Geosci. Model Dev. Discuss.,
https://doi.org/10.5194/gmd-2017-95-AC2, 2017

Geoscientific Model Development Discussions

Interactive comment on "Multivariable Integrated Evaluation of Model Performance with the Vector Field Evaluation Diagram" *by* Zhongfeng Xu et al.

Zhongfeng Xu et al.

xuzhf@tea.ac.cn

Received and published: 28 July 2017

We would like to thank the reviewer's comments. Our point-by-point responses are listed below:

==========================

Reviewer #2

This paper is closely related to an earlier paper published by the authors in GMD (doi:10.5194/gmd-9-4365-2016), but in this new paper I believe there is a fatal flaw in that the "vectors" considered are constructed from components representing individual fields which are in general not independent. The fields produced by climate models (and indeed the fields observed in the physical world) are rarely truly indepen-

Printer-friendly version

Discussion paper

dent. Consider, for example, the trivial case of temperature field at 900 hPa and the temperature field at 850 hPa. These fields would be very similar (with second one being slightly cooler than the first), and since they are not independent, they are unsuitable for use as components of a vector. Similarly there are relationships between specific humidity and temperature that yield high correlations between them. Thus, the "vectors" defined in this paper are based on dimensions that are not independent (i.e., not orthogonal). If I am correct that orthogonality of the vector components is a requirement, then the paper rests on unsound mathematics and should be rejected.

RESPONSE:

The definitions of the statistical quantities, i.e., RMSL, Rv, RMSVD, do request an orthogonal coordinates. In our study, we use the M-dimensional standard Cartesian coordinate. The Cartesian axes are mutually perpendicular to each other. Therefore, the axes are always orthogonal no matter whether the scalar fields are independent or not. We just assign each scalar variable to the corresponding component of the orthogonal coordinates and compare two constructed vector fields.

The statistics are still meaningful when the components of vector field are dependent. Consider an idealized case, assuming we constructed a modeled 2-dimensional vector fields (A) with exactly same x- and y- components (the x- and y- components are fully dependent). Consequently, all vectors in the vector fields point in the same direction (In contrast, the vectors point in different directions if the x- and y- components are not fully dependent). Assuming we also constructed an observed 2-dimensional vector fields (B) with exactly same x- and y- component, under such a circumstance, the vector similarity coefficient, Rv, between A and B is equal to the Pearson's correlation coefficient between the x- (or y-) component of A and B. This is still a meaningful measurement on the similarity of two vector fields even if the x- and y-components are fully dependent. As all vectors point in the same direction, one can only consider the variation of vector length. Under such a circumstance, the vector field can be treated as a scalar field (vector length) and one can use correlation coefficient to measure the relationship of two scalar fields, which is consistent with the vector similarity coefficient. Similarly, RMSL and RMSVD do not request the independence of various scalar fields, either.

If I am wrong, then the paper should be considered, but I'm not sure it adds much to what was already published in the earlier paper where the vector components were based on spatial direction (rather than variable). Isn't the present paper an obvious extension of the earlier paper (simply an application to a different vector, one based on dimensions defined by variables rather than spatial dimensions)?

RESPONSE:

Two papers do rest on the same mathematics except that the second paper generalizes the VFE diagram to evaluate vector fields in arbitrary dimensions against the vector fields in two-dimensions in the earlier paper. However, two papers present very different aspects of model evaluation. The earlier paper presented an evaluation method of model performance in simulating individual vector field, e.g., vector winds, temperature gradient. The current paper reports a multi-variable integrated evaluation (MVIE) method of model performance.

The scientific merits of this study relative to previous model evaluation methods mainly includes the following aspects:

(1) Our study provides a more comprehensive way to evaluate model performance in terms of multiple fields compared to the previous methods. For example, as discussed in the introduction section, the commonly used multivariable evaluation method, i.e. the portrait diagram (MCPI, Gleckler et al., 2008), only considers the root mean square errors (RMSEs) of various fields. Although RMSE takes both the correlation coefficient and standard deviation into account, the RMSE cannot explicitly measure the pattern similarity and variance. In contrast, the MVIE method includes multiple sta-

tistical quantities, e.g., RMSL, VSC, and RMSVD, can explicitly measure the pattern similarity, mean and variance, and the overall difference between two constructed vector fields. Therefore, the MVIE method can provide a more comprehensive evaluation of model performance.

(2) As shown in the pyramid chart (Fig. 4), the second paper constructs a hierarchical evaluation framework of model performance which is composed of commonly used statistical metrics for individual variables, VFE diagram, and multivariable integrated evaluation index (MIEI). The first level of metrics, i.e., correlation coefficient (R), RMS value, and RMSD, measures the model performance in terms of individual variables. The first level of metrics can provide more detailed information in specific aspect of model performance but is lack of summarization and cannot provide a quantitative evaluation on model performance in simulating multiple fields. The second level metrics, i.e., VSC, RMSL, standard deviation of RMS values (σRMS), and RMSVD in the VFE diagram, is derived from the first level of metrics and summarizes the overall performance of a climate model in simulating multiple fields. The VFE diagram well summarized multiple statistics but is still hard to rank model performance. The MIEI further summarizes the VSC, RMSL, and σRMS in the VFE diagram into a single index to rank various climate models in terms of the performance in simulating multiple fields. Thus, the hierarchical evaluation framework allows one to choose different levels of metrics, which will facilitate the evaluation and inter-comparison of model performance.

The abovementioned two aspects are new and constitute significant advances relative to previous model evaluation methods. Our first paper published in GMD in 2016 did not tackle any of abovementioned aspects. Instead, the first paper focused on the evaluation of model performance in simulating individual vector field, e.g., vector winds.

Interactive comment on Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2017-95, 2017.