EGUsphere, referee comment RC2
https://doi.org/10.5194/egusphere-2022-579-RC2, 2022

Comment on egusphere-2022-579

Anonymous Referee #2

Referee comment on "Anthropogenic climate change drives non-stationary phytoplankton variance" by Geneviève W. Elsworth et al., EGUsphere,
https://doi.org/10.5194/egusphere-2022-579-RC2, 2022

he manuscript, *Anthropogenic climate change drives non-stationary phytoplankton variance*, summarizes projected changes in global and regional phytoplankton variability using the NCAR CESM1 Large Ensemble under a high emissions scenario. The authors explore the key drivers of declining phytoplankton variability, highlighting the importance of top-down, zooplankton grazing in potentially driving future phytoplankton response.

Generally, the article concisely represents its findings but there are several points of clarification I would recommend. In particular, the use of specific statistical terminology could be more accurate. Multiple times throughout the text, the term "variance" is used when, I think, "variability" is intended. In many cases this "variability" is being assessed via the standard deviation of the large ensemble members which is similar to the variance but not the same. Additionally, I am not proficient in MLR, but the comments made in the prior review are troubling especially considering the results are key to the paper's conclusions regarding top-down controls but these results seem underrepresented in the primary manuscript text. I've included several additional minor comments and suggestions below pertaining to clarity and organization.

Specific Comments and Suggestions:

Lines 49-52: *Clarifying how variance in phytoplankton biomass may be changing over long time scales with climate change is important for fisheries management, especially at regional scales. Near- term predictions of phytoplankton biomass may also benefit from knowledge of the projected magnitude of internal variability, as the chaotic nature of internal variability hampers near-term predictions (Meehl et al., 2009, 2014).*

I think it's worth noting that the internal variability quantified using a large ensemble is Internal variability specific to the model and indicative of our uncertainty that results from its simplified representations of the real world processes and numerics. It doesn't necessarily have any bearing on real world manifestations of variability. Its primary utility to management and fisheries is in guiding our level of confidence in disentangling model signals from the noise.

Lines 103-104: *Six CESM1-LE members had corrupted ocean biogeochemistry*

I'm curious, what does "corrupted ocean biochemistry" mean? it might help to explain what makes an ensemble member usable versus not.

Figure 1. Add units: standard deviation should have the same units as the variable being assessed (i.e., phytoplankton carbon) but none appear in figure 1.

Lines 121-122: *Internal variability at each location (x,y) is approximated as the standard deviation across ensemble members (EMs) at a given time (t)*

The method described here indicates that the standard deviation is being used to quantify variability. However, throughout the paper, the authors reference the "variance" when I think they mean "variability". This is problematic because "variance" and "standard deviation", while related, are two different values and the way they are interchanged throughout the text is confusing. Please check all instances of "variance" in the paper for intended meaning and replace with "variability" where appropriate. I suggest including a description of the "coefficient of variance" method here, too.

Lines 142-143: *However, while the model ensemble captures regional patterns of observed variability, the CESM1-LE overestimates the magnitude of observed interannual variability.*

I may be mistaken but it seems this was determined using only a single ensemble member - is it appropriate for conclusions to be drawn for the full ensemble when only considering one ensemble member?

Lines 147: *A synthetic ensemble is a novel technique*

I don't think this technique can be called "novel" if it appears in two prior references

Lines 149-151: *Compared to the internal variability over the observational period (2002 to 2020) (purple circle, (Figure S2), the model ensemble slightly overestimates the magnitude of internal variability in chlorophyll observed in the real world.*

This seems like a result/ should appear in the result section. Also, it makes an assessment of the ensemble as a whole, but isn't it still based on the results from the single ensemble member? If not, this was a point of confusion for me, and I suggest clarifying.

Lines 153-154: *Annually averaged, global mean, upper-ocean (top 150m) integrated phytoplankton biomass across the model ensemble decreases from 76.1 mmol C m-2 to 66.2 mmol C m-2.*

It's not clear what timeframes these values represent. Is it 2006 vs. 2100? If so, it seems that such a narrow, 1-year window would risk aliasing higher frequency variability and potentially under- or overestimate the change in mean state. This is somewhat compensated for by the size of the ensemble but differs from the 10-year averaging described later in Line 223 Also, I suggest reporting the standard deviations for these numbers.

Lines 177-178: *we calculated the coefficient of variance as the standard deviation across the ensemble members for a given year (ensemble spread) divided by the ensemble mean.*

I suggest including this description in the methods section rather than the results.

Lines 178-180: *Figure 2b illustrates the change in the coefficient of variance from the historical period through the RCP8.5 forcing scenario (1920 to 2100).*

The results seem to jump from Figure 2a, to Figure 3, then back to 2b which is a bit confusing.

Line 180: *The coefficient of variance is relatively constant across the historical period (1920 to 2005), and then significantly declines by ~20% from 2006-2100.*

I'm not sure I agree with the assessment that the coefficient of variance is relatively constant across the historical period. 1920-1980 appears to have a positive trend with a range of about 6.1 to 7.3, which appears similar to the range of the time period covered by the dashed line in Figure 2b. I suggest testing the significance of the 1920-1980 trend.

Also, could the drop in coefficient of variance instead be explained by temporal distance from the perturbation that differentiates the ensemble members? If the 34 ensemble members differ in initial air temperature conditions, would the spread perhaps be expected to decrease as the simulation integrates further away from that initial discrepancy (i.e., solutions start to converge)?

Lines 190-193: *From 2006 to 2100, the coefficient of variance decreases by 3.3 x 10-5 yr-1 in the CESM1-LE, 2.0x10-4 yr in the MPI-ESM-LR1, 5.2x10-5 yr-1 in the CanESM2, and 3.9 x10-4 yr-1 in the GFDL-ESM2M. These declines are statistically significant in all model ensembles with the exception of the MPI-ESM-LR1 (Figure S2).*

It's not clear how these values across models are calculated, whether the end points of the time series or a range of years - the latter would be more appropriate (as done in Line 223) to avoid higher frequency variability and thus under- or overestimating the nature of the change. I also suggest reporting the specific statistical testing methods in the text if stating that the changes are significant.

Line 201: *We observe the largest magnitude decline in total phytoplankton carbon variance*

The table is reporting change in standard deviation, not variance. Standard deviation is expressed in the same units as the analyzed variable while variance is reported in the square of those units.

Figure 4: It's not clear what this figure adds to the discussion - it seems to be redundant with information in Figure 5. Perhaps if the outlines of the ecological regions were included?

Lines 219-221: *We quantified the relationship between phytoplankton carbon and the variables which contribute to changing phytoplankton biomass and its internal variability by performing a multiple linear regression (MLR) analysis. The MLR analysis was performed on linearly detrended annual anomalies using the ordinary least squares function of the Python package statsmodels.api*

This and the associated equations seem to belong in the methods section.

Line 274: *…and important global biogeochemical regions…*

What is considered an important biogeochemical region? This seems somewhat vague - I suggest elaborating to be a bit more specific.

Lines 278-280: *As such, the magnitude of changes in phytoplankton internal variance derived from the model ensemble should be interpreted as an overestimate when considering changes in phytoplankton internal variance driven by anthropogenic warming.*

Again, my impression was that this conclusion was derived from analyzing a single ensemble member which seems insufficient for assessing the entire ensemble.