RC1

Received and published: 30 November 2016

OVERVIEW

The manuscript describes the new version of the GLEAM dataset (v3) that includes land evaporation and root-zone soil moisture. The novelties in the retrieval algorithm and in the input datasets are firstly outlined. Secondly, three different datasets are generated by using different inputs and their quality is assessed through a comparison with in situ observations worldwide. Moreover, a cross comparison with respect to the previous version of GLEAM (v2) is carried out.

GENERAL COMMENTS

The manuscript is well written and clear. The new release of the dataset is surely of interest for many research applications both in the hydrology and climate disciplines. Moreover, the new release explicitly contains the root-zone soil moisture dataset that represent an additional benefit. The new release incorporates significant changes with respect to the previous version. Therefore, I believe the paper and the dataset deserve to be published on Geoscientific Model Development. Before the publication, I recognized some points that, in my opinion, need improvement and clarification.

We would like to thank the referee for reviewing the paper and giving some interesting comments and feedback. Below, we give a point-to-point reply to the comments posted by the reviewer.

1. MAJOR: In several Tables and Figures, the comparison between the three versions of v3 dataset, and against v2 dataset, is shown. In terms of soil moisture, it is highlighted that v3 performs better than v2 and that v03a is performing the best. However, I am wondering if the differences in the correlations between datasets are statistically significant. For instance, I believe that the differences reported in Table 3 for the overlap period between the three v3 datasets are not significant (median values between 0.61 and 0.65 for surface soil moisture). Therefore, I wouldn't stress too much that the new dataset is performing the best in terms of soil moisture, as the differences in the performance are quite small.

   We thank the referee for this comment and agree that we need to support the results with statistical significance tests. Therefore, in the revised version of the manuscript we will include the results of a statistical test to verify whether differences in correlations are significant or not. The discussion of the results will be based on these results as well.

2. MODERATE: It is underlined several times that v03b and v03c are "fully satellite-based" datasets. It is not correct. The satellite rainfall product used as input is the gauge-corrected version of TMPA. As it is well-known, in TMPA 3B42v7 dataset ground observations are used for correcting the monthly totals month-by-month. They are not used for correcting the long-term bias, as it reads at lines 1-2, page 12). Therefore, in the gauge-corrected product the contribution of ground observations is significant (note that it could happen that the seasonal cycle is inverted between the real-time and the gauge-corrected version). I suggest removing the definition of these datasets as "fully satellite-based".

We agree with the reviewer that the TMPA 3B42v7 product is not 'fully' satellite-based and that monthly totals are bias-corrected using gauge data. This is in clear contrast with the MSWEP dataset, where gauge-based products (e.g. the CPC-Unified dataset) are not used in a bias-correction step, but directly combined with other datasets using appropriate weights. This was indeed not clear from the original descriptions at P12-L1-4. We note however that the TMPA 3B42v7 product is described in the NASA website as a 'satellite precipitation product', and thus stating that the GLEAM output is solely based on satellite products as forcing is not – in our opinion – an overstatement. At the same time, we want to bring the reviewer's attention to the thin line separating what is a true satellite observation, since all the satellite observations used here have been calibrated using ground data after all... Nonetheless, and acknowledging the equivocality of this issue, we will avoid the use of 'fully' in the context of satellite-based forcing. We hope the reviewer can agree with this change.

3. MODERATE: It would be interesting to show a version "d" of the v3 dataset in which SMOS observations are assimilated in the product using MSWEP as rainfall input. It would allow to disentangle the impact of rainfall forcing and the assimilated soil moisture product on the final quality of GLEAM datasets.
   This is indeed an interesting experiment, which we have already done in the past to confirm our results about the quality of the different input datasets and the performance of the assimilation algorithm. Replacing for instance the TMPA 3B42v7 for MSWEP in the v3.0c dataset increases the average open loop correlation (i.e. without data assimilation) of the first layer soil moisture against the in situ measurements from 0.61 to 0.66 (note that these statistics might be slightly different from the ones reported in the manuscript due to a reprocessing of the in situ data). This clearly indicates the higher quality of the MSWEP dataset in reference to the TMPA 3B42v7. If SMOS soil moisture observations are assimilated, both soil moisture datasets consistently improve over the CONUS, resulting in a slight increase of the same statistics to 0.62 and 0.67, respectively. These results indicate the high quality of the SMOS soil moisture dataset and the efficiency of the simple Newtonian Nudging algorithm.

   Given that the paper is already quite extensive, including these results would not necessarily contribute to increasing the clarity of the manuscript. In addition, it is also not the main objective of the paper to focus on the quality of the different input datasets, neither to make strong claims about this. Therefore, the authors prefer not to include these results in the paper.

In the specific comments, I added some corrections and suggestions that should be implemented. On this basis, I believe the paper deserves to be published only after a moderate revision.

SPECIFIC COMMENTS (P: page, L: line or lines)

1. P5, L2: I missed how snowmelt is computed. Can the authors add some details?

This module of GLEAM is indeed not described in the paper. However, as this component of the model was not modified in reference to the original version, we would like to point readers to the first description of GLEAM in Miralles et al. (2011). A reference to the latter paper will be added to the revised version of the manuscript.

2. P6, L10: The paper by Lievens et al. (2016) is under review. As it is mentioned in the paper several times, and the readers do not have access to it, I believe some additional details should be included in this paper.
   The paper by Lievens et al. (2016) has recently been accepted and will be published online soon, so we will add the final citation to the revised manuscript.

3. P6, L17-18: This sentence is also repeated below, I suggest removing.
   Thanks. The second sentence will be removed.

4. P8, L12: I believe it should be specified that for w>wc S=1 and for w<wr S=0. Also for equation (5).
   We agree, yet this is already described in the original manuscript (P8-L1-4).

5. P10, L15-22: It is the third time in the paper that the three versions of v3 are described. Please try to avoid repetitions.
   This will be revised.

6. P11, L23-24: Why for a thicker model layer the representativeness of soil moisture measurements is lower? It should be explained.
   As the in situ soil moisture measurement is essentially a point measurement, it becomes less representative for the model if the volume to which it is compared gets larger (i.e. if the model layer gets thicker). When dealing with a 2D surface, the equivalent would be to think of the spatial representativeness of two different spatial resolutions (a coarse and a fine) and how they compare against a point measurement.

7. P11, L32: Likely, it should be stressed also in the abstract that the quality assessment of root-zone soil moisture products is mostly carried out in CONUS region.
   We agree with the reviewer and will include this information in the abstract.

8. P12, L8-9: Strictly speaking, also the first model layer (10 cm) is thicker than the sensing depth of SMOS and ESA CCI soil moisture products. It should be acknowledged.
   This is true and the resulting mismatch should be partly mitigated by the a priori bias removal. However, this is indeed not acknowledged in the paper and will be added in the revised version of the manuscript. We also acknowledge that the penetration depth of these sensors is variable, and can easily exceed 10 cm as well (see e.g.: "*de Jeu, R.A.M and Holmes, T. Derivation of soil moisture sensing depth from microwave satellite sensors, Poster Presentation at the European Geosciences Union General Assembly 2015*").

9. P13, L17: Figure 6 is not described in the text. Remove or add more details.
   The results in Figure 6 were only briefly referred to at P13-L17-18 of the original paper. In the new manuscript, we will further elaborate on the results in Figure 6, but only briefly, since the conclusions are analogous to the ones that may be drawn from Figure 5.

10. P13, L34: The possibility to correct for irrigation that is not modelled in GLEAM is highly interesting. However, it is not shown in the paper and, hence, the sentence should be smoothed.
    The effects of irrigation on soil moisture should be partly captured by satellite-derived soil moisture datasets. As a result, a temporary increase in observed satellite soil moisture will likely result in an increase of the modelled soil moisture after data assimilation. However, since with the current validation data we are unable to detect this effect, we agree with the referee that this statement should be smoothed.

11. Figure 3: Specify explicitly which plot refers to tall and short vegetation.
    We would like to emphasize that the same stress function for short and tall vegetation is implemented in GLEAM v3. Therefore, the panels in Figure 3 do not necessarily refer to either short or tall vegetation, but rather show the effect of the VOD on the stress (a large range in VOD vs. a small range in VOD). We will update the caption in Figure 3 to make this clear.