School of Aquatic and Fishery Sciences

University of Washington

Box 355020

Seattle, WA 98195-5020

(360) 710-9887

(206) 685-7471 (fax)

01 June, 2016

Dear Iyshwarya Gopinath,

I have submitted a revision to the manuscript FISH6657 entitled “Can autocorrelated recruitment be estimated using integrated assessment models and how does it affect population forecasts?” The revision addresses each of the concerns listed by the reviewers, and changes made to the manuscript and responses to each comment are specified below in bold. The manuscript benefited from the additional simulations suggested by Reviewer 2 and the editorial changes suggested by both Reviewers. I hope the revision meets your expectations.

All changes to the manuscript are tracked using track changes, and can be viewed in the submitted file.

Sincerely,

Kelli Faye Johnson

kfjohns@uw.edu

Reviewer 1

Line 38: What does 'forecasted waded biomass' mean?

**The typo has been replaced with “true biomass outside the predictive interval for the forecasted biomass”.**

Line 175: Should this be CV or standard error?

**No, the index is fitted using a lognormal distribution, and the variation about the true index of abundance (in log-space) is specified as a standard deviation. Although, the CV ~ sd when because with a lognormal distribution.**

Line 205: What is tfirst?

***tfirst* is now properly defined following the equation by the statement “*tfirst* refers to the first year that recruitment deviations are estimated.”**

Line 310: It doesn't really look like estimates become more precise with more data to me?

**The distribution width for estimation of autocorrelation decreased with increasing amounts of data (i.e., the width of the distributions in the top row of Fig 3 are smaller than the bottom row, where the length of the time series decreases from top to bottom row).**

Reviewer 2

However, I have one major concern regarding the main conclusion of the ms, that the external method is the most suitable method for accounting for autocorrelated recruitment. My concern relates to the fact that the model is perfectly specified (except the autocorrelation in recruitment) and to me it is unclear if the main conclusion holds if this major assumption is not met. I realize that the authors already partly address this issue, but this is not sufficient. Either tone down the discussion and conclusion or address this quantitatively.

**Thank you for the suggestion to investigate a more realistic scenario to help improve the applicability of our results to stock assessment scientists. We added a scenario “test” where we estimate steepness (rather than fix it at its true value), generate recruitment deviations with a marginal log-standard deviation of recruitment 0.6 (rather than 0.4), provide data starting in year 41, and specify the log standard deviation regarding the index of abundance at 0.25 (rather than 0.1). Although the model is still correctly specified, we believe that the increased uncertainty in the data and the additional complexity of estimating steepness more closely represents an actual stock assessment and provides further justification for our conclusions, which remain unchanged. We will leave it to further research to test increasingly complex scenarios.**

**The following text was added to the manuscript:**

**Methods “Additionally, a ‘less-information’ scenario was simulated and fitted using each estimation while also estimating stock-recruit steepness to facilitate evaluating performance in a more realistic environment”**

**2.1 “An autocorrelation level of 0.5 and a marginal log-standard deviation of recruitment of 0.6 (0.2 higher than all other scenarios) was used for a ‘less-information’ scenario.**

**2.2 “Data collection for the ‘less-information’ scenario started in year 41 and the log-standard deviation of the index of abundance was 0.25.”**

**2.3.1 “Steepness was estimated in the ‘less-information’ scenario using a beta prior (mean = 0.65, sd = 0.147) and fixed at the true value for all other scenarios.”**

**3.1 “Additionally, ‘External’ estimation of ρ was on average less biased than ‘Internal’ estimation for the ‘less-information’ scenario ( -0.21 and 0.42, respectively).”**

**3.2 “Coverage was less than expected for all estimation methods in the ‘less-information’ scenario (Fig. 7).”**

**Discussion “Estimating steepness proved to be difficult no matter which estimation method was used to account for autocorrelated recruitment deviations, reminding us that poor forecast coverage can arise from causes other autocorrelated recruitment.” “with more-informative data (e.g., catches from a stock experiencing a large contrast in spawning biomass)” was added to the next sentence as well.**

**Fig. 7 was added to show the forecast coverage of the “less-information” scenario.**

Minor comments:

L111 Give a (short) description of the main components of the SS tool

**The following text was added to line 114 “is an age-structured forward-projection single-species stock assessment framework that” to describe Stock Synthesis.**

L166-169 why not cover the full range of values (-0.9 to 0.9)?

**Thorson *et al*. (2014) estimated autocorrelation for a range of species, and the lowest estimated value was -0.06. Therefore, we chose to create our range of scenarios based on what we thought to be plausible, rather than performing a full-factorial design. The citation is listed in the text.**

L173 Please specify the multinomial model

**We do not believe that this addition will substantially improve the text.**