Associate Editor comments to author:

Multiple reviewers mention sloppy notation, question the meaning/use of “<<” in Line 105, and have concerns about dropping unconverged runs and simulating again until 50 runs were converged for each scenario. Also, there is general consensus that you should discuss your convergence criteria. Some suggestions for improving Figures and Tables are provided. Two reviewers mention Aldrin et al. (2020), which seems relevant and the authors should discuss relative to differences in finding/guidance. Reviewer 4 raises a good conceptual question about section 2.3, please expand on this point.

Figures S10+S11 – Are you isolating the effect of observation error or process error by setting the CV so low? Or do you mean that you are minimizing the effect of observation error?

Primary finding is that biased estimation of variance parameters, primarily sigmaNAA, produces bias in model estimated SSB and NAA when bias correction is on. But this is not highlighted in the Abstract and should be.

Comment on the role that the ar1\_y parameter might contribute to the bias resulting for bias correction. The value is quite large for flounder, nearly random walk, and this is the one stock with the strongest contraindication for bias correction.

The paragraph on Lines 296-307 was not a well developed line of analysis, not mentioned in the methods, and the reader has no information about what improvement in data was relative to earlier years. If this is an important part of the analysis, I suggest developing it further and providing more details that quantify data improvements. At present, the ‘conclusion’ that data quality and quantity are less influential is more of a hypothesis, but stating it as such without a well designed test could work against efforts to maintain high quality data streams.

In Lines 310-311, the authors state “In the absence of strong evidence in support of bias correction…” – what would such evidence look like? It would be good to give readers some insight on what they should be looking for.

Reviewer comments to author:

Reviewer: 1

Comments to the Author

See attached.

Reviewer: 2

Comments to the Author

Please see the attached PDF for comments.

Reviewer: 3

Comments to the Author

Comments on «Evaluating the impact of log-normal bias-correction on a

state-space stock assessment model»

This paper addresses an important question: bias-correction. I have a few concerns about the paper in its current form.

1. I do agree that bias-correction should be applied to observations, as in Section 2.3. Eqn (7) is saying that the observed catches C are biased estimates of the modelled catches \hat C. The authors should justify why this bias occurs. Further, from a theoretical statistical standpoint the maximum likelihood estimator is invariant to transformation of the data, i.e. we get the same parameter estimates whether we fit to C or log(C) (https://en.wikipedia.org/wiki/Maximum\_likelihood\_estimation: “The MLE is also equivariant with respect to certain transformations of the data….”). For normally distributed data (log(C)), nobody would suggest a bias correction. I may be wrong, but if I am right, the authors are bringing confusion into the discussion about bias correction. (My reading of the results and conclusion of the paper, bias correction applied to observations does not have a big effect)

2. My next point is similar, but applies to Section 2.2. I do not agree that there should be an operating model without bias correction. If the expectation of exp(\epsilon\_1) in eqn (1) is not 1, it affects the interpretation of the recruitment function f(SSB). The paper does not give details about f(SSB), so we do not know if f(SSB) will be correct for the missing bias correction. Hence, by opening up for E[exp(\epsilon\_1)] neq 1, I feel the authors are making the discussion too general.

3. The number of simulation replica is low only 50. By increasing by x10 would improve the precision of the results, but if the authors feel they can show their results with only 50 replica, I guess it is OK.

4. The paper is in general clearly written, but a bit sloppy written at some places:

a. Line 104: “larger than” followed by “neq”?

b. In eqn (8), is it median over y? Also, first it says theta\_{i,y}, and right after is says theta(i,y)

Reviewer: 4

log(Recruitment) = log(μ\_mean) - 0.5 \* σ\_Rec^2 + random\_error

In this formula, μ\_Rec is the parameter for the **true average (mean) recruitment**.

If **σ\_Rec** is underestimated... ...the bias correction term (-0.5 \* σ\_Rec^2) becomes a smaller negative number (i.e., it's too big). To compensate and get the right answer for log(Recruitment), the model **must underestimate μ\_Rec**.