**Consequences of implementing a 2-dimensional survival smoother into a state-space assessment model for Southern New England-Mid Atlantic yellowtail flounder**

**HAIKUN XU1\*, TIMOTHY J. MILLER2, JAMES T. THORSON3, JANET A. NYE4**

*1 School of Aquatic and Fishery Sciences, University of Washington, Seattle, Washington, USA*

*2 NOAA Northeast Fisheries Science Center, Woods Hole, Massachusetts, USA*

*3 NOAA Northwest Fisheries Science Center, Seattle, Washington, USA*

*4 School of Marine and Atmospheric Science, Stony Brook University, Stony Brook, New York, USA*

Manuscript submitted to *Fisheries Research*

*\* Corresponding author address*:

Haikun Xu, School of Aquatic and Fishery Sciences,

University of Washington, Seattle, WA 98195

E-Mail: [haikunxu@uw.edu](mailto:haikunxu@uw.edu)

# Abstract

Survival is an important population attribute in fisheries stock assessment models and is typically treated as a deterministic process without process uncertainty. A recently developed state-space assessment model, however, treated the variation in survival as stochastic and independent of age and year. Several studies have shown that some population attributes such as recruitment can be autocorrelated, and that not accounting for autocorrelated population processes in stock assessment models may result in notably biased estimates of population attributes. These studies as well as the strong retrospective pattern found in the last assessment of Southern New England yellowtail flounder motivated us to evaluate the two-dimensional (age and year) autocorrelations in survival. We found that survival deviations were first-order autocorrelated among both ages (autocorrelation = 0.56, standard error = 0.10) and years (autocorrelation = 0.33, standard error = 0.12). Moreover, the model which estimated the autocorrelations in survival deviations fitted data much better (AIC = 22) and had notably reduced retrospective patterns than the model which assumed independent survival. We also found that estimating the autocorrelations in survival deviations altered both the estimates and predictions of spawning biomass and fishing mortality by as much as 40%. Finally, by comparing the relative effect of an environmental covariate for recruitment and this 2-dimensional smoother for survival on model outcome, we showed that the survival smoother is notably more important to the assessment of Southern New England yellowtail flounder in terms of improving model fit according to AIC and providing reliable SSB prediction.

Keywords: survival; autocorrelation; State-space model; stock assessment; yellowtail flounder

# Introduction

Biological processes of a fish population usually, if not always, vary over age and time. Previous studies have shown that the variations in population processes can be autocorrelated among ages (Berg and Nielsen 2016; Nielsen and Berg 2014) or in time. For instance, a population process such as recruitment can be autocorrelated in time if the environmental or ecological process by which it is driven is autocorrelated in time (Johnson et al. 2016; Thorson et al. 2014). Johnson et al. (2016) found that in cases where recruitment is highly autocorrelated in time, ignoring this autocorrelation can potentially lead to large biases in model predictions as well as the associated uncertainty intervals. A population process, such as selectivity, can also be autocorrelated among ages (Nielsen and Berg 2014) probably because adjacent age classes have similar size, physiology, behavior, etc. When the variations in selectivity are autocorrelated among ages, the model including this autocorrelation was shown to fit to data better compared to that assuming mutually independent selectivity variations (Nielsen and Berg 2014). To date, however, evaluating a two-dimensional (2D) autocorrelation structure with age and year in population processes is rare. One exception is Cadigan (2015), who developed a state-space, age-structured assessment model for Northern Cod (*Gadus morhua*) where the among-age and among-year autocorrelations in natural mortality rate () were simultaneously estimated. However, the goodness-of-fit, retrospective pattern, and model estimates/predictions under alternative autocorrelation structures for were not compared.

Yellowtail flounder (*Limanda ferruginea*)is a commercially important demersal flatfish in the Northwest Atlantic ranging from the Labrador Sea in the North to the Chesapeake Bay in the South (NEFSC 2012). There are four stocks of yellowtail flounder delineated by the following areas: Canadian Grand Banks, Cape Cod-Gulf of Maine, Georges Bank, and Southern New England-Mid Atlantic. All four stocks experienced overfishing from the 1970s to the mid-1990s and since then all stocks have experienced some recovery (Stone et al. 2004) except for the southern-most Southern New England-Mid Atlantic (SNEMA) stock. The SNEMA stock is currently assessed using a statistical catch-at-age model (Age-Structured Assessment Program; Legault and Restrepo 1998) and the quality of recent assessments of the stock is claimed to be considerably degraded by the existence of two major sources of uncertainty (NEFSC, 2012 and L. Alade, NMFS NEFSC, per. comm). The first source of uncertainty is the inability to explain the dramatic decrease in recruitment since the 1990s. A recent study suggested that unfavorable environmental conditions due to a northward shift of the Gulf Stream path and the low spawning stock biomass (SSB) level were the two primarily reasons for persistently low recruitment in recent years (Xu et al. accepted). The second source of uncertainty is the strong retrospective pattern of spawning stock biomass (SSB) and fully-selected fishing mortality rate (*F*). The reason for the strong retrospective patterns is unclear, but no doubt, strong retrospective patterns in SSB and *F* can induce large uncertainty to the determination of stock and harvest status, respectively.

Retrospective pattern refers to the systematic inconsistency in estimates of fishery variables when addition years of data are added to the assessment model (Mohn 1999). It typically arises due to misspecifying the temporal changes in input data or biological parameters in stock assessment model (Hurtado-Ferro et al. 2014; Legault 2009). To address retrospective issue in the assessments of some New England fish stocks such as Georges Bank yellowtail flounder (Legault et al. 2012) and Gulf of Maine Atlantic cod (NEFSC 2013), stock assessment scientists sometimes impose a temporal trend to *M* in stock assessment models. *M* is an important parameter in stock assessment models because it directly influences stock productivity and moreover, misspecifying *M* can lead to biased estimation of population attributes (Miller and Legault 2017; Thorson et al. 2015) and key reference points such as virgin biomass and maximum sustainable yield (Johnson et al. 2015). However, *M* is often difficult to estimate accurately in stock assessment models because it is confounded with some other parameters such as fishing mortality and recruitment, so *M* is usually pre-specified as a time-invariant constant to simplify model estimation (Deroba and Schueller 2013; Johnson et al. 2015; Legault and Palmer 2015). Deroba and Schueller (2013) conducted a simulation experiment to evaluate the estimation biases in SSB and recruitment that are induced by misspecifying . They found that misspecifying the temporal variation in likely induces larger biases than misspecifying the age-variation in does, underling the importance of correctly specifying the temporal trend in to stock assessment models. Recognizing the empirical relationship that is an allometric function of fish size (Lorenzen 1996), time-invariant at age based on the allometric relationship that defines how declines with size was pre-specified in the last benchmark assessment of SNEMA yellowtail flounder (NEFSC 2012). In fact, the impacts of misspecifying are positively related to (Legault and Palmer 2015). Considering that were at the historical maximum in the last several years when was at the historical low level, misspecifying the temporal trend in can be a major reason for the severe retrospective problem in recent assessments of the stock.

Different from previous methods that imposed a temporal trend on , we explored a more flexible and objective method to address the retrospective problem via implementing a smoother on survival that was autoregressive over both age and year (i.e., two-dimensional). Particularly, we implemented an age-temporal first-order autoregressive (AR(1)) structure for survival into a state-space age-structured assessment model that was recently developed for SNEMA yellowtail flounder (Miller et al. 2016). Different from the more widely used statistical catch-at-age models that do not distinguish observation and process errors, state-space models are able to simultaneously estimate the process errors in unobserved population processes and the observation errors in associated data (Miller et al. 2016; Nielsen and Berg 2014). The simulation test by Miller and Hyun (in review) indicated that compared to statistical catch-at-age models, state-space models provide less biased estimation of population attributes and the associated uncertainty intervals, and a similar result has been demonstrated repeatedly for surplus production models (e.g., Ono et al. 2012). In the state-space age-structured assessment model developed by Miller et al. (2016), the environmental covariate incorporated in the stock-recruit function and annual survival are both assumed to be stochastic, and moreover, the stochastic variations in annual survival are assumed to be independent of age and time (Miller et al. 2016).

In this study, we evaluated the autocorrelation structure for SNEMA yellowtail flounder survival by modifying the likelihood function for survival deviations in the state-space model. With the development of software packages such as AD Model Builder (ADMB, Fournier et al., 2012) and Template Model Builder (TMB, Kristensen et al., 2015), various assumptions regarding survival autocorrelation structure can be compared objectively in the state-space model based on retrospective metrics such as Mohn’s rho (Mohn 1999) and some fitting criteria such as Akaike information criterion (AIC) (Burnham and Anderson 2002). We also assessed the degree to which the 2D survival smoother make significant differences in the estimates and near-term predictions of SB and *F* from the state-space model. Finally, we compared the relative importance of incorporating an environmental covariate into the stock-recruit function and implementing the proposed 2D survival smoother to SNEMA yellowtail flounder assessment.

# Material and Methods

## 2.1. 2D AR(1) survival smoother

We first compared various autocorrelation structures for survival deviations in the state-space model. For simplicity, we only considered the first-order autocorrelation structure that was widely used in previous studies (Cadigan 2015; Nielsen and Berg 2014). In the state-space model, the stochastic survival deviation () in age and year *y* () can be calculated by rewriting the stock equations:

where represents abundance at age and is the stock-recruit function in which an environmental time series () can be incorporated as a covariate. This state-space model is unique in its ability of incorporating the stochastic change of the environmental covariate over time, uncertainty in associated observations, and its effect on recruitment as a covariate in the stock-recruit function (Miller et al. 2016). is the total number of observation and prediction years and represents the plus group in the model. Strictly speaking, the survival deviation term () stands for population migration into or out of the stock because it does not alter either the *M* or *F* in the Baranov catch equation (Gudmundsson and Gunnlaugsson 2012), and in fact realized “survival” can be greater than one (i.e., whenever ). Indeed, population mixing between adjacent yellowtail flounder stocks has been observed in tagging studies, but the extent of which was not large enough to significantly affect the population dynamics of each individual stock, including the depleted SNEMA stock (Cadrin 2003; Goethel et al. 2015). Gudmundsson and Gunnlaugsson (2012) claimed that the survival deviation term can also be approximately treated as stochastic variation in , because impacts population dynamics primarily through stock equations. In fact, this survival deviation term can also be caused by deviations in fishing mortality or more generally deviations from the Baranov catch equation.

Miller et al. (2016) assumed that survival deviations are independent of age and time and normally distributed with zero means. In other words, for all *a* and *y*:

(2)

where for all ages were assumed to be the same but different from that for . It is based on the fact that survival variations for young-of-the-year (recruitment) are generally the largest. Another advantage of this state-space model over statistical catch-at-age models is that can be estimated internally in the model as a fixed effect parameter. The survival deviations of SNEMA yellowtail flounder or any other stock, however, are not necessarily independent. If survival deviations are autocorrelated among ages and years, they should follow a multivariate normal distribution instead:

(3)

where , is the  covariance matrix for the multivariate normal distribution and is calculated as the Kronecker product of the covariance matrix for the AR(1) process among ages () and the correlation matrix for the AR(1) process among years ():

where and are the two AR(1) coefficients in age and time, respectively. Either of them can be fixed at a constant between -1 and 1 or estimated in the state-space model as a fixed effect parameter. Note that when both and are fixed at 0, there is no survival autocorrelation in both dimensions so Eq. 3 is essentially the same as Eq. 2. In fact, is the likelihood distribution function for this covariance structure:

which means that the covariance between two survival deviations is positively related to how close the locations of the two survival deviations are on the age-time surface.

Age composition observations in the state-space model are assumed to follow a logistic-normal distribution (Atchison and Shen 1980):

where and are the vector of observed and predicted age proportions in year , respectively. For age composition data associated with indices of abundance (),

and for age composition data associated with catch (),

The variance in Eq. 6 is parameterized as

where is the estimated variance of observation error and is the effective sample size.

## 2.2. Application to SNEMA yellowtail flounder

We assessed the performance of our proposed 2D AR(1) survival smoother by using the data for SNEMA yellowtail flounder as a case study. It is worth noting that Xu et al. (accepted) found recruitment of SNEMA yellowtail flounder was significantly correlated with the GSI, and furthermore, incorporating the GSI into the stock-recruit function significantly improved model performance as measured by AIC. In detail, the GSI was found to affect the carrying capacity for pre-recruits and thereby the model in which the GSI was incorporated into the stock-recruit function as a limiting factor had the lowest AIC (Xu et al. accepted):

where is from Eq. 1, and , the environmental covariate for recruitment, is GSI. Here, we considered eight alternative models with differing stock-recruit functions or/and autocorrelation structures for survival deviations:

1. S0(base) assumes independent survival deviations and a Beverton-Holt stock-recruit function;
2. S0(age) is the same as S0(base) except with a survival AR(1) process among ages;
3. S0(year) is the same as S0(base) except with a survival AR(1) process among years;
4. S0(age, year) is the same as S0(base) except with survival AR(1) processes among both ages and years;
5. S(base) assumes independent survival deviations and incorporates the Gulf Stream Index (GSI) into the Beverton-Holt stock-recruit function, as specified in the best performing model found by Xu et al. (accepted);
6. S(age) is the same as S(base) except with a survival AR(1) process among ages;
7. S(year) is the same as S(base) except with a survival AR(1) process among years;
8. S(age, year) is the same as S(base) except with survival AR(1) processes among both ages and years.

The eight models all fitted to the same dataset that were used in the last benchmark assessment (NEFSC 2012) and made a three-year SSB prediction for 2012-2014. The observation data that were used to fit the state-space model include (1) three indices of abundance from the spring, fall, and winter bottom trawl surveys; (2) two spawning stock indices from ichthyoplankton surveys; (3) one commercial catch; and (4) annual age composition observations from the three bottom trawl surveys and the commercial catch (Miller et al. 2016). Empirical weight-at-age observation was used in the model as data, so were maturity and natural mortality at age. Maturity in the model was fixed at 0 for age 1, 0.47 for age 2, and 1 for age 3-6 (NEFSC 2012). All the observation data mentioned above can be accessed at [www.nefsc.noaa.gov/saw/sasi/sasi\_report\_options.php](http://www.nefsc.noaa.gov/saw/sasi/sasi_report_options.php). During prediction years, was fixed at 0 and weight, maturity, as well as natural mortality at age were fixed at the values used for the terminal year of the analysis. Hence, the difference in SSB predictions the eight models made should be solely caused by differing stock-recruit functions or/and autocorrelation structures for survival deviations.

Performance of the eight models was compared quantitatively with respective to model fit and retrospective pattern. Specifically, model fit was compared based on AIC, which is a linear function of the maximized marginal log-likelihood and number of fixed effect parameters in the state-space model. In TMB, the marginal likelihood of fixed effect parameters is calculated using the Laplace approximation to integrate across random effect parameters (Kristensen et al. 2015), and fixed effect parameters are then estimated via maximizing the marginal likelihood within the R computing environment (R Core Team 2015) . In this study, the “nlminb” function was used to minimize the negative of the marginal log-likelihood in R. After the fixed effect parameters being estimated, TMB predicts the random effect coefficients using empirical Bayes (Kristensen et al. 2015). The retrospective pattern was compared according to Mohn’s rho (Mohn 1999) of the seven retrospective peels for years 2004-2010, which is consistent with the process in the last benchmark assessment (NEFSC 2012):

where

is the estimate of SSB or in the last year of the data used in each retrospective peel.

For the two models with (S0(age, year)) and without (S0(base)) the 2D AR(1) survival smoother in particular, we compared the estimated and predicted survival deviations and abundance at age they provided. Also, the estimated and predicted SSB and , as well as the associated uncertainty intervals, from the two models were compared to evaluate the impacts of the 2D AR(1) survival smoother on the two population attributes that determine stock and harvest status. Considering that incorporating an environmental covariate into the stock-recruit function is rare for stock assessments, we compared S0(base) and S0(age, year) instead of S(base) and S(age, year) in this section for more general conclusions regarding how sensitive a stock assessment could be to the autocorrelation structure of survival deviations in the state-space model.

Finally, we compared the extent to which an environmental covariate for recruitment (S(base)) and the 2D AR(1) smoother for survival deviations (S0(age, year)) can influence model estimates and predictions. This comparison was made based upon the relative difference, , in SSB and . is model estimate from S(base) or S0(age, year) and is model estimate from S0(base), so the relative difference between S(base) and S0(base) should be solely due to incorporating the GSI into the stock-recruit function and that between S0(age, year) and S0(base) should be solely due to implementing the 2D AR(1) smoother for survival. The recruitment variation of SNEMA yellowtail flounder has been found to be strongly related to the environmental fluctuation caused by the latitudinal shift of the Gulf stream (Xu et al. accepted). Consequently, SNEMA yellowtail flounder is an ideal stock for comparing the relative importance of understanding the environment-recruitment relationship and evaluating the autocorrelations in survival deviations to a stock assessment. The comparison aimed at providing suggestions regarding the priorities of future stock assessment research for SNEMA yellowtail flounder and perhaps other similar fish stocks as well.

# Results

## 3.1. Goodness-of-fit

In general, incorporating the GSI into the stock-recruit function and implementing the 2D AR(1) smoother for survival both resulted in better model fit, but the improvement in model fit (as measured by AIC) that was due to implementing the 2D AR(1) survival smoother was considerably larger (Table 1). Specifically, relatively to S0(base) in which neither process was included, incorporating the GSI into the stock-recruit function (S(base)) reduced the AIC by 12.84, in contrast to by 32.48 due to implementing the 2D AR(1) survival smoother (S0(age, year)). The model in which both processes were included (S(age, year)) fitted the data best, as suggested by more than 2 unit decrease in AIC compared to any other model. Regardless of the form of the stock-recruit function, the model assuming 2D autocorrelated survival deviations fitted substantially better than the two corresponding models assuming 1D autocorrelated survival, indicating that survival deviations for SNEMA yellowtail flounder were autocorrelated among both ages and years. Moreover, the retrospective AIC value of the eight models showed that the ranking of eight model performances in the terminal year of data was persistent over at least the last eight years (Table 2). According to AIC and the estimates of and , the among-year autocorrelation in survival deviations was higher and had larger impact on model fit than the among-age autocorrelation in survival deviations, which, however, was not negligible (Table 1).

## 3.2. Retrospective pattern

In addition to improving model fit, implementing the 2D AR(1) survival smoother into the state-space model also reduced the retrospective patterns for SSB and to a large extent. In S0(base), the model without the 2D AR(1) survival smoother, the Mohn’s rhos for SSB and *F* were notably positive (0.11) and negative (-0.14), respectively (Fig. 1, first row). By contrast, in S0(age, year) where the 2D AR(1) survival smoother was implemented, the Mohn’s rhos for SSB (0.02) and (-0.04) became negligibly different from zero (Fig. 1, last row). In S0(age) and S0(year) where survival autocorrelation in only one of the two dimensions was accounted for, the Mohn’s rhos (in absolute) for SSB and *F* were smaller than those in S0(base) but larger than those in S0(age, year). The retrospective patterns for the four models with the GSI effect on recruitment were quantitatively similar (result not shown): the Mohn’s rhos for SSB and *F* were relatively large when the 2D AR(1) survival smoother was not implemented, but became negligible otherwise. The retrospective patterns for SSB and *F* further emphasized the necessity of implementing the 2D AR(1) survival smoother into the state-space model to account for autocorrelated survival of SNEMA yellowtail flounder.

## 3.3. The 2D AR(1) survival smoother changes model estimates and predictions

In this section we examined the extent to which the 2D AR(1) survival smoother impacts model estimates and predictions. First, we compared the estimated and predicted survival deviation at age from S0(base) and S0(age, year) to understand how the 2D AR(1) survival smoother affects the pattern of survival deviations on the age-time surface. Despite survival deviations in S0(base)was assumed be independent, the estimated survival deviations tended to have the same sign in adjacent ages and years (Fid. 2a), implying that survival was likely autocorrelated among both ages and years. As expected, the estimated survival deviations from S0(age, year) varied more smoothly over both age and time (Fig. 2b). Survival deviations in S0(base) were assumed to be mutually independent and normally distributed with mean zero, so their best prediction from S0(base) should be zero in 2012-2014 for all age classes (Fig. 2a). The survival deviations in S0(age, year), however, were estimated to be highly autocorrelated among years (), so the best prediction for survival deviations had the same sign (negative) as the terminal year estimate and a[symptot](javascript:void(0);)ically approached zero over time (Fig. 2b). That is to say, the predicted survival at age from S0(age, year) was lower than that from S0(base). Considering that the estimates of abundance at age from S0(base) and S0(age, year) were similar in the terminal year, 2011, the lower survival at age prediction in S0(age, year) corresponded to lower abundance at age prediction in S0(age, year) (Fig. 3).

The 2D AR(1) survival smoother also had substantial impacts on SSB and *F*. The estimates of SSB and from S0(base) and S0(age, year) were similar before 1990 but notably different thereafter (Fig. 4). In comparison with S0(base), S0(age, year) provided higher SSB estimates and lower estimates for 1990-2011. Although the terminal year SSB estimate from S0(age, year) was slightly higher than that from S0(base), the SSB prediction from S0(age, year) was considerably lower than that from S0(base). Again, this was due to that S0(age, year) accounted for the high among-year autocorrelation in survival deviations and thereby provided lower survival prediction than S0(base) did. Interestingly, although the SSB prediction from S0(age, year) was lower than that from S0(base), the associated uncertainty was larger.

## 3.4. Relative impact of recruitment covariate vs. survival smoother

Finally, we compared the extent to which model estimates and predictions are impacted by incorporating an environmental covariate into the stock-recruit function versus implementing the 2D AR(1) survival smoother (Fig. 5). Incorporating the GSI into the stock-recruit function (S(base)) minimally impacted the estimates of SSB and F, and its effect on SSB prediction was only noticeable in the second and third prediction years. Conversely, to compensate for the lower survival estimates after 1990 due to implementing the 2D AR(1) survival smoother, the estimates of SSB and *F* from S0(age, year) were as much as 60% higher and 40% lower than those from S0(base), respectively. Furthermore, implementing the 2D AR(1) survival smoother (S0(age, year)) also had a larger effect on SSB in all three prediction years compared to incorporating a climate covariate into the stock-recruit function (S(base)).

To understand why the recruitment covariate and survival smoother had considerably different impacts on model estimates and predictions, we compared the estimated and predicted SSB proportion at each age from S(age, year), the best fitting model previously found. Although interannual variations existed in SSB proportion at age, age 3-5 generally made the largest contributions to SSB around the terminal year of the data used in the assessment (Fig. 6). Age 1 contributed negligibly to SSB because few SNEMA yellowtail flounder become mature within their first year and the weight of each of these young fish is much less than older individuals.

As a result of the historically low F around the terminal year and zero F in prediction years, the contributions of old (or young) age classes became increasingly high (or low) after 2007. In the first prediction year (2012), incorporating the GSI into the stock-recruit function led to lower recruitment prediction due to unfavorable environmental conditions as suggested by the high GSI in 2011 (Fig. 3a in Xu et al. accepted). However, the SSB proportion at age pattern showed that the first age class made a negligible contribution to SSB, so the first SSB prediction S(base) provided was negligibly affected by the recruitment covariate. Later on, this negative recruitment signal due to the recruitment covariate propagated to age 2 and age 3 in the second and third prediction years, respectively, leading to increasingly lower SSB predictions compared to S0(base). Owing to that negative survival deviations were predicted for the oldest age classes, which are the primary SSB contributors, the survival smoother (S0(age, year)) immediately lower the first SSB prediction compared to S0(base) (Fig. 5). Note that the survival deviation at age was predicted to a[symptot](javascript:void(0);)ically approach zero over time (see Fig. 2), so the direct influence of the survival smoother on SSB predictions are expected to be increasingly weaker over time. However, the predicted negative survival deviations in the first prediction year still propagate to older age classes over time and indirectly influence SSB predictions in the second and third prediction years. In this case, the direct and indirect impacts of the survival smoother on SSB predictions both lead to increasingly lower SSB predictions from S0(age, year) compared to S S0(base) and even S(base).

# Discussion

Using a state-space, age-structured assessment model developed for SNEMA yellowtail flounder, we showed that implementing the 2D AR(1) survival smoother considerably improved model fit and reduced the retrospective patterns for SSB and . Different from some previous studies which addressed the retrospective problem in New England stock assessments by applying the temporal trend in , this paper provides a more objective, flexible, and generic approach to reduce retrospective pattern in stock assessments. Particularly, both the two AR(1) coefficients for the survival smoother can be pre-specified as a constant or estimated in the state-space model as fixed effect parameters. By comparing the AIC of four alternative models with differing autocorrelation assumptions for survival deviations, we showed that it is more important to estimate the among-year than among-age autocorrelation in survival deviations of SNEMA yellowtail flounder. The among-year autocorrelation in survival deviations was estimated to be high (0.56) in the best fitting model, indicating that biological processes for ages larger than 1 can also be autocorrelated across time. Moreover, the estimated survival deviations for age 5&6 were quantitatively comparable to those for age 1, informing us that the variation in adult survival is at least as influential as that in recruitment to the population dynamics of SNEMA yellowtail flounder.

We found that the improvement in model fit that was induced by the survival smoother alone was notably larger than that was induced by the recruitment covariate alone. We also found that the survival smoother had profound impacts on SSB and estimates as well as near-term SSB predictions, while the recruitment covariate negligibly impacted the estimates of SSB and . These two findings further underscored the importance of implementing the 2D AR(1) survival smoother to the assessment of SNEMA yellowtail flounder. Indeed, this 2D AR(1) smoother for survival deviations can also extended to three dimensions if, for example, survival is modelled to be sex-specific. Also, the generic 2D AR(1) smoother we introduce in this paper can be applied to other potentially autocorrelated biological processes (e.g., selectivity and growth) as well, which we predict would be an important research topic for continuously developing mixed-effect stock assessment models.

For SNEMA yellowtail flounder, implementing the 2D AR(1) survival smoother in the state-space model considerably reduced the severity of retrospective pattern and modified near-term SSB prediction as well as the associated uncertainty interval. Indeed, both the Mohn’s rhos for SSB and *F* became negligibly different from zero when the survival smoother was implemented, which means that the state-space model with the survival smoother is more reliable in terms of determining the stock and harvest status of SNEMA yellowtail flounder. Because the among-year autocorrelation in survival deviations was estimated to be high, the best predictions for survival deviation at age from the model with the survival smoother (S0(age, year)) have the same sign as those estimated for the terminal year and decline exponentially towards zero with slope determined by the autocorrelation coefficient. In contrast, the best predictions for survival deviation at age from the model without the survival smoother (S0(base)) are zero as a result of the independent assumption for survival deviations. Therefore, S0(age, year) provided lower SSB predictions than S0(base) for 2012-2014. In addition to the median SSB, the associated uncertainty predicted by the model with and without the survival smoother was also notably different. The SSB predictions from the model with the survival smoother have wider uncertainty interval, namely, not accounting for the positive autocorrelation in survival deviations corresponds to under-estimating the uncertainty in near-term SSB predictions.

Near-term SSB predictions changed substantially when the GSI was incorporated into the stock-recruit function or the 2D AR(1) smoother was implemented for survival. The reason that the recruitment covariate (GSI) only impacted the second- and third-year SSB predictions can be found in the maturity pattern. In 2012, the GSI-induced recruitment signal does not affect SSB prediction because the maturity at age 1 is zero. In 2013 and 2014, the GSI-induced recruitment signal propagates to age 2 and 3 at which maturity is 0.47 and 1, respectively, so the recruitment covariate made a significant difference in the second (-8%) and third (-24%) SSB predictions. It is worth noting that while the change in maturity from age 1 to older ages is dramatic, the increasingly amplified effect of the recruitment covariate on SSB predictions over time is also contributed by much larger weights at older ages than age 1. The survival smoother, in contrast, has notably (-10%) altered the first SSB prediction in 2012 already because the change in predicted survival it induces are large and even comparable to the specified in the model (0.23-0.41, corresponding to age 6-1). Later on, the survival smoother induces even larger changes in the second (-24%) and third (-30%) SSB predictions, both of which are larger than the corresponding ones due to the recruitment covariate. In terms of the accuracy of near-term SSB predictions, implementing the 2D AR(1) survival smoother is likely to be more important to SNEMA yellowtail flounder than incorporating an environmental effect on recruitment.

The 2D AR(1) survival smoother, however, does not make calculating biological reference points (BRPs) easier. We envision two different ways to calculate BRPs when using this survival smoother:

1. *Deterministic BRP*: The simplest calculation for BRPs is to ignore the stochastic variation in survival and calculate the yield curve given average survival. This procedure ignores the process error in biological processes and is typically used to calculate BRPs.
2. *Dynamic BRP*: Regime shifts in recruitment have been shown to trigger rebuilding rules that target biomass levels that are essentially impossible under new recruitment dynamics (Szuwalski et al. 2015). Therefore, BRPs can vary across time and be calculated annually based upon the survival deviation estimated for each year.

In the case study, we found that the estimated survival deviations were predominantly negative since the 1990s, so ignoring this trend in survival may result in largely biased BRPs estimation for the recent three decades. Whereas dynamic BRPs seems to be the more appropriate way of calculating BRPs, it is more challenging because an assumption regarding how survival deviations are attributed to fishing mortality, natural mortality, and migration is required before calculating the dynamic BRPs. These three attributors are generally confounded in stock assessment models, which means that it is difficult, if possible, to partition their influences on the estimates of population attributes including survival. We recommend future research exploring their relative performance when coupled with management procedures and given different forms of time-variation in survival.

Care should be taken when interpreting the main findings found in this study. First of all, the survival for the plus group is not strictly survival. It can be a caveat in this study considering that the predicted survival deviations for the plus group changed notably as the survival smoother implemented in the state-space model. Secondly, population dynamics in 2012-2014 was predicted in the absence of fishing (i.e., for ) to remove the potential impacts of differing F scenarios on near-term SSB predictions. This assumption likely results in an exaggerated shift of the predicted age composition towards older age classes. However, a parallel SSB prediction run assuming more realistic fishing mortality patterns (status quo fishing mortality) for 2012-2014 provided similar age structure predictions (not shown), because *F* reached the historically low level (~0.2) in 2011 due to the low stock productivity in recent years. Thus, the conclusions made under zero *F* for 2012-2014 are robust considering the current SNEMA yellowtail flounder harvest status. Lastly, all the conclusions made in this study are species-specific. For example, the relative importance of the recruitment covariate and survival smoother to SSB prediction is highly dependent upon the parameters (e.g., age at maturity, longevity, selectivity and weight-at-age) that influence the age structure and life history of the stock. Generally speaking, the survival smoother can directly impact the predicted survival at all ages, and thereby, is expected to be more important to near-term SSB prediction for late-matured and long-lived fish stocks. While a recruitment covariate can only directly impact near-term recruitment prediction. For late-matured and long-lived fish stocks, the covariate-induced change in recruitment prediction is not able to propagate to the majority of spawning age classes and notably impact SSB prediction in the near-term.

Lastly, it is important to note that implementing the 2D AR(1) survival smoother reduced the computation speed as it broke the sparseness of the Hessian matrix. In this case study with 69 fixed effects parameters and 287 random effect parameters, the implementation increased model convergence time from 10 seconds to 200 seconds in a laptop computer. The extent to which the convergence time increases in other cases due to the implementation is expected to be positively associated with the case-specific number of age classes and time steps (e.g., years) considered in the model. The implementation can actually be even more computational expensive when a better representation of the among-age survival autocorrelation structure is included, e.g., the irregular lattice correlation structure investigated in Berg and Nielsen (2016). However, with continuing improvement in computation ability, the consequence of reduced convergence speed is unlikely to be a hurdle in implementing this method in the future.

Table 1. Model results where only numbers-at-age (NAA) were estimated as random effects. NAA-6, where survival is a 2D AR1 process across years and ages, had the lowest AIC and Mohn’s . Mohn’s abbreviations: *R* = recruitment, *SSB* = spawning stock biomass, and *F* = fishing mortality averaged over ages 4-5. Maximum likelihood estimates of parameters constraining random effects are listed with standard error in parentheses.

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Estimated parameters | | | |  | Model fit | | |  | Mohn’s | | |
| Model |  |  |  |  |  | -log | AIC |  |  |  |  |  |
| NAA-1 | 1.18 (0.13) | — | — | — |  | -934.024 | -1712.0 | 221.6 |  | 5.86 | 0.99 | -0.42 |
| NAA-2 | 0.66 (0.08) | — | 0.92 (0.05) | — |  | -957.164 | -1756.3 | 177.3 |  | 4.56 | 0.90 | -0.38 |
| NAA-3 | 1.01 (0.12) | 0.58 (0.05) | — | — |  | -1027.684 | -1897.4 | 36.2 |  | 0.97 | 0.27 | -0.17 |
| NAA-4 | 0.86 (0.12) | 0.53 (0.05) | — | 0.46 (0.10) |  | -1036.859 | -1913.7 | 19.9 |  | 0.55 | 0.12 | -0.06 |
| NAA-5 | 0.76 (0.11) | 0.48 (0.05) | 0.60 (0.09) | — |  | -1044.252 | -1928.5 | 5.1 |  | 0.73 | 0.14 | -0.08 |
| NAA-6 | 0.72 (0.10) | 0.47 (0.05) | 0.52 (0.10) | 0.34 (0.12) |  | -1047.803 | -1933.6 | 0.0 |  | 0.56 | 0.09 | -0.03 |

Table 2. Model results where only recruitment and natural mortality (M) were estimated as random effects. M-10, which estimated mean M, , as well as 2D AR1 random effects on M, had the lowest AIC but not Mohn’s . M-5 and M-7 had lower Mohn’s than M-10, but higher AIC. Numbers-at-age (NAA) were treated as in model NAA-1. Only estimating , without any random effects on NAA or M, substantially reduced AIC and Mohn’s (compare M-1 to M-6). Mohn’s abbreviations: *R* = recruitment, *SSB* = spawning stock biomass, and *F* = fishing mortality averaged over ages 4-5. Maximum likelihood estimates of parameters constraining random effects are listed with standard error in parentheses.

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Estimated parameters | | | |  | Model fit | | |  | Mohn’s | | |
| Model |  |  |  |  |  | -log | AIC |  |  |  |  |  |
| M-1 | — | — | — | — |  | -934.024 | -1712.0 | 227.5 |  | 5.86 | 0.99 | -0.42 |
| M-2 | — | 1.18 (0.09) | — | — |  | -1037.981 | -1918.0 | 21.5 |  | 0.12 | 0.20 | -0.13 |
| M-3 | — | 1.15 (0.43) | — | 0.25 (0.49) |  | -979.398 | -1798.8 | 140.7 |  | 2.06 | 0.10 | -0.08 |
| M-4 | — | 0.13 (0.06) | 0.98 (0.02) | — |  | -988.290 | -1816.6 | 122.9 |  | 1.19 | -0.12 | 0.30 |
| M-5 | — | 0.89 (0.12) | 0.47 (0.21) | 0.39 (0.14) |  | -1042.965 | -1923.9 | 15.6 |  | -0.07 | 0.10 | -0.06 |
| M-6 | 0.79 (0.04) | — | — | — |  | -976.715 | -1795.4 | 144.1 |  | 2.36 | 0.15 | -0.09 |
| M-7 | 0.70 (0.12) | 0.71 (0.10) | — | — |  | -1045.295 | -1930.6 | 8.9 |  | -0.13 | -0.03 | 0.09 |
| M-8 | 0.81 (0.11) | 0.42 (0.16) | — | -0.86 (0.19) |  | -983.717 | -1805.4 | 134.1 |  | 1.98 | 0.12 | -0.07 |
| M-9 | 0.36 (0.22) | 0.12 (0.07) | 0.98 (0.03) | — |  | -983.959 | -1805.9 | 133.6 |  | 1.55 | -0.10 | 0.26 |
| M-10 | 0.83 (0.16) | 0.56 (0.08) | 0.37 (0.14) | 0.30 (0.16) |  | -1051.753 | -1939.5 | 0.0 |  | -0.31 | -0.17 | 0.27 |

Table 3. Model results where both numbers-at-age (NAA) and natural mortality (M) were estimated as random effects. NAA-M-5 had the lowest AIC but not Mohn’s . NAA-M-2 had higher AIC but the lowest Mohn’s . Four of the models did not converge. Mohn’s abbreviations: *R* = recruitment, *SSB* = spawning stock biomass, and *F* = fishing mortality averaged over ages 4-5.

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Estimated parameters | | |  | Model fit | | |  | Mohn’s | | |
| Model | NAA random effects | M random effects |  |  | -log | AIC |  |  |  |  |  |
| NAA-M-1 | , |  |  |  | -1051.090 | -1942.2 | 17.3 |  | 0.20 | 0.17 | -0.09 |
| NAA-M-2 | , | , , |  |  | -1048.691 | -1933.4 | 26.1 |  | -0.22 | 0.02 | 0.04 |
| NAA-M-3 | , |  | Yes |  |  |  |  |  |  |  |  |
| NAA-M-4 | , | , , | Yes |  | -1054.627 | -1943.3 | 16.2 |  | -0.52 | -0.28 | 0.54 |
| NAA-M-5 | , , , |  |  |  | -1061.727 | -1959.5 | 0.0 |  | 0.52 | 0.10 | -0.04 |
| NAA-M-6 | , , , | , , |  |  |  |  |  |  |  |  |  |
| NAA-M-7 | , , , |  | Yes |  |  |  |  |  |  |  |  |
| NAA-M-8 | , , , | , , | Yes |  |  |  |  |  |  |  |  |

Table 4. Results from models that estimated an effect of the Cold Pool Index (CPI) on recruitment, in addition to random effects on numbers-at-age (NAA) and natural mortality (M). NAA-M-CPI-1 had the lowest AIC but not Mohn’s . NAA-M-CPI-2 had higher AIC but the lowest Mohn’s of all models considered. Mohn’s abbreviations: *R* = recruitment, *SSB* = spawning stock biomass, and *F* = fishing mortality averaged over ages 4-5.

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Estimated parameters | | |  | Model fit | | |  | Mohn’s | | |
| Model | NAA random effects | M random effects |  |  | -log | AIC |  |  |  |  |  |
| NAA-M-CPI-1 | , |  |  |  | -996.345 | -1824.7 | 3.7 |  | 0.16 | 0.17 | -0.09 |
| NAA-M-CPI-2 | , | , , |  |  | -994.979 | -1818.0 | 10.4 |  | -0.11 | 0.02 | 0.06 |
| NAA-M-CPI-3 | , | , , | Yes |  | -1001.182 | -1828.4 | 0.0 |  | -0.42 | -0.25 | 0.47 |

A screenshot of a cell phone

Description automatically generated

Figure 1. Survival deviations by year and age estimated by models in which only numbers-at-age (NAA) were random effects.

A screenshot of a social media post

Description automatically generated

Figure 2. Deviations in natural mortality (M) by year and age estimated by models without numbers-at-age (NAA) random effects. Estimating mean M (right column) reduced the magnitude of M deviations for AR1 models (M-8 and M-9 vs. M-3 and M-4), but did not reduce the M deviations in models where the deviations varied by year and age (M-7 and M-10 vs. M-2 and M-5).

A picture containing bird

Description automatically generated

Figure 3. Reduction in retrospective patterns by including an effect of the Cold Pool Index (CPI) on recruitment, measured as the difference in Mohn’s between otherwise equivalent models. Across models with different random effects on numbers-at-age (NAA) and natural mortality (M), including CPI effects on recruitment reduced Mohn’s by about 0.1 on average and did not have much influence on or .

A picture containing brick

Description automatically generated

Figure 4. Deviations in natural mortality (M, top panel) and numbers-at-age (NAA, bottom panel) from the final model with lowest Mohn’s , NAA-M-CPI-2.

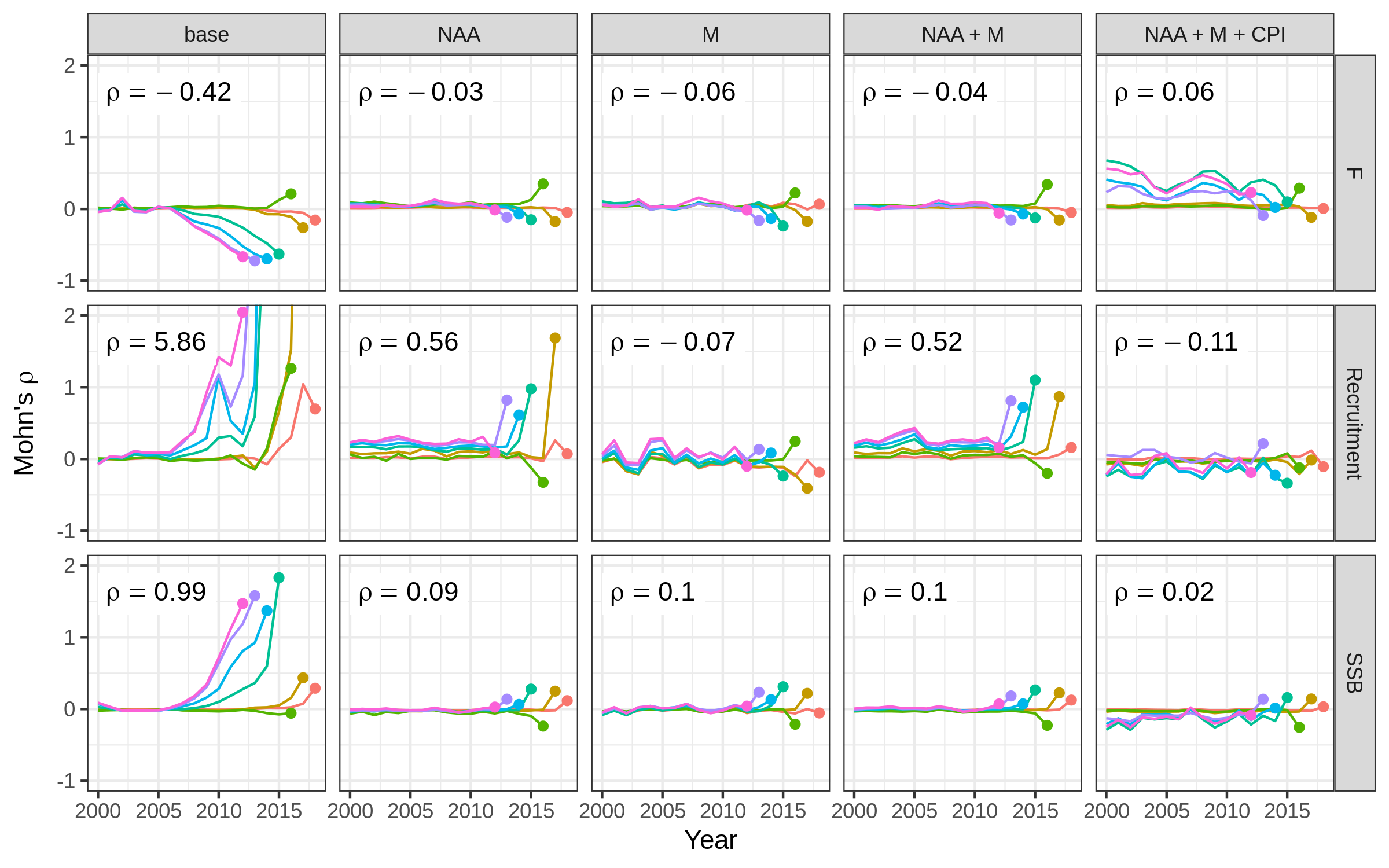


Figure 5. Retrospective patterns in fishing mortality (F, top row), recruitment (middle row), and spawning stock biomass (SSB, bottom row). Lines and points depict Mohn’s from seven peels, and the average Mohn’s is given in each panel. Columns show results by model: the statistical catch-at-age model, NAA-1, is presented as ‘base,’ and the rest are the models from Tables 1-4 with lowest Mohn’s (NAA-6, M-5, NAA-M-2, and NAA-M-CPI-2).

A close up of a map

Description automatically generated

Figure 6.

A close up of a map

Description automatically generated

Figure 7.

# References

Atchison, J.; Shen, S.M. Logistic-normal distributions: Some properties and uses. Biometrika. 67:261-272; 1980

Berg, C.W.; Nielsen, A. Accounting for correlated observations in an age-based state-space stock assessment model. ICES Journal of Marine Science: Journal du Conseil:fsw046; 2016

Burnham, K.P.; Anderson, D.R. Model selection and multimodel inference: a practical information-theoretic approach: Springer Science & Business Media; 2002

Cadigan, N.G. A state-space stock assessment model for northern cod, including under-reported catches and variable natural mortality rates 1. Canadian Journal of Fisheries and Aquatic Sciences. 72:1-13; 2015

Cadrin, S.X. Stock structure of yellowtail flounder off the northeastern United States. 2003

Deroba, J.J.; Schueller, A.M. Performance of stock assessments with misspecified age-and time-varying natural mortality. Fisheries Research. 146:27-40; 2013

Fournier, D.A.; Skaug, H.J.; Ancheta, J.; Ianelli, J.; Magnusson, A.; Maunder, M.N.; Nielsen, A.; Sibert, J. AD Model Builder: using automatic differentiation for statistical inference of highly parameterized complex nonlinear models. Optimization Methods and Software. 27:233-249; 2012

Goethel, D.R.; Legault, C.M.; Cadrin, S.X. Demonstration of a spatially explicit, tag-integrated stock assessment model with application to three interconnected stocks of yellowtail flounder off of New England. ICES Journal of Marine Science: Journal du Conseil. 72:164-177; 2015

Gudmundsson, G.; Gunnlaugsson, T. Selection and estimation of sequential catch-at-age models. Canadian Journal of Fisheries and Aquatic Sciences. 69:1760-1772; 2012

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. Looking in the rear-view mirror: bias and retrospective patterns in integrated, age-structured stock assessment models. ICES Journal of Marine Science: Journal du Conseil:fsu198; 2014

Johnson, K.F.; Councill, E.; Thorson, J.T.; Brooks, E.; Methot, R.D.; Punt, A.E. Can autocorrelated recruitment be estimated using integrated assessment models and how does it affect population forecasts? Fisheries Research. 183:222-232; 2016

Johnson, K.F.; Monnahan, C.C.; McGilliard, C.R.; Vert-pre, K.A.; Anderson, S.C.; Cunningham, C.J.; Hurtado-Ferro, F.; Licandeo, R.R.; Muradian, M.L.; Ono, K. Time-varying natural mortality in fisheries stock assessment models: identifying a default approach. ICES Journal of Marine Science: Journal du Conseil. 72:137-150; 2015

Kristensen, K.; Nielsen, A.; Berg, C.; Skaug, H. Template model builder TMB. J Stat Softw; 2015

Legault, C.M. Report of the retrospective working group. NOAA NMFS Northeast Fisheries Science Center Reference Document:09-01; 2009

Legault, C.M.; Alade, L.; Stone, H.H.; Gross, W.E. Stock assessment of Georges Bank yellowtail flounder for 2012. TRAC Ref Doc. 2:133; 2012

Legault, C.M.; Palmer, M.C. In what direction should the fishing mortality target change when natural mortality increases within an assessment? Canadian Journal of Fisheries and Aquatic Sciences. 73:1-9; 2015

Legault, C.M.; Restrepo, V.R. A flexible forward age-structured assessment program. ICCAT Col Vol Sci Pap. 49:246-253; 1998

Lorenzen, K. The relationship between body weight and natural mortality in juvenile and adult fish: a comparison of natural ecosystems and aquaculture. Journal of fish biology. 49:627-642; 1996

Miller, T.J.; Hare, J.A.; Alade, L.A. A state-space approach to incorporating environmental effects on recruitment in an age-structured assessment model with an application to Southern New England yellowtail flounder. Canadian Journal of Fisheries and Aquatic Sciences. 73:1261-1270; 2016

Miller, T.J.; Legault, C.M. Statistical behavior of retrospective patterns and their effects on estimation of stock and harvest status. Fisheries Research. 186:109-120; 2017

Mohn, R. The retrospective problem in sequential population analysis: An investigation using cod fishery and simulated data. ICES Journal of Marine Science: Journal du Conseil. 56:473-488; 1999

NEFSC. 54th Northeast Regional Stock Assessment Workshop (54th SAW) Assessment Report. US Dept Commer, Northeast Fish Sci Cent Ref Doc 12-18; 600 p; 2012

NEFSC. 57th Northeast Regional Stock Assessment Workshop (57th SAW) Assessment Report. US Deptartment of Commerce, Northeast Fisheries Science Center Reference Document; 2013

Nielsen, A.; Berg, C.W. Estimation of time-varying selectivity in stock assessments using state-space models. Fisheries Research. 158:96-101; 2014

Ono, K.; Punt, A.E.; Rivot, E. Model performance analysis for Bayesian biomass dynamics models using bias, precision and reliability metrics. Fisheries Research. 125:173-183; 2012

R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>. 2015

Stone, H.H.; Gavaris, S.; Legault, C.M.; Neilson, J.D.; Cadrin, S.X. Collapse and recovery of the yellowtail flounder (Limanda ferruginea) fishery on Georges Bank. Journal of Sea Research. 51:261-270; 2004

Szuwalski, C.S.; Vert‐Pre, K.A.; Punt, A.E.; Branch, T.A.; Hilborn, R. Examining common assumptions about recruitment: a meta‐analysis of recruitment dynamics for worldwide marine fisheries. Fish and Fisheries. 16:633-648; 2015

Thorson, J.T.; Jensen, O.P.; Zipkin, E.F. How variable is recruitment for exploited marine fishes? A hierarchical model for testing life history theory. Canadian Journal of Fisheries and Aquatic Sciences. 71:973-983; 2014

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