Geosci. Model Dev. Discuss.,
doi:10.5194/gmd-2016-266-RC2, 2017

Geoscientific Model Development Discussions

Interactive comment on "A non-linear Granger causality framework to investigate climate–vegetation dynamics" *by* Christina Papagiannopoulou et al.

Anonymous Referee #2

Received and published: 21 February 2017

Title: A non-linear Granger causality framework to investigate climate–vegetation dynamics

Authors: Christina Papagiannopoulou, Diego G. Miralles, Niko E. C. Verhoest, Wouter A. Dorigo, and Willem Waegeman

General Comments: Reviewer summary: The manuscript presents a non-linear Granger causality analysis to investigate climate-vegetation interactions. Anomalies of the normalized vegetation index (NDVI) are analyzed in conjunction with a full set of climate variables taken from re-analysis, in situ, and satellite observations. The data provide multi-decadal global coverage for water availability (precipitation, snow water equivalent and soil moisture data), temperature, and radiation. All data spans the

period 1981-2010 at the global scale and has been converted to a common monthly temporal resolution and 1x1 degree spatial resolution. At each pixel the NDVI data is considered the response and the climate data the predictor variables. A moving window of twelve months is used to determine if the climate data granger-causes the NDVI value. Analysis is performed on NDVI anomalies computed by subtracting the corresponding monthly expectation from the de-trended time series. The climate data as well as cumulative values and extreme indices calculated from the climate data were included as predictor variables. The non-linear Granger causality uses a non-linear random forest model, and is shown to explain more of the variance than the linear granger analysis.

Article contribution and overall impact: This study makes an effort to use multiple climate data sources to tease out predictability for vegetation anomalies. The authors highlight improvements with the non-linear method compared to traditional granger causality, as well as the importance of using extreme events. The discussion would benefit from a more explicit discussion of the uncertainty associated with the climate datasets used as predictors. Given that this study precedes or supports Papagiannopoulou et al (in review), more discussion of those results and their importance would be useful as that study is not available to the reader. Specifically, the follow-on study highlights the importance of specific climate predictors for particular regions. It is not clear how those variables are chosen from the many climate predictors, and it would be useful to provide an example in this manuscript to highlight the strength of this method with a clear detailed regional example.

Detailed comments:

Page 1 line 17-18: Should this read "predictions of vegetation in response to future climate can be improved through a better understanding. . ." ? as you are looking for climate drivers of vegetation.

Page 2 line 22: define "higher-level features" here and throughout manuscript. It is not

clear what these are. (Pg 11 line 4, pg.15 line 2)

Page 2 line 24: define “higher-level climate variables” not clear what this is.

Page 3 line 2-7: May not be necessary to include full definition of R2.

Page 3 line 30: update “might lead to wrong” to “might lead to incorrect”

Page 12 line 15-23: Are the results for all variables, or the most predictive variable, or a set of variables at each pixel?

Page 12 line 26-27: Why is this chosen as the minimum? Please explain or provide citation.

Page 13 line 10: by what margin is the uncertainty larger in these regions, and for what reasons? Are you referring to all the climate variables, if not please qualify. The citation references error for soil moisture. Add citations, which support the amount of uncertainty in these regions for the remaining data types.

Page 13 line 7 to bottom and page 14 line 1-4: Move this to discussion.

Page 14 line 1: Update to “vegetation anomalies are not necessarily”

Page 14 line 7: Use different phrasing for “unambiguous”

Page 14 line 7-12: move to discussion.

Page 14 line 8-10: Recommend re-wording this. The limit for figure 5 and the presentation of the non-linear analysis is still to a limit of R2 = 0.4 as in figure 4? An R2 of 0.4 does not seem like a strong correlation. Though figure 5 is improved from figure 4 there are large portions that show no improvement, and the overall explained variance is below 40% in most regions.

Page 14 line 10: “comparison between figs 4b and 5b” explain in more detail. It would be easier for the reader to compare these if they were in one figure block, or on the same page.

Page 15 line 5: Please provide more detail about this study. It comes up frequently in the manuscript, and a larger summary with details (supportive numbers or examples from regions) would be helpful since we do not have access to the manuscript.

Page 15 line 11: Has a test been run with only the anomalies and extremes? Would that sub-set of predictors provide strong predictive performance?

Page 16 line 1-2: Provide more detail from supporting manuscript for current manuscript. It is necessary to support this analysis that you can separate specific drivers.

Page 16 line 3-6: Connect this sentence to the following paragraph.

Page 16 line 17: Is the “framework” the non-linear component? Maybe just call it that – non-linear, rather than a framework. This implies a more complex process.

Page 17 line 11: explain “feature construction”

Page 17 line 16: update word order to read “causality based approaches indicate”

Interactive comment on Geosci. Model Dev. Discuss., doi:10.5194/gmd-2016-266, 2016.