Exploring selectivity, growth, and natural mortality in the assessment for snow crab in the eastern Bering Sea

Cody Szuwalski and Jack Turnock

May 7, 2017

Contents

CPT September 2016 comments, SSC comments, and author response:

CPT and SSC comments

The CPT had several comments and questions related to formatting and presentation from the September 2016 meeting:

- Review SAFE guidelines to make sure required tables and figures are present
- Plot the relative proportion of new to old shell males to see how important the lack of fit to old shell males really is
- Plot Bayesian posterior intervals for growth parameters
- Model 0 has to be last year's accepted model

These issues will be corrected in the SAFE document presented at the September meeting. The CPT also had several suggestions for potential model runs and expanded analyses, including:

- Estimate M for mature females
- Document rationale for prior on M for immature crab
- Try starting the assessment in 1982 to check the behavior of the survey qs when the first survey stanza is excluded
- Apply priors to the survey qs so they are somewhat constrained
- Provide more detailed MCMC chain diagnostics
- Extract bycatch mortality from the Tanner crab directed fisheries that is currently lumped into the groundfish trawl bycatch (in a table in the assessment chapter, not necessarily in the model)

These issues can be roughly grouped by the population processes affected by the request: natural mortality, selectivity/catchability, and growth. Each of the following sections will focus on these processes. The table listing the bycatch mortality from the Tanner crab directed fishery is not included here yet, but it will be in the September document. The SSC endorsed the above recommendations of the CPT and added evaluations of the utility of MCMC as an alternative to jittering and performing a spatial assessment to the list of topics to be considered. These will be discussed below.

Summary of analysis and author recommendations

MCMC diagnostics indicated processes that need attention regardless of whether Bayesian or frequentist methods are to be used, including growth and selectivity. Removing era 1 survey data reduced the number of parameters on their bounds and did not impact management quantities drastically, but it is still not clear how best to use the BSFRF data given implied selectivities and currently estimated selectivities. Shifting back to a linear growth model from the piece-wise linear model eliminated problem parameters, but resulted in poorer fits to survey and growth data. Other growth models could be considered. Estimating mature female M improved fits and returned estimates of mature female and male natural mortality with the expected relationship (i.e. female higher than male). However, immature natural mortality was lower than both, which was unexpected.

Bayesian vs. Maximum likelihood

Background

Each year the total allowable catch (TAC) is calculated from a distribution of the overfishing level (OFL). The methods for which these distributions are estimated varies among assessments, but the goal of producing a distribution is to account for uncertainty in the point estimates of management quantities associated with the data, assessment, and management process. In the past, this was accomplished in the snow crab model by fitting the model to the data, plugging the estimated parameter values into a separate projection model, and initiating the projection model with the numbers-at-length from the final year of assessment. This was done thousands of times, with random error added to the numbers at length to develop a distribution for the OFL. However, this method had some shortcomings. First, the parameter values are not perfectly known, but they were assumed to be so in the calculation of the OFL. Second, the amount of error added to the initial numbers at length was arbitrary and the distribution of the OFL is a direct result of this error. In addition, the practice of copying and pasting output from one model into another can be error prone. Finally, given the structure of the model and the data available, a time-consuming jittering process was necessary to ensure the maximum likelihood estimates were truly the 'best' estimates available.

A Bayesian framework can directly address almost all of these shortcomings. The work flow of a Bayesian approach to developing a distribution for the OFL uses exactly the same model and begins in the same way as a maximum likelihood approach–it fits the model to the data and produces a variance-covariance matrix. Functions to calculate the reference points and the OFL are written into the assessment so that they can be calculated without copying and pasting to another script (this can also be done for a maximum likelihood approach). Then a Markov chain Monte Carlo (MCMC) algorithm is applied to the model that uses the estimated variance-covariance matrix to determine a vector of parameter jumps to explore the posterior distribution of the model. If the likelihood of new parameter vector is higher than the previous vector, the new vector is retained. If the likelihood of the new parameter vector is lower, it is accepted with some probability (exactly what that probability varies based on the algorithm implemented). This process is repeated millions of times and allows characterization of the 'posterior' distribution of the model and parameters which explicitly incorporates uncertainty in all parameters and whatever prior information is available into management quantities. The parameters and derived quantities are all treated as distributions and these distributions are a function of the data (rather than an arbitrary constant). A Bayesian framework also obviates the need for 'jittering' because it does not rely on finding the absolute best values to fit the data, but rather on sufficient length of the 'chain' (i.e. the number of iterations of the MCMC algorithm) and appropriate model structure to allow the algorithm to appropriately explore the likelihood space an develop a distribution of parameters and derived quantities.

If the goal of assessment is to produce a distribution of the OFL to be used in management that incorporates uncertainty in the model and data, Bayesian methods are a natural choice.

Diagnostics

Although Bayesian methods are used specifically to produce distributions of parameters and derived quantities within complex models, they are somewhat more difficult to fit than frequentist models and require extensive diagnostics to ensure appropriate inference. The var/covar matrix plays a central role in MCMC and therefore needs to be appropriately estimated. Parameters hitting their bounds can be problematic for this process, so parameters estimated away from their bounds is one of the first diagnostics to be examined (and should also be done within ML frameworks that use the resulting var/covar matrices). In the snow crab assessment, there are several parameters that hit their bounds of estimation, including parameters associated with growth, survey selectivity in the first era, and the parameters associated with the BSFRF survey data (Figure 1).

The 'mixing' of parameters also needs to be considered. "Fast mixing" chains have parameter traces that reach a stationary distribution quickly, regardless of starting point. ADMB starts from the MLEs from fitting the model (unless you tell it otherwise), so one would hope that the parameters of the snow crab assessment would reach stationary distributions quickly. Most do, but some do not (Figure 2). Several of the parameters that have issues with hitting bounds also have slow mixing (growth parameters), but others associated with recruitment deviations and the initial numbers at length are also problematic.

Recommendations

Ultimately, the historically used methods did not account for parameter uncertainty well and somewhat arbitrarily defined the distribution of the OFL. Bayesian methodologies are designed to produce distributions of management quantities that incorporate parameter uncertainty and prior information. Still, as seen through the diagnostics presented here for the Bayesian interpretation of the 2016 snow crab assessment, some improvements should be made while pursuing Bayesian methods. These changes in methodology will likely result in larger changes to management quantities than the difference between ML and Bayesian methods (last year this difference was 3% of the OFL). Consequently, the remainder of this document will be directed at addressing questions related to the poorly behaved parameters and the processes they determine–primarily selectivity and growth.

Survey Selectivity

History, importance, and expectations of survey selectivity

Survey selectivity has been broken into three 'eras' in which separate selectivities and catchability coefficients have been estimated since at least 2005. The first era begins in 1978 and ends in 1981, the second era begins in 1982 and ends in 1988, and the third era begins in 1989 and runs until the present year (Figure 3). The stated rationale behind these divisions are a change in survey gear after 1981 and the survey began sampling stations farther north beginning in 1989 than previous years (in which it only reached to 61.2 N). The estimation of the catchability coefficient (often referred to as 'q') in each era is particularly important because it determines the scale of the population as seen through the survey index (all other parameters being equal). The value of the catchability coefficient ranges from 0 to 1 in this assessment (though values of q greater than 1 might be plausible if, for example, individuals of the surveyed population are somehow attracted to or herded into the survey gear) and represents the probability of capture for fully selected individuals. Shifting catchability up or down impacts how much the catch is inferred to affect the survey biomass (Figure 3). For example, in a system where q equals 1, the catch would appear to have a much larger influence on changes in the survey index than a system in which q equals 0.3.

The estimated catchability coefficients for each era in the snow crab assessment have varied over time (Figure 4). From the 2006 to 2010 assessments, q for all eras was estimated (or fixed) at values between 0.8 and 1. After the incorporation of the data from BSFRF experiments aimed at understanding the selectivity of NMFS survey gear (more on this below) in the 2011 assessment, catchability coefficients for era 2 and era 3 decreased for both males and females to somewhere between 0.6 and 0.8. The relationship between the estimated qs in era 2 and era 3 have been relatively consistent since 2010–era 2 has a larger q than era 3. This difference presumably allows the model to fit the large survey biomasses in the beginning of era 3 better. Interestingly, this relationship was reversed prior to the 2010 assessment. Catchability in the first era for males remained pinned at one (on its upper bound of estimation) for all of the historical assessments considered here. An estimated catchability coefficient of 1 is of concern because parameters on their bounds can bias the estimation of the variance covariance matrix (which can be an issue when developing distributions of derived quantities of the OFL through either maximum likelihood approaches or Bayesian approaches). It is also difficult to understand why catchability would be one in the first era, but not in subsequent eras.

Given these changes in catchability over time and its importance in determining quantities used in management, it may be useful to ask what the expected relationships between estimated catchability coefficients in different eras should be. One important factor to be considered is the number of stations sampled in a given year (e.g. Figure 5, Figure 6, and Figure 7) . In addition to a change in survey gear, the number of stations covered during the years 1978-1981 was much smaller than later years. Coverage among the years 1983-1986 was similar and then was expanded in 1987 and 1988 to the currently surveyed stations. Considering that a smaller area was surveyed in the first era, one might expect that catchability should be smaller in era 1 than in era 2, even if the change in gear was to a less efficient gear type. However, a scenario providing the opposite expectation may also make sense. The survey biomass is calculated by finding an average density and then scaling it to the entire survey area (which is constant regardless of the area actually surveyed). So, if the average density was higher in era 1 than in era 2 even though a smaller area was surveyed in era 1, one might expect that estimated catchability in era 1 to be higher. This said, we do not have to rely only on theorizing about what one might expect for catchability in each era; there are several data sources that provide information.

The Bering Sea Fisheries Research Foundation (BSFRF) conducted a survey of 108 tows in 27 survey stations in the Bering Sea in summer 2009. The BSFRF performed a similar study during 2010 in which the study area covered a larger portion of the distribution of snow crab than the 2009 study area. The mature biomass and size composition data gleaned from each of these experiments (and their complimentary NMFS survey observations; Figure 8 & Figure 9) are incorporated into the model by fitting them as an extra survey that is linked to the NMFS survey through a shared selectivity (see appendix A for a description of the way in which the surveys are related in the assessment model). Abundances estimated by the industry surveys were generally higher than the NMFS estimates, which provides evidence that the catchability of the NMFS survey gear is less than 1. Larger females are an exceptions to this observation, but this difference may be due to different towing locations for the two nets within the study area, or to variable catchability of females due to aggregation behavior. The BSFRF data can also be used to produce an empirical measure of selectivity and catchability, if the selectivity of the BSFRF is assumed to be 1 for all size classes (Figure 10) The ‘observed’ selectivity for males increases from <0.05 for the 27.5mm length bin for males to 0.58 at the 47.5mm length bin. Thereafter, selectivity decreases to approximately 0.30 until 112.5mm, after which it linearly increases to 1.0 at 132.5mm.

Model runs

Given uncertainty around appropriate catchability for era 1 and the availability of inferred selectivity for the most recent era, we undertook several model runs with different data and model configurations to better understand the behavior of survey selectivity (particularly q) in the assessment. First, we compared the accepted model from September 2016 (“Base”) to a model in which the data from 1978-1981 were removed (“Trim data”). Next, we fixed the survey in both era 2 and era 3 (era 1 now being eliminated) selectivity to the survey selectivity inferred from the BSFRF data (“Fixed obs sel”). Finally, we removed the BSFRF data to understand its influence on model outcomes (“No BSFRF”). We compared the fits to the data for each model and the resulting quantities used in management (e.g. reference biomasses and fishing mortalities, MMB, and the OFL).

Results

Fits to the data are difficult to compare among the scenarios by the value of the objective function (or contribution of individual likelihood components to the objective function Table 3) because the data changed for each scenario. However, visually, the fits for the models are similar, save “Fixed obs sel”. The terminal year of survey mature male biomass is considerably higher for “Fixed obs sel” and molt increment at lower size classes (up to around 60mm) is higher.

The “Base”" model estimated q in era 1 at 0.999 (0.999 for females), era 2 at 0.49 (0.32 for females), and era 3 at 0.62 (0.49 for females). After removing the first 4 years of data (model “Trim data”), q was 0.43 in era 2 for males (0.46 for females) and shifted to 0.67 in era 3 for males (0.54 for females) (Table 3). An upward shift occurred in q from era 2 to era 3 in both models (and for both sexes, but more so for males) that allows the higher survey estimates of MMB in the early 1990s to be fit better (Figure 11). The estimated qs for the “Trim data” and “Base” resulted in estimated selectivity that was higher than the inferred selectivity from the

BSFRF data (Figure 12). Removing the BSFRF data from the assessment ("No BSFRF") did not influence estimated q for males in era 2 greatly (0.47 vs 0.43), but it increased estimate q in era 3 to 0.79 from 0.67.

Removing the first four years of data resulted in a decrease in the OFL by 19% compared to "Base"", primarily because of a decrease in F35% and estimated MMB (Table 1). Fixing selectivity in all eras to the 'observed' survey selectivity increased the OFL by 236% compared to the 'Base' case, due to large changes in MMB, B35%, and F35%. These changes occurred primarily because of large differences in catchability and the associated change in estimated biomass, and changes in the probability of maturing (Figure 13). Eliminating the BSFRF data resulted in a decrease in the OFL of 53% compared to"Base", due to changes in F35% and MMB. A low estimated MMB was primarily related to an increase in estimated catchability.

Recommendations

We recommend dropping data from the years 1978-1981 given the uncertainty around gear selectivity during era 1, the models proclivity to estimate the catchability coefficient on its bound of 1, and the potential knock-on effects of this (both in estimating other qs and the Hessian and its role in MCMC). We do not recommend fitting directly to the BSFRF inferred selectivity given the very drastic changes in management quantities, but more consideration on the best way to use the BSFRF data is likely advisable given the potential effects hitting bounds on the estimation of parameters. Although it appeared that inclusion of the BSFRF data had some effect on estimates and derived quantities, the effect did not match the effect implied by the inferred selectivity. Depending on how the BSFRF data is treated, it may be prudent to revisit the designation of the timing of eras given the survey coverage, or explore time-varying selectivity. It was also clear that the probability of maturing is sensitive to assumptions about selectivity, so experiments designed to collect data to inform maturity may be useful to anchor this process. Each of the models presented after this section used "Trim data" as their starting point.

Natural mortality

Estimation and background

Natural mortality for snow crab in the Bering Sea is poorly known, due to relatively few targeted studies. In one of these studies, Nevissi, et al. (1995) used radiometric techniques to estimate shell age from last molt. The total sample size was 21 male crabs (a combination of Tanner and snow crab) from a collection of 105 male crabs from various hauls in the 1992 and 1993 NMFS Bering Sea survey. Representative samples for the 5 shell condition categories were collected that made up the 105 samples. The oldest looking crab within shell conditions 4 and 5 were selected from the total sample of SC4 and SC5 crabs to radiometrically age (Orensanz, Univ. of Washington, pers comm.). Shell condition 5 crab (SC5 = very, very old shell) had a maximum age of 6.85 years (s.d. 0.58, 95% CI approximately 5.69 to 8.01 years). The average age of 6 crabs with SC4 (very old shell) and SC5, was 4.95 years (range: 2.70 to 6.85 years). Given the small sample size, this maximum age may not represent the 1.5% percentile of the population that is approximately equivalent to Hoenig's method (1983). Maximum life span defined for a virgin stock is reasonably expected to be longer than these observed maximum ages from exploited populations. Particularly because fishing mortality was high before and during the time period during which this study was performed. Radiometric ages estimated by Nevissi, et al. (1995) may also be underestimated by several years, due to the continued exchange of material in crab shells even after shells have hardened (Craig Kastelle, pers. comm., Alaska Fisheries Science Center, Seattle, WA).

Tag recovery evidence from eastern Canada revealed observed maximum ages in exploited populations of 17-19 years (Nevissi, et al. 1995, Sainte-Marie 2002). A maximum time at large of 11 years for tag returns of terminally molted mature male snow crab in the North Atlantic has been recorded since tagging started about 1993 (Fonseca, et al. 2008). Fonseca, et al. (2008) estimated a maximum age of 7.8 years post terminal molt using data on dactal wear.

The mean for the prior for natural mortality used in this assessment is based on the assumption (informed by the studies above) that longevity would be at least 20 years in a virgin population of snow crab. Under negative exponential depletion, the 99th percentile corresponding to age 20 of an unexploited population corresponds to a natural mortality rate of 0.23. Using Hoenig's (1983) method a natural mortality equal to 0.23 corresponds to a maximum age of 18 years. Consequently, natural mortality for mature females was set to 0.23 yr^{-1}. Mature male natural mortality was estimated in the model with a prior constraint of mean of 0.23 yr^{-1} with a standard error equal to 0.054 (estimated from using the 95% CI of +-1.7 years on maximum age estimates from dactal wear and tag return analysis in Fonseca, et al. (2008)). Natural mortality for immature males and females was estimated in the 2016 assessment with a mean of 0.23 yr^{-1} and a standard error of 0.15.

Model runs

In the 2016 assessment, a multiplier for mature females was not estimated, although multipliers for mature males and immature females and males (together) were estimated. This omission resulted in estimates of mature female natural mortality being less than mature males, which is not intuitive biologically. Consequently, a model in which a multiplier for mature females was estimated ("Est female M") is presented here. Additionally, the prior for the multiplier on immature natural mortality for immature individuals was changed to 0.054 to match the assumptions for mature individuals (it hit its bound [2] if the prior was not more constrained).

Results

Nearly all likelihood components for the scenarios "Trim data" and "Est female M" can be directly compared because the data and weighting did not change (Table 4). The only exception is the M multiplier prior because the prior for the immature natural mortality was changed from 0.15 to 0.054 to match the prior for mature individuals. This change should result in a higher contribution of this likelihood component (and it did 16.74 vs. 79.34 for "Trim data" and "Est female M", respectively; Table 4). Even with this additional weighting in this component, the objective function value decreased by ~80 units for "Est female M". In particular, fits to female growth (Figure 14), survey lengths (Figure 15 & Figure 16), trawl lengths (Figure 17), and survey biomass (Figure 11) showed improvements, though they are sometimes difficult to see by eye. In "Est female M", natural mortality for mature females was estimated as 0.32^{-yr}; natural mortality for mature males was 0.26^{-yr} (Table 3). Estimated immature natural mortality was 0.29^{-yr}. The estimated OFL for "Est female M" was 7% less than "Trim data", which was the result of a lower estimated MMB and a smaller F35% (Table 1). Lower estimated MMB was a result of an increase in survey catchability for "Est female M" and slight changes in the probability of maturing (Figure 12 & Figure 13).

Discussion and recommendations

We recommend estimating natural mortality for mature females given the improvements in model fit and now biologically consistent relationship between mature male and mature female natural mortality. However, the prior on immature natural mortality may warrant revisiting and the longevity rationale for M should be reexamined under the newly estimated Ms.

Growth

Estimation and available data

Growth in the "Base" model is estimated as a piece-wise linear model in which two lines are fit to the data with an estimated break point determining where in the data the switch from one linear model to the next occurs. Unfortunately, data in the area in which the breakpoint occurs are few. Data to inform growth for snow crab in the Bering Sea are few. Tagging experiments were conducted on snow crab in 1980 with recoveries occurring in the Tanner crab (Chionoecetes bairdi) fishery in 1980 to 1982 (McBride 1982). However, data from this study are not used due to uncertainty about the effect of tagging on growth. Currently, 40 data points from 5 studies are used to estimate the post-molt length from pre-molt length for females and males (Figure 14). The studies include:

1. Transit study (Rugolo unpublished data, 2003); 14 crab
2. Cooperative seasonality study (Rugolo); 6 crab
3. Dutch harbor holding study; 9 crab
4. NMFS Kodiak holding study held less than 30 days; 6 crab
5. NMFS Kodiak holding study 2016; 5 crab

Data from the NMFS Kodiak holding study 2016 were new for the 2016 assessment. In the “Transit study”, pre- and post-molt measurements of 14 male crabs that molted soon after being captured were collected. The crabs were measured when shells were still soft because all died after molting, so measurements may be underestimates of postmolt width (Rugolo, pers. com.). The holding studies include only data for crab held less than 30 days because growth of crabs held until the next spring’s molting was much lower. Females molting to maturity were excluded from all data sets, since the molt increment is usually smaller. Crab missing more than two limbs were excluded due to other studies showing lower growth. Crab from Rugolo’s seasonal study were excluded that were measured less than 3 days after molting due to difficulty in measuring soft crab accurately. In general, growth of snow crab in the Bering Sea appears to be greater than growth of some North Atlantic snow crab stocks (Sainte-Marie 1995).

MCMC diagnostics (Figure 1) showed that parameters associated with growth are poorly behaved and this was also known from the difficulties with local minima in the likelihood surface explored by Turnock (2016). The response to the poor behavior of the growth parameters in the past was to perform a lengthy jittering process (in which the model was fit to the data hundreds of times with different starting points) to ensure that the estimates of parameters used in the projection model were from a run that returned the smallest negative log likelihood possible. A potential issue with this approach is that it assumes that the value of the poorly behaved growth parameters are ‘known’ in the projections to calculate the OFL, when there are actually no data to determine their value. Furthermore, the location of the breakpoint is such that the lower end of the breakpoint function for growth influences at most two length bins in the model (Figure 14). The fits of the breakpoint models to the range of lengths that are included in the assessment were also were also curious, with strange residual patterns (Figure 14).

Model runs

A model in which only a linear function for male and female growth was estimated (“Chop growth”) is presented for comparison. Simplifying the growth model seemed an appropriate consideration, given the difficulties with local minima and parameter behavior, the potentially small influence of a breakpoint model on length bins included in the model, and the curious fits to the growth data of the breakpoint model. Observations of growth increment for pre-molt lengths less than the smallest length bin in the model were eliminated before fitting the linear models to the data. A model in which natural mortality for mature females was estimated in addition to simplifying the growth function (“Growth + M”) is also presented. Both models are based off of “Trim data” in that the data for years 1978-1981 were removed.

Results

The overall objective function value for “Trim data” and “Chop growth” cannot be directly compared because the data changed between scenarios. However, given that there was less data for ‘Chop growth’, one would expect that all other things remaining the same, its objective function would decrease relative to “Trim data”. It did not (Table 4). Fits to male growth (Figure 14) and survey lengths (Figure 15 & Figure 16) showed the largest declines in fit. Allowing natural mortality for mature females to be estimated did not markedly improve overall fit (Table 4). The estimated OFL for "Chop growth" was 7% lower than "Trim data", resulting from decreases in estimated MMB, F35% and B35% (Table 1). Declines in MMB likely stem from smaller growth increments at lower sizes (up to about 55mm; Figure 14) and changes in catchability (Figure 12).

Discussion and recommendations

The sharp declines in fit as seen through the total likelihood suggest moving to a single linear model is a step backward (though the growth models fits were not compared via an information criterion that incorporates the cost of new parameters). However, the stability of the growth model in "Trim data" is poor. A potential fix for this is removing length bins from the small sizes until a single linear model is appropriate. Another potential fix is predicting growth increment instead of post-molt size and use a Beverton Holt shaped curve. This would cut the number of parameters used drastically and hopefully add stability to the model.

Data gaps and research priorities

The below notes and recommendations are the subset of issues included in the September 2016 SAFE that were not addressed by the above analyses. They are repeated here for discussion.

Data sources

As many raw data sources as possible should be included in the assessment. Estimating parameters outside of the model and inputting them as 'known' artificially decreases the uncertainty represented in the posteriors (and estimated standard errors) of management quantities. Weight at length data, data used to develop priors for natural mortality and maturity, and the selectivities calculated from the BSFRF data should be considered for inclusion in the model to comprehensively represent the uncertainty in management quantities. In addition to pulling as much data into the model as possible, standardizing and automating the creation of data files from the survey and catch databases would be very useful given the short time frame of the assessment cycle.

Modeling

When incorporating weight at length data into the assessment, it may be useful to consider a split in parameters for mature and immature males as is done for females. When splitting the mature males in the first year, it is assumed that they are all new shell, but the females are split out between new and old shell condition. Finally, considering the impact of basing natural mortality off of longevity and then splitting it into immature and mature M on the calculation of reference points may improve the interpretability of estimates of natural mortality.

Weighting

Different weighting of likelihood components can have drastic impacts on the management advice provided from an assessment. A close look at the way CVs, sample sizes, and other weighting factors are calculated and their influence on assessment results could provide better understanding of how well the model is balanced. Standardization of the weighting schemes would also improve readability of the code (for example, some size composition data have both 'weights' and 'sample sizes').

Scientific uncertainty

Natural mortality exerts a large influence over estimated management quantities, but is poorly known. Tagging studies targeted at estimating natural mortality could be very useful and could also shed light on the migration patterns, which could help us understand the impact of the fishery (e.g. centroids of large male abundance in the survey and catch do not match–is this because the crab are moving or because the fishery operates in a specific place? The answer to this question could influence priors on catchability.) Similarly, establishing measures of reproductive capacity that include females, the spatial overlap of mature individuals, the role water temperature plays in biennial spawning, and the effectiveness of mating by size for males may allow for relationships between recruitment and mature biomass to be found. In general, exploring the spatial dynamics of the population may allow for patterns and influences of the fishery and environment on the productivity of the stock to be more easily identified. Preliminary analyses suggest that retrospective biases may be a problem for the snow crab assessment. Retrospective biases can result from unaccounted for time-varying processes in the population dynamics of the model (Hurtado et al., 2015). Focused research on the potential for retrospective biases in the snow crab assessment should be pursued.

Literature cited

Chilton, E.A., C.E. Armisted and R.J. Foy. 2009. Report to industry on the 2009 Eastern Bering Sea crab survey. AFSC Processed Report 2009-XX.

Clark, W.G. 1991. Groundfish exploitation rates based on life history parameeters. Can. J. fish. Aquat. Sci. 48: 734-750.

Dawe, E.G., D.M. Taylor, J.M. Hoenig, W.G. Warren, and G.P. Ennis. 1991. A critical look at the idea of terminal molt in male snow crab (Chionoecetes opilio). Can. J. Fish. Aquat. Sci. 48: 2266-2275.

Ernst, B, J.M.(Lobo) Orensanz and D.A. Armstrong. 2005. Spatial dynamics of female snow crab (Chionoecietes opilio) in the eastern Bering Sea. Can. J. Fish. Aquat. Sci. 62: 250–268.

Fonseca, D. B., B. Sainte-Marie, and F. Hazel. 2008. Longevity and change in shell condition of adult male snow crab Chionoecetes opilio inferred from dactyl wear and mark-recapture data. Transactions of the American Fisheries Society 137:1029–1043.

Fournier, D.A. and C.P. Archibald. 1982. A general theory for analyzing catch-at-age data. Can.J.Fish.Aquat.Sci. 39:1195-1207.

Greiwank, A. and G.F. Corliss(eds). 1991. Automatic differentiation of algorithms: theory, implementation and application. Proceedings of the SIAM Workshop on the Automatic Differentiation of Algorithms, held Jan. 6-8, Breckenridge, CO. Soc. Indust. And Applied Mathematics, Philadelphia.

Hoenig, J. 1983. Empirical use of longevity data to estimate mortality rates. Fish. Bull. 82: 898-903.

Mcbride (1982). Tanner crab tag development and tagging experiments 1978-1982. In Proceedings of the International Symposium of the Genus Chionoecetes. Lowell Wakefield Fish. Symp. Ser., Alaska Sea Grant Rep. 82-10. University of Alaska, Fairbanks, Alaska. Pp. 383-403.

Methot, R. D. 1990. Synthesis model: An adaptable framework for analysis of diverse stock assessment data. Int. N. Pac. Fish. Comm. Bull. 50:259-277.

Murphy, J. T., A. B. Hollowed, J. J. Anderson. 2010. Snow crab spatial distributions: examination of density –dependent and independent processes. Pp. 49-79. In G. Kruse, G. Eckert, R. Foy, G. Kruse, R. Lipcius, B. St. Marie, D. Stram, D. Woodby (Eds.), Biology and management of Exploited Crab Populations Under Climate Change. Alaska Sea Grant Program Report AK-SG-10-01, University of Alaska Fairbanks, AK. Doi:10.4027/bmecppc.2010.19

Nevissi, A.E., J.M. Orensanz, A.J.Paul, and D.A. Armstrong. 1995. Radiometric Estimation of shell age in Tanner Crab, Chionoecetes opilio and C. bairdi, from the eastern Bering Sea, and its use to interpret indices of shell age/condition. Presented at the International symposium on biology, management and economics of crabs from high latitude habitats October 11-13, 1995, Anchorage, Alaska.

NPFMC (North Pacific Fishery Management Council). 2007. Environmental Assesment for Amendment 24. Overfishing definitions for Bering Sea and Aluetian Islands King and Tanner crab stocks. North Pacific Fishery Management Council,Anchorage, AK, USA..

NPFMC (North Pacific Fishery Management Council). 2000. Bering Sea snow crab rebuilding plan. Amendment 14. Bering Sea Crab Plan Team, North Pacific Fishery Management Council,Anchorage, AK, USA..

NPFMC 1998. Bering Sea and Aluetian Islands Crab FMP. Bering Sea Crab Plan Team, North Pacific Fishery Management Council, P. O. Box 103136, Anchorage, Ak 99510.

Orensanz, J.M., J. Armstrong, D. Armstrong and R. Hilborn. 1998. Crustacean resources are vulnerable to serial depletion – the multifaceted decline of crab and shrimp fisheries in the Greater Gulf of Alaska. Reviews in Fish Biology and Fisheries 8:117-176.

Otto, R.S. 1998. Assessment of the eastern Bering Sea snow crab, Chionoecetes opilio, stock under the terminal molting hypothesis. In Proceedings of the North Pacific Symposium on Invertebrate Stock Assessment and Management. Edited by G.S. Jamieson and A. Campbell. Can. Spec. Publ. Fish. Aquat. Sci. 125. pp. 109-124.

Paul, A.J., J.M. Paul and W.E. Donaldson. 1995. Shell condition and breeding success in Tanner crabs. Journal of Crustacean Biology 15: 476-480.

Rugolo, L.J., D. Pengilly, R. MacIntosh and K. Gravel. 2005. Reproductive dynamics and life-history of snow crab (Chionoecetes opilio) in the eastern Bering Sea. Final Completion Report to the NOAA, Award NA17FW1274, Bering Sea Snow Crab Fishery Restoration Research.

Rugolo, L.J., R.A. MacIntosh, C.E. Armisted, J.A. Haaga and R.S. Otto. 2003. Report to industry on the 2003 Eastern Bering Sea crab survey. AFSC Processed Report 2003-11.

Rodionov, S. 2004. A sequantial algorithm for testing climate regime shifts. Geophysical Research Letters 21: L09204.

Sainte-Marie, B., Raymond, S., and Brethes, J. 1995. Growth and maturation of the male snow crab, Chionoecetes opilio (Brachyura: Majidae). Can.J.Fish.Aquat.Sci. 52:903-924.

Sainte-Marie, B., J. Sevigny and M. Carpentier. 2002. Interannual variability of sperm reserves and fecundity of primiparous females of the snow crab (Chionoecetes opilio) in relation to sex ratio. Can.J.Fish.Aquat.Sci. 59:1932-1940.

Szuwalski, C.S. and Punt, A.E. 2013. Regime shifts and recruitment dynamics of snow crab, Chionoecetes opilio, in the eastern Bering Sea. Fisheries Oceanography, 22: 345-354.

Szuwalski, C.S. and Punt, A.E. 2012. Fisheries management for regime-based ecosystems: a management strategy evaluation for the snow crab fishery in the eastern Bering Sea. ICES Journal of Marine Science. 70: 955-967.

Tamone, S.L., M. Adams and J.M. Dutton. 2005. Effect of eyestalk ablation on circulating ecdysteroids in hemolymph of snow crab Chionoecetes opilio: physiological evidence for a terminal molt. Integr. Comp. Biol., 45(120), p.166-171.

Zheng, J., S. Siddeek, D. Pengilly, and D. Woodby. 2002. Overview of recommended harvest strategy for snow crabs in the Eastern Bering Sea. Regional Information Report No. 5J02-03. Alaska Department of Fish and Game. Juneau, Alaska.

Zheng, J., G.H. Kruse, and D.R. Ackley. 2001. Spatial distribution and recruitment patterns of snow crabs in the eastern Bering Sea. Spatial Processes and management of marine populations. Alaska sea grant college program. AK-SG-01-02, 2001.

Appendix A: Model structure

Population dynamics

Numbers of sex s of shell condition v and maturity state m at length l in the initial year of the assessment, N$_{s,v,m,y=1,l}$, were calculated from an estimated vector of numbers at length l by sex s and maturity state m for males, $\lambda_{s,m,l}$ and numbers at length l by sex s and shell condition v for females (i.e. 2 vectors for each sex were estimated). Estimated vectors of initial numbers at length by maturity for females were calculated by splitting the estimated vectors at length by the observed proportion mature in the first year of the survey.

$$N_{s,v,m,y=1,l} = \begin{cases} \Omega^{obs}_{s,l}\lambda_{s,1,l} & \text{if v = new; m = mat, s = fem} \\ 1-\Omega^{obs}_{s,l}\lambda_{s,1,l} & \text{if v = new; m = imat, s = fem} \\ \lambda_{s,2,l} & \text{if v = old; m = mat, s = fem} \\ 0 & \text{if v = old; m = imat} \end{cases} \tag{1}$$

Initial numbers at length for males were all assumed to be new shell.

$$N_{s,v,m,y=1,l} = \begin{cases} \lambda_{s,1,l} & \text{if v = new; m = mat, s = male} \\ \lambda_{s,2,l} & \text{if v = new; m = imat, s = male} \\ 0 & \text{if v = old; m = mat, s = male} \\ 0 & \text{if v = old; m = imat, s = male} \end{cases} \tag{2}$$

The dynamics after the initial year were described by:

$$N_{s,v,m,y+1,l} = \begin{cases} \Omega_{s,l}\kappa_{s,l'}Q_{s,imat,y,l'}X_{s,l',l} & \text{if v = new; m = mat} \\ 1-\Omega_{s,l}\kappa_{s,l'}Q_{s,imat,y,l'}X_{s,l',l} + Rec^{\epsilon}_{y}Pr_{l} & \text{if v = new; m = imat} \\ Q_{s,mat,y,l'} & \text{if v = old; m = mat} \\ (1-\kappa_{s,l'})Q_{s,imat,y,l'} & \text{if v = old; m = imat} \end{cases} \tag{3}$$

Where $\Omega_{s,l}$ was the probability of maturing at length l for sex s (a freely estimated vector for both males and females constrained by penalties on smoothness and a prior in some scenarios), $\kappa_{s,l'}$ was the probability of molting for an immature crab of sex s at length *l'* (set to 1 for all immature crab), and X$_{s,l,l'}$ was the size transition matrix describing the probability of transitioning from size *l'* to size l for sex s. Q$_{s,m,y,l'}$ was the number of crab of sex s, maturity state m, and length *l'* surviving natural and fishing mortality during year y:

$$Q_{s,m,y,l} = \sum_{v} N_{s,v,m,y,l}e^{Z_{s,v,m,y,l}} \tag{4}$$

Where N$_{s,v,m,y,l}$ represented the numbers, *N*, of sex s during year y of shell condition v and maturity state m at length l. Z$_{x,v,m,y,l}$ represented the total mortality experienced by the population and consisted of the sum of instantaneous rates of natural mortality by sex and maturity state, M$_{s,m}$, and fishing mortality, F$_{s,f,y,l}$ from each fishery. Each fishing mortality was subject to selectivity by length l, which varied between sexes s and fisheries f (and by year y if specified) . M$_{s,m}$ was specified in the model and a multiplier $\gamma_{natM,m}$

was estimated subject to constraints (see Table 2; this formulation effectively specified a mean and standard deviation for a prior distribution for M).

$$Z_{s,v,m,y,l} = \gamma_{natM,m} M_{s,m} + \sum_{f} S_{s,f,y,l} F_{s,f,y,l} \tag{5}$$

Selectivities in the directed and bycatch fisheries were estimated logistic functions of size. Different selectivity parameters were estimated for females and males in the directed fisheries ($S_{fem,dir,l}$ and $S_{male,dir,l}$, respectively), a single selectivity for both sexes was estimated for bycatch in the groundfish trawl fishery ($S_{trawl,l}$), and a retention selectivity was estimated for the directed fishery for males ($R_{dir,l}$; all females were discarded).

$$S_{male,dir,l} = \frac{1}{1+e^{-S_{slope,m,d}(L_l - S_{50,m,d}}}) \tag{6}$$

$$S_{fem,dir,l} = \frac{1}{1+e^{-S_{slope,f,d}(L_l - S_{50,f,d}}}) \tag{7}$$

$$S_{trawl,l} = \frac{1}{1+e^{-S_{slope,t}(L_l - S_{50,t}}}) \tag{8}$$

$$R_{dir,l} = \frac{1}{1+e^{-S_{slope,m,d}(L_l - S_{50,m,d}}}) \tag{9}$$

Where $S_{slope,s,f}$ was the slope of the logistic curve for sex *s* in fishery *f* and $S_{50,s,f}$ was the length at 50% selection for sex *s* in fishery *f*. Catches for all fisheries were modeled as pulse fisheries in which all catch was removed instantaneously (i.e. no natural mortality occurred during the fishery). Catch in fishery *f* during year *y* was calculated as the fraction of the total fishing mortality, $F_{s,f,y,l}$, applied to a given sex *s* in a fishery *f* times the biomass removed by all fisheries for that sex.

$$C_{male,dir,y} = \sum_{l}\sum_{v}\sum_{m} w_{male,l} \frac{R_l F_{male,dir,y,l}}{F_{male,dir,y,l+F_{trawl,y,l}}} N_{male,v,m,y,l} e^{-\delta_y M_{s,m}} (1 - e^{-(F_{male,dir,y,l}+F_{trawl,y,l})}) \tag{10}$$

$$C_{male,tot,y} = \sum_{l}\sum_{v}\sum_{m} w_{male,l} \frac{F_{male,dir,y,l}}{F_{male,dir,y,l+F_{trawl,y,l}}} N_{male,v,m,y,l} e^{-\delta_y M_{s,m}} (1 - e^{-(F_{male,dir,y,l}+F_{trawl,y,l})}) \tag{11}$$

$$C_{fem,dir,y} = \sum_{l}\sum_{v}\sum_{m} w_{fem,l} \frac{F_{fem,dir,y,l}}{F_{fem,dir,y,l+F_{trawl,y,l}}} N_{fem,v,m,y,l} e^{-\delta_y M_{s,m}} (1 - e^{-(F_{fem,dir,y,l}+F_{trawl,y,l})}) \tag{12}$$

$$C_{m+f,trawl,y} = \sum_{s}\sum_{l}\sum_{v}\sum_{m} w_{s,l} N_{s,v,m,y,l} e^{-\delta_y M_{s,m}} (1 - e^{-(F_{trawl,y,l})}) \tag{13}$$

Where δ_y was the mid point of the fishery (all fisheries were assumed to occur concurrently and the midpoint was based on the directed fishery, which accounts for the vast majority of the fishing mortality) and $w_{s,l}$ was the weight at length *l* for sex *s*. Trawl data and discard data were entered into the model with an assumed mortality of 80% and 30%, respectively. Fully-selected fishing mortality parameters for fishery *f* were estimated as a logged average over a given time period (F^{log}_{avg}) with yearly deviations around that mean ($F^{log}_{dev,y}$).

$$F_{f,y} = e^{(F^{log}_{avg,f} + F^{log}_{dev,f,y})} \tag{14}$$

Selectivity for the survey was estimated for 3 eras: 1978-1981, 1982-1988, and 1989-present. Selectivity was assumed to be logistic and separate parameters representing the length at which selection probability equal 50% and 95% ($s_{50,s,e}$ and $s_{95,s,e}$, respectively) were estimated for males and females in the third era (1989-present). Separate catchability coefficients ($q_{s,e}$) were estimated for males and females in all eras.

$$S_{surv,s,l,e} = \frac{q_{s,e}}{1 + e^{-log(19)\frac{L_l - s_{50,s,e}}{s_{95,s,e} - s_{50,s,e}}}}) \tag{15}$$

Survey selectivity was informed by experimental surveys during the years 2009 and 2010. A portion of the NMFS summer survey tows were accompanied by an industry vessel using nephrops trawls with an assumed selectivity of 1 for all size classes. To represent the proportion of the population covered by the experiment, a vector was freely estimated for males, S_y^{free} (subject to a scaling parameter), and a logistic curve was estimated for females.

$$S_{ind,s,l,y} = \begin{cases} \frac{q_{ind,s,y}}{1+e^{-log(19)\frac{L_l - s_{50,s,y}}{s_{95,s,y} - s_{50,s,y}}}}) & \text{if s = female} \\ q_{ind,s,y} S_y^{free} & \text{if s = male} \end{cases} \tag{16}$$

Based on this logic, after identifying the fraction of the crab at length covered by the experimental surveys, the length frequencies of the NMFS data collected simultaneously with the experimental trawls can be calculated by multiplying the numbers at length 'available' to the experimental trawls by the overall survey selectivity, $S_{surv,s,l,y}$. The predicted numbers at length for the NMFS and industry data from the selectivity experiment were calculated by multiplying the respective selectivities by the survey numbers at length.

$$S_{nmfs,s,l,y} = S_{ind,s,l,y} S_{surv,s,l,y} \tag{17}$$

Mature male and female biomass (MMB and FMB, respectively) were fitted in the objective function and were the product of mature numbers at length during year y and the weight at length, $w_{s,l}$:

$$MMB_y = \sum_{l,v} w_{male,l} N_{male,v,mat,y,l} \tag{18}$$

$$FMB_y = \sum_{l,v} w_{fem,l} N_{fem,v,mat,y,l} \tag{19}$$

$$w_{s,l} = \alpha_{wt,s} L_l^{\beta_{wt,s}} \tag{20}$$

Mature biomass can be calculated for different time through out the year, in which case the numbers at length are decremented by the estimated natural mortality. Parameters $\alpha_{wt,s}$ and $\beta_{wt,s}$ were estimated outside of the assessment model and specified in the control file.

Molting and growth occur before the survey. Immature crab were assumed to molt every year with an estimated probability of molting to maturity based on length l (in all the scenarios presented here, the probability of molting was 1 for all immature animals). For crab that do molt, the growth increment within the size-transition matrix, $X_{s,l,l'}$, was based on a piece-wise linear relationship between predicted pre- and post-molt length, ($\hat{L}_{s,l}^{pred}$ and $\hat{L}_{s,l}^{post}$, respectively) and the variability around that relationship was characterized by a discretized and renormalized gamma function, $Y_{s,l,l'}$.

$$X_{s,l,l'} = \frac{Y_{s,l,l'}}{\sum_{l'} Y_{s,l,l'}} \tag{21}$$

$$Y_{s,l,l'} = (\Delta_{l,l'})^{\frac{L\hat{}_{s,l} - (\bar{L}_l - 2.5)}{\beta_s}} \tag{22}$$

$$\hat{L}_{s,l}^{post,1} = \alpha_s + \beta_{s,1} L_l \tag{23}$$

$$\hat{L}^{post,2}_{s,l} = \alpha_s + \delta_s(\beta_{s,1} - \beta_{s,2}) + \beta_{s,2}L_l \tag{24}$$

$$\hat{L}^{post}_{s,l} = \hat{L}^{post,1}_{s,l}(1 - \Phi(\frac{L_l - \delta_{a,x}}{stgr})) + \hat{L}^{post,2}_{s,l}(\Phi(\frac{L_l - \delta_{a,x}}{stgr})) \tag{25}$$

$$\Delta_{l,l'} = \bar{L}_{l'} + 2.5 - L_l \tag{26}$$

$\hat{L}^{post,1}_{s,l}$ and $\hat{L}^{post,2}_{s,l}$ were predicted post-molt lengths from each piece of the piece-wise relationship, and $\Phi()$ was a cumulative normal distribution in which $\delta_{a,x}$ was an estimated change point. For "Growth chop" and "Growth + M", only $\hat{L}^{post,1}_{s,l}$ was used to predict post-molt lengths.

An average recruitment for the assessment period (1978-present) and yearly deviations around this average were estimated within the assessment. The sex ratio of recruitment was assumed to be 50/50 male to female. Each year's estimated recruitment was allocated to length bins based on a discretized and renormalized gamma function with parameters specified in the control file.

$$Rec_y = e^{(Rec_{avg} + Rec_{dev,y})} \tag{27}$$

$$Pr_l = \frac{(\Delta_{1,l})^{\alpha_{rec}/\beta_{rec}} e^{-\Delta_{1,l'}/\beta_{rec}}}{\sum_{l'}(\Delta_{1,l'})^{\alpha_{rec}/\beta_{rec}} e^{(-\Delta_{1,l'}/\beta_{rec})}} \tag{28}$$

Likelihood components

Three general types of likelihood components were used to fit to the available data (Table 5). Multinomial likelihoods were used for size composition data, log-normal likelihoods were used for indices of abundance data, and normal likelihoods were used for catch data, growth data, priors, and penalties. Multinomial likelihoods were implemented in the form:

$$L_x = \lambda_x \sum_y N^{eff}_{x,y} \sum_l p^{obs}_{x,y,l} ln(\hat{p}_{x,y,l}/p^{obs}_{x,y,l}) \tag{29}$$

L$_x$ was the likelihood associated with data component x, where λ_x represented an optional additional weighting factor for the likelihood, $N^{eff}_{x,y}$ was the effective sample sizes for the likelihood, $p^{obs}_{x,y,l}$ was the observed proportion in size bin l during year y for data component x, and $\hat{p}_{x,y,l}$ was the predicted proportion in size bin l during year y for data component x. 10 multinomial likelihood components were included in the assessment (see Table 5 for descriptions, weighting factors, and effective sample sizes). When likelihood components were removed from the analysis (e.g. "No BSFRF"), the weighting for that component was set to 0.

Iterative methods for determining appropriate effective samples sizes for composition data are suggested to avoid over-weighting the size composition data and washing out the signal from the indices of abundance. The method of implementation used here is discussed below.

Log normal likelihoods were implemented in the form:

$$L_x = \lambda_x \sum_y \frac{(ln(\hat{I}_{x,y}) - ln(I_{x,y}))^2}{2(ln(CV^2_{x,y} + 1))} \tag{30}$$

L_x was the contribution to the objective function of data component x, λ_x was any additional weighting applied to the component, $\hat{I}_{x,y}$ was the predicted value of quantity I from data component x during year y,

$\mathrm{I}_{x,y}$ was the observed value of quantity I from data component x during year y and $\mathrm{CV}_{x,y}$ was the coefficient of variation for data component x during year y. 5 log normal likelihood components were included in this assessment (see Table 5 for descriptions, weighting factors, and CVs).

Normal likelihoods were implemented in the form:

$$L_x = \lambda_x \sum_y (\hat{I}_{x,y} - I_{x,y})^2 \tag{31}$$

L_x was the contribution to the objective function of data component x, λ_x was represents the weight applied to the data component (and can be translated to a standard deviation), $\hat{I}_{x,y}$ was the predicted value of quantity I from data component x during year y, $\mathrm{I}_{x,y}$ was the observed value of quantity I from data component x during year y. 12 normal likelihood components were included in the "Base" assessment (see Table 5 for descriptions, weighting factors, and translated standard deviations).

Smoothing penalties were also placed on some estimated vectors of parameters in the form of normal likelihoods on the second differences of the vector. Code for this assessment can be found on github.com/szuwalski/SnowCrab.

Table 1: Changes in management quantities for each scenario considered. Reported quantites are the MLEs because running MCMC for every model was prohibitively time-consuming. The MLEs for scenarios in which MCMCs were performed are very close to the medians of the posterior distributions.

Model	MMB	B35	F35	FOFL	OFL
Base	92.09	152.3	1.91	1.14	24.59
Trim data	83.8	152.3	1.42	0.81	19.97
Fixed obs sel	221.6	215.3	3.49	2.62	82.68
No BSFRF	60.86	142.3	1.17	0.56	11.5
Est female M	74.29	139.6	1.21	0.68	18.38
Chop growth	79.57	149.8	1.34	0.75	18.66
Growth + M	70.89	137.4	1.17	0.64	17.35

Table 2: Parameter bounds and symbols

Parameter	Lower	Upper	Symbol
af	-100	0	α_f
am	-50	0	α_m
bf	1	10	$\beta_{f,1}$
bm	1	5	$\beta_{m,1}$
b1	1	1.5	$\beta_{f,2}$
bf1	1	2	$\beta_{m,2}$
deltam	10	50	δ_m
deltaf	5	50	δ_f
st_gr	0.5	0.5	stgr
growth_beta	0.749	0.751	β_g
mateste	-6	-1e-10	$\Omega_{m,l}$
matestfe	-6	-1e-10	$\Omega_{f,l}$
mean_log_rec	"-inf"	Inf	Rec_{avg}
rec_devf	-15	15	$Rec_{f,dev,y}$
alpha1_rec	11.49	11.51	α_{rec}
beta_rec	3.99	4.01	β_{rec}
mnatlen_styr	-3	15	$\lambda_{male,v,l}$
fnatlen_styr	-10	15	$\lambda_{fem,v,l}$
log_avg_fmort	"-inf"	Inf	$F^{log}_{avg,dir}$
fmort_dev	-5	5	$F^{log}_{dev,dir,y}$
log_avg_fmortdf	-8	-1e-04	$F^{log}_{avg,disc}$
fmortdf_dev	-15	15	$F^{log}_{dev,disc,y}$
log_avg_fmortt	-8	-1e-04	$F^{log}_{avg,trawl}$
fmortt_dev_era1	-15	15	$F^{log}_{dev,trawl,era1}$
fmortt_dev_era2	-15	15	$F^{log}_{dev,trawl,era2}$
log_avg_sel50_mn	4	5	$S_{50,new,dir}$
log_avg_sel50_mo	4	5	$S_{50,old,dir}$
fish_slope_mn	0.1	0.5	$S_{slope,m,d}$
fish_fit_slope_mn	0.05	0.5	$S_{slope,m,d}$
fish_fit_sel50_mn	85	120	$S_{50,old,dir}$
fish_slope_mo2	1.9	2	$S_{slope,m,d}$
fish_sel50_mo2	159	160	$S_{50,old,dir}$
fish_slope_mn2	0.01	2	$S_{slope,m,d}$
fish_sel50_mn2	100	160	$S_{50,old,dir}$
fish_disc_slope_f	0.1	0.7	$S_{slope,m,d}$
fish_disc_sel50_f	1	5	$S_{50,old,dir}$
fish_disc_slope_tf	0.01	0.3	$S_{slope,trawl}$
fish_disc_sel50_tf	30	120	$S_{50,trawl}$
srv1_q	0.2	1	$q_{m,era1,surv}$
srv1_q_f	0.2	1	$q_{f,era1,surv}$
srv1_sel95	30	150	$S_{95,era1,surv}$
srv1_sel50	0	150	$S_{50,era1,surv}$
srv2_q	0.2	1	$q_{m,era2,surv}$
srv2_q_f	0.2	1	$q_{f,era2,surv}$
srv2_sel95	50	160	$S_{95,era2,surv}$
srv2_sel50	0	80	$S_{50,era2,surv}$
srv3_q	0.2	1	$q_{m,era3,surv}$
srv3_sel95	40	200	$S_{95,m,era2,surv}$
srv3_sel50	25	90	$S_{50,m,era2,surv}$

Parameter	Lower	Upper	Symbol
srv3_q_f	0.2	1	$q_{f,era3,surv}$
srv3_sel95_f	40	150	$S_{95,f,era2,surv}$
srv3_sel50_f	0	90	$S_{50,f,era2,surv}$
srvind_q	0.1	1	$q_{m,09,ind}$
srvind_q_f	0.01	1	$q_{f,09,ind}$
srvind_sel95_f	55	120	$S_{95,f,09,ind}$
srvind_sel50_f	-50	55	$S_{50,f,09,ind}$
srv10in_q	0.1	1	$q_{m,10,ind}$
srv10ind_q_f	0.01	1	$q_{f,10,ind}$
selsmo10ind	-4	-0.001	SelVecMaleInd09
selsmo09ind	-4	-0.001	SelVecMaleInd10
Mmult_imat	0.2	2	$\gamma_{natM,imm}$
Mmult	0.2	2	$\gamma_{natM,mat,m}$
Mmultf	0.2	2	$\gamma_{natM,mat,f}$
cpueq	0.0000877	0.00877	q_{cpue}

Table 3: Estimated parameter values by scenario (these are maximum likelihood estimates)

Parameter	Base	Trim data	Fixed obs sel	No BSFRF	Est female M	Chop growth	Growth + M
af	-5.08	-5.06	-4.1	-5.08	-5.29	9.72	9.87
am	-5.74	-5.83	-7.48	-12.2	-5.61	4.89	4.84
bf	1.53	1.52	1.48	1.53	1.53	1.09	1.09
bm	1.54	1.54	1.62	1.83	1.53	1.19	1.19
b1	1.15	1.15	1.12	1.16	1.15	1.25	1.25
bf1	1.02	1.03	1	1.03	1.04	1.5	1.5
deltam	32.2	32.25	32.37	27.47	32.19	30	30
deltaf	34.37	34.29	36.51	34.33	34.13	27.5	27.5
st_gr	0.5	0.5	0.5	0.5	0.5	0.5	0.5
growth_beta	vector	vector	vector	vector	vector	vector	vector
mateste	vector	vector	vector	vector	vector	vector	vector
matestfe	vector	vector	vector	vector	vector	vector	vector
rec_devf	vector	vector	vector	vector	vector	vector	vector
alpha1_rec	11.5	11.5	11.5	11.5	11.5	11.5	11.5
beta_rec	4	4	4	4	4	4	4
mnatlen_styr	vector	vector	vector	vector	vector	vector	vector
fnatlen_styr	vector	vector	vector	vector	vector	vector	vector
log_avg_fmort	-0.15	-0.24	-0.52	-0.13	-0.15	-0.24	-0.15
fmort_dev	vector	vector	vector	vector	vector	vector	vector
log_avg_fmortdf	-6.42	-6.36	-6.76	-6.36	-5.73	-6.41	-5.8
fmortdf_dev	vector	vector	vector	vector	vector	vector	vector
log_avg_fmortt	-4.21	-4.65	-4.12	-4.6	-4.49	-4.7	-4.55
fmortt_dev_era1	vector	vector	vector	vector	vector	vector	vector
fmortt_dev_era2	vector	vector	vector	vector	vector	vector	vector
log_avg_sel50_mn	4.67	4.67	4.69	4.67	4.67	4.66	4.66
log_avg_sel50_mo	4.5	4.5	4.5	4.5	4.5	4.5	4.5
fish_slope_mn	0.19	0.19	0.18	0.19	0.19	0.19	0.19
fish_fit_slope_mn	0.42	0.42	0.42	0.43	0.43	0.43	0.43
fish_fit_sel50_mn	95.78	96	96.28	95.87	95.96	95.98	95.94
fish_slope_mo2	1.95	1.95	1.95	1.95	1.95	1.95	1.95
fish_sel50_mo2	159.5	159.5	159.5	159.5	159.5	159.5	159.5
fish_slope_mn2	1	1	1	1	1	1	1
fish_sel50_mn2	130	130	130	130	130	130	130
fish_disc_slope_f	0.24	0.24	0.24	0.24	0.25	0.24	0.25
fish_disc_sel50_f	4.26	4.25	4.23	4.24	4.24	4.25	4.24
fish_disc_slope_tf	0.08	0.09	0.07	0.09	0.08	0.09	0.08
fish_disc_sel50_tf	114.18	110.85	130	110.25	113.37	108.78	111.04
srv1_q	1	0.6	0.6	0.6	0.6	0.6	0.6
srv1_q_f	1	0.6	0.6	0.6	0.6	0.6	0.6
srv1_sel95	59.89	60.19	60	60.74	60.21	60.34	60.35
srv1_sel50	42.66	40.18	40	40.69	40.2	40.32	40.33
srv2_q	0.49	0.43	0.6	0.47	0.54	0.45	0.56
srv2_q_f	0.32	0.46	0.6	0.49	0.63	0.46	0.63
srv2_sel95	61.3	57.05	105	57.32	55.24	58.05	56.01
srv2_sel50	41.32	41.18	40	40.84	39.82	41.35	39.86
srv3_q	0.62	0.68	0.6	0.79	0.77	0.7	0.79
srv3_sel95	57.24	57.63	120	59.43	49.53	59.37	50.62
srv3_sel50	38.42	38.59	57.5	38.78	34.78	39.15	34.94

Parameter	Base	Trim data	Fixed obs sel	No BSFRF	Est female M	Chop growth	Growth + M
srv3_q_f	0.49	0.54	0.6	0.62	1	0.54	1
srv3_sel95_f	43.09	43.42	95	42.85	45.23	42.84	44.8
srv3_sel50_f	33.27	33.47	45	32.97	34.73	32.94	34.27
srvind_q	0.36	0.39	0.3	0.55	0.41	0.39	0.41
srvind_q_f	0.11	0.12	0.1	0.5	0.17	0.11	0.17
srvind_sel95_f	55	55	55	87.5	55	55	55
srvind_sel50_f	49.21	49.23	49.07	2.5	49.33	49.25	49.34
srv10ind_q_f	1	1	1	0.5	1	1	1
selsmo10ind	vector	vector	vector	vector	vector	vector	vector
selsmo09ind	vector	vector	vector	vector	vector	vector	vector
Mmult_imat	1.8	1.81	1.22	1.74	1.28	1.81	1.27
Mmult	1.13	1.08	1.13	1.06	1.15	1.09	1.16
Mmultf	1	1	1	1	1.41	1	1.42
cpueq	0	0	0	0	0	0	0

Table 4: Contribution to the objective function by individual likelihood component by modeling scenario. Values in columns after Model 0 are the likelihood contribution of Model 0 minus the likelihood contribution of the model in the column. Positive values represent improvements in fit. Note that some of the model scenarios involve changing the weightings of data sources which invalidate the comparison of likelihoods for a data source among models.

Likelihood component	Base	Trim data	Fixed obs sel	No BSFRF	Est female M	Chop growth	Growth + M
Recruitment deviations	40.33	39.38	38.73	38.62	39.24	38.4	37.95
Initial numbers old shell males small length bins	2.18	4.8	5.27	4.86	4.58	4.82	4.6
ret fishery length	382.62	292.49	300.94	294.43	291.49	304.35	303.67
total fish length (ret + disc)	821.72	820.67	828.46	819.21	819.98	827.64	827.6
female fish length	221.96	221.44	220.12	221.82	220	221.06	219.82
survey length	4639.21	4225.31	4300.94	4235.15	4171.65	4258.1	4196.8
trawl length	290.39	294.33	264.09	304.51	261.51	299.16	263.01
2009 BSFRF length	-82.94	-82.86	-90.9	0	-89.75	-82.63	-89.88
2009 NMFS study area length	-67.8	-67.49	-73.87	0	-73	-66.79	-72.56
M multiplier prior	19.81	16.74	7.28	13.41	79.34	17.25	80.91
maturity smooth	40.77	37.58	36.82	39.6	37.45	44.61	45.69
growth males	38.4	35.71	58.56	39.36	35.72	145.02	144.76
growth females	133.39	125.71	110.64	123.58	118.63	117.3	114.46
2009 BSFRF biomass	0.21	0.28	0.16	0	0.36	0.31	0.39
2009 NMFS study area biomass	0.09	0.15	0.12	0	0.3	0.18	0.33
cpue q	0.2	0.18	0.13	0.2	0.17	0.18	0.17
retained catch	4.06	3.86	3.9	3.57	3.75	4.12	3.98
discard catch	152.45	146.09	198.01	119.68	137.9	155.09	145.63
trawl catch	8.63	6.4	5.89	7.06	6.29	6.23	6.1
female discard catch	6.12	5.55	5.1	5.7	5.53	5.55	5.55
survey biomass	365.81	326	368.96	323.32	295.52	329.74	299.75
F penalty	37.39	24.32	21.62	25.81	24.85	24.62	25.27
2010 BSFRF Biomass	2.88	4.23	0.86	0	20.59	4.44	21.17
2010 NMFS Biomass	0.87	1.29	0.17	0	1.66	1.42	1.74
Extra weight survey lengths first year	510.42	566.02	563.82	566.53	556.15	564.37	554.64
2010 BSFRF length	-54.58	-54.21	-53.02	0	-52.56	-54.07	-52.83
2010 NMFS length	-59.21	-58.79	-70.96	0	-61.2	-58.89	-61.51
smooth selectivity	3.3	3.35	3.8	0	3.55	3.35	3.56
smooth female selectivity	0	0	0	0	0	0	0
init nos smooth constraint	40.44	47.62	51.34	46.87	45.54	47.59	45.68
Total	7499.12	6986.15	7106.98	7233.29	6905.24	7162.52	7076.45

Table 5: Likelihoods form and weighting for each likelihood component for models in the analysis (continued below)

Likelihood component	Form	Base	Trim data	Fixed obs sel
Recruitment deviations	normal	0.71	0.71	0.71
Initial numbers old shell males small length bins	normal	707.1	707.1	707.1
ret fishery length	multinomial	200	200	200
total fish length (ret + disc)	multinomial	200	200	200
female fish length	multinomial	200	200	200
survey length	multinomial	200	200	200
trawl length	multinomial	200	200	200
2009 BSFRF length	multinomial	200	200	200
2009 NMFS study area length	multinomial	200	200	200
M multiplier prior	normal	0.23	0.23	0.23
maturity smooth	normal	3.16	3.16	3.16
growth males	normal	0.71	0.71	0.71
growth females	normal	0.32	0.32	0.32
2009 BSFRF biomass	lognormal	NA	NA	NA
2009 NMFS study area biomass	lognormal	NA	NA	NA
cpue q	normal	0.32	0.32	0.32
retained catch	normal	0.22	0.22	0.22
discard catch	normal	3	3	3
trawl catch	normal	0.22	0.22	0.22
female discard catch	normal	17	17	17
survey biomass	lognormal	NA	NA	NA
F penalty	normal	0.5	0.5	0.5
2010 BSFRF Biomass	lognormal	NA	NA	NA
2010 NMFS Biomass	lognormal	NA	NA	NA
Extra weight survey lengths first year	multinomial	200	200	200
2010 BSFRF length	multinomial	200	200	200
2010 NMFS length	multinomial	200	200	200
smooth selectivity	norm2(firstdiff(firstDiff))	2	2	2
smooth female selectivity	norm2(firstdiff(firstDiff))	3	3	3

Likelihood component	Form	Base	Trim data	Fixed obs sel
init nos smooth constraint	norm2(firstdifference)	1	1	1

No BSFRF	Est female M	Chop growth	Growth + M
0.71	0.71	0.71	0.71
707.1	707.1	707.1	707.1
200	200	200	200
200	200	200	200
200	200	200	200
200	200	200	200
200	200	200	200
200	200	200	200
200	200	200	200
0.23	0.23	0.23	0.23
3.16	3.16	3.16	3.16
0.71	0.71	0.22	0.22
0.32	0.32	0.22	0.22
NA	NA	NA	NA
NA	NA	NA	NA
0.32	0.32	0.32	0.32
0.22	0.22	0.22	0.22
3	3	3	3
0.22	0.22	0.22	0.22
17	17	17	17
NA	NA	NA	NA
0.5	0.5	0.5	0.5
NA	NA	NA	NA
NA	NA	NA	NA
200	200	200	200
200	200	200	200
200	200	200	200
2	2	2	2
3	3	3	3
1	1	1	1

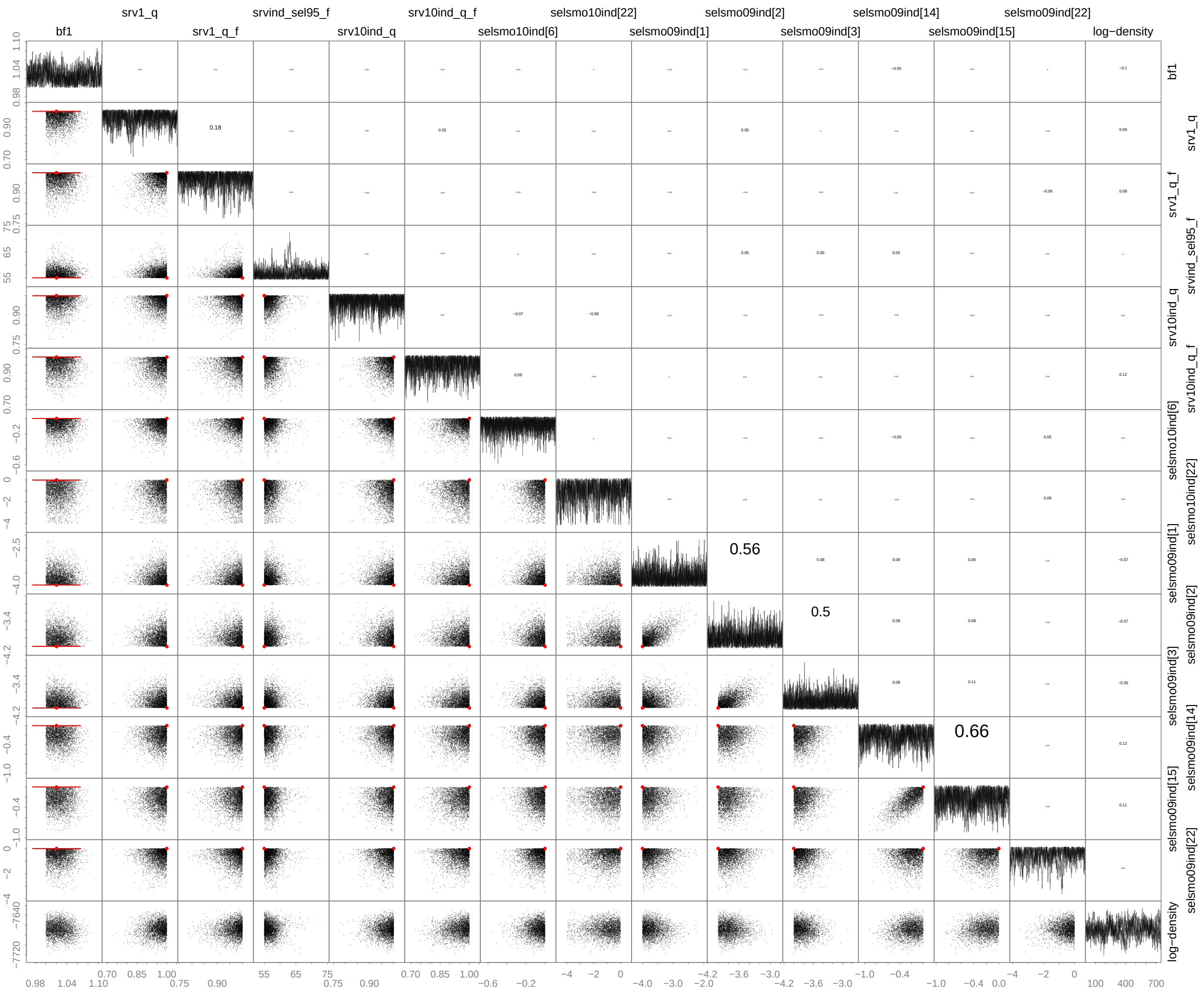

Figure 1: Slow mixing parameters from the base model of MCMC. Diagonal is the trace of the parameter over multiple chains. Lower triangular is a pairs plot of parameters on the x and y axes. Ellips are the 95% bivariate confidence intervals calculated from the .cor file. Frequentist estimates of uncertainty assume that the points are mulitvariate normal with the estimated covariance structure. Where points do not lie in the ellipses, this assumption is violated. Upper triangular values are the empirical correlations from the MCMC.

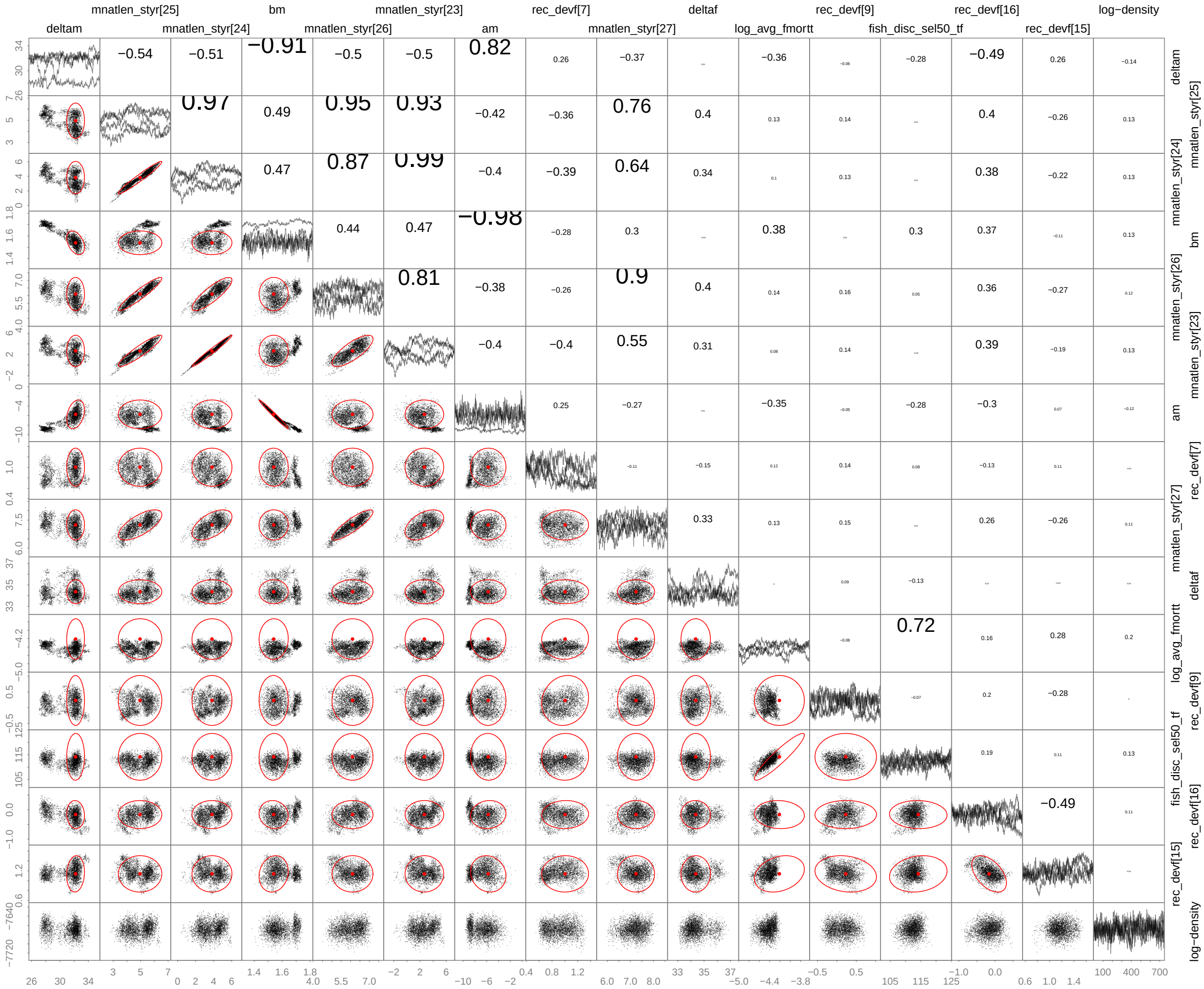

Figure 2: Bound hitting parameters from the base model of MCMC. Diagonal is the trace of the parameter over multiple chains. Lower triangular is a pairs plot of parameters on the x and y axes. Ellips are the 95% bivariate confidence intervals calculated from the .cor file. Frequentist estimates of uncertainty assume that the points are mulitvariate normal with the estimated covariance structure. Where points do not lie in the ellipses, this assumption is violated. Upper triangular values are the empirical correlations from the MCMC.

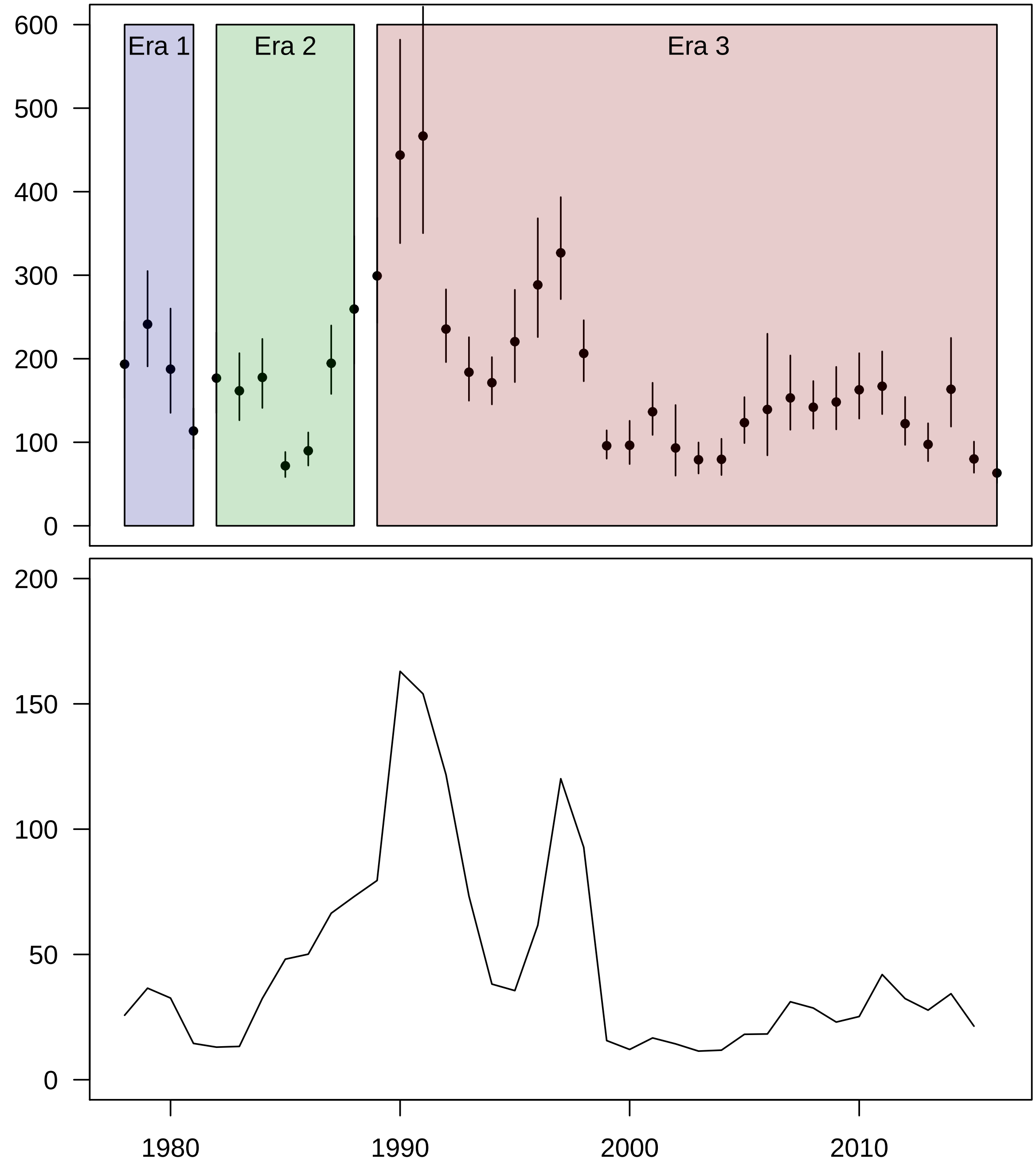

Figure 3: Divisions of survey data for estimation of q (MMB shown for reference) and total catches

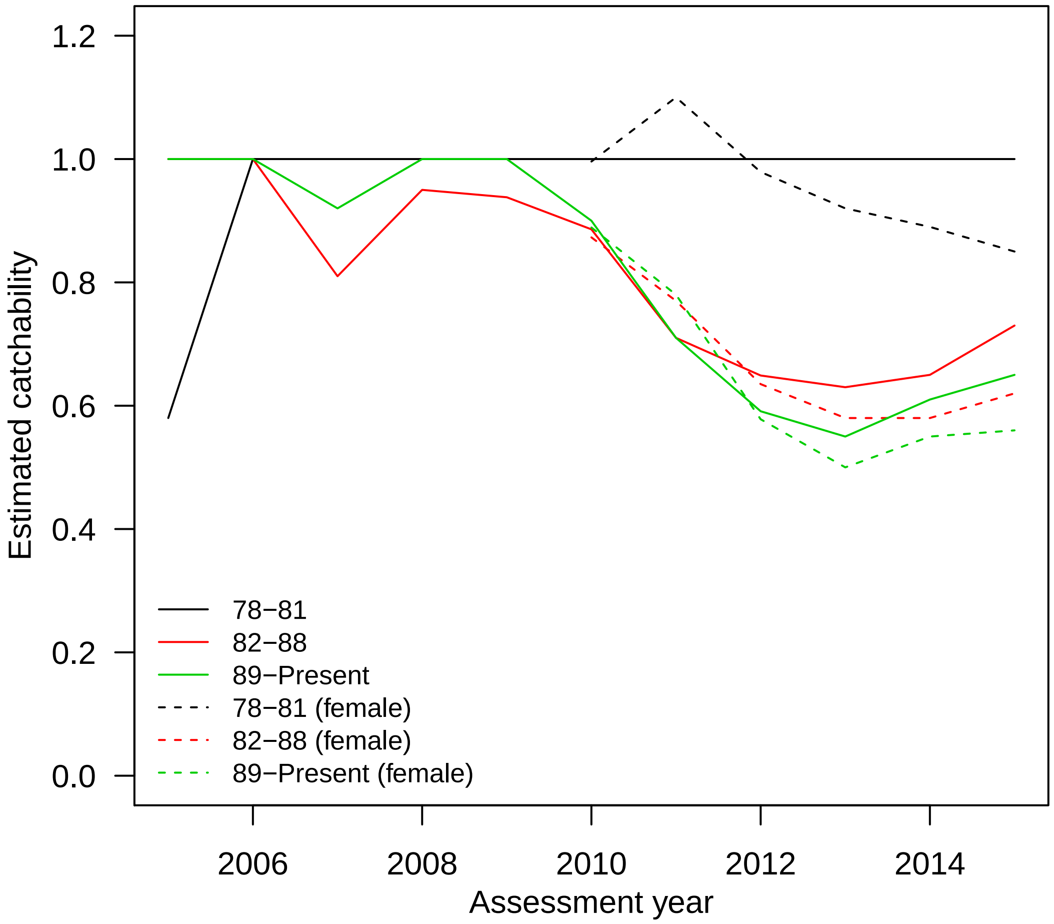

Figure 4: Changes in estimates of q within the assessment from 2005-2015

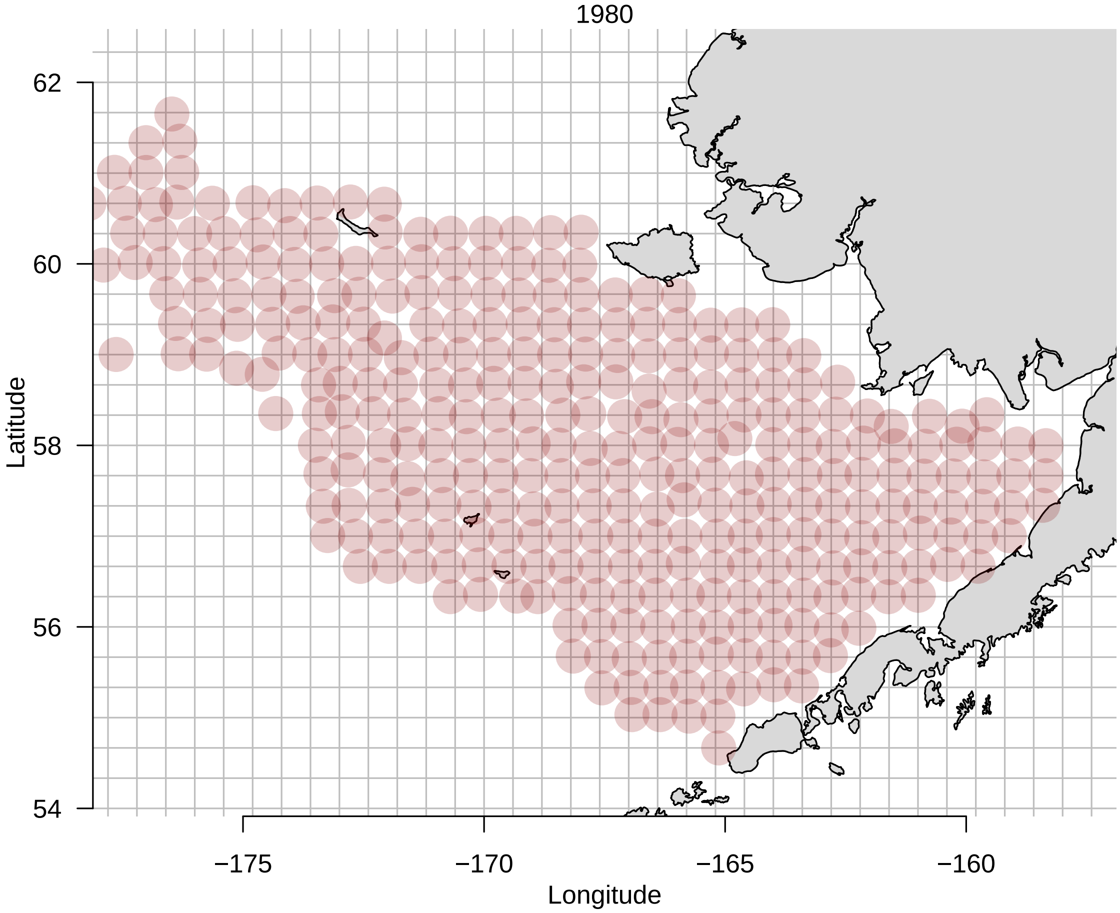

Figure 5: Distribution of survey locations in 1980

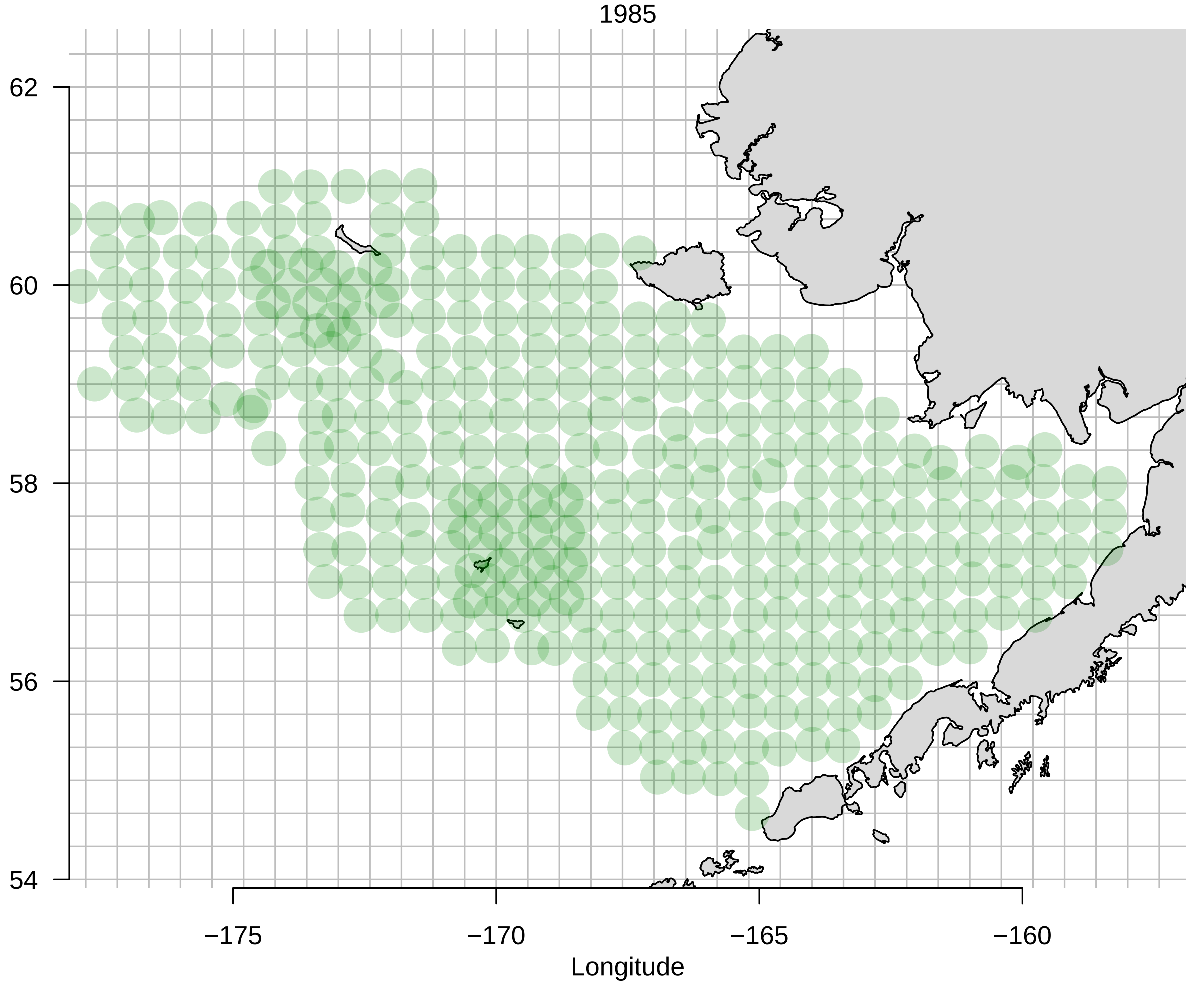

Figure 6: Distribution of survey locations in 1985

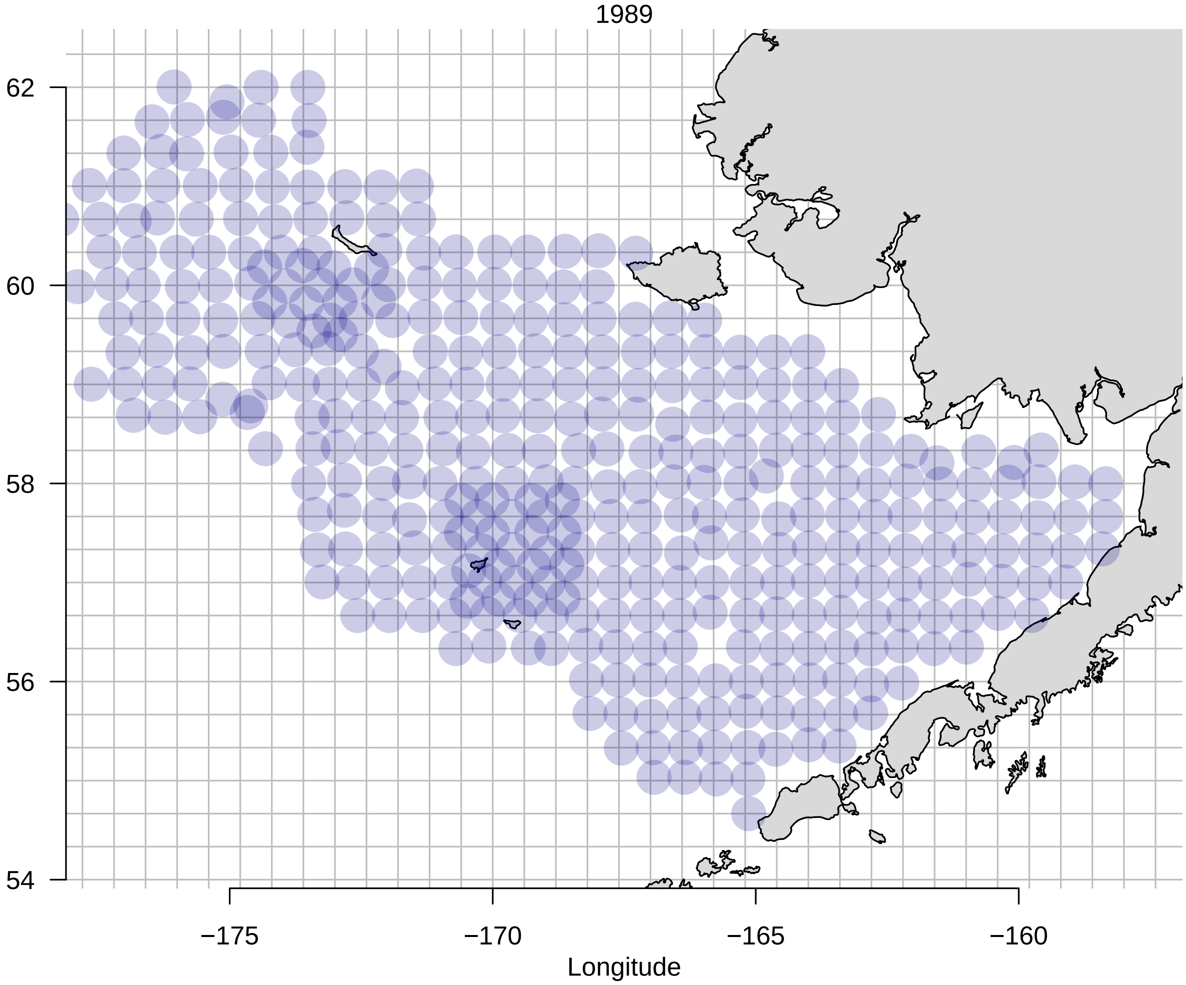

Figure 7: Distribution of survey locations in 1989

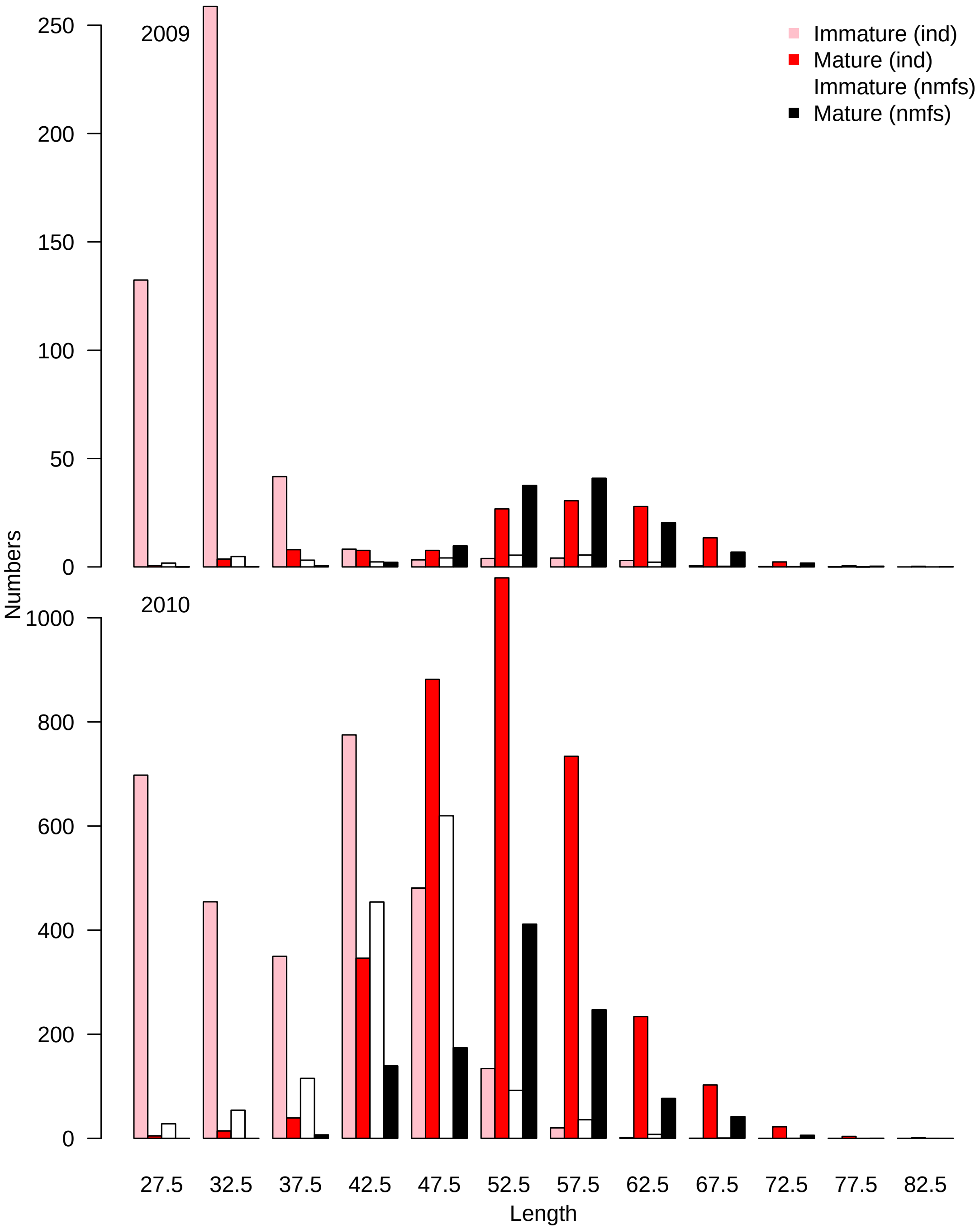

Figure 8: Raw female numbers from BSFRF survey selectivity experiments (2009 & 2010). Note a change in scale on the y-axis from 2009 to 2010

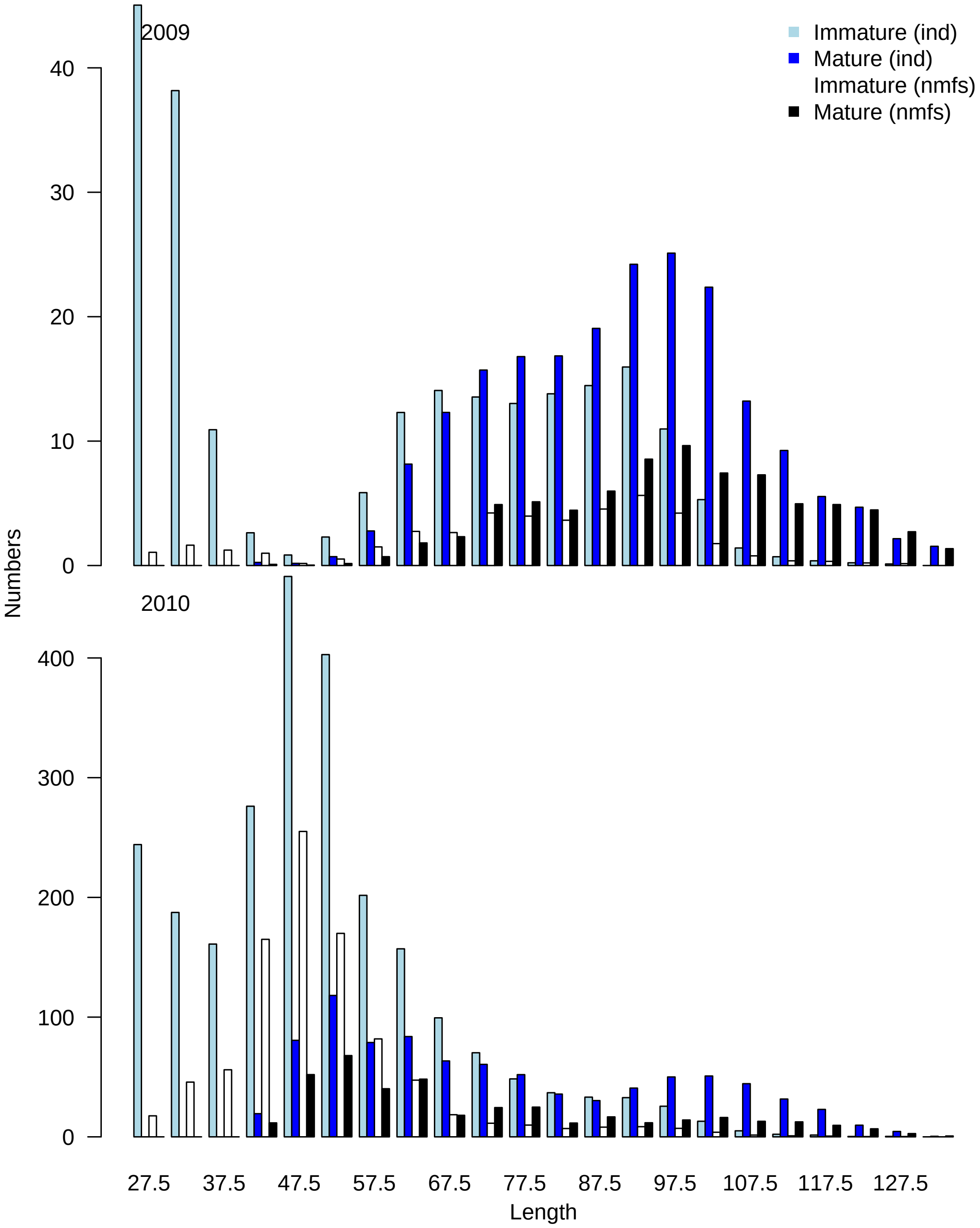

Figure 9: Raw male numbers from BSFRF survey selectivity experiments (2009 & 2010). Note a change in scale from 2009 to 2010 on the y-axis.

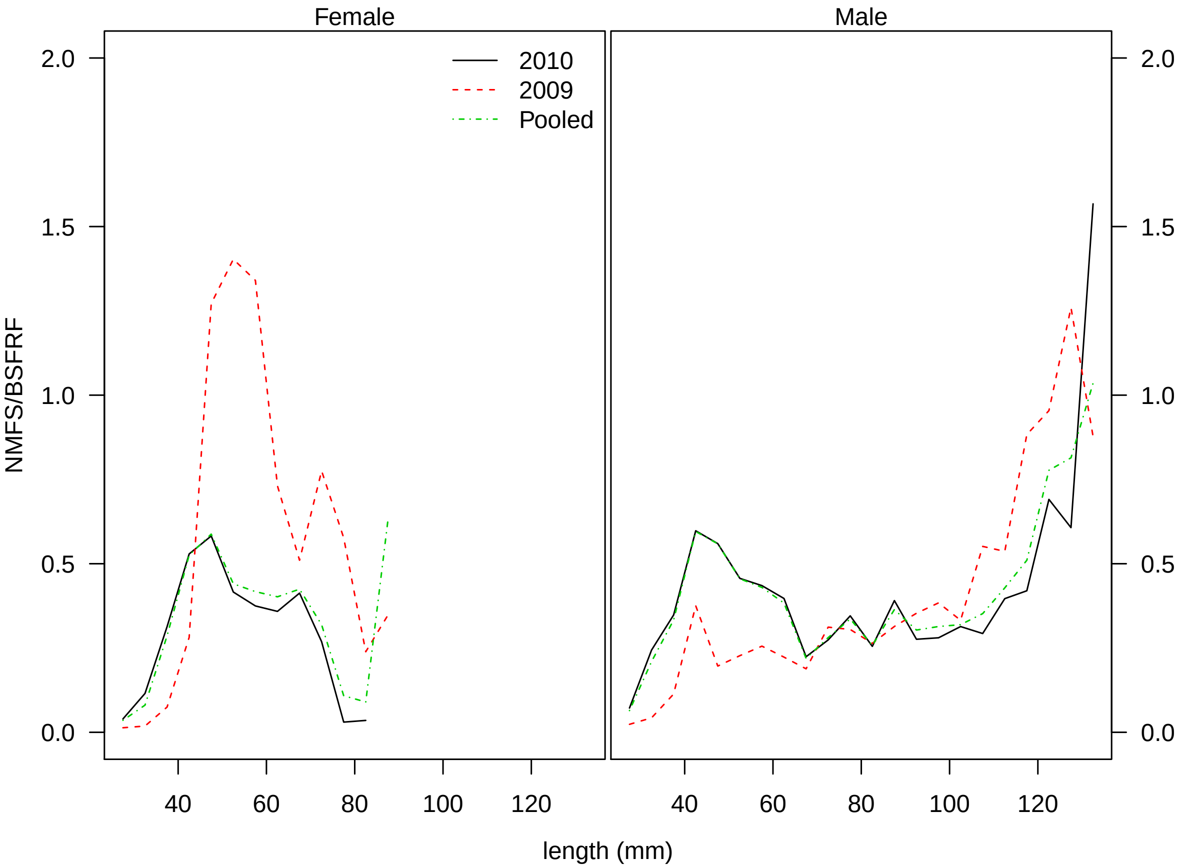

Figure 10: Ratio of observed densities at length in the NMFS survey vs. the BSFRF survey

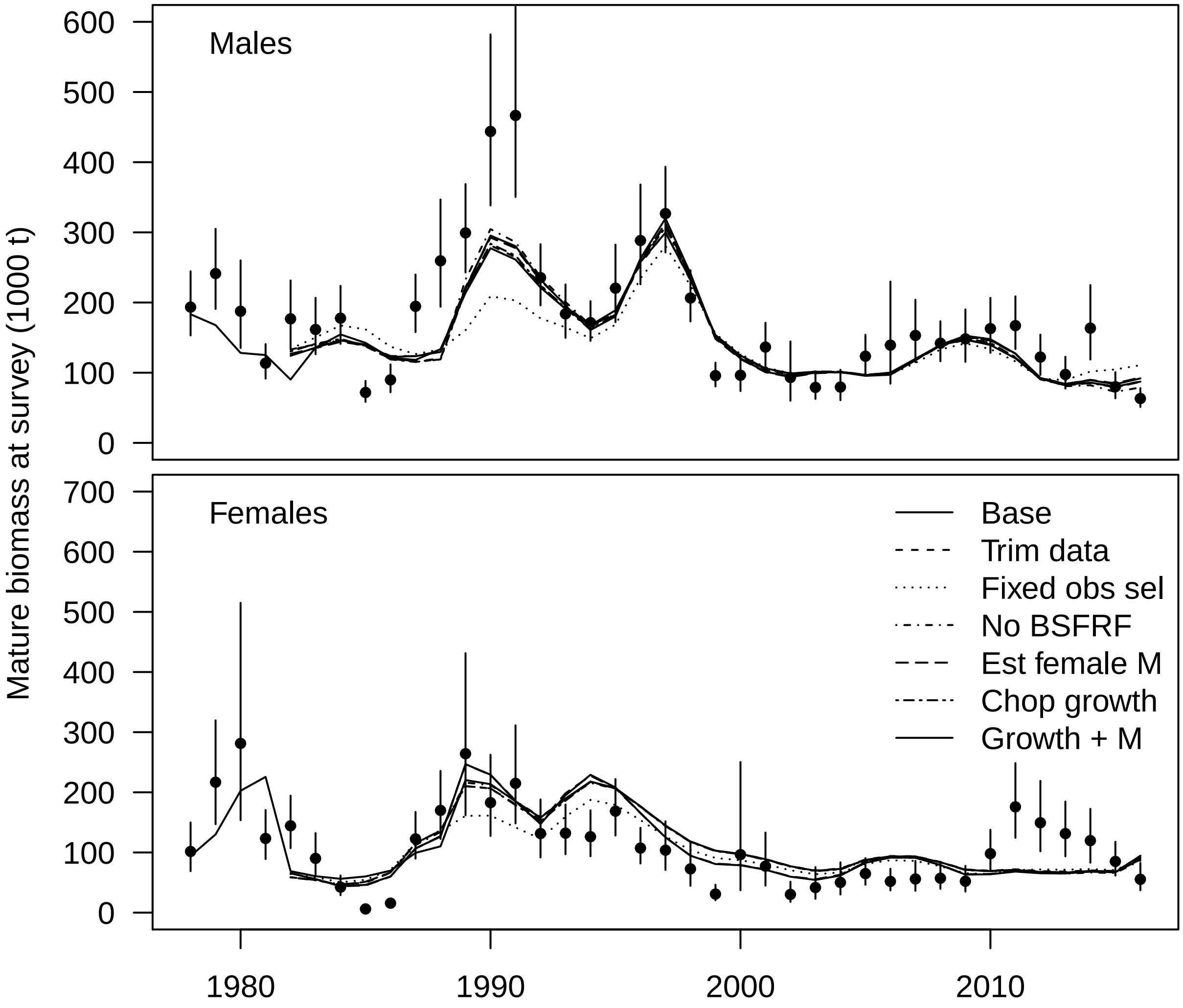

Figure 11: Model fits to the observed mature biomass at survey

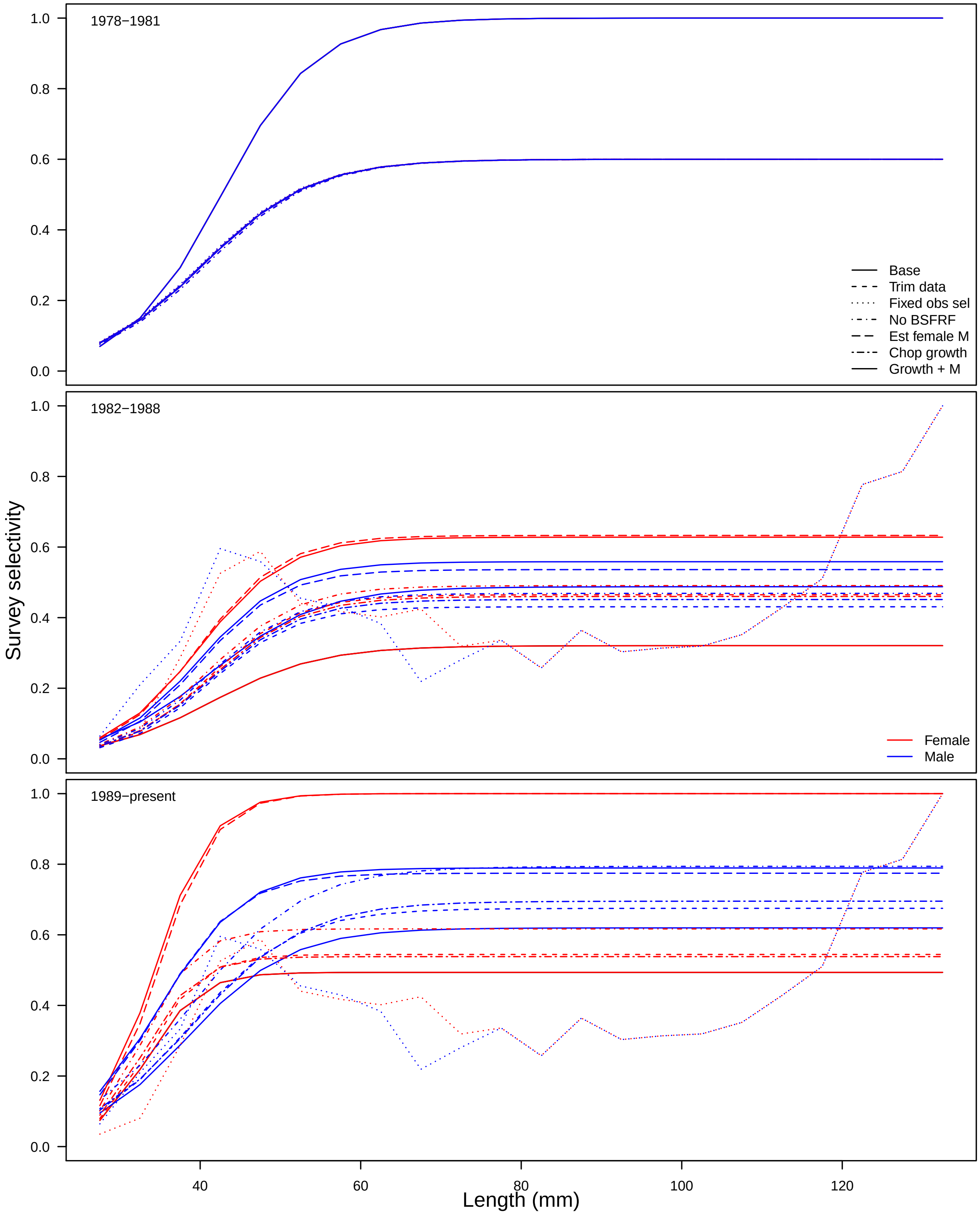

Figure 12: Estimated survey selectivity

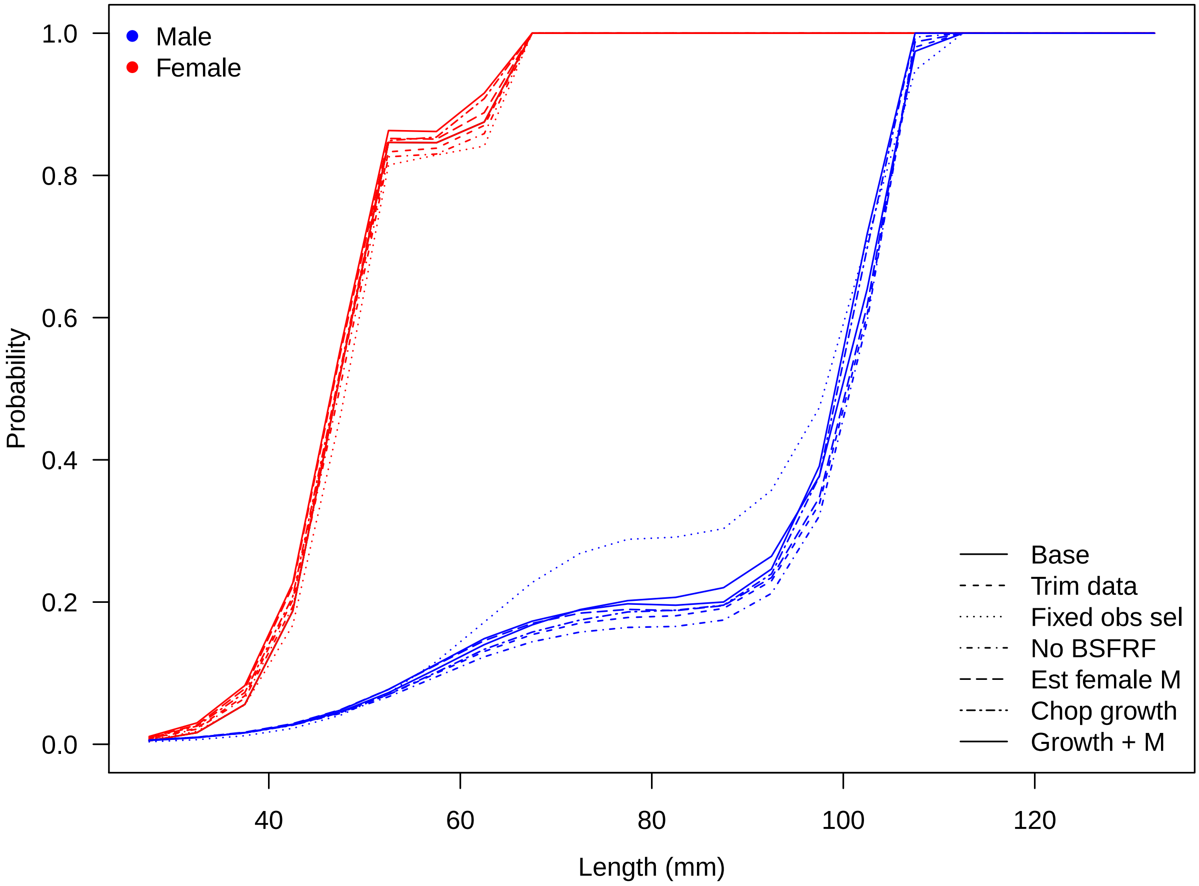

Figure 13: Estimated probability of maturing

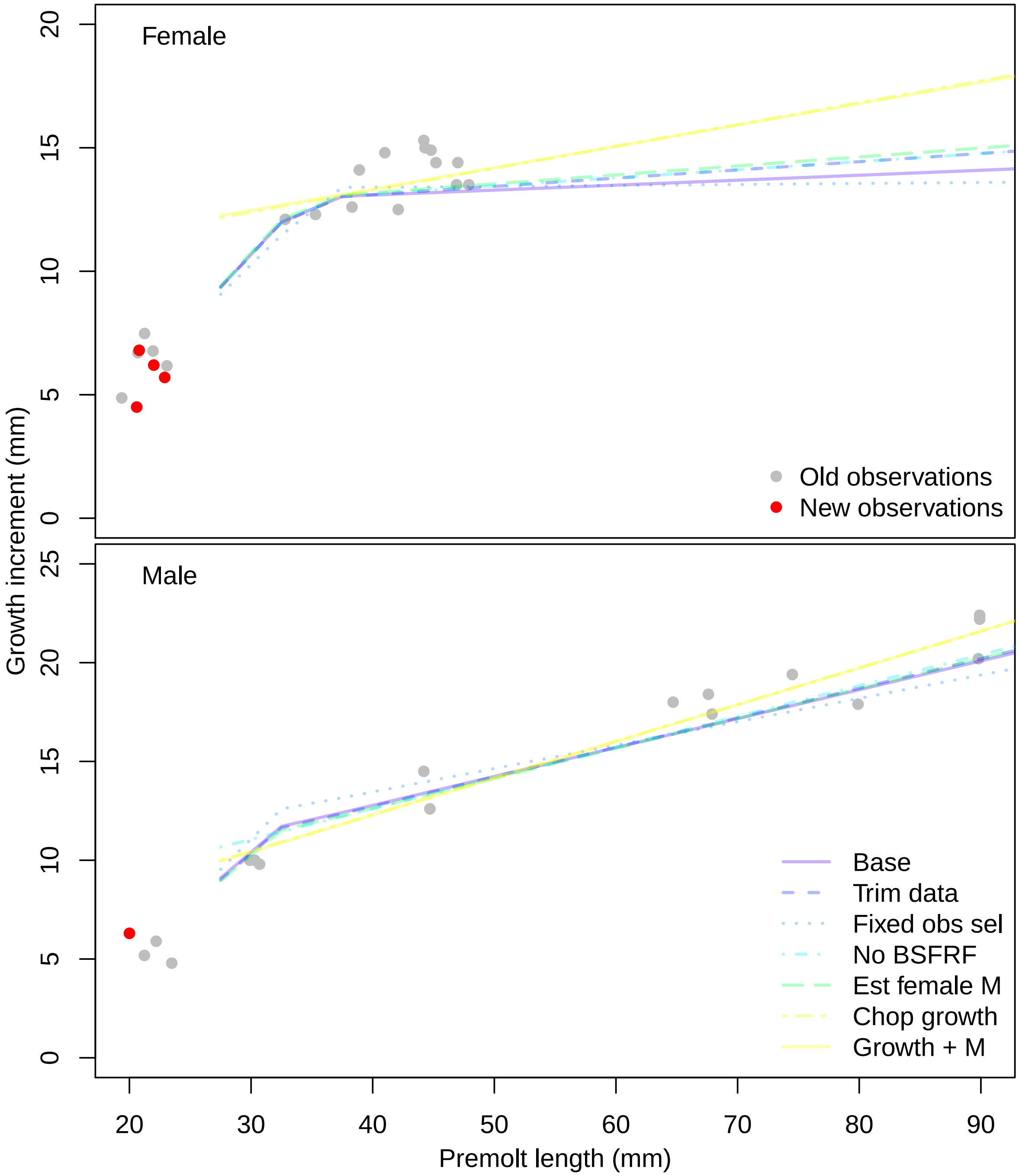

Figure 14: Model fits to the growth data

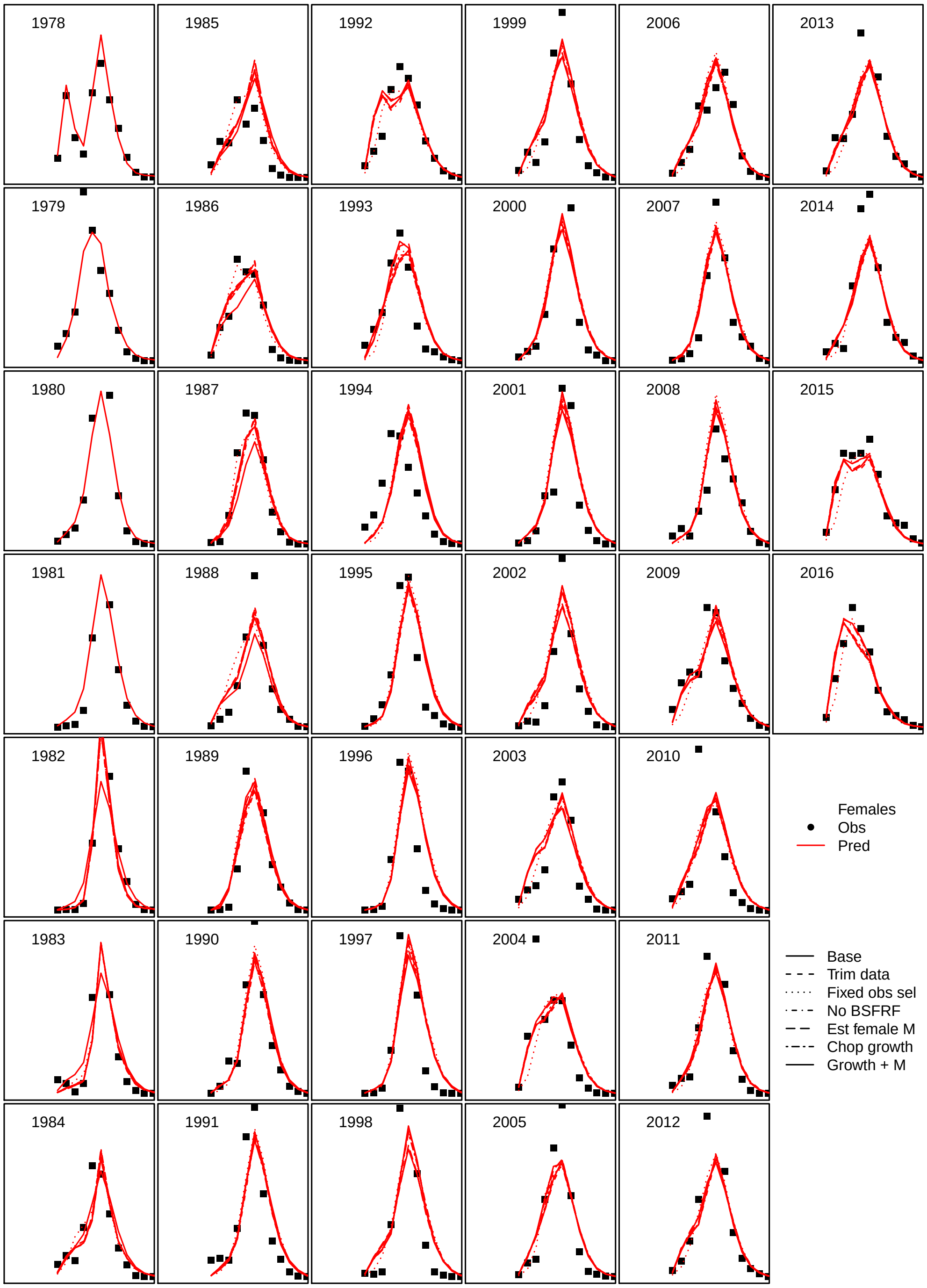

Figure 15: Model fits to female survey size composition data. Note that male and female survey selectivity proportions at length in a given year sum to 1. Consequently, the integral of predicted length compositions may appear to be different than the integral of the observed length composition data.

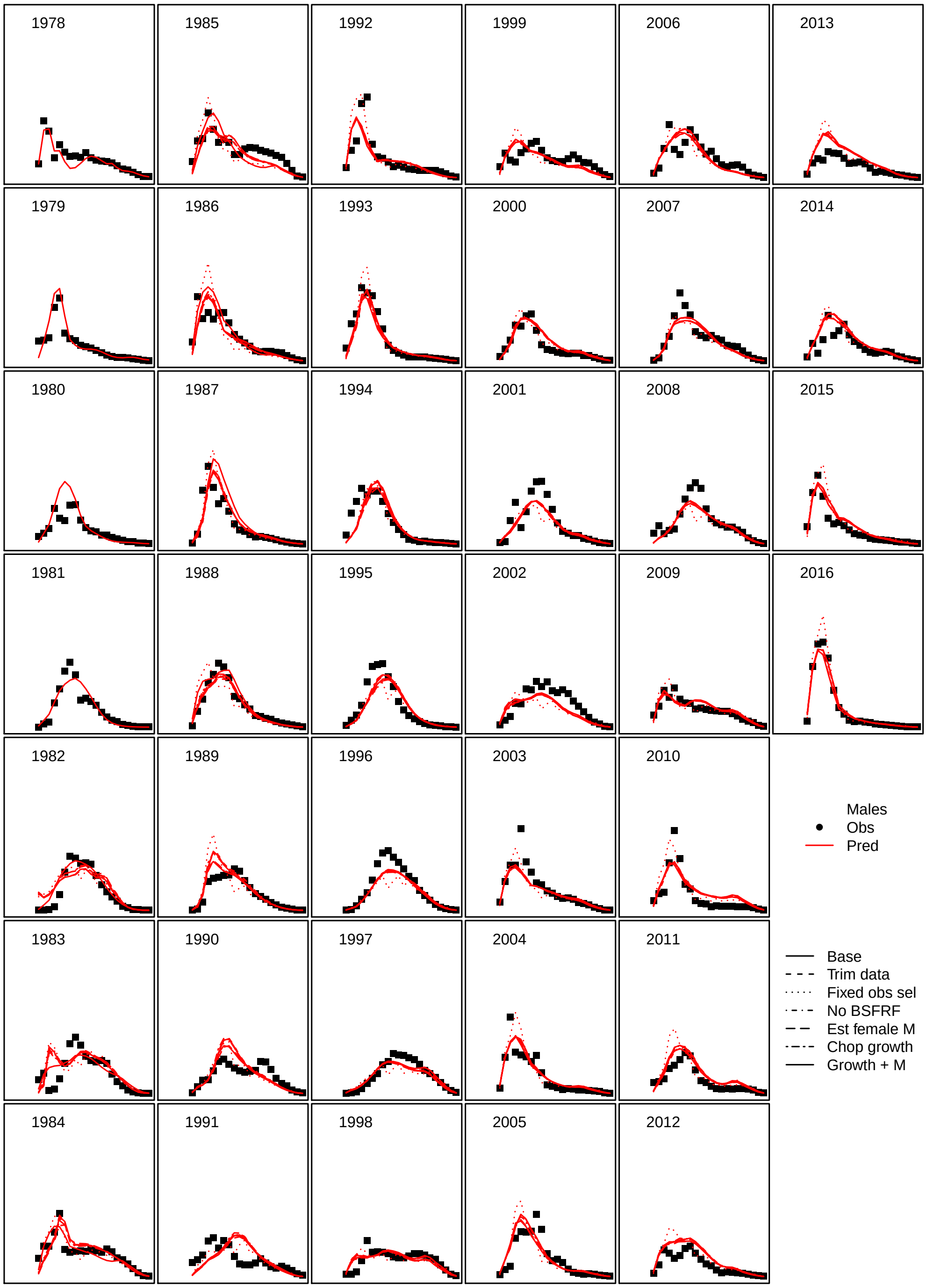

Figure 16: Model fits to male survey size composition data. Note that male and female survey selectivity proportions at length in a given year sum to 1. Consequently, the integral of predicted length compositions may appear to be different than the integral of the observed length composition data.

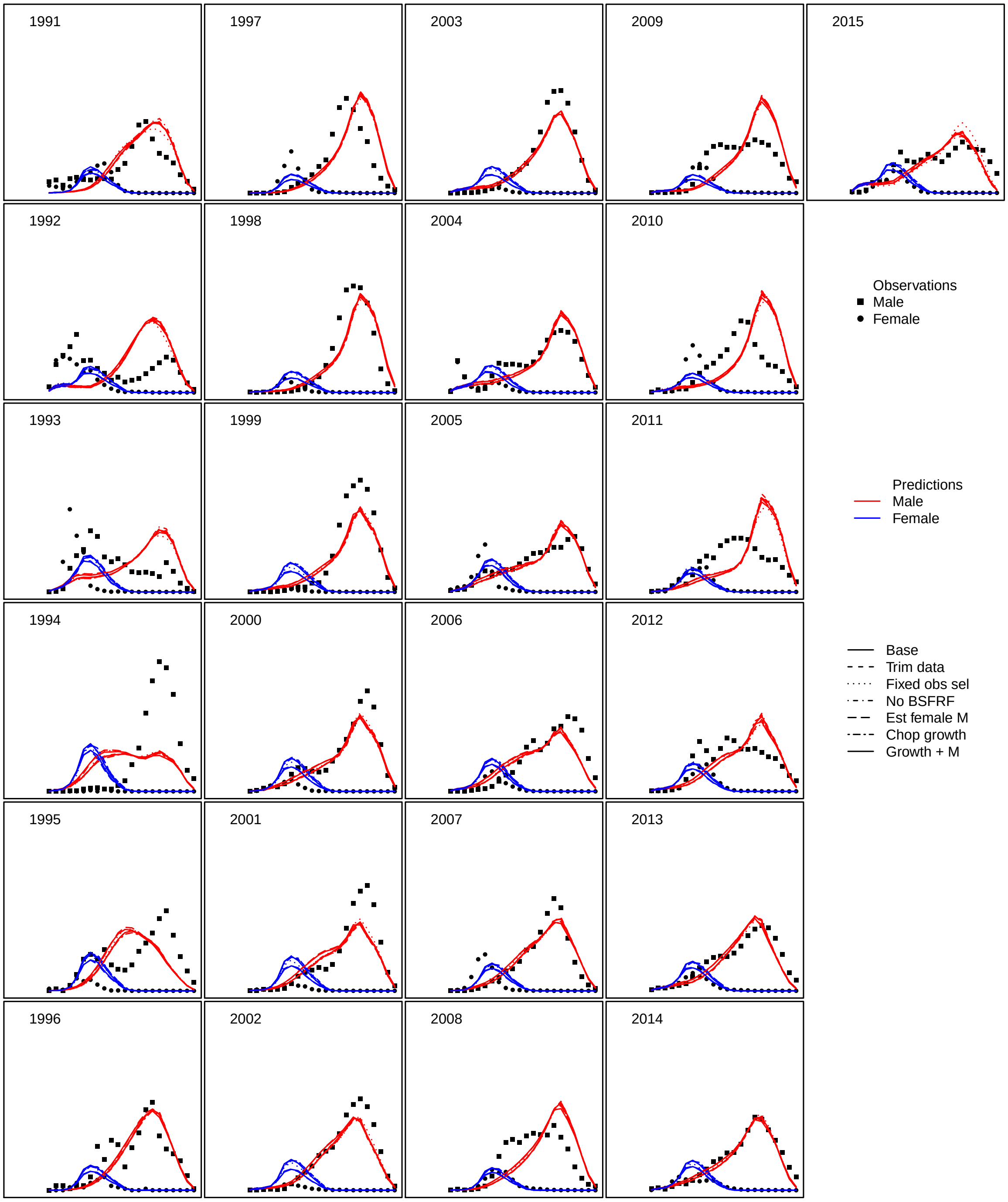

Figure 17: Model fits to trawl catch size composition data

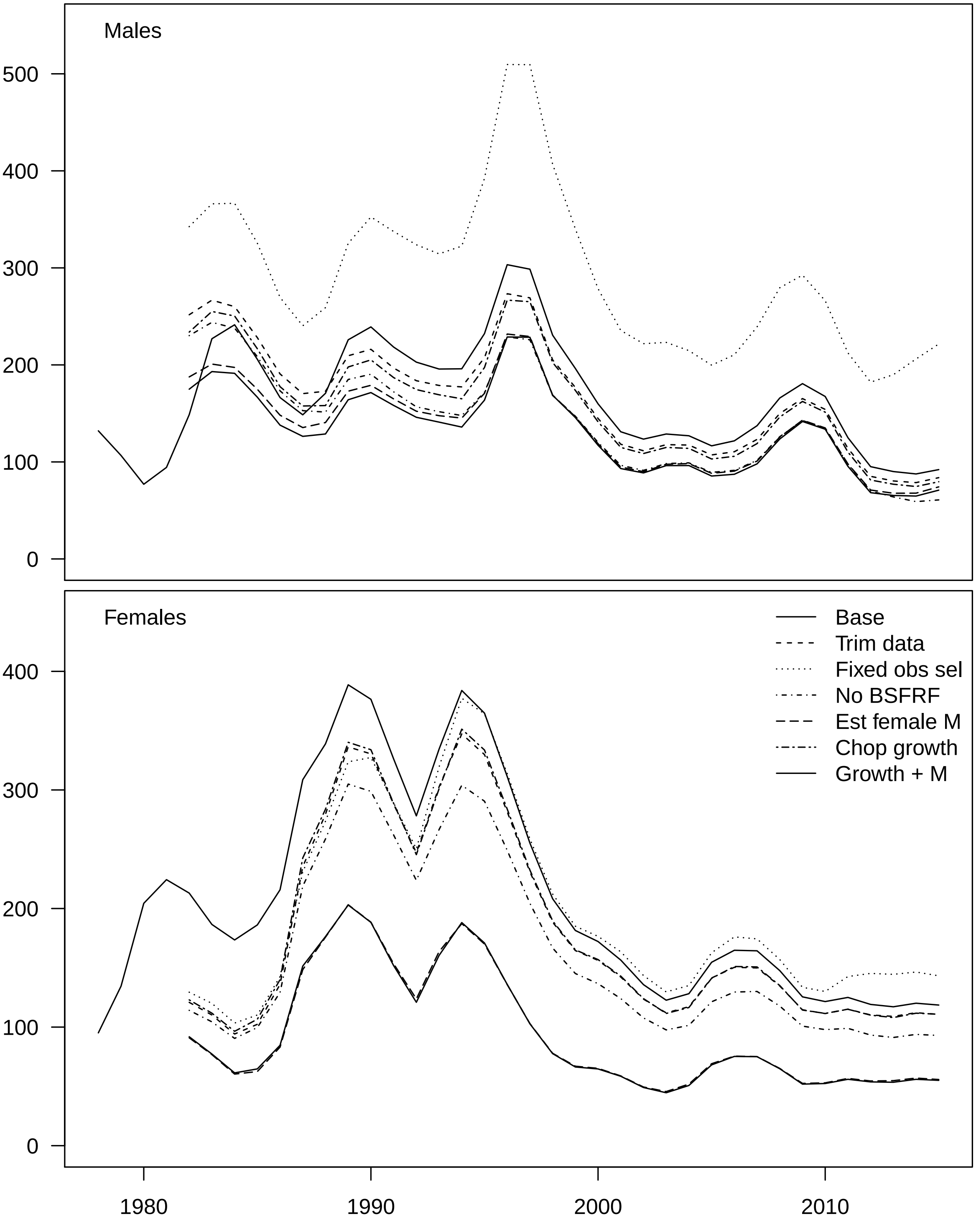

Figure 18: Model predicted mature male biomass at mating time

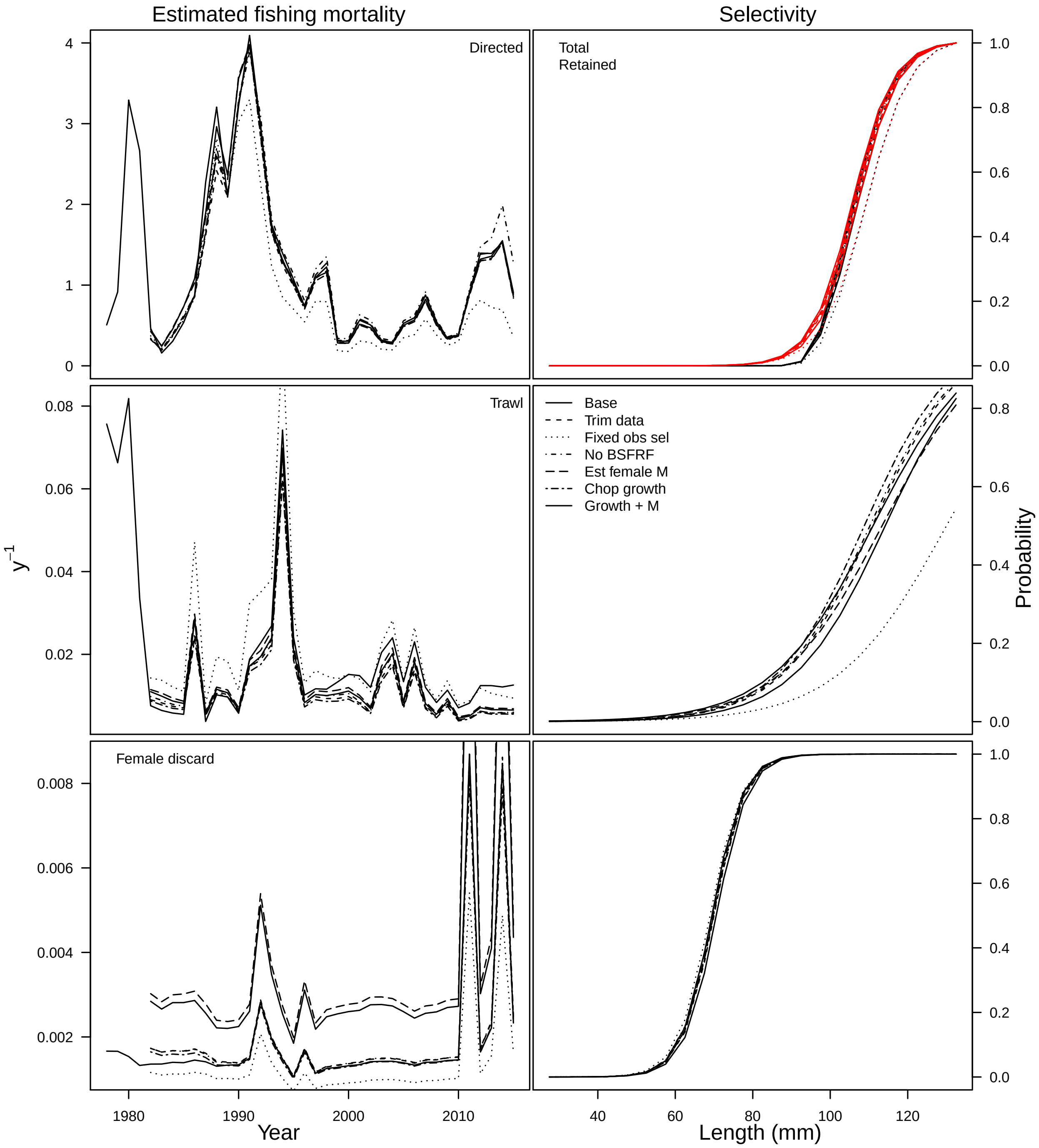

Figure 19: Model predicted fishing mortalities and selectivities for all sources of mortality

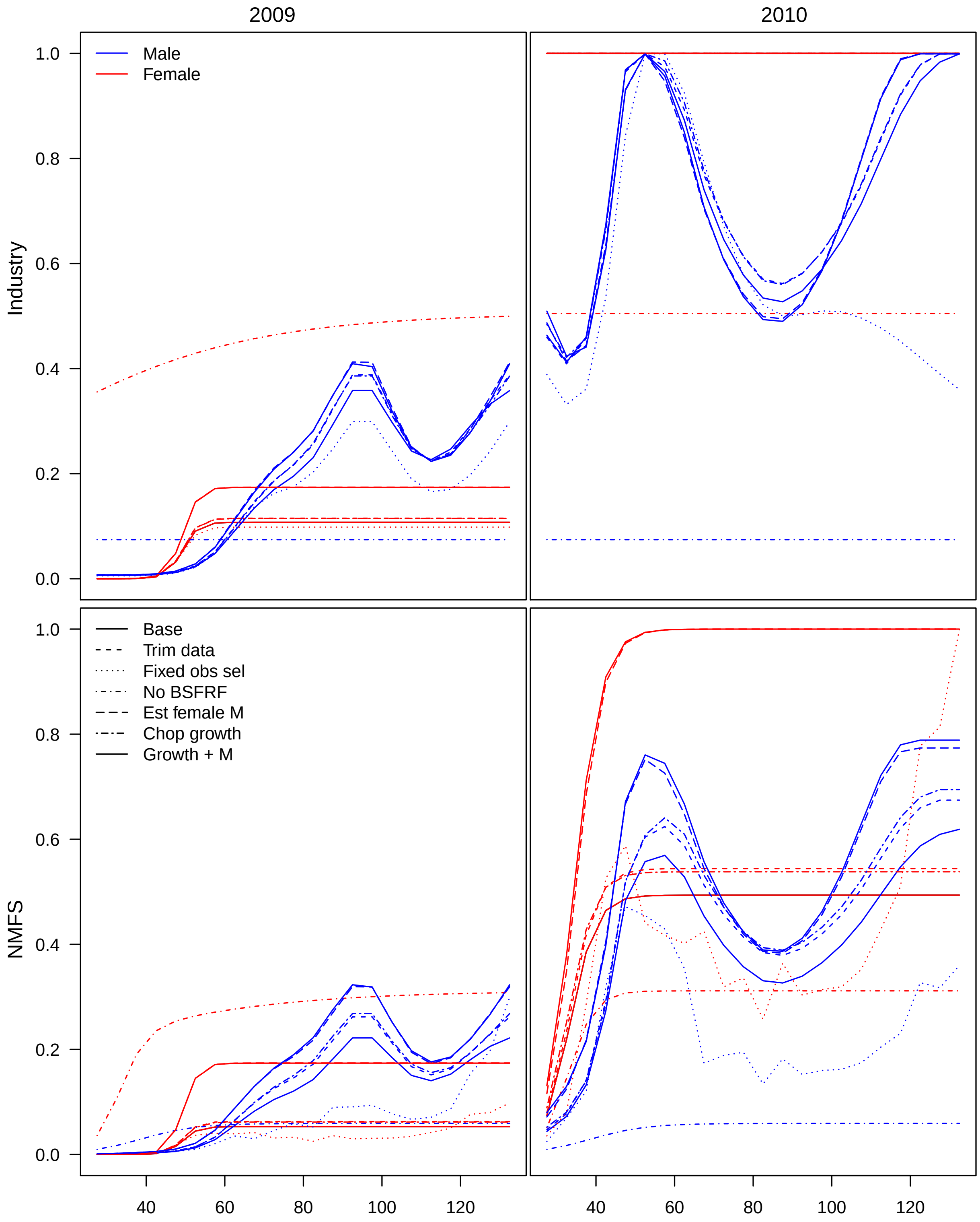

Figure 20: Estimated experimental survey selectivity (availability * survey selectivity)

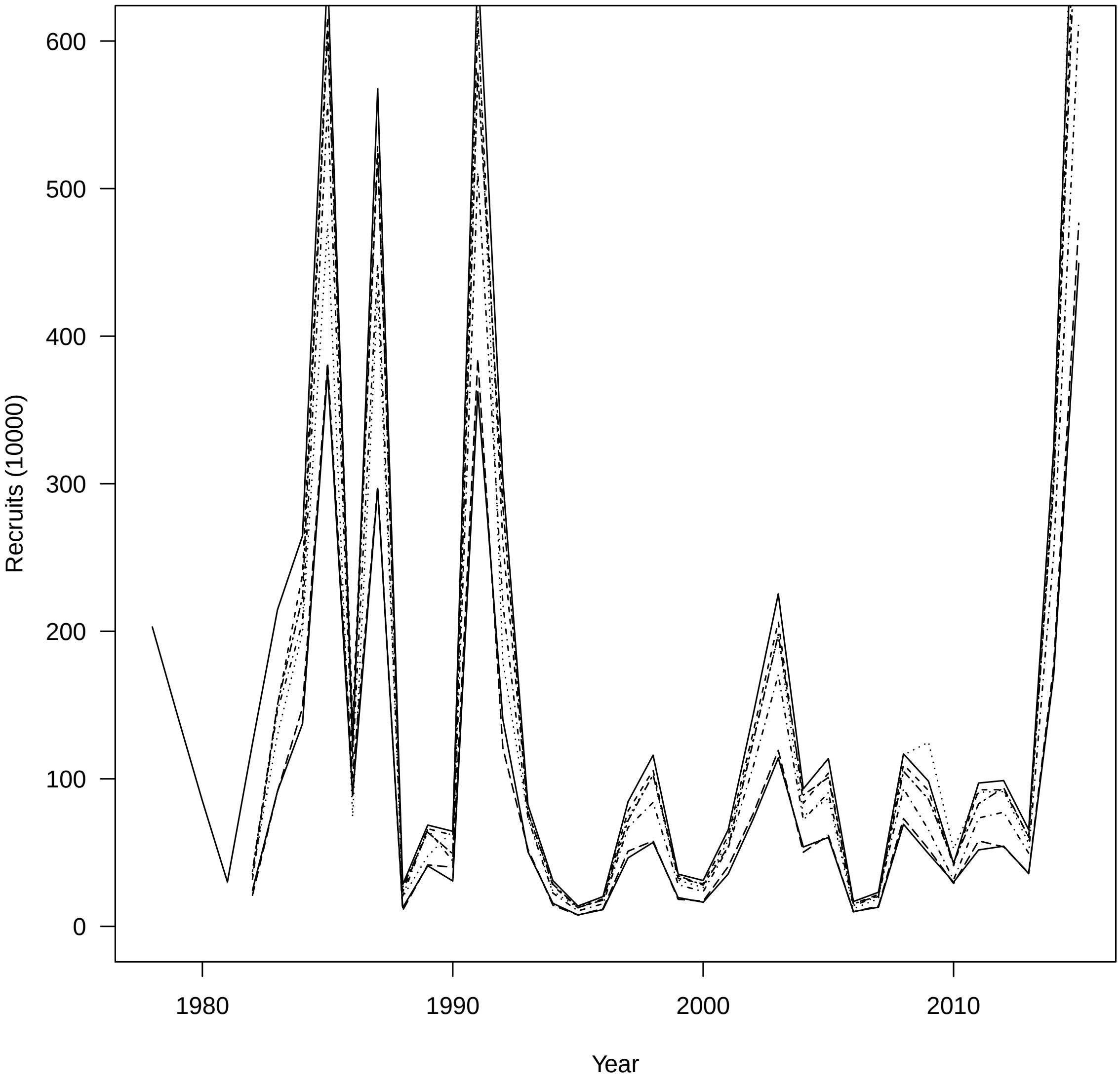

Figure 21: Estimated recruitment, fits to stock recruit curve (MMB lagged 5 years), and proportions recruiting to length bin

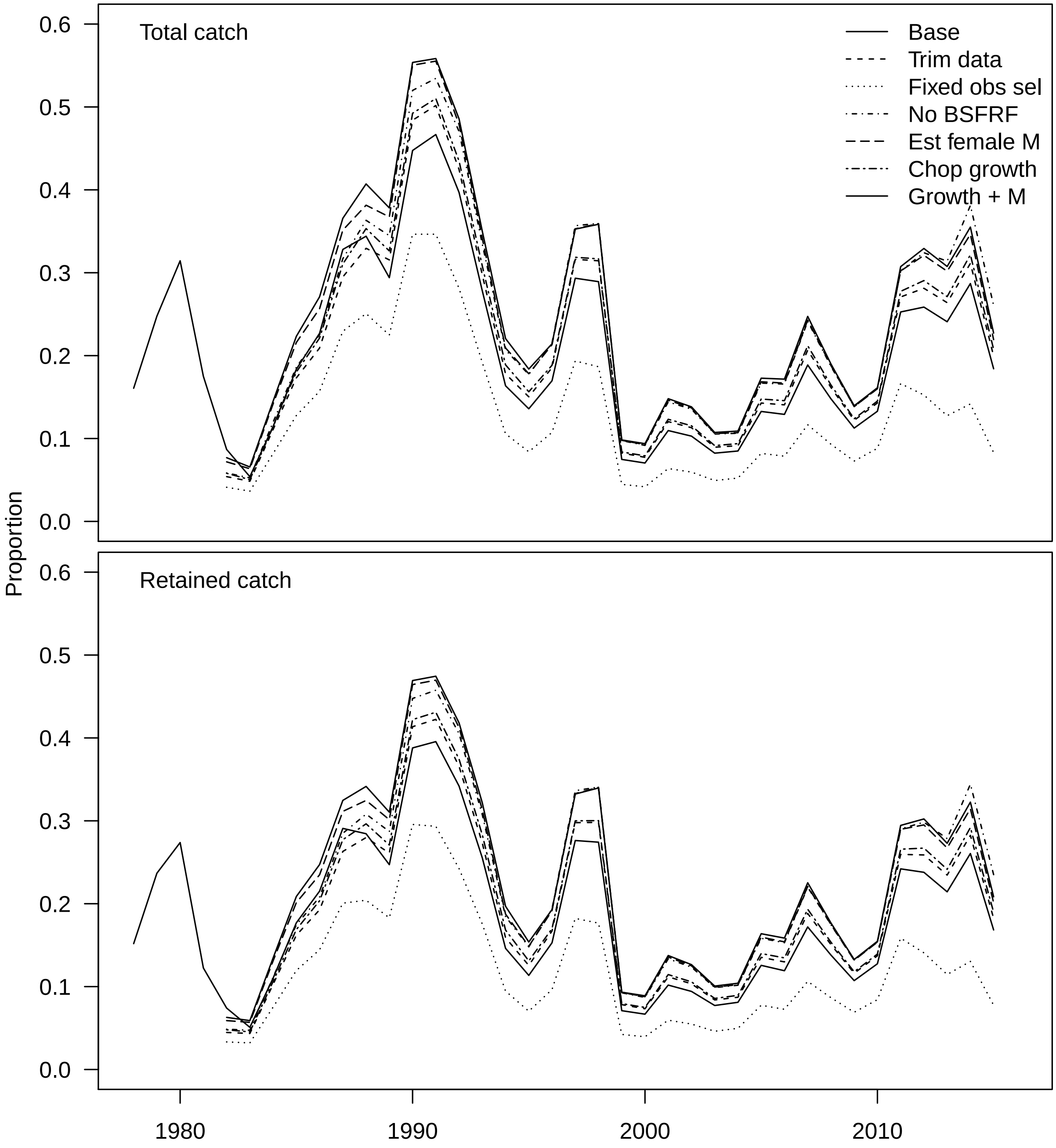

Figure 22: Model predicted ratio of catch to mature male biomass

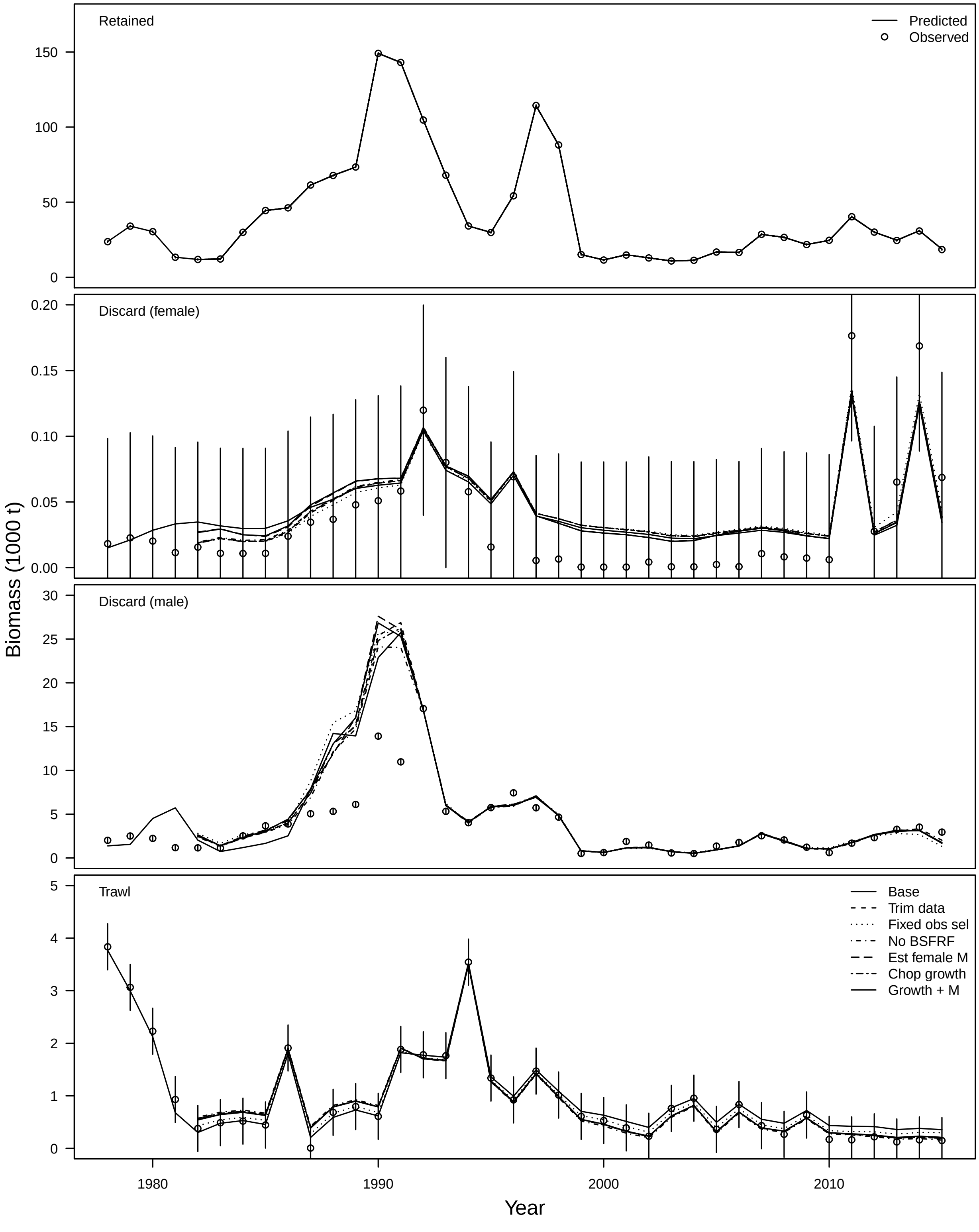

Figure 23: Model fits to catch data

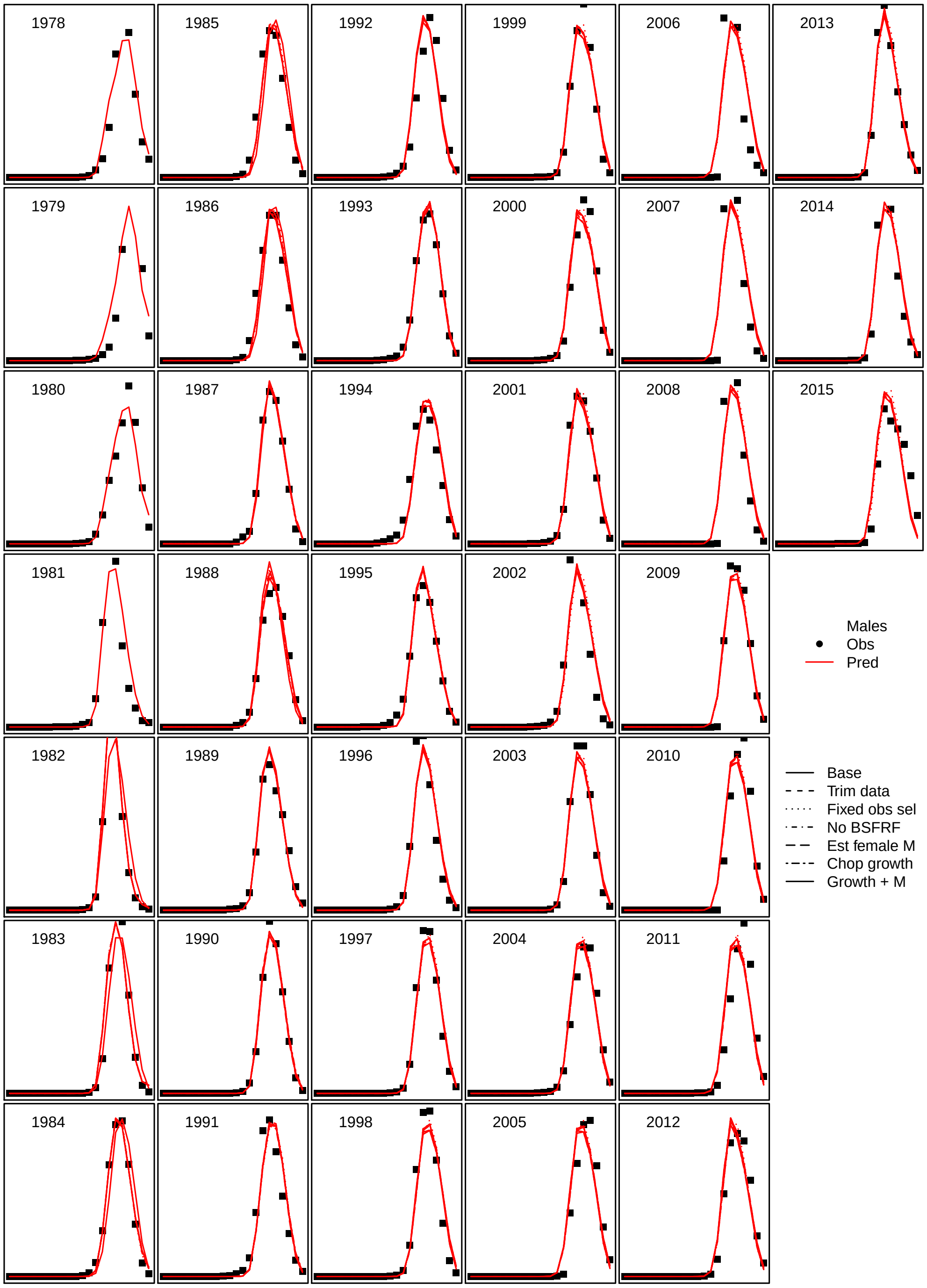

Figure 24: Model fits to retained catch size composition data