Fisheries and Oceans Canada | Pêches et Océans Canada
Ecosystems and Oceans Science | Sciences des écosystèmes et des océans

Canadian Science Advisory Secretariat (CSAS)

Research Document 2026/050

Gulf Region

State-Space Models for NAFO Division 4TVn Fall Spawning Atlantic Herring (*Clupea harengus*) Subpopulations

François Turcotte

Fisheries and Oceans Canada
Gulf Fisheries Centre
343 Université Avenue, P.O. Box 5030
Moncton, NB E1C 9B6

August 2026

Foreword

This series documents the scientific basis for the evaluation of aquatic resources and ecosystems in Canada. As such, it addresses the issues of the day in the time frames required and the documents it contains are not intended as definitive statements on the subjects addressed but rather as progress reports on ongoing investigations.

Published by:

Fisheries and Oceans Canada
Canadian Science Advisory Secretariat
200 Kent Street
Ottawa ON K1A 0E6

http://www.dfo-mpo.gc.ca/csas-sccs/
DFO.CSAS-SCAS.MPO@dfo-mpo.gc.ca

© His Majesty the King in Right of Canada, as represented by the Minister of the Department of Fisheries and Oceans, 2026

This report is published under the Open Government Licence - Canada

ISSN 1919-5044
ISBN 978-0-662-44065-9 Cat. No. Fs70-5/2026-050E-PDF

Correct citation for this publication:

Turcotte, F. 2026. State-Space Models for NAFO Division 4TVn Fall Spawning Atlantic Herring (*Clupea harengus)* Subpopulations. DFO Can. Sci. Advis. Sec. Res. Doc. 2026/050. iv + 98 p.

Aussi disponible en français :

*Turcotte, F. 2026. Modèles état-espace pour les sous-populations de hareng atlantique (*Clupea harengus*) frayant à l'automne dans la division 4TVn de l'OPANO. Secr. can. des avis sci. du MPO. Doc. de rech. 2026/050. iv + 104 p.*

TABLE OF CONTENTS

ABSTRACT

The last stock assessment of the Northwest Atlantic Fisheries Organization (NAFO) Division 4TVn fall spawning Atlantic Herring raised significant concerns that triggered a review of the assessment framework. This assessment framework review evaluates a state-space modeling approach to better capture time-varying productivity and selectivity and improving robustness over prior models. Key advancements include transitioning from statistical catch-at-age models to a state-space framework, developing separate models for each subpopulation (North, Middle, South) rather than a single aggregated model and disaggregating catch by fixed and mobile gear fleets with independently estimated selectivity. These structural changes enhance biological realism, reduce parameter confounding, and improve diagnostic performance. For each subpopulation, objectives included developing base and sensitivity models and identifying best-fitting models. Overall, the state-space models performed well and are recommended for future stock assessments and management.

INTRODUCTION

The last stock assessment of the Northwest Atlantic Fisheries Organization (NAFO) Division 4TVn fall spawning Atlantic Herring raised significant concerns that triggered a review of the assessment framework (Maynard et al. 2024). The main issues identified were the unrealistic declines in time-varying natural mortality for young fish (ages 2 to 6) and inflated cohort strength in population forecasts (Maynard et al. 2024). The statistical catch-at-age (SCA) model used in the previous assessments was developed for the 2020 stock assessment and implemented in AD Model Builder (ADMB) (Turcotte et al. 2021a, Fournier et al. 2012). This model improved the retrospective pattern by allowing for time variation in processes that are often treated as constant in stock assessment (i.e., catchability to the catch-per-unit effort (CPUE) index of biomass and natural mortality). Retrospective patterns generally suggest that the estimation model is misspecified, possibly because of incorrect model assumptions such as constant natural mortality or constant catchability (Cadrin 2025). However, the SCA model remained limited by subjective likelihood penalty tuning and insufficient flexibility for complex time-varying processes.

SCA models typically only estimate year-specific fishing mortality and recruitment due to limited degrees of freedom, often relying on penalized deviations with fixed or iteratively tuned penalties that introduce subjectivity. State-space models, such as those implemented in the Woods Hole Assessment Model (WHAM) R package (Stock and Miller 2021), address this using random effects, estimating standard deviation (SD) components to objectively constrain temporal variation while reducing the total number of parameters. The flexibility of WHAM allows the inclusion of random effects for key processes including natural mortality, numbers-at-age, selectivity, and catchability, while also supporting the inclusion of environmental covariates to link population dynamics to external drivers.

This assessment framework review explores a state-space modeling framework that integrates mechanistic (environmental covariate-driven) and empirical (stochastic random effects) approaches to better capture time-varying productivity and selectivity, improving robustness over prior models. Key advancements include transitioning from SCA models to a state-space framework, developing separate models for each subpopulation (North region, Middle region, South region) rather than a single aggregated model using shared parameters and indices, disaggregating catch into fixed and mobile gear fleets with independently estimated selectivity, and incorporating predator abundance effects on natural mortality for older ages. These structural changes are expected to enhance biological realism, reduce parameter confounding, and improve diagnostic performance.

METHODS

STOCK STRUCTURE

Distinct models were developed under the assumption that NAFO 4TVn fall-spawning herring constitutes a metapopulation, assessable and manageable as three subpopulations (North, Middle, and South) defined by spatially distinct spawning grounds identified in prior assessments. Spawning occurs predictably on these spatially and temporally defined grounds, which support the majority of the gillnet fishery. Atlantic herring exhibit high spawning site fidelity (natal homing) with limited straying, consistent with a metapopulation structure (McQuinn 1997; Brophy et al. 2006; Moll et al. 2025). Accordingly, surveys were assigned selectively to the most relevant subpopulation based on available movement data: Chaleur Bay/Miscou (BDCM) acoustic survey and the multispecies bottom trawl research vessel (RV) survey were applied

only where biologically appropriate (Turcotte et al. In press). In previous assessment frameworks, these surveys were uniformly applied across all subpopulations, which may have introduced misleading dynamics information in subpopulations where these surveys were less representative.

TIME VARYING PARAMETERS

Natural mortality

Evidence indicates that the lack of recovery of southern Gulf of St. Lawrence (sGSL) groundfish reflects substantial increases in natural mortality among large individuals across the demersal fish community and these mortality patterns represent a broad, ecosystem-wide phenomenon (Swain & Benoît 2015). A major driver of this elevated natural mortality is predation by grey seal (*Halichoerus grypus*), which has been linked to widespread mortality increases among adult groundfish (Benoît & Swain 2008; Swain & Benoît 2015; Neuenhoff et al. 2019; Rossi et al. 2021). Grey seal predation risk has also been shown to drive spatial distribution shifts of many sGSL groundfish, with individuals moving into deeper, lower-risk areas as predation pressure increased in traditional habitats (Swain et al. 2015).

The abundance of major herring predators in the sGSL has shifted markedly in recent decades (Benoît and Rail 2016). sGSL spring herring are a key prey species for grey seals (*Halichoerus grypus*), northern gannets (*Morus bassanus*), cetaceans, Atlantic Cod (*Gadus morhua*), White Hake (*Urophycis tenuis*), and Atlantic Bluefin Tuna (*Thunnus thynnus*). Of these, cod, seals, tuna, and gannets have all experienced major abundance changes, with increasing trends in seals, tuna, and gannets. Recent increases in model-estimated natural mortality for older herring (Turcotte et al. 2021b; Rossi and Turcotte In press) align temporally with rising abundances of major predators, notably Atlantic Bluefin Tuna (Turcotte et al. 2021b; 2021c; 2023) and grey seals (Hammill et al. 2023). Alternative explanations for elevated M, such as unreported catches or disease, appear negligible based on available evidence (Turcotte et al. 2021b).

Trends in estimated time-varying natural mortality of spring spawning herring younger fish (age group 2-6) were correlated with the decline in cod abundance in previous assessments (Turcotte et al. 2021b). However, in fall spawning herring, estimated time-varying natural mortality for this younger age group declined to near zero, which is unrealistic and likely to introduce correlated effects on recruitment and impact natural mortality estimates of older fish. Consequently, state-space models developed here were built under the assumptions that natural mortality for younger fish is not well informed by the available age composition data and should be fixed or estimated at a static mean level. Time-varying natural mortality for older fish is assumed to be reliably estimable either by the age composition data or with the inclusion of the major predators, Atlantic Bluefin Tuna and grey seals.

WHAM allows linear covariate effects on natural mortality (M), with the mean natural mortality either fixed or estimated, and the effects of environmental covariates estimated from the data. Covariate effects can be applied to selected ages through parameter mapping. In addition, process error in M can be included via random effects to account for deviations in natural mortality that are not explained by the environmental covariates.

Gillnet Catch-Per-Unit-Effort (CPUE) catchability

Time-varying catchability is common in fisheries and surveys (Wilberg et al. 2010) and often varies with stock abundance or distribution (e.g., Peterman and Steer 1981; Harley et al. 2001). A typical outcome is hyperstability, where catchability increases as abundance declines,

causing CPUE to remain high despite population decreases (Hilborn and Walters 1992). This pattern is common in small pelagic fish due to schooling and targeting of spawning aggregations, which can mask declines and delay recognition of collapse (Erisman et al. 2011). In some cases, density-dependent catchability can even drive “catchability-led stock collapse” (Pitcher 1995; Mackinson et al. 1997). Winters and Wheeler (1985) demonstrated that catchability in Northwest Atlantic herring stocks is inversely proportional to stock area. In Fortune Bay, reductions in abundance were associated with contractions in distribution rather than changes in density in a stable area.

Two versions of the CPUE index were evaluated in sensitivity runs for all subpopulation models. CPUE-GLM is the index used in previous assessments, whereas CPUE-GAM is an updated index produced by Maynard et al (In press). However, substantial uncertainty remains regarding the reliability of both indices because their values are highly sensitive to the data and assumptions used to reconstruct effort data (Maynard et al In press).

Selectivity

Considerable variation in the age composition of gillnet catches occurs in the three subpopulations data. The underlying causes were identified as changes in size-at-age and in gillnet mesh sizes over years (Turcotte et al. In press). Consequently, the commercial gillnet fishery selectivity should be modeled using random effects in the population model. The expectation is that selectivity for younger fish would decrease over time. A mean selectivity curve is anticipated to be dome-shaped as the peak selectivity from the selectivity at length modeling corresponds to age 10 on average over years (Turcotte et al. In press).

The spawning-grounds acoustic survey does not provide age-composition data. To define a selectivity function for this survey, the mean maturity-at-age schedule averaged across all years was applied. This approach assumes that immature fish do not aggregate on the spawning grounds and therefore are not available to the acoustic survey. This assumption is supported by the age and maturity composition of experimental gillnet catches on the spawning grounds, in which age-2, age-3, and immature fish are largely absent (Turcotte et al. In press).

DATA

Data inputs were selected based on the region-specific data review (Turcotte et al. In press). Departures from these recommendations were made when model diagnostics were unsatisfactory or when the parameters associated with the data sources were not estimable. Two fleet models (distinct fixed gear and mobile gear fleets) were analyzed first. The mobile gear catch was a combination of the regional inshore fleet catches and the overwinter fishery catches. Because no information is available to partition catches based on biological characteristics of the fish or by subpopulation, overwinter fishery catches were allocated to regions according to each region’s proportion of total gillnet catches, consistent with previous assessments (Turcotte et al. In press). The overwinter fishery proportions- and weight-at-age were combined proportionally to the regional inshore total catch and regional overwinter total catch. Coefficient of variation (CV) for sources of catch were based on a subjective assessment how reliable the indices were judged to be in terms of precision and the extent to which it is likely to be proportional to biomass (Francis et al. 2002; Francis 2011).

Data inputs

- Fishery catch-at-age: total catch, proportions-at-age and weight-at-age by fishing fleet (gillnets: fixed gear; seiners: mobile gear);
- DFO Chaleur Bay/Miscou acoustic survey (BDCM);

- DFO multispecies bottom trawl research vessel survey (RV);
- Spawning ground acoustic survey (SG);
- Experimental gillnet survey;
- Age-at-maturity ogives;
- Predator abundance indices;
- CPUE: historical GLM based (CPUE-glm) and new GAM based (CPUE-gam);
- Stock weight-at-age (September 1st): fishery weight-at-age.

SUBPOPULATIONS MODELS

A complete description of the WHAM package structure and model equations can be found in Stock and Miller (2021). To establish a base model with acceptable diagnostics for each subpopulation, initial model parameterizations were tested against model diagnostics for the following model assumptions:

1. Age composition likelihoods options: Dirichlet pooling zero observations with adjacent ages, Dirichlet treating zero observations as missing, Logistic normal treating zero observations as missing, Logistic normal pooling zero observations with adjacent ages and Logistic normal treating zero observations as missing with first-order autoregressive (AR1) correlation;
2. Numbers-at-age (NAA) options: random effects on all NAA (full state-space) or on recruitment only, independent and identically distributed (IID) random effects, AR1 correlated by age (AR1a), AR1 correlated by year (AR1y), or 2DAR1 correlated by age and year;
3. Numbers-at-age in year 1 options: age specific fixed effects, equilibrium, age specific independent random effects and age specific AR1 random effects;
4. Recruitment model: annual recruitment as a random walk with annual deviations as random effects, mean and annual recruitment as random effects, Beverton-Holt stock-recruitment, Ricker stock-recruitment.

Specific natural mortality and selectivity assumptions for each subpopulation were defined as described in the following sections. Once base models were established for each subpopulation, sensitivity analyses of models assuming different data sources and parameterization of key processes were performed. The main sensitivity analyses explored alternative structural assumptions about natural mortality, selectivity, process error in numbers-at-age, predator abundance effects on M and the use of the fixed gear fishery CPUE as an age aggregated index of abundance with time varying catchability.

Model diagnostics were assessed through visual inspections of fit to the data, convergence checks, delta Akaike Information Criterion (deltaAIC), standard error and confidence intervals of parameter estimates, examination of raw and one step ahead residuals, jitter and self-test analyses, retrospective patterns, Mohn's rho and biological credibility of estimates and derived quantities.

North

Data inputs to the base north model were the fixed gear total catch, proportions- and weight-at-age, the mobile gear total catch, proportions- and weight-at-age, the BDCM acoustic survey age aggregated biomass index, proportions-at-age and weight-at-age, the SG acoustic survey biomass index and annual maturity-at-age ogives. Coefficients of variation for aggregated catch

sources were 0.1 for the total catch of both fleets, 0.5 for the BDCM acoustic survey and 0.25 for the SG acoustic survey. Bias correction for NAA process error was not used.

In the base model, M was modeled as a fixed mean value of 0.3 and process error was modeled using random effects with AR1 correlation structure over years. The mean fixed gear fishery selectivity was modeled as age-specific and process error was modeled using an AR1 process over years. Parameters were freely estimated except for ages 8 and 9 that were fixed at 1 to avoid bound issues after examining sensitivity runs without fixed parameters. The mobile gear fishery selectivity was modeled as logistic curves in three time blocks (1978-1989; 1990-2005; 2006-2023). The BDCM acoustic survey selectivity was modeled as a logistic curve. The SG acoustic survey selectivity was modeled as fixed age-specific selectivity using the mean maturity-at-age schedule across regions and years.

Sensitivity analyses allowed deviations from the base model. M models explored estimating a mean M or fixing mean M to values other than 0.3 (0.2 or 0.4) using either IID or AR1 over years random effects, and using different age grouping. Selectivity models explored using a correlation structure over both years and ages (2DAR1) for random effects on the fixed gear fishery instead of an AR1 correlation over years. Predation models explored using predator abundance effects on M using either a random walk or an AR1 correlation over years to model the predator abundance. The predator effect was modeled as linear without year lag and shared on natural mortality estimates of the older age group only (7 to 11+). Lastly, gillnet CPUE-gam and CPUE-glm models explored adding the age aggregated CPUE index of biomass to the model and the estimation of time variation in the survey catchability using either a random walk or an AR1 process. SD parameters for CPUE q random effects over years were also fixed to different values in sensitivity runs.

Middle

Data inputs to the base middle model were the fixed gear total catch, proportions- and weights-at-age, the mobile gear total catch, proportions- and weights-at-age, the experimental gillnets survey aggregated biomass index and proportions-at-age, the SG acoustic survey biomass index and annual maturity-at-age ogives. Coefficients of variation for aggregated catch sources were 0.1 for the total catch of the fixed gear fleet and 0.3 for the total catch of the mobile fleet, 0.3 for the experimental nets survey and 0.5 for the SG acoustic survey. Bias correction for NAA process error was not used.

In the base model, M was modeled as a fixed mean value of 0.3 and process error was modeled using IID random effects over years. The mean fixed gear fishery selectivity was modeled as age-specific and process error was modeled using an AR1 process over years. Parameters for ages 9 and 10 were fixed at 1 to avoid bound issues after examining sensitivity runs without fixed parameters. The mobile gear fishery selectivity was modeled as logistic curves in three time blocks (1978-1989; 1990-2005; 2006-2023). The experimental nets survey selectivity was modeled as a logistic curve. The SG acoustic survey selectivity was modeled as fixed age-specific selectivity using the mean maturity-at-age schedule across regions and years.

The sensitivity analyses allowing deviations from the base middle model were identical to the north model sensitivity runs.

South

Data inputs to the base middle model were the fixed and mobile gear total catches, proportions- and weights-at-age, the experimental nets survey biomass index and proportions-at-age, the SG acoustic survey biomass index and annual maturity-at-age ogives. Coefficients of variation

for aggregated catch sources were 0.1 for each fleet, 1 for the experimental nets survey and 0.2 for the SG acoustic survey. Bias correction for NAA process error was not used.

In the base model, M was modeled as a fixed mean value of 0.3 and process error was modeled using random effects with AR1 correlation structure over years. The mean fixed gear fishery selectivity was modeled as age-specific and process error was modeled using an AR1 process over years. Parameters for ages 8 and 9 were fixed at 1 to avoid bound issues after examining sensitivity runs without fixed parameters. The experimental nets selectivity was modeled as a logistic curve. The SG acoustic survey selectivity was modeled as fixed age-specific selectivity using the mean maturity-at-age schedule across regions and years.

The sensitivity analyses allowing deviations from the base middle model were identical to the north model sensitivity runs. Additional sensitivity runs were performed by adding the RV survey biomass index and age composition, or allowing process error on the RV survey q.

RESULTS

NORTH

The two-fleet model converged and all parameters were estimable. It was selected as the initial base model configuration, which also included treating numbers-at-age in the first year as age-specific fixed effects, modelling process error in numbers-at-age as IID random effects across years and ages, and using a Dirichlet age-composition distribution that treated missing observations as random effects across all age composition datasets. These model assumptions were favored by AIC comparisons and diagnostic checks. The best-performing model from the sensitivity analyses is referred to as the north model, and all other configurations in the following sensitivity runs were compared against it.

The first set of sensitivity runs focused on alternative formulations of natural mortality (Appendix 1). In the north model, mean M was fixed at 0.3 across ages and years, and process error on M was modeled as random effects following an AR1 correlation structure over years and shared across ages 7–11+. A sensitivity run not allowing process error on M had higher AIC and more bias in retrospective analysis, even if NAA process error was allowed as IID random effects. Sensitivity runs that allowed process error on M but instead fixed mean M to 0.2 or 0.4 either failed to converge or resulted in higher AIC values than the north model. Models estimating mean M also failed to converge when using an AR1 process for older-age M process error, or produced strong retrospective patterns when process error was assumed to be IID random effects over years.

For the fixed-gear fishery, the north model used age-specific selectivity with an AR1 correlation structure on annual random effects. A more complex 2DAR1 correlation structure produced identical AIC and Mohn's rho values; therefore, the simpler AR1 model was retained.

Models allowing predator abundance effects on M converged but the covariates effects on M had high relative error. Of the two sensitivity runs for predator abundance effects, the model using a random walk to model predator abundance had lower AIC than the model using an AR1 correlation structure. However, both models effects on M had high relative error. The best performing model estimated a mean Grey seal effect β = 0.419 (SE = 0.168; 95% CI: 0.091 – 0.748) and a mean ABFT effect β = -0.022 (SE = 0.153; 95% CI: -0.321 – 0.277). The resulting M at age estimates were almost identical to M estimates without the predation effect. Consequently, the models including predator abundance effects were not retained.

Lastly, models that incorporated the gillnet CPUE biomass index as an additional biomass index performed poorly overall. In a model where catchability process error was modeled as AR1 random effects, the estimated process error SD was $\sigma = 0.25$ (SE = 0.054; 95% CI: 0.122–0.345) and the autocorrelation was $\rho = 0.985$ (SE = 0.021; 95% CI: 0.783–0.999). The corresponding annual q estimates spanned from 0.25 to 3.49, an increase of 1296%. By comparison, the ADMB SCA model with priors and likelihood penalties on q produced q estimates that varied by only about 100%. To assess whether restricting q deviations in the state-space models would improve performance, models were run with fixed AR1 SD values or assuming no process error on q (i.e., stationary q). None of these models converge. To achieve convergence, the BDCM acoustic survey selectivity parameters had to be fixed. The sensitivity run using the CPUE-glm and an AR1 process on q random effects generated SD and correlation estimates highly similar to the CPUE-gam models, and a similar rate of increase in q estimates to the CPUE-gam over time (a 1112% increase).

The assumptions about the catchability process error, either the correlation structure or the magnitude of the SD, had a strong influence on the estimated spawning stock biomass (SSB) (Figure 1). Models without process error on CPUE q or models restricting the q deviations generated large differences in SSB dynamics compared to a model estimating q SD and correlation or a model without the CPUE index. The model fit without the CPUE index and the model fit without restrictions on q process error yielding almost identical SSB. This implies in order to use the include the CPUE index to the stock assessment model, q process error has to be constrained using subjective values in order to achieve converge, yielding highly different stock dynamics. Hence, the introduced subjectivity is dictating the stock dynamics.

To investigate the biomass-CPUE relationship, GAM models were fit to the CPUE biomass indices and SSB estimates from models that did not use the CPUE index in model fitting. The expected relationship is a power curve if the CPUE is hyperstable to SSB variations, or a linear relationship is the CPUE is directly proportional to SSB (Harley et al. 2001). For the CPUE-glm, there was no relationship to SSB (Figure 2). For the CPUE-gam, the relationship was opposite to the expectation, where CPUE increased as SSB decreased. Consequently, the CPUE-glm can be considered to be uninformative.

Given the uncertainty in the CPUE index data itself due to high sensibility to effort assumptions (Maynard et al In press) and the poor performance of models estimating process error in CPUE q, these model structures were not retained.

Over all models tested, the base model performed best and was used to estimate the stock states and processes described below. The maximum gradient was 9.25e-12. The jitter analysis produced negative log-likelihoods, SSB and fishing mortality (F) estimates that were identical to the base model. The self-simulation study lower 2.5% and upper 97.5% intervals for both SSB and F covered zero in all years, indicating that the simulated samples produced estimates that were similar to the base estimates.

The base model fit to age composition data was good for the fixed gear fishery with small randomly distributed residuals and one-step-ahead (OSA) residuals without patterns by year, age, or cohort (Figures 4 and 5). The mobile gear fishery displayed some groupings of larger positive residuals at ages 3 and 4 and larger negative residuals at ages 5 and 6 (Figure 6). OSA residuals by year, age, or cohort were without major patterns (Figure 7). The BDCM acoustic survey age composition displayed larger residuals at ages 2 to 4 but overall, residuals were without apparent blocking (Figure 8). OSA residuals by year, age, or cohort were without patterns (Figure 9).

Model fit to both the fixed gear and mobile gear catch was acceptable (Figures 10 and 11). Model fit to the BDCM acoustic survey was acceptable considering the high interannual

variations in total biomass from this survey (Figure 12). The largest residuals occurred in two groups of years; in the late 1990s where model estimates were lower than observations and in the late 2000s when model estimates were higher than observations. However the log-scale standardized residuals were all within +/- 2. Model fit to the SG acoustic survey was acceptable considering the high interannual variations in total biomass from this survey (Figure 13). The confidence bands of model fit to aggregated catch and indices included zero in all years and showed no systematic trend in model fit over time (Figure 14).

The retrospective analysis revealed moderate bias in mean F (Figure 15), with a Mohn's ρ of 0.151. The initial peel exhibited underestimation (negative bias), whereas the subsequent three peels showed consistent overestimation (positive bias). Recruitment retrospective patterns were similar (Figure 16), with a Mohn's ρ of -0.156. SSB displayed the least retrospective pattern (Figure 17), with small deviations and a Mohn's ρ of -0.073. Overall, while mean F and recruitment shows modest systematic bias, SSB estimates are robust to terminal-year omissions, supporting model reliability.

Overall, model parameters were well estimated, with acceptable SE and CI (Table 1). The SD in NAA at age 2 random effects was estimated at 0.42 and estimated at 0.08 for ages 3 to 11+. Accordingly, deviations in NAA at ages 3 to 11+ were small and showed no major patterns (Figure 18).

Catchability estimates were reliably estimated, with BDCM and SG surveys q (0.34 and 0.23, respectively) showing tight SE and CI.

Fixed gear mean selectivity was dome-shaped across ages (Figure 19), with precise estimation for the youngest and oldest ages but moderate uncertainty at age 5. Ages 2 and 3 had very low mean selectivity with relatively tight confidence intervals (CI) and small standard errors (SE). Ages 8 and 9 were fixed at 0.99 to avoid bounding issues, and representing full selection. Age 10 remained highly selected and precisely estimated, whereas age 11+ mean selection dropped to 0.53 with high uncertainty. This mean selectivity curve represents an average over the full time series; however, substantial year-to-year variation in selectivity-at-age driven by changes in fish size-at-age and gillnet mesh configurations is explicitly modeled via random effects (Figure 20). Accordingly, the estimated AR1 process parameters showed high variation ($\sigma = 0.78$) and correlation ($\rho = 0.90$) and had low relative error. The general trend in the fixed gear selectivity was a decline in fish ages 2 to 7 selectivity. Fish ages 10 and 11+ also declined in selectivity as their abundance in population declined over time.

The mobile gear logistic selectivity (Figure 21) and the BDCM acoustic survey logistic selectivity (Figure 22) parameters were precisely estimated. The SG acoustic survey selectivity was fixed to the maturity-at-age schedule (Figure 23).

The natural mortality AR1 process parameters showed moderate precision (relative error approximately 40% for both σ and ρ), with SEs and confidence intervals indicating acceptable but not strong parameter identification. The mean natural mortality was fixed at 0.3, and point estimates for the age group 7 to 11+ ranged between 0.17 and 0.86 (Figure 24). M remained low between 1978 and 2010 except for a small increase in the late 1990s. M then increased rapidly to reach a maximum in 2014 and slowly decrease until 2021. M removals from the 7-11+ age group ranged between 1,591 and 58,749 tonnes (t) annually, with a mean of 18,625 t over the time series.

The uncertainty in SSB and F trends was low (Figure 25). Two periods of higher SSB occurred, in the mid- 1980s and around 2010. SSB declined between 2010 and 2022. Fully selected F was near 1.0 for the first three years of the time series then rapidly declined. F increased slowly to approximately 0.5 in 2000, then averaged 0.3 between 2014 and 2022. F was 0.13 in 2023.

Table 1: Parameter estimates, standard errors, and confidence intervals for the north model, – indicates fixed parameters.

Parameter	Estimate	Std. Error	95% CI lower	95% CI upper
NAA σ (age 2)	0.42	0.06	0.32	0.54
NAA σ (ages 3-11+)	0.08	0.03	0.04	0.16
BDCM fully selected q	0.34	0.05	0.25	0.46
SG fully selected q	0.23	0.04	0.16	0.33
Fixed gear mean Selectivity for age 2 (Block 1)	3.85E-04	3.83E-04	5.47E10-5	0.00
Fixed gear mean Selectivity for age 3 (Block 1)	0.01	0.01	9.35E10-4	0.04
Fixed gear mean Selectivity for age 4 (Block 1)	0.11	0.09	0.02	0.44
Fixed gear mean Selectivity for age 5 (Block 1)	0.46	0.24	0.11	0.85
Fixed gear mean Selectivity for age 6 (Block 1)	0.81	0.15	0.39	0.97
Fixed gear mean Selectivity for age 7 (Block 1)	0.96	0.04	0.78	0.99
Fixed gear mean Selectivity for age 8 (Block 1)	0.99	–	–	–
Fixed gear mean Selectivity for age 9 (Block 1)	0.99	–	–	–
Fixed gear mean Selectivity for age 10 (Block 1)	0.95	0.05	0.72	0.99
Fixed gear mean Selectivity for age 11+ (Block 1)	0.53	0.25	0.14	0.89
Mobile gear a50 (Block 2)	3.97	0.32	3.36	4.60
Mobile gear 1/slope (increasing) (Block 2)	0.87	0.09	0.71	1.06
Mobile gear a50 (Block 3)	3.54	0.20	3.15	3.94
Mobile gear 1/slope (increasing) (Block 3)	0.68	0.06	0.58	0.81
Mobile gear a50 (Block 4)	3.23	0.18	2.89	3.58
Mobile gear 1/slope (increasing) (Block 4)	0.69	0.06	0.57	0.82
BDCM a50 (Block 5)	1.30	0.10	1.12	1.50
BDCM 1/slope (increasing) (Block 5)	0.42	0.09	0.27	0.64
SG selectivity for age 2 (Block 6)	0.03	–	–	–
SG selectivity for age 3 (Block 6)	0.48	–	–	–
SG selectivity for age 4 (Block 6)	0.96	–	–	–
SG selectivity for age 5 (Block 6)	1.00	–	–	–
SG selectivity for age 6 (Block 6)	1.00	–	–	–
SG selectivity for age 7 (Block 6)	1.00	–	–	–
SG selectivity for age 8 (Block 6)	1.00	–	–	–
SG selectivity for age 9 (Block 6)	1.00	–	–	–
SG selectivity for age 10 (Block 6)	1.00	–	–	–
SG selectivity for age 11+ (Block 6)	1.00	–	–	–
Fixed gear selectivity RE σ (Block 1)	0.78	0.11	0.60	1.02
Fixed gear selectivity RE AR1 ρ (year) (Block 1)	0.90	0.07	0.66	0.97
Fixed gear age comp, Dirichlet: dispersion (ϕ)	215.32	24.86	171.72	270.00
Mobile gear age comp, Dirichlet: dispersion (ϕ)	46.84	3.72	40.09	54.72
BDCM age comp, Dirichlet: dispersion (ϕ)	31.73	3.35	25.80	39.01
M RE σ	0.36	0.16	0.15	0.86
M RE AR1 ρ (year)	0.65	0.28	−0.170	0.94

The relationship between recruitment and SSB did not seem to suggest a stock recruit relationship could describe the recruitment dynamics, as supported by the higher AIC of models using Ricker or Beverton-Holt stock recruit models (Figure 26). Recruitment was highly variable but generally higher between 1978 and 1990 and lower between 1991 and 2004 (Figure 27). Multiple years of high recruitment occurred between 2005 and 2010. Recruitment then declined to low levels and increased slightly in recent years. Estimated SSB in 2023 was 54,232 t (95% CI: 35,240 – 83,458).

MIDDLE

The two-fleet model converged and all parameters were estimable. It was selected as the initial base model configuration, which also included treating numbers-at-age in the first year as age-specific fixed effects, allowing process error in numbers-at-age 3 to 11+, and employing a Dirichlet age-composition distribution that treated missing observations as random effects across all age composition datasets. These model assumptions were favored by AIC comparisons and diagnostic checks. The best-performing model from the sensitivity analyses is referred to as the middle model, and all other configurations are compared against it.

The first set of sensitivity runs focused on alternative formulations of natural mortality (Appendix 1). In the middle model, mean M was fixed at 0.3 across ages and years. Negative deviations in numbers at ages 9 to 11+ from a model allowing process error in NAA as IID random effects over years suggested that process error on M at ages 9 to 11+ would likely be appropriate. Sensitivity runs using different age groupings for process error on M confirmed that fixing M for ages 2-8 and allowing process error on M for an age group of ages 9 to 11+ had the lowest bias in self-tests. To help convergence, the natural mortality process error sigma parameter was fixed to 0.35, based on estimated values in the north and south models. M process error was modeled as AR1 random effects over years, shared across ages 9–11+. A sensitivity run not allowing process error on M had higher AIC and more bias in retrospective analysis. A sensitivity run not allowing process error on M but allowing IID random effects on NAA had lower AIC than the middle model but more bias in the retrospective analysis of SSB. Sensitivity runs estimating mean M failed to converge with IID or AR1 random effects over years. Models that fixed mean M to 0.4 were not favored by AIC over the middle model. The model that fixed M to 0.2 with IID random effects had a deltaAIC under 2.0 compared to the middle model but had worse retrospective patterns and SD had high relative error. The model that fixed M to 0.2 with AR1 random effects over years had a negative deltaAIC compared to the middle model but estimated M went to 0, which is not biologically credible. Finally the model that fixed M to 0.3 but used AR1 correlated random effects over years had higher AIC than the middle model.

The sensitivity runs targeting selectivity, predator effects and the CPUE index with process error on catchability, described for the north model, were also used for the south model. These runs provided similarly limited support for alternative model structures. Selectivity sensitivity models failed to converge, predator effects were estimated with low precision and the CPUE catchability process error parameters either had high relative error or led to non-convergence when SD was fixed, when process error was not allowed, or when the CPUE-glm index was used.

Over all models tested, the base model performed best and was used to estimate the stock states and processes described below. The maximum gradient was 5.58e-08. The jitter analysis produced negative log-likelihoods, SSB and F estimates that were identical to the base model. The self-simulation study lower 2.5% and upper 97.5% intervals for both SSB and F covered zero in all years, indicating that the simulated samples produced estimates that were similar to the base estimates.

The base model fit to age composition data was good for the fixed gear fishery with small randomly distributed residuals and OSA residuals without patterns by year, age, or cohort (Figures 28 and 29). The mobile gear fishery displayed some groupings of larger positive residuals at ages 4 and 5 and larger negative residuals at ages 6 and 7 (Figure 30). OSA residuals by year, age, or cohort were without major patterns (Figure 31). The experimental nets survey age composition displayed small and random residuals (Figure 32). OSA residuals by year, age, or cohort were without patterns, while residuals in function of the observed value were generally larger and positive as observed values increased (Figure 33).

Model fit to both the fixed gear and mobile gear catch was acceptable (Figures 34 and 35). Two larger residuals occurred in years 1979 and 1980 in the mobile gear catch. Model fit to the experimental net survey index was poor as this survey contains little information on aggregated biomass trends is not expected to be a reliable indicator the population biomass (Turcotte et al. 2021a). Model fit to the SG acoustic survey was acceptable considering the high interannual variations in total biomass from this survey (Figure 37). The confidence bands of model fit to aggregated catch and indices included zero in all years and showed no systematic trend in model fit over time (Figure 38).

The retrospective analysis revealed moderate bias in mean F (Figure 39), with a Mohn's ρ of 0.263. The first peel exhibited small underestimation (negative bias), while the following three were positive. Recruitment retrospective patterns were underestimating recruitment but bias was larger (Figure 40), with a Mohn's ρ of -0.584. SSB displayed the least retrospective pattern (Figure 41), with a positive deviation followed by three negative and a Mohn's ρ of -0.201. F and SSB estimates were moderately robust to terminal-year omissions, supporting model reliability. The high bias in recruitment was expected as there is little information about abundance at age two and three in the catch or experimental nets age composition data, as both occur on spawning grounds.

Overall, model parameters were well estimated, with low SE and relative error, and narrow CI (Table A2). The SDs in NAA random effects were estimated at 0.36 for age 2 and 0.191 for ages 3+ (Figure 42). Catchability estimates were reliably estimated, with the experimental nets and SG surveys q (0.034 and 0.017, respectively) showing low SE.

Fixed gear mean selectivity was dome-shaped across ages (Figure 43). The uncertainty in mean estimates was high at young ages where selectivity was near zero. The uncertainty decreased as selectivity increased at older ages. The relative SEs were 0.83 for age 5, 0.30 for age 6 and 0.08 for age 7. Ages 8 to 10 selectivity were fixed at 0.99, representing full selection. Age 11+ mean selection dropped to 0.89 with low uncertainty (Residual Standard Error (RSE) = 0.16). This mean selectivity curve represents an average over the full time series; however, substantial year-to-year variation in selectivity-at-age driven by changes in fish size-at-age and gillnet mesh configurations is explicitly modeled via random effects (Figure 44). Accordingly, the estimated AR1 process parameters showed high variation (σ = 0.642) and correlation (ρ = 0.948) and were well estimated.

The mobile gear logistic selectivity (Figure 45) and the experimental nets survey age specific selectivity (Figure 46) parameters were precisely estimated. The SG acoustic survey selectivity was fixed to the maturity-at-age schedule (Figure 47).

Table 2: Parameter estimates, standard errors, and confidence intervals for the middle model, – indicates fixed parameters.

Parameter	Estimate	Std. Error	95% CI lower	95% CI upper
NAA σ (age 2)	0.363	0.070	0.249	0.531
NAA σ (ages 3-11+)	0.191	0.029	0.142	0.256
Experimental nets fully selected q	0.034	0.008	0.021	0.055
SG fully selected q	0.017	0.005	0.010	0.029
Fixed gear Mean Selectivity for age 2 (Block 1)	0.001	0.001	6.768×10−5	0.015
Fixed gear Mean Selectivity for age 3 (Block 1)	0.008	0.011	5.530×10−4	0.107
Fixed gear Mean Selectivity for age 4 (Block 1)	0.096	0.119	0.007	0.610
Fixed gear Mean Selectivity for age 5 (Block 1)	0.399	0.329	0.043	0.907
Fixed gear Mean Selectivity for age 6 (Block 1)	0.780	0.236	0.193	0.981
Fixed gear Mean Selectivity for age 7 (Block 1)	0.937	0.082	0.497	0.996
Fixed gear Mean Selectivity for age 8 (Block 1)	1.000	–	–	–
Fixed gear Mean Selectivity for age 9 (Block 1)	1.000	–	–	–
Fixed gear Mean Selectivity for age 10 (Block 1)	1.000	–	–	–
Fixed gear Mean Selectivity for age 11+ (Block 1)	0.893	0.147	0.292	0.994
Mobile gear a50 (Block 2)	3.733	0.257	3.244	4.249
Mobile gear 1/slope (increasing) (Block 2)	0.693	0.071	0.566	0.846
Mobile gear a50 (Block 3)	5.786	0.266	5.259	6.297
Mobile gear 1/slope (increasing) (Block 3)	0.922	0.056	0.818	1.039
Mobile gear a50 (Block 4)	4.699	0.334	4.052	5.355
Mobile gear 1/slope (increasing) (Block 4)	0.827	0.088	0.670	1.017
Experimental nets Selectivity for age 2 (Block 5)	0.002	7.000×10−4	0.001	0.004
Experimental nets Selectivity for age 3 (Block 5)	0.027	0.005	0.018	0.039
Experimental nets Selectivity for age 4 (Block 5)	0.171	0.025	0.127	0.227
Experimental nets Selectivity for age 5 (Block 5)	0.340	0.047	0.255	0.438
Experimental nets Selectivity for age 6 (Block 5)	0.647	0.080	0.481	0.784
Experimental nets Selectivity for age 7 (Block 5)	0.914	0.093	0.508	0.991
Experimental nets Selectivity for age 8 (Block 5)	1.000	–	–	–
Experimental nets Selectivity for age 9 (Block 5)	1.000	–	–	–
Experimental nets Selectivity for age 10 (Block 5)	1.000	–	–	–
Experimental nets Selectivity for age 11+ (Block 5)	0.621	0.218	0.210	0.910
Fixed gear Selectivity RE σ (Block 1)	0.642	0.104	0.468	0.881
Fixed gear Selectivity RE AR1 ρ (year) (Block 1)	0.948	0.049	0.697	0.992
Fixed gear age comp, Dirichlet: dispersion (ϕ)	188.490	30.900	136.694	259.914
Mobile gear age comp, Dirichlet: dispersion (ϕ)	31.080	2.630	26.331	36.686
Experimental nets age comp, Dirichlet: dispersion (ϕ)	77.457	12.485	56.476	106.233
M RE ρ (year)	0.897	0.078	0.586	0.978

The natural mortality process error sigma parameter was fixed to 0.35 while the correlation parameter (rho) was estimated at 0.897 with good precision (SD RSE = 9%), with SEs and confidence intervals indicating acceptable parameter identification. The mean natural mortality was fixed at 0.3, and values for the age group 9 to 11+ ranged between 0.19 and 1.38

(Figure 48). M remained low between 1978 and 2010. M then increased rapidly to reach a maximum in 2016 and slowly decrease until 2021. M removals from the 9-11+ age group ranged between 2,369 and 26,675 t annually, with a mean of 13,040 t over the time series.

The uncertainty in SSB and F trends was low (Figure 49). Two periods of higher SSB occurred, in the mid- 1980s and around 2010. SSB declined between 2010 and 2020. SSB increased every year between 2021 and 2023. Fully selected F was high for the first three years of the time series then rapidly declined. F was low and stable until the mid 1990s and averaged 0.42 between 1997 and 2017. F declined to lower levels between 2018 and 2022. F was 0.09 in 2023.

The relationship between recruitment and SSB did not seem to suggest a stock recruit relationship could describe the recruitment dynamics, as supported by the higher AIC of models using Ricker or Beverton-Holt stock recruit models (Figure 50). Recruitment was highly variable but generally higher between 1980 and 1990 and lower between 1991 and 2004 (Figure 27). Multiple years of high recruitment occurred between 2005 and 2010. Recruitment then declined to low levels and increased rapidly to high levels in recent years.

SOUTH

The two-fleet model converged and all parameters were estimable. It was selected as the initial base model configuration, which also included treating numbers-at-age in the first year as age-specific fixed effects, not allowing process error in numbers-at-ages 3 to 11+, and employing a Dirichlet age-composition distribution that treated missing observations as random effects for the all age composition data. These model assumptions were favored by AIC comparisons and diagnostic checks. The best-performing model from the sensitivity analyses is referred to as the south model, and all other configurations are compared against it.

The first set of sensitivity runs focused on alternative formulations of natural mortality (Appendix 1). In the south model, mean M was fixed at 0.3 across ages and years, and process error was modeled as IID random effects shared across ages 7–11+. Negative deviations in numbers at ages 9 to 11+ from a model allowing process error in NAA as IID random effects over years suggested that process error on M at ages 9 to 11+ would likely be appropriate. Sensitivity runs using different age groupings for process error on M confirmed that fixing M for ages 2-8 and allowing process error on M for an age group of ages 9 to 11+ had the lowest bias in self-tests. A model fixing mean M at 0.3 without process error failed to converge and estimate all parameters. A model fixing mean M at 0.3 without process error but allowing process error on NAA as IID random effects converged and all parameters were estimable. This model had lower AIC than the base model but had more bias in retrospective analyses of SSB and F. Most M sensitivity runs performed well, indicating greater uncertainty about the mean M value (0.2 to 0.4) in this region. However, one model fixing mean M to 0.2 and allowing IID random effects over years could not estimate the SD. Hence, mean M values lower than 0.3 are unlikely. Other models using 0.3 to 0.5 as fixed M and either AR1 correlated or IID random effects converged and successfully estimated the SD. Consequently, for consistency with other regions and considerations of population quantities scaling, 0.3 was retained as the fixed mean M value, and AR1 random effects over years were used as in the base model.

Sensitivity runs were performed using the RV survey biomass and age composition data, allowing process error on the RV survey catchability or not (see Turcotte et al. In press). These models performed poorly in most diagnostic tests. The model converged but M parameters could not be estimated and q was estimated over an improbable range of values, presumably taking all the variation in data, a behavior similar to the CPUE q process error results described

for the North model. No prior information was available to constrain the process error on RV survey q and is not a desirable model configuration. These models were not retained.

The sensitivity runs targeting predator effects and the CPUE index with process error on catchability, described for the north model, were also used for the south model. These runs provided similarly limited support for alternative model structures. Predator effects were estimated with low precision and the CPUE catchability process error parameters were either had high relative error or led to non-convergence when SD was fixed, when process error was not allowed, or when the CPUE-glm index was used.

Over all models tested, the base model performed best and was used to estimate the stock states and processes described below. The maximum gradient was 9.21e-11. The jitter analysis produced negative log-likelihoods, SSB and F estimates that were identical to the base model. The self-simulation study generated parameter estimates that were highly similar between runs. However there was bias in the last decade of F and SSB estimates. As in the retrospective analysis (see below), adding years of data up to 2025 got rid of the bias (not shown).

The base model fit to age composition data was good for the fixed and mobile gear fleets with small randomly distributed residuals and OSA residuals without patterns by year, age, or cohort (Figures 52-55). The experimental nets age composition displayed larger residuals at ages 3 and 4 (Figure 56). Larger residuals were positive and occurred only sporadically over the time-series, and not in recent years. OSA residuals by year, age, or cohort were without patterns (Figure 57).

Model fit to the fixed and mobile gear catch was acceptable (Figures 58 and 59). Residuals were small early in the time series and increased over time. The residuals seemed slightly correlated over years with group of years alternating between negative and positive residuals. Overall the scale of residuals was acceptable. Model fit to the experimental nets survey index was poor, as expected using a very large CV (Figure 60). The survey is not expected to track changes in population biomass. Model fit to the SG acoustic survey was good (Figure 61). The confidence bands of model fit to aggregated catch and indices included zero in all years and showed no systematic trend in model fit over time (Figure 62).

The retrospective analysis revealed important bias in mean F (Figure 63), with a Mohn's ρ of 0.89. Recruitment retrospective patterns first peels were small but the third and fourth were largely negative (Figure 64), with a Mohn's ρ of -0.415. SSB displayed similar retrospective patterns (Figure 65), with a Mohn's ρ of -0.337. The important patterns reflect the high variation in the SG survey and the associated population scale changing when years of data are peeled. This is likely due to the variable nature of acoustic surveys, and the short time series for this survey in particular. However, when years of data are added to 2025, the survey signal stabilizes and the retrospective patterns are small (results not shown).

Overall, model parameters were well estimated, with acceptable SE and CI (Table 3). The SD in NAA for age 2 random effects was estimated at $\sigma = 0.36$ (Figure 66). Catchability estimates were reliably estimated, with the experimental nets and SG surveys q (0.60 and 0.24, respectively) showing tight SE and CI.

The fixed gear mean selectivity (Figure 67) parameters had large confidence intervals at all ages. Ages 2 and 3 had very low mean selectivity, and men selectivity-at-age increase until age 8. Ages 8 to 10 were fixed at 1, representing full selection. Age 11+ mean selectivity was estimated at 0.947. This mean selectivity curve represents an average over the full time series; however, substantial year-to-year variation in selectivity-at-age driven by the annual proportion of fixed gear catch, changes in fish size-at-age and gillnet mesh configurations. This variation is explicitly modeled via random effects (Figure 68). Accordingly, the estimated AR1 process

parameters showed high variation ($\sigma = 1.081$) and correlation ($\rho = 0.858$). The AR1 process parameters were well estimated.

The mobile gear (Figure 69) and experimental nets survey (Figure 70) logistic selectivity parameters were precisely estimated. The SG acoustic survey selectivity was fixed to the maturity-at-age schedule (Figure 71).

The natural mortality process error parameter showed good precision (SD relative standard error 36%), with SEs and confidence intervals indicating acceptable parameter identification. The mean natural mortality was fixed at 0.3, and values for the age group 9 to 11+ ranged between 0.22 and 0.79 (Figure 72). M gradually increased, with variation, from 0.24 in 1978 to 0.47 in 2000. In 2010, M rapidly increased to 0.79 in 2016. M stayed high until 2020 and declined again. Terminal year M was 0.39. M removals from the 9-11+ age group ranged between 990 and 50,471 t annually, with a mean of 16,743 t over the time series.

The CV in SSB and F remained under 0.4 over the time series. Two periods of higher SSB occurred, in the mid- 1980s and around 2010 (Figure 73). SSB declined between 2010 and 2020. SSB increased every year between 2021 and 2023. Fully selected F decreased rapidly between 1978 and 1981 and remained low until the early 1990s. F then increased and remained between 0.2 and 0.3 until 2010 where it decreased again. F remained low until the terminal year. Fbar (mean F for ages 4 to 8) followed a dynamic to fully selected F but at a lower scale.

The relationship between recruitment and SSB did not seem to suggest a stock recruit relationship could describe the recruitment dynamics, as supported by the higher AIC of models using Ricker or Beverton-Holt stock recruit models (Figure 74). Recruitment was highly variable (Figure 75). Multiple years of high and low recruitment occurred between 1978 and 2023.

Table 3: Parameter estimates, standard errors, and confidence intervals for the south model, – indicates fixed parameters.

Parameter	Estimate	Std. Error	95% CI lower	95% CI upper
NAA σ (age 2)	0.357	0.044	0.28	0.455
SG fully selected q	0.06	0.017	0.035	0.105
Experimental nets fully selected q	0.24	0.066	0.139	0.412
Fixed gear mean Selectivity for age 2 (Block 1)	7.929×10−4	8.266×10−4	1.027×10−4	0.006
Fixed gear mean Selectivity for age 3 (Block 1)	0.011	0.011	0.001	0.078
Fixed gear mean Selectivity for age 4 (Block 1)	0.136	0.12	0.021	0.54
Fixed gear mean Selectivity for age 5 (Block 1)	0.485	0.257	0.111	0.876
Fixed gear mean Selectivity for age 6 (Block 1)	0.832	0.145	0.395	0.974
Fixed gear mean Selectivity for age 7 (Block 1)	0.971	0.03	0.808	0.996
Fixed gear mean Selectivity for age 8 (Block 1)	1	–	–	–
Fixed gear mean Selectivity for age 9 (Block 1)	1	–	–	–
Fixed gear mean Selectivity for age 10 (Block 1)	1	–	–	–
Fixed gear mean Selectivity for age 11+ (Block 1)	0.947	0.059	0.644	0.994
Mobile a50 (Block 2)	3.015	0.282	2.492	3.594
Mobile 1/slope (increasing) (Block 2)	0.656	0.089	0.501	0.853
Mobile a50 (Block 3)	4.092	0.255	3.603	4.598
Mobile 1/slope (increasing) (Block 3)	0.786	0.066	0.666	0.926
Mobile a50 (Block 4)	4.428	0.217	4.008	4.856
Mobile 1/slope (increasing) (Block 4)	0.701	0.063	0.588	0.834
SG Selectivity for age 2 (Block 2)	0.030	–	–	–
SG Selectivity for age 3 (Block 2)	0.480	–	–	–
SG Selectivity for age 4 (Block 2)	0.960	–	–	–
SG Selectivity for age 5 (Block 2)	1.000	–	–	–
SG Selectivity for age 6 (Block 2)	1.000	–	–	–
SG Selectivity for age 7 (Block 2)	1.000	–	–	–
SG Selectivity for age 8 (Block 2)	1.000	–	–	–
SG Selectivity for age 9 (Block 2)	1.000	–	–	–
SG Selectivity for age 10 (Block 2)	1.000	–	–	–
SG Selectivity for age 11+ (Block 2)	1.000	–	–	–
Experimental nets a50 (Block 3)	4.427	0.22	4.002	4.862
Experimental nets 1/slope (increasing) (Block 3)	0.785	0.06	0.675	0.91
Fixed gear Selectivity RE σ (Block 1)	1.081	0.138	0.842	1.388
Fixed gear Selectivity RE AR1 ρ (year) (Block 1)	0.858	0.087	0.563	0.959
Fixed gear age comp, Dirichlet: dispersion (ϕ)	110.397	13.139	87.429	139.4
Mobile age comp, Dirichlet: dispersion (ϕ)	29.179	2.293	25.014	34.037
Experimental nets age comp, Dirichlet: dispersion (ϕ)	21.264	2.27	17.25	26.213
M RE σ	0.565	0.134	0.356	0.898
M RE AR1 ρ (year)	0.893	0.093	0.49	0.981

DISCUSSION

MODELS PERFORMANCE

Models fit to the age composition data was good overall. Raw residuals indicated a very good fit for the fishery catch-at-age data. While still acceptable, residuals were generally larger for mobile fleets, which was expected as they have lower sample sizes and rely on lesser quality samples compared to fixed gear catch (Turcotte et al. In press). Residuals for the survey age compositions were generally larger, which is expected given the smaller sample sizes relative to the commercial fishery. Residuals were also typically higher at younger ages, reflecting the limited information available on young fish in the survey data. OSA residuals were broadly satisfactory, with only minor tail deviations that are not concerning and should not impair the model's ability to support inference (Stewart et al. 2025). It is also important to note that normality tests applied to OSA residuals may fail to detect certain forms of model mis-specification (Babyn and Cadigan 2025).

Fit to the age-aggregated indices was acceptable across all models. Despite substantial year-to-year variability in most indices, model fits captured the overall trends well. The regional SG acoustic surveys were included in all models. Given their high spatial and temporal coverage of herring aggregations on the spawning grounds, these surveys are expected to provide reliable indicators of both SSB and the biomass available to the fishery, which also operates on the spawning grounds. The use of additional age-aggregated surveys varied among regions, but this approach should better reflect regional population dynamics compared with previous assessments, which shared RV and BDCM survey data across regions. Model fit to experimental nets biomass indices were poor, which was intentionally done by using large CVs for these surveys as their not expected to reflect population abundance. These gillnet surveys suffer from the same issues identified for the CPUE indices and the apparent impossibility to estimate the required time variation in the catchability.

Retrospective patterns were acceptable overall and did not suggest model mis-specification. The largest biases occurred in the Middle region recruitment retrospective patterns, and in the South model SSB and F, although resolved when additional years of data were provided to the model. Retrospective patterns are often misinterpreted as a measure of estimation bias that can be corrected. Retrospective patterns can be an informative diagnostic to identify and confront model misspecification, and if retrospective patterns cannot be reduced with respecified models, they can be communicated as measure of uncertainty for consideration in the precautionary management (Cadrin 2025). Here all sensitivity runs for the Middle model displayed important retrospective patterns in recruitment, but not in SSB and F. The lack of young fish information in catch and survey data is likely the source of the patterns.

Jittering initial parameter values yielded identical model results in the three subpopulations. Self- test simulation overall generated little to no bias, except in the South model SSB and F estimates. Again, the bias was nil when additional years of data were provided to the model (as in the retrospective patterns). It is likely that source of the bias in both diagnostics is the noisy SG acoustic survey data which the model heavily relies on.

The North and Middle models were able to estimate process error in selectivity, M and NAA. The NAA deviations were small and approximately IID, and the estimated deviations across all processes appeared reasonable. Of the three regions, the North had the highest data quality based on a recent review (Turcotte et al. In press), which likely enabled the successful estimation of these processes. In contrast, the South model could estimate process error in either M or NAA, but not both simultaneously. The estimated sigma in NAA deviation was large and was likely taking a lot the variation in the data, as other processes and parameters were

poorly estimated when NAA process error was allowed. This outcome is expected because process errors in one component can alias those in others (Stock and Miller 2021), and NAA deviations may arise from fish movement or changes in selectivity or M (Gudmundsson and Gunnlaugsson 2012). Identifying the appropriate model configuration is critical because misattributing the source of process error can bias key management quantities (Szuwalski et al. 2018). For example, incorrectly assigning natural mortality deviations to other processes, or vice versa, has been shown to generate substantial bias in management metrics (Li et al. 2024). Given the evidence for temporal changes in M and the consistency across regional models, process error was preferentially assigned to M rather than NAA. This choice also helped reduce retrospective patterns in the South model. In the Middle model, only selectivity and M process error was able to be estimated.

Diagnostics supported estimating process error in M across all models. However, predator abundance did not provide sufficient information to reliably estimate its assumed effects on M. This is unsurprising, as numerous processes intervene between predator abundance and actual prey consumption, making abundance alone an unreliable predictor of prey mortality. Even when incorporating distributional overlap, diet composition, and prey selectivity, estimating predator consumption and its effect on sGSL herring M remains challenging, and does not meaningfully alter estimates of stock biomass or M (Rossi and Turcotte In press).

Independent modeling showed that trends and magnitudes in time-varying natural mortality for older ages are nearly identical whether estimated solely from herring age composition data or by explicitly incorporating predator consumption as removals (Rossi and Turcotte In press). This consistency indicates that herring age composition alone likely provides sufficient signal to reliably estimate predation-driven changes in natural mortality. Time variation in natural mortality is notoriously difficult to estimate. However, the estimation here was satisfactory and the estimated M removals matched expectations from available predator consumption information. Excluding the mid-1980s peak value of approximately 100,000 t, which was largely driven by cod consumption, estimates by Benoît and Rail 2016 suggest annual predator consumption of herring between 25,000 and 40,000 t in most years post 1990. This was excluding whales, which may represent an important additional source of mortality. Hence, the here estimated M removals for the age groups were M was time varying and reached high values matched the scale of possible herring consumption by predators. It should be noted that uncertainty is substantial for both estimated M removals and predator consumption. Nevertheless, the order of magnitude—ranging from tens of thousands to about one hundred thousand tonnes—shows broad agreement between estimated natural mortality and predator consumption.

Separating the catch by fleet allowed for more precise estimation of fleet-specific selectivity compared with previous assessments, potentially reducing confounding with other processes. Process error on fixed gear selectivity allowed to capture the variation in fish size at age and gillnet mesh size usage over time. The resulting annual selectivity vectors matched the expected age and size selection in the catch from the available mesh size changes, size at age changes, and mesh selectivity at age and size information (Turcotte et al. In press).

Finally, process error in gillnet CPUE q performed poorly in all models and substantially affected SSB estimates when random effects were constrained to produce credible q values, values that the freely estimated process error parameters could not generate. This represents a notable departure from previous assessments, in which the CPUE index was historically relied upon. However, the SG acoustic survey is expected to provide a much more reliable indicator of SSB on the spawning grounds. Unlike CPUE, it is not subject to hyperstability due to its sampling design, also directly involves industry patterns because fishers participate in the sampling, and is therefore likely a substantial improvement over the CPUE index, historically produced jointly by DFO and industry, for estimating biomass on the spawning grounds.

UNCERTAINTY

Uncertainty in the catch depends on the reliability of the split between spring and fall spawners, which in turn relies on biological sampling. The number of fish sampled varies by fishery and year, generally being high for the fixed-gear fishery but lower for the mobile-gear fisheries (except Inshore North which had high quality data; Turcotte et al. In press). To account for this lower precision, mobile-gear catches in the Middle region and combined catches in the South region have been assigned higher global CVs. Nevertheless, annual effective sample sizes and total catch CVs could be calculated and incorporated into the assessment models to explicitly account for year-to-year variation in catch uncertainty.

Another source of uncertainty lies in the scale of the population. Here, the total SSB reached a maximum of nearly 750,000 t in 1986. This estimate is approximately 2.5 times larger than the SSB from the statistical catch-at-age model used in the previous assessment (Maynard et al. 2024), yet about half of the SSB estimated for the same period by the TMB-based assessment of Rossi and Turcotte (In press). For context, the scaling of SSB for this population is comparable to other species, such as mackerel, which reached around 400,000 t in 1985–1990 (Van Beveren et al. 2023), and spring-spawning herring, which were around 175,000 t in 1985–1995.

The South model showed bias in self-test and retrospective patterns that were resolved when additional years of data were provided to the model. Nonetheless, these routine diagnostics should be performed and analyzed with scrutiny at every assessment cycle to avoid biased population estimates.

Differences among the three herring assessment models in data treatment, model structure, and estimation framework highlight the well-known challenge of estimating the absolute scale of a population. SSB estimates are highly sensitive to assumptions regarding selectivity, natural mortality, catchability, and the treatment of process versus observation error, and different modeling choices can produce substantially different absolute biomass estimates. In particular, assumptions about natural mortality alone can scale the herring population biomass considerably (Rossi and Turcotte In press). Consistent with this, Szuwalski 2023 found that in eastern Berring sea snow crab, allowing natural mortality or catchability to vary over time improved model fits and reduced retrospective patterns, yet the resulting management advice differed by almost an order of magnitude depending on which processes were allowed to vary and how they were constrained within the estimation framework. Here, the processes were not constrained, which can increase our confidence in the in the scale of estimates.

Improvements to the models could be realized in the short-term for the following.

1. Features were recently added to WHAM that allow to model multiple stocks with movement according to season and regions. This could be well suited for this stock and could potentially allow to separate the overwinter catch among the three regions.
2. A model assuming a mixed population should also be attempted as an alternative stock structure exploration.
3. Environmental effects of recruitment should be explored as research is produced to better inform recruitment dynamics and reduce the uncertainty in recruitment estimates and population forecasts.
4. Removing from the likelihood the proportions at age from the commercial catch where the data was not from the most optimal selection algorithm level in a given year (Turcotte et al. In press).
5. Calculating CVs for the various biomass indices and use as input to the models.

CONCLUSION

The state-space models performed well and are recommended for use in future stock assessments and management. WHAM is a R package widely used in stock assessments in the US and is increasingly adopted within DFO. Its broad user base provides several advantages over in-house models, including ongoing maintenance, community support for troubleshooting, and continuous development of improvements.

REFERENCES CITED

Babyn, J., and Cadigan, N.G. 2025. A PSA for OSA residuals: residual diagnostics for state–space stock assessment models. Canadian Journal of Fisheries and Aquatic Sciences. 82: 1-18.

Benoît, H.P., and Swain, D.P. 2008. Impacts of environmental change and direct and indirect harvesting effects on the dynamics of a marine fish community. Can. J. Fish. Aquat. Sci. 65: 2088–2104.

Benoît, H.P., and Rail, J.-F. 2016. Principal predators and consumption of juvenile and adult Atlantic Herring (*Clupea harengus*) in the southern Gulf of St. Lawrence. DFO Can. Sci. Advis. Sec. Res. Doc. 2016/065. viii + 42 p.

Brophy, D., Danilowicz, B. S., and King, P. A. 2006. Spawning season fidelity in sympatric populations of Atlantic Herring (*Clupea harengus*). Can. J. Fish. Aquat. Sci. 63: 607–616.

Cadrin, S.X. 2025. Misinterpreting retrospective patterns in fishery stock assessment. ICES J. Mar. Sci. 82(2).

Erisman, B.E., Allen, L.G., Claisse, J.T., Pondella, D.J., Miller, E.F., Murray, J.H., and Walters, C. 2011. The illusion of plenty: hyperstability masks collapses in two recreational fisheries that target fish spawning aggregations. Can. J. Fish. Aquat. Sci. 68(10): 1705–1716.

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.

Francis, C.R.I.C., Hurst, J.R., Renwick, J.A. 2002. Quantifying annual variation in catchability for commercial and research fishing. Fishery Bulletin. 101(2).

Francis, C.R.I.C. 2011. Data weighting in statistical fisheries stock assessment models. Can. J. Fish. Aquat. Sci. 68: 1124–1138.

Gudmundsson, G., and Gunnlaugsson, T. 2012. Selection and estimation of sequential catch-at- age models. Can. J. Fish. Aquat. Sci. (69): 1760–1772.

Hammill, M.O., Rossi, S.P., Mosnier, A., den Heyer, C.E., Bowen, W.D., and Stenson, G.B. 2023. Grey Seal Abundance in Canadian Waters and Harvest Advice. DFO Can. Sci. Advis. Sec. Res. Doc. 2023/053. iv + 40 p.

Harley, S. J., Myers, R. A. and Dunn, A. 2001. Is catch-per-unit-effort pro- proportional to abundance? Can. J. Fish. Aquat. Sci. 58: 1760–1772.

Hilborn, R. and Walters, C. 1992. Quantitative Fisheries Stock Assessment: Choice, Dynamics and Uncertainty; Springer Science: Boston, MA, USA, 1992; pp. 453–470.

Li, C., Deroba, J.J., Miller, T.J., Legault, C.M., and Perretti, C.T. 2024. An evaluation of common stock assessment diagnostic tools for choosing among state-space models with multiple random effects processes. Fish. Res. 273 (February): 106968. Elsevier B.V.

Mackinson, S., Sumaila, U. R. and Pitcher, T. J. 1997. Bioeconomics and catchability: Fish and fishers behavior during stock collapse. Fish. Res. 31: 11–17.

Maynard, L, Burbank, J., DeJong, R., Turcotte, F., McDermid, J., Sylvain, F.-E. and Rolland, N. 2024. Supporting Material for Stock Assessment of NAFO Division 4TVn Southern Gulf of St. Lawrence Atlantic Herring (*Clupea harengus*) in 2022-2023. DFO Can. Sci. Advis. Sec. Res. Doc. 2024/058. xi + 147 p.

Maynard, L.D., Landry, L., and Burbank, J. In press. NAFO Division 4TVn Fall Spawning Atlantic Herring (*Clupea harengus*) Assessment Framework Review: Catch per Unit Effort Index. DFO Can. Sci. Advis. Sec. Res. Doc.

McQuinn, I.H. 1997. Metapopulations and the Atlantic herring. Rev. Fish Biol. Fish. 7: 297–329.

Moll, D., Polte, P., Jochum, K.P., Gröhsler, T., Bekkevold, D., Mcquinn, I., Möllmann, C., Zimmermann, C., and Kotterba, P. 2025. Direct evidence of natal homing in an Atlantic herring metapopulation. Sci. Adv. 11.

Neuenhoff, R.D., Swain, D.P., Cox, S.P., McAllister, M.K., Trites, A.W., Walters, C.J., and Hammill, M.O. 2019. Continued decline of a collapsed population of Atlantic cod (*Gadus morhua*) due to predation-driven Allee effects. Can. J. Fish. Aquat. Sci. 76: 168–184.

Peterman, R.M., and G.J. Steer. Relation between sport-fishing catch- ability coefficients and salmon abundance. Trans. Am. Fish. Soc. 110: 585–593 (1981).

Pitcher, T.J. 1995. The impact of pelagic fish behavior on fisheries. Scientia Marina. 59: 295–306.

Rossi, S.P., Cox, S.P., Hammill, M.O., den Heyer, C.E., Swain, D.P., Mosnier, A. and Benoît, H.P., 2021. Forecasting the response of a recovered pinniped population to sustainable harvest strategies that reduce their impact as predators. ICES Journal of Marine Science. 78: 1804-1814.

Rossi, S.P., and Turcotte, F. In press. A Spatially-Disaggregated Statistical Catch-at-Age Model for Fall Spawning Atlantic Herring (*Clupea harengus*) in the Southern Gulf of St. Lawrence (NAFO Division 4TVn) Accounting for Predation. DFO Can. Sci. Advis. Sec. Res. Doc.

Stewart, I.J., and Monnahan, C.C. 2025. Diagnosing common sources of lack of fit to composition data in fisheries stock assessment models using One-Step-Ahead (OSA) residuals. CJFAS: 0–3.

Stock, B.C., and Miller, T.J. 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. Fish. Res. 240(April): 105967. Elsevier B.V.

Swain, D.P., and Benoît, H.P. 2015. Extreme increases in natural mortality prevent recovery of collapsed fish populations in a Northwest Atlantic ecosystem. Marine Ecology Progress Series. 519: 165–182.

Swain, D.P, Benoît, H.P, and Hammill, M.O. 2015. Spatial distribution of fishes in a Northwest Atlantic ecosystem in relation to risk of predation by a marine mammal. J. Anim. Ecol. 84: 1286-1298.

Szuwalski, C.S., Ianelli, J.N., Punt, A.E., 2018. Reducing retrospective patterns in stock assessment and impacts on management performance. ICES J. Mar. Sci. 75: 596–609.

Szuwalski, C.S., Hollowed, A.B., Holsman, K.K., Ianelli, J.N., Legault, C.M., Melnychuk, M.C., Ovando, D., and Punt, A.E. 2023. Unintended consequences of climate-adaptive fisheries management targets. Fish and Fisheries. 24: 439–453.

Turcotte, F., Swain, D. P., and McDermid, J. L. 2021a. NAFO 4TVn Atlantic herring population models: from Virtual Population Analysis to Statistical Catch-at-Age estimating time-varying natural mortality. DFO Can. Sci. Advis. Sec. Res. Doc. 2021/029. vi + 52 p.

Turcotte, F., Swain, D.P., McDermid, J.L., and DeJong, R. 2021b. Assessment of the NAFO Division 4T southern Gulf of St. Lawrence Atlantic herring (*Clupea harengus*) in 2018 and 2019. DFO Can. Sci. Advis. Sec. Res. Doc. 2021/030. xiv + 158 p.

Turcotte, F., McDermid, J.L., Tunney, T.D., and Hanke, A. 2021c. Increasing occurrence of Atlantic bluefin tuna on Atlantic herring spawning grounds: a signal of escalating pelagic predator–prey interaction? Marine and Coastal Fisheries. 13: 240–252.

Turcotte, F. 2022. Environmental effects on recruitment dynamics and population projections of NAFO Division 4TVn Spring Spawning Atlantic Herring. DFO Can. Sci. Advis. Sec. Res. Doc. 2022/047. iv + 25 p.

Turcotte, F., Hanke, A., and McDermid, J.L. 2023. Atlantic bluefin tuna diet variability in the southern Gulf of St. Lawrence, Canada. Mar. Environ. Res. 187: 105949.

Turcotte, F., Landry, L., McDermid, J.l., Maynard, L., and Burbank, J. In press. NAFO Division 4TVn Fall Spawning Atlantic Herring (*Clupea harengus*) Assessment Framework Review: Fisheries and Surveys Data. DFO Can. Sci. Advis. Sec. Res. Doc.

Van Beveren, E., Boudreau, M., Lévesque, L., Lehoux, C., Boudreau, M., and Plourde, S. 2023. Assessment of the Northern Contingent of Atlantic Mackerel (*Scomber scombrus*) in 2022. DFO Can. Sci. Advis. Sec. Res. Doc. 2023/080. v + 48 p.

Wilberg, M.J., Thorson, J.T., Linton, B.C., and Berkson, J. 2010. Incorporating time-varying catchability into population dynamic stock assessment models. Rev. Fish. Sci. 18(1): 7–24.

Winters, G.H., and Wheeler, J.P. 1985. Interaction Between Stock Area, Stock Abundance, and Catchability Coefficient. Can. J. Fish. Aquat. Sci. 42(5): 989–998.

FIGURES

NORTH MODEL

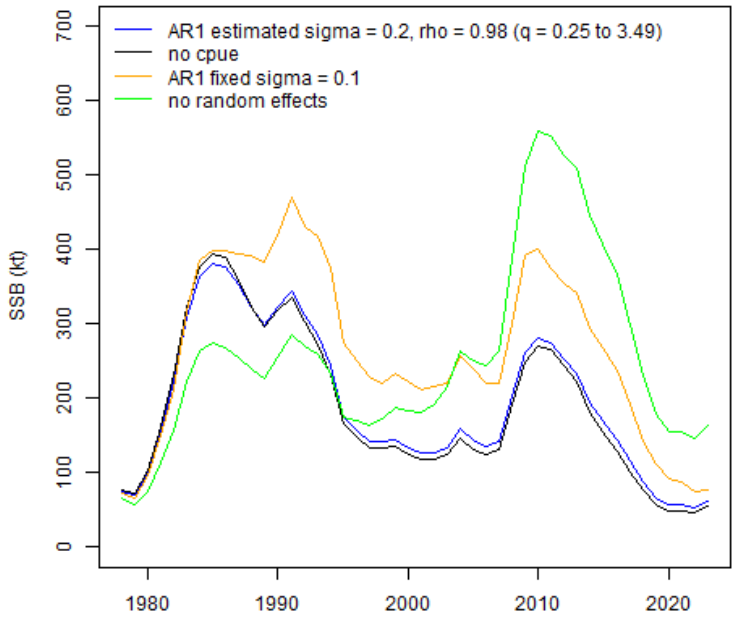

Figure 1: Spawning stock biomass (SSB kt) from the North model sensitivity runs for the CPUE index of biomass and time varying catchability (q). Base model without the CPUE index (no cpue; black line), Base model with CPUE q process error modeled as AR1 process (blue line), AR1 SD parameter fixed at 0.1 (red line) and base model with the CPUE without random effects (green line).

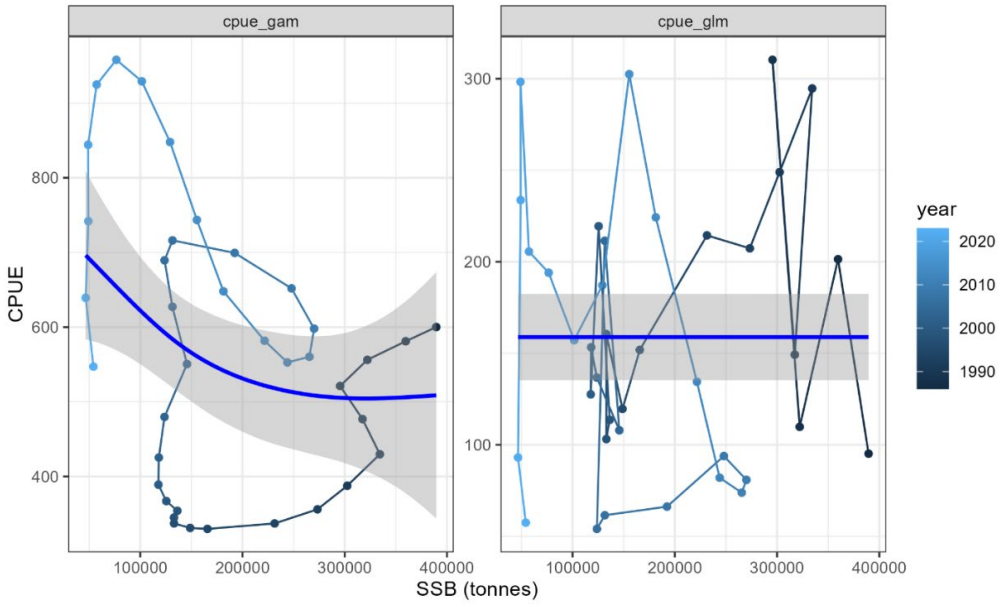

Figure 2: Gillnet CPUE biomass indices (y axis, CPUE-gam, left panel and CPUE-glm, right panel) relationship to spawning stock biomass (SSB; x axis, tonnes) by year (colored points and line) between 1978 in 2023 for the North model. Blue lines and grey shading are GAM model fit to each relationship and standard error.

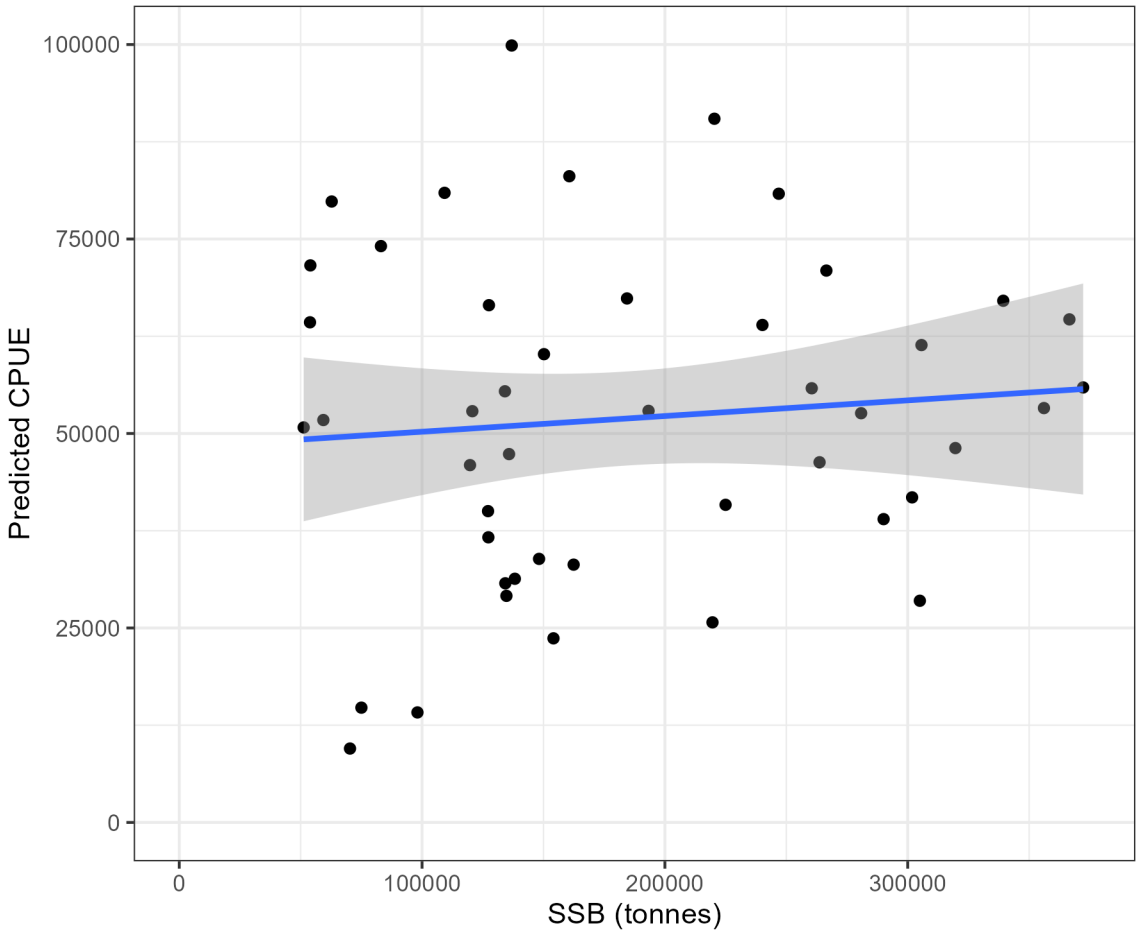

Figure 3: Predicted gillnet CPUE biomass index (CPUE-gam) relationship to spawning stock biomass (SSB; x axis, tonnes) between 1978 in 2023 for the North model sensitivity run. Blue line and grey shading is a GAM model fit and standard error.

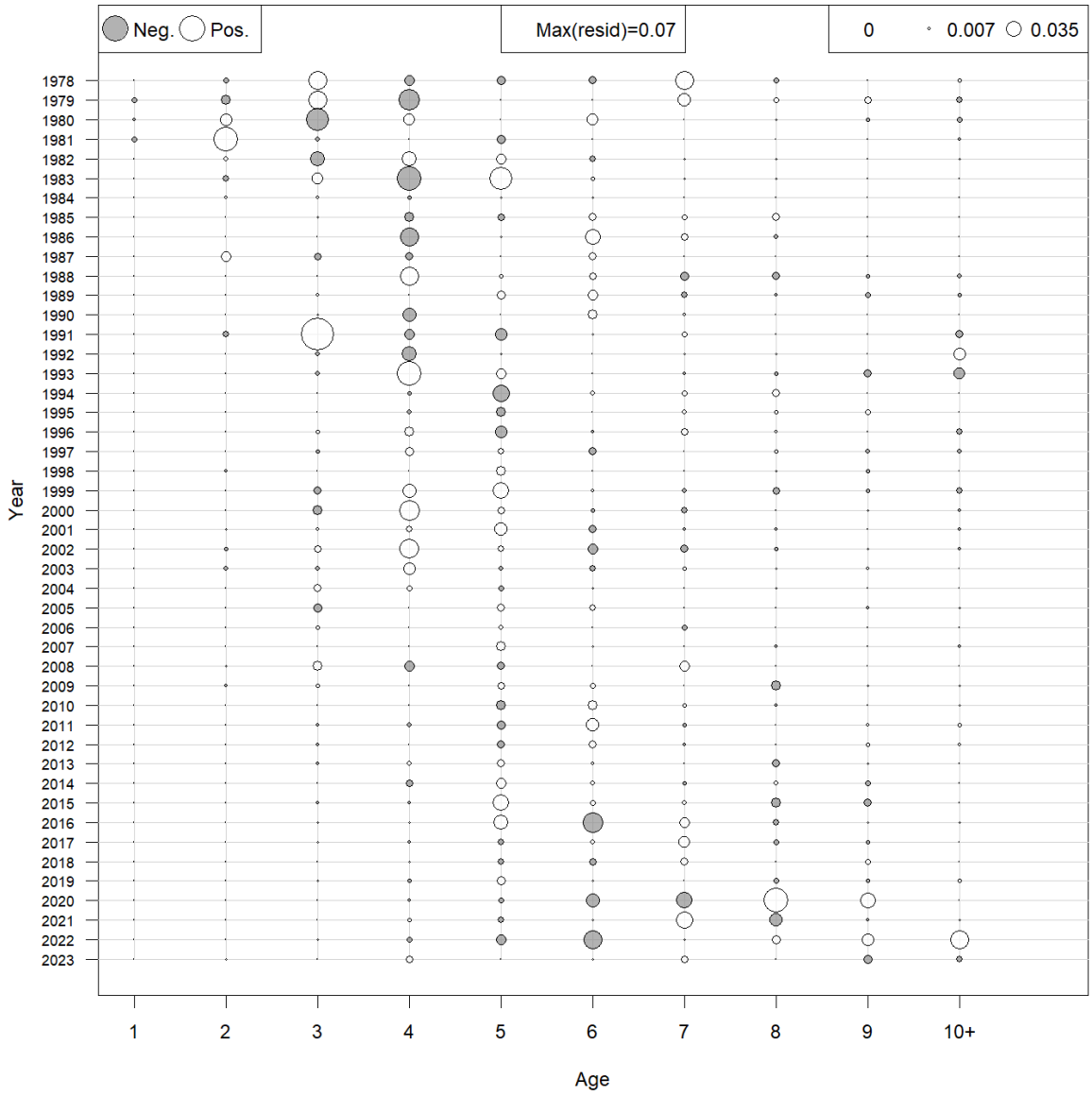

Figure 4: North model fixed gear fishery age composition raw residuals (observed-predicted) between 1978 and 2023. White circles are positive residuals and grey circles are negative residuals. Circle size is proportional to the residual value.

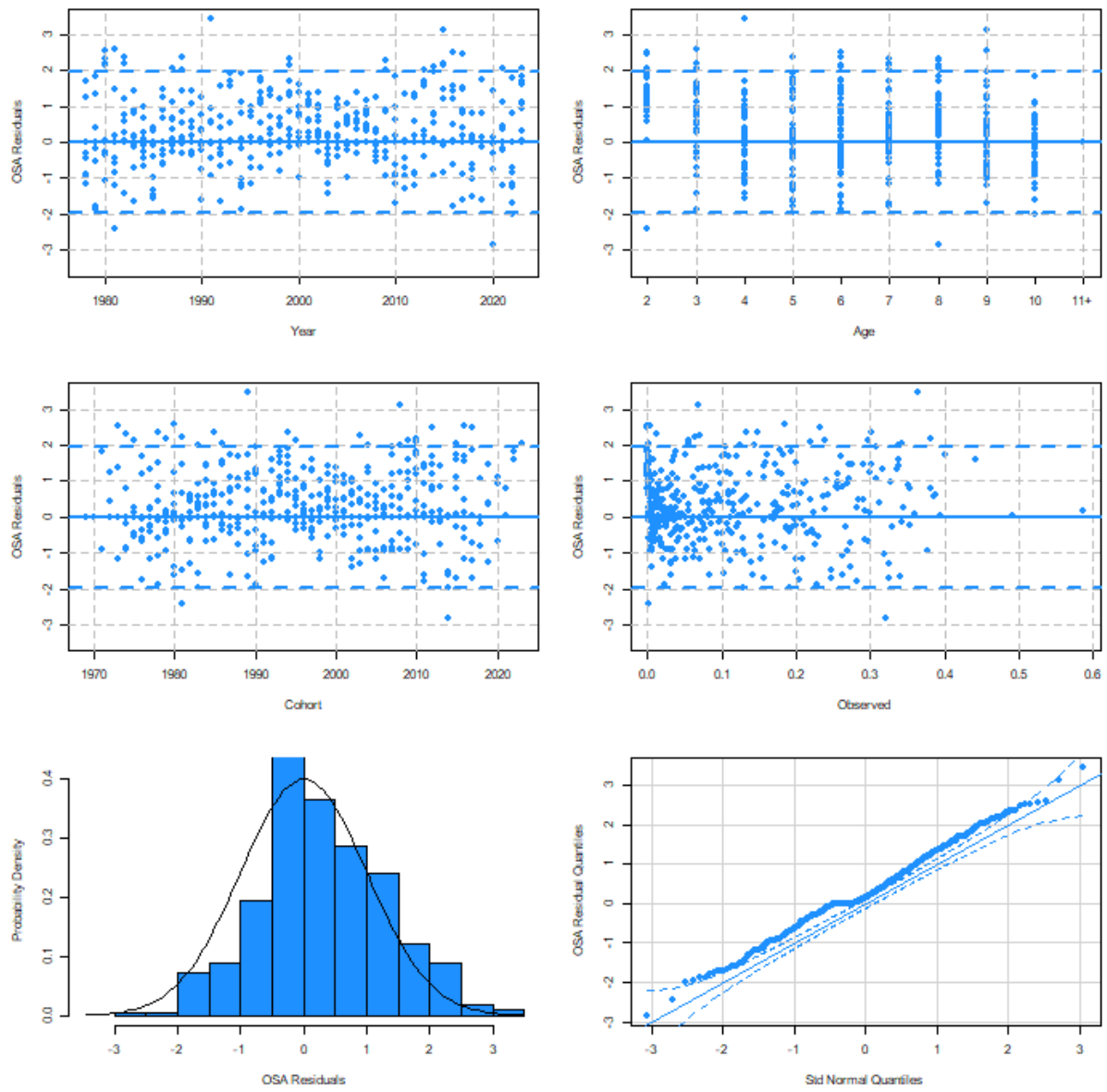

Figure 5: North model fixed gear age composition one-step-ahead (OSA) residuals by (panels from left to right top to bottom) year, age, cohort and by observed value, probability density and versus normal quantiles.

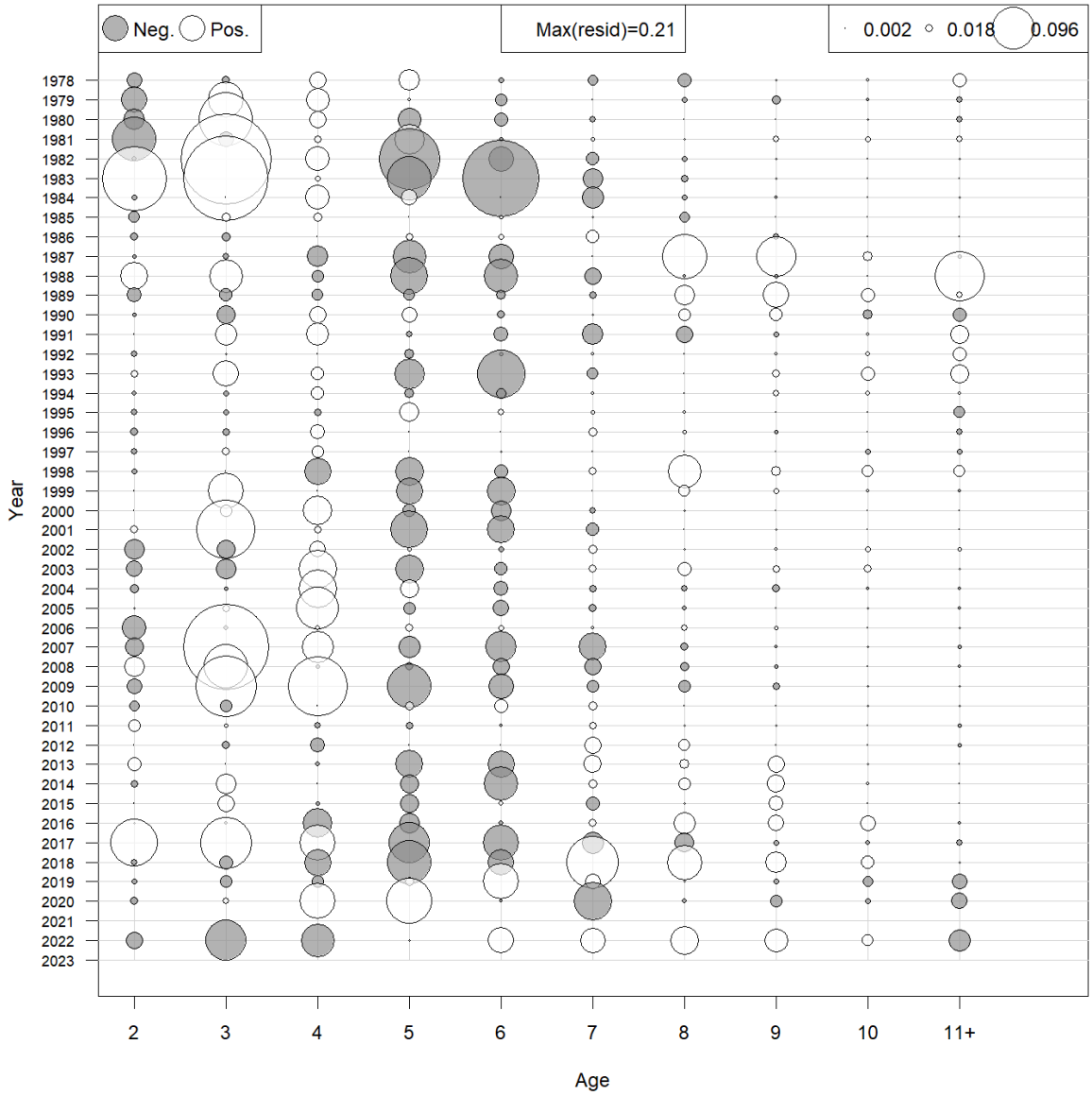

Figure 6: North model mobile gear fishery age composition raw residuals (observed-predicted) between 1978 and 2023. White circles are positive residuals and grey circles are negative residuals. Circle size is proportional to the residual value.

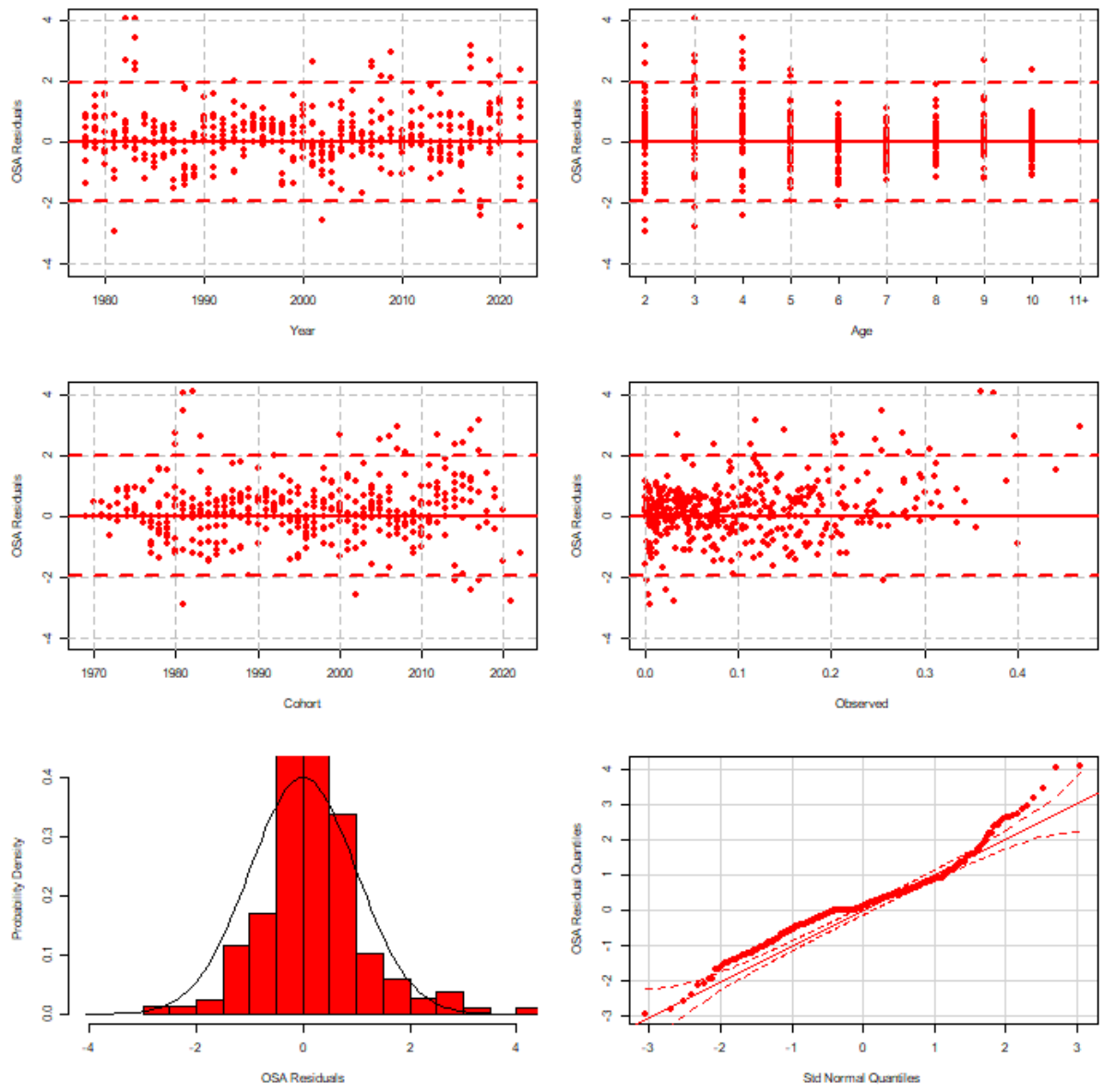

Figure 7: North model mobile gear age composition one-step-ahead (OSA) residuals by (panels from left to right top to bottom) year, age, cohort and by observed value, probability density and versus normal quantiles.

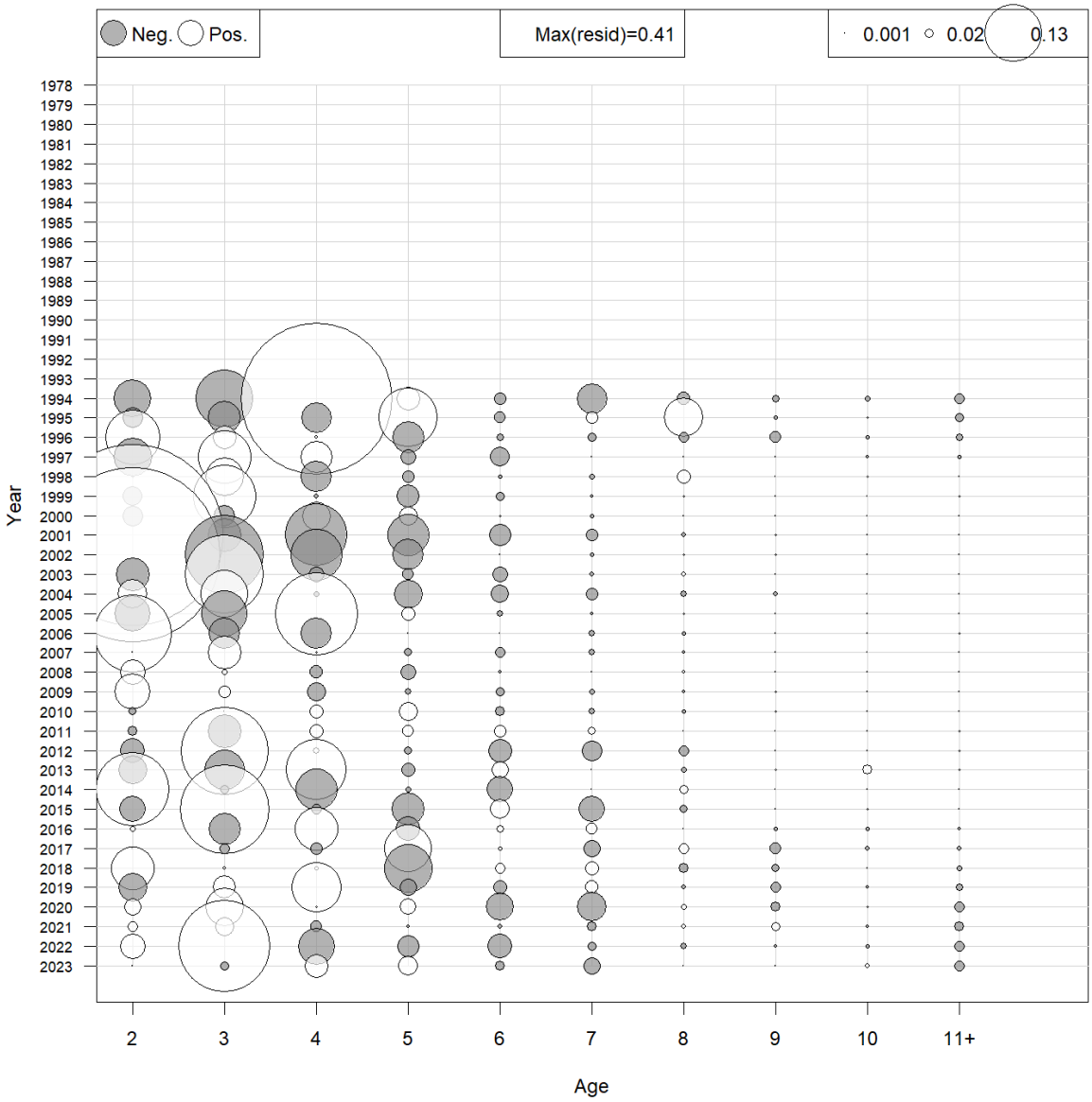

Figure 8: North model BDCM acoustic survey age composition raw residuals (observed-predicted) between 1978 and 2023. White circles are positive residuals and grey circles are negative residuals. Circle size is proportional to the residual value.

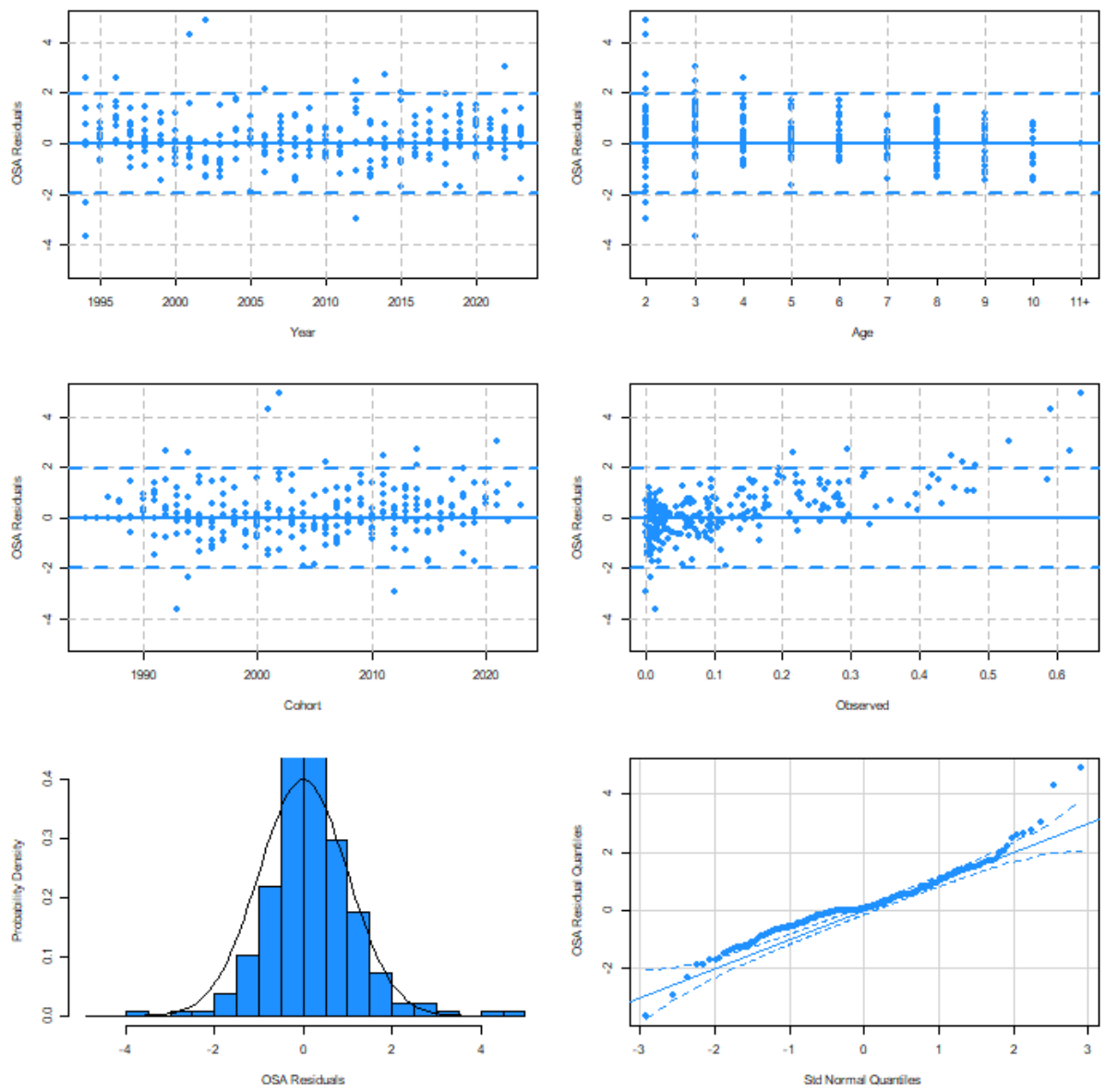

Figure 9: North model BDCM acoustic survey age composition one-step-ahead (OSA) residuals by (panels from left to right top to bottom) year, age, cohort and by observed value, probability density and versus normal quantiles.

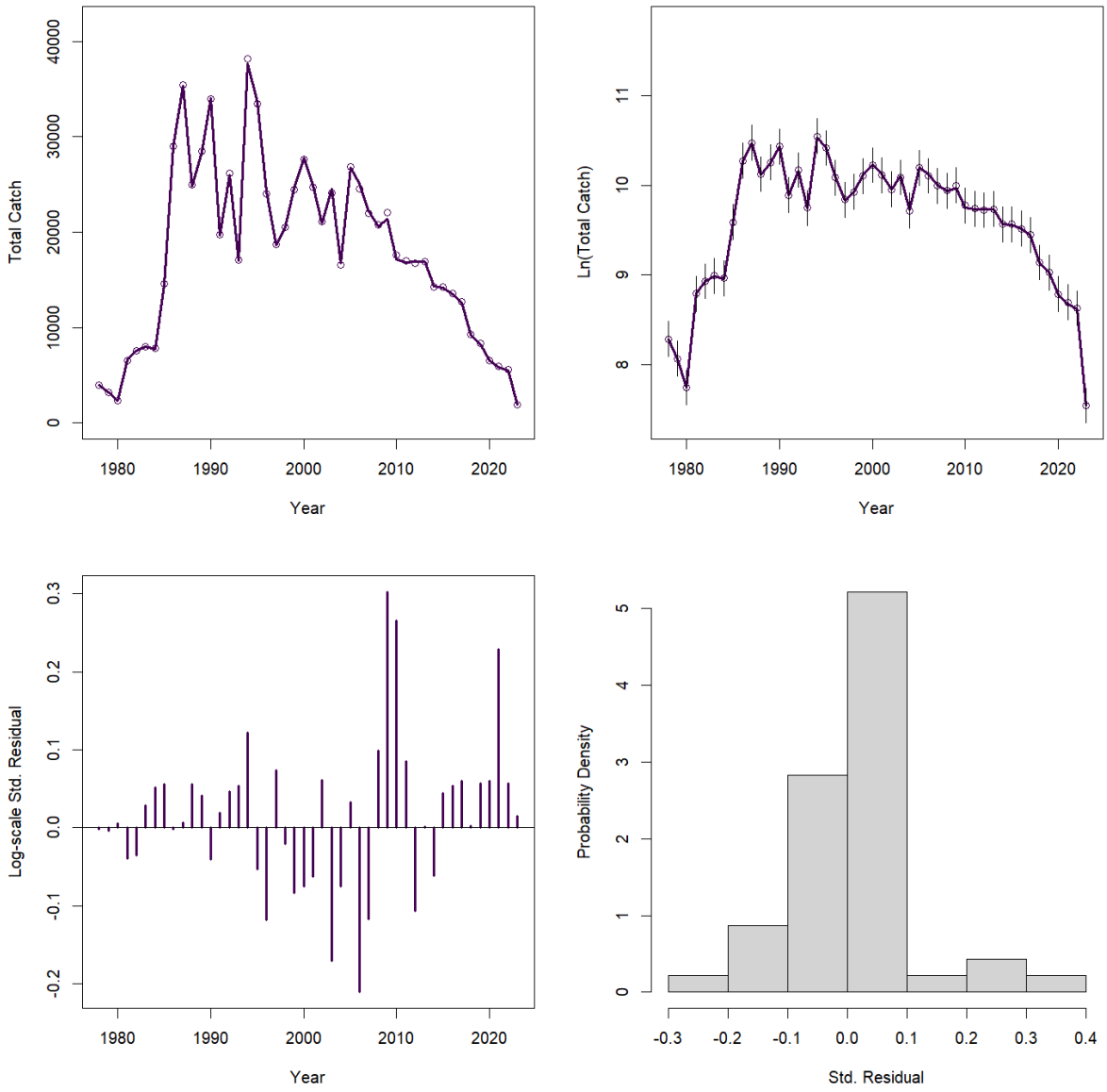

Figure 10: North model fit to the fixed gear total catch (top left), and to the log of the catch (top right), standardized residuals by year (bottom left) probability density of standardized residuals (bottom right).

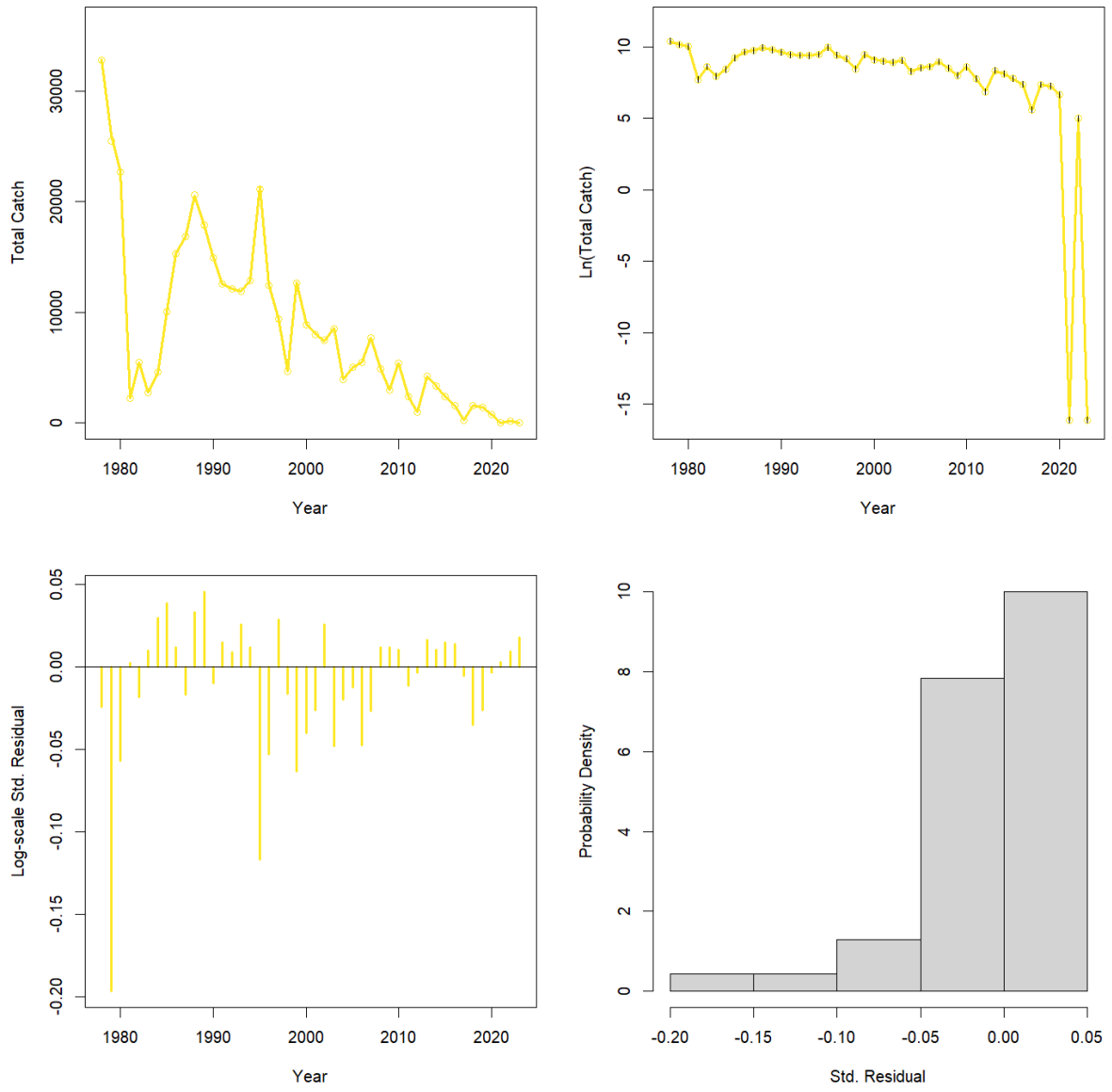

Figure 11: North model fit to the mobile gear total catch (top left), and to the log of the catch (top right), standardized residuals by year (bottom left) probability density of standardized residuals (bottom right).

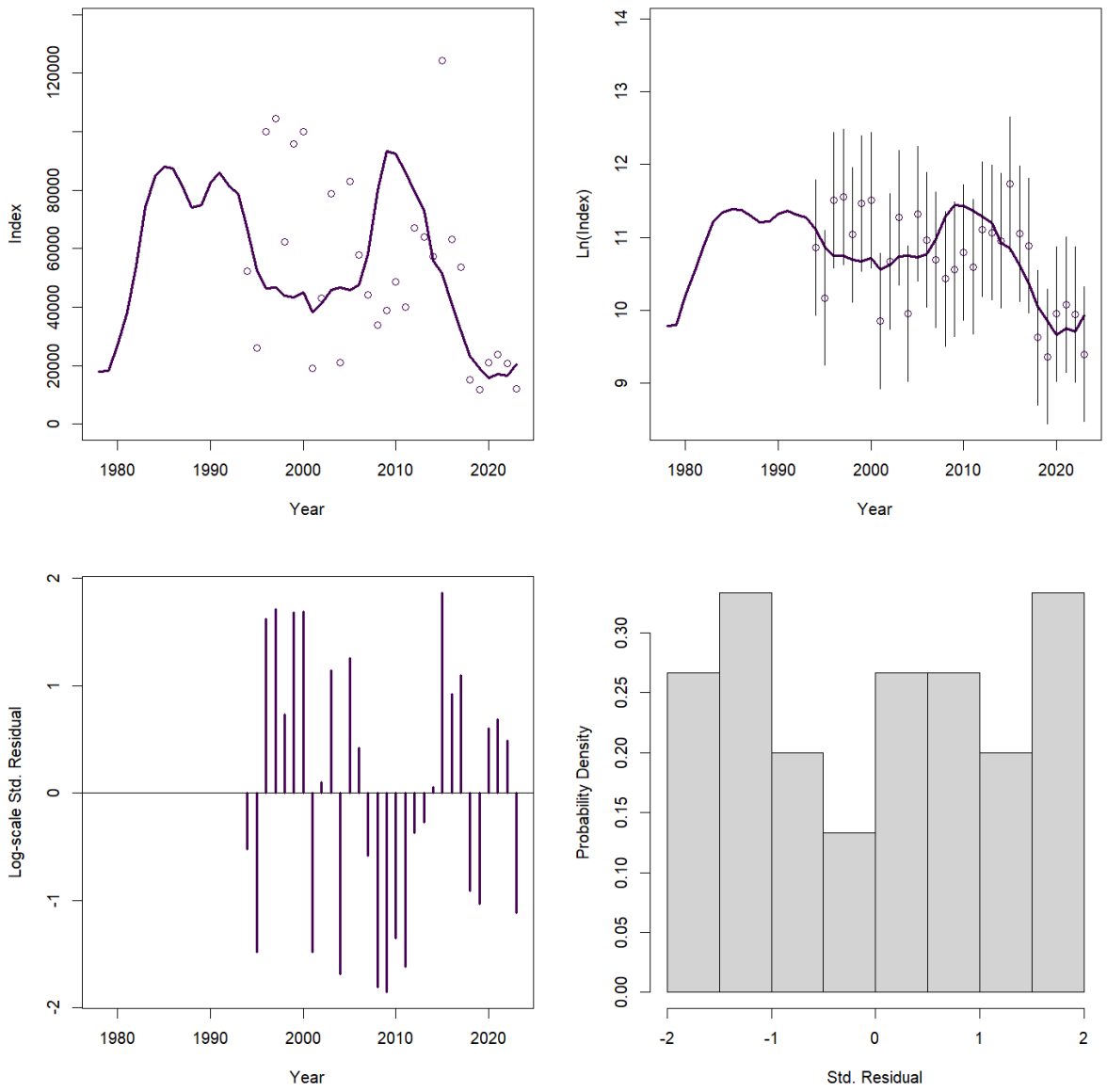

Figure 12: North model fit to the BDCM acoustic survey biomass index (top left), and to the log of the index (top right), standardized residuals by year (bottom left) probability density of standardized residuals (bottom right).

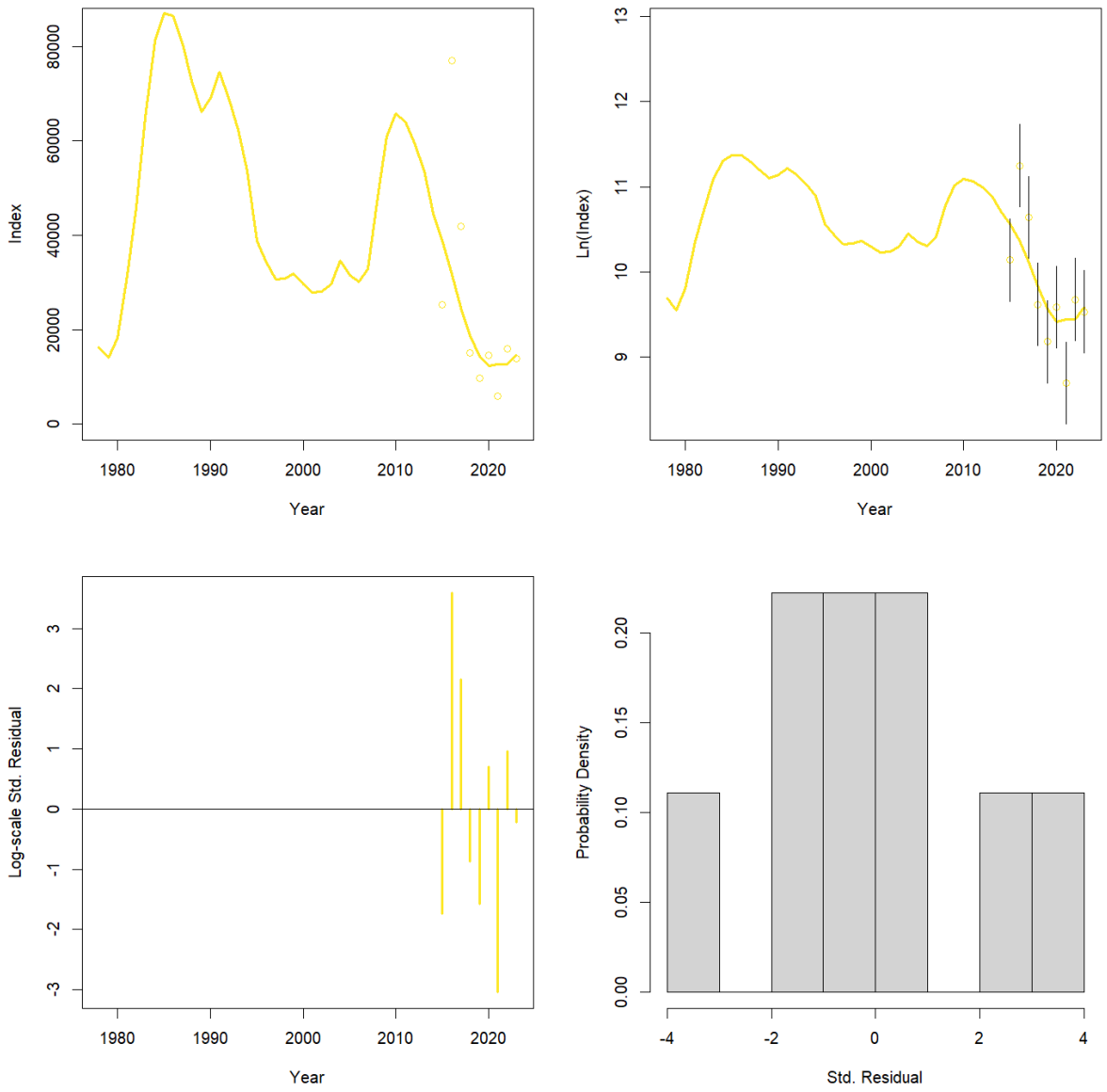

Figure 13: North model fit to the SG acoustic survey biomass index (top left), and to the log of the index (top right), standardized residuals by year (bottom left) probability density of standardized residuals (bottom right).

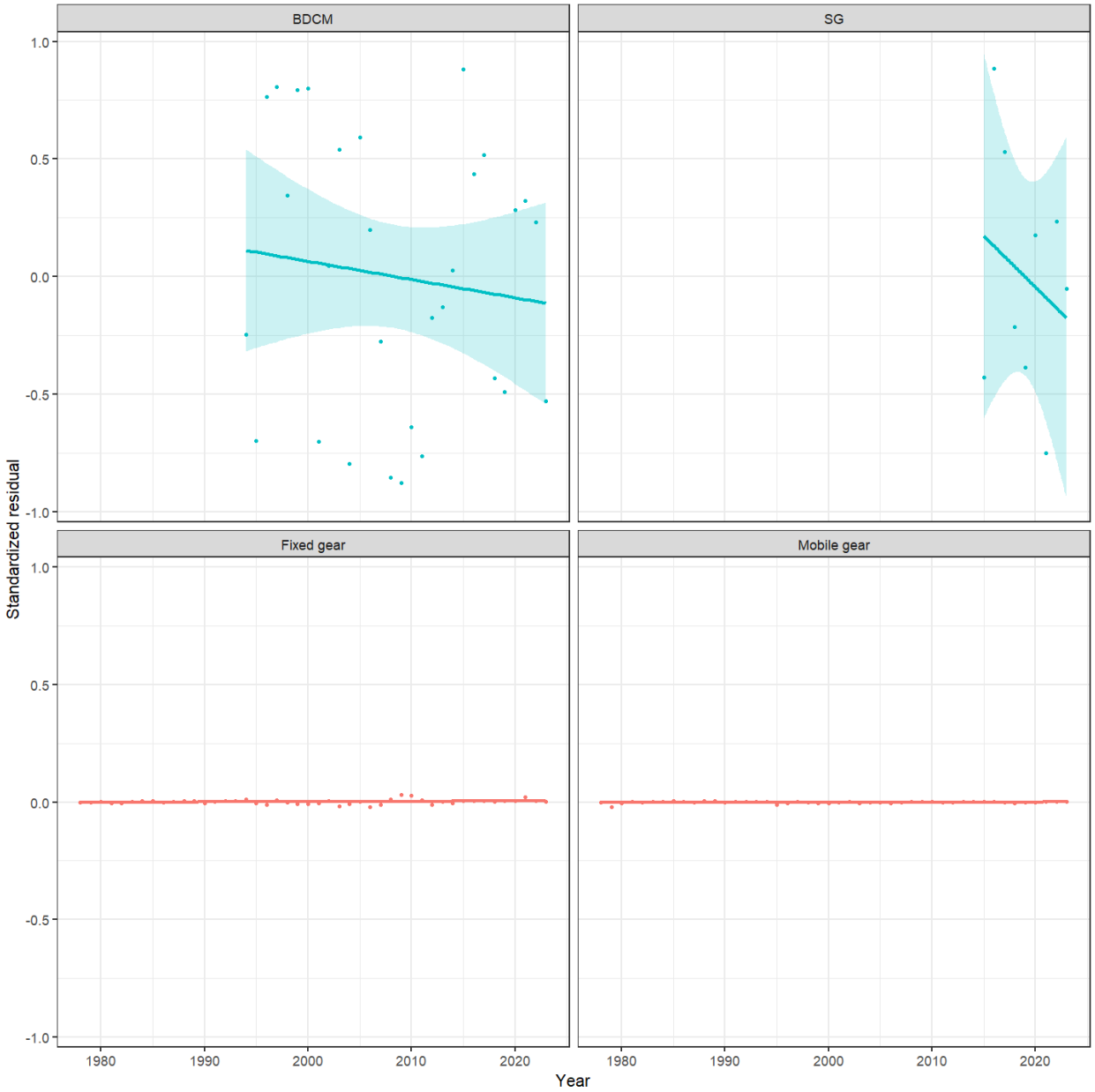

Figure 14: North model standardized residuals by year for the BDCM acoustic survey, SG acoustic survey, fixed gear catch and mobile gear catch.

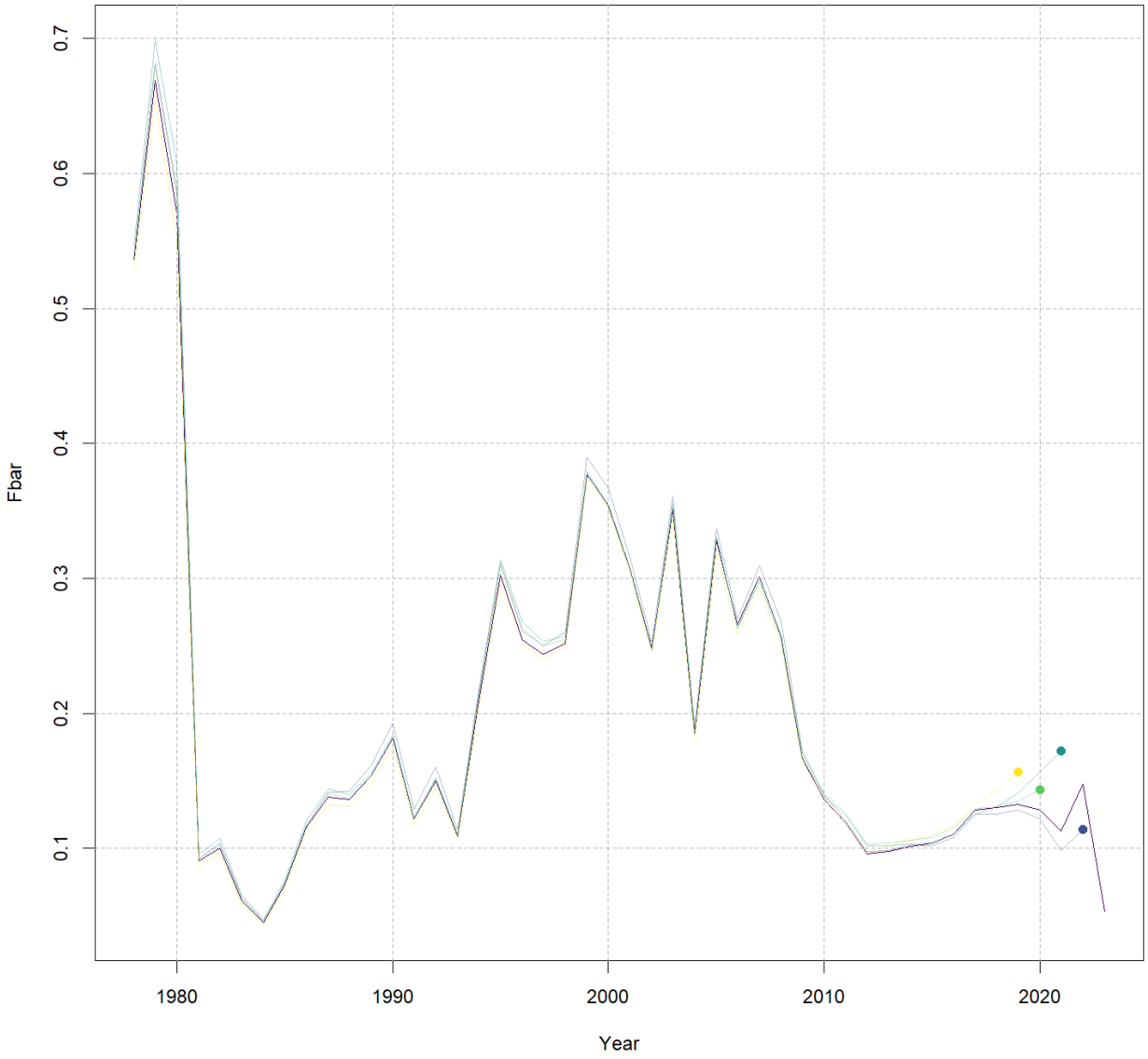

Figure 15: North model retrospective analysis for average fishing mortality (Fbar) for 4 peels.

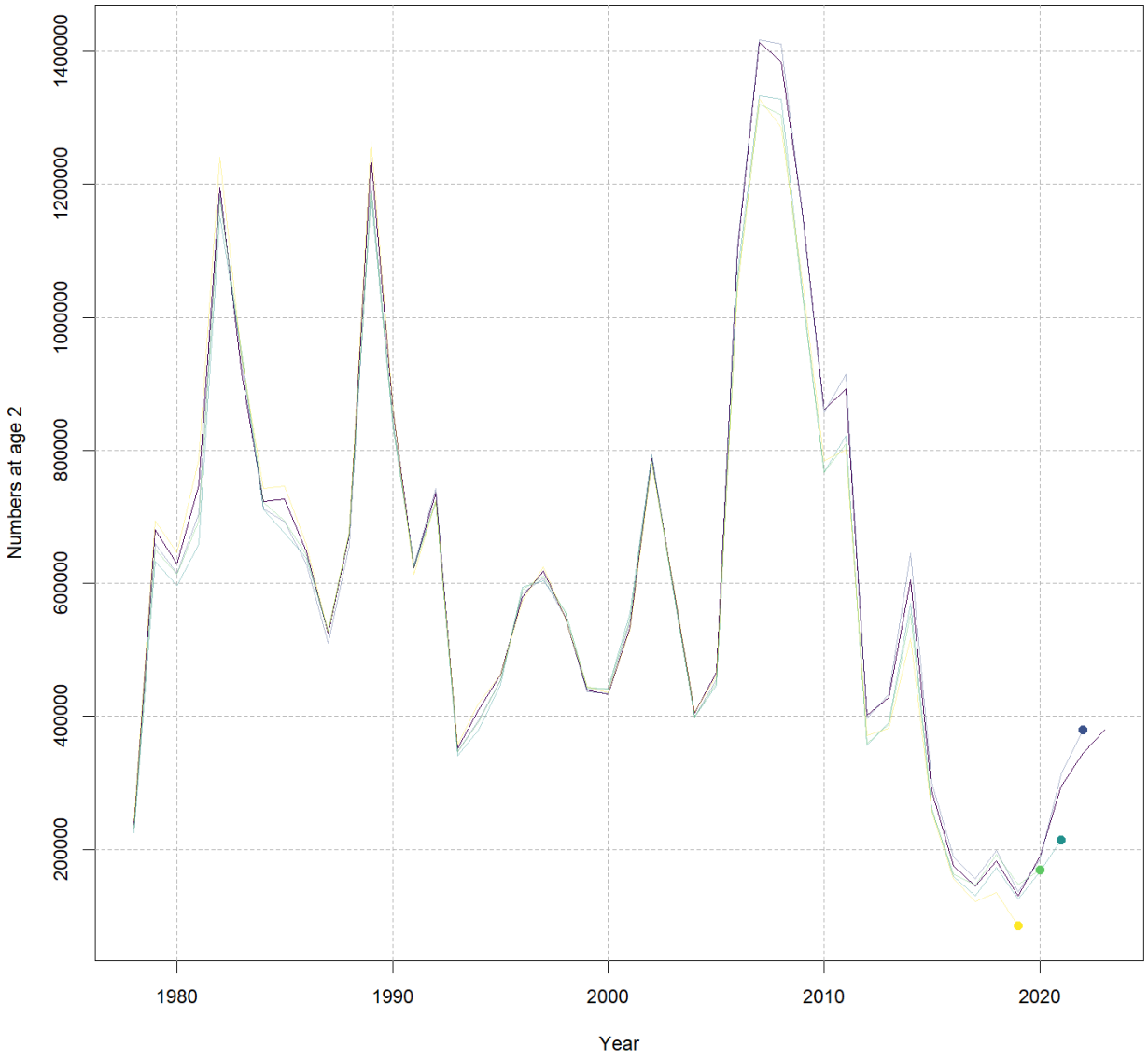

Figure 16: North model retrospective analysis for recruitment (numbers at age 2) for 4 peels.

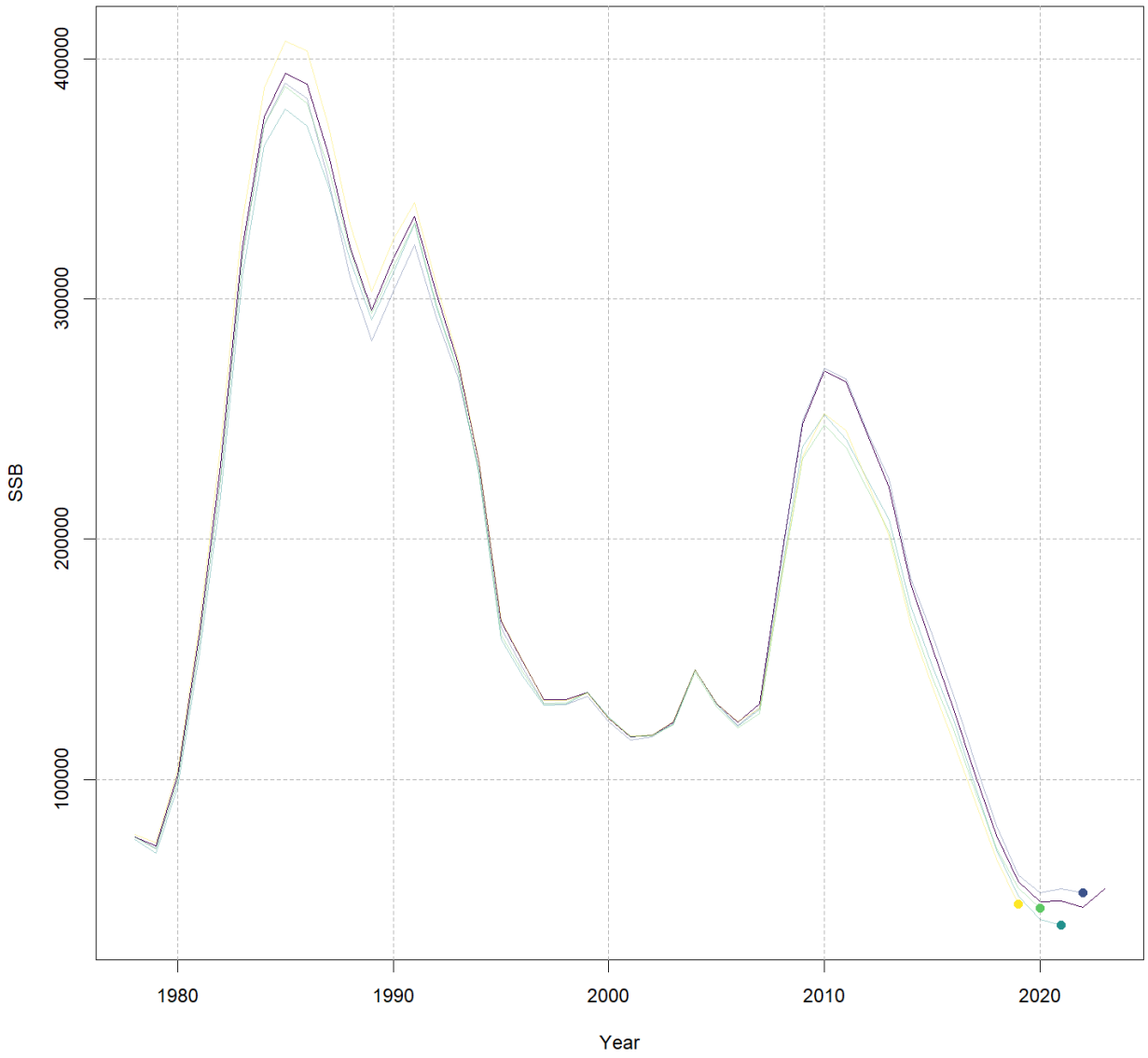

Figure 17: North model retrospective analysis for spawning stock biomass (SSB; tonnes) for 4 peels.

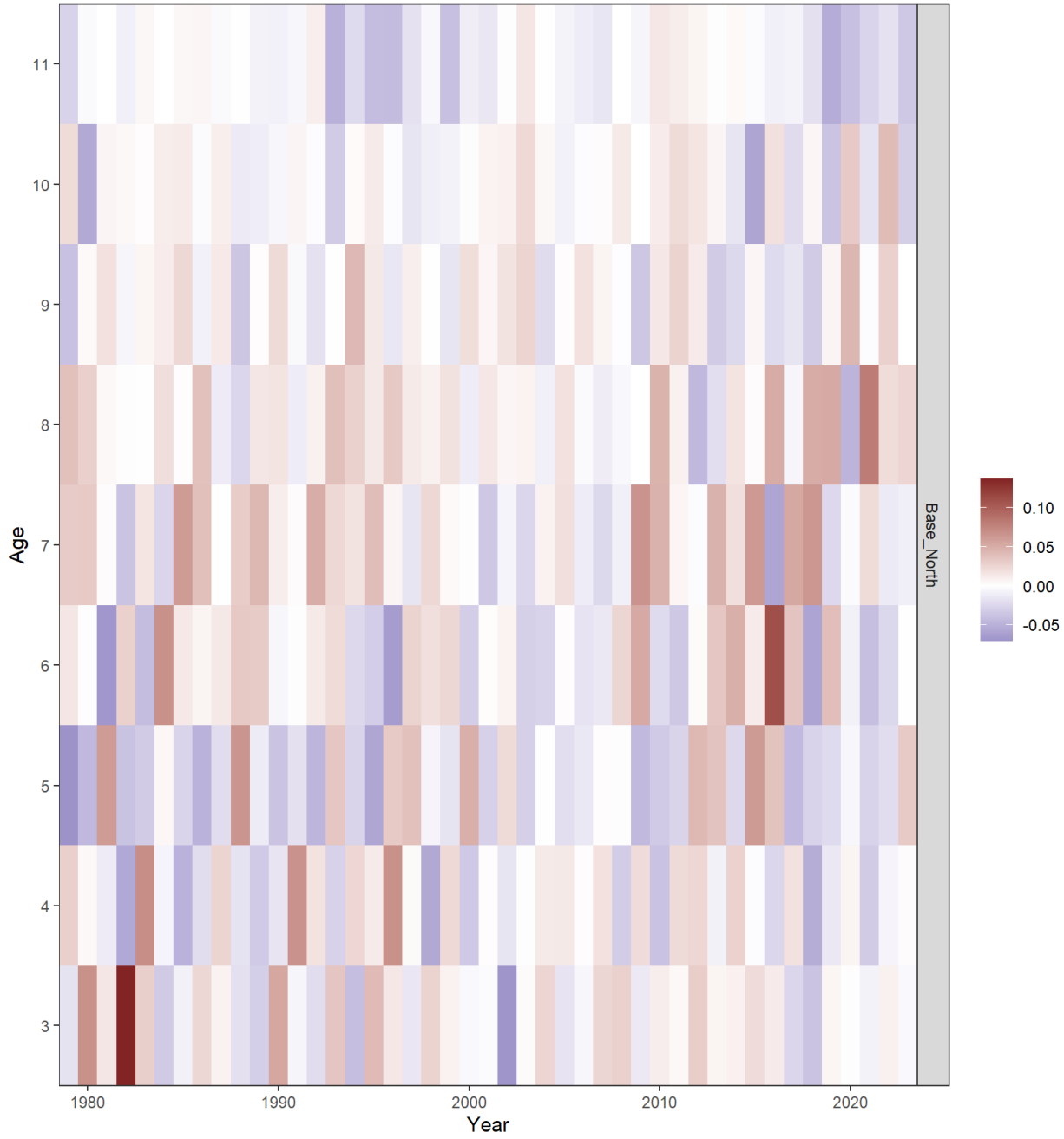

Figure 18: North model deviations in numbers-at-age (NAA) by age and year, for ages 3 to 11+.

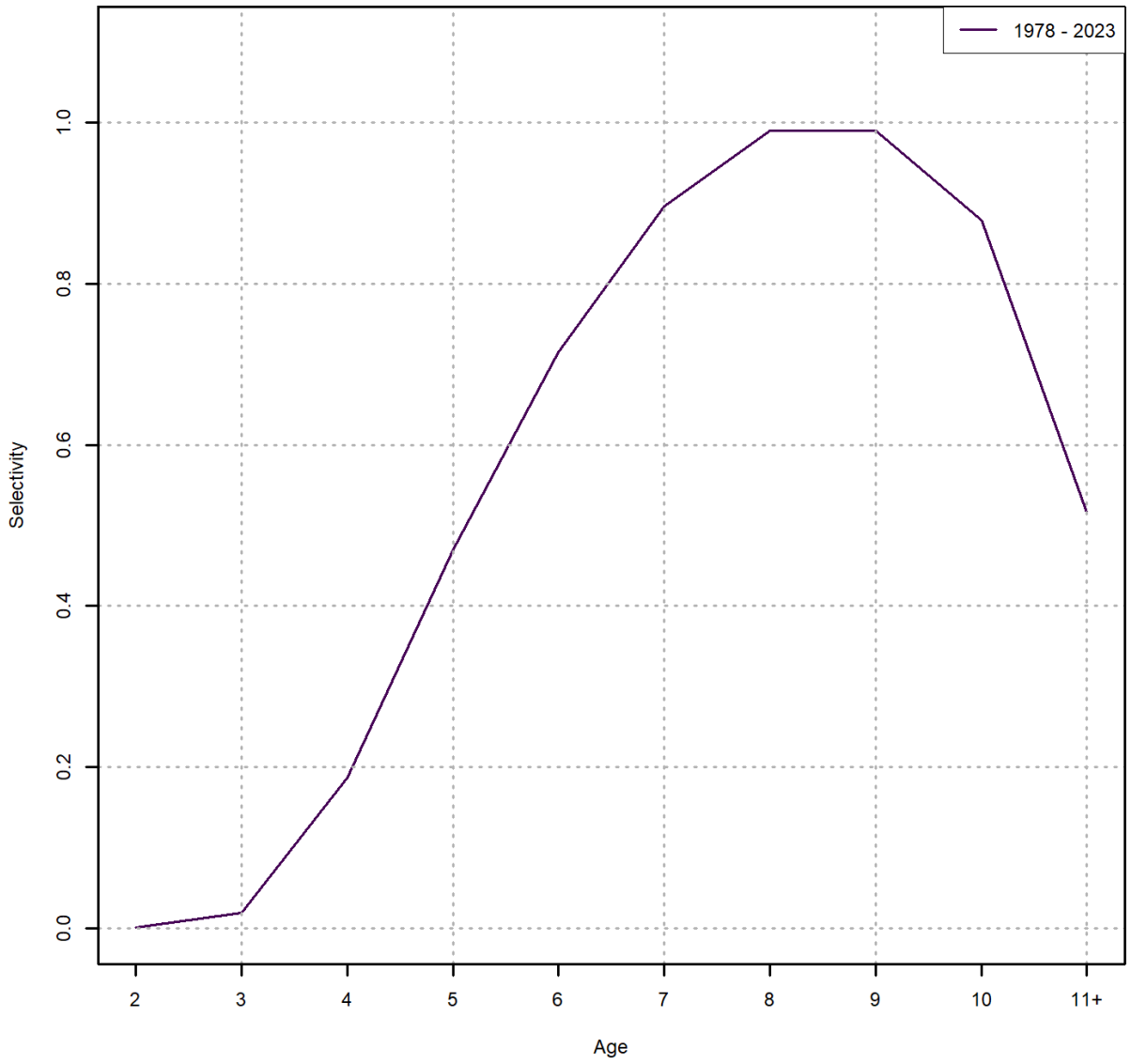

Figure 19: North model mean selectivity-at-age for the fixed gear fishery.

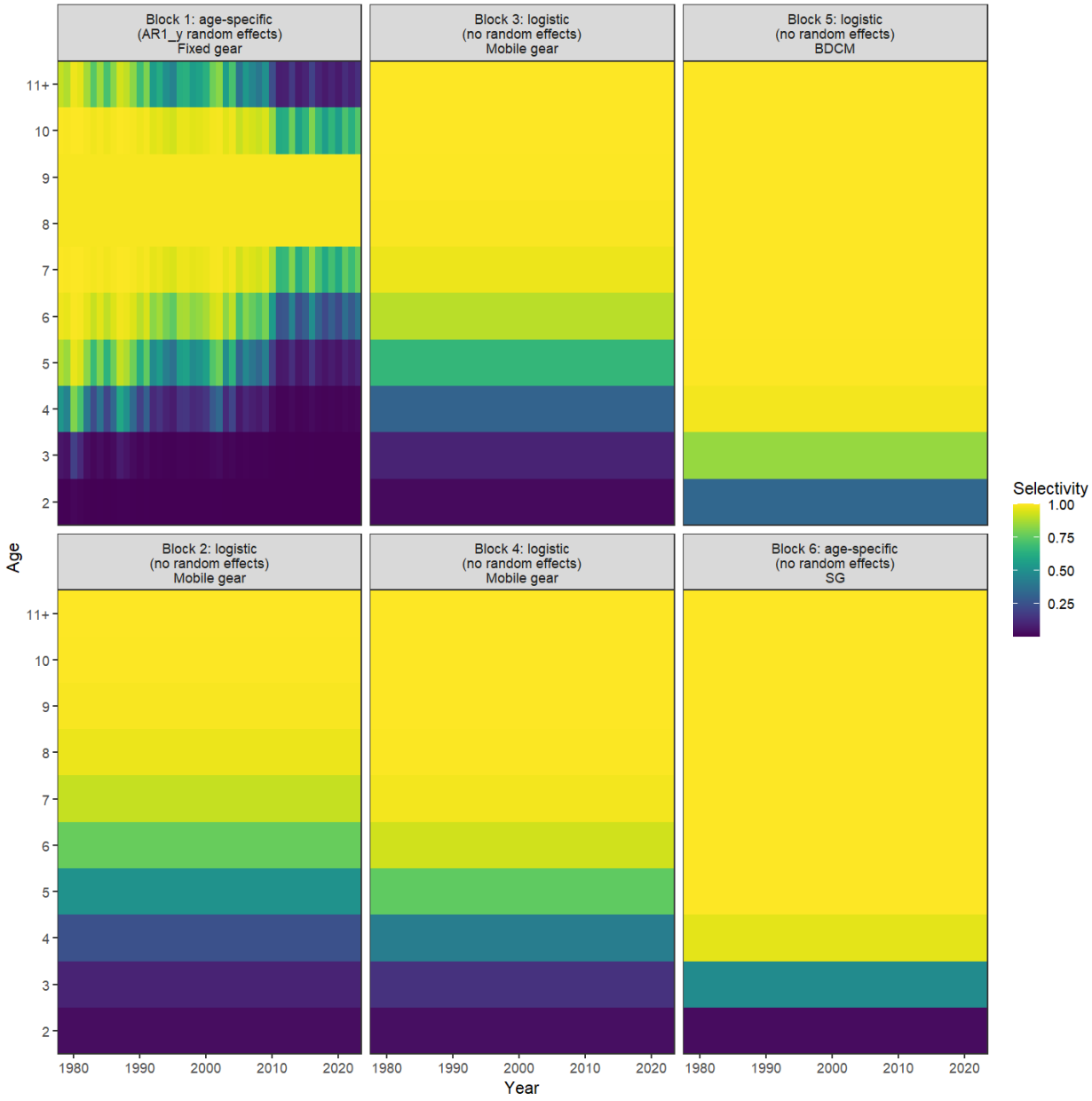

Figure 20: North model realized selectivity including random effects were applicable, for all sources of catch, surveys and time blocks.

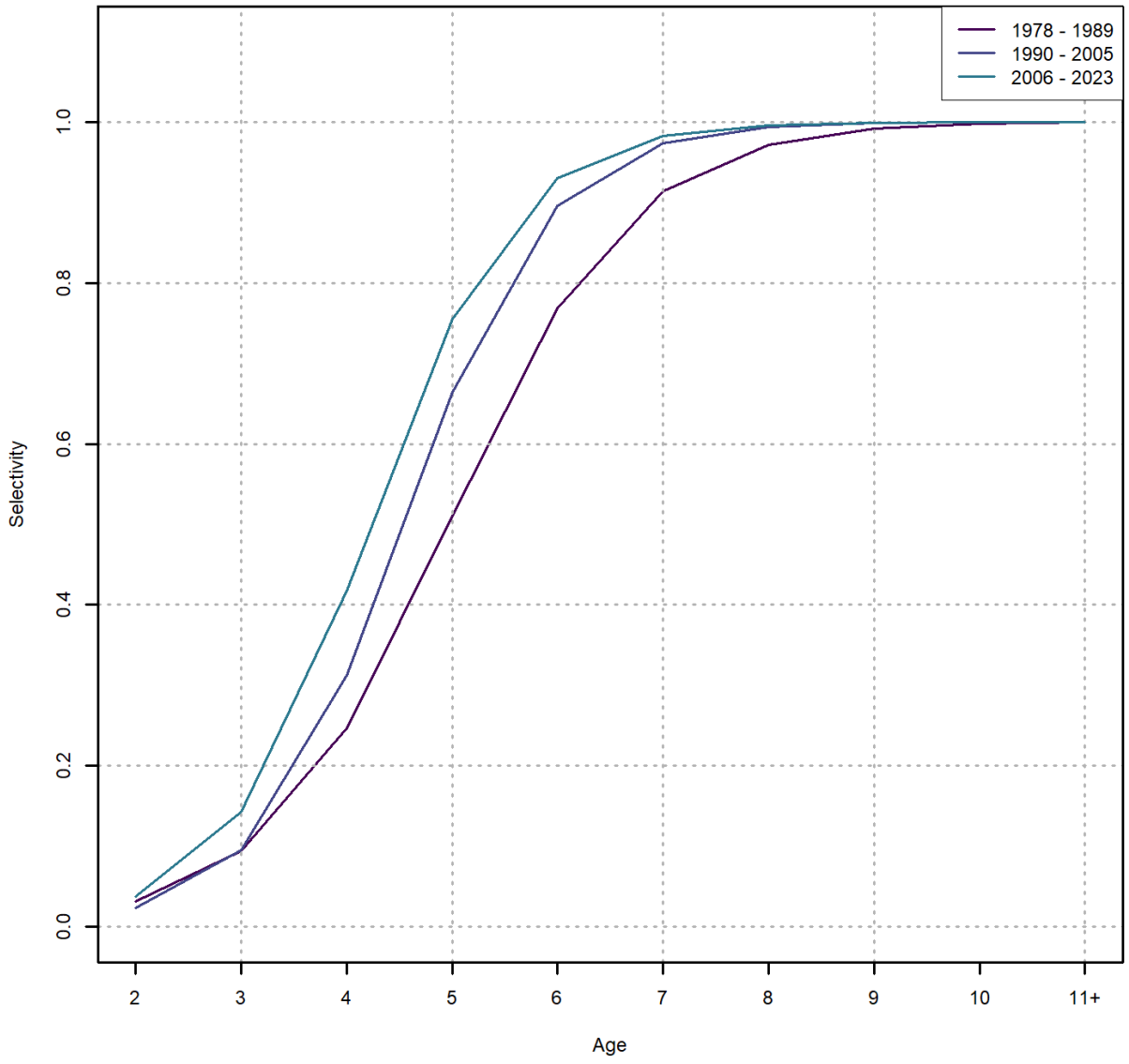

Figure 21: North model mean selectivity-at-age for the mobile gear fishery in three time blocks (1978-1989; 1990-2005; 2006-2023).

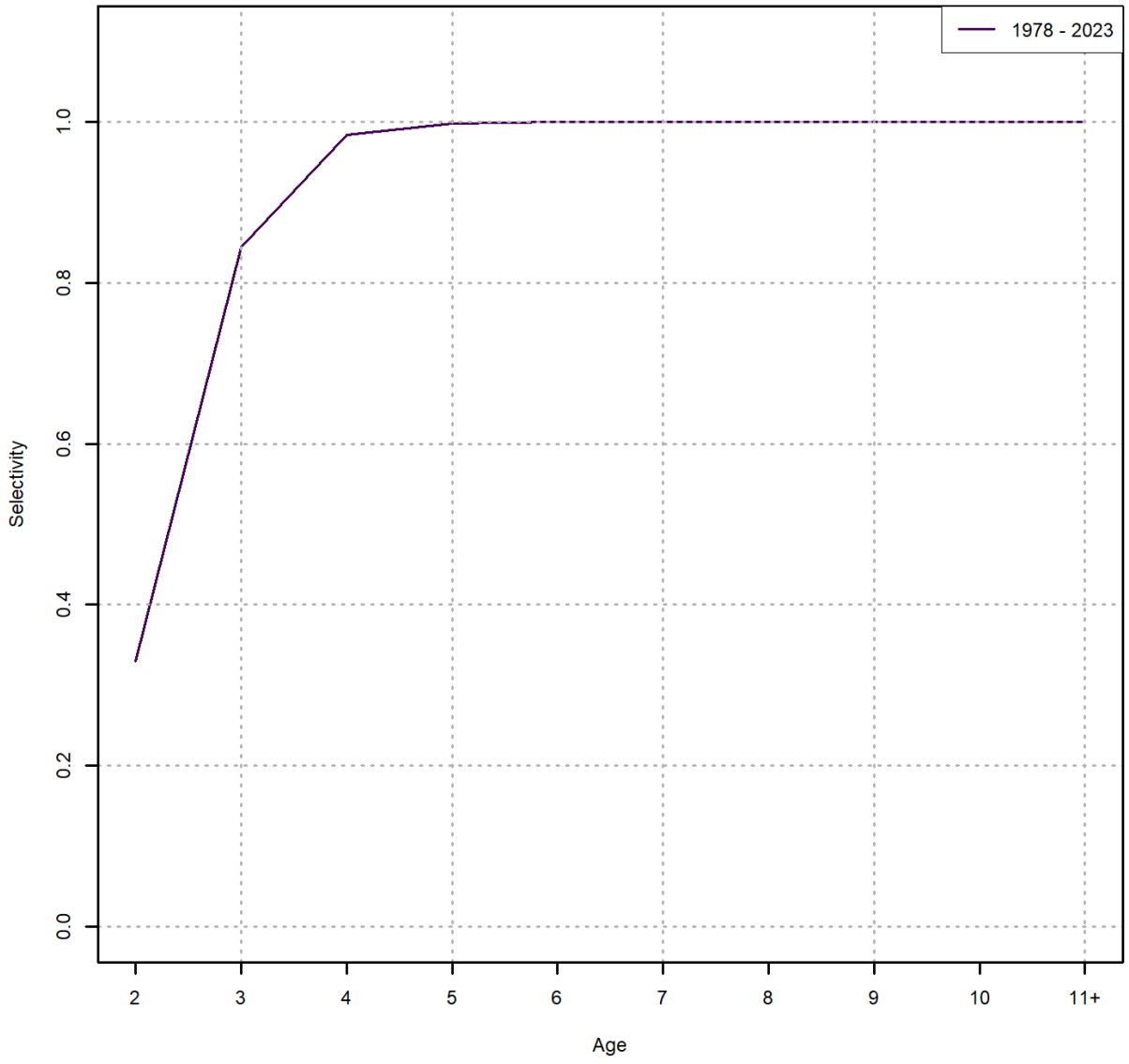

Figure 22: North model BDCM acoustic survey mean selectivity-at-age modeled as logistic curve.

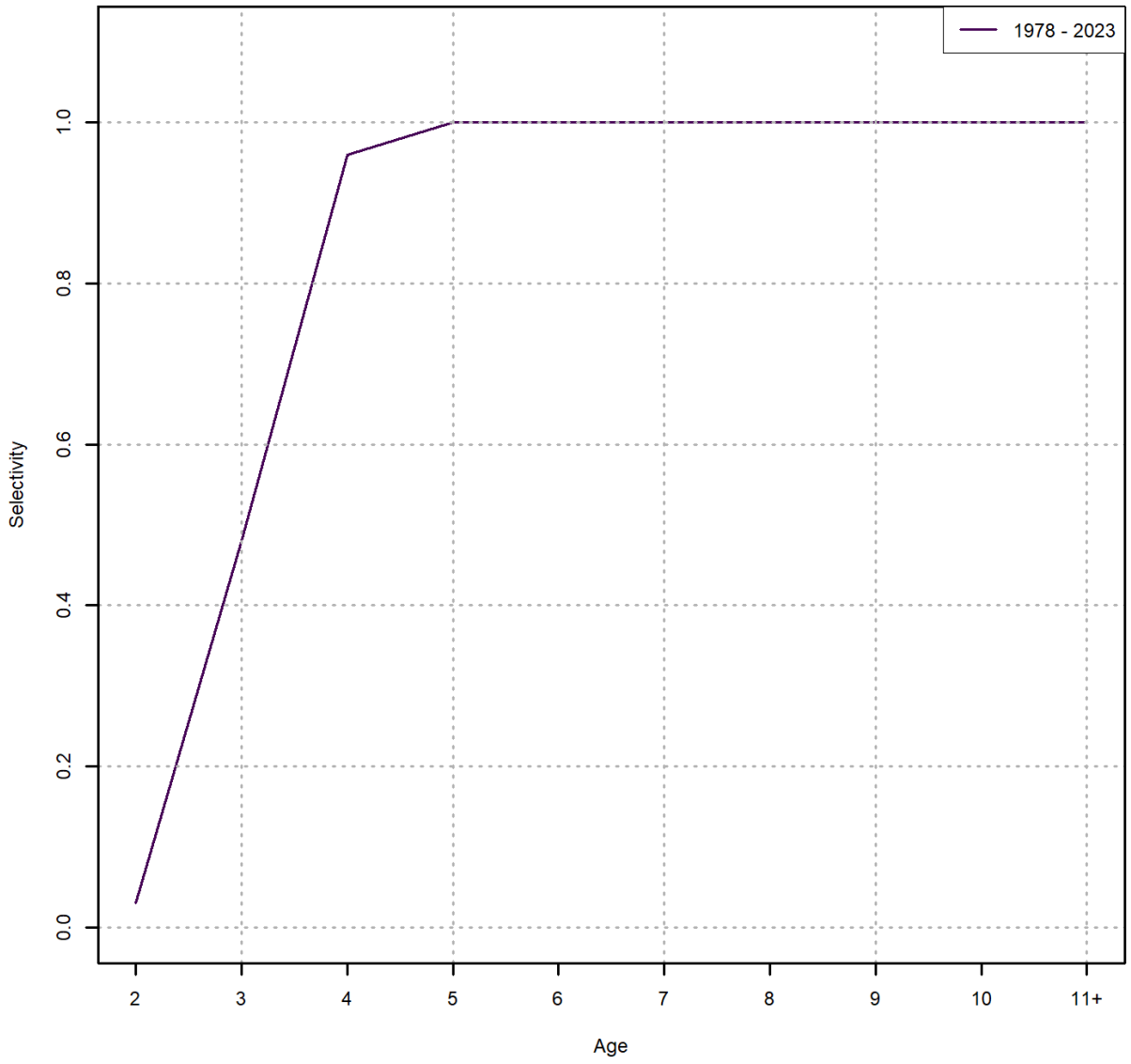

Figure 23: North model SG acoustic survey mean selectivity-at-age modeled as the mean maturity-at-age schedule.

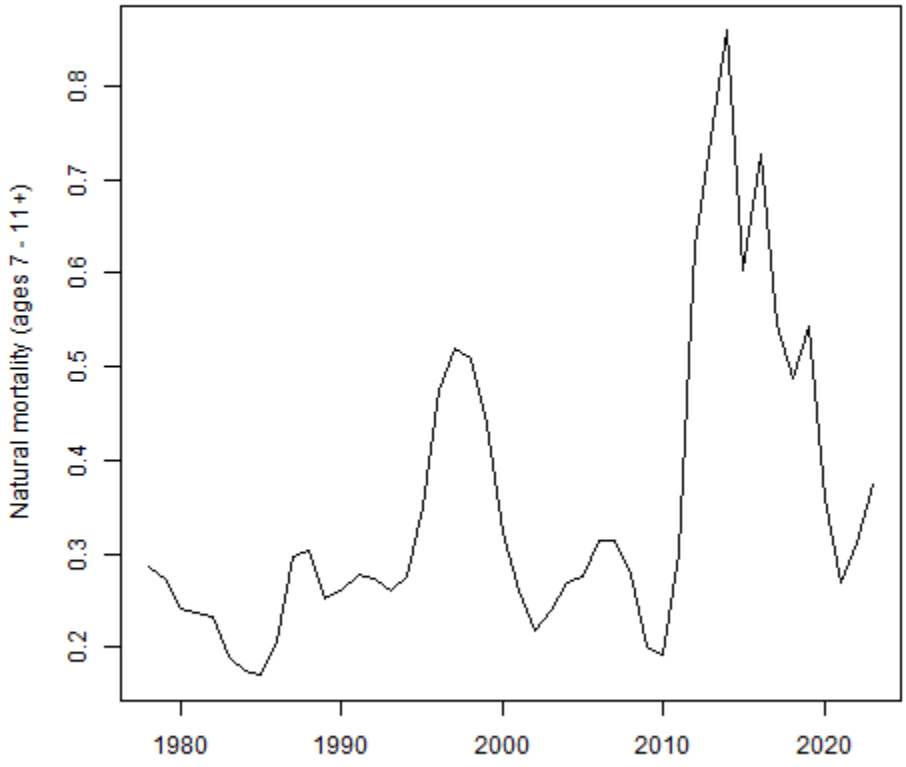

Figure 24: North model natural mortality for the age group 7 to 11+.

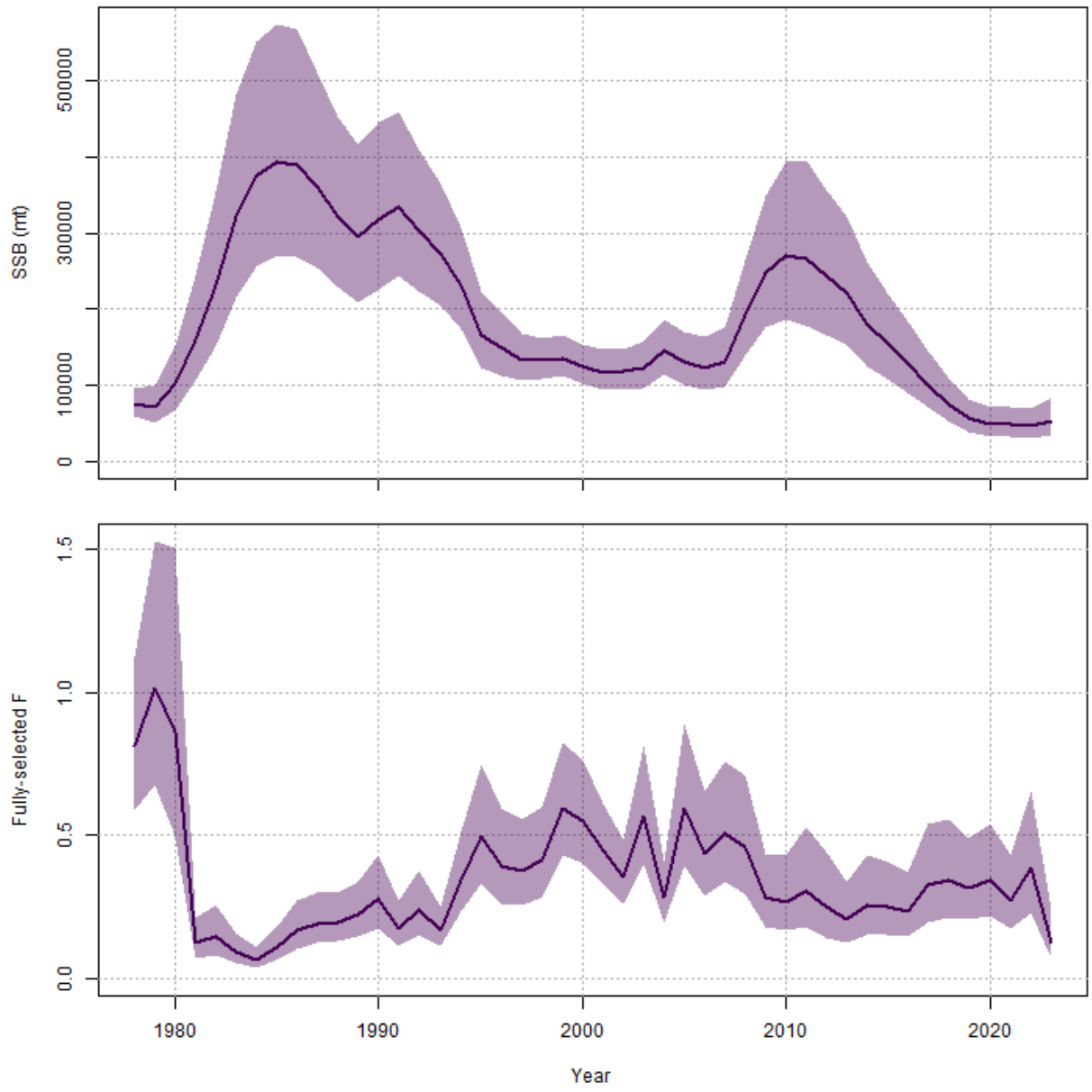

Figure 25: North model spawning stock biomass (SSB; upper panel) and fully-selected fishing mortality (F; lower panel) between 1978 and 2023. Solid lines are means and shading are 95% confidence intervals.

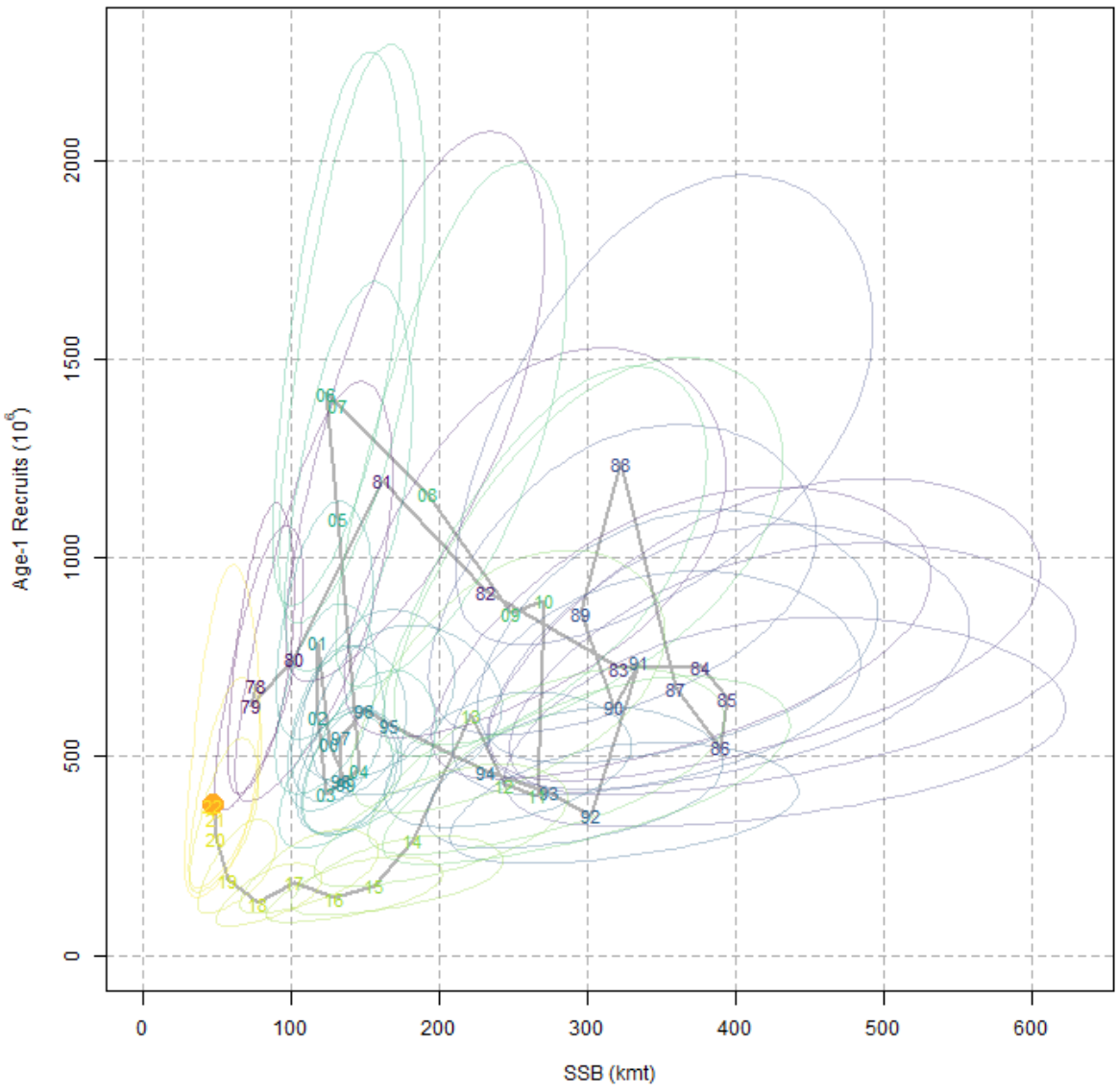

Figure 26: North model recruitment (y axis, numbers at age 2 in hundred thousands) and spawning stock biomass (x axis; SSB in metric kilotonnes). The solid line joins points by year, indicated by the label where the two digits represent the last two digits of the year. Large circles are uncertainty around SSB and recruitment estimates.

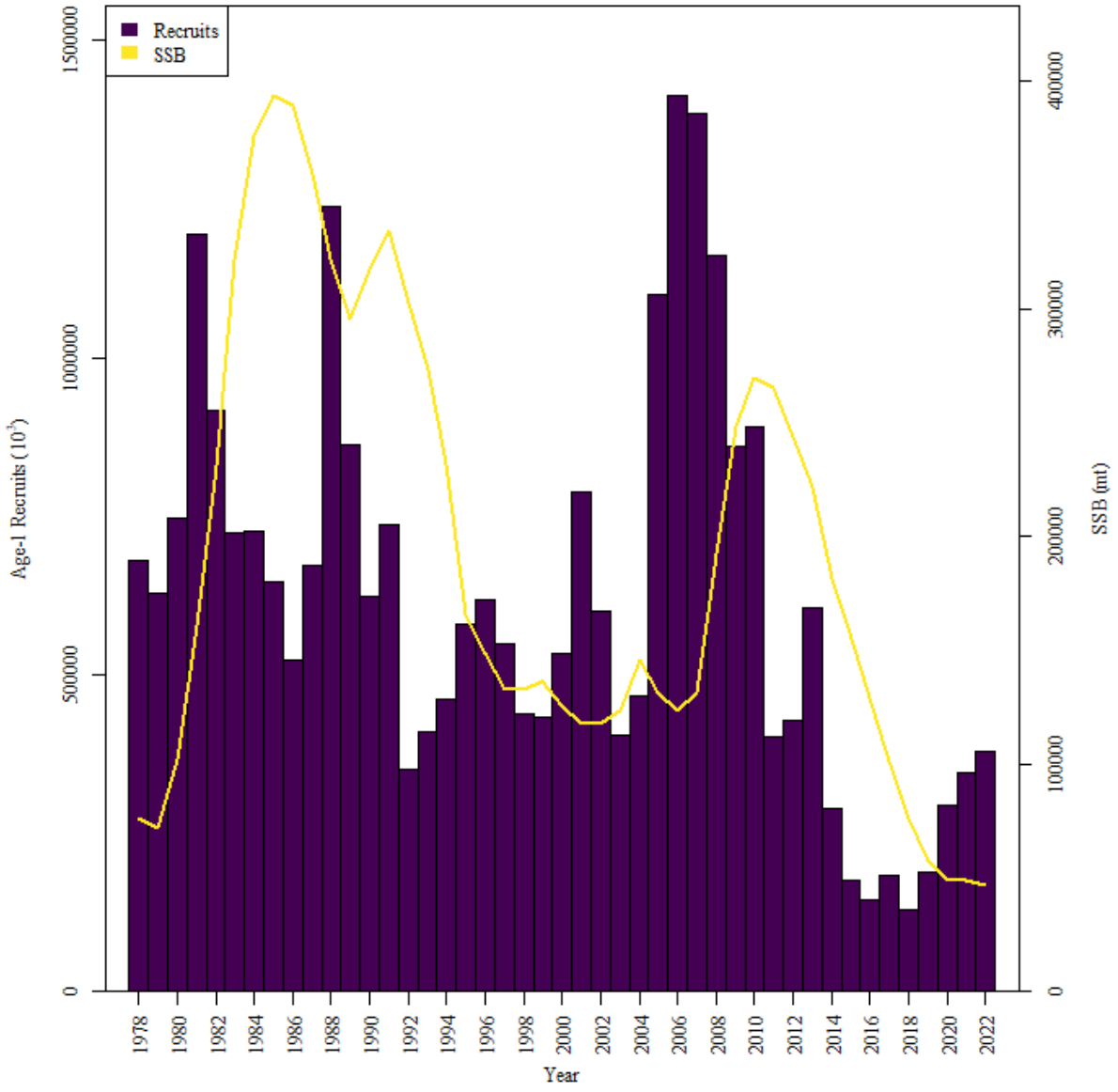

Figure 27: North model recruitment (left axis, purple bars, numbers at age 2 in thousands) and spawning stock biomass (right axis; yellow line, SSB) between 1978 and 2022.

MIDDLE MODEL

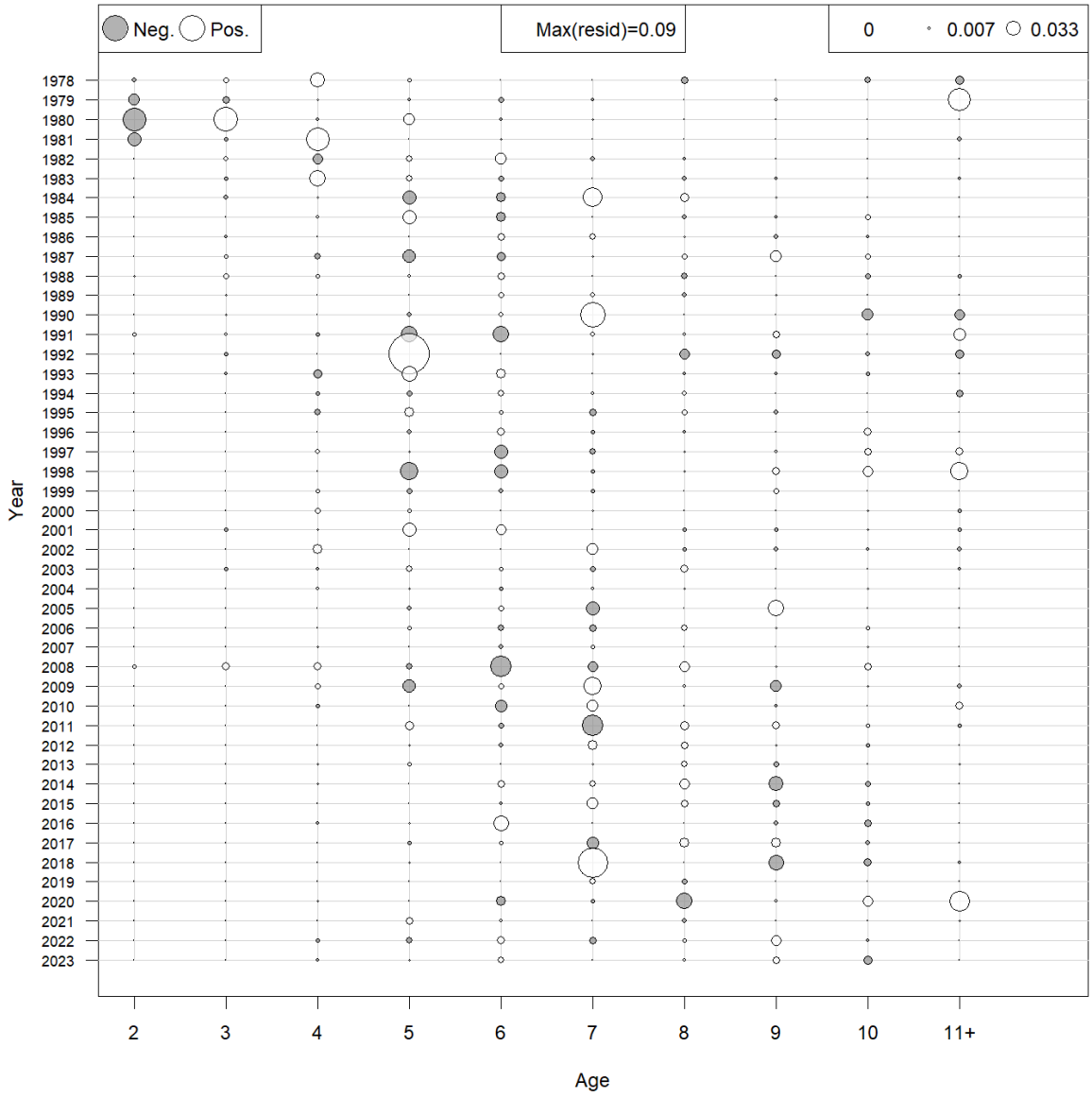

Figure 28: Middle model fixed gear fishery age composition residuals (observed-predicted) between 1978 and 2023. White circles are positive residuals and grey circles are negative residuals. Circle size is proportional to the residual value.

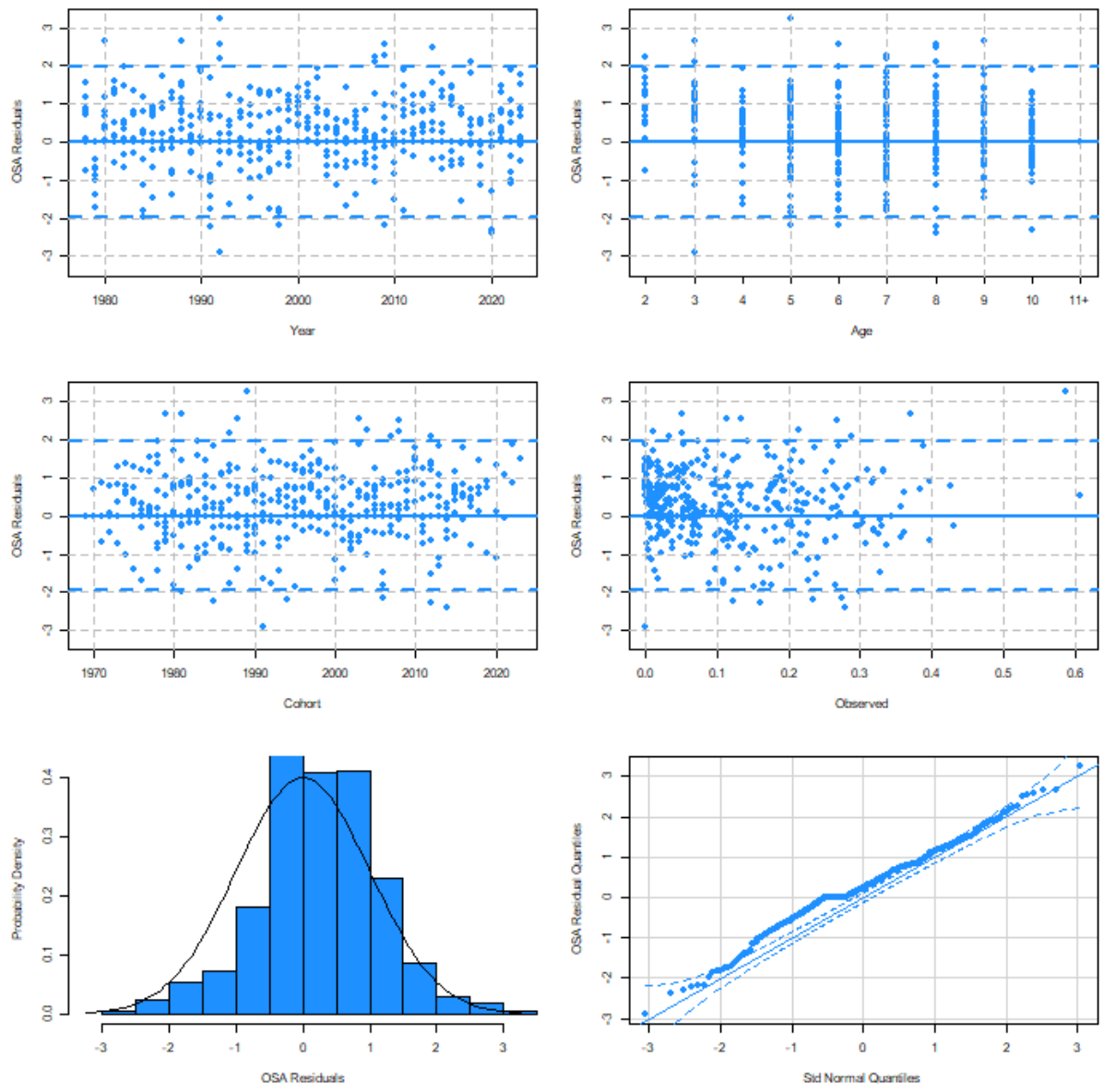

Figure 29: Middle model fixed gear age composition one-step-ahead (OSA) residuals by (panels from left to right top to bottom) year, age, cohort and by observed value, probability density and versus normal quantiles.

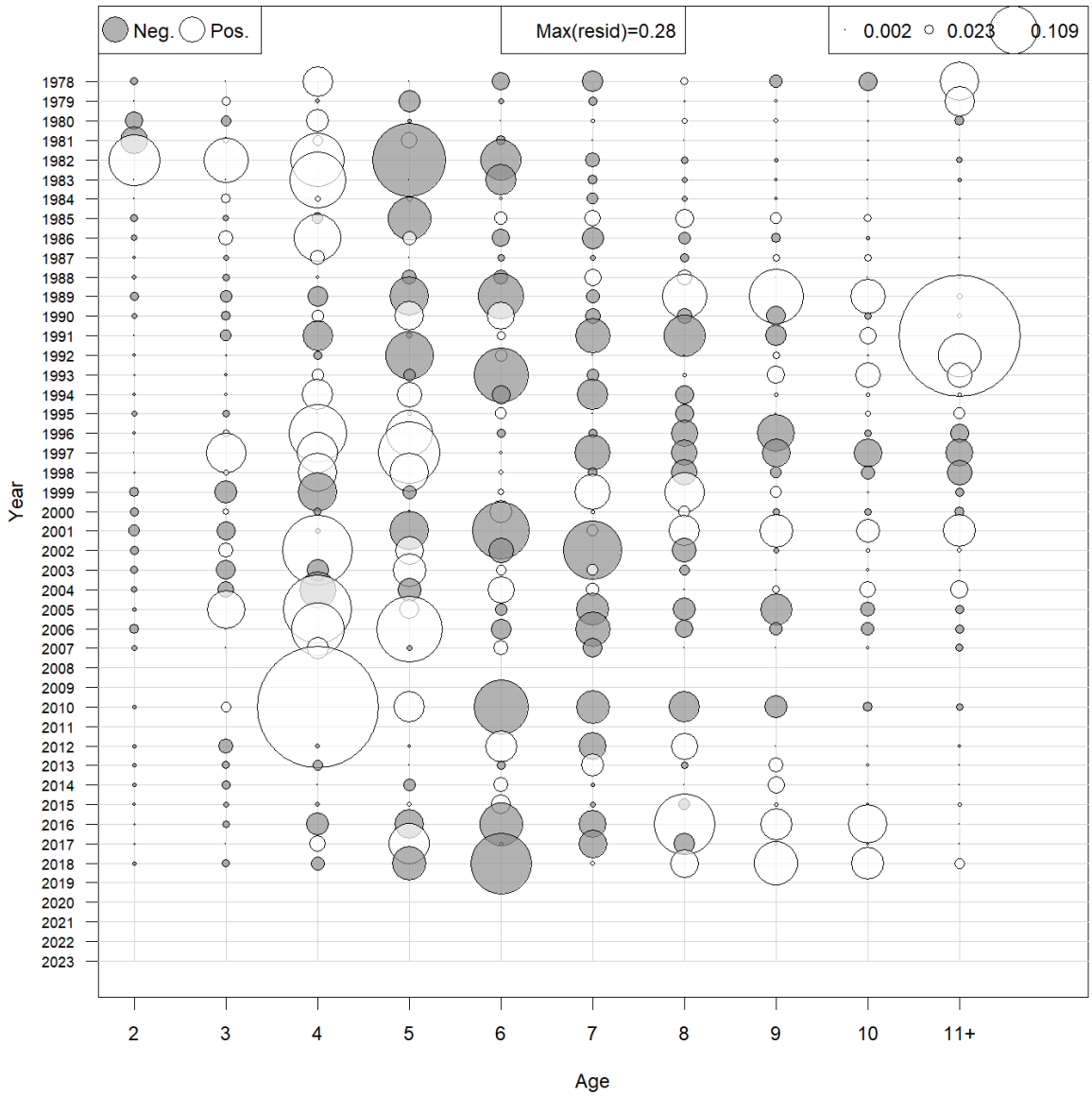

Figure 30: Middle model mobile gear fishery age composition raw residuals (observed-predicted) between 1978 and 2023. White circles are positive residuals and grey circles are negative residuals. Circle size is proportional to the residual value.

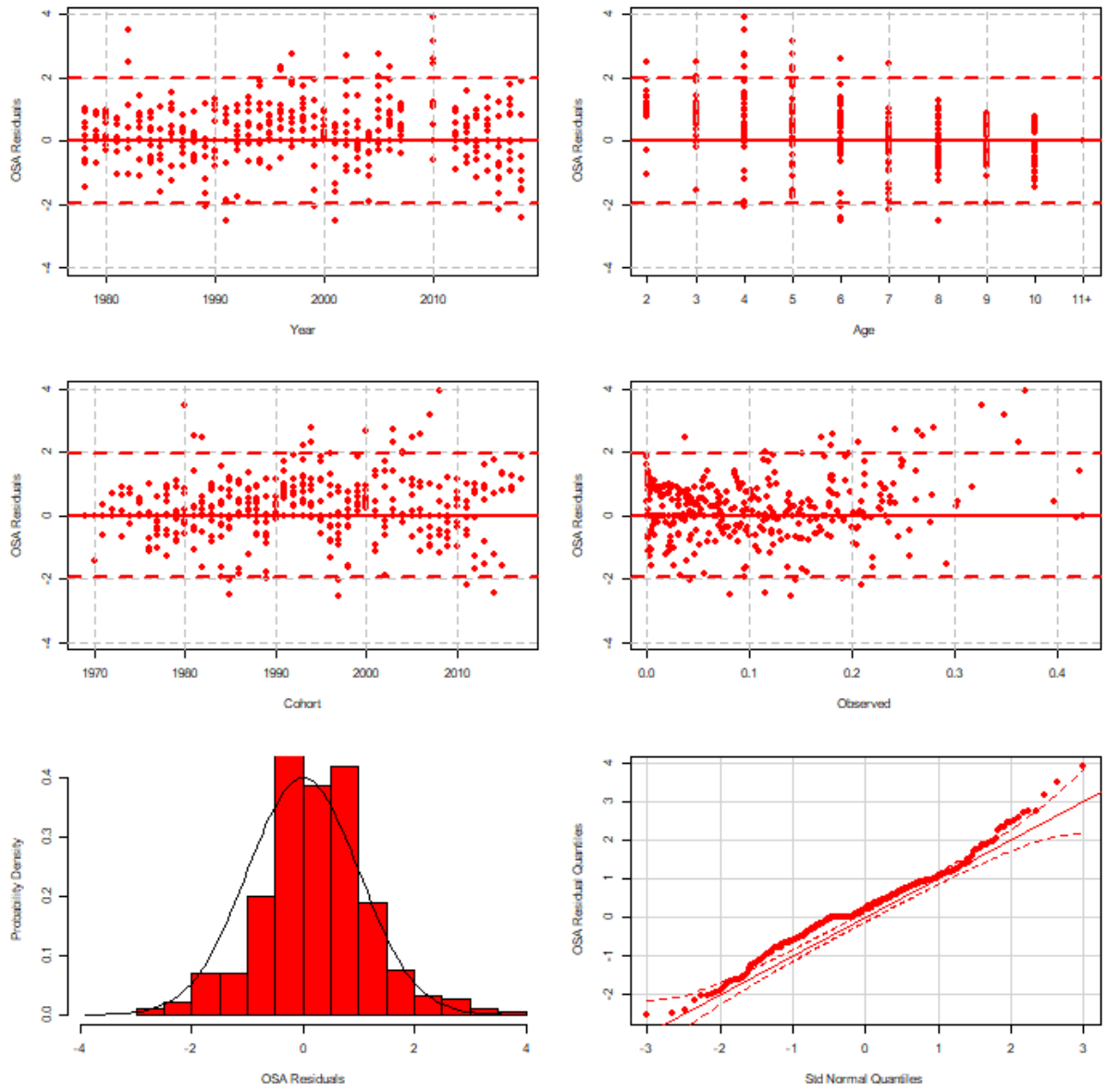

Figure 31: Middle model mobile gear age composition one-step-ahead (OSA) residuals by (panels from left to right top to bottom) year, age, cohort and by observed value, probability density and versus normal quantiles.

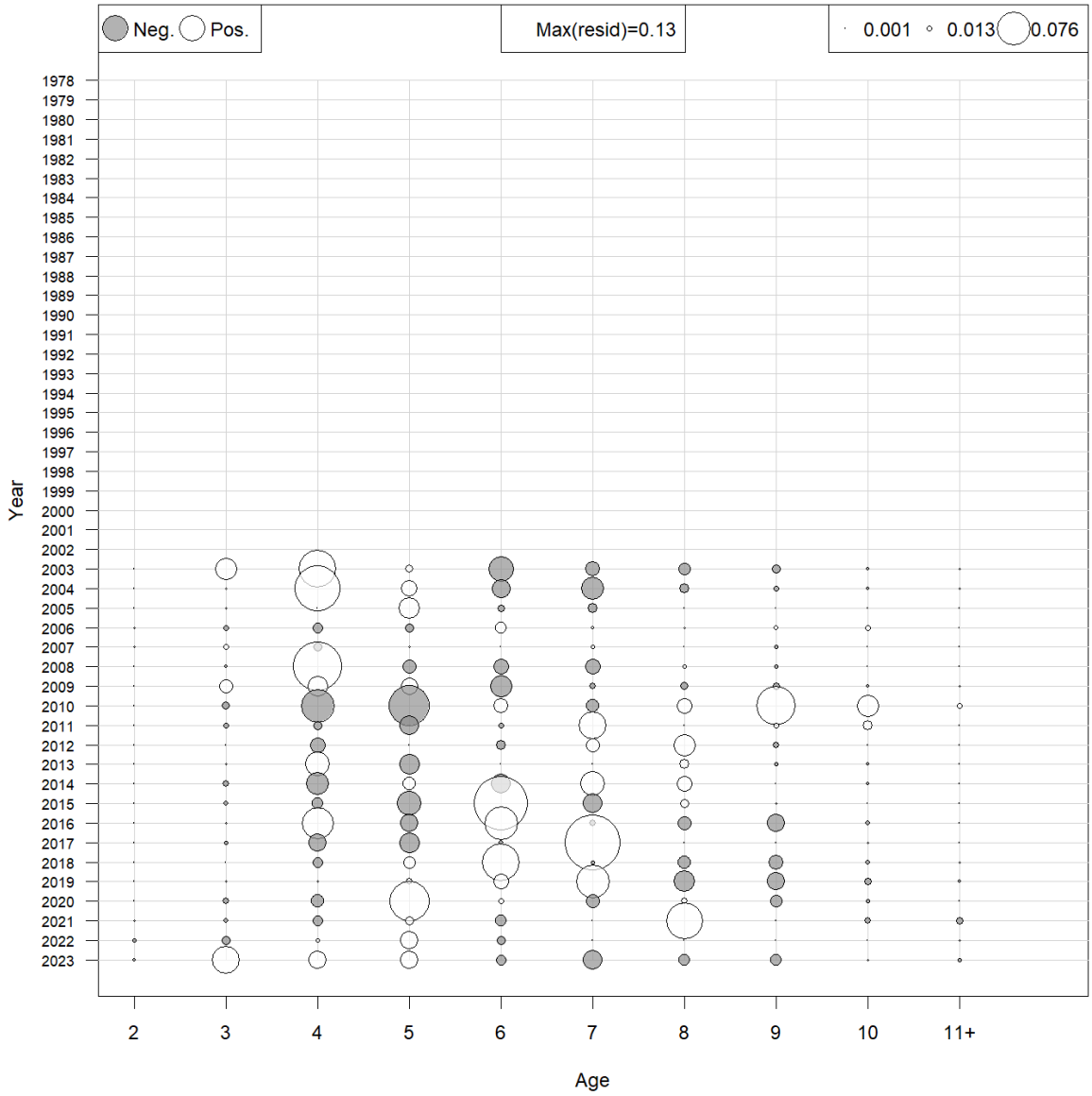

Figure 32: Middle model experimental nets survey age composition raw residuals (observed-predicted) between 1978 and 2023. White circles are positive residuals and grey circles are negative residuals. Circle size is proportional to the residual value.

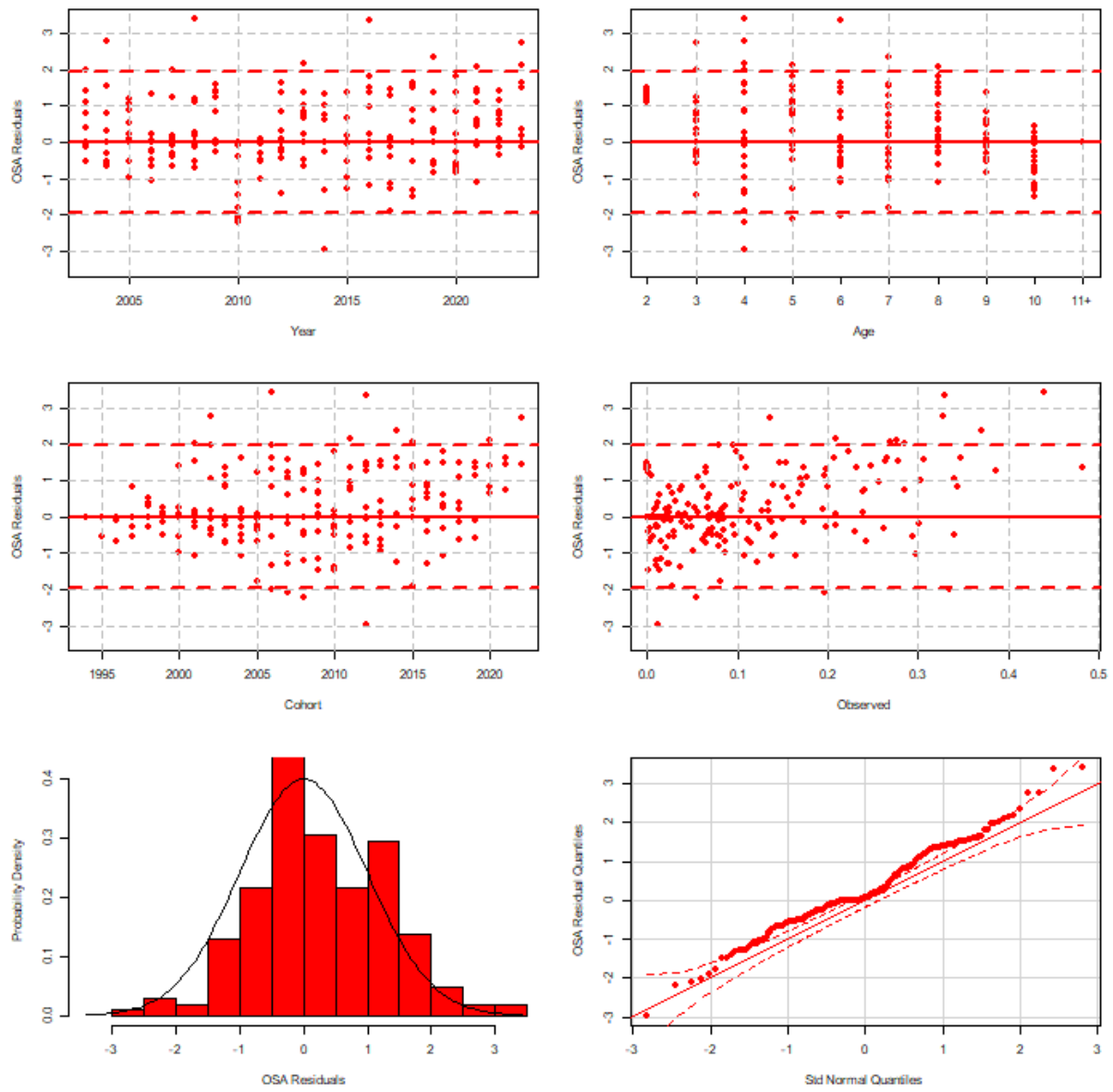

Figure 33: Middle model experimental nets survey age composition one-step-ahead (OSA) residuals by (panels from left to right top to bottom) year, age, cohort and by observed value, probability density and versus normal quantiles.

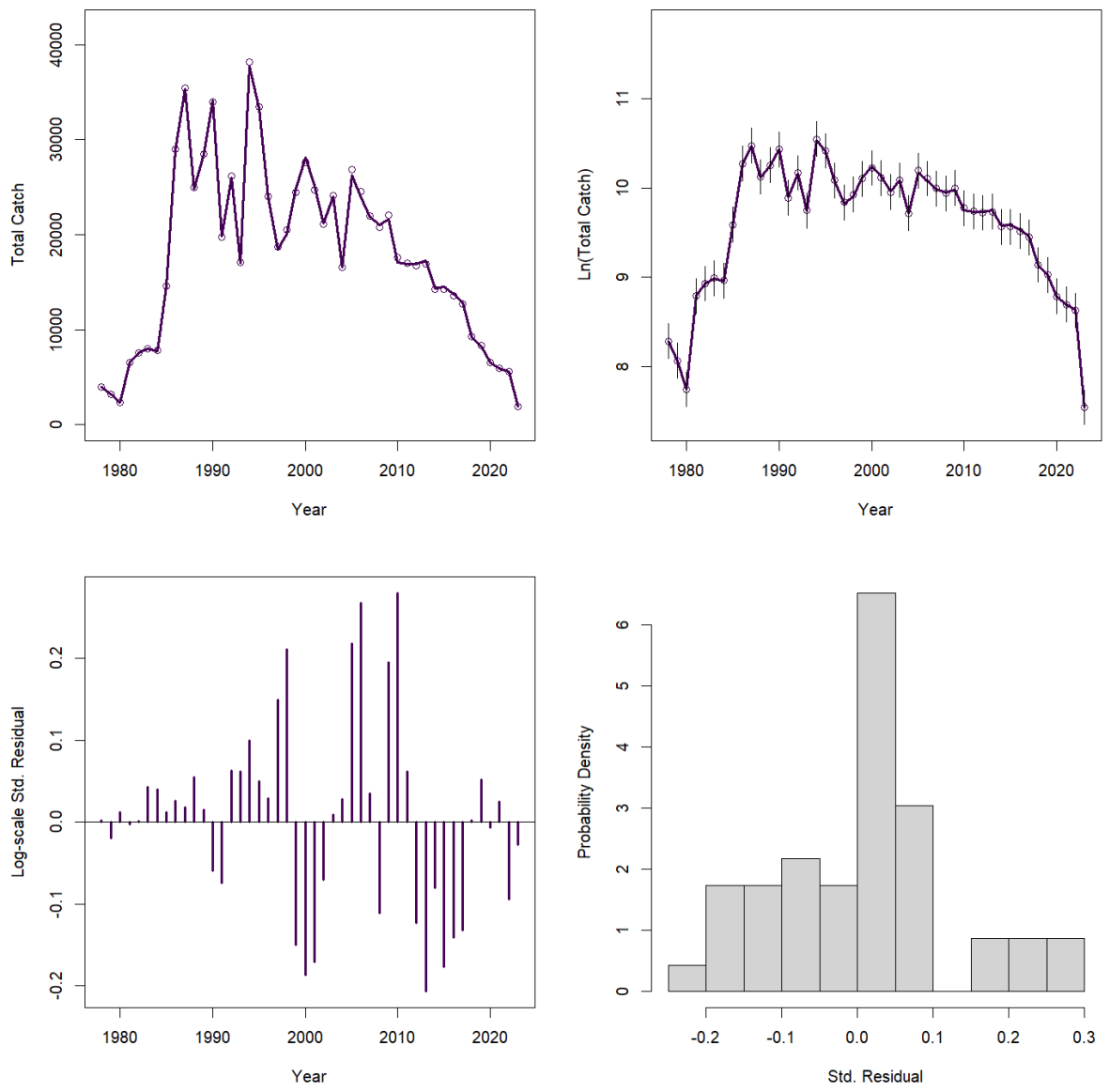

Figure 34: Middle model fit to the fixed gear total catch (top left), and to the log of the catch (top right), standardized residuals by year (bottom left) probability density of standardized residuals (bottom right).

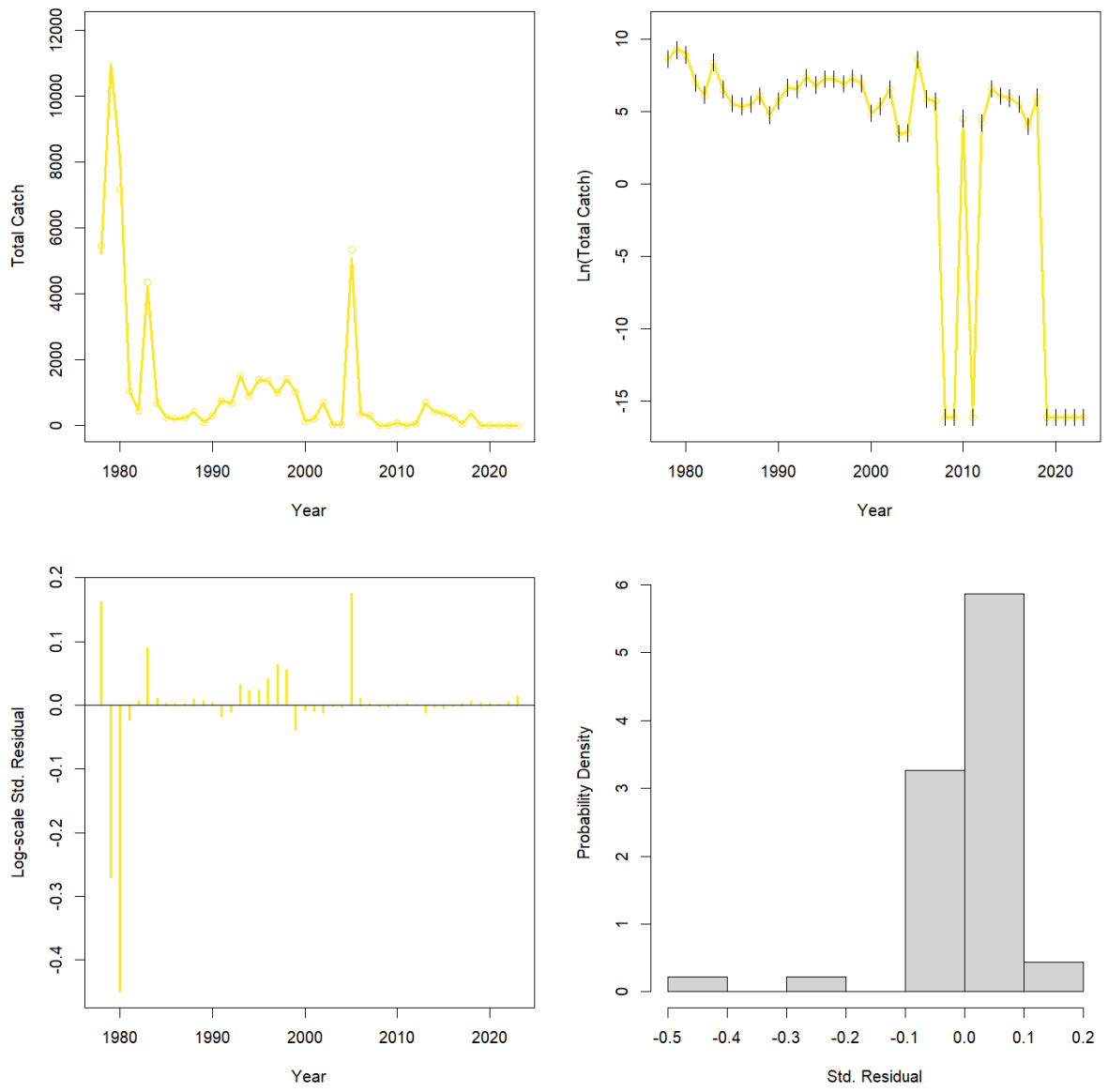

Figure 35: Middle model fit to the mobile gear total catch (top left), and to the log of the catch (top right), standardized residuals by year (bottom left) probability density of standardized residuals (bottom right).

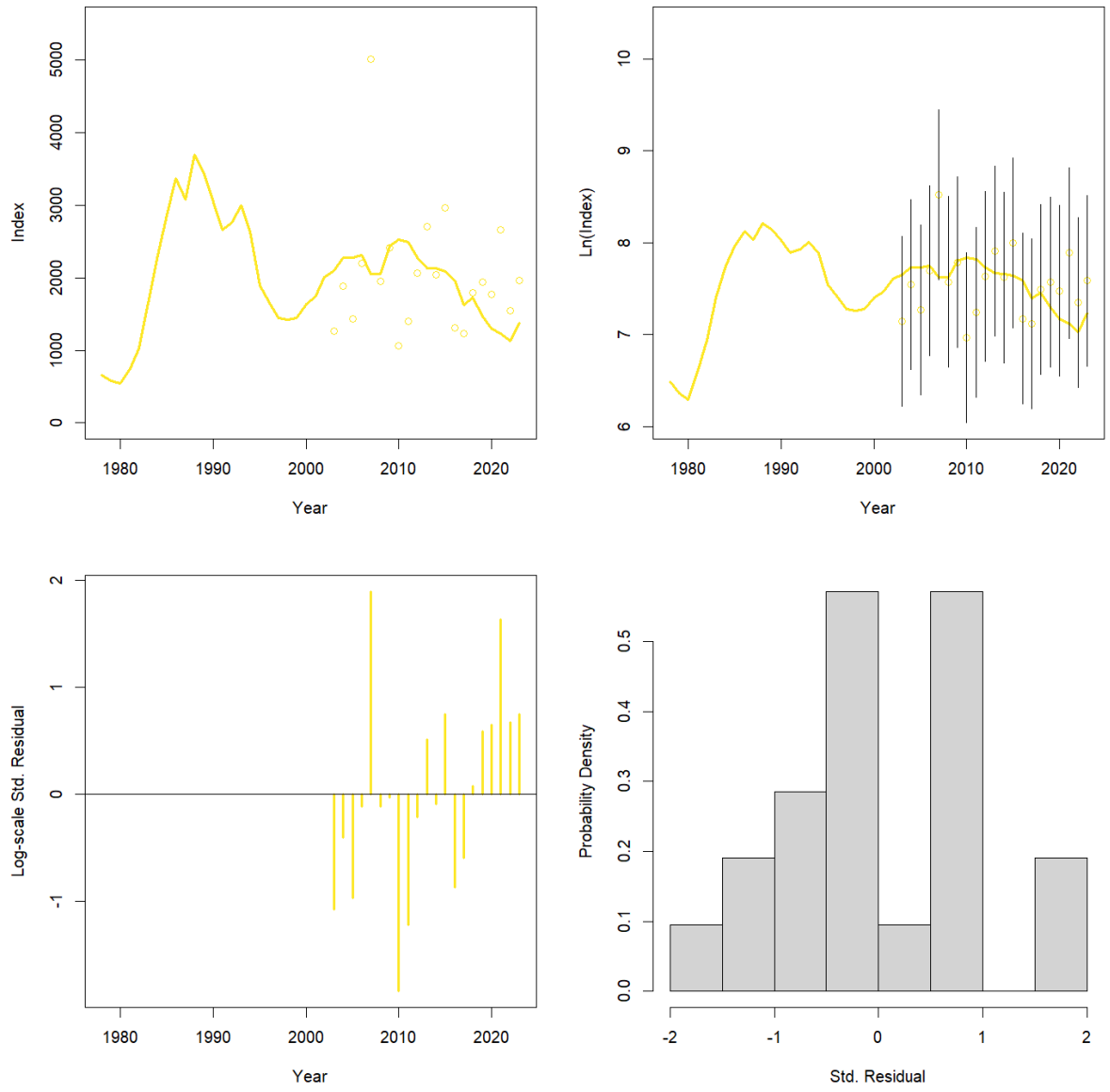

Figure 36: Middle model fit to the experimental nets survey biomass index (top left), and to the log of the index (top right), standardized residuals by year (bottom left) probability density of standardized residuals (bottom right).

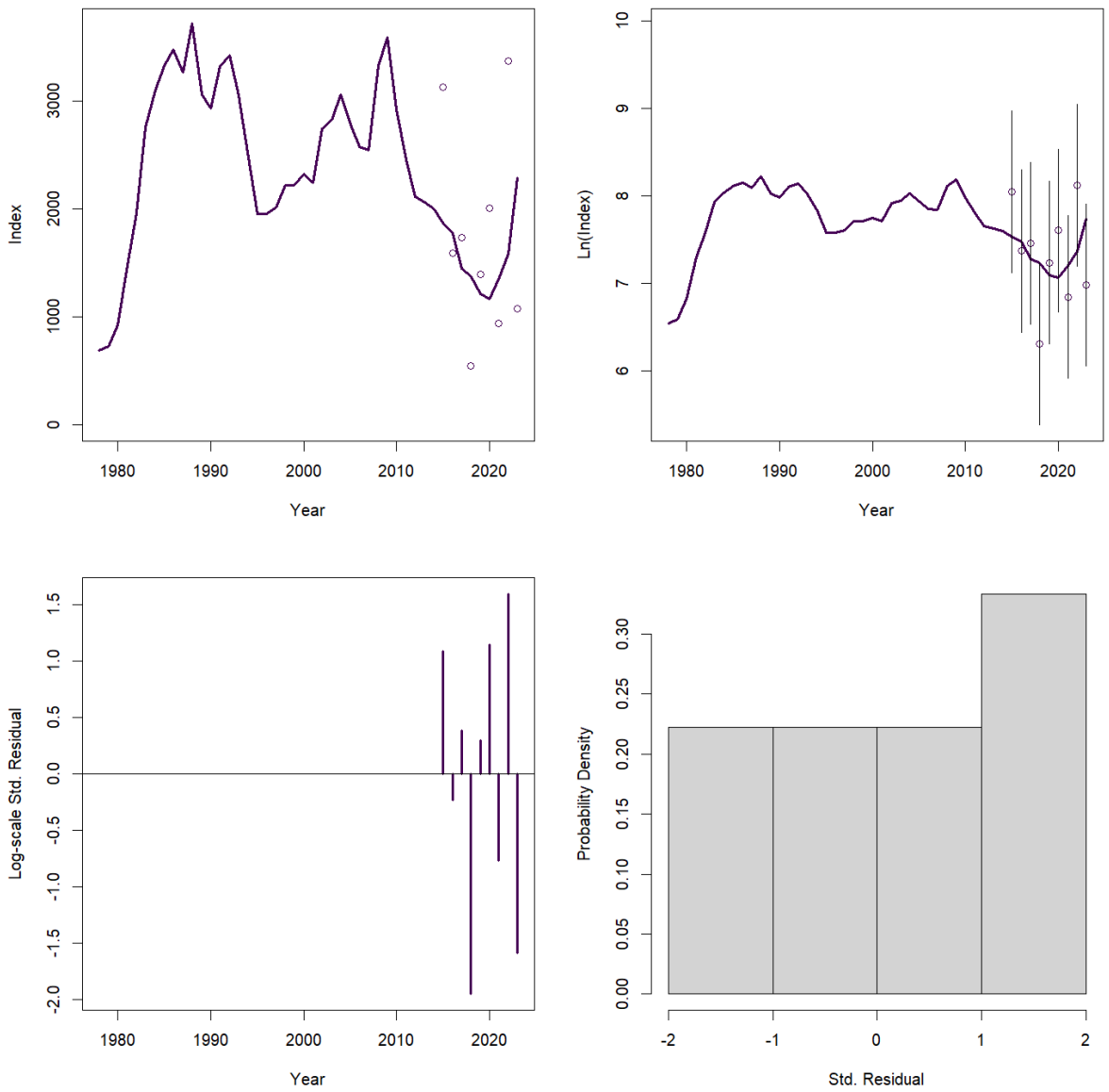

Figure 37: Middle model fit to the SG acoustic survey biomass index (top left), and to the log of the index (top right), standardized residuals by year (bottom left) probability density of standardized residuals (bottom right).

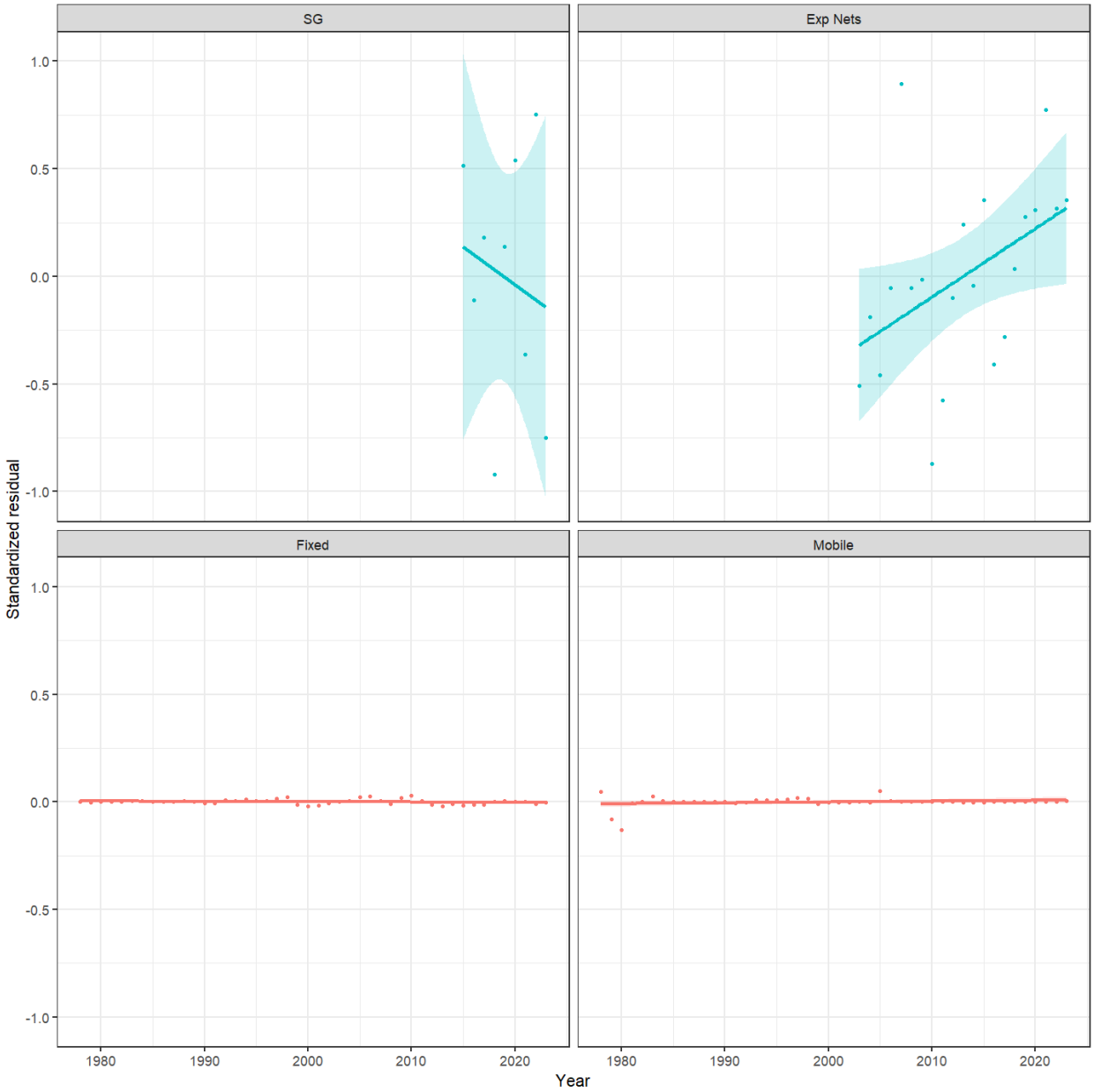

Figure 38: Middle model standardized residuals by year for the SG acoustic survey, experimental nets survey, fixed gear catch and mobile gear catch.

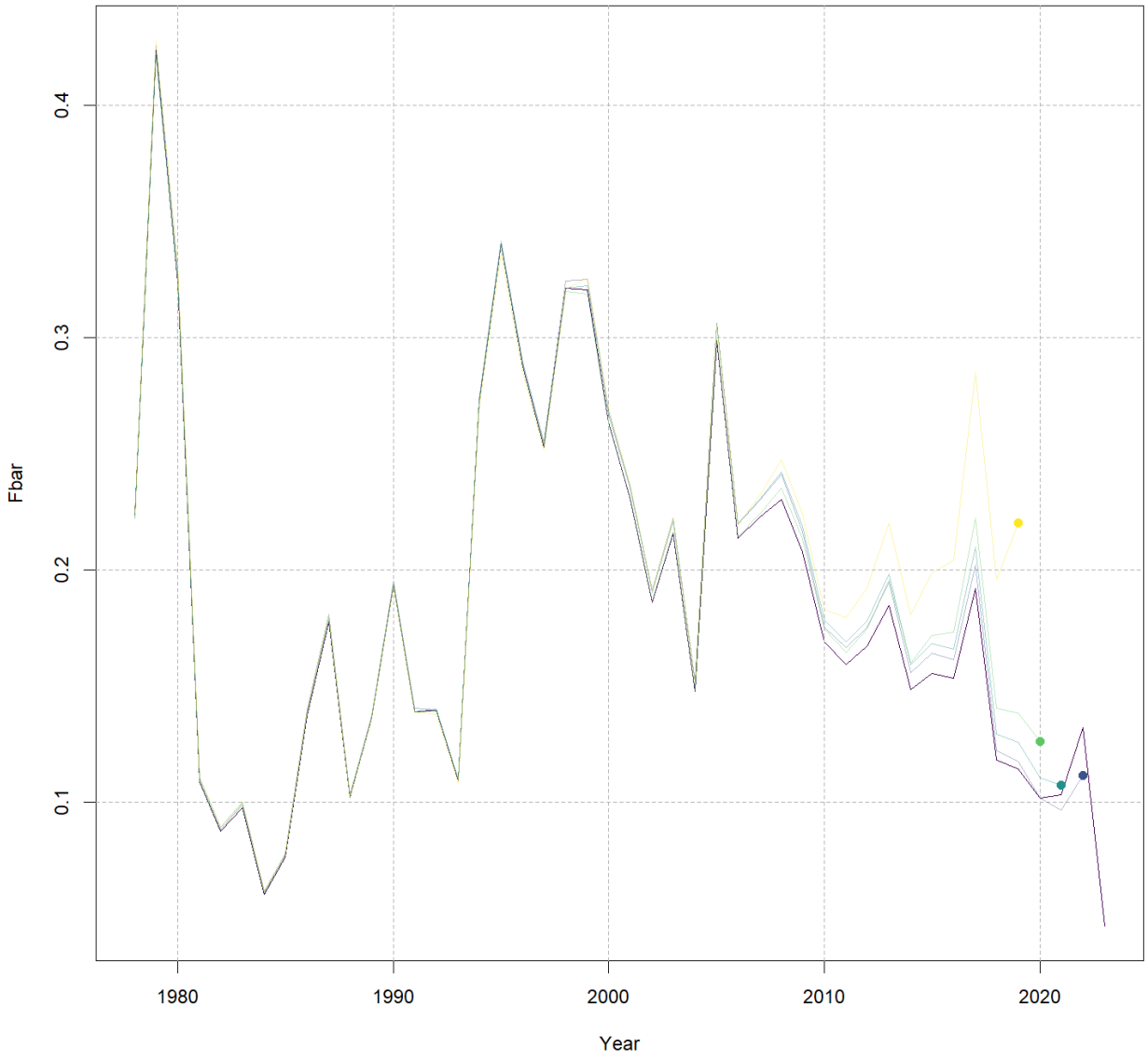

Figure 39: Middle model retrospective analysis for average fishing mortality (Fbar) for 4 peels.

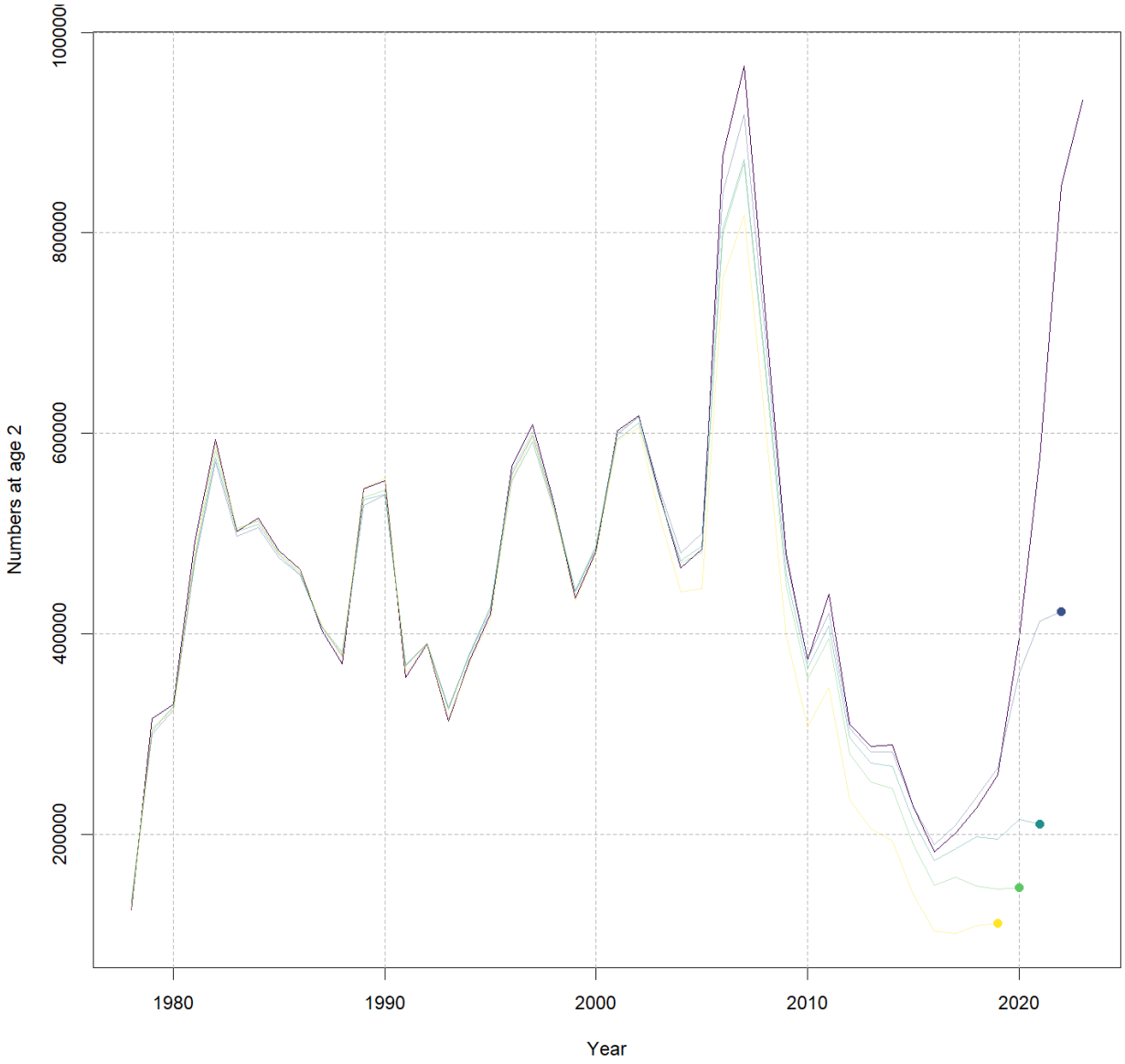

Figure 40: Middle model retrospective analysis for recruitment (numbers at age 2) for 4 peels.

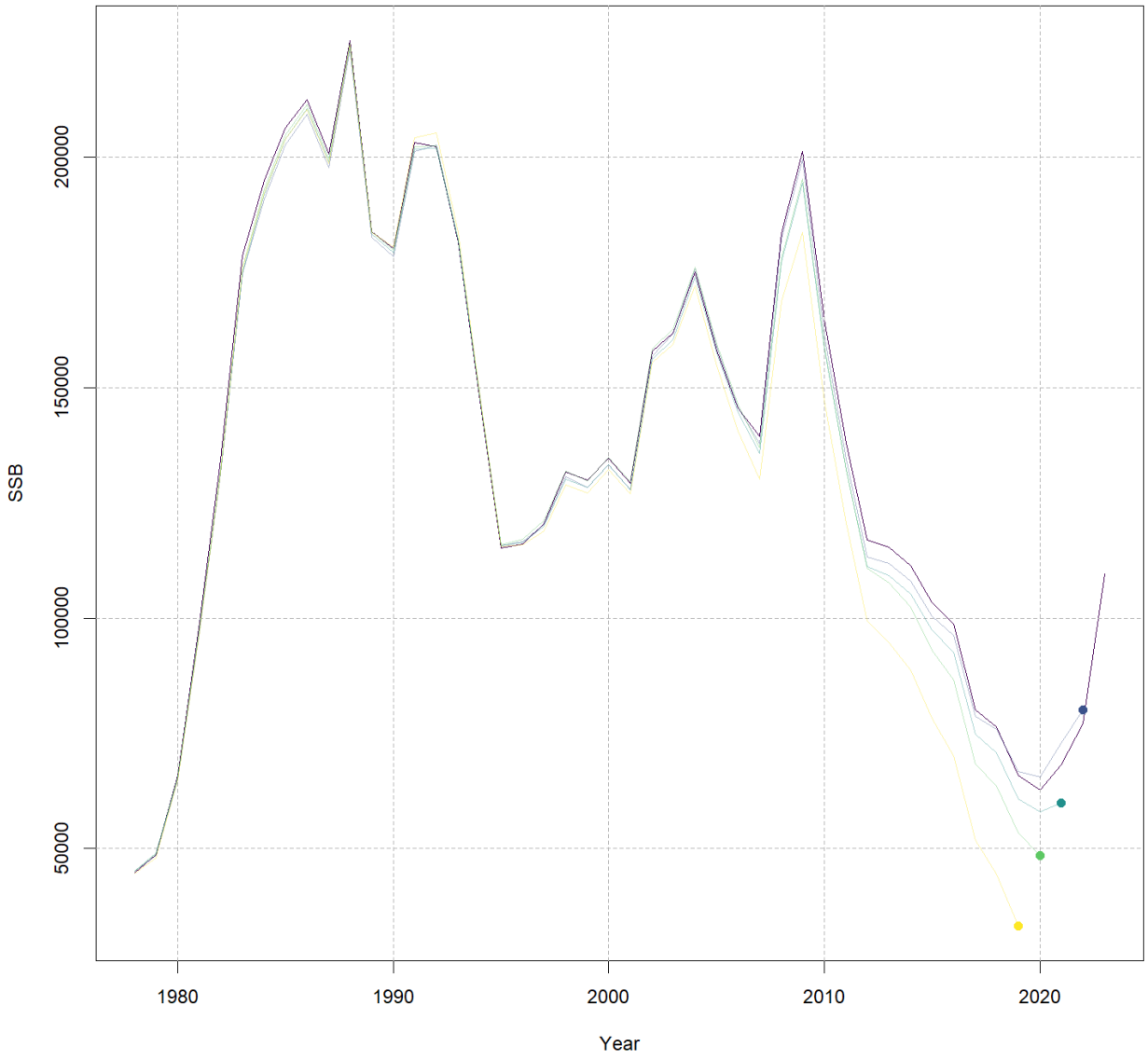

Figure 41: Middle model retrospective analysis for spawning stock biomass (SSB; tonnes) for 4 peels.

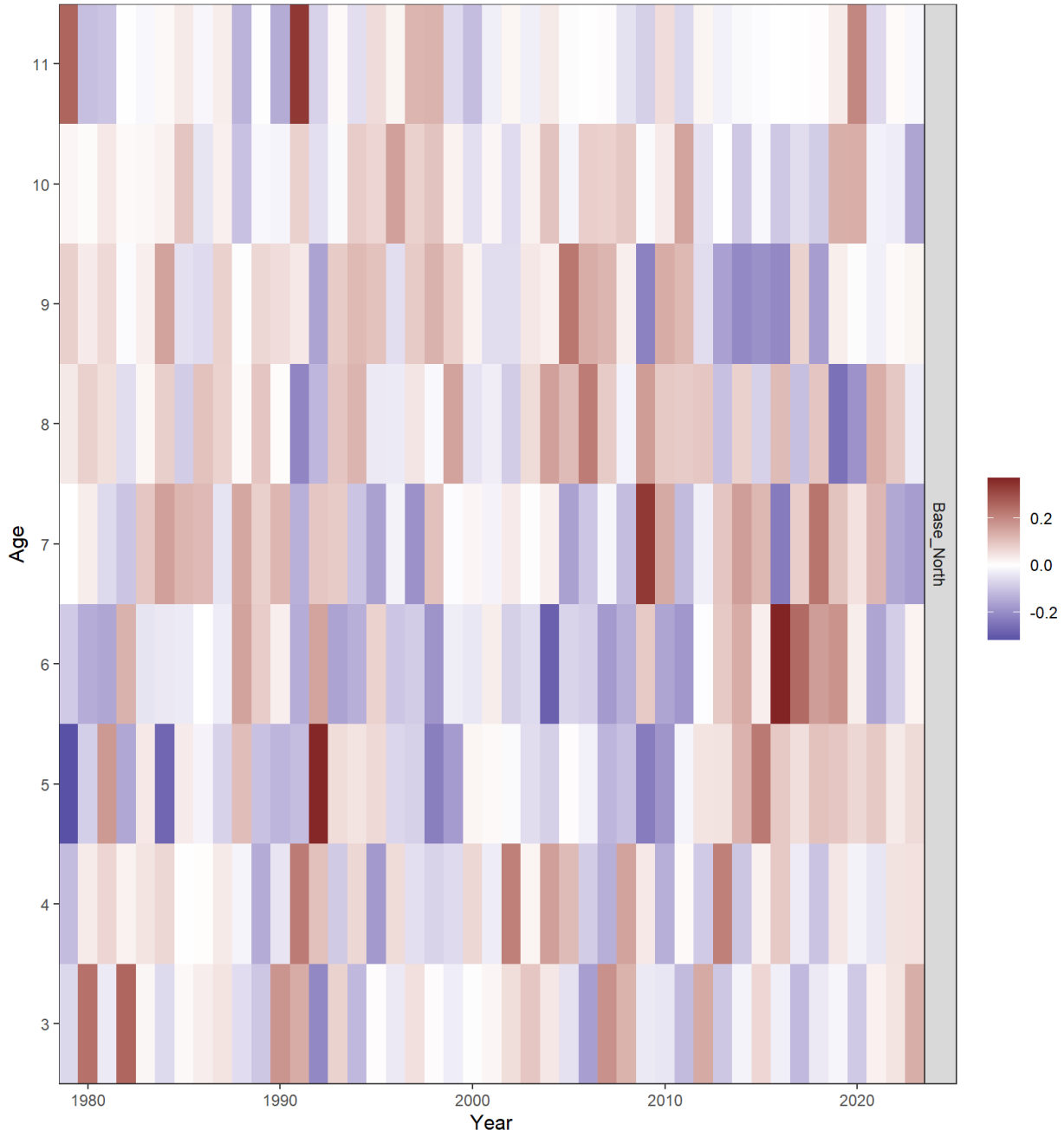

Figure 42: Middle model deviations in numbers-at-age (NAA) by age and year, for ages 3 to 11+.

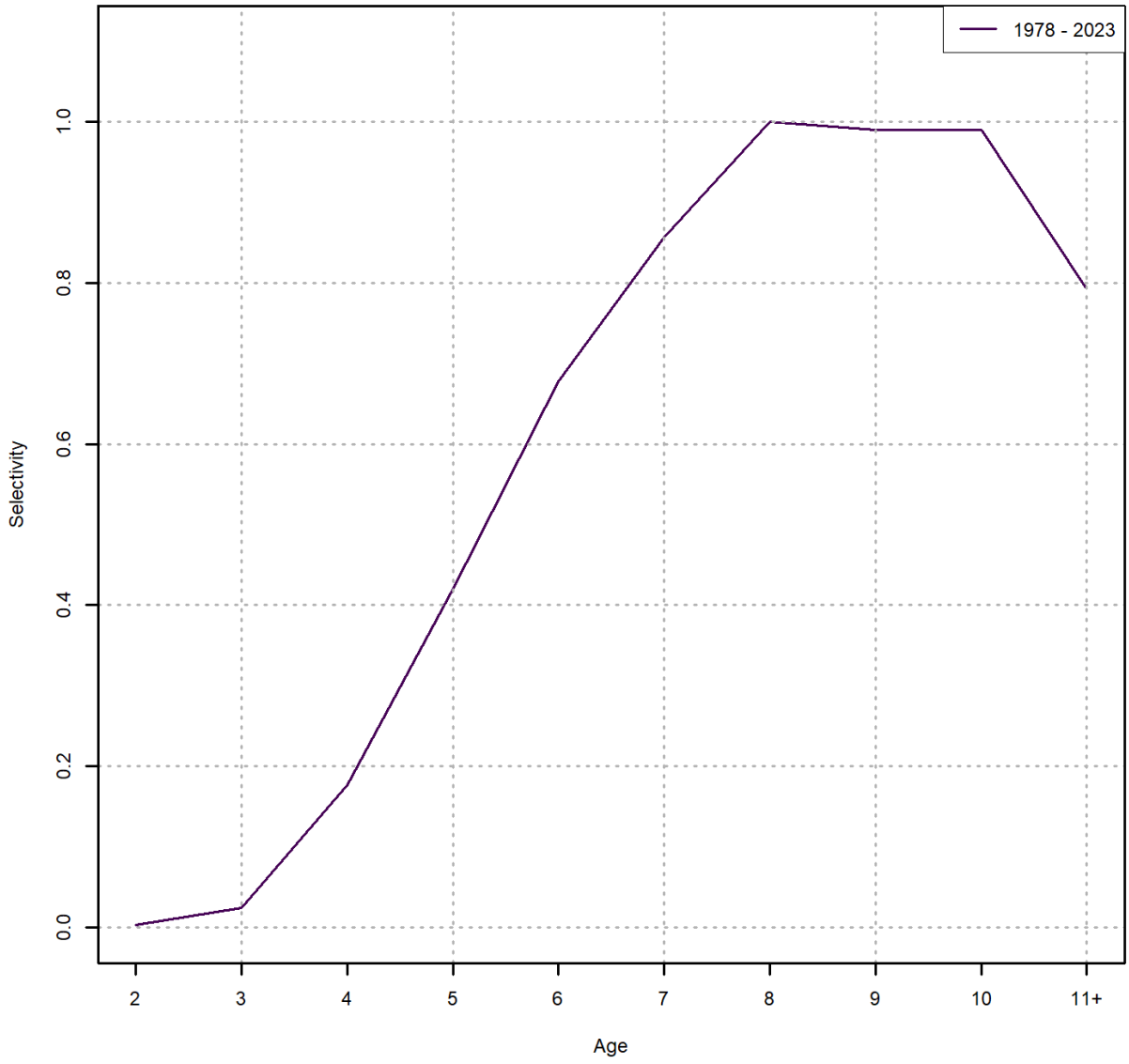

Figure 43: Middle model mean selectivity-at-age for the fixed gear fishery.

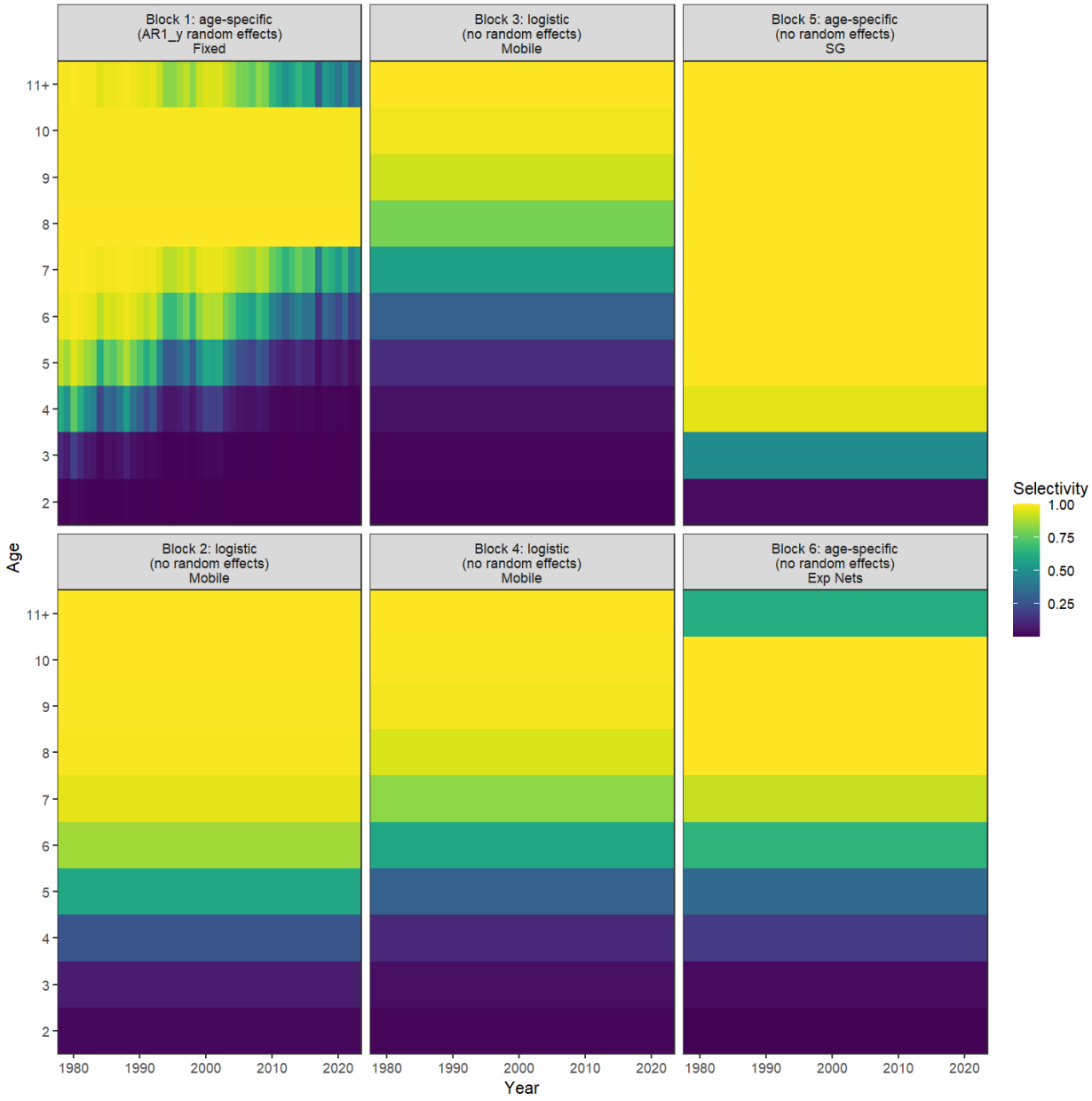

Figure 44: Middle model realized selectivity including random effects were applicable, for all sources of catch, surveys and time blocks.

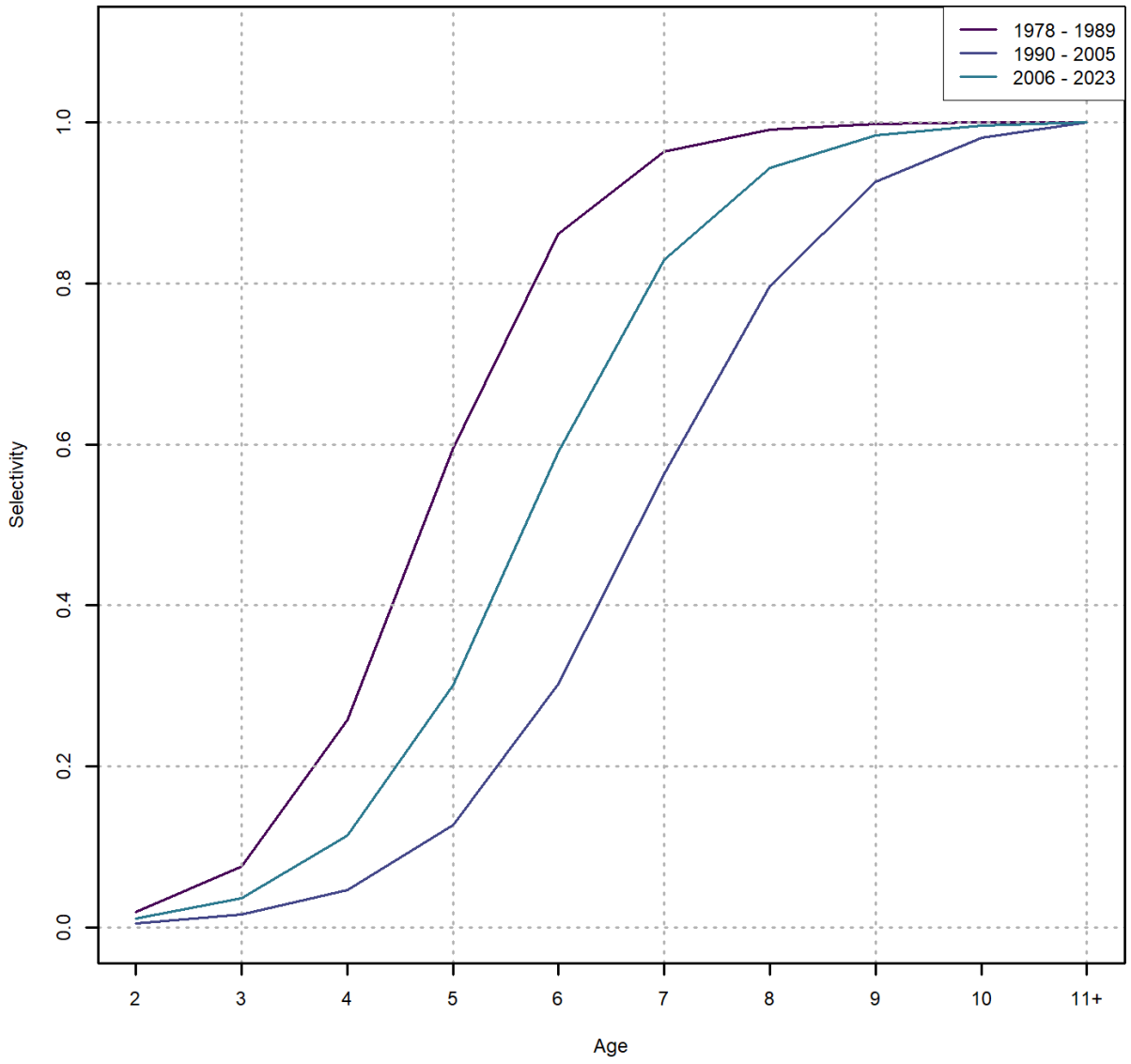

Figure 45: Middle model mean selectivity-at-age for the mobile gear fishery in three time blocks (1978-1989; 1990-2005; 2006-2023).

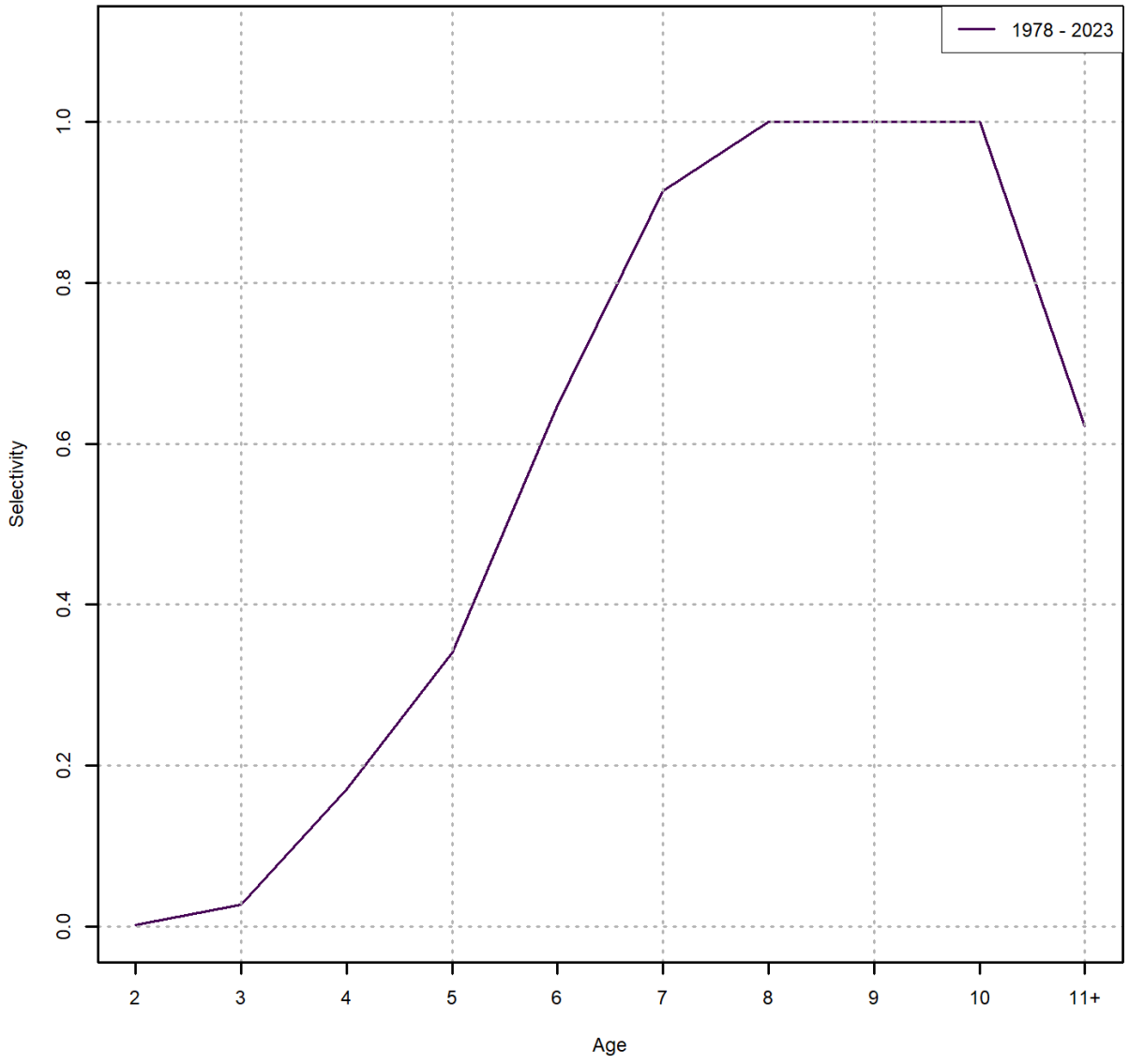

Figure 46: Middle model experimental gillnets survey mean selectivity-at-age modeled as logistic curve.

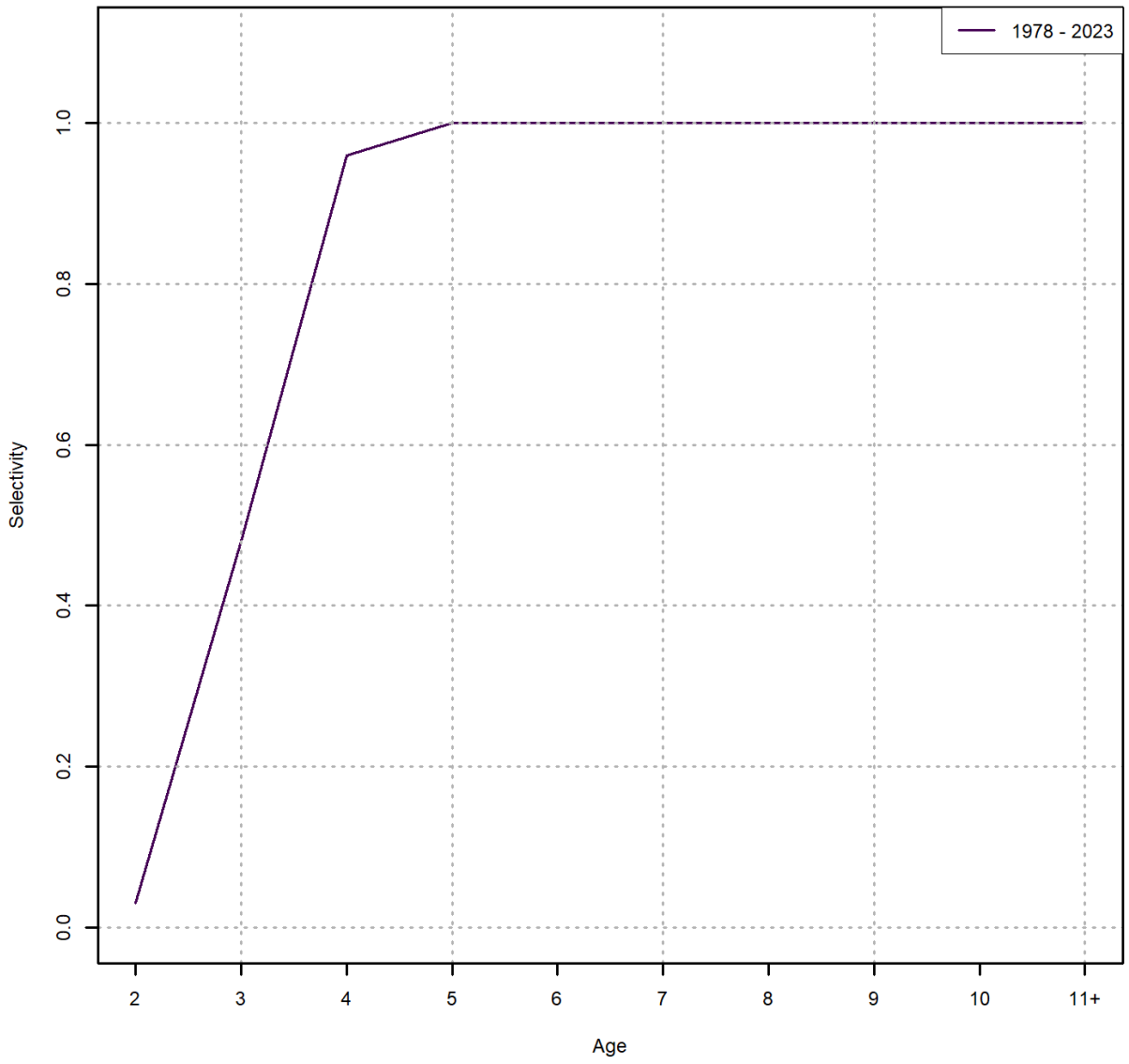

Figure 47: Middle model SG acoustic survey mean selectivity-at-age modeled as the mean maturity-at-age schedule.

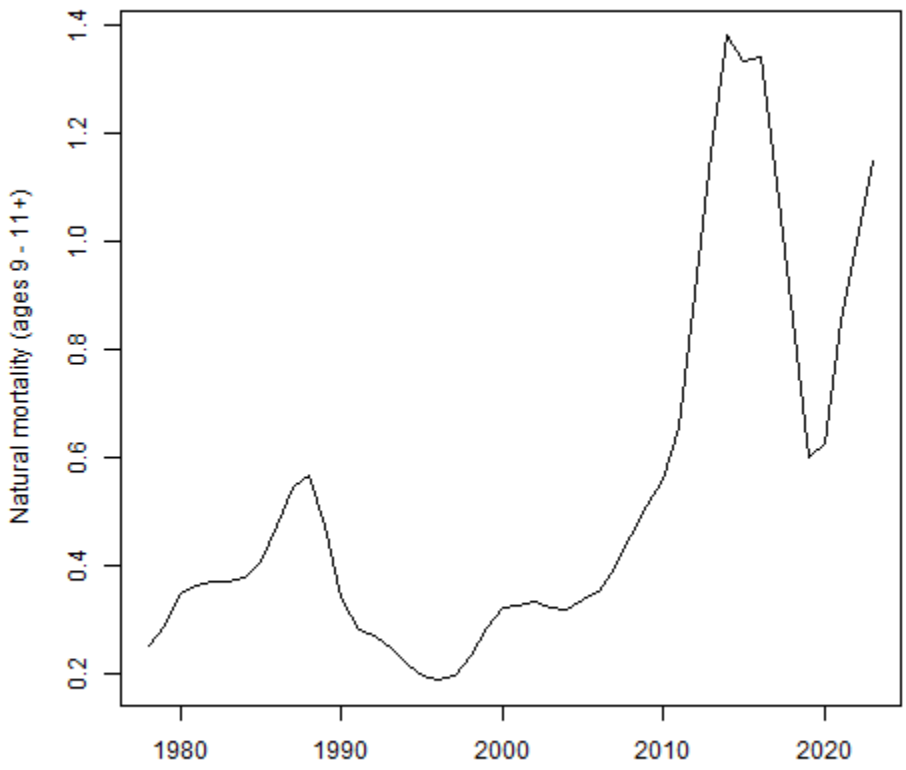

Figure 48: Middle model mean natural mortality (fixed = 0.3) and shared deviations in natural mortality at age (MAA) by age and year, for the age group 9 to 11+.

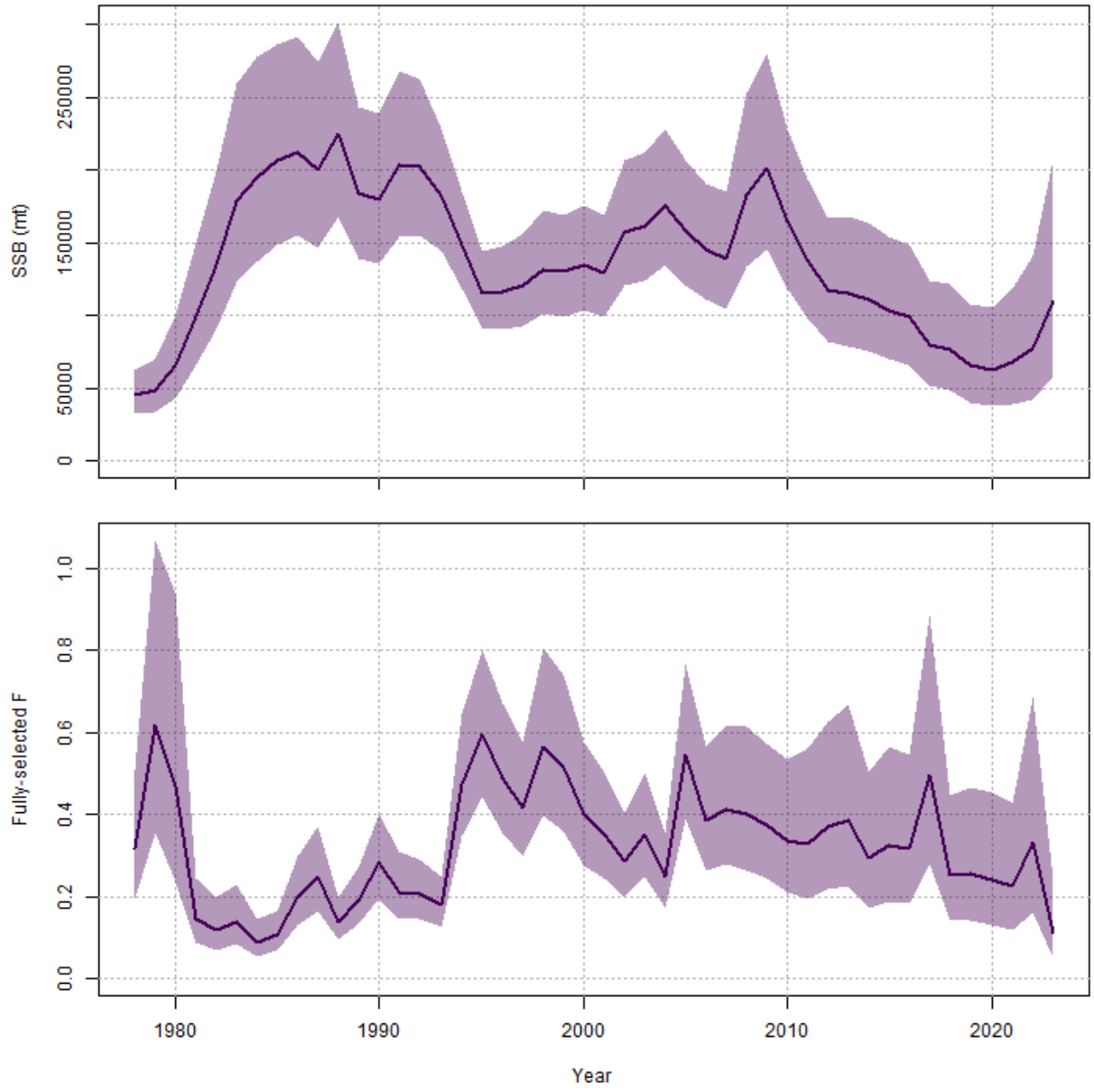

Figure 49: Middle model spawning stock biomass (SSB; upper panel) and fully-selected fishing mortality (F; lower panel) between 1978 and 2023. Solid lines are means and shading are 95% confidence intervals.

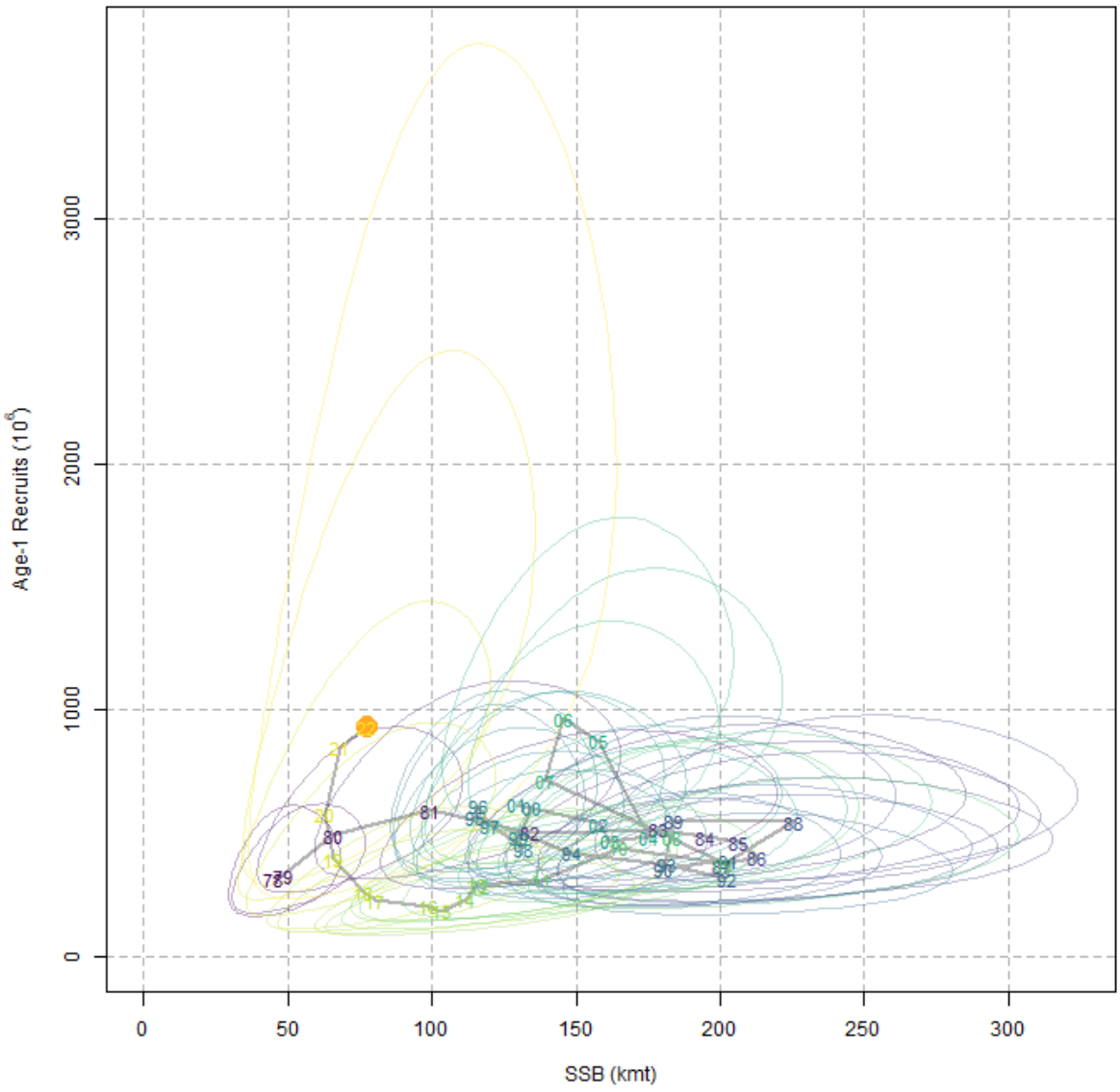

Figure 50: Middle model recruitment (y axis, numbers at age 2 in hundred thousands) and spawning stock biomass (x axis; SSB in metric kilotonnes). The solid line joins points by year, indicated by the label where the two digits represent the last two digits of the year. Large circles are uncertainty around SSB and recruitment estimates.

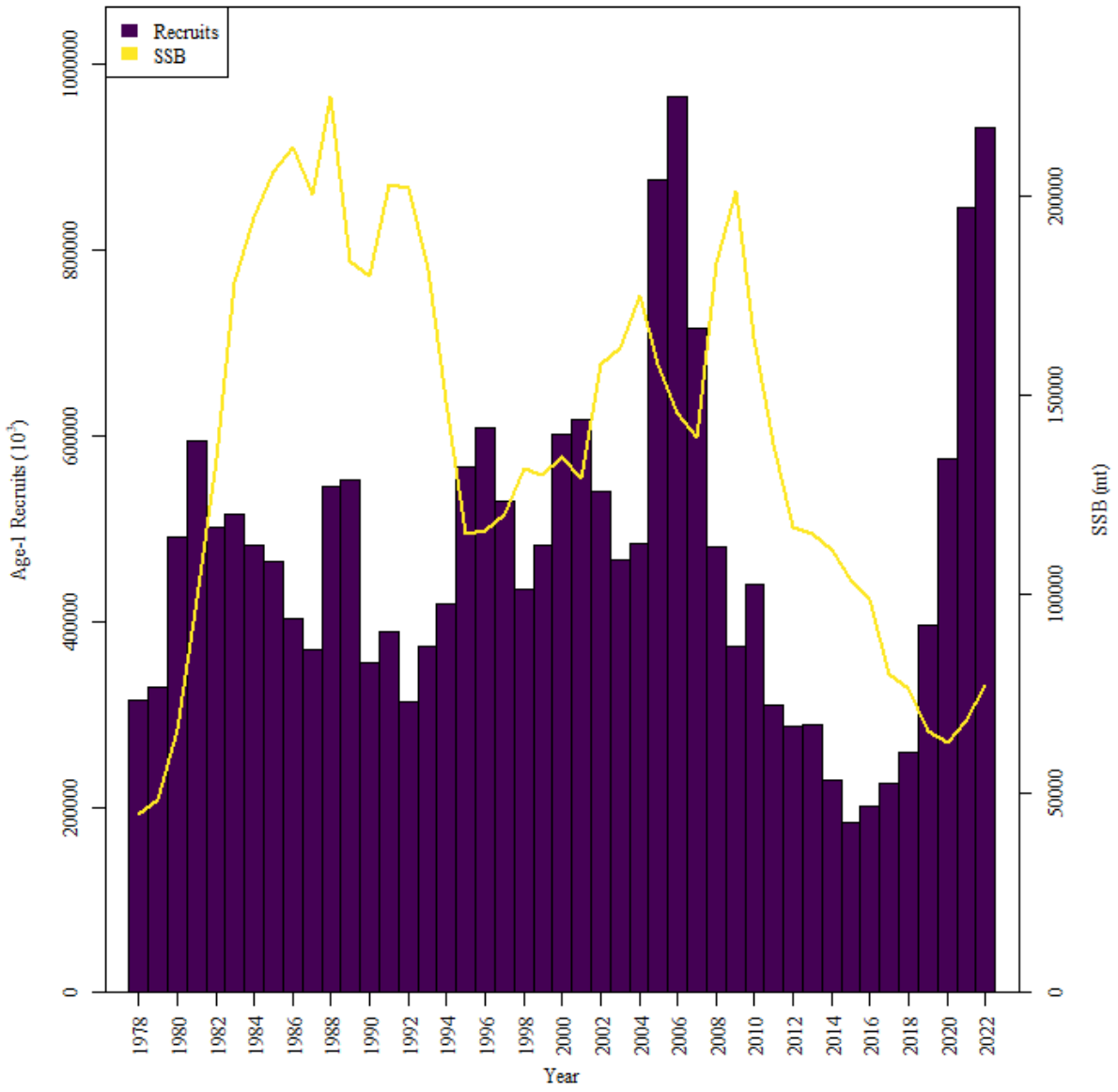

Figure 51: Middle model recruitment (left axis, purple bars, numbers at age 2 in thousands) and spawning stock biomass (right axis; yellow line, SSB) between 1978 and 2022.

SOUTH MODEL

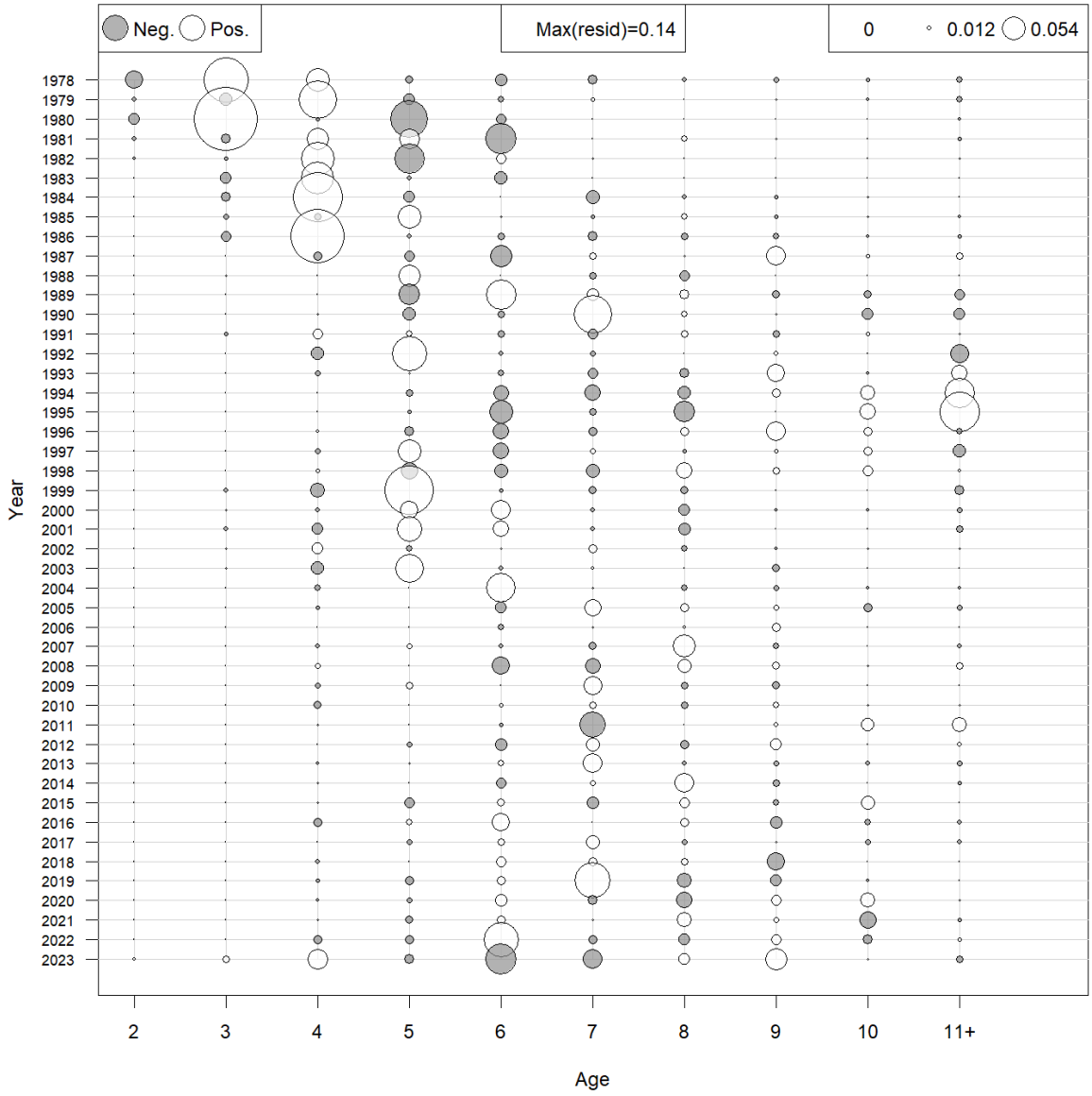

Figure 52: South model fixed gear fishery age composition residuals (observed-predicted) between 1978 and 2023. White circles are positive residuals and grey circles are negative residuals. Circle size is proportional to the residual value.

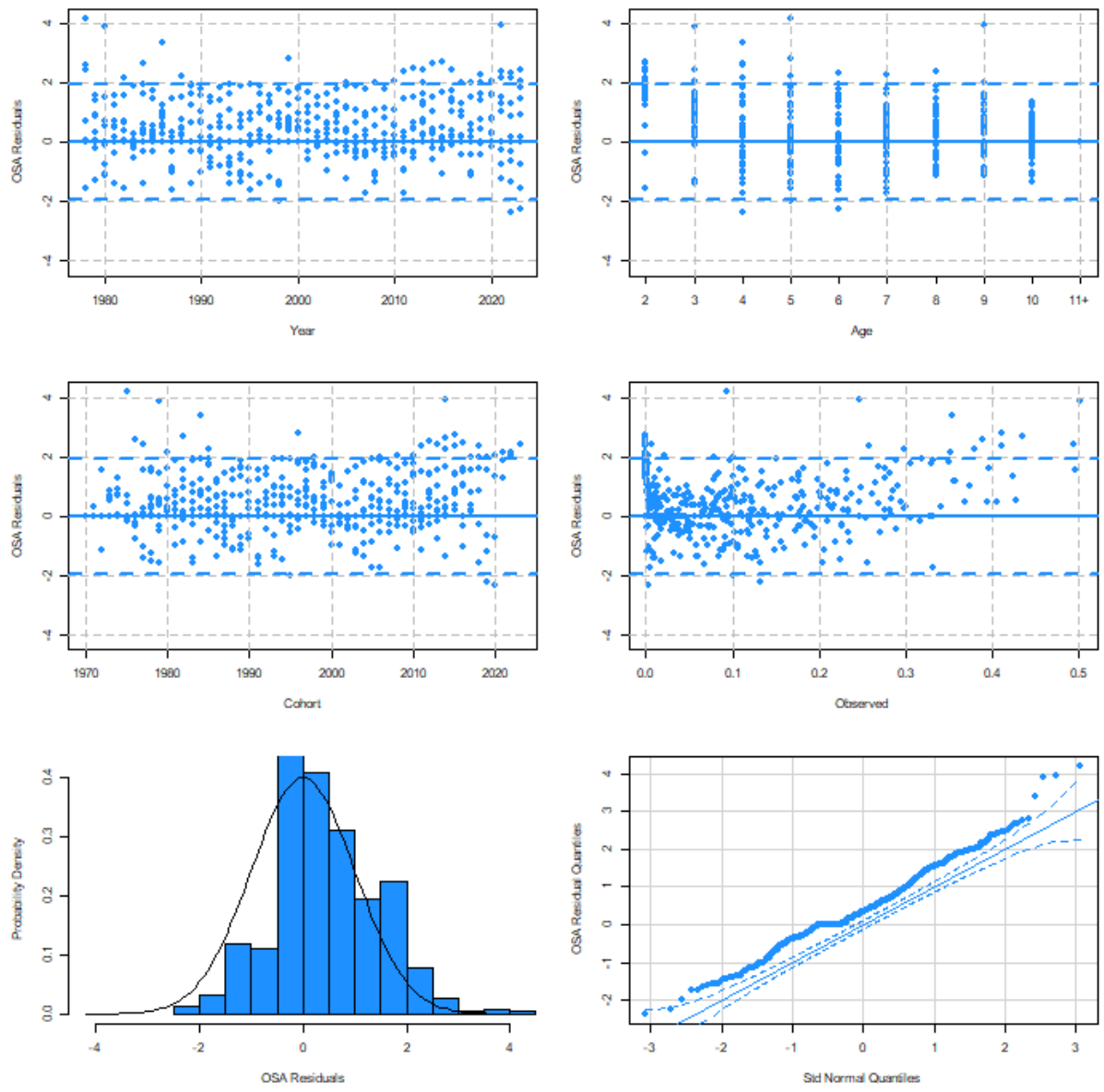

Figure 53: South model fixed gear age composition one-step-ahead (OSA) residuals by (panels from left to right top to bottom) year, age, cohort and by observed value, probability density and versus normal quantiles.

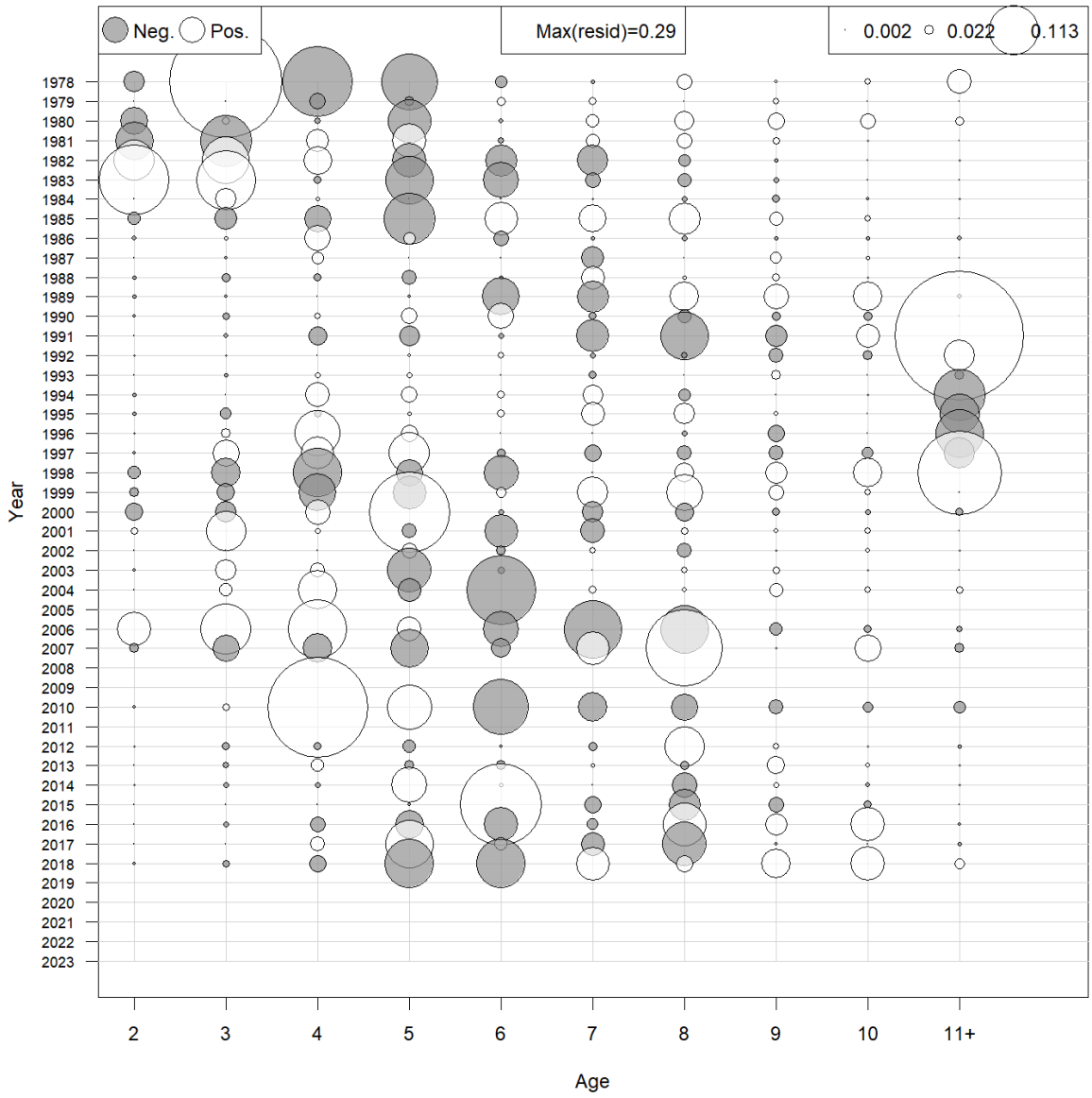

Figure 54: South model mobile gear fishery age composition residuals (observed-predicted) between 1978 and 2023. White circles are positive residuals and grey circles are negative residuals. Circle size is proportional to the residual value.

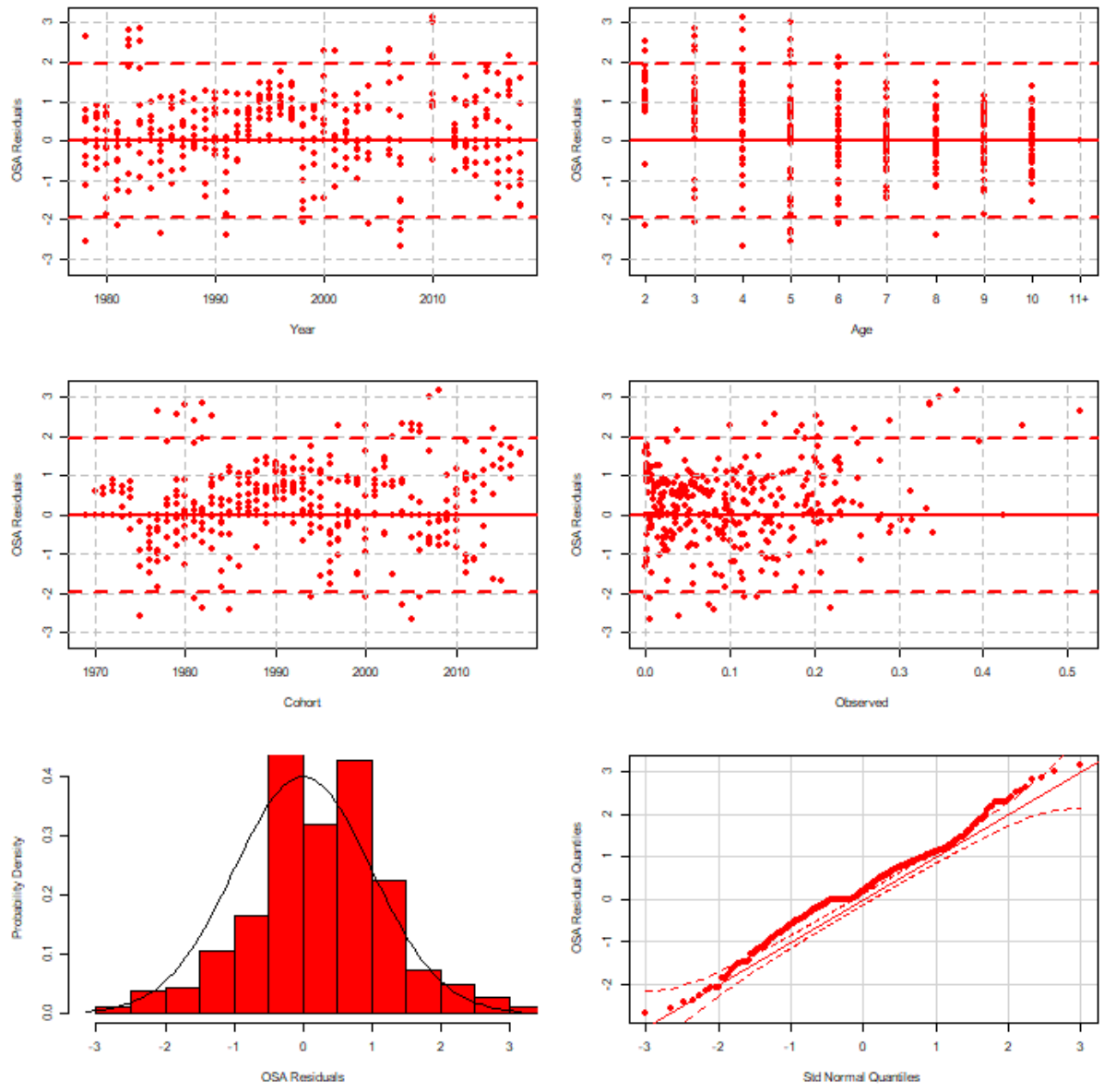

Figure 55: South model mobile gear age composition one-step-ahead (OSA) residuals by (panels from left to right top to bottom) year, age, cohort and by observed value, probability density and versus normal quantiles.

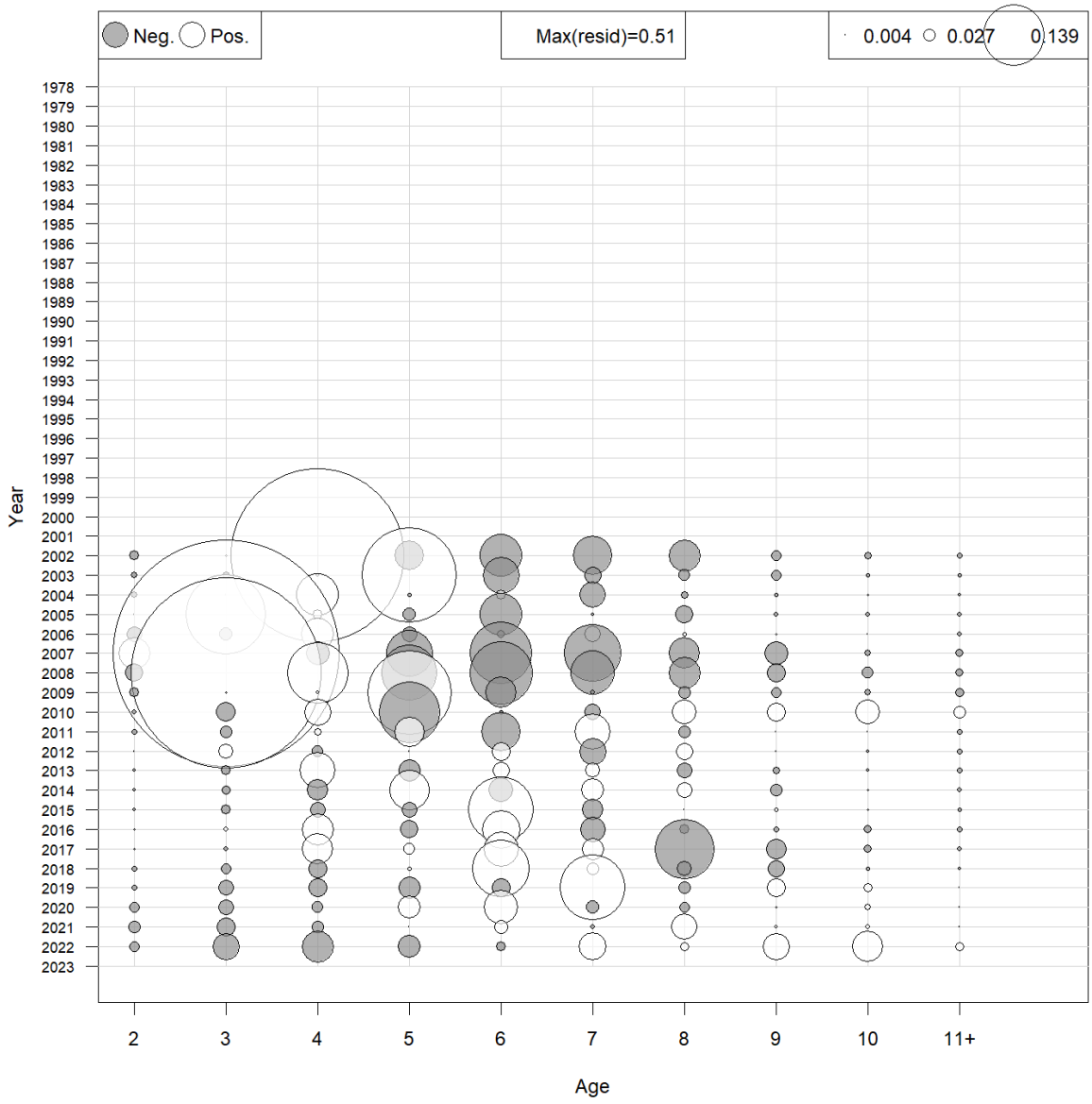

Figure 56: South model RV survey age composition residuals (observed-predicted) between 1978 and 2023. White circles are positive residuals and grey circles are negative residuals. Circle size is proportional to the residual value.

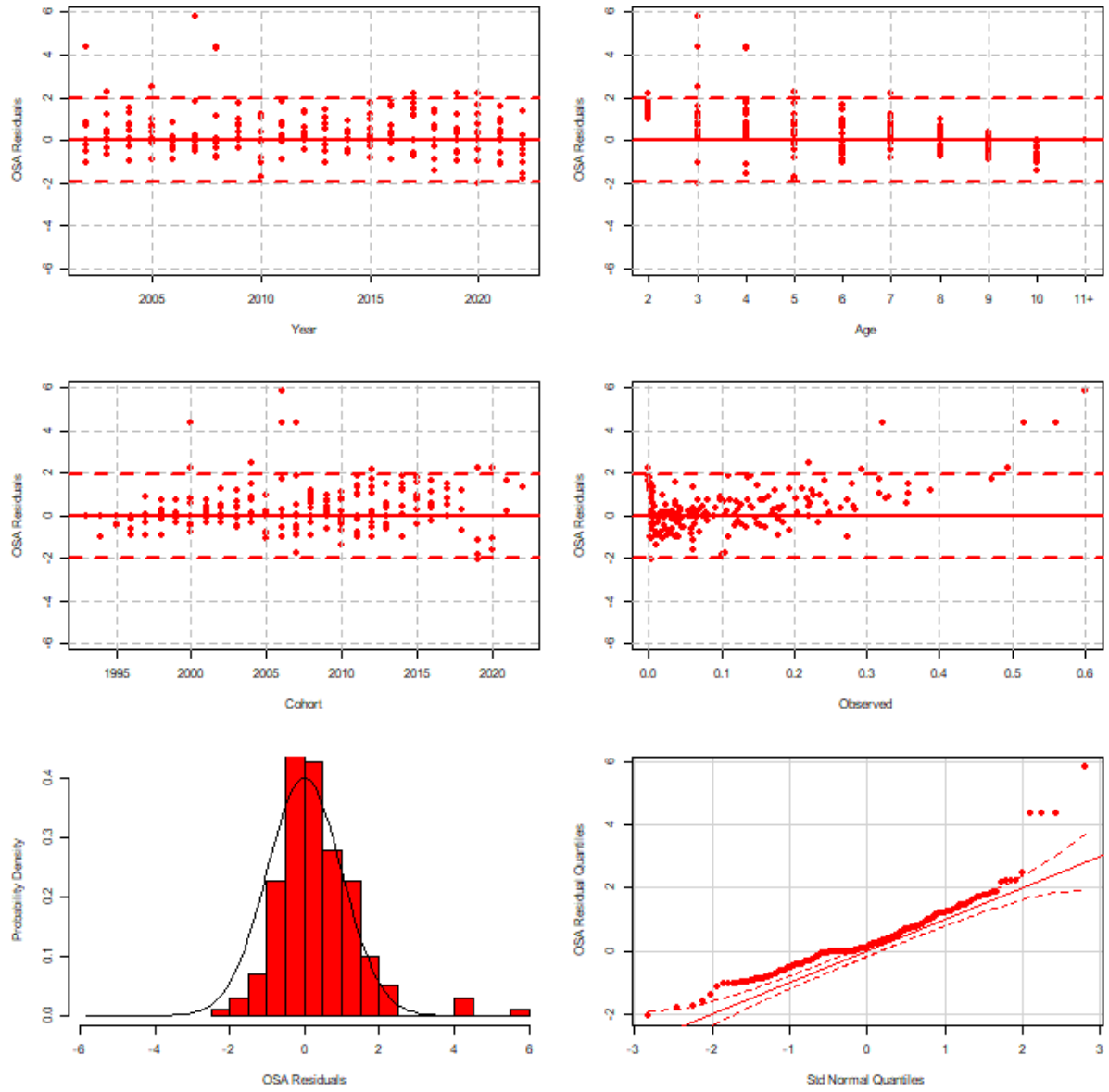

Figure 57: South model experimental nets survey age composition one-step-ahead (OSA) residuals by (panels from left to right top to bottom) year, age, cohort and by observed value, probability density and versus normal quantiles.

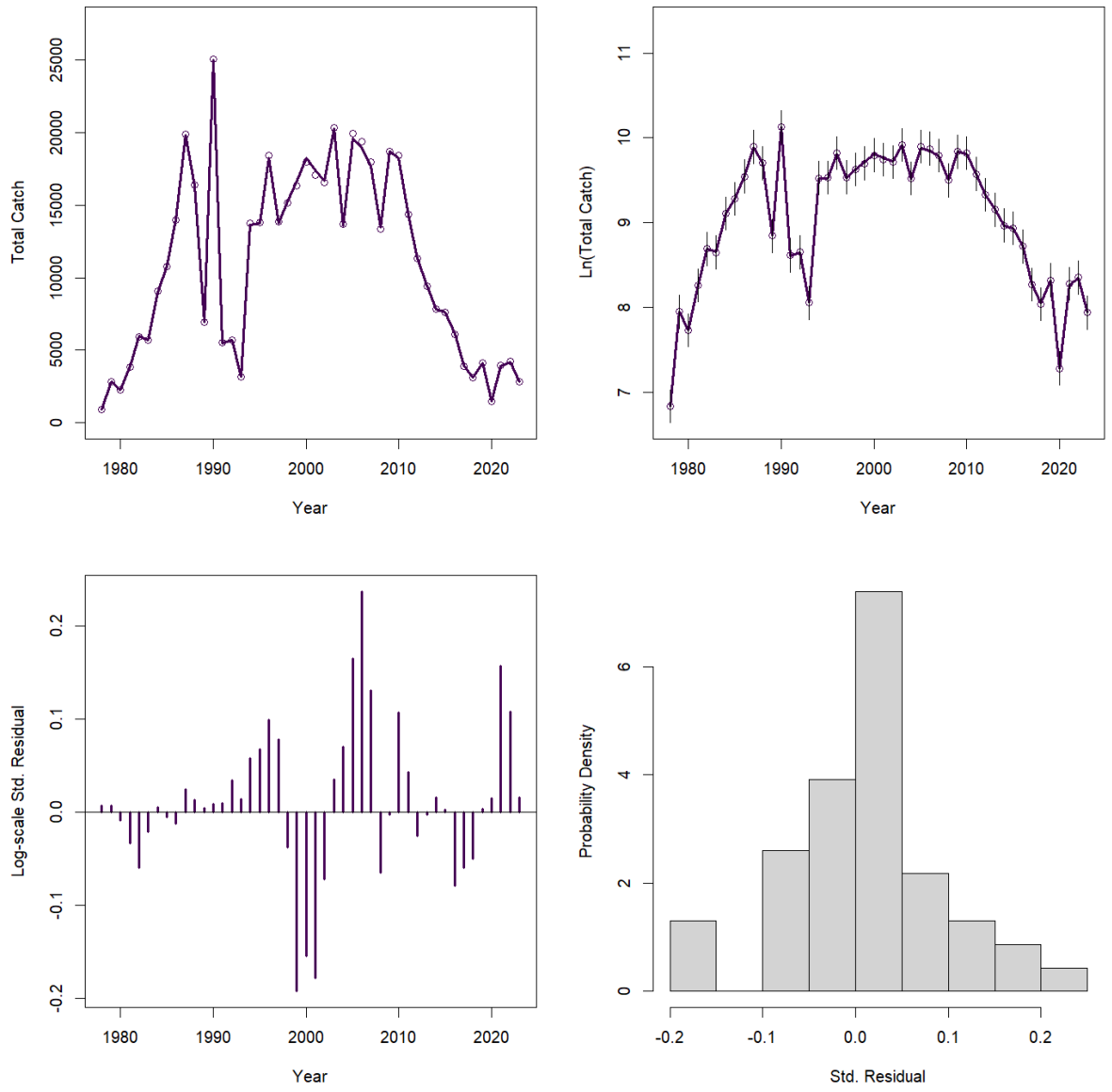

Figure 58: South model fit to the fixed gear total catch (top left), and to the log of the catch (top right), standardized residuals by year (bottom left) probability density of standardized residuals (bottom right).

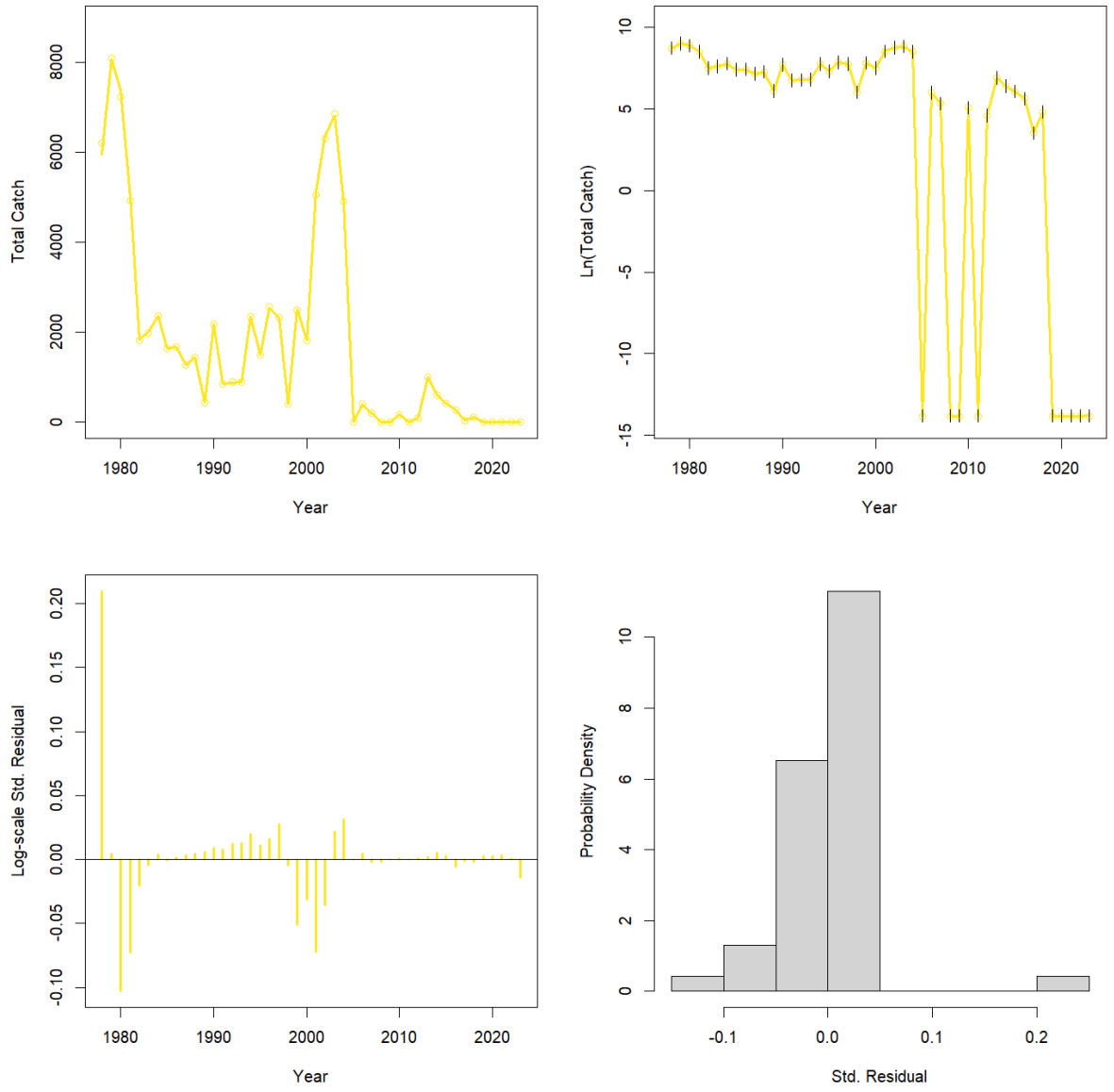

Figure 59: South model fit to the mobile gear total catch (top left), and to the log of the catch (top right), standardized residuals by year (bottom left) probability density of standardized residuals (bottom right).

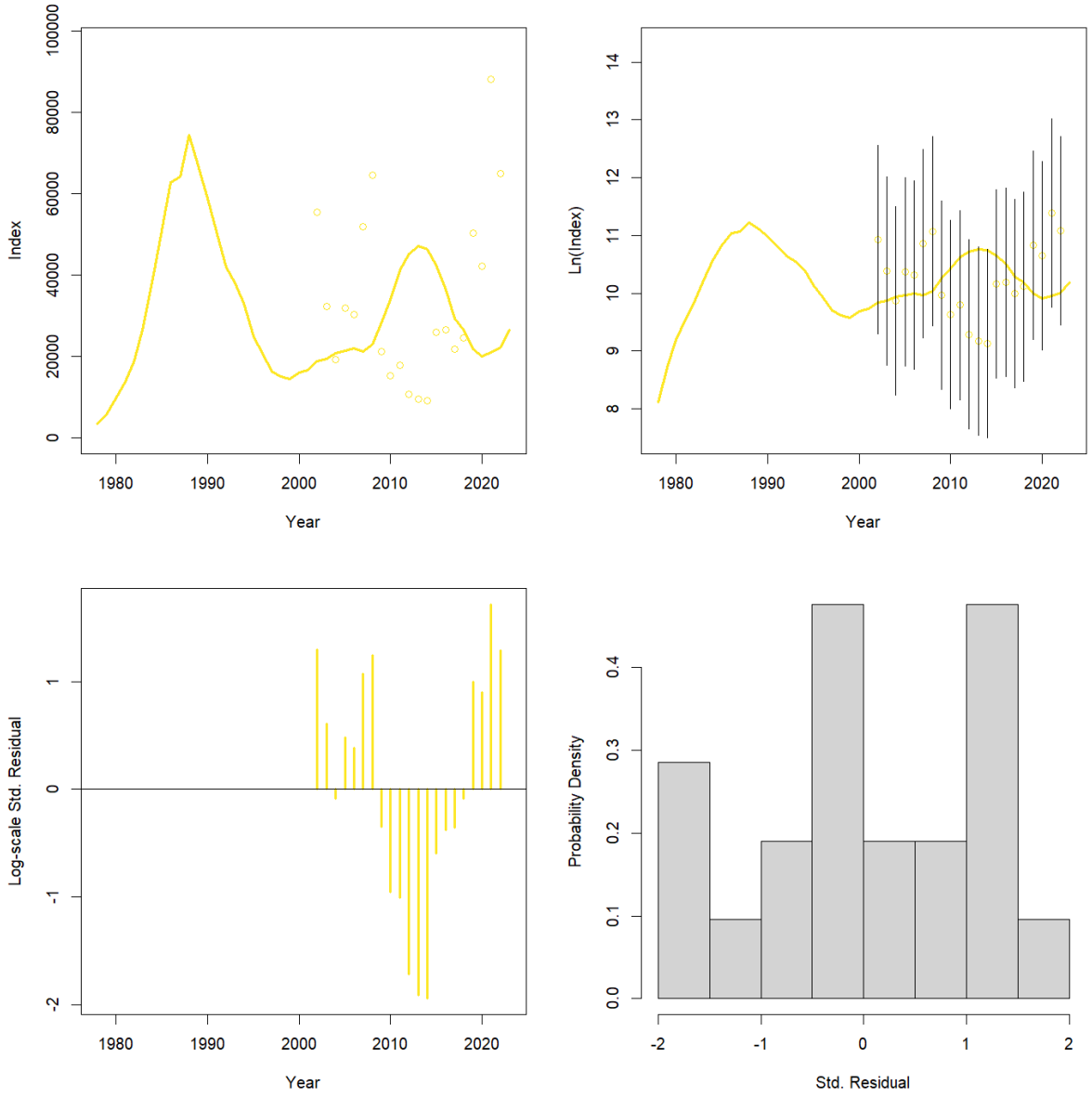

Figure 60: South model fit to the experimental nets survey biomass index (top left), and to the log of the index (top right), standardized residuals by year (bottom left) probability density of standardized residuals (bottom right).

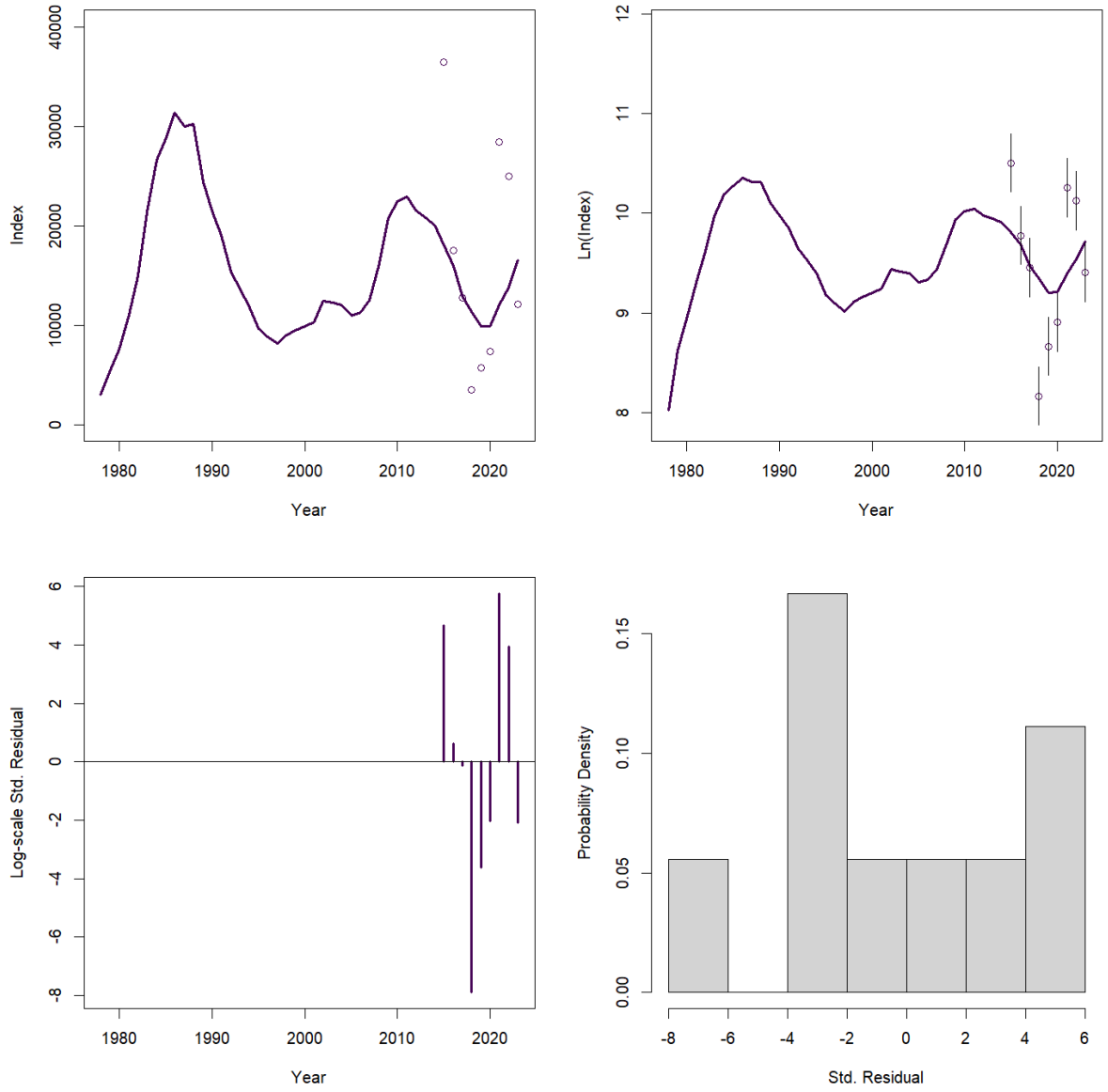

Figure 61: South model fit to the SG acoustic survey biomass index (top left), and to the log of the index (top right), standardized residuals by year (bottom left) probability density of standardized residuals (bottom right).

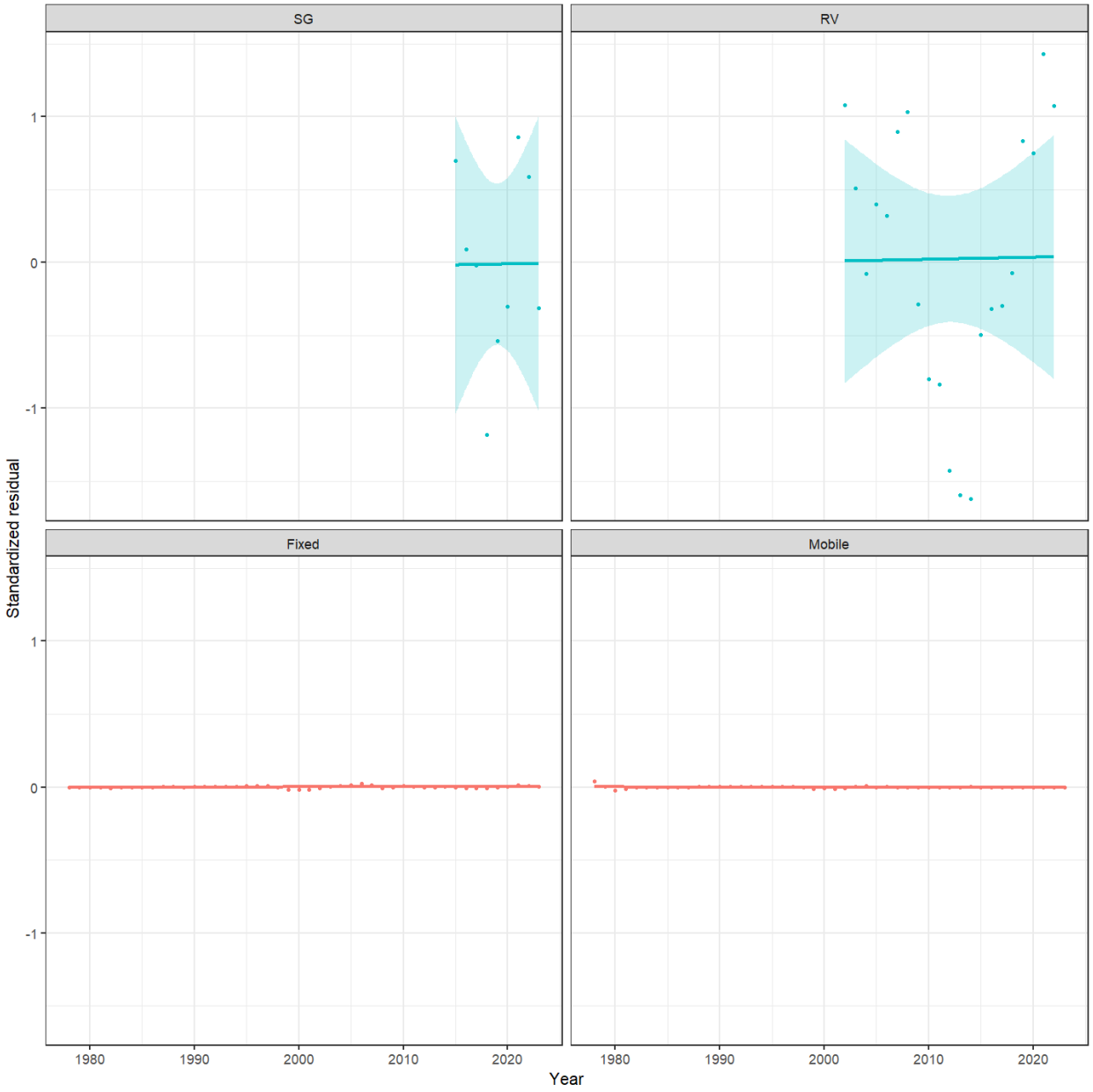

Figure 62: South model standardized residuals by year for the SG acoustic survey, experimental nets survey and fixed and mobile gears catch.

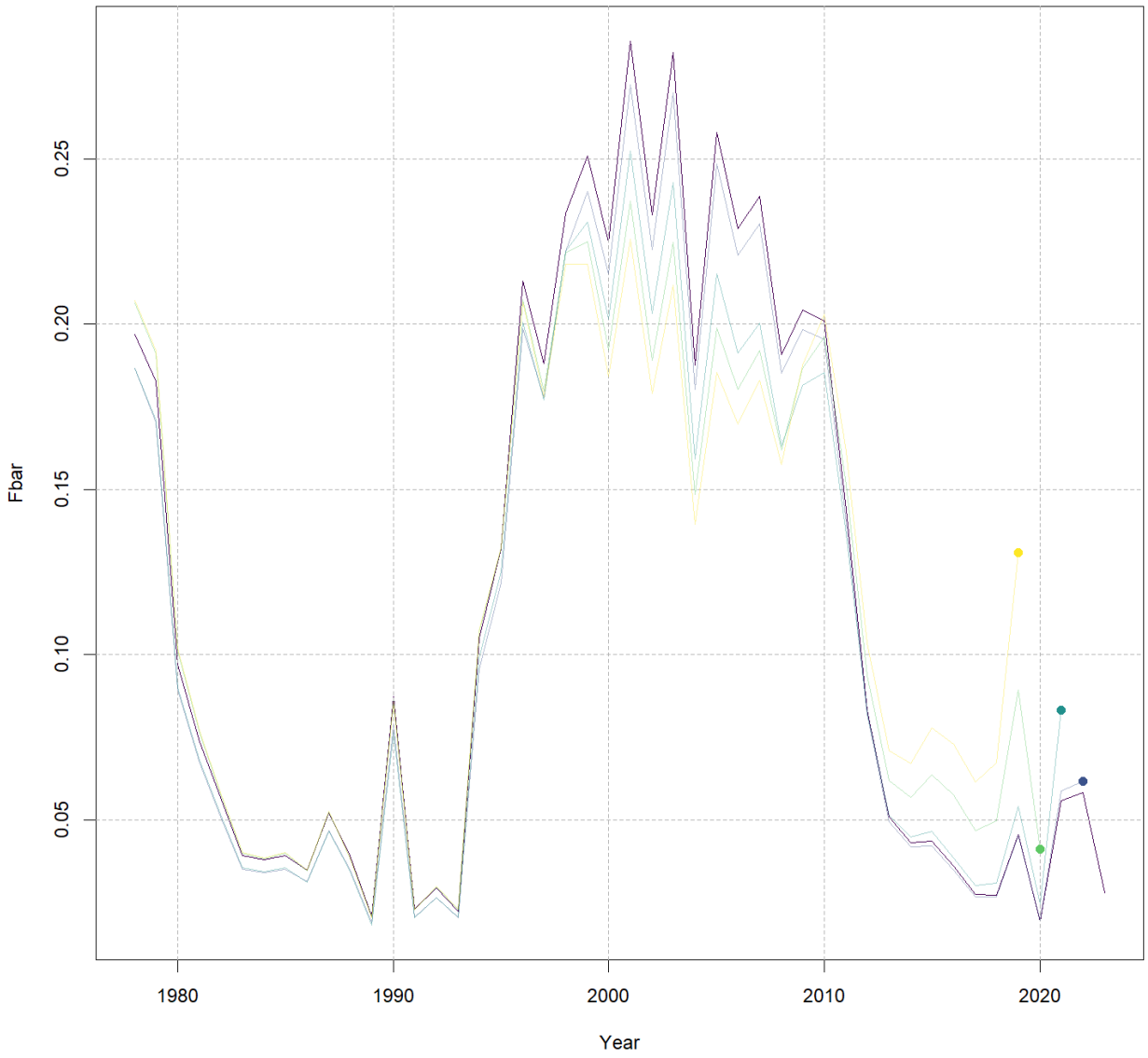

Figure 63: South model retrospective analysis for average fishing mortality (Fbar) for 4 peels.

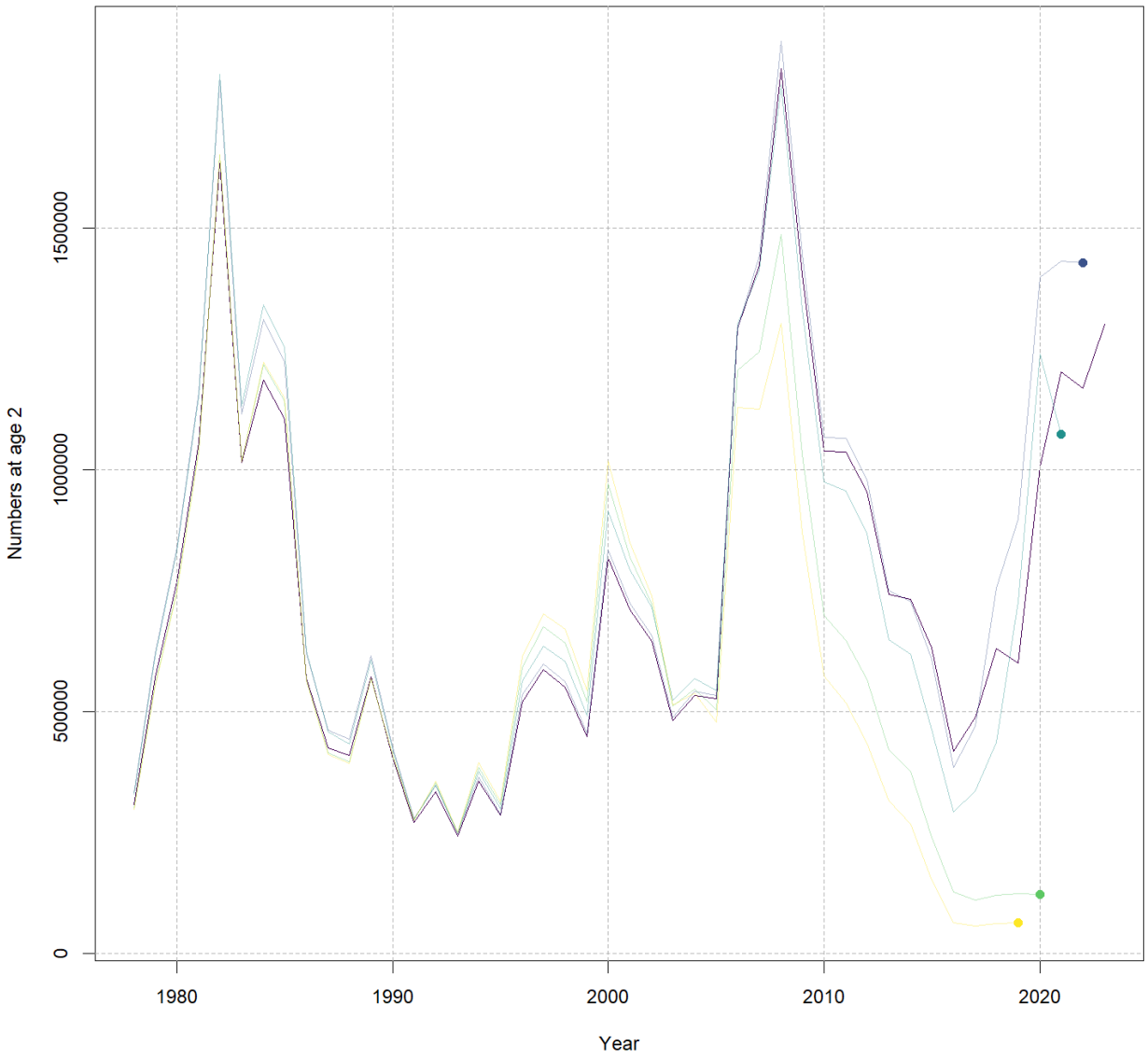

Figure 64: South model retrospective analysis for recruitment (numbers at age 2) for 4 peels.

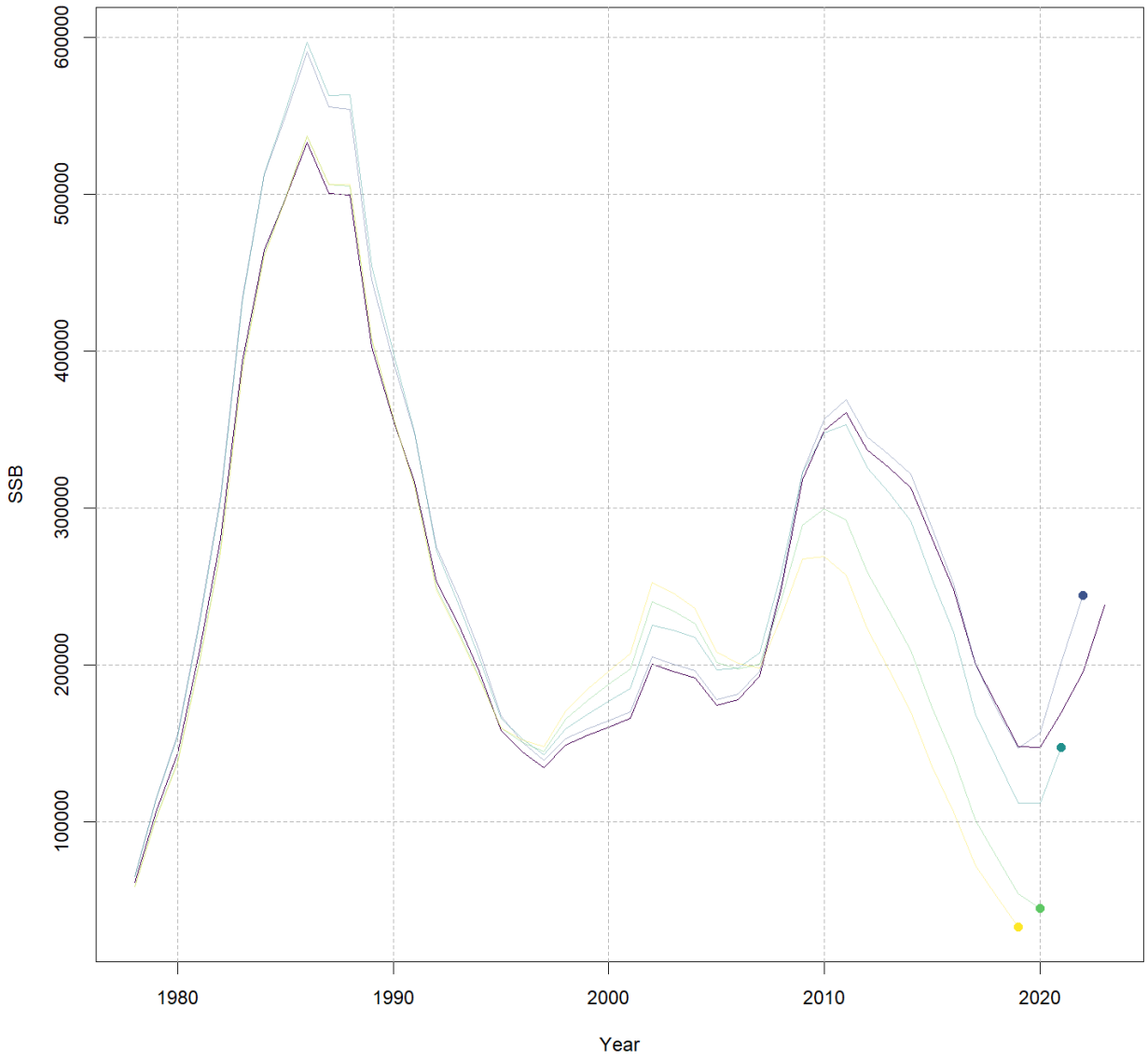

Figure 65: South model retrospective analysis for spawning stock biomass (SSB; tonnes) for 4 peels.

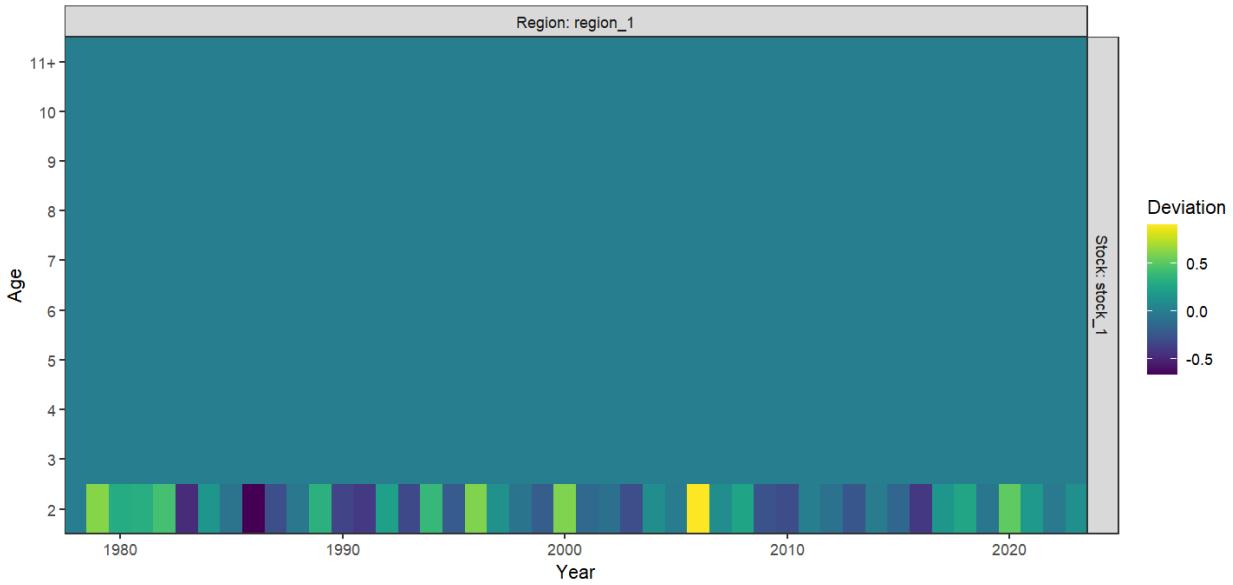

Figure 66: South model deviations in numbers-at-age (NAA) by age and year, with one SD for age 2. The South mode didn't allow process error in NAA for ages 3 to 11+.

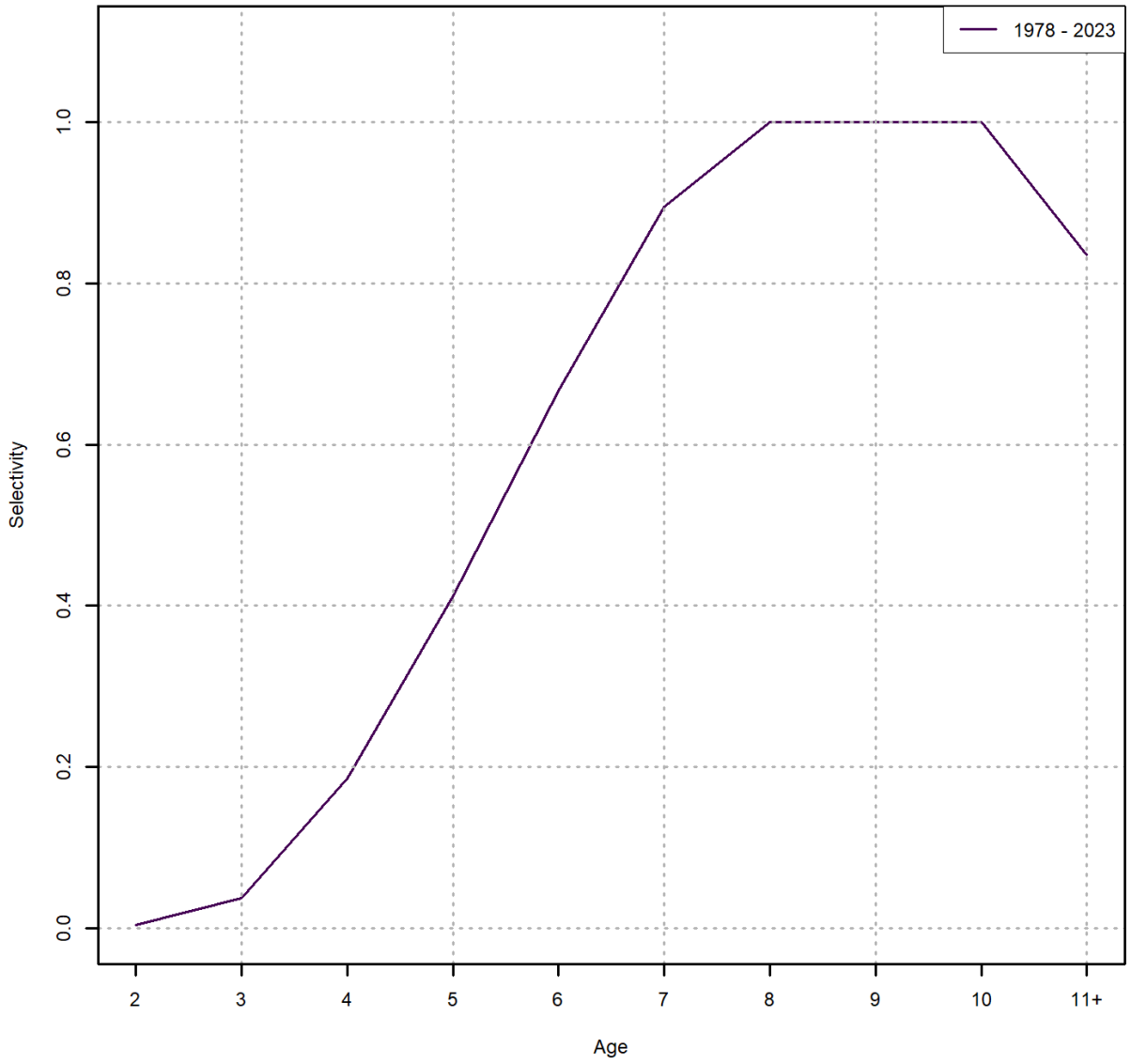

Figure 67: South model mean selectivity-at-age for the combined fixed and mobile gears fishery.

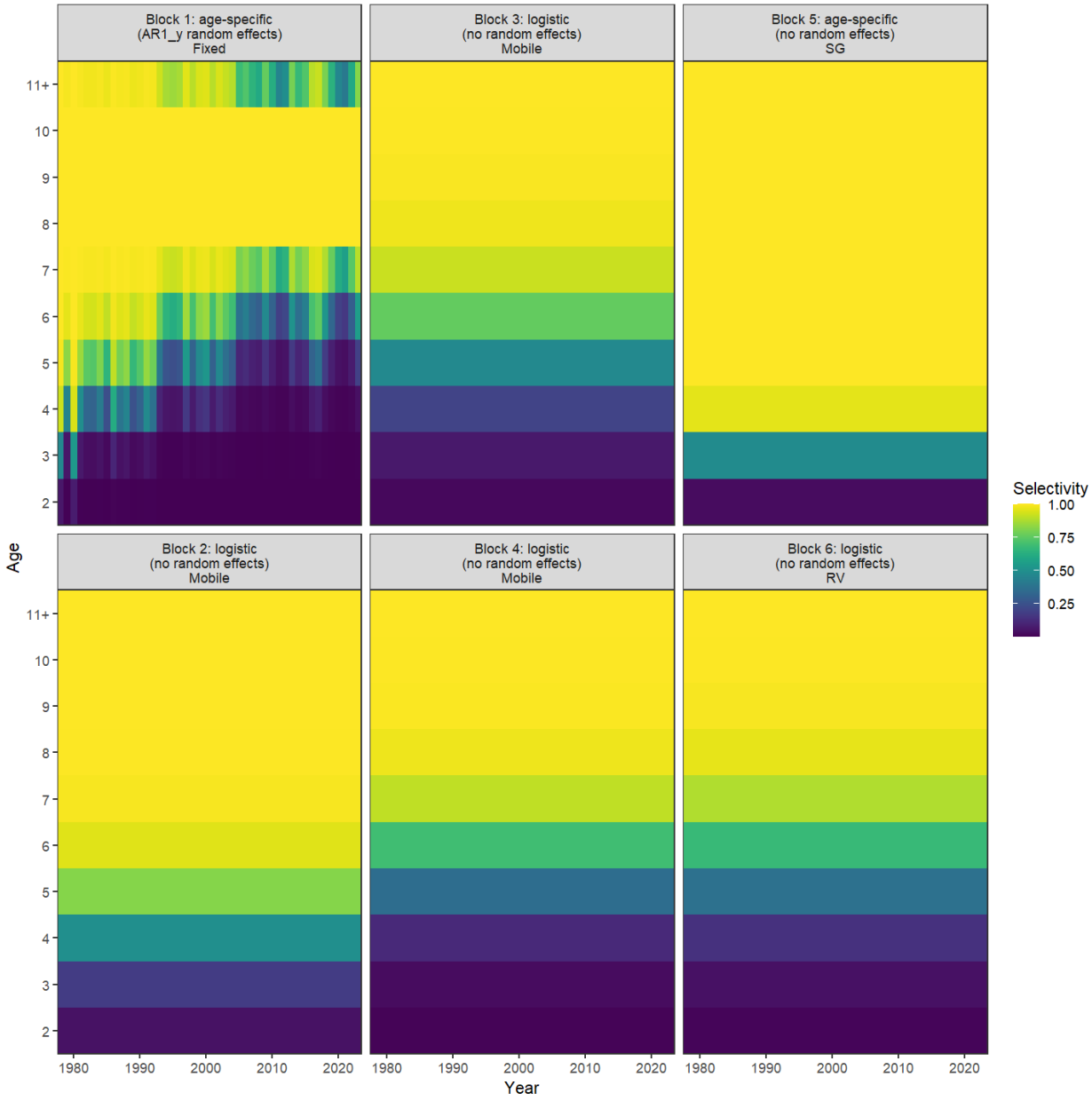

Figure 68: South model realized selectivity including random effects were applicable, for all sources of catch, surveys and time blocks.

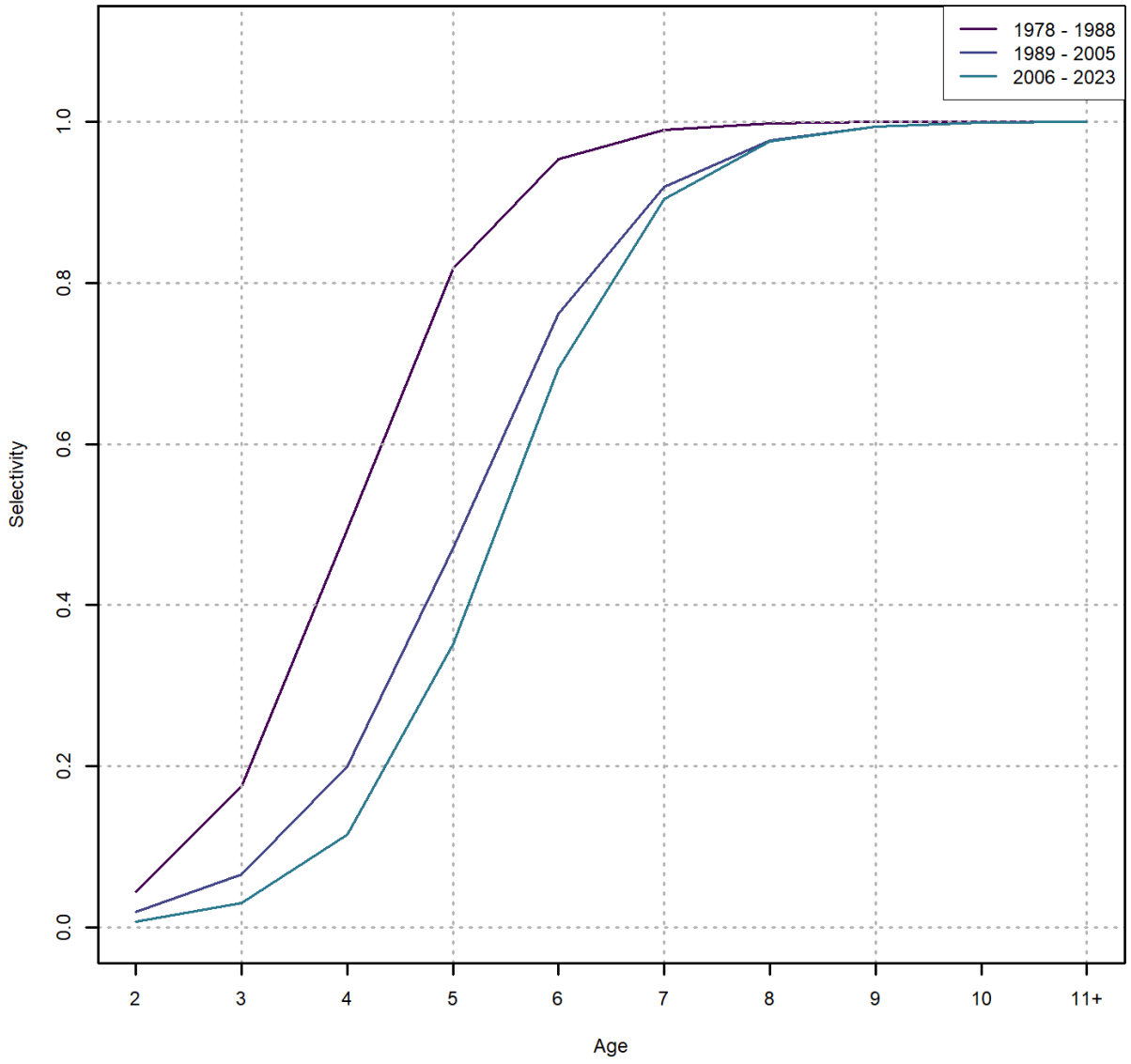

Figure 69: South model mean selectivity-at-age for the combined fixed and mobile gears fishery.

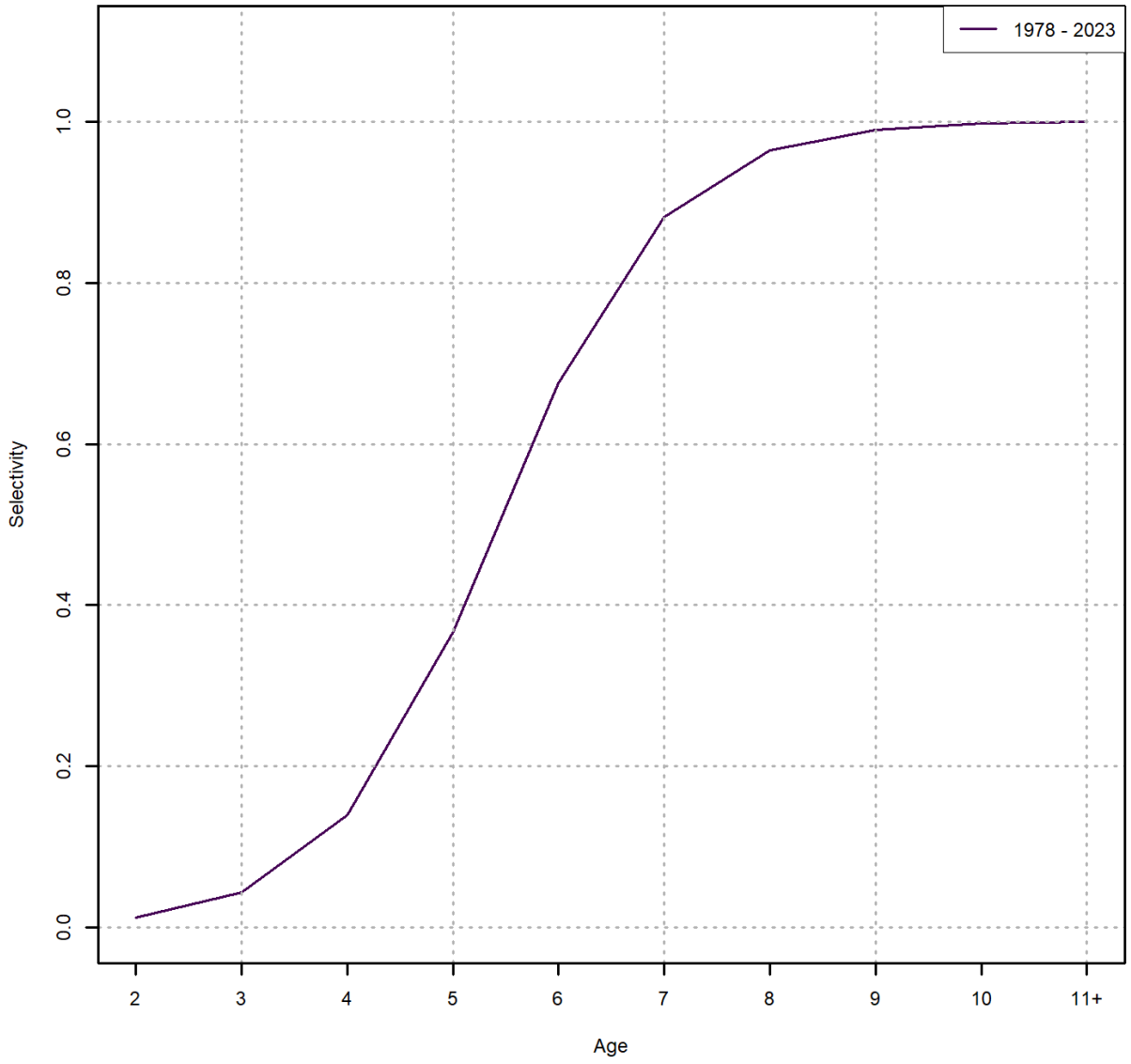

Figure 70: South model experimental nets survey mean selectivity-at-age modeled as logistic curve.

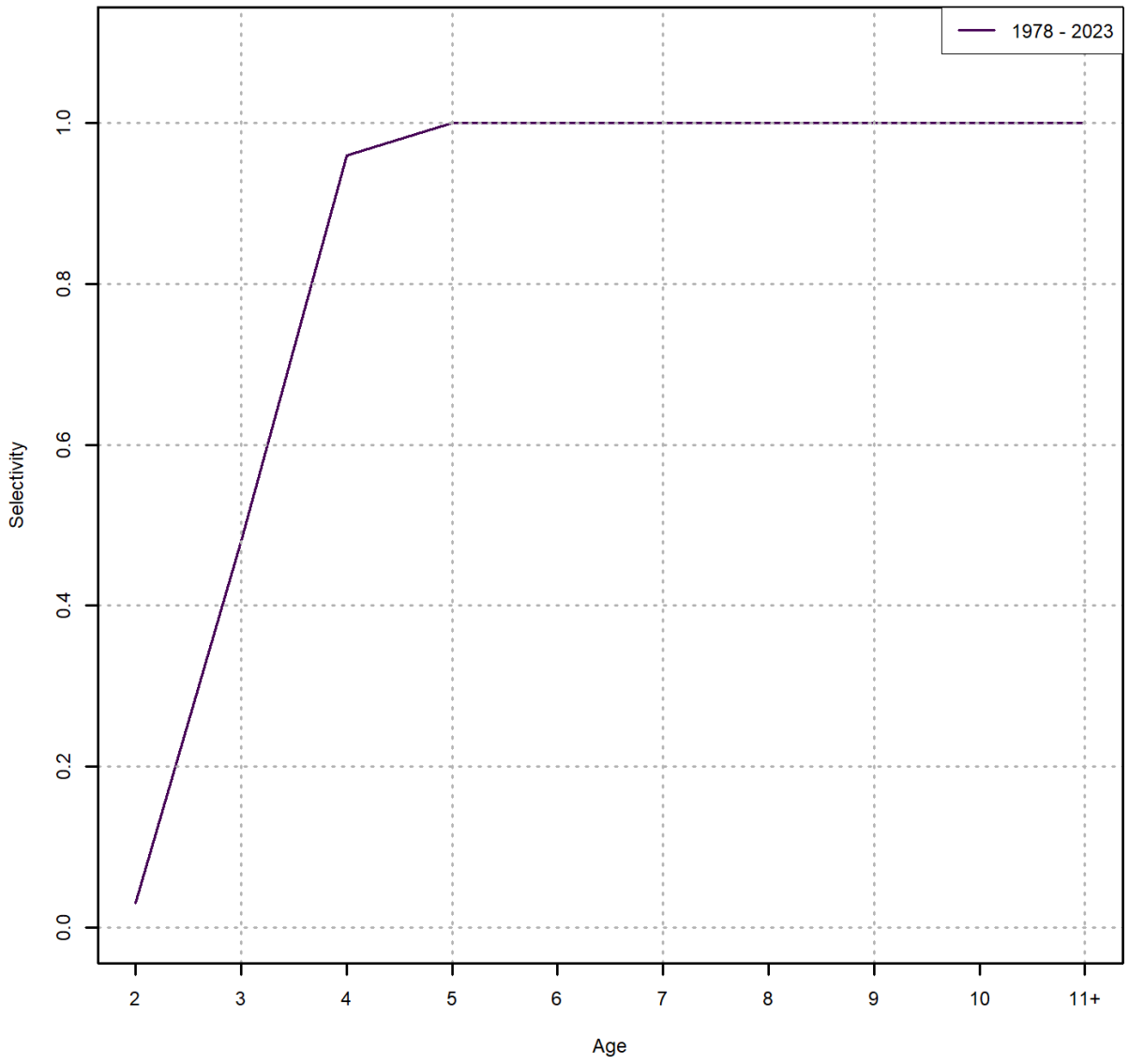

Figure 71: South model SG acoustic survey mean selectivity-at-age modeled as the mean maturity-at-age schedule.

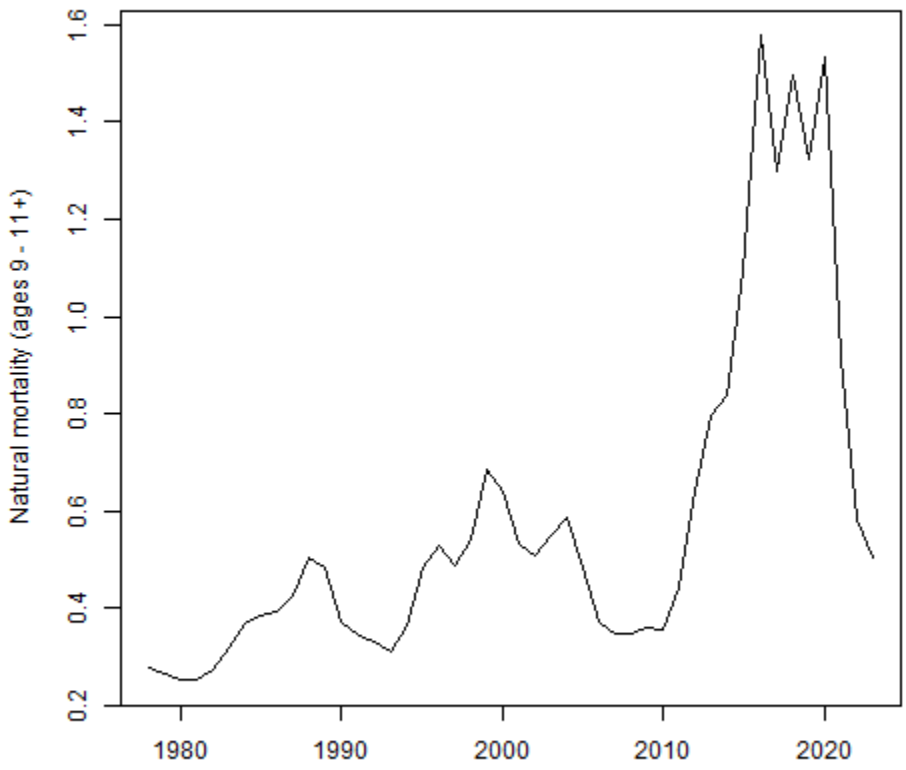

Figure 72: South model mean natural mortality (fixed = 0.3) and shared deviations in natural mortality at age (MAA) by age and year, for the age group 9 to 11+.

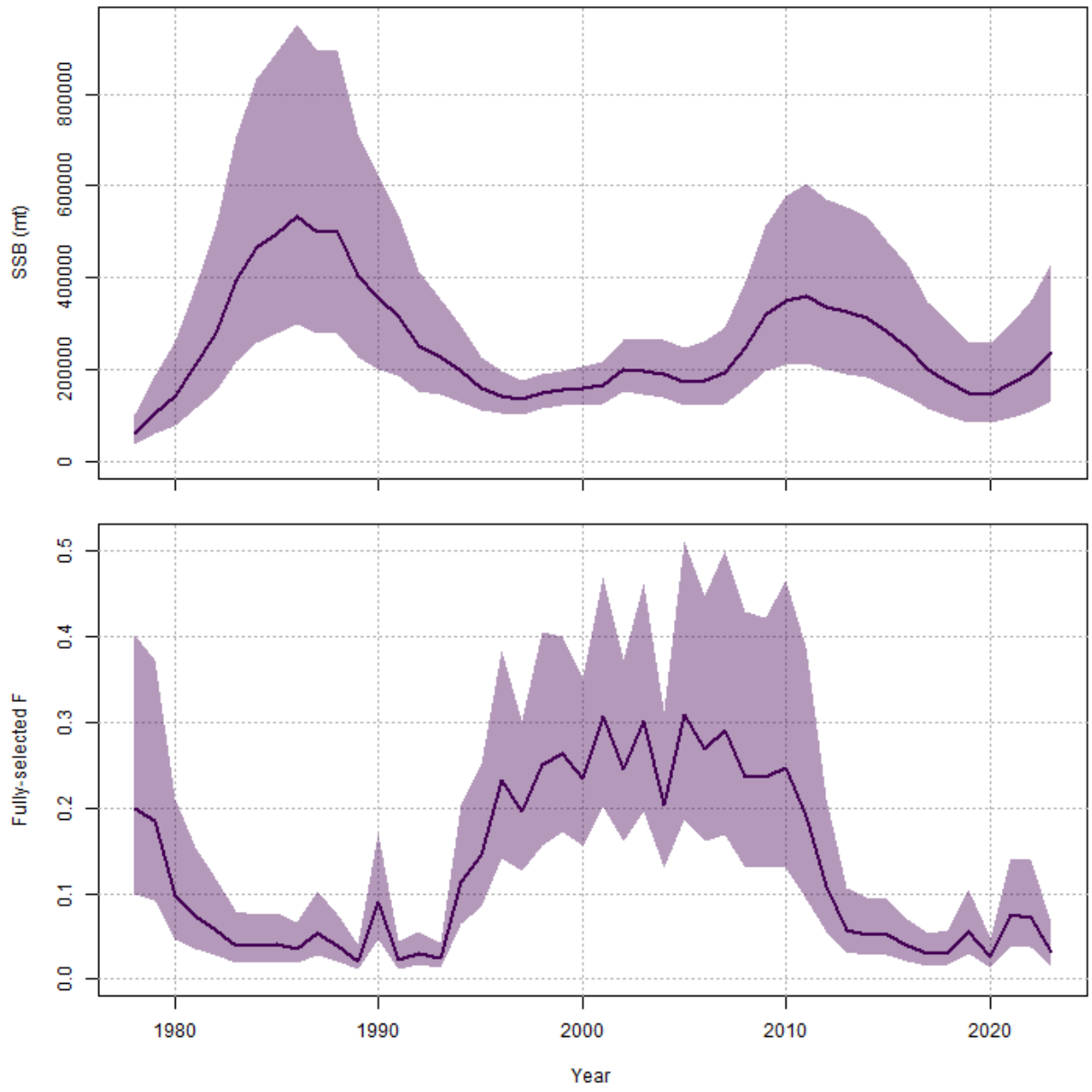

Figure 73: South model spawning stock biomass (SSB; upper panel) and fully-selected fishing mortality (F; lower panel) between 1978 and 2023. Solid lines are means and shading are 95% confidence intervals.

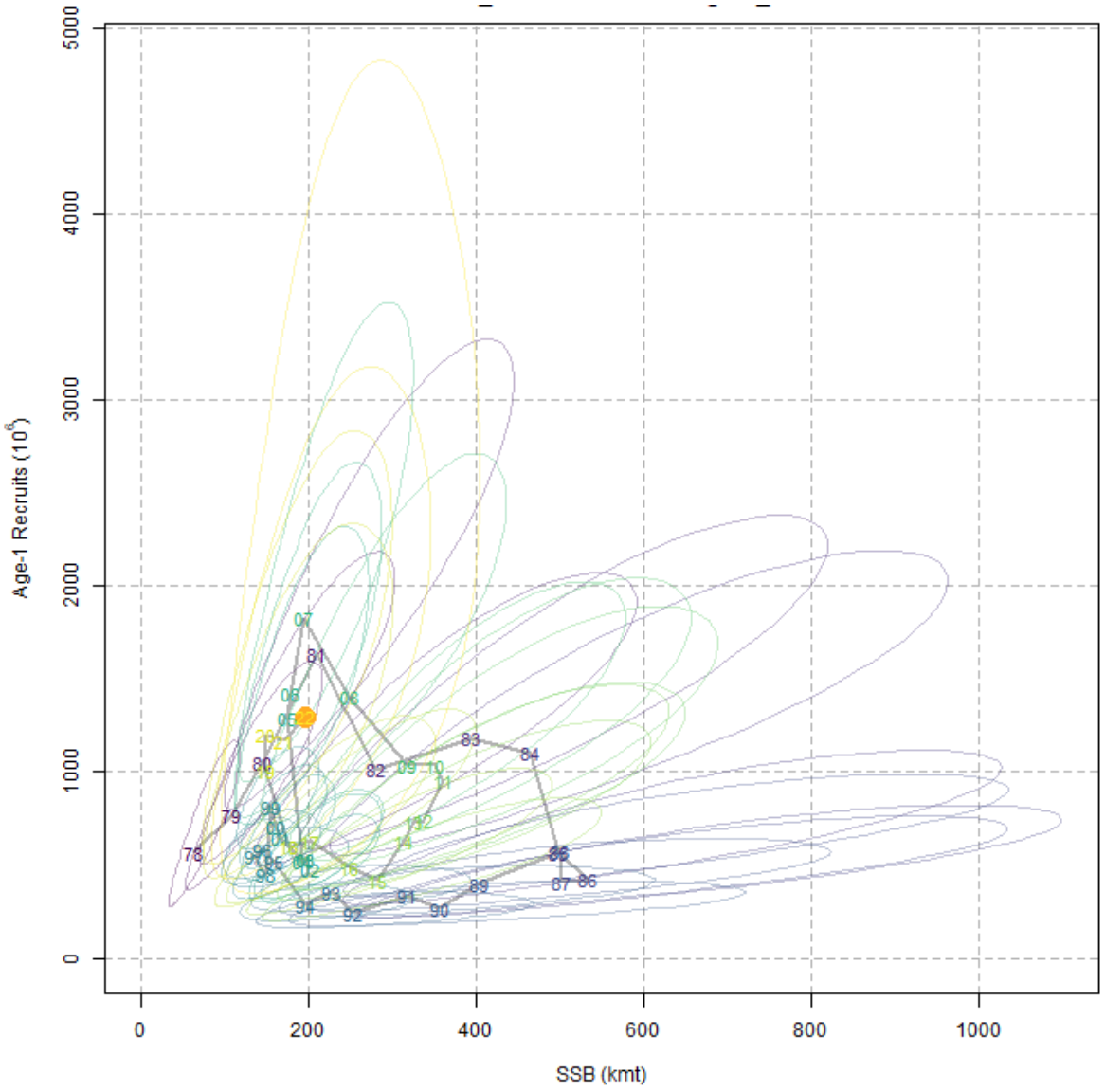

Figure 74: South model recruitment (y axis, numbers at age 2 in hundred thousands) and spawning stock biomass (x axis; SSB in metric kilotonnes). The solid line joins points by year, indicated by the label where the two digits represent the last two digits of the year. Large circles are uncertainty around SSB and recruitment estimates.

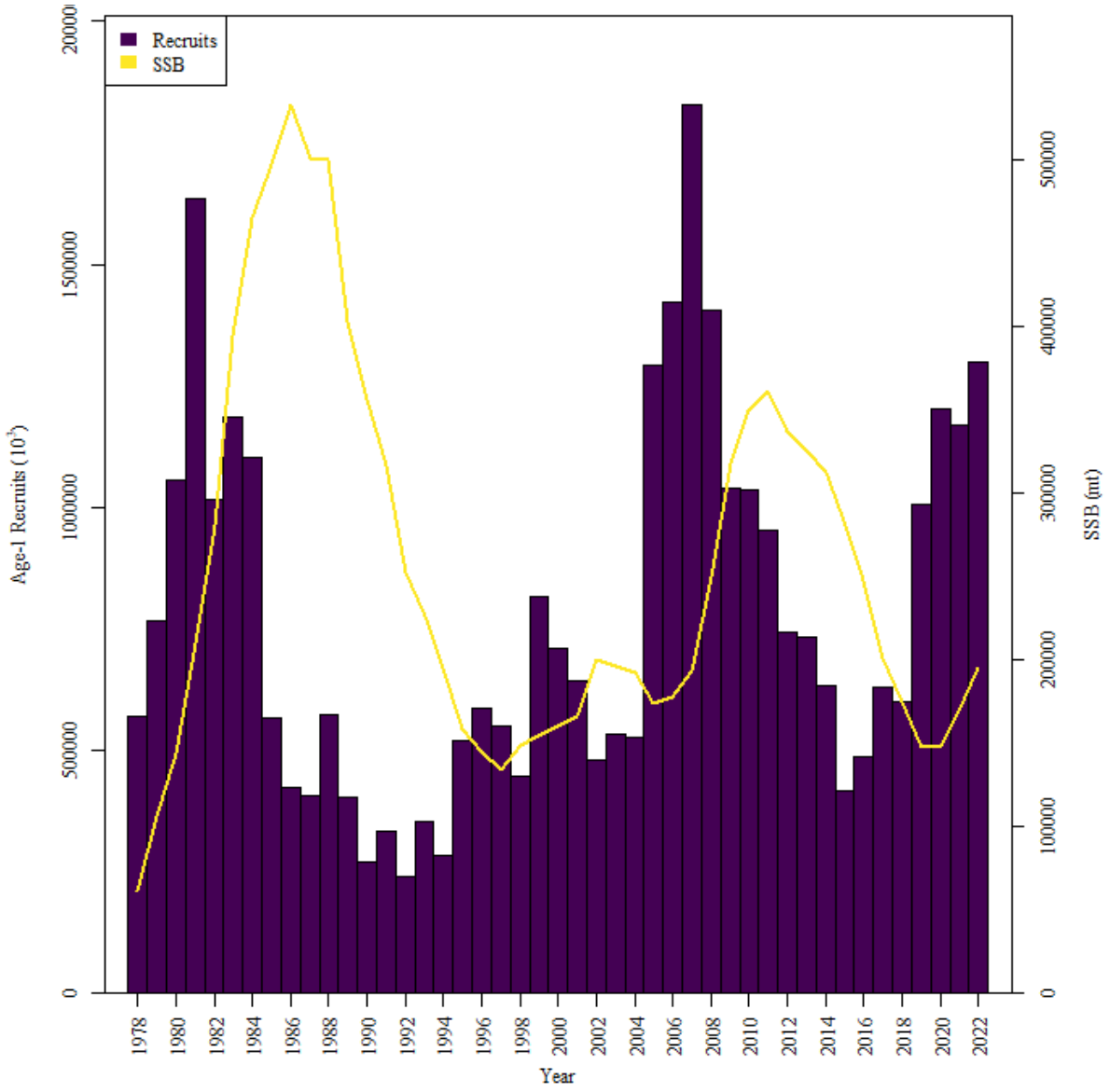

Figure 75: South model recruitment (left axis, purple bars, numbers at age 2 in thousands) and spawning stock biomass (right axis; yellow line, SSB) between 1978 and 2022.

APPENDIX 1: SENSITIVITY RUNS

Table A1: North model sensitivity runs for mean natural mortality (M), random effects (RE), delta Akaike Information Criterion (dAIC), retrospective analysis Mohn's rho values for recruitment (R), spawning stock biomass (SSB) and average F (Fbar).

Model name	Mean M	RE	Converged	All parameters estimable	dAIC	rho R	rho SSB	rho Fbar
north	fixed 0.3	ar1_y	yes	yes	0	-0.156	-0.073	0.151
north_fixM0.3	fixed 0.3	-	yes	yes	14.9	0.128	-0.196	0.283
north_estMiid	estimated	iid	yes	yes	-1.9	-0.380	-0.147	0.228
north_estMar1y	estimated	ar1_y	no	no	-	-	-	-
north_fixM0.4ar1y	fixed 0.4	ar1_y	yes	yes	10.7	-0.185	-0.109	0.195
north_fixM0.2ar1y	fixed 0.2	ar1_y	no	no	-	-	-	-

Table A2 : Middle model sensitivity runs for mean natural mortality (M), random effects (RE), delta Akaike Information Criterion (dAIC), retrospective analysis Mohn's rho values for recruitment (R), spawning stock biomass (SSB) and average F (Fbar).

Model name	Mean M	M RE	NAA RE	Converged	All parameters estimable	dAIC	rho R	rho SSB	rho Fbar
middle	fixed 0.3	ar1_y	iid	yes	yes	0	-0.486	-0.019	-0.141
middle_fixM0.3	fixed 0.3	-	-	yes	yes	1.1	-0.522	0.161	0.0985
middle_fixM0.3_NAAiid	fixed 0.3	-	iid	yes	yes	-9.4	-0.558	-0.095	-0.031
middle_estM_iid	estimated	iid_y	-	no	-	-	-	-	-
middle_estM_ar1y	estimated	ar1_y	-	no	-	-	-	-	-
middle_fixM0.4_ar1y	fixed 0.4	ar1_y	-	yes	yes	5.1	-0.492	-0.074	-0.085
middle_fixM0.2_ar1y	fixed 0.2	ar1_y	-	yes	yes	-9.9	-0.517	-0.118	-0.021
middle_fixM0.4_iidy	fixed 0.4	iid_y	-	yes	yes	3.1	-0.499	-0.085	-0.085
middle_fixM0.2_iidy	fixed 0.2	iid_y	-	yes	yes	-1.4	-0.558	-0.098	0.002
middle_fixM0.3_ar1y	Fixed 0.3	ar1_y	-	yes	yes	37.3	-0.514	-0.141	-0.014

Table A3: South model sensitivity runs for mean natural mortality (M), random effects (RE), delta Akaike Information Criterion (dAIC), retrospective analysis Mohn's rho values for recruitment (R), spawning stock biomass (SSB) and average F (Fbar).

Model name	Mean M	M RE	NAA RE	Converged	All parameters estimable	dAIC	rho R	rho SSB	rho Fbar
south	fixed 0.3	iid_y	-	yes	yes	0	0.1085	0.0356	0.0312
south_fixM0.3	fixed 0.3	-	-	no	no	-	-	-	-
south_fixM0.3_NAAiid	fixed 0.3	-	iid	yes	yes	-33.4	-0.031	0.0579	0.1108
south_fixM_ar1y	fixed 0.3	ar1_y	-	yes	yes	-16.1	0.0922	0.0783	-0.07
south_estM_ar1y	estimated	ar1_y	-	yes	no	-	-	-	-
south_estM_iidy	estimated	iid_y	-	yes	no	-	-	-	-
south_fixM0.4_ar1y	fixed 0.4	ar1_y	-	yes	yes	-17.9	0.1359	0.0786	-0.062
south_fixM0.2_ar1y	fixed 0.2	ar1_y	-	yes	yes	-8.6	0.0423	0.0809	-0.082
south_fixM0.4_iidy	fixed 0.4	iid_y	-	yes	yes	1	0.1842	0.0671	-0.018
south_fixM0.2_iidy	fixed 0.2	iid_y	-	yes	yes	-1.2	-0.014	-0.008	0.1409
south_fixM0.5_ar1y	fixed 0.5	ar1_y	-	yes	yes	-20.3	-0.175	0.0842	-0.06