Review of gmd-2018-92

The work presented by Papagiannopoulou et al. in this manuscript is of interest for the reader of GMD and is also very relevant for the ecosystem and climate research community. Overall the manuscript is well structured and the methodology section generally well documented. Knowing that the focus of GMD is on the progress and novelty in computation and model development, I support the need for in-depth description of the MTL model and its performances (e.g. SLT vs MLT, capability to detect Granger causality, etc.). However I believe that the manuscript would be strengthened and results better supported if the authors could really demonstrate that the new product (i.e. map of hydro-climatic biomes) is outperforming other bioclimatic maps that did not consider in their design the vegetation response to climate variability. This is still lacking in the current manuscript. In addition some methodological aspects that led to the final design of the MLT and clustering should also be improved to backup the authors' statement on the performances of the final models and derived product. Based on these observations and on the detailed comments provided below I recommend the paper for major revision.

Specific comments

Introduction:

- Studying vegetation response to climate variability is and has been the focus of numerous researches. I know the objective of the authors is to create a new bio-climatic map, however I can imagine that their work build up on recent developments in science regarding ecosystem response to climate variability. This is not well reflected in the introduction. Please add some references to key papers, studies in the matter. Some suggestions below:

Liu L, Zhang Y, Wu S, Li S, Qin D (2018) Water memory effects and their impacts on global vegetation productivity and resilience. Sci Rep, 8, 2962.

Seddon AW, Macias-Fauria M, Long PR, Benz D, Willis KJ (2016) Sensitivity of global terrestrial ecosystems to climate variability. Nature, 531, 229-232.

De Keersmaecker W, Lhermitte S, Tits L, Honnay O, Somers B, Coppin P (2015) A model quantifying global vegetation resistance and resilience to short-term climate anomalies and their relationship with vegetation cover. Global Ecology and Biogeography, 24, 539-548.

Nemani RR, Keeling CD, Hashimoto H et al. (2003) Climate-driven increases in global terrestrial net primary production from 1982 to 1999. Science, 300, 1560-1563.

- The authors claim (p2, l23) that it is the first time that ecoregions are being defined based on the analysis of vegetation response to climate variability. I agree that the idea is relatively novel and definitely relevant. Yet previous attempts have been made, notably by combining PCA and clustering techniques applied to climate and vegetation dataset. See the following reference as an example: Ivits E, Horion S, Fensholt R, Cherlet M (2014) Global Ecosystem Response Types Derived from the Standardized Precipitation Evapotranspiration Index and FPAR3g Series. Remote Sensing, 6, 4266-4288.

Methodology

- **Sect. 2.4.** The authors mentioned that the ASO method used here should not be confused with PCA. It would be useful to develop this statement. Indeed for both techniques orthonormal vectors are derived from the high dimensional feature space, creating a new 'optimized' low-dimensional feature space. The authors mentioned that the goal of the ASO method is to detect the PC of the predictive structure. Knowing that PCA can be performed in two ways (t-mode and s-mode), the t-mode being the most frequently used by climatologist to identify recurrent spatial patterns over time, whereas the S-mode allows for identifying recurrent temporal patterns over space. How would the current method differ from an extended PCA in S-mode? I can imagine that using EPCA over a dataset as large as the one used here could be a real challenge for example. But I would like the authors to elaborate on the pros and cons of the new method as compared to already established techniques in the climate research such as PCA/EPCA for example.
- **Sect. 2.5.** The authors do not give any name or reference for the clustering technique used here. Please clarify if a new algorithm has been developed for the study or if an already developed clustering technique was applied.
- **General comment on the use of R2 for assessing the model performance**: at several occasions (in the manuscript and in the supplementary material), the authors used r2 to quantify the performance of different models (MLT vs. SLT, models with and without Granger causality, inclusion of higher-level features in the input dataset, final decision on the number of clusters). They generally conclude that the best model is the one with the highest r2. I agree on the principle, however looking at the differences between r2 (e.g. figures 3b and 3d, large areas present difference in r2 below 0.1), I wonder whether all these differences are statistically significant. As based on the analysis of r2, the authors are deciding on the final set of input data, the final design of the MLT model, and the final number of clusters, I would really urge the need for further statistical assessment of the model performances. One first analysis could simply be to estimate the percentage area of pixels with statistically significant increase in r2.

Results:

- General comment on the final number of clusters: the fact that the majority of the Iberian Peninsula is included in the transitional energy driven cluster together with Ireland, an important part of SE Asia, part of Brasil and Venezuela – Colombia makes me wonder if a higher number of clusters would not be more appropriate. The authors mentioned already in Figure S2 that the differences in the predictive performance for h= 6 - 15 are marginal. Further assessments should therefore be performed in order to identify the optimal number of hydro-climatic biomes. Part of this assessment should be dedicated to the understanding of the actual drivers (main predictors) for each biome. I believe providing a solid justification for the naming of the different biomes (by referring back to the main predictors) would be beneficial for the paper.
- In relation to the previous comment, how does the new global map of hydro-climatic biomes perform as compared to previous ones (not including information of vegetation condition and response to climate)? It would be really interesting if the authors could showcase for one (or more) bio-climatic

zone how the new bio-climatic zone provide a finer, more accurate picture of global terrestrial biomes by analysis the specific (/sub-local) ecosystem response to climate variability. To this regard, the bio-climatic map produced by Metzger et al. (see reference below) could also be of interest for comparison.

Metzger MJ, Bunce RGH, Jongman RHG et al. (2012) A high-resolution bioclimate map of the world: a unifying framework for global biodiversity research and monitoring. Global Ecology and Biogeography, 22, 630–638.

- Figure 4. (c) The Koeppen classification divides the world into 5 main classes and 29 sub-classes. The authors should justify the use of 10 classes in the figure. This can be very misleading when looking and interpreting the results. An example: I do not think that the statement p14, l21-23 '… the region of North Asia is coherent in terms of climate, but quite diverse in terms of vegetation types ; the hydro-climatic biomes show a clear distinction from shrublands (…) to coniferous …' holds entirely when looking at the high level details (29 classes) of the Koeppen classification. Please justify your choice here.

- Supplementary material S4. The authors indicate that the best-formed clusters are depicted in FigS4a (hence by the hydro-climatic biomes). I find very difficult to make any final judgment of the best "depiction" (/detection) of biomes based on the 2dimensional graphs provided here.

Technical comments

- P5, l14: please add a reference for the statement: '… this kind of modelling is becoming more common in climate science…'

- P10, l10: please clarify what you mean by multi-month vegetation dynamics. Is it seasonal, sub-seasonal, yearly?

- P12, l5: please correct 'Geanger' with 'Granger'

- Figure 4. (a) the color code for the clusters sub-tropical energy driven and mid-latitude temperature driven are too similar. It is difficult to differentiate them. Please adjust the color scheme of the legend.

- p15, l22: The term 'turning point' has only been introduced recently in ecosystem and climate science so for clarity, you can refer to:

Horion S, Prishchepov AV, Verbesselt J, De Beurs K, Tagesson T, Fensholt R (2016) Revealing turning points in ecosystem functioning over the Northern Eurasian agricultural frontier. Glob Chang Biol, 22, 2801-2817.