Geosci. Model Dev. Discuss.,
doi:10.5194/gmd-2015-250-RC2, 2016


Geoscientific Model Development Discussions
Open Access



# *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 #2**

Received and published: 11 March 2016

This study computes multivariate cost functions which can be used to choose optimal values of climate-model tuning parameters. These cost functions are special because they take correlations between neighboring points and between the variables analyzed into account. Using these special cost functions is shown to make little difference when optimizing 2 cloud parameters in CAM3.1 for 2m T, 200mb zonal winds, surface pressure, and surface precipitation. On one hand, lack of sensitivity implies the methodology is probably implemented reasonably well. On the other hand, lack of sensitivity begs the question of whether the sophistication is necessary.

Correlations in space and across variables were handled by Gaussian Markov Random Fields (GMRFs). I had a hard time understanding whether this is an appropriate technique or whether it was implemented correctly. In particular:

1. Equation 1 is introduced in the introduction and is said to be the culmination of the subsequent derivations but is never fully explained. Better explanation is needed. In particular, I don't think it makes sense to provide this equation in the introduction.

2. I think Eq. 1 is a log-likelihood function derived from assuming model errors follow a multivariate Gaussian distribution (eq. 2) with the inverse covariance matrix $\Sigma^{-1}$ replaced by GMRF precision matrix. These points need clarification and the reasonableness of assuming a multivariate Gaussian distribution for model output and for approximating the covariance matrix with a GMRF precision matrix both require further justification.

3. I think the log-likelihood function in eq 1 is missing the following term: $ln((2\pi)^-n/2\text{tr}(\Sigma)^{-1/2})$. Is this true? In any case, this derivation needs to be more clear.

4. The precision matrix is only described for the 2x2 case. Are the "rules" on page 6 followed only for the 2x2 case, or are they followed for all cases?

5. Because Q seems to be defined independently of the spatial autocorrelation in the actual data, I find it hard to believe that it can be a good approximation for $\Sigma^{-1}$ except by chance. In particular, I bet the "witch hat graph" for surface precipitation alone (which has short autocorrelation lengthscales) looks very different than that for surface temperature (which has long autocorrelation lengthscales) and that Fig. 4 only looks reasonable for quantities which happen to have the autocorrelation structure matching the precision matrix assumptions. I would like to see the comparison between $Q^{-1}$ and $\Sigma$ (note I'm asking for things in correlation-matrix space rather than precision matrix space because the former is easier to interpret physically) for several different output fields to gain confidence in the method. The fact that Q is defined independently of autocorrelation in the actual data is my single biggest concern with this paper.

5. The fact that the precision matrix has a zero eigenvalue seems to be an obvious result of the fact that Q indicates the neighbors of each cell and neighbors of the last cell can be predicted from the others (because the cells which are its neighbors have already tagged it as being their neighbor). I am surprised and alarmed that your solution to this problem is to add a small perturbation to make your singular matrix merely nearly-singular. It seems like this nearly-singular matrix will at best have numerical issues and at worst isn't actually solving the system you meant to solve. Wouldn't it make more sense to replace the system with a matrix of 1 lower dimension?

*Other Comments:*

1. In the title and elsewhere, you call your method a "metric". I think the benefit of your approach is that it allows you to evaluate a log-likelihood function in order to choose the best parameter settings fo an uncertainty-quantification problem. While the log-likelihood function does give you a scalar value for a particular set of parameters and is therefore a metric of sorts, I think emphasizing that you're defining a metric is kind of missing the main point of what you're doing. In particular, you have to specify exactly what output you want to use to define a metric and I think a benfit of your method is that it should work on a wide variety of output data choices. In short, I'd suggest changing "metric" to "method" throughout the text.

2. using CAM3.1 is odd and detracts from the publication-worthiness of the paper because it is an ancient model which nobody cares about anymore. Can you really not find data from more recent model runs? It would be worth the effort.

3. In eq. 2, you need to indicate that |x| is the determinant of x.

4. p. 6 line 8: "fuller" should be "more full"

5. You should define what the Kronecker product is for climate people, who may not know off the top of their heads.

6. p. 7, line 18: you're missing a word between supplemental and carries.





Printer-friendly version

Discussion paper

7. p. 11 line 25: low pressure cooling the underlying surface would be a *positive* correlation. Perhaps you're seeing a "thermal low" effect?

8. You show in Fig. 5 that using GMDV or not doesn't make a big difference. Is this the result of your particular choice of parameters and/or model version and/or output variables? Taking the time to test your method in other cases would at a minimum make your conclusions more robust and could potentially show that your method has an important impact in certain circumstances.

9. In the supplementary material, why assume x has means which are all zero?

*Recommendation*:

In summary, I don't have confidence based on the information conveyed in the current draft that this approach makes sense. Part of this is that the methodology is not clearly explained (which can be fixed by rewriting). The other part is that the method seems to assume a spatial autocorrelation structure without any consideration of the covariance structure of the actual data. For these reasons I think the paper should be rejected with encouragement to resubmit once these problems are fixed. Another problem with this paper is that its results are not particularly exciting - using the complicated methodology which is the core of the paper has little impact on model results. I don't think this is reason to reject the paper (because negative results are important to present so others don't repeat them), but is reason to be less enthusiastic in encouraging a resubmit.

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