May 2016 Tanner Crab Stock Assessment Activities Report: Part 2

William T. Stockhausen
Alaska Fisheries Science Center
April 2016

New Model Code: TCSAM2015

TCSAM2015, the new version of the Tanner crab model, is an integrated assessment model like the previous version, TCSAM2013, that is fit to multiple data sources. It was developed by the author in C++ using AD Model Builder (Fournier et al., 2012) libraries. TCSAM2015 is based on the Tanner crab model used in the 2015 stock assessment (Stockhausen, 2015; Appendix A in Part 1 of this report), but it differs from it in several respects and is completely new model code. Appendix A to Part 2 describes the TCSAM2015 model in detail. The six most substantial differences between TCSAM2015 and the model configuration using TCSAM2013 adopted for the 2015 stock assessment are:

- ability to define multiple time periods for any model process via input files
- ability to specify data aggregation level via input files
- exclusive use of the Gmacs fishing mortality model (FMM).
- re-parameterization of some model processes to (hopefully) improve convergence properties
- ability to specify of Bayesian-like priors on any parameters
- ability to do Tier 3-type OFL calculations directly within a model run, rather post-processing model results using standalone code

The new code also eliminates all of the "hard-wired" components in TCSAM2013, such as the number and range of model time periods. This makes developing alternative model configurations, particularly defining multiple time periods for model processes, much simpler and faster. The new code also provides more extensive options for data types, model fitting, and selectivity functions.

As with TCSAM2013, model parameters in TCSAM2015 are estimated using a maximum likelihood approach. Data components entering the likelihood potentially include survey abundance and/or biomass, survey size compositions, retained catch abundance and/or biomass, retained catch size compositions, total catch or discard catch abundance and/or biomass from at-sea observer sampling, and total catch size compositions from at-sea observer sampling.

An R-based package for data simulation, rsimTCSAM, has been developed to provide a completely independent code basis for testing TCSAM2015 features, functionality and estimation performance. In addition, another R-based package, rTCSAM2015, has been developed to simplify making multiple model runs (e.g., jittering initial parameter values, varying fixed model parameters in a sequence), plotting individual model output, and making comparisons among multiple model runs or between rsimTCSAM and TCSAM2015 models, and comparison of multiple TCSAM2015 models. ADMB model code for TCSAM2015 is publicly available on github[1]. The current version of the code is on the

---

[1] https://github.com/wStockhausen/tcsam2015.

"devFisheryFs" branch. It also requires the wtsADMB library of ADMB and C++ functions[2]. R code for the rsimTCSAM package[3] and rTCSAM2015[4] package are also available on github.

It is anticipated that future improvements to the Tanner crab model (e.g., incorporating BSFRF surveys, chela height data, and growth data) will be incorporated into TCSAM2015 and that the TCSAM2013 code will not be further updated.

Testing with simulated data

As noted above, code development and testing for TCSAM2015 has been facilitated by the associated rsimTCSAM package. Results from a recent series of simulation tests of the TCSAM2015 code are presented in this section.

The rsimTCSAM package was used to simulate population, fishery, and survey dynamics for a Tanner crab-like stock, using many parameter values and fishery and survey configuration details from the Tanner crab assessment model in order to "exercise" the TCSAM2015 model code in a realistic fashion. The simulation was run from 1950 to 2014 using 5 mm CW size bins with left cutpoints from 25 mm CW to 180 mm CW. Weight-at-size relationships were based on the standard power law formula for Tanner crab and were constant over the model time period:

sex	maturity	parameters a	b
female	immature	0.000637	2.794
	mature	0.000344	2.956
male	immature	0.000163	3.136
	mature	0.000163	3.136

Natural mortality was modeled using two time periods, with enhanced rates during 1980-1984 compared with the remainder of the simulation:

sex	maturity	standard M	cv	1980-1984 M	cv
male	mature	0.23	0	0.4	0
male	immature	0.23	0	0.4	0
female	mature	0.23	0	0.4	0
female	immature	0.23	0	0.4	0

Although no process error due to random variation in natural mortality rates was included in this simulation (cv's = 0), it remains an option for future testing.

Immature crab molted every year. The probability of molting to maturity was constant over the modeled time period. It was described using sex-specific ascending logistic functions parameterized by size at 50% probability ($z50$) and a standard width:

---

[2] https://github.com/wStockhausen/wtsADMB. Current branch is "master".
[3] https://github.com/wStockhausen/rsimTCSAM. Current branch is "master".
[4] https://github.com/wStockhausen/rTCSAM2015. Current branch is "master".

sex	z50	width
male	100	10
female	80	10

Growth was also constant over the modeled time period. It was modeled using cumulative gamma distribution functions parameterized using:

sex	ln(a)	ln(b)	ln(scale)
male	log(0.43)	log(0.97)	log(0.75)
female	log(0.70)	log(0.88)	log(0.75)

where a and b define the sex-specific relationship between mean molt increment and size, while the scale parameter determines the width of resulting molt increment distributions. Parameters were specified on the log-scale to be consistent with the parameterization used in TCSAM2015.

Parameters determining recruitment were constant over the modeled time period, so no "regime shifts" were included in this simulation (although they could be in future testing). Median recruitment was set at 71 million crab, with an annual cv of 0.5 to add a component of process error to the simulation. The sex ratio (1:1) and size distributions at recruitment were set similarly to those in the 2015 assessment model.

Four fisheries, one directed fishery ("TCF") and three bycatch fisheries ("SCF", "RKF", and "GTF"), were defined for the simulation. The directed TCF fishery "resembles" the directed Tanner crab fishery, while the SCF, RKF and GTF bycatch fisheries resemble the snow crab, Bristol Bay red king crab, and groundfish bycatch fisheries included in the stock assessment. The TCF is the only simulated fishery that retains male crab, while all four fisheries catch and discard both male and female crab. Handling mortality rates of 0.321 and 0.8 were applied to discarded crab in the crab pot and groundfish (trawl) fisheries, respectively. For each fishery, sex-specific log-scale mean capture rates and standard deviations, as well as capture selectivity functions, were applied over the different time periods defined for each fishery, as defined in the following table:

fishery	time period	sex	capture rates		capture selectivity curves				
			log-scale mean	std dev	type	az50	az95-az50	az95-dz95	dz95-dz50
TCF	1965-1984	male	ln(0.30)	0.4	logistic	110	15		
		female	ln(0.10)		logistic	80	15		
	1978-1996, 2005-2009, 2013	male	ln(0.30)	0.4	logistic	125	15		
		female	ln(0.10)		logistic	80	15		
SCF	1978-1996	male	ln(0.15)	0.4	double logistic	80	15	15	15
		female	ln(0.08)		logistic	80	30		
	1997-2004	male	ln(0.10)	0.4	double logistic	90	15	15	15
		female	ln(0.05)		logistic	75	20		
	2005-2014	male	ln(0.10)	0.4	double logistic	100	15	15	15
		female	ln(0.05)		logistic	85	20		
RKF	1978-1983, 1986-1993,	male	ln(0.10)	0.4	logistic	150	30		
		female	ln(0.02)		logistic	90	15		
	1997-2004	male	ln(0.05)	0.4	logistic	130	30		
		female	ln(0.01)		logistic	80	15		
	2005-2014	male	ln(0.010)	0.4	logistic	150	30		
		female	ln(0.002)		logistic	85	15		
GTF	1973-1986	male	ln(0.15)	0.4	logistic	55	30		
		female	ln(0.03)		logistic	65	50		
	1987-1996	male	ln(0.10)	0.4	logistic	95	30		
		female	ln(0.02)		logistic	55	50		
	1997-2014	male	ln(0.05)	0.4	logistic	65	100		
		female	ln(0.01)		logistic	75	40		

In the table above, the capture rate standard deviation determines the annual log-scale variability (~cv) in the applied capture rates; deviations are not sex-specific. The logistic selectivity functions are defined in terms of the size at 50% selection on the ascending limb of the curve (az50) and the increment to 95% selection (az95-az50), while the double logistic curves are additionally defined by subsequent increments to 95% selection on the descending limb (dz95-az95) and to 50% selection on descending limb (dz50-dz95). In addition, a single retention curve for the TCF was defined over the modeled period using a logistic curve with 50% selection occurring at 138 mm CW and 95% selection occurring at 153 mm CW.

Finally, characteristics for a survey reminiscent of the annual NMFS trawl survey were defined in two time periods using:

Survey	time period	sex	catchability		capture selectivity curves		
			log-scale mean	std dev	type	az50	az95
NMFS Trawl Survey	1975-1981	male	ln(1.0)	0	logistic	50	120
		female	ln(0.8)	0	logistic	60	170
	1982-2015	male	ln(1.0)	0	logistic	30	120
		female	ln(0.8)	0	logistic	-5	90

where survey catchability (Q) was defined on the log-scale, smaller for females, and the same in both time periods while survey selectivities were also sex-specific and differed between the two periods.

Although it is possible to inject survey-related process error into the simulations through annual variation in Q, this was not done for this simulation (Q standard deviations = 0).

The rsimTCSAM package produces output data files and a model configuration file suitable for use in running a TCSAM2015 model. Although observation error can be added to the fishery and survey data produced in the simulation, this was not done for this exercise. The model configuration file for the simulation ("rsimTCSAM.Configuration.dat"), as well as the simulation model output files and plots of the simulation results, can be reviewed in the accompanying online file "SimulationModelFiles.zip". The data types, data sources, and time frames for the simulated output data used below are indicated in Figure 1.

The TCSAM2015 model was run for six scenarios (A4A-F), each with different sets of model processes fixed or estimated, to make an initial test of estimation performance with "perfect" data, as well as to verify that the model ran correctly. The following table indicates which model processes were estimated for each scenario, where "E" indicates the parameters were estimated and "NE" indicates they were not estimated:

Process	parameter type	Model Scenario					
		A4A	A4B (A)	A4C (B)	A4D (B)	A4E (D)	A4F (E)
Recruitment	log-scale mean	E	E	E	E	E	E
	annual deviations	E	E	E	E	E	E
Natural mortality	log-scale offsets	NE	NE	NE	E	E	E
Growth	log-scale coefficients	NE	NE	NE	NE	E	E
Molt-to-maturity	logit-scale parameters	E	E	E	E	E	E
Fisheries	handling mortality	NE	NE	NE	NE	NE	NE
	log-scale mean	E	E	E	E	E	E
	annual deviations	E	E	E	E	E	E
	female offset	E	E	E	E	E	E
	capture selectivity	E	E	E	E	E	E
	retention curve	E	E	E	E	E	E
Surveys	log-scale base Q	NE	NE	E	NE	NE	E
	Q offsets	NE	NE	E	NE	NE	E
	selectivity	NE	E	E	E	E	E

In the model scenario labels above, the "base" scenario from which the scenario in question was derived is given in parentheses. The highlighting indicates the changes made to obtain each scenario from its base. Model parameters that were not estimated were fixed at values equivalent to those used in the simulation.

Previous testing with other simulated data indicated almost all model runs converged to the same objective function value so, due to time constraints, each scenario was only run once. This is not a recommended practice, particularly when fitting noisy data (which this was not); however, in this case it has minimal effect on the outcome. Initial values for "dev" and offset parameters were set to 0; molt-to-maturity parameters were set using a nominal, linearly increasing trend with size on the logit scale (since molt-to-maturity curves were assumed to be non-decreasing functions of size). All other estimated parameters were started at randomly-selected (jittered) values.

Final objective values for each scenario are given in the following table:

| | Model Scenario | | | | | |
	A4A	A4B	A4C	A4D	A4E	A4F
objective function	313.309	313.301	313.299	313.259	312.817	312.783
maximum gradient	3.37E-05	4.32E-05	2.03E-05	4.21E-05	1.61E-04	1.41E-04

From a theoretical standpoint, each scenario should result in identical values for the estimated parameters, which would also be identical to the corresponding values used to run the simulation, and thus each should achieve the same final value for the objective function. However, (very) slightly better fits are achieved as more parameters are estimated across the scenarios, but this is probably a consequence of numerical round-off in the input data files—estimating more parameters gives just a slightly better fit to not-quite-perfect data.

Graphical comparisons of final values in each scenario, as well as estimated uncertainties, are made in the accompanying online material in the file "ModelComparisons.SimulatedData.pdf" on pages 1-31. For each parameter, the estimated (or fixed, depending on scenario) value is shown as a solid vertical line and the associated uncertainty, based on the std file produced after inverting the hessian during the model run, is shown as a normal distribution of the same color--centered on the estimated value, with width based on the estimated standard deviation. For the most part, parameter estimates are almost identical across the scenarios and estimated uncertainties are small. Not surprisingly, though, uncertainties are larger for estimated parameters related to natural mortality (pLnDMX, pLnDMM, and pLnDMXM; pp. 7-8), growth (pLnGrA and pLnGrB, p. 8), double-logistic selectivity functions (pS3 and pS4; p. 17) and survey catchability (pLnDQX, p. 31). More surprising, however, is the increase in uncertainty in female survey selectivity parameters (pS1[02] and pS1[04], p. 13) when survey Q-related parameters are estimated (scenarios A4C and A4F).

Models were fit to "survey" aggregated abundance data, aggregated biomass data, and size compositions (Figure 1) by sex, maturity (immature, mature) and shell condition (new shell, old shell) combinations. For the "directed" fishery TCF, the models were fit by sex and shell condition to retained catch, discard catch and total catch data for aggregated abundance, aggregated biomass, and size compositions. The bycatch fisheries were fit by sex to aggregated total (i.e., discard) catch abundance, biomass and size compositions. Lognormal likelihoods were used to fit aggregated abundance and biomass data, while multinomial likelihoods were used to fit size compositions. Comparison across the scenarios for the values of various penalties, priors and data likelihood components to the objective function reveal very small changes between scenarios ("ModelComparisons.SimulatedData.pdf", pp. 32-46).

Estimated values are very similar to true values (indicated by "rsim") across all estimating scenarios for the probability of molt-to-maturity ("ModelComparisons.SimulatedData.pdf", p. 50; all scenarios), natural mortality (p. 47;and scenarios A4D-F), and growth (pp. 48-49, scenarios A4E and F). Similarly, survey and fishery selectivity curves appear to be well-estimated (pp. 64-69) and almost identical to the true ("rsim"), with the largest (but small) differences occurring for females in the 1982+ surveys (curve 04; p. 64) for scenario A4F and males in the SCF in 1978-1996 (curve 09; p. 65) for all scenarios. However, fully-selected Q in the survey is somewhat overestimated for females (and slightly underestimated for males) in both time periods for scenario A4F, potentially indicating difficulties when attempting to simultaneously estimate survey Q and growth (Figure 2).

The recruitment time series estimated by the models is virtually identical across all scenarios (Figure 3). Once the models are informed with data (catch information from the directed fishery begins in 1965), TCSAM2015 appears to do a very good, but not perfect, job of estimating recruitment to the population.

The models do not track variation in recruitment prior to about 1960 (the model starts in 1950), but follow an averaged trajectory until about 1958, when they start to track annual fluctuations in recruitment with increasing accuracy until 1970, after which they track recruitment with very little error until the final two model years, which correspond to a large swing in recruitment. The inability to track recruitment in the earliest years leads to all scenarios slightly overestimating the true initial population abundance-at-size (Figure 4). As with recruitment, population abundance and biomass time series appear to track a smoothed average of the true time series until into the 1960s, but then all six scenarios track the true time series very well (Figures 5 and 6).

As a consequence of the results presented here, as well as other testing TCSAM2015 has undergone, it appears the model code is functioning correctly (e.g., handling multiple data sources and model time blocks correctly) and is capable of population-related quantities and time series in an unbiased fashion, at least with perfect data.

Model runs with real data

Model runs have also been made using real data from the 2015 assessment. Ideally, a comparison between "equivalent" TCSAM2015 and TCSAM2013 model runs would result in identical results, and TCSAM2015 could be adopted immediately as the standard code used in the stock assessment. At present, however, it may not be possible to achieve enough "equivalency" between model runs using the different codes because, although many of the major features of the models can be (by selecting appropriate options) duplicated, some of the "details" (like the parameterizations for natural mortality) differ. Certainly, though, a better attempt can be made than the one for which results are presented here.

As a first attempt at comparing equivalent models, the TCSAM2015 code was run using the 2015 assessment data in two configurations, "4GC" and "4G", analogous to TCSAM2013 models A-J and A-J.L0 described in Part 1. Like A-J, scenario 4GC was run with the Gmacs FMM fitting fishery catch mortality (retained and discard) using normal likelihoods and assumed error variances of 2 while, like A-J.L0, scenario 4G was also run with the Gmacs FMM but fitting fishery catch mortality using lognormal likelihoods with assumed error cv's of 0.05 for retained catch and 0.20 for discard mortality. The TCSAM2015 scenarios incorporated the same time periods as the TCSAM2013 scenarios for survey catchabilities and selectivities, fishery selectivities, and natural mortality. However, the TCSAM2015 scenarios started in 1949 and employed the Gmacs cumulative gamma pdf growth function, while the TCSAM2013 scenarios started in 1930 and used a normalized gamma pdf growth function. Relative weightings between likelihood components may not have been identical, as well.

The TCSAM2015 scenarios were each run over 150 times using randomly-selected (jittered) initial values for all estimated parameters, excluding "dev" parameters and parameters related to the probability of molt-to-maturity (terminal molt). Dev parameters were initialized to 0 and the parameters related to the probability of molt-to-maturity were initialized as (logit-scale) linear functions of size. For each scenario, the model run with the smallest final objective function value and maximum gradient was assumed to have converged at the global minimum objective function value for the given model. Objective function values were not comparable between models.

In 1975, the first year for the NMFS trawl survey data, the TCSAM2015 scenarios fit mature male biomass better than mature female biomass, while the opposite is true for the TCSAM2013 scenarios (Figure 7). Otherwise, the fits to mature biomass from the NMFS trawl surveys are reasonably similar for all four scenarios, and the model estimates are more similar to each other than they are to the observations. Fits to retained catch are quite good for scenario 4G, and appear to be better than the equivalent TCSAM2013 model (A-J.L0), while they are rather poor for scenario 4GC, which appears to also fit more poorly than TCSAM2013 scenario A-J (Figure 8). After 1991, when observations became available, similar results were obtained for total catch mortality for males in the directed fishery (Figure

9). Excluding the 1975-1980 period, the estimated trends prior to 1992 for total catch mortality are remarkably similar across all four scenarios. During the 1975-1980 period, however, the TCSAM2015 scenarios estimate peaks in male catch mortality in 1980, while the TCSAM2013 scenarios estimate peaks in 1978.

Estimated time series for MMB from the four models exhibit substantial differences prior to 1990, but comparatively small differences after 1990 (Figure 10). Similar observations hold for estimated recruitment time series, although there is also clearly a scaling bias: the trends (post-1990) are similar but the scales are different. Clearly, though, the two sets of model codes do not yet yield equivalent results.

These differences between the TCSAM2015 and TCSAM2013 model runs may partly be due to differences in when the models were started, although the small differences between Scenarios A-G and A-H in Part 1 would suggest this is not the case. Alternatively, the manner in which they were initialized may play a role: the TCSAM2013 scenarios fixed fishing mortality rates to small values (≤ 0.05) for each fishery from the model start up to the first year data (catch or effort) were available, while the TCSAM2015 scenarios assumed fishing mortality rates were zero when no data was available. Ideally, model results should not be sensitive to such assumptions regarding initialization but whether or not this is the case remains to be tested. It may also be that the weightings, penalties and priors applied in the objective function in the two model codes are different enough to produce the inequality of the models compared here. This matter will be looked into more carefully and resolved before the TCSAM2015 model code, despite its advantages over the current code, should be adopted for the Tanner crab stock assessment.

Recommendations
- The factors contributing to the differences between "equivalent" TCSAM2015 and TCSAM2013 models should be identified and resolved.
- Assuming the previous bullet has been addressed, the author's preferred model for the 2016 stock assessment should be run using the TCSAM2015 code as well as the TCSAM2013 code.

Literature Cited
Stockhausen, W. 2015. 2015 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: 2015 Final Crab SAFE. North Pacific Fishery Management Council. Anchorage, AK.
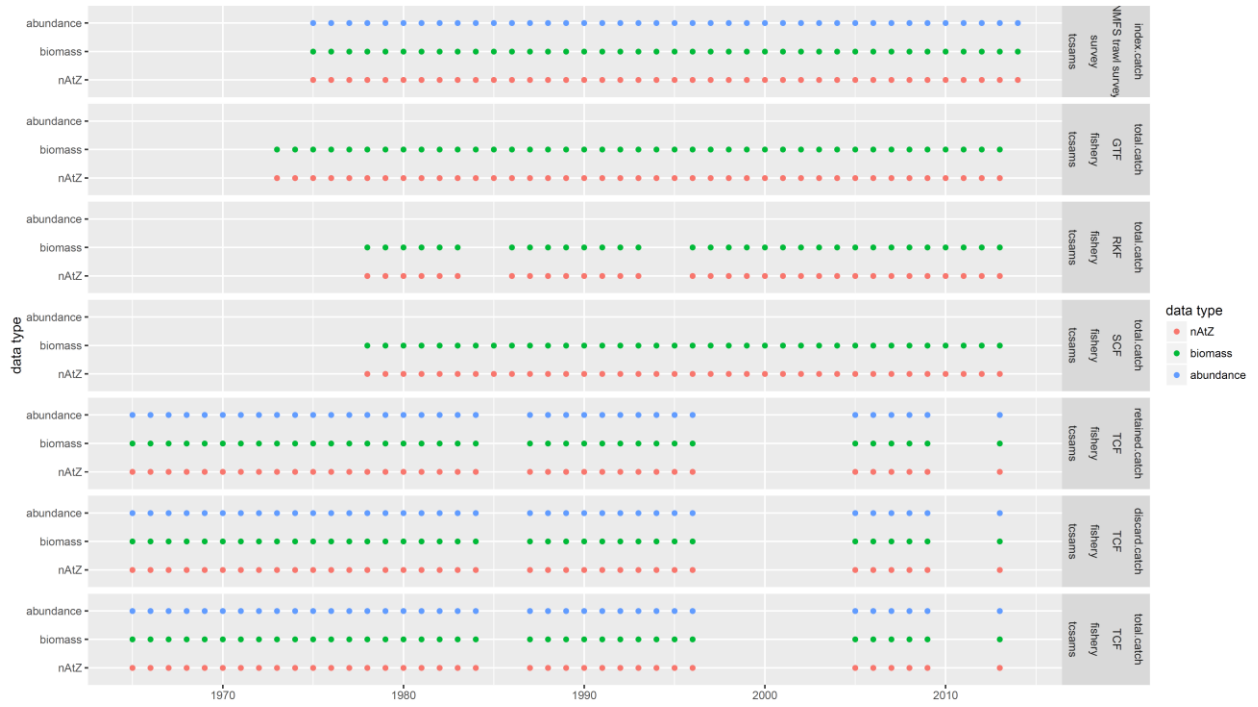
Figures

Figure 1. Simulated data used in testing TCSAM2015. "nAtZ" refers t0 numbers-at-size data (size composition data).
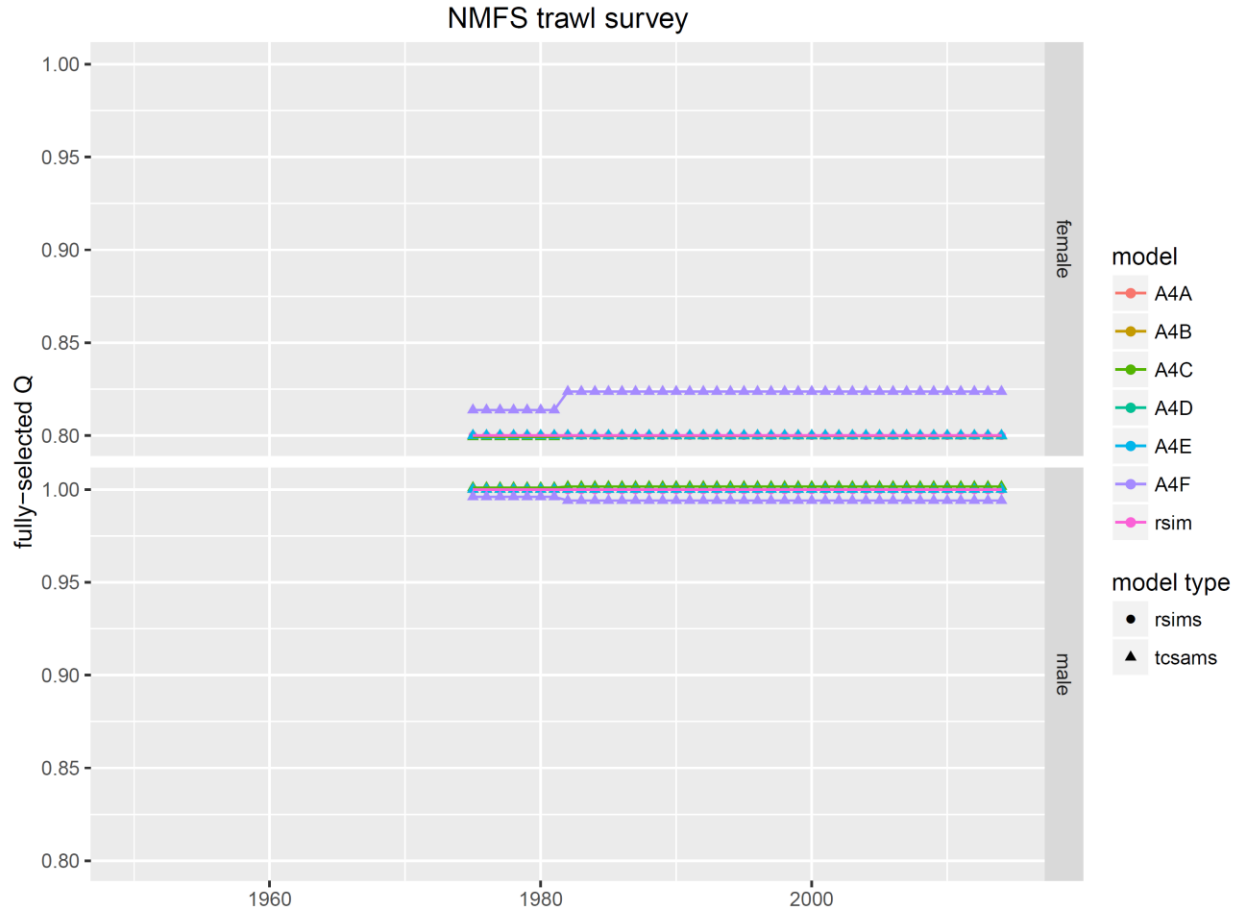
Figure 2. Comparison of the estimated fully-selected survey Q by sex for the two survey time periods across scenarios for simulated data. "rsim" indicates the true values used in the simulation. Related parameters were estimated only in scenarios A4C and A4F.

Figure 3. Comparison of the estimated recruitment time series across scenarios for simulated data. "rsim" indicates the true time series from the simulation.
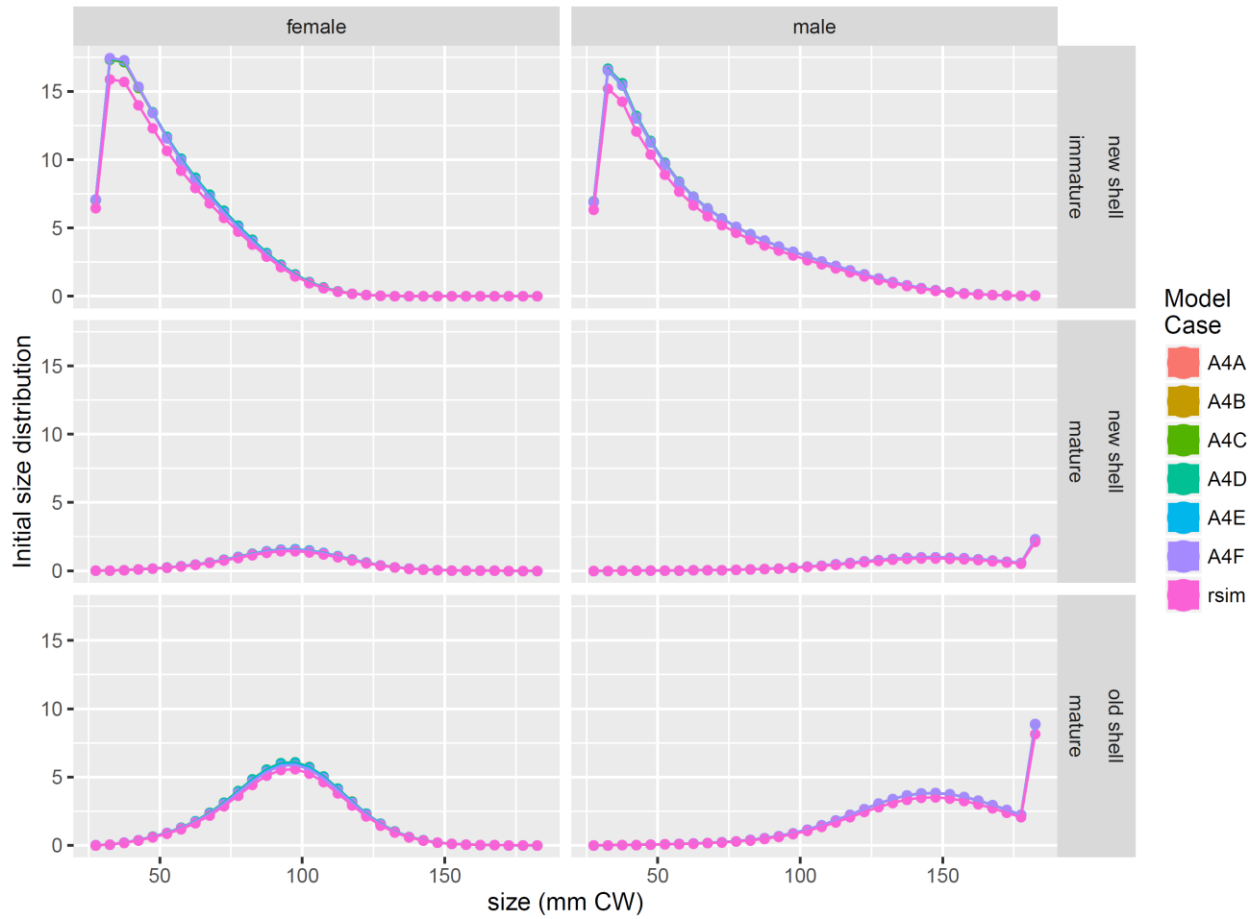
Figure 4. Comparison of the estimated initial population abundance by size across scenarios for simulated data. "rsim" indicates the true initial abundance at size in the simulation.
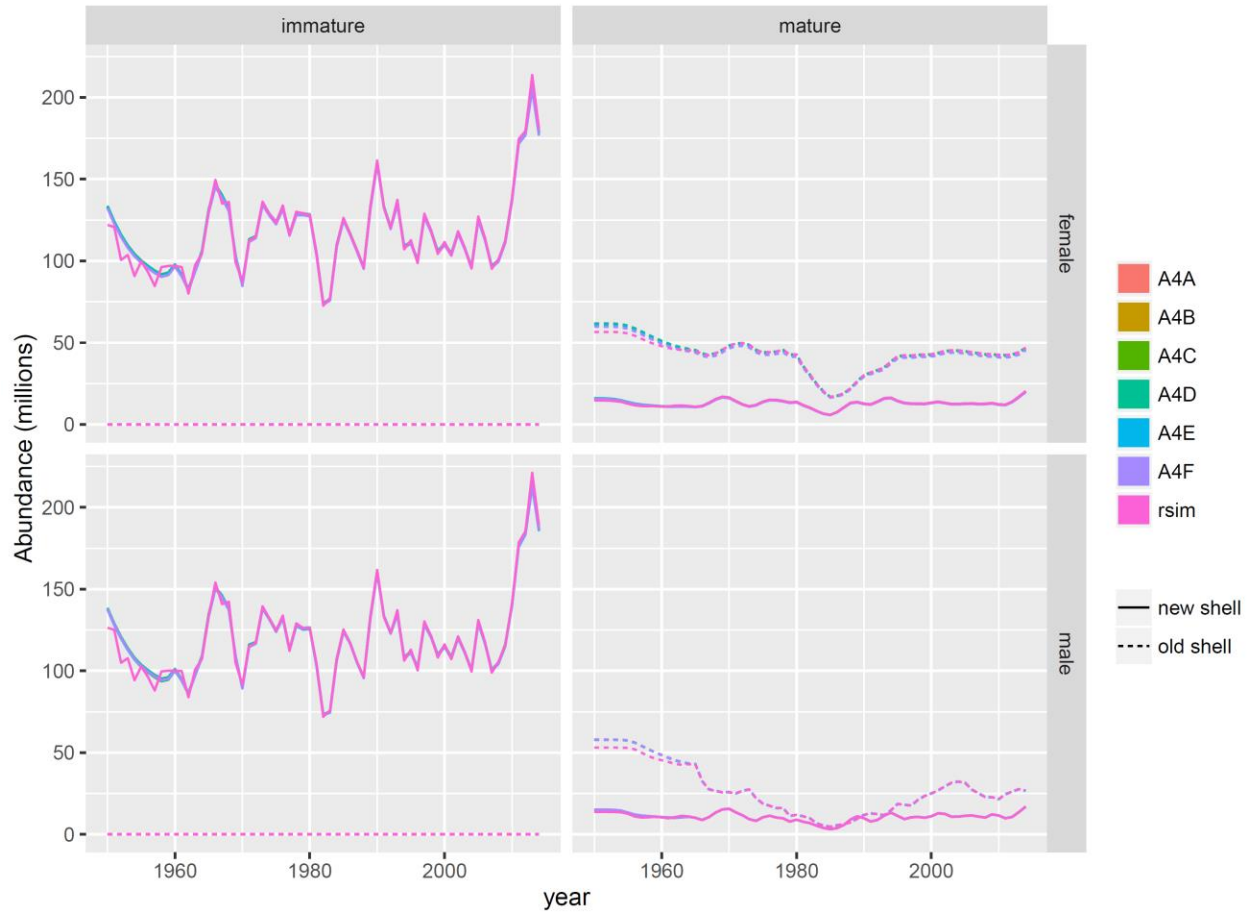
Figure 5. Comparison of the estimated population abundance time series by sex *x* maturity *x* shell condition across scenarios for the simulated data. "rsim" indicates the true values from the simulation.
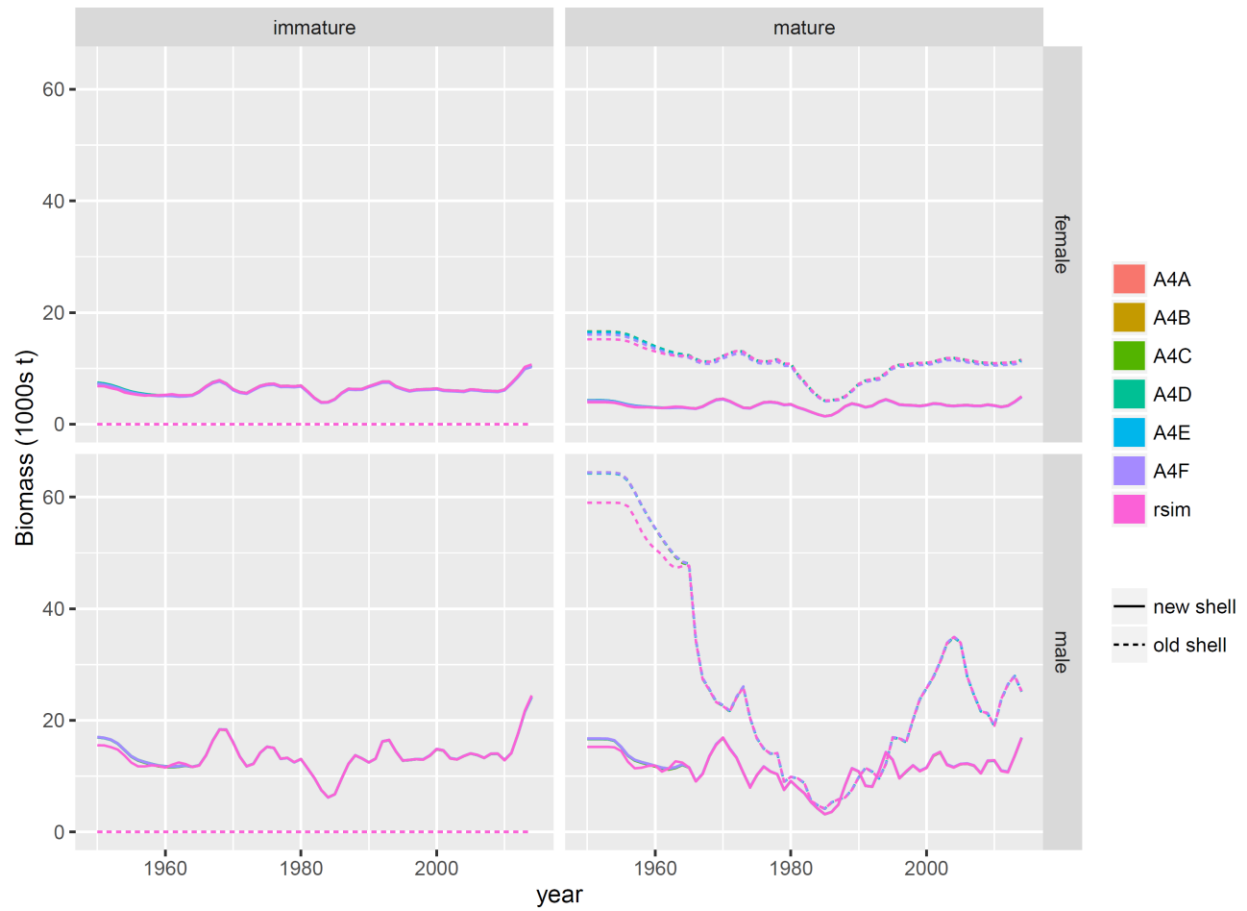
Figure 6. Comparison of the estimated population biomass time series by sex *x* maturity *x* shell condition across scenarios for simulated data. "rsim" indicates the true time series values from the simulation.
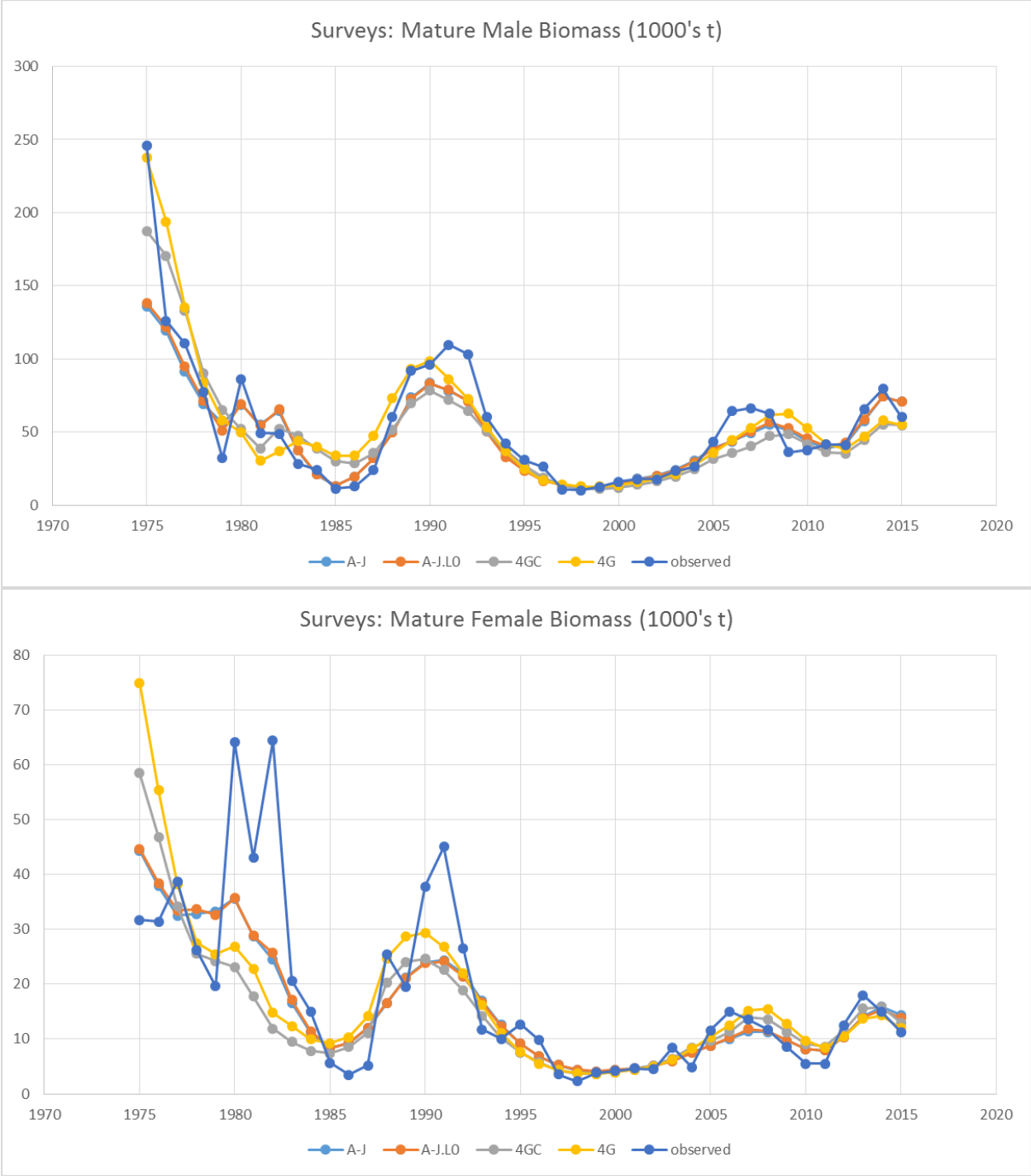
Figure 7. Comparison of fits to survey biomass for mature males (upper plot) and mature females (lower plot) across scenarios for TCSAM2013 (A-J, A-J.L0) and TCSAM2015 (4GC, 4G).
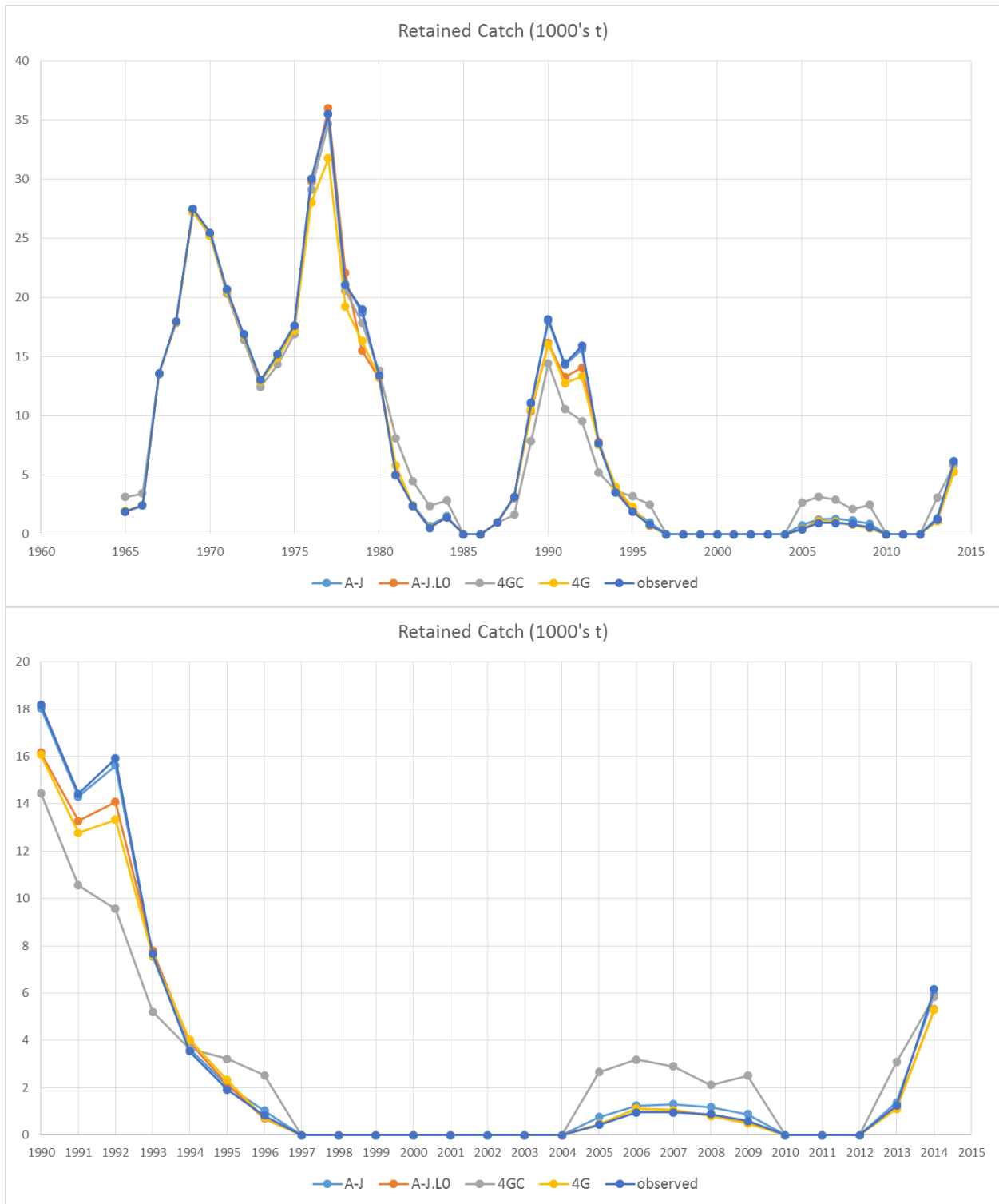
Figure 8. Comparison of fits to retained catch across scenarios for TCSAM2013 (A-J, A-J.L0) and TCSAM2015 (4GC, 4G).

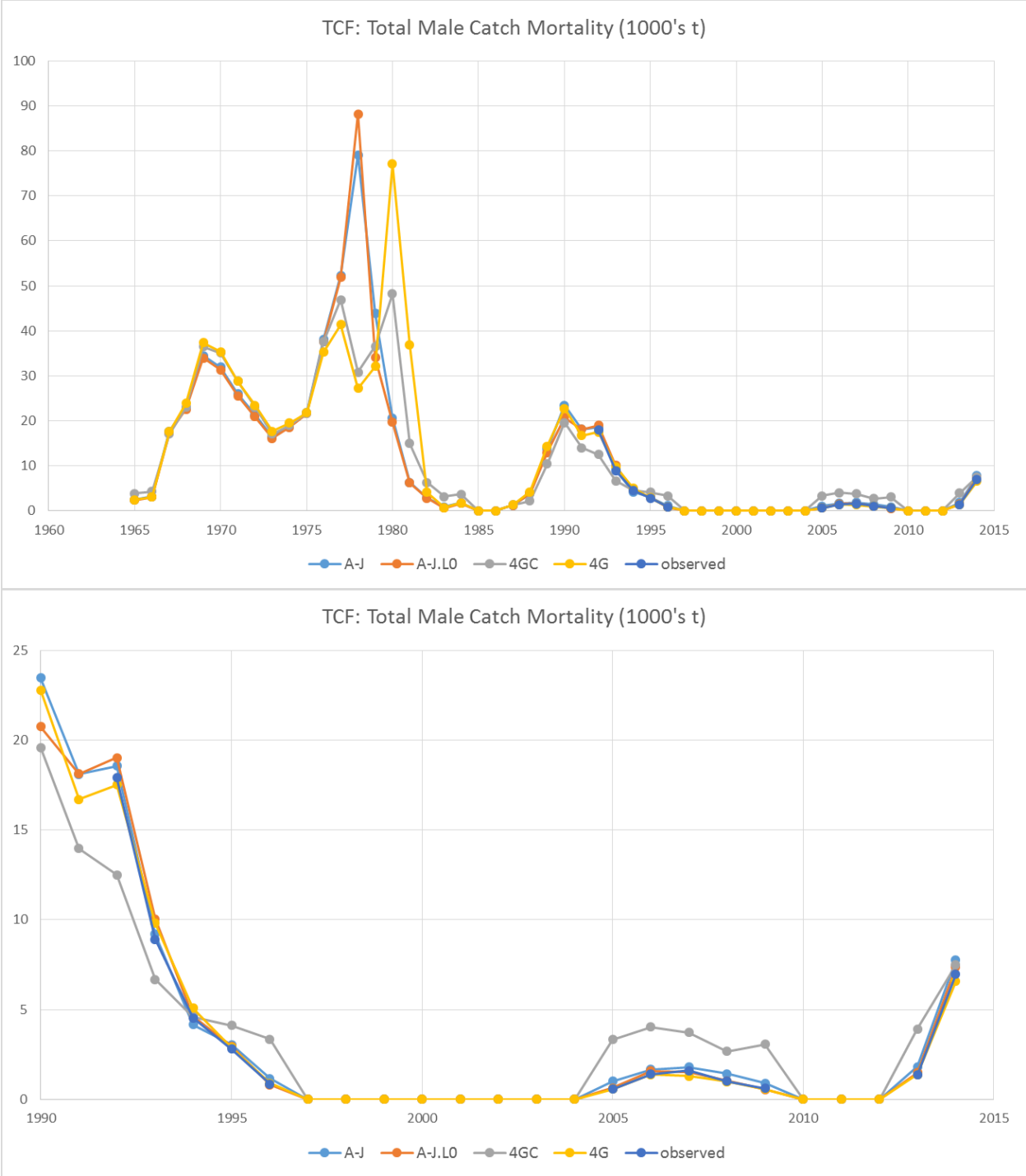Figure 9. Comparison of fits to total catch mortality (retained + discard) for males in the directed fishery across scenarios for TCSAM2013 (A-J, A-J.L0) and TCSAM2015 (4GC, 4G).
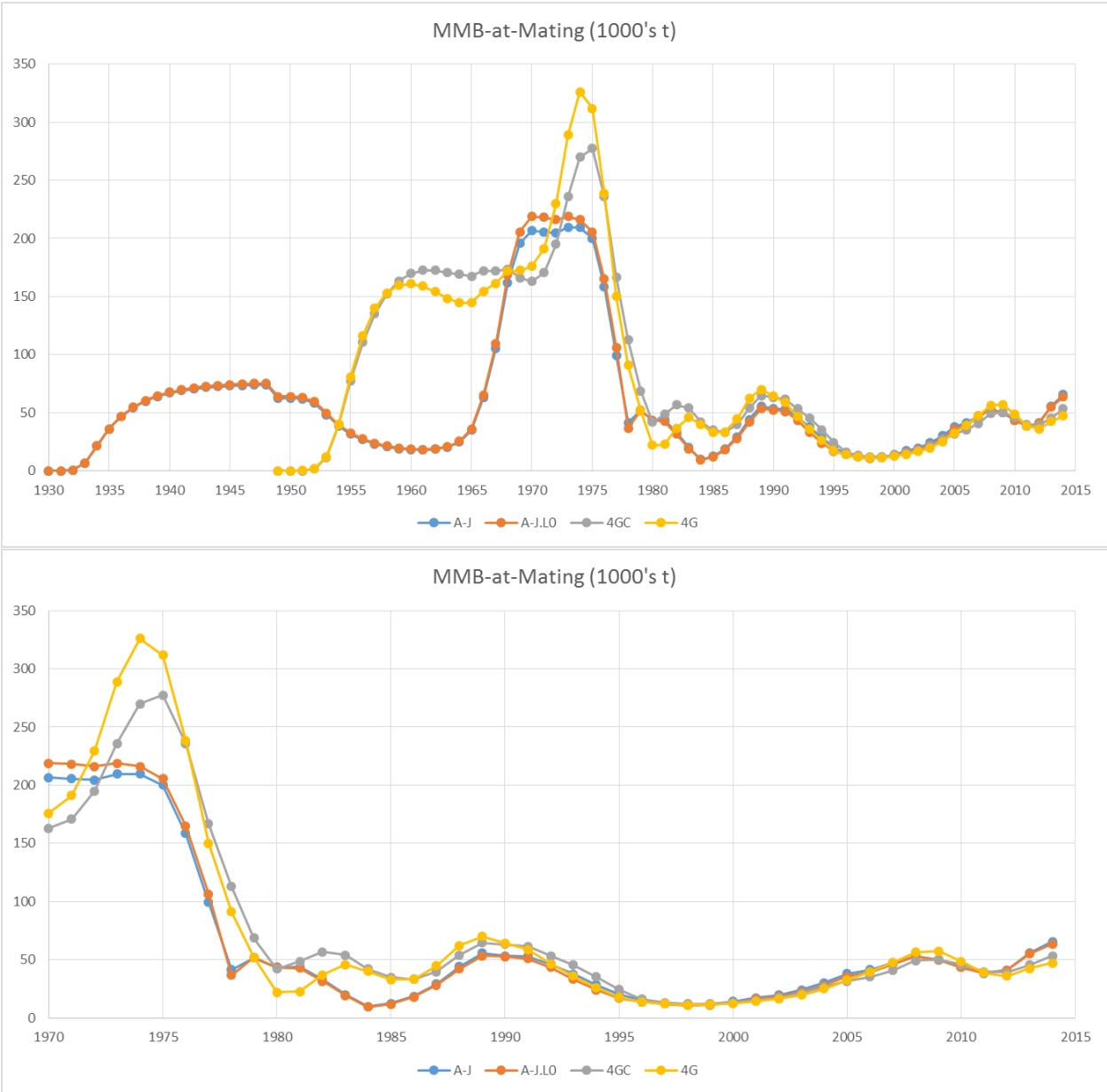
Figure 10. Comparison of the estimated MMB at mating across scenarios for TCSAM2013 (A-J, A-J.L0) and TCSAM2015 (4GC, 4G).
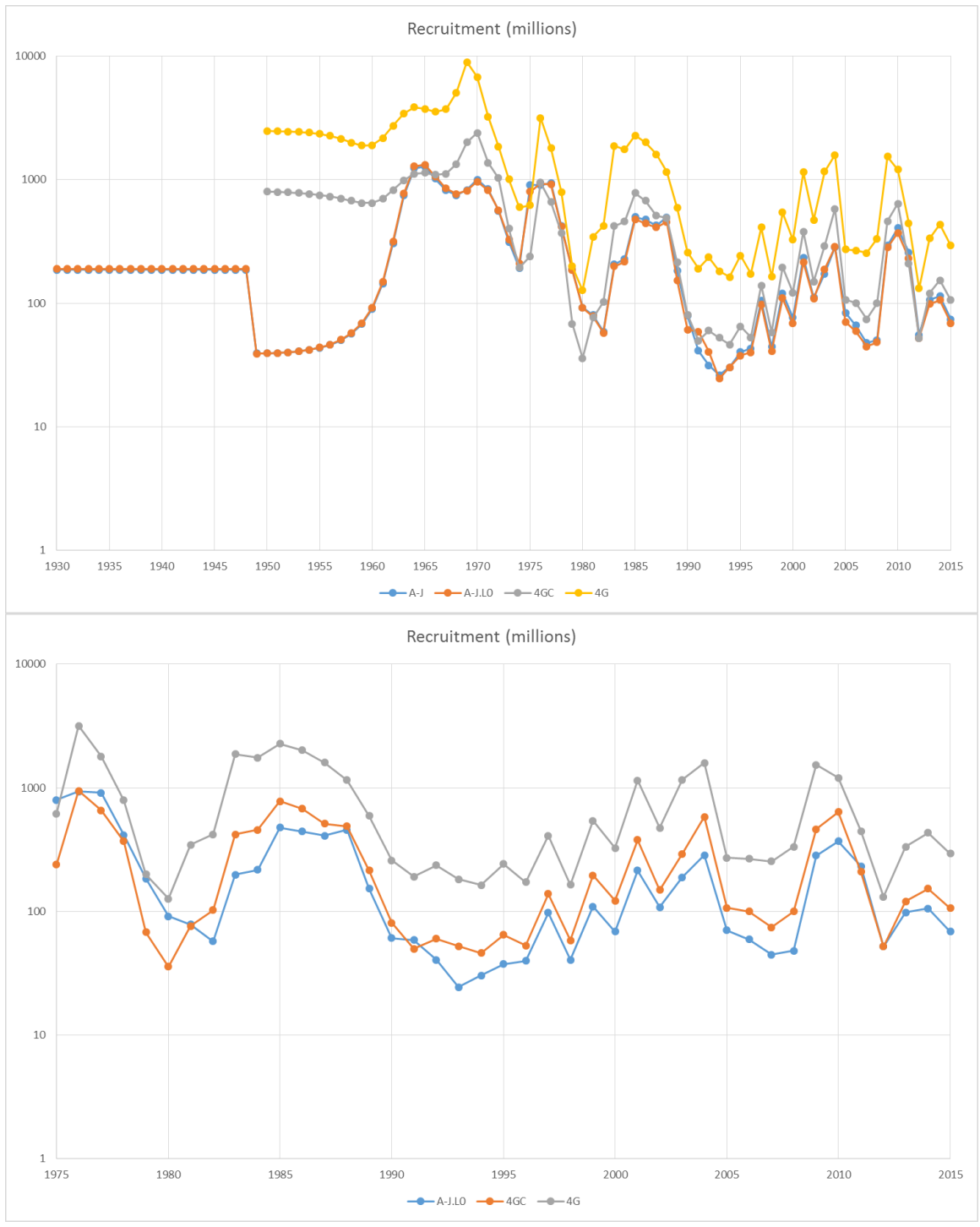
Figure 11. Comparison of the estimated recruitment across scenarios for TCSAM2013 (A-J, A-J.L0) and TCSAM2015 (4GC, 4G).

Appendix A: TCSAM (Tanner Crab Stock Assessment Model) 2015 Description

A. General population dynamics

Population abundance at the start of year y in the model, $n_{y,x,m,s,z}$, is characterized by sex x (male, female), maturity state m (immature, mature), shell condition s (new shell, old shell), and size z (carapace width, CW). Changes in abundance due to natural mortality, molting and growth, maturation, fishing mortality and recruitment are tracked on an annual basis. Because the principal crab fisheries occur during the winter, the model year runs from July 1 to June 30 of the following calendar year.

The order of calculation steps to project population abundance from year y to $y+1$ depends on the assumed timing of the fisheries (δt_y^F) relative to molting (δt_y^m) within year y. The steps when $\delta t_y^F \leq \delta t_y^m$ are outlined below first (Steps A1.1-A1.4), followed by the steps when $\delta t_y^m < \delta t_y^F$. (Steps A2.1-A2.4).

A1. Calculation sequence when $\delta t_y^F \leq \delta t_y^m$

Step A1.1: Survival prior to fisheries
Natural mortality is applied to the population from the start of the model year (July 1) until just prior to prosecution of pulse fisheries for year y at δt_y^F. The numbers surviving at δt_y^F in year y are given by:

$$n_{y,x,m,s,z}^1 = e^{-M_{y,x,m,s,z} \cdot \delta t_y^F} \cdot n_{y,x,m,s,z} \qquad \text{A1.1}$$

where M represents the annual rate of natural mortality in year y on crab classified as x, m, s, z.

Step A1.2: Prosecution of the fisheries
The directed fishery and bycatch fisheries are modeled as pulse fisheries occurring at δt_y^F in year y. The numbers that remain after the fisheries are prosecuted are given by:

$$n_{y,x,m,s,z}^2 = e^{-F_{y,x,m,s,z}^T} \cdot n_{y,x,m,s,z}^1 \qquad \text{A1.2}$$

where $F_{y,x,m,s,z}^T$ represents the total fishing mortality (over all fisheries) on crab classified as x, m, s, z in year y.

Step A1.3: Survival after fisheries to time of molting/mating
Natural mortality is again applied to the population from just after the fisheries to the time at which molting/mating occurs for year y at δt_y^m (generally Feb. 15). The numbers surviving at δt_y^m in year y are then given by:

$$n_{y,x,m,s,z}^3 = e^{-M_{y,x,m,s,z} \cdot (\delta t_y^m - \delta t_y^F)} \cdot n_{y,x,m,s,z}^2 \qquad \text{A1.3}$$

where, as above, M represents the annual rate of natural mortality in year y on crab classified as x, m, s, z.

Step A1.4: Molting, growth, and maturation
The changes in population structure due to molting, growth and maturation of immature (new shell) crab, as well as the change in shell condition for new shell mature crab due to aging, are given by:

$$n_{y,x,MAT,NS,z}^4 = \sum_{z'} \Theta_{y,x,z,z'}^{MAT} \cdot \phi_{y,x,z'} \cdot n_{y,x,IMM,NS,z'}^3 \qquad \text{A1.4a}$$

$$n^4_{y,x,IMM,NS,z} = \sum_{z'} \Theta^{IMM}_{y,x,z,z'} \cdot (1 - \phi_{y,x,z'}) \cdot n^3_{y,x,IMM,NS,z'} \qquad \text{A1.4b}$$

$$n^4_{y,x,MAT,OS,z} = n^3_{y,x,MAT,OS,z} + n^3_{y,x,MAT,NS,z} \qquad \text{A1.4c}$$

where $\phi_{y,x,z}$ is the probability that an immature (new shell) crab of sex x and size z will undergo its terminal molt to maturity and $\Theta^m_{y,x,z,z'}$ is the growth transition matrix from size z' to z for that crab, which may depend on whether ($m=MAT$; eq. A1.4a) or not ($m=IMM$; eq. A1.4b) the terminal molt to maturity occurs. Additionally, crabs that underwent their terminal molt to maturity the previous year are assumed to change shell condition from new shell to old shell (A1.4c). Note that the numbers of immature old shell crab are identically zero in the current model because immature crab are assumed to molt each year until they undergo the terminal molt to maturity, consequently the corresponding equation for $m=IMM$, $s=NS$ above is unnecessary.

Step A1.5: Survival to end of year, recruitment, and update to start of next year
Finally, population abundance at the start of year $y+1$ due to natural mortality on crab from the time of molting in year y until the end of the model year (June 30) and recruitment of immature new shell (IMM, NS) crab at the end of year y ($R_{y,x,z}$) are given by:

$$n_{y+1,x,m,s,z} = \begin{cases} e^{-M_{y,x,IMM,NS,z} \cdot (1 - \delta t^m_y)} \cdot n^4_{y,x,IMM,NS,z} + R_{y,x,z} & m = IMM, s = NS \\ e^{-M_{y,x,m,s,z} \cdot (1 - \delta t^m_y)} \cdot n^4_{y,x,m,s,z} & otherwise \end{cases} \qquad \text{A1.5}$$

A2. Calculation sequence when $\delta t^m_y < \delta t^F_y$

Step A2.1: Survival prior to molting/mating
As in the previous sequence, natural mortality is first applied to the population from the start of the model year (July 1), but this time until just prior to molting/mating in year y at δt^m_y (generally Feb. 15). The numbers surviving at δt^m_y in year y are given by:

$$n^1_{y,x,m,s,z} = e^{-M_{y,x,m,s,z} \cdot \delta t^m_y} \cdot n_{y,x,m,s,z} \qquad \text{A2.1}$$

where M represents the annual rate of natural mortality in year y on crab classified as x, m, s, z.

Step A2.2: Molting, growth, and maturation
The changes in population structure due to molting, growth and maturation of immature (new shell) crab, as well as the change in shell condition for new shell mature crab due to aging, are given by:

$$n^2_{y,x,MAT,NS,z} = \sum_{z'} \Theta^{MAT}_{y,x,z,z'} \cdot \phi_{y,x,z'} \cdot n^1_{y,x,IMM,NS,z'} \qquad \text{A2.2a}$$

$$n^2_{y,x,IMM,NS,z} = \sum_{z'} \Theta^{IMM}_{y,x,z,z'} \cdot (1 - \phi_{y,x,z'}) \cdot n^1_{y,x,IMM,NS,z'} \qquad \text{A2.2b}$$

$$n^2_{y,x,MAT,OS,z} = n^1_{y,x,MAT,OS,z} + n^1_{y,x,MAT,NS,z} \qquad \text{A2.2c}$$

where $\phi_{y,x,z}$ is the probability that an immature (new shell) crab of sex x and size z will undergo its terminal molt to maturity and $\Theta^m_{y,x,z,z'}$ is the growth transition matrix from size z' to z for that crab, which may depend on whether ($m=MAT$; eq. A2.2a) or not ($m=IMM$; eq. A2.2b) the terminal molt to maturity occurs. Additionally, crabs that underwent their terminal molt to maturity the previous year are assumed to change shell condition from new shell to old shell (A2.2c). Again, the numbers of immature old shell crab are identically zero in the current model because immature crab are assumed to molt each year until they undergo the terminal molt to maturity, consequently the corresponding equation for $m=IMM$, $s=NS$ above is unnecessary.

Step A2.3: Survival after molting/mating to prosecution of fisheries
Natural mortality is again applied to the population from just after molting/mating to the time at which the fisheries occur for year y (at δt^F_y). The numbers surviving at δt^F_y in year y are then given by:

$$n^3_{y,x,m,s,z} = e^{-M_{y,x,m,s,z} \cdot (\delta t^F_y - \delta t^m_y)} \cdot n^2_{y,x,m,s,z}$$ A2.3

where, as above, *M* represents the annual rate of natural mortality in year y on crab classified as *x, m, s, z*.

Step A2.4: Prosecution of the fisheries
The directed fishery and bycatch fisheries are modeled as pulse fisheries occurring at δt^F_y in year y. The numbers that remain after the fisheries are prosecuted are given by:

$$n^4_{y,x,m,s,z} = e^{-F^T_{y,x,m,s,z}} \cdot n^3_{y,x,m,s,z}$$ A2.4

where $F^T_{y,x,m,s,z}$ represents the total fishing mortality (over all fisheries) on crab classified as *x, m, s, z* in year y.

Step A2.5: Survival to end of year, recruitment, and update to start of next year
Finally, population abundance at the start of year y+1 due to natural mortality on crab from just after prosecution of the fisheries in year y until the end of the model year (June 30) and recruitment of immature new (IMM, NS) shell crab at the end of year y ($R_{y,x,z}$) and are given by:

$$n_{y+1,x,m,s,z} = \begin{cases} e^{-M_{y,x,IMM,NS,z} \cdot (1-\delta t^F_y)} \cdot n^4_{y,x,IMM,NS,z} + R_{y,x,z} & m = IMM, s = NS \\ e^{-M_{y,x,m,s,z} \cdot (1-\delta t^F_y)} \cdot n^4_{y,x,m,s,z} & otherwise \end{cases}$$ A2.5

B. Model processes: natural mortality
At its most general, natural mortality $M_{y,x,m,s,z}$ is parameterized as a time-varying (in blocks of years) function of sex, maturity state, and size using the following functional form:

$$lnM_{y,x,m} = \mu^0 + \mu^0_t + \delta_{m,MAT} \cdot \mu^{MAT}_t + \delta_{x,FEM} \cdot \mu^{FEM}_t + \delta_{x,FEM} \cdot \delta_{m,MAT} \cdot \mu^{FEM,MAT}_t$$ B.1

$$M_{y,x,m,s,z} = \begin{cases} \exp(lnM_{y,x,m}) & if\ Lorenzen\ option\ is\ not\ selected\ for\ block\ t \\ \exp(lnM_{y,x,m}) \cdot \frac{z_{base}}{z} & if\ Lorenzen\ option\ is\ selected\ for\ block\ t \end{cases}$$ B.2a / B.2b

where y falls into time block *t*, the μ's are (potentially) estimable parameters on the ln-scale, , $\delta_{i,j}$ is 1 if $i=j$ and 0 otherwise. μ^0 represents the baseline (ln-scale) natural mortality rate on immature males, while

μ_t^0 is the offset on immature males in time block t, μ_t^{MAT} is the offset for mature crab in time block t, μ_t^{FEM} is the offset for females in time block t, and $\mu_t^{FEM,MAT}$ is the offset for mature females in time block t. As an option, one can include (by time block) size dependence in natural mortality using Lorenzen's approach (eq. B.2b), where z_{base} is a specified reference size (mm CW).

This parameterization for natural mortality differs from that in TCSAM2013 (Part 1, Appendix A, Section B). In TCSAM2013, sex/maturity-state variations to the base mortality rate are estimated on the arithmetic scale, whereas here they are estimated on the ln-scale. The latter approach may be preferable in terms of model convergence properties because the arithmetic-scale parameter values must be constrained to be positive by placing limits on their values whereas the ln-scale parameter values do not. However, the use of strong priors on the arithmetic-scale parameters in TCSAM2013 (Part 1, Appendix A, eq. B3) probably addresses this issue satisfactorily. TCSAM2013 also incorporates the ability to estimate additional effects on natural mortality during the 1980-1984 time period, but this time block is hard-wired in the code; thus investigating how changes to this time block affect the assessment require modifying and recompiling the code for every alternative time block considered. A similar study using TCSAM2015 would not require modifying the model code because time blocks can be defined for any model process (e.g., natural mortality) in the model input files.

C. Model processes: growth

Annual growth of immature crab in TCSAM2015 can be based on the same approach used in TCSAM2013, except that growth can vary by time block. As such, growth is expressed by sex-specific transition matrices $\Theta_{t,x,z,z'}$ that specify the probability that crab in pre-molt size bin z grow to post-molt size bin z' during time block t. The sex-specific growth matrix $\Theta_{t,x,z,z'}$ is given by

$\Theta_{t,x,z,z'} = c_{t,x,z} \cdot \Delta_{z,z'}{}^{\alpha_{t,x,z}-1} \cdot e^{-\frac{\Delta_{z,z'}}{\beta_{t,x}}}$	Sex-specific (x) transition matrix for growth from pre-molt z to post-molt z', with $z' \geq z$	C.1
$c_{x,z} = \left[\sum_{z'} \Delta_{z,z'}{}^{\alpha_{t,x,z}-1} \cdot e^{-\frac{\Delta_{z,z'}}{\beta_{t,x}}} \right]^{-1}$	Normalization constant so $1 = \sum_{z'} \Theta_{x,z,z'}$	C.2
$\Delta_{z,z'} = z' - z$	Actual growth increment	C.3
$\alpha_{t,x,z} = [\bar{z}_{t,x,z} - z]/\beta_{t,x}$	Mean molt increment, scaled by β_x	C.4
$\bar{z}_{t,x,z} = e^{a_{t,x}} \cdot z^{b_{t,x}}$	Mean size after molt, given pre-molt size z	C.5

where the $a_{t,x}$, $b_{t,x}$, and $\beta_{t,x}$ (parameters in TCSAM2013) are arithmetic-scale versions of the ln-scale model parameters $pLnA_{t,x}$, $pLnB_{t,x}$, and $pLn\beta_{t,x}$:

$a_{t,x} = e^{pLnA_{t,x}}$		C.6
$b_{t,x} = e^{pLnB_{t,x}}$		C.7

$\beta_{t,x} = e^{pLn\beta_{t,x}}$		C.8

Again, because $a_{t,x}$, $b_{t,x}$, and $\beta_{t,x}$ must be non-negative, the associated parameters in TCSAM2015 are estimated on the ln-scale and transformed to the arithmetic scale. It should be noted that C.1-2 provides an approximation to calculating the cumulative gamma distribution over each size bin. As a preferred alternative, the option is also provided to calculate $\Theta_{t,x,z,z'}$ using ADMB's cumulative gamma function. This option is preferred because it appears to have better numerical properties than the TCSAM2013 approach.

$\Theta_{t,x,z,z'}$ is used to update the numbers-at-size for immature crab, $n_{y,x,z}$, from pre-molt size z to post-molt size z' using:

$n^+_{y,x,z'} = \sum_z n_{y,x,z} \cdot \Theta_{t,x,z,z'}$		C.9

where y falls within time block t.

Priors using normal distributions are imposed on $a_{t,x}$ and $b_{t,x}$ in TCSAM2013, with the values of the hyper-parameters hard-wired in the model code (App. 1, Section C). While priors may be defined for the associated parameters here, these are identified by the user in the model input files and are not hard-wired in the model code.

D. Model processes: maturity

Maturation of immature crab in TCSAM2015 is based on a similar approach to that taken in TCSAM2013, except that the sex- and size-specific probabilities of maturation, $\phi_{t,x,z}$ (where size z is pre-molt size), can vary by time block. After molting, but before assessing growth, the numbers of (new shell) crab remaining immature, $n^+_{y,x,IMM,NS,z}$, and those maturing, $n^+_{x,MAT,NS,z}$, at pre-molt size z are given by:

$n^+_{y,x,IMM,NS,z} = (1 - \phi_{t,x,z}) \cdot n_{y,x,IMM,NS,z}$		D.1a
$n^+_{y,x,MAT,NS,z} = \phi_{t,x,z} \cdot n_{y,x,IMM,NS,z}$		D.1b

where y falls in time block t and $n_{y,x,IMM,NS,z}$ is the number of immature, new shell crab of sex x at pre-molt size z.

The sex- and size-specific probabilities of maturing, $\phi_{t,x,z}$, are related to the logit-scale model parameters $p^{mat}_{t,x,z}$ by:

$\phi_{t,FEM,z} = \begin{cases} \dfrac{1}{1 + e^{p^{mat}_{t,FEM,z}}} & z \le z^{mat}_{t,FEM} \\ 1 & z > z^{mat}_{t,FEM} \end{cases}$	female probabilities of maturing at pre-molt size z	D.2a
$\phi_{t,MALE,z} = \begin{cases} \dfrac{1}{1 + e^{p^{mat}_{t,MALE,z}}} & z \le z^{mat}_{t,MALE} \\ 1 & z > z^{mat}_{t,MALE} \end{cases}$	male probabilities of maturing at pre-molt size z	D.2b

where the $z^{mat}_{t,x}$ are constants specifying the minimum pre-molt size at which to assume all immature crab will mature upon molting. The $z^{mat}_{t,x}$ are used here pedagogically; in actuality, the user specifies the

number of logit-scale parameters to estimate (one per size bin starting with the first bin) for each sex, and this determines the $z_{t,x}^{mat}$ used above.

This parameterization differs from that used in TCSAM2013 (App. 1, Section D). In TCSAM2013, the model parameters are estimated on the ln-scale and constrained to be less than 0 so that the resulting maturation probabilities are between 0 and 1. However, the parameters associated with larger size bins frequently hit the 0 upper bound in TCSAM2013, which may affect overall model convergence and stability. The logit-scale parameters used here may be less problematic in this respect.

Second difference penalties are applied to the parameter estimates in TCSAM2013's objective function to promote relatively smooth changes in these parameters with size. Similar penalties (smoothness, non-decreasing) can be applied in TCSAM2015.

E. Model processes: recruitment

Recruitment of immature (new shell) crab in TCSAM2015 has a similar functional form to that used in TCSAM2013(App. 1, Section E), except that the sex ratio at recruitment is not fixed at 1:1 and multiple time blocks can be specified in the new model (not just the "historical" and "current" blocks defined in TCSAM2013). Recruitment in year y of sex *x* crab at size *z* is specified as

$R_{y,x,z} = \dot{R}_y \cdot \ddot{R}_{y,x} \cdot \dddot{R}_{y,z}$	recruitment of immature, new shell crab	E.1

where $\dot{R}_y$ represents total recruitment in year *y* and $\ddot{R}_{y,x}$ represents the fraction of sex x crab recruiting, and $\dddot{R}_{y,z}$ is the size distribution of recruits, which is assumed identical for males and females.

Total recruitment in year *y*, $\dot{R}_y$, is parameterized as

$\dot{R}_y = e^{pLnR_t + \delta R_{t,y}} \quad y \in t$	total recruitment	E.2

where *y* falls within time block *t*, $pLnR_t$ is the ln-scale mean recruitment parameter for *t*, and $\delta R_{t,y}$ is an element of a "devs" parameter vector for *t* (constrained such that the elements of the vector sum to zero).

The fraction of crab recruiting as sex *x* in year *y* in time block *t* is parameterized using the logistic model

$\ddot{R}_{y,x} = \begin{cases} \dfrac{1}{1 + e^{pLgtRx_t}} & x = MALE \\ 1 - \ddot{R}_{y,MALE} & x = FEMALE \end{cases} \quad y \in t$	sex-specific fraction recruiting	E.3

where $pLgtRx_t$ is the logit-scale parameter determining the sex ratio in time block *t*.

The size distribution for recruits in time block *t*, $\dddot{R}_{t,z}$, is based on a gamma-type distribution and is parameterized as

$\dddot{R}_{t,z} = c^{-1} \cdot \Delta_z^{\frac{\alpha_t}{\beta_t} - 1} \cdot e^{-\frac{\Delta_z}{\beta_t}}$	size distribution of recruiting crab	E.4
$c = \sum_z \Delta_z^{\frac{\alpha_t}{\beta_t} - 1} \cdot e^{-\frac{\Delta_z}{\beta_t}}$	normalization constant so that $1 = \sum_z \dddot{R}_{t,z}$	E.5
$\Delta_z = z + \delta z/2 - z_{min}$	offset from minimum size bin	E.6

$\alpha_t = e^{pLnRa_t}$	gamma distribution location parameter	E.7
$\beta_t = e^{pLnRb_t}$	gamma distribution shape parameter	E.8

where $pLnRa_t$ and $pLnRb_t$ are the ln-scale location and shape parameters and the constant δz is the size bin spacing.

A final time-blocked parameter, $pLnRCV_t$, is associated with the recruitment processes. This parameter represents the ln-scale coefficient of variation (cv) in recruitment variability in time block t. These parameters are used in a penalty/prior on the recruitment "devs" in the model likelihood function.

F. Selectivity and retention functions
Selectivity and retention functions in TCSAM2015 are specified independently from fisheries and surveys in TCSAM2015, but subsequently assigned to them. This allows a single selectivity function to be "shared" among multiple fisheries and/or surveys, and among time blocks and sexes, if so desired.

Currently, the following selectivity/retention functions are available for use in the model:

$S_z = \left\{1 + e^{-\beta \cdot (z - z_{50})}\right\}^{-1}$	standard logistic	F.1
$S_z = \left\{1 + e^{-\ln(19) \cdot \frac{(z - z_{50})}{\Delta z_{95-50}}}\right\}^{-1}$	logistic w/ alternative parameterization	F.2
$S_z = \left\{1 + e^{-\beta \cdot (z - \exp(lnZ_{50}))}\right\}^{-1}$	logistic w/ alternative parameterization	F.3
$S_z = \left\{1 + e^{-\ln(19) \cdot \frac{(z - z_{50})}{\exp(ln\Delta z_{95-50})}}\right\}^{-1}$	logistic w/ alternative parameterization	F.4
$S_z = \dfrac{1}{1 + e^{-\beta_a \cdot (z - z_{a50})}} \cdot \dfrac{1}{1 + e^{\beta_d \cdot (z - z_{d50})}}$	double logistic	F.5
$S_z = \dfrac{1}{1 + e^{-\ln(19) \cdot \frac{(z - z_{a50})}{\Delta z_{a(95-50)}}}} \cdot \dfrac{1}{1 + e^{\ln(19) \cdot \frac{(z - z_{d50})}{\Delta z_{d(95-50)}}}}$	double logistic with alt. parameterization	F.6

A double normal selectivity function (requiring 6 parameters to specify) has also be implemented as an alternative to the double logistic functions. In the above functions, all symbols (e.g., β, Δz_{95-50}) represent parameter values, except "z" which represents crab size.

Selectivity parameters are defined independently of the functions themselves, and subsequently assigned. It is thus possible to "share" parameters across multiple functions. The "parameters" used in selectivity functions are further divided into mean parameters across a time block and annual deviations within the time block. Thus, for example, z_{50} in eq. F1 is actually expressed as $z_{50,y} = \bar{z}_{50} + \delta z_{50,y}$ in terms of model parameters $pS1$ and $pDevsS1_y$, where $\bar{z}_{50} = pS1$ is the mean size-at-50%-selected over the time period and $\delta z_{50,y} = pDevsS1_y$ is the annual deviation. To accommodate the 6-parameter double normal equation, six "mean" parameter sets ($pS1$, $pS2$,..., $pS6$) and six associated sets of "devs" parameter vectors ($pDevsS1$, $pDevsS2$,..., $pDevsS6$) are defined in the model to specify the parameterization of individual selectivity/retention functions.

Finally, three different options to normalize individual selectivity curves are provided: 1) no normalization, 2) specifying a fully-selected size, and 3) re-scaling such that the maximum value of the re-scaled function is 1. A normalization option must be specified in the model input files for each defined selectivity/retention curve.

G. Fisheries

Unlike TCSAM2013, which explicitly models 4 fisheries that catch Tanner crab (one as a directed fishery, three as bycatch), there is no constraint in TCSAM2015 on the number of fisheries that can be incorporated in the model. The only requirement is that each model fishery defined in the input files has a corresponding data component from which parameters can be estimated.

TCSAM2015 uses the Gmacs approach to modeling fishing mortality. The total (retained + discards) fishing mortality rate, $F_{f,y,x,m,s,z}$, in fishery f during year y on crab in state x, m, s, and z (i.e., sex, maturity state, shell condition, and size) is related to the associated fishery capture rate $\phi_{f,y,x,m,s,z}$ by

$F_{f,y,x,m,s,z} = \left[h_{f,t} \cdot (1 - \rho_{f,y,x,m,s,z}) + \rho_{f,y,x,m,s,z}\right] \cdot \phi_{f,y,x,m,s,z}$	fishing mortality rate	G.1

where $h_{f,t}$ is the handling (discard) mortality for fishery f in time block t (which includes year y) and $\rho_{f,y,x,m,s,z}$ is the fraction of crabs in state x, m, s, z that were caught and retained (i.e., the retention function). The retention function is identically 0 for females in a directed fishery and for both sexes in a bycatch fishery. For a directed fishery, the retention function for males is selected from one of the selectivity/retention functions discussed in the previous section.

If $n_{y,x,m,s,z}$ is the number of crab classified as x, m, s, z in year y just prior to the prosecution of the fisheries, then

$c_{f,y,x,m,s,z} = \dfrac{\phi_{f,y,x,m,s,z}}{F^T_{y,x,m,s,z}} \cdot \left[1 - e^{-F^T_{y,x,m,s,z}}\right] \cdot n_{y,x,m,s,z}$	number of crab captured	G.2

is the number of crab classified in that state that were *captured* by fishery f, where $F^T_{y,x,m,s,z} = \sum_f F_{f,y,x,m,s,z}$ represents the total (across all fisheries) fishing mortality on those crab. The number of crab retained in fishery f classified as x, m, s, z in year y is given by

$r_{f,y,x,m,s,z} = \dfrac{\rho_{f,y,x,m,s,z} \cdot \phi_{f,y,x,m,s,z}}{F^T_{y,x,m,s,z}} \cdot \left[1 - e^{-F^T_{y,x,m,s,z}}\right] \cdot n_{y,x,m,s,z}$	number of retained crab	G.3

while the number of discarded crab, $d_{f,y,x,m,s,z}$, is given by

$d_{f,y,x,m,s,z} = \dfrac{(1 - \rho_{f,y,x,m,s,z}) \cdot \phi_{f,y,x,m,s,z}}{F^T_{y,x,m,s,z}} \cdot \left[1 - e^{-F^T_{y,x,m,s,z}}\right] \cdot n_{y,x,m,s,z}$	number of discarded crab	G.4

and the discard mortality, $dm_{f,y,x,m,s,z}$, is

$dm_{f,y,x,m,s,z} = \dfrac{h_{f,y} \cdot (1 - \rho_{f,y,x,m,s,z}) \cdot \phi_{f,y,x,m,s,z}}{F^T_{y,x,m,s,z}} \cdot \left[1 - e^{-F^T_{y,x,m,s,z}}\right] \cdot n_{y,x,m,s,z}$	discard mortality (numbers)	G.5

The biomass associated with the above components is obtained by multiplying each by $w_{x,m,z}$, the associated individual crab weight (estimated outside the model).

The capture rate $\phi_{f,y,x,m,s,z}$ (*not* the fishing mortality rate $F_{f,y,x,m,s,z}$) is modeled as a function separable into separate year and size components such that

$\phi_{f,y,x,m,s,z} = \phi_{f,y,x,m,s} \cdot S_{f,y,x,m,s,z}$	fishing capture rate	G.6

where $\phi_{f,y,x,m,s}$ is the fully-selected capture rate in year y and $S_{f,y,x,m,s,z}$ is the size-specific selectivity.

The fully-selected capture rate $\phi_{f,y,x,m,s}$ for y in time block t is parameterized in the following manner:

$\phi_{f,y,x,m,s} = \exp(\overline{lnC}_{f,t,x,m} + pDevsC_{f,t,y})$	G.7

where the $pDevsC_{f,t,y}$ are elements for year y of time block t of model parameter "devs" vectors representing annual variations from the ln-scale mean fully-selected capture rate $\overline{lnC}_{f,t,x,m}$. The latter is expressed in terms of model parameters as

$\overline{lnC}_{f,t,x,m} = pLnC_f + pLnDCT_{f,t} + \delta_{m,IMM} \cdot pLnDCM_{f,t} + \delta_{x,FEM} \cdot pLnDCX_{f,t}$ $\qquad + \delta_{x,FEM} \cdot \delta_{m,IMM} \cdot pLnDCXM_{f,t}$	G.8

where $pLnC_f$ is the baseline ln-scale capture rate (for mature males), $pLnDCT_{f,t}$ is an additive modifier for time block t, $pLnDCM_{f,t}$ is an additive modifier for immature crab, $pLnDCX_{f,t}$ is the additive modifier for females, and $pLnDCXM_{f,t}$ is the additive modifier for immature females.

H. Surveys
If $n_{y,x,m,s,z}$ is the number of crab classified as x, m, s, z in year y just prior to the prosecution of a survey, then the abundance, $a_{v,y,x,m,s,z}$, and biomass, $b_{v,y,x,m,s,z}$, for crab classified in that state by survey v is given by

$a_{v,y,x,m,s,z} = q_{v,y,x,m,s,z} \cdot n_{y,x,m,s,z}$	survey abundance	H.1
$b_{v,y,x,m,s,z} = w_{x,m,z} \cdot q_{v,y,x,m,s,z} \cdot n_{y,x,m,s,z}$	survey biomass	H.2

where $q_{v,y,x,m,s,z}$ is the size-specific survey catchability on this component of the population and $w_{x,m,z}$ is the associated individual crab weight (estimated outside the model).

The survey catchability $q_{v,y,x,m,s,z}$ is decomposed in the usual fashion into separate time block and size components such that, for y in time block t:

$q_{v,y,x,m,s,z} = q_{v,t,x,m,s} \cdot S_{v,t,x,m,s,z}$	survey catchability	H.3

where $q_{v,t,x,m,s}$ is the fully-selected catchability in time block t and $S_{v,t,x,m,s,z}$ is the size-specific survey selectivity.

The fully-selected catchability $q_{v,t,x,m,s}$ is parameterized in a fashion similar to that for fully-selected fishery capture rates (except that annual "devs" are not included) in the following manner:

$q_{v,t,x,m,s} = \exp(pLnQ_v + pLnDQT_{v,t} + \delta_{m,IMM} \cdot pLnDQM_{v,t} + \delta_{x,FEM} \cdot pLnDQX_{v,t}$ $\qquad + \delta_{x,FEM} \cdot \delta_{m,IMM} \cdot pLnDQXM_{v,t})$	H.4

where $pLnQ_v$ is the baseline ln-scale capture rate (for mature males), $pLnDQT_{v,t}$ is an additive modifier for time block t, $pLnDQM_{v,t}$ is an additive modifier for immature crab, $pLnDQX_{v,t}$ is an additive ln-scale modifier for females, and $pLnDQXM_{v,t}$ is an additive modifier for immature females.

I. Model fitting: objective function equations

The TCSAM2015 model is fit by minimizing an objective function, $\mathcal{O}$, with additive components consisting of: 1) negative log-likelihood functions based on specified prior probability distributions associated with user-specified model parameters, and 2) several negative log-likelihood functions based on input data components, of the form:

| $$\mathcal{O} = -2 \sum_p \lambda_p \cdot \ln(\wp_p) - 2 \sum_l \lambda_l \cdot \ln(\mathcal{L}_l)$$ | model objective function | I.1 |

where $\wp_p$ represents the pth prior probability function, $\mathcal{L}_l$ represents the lth likelihood function, and the λ's represent user-adjustable weights for each component.

Prior Probability Functions
Prior probability functions can be associated with each model parameter or parameter vector by the user in the model input files (see Section K below for examples on specifying priors).

Likelihood Functions
The likelihood components included in the model's objective function are based on normalized size frequencies and time series of abundance or biomass from fishery or survey data. Survey data optionally consists of abundance and/or biomass time series for males, females, and/or all crab (with associated survey cv's), as well as size frequencies by sex, maturity state, and shell condition. Fishery data consists of similar data types for optional retained, discard, and total catch components.

Size frequency components
Likelihood components involving size frequencies are based on multinomial sampling:

| $$\ln(\mathcal{L}) = \sum_y n_{y,c} \cdot \sum_z \left\{ p_{y,c,z}^{obs} \cdot \ln\left(p_{y,c,z}^{mod} + \delta\right) - p_{y,c,z}^{obs} \cdot \ln\left(p_{y,c,z}^{obs} + \delta\right) \right\}$$ | multinomial log-likelihood | I.2 |

where the y's are years for which data exists, "c" indicates the population component classifiers (i.e., sex, maturity state, shell condition) the size frequency refers to, $n_{y,c}$ is the classifier-specific effective sample size for year y, $p_{y,c,z}^{obs}$ is the observed size composition in size bin z (i.e., the size frequency normalized to sum to 1 across size bins for each year), $p_{y,c,z}^{mod}$ is the corresponding model-estimated size composition, and δ is a small constant. The manner in which the observed and estimated size frequencies for each data component are aggregated (e.g., over shell condition) prior to normalization is specified by the user in the model input files. Data can be entered in input files at less-aggregated levels of than will be used in the model; it will be aggregated in the model to the requested level before fitting occurs.

Aggregated abundance/biomass components
Likelihood components involving aggregated (over size, at least) abundance and or biomass time series can be computed using one of three potential likelihood functions: the normal, the lognormal, and the "norm2". The likelihood function used for each data component is user-specified in the model input files.

The ln-scale normal likelihood function is

$\ln(\mathcal{L}^N)_c = -\dfrac{1}{2}\sum\limits_{y}\left\{\dfrac{\left[a_{y,c}^{obs}-a_{y,c}^{mod}\right]^2}{\sigma_{y,c}^2}+\ln\left[\sigma_{y,c}^2\right]\right\}$	normal log-likelihood	I.3

where $a_{y,c}^{obs}$ is the observed abundance/biomass value in year y for aggregation level c, $a_{y,c}^{mod}$ is the associated model estimate, and $\sigma_{y,c}^2$ is the variance associated with the observation.

The ln-scale lognormal likelihood function is

$\ln(\mathcal{L}^{LN})_c = -\dfrac{1}{2}\sum\limits_{y}\left\{\dfrac{\left[ln\left(a_{y,c}^{obs}+\delta\right)-ln\left(a_{y,c}^{mod}+\delta\right)\right]^2}{\sigma_{y,c}^2}+\ln\left[\sigma_{y,c}^2\right]\right\}$	lognormal log-likelihood	I.4

where $a_{y,c}^{obs}$ is the observed abundance/biomass value in year y for aggregation level c c, $a_{y,c}^{mod}$ is the associated model estimate, and $\sigma_{y,c}^2$ is the ln-scale variance associated with the observation.

For consistency with TCSAM2013, a third type, the "norm2", may also be specified

$\ln(\mathcal{L}^{N2})_x = -\sum\limits_{y}\left[a_{y,x}^{obs}-a_{y,x}^{mod}\right]^2$	"norm2" log-likelihood	I.5

This is equivalent to specifying a normal log-likelihood with $\sigma_{y,x}^2 \equiv 0.5$. This is the standard likelihood function applied tin TCSAM2013 to fishery catch time series.

Aggregation fitting levels
A number of different ways to aggregate input data and model estimates prior to fitting likelihood functions have been implemented in TCSAM2015. These include:

Abundance/Biomass	Size Conpositions	
by	by	extended by
total	total	x
x		x, m
x, mature only	x	--
x, m		m
x, s		s
x, m, s	x, m	--
		s
	x, s	
	x, m, s	

where x, m, s refer to sex, maturity state and shell condition and missing levels are aggregated over. For size compositions that are "extended by" x, m, s, or x, m, this involves appending the size compositions corresponding to each combination of "extended by" factor levels, renormalizing the extended composition to sum to 1, and then fitting the extended composition using a multinomial likelihood.

K. Parameter specification for model processes
Parameter specification in TCSAM2015 occurs entirely within the model input files and is extremely flexible in terms of setting initial values, defining upper and lower limits on estimated parameter values,

specifying prior distributions and hyper-prior parameters for use in the model likelihood function, and defining time blocks across which parameters related to a given model process are combined. Parameters are organized in the input files to the model according to the model process (e.g., recruitment, fishing mortality, etc.) the parameter group affects.

Two types of parameters are currently incorporated in TCSAM2015, "number_vector"s and "vector_vector"s. Parameters of the first kind, number_vectors (i.e., a vector of parameter numbers), are used to define and estimate different values (numbers) associated with the same parameter in different time blocks. Different characteristics (e.g., upper and lower limits, initial value, estimation phase) can be associated with each value of a number_vector-type parameter. Parameters of the second kind, vector_vectors (i.e., a vector of parameter vectors), are used to define and estimate different vectors associated with a parameter vector (e.g., a "devs" vector) across different time blocks. Different characteristics (e.g., upper and lower limits, initial value, estimation phase) can be associated with each vector of a vector_vector-type parameter. There are no parameters of the third kind (yet!). Hopefully these terms will be clarified by the following example.

Text Box 1 illustrates an example specification for the recruitment process involving the model parameters *pLnR*, *pLnRCV*, *pLgtRX*, *pLnRa*, *pLnRb* (all number_vectors) and the "devs" parameter vector *pDevsLnR* (a vector_vector). Time blocks are defined for the recruitment process, not for individual parameters. The latter can be used across multiple time blocks. Time blocks are defined in the PARAMETER_COMBINATIONS section (lines 2-6 in the example), and individual parameters are assigned using indices. In the example, two parameter combinations are defined, specifying combinations of the recruitment-associated parameters to two time blocks ("[-1:1974]", i.e. model start year to 1974, and "[1975:-1]", i.e. 1975 to model end year). Recruitment in the first time block is a function of the first parameter definition (id=1) for each of the recruitment parameters, while it is a function of the second parameter definitions (id=2) for *pLnR* and *pDevsLnR* and the first definition for the remaining parameters in the second time block. In the example, the two time blocks are continuous, but it is also possible to define discontinuous blocks (e.g., "[1965:1971; 1980:1990]"). Default index values (-1) correspond to the minimum or maximum index value used for the index type in the model, depending on position in the block definition. For year indices, it is also possible to use "-2" to refer to the current assessment year (-1 refers to the current fishery year).

For each number_vector-type parameter (e.g., *pLnR*, starting at line 8), the user specifies (line 9) the number of different values that will be assigned in the PARAMETER_COMBINATIONS section. For each number, the user specifies (e.g. line 11) the "lower" and "upper" bounds on the value, the default initial value ("init_val"), the "phase" in the model convergence scheme at which the value is first estimated, the likelihood multiplier ("prior_wgt") on the prior associated with the value, the name of the prior to use ("prior_type"; e.g. 'normal' or 'none'), the hyper-parameters associated with the prior ("prior_params"; e.g., mean and standard deviation for a 'normal' prior) and any additional constants required for the function used as the prior. In addition, options ("jitter?", "resample?") for setting the initial value can be turned on or off. If both are "OFF", then the default ("init_val") is used. If jittering is "ON", the initial value will be a random draw between the lower and upper bounds set for the number. If resampling is turned "ON", the initial value will be a random draw based on the prior distribution.

A similar logic applies to parameter vector_vectors (e.g., *pDevsLnR*), except that the user must also specify the type of indexing ("idx.type"; e.g., line 32) used for each vector (one of the model index types: "YEAR", "SEX", "MATURITY_STATE", "SHELL_CONDITION", "SIZE", "FISHERY" or "SURVEY") and define the range for the indices as a "block". The indices the block defines need not be continuous.

Text Boxes

```
 1| recruitment #parameter group name
 2| PARAMETER_COMBINATIONS #required keyword
 3| 2  #number of rows defining parameter combinations
 4| #id  YEAR_BLOCK  pLnR   pLnRCV  pLgtRX    pLnRa    pLnRb  pDevsLnR
 5|  1  [-1:1974]    1      1       1         1        1        1      #model spin-up period
 6|  2  [1975:-1]    2      1       1         1        1        2      #data-informed model
period
 7|. PARAMETERS #required keyword
 8| pLnR #parameter name; ln-scale mean recruitment parameter
 9| 2  #number of parameters
10| #id lower upper jitter? init_val  phase resample? prior_wgt prior_type prior_params
prior_consts
11|  1    0    20    OFF       8       1     OFF        1         normal    10 3
#spin-up period
12|  2    0    20    OFF    11.4       1     OFF        1         normal    10 3
#model period
13| pLnRCV #parameter name; ln-scale parameter for cv of recruitment
14| 1  #number of parameters
15| #id lower  upper jitter? init_val    phase    resample?   prior_wgt   prior_type
prior_params  prior_consts
16|  1   -2.0   2.0   OFF  -0.43275213   -1        OFF           1          none    #full model
period (init_val equiv. to var=0.5)
17| pLgtRX #parameter name; logit-scale parameter for male sex ratio
18| 1  #number of parameters
19| #id lower  upper jitter? init_val  phase   resample? prior_wgt   prior_type  prior_params
prior_consts
20|  1   -1      1    OFF      0       -1        OFF        1         normal     0  0.2
#full model period
21| pLnRa #ln-scale gamma distribution location parameter for pr(size-at-recruitment)
22| 1  #number of parameters
23| #id  lower  upper jitter? init_val    phase   resample?   prior_wgt   prior_type
prior_params   prior_consts
24|  1    1      4     ON   2.442347     -1        OFF           1          normal     2.5 1
#init_val = ln(11.50)
```

Text Box 1. Example parameter specification for recruitment in TCSAM2015. Input values are in black text, comments are in green. Line numbers (text in blue) are shown for reference purposes.