Geosci. Model Dev. Discuss.,
doi:10.5194/gmd-2016-172-RC1, 2016

Geoscientific Model Development Discussions

Interactive comment on "A diagram for evaluating multiple aspects of model performance in simulating vector fields" *by* Zhongfeng Xu et al.

Anonymous Referee #1

Received and published: 5 September 2016

A diagram for evaluating multiple aspects of model performance in simulating vector fields by Z. Xu, Z. Hou, Y. Han, and W. Guo

The authors describe an extension of the Taylor diagram (Taylor, 2001) to compare 2-dimensional quantities such as horizontal wind fields from different sources in an analogous way to that described by Taylor (2001) for scalar quantities. This can be a useful extension to the widely Taylor diagrams when evaluating the output of climate models with observations or when comparing models again predecessor versions or other models.

I suggest major revisions to the manuscript before publication in Geoscientific Model Development addressing the points given below.

General comments

- A paragraph giving some guidance on the scientific interpretation of the proposed statistics would be helpful, possibly with examples demonstrating how "good" and "not so good" agreement between two data sets looks like. It would also be helpful to explicitly point to possible issues and limitations to keep in mind when looking at complex quantities such as the skill scores proposed by the authors (equations 13 and 14).
- An application of the proposed VFE diagram to evaluate the performance of models usually requires taking into account observational uncertainties as the reference data (at x=1, y=0; e.g., figure 9) are usually not the truth. This is particularly the case for quantities that have larger uncertainties than 850-hPa wind speed. I would like to see a discussion of possible issues and limitations as well as thoughts of how to deal with observational uncertainties in this type of diagram.
- What are the key messages of section 3.2 (relationship of VSC with MDA)? Things need to be put into context by providing a motivation for this analysis. The statements made have to be more precise, for example, "[...] a smaller MDA generally corresponds to a larger R_v, and vice versa." does not seem to provide a lot of useful information for the application and interpretation of the VFE diagram. Again, guidelines of how to interpret the proposed statistics in a scientifically meaningful way would be helpful.
- Section 5.1 gives an example for the application of the VFE diagram using 850-hPa wind speed from 19 CMIP5 models. There are, however, no details on the

model runs used (model experiment, ensemble members, etc.). It is also not clear to me what exactly you are comparing (multi-year annual means, monthly means, etc.). I presume the models have been regridded to a common grid? If so, which grid and which interpolation method has been applied? Also, I am missing a reference for the NCEP reanalysis data used. This section needs some rewriting to make clear what exactly has been done and what is being compared here. The current description is not sufficient to reproduce any of the results shown here, which is not acceptable for a scientific paper.

- Section 5.2.2 (statistical significance of differences): I am missing a clear definition of what the authors mean by "statistical significance of differences" in the context of 2-dim vector fields. The argument of "separated groups" without further explanations or governing equations is not precise enough for a scientific paper. I have the impression that the authors are rather speculating here than presenting any scientific evidence. This leads to contradictions within the section that need to be addressed. For example, the authors claim on p. 10, l. 27-28: "Thus, the differences between models 12, 13 [...] are likely to be significant.". The authors contradict this statement a few lines later (p. 11, l. 1-2): "[...] which may not be sufficient to conclude a significant difference between the three models, especially for models 12 and 13." leaving the reader confused. Also, it is once again not clear what data have been used and what quantities are being compared (models, model experiments, ensemble members, time period, averaging, time resolution, regridding, etc.).

Specific comments

- p. 5, l. 24, give equation numbers for RMSL
- p. 7, l. 8: replace "for a certain angle" with "by a certain angle"

- p. 7, l. 9: production → generation
- p. 7, l. 21: what do you mean by "the performance of R_v"?
- p. 9, l. 8: replace "a dotted contour" with, for instance, "dotted circles"
- p. 9, l. 9: "line" → "lines"
- p. 9, l. 15, insert "VFE" before "diagram"
- p. 11, l. 30: "anomalous scalar fields" → "scalar anomaly fields"
- p. 13, equation (A1): how did you get from line 2 to line 3? Shouldn't $(\bar{x}_{ai} + x\prime_{ai})^2 = \bar{x}^2_{ai} + 2\bar{x}_{ai}x\prime_{ai} + x\prime^2_{ai}$? I.e., what happened to the term $2\bar{x}_{ai}x\prime_{ai}$?
- p. 14, equation (A4): similar to eqn. (A1), how did you get from line 2 to 3?
- p. 19: the added value of figure 2 seems rather limited, this figure could be deleted
- p. 20, "[...] and each randomly produced vector field" → "[...] and randomly generated vector fields"
- p. 20, delete "are" before "included in the statistics"
- p. 21, "vector similar coefficients" → "vector similarity coefficients"
- p. 21, "between January and 12 months (Solid line)": this formulation is not clear, please rephrase

Interactive comment on Geosci. Model Dev. Discuss., doi:10.5194/gmd-2016-172, 2016.