Biogeosciences

Open Access

Discussions

EGU

BGD

Interactive comment

Interactive comment on "Summarizing the state of the terrestrial biosphere in few dimensions" *by* Guido Kraemer et al.

Anonymous Referee #2

Received and published: 19 September 2019

This is a very interesting paper addressing some important issues of big data analysis for ecology studies. It is rich in analyses and provides some new views on an old method (PCA). I particularly liked the analysis of trajectories that I found quite powerful, notably for case studies. Yet I found it difficult to understand what key research questions are addressed in this paper. This is important to clarify at the end of the introduction as the authors is providing us with a suit of analyses that may resemble (for non PCA-expert) an attempt of addressing many (all?) questions without real rationale. The readers need to have a clear (concise) view of the objectives of this paper, and they need to be guided through the analyses by referring back to the main research questions.

In addition, I also have a major concern related to the set of inputs data used to feed

[Printer-friendly version]

[Discussion paper]

the PCA. I agree that PCA is a powerful tool to deal with correlated variables, yet I have difficulties understanding why the authors have decided to include variables that are obviously highly correlated. To my opinion, vegetation productivity proxies are overrepresented as well as those related to water availability and stress. It puts some doubts in my head as to whether the finding of PC1 (primary productivity) and PC2 (surface hydrology) driving the state of the biosphere in space and time is truly original (or just purely mathematical). It is therefore important for the authors to justify the set of original variables. A suggestion could also be to decrease the number of input variables (removing obvious redundant proxies) as the amount of data to be condensed is mainly coming from the 8days interval used for the analysis.

Finally I also have other comments and concerns - notably related to the structure of the manuscript - that would need to be addressed by the authors prior publication of their research (see attached report for details).

— Detailed comments

(1) Abstract

The authors start off the abstract by mentioning the importance of detecting abrupt and gradual changes in terrestrial ecosystem but do not develop further in the introduction. In the method section, the detection of breakpoints reappears but no results are presented or discussed (except for the appendix A). The authors should decide whether to consider the detection of abrupt changes as a real research question for this study.

(2) Introduction

As stated in my main comment, I find that there is somewhat a mismatch between the introduction and the method section. In the introduction, the authors touch upon many issues related to assessing and attributing changes of biosphere properties. However apart from creating a new set of independent, 'essential' variables, they do not clearly mention what other research questions this study is going to address ; whereas in

the methods they mention PCA, trend and breakpoints analyses. Clearly stating the research questions for this study would help the readers to understand the rationale behind each analysis.

(3) Data and methods

The description of the data slightly too minimalistic, including in the appendix F. Mentioning the input data (satellite, climate or others) feeding into each dataset would be helpful. The observation period used for this study is also not mentioned.

L75. This statement is not always valid (e.g. in the case of equal-area projection). The sentence would be clearer if the authors would mention the projection system used here.

L77. The authors mentioned that they used a modified PCA, reading from the description given in the following lines, the PCA applied here seems to be standard. Could the authors provide some explanations to why / how the PCA has been modified? It should also clarify whether they applied the PCA in s or t-mode.

Per-pixel analysis. It would be nice here to make a link to the (extended – see comment above) research questions in order to understand directly the rationale for such analyses.

(4) Results

General comment: I highly suggest to split the results and discussion into two separate sections. It will facilitate the reading and will allow the authors to emphasise better the originality of their work. Example: L155-161, L164-173, L175-182, L235-246, etc. should not be in a results section s.s., but would rather belong to a discussion (or even introduction or method). Please consider at least moving all methods description and introduction to new concepts to the respective adequate sections.

L153 and Figure 1. The authors mentioned that there is a knee at component 2. I believe it is rather at component 3. This component still contribute to the total variability

to a share of almost 10%, therefore the authors should either include it in the rest of the analysis or provide an adequate justification not to.

Also I generally miss a figure presenting together the temporal and the spatial patterns for the main PCs. This could be put as supplementary material.

In the caption of Fig.1 I would recommend to change the term axis 1 and 2 by PC1 and 2. The comment also applies to the text itself (Ex. L190).

L183. Please describe in the first sentence what the triangle is made of.

L203. 'movement of a spatiotemporal pixel in variable space', please rephrase. A pixel cannot be moving spatiotemporally, like in a sliding puzzle.

L221-224. This should be described in the methods section and should be linked to a key research questions.

(5) Conclusion

L341. The results of the breakpoints analyses were not reported or discussed in the main text, therefore the statement 'To monitor gradual and abrupt changes in times of global change' do not hold.

Appendixes. Some results presented in the appendixes do not appear in the main text, e.g. Figures A1 and B1. The authors should maybe decide on the key results to be presented here and maybe save some others for a follow-up paper?

(6) Two final comments for reflexion:

- The authors have applied PCA on time series of 8day variables without considering any lag or accumulation effect in the response of a given variable. Would it be fair to say that legacy effects might not be captured adequately by such analysis?

- The authors refer to the MEI in the introduction as an example of a successful PCA-based indicator. Could the authors elaborate on the requirement for operationalising

their methods (e.g. if one would like to use the new indicators operationally, how frequently should the PCA be updated?).

Interactive comment on Biogeosciences Discuss., https://doi.org/10.5194/bg-2019-307, 2019.