

# **Discussion paper for CIE review of the eastern Bering Sea walleye pollock stock assessment**

**May 2025, CIE review**

Jim Ianelli, Carey McGilliard, Sophia Wassermann

2025-05-01 00:00

# Executive Summary

This report builds on the 2024 pollock stock assessment, using the same datasets but extending the evaluation through new tools, data comparisons, and sensitivity analyses. The primary goal is to evaluate how alternative data sources, modeling approaches, and assumptions affect the perception of stock status and management advice.

The first section introduces the **R library ebswp** (see 1), which streamlines compilation and visualization of assessment results. This provides a foundation for reproducible workflows across different model scenarios.

The second section presents **comparisons of FT-NIR and traditional microscope age estimates** (see 2). FT-NIR offers efficiencies but showed lower cohort consistency in bottom-trawl data compared to traditional ageing. Model runs incorporating FT-NIRs demonstrated that treatment of ageing error (fleet-aggregated vs. fleet-specific) influences perception of spawning biomass and reference points, though management quantities remained broadly comparable.

In the next section we explore **leave-one-out survey sensitivity analyses** (see 3). Results indicated that the influence of the bottom-trawl survey (BTS) was largest. Spawning biomass and recruitment estimates changed considerably when BTS data were downweighted. Other data sources, such as ATS and AVO indices, contributed less to overall stock trends.

The next sections examine incorporation of **multi-species natural mortality (M-at-age and year) estimates and revised acoustic-trawl survey uncertainty** (see 4). CEATTLE-based M estimates led to higher mortality at many ages, resulting in higher recruitment but lower near-term spawning biomass relative to the 2024 base model. Alternative acoustic estimates (Urmy, Ressler, and De Robertis (2025)) had smaller effects on management benchmarks, but highlight areas for further refinement.

In the fifth section we evaluated some **alternative assessment platforms** (see 5), including SS3, SAM, Rceattle, and a spatial RTMB model. Results demonstrated broad consistency in stock trends but revealed platform-specific differences in selectivity, recruitment, and uncertainty. Early applications of Rceattle and spatial models suggest promising avenues for bridging across platforms and explicitly modeling spatial or sex-specific processes.

The next section we presented **bottom-trawl survey spatial distribution modeling** (see 6), using sdmTMB and tinyVAST. Model-based indices and age compositions incorporated environmental covariates such as cold-pool extent and bottom temperature, improving

robustness in years with sparse coverage. The resulting indices were consistent with design-based estimates, but offered improved diagnostics and continuity across survey years.

The seventh section details **fishery selectivity parameterization** (see 7). A formal derivation and implementation of time-varying selectivity highlights how plateau assumptions, normalization constraints, and penalized deviations are used to ensure interpretable and stable selectivity estimates. This clarifies the role of selectivity in driving historical and current exploitation patterns.

Overall, the analyses underscore three key points:

1. *Emerging methods:* FT-NIR ageing, multi-species mortality, and model-based survey products hold promise but require careful validation.
2. *Data dependence:* The BTS survey remains the critical data source, with model outcomes highly sensitive to its inclusion.
3. *Platform diversification:* Alternative modeling frameworks provide useful cross-checks and avenues for future development, enhancing confidence in assessment results.

# 1 R Library for assessment compilations (ebswp)

The **ebswp** library is a collection of functions for compiling and analyzing results from the stock assessment models for the eastern Bering Sea walleye pollock (Ianelli (2025)). The library is designed to work with the ADMB code and depends on a consistent format for the report files generated. The link in the reference provides access to some notes on applications as vignettes to the package (e.g., . [here](#)).

## 2 Comparing FT-NIR compositions with traditional microscope age estimations

The fish age and growth lab at the AFSC have endeavored to expand on new technologies (Helser et al. (2019)).

### 2.1 FT-NIRS data consistency comparisons

Examining the abundance-at-age (for design-based estimates of the bottom-trawl age compositions) we show that in general, the TMA estimates show a high level of consistency by cohort (Figure 1). For these same data, when applied to the FTNIRs age-estimation method, the consistency was lower (Figure 2). The extent that this is a concern is unknown.

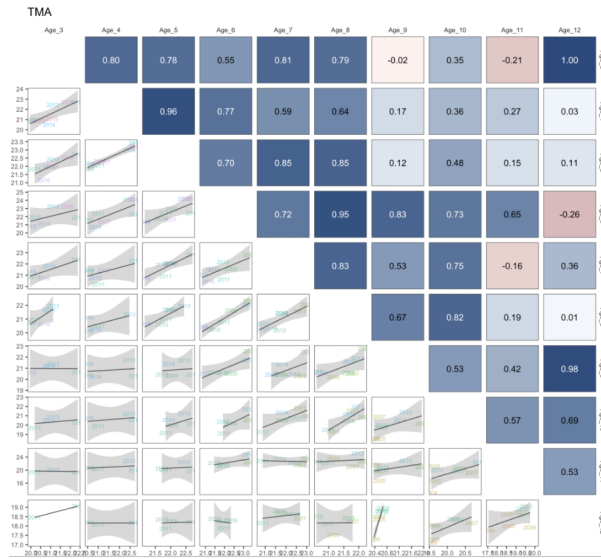


Figure 1: Cohort consistency over different ages based on log-abundance from the bottom-trawl survey data for pollock in recent years for traditional microscope age estimates. The cohorts are shown in the labels.

When preparing the data for the assessment, we plotted out the proportions-at-age for the TMA data and noted the differences in the estimates using the FTNIRs age estimates (Figure 3,

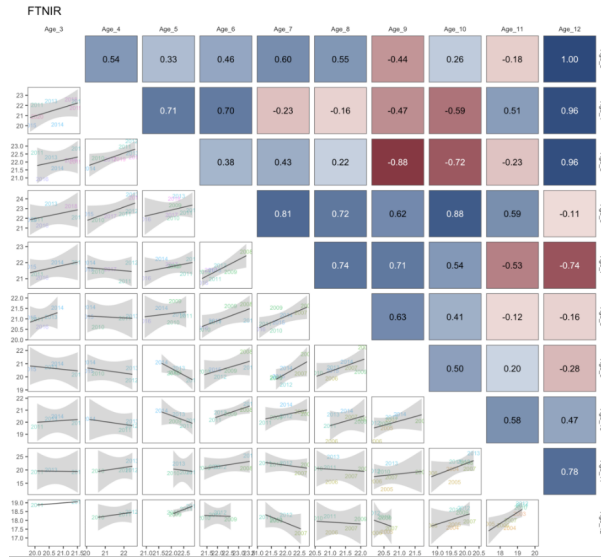


Figure 2: Cohort consistency over different ages based on log-abundance from the bottom-trawl survey data for pollock in recent years for FTNIRs age estimates. The cohorts are shown in the labels.

Figure 4, and Figure 5). Other preparations involved estimating the age-determination errors for each method following that of Punt et al. (2008). For the FTNIRs data, we optionally estimated separate age-estimation error matrices by gear type. The appearance of the fleet-aggregated and fleet-specific input data files is shown in Figure 6.

For the assessment model, the alternatives are shown based on the coverage of data and availability among types of ageing is shown in Figure 7. These figures are intended to show what data are available for the model testing and the different ways of treating age determination error matrices.

## 2.2 Comparisons of fit to FT-NIRS models

The layout of model runs comparing the traditional microscope ages (TMA) with the FT-NIRS ages is shown in Table 2.1. The models are run with the FT-NIRS data with different age-error matrix specifications. The model fits to the data are shown in Table 2.2. The table shows the goodness-of-fit

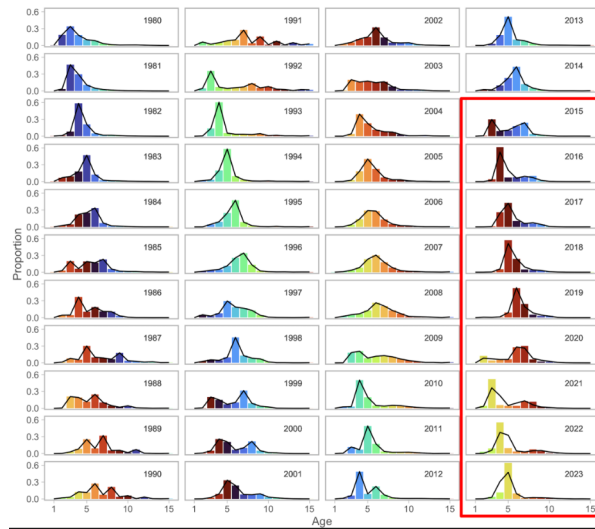


Figure 3: Proportions at age for recent years for the fishery data where the columns represent the traditional microscope age estimates (TMA) compared to the FTNIRs age estimates (lines for the red-outlined panels).

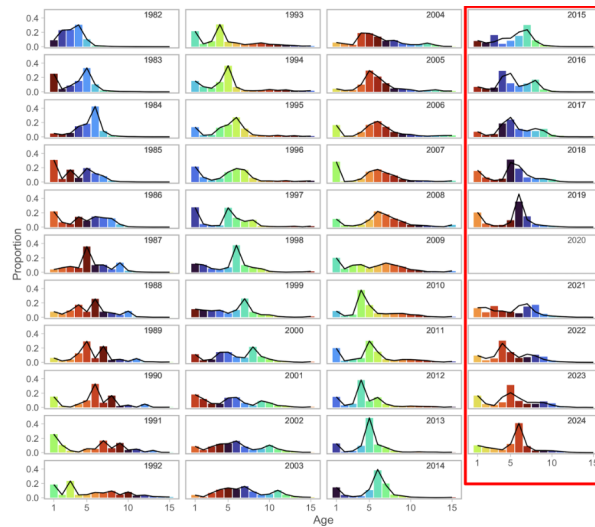


Figure 4: Proportions at age for recent years for the bottom-trawl survey data where the columns represent the traditional microscope age estimates (TMA) compared to the FTNIRs age estimates (lines for the red-outlined panels).

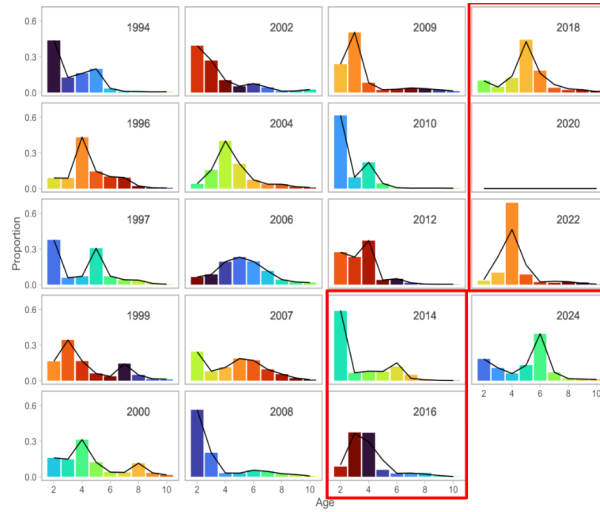


Figure 5: Proportions at age for recent years for the acoustic-trawl survey data where the columns represent the traditional microscope age estimates (TMA) compared to the FTNIRs age estimates (lines for the red-outlined panels).

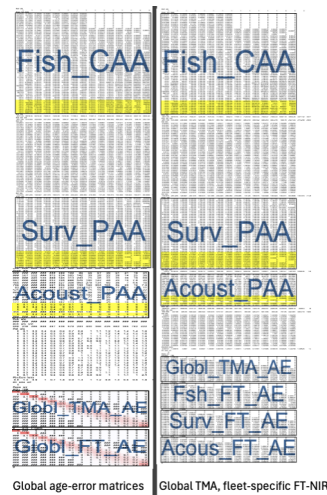


Figure 6: Schematic of input datafile showing the breakout of the fleet-specific FTNIRs age-error matrices (right panel) compared to that where the global FTNIRs age-error matrices are used (left panel).

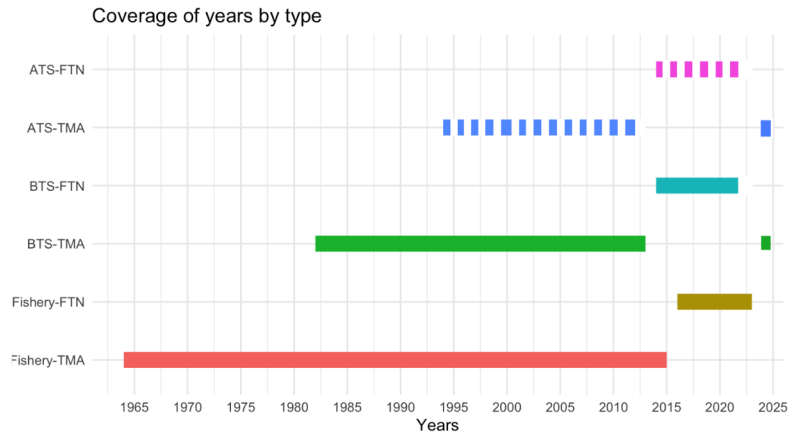


Figure 7: Data extent by age composition for the models with FTNIRs age estimates. The data types are for the Fishery, bottom trawl survey (BTS), and acoustic trawl survey (ATS) and the years of coverage are shown in the x-axis.

Table 2.1: Configuration of different model options testing the impact of different treatment of the age composition data and ageing error.

Name	Model description	
	Age type	Age-error
Base	TMA	Global TMA
FT-NIR 1	TMA + FT-NIR	Global TMA, global FT-NIR
FT-NIR 2	TMA + FT-NIR	Global TMA, disaggregated FT-NIR

Table 2.2: Goodness-of-fit measures to primary data for different data and age-error applications. RMSE=root-mean square log errors, NLL=negative log-likelihood, SDNR=standard deviation of normalized residuals, Eff. N=effective sample size for composition data). See text for incremental model descriptions.

Component	Base	FT-NIR fleets aggregated	FT-NIR fleet-specific
RMSE BTS	0.150	0.151	0.152
RMSE ATS	0.167	0.163	0.158
RMSE AVO	0.216	0.217	0.215
RMSE CPUE	0.097	0.097	0.097
SDNR BTS	0.890	0.900	0.880
SDNR ATS	0.930	0.910	0.880
SDNR AVO	0.950	0.960	0.940
Eff. N Fishery	358	336	381
Eff. N BTS	159	148	153
Eff. N ATS	233	221	233
Catch NLL	3	4	3
BTS NLL	24	24	23
ATS NLL	8	7	7
ATS Age1 NLL	7	8	9
AVO NLL	8	8	8
Fish Age NLL	417	559	614
BTS Age NLL	286	330	313
ATS Age NLL	43	42	41
NLL selectivity	213	216	215
NLL Priors	20	20	20
Data NLL	807	995	1,032
Total NLL	1110.900	1298.930	1332.810

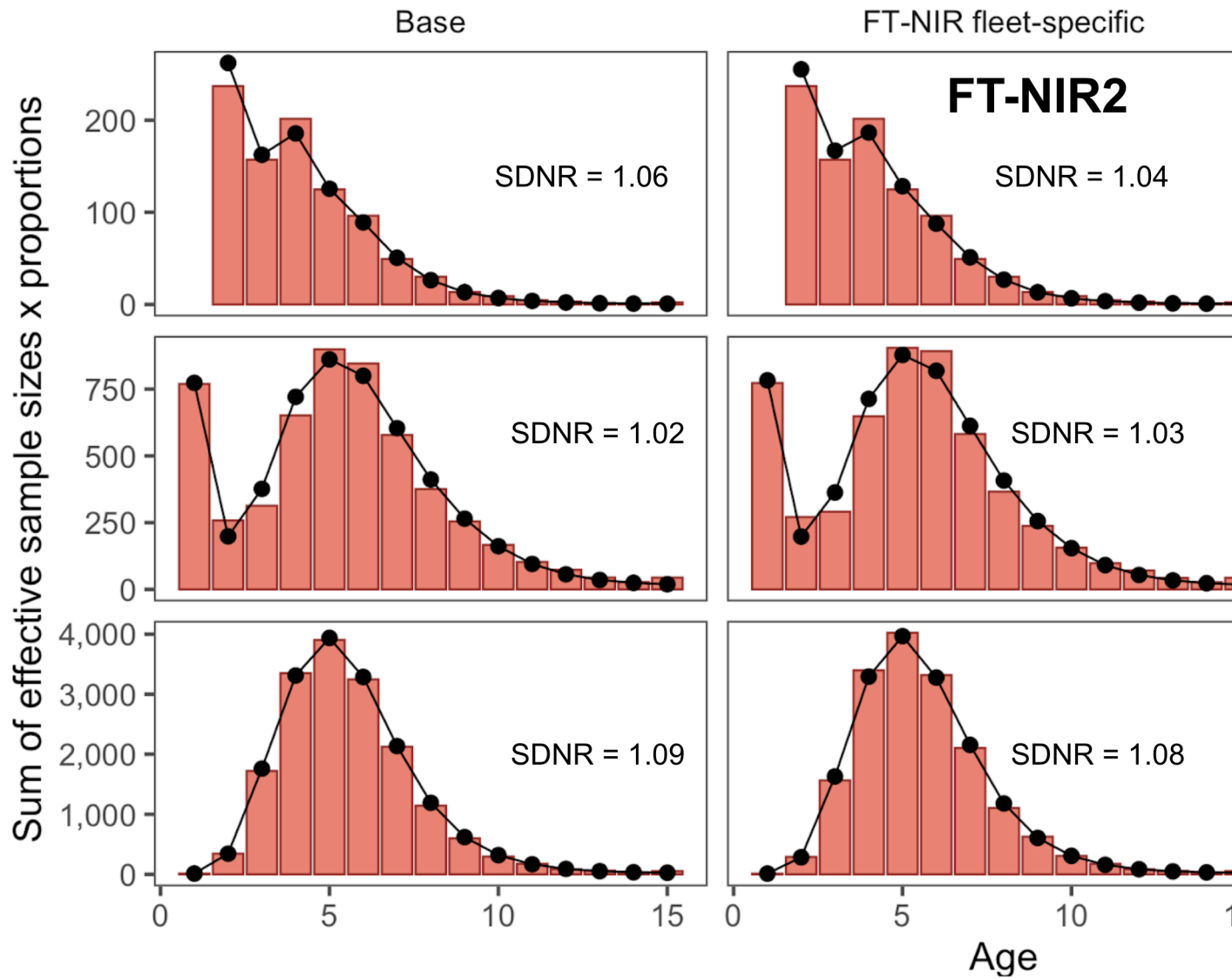


Figure 8: One-step ahead residuals for different model specifications for the fishery composition data

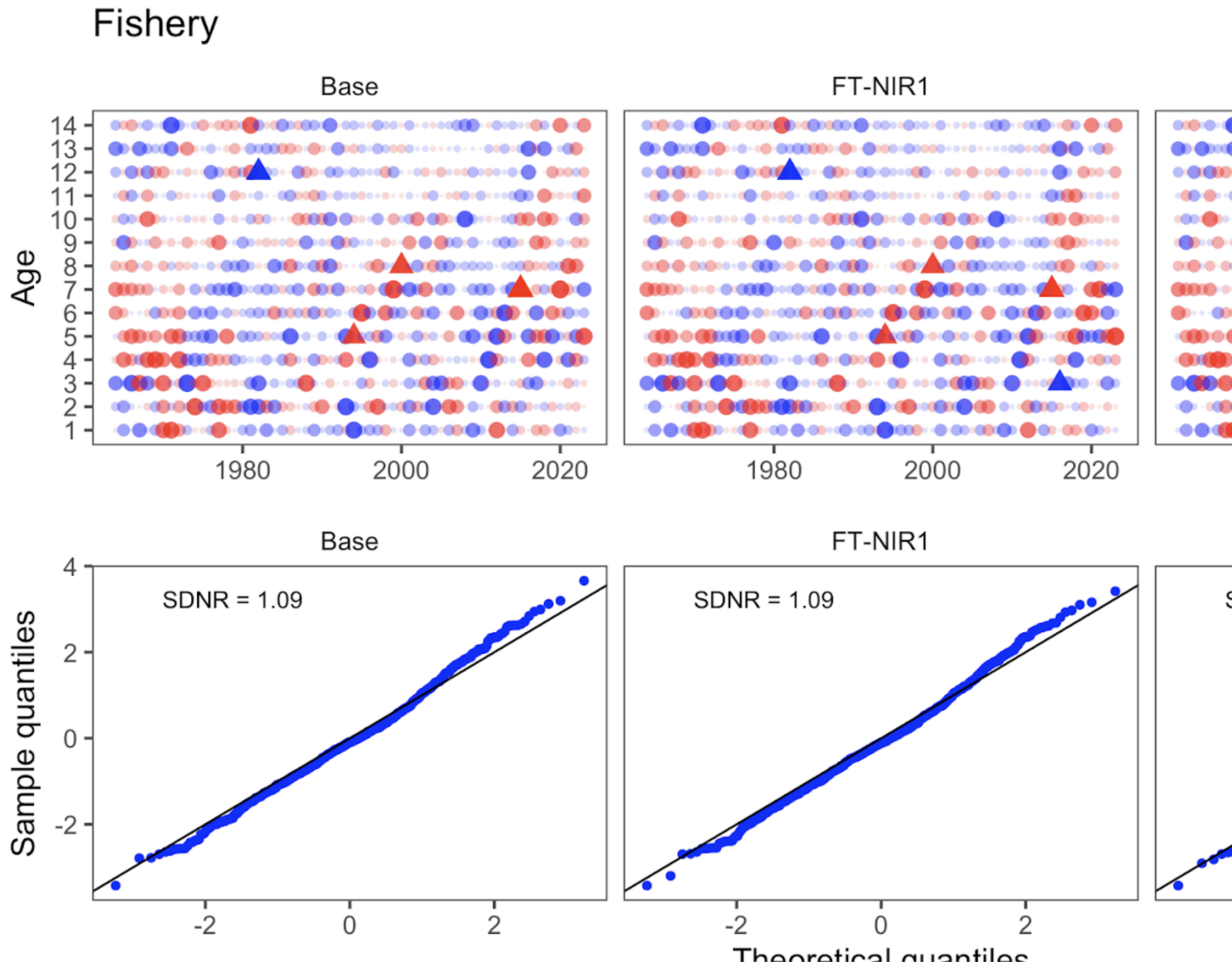


Figure 9: One-step ahead residuals for different model specifications for the fishery composition data

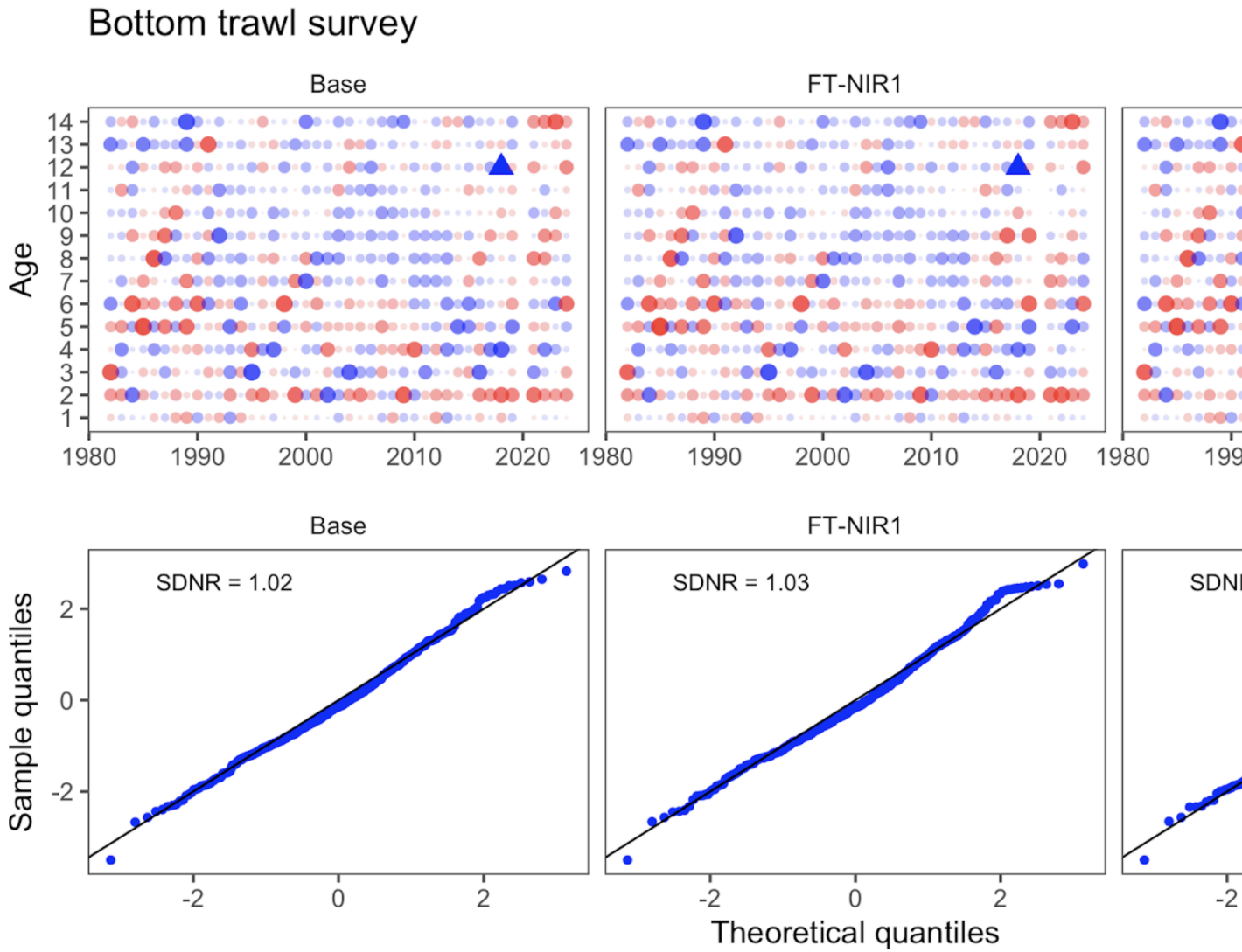


Figure 10: One-step ahead residuals for different model specifications for the bottom-trawl survey composition data

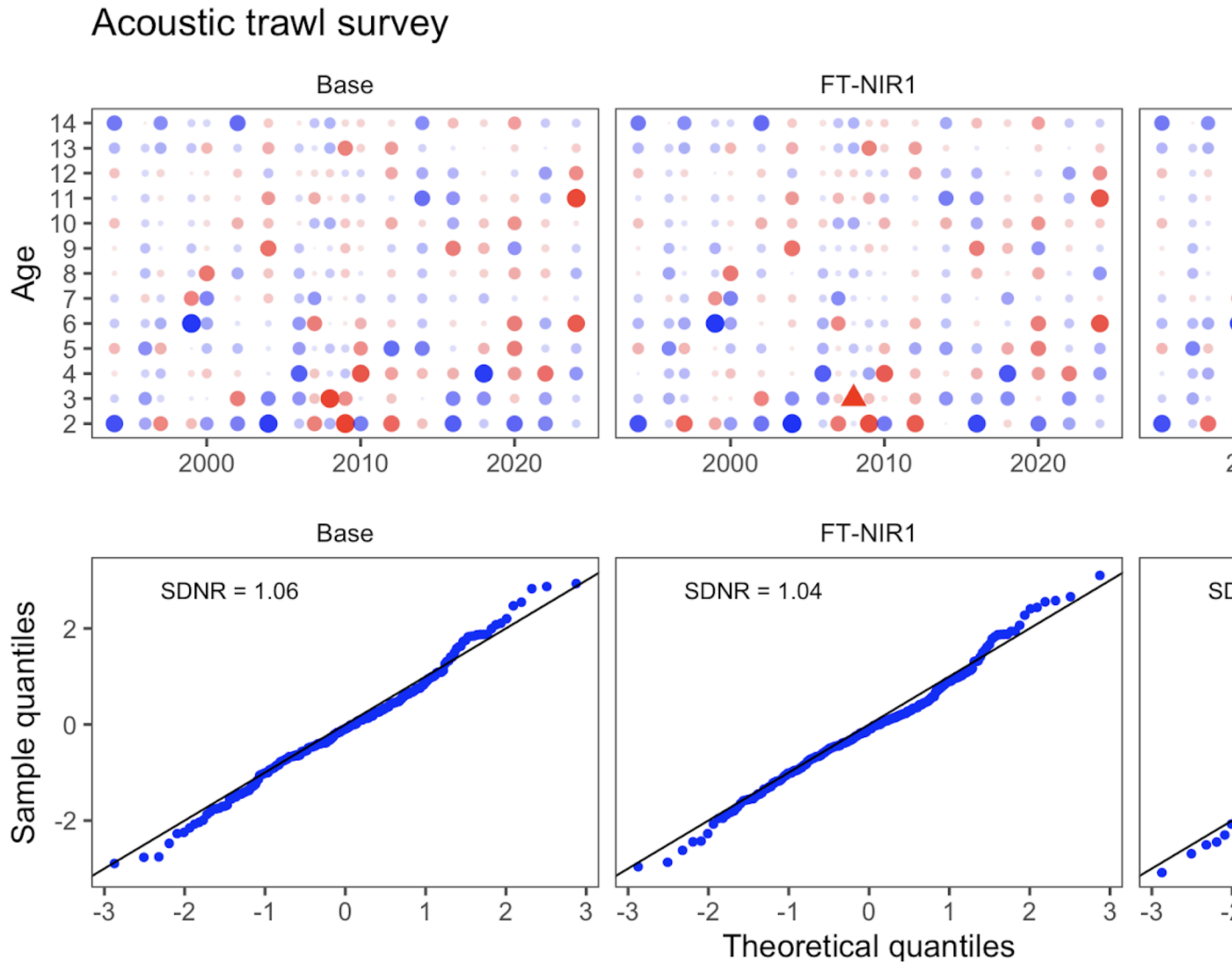


Figure 11: One-step ahead residuals for different model specifications for the acoustic-trawl survey composition data

Table 2.3: Estimated derived quantities and selected corresponding CVs for stock status and reference point estimates to primary data (base) and for different FT-NIR data treatments for age-error estimates. See text for incremental model descriptions.

Component	Base	FT-NIR fleets aggregated	FT-NIR fleet-specific
$B_{2025}$	2,800	2,500	2,700
$CV_{B_{2025}}$	0.14	0.14	0.14
$B_{MSY}$	1,667	1,652	1,670
$CV_{B_{MSY}}$	0.23	0.24	0.23
$B_{2025}/B_{MSY}$	165%	153%	162%
$B_0$	4,318	4,360	4,436
$B_{35\%}$	1,935	1,898	1,931
SPR rate at $F_{MSY}$	31%	30%	29%
Steepness	0.63	0.64	0.64
Est. $B_{2024}/B_{2024, no fishing}$	0.65	0.62	0.64
$B_{2024}/B_{MSY}$	182%	169%	178%

### 3 Leave-one-out survey sensitivity analysis

Here we present model runs that downweight one or more sources of survey data at a time. Table 3.1 shows the combinations of data types included for this analysis, which includes six model runs: (1) downweight all Acoustic Trawl Survey (ATS) data (age 1 survey index, age 2+ survey biomass index, and age composition data), (2) downweight the ATS age composition data, (3) downweight the Acoustic Vessels of Opportunity (AVO) data (survey biomass index), (4) downweight all Eastern Bering Sea Trawl Survey (BTS) data (age 1+ survey biomass index and age composition data), (5) downweight the BTS age composition data, and (6) downweight all survey biomass indices (BTS, ATS, and AVO survey biomass indices). The ATS age 1 data are included as a separate survey index in the assessment model, while BTS age 1 data are included with the age composition data in the multinomial objective function component.

Among these models, downweighting the BTS data (both solely downweighting the BTS age composition data and downweighting all BTS data) led to changes in trends over time in estimated spawning biomass Figure 12 . The model estimates of recruitment were variable across sensitivity runs, but estimates differed the most from the 2024 accepted model when all BTS data were downweighted Figure 13. Estimates of BTS survey biomass fit the BTS data very poorly in most years when the BTS data were downweighted (Figure 14). Additionally, the fits to the BTS survey biomass were affected by downweighting the BTS age composition data (without downweighting the BTS survey index). The model was relatively insensitive to other data being excluded, and fits to the ATS and AVO indices were relatively insensitive to downweighting the BTS data (see Figure 14, Figure 15, and Figure 16). The root mean square error (RMSE) in Table 3.2 for the model run that downweighted all BTS data was large and the and effective sample size (Eff. N) was low as compared to the 2024 accepted model.

Table 3.1: Configuration of different model options testing the impact of different data components

Dataset	Model description					
	Drop all BTS	Drop all ATS	Drop all indices	Drop AVO	Drop BTS Age composition	Drop ATS Age composition
BTS age 2+ Index	X		X			
ATS age 2+ Index		X	X			
ATS age 1 Index		X	X			
AVO Index			X			
BTS comps	X				X	
ATS comps		X				

Additionally, downweighting solely the BTS age composition data led to low effective sample size for the BTS.

The importance of the survey data can be seen in some of the reference calculations where the uncertainty increases dramatically and the stock-status perception was quite different across sensitivity runs Table 3.3. Most dramatically, dropping all survey indices led to a CV in “ $B_{MSY}$ ” estimates of 1.26 as compared to 0.13 in the 2024 accepted model.

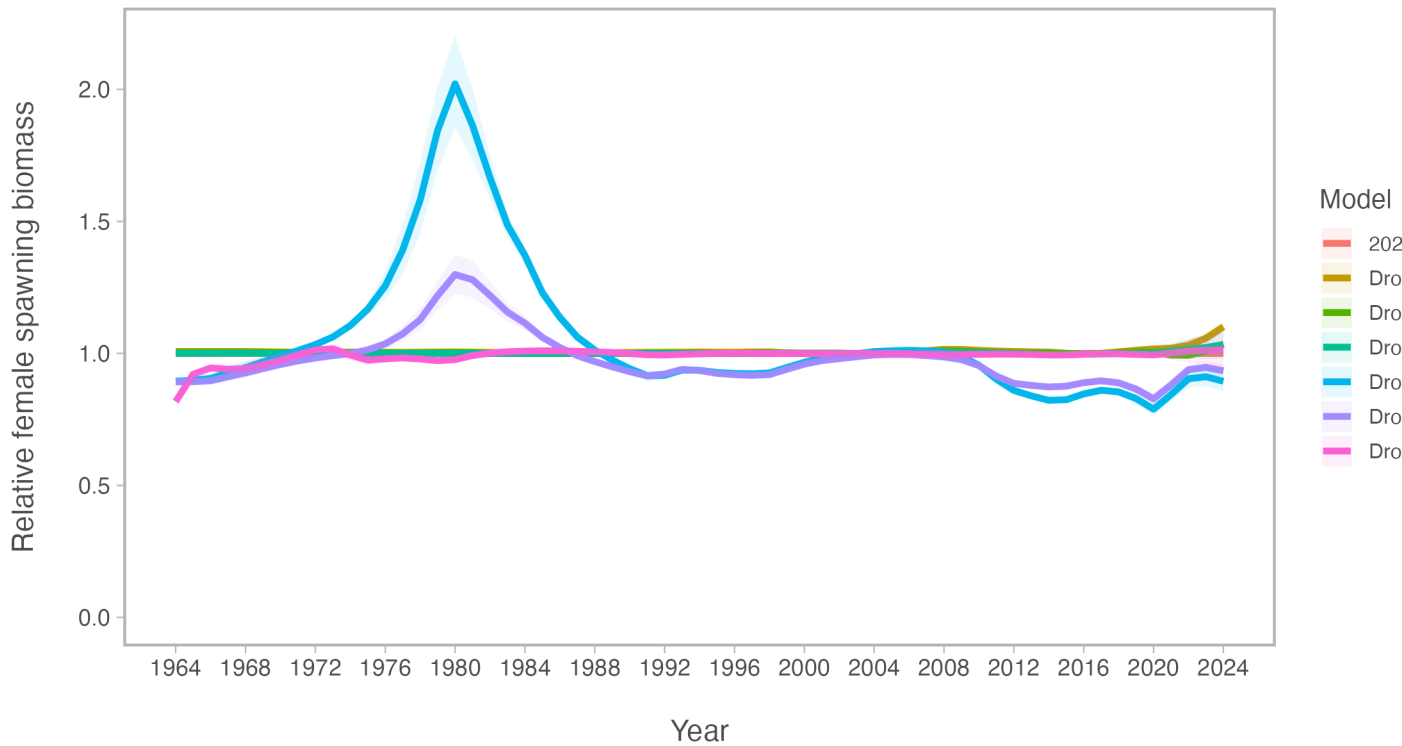


Figure 12: Comparison of spawning biomass (SSB) relative to the 2024 accepted model for models that downweight one data source at a time

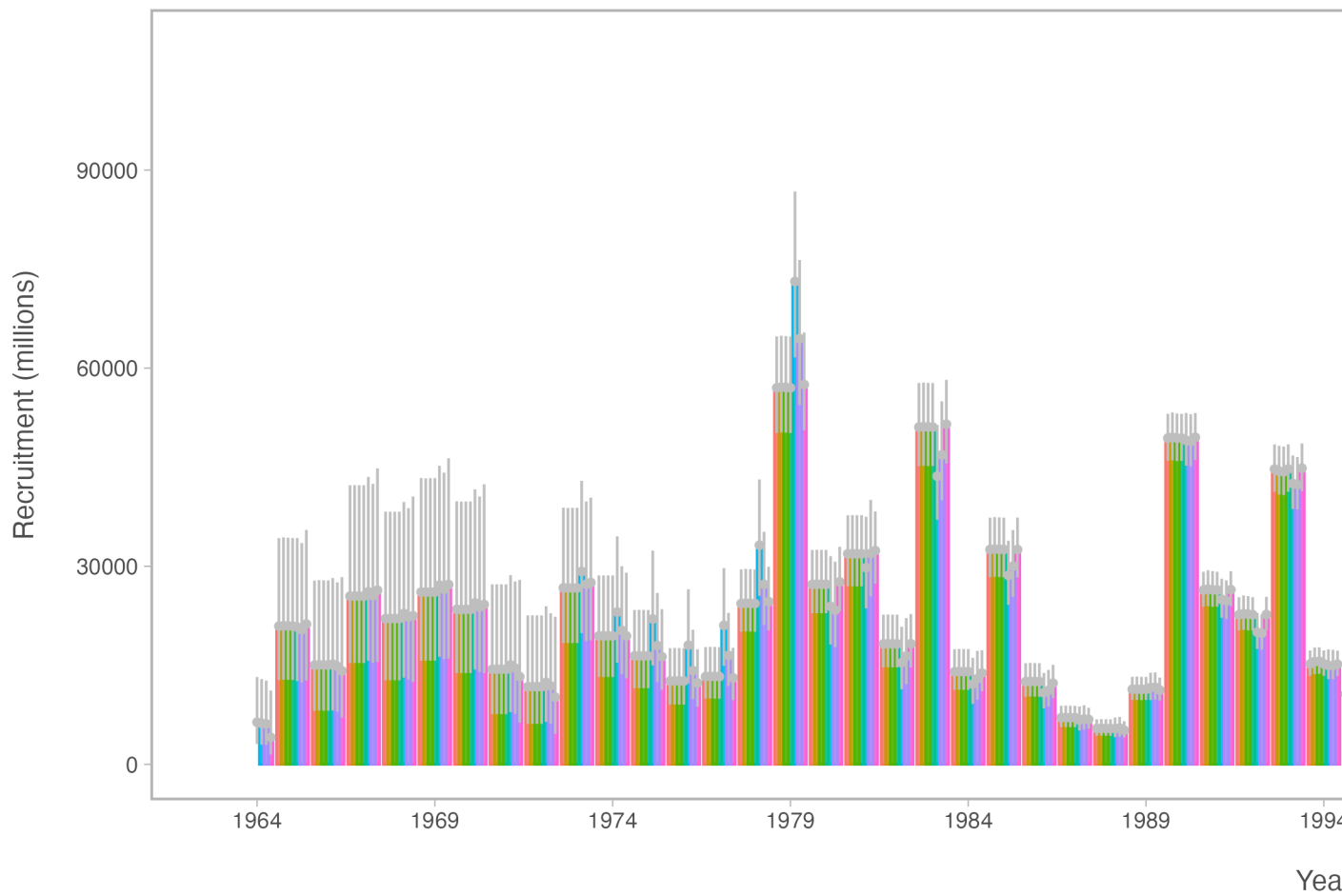


Figure 13: Comparison of recruitment estimates for models that downweight one data source at a time

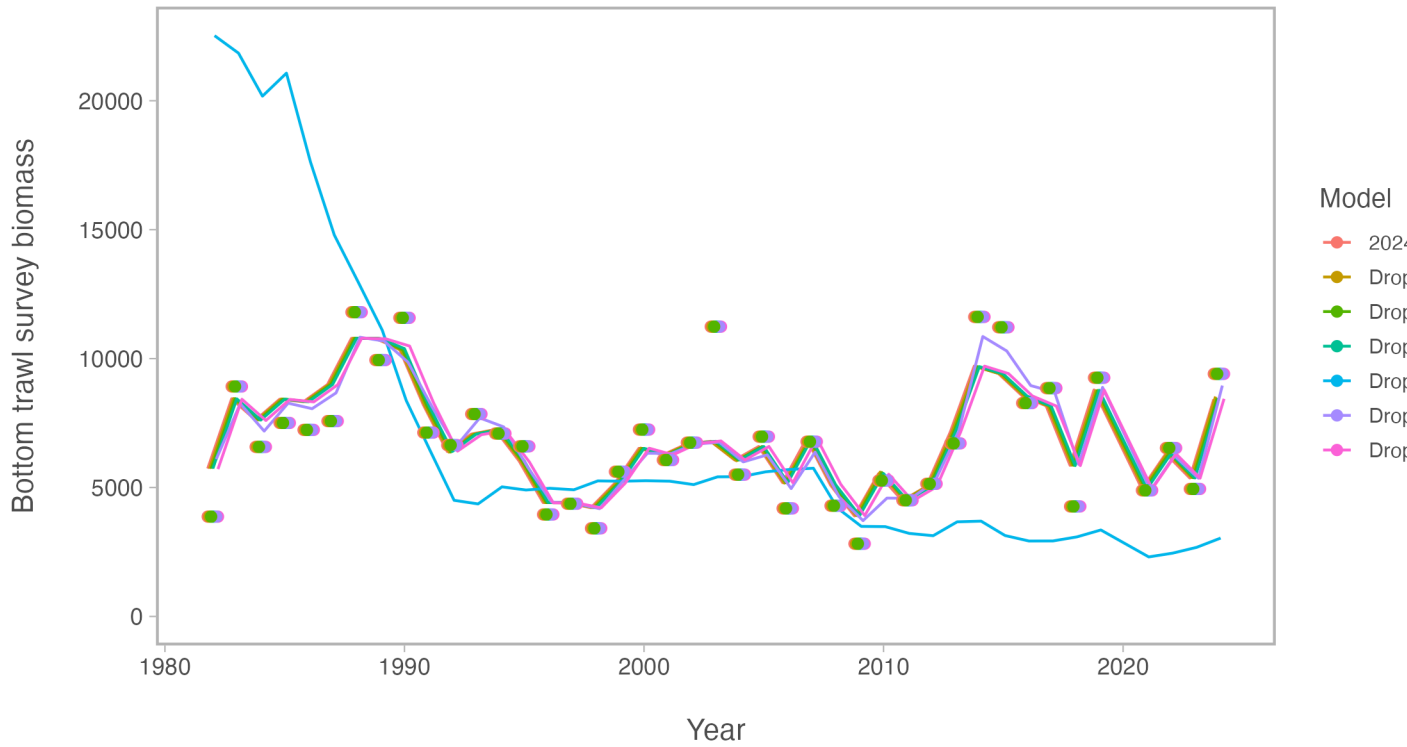


Figure 14: Comparison of fits to the Eastern Bering Sea Bottom Trawl Survey (BTS) biomass index for models that downweight one data source at a time

Table 3.2: Goodness-of-fit measures to primary data for leave-one-out model sensitivity runs. RMSE=root-mean square log errors, Eff. N=effective sample size for composition data).

Component	2024 accepted	Drop BTS	Drop ATS	Drop AVO	Drop indices	Drop BTS age
RMSE BTS	0.157	0.689	0.157	0.157	0.157	0.152
RMSE ATS	0.180	0.182	0.188	0.191	0.187	0.178
RMSE AVO	0.229	0.237	0.238	0.241	0.238	0.231
RMSE CPUE	0.093	0.090	0.093	0.093	0.098	0.092
Eff. N Fishery	1,182	1,379	1,190	1,185	1,176	1,344
Eff. N BTS	248	95	254	249	245	84
Eff. N ATS	269	230	164	269	273	228

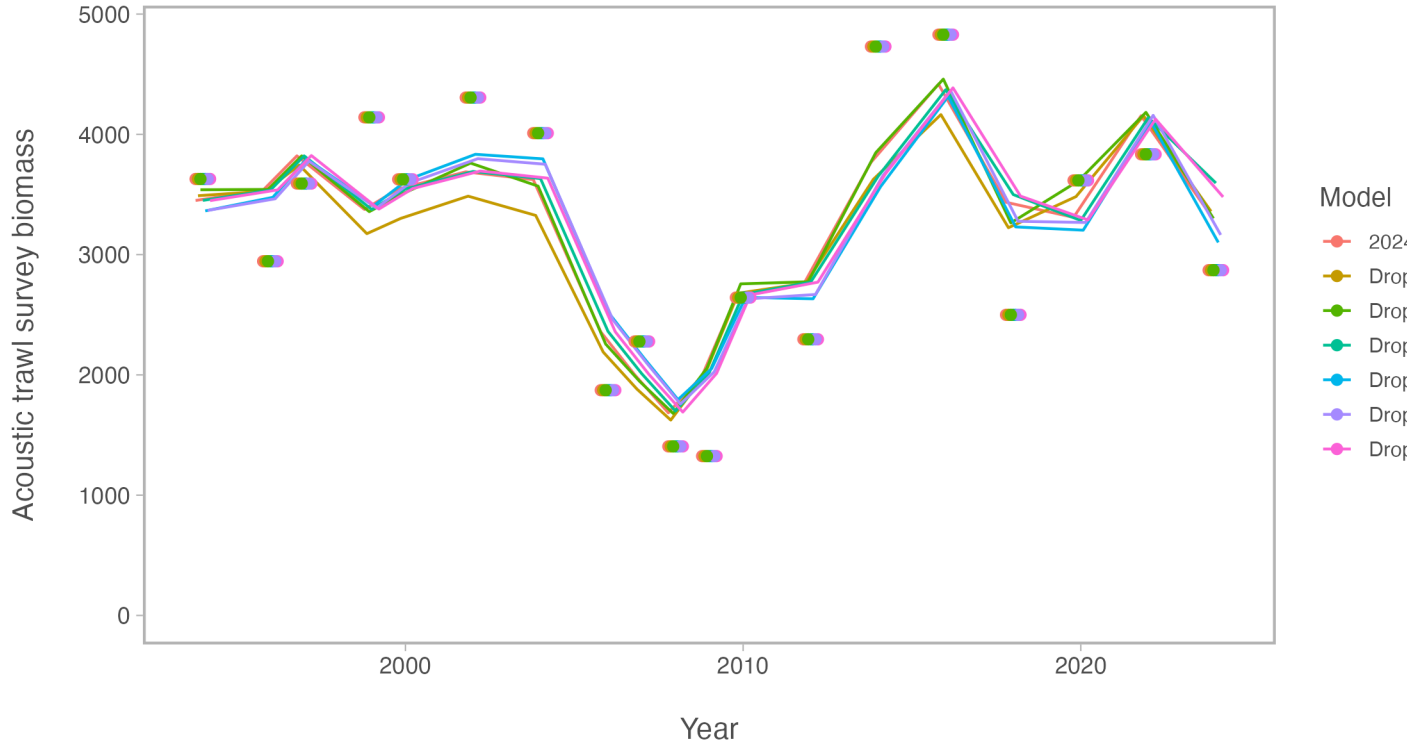


Figure 15: Comparison of fits to the Acoustic Trawl Survey (ATS) biomass index for models that downweight one data source at a time

Table 3.3: Estimated derived quantities and selected corresponding CVs for leave-one-out sensitivity analysis runs. See text for incremental model descriptions.

Estimated Quantity	2024 accepted	Drop BTS	Drop ATS	Drop AVO	Drop indices	Drop
$B_{2025}$	3,100	2,700	3,600	3,300	3,200	
$CV_{B_{2025}}$	0.13	0.16	0.14	0.16	0.18	
$B_{MSY}$	2,310	2,415	2,360	2,322	3,677	
$CV_{B_{MSY}}$	0.3	0.31	0.31	0.31	1.26	
$B_{2025}/B_{MSY}$	135%	111%	151%	141%	86%	
$B_0$	5,975	6,136	6,120	6,009	8,934	
$B_{35\%}$	2,066	1,988	2,101	2,080	2,074	
SPR rate at $F_{MSY}$	31%	35%	31%	31%	42%	
Steepness	0.62	0.59	0.62	0.62	0.52	
Est. $B_{2024}/B_{2024, no fishing}$	0.61	0.51	0.63	0.61	0.42	
$B_{2024}/B_{MSY}$	147%	126%	158%	151%	93%	

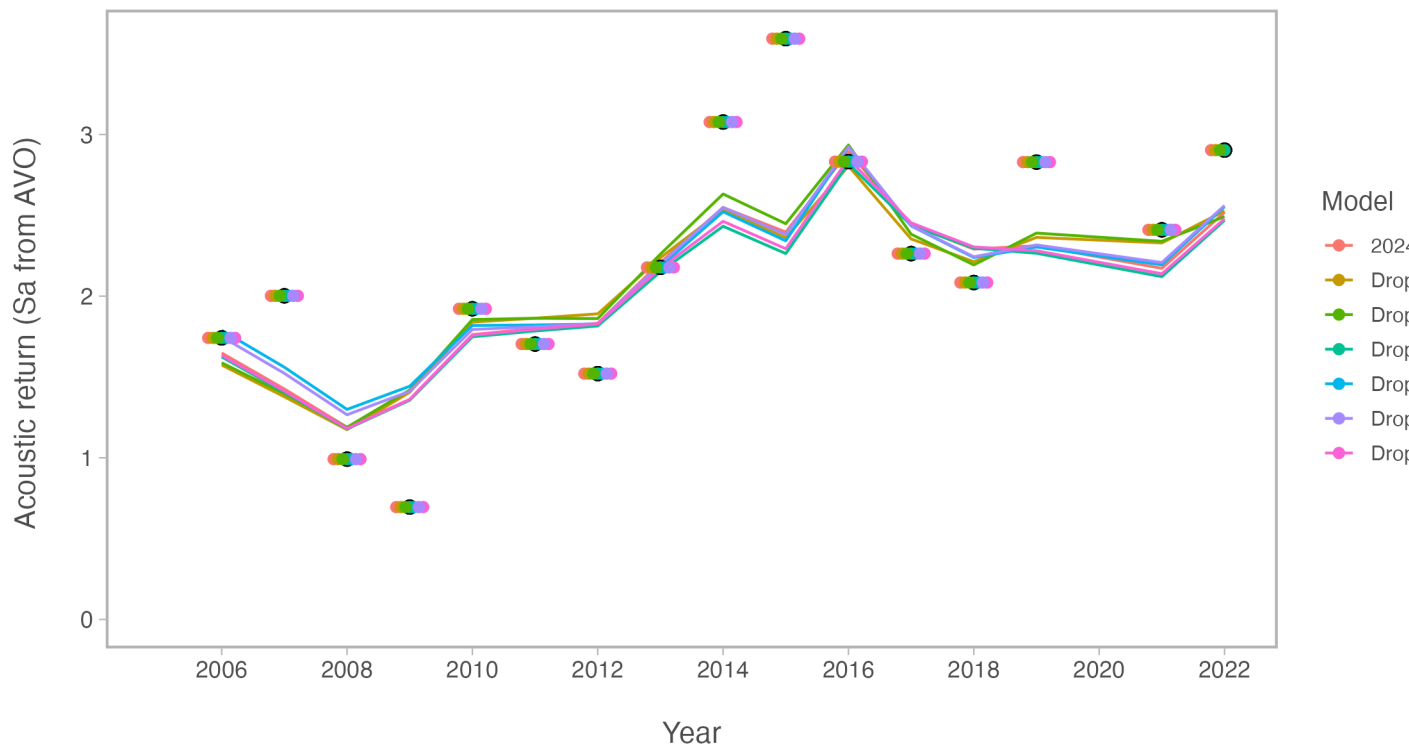


Figure 16: Comparison of fits to the Acoustic Vessels of Opportunity (AVO) biomass index for models that downweight one data source at a time

## 4 Other model evaluations for base model

### 4.1 Multi-species M-at-age and year estimates

A commonly adopted approach is to use estimated natural mortality-at-age and year from multispecies trophic interaction models (e.g., Trijoulet, Fay, and Miller (2020)). The current assessment has an option to include the 2024 CEATTLE estimates of natural mortality (Table 4.1; Figure 17). This resulted in slightly higher recruitment (Figure 18) lower spawning biomass in the near term (Figure 19). This is due to the higher natural mortality for most ages and years compared to the base 2024 model.

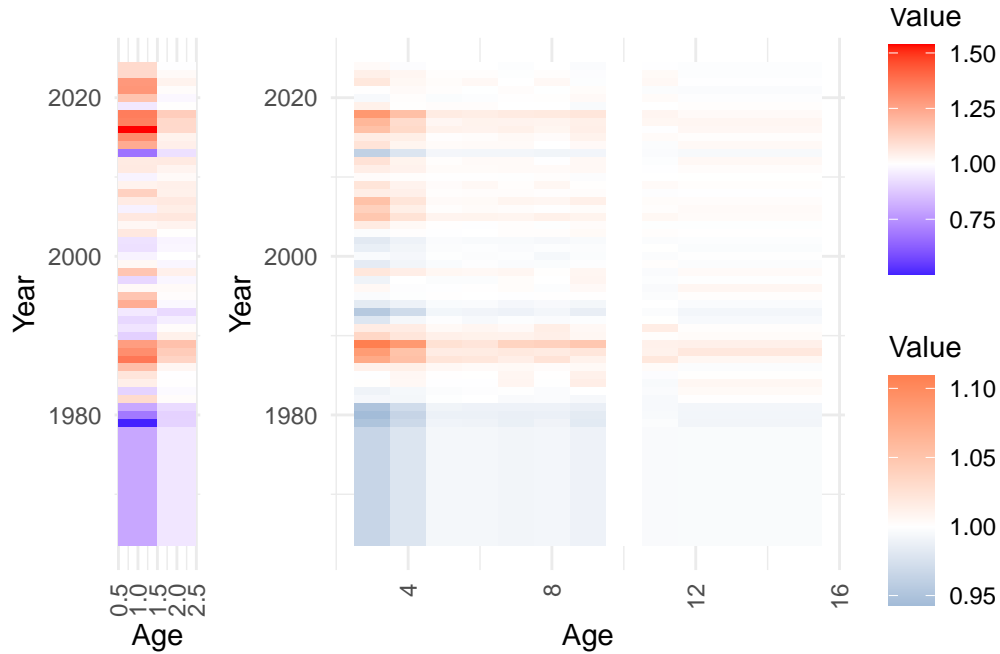


Figure 17: Depiction of the natural mortality (variability from mean value at age) as estimated from CEATTLE.

Table 4.1: Natural mortality-at-age used over ages and time based on the CEATTLE 2024 model.

	1	2	3	4	5	6	7	8	9	10
1964	1.024	0.341	0.320	0.311	0.303	0.304	0.303	0.304	0.304	0.3001
1965	1.024	0.341	0.320	0.311	0.303	0.304	0.303	0.304	0.304	0.3001
1966	1.024	0.341	0.320	0.311	0.303	0.304	0.303	0.304	0.304	0.3001
1967	1.024	0.341	0.320	0.311	0.303	0.304	0.303	0.304	0.304	0.3001
1968	1.024	0.341	0.320	0.311	0.303	0.304	0.303	0.304	0.304	0.3001
1969	1.024	0.341	0.320	0.311	0.303	0.304	0.303	0.304	0.304	0.3001
1970	1.024	0.341	0.320	0.311	0.303	0.304	0.303	0.304	0.304	0.3001
1971	1.024	0.341	0.320	0.311	0.303	0.304	0.303	0.304	0.304	0.3001
1972	1.024	0.341	0.320	0.311	0.303	0.304	0.303	0.304	0.304	0.3001
1973	1.024	0.341	0.320	0.311	0.303	0.304	0.303	0.304	0.304	0.3001
1974	1.024	0.341	0.320	0.311	0.303	0.304	0.303	0.304	0.304	0.3001
1975	1.024	0.341	0.320	0.311	0.303	0.304	0.303	0.304	0.304	0.3001
1976	1.024	0.341	0.320	0.311	0.303	0.304	0.303	0.304	0.304	0.3001
1977	1.024	0.341	0.320	0.311	0.303	0.304	0.303	0.304	0.304	0.3001
1978	1.024	0.341	0.320	0.311	0.303	0.304	0.303	0.304	0.304	0.3001
1979	0.644	0.326	0.315	0.308	0.303	0.303	0.302	0.303	0.302	0.3001
1980	0.883	0.325	0.313	0.307	0.302	0.302	0.302	0.303	0.302	0.3001
1981	1.022	0.333	0.315	0.309	0.303	0.303	0.303	0.303	0.303	0.3001
1982	1.415	0.365	0.329	0.316	0.305	0.306	0.305	0.305	0.305	0.3001
1983	1.156	0.358	0.328	0.316	0.305	0.306	0.306	0.305	0.307	0.3001
1984	1.342	0.363	0.332	0.320	0.305	0.306	0.308	0.306	0.312	0.3001
1985	1.380	0.362	0.331	0.320	0.306	0.306	0.308	0.307	0.311	0.3001
1986	1.517	0.371	0.336	0.322	0.307	0.307	0.308	0.307	0.309	0.3001
1987	1.756	0.406	0.356	0.335	0.312	0.312	0.310	0.314	0.311	0.3001
1988	1.684	0.415	0.360	0.335	0.310	0.311	0.312	0.312	0.314	0.3001
1989	1.632	0.425	0.368	0.346	0.313	0.313	0.317	0.318	0.322	0.3001
1990	1.145	0.377	0.343	0.325	0.309	0.310	0.309	0.311	0.309	0.3001
1991	1.206	0.365	0.337	0.323	0.307	0.308	0.307	0.311	0.309	0.3001
1992	1.169	0.348	0.325	0.316	0.305	0.305	0.305	0.305	0.307	0.3001
1993	1.220	0.332	0.316	0.309	0.303	0.303	0.303	0.304	0.303	0.3001
1994	1.576	0.357	0.326	0.315	0.304	0.305	0.304	0.305	0.305	0.3001
1995	1.493	0.365	0.331	0.318	0.306	0.306	0.306	0.307	0.307	0.3001
1996	1.306	0.367	0.334	0.318	0.305	0.306	0.307	0.306	0.309	0.3001
1997	1.168	0.355	0.328	0.318	0.305	0.305	0.307	0.306	0.310	0.3001
1998	1.492	0.377	0.338	0.323	0.308	0.308	0.308	0.307	0.310	0.3001
1999	1.311	0.355	0.326	0.316	0.304	0.306	0.305	0.306	0.307	0.3001
2000	1.249	0.362	0.330	0.316	0.305	0.305	0.305	0.305	0.306	0.3001
2001	1.194	0.355	0.327	0.316	0.305	0.305	0.305	0.306	0.306	0.3001
2002	1.199	0.353	0.325	0.315	0.304	0.305	0.305	0.305	0.306	0.3001
2003	1.367	0.361	0.331	0.319	0.305	0.306	0.306	0.307	0.309	0.3001 <sup>23</sup>
2004	1.312	0.374	0.337	0.320	0.306	0.306	0.306	0.306	0.308	0.3001
2005	1.366	0.387	0.348	0.326	0.308	0.309	0.309	0.310	0.310	0.3001
2006	1.243	0.380	0.345	0.323	0.306	0.307	0.307	0.307	0.308	0.3001
2007	1.355	0.385	0.349	0.325	0.307	0.307	0.309	0.309	0.311	0.3001
2008	1.450	0.377	0.336	0.322	0.306	0.307	0.307	0.307	0.308	0.3001
2009	1.328	0.378	0.339	0.321	0.306	0.308	0.306	0.308	0.307	0.3001

DPS pollack assessment discussion paper

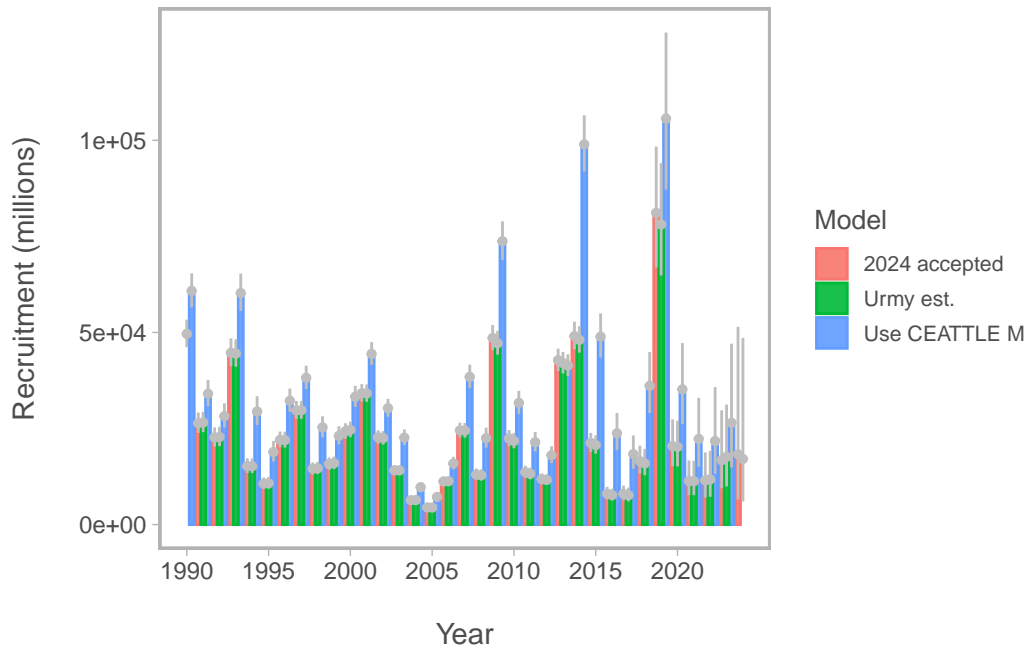


Figure 18: Comparison of recruitment estimates comparing 2024 assessment results with one where M-at-age and year is specified from CEATTLE.

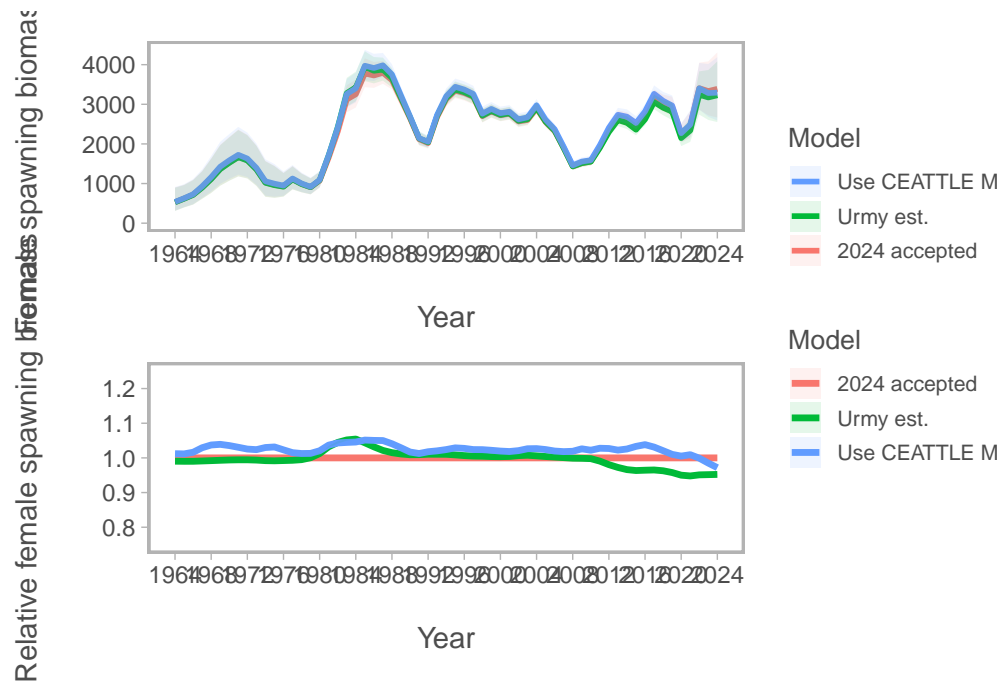


Figure 19: Comparison of spawning biomass (SSB) and SSB relative to the 2024 accepted model for model that uses the M-at-age matrix from CEATTLE.

## 4.2 Total uncertainty in acoustics

We also include a comparison of revised acoustic-trawl estimates (Urmy, Ressler, and De Robertis (2025); Figure 20). The natural mortality alternative run resulted in generally better fits to most data components whereas the “untuned” run with the Urmy, Ressler, and De Robertis (2025) estimates were generally poorer (Table 4.2). Results from the Urmy, Ressler, and De Robertis (2025) had minor impact on management-related quantities whereas those with the alternative natural mortality specification differed in a number of ways (Table 4.3). The most notable difference was the estimated  $B_{MSY}$  and  $B_{35\%}$  values.

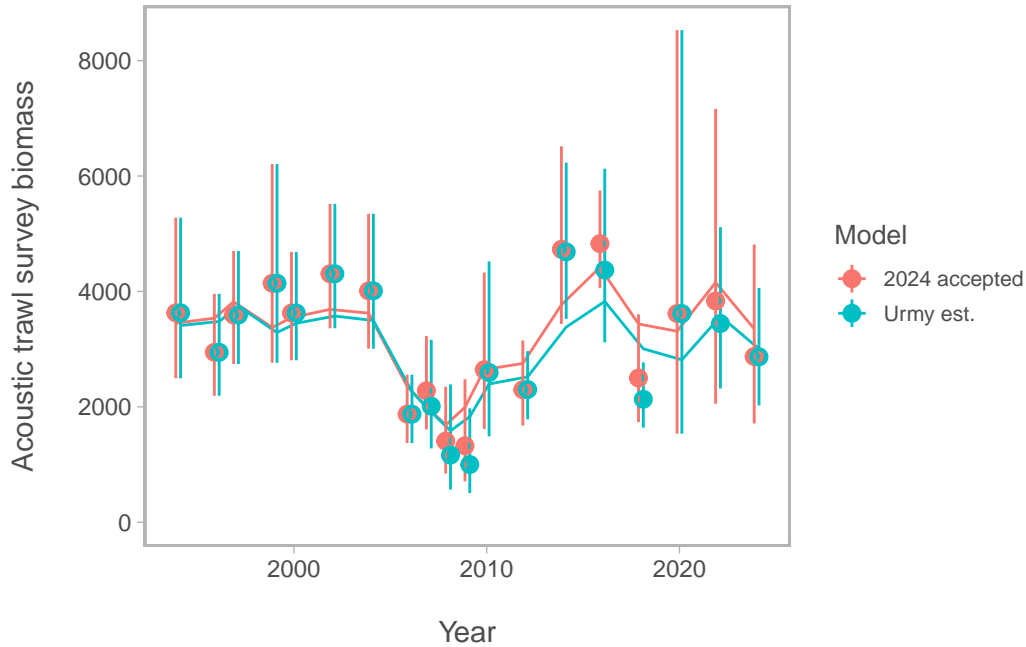


Figure 20: Comparison of fits to the Eastern Bering Sea acoustic-Trawl Survey (ATS) biomass index for the accepted 2024 model and one with revised estimates of index values and uncertainty from Urmy et al., (in prep).

Table 4.2: Goodness-of-fit measures to primary data for model sensitivity to applying the matrix of M-at-age and year and acoustic uncertainty estimates. RMSE=root-mean square log errors, SDNR=standard deviation of normalized residuals, Eff. N=effective sample size for composition data).

Component	2024 accepted	Use CEATTLE M	Urmy est.
RMSE BTS	0.157	0.157	0.161
RMSE ATS	0.180	0.179	0.226
RMSE AVO	0.229	0.230	0.236
RMSE CPUE	0.093	0.093	0.093
SDNR BTS	0.950	0.960	0.970
SDNR ATS	1.000	0.990	1.220
SDNR AVO	0.980	0.960	1.020
Eff. N Fishery	1,182	1,176	1,188
Eff. N BTS	248	241	221
Eff. N ATS	269	267	278
Catch NLL	3	3	3
BTS NLL	31	31	30
ATS NLL	9	9	13
ATS Age1 NLL	11	12	14
AVO NLL	9	8	9
Fish Age NLL	167	168	162
BTS Age NLL	194	196	205
ATS Age NLL	30	31	33
NLL selectivity	196	196	199
NLL Priors	20	20	20
Data NLL	465	468	480
Total NLL	718	713	735

Table 4.3: Estimated derived quantities and selected corresponding CVs for sensitivity to natural mortality specification and acoustic uncertainty estimates.

Estimated Quantity	2024 accepted	Use CEATTLE M	Urmy est.
$B_{2025}$	3,100	3,000	3,000
$CV_{B_{2025}}$	0.13	0.13	0.13
$B_{MSY}$	2,310	2,584	2,290
$CV_{B_{MSY}}$	0.3	0.35	0.3
$B_{2025}/B_{MSY}$	135%	116%	131%
$B_0$	5,975	6,568	5,929
$B_{35\%}$	2,066	1,844	2,057
SPR rate at $F_{MSY}$	31%	35%	31%
Steepness	0.62	0.59	0.62
Est. $B_{2024}/B_{2024, no fishing}$	0.61	0.54	0.59
$B_{2024}/B_{MSY}$	147%	127%	141%

## 5 Alternative assessment software

### 5.1 Stock-synthesis 3

Stock synthesis 3 (SS3) is a well establish framework and is common in many diverse settings. (Methot and Wetzel (2013)) For EBS pollock, we extended an initial framework based on the configurations used for Pacific hake (Grandin et al. (2024)) since there are many similarities in the type of data available. Namely, empirical weight-at-age, acoustic survey data, and indices of 1-year olds.

Within the SS3 framework, we evaluated different configurations for fishery selectivity variability (Figure 21). These showed the impact on the uncertainty and stock trends (Figure 22 for SSB and Figure 23 for recruitment).

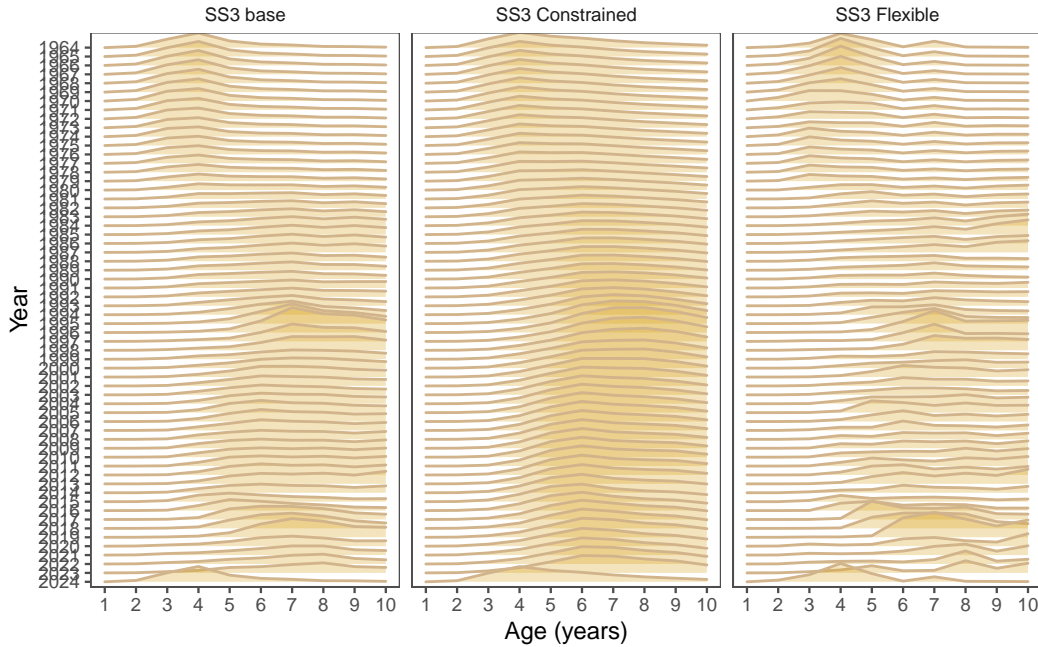


Figure 21: Model results comparing three different SS3 model configurations. The models are the base model (Base), the constrained model (Constrained), and the flexible model (Flexible).

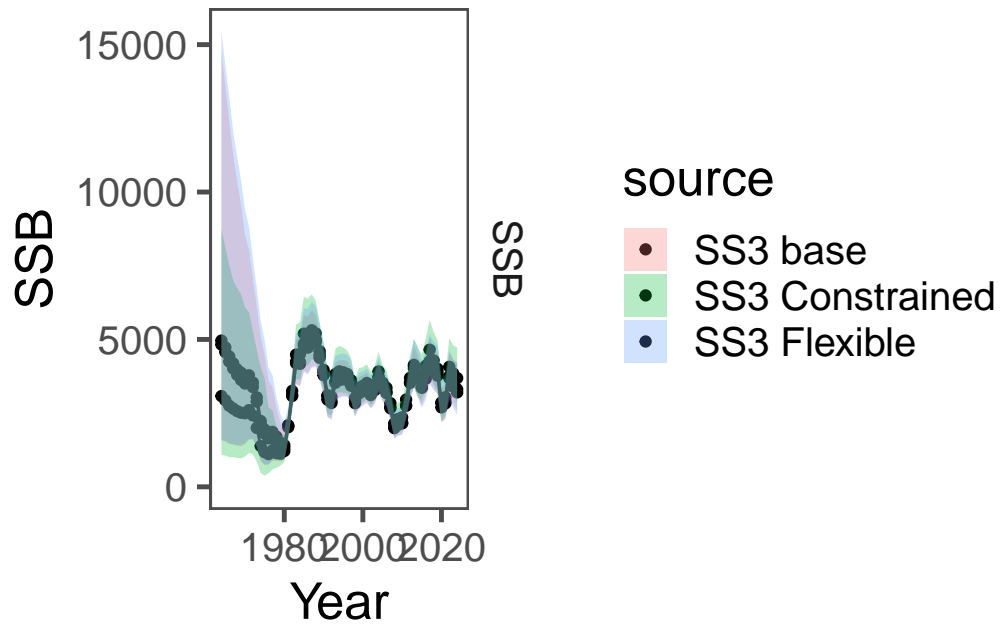


Figure 22: Model results comparing the time-series for three different SS3 model configurations. The models are the base model (Base), the constrained model (Constrained), and the flexible model (Flexible).

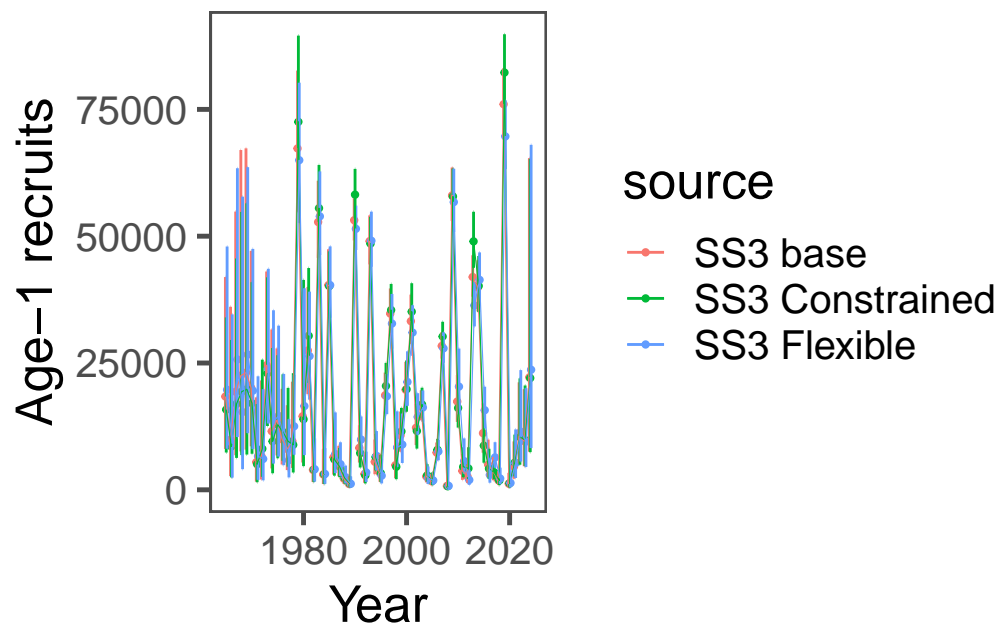


Figure 23: Model results comparing the time-series for three different SS3 model configurations. The models are the base model (Base), the constrained model (Constrained), and the flexible model (Flexible).

Comparing the “base” SS3 model with the pollock model showed different patterns in the early periods for selectivity (Figure 24) but slight differences in SSB and recruitment estimates (Figure 25 and Figure 26).

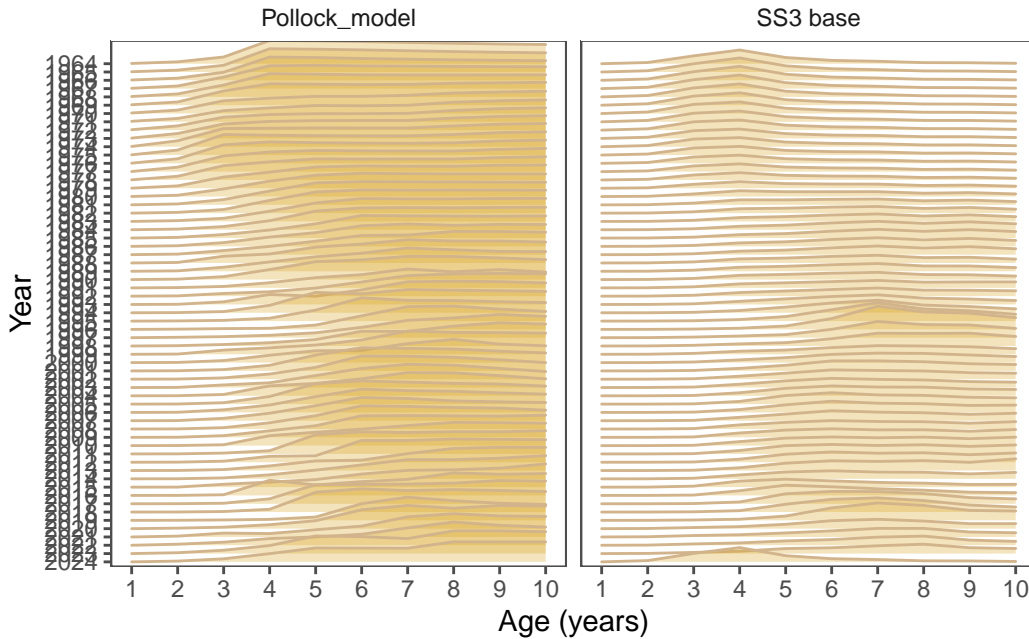


Figure 24: Model results comparing selectivity estimates between the base SS3 model configuration and the pollock model accepted in 2024.

## 5.2 SAM

The Stock Assessment Model (SAM) is a state-space model written in TMB (Nielsen and Berg (2014), Nielsen et al. (2021)). We attempted to configure the settings to have similar levels of flexibility to the reference model used for catch advice in 2024 and the settings and results can be found [here](#)). Comparing “selectivity” in the SAM model to the reference model, we see broad similarities but slightly lower inter-annual variability in the SAM run (Figure 27). The spawning biomass trend largely overlaps in the uncertainty between the SAM and pollock model (Figure 28a). The recruitment estimates are similar as well (Figure 28b).

## 5.3 Other models in development

The “Woods Hole Assessment model” (WHAM) is written in TMB and is a fully state-space model (Stock and Miller (2021)). Initial trials are available on the [github repository](#). We abandoned this effort given difficulties in some aspects of specifications and time constraints.

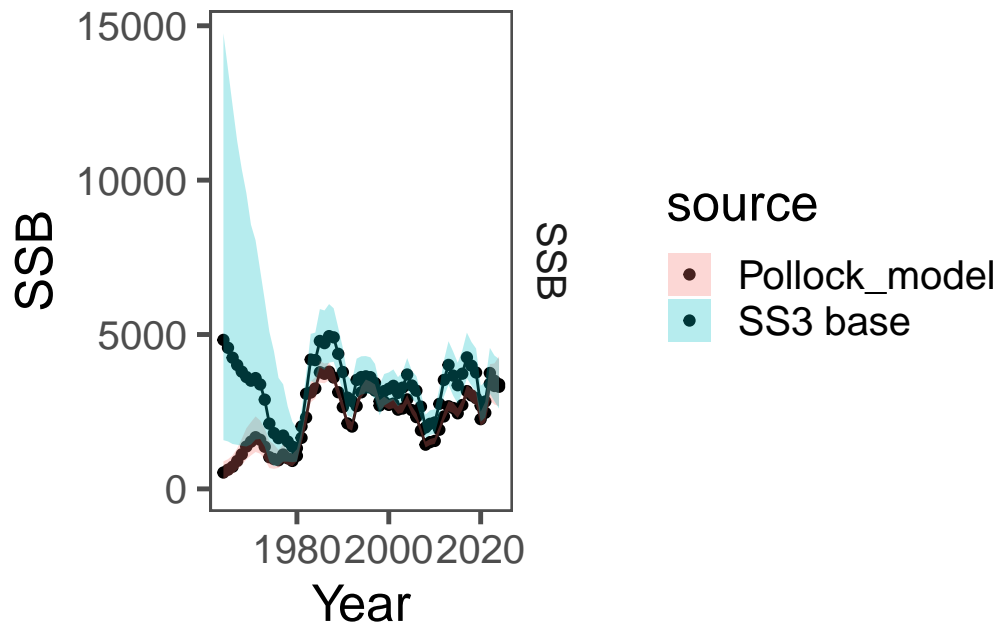


Figure 25: Model results comparing the base SS3 model configuration with the model used for pollock in 2024.

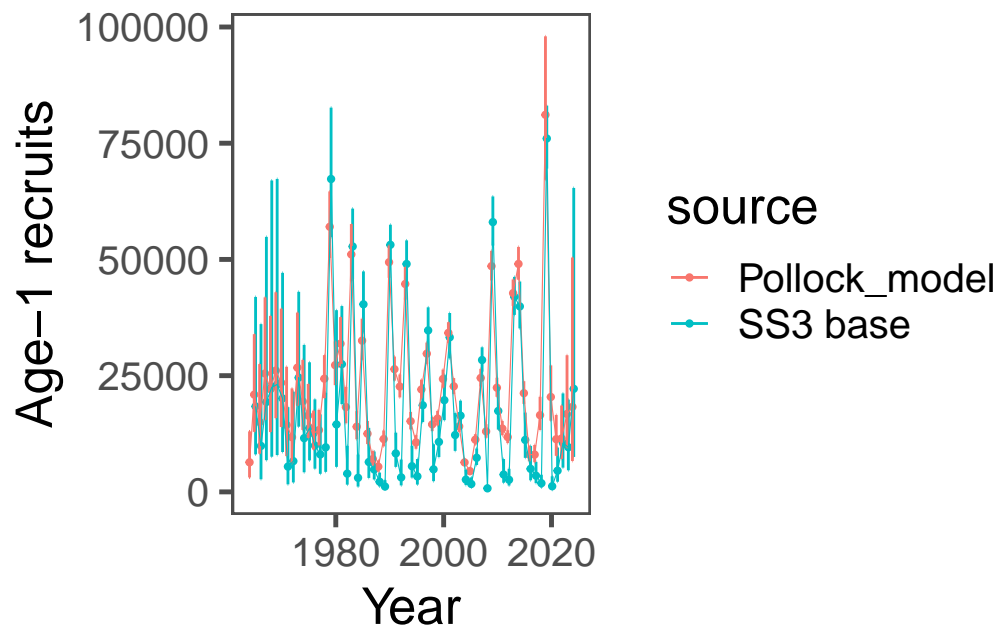


Figure 26: Model results comparing the base SS3 model configuration with the model used for pollock in 2024.

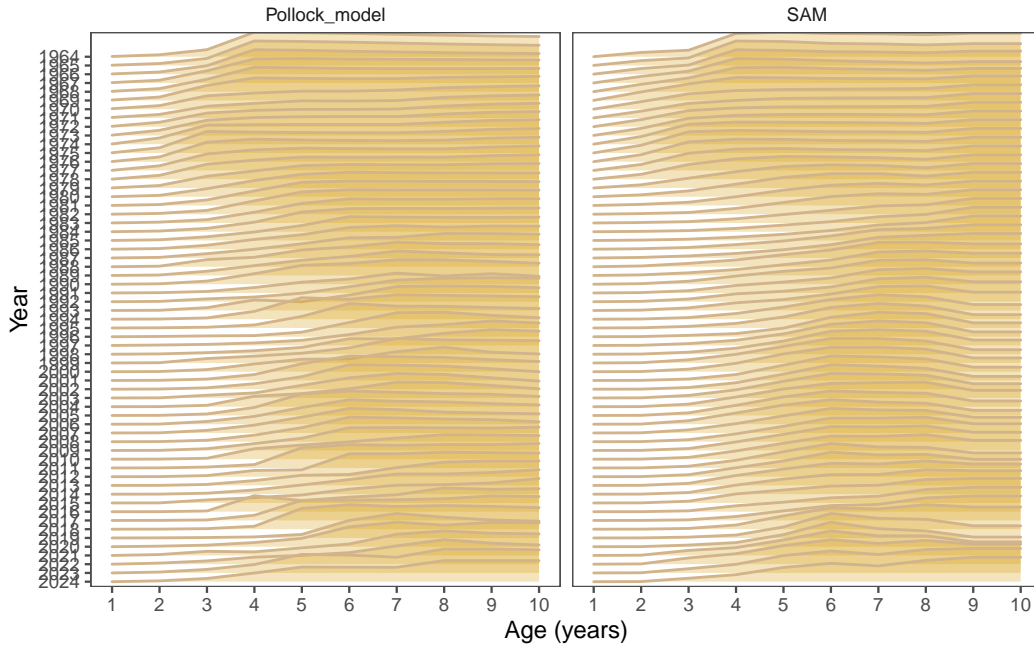
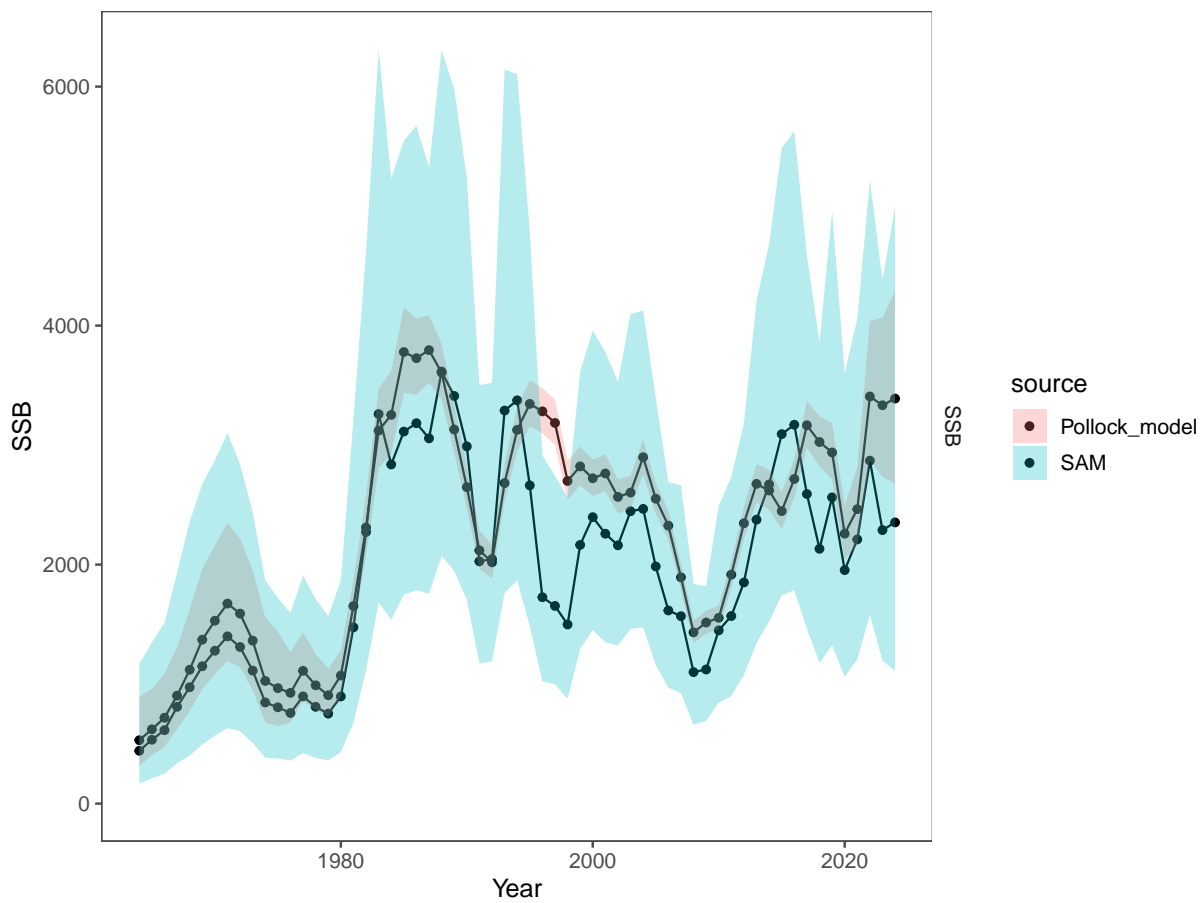


Figure 27: Model results comparing last year's model with the SAM model.

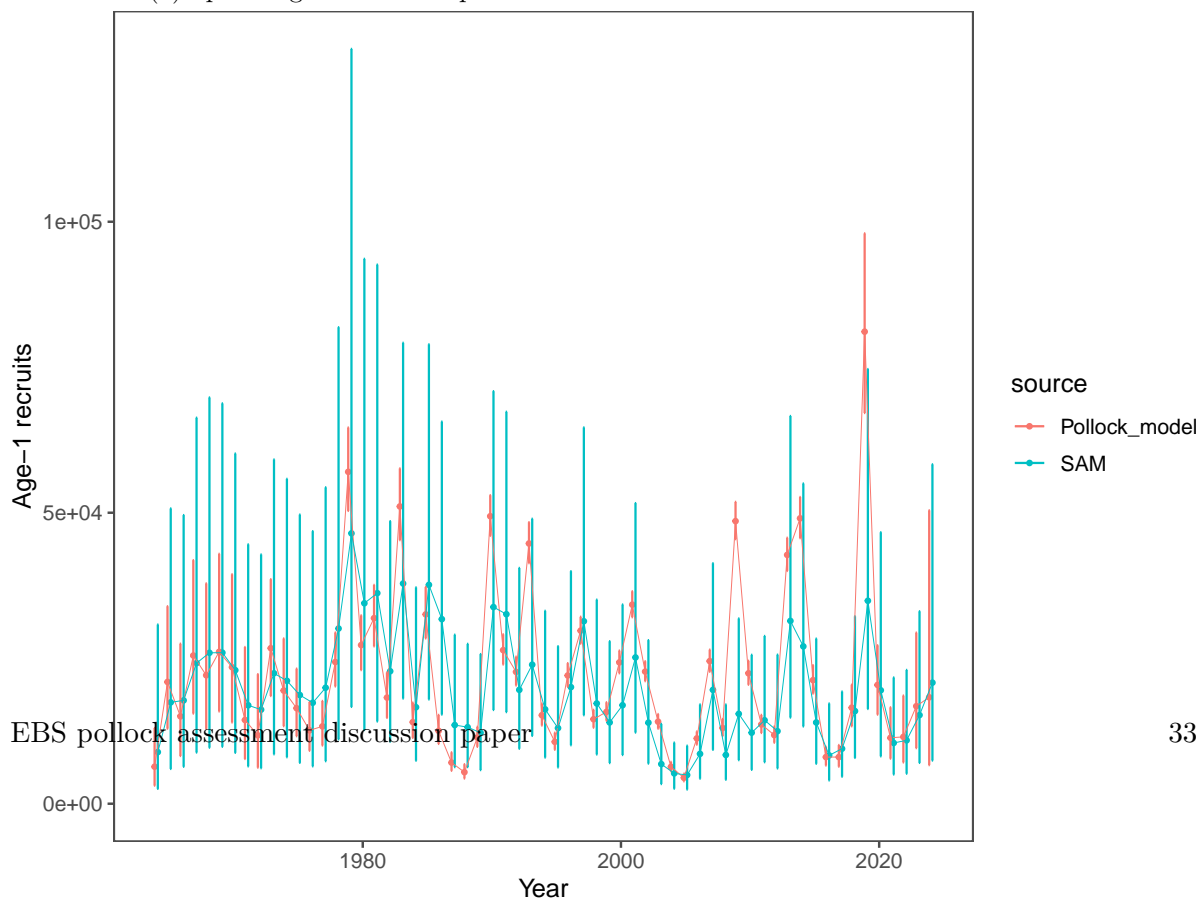
The Rceattle model is developed by the AFSC and is written in TMB (Adams et al. (2022)). The model is an extension of the ADMB version used for multi-species model in the Eastern Bering Sea (Jurado-Molina, Livingston, and Ianelli (2005), Holsman et al. (2024)) and has a number of newer features. It is flexible enough to include a variety of data types using lengths and can be set up for species tracked for sexually dimorphic growth and other processes. Since it has been recast into TMB, it allows for the ability to specify random effects in a number of processes. One development of the model is to have the flexibility to bridge among extant single-species modeling platforms by using it in “single-species” mode. To illustrate that, we introduce an application of Rceattle to the EBS pollock data set. The model is still under development, but we present preliminary results for consideration. Initial runs are favorable and we were able to have time-varying selectivity treated as fixed-effects (meaning penalty weights pre-specified) but also with them specified as random effects (with the penalty “variances” estimated (Figure 29)). The model estimates of SSB are similar to the 2024 model but more work is needed to understand where differences remain in the model specification (Figure 30).

## 5.4 Spatial modeling of the EBS pollock stock

In development is a spatially disaggregated general model by colleagues at the Auke Bay lab and the University of Alaska, Fairbanks (M. Cheng (2025)). This general model which has



(a) Spawning biomass comparison of the 2024 assessment and SAM estimates.



(b) Age-1 recruitment comparison of the 2024 assessment and SAM estimates.

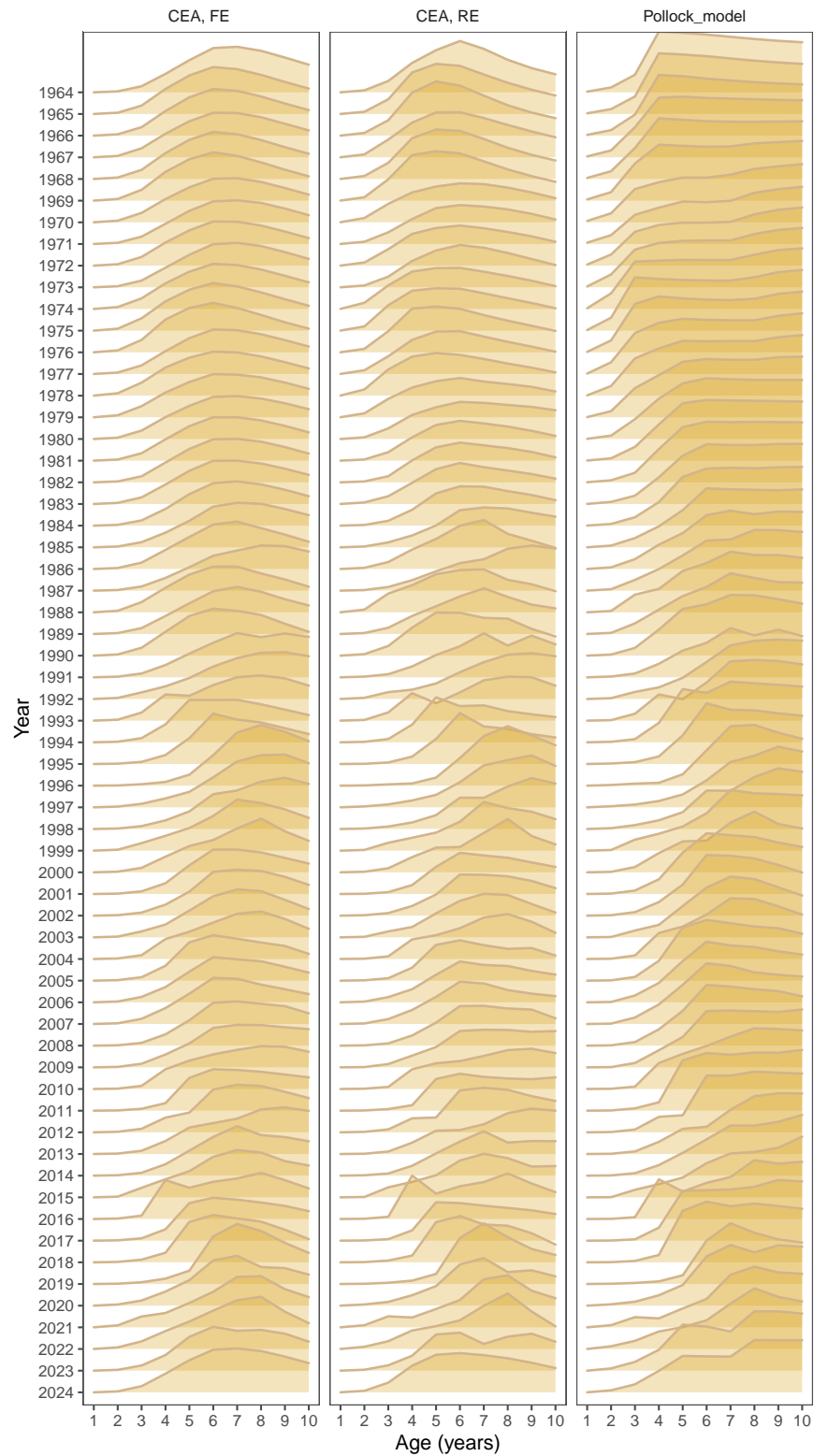


Figure 29: Model selectivity estimates comparing last year's model with the Rceattle model with fixed-effect time varying fishery selectivity (CEA, FE) and random effects (CEA, RE).  
 EBS pollock assessment discussion paper

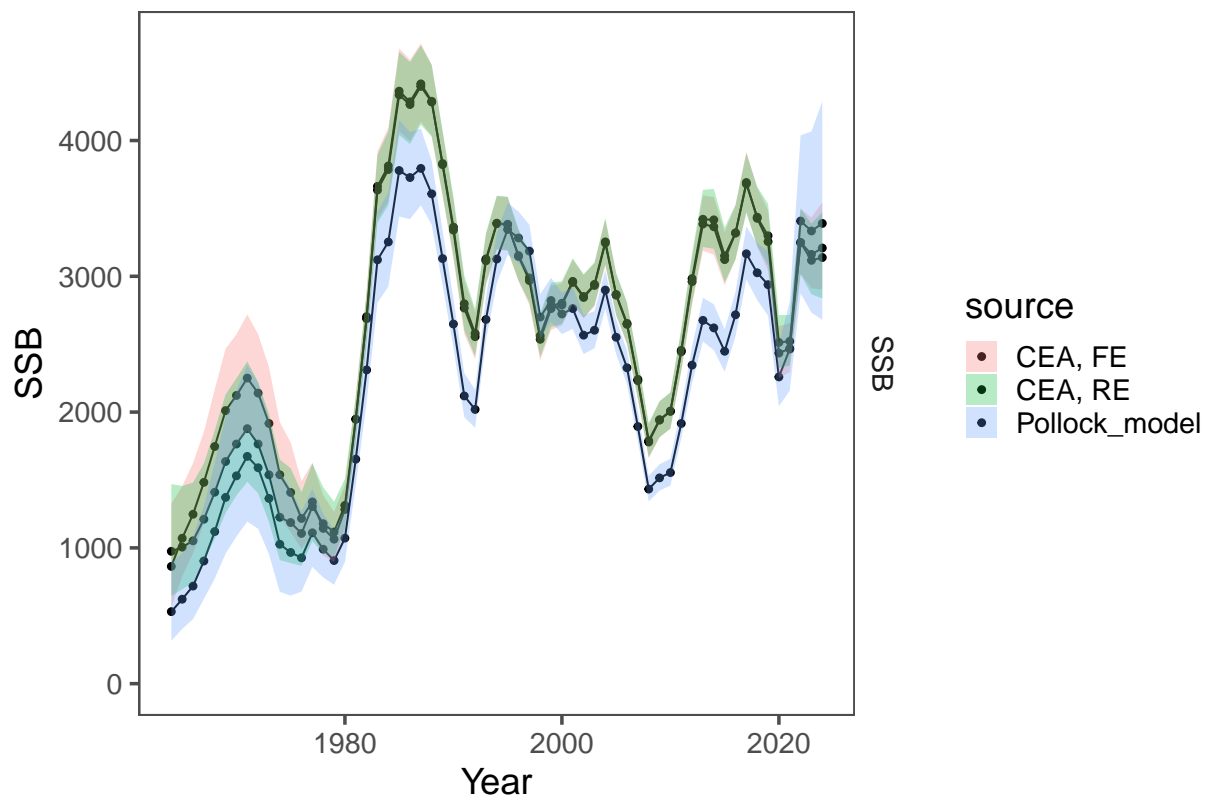
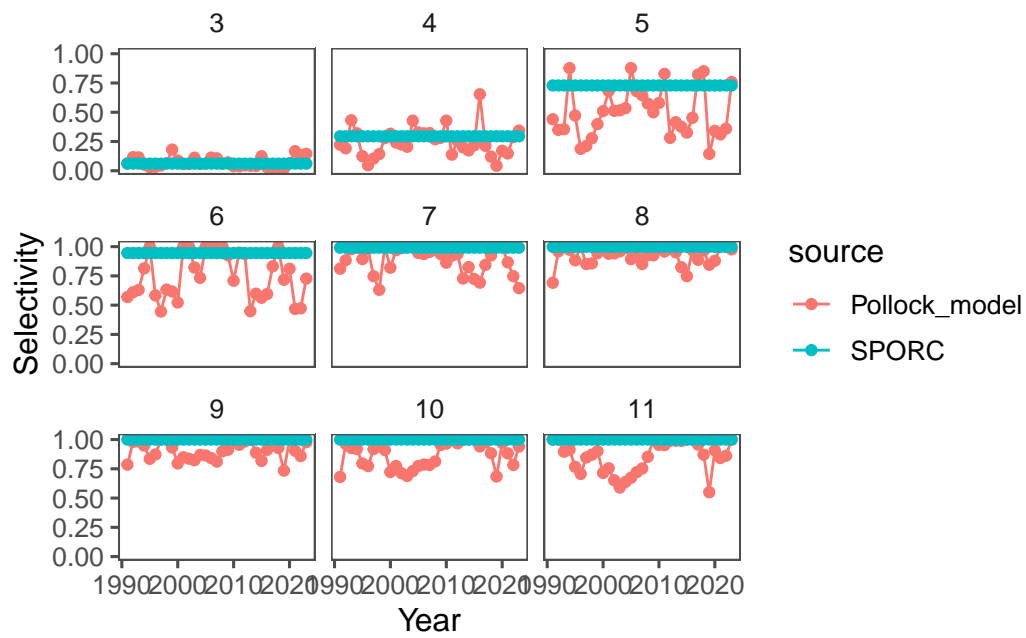


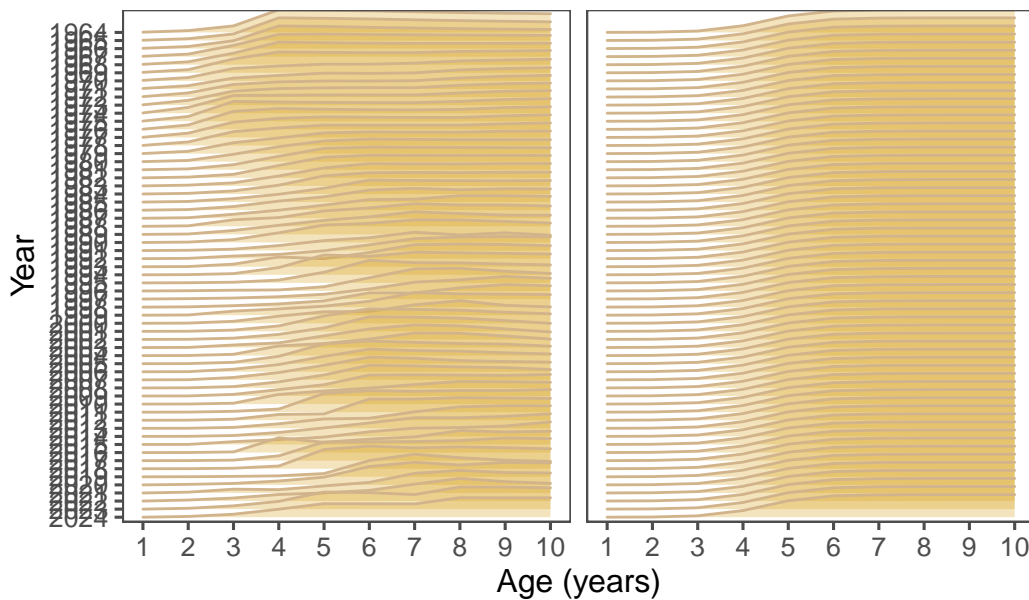
Figure 30: Model estimates of spawning biomass (SSB) comparing last year's model with the Rceattle model with fixed-effect time varying fishery selectivity (CEA, FE) and random effects (CEA, RE).

flexibility to specify regions and a variety of options including random effects. While this model is not yet fully developed, we present preliminary results here. The model is written in RTMB and is similar to the WHAM model. We pursue this to explore the potential for a more flexible model that can be used for more explicit considerations of stocks linked to different regions (e.g., the US and Russian portions of the EBS). Also, this model to add sex-specificity and random effects (M. L. H. Cheng et al. (2025)).

Figure 31 shows the model results comparing time-varying selectivity whereas Figure 32 shows the model results time series. The difference in the age-1 recruits (Figure 33) is likely due to the fact that the spatial model currently has constant M-at-age compared to the base model.



(a) Model estimates of selectivity comparing last year's model with the SPock model.



(b) Model estimates of selectivity comparing last year's model with the SPock model.

Figure 31: Model estimates of selectivity comparing last year's model with the SPock model.

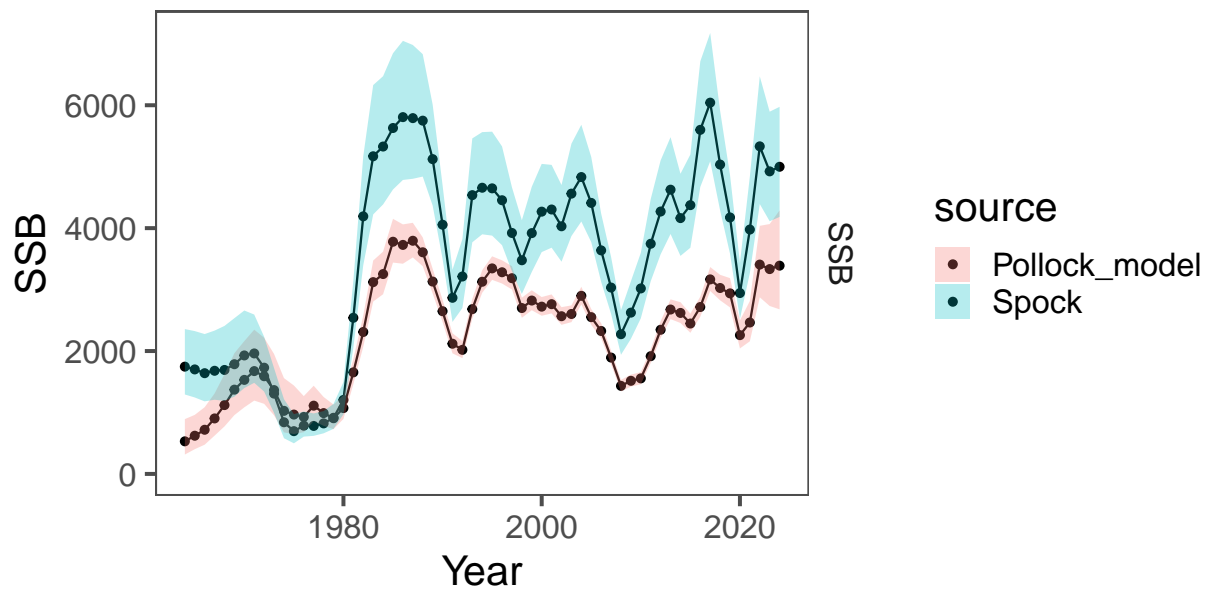


Figure 32: Model results comparing last year's model with the SPock model.

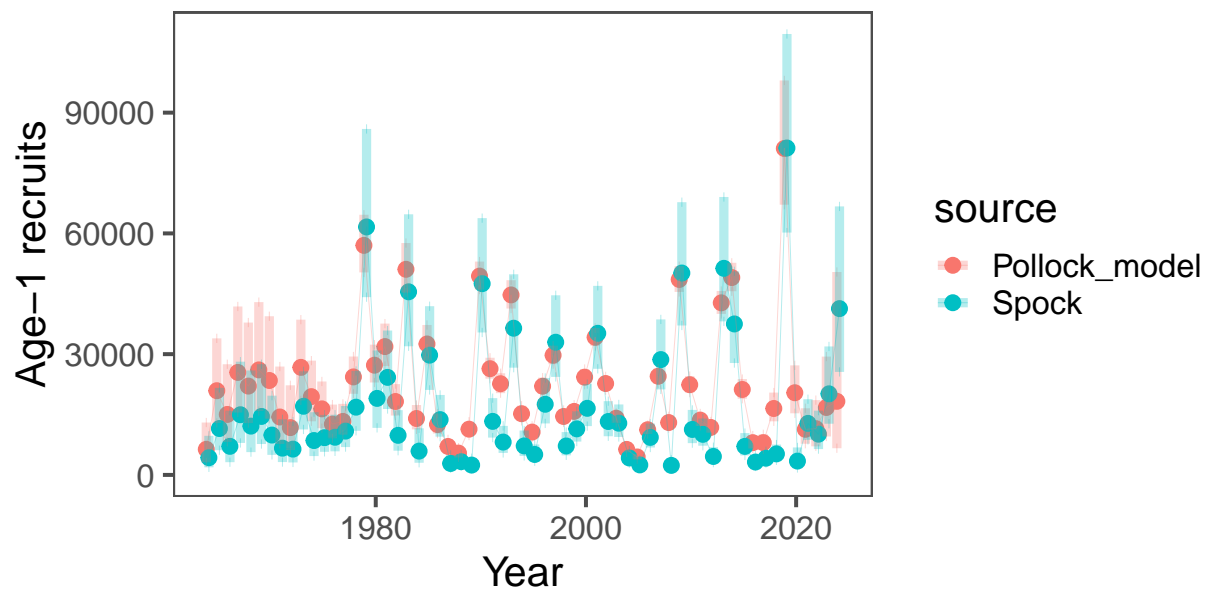


Figure 33: Model results comparing last year's model with the SPock model.

## 6 Bottom-trawl survey spatial distribution modeling

### 6.1 Abundance Index Methods

For model-based indices in the Bering Sea, we fitted observations of numerical abundance or biomass per unit area (where the choice of abundance or biomass varied by stock at the request of assessors) from all sampling locations in the 83-112 bottom trawl survey of the EBS, 1982-2025, including:

- Exploratory northern extension samples in 2001, 2005, and 2006
- 83-112 samples available in the NBS in 1982, 1985, 1988, 1991, 2010, 2017-2019, 2021-2023, and 2025

NBS samples collected prior to 2010 and in 2018 did not follow the 20 nautical mile sampling grid that was used in other survey years. Assimilating these unbalanced data requires extrapolating into unsampled areas, facilitated by including a spatially varying response to cold-pool extent (Thorson 2019). This spatially varying coefficient term was estimated for both linear predictors of a delta-model, and detailed comparison of results for EBS pollock has shown that it has a small but notable effect on these indices and resulting stock assessment outputs (O’Leary et al. 2020).

Rather than using the cold-pool covariate for yellowfin sole, we instead used the mean bottom temperature within the inner and middle domain strata from an interpolated temperature product. All models were fitted in the sdmTMB R package (<https://pbs-assess.github.io/sdmTMB>; Anderson et al. (2022)). All environmental data used as covariates were computed within the coldpool R package (<https://github.com/afsc-gap-products/coldpool>; Rohan, Barnett, and Charriere (2022)).

We used a spatiotemporal Poisson-link delta-model (Thorson 2018) involving two linear predictors and a gamma distribution to model positive catch rates. We predicted population density across the entire EBS and NBS in each year, using AFSC GAP-approved prediction grids. Spatial and spatiotemporal fields were estimated using a spatial mesh including 250 “knots” whose locations were determined using a K-means algorithm to place the knots evenly over space, in proportion to the dimensions of the prediction grid.

While the prior VAST model was fitted with a 750 knot mesh, we found that sdmTMB models fitted with mesh resolutions from 250-750 knots all provided extremely similar index estimates and that reducing the number of knots dramatically improved computation time. We estimated geometric anisotropy (how spatial autocorrelation declines with differing rates over distance in some cardinal directions than others) and included a spatial and spatiotemporal term for both linear predictors.

To facilitate prediction of density in regions with unsampled years, we specified that the spatiotemporal fields were structured over time as an AR(1) process (where the magnitude of autocorrelation was estimated as a fixed effect for each linear predictor). However, in cases where the AR(1) correlation parameter was estimated to be close to 1 for either model component, the model would be collapsed into a simpler structure by specifying  $\rho = 1$ , i.e., modeling spatiotemporal variation as a random walk. We did not include any temporal correlation for intercepts, which we treated as fixed effects for each linear predictor and year. Finally, we used epsilon bias-correction to correct for retransformation bias (Thorson and Kristensen 2016).

We checked model fits for convergence by confirming that the derivative of the marginal likelihood with respect to each fixed effect was sufficiently small (less than  $\sim 0.001$ ) and that the Hessian matrix was positive definite, along with other model checks produced by the function `sdmTMB::sanity()`. We then checked goodness of model fit by computing Dunn-Smyth randomized quantile residuals (Dunn and Smyth 1996) and visualizing these using a quantile-quantile plot within the DHARMA R package (Hartig 2021). We also evaluated the distribution of these residuals over space in each year, and inspected them for evidence of residual spatiotemporal patterns. Results compared favorably among methods, and the model index result was very similar to previous estimates (Figure 34).

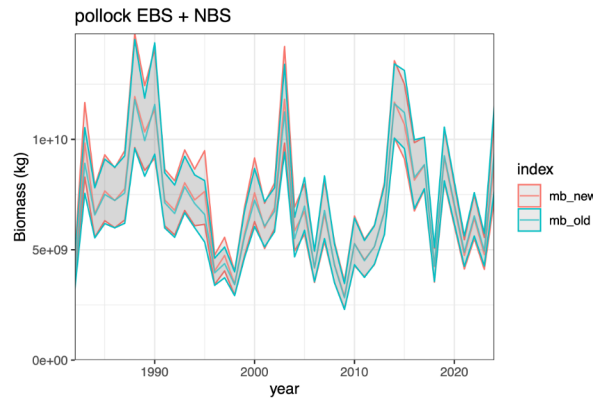


Figure 34: Model results comparing last year’s model with the revised sdmTMB model.

## 6.2 Model-based Age Composition Methods

For model-based estimation of age compositions in the Bering Sea, we fitted observations of numerical abundance-at-age at each sampling location. This was made possible by applying a year-specific, region-specific (EBS and NBS) age-length key to records of numerical abundance and length composition. In subcategories (combinations of year, length, age, sex) that contained insufficient data, age composition was computed from length composition given a globally pooled age-length key.

We computed these estimates in the `tinyVAST` R package (<https://vast-lib.github.io/tinyVAST>; Thorson et al. (2024)), assuming a Poisson-link delta-model (Thorson 2018) involving two linear predictors, and a gamma distribution to model positive catch rates. We did not include any density covariates in estimation of age composition for consistency with models used in the previous assessments and due to computational limitations.

We estimated the spatial range assuming geometric isotropy (i.e., spatial autocorrelation declines with equal rates over distance in all cardinal directions). We used the same extrapolation grid as implemented for abundance indices, but here we modeled spatial and spatiotemporal fields with a mesh with a coarser spatial resolution than the index model, here using 50 “knots”. This reduction in the spatial resolution of the model, relative to that used for the abundance indices, was necessary due to the increased computational load of fitting multiple age categories.

The estimates of abundance at age were bias-corrected with the same method as the abundance index. We implemented the same diagnostics to check convergence and model fit as those used for the abundance index. Results compared favorably among methods (Figure 35).

Both the model-based bottom trawl index and age composition estimates were calculated using code in the AFSC GAP model-based indices GitHub repository (<https://github.com/afsc-gap-products/model-based-indices>). Design-based estimates of bottom trawl products were calculated using code in the `gapindex` R package (<https://github.com/afsc-gap-products/gapindex>).



Figure 35: Model results comparing last year's model with the revised sdmTMB model.

## 7 Pollock model fishery selectivity notes

The following provides more detail based on the actual code as adopted from Butterworth, Ianelli, and Hilborn (2003). During the CIE review this came up as a point of clarification and we provide the details here for reference.

### 7.1 Initial Setup

#### 7.1.1 Average Selectivity

$$\overline{sel} = \log(\text{mean}(\exp(\mathbf{c})))$$

where  $\mathbf{c}$  is the vector of base selectivity coefficients.

#### 7.1.2 Base Year Selectivity ( $t = \text{stsel}$ )

For ages 1 to  $n_{\text{sel}}$ :

$$\log(S_{t,a}) = c_a \quad \text{for } a = 1, 2, \dots, n_{\text{sel}}$$

For older ages (plateau assumption):

$$\log(S_{t,a}) = c_{n_{\text{sel}}} \quad \text{for } a = n_{\text{sel}} + 1, \dots, n_{\text{ages}}$$

#### 7.1.3 Initial Normalization

$$\log(S_{t,a}) = \log(S_{t,a}) - \log\left(\frac{1}{n_{\text{ages}}} \sum_{a=1}^{n_{\text{ages}}} \exp(\log(S_{t,a}))\right)$$

This ensures:  $\prod_{a=1}^{n_{\text{ages}}} S_{t,a}^{1/n_{\text{ages}}} = 1$

## 7.2 Time-Varying Component

### 7.2.1 Selectivity Evolution ( $t > \text{stsel}$ )

For years without regime changes:

$$\log(S_{t+1,a}) = \log(S_{t,a}) \quad \text{for all ages}$$

For years with regime changes (when  $t + 1 \in \text{yrs\_ch\_fsh}$ ):

**Ages 1 to  $\text{nsel}$ :**

$$\log(S_{t+1,a}) = \log(S_{t,a}) + \delta_{ii,a} \quad \text{for } a = 1, 2, \dots, \text{nsel}$$

**Older ages (plateau maintained):**

$$\log(S_{t+1,a}) = \log(S_{t+1,\text{nsel}}) \quad \text{for } a = \text{nsel} + 1, \dots, \text{nages}$$

where: -  $\delta_{ii,a}$  is the selectivity deviation for change event  $ii$  and age  $a$  -  $ii$  is the index of the current selectivity change event

### 7.2.2 Annual Normalization

For each year  $t$ :

$$\log(S_{t,a}) = \log(S_{t,a}) - \log\left(\frac{1}{\text{nages}} \sum_{a=1}^{\text{nages}} \exp(\log(S_{t,a}))\right)$$

and the actual selectivity is:

$$S_{t,a} = \exp(\log(S_{t,a}))$$

## 7.3 Constraints and Properties

1. **Normalization Constraint:**

$$\prod_{a=1}^{\text{nages}} S_{t,a}^{1/\text{nages}} = 1 \quad \forall t$$

2. **Dome-shaped Constraint:**

$$S_{t,a} = S_{t,\text{nsel}} \quad \text{for } a > \text{nsel}$$

3. **Continuity:** Selectivity changes only occur in specified years (`yrs_ch_fsh`)
4. **Deviation Structure:**

$\delta_{ii} \sim$  estimated deviations with penalties

### 7.3.1 Implementation Notes

- Base coefficients  $\mathbf{c}$  are estimated parameters
- Deviations  $\delta$  are typically penalized to prevent overfitting
- The plateau assumption prevents unrealistic selectivity patterns for older ages
- Geometric mean normalization maintains interpretability relative to fishing mortality

## 7.4 Regularization and Penalty Terms

The selectivity model includes several penalty terms to ensure realistic and stable estimates:

Penalty Term 1: Deviation Magnitude Penalty  $P_1 = \frac{\lambda_1}{N_{groups}} \sum_{i=1}^{N_{changes}} \sum_{a=1}^{n_{sel}} \delta_{i,a}^2$  where:

$\lambda_1 = \text{ctrl\_flag}(10)$  (penalty weight parameter)  $N_{groups} = \text{group\_num\_fsh}$  (number of fishery groups) This shrinks deviations toward zero when not informed by data

Penalty Term 2: Base Selectivity Smoothness Penalty  $P_2 = \frac{\lambda_2}{N_{changes}} \sum_{a=1}^{n_{sel}-2} [\nabla^2 \log(S_{styr,a})]^2$  where:

$\lambda_2 = \text{ctrl\_flag}(11)$  (smoothness penalty weight)  $\nabla^2 \log(S_{t,a}) = \log(S_{t,a+2}) - 2\log(S_{t,a+1}) + \log(S_{t,a})$  (second difference) This penalizes high curvature in the base selectivity pattern

Penalty Term 3: Change Year Smoothness Penalty  $P_3 = \frac{\lambda_2}{N_{changes}} \sum_{i=1}^{N_{changes}} \sum_{a=1}^{n_{sel}-2} [\nabla^2 \log(S_{t_i,a})]^2$

where  $t_i$  represents each change year. This maintains smooth selectivity curves in all change years. Penalty Term 4: Temporal Stability Penalty (Random Walk)

$P_4 = \sum_{i=1}^{N_{changes}} \frac{1}{2\sigma_i^2} \sum_{a=1}^{n_{ages}} [\log(S_{t_i-1,a}) - \log(S_{t_i,a})]^2$  where:

$\sigma_i = \text{sel\_ch\_sig\_fsh}(i)$  (change magnitude parameter for change event  $i$ ) This penalizes large year-to-year changes in selectivity Only applied when  $t_i \neq styr$  (change year is not the start year)

Total Penalty Function  $P_{total} = P_1 + P_2 + P_3 + P_4$  The total negative log-likelihood becomes:  $-\log L = -\log L_{data} + P_{total}$  Penalty Interpretation

$P_1$ : Prevents overfitting by shrinking deviations toward zero  $P_2$  and  $P_3$ : Enforce smooth, realistic selectivity curves (penalize jagged patterns)  $P_4$ : Constrains temporal evolution (selectivity changes smoothly over time)

The penalty weights ( $\lambda_1, \lambda_2$ ) and variance parameters ( $\sigma_i$ ) control the balance between fitting data and maintaining realistic selectivity patterns. Implementation Notes

Base coefficients  $\mathbf{c}$  are estimated parameters Deviations  $\delta$  are penalized to prevent overfitting The plateau assumption prevents unrealistic selectivity patterns for older ages Geometric mean normalization maintains interpretability relative to fishing mortality Second differences approximate curvature (second derivatives) for smoothness penalties

# References

- Adams, Grant D., Kirstin K. Holsman, Steven J. Barbeaux, Martin W. Dorn, James N. Ianelli, Ingrid Spies, Ian J. Stewart, and André E. Punt. 2022. “An Ensemble Approach to Understand Predation Mortality for Groundfish in the Gulf of Alaska.” *Fisheries Research* 251: 106303. <https://doi.org/https://doi.org/10.1016/j.fishres.2022.106303>.
- Anderson, S. C., E. J. Ward, P. A. English, and L. A. K. Barnett. 2022. “sdmTMB: An R Package for Fast, Flexible, and User-Friendly Generalized Linear Mixed Effects Models with Spatial and Spatiotemporal Random Fields.” *BioRxiv*, 2022–03. <https://doi.org/10.1101/2022.03.24.485545>.
- Butterworth, Doug S, James N Ianelli, and Ray Hilborn. 2003. “A Statistical Model for Stock Assessment of Southern Bluefin Tuna with Temporal Changes in Selectivity.” *Afr. J. Mar. Sci.* 25: 331–61.
- Cheng, Matthew. 2025. *SPoCK: Spatial (s) Population (Po) Model with Connectivity (c) Across Generalized Components (k)*.
- Cheng, Matthew L. H., Daniel R. Goethel, Peter-John F. Hulson, Michael J. Wilberg, Craig Marsh, and Curry J. Cunningham. 2025. “Misspecifying Sex-Structured Dynamics in Stock Assessment Models.” *Fish and Fisheries*, February, 1–19. <https://doi.org/10.1111/faf.12891>.
- Dunn, K. P., and G. K. Smyth. 1996. “Randomized Quantile Residuals.” *Journal of Computational and Graphical Statistics* 5 (1): 1–10. <https://doi.org/10.1080/10618600.1996.10474708>.
- Grandin, C. J., K. F. Johnson, A. M. Edwards, and A. M. Berger. 2024. “Status of the Pacific Hake (Whiting) Stock in U.S. And Canadian Waters in 2024.” Joint Technical Committee of the U.S.; Canada Pacific Hake/Whiting Agreement, National Marine Fisheries Service; Fisheries; Oceans Canada.
- Hartig, Florian. 2021. *DHARMA: Residual Diagnostics for Hierarchical (Multi-Level / Mixed) Regression Models*. <http://florianhartig.github.io/DHARMA/>.
- Helser, Thomas E., Irina Benson, Jason Erickson, Jordan Healy, Craig Kastle, and Jonathan A. Short. 2019. “A Transformative Approach to Ageing Fish Otoliths Using Fourier Transform Near Infrared Spectroscopy: A Case Study of Eastern Bering Sea Walleye Pollock (*Gadus Chalcogrammus*).” *Canadian Journal of Fisheries and Aquatic Sciences* 76 (5): 780–89. <https://doi.org/10.1139/cjfas-2018-0112>.
- Holsman, K. K., J. Ianelli, K. Shotwell, S. Barbeaux, K. Aydin, G. Adams, K. Kearney, A. Sulc, and S. Wassermann. 2024. “Climate-Enhanced Multispecies Stock Assessment for Walleye Pollock, Pacific Cod, and Arrowtooth Flounder in the Eastern Bering Sea.”

- Technical Report. Anchorage, AK: North Pacific Fishery Management Council. <https://www.npfmc.org/wp-content/PDFdocuments/SAFE/2024/EBSmultispp.pdf>.
- Ianelli, Jim. 2025. *Ebswp: Facilitates Model Running for EBS Pollock*. <http://afsc-assessments.github.io/EBSpollock/>.
- Jurado-Molina, J., P. A. Livingston, and J. N. Ianelli. 2005. “Incorporating Predation Interactions to 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): 1865–73.
- Methot, Richard D., and Chantell R. Wetzel. 2013. “Stock Synthesis: A Biological and Statistical Framework for Fish Stock Assessment and Fishery Management.” *Fisheries Research* 142 (May): 86–99. <https://doi.org/10.1016/j.fishres.2012.10.012>.
- Nielsen, Anders, and Casper W. Berg. 2014. “Estimation of Time-Varying Selectivity in Stock Assessments Using State-Space Models.” *Fisheries Research* 158: 96–101. <https://doi.org/10.1016/j.fishres.2014.01.014>.
- Nielsen, Anders, Niels T Hintzen, Henrik Mosegaard, Vanessa Trijoulet, and Casper W Berg. 2021. “Multi-Fleet State-Space Assessment Model Strengthens Confidence in Single-Fleet SAM and Provides Fleet-Specific Forecast Options.” *ICES Journal of Marine Science* 78 (6): 2043–52. <https://doi.org/10.1093/icesjms/fsab078>.
- O’Leary, C. A., J. T. Thorson, J. N. Ianelli, and S. Kotwicki. 2020. “Adapting to Climate-driven Distribution Shifts Using Model-based Indices and Age Composition from Multiple Surveys in the Walleye Pollock (*Gadus Chalcogrammus*) Stock Assessment.” *Fisheries Oceanography* 29 (6): 541–57. <https://doi.org/10.1111/fog.12494>.
- Punt, A. E., D. C. Smith, K. KrusicGolub, and S. Robertson. 2008. “Quantifying Age-Reading Error for Use in Fisheries Stock Assessments, with Application to Species in Australia’s Southern and Eastern Scalefish and Shark Fishery.” *Can. J. Fish. Aquat. Sci.* 65: 1991–2005.
- Rohan, S. K., L. A. K. Barnett, and N. Charriere. 2022. “Evaluating Approaches to Estimating Mean Temperatures and Cold Pool Area from AFSC Bottom Trawl Surveys of the Eastern Bering Sea.” Tech. Mem. NMFS-AFSC-456. U.S. Dep. Commer., NOAA. <https://doi.org/10.25923/1wwh-q418>.
- Stock, Brian C., and Timothy J. Miller. 2021. “The Woods Hole Assessment Model (WHAM): A General State-Space Assessment Framework That Incorporates Time- and Age-Varying Processes via Random Effects and Links to Environmental Covariates.” *Fisheries Research* 240 (August): 105967. <https://doi.org/10.1016/j.fishres.2021.105967>.
- Thorson, J. T. 2018. “Three Problems with the Conventional Delta-Model for Biomass Sampling Data, and a Computationally Efficient Alternative.” *Canadian Journal of Fisheries and Aquatic Sciences* 75 (9): 1369–82. <https://doi.org/10.1139/cjfas-2017-0266>.
- . 2019. “Measuring the Impact of Oceanographic Indices on Species Distribution Shifts: The Spatially Varying Effect of Cold-pool Extent in the Eastern Bering Sea.” *Limnology and Oceanography* 64 (6): 2632–45. <https://doi.org/10.1002/lno.11238>.
- Thorson, J. T., S. C. Anderson, P. Goddard, and C. N. Rooper. 2024. “tinyVAST: R Package with an Expressive Interface to Specify Lagged and Simultaneous Effects in Multivariate Spatio-Temporal Models.” *arXiv Preprint* arXiv:2401.10193.
- Thorson, J. T., and K. Kristensen. 2016. “Implementing a Generic Method for Bias Correc-

- tion in Statistical Models Using Random Effects, with Spatial and Population Dynamics Examples.” *Fisheries Research* 175: 66–74. <https://doi.org/10.1016/j.fishres.2015.11.016>.
- Trijoulet, Vanessa, Gavin Fay, and Timothy J. Miller. 2020. “Performance of a State-Space Multispecies Model: What Are the Consequences of Ignoring Predation and Process Errors in Stock Assessments?” *Journal of Applied Ecology* 57 (1): 121–35. <https://doi.org/https://doi.org/10.1111/1365-2664.13515>.
- Urmy, Samuel S., Patrick Ressler, and Alex De Robertis. 2025. “Estimating Uncertainty from Multiple Sources in Acoustic-Trawl Surveys.” *Unpublished Manuscript*, May.