Uncertainty estimation for a new exponential filter-based long-term root-zone soil moisture dataset from C3S surface observations

Adam Pasik[1], Alexander Gruber[1], Wolfgang Preimesberger[1], Domenico De Santis[2], and Wouter Dorigo[1]
[1]Department of Geodesy and Geoinformation, TU Wien, Wiedner Hauptstraße 8, 1040 Vienna, Austria
[2]Research Institute for Geo-Hydrological Protection, National Research Council, Via Cavour 4-6, 87036 Rende (CS), Italy
Correspondence: Adam Pasik (adam.pasik@geo.tuwien.ac.at)

Abstract. Soil moisture is a key variable in monitoring climate and an important component of the hydrological, carbon, and energy cycles. Satellite products ameliorate the sparsity of field measurements but are inherently limited to observing the near-surface layer, while water available in the unobserved root zone controls critical processes like plant water uptake and evapotranspiration. A variety of approaches exists for modelling root-zone soil moisture (RZSM), including approximating it from surface layer observations. While the number of available RZSM datasets is growing, they usually do not contain estimates of their uncertainty. In this paper we derive a long-term RZSM dataset (2002–2020) from the Copernicus Climate Change Service (C3S) surface soil moisture (SSM) COMBINED product via the exponential filter (EF) method. We identify the optimal value of the method's model parameter T, which controls the level of smoothing and delaying applied to the surface observations, by maximizing the correlation of RZSM estimates with field measurements from the International Soil Moisture Network (ISMN). Optimized T-parameter values were calculated for four soil depth layers (0–10 cm, 10–40 cm, 40–100 cm, and 100–200 cm) and used to calculate a global RZSM dataset. The quality of this dataset is then globally evaluated against RZSM estimates of the ERA5-Land reanalysis. Results of the product comparison show satisfactory skill in all four layers with median Pearson correlation ranging from 0.54 in the topmost to 0.28 in the deepest soil layer. Temporally-dynamic product uncertainties for each of the RZSM product layers are estimated by applying standard uncertainty propagation to SSM input data and by estimating structural uncertainties of the EF method from ISMN ground reference measurements taken at the surface and in varying depths. Uncertainty estimates were found to exhibit both realistic absolute magnitudes as well as temporal variations. The product described here is, to our best knowledge, the first global, long-term, uncertainty-characterized, and purely observation-based product for RZSM estimates up to 2 m depth.

1 Introduction

Soil moisture (SM) is an essential climate variable (ECV) crucial for understanding and modelling the Earth's climate, and an important control of hydrological, energy, and carbon fluxes (GCOS, 2022; Dorigo et al., 2021a). Global monitoring of SM is necessary for a variety of applications such as meteorological modelling (Albergel et al., 2008), monitoring drought (Tobin et al., 2017), and modelling groundwater recharge (Bouaziz et al., 2020), runoff, and catchment response to storms (Brocca et al., 2010).

In situ SM measurements are considered to provide the most accurate SM data but can differ greatly in measuring equipment and usually lack estimates of their uncertainties (Dorigo et al., 2011). Widely distributed SM field measurements are available from centralized platforms such as the International Soil Moisture Network (ISMN) (Dorigo et al., 2021b). While being essential for satellite and model product calibration and validation, in situ measurements lack the spatial coverage necessary for large-scale applications, especially in the Global South (see Figure A1; Dorigo et al. (2021a); Mishra et al. (2020)). Quasi-global SM information is available from modelled and satellite products, but their spatial resolution is very coarse (usually tens to hundreds of square kilometers) and usually insufficient to resolve the significant spatio-temporal heterogeneity of SM, which poses challenges to large-scale monitoring (Brocca et al., 2010). Global land surface model products provide gap-free and long-term SM estimates at various depths and chosen time intervals, but are computationally expensive and may depend on many auxiliary inputs that are not always available globally or in sufficient quality or resolution (Mishra et al., 2020; Albergel et al., 2008). In contrast, remote sensing retrievals are available only at satellite overpass times and unreliable under various conditions including frozen ground, dense vegetation, or radio frequency interference (RFI) (Gruber et al., 2019; Dorigo et al., 2017). Moreover, microwaves used for SM retrieval contain mainly information on water content in the surface layer, hampering their usability for studying or modelling processes in the soil root zone. Root-zone soil moisture (RZSM), often defined as the water present in the top meter of the soil column (Mishra et al., 2020; Baldwin et al., 2017; de Lange et al., 2008), is a component of the Global Climate Observing System (GCOS) ECV portfolio and a necessary variable for closing the water cycle (GCOS, 2016, 2022). RZSM also represents the water available for plant water uptake and thus affects evapotranspiration rates (Martens et al., 2017; Ford et al., 2014; Albergel et al., 2008) and plays a critical role in agricultural productivity forecasting (Wang et al., 2017) and drought monitoring (Vreugdenhil et al., 2022; Tobin et al., 2017).

The existing link between SM dynamics in the surface layer and the root zone (Albergel et al., 2008; Wang et al., 2017; Ford et al., 2014; Sure and Dikshit, 2019) allows for estimating RZSM from surface SM (SSM) observations via a variety of hydrological models~~(Manfreda et al., 2014; Bouaziz et al., 2020), including reanalysis (Muñoz Sabater et al., 2021; Rodell et al., 2004) and data assimilation techniques~~. These include relatively simple two-layer approaches approximating RZSM as a function of SSM (Manfreda et al., 2014), compound process-based models requiring sophisticated parameter calibration (Bouaziz et al., 2020) , as well as immensely complex and computationally expensive land surface models requiring many auxiliary inputs (Muñoz Sabater et al., 2 . Satellite-based SSM observations can also be assimilated into a land surface model to produce estimates of RZSM with global coverage, as in the case of SMAP L4 RZSM product (Reichle et al., 2017a). An alternative, less complex approach that approximates RZSM solely from SSM estimates—and can thus be readily applied to satellite retrievals—is the so-called exponential filter (EF) method (Wagner et al., 1999; Albergel et al., 2008). In essence, the EF method approximates conditions in the root zone by smoothing and delaying SSM, which is generally characterized by greater fluctuations (Beck et al., 2009; Mahmood and Hubbard, 2007). Even though the coupling strength between the surface and root-zone layers decreases with depth (Mahmood and Hubbard, 2007; Ford et al., 2014; Mishra et al., 2020) and the skill of the method in predicting RZSM has been demonstrated to deteriorate accordingly (Paulik et al., 2014; Brocca et al., 2010; Sure and Dikshit, 2019), it is still widely used due to its relatively good performance and independence of ancillary inputs as well as its low computational cost and overall

simplicity. However, the EF method is susceptible to prolonged data gaps in SSM data and thus requires an adequate number of input observations within a time interval consistent with the temporal scale of RZSM dynamics.

Regardless of the method used to derive RZSM estimates, most products do not provide information about the magnitude of random errors such as the standard deviation of their distribution, hereinafter referred to as uncertainties (Gruber et al., 2020). Two approaches have been proposed to characterize the time-variant quality of RZSM estimates derived with the EF method. The first approach, reported in Bauer Marschallinger (2018) and also utilized in this study, is a quality flag that is derived from the number of valid SSM estimates available within a specific time window preceding a specific EF-based RZSM estimate. The second approach, proposed by De Santis and Biondi (2018), uses the standard law of uncertainty propagation (Taylor, 1997) in order to characterize the random error variances of EF-based RZSM estimates. This approach takes into account the uncertainties of both the SSM input data and the EF model parameter, but does not consider the model structural ~~error (Beven, 2005)~~ uncertainty (Beven, 2005) of the EF method. The latter, due to the simplistic nature of the EF method and the limited surface–root zone coupling, can also contribute significantly to the uncertainty budget and thus must not be neglected when characterizing product errors.

In this paper, we propose to estimate the model structural uncertainty of the EF using in situ measurements of surface and root-zone SM from the ISMN. We then use these estimates together with the law for the propagation of uncertainties (similar to De Santis and Biondi (2018)) to produce a global, fully error-characterized RZSM data set for four soil layers (0–10 cm, 10–40 cm, 40–100 cm, and 100–200 cm) between 2002 and 2020, taking C3S soil moisture as input to the model. ~~To~~ While other EF-based datasets exist (e.g., the SMOS L4 product), they offer limited spatio-temporal coverage and lack quantitative uncertainty information (**??**). The focus and novelty of this paper lie in quantifying, rather than reducing, the EF model's known limitations by providing a methodology for comprehensive uncertainty estimation for the EF method. Additionally, to our best knowledge, this dataset is, as yet, the longest available solely observation-based, error-characterized global RZSM product.

2 Datasets and data pre-processing

2.1 C3S surface soil moisture

Global input satellite surface observations were obtained from the Copernicus Climate Change Service (C3S) Surface Soil Moisture COMBINED product v202012, hereinafter referred to as C3S SSM. C3S SSM is a merged product that combines satellite SSM retrievals from four active and ten passive microwave sensors (see Figure 1) into a daily global dataset on a regular 0.25 degree grid, expressed in volumetric units (m^3/m^3) (C3S, 2020). Invalid retrievals due to frozen ground, dense vegetation, RFI, and other factors are masked out. Although the C3S product provides SSM data from 1978 onward, their quality and spatio-temporal coverage increases significantly in more recent periods when sensors measuring in frequency domains better suitable for SSM retrieval are available. Therefore, only C3S SSM data for the period 2001-2020 were used in this study. Note that data from the first year of this period was used only as the model adjustment period and not included in later analyses.

The uncertainty estimates provided for the merged SSM retrievals in the C3S SSM product were computed by means of Triple Collocation Analysis (TCA)~~(Gruber et al., 2017)~~. More specifically, (stationary) uncertainties were estimated for each satellite sensor separately and used to calculate the merging weights. Uncertainties of the merged SSM estimates were then calculated from the law for the propagation of uncertainties ~~to account for the quality improvement~~ (i.e., predicting the uncertainty reduction due to the ~~merging~~weighted averaging, assuming that merging weights are correct; see (Gruber et al., 2017)). Note that the distinctive life spans and spectral bands of the used satellite missions ~~, therefore, lead also~~ (e.g., C and X-bands used by AMSR-E, L-band used by SMOS and SMAP) can potentially also lead to distinctive changes in the data quality of the merged product via the differences in their sensitivity to precipitation or evaporation. These sudden changes in ~~product uncertainty~~ SSM and uncertainty data are hereinafter referred to as ~~structural~~ systemic breaks (Preimesberger et al., 2021). Although said breaks have a marginal impact on the SSM signal itself due to the inter-calibration of sensors, they are distinct in the uncertainty estimates. As more and newer sensors provide ~~better-quality~~ better retrievals, mean uncertainty values ~~after each structural break typically decrease (Gruber et al., 2017). This is apparent, e.g., in the shift in C3S SSM uncertainty values after the introduction of AMSR-E in 2002 (van der Schalie et al., 2017; Gruber et al., 2019).~~ typically decrease distinctively with every new satellite launch in more recent periods (Gruber et al., 2017).

C3S data are readily available from the Copernicus Climate Data Store (CDS) and detailed information on the C3S dataset and its underlying ESA CCI v5 merging algorithm can be found in the relevant documentation (C3S, 2020; Dorigo et al., 2021c).

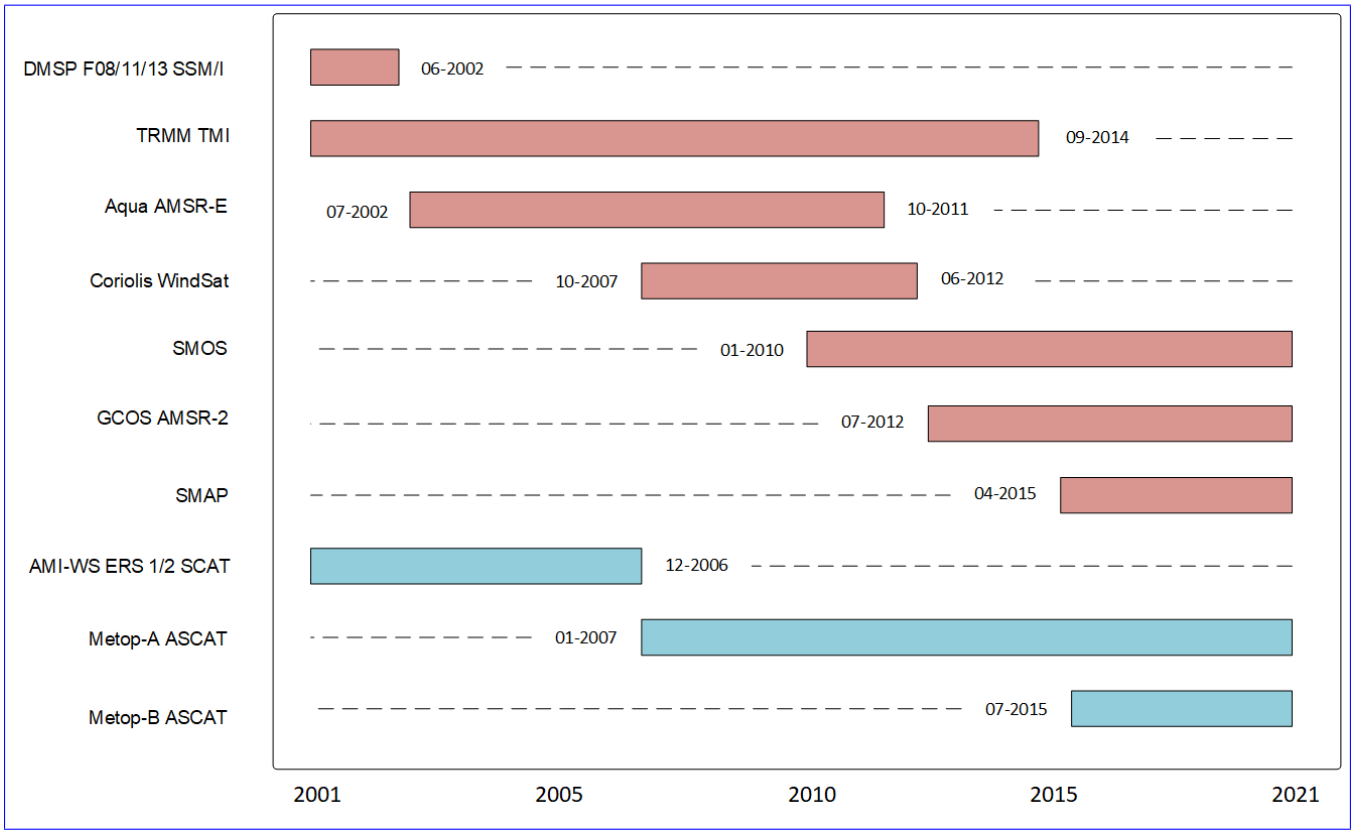

Figure 1. A timeline of satellite missions utilized in the C3S v202012 dataset during the period 2001-2020. Passive sensors (radiometers) are represented by red bars, while active sensors (scatterometers) are represented by blue bars.

2.2 Soil moisture field measurements

Field measurements for optimizing the model parameters of the EF method and for estimating its uncertainties were obtained from the International Soil Moisture Network (ISMN) for the period 2002-2020 (Dorigo et al., 2021b). Only data from sensors with a measuring depth ≤ 200 cm and internally flagged as reliable (Dorigo et al., 2013) were considered. Measuring depths of SM sensors placed vertically in a depth range, e.g., 10–40 cm, refer to their mean measuring depth. Data from multiple sensors installed at the same location and depth were averaged. ISMN data, typically available as hourly readings, were aggregated to mean daily values to match the temporal sampling of satellite observations. Furthermore, we only used ISMN stations where at least 100 data points concurrent with C3S SSM retrievals were available. Notably, approximately 80% of selected ISMN time series originate from North America and Europe (Figure A1) and the availability of data declines with depth.

2.3 ERA5-Land soil moisture

ERA5-Land (E5L) is a multi-decadal climate reanalysis with an extensive portfolio of land variables computed by the assimilation of ERA5 atmospheric variables into the H-TESSEL land surface model (Muñoz Sabater et al., 2021). Modelled SM data are available for four depth layers (0–7 cm, 7–28 cm, 28–100 cm, and 100–289 cm) on a regular 0.1 degree grid and accessible via the Copernicus Climate Data Store (CDS) (Muñoz Sabater, 2019, 2021). We used E5L for a product intercomparison with the RZSM product developed in this study, carried out for the period 2002–2020 within the Quality Assurance for Soil Moisture framework (QA4SM; https://qa4sm.eu), which automatically resamples and matches observations of the compared datasets and delivers a wide range of validation metrics.

3 Methods

3.1 Exponential filter

The EF method (Wagner et al., 1999) relies on a simple two-layer water balance model where the only considered exchange between the surface layer and the reservoir below it is infiltration. The method assumes that the fluxes from the surface to the sub-surface layers are proportionate to the difference in SM content between both layers. In this study, we utilize the recursive formulation of the method (Albergel et al., 2008):

$$RZSM(t_n) = RZSM(t_{n-1}) + K_n \cdot (SSM(t_n) - RZSM(t_{n-1})) \tag{1}$$

t_n and t_{n-1} denote timestamps (in days) of the current and previous SSM observations, respectively. Conditions in the root zone are approximated by a weighted combination of the new input SSM observation and past model estimates, with more recent estimates receiving higher weights on a time scale defined by the method's only parameter *T* (temporal length, typically in days). Weights are controlled by the gain term *K*, which ranges from 0 to 1 and is calculated as follows:

$$K_n = \frac{K_{n-1}}{K_{n-1} + e^{-\frac{t_n - t_{n-1}}{T}}} \tag{2}$$

At initialization, when no preceding estimates are available, the EF calculation is started with $K_0 = 1$ and $RZSM(t_0) = SSM(t_0)$.

Temporal variability in the root zone is generally smaller than at the surface, hence the T-value and its associated level of smoothing applied to the SSM data increase with depth (Wagner et al., 1999; Paulik et al., 2014; Wang et al., 2017; Beck et al., 2009; Mahmood and Hubbard, 2007). The optimal T-value (T_{opt}, the value that leads to the best possible representation of RZSM at a certain location using the EF), has been related to differences in utilized SSM sensors (Bouaziz et al., 2020; Sure and Dikshit, 2019), SSM sampling frequency (Brocca et al., 2010; Pellarin et al., 2006), and land surface features (Albergel et al., 2008; de Lange et al., 2008). In particular, T acts as a conglomerate proxy for various environmental factors assumed

to rule the infiltration process (e.g., soil texture, evapotranspiration, and climate), but past research on the importance of the exact driving factors is inconclusive and even contradictory (Wang et al., 2017; Bouaziz et al., 2020). To optimize the T-parameter, numerous control factors have been tested (Bouaziz et al., 2020; Mishra et al., 2020; Stefan et al., 2021) and ever more sophisticated methods been employed, including machine learning approaches (Grillakis et al., 2021). Other limitations of the method include generally poorer performance in arid zones and when soil texture is not homogeneous throughout the soil column (Yang et al., 2022; Ford et al., 2014).

Due to the high spatio-temporal heterogeneity of SM (Famiglietti et al., 2008) and its surface–root zone coupling—and hence the difficulty in properly estimating the T-parameter accurately—an uncalibrated value of T=20 has sometimes been used to describe all of the water content in the first 100 cm of the soil column (Wagner et al., 1999; de Lange et al., 2008). Results obtained by using a constant value T=20 were similar to those obtained with T-values calibrated for soil texture (de Lange et al., 2008). Limited sensitivity of the EF to T due to different environmental factors was also observed by other studies, which supports choosing a single value for T_{opt} to represent a particular depth for large areas or even globally (Albergel et al., 2008; Brocca et al., 2010, 2011; Grillakis et al., 2021). It is precisely such limitations that we attempt to describe with the uncertainty estimation scheme developed in this study, and hence advance the understanding of the EF method's performance.

3.1.1 RZSM quality flags

Prolonged temporal data gaps will cause K to increase and may cause the EF to put excessive weight on new SSM input. In the extreme case, a very long data gap (whose duration depends on the chosen T-value) can reset the EF to the initial state of $K_n = 1$ and *RZSM(t_n) = SSM(t_n)* (see above). We run a one-year adjustment period (2001) for K to reach an equilibrium state, and utilize the EF quality flag ($qflag$) described in Bauer Marschallinger (2018) to avoid such re-initializations due to frequent and/or persistent data gaps. The $qflag$ is recursively calculated for each RZSM estimate and reflects the availability of SSM input data in the preceding time period.

$$qflag(t_n) = \begin{cases} 1 + qflag(t_{n-1}) \cdot e^{-\frac{t_n - t_{n-1}}{T}}, & \textit{if SSM at } t_n \textit{ is available} \\ qflag(t_{n-1}) \cdot e^{-\frac{t_n - t_{n-1}}{T}}, & \textit{if SSM at } t_n \textit{ is unavailable} \end{cases} \tag{3}$$

The quality flag calculation is initialized with $qflag(t_n) = 1$. A normalization factor of $\sum_{j=0}^{\infty} e^{-\frac{j}{T}}$ is used to express the calculated flag values in percentages with higher values indicating a greater density of SSM data available for calculation. If the quality flag falls below a T-specific threshold, RZSM estimates are masked out. The thresholds used here have been interpolated from those empirically determined by Bauer Marschallinger (2022) for a set of discrete T-values (35%, 40%, 45%, 50%, 55%, 60%, 65%, and 70% for the T-values 2, 5, 10, 15, 20, 40, 60, and 100, respectively). If input data is unavailable but satisfactory data density has been achieved in preceding days, the latest RZSM estimate is propagated forward until new input data becomes available or the quality flag drops below its respective threshold. In the latter case, the output value is masked out. Importantly, even if new SSM input becomes available to the EF after prolonged data gaps, RZSM estimates derived from it remain masked until the $qflag$ exceeds the aforementioned threshold again.

3.1.2 *T*-parameter optimization

We optimize T for a particular depth of the soil column by maximizing the correlation between the satellite-based RZSM estimates and the in situ measurements (Paulik et al., 2014; Grillakis et al., 2021). Satellite and in situ data are matched in space by means of the nearest neighbour method. The impact of the spatial mismatch error between the large footprint of the satellite-based product and point-scale field measurement is mitigated by excluding time series that exhibit a correlation coefficient (Pearson's r) lower than 0.5 (Grillakis et al., 2021) or are not statistically significant ($p \geq 0.05$).

EF calculations are repeated for *T*-values 1–100 and T_{opt} is selected for each of the available ISMN time series based on the highest correlation coefficient. We then group T_{opt} values based on the measurement depth of the respective in situ sensor into four bins corresponding to the RZSM target layers. These depth layers, chosen to be 0–10 cm, 10–40 cm, 40–100 cm, and 100–200 cm, were defined to reflect those in common model-based RZSM products (Rodell et al., 2004; Muñoz Sabater et al., 2021). Finally, the median value of T_{opt} from each bin is chosen to compute a global RZSM product from the C3S SSM dataset.

A cross-validation is carried out to verify that T_{opt} values were not over-fitted to the local ISMN site conditions. Therefore, the sample set is randomly divided into 5 subsets of equal size (per bin), then each of the subsets was used once to validate the method fit to the remaining 4 bins.

3.2 Uncertainty estimation

3.2.1 Baseline method

In De Santis and Biondi (2018), the standard law for the propagation of uncertainties is applied to the EF method, assuming the errors in the SSM inputs and T-parameter to be normally distributed and uncorrelated. We use this approach as a baseline for our analyses. The recursive formulation of this baseline method is as follows:

$$\sigma(RZSM_n) = \sqrt{\Delta_n^2 + \left(\frac{\partial RZSM_n}{\partial T}\right)^2 \sigma^2(T)} \tag{4}$$

where:

$$\Delta_n^2 = K_n^2 \sigma^2(SSM_n) + (1 - K_n)^2 \Delta_{n-1}^2 \tag{5}$$

and:

$$\frac{\partial RZSM_n}{\partial T} = \frac{K_n}{T}\left[G_n(RZSM_{n-1} - RZSM_n) + e^{-\frac{t_n - t_{n-1}}{T}} \frac{T}{K_{n-1}} \frac{\partial RZSM_{n-1}}{\partial T}\right] \tag{6}$$

with G_n defined as:

$$G_n = e^{-\frac{t_n - t_{n-1}}{T}}\left(G_{n-1} + \frac{1}{K_{n-1}} \frac{t_n - t_{n-1}}{T}\right) \tag{7}$$

$\sigma(RZSM)$ and $\sigma(T)$ denote the uncertainty of the RZSM estimates (in m^3/m^3) and the EF model parameter T ~~,~~(a unit of time, in days), respectively. The equation is initialized as $\Delta_0 = \sigma(\text{SSM}_0)$, $\partial\text{RZSM}_0 / \partial T = 0$ and $G_0 = 0$. Uncertainties of the SSM input data are considered by the Δ term (in m^3/m^3), which also takes into account the effect of possible prolonged input data gaps dependent on the T-value. The Jacobian term ∂RZSM/∂T assumes high values proportional to the latest SSM input variability on a time scale related to the T-parameter (expressed as m^3/m^3 over time). This is reflected in significant changes in the RZSM value associated with wetting or drying of the soil. Finally, the term G (dimensionless) weighs the contribution of change recorded between the latest and penultimate RZSM estimates.

3.2.2 *T*-parameter uncertainty

De Santis and Biondi (2018) used an arbitrary value of $\sigma(T)$ equal to 10% of locally calibrated T_{opt}. This is in line with other studies on SM uncertainty propagation (Parinussa et al., 2011; Pathe et al., 2009), who used this uncertainty percentage for parameters without well-defined accuracy. In our study, we determine T_{opt} values based on a limited number of available in situ time series and apply these values to estimate RZSM globally. Consequently, $\sigma(T)$ is likely to be greater due to a variety of environmental conditions not accounted for or underrepresented in the available in situ sample. We therefore propose the median absolute deviation (MAD) of T_{opt} (Section 3.1.2) as a more appropriate proxy for $\sigma(T)$. In this case, the MAD is preferred over the variance because the sampling distribution of T_{opt} is both non-Gaussian and bounded (Leys et al., 2013).

3.2.3 EF model structural uncertainty

Recall that the standard law for the propagation of uncertainty (which is used in the baseline method) does not account for model structural uncertainty in the EF, which, due to the simplistic nature of the method and the limited surface–root zone coupling, can account for a significant portion of the overall uncertainty budget.

We propose to estimate model structural uncertainty ($\sigma(EF)$) from in situ data using stations that operate sensors both at the surface and in the root zone. At these stations, we derive RZSM estimates from the SSM measurements using the EF method and then compare them to actual RZSM station measurements. For this analysis, the T-value was optimized for each station and depth individually to minimize its influence on the estimation of $\sigma(EF)$. This provides direct estimates for $\sigma(EF)$ as:

$$\sigma(EF) = ubRMSD(RZSM_{EF}, RZSM_{ISMN}) \tag{8}$$

where $ubRMSD$ denotes the unbiased root-mean-square difference. Note that 'unbiased', in this case, does not only refer to a correction for bias in the mean (as is most commonly done) but also to a correction for bias in variance, which also constitutes an unintended systematic component in the $RMSE$ (Gupta et al., 2009). Only sites with measurements from more than a single depth and at least one sensor within the surface layer (≤ 10 cm) were selected. Time series with negative correlation between EF-based RZSM estimates and in situ RZSM measurements were disregarded. As a result, a total of 1509 in situ sites were considered. Note that the EF model structural uncertainty computed at the point scale is assumed to be representative for the coarse scale as well.

Finally, the EF structural uncertainties obtained from Eq. (8) add to the propagated RZSM uncertainty budget (Eq. (4)) as:

$$\sigma(RZSM_n) = \sqrt{\Delta_n^2 + \left(\frac{\partial RZSM_n}{\partial T}\right)^2 \sigma^2(T) + \sigma^2(EF)} \tag{9}$$

4 Results and discussion

In this section, we first show results of the point-scale T-parameter optimization. Next, we compare the gridded RZSM product globally to E5L. We then discuss the estimates for EF model structural uncertainties. Finally, we compare our RZSM uncertainty estimates with those obtained with the baseline method.

4.1 T-parameter optimization

After filtering out unreliable data (see Section 2.2), 3901 ISMN time series from 67 different measuring depths between 0 and 200 cm were available for the T-optimization process. Figure 2 shows the distribution of T_{opt} values binned into our four chosen RZSM layers (0–10 cm, 10–40 cm, 40–100 cm, and 100–200 cm). The median T_{opt} values for these layers were 6, 15, 48 and 70 days, increasing with soil depth as expected (Paulik et al., 2014; Wang et al., 2017). These median T_{opt} values were then used to compute RZSM globally.

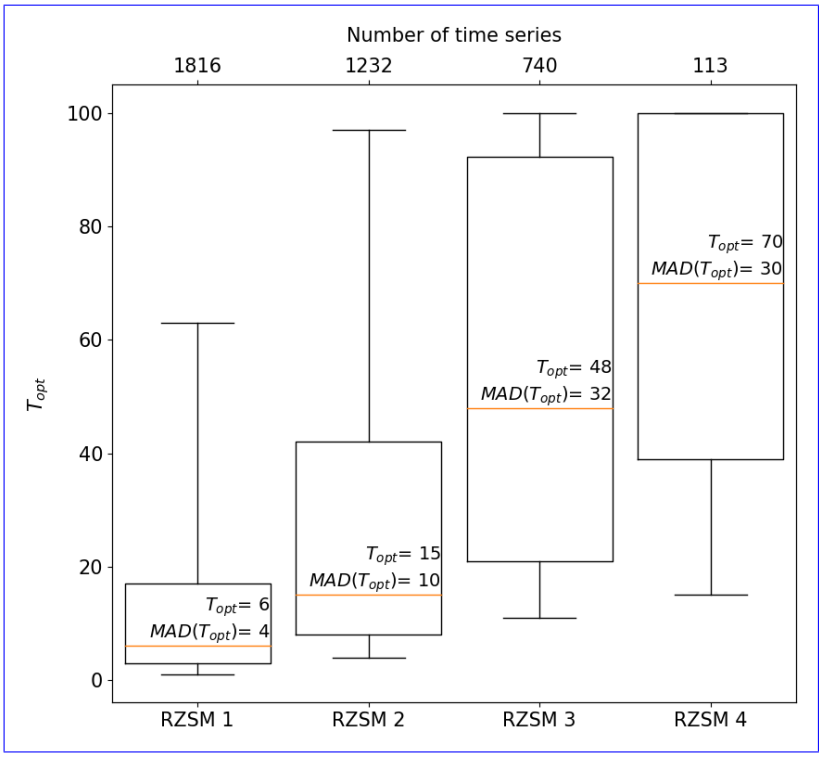

Figure 2. T_{opt} values calibrated on 3901 in situ time series and binned per RZSM layers 1–4. Median values (represented by orange lines) from each bin were used to compute a global RZSM product. Median absolute deviations ($MAD(T_{opt})$) were used in estimating RZSM uncertainties.

A five-fold cross-validation was performed to verify the robustness of this approach. The variability in median T_{opt} values per soil layer is increasing with depth but remains negligibly small in all layers, with 6, 15–16, 47–50 and 67–72 for soil layers 1–4, respectively (Figure 3a). Subsequently, the five median T_{opt} values derived from the training subsets were used to estimate RZSM for the different layers of the respective validation sets (Figure 3c) and resulted in Pearson's r of 0.64–0.67, 0.64–0.65, 0.57–0.6, and 0.48–0.6 for soil layers 1–4, respectively. When evaluating each training set directly, correlations were 0.65–0.66, 0.65, 0.58–0.59, and 0.53–0.56 for soil layers 1–4, respectively (Figure 3b).

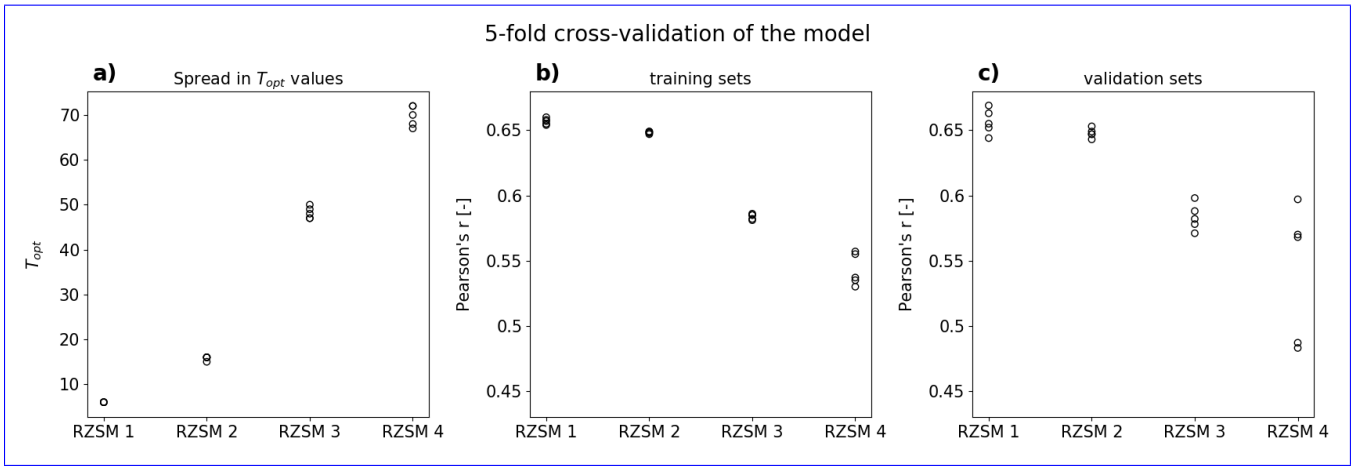

Figure 3. Cross validation results showing the spread in T_{opt} values (a), and agreement of the training (b) and validation (c) sets with in situ data.

The little variability between the validation and training sets suggests that T_{opt} values are not over-fitted to ISMN site conditions and can be used robustly in other regions as well. Notably, the spread in median T_{opt} values increases with soil depth while the correlation scores decrease. This indicates reduced reliability of the method in deeper soil layers, which is in line with the assumption that the coupling between the surface and root-zone SM decreases with depth. Note, however, that results for deeper layers are also affected by the smaller sample sizes at greater depths.

4.2 Global RZSM product quality assessment

A global SM dataset spanning the 2002–2020 period was computed using the EF method and T-parameters optimized at point scale with the approach described in section 4.1. Figures 4a–e) show correlation maps of each of the RZSM product layers as well as the input C3S SSM dataset with E5L. The spatial patterns observed in the C3S SSM data (Figure 4a) are strikingly similar to those in RZSM layer 1 (Figure 4b) with slight to moderate deterioration in performance over the high latitudes ($> 60° N$). This is not surprising given that both products differ only by a small degree of smoothing applied to RZSM layer 1 and are compared to the same E5L layer (0-7 cm). RZSM layers 2 and 3 (Figure 4c-d) are compared to E5L layers 7-28 and 28-100 cm, respectively, and largely preserve good performance in regions where the input C3S SSM product also performs well, i.e., in Europe (bar Scandinavia), the Caspian and Aral Sea basins, the Eastern United States, India, Southeast Asia, South America, Sub-Saharan Africa, and Australia. At the same time, deterioration of performance is observed in high latitudes and in arid environments such as the Sahara desert and the Arabian Peninsula where the reduced strength of coupling between the surface and root-zone dynamics may hinder the EF performance (Yang et al., 2022). The patterns of good and poor performance visible in RZSM layers 1-3, are not replicated in RZSM layer 4 (Figure 4e) where the agreement with the reference E5L 100-289 cm layer is spatially very heterogeneous and worse overall. The few regions where the good performance observed in shallower layers is preserved include India, Southeast Asia, and the Eastern United States.

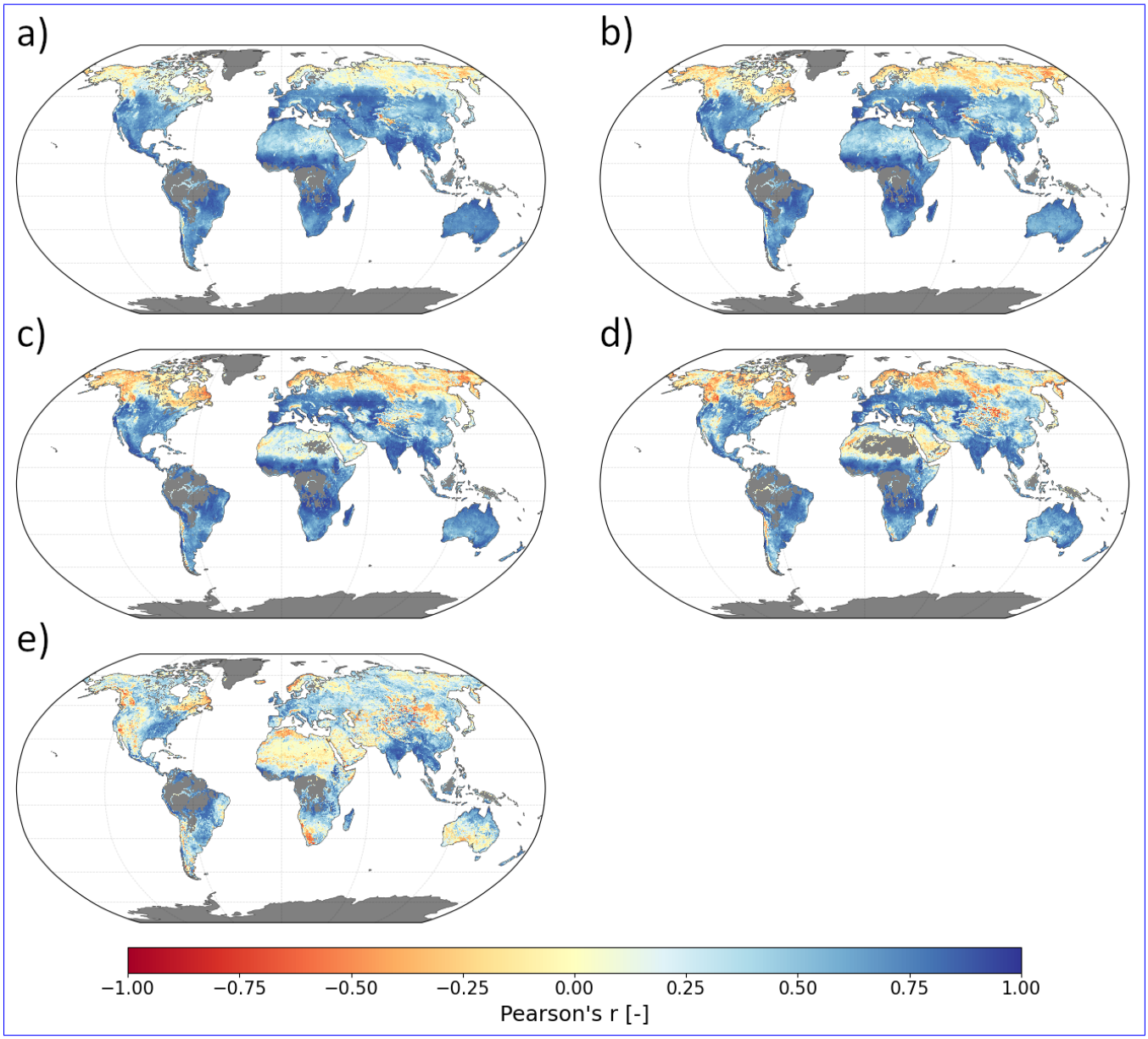

Figure 4. Spatial correlation maps of the C3S SSM (a) and RZSM products (b-e) with E5L layer 0–7 cm (a-b), 7–28 cm (c), 28–100 cm (d) and 100–289 cm (e).

Figure 5 shows a comparison between all of the RZSM product layers as well as the input C3S SSM dataset with E5L. RZSM product layers agree best with the (approximately) matching E5L depth layers in all but one case. The highest median Pearson correlations with the E5L reference layers (0–7 cm, 7–28 cm, 28–100 cm, and 100–289 cm) were obtained by C3S SSM (0.55), RZSM layer 1 (0.49), RZSM layer 3 (0.41), and RZSM layer 4 (0.28), respectively. Even though C3S SSM correlates best with

the E5L surface layer, the correlation score for RZSM layer 1 (0–10 cm) is only insignificantly smaller (0.54). Similarly, the second E5L layer (7–28 cm) best agrees with RZSM layer 1 (r=0.49) but the layer most congruent in depth (RZSM 2) is a close second (r=0.47). In the remaining depth layers, correlations between C3S SSM and E5L are substantially lower than for the RZSM product, which proves the ability of our EF approach to approximate SM below the surface layer. While RZSM layer 1 (0–10 cm) shows the best agreement with E5L layer 2 (7–28 cm; r=0.49) RZSM layer 2 (10–40 cm) correlates only slightly less with that layer (r=0.47).

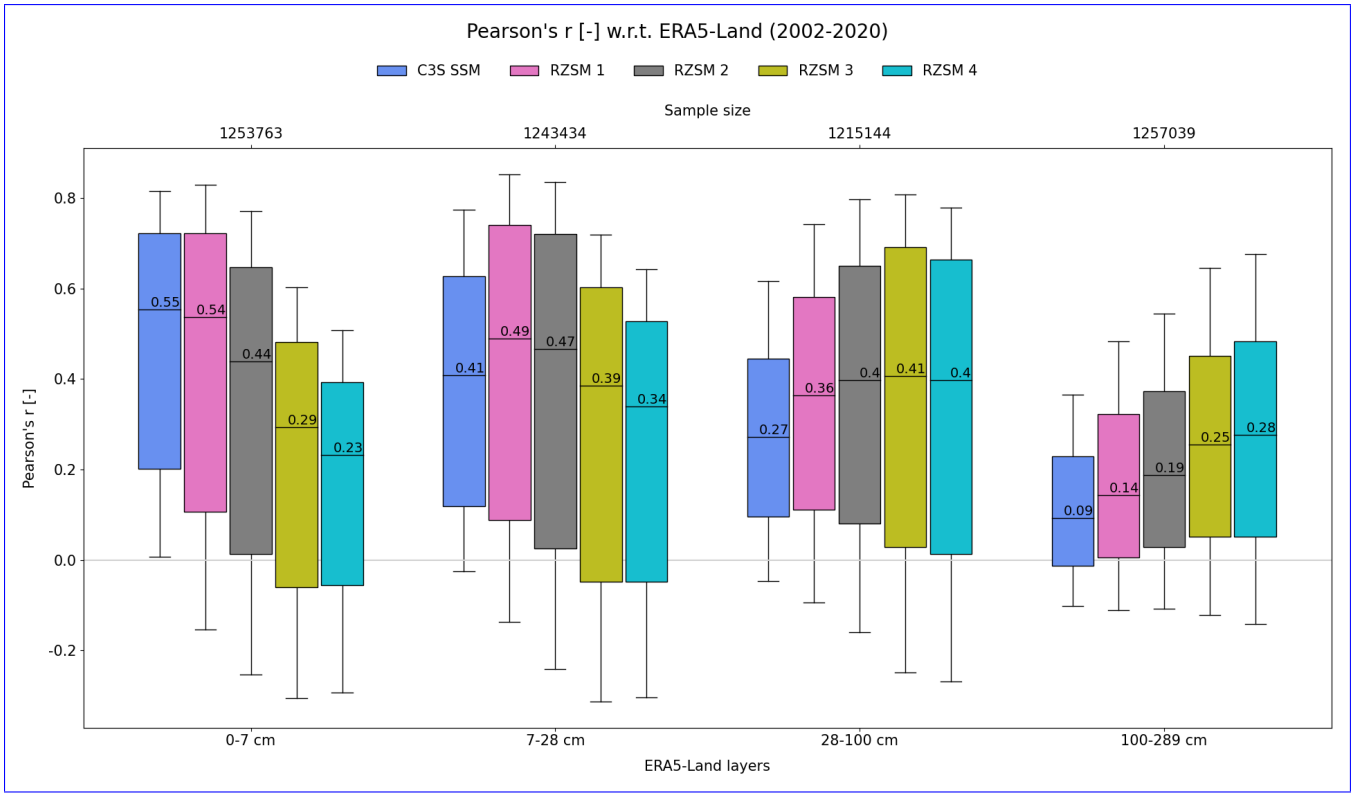

Figure 5. Product intercomparison of the C3S SSM and RZSM products against E5L SM.

Results are also consistent with the assumption of the EF model that SM dynamics decrease with depth and that T_{opt} ought to increase accordingly, as was also found by other studies (Wagner et al., 1999; Paulik et al., 2014; Wang et al., 2017; Beck et al., 2009; Mahmood and Hubbard, 2007). At the same time, the maximum correlation values decrease with depth confirming the diminishing coupling between the surface and root-zone layers, as also found at in situ station level 3 and demonstrated by others (Paulik et al., 2014; Brocca et al., 2010; Sure and Dikshit, 2019). The performance of our product is similar to that of other satellite-based RZSM products found in other studies, especially when considering the same regions for assessment (Reichle et al., 2017b; Xu et al., 2021). While the data set presented here does not outperform other existing RZSM products, it distinguishes itself as the only purely observation-based global product covering such a long time period, and the only EF-based product that has uncertainty estimates provided with it.

4.3 EF model structural uncertainty

Figure 6 shows estimates for the model structural uncertainties (Section 3.2.3) obtained at all available in situ sites, binned into the four RZSM product layers. Their median values (represented by orange lines and annotated) were used as estimates for $\sigma(EF)$. Note that in situ measurement errors were assumed to be negligible and thus not influencing ubRMSD estimates, which likely causes model structural uncertainties to be overestimated. Also, structural uncertainties are assumed to be constant in time.

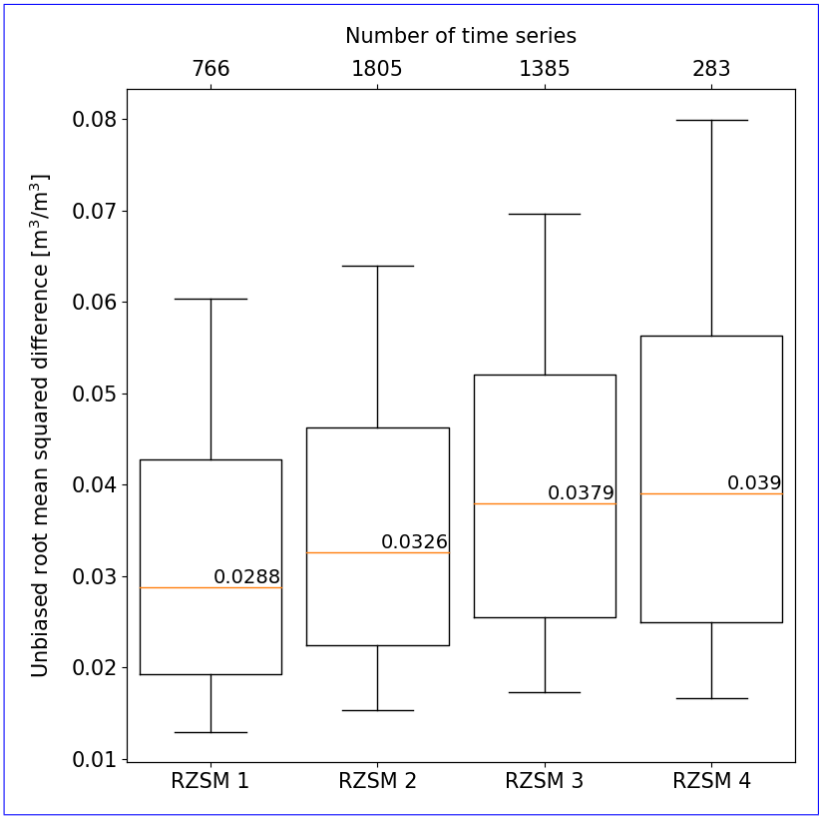

Figure 6. The ubRMSD between propagated RZSM from in situ SSM using the EF model and measurements of RZSM at the same location and in the same depth, calculated at 1509 different sites. The median ubRMSD value for each bin (represented by orange lines and annotated) represents $\sigma(EF)$ for the according T_{opt}.

As anticipated, an increase in $\sigma(EF)$ corresponds to the growing distance between the surface and the root-zone measurements, demonstrating the decreasing coupling strength between both layers. Note that $\sigma(EF)$ shows significant variability within RZSM layers which is likely, at least to some degree, related to variations in local conditions. However, as with the T-parameter optimization, we estimate structural uncertainties based only on a limited number of in situ stations and, therefore, use the median to extrapolate globally.

4.4 RZSM uncertainty budget calculation

Figure 7 compares RZSM uncertainty estimates obtained from the baseline method (De Santis and Biondi, 2018) with those from the approach proposed here. Figure 7a shows a time series of RZSM uncertainties from the baseline method at an arbitrary location in Benin (9.875N, 1,625E). Figure 7b shows the effect of changing $\sigma(T)$ from 10% of T_{opt} to the median absolute deviation of T_{opt}, which is an amplified temporal variability. Simultaneously, mean uncertainty values increase and get closer to the magnitudes of the input SSM dataset. Moreover, they no longer diminish with increasing T-values (i.e., depth) as is the case in the baseline formulation. This is presumably more realistic since the progressive decoupling between the surface and deeper soil layers can be expected to cause uncertainties to increase rather than to decrease.

Figure 7c shows the impact of accounting for $\sigma(EF)$ in the total uncertainty budget when using 10% of T_{opt} as T-parameter uncertainty ($\sigma(T)$). Considering this term substantially increases the magnitude of the propagated uncertainties and leads them to increase with depth (as does $\sigma(EF)$). However, the uncertainties' temporal variability is reduced substantially as the effect of $\sigma(T)$ is overshadowed by that of $\sigma(EF)$. Finally, Figure 7d shows the combined effect of using the MAD of T_{opt} as its parameter noise $\sigma(T)$ and accounting for model structural uncertainty $\sigma(EF)$. Compared to the baseline (Figure 7a), this yields an increased overall magnitude of the uncertainties, a more realistic increase in (temporal average) uncertainties with depth, and an amplified temporal variability in all layers during transitions between dry and wet conditions ~~, which is also expected~~ (see Figure 8). The latter effect is caused by the simplistic nature of the model, which essentially operates as a smoother and therefore attenuates sudden variations in the SSM signal which in reality may be transmitted into the deeper layers in a more significant manner. The reduced accuracy of the EF method during soil wetting and drying phases was also observed by others (Ford et al., 2014).

4.5 Assessment of uncertainty estimates

Similar as in De Santis and Biondi (2018), we assess the use of the proposed MAD estimates for $\sigma(T)$ by computing Pearson's *r* and root-mean-square differences (RMSD) with respect to in situ data before and after removing a fixed percentage of the data (5, 10, 15, and 20%) with highest uncertainty estimates. In case of effective correspondence between high values of both the estimated RZSM uncertainties and the observed RZSM deviations from reference in situ measurements, it is expected that the skill metrics will improve due to the masking. This hypothesized correspondence holds well as long as the difference between in situ and satellite-based RZSM values is mainly due to the random errors of the latter. Note that this analysis is only sensitive to the impact of using different values for $\sigma^2(T)$ ($(T_{opt}/10)^2$ versus $MAD(T_{opt})^2$) since the estimated structural uncertainty $\sigma(EF)$ is constant in time and therefore cannot change the ranking of the total uncertainties.

~~Figure 8shows both SM and uncertainty estimates for C3S SSM and RZSM layer 2 estimates for the same site as in Figure 7 (9.875N, 1,625E) and highlighting~~ Figures 8a) and d) indicate (in magenta shading) 20% of RZSM layer 2 data with the highest uncertainties masked out in the experiment described above based on uncertainties estimated with the baseline (b), and our method (d), respectively. Overall, despite the differences in magnitude and amplitude, both our and the baseline method assign the highest uncertainty values to timestamps corresponding to significant soil wetting or drying events. However, in the

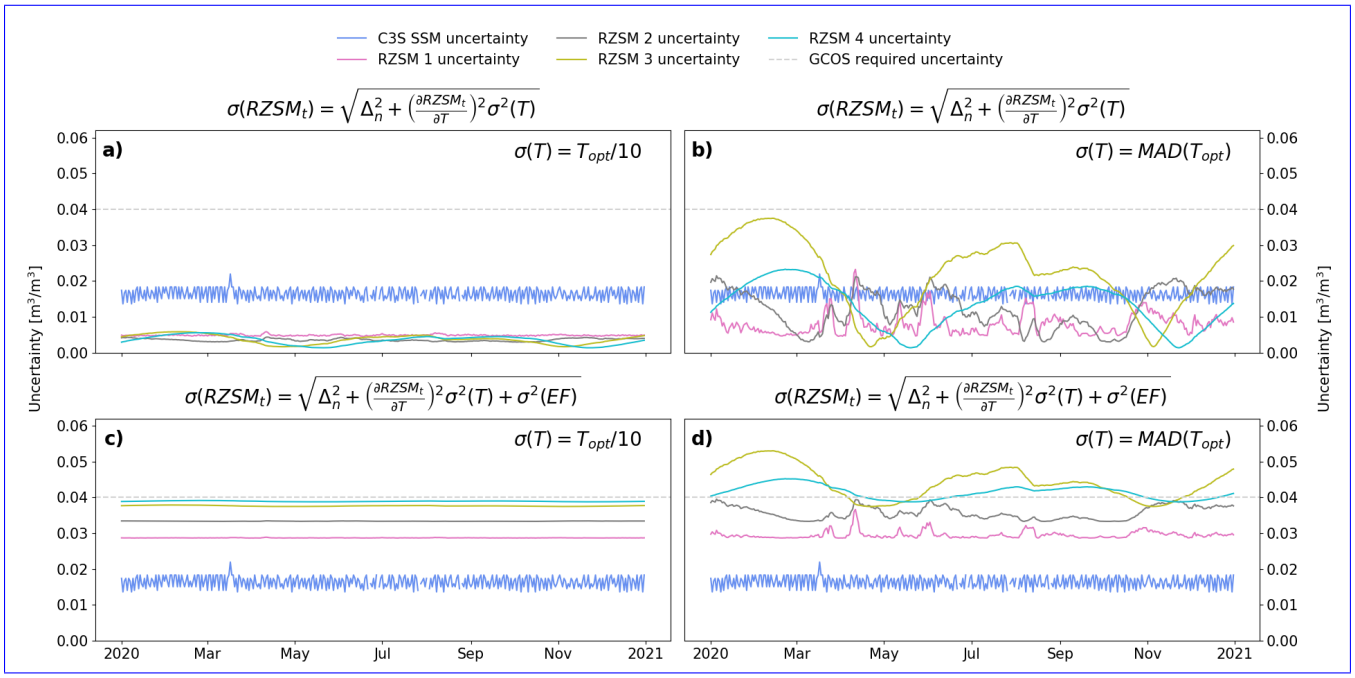

Figure 7. Evaluation of the impact of changes to the baseline method illustrated on an example 2020 time series from 9.875N, 1,625E. C3S SSM uncertainties were propagated with the baseline scheme in a), while b) and c) show the individual impacts of increasing the noise of T from 10% of T_{opt} to $MAD(T_{opt})$ and adding the term $\sigma^2(EF)$, respectively. Combined effects of both changes are shown in d). The dashed grey line indicates ~~target RZSM~~ the uncertainty level ~~as required~~ defined by GCOS (2022) as an accuracy goal for RZSM products.

baseline method the average magnitude of SSM input uncertainty appears to have a greater influence on the calculated RZSM uncertainty estimates. This is most evident when comparing values before and after the inclusion of Metop-A ASCAT into the C3S product in January 2007 (indicated by the dashed vertical line in Figure 8), which substantially improved data quality thereafter. Specifically, mean C3S SSM uncertainty dropped from 0.029 m^3/m^3 to 0.018 m^3/m^3. Such a clear shift is also visible in the uncertainty values propagated with the baseline method (from 0.008 m^3/m^3 before to 0.004 m^3/m^3 after the introduction of Metop-A ASCAT). This causes the baseline method to predict the majority of the 20% most uncertain SM values to occur in the pre-ASCAT period. In contrast, in our approach, average uncertainties remain stable (at 0.036 m^3/m^3) over the entire time period. This suggests that the use of $MAD(T_{opt})^2$ as an estimate for T-parameter uncertainty reduces the sensitivity to ~~structural~~ systemic breaks, i.e., large variations between the uncertainties of the C3S SSM input sensors, and improves the method's capability to predict day-to-day uncertainty variations. Lastly, after the introduction of ASCAT, both schemes consistently assign higher uncertainties to timestamps characterized by large SM changes. Taken together, while the use of $T_{opt}/10$ as T parameter uncertainty seems to yield realistic estimates for uncertainty variations due to the use of different C3S SSM input sensors, using $MAD(T_{opt})$ as T parameter uncertainty seems to better predict day-to-day uncertainty variations in the RZSM estimates.

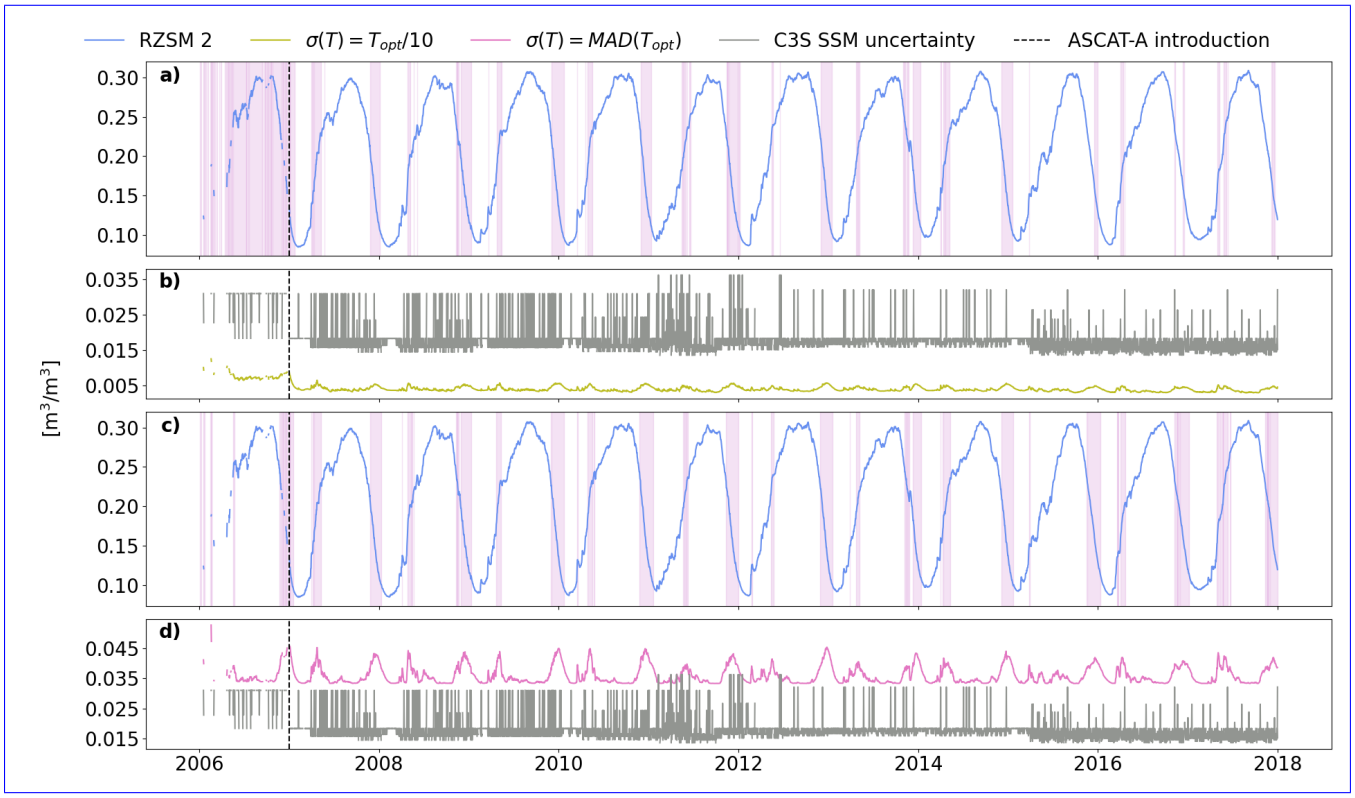

Figure 8. Differences in uncertainty variations of the baseline (~~a~~a-b) and our proposed uncertainty estimation approach (~~b~~c-d). Illustrated on the example of ~~AMMA-CATCH: Belefoungou-Mid~~ RZSM layer 2 at an arbitrary location in Benin (9.875N, 1,625E)~~field measurements from 20 cm depth~~.

Figure 9 shows the results of the data removal experiment described above, summarized for all considered ISMN stations. To compare the performance with and without the effect of C3S ~~structural~~ systemic breaks on the uncertainty values (see above), results are shown for both the full product period (2002–2020; Figure 9a–d) and a sub-period without breaks, i.e., from the inclusion of SMAP data onward (April 1st 2015–2020; Figure 9e–h). In both cases, correlation coefficients obtained for the complete time series were compared to those obtained after removing 5, 10, 15, and 20 % of data with the highest associated uncertainties.

In case of the full product period (Figure 9a–d), using $\sigma(T) = T_{opt}/10$ as T~~parameter~~ parameter uncertainty seems to ~~consistently yield more realistic estimates of temporal uncertainty variations~~ yield more consistent improvements in correlation with the in situ reference after removing a percentage of the most uncertain data, than using $\sigma(T) = MAD(T_{opt})$. This is true for all four soil layers. Masking out more uncertain data indicated by either method consistently improves agreement with in situ reference data in the first two product layers. This improvement increases the more data are masked out, as is expected. In the absence of such breaks (Figure 9e–h) RZSM uncertainty variations seem to be better predicted when using

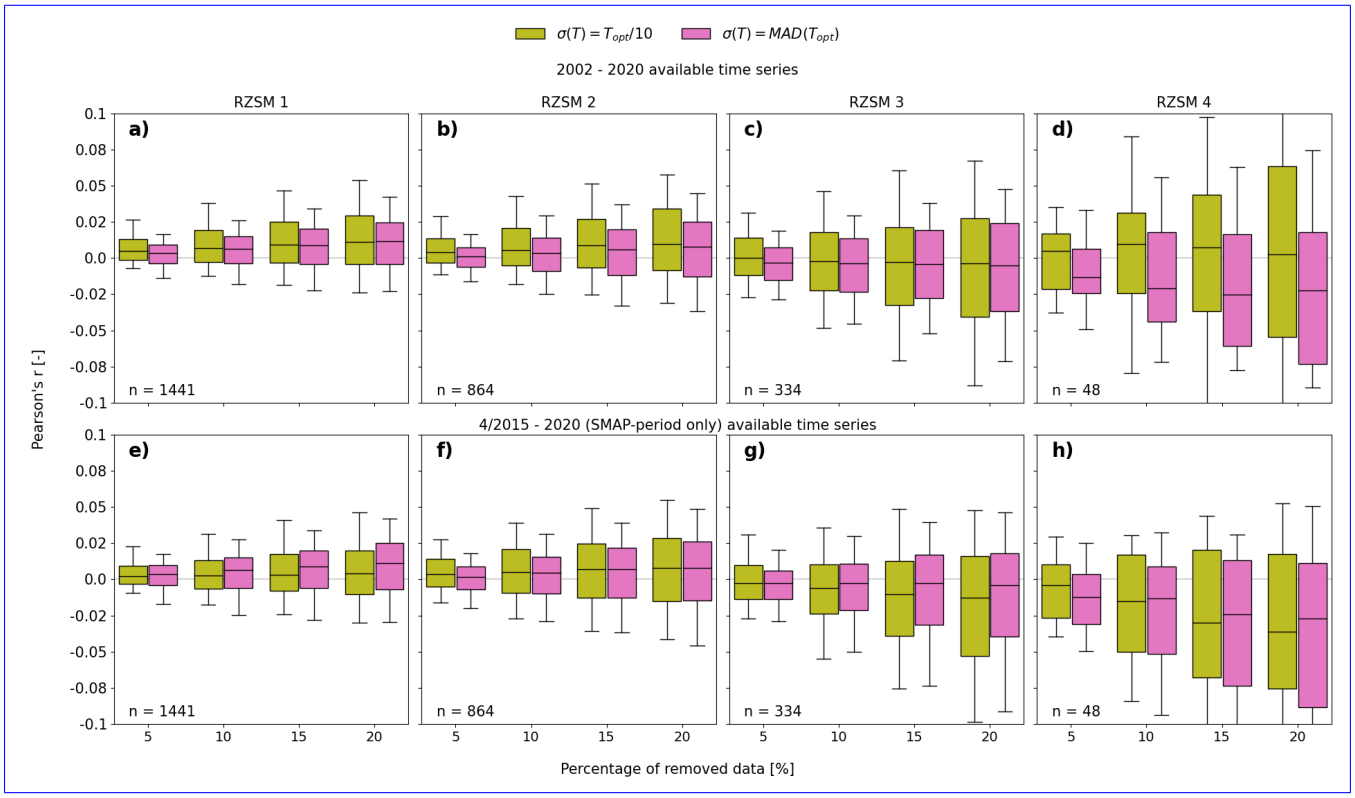

Figure 9. Correlations with in situ measurements (y-axis) before and after removing a fixed percentage of data with the highest uncertainty (x-axis) for the period 2002–2020 (a–d) and 2015–2020 (e–h). Uncertainties were calculated using either $\sigma(T) = T_{opt}/10$ (olive colors) or $\sigma(T) = MAD(T_{opt})$ (orchid colors).

$\sigma(T) = MAD(T_{opt})$ as T parameter uncertainty in almost all cases. Notably, in layers 3 and 4, data removal according to either method degraded the agreement with field measurements.

At greater depths, the contribution of the model structural uncertainty to the total uncertainty budget has been shown to increase. In the circumstances where the EF model appears to be inadequate, for example due to poor coupling between the root zone in consideration and the surface layer, it can be assumed that the model structural uncertainty is so predominant as to make the temporal patters of the other uncertainty components marginal in practice. However, in circumstances where the magnitude of the real uncertainty is such as to make the EF-based RZSM so unreliable, the lack of ability to reproduce the temporal variations of the estimated uncertainty becomes less relevant.

In summary, the propagation of C3S SSM input uncertainties yields accurate predictions of temporal uncertainty variations of RZSM estimates obtained with the EF method for the first two layers (0–10 cm and 10–40 cm). This is no longer the case for deeper layers (40–100 cm and 100-200 cm). Note, however, that the RZSM estimates in these layers themselves still exhibit reasonable skill when evaluated against E5L (see Figure 5).

5 Summary and Conclusions

In this study, we computed root-zone soil moisture (RZSM) globally in four depth layers (0–10 cm, 10–40 cm, 40–100 cm, and 100-200 cm) from merged satellite surface soil moisture (SSM) retrievals of the Copernicus Climate Change Service (C3S) COMBINED product v202012 using the exponential filter (EF) method. The EF model parameter T has been optimized at point scale by maximizing the correlation against globally-distributed in situ SM measurements from the International Soil Moisture Network (ISMN). The median of the optimized T values at each layer have been used to compute the global product. A global product intercomparison with ERA5-Land (E5L) reanalysis SM data has shown a satisfactory level of agreement in all layers (global median correlations of the four above-mentioned product layers against E5L reference layers 0–7 cm, 7–28 cm, 28-100 cm, and 100-289 cm were 0.54, 0.47, 0.41, and 0.28, respectively).

Uncertainties in the RZSM estimates obtained with the EF method were calculated using the law for the propagation of uncertainties. Uncertainties of the input SSM data were available in the C3S product and have been calculated by the data producers using Triple Collocation Analysis (TCA). We tested the use of the median absolute deviation of optimized T parameters at the available ISMN locations ($MAD(T_{opt})$) as a proxy for T parameter noise. Results obtained using $MAD(T_{opt})$ in uncertainty propagation were compared with results obtained using 10% of the optimized T parameter itself ($T_{opt}/10$) used in earlier studies. While the use of $T_{opt}/10$ as T parameter uncertainty seems to yield realistic estimates for uncertainty variations due to the use of different C3S SSM input sensors, using $MAD(T_{opt})$ as T parameter uncertainty seems to better predict day-to-day uncertainty variations in the RZSM estimates. A higher value assumed by $\sigma(T)$ (in this case $MAD(Topt)$) places more weight on short-term significant variations in RZSM values (accounted for by the Jacobian term ∂RZSM/∂T) and overshadows the contribution of the input uncertainties (Δ) to the overall uncertainty budget. This approach results in higher uncertainty outputs paralleling significant changes in RZSM signal (e.g., soil wetting/drying events) and is generally better suited to describe day-to-day uncertainty variations. Meanwhile, lower value of $\sigma(T)$ (here $Topt/10$) favors the impact of the input uncertainties and appears to be more skillful in detecting sudden shifts in the magnitude of the input uncertainties due to C3S SSM sensor changes. While both the significant variations in RZSM values and the magnitude shifts in the input uncertainties are crucial elements of the overall uncertainty budget, there appears to be a trade-off in favoring the impact of one or the other based on the value assumed by $\sigma(T)$.

Even though propagating SSM input and model parameter uncertainties yields credible predictions of temporal uncertainty variations, absolute uncertainty magnitudes appear unrealistically small (below 0.01 m^3/m^3). This is because the propagation of uncertainty only accounts for uncertainties in the data and parameters input to the EF method, but not for limitations of the EF method itself (e.g., the progressive inability of the method to model deeper-layer RZSM due to vanishing surface–root zone coupling). We proposed to estimate these EF model structural uncertainties as the unbiased root-mean-square differences between RZSM estimates for each of our four product depth layers obtained by applying the EF method to in situ SSM measurements, and actual in situ RZSM measurements taken at the same location and depth. This was done at all available ISMN sites and the median of these estimates used as a global proxy of EF structural uncertainty for each of the four product depth layers, respectively. Combined, propagated SSM input and model parameter uncertainties and EF structural uncertainties were

considered to yield realistic estimates of the total RZSM product uncertainty budget in all layers (global mean uncertainties of the four product layers are 0.031, 0.035, 0.04, and 0.04 m^3/m^3). Note, however, that a quantitative validation of uncertainty magnitudes is still pending due to the lack of reliable uncertainty reference data on a global scale and for different RZSM depth layers.

The EF parameter uncertainty was estimated on a global scale and can be expected to differ for smaller scales, especially where the variability in environmental conditions is lower. Similarly, estimates of the EF model structural uncertainty are likely to differ on local to regional scales. Also, the structural uncertainty of the EF, here assumed to be constant in time, could in fact vary on a sub-seasonal scale given the phenomena that regulate the process of water transfer in the soil. Moreover, random errors of the in situ measurements were assumed to be negligible and were not accounted for in estimating the structural uncertainty of the model. Nonetheless, it is plausible that the EF structural uncertainty is much greater than the random uncertainty of the in situ sensors. Estimates of the random uncertainty of in situ sensors could allow for a more accurate estimation of the EF structural uncertainty in the future.

Further insights could also be gained by evaluating the behaviour of the proposed method in propagating uncertainties of different SSM input data, i.e., single-sensor products without ~~structural~~ systemic breaks and non-static input SSM uncertainties obtained by different means than TCA. Nonetheless, this study is an important step towards understanding and describing the uncertainties of EF-based RZSM products.

6 Code and data availability

Python package used in the computation of the root-zone soil moisture data and its associated uncertainties from surface soil moisture observations by means of exponential filter: https://github.com/TUW-GEO/pyswi

Global root-zone soil moisture data produced and utilized in this study, available for 2002-2020 period as daily image files in netCDF4 format: https://doi.org/10.48436/9gsg6-nn854

Appendix A: ISMN references

Table A1: ISMN networks used in this study.

Network	Time series used for T-parameter optimization	Time series used for EF model structural uncertainty estimation	Reference
AMMA-CATCH	31	27	Mougin et al. (2009); Cappelaere et al. (2009); de Rosnay et al. (2009); Lebel et al. (2009); Galle et al. (2015)
ARM	90	113	Cook (2016a, b, 2018)
AWDN	112	148	-
BIEBRZA_S-1	15	18	Musial et al. (2016)
BNZ-LTER	7	22	Van Cleve et al. (2015)
CALABRIA	12	-	Brocca et al. (2011)
CAMPANIA	1	-	Brocca et al. (2011)
COSMOS	65	2	Zreda et al. (2008, 2012)
CTP-SMTMN	147	167	Yang et al. (2013)
DAHRA	4	4	Tagesson et al. (2015)
FLUXNET-AMERIFLUX	22	16	-
FMI	16	37	Ikonen et al. (2016, 2018)
FR_Aqui	28	23	Al-Yaari et al. (2018); Wigneron et al. (2018)
GROW	118	-	Xaver et al. (2020); Zappa et al. (2019, 2020)
GTK	-	24	-
https://www.overleaf.com/project/6194e50984e1f0b1f43a008b HiWATER_EHWSN			Kang et al. (2014); Jin et al. (2014)
HOAL	90	97	Blöschl et al. (2016); Vreugdenhil et al. (2013)
HOBE	64	60	Jensen and Refsgaard (2018); Bircher et al. (2012)

HSC_SEOLMACHEON	1	-	-
HYDROL-NET_PERUGIA	4	6	Flammini et al. (2018a, b); Morbidelli et al. (2011, 2014, 2017)
ICN	-	24	Hollinger and Isard (1994)
IIT_KANPUR	-	3	-
IMA_CAN1	9	-	Biddoccu et al. (2016); Raffelli et al. (2017); Capello et al. (2019)
IPE	1	-	Alday et al. (2020)
iRON	5	16	Osenga et al. (2019, 2021)
KIHS_CMC	54	38	-
KIHS_SMC	51	32	-
LAB-net	2	1	Mattar et al. (2014, 2016)
MAQU	53	62	Su et al. (2011); Dente et al. (2012)
MOL-RAO	9	10	Beyrich and Adam (2007)
MySMNet	15	11	Kang et al. (2019)
NAQU	5	31	Su et al. (2011); Dente et al. (2012)
NGARI	5	84	Su et al. (2011); Dente et al. (2012)
NVE	-	10	-
ORACLE	24	32	-
OZNET	101	105	Young et al. (2008); Smith et al. (2012)
PBO_H2O	115	-	Larson et al. (2008)
PTSMN	80	60	Hajdu et al. (2019)
REMEDHUS	22	-	González-Zamora et al. (2019)
RISMA	51	62	Canisius (2011); L'Heureux (2011); Ojo et al. (2015)
RSMN	13	-	-
SASMAS	27	13	Rüdiger et al. (2007)
SCAN	575	806	Schaefer et al. (2007)
SKKU	56	42	Nguyen et al. (2017)
SMN-SDR	76	127	Zhao et al. (2020); Zheng et al. (2022)
SMOSMANIA	79	66	Calvet et al. (2007); Albergel et al. (2008); Calvet et al. (2016)
SNOTEL	788	942	Leavesley et al. (2008)

SOILSCAPE	385	247	Moghaddam et al. (2010, 2016); Shuman et al. (2010)
SWEX_POLAND	6	17	Marczewski et al. (2010)
TAHMO	68	10	-
TERENO	14	10	Zacharias et al. (2011); Bogena et al. (2012, 2018); Bogena (2016)
UDC_SMOS	16	11	Loew et al. (2009); Schlenz et al. (2012)
UMBRIA	37	28	Brocca et al. (2008, 2009, 2011)
UMSUOL	4	6	-
USCRN	309	358	Bell et al. (2013)
USDA-ARS	4	-	Jackson et al. (2010)
VAS	1	-	-
VDS	12	8	-
WEGENERNET	1	-	Kirchengast et al. (2014); Fuchsberger et al. (2021)
WSMN	1	-	Petropoulos and McCalmont (2017)

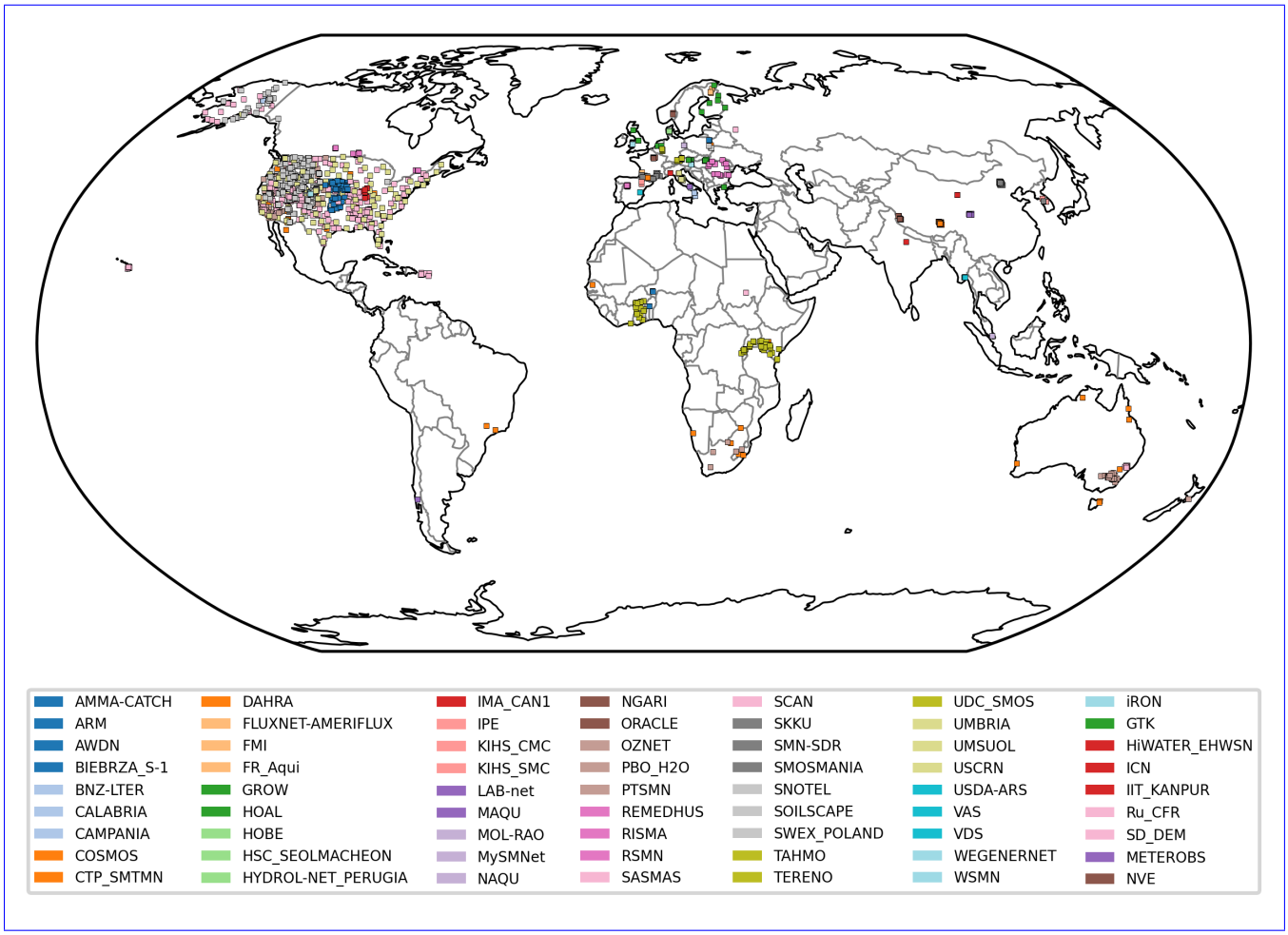

Figure A1. Location map of the ISMN in situ stations used in this study and listed in TableA1.

Author contributions. Adam Pasik: Conceptualization, Formal analysis, Investigation, Visualization, Software, Writing - original draft preparation. Alexander Gruber: Conceptualization, Methodology, Supervision, Writing - review & editing. Wolfgang Preimesberger: Data curation, Software, Validation, Writing - review & editing. Domenico De Santis: Methodology, Software, Writing - review & editing. Wouter Dorigo: Conceptualization, Methodology, Funding acquisition, Supervision, Writing - review & editing.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. This study received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement nº 870353 - Global Gravity-based Groundwater Product (G3P) project. G3P is funded in response to the Earth observation call

LC-SPACE-04-EO-2019-2020 “Copernicus evolution – Research activities in support of cross-cutting applications between Copernicus services”, as part of the H2020-SPACE-2018-2020 activity “Leadership in Industrial Technologies - Space Part”. Please visit https://g3p.eu for further information on the project.

References

Al-Yaari, A., Dayau, S., Chipeaux, C., Aluome, C., Kruszewski, A., Loustau, D., and Wigneron, J.-P.: The AQUI Soil Moisture Network for Satellite Microwave Remote Sensing Validation in South-Western France, Remote Sensing, 10, https://doi.org/10.3390/rs10111839, 2018.

Albergel, C., Rüdiger, C., Pellarin, T., Calvet, J.-C., Fritz, N., Froissard, F., Suquia, D., Petitpa, A., Piguet, B., and Martin, E.: From near-surface to root-zone soil moisture using an exponential filter: an assessment of the method based on in-situ observations and model simulations, Hydrology and Earth System Sciences, 12, 1323–1337, https://doi.org/10.5194/hess-12-1323-2008, 2008.

Alday, J. G., Camarero, J. J., Revilla, J., and Resco de Dios, V.: Similar diurnal, seasonal and annual rhythms in radial root expansion across two coexisting Mediterranean oak species, Tree Physiology, 40, 956–968, https://doi.org/10.1093/treephys/tpaa041, 2020.

Baldwin, D., Manfreda, S., Keller, K., and Smithwick, E.: Predicting root zone soil moisture with soil properties and satellite near-surface moisture data across the conterminous United States, Journal of Hydrology, 546, 393–404, https://doi.org/10.1016/j.jhydrol.2017.01.020, 2017.

Bauer Marschallinger, B.: Product User Manual CGLOPS1_PUM_SWIV3-SWI10-SWI-TS l2.60, Tech. rep., Copernicus Global Land Operations, 2018.

Bauer Marschallinger, B.: Algorithm Theoretical Basis Document, CGLOPS1_ATBD_SWI1km-V1 l1.30, Tech. rep., Copernicus Global Land Operations, 2022.

Beck, H. E., de Jeu, R. A. M., Schellekens, J., van Dijk, A. I. J. M., and Bruijnzeel, L. A.: Improving Curve Number Based Storm Runoff Estimates Using Soil Moisture Proxies, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2, 250–259, https://doi.org/10.1109/JSTARS.2009.2031227, 2009.

Bell, J. E., Palecki, M. A., Baker, C. B., Collins, W. G., Lawrimore, J. H., Leeper, R. D., Hall, M. E., Kochendorfer, J., Meyers, T. P., Wilson, T., and Diamond, H. J.: U.S. Climate Reference Network Soil Moisture and Temperature Observations, Journal of Hydrometeorology, 14, 977 – 988, https://doi.org/10.1175/JHM-D-12-0146.1, 2013.

Beven, K.: On the concept of model structural error, Water Science and Technology, 52, 167–175, https://doi.org/10.2166/wst.2005.0165, 2005.

Beyrich, F. and Adam, W.: Site and Data Report for the Lindenberg Reference Site in CEOP - Phase 1, Tech. Rep. 230, Deutscher Wetterdienst, Offenbach am Main, 2007.

Biddoccu, M., Ferraris, S., Opsi, F., and Cavallo, E.: Long-term monitoring of soil management effects on runoff and soil erosion in sloping vineyards in Alto Monferrato (North–West Italy), Soil and Tillage Research, 155, 176–189, https://doi.org/10.1016/j.still.2015.07.005, 2016.

Bircher, S., Skou, N., Jensen, K. H., Walker, J. P., and Rasmussen, L.: A soil moisture and temperature network for SMOS validation in Western Denmark, Hydrology and Earth System Sciences, 16, 1445–1463, https://doi.org/10.5194/hess-16-1445-2012, 2012.

Blöschl, G., Blaschke, A. P., Broer, M., Bucher, C., Carr, G., Chen, X., Eder, A., Exner-Kittridge, M., Farnleitner, A., Flores-Orozco, A., Haas, P., Hogan, P., Kazemi Amiri, A., Oismüller, M., Parajka, J., Silasari, R., Stadler, P., Strauss, P., Vreugdenhil, M., Wagner, W., and Zessner, M.: The Hydrological Open Air Laboratory (HOAL) in Petzenkirchen: a hypothesis-driven observatory, Hydrology and Earth System Sciences, 20, 227–255, https://doi.org/10.5194/hess-20-227-2016, 2016.

Bogena, H., Kunkel, R., Puetz, T., Vereecken, H., Krueger, E., Zacharias, S., Dietrich, P., Wollschlaeger, U., Kunstmann, H., Papen, H., Schmid, H. P., Munch, J. C., Priesack, E., Schwank, M., Bens, O., Brauer, A., Borg, E., and Hajnsek, I.: TERENO - Long-term monitoring network for terrestrial environmental research, Hydrologie und Wasserbewirtschaftung, 56, 138–143, 12.02.03; LK 01, 2012.

Bogena, H., Montzka, C., Huisman, J., Graf, A., Schmidt, M., Stockinger, M., von Hebel, C., Hendricks-Franssen, H., van der Kruk, J., Tappe, W., Lücke, A., Baatz, R., Bol, R., Groh, J., Pütz, T., Jakobi, J., Kunkel, R., Sorg, J., and Vereecken, H.: The TERENO-Rur Hydrological Observatory: A Multiscale Multi-Compartment Research Platform for the Advancement of Hydrological Science, Vadose Zone Journal, 17, 180 055, https://doi.org/10.2136/vzj2018.03.0055, 2018.

Bogena, H. R.: TERENO: German network of terrestrial environmental observatories, Journal of large-scale research facilities, 2, https://doi.org/10.17815/jlsrf-2-98, 2016.

Bouaziz, L. J. E., Steele-Dunne, S. C., Schellekens, J., Weerts, A. H., Stam, J., Sprokkereef, E., Winsemius, H. H. C., Savenije, H. H. G., and Hrachowitz, M.: Improved Understanding of the Link Between Catchment-Scale Vegetation Accessible Storage and Satellite-Derived Soil Water Index, Water Resources Research, 56, e2019WR026 365, https://doi.org/10.1029/2019WR026365, 2020.

Brocca, L., Melone, F., and Moramarco, T.: On the estimation of antecedent wetness conditions in rainfall–runoff modelling, Hydrological Processes, 22, 629–642, https://doi.org/10.1002/hyp.6629, 2008.

Brocca, L., Melone, F., Moramarco, T., and Morbidelli, R.: Antecedent wetness conditions based on ERS scatterometer data, Journal of Hydrology, 364, 73–87, https://doi.org/10.1016/j.jhydrol.2008.10.007, 2009.

Brocca, L., Melone, F., Moramarco, T., Wagner, W., and Hasenauer, S.: ASCAT soil wetness index validation through in situ and modeled soil moisture data in central Italy, Remote Sensing of Environment, 114, 2745–2755, https://doi.org/10.1016/j.rse.2010.06.009, 2010.

Brocca, L., Hasenauer, S., Lacava, T., Melone, F., Moramarco, T., Wagner, W., Dorigo, W., Matgen, P., Martínez-Fernández, J., Llorens, P., Latron, J., Martin, C., and Bittelli, M.: Soil moisture estimation through ASCAT and AMSR-E sensors: An intercomparison and validation study across Europe, Remote Sensing of Environment, 115, 3390–3408, https://doi.org/10.1016/j.rse.2011.08.003, 2011.

C3S: Algorithm Theoretical Baseline Document (ATBD) - Soil Moisture Service D1.SM.2-v3.0, Tech. rep., EODC, https://doi.org/10.24381/cds.d7782f18, 2020.

Calvet, J.-C., Fritz, N., Froissard, F., Suquia, D., Petitpa, A., and Piguet, B.: In situ soil moisture observations for the CAL/VAL of SMOS: the SMOSMANIA network, in: 2007 IEEE International Geoscience and Remote Sensing Symposium, pp. 1196–1199, https://doi.org/10.1109/IGARSS.2007.4423019, 2007.

Calvet, J.-C., Fritz, N., Berne, C., Piguet, B., Maurel, W., and Meurey, C.: Deriving pedotransfer functions for soil quartz fraction in southern France from reverse modeling, SOIL, 2, 615–629, https://doi.org/10.5194/soil-2-615-2016, 2016.

Canisius, F.: Calibration of Casselman, Ontario Soil Moisture Monitoring Network, Tech. rep., Agriculture and Agri-food Canada, 2011.

Capello, G., Biddoccu, M., Ferraris, S., and Cavallo, E.: Effects of Tractor Passes on Hydrological and Soil Erosion Processes in Tilled and Grassed Vineyards, Water, 11, https://doi.org/10.3390/w11102118, 2019.

Cappelaere, B., Descroix, L., Lebel, T., Boulain, N., Ramier, D., Laurent, J.-P., Favreau, G., Boubkraoui, S., Boucher, M., Bouzou Moussa, I., Chaffard, V., Hiernaux, P., Issoufou, H., Le Breton, E., Mamadou, I., Nazoumou, Y., Oi, M., Ottlé, C., and Quantin, G.: The AMMA-CATCH experiment in the cultivated Sahelian area of south-west Niger – Investigating water cycle response to a fluctuating climate and changing environment, Journal of Hydrology, 375, 34–51, https://doi.org/10.1016/j.jhydrol.2009.06.021, 2009.

Cook, D. R.: Soil Water and Temperature System (SWATS) Instrument Handbook, Tech. rep., US Department of Energy, https://doi.org/10.2172/1251383, 2016a.

Cook, D. R.: Soil Temperature and Moisture Profile (STAMP) System Handbook, Tech. rep., US Department of Energy, https://doi.org/10.2172/1332724, 2016b.

Cook, D. R.: Surface Energy Balance System (SEBS) Instrument Handbook, Tech. rep., US Department of Energy, https://doi.org/10.2172/1004944, 2018.

de Lange, R., Beck, R., van de Giesen, N., Friesen, J., de Wit, A., and Wagner, W.: Scatterometer-Derived Soil Moisture Calibrated for Soil Texture With a One-Dimensional Water-Flow Model, IEEE Transactions on Geoscience and Remote Sensing, 46, 4041–4049, https://doi.org/10.1109/TGRS.2008.2000796, 2008.

de Rosnay, P., Gruhier, C., Timouk, F., Baup, F., Mougin, E., Hiernaux, P., Kergoat, L., and LeDantec, V.: Multi-scale soil moisture measurements at the Gourma meso-scale site in Mali, Journal of Hydrology, 375, 241–252, https://doi.org/10.1016/j.jhydrol.2009.01.015, surface processes and water cycle in West Africa, studied from the AMMA-CATCH observing system, 2009.

De Santis, D. and Biondi, D.: Error Propagation from Remotely Sensed Surface Soil Moisture Into Soil Water Index Using an Exponential Filter, in: HIC 2018. 13th International Conference on Hydroinformatics, edited by Loggia, G. L., Freni, G., Puleo, V., and Marchis, M. D., vol. 3 of *EPiC Series in Engineering*, pp. 520–525, EasyChair, https://doi.org/10.29007/kvhb, 2018.

Dente, L., Su, Z., and Wen, J.: Validation of SMOS Soil Moisture Products over the Maqu and Twente Regions, Sensors, 12, 9965–9986, https://doi.org/10.3390/s120809965, 2012.

Dorigo, W., Xaver, A., Vreugdenhil, M., Gruber, A., Hegyiová, A., Sanchis-Dufau, A., 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 Journal, 12, vzj2012.0097, https://doi.org/10.2136/vzj2012.0097, 2013.

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 Sensing of Environment, 203, 185–215, https://doi.org/https://doi.org/10.1016/j.rse.2017.07.001, 2017.

Dorigo, W., Dietrich, S., Aires, F., Brocca, L., Carter, S., Cretaux, J.-F., Dunkerley, D., Enomoto, H., Forsberg, R., Güntner, A., Hegglin, M. I., Hollmann, R., Hurst, D. F., Johannessen, J. A., Kummerow, C., Lee, T., Luojus, K., Looser, U., Miralles, D. G., Pellet, V., Recknagel, T., Vargas, C. R., Schneider, U., Schoeneich, P., Schröder, M., Tapper, N., Vuglinsky, V., Wagner, W., Yu, L., Zappa, L., Zemp, M., and Aich, V.: Closing the Water Cycle from Observations across Scales: Where Do We Stand?, Bulletin of the American Meteorological Society, 102, E1897 – E1935, https://doi.org/10.1175/BAMS-D-19-0316.1, 2021a.

Dorigo, W., Himmelbauer, I., Aberer, D., Schremmer, L., Petrakovic, I., Zappa, L., Preimesberger, W., Xaver, A., Annor, F., Ardö, J., Baldocchi, D., Bitelli, M., Blöschl, G., Bogena, H., Brocca, L., Calvet, J.-C., Camarero, J. J., Capello, G., Choi, M., Cosh, M. C., van de Giesen, N., Hajdu, I., Ikonen, J., Jensen, K. H., Kanniah, K. D., de Kat, I., Kirchengast, G., Kumar Rai, P., Kyrouac, J., Larson, K., Liu, S., Loew, A., Moghaddam, M., Martínez Fernández, J., Mattar Bader, C., Morbidelli, R., Musial, J. P., Osenga, E., Palecki, M. A., Pellarin, T., Petropoulos, G. P., Pfeil, I., Powers, J., Robock, A., Rüdiger, C., Rummel, U., Strobel, M., Su, Z., Sullivan, R., Tagesson, T., Varlagin, A., Vreugdenhil, M., Walker, J., Wen, J., Wenger, F., Wigneron, J. P., Woods, M., Yang, K., Zeng, Y., Zhang, X., Zreda, M., Dietrich, S., Gruber, A., van Oevelen, P., Wagner, W., Scipal, K., Drusch, M., and Sabia, R.: The International Soil Moisture Network: serving Earth system science for over a decade, Hydrology and Earth System Sciences, 25, 5749–5804, https://doi.org/10.5194/hess-25-5749-2021, 2021b.

Dorigo, W., Preimesberger, W., Moesinger, L., Pasik, A., Scanlon, T., Hahn, S., Van der Schalie, R., Van der Vliet, M., De Jeu, R., Kidd, R., Rodriguez-Fernandez, N., and Hirschi, M.: ESA Soil Moisture Climate Change Initiative: COMBINED Product, Version 05.3, Centre for Environmental Data Analysis, https://catalogue.ceda.ac.uk/uuid/e43aead9947549078c2d108b2c3632b2, 2021c.

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, Hydrology and Earth System Sciences, 15, 1675–1698, https://doi.org/10.5194/hess-15-1675-2011, 2011.

Famiglietti, J. S., Ryu, D., Berg, A. A., Rodell, M., and Jackson, T. J.: Field observations of soil moisture variability across scales, Water Resources Research, 44, https://doi.org/10.1029/2006WR005804, 2008.

Flammini, A., Corradini, C., Morbidelli, R., Saltalippi, C., Picciafuoco, T., and Giráldez, J. V.: Experimental Analyses of the Evaporation Dynamics in Bare Soils under Natural Conditions, Water Resources Management, 32, 1153–1166, https://doi.org/10.1007/s11269-017-1860-x, 2018a.

Flammini, A., Morbidelli, R., Saltalippi, C., Picciafuoco, T., Corradini, C., and Govindaraju, R. S.: Reassessment of a semi-analytical field-scale infiltration model through experiments under natural rainfall events, Journal of Hydrology, 565, 835–845, https://doi.org/10.1016/j.jhydrol.2018.08.073, 2018b.

Ford, T. W., Harris, E., and Quiring, S. M.: Estimating root zone soil moisture using near-surface observations from SMOS, Hydrology and Earth System Sciences, 18, 139–154, https://doi.org/10.5194/hess-18-139-2014, 2014.

Fuchsberger, J., Kirchengast, G., and Kabas, T.: WegenerNet high-resolution weather and climate data from 2007 to 2020, Earth System Science Data, 13, 1307–1334, https://doi.org/10.5194/essd-13-1307-2021, 2021.

Galle, S., Grippa, M., Peugeot, C., Bouzou Moussa, I., Cappelaere, B., Demarty, J., Mougin, E., Lebel, T., and Chaffard, V.: AMMA-CATCH a Hydrological, Meteorological and Ecological Long Term Observatory on West Africa : Some Recent Results, in: AGU Fall Meeting Abstracts, vol. 2015, pp. GC42A–01, 2015.

GCOS: The Global Observing System for Climate: Implementation needs, World Meteorological Organization, 214, https://public.wmo.int/en/resources/library/global-observing-system-climate-implementation-needs, 2016.

GCOS: The 2022 GCOS ECVs Requirements, World Meteorological Organisation, 245, https://public.wmo.int/en/resources/library/global-observing-system-climate-implementation-needs, 2022.

González-Zamora, A., Sánchez, N., Pablos, M., and Martínez-Fernández, J.: CCI soil moisture assessment with SMOS soil moisture and in situ data under different environmental conditions and spatial scales in Spain, Remote Sensing of Environment, 225, 469–482, https://doi.org/10.1016/j.rse.2018.02.010, 2019.

Grillakis, M. G., Koutroulis, A. G., Alexakis, D. D., Polykretis, C., and Daliakopoulos, I. N.: Regionalizing Root-Zone Soil Moisture Estimates From ESA CCI Soil Water Index Using Machine Learning and Information on Soil, Vegetation, and Climate, Water Resources Research, 57, e2020WR029 249, https://doi.org/10.1029/2020WR029249, e2020WR029249 2020WR029249, 2021.

Gruber, A., Dorigo, W., Crow, W., and Wagner, W.: Triple Collocation-Based Merging of Satellite Soil Moisture Retrievals, IEEE Transactions on Geoscience and Remote Sensing, PP, 1–13, https://doi.org/10.1109/TGRS.2017.2734070, 2017.

Gruber, A., Scanlon, T., van der Schalie, R., Wagner, W., and Dorigo, W.: Evolution of the ESA CCI Soil Moisture climate data records and their underlying merging methodology, Earth System Science Data, 11, 717–739, https://doi.org/10.5194/essd-11-717-2019, 2019.

Gruber, A., De Lannoy, G., Albergel, C., Al-Yaari, A., Brocca, L., Calvet, J.-C., Colliander, A., Cosh, M., Crow, W., Dorigo, W., Draper, C., Hirschi, M., Kerr, Y., Konings, A., Lahoz, W., McColl, K., Montzka, C., Muñoz-Sabater, J., Peng, J., Reichle, R., Richaume, P., Rüdiger,

C., Scanlon, T., van der Schalie, R., Wigneron, J.-P., and Wagner, W.: Validation practices for satellite soil moisture retrievals: What are (the) errors?, Remote Sensing of Environment, 244, 111 806, https://doi.org/10.1016/j.rse.2020.111806, 2020.

Gupta, H. V., Kling, H., Yilmaz, K. K., and Martinez, G. F.: Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling, Journal of Hydrology, 377, 80–91, https://doi.org/10.1016/j.jhydrol.2009.08.003, 2009.

Hajdu, I., Yule, I., Bretherton, M., Singh, R., and Hedley, C.: Field performance assessment and calibration of multi-depth AquaCheck capacitance-based soil moisture probes under permanent pasture for hill country soils, Agricultural Water Management, 217, 332–345, https://doi.org/10.1016/j.agwat.2019.03.002, 2019.

Hollinger, S. E. and Isard, S. A.: A Soil Moisture Climatology of Illinois, Journal of Climate, 7, 822 – 833, https://doi.org/10.1175/1520-0442(1994)007<0822:ASMCOI>2.0.CO;2, 1994.

Ikonen, J., Vehviläinen, J., Rautiainen, K., Smolander, T., Lemmetyinen, J., Bircher, S., and Pulliainen, J.: The Sodankylä in situ soil moisture observation network: an example application of ESA CCI soil moisture product evaluation, Geoscientific Instrumentation, Methods and Data Systems, 5, 95–108, https://doi.org/10.5194/gi-5-95-2016, 2016.

Ikonen, J., Smolander, T., Rautiainen, K., Cohen, J., Lemmetyinen, J., Salminen, M., and Pulliainen, J.: Spatially Distributed Evaluation of ESA CCI Soil Moisture Products in a Northern Boreal Forest Environment, Geosciences, 8, https://doi.org/10.3390/geosciences8020051, 2018.

Jackson, T. J., Cosh, M. H., Bindlish, R., Starks, P. J., Bosch, D. D., Seyfried, M., Goodrich, D. C., Moran, M. S., and Du, J.: Validation of Advanced Microwave Scanning Radiometer Soil Moisture Products, IEEE Transactions on Geoscience and Remote Sensing, 48, 4256–4272, https://doi.org/10.1109/TGRS.2010.2051035, 2010.

Jensen, K. H. and Refsgaard, J. C.: HOBE: The Danish Hydrological Observatory, Vadose Zone Journal, 17, 180 059, https://doi.org/10.2136/vzj2018.03.0059, 2018.

Jin, R., Li, X., Yan, B., Li, X., Luo, W., Ma, M., Guo, J., Kang, J., Zhu, Z., and Zhao, S.: A Nested Ecohydrological Wireless Sensor Network for Capturing the Surface Heterogeneity in the Midstream Areas of the Heihe River Basin, China, IEEE Geoscience and Remote Sensing Letters, 11, 2015–2019, https://doi.org/10.1109/LGRS.2014.2319085, 2014.

Kang, C. S., Kanniah, K. D., and Kerr, Y. H.: Calibration of SMOS Soil Moisture Retrieval Algorithm: A Case of Tropical Site in Malaysia, IEEE Transactions on Geoscience and Remote Sensing, 57, 3827–3839, https://doi.org/10.1109/TGRS.2018.2888535, 2019.

Kang, J., Li, X., Jin, R., Ge, Y., Wang, J., and Wang, J.: Hybrid Optimal Design of the Eco-Hydrological Wireless Sensor Network in the Middle Reach of the Heihe River Basin, China, Sensors, 14, 19 095–19 114, https://doi.org/10.3390/s141019095, 2014.

Kirchengast, G., Kabas, T., Leuprecht, A., Bichler, C., and Truhetz, H.: WegenerNet: A Pioneering High-Resolution Network for Monitoring Weather and Climate, Bulletin of the American Meteorological Society, 95, 227 – 242, https://doi.org/10.1175/BAMS-D-11-00161.1, 2014.

Larson, K. M., Small, E. E., Gutmann, E. D., Bilich, A. L., Braun, J. J., and Zavorotny, V. U.: Use of GPS receivers as a soil moisture network for water cycle studies, Geophysical Research Letters, 35, https://doi.org/10.1029/2008GL036013, 2008.

Leavesley, G., David, O., Garen, D., Lea, J., Marron, J., Pagano, T., Perkins, T., and Strobel, M.: A modeling framework for improved agricultural water supply forecasting, in: AGU Fall Meeting Abstracts, vol. 1, San Francisco, CA, USA, 2008.

Lebel, T., Cappelaere, B., Galle, S., Hanan, N., Kergoat, L., Levis, S., Vieux, B., Descroix, L., Gosset, M., Mougin, E., Peugeot, C., and Seguis, L.: AMMA-CATCH studies in the Sahelian region of West-Africa: An overview, Journal of Hydrology, 375, 3–13, https://doi.org/10.1016/j.jhydrol.2009.03.020, surface processes and water cycle in West Africa, studied from the AMMA-CATCH observing system, 2009.

Leys, C., Ley, C., Klein, O., Bernard, P., and Licata, L.: Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median, Journal of Experimental Social Psychology, 49, 764–766, https://doi.org/10.1016/j.jesp.2013.03.013, 2013.

L'Heureux, J.: 2011 Installation Report for AAFC- SAGES Soil Moisture Stations in Kenaston, SK, Tech. rep., Agriculture and Agri-food Canada, 2011.

Loew, A., Dall'Amico, J. T., Schlenz, F., and Mauser, W.: The Upper Danube soil moisture validation site: measurements and activities, The EGU General Assembly, 674, 2924, 2009.

Mahmood, R. and Hubbard, K. G.: Relationship between soil moisture of near surface and multiple depths of the root zone under heterogeneous land uses and varying hydroclimatic conditions, Hydrological Processes, 21, 3449–3462, https://doi.org/10.1002/hyp.6578, 2007.

Manfreda, S., Brocca, L., Moramarco, T., Melone, F., and Sheffield, J.: A physically based approach for the estimation of root-zone soil moisture from surface measurements, Hydrology and Earth System Sciences, 18, 1199–1212, https://doi.org/10.5194/hess-18-1199-2014, 2014.

Marczewski, W., Slominski, J., Slominska, E., Usowicz, B., Usowicz, J., Romanov, S., Maryskevych, O., Nastula, J., and Zawadzki, J.: Strategies for validating and directions for employing SMOS data, in the Cal-Val project SWEX (3275) for wetlands, Hydrology and Earth System Sciences Discussions, 7, 7007–7057, https://doi.org/10.5194/hessd-7-7007-2010, 2010.

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, Geoscientific Model Development, 10, 1903–1925, https://doi.org/10.5194/gmd-10-1903-2017, 2017.

Mattar, C., Santamaría-Artigas, A., Durán-Alarcón, C., Olivera-Guerra, L., Fuster, R., and Borvarán, D.: LAB-net the first Chilean soil moisture network for remote sensing applications, in: Quantitative Remote Sensing Symposium (RAQRS), pp. 22–26, 2014.

Mattar, C., Santamaría-Artigas, A., Durán-Alarcón, C., Olivera-Guerra, L., Fuster, R., and Borvarán, D.: The LAB-Net Soil Moisture Network: Application to Thermal Remote Sensing and Surface Energy Balance, Data, 1, https://doi.org/10.3390/data1010006, 2016.

Mishra, V., Ellenburg, W. L., Markert, K. N., and Limaye, A. S.: Performance evaluation of soil moisture profile estimation through entropy-based and exponential filter models, Hydrological Sciences Journal, 65, 1036–1048, https://doi.org/10.1080/02626667.2020.1730846, 2020.

Moghaddam, M., Entekhabi, D., Goykhman, Y., Li, K., Liu, M., Mahajan, A., Nayyar, A., Shuman, D., and Teneketzis, D.: A Wireless Soil Moisture Smart Sensor Web Using Physics-Based Optimal Control: Concept and Initial Demonstrations, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 3, 522–535, https://doi.org/10.1109/JSTARS.2010.2052918, 2010.

Moghaddam, M., Silva, A., Clewley, D., Akbar, R., Hussaini, S., Whitcomb, J., Devarakonda, R., Shrestha, R., Cook, R., Prakash, G., Santhana Vannan, S., and Boyer, A.: Soil Moisture Profiles and Temperature Data from SoilSCAPE Sites, USA, https://doi.org/10.3334/ORNLDAAC/1339, 2016.

Morbidelli, R., Corradini, C., Saltalippi, C., Flammini, A., and Rossi, E.: Infiltration-soil moisture redistribution under natural conditions: experimental evidence as a guideline for realizing simulation models, Hydrology and Earth System Sciences, 15, 2937–2945, https://doi.org/10.5194/hess-15-2937-2011, 2011.

Morbidelli, R., Saltalippi, C., Flammini, A., Rossi, E., and Corradini, C.: Soil water content vertical profiles under natural conditions: matching of experiments and simulations by a conceptual model, Hydrological Processes, 28, 4732–4742, https://doi.org/10.1002/hyp.9973, 2014.

Morbidelli, R., Saltalippi, C., Flammini, A., Cifrodelli, M., Picciafuoco, T., Corradini, C., and Govindaraju, R. S.: In situ measurements of soil saturated hydraulic conductivity: Assessment of reliability through rainfall–runoff experiments, Hydrological Processes, 31, 3084–3094, https://doi.org/10.1002/hyp.11247, 2017.

Mougin, E., Hiernaux, P., Kergoat, L., Grippa, M., de Rosnay, P., Timouk, F., Le Dantec, V., Demarez, V., Lavenu, F., Arjounin, M., Lebel, T., Soumaguel, N., Ceschia, E., Mougenot, B., Baup, F., Frappart, F., Frison, P., Gardelle, J., Gruhier, C., Jarlan, L., Mangiarotti, S., Sanou, B., Tracol, Y., Guichard, F., Trichon, V., Diarra, L., Soumaré, A., Koité, M., Dembélé, F., Lloyd, C., Hanan, N., Damesin, C., Delon, C., Serça, D., Galy-Lacaux, C., Seghieri, J., Becerra, S., Dia, H., Gangneron, F., and Mazzega, P.: The AMMA-CATCH Gourma observatory site in Mali: Relating climatic variations to changes in vegetation, surface hydrology, fluxes and natural resources, Journal of Hydrology, 375, 14–33, https://doi.org/10.1016/j.jhydrol.2009.06.045, 2009.

Muñoz Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., Boussetta, S., Choulga, M., Harrigan, S., Hersbach, H., Martens, B., Miralles, D. G., Piles, M., Rodríguez-Fernández, N. J., Zsoter, E., Buontempo, C., and Thépaut, J.-N.: ERA5-Land: a state-of-the-art global reanalysis dataset for land applications, Earth System Science Data, 13, 2021.

Musial, J., Dabrowska-Zielinska, K., Kiryla, W., Oleszczuk, R., Gnatowski, T., and Jaszczynski, J.: Derivation and validation of the high resolution satellite soil moisture products: a case study of the Biebrza Sentinel-1 validation sites, online, 2016.

Muñoz Sabater, J.: ERA5-Land hourly data from 1981 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS), https://doi.org/10.24381/cds.e2161bac, 2019.

Muñoz Sabater, J.: ERA5-Land hourly data from 1950 to 1980. Copernicus Climate Change Service (C3S) Climate Data Store (CDS), https://doi.org/10.24381/cds.e2161bac, 2021.

Nguyen, H. H., Kim, H., and Choi, M.: Evaluation of the soil water content using cosmic-ray neutron probe in a heterogeneous monsoon climate-dominated region, Advances in Water Resources, 108, 125–138, https://doi.org/10.1016/j.advwatres.2017.07.020, 2017.

Ojo, E. R., Bullock, P. R., L'Heureux, J., Powers, J., McNairn, H., and Pacheco, A.: Calibration and Evaluation of a Frequency Domain Reflectometry Sensor for Real-Time Soil Moisture Monitoring, Vadose Zone Journal, 14, vzj2014.08.0114, https://doi.org/10.2136/vzj2014.08.0114, 2015.

Osenga, E. C., Arnott, J. C., Endsley, K. A., and Katzenberger, J. W.: Bioclimatic and Soil Moisture Monitoring Across Elevation in a Mountain Watershed: Opportunities for Research and Resource Management, Water Resources Research, 55, 2493–2503, https://doi.org/10.1029/2018WR023653, 2019.

Osenga, E. C., Vano, J. A., and Arnott, J. C.: A community-supported weather and soil moisture monitoring database of the Roaring Fork catchment of the Colorado River Headwaters, Hydrological Processes, 35, e14 081, https://doi.org/10.1002/hyp.14081, 2021.

Parinussa, R. M., Meesters, A. G. C. A., Liu, Y. Y., Dorigo, W., Wagner, W., and de Jeu, R. A. M.: Error Estimates for Near-Real-Time Satellite Soil Moisture as Derived From the Land Parameter Retrieval Model, IEEE Geoscience and Remote Sensing Letters, 8, 779–783, https://doi.org/10.1109/LGRS.2011.2114872, 2011.

Pathe, C., Wagner, W., Sabel, D., Doubkova, M., and Basara, J.: Using ENVISAT ASAR global mode data for surface soil moisture retrieval over oklahoma, usa, IEEE Trans. Geosci. Rem. Sens., 47, 468–480, https://doi.org/10.1109/TGRS.2008.2004711, 2009.

Paulik, C., Dorigo, W., Wagner, W., and Kidd, R.: Validation of the ASCAT Soil Water Index using in situ data from the International Soil Moisture Network, International Journal of Applied Earth Observation and Geoinformation, 30, 1–8, https://doi.org/10.1016/j.jag.2014.01.007, 2014.

Pellarin, T., Calvet, J.-C., and Wagner, W.: Evaluation of ERS scatterometer soil moisture products over a half-degree region in southwestern France, Geophysical Research Letters, 33, https://doi.org/10.1029/2006GL027231, 2006.

Petropoulos, G. P. and McCalmont, J. P.: An Operational In Situ Soil Moisture & Soil Temperature Monitoring Network for West Wales, UK: The WSMN Network, Sensors, 17, https://doi.org/10.3390/s17071481, 2017.

Preimesberger, W., Scanlon, T., Su, C.-H., Gruber, A., and Dorigo, W.: Homogenization of Structural Breaks in the Global ESA CCI Soil Moisture Multisatellite Climate Data Record, IEEE Transactions on Geoscience and Remote Sensing, 59, 2845–2862, https://doi.org/10.1109/TGRS.2020.3012896, 2021.

Raffelli, G., Previati, M., Canone, D., Gisolo, D., Bevilacqua, I., Capello, G., Biddoccu, M., Cavallo, E., Deiana, R., Cassiani, G., and Ferraris, S.: Local- and Plot-Scale Measurements of Soil Moisture: Time and Spatially Resolved Field Techniques in Plain, Hill and Mountain Sites, Water, 9, https://doi.org/10.3390/w9090706, 2017.

Reichle, R., DeLannoy, G., Koster, R. D., Crow, W. T., and Kimball, J.: SMAP L4 9 km EASE-Grid Surface and Root Zone Soil Moisture Geophysical Data, Version 3, Tech. rep., TU Vienna, https://doi.org/10.5067/B59DT1D5UMB4, 2017a.

Reichle, R. H., Lannoy, G. J. M. D., Liu, Q., Ardizzone, J. V., Colliander, A., Conaty, A., Crow, W., Jackson, T. J., Jones, L. A., Kimball, J. S., Koster, R. D., Mahanama, S. P., Smith, E. B., Berg, A., Bircher, S., Bosch, D., Caldwell, T. G., Cosh, M., Ángel González-Zamora, Collins, C. D. H., Jensen, K. H., Livingston, S., Lopez-Baeza, E., Martínez-Fernández, J., McNairn, H., Moghaddam, M., Pacheco, A., Pellarin, T., Prueger, J., Rowlandson, T., Seyfried, M., Starks, P., Su, Z., Thibeault, M., van der Velde, R., Walker, J., Wu, X., and Zeng, Y.: Assessment of the SMAP Level-4 Surface and Root-Zone Soil Moisture Product Using In Situ Measurements, Journal of Hydrometeorology, 18, 2621 – 2645", https://doi.org/doi.org/10.1175/JHM-D-17-0063.1, 2017b.

Rodell, M., Houser, P. R., Jambor, U., Gottschalck, J., Mitchell, K., Meng, C.-J., Arsenault, K., Cosgrove, B., Radakovich, J., Bosilovich, M., Entin, J. K., Walker, J. P., Lohmann, D., and Toll, D.: The Global Land Data Assimilation System, Bulletin of the American Meteorological Society, 85, 381 – 394, https://doi.org/10.1175/BAMS-85-3-381, 2004.

Rüdiger, C., Hancock, G., Hemakumara, H. M., Jacobs, B., Kalma, J. D., Martinez, C., Thyer, M., Walker, J. P., Wells, T., and Willgoose, G. R.: Goulburn River experimental catchment data set, Water Resources Research, 43, https://doi.org/10.1029/2006WR005837, 2007.

Schaefer, G. L., Cosh, M. H., and Jackson, T. J.: The USDA Natural Resources Conservation Service Soil Climate Analysis Network (SCAN), Journal of Atmospheric and Oceanic Technology, 24, 2073 – 2077, https://doi.org/10.1175/2007JTECHA930.1, 2007.

Schlenz, F., dall'Amico, J. T., Loew, A., and Mauser, W.: Uncertainty Assessment of the SMOS Validation in the Upper Danube Catchment, IEEE Transactions on Geoscience and Remote Sensing, 50, 1517–1529, https://doi.org/10.1109/TGRS.2011.2171694, 2012.

Shuman, D. I., Nayyar, A., Mahajan, A., Goykhman, Y., Li, K., Liu, M., Teneketzis, D., Moghaddam, M., and Entekhabi, D.: Measurement Scheduling for Soil Moisture Sensing: From Physical Models to Optimal Control, Proceedings of the IEEE, 98, 1918–1933, https://doi.org/10.1109/JPROC.2010.2052532, 2010.

Smith, A. B., Walker, J. P., Western, A. W., Young, R. I., Ellett, K. M., Pipunic, R. C., Grayson, R. B., Siriwardena, L., Chiew, F. H. S., and Richter, H.: The Murrumbidgee soil moisture monitoring network data set, Water Resources Research, 48, https://doi.org/10.1029/2012WR011976, 2012.

Stefan, V.-G., Indrio, G., Escorihuela, M.-J., Quintana-Seguí, P., and Villar, J. M.: High-Resolution SMAP-Derived Root-Zone Soil Moisture Using an Exponential Filter Model Calibrated per Land Cover Type, Remote Sensing, 13, https://doi.org/10.3390/rs13061112, 2021.

Su, Z., Wen, J., Dente, L., van der Velde, R., Wang, L., Ma, Y., Yang, K., and Hu, Z.: The Tibetan Plateau observatory of plateau scale soil moisture and soil temperature (Tibet-Obs) for quantifying uncertainties in coarse resolution satellite and model products, Hydrology and Earth System Sciences, 15, 2303–2316, https://doi.org/10.5194/hess-15-2303-2011, 2011.

Sure, A. and Dikshit, O.: Estimation of root zone soil moisture using passive microwave remote sensing: A case study for rice and wheat crops for three states in the Indo-Gangetic basin, Journal of Environmental Management, 234, 75–89, https://doi.org/10.1016/j.jenvman.2018.12.109, 2019.

Tagesson, T., Fensholt, R., Guiro, I., Rasmussen, M. O., Huber, S., Mbow, C., Garcia, M., Horion, S., Sandholt, I., Holm-Rasmussen, B., Göttsche, F. M., Ridler, M.-E., Olén, N., Lundegard Olsen, J., Ehammer, A., Madsen, M., Olesen, F. S., and Ardö, J.: Ecosystem properties of semiarid savanna grassland in West Africa and its relationship with environmental variability, Global Change Biology, 21, 250–264, https://doi.org/10.1111/gcb.12734, 2015.

Taylor, J. R.: An Introduction to Error Analysis: The Study of Uncertainties in Physical Measurements., 2nd ed. Sausalito: University Science Books., 1997.

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, Hydrology and Earth System Sciences, 21, 4403–4417, https://doi.org/10.5194/hess-21-4403-2017, 2017.

Van Cleve, K., Chapin, F., Stuart, R., and W., R.: Bonanza Creek Long Term Ecological Research Project Climate Database, online, https://www.lter.uaf.edu/, 2015.

van der Schalie, R., de Jeu, R., Kerr, Y., Wigneron, J., Rodríguez-Fernández, N., Al-Yaari, A., Parinussa, R., Mecklenburg, S., and Drusch, M.: The merging of radiative transfer based surface soil moisture data from SMOS and AMSR-E, Remote Sensing of Environment, 189, 180–193, https://doi.org/https://doi.org/10.1016/j.rse.2016.11.026, 2017.

Vreugdenhil, M., Dorigo, W., Broer, M., Haas, P., Eder, A., Hogan, P., Bloeschl, G., and Wagner, W.: Towards a high-density soil moisture network for the validation of SMAP in Petzenkirchen, Austria, in: 2013 IEEE International Geoscience and Remote Sensing Symposium - IGARSS, pp. 1865–1868, https://doi.org/10.1109/IGARSS.2013.6723166, 2013.

Vreugdenhil, M., Greimeister-Pfeil, I., Preimesberger, W., Camici, S., Dorigo, W., Enenkel, M., van der Schalie, R., Steele-Dunne, S., and Wagner, W.: Microwave remote sensing for agricultural drought monitoring: Recent developments and challenges, Frontiers in Water, 4, https://doi.org/10.3389/frwa.2022.1045451, 2022.

Wagner, W., Lemoine, G., and Rott, H.: A Method for Estimating Soil Moisture from ERS Scatterometer and Soil Data, Remote Sensing of Environment, 70, 191–207, https://doi.org/10.1016/S0034-4257(99)00036-X, 1999.

Wang, T., Franz, T. E., You, J., Shulski, M. D., and Ray, C.: Evaluating controls of soil properties and climatic conditions on the use of an exponential filter for converting near surface to root zone soil moisture contents, Journal of Hydrology, 548, 683–696, https://doi.org/10.1016/j.jhydrol.2017.03.055, 2017.

Wigneron, J.-P., Dayan, S., Kruszewski, A., Aluome, C., AI-Yaari, M. G.-E. A., Fan, L., Guven, S., Chipeaux, C., Moisy, C., Guyon, D., and Loustau, D.: The Aqui Network: Soil Moisture Sites in the "Les Landes" Forest and Graves Vineyards (Bordeaux Aquitaine Region, France), in: IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium, pp. 3739–3742, https://doi.org/10.1109/IGARSS.2018.8517392, 2018.

Xaver, A., Zappa, L., Rab, G., Pfeil, I., Vreugdenhil, M., Hemment, D., and Dorigo, W. A.: Evaluating the suitability of the consumer low-cost Parrot Flower Power soil moisture sensor for scientific environmental applications, Geoscientific Instrumentation, Methods and Data Systems, 9, 117–139, https://doi.org/10.5194/gi-9-117-2020, 2020.

Xu, L., Chen, N., Zhang, X., Moradkhani, H., Zhang, C., and Hu, C.: In-situ and triple-collocation based evaluations of eight global root zone soil moisture products, Remote Sensing of Environment, 254, 112 248, https://doi.org/https://doi.org/10.1016/j.rse.2020.112248, 2021.

Yang, K., Qin, J., Zhao, L., Chen, Y., Tang, W., Han, M., Lazhu, Chen, Z., Lv, N., Ding, B., Wu, H., and Lin, C.: A Multiscale Soil Moisture and Freeze–Thaw Monitoring Network on the Third Pole, Bulletin of the American Meteorological Society, 94, 1907 – 1916, https://doi.org/10.1175/BAMS-D-12-00203.1, 2013.

Yang, Y., Bao, Z., Wu, H., Wang, G., Liu, C., Wang, J., and Zhang, J.: An Exponential Filter Model-Based Root-Zone Soil Moisture Estimation Methodology from Multiple Datasets, Remote Sensing, 14, https://doi.org/10.3390/rs14081785, 2022.

Young, R., Walker, J., Yeoh, N., Smith, A.and Ellett, K., Merlin, O., and Western, A.: Soil moisture and meteorological observations from the murrumbidgee catchment, Tech. rep., Department of Civil and Environmental Engineering, The University of Melbourne, 2008.

Zacharias, S., Bogena, H., Samaniego, L., Mauder, M., Fuß, R., Pütz, T., Frenzel, M., Schwank, M., Baessler, C., Butterbach-Bahl, K., Bens, O., Borg, E., Brauer, A., Dietrich, P., Hajnsek, I., Helle, G., Kiese, R., Kunstmann, H., Klotz, S., Munch, J. C., Papen, H., Priesack, E., Schmid, H. P., Steinbrecher, R., Rosenbaum, U., Teutsch, G., and Vereecken, H.: A Network of Terrestrial Environmental Observatories in Germany, Vadose Zone Journal, 10, 955–973, https://doi.org/10.2136/vzj2010.0139, 2011.

Zappa, L., Forkel, M., Xaver, A., and Dorigo, W.: Deriving Field Scale Soil Moisture from Satellite Observations and Ground Measurements in a Hilly Agricultural Region, Remote Sensing, 11, https://doi.org/10.3390/rs11222596, 2019.

Zappa, L., Woods, M., Hemment, D., Xaver, A., and Dorigo, W.: Evaluation of remotely sensed soil moisture products using crowdsourced measurements, in: Eighth International Conference on Remote Sensing and Geoinformation of the Environment (RSCy2020), edited by Themistocleous, K., Papadavid, G., Michaelides, S., Ambrosia, V., and Hadjimitsis, D. G., vol. 11524, p. 115241U, International Society for Optics and Photonics, SPIE, https://doi.org/10.1117/12.2571913, 2020.

Zhao, T., Shi, J., Lv, L., Xu, H., Chen, D., Cui, Q., Jackson, T. J., Yan, G., Jia, L., Chen, L., Zhao, K., Zheng, X., Zhao, L., Zheng, C., Ji, D., Xiong, C., Wang, T., Li, R., Pan, J., Wen, J., Yu, C., Zheng, Y., Jiang, L., Chai, L., Lu, H., Yao, P., Ma, J., Lv, H., Wu, J., Zhao, W., Yang, N., Guo, P., Li, Y., Hu, L., Geng, D., and Zhang, Z.: Soil moisture experiment in the Luan River supporting new satellite mission opportunities, Remote Sensing of Environment, 240, 111 680, https://doi.org/10.1016/j.rse.2020.111680, 2020.

Zheng, J., Zhao, T., Lü, H., Shi, J., Cosh, M. H., Ji, D., Jiang, L., Cui, Q., Lu, H., Yang, K., Wigneron, J.-P., Li, X., Zhu, Y., Hu, L., Peng, Z., Zeng, Y., Wang, X., and Kang, C. S.: Assessment of 24 soil moisture datasets using a new in situ network in the Shandian River Basin of China, Remote Sensing of Environment, 271, 112 891, https://doi.org/10.1016/j.rse.2022.112891, 2022.

Zreda, M., Desilets, D., Ferré, T. P. A., and Scott, R. L.: Measuring soil moisture content non-invasively at intermediate spatial scale using cosmic-ray neutrons, Geophysical Research Letters, 35, https://doi.org/10.1029/2008GL035655, 2008.

Zreda, M., Shuttleworth, W. J., Zeng, X., Zweck, C., Desilets, D., Franz, T., and Rosolem, R.: COSMOS: the COsmic-ray Soil Moisture Observing System, Hydrology and Earth System Sciences, 16, 4079–4099, https://doi.org/10.5194/hess-16-4079-2012, 2012.