A Transition of the Flathead Sole-Bering Flounder Stock Assessment in the Bering Sea and Aleutian Islands to the Stock Synthesis Framework

Carey McGilliard, September 2018

Introduction

The purpose of this document is to outline a proposed change from conducting assessments using the previous BSAI flathead sole assessment model framework to conducting assessments using Stock Synthesis versions 3.24o and 3.30.12 (SS3; Methot and Wetzel 2013); the two versions of SS3 yield identical results.

Previous assessments were conducted using an ADMB-based age- and sex-structured population dynamics model with length-at-age, weight-at-length, maturity-at-age, and age-length transition matrices estimated outside of the model (referred to as "the 2016 model" or "the 2016 accepted model" in this document). The 2016 model estimated the log of mean recruitment, the log of historical (pre-1977) mean recruitment, the log of mean historical fishing mortality (pre-1977), parameters for logistic length-specific (but not sex-specific) selectivity curves for the fishery and survey, recruitment deviations, and yearly fishing mortality rates. The model included ages 3-21 (age 21 was a plus group) and excluded data for fish younger than age 3 and smaller than 6cm in length. In addition, the 2016 model estimated recruitment deviations beginning in the starting year of the model. The initial conditions assumed that the stock was at a fished equilibrium, based on values estimated within the model for mean historical recruitment and mean historical fishing mortality. A recruitment deviation was estimated for age 3 individuals in the initial year of the model. Additionally, the model assumed that spawning occurred in March of each year (but with spawning biomass calculated based on weight-at-age at the beginning of the year), and that fishing mortality and natural mortality occurred throughout the year. Numbers-at-length were used in equations for predicted catch and biomass and were based on the numbers-at-age and the age-length transition matrix (representative of mid-year lengths).

SS3 is a flexible assessment model framework that extends the capabilities of the old model code to address the concerns expressed by the BSAI Plan Team, the SSC, and previous assessment authors. As an initial effort towards addressing these concerns, this document outlines a framework designed to begin to resolve these issues and transition the assessment to the SS3 framework. SS3 allows for clear specification of alternative models that can easily deal with concerns and issues that have been raised. In particular:

(1) Additional assumptions about the initial conditions, including early recruitment deviations, can be included in the model.
(2) Data on the age distribution within each length bin or mean weight-at-age data can be included and used to estimate growth within the assessment model.
(3) Alternative functional forms of selectivity curves are available and can be used to explore age-based and dome-shaped fishery selectivity, as well as time-varying approaches; the previous model used logistic length-based selectivity.
(4) Multiple survey and fishing fleets can easily be included in the model and hence allow for easy explorations of selectivity and catchability for shelf, slope, and Aleutian Islands survey data, separately, as well as consideration for selectivity of pelagic vs. non-pelagic fishing gear, sector-specific fishery selectivity, or selectivity specific to intended target species.

(5) Alternative equations for including an environmental linkage between temperature and catchability, and for modeling time-varying catchability, can be explored.
(6) The timing of population dynamics and fishing processes is modeled with more attention to detail.
(7) Including ages 0 to 2 in the model, and corresponding data, informs the shape of the survey selectivity curve at its lower bound, and may help to resolve issues with unrealistic survey selectivity estimates.
(8) Area-specific fleet dynamics and/or population dynamics can be modeled if substantial differences are found in stock structure, age-distributions, or fishing characteristics between areas.
(9) Seasons can be modeled if it is found that the characteristics of fishing are substantially different in different parts of the year and that the year-long time step for selectivity cannot represent the dynamics well.

SSC and Plan Team Comments on Previous Assessments

GOA Plan Team comment 11/2016: The Team recommends examining the use of time blocks in selectivity due to changes in fishing practices.

Author Response: CRM completed a transition of the model to the SS3 framework and an SS3 model was run done using time blocks for fishery selectivity from 1964 (the model start year) to 1991, 1992-2007 (when the structure of our current halibut bycatch regulations were implemented), and 2008-present, when the BSAI groundfish trawl fishery was rationalized.

Data used in SS3 and the Old Model

An important difference between the old model code and SS3 is that the youngest age class in the old model code (age 3) represents only age 3 individuals, while SS3 population dynamics begin at age 0 and consider the lowest age and length bins of data to be the proportion of individuals ages 0-3 and lengths 0-the upper limit of the lowest length bin, respectively. Therefore, age- and length-composition data must include ages 0-2 and any lengths no matter how small in SS3, while the old model code omitted data on ages 0-2 (and excluded data on fish smaller than 6 cm). Ignoring this difference between models will result in differences between expected and observed age- and length-compositions for the youngest age and length bins when selectivity at these ages and lengths is estimated to be greater than 0 in SS3. The data on ages 0-2 that are included in SS3 can inform estimates of selectivity at the lowest ages (even, or especially, if they are all zeros). These data may also improve recruitment estimates in the most recent years if age 0-2 fish were captured by the survey or the fishery.

With the exception of age 0-2 individuals, the same data used in the 2016 accepted model (McGilliard et al. 2016) were used in the SS3 model runs, as listed in the table below.

Source	**Data**	**Species Included**	**Years**
NMFS Aleutian Islands Groundfish Trawl Survey	Survey biomass (linear regression used to combine BS shelf survey estimates with AI survey estimates for a single survey biomass index)	Flathead only; no Bering flounder were caught in the Aleutian Islands	1980, 1983, 1986, 1991-2000 (triennial), 2002-2006 (biennial), 2010-2016 (biennial)
NMFS Bering Sea Shelf Groundfish Survey (standard survey area only[1])	Survey biomass (linear regression used to combine BS shelf survey estimates with AI survey estimates for a single survey biomass index)	Flathead sole and Bering flounder combined	1982-2016
	Age Composition	Flathead sole only	1982, 1985, 1992-1995, 2000-2015
	Length Composition	Flathead sole only	1983, 1984, 1986-1991, 1996-1999, 2016
U.S. trawl fisheries	Catch (Bering Sea and Aleutian Islands; pelagic and non-pelagic trawl[2])	Flathead sole and Bering flounder combined	1977-2016
	Age Composition (Bering Sea only; non-pelagic trawl only)	Flathead sole only	1994, 1995, 1998, 2000, 2001, 2004-2007, 2009-2015
	Length Composition (Bering Sea only; non-pelagic trawl only)	Flathead sole only	1977-1993, 1994, 1996-1997, 1999, 2002-2003, 2008, 2016

1. Excludes survey strata 70, 81, 82, 90, 140, 150, and 160
2. A very small amount of catch is taken with hook and line and is included in the total catch biomass

Description of differences between the 2016 model and the SS3 framework

There are fundamental differences between the 2016 accepted model and the SS3 modeling framework and that make it impossible to configure a fully matching model using SS3. The table below lists the differences, whether the way that each of these factors is modeled is effective within the old model code, and whether SS3 offers a more effective way to model each of these factors than does the old model code. While the phrase “the 2016 model” refers to the combination of the old model code and one particular configuration of the input files for running the code that was used to produce the 2016 accepted model, “the old model code” refers to any run configuration of the input files (any model runs) that could be created using the old model’s ADMB code.

Source of Difference between Models	Is this a positive attribute of the 2016 model?	Would SS3 do a better job of modeling this attribute?
2016 model starts at age 3, SS3 model starts at age 0	No. Data from ages 0-2 cannot be used in the 2016 model (or the old model code in general)	Yes, SS3 would allow for data from ages 0-2 to be used, potentially informing selectivity curves at low ages.
Length-based, logistic survey selectivity estimates do not match	No. The survey selectivity in the 2016 model was problematic, as the curve estimated was shallow, never reaching 0, and only reaching 1 at very old ages, which was not believable. In addition, there was a strong retrospective pattern in survey selectivity parameter estimates in the 2016 model.	Yes, SS3 offers a large suite of options for modeling selectivity, including sex-specific, age-based, double-normal, and non-parametric options. In addition, SS3 starts at age 0, and can fit to data for ages 0-2, which would help to define the selectivity curves at low ages.
Growth models, including age-length transition matrices both use the von-Bertalanffy growth curve, but cannot be matched exactly between the 2016 model and an SS3 model. There are small differences in modeling growth, leading to differences in all biomass estimates between the models, even when numbers-at-age match exactly.	Neutral. However, the old model code doesn't include as many options as SS3 to fine-tune the calculations involving growth to be specific to the timing of events in the model.	Yes, SS3 would allow for use of data on the distribution of ages within each length bin, which would, in turn, allow for estimation of growth parameters, including the CV of the youngest and oldest fish, which defines the age-length transition matrix. Likewise, a weight-at-age vector can be input for each year of the model, if desired, to account for time-varying weight-at-age and to avoid defining relationships with length altogether.

Historical mean recruitment can be modeled in both frameworks (the old model code and SS3), but when it is estimated or fixed to a different value than main-period mean recruitment, the model behavior in initial years is different in the old model framework than in SS3. This occurs because a recruit is an age 3 individual in the old model code and an age 0 individual in SS3. Therefore, in the initial model year, age 3s in SS3 were born three years earlier under the assumption of historical mean recruitment, while age 3s in the old model recruited in the initial year of the main period recruitment regime. Also, SS3 models a likelihood penalty for historical mean recruitment values to prevent it from being too high, while the old model code estimates historical mean recruitment without upper or lower bounds or penalty.	There is not a specific problem with the methods for modeling historical recruitment in the old model code, although parameter estimates may be unrealistic because the parameter unbounded. The most recent CIE review identified the estimates of F and recruitment in the initial model years as an area of substantial uncertainty, needing further investigation.	Yes, SS3 would allow for estimation of early recruitment deviations, which would be informed by the length and age composition data on fish that were recruits in this early period. This may allow the model to better distinguish between historical recruitment and historical fishing mortality. Using either modeling framework, foreign reported data on BSAI flathead sole exists back to 1964 and could be investigated for use in better informing initial conditions as well.
Timing of population dynamics is different between the model frameworks. The old model code specifies that the survey occurred mid-year by using mid-year weight-at-length and -age, but applies continuous mortality for the entire year to numbers-at-age when calculating survey biomass and predicted length and age distributions. The spawning month is specified as March in the 2016 model, but spawning biomass is specified using beginning-of-the-year weight-at-age.	No. It would be more accurate if half of the year's natural and fishing mortality were applied when calculating mid-year survey biomass, length, and age distributions.	Yes, the timing of population dynamics in SS3 can be modeled accurately, according to user-specified inputs.

Modeling of the relationship between temperature and catchability cannot follow the exact same equation in SS3 as in the old model code (though it can still be modeled in SS3)	No, the estimate of the temperature-catchability relationship in the 2016 model is close to 0 (no relationship), which leads to a similar fit to survey biomass as for the same model without this relationship estimated, but includes lots of tiny deviations from the trend line that have little meaning.	Possibly. SS3 offers the ability to link an environmental index to catchability (or to any other parameter) through a multiplicative or additive relationship. This is a different equation than the one in the old model code.
Predicted survey and catch biomass are based on mean numbers-at-length and catch-at-length, which means that the numbers-at-age were converted to numbers-at-length by multiplying by the age-length transition matrix (which differs between models). The predicted catch biomass in SS3 is based on numbers-at-age and catch-at-age.	Not clear. The underlying population dynamics in the old model code are age-structured, and numbers-at-age are multiplied by mean weight-at-age to calculate predicted biomass proportions, rather than translating the information through the age-length transition matrix to numbers-at-length, and multiplying by the weight-length relationship. *Note that CRM coded up and used standard predicted survey and catch biomass based on numbers-at-age and catch-at-age for several runs of the old model framework in this exercise to facilitate comparisons with SS3 runs.*	The likelihood components used in SS3 are standard for statistical catch-at-age models. In addition, because the age-length transition matrix is calculated or estimated internally, it can be adjusted to reflect the timing of the calculations for which it is used.

Steps to understanding differences between model frameworks to construct a model in SS3 that best matched the 2016 model

Given that SS3 could not be configured to fully match the 2016 accepted model, an age-based selectivity option and age-based survey biomass and catch biomass prediction options were added to the old model code so that a few variants of the 2016 model could be configured to better match equivalent configurations in SS3. The main goal of adding to the old model code was to start with a configuration of the old model code and SS3 where population dynamics and likelihood equations matched exactly, and from there, to demonstrate the irreconcilable differences in the two model frameworks one piece at a time. Once differences are demonstrated, an SS3 configuration that best matches the 2016 model is shown and compared to the 2016 model. Below is a list of old and SS3 model configurations that are compared to demonstrate where the models match or fail to match, and the corresponding figures. The results section below leads the reader through each of the figures and the reasons for matching or non-matching results in each step.

Old model variant	SS3 configuration	Corresponding Figure(s)
(old_a) As for 2016 model, but with age-based selectivity and	(SS3_a) A model with age-based selectivity and all other	Figure 1 and Figure 2

all parameters fixed, except for fishing mortality. Also, predicted catch and survey biomass were calculated from numbers-at-age (not numbers-at-length)	parameters fixed to the same values as for the old model code variant, with only fishing mortality estimated. Growth schedules (rather than parameters) are specified to match those of old_a	
(old_b) as for old_a	(SS3_b) As for SS3_a, but with growth specified as parameters (rather than schedules), to match those of old_b as closely as possible.	Figure 3 and Figure 4
(old_c) as for old_a, but with historical R fixed to its MLE from the 2016 model	(SS3_c) As for SS3_a, but with R1 (historical R in SS3 terms) fixed to an equivalent value to that in old model variant old_c.	Figure 5
(old_d) as for old_a, but with mean recruitment and recruitment deviations estimated (in addition to estimating fishing mortality, as was done in old_a)	(SS3_d) As for SS3_b, but with R0 (mean recruitment in SS3 terms) estimated and recruitment deviations estimated (in addition to estimating fishing mortality, as was done in SS3_b)	Figure 6
(old_e) as for old_d, but with the age-based selectivity curves estimated, rather than being fixed.	(SS3_e) As for SS3_d, but with the age-based selectivity curves estimated instead of being fixed.	Figure 7
(old_f) as for old_d, but with length-based selectivity curves fixed.	(SS3_f) As for SS3_d, but with length-based selectivity curves fixed to match those of old_f (which were chosen such that no fish under age 3 would be selected in SS3)	Figure 8
(old_g) as for old_f, but with the length-based selectivity parameters estimated instead of being fixed	(SS3_g) As for SS3_f, but with the length-based selectivity parameters estimated instead of being fixed.	Figure 9
(old_h) as for old_g, but with catch and survey biomass calculated from numbers-at-length and length-based selectivity (which requires the age-length transition matrix)	(SS3_h) As for SS3_g	Figure 10
(old_i) as for old_g (identical to the 2016 model, except that catch and survey biomass are calculated based on numbers-at-age and selectivity-at-age)	(SS3_i) As for SS3_g and SS3_h, except that survey selectivity (which is estimated) is an age-based, double-normal curve without a descending limb (forced to be asymptotic) and catchability is fixed to be 0.7	Figure 11

	(0.7 is the maximum derived age-based survey selectivity from the 2016 model's survey selectivity curve). Therefore, this SS3 configuration acts to mimic the 2016 model.	
2016 model	(SS3_j) as for SS3_i.	Figure 12-Figure 20

The following sections offer more details on the population and observation models within both the old model framework and SS3.

Description of population and observation models within both modeling frameworks

Mean recruitment and historical recruitment
Several steps were taken to build an SS3 model with population dynamics that best matched those of the 2016 model using configurations of the old model code and SS3 that were deterministic (Figure 1-Figure 6). First, the relationship between the log of mean recruitment estimated in the 2016 model $(\ln(\bar{R}))$ and the log of R_0 (unfished recruitment $(\ln(R_0))$ that is estimated in SS3 was determined (Equation 1), where M is natural mortality. Note that $\ln(R_0)$ in SS3 operates as the log of mean recruitment when no stock-recruitment relationship is assumed.

(1) $$\ln(R_0) = \ln(\bar{R}) + \ln(1000) + 3M$$

The $\ln(\bar{R})$ estimated in the 2016 model refers to total mean recruitment of age 3 individuals (males and females), while $\ln(R_0)$ refers to total recruitment of age 0 individuals in thousands. Both models are able to estimate a separate historical mean recruitment parameter. The 2016 model estimates an $\ln(R_{hist})$ parameter, which translates to the SS3 parameter, $\ln(R_{offset})$, as follows:

(2) $$\ln(R_{offset}) = \ln(R_{hist}) + \ln(1000) + 3M - \ln(R_0)$$

One difference between the two models is that $\ln(R_{hist})$ is an unbounded parameter in the old model, while $\ln(R_{offset})$ is a bounded parameter with user-specified bounds and, in addition, there is a likelihood component associated with $\ln(R_{offset})$ in SS3 to prevent it from becoming too large. In the section below entitled "Initial Conditions," a timing mismatch in the application of historical mean recruitment is explained.

Both models assume a 1:1 sex ratio.

Initial Conditions
Initial conditions are identical between model frameworks when historical mean recruitment is specified to be equal to the main-period mean recruitment $(\ln(\bar{R}) = \ln(R_{hist}))$ and historical fishing mortality is set equal to 0 in old model variants and in SS3. When historical fishing mortality is added to the model, the models are identical as long as selectivity below age 3 in SS3 is equal to 0. When historical mean recruitment differs from the main period mean recruitment, numbers-at-age 3 differ for the first three years of the model (1977-1979) because the historical mean recruitment is applied to age 0 fish in 1977-1979 in SS3 and to age 3 fish in the old model. Hence, in 1977-1979, SS numbers-at-age 3 reflect the

historical-period's mean recruitment, while the old model's numbers-at-age 3 reflect the main period mean recruitment.

Growth
The old model framework used externally estimated maturity-at-age and weight-at-age schedules. Weight-at-age at the beginning of the year and mid-year were both specified as vector inputs to the model. The maturity and weight-at-age schedules can be input into SS3 to be identical between the old model and SS3 with a setting in SS3 to bypass specifying (or estimating) growth parameters (Figure 1). However, the old model code reads in an externally-calculated age-length transition matrix, while SS3 must internally calculate the age-length transition matrix based on the parameters specified for the von-Bertalanffy growth curve, the allometric length-weight relationship, and the CVs in length-at-age of the youngest and oldest age classes modeled. The age-length transition matrix is used in both model frameworks to translate length-based selectivity into age-based selectivity for application in calculating the numbers-at-age, and to calculate predicted survey and fishery length distributions. When growth parameters are specified in SS3 (instead of these schedules being input as vectors), small differences arise between models. Most runs in this matching exercise specify growth parameters (rather than vectors/schedules) in order to match the 2016 model's specification of length-based selectivity (to allow SS3 to calculate the age-length transition matrix required when using length-based selectivity). However, a few comparisons were done where growth schedules were specified (along with age-specific selectivity) to achieve a match between model frameworks in deterministic population dynamics. This approach allowed for some other aspect of the frameworks (such as initial conditions) to be compared without the confounding influence of slightly different growth curve estimates and age-length transition matrices between the old model framework and SS3.

Selectivity
The 2016 model assumed length-based logistic selectivity curves for fishery and survey selectivity, with the same selectivity curves for males and females. Although selectivity was configured in the same way in SS3, estimates of survey selectivity were different between the two models for several reasons. First, the data used only by SS3 on ages 0-2, and the modeling of ages 0-2 informs SS3's selectivity curves. Second, the length-based selectivity curves were converted into age-based selectivity within the models before use in the numbers-at-age equations by multiplying by the age-length transition matrix, and the age-length transition matrices were a mismatch between models.

In the interest of distinguishing the effects of various differences between the models, some SS3 and old model variant runs were done using age-based, logistic selectivity curves, configured such that selectivity below age 3 was zero. To show the effects of the different age-length transition matrices on the calculation of derived age-based selectivity from length-based selectivity curves, some comparisons were done using length-based selectivity, fixed to the same values in both model frameworks, and set to 0 at lengths that could be associated with fish under the age of 3. In addition, some SS3 runs used double normal, age-based, sex-specific survey selectivity curves with selectivity below age 3 set to 0 and restricting the curves from becoming dome-shaped. The rationale for this approach is described in the results section, and the SS3 model that best matched the dynamics of the 2016 configuration of the old model used this double-normal age-based and sex-specific survey selectivity.

The fishery selectivity curves estimated in SS3 and the old model were similar to each other.

Biomass and Timing
SSB and survey biomass were shown to be very similar in deterministic model variant comparisons when selectivity curves were identical (and equal to 0 below age 3) and the empirical values for beginning and

mid-year weight-at-age were used, along with age-based selectivity (which does not require the age-length transition matrix to be used to calculate population dynamics; Figure 2). SSB and survey biomass differ slightly between otherwise comparable deterministic model variants when growth parameters are internally specified in SS3 (Figure 4). Specifying growth internally in SS3 is a requirement when using length-based selectivity curves because SS3 calculates the age-length transition matrix internally. It is then used to translate length-based selectivity into age-based selectivity to apply selectivity to numbers-at-age and to translate predicted numbers-at-age into predicted proportions-at-length to fit to the length composition data. Therefore, this slight difference in biomass estimates between the 2016 model and SS3 is unavoidable because the 2016 model used length-based selectivity. In addition, in the old model code's spawning biomass was calculated based on numbers-at-age in March, but using January weight-at-age. The survey biomass in the 2016 model was calculated using mid-year weight-at-age and assuming mortality throughout the year. Several model runs were done specifying spawning in old model variants as occurring at the beginning of the year to minimize differences in spawning biomass.

Recruitment Deviations

Recruitment deviations in the old model were estimated from the first to last model year (1977-2016) and applied to age 3 recruits. Recruitment deviations in SS3 were matched to the 2016 model by estimating age 0 recruits beginning in 1974 until 2013.

Yearly and Historical Fishing Mortality

Yearly apical fishing mortality and average historical (pre-1977) fishing mortality were estimated in both model frameworks. The population dynamics associated with fishing mortality were identical in the two models. Estimates of initial numbers-at-age 3 will differ between models when historical F and fishery selectivity for ages 0-2 are both greater than 0. This occurs because SS3's age 3 fish were subject to fishing mortality in the historical period, whereas the old model's age 3 fish are considered new recruits and were not modeled in the historical period, nor subject to the historical F.

Stock-Recruitment

The 2016 model and SS3 estimated recruits as mean-unbiased recruitment deviations from an estimated mean value with a σ_R set to 0.5. The 2016 model estimated recruitment at age 3 and SS3 estimated recruitment at age 0. Numbers-at-age-3 were compared between the 2016 model and SS3 and were similar for most configurations of the two models.

Temperature-catchability parameter

The old model estimated a parameter relating summer bottom temperature to catchability. This relationship was omitted from model runs comparing the old model variants to equivalent SS3 configurations because it was estimated to have a negligible effect on the population dynamics and fit of the 2016 model. Though visually it appears that there may be a relationship between survey biomass and temperature, the relationship included in the old model was ineffective and should not be a barrier to moving to a different assessment framework. SS3 has the ability to estimate a multiplicative or additive linkage between an environmental index and any model parameter (including catchability). In addition, SS3 can estimate time-varying parameters, such as time-varying catchability.

Likelihood Components

Table 2 lists the equations for each likelihood component used in SS3 and in the 2016 model (and variants).

Catch biomass
The 2016 model translated numbers-at-age into numbers-at-length by multiplying by the age-length transition matrix as an input to the Baranov catch equation in terms of length, multiplied by the weight-length relationship, and summed over length bins to calculate the predicted catch and survey biomass. SS3 uses numbers-at-age in the Baranov catch equation to calculate catch-at-age, multiplies by the weight-at-age relationship, and sums ages to calculate predicted catch biomass. However, the age-length transition matrices used for these processes are mismatched between the 2016 model and SS3, leading to differences in predicted catch biomass for the same fishing mortality estimates.

In the interest of distinguishing the role of the of the mismatched age-length transition matrix between the 2016 model and SS3 from other model differences, an alternative calculation of catch biomass based on numbers-at-age and mid-year weight-at-age was incorporated into the old model code as an option for use (maintaining the original catch biomass equation as an option as well).

Survey biomass

Predicted survey biomass is calculated similarly in both model frameworks, except that SS3 assumes that the survey occurs mid-year (which it does), while the 2016 model uses continuous, year-round mortality (calculated at the end of the year) along with the mid-year weight-length relationship in calculations. As for catch biomass, the 2016 model predicts survey biomass based on numbers-at-length (as calculated from numbers-at-age and the age-length transition matrix), and we know that the age-length transition matrix is a mismatch between the 2016 model and SS3.

As for the predicted catch biomass, in the interest of distinguishing the role of the mismatch in predicted survey biomass between models, an additional option to calculate survey biomass based on numbers-at-age and weight-at-age (leaving the age-length transition matrix out of the calculations) was added to the old model and used in some old model variants to compare to SS3 model configurations.

Age- and length-composition likelihood components
The age- and length-composition likelihood components in SS3 are identical to those in the 2016 model. However, as noted above, the observations of survey proportions-at-age and proportions-at-length differ among models in that the data given to SS3 includes the data given to the old model code in addition to the proportions of age 0-2 fish and lengths below 6 cm.

Recruitment

The 2016 model and SS3 estimate recruitment deviations that are constrained by the σ_R value specified. The likelihood components are slightly different (Table 2), but the two model frameworks estimate similar recruitment patterns under a number of alternative model configurations.

Alternative SS3 model configuration to consider for the 2018 assessment

Two alternative models in SS3 are proposed to address some of the shortcomings of the 2016 model and the best matching SS3 model. Changes to the model are:

(1) Foreign reported catches were included in the data from 1964-1987 (Table 3). The most recent CIE review feedback listed the initial conditions of the 2016 model as uncertain and in need of investigation. The 2016 model estimated historical recruitment to be 58 million recruits, and main period mean recruitment to be 835 million recruits. Given the data on foreign reported catches

that was not considered in the 2016 assessment (and even without that data), it is hard to believe that mean recruitment in the historical period may have been that low.

(2) The recruitment likelihood function used a sum-to-zero constraint
(3) Recruitment was fixed to its mean value for the last 4 model years due to lack of non-zero observations of young fish.
(4) Recruitment deviations were estimated dating back to 1961.
(5) Survey selectivity was changed to be age-based and sex-specific, using a double-normal selectivity curve. Derived age-based selectivity from the length-based curves estimated in the variant SS3_g indicated that age-based selectivity reached 1 – the model estimated that even small, old fish were fully selected by the survey, and therefore survey selectivity can easily be estimated as length- or age-based using a curve with an asymptote at 1. Use of age-based survey selectivity avoids the need to translate selectivity through the age-length transition matrix before being applied to numbers-at-age.
(6) A data weighting scheme developed by Francis, 2011 was used.

The above list describes Model 18.0.

An alternative model, Model 18.0b, is as for Model 18.0, but estimated separate fishery selectivity curves for each of 3 distinct time periods: 1964-1988, 1989-2007, and 2008-2016. These time blocks represent major change-points in the management of the flatfish trawl fishery.

Additional model runs that are not presented

In addition, an SS3 model was configured with an additive (log-based) relationship between summer bottom temperature and catchability (as for the old model), but (as for the 2016 model), the relationship lead to only very small adjustments in the survey biomass estimates. This model was not considered further.

A model was considered to estimate a male offset parameter from female natural mortality in a configuration where fishery length-based selectivity was estimated for males and females together. This was done because it is thought that natural mortality for males is different for that of females for many flatfish populations and the sex-specific selectivity curves in Models 2018.0 and 2018.0b fit the comp data by estimating that males recruit to the fishery at smaller sizes than do females. This seems unlikely unless there is some sex-specific spatial behavior driving this pattern and it was thought to be more likely that male natural mortality is different from female natural mortality, but the model estimates male natural mortality to be almost exactly the same as female natural mortality, so this model is not shown.

Results

Transition of Old Model into an Equivalent SS3 Model

To explore the effects of each of the non-matching factors listed in the table above, a variant of the 2016 accepted model was formulated as described in the section above entitled “Steps to understanding differences between model frameworks” to provide a starting place where the old model variant matched an SS3 model as closely as possible without investing a substantial amount of time in re-coding the old model framework. We start with an old model variant and an SS3 model configuration like the 2016 accepted model, but without historical F or historical catches, and with historical mean recruitment fixed to be equal to mean recruitment, and age-based logistic selectivity curves fixed and equal to zero below age 3. In addition, maturity-at-age and weight-at-age schedules (rather than parameters) were specified in both models, and the models were run deterministically with recruitment deviations fixed to 0, estimating only yearly fishing mortality. Differences in the biomass likelihood components were eliminated by adding an option to the code of the old model to use predicted catch biomass based on catch-at-age, rather than catch-at-length, eliminating the use of the age-length transition matrix from the calculation of the catch biomass from numbers-at-age. This is a standard method for calculating the predicted catch for the catch biomass likelihood component in statistical catch-at-age models. Maturity, weight-at-age, age-based selectivity, mean recruitment, numbers at age 3 (which are new recruits in the old model), and initial numbers-at-age in the absence of historical F and historical mean recruitment can all be matched exactly between the models (Figure 1 and Figure 2). The 2016 model used length-based survey and fishery selectivity. For the model variants shown in Figure 2, it was necessary to add age-based, sex-specific selectivity to the old model code to achieve a close match between model frameworks. This is the case because the age-length transition matrix differed between the 2016 model and SS3 and use of length-based selectivity requires that the age-length transition matrix be used to convert the length-based selectivity into age-based selectivity for calculation of numbers-at-age. Growth parameters (rather than schedules) need to be specified to calculate or estimate an age-length transition matrix within SS3 (SS3 lacks an option to specify an age-length transition matrix that is calculated externally); this leads to small mismatches in the weight-at-age schedules, but maturity-at-age remains exactly the same between models (Figure 3: growth as specified in the 2016 model and within SS3 using parameters rather than schedules). In addition, the calculation methods for the age-length transition matrix differ slightly from those input to the 2016 model as well, as described above. Hence, moving from the near-perfect match of models shown in Figure 2, incorporating the growth estimates from Figure 3 into the SS3 model and specifying identical length-based selectivity curves in both model frameworks leads to small mismatches in biomass quantities from the two model runs, even in a run with only fishing mortality estimated, and using standard catch-at-age likelihood equations (Figure 4).

Figure 5 shows model configurations identical to that shown in Figure 2 (where population dynamics between the two model runs matched almost exactly), except that here, the parameter in each model run determining historical recruitment was fixed to the old model's 2016 estimate (54 million age 3 recruits), which was very low relative to the main period recruitment estimate (834 million age 3 recruits). The numbers-at-age 3 in 1977-1979 are dramatically different between models in Figure 5. This occurs because SS3 estimates age 0 recruits and the old model framework estimates age 3 recruits. Therefore, in 1977, the age 3 individuals from SS3 recruited in 1974 under low historical mean recruitment, while the age 3 individuals from the old model variant recruited in 1977 under much higher main-period mean recruitment. The historical period effectively ends three years earlier in the old model framework, which can be seen in the plots of yearly fishing mortality, spawning stock biomass, and survey biomass. This timing mismatch disappears later in the time series after many years in which mean recruitment was high.

A comparison of model variants like those in Figure 2, but with growth specified internally, and mean recruitment and recruitment deviations estimated is shown in Figure 6. Fishing mortality, spawning stock biomass, and survey biomass match almost exactly between these models, and estimates of numbers-at-age 3 are very similar between models. Both models estimate very large numbers-at-age 3 in the last 2 years, which is likely unrealistic and uninformed by the data. While the old model framework is currently hard-wired to estimate recruitment through the end year that is modeled, SS3 options include specifying the first and last year in which recruitment deviations are estimated, and the user can choose to fix the recruitment in the last years of the model to mean recruitment for the purpose of projections.

Figure 7 shows a comparison of model configurations that are like those shown in Figure 6, except that the models estimate age-based selectivity (as well as mean recruitment, recruitment deviations, and fishing mortality). Here, fishery selectivity is very similar between the two models, but survey selectivity has a much more shallow slope in the old model than in SS3. SS3 estimates a selectivity of 0 at the lowest ages, which is informed by data indicating zero (or very low) catches of age 0-2 fish in the survey; these data cannot be included in the old model. In the old model, the youngest age modeled is age 3 and it is assumed that the lowest age bin of data (for age 3s) only includes age 3 fish and not age 0-3 fish.

Figure 8 shows a model comparison like that in Figure 6 where mean recruitment, recruitment deviations, and fishing mortality were estimated and age-based selectivity was fixed, but here the survey and fishery selectivity are length-based (as in the 2016 configuration of the old model) and fixed. As in Figure 6, the selectivity parameters were chosen such that selectivity below age 3 would be equal to 0 to minimize differences due to the different age-at-recruitment between models. Given that the selectivity curves match exactly between models, it is the influence of the mismatch between the age-length transition matrices used in the two models that leads to a scale difference in biomass estimates between models (for both spawning and survey biomass) and small mismatches in yearly estimates of fishing mortality. The estimates of numbers-at-age are generally similar in most years.

Adding another layer of the 2016 model configuration back in to the comparison, Figure 9 shows the same comparison as for Figure 8, but with the length-based selectivity estimated instead of being fixed. As for the model runs where age-based selectivity parameters were estimated (Figure 7), the fishery selectivity curves estimated were very similar between the two models. Likewise, the old model estimated a shallow survey selectivity slope that does not reach 0 at small lengths. Again, the old model is informed only by data from age 3+ individuals, while the SS3 model is informed by observations of 0 (or few) individuals of age 0-2. Essentially, the shallow slope estimate in the old model is an estimate of catchability that is below 1, and this shows up as a scale mismatch in spawning biomass, where the spawning biomass of the old model is inflated because the old model estimates that large, mature fish occur in the population that aren't selected by the survey. The SS3 model survey selectivity indicates that the large, mature fish are fully selected.

The scale mismatch in spawning biomass between models becomes larger when, in addition to differences in survey selectivity estimates, and growth, the old model's length-based biomass likelihood equations are used in the objective function (Figure 10). The length-based biomass likelihood equations in the old model make use of the mismatching age-length transition matrix to calculate predicted fishery and survey catches in biomass, whereas the SS3 model calculates these quantities using the weight-at-age schedule (which is a better match between models than the age-length transition matrix). To minimize the objective function such that catch biomass matches the data, a different scale of yearly fishing mortalities must be estimated by the old model than by SS3, and other parameter estimates between models must then differ in order to fit predicted proportions-at-age to the age composition data in both models, given the differences in Fs.

The SS3 model that best mimicked the 2016 model

It is possible to configure SS3 into a model that better matches the 2016 configuration of the old model than does the SS3 model shown in Figure 10 with historical mean recruitment and historical fishing mortality added. Two main sources of mismatch between the models are (1) the age at recruitment and (2) the age-length transition matrix used to convert length-based selectivity to age-based selectivity for calculation of numbers-at-age. This conversion to age-based selectivity within the model meant that an equivalent model could be configured using sex-specific, age-based selectivity, and this was done using the following approach:

1. Double-normal survey selectivity was specified, fixing the parameter defining the descending limb of the curve to a large value such that selectivity was asymptotic (as selectivity in the old model was asymptotic)

2. Selectivity below age 3 was set to 0 (which is an option that can be specified when specifying the double-normal selectivity curve in SS3).
3. The old model's age-length transition matrix was used to convert the old model's estimated length-based survey selectivity to age-based survey selectivity and the selectivity at which an asymptote (maximum age-based selectivity) occurred was noted. If selectivity is below 1, this means that some old individuals are never subject to being caught by the survey because they never grew large enough to be selected. In function, an age-based selectivity with an asymptote below 1 lowers catchability. In the 2016 model, the asymptote occurs around age 7 and the average age-based selectivity between ages 7 and 21 was 0.7. An assumption of age-based selectivity reaching an asymptote at 1 with a catchability of 0.7 is equivalent to an assumptions of age-based selectivity with an asymptote at 0.7 and a catchability of 1.
4. The SS3 model was configured with catchability fixed at 0.7 (instead of 1) and an age-based, sex-specific, double-normal, asymptotic survey selectivity curve (the descending limb was fixed to a large value).

A model comparison between the SS3 model configured using this approach and the 2016 model configuration was then done, replacing SS3's length-based survey selectivity curve with this age-based curve. Here, estimation of historical fishing mortality and historical mean recruitment were included, and standard catch-at-age likelihood equations in the old model. The comparison is shown in Figure 11, using standard catch-at-age likelihood questions in the old model. The estimation of historical recruitment leads to a similar mismatch of fishing mortality and numbers-at-age in the initial years of the model to that seen in Figure 5 (a deterministic model with historical recruitment fixed to the value estimated in the 2016 old model configuration). Aside from these initial years, the fishing mortality, numbers-at-age, survey biomass, and fishery selectivity are all similar for the two models. The survey selectivity is somewhat similar when remembering that catchability in the SS3 model is set at 0.7, so one can imagine the SS3 survey selectivity shifted downwards on the y-axis to a maximum value of 0.7. The spawning biomass is a little bit smaller in SS3 than in the old model and this is due to the difference in shape between the estimates of selectivity where the asymptote occurs at an earlier age in SS3 than in the old model.

Finally, Figure 12 shows the same comparison as for Figure 11, but using the length-based likelihood equations that were used in the 2016 configuration of the old model such that this is a true comparison to the 2016 accepted model. The mismatch between models is more evident when using the length-based likelihood equations in the old model. The yearly fishing mortality rates are similar between models, but the equations for the predicted catch biomass and survey biomass yield different results between models. To match these values to the catch biomass and survey biomass data, the models must adjust a parameter or parameters that have an influence on the scale of the biomass, such as mean recruitment and/or recruitment deviations. The plot of numbers-at-age-3 shows a difference in scale that is then amplified by multiplying by weight-at-age to calculate the spawning biomass. Figure 13-Figure 20 show a comparison between the SS3 model and the 2016 model of fishery and survey proportions-at-length and proportions-at-age. The difference in the models caused by estimating historical mean recruitment for a model with recruits at age 3 vs age 0 can be seen in fits to age- and length-compositions in the initial model years, but in general, the proportions-at-age/-length are very similar for the two models.

Alternative SS3 models for the 2018 assessment

Differences from the SS3 2016 model match and the two 2018 models

There is a scale difference in spawning biomass between the 2016 best matched model and Models 2018.0 and 2018.0b because the 2016 model match fixed catchability to 0.7 (Figure 21-Figure 23). This catchability specification was a hack to match SS3 to the 2016 model that is not necessary moving forward because SS3 (making use of data on age 0-2s) estimates a derived age-based selectivity curve that reaches an asymptote at 1 (the 2016 model with the old code estimated this asymptote at 0.7, which prompted this "hack" of setting catchability in SS3 to 0.7 to best mimic the old model). In addition, the old model did not include foreign-reported catch data, and estimated simple deviations, which led to slightly different estimates of recruitment from the 2018.0-series models.

Comparing Models 2018.0 and 2018.0b
Models 2018.0 and 2018.0b estimate nearly identical spawning biomass and survey biomass (Figure 21 and Figure 22). The estimation of different selectivity curves in different time periods led to differences in the scale of fishing mortality in the early years of the model, with model 2018.0b estimating lower fishing mortality in the 1st and 3rd time blocks (prior to 1988 and after 2007; Figure 23-Figure 25, Table 6). Selectivity during the middle time block is similar between models (Table 6, Figure 25).

When sex-specific selectivity is estimated (as it is in Models 2018.0 and 2018.0b), the resulting selectivity curves indicate that males recruit to the fishery at smaller sizes than do females. This could occur if smaller females were for some reason less vulnerable to the fishery due to spatial or some other behavior. However, it is likely that some other difference between males and females is not represented by these models. Growth and/or variation in growth may be mis-specified, as it was estimated outside of the model (and therefore unable to account for the effect of selectivity on length-at-age samples, for instance). It may also be the case that growth has varied over time or space in a different way for males than for females. A model run was done with selectivity for males and females estimated together, and where male natural mortality was estimated; this model run estimated male natural mortality to be almost identical to female selectivity. Further research is needed to understand why smaller males are found in the length composition data as compared to females. Although selectivity may not be a full or accurate explanation for these differences in the data between males and females, it leads to substantially improved fits to age-composition data over the best matching model from 2016 (Figure 26).

Model 2018.0 and 2018.0b show similar fits to age composition data, in aggregate over years (Figure 26). Model 2018.0 showed very poor fits to fishery length composition data during the early years of the model (prior to the mid-80s; Figure 30), which was part of the motivation to implement Model 2018.0b, which estimates a separate selectivity curve for this early time period and for the most recent years after the groundfish fishery was rationalized in 2008. Figure 26 and Figure 34 shows that fits to fishery length data were improved by including these time blocks for selectivity. Detailed plots of yearly fits to survey and fishery age composition data are shown in Figure 28-Figure 31 for Model 2018.0 and in Figure 32-Figure 35 for Model 2018.0b.

Table 4 shows likelihood components for the three models. Most likelihood components (except for the survey biomass likelihood component) cannot be compared between the 2016 matching model and the other two models because a different data weighting scheme was implemented in Models 2018.0 and 2018.0b. The survey likelihood components for Models 2018.0 and 0b were lower than for the 2016 SS3 matching model. In addition, the overall likelihood and all components were slightly lower (better) for Model 2018.0b than for Model 2018.0 (but were within a similar range of one another).

Table 6 shows the survey and fishery selectivity parameter estimates for each model. The parameter estimates for fishery selectivity had higher standard deviations for the earliest time block (1964-1977) when data were more sparse. Only catches were available for 1964-1976, and only catches and length-composition data were available until the year 2000.

Literature Cited

Methot, R. D. and C. R. Wetzel. 2013. Stock synthesis: A biological and statistical framework for fish stock assessment and fishery management. Fisheries Research **142**:86-99.

McGilliard, C.R., Nichol, D., and Palsson, W. 2016. 9. Assessment of the Flathead Sole-Bering Flounder Stock in the Bering Sea and Aleutian Islands. In Stock Assessment and Fishery Evaluation Report for the Groundfish Resources of the Bering Sea and Aleutian Islands. pp. 1229-1318. North Pacific Fishery Management Council, P.O. Box 103136, Anchorage AK 99510.

Tables

Table 1. Symbols used in this document.

Symbol	Meaning
x	sex
a	age
l	length
f	fleet (fishery or survey)
t	time
$S_{f,x,a}$	Selectivity for fleet *f*, sex *x*, and age *a*
$N_{t,x,a}$, $\boldsymbol{N_{t,x}}$	Numbers at age *a*, time *t*, and sex *x*, and vector of numbers-at-age
$w_{a,x}$, w_l	Mid-year weight at age *a* for sex *x* and weight at length *l*
$Z_{t,x,a}$	Total mortality at age *a*, sex *x*, and time *t*
$I_{t,f}$	Observed survey biomass at time *t* for fleet *f*
$SB_{t,f}$	Predicted survey biomass at time *t* for fleet *f*
$\sigma_{survey,t,f}$	Standard deviation of observed survey biomass at time *t* for fleet *f*
$n_{t,x,f}$	Number of age-composition observations at time *t* for sex *x* and fleet *f*
$p_{t,x,f,a}$	Observed proportion at age *a*, time *t*, fleet *f*, and sex *x*
$\hat{p}_{t,x,f,a}$	Predicted proportion at age *a*, time *t*, fleet *f*, and sex *x*
$n_{2,t,x,f}$	Number of length-composition observations at time *t* for sex *x* and fleet *f*
$p_{t,x,f,l}$	Observed proportion at length *l*, time *t*, fleet *f*, and sex *x*
$\hat{p}_{t,x,f,l}$	Predicted proportion at length *l*, time *t*, fleet *f*, and sex *x*
$\tilde{R}_t$	Estimated mean recruitment in year *t*
σ_R	Recruitment CV
b_t	Bias adjustment factor at time *t* (specified in SS3 only)
C_t^{obs}	Observed catch at time *t*
$\hat{C}_t$	Predicted catch at time *t*
$\sigma_{t,f}$	Standard error of catch at time *t* for fleet *f* (specified for SS3 only)
$\boldsymbol{\varphi_x}$	Age-length transition matrix (rows are ages, columns are lengths)

Table 2. Likelihood components used in the old model and SS3 model. Numbers in the component column are likelihood component weightings for: (SS3, old model).

Component	SS3	Old Model
Survey biomass $(SB_{t,f})$ equation	$\sum_{x,l} w_l \sum_a S_{f,x,l} S_{f,x,a} \varphi_{t,x,a,l} N_{t,x,a} e^{-0.5(Z_{t,x,a})}$	$N_{t,x,l} = \boldsymbol{N_{t,x}\varphi_x}$; $\sum_{x,l} S_{f,x,l} N_{t,x,l} w_l \frac{1-(e^{-(S_{f,x,l}F+M)})}{(S_{f,x,l}F+M)}$ Alt Option: Age-based
Survey biomass likelihood (1,1)	$\sum_{t \in survey} \frac{\left(\ln(I_{t,f}) - \ln(SB_{t,f})\right)^2}{2\sigma^2_{survey}}$	As for SS3
Age composition (0.93, 0.93) survey (0.52, 0.52) fishery	$\sum_t \sum_x \sum_a n_{t,x,f}\, p_{t,x,f,a} \ln\left(\frac{p_{t,x,f,a}}{\hat{p}_{t,x,f,a}}\right)$	As for SS3
Length Composition (0.42, 0.42)	$\sum_t \sum_x \sum_l n_{2,t,x,f}\, p_{t,x,f,l} \ln\left(\frac{p_{t,x,f,l}}{\hat{p}_{t,x,f,l}}\right)$	As for SS3
Recruitment (1,1)	$\frac{1}{2}\left(\sum_{t=1977}^{2016} \frac{\tilde{R}_t^2}{\sigma_R^2} + b_t \ln(\sigma_R^2)\right)$ (sum to 0 constraint possible, but not used)	$\frac{1}{2\sigma_R^2} \sum_{t=1977}^{2016} \left(\ln(N_{t,1}) - \ln(\tilde{R}_t\ ) + \frac{\sigma_R^2}{2}\right) +$ $n_t \ln(\ \sigma_R\ )$, (sum to 0 constraint possible, but not used)
Catch biomass ($\hat{C}_t$) equation	$\sum_{a,x} \left(\frac{S_{a,x}\, F_t}{(S_{a,x}\, F_t + M)}\ N_{t,x,a}(1 - e^{-(S_{a,f}\, F_t)}) w_{x,a}\right)$	$N_{t,x,l} = \boldsymbol{N_{t,x}\varphi_x}$; $\sum_{l,x} \left(\frac{S_{f,l}\, F_t}{(S_{f,l}\, F_t + M)}\ N_{t,x,l}(1 - e^{-(S_{f,l}\, F_t)}) w_{x,l}\right)$ Alt option: as for SS3
Catch (50,50)	$\sum_t \frac{(\ln(C_t^{obs}) - \ln(\hat{C}_t))^2}{2\sigma^2_{t,f}}$	$\sum_t (\ln(C_t^{obs}) - \ln(\hat{C}_t))^2$

Table 3. Catch biomass as used in the 2016 model and in the alternative 2018 SS3 model

Year	Catch in 2018 SS3 Model	Catch in 2016 Model	Year	Catch in 2018 SS3 Model	Catch in 2016 Model
Initial Catch	11659		1993	13574	13574
1964	12315		1994	17006	17006
1965	3449		1995	14715	14715
1966	5086		1996	17346	17346
1967	11218		1997	20683	20683
1968	12606		1998	24387	24387
1969	9610		1999	18573	18573
1970	21050		2000	20441	20441
1971	26108		2001	17811	17811
1972	10380		2002	15575	15575
1973	17715		2003	13785	13785
1974	13198		2004	17398	17398
1975	5011		2005	16108	16108
1976	7565		2006	17981	17981
1977	7909	7909	2007	18958	18958
1978	13864	6957	2008	24540	24540
1979	6042	4351	2009	19558	19558
1980	8600	5247	2010	20128	20128
1981	10609	5218	2011	13559	13559
1982	8417	4509	2012	11367	11367
1983	5518	5240	2013	17355	17354
1984	4458	4458	2014	16512	16512
1985	5636	5636	2015	11307	11307
1986	5208	5208	2016	8321	8321
1987	3595	3595			
1988	6783	6783			
1989	3604	3604			
1990	20245	20245			
1991	14197	14197			
1992	14407	14407			

Table 4. Likelihood component values for the 2016 SS best matching model and for Models 2018.0 and 2018.0b. Only the 2016 SS3 best matching model survey biomass likelihood component can be compared to Models 2018.0 and 0b because the data weighting approach was different in 2016 SS3 best matching model than for the other two models.

Likelihood Component	2016 SS3 best match	Model 2018.0	Model 2018.0b
TOTAL	1,748	456	435
Survey	-21.66	-35.62	-36.36
Length_comp	1,020	159	147
Age_comp	605	312	307
Recruitment	128.991	20.694	16.959

Table 5. Parameter estimates for key scale-related parameters for the SS3 best model match, and Model 2018.0 and 0b. Historical mean recruitment = exp(log mean recruitment + log historical recruitment offset). Initial fishing mortality and historical recruitment was applied in 1977 for the 2016 best match and in 1964 for Models 2018.0 and 2018.0b.

	2016 SS3 Best Match		**Model 2018.0**		**Model 2018.0b**	
Parameter	**Est**	**sd**	**Est**	**sd**	**Est**	**sd**
log mean recruitment	13.25	0.07	13.63	0.03	13.64	0.03
log catchability	-0.36	fixed	0	fixed	0	fixed
Initial fishing mortality	0.03	0.00	0.07	0.01	0.03	0.00
log historical recruitment offset	-1.08	0.08	0	fixed	0	fixed

Table 6. Selectivity parameter estimates for Models 2018.0 and 2018.0b.

	Model 2018.0, 1964-2016		Model 2018.0b, 2008-2016		Model 2018.0b, 1988-2007		Model 2018.0b 1964-1987	
	Est	StDev	Est	StDev	Est	StDev	Est	StDev
Size at 50% selectivity (f)	39.379	1.084	37.196	1.091	39.429	1.522	28.312	3.522
Slope (f)	7.994	0.870	4.339	1.449	7.050	1.346	9.147	3.688
Size at 50% selectivity (m) (offset from f)	-3.608	0.595	-2.833	0.709	-3.985	0.821	-0.952	3.359
Slope (m) (offset from f)	-1.059	0.939	0.547	1.602	-0.873	1.355	-1.231	4.718
Peak: beginning age for the plateau (f)	7.162	0.312	7.208	0.303	7.208	0.303	7.208	0.303
Ascending width (f; ln)	2.324	0.132	2.338	0.127	2.338	0.127	2.338	0.127
Male peak offset	-0.602	0.319	-0.606	0.315	-0.606	0.315	-0.606	0.315
Male ascending width offset (ln)	-0.188	0.152	-0.188	0.149	-0.188	0.149	-0.188	0.149

Figures

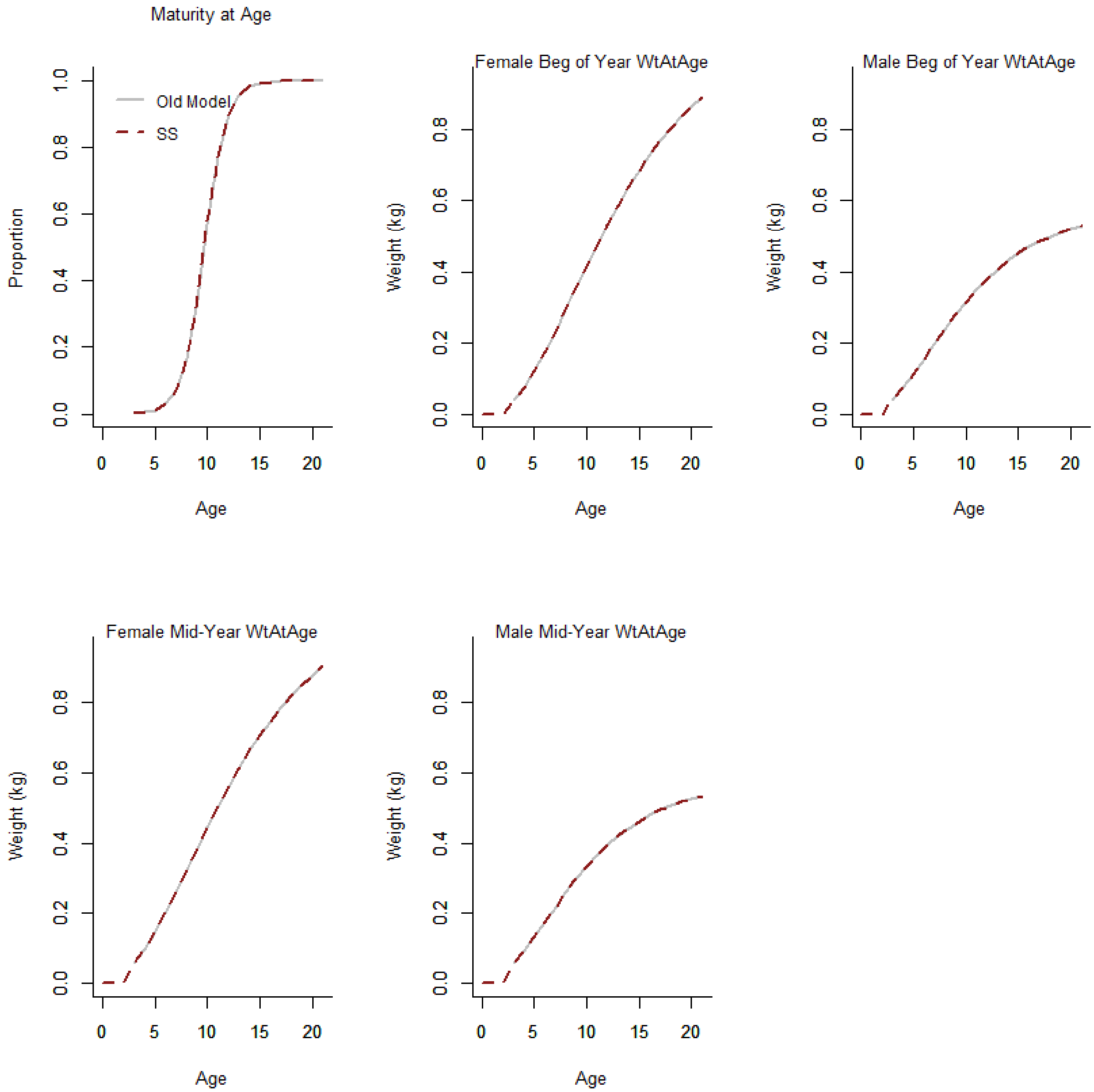


Figure 1. Maturity and weight-at-age schedules for the old model (grey lines) and SS3 (dotted red lines) when specified as vectors input to the SS3 model. This method of specifying growth in SS3 cannot be used if length-based selectivity is also used because it does not allow the user to specify an age-length transition matrix.

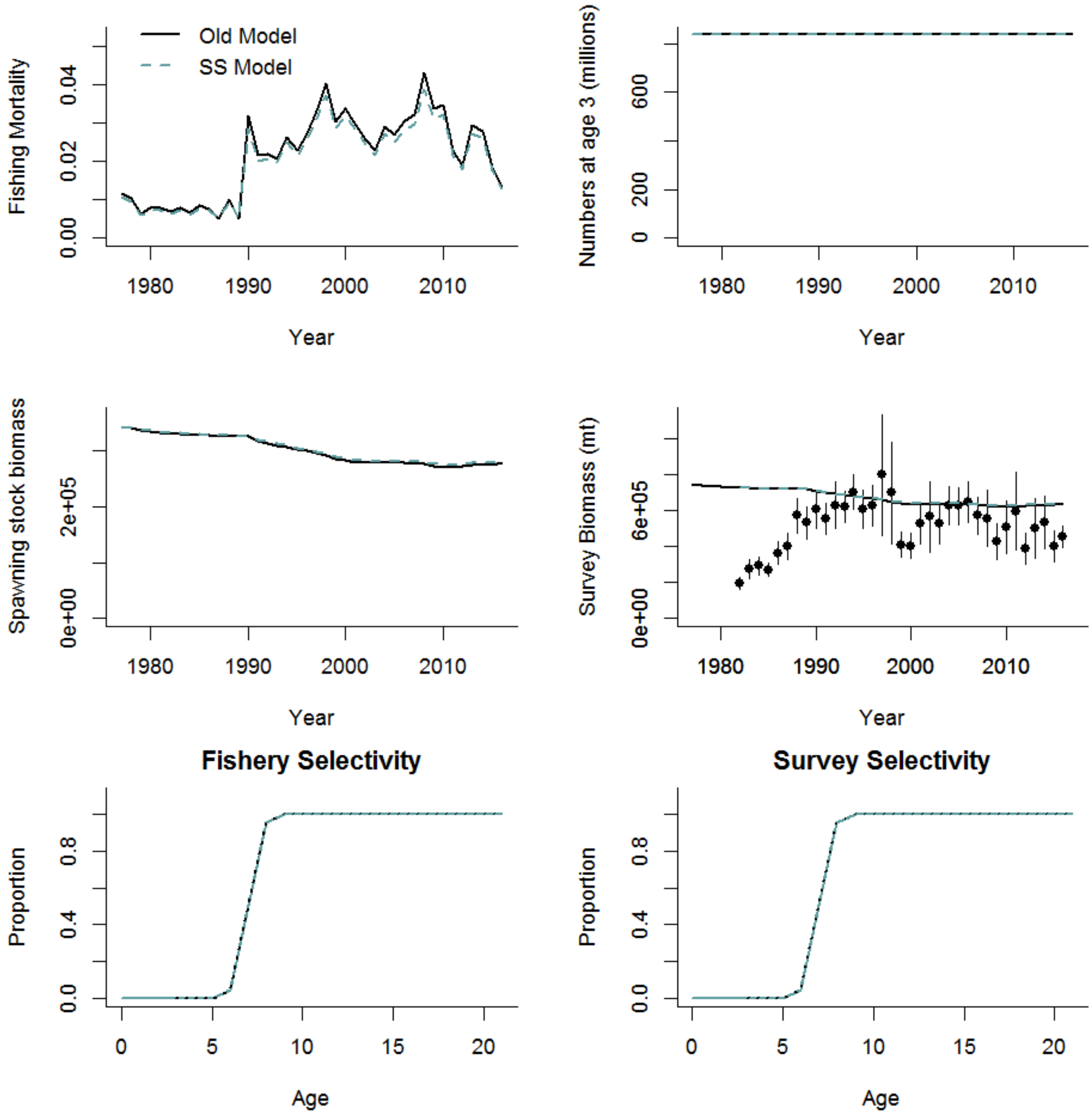


Figure 2. Population dynamics of the old model and SS3 in a deterministic run with fishing mortality estimated and all other parameters fixed. Empirical maturity and weight-at-age schedules were input to SS3 and not specified internally. Age-based selectivity was modeled and was chosen such that selectivity at young ages was equal to 0. Biomass likelihood equations were age-based.

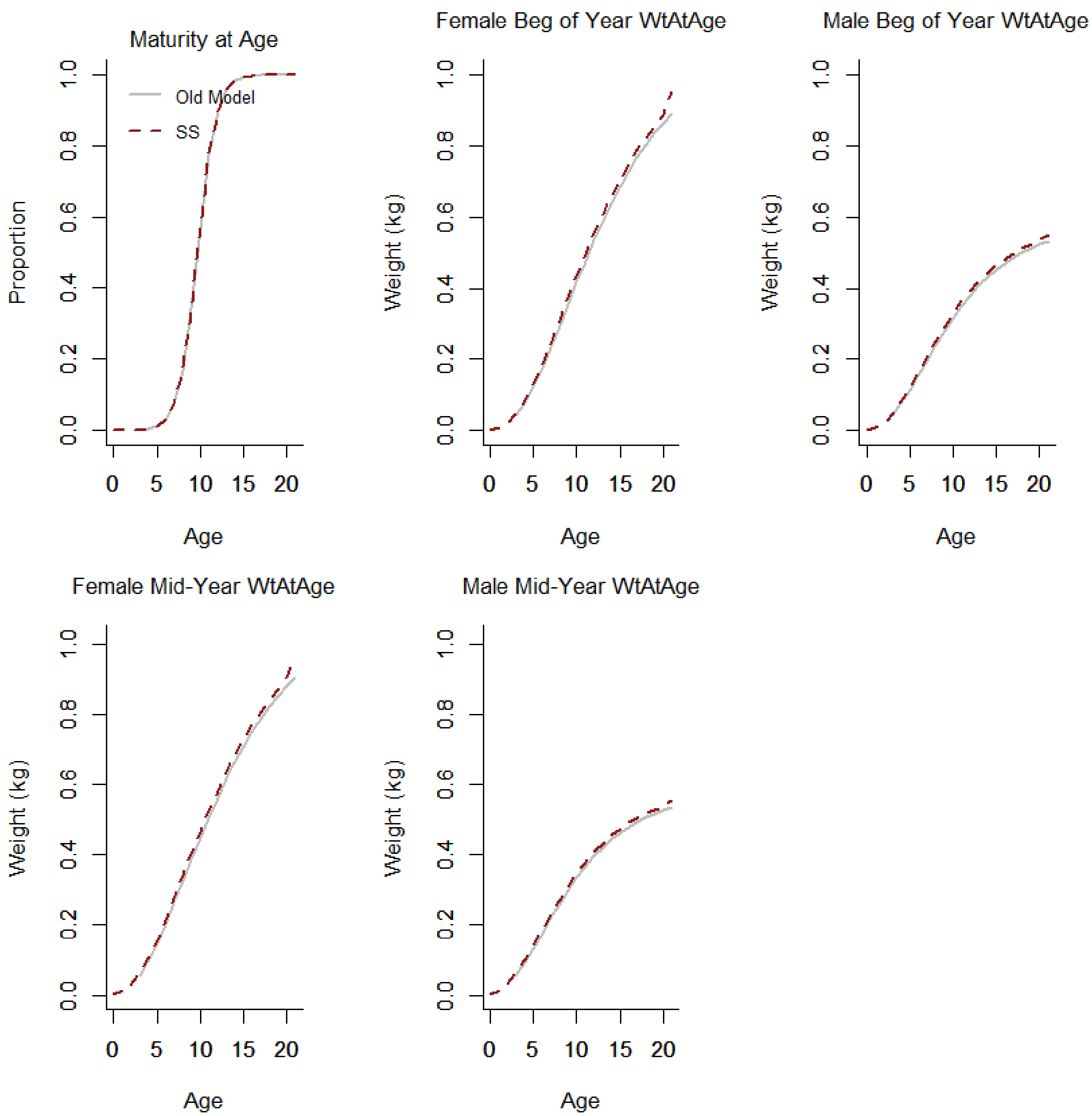


Figure 3. Maturity and weight-at-age schedules for the old model (grey lines) and SS3 (dotted red lines) when specified internally in the SS3 model. This method of specifying growth must be used when length-based selectivity is used because it provides SS3 with the information necessary to populate an age-length transition matrix internally, which is needed to apply length-based selectivity to numbers-at-age and to calculate predicted proportions-at-length from predicted numbers-at-age.

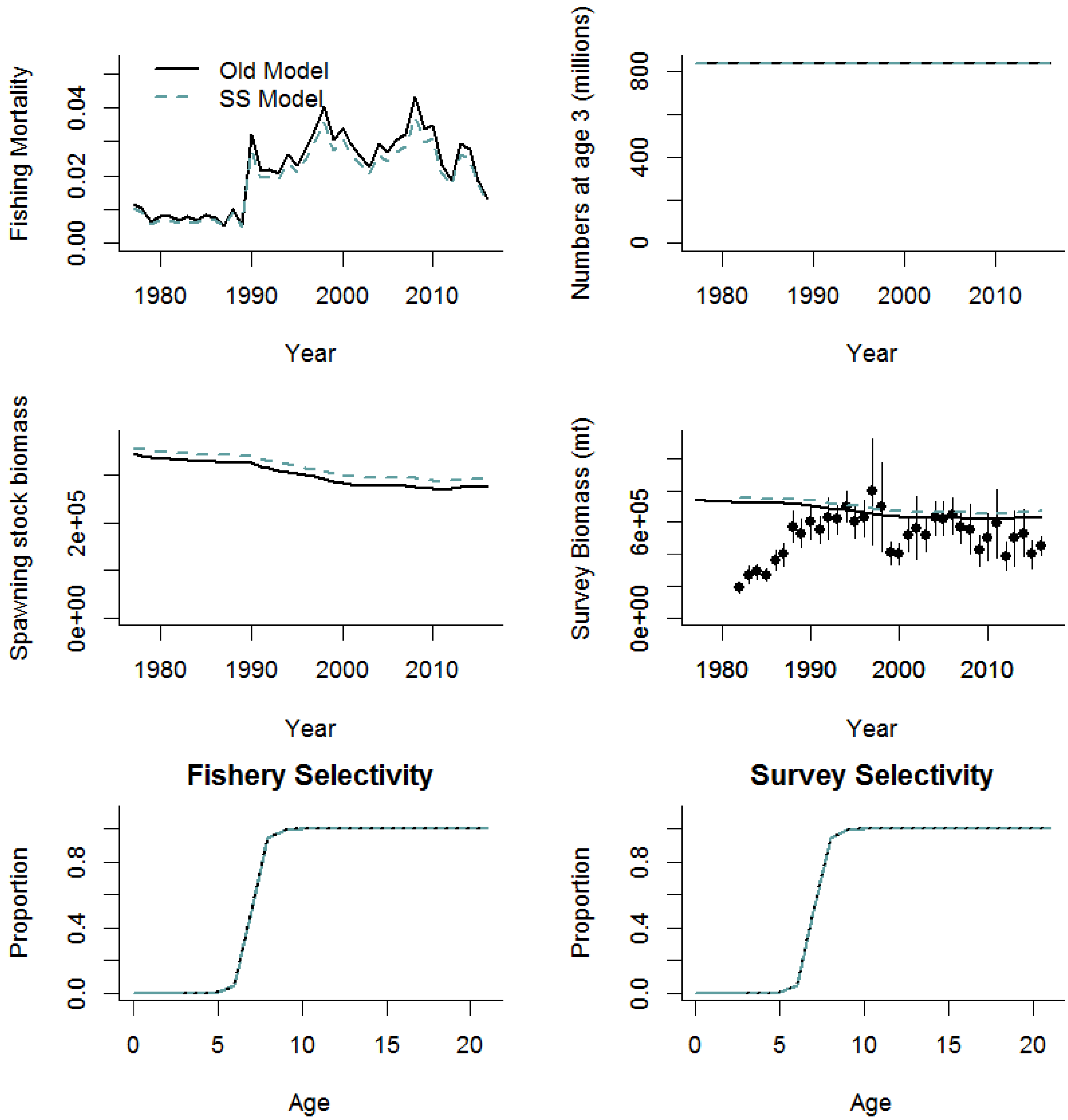


Figure 4. Population dynamics of the old model and SS3 in a deterministic run with fishing mortality estimated and all other parameters fixed. Maturity and weight-at-age schedules were specified internally in SS3. Age-based selectivity was modeled and was chosen such that selectivity at young ages was equal to 0. Biomass likelihood equations were age-based.

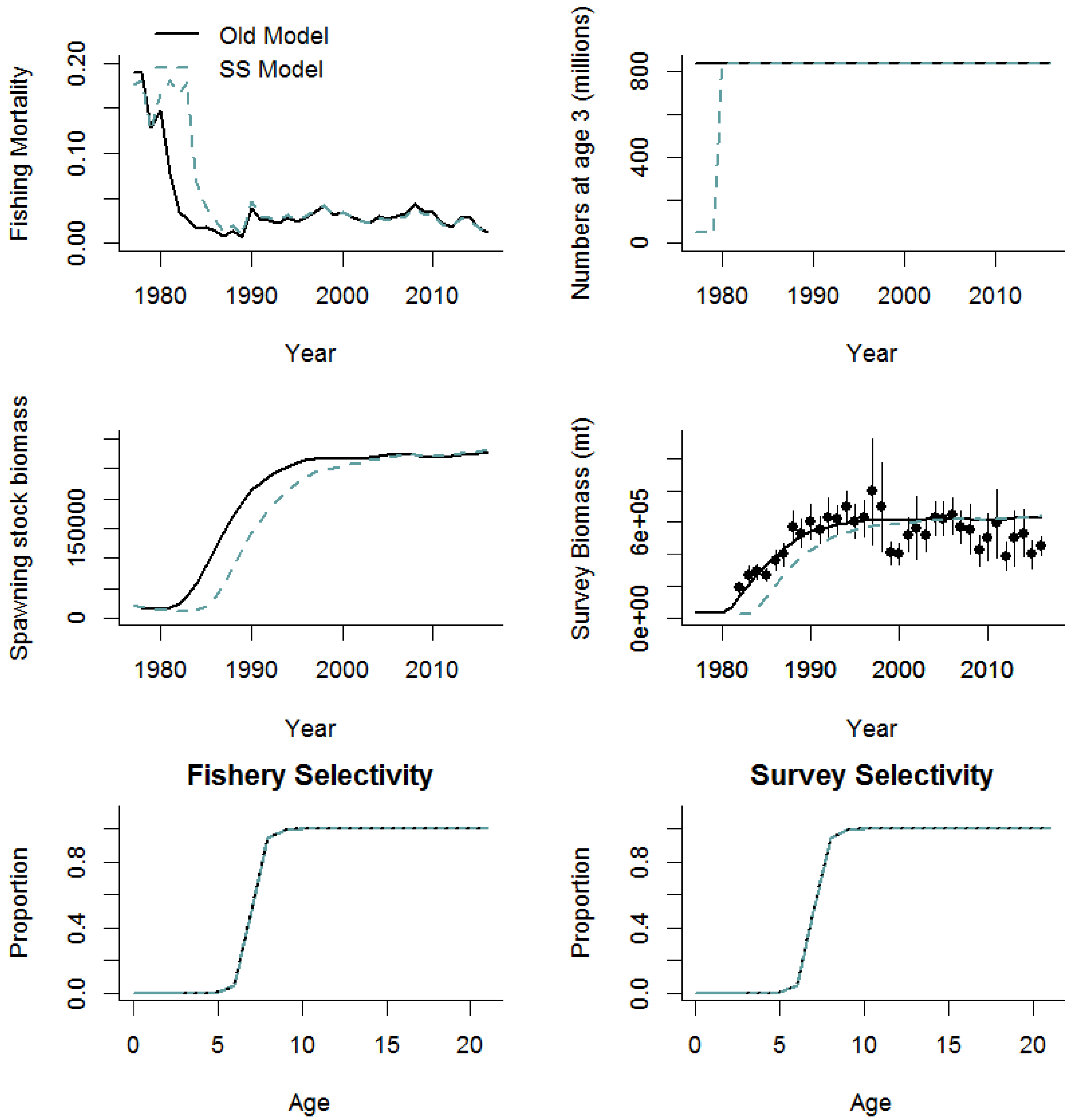


Figure 5. Population dynamics of the old model and SS3 in a deterministic run with historical R fixed and equal to 55 million age 3 recruits (mean recruitment from 1977 onwards in the model is equal to 835 million recruits). Fishing mortality was estimated and all other parameters were fixed. Maturity and weight-at-age schedules were specified empirically in SS3 to match exactly between models. Age-based selectivity was modeled and was chosen such that selectivity at young ages was equal to 0. Biomass likelihood equations were age-based.

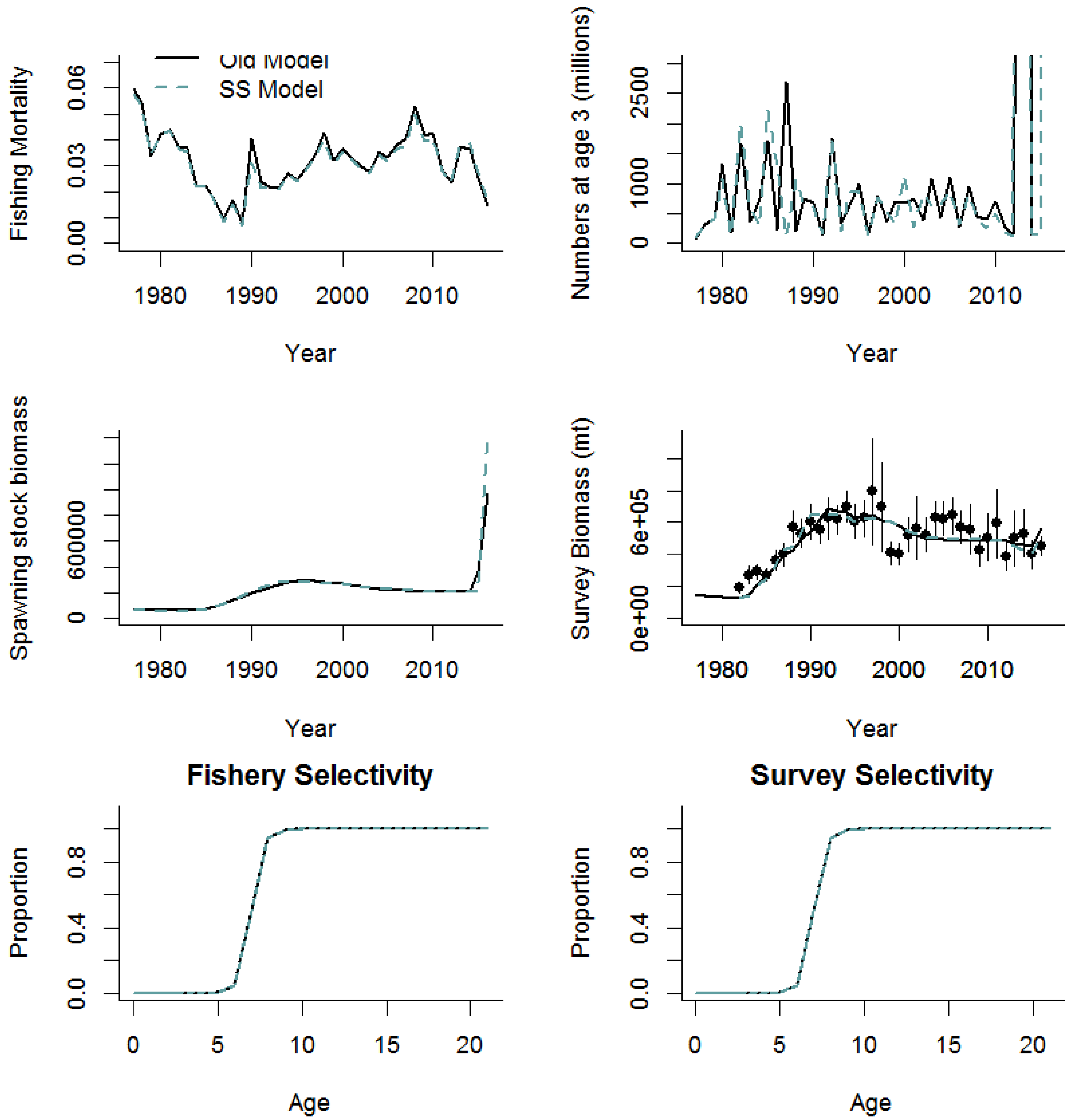


Figure 6: Population dynamics of the old model and SS3 in a run with mean log recruitment (R0 in SS3) and recruitment deviations estimated, as well as fishing mortality. Maturity and weight-at-age schedules were specified internally in SS3. Age-based selectivity was modeled and was chosen such that selectivity at young ages was equal to 0. Biomass likelihood equations were age-based.

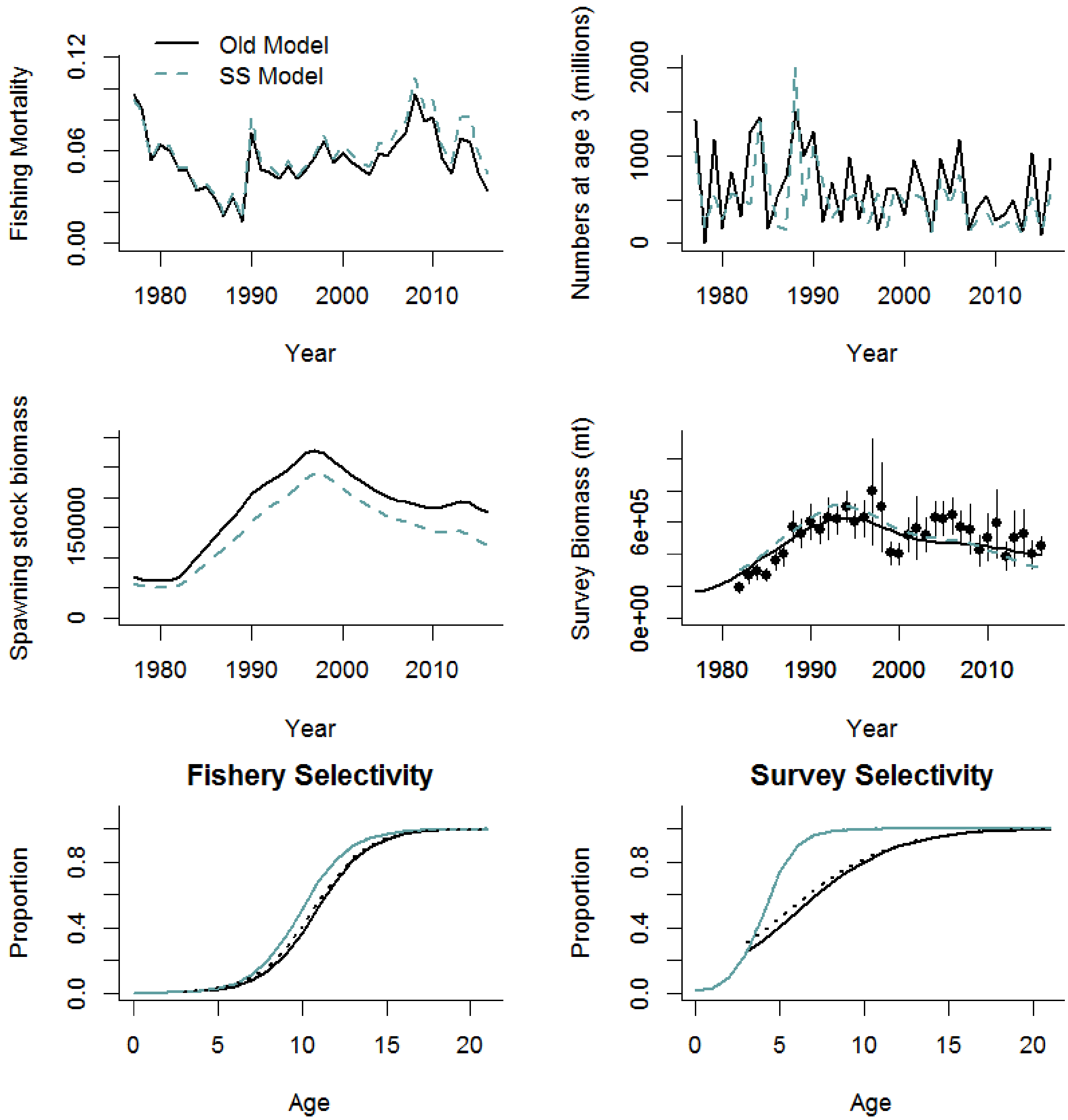


Figure 7. As for Figure 6, except that age-based selectivity is estimated instead of being fixed.

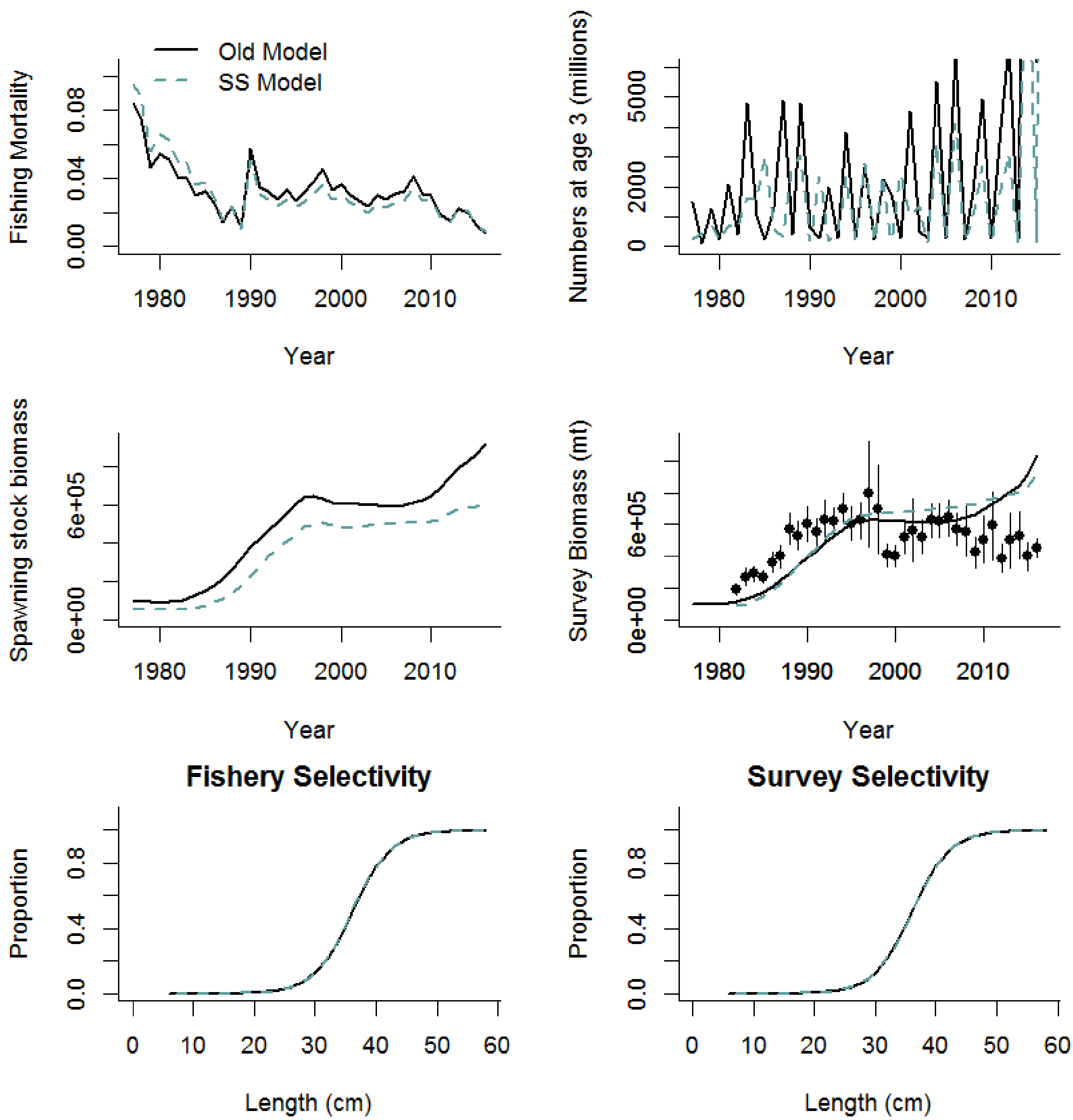


Figure 8: As for Figure 6, but length-based selectivity was modeled (and fixed) and was chosen such that selectivity at young ages was equal to 0.

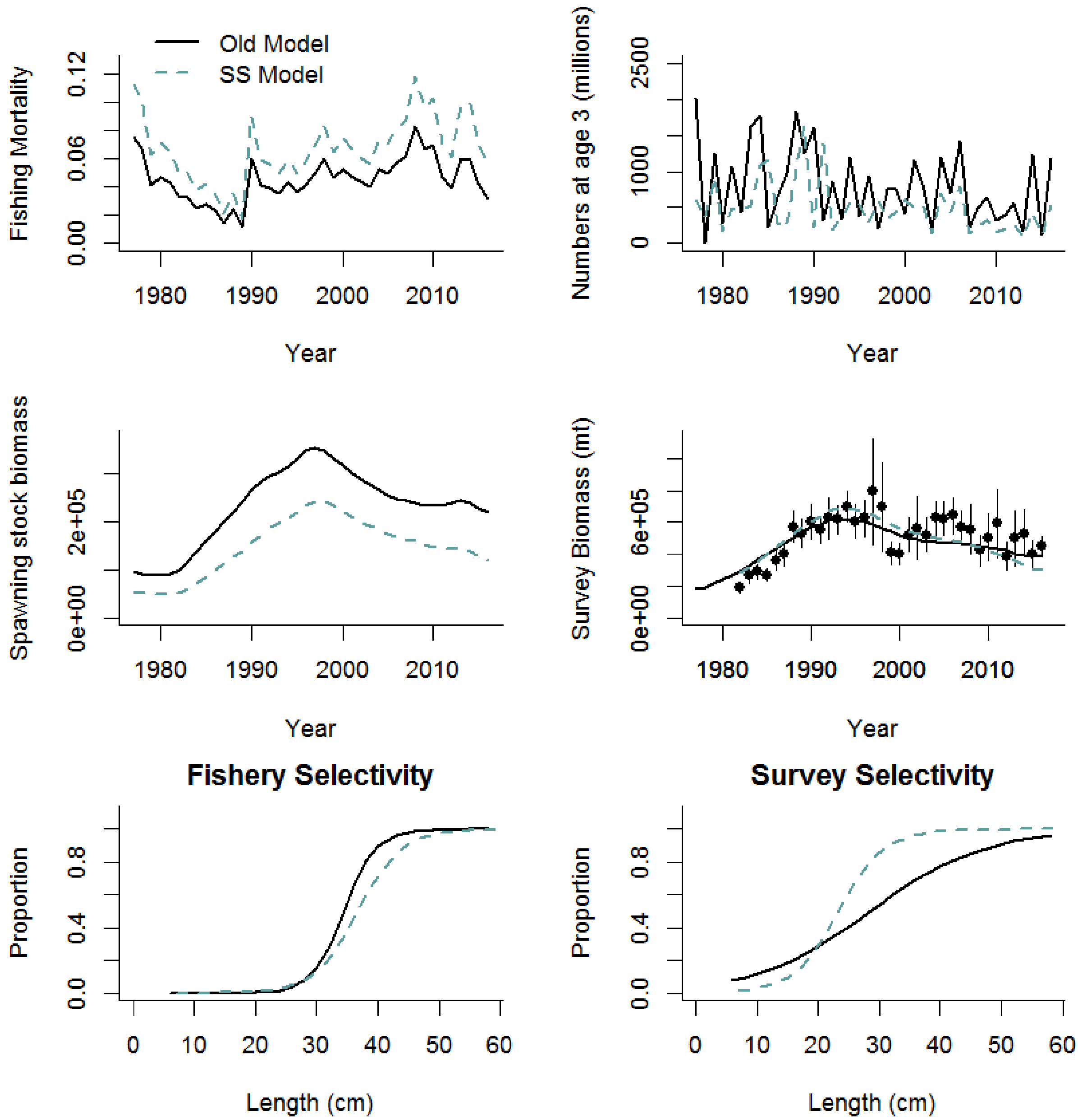


Figure 9: As for Figure 8, except that the length-based selectivity was estimated and not fixed. The length-based selectivity requires the use the age-length transition matrix to convert length-based selectivity to age-based selectivity, which is required to calculate numbers-at-age. The age-length transition matrix is not an exact match between models, and the result is a difference in the scale of F between models. The survey selectivity in the old model is not informed with data for ages 0-2. This leads to differences in estimates of selectivity, which leads to a bigger mismatch in spawning stock biomass between the two models.

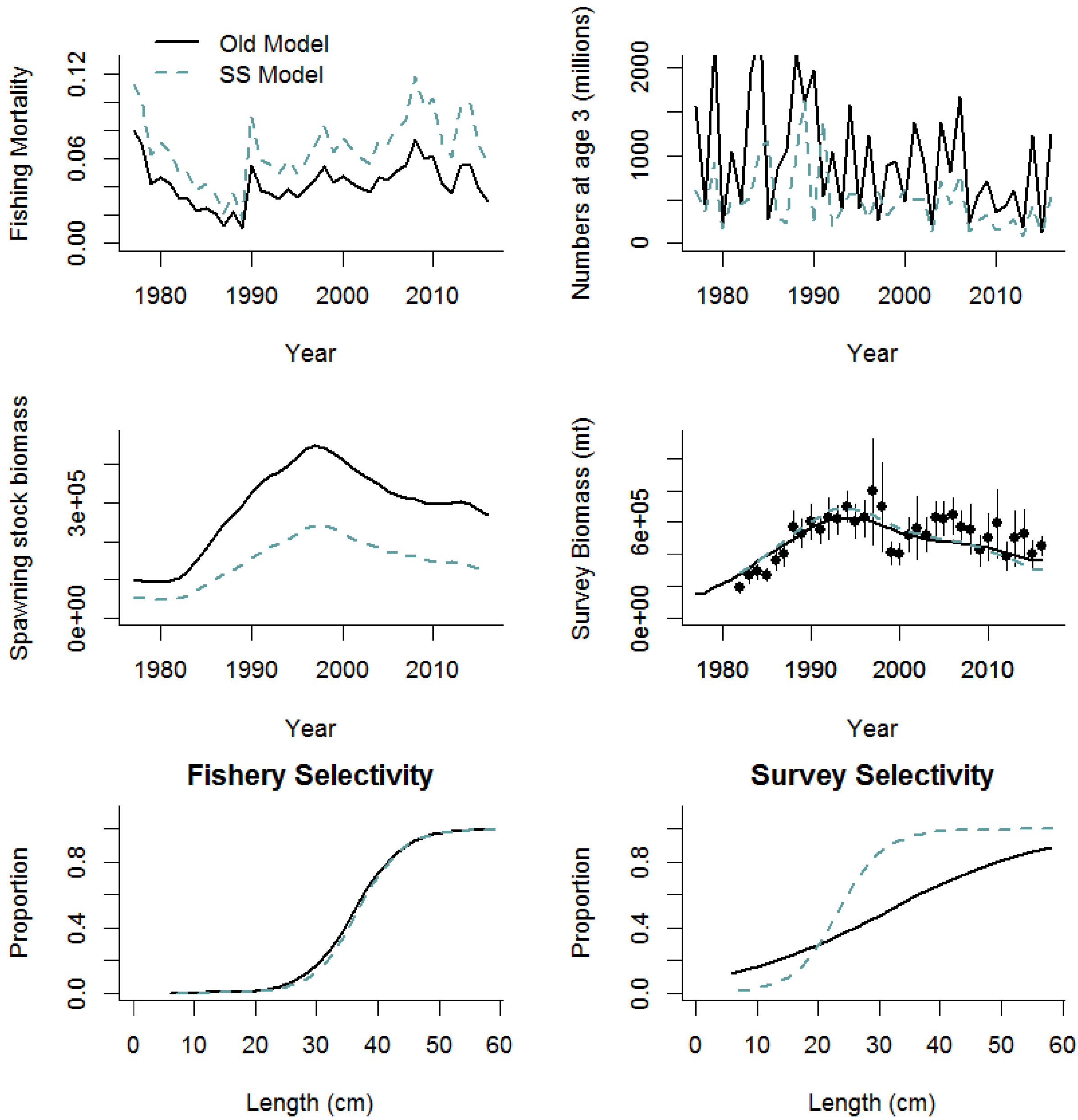


Figure 10. As for Figure 9, except that the old model uses length-based likelihood formulations, calculating catch and survey biomass from numbers-at-length and length-based selectivity (which requires the age-length transition matrix).

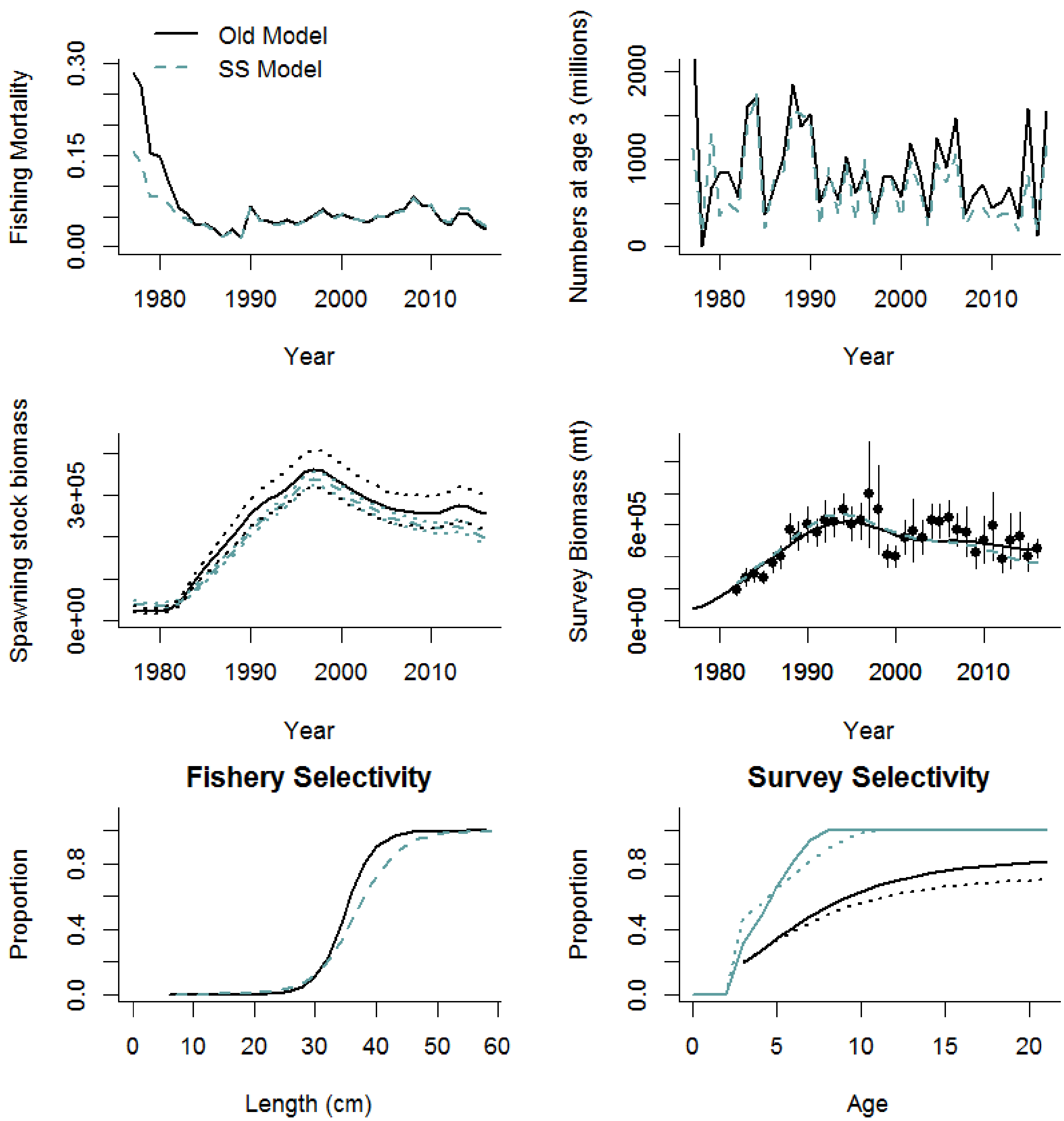


Figure 11. An SS3 model that uses double-normal age-based survey selectivity with selectivity below age 3 fixed to 0, and catchability fixed to 0.701 (the mean derived age-based average selectivity in the old model for ages 7-21, when the curve had come to a plateau), compared to the old model, as implemented in 2016, but with standard age-based likelihood formulations.

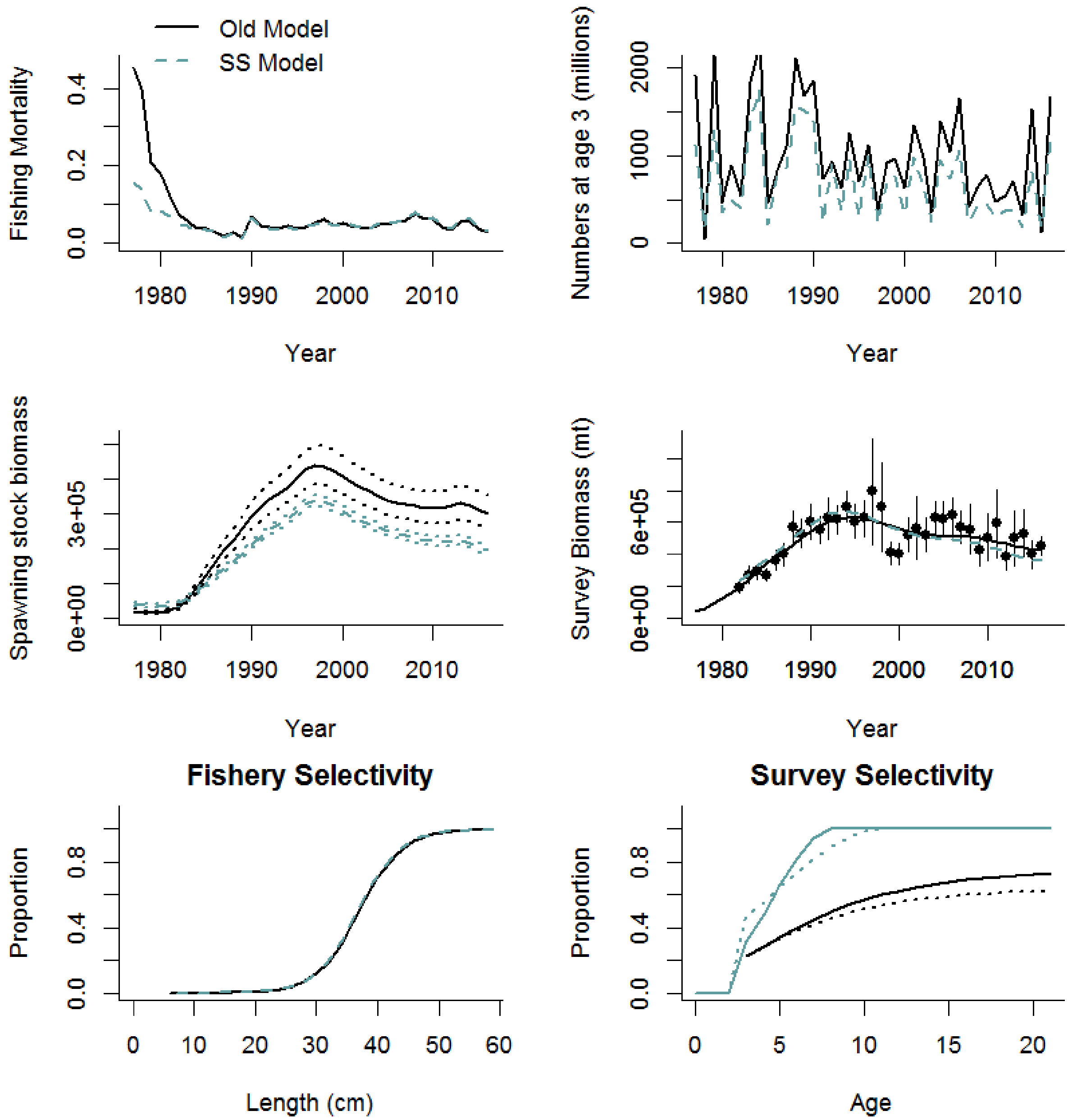


Figure 12. As for Figure 11, but the old model used is the 2016 model, which calculates predicted catch and survey biomass from numbers-at-length and length-based selectivity.

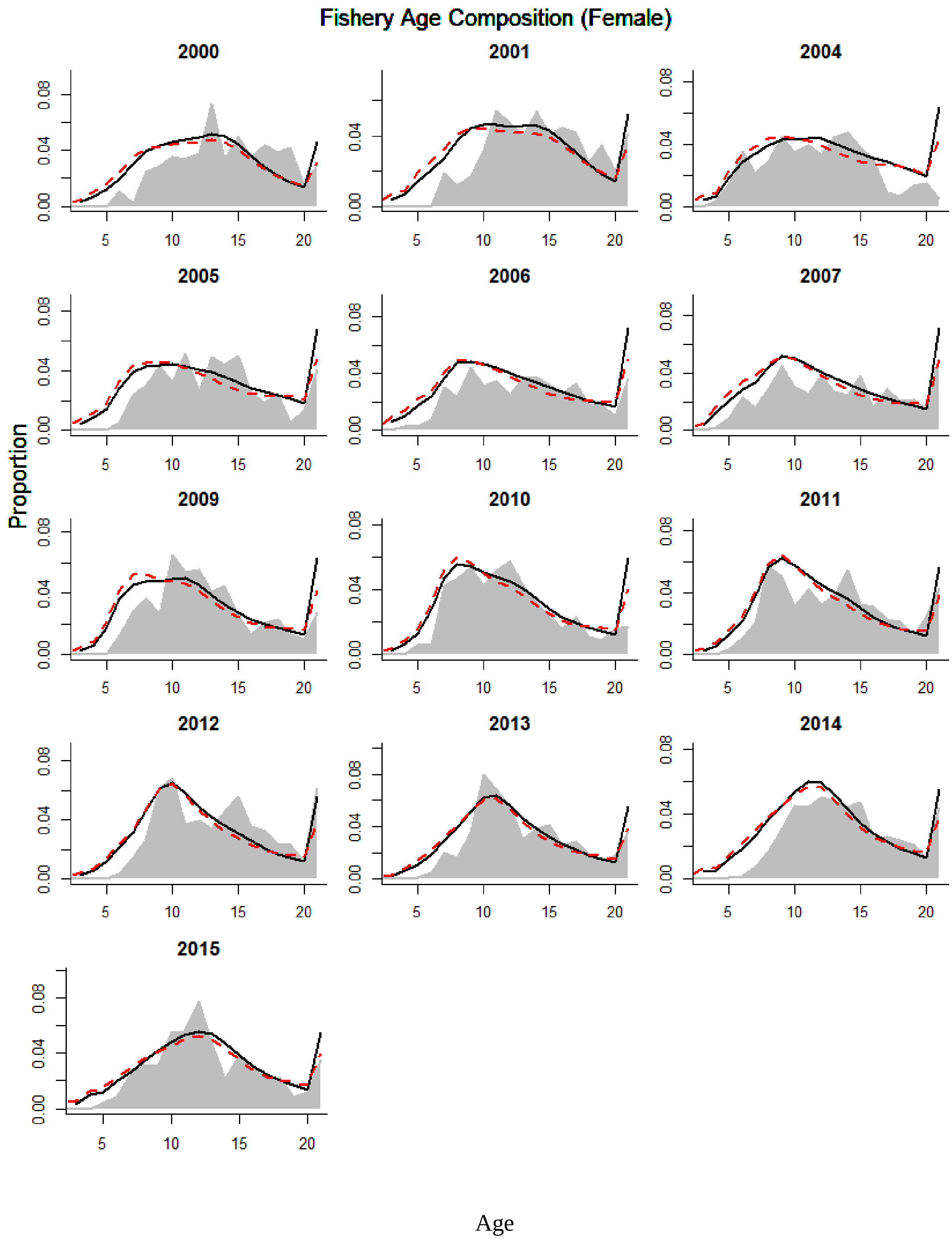


Figure 13. Yearly female fishery age composition for the 2016 model (black lines), the best matching SS3 model (with age-based selectivity and catchability configured to mimic the length-based catchability of the old model) shown in Figure 12 (dashed red lines), and the data (grey shaded areas).

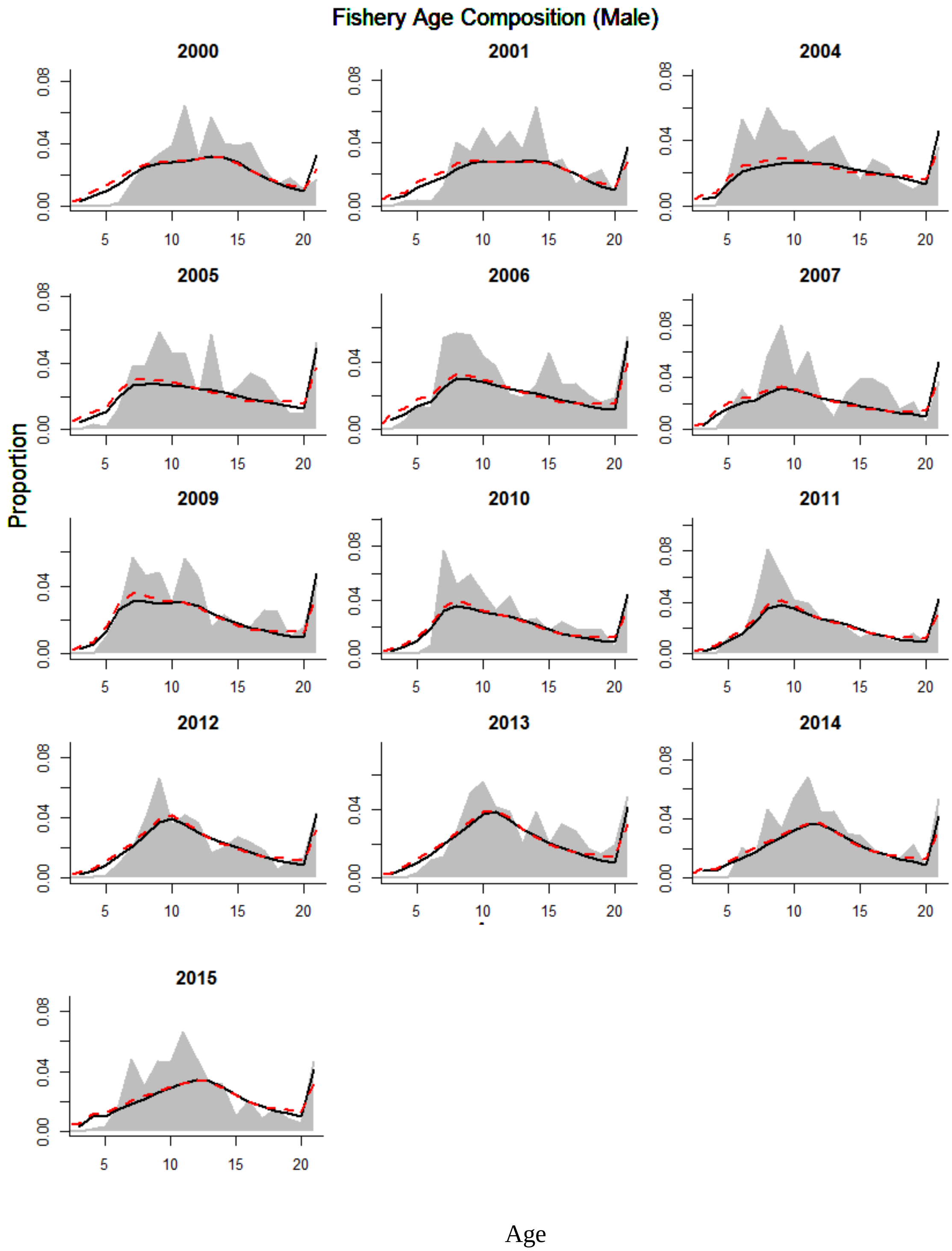


Figure 14. Yearly male fishery age composition for the 2016 configuration of the old model (black lines), the best matching SS3 model (with age-based selectivity and catchability configured to mimic the length-based catchability of the old model) shown in Figure 12 (dashed red lines), and the data (grey shaded areas).

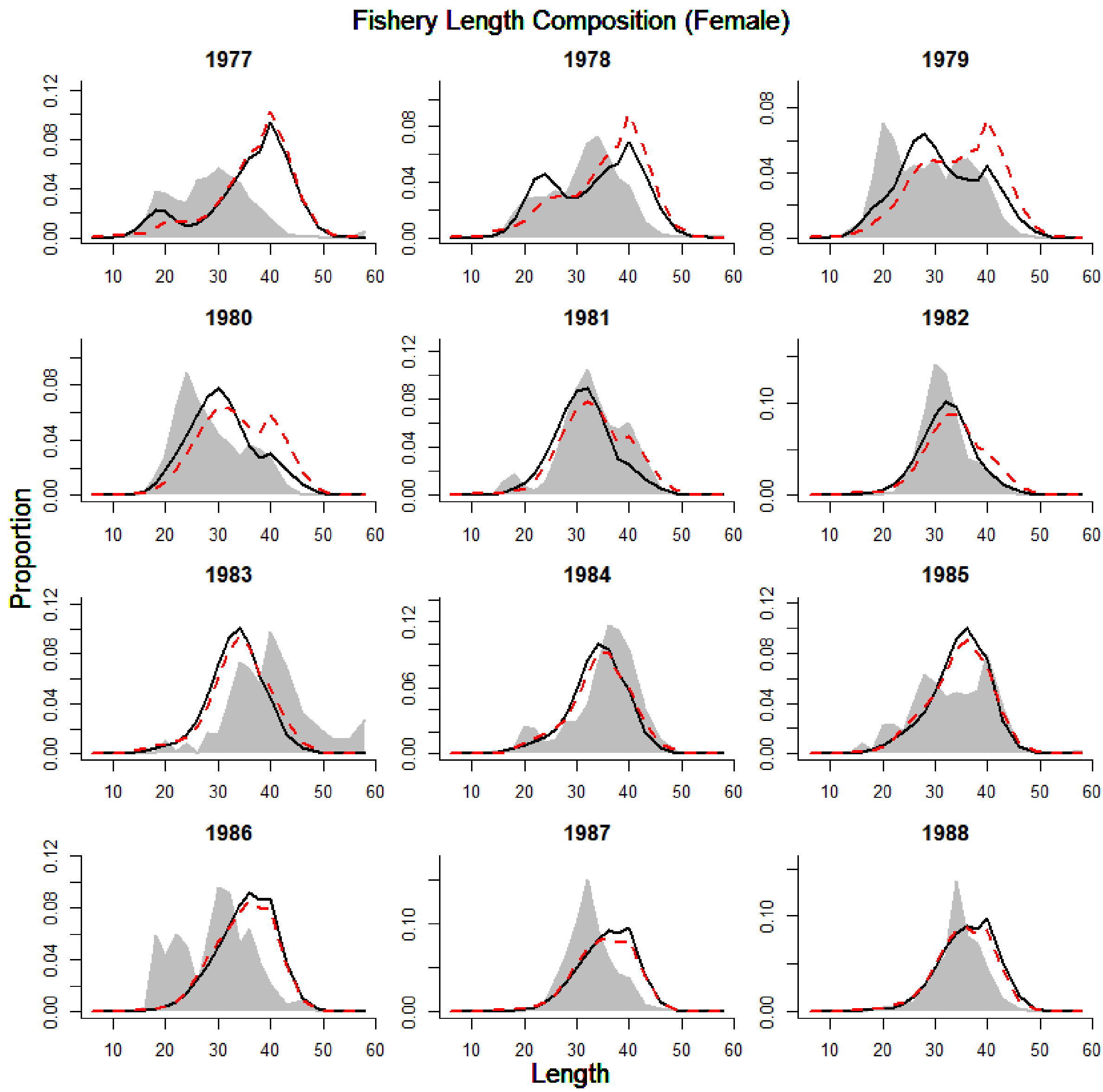


Figure 15. Yearly female fishery length composition for the 2016 configuration of the old model (black lines), the best matching SS3 model (with age-based selectivity and catchability configured to mimic the length-based catchability of the old model) shown in Figure 12 (dashed red lines), and the data (grey shaded areas).

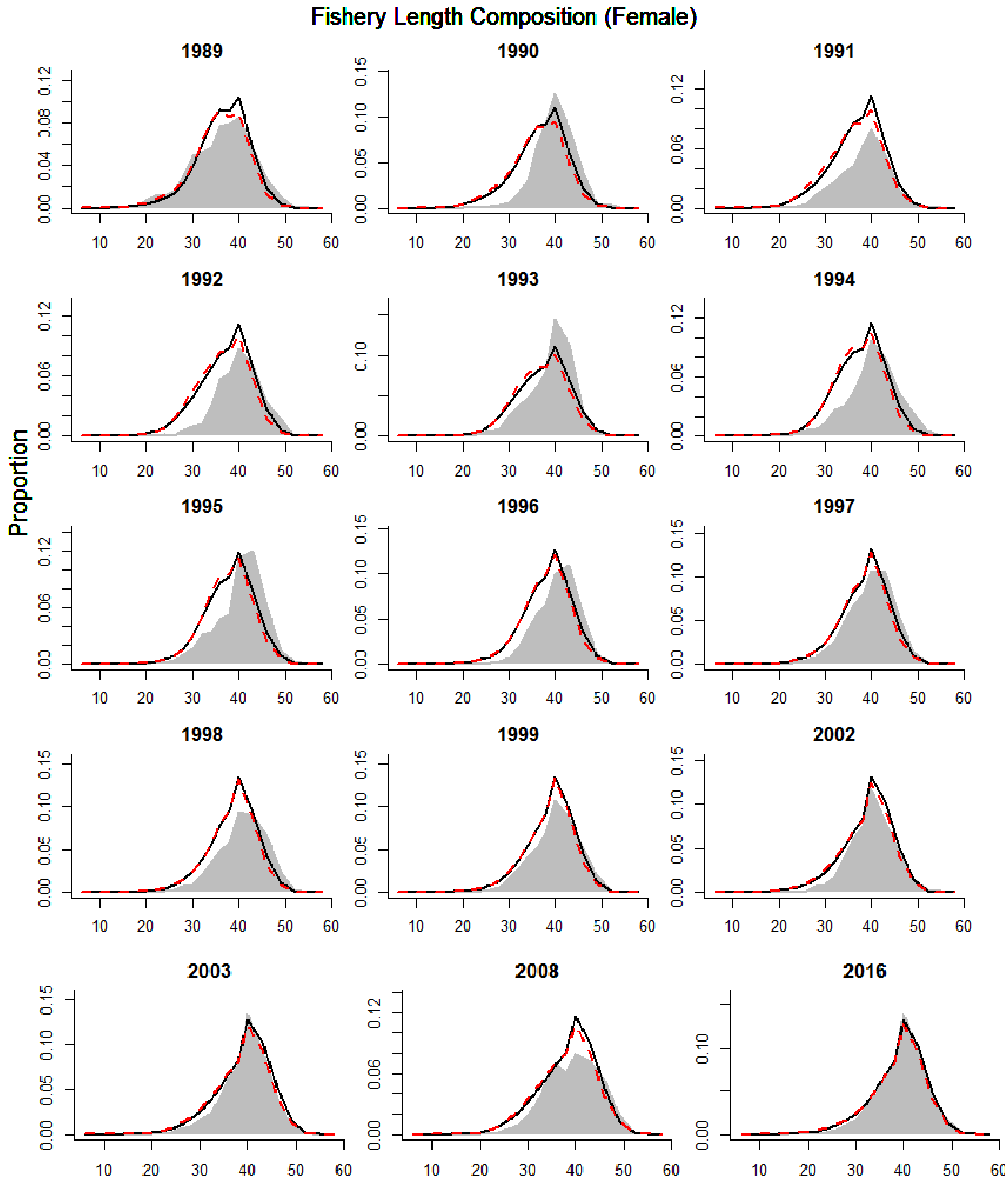


Figure 15, continued.

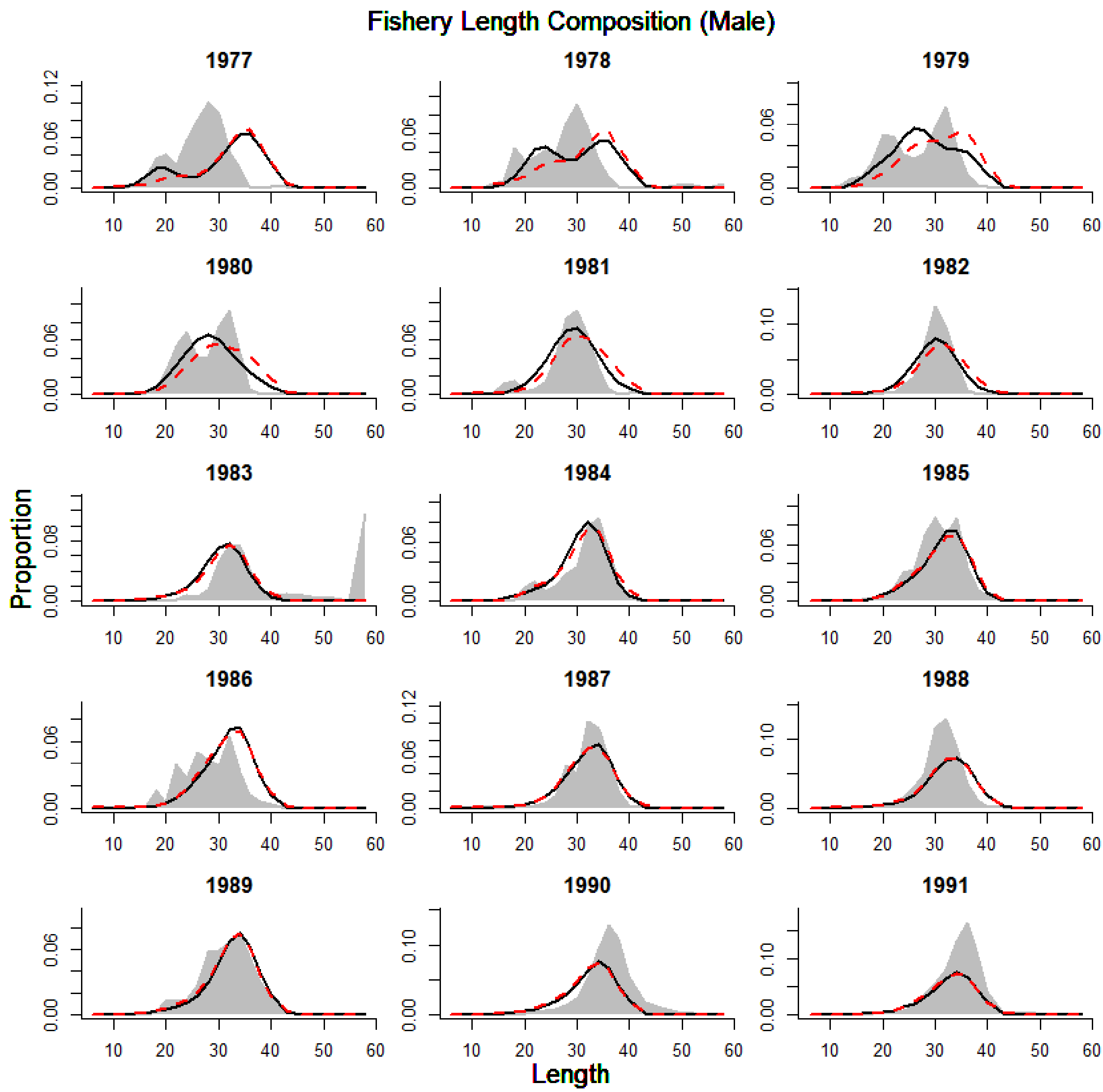


Figure 16. Yearly male fishery length composition for the 2016 configuration of the old model (black lines), the best matching SS3 model (with age-based selectivity and catchability configured to mimic the length-based catchability of the old model) shown in Figure 12 (dashed red lines), and the data (grey shaded areas).

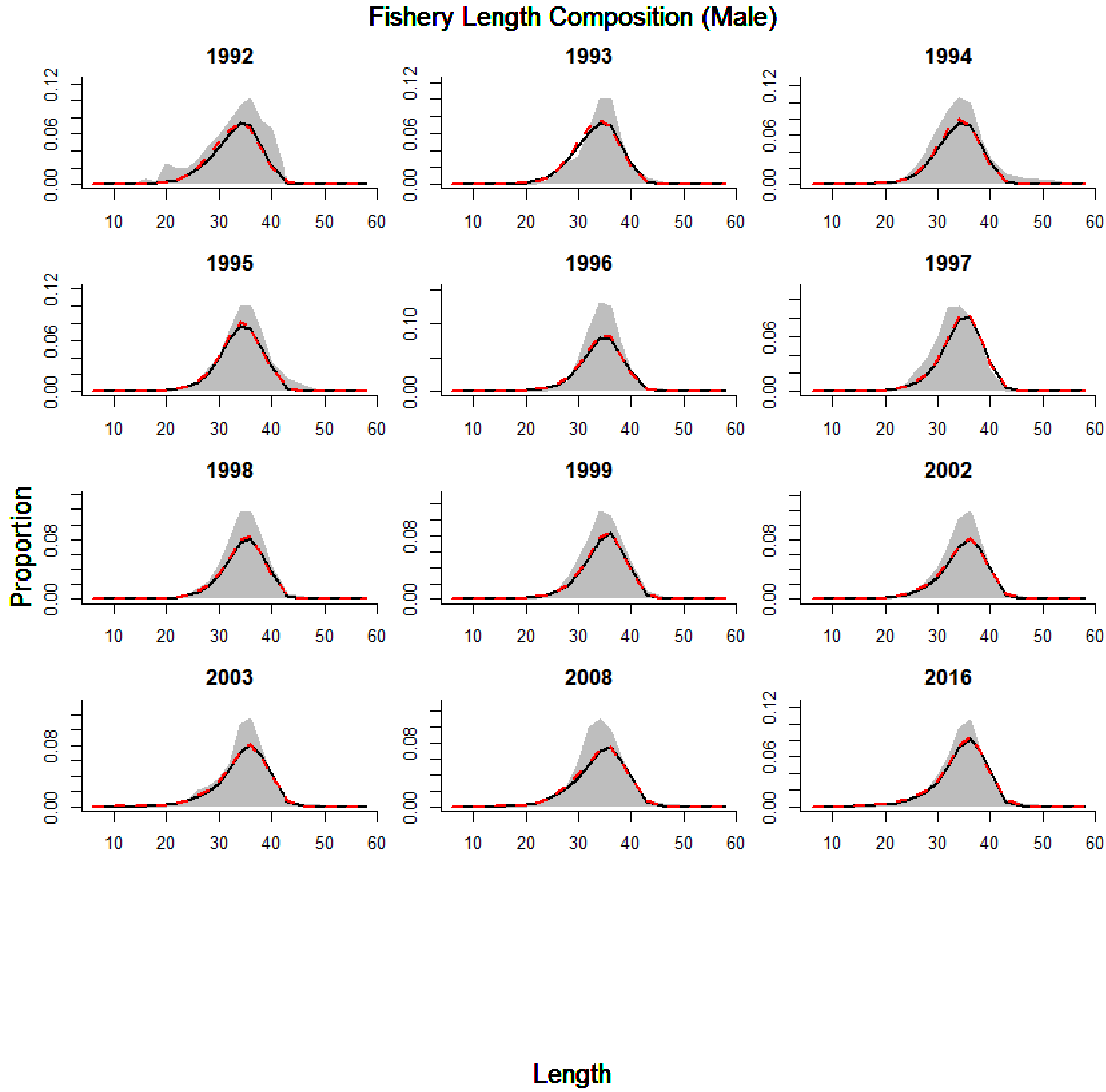


Figure 16, continued.

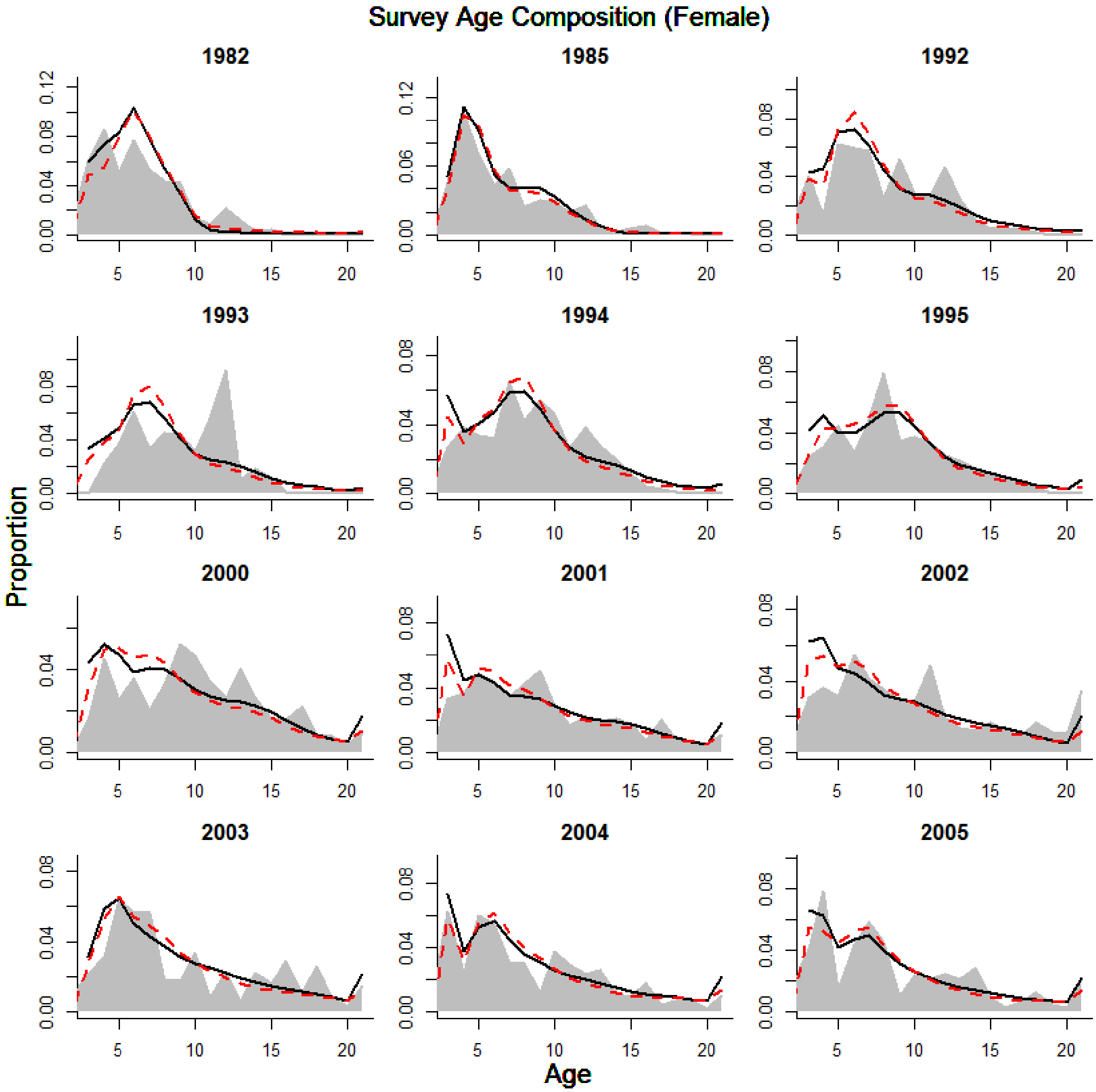


Figure 17. Yearly female survey age composition for the 2016 configuration of the old model (black lines), the best matching SS3 model (with age-based selectivity and catchability configured to mimic the length-based catchability of the old model) shown in Figure 12 (dashed red lines), and the data (grey shaded areas).

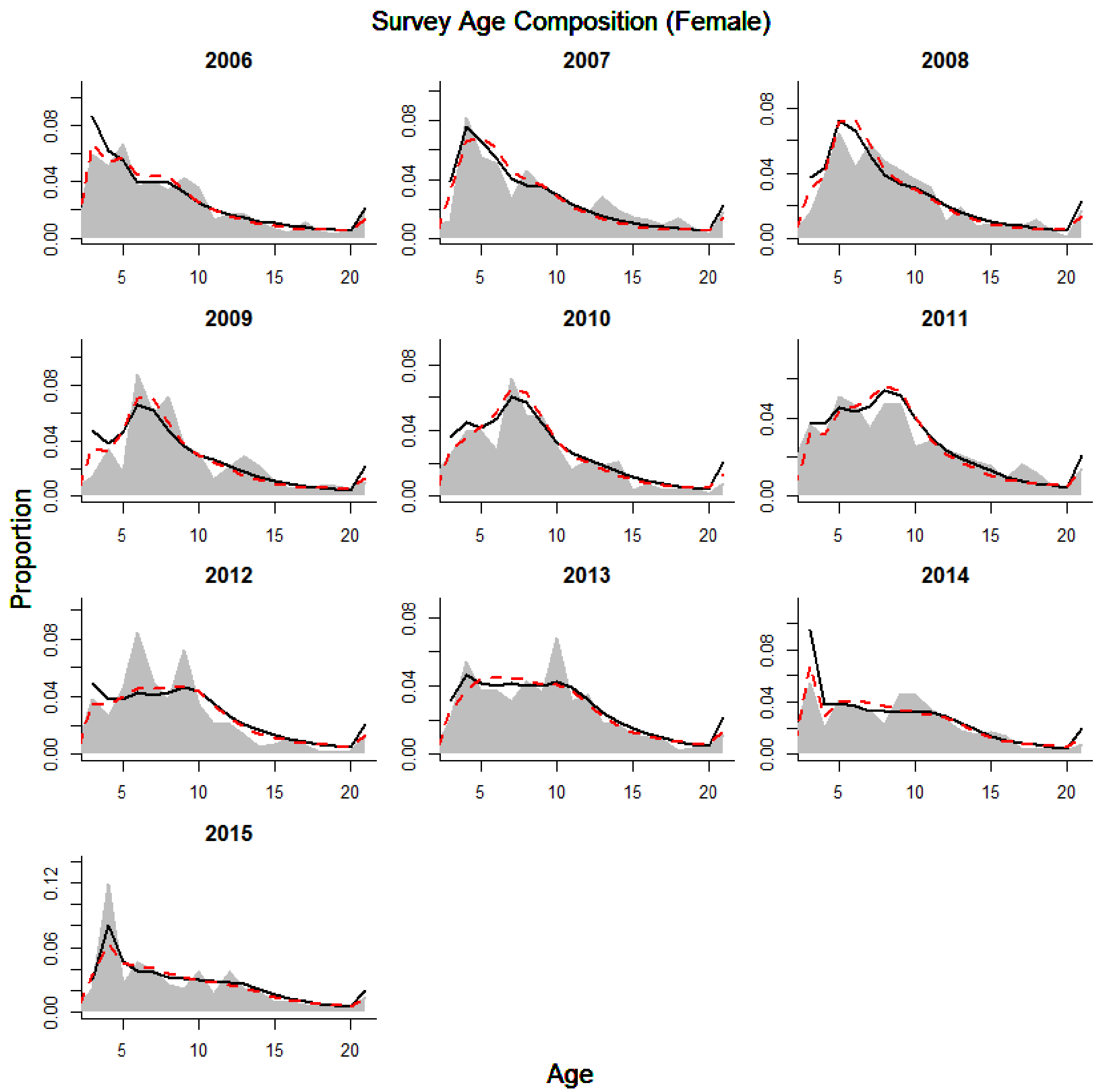


Figure 17, continued

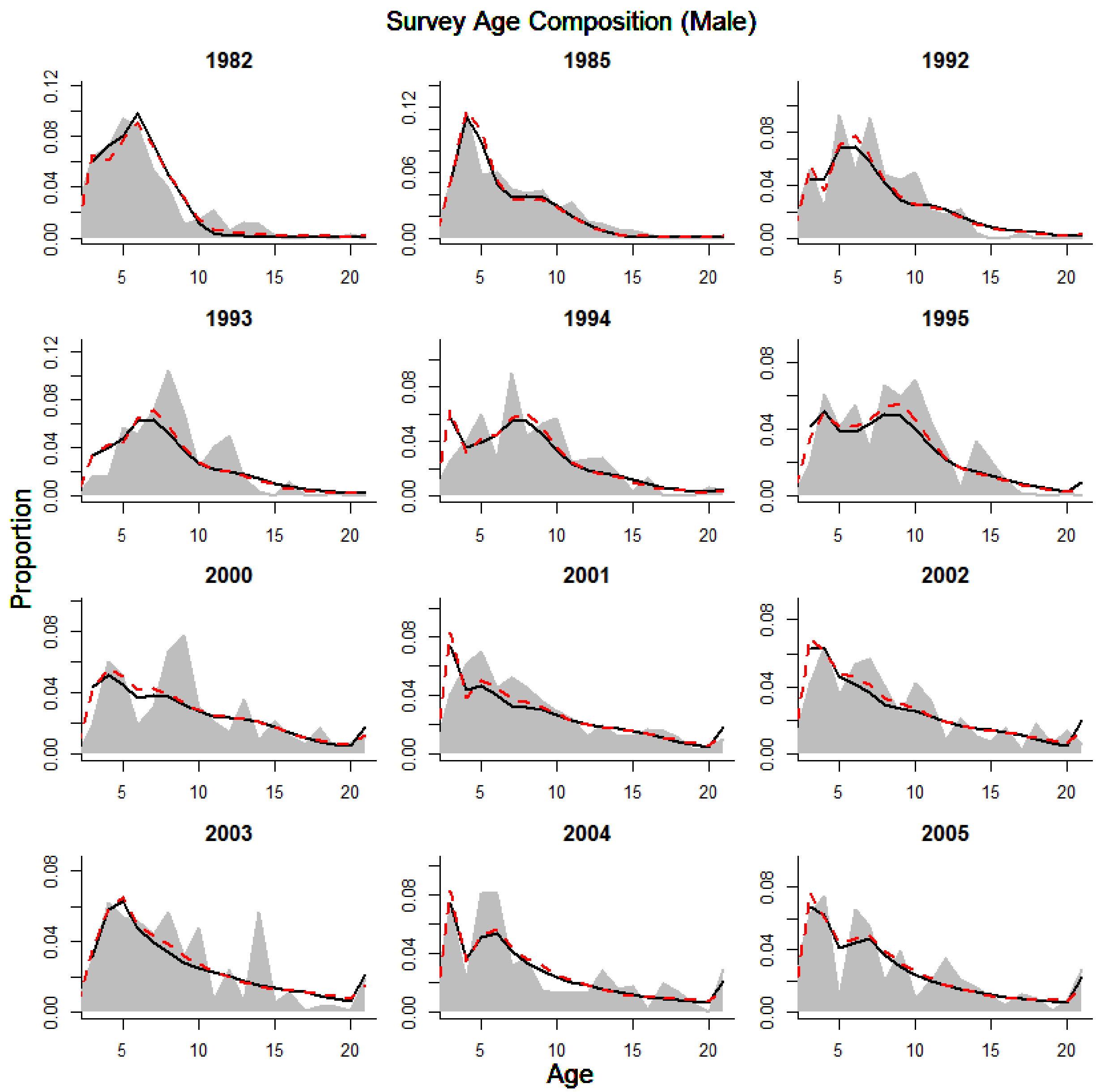


Figure 18. Yearly male survey age composition for the 2016 configuration of the old model (black lines), the best matching SS3 model (with age-based selectivity and catchability configured to mimic the length-based catchability of the old model) shown in Figure 12 (dashed red lines), and the data (grey shaded areas).

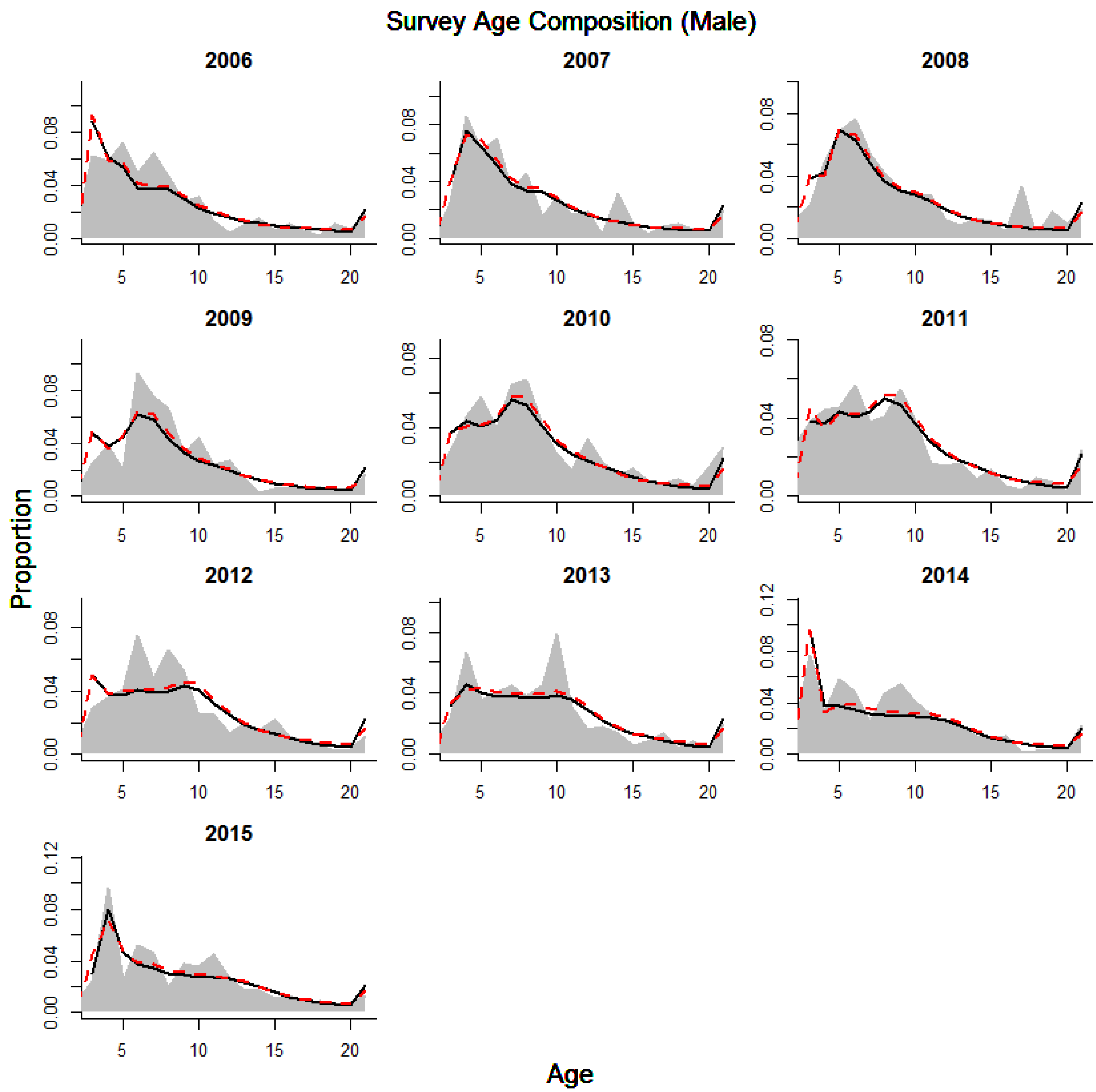


Figure 18, continued.

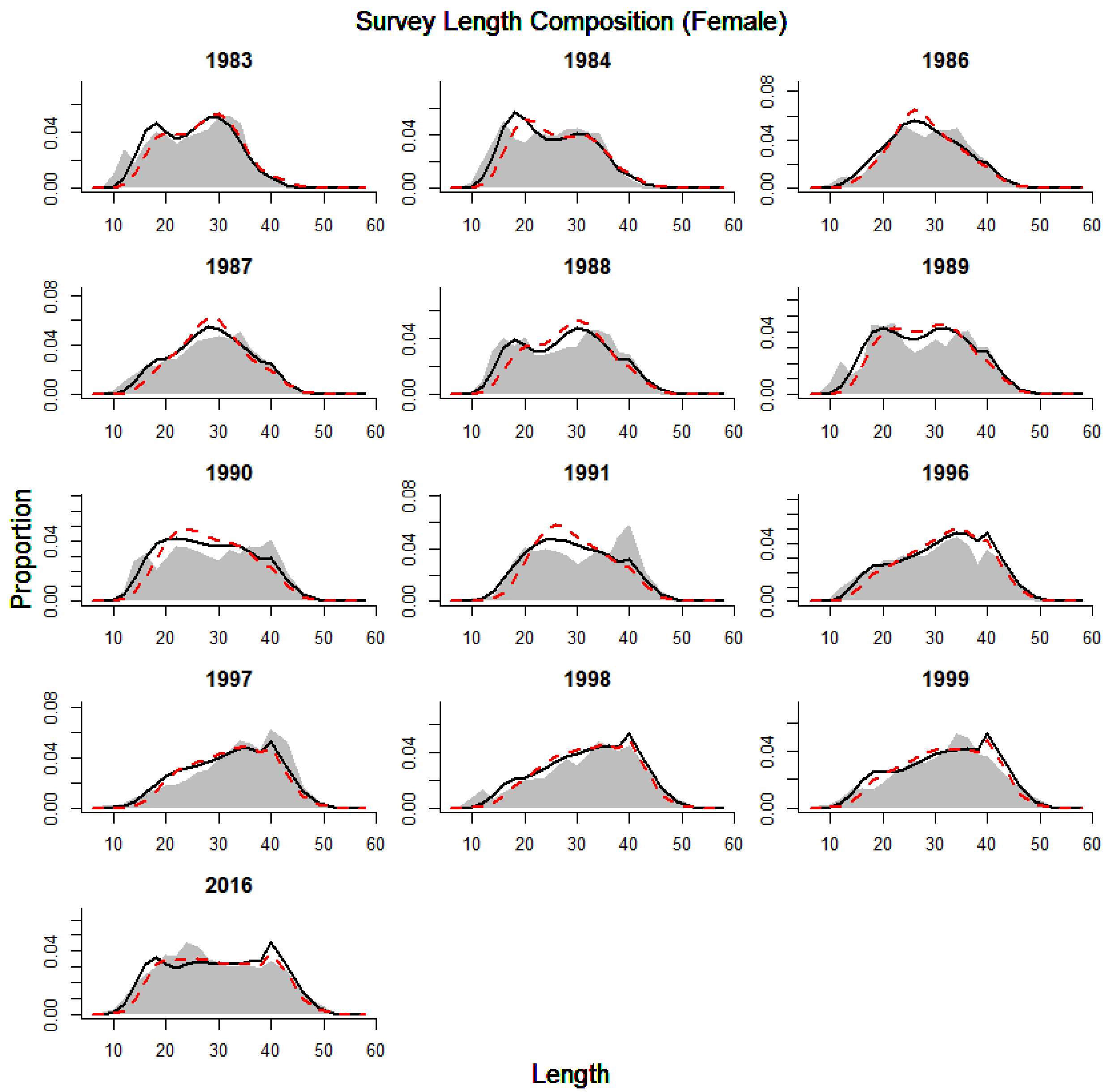


Figure 19. Yearly female survey length composition for the 2016 configuration of the old model (black lines), the best matching SS3 model (with age-based selectivity and catchability configured to mimic the length-based catchability of the old model) shown in Figure 12 (dashed red lines), and the data (grey shaded areas).

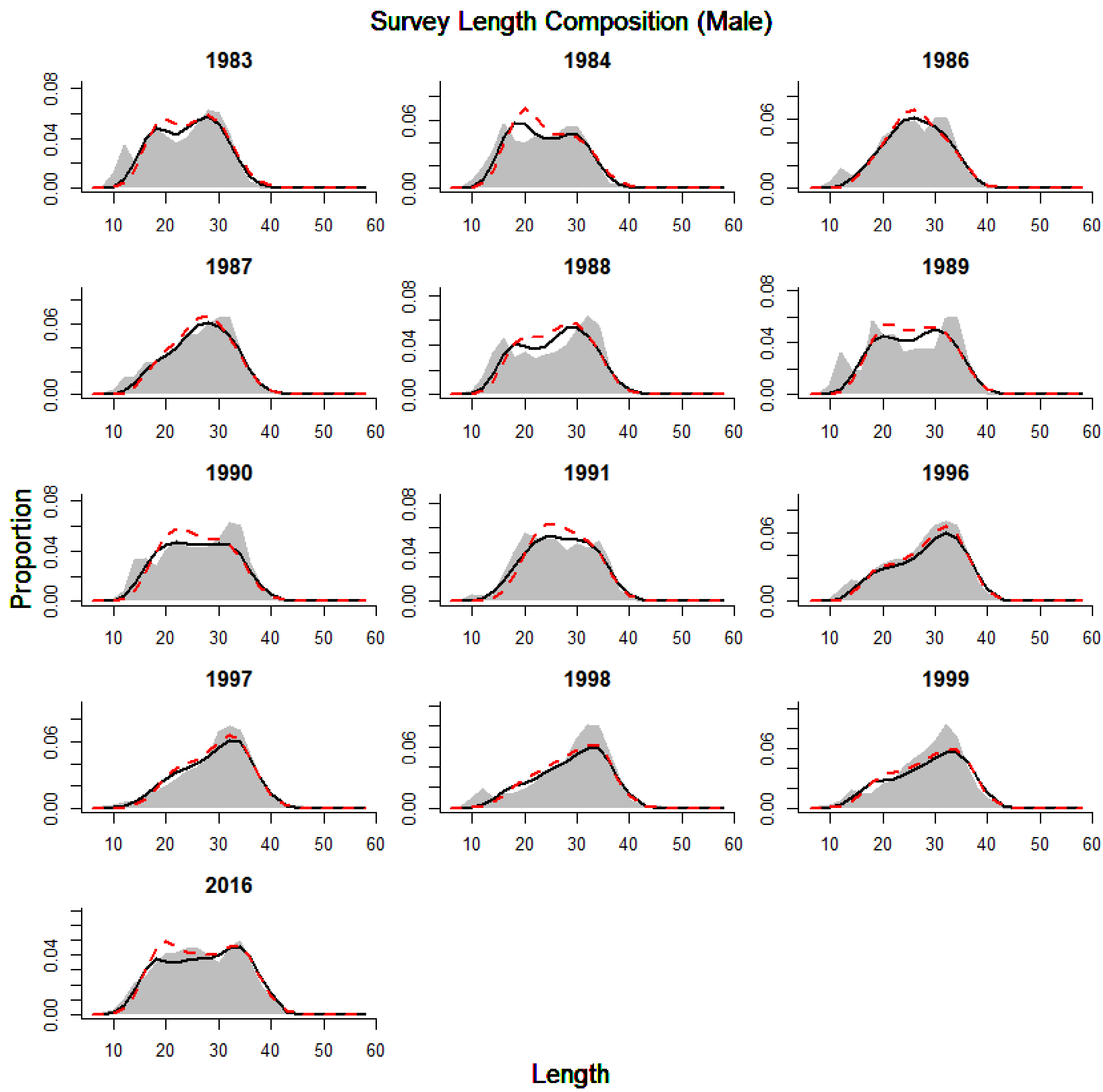


Figure 20. Yearly male survey length composition for the 2016 configuration of the old model (black lines), the best matching SS3 model (with age-based selectivity and catchability configured to mimic the length-based catchability of the old model) shown in Figure 12 (dashed red lines), and the data (grey shaded areas).

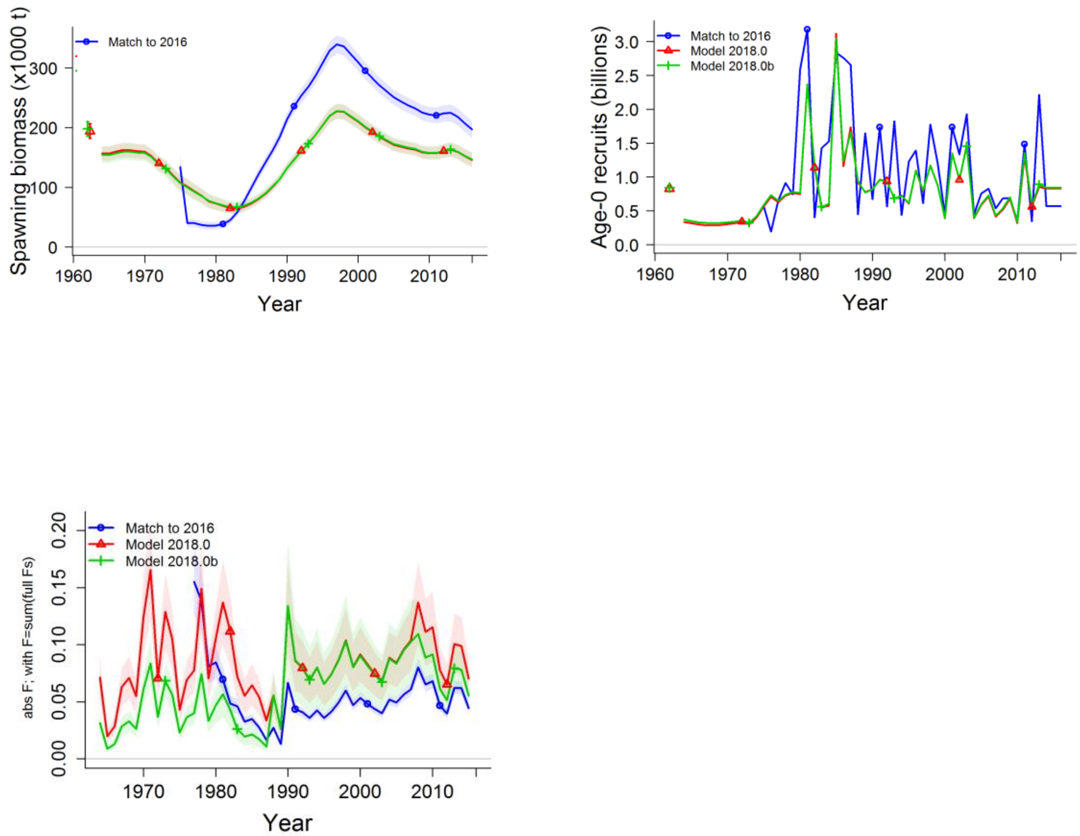


Figure 21. A comparison of the best matching SS3 model, and two proposed models for 2018.0 and 2018.0b. Both 2018.0 and 2018.0b. 2018.0 and 2018.0b are identical, except that 2018.0b estimates different selectivity curves for 3 distinct time periods: 1964-1987, 1988-2007, and 2008-2016.

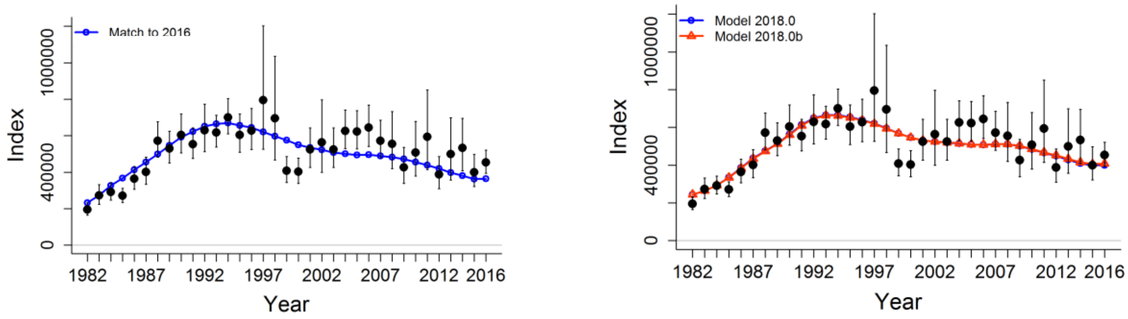


Figure 22. Observed (black dots) and predicted index of survey biomass for the best matching SS3 model to the 2016 model (left panel) and for Models 2018.0 and 2018.0b (right panel). Vertical black lines show 95% confidence intervals about the observations.

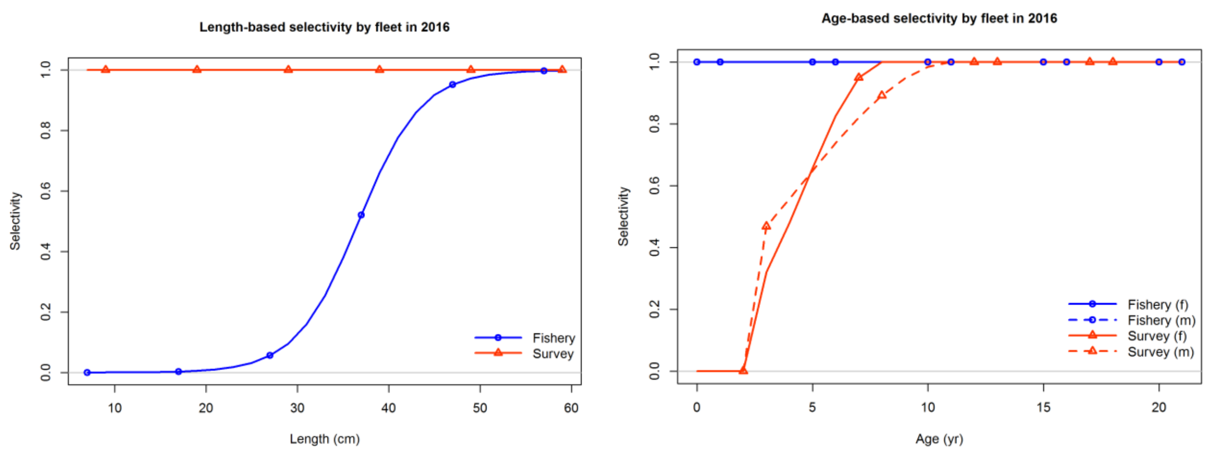


Figure 23. Selectivity for the SS3 model that best matched the 2016 model (note that catchability was fixed at 0.7 in this model). Left panel: length-based fishery selectivity (applies to both males and females), right panel: age-based, sex-specific survey selectivity.

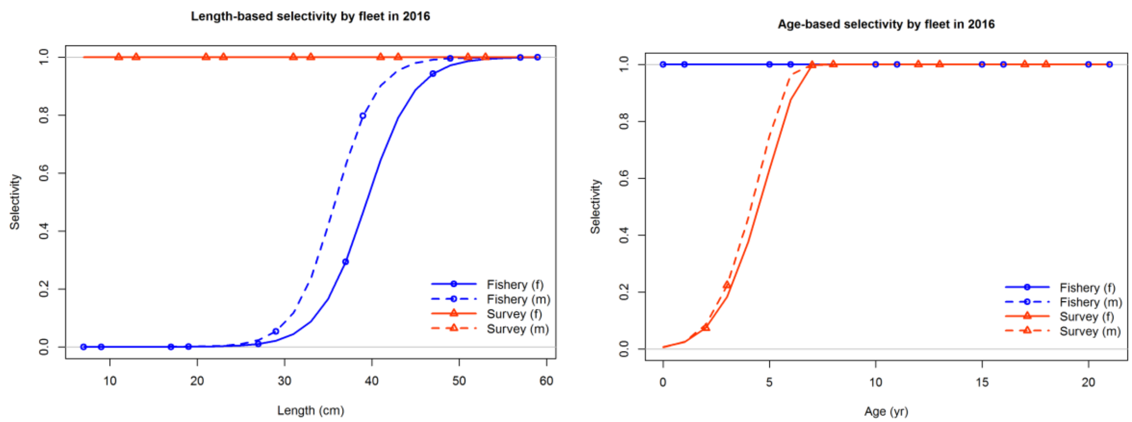


Figure 24. Selectivity for Model 2018.0. Left panel: length-based, sex-specific fishery selectivity, right panel: age-based, sex-specific survey selectivity.

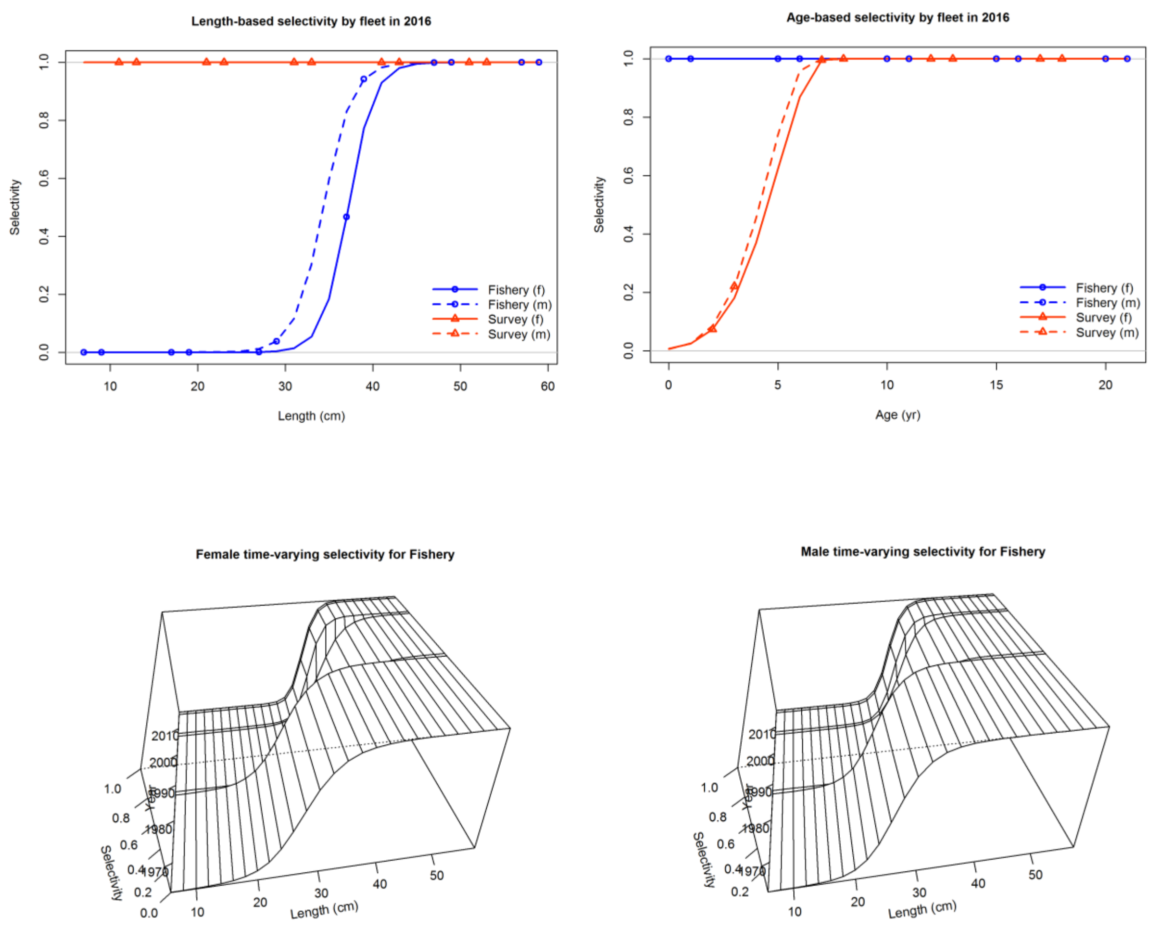


Figure 25. Fishery and survey selectivity curves for Model 2018.0b, where separate selectivity cuves were estimated for 3 time periods: 1964-1987, 1988-2007, and 2008-2017.

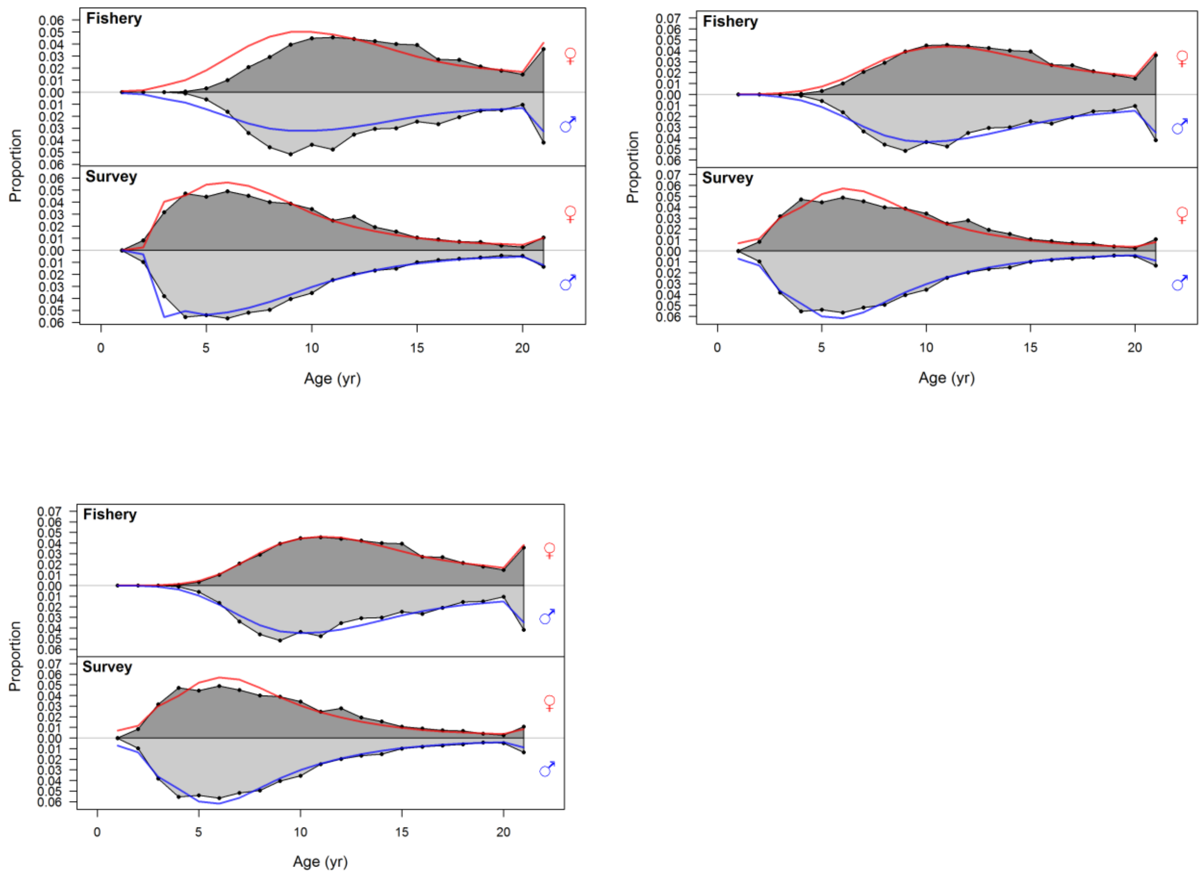


Figure 26. A comparison of aggregated fits to age composition data (grey) for the SS3 best model match to the 2016 model (top left), Model 2018.0 (top right), and Model 2018.0b (bottom left).

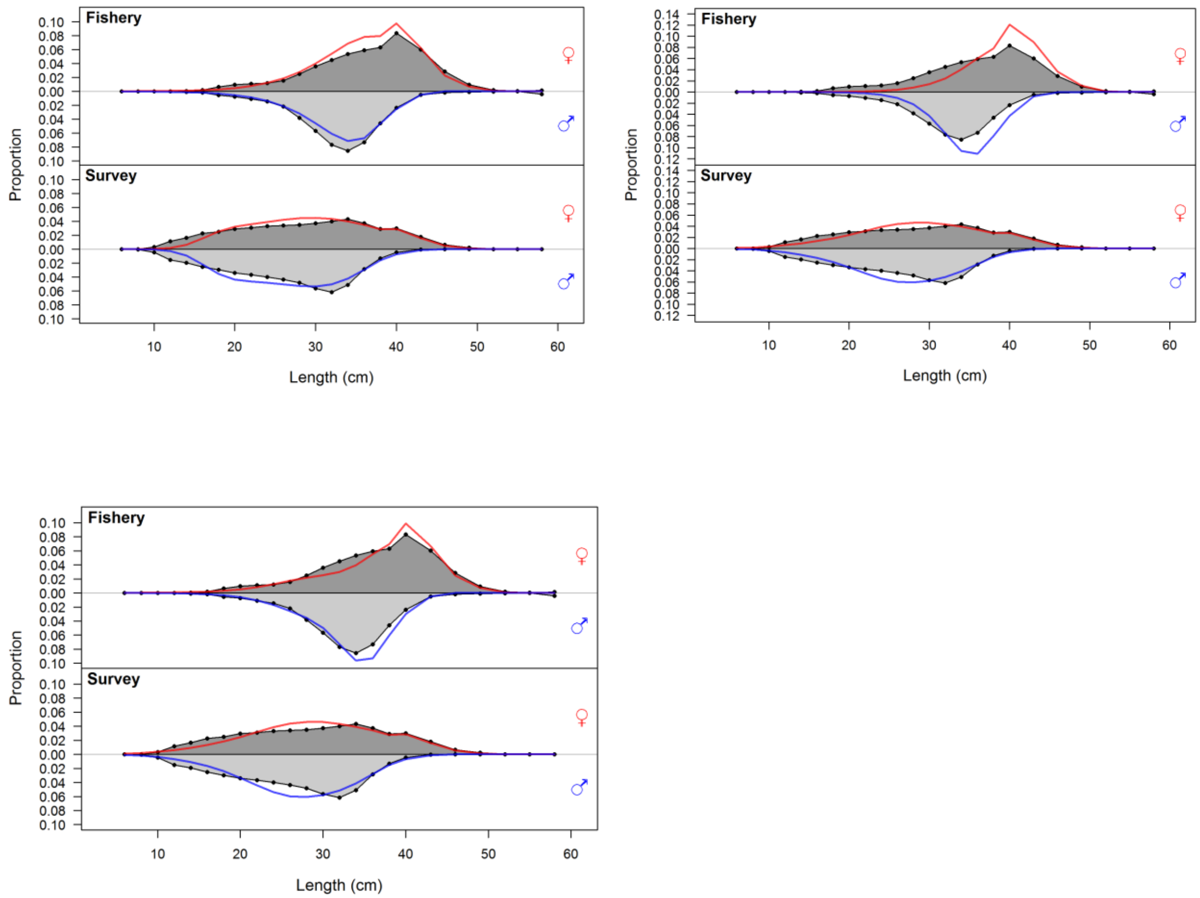


Figure 27. A comparison of aggregated fits to length composition data (grey) for the SS3 best model match to the 2016 model (top left), Model 2018.0 (top right), and Model 2018.0b (bottom left).

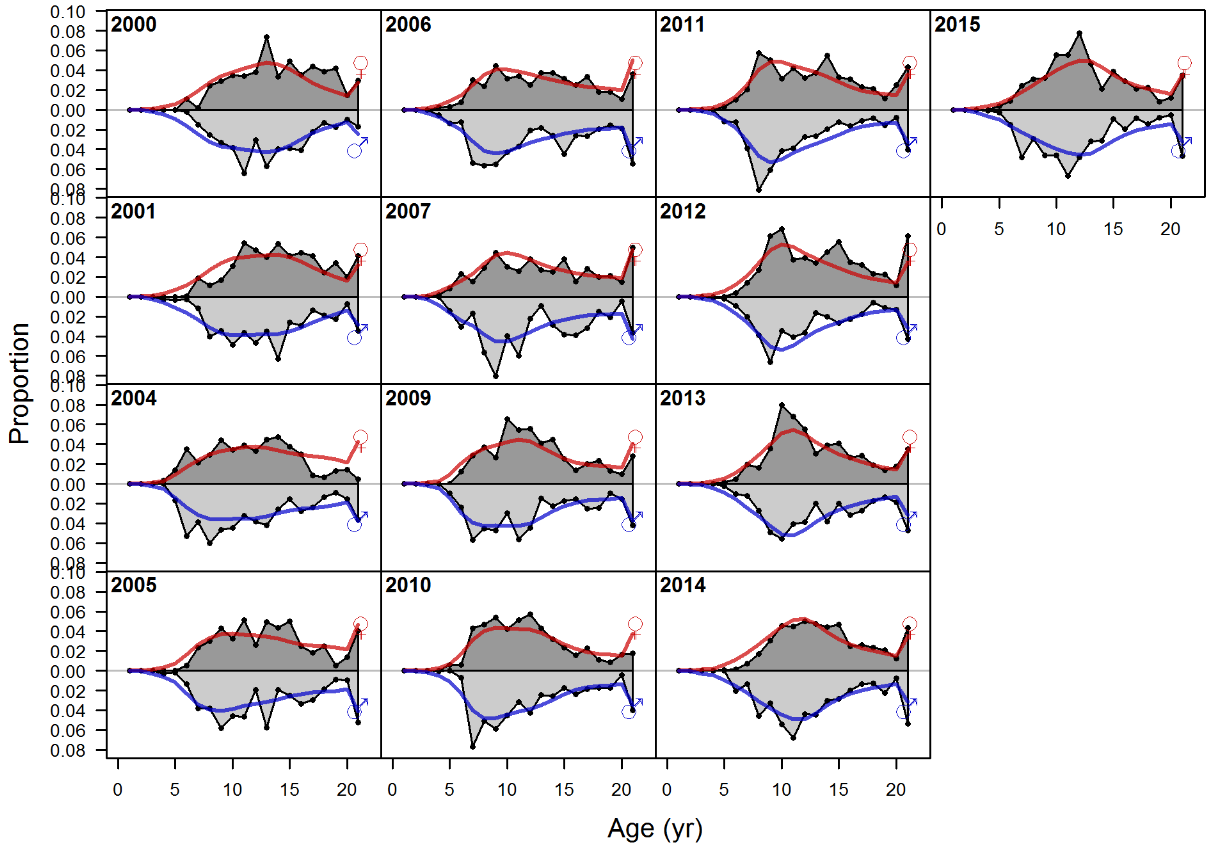


Figure 28. Fits to fishery age composition data (grey) by year for Model 2018.0.

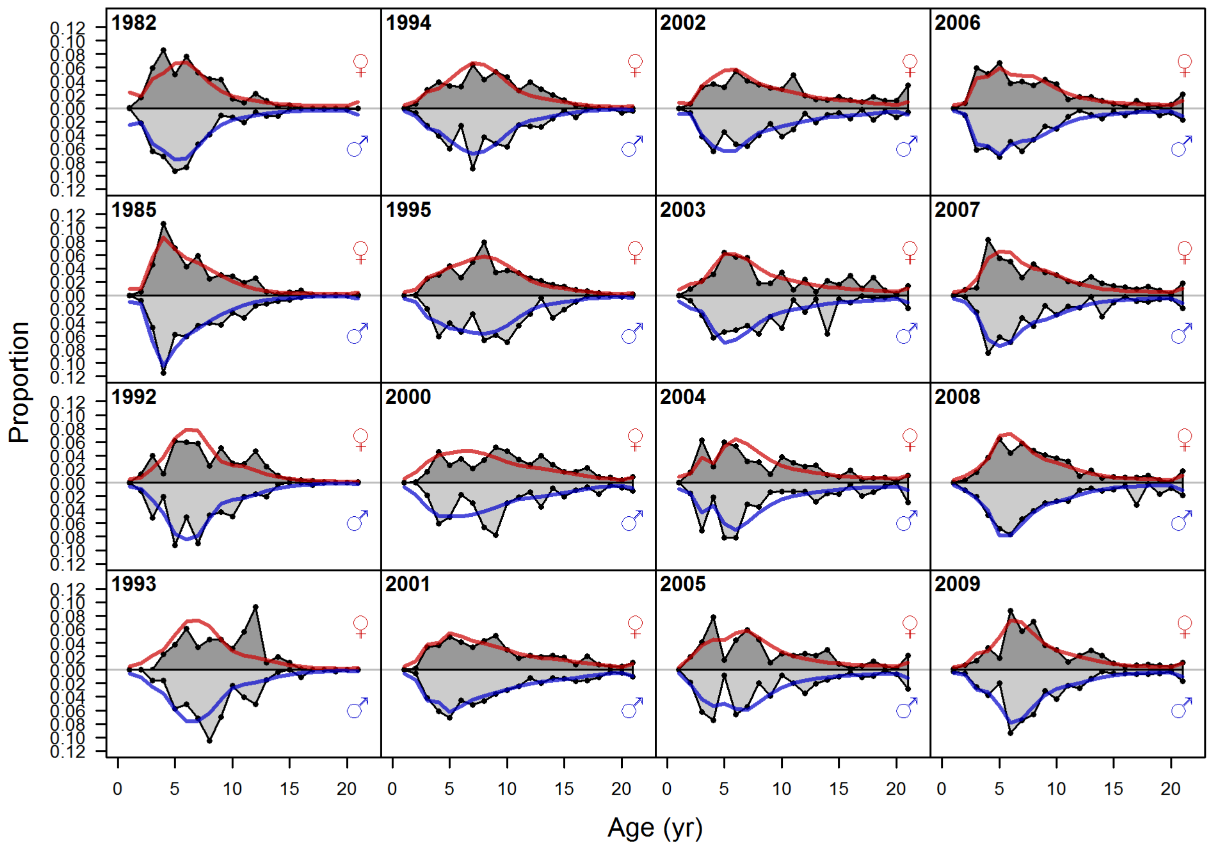


Figure 29. Fits to survey age composition data (grey) by year for Model 2018.0 (part 1 of 2).

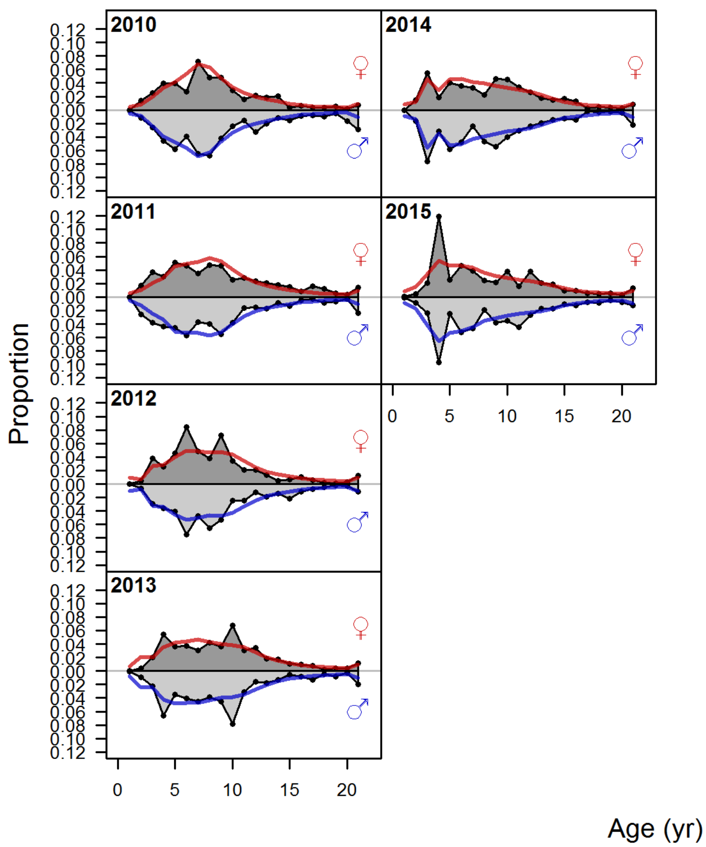


Figure 29, continued (part 2 of 2).

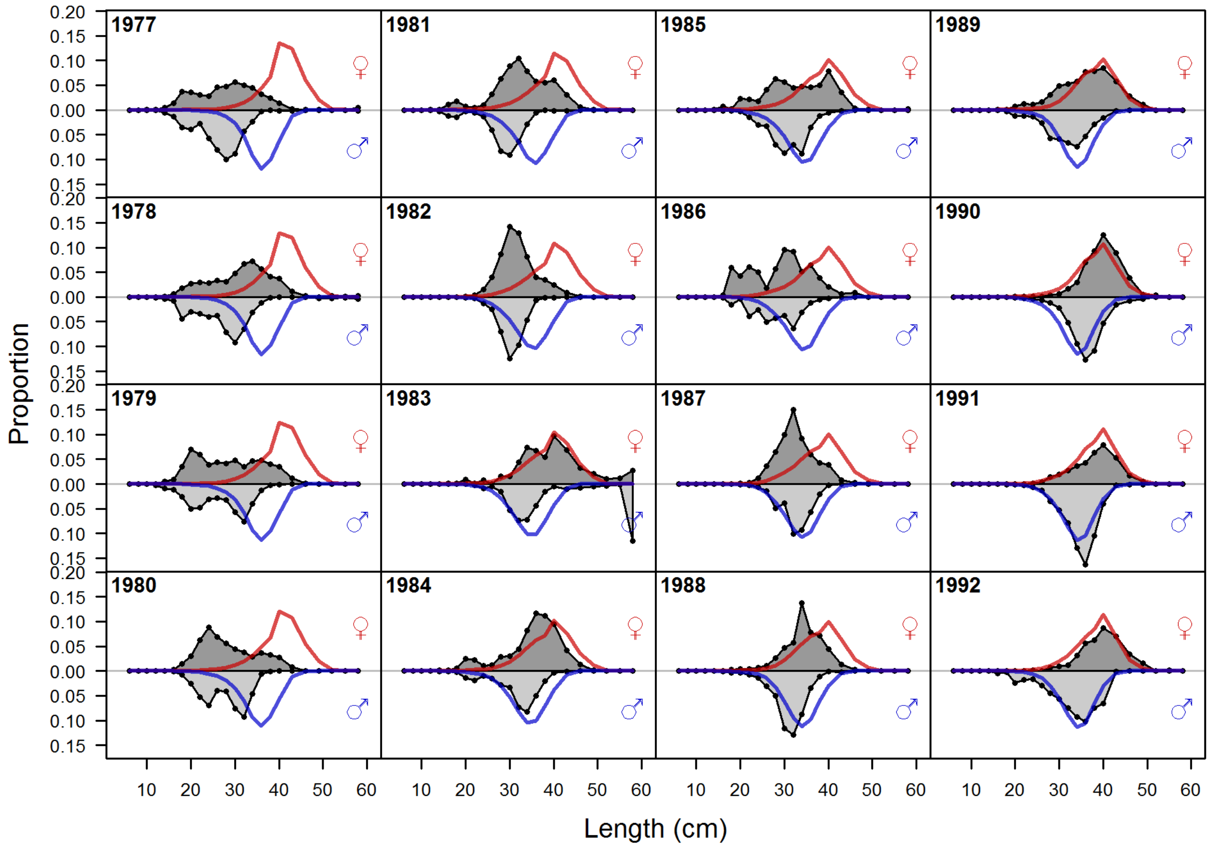


Figure 30. Fits to fishery length composition data (grey) by year for Model 2018.0. In years for which age composition data existed, length composition data were given an effective sample size of 1 (part 1 of 2).

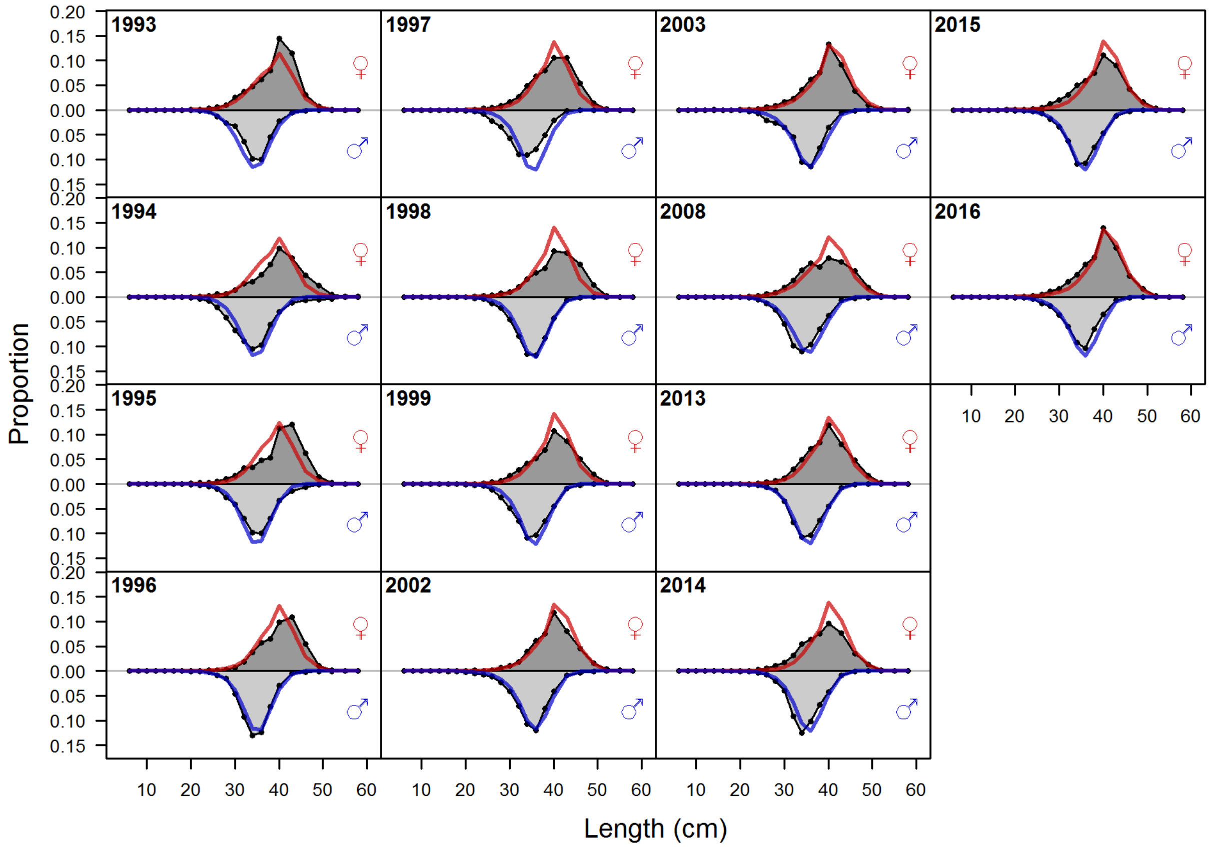


Figure 30, continued (part 2 of 2).

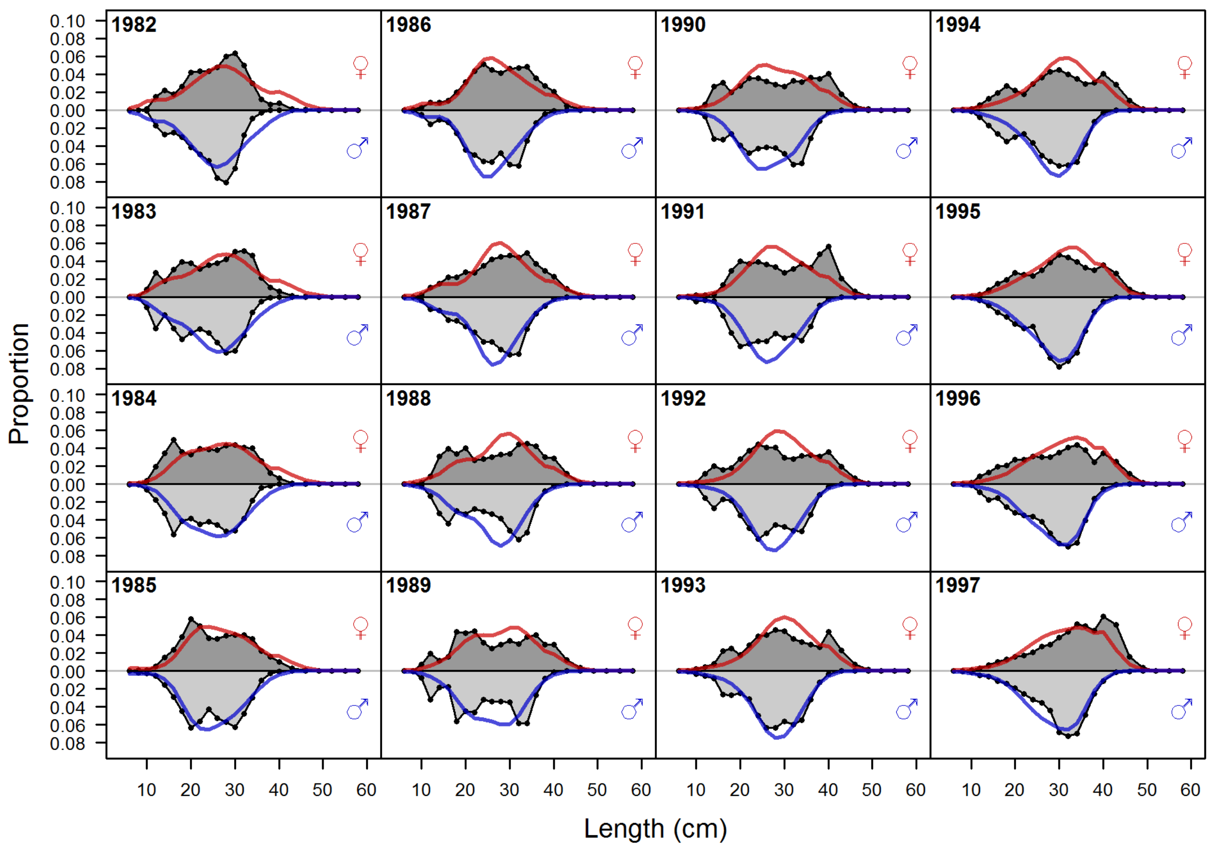


Figure 31. Fits to survey length composition data (grey) by year for Model 2018.0. In years for which age composition data existed, length composition data were given an effective sample size of 1 (part 1 of 3).

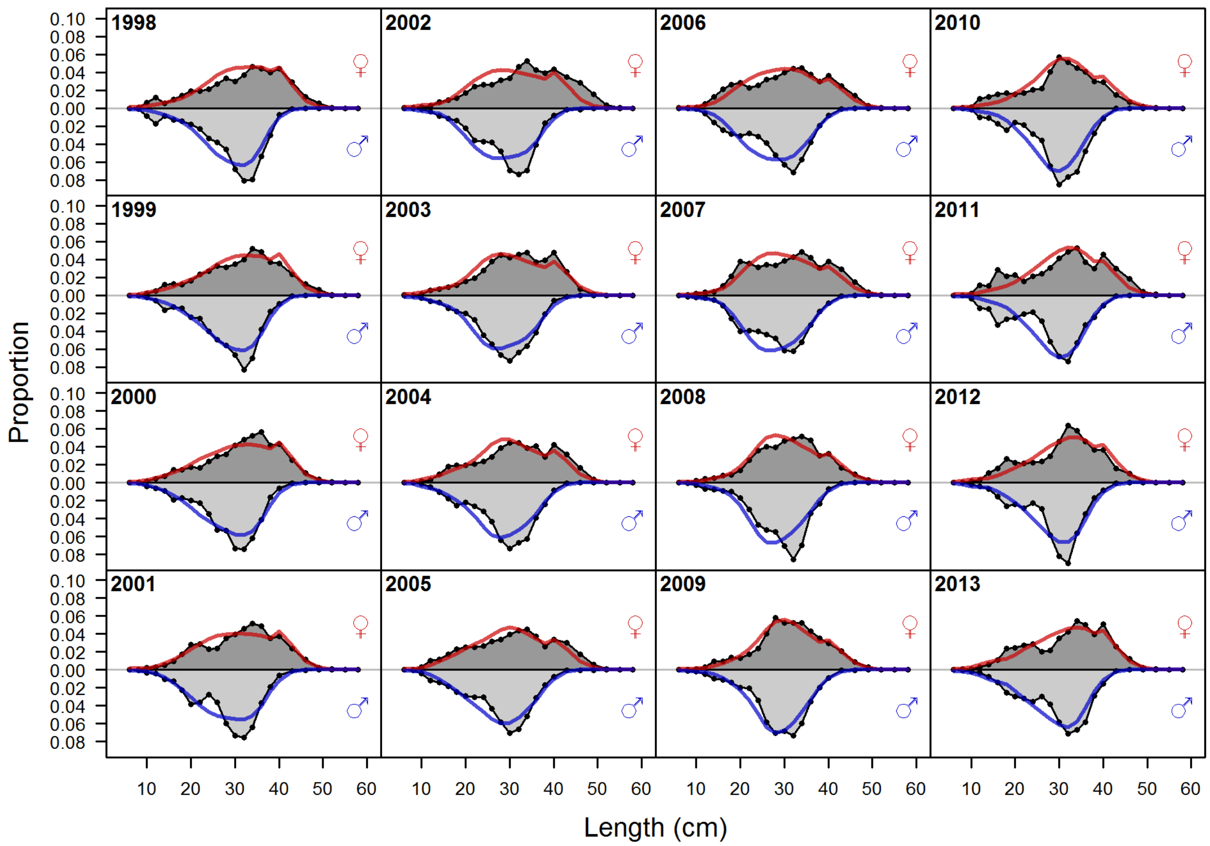


Figure 31, continued (part 2 of 3).

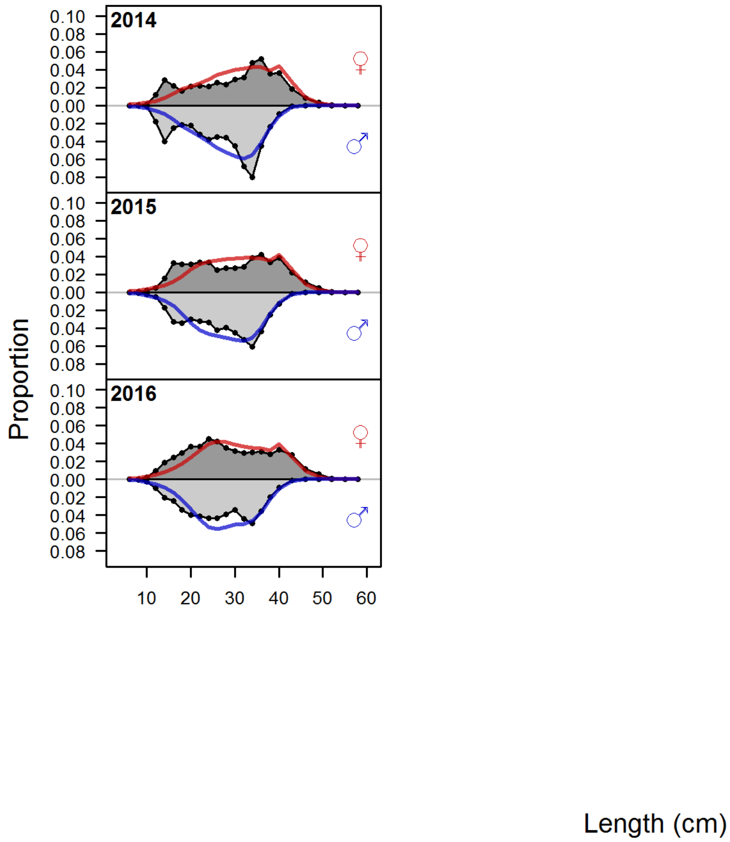


Figure 31, continued (part 3 of 3).

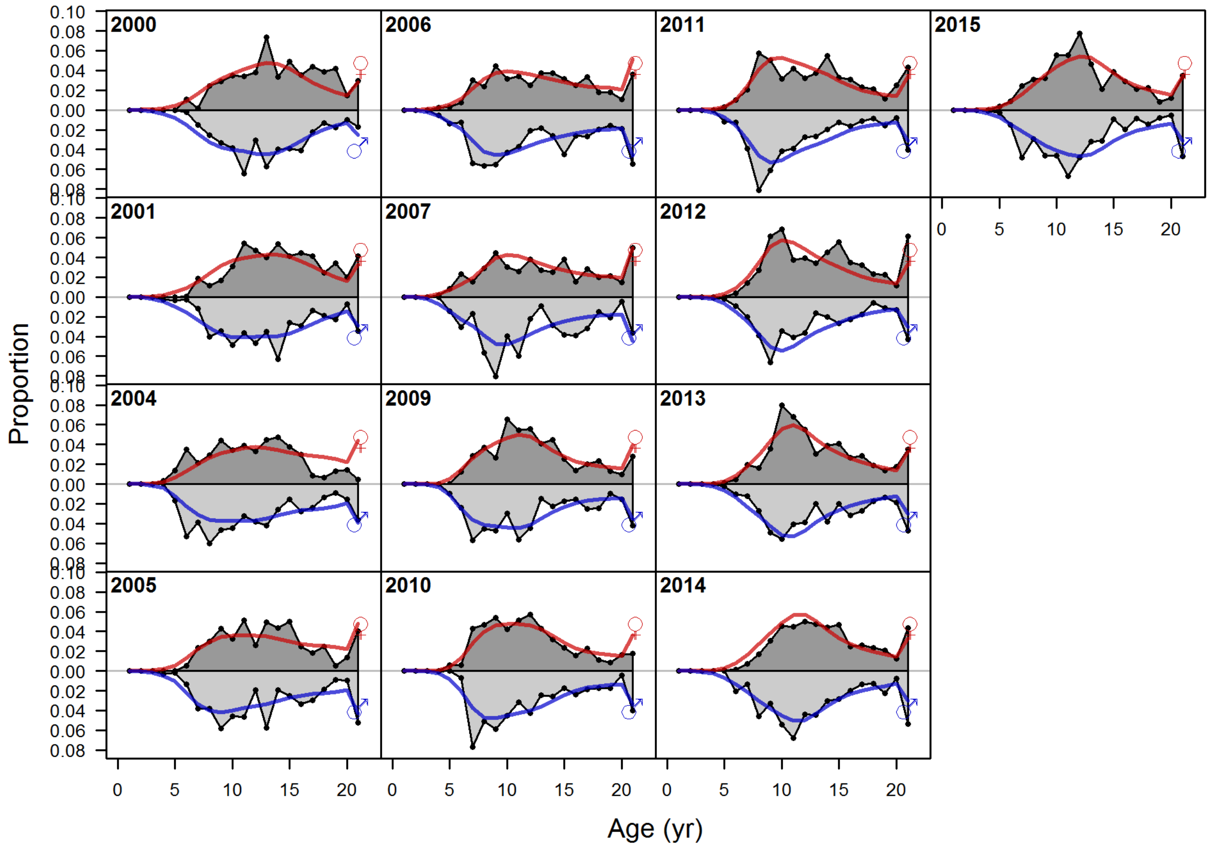


Figure 32. Fits to fishery age composition data (grey) by year for Model 2018.0b.

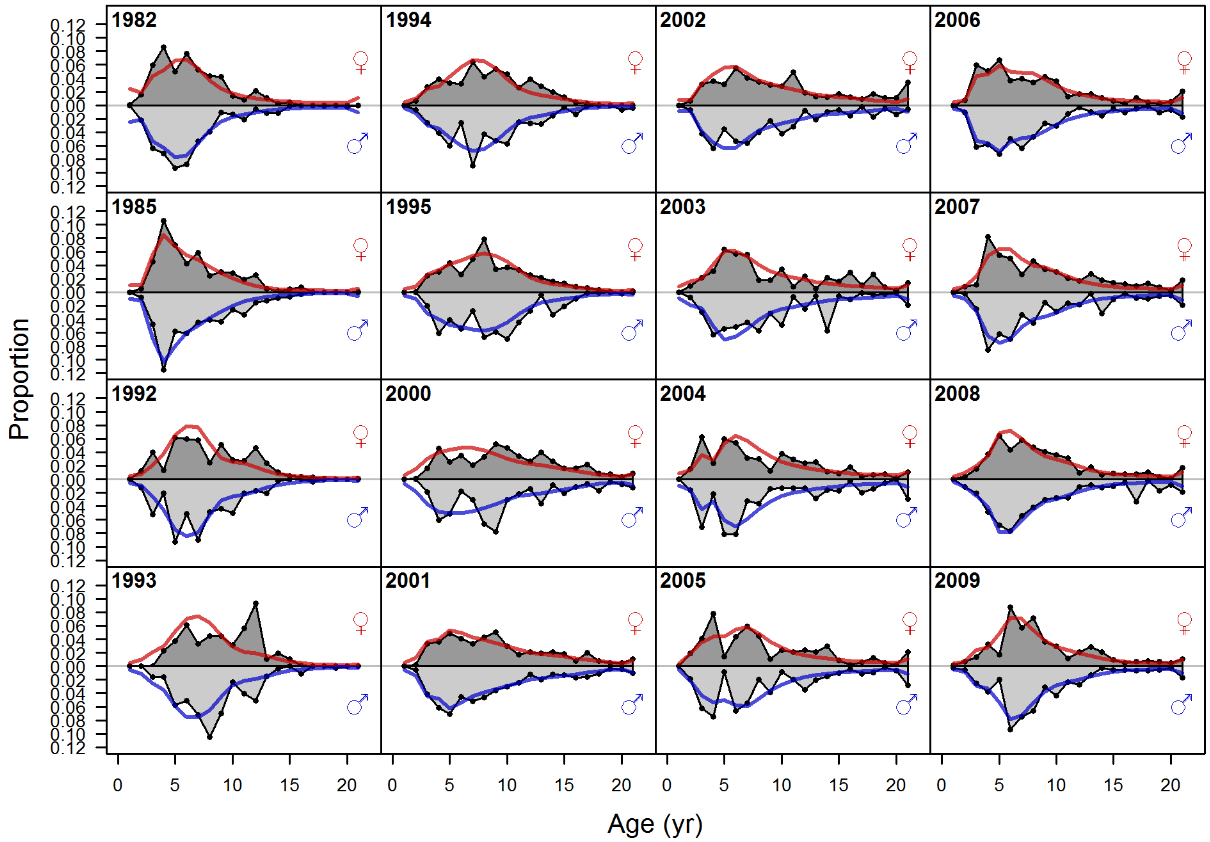


Figure 33. Fits to survey age composition data (grey) by year for Model 2018.0b (part 1 of 2).

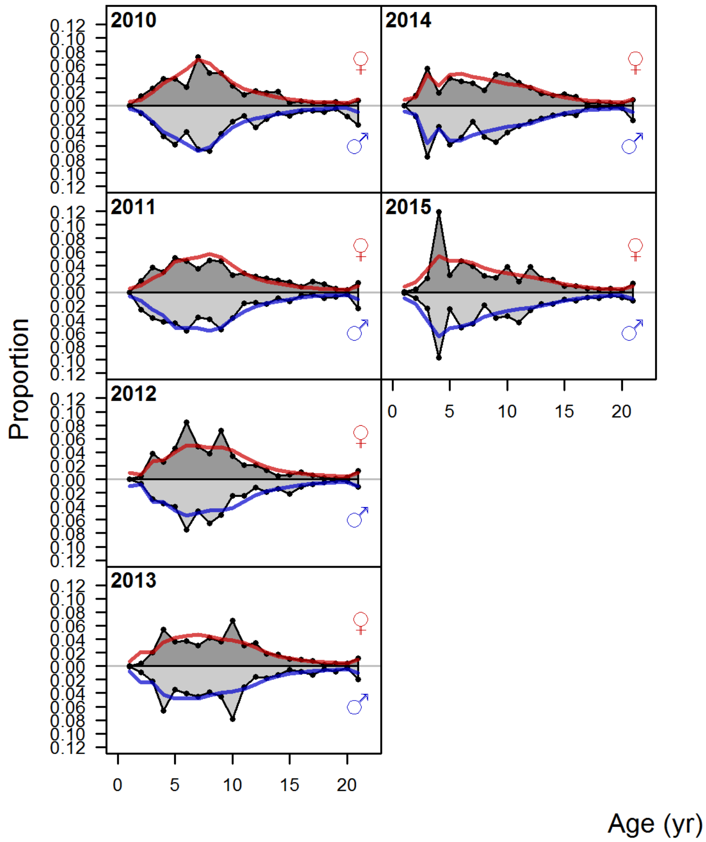


Figure 33, continued (part 2 of 2).

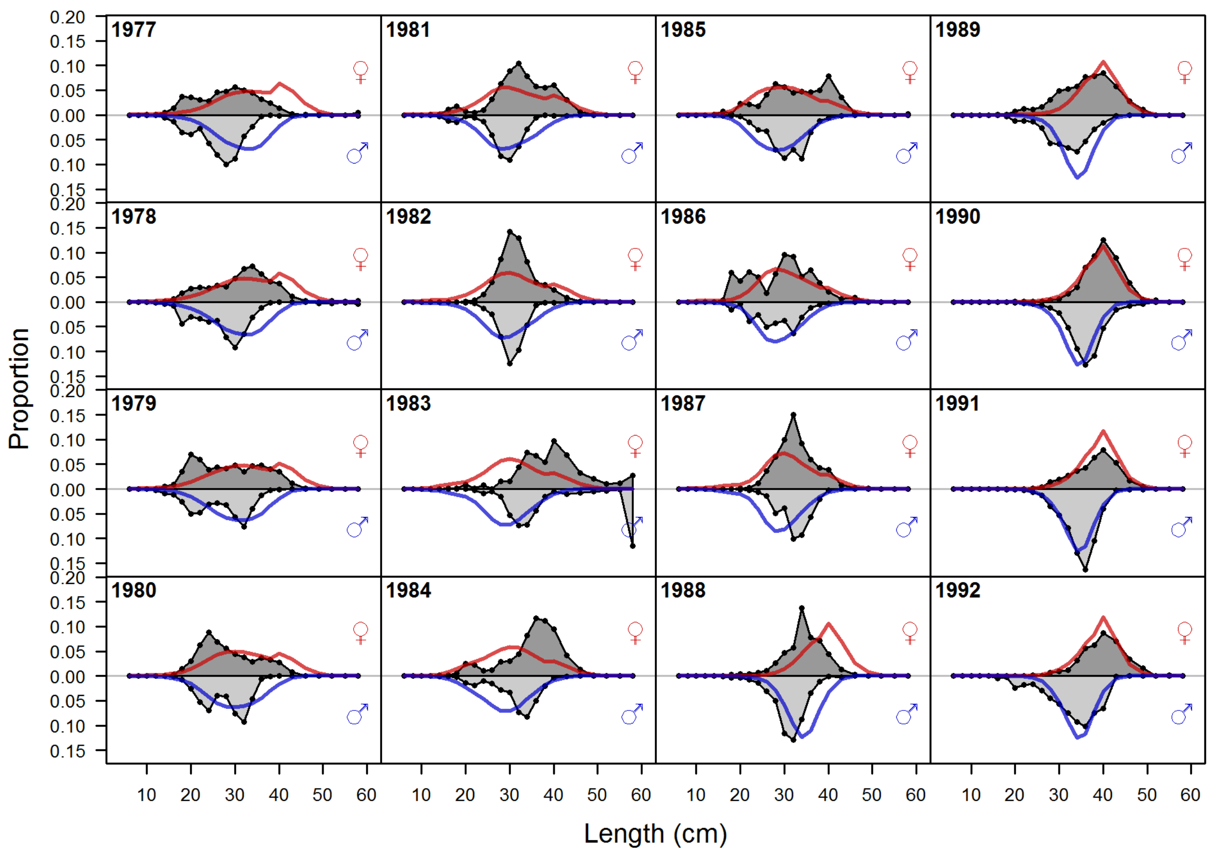


Figure 34. Fits to fishery length composition data (grey) by year for Model 2018.0b. In years for which age composition data existed, length composition data were given an effective sample size of 1 (part 1 of 2).

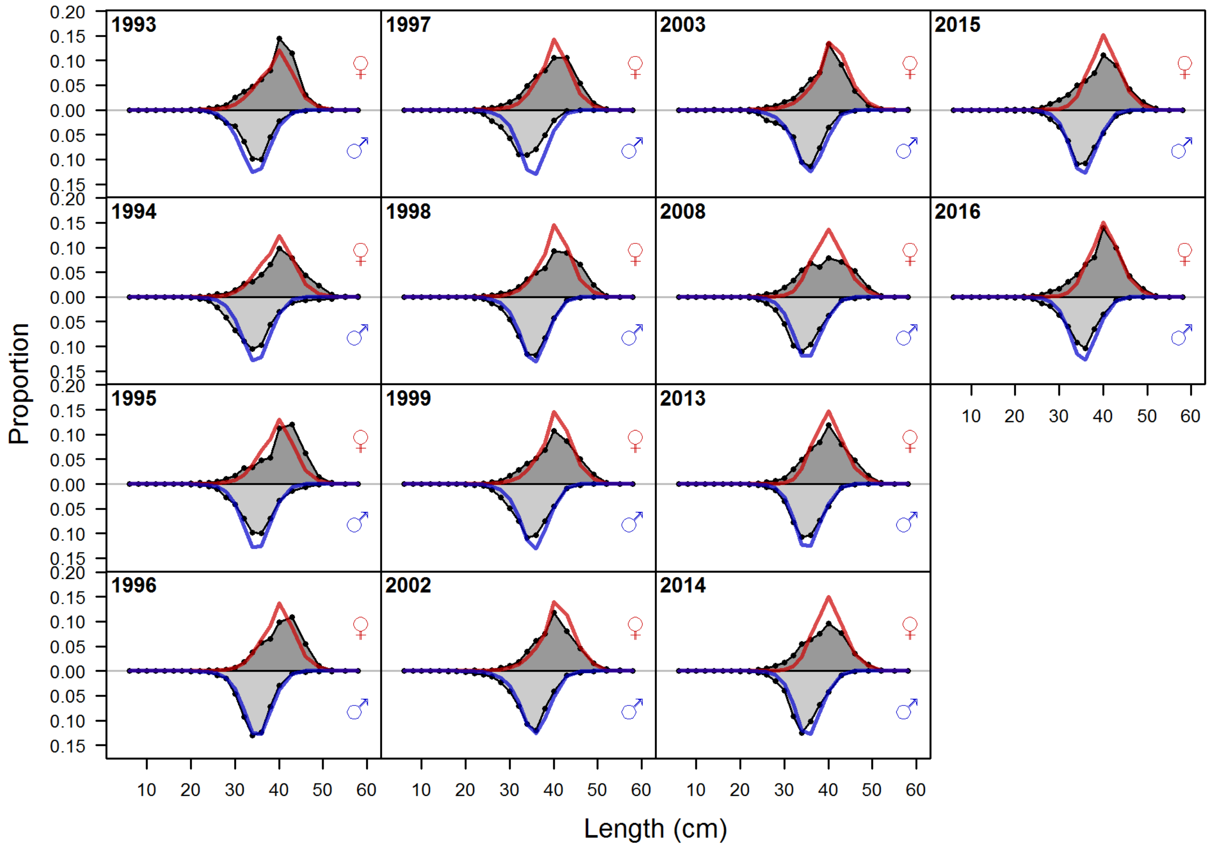


Figure 34, continued (page 2 of 2)

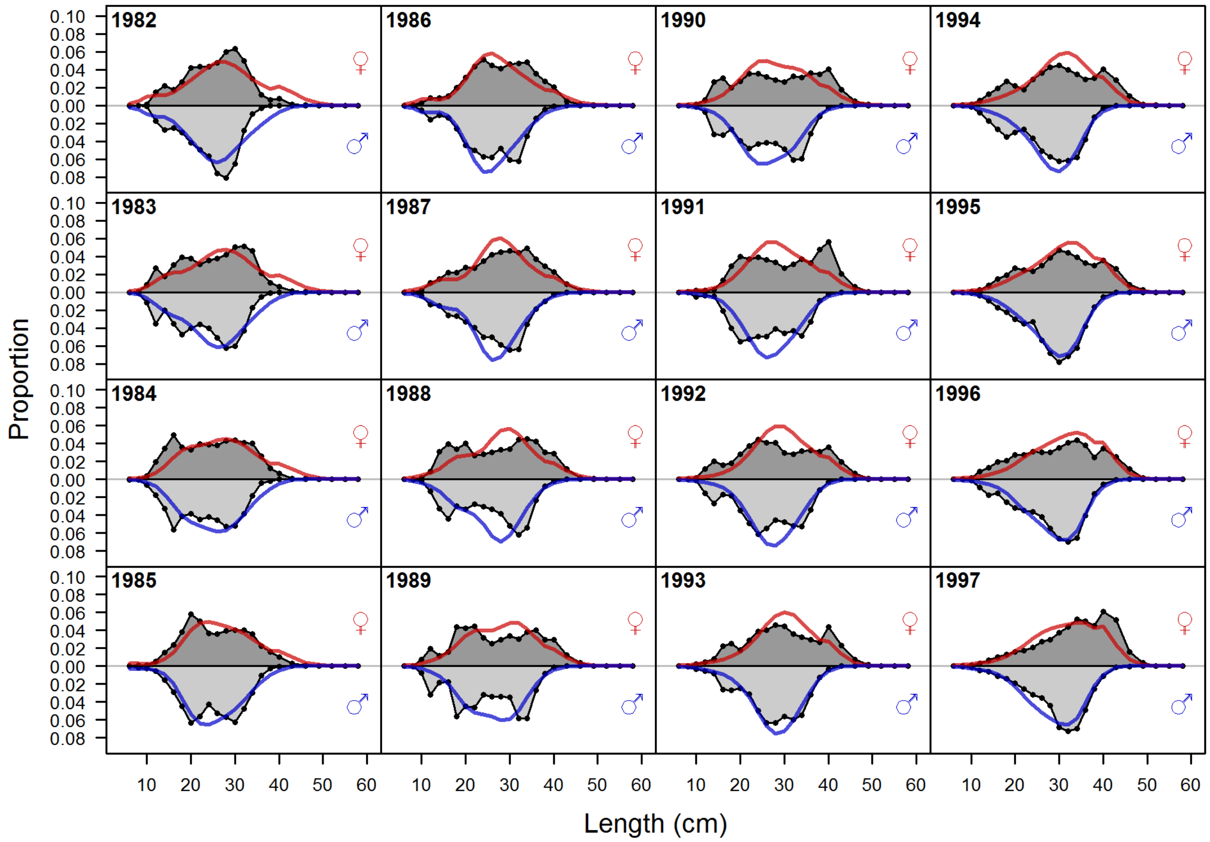


Figure 35. Fits to survey length composition data (grey) by year for Model 2018.0b. In years for which age composition data existed, length composition data were given an effective sample size of 1 (part 1 of 3).

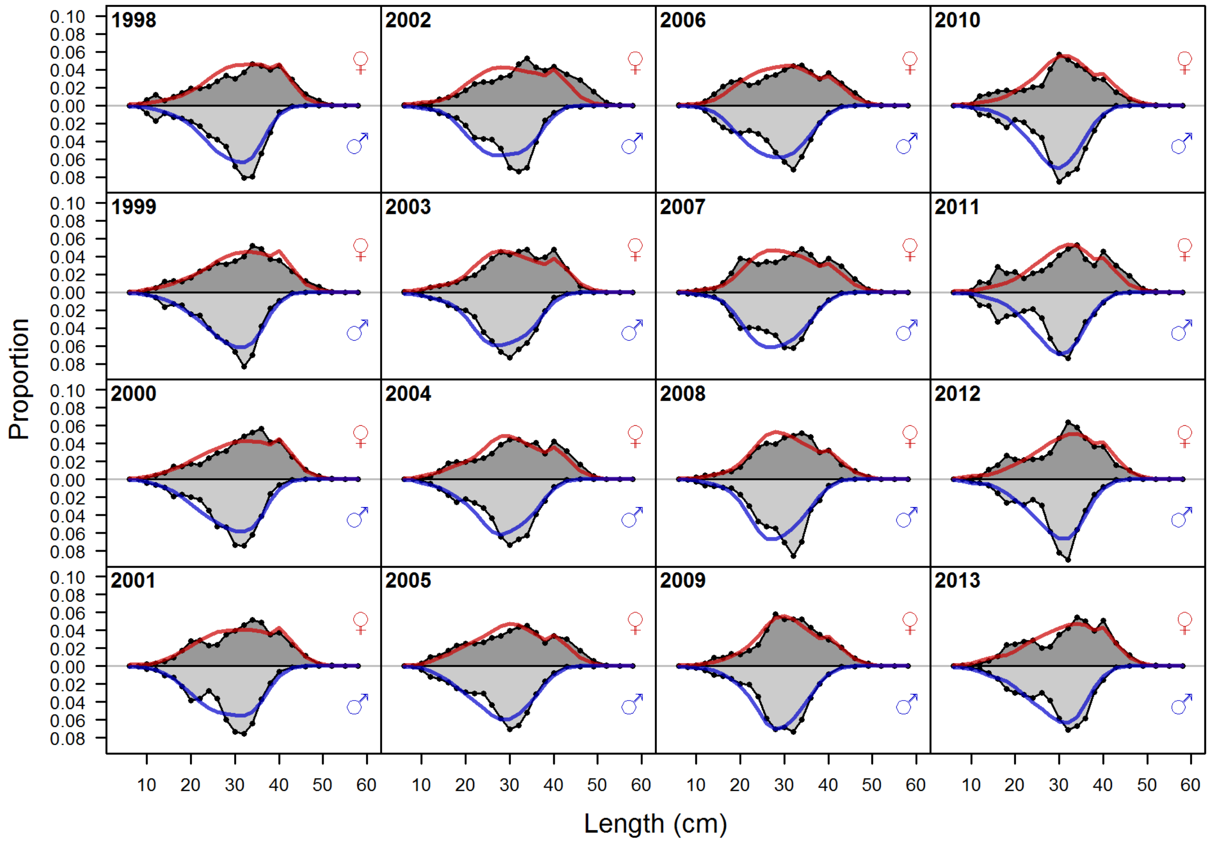
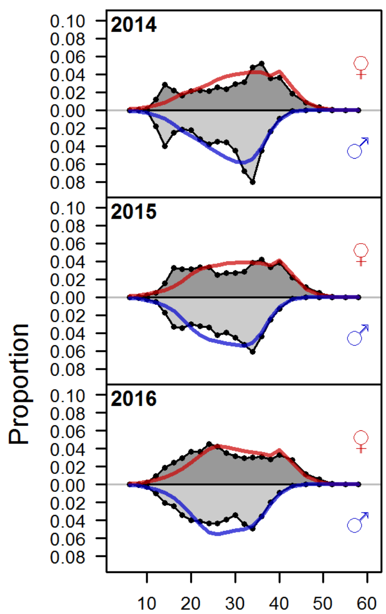


Figure 35, continued (part 2 of 3).



Length (cm)

Figure 35, continued (part 3 of 3).