

Earth Syst. Sci. Data Discuss., author comment AC1  
<https://doi.org/10.5194/essd-2021-84-AC1>, 2021  
© Author(s) 2021. This work is distributed under  
the Creative Commons Attribution 4.0 License.



## Reply on RC1

Yaoping Wang et al.

---

Author comment on "Development of observation-based global multilayer soil moisture products for 1970 to 2016" by Yaoping Wang et al., Earth Syst. Sci. Data Discuss., <https://doi.org/10.5194/essd-2021-84-AC1>, 2021

---

### Reviewer #1

By synthesizing a wide range of SM datasets using statistical methods, this study developed and evaluated seven global, long-term, multi-layer monthly SM products, which are crucial for many research fields. I appreciate the great efforts made by the authors and I believe that this study will definitely contribute to better understanding the global water, energy and biogeochemical cycles. Generally, the topic is very interesting and the material was well organized. However, some clarifications are needed to improve the quality of the manuscript, and I'd like to recommend an acceptance of the manuscript for publication after minor revisions.

**Response:** Thank you for the encouragement.

Comments and suggestions:

- Section 2.1, three time ranges have been used in OLC method, please give necessary explanation on this.

**Response:** Thank you for the comment. An explanation is now provided in Section 2.1 in lines 105–107 (untracked version; tracked version lines 113–116):

"For the OLC method, the ORS datasets were grouped based on three time ranges (1970–2010, 1981–2010, and 1981–2016) that were selected to maximize the available ORS datasets in each time range."

- Section 2.2, please describe the details about the usage of the ISMN observation to train the OLC method.

**Response:** Thank you for the comment. We added some descriptions and a reference to Sect. 2.7, where the OLC method was described fully, in lines 142–144 (untracked version; tracked version lines 151–154):

"After the ISMN observations were aggregated to monthly 0.5° resolutions, 60% of the month-grids were used as the observed SM values in the OLC method (the variable; see

Sect. 2.7), and the remainder were reserved for evaluating all the merged products.”

- Figure 3, what is the purpose of using SMOS L3, SoMo, SMVERGE v2, SOMS L4 and GLEAM v3.3a in the analysis? The relevant discussion is not clear enough.

**Response:** Thank you for your question. We have recognized that the original description for why and how these datasets were used was too brief. We expanded Sect. 2.5 to give a more detailed description of the rationale for using these datasets in lines 188–196 (untracked version; tracked version lines 197–206):

“A recent discussion on the evaluation of coarse-scale soil moisture datasets noted that neither in situ observations, which have limited coverage and small spatial footprint, nor satellite and LSMs, which have retrieval or modeling errors, can be considered fully adequate for evaluation at the global scale; as such, a sound evaluation practice would require combining multiple sources of data (Gruber et al. 2020). Following this recommendation, the merged SM products were evaluated against the reserved 40% in situ observations (Sect. 2.2), as well as a few gridded reanalysis, satellite, and machine learning-upscaled SM datasets. Although the merging process aimed to use as many existing SM datasets as possible, the gridded datasets in Table 1 were not used in the merging because of incompatible vertical resolution, non-global spatial coverage, or short temporal coverage (Table 1). Such evaluation against multi-source gridded datasets complements the evaluation against in situ observations by providing sanity checks on the behavior of the merged products at large scales.”

We also provided a more detailed description of how the evaluation was performed using the SMOS L3, SoMo, SMVERGE v2, SOMS L4, and GLEAM v3.3a datasets in Sect. 2.5 in lines 211–224 (untracked version; tracked version lines 221–234):

“The merged products were evaluated against the validation set of in situ observations and the gridded SM datasets using three common metrics: mean bias (Bias), root mean squared error (RMSE), and Pearson correlation coefficient (Corr). For evaluation against the in situ observations, the metrics were calculated both for the whole validation set and for each land cover type in consideration of the uneven distribution of ISMN observations across land cover types (Figure A1). The observational values used in each calculation were the land-cover-weighted averages (see Sect. 2.2), and the merged values were from the grids and time steps that have the observational values. For evaluation against the SMOS L3 gridded dataset, the 0–10 cm layer of the merged products and the source datasets (ORS, CMIP5, and CMIP6) were used. For evaluation against the other evaluation datasets, the merged and source datasets were linearly interpolated to depths of the evaluation datasets. The annual climatology, mean seasonal anomalies (i.e., the climatology of individual months minus the annual climatology), least-squares linear trends, and anomalies (i.e., the original values minus the mean seasonal cycle and trends) were calculated for each common grid cell and over the common time period between each pair of evaluated and evaluation datasets. Then, for each characteristic (climatology, seasonal cycle, linear trends, or anomalies), the Bias, RMSE, and Corr were calculated using the values of the characteristic pooled over all the common grid cells. When calculating the Bias, RMSE, and Corr for the trends, the insignificant trends at  $p = 0.1$  were set to zero to prevent small random variability from influencing the results.”

- Figure 4, Mean ORS, OLC ORS and EC ORS show large bias after 1988, however, they give high correlations. Why?

**Response:** Thank you for pointing this out. Correlations and bias do not depict the same aspect of errors and can be unrelated to each other. In Figure 5 (originally Figure 4), we used the Spearman rank correlation, which is Pearson correlation between ranks. That is, for a pair of time series, and ( $j = 1, 2, \dots, N$ ), the calculation procedure converts each into

its rank among all the  $x$ -values, and each into its rank among all the  $y$ -values, and then calculates the Pearson correlation between the ranks. Therefore, the Spearman correlation is not sensitive to the original magnitudes of the  $x$ - and  $y$ -series. We chose the Spearman correlation because scPDSI is a normalized index, and its magnitude is not comparable to the magnitude of soil moisture. To clarify the rationale of this choice and its consequence, we added the following text to the description of Sect. 3.4 (originally Sect. 3.3), in lines 441–446 (untracked version; tracked version 443–477):

“To better quantify the similarity between the scPDSI and SM anomalies, Spearman correlations (Hollander et al., 2013) were calculated and are shown above each panel in Figures 4 and A6–A8. The Spearman correlation metric was deemed suitable for measuring the similarity because the magnitudes of scPDSI, which is a unitless standardized index, and of SM anomalies ( $\text{m}^3/\text{m}^3$ ), are not comparable. Spearman correlation is not sensitive to magnitudes because the metric is calculated using the rank of each  $x$ -value among all the  $x$ -values, and the rank of each  $y$ -value among all the  $y$ -values, for an  $x$ - $y$  pair of time series (Hollander et al., 2013).”

We also generally revised Sect. 3.4 (originally Sect. 3.3) to describe Figures 5 (originally Figure 4) and A6–A8 (originally A6–A8) in more detail and clarity.

- Figure 5, too many plots present limited useful information, which should be improved.

**Response:** Thank you for the suggestion. The original intention of Figure 5 was to show that the time series of the merged products did not have temporal discontinuity and had reasonable temporal variability. In recognition of this comment, and the fact that time series plots did not provide quantifiable measures on temporal discontinuity and variability, we added a homogeneity test of whether temporal discontinuity existed in the concatenated datasets (OLC ORS, EC ORS, and EC ALL), replaced Figure 5 with a spectrum analysis, and reduced the number of panels shown in the main text.

The method of the added homogeneity test is described at the end of Sect. 2.9, lines 336–352 (untracked version; tracked version lines 346–362), and the results of the homogeneity test are described in the new Sect. 3.3, lines 407–425 (untracked version; tracked version lines 435–460). In summary, no discontinuities in mean were identified in the merged products. Some discontinuities in variance were identified, but the concatenated datasets (OLC ORS, EC ORS, and EC ALL), which used a few different sets of source datasets for different time periods, were not found to be more discontinuous than the non-concatenated datasets (Mean ORS, EC CMIP5, EC CMIP6, EC CMIP5+6), which used the same source datasets throughout. This result demonstrated that the concatenation practice did not introduce additional discontinuities into the merged data. Discussion about the potential causes of discontinuity in the merged products are added to lines 598–605.

The new spectrum analysis is shown in Figure 6 (originally Figure 5), and the description of Figure 6 in Sect. 3.5 (originally Sect. 3.4), lines 459–476 (untracked version; tracked version 436–454), was re-written to reflect the new figure. The power spectral densities of the ORS-based merged products generally fall within the envelopes of the source datasets. The same was true for the CMIP5- and CMIP6-based merged products, except at the 50–100 cm depth. These results demonstrated that the ORS-based products were better than the CMIP5- and CMIP6-based merged products, and had reasonable temporal variability.

A few lines were added to the abstract to reflect the added homogeneity test and spectrum analysis, lines 25–28 (untracked version; tracked version lines 27–31):

“The merged products generally showed reasonable temporal homogeneity and physically

plausible global sensitivities to observed meteorological factors, except that the ESM-dependent products underestimated the low-frequency temporal variability in SM, and over-estimated the high-frequency variability for the 50–100 cm depth.”

- Section 3.5, the discussion is not deep enough, which provides limited information to the readers.

**Response:** Thank you for the comment. The original figure accompanying the original Sect. 3.5 (new Sect. 3.6) shows globally aggregated probability density distributions, which has limited information content that precluded detailed discussion. We replaced it with a summary of the dominant control on soil moisture at the grid scale (Figure 7, originally Figure 6), and rewrote the entire section (untracked version 504–521; tracked version 548–582). The results showed that precipitation was the dominant control of soil moisture and had generally significant positive correlations with soil moisture. Temperature was a more important control of soil moisture in the CMIP5- and CMIP6-based datasets than in the ORS-based datasets, but in all the merged datasets, any significant correlations between temperature and soil moisture tended to be negative. Solar radiation was an important control of soil moisture only in the CMIP5- and CMIP6-based datasets, in the northern mid-to-high latitudes and tropical forests, which might be caused by light limitation on ecosystems. These results are plausible compared to physical mechanisms and show that the merged products had reasonable sensitivities to meteorological forcings.

Although the original Figure 6 showed seasonal variations in partial correlations, we chose to no longer show the seasonal variations in the new Figure 7, because they will result in too many graphs and are not of any particular importance to the evaluation.