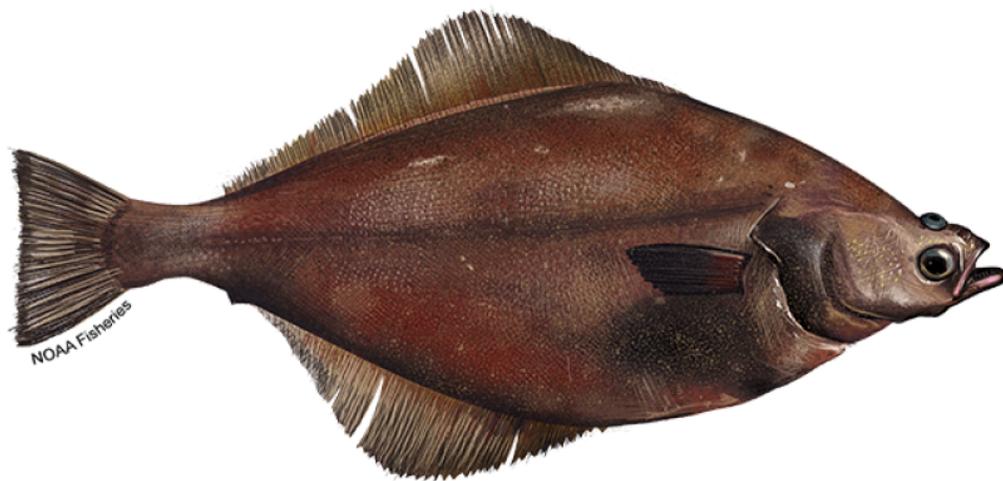


7. Assessment of the arrowtooth flounder stock in the Gulf of Alaska - Data Review

S. Kalei Shotwell, Grant Adams, Dana Hanselman

September 2025



With Contributions from:

Ned Laman, Beth Matta, Margaret Siple, Jane Sullivan

This report may be cited as: Shotwell, S.K., Adams, G. and Hanselman, D. 2025. Assessment of the arrowtooth flounder stock in the Gulf of Alaska - Data Review. North Pacific Fishery Management Council, Anchorage, AK. Available from <https://meetings.npfmc.org/Meeting/Details/3099>.

This information is distributed solely for the purpose of pre-dissemination peer review under applicable information quality guidelines. It has not been formally disseminated by the National Marine Fisheries Service and should not be construed to represent any agency determination or policy.

Table of Contents

7. Assessment of the arrowtooth flounder stock in the Gulf of Alaska - Data Review.....	1
Table of Contents.....	2
Executive Summary.....	4
Summary of Changes in Assessment Inputs:.....	4
Changes in the input data:.....	4
Changes in the assessment methodology (see Appendix for bridging of these models):	4
Responses to SSC and Plan Team Comments:.....	5
General Stock Assessment Comments:.....	5
Comments Specific to this Assessment:.....	5
Introduction.....	6
Data.....	6
Fishery Data.....	7
Catch.....	7
Fishery Age and Length Compositions.....	7
Survey Data.....	8
AFSC Bottom Trawl Survey Biomass Estimates.....	8
AFSC Bottom Trawl Survey Age and Length Compositions.....	8
Other Data.....	9
AFSC Longline Survey.....	9
International Pacific Halibut Commission Survey.....	9
Analytical Approach.....	10
Parameters Estimated Outside the Assessment Model.....	10
Natural Mortality.....	10
Growth.....	11
Length at Age.....	11
Weight at Age.....	11
Ageing Error.....	12
Maturity.....	13
Comparison of Data Models.....	13
Description of Alternative Models.....	14
Conclusion.....	14
Acknowledgements.....	14
References.....	15
Tables.....	18
Figures.....	20
Executive Summary.....	43
Background.....	43
Justification.....	44
Methods.....	45
Data.....	46

Model Structure.....	46
Models.....	47
Results.....	47
Discussion and Recommendation.....	48
Code and data.....	49
References.....	49
Tables.....	52
Figures.....	60
Supplementary Material.....	75
Model bridging.....	75
Results.....	75

Executive Summary

The following substantive changes have been made in the Gulf of Alaska (GOA) arrowtooth flounder assessment relative to the last full stock assessment and fishery evaluation (SAFE) report (Shotwell et al., 2021). This report focuses on data updates since the bridging exercise, completed in 2024 (Appendix 7A), and provides the list of alternative models accepted by the GOA Plan Team and Scientific and Statistical Committee (SSC).

Summary of Changes in Assessment Inputs:

Changes in the input data:

- Updated estimates of catch through August 23, 2025.
- Updated all fishery length compositions and included new compositions 2021-2024
- Updated all AFSC trawl survey biomass point-estimates and standard errors and included new estimates from 2023
- Updated all AFSC trawl survey age compositions and included new compositions from 2021, 2023
- Removed AFSC trawl survey estimates and associated age compositions from non-standardized AFSC surveys and during years with low confidence in arrowtooth identification (pre-1993)
- Removed all fishery length compositions prior to the Observer Program (pre-1991) and years with fewer than 300 samples
- Changed fishery length composition bins to match observed length distribution
- Updated the length-at-age transition matrices with new bin structure and data through 2024
- Updated the weight-at-age vectors with data through 2024
- Updated the aging error matrix with all available age precision information (through 2021)

Changes in the assessment methodology (see Appendix for bridging of these models):

- Model software change from Automatic Differentiation Model Builder (ADMB) to Template Model Builder (TMB) within the Rceattle platform.
- Model 25.0: single-species TMB based Rceattle model with data improvements that fixes sex-specific M (females = 0.2 and males = 0.35) and treats annual recruitment as random effects.
- Model 25.1: Model 25.0 but estimates sex-specific M instead of fixing M.
- Research: Cannibalism-enhanced single species model which is the multi-species TMB based Rceattle that estimates sex-, age-, and time-varying M due to cannibalism from arrowtooth (M2), sex-specific residual mortality (M1), and treats annual recruitment as random effects.

Responses to SSC and Plan Team Comments:

General Stock Assessment Comments:

This section will be completed for the November assessment.

Comments Specific to this Assessment:

"There appears to be a shift to lower recruitment in recent years, beginning in 2006 (i.e., the 2005 year class). The Team recommends investigating whether these lower recruitments are related to environmental conditions in the GOA."

"The Team noted that the decrease in biomass began before the recent heatwaves in the north Pacific and is similar to drops observed in other flatfish during this time and may be potentially linked to extended poor recruitment during cold pattern in 2006-2007." (GOA Plan Team, November 2019)

We investigate these lower recruitment trends through the Ecosystem and Socioeconomic Profile (ESP) framework and provide a draft ESP at the September Plan Team for review of progress. An external Dynamic Structural Equation Model (DSEM) has been developed for this stock and is included in the draft ESP report. We plan to run an integrated DSEM within the Rceattle platform to present as a research ecosystem model in the final ESP to be submitted for review at the November Plan Team.

"The Team also noted the potential of using AFSC longline survey data for Arrowtooth flounder as they are caught in significant numbers on that survey."

"The SSC requests the authors investigate including IPHC survey data in this assessment, and whether fishery catch-at-age information is available for inclusion in the model." (SSC, December 2019)

We will provide a review of the available data from the Alaska Fisheries Science Center (AFSC) longline survey, the International Pacific Halibut Commission (IPHC) survey, and fishery ages in the full November assessment. At this time we provide the review from the 2021 SAFE assessment as a starting point. We also have information from a nearshore survey from the Alaska Department of Fish and Game (ADF&G) that is currently reported in the ESP. We may include this in the potential list of additional surveys in the "Other Data" section.

"The Team recommended that the TMB version of the single species model be used in 2025 and brought forward to the September meeting." (GOA Plan Team, September 2024)

"The SSC supports the GOA GPT recommendation that the bridging exercise was sufficient to go forward with the TMB single-species model using CEATTLE in 2025." (SSC, October 2024)

The GOA Plan Team and the SSC both recommend the transition of the GOA arrowtooth flounder model in ADMB to TMB via the Rceattle framework. They also determined that the bridging exercise we provided in September and October of 2024 was sufficient and we did not need to bring forward the ADMB model prior to the full assessment in 2025. There was a bit of confusion in the above recommendation regarding the need to present the models again that were already recommended. In an effort to not be redundant, but also provide documentation of our progress to date, we provide a review of the data improvements to be implemented in the full assessment for 2025 and include the 2024 bridging document as an appendix.

Introduction

Arrowtooth flounder (*Atheresthes stomias*) are relatively large flatfish that range from central California to the eastern Bering Sea (EBS), and as far west as the Kuril Islands (Orlov 2004). Arrowtooth flounder occur in waters from about 20m to 800m, although catch per unit effort (CPUE) from survey data is highest between 100m and 300m. Spawning occurs in deep water (>400 meters) in the Gulf of Alaska GOA and along the shelf break in the eastern Bering Sea (Doyle et al. 2018). Migration patterns are not well known for Arrowtooth flounder; however, there is some indication that Arrowtooth flounder move into deeper water as they grow, similar to other flatfish (Zimmerman and Goddard 1996). Fisheries data off Washington suggest that larger fish may migrate to deeper water in winter and shallower water in summer (Rickey 1995). Arrowtooth flounder spawn in deep waters (>400m) along the continental shelf break in winter (Blood et al. 2007). They are batch spawners, spawning from fall to winter off Washington State at depths greater than 366m (Rickey 1995).

The age composition of the species shows fewer males relative to females as fish increase in age, which suggests higher natural mortality (M) for males (Wilderbuer and Turnock 2009). To account for this process, natural mortality has typically been fixed at 0.2 for females and 0.35 for males in the model. Different options for natural mortality were considered in the 2017 assessment, which evaluated natural mortality as a function of the size of the fish (Charnov 1982, Gislason et al. 2010, Lorenzen 1996). The distribution of ages appears to vary by region and sex. Males were not observed older than age 20 prior to 2005 in the GOA; however, males age 21 have been observed in every survey since that time. The sex ratio of arrowtooth flounder also varies by region. In the GOA, the observed ratio from fishery observer length frequency collections is 69% female, 31% male. Survey length compositions from the Bering Sea indicate that the proportion female is 70% on the Bering Sea shelf, 72% on the Bering Sea slope, and 62% in the Aleutian Islands. Information concerning the genetic stock structure of arrowtooth flounder is not currently available, although efforts are underway to initiate research.

Data

New data used in this assessment include estimates of total catch, trawl survey biomass estimates and standard errors from the Alaska Fisheries Science Center (AFSC) GOA bottom trawl survey, and sex-specific trawl survey age and fishery length-frequencies from observer sampling. Length composition data were available from each survey; however, length data are only used in the model for each year when age composition data are not available. Age compositions for GOA arrowtooth flounder were available or will be available for all survey years; therefore survey length compositions are not used in the model. They are only included in this document for illustration purposes. The model simulates the dynamics of the population and compares the expected values of the population characteristics to data observed from AFSC surveys and fishery sampling programs.

The following data sources were available for use in the model. Bolded years indicate new data inputs for this year's assessment. Please note, at this time the most recent AFSC bottom trawl survey estimate was not available but this will be included in the full November assessment.

Source	Data	Years
AFSC GOA bottom trawl survey	Survey biomass and standard error	1993,1996,1999,2001,2003,2005,2007,2009 2011,2013,2015,2017,2019,2021, 2023,2025
	Age Composition	1993,1996,1999,2001,2003,2005,2007,2009 2011,2013,2015,2017,2019, 2021,2023
U.S. fisheries	Catch Biomass	1977-2021, 2022-2025
	Length composition	1991-2020, 2021-2024

*Note catch is current up to 8/23/2025

Fishery Data

Catch

Joint venture fisheries started in the late 1970s and the U.S. domestic Observer Program was phased in starting in 1987 and was completely domestic by 1990. Reliable cruise and vessel records started in 1991. Arrowtooth flounder catch was available from observer at-sea sampling applied to the Alaska Regional Office blend estimates for 1991 to the present. The estimate of annual arrowtooth flounder catch between 1960 and 1990 was extrapolated from total flatfish catch by multiplying the proportion of arrowtooth in observer sampled flatfish catches (nearly 50%) by the reported flatfish catch (1960-1977 from Murai et al. 1981 and 1978-1993 from Wilderbuer and Brown (1993) (Table 7.1).

Fishery Age and Length Compositions

Otoliths have been collected sporadically in the fishery since 1982 but sample sizes are generally low following the initiation of the Observer Program in 1990 (Table 7.2). It may be possible to age some of these otoliths during years when the samples were higher (>100) through the AFSC Age and Growth prioritization system; however, the ageing request would need to be evaluated within the scope of the AFSC Age and Growth available staff time and resources (J. Short and B. Matta, *pers. commun.*). Also, the years when otoliths were higher are fairly sporadic, and aging of these otoliths may not be worth the extra effort given the amount of otoliths aged by the AFSC Age and Growth program each year. Future ageing of available otoliths for this stock will be determined based on the AFSC and the Age and Growth Program priorities.

The number of fisheries otoliths and length observations taken by fisheries observers by year, 1991 to present, are provided in Table 7.2. Sample sizes (number of individual fish) for the fishery length data by sex were under 300 in 1992, 1994, 1996, and 1998; therefore, we do not use this information in the model. The distribution of male and female length observations over all available years since the Observer Program began are shown in Figure 7.1. There are no clear directional long-term trends in the length data but there is variation over time. Also data between 1991 and early 2000s appears to contain a lot of larger females than any following year in the time series. Observers did not consistently begin identifying arrowtooth flounder and Kamchatka flounder to species until 2008, and the catch accounting system did not speciate Kamchatka flounder until 2011 (Bryan et al., 2020). Generally there are very few Kamchatka flounder caught in the GOA bottom trawl surveys, but there may be some in the fishery and we plan to investigate this further to determine if this may be a reason to not use this early fishery length data prior to consistent identification in 2008. We will include results of this investigation in the November assessment and may further limit the fishery length composition used in the model.

Survey Data

AFSC Bottom Trawl Survey Biomass Estimates

Bottom trawl surveys were conducted on a triennial basis in the Gulf of Alaska starting in 1984 through 1999 and became biennial starting in 2001. The triennial AFSC surveys used a nor'easter trawl. A different survey design was used in the eastern GOA in 1984 and much of the survey effort in the western and central GOA in 1984 and 1987 was conducted by Japanese vessels that used a different net design than the U.S. vessels. Additionally, the trawl used in the 1984 and 1987 surveys had no bobbin or roller gear, which would cause the gear to be more in contact with the bottom than current trawl gear, and may have restricted the locations of trawl sites to more trawlable areas. The triennial surveys covered all areas of the GOA out to a depth of 500 m (in some surveys to 700 m or 1,000 m), but the 2001 biennial survey did not sample the eastern GOA. We use an average estimate of the previous three surveys (1993, 1996, 1999) to estimate the eastern GOA biomass for 2001 as in past assessments. Analyses of species identification confidence for the bottom trawl survey data (Stevenson and Hoff, 2009; Orr et al., 2014) indicated that a moderate level of confidence was attained for arrowtooth flounder starting in 1980 for the Gulf of Alaska and Aleutian Islands survey and in 1992 for the eastern Bering Sea shelf survey. Additionally, recently survey strata adjustments have occurred in a data cleaning effort of the bottom trawl survey data and consistency in species identification confidence estimates have resulted in data only available since 1992 for arrowtooth flounder. There is also now a new species code for unidentified arrowtooth/Kamchatka flounder which has been applied to data prior to 1992. We, therefore, use information starting from 1993 for the GOA bottom trawl surveys.

The survey catchability coefficient (q) in the assessment model was assumed to be 1.0. AFSC has conducted studies to estimate the escapement under the survey net and herding of fish into the net. The percent of arrowtooth flounder caught that were in the path of the net varies by size from about 80% at 27 cm (about age 3) to about 96% at greater than 45cm (equal to or greater than age 7 for females and age 10 for males) (Somerton et al. 2007). Somerton et al. (2007) estimated the effect of herding combined with escapement under the net to be an effective multiplier of about 1.3 on survey catch for arrowtooth flounder. The combination of escapement under the net and herding into the net indicates that abundance would be about 23% less than the estimated survey abundance. At this time we assume q to be 1.0 but may explore estimating q in the future given this herding experiment.

Survey biomass estimates increased to 3 million tons in 2003, and have since declined to approximately 1 million metric tons from 2017 to present. Survey biomass has generally been declining since 2003, and the 2017 estimate was the lowest estimate on record. The 2019 to the present estimates were each slightly higher than the previous survey (Figure 7.2). Survey estimates of biomass by area are generally highest in the central GOA (Chirikof and Kodiak regions), and the eastern (southeastern and Yakutat) and western GOA (Shumagin) have similar biomass of arrowtooth flounder (Figure 7.3). The Kodiak areas has experienced the greatest declines in arrowtooth flounder biomass since 2003, but has recovered slightly since 2017.

AFSC Bottom Trawl Survey Age and Length Compositions

Otoliths from the 1993 to 2023 AFSC trawl surveys have been aged and are used in the model. The most recent survey ages are not yet available but have been requested for priority ageing. Differences in ageing methodology exist but are not expected to bias results (D. Anderl, *pers. commun.*). Length composition data are not used when age data are available or anticipated to be available in the following assessment year. Length frequency data were collected opportunistically for arrowtooth flounder on three GOA surveys conducted in 1985, 1986, and 1989. These surveys were not part of the standard AFSC GOA bottom surveys but the length frequency data have been included in previous assessments. As these opportunistically collected length compositions are not part of the standard survey protocol, we no longer

use this information in the model. Length frequency data from all AFSC surveys indicates no long term trends, and that females are larger than males (Figure 7.4a,b).

Other Data

AFSC Longline Survey

The AFSC longline survey has been conducted annually since 1988, and relative population numbers (RPNs) and relative population weights (RPWs) have been computed for each year and are available since 1992 for Arrowtooth and Kamchatka flounder combined (Figure 7.5). The AFSC longline survey is conducted annually over the continental slope region of the BSAI and the GOA. The GOA stations are sampled each year while the Bering Sea is sampled on odd years and the Aleutian Islands in even years. This survey provides data on the relative abundance of Arrowtooth flounder in the form of RPNs and RPWs for fish on the continental slope as indices of stock abundance. Relative population abundance indices are computed annually using survey catch per unit of effort (CPUE) rates that are multiplied by the area size of the stratum within each geographic area. These relative population indices are available by numbers (RPN) and weights (RPW) for a given species (Rodgveller et al. 2011). The survey is primarily directed at sablefish, but also catches considerable numbers of arrowtooth flounder. Also, historically, arrowtooth flounder and Kamchatka flounder were not separated by species and were just recently separated out in 2019. Therefore, we provide RPNs for arrowtooth and Kamchatka flounder combined to see changes throughout the entire time series (Figure 7.5). Results for this survey concerning flatfish, should also be viewed with some caution, as the RPNs and RPWs do not take into account possible effects of competition for hooks with other species caught on the longline, especially sablefish.

RPNs in the GOA show a somewhat decadal cyclic pattern since the mid-1990s to about 2010 and then have declined to present low values, with a steep decline at the onset of the 2014 marine heatwave, similar to the bottom trawl survey estimates. Values range from a high in 1999 to a low in 2020. This same pattern is evident in the BSAI time series for Arrowtooth and Kamchatka flounder except the pattern is more variable and the decline is less steep. Some of the fluctuations may be related to changes in the abundance of sablefish regarding competition for hooks among species. The 2021 longline survey RPN value for Arrowtooth and Kamchatka flounder combined is up 64% from 2020 (Figure 7.5), but is still 71% below the long term mean of the time series.

Length data are also collected for Arrowtooth flounder during longline surveys and compositions are available since 1992. A clear shift in size has occurred throughout the time series with increasing abundance of larger fish sampled until the mid-2000s, after which a shift to small fish occurs until about 2017. In recent years there are fewer fish in the survey which may have to do with the declines in the population and less of the stock in the slope environment where the survey primarily samples.

International Pacific Halibut Commission Survey

The International Pacific Halibut Commission (IPHC) conducts a longline survey each year to assess Pacific halibut. This survey differs from the AFSC longline survey in gear configuration and sampling design, but also catches arrowtooth flounder. More information on this survey can be found in Soderlund et al. (2009). A major difference between the AFSC and the IPHC surveys is that the IPHC survey samples the shelf consistently from 1-500 meters, whereas the AFSC longline survey samples the slope and select gullies from 150 to 1000 meters. Because the majority of effort occurs on the shelf in shallower depths, the IPHC survey samples more suitable arrowtooth flounder habitat than the AFSC longline survey and is similar to the AFSC bottom trawl survey; however, lengths of arrowtooth flounder are not taken on the IPHC survey.

RPNs have been computed for each year of the IPHC survey and are available since 1998 to 2021 for arrowtooth flounder (Figure 7.6). RPNs in the GOA have ranged from a low RPN in 2017 to a high in 2000 and also generally follow the trajectory of the AFSC bottom trawl survey since 2005 when the population started to decline. RPNs are generally higher in the CGOA as with the AFSC bottom trawl survey. RPNs increased in 2021 in all areas except the WGOA when compared to 2020. The 2021 GOA estimate is 44% above the 2019 estimate (2020 was not completely sampled) and is now 48% below the long-term average for the time series. No length data are collected for arrowtooth flounder on the IPHC survey. Changes to the sampling protocol and coverage of the IPHC longline survey have recently occurred. As such, we do not recommend using this survey in the model as these changes will limit the survey utility moving forward.

Analytical Approach

We use a previously developed climate-enhanced multi-species statistical catch-at-age modeling framework called CEATTLE (Climate-Enhanced, Age-based model with Temperature-specific Trophic Linkages and Energetics, Holsman et al., 2016) that has been expanded for groundfish in the GOA using TMB (Adams et al., 2022, Kristensen et al., 2016) as the basis for the new assessment model and alternative explorations for GOA arrowtooth flounder. Developed in part from the underpinnings of multi-species statistical catch-at-age analysis (MSCAA, Jurado-Molina et al., 2005) and multi-species virtual population analysis (MSVPA; Jurado-Molina et al., 2005; Magnusson, 1995), CEATTLE links single-species age-structured models (Appendix 7a, Table 1) through predation mortality conditioned on the temperature-dependent bioenergetic demand and diet-based prey-selectivity patterns of predators (Curti et al., 2013; Holsman et al., 2016; Kinzey and Punt, 2009). Predation mortality assumes a Holling Type II functional response and parameters are either pre-specified or estimated by fitting to survey and fishery data (Appendix 7A, Table 2). CEATTLE can be run in single-species mode by “turning-off” time-varying predation mortality or in multi-species mode by “turning-on” time-varying predation mortality. Please see the GitHub repository on Rceattle (<https://github.com/grantdadams/Rceattle>) for more details.

Parameters Estimated Outside the Assessment Model

Natural Mortality

Natural mortality (M) rates for Gulf of Alaska Arrowtooth flounder were estimated using the methods of Wilderbuer and Turnock (2009). A higher natural mortality for males than females was used to fit the age and size composition data, which are about 70% females. A value of $M=0.35$ for males was chosen so that the survey selectivities for males and females both reached a maximum close to 1.0. A likelihood profile on male natural mortality resulted in a mean and mode of 0.354 with 95% confidence intervals of 0.32 to 0.38 (Turnock et al. 2002, Figure 10.14). Model runs examining the effect of different natural mortality values for male Arrowtooth flounder can be found in the Appendix of the 2000 SAFE (https://www.afsc.noaa.gov/REFM/stocks/Historic_Assess.htm). Differential natural mortality by sex can be a factor that needs consideration in management of targeted fish stocks, however, since GOA arrowtooth flounder is currently exploited at low levels, this effect is not a concern for this stock (Wilderbuer and Turnock 2009).

A catch-curve analysis was conducted to provide sex-specific estimates of total mortality (Z) from AFSC bottom trawl survey specimen data from 1993 - 2023 (Figure 7.7). Numbers-at-age were tallied separately for males and females. To reduce the influence of selectivity and variable year-class strength at younger ages, the analysis was restricted to ages 6 and older. At older ages, bins with very low sample sizes (<5 fish) were excluded to avoid undue leverage. This analysis of specimen data suggests a total mortality of 0.22 for females and 0.28 for males which is in line with the previous estimates of natural mortality

(Wilderbuer and Turnock, 2009) given the low exploitation rate of arrowtooth flounder. This also suggests that the natural mortalities estimated in the CEATTLE model (Appendix 7A) might be too high to be biologically plausible.

Growth

Growth was estimated from length and age data from AFSC GOA bottom trawl surveys from 1993 to 2023 and incorporated in the assessment using a length-age conversion matrix.

Length at Age

There are two length-age conversion matrices (one for females and one for males) that are generated from AFSC GOA bottom trawl survey specimen data from 1993 to 2023.

Length bins were defined in 2 cm increments spanning from 20 cm to 75 cm, with the uppermost bin treated as a plus group (≥ 75 cm). This is a change from the non-sequential length bin structure initiated at the start of this assessment in 2000 (Turnock et al., 2002). We determined that this new length bin structure adequately captures the distribution of length samples present in the fishery compositions, which are the only length compositions used in this model at this time. The previous length bin structure was initiated when fewer ages were available for this model and it relied more on the survey length compositions. Age bins were defined for ages 1–20, with an additional plus group for ages 21 and older. These are the same age bins that have been used since first initiated in 2015 (Spies et al., 2015).

For each sex, mean length-at-age and standard deviations were estimated from survey specimen data. The distribution of length at each age was assumed to follow a normal distribution with these estimated means and standard deviations. Conditional probabilities for each age–length cell were computed by integrating the normal distribution between successive bin boundaries. Specifically, the probability of observing length bin l given age a was calculated as:

$$P(L \in \text{bin } l | a) = \Phi(L_{l+1}; \mu_a, \sigma_a) - \Phi(L_l; \mu_a, \sigma_a)$$

where Φ is the cumulative normal distribution function, L_l is the lower boundary of length bin l , L_{l+1} is the boundary of the next bin, μ_a is the mean length at age a , and σ_a is the standard deviation at age a . The first bin was defined from $-\infty$ to the second bin boundary, and the final bin was defined from the last boundary to $+\infty$ to form a proper probability distribution. Matrices were calculated separately for males and females to account for sex-specific growth differences (Figure 7.8a,b).

Weight at Age

Weight at age used in the model is based on a von Bertalanffy relationship determined from the stratified age collection on GOA bottom trawl surveys from 1993–2023 (Figure 7.9). First, weight–length relationships were estimated separately for males and females by fitting the log–linear model

$$\ln(W) = \alpha + \beta \ln(L),$$

where W is whole body weight (kg) and L is fork length (cm). These regressions provided sex-specific scaling exponents (β) that capture allometric growth in weight relative to length.

Second, a von Bertalanffy growth function was fit to observed length-at-age for each sex from the survey specimen data used in the previous section,

$$L(a) = L_{\infty} \left(1 - e^{-k(a-t_0)} \right)$$

where a is age (years), L_{∞} is the asymptotic length, k is the growth coefficient, and t_0 is the theoretical age at zero length.

Predicted mean weight-at-age was then computed by combining the two models,

$$\hat{W}(a) = \exp(\alpha) [L(a)]^{\beta},$$

$$W_a = e^{\alpha} L_a^{\beta}$$

using the sex-specific von Bertalanffy parameters (L_{∞} , k , t_0) and the corresponding weight-length exponent (β). This two-step approach stabilizes estimation by constraining the weight-at-age function to the empirically observed weight-length relationship.

Final weight-at-age vectors were tabulated for ages 1–20, with an additional 21+ plus group. The plus group value was calculated as the weighted mean of predicted weights for all observed ages ≥ 21 , using sample frequencies as weights. These vectors were used as model input in the assessment (Figure 7.10).

Ageing Error

An ageing error transition matrix was first implemented for GOA arrowtooth flounder in 2017 and used the estimated percent agreement among two age readers for age samples from 1991-2015. Ageing for arrowtooth flounder began in the early 1990s in the Bering Sea and the GOA ageing started several years later. The initial legacy matrix for GOA arrowtooth flounder includes data from the Bering Sea, possibly due to limited data for the GOA at the time this matrix was included in the model. This initial legacy matrix was computed using a model that incorporates a linear increase in the standard deviation of ageing error and assumed ageing error was normally distributed (Dorn et al., 2003, Methot, 2000). Percent agreement was predicted by the sum probability that both readers were correct, that both readers were off by one year in the same direction, and the probability that both age readers were off by two years in the same direction (Methot, 2000).

For 2025, we updated our ageing error matrices with new reader testing data up to 2021 and only using data from the GOA. Sample sizes of the raw paired age readings from the GOA bottom trawl survey are provided in Figure 7.11. There are four possible age fields for all ageing collections: `read_age`, `test_age`, `update_age`, and `final_age`. `read_age` is the estimate made by the main reader for a given collection, and `test_age` is the estimate made by a second independent reader to assess precision and relative bias between readers. `final_age` is the final resolved age that is passed along to end users and is usually the same as the `read_age`. The `read_age` can differ from the `final_age` for a number of reasons such as if the reader and tester resolved discrepancies to the tester age or a new third age, or if the reader re-examined the otolith and determined that their initial age was wrong (as sometimes happens during outlier evaluation). The `update_age` captures how discrepancies were resolved and whether outliers were re-examined (Beth Matta, *pers. commun.*).

We used the `test_age` and the `final_age` data and evaluated three approaches for constructing the ageing error matrix (AEM) used to account for misclassification of otolith ages: (1) an empirical matrix derived directly from paired test-final readings on recent collections, (2) a modeled matrix following the framework of Candy et al. (2012), and (3) the legacy matrix previously used in the assessment (Figure 7.12a,b). The empirical matrix is simple and transparent, directly reflecting observed discrepancies between independent test ages and final resolved ages. It shows relatively high agreement on the diagonal

for younger ages, with error increasing at older ages. However, the empirical matrix can be noisy in sparsely sampled age classes, particularly in the plus group. The Candy-style modeled matrix smooths the empirical error patterns by fitting statistical models to the absolute and signed error distributions. This approach enforces logical constraints (e.g., no negative ages, capping at 21+) and provides more stable estimates at older ages, where paired samples are limited. It also allows for asymmetric error, consistent with a tendency toward under-ageing at advanced ages. Future efforts may use the AFSC “read_codes” in the Candy model which it was designed to use.

The old matrix exhibited lower diagonal probabilities at ages 2–10 compared to the new empirical and Candy matrices, implying a higher assumed rate of misclassification at young ages (Figure 7.13). In contrast, both the new empirical and Candy-modeled matrices indicated stronger reader agreement at younger ages, with diagonal probabilities generally above those of the old matrix. At intermediate and older ages, all matrices showed declining correct-age probabilities with greater spread into adjacent bins, but the patterns differed. The new empirical matrix showed more variability and some asymmetry, consistent with observed under-ageing at older ages. Overall, the new empirical matrix best reflects current reader performance, showing improved agreement at young ages relative to the old matrix. The empirical matrix best represents the current observed ageing performance, while the Candy-style matrix provides a robust, smoothed alternative less affected by small sample sizes. For this assessment we recommend using the new empirical matrix as the base case, with the Candy-style matrix as a sensitivity run to evaluate robustness. Both improve upon the old matrix by incorporating updated reader performance and contemporary data.

Maturity

Maturity at age was based on a maturity-at-length study by Zimmerman (1997) through 2013. Length at 50% maturity was estimated at 47 cm with a logistic slope of -0.3429 from Arrowtooth flounder sampled in hauls from in the September 1993 bottom trawl survey (Zimmerman 1997). Elsewhere in their range, length at 50% maturity was 36.8 cm for females and 28.0 cm for males from survey data in 1992 off Washington, with logistic slopes of -0.54 and -0.893 respectively (Rickey 1995). Arrowtooth flounder had length at 50% maturity of 44 cm for females and 29 cm for males off the coast of Oregon (Rickey 1995). Spawning fish were found in depths from 108m to 360m in March to August in the Gulf of Alaska (Hirshberger and Smith 1983) from analysis of trawl surveys from 1975 to 1981. Most observations of spawning fish have been in the northeastern Gulf, off Prince William Sound, off Cape St. Elias, and Icy Bay.

A study was conducted in 2008 that examined maturity-at-age that estimates age at maturity rather than length at maturity (Stark 2008). In this study, a sample of 301 fish was taken in February 2002 and a separate collection (226 fish) was taken in July 2003, both from the central GOA. Parameter estimates based on the February sample were used in the current study because Arrowtooth flounder spawn during winter months. The estimate of logistic 50% maturity was 7 years, the logistic slope (B) was 1.3817 and the intercept (A) was -9.6183. Fish matured at a slightly younger age in the 2008 study compared to the 1997 study. This maturity ogive (Stark 2008) has been used in the model since 2015. Age at 50% maturity is age 7 in females, and is 20% in age 6 fish.

Comparison of Data Models

We provide a bridging of these data improvements with a comparison of total biomass (age 1+), spawning stock biomass, age-1 recruitment, and AFSC bottom trawl survey estimates across the sequential addition of the new data, cleaning, and updates of the transition and error matrices (Figure 7.14a-d). These initial model runs resulted in similar trajectories across models with the larger changes being at the beginning of the time series due to the reduction in fishery length compositions prior to the start of the Observer Program and successively lower period of high abundance from 2000 through about 2015 (Figure

7.14a,b). This successive decrease is also clear in the 2000 recruitment estimate (Figure 7.14c). The fit to the AFSC trawl survey biomass estimates is very consistent across models (Figure 7.14d). The terminal year estimates were fairly similar for all models. The final data model with all the new cleaned data and updated transition and error matrices will be the base model moving forward for November. We call this Model 25.0. Fit to the AFSC bottom trawl survey age compositions for this Model 25.0 (Figure 7.15a,b) was improved from last year's model (Figure 7A.4). Fit to the fishery lengths for Model 25.0 (Figure 7.16a,b) are not comparable with the 2024 bridged model because of the new length bins. However, we do see some lack of fit in the plus group which could be due to the early fishery length data when specciation was not consistently conducted. We will investigate this further for the November assessment.

Description of Alternative Models

The following are the alternative models that will be explored for the full assessment in November. Please see the Appendix 7A for the complete bridging from the last full assessment. These alternative models were approved by the Groundfish Plan Team and the Scientific and Statistical Committee (SSC) for use in the 2025 full assessment ([September GOA Plan Team Minutes 2024](#), [October SSC Minutes 2024](#)).

- Model 25.0: single-species TMB based Rceattle model with data improvements that fixes sex-specific M (females = 0.2 and males = 0.35) and treats annual recruitment as random effects.
- Model 25.1: Model 25.0 but estimates sex-specific M instead of fixing M.

We also have a research cannibalism-enhanced single species model that is the multi-species TMB based Rceattle cannibalism model that estimates sex-, age-, and time-varying M due to cannibalism from arrowtooth (M2), sex-specific residual mortality (M1), and treats annual recruitment as random effects.

Conclusion

We recommend these data improvements for the November assessment along with the models presented and approved in the 2024 bridging document (Appendix 7A). This will allow for a new model platform to be implemented and include necessary data updates to allow for the best scientific information available.

Acknowledgements

We would like to thank the many contributors to this assessment including survey biologists and fishery observers who have collected the raw data used in this assessment. Many thanks to different programs at the AFSC that responded to our questions regarding available data for GOA arrowtooth flounder. Also, we thank the Alaska Regional Office for their help with understanding recent fishery trends. We also thank our internal reviewer, Lee Cronin-Fine, for his thorough and insightful review of this report and to the GOA Groundfish Plan Team and SSC for their reviews and support.

References

- Adams, G. D., Holsman, K. K., Barbeaux, S. J., Dorn, M. W., Ianelli, J. N., Spies, I., Stewart, I. J., & Punt, A. E. (2022). An ensemble approach to understand predation mortality for groundfish in the Gulf of Alaska. *Fisheries Research*, 251(October 2021), 106303. <https://doi.org/10.1016/j.fishres.2022.106303>
- Adams, G., Holsman, K., Hulson, P., Monnahan, C., Shotwell, K., Stewart, I., and Punt, A. (2023) Multispecies model estimates of time-varying natural mortality of groundfish in the Gulf of Alaska. In: Ferriss, B.E. 2023. Ecosystem Status Report 2023: Gulf of Alaska, Stock Assessment and Fishery Evaluation Report, North Pacific Fishery Management Council, 1007 West Third, Suite 400, Anchorage, Alaska 99501.
- Anderson, P. J., & Piatt, J. F. (1999). Community reorganization in the Gulf of Alaska following ocean climate regime shift. *Marine Ecology Progress Series*, 189, 117-123.
- Barnes, C. L., Beaudreau, A. H., Hunsicker, M. E., & Ciannelli, L. (2018). Assessing the potential for competition between Pacific Halibut (*Hippoglossus stenolepis*) and Arrowtooth Flounder (*Atheresthes stomias*) in the Gulf of Alaska. *PloS one*, 13(12), e0209402.
- Blood, D. M., A. C. Matarese, and M. S. Busby. 2007. Spawning, egg development, and early life history dynamics of Arrowtooth Flounder (*Atheresthes stomias*) in the Gulf of Alaska. U.S. Dep. Commer., NOAA Prof. Pap. NMFS 7, 28 p.
- Bryan, M. Dorn, M., Stockhausen, W., Barbeaux, S., Ianelli, J., Lowe, S., McGilliard, C., Monnahan, C., Spencer, P., Spies, I. and Thompson, T. 2020. Evaluating the impact of not having recent survey data in Alaska Fisheries Science Center groundfish and crab stock assessment models.
- Candy, S.G., Nowara, G.B., Welsford, D., McKinlay, J.P., 2012. Estimating an ageing error matrix for Patagonian toothfish (*Dissostichus eleginoides*) otoliths using between-reader integer errors, readability scores, and continuation ratio models. *Fish. Res.* 115, 14–23. <https://doi.org/10.1016/j.fishres.2011.11.005>.
- Charnov, E. L. *The Theory of Sex Allocation* (Princeton Univ. Press, Princeton, 1982.
- Clark, W. G., & Hare, S. R. (2002). Effects of climate and stock size on recruitment and growth of Pacific halibut. *North American Journal of Fisheries Management*, 22(3), 852-862.
- Curti, K. L., Collie, J. S., Legault, C. M., & Link, J. S. (2013). Evaluating the performance of a multispecies statistical catch-at-age model. *Canadian Journal of ...*, 484(January), 470–484. <https://doi.org/10.1139/cjfas-2012-0229>
- Dorn, M.W., 1992. Detecting environmental covariates of Pacific whiting *Merluccius productus* growth using a growth-increment regression model. *Fishery Bulletin* 90: 260-275.
- Doyle, M.J., Debenham, C., Barbeaux, S.J., Buckley, T.W., Pirtle, J.L., Spies, I.B., Stockhausen, W.T., Shotwell, S.K., Wilson, M.T., Cooper, D.W., 2018. A full life history synthesis of Arrowtooth Flounder ecology in the Gulf of Alaska: exposure and sensitivity to potential ecosystem change. *J. Sea Res.* 142, 28–51. <https://doi.org/10.1016/j.seares.2018.08.001>.
- Ferriss, B.E. (2023) Ecosystem Status Report 2023: Gulf of Alaska, Stock Assessment and Fishery Evaluation Report, North Pacific Fishery Management Council, 1007 West Third, Suite 400, Anchorage, Alaska 99501.
- Fournier, D.A., Skaug, H.J., Ancheta, J., Ianelli, J., Magnusson, A., Maunder, M.N., Nielsen, A., and Sibert, J. 2012. AD Model Builder: using automatic differentiation for statistical inference of highly parameterized complex nonlinear models. *Optim. Methods Softw.* 27:233-249.
- Gislason, H., Daan, N., Rice, J.C. and Pope, J.G., 2010. Size, growth, temperature and the natural mortality of marine fish. *Fish and Fisheries*, 11(2), pp.149-158.
- Hirshberger, W.A., and G.B. Smith. 1983. Spawning of twelve groundfish species in the Alaska and Pacific coast regions, 1975-81. U.S. Dep. Commer., NOAA Tech. Memo. NMFS F/NWC:44, 50p.
- Holsman, K. K., & Aydin, K. (2015). Comparative methods for evaluating climate change impacts on the foraging ecology of Alaskan groundfish. *Marine Ecology Progress Series*, 521, 217-235.

- Holsman, K. K., Ianelli, J., Aydin, K., Punt, A. E., & Moffitt, E. A. (2016). A comparison of fisheries biological reference points estimated from temperature-specific multi-species and single-species climate-enhanced stock assessment models. *Deep Sea Research Part II: Topical Studies in Oceanography*, 134, 360–378. <https://doi.org/10.1016/j.dsr2.2015.08.001>
- Holsman, K. K., Barnes, C., Aydin, K., Laurel, B., Hurst T., & Heintz T. (2023). Patterns in Foraging and Energetics of Walleye Pollock, Pacific Cod, Arrowtooth Flounder, and Pacific Halibut. In: Elizabeth Sidden (Ed.), *Ecosystem Status Report 2023: Eastern Bering Sea, Stock Assessment and Fishery Evaluation Report*, North Pacific Fishery Management Council, 1007 West 3rd Ave., Suite 400, Anchorage, Alaska 99501.
- Hulson, P.-J. F., Barbeaux, S. J., Ferriss, B., Echave, K., Nielsen, J., Shotwell, S. K., Laurel, B., Spies, I., Litzow, M., Mcdermott, S., Rand, K., Levy, C., Wang, M., & Rogers, L. (2023). *Assessment of the Pacific cod stock in the Gulf of Alaska Pacific cod*.
- Jurado-Molina, J., Livingston, P. A., & Ianelli, J. N. (2005). Incorporating predation interactions in a statistical catch-at-age model for a predator-prey system in the eastern Bering Sea. *Canadian Journal of Fisheries and Aquatic Sciences*, 62(8), 1865–1873. <https://doi.org/10.1139/f05-110>
- Kinzey, D., & Punt, A. E. (2009). Multispecies and Single-Species Models of Fish Population Dynamics: Comparing Parameter Estimates. *Natural Resource Modelling*, 22(1), 67–104.
- Kristensen, K., Nielsen, A., Berg, C. W., Skaug, H., & Bell, B. M. (2016). TMB: Automatic Differentiation and Laplace Approximation. *Journal of Statistical Software*, 70(5), 1–21. <https://doi.org/10.18637/jss.v070.i05>
- Lorenzen, K., 1996. The relationship between body weight and natural mortality in juvenile and adult fish: a comparison of natural ecosystems and aquaculture. *Journal of fish biology*, 49(4), pp.627-642.
- Magnusson, K. (1995). An overview of the multispecies VPA—theory and applications. *Reviews in Fish Biology and Fisheries*, 212, 195–212. <https://doi.org/10.1007/BF00179756>
- McAllister, M. K., & Ianelli, J. N. (1997). Bayesian stock assessment using catch-age data and the sampling-importance resampling algorithm. *Canadian Journal of Fisheries and Aquatic Sciences*, 54(2), 284-300.
- Methot, R.D. 2000. Technical description of the stock synthesis assessment program. U.S. Dept. Commer., NOAA Tech. Memo. NMFS-NWFSC-43, 46 p.
- NPFMC. (2017). *Fishery Management Plan for Groundfish of the Bering Sea and Aleutian Islands Management Area*.
- NPFMC. (2019). *Fishery Management Plan for Groundfish of the Gulf of Alaska. August*.
- Orlov, A.M., 2004. Migrations of various fish species between Asian and American waters in the North Pacific Ocean. *Aqua, Journal of Ichthyology and Aquatic Biology*, 8(3), pp.109-124.
- Orr, J. W., D. T. Drumm, E. A. Laman, D. E. Stevenson, and G. R. Hoff. 2014. Species identification confidence in the Gulf of Alaska and Aleutian Islands surveys (1980-2011). AFSC Processed Rep. 2014-01, 258 p. Alaska Fish. Sci. Cent., NOAA, Natl. Mar. Fish. Serv., 7600 Sand Point Way NE, Seattle WA 98115.
- Rickey, M.H. 1995. Maturity, spawning, and seasonal movement of Arrowtooth Flounder, *Atheresthes stomias*, off Washington. *Fish. Bull.*, U.S. 93(1):127-138.
- Shotwell, S. K., Spies, I., Ianelli, J. N., Aydin, K., Hanselman, D. H., Palsson, W., Siwicke, K., Sullivan, J., & Yasumiishi, E. (2021). *Assessment of the Arrowtooth flounder stock in the Gulf of Alaska*. <https://apps-afsc.fisheries.noaa.gov/refm/docs/2019/GOAatf.pdf>
- Shotwell, S.K., J.L. Pirtle, J. Watson, A. Deary, M. Doyle, S.J. Barbeaux, M. Dorn, G.A. Gibson, E. Goldstein, D.H. Hanselman, A.J. Hermann, P.J.F. Hulson, B. Laurel, J.H. Moss, O. Ormseth, D. Robinson, L.A. Rogers, C.N. Rooper, I. Spies, W. Strasburger, R.M. Suryan, J.J. Vollenweider (2022) Synthesizing integrated ecosystem research to create informed stock-specific indicators for next generation stock assessments. *Deep-Sea Res. Pt. II: Top. Studies. Oceanogr.* 198. DOI: <https://doi.org/10.1016/j.dsr2.2022.105070>
- Shotwell, S., Hanselman, D., Siwicke, K., Spies, I. and Sullivan, J. (2023) Assessment of the arrowtooth flounder stock in the Gulf of Alaska. In: Stock assessment and fishery evaluation report for the

- groundfish resources of the Gulf of Alaska. North Pacific Fishery Management Council, Anchorage, AK.
- Shotwell, S.K., K., Blackhart, C. Cunningham, E. Fedewa, D., Hanselman, K., Aydin, M., Doyle, B., Fissel, P., Lynch, O., Ormseth, P., Spencer, S., Zador. 2023. Introducing the Ecosystem and Socioeconomic Profile, a proving ground for next generation stock assessments. *Coastal Management*. 51:5-6, 319-352, DOI: 10.1080/08920753.2023.2291858.
- Soderlund, E., E. Anderson, C. L. Dykstra, T. Geernaert, and A. M. Ranta. 2009. 2008 standardized stock assessment survey. *Int. Pac. Comm. Report of Assessment and Research Activities 2008* 469-496.
- Spies, I., K. Aydin, J.N. Ianelli, and W. Palsson. 2015. Assessment of the arrowtooth flounder stock in the Gulf of Alaska. In *Stock assessment and fishery evaluation report for the groundfish resources of the Gulf of Alaska*. North Pacific Fishery Mngt. Council, 605 W 4th Ave, Suite 306 Anchorage, AK 99501.
- Stark, J. 2008. Age- and length-at-maturity of female Arrowtooth Flounder (*Atheresthes stomias*) in the Gulf of Alaska. *Fish. Bull* 106: 328-333.
- Stevenson, D. E., and G. R. Hoff. 2009. Species identification confidence in the eastern Bering Sea shelf survey (1982-2008). AFSC Processed Rep. 2009-04, 46 p. Alaska Fish. Sci. Cent., NOAA, Natl. Mar. Fish. Serv., 7600 Sand Point Way NE, Seattle WA 98115.
- Turnock, B.J., T.K. Wilderbuer and E.S. Brown. 2002. Arrowtooth Flounder. In *Stock Assessment and Fishery Evaluation Report for the 2002 Gulf of Alaska Groundfish Fishery*. Gulf of Alaska Groundfish Plan Team, North Pacific Fishery Management Council, P. O. Box 103136, Anchorage, Ak 99510.
- Wilderbuer, T. K., and E. S. Brown. 1993. Flatfish. In T. K. Wilderbuer (editor), *Condition of groundfish resources of the Gulf of Alaska as assessed in 1993*. U.S. Dep. Commer., NOAA Tech. Memo, NMFS F/NWC-165.
- Wilderbuer, T.K. and B.J. Turnock. 2009. Differential Sex-Specific Natural Mortality of Arrowtooth Flounder in Alaska: Implications for Exploitation and Management. *North American Journal of Fisheries Management* Vol. 29 (2).
- Zimmerman, M., and P. Goddard. 1996. Biology and distribution of Arrowtooth Flounder, *Atheresthes stomias*, and Kamchatka flounders (*A. evermanni*) in Alaskan waters. *Fish. Bull.*, U.S.
- Zimmerman, M. 1997. Maturity and fecundity of Arrowtooth Flounder, *Atheresthes stomias*, from the Gulf of Alaska. *Fish. Bull.* 95:598-611.

Tables

Table 7.1. Catch, OFL, ABC, and TAC for arrowtooth flounder in the Gulf of Alaska from 1964 to the present. Values are in metric tons. Arrowtooth flounder ABC was separated from the flatfish ABC after 1990. Source: AKFIN database (<https://akfinbi.psmfc.org/analytics/>). Catch value for 2025 is accurate as of August 23, 2025.

Year	Catch	OFL	ABC	TAC	Year	Catch	OFL	ABC	TAC
1977	9,449				2012	20,565	250,100	212,882	103,300
1978	8,409				2013	21,612	247,196	210,451	103,300
1979	7,579				2014	36,300	229,248	195,358	103,300
1980	7,848				2015	19,056	226,390	192,921	103,300
1981	7,433				2016	19,835	219,430	186,188	103,300
1982	4,639				2017	26,866	219,327	186,083	103,300
1983	6,331				2018	18,873	180,697	150,945	76,300
1984	3,457				2019	20,061	174,598	145,841	99,295
1985	1,539				2020	21,122	153,017	128,060	96,969
1986	1,221				2021	9,103	151,723	126,970	97,372
1987	4,963				2022	11,627	143,100	119,779	96,501
1988	5,138				2023	9,194	142,749	119,485	94,286
1989					2024	18,180	142,485	119,249	94,141
1990	7,706		343,300		2025	11,272	142,832	119,547	100,454
1991	10,034		340,100	20,000					
1992									
1993	15,970	427,220	303,889	25,000					
1994	15,559	451,690	321,287	30,000					
1995	23,560	275,930	236,240	30,000					
1996	22,583	231,420	198,130	35,000					
1997	16,319	280,800	197,840	35,000					
1998	12,975	295,970	208,337	35,000					
1999	16,207	308,875	217,106	35,000					
2000	24,252	173,915	145,361	35,000					
2001	19,964	173,546	148,151	38,000					
2002	21,231	171,057	146,264	38,000					
2003	29,994	181,394	155,139	38,000					
2004	15,304	228,134	194,900	38,000					
2005	19,770	228,134	194,900	38,000					
2006	27,653	207,700	177,800	38,000					
2007	25,494	214,828	184,008	43,000					
2008	29,293	266,914	226,470	43,000					
2009	24,937	261,022	221,512	43,000					
2010	24,268	254,271	215,882	43,000					
2011	30,903	251,068	213,150	43,000					

Table 7.2. The number of fisheries length and otolith (none aged) samples collected in each year from 1991 to the present. Source: AKFIN NMFS AKRO BLEND/Catch Accounting System.

Year	Number of lengths	Number of otoliths
1991	2770	100
1992	196	
1993	648	
1994	59	
1995	1242	160
1996	647	
1997	1317	50
1998	1390	
1999	2124	2
2000	5115	9
2001	5134	2
2002	6732	29
2003	12629	
2004	3947	1
2005	5631	
2006	6606	
2007	7960	4
2008	7409	33
2009	6647	
2010	5946	23
2011	8428	8
2012	7204	44
2013	6235	
2014	11918	52
2015	8737	
2016	8368	
2017	12220	2
2018	5710	80
2019	6888	110
2020	5418	111
2021	5332	96
2022	6376	108
2023	4538	84
2024	7924	124

Figures

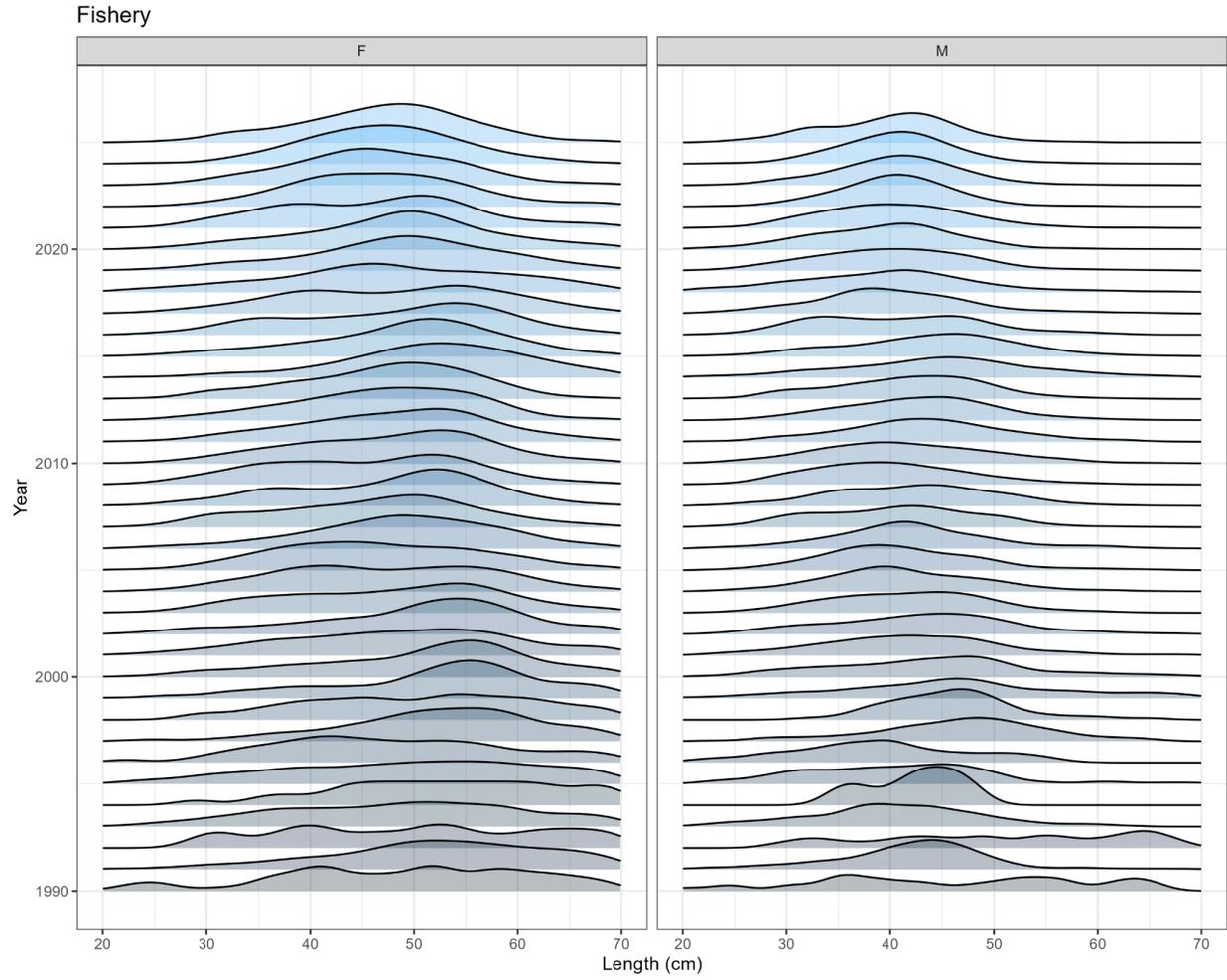


Figure 7.1. Length distribution from fisheries observers from 1990 to present.

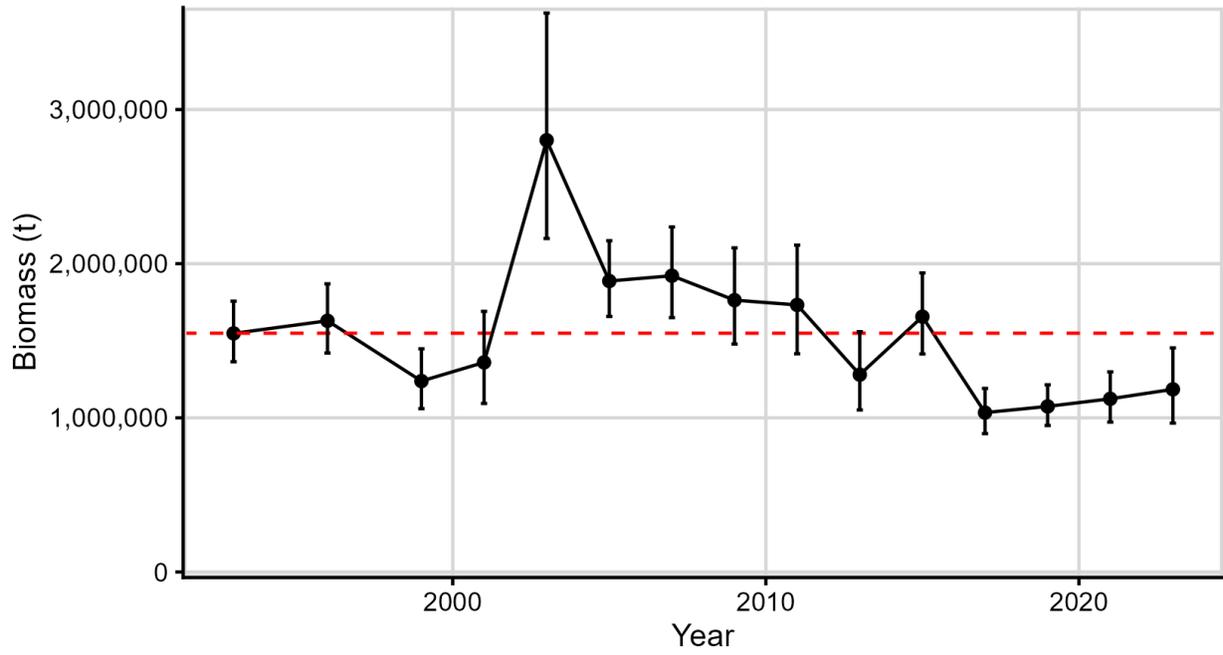


Figure 7.2. Biomass estimates of arrowtooth flounder from the AFSC bottom trawl survey 1993 to the present.

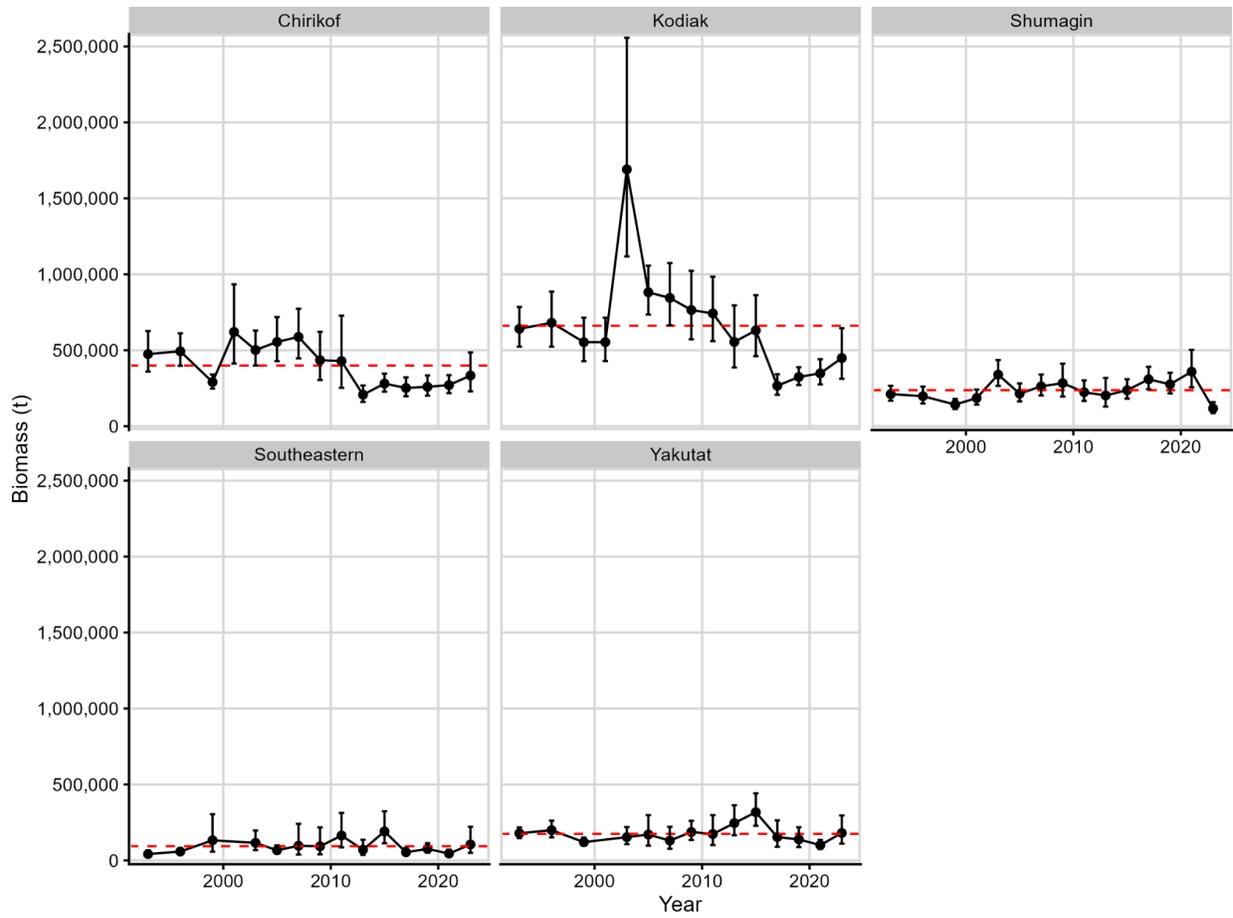


Figure 7.3. Area specific biomass estimates of arrowtooth flounder from the AFSC bottom trawl survey 1993 to the present.

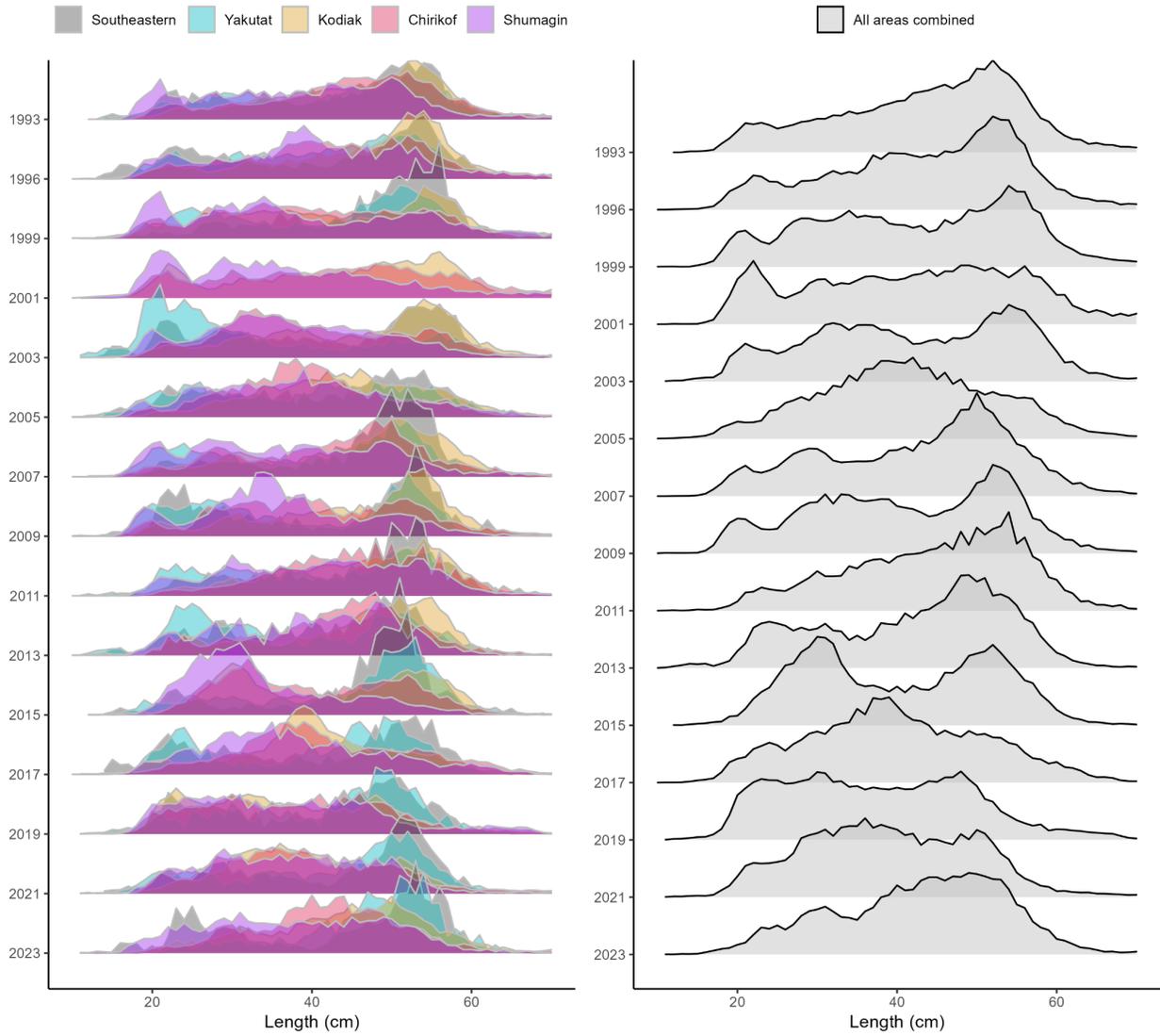


Figure 7.4a. Length composition data for female arrowtooth flounder from the AFSC bottom trawl surveys 1993 to the present.

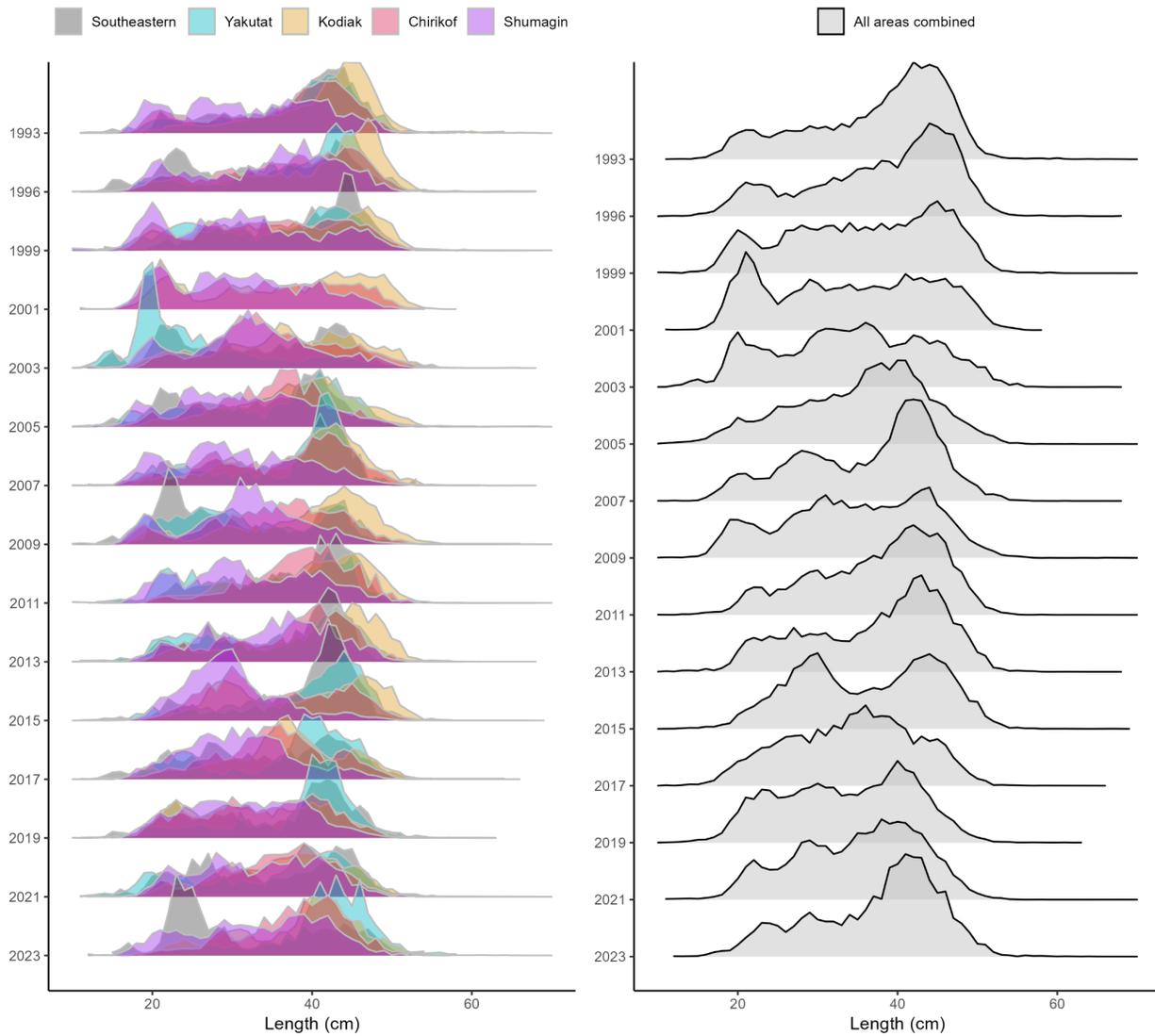


Figure 7.4b. Length composition data for male arrowtooth flounder from the AFSC bottom trawl surveys 1993 to the present.

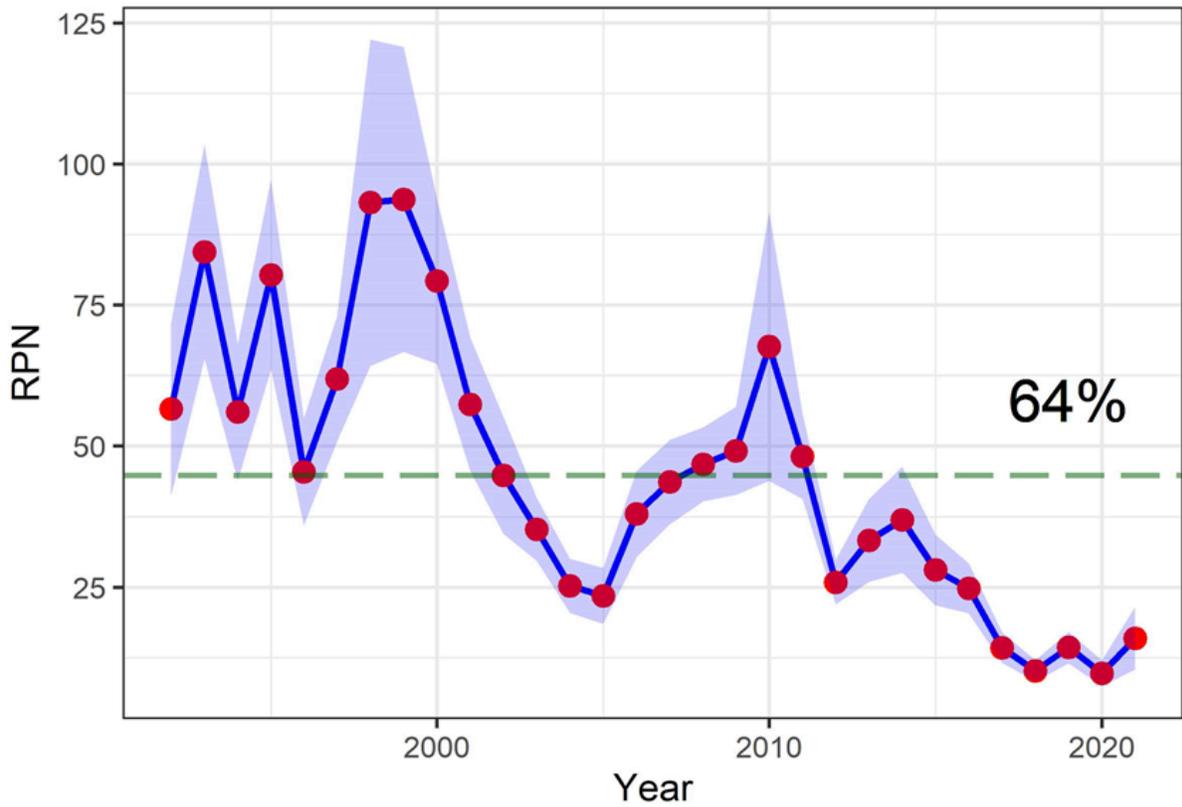


Figure 7.5. Relative population numbers for arrowtooth flounder and Kamchatka flounder combined from the AFSC longline survey 1992 to 2021 (Note: this will be updated in November). The percentage in bold represents the change to the most recent survey estimate from the previous survey estimate.

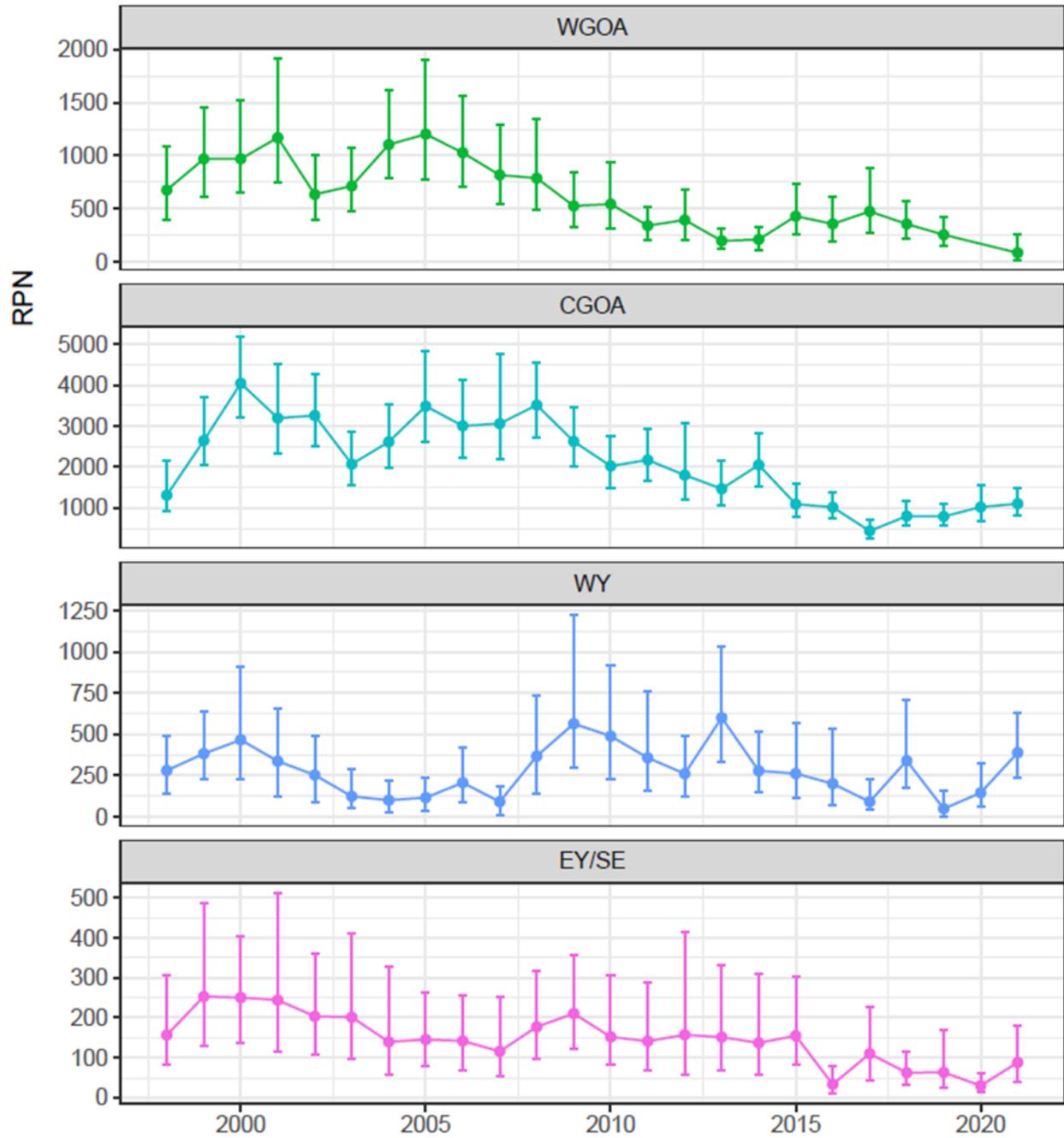


Figure 7.6. Relative population numbers (RPN) for arrowtooth flounder from the International Pacific Halibut Commission (IPHC) survey 1998 to 2021 (Note: this will be updated in November). WGOA = western Gulf of Alaska, CGOA = central Gulf of Alaska, WY = west yakutat, EY/SE = east yakutat/southeast Alaska.

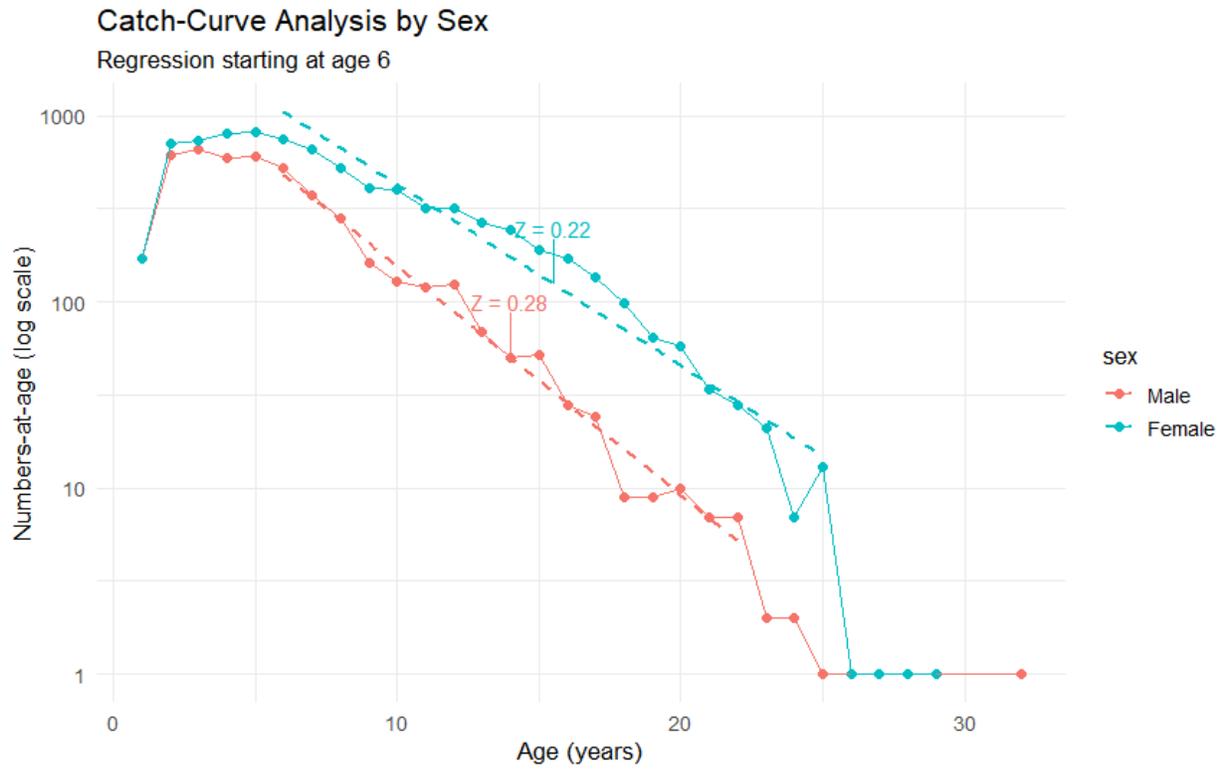


Figure 7.7. Catch curves for male and female arrowtooth flounder with associated estimates of total mortality.

Age → Length Conversion Matrix (Female)

Ages 1–20 with 21+; length bins 2 cm from 20–74 cm, plus ≥75 cm
X-axis shows bin midpoints; first label 20, last label 75+

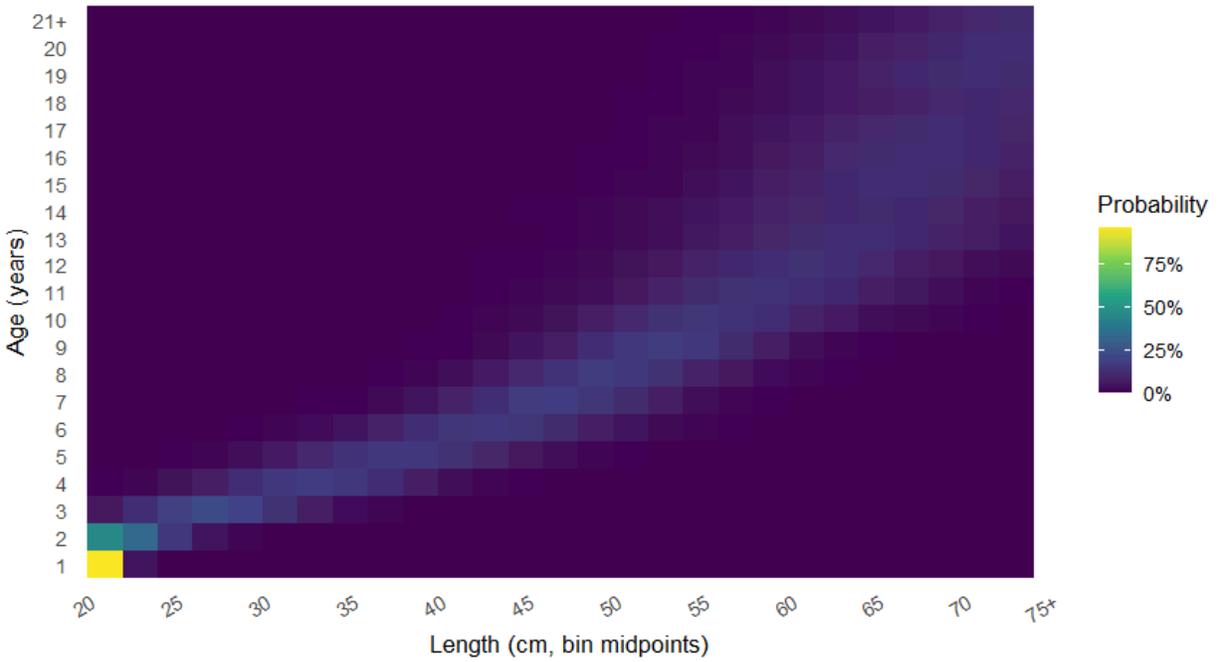


Figure 7.8a. Length conversion matrix for female arrowtooth flounder.

Age → Length Conversion Matrix (Male)

Ages 1–20 with 21+; length bins 2 cm from 20–74 cm, plus ≥75 cm
X-axis shows bin midpoints; first label 20, last label 75+

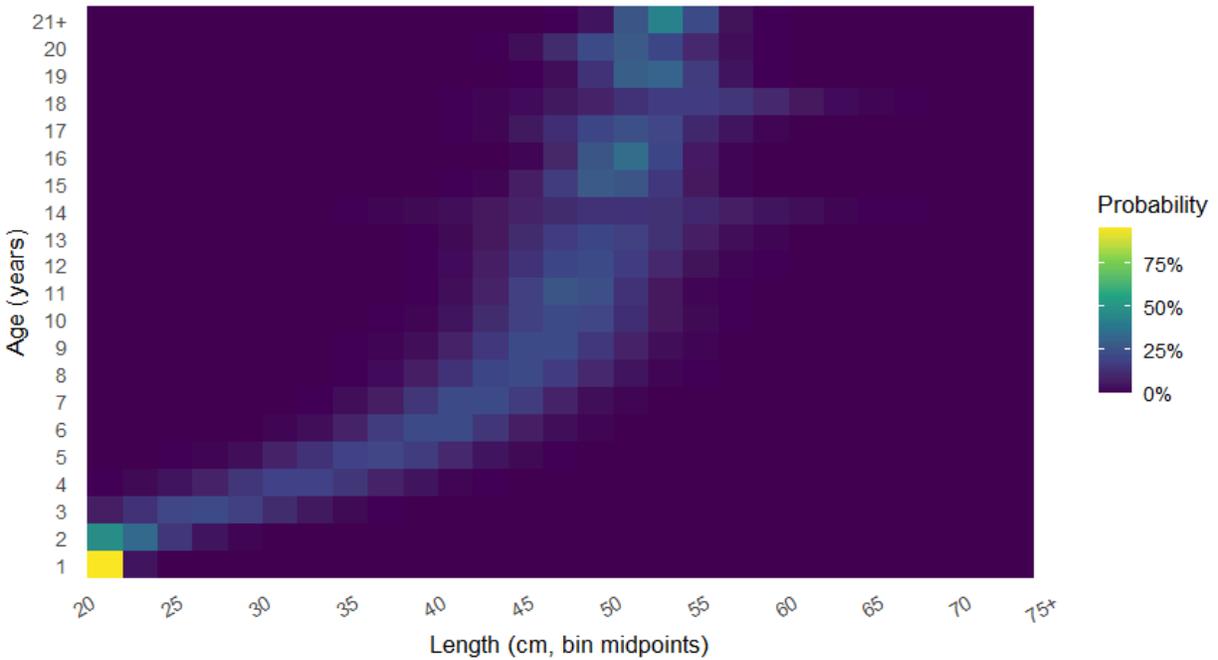


Figure 7.8b. Length conversion matrix for female arrowtooth flounder.

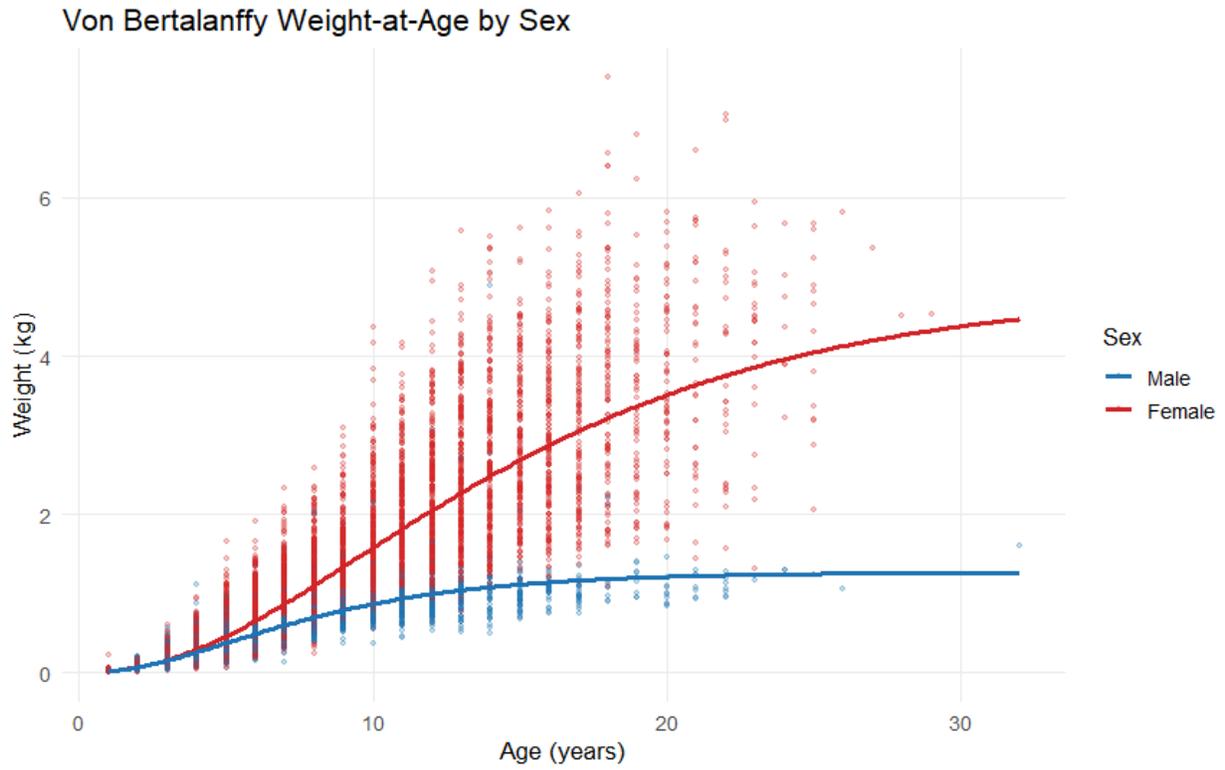


Figure 7.9. Von Bertalanffy relationship for weight at age of female and male arrowtooth flounder from AFSC bottom trawl survey 1993 to 2023.

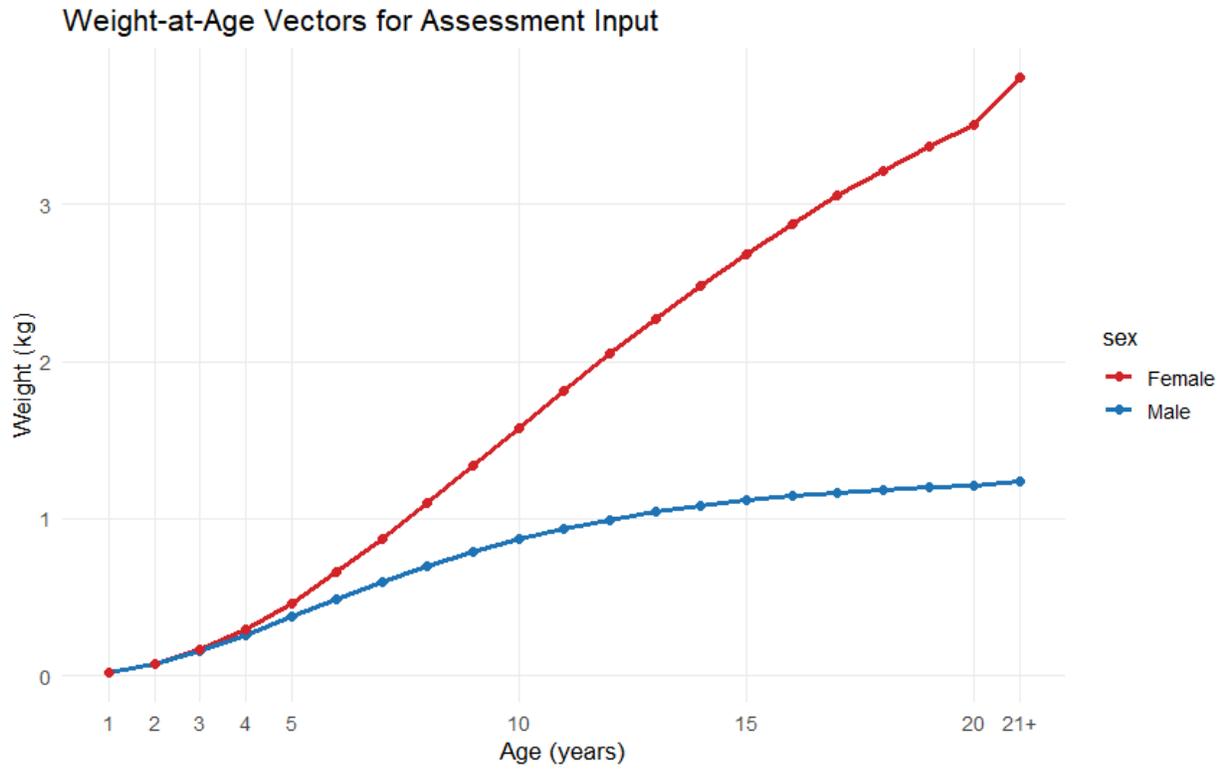


Figure 7.10. Weight at age vectors for female and male arrowtooth flounder.

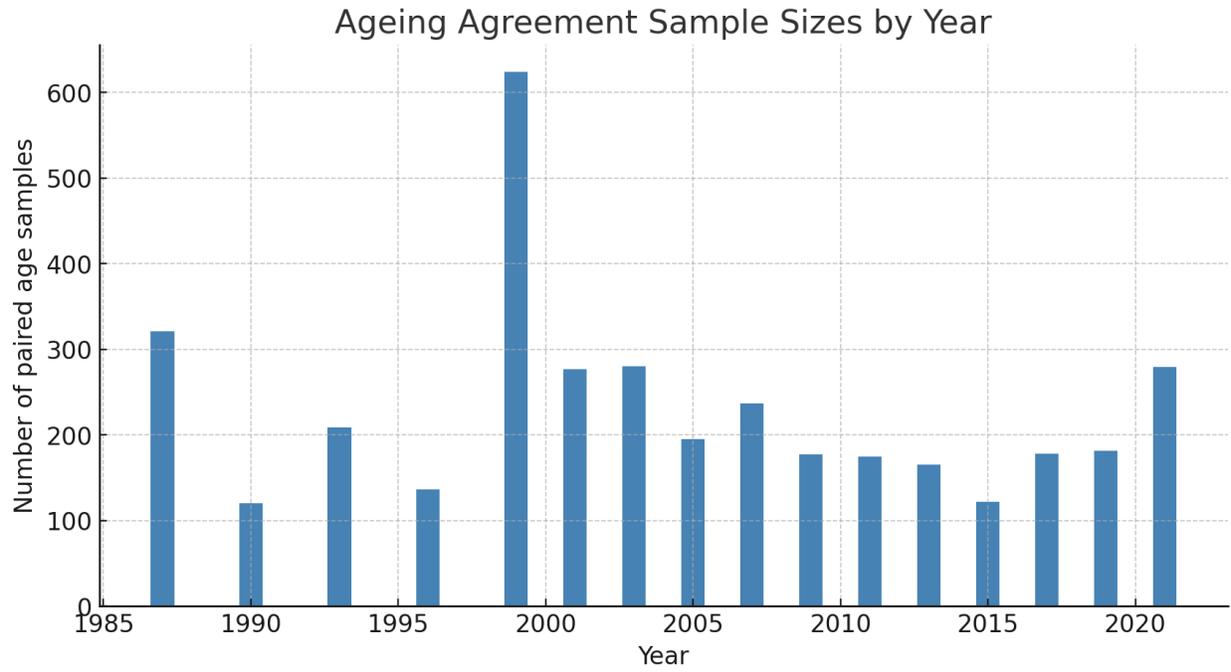


Figure 7.11. Sample sizes of raw paired age reading for arrowtooth flounder from Gulf of Alaska AFSC bottom trawl surveys through 2021.

Ageing Error Matrices (Ages 1–20, 21+)

Rows: true (final) age; Columns: observed age

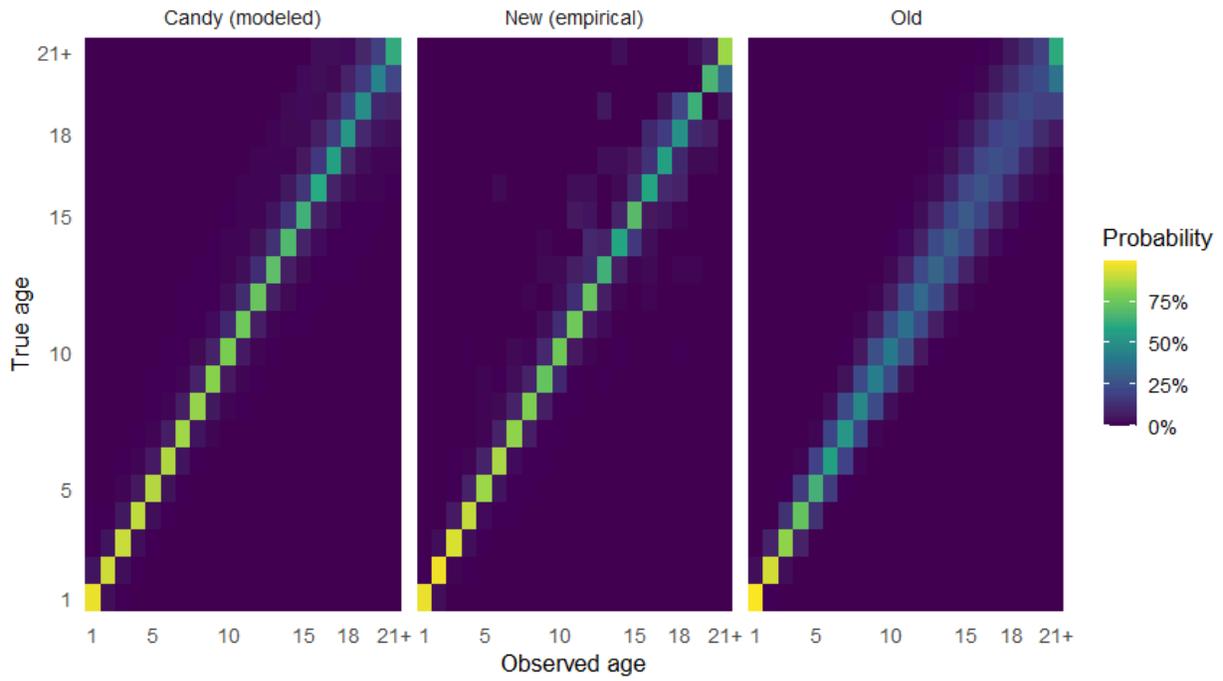


Figure 7.12a. Comparison of three ageing error matrices from the previous stock assessment, empirically-derived, and Candy et al., 2012 framework.

Differences Between Ageing Error Matrices

Positive = first matrix has higher probability

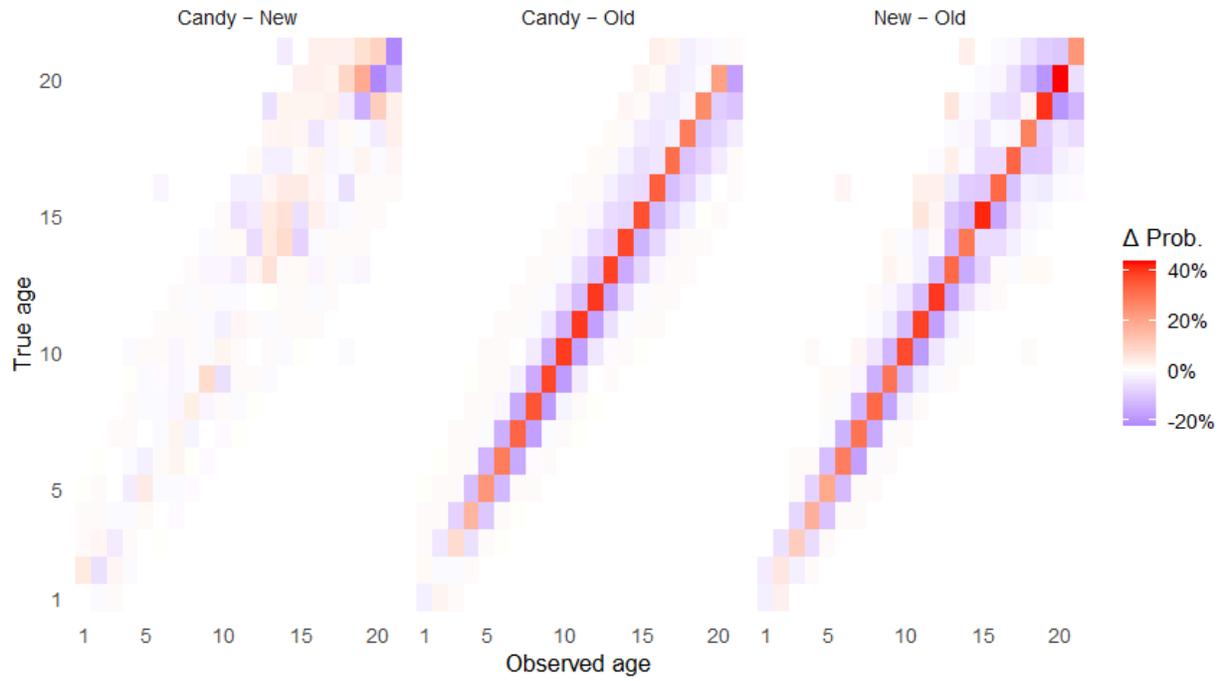


Figure 7.12b. Differences between three ageing error matrices for arrowtooth flounder from the previous stock assessment, empirically-derived, and Candy et al., 2012 framework.

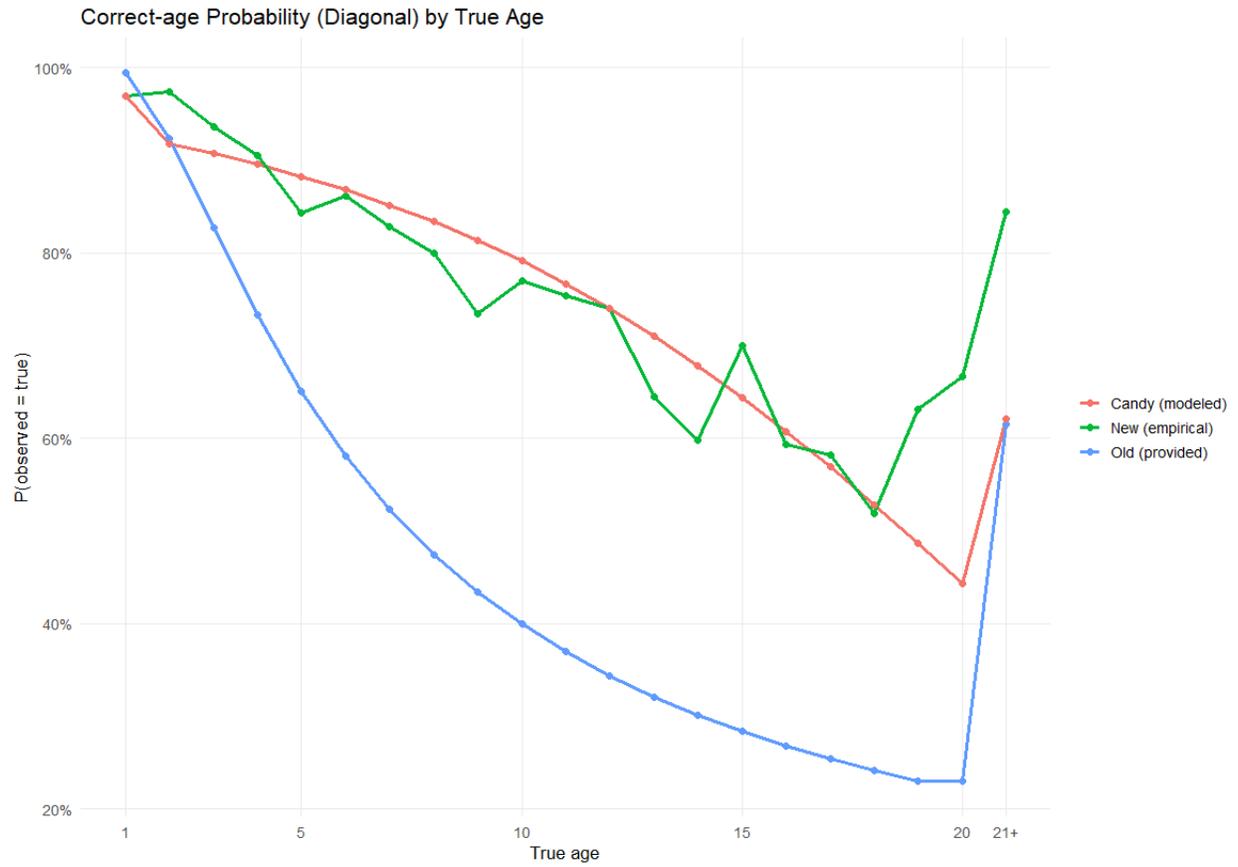


Figure 7.13 Diagonal probabilities for three ageing error matrices of arrowtooth flounder from the previous stock assessment, empirically-derived, and Candy et al., 2012 framework.

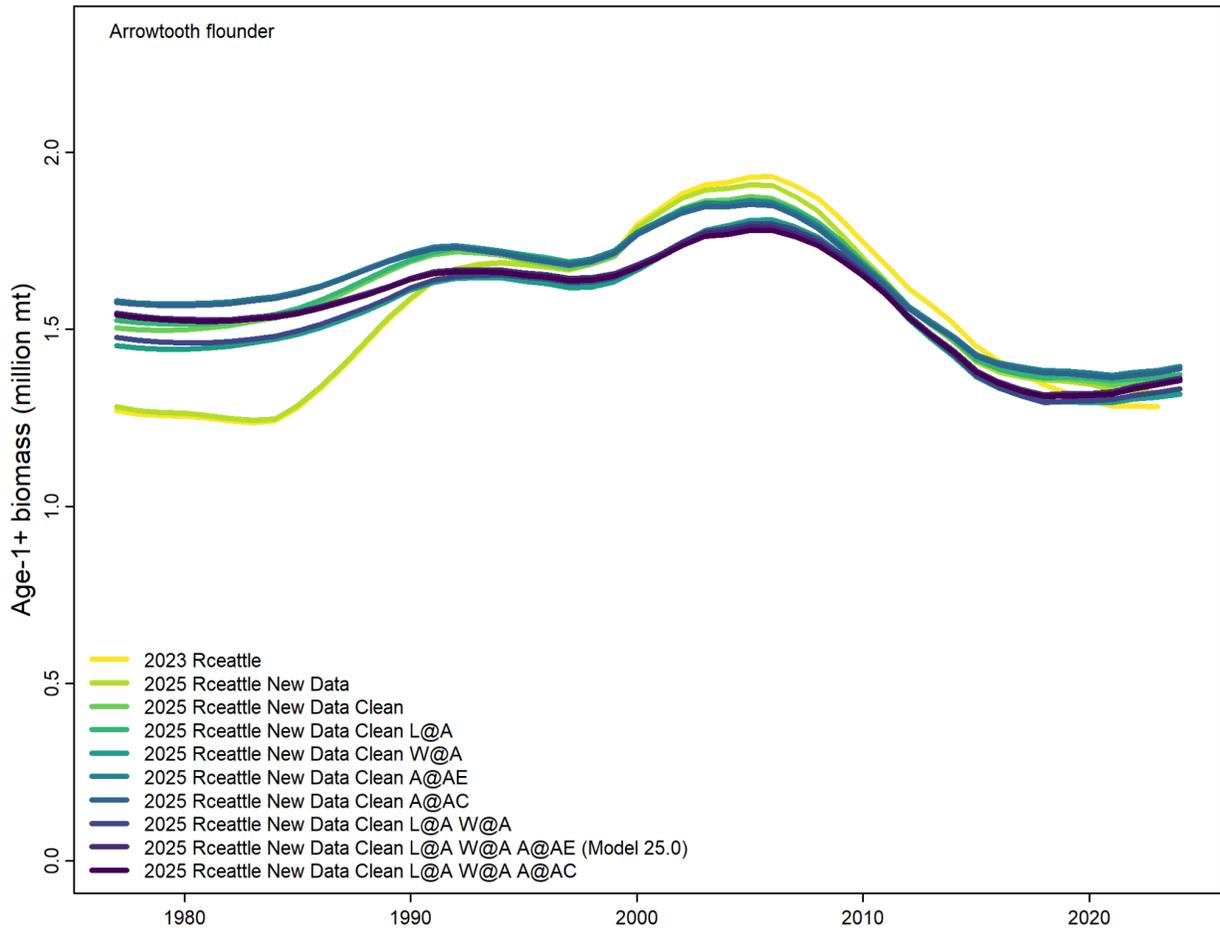


Figure 7.14a: Comparison of total biomass (age 1+) from initial model runs of the data changes for arrowtooth flounder. L = length, A = age, W = weight, E = empirical, and C = Candy method.

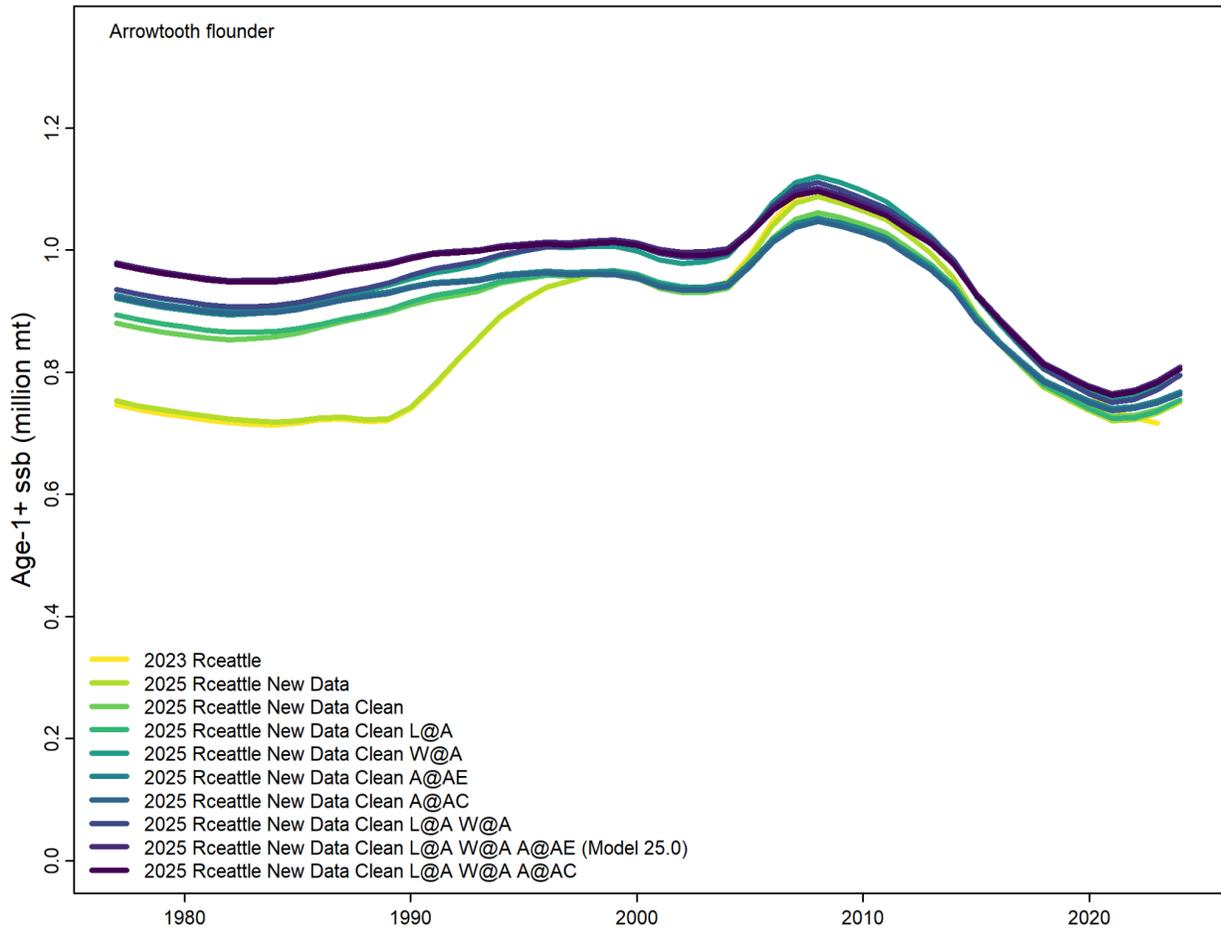


Figure 7.14b: Comparison of spawning stock biomass (SSB) from initial model runs of the data changes for arrowtooth flounder. L = length, A = age, W = weight, E = empirical, and C = Candy method.

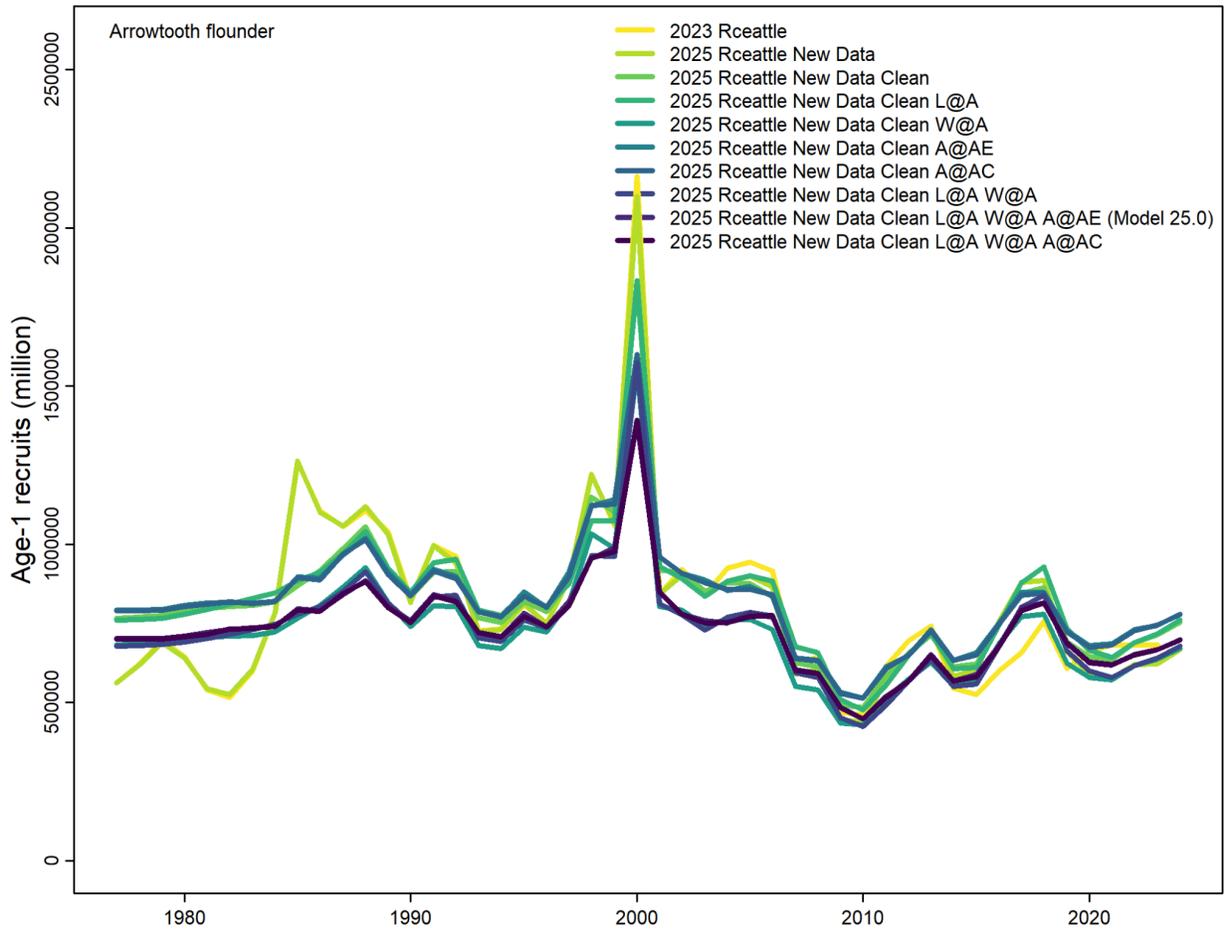


Figure 7.14c: Comparison of recruitment (age 1+) from initial model runs of the data changes for arrowtooth flounder. L = length, A = age, W = weight, E = empirical, and C = Candy method.

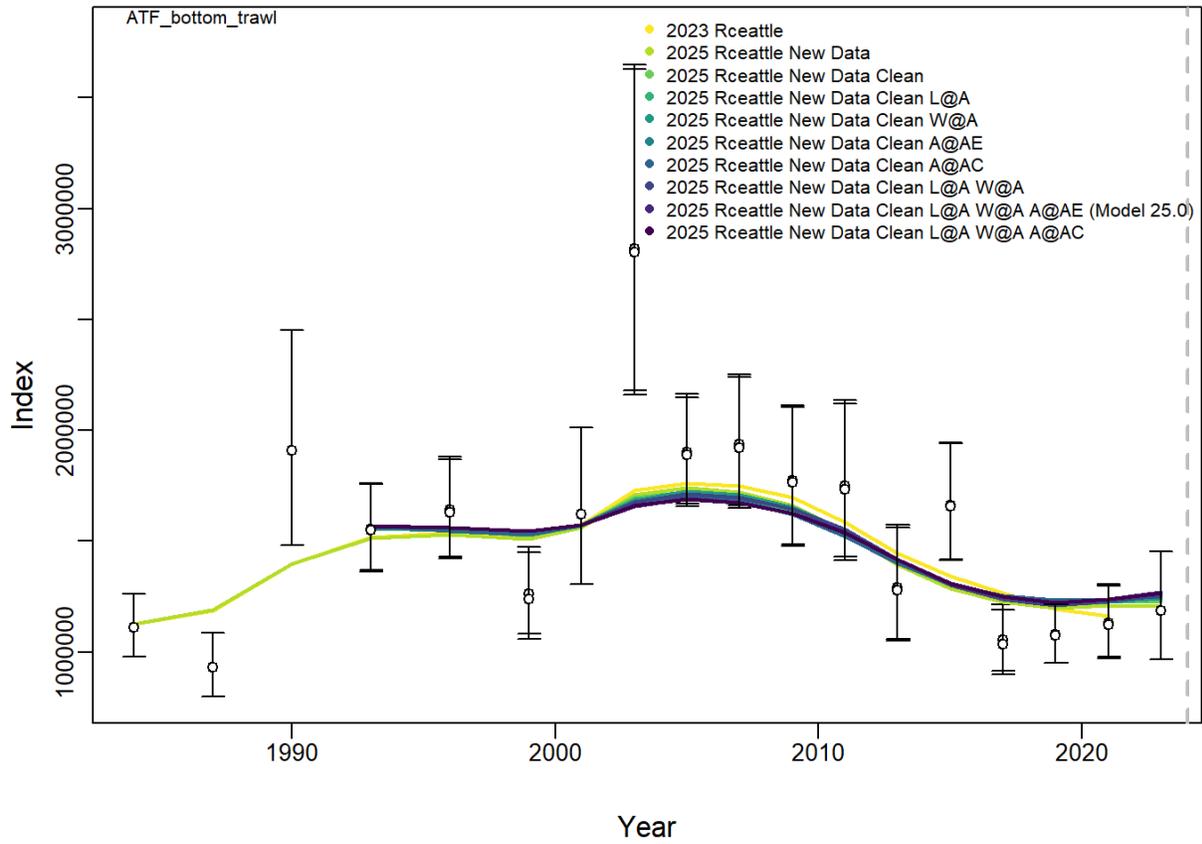


Figure 7.14d: Comparison of AFSC trawl survey biomass from initial model runs of the data changes for arrowtooth flounder. L = length, A = age, W = weight, E = empirical, and C = Candy method.

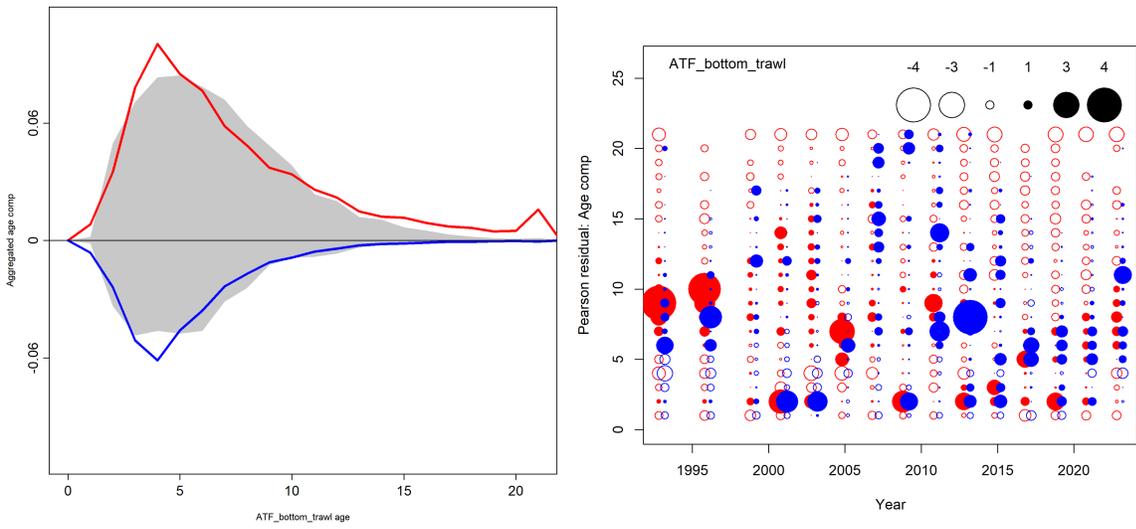


Figure 7.15a: Fits to aggregated and pearson residuals of observed and expected bottom trawl survey age-composition data for Model 25.0. Females are red and males are blue.

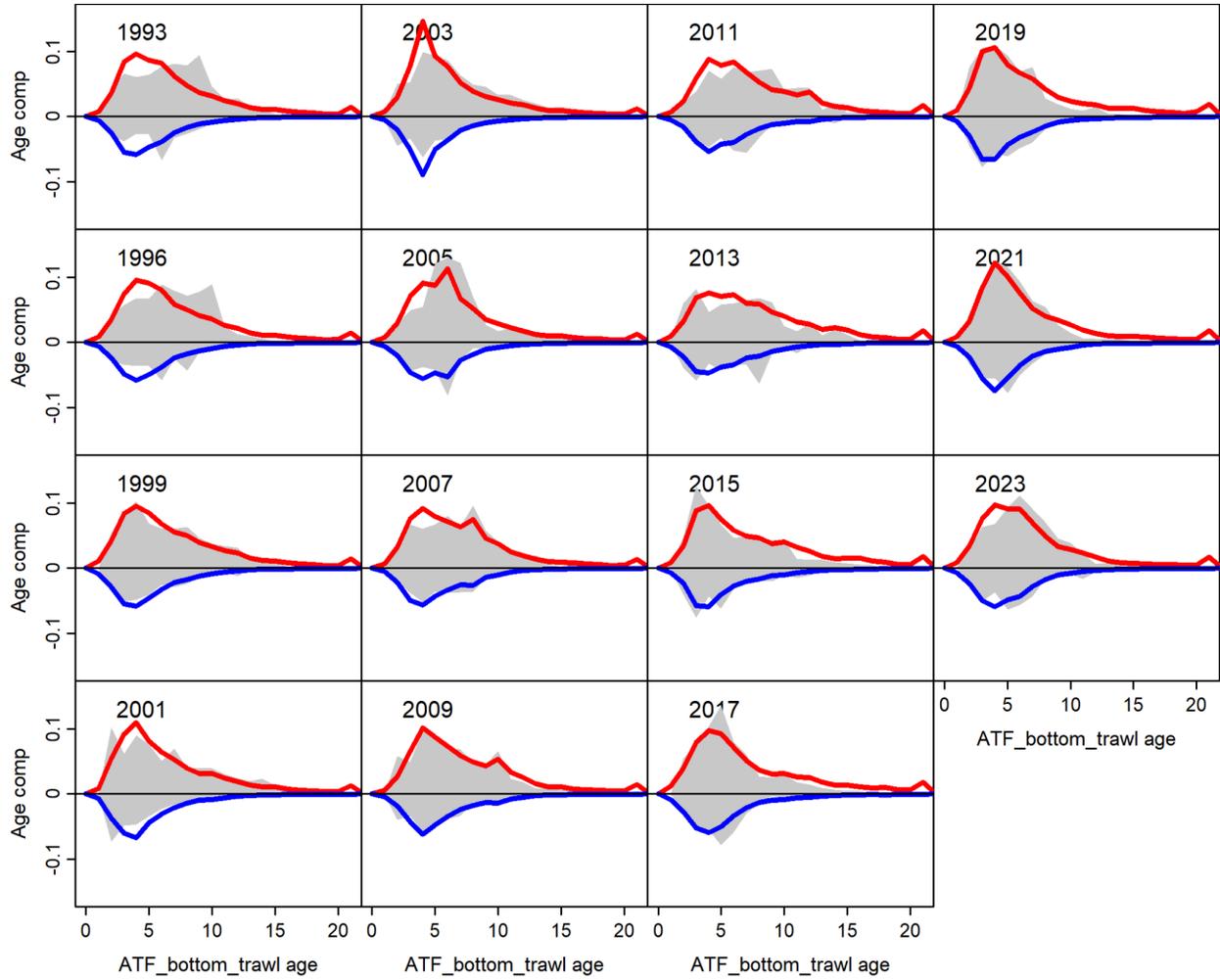


Figure 7.15b: Annual fits to AFSC summer bottom trawl age composition data for Model 25.0.

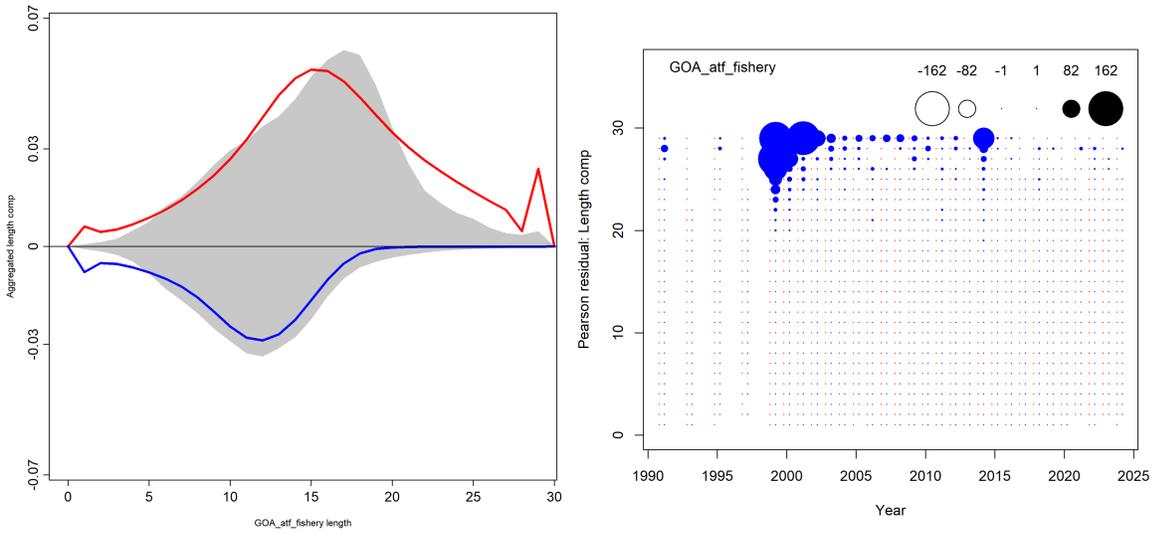


Figure 7.16a: Fits to aggregated and pearson residuals of observed and expected fishery length-composition data for Model 25.0. Females are red and males are blue.

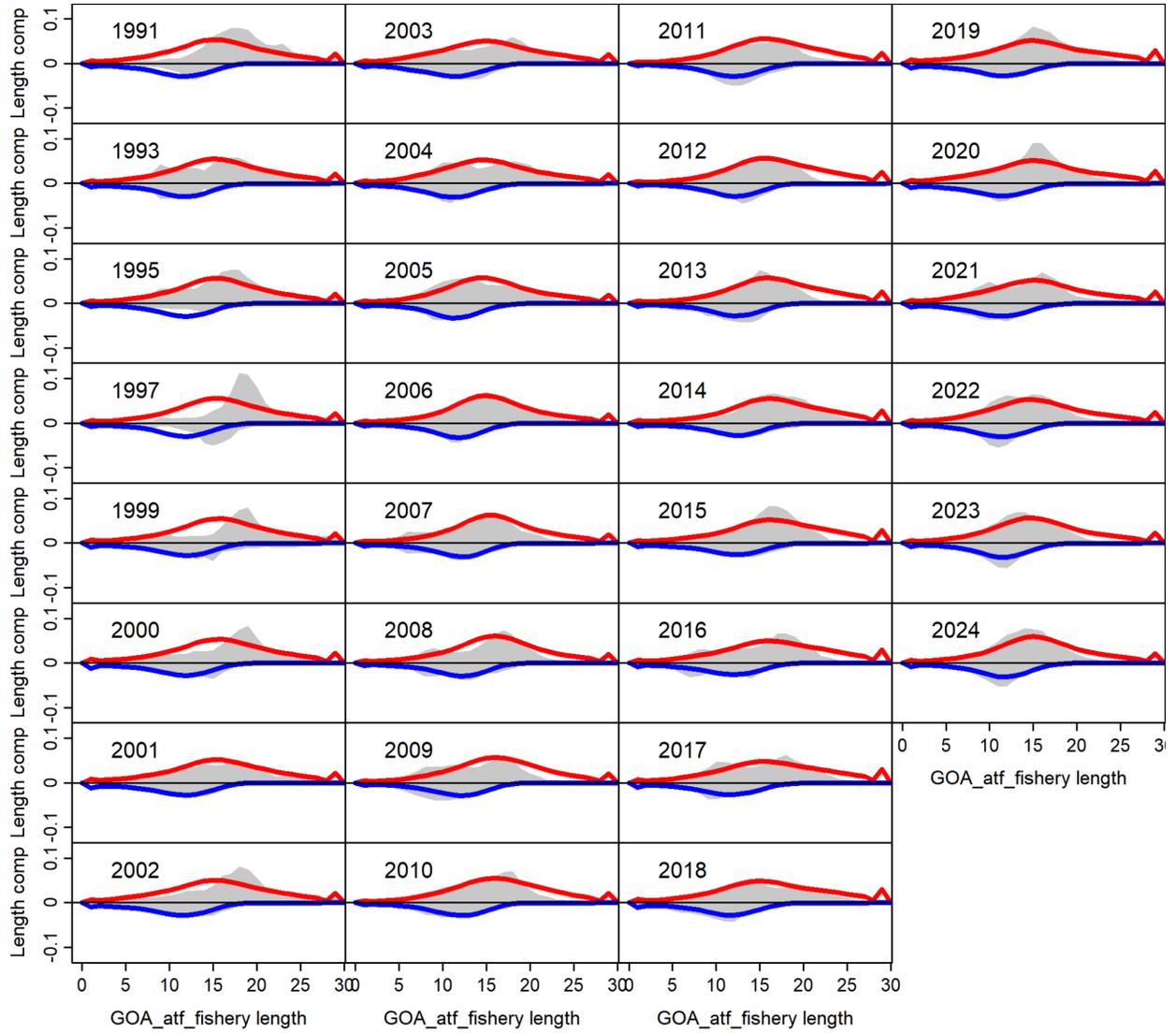


Figure 17.16b: Annual fits to fishery length compositions data for Model 25.0.

Appendix 7A: Bridging the Gulf of Alaska arrowtooth flounder assessment to CEATTLE to improve parameter estimation, account for cannibalism, and increase climate readiness

Grant Adams and Kalei Shotwell

September 2024

Executive Summary

Recommendations from the Groundfish Plan Team and the Scientific and Statistical Committee (SSC) for the Gulf of Alaska (GOA) arrowtooth flounder (ATF) stock assessment and fishery evaluation (SAFE) over the past several years included investigation of the following elements:

- Recent lower recruitment trends and relationship to environmental conditions in the GOA, including the development of an ESP
- Lack of fit in female survey age and fishery length compositions, including interactions between female natural mortality and selectivity
- Incorporation of predation mortality estimates from the GOA CEATTLE model
- Update growth and age-length conversion matrices

In response to these recommendations and to improve our ability to investigate these concerns, we plan to improve the current GOA ATF assessment model in two main ways. First, we want to bridge from the current Automatic Differentiation Model Builder (ADMB) software to Template Model Builder (TMB) and second we want to include the impact of predation mortality by exploring the multi-species age-structured stock assessment model. We plan to do this by using the CEATTLE (Climate-Enhanced, Age-based model with Temperature-specific Trophic Linkages and Energetics) model that has been expanded for groundfish in the GOA (Holsman et al., 2016, Adams et al., 2022) using TMB (Kristensen et al., 2016).

Background

The current GOA ATF stock assessment model is based on an age-structured model using ADMB software (Fournier et al. 2012). The framework uses automatic differentiation and allows estimation of highly-parameterized and non-linear models. The approach consists of an assessment model, which uses survey and fishery data to generate a historical time series of population estimates, and a projection model which uses results from the assessment model to predict future population estimates and recommended harvest levels. The assessment model is fit to survey abundance data, survey age data, and fishery length composition data with a harvest control rule to model the status and productivity of these stocks and set quotas.

The current operational model using ADMB has been used for the GOA and the Bering Sea and Aleutian Islands (BSAI) ATF stock assessments since 2015 and stems from the "generalized" ADMB model developed for Alaska flatfish (J. Ianelli, pers. comm.). The model incorporates ages 1-20 with 21 and greater in a plus group, estimates age-based selectivity up to age 19 with full selectivity after age 19, and uses 26 length bins starting at 10 cm to a 75 cm plus group and defined at 2 and 3 cm intervals (see last full assessment for more details, Shotwell et al., 2021). A Markov chain Monte Carlo (MCMC) is performed in ADMB to capture variability in recruitment, female spawning biomass, and total (age 1+) biomass. No spawner-recruit curve is used in the model. Instead, we calculated average recruitment with

an estimated lognormal deviation for each year of the model with the exception of the final year. In the final modeled year, recruitment is set to median recruitment. Recruitment deviations are freely estimated but with a modest penalty on extreme deviations from the mean value (σ_r value). Age at recruitment is set to one in the model. Recruitment is informed by subsequent age and length composition and there is little information to inform recruitment in the final few years because selectivity is low for younger ATF. Ages 2-21 are subject to a sex-specific vector of instantaneous rates of natural mortality. Natural mortality is subscripted for sex, as males appear to have higher natural mortality than females in this species (Wilderbuer and Turnock 2009, 0.2 for females, 0.35 for males). Fishery selectivity is estimated as a smooth, age- and sex-specific non-parametric function through age 19. Survey selectivity is modeled using a two parameter ascending logistic function for males and females (four parameters total). Please refer to the most recent full SAFE report (Shotwell et al., 2021) for more details on the current assessment model (hereafter referred to as the ADMB model).

Justification

Since 2019, recommendations from the Groundfish Plan Team and the SSC for the GOA ATF SAFE have included investigation of the following elements:

- Recent lower recruitment trends and relationship to environmental conditions in the GOA, including the development of an ESP
- Lack of fit in female survey age and fishery length compositions, including interactions between female natural mortality and selectivity
- Incorporation of predation mortality estimates from the GOA CEATTLE model
- Update growth and age-length conversion matrices

The goals of this document are to describe how we intend to start addressing these recommendations by 1) updating the current ADMB model to TMB to improve parameter estimation and 2) accounting for both the impacts of cannibalism and fishery removals in the population dynamics assessment model used for arrowtooth flounder.

TMB (Kristensen et al., 2016) is based on ADMB, but allows for improved estimation of hierarchical non-linear models. TMB, often considered a successor to ADMB, can estimate time-varying parameters (e.g. recruitment deviations) as random effects using marginal maximum likelihood rather than the penalized likelihood used in ADMB. This allows proper estimation of associated variance parameters, which can bias management quantities if misspecified. For example, the current ADMB assessment does not have an explicit recruitment variance parameter, rather the ADMB assessment penalizes annual recruitment deviates by squaring them in the log-likelihood (*i.e.* rec_dev^2). This parameterization implicitly assumes a variance of 0.707, which may not accurately reflect the amount of variability in recruitment. Moving to TMB will allow estimation of recruitment deviates as random effects and allow estimation of age- and sex-specific non-parametric selectivity penalties or other selectivity forms.

Arrowtooth flounder are generalist predators. Analyses from bottom trawl survey diet data show a non-trivial amount of cannibalism particularly on the small to medium (30-299 mm) sizes (Doyle et al., 2018). An ontogenetic shift occurs in ATF diet with smaller fish (<200 mm) consuming primarily euphausiids and capelin to larger fish consuming mostly walleye pollock and increasing amounts of flatfish as they age. Predation on ATF varies with their habitat utilization from pelagic zooplanktivorous predators (e.g., walleye pollock, northern rockfish, Pacific ocean perch) while ATF are small (<30 mm) to increasing amounts of benthic predation (particularly from Pacific cod, Pacific halibut, and larger ATF) as ATF age. The larger ATF (>400 mm) are consumed by the largest predators (e.g., sleeper sharks) (Doyle et al., 2018). Adams et al. (2022) estimated that on average approximately 27,000 to 37,000 t of ATF were consumed annually by predators in the models explored and that of those species considered (Pacific

halibut, ATF, Pacific cod, walleye pollock), ATF was the primary consumer. The fishery has caught between 1,200 to 37,000 mt of ATF between 1977 and 2023. Catch was on average around 17,000 mt since 1991 and has been decreasing from the maximum in 2014 at 37,000 mt to about 10,000 mt since 2021 (see Table 7.3 in Shotwell et al., 2021). In some years cannibalism can largely exceed the removals by the fishery and potentially impact harvest recommendations, highlighting the need to account for the impacts of cannibalism on the population dynamics of ATF (Adams et al., 2022).

A preliminary ecosystem and socioeconomic profile or ESP has recently been created for GOA ATF that includes several mechanistically derived physical and biological indicators to evaluate the impacts of the ecosystem on ATF (Doyle et al., 2018; Shotwell et al., 2022; Ferriss, 2023). Indicators in this report include estimated ATF annual ration, ATF biomass eaten due to cannibalism, and sex-specific natural mortality (M) for ages 1 and 2 from the most recent GOA CEATTLE model (Adams et al., 2023) over the years 1977 to 2023. ATF ration averages 9.7 million t and ranges from 8.2-11.4 million t, clearly showing the impact this predator has in the GOA ecosystem. Biomass eaten due to cannibalism ranges from 13,000 t to 49,000 t and averages 26,000 t, supporting the Adams et al., 2022 results. Cannibalism exceeds catch in all years up until 2011 where it becomes comparable to catch. Since 2021, cannibalism has again exceeded catch due to the recent lower harvest levels. Sex-specific M for ages 1 and 2 are much higher than the ADMB model fixed estimates of 0.2 for females and 0.35 for males (Shotwell et al., 2021). Average M at age 1 was equal to 0.335 for females (range 0.326 to 0.349) and 0.43 for males (range 0.417 to 0.451). Average M at age 2 was equal to 0.356 for females (range 0.343 to 0.374) and 0.458 for males (range 0.439 to 0.485). The variable M by sex and age from the multi-species models from Adams et al., 2023 supports exploring the estimation of sex- and age-specific M for ATF.

In the GOA, climate can impact the population dynamics of groundfish related to stock assessment through recruitment, bioenergetics/growth, and mortality (Anderson and Piatt, 1999; Clark and Hare, 2022). Additionally, this can have impacts on cannibalism in arrowtooth flounder via temperature-dependent foraging rates, changing species distributions, and shifts in abundance of conspecific predators (Barnes et al., 2018; Holsman and Aydin, 2015). These shifts in predator-prey interactions undermine the assumptions of time-invariant natural mortality (and the equilibrium assumptions of reference points), highlighting the need to evaluate the impact of climate-driven dynamics on management recommendations for ATF. Currently the ADMB model for arrowtooth is not configured for exploration and estimation of climate-linkages. Development of models that can allow incorporation and testing of multiple climate-linkages would be valuable for evaluating future management performance and explaining historical population variability.

Methods

We use a previously developed climate-enhanced multi-species statistical catch-at-age modeling framework called CEATTLE (Climate-Enhanced, Age-based model with Temperature-specific Trophic Linkages and Energetics, Holsman et al., 2016) that has been expanded for groundfish in the GOA using TMB (Adams et al., 2022, Kristensen et al., 2016) as the basis for the new assessment model explorations for ATF. Developed in part from the underpinnings of multi-species statistical catch-at-age analysis (MSCAA, Jurado-Molina et al., 2005) and multi-species virtual population analysis (MSVPA; Jurado-Molina et al., 2005; Magnusson, 1995), CEATTLE links single-species age-structured models (Table 7A.1) through predation mortality conditioned on the temperature-dependent bioenergetic demand and diet-based prey-selectivity patterns of predators (Curti et al., 2013; Holsman et al., 2016; Kinzey and Punt, 2009). Predation mortality assumes a Holling Type II functional response and parameters are either pre-specified or estimated by fitting to survey and fishery data (Table 7A.2). CEATTLE can be run in single-species mode by “turning-off” time-varying predation mortality or in multi-species mode by “turning-on” time-varying predation mortality.

Data

CEATTLE uses the same inputs as the ADMB single-species assessment model (ADMB model, Shotwell et al., 2021) used to provide management advice in addition to diet and bioenergetics data (Adams et al., 2022) (Table 7A.5). Catchability for the AFSC bottom trawl survey is assumed to be 1, and age-at-maturity, weight-at-age, age-length conversion matrix, and ageing error matrix are all estimated outside the model. All data and assumed uncertainties are consistent with the ADMB model. Diet and bioenergetics data were derived from the AFSC stomach sampling program from 1990 to 2015 (Adams et al., 2022, Holsman et al. 2023, Holsman & Aydin 2015). Time-series of bottom temperature derived from the Climate Forecast System Reanalysis for the Pacific cod assessment (Hulson et al., 2023) was used to parametrize consumption and diet composition.

Model Structure

For arrowtooth flounder in the GOA, the single-species CEATTLE spans 1977 to 2023 and is parameterized similarly to the ADMB model (Adams et al., 2022; Shotwell et al., 2021). The effects of changing climate are accounted for by conducting projections with temperature that determine consumption rates over time. Parameters estimated inside the model include the number-at-age in the first year of the assessment, the number of recruits in subsequent years, the fishing mortality rates for each year, and survey/fishery selectivity. Similar to the ADMB model, separate fishery selectivities were estimated non-parametrically for each age, up to age-19, and the shape of the selectivity curve was constrained to be a smooth function. Survey selectivities were modeled using a two-parameter ascending logistic function. The selectivities by age were estimated separately for females and males. A differential age-invariant natural mortality is assumed or estimated for each sex. CEATTLE assumes multinomial likelihoods for composition data and log-normal likelihoods for index and catch data (Table 7A.3). Despite similarities, the assessment model and CEATTLE have the following differences:

1. In the case of age-composition data, the multinomial in the ADMB model is defined as:

$$NLL = - \sum_i \sum_y n_{f_i 1y} \sum_a \left(\hat{O}_{f_i 1ay} + 0.00001 \right) * \ln \left(O_{f_i 1ay} + 0.00001 \right) - \sum_i \sum_y n_{f_i 2y} \sum_a \left(\hat{O}_{f_i 2ay} + 0.00001 \right) * \ln \left(O_{f_i 2ay} + 0.00001 \right)$$

where $\hat{O}_{f_i 1ay}$ is the predicted age composition from fleet f_i for species i for sex (1 = females; 2 = males) for true age a and year y , $O_{f_i 1ay}$ is the observed true age composition, and $n_{f_i 1y}$ is the sex-specific sample size. The ADMB model, therefore, assumes the input sample sizes are sex-specific, but the observed and true age compositions are calculated relative to both sexes as in Table 7A.1, resulting in a mis-specified multinomial distribution. CEATTLE, instead assumes a single sex-combined input sample size ($n_{f_i y}$) avoiding the mis-specification and allowing the sex ratio to be correctly estimated. Additionally, CEATTLE uses the predicted and observed age composition for observed ages in the multinomial likelihood (accounting for ageing error; see point 4 below).

2. The non-parametric fishery selectivity penalties (χ_{f_i}) in CEATTLE are sex-invariant (Table 7A.3), while in the ADMB model, these penalties are different for males and females.
3. The log-normal likelihoods used by the CEATTLE model include a log-normal bias correction and exclude an added constant of 0.0001 that was included in the ADMB model. CEATTLE also

utilizes the complete probability distribution rather than the simplified form used in the ADMB model.

4. To account for ageing error, the ADMB model multiplies the observed age composition data by the ageing error matrix $\hat{O}_{f_i 1ay} = \hat{O}_{f_i 1ay} AE_{iaa}^{\hat{}}$. However, CEATTLE multiplies the expected true age composition by the ageing error matrix to derive the expected observed age-composition $O_{f_i 1ay}^{\hat{}} = O_{f_i 1ay} AE_{iaa}^{\hat{}}$.
5. Multinomial distribution weights for the ADMB model are not updated with new data (weights in the 2017 assessment are the same in 2021). However, CEATTLE calculates McAllister and Ianelli (1997) weights allowing for easy sample size tuning. CEATTLE can also be easily updated so that composition weights are directly estimated using the dirichlet-multinomial distribution.

Models

We present four bridging models to demonstrate the major transitions from the ADMB platform to the multi-species CEATTLE model:

1. ADMB model: current operational single-species ADMB based assessment from the 2021 SAFE (Shotwell et al., 2021) and updated catch to 2023 (Shotwell et al., 2023).
2. TMB single-spp (species) fixed natural mortality (M) model: the single-species TMB based CEATTLE model that fixes sex-specific M (females = 0.2 and males = 0.35) and treats annual recruitment as random effects.
3. TMB single-spp (species) estimated M model: the single-species TMB based CEATTLE model that estimates sex-specific M and treats annual recruitment as random effects.
4. TMB multi-spp (species) model: the “multi-species” CEATTLE based cannibalism model that estimates sex-, age-, and time-varying M due to cannibalism from ATF ($M2$), sex-specific residual mortality (MI), and treats annual recruitment as random effects.

The sex-specific residual mortality (MI) in the “multi-species” CEATTLE model represents mortality due to predation from other species, disease, senescence, etc. We also created a series of sub-models to demonstrate bridging between the ADMB model and the TMB model in CEATTLE (between models 1 and 2 listed above). Please see the appendix for details on this section of the model bridging exercise. Due to differences in likelihood specifications between the ADMB model and CEATTLE, likelihood components are not directly comparable, however, see the appendix for a more extensive bridging.

Results

Both the ADMB and single-species CEATTLE model that fixed sex-specific M (models 1 & 2) showed similar if not exactly similar trends in recruitment, spawning stock biomass, and biomass (Figure 7A.1). The single-species model that estimated M and the multi-species CEATTLE model that included cannibalism (models 3 & 4) led to higher estimates of recruitment and biomass, owing to higher estimates of total natural mortality for younger age-classes (Figure 7A.2). The single-species model that estimated M estimated it to be 0.296 for females and 0.383 for males, which is higher than the 0.2 and 0.35 assumed by the current assessment. The multi-species model that estimated sex-, age-, and time-varying mortality, mortality due to cannibalism ($M2$) only impacted ages 1-6 for females and ages 1-7 for males. For females, total mortality estimated on average as age-1 = 0.335, age-2 = 0.356, age-3 = 0.351, age-4 = 0.321, age-5 = 0.31, and age-6 = 0.307 between 1977 and 2023. For males, total mortality estimated on average as age-1 = 0.43, age-2 = 0.458, age-3 = 0.455, age-4 = 0.408, age-5 = 0.395, age-6 = 0.39, and age-7 = 0.389 between 1977 and 2023 (Figure 7A.2). However, both the single-species model that estimated M and the multi-species model led to lower estimates of spawning stock biomass than the

ADMB model due to higher estimates of mortality for older fish. In the multi-species model residual mortality (*e.g.* total natural mortality for age-8+ fish) was estimated as $MI = 0.307$ for females and $MI = 0.388$ for males. This was also higher than the natural mortality of 0.2 and 0.35 assumed by the assessment for females and males, respectively. The multi-species model also estimated that cannibalism peaked in 2001 to 49,122 t, but has decreased in recent years as biomass has decreased (Figure 7A.3).

The single-species CEATTLE model that estimated M led to the lowest log-likelihood among the CEATTLE models evaluated (Table 7A.6). However, the multi-species CEATTLE model led to a lower log-likelihood than the single-species model that fixed M (Table 7A.6), owing to an improved fit to the survey biomass, survey age-composition, and fishery length-composition data (Figures 7A.4-6). Both the single-species model that estimated M and the cannibalism model had similar log-likelihoods for the data components, but differed in likelihood penalties. Both the ADMB and single-species CEATTLE model that fixed M (models 1 & 2) had similar Pearson and OSA residual patterns (Figure 7A.4 & 7A.5). OSA residuals from the survey age composition data were smaller for the single-species model that estimated M and the cannibalism model (Figure 7A.5). Females from the survey age-data early in the time series had a positive trend for the 1979 cohort across all models (Figure 7A.4). All models fit the 22nd length bin from males poorly from the fishery length composition data (Figure 7A.6). Due to the large residuals from the fishery length-composition data, OSA residuals did not converge. Both the ADMB and single-species CEATTLE model that fixed M (models 1 & 2) had similar estimates of survey and fishery selectivity (Figure 7A.7). Selectivity in the single-species model that estimated M and the multi-species model (Models 3 & 4) was more asymptotic for the fishery than in the single-species models that fixed M (Models 1 & 2; Figure 7A.7). All models had similar fits to the survey index data (Figure 7A.8). There were no severe retrospective patterns in the single- or multi-species models across 10 annual peels (Figure 7A.9). The single-species model that estimated M and the multi-species model were consistent in their estimates of M and MI , respectively, across peels (Figure 7A.10). The single-species model that estimated M had slightly lower Mohn's rho (0.08) than the single-species model that fixed M and the multi-species model (both at 0.09). However, across both single- and multi-species CEATTLE models, treating annual recruitment deviates as random effects led to lower retrospective bias than when treating them as penalized deviates assuming $\sigma_R = 0.707$. Profiling σ_R when treating recruitment deviates as random effects or as penalized deviates indicated an inability to estimate σ_R in ADMB (Figure 7A.11) because the highest density was at $\sigma_R = 0$.

Discussion and Recommendation

Estimating the single-species arrowtooth flounder stock assessment in TMB using CEATTLE has multiple benefits when compared to the current ADMB model. CEATTLE was developed with three goals in mind: 1) flexibly and easily recreate single-species AFSC groundfish assessments in TMB, 2) allow easy model development and reparameterization (*i.e.* multiple observation models can easily be estimated and compared), and 3) allow estimation of time- and age-varying predation mortality for trophically linked species. Currently, the ADMB assessment has a few elements that are mis-specified that are not present in CEATTLE (differences 1, 4, and 5 above). These can be easily corrected in ADMB, however, CEATTLE also allows for correct estimation of recruitment by treating recruitment deviates as random effects. Likelihood profiles of σ_R show that estimating variance in recruitment is not possible in ADMB. Additionally, comparison of CEATTLE models estimated using penalized maximum likelihood and marginal maximum likelihood show that estimation of σ_R has impacts on derived quantities used for management. Treating recruitment deviates as random effects also leads to lower retrospective bias than estimating the models using penalized maximum likelihood. Bridging also suggests that estimation using CEATTLE is able to find a better fit to the data than the current ADMB model (see Appendix). Using CEATTLE, improvements to the model could be explored including using the marginal maximum likelihood approach to estimate the non-parametric selectivity penalties or explore other time-varying

selectivity functions that may better fit the composition data. We therefore recommend using TMB as the assessment platform via CEATTLE moving forward for GOA ATF.

Accounting for cannibalism in the population dynamics of arrowtooth flounder using CEATTLE demonstrated a better fit to the observed data than the single-species CEATTLE and ADMB models that fixed M . Natural mortality was estimated to be higher than the values assumed by the assessment when estimated in both single- and multi-species frameworks. As a result fishery selectivity was estimated to be more asymptotic than when M was fixed. Additionally, cannibalism in most years has been more than removals by the fishery. However, given the difficulty in defining Tier 3 reference points for multi-species models, using the multi-species CEATTLE model for the assessment may not make sense. Rather, the multi-species model could be used to inform estimates of mortality in the single-species assessment and supplement the assessment by informing on the impacts of both fishery removals and cannibalism on the stock (*i.e.* project the multi-species model with the recommended ABC). For future iterations, it would be useful to update the time-series of diet data in the model, and explore including other time-varying elements (e.g., weight-at-age data, age-length conversion matrices, diet data). All of which can be easily implemented in CEATTLE and are currently used for other models. Therefore, our recommendations for future assessments include:

- Move the stock assessment model to Template Model Builder (TMB) to correctly treat recruitment deviates as random effects and explicitly estimate the associated variance parameter.
- Estimate sex-specific, but age- and time-invariant natural mortality (M).
- Update the weight-at-age and length-at-age transition matrices with recent data and explore the potential for time-varying growth in the assessment model.
- Update the aging error matrix with recent age data.
- Utilize the multi-species model to inform age- and (possibly) time-varying M .
- Explore model sensitivity to different assumptions of catchability (*i.e.* catchability is estimated).

Code and data

<https://github.com/grantdadams/GOA-ATF-ESP/tree/master>

References

- Adams, G. D., Holsman, K. K., Barbeaux, S. J., Dorn, M. W., Ianelli, J. N., Spies, I., Stewart, I. J., & Punt, A. E. (2022). An ensemble approach to understand predation mortality for groundfish in the Gulf of Alaska. *Fisheries Research*, 251(October 2021), 106303. <https://doi.org/10.1016/j.fishres.2022.106303>
- Adams, G., Holsman, K., Hulson, P., Monnahan, C., Shotwell, K., Stewart, I., and Punt, A. (2023) Multispecies model estimates of time-varying natural mortality of groundfish in the Gulf of Alaska. In: Ferriss, B.E. 2023. Ecosystem Status Report 2023: Gulf of Alaska, Stock Assessment and Fishery Evaluation Report, North Pacific Fishery Management Council, 1007 West Third, Suite 400, Anchorage, Alaska 99501.
- Anderson, P. J., & Piatt, J. F. (1999). Community reorganization in the Gulf of Alaska following ocean climate regime shift. *Marine Ecology Progress Series*, 189, 117-123.
- Barnes, C. L., Beaudreau, A. H., Hunsicker, M. E., & Ciannelli, L. (2018). Assessing the potential for competition between Pacific Halibut (*Hippoglossus stenolepis*) and Arrowtooth Flounder (*Atheresthes stomias*) in the Gulf of Alaska. *PloS one*, 13(12), e0209402.
- Clark, W. G., & Hare, S. R. (2002). Effects of climate and stock size on recruitment and growth of Pacific halibut. *North American Journal of Fisheries Management*, 22(3), 852-862.

- Curti, K. L., Collie, J. S., Legault, C. M., & Link, J. S. (2013). Evaluating the performance of a multispecies statistical catch-at-age model. *Canadian Journal of ...*, 484(January), 470–484. <https://doi.org/10.1139/cjfas-2012-0229>
- Doyle, M.J., Debenham, C., Barbeaux, S.J., Buckley, T.W., Pirtle, J.L., Spies, I.B., Stockhausen, W.T., Shotwell, S.K., Wilson, M.T., Cooper, D.W., 2018. A full life history synthesis of Arrowtooth Flounder ecology in the Gulf of Alaska: exposure and sensitivity to potential ecosystem change. *J. Sea Res.* 142, 28–51. <https://doi.org/10.1016/j.seares.2018.08.001>.
- Ferriss, B.E. (2023) Ecosystem Status Report 2023: Gulf of Alaska, Stock Assessment and Fishery Evaluation Report, North Pacific Fishery Management Council, 1007 West Third, Suite 400, Anchorage, Alaska 99501.
- Fournier, D.A., Skaug, H.J., Ancheta, J., Ianelli, J., Magnusson, A., Maunder, M.N., Nielsen, A., and Sibert, J. 2012. AD Model Builder: using automatic differentiation for statistical inference of highly parameterized complex nonlinear models. *Optim. Methods Softw.* 27:233-249.
- Holsman, K. K., & Aydin, K. (2015). Comparative methods for evaluating climate change impacts on the foraging ecology of Alaskan groundfish. *Marine Ecology Progress Series*, 521, 217-235.
- Holsman, K. K., Ianelli, J., Aydin, K., Punt, A. E., & Moffitt, E. A. (2016). A comparison of fisheries biological reference points estimated from temperature-specific multi-species and single-species climate-enhanced stock assessment models. *Deep Sea Research Part II: Topical Studies in Oceanography*, 134, 360–378. <https://doi.org/10.1016/j.dsr2.2015.08.001>
- Holsman, K. K., Barnes, C., Aydin, K., Laurel, B., Hurst T., & Heintz T. (2023). Patterns in Foraging and Energetics of Walleye Pollock, Pacific Cod, Arrowtooth Flounder, and Pacific Halibut. In: Elizabeth Sidden (Ed.), *Ecosystem Status Report 2023: Eastern Bering Sea, Stock Assessment and Fishery Evaluation Report*, North Pacific Fishery Management Council, 1007 West 3rd Ave., Suite 400, Anchorage, Alaska 99501.
- Hulson, P.-J. F., Barbeaux, S. J., Ferriss, B., Echave, K., Nielsen, J., Shotwell, S. K., Laurel, B., Spies, I., Litzow, M., Mcdermott, S., Rand, K., Levy, C., Wang, M., & Rogers, L. (2023). *Assessment of the Pacific cod stock in the Gulf of Alaska Pacific cod*.
- Jurado-Molina, J., Livingston, P. A., & Ianelli, J. N. (2005). Incorporating predation interactions in a statistical catch-at-age model for a predator-prey system in the eastern Bering Sea. *Canadian Journal of Fisheries and Aquatic Sciences*, 62(8), 1865–1873. <https://doi.org/10.1139/f05-110>
- Kinzey, D., & Punt, A. E. (2009). Multispecies and Single-Species Models of Fish Population Dynamics: Comparing Parameter Estimates. *Natural Resource Modelling*, 22(1), 67–104.
- Kristensen, K., Nielsen, A., Berg, C. W., Skaug, H., & Bell, B. M. (2016). TMB: Automatic Differentiation and Laplace Approximation. *Journal of Statistical Software*, 70(5), 1–21. <https://doi.org/10.18637/jss.v070.i05>
- Magnusson, K. (1995). An overview of the multispecies VPA—theory and applications. *Reviews in Fish Biology and Fisheries*, 212, 195–212. <https://doi.org/10.1007/BF00179756>
- McAllister, M. K., & Ianelli, J. N. (1997). Bayesian stock assessment using catch-age data and the sampling-importance resampling algorithm. *Canadian Journal of Fisheries and Aquatic Sciences*, 54(2), 284-300.
- NPFMC. (2017). *Fishery Management Plan for Groundfish of the Bering Sea and Aleutian Islands Management Area*.
- NPFMC. (2019). *Fishery Management Plan for Groundfish of the Gulf of Alaska. August*.
- Shotwell, S. K., Spies, I., Ianelli, J. N., Aydin, K., Hanselman, D. H., Palsson, W., Siwicke, K., Sullivan, J., & Yasumiishi, E. (2021). *Assessment of the Arrowtooth flounder stock in the Gulf of Alaska*. <https://apps-afsc.fisheries.noaa.gov/refm/docs/2019/GOAatf.pdf>
- Shotwell, S.K., J.L. Pirtle, J. Watson, A. Deary, M. Doyle, S.J. Barbeaux, M. Dorn, G.A. Gibson, E. Goldstein, D.H. Hanselman, A.J. Hermann, P.J.F. Hulson, B. Laurel, J.H. Moss, O. Ormseth, D. Robinson, L.A. Rogers, C.N. Rooper, I. Spies, W. Strasburger, R.M. Suryan, J.J. Vollenweider (2022) Synthesizing integrated ecosystem research to create informed stock-specific indicators for

next generation stock assessments. *Deep-Sea Res. Pt. II: Top. Studies. Oceanogr.* 198. DOI: <https://doi.org/10.1016/j.dsr2.2022.105070>

Shotwell, S., Hanselman, D., Siwicke, K., Spies, I. and Sullivan, J. (2023) Assessment of the arrowtooth flounder stock in the Gulf of Alaska. In: Stock assessment and fishery evaluation report for the groundfish resources of the Gulf of Alaska. North Pacific Fishery Management Council, Anchorage, AK.

Wilderbuer, T.K. and B.J. Turnock. 2009. Differential Sex-Specific Natural Mortality of Arrowtooth Flounder in Alaska: Implications for Exploitation and Management. *North American Journal of Fisheries Management* Vol. 29 (2).

Tables

Table 7A.1. General population dynamics equations in CEATTLE for species i , sex s , age a , and length l in year y , and for survey or fishery f_i . See Table 4 for parameter definitions.

Definition	Equation
Initial abundance	$N_{isa1} = \left\{ \begin{array}{l} R_{-F_i} e^{\left(-\sum_{j=1}^{a-1} M1_{isj} + \tau_{ia} \right)} \rho_{i1} \text{ if } a < A_i \\ \frac{R_{-F_i} e^{\left(-\sum_{j=1}^{a-1} M1_{isj} + \tau_{ia} \right)} \rho_{i1}}{\left(1 - e^{\left(-M1_{isa} \right)} \right)} \text{ if } a = A_i \end{array} \right.$
Recruitment	$N_{is1y} = R_{isy} = R_{-i} e^{\tau_{iy}} \rho_{i1}$
Numbers at age	$N_{isa+1y+1} = N_{isay} e^{-Z_{isay}}$ $N_{isA_y+1} = N_{isA_{-1}y} e^{-Z_{isA_{-1}y}} + N_{isA_y} e^{-Z_{isA_y}}$
Survey biomass (kg)	$CPUE: \hat{B}_{f_i y} = \sum_s \sum_a N_{isay} e^{-Month_{f_i} * Z_{isay}} W_{f_i say} sel_{f_i say} q_{f_i y}$
Catch-at-age	$C_{f_i say} = \frac{F_{f_i say}}{Z_{isay}} \left(1 - e^{-Z_{isay}} \right) N_{isay}$
Total catch (kg)	$C_{f_i y}^* = \sum_s \sum_j \frac{F_{f_i say}}{Z_{isay}} \left(1 - e^{-Z_{isay}} \right) N_{isay} W_{f_i say}$
Age-1+ biomass (kg)	$B_{iy} = \sum_s \sum_{a>0} N_{isay} W_{B_{ay}}$
Spawning biomass at age (kg)	$SB_{iay} = N_{i1ay} e^{-Month_{ssb_i} / 12 * Z_{i1ay}} W_{ssb_{i1ay}} \delta_{ia}$
Total mortality at age (single-species)	$Z_{isay} = M1_{isa} + \sum_{f_i} F_{f_i say}$
Total mortality at age (multi-species)	$Z_{isay} = M1_{isa} + M2_{isay} + \sum_{f_i} F_{f_i say}$
Fishing mortality at age and fleet	$F_{f_i say} = F_{-f_i} e^{\epsilon_{f_i y}} sel_{f_i say}$
Projected fishing mortality	$F_{f_i say} = F_{Target_i} sel_{f_i say} \Psi_{f_i}$
Predicted age composition	$\hat{O}_{f_i say} = \frac{\sum_a C_{f_i say} AE_{iaa}}{\sum_s \sum_a \sum_{f_i} C_{f_i say} AE_{iaa}} \quad \text{if } f_i = \text{fishery}$ $\hat{O}_{f_i say} = \frac{\sum_a N_{isay} e^{-Month_{f_i} * Z_{isay}} sel_{f_i say} AE_{iaa}}{\sum_s \sum_a \sum_{f_i} N_{isay} e^{-Month_{f_i} * Z_{isay}} sel_{f_i say} AE_{iaa}} \quad \text{if } f_i = \text{survey}$
Predicted length composition	$\hat{O}_{f_i sly} = \frac{\sum_a \sum_{f_i say} C_{f_i say} AE_{iaa} \cdot ATM_{isal}}{\sum_s \sum_l \sum_a \sum_{f_i say} C_{f_i say} AE_{iaa} \cdot ATM_{isal}} \quad \text{if } f_i = \text{fishery}$

$$\hat{O}_{f_i, sly} = \frac{\sum_a \sum_{isay} N_{isay} e^{-\text{Month}_{f_i} \cdot Z_{isay}} \text{sel}_{f_i, say} AE_{iaa} \cdot ATM_{isal}}{\sum_s \sum_l \sum_a \sum_{isay} N_{isay} e^{-\text{Month}_{f_i} \cdot Z_{isay}} \text{sel}_{f_i, say} AE_{iaa} \cdot ATM_{isal}} \text{ if } f_i = \text{survey}$$

Table 7A.2. Predation mortality equations for predator species p , sex b , age j , and prey species i , sex s and age a . For parameter definitions see Table 4.

Definition	Equation
Predation mortality	$M2_{isay} = \sum_{pbj} \left(\frac{N_{pbjy} \delta_{pbjy} \hat{S}_{pbjisa}}{\left(\sum_{isa} \left(\hat{S}_{pbjisa} B_{isay} \right) + B_p^{other} \left(1 - \sum_{isa} \left(\hat{S}_{pbjisa} \right) \right) \right)} \right)$
Predator-prey suitability	$S_{pbjisa}^{\wedge} = \frac{1}{n_y} \sum_y \left(\frac{\frac{U_{psaisa}}{w_{isay} N_{isay}^{n_i}}}{\sum_{isa} \left(\frac{U_{pbjisa}}{w_{isay}} \right) + \frac{1 - \sum_{isa} U_{pbjisa}}{B_p^{other}}} \right)$
Individual specific ration ($kg\ kg^{-1}\ yr^{-1}$)	$\delta_{pbjy} = \varphi_{pbjy} \alpha_p^{\delta} W_{pbjy}^{(1+\beta_p)} f(T_y)_p$
Temperature scaling algorithms	$f(T_y)_p = V^X e^{(X(1-V))}$
...	$V = (T_p^{cm} - T_y) / (T_p^{cm} - T_p^{co})$
...	$X = \left(Z^2 (1 + (1 + 40/Y)^{0.5})^2 \right) / 400$
...	$Z = \ln(Q_p^c) (T_p^{cm} - T_p^{co})$
...	$Y = \ln(Q_p^c) (T_p^{cm} - T_p^{co} + 2)$

Table 7A.3. Components of the likelihood function for species i , sex s , age a , or length l during year y for survey s or fishery f . For parameter definitions see Table 4.

Description	Equation
Data components	
Survey biomass	$\ln(B_{f,y}) \sim N\left(\ln(\hat{B}_{f,y}) - \frac{\sigma_{f,y}^2}{2}, \sigma_{f,y}^2\right)$
Total catch	$\ln(C_{f,y}^*) \sim N\left(\ln(\hat{C}_{f,y}^*) - \frac{\sigma_{f,y}^2}{2}, \sigma_{f,y}^2\right)$
Age composition	$-\sum_i n_{f,y} \sum_s \sum_{\hat{a}} \sum_y \left(\hat{O}_{f_i s \hat{a} y} + 0.00001\right) * \ln\left(O_{f_i s \hat{a} y} + 0.00001\right)$
Length composition	$-\sum_i n_{f,y} \sum_s \sum_l \sum_y \left(\hat{O}_{f_i s l y} + 0.00001\right) * \ln\left(O_{f_i s l y} + 0.00001\right)$
Penalties	
Non-parametric selectivity	$\sum_i \sum_s \sum_a^{A_i-1} \chi_{f_i} \left[\ln\left(\frac{e^{\psi_{f_i s a}}}{e^{\psi_{f_i s a+1}}}\right) - \ln\left(\frac{e^{\psi_{f_i s a+1}}}{e^{\psi_{f_i s a+2}}}\right) \right]^2$ $\chi_{f_i} = 0 \quad \text{if } e^{\psi_{f_i s a}} \leq e^{\psi_{f_i s a+1}}$
Recruitment deviate	$\tau_{iy} \sim N\left(-\frac{\sigma_{R_i}^2}{2}, \sigma_{R_i}^2\right)$
Fishing mortality deviate	$\epsilon_{f,y} \sim \epsilon_{f,y} * \epsilon_{f,y}$

Table 7A.4. Parameter definition for CEATTLE.

Category	Parameter	Definition
Index	i	Species
	s	Sex
	a	Age
	\hat{a}	Observed age
	A	Plus group
	l	Length
	y	Year
	Y	Last year of estimation (not projection)
	p	Predator species
	b	Predator sex
	j	Predator age
	k	Predator length
	f_i	Fleet/survey
Population model	R_{isy}	Recruitment
	\bar{R}_i	Mean recruitment
	τ_{iy}	Annual recruitment deviate
	δ_{ia}	Maturity-at-age
	ρ_{ia}	Sex-ratio
	$M1_{isa}$	Residual mortality yr ⁻¹
	$M2_{isay}$	Predation-mortality yr ⁻¹
	$F_{f_i say}$	Fleet-specific fishing mortality yr ⁻¹
	F_{isay}	Total fishing mortality yr ⁻¹
	Z_{isay}	Total mortality yr ⁻¹
	\bar{F}_{f_i}	Mean fishing mortality yr ⁻¹
	$\epsilon_{f_i y}$	Annual fishing mortality deviate
	ψ_{f_i}	Fishing mortality apportionment per fleet
	N_{isay}	Numbers-at-age
	\bar{N}_{isay}	Average numbers-at-age
	SB_{ijy}	Spawning-stock-biomass (kg)
	B_{isay}	Biomass (kg)
	\bar{B}_{isay}	Average biomass (kg)
Predation model	\hat{S}_{pbjisa}	Predator-prey suitability
	H_i	Holling functional response parameter
	B_p^{other}	Biomass of “other prey” (kg)
	T_y	Observed bottom temperature (°C)
	T_p^{co}	Thermal optima for consumption (°C)
	T_p^{cm}	Thermal limit of consumption (°C)

	Q_p^c	Rate at which consumption increased over relatively low water temperatures
	$\alpha_p^\delta, \beta_p^\delta$	Weight specific intercept and slope of maximum consumption
	φ_{pbj}	Scalar for maximum to observed consumption
	$\frac{U}{pbj,isa}$	Average proportion of prey-at-age in the stomach of a predator-at-age
Observation model		
	$CPUE: B_{f,y}$	Survey/fishery relative biomass (kg)
	$sel_{f,say}$	Selectivity
	$q_{f,y}$	Catchability
	W	Weight (kg)
	$C_{f,say}$	Catch-at-age
	$C_{f,y}^*$	Total catch (kg)
	$\hat{O}_{f,say} / \hat{O}_{f,sly}$	Predicted age/length composition
	AE_{iaa}	Ageing error matrix
	ATM_{isal}	Age-transition matrix (growth trajectory)
	\hat{S}_{pbjisa}	Predator-prey suitability

Table 7A.5. Time-series and sources of data used in CEATTLE model

Source	Data type	Data collection cycle
NMFS bottom trawl survey	Absolute biomass	Triennial and Biennial (1984-2021)
	Age composition	1984-2019
	Diet composition	1990-2015
Holsman and Aydin (2015)	Bioenergetic demand	1990-2015
Fishery	Total catch	Annual (1977-2023)
	Length composition	1977-2020

Table 7A.6. Model comparison table. Note, likelihoods from the ADMB model are not comparable with CEATTLE.

	ADMB	CEATTLE Single-spp (fix M)	CEATTLE Single-spp (est M)	CEATTLE multi-spp
Platform	ADMB	TMB	TMB	TMB
<i>Likelihoods</i>				
Catch	4.31803e-08	-97.51	-97.38	-94.38
Fishery length composition	816.17	83.32	63.94	63.76
Survey biomass	27.86	1.16	-3.94	-4.44
Survey age composition	276.91	43.36	40.49	40.54
Recruit deviations	5.04	24.30	19.26	21.96
Selectivity penalties	7.30	9.91	6.48	6.95
F regularity	21.20	20.44	21.08	22.20
Joint NLL	177.67	84.99	49.94	56.61
Marginal NLL		96.47	66.18	68.92
Number of parameters	158	92	94	94
<i>Estimates</i>				
q -trawl	1	1.0	1.0	1.0
Mean Recruitment (thousand)	801,987	806,846	1,921,099	2,451,550
σ_r	0.707	0.477	0.445	0.462
2023 total biomass (t)	1,276,460	1,273,952	1,631,329	1,747,704
2023 spawning biomass (t)	696,871	718,708	550,638	548,410
$B_{100\%}$ (t)	1,018,700	1,684,452	2,045,104	2,141,534
$B_{40\%}$ (t)		673,781	818,042	856,614
$SSB_{100\%}$ (t)		984,431	752,438	738,569
$SSB_{40\%}$ (t)	407,478	393,772	300,975	295,428

Figures

Figure 7A.1. Time-series of estimated age-1 recruitment (millions; top panel), total biomass (million mt; middle panel), spawning stock biomass (million mt; bottom panel) as determined from the 2023 ADMB assessment, CEATTLE single-species fixed M and estimated M, and multi-species (cannibalism) models.

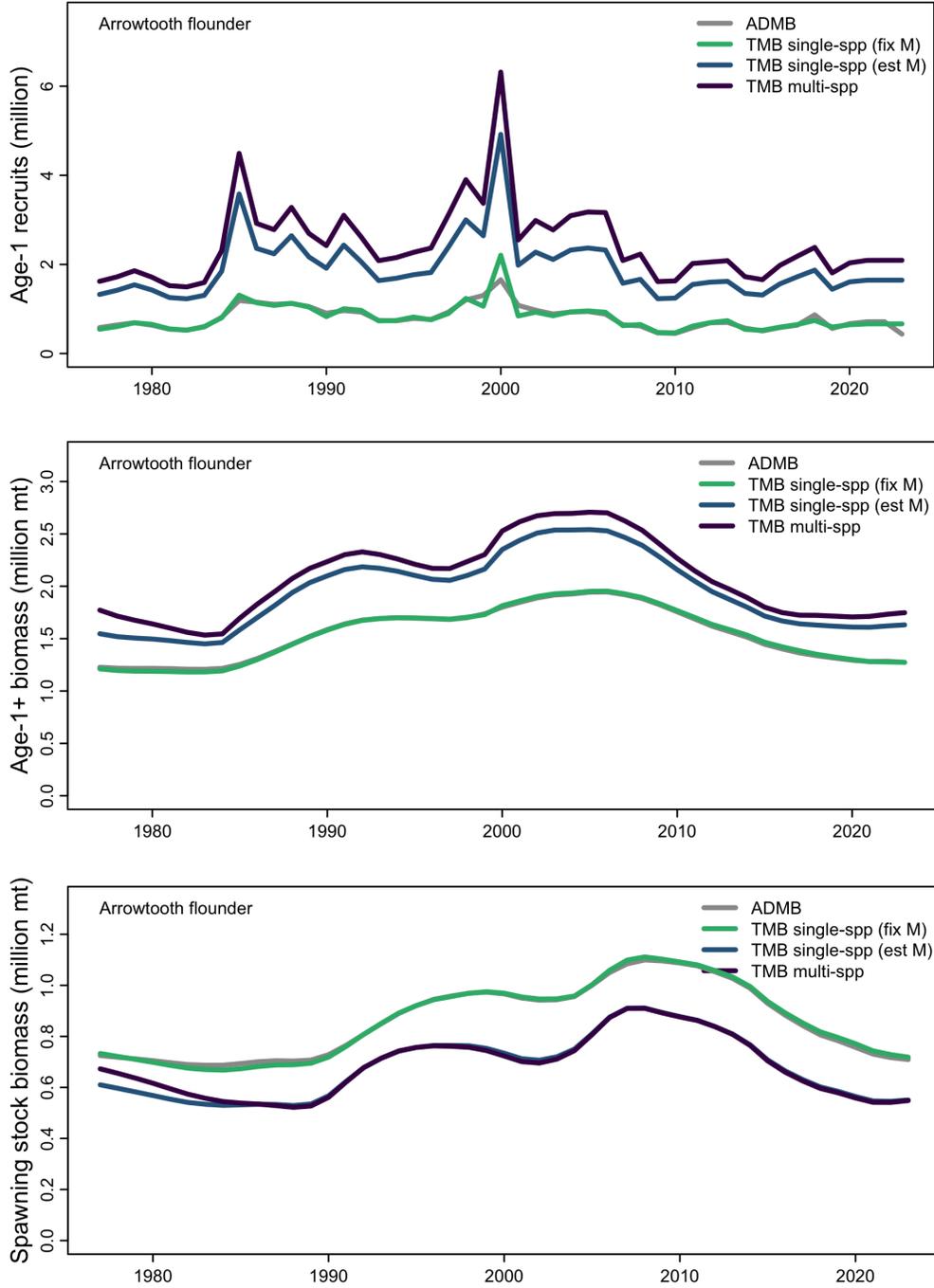


Figure 7A.2a. Time-series of total natural mortality at age-1 (top-left panel), age-2 (top-right panel), age-3 (bottom-left panel), and age-4 (bottom-right panel) from the single- and multi-species (cannibalism) CEATTLE models. For the single-species ADMB and CEATTLE single-species with fixed M , natural mortality is set at 0.2 for females and 0.35 for males.

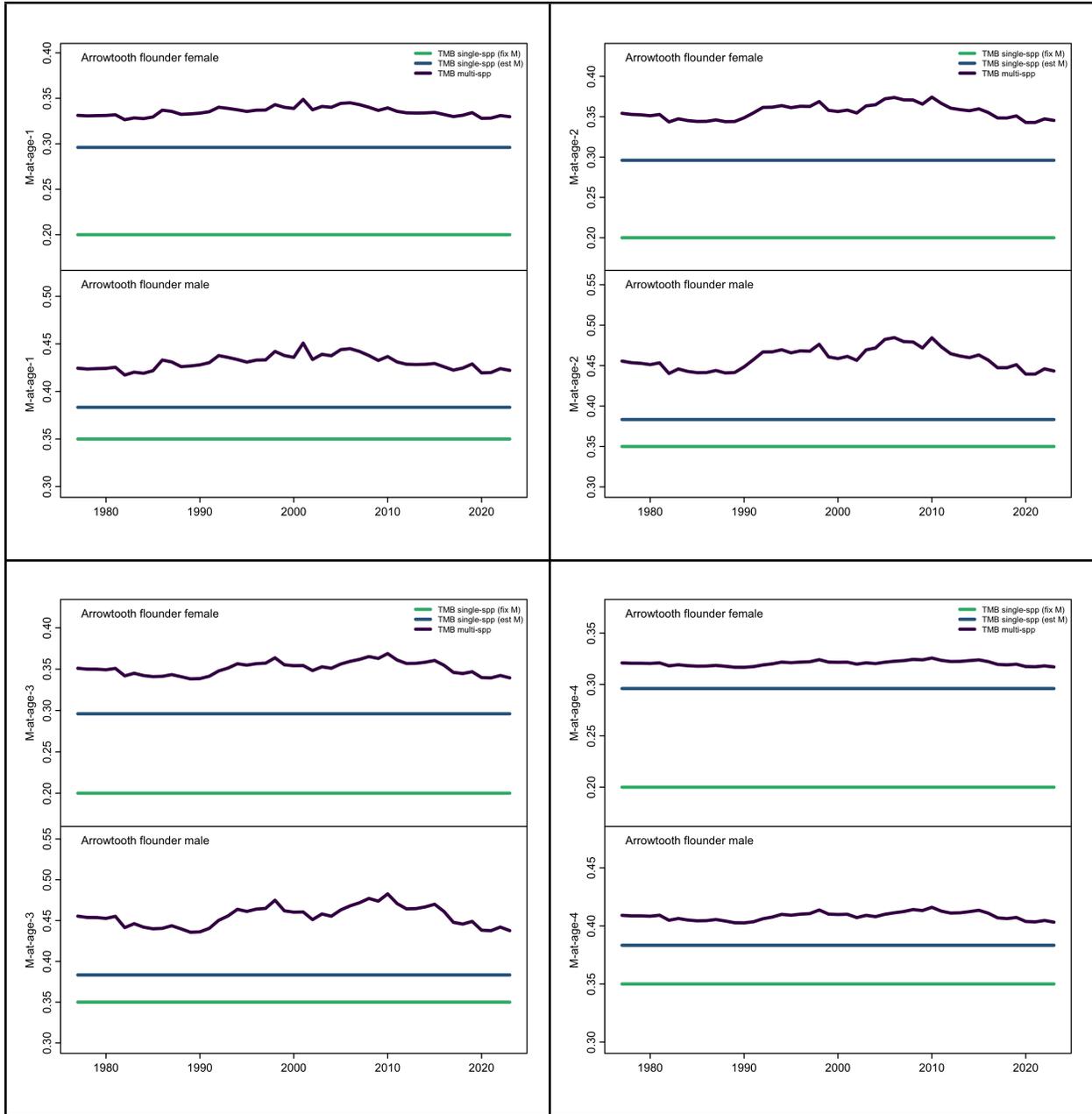


Figure 7A.2b. Time-series of total natural mortality at age-5 (top-left panel), age-6 (top-right panel), age-7 (bottom-left panel), and age-8 (bottom-right panel) from the single- and multi-species (cannibalism) CEATTLE models. For the single-species ADMB and CEATTLE single-species with fixed M , natural mortality is set at 0.2 for females and 0.35 for males.

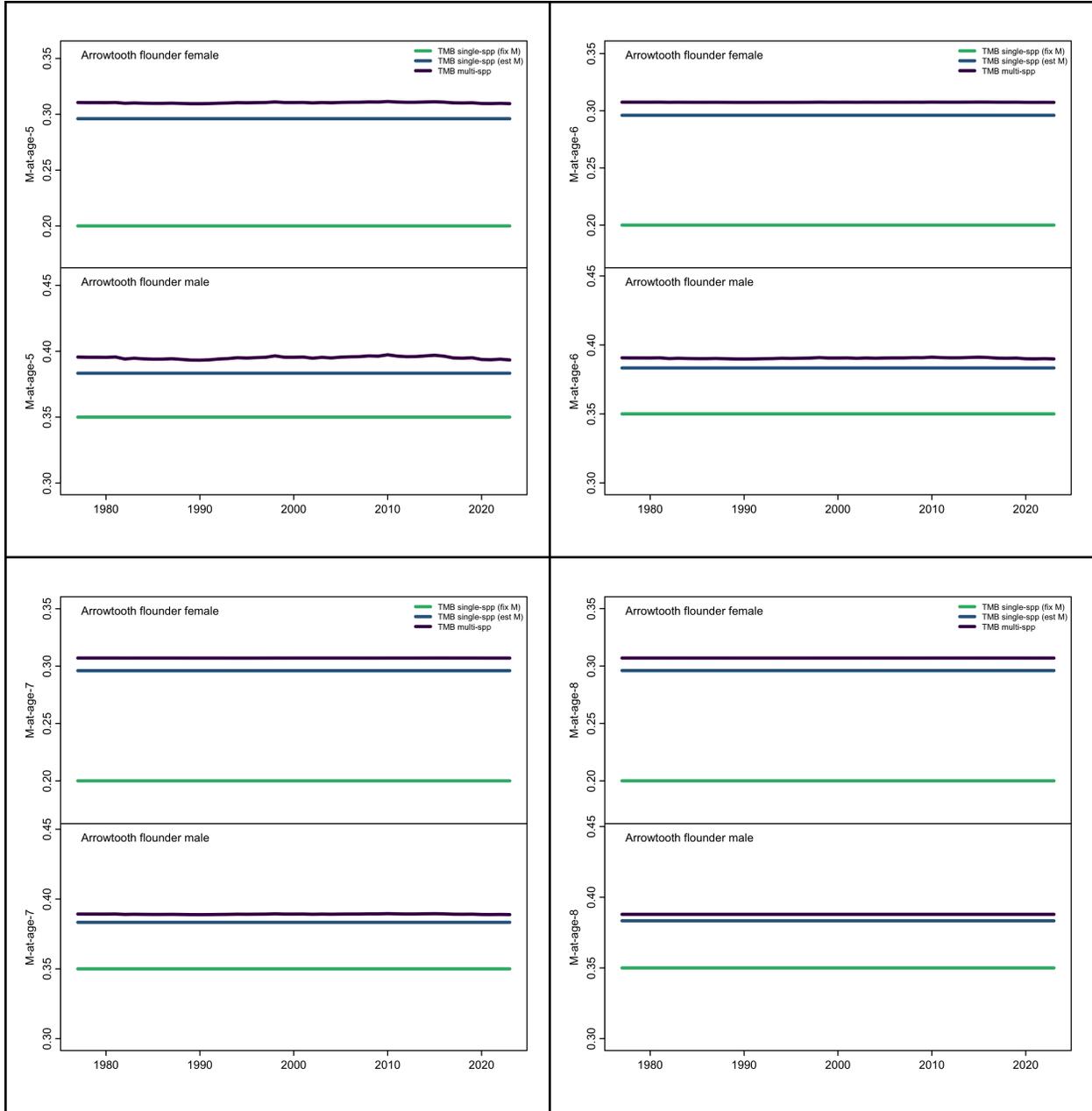


Figure 7A.3. Time-series of estimated and observed catch (mt) and biomass consumed (mt) as prey as determined from the 2023 ADMB assessment and CEATTLE single- and multi-species (cannibalism) models.

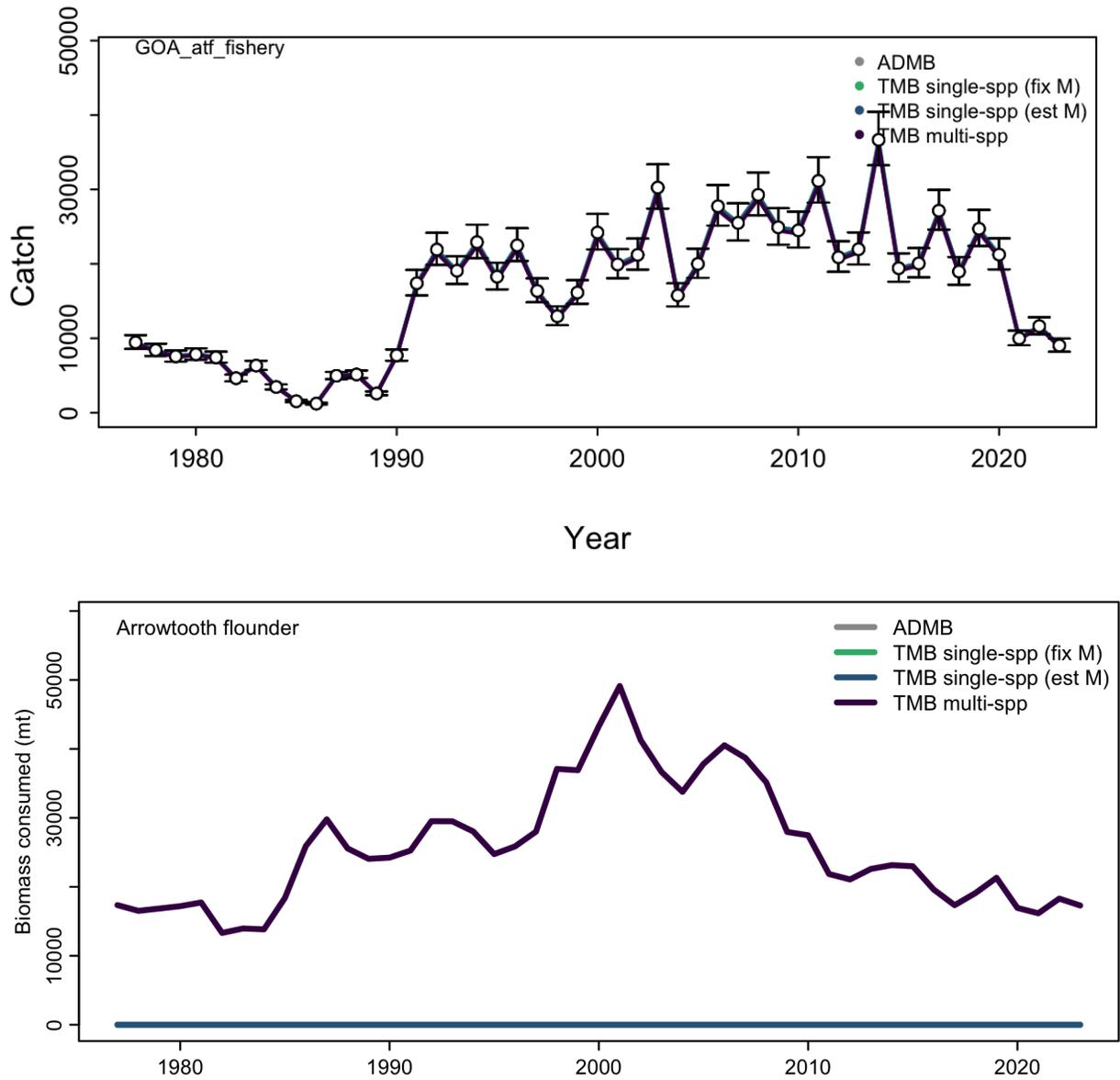
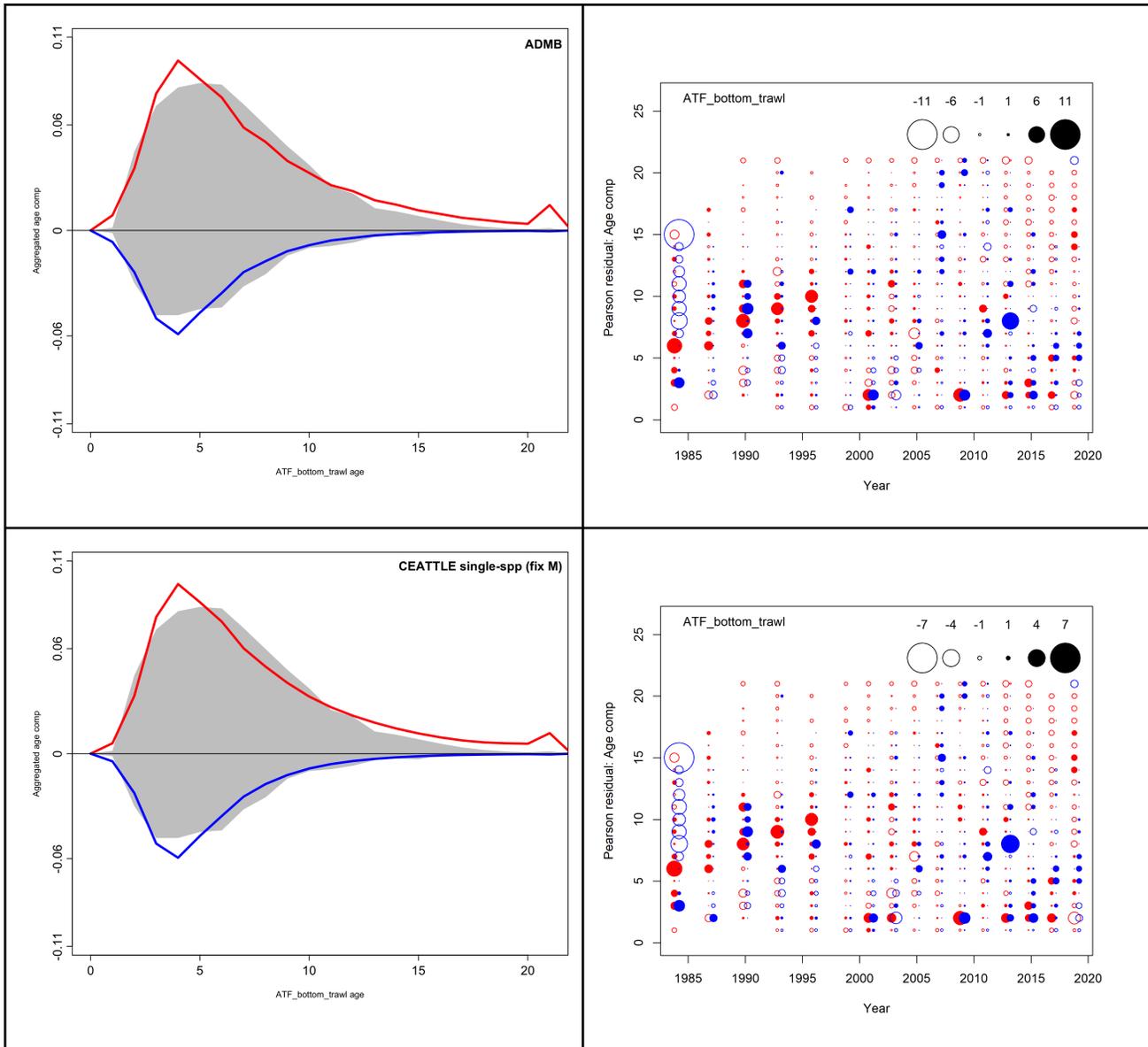


Figure 7A.4. Fits to aggregated and pearson residuals of observed and expected bottom trawl survey age-composition data between the ADMB model and CEATTLE single- and multi-species models. Females are red and males are blue. NOTE: difference in z-scale.



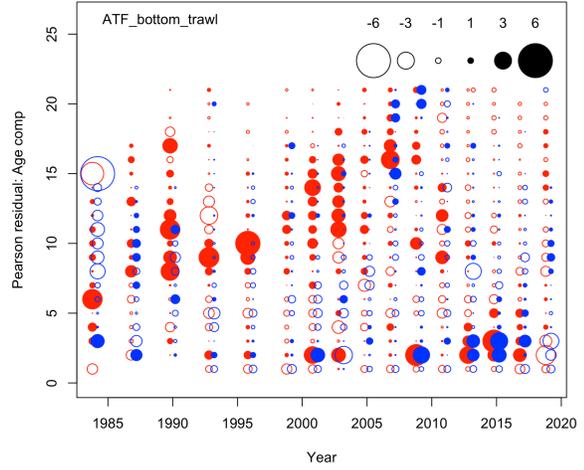
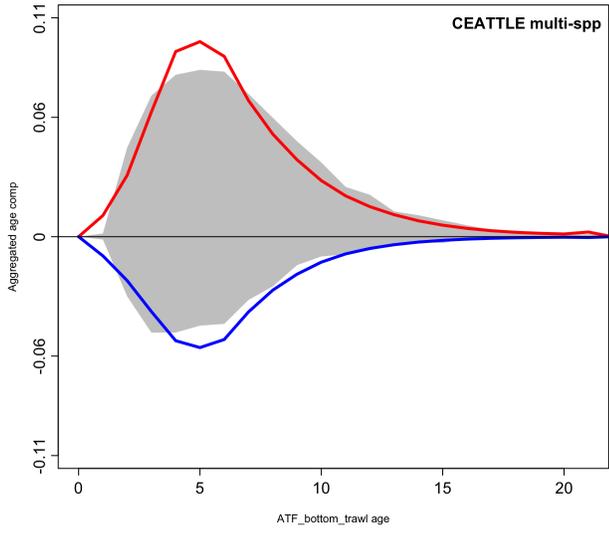
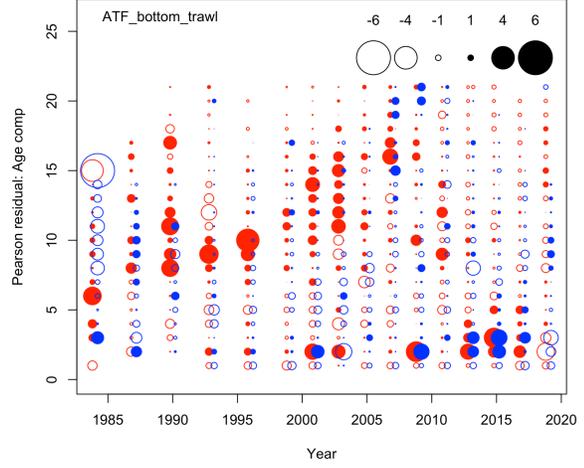
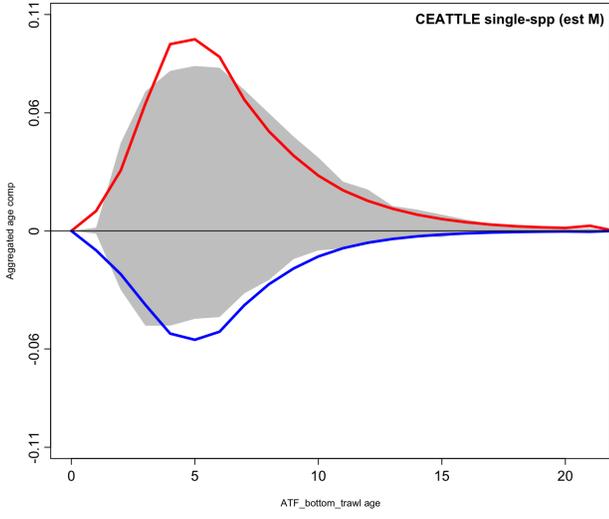


Figure 7A.5. One step ahead (OSA) residuals of observed and expected bottom trawl survey age-composition data between the ADMB model and CEATTLE single- and multi-species models. NOTE: difference in z-scale.

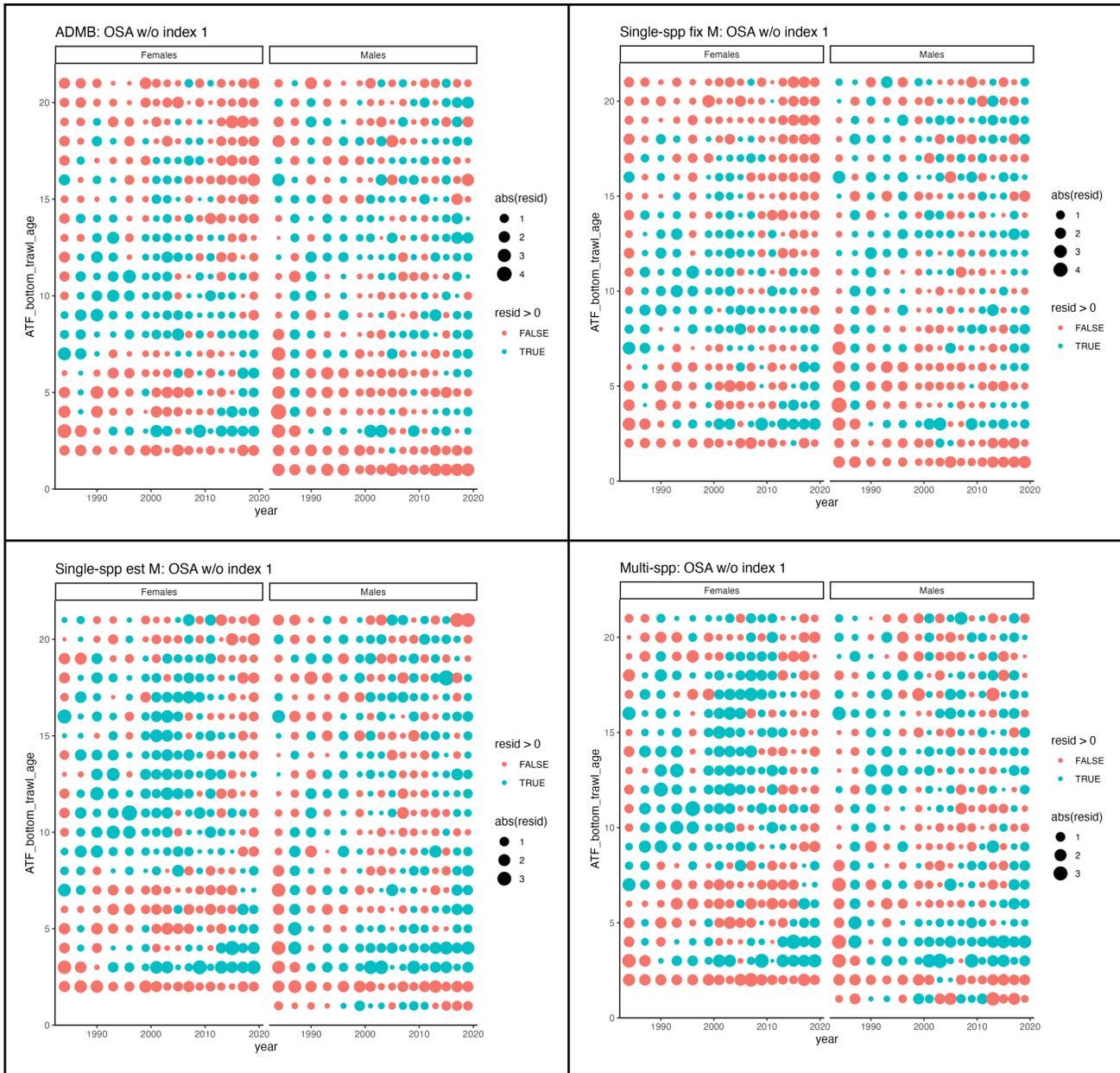
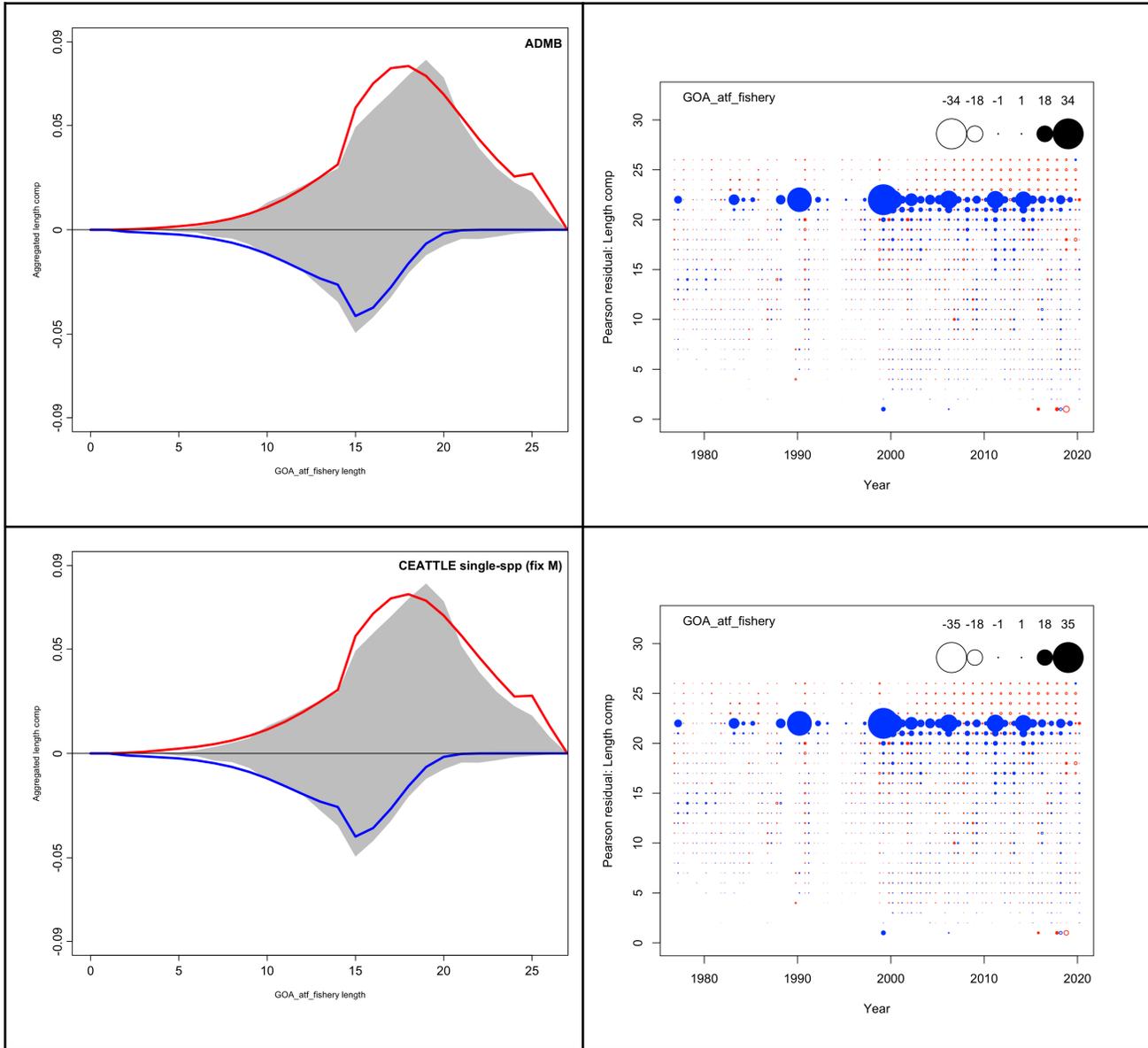


Figure 7A.6. Fits to aggregated and pearson residuals of observed and expected fishery length-composition data between the ADMB model and CEATTLE single- and multi-species models. Females are red and males are blue. NOTE: difference in z-scale.



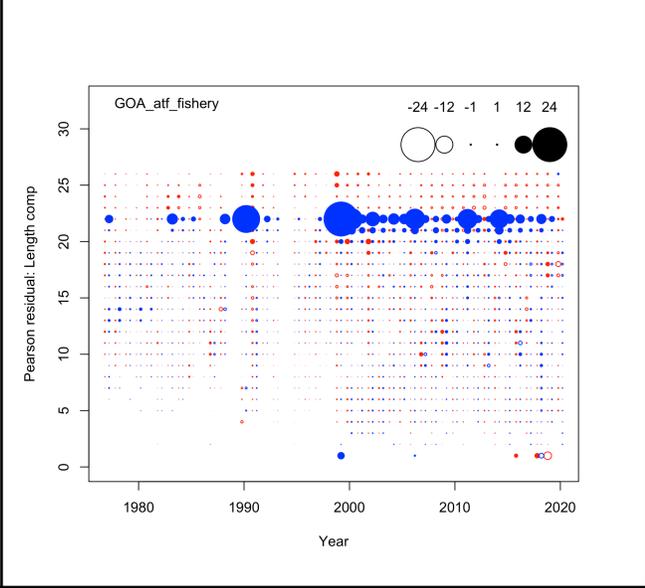
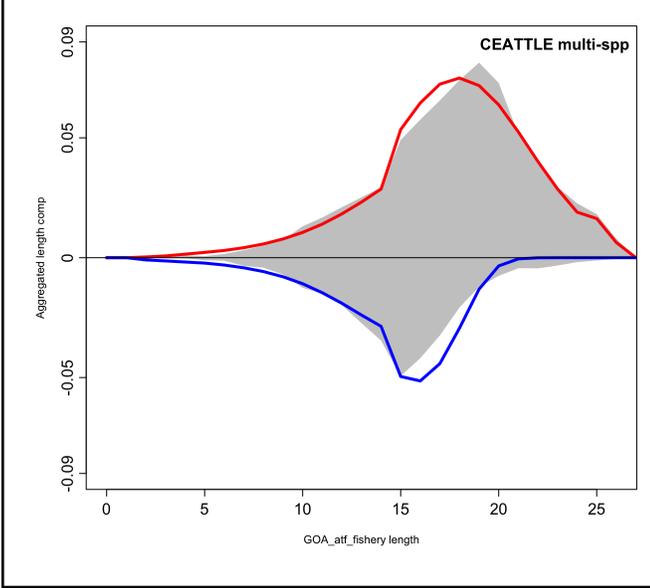
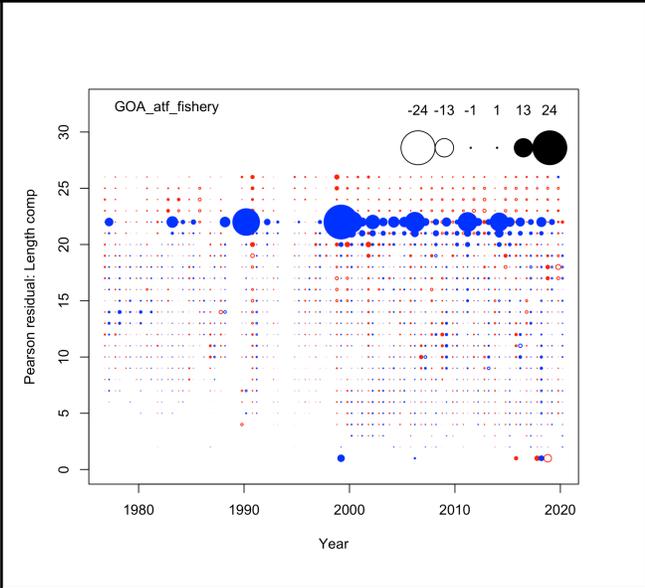
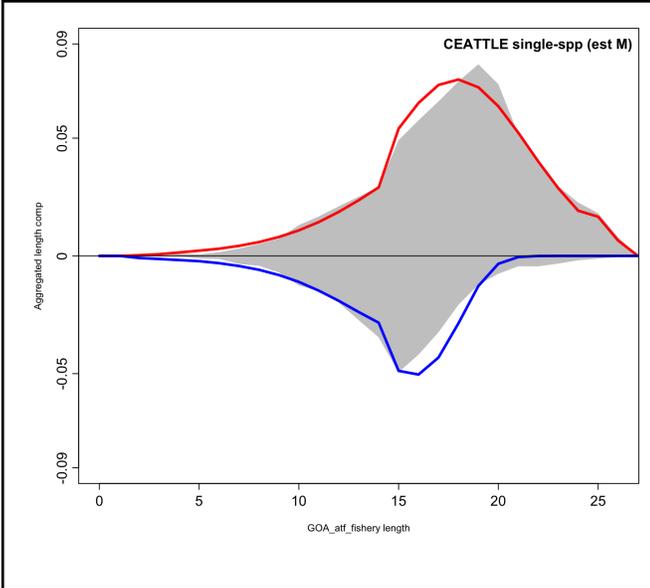
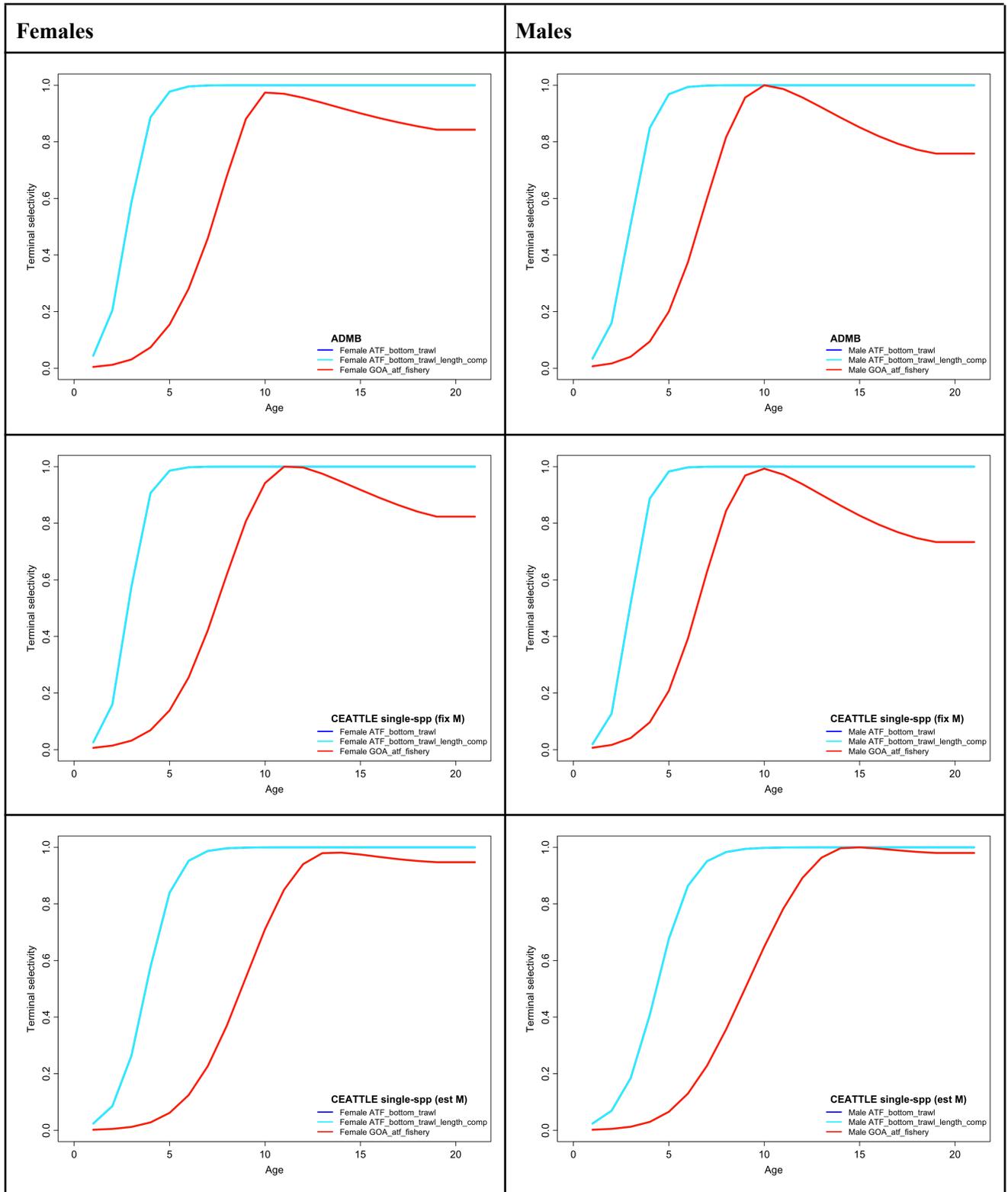


Figure 7A.7. Estimated selectivity between the ADMB model and CEATTLE single- and multi-species models.



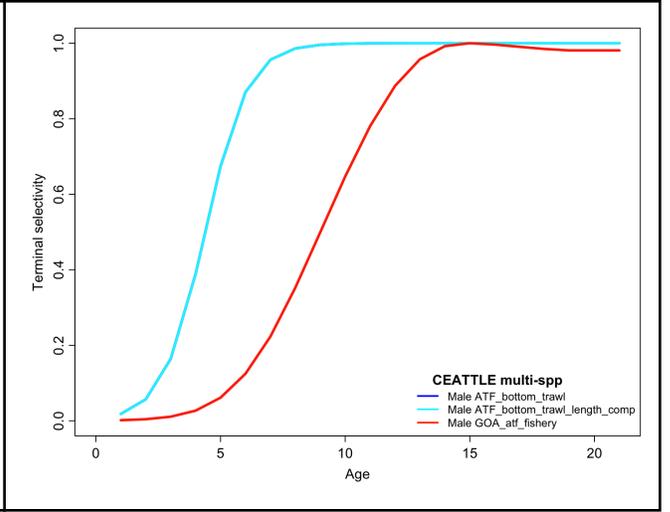
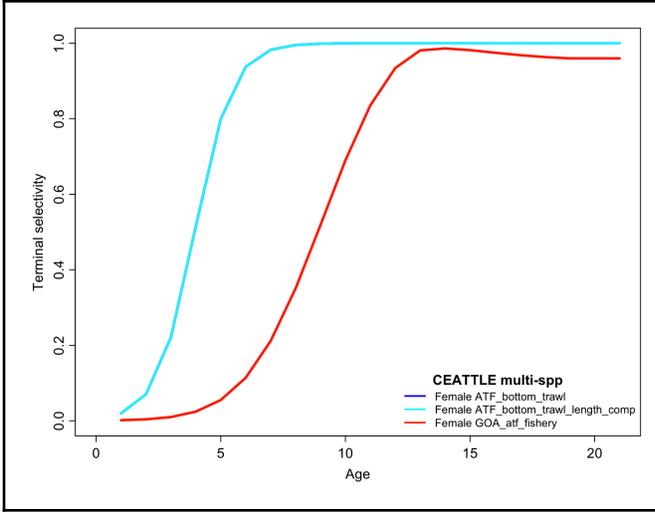


Figure 7A.8. Observed (circles) and expected (lines) AFSC trawl survey biomass across the ADMB model and CEATTLE single- and multi-species models.

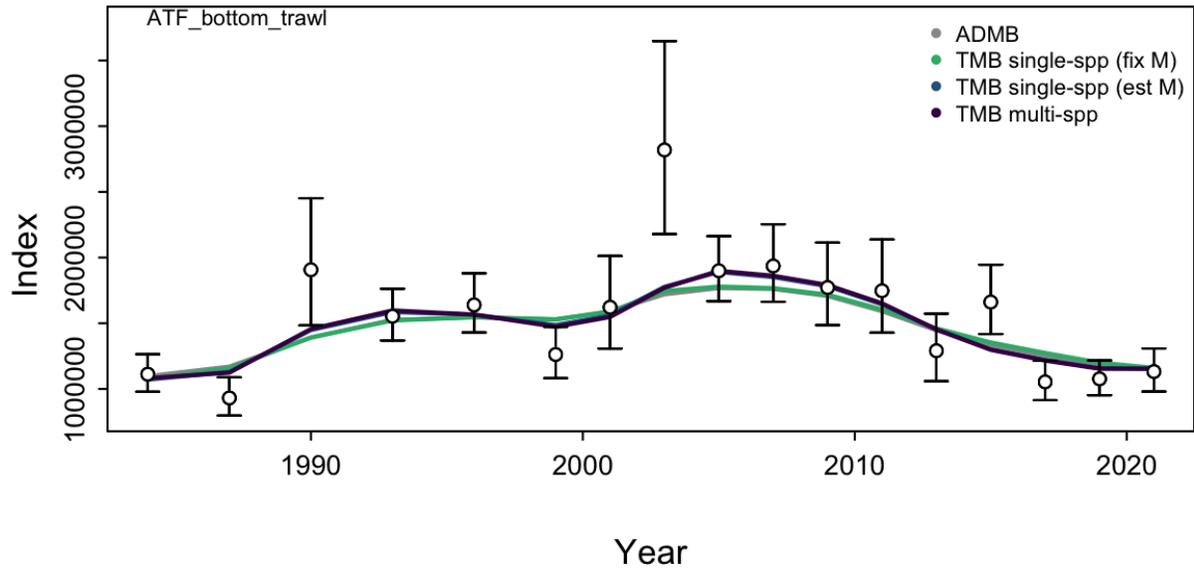


Figure 7A.9. Retrospective spawning age-1+ total biomass estimates across CEATTLE single- and multi-species models. Mohn's rho is indicated in the upper right hand corner. The left column are models where recruitment deviates are treated as penalized deviates (similar to ADMB) and the right column are models where recruitment deviates are treated as random effects.

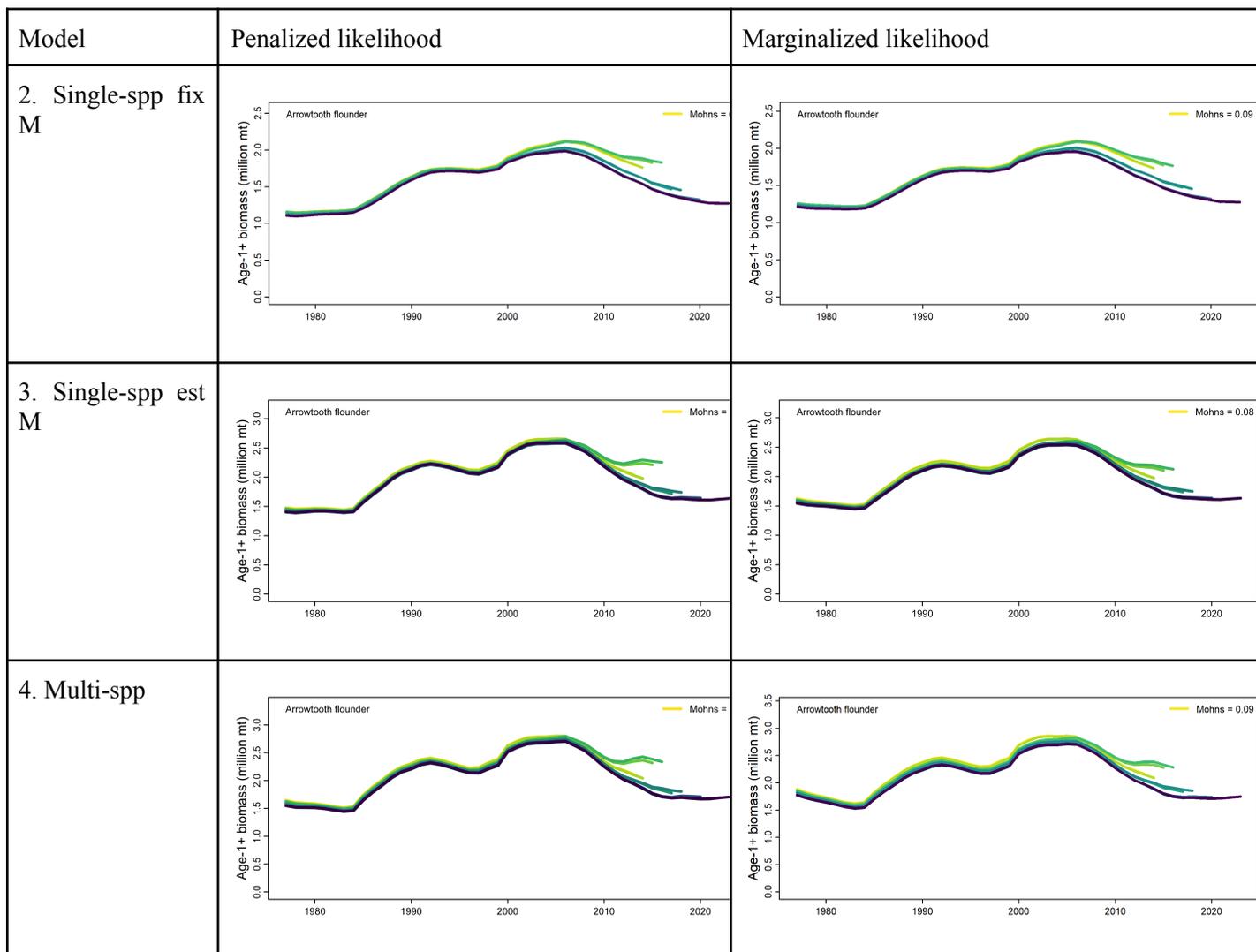


Figure 7A.10. Retrospective estimates of total natural mortality (M) and residual natural mortality ($M1$) for CEATTLE single-species models that estimated M and multi-species models where recruitment deviates are treated as random effects.

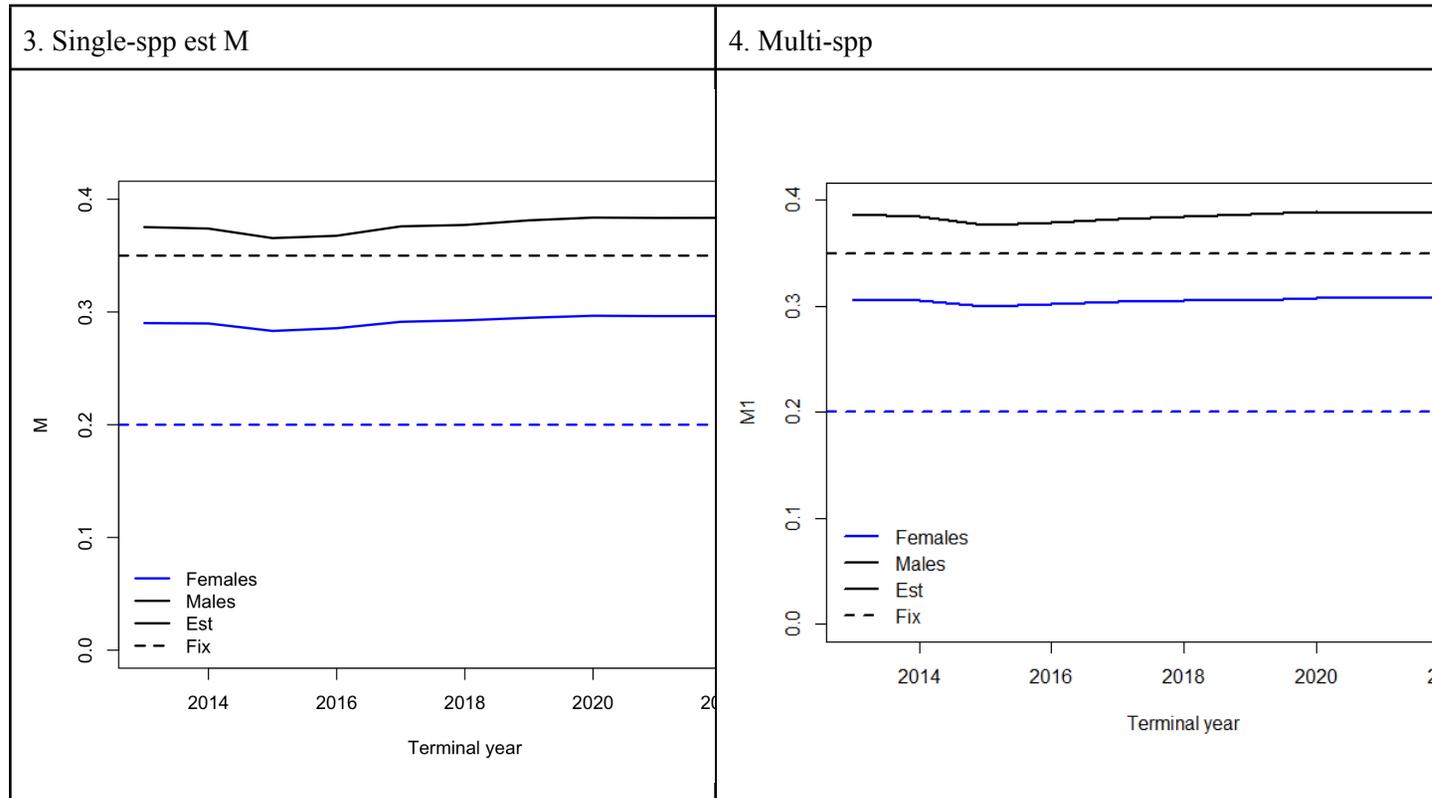
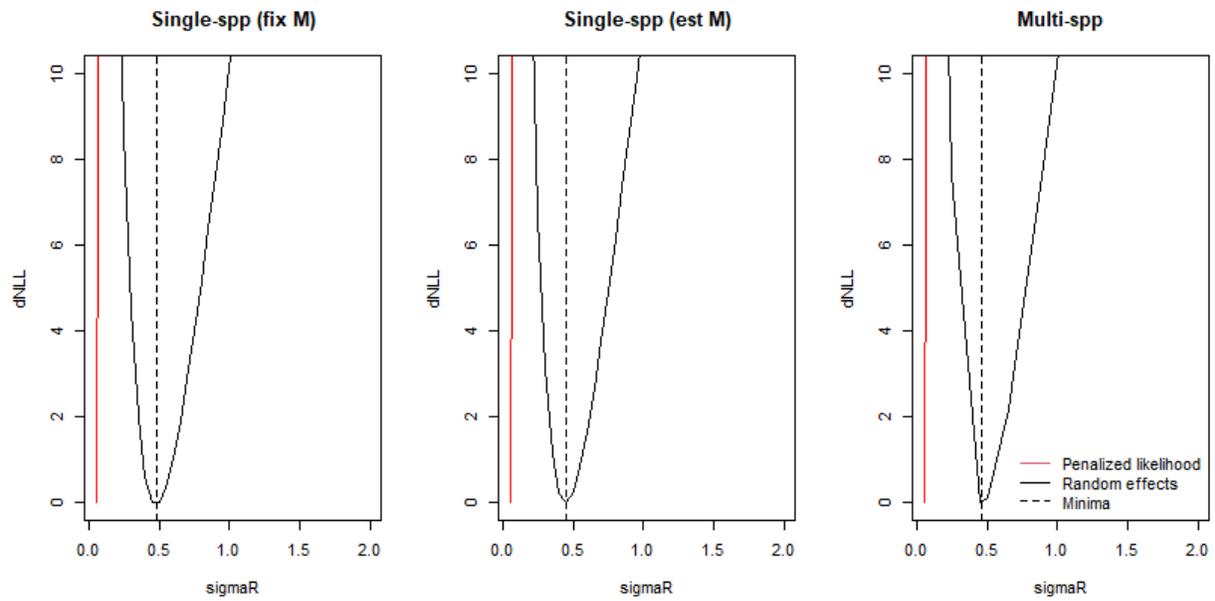


Figure 7A.11. Likelihood profiles for variance in recruitment (σ_R) across CEATTLE single- and multi-species models where recruitment deviates are treated as penalized deviates (similar to ADMB) or as random effects.



Supplementary Material

Model bridging

A series of models were developed to bridge the ADMB model to CEATTLE and evaluate the impacts of the previous four differences on model outputs:

- *Base model*: a base 2023 ADMB model. This model uses the 2021 SAFE assessment parametrization and extends the model and catch series to 2023, rather than use a separate projection module.
- *Bridging model 1*: The 2023 ADMB model with the multinomial likelihood correctly specified by setting the fishery length sample size for males to the same value as females and where the non-parametric selectivity penalties are set the same for males and females (40 and 100). Note, setting the penalties to 10 and 200 (female specification) for both sexes did not result in a converged model.
- *Bridging model 2*: CEATTLE where the maximum likelihood parameter estimates from *bridging model 1* are used as fixed inputs. Structurally, this model is exactly the same as *bridging model 1* but likelihood components are fully specified and include the log-normal bias correction.
- *Bridging model 3*: CEATTLE where the model parameters are estimated using TMB.
- *Bridging model 4*: ADMB model where the maximum likelihood parameter estimates from *bridging model 3* (CEATTLE) are used as fixed inputs.
- *Bridging model 5*: Same as model 3, however, the ageing error matrix is correctly applied to the expected true age-composition.
- *Bridging model 6*: Same as model 4, however, the recruitment deviates are treated as random effects and the associated variance parameter is estimated.

Results

All single-species bridging models showed similar if not exactly similar trends in recruitment, spawning stock biomass, and biomass (Figure S1). Bridging model 1 and 2 had the same derived quantities, given that the parameters from the ADMB model were input into CEATTLE. This included total predicted catch, selectivity, and predicted survey biomass. Similarly, the likelihood component for non-parametric selectivity penalties were the same between bridging model 1 and 2 (Table S1). However, other likelihood components differed because CEATTLE uses the full probability distribution rather than simplified forms. Bridging model 2 and 3 had overlapping estimates of recruitment, spawning stock biomass, and biomass indicating that the minor changes in the log-likelihood formulations and optimization between ADMB and CEATTLE. Bridging model 3 resulted in a lower negative log-likelihood than when using the maximum likelihood parameter estimates from ADMB, suggesting that optimization is improved in TMB. Similarly, bridging model 4 (CEATTLE parameters input into ADMB) had lower negative log-likelihood components for survey index and fishery length data than bridging model 2, but due to high penalties on catch fits in ADMB, had a higher total negative log-likelihood. Correcting the specification of the ageing error matrix resulted in a higher log-likelihood, but estimates of total biomass and spawning biomass from bridging model 5 were extremely similar to bridging model 3. For bridging model 6, treating recruitment deviates as random effects resulted in a lower variance parameter ($\sigma_R = 0.48$) than the value implicitly assumed by ADMB ($\sigma_R = 0.707$). This resulted in slightly elevated estimates of biomass and spawning stock biomass in the first 10 years of the time series.

Table S1. Model comparison of the bridging models.

	<i>Base</i>	<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>	<i>5</i>	<i>6</i>
Platform	<i>ADMB</i>	<i>ADMB</i>	<i>TMB</i>	<i>TMB</i>	<i>ADMB</i>	<i>TMB</i>	<i>TMB</i>
<i>Likelihoods</i>							
Catch	4.32e-08	6.41e-07	-97.28	-97.52	0.00205	-97.52	-97.51
Fishery length composition	816.17	841.17	85.23	83.08	823.03	83.25	83.32
Survey biomass	27.86	28.16	1.84	0.82	27.19	-0.31	1.16
Survey age composition	276.91	275.98	32.67	32.93	278.16	42.80	43.36
Recruit deviations	5.04	5.03	47.57	44.41	6.06	45.51	24.30
Selectivity penalties	7.30	8.72	8.73	9.91	9.91	10.39	9.91
F regularity	21.20	21.31	21.31	19.35	19.35	19.19	20.44
Joint NLL	177.67	181.95	100.08	92.97	195.81	103.31	84.99
Marginal NLL							96.47
Number of parameters	158	158	158	158	158	158	92
<i>Estimates</i>							
<i>q</i> -trawl	1	1	1.0	1.0		1.0	1.0
Mean Recruitment (thousand)	801,987	800,286	806,206	815,324	800,286	814,450	806,846
σ_r	0.707	0.707	0.707	0.707	0.707	0.707	0.48
$F_{40\%}$							
2023 total biomass (t)	1,273,800	1,270,450	1,284,926	1,279,433	1,265,700	1,272,617	1,273,952
2023 spawning biomass (t)	709,254	710,935	710,935	713,886	713,886	712,504	718,708
$B_{100\%}$ (t)			1,684,224	1,702,520		1,700,424	1,684,452
$B_{40\%}$ (t)			673,690	681,008		680,170	673,781
$SSB_{100\%}$ (t)			984,753	995,141		993,804	984,431
$SSB_{40\%}$ (t)			393,901	398,056		397,522	393,772
$ABC_{F40\%}$ (t)							

Figure S1. Time-series of age-1 recruitment (millions; top panel), total biomass (million mt; upper middle panel), spawning stock biomass (million mt; lower middle panel), and catch (mt; bottom panel) as determined from the ADMB and bridging models.

