Geosci. Model Dev. Discuss., 8, C198–C200, 2015
www.geosci-model-dev-discuss.net/8/C198/2015/

Geoscientific Model Development Discussions

Interactive comment on "Validating a 1-D SVAT model in a range of USA and Australian ecosystems: evidence towards its use as a tool to study Earth's system interactions" *by* G. P. Petropoulos et al.

Anonymous Referee #1

Received and published: 20 March 2015

Comments/questions for Petropoulos et al. 2015

The introduction could be greatly shortened. The authors give a motivation for the study of land-atmosphere interactions, as well as a history of land surface modeling, neither of which are necessary. Basically the last paragraph of the introduction (p2444) would suffice. It describes the paper's objectives and the methodology used to meet those objectives. The previous few pages are mainly superfluous.

The authors clearly describe the metrics they are using in section 4.3, but none of the metrics are relative, which I found an impediment to judging the quality of simulations.

Specifically, what does a mean bias error of say 50 W/m2 imply? If the mean energy flux is 500 W/m2, then perhaps that is a 'good' simulation; if the mean is 50 W/m2, it is likely not. Without a reference to the actual observed value, the reader cannot determine this. For example, in section 5.3, the authors state that because the RMSD was the lowest of the examined fluxes, and the average mean bias error is 2.84 W/m2, latent heat (LE) was well reproduced. However, an examination of figure 6 shows LE to exhibit quite a lot of scatter compared to some of the other variables. In addition, most of the values are 150 W/m2 or less, indicating that the RMSD of $\sim$40 W/m2 is quite significant. This undermines statements such as "The model showed excellent precision in reproducing daily trends of LE fluxes in most sites evaluated"

Some statements were confusing: p 2454 "A systematic underestimations of Rnet was evident, leading to an overall satisfactory agreement between the model predictions and in situ observations"; why is a systematic underestimation considered satisfactory?

Another issue is the use of correlation; when looking at signals having a strong diurnal cycle, high correlations are to be expected (incidentally, are three significant digits for R2 and NASH necessary?). Are there metrics that the authors could use that would account for this effect?

End of section 5.1, the authors note that larger errors for Oz sites occur feb-june, while the converse is true in the Ameriflux sites; the authors may wish to add that these time periods correspond to summer for each region, and are therefore consistent.

The paper contains no figures showing actual simulated fluxes. This would helpful in understanding the characteristics of the errors, as well as the observations.

The authors use subjective descriptions throughout the paper. For example, in the abstract: "A good to excellent agreement between the model predictions and the in situ measurements was reported,..." good/excellent are not defined, and the reader may not agree with these subjective measures. Another, pg 2454: "leading to an overall satisfactory agreement", what is the authors' definition of satisfactory? p2456:

"The latter was suggestive that model predictions were in good to excellent agreement to the in situ measurements", what differentiates "good" from "excellent"? I would encourage the authors to replace such statements in the text with statements having clearly defined meaning (e.g. "within 10 per cent of observed", etc...).

Interactive comment on Geosci. Model Dev. Discuss., 8, 2437, 2015.