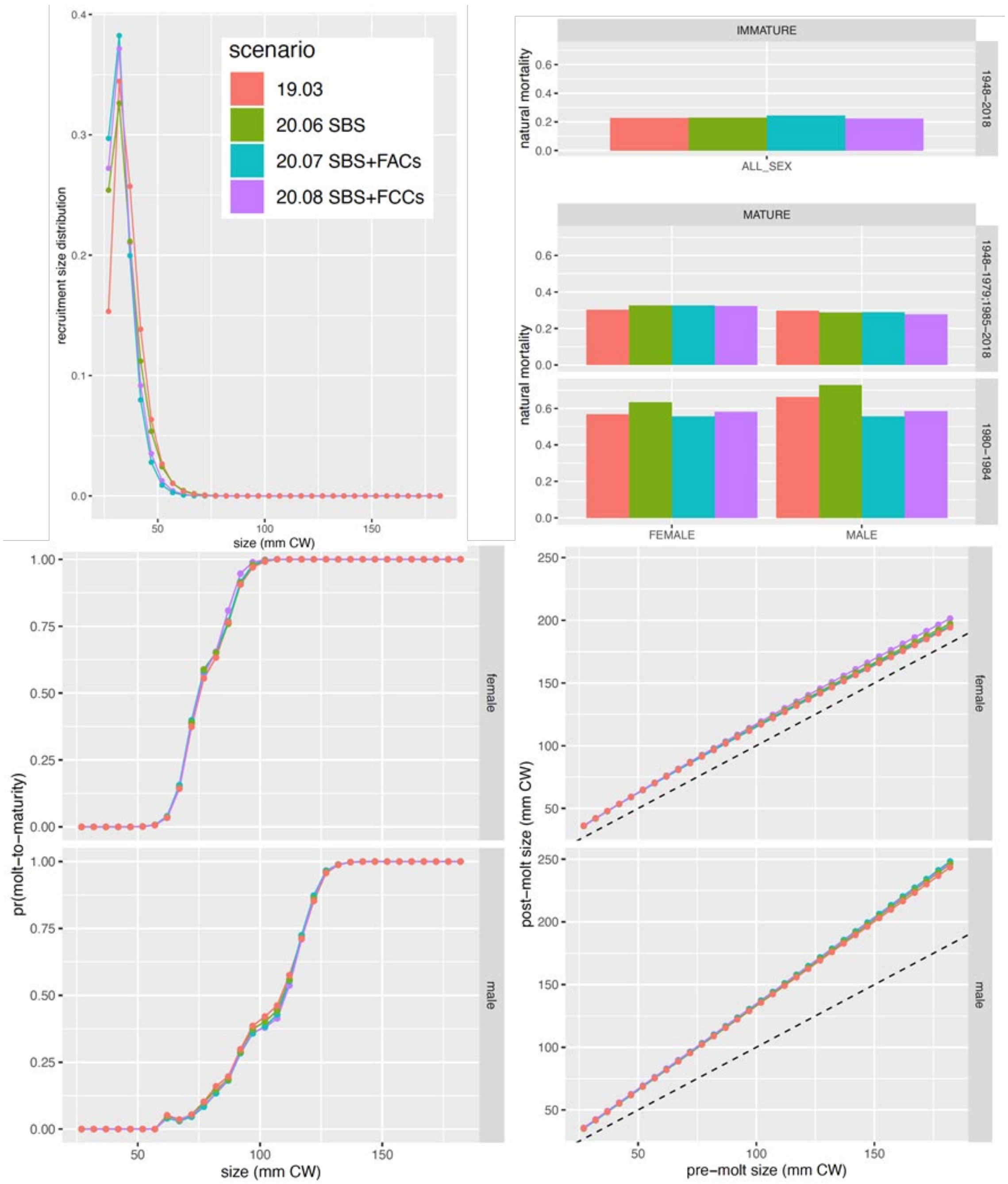

Figure 6.4.7. Comparison of estimated recruitment size distributions (upper left), natural mortality rates (upper right), probabilities of terminal molt (lower left) and mean post-molt size (lower right) for Scenarios 19.03, 20.06, 20.07, and 20.08.

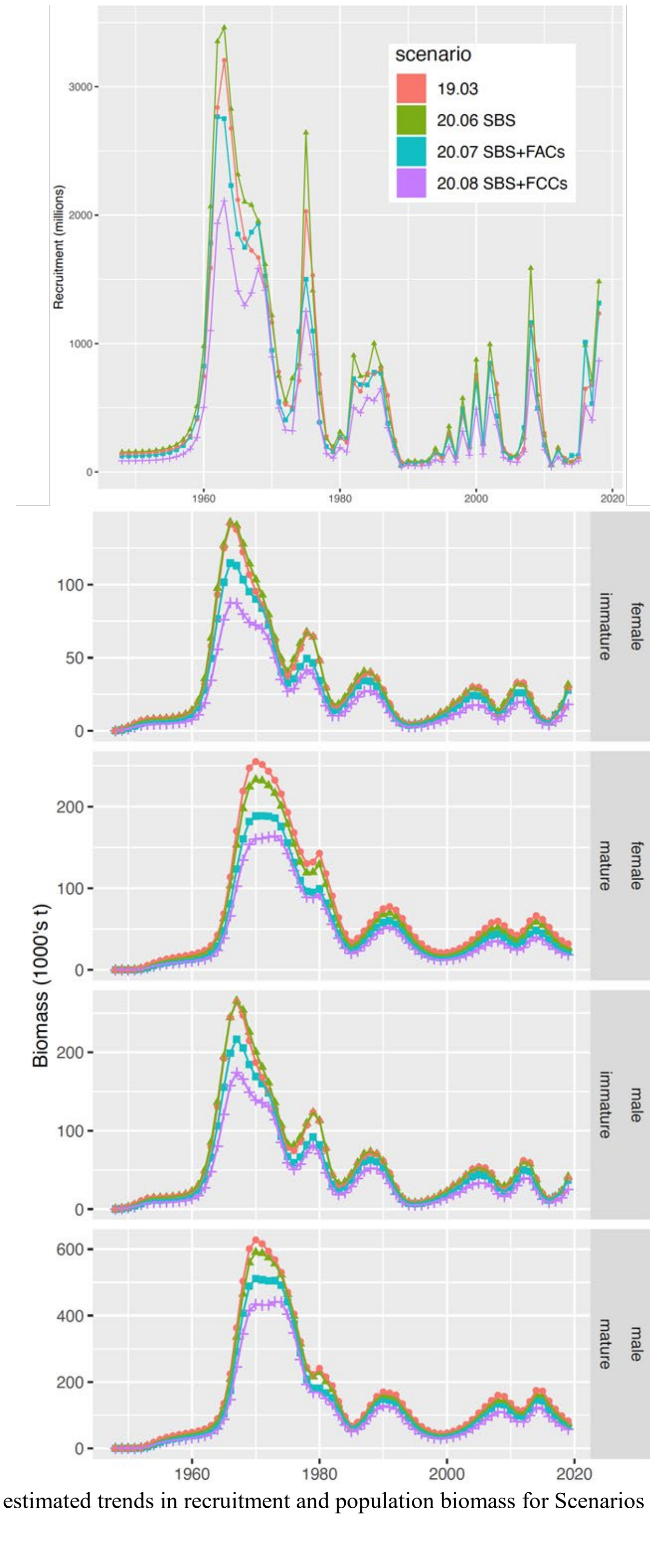

Figure 6.4.8. Comparison of estimated trends in recruitment and population biomass for Scenarios 19.03, 20.06, 20.07, and 20.08.

7.0 Recommended model scenarios for Sept 2020

Based on the scenarios presented here, the following, with data updated for 2019/20, are recommended to be considered for status determination and setting OFL/ABC:

- 19.03 (the 2019 assessment model updated with 2019/2020 data)
- 20.01 (19.03 + estimating the recruitment size distribution)
- 20.04* (20.01 [not 19.03] + VAST model-based survey biomass estimates)
- 20.06 (20.01 + SBS BSFRF and NMFS data)
- 20.07 (20.01 + SBS BSFRF data and fixed availability curves)

It should be noted that VAST estimates have been provided based on the entire EBS survey area for Tanner crab. In order to combine the VAST estimates with estimates from the SBS studies, VAST should probably be applied to these data as well.

8.0 Literature Cited

Lang, C. A., J. I. Richar, R. J. Foy. 2018. The 2017 eastern Bering Sea continental shelf and northern Bering Sea bottom trawl surveys: Results for commercial crab species. U.S. Dep. Commer., NOAA Tech. Memo. NMFS-AFSC-372, 233 p.

Stockhausen, W. 2018a. 2018 Stock Assessment and Fishery Evaluation Report for the Tanner Crab Fisheries of the Bering Sea and Aleutian Islands Regions. In: Stock Assessment and Fishery Evaluation Report for the King and Tanner Crab Fisheries of the Bering Sea and Aleutian Islands: 2018 Final Crab SAFE. North Pacific Fishery Management Council. Anchorage, AK.

Thorson, J. 2019. Guidance for decisions using the Vector Autoregressive Spatio-Temporal (VAST) package in stock, ecosystem, habitat and climate assessments. Fisheries Research. 210:143-161. 10.1016/j.fishres.2018.10.013.

Wood, S.N. 2011. Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. Journal of the Royal Statistical Society (B). 73(1):3-36.

Wood, S.N. 2017. Generalized Additive Models: An Introduction with R (2nd edition). Chapman and Hall/CRC.