Geosci. Model Dev. Discuss.,
doi:10.5194/gmd-2015-250-RC1, 2016


Geoscientific Model Development Discussions
Open Access
EGU



# *Interactive comment on* "A new metric for climate models that includes field and spatial dependencies using Gaussian Markov Random Fields" *by* A. Nosedal-Sanchez et al.

**Anonymous Referee #1**

Received and published: 26 February 2016

Summary Statement: Nosedal-Sanchez et al introduce Gaussian Markov random fields (GMRF) to account for co-variations across space and between fields while comparing climate model output to observations. These types of co-variations are potentially interesting to GMD readers and important to consider in evaluating climate model performance because assumptions of independence can lead to overconfidence and incorrect statistical inferences. While the presentation of the GMRF is reasonable and the illustration of the method to CAM3 data informative, there are areas where the manuscript can be improved, both in terms of readability and technical clarity. A number of items are listed below in random order. After addressing these items, some of which may require revisions, I should be in a position to recommend this manuscript



Printer-friendly version

Discussion paper

for publication in GMD.

Item 1. Conditional independence is an assumption underlying Markov random fields. For three variables A, B, and C, the joint probability distribution of A and B conditioned on C, written as p(A,B|C), can be factored into the product p(A|C) times p(B|C) for all values of C if A and B are conditionally independent of C. The authors should argue, or preferably demonstrate, that the necessary conditional independence properties approximately hold for the application of their method to climate model fields. The feedbacks across scales in the climate system and the coupled nature of the physical equations may serve as a basis for some degree of conditional independence, though I expect some cases where p(A|C) is a poor approximation of p(A|B,C) as implied by conditional independence.

Item 2. The authors appear to neglect temporal relationships in the method and example, even though such relationships are prevalent in the climate system. A perturbation in the pattern of sea surface temperature in the tropics, for example, may take months before the signal shows up in the spatial distribution of precipitation in the mid-latitudes. While introducing temporal correlations into their method is beyond the scope of the manuscript and not required at this stage, it would still be beneficial to readers if the authors described how their method could be extended is this way.

Item 3. The opening paragraph states that there is skepticism in using a scalar metric to assess climate model performance. This gives the impression that everything gets boiled down to a single number, which isn't the case. Climate models are often assessed using a vector of scalar quantities (e.g. as in Gleckler et al), a scalar measure of a vector field, or combinations of these and other metrics. A single field projected on a Taylor diagram, for example, considers two orthogonal scalar quantities (centered rms and correlation). Please clarify the description.

Item 4. The opening paragraph also describes the need to account for spatial and field dependencies. Field dependence is an essential feature of your methodology, so it would be useful to readers to provide a specific example of what you mean by field dependence early in the introduction.

Item 5. The first sentence in the second paragraph in the introduction is a little awkward and should be rewritten (lines 12-15, page 3). There is an observational record of climate, but not an observational record of a climate model. Moreover, this statement seems to suggest that data assimilation is primarily a data imputation method, which it really isn't. Data assimilation minimizes the differences between the model state and observations, while insuring that the state fields abide by conservation laws (mass, energy, and momentum) and other important physical dependencies. This paragraph overall seems to imply that the models don't do a very good job with the dependencies, which is arguable. I have confidence that the models are getting many of the large scale dependencies about right (e.g. equator to pole gradients, land-ocean contrasts, temperature dependence of water vapor through Clausius-Clapeyron, etc).

Item 6. It's a good idea to present the general idea behind the metric in equation (1) in the introduction, though I found myself flipping back and forth between the introduction and section 2 to make better sense of the information. To make it easier for readers to get through the introduction without getting hung up on details, perhaps you could introduce the concept in more general terms. Also, the symbol Z is used for the metric in this section, but it doesn't appear elsewhere in the manuscript and should be dropped. And the times symbol in 'n_obs x n_pts' on line 3, page 4 suggests that v is a matrix with n_obs rows and n_pts columns. I recommend changing it to n_obs n_pts.

Item 7. There is a typo in the lower right element of the S inverse matrix on line 11, page 7. The sigma index should be 22, instead of 11. For consistency, use the same indices for off-diagonal terms (e.g. S12 is used for the lower left term on line 11, while S21 is used on line 13 page 7).

Item 8. Regarding the alpha parameter, please provide references or further information about the statement that alpha depends only on the geometry of the neighborhood and not on the details of the fields. I have other questions about alpha. How much does it vary going from a first order neighborhood to a higher order neighborhood? Can alpha be extended from a scalar to a vector to optimize the covariances in different regions in the neighborhood?

Item 9. The witch hat plots are convenient, but take some effort to get used to. It would be useful it you first stepped the readers through the concept with a simple example. How much does the shape of the witch hat depend on the selected indexing for the neighborhood? E.g. swapping x3 and x4 in figure 1 appears arbitrary, but results in a different Q. Does the averaging of the cells for a given distance from the diagonal hide information that could be important? Are there other simple ways to show the differences between the empirical and GMRF estimates (e.g. Hinton diagrams)?

Item 10. Figures 3 and 4 are positioned before section 4 in the manuscript, but the figures rely on information about the climate model data from that section (e.g. the estimates are from 15 samples). Please cross reference the material from section 4 where needed to avoid confusion.

Item 11. In the last paragraph on page 11, the authors state that the only meaningful correlations are of TREFHT with PSL and PRECT with PSL. However, if TREFHT and PRECT are individually correlated with PSL, shouldn't TREFHT and PRECT also be correlated to each other? Moreover, there is a contradiction between the physical explanation on lines 23-25, page 11 and the sign of the correlation between PSL and PRECT (low pressure systems increase precipitation).

Item 12. The model simulations use prescribed sea surface temperatures, which strongly constrain the near surface air temperature, so it seems surprising that the biggest changes in cost are associated with the 2-m air temperature. Can the authors provide a physical explanation for their finding?

Item 13. The authors find that spatial dependencies are more important to capture than field dependencies for the four selected outputs (PSL, TREFHT, U, PRECT). Do they have any reason to suspect (or can they show) that the field dependencies will dominate over spatial dependencies for other fields? If not, then this suggests that it may not be critical to capture the field dependencies and that their method does not offer many clear benefits over standard model assessment techniques. From their example in figure 5 of optimizing model performance by changing two parameters (c0 and Ke), it even looks like adding the spatial dependence alone would not greatly affect the conclusions drawn from assuming spatial independence (i.e., that high values of c0 and low values of Ke are best).

Item 14. Observational uncertainty does not appear to be taken into account in their method. Can the authors comment on and suggest ways to incorporate observational uncertainty into their test statistic?

---

Interactive comment on Geosci. Model Dev. Discuss., doi:10.5194/gmd-2015-250, 2016.