https://doi.org/10.5194/essd-2021-84
Preprint. Discussion started: 16 March 2021

Open Access Earth System Science Data Discussions

Development of Observation-based Global Multi-layer Soil Moisture Products for 1970 to 2016

Yaoping Wang[1,2], Jiafu Mao[2,*], Mingzhou Jin[1,3], Forrest M. Hoffman[4], Xiaoying Shi[2], Stan D. Wullschleger[2], and Yongjiu Dai[5]

[1] Institute for a Secure and Sustainable Environment, University of Tennessee, Knoxville, TN, USA 37902
[2] Environmental Sciences Division and Climate Change Science Institute, Oak Ridge National Laboratory, Oak Ridge, TN, USA 37830
[3] Department of Industrial and Systems Engineering, University of Tennessee, Knoxville, USA 37996
[4] Computational Sciences and Engineering Division and Climate Change Science Institute, Oak Ridge National Laboratory, Oak Ridge, TN, USA 37830
[5] School of Atmospheric Sciences, Sun Yat-sen University, Guangzhou, China 519082

Correspondence to: Jiafu Mao (1 Bethel Valley Rd, Oak Ridge, TN 37830; *maoj@ornl.gov*)

Abstract. Soil moisture (SM) datasets are critical to understanding the global water, energy, and biogeochemical cycles and benefit extensive societal applications. However, individual sources of SM data (e.g., in situ and satellite observations, reanalysis, offline land surface model simulations, Earth system model simulations) have source-specific limitations and biases related to the spatiotemporal continuity, resolutions, and modeling/retrieval assumptions. Here, we developed seven global, gap-free, long-term (1970–2016), multi-layer (0–10, 10–30, 30–50, and 50–100 cm) SM products at monthly 0.5° resolution (available at https://doi.org/10.6084/m9.figshare.13661312.v1) by synthesizing a wide range of SM datasets using three statistical methods (unweighted averaging, optimal linear combination, and emergent constraint). The merged products outperformed their source datasets when evaluated with in situ observations and the latest gridded datasets that did not enter merging because of insufficient spatial, temporal, or soil layer coverage. Assessed against in situ observations, the global mean bias of the synthesized SM data ranged from −0.044 to 0.033 m^3/m^3, root mean squared error from 0.076 to 0.104 m^3/m^3, and Pearson correlation from 0.35 to 0.67. The merged SM datasets also showed the ability to capture historical large-scale drought events and physically plausible global sensitivities to observed meteorological factors. Three of the new SM products, produced by applying any of the three merging methods onto the source datasets excluding the Earth system models, were finally recommended for future applications because of their better performances than the Earth system model–dependent merged estimates. Despite uncertainties in the raw SM datasets and fusion methods, these hybrid products create added value over existing SM datasets because of the performance improvement and harmonized spatial, temporal, and vertical coverages, and they provide a new foundation for scientific investigation and resource management.

https://doi.org/10.5194/essd-2021-84
Preprint. Discussion started: 16 March 2021

1 Introduction

High-quality global soil moisture (SM) datasets benefit many applications, such as understanding drought changes and ecosystem dynamics (Green et al., 2019; Kumar et al., 2019), studying land-atmosphere feedbacks (Li et al., 2020a), benchmarking model capabilities (Loew et al., 2013), and initializing weather and climate forecast systems (Sospedra-Alfonso and Merryfield, 2018). The majority of SM products fall into five categories: in situ measurements, satellite observations, offline land surface model (LSM) simulations, reanalysis, and Earth system model (ESM) simulations. In situ measurements provide the most direct SM observations at the point scale but are too sparse to be interpolated to the global level (spatial autocorrelation dies ~300 km (Gruber et al., 2016)). Satellite-derived SM records only penetrate the top few centimeters of soil and contain errors and spatial gaps typically caused by factors such as a change in path, dense vegetation, frozen soil, water bodies, and radio frequency interference (Llamas et al., 2020; Wang et al., 2012). Although a long-term (1979–present) concatenated SM dataset was developed by merging data from multiple satellites, its spatial gaps remain unresolved (Dorigo et al., 2012). The SM in LSM simulations have complete spatial and temporal coverage, which is convenient for regional and global analysis (Gu et al., 2019); however, LSM simulations may contain considerable uncertainties because of inadequacies in the model physics, parameterization, and drivers (Andresen et al., 2020). Reanalysis datasets assimilate observations into LSMs or coupled forecast systems that have LSMs as a component, and are gap-free. Direct assimilation of remote-sensing SM, which has been the practice for some recent reanalysis—such as ECMWF Reanalysis 5 (ERA5) (de Rosnay et al., 2013) and Global Land Evaporation Amsterdam Model (GLEAM) (Martens et al., 2017)—is likely to improve the performance relative to free-running LSMs. Still, many reanalyses do not directly assimilate the observational SM, such as the Japanese 55-Year Reanalysis (JRA55) (Kobayashi et al., 2015) and Modern-Era Retrospective Analysis for Research and Applications Version 2 (MERRA2) (McCarty et al., 2016). Also, the meteorological variables, especially precipitation, simulated by the atmosphere model of the coupled reanalysis system may be biased, leading to inaccurate SM estimates by the intrinsic LSM component (Balsamo et al., 2015). Fully coupled ESMs, such as those for the Coupled Model Intercomparison Project phases 5 and 6 (CMIP5 and CMIP6) (Eyring et al., 2016; Taylor et al., 2012), provide SM simulations for both historical and future periods. ESMs, however, share the same uncertainty sources for the SM estimates as the LSMs; moreover, the SM in ESM simulations have internal variability-related uncertainties induced by unrealistic initialization from the preindustrial conditions rather than the real world (Eyring et al., 2016; Taylor et al., 2012).

There is active development toward generating more accurate, gap-free SM datasets. Methods for filling the spatial gaps in satellite observations have been under investigation, but the resulting estimates either cover short time periods or target only parts of the globe (Llamas et al., 2020; Wang et al., 2012). One global multi-layer SM product was generated by upscaling in situ observations using machine learning and selected SM predictors; however, it only focused on 2000–2019 (O and Orth, 2020). Unlike current global reanalyses that directly assimilate satellite SM (Martens et al., 2017; de Rosnay et al., 2013), some studies merged in situ observations, or both in situ and satellite observations, with offline LSMs to improve accuracy

https://doi.org/10.5194/essd-2021-84
Preprint. Discussion started: 16 March 2021

Open Access Earth System Science Data Discussions

while retaining complete spatiotemporal coverage. Nevertheless, these efforts were mainly conducted on the regional scale using limited sets of LSMs (Wu et al., 2018; Zeng et al., 2016; Gruber et al., 2016).

Therefore, the need exists to develop global merged SM products that comprehensively combine the information from the latest in situ and satellite observations, offline LSMs, reanalysis, and ESMs using advanced fusion methods. Constrained by various types of observations, the merged products would likely perform better than the SM in the original LSMs or ESMs while being gap-free in space and having long temporal and multi–soil-layer coverage. The fusion of multiple LSMs, reanalysis, and ESMs also involves ensemble averaging, which may reduce the SM uncertainties from individual models by cancelling the model-specific errors (Giorgi and Mearns, 2002). This study presents a group of SM products derived using three merging methods: unweighted averaging, optimal linear combination (OLC), and emergent constraint (EC). Unweighted averaging assigns equal weight to all the source datasets and does not use in situ information. The OLC is an ensemble weighting and rescaling algorithm that is optimal in the sense that the weighted average minimizes the mean squared difference with respect to the site-level observations (Bishop and Abramowitz, 2013). The OLC method was previously found to lead to improved performance in the merged product relative to the source datasets in terms of the global evapotranspiration and runoff (Hobeichi et al., 2018, 2019). The EC method is common for reducing uncertainty in future ESM simulations (Mystakidis et al., 2016; Padrón et al., 2019). This method first establishes physically meaningful and statistically significant relationship between the constraint variables that have observations and a target variable that has no observations across multiple ESMs, and then uses the relationship and actual observations to constrain the target variable (Mystakidis et al., 2016; Padrón et al., 2019). Given the clear physical relationships between the SM and meteorological variables, we hypothesized in this study that the EC method can be applied to reduce forcing-related biases in offline LSMs and reanalysis, and to align the natural internal variability in ESMs with the real world. Seven new monthly multi-layer SM datasets at 0.5°×0.5° resolution for 1971–2016 were produced by implementing these merging algorithms onto different combinations of the mentioned raw SM estimates. The merged products with different setups were then systematically evaluated against in situ measurements that were reserved for validation, semi-independent gridded SM datasets, drought indices, and meteorological variables.

2 Methods and source datasets

2.1 Overview

Figure 1 shows the schematic of the merging procedure to create the seven SM products. The unweighted averaging and OLC (Hobeichi et al., 2018, 2019) methods were applied over the observational or observation-forced datasets (i.e., **o**ffline LSMs, **r**eanalysis, **s**atellite [ORS]). The EC method (Mystakidis et al., 2016) was applied over both the ORS datasets and the CMIP5/CMIP6 simulations (Eyring et al., 2016; Taylor et al., 2012). This differential treatment between the ORS and ESM datasets was because the OLC algorithm, which only uses sparse in situ SM observations, provides weaker constraint than the EC method, which uses gridded global meteorological observations. The weaker constraint was found to be inadequate for the ESM simulations in a preliminary analysis (results not shown). Because the ORS datasets do not have uniform temporal

https://doi.org/10.5194/essd-2021-84
Preprint. Discussion started: 16 March 2021

coverage, the unweighted averaging only used the ORS datasets that cover 1970–2016. For the OLC method, the ORS datasets were grouped into three time ranges (1970–2010, 1981–2010, and 1981–2016), separately processed following a previous method for concatenating the remote-sensing SM (Dorigo et al., 2017; Liu et al., 2011, 2012). The CMIP5 and CMIP6 datasets always cover 1970–2016; when they were used jointly with the ORS datasets to produce the EC ALL (i.e., including all the source datasets) product, they were subset to the same time ranges as the ORS datasets, separately processed and concatenated (bottom of Figure 1). All the synthesized monthly SM datasets are at 0.5° resolution, cover 1970–2016, and contain four depths (0–10, 10–30, 30–50, and 50–100 cm). The following sections provide more details of the datasets, merging methods, and processing procedures.

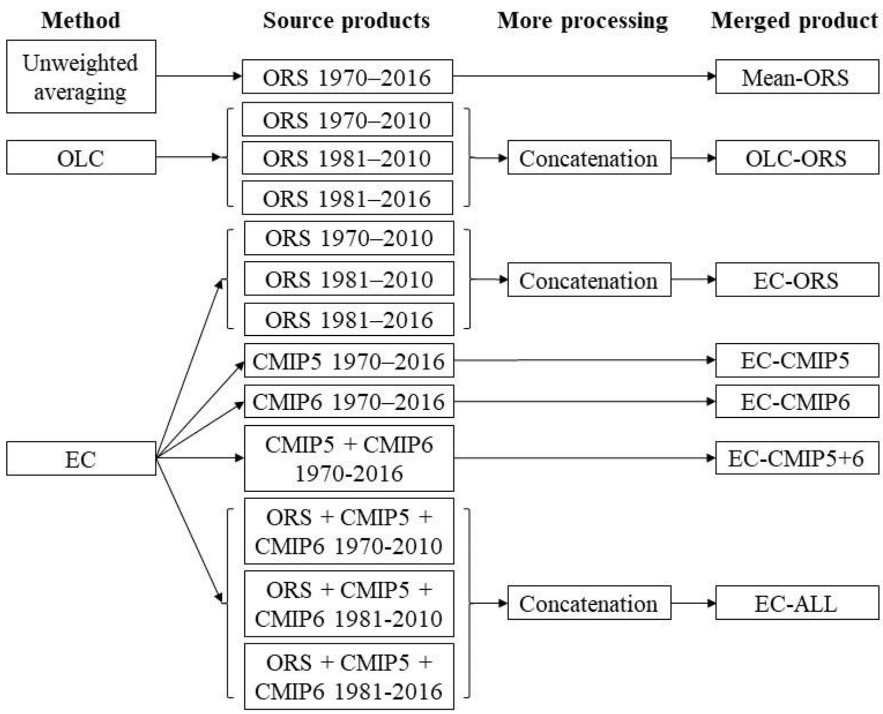

Figure 1: Procedure of creating the merged SM products using different methods and source products.

2.2 In situ SM observations

In situ SM observations were obtained from the International Soil Moisture Network (ISMN) (Dorigo et al., 2011, 2013). Only the observations associated with the ISMN quality flags “G” (good) or “M” (parameter value missing) were retained. The resulting dataset contains ~1,400 stations worldwide and spans from 1964 to the present. Because only a few stations were available at the beginning of the time period, only the observations in 1970 or later were used. To facilitate processing by the OLC method, the ISMN observations were aggregated to monthly 0.5° resolution and regular depths (0–10, 10–30, 30–50, and 50–100 cm). The aggregation to monthly resolution was simply averaging over all the available observations in each month at each station. Although it is desirable to apply a stricter criterion in the monthly averaging, such as treating a month

https://doi.org/10.5194/essd-2021-84
Preprint. Discussion started: 16 March 2021

as missing if observations exist fewer than 15 days in the month, applying such a criterion would exclude most of the stations in northern and eastern Asia, which only have 1–3 observations per month. Multiple methods were tested for the aggregation to 0.5° resolution: (1) simply averaging all the stations in each grid; (2) weighted-averaging the stations based on the percentage of grid area that the land cover of each station represents, using the Moderate Resolution Imaging Spectroradiometer (MODIS) MCD12C1 product (Friedl and Sulla-Menashe, 2015); and (3) the same as (2) except that if the total area of all the land covers that the stations represent does not account for a sufficient percentage of the grid area (40%), the grid was set to missing. Because all three methods resulted in similar performance in the merged products (results not shown), only the second method was adopted for the final products. The aggregation to regular depths was simply to average over all the available observations in each depth interval. When an observation was taken on the interface of two depth intervals (e.g., exactly at 10 cm), the observation was assigned to the shallower depth interval (e.g., 0–10 cm). Figure S1 shows the aggregated ISMN observations at the 0.5° grid scale, and the number of observations that falls into each land cover type is displayed in Figure S2. The observations are overrepresented in the developed part of the world; the most overrepresented land cover types include the deciduous broadleaf forests, grasslands, and croplands, whereas the most underrepresented land cover types are evergreen broadleaf forests, mixed forests, closed shrublands, and permanent wetlands. The number of available monthly observations did not decrease with deeper soil layers (global totals are 19,317 station-months for the 0–10 cm layer, 25,307 for the 10–30 cm layer, 25,011 for the 30–50 cm layer, and 20,660 for the 50–100 cm layer), although no observations were available for the closed shrublands and permanent wetlands across the deeper soil layers (10–30, 30–50, and 50–100 cm). After the ISMN observations were aggregated to monthly 0.5° resolutions, 60% of the month-grids were used to train the OLC method, and the remainder were reserved for validating all the merged products. The training month-grids were uniformly randomly selected without distinguishing between space and time.

2.3 ORS, CMIP5, and CMIP6 SM products

Tables S1–S4 list the monthly gridded ORS, CMIP5, and CMIP6 results that were used as the SM source datasets for the merged products, as well as the depths (0–10, 10–30, 30–50, or 50–100 cm) and time ranges (1970–2016, 1970–2010, 1981–2010, or 1981–2016) for which the datasets were used. The name of the used SM variable was "mrlsl" in the CMIP5 collection and "mrsol" in the CMIP6 collection. Prior to merging, all the ORS, CMIP5, and CMIP6 datasets were bilinearly interpolated to 0.5° resolution, linearly interpolated to the target soil depths (0–10, 10–30, 30–50, or 50–100 cm), and masked with a common land mask using the National Center for Atmospheric Research (NCAR) Command Language 6.6.2 (UCAR/NCAR/CISL/TDD, 2019). Linear interpolation to all four soil depths could not be achieved for all the source datasets because the soil layers in some models are too shallow or too coarse. For example, the bottom depths of the Joint UK Land Environment Simulator (JULES) LSM are 10 cm, 35 cm, 1 m, and 3 m (Sebastian Lienert, 2019, Personal communication). Therefore, the 10–30 and 50–100 cm depths are contained within single layers in the JULES model and could not be directly interpolated. The four target soil depths here were selected to maximize the number of source datasets that could be interpolated to each depth. Although the Community Land Model version 4 (CLM4) simulates SM as deep as 421 cm, the model was only used for 0–10 and 10–30 cm (Table S3). This was because the SM values at deeper than 38 cm were very low (<0.0005 m^3/m^3 on global average), and even lower than the SM values at shallower than 38 cm in the same model (>0.0025 m^3/m^3 on global average). The common land mask was created by intersecting the grid points that satisfy two criteria: (1) all the datasets except the European Space Agency Climate Change Initiative (ESA CCI) v4.5, after being interpolated to 0.5°, have valid values; and (2) at least 50% of the land cover is not water bodies or permanent snow and ice in the MODIS MCD12C1 product (Friedl and Sulla-Menashe, 2015). Because of many spatial gaps, the ESA CCI v4.5 dataset was not used to create the common land mask and received special handling. In addition to being masked with the common land mask, the ESA CCI v4.5 dataset was masked using its accompanying quality flags, meaning at the time steps and grids that have snow coverage or temperature below zero (flag = 1), dense vegetation (flag = 2), no valid SM estimates (flag = 4), SM values above physical boundary (flag = 8), or only unreliable SM values (flag = 16). The ESA CCI v4.5 dataset was excluded from merging at the time steps and grids in which the missing values exist.

2.4 Observed, CMIP5, and CMIP6 temperature and precipitation

The EC method, as implemented in this study (Sect. 2.8), requires the air temperature and precipitation forcings that correspond to each SM dataset, and observed air temperature and precipitation as inputs. The temperature and precipitation forcings that correspond to the ORS SM datasets are listed in Table S5. Because the ESA CCI v4.5 dataset is observational and the GLEAM v3.3a directly assimilates the ESA CCI dataset (Dorigo et al., 2017), these two datasets were assumed to correspond to the observed temperature and precipitation in the Climate Research Unit (CRU) TS v4.03 dataset (Harris et al., 2014). The temperature and precipitation forcings for the various reanalysis datasets were obtained from the same reanalysis. The temperature and precipitation forcings for the Multi-scale Synthesis and Terrestrial Model Intercomparison Project (MsTMIP) collection of LSMs were from the CRU NCEP v4 dataset, which, at monthly level, is equal to the CRU TS v3.20 dataset (Huntzinger et al., 2018). The temperature and precipitation forcings for the Trends and Drivers of the Regional Scale Sources and Sinks of Carbon Dioxide version 7 (TRENDY v7) collection of LSMs were from the CRU TS v3.26 dataset (Sitch and Friedlingstein, 2019, Personal Communication). The temperature and precipitation datasets that correspond to the CMIP5 and CMIP6 SM were from the same ESMs and ensemble members (Table S4). For observed air temperature and precipitation, the CRU TS v4.03 dataset (Harris et al., 2014) was used. All the temperature datasets were bilinearly interpolated, and the precipitation datasets conservatively interpolated, to 0.5° resolution using the NCAR Command Language 6.6.2 (UCAR/NCAR/CISL/TDD, 2019). All the temperature and precipitation datasets were limited to the same common land mask as the SM products.

2.5 In situ and gridded SM datasets for evaluation of the merged products

In addition to the validation on the reserved 40% in situ observations (Sect. 2.2), the merged SM products were evaluated against a few global and regional gridded SM datasets that were not used in the merging process because of incompatible vertical resolution or short temporal coverage (Table 1). These evaluations complement the evaluation against in situ

https://doi.org/10.5194/essd-2021-84
Preprint. Discussion started: 16 March 2021

observations by providing sanity checks on the behavior of the merged products at a large scale. All the evaluation datasets were bilinearly interpolated to 0.5°×0.5° and aggregated to a monthly level using the NCAR Command Language 6.6.2 (UCAR/NCAR/CISL/TDD, 2019). For the Soil Moisture and Ocean Salinity (SMOS) L3 dataset, which was available as monthly aggregates (https://www.catds.fr/sipad/), only the data points with retrieval error (i.e., the "DQX" field) < 0.07 m^3 m^{-3} were used, and the ascending (MIR_CLM4RA) and descending (MIR_CLM4RD) orbits were averaged. For the SMOS L4 dataset, which was only available at daily resolution, the days with a quality index of 1 (highest quality) were used, and the monthly averaging was restricted to the months with less than 13 missing days to be consistent with the downloaded monthly SMOS L3 dataset. The other evaluation datasets do not have gaps, so the aggregation to monthly level was straightforward. The SoMo.ml SM is mainly upscaled from the ISMN dataset (O and Orth, 2020) and therefore not totally independent from the OLC ORS merged product, but independent from the EC-based merged products. The GLEAM v3.3a 0–100 cm dataset is not independent from the ORS- or ALL-based merged products, which use the 0–10 cm part of the GLEAM v3.3a dataset (Table S2), but is independent from the CMIP5- or CMIP6-based merged products.

Table 1: Global and regional datasets that were compared against the merged SM products.

Dataset	Type	Period	Depth (cm)	Resolution	Coverage	Reference
SMOS L3 RE04 MIR_CLF3MA, MIR_CLF3MD	Satellite	2010–2020	Surface (0–5)	~25 km	Global with missing values	(Al Bitar et al., 2017)
SMOS L4 SCIE MIR_CLM4RD	Reanalysis	2010–2020	0–100	~25 km	Global with missing values	(Al Bitar et al., 2013)
GLEAM v3.3a	Reanalysis	1980–2018	0–100	0.25°	Global	(Martens et al., 2017)
SMERGE v2	Reanalysis	1979–2019	0–40	0.125°	Contiguous United States	(Tobin et al., 2017)
SoMo.ml	Machine learning upscaled from in situ observations	2000–2019	0–10, 10–30, 30–50	0.25°	Global	(O and Orth, 2020)

2.6 Drought and meteorological datasets for evaluation of the merged products

In situ observations are sparse and represent much smaller spatial scales than the 0.5° grid of the merged products. Other global and regional SM datasets have short temporal coverage and spatial gaps in the evaluation datasets, and non-independence between the merged products and evaluation datasets. Given these limitations, and to further ascertain the quality of the merged products, the new SM products were evaluated using process-based observational metrics, including the responses to prominent historical drought events and historical climate change (e.g., precipitation, temperature, downwelling shortwave radiation). The selected historical drought events were the US drought of 1985–1992 and the Australian millennium drought of 2002–2009 (Spinoni et al., 2019). A Self-Calibrated Palmer Drought Severity Index (scPDSI) dataset (Dai et al., 2004) was used as the benchmark for the evolution of these two drought events. The precipitation, temperature, and downwelling shortwave radiation datasets are from the Global Soil Wetness Project (GWSP) version 3 reanalysis (Dirmeyer et al., 2006), which provide some independence from the CRU TS v4.03 temperature and precipitation used in the EC method (Sect. 2.4).

https://doi.org/10.5194/essd-2021-84
Preprint. Discussion started: 16 March 2021

The SM climatic sensitivities were derived using the partial correlations with each meteorological variable calculated conditional on the other two variables.

2.7 OLC method

Let x_k^{tj} stand for the SM value of the source dataset k (k = 1, 2, …, K) at time step t (t = 1, 2, …, T) and grid j (j = 1, 2, …, S), and o^{tj} be the observed SM values at time step t and grid j where $(t, j) \in V$ and V is the subset of grids and time steps that have observed SM. OLC calculates the final estimated SM (μ_e^{tj}) as a weighted average, with w_k denoting the weight of source dataset k, using Eq. (1):

$$\mu_e^{tj} = \sum_{k=1}^{K} w_k x_k^{tj} \tag{1}$$

The optimal vector of weights for the source datasets, $\boldsymbol{w} = [w_1, w_2, \dots, w_K]^T$, which minimizes the mean squared error subject to $\sum_{k=1}^{K} w_k = 1$, is a function of the error covariance matrix of the source datasets ($\boldsymbol{A}$) following Eq. (2):

$$\boldsymbol{w} = \frac{\boldsymbol{A}^{-1}\mathbf{1}}{\mathbf{1}^T\boldsymbol{A}^{-1}\mathbf{1}} \tag{2}$$

The OLC procedure without a constant term requires the source datasets to be unbiased (Bishop and Abramowitz, 2013), but the ORS datasets are biased relative to the in situ observations. Therefore, to prevent the biases from influencing the weights, the error covariance matrix $\boldsymbol{A}$ was the covariance matrix of locally centered errors (e_k^{tj}), which were calculated following Eq. (3):

$$e_k^{tj} = \left(x_k^{tj} - x_k^{\cdot j}\right) - \left(o^{tj} - o^{\cdot j}\right) \quad , \quad (t, j) \in V \tag{3}$$

where $x_k^{\cdot j}$ is the time-averaged SM value of the source dataset k at grid j, and $o^{\cdot j}$ is the time-averaged observed SM value at grid j. The time averaging was over the time steps in which observed SM values exist in each grid j. The grids and time steps for which the centered errors exist were pooled together to create a single vector of errors for each source dataset. This vector of errors e_k^{tj} was then used to calculate the error covariance matrix $\boldsymbol{A}$. Therefore, the derived weights were optimal with regard to the locally centered errors, not the un-centered errors $\left(x_k^{tj} - o^{tj}\right)$. However, this limitation had to be accepted because the ISMN observations were too sparse to enable estimating the biases at every grid in space, and a constant bias could not be assumed over the large spatial domain of the study, which spans very dry to very wet climate zones.

The OLC procedure was implemented in Python 3.6.3 under a CentOS Linux environment. Different functions for calculating the error covariance matrix in Python were compared to eliminate potential numerical instability in matrix inversion. The results were found to be similar, but the ShrunkCovariance function in Scikit-learn v0.21.3 (Pedregosa et al., 2011) generated slightly better validation performance for the estimated SM than the other tested functions. Therefore, ShrunkCovariance was selected for calculating $\boldsymbol{A}$.

https://doi.org/10.5194/essd-2021-84
Preprint. Discussion started: 16 March 2021

In addition to estimating SM, the OLC procedure also calculates the associated uncertainty in the form of standard deviation (σ_e^{tj}) based on adjusted weights ($\tilde{w}_k$), adjusted source SM datasets ($\tilde{x}_k^{tj}$), and the estimated SM (μ_e^{tj}), following Eq. (4):

$$\sigma_e^{tj} = \sqrt{\sum_{k=1}^{K} \tilde{w}_k \left(\tilde{x}_k^{tj} - \mu_e^{tj}\right)^2} \tag{4}$$

The adjusted weights are a function of the original weights (w_k) and a parameter α to have $\tilde{w}_k \geq 0$ and maintain $\sum_{k=1}^{K} \tilde{w}_k = 1$, following Eqs. (5) and (6):

$$\tilde{w}_k = \frac{w_k + (\alpha - 1)\frac{1}{K}}{\alpha} \tag{5}$$

$$\alpha = \begin{cases} 1 & \text{if all } w_k \text{ are nonnegative} \\ 1 - K\min(w_k) & \text{if the minimum } w_k \text{ is negative} \end{cases} \tag{6}$$

The adjusted source SM datasets ($\tilde{x}_k^{tj}$) are linear functions of the source SM datasets (x_k^{tj}) and the estimated SM (μ_e^{tj}) through parameters α and β, where the parameter β is a function of the discrepancy between the observations and the estimated SM (s_e^2), following Eqs. (7)–(9):

$$\tilde{x}_k^{tj} = \mu_e^{tj} + \beta\left(x_k^{\cdot j} + \alpha\left(x_k^{tj} - x_k^{\cdot j}\right) - \mu_e^{tj}\right) \tag{7}$$

$$\beta = \sqrt{\frac{s_e^2}{\frac{1}{N}\sum_{j,t\in V}\sum_{k=1}^{K} \tilde{w}_k \left(x_k^{\cdot j} + \alpha\left(x_k^{tj} - x_k^{\cdot j}\right) - \mu_e^{tj}\right)^2}} \tag{8}$$

$$s_e^2 = \frac{\sum_{j,t\in V}\left(\mu_e^{tj} - o^{tj}\right)^2}{N - 1} \tag{9}$$

where V is the subset of grid and time step combinations that have observed SM, and N is the total number of grid-time steps that have observed SM (i.e., $N = |V|$).

2.8 EC method

For establishing the EC relationship, temperature and precipitation were selected to be the constraint variables because of their significant roles in controlling evapotranspiration from and recharge to the soil water. SM was the target variable. For each target year, month, grid, and soil depth, a linear regression relationship was fitted using SM anomalies as the predictand, and temperature and precipitation anomalies as the predictors. The SM, temperature, and precipitation anomalies of each source dataset (i.e., the datasets in the ORS, CMIP5, CMIP6, CMIP5+6, or ALL group; Figure 1) were calculated by removing the monthly climatology of 1981–2010. The vectors of SM, temperature, and precipitation anomalies in each regression relationship consisted of the anomalies for the target year, month, and soil depth over the 9 nearest grids to the target grid and over all the source datasets. If the fitted regression slopes of both temperature and precipitation anomalies were significant at

https://doi.org/10.5194/essd-2021-84
Preprint. Discussion started: 16 March 2021

the $p = 0.05$ level, the fitted regression was used as the EC relationship. If either slope was insignificant, the regression was refitted using only precipitation or temperature anomalies as the predictor. If the refitted slope of precipitation (temperature) anomalies was significant at the $p = 0.05$ level and had lower p-value than the refitted slope of temperature (precipitation) anomalies, the refitted regression with precipitation (temperature) anomalies was used as the EC relationship. If neither of the refitted slopes was significant at the $p = 0.05$ level, the EC relationship was deemed insignificant for this year, month, grid, and soil depth. After the significant EC relationships were obtained, the observed temperature and precipitation were converted to anomalies relative to the monthly climatology of 1981–2010, and fed into the EC relationships to generate constrained SM anomalies. Finally, the constrained SM anomalies were added to the mean monthly climatology over all the source datasets to generate constrained SM values. For the combinations of year, month, grid, and soil depth that did not have significant EC relationships, the mean monthly climatology over all the source datasets was used as the constrained SM values. Uncertainties in the EC-constrained SM values were estimated using standard deviation of the prediction of the linear regressions, calculated by the "wls_prediction_std" function of the Python package statsmodels (Seabold and Perktold, 2010). Uncertainties in which there were no significant EC relationships were estimated using the standard deviation of the source datasets.

The fitted EC relationships are summarized in Figures S3 and S4 using the average values of the significant regression coefficients, and the percentage of significant regression coefficients for temperature and precipitation, respectively. The regression coefficients for temperature were mostly negative, and for precipitation mostly positive, which can be explained by the fact that higher temperature causes higher evaporative loss of water from soil, and higher precipitation causes more recharge of water into soil. In the Sahara region, the average regression coefficients were mostly positive for temperature (Figure S3), which might be related to interannual correlation between precipitation and temperature caused by the West African monsoon (Zhang and Cook, 2014). For the ORS datasets at 30–50 and 50–100 cm, the regression coefficients were also mostly positive for temperature (Figure S3). The ORS datasets at these depths only represent three sets of meteorological forcings (GLDAS NOAH025_M2.0, CRU TS v3.20 for MsTMIP, and CRU TS v3.26 for TRENDY v7; Tables S1–S3 and S5). Therefore, the EC relationships for the ORS datasets at these depths likely have high uncertainty. Because only small percentages of EC relationships were significant at these depths (Figures S3 and S4), the counterintuitive regression coefficients were unlikely to have a large impact on the merged product. In general, the percentages of significant EC relationships were higher for the CMIP5, CMIP6, CMIP5+6, and ALL datasets than for the ORS datasets, suggesting that more diverse source datasets lead to stronger EC relationships (Figures S3 and S4). In the preliminary analysis, additional setups for the regression were tested, including whether the regression should use the actual values of SM, temperature, and precipitation or the anomalies, and whether the regression should use only the target grid or the 9 nearest grids. The setup with anomalies and the 9 nearest grids was found to result in more significant regression coefficients and better performance (results not shown).

https://doi.org/10.5194/essd-2021-84
Preprint. Discussion started: 16 March 2021

2.9 Temporal concatenation

Because some ORS datasets were not available for the entire 1970–2016 period, separate datasets were produced for three different periods (1970–2010, 1981–2016, and 1981–2010) and were concatenated into a continuous 1970–2016 product using a previous intercalibration approach (Figure 1) (Dorigo et al., 2017; Liu et al., 2011, 2012). To perform the concatenation on the estimated SM values, the merged product for each of the three periods was decomposed into monthly climatology and monthly anomalies, with the monthly climatology being calculated on the overlapping period (1981–2010). Then, the anomalies of the 1970–2010 and 1981–2010 product were rescaled to have the same cumulative distribution function (CDF) as the anomalies of the 1981–2016 product during their overlapping period (1981–2010) using the piecewise linear CDF matching technique (Liu et al., 2011). In the first step of the CDF matching technique, the 0th, 5th, 10th, 20th, 30th, 40th, 50th, 60th, 70th, 80th, 90th, 95th, and 100th percentiles of the anomalies of each product during the overlapping period were identified on their CDF curves. In the second step, the percentiles of the 1970–2010 and 1981–2010 products were plotted against the percentiles of the 1981–2016 product. A linear line was drawn between each two adjacent percentiles (e.g., the 5th and 10th percentiles), resulting in 12 linear segments. In the last step, the anomalies of the 1970–2010 and 1981–2010 datasets that fell into each interval of percentiles (e.g., 5th–10th) were rescaled using the equations of the linear segments. Values outside the range of the monthly anomalies during the overlapping period were rescaled using the equation of the closest linear segment. A graphic illustration of the CDF matching technique can be found in Fig. 3 of Liu et al. (2011). The CDF matching was conducted for all the months as a whole, rather than for each month separately, because the latter setup would result in too few data points (36 data points during 1981–2016) to robustly determine the percentiles. The rescaled anomalies were added back to the monthly climatology of each product. Finally, the three added-back products were concatenated by using the 1970–2010 product for 1970–1980, the 1981–2010 product for 1981–2010, and the 1981–2016 product for 2010–2016. To concatenate the estimated SM uncertainty of the OLC method (σ_e^{tj}; see Sect. 2.7) and the EC method (see Sect. 2.8), the uncertainty values of 1970–2010 and 1981–2010 were directly rescaled to the 1981–2016 values using the CDF matching without prior conversion to monthly anomalies, and the rescaled uncertainty values concatenated like the mean values.

3 Results

3.1 Evaluation against the validation set of ISMN SM observations

The merged products were evaluated against the validation set of ISMN observations using common metrics: mean bias (Bias), root mean squared error (RMSE), and Pearson correlation coefficient (Corr). The metrics were calculated both for the whole validation set and for each land cover type separately in view of the uneven distribution of ISMN observations across land cover types (Figure S1). When evaluated on the whole validation set, the Bias of the merged products ranged from −0.044 to 0.033, the RMSE ranged from 0.076 m^3/m^3 to 0.104 m^3/m^3, and the Corr ranged from 0.35 to 0.67, across the four soil depths (Figure 2). The merged products generally showed smaller magnitude of Bias, smaller RMSE, and higher Corr than the source

https://doi.org/10.5194/essd-2021-84
Preprint. Discussion started: 16 March 2021

Open Access Earth System Science Data Discussions

datasets from which the products were merged (Figure 2). The Bias of both the source and merged datasets shifted from mostly positive to mostly negative from the shallowest to the deepest soil layer, indicating a tendency toward the overestimation of vertical SM gradient; the shallower soil layers also tended to have lower RMSE and higher Corr than the deeper layers (Figure 2). The Bias values of individual merged products were similar; the RMSE and Corr values of the ORS-based merged products (Mean ORS, OLC ORS, EC ORS) were better than the EC ALL product; and the RMSE and Corr values of the EC ALL product were better than the CMIP5- or CMIP6-based merged products (EC CMIP5, EC CMIP6, EC CMIP5+6) (Figure 2).

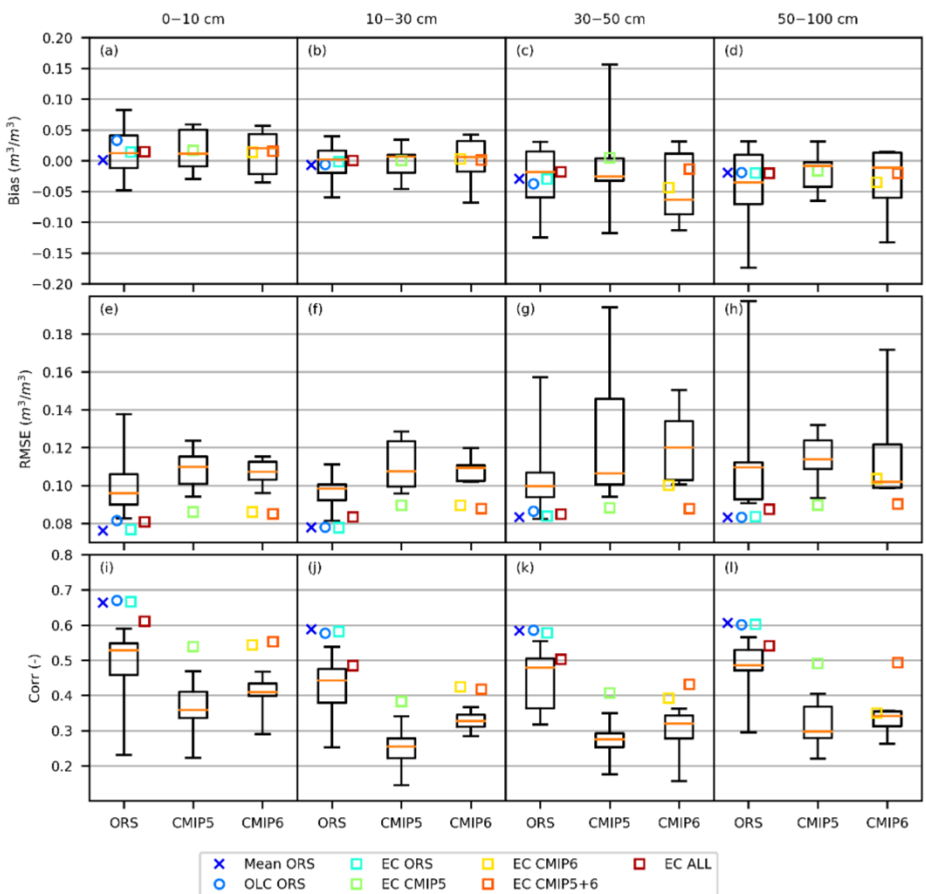

Figure 2: Performance of the original ORS, CMIP5, and CMIP6 datasets (boxplots) and the merged products (scatter plots) on the validation set of observations. The boxplots show (from top to bottom) maximum, 75th percentile, median, 25th percentile, and minimum. The ORS boxplot includes all the ORS datasets evaluated on their available years.

The merged products showed lower magnitude of Bias, lower RMSE, and higher Corr than the source datasets across most of the land cover types (Figure S5). The exceptions were the shallower (0–10 and 10–30 cm) soil layers over the water bodies, evergreen needleleaf forests, and evergreen broadleaf forests, where the merged products produced similar RMSE to the bulk of the source datasets (Figure S5e–h), and the deeper (30–50 and/or 50–100 cm) soil layers over the open shrublands, urban and built-up lands, cropland/natural vegetation mosaics, and barren lands, where the merged products showed similar or even

https://doi.org/10.5194/essd-2021-84
Preprint. Discussion started: 16 March 2021

lower Corr, and similar RMSE to the bulk of the source datasets (Figure S5e–i). Although the merged datasets considerably overestimated the SM in the water bodies and evergreen needleleaf forests (Bias = 0.016 to 0.146 m^3/m^3; Figure S5a–d), the high Corr for these two land cover types (0.30 to 0.72 for the ORS-based merged products, −0.27 to 0.55 for the other merged products; Figure S5i–l) indicated good ability to track the spatio-temporal variability. The hybrid products underestimated the SM in the 0–10 cm layer of the evergreen broadleaf forests (Bias = −0.174 to −0.095 m^3/m^3), the deciduous needleleaf forests (Bias = −0.162 to −0.055 m^3/m^3), and the deeper soil layers of many other land cover types (Figure S5a–d). The Corr values of the merged products were also very low in the evergreen broadleaf forests (−0.81 to 0.05; Figure S5i–l). Similar to at the global level, the ORS-based merged SM tended to outperform the EC ALL and the CMIP5- and CMIP6-based merged products (Figure S5). The three merging methods performed similarly over most land cover types, but the OLC method (OLC ORS product) had lower RMSE and higher Corr than the other two methods (Mean ORS and EC ORS products) over the urban and built-up lands, crop/natural vegetation mosaic, and barren land cover types (Figure S5).

3.2 Evaluation against global and regional gridded SM datasets

The evaluating results for the merged SM products against multiple datasets were more dependent on the evaluation dataset than the merged product (Figure 3). Nevertheless, if each evaluation dataset was viewed separately, the merged products generally had lower RMSE and higher Corr than the average RMSE and Corr of corresponding source datasets. When evaluated against the SMOS L3 and L4 SM, the merged products had high Bias in climatology (0.097 to 0.17 m^3/m^3), generally high RMSE (climatology: 0.122 to 0.194 m^3/m^3, seasonality: 0.024 to 0.049 m^3/m^3, trend: 0.047 to 0.207 m^3/m^3 per 50 years, anomalies: 0.040 to 0.045 m^3/m^3), and low Corr (climatology: 0.07 to 0.35, seasonality: 0.28 to 0.56, trend: 0.006 to 0.13, anomalies: 0.03 to 0.28). When evaluated against the global machine learning–derived SoMo.ml and the GLEAM v3.3a root zone SM, the merged products had high RMSE (0.058 to 0.092 m^3/m^3) and high Corr (0.67–0.86) in climatology; the lowest Corr occurred in spatial trends (−0.016 to 0.46). When evaluated against the US-only Soil MERGE (SMERGE) v2 SM, the merged products had low RMSE (0.042 to 0.051 m^3/m^3) and intermediate Corr (0.50 to 0.62) in climatology; the lowest Corr occurred in the spatial trends (0.12 to 0.34).

https://doi.org/10.5194/essd-2021-84
Preprint. Discussion started: 16 March 2021

Open Access Earth System Science Data Discussions

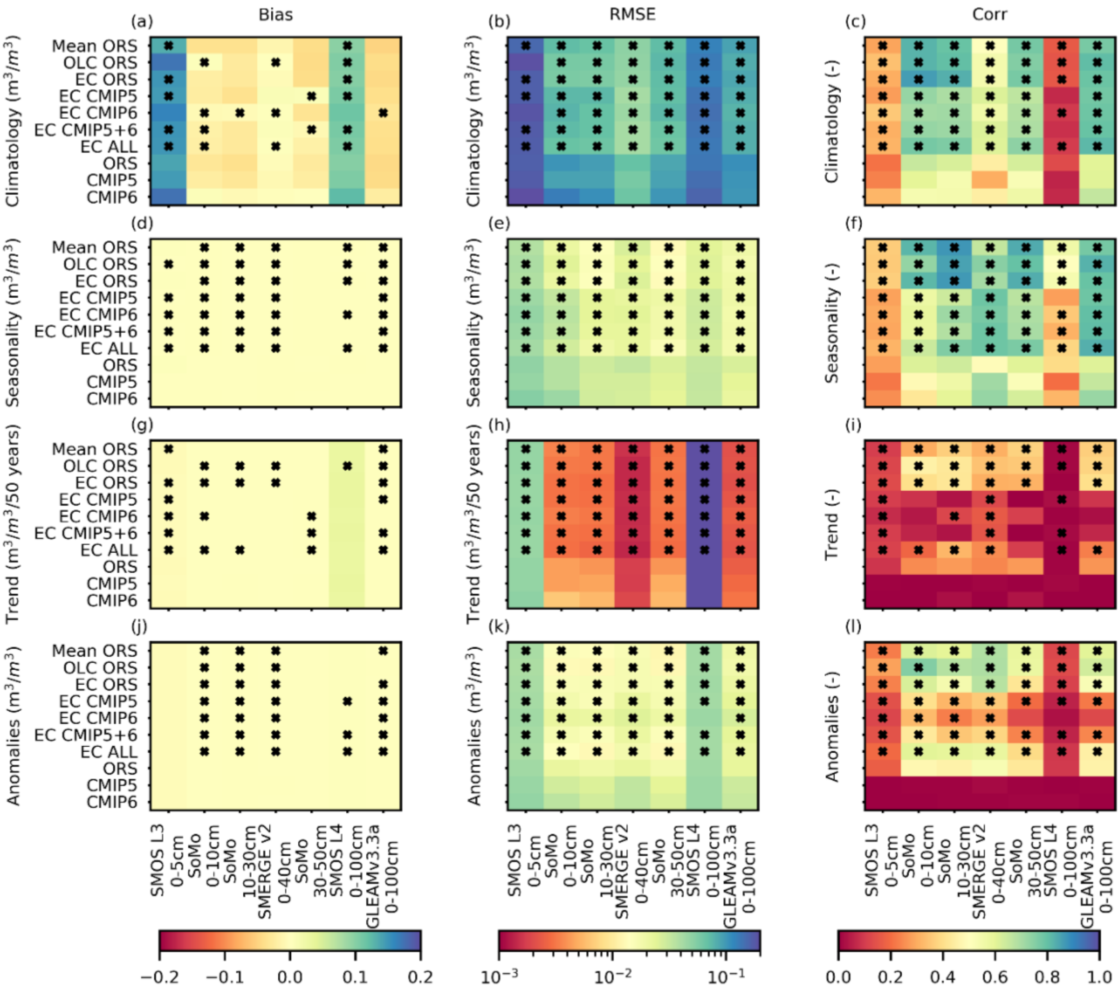

Figure 3: The Bias, RMSE, and Corr, calculated over space, between the climatology, seasonal cycle, linear trends, and anomalies of the individual merged products (Mean ORS through EC ALL) and source datasets (ORS, CMIP5, CMIP6), and the global and regional datasets for evaluation (SMOS L3, SoMo, SMERGE v2, SMOS L4, GLEAMv3.3a). Hatching indicates that the magnitude of the Bias, RMSE, or Corr of the merged product is better than the product's source datasets in the same column (for EC CMIP5+6, the metric was the average of CMIP5 and CMIP6; for ALL, the metric was the average of ORS, CMIP5, and CMIP6). The climatology, seasonal cycle, linear trends, and anomalies were calculated over the overlapping time periods and soil layer depths between the evaluated and evaluation datasets. The anomalies were calculated by removing the seasonal cycle and linear trends from the SM values. Vertical interpolation of the merged and source datasets to 0–40 and 0–100 cm were calculated linearly. When calculating the evaluation metrics of linear trends, the insignificant trends at *p* = 0.1 level were set to zero to prevent small random variability from influencing the results. The evaluation metrics of the source datasets were calculated for each individual source dataset and averaged over the three groups (ORS, CMIP5, and CMIP6).

3.3 Evaluation against selected drought events

Both the 0–10 cm (Figures 4 and A6) and the 0–100 cm (Figures S7 and S8) SM anomalies displayed the ability to capture the evolution of the US drought during 1985–1992 and the Australian millennium drought during 2002–2009. The spatial patterns were consistent with those indicated by the scPDSI, with Spearman correlations mainly falling in the range of 0.5 to 0.8. The

https://doi.org/10.5194/essd-2021-84
Preprint. Discussion started: 16 March 2021

CMIP5- and CMIP6-based merged products did not capture the negative anomalies in 2007–2008 in Australia, which were reproduced by the other merged products (Figures S6 and S8). The better performances of the ORS-based than the CMIP5- and CMIP6-based merged products were consistent with the evaluation results against in situ observations (Figures 2 and S5).

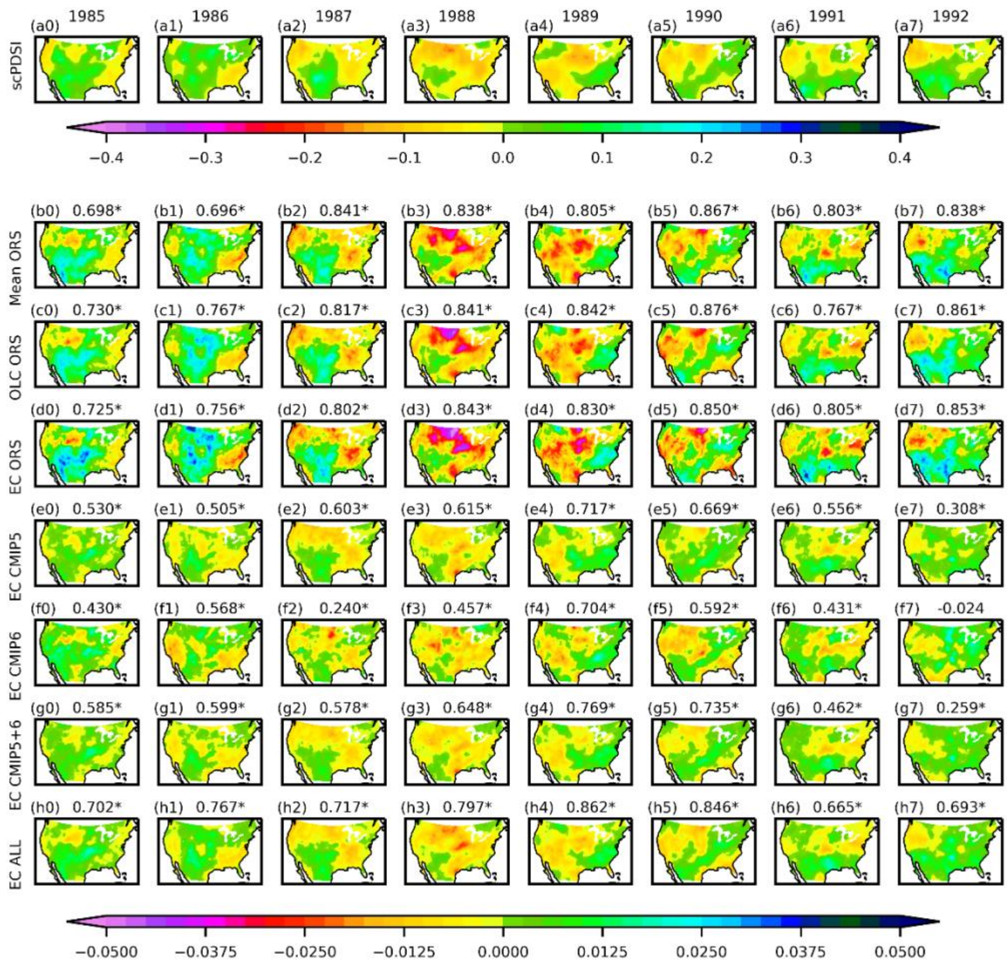

Figure 4: The annual mean scPDSI anomalies (no unit) and annual mean SM anomalies (m^3/m^3) during the US drought 1985–1992. The numbers on top of the plots are the Spearman correlation between the anomalies of the merged product and scPDSI, and the asterisk (*) indicates that the correlation is significant at $p = 0.05$. The anomalies were calculated relative to the climatology of 1970–2016. The SM anomalies are for 0–10 cm.

3.4 The spatial and temporal characteristics of the merged SM products

Averaged over the globe and large geographical regions, the time series of SM anomalies for different merged products corresponded well with one another, except in the high-latitude regions (SREX ID = 1, 2) and the arid Sahara region (SREX ID = 14) (Figures 5 and S9). The interannual variability of the 50–100 cm global mean time series of the EC CMIP6 product appeared too small compared with the other products (Figure 5). The good correspondence indicates that the temporal

https://doi.org/10.5194/essd-2021-84
Preprint. Discussion started: 16 March 2021

Open Access Earth System Science Data Discussions

concatenation involved in creating the OLC ORS and EC ALL products (Sect. 2.9) did not create any temporal discontinuities distinguishing them from the other products, and that the temperature- and precipitation-based constraint in the EC method induced natural internal variability similar to the real world.

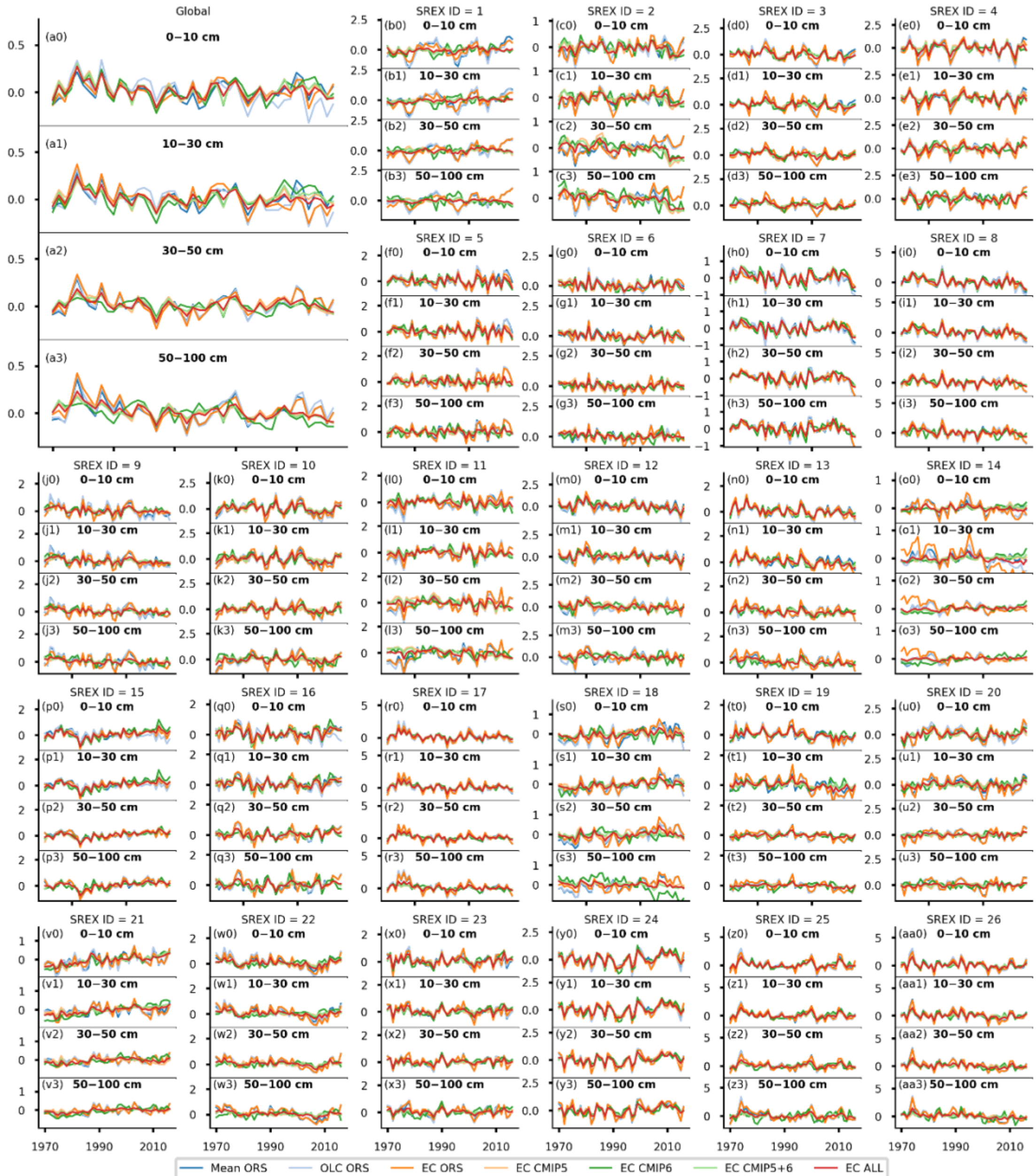

Figure 5: The time series of annual mean SM anomalies (10^{-2} m^3/m^3) relative to the 1981–2010 climatology averaged for the globe and the IPCC SREX regions (Field et al., 2012).

https://doi.org/10.5194/essd-2021-84
Preprint. Discussion started: 16 March 2021

The SM climatology showed reasonable spatial patterns in all the merged products, with the lowest values occurring in the arid regions of Sahara, western United States, central Asia, and interior Australia, and highest values occurring in the high latitudes and tropical rainforest regions (Figure S10). The OLC merging method caused an increase in absolute SM values, especially in the 0–10 and 10–30 cm soil layers, relative to unweighted averaging (Figure S10, 1st and 2nd rows). The EC method did not induce a similar increase (Figure S10, 1st and 3rd rows), which was expected because the procedure did not change the 1981–2010 climatology of the source datasets (Sect. 2.8). The Mean ORS, EC ORS, EC CMIP5+6, and EC ALL products showed little difference in SM climatology across the soil layers, but the OLC ORS and EC CMIP6 products showed decreased SM from the shallower to deeper soil layers (Figure S10). The EC CMIP5 had the highest SM values at the 30–50 cm soil layer (Figure S10).

The timings of annual maximum SM were mostly consistent across different merged products, with exceptions occurring in northeast Asia, eastern Canada, and Alaska (Figure S11). The maximum SM occurred around February in the southern subtropics, southern North America, and the Mediterranean; around September in the monsoonal regions of Africa and southern and eastern Asia; and around May in northern North America and most of Eurasia. At deeper soil layers (30–50 and 50–100 cm), the CMIP5- and CMIP6-based merged SM showed earlier occurrence of annual maximum SM (around June) than the other merged datasets (around September) in eastern Asia.

All the merged products showed increasing SM trends in the northern high-latitudes, central Eurasia, and northern Africa, and decreasing trends in eastern South America, southern Africa, and eastern Australia (Figure S12). The CMIP5- and CMIP6-based merged datasets showed greater drying in eastern North America and Europe than the ORS-based estimates, and less drying near the North China Plain than the ORS-based products. A major difference existed between the CMIP6-based merged products (EC CMIP6, EC CMIP5+6, EC ALL) and the other products in northeast Asia in the 50–100 cm soil layer, where the former displayed strong drying trends and the latter did not. The estimated uncertainty intervals of the merged products were slightly larger for the OLC method than the unweighted standard deviation of the source datasets, and both were considerably larger than the EC method (Figure S13). For all the methods, the uncertainty intervals were greater in the temperate regions than in the arid regions, which was consistent with the higher SM absolute values in the temperature regions.

3.5 Sensitivity to precipitation, air temperature, and surface downwelling shortwave radiation

The probability distributions of the significant partial correlations showed that the merged SM products were primarily positively correlated with precipitation, and negatively correlated with air temperature and surface downwelling shortwave radiation for either annual, June–August, or December–February averages (Figure 6). The partial correlations between SM and precipitation were significant in a high fraction of grid cells, and the fraction decreased toward deeper soil layers (Figure S14). A moderate number of grid cells had significant partial correlations between the SM and air temperature, and few grid cells had significant partial correlations between the SM and surface downwelling shortwave radiation (Figures S15 and S16). Again, the CMIP5- and CMIP6-based merged products showed some inconsistencies with the ORS-based products, with the former showing more negative and significant partial correlations with air temperature in the southern United States,

https://doi.org/10.5194/essd-2021-84
Preprint. Discussion started: 16 March 2021

Mediterranean, and southern hemisphere (Figure S15), less negative or significant partial correlations with shortwave radiation across the globe, and less positive or significant partial correlations with precipitation across the globe (Figures S16 and S14).

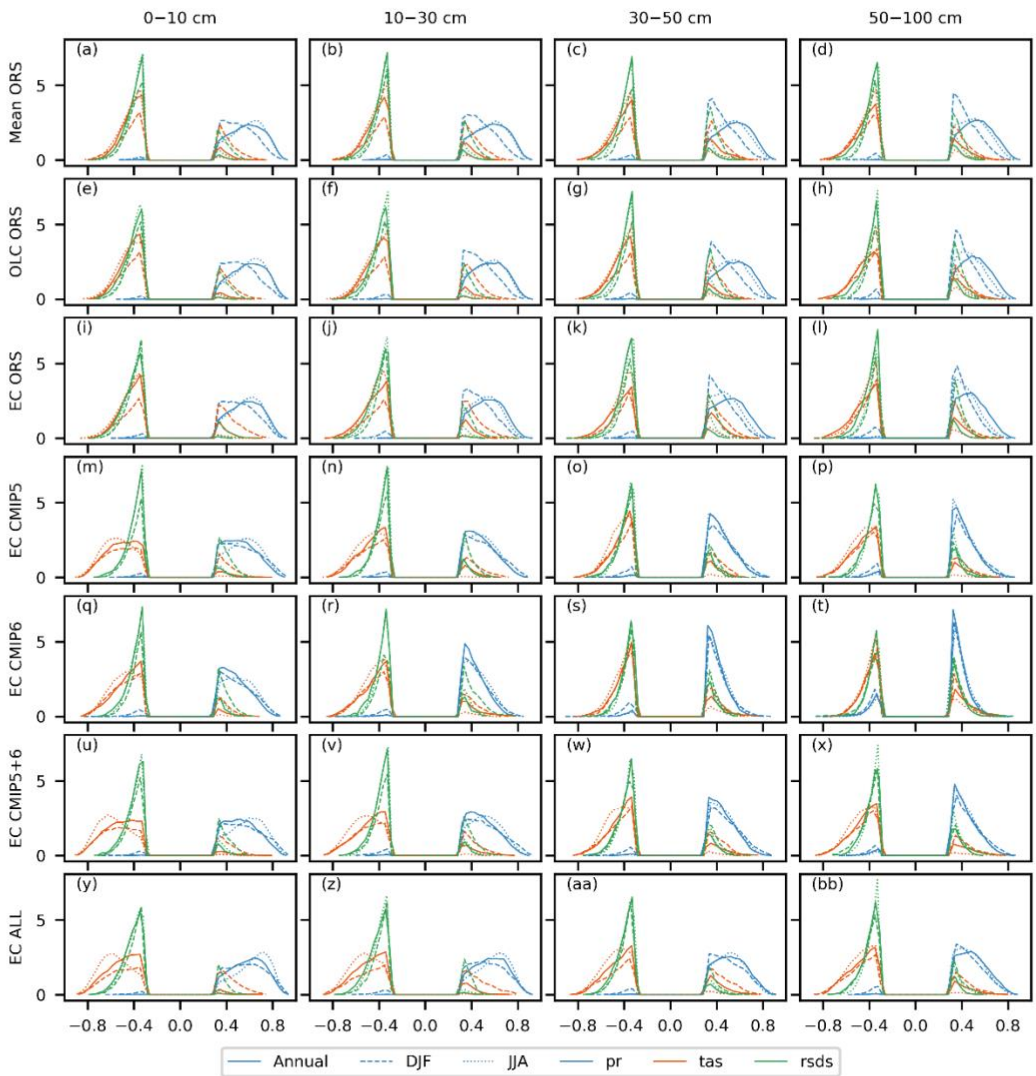

Figure 6: Probability distributions of the partial correlation coefficients (no unit) among annual, boreal winter (DJF), and boreal summer (JJA) mean SM in the merged products and observed precipitation (pr), air temperature (tas), and shortwave radiation (rsds). Only the significant partial correlation coefficients at the *p* = 0.05 level were used in calculating the probability distributions.

4 Discussion

Overall, the merged SM products showed better performances than their source datasets; they also showed the ability to capture large-scale drought events, as well as reasonable spatiotemporal patterns and climatic response characteristics. The ranges of

https://doi.org/10.5194/essd-2021-84
Preprint. Discussion started: 16 March 2021

performance metrics of the new datasets (Figures 2 and S5) were broadly within the estimates reported by previous SM evaluations, although making a strict comparison is difficult because of the widely different spatiotemporal coverages and resolutions (Beck et al., 2021; Karthikeyan et al., 2017; Li et al., 2020b; Wang et al., 2021a; Yuan and Quiring, 2017). These results (Figures 2 and S5) demonstrated that the merging procedures (unweighted averaging, OLC, EC) used were effective in creating relatively accurate long-term multi-layer SM data at the global scale.

Regarding the three merging methods, the OLC method only showed better performance than unweighted averaging over the urban and built-up lands, crop/natural vegetation mosaic, and barren land cover types (Figure S5), which may be a benefit rendered by the overrepresentation of these land cover types in the in situ observations (Figure S1). The ISMN stations are very sparse (Figure S1), and a previous study suggested that denser observations may lead to better-performing merged product (Gruber et al., 2018). In the future, data sources such as FLUXNET and local SM networks that are not included in the ISMN may be exploited to improve the OLC-ORS product. Future extension of the OLC method may aim to account for the spatial representativeness of individual stations (Molero et al., 2018), and to test alternative error-estimation methods such as the extended collocation (Gruber et al., 2016). The EC method showed similar performance to the unweighted averaging when applied onto the ORS source datasets, which may be because the meteorological forcings for these datasets were already realistic (Table S5). However, the effectiveness of the EC method was clear when applied onto the online CMIP5 and CMIP6 simulations. Despite such EC-based improvement, the ORS-based merged products tended to perform better than the EC CMIP5, CMIP6, CMIP5+6, and ALL products (Figures 2 and S5). The current EC procedure used simple linear regression form and only temperature and precipitation as constraint variables (Sect. 2.8). In addition to temperature and precipitation, the SM was influenced by other atmospheric and land conditions (e.g., wind, leaf area and stomata closure, snow cover and melt, groundwater). Therefore, future studies may achieve better EC outcomes by incorporating more influencing factors into the EC procedure, and by using nonlinear regression methods such as machine learning. Another drawback of the current EC method is the low uncertainty interval (Figure S13), which is likely an underestimation. Whereas the OLC method accounts for the difference between in situ and source datasets in estimating the uncertainty interval (Sect. 2.7), the EC method does not. Future studies should also aim to better incorporate this information into the estimation of EC-based uncertainty interval, and to better account for the structural uncertainty introduced by the regression form and limited range of predictors in the EC procedure.

The high performance variability of merged products across space (Figure S5) is consistent with previous studies (Beck et al., 2021; Karthikeyan et al., 2017; Li et al., 2020b; Wang et al., 2021a; Yuan and Quiring, 2017). The high RMSE of the merged products in the shallower soil layers across the water bodies and evergreen needleleaf forests (Figure S5e–h) were likely caused by the high positive Bias in these land cover types (Figure S5a–d) since the corresponding Corr values were relatively high (Figure S5i–l). The positive Bias over water bodies may be caused by inaccurate land-water classification at the resolution of the source datasets (Tables S1–S4). The positive and negative Bias over the forested land cover types in high-latitude and tropical regions (e.g., evergreen needleleaf forests, evergreen broadleaf forests, and deciduous needleleaf forests; Figure S2) may be due to biases in evapotranspiration and leaf area index in the source LSMs, reanalysis, and ESMs (Tables S2–S4), and further related to processes such as rooting depth and hydraulic redistribution (Pan et al., 2020; Wang et al., 2021b). Low Corr occurred over some land cover types in high latitudes, semi-arid to arid regions, and urban areas (e.g., open shrublands, urban and built-up lands, cropland/natural vegetation mosaics, and barren lands; Figure S2). In the high latitudes, the low Corr may be associated with inadequate frozen soil processes in the source LSMs, reanalysis, and ESMs (Andresen et al., 2020). In the semi-arid to arid regions, the low Corr may be due to random errors in SM observations and simulated values, which would be comparatively large for low SM values. In urban areas, the low performance may be caused by the radio frequency interference of satellite observations (Wang et al., 2012) and inadequacies in the representation of urban areas at the resolution of the source model products (Table S2–Table S4).

When evaluated against the global and regional gridded datasets that were not used in the merging, the merged products showed the highest RMSE in climatology and lowest Corr in spatial trends (Figure 3). Such results were likely because the climatology SM values had higher magnitudes than the seasonal and interannual anomalies or trends, and the historical SM trends were highly uncertain. The low similarity between the SMOS L3/L4 and the synthesized SM products may be caused by the short overlapping period (2010–2016, Table 1). A previous study also found systematic differences between the climatology of satellite-observed and simulated SM (Piles et al., 2019). The high similarity between the SoMo.ml and GLEAM v3.3a root zone SM datasets and the merged products may be because the former depend on the same ISMN stations, from which the OLC-ORS product was derived (Sect. 2.2), and the latter were from the same reanalysis as the GLEAM v3.3a surface dataset used in the merging (Table S2). In general, because these evaluation datasets are not ground truths like in situ observations, the identified differences in evaluation metrics should not be viewed as an absolute indicator of unreliability in the merged products. Similarly, the benchmarking against scPDSI (Sect. 3.3) only provided qualitative rather than quantitative indicators of performance for the merged products because scPDSI is essentially a different variable from SM.

The vertical gradient in Bias (Figure 2, Figure S5), the high uncertainty in the vertical gradient in the climatology of merged SM (Figure S10), and the divergent trends in the 50–100 cm SM in northeast Asia across the merged products (Figure S12) point to the need to reduce uncertainties in vertical distribution and dynamics of SM in the merged products. The high SM values for the HadGEM2-CC and HadGEM-ES datasets may be the reason why the highest SM occurred in the 30–50 cm layer in the EC CMIP5 product (Figure S17). All the source datasets for the EC CMIP6 SM showed negative trends in northeast Asia in the 50–100 cm soil layer, but this feature does not exist in the original ORS or CMIP5 datasets (results not shown). All the source datasets do not have consistent SM vertical gradients, with the maximum value falling at either the surface, deepest, or intermediate soil layers (Figure S17). Such vertical inconsistencies may be related to inconsistencies in the vertical discretization of the soil column (Tables S2–S4), soil properties in each layer, modeling of lateral flow and drainage, or other factors (e.g., Balsamo et al., 2009; Best et al., 2011; Melton et al., 2019). Previous regional or global SM evaluations (e.g., Beck et al., 2021; Karthikeyan et al., 2017; Li et al., 2020b; Wang et al., 2021a; Yuan and Quiring, 2017) rarely focused on the performance on vertical gradient, and such a limitation should be better addressed in future analyses and dataset development.

The temporal continuity of the merged products was only examined visually in this study (Figure 5). Although the applied concatenation procedure (Sect. 2.9) was previously demonstrated to lead to good temporal continuity (Su et al., 2016), the merged products should be subject to more rigorous statistical testing in the future. The SM seasonality in the merged products (Figure S11) was broadly consistent with previously reported timing of annual maximum precipitation (Knoben et al., 2019). Differences at the deeper soil layers between the CMIP5- and CMIP6-based merged products and the ORS-based products may be partially caused by the lack of consideration of lagged SM response to meteorological drivers, especially at the deeper layers, in the EC method (Sect. 2.8). Uncertainty at the deeper layers would also be high because fewer source datasets were available than for the shallower layers (Tables S1–S4). The SM trends (Figure S12) were broadly consistent with previous reports on historical changes in agricultural droughts (Dai and Zhao, 2017; Liu et al., 2019; Lu et al., 2019).

The primarily positive partial correlations between the SM and precipitation, and the primarily negative partial correlations between the SM and air temperature or shortwave radiation, were consistent with expectations from physical processes. The existence of significant positive partial correlations between air temperature and SM might be caused by less precipitation falling as snow at higher temperatures. The weaker relationships between precipitation and SM in the CMIP5- and CMIP6-based merged products than these in the ORS-based products (Figure S14) may be because the EC method did not fully explain the temporal mismatch between the source datasets and the real world because the relationships were insignificant in various grids and time steps (Figure S4). The stronger relationships between air temperature and SM of the CMIP5- and CMIP6-based merged products than these of the ORS-based products (Figure S15) may be partially caused by compensation for the weaker relationships between shortwave radiation and SM (Figure S16). Because the temperature and shortwave radiation tend to be highly correlated, shortwave radiation was not considered as a predictor for the EC method in the present research.

5 Data availability

The seven SM products, including the estimated SM values and uncertainty intervals, are available from https://doi.org/10.6084/m9.figshare.13661312.v1 (Wang and Mao, 2021). The files are in NetCDF4 format.

6 Code availability

The source codes for developing all the SM datasets are available at https://ywang11@bitbucket.org/ywang11/soil_moisture_merge.git

7 Conclusions

This study achieved the goal of creating long-term, gap-free, multi-layer SM products (1970–2016, 0.5°, monthly, 0–10, 10–30, 30–50, and 50–100 cm) that displayed realistic temporal evolutions and spatial patterns and outperformed the source SM datasets in the systematic evaluations against independent in situ measurements and semi-independent gridded SM estimates. Three new SM products (mean ORS, OLC ORS, and EC ORS) developed from the satellite observations, reanalysis, and offline LSMs were shown to perform better than those based on the ESMs. Therefore, they are recommended for future applications, such as the detection and attribution of historical changes of SM and associated extreme events, providing the initial and boundary conditions for atmospheric models, benchmarking various types of models, and managing drought and flood risks. By comparing three different merging methods (unweighted averaging, OLC, and EC), this study further denoted that the OLC method may require more in situ observations to exceed the unweighted averaging, and that linear regression-based EC with a limited range of un-lagged predictors was inadequate in correcting all the ESM errors. Future SM developments may aim to assemble more in situ SM datasets and to implement other advanced fusion algorithms (e.g., extended collocation, machine learning).

8 Author contributions

J.M. conceived the research; Y.W. and J.M. performed the analyses and drafted the figures; Y.W. and J.M. wrote the first draft of the manuscript; M.J., F.M.H., X.S., S.D.W., and Y.D. reviewed and edited the manuscript before submission. All authors made substantial contributions to the discussion of content.

9 Competing interests

The authors declare that they have no conflict of interest.

Acknowledgements

This research was supported by an Oak Ridge National Laboratory (ORNL) subcontract (4000169153), and the Reducing Uncertainties in Biogeochemical Interactions through Synthesis and Computation Science Focus Area (RUBISCO SFA) project in the Earth and Environmental Systems Sciences Division (EESSD) of the Biological and Environmental Research (BER) office in the US Department of Energy (DOE) Office of Science. ORNL is managed by UT-BATTELLE, LLC, for DOE under Contract No. DE-AC05-00OR22725. The Dai self-calibrated Palmer Drought Severity Index data was provided by the NOAA/OAR/ESRL PSL, Boulder, Colorado, USA, from their website at https://psl.noaa.gov/. The GSWP3 meteorological data was downloaded as the prepared historical input datasets for ISIMIP3a. For their roles in producing, coordinating, and making available the ISIMIP input data and impact model output, we acknowledge the modeling groups, the ISIMIP sector coordinators and the ISIMIP cross-sectoral science team.

References

Al Bitar, A., Kerr, Y. H., Merlin, O., Cabot, F., Wigneron, J.-P., Froissard, F., Suquia, D., Petitpa, A., Piguet, B., and Martin, E.: Global drought index from SMOS soil moisture, in: IEEE International Geoscience and Remote Sensing Symposium, IGARSS, Melbourne, Australia, n.p., 2013.

Al Bitar, A., Mialon, A., Kerr, Y. H., Cabot, F., Richaume, P., Jacquette, E., Quesney, A., Mahmoodi, A., Tarot, S., Parrens, M., Al-Yaari, A., Pellarin, T., Rodriguez-Fernandez, N., and Wigneron, J.-P.: The global SMOS Level 3 daily soil moisture and brightness temperature maps, Earth Syst. Sci. Data, 9, 293–315, https://doi.org/10.5194/essd-9-293-2017, 2017.

Andresen, C. G., Lawrence, D. M., Wilson, C. J., McGuire, A. D., Koven, C., Schaefer, K., Jafarov, E., Peng, S., Chen, X., Gouttevin, I., Burke, E., Chadburn, S., Ji, D., Chen, G., Hayes, D., and Zhang, W.: Soil moisture and hydrology projections of the permafrost region – a model intercomparison, The Cryosphere, 14, 445–459, https://doi.org/10.5194/tc-14-445-2020, 2020.

Balsamo, G., Beljaars, A., Scipal, K., Viterbo, P., van den Hurk, B., Hirschi, M., and Betts, A. K.: A revised hydrology for the ECMWF model: Verification from field site to terrestrial water storage and impact in the integrated forecast system, J. Hydrometeorol., 10, 623–643, https://doi.org/10.1175/2008JHM1068.1, 2009.

Balsamo, G., Albergel, C., Beljaars, A., Boussetta, S., Brun, E., Cloke, H., Dee, D., Dutra, E., Muñoz-Sabater, J., Pappenberger, F., de Rosnay, P., Stockdale, T., and Vitart, F.: ERA-Interim/Land: a global land surface reanalysis data set, Hydrol. Earth Syst. Sci., 19, 389–407, https://doi.org/10.5194/hess-19-389-2015, 2015.

Beaudoing, H., Rodell, M., and NASA/GSFC/HSL: GLDAS Noah Land Surface Model L4 3 hourly 0.25 x 0.25 degree V2.0, Goddard Earth Sciences Data and Information Services Center (GES DISC), Greenbelt, MD, USA, 2019.

Beck, H. E., Pan, M., Miralles, D. G., Reichle, R. H., Dorigo, W. A., Hahn, S., Sheffield, J., Karthikeyan, L., Balsamo, G., Parinussa, R. M., Du, J., Kimball, J. S., Vergopolan, N., and Wood, E. F.: Evaluation of 18 satellite- and model-based soil moisture products using in situ measurements from 826 sensors, Hydrol. Earth Syst. Sci., 25, 17–40, https://doi.org/10.5194/hess-25-17-2021, 2021.

Best, M. J., Pryor, M., Clark, D. B., Rooney, G. G., Essery, R. . L. H., Ménard, C. B., Edwards, J. M., Hendry, M. A., Porson, A., Gedney, N., Mercado, L. M., Sitch, S., Blyth, E., Boucher, O., Cox, P. M., Grimmond, C. S. B., and Harding, R. J.: The Joint UK Land Environment Simulator (JULES), model description – Part 1: Energy and water fluxes, Geosci. Model Dev., 4, 677–699, https://doi.org/10.5194/gmd-4-677-2011, 2011.

Bishop, C. H. and Abramowitz, G.: Climate model dependence and the replicate Earth paradigm, Clim. Dyn., 41, 885–900, https://doi.org/10.1007/s00382-012-1610-y, 2013.

Dai, A. and Zhao, T.: Uncertainties in historical changes and future projections of drought. Part I: estimates of historical drought changes, Climatic Change, 144, 519–533, https://doi.org/10.1007/s10584-016-1705-2, 2017.

Dai, A., Trenberth, K. E., and Qian, T.: A global dataset of Palmer Drought Severity Index for 1870–2002: Relationship with soil moisture and effects of surface warming, J. Hydrometeorol., 5, 1117–1130, https://doi.org/10.1175/jhm-386.1, 2004.

https://doi.org/10.5194/essd-2021-84
Preprint. Discussion started: 16 March 2021

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, Q.J.R. Meteorol. Soc., 137, 553–597, https://doi.org/10.1002/qj.828, 2011.

Dirmeyer, P. A., Gao, X., Zhao, M., Guo, Z., Oki, T., and Hanasaki, N.: GSWP-2: Multimodel analysis and implications for our perception of the land surface, B. Am. Meteorol. Soc., 87, 1381–1398, https://doi.org/10.1175/bams-87-10-1381, 2006.

Dorigo, W., de Jeu, R., Chung, D., Parinussa, R., Liu, Y., Wagner, W., and Fernández-Prieto, D.: Evaluating global trends (1988-2010) in harmonized multi-satellite surface soil moisture, Geophys. Res. Lett., 39, https://doi.org/10.1029/2012GL052988, 2012.

Dorigo, W., Wagner, W., Albergel, C., Albrecht, F., Balsamo, G., Brocca, L., Chung, D., Ertl, M., Forkel, M., Gruber, A., Haas, E., Hamer, P. D., Hirschi, M., Ikonen, J., de Jeu, R., Kidd, R., Lahoz, W., Liu, Y. Y., Miralles, D., Mistelbauer, T., Nicolai-Shaw, N., Parinussa, R., Pratola, C., Reimer, C., van der Schalie, R., Seneviratne, S. I., Smolander, T., and Lecomte, P.: ESA CCI soil moisture for improved Earth system understanding: State-of-the art and future directions, Remote Sens. Environ., 203, 185–215, https://doi.org/10.1016/j.rse.2017.07.001, 2017.

Dorigo, W. A., Wagner, W., Hohensinn, R., Hahn, S., Paulik, C., Xaver, A., Gruber, A., Drusch, M., Mecklenburg, S., van Oevelen, P., Robock, A., and Jackson, T.: The International Soil Moisture Network: a data hosting facility for global in situ soil moisture measurements, Hydrol. Earth Syst. Sci., 15, 1675–1698, https://doi.org/10.5194/hess-15-1675-2011, 2011.

Dorigo, W. A., Xaver, A., Vreugdenhil, M., Gruber, A., Hegyiová, A., Sanchis-Dufau, A. D., Zamojski, D., Cordes, C., Wagner, W., and Drusch, M.: Global automated quality control of in situ soil moisture data from the International Soil Moisture Network, Vadose Zone J., 12, 1-21 vzj2012.0097, https://doi.org/10.2136/vzj2012.0097, 2013.

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., and Taylor, K. E.: Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization, Geosci. Model Dev., 9, 1937–1958, https://doi.org/10.5194/gmd-9-1937-2016, 2016.

Field, C. B., Barros, V., Stocker, T. F., and Dahe, Q. (Eds.): Managing the risks of extreme events and disasters to advance climate change adaptation: Special report of the intergovernmental panel on climate change, Cambridge University Press, Cambridge, https://doi.org/10.1017/CBO9781139177245, 2012.

Friedl, M. and Sulla-Menashe, D.: MCD12C1 MODIS/Terra+Aqua land cover type L3 global 0.05Deg CMG V006 [Data set], https://doi.org/10.5067/MODIS/MCD12C1.006, 2015.

Giorgi, F. and Mearns, L. O.: Calculation of average, uncertainty range, and reliability of regional climate changes from AOGCM simulations via the “reliability ensemble averaging” (REA) method, J. Climate, 15, 1141–1158, https://doi.org/10.1175/1520-0442(2002)015<1141:COAURA>2.0.CO;2, 2002.

Green, J. K., Seneviratne, S. I., Berg, A. M., Findell, K. L., Hagemann, S., Lawrence, D. M., and Gentine, P.: Large influence of soil moisture on long-term terrestrial carbon uptake, Nature, 565, 476–479, https://doi.org/10.1038/s41586-018-0848-x, 2019.

Gruber, A., Su, C.-H., Crow, W. T., Zwieback, S., Dorigo, W. A., and Wagner, W.: Estimating error cross-correlations in soil moisture data sets using extended collocation analysis, J. Geophys. Res. Atmos., 121, 1208–1219, https://doi.org/10.1002/2015JD024027, 2016.

Gruber, A., Crow, W. T., and Dorigo, W. A.: Assimilation of spatially sparse in situ soil moisture networks into a continuous model domain, Water Resour. Res., 54, 1353–1367, https://doi.org/10.1002/2017WR021277, 2018.

Gu, X., Li, J., Chen, Y. D., Kong, D., and Liu, J.: Consistency and discrepancy of global surface soil moisture changes from multiple model-based data sets against satellite observations, J. Geophys. Res. Atmos., 124, 1474–1495, https://doi.org/10.1029/2018JD029304, 2019.

Harris, I., Jones, P. D., Osborn, T. J., and Lister, D. H.: Updated high-resolution grids of monthly climatic observations – the CRU TS3.10 dataset, Int. J. Climatol., 34, 623–642, https://doi.org/10.1002/joc.3711, 2014.

Hauser, M.: Regionmask: plotting and creation of masks of spatial regions, Zenodo, https://doi.org/10.5281/zenodo.3585543, 2019.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J.: The ERA5 global reanalysis, Q.J.R. Meteorol. Soc., 146, 1999–2049, https://doi.org/10.1002/qj.3803, 2020.

Hobeichi, S., Abramowitz, G., Evans, J., and Ukkola, A.: Derived Optimal Linear Combination Evapotranspiration (DOLCE): a global gridded synthesis ET estimate, Hydrol. Earth Syst. Sci., 22, 1317–1336, https://doi.org/10.5194/hess-22-1317-2018, 2018.

Hobeichi, S., Abramowitz, G., Evans, J., and Beck, H. E.: Linear Optimal Runoff Aggregate (LORA): a global gridded synthesis runoff product, Hydrol. Earth Syst. Sci., 23, 851–870, https://doi.org/10.5194/hess-23-851-2019, 2019.

Huntzinger, D. N., Schwalm, C. R., Wei, Y., Cook, R. B., Michalak, a M., Schaefer, K., Jacobson, a R., Arain, M. a, Ciais, P., Fisher, J. B., Hayes, D. J., Huang, M., Huang, S., Ito, A., Jain, a K., Lei, H., Lu, C., Maignan, F., Mao, J., Parazoo, N. C., Peng, C., Peng, S., Poulter, B., Ricciuto, D. M., Tian, H., Shi, X., Wang, W., Zeng, N., Zhao, F., Zhu, Q., Yang, J., and Tao, B.: NACP MsTMIP: Global 0.5-degree model outputs in standard format, Version 1.0, ORNL Distributed Active Archive Center, https://doi.org/10.3334/ornldaac/1225, 2018.

Karthikeyan, L., Pan, M., Wanders, N., Kumar, D. N., and Wood, E. F.: Four decades of microwave satellite soil moisture observations: Part 2. Product validation and inter-satellite comparisons, Advances in Water Resources, 109, 236–252, https://doi.org/10.1016/j.advwatres.2017.09.010, 2017.

Knoben, W. J. M., Woods, R. A., and Freer, J. E.: Global bimodal precipitation seasonality: A systematic overview, Int. J. Climatol., 39, 558–567, https://doi.org/10.1002/joc.5786, 2019.

Kobayashi, S., Ota, Y., Harada, Y., Ebita, A., Moriya, M., Onoda, H., Onogi, K., Kamahori, H., Kobayashi, C., Endo, H., Miyaoka, K., and Takahashi, K.: The JRA-55 Reanalysis: General specifications and basic characteristics, J. Meteor. Soc. Japan, 93, 5–48, https://doi.org/10.2151/jmsj.2015-001, 2015.

Kumar, S., Newman, M., Wang, Y., and Livneh, B.: Potential reemergence of seasonal soil moisture anomalies in North America, J. Climate, 32, 2707–2734, https://doi.org/10.1175/JCLI-D-18-0540.1, 2019.

Laloyaux, P., de Boisseson, E., and Dahlgren, P.: CERA-20C: An Earth system approach to climate reanalysis, ECMWF Newsl., 150, 25–30, https://doi.org/10.21957/FFS36BIRJ2, 2017.

Li, L., Shangguan, W., Deng, Y., Mao, J., Pan, J., Wei, N., Yuan, H., Zhang, S., Zhang, Y., and Dai, Y.: A causal inference model based on random forests to identify the effect of soil moisture on precipitation, J. Hydrometeorol., 21, 1115–1131, https://doi.org/10.1175/JHM-D-19-0209.1, 2020a.

Li, M., Wu, P., and Ma, Z.: A comprehensive evaluation of soil moisture and soil temperature from third-generation atmospheric and land reanalysis data sets, Int. J. Climatol., 40, 5744–5766, https://doi.org/10.1002/joc.6549, 2020b.

Liu, Y., Liu, Y., and Wang, W.: Inter-comparison of satellite-retrieved and Global Land Data Assimilation System-simulated soil moisture datasets for global drought analysis, Remote Sens. Environ., 220, 1–18, https://doi.org/10.1016/j.rse.2018.10.026, 2019.

Liu, Y. Y., Parinussa, R. M., Dorigo, W. A., De Jeu, R. A. M., Wagner, W., van Dijk, A. I. J. M., McCabe, M. F., and Evans, J. P.: Developing an improved soil moisture dataset by blending passive and active microwave satellite-based retrievals, Hydrol. Earth Syst. Sci., 15, 425–436, https://doi.org/10.5194/hess-15-425-2011, 2011.

Liu, Y. Y., Dorigo, W. A., Parinussa, R. M., de Jeu, R. A. M., Wagner, W., McCabe, M. F., Evans, J. P., and van Dijk, A. I. J. M.: Trend-preserving blending of passive and active microwave soil moisture retrievals, Remote Sens. Environ., 123, 280–297, https://doi.org/10.1016/j.rse.2012.03.014, 2012.

Llamas, R. M., Guevara, M., Rorabaugh, D., Taufer, M., and Vargas, R.: Spatial gap-filling of ESA-CCI satellite-derived soil moisture based on geostatistical techniques and multiple regression, Remote Sens., 12, 665, https://doi.org/10.3390/rs12040665, 2020.

Loew, A., Stacke, T., Dorigo, W., de Jeu, R., and Hagemann, S.: Potential and limitations of multidecadal satellite soil moisture observations for selected climate model evaluation studies, Hydrol. Earth Syst. Sci., 17, 3523–3542, https://doi.org/10.5194/hess-17-3523-2013, 2013.

Lu, J., Carbone, G. J., and Grego, J. M.: Uncertainty and hotspots in 21st century projections of agricultural drought from CMIP5 models, Sci. Rep., 9, 4922, https://doi.org/10.1038/s41598-019-41196-z, 2019.

Martens, B., Miralles, D. G., Lievens, H., van der Schalie, R., de Jeu, R. A. M., Fernández-Prieto, D., Beck, H. E., Dorigo, W. A., and Verhoest, N. E. C.: GLEAM v3: satellite-based land evaporation and root-zone soil moisture, Geosci. Model Dev., 10, 1903–1925, https://doi.org/10.5194/gmd-10-1903-2017, 2017.

McCarty, W., Coy, L., Gelaro, R., Huang, A., Merkova, D., Smith, E. B., Sienkiewicz, M., and Wargan, K.: MERRA-2 input observations: summary and assessment, Goddard Space Flight Center, Greenbelt, MD, USA, 2016.

Melton, J. R., Verseghy, D. L., Sospedra-Alfonso, R., and Gruber, S.: Improving permafrost physics in the coupled Canadian Land Surface Scheme (v.3.6.2) and Canadian Terrestrial Ecosystem Model (v.2.1) (CLASS-CTEM), Geosci. Model Dev., 12, 4443–4467, https://doi.org/10.5194/gmd-12-4443-2019, 2019.

Molero, B., Leroux, D. J., Richaume, P., Kerr, Y. H., Merlin, O., Cosh, M. H., and Bindlish, R.: Multi-timescale analysis of the spatial representativeness of in situ soil moisture data within satellite footprints: Soil moisture time and spatial scales, J. Geophys. Res. Atmos., 123, 3–21, https://doi.org/10.1002/2017JD027478, 2018.

Mystakidis, S., Davin, E. L., Gruber, N., and Seneviratne, S. I.: Constraining future terrestrial carbon cycle projections using observation-based water and carbon flux estimates, Glob. Change Biol., 22, 2198–2215, https://doi.org/10.1111/gcb.13217, 2016.

O, S. and Orth, R.: Global soil moisture from in-situ measurements using machine learning -- SoMo.ml, arXiv:2010.02374 [physics], 2020.

Padrón, R. S., Gudmundsson, L., and Seneviratne, S. I.: Observational constraints reduce likelihood of extreme changes in multidecadal land water availability, Geophys. Res. Lett., 46, 736–744, https://doi.org/10.1029/2018GL080521, 2019.

Pan, S., Pan, N., Tian, H., Friedlingstein, P., Sitch, S., Shi, H., Arora, V. K., Haverd, V., Jain, A. K., Kato, E., Lienert, S., Lombardozzi, D., Nabel, J. E. M. S., Ottlé, C., Poulter, B., Zaehle, S., and Running, S. W.: Evaluation of global terrestrial evapotranspiration using state-of-the-art approaches in remote sensing, machine learning and land surface modeling, Hydrol. Earth Syst. Sci., 24, 1485–1509, https://doi.org/10.5194/hess-24-1485-2020, 2020.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E.: Scikit-learn: machine learning in Python, J. Mach. Learn. Res., 12, 2825–2830, 2011.

Piles, M., Ballabrera-Poy, J., and Muñoz-Sabater, J.: Dominant features of global surface soil moisture variability observed by the SMOS satellite, Remote Sens., 11, 95, https://doi.org/10.3390/rs11010095, 2019.

Poli, P., Hersbach, H., Dee, D. P., Berrisford, P., Simmons, A. J., Vitart, F., Laloyaux, P., Tan, D. G. H., Peubey, C., Thépaut, J.-N., Trémolet, Y., Hólm, E. V., Bonavita, M., Isaksen, L., and Fisher, M.: ERA-20C: An atmospheric reanalysis of the twentieth century, J. Climate, 29, 4083–4097, https://doi.org/10.1175/JCLI-D-15-0556.1, 2016.

de Rosnay, P., Drusch, M., Vasiljevic, D., Balsamo, G., Albergel, C., and Isaksen, L.: A simplified Extended Kalman Filter for the global operational soil moisture analysis at ECMWF, Q.J.R. Meteorol. Soc., 139, 1199–1213, https://doi.org/10.1002/qj.2023, 2013.

Rui, H. and Beaudoing, H.: README document for NASA GLDAS version 2 data products, NASA Goddard Space Flight Center (GES DISC), Greenbelt, MD, USA, 2020.

Seabold, S. and Perktold, J.: Statsmodels: Econometric and statistical modeling with Python, in: 9th Python in Science Conference, https://doi.org/10.25080/majora-92bf1922-011, 2010.

Sitch, S., Huntingford, C., Gedney, N., Levy, P. E., Lomas, M., Piao, S. L., Betts, R., Ciais, P., Cox, P., Friedlingstein, P., Jones, C. D., Prentice, I. C., and Woodward, F. I.: Evaluation of the terrestrial carbon cycle, future plant geography and climate-carbon cycle feedbacks using five Dynamic Global Vegetation Models (DGVMs), Glob. Change Biol., 14, 2015–2039, https://doi.org/10.1111/j.1365-2486.2008.01626.x, 2008.

Sospedra-Alfonso, R. and Merryfield, W. J.: Initialization and potential predictability of soil moisture in the Canadian Seasonal to Interannual Prediction System, J. Climate, 31, 5205–5224, https://doi.org/10.1175/JCLI-D-17-0707.1, 2018.

Spinoni, J., Barbosa, P., De Jager, A., McCormick, N., Naumann, G., Vogt, J. V., Magni, D., Masante, D., and Mazzeschi, M.: A new global database of meteorological drought events from 1951 to 2016, J. Hydrol. Reg. Stud., 22, 100593, https://doi.org/10.1016/j.ejrh.2019.100593, 2019.

Su, C.-H., Ryu, D., Dorigo, W., Zwieback, S., Gruber, A., Albergel, C., Reichle, R. H., and Wagner, W.: Homogeneity of a global multisatellite soil moisture climate data record, Geophys. Res. Lett., 43, 11245–11252, https://doi.org/10.1002/2016GL070458, 2016.

Taylor, K. E., Stouffer, R. J., and Meehl, G. A.: An overview of CMIP5 and the experiment design, B. Am. Meteorol. Soc., 93, 485–498, https://doi.org/10.1175/BAMS-D-11-00094.1, 2012.

https://doi.org/10.5194/essd-2021-84
Preprint. Discussion started: 16 March 2021

Earth System Science Data
Open Access Discussions

Tobin, K. J., Torres, R., Crow, W. T., and Bennett, M. E.: Multi-decadal analysis of root-zone soil moisture applying the exponential filter across CONUS, Hydrol. Earth Syst. Sci., 21, 4403–4417, https://doi.org/10.5194/hess-21-4403-2017, 2017.

UCAR/NCAR/CISL/TDD: The NCAR Comannd Language (Version 6.6.2) [Software], Boulder, CO, USA, https://doi.org/10.5065/D6WD3XH5, 2019.

Wang, G., Garcia, D., Liu, Y., de Jeu, R., and Johannes Dolman, A.: A three-dimensional gap filling method for large geophysical datasets: Application to global satellite soil moisture observations, Environ. Model. Softw., 30, 139–142, https://doi.org/10.1016/j.envsoft.2011.10.015, 2012.

Wang, Y. and Mao, J.: Global multi-layer soil moisture products, https://doi.org/10.6084/m9.figshare.13661312.v1, 2021.

Wang, Y., Leng, P., Peng, J., Marzahn, P., and Ludwig, R.: Global assessments of two blended microwave soil moisture products CCI and SMOPS with in-situ measurements and reanalysis data, 94, 102234, https://doi.org/10.1016/j.jag.2020.102234, 2021a.

Wang, Z., Zhan, C., Ning, L., and Guo, H.: Evaluation of global terrestrial evapotranspiration in CMIP6 models, Theor. Appl. Climatol., 143, 521–531, https://doi.org/10.1007/s00704-020-03437-4, 2021b.

Wu, Z., Zhou, J., He, H., Lin, Q., Wu, X., and Xu, Z.: An advanced error correction methodology for merging in-situ observed and model-based soil moisture, J. Hydrol., 566, 150–163, https://doi.org/10.1016/j.jhydrol.2018.09.018, 2018.

Yuan, S. and Quiring, S. M.: Evaluation of soil moisture in CMIP5 simulations over the contiguous United States using in situ and satellite observations, Hydrol. Earth Syst. Sci., 21, 2203–2218, https://doi.org/10.5194/hess-21-2203-2017, 2017.

Zeng, Y., Su, Z., van der Velde, R., Wang, L., Xu, K., Wang, X., and Wen, J.: Blending satellite observed, model simulated, and in situ measured soil moisture over Tibetan plateau, Remote Sens., 8, 268, https://doi.org/10.3390/rs8030268, 2016.

Zhang, G. and Cook, K. H.: West African monsoon demise: Climatology, interannual variations, and relationship to seasonal rainfall, J. Geophys. Res. Atmos., 119, 10,175-10,193, https://doi.org/10.1002/2014JD022043, 2014.