

## ***Interactive comment on “Benchmark products for land evapotranspiration: LandFlux-EVAL multi-dataset synthesis” by B. Mueller et al.***

**Anonymous Referee #1**

Received and published: 30 April 2013

General The authors have compiled an impressive collection of ET estimates to put forth a benchmarking product. Overall, the analysis contains thoughtful commentary and may be of value to several categories of users within the scientific community. It is this last point that the authors need to consider more carefully in their presentation. Given the array of uncertainties in each product, mostly driven by forcing issues, but also perhaps related to divergent land cover classification, it is not immediately clear how a potential user of these data would select among the different data sources, or towards a blended merger of multiple products. Given the broad scope and promise of this work, I feel that it is of value and worth publication. However, there are several major issues within the manuscript that need to be resolved/expanded upon prior to publication.

C1302

Major Beyond the controls of model forcing, the other dominant mechanism behind ET will be land-cover (i.e. vegetation). Two major limitations of the present analysis must be clarified addressed in this regard. First the issue of land-cover agreement between products/models needs to be addressed, since this alone could be responsible for differences among products, e.g. if one product is assigned forest over a pixel where another product is assigned grassland. The second issue is spatial aggregation, which is related to the first issue. How are land-cover classes aggregated? Was interpolation linear, and how were land-cover discrepancies handled? These issues may provide additional insights into inter-product discrepancies (e.g. differences in albedo, roughness must affect latent heating). Data constraints: the authors constrained ET by net radiation. Is this a physically reconcilable approach? Can the authors cite another study employing this method, or would a Bowen ratio approach be more suitable? PG9, L8: Why can ground heat flux not be neglected for ET values less than 0.3 mm/day? Clearly there is a very large coefficient of variation among ET estimates with respect to precipitation (Figure 5). In addition there are frequent conflicting trends (Table 3,4) among products. Given such large disagreements in the data, how can a potential user of such a dataset be confident in their hydrological-consistency? Should the ensemble mean, median be used? Should certain products be thrown out? A much broader discussion of these points is needed in order to provide the context of this benchmarking effort for the scientific community. Introduction: Page 5, line 1-2: The authors describe a precipitation trend 1900-1988 as important, but then only offer unsubstantiated explanations, such as intensification of the hydrologic cycle. This issue is certainly linked to changes in ET, which the authors use to reconcile potential trends. A simulated ET product is certainly a product of it's forcing, such that the issue of changes in radiative forcing due to solar cycling seems relevant in this context and should be included by the authors.

Minor Several grammatical issues: i.e. pg5 line 24, Pg9,L25,26 Figure 3: Using a yellow time series (i.e. VIC) is very difficult to see and the authors should consider a different color scheme. Figure 4: It is not clear that this figure is necessary.

C1303

C1304