Review report on “Anthropogenic climate change drives non-stationary phytoplankton variance”

This is a very interesting study on the impacts of climate change on the internal variability of phytoplankton, based on an Earth System Model ensemble. While I believe this study is quite important and of interest in terms of results (the main result being a reduction in phytoplankton internal variability under anthropogenic driven climate change) as in terms of implications (especially the link with fisheries stock assessment, whose uncertainty could be reduced as a consequence of this reduced phytoplankton internal variability), I have major concerns about the MLR method used to estimate the physical and biogeochemical drivers of trends in phytoplankton internal variability. I elaborate on these and other issues below.

MAJOR COMMENTS:

My first concern is the choice of explanatory variables for the MLR:

MLR is a great tool for exploring relationships between variables, but as you have indicated in the text, it is unable to distinguish between bottom-up and top-down relationships that link two variables. For these reasons, in order to be able to interpret the results with causality relationship, you should :

1) Use only variables for which the causal relationship with phytoplankton biomass is known (or for which the first order of this relationship is known), e.g., SST (the first order is a bottom-up relationship: warming drives phytoplankton biomass by increasing metabolic rates, a positive effect, and by increasing nutrient stratification, a negative effect. At second order, one could have a top-down feedback of phytoplankton biomass change modifying carbon cycling and indirectly temperature, but one would neglect this effect). For this reason, I think that including zooplankton/zooplankton grazing in such an analysis is not appropriate because you are not able to separate top-down and bottom-up effects on phytoplankton biomass.
2) Use only variables for which the causal relationship with the target variable is the same. In your case, use the variables for which phytoplankton biomass is a consequence, not a cause. Again, while I believe that zooplankton do exert top-down control over phytoplankton (so that phytoplankton biomass would be a consequence), PFT models with small numbers of zooplankton are likely to be dominated by bottom-up control (so that zooplankton would be the consequence and phytoplankton the cause). To support this claim, trophic amplification under climate change in these models has been described as a good indicator of bottom-up control of zooplankton by phytoplankton (Chust et al., 2014, Kwiatkowski et al., 2019), a pattern that is altered when higher trophic levels are considered (Dupont et al., 2022).
3) MLR analysis assumes independence of explanatory variables, which is clearly not the case (e.g., MLD and Nutrient are highly correlated). I agree that this is a classic problem in multivariate analysis on climate variables, but this point should be discussed further, by providing at least one correlation matrix between all explanatory variables.

Nevertheless, your signal on zooplankton is clearly related to the strong relationship between zooplankton and phytoplankton, which is expected but clearly interesting. I think you should analyze (with simple linear regression) the relationship between phytoplankton and zooplankton separately from the other variables, which would clearly fit the main message of your paper: showing that the effect of climate change on the internal variability of phytoplankton is transferred to the internal variability of zooplankton would demonstrate a transfer of the trend in internal variability to the higher trophic levels (in this case, zooplankton), which you could then extrapolate in the discussion to even higher trophic levels (e.g. fish). It would also be interesting to compare trends in phytoplankton and zooplankton internal variability. Is it higher? Lower? Why? You could also do the same MLR analysis with trends in total plankton with the bottom-up effect.

My second concern is about the MLR method itself, which I think is wrong in its current form:

First of all, I am missing some details to understand what exactly you did with the MLR. You would need to make it clear which variables are used for each step of the method. In particular, it is not clear which variable depends on i) time, ii) space, and iii) the member of the model set. For the rest of my argument, *t* will refer to time, x to the grid cell (spatial position) and i to the model set member. Y will refer to the phytoplankton biomass and X to any explanatory variable.

A) So, for what I understand, your first step was to prepare linearly detrended annual anomalies. So, for a variable X, with a trend a, the considered variable in the MLR is

$X^d(t,x,i)=(X(t,x,i)-X(0,x,i))-a(x)*t$

, a field with 3 dimensions : space, time and model ensemble member. The same calculation gives you $Y^d(t,x,i)$. If it is based on globally averaged values, I would recommend to keep the space dimension.

B) With the MLR, you fitted the following relationship and thus estimated the coefficients dY^d/dX^d :

$Y^d(t,x,i) = sum ((dY^d/dX^d)* X^d(t,x,i)))$

An approximation of the first order taylor development which would give

$Y^d(t,x,i) = sum ((dY^d/dX^d)(t,x,i) * X^d (t,x,i))$

C) Then, by linearity, you compute (t being a 10-year average)

$Sigma\_i(Y^d(t,x,i))= sum (dY^d/dX^d\ sigma\_(X^d(t,x,i)))$

BUT : **Variance isn't linear (neither is the standard deviation).** Even if two variables are independent (which is definitely not the case), $VAR(aX+bY)= a^2 Var (X) + b^2Var(Y)$.

In your case, if you wanted to reconstruct your variance, you would use the following formula (a_i being your MLR coefficient dY^d/dX^d and Cov being the covariance and not the coefficient of variance here):

$$\begin{aligned}\operatorname{Var}\left(\sum_{i=1}^N a_iX_i\right) &= \sum_{i,j=1}^N a_ia_j \operatorname{Cov}(X_i,X_j)\\ &= \sum_{i=1}^N a_i^2 \operatorname{Var}(X_i) + \sum_{i\neq j} a_ia_j \operatorname{Cov}(X_i,X_j)\\ &= \sum_{i=1}^N a_i^2 \operatorname{Var}(X_i) + 2\sum_{1\le i<j\le N} a_ia_j \operatorname{Cov}(X_i,X_j).\end{aligned}$$

You can calculate this value perfectly well, but I'm not sure that's what you want to do.

Indeed, I'm not sure what kind of information you expect from the relationships between linearly detrended variables and linearly detrended phytoplankton carbon biomass: do you want to explain the internal phytoplankton variability by the internal variability of others variables or do you want to explain it by the trends of other variables ? I think the second option, or both, would be more appropriate (e.g., is the increase in temperature related to the reduction in internal variability of phytoplankton biomass?)

While I think your current method is wrong, I keep in mind that too few details have been provided to be certain, and perhaps I will be convinced of your method when more details are included. Nevertheless, I suggest another approach:

You have introduced CoV (Coefficient of Variance), and I think this variable is indeed more appropriate than standard deviation because it removes the effect of reduced mean state values on the change in internal variability (i.e., a lower mean state will lead to lower internal variability in absolute magnitude, but not necessarily to a reduced coefficient of variance).

You could perform the MLR on CoV directly (which would mean using the mean anomalies of the entire model ensemble), i.e., estimate a linear relationship between *CoV_phyto(t,x)* and other variables:

*CoV_phyto(t,x) = CoV_phyto(0,x) + a * anomaly_var1(t,x)+ b * anomaly_var2(t,x)…*

Or if you want-to keep detrended variables:

CoV_phyto(t,x) = CoV_phyto(0,x) + a CoV_var1 (t,x)+ b CoV_var2(t,x)…

MINOR COMMENTS :

Discussion : The discussion is quite short, I would like to see a discussion of the mechanisms that might lead to this reduction in internal phytoplankton variance. In addition, the discussion focuses on the top-down control of zooplankton on phytoplankton. Although the authors no longer assert in the current version that zooplankton are a driver of trends in internal phytoplankton variability, they continue to discuss it, which is not necessarily relevant. Given my main comment on how to study the relationship between phytoplankton and zooplankton variability, I would focus on the bottom-up effect of phytoplankton on zooplankton to support the impact of changes in phytoplankton variability on higher trophic levels, and then discuss top-down effects as a limitation to the interpretation of your results.

Wording : Consider using “internal variability” instead of “variance” throughout the text, starting with the title. While variance can refer to many temporal scales (seasonal, interannual,…), I think internal variability is much more accurate (e.g., L7, “internal variability” instead of "internal variance")

Other comments :

L1: Bopp et al., 2001, 2013; Laufkötter et al., 2015; Kwiatkowski et al., 2020 are model studies. If you want to keep past tense, please add data-based reference. Or use another tense.

L4: I would mention the impact of phytoplankton on the carbon cycle. Also in the discussion.

L45-46: As formulated, the results of the Resplandy’s study are not clear.

L54: The last sentence of the paragraph does not flow well with the rest.

Fig 1 and L160: “by ensemble member 1 of the CESM1-LE”: why not using the average of the model ensemble members ?

Eq. 1: use a separate symbol for internal variability (you have twice sigma)

Eq. 2: Specify on which variable your mean LE is computed (time, space and model ensemble members)

L132: unclear, why not using the same term as above, i.e., "ensemble mean", i guess that LE(x,y,t) is the same as LE but this needs to be specified.

L134: Please avoid using variance if you refer to internal variability.

L164: use interannual instead of temporal which is no precise

L179 (and L290): I disagree with this statement: the comparison between the observed variance and the modeled variance does not give any information about the trends in the variance.

Fig 3. Specify the variables on which you have applied a t-test, at least as supplementary material. The reader should be able to assess the validity of your test (sample size, normality assumption, etc.).

L239: Why not use the total biomass of phytoplankton everywhere?

Fig 4: Please add the regions to the map (at least ASP, SAP, SOC, and EQP), and consider showing diatom biomass over small phytoplankton biomass ratio, as diazotrophs are not dominant anywhere and to be more quantitative.

L260: The wording of the sentence suggests that zooplankton exert top-down control over phytoplankton, which is uncertain or even false.

L275-279: I think this result (with an appropriate MLR method) is very interesting in explaining what is driving the zooplankton variability (see main comment) and should be interpreted from a zooplankton perspective.

L287: I do not believe that either of these references is relevant to this statement.

Références:

Chust, G., Allen, J.I., Bopp, L., Schrum, C., Holt, J., Tsiaras, K., Zavatarelli, M., Chifflet, M., Cannaby, H., Dadou, I., Daewel, U., Wakelin, S.L., Machu, E., Pushpadas, D., Butenschon, M., Artioli, Y., Petihakis, G., Smith, C., Garçon, V., Goubanova, K., Le Vu, B., Fach, B.A., Salihoglu, B., Clementi, E. and Irigoien, X. (2014), Biomass changes and trophic amplification of plankton in a warmer ocean. Glob Change Biol, 20: 2124-2139. https://doi.org/10.1111/gcb.12562

Kwiatkowski, L, Aumont, O, Bopp, L. Consistent trophic amplification of marine biomass declines under climate change. *Glob Change Biol*. 2019; 25: 218– 229. https://doi.org/10.1111/gcb.14468

Dupont, L., Le Mézo, P., Aumont, O., Bopp, L., Clerc, C., Ethé, C., & Maury, O. (2022). High trophic level feedbacks on global ocean carbon uptake and marine ecosystem dynamics under climate change. *Global Change Biology*.