**Biogeosciences
Discussions**

Interactive
Comment

Interactive comment on "A comprehensive benchmarking system for evaluating global vegetation models" *by* D. I. Kelley et al.

Anonymous Referee #1

Received and published: 11 January 2013

The paper by Kelley et al. reports on a very well designed model-benchmarking study. The paper is well written, and highly relevant to the audience of Biogeosciences. Also I want to mention that the paper is very timely and certainly helpful for related activities, most prominently the "International Land Model Benchmarking Project" (ILAMB, www.ilamb.org). I see a very innovative aspect in the fact that the paper applies a wide range of benchmarks on different aspect that are relevant to land surface processes. The paper could be elaborated a bit more in the introduction in order to better discuss related papers and embed the present study in the context of model evaluation activities.

I have two remarks that may not necessarily be included in the revisions, but could be useful to subsequent analysis:

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

1. In the list of principles (2.1) I miss one important criterion: Any data set used in a transparent benchmarking should be free to the scientific community. I would suggest to only use benchmarking data set that are usable by other modeling groups - otherwise, different benchmarking exercises cannot be compared on the same grounds.

2. The comparisons of the seasonality (2.3.3) looks a bit complicated to me. Kobayashi & Salam (2000); van Oijen *et al.* (2011) have shown that the MSE (the squared part of the RMSE) can be decomposed to three elements:

$$\text{MSE} = \left\langle (X_i - X_{i'})^2 \right\rangle = \underbrace{(\langle X_i \rangle - \langle X_{i'} \rangle)^2}_{\text{bias}^2} + \underbrace{(\sigma_i - \sigma_{i'})^2}_{\text{variance diff.}^2} + \underbrace{2(\sigma_i \sigma_{i'})(1 - r)}_{\text{phase error}}. \quad (1)$$

The meaning of the squared data bias is obvious, the second term indicates differences in the fast variability, and the lack of correlation r between X_i and $X_{i'}$ is a very simple estimator for phase errors. Wouldn't this last term do the job in this benchmarking exercise?

References

Kobayashi S, Salam MU (2000) Comparing simulated and measured values using mean squared deviations and its componenets. *Agronomy Journal*, **92**, 345–352.

van Oijen M, Cameron DR, Butterbach-Bahl K, *et al.* (2011) A bayesian framework for model calibration, comparison and anaylsis: Application to four models for the biogeochemistry of a a norway spruce forest. *Agricultural Forest Meteorology*, **151**, 1609–1621.

Interactive comment on Biogeosciences Discuss., 9, 15723, 2012.