Review 1 of Elsworth et al. (2022) "Anthropogenic climate change drives non-stationary phytoplankton variance", submitted to Biogeosciences.

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. 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. 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.

We thank the reviewer for their constructive comments and appreciate their careful reading of the manuscript. In response to this and another reviewer's feedback, we took a different approach to develop an understanding of the drivers of changing phytoplankton coefficient of variation (CoV) with anothropogenic climate change. We now use a machine learning approach in which we generate an ensemble of boosted regression trees to quantify drivers in changing phytoplankton carbon biomass CoV. Unlike linear models, boosted trees are able to capture non-linear interaction between the explanatory variables and the target variable. At every step, the ensemble fits a new learner to the difference between the observed response and the aggregated prediction of all learners grown previously, aiming to minimize mean-squared error.

We use the Matlab function *predictorImportance* to estimate the importance of the predictors for each tree learner in the ensemble, which computes the importance of the predicotrs in the tree by summing changes due to splits on every predictor and dividing the sum by the total number of branches. Revised estimates of the relative importance of predictor variables for changing phytoplankton CoV for each of the four key ocean regions are illustrated in a new version of Figure 5 and elaborated upon in the manuscript text.



"Figure 5: Relative importance of predictor variables on phytoplankton biomass coefficient of variance across the RCP8.5 forcing scenario (2006 to 2100). Marine ecological regions are defined in Tagliabue et al. (2021). Regions were selected which aligned with the highest fisheries catch in the (a) Atlantic and (b) Pacific basins and the biogeochemically important (c) Southern Ocean and (d) Equatorial Pacific regions. The dominant phytoplankton functional type is considered in each region. In regions with a mixed ecological assemblage, total phytoplankton carbon is considered."

L133: "We quantified the drivers of phytoplankton carbon biomass CoV in key ocean regions by generating an ensemble of boosted regression trees. Unlike linear models, boosted trees are able to capture non-linear interaction between the predictors and the response. A regression tree ensemble is a predictive model composed of a weighted combination of multiple regression trees. At every step, the ensemble fits a new learner to the difference between the observed response and the aggregated prediction of all learners grown previously, aiming to minimize mean-squared error. We generate an ensemble of boosted regression trees (maximum tree depth = 10) using the Matlab function *fitrensemble*. Our predictor variables are the regional mean, ensemble mean temperature, mixed layer depth, incoming shortwave radiation, physically mediated iron, physically mediated phosphate, zooplankton carbon, and zooplankton grazing (diatom, small phytoplankton, or their sum) annually resolved from 2006 to 2100. We use the Matlab function *predictorImportance* to estimate the importance of the predictors for each tree learner in the ensemble; it computes the importance of the predictors in a tree by summing changes due to splits on every predictor and dividing the sum by the total number of branches."

L246: "We identify the importance of different predictors to changing phytoplankton biomass CoV in four distinct ecological regions using a machine learning (boosted regression tree) approach. In the subpolar Atlantic (ASP) and subpolar Pacific (SAP) ecological provinces (Figure 4 diatom

biomass CoV declines between the beginning and end of the century (Table 1). In the Atlantic subpolar region, the most important predictor of diatom biomass CoV is phosphate advection, with smaller contributions from zooplankton carbon (Figure 5a). In the subarctic Pacific region, sea surface temperature is the most important predictor of diatom biomass CoV, with phosphate advection playing a secondary role (Figure 5b)...'

L270: "Using a machine learning approach, we identify the importance of different predictors to changing phytoplankton biomass internal variability. In all four ecological provinces, a combination of bottom-up controls (e.g., nutrient supply, light availability) and top-down controls (e.g., grazer biomass) predict the decline in phytoplankton biomass CoV with anthropogenic warming..."

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 dYd/dXd:

 $Y^{d}(t,x,i) = sum\left((dY^{d}/dX^{d}) * X^{d}(t,x,i)\right)$

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), $VAR(aX+bY) = a^2 Var(X) + b^2 Var(Y)$.

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

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).

$$egin{aligned} \operatorname{Var}\!\left(\sum_{i=1}^N a_i X_i
ight) &= \sum_{i,j=1}^N a_i a_j \operatorname{Cov}(X_i,X_j) \ &= \sum_{i=1}^N a_i^2 \operatorname{Var}(X_i) + \sum_{i
eq j} a_i a_j \operatorname{Cov}(X_i,X_j) \ &= \sum_{i=1}^N a_i^2 \operatorname{Var}(X_i) + 2\sum_{1\leq i < j \leq N} a_i a_j \operatorname{Cov}(X_i,X_j). \end{aligned}$$

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:

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

We thank the reviewer for this helpful insight. To address this suggestion, we revised our statistical analysis approach to use a machine learning approach in which we generate an ensemble of boosted regression trees to quantify drivers in changing phytoplankton carbon biomass CoV. We have replaced Figure 5 with the results from our machine learning analysis and have modified the

associated text to reflect this new analytical framework. Please see the manuscript modifications above.

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 variability" variance")

We agree with the reviewer and have replaced "variance" with "internal variability" both in the title and throughout the manuscript.

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.

Thank you for this suggestion. We have modified the tense in which we refer to modelling studies both in the abstract and in the introduction.

L1: "Earth System Models suggest that anthropogenic climate change will influence marine phytoplankton over the coming century, with light limited regions becoming more productive and nutrient limited regions less productive."

L15: "Anthropogenic climate change significantly impacts marine ecosystems from phytoplankton (Bopp et al., 2001, 2013; Laufkötter et al., 2015; Kwiatkowski et al., 2020) to fish (Perry et al., 2005; Cheung et al., 2009, 2010; Mills et al., 2013; Wernberg et al., 2016; Flanagan et al., 2018; Staudinger et al., 2019)."

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

We agree that the manuscript would benefit from elaborating on the role of phytoplankton in the carbon cycle. We have added text to the introduction to reflect this important role.

L11: "The abundance and distribution of phytoplankton, the base of the marine food web and an important component of the marine carbon cycle, will likely change with anthropogenic warming."

L307: "In this context, a decline in phytoplankton internal variability with anthropogenic climate change may improve the accuracy of near-term predictions of phytoplankton biomass, producing more reliable forecasts of fisheries productivity and marine carbon cycling."

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

Thank you for this clarification. We have revised the text to accurately describe the results of the Resplandy study.

L45: "For example, Resplandy et al. (2015) examined the contribution of internal variability to air-sea CO_2 and O_2 fluxes with climate change using a suite of six ESMs. Their analyses revealed distinct regional differences in variability of air-sea gas fluxes, with the Southern Ocean and the tropical Pacific playing a significant role."

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

Thank you for this comment. We have removed this sentence to enhance the flow of the paragraph.

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

When using ensembles of ESMs, the spread of the ensemble members represents the range of internal variability while the mean of the ensemble members represents the externally forced component. We have compared the observational record to that of ensemble member 1, because it allows for a like-for-like comparison of temporal standard deviation (both member 1 and real-world observations are simultaneously influenced by internal and external factors). However, if we compared the observational record to the ensemble mean we would *instead* be evaluating changes in the externally forced temporal variability over this period.

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

As sigma (σ) is the typical symbol for standard deviation, we have opted to keep this equation asis.

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

This is an excellent suggestion. We have revised Eq. 2 to reflect the precise variables for which the mean LE is computed.

We have added the abbreviation 'EM' to the equation to clarify that we are averaging across ensemble members.

$$CoV(x, y, t) = \frac{\sigma(EM(x, y, t))}{\overline{LE}^{EM}}$$

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.

We have added the abbreviation 'EM' to the overbar to clarify that we are averaging across ensemble members.

$$\overline{LE}^{EM}(x,y,t)$$

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

We have modified the text accordingly.

L134: "We quantified the drivers of changing phytoplankton carbon biomass CoV in key ocean regions by generating an ensemble of boosted regression trees."

L164: use interannual instead of temporal which is no precise

Thank you for this suggestion. We have modified the text to use interannual variability rather than temporal variability.

L154: "We use the observational phytoplankton carbon dataset of Bellacicco et al. (2020), annually averaged and interpolated onto a 1 degree grid, to evaluate interannual variability in phytoplankton biomass in a single model ensemble member."

L159: "However, while the model ensemble captures regional patterns of observed variability, the CESM1-LE overestimates the magnitude of observed interannual variability. Some regions of the global ocean display a substantial mismatch in interannual variability between the model and that estimated from observations (Figure 1, Table S1 While the differences can be quite large in some regions, we note that this is an evaluation of interannual variability (rather than internal variability, the focus of this study), and that estimates from the satellite product derive from a collection of data products which may also display biases (Table S1)."

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.

Thank you for this feedback. We have removed the final sentence of this paragraph to reflect this comment.

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.).

We agree that the manuscript would benefit from elaborating the precise variables on which we have applied a t-test. We have revised the supplementary material to provide more information on the t-test in Table S2 and have referred to this table in the caption of Figure 3.

Table S2. Summary statistics for the t-test performed on total phytoplankton biomass to determine trend significance across the RCP8.5 forcing scenario (2006 to 2100). Datasets are normally distributed.

Variable	Sample Size	Mean	Standard Error	95% CI
Total Phytoplankton Biomass Mean Trend	94	-0.0697	0.00459	-0.0743 to -0.0651
Total Phytoplankton Biomass Standard Deviation Trend	94	-0.0164	0.00323	-0.0196 to -0.0132

L209: "Figure 3. (a) Percentage change in annual total phytoplankton carbon concentration over the RCP8.5 forcing scenario (2006 to 2100) simulated by the CESM1-LE. (b) Percentage change in annual total phytoplankton internal variability over the same period. The change in the mean and the variability are calculated using averages across the first (2006 to 2016) and last (2090 to 2100) decades of the RCP8.5 forcing scenario. Hatched areas indicate regions of trend insignificance determined by a t-test with a p value greater than 0.05. Summary statistics for the t-test are available in the supplemental information (Table S2)."

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

For our regional analyses, we focused on the phytoplankton functional type (PFT) that dominates each region. In the Atlantic subpolar and the subarctic Pacific regions diatoms dominate the regional phytoplankton composition. However, in the Southern Ocean and Equatorial Pacific, both diatoms and small phytoplankton coexist in a mixed assemblage. This nuance allowed us to consider the specific grazing terms that were relevant given the regional PFT distribution.

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.

We thank the reviewer for this helpful insight. We have revised Figure 4 to include the overlayed ecological regions considered in the manuscript. The full map of all 11 ecological provinces is provided in the Supplemental Information (Figure S2).



"Figure 4: Distribution of the dominant phytoplankton functional type in biomass carbon averaged across the RCP8.5 forcing scenario (2006 to 2100). The CESM1-LE simulates three phytoplankton functional types: diatoms, diazotrophs, and small phytoplankton. Regions where diatoms dominate are shown in yellow and regions where small phytoplankton dominate are shown in purple. Diazotrophs do not dominate in any region of the global ocean. The four ecological provinces are shown: subpolar Pacific (SAP), subpolar Atlantic (ASP), Equatorial Pacific (EQP), and Southern Ocean (SOC)."

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

Thank you for this clarification. We have removed this text to reflect the results of our new machine learning analysis.

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.

Absolutely. We agree that the manuscript would benefit from applying a more statistically robust approach. We revised our statistical analysis approach to use a machine learning approach in which we generate an ensemble of boosted regression trees to quantify drivers in changing phytoplankton carbon biomass CoV. We have replaced Figure 5 with the results from our machine learning analysis and have modified the associated text to reflect this new analytical framework.

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

Thank you for this comment. We have removed this text to reflect the results of our new machine learning analysis.

Review 2 of Elsworth et al. (2022) "Anthropogenic climate change drives non-stationary phytoplankton variance", submitted to Biogeosciences.

My only remaining recommendation is to remove Diazotrophs from the color bar for Figure 4. If Diazotrophs to not dominate in any region (as specified in the authors' responses to the initial review), it seems unnecessary to include them in the figure. If the authors have some other reason for including Diazotrophs, it should at least be specified in the caption that Diazotrophs are not actually visible in the figure.

Thank you for this suggestion. We have removed diazotrophs from the color bar in Figure 4 and have revised the figure caption. We have also revised Figure 4 to include the overlayed ecological regions considered in the manuscript.



"Figure 4: Distribution of the dominant phytoplankton functional type in biomass carbon averaged across the RCP8.5 forcing scenario (2006 to 2100). The CESM1-LE simulates three phytoplankton functional types: diatoms, diazotrophs, and small phytoplankton. Regions where diatoms dominate are shown in yellow and regions where small phytoplankton dominate are shown in purple. Diazotrophs do not dominate in any region of the global ocean. The four ecological provinces are shown: subpolar Pacific (SAP), subpolar Atlantic (ASP), Equatorial Pacific (EQP), and Southern Ocean (SOC)."

Review 3 of Elsworth et al. (2022) "Anthropogenic climate change drives non-stationary phytoplankton variance", submitted to Biogeosciences.

This work investigates the internal variability of phytoplankton biomass in an Earth System Model large ensemble (CESM1-LE). The results show a decrease in internal variability under RCP8.5.

The manuscript is well written and well laid out and the results are novel and important. I have only some minor suggestions.

We thank the reviewer for their constructive comments and appreciate their careful reading of the manuscript.

Comments

Line 128: The coefficient of variance seems to be much more commonly referred to as the coefficient of variation. Is there any reason why you use the term "coefficient of variance"?

We thank the reviewer for this clarification. We have revised the text throughout to replace coefficient of variance with coefficient of variation.

Line 130: Should the nominator in this equation read LE(x,y,t) (with a line above) as it is defined on line 132? In that case I would rephrase the sentence on line 131 to say something like: where LE(x,y,t), the forced response of the large ensemble, is calculated as the mean of ensemble members at a given location and time.

We appreciate this suggestion. We have revised the *denominator* of Eq. 2 to reflect that we are calculating the mean of the ensemble members at a given location and time.

$$CoV(x, y, t) = \frac{\sigma(EM(x, y, t))}{\overline{LE}^{EM}}$$

Line 140: Should be biomass standard deviation instead of variance? Also on line 145. Also check other places throughout the manuscript like on lines 225-229.

Thank you for noting this discrepancy. In response to two of the other reviewers' comments, we now use a machine learning approach to diagnose the drivers of changing phytoplankton biomass coefficient of variation. Therefore, this text has been removed throughout.

Line 169 - 175: It is not specified in the text that it is surface chlorophyll that you are using. Also, could you please clarify why you used surface chlorophyll for the validation of internal variability instead of biomass as in the rest of the analysis?

Thank you for this comment. We use surface chlorophyll for the validation of internal variability, because we are building on previous work in which we created a synthetic ensemble of observed

surface chlorophyll to emulate observed variability. We have revised the text to clarify that we used surface chorophyll in our analysis.

L169: "As an evaluation of the model's ability to represent internal variability (ensemble spread), we compare the internal variance in chlorophyll simulated in the CESM1-LE to a synthetic ensemble generated from observed surface chlorophyll concentrations over the MODIS remote sensing record (Elsworth et al. 2020, 2021) (Figure S4; chlorophyll was readily available in the CESM1-LE and can be directly compared with our synthetic ensemble of observed surface chlorophyll)."

212: Specify that it is surface chlorophyll.

We have specified that we are referring to surface chlorophyll throughout the manuscript.

L212: "We illustrate this by analyzing surface phytoplankton chlorophyll (rather than biomass; surface chlorophyll was readily available in the CMIP5 archive) from three other CMIP5 ESM large ensembles which include representation of ocean biogeochemistry..."

Review 4 of Elsworth et al. (2022) "Anthropogenic climate change drives non-stationary phytoplankton variance", submitted to Biogeosciences.

The authors are to be commended for taking on an interesting and important scientific challenge. However I do have a primary concern with the analytical framework applied using an MLR approach to evaluate the drivers of future changes in biomass variance. I believe that the methodology applied is sufficiently problematic that the primary result (grazing) is not convincing.

A biomass framework has been developed and applied for looking at marine ecosystems, namely that of Behrenfeld and Boss (2014; Annual Reviews), or more concisely the Behrenfeld framework. This method provides a quantitative and mechanistically-based means to connect biomass fluctuations with the underlying drivers at the timescales resolved by the model output (presumably monthly here). Within the Behrenfeld framework, decomposition into the underlying drivers is based on mass conservation using fields that should be saved for the model. For the case of the case of entrainment as it impacts biomass (a viable mechanism) there is not a convincing case that this will be properly represented by the fluctuations in annual mean MLD. This may be the case, but it is important for the authors to demonstrate this rather than simply assume it to be the case. Likewise temperature and nutrient dependencies in the model work into growth rates for phytoplankton, and the representation in the MLR approach is not convincing or state-of-the-art. The burden is really on the authors to justify their decision to use an MLR approach, rather than the mechanistically-based biomass framework of Behrenfeld.

I would urge the authors to apply the biomass framework of Behrenfeld to identify properly the relative roles of light, nutrient, and temperature drivers, with the analysis performed at the (monthly?) timescales resolved in the model output. Otherwise in my view the main ideas emphasized in the interpretation of the results are somewhat unsubstantiated.

We thank the reviewer for their constructive comments and appreciate their careful reading of the manuscript. The reviewer makes a general comment to revise our MLR analysis to make it more statistically robust. This also reflects the comments of Reviewer 1. We have addressed this by revising our MLR analysis to instead employ a machine learning approach and have extended the discussion section of the manuscript to reflect this revised analysis.



"Figure 5: Relative importance of predictor variables on phytoplankton biomass coefficient of variance across the RCP8.5 forcing scenario (2006 to 2100). Marine ecological regions are defined in Tagliabue et al. (2021). Regions were selected which aligned with the highest fisheries catch in the (a) Atlantic and (b) Pacific basins and the biogeochemically important (c) Southern Ocean and (d) Equatorial Pacific regions. The dominant phytoplankton functional type is considered in each region. In regions with a mixed ecological assemblage, total phytoplankton carbon is considered."

L133: "We quantified the drivers of phytoplankton carbon biomass CoV in key ocean regions by generating an ensemble of boosted regression trees. Unlike linear models, boosted trees are able to capture non-linear interaction between the predictors and the response. A regression tree ensemble is a predictive model composed of a weighted combination of multiple regression trees. At every step, the ensemble fits a new learner to the difference between the observed response and the aggregated prediction of all learners grown previously, aiming to minimize mean-squared error. We generate an ensemble of boosted regression trees (maximum tree depth = 10) using the Matlab function *fitrensemble*. Our predictor variables are the regional mean, ensemble mean temperature, mixed layer depth, incoming shortwave radiation, physically mediated iron, physically mediated phosphate, zooplankton carbon, and zooplankton grazing (diatom, small phytoplankton, or their sum) annually resolved from 2006 to 2100. We use the Matlab function *predictorImportance* to estimate the importance of the predictors for each tree learner in the ensemble; it computes the importance of the predictors in a tree by summing changes due to splits on every predictor and dividing the sum by the total number of branches."

L246: "We identify the importance of different predictors to changing phytoplankton biomass CoV in four distinct ecological regions using a machine learning (boosted regression tree) approach. In the subpolar Atlantic (ASP) and subpolar Pacific (SAP) ecological provinces (Figure 4 diatom

biomass CoV declines between the beginning and end of the century (Table 1). In the Atlantic subpolar region, the most important predictor of diatom biomass CoV is phosphate advection, with smaller contributions from zooplankton carbon (Figure 5a). In the subarctic Pacific region, sea surface temperature is the most important predictor of diatom biomass CoV, with phosphate advection playing a secondary role (Figure 5b)...'

L270: "Using a machine learning approach, we identify the importance of different predictors to changing phytoplankton biomass internal variability. In all four ecological provinces, a combination of bottom-up controls (e.g., nutrient supply, light availability) and top-down controls (e.g., grazer biomass) predict the decline in phytoplankton biomass CoV with anthropogenic warming..."

While the Behrenfeld and Boss, 2014 analytical framework is applicable to the study of bloom dynamics on the scale of days to months, our analysis relates to drivers of phytoplankton variance on the scale of centuries. The difference in relevant timescales between the studies result in different analytical approaches. This study does not attempt to address drivers of bloom dynamics.

The CESM1-LE does not provide output for phytoplankton growth rate. Without information about phytoplankton growth rate, we cannot reconstruct the change in phytoplankton or the change in phytoplankton variance over time using the analytical approach suggested from Behrenfeld and Boss, 2014. However, we have explored a more statistically robust approach to reconstruct predictors of phytoplankton variance in response to Reviewer 1's comment.