Biogeosciences Discuss.,
https://doi.org/10.5194/bg-2019-252-RC1, 2019

Biogeosciences Discussions

Interactive comment on "Spatio-Temporal Variations and Uncertainty in Land Surface Modelling for High Latitudes: Univariate Response Analysis" *by* Didier G. Leibovici et al.

Anonymous Referee #1

Received and published: 13 August 2019

General comments

The authors present an analysis of spatio-temporal agreement and differences between predicted Net Primary Production (NPP) from four land surface models (LSMs). The main method is an extension of singular value decomposition to multi-dimensional matrices/tensors. The authors consider how the method provides a dimensionality reduction of the data to provide Principal Tensors (PT) which represents the dominant patterns in the data set. The manuscript is generally well written with a good description of the application and about the right level of information on the four LSM models.

I think the balance between the description of the method, presentation of the results,

and the interpretation of the results. First, I'm not sure whether such a lengthy description of SVD and then the PTA k method is required for the intended audience. Inconsistent naming also complicates the methods section, for example, referring to the data set as a "data table" and then to as "tensors". The term Principal tensor (PT) is also not introduced in the methods (e.g Pg 8 L12) but is referred to throughout the text. Introducing PT as the extension of the singular vectors to higher dimensions around Pg 8 L12 would help clarify this.

More importantly, the manuscript would benefit from a greater focus on interpreting the PTs and how these explain "uncertainty" in LSMs. There is not much reference to uncertainty in the manuscript. In the presentation of the results there is a good description of the spatial patterns in each PT but not the interpretation of what each PT and it's singular value means for the uncertainty arising from the different LSMs. For example, that one of the PTs is essentially a "spatial correction applied to ORCHIDEE-HLveg" provides information on the biases between LSMs but this point doesn't come through very strongly. The discussion about this (Pg13 L10) makes the point that OR_HL has a temporal evolution which contrasts to the other three LSMs and this discussed somewhat in the conclusions again. But so the question is this a bias in OR_HL or a representation of the uncertainty (in terms of the spread) in the temporal evolution of NPP?

Similarly, if the first principal component explains around 90% of the variance in the data and it has "only a weak dependence on the LSM", I assume that there is quite good agreement between the LSMs and therefore the uncertainty (i.e. spread) arising from different LSMs isn't that large? If that's true, I think considering the LSM-weights for each PT provides more information on the spatio-temporal patterns of uncertainties. Would this reveal whether the main patterns are driven by the difference in LSMs (weights) or other variances (annual variability in the forcing data for example)?

Section 5 focuses on more on the differences between the LSMs by repeating the method on normalised NPP differences which provides the link to the relative spread

of the NPP estimates. But I still think the manuscript is missing that final link between what the PTs explain about different types of uncertainties (i.e bias and precision) and their relative importance (as indicated by the explained variance?). Some of these inferences are found throughout the manuscript but a dedicated section for relating the PTs to uncertainties (in the most common sense as the precision and any biases of an estimate) would help this.

Relatedly, it would be interesting to know how much more informative the method is for inferring this information on uncertainty than something comparatively simpler such as the spatial and temporal patterns in the coefficient of variation (or some other dispersion metric). This could potentially be just a qualitative comparison to show how the method provides more information about the LSM differences than a quantity which is more commonly considered an "uncertainty".

Finally, a greater discussion of how the PTs and information could be used for quantifying uncertainty for CSI models would be a good addition. For example, does the method essentially provide a representation of the data (in terms of uncertainty ie mean and covariance matrix) which could be used with these models? Or would the original NPPl data have to be propagated through the CSI models to provide uncertainties (i.e. ensemble-based or some moment propagation method)?

Specific comments and minor corrections

Pg.1 L.5 " ...will have different impacts". Impacts on what?

Pg 6 L 2 multi-variate → multivariate

Pg.11 L.4 singulat → singular

Pg.26 L13 Although this analysis was only carried out for JULES, there is no reason to expect different findings for the other LSMs". Why not?

Interactive comment on Biogeosciences Discuss., https://doi.org/10.5194/bg-2019-252, 2019.