Fisheries and Oceans Canada / Pêches et Océans Canada

NOAA FISHERIES
NATIONAL MARINE FISHERIES SERVICE

CERT

Comité d'évaluation des ressources transfrontalières

Document de travail 2017/06

Ne pas citer sans
autorisation des auteurs

TRAC

Transboundary Resources Assessment Committee

Working Paper 2017/06

Not to be cited without
permission of the authors

Sensitivity analysis of terminal year versus three year moving average estimation of survey biomass using an empirical model and a state-space model

Charles T. Perretti, Michael J. Fogarty, Christopher M. Legault

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

Canada

TABLE OF CONTENTS

ABSTRACT

A simulation study was performed to compare the performance of the previously used empirical assessment model to a state space model for estimating survey biomass using either the terminal year estimate or a three year moving average. An age-structured model of Georges Bank yellowtail flounder was used to evaluate the two models under three future scenarios: no change in biomass, increasing biomass, and decreasing biomass. Each scenario was driven by one of two drivers: recruitment, or fishing. In all scenarios and for both models, the terminal year estimate was found to be more accurate than the three-year average. In addition, the state space model provided more accurate biomass estimates than the empirical model in four out of the five scenarios. The confidence interval for the state space model was generally accurate, but was biased towards underestimating the true confidence interval when biomass was trending up, and overestimating it when biomass was trending down. Similarly, the state space model exhibited a positive retrospective pattern in the declining biomass scenarios, and a negative retrospective pattern in the increasing biomass scenarios. Overall, the largest gains in accuracy were achieved by using the terminal year estimate rather than a three year smooth, and smaller gains were achieved by using the state space model rather than the empirical model.

RÉSUMÉ

INTRODUCTION

The most recent assessment for Georges Bank Yellowtail Flounder (GBYT) used the average biomass of the three bottom trawl surveys (NEFSC Spring, NEFSC Fall, and DFO spring) as the estimated stock biomass. We compare the performance of this model to a multivariate, random walk state-space model. We also compare the estimation accuracy of each model when using only the terminal year as the estimated biomass versus using an average of the final three years of the time series. We use an age-structured simulation model of GBYT to simulate trawl survey time series with similar characteristics to the real time series. We then compare the two models and two estimates under three future scenarios: increasing biomass, decreasing biomass, and no change. We simulate two drivers of the scenarios: recruitment and fishing. For the state space model, we examine the confidence interval coverage and the magnitude of the retrospective pattern.

METHODS

SIMULATION MODEL

The population was modeled as an age-structured model with ten ages including a plus group. The simulation was initialized using a 100 year burn-in period which ends in year 1972. From 1973 to 2013 the model uses recruitment and fishing mortality as estimated in the 2014 GBYT assessment (Legault et al. 2014).

Recruitment was modeled as a log-normal random variable with mean equal to the estimated recruitment in the stock assessment with standard deviation 0.1 (σ_p). The fishing mortality of the burn-in period was set to the fishing mortality estimated in the first year of the assessment (1973). Similarly, the mean recruitment of the burn-in was set to the estimated recruitment in 1973.

$$N_{t,1} \sim lnN(\mu_t, \sigma_r)$$

$$\mu_t = \log(r_t) - 0.5{\sigma_r}^2$$

$$N_{t+1,2:9} = \mathrm{N}_{t,1:8}\mathrm{e}^{-(\mathrm{f}_{\mathrm{t},1:8}\mathrm{s}_{1:8}+m)}$$

$$N_{t+1,10+} = \mathrm{N}_{t,9}\mathrm{e}^{-(\mathrm{f_t s_9}+m)} + \mathrm{N}_{t,10+}\mathrm{e}^{-(\mathrm{f_t s_{10+}}+m)}$$

$$B_{t,a} = w_a N_{t,a}$$

Where $N_{t,1}$ is the abundance of age-1 fish at time *t*, and is drawn from a lognormal distribution with standard deviation σ_r and mean μ_t. Mean μ_t is set to the stock assessment estimate for year *t*. The abundance of fish age 2 to 9 in year *t+1* ($N_{t+1,2:9}$) is the fraction of individuals that survive after fishing mortality (f_t) and natural mortality (*m*). Fishing selectivity on age *a* is denoted by s_a. The plus group abundance in year t+1 ($N_{t+1,10+}$) is the sum of age-9 fish that survive through year *t* plus the abundance of surviving age 10+ fish. Biomass is calculated as abundance at age multiplied by weight at age (w_a). All parameter values are listed in Table 1.

The NMFS spring, NMFS fall and DFO spring bottom trawl surveys were simulated. For simplicity, all ages were assumed to be fully selected and all surveys to occur on January 1 to allow the observations to be directly related to the true biomass.

Survey observation error was simulated as independent log-normal error with standard deviation 0.2 (σ_s, Table 1). For each simulation, one year was randomly selected as an outlier year in which the standard deviation of the log-normal error was set to $4\sigma_s$ for all surveys. This more closely mimics the actual time series for Georges Bank yellowtail flounder than assuming a constant standard deviation for all years and creates a model mis-specification relative to the multivariate random walk state-space model.

After the burn-in and the simulation period ending in 2013, an additional five year scenario period was simulated to examine the performance of the estimation models under three different trends, and two drivers. The three trends were: 1) no change in biomass, 2) increasing biomass, and 3) decreasing biomass. The two drivers of the trends were: 1) change in recruitment, and 2) change in fishing. The no change trend was simulated by keeping both recruitment and fishing fixed at the 2013 value. Increasing biomass was simulated by either increasing the mean of the log-normal recruitment by 50% each year, or by decreasing the fishing mortality rate by 50% each year. Similarly, the decreasing biomass scenario was simulated by either decreasing the mean recruitment by 50% each year, or increasing the fishing mortality by 50% each year. There were a total of five scenarios: no change, and two trends driven by two drivers. Each scenario was simulated 1000 times to estimate the performance of the estimation models.

ESTIMATION MODELS

GBYT biomass in the most recent assessment was estimated as the mean of the three survey estimates, which we label the “empirical” model. We compare the empirical model to a multivariate random walk state-space model, which we label the “state space” model. A state-space model includes both a model of the observation process and the underlying population process. Here, the biomass observations are modeled as three independent, log-normally distributed random variables with mean equal to the estimated (unobserved) true biomass, while the population process is modeled as a univariate random walk on a log-scale. Under this model, changes in state from one time point to the next are normally distributed on a log-scale and uncorrelated over time.

$$B_{t+1} = B_t e^{\varepsilon_t}$$

$$\boldsymbol{O}_t = B_t e^{\boldsymbol{\omega}_t}$$

$$\varepsilon \sim N(0, \sigma_p)$$

$$\boldsymbol{\omega} \sim MVN(0, I\boldsymbol{\sigma_o})$$

Where B_t is the true biomass at time t, and ε is a normally distributed random variable, $\boldsymbol{O}_t$ is the vector of observations at time t (i.e., the three surveys), $\boldsymbol{\omega}$ is a multivariate normal random variable with zero off-diagonal covariance, and $\boldsymbol{\sigma_o}$ is the observation error variances. Intuitively, the observations are modeled as the product of random population fluctuations and random observation errors where the errors are log-normal for both processes. Standard errors of the parameter and state estimates were calculated as the inverse of the negative Hessian matrix of the likelihood function. The model was fit using Template Model Builder (Kristensen et al., 2016) in the *R* statistical language, and all code is available at https://github.com/perretti/index_methods.

PERFORMANCE EVALUATION

The likelihood function of the state-space model uses observations both before and after a given year to estimate that year’s biomass. Therefore, the terminal year, which has no observations after it, will have higher estimation error than non-terminal years. Estimation error statistics that include non-terminal years will therefore have lower error than those using only terminal years. In contrast, the empirical model uses only a single year’s observations to estimate that year's biomass, therefore its error is unaffected by terminal years. To make a fair comparison between the two models, we evaluate their performance only on the terminal year of the simulated time series. Error was measured on the five-year scenario period, where the state space model was re-fit for each of the five years. This allows for five terminal year comparisons per simulation, and an estimate of the retrospective pattern of the state space model.

Estimation error was quantified as the mean absolute error (MAE) of the terminal year fitted biomass compared to the true biomass. We also calculated the MAE of each model when using the arithmetic mean of the final three years as the estimated biomass. This allows us to compare a temporally smoothed estimate (three year mean) to an unsmoothed estimate (terminal year only) for each model.

For the state space model we estimated the accuracy of the confidence interval for the terminal year using a decile coverage histogram. A decile histogram shows how often the true biomass falls within each of the ten deciles of the confidence interval distribution, where perfect performance is represented by having 10% of the observations fall in each of the ten deciles. We also estimated the retrospective pattern of the state-space model using Mohn's rho statistic (Mohn 1999 as modified in Miller and Legault 2017).

RESULTS AND DISCUSSION

The scale and trajectory of the simulated time series broadly matches that of the real time series (Figures 1, 2a). The no change scenario leads to an equilibrium biomass of approximately 4000 metric tons (mt) (Figure 2a). In both the increasing and decreasing biomass scenario, the change in biomass is larger when driven by recruitment as opposed to fishing. In the increasing biomass scenario, biomass reaches approximately 10,000mt in 2018 when driven by increasing recruitment, compared to approximately 5,000mt when driven by decreasing fishing (Figures 2b, 2c). Similarly, in the decreasing biomass scenario, biomass reaches approximately 1,000mt when driven by decreasing recruitment, and 2,000mt when driven by increasing fishing (Figures 2d, 2e).

In all scenarios and for both models, the terminal year estimate was more accurate that the three year mean (Table 2 and Figure 3). The terminal year estimate outperformed the three year mean most strongly for the state space model, which had 54.0% lower error when using the terminal year, while the empirical model had 47.9% lower error when using the terminal year.

The state space terminal year model was more accurate than the empirical terminal year model in all scenarios except for the decreasing biomass driven by recruitment scenario. MAE for each model and scenario is given in Table 2 and Figure 3. The state space model outperformed the empirical model by the largest margin in the no change scenario (11.7% improvement), and the fishing-driven trend scenarios (7.5% and 14.0% in the decreasing and increasing trends, respectively). The improvement was smallest in the increasing biomass driven by recruitment scenario (3.3%), and negative in the decreasing biomass driven by recruitment scenario (-5.0%). Overall, there was a 6.3% average percent error reduction when using the state space model with the terminal year estimate rather than the empirical model with the terminal year estimate.

The decile coverage histograms show that the state space model confidence interval was most accurate in the scenarios where the trend was the weakest (i.e., the no change scenario and the fishing driven scenarios). In general, the model tended to underestimate the true trend, and this lead to a downward bias in the confidence interval during increasing trends, and an upward bias during decreasing trends. For example, in the no change scenario, where biomass is gradually declining, the confidence interval slightly overestimates the true state, with more than 10% of the observations in each of the low deciles and less than 10% in the high deciles (Figure 4). In the increasing biomass driven by fishing scenario, where biomass does not show a strong trend, the confidence interval showed the most accurate coverage across all deciles. In the increasing biomass driven by recruitment the coverage was biased downward, and inversely in both of the decreasing scenarios the model was biased upward. This underestimation of trend would be expected to be even stronger in the three year mean estimates, but was not computed due to difficulty in estimating an appropriate uncertainty estimate for the three year mean of the state-space model.

The Mohn’s rho calculations for the state space model show a generally weak retrospective pattern (usually less than 10%) matching the pattern described by the decile coverage plots (Table 3, Figure 5). A positive retrospective pattern, in which estimates for a given year decline as additional years are added, was found in the decreasing biomass scenarios, while a negative retrospective pattern was found in the increasing biomass scenarios. In the no change scenario, which has a slight downward biomass trend, Mohn’s rho was slightly positive. Also following the decile results, the retrospective pattern was strongest in the recruitment driven scenarios (which exhibited the largest trend) and weakest in the fishing driven scenarios.

Overall, we find that the largest gains in accuracy are achieved by using the terminal year estimate rather than the three year moving average. Smaller gains are achieved by using the state space model rather than the empirical model, with the state space model outperforming the empirical model in four out of five scenarios.

LITERATURE CITED

Kristensen K, Nielsen A, Berg C.W., Skaug H, Bell B.M. 2016. TMB: Automatic differentiation and Laplace approximation. Journal of Statistical Softwware 70: 1−21.

Legault, C.M., L. Alade, W.E. Gross, and H.H. Stone. 2014. Stock Assessment of Georges Bank Yellowtail Flounder for 2014. TRAC Ref. Doc. 2014/01. 214 p.

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

Mohn, R. 1999. The retrospective problem in sequential population analysis: an investigation using cod fishery and simulated data. ICES Journal of Marine Science 56(4): 473-488.

TABLES

Table 1. Values and descriptions of the simulation parameters.

Parameter	Value	Description
σ_r	0.1	Process error standard deviation.
σ_s	0.2	Observation error standard deviation.
s_a	1, 1, 1, …, 1	Selectivity-at-age
w_a	0.148, 0.317, 0.453, 0.588, 0.724, 0.921, 0.921, …, 0.921	Weight-at-age (kg)
m	0.4	Natural mortality rate

Table 2. MAE for each model. The lowest MAE for each scenario is bolded.

Trend	Driver	Empirical model 3yr mean	State space model 3yr mean	Empirical model terminal year	State space model terminal year
No change	NA	0.489	0.485	0.328	**0.290**
Increasing biomass	Recruitment	1.212	1.360	0.547	**0.529**
Increasing biomass	Fishing	0.509	0.528	0.397	**0.342**
Decreasing biomass	Recruitment	0.819	0.865	**0.212**	0.222
Decreasing biomass	Fishing	0.616	0.628	0.275	**0.254**

Table 3. Mean Mohn's rho of the state space model in each scenario.

Trend	Driver	Mohn's rho
No change	NA	0.006
Increasing biomass	Recruitment	-0.040
Increasing biomass	Fishing	-0.011
Decreasing biomass	Recruitment	0.066
Decreasing biomass	Fishing	0.027

FIGURES

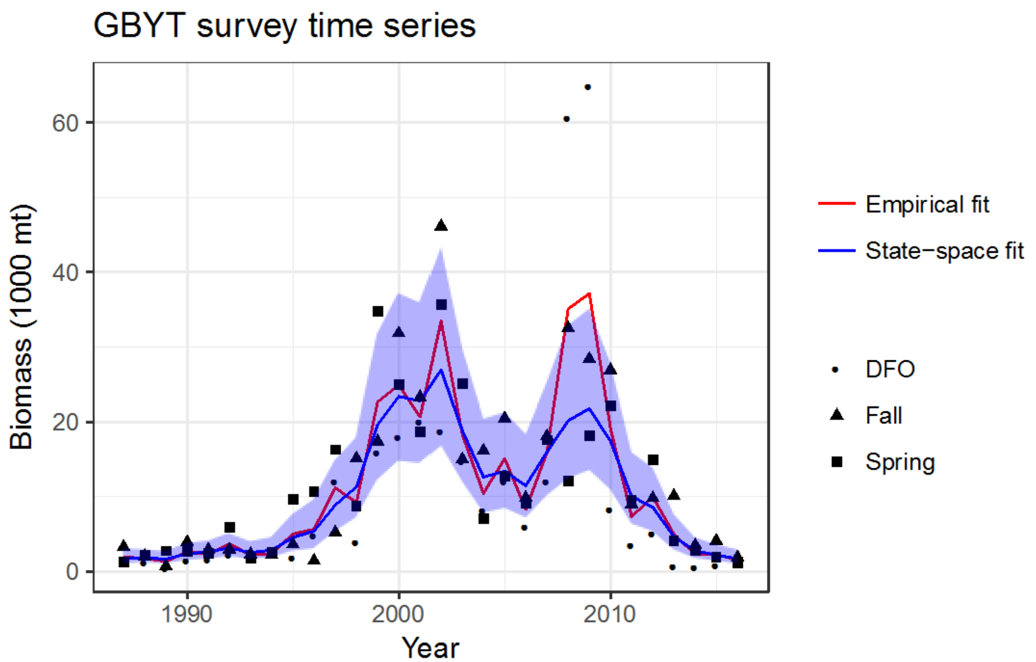

Figure 1. GBYT survey biomass with fitted empirical model and state space model.

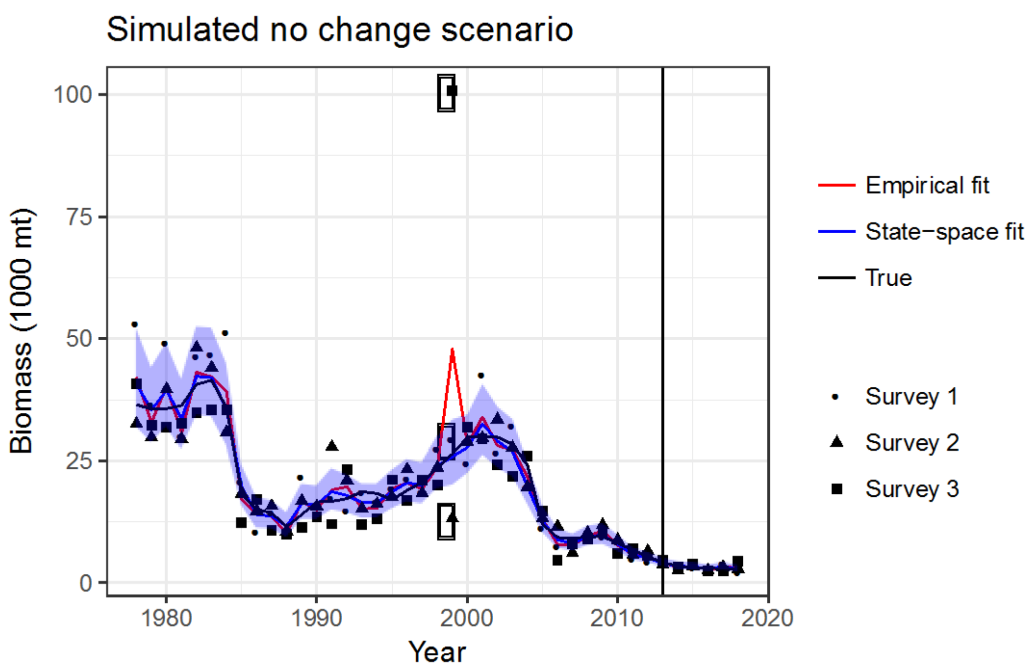

Figure 2a. Simulated survey biomass with model fits for the no change scenario. Vertical line denotes the start of the scenario, and outlier surveys are circled.

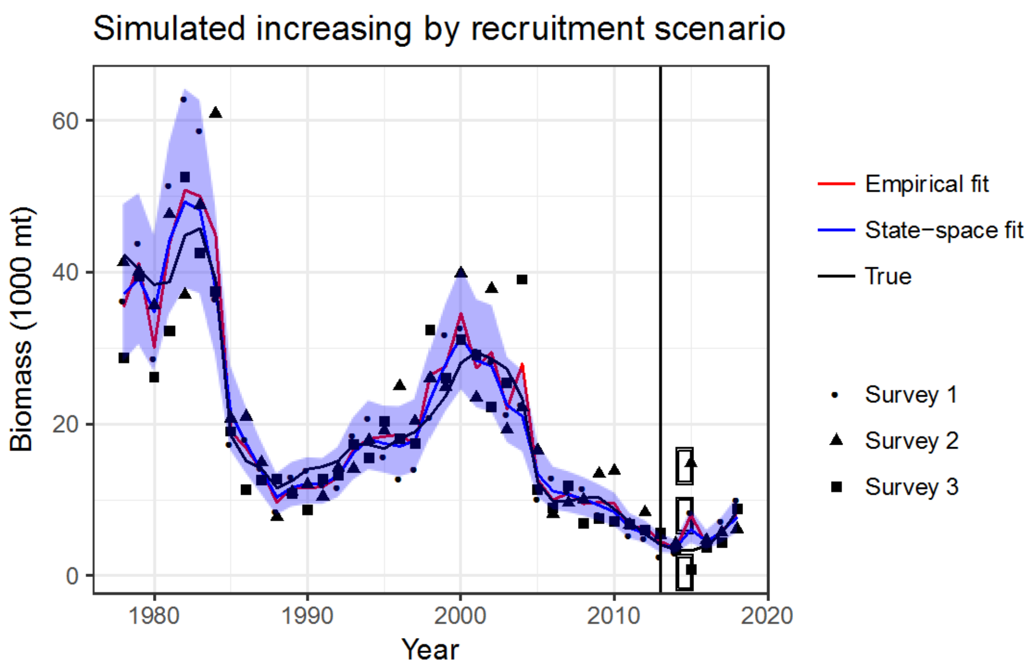

Figure 2b. Simulated survey biomass with model fits for the increasing by recruitment scenario. Vertical line denotes the start of the scenario, and outlier surveys are circled.

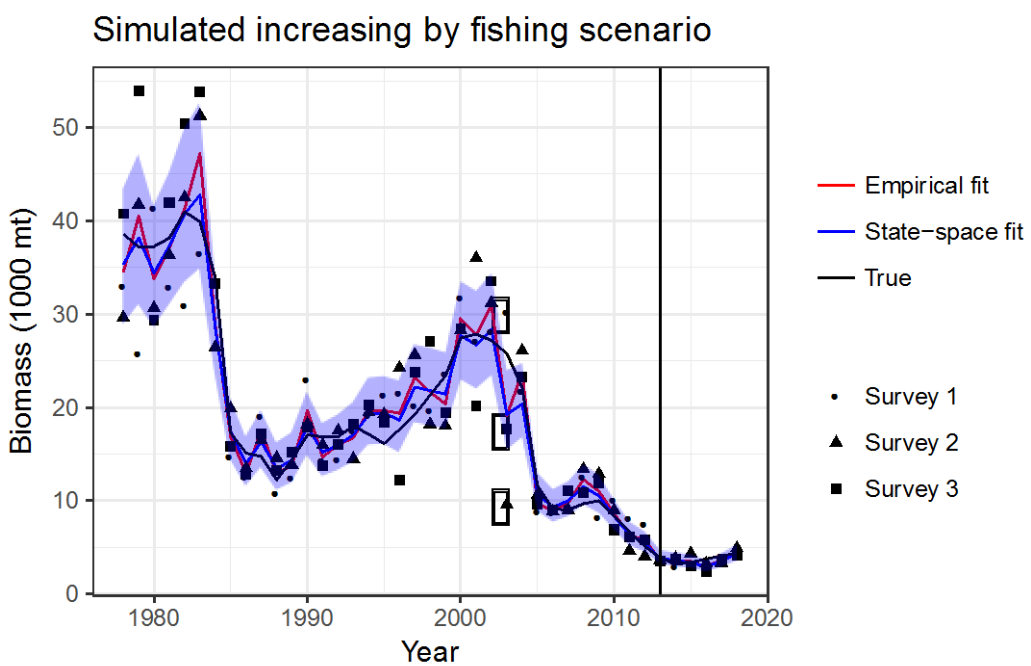

Figure 2c. Simulated survey biomass with model fits for the increasing by fishing scenario. Vertical line denotes the start of the scenario, and outlier surveys are circled.

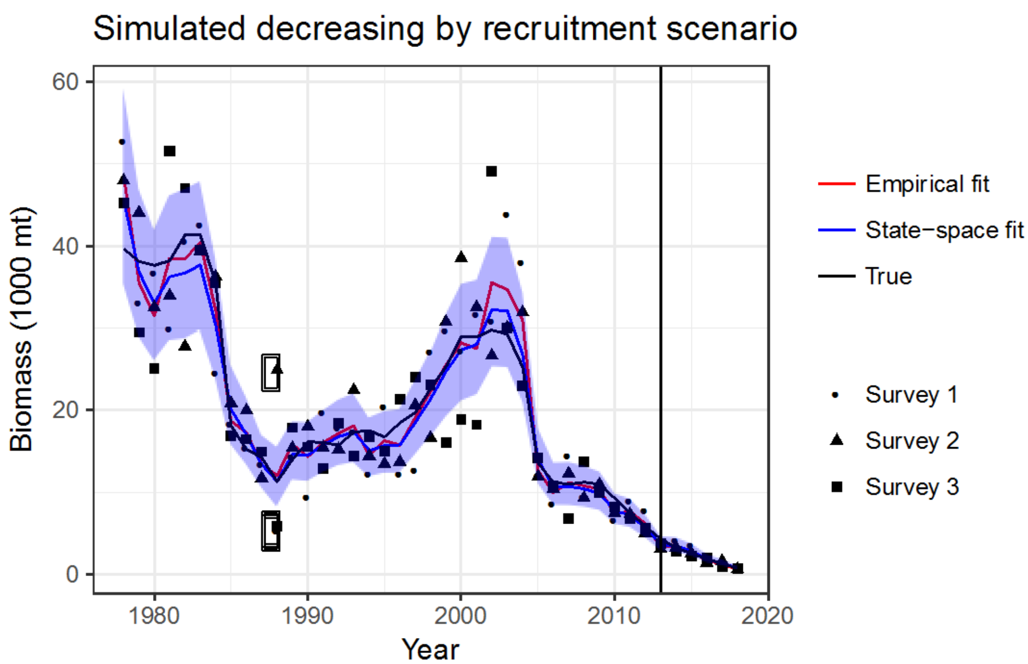

Figure 2d. Simulated survey biomass with model fits for the decreasing by recruitment scenario. Vertical line denotes the start of the scenario, and outlier surveys are circled.

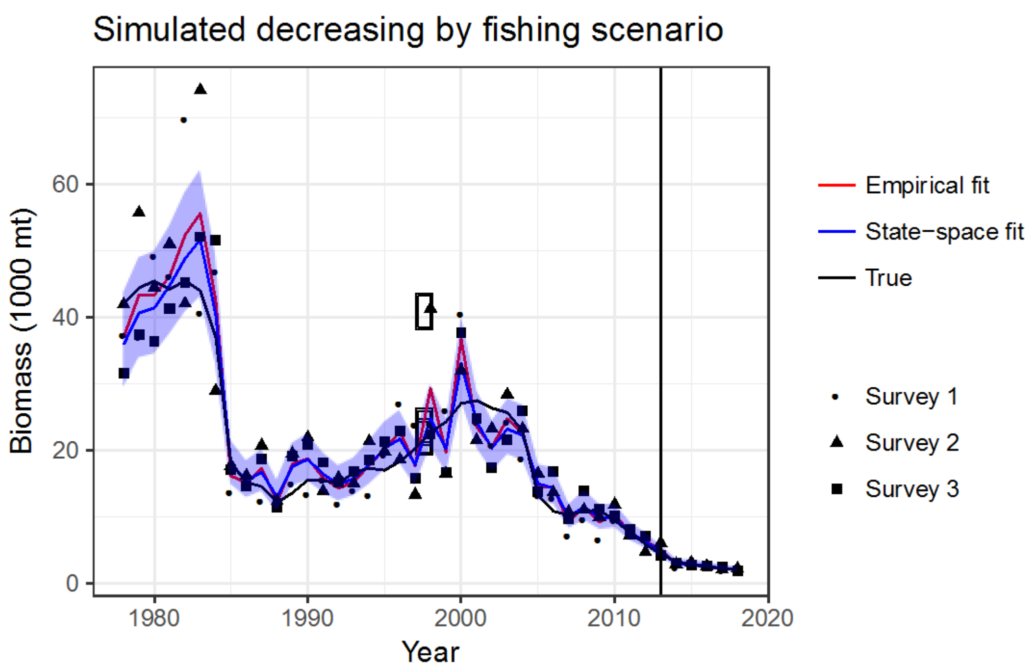

Figure 2e. Simulated survey biomass with model fits for the decreasing by fishing scenario. Vertical line denotes the start of the scenario, and outlier surveys are circled.

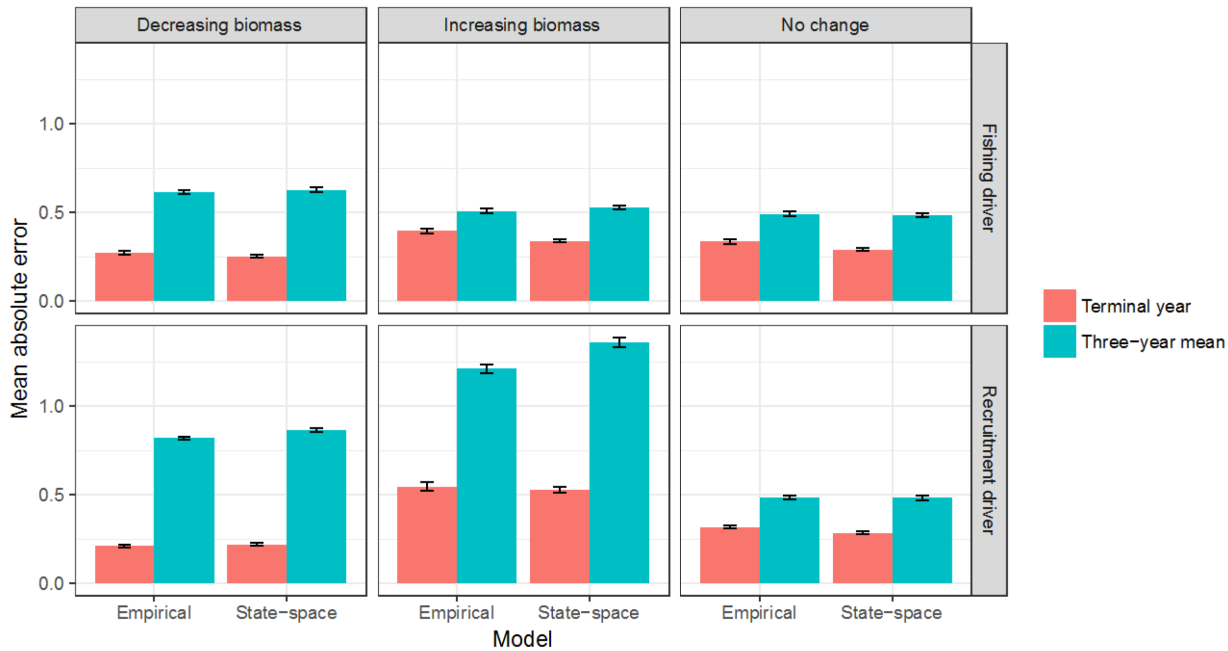

Figure 3. Model estimation error (MAE) for each model, trend, and driver.

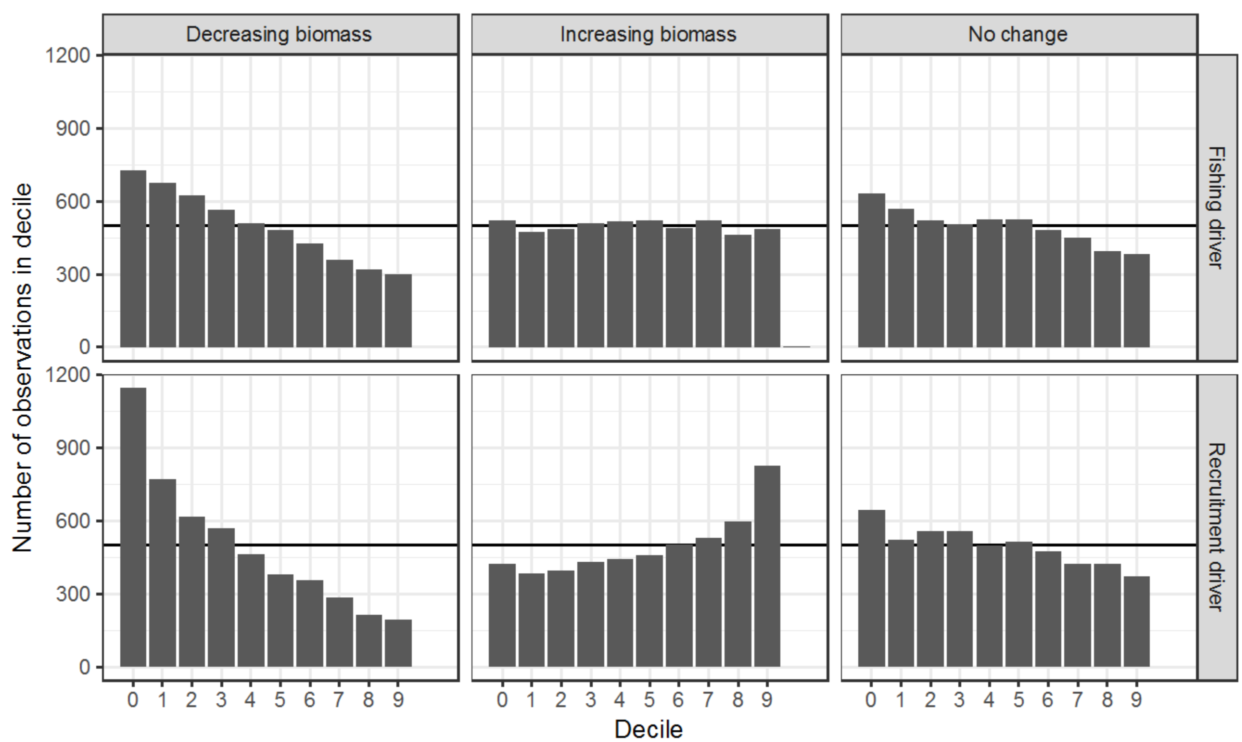

Figure 4. Decile coverage plots for the confidence interval of the state space model for each trend and driver. Bars show the number of observations within each decile. Perfect coverage is represented by the horizontal line.

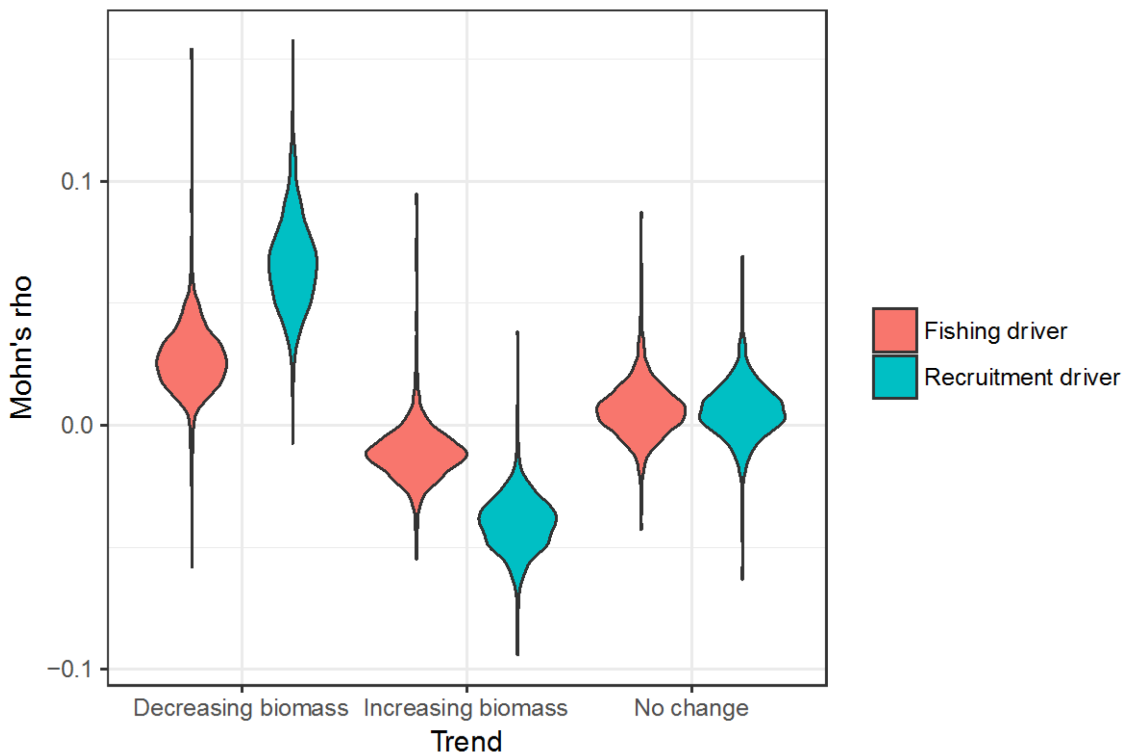

Figure 5. Mohn’s rho for the state-space model for each scenario. The shaded area shows the distribution of Mohn’s rho for each trend and driver.