Biogeosciences Discuss.,
https://doi.org/10.5194/bg-2019-297-AC3, 2019

Biogeosciences Discussions
Open Access
EGU

Interactive comment on "Causal networks of biosphere–atmosphere interactions" *by* Christopher Krich et al.

Christopher Krich et al.

ckrich@bgc-jena.mpg.de

Received and published: 16 October 2019

Response to reviewer 1

The manuscript presents the application of the newly developed PCMCI algorithm for the detection of causal links in geophysical data, focusing on biosphere/atmosphere interactions. Applications of the algorithm is done using flux tower eddy covariance data, for fine temporal scale analysis, and satellite derived NDVI along with climate data are used for global scale analysis at coarser scales. The topic is important and clearly within the scope of Biogeosciences. Identifying and quantifying causation in geophysical data is crucial for understanding the interplay of the processes involved and developing models. The

Printer-friendly version

Discussion paper

present manuscript is intended to my understanding to be primarily a proof of concept of the applicability of the PCMCI algorithm. While the paper is well written several parts need to be better clarified.

We thank the reviewer for the support of the manuscript and further suggestions and address the points below.

Specific comments:

- **I find the description of the algorithm on the paper slightly confusing (in particular for a Biogeosciences audience). I believe that the reader must refer to (Runge et al.,2018) to understand the basic principles behind the algorithm. I strongly suggest the authors to restructure and clarify this section. Simplifying the description and reporting the algorithm details as supplementary information could benefit the fluency of the manuscript.**

  We will improve the accessibility of the method by adding an introductory paragraph to the method section. Here we will explain how PCMCI relates to existing methods, which will help the reader gain a more intuitive understanding of the aim and concept of PCMCI. A detailed description including mathematical notations will be then given in the following subsections.

- **The synthetic test developed to quantify the skill of the algorithm, when its assumptions are not valid (e.g. seasonality, heteroscedacity) is clearly important. However, it is not currently clear how the results derived from this analysis can be generalized beyond the Hainich site. Emphasis should be given on the transferability of the magnitude of expected biases at a global scale.**

  We believe this may be a misunderstanding. The synthetic test did not focus solely on the Hainich site. The impression might be given, as a few plots (time series and networks) are showing as an example the results from Hainich. We

used radiation data from 72 FLUXNET sites to run the model. Therefore, the results are not bound to the conditions at Hainich.

To clarify this we will change within the first paragraph of section '2.3.1 Artificial Dataset - Test Model' the sentence:

> Using time series of measured global radiation (Rg) we created three artificial time series that conceptually represent temperature (T), gross primary production (GPP) and ecosystem respiration (Reco)."

To:

> The artificial dataset was created using a test model which takes time series of measured global radiation (Rg) and creates three artificial time series that conceptually represent air temperature (T), gross primary production (GPP) and ecosystem respiration (Reco).[....]

And we will add as well:

> The model was fitted to real observational data (using radiation, temperature and land-atmosphere fluxes) of daily time resolution, measured by eddy-covariance method (Baldocchi et al., 1988; Baldocchi, 2003) from FLUXNET, by minimizing the sum of squared residuals using the gradient descent implemented in the Optim.jl package (Mogensen and Riseth, 2018). We fitted the model to 72 sites listed in Table B1 given in the Supplementary Material section.

We think that this adjustment should clarify that we are not focusing on the Hainich site only.

- **In a broader sense, a key question is why would the authors use a procedure for the identification of causal links, when the basic assumptions of the procedure are violated by the data?**

  Every statistical method comes with a set of assumptions which are required to assure interpretability of the results. In a real world study case, those assumptions are typically not fully met. That is in fact a justification for why we conducted the present study. We tested the limits of applicability of a causal inference method in the case of two commonly measured environmental variables. PCMCI in conjunction with ParCorr possesses several attributes which favour its use, as shown in our analysis: high detection power, interpretability, and computational efficiency. We aimed to identify whether PCMCI and ParCorr could in fact deliver reliable results using artificial and real data.

- **I believe the authors should better explain why the proposed algorithm in more efficient than e.g. the spectral Granger causality algorithm proposed earlier by Detto et al.(2012), which, to my understanding, is inherently non-parametric and not sensitive to periodicities.**

  Thanks for pointing out the need for a more detailed comparison to other methods. This is basically in line with Reviewer 3 who, however, suggested the comparison to another approach. We totally agree on the importance of such comparisons but regard numerical analyses evaluating other methods to be out of scope of this paper. Nevertheless, we agree that we could better describe the specific characteristics and benefits of PCMCI in conjunction with the independence test ParCorr.

  PCMCI is computationally efficient, has proper significance testing and attribution of link strengths. As we observe, the seasonality is an obstacle for PCMCI and might be better handled with a spectral method. But, to the best of our knowledge, spectral Granger causality is also bound to causal stationarity. Thus, within an analysis spanning (possibly) several years and seasons, one would assume a

constant network structure throughout time (same network structure in winter as in summer). However, this is not the case, as we could show in our analysis of the Majadas ecosystem.

We will stress the specific characteristics of PCMCI in a newly added paragraph of the method section.

- **In Fig2 it is shown that with an increasing sample size, the fraction of falsely identified causal links increases, when the algorithm's assumptions are not valid. This can be a significant drawback as the best datasets (i.e. with long records), are more prone to errors. The authors should better discuss this.**

  To analyse long record time series data one has to question whether causal stationarity is fulfilled. As we have seen in the example of the Majadas dataset, strong differences in the dependence structure occur for different months of the year. Accounting for causal stationarity still does not conflict with analysing long data records with PCMCI, because by applying a proper mask one can estimate the network structure for one month or season within multi year time series (which is similar to our mask used in the Majadas dataset: taking only noon values but for consecutive days). Such a mask increases stationarity and therefore reduces an inflated false positive rate.

  This is partially discussed already on page 20 line 3 to 20. To further clarify we will modify the submitted paper as follows:

  We replace line 13 and 15:

  > Much of the influence of heteroscedasticity is also removed when limiting the analysis to a specific time period, i.e.season, which makes the data causally stationarity (cf. Sect. 2.1). For example, the link from radiation to GPP vanishes in winter as there is mostly no active plant material left.

With:

> The increasing FPR with increasing time series length can further raise doubts regarding the analysis of long time series. For such an analysis, though, the assumption of causal stationarity should first be assessed. This is typically done by limiting the analysis to a specific time period (i.e. 'masking'), e.g., a specific season, month, or time of the day. For example, the link from radiation to GPP vanishes in winter as there is mostly no active plant material left. Such masking reduces further influences of remaining seasonality or heteroscedasticity.

- **If I understood correctly, it is shown that for the baseline case, the algorithm cannot robustly identify the true causal non-linear links (as expected). However, the identification rate increases incorporating seasonality (which I presume also violates the stationarity assumption of the algorithm), which is counter intuitive. The authors in their discussion attribute this behaviour to the variance of the parent variable. Can the authors discuss how this artefact can limit the range of applicability of the procedure for global scale applications (i.e. differences in regions with distinct seasonality or not)?**

  Seasonality constitutes a common driver in this model. In general, such common drivers increase the dependence among the variables and, hence, lead to a higher detection rate for true links (TPR) as well as a higher false positive rate (FPR) for absent links since this driver is not conditioned out properly. Therefore it is not counter intuitive that both the TPR and the FPR rate increase in the seasonality model. To reduce the effect of seasonality further, we suggest to use a mask or use deseasonalized time series.

  We will add such a explanation in section 4.1.

Interactive comment on Biogeosciences Discuss., https://doi.org/10.5194/bg-2019-297, 2019.