Fisheries Research 207 (2018) 1–8

Contents lists available at ScienceDirect

Fisheries Research

journal homepage: www.elsevier.com/locate/fishres

When does smoothing the output of index-based stock assessments improve estimates of fish population biomass?

Charles T. Perretti*, Christopher M. Legault

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

ARTICLE INFO

Handled by A.E. Punt

Keywords:
Stock assessment
Yellowtail flounder
Time series
State-space model

ABSTRACT

Index-based stock assessments often use time series smoothing to reduce estimation error in their output. However, this may increase error if actual changes in the population are masked by the smoothing. Therefore, it is important to understand how smoothing assessment output will affect estimation accuracy under conditions commonly observed in fish survey time series. Here, we simulated the Georges Bank yellowtail flounder survey time series and assessment in which three bottom trawl surveys are used to estimate the state of the population. We compared the accuracy of the assessment output when it is unsmoothed versus when is smoothed by applying a three-year moving average to the output index. We did so for two estimation models: the currently used empirical assessment model, and a multivariate state-space random walk model. We simulated five biomass trends under two drivers, and examined model performance when one, two, or three surveys were used to fit the model. Overall, the unsmoothed state-space model consistently outperformed the other methods, particularly when there was a rapid change in biomass in the final years, and all three surveys were used to fit the model. The benefit of using the unsmoothed estimate also increased with the underlying rate of change in the population. The advantage gained by using the unsmoothed estimate also increased as additional surveys were used to fit the model. The benefit of using the state-space model versus the empirical model was similar to the benefit of using the unsmoothed versus smoothed estimate. Our results suggest that an unsmoothed state-space model should be used when multiple survey time series are available and population biomass appears to be rapidly changing in recent years.

1. Introduction

A common goal in stock assessment is to estimate population status from fishery-independent surveys. This is particularly important in index-based stock assessments which rely on surveys as their primary indicator of population status (Legault et al., 2014; NEFSC, 2015). In the Northeast US, index assessments use multi-species bottom trawl surveys, which implement design-based sampling to quantify fish abundance and biomass (Azarovitz, 1981). These surveys yield valuable information. However their results often exhibit high levels of temporal variability. This variability may be driven by factors other than changes in population biomass, such as observation error, changes in catchability, and changes in species availability to the survey. Since these factors do not represent true fluctuations in population biomass, index assessments often apply a smoothing statistic to their output to provide a more accurate estimate of population size or trend.

A common approach to smoothing index-based assessment output is to replace the most recent year estimate with an average of several years. Different time spans for the smooth are used in different assessments, but a three-year average is common (e.g., NEFSC, 2009). Estimation error is driven by both bias and variance (Hastie et al., 2016), and, although smoothing will reduce variance, it may also increase bias if the true population state is distinctly different in sequential time points. Therefore, there is a possible tradeoff between variance and bias, where smoothed estimates are low in variance but potentially high in bias, while unsmoothed estimates are high in variance but potentially low in bias. Therefore, it is important to understand how smoothing affects estimation error under conditions commonly observed in index-based stock assessments.

To investigate this question, we performed a simulation study using Georges Bank yellowtail flounder (GBYT) as our focal stock. The GBYT stock is a transboundary resource in Canadian and US waters, which ranges from Chesapeake Bay to southern Labrador. Biomass of GBYT has reached historic lows in recent years (Legault and McCurdy, 2017), and in 2014 the virtual population assessment (VPA) model was replaced with an index-based assessment model (Legault et al., 2014),

* Corresponding author.
E-mail address: charles.perretti@noaa.gov (C.T. Perretti).

https://doi.org/10.1016/j.fishres.2018.05.024
Received 14 January 2018; Received in revised form 25 May 2018; Accepted 26 May 2018
0165-7836/

which we refer to as the "empirical assessment." The VPA had suffered from strong retrospective patterns for several years before being replaced by the empirical assessment. The most recent GBYT empirical assessment estimated population biomass as the average catchability-adjusted biomass of three bottom trawl survey time series— the US Northeast Fisheries Science Center (NEFSC) spring survey, the NEFSC fall survey, and the Canadian Department of Fisheries and Oceans spring survey. Each data point in each time series is the stratified mean catchability-adjusted biomass from the survey. The assessment takes the simple average of the three time series to condense them into a single index, and then applies an exploitation rate to generate catch advice for the upcoming year. The survey indices are highly variable, with some outlier years. Therefore, smoothing of the assessment output is a potentially useful method for reducing observation error.

An alternative to taking a simple average of the three time series is to use a statistical model to estimate a single latent time series from the three observed time series (i.e., a state-space model). State-space models have become increasingly popular in ecology due to their ability to partition sources of error, estimate missing observations, and model hierarchical processes (Durbin and Koopman, 2012). They have been used in stock assessments to estimate, for example, environmental effects on recruitment (Miller et al., 2016), time-varying selectivity (Nielsen and Berg, 2014), and uncertainty in natural mortality (Cadigan, 2015). Therefore, we compare the currently-used empirical assessment model to a multivariate random walk state-space assessment model and evaluate whether smoothing the output index generated by the state-space model improves its accuracy.

We hypothesize that the trend in the underlying population state in the final years of the index will have a strong effect on the usefulness of smoothing. Therefore, we simulate a range of trends in the final five years of the time series, and examine whether driving the trend by either recruitment or fishing has an impact. Although the most recent GBYT flounder assessment utilized three surveys, many stocks are assessed with only a single survey. Therefore, we also examine how performance is impacted by using one, two, or three surveys to fit the model. Finally, for the state-space model, we examine the retrospective pattern in each scenario and the accuracy of the confidence interval.

2. Methods

2.1. Operating model

The GBYT stock was simulated using an age-structured model with ten ages, including a plus group. A 100-year burn-in period ending in 1972, with the fishing mortality rate (F) set to the 1973 value, was used to create variation in initial biomass and allows us to evaluate whether the stationary distribution of the model approximates that of the first available observations. From 1973 onward the model uses recruitment and fishing mortality estimates from the 2014 GBYT VPA assessment model (Legault et al., 2014). The mean simulated abundance-at-age in 1973 for all age classes was within a factor of 2.5 of the estimated age structure of the VPA. Although the VPA model is no longer used for management, it represents our best estimates for recruitment and mortality over that time. The fishing mortality during 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.

Recruitment was modeled as a log-normal random variable with mean (μ_t) equal to the estimated recruitment in the stock assessment and standard deviation of 0.3 (σ_r).

$$N_{t,1} \sim lnN(\mu_t, \sigma_r^2)$$

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

where $N_{t,1}$ is recruitment, which is the abundance of age-1 fish at time t, and is drawn from a lognormal distribution with variance σ_r^2 and mean μ_t.

The abundance of fish aged 2–9 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.

$$N_{t+1,2:9} = N_{t,1:8}e^{-(F_t s_{1:8}+M)}$$

$$N_{t+1,10+} = N_{t,9}e^{-(F_t s_9+M)} + N_{t,10+}e^{-(F_t s_{10+}+M)}$$

The observed abundance for each age-class is lognormally distributed with mean equal to the true abundance, and standard deviation of 0.3 (σ_s)

$$Nobs_{t,a} \sim lnN(\mu obs_t, \sigma_s^2)$$

$$\mu obs_t = \log(N_{t,a}) - 0.5\sigma_s^2$$

Biomass-at-age ($B_{t,a}$) is the product of abundance-at-age and weight-at-age (w_a).

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

Parameter values are listed in Table 1.

2.2. Data generation

The NEFSC spring, NEFSC 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 of 0.3 (σ_s, Table 1), and applied to the abundance of each age class. The assumed standard deviations result in average CV over all replicates of 1.06 which matches the actual data to within 0.01. In the real time series, there are instances in which several surveys exhibit outlier observations simultaneously (e.g., 2008 DFO and Fall NEFSC). To emulate this effect, in each simulation one year was randomly selected as an outlier year in which the standard deviation of the log-normal error was

Table 1
Simulation parameter values and descriptions. Fishery selectivity-at-age, weight-at-age, and the natural mortality rate are from Legault et al. (2014).

Parameter	Value	Description
σ_r ($N_{t,a}$)	0.3	Standard deviation of process error
σ_s ($Nobs_{t,a}$)	0.3	Standard deviation of observation error
s_a (yr^{-1})	0.01, 0.2, 0.6, 1.0, …., 1.0	Fishing selectivity at age
w_a (kg)	0.148, 0.317, 0.453, 0.588, 0.724, 0.921, 0.921, …, 0.921	Weight-at-age
M (yr^{-1})	0.4	Natural mortality rate
$F_{multiplier}$	0.72, 0.25, 0.0, 2.15, 11.9	Fishing mortality rate multiplier for the no change, increasing slowly, increasing rapidly, decreasing slowly, and decreasing rapidly scenarios, respectively
$R_{multiplier}$	1.1, 1.39, 1.7, 0.8, 0.5	Recruitment multiplier in the no change, increasing slowly, increasing rapidly, decreasing slowly, and decreasing rapidly scenarios, respectively