Geosci. Model Dev. Discuss., 8, C3341–C3345, 2015
www.geosci-model-dev-discuss.net/8/C3341/2015/

Geoscientific Model Development Discussions
Open Access

Interactive comment on "The GEWEX LandFlux project: evaluation of model evaporation using tower-based and globally-gridded forcing data" *by* M. F. McCabe et al.

M. F. McCabe et al.

matthew.mccabe@kaust.edu.sa

Received and published: 13 December 2015

Author Introduction. We thank the anonymous reviewer for their thoughtful comments on our manuscript. Following is a point-by-point response to the issues and suggestions raised.

Comment 1. Address footprint of tower flux observations and how those model is capable to explicitly treat sub-grid variability. If not, the evaluation cannot guarantee the comparisons are not biased. Screening out tower flux measurements over heterogeneous footprint also helps.

Author Response. The issue of footprint analysis is certainly of interest in high-resolution flux retrievals, especially where spatial variations can be resolved at the pixel scale. However, in our analysis, the gridded data completely covers any within-footprint scale impacts: by at least one order of magnitude. Tower data are influenced by the tower-footprint (which in addition changes per sensor, thus per measured variable), but this is an inherent element of flux-tower analysis. Sub-grid scale variability is not explored here, since it is well beyond the scope of the analysis. Regardless, none of the models explicitly account for this, beyond those that include fractional cover measurements (which likely represent most subgrid variability). Point-scale tower meteorological forcing is assumed to be areally representative (as there is no other alternative).

One of the unique aspects of this work is that tower data are consistent across all model simulations: that is, tower-bias is minimized, by ensuring that all models are assessed against the same tower records. Further, the tower-to-grid scale analysis also acts as a diagnostic of representativeness and point-to-pixel error (apart from forcing uncertainty). Screening out "heterogeneous" landscapes (most are heterogeneous even at the tower scale) would unreasonably restrict any evaluation of flux retrievals from ever being undertaken: including those performed here. Here we accept the inevitability of real-world heterogeneity and include this as an inherent element of our intercomparison study.

Author Changes. Table A1 has been updated to include tower height, in addition to ground elevation, as an indirect measure of the tower fetch. We also highlight the inevitable nature of heterogeneity in the landscape in the Introduction through the following sentence: "One of the unique aspects of the present study is that tower data are consistent across all model simulations: that is, tower-bias is minimized, by ensuring that all models are assessed against the same tower records. Further, even though sub-grid scale variability is not explored here (since none of the models explicitly account for this), the tower-to-grid scale analysis acts as a diagnostic of representativeness and point-to-pixel error."

GMDD
8, C3341–C3345, 2015

Comment 2. How climate zones are determined? Since those models are mainly based on energy and water balance and availability, physical criteria should be explicitly provided. In that sense, additional analysis or replacement for 3.3 over Budyko climate regime and the curve.

Author Response. Climate zones are specified in the metadata of each of the Fluxnet tower locations, which are ultimately derived from Rubel and Kottek (2010) "Observed and projected climate shifts 1901-2100 depicted by world maps of the Köppen-Geiger climate classification", Meteorol. Z., 19, 135-141. DOI: 10.1127/0941-2948/2010/0430. This is the standard approach to climate classification, so has been retained for this analysis. A statement describing the classification has been included in Section 2.1.4.

Author Changes. The following statement is included in Section 2.1.4: "These zones were prescribed from the tower-specific metadata, which were in turn derived from Rubel and Kottek (2010) based on a Köppen-Geiger climate classification".

Comment 3. Provide additional information for evaluation of grid-based forcing data quality. Knowing the character of error (i.e., phase mismatch or persistent offset) would be also very important.

Author Response. The emphasis here has been placed on examining the variability in model response, rather than on variability in model forcing: principally because characterizing the impact of forcing uncertainty on the models themselves is a complicated (and under-studied) task. Within the context of this study, attributing model error to forcing error requires that the "true" forcings be known, which of course isn't the case. Nonetheless, the reviewer correctly identifies an area of much needed further study.

The focus of our analysis is to examine model performance in response to an operational forcing set, using tower data as a benchmark. Given that the analysis encompasses record lengths of (on average) 4 years (see Figure A1) and that the 15 or 30 minute tower data was upscaled to 3 hourly, a focused assessment on the forcing data is impractical. Importantly, any potential lag or phase mismatch would likely

GMDD

8, C3341–C3345, 2015

be removed in the temporal upscaling. Figure 7 and Figure 10 provide both R2 and Nash-Sutcliffe efficiency metrics of comparisons between the 3 hourly gridded data and the temporally equivalent local-scale forcing across all of the 45 tower sites, for both climate zone and biome type. As can be seen, the correlation is very strong for all variables except for wind-speed (which is certainly one of the most challenging variable to match between scales).

The individual gridded forcing data products have been used previously in a range of analyses. As noted above, the purpose of our investigation is not to assess the quality of the forcing, but to examine the impact of these widely available datasets on the estimation of model derived evaporation. The issue of forcing quality and the impacts of this on product accuracy are regularly highlighted throughout the manuscript (see e.g. Abstract Line 5, Section 3.1 and elsewhere) and form key elements of the presented results. As highlighted in Section 3.1 (from Line 20), while we acknowledge that spatial scale plays a role, the impact of internal inconsistencies within the gridded data (i.e. the fact that data sets are derived from different sources) drive the inter-product differences.

Comment 4. If possible, aligning the analysis in previous researches with different types for models (e.g., land surface models and hydrologic models) will introduce additional values.

Author Response. We are unclear what the reviewer is recommending here. In response to a query by Reviewer 2, we have provided some additional details on the differentiation of this work to previous analysis. This provides a clearer rationale for the current effort and places it in the context of prior work. An additional recommendation of Reviewer 2 was to provide a synthesis study that overviews the considerable research efforts that have driven global flux studies over the last few years: this will form the focus of future efforts. In terms of benchmarking the performance of evaporation models against GCM, LSM or reanalysis type output, we refer to the earlier and thematically related studies by Jimenez et al (2011) and Mueller et al. (2011, 2013).

GMDD

8, C3341–C3345, 2015

Jiménez et al. (2011). "Global intercomparison of 12 land surface heat flux estimates." J. Geophys. Res. 116(D2): D02102

Mueller et al. (2011). "Evaluation of global observations-based evapotranspiration datasets and IPCC AR4 simulations." Geophysical Research Letters 38(6)

Mueller et al. (2013). "Benchmark products for land evapotranspiration: LandFlux-EVAL multi-data set synthesis." Hydrology and Earth System Sciences 17(10): 3707-3720.

Comment 5. Additionally, this manuscript can be more concise. Remove or rephrase of speculations without any robust evidence (e.g., some sentences including ‘may’ or ‘perhaps’) would help.

Author Response. We have attempted to avoid pure speculations. However, there remains much additional work that needs to be done to untangle some of the numerous issues related to global flux development. Speculation is an inevitable by-product of this. Where questions on the intangibles of our analysis remain, the intent is to highlight these issues and to assist in spurring on future research efforts. We have rephrased such terms as mentioned by the reviewer (“perhaps” now appears only once).

Author Changes. We have removed the line “likely due to variability in performance as a response to the land surface type or climate condition” on Page 6825, Line 24. We have replaced “perhaps” with “represents one of” in the Discussion on Page 6838, Line 23. “Perhaps” has been removed on Page 6838, Line 27.

Interactive comment on Geosci. Model Dev. Discuss., 8, 6809, 2015.