Fisheries and Oceans
Canada

Pêches et Océans
Canada

Ecosystems and
Oceans Science

Sciences des écosystèmes
et des océans

NAFO 4TVn Atlantic herring population models: from Virtual Population Analysis to Statistical Catch-at-Age estimating time-varying natural mortality

F. Turcotte, D. P. Swain, J. L. McDermid

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

May 2021

Canada

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/
csas-sccs@dfo-mpo.gc.ca

TABLE OF CONTENTS

LIST OF FIGURES

ABSTRACT

The most recent assessment of NAFO Division 4TVn spring and fall spawning Atlantic herring stocks was conducted in 2018 using Virtual Population Analysis models (VPA) with time-varying catchability. The increase in magnitude of a retrospective pattern in the spawning stock biomass (SSB) of the fall spawning herring models suggested that the model failed to incorporate one or more non-stationary processes in the population dynamics of this stock or in the observation model relating indices of abundance to population abundance. The 2018 spring spawning herring VPA also showed an increase in residual patterns in the catch-per-unit-effort (CPUE) and acoustic indices. Abundances of major predators of herring have changed drastically in the sGSL in the last decades, potentially generating important changes in herring natural mortality. Failure to account for changes in natural mortality due to changes in predator-prey interactions can result in biased estimates of population parameters and vital rates in stock assessments. Hence, estimating natural mortality was another motivation to re-explore 4T herring population models. The objectives of this paper are to perform a comparison of the VPA used in the most recent stock assessment and a series of Statistical Catch-at-Age (SCA) models with different assumptions about temporal variation in population processes (i.e., natural mortality) and/or observation processes (i.e., catchability in the fixed gear fishery). The aim is to determine the best performing model for the 2020 stock assessments. SCA models performed better than VPA models for both herring stocks. For the spring spawning herring stock, the SCA model estimating time-varying natural mortality and catchability to the CPUE index in the gillnet fishery (the qmSCA model) was the best performing model. In the fall spawning stock, the qSCA and the qmSCA models performed best, but the qmSCA was selected as the best model as it provided natural mortality estimates, an important parameter in 4TVn herring stock assessment. Retrospective patterns in SSB from this model must be monitored and the source of the pattern will be investigated using new data sources. Overall, the selected models offered improvements over VPA models used in previous assessments.

INTRODUCTION

Two Atlantic herring (*Clupea harengus*; hereafter herring) stocks occur in the southern Gulf of Saint-Lawrence (sGSL), a spring spawning and a fall spawning herring stock. Spring spawning occurs primarily from April to May but extends to June 30 at depths less than 10 m. Fall spawning occurs from mid-August to mid-October at depths of 5 to 20 m, but can occur as early as July 1. The spring spawning and fall spawning herring in Northwest Atlantic Fisheries Organization (NAFO) Division 4TVn are genetically distinct stocks and are assessed separately (Lamichhaney et al. 2017; Kerr et al. 2019; Fuentes-Pardo et al. 2019). Herring also show high spawning site fidelity (Wheeler and Winters 1984; McQuinn 1997; Brophy et al. 2006). Fall spawning herring in sGSL are therefore assessed using regionally-disaggregated assessment models (North, Middle, South regions; Figure 1).

The sGSL herring are harvested on spawning grounds by a gillnet fleet (referred to as "fixed" gear fleet) and a purse seine fleet ("mobile" gear fleet). Over time, the mobile gear fleet has consisted of varying numbers of large southern Gulf vessels (>19.8 m), and small seiners (<19.8 m) that can participate in the inshore fishery. The fixed gear fishery is focused in NAFO Div.4T, whereas the mobile gear fishery occurs in NAFO Div.4T and in Div.4Vn (1978-1998 only, Figure 1). During the spring and fall fishing seasons, the mobile fleet are prohibited from fishing in areas set aside exclusively for the fixed gear fleet (Claytor et al. 1998). In the spring fishery, mobile gear fleets fish along the northern boundary of NAFO region 4Tf, this is referred to as the "Edge" fishery. A global allocation or Total Allowable Catch (TAC) was first introduced in 1972 at 166,000 t, and reduced to 40,000 t in 1973. Separate TACs for the spring and fall spawning herring components began in 1985. The TACs were first allotted by fishing season and later attributed to spring or fall spawning herring landings based on biological samples taken during the fishery. The percentage of spring and fall spawning herring in the catch varies according to season and gear type. Both spring and fall spawning herring are harvested in the spring and fall fisheries and must therefore be separated into the appropriate groups to determine if the TAC for these groups has been attained. Since 1981, the fixed gear fleet has accounted for most of the catch of spring and fall spawning herring (McDermid et al. 2018).

The April 2015 framework for fall spawning herring reviewed and adopted new indices including an experimental net index, an acoustic index, and a multi-species bottom trawl index. The framework also adopted the use of time-varying catchability in the virtual population analysis (VPA) for fall spawning herring (Swain 2016b; Surette 2016; Surette et al. 2016). The 2016 herring stock assessment also adopted time-varying catchability for the spring herring VPA model (Swain 2016a). The most recent assessment of NAFO 4TVn spring and fall spawning Atlantic herring stocks was conducted in 2018 (McDermid et al. 2018) using VPA models with time-varying catchability. The increase in magnitude of a retrospective pattern in the spawning stock biomass (SSB) of the fall spawning herring models suggested that the model failed to incorporate one or more non-stationary processes in the population dynamics of this stock or in the observation model relating indices of abundance to population abundance (Appendix 1). The 2018 spring spawning herring VPA also started to show increasing residual patterns in the catch-per-unit-effort (CPUE) and acoustic indices (Appendix 1).

The abundance of major predators of herring (Benoît and Rail 2016) has changed drastically in the sGSL in the last decades (Atlantic cod, Swain et al. 2019; Grey seals, Hammill et al. 2014a and Atlantic Bluefin tuna, ICCAT 2017), potentially generating important changes in herring natural mortality. Failure to account for increases in natural mortality due to changes in predator-prey interactions can result in biased estimates of population parameters and vital rates in stock assessments (Overholtz et al. 2008; Legault and Palmer 2016; Jacobsen and

Essington 2018; Jacobsen et al. 2019). Hence, estimating natural mortality was another motivation to re-explore 4TVn herring populations models.

All previous population models of 4TVn herring have been VPA models. Statistical Catch at Age (SCA) provides an alternative modelling framework. VPA is backward projecting from abundance at age in the terminal (most recent) year; terminal abundances at age are parameters estimated in the model. SCA is forward projecting from abundance at age in the first year and at the first age in all years; these are estimated in the model, either as parameters (the approach used here) or by fitting a stock-recruit relationship.

SCA models offer statistical advantages over VPA, such as: 1) VPA assumes that the fishery catch-at-age is known without error, whereas SCA assumes that there is observation error in the proportions at age in the fishery catches. 2) VPA fits to the abundance indices at age and assumes that indices at different ages in the same year are independent, whereas SCA fits to the age-aggregated biomass indices and to the proportions at age in the fishery and survey catches. This accounts for the lack of independence between catches at different ages in the same year.

The objectives of this paper are to perform a comparison of the VPA used in the most recent stock assessment and a series of SCA models with different assumptions about temporal variation in population processes (i.e., natural mortality) and/or observation processes (i.e., catchability in the fixed gear fishery). The aim is to determine the best performing model for use in the 2020 stock assessments.

SPRING SPAWNING ATLANTIC HERRING MODELS

METHODS

Data

Data inputs for both the spring and fall spawning components are described in Turcotte et al. (2020).

Catch data comes from commercial fishery catches that are sampled dockside by the Department of Fisheries and Oceans (DFO) scientific personnel for the fixed and mobile fisheries, and at sea by fisheries observers in the mobile fishery. Catch-at-age and weight-at-age matrices (ages 4–11+,1978–2019) are derived for each of the fixed and mobile gear fleet catches. These were calculated using age-length keys and length-weight relationships for each spawning component, gear type and fishing season. The model inputs are fishery catch total biomass, proportions-at-age and weight-at-age.

The first index of abundance is a CPUE abundance index in the fixed gear fishery. Effort data comes from the dockside monitoring program and a telephone survey of harvesters. Fixed gear catch and effort data were used to construct catch-at-age and weight-at-age matrices (ages 4-10,1990–2019). The fixed gear fisheries occur on the spawning grounds and landings from this fishery account for approximately 60 % of the spring spawning herring catch. The model inputs are the fishery CPUE total biomass, proportions-at-age and weight-at-age.

The second index of abundance is generated from the annual fishery-independent acoustic/pelagic trawl survey. This index includes catch-at age and weight-at-age data from NAFO areas 4Tmno where both stocks mix at different life-stages and for different time frames. For spring spawning herring, model inputs are for ages 4 to 8 for years 1994 to 2019. The model inputs are the acoustic total biomass, proportions-at-age and weight-at-age.

Model structure

Age-structured population models (VPA and SCA), implemented in AD Model Builder (Fournier et al. 2012), were fit to the sGSL herring data. In this analysis, five models are compared for spring spawning herring stocks: a VPA model and four SCA models with and without time-varying fully-recruited catchability (q) to the fishery and/or natural mortality (M) parameters. The VPA model was used in the last assessment and is described in Swain 2016a, Swain 2016b and McDermid et al. 2018. Models are named as follows:

- qVPA (with time varying q, stationary M)

- qSCA (with time-varying q and stationary M)

- mSCA (with time-varying M and stationary q)

- qmSCA (with time-varying q and M)

- statSCA (with stationary q and M)

For the SCA models, estimated parameters include the numbers-at-age in the initial year (1978), yearly recruitment (average recruitment and recruitment deviations in numbers of age 2 fish), selectivity parameters in three time blocks (to account for changes in selectivity due to changes in length-at-age and mesh sizes used in the fishery), initial fishing mortality prior to 1978 (used to initialize abundance at age in the first model year), CPUE and acoustic survey q and yearly q deviations for the CPUE index, initial M and yearly M deviations for two age groups (2-6 and 7-11+) and the observation error in the indices.

VPA model parameters included log abundance at ages 5 to 11+ in the "terminal" year (i.e. 2020, the last model year plus 1), initial (1978-1990) fully-recruited q to the CPUE index, yearly (1991-2019) q deviations for the CPUE index, q to the acoustic index, and the standard deviation (SD) of observation error at age for each of the indices. In addition, two additional parameters were required to model selectivity-at-age to the CPUE index, which was modelled as a logistic function of age.

In VPA and SCA models where natural mortality was stationary, M was assumed to be constant at 0.2 for all years and ages. In models where q was stationary, a constant q parameter was estimated over years for each abundance index.

Natural mortality

For models allowing non-stationarity in M, independent time-series of M were estimated for two age groups: ages 2-6 (j = 1) and 7-11+ (j = 2). These time series were estimated on the log scale as random walks:

$$\log(M_{j,t}) = \log M_j^{init} \text{ where t= 1978}$$

$$\log(M_{j,t}) = \log(M_{j,t-1}) + Mdev_{j,t,} \text{ where t>1978}$$

$$Mdev_{j,t} \sim Normal(0, \sigma_j^M)$$

where $\log(M_j^{init})$ and $Mdev_{j,t,}$ are parameters estimated by the model. The M deviations ($Mdev_{j,t}$) were assumed to be normally distributed with a mean of 0 and SD σ_j^M (fixed at 0.075 for all j). The random walk started in 1979. Priors were supplied for M^{init}. These priors were normally distributed with means of 0.2 and standard deviations of 0.1 for both age groups (i.e., $M_j^{init} \sim N(0.2, 0.1)$).

The model likelihood included penalty terms due to the priors on M:

$$0.5 \sum_{j,y} (Mdev_{j,t}^2)/(\sigma_j^M)^2 + 0.5 \sum_J \exp(\log(M_j^{init}) - 0.2)^2/0.1^2$$

Catchability

For models that allow for process error in fully-recruited catchability (q) to the fixed gear fishery, the initial value of q in 1990 (the first year with CPUE data) was a model parameter and the subsequent values of q were estimated as a random walk:

$$q_t = \exp(\log q) \text{ where t = 1990}$$

$$q_t = q_{t-1} * \exp(qdev_{t,)} \text{ where t > 1990}$$

$$qdev_t \sim Normal(0, \sigma^q)$$

where $\log(q_t)$ and $qdev_{t,}$ are parameters estimated by the model. The q deviations ($qdev_t$) were assumed to be normally distributed with a mean of 0 and a SD σ^q fixed at 0.1.

The model likelihood included a penalty term due to the prior on the q deviations:

$$0.5 \sum_t (qdev_t^2)/(\sigma^q)^2$$

Selectivity

In the SCA models selectivity $S_{g,a,t}$ was indexed by catch source g, age a and year t. Fishery selectivity (g =1), selectivity to the CPUE in the gillnet fishery (g =2) and to the acoustic survey (g =3) were assumed to be logistic functions of age. For the commercial fishery, separate functions were fit to three time periods:

1. 1978 to 1989 (p =1),

2. 1990 to 2004 (p =2), and

3. 2005 to 2019 (p =3) (i.e $S_{1,a,p} = f(s_{1,a,t})$ and t ∈ 1978,1979, ... ,1989 for p =1, etc.). These time periods were chosen based on an examination of the yearly fixed/mobile gear proportions in the commercial fishery.

It could be argued that selectivity to the CPUE index and to the fishery may be dome shaped due to the use of gillnets. Selectivity models that allowed for a dome shape (e.g., double logistic, gamma, exponential logistic) were also examined and they did estimate that selectivity was dome shaped. The descending limb of the dome was steeper and declined to a lower level in the 2005-2017 period than in the 1990-2004 period. For example, using the above three selectivity models, selectivity at age 10 in the gillnet fishery was estimated to be about 0.5, 0.8 or 0.9 in 1990-2004 respectively and 0.2, 0.2 and 0.8 in 2005 to 2017 (see Appendix 2 for details). However, size at age of herring has been declining since the mid-1980s (Figure A2.1.1, Appendix 2). If selectivity was dome-shaped, old herring (e.g., age-10) would be on the descending limb. Consequently, decreases in size at age would increase their selectivity to the gillnet gear, not decrease it. Independent estimates of relative selectivity at age of fall spawners confirms that their selectivity at older ages has increased, not decreased, as their size at age has declined. Declining abundance at old ages that is not accounted for by fishery catches and estimated natural mortality can be spuriously accounted for by estimating declining selectivity at old ages. Consequently, these estimates of declining selectivity for older herring in recent years were judged to be spurious and the decision was made to use logistic selectivity models combined with the empirical estimates of relative selectivity at length. For the acoustic survey

the logistic model estimated selectivity to be 1.0 for all calibrated ages, consistent with the expectation that selectivity of acoustic estimates might be expected to be largely independent of age.

The qVPA selectivity to the CPUE in the gillnet fishery was modelled as a logistic function of age for ages 4 to 9, but was freely estimated for age 10 to allow for dome-shaped selectivity at age.

Initial abundance and recruitment

In SCA models, population abundance at age 2 (recruitment) in year t was estimated based on log average recruitment ($\bar{R}$) and annual recruitment deviations $Rdev_t$:

$$R_t = \exp(\bar{R} + Rdev_t)$$

$$Rdev_t \sim Normal(0, \sigma^R)$$

where $\bar{R}$ and $Rdev_t$ are parameters estimated by the model. The recruitment deviations ($Rdev_t$) were assumed to be normally distributed with a mean of 0 and σ^R fixed at 0.5. For older ages a ($a \in 3, 4, ... 11+$) in year 1, population abundance was estimated by projecting cohorts forward from age 2 in year 1 minus (a-2) to their age in year 1, as follows.

For abundance at age $a \in 3,4, ... A$-1 in year 1, where A is the last age (11+):

$$N_{a,1} = \exp(\bar{R} + Rdev_a^{r1} - \sum_{b=2}^{b=a-1} (s_{b,1}Fi + M_{b,1}))$$

For abundance at age A in year 1:

$$N_{A,1} = \frac{\exp(\bar{R} + Rdev_A^{r1} - \sum_{b=2}^{b=A-1}(s_{b,1}Fi + M_{b,1}))}{1 - \exp(-(s_{A,1}Fi + M_{A,1}))}$$

where $N_{a,1}$ is abundance at age a in year 1, $Rdev_a^{r1}$ are recruitment deviations used to initialize abundance at age a in year 1, $s_{b,1}$ is fishery selectivity at age b in year 1, Fi is fully-recruited fishing mortality for initializing abundance at age in year 1, $M_{b,1}$ is natural mortality at age b in year 1, and b indexes age in the summations.

The model likelihood included penalty terms due to the priors on the recruitment deviations used to initialize abundance at age 2 in all years and at older ages in year 1:

$$0.5 \sum_t (Rdev_t^2)/(\sigma^R)^2 + 0.5 \sum_a (Rdev_a^{ri})^2/(s^R)^2$$

After recruitment to age 2, cohorts were projected forward in the usual manner:

$$N_{a,t} = N_{a-1,t-1} \times \exp(-Z_{a-1,t-1})$$

$$Z_{a,t} = s_{1,a,t} \times F_t + M_{a,t}$$

where a and t index age and year, N denotes abundance, Z is total mortality, M denotes natural mortality, F is fully-recruited fishing mortality and $s_{1,a,t}$ is selectivity at age a in year t in the fishery.

In the qVPA model, because the youngest age in the abundance indices is age 3, it is not possible to obtain direct estimates of abundance at ages 2 to 3 in the terminal year T+1 and at age 2 in year T, where T is the last year with indices. These were obtained using the estimated average recruitment rate in the most recent five years and the estimated SSB producing a particular cohort.

Objective function

Depending on model structure (SCA, VPA, *q* and/or *M* estimation), the objective function for the models included the following components:

- All SCA models: discrepancies between observed and predicted values of the age-aggregated biomass indices for the CPUE in the gillnet fishery and acoustic survey. Indices were assumed to be lognormally distributed with standard deviations estimated by the model. Models allowed for weighting of the likelihoods of the biomass indices.

- All SCA models: discrepancies between observed and predicted proportions at age (PAA) in the fishery, CPUE and acoustic survey catches. The PAA were assumed to follow a multivariate logistic distribution, which estimates data variances. qVPA model: discrepancies between observed and predicted abundance at age (NAA) in the CPUE and acoustic survey catches. The NAA indices were assumed to be lognormally distributed with standard deviations estimated by the model.

In addition, particular models included the following likelihood components:

- SCA M models: a normal prior for the log M deviations,

- SCA M models: a normal prior for the initial values of M,

- SCA and VPA q models: a normal prior for the log q deviations,

- SCA models: a normal prior for the log recruitment deviations and

- SCA models: a normal prior for the log recruitment deviations used to calculate abundance at age in 1978.

SCA models index likelihoods were given different weights in the objective function calculation. Based on preliminary analysis of model fit to the age-aggregated indices with different weights, the CPUE biomass index likelihood was given a weight of one, while the acoustic biomass index likelihood was given a weight of three.

Approximate 95 % credible intervals were estimated based on 210,000 Markov chain Monte Carlo (MCMC) samples with the first 10,000 samples discarded and every 40th of the subsequent samples saved. All population estimates are posterior medians based on the MCMC sampling.

Model goodness-of-fit

Goodness-of-fit to indices was assessed by visual examination of estimated and observed aggregated biomass or abundance plots. Discrepancies between predicted PAA or NAA and observed PAA or NAA were assessed by plotting the residuals by year and age, and looking for "blocking" through ages or years. The sum of squares of the residuals were calculated for each index of abundance and compared between models (VPA residuals were transformed to PAA outside of the model). Retrospective patterns in SSB estimates were assessed by plotting SSB time-series estimated by sequentially removing the terminal year of data, for 4 years (2015 to 2019).

Important predators of herring for which abundance has changed over the time series were obtained from Swain et al. 2019 for age 5+ Atlantic cod, Hammill et al. 2014a for Grey seals and from the ICCAT 2017 rod and reel CPUE for the sGSL. As predator data were in different units, abundance indices for each predator and natural mortality estimates were rescaled between 0 and 1 to be comparable between different units of measurement, allowing to compare the timing and direction of changes in values.

RESULTS

Model fit

The qVPA model offered overall good fit to the age-aggregated CPUE abundance index (Figure 2). However, for the acoustic index, predicted values tended to overestimate observed values early in the time series (1995-2002) and underestimate observed values in recent years (2010-2019). The qSCA model also offered overall good fit to indices except for recent years (2018-2019), where fitted values are larger than observed values. The mSCA offered poor fit to the CPUE index, especially between 1990-1999, where fitted values are greater than the observed values. The mSCA model underestimated the data in the early years of the acoustic index. The stationary SCA offered the worst fit, with issues similar to both the qSCA and the mSCA, i.e. fitted values greater than data for the CPUE index in years 1990-1999, and fitted values greater than data in 2018-2019 for both indices. It also consistently underestimated the CPUE data from 2003 to 2015. The qmSCA offered the best fit to the age-aggregated biomass indices.

Figure 3 shows the residuals (observed–predicted) in proportions at age. The qVPA residuals from the CPUE index showed severe blocking and the sum of squared residuals was higher than all other models. Residuals were almost all negative in the beginning of the time series for older ages, and at the end of the time series for young ages. Residuals were almost all positive for ages 4 and 5 between 1990 and 2009, and for ages 7 to 10 between 2010 and 2019. Residuals form the acoustic survey did not show major patterns and were similar to other models. CPUE residuals in the SCA models also showed some blocking, although not as severe (i.e. compared to the CPUE residuals in the qVPA model). The CPUE residual patterns in the SCA models were very similar between the q and stationary models, and between the m and qm models. The sum of squared residuals (SSR) were greater for the former models (83) than the latter models (60-61). Residual patterns in the acoustic survey data were very similar between all SCA models and showed little blocking. The SSR in acoustic residuals were very similar among all these models (48-52).

Retrospective patterns refer to systematic changes (e.g., regular decreases or increases) in model estimates as years of data are added to the analysis. The occurrence of these patterns suggests that the model fails to take into account non-stationarities in population dynamics (e.g., time-varying M) or the observation process (e.g., time-varying catchability). The qVPA, qSCA and stationary SCA all showed retrospective patterns in constant declining direction in SSB estimates (Figure 4). There were no retrospective patterns in the mSCA and qmSCA model estimates of SSB.

Model estimates

The scale of spawning stock biomass (SSB, ages 4-11+) estimates changed with the inclusion of time-varying natural mortality (Figure 5). SSB estimates were higher in models m SCA and qm SCA than other models, mostly for years 1985 to 1995, and 2003 to 2015. SSB estimates for the terminal year (2019) were highest for the mSCA, qmSCA and statSCA models, intermediate for the qSCA model, and lowest for the qVPA model. Overall, SSB trends were similar for all models from 1978 until 2003, then only models estimating *M* showed a modest increase in SSB up to 2010.

The addition of time varying *M* also increased the estimated scale of recruitment (the number of age-2 fish, Figure 5). The temporal trends in recruitment were generally similar among models except that estimated number of recruits were much greater in the 2005-2015 period in the

mSCA and qmSCA models relative to the other models. Estimated recruitment in the qVPA model also remains very low in recent years compared to the other models.

Trends in estimated abundance-weighted fishing mortality (ages 6-8) were similar between models (Figure 5). Estimated F were lowest for the mSCA and qmSCA models, especially since 2009. This reflects the high estimated M during this period. In the recent period, estimated F was highest for the qVPA model, reflecting the very low estimates of recent SSB in this model.

Estimated fully-recruited CPUE q was stationary at 0.00519 for the mSCA model, and at 0.00851 for the statSCA model. In other models CPUE q was allowed to vary beginning in 1991, the second year of the CPUE index time series. Estimated CPUE q in the qVPA model increased from 0.00327 in 1990 to 0.03597 in 2019, much higher than the maximum estimates in the other q models. Estimated CPUE q in the qSCA model increased from 0.00361 in 1990 to a peak of 0.01678 in 2009 and then declined to 0.00938 in 2019. Among models with time-varying CPUE q, estimates were lowest in the qmSCA model in all years. For this model, q increased from 0.00222 in 1990 to a peak of 0.00896 in 2007 and then declined to about 0.008 in the 2010-2019 period (Figure 5). In all models estimating time-varying q, catchability increased as SSB declined (Figure 6).

Natural mortality was fixed at 0.2 for the qVPA, the qSCA, and the statSCA models. Estimated M trends were similar within age groups between the mSCA and qmSCA models (Figure 5). M estimates for the age group 2-6 varied between 0.22 and 0.46 over the time series, with highest estimates around 1980 and between 2000 and 2012, and lowest estimates between years 1988 and 1995. M estimates in 2019 for the age group 2-6 were about 0.33. M estimates for the age group 7-11+ increased gradually from 0.21 to 0.59 between 1978 and 2010. Starting in 2011, estimates sharply increased to reach 1.37 in 2016. Estimates were stable at around 1.32 in recent years (2017 to 2019). Natural mortality for age group 2-6 did not seem to correlate with Atlantic cod abundance (Pearson's r = -0.21), while the estimated M trend for age group 7-11+ significantly correlated with the summed grey seal and Atlantic bluefin tuna relative abundance trend over the same time period (Pearson's r = 0.91, Figure 19).

FALL SPAWNING ATLANTIC HERRING MODELS

METHODS

Data

For the fall spawning herring stock, a regionally-disaggregated model for three regions (North, Middle, South) that encompass the entire NAFO Div.4T area was used. The regions are defined on the basis of traditional herring spawning beds and fishing areas (Figure 1):

- North (Gaspé and Miscou; 4Tmnopq)

- Middle (Escuminac-Richibucto and west Prince Edward Island; 4Tkl)

- South (east Prince Edward Island and Pictou; 4Tfghj)

Catch data comes from commercial fishery catches that are sampled dockside by the DFO scientific personnel for the fixed and mobile fisheries, and at sea by fisheries observers in the mobile fishery. Catch-at-age and weight-at-age matrices (ages 4-11+,1978–2018) are derived for each of the fixed and mobile gear fleet catches. These were calculated using age-length keys and length-weight relationships for each spawning component, gear type, fishing season, and region. The model inputs are fishery catch total biomass, proportions-at-age and weight-at-age. Starting in this assessment, historical seiner catch from 4Vn was re-distributed in the

North, Middle and South regions in proportion to the region's fixed gear landing. In previous assessment, re-distribution was based on seiner landings in each region, resulting in regions without seiner landings not receiving catch redistribution from 4Vn seiner landings. Also starting in this assessment, seiner catch from the edge fishery was re-distributed in North, Middle and South regions proportionally to fixed gear landings. Previously, landings in the edge fishery were all attributed to the south region.

The first index of abundance is a CPUE abundance index in the fixed gear fishery. Effort data comes from dockside monitoring program and a telephone survey of harvesters. Fixed gear catch and effort data were used to construct catch-at-age and weight-at-age matrices (ages 4-10,1986–2018) for all stocks. The fixed gear fisheries occur on the spawning grounds and landings from this fishery account for more than 90 % of the fall spawning herring catch. The model inputs are CPUE total biomass, proportions-at-age and weight-at-age.

The experimental net (multiple panels of varying mesh size) sampling program provides mean abundance-at-age and weight-at-age by year and region (North, Middle, South, ages 3 to 9, 2002 - 2018). Some changes were made to the index since the 2018 assessment. For this assessment, the catch at length of each mesh size was summed per day per region, and then the mean catch at length per region per year was calculated. The catch at age data was then constructed using age-length keys. Samples with zero catches were included in the analysis and no correction factor was applied to the catch at age to account for soak time. The experimental nets data was also used to estimate changes in selectivity at age to the commercial gillnet fishery (Surette et al. 2016). The model inputs are experimental nets catches total biomass, proportions-at-age and weight-at-age.

The fishery-independent annual multi-species bottom trawl survey (hereafter RV survey) is conducted each September since 1971. This survey provides catch-at- age (mean number per standardized tow) and weight-at-age of FS herring (age 4-6, 2002-2018). Prior to this assessment, the catch rates used for this index were based on predicted values from a model external to the assessment model. In this assessment, the actual observed data was used, rather than model predictions. The model inputs are bottom trawl catch total biomass, proportions-at-age and weight-at-age.

The fourth index of abundance is generated from the annual fishery-independent acoustic/pelagic trawl survey. This index includes catch-at age data from NAFO areas 4Tmno where spring and fall spawning herring mix at different life-stages and for different time frames. Juvenile fall spawners concentrate in the area in late summer and autumn. Hence, this index is considered a recruitment index (ages 2 and 3) for fall spawners, for years 1994 to 2018. The model inputs are acoustic total biomass, proportions-at-age and weight-at-age.

Models structure

Age-structured population models (VPA and SCA), implemented in AD Model Builder (Fournier et al. 2012), were fit to the sGSL herring data. In this analysis, five models are compared for fall spawning herring stocks: a VPA model and four SCA models with and without time-varying fully-recruited catchability (q) to the fishery and/or natural mortality (M) parameters. The VPA model was used in the last assessment and described in Swain 2016a, Swain 2016b and McDermid et al. 2018. Models are named as follows:

- qVPA (with time varying q, stationary M)
- qSCA (with time-varying q and stationary M)
- mSCA (with time-varying M and stationary q)

- qmSCA (with time-varying *q* and *M*)

- statSCA (with stationary *q* and *M*)

For each region (North, Middle, South), the SCA models included the following estimated parameters: the numbers-at-age in the initial year (1978), yearly recruitment (average log recruitment and recruitment deviations in numbers of age 2 fish), selectivity parameters for each source of catch, initial fishing mortality prior to 1978, *q* for each index and yearly *q* deviations for the CPUE index, *M* and yearly *M* deviations for two age groups (2-6 and 7-11+) and the observation error to the indices.

The fall stock VPA was a model including time-varying *q* and stationary *M*. VPA model parameters are, for each region (North, Middle, South), abundance at ages 4 to 11+ at the beginning of the most recent year (2019), *q* at age to the experimental nets, the RV survey and the acoustic survey, the SD of observation error at age for each of these indices, fully-recruited *q* to the CPUE index in 1986 and earlier years, annual *q* deviations to the CPUE index for 1987 to 2018, 2 parameters for an average selectivity-at-age logistic function for the CPUE index, and the SD of observation error at age for this index. For terminal abundance, all CPUE parameters and catchability at age to the experimental nets, separate parameters were estimated for each of the three populations.

In VPA and SCA models where natural mortality was stationary, *M* was assumed to be constant at 0.2 for all years and ages. In models where *q* was stationary, a single *q* parameter was estimated over years for each abundance index.

Time-varying natural mortality *M* and catchability to the CPUE gillnet fishery *q*, initial abundance in 1978 and recruitment in 1979 to 2019 were all estimated as described for the spring spawning models, with parameters independently estimated for each region (North, Middle, South). The population was projected forward as described for the spring spawning herring assessment, except that the beginning of the fishing season was set at August 1 instead of April 1. Fall SCA models had the same objective function components has described for the spring spawning herring assessment models.

Size-at-age of 4TVn herring has been declining since at least the mid-1980s (Figure 7). This is expected to result in changes in the selectivity-at-age of herring to the gill-net fishery. Historically, two mesh sizes have been used in this fishery, 2 5/8" and 2 3/4". Changes in selectivity-at-age to these mesh sizes were estimated as follows. First, relative selectivity at length was estimated for these mesh-sizes using data from the experimental nets (Surette et al. 2016). These nets consisted of a range of mesh sizes from 2" to 2 3/4". Then selectivity-at-length was converted to relative selectivity-at-age in each year based on the age-length keys for each year. Annual selectivity-at-age functions for the CPUE indices ($S_{t,a}^{Ca}$,Figure 8) were incorporated in the models as follows:

$$S_{p,t,a}^{Ca} = S_{p,a}^{C} * \left(\left(Pr_{p,t}^{258} * rS_{t,a}^{258} \right) + \left(1 - Pr_{p,t}^{258} \right) * rS_{t,a}^{234} \right)$$

where $S_{p,a}^{C}$ is a time-invariant population-specific logistic selectivity curve for the CPUE fishery, $Pr_{p,t}^{258}$ is the proportion of nets in year *t* and population *p* that are of mesh size 2 5/8", $rS_{t,a}^{258}$ is relative selectivity to mesh size 2 5/8" for age *a* in year t, $rS_{t,a}^{234}$ is relative selectivity to mesh size 2 3/4" for age *a* in year *t*, and $S_{p,t,a}^{Ca}$ is selectivity to the fishery CPUE index for age *a* in population *p* and year *t*. $S_{p,a}^{C}$ was included in the equation to convert from the relative to absolute scale. A similar procedure was used to adjust selectivity of the multimesh experimental nets and the fishery for changes in size-at-age. For the experimental nets, selectivity at length was the average of the values for the seven mesh sizes used. For the commercial fishery, $S_{p,a}^{C}$ was

estimated separately for three time periods to take into account changes in the proportion of mobile gear catches in the fishery.

The procedure for converting selectivity at length to annual selectivity at age changed slightly from earlier years. Previously, selectivity at age *a* in year t was based on the length distribution of the CPUE catch at age *a* in year t. However, this calculation should be based on the length distribution at age *a* before selection by the fishery. We adopted this approach here, calculating the length distribution at age in the population before fishing by dividing the length distribution at age in the catch by the selectivity at length.

Based on preliminary analysis of model fit to the age-aggregated indices and retrospective analysis, likelihoods for the biomass indices were given different weights depending on the model used. For the qSCA and qmSCA models, likelihood weights were 4 for the CPUE biomass index, 0 for the experimental nets index, and 1 for the RV and acoustic survey biomass indices. This improved fit to the indices and reduced retrospective patterns. For the statSCA, mSCA and qVPA models, likelihoods were given a weight of 1 for all biomass indices to avoid Hessian matrices that were not positive definite.

Approximate 95 % credible intervals were estimated based on 210,000 MCMC samples with the first 10,000 samples discarded and every 40th of the subsequent samples saved. All population estimates are posterior medians based on the MCMC sampling. Goodness-of-fit was assessed as described for spring models, but retrospective analysis results was also assessed using Mohn's Rho, using the icesAdvice R package (Magnusson et al. 2018).

RESULTS

Model fit

The qSCA and qmSCA models offered overall good fit to the age-aggregated CPUE biomass index for all regions. The qVPA model offered an acceptable but less good fit for the North and South region for most years. The mSCA offered poor fit to the CPUE index especially in the North and Middle region. The stationary SCA offered the worst fit, especially in the first years in the North and South region, where fitted values were all larger than observed values and in the South between 2007 and 2015 where all fitted values were smaller than observed values (Figure 9).

Preliminary analysis showed that predicted annual age-aggregated experimental nets catch biomass showed little correspondence with the observed indices, and thus the age-aggregated biomass index based on the experimental nets was not used in this assessment. The qSCA and qmSCA experimental net index likelihood were given a weight of zero in the objective function (the equivalent of removing the data from the model). As a result, fit to the age-aggregated biomass index was poor for those two models, and uncertainty was very large. All other models generally showed mediocre to poor fits to the index in all regions. The statSCA showed poor fit for the North region, were predicted values were either larger or smaller than observed values on most years except the first 5 years of the time-series. Fit was acceptable in the Middle region. In the South region, predicted values were larger than observed values in most of the years until 2013, and then predicted values were smaller than observed values in subsequent years. The mSCA showed poor fit, especially in the Middle and South regions, where fitted values were larger than observed values around 2010, and smaller than observed values in the South region for all recent years. The observed values in the South region showed a V-shaped pattern, which was not reflected in the predicted pattern for any of the models. The qVPA predicted values showed a flat trend through the observed values for the North and Middle regions, whereas the predicted values in the South roughly matched the declining trend in

observed values prior to 2015 but failed to match the subsequent increase in observed values. Overall, fit to the age-aggregated biomass index was poor for the experimental nets in all models (Figure 10).

Fit to the age-aggregated RV survey biomass index was poor for all models except the mSCA and qmSCA. In these two models, fit was good between years 2005 and 2018 whereas predicted values were often smaller than observed values in early years (Figure 11). Fit to the age-aggregated acoustic survey biomass index was generally poor for all models, except for the mSCA and qmSCA for years between 2005 and 2018 (Figure 11).

The qmSCA model showed a better fit to indices overall, equal to the qSCA for the fit to the CPUE index, but better fit to the RV and acoustic survey indices.

Residual patterns

Some blocking was evident between observed and predicted fishery proportions at age in the statSCA and qSCA model residuals. In the North region, residuals were mostly positive for ages 2 and 6 to 9 between 1980 and 2008 and mostly negative for ages 3 and 4. All models showed larger negative residuals for younger and older ages in recent years. All models also showed larger residuals for ages 2, 3 and 8 to 10. Residuals were generally smaller for ages between 3 and 7. Overall, the mSCA and qmSCA showed the smallest SSR of fishery proportions at age residuals (Figure 12). VPA assumes catch is known without error, so there were no fishery residuals from that model.

SCA models showed no major blocking between the observed and predicted values residuals of the CPUE index proportions at age. The qVPA model showed similar residual patterns across regions. Residuals were positive for ages 4 to 5 between 1986 and 2005, and for ages 7 to 11+ in recent years. Residuals were generally negative for ages 7 to 11+ between 1986 and 2005, and for ages 4 to 6 in recent years. Between the SCA models, the mSCA and qmSCA showed the lower SSR in all regions, though differences were small (Figure 13).

Some blocking was evident between observed and predicted experimental nets proportions at age in all SCA models. In all regions but especially in the Middle and South regions, residuals were negative for ages 5 to 7 in years 2002 to 2010, positive for ages 3 and 4 for the same years, and then negative for ages 3 and 4 between 2010 and 2018. The qVPA showed the most severe residuals pattern with mostly positive residuals in ages 3 to 5 in early years, and ages 7 to 9 in recent years. SSR was also higher in the qVPA compared to all SCA models. Overall SCA models, the mSCA and qmSCA had the lowest SSR in all regions (Figure 14).

No residual patterns were apparent in the RV survey and acoustic survey proportions at age. All models show large negative residuals to the acoustic survey age 2 for years 2011 and 2012. Of all SCA models, the statSCA and qSCA models had the lowest SSR in the RV survey proportions at age, whereas there was very little difference between models in the acoustic survey SSR. The qVPA SSR was higher than for all SCA models for both indices (Figure 15).

The mSCA and qmSCA showed the least residual patterns and lowest SSR of all models.

Retrospective analysis

All models showed retrospective patterns in SSB estimates, with various magnitudes and directions in the patterns. The stationary SCA showed mostly negative peels and rho values around -0.2 in all regions. The qSCA showed similar patterns but with more positive peels in the North and South regions, hence showing rho values closer to zero for those regions. The mSCA showed negative peels in all regions, the magnitude of the negative pattern being greater in the North and Middle regions, with rho values of -0.48 and -0.41, respectively. The qmSCA also

showed negative patterns for all regions, but of lesser magnitude than the mSCA for the North and Middle regions, but of greater amplitude for the South region. Rho values for the qmSCA ranged between -0.35 and -0.38. The qVPA showed a negative peel followed by positive peels in the North region with a rho of -0.14. The Middle region showed a mix of negative and positive peels with a rho value -0.19 while the pattern retrospective pattern was strongly negative in the South region with a rho of -0.36 (Figure 16).

For the North region, the qSCA had least retrospective patterns of all models. The qVPA showed a low rho value but a single peel is negative while all following peels are positive. Of SCA models estimating natural mortality, the qmSCA model had lesser retrospective patterns. For the Middle region, the qVPA, qSCA and statSCA showed retrospective patterns of lesser magnitude. Of SCA models estimating natural mortality, the qmSCA model again had lesser retrospective patterns. For the South region, the qSCA showed the least retrospective patterns, while the statSCA and mSCA showed negative patterns, and the qmSCA and qVPA negative patterns of greater magnitude (Figure 16).

Overall, the qSCA performed best in regard to the retrospective analysis. Of models estimating time-varying natural mortality, the qmSCA performed best, while showing consistent negative retrospective patterns.

Model estimates

Estimated scale of spawning stock biomass (SSB, ages 4-11+) varied between models, as expected when the estimated natural mortality in some models exceeded the value assumed in other models. SSB estimates were higher in models mSCA and qmSCA compared to other models, mostly for years between the 2000s and 2018, in all regions. All models estimating time-varying natural mortality showed a sharp increase in SSB between the 2000s and 2010-2012, followed by a sharp decline in SSB until the terminal year (Figure 17).

In the North region, the mSCA model showed the highest SSB estimates in the 2008-2018 period, reaching 212 kt. The qmSCA showed a similar trend, but with a lower maximum SSB at 157 kt. The qSCA and statSCA both showed stable SSB around 65 kt for that time period, while the qVPA showed a decreasing trend from 54 to 26 kt. SSB estimates for the terminal year (2018) in the North region were similar for all SCA models (between 54 and 64 kt), but lower for the qVPA (26 kt). SSB estimates for years prior to 2008 varied only slightly between models (Figure 17).

In the Middle region, the mSCA and qmSCA showed similar trends in SSB, with the mSCA showing a slightly higher maximum value in 2011 at 122 kt. Both models showed a sharp increase in SSB starting in 2008, followed by a sharp decrease starting in 2011. The qSCA, statSCA and qVPA all showed similar trends between them for that time-period, with a small increase in SSB starting in 2008 reaching around 50 kt, followed by a decline until the terminal year. SSB estimates for the terminal year were highest for the mSCA (28 kt) and qmSCA (38 kt), and lower but similar for the qSCA (13 kt), statSCA (13 kt) and qVPA (18 kt). SSB estimates for years prior to 2008 varied only slightly between models (Figure 17).

In the South region, the mSCA and qmSCA showed similar trends in SSB. SSB increased following 2002 to reach around 300 kt, and then decreased sharply after 2011. The qSCA, statSCA and qVPA showed similar trends between them, with a generally constant decline in SSB starting in 2002. SSB estimates for the terminal year were highest for the mSCA (53 kt) and qmSCA (49 kt), and lower for the qSCA (11 kt), statSCA (18 kt) and qVPA (27 kt). SSB estimates for years prior to 2002 varied only slightly between models (Figure 17).

Estimated recruit abundance varied between regions, with the North region showing the greatest variation over the whole time-series. The number of recruits is lower for the mSCA and qmSCA between 1978 and 1994, and then all SCA models showed higher number of recruits than the qVPA, in all years until the terminal year. A peak in recruitment in the late 2000s was only present in the SCA models estimating natural mortality. In all models, number of recruits declined around 2010 until the terminal year. In the Middle and South regions, all models showed similar trends and scale in number of recruits, except slightly higher values for the mSCA and qmSCA between 1978-1988 and in the late 2000s. In the Middle region, estimates in recent years showed an increase in SCA models, whereas the qVPA model showed a decline. In the South region, all SCA models showed an increase in recruitment in the last two years, while the qVPA shows a decline (Figure 17).

Estimated abundance weighted fishing mortality (ages 6-8) trends were similar between models in all regions between years 1978 and 2000s. From that period on, SCA models with time-varying M showed lower estimated fishing mortality than the stationary SCA, the qSCA and the qVPA. These three models showed similar trends, except in the North region where the qVPA showed stable values between 2010 and the terminal year, and the qSCA and statSCA showed a decline in fishing mortality over the same time period. In the North region, terminal year estimates for all SCA models were around 0.2, and at 0.6 for the qVPA. In the Middle region, terminal year estimates for SCAs estimating M were between 0.15 and 0.2, and between 0.32 and 0.49 for other models. In the South region, terminal year estimates were near 0.08 for the SCA models estimating M, 0.13 for the qVPA, 0.21 for the statSCA, and 0.38 for the qSCA (Figure 18).

Estimated fully-recruited CPUE q was stationary at values between 0.020 and 0.028 for the statSCA depending on regions, whereas estimated values for the mSCA model were stationary between 0.010 and 0.013. The models estimating time-varying q showed different trends between regions and between models. In the North region, the qSCA and the qmSCA showed similar trends with an increase until 2000 followed by a decrease, but the qSCA estimates were larger over the time-series. The qVPA trend was different from SCA models, showing a gradual increase from the beginning of the time-series until a plateau was reached in 2013, at the highest values of all models.

In the Middle region, models estimating time-varying q all showed similar trends until the early 2000s where the qmSCA estimates declined to lowest values around 2010, and increases slightly afterwards. The qSCA and qVPA showed similar trends with some variation between 2000 and 2015, where the qSCA estimates increased rapidly, reaching a higher value than the qVPA for the terminal year.

In the South region, the qmSCA and the qSCA increased until the late 1990s, followed by a decline until the terminal year for the qmSCA and by an increase until the terminal year for the qSCA. The qSCA and the qVPA showed similar trends with some variation around the year 2000 and in the terminal years where the qVPA estimates were lower (Figure 18). In the qmSCA model, catchability increased as SSB declined, with some variation (Figure 20). In the qSCA and qVPA models, catchability increased as SSB declined with some variation in the North and South regions, but seemed to vary independently of SSB in the Middle region.

Natural mortality was fixed at 0.2 for the qVPA, the qSCA, and the statSCA. Estimated M trends were similar within age groups between M models and between regions. For ages 2-6 in the North region, natural mortality estimates were low and slightly increasing or decreasing for the qmSCA and mSCA, respectively, to reach similar low values around 0.007 in recent years. In the Middle and South regions, estimates were stable around 0.3 or 0.4 depending on region, followed by a gradual decline around 1990, to reach similar values than in the North region in

the terminal year. For the age group 7-11+, estimates from all regions increased gradually from around 0.2 in the beginning year to around 0.3 in 2005. Starting in 2006, estimates sharply increased to reach a plateau between 0.8 and 1.0, depending on the region (Figure 18). The M estimates for age group 2-6 significantly correlated with the trend in Atlantic cod sGSL abundance in the North (Pearson's r = 0.95), Middle (Pearson's r = 0.94) and South (Pearson's r = 0.90) regions. For the age group 7-11+, the estimated M trends significantly correlated with the summed grey seal and Atlantic bluefin tuna relative sGSL abundance indices in the North (Pearson's r = 0.98), Middle (Pearson's r = 0.98) and South (Pearson's r = 0.99) regions (Figure 19).

DISCUSSION

BEST MODEL

For the spring spawning 4TVn Atlantic herring stock, the qmSCA was the best model based on its better fit to indices, absence of blocking in residuals, low sum of squares of residuals and absence of retrospective patterns. This model accounted for non-stationarity by allowing catchability to the gillnet fishery and natural mortality to vary over time.

For the fall spawning herring stock, the qSCA and qmSCA were the best performing models. They both showed a better fit to the CPUE index, and the qmSCA showed a better fit to the RV and acoustic survey indices. The qmSCA also showed the least sum of squares of residuals. However, retrospective patterns in the qmSCA SSB estimates were second to worst of all models. The patterns were negative and in a constant direction, indicating that the quantity being evaluated is consistently underestimated. The qSCA retrospective patterns were less severe. However, the qSCA model doesn't provide natural mortality estimates and if M did increase over 0.2 in the time-series, this model certainly underestimates SSB and potentially presents biased recruitment and fishing mortality estimates.

The rationale for choosing the qmSCA as the favored model even in the presence of a retrospective pattern is based on the following points:

1. In the qmSCA models used here, time-varying M was estimated independently for four different population models (spring spawners, North, Middle and South region fall spawners), all showing very similar trends.

2. M for age group 7-11+ showed important changes over the time series in all qmSCA models, as expected from predator abundance information. The timing, direction, and rate of change in M7-11+ trends for all models are very similar to the trend in combined major predator abundance change for the same time period. The trends in younger herring (ages 2-6) were also similar to Atlantic cod abundance. It is possible to use auxiliary information or covariates to assist M estimation from the model (e.g. Marty et al. 2003; Deriso et al. 2008). Here, no covariate was provided to the model. However the estimates from all models were estimated independently and correlated well with major herring predator abundance in the sGSL.

3. The qmSCA model for the spring spawners did not show a retrospective pattern, and the M trends in the fall stock models are almost identical. Predator effects on spring and fall herring stocks were expected to be fairly similar.

4. Fisheries management is often based on the assumption that natural mortality is constant through time, yet numerous examples show that predator-prey interactions are dynamic (Lee et al. 2011, Thorson et al. 2015; Skern-Mauritzen et al. 2016; Jacobsen and Essington 2018; Siple et al. 2018). Failure to account for increases in natural mortality due

to changes in predator-prey interactions in a stock assessment will result in biased estimates of population parameters and vital rates (Overholtz et al. 2008; Legault and Palmer 2016; Jacobsen and Essington 2018; Jacobsen et al. 2019).

In the case of choosing between the qSCA and the qmSCA as the best model for 4TVn fall spawning herring, the tradeoff is accepting a model with a retrospective pattern underestimating SSB from year to year (qmSCA), or accepting a model not accounting for natural mortality (qSCA). Changes in the natural mortality of herring in the sGSL are expected to have occurred given the large changes observed in the abundances of important predators of herring. In all four herring populations examined here, the estimated changes in natural mortality were very similar among populations and were strongly correlated with the changes in predator abundances. These indepenedent observations provide strong support for the changes in natural mortality estimated by the qmSCA model, and we have thus chosen this model as the basis for advice.

The scale of underestimation of SSB by qmSCA can be used in the risk assessment when considering uncertainty. The "true" SSB value within the MCMC confidence interval will more than likely reside above the median estimate, in a scale that is proportional to Mohn's rho. As Mohn's rho is similar between the three regions, the scale of the bias towards SSB underestimation can be expected to be similar. Retrospective analysis and Mohn's rho should be investigated every year to detect changes in the direction and scale of patterns. A negative value for the rho statistic means that the quantity being evaluated is consistently being underestimated (when compared with the estimate from the full time-series) and is potentially less problematic than overestimation in terms of sustainability (Hurtado-Ferro et al. 2015).

Different strategies should be explored to identify the source of the residual patterns. As recruitment, catchability to the fishery and natural mortality are all allowed to be time-varying and selectivity is estimated in time-blocks, it is likely that any change in population dynamics can be accounted for in these models. Thus, the source of the retrospective pattern may be a conflict between the catch and the biomass indices and the age data or the lack of sufficient data to calibrate the population dynamics in the M models. The SSB retrospective pattern may be a consequence in the delay of estimating changes in M because of the penalty on non-zero M deviations. As new years of data supporting a change in M are added to the model, the penalty is out-weighted by the data, and M is allowed to change, generating a change in SSB. It may be inevitable unless highly informative data is added to the model to support quicker detections of M and SSB changes.

Additional data that could be used to address this issue for fall stock models include: 1) including a broader range of ages in the acoustic survey data, and 2) incorporating the spawning ground acoustic survey data. The current acoustic survey index and age proportions include ages 2 and 3 only, but information is available for older ages. The spawning ground acoustic surveys started in 2015 and now include five years of data. This industry collaborative survey provides an average nightly biomass estimate on each spawning ground, surveyed up to five times during the spawning season. Due to its large spatial and temporal coverage of biomass dynamics on all major spawning grounds, the addition of these data to population models will provide a well-informed biomass index. Age-composition for the index will be obtained from the experimental nets survey, sampled at the same locations at the same frequency.

MODEL ESTIMATES

Potential sources of natural mortality for both stocks include unreported catches, disease and predation. Unreported catches of herring probably mostly come from the bait fisheries. Catches

in bait fisheries were historically not accounted for in the assessments of either spring or fall spawning herring components. Catches in these fisheries are meant to be recorded in harvester logbooks but compliance with the requirement to complete and return logbooks to DFO is low. Catches of herring in the bait fishery are expected to be much lower than landings in the commercial fishery. Nonetheless, this unaccounted fishing mortality is now accounted for in the natural mortality estimates. Disease mortality is expected to be relatively small in 4TVn herring, as no disease-related mortality events were recorded in the time period covered by the assessment.

Herring is a vital pelagic prey species for numerous predators in the sGSL including grey seals (*Halichoerus grypus,* Hammill and Stenson 2000; Hammill et al. 2007, 2014b), seabirds (Cairns et al. 1991), cetaceans (Fontaine et al. 1994; Benoît and Rail 2016), Atlantic cod (*Gadus morua*, Hanson and Chouinard 2002), white hake (*Urophycis tenuis*, Benoît and Rail 2016) and Atlantic Bluefin tuna (*Thunnus thynnus*, Pleizier et al. 2012). Of these major predators, abundances of cod, grey seals and Bluefin tuna have changed drastically in the sGSL in the last decades. Hence, herring natural mortality was expected to change over time. Grey seals are the main pinniped predators of marine fish in the sGSL (Hammill and Stenson 2000). Increases in the abundance of Grey Seal occurring in the sGSL have been linked with important increases in the mortality of several demersal fish stocks that are declining in abundance or failing to recover from fishery-induced collapse (Benoît et al. 2011; Swain and Benoît 2015; Neuenhoff et al. 2019). The West Atlantic tuna stock biomass declined in the 1970s to its lowest level where it remained for more than two decades, then began a gradual increase from 2004 to reach 60 % of the 1974 biomass in 2013 (ICCAT 2017). Abundance of cod ages 5+ was high in the late 1970s before the stock collapsed in the late 1980s and early 1990s, and has kept declining since (Neuenhoff et al. 2019). Seabird abundance (northern gannets (*Morus bassanus*), double-crested cormorants (*Phalacrocorax auritus*) and great cormorants (*P. carbo*)) also increased in the sGSL between the 1970s and the 2000s, and all are herring consumers (Benoit and Rail 2016). However, more analyses of their distribution, diet and the scale of the increase in abundance (cormorants) are necessary before drawing links with estimated herring natural mortality. Information on consumption by cetaceans is also very scarce.

Catchability to commercial fisheries is expected to increase over time as technological improvements are implemented. Herring harvesters have reported changes in fishing procedures which have not been incorporated in effort standardization. Increases in fishing efficiency resulting from these changes in fishing procedures may underlie the estimated increase in catchability. Catchability to fisheries may also be density-dependent, increasing as population size decreases (Winters and Wheeler 1985). For the spring spawning herring models, *q* increased as SSB declined, suggesting a density-dependent effect. In the fall spawning herring models, the qmSCA model suggested mostly density-dependent effects but with more variation in the relationship, suggesting the influence of technological improvements in *q* estimates. This mix of the effects of density-dependence and technological improvements was greater in qVPA and qSCA models.

CONCLUSION

SCA models performed better than VPA models for both herring stocks. For the spring spawning herring stock, the SCA model estimating time-varying natural mortality and catchability to the gillnet fishery was the best performing model. In the fall spawning stock, the qSCA and the qmSCA performed best, but the qmSCA was selected as the best model as it offered natural mortality estimates, an important parameter in 4TVn herring stocks population dynamics. Retrospective patterns in SSB from this model must be monitored and the source of

the pattern will be investigated using new data sources. Overall, the selected models offered improvements over VPA models used in previous assessments.

REFERENCES CITED

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.

Benoît, H. P., Swain, D. P., Bowen, W. D., Breed, G. A., Hammill, M. O., Harvey, V. 2011. Evaluating the potential for grey seal predation to explain elevated natural mortality in three fish species in the southern Gulf of St. Lawrence. Mar Ecol Prog Ser 442:149-167.

Brophy, D., Danilowicz, B. S., and King, P. A. 2006. Spawning season fidelity in sympatric populations of Atlantic herring (*Clupea harengus*). Canadian Journal of Fisheries and Aquatic Sciences, *63*(3), 607–616.

Cairns, D. K., Chapdelaine, G., and Montevecchi, W. A. 1991. Prey exploitation by seabirds in the Gulf of St. Lawrence. *In* The Gulf of St. Lawrence: small ocean or big estuary? pp. 277-291. Ed by J. -C. Therriault. Canadian Special Publication of Fisheries and Aquatic Sciences. 113.

Claytor, R., LeBlanc, C., MacDougall, C., and Poirier, G. 1998. Assessment of the NAFO Division 4T southern Gulf of St. Lawrence Herring stock, 1997. DFO Can. Sci. Advis. Sec. Res. Doc. 98/47. 154 p.

Deriso, R. B., Maunder, M. N., and Pearson, W. H. 2008. Incorporating covariates into fisheries stock assessment models with application to Pacific herring of Prince William Sound, Alaska. Ecological Applications, 18: 1270–1286.

Fontaine, P. -M., Hammill, M. O., Barrette, C., and Kingsley, M. C. S. 1994. Summer diet of the harbour porpoise (*Phocoena phocoena*) in the estuary and the northern Gulf of St. Lawrence. Canadian Journal of Fisheries and Aquatic Sciences, 51: 172–178.

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. Optimization Methods & Software. 27:2, 233-249.

Fuentes-Pardo, A. P., Bourne, C., Singh, R., Emond, K., & Pinkham, L. 2019. Adaptation to seasonal reproduction and thermal minima-related factors drives fine-scale divergence despite gene flow in Atlantic herring populations Affiliations : *BioRxiv Preprint*, *1*(902), 1–48.

Hammill, M. O., and Stenson, G. B. 2000. Estimated prey consumption by harp seals (*Phoca groenlandica*), hooded seals (*Cystophora cristata*), grey seals (*Halichoerus grypus*) and harbour seals (*Phoca vitulina*) in Atlantic Canada. Journal Northwest Atlantic Fishery Science, 26: 1-23.

Hammill, M. O., Stenson, G. B., Proust, F., Carter, P., and McKinnon, D. 2007. Feeding by grey seals in the Gulf of St. Lawrence and around Newfoundland. *In* Grey seals in the North Atlantic and the Baltic, pp. 135–152. Ed. T. Haug, M. Hammill, D. Olafsdottir. NAMMCO Scientific Publication 6.

Hammill, M.O., den Heyer, C.E., and Bowen, W.D. 2014. Grey Seal Population Trends in Canadian Waters, 1960-2014. DFO Can. Sci. Advis. Sec. Res. Doc. 2014/037.

Hammill, M. O., Stenson, G., Swain, D. P., and Benoît, H. P. 2014b. Feeding by grey seals on endangered stocks of Atlantic cod and white hake. ICES Journal of Marine Science 71: 1332-1341.

Hanson, J. M., and Chouinard, G. A. 2002. Diet of Atlantic cod in the southern Gulf of St. - Lawrence as an index of ecosystem change, 1959-2000. Journal of Fish Biology, 60: 902–922.

Hurtado-Ferro, F., Szuwalski, C. S., Valero, J. L., Anderson, S. C., Cunningham, C. J., Johnson, K. F., Licandeo, R., McGilliard, C. R., Monnahan, C. C., Muradian, M. L., Ono, K., Vert-Pre, K. A., Whitten, A. R., & Punt, A. E. 2015. Looking in the rear-view mirror: Bias and retrospective patterns in integrated, age-structured stock assessment models. ICES Journal of Marine Science, *72*(1), 99–110.

ICCAT. 2017. Report of the 2017 ICCAT Bluefin stock assessment meeting. Madrid, Spain, 20-28 July 2017.

Jacobsen, N. S., and Essington, T. E. 2018. Natural mortality augments population fluctuations of forage fish. Fish and Fisheries, 19: 791–797.

Jacobsen, N. S., Thorson, J. T., and Essington, T. E. 2019. Detecting mortality variation to enhance forage fish population assessments. ICES Journal of Marine Science, 76: 124-135.

Kerr, Q., Fuentes-Pardo, A. P., Kho, J., McDermid, J. L., and Ruzzante, D. E. 2019. Temporal stability and assignment power of adaptively divergent genomic regions between herring (*Clupea harengus*) seasonal spawning aggregations. Ecology and Evolution, 9, 500-510.

Lamichhaney, S., Fuentes-Pardo, A. P., Rafati, N., Ryman, N., McCracken, G. R., Bourne, C., Singh, R., Ruzzante, D. E., and Andersson, L. 2017. Parallel adaptive evolution of geographically distant herring populations on both sides of the North Atlantic Ocean. Proc. Nat. Acad. Sci. 114(17), E3452-E3461.

Lee, H. H., Maunder, M. N., Piner, K. R., and Methot, R. D. 2011. Estimating natural mortality within a fisheries stock assessment model: an evaluation using simulation analysis based on twelve stock assessments. Fisheries Research, 109: 89-94.

Legault, C. M., and Palmer, M. C., 2016. In what direction should the fishing mortality target change when natural mortality increases within an assessment. Canadian Journal of Fisheries and Aquatic Sciences, 73: 349-357.

Magnusson, A., Millar, C., and Cooper, A. 2018. icesAdvice: Functions Related to ICES Advice. R package version 2.0-0.

Marty, G. D., Quinn, T. J., Carpenter, G., Meyers, T. R., and Willits, N. H. 2003. Role of disease in abundance of a Pacific herring (*Clupea pallasi*) population. Canadian Journal of Fisheries and Aquatic Sciences, 60: 1258–1265.

McDermid, J.L., Swain, D.P., Turcotte, F., Robichaud, S.A., and Surette, T. 2018. Assessment of the NAFO Division 4T southern Gulf of St. Lawrence Atlantic Herring (*Clupea harengus*) in 2016 and 2017. DFO Can. Sci. Advis. Sec. Res. Doc. 2018/052. xiv + 122 p.

McQuinn, I. H. 1997. Metapopulations and the Atlantic herring. *Reviews in Fish Biology and Fisheries*, 7, 297–329.

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. Canadian Journal of Fisheries and Aquatic Sciences, 76: 168–184.

Overholtz, W. J., Jacobson, L. D., and Link, J. S. 2008. An ecosystem approach for assessment advice and biological reference points for the Gulf of Maine–Georges Bank Atlantic herring complex. North American Journal of Fisheries Management, 28: 247–257.

Pleizier, N. K., Campana, S. E., Schallert, R. J., Wilson, S. G., and Block, B. A. 2012. Atlantic Bluefin tuna (*Thunnus thynnus*) diet in the Gulf of St. Lawrence and on the Eastern Scotian shelf. Journal of Northwest Atlantic Fishery Science, 44: 67–76.

Siple, M. C., Shelton, A. O., Francis, T. B., Lowry, D., Lindquist, A. P., and Essington, T. E. 2018. Contributions of adult mortality to declines of Puget Sound Pacific herring. ICES Journal of Marine Science, 75: 319–329.

Skern-Mauritzen, M., Ottersen, G., Handegard, N. O., Huse, G., Dingsør, G. E., Stenseth, N. C., and Kjesbu, O. S. 2016. Ecosystem processes are rarely included in tactical fisheries management. Fish and Fisheries, 17:165-175.

Surette, T. J. 2016. Abundance indices of Atlantic herring (*Clupea harengus*) from the southern Gulf of St. Lawrence based on the September multispecies bottom trawl survey. DFO Can. Sci. Advis. Sec. Res. Doc. 2016/064. vii + 33 p.

Surette, T. J., LeBlanc, C. H., and Mallet, A. 2016. Abundance indices and selectivity curves from experimental multi-panel gillnets for the southern Gulf of St. Lawrence fall herring fishery. DFO Can. Sci. Advis. Sec. Res. Doc. 2016/067. vi + 23 p.

Swain, D.P. 2016a. Population modelling results for the assessment of Atlantic Herring (*Clupea harengus*) stocks in the southern Gulf of St. Lawrence (NAFO Division 4T) to 2015. DFO Can. Sci. Advis. Sec. Res. Doc. 2016/061. x + 53 p.

Swain, D.P. 2016b. Assessment framework for fall-spawning Atlantic herring (*Clupea harengus*) in the southern Gulf of St. Lawrence (NAFO Div. 4T): Population models and status in 2014. DFO Can. Sci. Advis. Sec. Res. Doc. 2016/066. x + 58 p.

Swain, D. P. Benoît, H. P. 2015. Extreme increases in natural mortality prevent recovery of collapsed fish populations in a Northwest Atlantic ecosystem. Mar Ecol Prog Ser 519:165-182.

Swain, D.P., Ricard, D., Rolland, N., and Aubry, É. 2019. Assessment of the southern Gulf of St. Lawrence Atlantic Cod (*Gadus morhua*) stock of NAFO Div. 4T and 4Vn (November to April), March 2019. DFO Can. Sci. Advis. Sec. Res. Doc. 2019/038. iv + 105 p.

Thorson, J. T., Monnahan, C. C., and Cope, J. M. 2015. The potential impact of time-variation in vital rates on fisheries management targets for marine fishes. Fisheries Research, 169: 8–17.

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

Wheeler, J. P., and Winters, G. H., 1984. Homing of Atlantic herring (*Clupea harengus*) in Newfoundland waters as indicated by tagging data. Can. J. Fish. Aquat. Sci. 41: 108-117.
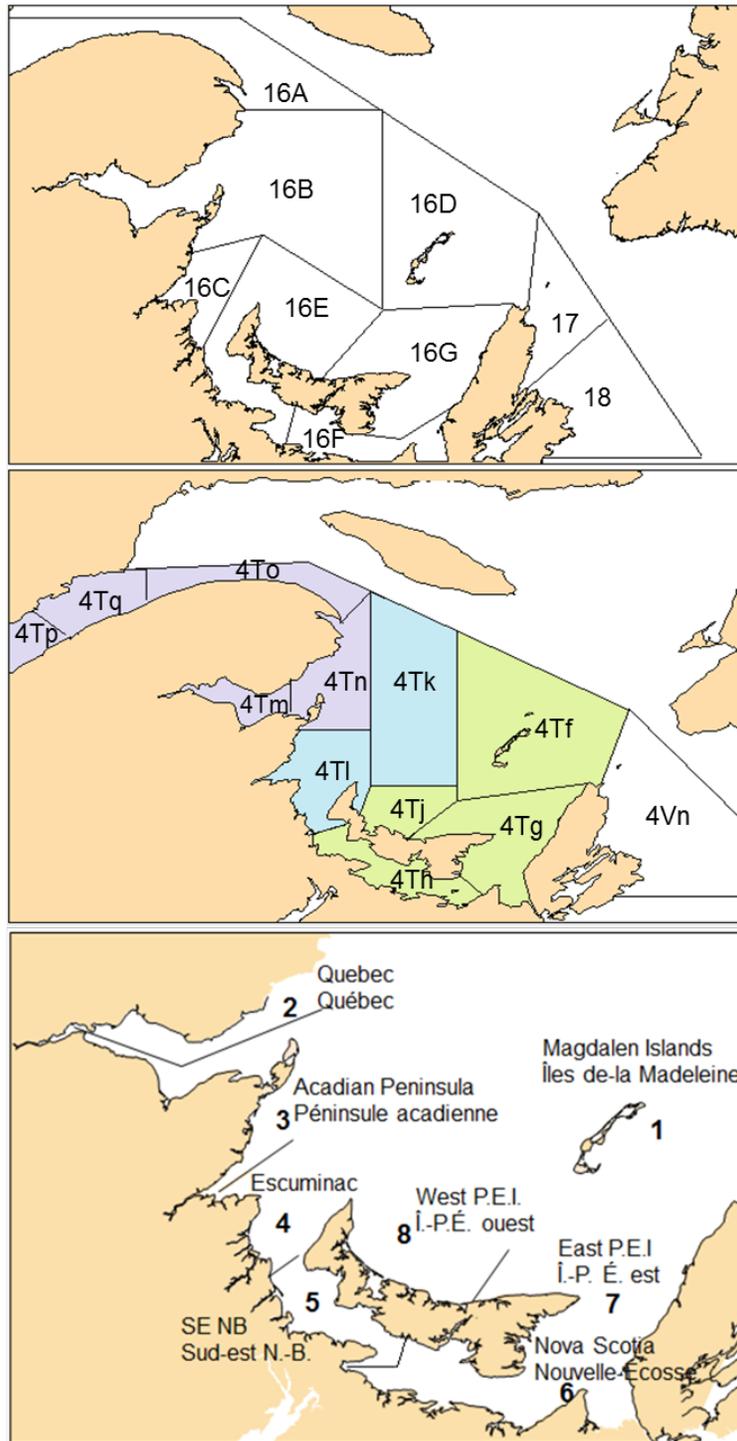
Figure 1. Southern Gulf of St. Lawrence herring fishery management zones (upper panel), Northwest Atlantic Fisheries Organization (NAFO) divisions 4T and 4Vn, where purple represents the North region, blue = Middle region, and green = South region (middle panel), and geographic areas used in the telephone survey of the herring gillnet fishery (lower panel).
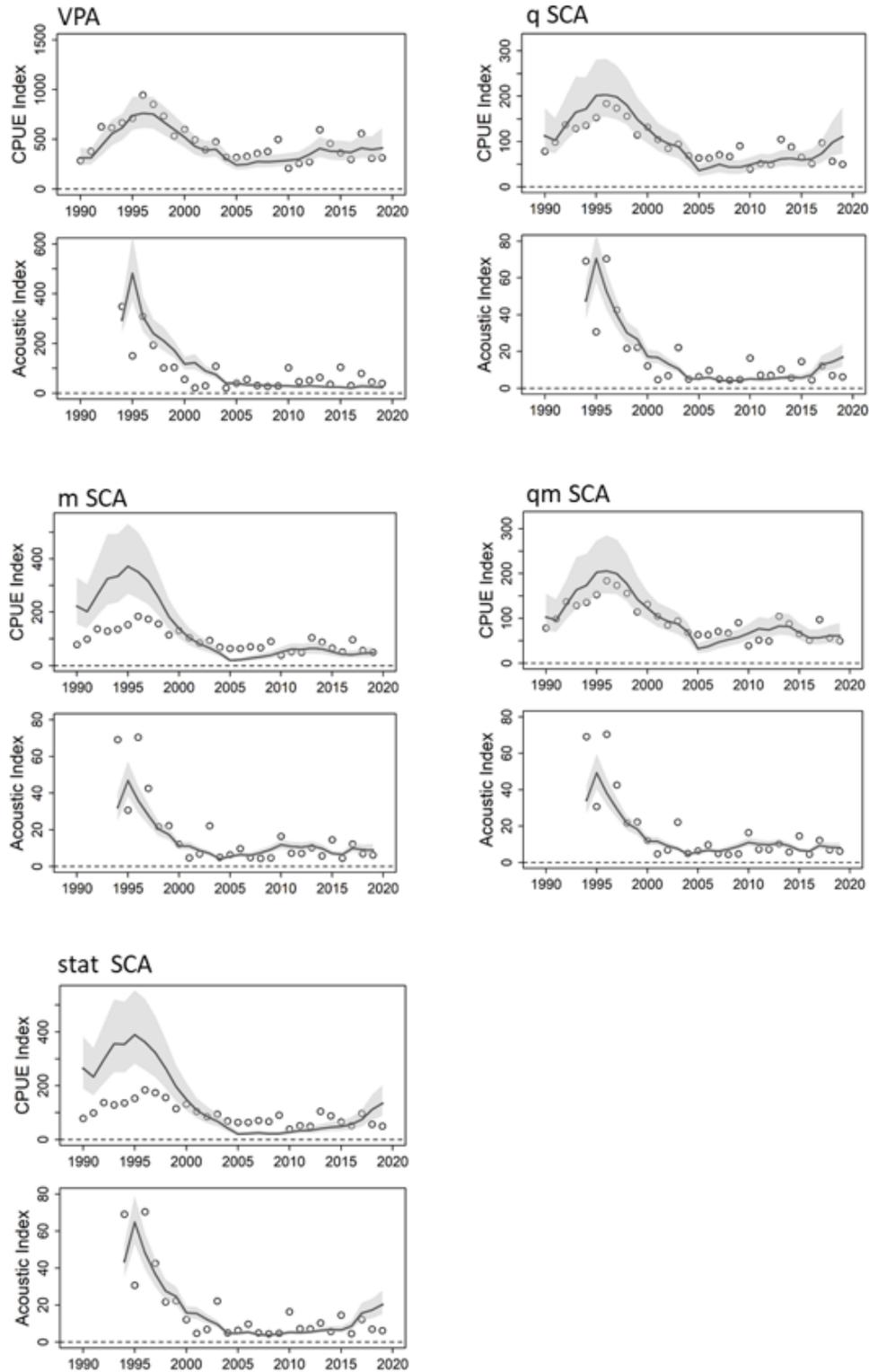
Figure 2. Observed (circles) and predicted (lines and shading) age-aggregated CPUE (upper panels) and acoustic (lower panels) fit to indices for spring spawning Atlantic herring models (VPA, qSCA, mSCA, qmSCA and statSCA). The lines show the median predicted indices and the shading the 95 % confidence intervals of the predictions based on MCMC sampling.
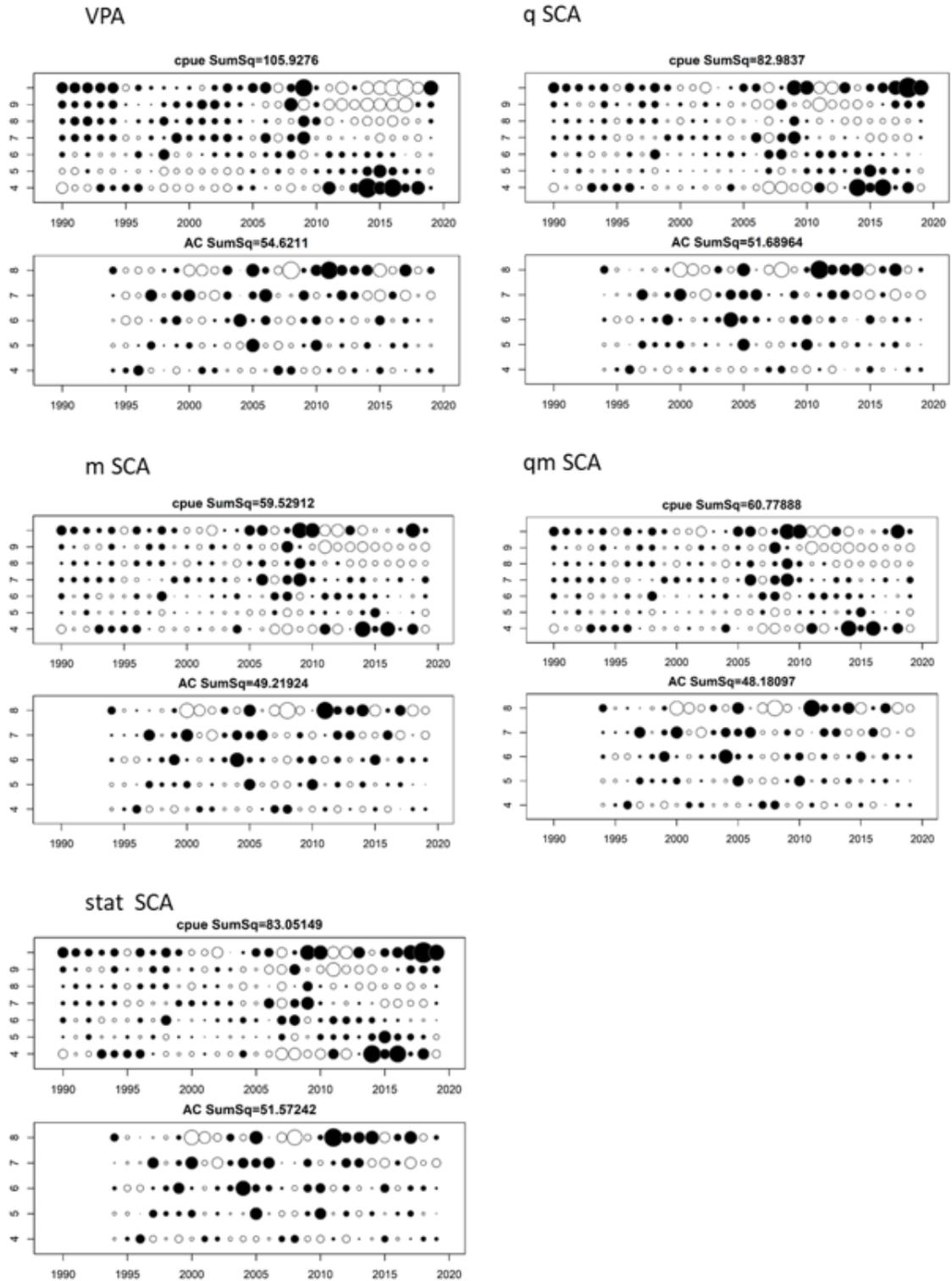
Figure 3. Residuals in proportions at age for spring spawning Atlantic herring population models (VPA, qSCA, mSCA, qmSCA and statSCA). The upper panel shows residuals for the CPUE index and the bottom panel shows residuals for the acoustic index. Rows are for ages and columns are for years. Circle radius is proportional to the absolute value of residuals. Black circles indicate negative residuals (i.e., observed < predicted) and white circles indicate positive residuals.
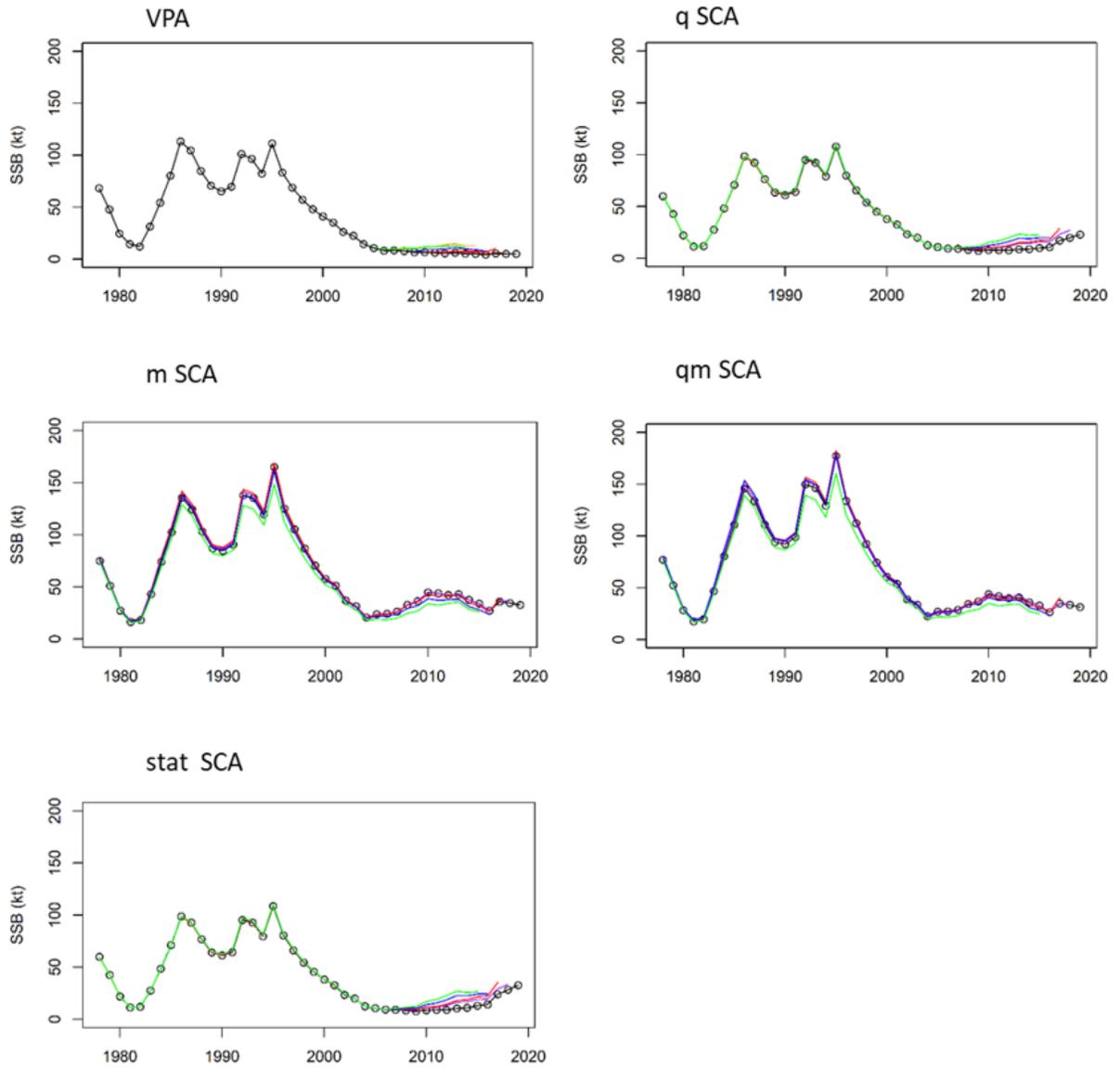
Figure 4. Retrospective patterns in estimated spawning stock biomass (SSB, kt) of ages 4 to 11+ for spring spawning Atlantic herring population models (VPA, qSCA, mSCA, qmSCA and statSCA).
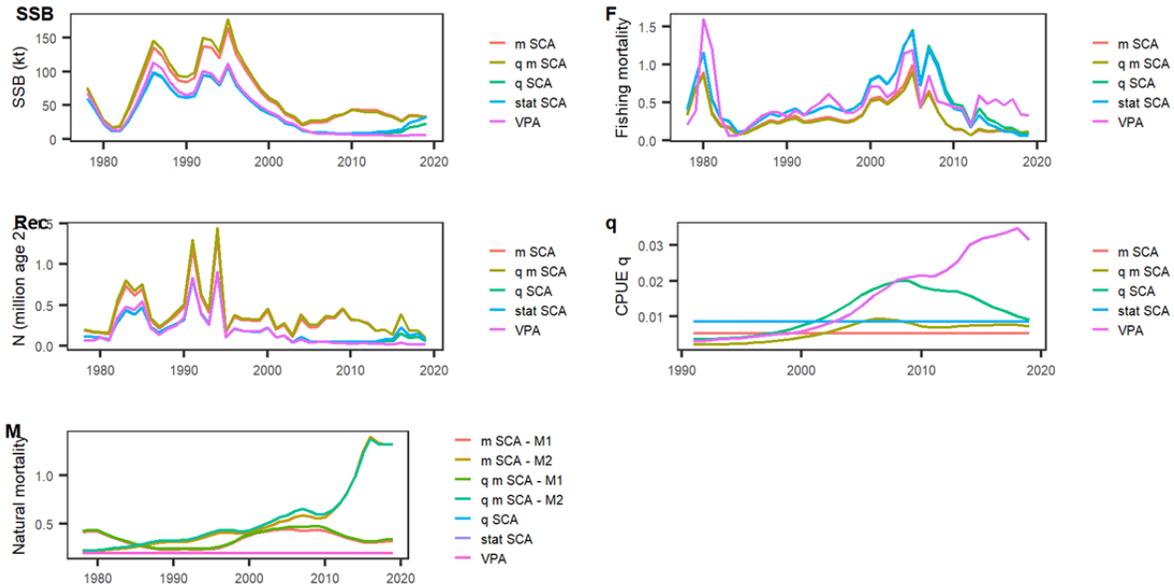
Figure 5. Median MCMC estimates plots. SSB plot: Spawning Stock Biomass (SSB kt), F plot: Fishing mortality, Rec plot: Number of age 2 fish (recruitment, in millions), q plot: fully-recruited catchability to CPUE q, M plot: Natural mortality for ages 2-6 (M1) and 7-11+ (M2), for spring spawning Atlantic herring population models (VPA, qSCA, mSCA, qmSCA and statSCA).
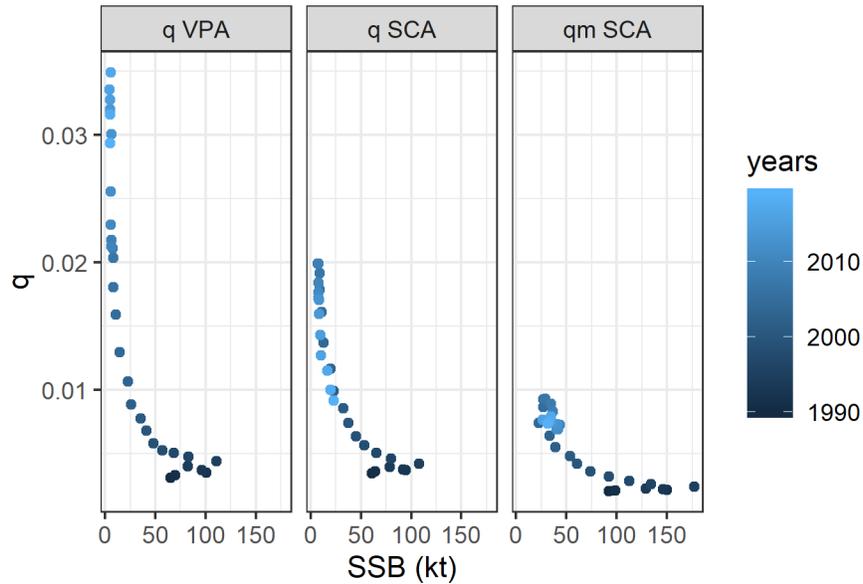
Figure 6. Fully-recruited catchability to the CPUE gillnet fishery (q) in function of SSB (tons) for 4TVn spring spawning herring population models estimating time-varying q.
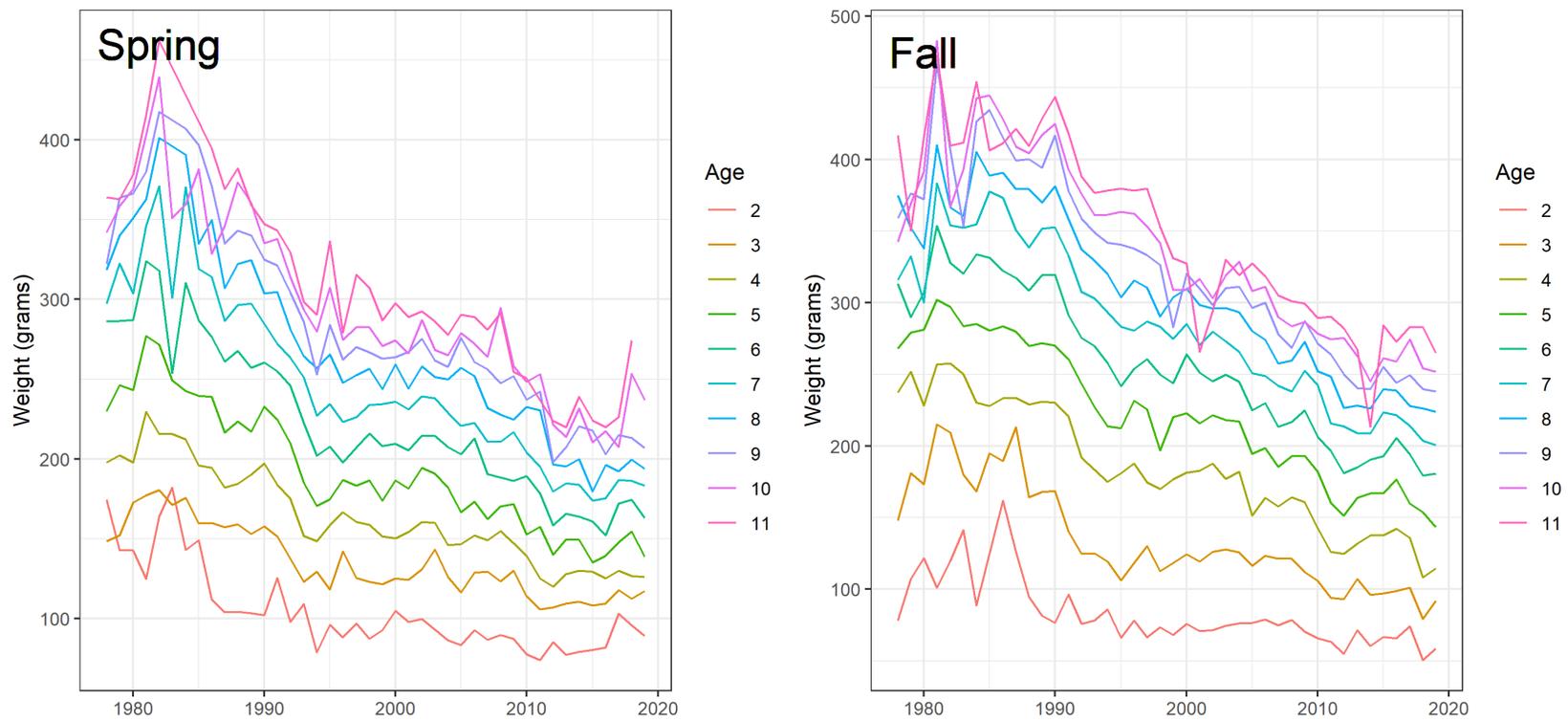
Figure 7. Weight at age of spring spawning (left panel) and fall spawning (right panel) Atlantic herring between 1978 and 2019.
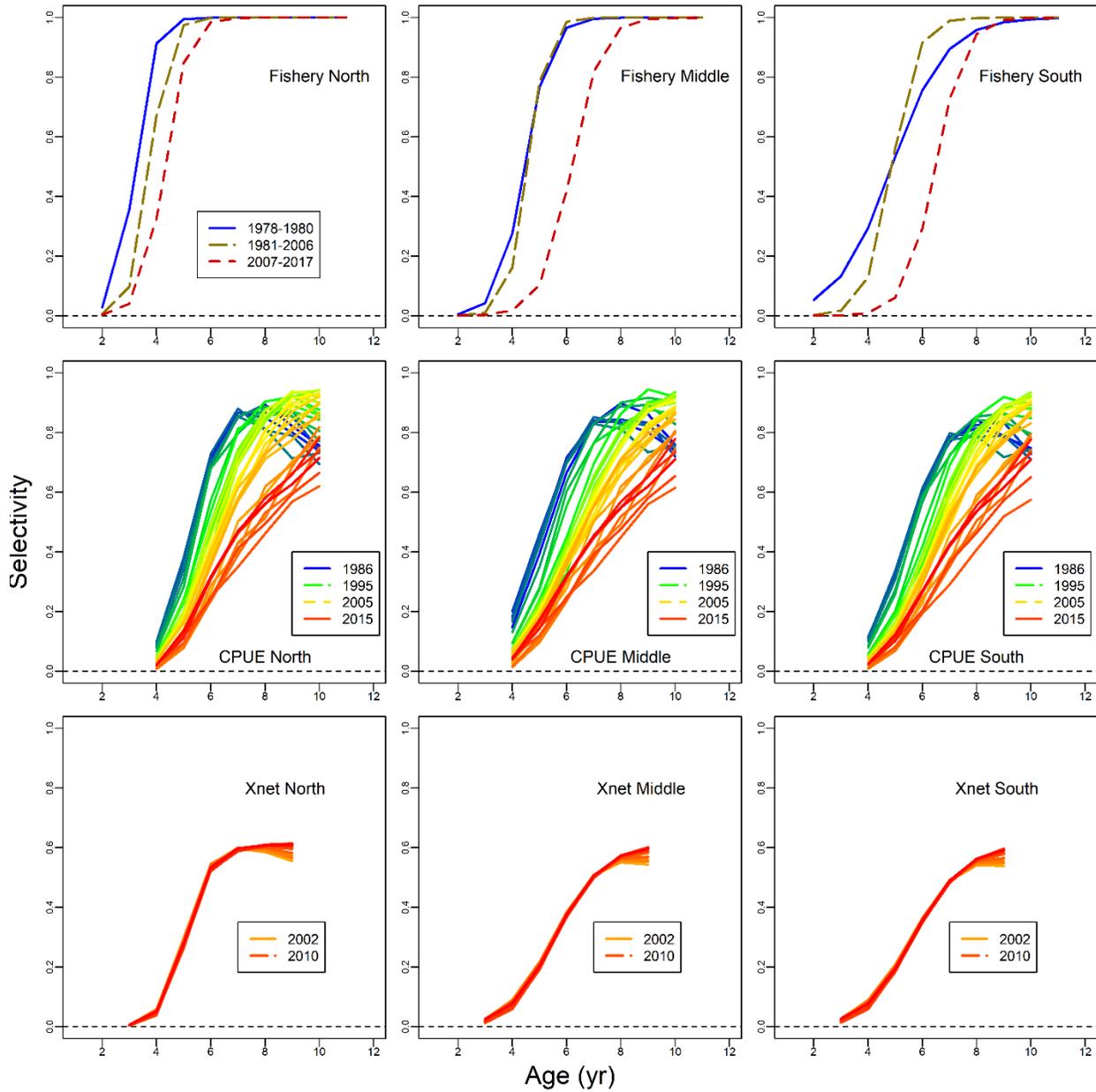
Figure 8. Estimated selectivity-at-age to the total commercial fishery (top row), the CPUE index for the fixed-gear portion of the fishery (middle row) and the experimental nets (bottom row). Results are shown for the qSCA model of the North (left column), Middle (center column) and South (right column) fall spawning Atlantic herring populations.
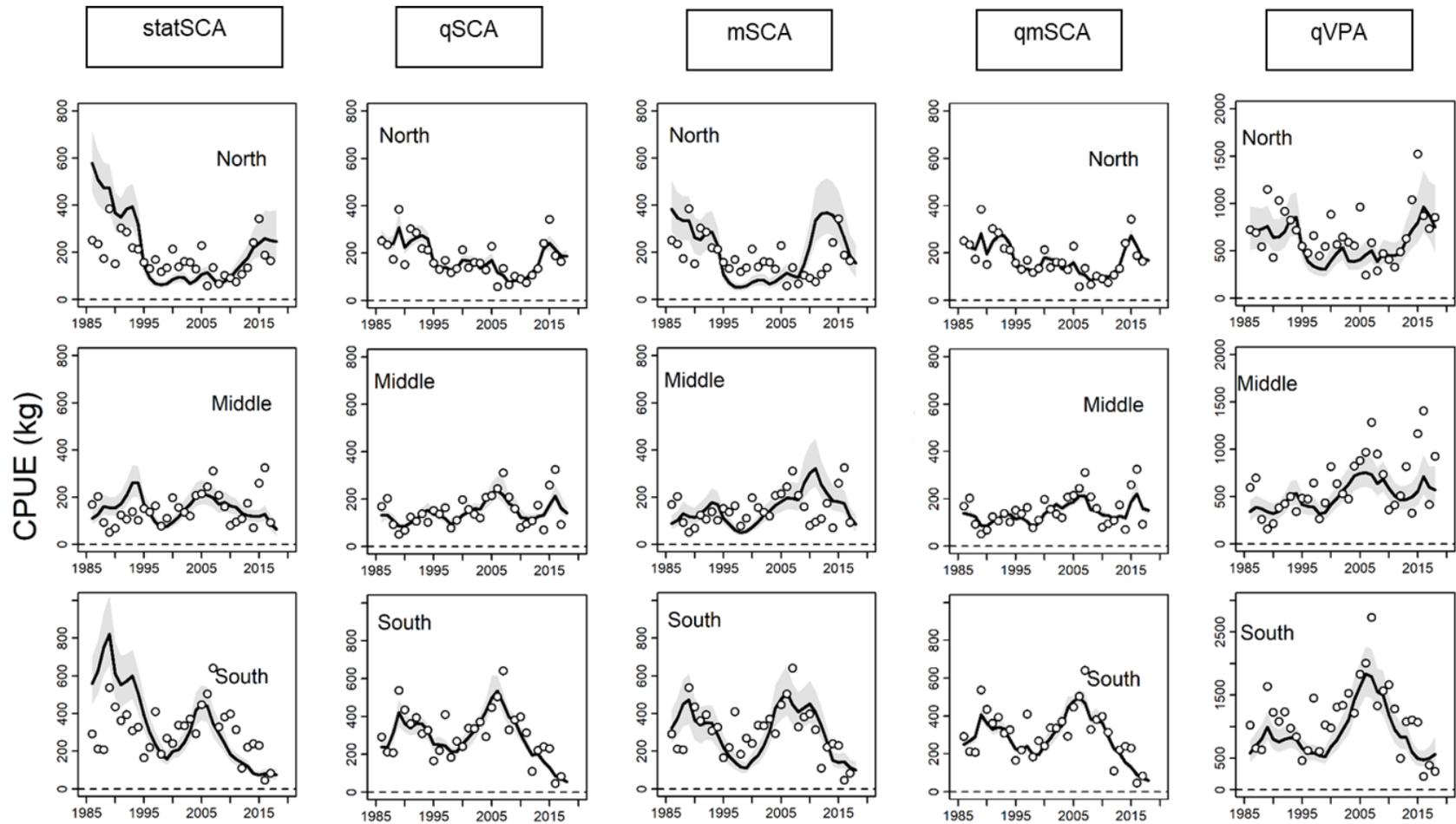
Figure 9. Observed (circles) and predicted (lines) age-aggregated CPUE fit to indices for 4TVn fall spawning Atlantic herring models (statSCA, qSCA, mSCA, qmSCA and qVPA). The lines show the maximum likelihood predicted indices of biomass (kg). The lines show the median predicted indices and the shading the 95 % confidence intervals of the predictions based on MCMC sampling.
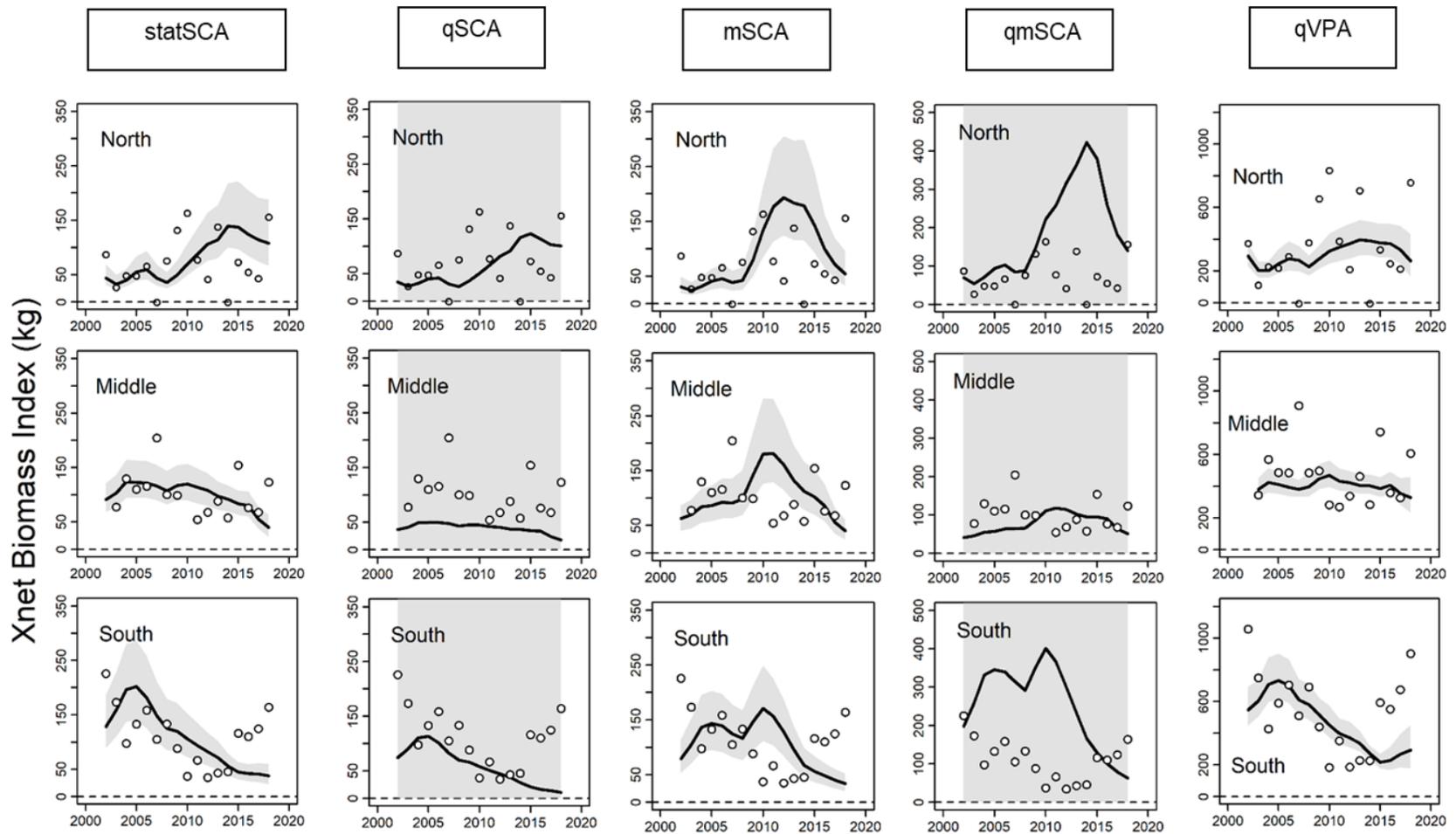
Figure 10. Observed (circles) and predicted (lines) age-aggregated experimental nets biomass fit to indices for 4TVn fall spawning Atlantic herring models (statSCA, qSCA, mSCA, qmSCA and qVPA). The lines show the maximum likelihood predicted indices of biomass (kg). The lines show the median predicted indices and the shading the 95 % confidence intervals of the predictions based on MCMC sampling
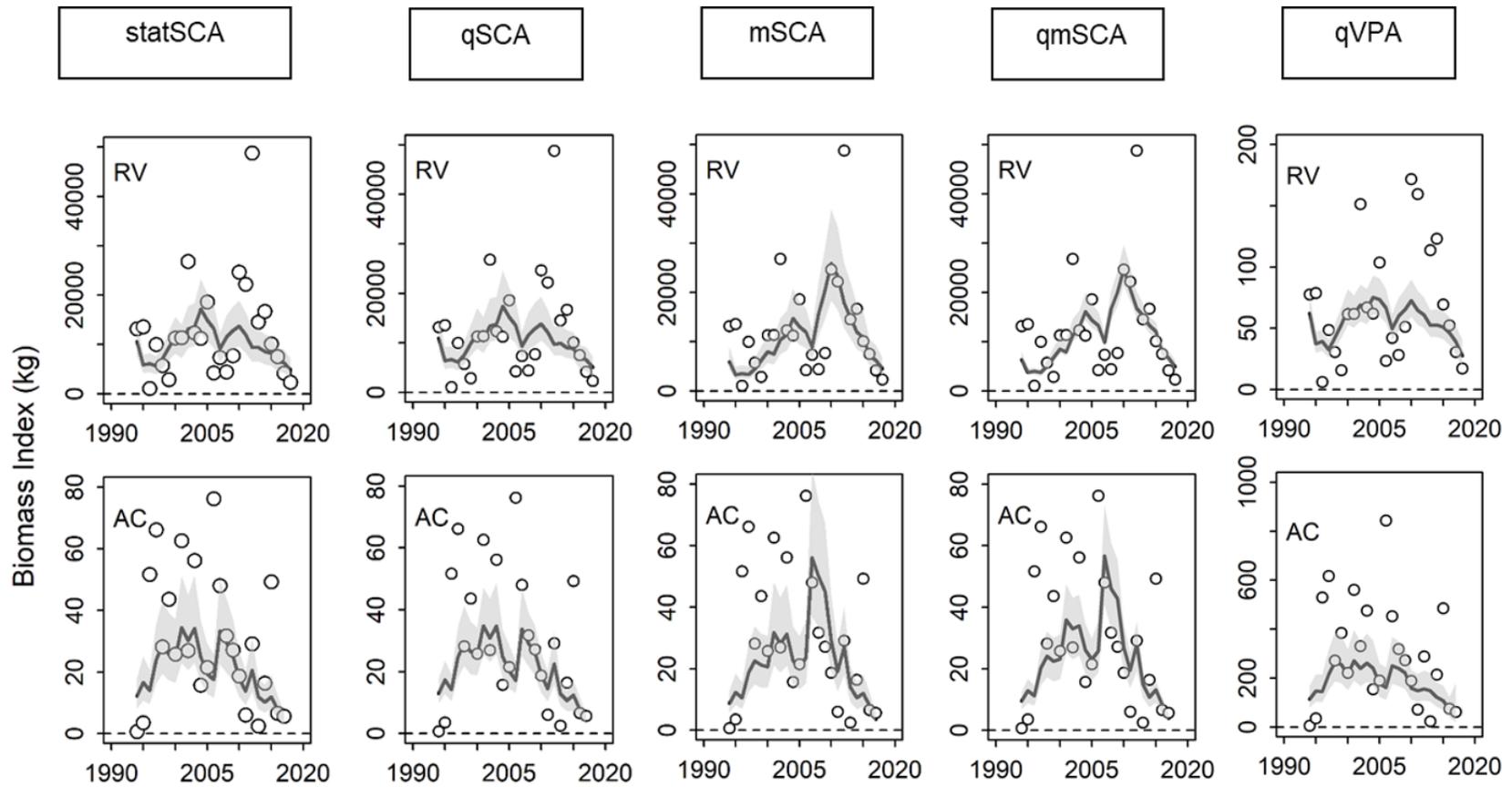
Figure 11. Observed (circles) and predicted (lines) age-aggregated RV and acoustic surveys fit to indices for 4TVn fall spawning Atlantic herring models (statSCA, qSCA, mSCA, qmSCA and qVPA). The lines show the maximum likelihood predicted indices of biomass (kg). The lines show the median predicted indices and the shading the 95 % confidence intervals of the predictions based on MCMC sampling
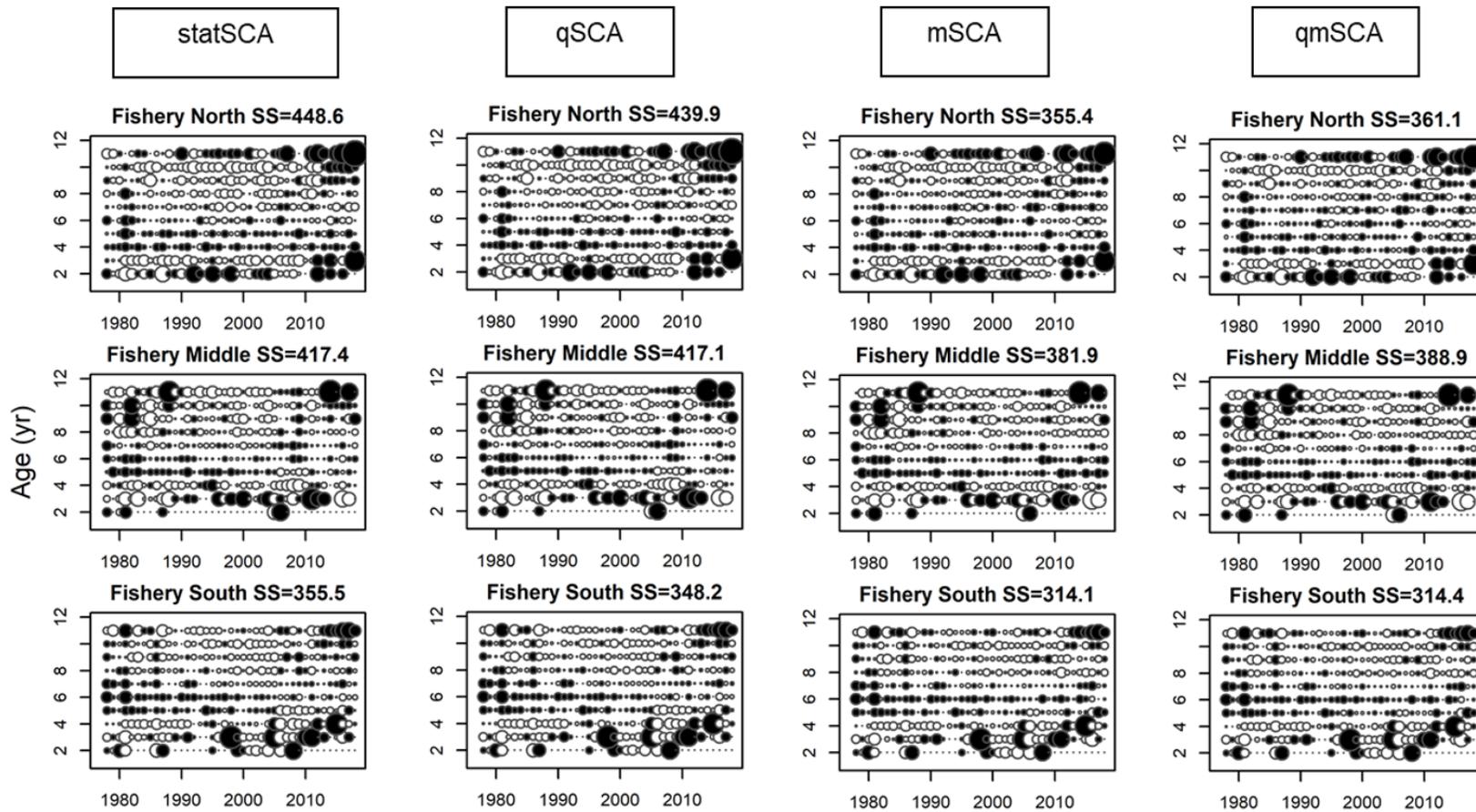
Figure 12. Residuals in fishery catch proportions at age for SCA population models (statSCA, qSCA, mSCA, qmSCA) of fall spawning Atlantic herring in the North, Middle and South regions. Rows are for ages and columns are for years. Circle radius is proportional to the absolute value of residuals. Black circles indicate negative residuals (i.e., observed < predicted) and white circles indicate positive residuals.
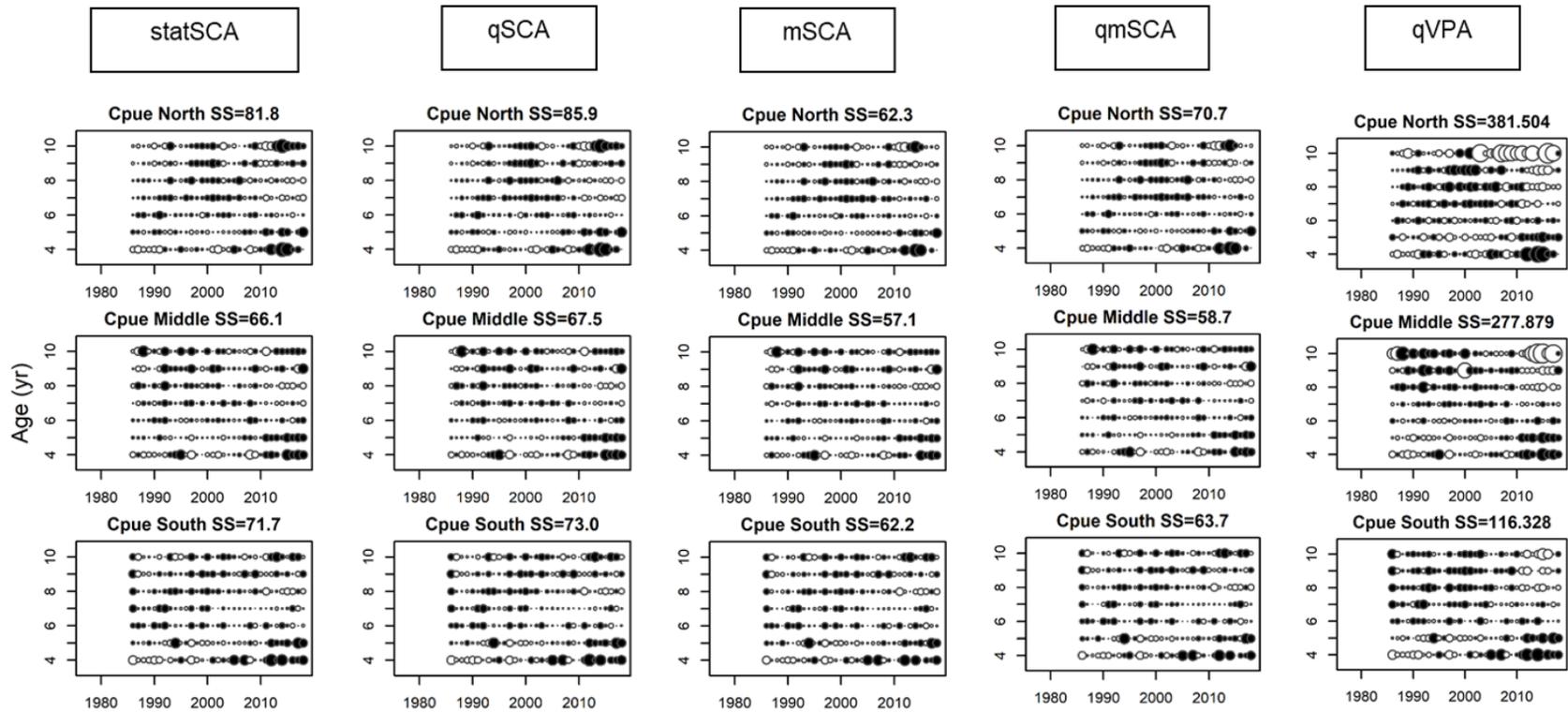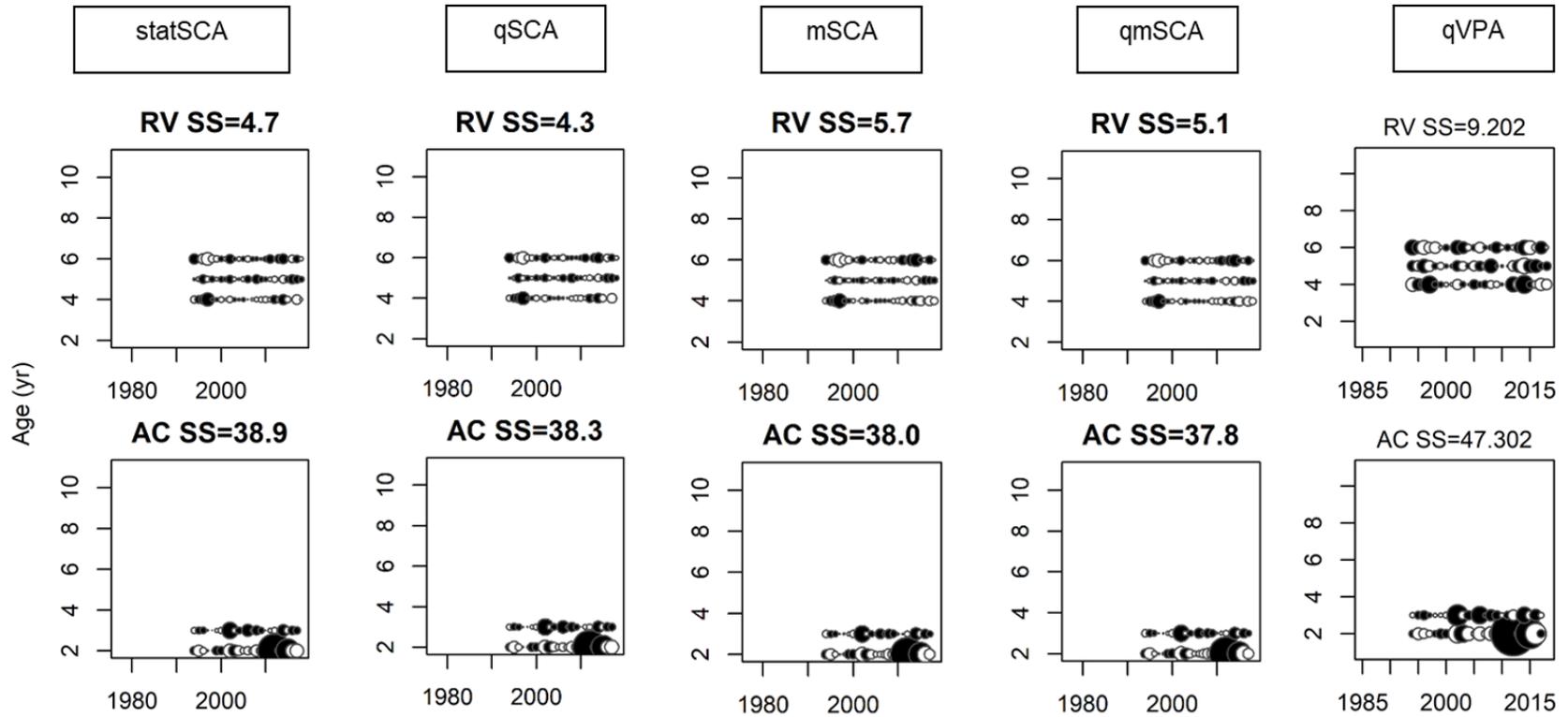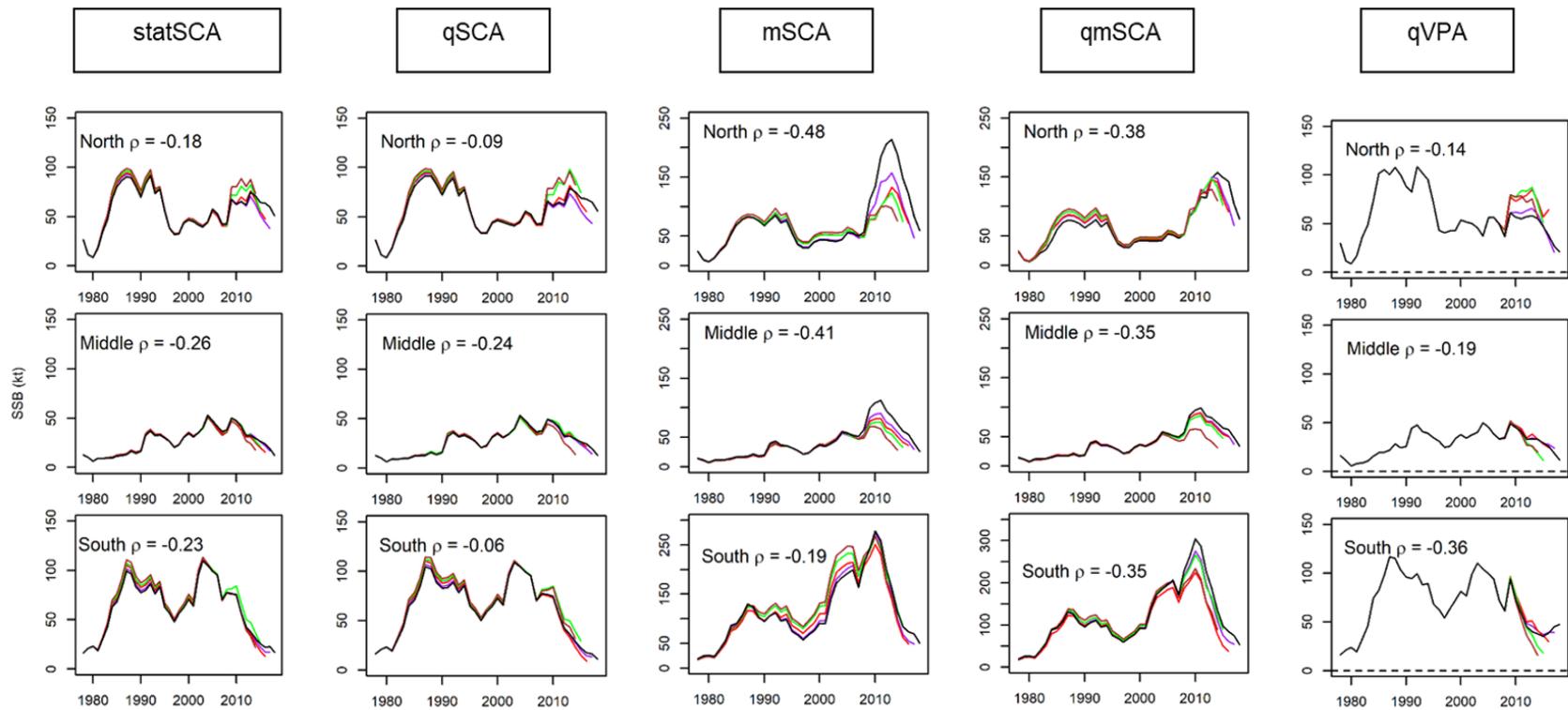
Figure 13. Residuals in catch per unit of effort (CPUE) in the gillnet fishery proportions at age for population models (statSCA, qSCA, mSCA, qmSCA and qVPA) of fall spawn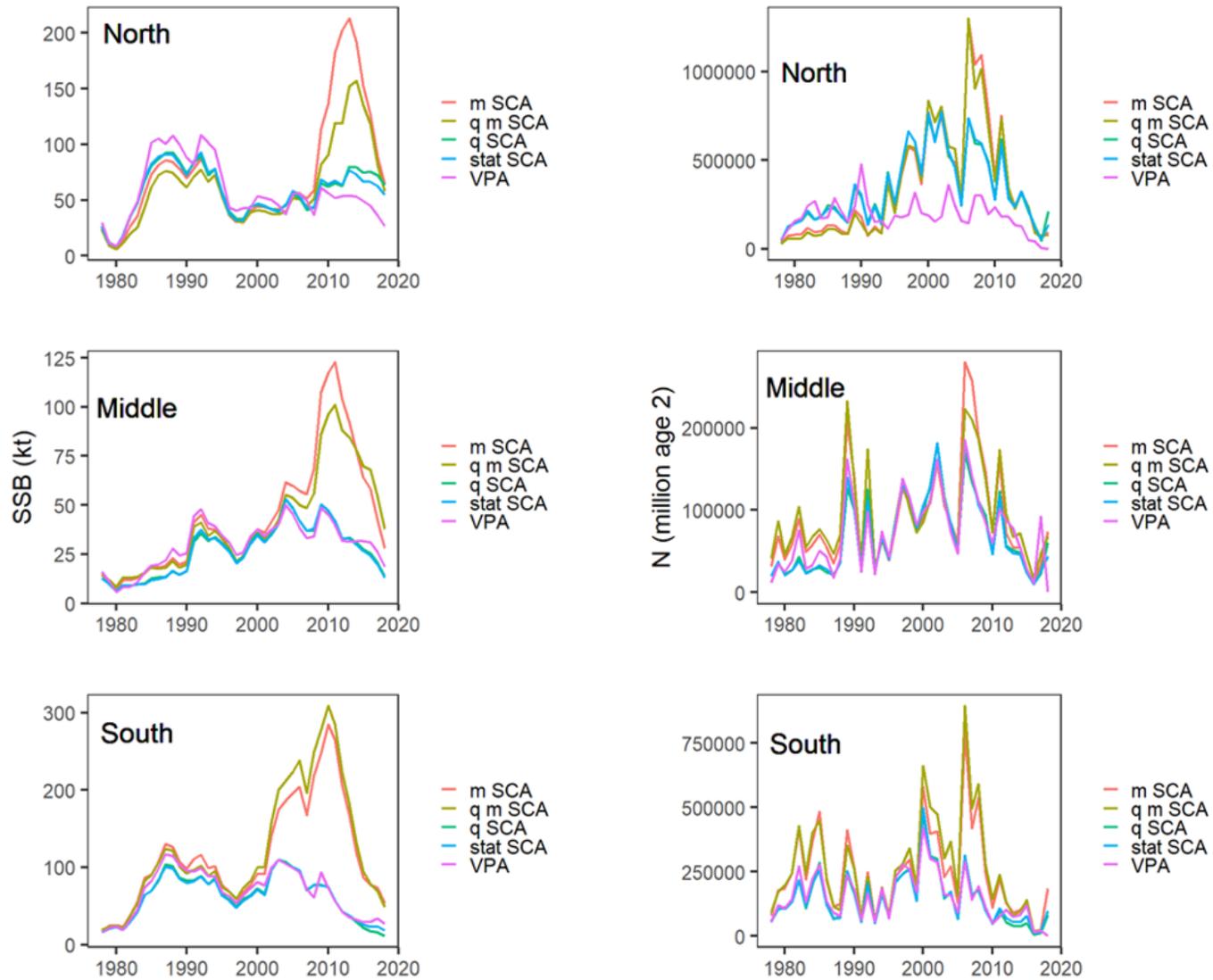ing Atlantic Herring in the North, Middle and South regions. Rows are for ages and columns are for years. Circle radius is proportional to the absolute value of residuals. Black circles indicate negative residuals (i.e., observed < predicted) and white circles indicate positive residuals.
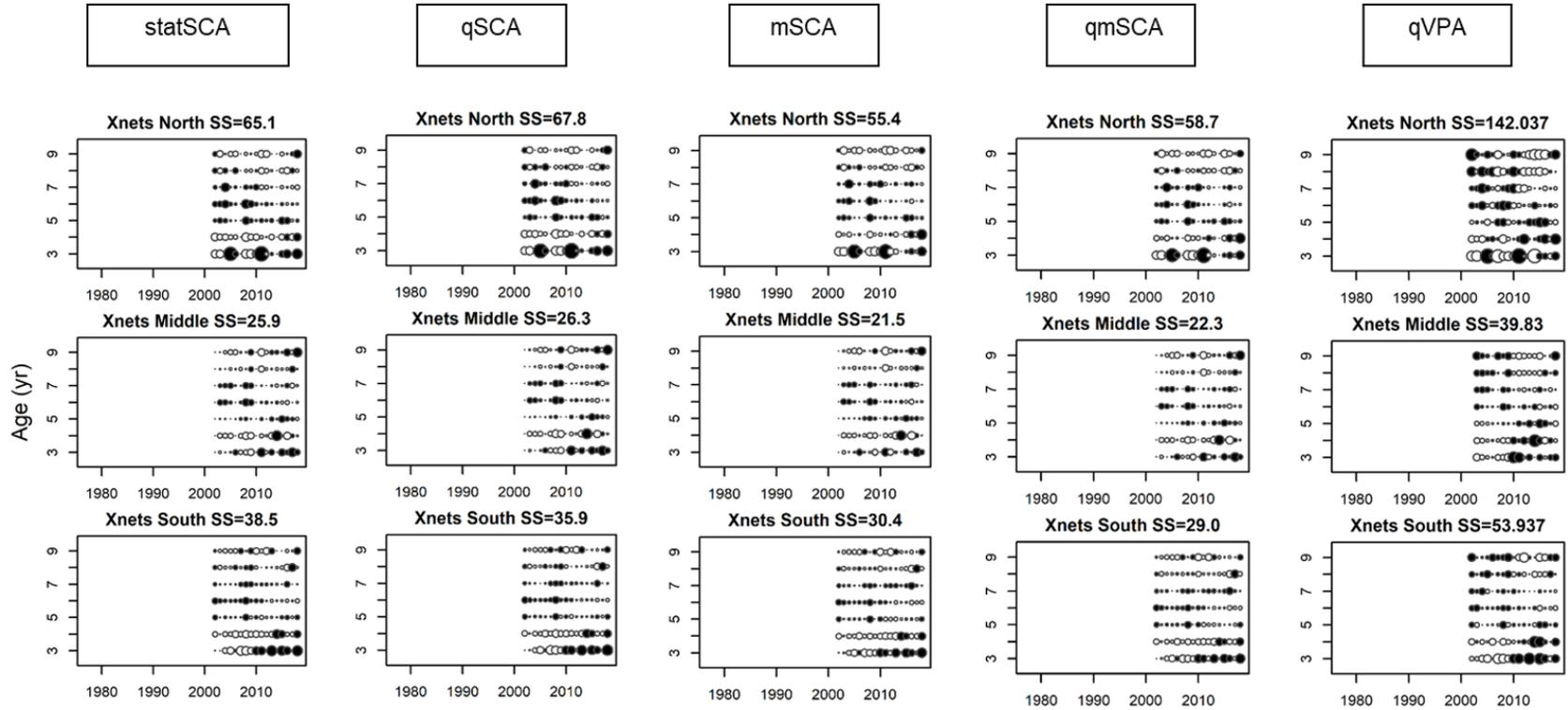
Figure 14. Residuals in the experimental gillnet index catch proportions at age for population models (statSCA, qSCA, mSCA, qmSCA and qVPA) of fall spawning Atlantic Herring in the North, Middle and South regions. Rows are for ages and columns are for years. Circle radius is proportional to the absolute value of residuals. Black circles indicate negative residuals (i.e., observed < predicted) and white circles indicate positive residuals.

Figure 15. Residuals in the RV and acoustic surveys proportions at age for population models (statSCA, qSCA, mSCA, qmSCA and qVPA) of fall spawning Atlantic Herring in the North, Middle and South regions. Rows are for ages and columns are for years. Circle radius is proportional to the absolute value of residuals. Black circles indicate negative residuals (i.e., observed < predicted) and white circles indicate positive residuals

Figure 16. Retrospective patterns and Mohn's rho in estimated spawning stock biomass (SSB, kt) of ages 4 to 10 for five population models (statSCA, qSCA, mSCA, qmSCA and qVPA) of fall spawning Atlantic Herring for the North, Middle and South regions.

Figure 17. Median MCMC estimates of spawning stock biomass (SSB kt, left) and recruitment (Number of age 2 fish, millions, right) for five population models (statSCA, qSCA, mSCA, qmSCA and qVPA) of fall spawning Atlantic herring for the North, Middle and South regions.
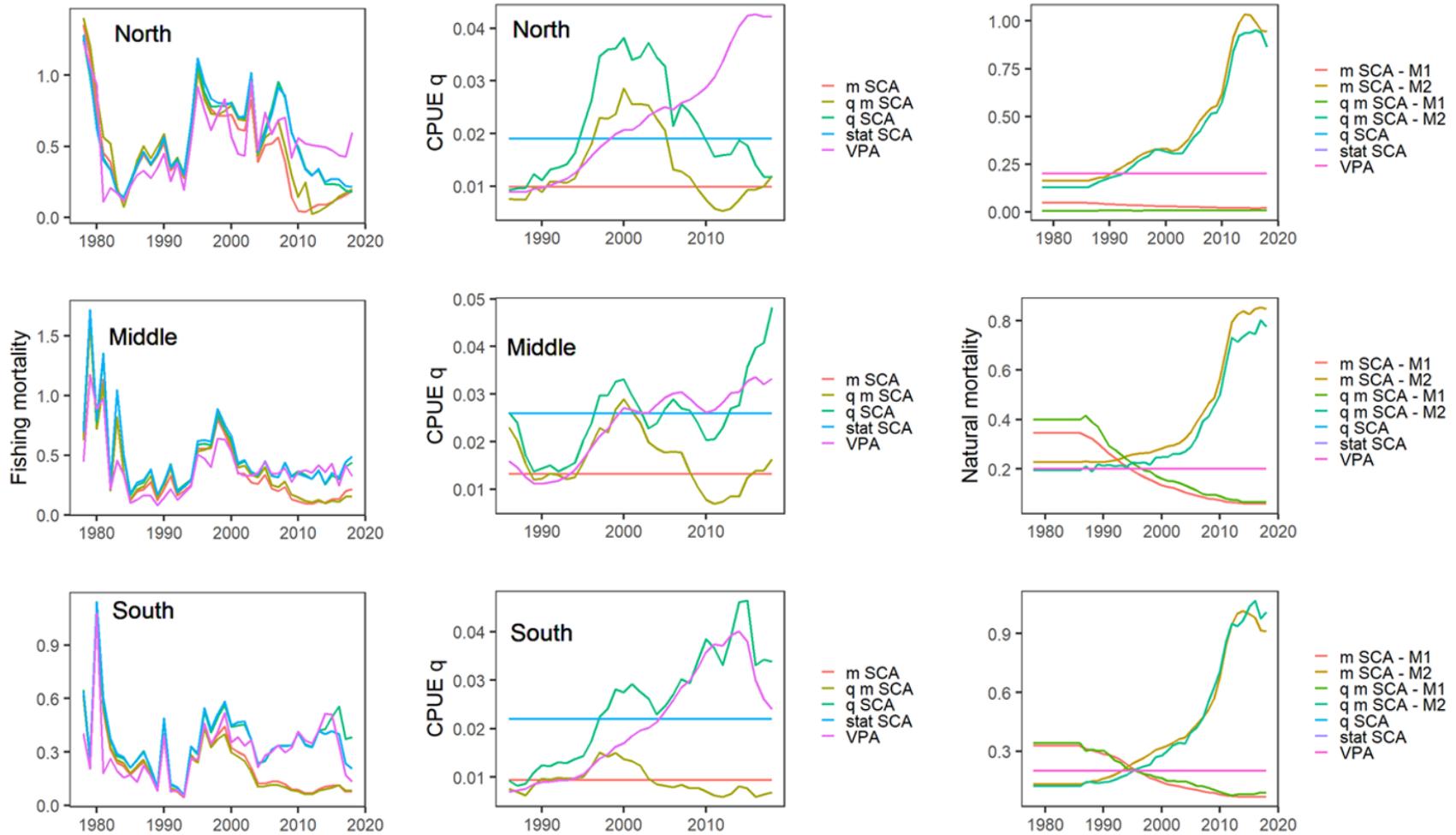
Figure 18. Median MCMC estimates of fishing mortality (left), fully-recruited catchability to the CPUE in the gillnet fishery (CPUE q, center) and Natural mortality (right, M1 = ages 2-6, M2 = ages 7 – 11+) for five population models (statSCA, qSCA, mSCA, qmSCA and qVPA) of fall spawning Atlantic herring in the North, Middle and South regions.
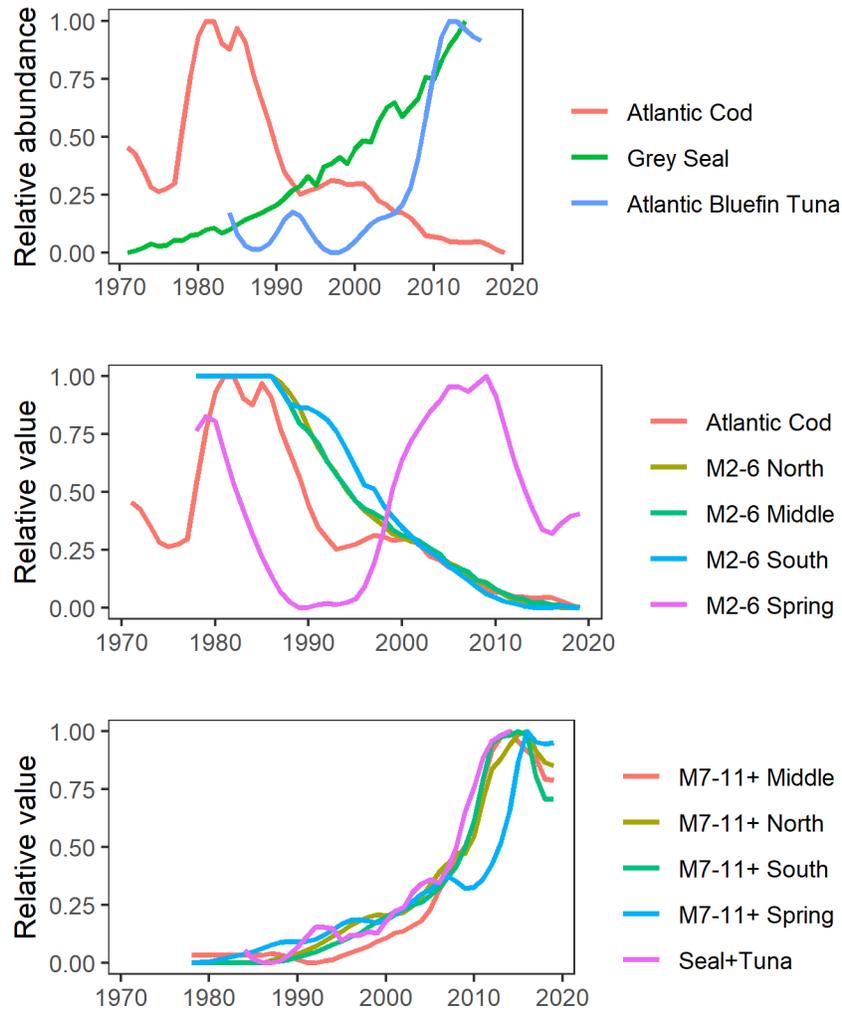
Figure 19. Scaled (0-1) relative abundance indices for herring major predators (Atlantic cod, grey seal, Atlantic bluefin tuna) between 1970-2019 (upper panel). Scaled relative value of Atlantic cod sGSL abundance and natural mortality estimates for age group 2-6 in qmSCA spring and fall herring stock models (middle panel). Scaled relative value of the summed sGSL indices of abundance for grey seals and Atlantic bluefin tuna, and natural mortality estimates for age group 7-11+ in qmSCA spring and fall herring stock models (lower panel). Natural mortality estimates are median MCMC estimates.
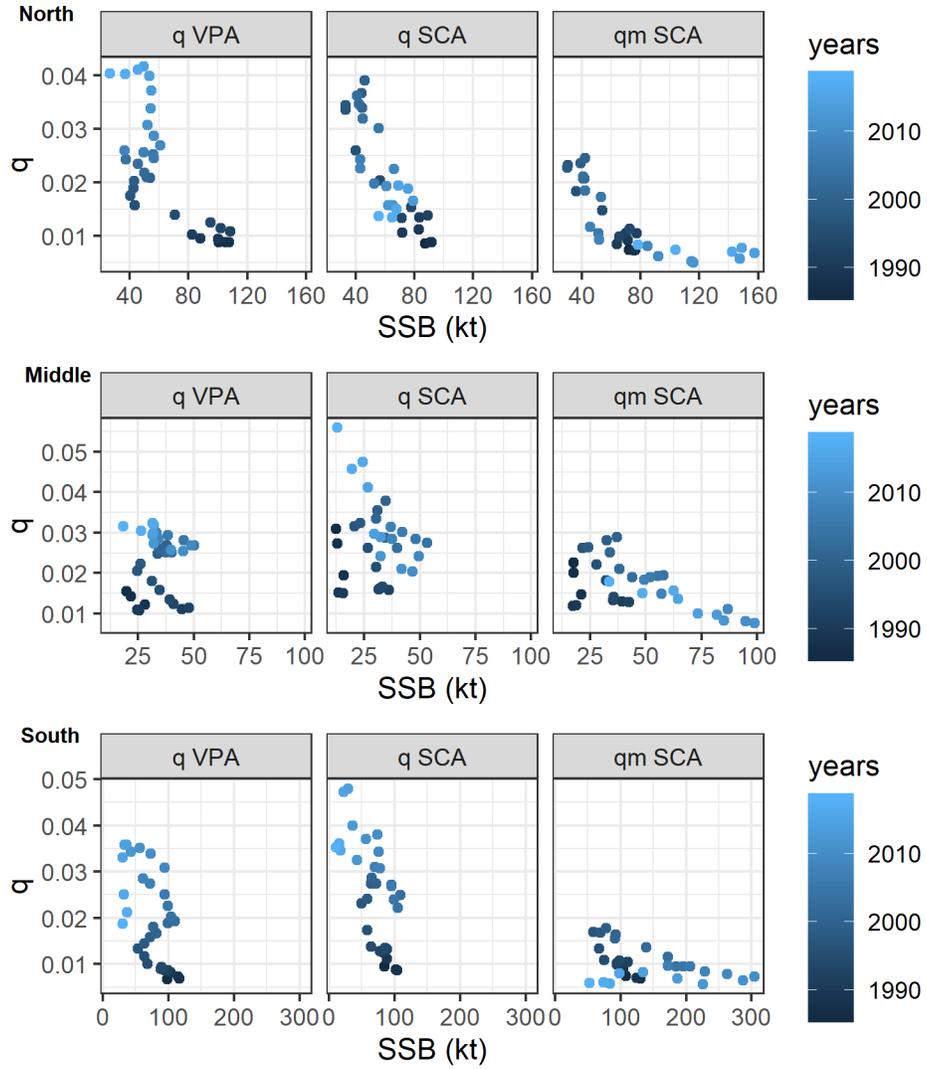
Figure 20. Fully-recruited catchability to the CPUE gillnet fishery (q) in function of SSB (tons) for fall spawning herring population models estimating time-varying q (qVPA, qSCA, qmSCA) in the North, Middle and South regions.

APPENDIX 1. RETROSPECTIVE ANALYSES IN EARLIER ASSESSMENTS

This appendix shows the retrospective patterns in the estimates of SSB in "accepted" models of spring and fall herring in the 2014 to 2018 assessment models.
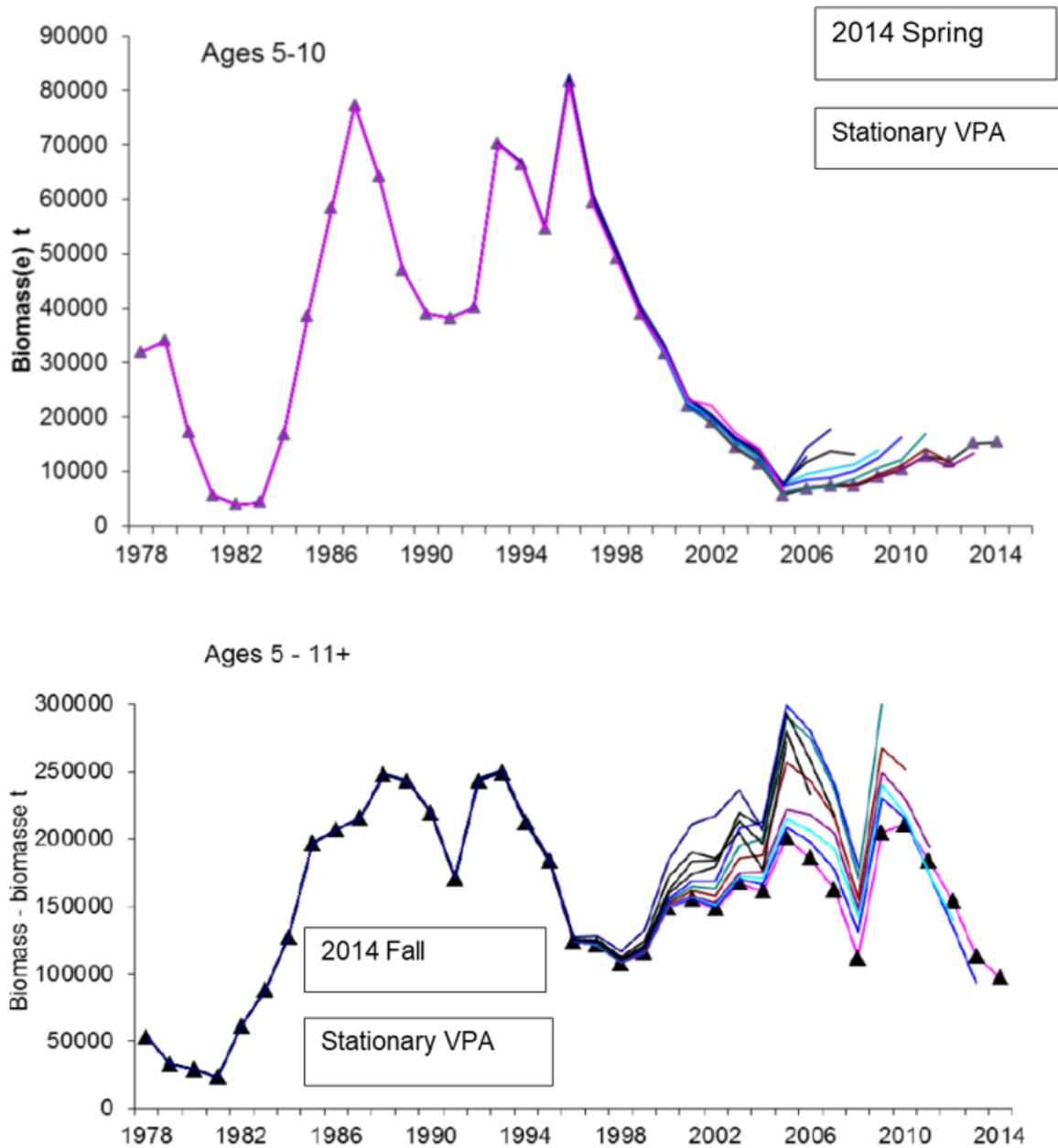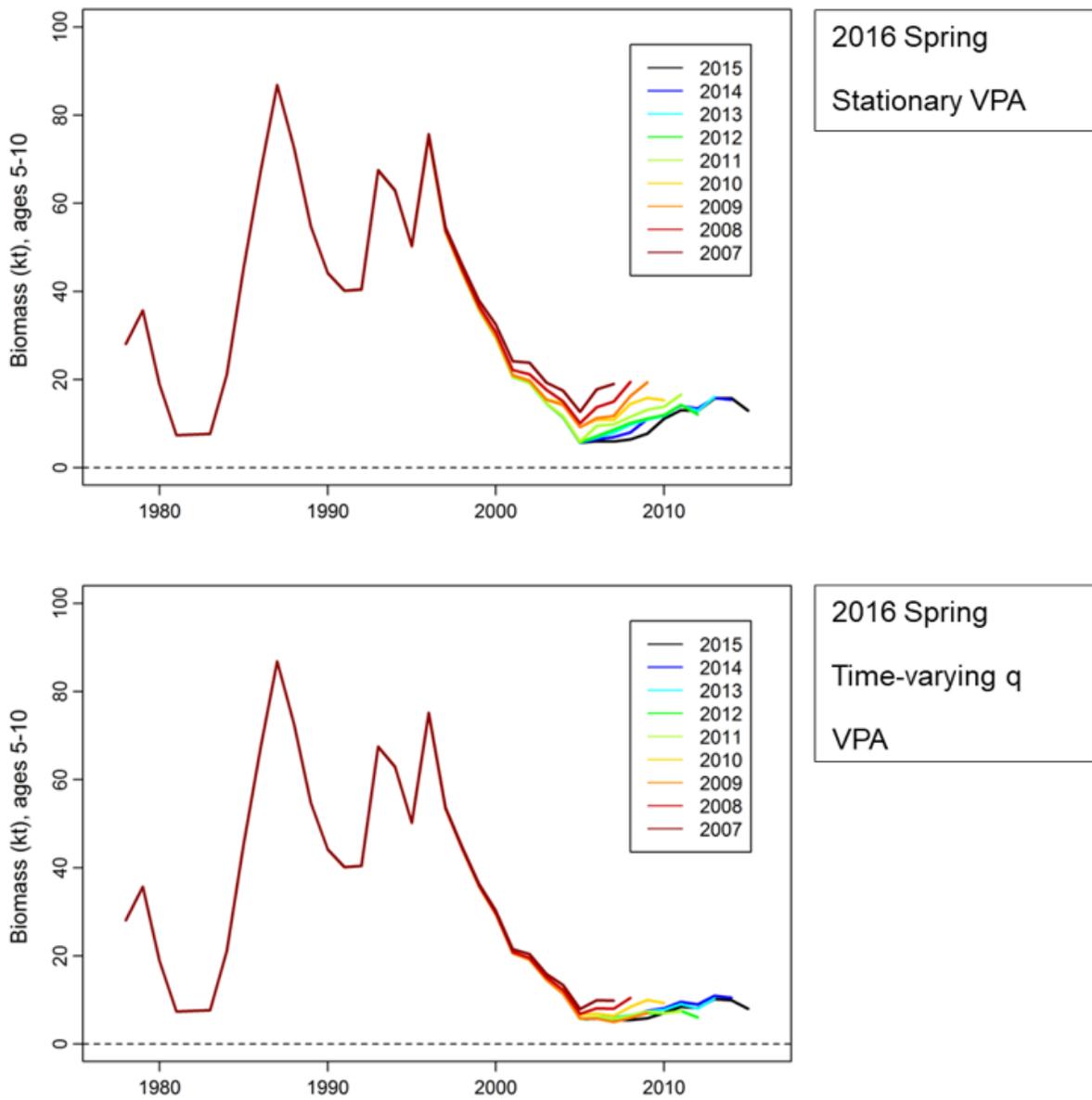

Figure A1.1. Retrospective patterns in the estimates of SSB in the assessment of the NAFO Division 4T spring spawning Herring (upper panel) and fall spawning Herring (lower panel) in 2014.
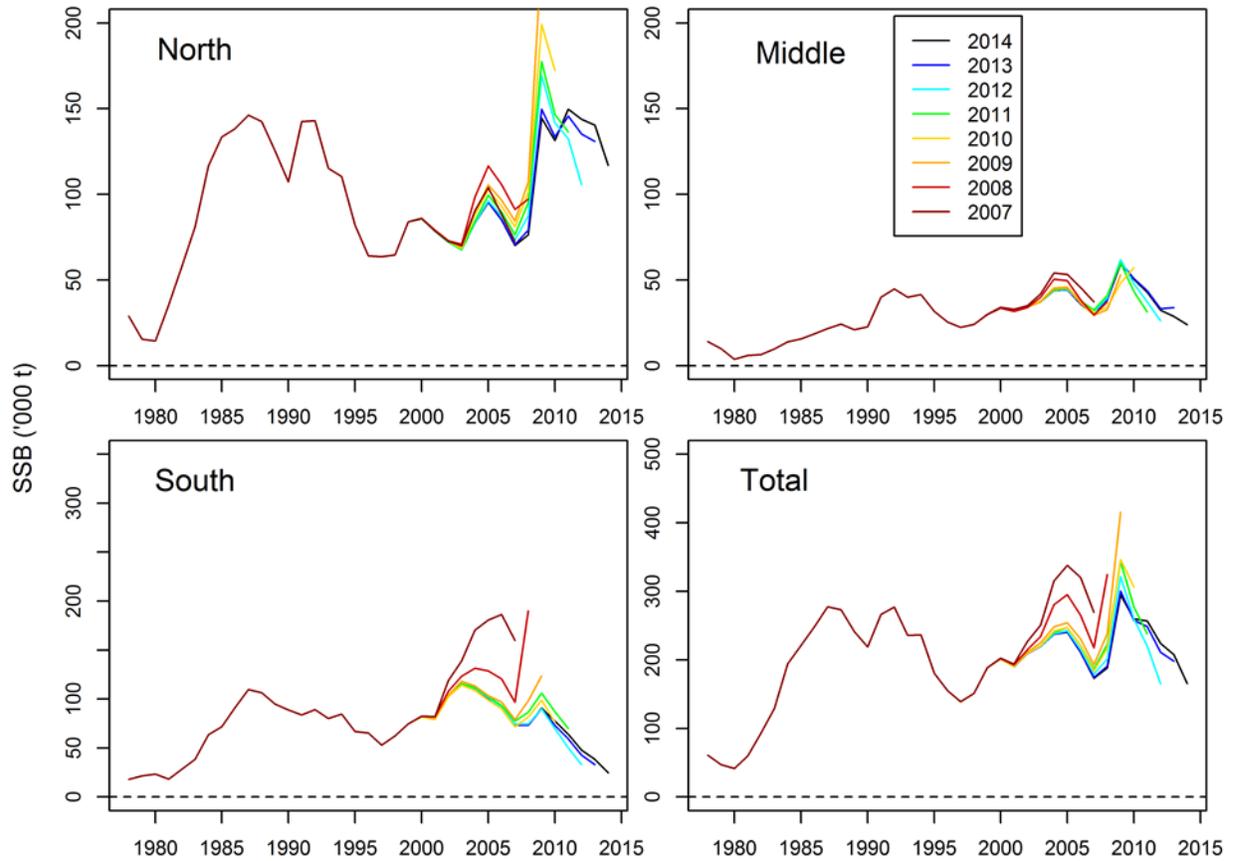
Figure A1.2. Retrospective patterns in the estimates of SSB in the assessment of the NAFO Division 4T fall spawning Herring in 2014.
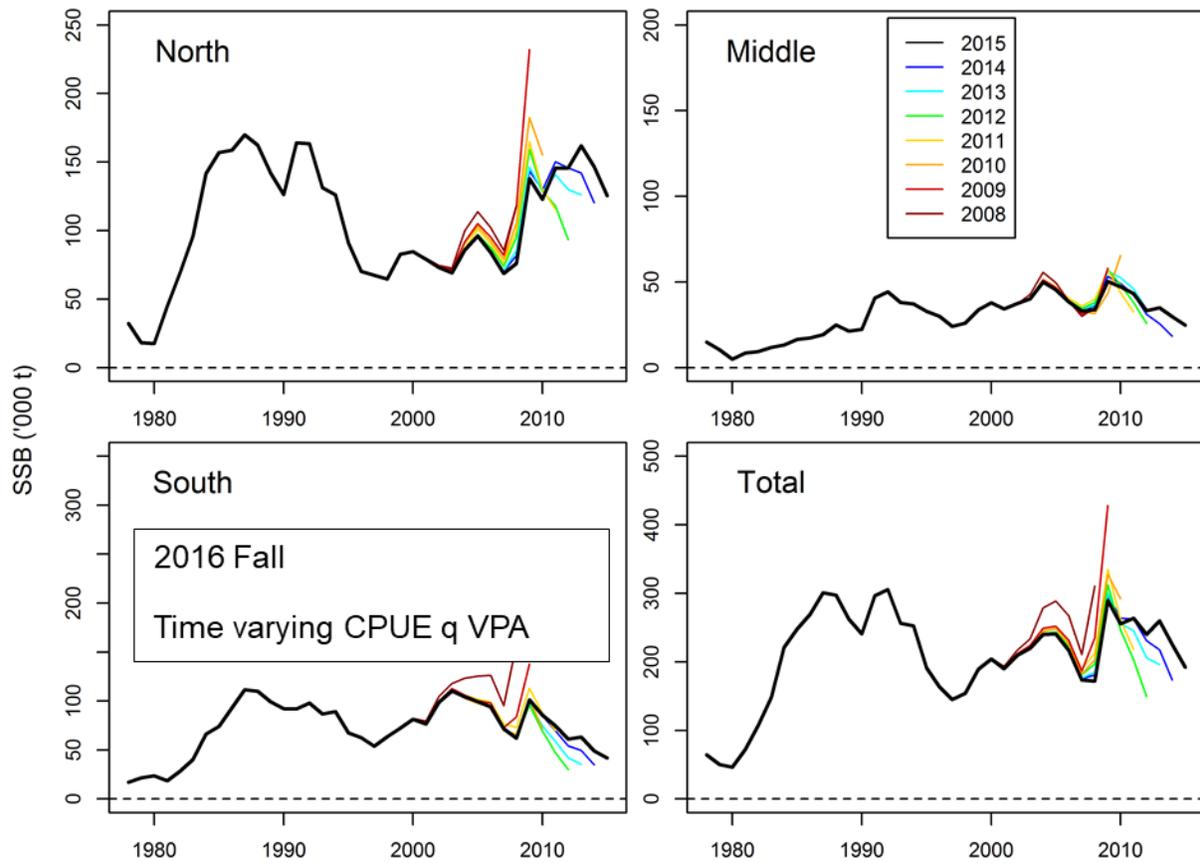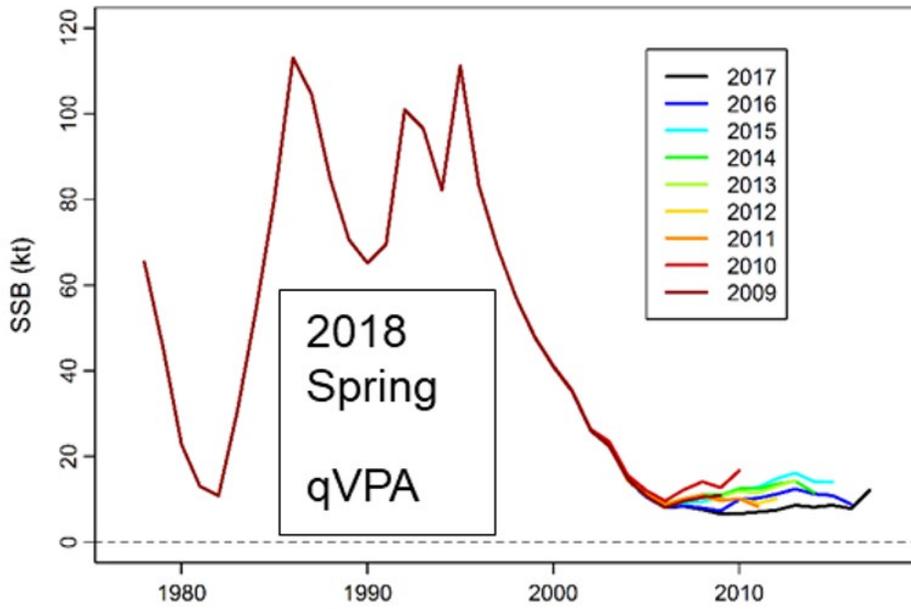
Figure A1.3. Retrospective patterns in the estimates of SSB in the assessment of the NAFO Division 4T fall spawning Herring using a time-varying q VPA in 2016.

Figure A1.4. Retrospective patterns in the estimates of SSB in the assessment of the NAFO Division 4T spring spawning Herring using a stationary VPA (upper panel) and a time-varying q VPA (lower panel) in 2016.

Figure A1.5. Retrospective patterns in the estimates of SSB in the assessment of the NAFO Division 4T spring spawning Herring using a time-varying q VPA in 2018.
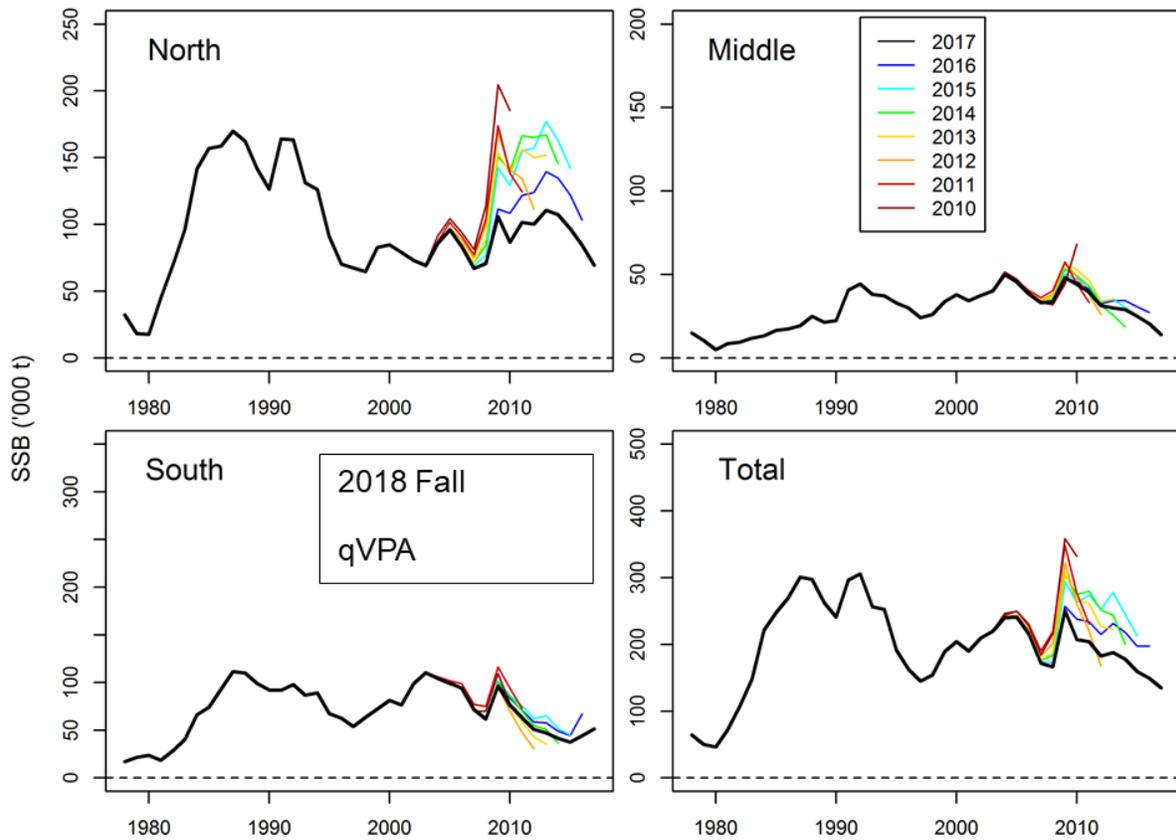

Figure A1.6. Retrospective patterns in the estimates of SSB in the assessment of the NAFO Division 4T fall spawning Herring using a time-varying q VPA in 2018.

APPENDIX 2. CHOICE OF SELECTIVITY FUNCTIONS FOR SPRING HERRING

Background information indicates that the selectivity curves for the spring herring fishery and the gillnet CPUE index should not be sharply domed, and if they are domed the dome should become less severe over time (Appendix 2.1). Spring herring size at age has steadily declined over time. Estimates of relative selectivity over time and age (based on multi-mesh experimental nets) indicate that selectivity at age was domed shaped with a very shallow descending limb in the 1980s. The selectivity curve then lost its dome, initially becoming flat-topped and more recently steadily increasing. These changes involve decreasing selectivity for young fish and increasing selectivity for older fish, consistent with the declining size-at-age. In recent years selectivity of older fish is also decreasing, but this occurs before a dome is reached (i.e., selectivity steadily increases with age, with no dome or flat-top).

Consistent with this, the partial recruitment curve to the fishery in recent VPAs is flat-topped. In earlier VPAs the calibration coefficients (q-at-age) for the CPUE index was dome-shaped, but the dome was only slight. Values of q-at-age to the acoustic survey is roughly flat. (See Appendix 2.1 for details). Using the recent VPA model, but estimating selectivity at age independently for each age, again yields a flat-topped fishery partial recruitment curve, flat-topped selectivity for the CPUE index in 1990-2004 and a "slight" dome in 2005-2017 (age 10 selectivity = age 8 selectivity) (Appendix 2.2). Using all selectivity functions (except logistic), the recent SCA model estimates dome-shaped selectivity for the fishery and CPUE index (Appendix 2.2). Except for exponential-logistic selectivity curves, all these domes are sharp, and they get sharper in the most recent period when selectivity at ages 10 or 11+ can be as low as that at ages 2 to 4 years old. This result is contrary to expectation given the results discussed above. This may reflect model misspecification (i.e., constant M of 0.2). This may force the model to deal with increasing M by estimating very sharp selectivity domes (i.e. the old fish are there but can't be caught).

This hypothesis is supported by the selectivity curve for the acoustic survey. Size composition in the acoustic survey is estimated by mid-water trawling of herring schools. Previous VPAs estimated a roughly flat q-at-age to the acoustic survey, or fluctuation without a trend between levels of 0.6 and 1.0. In contrast, SCA models with selectivity functions that allow a dome estimate sharply declining selectivity-at-age, from 1 at age 4 to as low as 0.2 at age 8 (only ages 4 to 8 are used for this index). When acoustic survey selectivity is estimated in 2 time blocks, the estimated decline in selectivity with age is much steeper in the recent period. This is again consistent with model misspecification (M assumed to be constant when it is instead increasing). See Appendices 2.1 and 2.2 for details.

There is almost no difference in the PAA residual patterns between models (Appendix 2.3).

Conclusion: Given the ancillary information, the sharp decline in selectivity of old herring estimated by most selectivity functions is not plausible. This is partly addressed by allowing M to vary over time. The estimated changes in M are plausible given the observed changes in the abundances of predators of herring.

APPENDIX 2.1. BACKGROUND INFORMATION

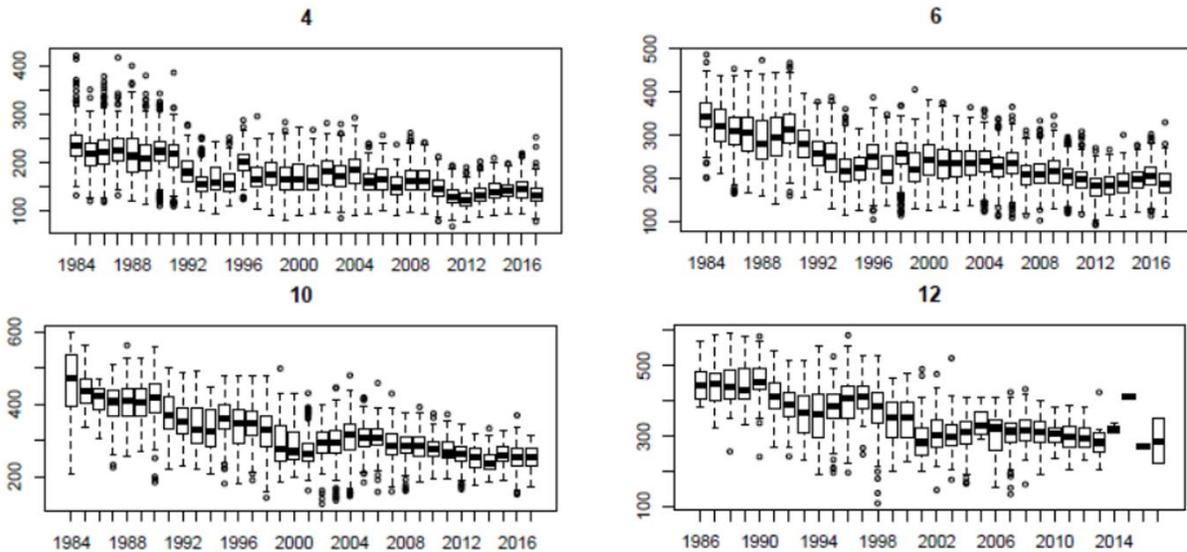A. Changes in mean weight at age of spring herring

Figure A2.1.1. Time series of weight at age (grams) boxplots of NAFO Division 4T spring spawning Herring for ages 4, 6, 10 and 12.

Size-at-age has been declining over time. If the selectivity curve is dome-shaped, selectivity should be decreasing for young herring and increasing for old herring.

B. Changes in relative selectivity by age and year based on an analysis of catches in the multi-mesh size experimental nets. Results are for fall herring (which show similar declines in size-at-age).
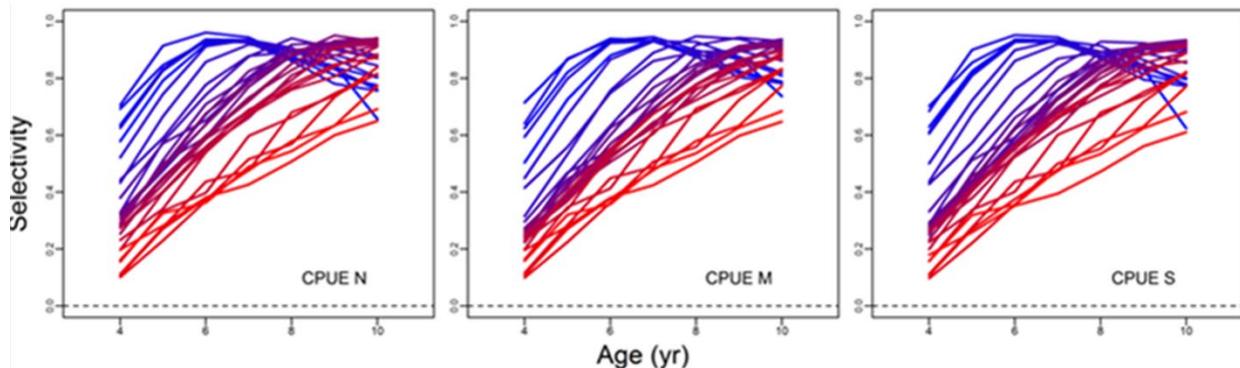
Figure A2.1.2. Estimated relative selectivity (y axis) in the multimesh experimental nets by age (x axis) and year (colored lines) for NAFO Division 4T fall spawning Herring in the North, Middle and South subpopulations over time (continuous scale from blue in 1978 to red in 2015).

These curves of relative selectivity at age were estimated using multimesh experimental nets. This suggests a slightly domed curve early in the time series (when size-at-age was high transitioning to a logistic curve later in the time series (when size-at-age was small). Selectivity decreased for young fish and increased for old fish as size-at-age decreased.

C. Partial recruitment vector (F_a/max(F_a)) from a VPA model for spring spawners.

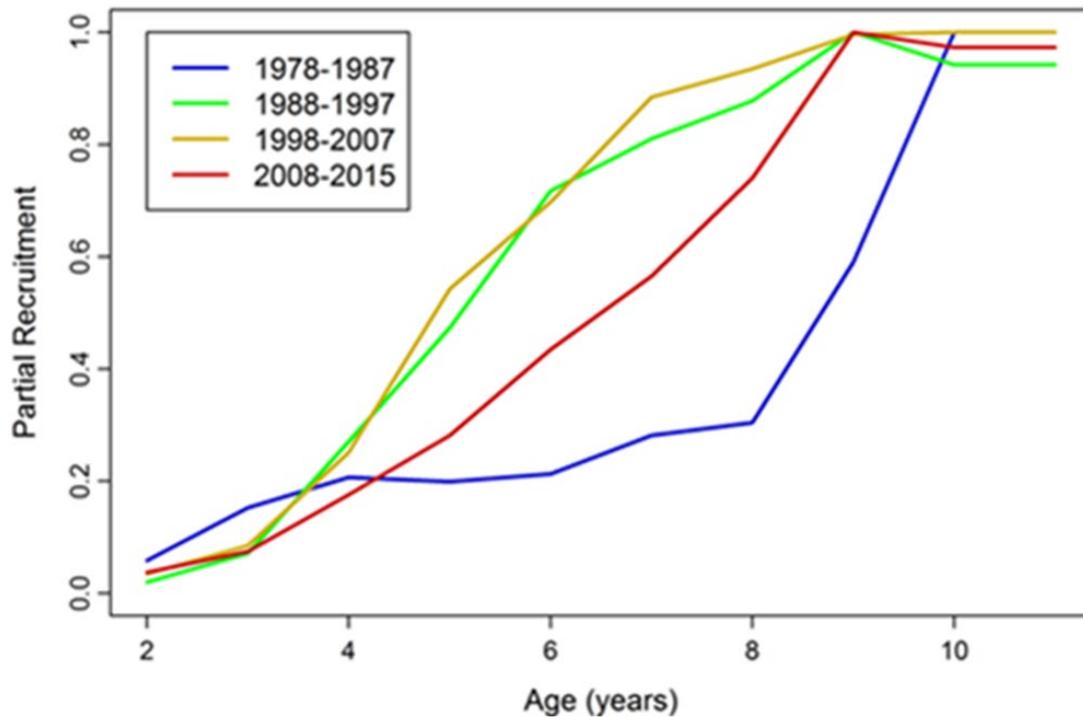*Figure A2.1.3. Partial recruitment vector (Fa/max(Fa)) for ages 2 to 11+ from a VPA model for NAFO Division 4T spring spawning Herring for blocks of years. Time periods are identified by colored lines.*

This suggests a "flat-topped" or logistic selectivity at age function.

D. Calibration coefficients (q-at-age) from previous VPAs for spring spawners.

Table A2.1.1. Calibration coefficients (q-at-age) in the CPUE and Acoustic indices from previous VPAs for NAFO Division 4T spring spawning Herring.

Index	age	2007	2008	2009	2011	2013
CPUE	4	0.0011	0.0015	0.000328	0.000197	0.000152
	5	0.0028	0.0037	0.000994	0.000584	0.000554
	6	0.0037	0.0047	0.001406	0.000828	0.000865
	7	0.004	0.0053	0.001682	0.001023	0.001118
	8	0.0041	0.0055	0.001648	0.001031	0.001128
	9	0.004	0.0051	0.001490	0.000953	0.001073
	10	0.0035	0.0043	0.001489	0.000953	0.001057
Acoustic	4	0.0035	0.0036	-	-	-
	5	0.0031	0.0037	-	-	-
	6	0.0029	0.0033	-	-	-
	7	0.0028	0.003	-	-	-
	8	0.0047	0.0044	-	-	-

APPENDIX 2.2. SELECTIVITY MODELS CONSIDERED

- q is estimated separately for each CPUE time block to account for increasing CPUE q.

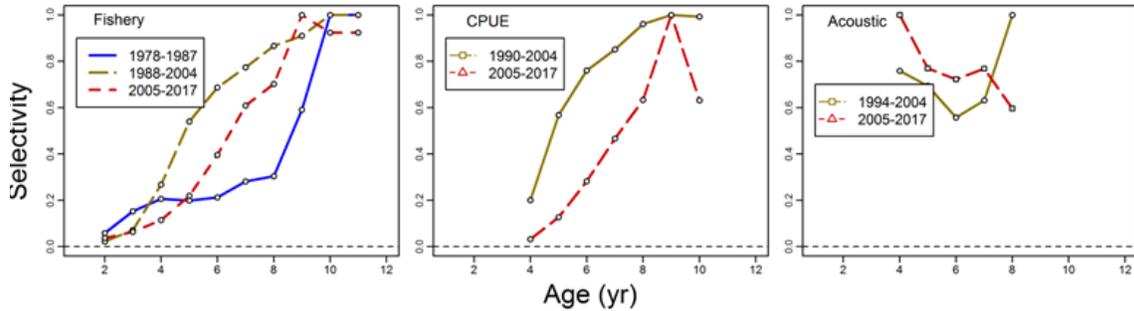Model 1: VPA, no selectivity function, selectivity estimated for each age

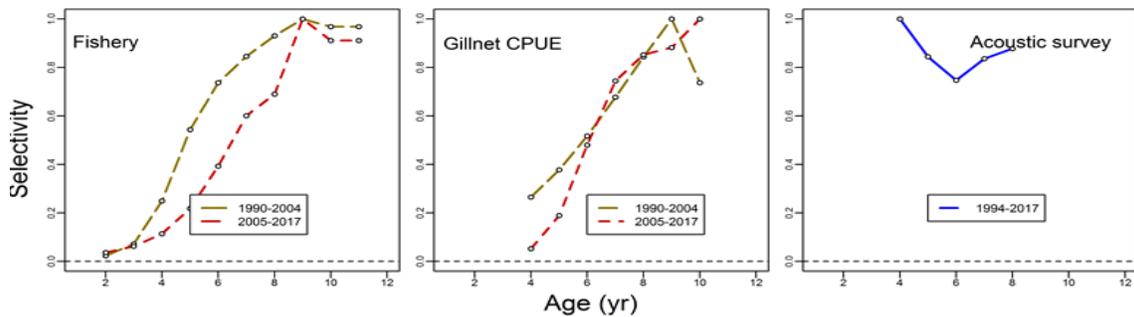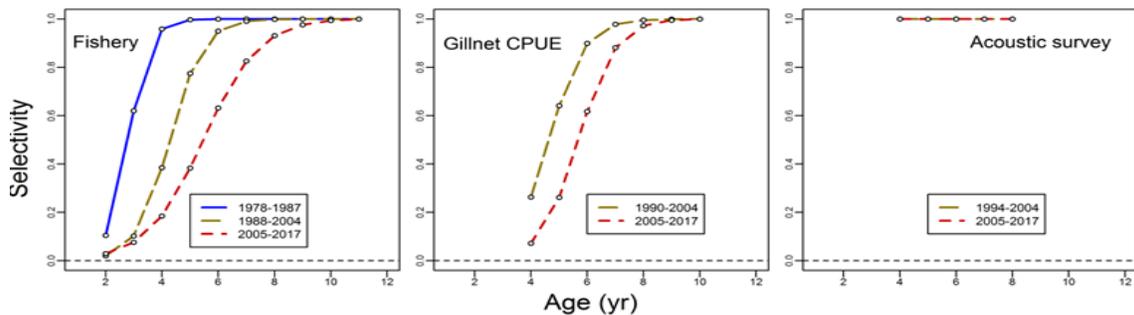*Figure A2.2.1. Model 1: VPA, no selectivity function, selectivity estimated for each age: Estimated selectivity (y axis) at age (x axis) for the fishery (left panel), CPUE (center panel) and acoustic survey (right panel) catches. Colored lines identify time periods.*

Model 2: VPA, Logistic Plus

Figure A2.2.2. Model 2: VPA, logistic plus selectivity function, selectivity estimated for each age: Estimated selectivity (y axis) at age (x axis) for the fishery (left panel), CPUE (center panel) and acoustic survey (right panel) catches. Colored lines identify time periods.

Model 3: SCA, Logistic selectivity

Figure A2.2.3. Model 3: SCA, logistic selectivity function, selectivity estimated for each age: Estimated selectivity (y axis) at age (x axis) for the fishery (left panel), CPUE (center panel) and acoustic survey (right panel) catches. Colored lines identify time periods.

Model 4: SCA, "Hybrid" Logistic selectivity

Models selectivity as a logistic function up to age 8 and freely estimates selectivity at ages 9-11+.
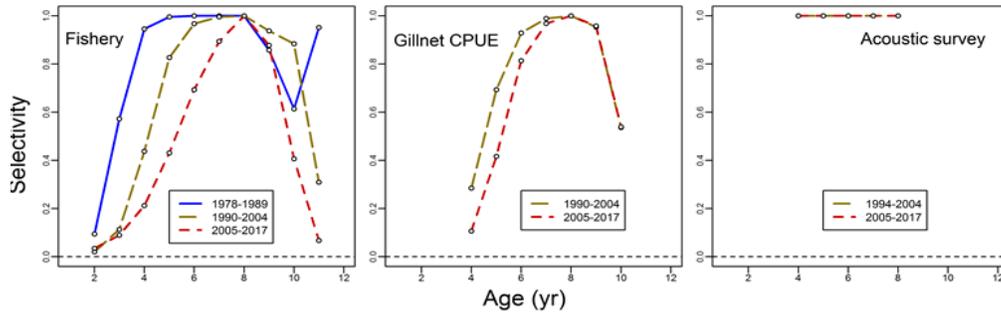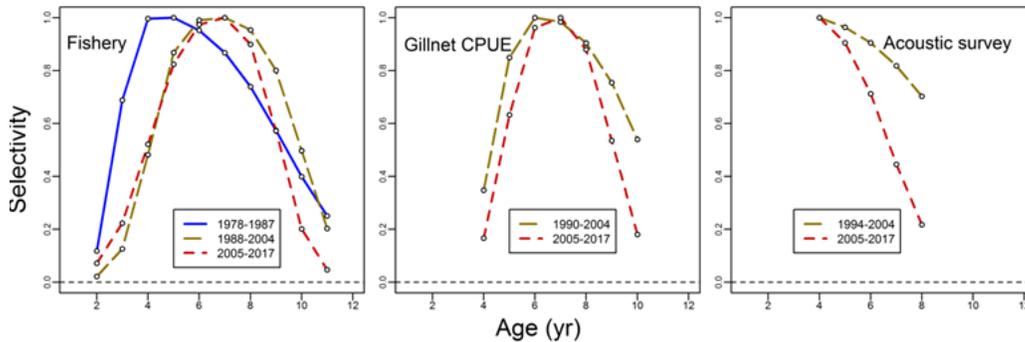

Figure A2.2.4. Model 4: SCA, hybrid logistic selectivity function, selectivity estimated for each age: Estimated selectivity (y axis) at age (x axis) for the fishery (left panel), CPUE (center panel) and acoustic survey (right panel) catches. Colored lines identify time periods.

Model 5: SCA, Double Logistic selectivity

Figure A2.2.5. Model 5: SCA, double logistic selectivity function, selectivity estimated for each age: Estimated selectivity (y axis) at age (x axis) for the fishery (left panel), CPUE (center panel) and acoustic survey (right panel) catches. Colored lines identify time periods.

Model 6: SCA, Gamma selectivity
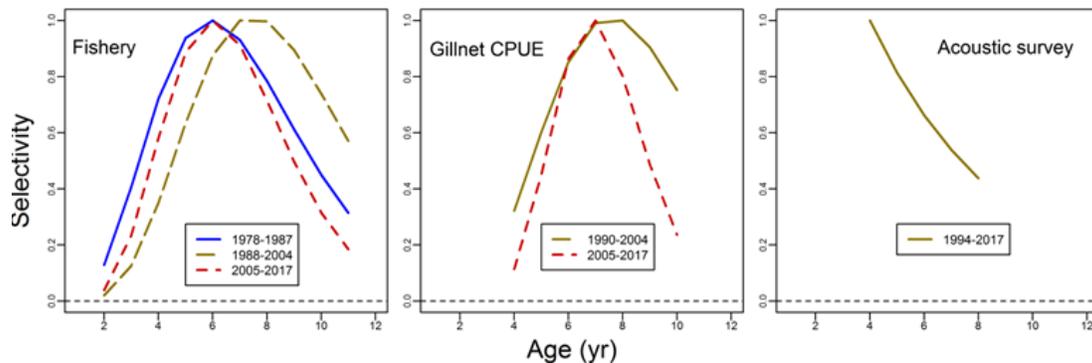
Maximum gradient component is 40.

Figure A2.2.6. Model 6: SCA, gamma logistic selectivity function, selectivity estimated for each age: Estimated selectivity (y axis) at age (x axis) for the fishery (left panel), CPUE (center panel) and acoustic survey (right panel) catches. Colored lines identify time periods.

Model 7: SCA, Exponential-logistic selectivity

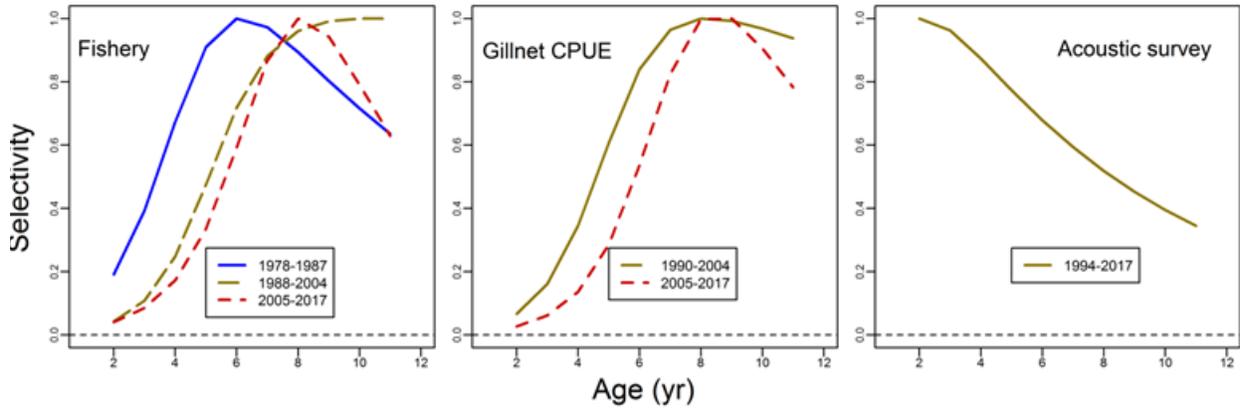Estimates for selectivity parameter p1 are at the upper bound.

Figure A2.2.7. Model 7: SCA, exponential logistic selectivity function, selectivity estimated for each age: Estimated selectivity (y axis) at age (x axis) for the fishery (left panel), CPUE (center panel) and acoustic survey (right panel) catches. Colored lines identify time periods.

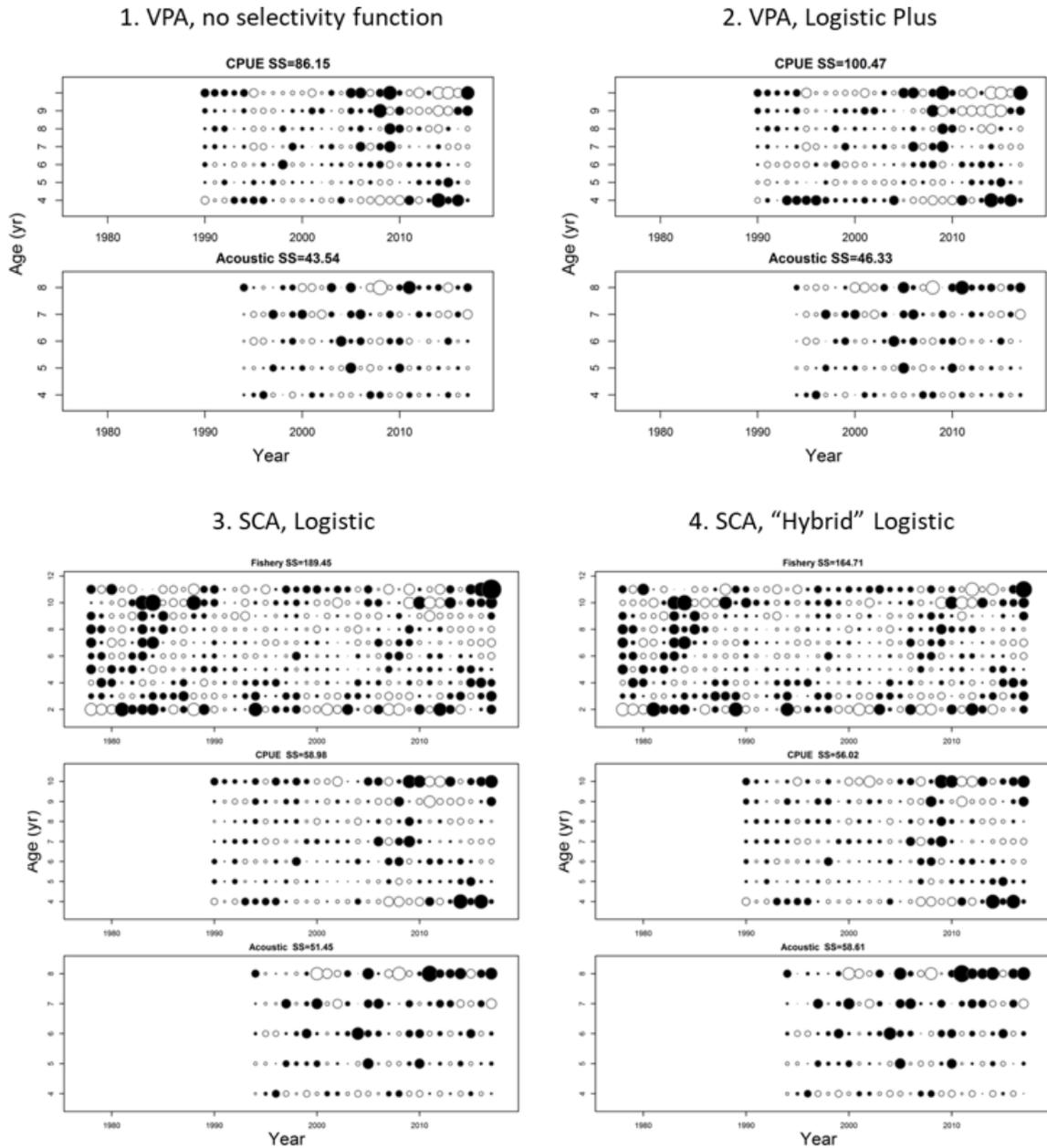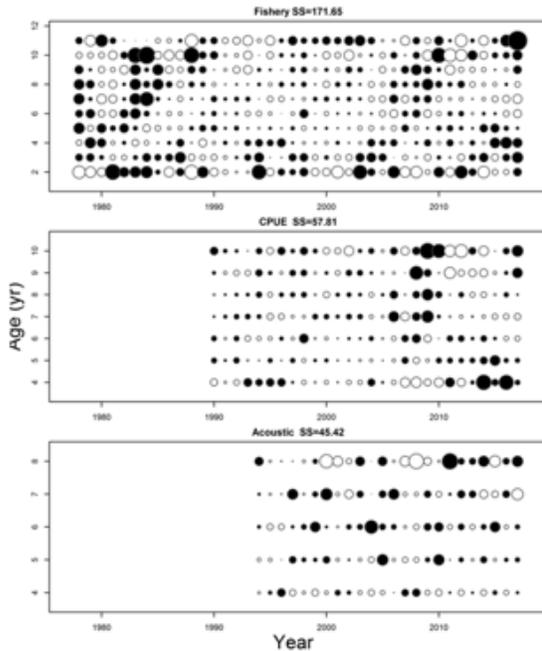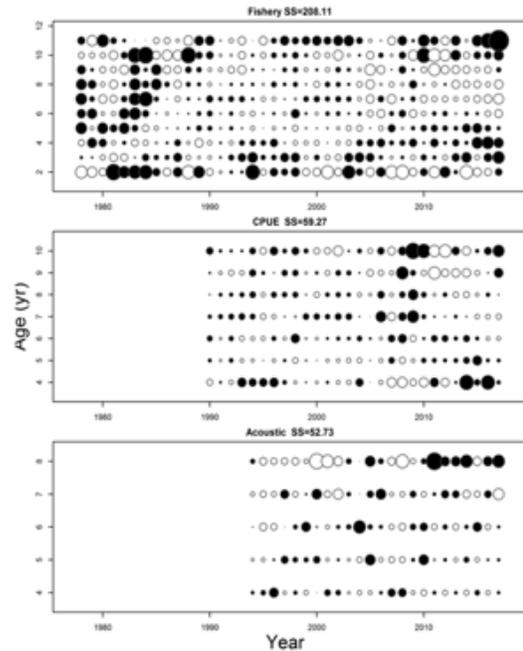# APPENDIX 2.3. MODEL FITS TO THE PROPORTIONS-AT-AGE

Figure A2.3.1. Model fits to the abundance at age for Model 1 VPA no selectivity function (top-left panel), Model 2 VPA logistic plus selectivity function (top-right panel), and fit to proportions at age for Model 3 SCA logistic selectivity function (lower left panel) and Model 4 SCA hybrid logistic selectivity function (lower right panel) in fishery and index catches. Rows are for ages and columns are for years. Circle radius is proportional to the absolute value of residuals. Black circles indicate negative residuals (i.e., observed < predicted) and white circles indicate positive residuals. SS value is sum of squared residuals.

5. SCA, double Logistic
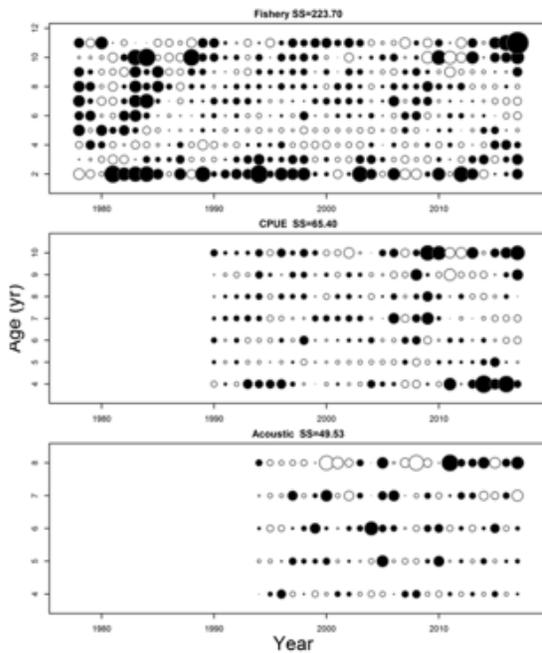
6. SCA, Gamma

7. SCA, Exponential-Logistic

Figure A2.3.2. Model fits to the proportions at age for Model 5 SCA double logistic selectivity function (top-left panel), Model 6 SCA gamma selectivity function (top-right panel) and Model 3 SCA exponential-logistic selectivity function (lower left panel) in fishery and index catches. Rows are for ages and columns are for years. Circle radius is proportional to the absolute value of residuals. Black circles indicate negative residuals (i.e., observed < predicted) and white circles indicate positive residuals. SS value is sum of squared residuals.