Geosci. Model Dev. Discuss., 8, C3326–C3333, 2015
www.geosci-model-dev-discuss.net/8/C3326/2015/

Geoscientific
Model Development
Discussions

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 paper. In responding to these, we have itemized the various comments in order of their appearance in the response letter.

Comment 1. While I tend to agree this is an interesting and important study as we continue on the way of better global ET mapping, it has several issues. First of all, I do not think it is appropriate to compare ET estimates at the grid-cell scale (i.e., 0.5 in this case, or âĹij 55 by 55 km) with flux-tower measurements (usually have a footprint of 1âĹij2 km). The spatial heterogeneity in soil, vegetation and micro-climate

can lead to large spatial variation in ET within one grid-box. This issue can be more strongly manifested if there were large variation in topography (leads to strong climate gradients) or vegetation type (the grid-box sits on the boundary of two biome types). Therefore, their conclusion regarding grid-box evaluation is no longer stand.

Author Response. The reviewer raises an important point regarding the issue of scale in the evaluation of global flux products. Indeed, this largely reflects one of the motivations for undertaking this work: how to evaluate large-scale flux retrievals in the face of limited observations. The overriding rationale of this paper is not to use tower-based fluxes to "validate" grid-scale model simulations, but rather to understand how grid-scale fluxes compare with a parallel simulation exercise run at the tower of the scale. As noted in the Introduction (see line 29), "Understanding how application of gridded forcing data might influence the performance of the different models, relative to their performance when forced with (presumably) higher-quality tower data, is a motivating rationale for this work". The models are deliberately run with both sets of forcing to examine the impact of scaling issues on flux response: a very standard approach to assess such behavior. Indeed, this provides an indirect assessment of the footprint mismatch, which is explicitly recognized in Paragraph 2 of Section 3.1, as well as in Figures 7 and 10, which highlight the strong correlation between variables at these opposing scales.

In the absence of any alternative means of assessment, making use of an extensive record of tower-based fluxes is certainly an appropriate method by which the behavior of large-scale fluxes can be examined. Determining alternative metrics of performance (i.e. the hydrological consistency approach of McCabe et al. 2008) is an area of ongoing investigation.

McCabe MF, EF Wood, R Wójcik, M Pan, J Sheffield, H Su and H Gao (2008) "Hydrological consistency using multi-sensor remote sensing data for water and energy cycle studies" Remote Sensing of Environment, 112(2): 430-444

Comment 2. In this sense, secondly, a big concern is what is the advancement of this study over a recently published paper in AFM (also conducted by the same group of authors, i.e., Ershadi et al., 2014, Multi-site evaluation of terrestrial evaporation models using FLUXNET data, AFM). In Ershadi et al. (2014), four models (three of them are used in this study, including PT-LPJ, SEBS, and PM-Mu) were evaluated at 20 flux sites, most of which (or all, I did not check) are also used in this study. What is the difference between using 20 and 45 sites, if they both intent to represent a “global situation”.

Author Response. To clarify, only two models overlap with the current study (SEBS and PT-JPL). The Penman-Monteith model was a single-source approach. Regardless, there are considerable advances in this current contribution. Most notably, the present study focuses on the scaling between the tower and the overlying grid, which is a critical scale of understanding in assessing the quality of our global ET estimates (for additional details, see our more detailed response in reply to Comment 1). In terms of more specific differences, the study of Ershadi et al. (2014) focused solely on a tower-based comparison of four evaporation models. Here we extend the analysis to include more than double the number of towers, an analysis encompassing a longer time period, and also examine a range of additional biome types as well as climate zones.

Comment 3. On to content there is not much difference between the two studies, given that the grid-box scale evaluation performed here is not valid. Moreover, these two studies even come to a similar conclusion that PT-JPL seems to perform generally better than the other models.

Author Response. We completely disagree with the reviewers position that the grid-scale evaluation is invalid. Apart from providing no justification for such a statement, the reviewer offers no alternative approach to evaluating global flux simulations (despite recognizing the importance of such research). In terms of the many differences between the works, as well as the fundamental scientific rationale which differentiates these, we refer the reader to our previous responses to Comments 1 and 2. That the

papers arrive at a similar conclusion is a verification of the robustness of findings from the earlier work: a very appropriate outcome.

Comment 4. Page 6823-Line 1. Soil evaporation in PM-MU is not entirely based on the Penman-Monteith equation, although it uses PM to estimate potential ET over unsaturated surface. Then, the effects of soil moisture restriction on soil evaporation are reflected by meteorological forcing based on the complementary relationship.

Author Response. The reviewer is correct in that PM-Mu used a hybrid approach to soil evaporation. Specifically, the PM-Mu actual soil evaporation flux is calculated using potential soil evaporation and a soil moisture constraint function from the Fisher et al. (2008) ET model. This function is based on the complementary hypothesis (Bouchet, 1963), which defines land-atmosphere interactions from air VPD and relative humidity (RH, %). The paper has been modified to reflect this. We had deliberately minimised model descriptions, given the large body of literature available to these models, but have adjusted the text to reflect these additional details.

Author Changes: We have changed the relevant lines in Section 2.2.3 to reflect the following: “Estimation of evaporation for the interception and transpiration components is based on the Penman-Monteith equation (Monteith, 1965). Actual soil evaporation is calculated using potential soil evaporation and a soil moisture constraint function from the Fisher et al. (2008) ET model. This function is based on the complementary hypothesis (Bouchet, 1963), which defines land-atmosphere interactions from air VPD and relative humidity (RH, %). Evaporation components are weighted based on the fractional vegetation cover, relative surface wetness and available energy.”

Comment 5. Page 6824-Line 25. Ep is the annual potential evaporation.

Author Response. Noted. Thanks for this editorial correction.

Author Changes. The sentence has now been changed to state potential evaporation.

Comment 6. The PM-Mu model was designed to run at a time scale longer than a day, and so the model parameters were calibrated at that time scale. If using the default parameter value at the instantaneous time scale, there can be some uncertainties. This issue needs to be mentioned and discussed.

Author Response. In the PM-MU (MOD-16) ATBD document it is stated: "The MOD16 algorithm computes ET at a daily time step. This is made possible by the daily meteorological data, including average and minimum air temperature, incident PAR and specific humidity, provided by NASA's Global Modeling and Assimilation Office (GMAO or MERRA GMAO)". If this is the case, one assumes that the parameters used are appropriate for that time scale. It is worth noting that the developer of the model (Qiaozhen Mu) is a co-author on a related publication where the model is used at the 3-hourly time scale! Nonetheless, the possible influence of parameter specification has now been mentioned explicitly at the end of Section 2.2.3.

Author Changes. Lines added at the end of Section 2.2.3. "One consideration that may influence model simulations is that this parameterization approach was developed at the daily-scale. However, both the present and also a recent related study (Miralles et al. 2015) suggest no obvious impact for sub-daily application."

Comment 7. Page 6826-Line6. Figure 4 should be Figure3.

Author Response. Noted. Thanks for this editorial correction.

Author Changes. The sentence has been adjusted to refer to Figure 3.

Comment 8. Page 6826-Line13-15. This is not supervising, because primary use of meteorological data does not allow PM-Mu to effectively capture the soil moisture restriction on ET on a physical basis. The consequences of that are a slower response of variations in energy and heat fluxes than the thermal remote sensing-based ET models and unreasonable spatial ET patterns (because the spatial variation in soil moisture is usually much larger than that in climate variables and vegetation types).

Author Response. It is unclear what the reviewer wishes to convey in this comment,

particularly in regards to "supervising". In terms of the PM-Mu approach, we have not previously observed (or are aware of) any inferred lag in response. Similarly, no such lag has been observed in either PT-JPL or SEBS in a number of intercomparison exercises. Only GLEAM makes explicit use of a soil moisture term: none of the other models incorporate moisture dynamics.

Comment 9. Line 6827-Line 15. Why the aridity index varies between 0 to 1? P can be larger than Ep, for example, in Amazon.

Author Response. The author is correct in that the aridity index can go beyond a value of 1. As we are not exploring the concept of aridity in tropical environments, we had limited the scale to fall between 0-1. The purpose of this overlay was to show the performance of the models in dryer-conditions. I would note that the maximum value is 2.16. To avoid any potential confusion, we have adjusted the secondary axis to vary between 0-2 to better reflect the full range of aridity values.

Author Changes. Figures have been updated to reflect the full range of aridity index (varying between 0-2) and have been placed on a secondary y-axis (to remove the 0-1 range conflict with the r statistic)

Comment 10. Page 6835-Line 24. I am not sure if the performance of SEBS is that highly dependent on vegetation height. As shown in this study, SEBS also fails in estimating ET over shrubland, where vegetation height is generally low (shorter than 3 m). To me, the key issue here is single-source vs multi-source. SEBS is a single source model that does not allow partitioning between soil evaporation and plant transpiration. In shrubland ecosystems where vegetation cover is usually low and highly non-uniform, soil evaporation account for a large proportion of the total ET. In this case, I would expect that SEBS overestimates ET, as demonstrated by the authors. Similar conclusion was made in a comparison study by Choi et al. (2009), in which they found that multi-source models consistently outperform single-sources models, and the largest discrepancy was found in areas with LAI smaller than 2 (indicate low vegetation cover and possibly non-uniform cover).

Author Response. The reviewer offers some excellent additional insight into the SEBS model. However, it is worth noting that Choi et al. (2009) did not specifically study the SEBS model, so it is hard to draw definitive conclusions as to that studies translation to the present work, particularly as it focused only on corn and soybean agricultural fields (i.e. croplands in our study). But the reviewers point remains, although not necessarily in the manner described: in a recent study focusing on the Penman-Monteith model (Ershadi et al. 2015), a single-source PM model outperformed other dual and three source configurations over grasslands and shrublands, highlighting the complexity of these interactions as well as model specific performance.

Choi, M., W. P. Kustas, M. C. Anderson, R. G. Allen, F. Li, and J. H. Kjaersgaard (2009) "An intercomparison of three remote sensing-based surface energy balance algorithms over a corn and soybean production region (Iowa, U.S.) during SMACEX", Agr. Forest Meteorol., 149, 2082–2097, doi:10.1016/j.agrformet.2009.07.002.

Comment 11. Page 6840-Line10. While I tend to agree that coarse resolution meteo files will reduce the model performance, I do not see this point is well supported by the analysis here. Again, this is due to inappropriate validation of 0.5 degree ET estimates by using 1âĹij2 km2 flux-tower observation. This is wrong.

Author Response. The analysis performed here could not be more conclusive that simulations using the coarse-resolution forcing data perform worse than simulations run with the corresponding tower data. In the absence of other performance metrics, the tower fluxes offer an appropriate means through which to assess model behavior. Indeed, this is a standard approach, so it is unclear precisely what the reviewer is challenging here. We are afraid that this issue must remain a matter of opinion, because we cannot see a way of substantiating your argument without having validation data at 0.5 degrees resolution (and of course if we did have those data, the disagreement would not be unnecessary). These points are well explained in the Discussion section.

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