ICES Journal of
Marine Science

ICES Journal of Marine Science (2020), 77(3), 911–920. doi:10.1093/icesjms/fsaa034

Original Article

Simulation testing methods for estimating misreported catch in a state-space stock assessment model

Charles T. Perretti [1]*, Jonathan J. Deroba[1], and Christopher M. Legault[1]

[1]*National Marine Fisheries Service, Northeast Fisheries Science Center, 166 Water Street, Woods Hole, MA 02543, USA*

**Corresponding author: tel: 508-495-2323; e-mail: charles.perretti@noaa.gov.*

Perretti, C. T., Deroba, J. J., and Legault, C. M. Simulation testing methods for estimating misreported catch in a state-space stock assessment model. – ICES Journal of Marine Science, 77: 911–920.

Received 13 December 2019; revised 7 February 2020; accepted 10 February 2020; advance access publication 12 March 2020.

State-space stock assessment models have become increasingly common in recent years due to their ability to estimate unobserved variables and explicitly model multiple sources of random error. Therefore, they may be able to better estimate unobserved processes such as misreported fishery catch. We examined whether a state-space assessment model was able to estimate misreported catch in a simulated fishery. We tested three formulations of the estimation model, which exhibit increasing complexity: (i) assuming no misreporting, (ii) assuming misreporting is constant over time, and (iii) assuming misreporting follows a random walk. We tested these three estimation models against simulations using each of the three assumptions and an additional fourth assumption of uniform random misreporting over time. Overall, the worst estimation errors occurred when misreporting was ignored while it was in fact occurring, while there was a relatively small cost for estimating misreporting when it was not occurring. Estimates of population scale and fishing mortality rate were particularly sensitive to misreporting assumptions. Furthermore, in the uniform random scenario, the relatively simple model that assumed misreporting was fixed across ages and time was more accurate than the more complicated random walk model, despite the increased flexibility of the latter.

Keywords: catch misreporting, fisheries, state space, stock assessment.

Introduction

State-space models, and hierarchical models in general, have become increasingly common in fisheries science (Thorson and Minto, 2015) due to their ability to estimate unobserved processes and multiple sources of random variability. State-space models have been used in stock assessment for at least three decades (Mendelssohn, 1988; Sullivan, 1992; Gudmundsson, 1994) but recently have gained use due to advancements in statistical software that allow for the estimation of complex latent processes (Nielsen and Berg, 2014; Cadigan, 2016; Miller and Hyun, 2018). This ability to model unobserved processes (often as random effects) is potentially useful in stock assessment as it is known that fish populations are influenced by multiple sources of random variability and a variety of unobserved variables.

One particularly important unobserved process is misreported fishery catch. Misreported catch in the Northwest Atlantic has been documented in the form of mislabelled landings (*United States of America, v. Carlos A. Rafael*, 2017; Bellanger *et al.*, 2019), and misreported location of catch leading to incorrect stock attribution (Palmer, 2017). Misreporting has also been documented in the Northeast Atlantic in the form of underreported discards (ICES, 2014). Persistent misreporting can lead to systematic errors in a stock assessment, sometimes manifested as retrospective patterns (Legault, 2009) that can then lead to biased management advice, overfishing, and failure to meet rebuilding targets. Therefore, developing and testing methods for estimating misreporting is important.

A variety of methods have been proposed for estimating misreported catch in stock assessments. Excluding observed catch completely has been suggested and shown to provide improved estimates under certain circumstances (Cook, 2013). Alternatively, misreported catch can be estimated outside of the assessment model using, for example observer reports and perceived misreporting incentives to scale reported catch (Pitcher *et al.*, 2002). Another approach, which has recently been adapted to state-space models, is to treat catch as a censored observation, where the observed catch is a lower bound of the true catch, and the upper bound is typically set as a multiple of the observed catch (Hammond and Trenkel, 2005; Bousquet *et al.*, 2010; Cadigan, 2016).

Here, we used simulations to test an alternative state-space approach to estimating misreporting catch, and the extent to which misreported catch affects stock assessment output. We focused on the state-space assessment model (SAM, Nielsen and Berg, 2014) due to its current use as the assessment model for several stocks managed by the International Council for the Exploration of the Sea (e.g. ICES, 2017, 2018) and increased consideration for stock assessments in the US Northeast (Northeast Fisheries Science Center, 2018a, b).

Using a simulation model conditioned on the SAM fit for North Sea cod, we tested three formulations of the SAM estimation model, each with increasing complexity. The estimation models either assumed, (i) no misreporting, (ii) misreporting that is fixed over time, or (iii) misreporting that follows a random walk process. We tested these three estimation models against simulation scenarios using each of the assumptions and an additional simulation scenario that assumes uniform random misreporting over time. We examined each estimation model's ability to estimate important variables and commonly used model outputs.

Methods

SAM estimation model structure

We briefly describe the model structure of SAM (for a detailed description, see Berg and Nielsen, 2016). The observation model consists of fishery catch-at-age and survey index-at-age modelled as

$$\log C_{a,y} = \log\left(\frac{F_{a,y}}{Z_{a,y}}(1-\mathrm{e}^{-Z_{a,y}})N_{a,y}\right) + \epsilon^{(C)}_{a,y},$$
$$\log I^{(v)}_{a,y} = \log\left(q^{(v)}_a \mathrm{e}^{-Z_{a,y}\left(\frac{D^{(v)}}{365}\right)} N_{a,y}\right) + \epsilon^{(I(v))}_{a,y},$$

where y is the year index, C is the catch-at-age, F and Z are the fishing and total mortality-at-age, N is the population abundance-at-age, v is the survey index, I is the index catch-at-age, q is the catchability parameter, and D is the day of the year that the survey occurs. The random variability associated with each observation equation is denoted by ϵ and modelled as a multivariate normal random variable with possible covariance across ages.

The process model estimates population abundance-at-age and fishing mortality-at-age as unobserved states. Assuming the oldest age is a plus group and log-recruitment follows a random walk we have

$$\log N_{1,y} = \log N_{1,y-1} + \eta_{1,y},$$
$$\log N_{a,y} = \log N_{a-1,y-1} - F_{a-1,y-1} - M_{a-1,y-1} + \eta_{a,y}, 2 \le a < A,$$
$$\log N_{A,y} = \log\left(N_{A-1,y-1}\mathrm{e}^{-F_{A-1,y-1}-M_{A-1,y-1}} + N_{A,y-1}\mathrm{e}^{-F_{A,y-1}-M_{A,y-1}}\right) + \eta_{A,\ y},$$

where a is the age index with plus group A, y is the year index, N is population abundance, M is the natural mortality rate, and η is a normally distributed random variable that is independent across ages. The fishing mortality, F, is assumed to follow a random walk process on the log scale with possible correlation across ages

$$\log F_{a,y} = \log F_{a,y-1} + \omega_{a,y},\ \ \omega \sim \mathrm{MVN}(0,\ \Sigma).$$

Parameter values are estimated via maximum likelihood using Template Model Builder (TMB, Kristensen *et al.*, 2016). SAM treats both fishing mortality and numbers-at-age as random effects, which are estimated by TMB using the Laplace approximation. Model convergence was checked for each simulation, and it was confirmed that the Hessian of the likelihood function was positive definite allowing for standard errors to be estimated. Simulation code for this study is available at https://github.com/perretti/state-space-tests, and the amended SAM code can be found at https://github.com/perretti/SAM.

Misreporting estimation

The default method for estimating misreported catch in SAM is as a vector of fixed effects on the reported catch, where a value of >1 indicates underreporting. That is,

$$\log(C_{a,y}s_{a,y}) = \log\left(\frac{F_{a,y}}{Z_{a,y}}(1-e^{-Z_{a,y}})N_{a,y}\right) + \epsilon^{(C)}_{a,y},$$

where s is the scale parameter that scales the true catch to the observed catch (C). The scaling can be age specific and year specific where the user must specify which ages and years to estimate.

Although SAM allows for age- and year-specific estimation of misreporting, the fixed-effect approach ignores correlation in misreporting over time and can lead to long run times when attempting to estimate both age- and year-specific scale parameters over many years. To overcome this problem, we also tested whether estimating the scale parameters as a random effect could provide a more accurate method of estimating age-specific, time-varying misreporting. Similar to the estimation approach for N and F, we treated the scaling parameter as a latent random walk process on the log scale.

$$\log(s_{a,y}) = \log(s_{a,y-1}) + \eta_{a,y-1}, \eta \sim N\left(0, \sigma^2\right).$$

For simplicity, we treated the misreporting process as independent across ages, although future work could explore the utility of estimating covariance in misreporting across ages.

We compared the random walk misreporting estimation method to either ignoring misreporting (labelled "no misreporting") or using SAM's built-in feature of estimating it as a fixed effect (labelled "fixed misreporting"). To compare the relatively complex random walk model to a simpler approach, we estimated the fixed misreporting as a single scaling parameter for all ages and years.

Simulation model

Our simulation model is identical in structure to that of SAM except for the generation of misreported catch, which we describe below. We conditioned the simulation model using fixed-effect parameter estimates from the SAM fit to North Sea cod because this model was freely available (https://github.com/fishfollower/SAM/tree/master/testmore/nscod) and North Sea cod has a similar life history to gadoids in the Northwest Atlantic. The North

Sea cod model has six age classes with the oldest age treated as a plus group, and the input data are survey catch-at-age from the International Bottom Trawl Survey, which occurs in the first quarter of the year and the third quarter, and one fishery fleet (Figure 1), where the two oldest age groups within a survey are assumed to have the same catchability. The model assumes observation error, and fishing mortality variability is independent across ages; therefore, the estimation model was also set up to estimate observation error and fishing mortality variability as independent across ages. Observation errors, process errors and misreporting values were re-simulated in each replicate simulation. Model parameter values are listed in Table 1.

We simulated four scenarios that represent increasingly complex forms of misreporting: (i) no misreporting (i.e. $s = 1$), (ii) misreporting that is constant over time and ages, which we refer to as "fixed misreporting", (iii) misreporting that follows a random walk between years and is independent across ages, and (iv) misreporting that follows a uniform random distribution across ages and years. This set-up allowed us to test each estimation model against a correctly specified model, and in the case of the uniform random scenario, we tested how they perform when they are all incorrectly specified. For each variant of the simulation model, 300 datasets were simulated and the estimation models applied in a full factorial design.

We simulated misreporting primarily as underreported catch because that is thought to be the most common form of misreporting. All misreporting occurred in the last 10 years across all ages in the simulations. For the fixed misreporting scenario, we simulated a range of possible underreporting by drawing the scale parameter for each simulation from a uniform distribution bounded between 1.5 and 10. In other words, each simulation had unreported catch that ranged from 33% to 90% of the total catch (Table 2). The random walk misreporting was simulated on the log scale with a 0.2 *SD*, which results in the catch observations being underreported by 67% on average. Similar to the fixed scenario, in the uniform random misreporting scenario, the scale parameter value in each year is bounded between 33% and 90% underreporting. The misreporting estimation models correctly assume misreporting occurs in the last 10 years of the time series.

We note that a property of a random walk process is that the variance of the distribution of possible states in time t increases at $t\sigma^2$ where σ^2 is the variance of the error term. This increase in variance of possible states over time is further magnified in SAM where the random walk is modelled on the log scale. When converted to the natural scale, there is a possibility for the model to exhibit unrealistically high levels of abundance and *F*. Although this did not cause model-convergence problems for North Sea cod, it may be problematic for species with higher estimated process error variances.

Performance metrics

To evaluate model performance, we examined the estimation error of catch, *F*, abundance (*N*), and spawning stock biomass (SSB). We chose the mean absolute percentage error (MAPE) as the error summary statistic because it provides a scale-free estimate of error across ages, which have varying scales (e.g. age-0 abundance is much higher than age-6+ abundance). Absolute percentage error is defined as $\text{abs}(100\ (\hat{x} - x)/x)$, where $\hat{x}$ is the

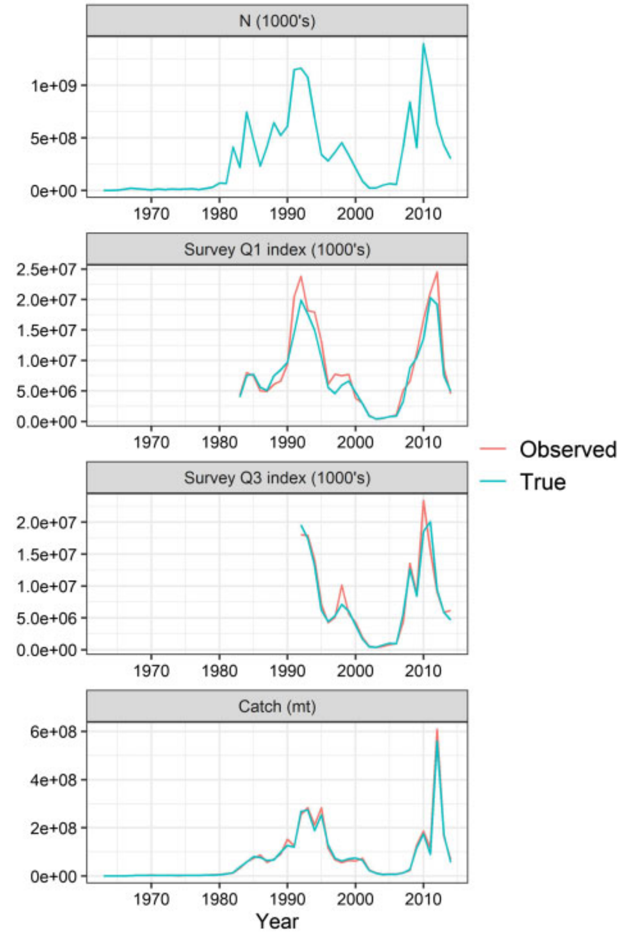

Figure 1. An example simulation of population abundance (*N*), survey catch, and total fishery catch for the no misreporting scenario. Observed is the simulated time series with observation and process errors, and the true line is the time series without observation error.

Table 1. Parameter values used in the simulations.

Parameters	Description	Values
σ_{s1}	Survey Q1 observation error standard deviation	0.28
σ_{s2}	Survey Q3 observation error standard deviation	0.23
σ_c	Fishery catch observation error standard deviation	0.24
σ_{n0}	Age-0 (recruitment) process error standard deviation	1.06
σ_{n1}	Age-1+ process error standard deviation	0.19
σ_F	Fishing mortality process error standard deviation	0.13
q_0^1	Catchability of age-0 fish in survey Q1	0.01
q_1^1	Catchability of age-1 fish in survey Q1	0.04
q_2^1	Catchability of age-2 fish in survey Q1	0.07
q_{3-4}^1	Catchability of age-3 and age-4 fish in survey Q1	0.08
q_0^2	Catchability of age-0 fish in survey Q3	0.03
q_1^2	Catchability of age-1 fish in survey Q3	0.05
q_{2-3}^2	Catchability of age-2 and age-3 fish in survey Q3	0.07

Q1, first quarter of the year; Q3, third quarter.

Table 2. Scale parameter specifications for the misreporting simulation scenarios.

Scenario	Scale parameter(s)
No misreporting	$s_{ay} = 1, \forall a,y$
Fixed	$s_{ay} = c, \forall a,y, c \sim \mathrm{Unif}(1.5, 10)$
Random walk	$\log(s_{a,y}) = \log(s_{a,y-1}) + \eta_{a,y-1}, \ \ \eta \sim N(0, 0.2),$ $s_{a,1} \sim N(1, \ 0.2)$
Uniform random	$s_{a,y} \sim \mathrm{Unif}(1.5, \ 10)$

estimated value and x is the true value. The mean was taken over all replicates and age classes, although age-specific results are similar. We also examined the magnitude of the retrospective error for each estimation model as quantified by the standard modification of Mohn's ρ (Mohn, 1999; Brooks and Legault, 2016) using a 7-year peel. We calculated the Mohn's ρ of average F (ages 4–6+), SSB, and recruitment. We also calculated the frequency with which each model's ρ-adjusted values fell outside the 90% confidence interval of the unadjusted estimates, which is the level at which the retro-adjusted values would be used in a stock assessment in the US Northeast (Brooks and Legault, 2016). The ρ-adjusted value is calculated as $\hat{x}/(1+\rho)$ where $\hat{x}$ is the unadjusted value and ρ is the Mohn's ρ value.

In addition, we examined the accuracy of the F40% catch advice, which is the catch that occurs when F is set such that it reduces SSB per recruit to 40% of unfished SSB per recruit. The accuracy of the catch advice was evaluated by comparing the estimated vs. true catch corresponding to F40% in the year after the terminal year of the simulated data set. The estimated catch is obtained by applying the estimated F corresponding to F40% to the terminal year's estimated exploitable biomass, and this is then compared with the true catch corresponding to F40%. Although the akaike information criterion (AIC) is another possible metric with which to compare models, it was not used here because it has been shown that AIC is not a reliable method of comparing hierarchical models (Gelman *et al.*, 2014).

Results

Overall, the best performing estimation model in each scenario was the one that used the correct assumption for misreporting (i.e. the correctly specified estimation model). However, the worst errors consistently occurred when misreporting was occurring but it was ignored in the estimation model (Figure 2, bottom three rows). In these cases, the MAPE of the no misreporting estimation model often exceeded 50%, and for some variables, it exceeded 100%. On the other hand, when misreporting was not occurring but it was estimated (Figure 2, top row), estimation errors were comparatively low (always <30%), and for the fixed estimation model for F, N, and SSB, it was nearly identical to the error of the no misreporting estimation model. In the scenarios with misreporting, the no misreporting estimation model tended to underestimate catch and F while overestimating N and SSB (Supplementary Figure S1). The correctly specified estimation model is approximately unbiased in each scenario, with mean percentage errors near zero (Supplementary Figure S1), although, for some variables, there is evidence of bias up to 10%, particularly in the initial years of the time series.

In Figure 2, it is sometimes difficult to see differences in error between the two misreporting estimation models; however, the fixed model outperformed the random walk model across all of the variables in the plot when the fixed and random walk estimation models were misspecified (i.e. the no misreporting and uniform scenario, Supplementary Figure S2). In these two scenarios, when averaged over all variables, ages, and years with misreporting, the fixed estimation model MAPE was 14.5% while the random walk model MAPE was 19.2%. The difference in error was largest in the estimation of F, where the fixed model MAPE was 7.9% points lower than the random walk model. In the uniform random scenario, where the increased flexibility of the random walk estimation model was expected to generate more accurate estimates of the estimation variables, we found that the random walk model actually had higher error than the fixed model, particularly when estimating F.

The fixed estimation model was able to estimate the scale parameter (which controls misreporting estimates) more accurately than the random walk model in both the no misreporting and fixed scenarios, while the random walk model was more accurate in the random walk and uniform random scenarios (Figure 3). Although the random walk estimation model more accurately estimated the scale parameters in the uniform random scenario and fixed-effect estimates were similar to the fixed model (Supplementary Figures S3–S6), the random walk model still provided less accurate estimates of the assessment outputs of interest (Figure 4). This is likely due to the increased uncertainty of the random walk scale estimates (represented by larger confidence intervals in Figure 5), which resulted in less accurate estimates of the other random effects (i.e. N and F).

In agreement with previous work, the AIC often does not select the model with the lowest estimation error, particularly in the uniform random scenario where it always selects the random walk estimation model despite the fixed estimation model having lower estimation error (Supplementary Table S1).

Retrospective errors were generally low in all simulation scenarios. As a result, the retro-adjusted estimates were rarely outside the 90% confidence interval of the unadjusted estimates (Table 3), which is the level at which the retro-adjusted values would be used in a stock assessment in the US Northeast (Brooks and Legault, 2016). This is true even for the no misreporting model that often had high estimation errors in the scenarios with misreporting. Overall, the random walk and fixed estimation models tended to have the lowest mean retrospective error, while the no misreporting model usually had the highest retrospective error (Figure 6). Retrospective error among the two misreporting estimation models was often unrelated to estimation error. For example, in the uniform random scenario, the random walk estimation model had nearly identical mean retrospective errors to the fixed estimation model, despite having higher estimation error.

Catch advice error generally followed estimation error patterns, with the correctly specified model always having the lowest catch advice error, and the fixed model outperforming the random walk model in the no misreporting scenario and the uniform random scenario (Figure 7). In general, the catch advice generated by the no misreporting model was negatively biased when misreporting was occurring, with an average percentage error of −14% over all misreporting scenarios. This follows from the tendency of the no misreporting model to underestimate catch, N, and SSB, which results in a negative bias of estimates of catch at F40% (Table 4).

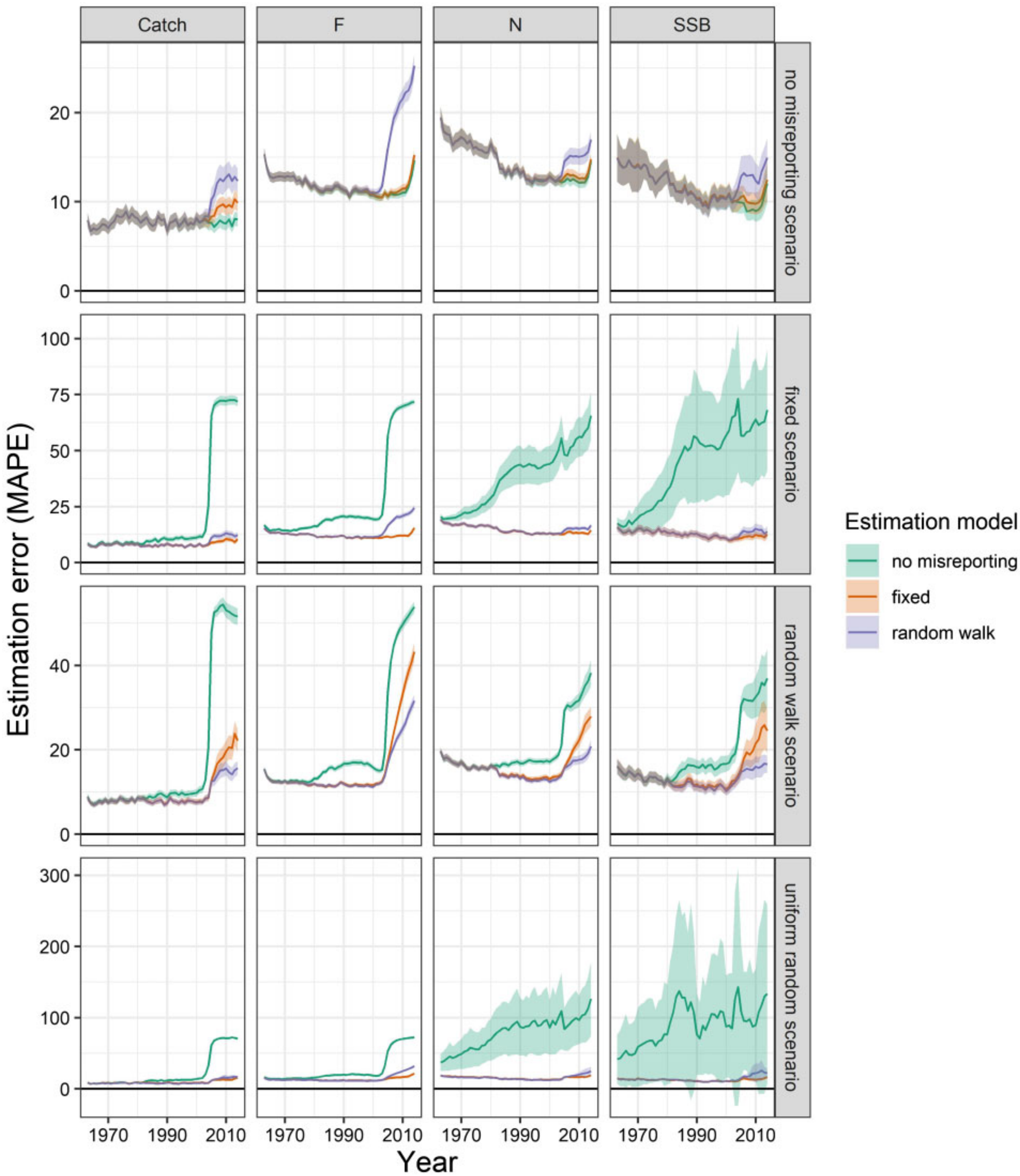

Figure 2. Estimation error results for all estimation models under each scenario. The shaded region is the 95% confidence interval.

Discussion

Overall, we found that the largest estimation errors occurred when misreporting was occurring but it was not estimated. On the other hand, there was a comparatively small increase in error when misreporting was not occurring but it was estimated. This risk asymmetry suggests that it is best to estimate misreporting in an assessment even if there is only a small possibility that it is occurring, as opposed to only estimating misreporting when there is certainty that it is occurring.

For each simulation scenario, the estimation model that was correctly specified always produced the most accurate estimates. However, in reality, an assessment model is never completely correctly specified; therefore, it is noteworthy that in the uniform random scenario, where all estimation models were misspecified, the random walk model produced less accurate estimates of important assessment outputs despite producing more accurate estimates of misreporting. As there was no evidence of parameter confounding in any estimation model (Supplementary Figures

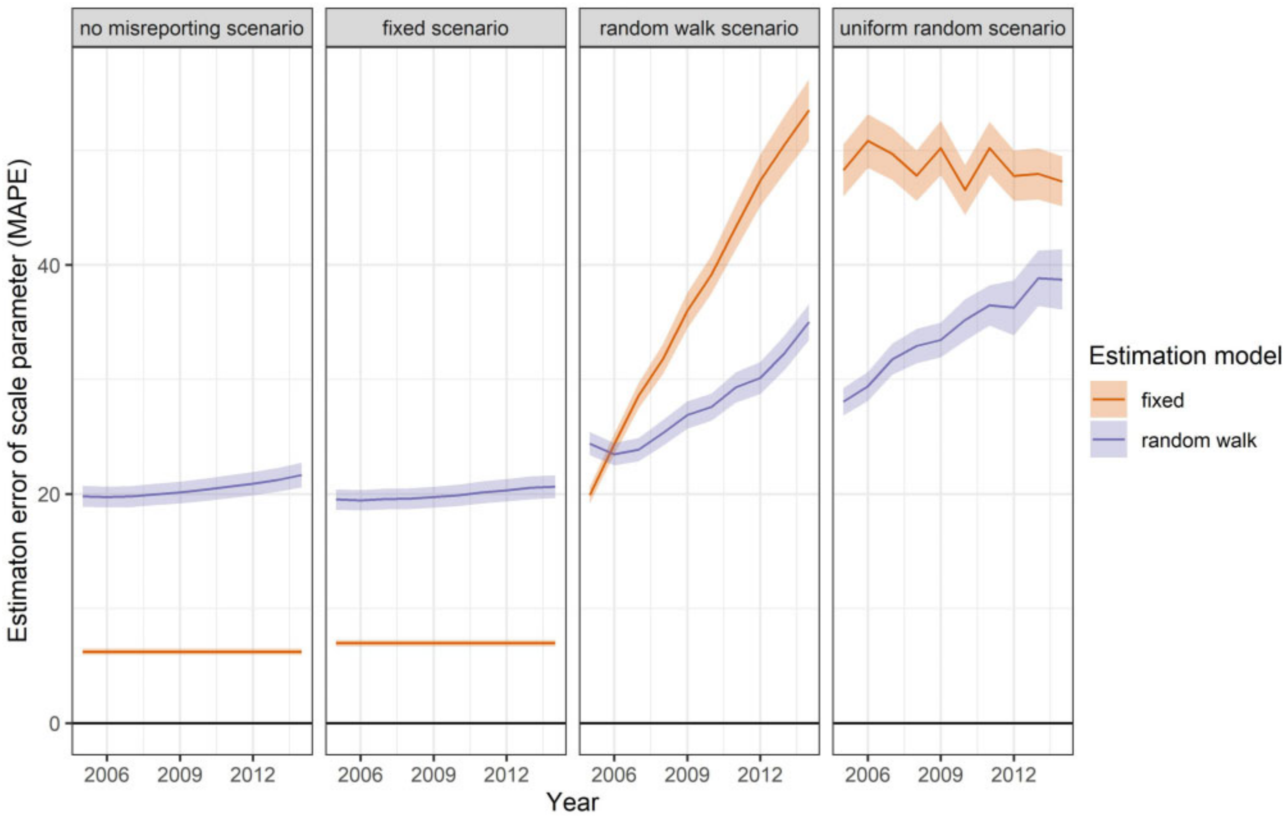

Figure 3. Estimation error of the scale parameter(s) under each simulation scenario. The shaded interval is the 95% confidence interval.

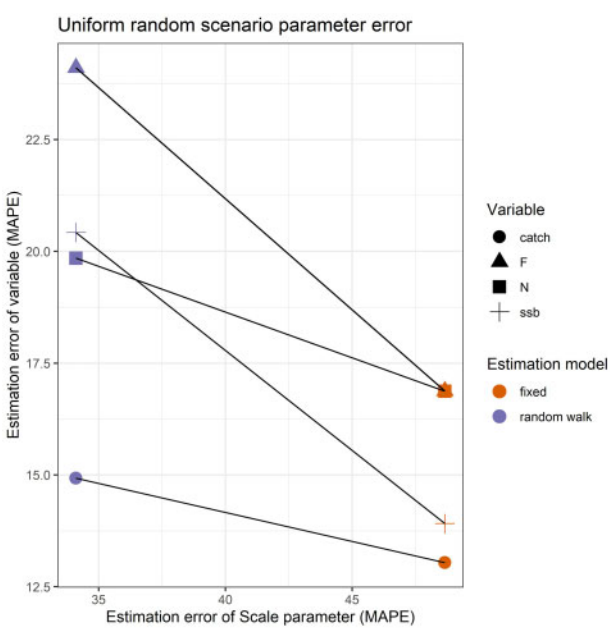

Figure 4. Scale parameter error vs. error of other variables for the uniform random scenario.

S7–S9), this appears to be due to the higher uncertainty associated with the more complicated misreporting function estimated by the random walk model, which leads to less accurate random effect estimates overall. This result of more complicated functions having higher uncertainty is not unique to stock assessment models [see bias-variance trade-off in Hastie *et al.* (2009)]; however, future research could investigate whether this result also occurs when estimating other latent variables in state-space stock assessments (e.g. environmental effects). Future research could also investigate whether estimation models that are intermediate between the fixed model and the random walk model provide more accurate estimates (e.g., a random walk that is fixed across ages, or a fixed model that varies across ages).

The retrospective errors observed here tended to be low compared with what is commonly seen in assessments using statistical catch-at-age models (e.g. Northeast Fisheries Science Center, 2017). This was true even for the no misreporting estimation model despite often having large estimation errors, although the no misreporting model did tend to have the highest retrospective errors in the scenarios with misreporting. Currently used rules in the US Northeast determining whether retrospective adjustments should be performed would need to be revised when using the SAMs tested here because the retrospective error was rarely large enough to trigger an adjustment, despite occasionally large estimation errors. Furthermore, there was not always a clear relationship between the magnitude of the retrospective error and the magnitude of the estimation error. It appears that the added flexibility of the latent process in the random walk model was able to keep retrospective errors low, even though estimation error was often higher than the fixed model. This suggests that state-space models with more complicated latent processes will tend to have lower retrospective errors regardless of estimation

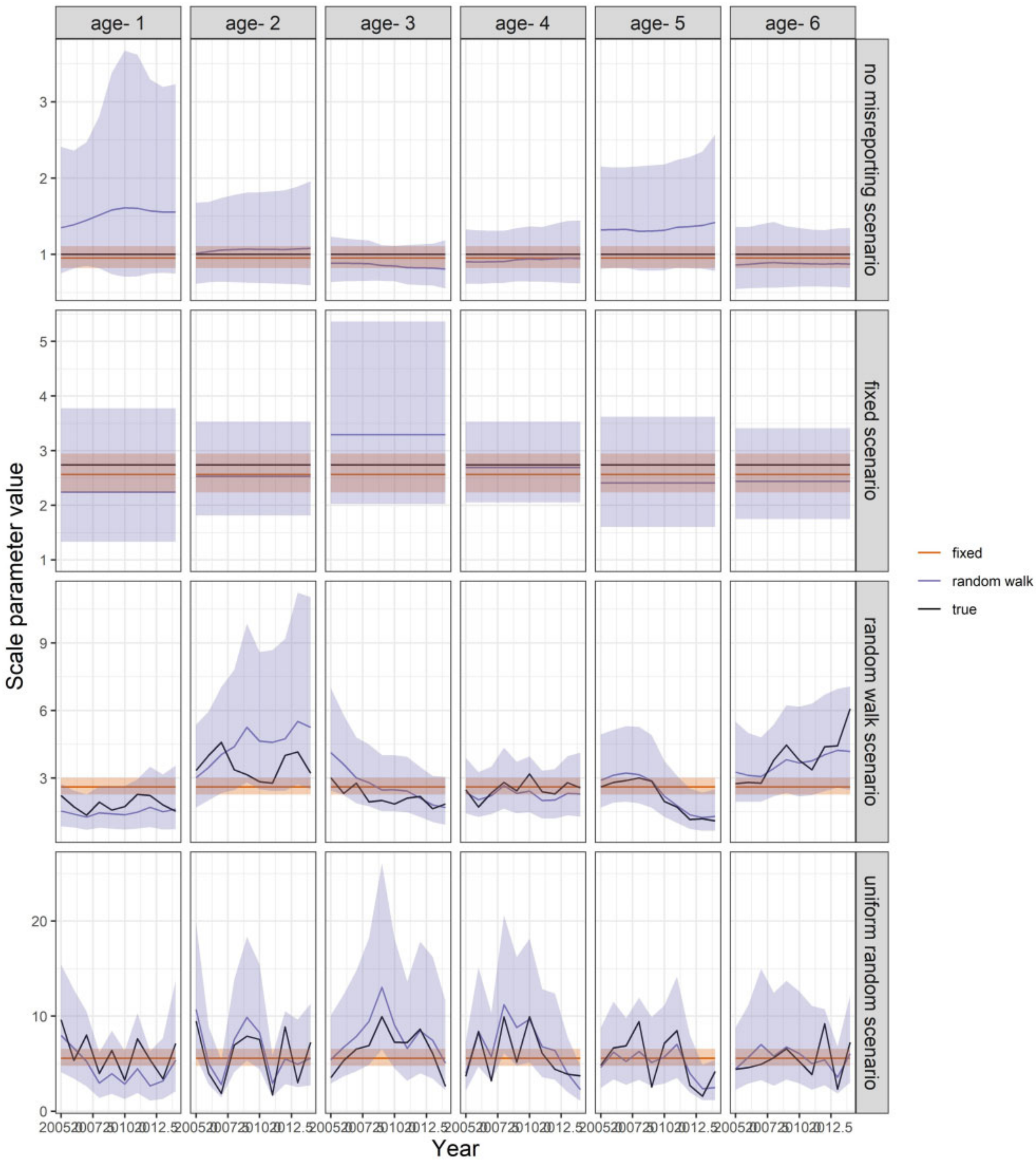

Figure 5. An example scale simulation and fit for each scenario. The shaded interval is the 95% confidence interval of the fit.

accuracy, although a study focused specifically on this question is warranted.

Using SAM as a simulation model (also known as an operating model) is complicated by the fact that the expected value of both *F* and recruitment increases with time (as noted in Nielsen and Berg, 2019). This can lead to unrealistically high levels of *F* and *N* in some simulations. Although this can be mitigated by simulating conditional on a single set of estimated random effects, as is done by default in SAM, this is not an option if one is interested in simulating all sources of variability (including process error).

Table 3. Percentage of simulations that would require a retrospective adjustment based on whether the retro-adjusted estimate falls outside the 90% confidence interval of the estimate.

Simulation scenario	No misreporting estimation model	Fixed estimation model	Random walk estimation model
No misreporting	0.0	0.0	1.2
Fixed	9.5	0.0	1.4
Random walk	11.1	2.1	1.8
Uniform random	8.7	0.7	0.5

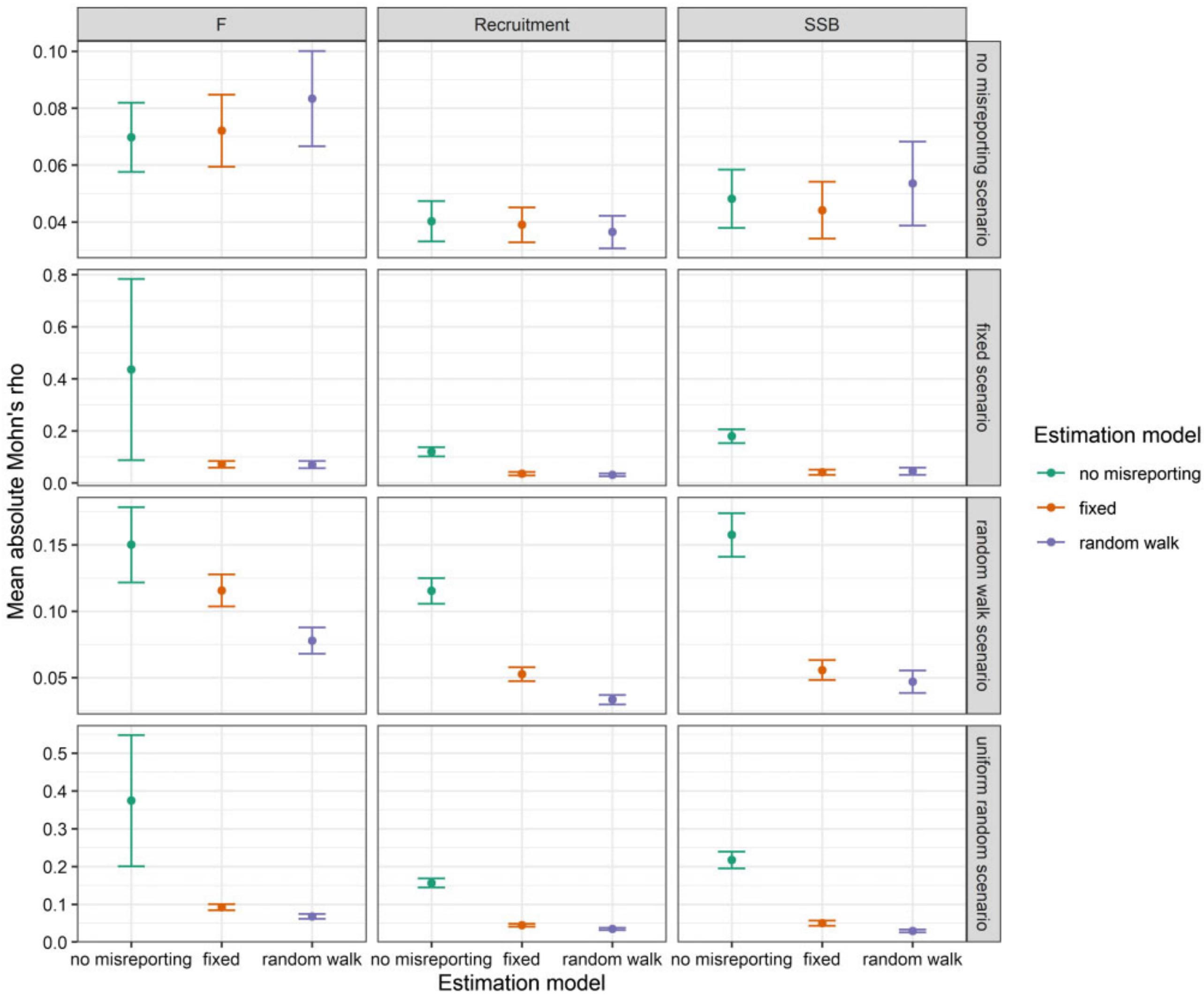

Figure 6. Mean retrospective error as measured by Mohn's ρ. Error bars are the 95% confidence interval.

Although the high levels of *F* and recruitment did not cause numerical problems in our simulations, it may be problematic for species in which the estimated process error is higher, or if a researcher is interested in simulating all sources of error and constraining the simulations to biologically reasonable levels. A possible solution to this problem would be to implement a mean-reverting random effect process in SAM such as an autoregressive process.

We note that the misreporting estimation models in this study correctly assumed that the misreporting only occurs in the final 10 years of the time series. In practice, the timeframe of misreporting will not be known and estimating misreporting in all years is not possible due to confounding between the scale parameter(s), *N* and *F*. Future research should examine the performance of different estimation models when the timeframe of misreporting is incorrectly specified, as well as methods for estimating the misreporting timeframe in real time series.

We found that implementing random walk estimation of misreporting did not improve performance relative to the default SAM configuration of fixed misreporting estimation in the uniform random misreporting scenario. Although the uniform random scenario is a specific form of misspecification, it suggests that, under certain circumstances, the increased uncertainty associated with estimating the more flexible, but more highly parametrized, random walk misreporting latent process may lead to less accurate estimates than the simple fixed approach. Future work should investigate whether this result is specific to the uniform random scenario or if it occurs in other misspecified scenarios. One way to potentially address this problem is to reduce the flexibility of the random walk model by estimating the covariance in misreporting across ages. This would mirror the approach already implemented in SAM for estimating *F* and observation error. Estimating the covariance across ages should result in the partial pooling of the scale parameter estimates, which would reduce the variability in the estimates overall. However, this would introduce additional parameters to be estimated in the form of a covariance matrix, potentially increasing uncertainty overall. Future research could investigate this approach directly, as well as the generality of the result that more flexible latent processes can lead to worse estimates of key output variables.

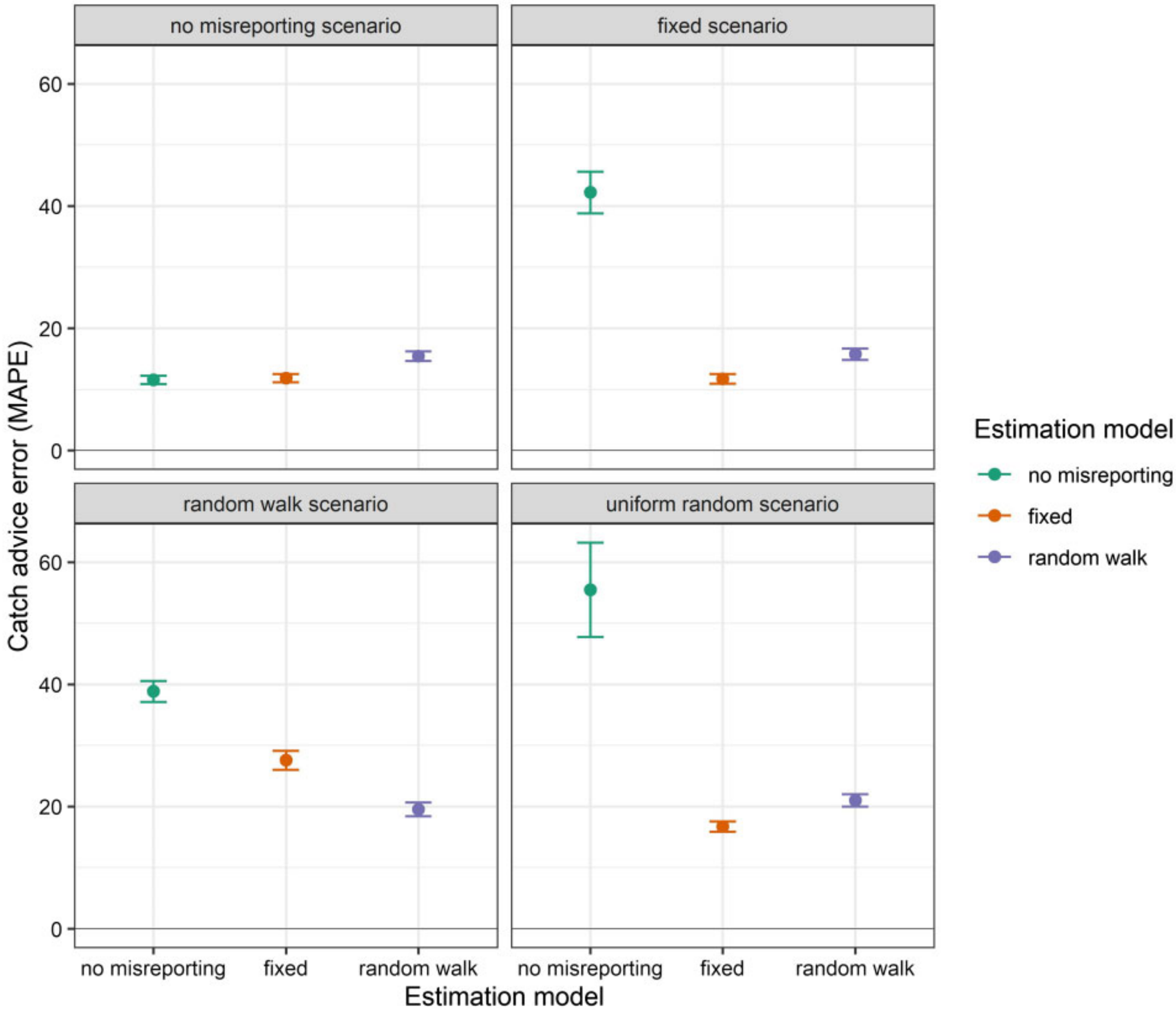

Figure 7. Catch advice error when averaged over all replicates for each scenario. Error bars represent the 95% confidence interval.

Table 4. Bias of catch advice for each estimation model as quantified by mean percentage error (see Figure 7 for mean *absolute* percentage error).

Simulation scenario	No misreporting estimation model	Fixed estimation model	Random walk estimation model
No misreporting	−1.4	−0.5	−0.9
Fixed	−12.9	−0.3	−1.4
Random walk	−24.0	1.3	−3.6
Uniform random	−18.0	0.7	−0.8

Supplementary data

Supplementary material is available at the *ICESJMS* online version of the manuscript.

References

Bellanger, M., Holland, D. S., Anderson, C. M., and Guyader, O. 2019. Incentive effect of joint and several liability in fishery cooperatives on regulatory compliance. Fish and Fisheries, 20: 715–728.

Berg, C. W., and Nielsen, A. 2016. Accounting for correlated observations in an age-based state-space stock assessment model. ICES Journal of Marine Science, 73: 1788–1797.

Bousquet, N., Cadigan, N., Duchesne, T., and Rivest, L.-P. 2010. Detecting and correcting underreported catches in fish stock assessment: trial of a new method. Canadian Journal of Fisheries and Aquatic Sciences, 67: 1247–1261.

Brooks, E. N., and Legault, C. M. 2016. Retrospective forecasting—evaluating performance of stock projections for New England groundfish stocks. Canadian Journal of Fisheries and Aquatic Sciences, 73: 935–950.

Cadigan, N. G. 2016. A state-space stock assessment model for northern cod, including under-reported catches and variable natural mortality rates. Canadian Journal of Fisheries and Aquatic Sciences, 73: 296–308.

Cook, R. M. 2013. A fish stock assessment model using survey data when estimates of catch are unreliable. Fisheries Research, 143: 1–11.

Gelman, A., Hwang, J., and Vehtari, A. 2014. Understanding predictive information criteria for Bayesian models. Statistics and Computing, 24: 997–1016.

Gudmundsson, G. 1994. Time series analysis of catch-at-age observations. Journal of the Royal Statistical Society: Series C (Applied Statistics), 43: 117–126.

Hammond, T. R., and Trenkel, V. M. 2005. Censored catch data in fisheries stock assessment. ICES Journal of Marine Science, 62: 1118–1130.

Hastie, T., Tibshirani, R., and Friedman, J. 2009. The Elements of Statistical Learning. Springer, New York.

ICES. 2014. Report of the Benchmark Workshop on Pelagic Stocks. ICES Document CM 2014/ACOM:43. ICES, Copenhagen, Denmark.

ICES. 2017. Report of the Arctic Fisheries Working Group (AFWG). Technical Report ICES Document CM 2017/ACOM:06. ICES, Copenhagen, Denmark.

ICES. 2018. Report of the Working Group on Widely Distributed Stocks (WGWIDE). Technical Report ICES Document CM 2018/ACOM:23. ICES, Copenhagen, Denmark.

Kristensen, K., Nielsen, A., Berg, C. W., Skaug, H., and Bell, B. 2016. TMB: automatic differentiation and Laplace approximation. Journal of Statistical Software, 70:

Legault, C. M. 2009. Report of the Retrospective Working Group 09-01. NOAA Fisheries, Northeast Fisheries Science Center, Woods Hole, MA.

Mendelssohn, R. 1988. Some problems in estimating population sizes from catch-at-age data. Fishery Bulletin, 86: 617–630.

Miller, T. J., and Hyun, S.-Y. 2018. Evaluating evidence for alternative natural mortality and process error assumptions using a state-space, age-structured assessment model. Canadian Journal of Fisheries and Aquatic Sciences, 75: 691–703.

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

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

Nielsen, A., and Berg, C. W. 2019. Response to: comments on incongruous formulations in the SAM (state-space assessment model) model and consequences for fish stock assessment. Fisheries Research, 210: 228–229.

Northeast Fisheries Science Center. 2017. Operational Assessment of 19 Northeast Groundfish Stock, Updated Through 2016. Northeast Fisheries Science Center Reference Document 17–17. US Department of Commerce.

Northeast Fisheries Science Center. 2018a. 64th Northeast Regional Stock Assessment Workshop (64th SAW) Assessment Report. Northeast Fisheries Science Center Reference Document 18–06. US Department of Commerce.

Northeast Fisheries Science Center. 2018b. 65th Northeast Regional Stock Assessment Workshop (65th SAW) Assessment Report. Northeast Fisheries Science Center Reference Document 18–11. US Department of Commerce.

Palmer, M. 2017. Vessel Trip Reports Catch-area Reporting Errors: Potential Impacts on the Monitoring and Management of the Northeast United States Groundfish Resource 17–02. NOAA Fisheries, Northeast Fisheries Science Center, Woods Hole, MA.

Pitcher, T. J., Watson, R., Forrest, R., Valtýsson, H. Þ., and Guénette, S. 2002. Estimating illegal and unreported catches from marine ecosystems: a basis for change. Fish and Fisheries, 3: 317–339.

Sullivan, P. J. 1992. A Kalman filter approach to catch-at-length analysis. Biometrics, 48: 237–257.

Thorson, J. T., and Minto, C. 2015. Mixed effects: a unifying framework for statistical modelling in fisheries biology. ICES Journal of Marine Science, 72: 1245–1256.

United States of America, v. Carlos A. Rafael. 2017. 282 U.S. 407, United States District Court, D. Massachusetts.

Handling editor: Ernesto Jardim