Ecosystem physio-phenology revealed using circular statistics

Response to reviewers

Dear Editors,

Please find below our responses to the reviewers. The comments by the reviewers are very relevant and will certainly help us to improve the quality of the manuscript. In the following we repeat the comments by the reviewers in bold and our response (RS) to each one in normal font.

Responses to Reviewer 1

Reviewer 1: In "Ecosystem physio-phenology revealed using circular statistics", Pabon-Moreno et al. have analyzed how the timing of maximum gross primary productivity is related to climate variables such as air temperature, solar radiation, precipitation, and VPD. They have analyzed 52 FLUXNET sites with more than 7 years of data and applied a circular regression method to (a) understand which environmental variable best predicts the timing of GPPmax and (2) measure the sensitivity of the response to each variable and (c) evaluate the method for different plant functional types. The topics is interesting, and the questions are relevant. The authors have also performed a simulation analysis to compare linear and circular regression methods, in particular given that some of the sites are in the southern hemisphere and hence may not be on the same calendar year as the northern hemisphere sites, the authors have justified circular regression methods are more appropriate than linear regression methods. The manuscript is generally well written and presented, however I have a couple major concerns related to the methods and conclusion that I strongly recommend being addressed by the authors.

We thank the reviewer for the accurate summary of our paper.

I am not sure how finding shortwave radiation is related to the annual trend of GPP is surprising. Especially with the not particularly high correlation values from the model outputs. My concern is that what the model predicts may be actually the average seasonality of the site, which is generally represented/regulated by the annual variation of solar radiation. I think it would have been more convincing if the model could predict "weird years" rather than normal years. So, one might argue that the model is tuned to track the seasonality of the sites with an average predictability power. See my next comment which is related.

The reviewer is right that shortwave incoming radiation (SWin) driving GPP - confirming this would be indeed no surprise. But please note that this paper does not analyze GPP: we are predicting the timing when GPP is maximized (the units we predict are "day of year" and not "g C /(m2 day)"). GPPmax represents the "optimum" ecosystem state when ecosystems are maximizing the uptake of CO_2 per year. We would also clarify that maximizing "predictability" is not our main aim as we are primarily interested in understanding the sensitivity of this state of ecosystem physio-phenology to climate variability. Given that radiation typically has a very low interannual variability we expect that the timing of GPPmax should be sensitive to other factors.

My other concern is that DOY values were directly used in the model as response variables. However, to analyze the inter-annual variability of the response, the anomalies should be used in the model. This is somehow related to the previous comment, as using site-specific model and absolute response values may result in obtaining the average annual trend and not the year-to-year variabilities. I think it would be best if the authors could use anomalies for each site as "y" in equation 2.

Equation 2 describes the circular linear regression where μ is the mean angular direction of a Von Mises distribution. As we mentioned in line 108 the mean angular direction is estimated via maximum likelihood. All interannual observations of DOY_{GPPmax} are used on the model, and the final result is constrained to a Von Mises distribution. The μ parameter cannot be removed from the equation, on the other hand the anomalies are considered into the amplitude of the Von Mises distribution (β) that is estimated internally.

Note that using absolute values in a consolidated model (all sites together) is another potentially good idea but that would detect the spatial (or site-by-site) patterns in the data rather than the temporal trends (which is the main question here).

The use of DOY values is necessary to quantify the sensitivity to the climate variables. On the other hand, if only consider DOY_{GPPmax} anomalies (outliers) the main research question regarding climate sensitivity can not be solved given that we will not analyze a representative sample of the observations. Considering only DOY_{GPPmax} outliers the research question should be more related to extreme events or temporal anomalies that as we mentioned in the previous comment are not the main topic in our study.

Minor comments:

There are also a few minor comments that I came across:

1- There is extensive use of parentheses in the paper that sometimes make the narrative hard to follow. I suggest avoiding unnecessary parentheses in the manuscript.

The paper will be revised accordingly.

2- The authors have used present tense throughout the manuscript at many places where past tense verbs are recommended.

This is a matter of “style” and we would like to keep the writing in present as we feel it is better to read.

3- line 141, “closed parenthesis” that was never opened

Solved. The parenthesis was removed

4- the narrative of the Results section can be improved, especially because the reader has to go back to the method to remember the terminologies and acronyms related to the method.

Thank you for this observation, we will revise the manuscript accordingly.

5- line 277: "Although the sensitivity of the DOYGPPmax to the climate drivers is site specific, it is possible to extrapolate the circular regression model for different sites with the same vegetation type and similar latitudes." That's a big claim. I'm not sure if the manuscript has provided convincing evidence with only 52 sites to support this.

Given that we used cross-validation to measure the performance of the model per vegetation type. We consider 52 sites should be enough to provide a robust analysis. On the other hand, 52 sites are the data available globally with at least 7 years of records.

6- What are the temporal windows for each predictor variable?

In our study the temporal window for the predictors is given by the half-time parameter of the half-life decay function (See Supplement 1. Half-time sensitivity analysis (System memory to explain DOY_{GPPmax})). In this section we run a sensitivity analysis to quantify how the change of the half-time parameter affects the correlation coefficient between the observed and predicted DOY_{GPPmax}.

Reviewer 2

General comments: In the manuscript "Ecosystem physio-phenology revealed using circular statistics", Pabon-Moreno et al. used a new method – the circular linear regression to estimate the timing of the maximum gross primary productivity (DOYgppmax) at 52 eddy covariance towers, and further quantified the sensitivity of DOYgppmax to a range of climate variables based on the results from this new regression. The manuscript is relevant to the topics of the journal. While I agree with the authors that circular linear regression has the potential to be a framework of future generalized phenology models, I have some doubts about the advantages of circular regression over the conventional linear regression approach, as well as the interpretation of results. It may need some substantial revisions. I apologize that I cannot be more supportive at this stage. I hope the authors can find this review helpful (please see below)

We thank the reviewer for sharing her considerations. The comments provided below are indeed very helpful.

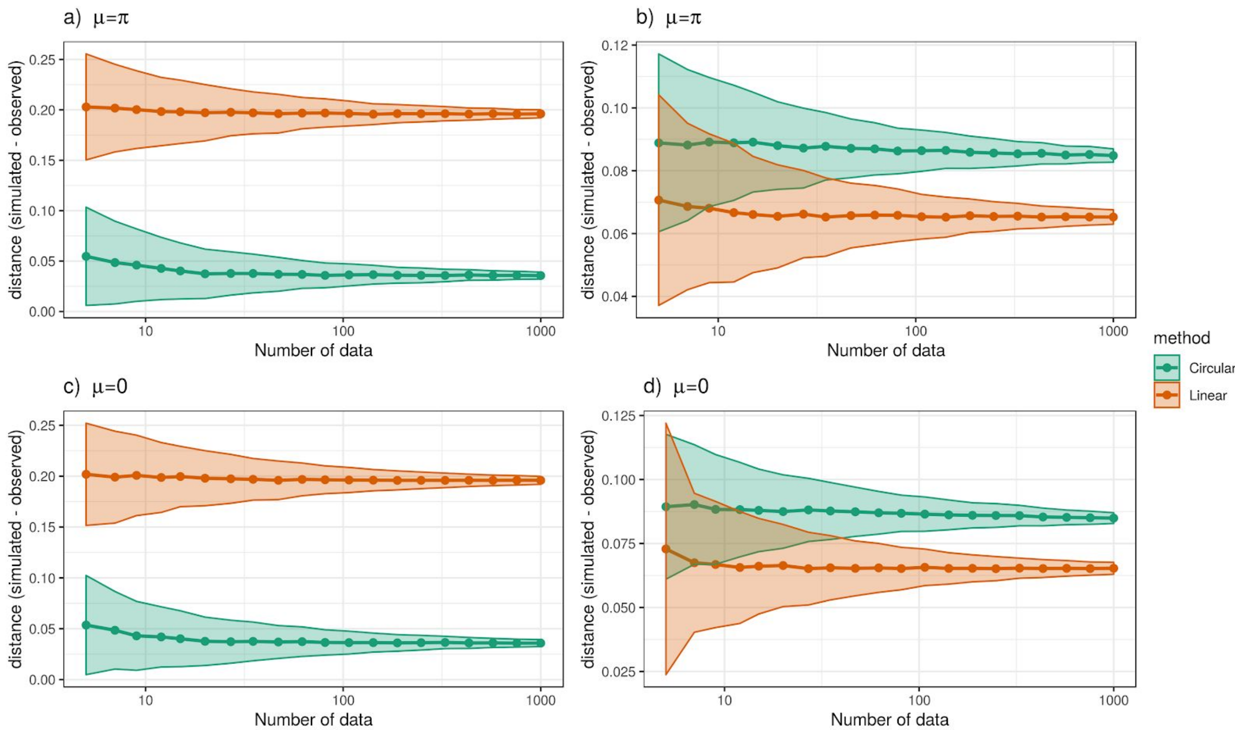
The authors introduced two advantages of circular regression 1) it is more accurate than linear regression 2) it can analyze the phenological event regardless of the locations of events, esp. for the southern hemisphere. For 1), I am concerned about circular reasoning, as the authors used two phenological events pre-defined by circular regression to compare

the performance of circular regression and linear regression, it is very likely the circular regression can outperform linear regression in this case. In addition, the author used the distance between observed beta and estimated beta to assess the efficiency of two models, and suggested that because the magnitude of distance for beta1 is larger than the distance for beta2, and the results on distance for beta1 favored circular regression, so circular regression is better. But the magnitude of distance for beta is also dependent on beta itself. Beta2 (0.3) is larger than beta1 (0.1), after normalize the distance of beta by beta, the result based on beta1 does not carry more weight than beta2, and the results on the distance of beta2 in fact favored linear regression.

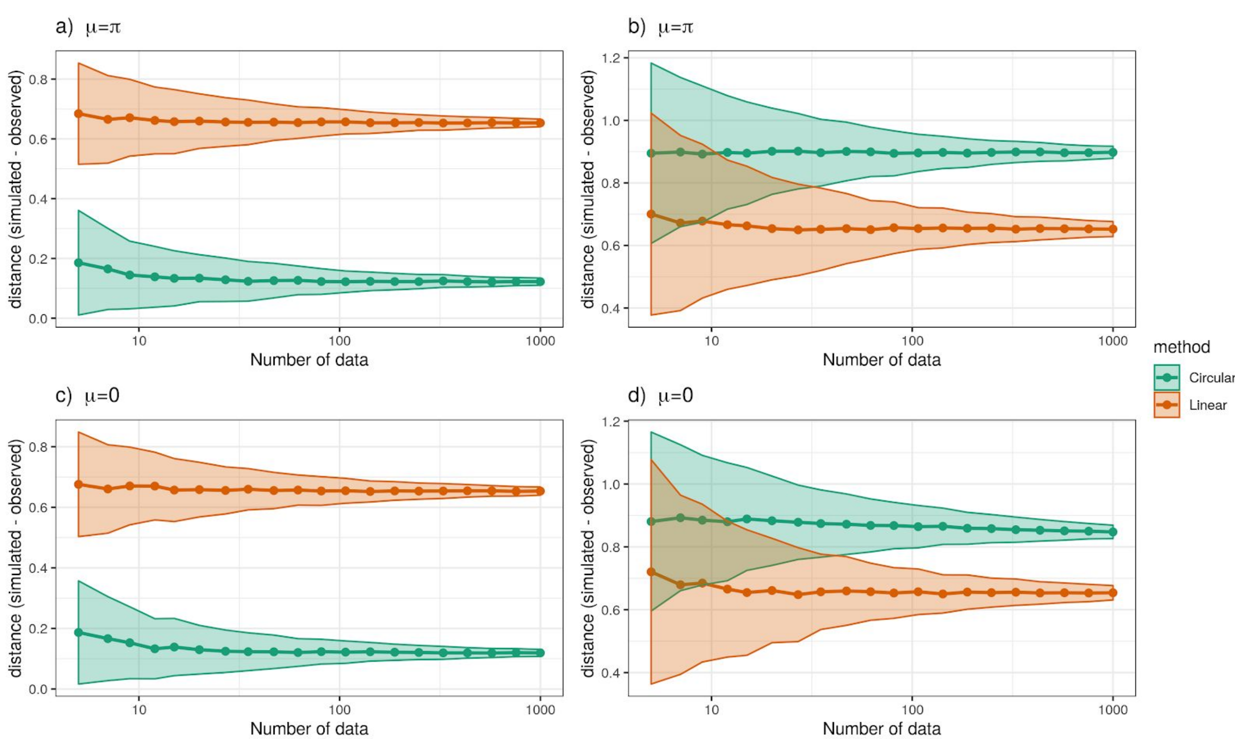
Regarding the first point. We used equation 2 to simulated the data. Nevertheless we are analyzing the performance to recover the original beta values of the equation and not the predictive power of the model. We used equation 2 given that linear regression does not allow to define the mean timing of phenological events. This is problematic especially when we want to analyze phenological events at the beginning and the end of the year.

Regarding the second point there is a misunderstanding related to the beta values. In our study $beta_1 = 0.3$ and $beta_2 = 0.1$ (Line 124). For this reason if we divide the distance by the beta values as suggested by the reviewer, the tendency of the results does not change (Please see the plots below). In both plots, circular regression has a better performance recovering $beta_1$, while linear regression has a better performance recovering $beta_2$ when the number of data is greater than 100. We will include the plot with the distances divided by the beta value to show that there is not a strong effect in the results. We will modify the line 135 of the methods by: "We estimate the difference between the recovered and the original coefficient divided by the beta value as the efficiency of the model (i.e. lower values mean higher efficiency).". And we will modify the line 166 of the results by: "Nevertheless, the differences between both regressions for $beta_2$ are of the order of 0.2 while the differences for $beta_1$ are of the order of 0.5."

Original Plot:



Plot normalizing the values per beta (distance / beta):



For 2), I am not sure why conventional phenology models cannot be used in the Southern Hemisphere (e.g. L208-209), say the degree-day model can be easily deployed if we know the temperature preceding budburst in Australia (e.g. Webb et al., 2008) and we can also get meaningful climate sensitivity of the event. Overall, I am not sure the circular model is superior to conventional models based on the evidence available in the manuscript.

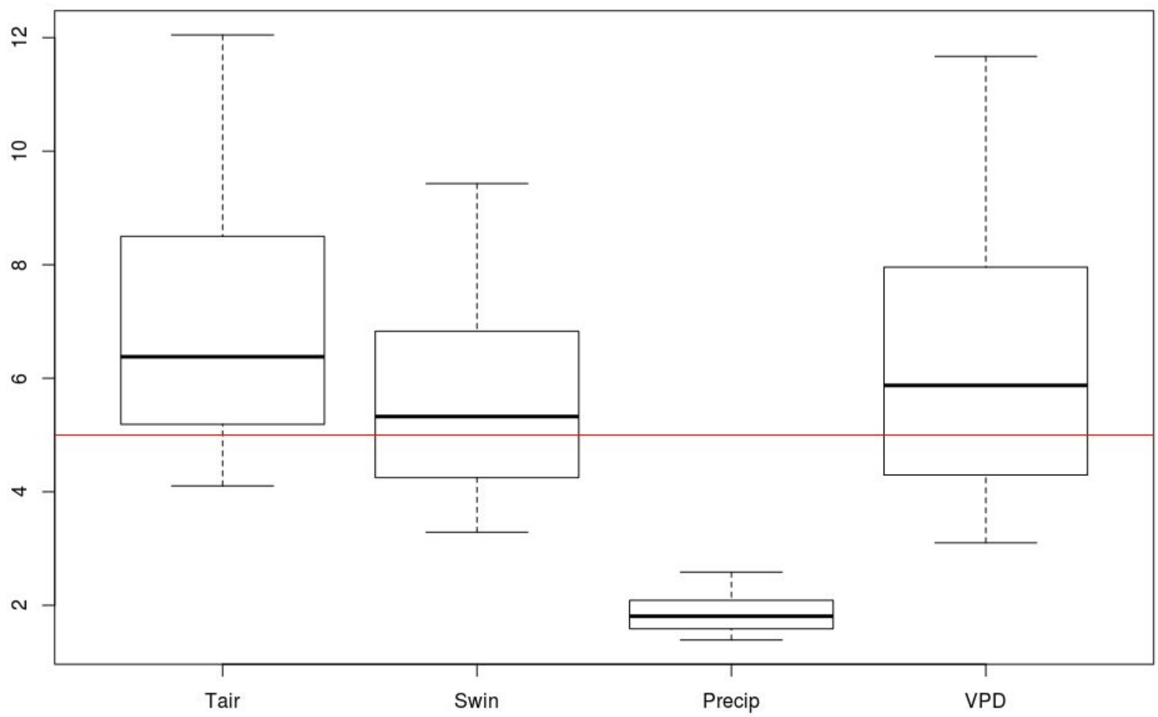
Please note that the only claim we make is that circular regression is more suitable than conventional linear models for analyzing phenological data - of course, process-based phenological models should outperform such statistical approaches. But our analysis reveals that we can learn the sensitivity to climate drivers in a purely empirical manner. In general, in any degree-day model there is a parameter to set an initial time to start accumulating warming. This will require again to define a t0 and in our view circular statistics could potentially avoid manual tunings of this kind.

Some questions over the interpretation of the results. First, I am a bit worried about overfitting of the model, as the leave-one-out validation suggest much less robust performance (r = -0.3 ~ 0.7) for PFTs compared to the r (r = 0.7 ~ 0.9 according to Table S1) we obtained using the training dataset.

As we mentioned in previous comments the main objective of the study was to analyze the sensitivity of DOY_{GPPmax} to different climate drivers. For this reason, each site has a unique "r" we consider that high r values are not an argument for dismissing the sensitivities of the climate variables. After cross-validation is expected that the predictive power of the model decreases, but the performance is not so bad considering it is estimated across vegetation types.

Second, at seasonal time scale, air temperature, radiation and VPD are all highly correlated with each other, how much can we trust their respective sensitivities estimated by circular regression. Wouldn't the sensitivity of air temperature be account for by the sensitivity of radiation if there is co-linearity between the two?

We performed a variance inflation factor analysis (VIF) for all sites-variables. The analysis shows (see plot below) that the colinearity of Air temperature, Shortwave Incoming Radiation, and VPD increases the variance of the regression coefficient. (VIF > 5). To solve this problem it is necessary to implement a PCA with these variables and run again all the analysis using the first axis of the PCA and precipitation as predictors of DOY_{GPPmax}. This means a major change in the manuscript given that the results of the sensitivity of DOY_{GPPmax} to climate variables will change. The respective discussion, and the conclusion need to be re-written. The revised version of the manuscript will contain these changes.



Third, I guess the so-called "memory effect" or "accumulated effect" of past climate is considered in circular regression through equation (1). Is this potentially one of the key differences between circular model and linear model? Does it mean the climate conditions closer to the event is more important than the climate conditions further back, and different climate variables are prescribed with different half-life here? I hope this part of the method is clearer.

We will add a better explanation:
"The idea of the decay function is that events in the present (DOY_{GPPmax}) are affected by past conditions (past climatic conditions). In this sense, the climatic conditions when DOY_{GPPmax} happens will have a weight of 1 to explain it. The day before will be less than the first day (e.g. weight of 0.8) and so on."

Fourth, the authors delegated the complex temperature sensitivity to consumption of available water (L240-). I am not sure there is a clear mechanistic underlying this link as there is no evidence supporting plant water uptake is related to temperature here. Soil water content may directly impact GPP (Stocker et al., 2018), it is not necessarily related to temperature, maybe VPD though. My major concern is about the robustness of the climate sensitivity identified in the manuscript.

The relationship temperature ~ water consumption is a hypothesis that we put forward to explain the non-predominant sign for the temperature coefficient. It is important to mention that GPPmax is different to DOY_{GPPmax}. The last one is the timing when GPP is maximized during the growing season. In this sense, the magnitude of GPP can decrease when soil water content decreases but this will not necessarily affect DOY_{GPPmax}. To clarify this point we will modify the legend of figure 8 from "theoretical" to "hypothetical"

Minor specific comments:

1. it is not accurate to say "(DOYgppmax). . . is the time the plants reach their maximum potential for CO2 absorption". GPP is the product of vegetation density (i.e. LAI) and the photosynthesis of individual leaves. When leaves have the maximum photosynthetic capacity/potential, it does not mean the whole canopy would be the most productive, as leaf photosynthesis can be downregulated by environment, and it also depends on how many leaves are there in the ecosystem.

The reviewer is right that our wording is not very accurate here. We will clarify that we are analyzing a metric at the canopy scale. We will write "the time when the ecosystem reaches its maximum potential for CO2 absorption".

2. Figure 1. In figure caption and in the text (L64), you mentioned each line represents the interannual variability. I feel it needs further clarification on how to read the figure. From what I understand, the distance between the line and the circle indicate the frequency of DOYgppmax, and the spread of linear may imply the variability of DOYgppmax.

We will modify the legend to “The distance between the color line and the circle represents the frequency of the DOY_{GPPmax} observations. The distance between the end and the beginning of the distribution represent the DOY_{GPPmax} interannual variability”

3. Method. Need more explanation about equation (1), as it not clear the meaning of x, N, N0, and the reason to include this half-life process here.

We will modify the manuscript accordingly.

4. I think the title of the paper might overshoot what in fact was done in the paper, since only one type of phenological event was studied, and I am not sure there is a pattern that really is “revealed” here that we can easily extrapolate for us to understand DOYgppmax due to the reported site-specific sensitivities. The concept of physiphenology is new to me, maybe the authors can provide a reference? I feel most conventional phenological events (e.g. budburst, leafout, leaf coloring, leaf senescence) are physiological changes of plants, so why they are not qualified as physi-phenology or do we really need this definition here. DOYgppmax sounds like a carbon uptake phenological phase.

We defined ecosystem physio-phenology as the temporal variability of optimum and basal ecosystem states in terms of the exchange of energy and matter between the ecosystem and the atmosphere. We defined DOY_{GPPmax} as a physio-phenological event because data is derived from the fluxes of the exchange of energy and matter between the ecosystems and the atmosphere and represents in a very accurate way the plants’ photosynthesis. Budburst, leaf coloring etc. are phenological state of plants that in specific cases not necessarily represent the physiological state of the plants (e.g. A green canopy does not necessarily mean that plants are photosynthetically active during winter). This limit between how much we can extrapolate between physiology and the light reflectance of the leaves is surpassed by the eddy covariance technique allowing us to quantify the ecosystem fluxes.

Although we only analyze one physio-phenological event (DOY_{GPPmax}), given that this study introduced the conceptual and methodological framework to analyze physio-phenological events we consider that the title is according to the research presented in the paper.

Regarding the comment: “I am not sure there is a pattern that really is “revealed” here that we can easily extrapolate for us to understand DOYgppmax due to the reported site-specific sensitivities". In section 4.2 "Sensitivity of DOY_{GPPmax} to climate variables," we summarized the effect of each climate variable at global scale.

Technical comments:

1. "2" in "CO2" is subscript

Fixed

2. please define "GPPmax" at its first appearance.

Fixed. Line 46

3. L201, according to Figure 7, GRA is -0.3 rather than 0?

Fixed

4. How to interpret the tendency in Figure 7?

In Figure 7 the tendency (blue line) represents the overestimation or underestimation of the model for specific DOY_{GPPmax} values.

5. L150, "leaf" to "leave"

Fixed

6. Table A1, maybe list the site according to their names or vegetation types. Now it is based on doi and not easy for readers to search sites.

We will modify the table showing the sites names by alphabetical order.

7. It would be helpful to condense figures in supplementary material 2 into a table, showing the sensitivity of each climate variable and significant level indicated by *. And please consider merging two supplementary materials into one.

Given that DOY_{GPPmax} sensitivity to the climate variables was estimated implementing bootstrapping, we consider that it is more important to show the distribution of the data than just the mean, also for the p-values. Regarding the second comment, we would like to keep the supplementary materials as separate.