Hydrol. Earth Syst. Sci. Discuss., 6, C2643–C2645, 2009
www.hydrol-earth-syst-sci-discuss.net/6/C2643/2009/

Hydrology and Earth System Sciences Discussions

Interactive comment on "Seasonal evaluation of the land surface scheme HTESSEL against remote sensing derived energy fluxes of the Transdanubian region in Hungary" *by* E. L. Wipfler et al.

Anonymous Referee #1

Received and published: 16 November 2009

In my opinion, this paper is not ready for publication. The subject matter (evaluation of model-derived fluxes against observations) is important, but the approach has some important flaws. When these flaws are addressed, the results may not look very good. 1. My main criticism is the implicit assumption in the text that the SEBAL fluxes (derived from remote sensing) are accurate enough to constitute "truth" for the evaluation. The SEBAL fluxes may take some inputs from remote sensing, but this doesn't mean that they're intrinsically accurate; on the contrary, by not explicitly accounting for water

stress impacts on evaporation (apparently using instead, at least in part, some kind of surface temperature - NDVI metric for water stress), and by making a host of other assumptions, the accuracy of the SEBAL fluxes is limited and probably questionable in many places. In fact, one could argue that the SEBAL flux accuracy shouldn't be expected to be greater necessarily than that of the HTESSEL fluxes being evaluated. The authors provide Figure 4 as a way of demonstrating the validity of the SEBAL fluxes. The evaporation fluxes from SEBAL are compared in the right panel of Figure 4 to measured fluxes at some meteorological towers, and at first glance, the two sets of fluxes do look highly correlated. Evaporation, though, almost certainly increases with net radiation, so all we're seeing here (I think) is the ability of SEBAL to capture the net radiation, which is inherently easier and is, in any case, already shown in the left panel. A *true* test of SEBAL against ground observations would involve constructing a scatter plot of lambdaE/Rn values, not E values, for Figure 4. I'm guessing that such a scatter plot wouldn't look so good. 2. The lambdaE/Rn ratio is actually used later in the work (e.g., Fig. 5), showing that the authors recognize this ratio as an important diagnostic. This, however, brings up a troubling issue. I'm sure that this is not true, but at least the appearance is there that the authors sometimes use E rather than lambdaE/Rn to make their statistics look better. I'm speaking now (in addition to Figure 4) of the final row of Table 4. Why is the correlation coefficient of E shown, rather than that of lambdaE/Rn? Again, you get excess correlation with E values just because E scales with the net radiation – it's not a true test of the model's ability to partition the radiation between latent and sensible heat. Again, I'm sure the authors don't mean to do this, but the table, as written, comes off looking misleading. 3. I'm especially confused by the statistics discussed in Table 2. The percentiles may match, but do the fluxes match in the *same geographical locations*? Isn't that the appropriate statistical measure of success? Couldn't you get the right percentile values with the low and high values from SEBAL in different locations than those of HTESSEL? The percentiles-based statistical analysis comes off looking strange and ineffective. Again, why don't the authors don't plot the SEBAL seasonally-averaged lambdaE/Rn (not E) at each grid cell against the corresponding HTESSEL values in a scatter plot and then compute the correlation between the values? 4. I'm a little surprised by the use of TRMM data to correct the RACMO data. TRMM is focused on tropical precipitation, and its accuracy over land, relative to that over the ocean, is small. Can the authors comment on the accuracy of TRMM data at 47 degrees north? Isn't there a gauge-based dataset available? Is the TRMM data used some kind of merged product? 5. To summarize some of my main concerns, there are, I think, at least four reasons for differences between the SEBAL fluxes and the HTESSEL fluxes: (a) the SEBAL fluxes are themselves estimates and are likely to be inaccurate; (b) the HTESSEL model probably has sub-optimal parameterizations; (c) the forcing used for HTESSEL may be off during the evaluation period; (d) the initial conditions for the HTESSEL runs may be inappropriate (only two configurations of initial conditions are examined here). The paper claims to make quantitative estimates for the HTESSEL errors (the second reason). After reading it, I'm not at all convinced that the authors have truly isolated the HTESSEL inaccuracies from the others.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 6, 6293, 2009.