Transforming in-situ measurements allows robust estimation of the spatial average of soil moisture despite sensor failures

Felix Pohl[1], Martin Schrön[1], Corinna Rebmann[2], Luis E. Samaniego[3], and Anke Hildebrandt[4]

[1]Helmholtz Centre for Environmental Research GmbH - UFZ
[2]Helmholtz Centre for Environmental Research - UFZ
[3]UFZ-Helmholtz Centre for Environmental Research
[4]Friedrich Schiller University Jena

November 24, 2022

Abstract

Robust estimation of average soil water content with spatial resolution of a few tens to a few hundreds of meters is essential for evaluating models or data assimilation products. Due to the high spatial variability of soil moisture at the point scale, sufficient coverage of spatial observations is required to estimate a robust field average. If sensors fail over time, averaging the remaining measurements risks the introduction of artificial shifts in the resulting time series. Here, we explore the problem of using incomplete soil moisture observations to estimate spatial averages and propose a correction accounting for temporal persistence of spatial patterns. By transforming, i.e. upscaling, each sensor measurement to the field scale using information from time periods with sufficient coverage, the dependence on full spatial coverage can be decreased. The transformed values allow to build a more robust approximation to the spatial mean, even when spatial coverage becomes sparse. We found that high temporal stability of the sensors does not necessarily guarantee that the transformed time series will provide a good estimate of the mean and therefore recommend the use of robust statistics to derive the field mean, which requires at least three estimates per observation time. The proposed protocol is applicable for observational time series with varying sample size across a given spatial extent, and it can be adopted for other variables exhibiting a temporally stable bias between the individual point observations and field scale average.

Transforming in-situ measurements allows robust estimation of the spatial average of soil moisture despite sensor failures

Felix Pohl[1], **Martin Schrön**[1], **Corinna Rebmann**[1], **Luis Samaniego**[1], **Anke Hildebrandt**[1,2,3]

[1]Helmholtz-Centre for Environmental Research, Permoserstraße 15, 04318 Leipzig
[2]Friedrich Schiller University Jena, Institute of Geoscience, Burgweg 11, 07749 Jena
[3]German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Puschstrasse 4, 04103 Leipzig, Germany

Key Points:

- The multi-sensor average of soil moisture data is prone to substantial bias as sensors fail over time.
- Reference estimates can be used to transform single sensor measurements, thus reducing the number of required sensors.
- CDF matching with dynamic piecewise linear regression can robustly transform measurements, also in the presence of extreme events.

Corresponding author: Felix Pohl, `felix.pohl@ufz.de`

Abstract

Robust estimation of average soil water content with spatial resolution of a few tens to a few hundreds of meters is essential for evaluating models or data assimilation products. Due to the high spatial variability of soil moisture at the point scale, sufficient coverage of spatial observations is required to estimate a robust field average. If sensors fail over time, averaging the remaining measurements risks the introduction of artificial shifts in the resulting time series. Here, we explore the problem of using incomplete soil moisture observations to estimate spatial averages and propose a correction accounting for temporal persistence of spatial patterns. By transforming, i.e. upscaling, each sensor measurement to the field scale using information from time periods with sufficient coverage, the dependence on full spatial coverage can be decreased. The transformed values allow to build a more robust approximation to the spatial mean, even when spatial coverage becomes sparse. We found that high temporal stability of the sensors does not necessarily guarantee that the transformed time series will provide a good estimate of the mean and therefore recommend the use of robust statistics to derive the field mean, which requires at least three estimates per observation time. The proposed protocol is applicable for observational time series with varying sample size across a given spatial extent, and it can be adopted for other variables exhibiting a temporally stable bias between the individual point observations and field scale average.

1 Introduction

Soil moisture is a key variable for the assessment of climate change effects on ecosystem functioning (Vereecken et al., 2008; Humphrey et al., 2018; Green et al., 2019; Humphrey et al., 2021). Although its share of the global water resources is small, soil moisture plays an essential role for maintaining transpiration, plant productivity and plant health (Jaleel et al., 2009; C. Wang et al., 2019). Especially the observation of changes in soil water balance and temporal trends are of key importance, e.g., for the further development of monitoring, early warning, and projection systems related to drought or flood events (Hao et al., 2018; Bordoni et al., 2021; Rakovec et al., 2022), for the identification of parameters in hydrological models (Cuntz et al., 2015), and for improving the parameterization of land surface models (Samaniego et al., 2017). Remote sensing products and land surface models can provide large-scale information (e.g., Babaeian et al., 2019; Yao et al., 2021), but they require robust reference data for validation and error quantification (Gruber et al., 2020).

Spatial reference estimates of soil moisture are usually derived from in-situ measurements (Gruber et al., 2020). Even modern techniques to directly measure field-average water content, such as cosmic-ray neutron sensing or remote sensing, typically require multiple in-situ measurements for calibration (Colliander et al., 2017; Schrön et al., 2017). To bridge the "support gap" between reference measurements (point scale) and target product (spatial scales from meters to kilometers) requires a transfer of the information from the lower hierarchical level to the grid scale of the target product (Y. Pachepsky & Hill, 2017). Depending on the heterogeneity of the area of interest, multiple spatial measurements are needed to estimate its spatial average soil water content reliably. Due to the strong effect of local factors on soil water dynamics, randomly located single sensor measurements are usually not representative of the entire extent (Brocca et al., 2009; Heathman et al., 2012; Zhu et al., 2018). Based on literature review, Crow et al. (2012) concluded that on an area of about $800\,\mathrm{m}^2$ on average 10–20 sensors are required to obtain the field mean with an accuracy of $2\,\mathrm{vol.\,\%}$ $(1\,\sigma)$ in the top layer. Depending on the site-specific characteristics, such as topography, vegetation or climate, the actual number of required sensors can range from 1–12 sensors (Hupet & Vanclooster, 2002) to 42 sensors (C. Wang et al., 2008) in extreme cases.

Even if sufficient coverage of an area is achieved by a spatial sensor setup, continuous monitoring is always prone to sensor failures or measurement errors (e.g., through frost, cracks or preferential flow). The corresponding gaps disrupt the integrity of the representative ensemble, which can lead to shifts, biases, and increased uncertainties in the determination of the field mean (Y. A. Pachepsky et al., 2005; Guber et al., 2008; Cosh et al., 2016). The reason is that soil moisture conditions at a given location are subject to local and non-local controls (Vereecken et al., 2014; Fatichi et al., 2015; Hu et al., 2017) that can cause drier or wetter conditions on the point scale compared to the spatial average. Vachaud et al. (1985) demonstrated that this bias between point and field scale can be persistent in time, a phenomenon commonly referred to as temporal stability (TS), or also temporal persistence, rank stability, or rank order (Chen, 2006; Vanderlinden et al., 2012). Since, numerous studies confirmed TS of soil water content (e.g., Kachanoski & de Jong, 1988; Rolston et al., 1991) and utilized it for various hydrological applications, e.g., for data assimilation (Pan et al., 2012; Baatz et al., 2021) and model development (Brocca et al., 2017). The finding of TS was also essential for developing strategies to reduce the need for multiple spatial measurements when deriving reference values for the spatial mean. Here, TS can be used to identify locations that are representative of the area of interest (Grayson & Western, 1998; Jacobs, 2004; Brocca et al., 2009; Ran et al., 2017) or to correct the individual point-to-field scale bias when measuring at non-representative locations (De Lannoy et al., 2007; Crow et al., 2012; K. C. Kornelsen et al., 2015)

In a comprehensive review on TS, Vanderlinden et al. (2012) found that 29 % of all investigated datasets had a bias in their calculation of the mean relative difference, and concluded that it was likely caused by incomplete observations. Several statistical and data-driven methods to fill missing values in soil moisture time series have been tested (Bárdossy et al., 2005; Dumedah & Coulibaly, 2011; K. Kornelsen & Coulibaly, 2014; Shao et al., 2017) ranging from fairly simple techniques such as monthly average replacement to more advanced approaches such as k-NN, local variance reducing techniques, artificial neural networks or evolutionary polynomial regression. While the performance of the studied methods differed, all have in common that they are only suitable for closing relatively short gaps. For example, K. Kornelsen and Coulibaly (2014)recommend only filling gaps that are no longer than 72–100 hours since accuracy decreases with increasing gap length. Similarly, Dorigo et al. (2013) reported that the automated quality control system from the International Soil Moisture Network is incapable of handling large data gaps.

The objective of this study is to assess a strategy to create robust spatial averages in the presence of spatially and temporally irregularly distributed data gaps. In particular, we demonstrate our approach based on soil moisture data from a distributed monitoring network ($\sim 1\,\mathrm{ha}$) installed in a deciduous forest in Germany, in which most of the sensors failed over time, resulting in spatial data gaps of $> 80\,\%$ in comparison to the originally installed setup. Previous work has shown that only a certain number of active sensors are required to reliably estimate the spatial mean (Brocca et al., 2010; Crow et al., 2012; Gao et al., 2013; S. Lv et al., 2020), suggesting that the mean can still be estimated after sensor failure, provided enough sensors remain active. By estimating the minimum number of required sample size (MNRS), the data set can be split into reference and application period. We hypothesize that a transformation of the measurements is required outside of the reference period because the temporal stability of the spatial patterns can lead to bias in the time series if sensors fail (Y. A. Pachepsky et al., 2005; Guber et al., 2008). We use the reference period to estimate parameters for the transformation of the remaining sensor measurements. The upscaled data can then be used to robustly approximate the spatial average, even from a small subset of the full monitoring network. An overview of our proposed procedure can be found in Fig. 1. Additionally, we also assess the temporal stability of the sensors and discuss how it affects the accuracy of the transformation.

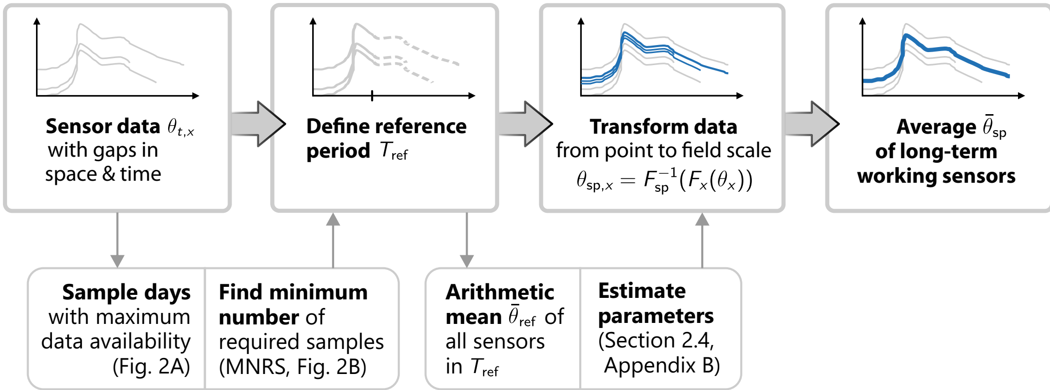

Figure 1. An overview of our proposed procedure to estimate the average field scale water content from in-situ measurements with gaps in space and time.

2 Methods and data

2.1 Soil moisture monitoring

Data is gathered within a 1 ha fenced area of the forest 'Hohes Holz' (DE-HoH, N52°05′ E11°13′, 193 m above sea-level) which is located in the northern area of the Bode water catchment near Magdeburg in Central Germany (Wollschläger et al., 2017). Soil water content sensors from a distributed monitoring network (SoilNet-WSN with SPADE sensors, sceme.de GmbH, Germany, Bogena et al., 2010) were originally installed in the frame of a trenching experiment (Marañón-Jiménez et al., 2021) in April 2014 and distributed considering patches with low and high tree (and thus root) density (21 nodes, see Fig. A1) for 15 locations. The six sensors of a node were installed in vertical profiles ranging from 10 cm to 60 cm depth. Additionally, sensors of six more nodes were distributed in the shallow layer between 10 cm and 30 cm to cover the higher soil moisture dynamics of this zone. Of the total setup described in Marañón-Jiménez et al. (2021), only sensors without soil treatment were used for our analysis. Data were acquired every 10 min via the network coordinator and stored on a field computer. In addition soil moisture was also measured with CS616 sensors (Campbell Scientific Inc., Logan, Utah, USA) in two additional profiles. Those data were also acquired and stored as 10 min averages by a CR1000 data logger (Campbell Scientific Inc., Logan, Utah, USA). More information on the research site are given in Appendix A.

Physically unrealistic data were removed by semi-automated procedures that check for limit exceedances (values below zero or above local average porosity) and spikes unrelated to precipitation. Daily averages were calculated per sensor if more than 20 % of data per day was available. For the present analysis we worked with the daily averages of the period from April 2014 to April 2021. Sensors that provided data on less than 30 days were omitted to avoid low statistical power. In total, the data used here consists of measurements from 30 sensors in 10 cm, 15 sensors in 20 cm, 24 sensors in 30 cm, 16 sensors in 40 cm and 16 sensors in 50 cm. After sequential sensor failures, between 1 and 10 sensors remained in operation per layer as of 2018. We present results mainly for the 10 cm and 50 cm layers because they show the largest differences. The 60 cm layer only consisted of very few sensors in the original setup and was therefore not considered in this work.

2.2 Evaluation of temporal stability

Given are a number A of active measurements of volumetric soil water content, θ, at locations $x \in (1, ..., A)$ and times $t \in T$ during a total measurement period T. The arithmetic spatial and temporal mean soil water contents are calculated as follows:

$$\bar{\theta}_t = \frac{1}{A} \sum_{x \in A} \theta_{t,x} \ , \tag{1}$$

where $\bar{\theta}_t$ is the spatial arithmetic mean over all active sensors at time t and

$$\bar{\theta}_x = \frac{1}{T} \sum_{t \in T} \theta_{t,x} \ , \tag{2}$$

where $\bar{\theta}_x$ is the temporal average of all observations by a sensor at location x.

To quantify temporal stability (TS), the mean relative difference (MRD) and its standard deviation (SDRD) are commonly used (Vachaud et al., 1985; Vanderlinden et al., 2012). MRD indicates the average deviation of the point measurement from the field mean, i.e., whether a particular location is drier or wetter on average than the field mean, and is defined as:

$$\mathrm{RD}_{x,t} = \frac{\theta_{x,t} - \bar{\theta}_t}{\bar{\theta}_t} \ , \tag{3}$$

$$\mathrm{MRD}_x = \frac{1}{n_x} \sum_{t=1}^{n_x} \mathrm{RD}_{x,t} \ , \tag{4}$$

where $\mathrm{RD}_{x,t}$ is the relative difference of θ at the location x and observation time t, and n_x is the number of observation days of each location. Small absolute values of $\mathrm{RD}_{x,t}$ indicate locations that are near the spatial average. The standard deviation of the relative difference (SDRD) can be used to describe the TS of a location, with lower values indicating high stability or temporal persistence of the soil moisture conditions at that location. SDRD is defined as:

$$\mathrm{SDRD}_x = \frac{1}{\sqrt{n_x - 1}} \sum_{t=1}^{n_x} (\mathrm{RD}_{x,t} - \mathrm{MRD}_x)^2 \ . \tag{5}$$

Jacobs (2004) defined a single metric that combines the information of MRD and SDRD, that can be used to define representative locations for the target area. We follow the suggestion of Zhao et al. (2010) and use the term index of time stability (ITS) instead of RMSE proposed by Jacobs (2004) to avoid confusion with the general RMSE. ITS can be calculated as:

$$\mathrm{ITS}_x = \sqrt{\mathrm{MRD}_x^2 + \mathrm{SDRD}_x^2} \ . \tag{6}$$

The smaller the value for ITS, the better a sensor location reflects the spatial average.

2.3 Definition of the reference period

We assume that the estimation error for $\overline{\theta_t}$ is small if the sample is sufficiently large, which implies that the identity of the sample (i.e., which sensors are active at time t) has little effect on the estimate of $\overline{\theta_t}$. Obviously, due to the TS of soil moisture patterns, with increasing sensor loss the fluctuating identity of the sample leads to different biases and increases the estimation error of $\overline{\theta_t}$. In order to investigate the relation between sample size and error, for each depth we randomly selected 20 days with the largest sample size, i.e., the amount of active sensors at time t. From those, we removed randomly (bootstrap with replacement, $n = 1000$) some of the active sensors ($b = 5\% \cdots 95\%$), and calculated spatial averages $\overline{\theta_{t,b}}$. We then related the coefficient of variation of the bootstrapped averages for each sampling stage b to the sample size. We determined the threshold for the minimum required sample size (MNRS) based on the ratio between increase in the coefficient of variation and the change in the sample size. Sample sizes at which the increase in CV was equal to or greater than the decrease in the sample size (Zanella et al., 2017) was used as the MNRS to estimate the spatial mean as reference. The ensemble of all measurement intervals t, at which the amount of active sensors is equal or greater than the estimated MNRS, forms the reference period (T_{ref}).

2.4 Statistical transformation

We use a non-linear transformation to estimate the field scale average from the point scale in situ measurements. This transformation is commonly known as "cumulative distribution function (CDF) matching" when its applied to soil moisture data (Reichle, 2004; Drusch, 2005; De Lannoy et al., 2007; Liu et al., 2011; Han et al., 2012; S. Wang et al., 2018) and as "quantile mapping" when it is used to correct output of climate models (Thrasher et al., 2012; Maraun, 2013; Cannon et al., 2015). Gudmundsson et al. (2012) discuss even more formulations that can be found in the literature. To avoid further confusion, we will use the term "statistical transformation" as it correctly represents the technical procedure without undermining previous studies, as suggested by Gudmundsson et al. (2012).

We attempt to correct for the point-to-field scale bias of each sensor by finding a function h that transforms the distribution of each sensor measurements to match the distribution of the observed spatial average:

$$\theta_{\mathrm{sp},x} = h(\theta_x) = F_{\mathrm{sp}}^{-1}(F_x(\theta_x)), \tag{7}$$

where F is the CDF of the spatial (sp) and point scale soil water content, respectively, and F^{-1} is the inverse CDF. We solve Eq. 7 by using the empirical CDF of the in situ measurements and the reference spatial average ($\bar{\theta}_{\mathrm{ref}}$). Previous soil moisture related works estimated h through least square fits of a third (Drusch, 2005; De Lannoy et al., 2007; Han et al., 2012; Gao et al., 2019; Tian et al., 2020) or fifth (Brocca et al., 2011; Gao et al., 2017; Zhuang et al., 2020) order polynomial equation or by a 2-parametric linear transformation (Scipal et al., 2008). Liu et al. (2009) realized the CDF matching by dividing the CDFs into eight segments with breaks at the 5th, 10th, 25th, 50th, 75th, 90th and 95th percentile, and then applying a simple linear regression for each segment to adjust the data. This approach has been adopted by, e.g., Liu et al. (2011) and (Xu & Cheng, 2021) with slightly different segments.

We adopt piecewise linear regression (PLR) to implement CDF matching because PLR has some advantages over polynomial models: (1) it is very flexible and therefore allows better fits when the data to be modeled do not follow a polynomial equation, and (2) it avoids the extrapolation problem of polynomial models since these can have strong inflections outside the domain of the data used for matching. However, instead of using fixed breaks for the segments, we estimated the breakpoints individually for each sensor because there is no objective reason why a break in the regression model should be expected at a certain percentile. This ensures that segments are built on breaks in the relationship of the data and are not limited to a specific percentile.

Breakpoint or change point detection can be realized in various ways (van den Burg & Williams, 2020). We used the r-package "dpseg" which offers a dynamic programming approach by incrementally finding local optima of a score function (Machne & Stadler, 2020):

$$S_j = \max_{i \leq j}(S_{i-\mathcal{J}} + \mathrm{score}(i,j)) - P \quad \text{with} \quad \mathcal{J} \in \{0,1\}\ , \tag{8}$$

where S is the j-th breakpoint, $\mathcal{J}$ is a binary jump parameter defining whether discontinuous jumps between adjacent segments are allowed, P is a penalty parameter tuning the allowed variance per segment, and $\mathrm{score}(i,j)$ is a scoring function quantifying the goodness-of-fit between points i and j. The negative variance of residuals is used as the scoring function:

$$\mathrm{score}(i,j) = -s_r^2 \; . \tag{9}$$

Examples of the derived breakpoints and transformation functions can be found in Appendix B.

3 Results

3.1 Minimum required sample size and temporal stability

We present spatially distributed soil moisture measurements at measurement depths of 10 cm and 50 cm. Detailed information about the particular structure of the available (respectively missing) data is given in panel A of Fig. 2. Measurements of 30 sensors in the 10 cm layer and 16 sensors in the 50 cm layer were available for our study. While some sensors delivered data for up to 98 % of the entire observation period, other sensors provided measurements only on up to 5 % of all days (or the data were rejected due to unrealistic values). Spatial data gaps are lowest at the beginning of the field study in April 2014, with most sensors failing especially during or after the winter of 2017. From 2018 on, about six to eight sensors were still operating in the 10 cm layer, while only three sensors remained in continuous operation in the 50 cm layer.

To get an idea of how the sensor failure affects the reliability of the spatial average (shown in panel A), we bootstrapped the mean of the sensors on the days with the highest data availability and then artificially reduced the sample size. The coefficient of variation (CV) of the mean is displayed in panel B in Fig. 2. The change in the CV with decreasing sensor availability shows that the CV hardly deteriorates when only a few sensors are removed, but then increases sharply when the number of sensors is small. We determined the threshold for the minimum number of required samples (MNRS) based on the ratio between increase in the coefficient of variation and the change in the sample size. The resulting MNRS is six sensors for 10 cm and five sensors for 50 cm. On these sample days with maximum data availability, the CV is less than 10 % when the MNRS is reached. It follows that the threshold ideally ensures that the CV does not exceed 10 % throughout the reference period. It also follows that in 10 cm the MNRS is given for for most of the observation period, while in 50 cm only data up to 2017 can be considered as reference.

Panel C in Fig. 2 presents the rank-ordered mean relative difference (MRD), its standard deviation (SDRD) and the index of time stability (ITS) of each sensor. Note that here we have only used the previously estimated reference period with days that meet the MNRS. At each depth, sensors can be identified that are close to the average for the entire site, and likewise, some locations are much wetter or drier than average. Deviations from the mean value can range from −51 % to 41 %, in relative terms. The comparison of 10 cm and 50 cm shows that the soil acts like a natural low-pass filter, causing the margins of the MRD to decrease with increasing depth. SDRD is also higher on average in the upper layer and sites with MRD close to zero can occur for both, smaller and larger SDRD, respectively.

3.2 Predicting the field average from in situ measurements

By mapping the distribution of each sensor to the distribution of the spatial reference mean, the measurements of each sensor are essentially transformed into a predictor of the field mean. In other words, they are rescaled (i.e., upscaled) from the lower hierarchical level, the point scale, to the field scale. Fig. 3 presents the results of this trans-

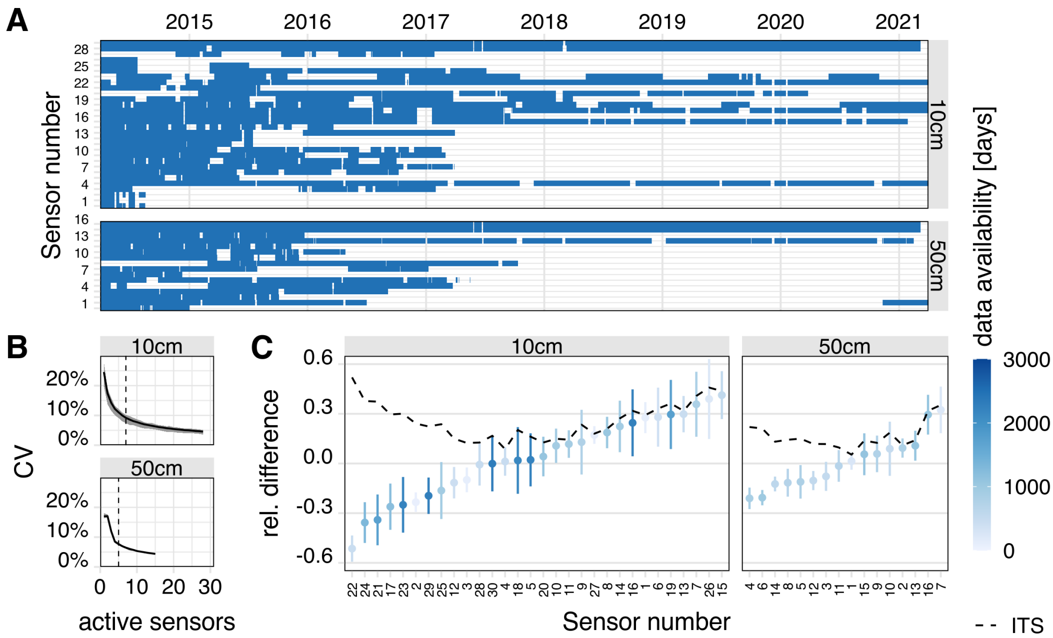

Figure 2. Quantitative information about the data used in this study. Panel A shows time series of data availability for two layers. Panel B shows the coefficient of variation (CV) of the bootstrapped mean soil moisture for various sample sizes at days with maximum data availability. The black line gives the average of all 20 days, and the gray area illustrates the range of CV over those days. Vertical dashed lines represent the threshold where the ratio between the change of CV and sample size becomes larger than unity, and which was taken to identify dates with sufficient data availability for a reliable estimate of the spatial average. Panel C shows the rank-ordered mean relative difference (MRD). Vertical bars are the standard deviation of the relative difference (SDRD) and the dashed line is the index of time stability (ITS). The colors refer to the number of days each sensor provided data.

formation. The sensors' point measurements are shown in gray and their transformed estimates in blue colour. Panel A of Fig. 3 shows the time series of all available measurements, both before and after transformation. The reference line is the arithmetic mean of the original sensor measurements in times with sufficient data (cf. Fig. 2).

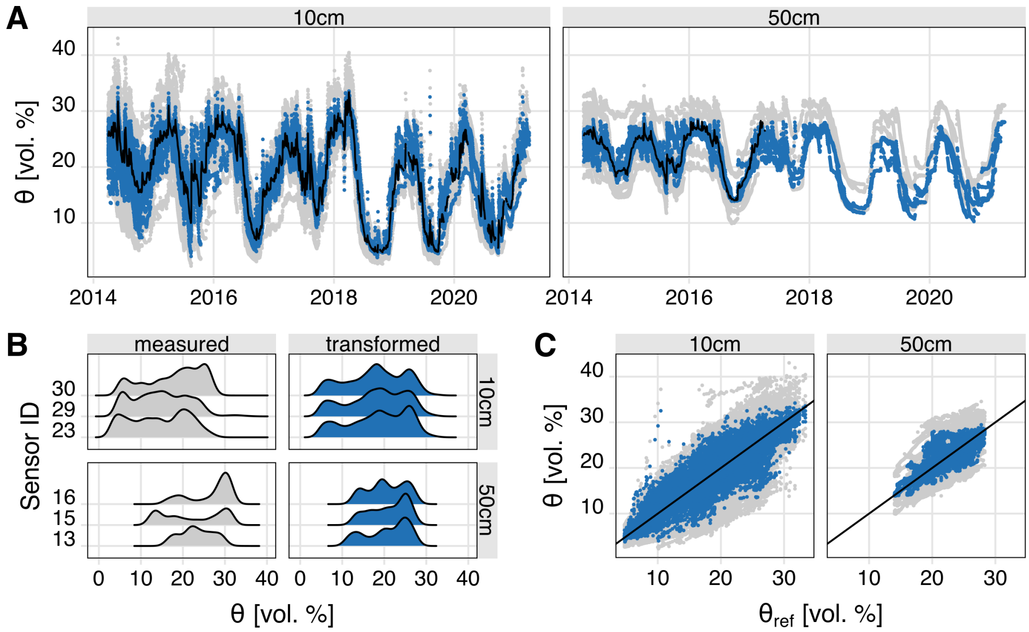

Figure 3. Information on the measurements in 10 cm and 50 cm before (gray) and after transformation (blue). Panel A shows the time series of all sensors, solid line is the arithmetic mean during the reference period. Panel B shows exemplary the empirical probability density function of three sensors with most available data before (left side) and after transformation (right side). Panel C shows the scatter plot of original and transformed sensor data versus the spatial average during the reference period.

Panel B of Fig. 3 shows the effect of the transformation on the probability density function (PDF) using the example of three sensors that provided most data at their respective depths. Overall, the PDFs of the original sensor measurements have quite different shapes, with those of the sensors at 10 cm depth being much more similar than those of the sensors at 50 cm depth. The PDFs of the sensors can be roughly summarized by a bimodal shape with a peak in the wet region and a peak in the dry region, with the exception of sensor id 13, which corresponds more to a unimodal distribution. After transformation, the PDF of the sensors at 10 cm are very similar and follow a trimodal distribution with a peak in the wet region, a second peak in the intermediate region and a less pronounced third peak in the dry region. The sensors in 50 cm roughly follow the same shape, but on a smaller range and with more variety.

Panel C of Fig. 3 complements the description of the transformed data with a scatter plot of reference versus sensor data, both for the original measurements and the transformed values. The rescaled values are much closer to the 1:1 line and less scattered than the original measurements. In comparison of 10 cm to 50 cm, in the lower layer both the variability and the range of the measurements is much smaller than in 10 cm. Note that here we present the reference to the transformation based on the same reference data. For a more detailed performance analysis with a test and training setup, the reader is referred to Appendix C.

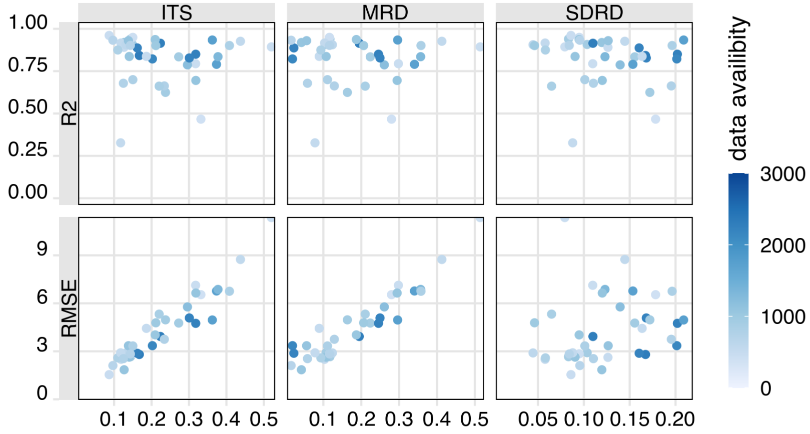

Figure 4. Scatter plot of R^2 and RMSE of rescaled measurements versus different indicators of temporal stability. Colours refer to the amount of data of each sensor.

Fig. 4 shows the relationship between various methods to characterize temporal stability (cf. Panel C, Fig. 2) and goodness of fit (gof, i.e., R^2 and RMSE) of the transformed measurements during the reference period, where an estimate of the real spatial average is available. Overall, a lower index of time stability (ITS) indicates a lower RMSE value after transformation, while R^2 remains largely unaffected. Splitting ITS into its two components, MRD and SDRD, shows that the clear relationship between ITS and RMSE is more dominated by MRD than SDRD. R^2 appears to be largely independent of the TS characteristics. However, it should be noted that R^2 is generally high regardless of TS, and the only sensors with an R^2 below 0.5 are those with low data availability. In these cases, the lower goodness of fit could simply be an artifact of the short measurement period.

3.3 Field average prediction with small sample sizes

We evaluated the effectiveness of upscaling sensors (Eq. 7) for field-scale soil water content estimation at small sample sizes during the reference period. For this, we calculated the arithmetic mean from combinations of one to six sensors (1000 repetitions) and plot the empirical CDF of RMSE and R^2 in Fig. 5. It is clear that the transformation drastically reduces the RMSE in both depth, while R^2 obtained from the means of the original measurements overall is larger than that of the transformed ones. At 10 cm, using more than two sensors does not appreciably improve the R^2 and RMSE when using the transformed data. Instead, for the original measurements, the RMSE improves significantly with each additional sensor. The same is true for 50 cm, where the RMSE is also overall smaller. With at least three sensors, an R^2 of more than 0.8 can be expected in most cases.

Three sensors remained active in the 50 cm layer, and therefore the three sensors with the highest data availability were also selected for 10 cm as a comparison. We benchmarked the performance of the three rescaled sensors in each layer as predictors of soil water content (Eq. 7) at the field scale against the estimate from the reference period, and present the goodness-of-fit (gof) in Tab. 1. In addition, we also considered the use of all three sensors and combined their estimates using the median as a robust metric of the center of the distribution. In both depths, the R^2 is higher and the RMSE lower for the median of all compared to the individual sensors. Information on the characteristics of TS of the sensors can also be found in Tab. 1.

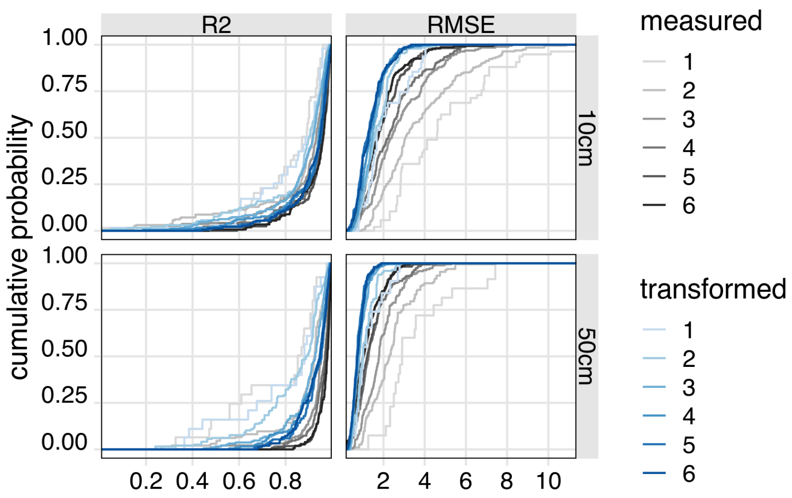

Figure 5. Empirical cumulative distribution function of R^2 and RMSE of the relation between the reference average and the average of transformed (blue) and measured (gray) SWC for random combinations of 1–6 sensors.

Table 1. Example of R^2 and RMSE for three sensors with most available data in both layers, and characteristics of TS for each sensor. In addition, R^2 and RMSE are presented for the median of the three sensors as a robust alternative to the individual predictions of the three sensors

Depth	Predictor	R^2	RMSE	|MRD|	SDRD	ITS
10 cm	Sensor 23	0.84	2.85	0.25	0.17	0.30
	Sensor 29	0.95	1.63	0.19	0.11	0.22
	Sensor 30	0.85	2.75	0.00	0.17	0.17
	Median	0.96	1.45			
50 m	Sensor 13	0.87	1.35	0.11	0.09	0.14
	Sensor 15	0.95	0.85	0.06	0.13	0.14
	Sensor 16	0.76	1.84	0.29	0.12	0.32
	Median	0.95	0.78			

The estimates for the field-scale soil water content for our research site using the transformed data is shown in Fig. 6. The difference between the simple mean of the original data (gray line) and the median of the transformed data (orange line) illustrates the effect of rescaling on field-scale soil water content. In winter and summer, the median of the transformed data are about 3–4 vol. % lower than the original data. The time series shifts towards an apparently wetter regime after the reference period, likely caused due to the sensor error than a change of the climate or soil characteristics.

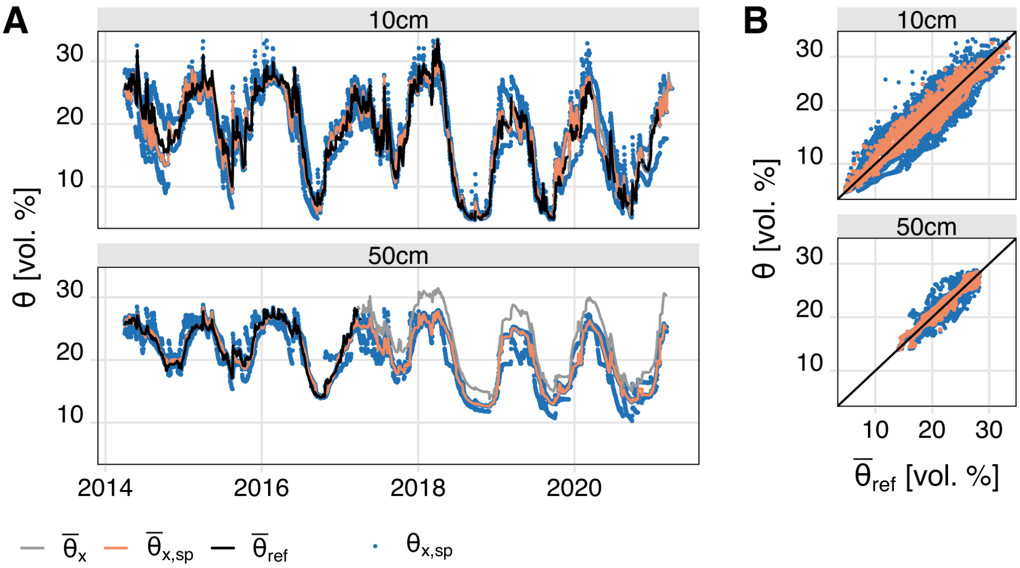

Figure 6. Estimated field scale soil water content in 10 cm and 50 cm. Three sensors remained working in 50 cm (cf. Panel A in Fig. 2), therefore three sensors were also selected in 10 cm (cf. Panel B in Fig. 3) based on highest data availability. Panel A shows the time series of the spatial average point estimates of the three sensors (Eq. 7, $\theta_{x,\mathrm{sp}}$) and their median $\bar{\theta}_{x,\mathrm{sp}}$, respectively, the reference spatial average ($\bar{\theta}_{\mathrm{ref}}$) and the arithmetic mean of the full uncorrected dataset ($\bar{\theta}_x$) outside of the reference period. Panel B shows the scatter plot of the point estimates of the three selected sensors (blue) and their median (orange), respectively, against the reference period true spatial average per layer.

4 Discussion

4.1 Challenges and limitations of working with gaps in soil moisture measurements

The spatial soil moisture characteristics of the research site exhibit strong heterogeneity, with differences of up to 20 vol. % between sensors at the same measurement time. The range of the mean relative difference (Panel C, Fig. 2) is comparable to results in other studies with similar forest and climate types (L. Lv et al., 2016; Wei et al., 2017; Zhu et al., 2021), indicating that the high degree of scatter is typically expected at such sites. The uneven data contribution of sensors along the dry-humid gradient within the data set creates a systematic problem when the remaining measurements of incomplete measurements are averaged: Due to the ordered structure of the data, a changing number of active sensors can lead to a systematic bias of the time series at the field scale (Y. A. Pachepsky et al., 2005; Guber et al., 2008). For example, if more sensors fail in drier locations, a time series of calculated spatial average using Eq. 1 would gradually shift more toward the wetter environment and thus does not reflect the entire study site anymore. This behavior was observed in the 50 cm layer (see Fig. 6). After the failure of most sensors in 2017, an artificial shift towards a wetter regime was observed.

In principle, it is always desirable to have complete measurement series, but data failure is not uncommon in field studies. However, the bootstrapping simulation shows that as long as a minimum number of sensors is active, the estimation uncertainty of the spatial mean is negligible. Bootstrapping is a robust method to estimate the sampling distribution if the true distribution is unknown, and therefore also commonly used to assess the statistical distribution of soil moisture measurements (Rowlandson et al., 2015; Singh et al., 2019; Fathololoumi et al., 2021). C. Wang et al. (2008) compared bootstrap-based estimation of required sample size with other geo-statistical and stratified sampling strategies and found similar results among the methods. Due to limited data availability, we only examined exemplary days with the greatest possible completeness. It is known that the spatial variability of soil water content changes with different phases (wetting, drying) (Illston et al., 2004; Vereecken et al., 2014), absolute water content (Brocca et al., 2010; Peng et al., 2016) and phenological state of the ecosystem (T. Wang et al., 2015). Therefore, in future studies, it might be advisable with better data availability to investigate MNRS separately for different phases. Because there may be seasonal controls and seasonal variation of the spatial dispersion (Hupet & Vanclooster, 2002; Illston et al., 2004; Biswas, 2014; Hu et al., 2017), predictions could be improved by estimating seasonal correction functions. However, this procedure might lead to jumps in the time series when moving from one season to the next, and was therefore not considered in this study.

The estimated threshold value of the MNRS for a reliable averaging of the measured values showed that only in the 10 cm layer sufficient sensors were consistently active (with the exception of a few days). In the other depths the threshold value was undercut, so that the derivation of an average without correction would lead to a biased time series. Especially obvious in 50 cm depth, a clear difference between measured and transformed data can be seen outside the reference period. Although a definitive assessment of goodness-of-fit in later years is not possible for our remaining measurements due to the lack of reference values, it shows how increasing bias threatens to manifest itself as a temporal trend. Furthermore, the proposed corrected time series is clearly a better estimate of the true spatial average for the following two reasons: First, the sudden increase in soil moisture in winter should be explainable by physical reasons by major changes in climatic conditions, since soil moisture was at a similar level in all previous winters. In fact, however, an extreme drought began in Central Europe in the summer of 2018, which also affected the study region. A daily time series of the Standardized Precipitation Evapotranspiration Index for a nearby research site can be found in Hermanns et al. (2021). Thus, an apparent shift toward a wetter regime is implausible.

Second, the measurements of the 10 cm layer can be used as a surrogate for a reference, since enough sensors were active during the entire study period. Although the soil acts as a low-pass filter, resulting in much less diurnal variability in the deeper soil layer than in the higher layers, there should still be a clear statistical relationship between the spatial averages between the corresponding depths. Correlating the time series of the spatial mean during the reference period (e.g. the best estimate of the true spatial mean) in 10 cm (cf. Fig. 5) with the field scale average time series in 50 cm yields an R^2 of 0.74 for the original measurements and 0.89 for the transformed measurements (not shown). Likewise, the RMSE decreases from 5.77 % for the original measurements to 4.06 % for the transformed measurements (not shown). Both statistics indicate a stronger relationship between layers when the transformed average is used in 50 cm.

4.2 Robust estimates of spatial soil moisture averages

The proposed statistical transformation is technically equivalent to other upscaling studies and several transformation approaches have been discussed in the literature, ranging from simple linear scaling (De Lannoy et al., 2007; Crow et al., 2012) to more advanced approaches such as Bayesian regression (Qin et al., 2013), block kriging (J. Wang et al., 2015), random forest (Clewley et al., 2017; Zappa et al., 2019) or deep learning methods (Zhang et al., 2017). For an extensive comparison of nonlinear rescaling functions see Afshar and Yilmaz (2017). Since many of these techniques involve rescaling of remote sensing products, the applicability of their results to upscaling of in situ measurements of individual sites needs to be examined.

To save cost and effort, it is generally desirable to measure at single points rather than with multiple, randomly distributed measurements. Many studies have investigated the feasibility to utilize TS of spatial patterns to use representative measurement locations (RML) for the spatial average (e.g., Rivera et al., 2014; Molero et al., 2018; Singh et al., 2019; Fry & Guber, 2020). On the other hand, there are studies that report that TS can change inter-seasonally (Zhao et al., 2010; Biswas, 2014; Dari et al., 2019) or that TS could not be confirmed depending on the type of measurement (Kirda & Reichardt, 2000; Heathman et al., 2012; Vanderlinden et al., 2012). Likewise, we found that with our data no single sensor could perfectly replace the reference measurements and that characteristics of temporal stability were only partly related to the accuracy of the rescaled measurements. We therefore deduce that, for deriving spatial averages from small samples, it is more reliable to use all available measurements and combine them by using robust estimates for the statistical location (Rousseeuw & Verboven, 2002) instead of working with single representative, potentially upscaled sensors.

5 Conclusions

In this study, we used a data set of continuous soil moisture measurements over seven years at a deciduous forest site with a large number of consecutive sensor failures to illustrate the problem of averaging incomplete observations. The characteristics of the temporal stability are comparable to other studies with similar forest and climate types. We found that as the number of sensor failures increases, the risk and magnitude of artificial shifts in the time series of the field mean increase due to the large spatial heterogeneity of soil moisture. Therefore, we adopted a strategy to cope with the spatial data gaps and temporally inconsistent sensor failures. First, we estimated the number of minimum required spatial sensor coverage to determine reference values for the spatial mean. In the second step, we corrected the point-to-field scale bias of the remaining sensor measurements outside the reference period. The corrected measurements could then be used to reliably determine the spatial mean despite extensive spatial data gaps. To estimate the spatial average from the upscaled data, we found that the median of the remaining measurements yields a higher accuracy rather than using single locations as representatives.

Overall, we emphasize the importance of making adequate adjustments for failed sensors when averaging spatial in situ measurements. Systematic spatial bias can introduce artificial trends in the spatial average time series that would affect interpretations regarding extreme events or regime shifts due to anthropogenic change. The results of this study can also be applied to other research areas where a temporally stable bias between point and spatial estimates can be expected. Furthermore, the results may be useful not only in the context of sensor failures, but also in reducing measurement effort. Once the spatial mean can be reliably estimated from a small number of sensors, it is possible to operate the network with a reduced setup. At the same time, the transformation of the measurements requires reference estimates, and so far too little is known about how long and to what extent these reference measurements have to be operated.

Future research should focus on sensitivity to the length and spatial scale of those reference estimates. At the same time, indirect measurements from remote sensing products as well as cosmic-ray neutron sensing measurements could be useful sources of information to further reduce the in situ effort required for reference determinations of soil water content.

Appendix A Research site description

The 'Hohes Holz' is a deciduous forest covering an area of around 15 km^2, dominated by sessile oak (Quercus petraea (Matt.) Liebl.), common beech (Fagus sylvatica L.), and hornbeam (Carpinus betulus L.). The climate is a temperate climate with a mean annual temperature of 9.1 °C and a mean annual precipitation of 563 mm (climate period 1981–2010, station Ummendorf of the German Weather Service). During the investigated period from 2014 until 2020 yearly precipitation sums ranged from 301 mm (2018) to 610 mm (2017). The bedrock is Pleistocene sandy loess above till and Mesozoic muschelkalk, with Haplic Cambisol as predominant soil type. Soil texture at 0–20 cm depth was 3.0 % (±1.8 %) sand, 87.1 % (±2.1 %) silt, and 10.0 % (±2.2 %) clay.

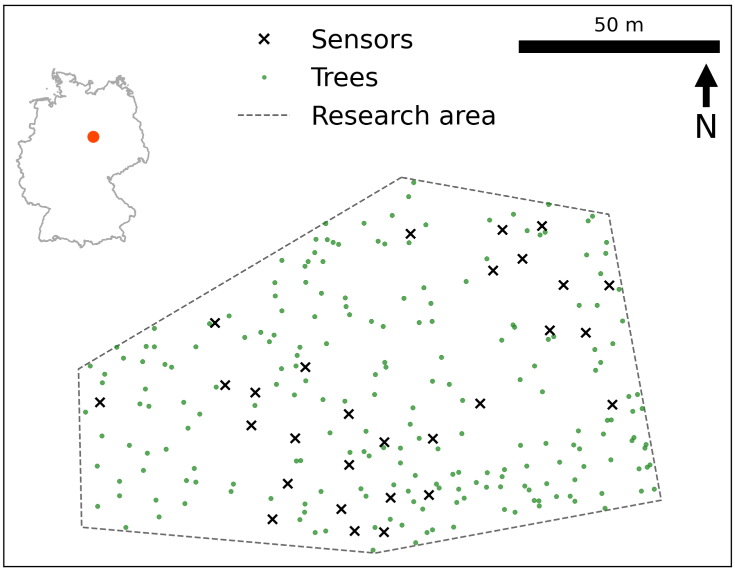

Figure A1. Sensor distribution in the study area 'Hohes Holz' (DE-HoH) with additional information on tree density and the location of the study area in Germany. Note that at each marker sensors are distributed over the depths from 10 cm to 60 cm in 10 cm intervals, but not all depths are covered at each marker.

Appendix B CDF matching with dynamic piecewise linear regression

We compare polynomial fits, which are traditionally used for cumulative distribution function (CDF) matching, against the piecewise linear regression (PLR) approach with flexible segments, proposed in this study. Fig. B1 shows exemplary the difference between the spatial reference mean and sensors with most available data. Sensor 16 in 50 cm and sensor 23 in 10 cm are good examples of how the polynomial fit can lead to large over- or underestimates, especially in the context of extrapolation. The PLR approach can theoretically be fit to any functional form and extrapolation can be realized by adopting the last known linear function at the minimum and maximum spatial reference.

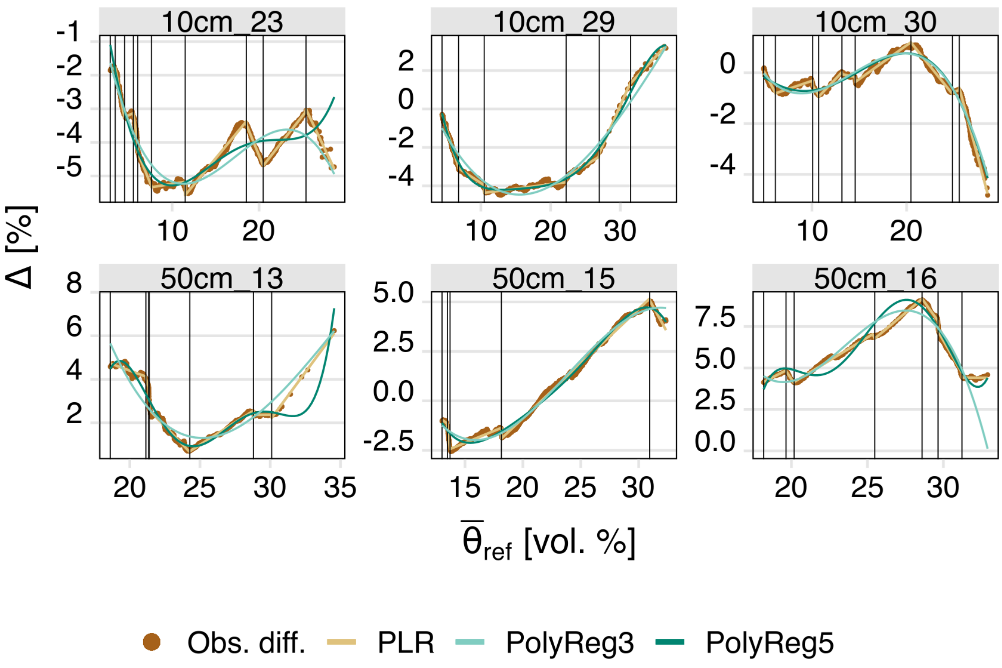

Figure B1. Transformation functions for exemplary sensors with the most available data (see Fig. 2). The dots are the observed difference between the CDFs and the lines are the derived functions for fitting the CDF of the sensors to the CDF of the spatial reference mean. CDFs before and after transformation are shown in Fig. 3. In this study, dynamic piecewise linear regression was used (PLR); for comparison, the traditionally used fits with polynomial regression (third and fifth order, respectively) are also shown. Vertical lines are the breaks of the PLR.

Appendix C Performance of statistical transformation

Similar to De Lannoy et al. (2007) and Gudmundsson et al. (2012), we benchmarked the proposed transformation of section 2.4 against the following parametric transformations:

$$\theta_{\mathrm{sp},x} = a + \theta_x \tag{C1}$$
$$\theta_{\mathrm{sp},x} = b \cdot \theta_x \tag{C2}$$
$$\theta_{\mathrm{sp},x} = a + b \cdot \theta_x \tag{C3}$$

where θ is the soil moisture on the point and field scale, respectively, and a, b are parameters to be estimated. We split the data of the reference period into two equally sized groups to test the performance of the transformations. The parameters of the scaling methods were estimated using the training data and then applied to the test data. The accuracy of the fit between the observed and estimated field average soil moisture was assessed with the following goodness-of-fit parameters: Mean absolute error (MAE), Root Mean Square Error (RMSE), Pearson correlation coefficient (R) and Nash-Suttcliffe efficiency (NS):

$$\mathrm{MAE} = |\bar{y} - \bar{\hat{y}}| \tag{C4}$$
$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2} \tag{C5}$$
$$\mathrm{R} = \frac{\mathrm{Cov}(y, \hat{y})}{s_y s_{\hat{y}}} \tag{C6}$$
$$\mathrm{NS} = 1 - \frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{N}(y_i - \bar{y}_i)^2}\,. \tag{C7}$$

Tab. C1 summarizes the results for each method as the mean of the gof of all sensors per layer. In every layer and with every gof, the non-linear CDF matching achieved best results. Among the linear methods, the linear regression usually performed best. Both, the MAE and RMSE decrease with depth, with a maximum MAE of 2.76 % for θ_{C2} in 30 cm and a minimum of 1.02 % for θ_{CDF} in 50 cm. The correlation coefficient shows narrow ranges per layer, with very strong correlations for all mθ_{CDF} in 40 cm (0.94) and lowest correlation for θ_{C1} in 50 cm (0.85). The Nash–Sutcliffe model efficiency coefficient (NS) shows lowest performance in 50 cm (0.712 for θ_{C1}) and highest performance in 10 cm (0.87 for θ_{CDF}). RMSE is lowest for θ_{C3} (1.50 % in 50 cm) and highest for θ_{C2} (3.66 % in 30 cm).

Table C1. Mean average error (MAE, [vol. %]), Nash–Suttcliffe criterium (NS, [–]), correlation coefficient (R, [–]) and root mean square error (RMSE, [vol. %]) between the spatial average soil moisture and transformed point measurements during the reference period. Scores are the average over all sensors per layer.

	MAE				NS				R				RMSE			
Layer	θ_{C1}	θ_{C2}	θ_{C3}	θ_{CDF}	θ_{C1}	θ_{C2}	θ_{C3}	θ_{CDF}	θ_{C1}	θ_{C2}	θ_{C3}	θ_{CDF}	θ_{C1}	θ_{C2}	θ_{C3}	θ_{CDF}
10 cm	2.03	2.15	1.88	1.75	0.84	0.84	0.84	0.87	0.92	0.91	0.93	0.93	2.69	2.87	2.43	2.41
20 cm	1.89	2.22	1.58	1.32	0.80	0.77	0.80	0.85	0.89	0.88	0.91	0.93	2.45	2.82	2.01	1.89
30 cm	2.02	2.76	1.70	1.51	0.80	0.74	0.80	0.84	0.90	0.87	0.91	0.92	2.79	3.66	2.26	2.22
40 cm	1.77	1.92	1.48	1.24	0.84	0.83	0.86	0.88	0.92	0.91	0.93	0.94	2.34	2.45	1.91	1.85
50 cm	1.56	1.54	1.11	1.02	0.72	0.74	0.75	0.80	0.85	0.86	0.90	0.90	2.07	2.00	1.50	1.51

Appendix D Open Research

The daily averaged soil moisture measurements used in the study are available at the Helmholtz-Centre for Environmental Research data archive via doi.org/10.48758/ufz.12770 under CC BY-NC-SA 4.0 (Rebmann et al., 2018). The code to reproduce all results and figures (except for Fig. 1 and Fig. A1) is preserved at doi.org/10.5281/zenodo.6653168 and available under CC BY-NC-SA 4.0 (Pohl et al., 2022).

Acknowledgments

A.H. gratefully acknowledges the support of iDiv funded by the German Research Foundation (DFG– FZT 118, 202548816) and CRC AquaDiva (SFB 1076 – Project Number 218627073). M.S. acknowledges support by the DFG (German Research Foundation) via the project 357874777, research unit FOR 2694 Cosmic Sense. The study has been made possible by the infrastructural funds of the Helmholtz Association and the Terrestrial Environmental Observatories (TERENO). The operation and data gathering of field data was supported by Sebastian Gimper and Patrick Schmidt. Juliane Mai and Matthias Cuntz supported the data-treatment procedures in previous versions. We thank Floris Hermanns for his support in Fig. A1.

References

Afshar, M., & Yilmaz, M. (2017). The added utility of nonlinear methods compared to linear methods in rescaling soil moisture products. *Remote Sensing of Environment*, *196*, 224–237. doi: 10.1016/j.rse.2017.05.017

Baatz, R., Hendricks Franssen, H. J., Euskirchen, E., Sihi, D., Dietze, M., Ciavatta, S., ... Vereecken, H. (2021). Reanalysis in Earth System Science: Toward Terrestrial Ecosystem Reanalysis. *Reviews of Geophysics*, *59*(3). doi: 10.1029/2020RG000715

Babaeian, E., Sadeghi, M., Jones, S. B., Montzka, C., Vereecken, H., & Tuller, M. (2019). Ground, Proximal, and Satellite Remote Sensing of Soil Moisture. *Reviews of Geophysics*, *57*(2), 530–616. doi: 10.1029/2018RG000618

Bárdossy, A., Pegram, G. G. S., & Samaniego, L. (2005). Modeling data relationships with a local variance reducing technique: Applications in hydrology. *Water Resources Research*, *41*(8). doi: 10.1029/2004WR003851

Biswas, A. (2014). Season- and depth-dependent time stability for characterising representative monitoring locations of soil water storage in a hummocky landscape. *CATENA*, *116*, 38–50. doi: 10.1016/j.catena.2013.12.008

Bogena, H., Herbst, M., Huisman, J., Rosenbaum, U., Weuthen, A., & Vereecken, H. (2010). Potential of Wireless Sensor Networks for Measuring Soil Water Content Variability. *Vadose Zone Journal*, *9*(4), 1002–1013. doi: 10.2136/vzj2009.0173

Bordoni, M., Inzaghi, F., Vivaldi, V., Valentino, R., Bittelli, M., & Meisina, C. (2021). A Data-Driven Method for the Temporal Estimation of Soil Water Potential and Its Application for Shallow Landslides Prediction. *Water*, *13*(9), 1208. doi: 10.3390/w13091208

Brocca, L., Ciabatta, L., Massari, C., Camici, S., & Tarpanelli, A. (2017). Soil Moisture for Hydrological Applications: Open Questions and New Opportunities. *Water*, *9*(2), 140. doi: 10.3390/w9020140

Brocca, L., Hasenauer, S., Lacava, T., Melone, F., Moramarco, T., Wagner, W., ... Bittelli, M. (2011). Soil moisture estimation through ASCAT and AMSR-E sensors: An intercomparison and validation study across Europe. *Remote Sensing of Environment*, *115*(12), 3390–3408. doi: 10.1016/j.rse.2011.08.003

Brocca, L., Melone, F., Moramarco, T., & Morbidelli, R. (2009). Soil moisture temporal stability over experimental areas in Central Italy. *Geoderma*, *148*(3-4), 364–374. doi: 10.1016/j.geoderma.2008.11.004

Brocca, L., Melone, F., Moramarco, T., & Morbidelli, R. (2010). Spatial-temporal variability of soil moisture and its estimation across scales. *Water Resources Research*, *46*(2). doi: 10.1029/2009WR008016

Cannon, A. J., Sobie, S. R., & Murdock, T. Q. (2015). Bias Correction of GCM Precipitation by Quantile Mapping: How Well Do Methods Preserve Changes in Quantiles and Extremes? *Journal of Climate*, *28*(17), 6938–6959. doi: 10.1175/JCLI-D-14-00754.1

Chen, Y.-j. (2006). Letter to the Editor on “Rank Stability or Temporal Stability”. *Soil Science Society of America Journal*, *70*(1), 306–306. doi: 10.2136/sssaj2005.0290l

Clewley, D., Whitcomb, J. B., Akbar, R., Silva, A. R., Berg, A., Adams, J. R., … Moghaddam, M. (2017). A Method for Upscaling In Situ Soil Moisture Measurements to Satellite Footprint Scale Using Random Forests. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *10*(6), 2663–2673. doi: 10.1109/JSTARS.2017.2690220

Colliander, A., Jackson, T., Bindlish, R., Chan, S., Das, N., Kim, S., … Yueh, S. (2017). Validation of SMAP surface soil moisture products with core validation sites. *Remote Sensing of Environment*, *191*, 215–231. doi: 10.1016/j.rse.2017.01.021

Cosh, M. H., Ochsner, T. E., McKee, L., Dong, J., Basara, J. B., Evett, S. R., … Sayde, C. (2016). The Soil Moisture Active Passive Marena, Oklahoma, In Situ Sensor Testbed (SMAP-MOISST): Testbed Design and Evaluation of In Situ Sensors. *Vadose Zone Journal*, *15*(4), 1–11. doi: 10.2136/vzj2015.09.0122

Crow, W. T., Berg, A. A., Cosh, M. H., Loew, A., Mohanty, B. P., Panciera, R., … Walker, J. P. (2012). Upscaling sparse ground-based soil moisture observations for the validation of coarse-resolution satellite soil moisture products. *Reviews of Geophysics*, *50*(2). doi: 10.1029/2011RG000372

Cuntz, M., Mai, J., Zink, M., Thober, S., Kumar, R., Schäfer, D., … Samaniego, L. (2015). Computationally inexpensive identification of noninformative model parameters by sequential screening. *Water Resources Research*, *51*(8), 6417–6441. doi: 10.1002/2015WR016907

Dari, J., Morbidelli, R., Saltalippi, C., Massari, C., & Brocca, L. (2019). Spatial-temporal variability of soil moisture: Addressing the monitoring at the catchment scale. *Journal of Hydrology*, *570*, 436–444. doi: 10.1016/j.jhydrol.2019.01.014

De Lannoy, G. J., Houser, P. R., Verhoest, N. E., Pauwels, V. R., & Gish, T. J. (2007). Upscaling of point soil moisture measurements to field averages at the OPE3 test site. *Journal of Hydrology*, *343*(1-2), 1–11. doi: 10.1016/j.jhydrol.2007.06.004

Dorigo, W., Xaver, A., Vreugdenhil, M., Gruber, A., Hegyiová, A., Sanchis-Dufau, A., … Drusch, M. (2013). Global Automated Quality Control of In Situ Soil Moisture Data from the International Soil Moisture Network. *Vadose Zone Journal*, *12*(3), vzj2012.0097. doi: 10.2136/vzj2012.0097

Drusch, M. (2005). Observation operators for the direct assimilation of TRMM microwave imager retrieved soil moisture. *Geophysical Research Letters*, *32*(15), L15403. doi: 10.1029/2005GL023623

Dumedah, G., & Coulibaly, P. (2011). Evaluation of statistical methods for infilling missing values in high-resolution soil moisture data. *Journal of Hydrology*, *400*(1-2), 95–102. doi: 10.1016/j.jhydrol.2011.01.028

Fathololoumi, S., Vaezi, A. R., Firozjaei, M. K., & Biswas, A. (2021). Quantifying the effect of surface heterogeneity on soil moisture across regions and surface characteristic. *Journal of Hydrology*, *596*, 126132. doi: 10.1016/j.jhydrol.2021.126132

Fatichi, S., Katul, G. G., Ivanov, V. Y., Pappas, C., Paschalis, A., Consolo, A., … Burlando, P. (2015). Abiotic and biotic controls of soil moisture spatiotemporal variability and the occurrence of hysteresis. *Water Resources Research*, *51*(5), 3505–3524. doi: 10.1002/2014WR016102

Fry, J. E., & Guber, A. K. (2020). Temporal stability of field-scale patterns in soil water content across topographically diverse agricultural landscapes. *Journal of Hydrology*, *580*, 124260. doi: 10.1016/j.jhydrol.2019.124260

Gao, X., Wu, P., Zhao, X., Wang, J., Shi, Y., Zhang, B., ... Li, H. (2013). Estimation of spatial soil moisture averages in a large gully of the Loess Plateau of China through statistical and modeling solutions. *Journal of Hydrology*, *486*, 466–478. doi: 10.1016/j.jhydrol.2013.02.026

Gao, X., Zhao, X., Brocca, L., Huo, G., Lv, T., & Wu, P. (2017). *Depth scaling of soil moisture content from surface to profile: Multistationtesting of observation operators* (Preprint). Vadose Zone Hydrology/Modelling approaches. doi: 10.5194/hess-2017-292

Gao, X., Zhao, X., Brocca, L., Pan, D., & Wu, P. (2019). Testing of observation operators designed to estimate profile soil moisture from surface measurements. *Hydrological Processes*, *33*(4), 575–584. doi: 10.1002/hyp.13344

Grayson, R. B., & Western, A. W. (1998). Towards area1 estimation of soil water content from point measurements: Time and space stability of mean response. *Water Resources Research*, *207*(1/2), 15.

Green, J. K., Seneviratne, S. I., Berg, A. M., Findell, K. L., Hagemann, S., Lawrence, D. M., & Gentine, P. (2019). Large influence of soil moisture on long-term terrestrial carbon uptake. *Nature*, *565*(7740), 476–479. doi: 10.1038/s41586-018-0848-x

Gruber, A., De Lannoy, G., Albergel, C., Al-Yaari, A., Brocca, L., Calvet, J.-C., ... Wagner, W. (2020). Validation practices for satellite soil moisture retrievals: What are (the) errors? *Remote Sensing of Environment*, *244*, 111806. doi: 10.1016/j.rse.2020.111806

Guber, A., Gish, T., Pachepsky, Y., van Genuchten, M., Daughtry, C., Nicholson, T., & Cady, R. (2008). Temporal stability in soil water content patterns across agricultural fields. *CATENA*, *73*(1), 125–133. doi: 10.1016/j.catena.2007.09.010

Gudmundsson, L., Bremnes, J. B., Haugen, J. E., & Engen-Skaugen, T. (2012). Technical Note: Downscaling RCM precipitation to the station scale using statistical transformations – a comparison of methods. *Hydrology and Earth System Sciences*, *16*(9), 3383–3390. doi: 10.5194/hess-16-3383-2012

Han, E., Heathman, G. C., Merwade, V., & Cosh, M. H. (2012). Application of observation operators for field scale soil moisture averages and variances in agricultural landscapes. *Journal of Hydrology*, *444–445*, 34–50. doi: 10.1016/j.jhydrol.2012.03.035

Hao, Z., Singh, V. P., & Xia, Y. (2018). Seasonal Drought Prediction: Advances, Challenges, and Future Prospects. *Reviews of Geophysics*, *56*(1), 108–141. doi: 10.1002/2016RG000549

Heathman, G. C., Cosh, M. H., Han, E., Jackson, T. J., McKee, L., & McAfee, S. (2012). Field scale spatiotemporal analysis of surface soil moisture for evaluating point-scale in situ networks. *Geoderma*, *170*, 195–205. doi: 10.1016/j.geoderma.2011.11.004

Hermanns, F., Pohl, F., Rebmann, C., Schulz, G., Werban, U., & Lausch, A. (2021). Inferring Grassland Drought Stress with Unsupervised Learning from Airborne Hyperspectral VNIR Imagery. *Remote Sensing*, *13*(10), 1885. doi: 10.3390/rs13101885

Hu, W., Si, B. C., Biswas, A., & Chau, H. W. (2017). Temporally stable patterns but seasonal dependent controls of soil water content: Evidence from wavelet analyses. *Hydrological Processes*, *31*(21), 3697–3707. doi: 10.1002/hyp.11289

Humphrey, V., Berg, A., Ciais, P., Gentine, P., Jung, M., Reichstein, M., ... Frankenberg, C. (2021). Soil moisture–atmosphere feedback dominates land carbon uptake variability. *Nature*, *592*(7852), 65–69. doi: 10.1038/s41586-021-03325-5

Humphrey, V., Zscheischler, J., Ciais, P., Gudmundsson, L., Sitch, S., & Seneviratne, S. I. (2018). Sensitivity of atmospheric CO2 growth rate to observed changes in terrestrial water storage. *Nature*, *560*(7720), 628–631. doi: 10.1038/s41586-018-0424-4

Hupet, F., & Vanclooster, M. (2002). Intraseasonal dynamics of soil moisture variability within a small agricultural maize cropped ®eld. *Journal of Hydrology*, 16.

Illston, B. G., Basara, J. B., & Crawford, K. C. (2004). Seasonal to interannual variations of soil moisture measured in Oklahoma. *Int. J. Climatol.*, 14.

Jacobs, J. (2004). SMEX02: Field scale variability, time stability and similarity of soil moisture. *Remote Sensing of Environment*, *92*(4), 436–446. doi: 10.1016/j.rse.2004.02.017

Jaleel, C. A., Manivannan, P., Wahid, A., Farooq, M., Al-Juburi, J., Somasundaram, R., & Panneerselvam, R. (2009). Drought Stress in Plants: A Review on Morphological Characteristics and Pigments Composition. *Int. J. Agric. Biol.*, *11*(1), 6.

Kachanoski, R. G., & de Jong, E. (1988). Scale dependence and the temporal persistence of spatial patterns of soil water storage. *Water Resources Research*, *24*(1), 85–91. doi: 10.1029/WR024i001p00085

Kirda, C., & Reichardt, K. (2000). Comparison of neutron moisture gauges with non-nuclear methods to measure field soil water status. *International Agrophysics*, *6*(1-2), 77–87.

Kornelsen, K., & Coulibaly, P. (2014). Comparison of Interpolation, Statistical, and Data-Driven Methods for Imputation of Missing Values in a Distributed Soil Moisture Dataset. *Journal of Hydrologic Engineering*, *19*(1), 26–43. doi: 10.1061/(ASCE)HE.1943-5584.0000767

Kornelsen, K. C., Cosh, M. H., & Coulibaly, P. (2015). Potential of bias correction for downscaling passive microwave and soil moisture data: Bias Correction Downscaling. *Journal of Geophysical Research: Atmospheres*, *120*(13), 6460–6479. doi: 10.1002/2015JD023550

Liu, Y. Y., Parinussa, R. M., Dorigo, W. A., & Evans, J. P. (2011). Developing an improved soil moisture dataset by blending passive and active microwave satellite-based retrievals. *Hydrol. Earth Syst. Sci.*, 12.

Liu, Y. Y., van Dijk, A. I. J. M., de Jeu, R. A. M., & Holmes, T. R. H. (2009). An analysis of spatiotemporal variations of soil and vegetation moisture from a 29-year satellite-derived data set over mainland Australia. *Water Resources Research*, *45*(7). doi: 10.1029/2008WR007187

Lv, L., Liao, K., Lai, X., Zhu, Q., & Zhou, S. (2016). Hillslope soil moisture temporal stability under two contrasting land use types during different time periods. *Environmental Earth Sciences*, *75*(7), 560. doi: 10.1007/s12665-015-5238-1

Lv, S., Schalge, B., Saavedra Garfias, P., & Simmer, C. (2020). Required sampling density of ground-based soil moisture and brightness temperature observations for calibration and validation of L-band satellite observations based on a virtual reality. *Hydrology and Earth System Sciences*, *24*(4), 1957–1973. doi: 10.5194/hess-24-1957-2020

Machne, R., & Stadler, P. F. (2020). *Piecewise Linear Segmentation by Dynamic Programming.*

Marañón-Jiménez, S., Radujković, D., Verbruggen, E., Grau, O., Cuntz, M., Peñuelas, J., … Rebmann, C. (2021). Shifts in the Abundances of Saprotrophic and Ectomycorrhizal Fungi With Altered Leaf Litter Inputs. *Frontiers in Plant Science*, *12*, 682142. doi: 10.3389/fpls.2021.682142

Maraun, D. (2013). Bias Correction, Quantile Mapping, and Downscaling: Revisiting the Inflation Issue. *Journal of Climate*, *26*(6), 2137–2143. doi: 10.1175/

JCLI-D-12-00821.1

Molero, B., Leroux, D. J., Richaume, P., Kerr, Y. H., Merlin, O., Cosh, M. H., & Bindlish, R. (2018). Multi-Timescale Analysis of the Spatial Representativeness of In Situ Soil Moisture Data within Satellite Footprints: Soil Moisture Time and Spatial Scales. *Journal of Geophysical Research: Atmospheres*, *123*(1), 3–21. doi: 10.1002/2017JD027478

Pachepsky, Y., & Hill, R. L. (2017). Scale and scaling in soils. *Geoderma*, *287*, 4–30. doi: 10.1016/j.geoderma.2016.08.017

Pachepsky, Y. A., Guber, A. K., & Jacques, D. (2005). Temporal persistence in vertical distributions of soil moisture contents. *SOIL SCI. SOC. AM. J.*, *69*, 6.

Pan, F., Pachepsky, Y., Jacques, D., Guber, A., & Hill, R. L. (2012). Data Assimilation with Soil Water Content Sensors and Pedotransfer Functions in Soil Water Flow Modeling. *Soil Science Society of America Journal*, *76*(3), 829–844. doi: 10.2136/sssaj2011.0090

Peng, Z., Tian, F., Hu, H., Zhao, S., Tie, Q., Sheng, H., ... Lu, H. (2016). Spatial Variability of Soil Moisture in a Forest Catchment: Temporal Trend and Contributors. *Forests*, *7*(12), 154. doi: 10.3390/f7080154

Pohl, F., Schrön, M., Rebmann, C., Samaniego, L., & Hildebrandt, A. (2022). *Transforming in situ measurements allows robust estimation of the spatial soil moisture average despite sensor failures: Supporting software code.* Zenodo. doi: 10.5281/zenodo.6669749

Qin, J., Yang, K., Lu, N., Chen, Y., Zhao, L., & Han, M. (2013). Spatial upscaling of in-situ soil moisture measurements based on MODIS-derived apparent thermal inertia. *Remote Sensing of Environment*, *138*, 1–9. doi: 10.1016/j.rse.2013.07.003

Rakovec, O., Samaniego, L., Hari, V., Markonis, Y., Moravec, V., Thober, S., ... Kumar, R. (2022). The 2018–2020 Multi-Year Drought Sets a New Benchmark in Europe. *Earth's Future*, *10*(3), e2021EF002394. doi: 10.1029/2021EF002394

Ran, Y., Li, X., Jin, R., Kang, J., & Cosh, M. H. (2017). Strengths and weaknesses of temporal stability analysis for monitoring and estimating grid-mean soil moisture in a high-intensity irrigated agricultural landscape. *Water Resources Research*, *53*(1), 283–301. doi: 10.1002/2015WR018182

Rebmann, C., Aubinet, M., Schmid, H., Arriga, N., Aurela, M., Burba, G., ... Franz, D. (2018). ICOS eddy covariance flux-station site setup: A review. *International Agrophysics*, *32*(4), 471–494. doi: 10.1515/intag-2017-0044

Reichle, R. H. (2004). Bias reduction in short records of satellite soil moisture. *Geophysical Research Letters*, *31*(19), L19501. doi: 10.1029/2004GL020938

Rivera, D., Lillo, M., & Granda, S. (2014). Representative locations from time series of soil water content using time stability and wavelet analysis. *Environmental Monitoring and Assessment*, *186*(12), 9075–9087. doi: 10.1007/s10661-014-4067-0

Rolston, D., Biggar, J., & Nightingale, H. (1991). Temporal persistence of spatial soil-water patterns under trickle irrigation. *Irrigation Science*, *12*(4). doi: 10.1007/BF00190521

Rousseeuw, P. J., & Verboven, S. (2002). Robust estimation in very small samples. *Computational Statistics & Data Analysis*, *40*(4), 741–758. doi: 10.1016/S0167-9473(02)00078-6

Rowlandson, T., Impera, S., Belanger, J., Berg, A. A., Toth, B., & Magagi, R. (2015). Use of in situ soil moisture network for estimating regional-scale soil moisture during high soil moisture conditions. *Canadian Water Resources Journal / Revue canadienne des ressources hydriques*, *40*(4), 343–351. doi: 10.1080/07011784.2015.1061948

Samaniego, L., Kumar, R., Thober, S., Rakovec, O., Zink, M., Wanders, N., ... Attinger, S. (2017). Toward seamless hydrologic predictions across spatial scales. *Hydrology and Earth System Sciences*, *21*(9), 4323–4346. doi: 10.5194/hess-21-4323-2017

Schrön, M., Köhli, M., Scheiffele, L., Iwema, J., Bogena, H. R., Lv, L., ... Zacharias, S. (2017). Improving calibration and validation of cosmic-ray neutron sensors in the light of spatial sensitivity. *Hydrology and Earth System Sciences*, *21*(10), 5009–5030. doi: 10.5194/hess-21-5009-2017

Scipal, K., Drusch, M., & Wagner, W. (2008). Assimilation of a ERS scatterometer derived soil moisture index in the ECMWF numerical weather prediction system. *Advances in Water Resources*, *31*(8), 1101–1112. doi: 10.1016/j.advwatres.2008.04.013

Shao, J., Meng, W., & Sun, G. (2017). Evaluation of missing value imputation methods for wireless soil datasets. *Personal and Ubiquitous Computing*, *21*(1), 113–123. doi: 10.1007/s00779-016-0978-9

Singh, G., Panda, R. K., & Mohanty, B. P. (2019). Spatiotemporal Analysis of Soil Moisture and Optimal Sampling Design for Regional-Scale Soil Moisture Estimation in a Tropical Watershed of India. *Water Resources Research*, *55*(3), 2057–2078. doi: 10.1029/2018WR024044

Thrasher, B., Maurer, E. P., McKellar, C., & Duffy, P. B. (2012). Technical Note: Bias correcting climate model simulated daily temperature extremes with quantile mapping. *Hydrology and Earth System Sciences*, *16*(9), 3309–3314. doi: 10.5194/hess-16-3309-2012

Tian, J., Han, Z., Bogena, H. R., Huisman, J. A., Montzka, C., Zhang, B., & He, C. (2020). Estimation of subsurface soil moisture from surface soil moisture in cold mountainous areas. *Hydrology and Earth System Sciences*, *24*(9), 4659–4674. doi: 10.5194/hess-24-4659-2020

Vachaud, G., Passerat De Silans, A., Balabanis, P., & Vauclin, M. (1985). Temporal Stability of Spatially Measured Soil Water Probability Density Function. *Soil Science Society of America Journal*, *49*(4), 822–828. doi: 10.2136/sssaj1985.03615995004900040006x

van den Burg, G. J. J., & Williams, C. K. I. (2020). An Evaluation of Change Point Detection Algorithms. *arXiv:2003.06222 [cs, stat]*.

Vanderlinden, K., Vereecken, H., Hardelauf, H., Herbst, M., Martínez, G., Cosh, M. H., & Pachepsky, Y. A. (2012). Temporal Stability of Soil Water Contents: A Review of Data and Analyses. *Vadose Zone Journal*, *11*(4), vzj2011.0178. doi: 10.2136/vzj2011.0178

Vereecken, H., Huisman, J., Pachepsky, Y., Montzka, C., van der Kruk, J., Bogena, H., ... Vanderborght, J. (2014). On the spatio-temporal dynamics of soil moisture at the field scale. *Journal of Hydrology*, *516*, 76–96. doi: 10.1016/j.jhydrol.2013.11.061

Vereecken, H., Huisman, J. A., Bogena, H., Vanderborght, J., Vrugt, J. A., & Hopmans, J. W. (2008). On the value of soil moisture measurements in vadose zone hydrology: A review. *Water Resources Research*, *44*(4). doi: 10.1029/2008WR006829

Wang, C., Fu, B., Zhang, L., & Xu, Z. (2019). Soil moisture–plant interactions: An ecohydrological review. *Journal of Soils and Sediments*, *19*(1), 1–9. doi: 10.1007/s11368-018-2167-0

Wang, C., Zuo, Q., & Zhang, R. (2008). Estimating the necessary sampling size of surface soil moisture at different scales using a random combination method. *Journal of Hydrology*, *352*(3-4), 309–321. doi: 10.1016/j.jhydrol.2008.01.011

Wang, J., Ge, Y., Heuvelink, G., & Zhou, C. (2015). Upscaling In Situ Soil Moisture Observations to Pixel Averages with Spatio-Temporal Geostatistics. *Remote Sensing*, *7*(9), 11372–11388. doi: 10.3390/rs70911372

Wang, S., Shan, H., Zhang, C., Wang, Y., & Shi, C. (2018). Bias Correction in Monthly Records of Satellite Soil Moisture Using Nonuniform CDFs. *Advances in Meteorology*, *2018*, 1–11. doi: 10.1155/2018/1908570

Wang, T., Wedin, D. A., Franz, T. E., & Hiller, J. (2015). Effect of vegetation on the temporal stability of soil moisture in grass-stabilized semi-arid sand dunes. *Journal of Hydrology*, *521*, 447–459. doi: 10.1016/j.jhydrol.2014.12.037

Wei, L., Dong, J., Gao, M., & Chen, X. (2017). Factors Controlling Temporal Stability of Surface Soil Moisture: A Watershed-Scale Modeling Study. *Vadose Zone Journal*, *16*(10), vzj2016.12.0132. doi: 10.2136/vzj2016.12.0132

Wollschläger, U., Attinger, S., Borchardt, D., Brauns, M., Cuntz, M., Dietrich, P., ... Zacharias, S. (2017). The Bode hydrological observatory: A platform for integrated, interdisciplinary hydro-ecological research within the TERENO Harz/Central German Lowland Observatory. *Environmental Earth Sciences*, *76*(1), 29. doi: 10.1007/s12665-016-6327-5

Xu, S., & Cheng, J. (2021). A new land surface temperature fusion strategy based on cumulative distribution function matching and multiresolution Kalman filtering. *Remote Sensing of Environment*, *254*, 112256. doi: 10.1016/j.rse.2020.112256

Yao, P., Lu, H., Shi, J., Zhao, T., Yang, K., Cosh, M. H., ... Entekhabi, D. (2021). A long term global daily soil moisture dataset derived from AMSR-E and AMSR2 (2002–2019). *Scientific Data*, *8*(1), 143. doi: 10.1038/s41597-021-00925-8

Zanella, P. G., de Carvalho, C. A. B., Ribeiro, E. T., Madeiro, A. S., & Gomes, R. D. S. (2017). Optimal quadrat area and sample size to estimate the forage mass of stargrass. *Semina: Ciências Agrárias*, *38*(5), 3165. doi: 10.5433/1679-0359.2017v38n5p3165

Zappa, L., Forkel, M., Xaver, A., & Dorigo, W. (2019). Deriving Field Scale Soil Moisture from Satellite Observations and Ground Measurements in a Hilly Agricultural Region. *Remote Sensing*, *11*(22), 2596. doi: 10.3390/rs11222596

Zhang, D., Zhang, W., Huang, W., Hong, Z., & Meng, L. (2017). Upscaling of Surface Soil Moisture Using a Deep Learning Model with VIIRS RDR. *ISPRS International Journal of Geo-Information*, *6*(5), 130. doi: 10.3390/ijgi6050130

Zhao, Y., Peth, S., Wang, X. Y., Lin, H., & Horn, R. (2010). Controls of surface soil moisture spatial patterns and their temporal stability in a semi-arid steppe. *Hydrological Processes*, *24*(18), 2507–2519. doi: 10.1002/hyp.7665

Zhu, X., He, Z., Du, J., Chen, L., Lin, P., & Tian, Q. (2021). Spatial heterogeneity of throughfall and its contributions to the variability in near-surface soil water-content in semiarid mountains of China. *Forest Ecology and Management*, *488*, 119008. doi: 10.1016/j.foreco.2021.119008

Zhu, X., Shao, M., & Liang, Y. (2018). Spatiotemporal characteristics and temporal stability of soil water in an alpine meadow on the northern Tibetan Plateau. *Canadian Journal of Soil Science*, CJSS-2017-0078. doi: 10.1139/CJSS-2017-0078

Zhuang, R., Zeng, Y., Manfreda, S., & Su, Z. (2020). Quantifying Long-Term Land Surface and Root Zone Soil Moisture over Tibetan Plateau. *Remote Sensing*, *12*(3), 509. doi: 10.3390/rs12030509