https://doi.org/10.5194/hess-2019-549
Preprint. Discussion started: 18 November 2019

Hydrology and Earth System Sciences
Discussions

Blending SMAP, Noah and In Situ Soil Moisture Using Multiple Methods

Ning Zhang[1], Steven M. Quiring[1], Trent W. Ford[2]

[1]: Department of Geography, Ohio State University;

[2]: Illinois State Water Survey, University of Illinois at Urbana-Champaign

Corresponding author: Ning Zhang (zhang.7819@osu.edu)

https://doi.org/10.5194/hess-2019-549
Preprint. Discussion started: 18 November 2019

Hydrology and Earth System Sciences
Discussions

Abstract

Soil moisture can be obtained from in-situ measurements, satellite observations, and model simulations. This study evaluates different methods of combining model, satellite, and *in-situ* soil moisture data to provide an accurate and spatially-continuous soil moisture product. Three independent soil moisture datasets are used, including an *in situ*-based product that uses regression kriging (RK) with precipitation, SMAP L4 soil moisture, and model-simulated soil moisture from the Noah model as part of the North American Land Data Assimilation System. Triple collocation (TC), relative error variance (REV), and RK were used to estimate the error variance of each parent dataset, based on which the least squares weighting (LSW) was applied to blend the parent datasets. These results were also compared with that using simple average (AVE). The results indicated no significant differences between blended soil moisture datasets using errors estimated from TC, REV or RK. Moreover, the LSW did not outperform AVE. The SMAP L4 data have a significant negative bias (-18%) comparing with *in-situ* measurements, and *in-situ* measurements are valuable for improving the accuracy of hybrid results. In addition, datasets using anomalies and percentiles have smaller errors than using volumetric water content, mainly due to the reduced bias. Finally, the *in situ*-based soil moisture and the simple-averaged product from *in situ*-based and Noah soil moisture are the two optimal datasets for soil moisture mapping. The *in situ*-based product performs better when the sample density is high, while the simple-averaged product performs better when the station density is low, or measurement sites are less representative.

Keywords: soil moisture; in situ network; remote sensing, regression kriging; triple collocation; relative error variance;

https://doi.org/10.5194/hess-2019-549
Preprint. Discussion started: 18 November 2019

Hydrology and Earth System Sciences Discussions

1. Introduction

Soil moisture is a critical component of the climate system. It modulates the exchange of water and energy between land and atmosphere through evapotranspiration (Seneviratne et al., 2010). Soil moisture has great value for understanding and predicting soil erosion and water quality (Keesstra et al., 2016;Abbaspour et al., 2015), agricultural and water resource management (Pittelkow et al., 2015;Dobriyal et al., 2012), runoff and flooding (Brocca et al., 2010;Wanders et al., 2014), drought monitoring (Dai, 2013;Wang et al., 2011) and weather and climate forecasting (Hirschi et al., 2011;Seneviratne et al., 2010). Despite the importance of soil moisture, accurate, spatially-continuous soil moisture datasets with high temporal and spatial resolution are elusive.

There are three primary sources of soil moisture information: remote sensing (RS) observations, Land Surface Models (LSMs), and in-situ measurements. Microwave remote sensing is responsive to surface (~5-cm) soil moisture in regions with sparse to moderate vegetation density. The passive microwave satellites that are currently in orbit include the Soil Moisture and Ocean Salinity (SMOS) satellite (launched 2009, 35 km resolution, Kerr et al. (2001)), the Advanced Microwave Scanning Radiometer 2 (AMSR2) (25 km resolution, Imaoka et al. (2010)) onboard the GCOM-W1 satellite, and the Soil Moisture Active Passive (SMAP) satellite (launched 2015, 9 km resolution, Entekhabi et al. (2010)). The Advanced Scatterometer-A/B (ASCAT-A/B) (Wagner et al., 2013) on board of the Meteorological Operational (METOP) satellite series (launched 2006 and 2012 respectively, 25 km resolution) is an active microwave satellite in orbit. The coarser spatial resolution of these sensors is compensated by their greater spatial coverage and more frequent revisit times. In contrast, the active synthetic aperture radar (SAR) systems, such as the one onboard the RADARSAT-2 satellite (launched 2007, 3 m resolution) (Lievens and Verhoest, 2012) and the ones onboard the Sentinel-1 (A/B) satellite constellation (launched in

https://doi.org/10.5194/hess-2019-549
Preprint. Discussion started: 18 November 2019

2014 and 2016, respectively, 5 m resolution) (Paloscia et al., 2013), provide soil moisture information at finer spatial resolution, but with limited spatial coverage and less frequent revisit times.

A limitation of all microwave RS soil moisture datasets is that they can only measure soil moisture in the top 5 cm (or less) of the soil due to the limited penetration depth of microwave signals. In addition, they cannot detect soil moisture under snow or ice, or in frozen soils. There are also challenges with retrievals in areas with complex topography, dense vegetation, near water bodies, or cities (Wagner et al., 1999;Parinussa et al., 2011). Ford and Quiring (2019) compared the RS soil moisture datasets from SMAP (SMAP L3 and SMAP L4), SMOS and ESA-CCI with in-situ measurements and found the SMAP L3 product consistently performed best among the four.

Model-simulated soil moisture is another source of spatially-continuous soil moisture. The NOAA Climate Prediction Center (CPC) (Huang et al., 1996), Global Land Data Assimilation System (GLDAS) (Rodell et al., 2004) and North American Land Data Assimilation System (NLDAS) (Mitchell et al., 2004) all provide simulated soil moisture operationally at various depths and time scales. Compared with in-situ measurements, Chen et al. (2013) found all four GLDAS LSMs systematically underestimate the surface soil moisture in the Tibetan Plateau. Bi et al. (2016) also found that all the GLDAS LSMs are strongly correlated with observations, but the Mosaic model consistently has larger biases than other LSMs (the largest bias reaches 0.192 m^3m^{-3}) in the Tibetan Plateau. By comparing GLDAS Noah model with the Standardized Precipitation Index and a multi-satellite surface soil moisture product, Spennemann et al. (2015) found that GLDAS Noah accurately captured the variability of soil moisture anomalies over southern South America, but the accuracy varied both regionally and seasonally. Soil moisture simulations from NLDAS phase 2 (NLDAS-2) are also found to have large biases when compared to in-situ observations

https://doi.org/10.5194/hess-2019-549
Preprint. Discussion started: 18 November 2019

(Xia et al., 2014). Specifically, the Noah and VIC models tend to overestimate soil moisture, while Mosaic and SAC models underestimate soil moisture when compared with in-situ observations (Xia et al., 2015). Ford and Quiring (2019) compared the modeled soil moisture from NLDAS-2 and CPC with in-situ measurements and that the found NLDAS-2 models consistently performed better than the CPC model.

Similar to RS soil moisture, model-simulated soil moisture is difficult to validate because of the scale mismatch and the in-situ networks are not dense enough to adequately resolve soil moisture variability within each LSM pixel. In addition, the reliability of model-simulated soil moisture varies significantly from model to model, and over time and space (Ford and Quiring, 2019;Spennemann et al., 2015). Models generally perform well in representing the variations in soil moisture and soil moisture anomalies (Downer and Ogden, 2003;Meng and Quiring, 2008;Albergel et al., 2012), but they tend to have large biases in simulating the absolute volumetric water content of the soil (Xia et al., 2015;Bi et al., 2016).

In-situ soil moisture measurements from individual field campaigns and regional and national soil moisture monitoring networks are invaluable for calibrating and validating LSMs and RS-based soil moisture datasets and other hydrological and climatological studies. Great efforts have been made to assemble, homogenize, and standardize in-situ soil moisture measurements from different networks, time frames, sensors, depths and format (Cosh et al., 2016;Dorigo et al., 2013;Zhang et al., 2017a;Ford and Quiring, 2014). Currently, the coordinated in-situ soil moisture networks include the International Soil Moisture Network (ISMN) (Dorigo et al., 2011) and the North American Soil Moisture Database (NASMD) (Quiring et al., 2016). Despite the coordinated and standardized in-situ measurements, the number of stations and networks measuring soil moisture continuously is still very limited at either regional or global scale. In addition, the small spatial representativeness of in-situ data (a point measurement) also limits its application at larger spatial scales.

To fully apply in-situ soil moisture on a continuous basis, a variety of approaches have been adopted to generate spatial continuous soil moisture datasets based on in-situ measurements. For example, Takagi and Lin (2012) used regression kriging (RK) along with five topographic variables (elevation, curvature, slope, upslope contributing area, and topographic wetness index) to generate soil moisture maps of the Shale Hills. The RMSE of RK predictions ranged from 0.03 to 0.08 m^3m^{-3} in the surface layer (0-10 cm). Yao et al. (2013) compared ordinary kriging (OK), inverse distance weighting (IDW), linear regression and regression kriging (RK) to estimate soil moisture at small catchment (2 km^2) with complex terrains. The auxiliary variables used in RK and linear regressions were land use types, slope, and annual average solar radiation. They found both OK and IDW did not perform well in complex terrains, while RK performed best with mean Nash-Sutcliffe Efficiency (NSE) of 0.69 followed by linear regression. Yuan and Quiring (2017) compared the reduced optimal interpolation (ROI) based on in-situ and simulated soil moisture from VIC model with co-kriging and IDW methods for surface soil moisture mapping in Oklahoma, and found that ROI performs better than the other two methods with a mean NSE around 0.58 and a mean absolute error 0.03 $m^3\,m^{-3}$.

In general, the RK method performs better than other geostatistical and non-geostatistical methods (Keskin and Grunwald, 2018), but it requires a relatively high density of soil moisture measurements, strong correlations between soil moisture and auxiliary variables and the accurate measurement of auxiliary variables (Li and Heap, 2011;Keskin and Grunwald, 2018). In addition, none of the RK-gridded soil moisture datasets have been compared with the model-simulated and satellite-derived soil moisture.

https://doi.org/10.5194/hess-2019-549
Preprint. Discussion started: 18 November 2019

Hydrology and Earth System Sciences
Discussions

In summary, each source of soil moisture data has its strengths and weaknesses. However, none of them, at least by themselves, are adequate for providing accurate soil moisture data at high temporal and spatial resolutions. Therefore, it is useful to combine these three independent data sources to capitalize on the strengths of each and to generate an optimal soil moisture product to facilitate real-world applications. There are a number of methods that are commonly used for blending together different soil moisture datasets, including triple collocation (TC) (Stoffelen, 1998) with least square weighting (LSW) and simple averaging (averaging parent datasets using equal weighting). For example, Yilmaz et al. (2012) generated a hybrid soil moisture anomaly product at 0.25° grid by merging model-derived soil moisture, thermal infrared RS-soil moisture, and microwave RS-based soil moisture using TC and LSW. The TC-merged product had less uncertainty, but it did not outperform the simple averaging method. Zeng et al. (2016) also used TC with LSW to blend the soil moisture from two satellite (AMSRE and ASCAT) and one reanalysis soil moisture product (ERA-Interim). Their merged product performed better than simple averaging in the sub-humid and semi-arid regions, but the performances of TC with LSW and simple averaging were similar in arid regions.

There are a number of knowledge gaps that still exist, including (1) of the lack of in-situ soil moisture inclusion in product blending. Current studies mainly focus on combining modeled and RS soil moisture, rather than combining all three sources (modeled, RS and in-situ). In-situ measurements can be useful for improving the accuracy of hybrid soil moisture datasets. (2) There is no comprehensive evaluation of different data blending methods. In addition to TC, there are a variety of other methods that are available for combining different datasets such as Kriging and Relative Error Variance (REV) (Vinnikov et al., 1996;Ford and Quiring, 2019). Therefore, it would be helpful to compare the accuracy of different blending methods to identify the optimal

https://doi.org/10.5194/hess-2019-549
Preprint. Discussion started: 18 November 2019

approach for soil moisture. (3) The impact of measurement units (e.g., volumetric water content, soil moisture anomalies, and percentiles) is unknown. For example, is it better to convert all of the soil moisture measurements to anomalies or percentiles before blending? (4) A simple and operational methodology is still needed for accurate daily soil moisture mapping with high spatial resolution. Current methods to generate gridded soil moisture data products cannot produce data with sufficient spatial resolution for many agricultural and hydrological applications.

This paper addresses all four of these knowledge gaps by assessing different blending methods to merge model-simulated, RS-based and in-situ soil moisture data into a 4-km soil moisture product. The impact of different measurement units (absolute, anomalies and percentiles) on the accuracy of the blended product are also investigated. Finally, two optimal datasets are identified and the utility of these datasets are demonstrated.

2. Study area and data

This study is conducted in the south-central region of the United States, covering four states, including Texas, Oklahoma, Arkansas, and Louisiana with a total area ~1,150,400 km^2. The south-central U.S. an important agricultural region in the U.S., but also one that is drought-prone (Tian and Quiring, 2019). For example, the four states account for about 10% of national winter wheat production in 2017 (National Agricultural Statistics Service). According to the Köppen climate classification, the climate of this region varies from warm temperate (about three-fourths of the region) in the east to the arid (about one-fourth of the region) in the west (Kottek et al., 2006). The annual average temperature gradually decreases from south (27 °C) to north (13 °C), and the mean annual precipitation gradually increases from west (<25 cm) to east (>190 cm).

This study uses in-situ measurements of soil moisture, satellite-observed soil moisture, model-simulated soil moisture, precipitation and air temperature data. Detailed information on the

https://doi.org/10.5194/hess-2019-549
Preprint. Discussion started: 18 November 2019

spatial and temporal resolution, period of record and measurement depths are listed in Table 1. To facilitate comparison, the common period of record from 2015/03/31 to 2018/12/31 (SMAP product) were extracted for all datasets.

Table 1. Summary of datasets used in this study

Data Sources		**Temporal Domain**	**Temporal resolution**	**Spatial resolution***	**Measurement depths (cm)**
In-situ	OKM	1998-present	Daily	115 out of 129 sites	5, 25, 60, 75
	WTM	2002-present	Daily	64 out of 70 sites	5, 20, 60, 75
	SCAN	1994-present	Daily	21 out of 219 sites	5, 10, 20, 50, 100
	CRN	2009-present	Daily	15 out of 147 sites	5, 10, 20, 50, 100
SMAP L4		2015/03/31 - present	3 hours; latency: 2 days	9-km	5, 0-100
NLDAS_V2 Noah Model		1979-present	Hourly	1/8°	0-10, 10-40, 40-100
PRISM	ppt, tmp	1981-present	Daily	4-km	

*: The in-situ measurements are point-based, thus the spatial resolution for in-situ data refers to the number of stations used in this study out of a total number of stations of the sparse network.

2.1 In-Situ Soil Moisture Measurements

The in-situ soil moisture data are collected from four sparse networks: Oklahoma Mesonet (OKM), West Texas Mesonet (WTM), Soil Climate Analysis Network (SCAN) and Climate Reference Network (CRN). Daily soil moisture measurements were obtained from North American Soil Moisture Database (NASMD) in the units of volumetric water content (m^3 m^{-3}) (Quiring et al., 2016). Since different networks collect data at different time intervals ranging from every 5 minutes to once per day, for consistency a single morning measurement (7 am LST) is extracted to represent the daily value. This is not ideal, but it is reasonable for applications in which diurnal variations in soil water content are inconsequential, such as drought monitoring. The raw

https://doi.org/10.5194/hess-2019-549
Preprint. Discussion started: 18 November 2019

measurements have passed through the Quality Assurance and Quality Control (QAQC) process (Ford and Quiring, 2014), with dubious or questionable values been removed and filled. The near-surface measurements (5 cm) from a total of 215 stations (Fig. 1) were obtained for this study.

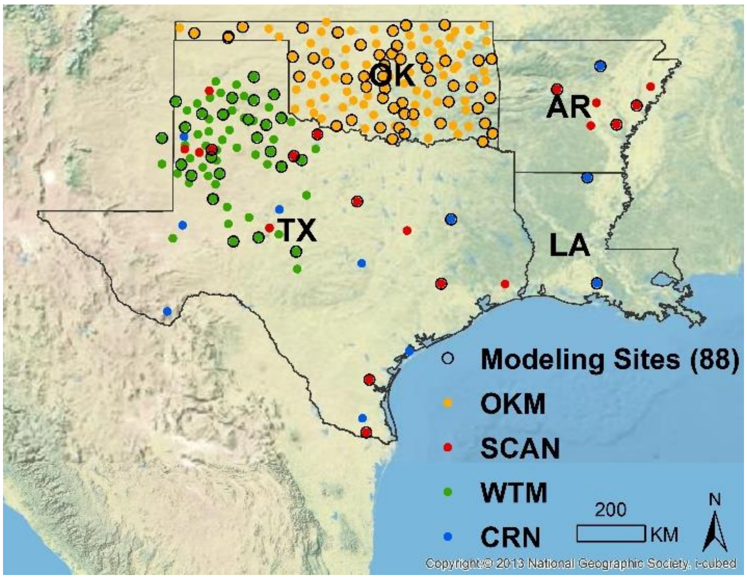

Fig. 1 Study area and stations for in-situ soil moisture measurements. Background map (USA Topographic Basemap) Copyright:© 2013 National Geographic Society, i-cubed

2.2 SMAP-L4 Soil Moisture

The SMAP Level-4 Surface and Root-Zone Soil Moisture are adopted in this study because it provides a temporally complete set of global soil moisture data. The SMAP L4 product is a merged soil moisture product from SMAP L-band brightness temperature observations and estimates from the NASA Catchment land surface model using a data assimilation system (Reichle et al., 2018). The L4 Geophysical Data are used, which are available from 31 March, 2015 to present (with 2-3 days latency). They include both surface (0-5 cm) and root-zone (0-100 cm) soil moisture every 3 hours at a spatial resolution of 9-km. The unbiased RMSE for SMAP L4_SM surface and root zone soil moisture are reported to be 0.038 $m^3\ m^{-3}$ and 0.030 $m^3\ m^{-3}$ respectively

https://doi.org/10.5194/hess-2019-549
Preprint. Discussion started: 18 November 2019

(Reichle et al., 2017). Finally, to be consistent with the in-situ measurements, the SMAP L4 product with time slot covering 7 am are extracted each day to represent the daily soil moisture from 2015/03/31 to 2018/12/31. The nearest neighbor assignment is used to resample SMAP L4 surface soil moisture from 9-km to 4-km to match the spatial resolution of other datasets (e.g., PRISM).

2.3 NLDAS-2 Noah Soil Moisture

This study uses the simulated soil moisture from the NLDAS-2 Noah model. The Noah model provides hourly soil moisture fields at 1/8° grid from 1979 to present. The Noah model has four soil layers: 0–10 cm, 10–40 cm, 40–100 cm, and 100–200 cm, but only the top layer is used in this study. Details about the NLDAS-2 configuration of the Noah LSM can be found in Xia et al. (2012). To be consistent with the in-situ measurements, the Noah output at 7 am are extracted each day to represent the daily soil moisture, and the data from 2015 to 2018 are adopted to match the record length of the SMAP data. Finally, the nearest neighbor method is used to resample the simulated soil moisture from 12.5-km to 4-km to match with other datasets.

2.4 PRISM Climate Data

The PRISM (Parameter-elevation Relationships on Independent Slopes Model) datasets are developed by Oregon State University's PRISM Climate Group (Daly et al., 2008). They are official climatological data sets of the USDA. PRISM use surface stations and a weighted regression scheme to generate daily updated spatial mapping of climate variables (e.g., precipitation, temperature, dew point, vapor pressure deficit) over the contiguous United States. There are more than 13,000 quality controlled surface stations used for precipitation interpolation and more than 10,000 stations used for temperature interpolation (Daly et al., 2008). The

https://doi.org/10.5194/hess-2019-549
Preprint. Discussion started: 18 November 2019

climatological normals (average monthly and annual conditions over 1981-2010) and monthly and daily data are available at 4-km and 800-m resolution from 1981 to present.

The 4-km daily precipitation from PRISM are used in this study. Since there is strong coupling between soil moisture and precipitation (Koster et al., 2004), precipitation has been widely used as an important input for soil moisture estimation in various LSMs (Liu et al., 2018;Xia et al., 2012;Liang et al., 1996). Here, the Antecedent Precipitation Index (API) is calculated based on precipitation and adopted for soil moisture interpolation using in-situ measurements and regression kriging. The API index is introduced in Section 2.5.2 and regression kriging is introduced in Section 3.

2.5 Data Preparation

2.5.1 Anomalies and Percentiles

The volumetric water content of the soil varies as a function of weather conditions, soil characteristics, vegetation, topography, among other factors, and so it cannot be directly compared between different locations. In contrast, relative measures of soil wetness, such as anomalies and percentiles can used to standardize soil moisture from different sensors and locations and make them comparable (Ford et al., 2015;Zhang et al., 2017a). In this study, anomalies and percentiles are calculated for all 3 datasets (In-situ, SMAP, NLDAS and PRISM). Anomalies are calculated by removing the seasonal climatology from the absolute soil moisture at each day (Crow and Van den Berg, 2010). The climatological mean is calculated using a moving-window approach (Chen et al., 2017), which averages all available soil moisture estimates across all years within a 31-day window (Dong et al., 2018) centered on the target day.

Percentiles are calculated using an empirical probability distribution function and moving window approach as well. At each day of the year, all the data fall within a 31-day window centered

https://doi.org/10.5194/hess-2019-549
Preprint. Discussion started: 18 November 2019

on that day was used to construct the empirical probability distribution function. Ford et al. (2016) found sample sizes of 93 to 186 daily soil moisture observations were required to generate robust percentiles. In our case, SMAP has the shortest data record (3 years), thus has 93 data points (31 days in window × 3 years) from which to build the distribution and compute the quartiles and percentiles. This has met the sampling size to generate robust percentiles. For other datasets, the total length of records is used to generate the percentiles (e.g., 20 years for in-situ, 40 years for NLDAS). Percentiles range from 0 or (0%) to 1 (or 100%), which corresponds to the driest (0%) and wettest (100%) soil conditions at a specific site over the entire study period.

2.5.2 Antecedent Precipitation Index (API)

The API is precipitation-based moisture index. It is used to indicate the wetness of a location and has been widely applied in drought monitoring (Crow et al., 2012a), runoff forecasting (Anctil et al., 2004), soil moisture estimation (Ochsner et al., 2019) and crop yield prediction (Zhang et al., 2017b). API takes preceding precipitation into account to estimate the current moisture status, and is formulated as (Kohler and Linsley, 1951):

$$API(i) = \mathrm{API(i-1)} * \mathrm{k} + \mathrm{PPT(i)} \qquad (1)$$

Where API(i) is the API at day i, PPT(i) is the precipitation occurring on day i; k is an empirical decay factor between 0.80 and 0.98 (Heggen, 2001). In this study, a set of k values (from 0.80 to 0.99) is tested to determine the optimal k value that results in the highest correlation between API and soil moisture based on 215 stations. Fig. S1 shows the variation in correlation as a function of different k values. The highest correlation (r = 0.4) is achieved at k = 0.92. Therefore, k = 0.92 is used in this study for API calculation.

https://doi.org/10.5194/hess-2019-549
Preprint. Discussion started: 18 November 2019

2.5.3 Site Selection

In this study, 40% of the stations with soil moisture measurements (88 sites) are used for modeling (black circles in Fig. 1), while the remaining 60% of stations (127 sites) are used for out-of-sample validation. The 88 modeling sites are selected based on the Index of Temporal Stability (ITS) (Jacobs et al., 2010;Zhao et al., 2010). ITS is an indicator of the temporal representative locations. The location with the lowest ITS value is the location with the highest temporal stability. The ITS at location i (ITS_i) is calculated as:

$$ITS_i = \sqrt{MRD_i{}^2 + SDRD_i{}^2} \quad (2)$$

$$MRD_i = \frac{1}{\mathrm{T}}\sum_{j=1}^{\mathrm{T}} RD_{i,j} \quad (3)$$

$$SDRD_i = \sqrt{\frac{1}{\mathrm{T}-1}\sum_{j=1}^{T}\left(RD_{ij} - MRD_i\right)^2} \quad (4)$$

$$RD_{ij} = \frac{\theta_{ij} - \overline{\theta_J}}{\overline{\theta_J}} \quad (5)$$

Where θ_{ij} is individual daily measurement of soil moisture at location $i \subset [1, \ldots, N]$ and time j $\subset [1, \ldots, T]$, and $\overline{\theta_J}$ is the spatial average of soil moisture at all locations at time j. $RD_{i,j}$ is the relative difference of location i at time j, which is introduced by Vachaud et al. (1985). MRD_i is the mean relative difference of location i. It averages the RD at location i across an entire period (T days), and represents the location's temporal bias or whether the location is wetter or drier than the average of the area during T days. $SDRD_i$ is the standard deviation of the RD at location i. It describes the degree of the temporal stability of a location, or whether a location is temporally stable. Therefore, a temporally representative site is one with a small mean bias and can be characterized by low values of both MSD and SDRD, and a low value of ITS (Cho and Choi, 2014;Penna et al., 2013;Brocca et al., 2012).

https://doi.org/10.5194/hess-2019-549
Preprint. Discussion started: 18 November 2019

Since anomalies can be negative, the absolute value of the difference between $A_{i,j}$ and $\overline{A_J}$ (Eq. 6) is adopted to represent the relative difference of anomalies at location i and time j (RD_A_{ij}) (Wang et al., 2017;Mittelbach and Seneviratne, 2012):

$$RD\_A_{ij} = |A_{i,j} - \overline{A_J}| \qquad (6)$$

where $\overline{A_J} = \frac{1}{N}\sum_{i=1}^{N} A_{ij}$, indicates the spatial average of anomalies of all stations at time j.

In this study, the 88 modeling sites are selected by three steps: (1) Calculate and rank the ITS of 215 stations in ascending order; (2) Evenly divide the ranked ITS into four groups; (3) Within each group, select the 22 sites with the smallest ITS values. The 88 sites are selected in this way to ensure an evenly sampled sites across the ITS range, which best mimic the reality that in-situ stations have different temporal representativeness. Although ITS ranking using absolute soil moisture, anomalies and percentiles are not exactly the same, the differences are minor. To be consistent across datasets and facilitate comparison, the same 88 (127) sites selected using the absolute soil moisture were used for calibration (validation) using soil moisture anomalies and percentiles, because the selected sites are also evenly distributed within the ITS range calculated by anomalies and percentiles (Fig. S2).

3. Blending Methods

The soil moisture blending schemes used in this study are summarized in Fig. 2. Two categories of parent datasets are adopted in this study (Fig. 2). The first category is consist of SMAP observations, NLDAS simulations, and RK-gridded soil moisture using in-situ soil moisture and API. The datasets from the first category are implemented with three data formats: absolute values (PP1), anomalies (PP2), and percentiles (PP3). The second category consists of

https://doi.org/10.5194/hess-2019-549
Preprint. Discussion started: 18 November 2019

the RK-gridded soil moisture (PP4) using in-situ soil moisture with the absolute values of API, SMAP, and NLDAS soil moisture respectively.

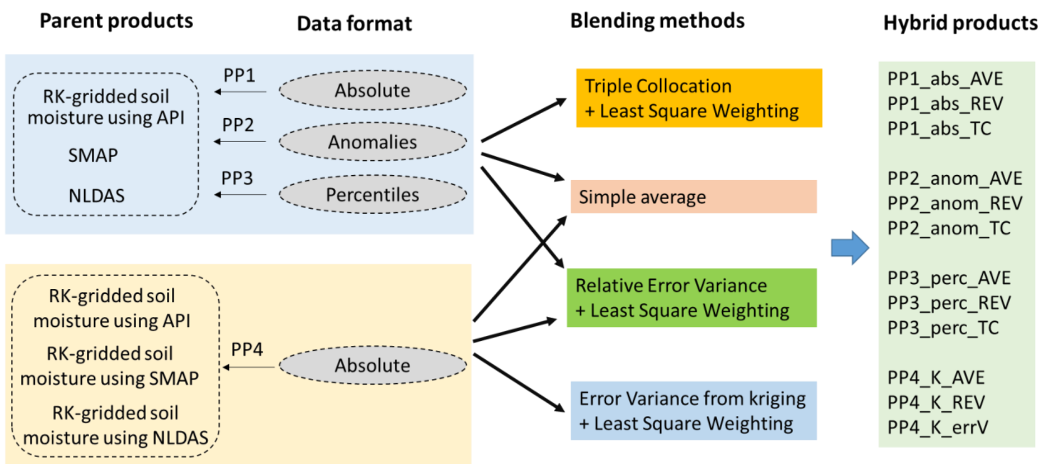

Fig. 2 Framework of soil moisture blending, and their associated parent datasets, data format, blending methods and output products.

3.1 Regression Kriging (RK)

Regression Kriging (RK) is one of the most popular and robust hybrid spatial interpolation techniques in the digital mapping of soil properties (Keskin and Grunwald, 2018). RK combines a regression between the target variable and auxiliary variables with simple kriging of the regression residuals (Hengl et al., 2007;Odeha et al., 1994). Previous studies revealed RK often outperforms non-geostatistical methods (Mishra et al., 2010;Yang et al., 2019;Li and Heap, 2011), ordinary kriging (Hengl et al., 2004), and co-kriging (Eldeiry and Garcia, 2010). The RK models can be expressed as two parts (Hengl et al., 2004):

$$\hat{z}(s_0) = \hat{m}(s_0) + \hat{e}(s_0) \quad (7)$$

Where $\hat{m}(s_0)$ is the fitted trend, $\hat{e}(s_0)$ is the interpolated residual. In this study, the trend term $\hat{m}(s_0)$ is fitted by a linear model between the auxiliary variable and soil moisture.

$$\hat{m}(s_0) = \hat{\beta} \cdot q(s_0) \quad (8)$$

https://doi.org/10.5194/hess-2019-549
Preprint. Discussion started: 18 November 2019

Where, $\hat{\beta}$ is the estimated model coefficients using generalized least squares, $q(s_0)$ is the auxiliary variable (e.g., API) at the target location s_0. The residual from the linear model is then interpolated by simple kriging with an assumed 0 mean.

$$\hat{e}(s_0) = \sum_{i=1}^{n} \lambda_i \cdot e(s_i) \quad (9)$$

Where, λ_i are kriging weights determined by the spatial dependence structure of the residual, and $e(s_i)$ is the residual at location s_i. By adding the kriging residuals to the predicted trend, the final RK prediction are obtained. RK also provide the error estimation of predicted values as (Hengl et al., 2007):

$$\sigma_{RK}^2(s_0) = (C_0 + C_0) - c_0^T \cdot C^{-1} \cdot c_0 + (q_0 - q^T \cdot C^{-1} \cdot c_0)^T \cdot (q^T \cdot C^{-1} \cdot q)^{-1} \cdot (q_0 - q^T \cdot C^{-1} \cdot c_0) \quad (10)$$

Where C is the covariance matrix of the residuals, $C_0 + C_0$ is the sill variation, c_0 is the vector of covariance of residuals at the unvisited locations, q is a matrix of predictors at the sampling locations, and q_0 is the vector of p+1 predictors (p=1 in our case).

In this study, two sets of auxiliary variables are tested for RK. The first set of auxiliary variables are API in the format of absolute values, anomalies, and percentiles respectively. Given that precipitation is the chief driver of soil moisture, and a strong positive relationship was observed between the soil moisture and API over the contiguous United States (Fig. S3), thus API can be used as an important predictor of soil moisture variations. The second set of auxiliary variables are respectively the SMAP L4 and the NLDAS surface soil moisture in the format of volumetric water content (Fig. 2).

3.2 Triple collocation (TC)

Triple collocation (TC) is a technique for estimating the error variance (errVar, $m^3\ m^{-3}$) of three independent datasets with respect to the unknown truth (Stoffelen, 1998). It assumes a linear error model (Eq. 11- Eq.13) between each product and the unknown truth (t). The errors from the independent sources are assumed to have zero mean ($E(e_i) = 0$) and are uncorrelated with each other ($Cov(e_i, e_j) = 0,\ i \neq j$) and with the truth ($Cov(e_i, t) = 0$). TC analysis has been widely used to estimate the errors of various measurement systems, such as the ocean waves (Caires and Sterl, 2003), wind fields (McColl et al., 2014;Stoffelen, 1998), leaf area index (Fang et al., 2012), precipitation (Roebeling et al., 2012), and soil moisture (Su et al., 2014;Yilmaz et al., 2012;Dorigo et al., 2010;Gruber et al., 2013). Gruber et al. (2016) reviewed the previous TC analysis on soil moisture, and found there are two different notations of TC formula, the difference notation (Stoffelen, 1998;Scipal et al., 2008;Yilmaz et al., 2012) and the covariance notation (Stoffelen, 1998;McColl et al., 2014). They demonstrated that two different notations are mathematically identical in the ideal case that each product is bias-free ($\alpha_i = 0$ in Eq. 11 to Eq. 13). However, in reality, there is always bias in each product, which results in a slightly different value of errVar estimated using the two notations. The difference notation format of TC accounts for the total errVar (including variance from both bias α_i and error term e_i), while the covariance notation of TC only focuses on the errVar from the error term (e_i). In this study, the difference notation of TC is adopted to account for the total error variance of parent datasets.

$$\theta_K = \alpha_K + \beta_K \theta_t + e_K \quad (11)$$

$$\theta_S = \alpha_S + \beta_S \theta_t + e_S \quad (12)$$

$$\theta_N = \alpha_N + \beta_N \theta_t + e_N \quad (13)$$

Where, θ_i (i ∈ (S, N, K)) are three collocated soil moisture datasets for SMAP, NLDAS and RK-gridded soil moisture, respectively; θ_t is the unknown true soil moisture; α_i (i ∈ (K, S, N)) and β_i (i ∈ (K, S, N)) are systematic additive and multiplicative biases of product i with respect to the truth, and e_i (i ∈ (K, S, N)) are the additive zero-mean random errors for each system. When anomalies or percentiles are used, the additive bias α_i can be deemed as zero, because these two methods either removed the climatology mean from each product or standardized each product.

A reference dataset must be selected from the three input datasets and rescaling is required to transfer the other two datasets into the same observation space of the reference dataset. Our preliminary results showed that the choice of reference dataset did not impact the final results, thus the RK-gridded soil moisture is selected as the reference dataset in this study ($\theta_R = \theta_K$). The rescaling method (Eq. 14) from Dorigo et al. (2010) is used.

$$\theta_i^* = \overline{\theta_R} + \sqrt{\frac{VAR(\theta_R)}{VAR(\theta_i)}} \cdot (\theta_i - \overline{\theta_i}) \quad (14)$$

Where, $\overline{\theta_R}$ and $VAR(\theta_R)$ are respectively the mean and variance of the reference soil moisture. After the rescaling of the parent datasets, Eq. (11) to (13) can be rewritten as:

$$\theta_K^* = \beta_K \theta_t + e_K^* \quad (15)$$

$$\theta_S^* = \beta_K \theta_t + e_S^* \quad (16)$$

$$\theta_N^* = \beta_K \theta_t + e_N^* \quad (17)$$

where, θ_i^* (i ∈ (K, S, N)) are the rescaled soil moisture datasets. Finally, the error variances can be estimated by averaging the cross-multiplied differences between the three datasets:

$$\sigma_K^{*2} = \overline{(\theta_K^* - \theta_S^*)(\theta_K^* - \theta_N^*)} \quad (18)$$

$$\sigma_S^{*2} = \overline{(\theta_S^* - \theta_K^*)(\theta_S^* - \theta_N^*)} \quad (19)$$

$$\sigma_N^{*2} = \overline{(\theta_N^* - \theta_S^*)(\theta_N^* - \theta_K^*)} \quad (20)$$

Different combinations of triplets are also tested in this study to examine the impact of triplets on TC estimates. The triplet candidates include the in-situ measurements (denoted by "I" in the following figure and text), the SMAP L-4 surface soil moisture (denoted by "S"), simulated soil moisture from NLDAS-2 Noah model (denoted by "N"), and the RK-gridded soil moisture

https://doi.org/10.5194/hess-2019-549
Preprint. Discussion started: 18 November 2019

using regression kriging (denoted by "K"). Four combinations of the candidates are tested, including (I, S, N), (I, K, S), (I, K, N), and (K, S, N). The triplet candidates are extracted from the 127 out-of-sample stations. Scipal et al. (2008) found at least 100 collocated triplet samples are required for a reliable estimation of the variance. In our case, the time series from 2015/03/31 to 2018/12/31 is used for TC analysis, which results in 1372 collocated triplet samples at every station with a serially complete record. The stations with less than 100 observations are removed from the TC error estimation.

3.3 Relative Error Variance (REV)

Relative Error Variance (REV) is the ratio of measurement error variance to real soil moisture variance. It measures the displacement of autocorrelation in a measured quantity. Delworth and Manabe (1988) recognized that a soil moisture time series behaves like a first-order Markov process. Later Vinnikov and Yeserkepova (1991) validated and confirmed this and noted the autocorrelation function of soil moisture can be expressed as an exponential form of lag length:

$$\gamma(\tau) = \exp(-\tau/T) \qquad (21)$$

where $\gamma(\tau)$ is the autocorrelation function, τ is the lag, and T is the decay time scale. Robock et al. (1995) also found a linear best fit of $\ln(\gamma)$ verse τ does not cross zero at a value of $\gamma(\tau = 0) = 1$. The displacement of the autocorrelation $\gamma(\tau)$ at $\tau = 0$ is related to the measurement error (a) as:

$$\gamma(\tau = 0) = 1 - a \qquad (22)$$

Successively, Vinnikov et al. (1996) partitioned the soil moisture variation into red noise (σ^2, actual variance of the soil moisture measurement) and white noise (δ^2, random error of measurements), and noted that the ratio δ^2/σ^2 can be used as a measure of the random error in the measurement. Dirmeyer et al. (2016) related the ratio δ^2/σ^2 to the measurement error (a), and used it to estimate the random measurement error of different sparse soil moisture networks. Ford

https://doi.org/10.5194/hess-2019-549
Preprint. Discussion started: 18 November 2019

and Quiring (2019) applied relative error variance (REV) to quantify the proportion of measurement error within real soil moisture variance, as:

$$REV = \frac{\delta^2}{\sigma^2} = \frac{a}{(1+a)} \qquad (23)$$

Thus, a higher REV value represents a larger proportion of random measurement error. REV is a powerful measurement of random measurement error or uncertainties, and it does not require independent data, unlike the TC method.

3.4 Least square weighting (LSW)

Yilmaz et al. (2012) adopted the least square framework to achieve an objective blending of satellite and modeled soil moisture. The same methodology was adopted by Zeng et al. (2016) to merge the satellite and reanalysis soil moisture data. In this study, the least square weighting (LSW) is used to blend soil moisture data from satellite, model and in-situ measurements based on error variances estimated from the TC, REV and RK methods, respectively (Fig. 2). The desired estimate of soil moisture (S_m) via blending different sources of data using least squares framework, can be expressed as:

$$S_m = w_x S_x + w_y S_y + w_z S_z \qquad (24)$$

Where, w_x, w_y and w_z are the relative weights of three parent datasets S_x, S_y and S_z respectively. Then a cost function (J) is constructed using the weights and the error variance of the parent datasets, such that:

$$J = \sigma_m^2 = w_x^2 \sigma_x^2 + w_y^2 \sigma_y^2 + w_z^2 \sigma_z^2 \qquad (25)$$

Where, σ_x^2, σ_y^2, and σ_z^2 are the estimated error variance for the three parent datasets. To have an unbiased estimation of S_m, the sum of weights should be 1 ($w_x + w_y + w_z = 1$), thus:

$$J = w_x^2 \sigma_x^2 + w_y^2 \sigma_y^2 + (1 - w_x - w_y)^2 \sigma_z^2 \qquad (26)$$

Finally, by minimizing the cost function and the partial derivative of function J with respect to w_x and w_y ($\frac{\partial J}{\partial w_x} = 0, \frac{\partial J}{\partial w_y} = 0$), the optimal estimation of the weights are obtained as

$$w_x = \frac{\sigma_y^2 \sigma_z^2}{\sigma_x^2 \sigma_y^2 + \sigma_x^2 \sigma_z^2 + \sigma_y^2 \sigma_z^2} \tag{27}$$

$$w_y = \frac{\sigma_x^2 \sigma_z^2}{\sigma_x^2 \sigma_y^2 + \sigma_x^2 \sigma_z^2 + \sigma_y^2 \sigma_z^2} \tag{28}$$

$$w_z = \frac{\sigma_x^2 \sigma_y^2}{\sigma_x^2 \sigma_y^2 + \sigma_x^2 \sigma_z^2 + \sigma_y^2 \sigma_z^2} \tag{29}$$

It can be seen that the weights are functions of the error variance of the parent datasets, and the product with larger error variance will be given smaller weights and vice versa. If only blending two soil moisture datasets, the least square method can be applied similar, with weights:

$$w_x = \frac{\sigma_y^2}{\sigma_x^2 + \sigma_y^2} \tag{30}$$

$$w_y = \frac{\sigma_x^2}{\sigma_x^2 + \sigma_y^2} \tag{31}$$

In this study, all three parent datasets (K, S, N) and combinations of two from them (KS, KN, SN) are tested to generate hybrid datasets. By comparing the hybrid results using all three and two out of three datasets, the optimal blending product can be identified using the least datasets while maintaining the accuracy.

3.5 Goodness of fit

In this study, 88 sites (40%) out of the total 215 stations were used for RK modeling, while the remaining 127 sites (60%) were used for out-of-sample validation. The Mean Absolute Error (MAE), Root Mean Square Error (RMSE), the Nash-Sutcliffe Efficiency (NSE) score, and the decomposition of Mean Square Error (MSE) by its mean difference (MSE_MD^2) and its pattern variation (MSE_VAR) were used for the validation and comparison of hybrid datasets. The

https://doi.org/10.5194/hess-2019-549
Preprint. Discussion started: 18 November 2019

decomposition MSE is helpful to diagnose whether the error is mainly due to the bias or variation. A detailed description of above mentioned indicators (including equations) are provided in Supplementary Text S1.

4. Results and Discussions

4.1 Patterns of parent datasets

Fig. 3 compares three statistic features, such as mean, standard deviation (STD) and coefficient of variation (CV) of the absolute values of four soil moisture datasets over 127 out-of-sample sites. The four datasets include the in-situ soil moisture measurements, the RK-gridded soil moisture from API (K-API) and the SMAP and NLDAS soil moisture. Compared with in-situ measurements, the three soil moisture datasets (K-API, SMAP and NLDAS) show an underestimation of soil moisture (Fig. 3a) with bias ratios of -9%, -18% and -8%, respectively. The large negative bias of SMAP L4 data indicates that the produce may overestimate dryness if used alone without animalization or standardization. The largest values of STD are observed for in-situ soil moisture, followed by SMAP and K-API, while NLDAS has the smallest STD values (Fig. 3b). This is reasonable since field measurements are point scale and contain more information on spatial heterogeneity and thus exhibit a higher degree of variability. As spatial resolution increases, a smoother pattern and less variability are expected. Another reason that NLDAS has the smallest STD is because the model-simulated soil moisture are solved at each grid cell using a land surface model. SMAP presents the highest (and significantly larger than others) CV among all datasets (Fig. 3c), which indicates that there is a large degree of variability in the SMAP soil moisture. The large CV of SMAP is jointly attributed to its small mean value and large STD. In contrast, the NLDAS has the smallest CV among all datasets (Fig. 3c), which is mainly due to its smallest STD among all datasets. K-API has the most comparable CV values and smaller range than that of in-situ measurements. This indicates that K-API is the product that is most similar to the in-situ measurements.

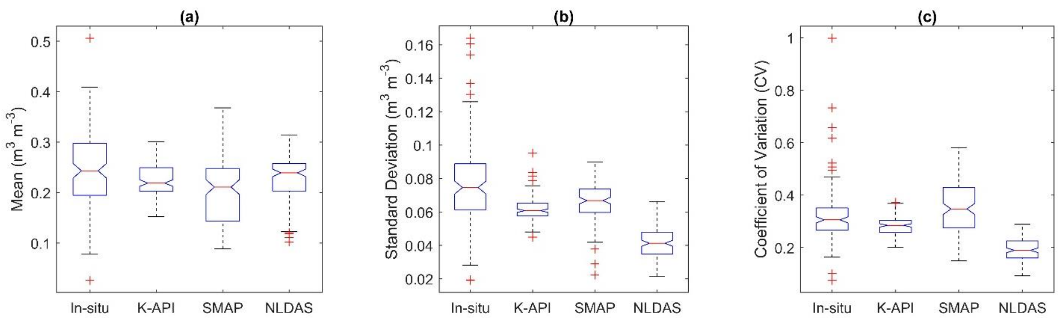

Fig. 3 Comparison of the (a) means, (b) standard deviation and (c) coefficient of variation of four datasets over 127 out-of-sample sites. The four data sets include the in-situ soil moisture measurements, the kriged soil moisture from API (K-API) and the SMAP and NLDAS soil moisture. All datasets are in the format of absolute values.

4.2 Errors of parent datasets

Fig. 4 shows the error (uncertainties) estimated using TC, REV and RK using different combinations of parent datasets. Fig. 4 (a)-(d) reveal that the error estimated from TC depends on the parent triplets used. Changing the parent triplets changes the magnitude and ranking of the parent datasets. For example, the errors estimated for in-situ data are higher when it is grouped with SMAP and NLDAS (Fig. 4a), than when it is grouped with K-API and NLDAS (Fig. 4b) or with K-API and SMAP (Fig. 4c). In addition, in-situ data have significantly higher errors than SMAP in Fig. 4(a), while Fig. 4c shows in-situ data have significantly lower errors than that of SMAP. Similar, contrasting results are also found between In-situ and NLDAS by comparing Fig. 4a and Fig. 4b. These results indicate that TC only provides a relative measure of accuracy. Yilmaz et al. (2012) also noted that TC is not ideal for capturing absolute error and can only estimate the relative error. They found that the absolute error depends on the reference dataset selected.

https://doi.org/10.5194/hess-2019-549
Preprint. Discussion started: 18 November 2019

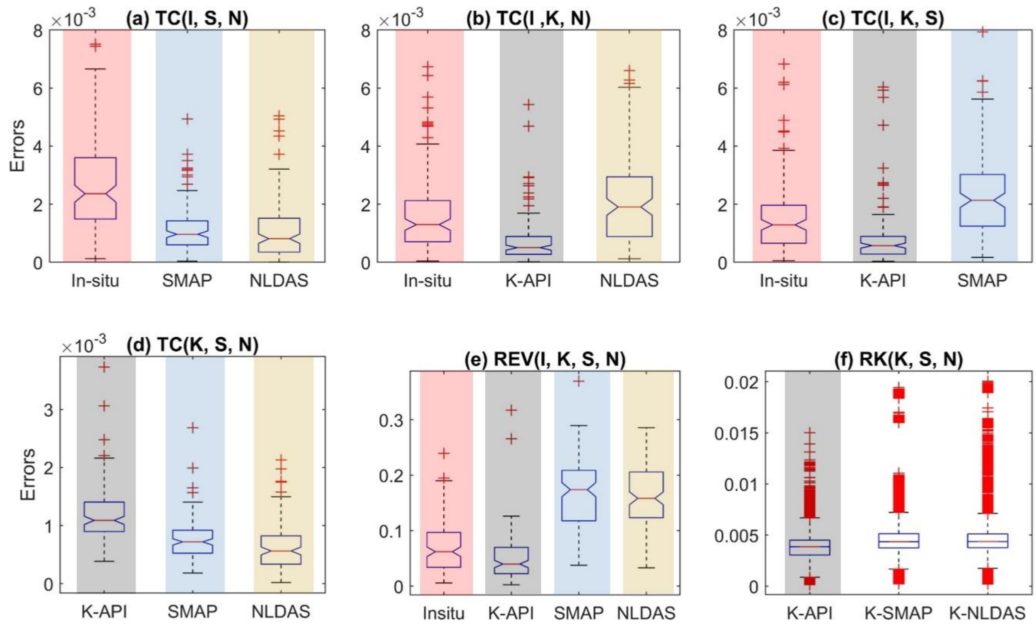

Fig. 4 (a)-(d) Estimated errors (uncertainties) using TC with different combination of four datasets (In-situ, K-API, SMAP and NLDAS soil moisture). (e) Estimated errors using REV for four datasets; and (f) Estimated errors using RK for three datasets (K-API, SMAP and NLDAS soil moisture). The datasets used here are absolute soil moisture over 127 out-of-sample stations. The same color indicates the same dataset used in TC and REV analysis. Note, both TC and REV provide one error estimation through the entire period (1372 days) for each station, while RK provides one error estimation at each day for each station. Therefore, there are 127 points within the boxplots using TC and REV (Fig. 6a-e), while there are 174244 points (127 sites*1372 days) within the boxplots using RK (Fig. 6f).

Moreover, our study reveals that the relative errors of a single soil moisture dataset, estimated from TC, are sensitive to the choice of input datasets (Fig. 4a-d). Thus, caution should be used when selecting the input datasets for TC analysis. In this study, K-API, SMAP and NLDAS are used for soil moisture blending with TC error estimation (Fig. 4d). The error ranking of the

https://doi.org/10.5194/hess-2019-549
Preprint. Discussion started: 18 November 2019

Hydrology and Earth System Sciences Discussions
Open Access
EGU

three datasets from TC are K-API>API>NLDAS (Fig. 4d), and the differences among the three datasets are statistically significant.

Our study also demonstrates that the measurement units (Fig. S4) do not impact the relative relationship (error ranking) between the different datasets. It is also interesting to note that the in-situ data always have relatively larger error when compared with other datasets using TC (Fig. 4a to 8c). This may due to its high spatial representativeness errors (Miralles et al., 2010;Crow et al., 2012b;Yilmaz et al., 2012). If this is true, then using in-situ data as the ground truth for validation may not be the best choice.

The REV (Fig. 4e) and RK (Fig. 4f) can also be used to estimate error in different datasets. REV (Fig. 4e) and RK (Fig. 4f) provide consistent results and both indicate that K-API has significantly smaller errors than SMAP and NLDAS, while SMAP and NLDAS are similar (i.e., there is not a statistically significant difference between the two). Although REV is a relative ratio between measurement error variance to real soil moisture variance, unlike TC, it does not depend on another dataset during calculation. Therefore, REV provides a consistent estimate for each product that does not change depending on the other datasets that are included. By comparing Fig. 4(b)-(c) with Fig. 4(e), in-situ data have larger errors than K-API based on both the TC and REV methods.

Fig. 5A and 5B illustrate the spatial distribution of errors estimated using TC (Fig. 5A) and REV (Fig. 5B) for the three parent datasets (K-API, SMAP and NLDAS). When using TC estimates, the results agree well with Fig. 4d. NLDAS has the smallest error among the three, with low errors found in the central and southwestern portions of the study area. The K-API has larger errors near the Gulf of Mexico and in Oklahoma, while SMAP has larger errors scattered throughout the study region. On the contrary, when using REV, the K-API has the smallest errors

https://doi.org/10.5194/hess-2019-549
Preprint. Discussion started: 18 November 2019

Hydrology and Earth System Sciences Discussions

among the three datasets over the entire study area, while SMAP data has larger errors in the western part of the study region and NLDAS has larger errors in the eastern part of the study region.

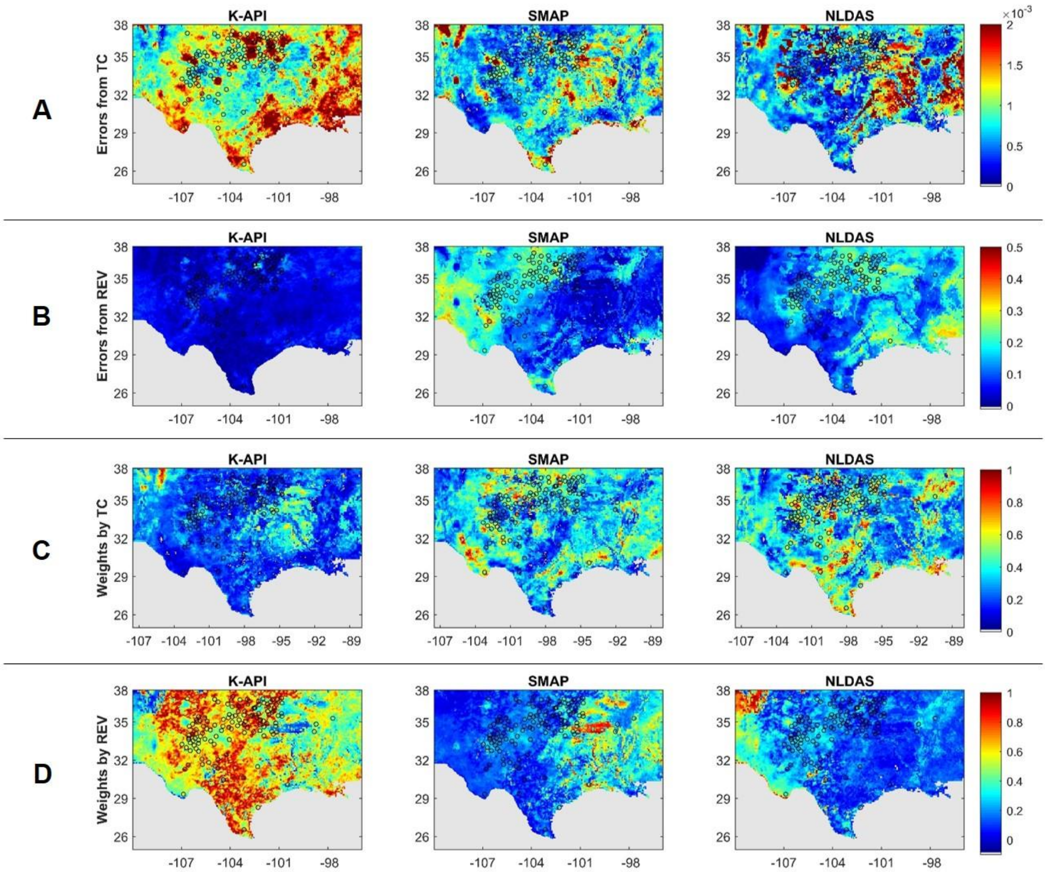

Fig. 5 Spatial maps of errors (A and B) and LSW weights (C and D) based on errors from TC and REV for each parent product. All products are in the format of absolute soil moisture. The black circles indicate the locations of 127 out-of-sample stations.

4.3 Weights of parent datasets

Fig. 5C and 5D reveal the spatial distribution of LSW weights calculated using the errors estimated from TC and REV. Fig. 6 compares the weights derived from TC, REV and RK based on 127 out-of-sample stations. Generally, larger weights are given to the product with smaller

https://doi.org/10.5194/hess-2019-549
Preprint. Discussion started: 18 November 2019

errors (Fig. 5), and the weighting of each product are not impacted by the data format used (Fig. S5). When using TC, larger weights are given to the NLDAS (median value of 0.42) and SMAP (median value of 0.32), while K-API tends to be given lower weight, with a median value about 0.2 (Fig. 6a). In contrast, higher weights are given to K-API (median value of 0.6), especially in the central part of study area, when REV is used, while smaller weights are given to SMAP and NLDAS with a median value of 0.2 for both (Fig. 6b). SMAP has higher weights in the eastern part of the study region than in the west. NLDAS is given higher weights in the western part of the study region (Fig. 5D). In general, the weighting scheme derived from TC (Fig. 5C and Fig. 6a) has patterns that are opposite to those based on REV (Fig. 5D and Fig. 6b). It is also interesting to note that the weighting scheme derived from RK (Fig. 6c) is similar to the mean weighting (0.33 weighting line in green). This analysis has demonstrated that the choice of weighting scheme can have a substantial influence on the relative weights that are assigned to each product.

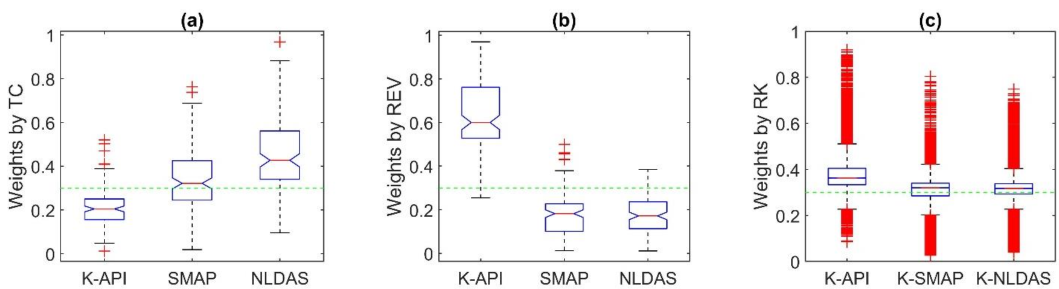

Fig. 6 Weights of each product (K-API, SMAP and NLDAS) based on LSW using errors estimated from (a) TC, (b) REV and (c) RK. The weights are calculated based on absolute soil moisture over 127 out-of-sample stations. The green lines indicate the averaging weighting, where each product is given the same weight of 0.33.

Considering each dataset has errors and the inclusion of additional datasets may increase the uncertainty, therefore it is important to evaluate whether it is necessary to use all three soil moisture datasets to achieve the highest accuracy. Therefore, we iteratively selected pairs of the

https://doi.org/10.5194/hess-2019-549
Preprint. Discussion started: 18 November 2019

Hydrology and Earth System Sciences
Open Access
Discussions

parent datasets and generated a hybrid soil moisture product. The results for each combination are provided in Fig. 7. Similar to the results from Fig. 6 and Fig. S5, the TC and REV provide opposite weighting results, and the weighting from RK is similar to the simple average (equal weight of two datasets). This analysis also demonstrates that the data format has little impact on the results, especially for REV (right column in Fig. 7).

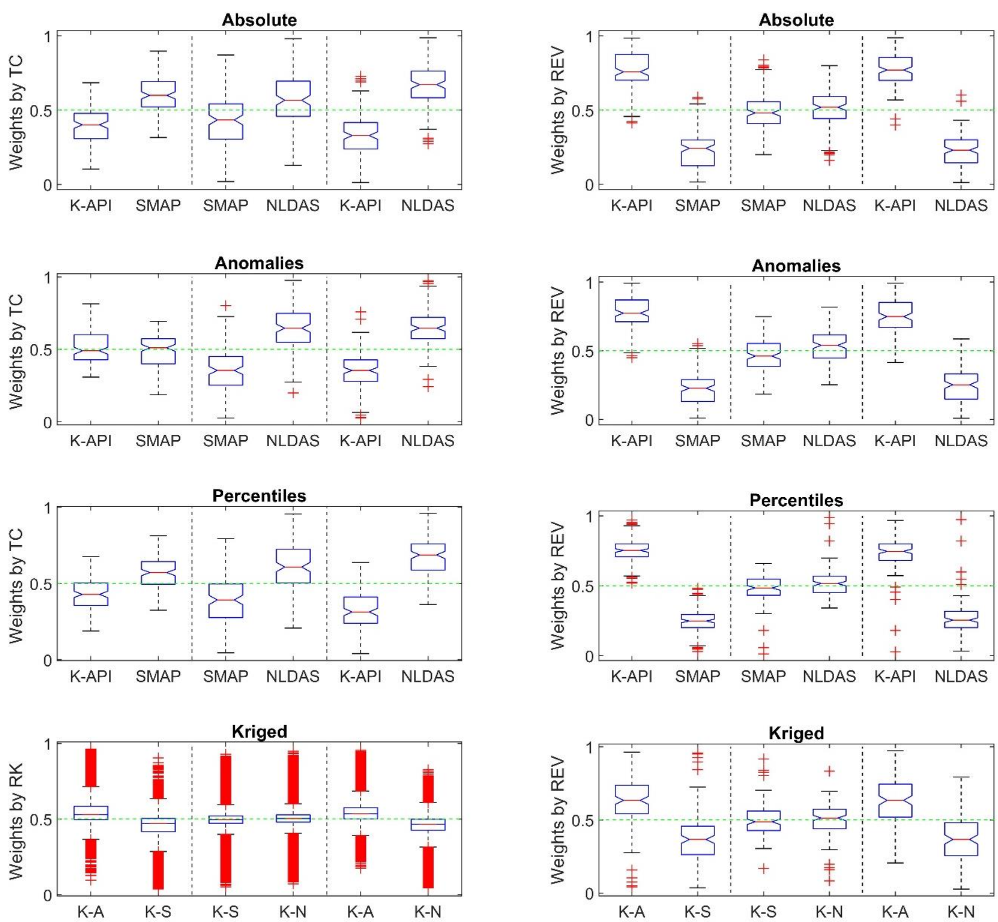

Fig. 7 Weights of soil moisture products in the format of anomalies (top row) and percentiles (bottom row) based on least square weighting using errors estimated from TC (left column) and REV (right column). The green line indicates the simple average weighting scheme with equal

https://doi.org/10.5194/hess-2019-549
Preprint. Discussion started: 18 November 2019

weight (0.5 here) for each product. K-A is short for regression kriging using API, K-S is short for regression kriging using SMAP, and K-N is short for regression kriging using NLDAS data.

4.4 Evaluation of hybrid results

Fig. 8 evaluates the hybrid results of soil moisture anomalies from different methods (simple average (AVE), REV and TC) based on MAE, RMSE, MSE_MD^2, MSE_VAR and NSE. The assessment of hybrid datasets in other formats (absolute values, percentiles and RK-gridded soil moisture) are presented in Fig. S6, S7 and S8, respectively. In terms of MAE (Fig. 8a), K-API has the smallest errors (MAE_{median} = 0.037) among the three parent datasets, while SMAP has the largest errors (MAE_{median} =0.050) and NLDAS falls in the middle (MAE_{median} =0.046). The analysis demonstrates that after blending the three parent datasets, the merged datasets do not outperform the parent products, especially in comparison to K-API. Although the MAEs of the AVE3, REV3 and TC3 are significantly smaller than that of SMAP and NLDAS (boxes' notches do not overlap in Fig. 8a), they are not statistically significantly different from K-API (overlapped notches in Fig. 8a). Our findings contrast with those of Yilmaz et al. (2012), who found that a merged soil moisture product generated from ALEXI, Noah and LPRM using TC is more accurate than the individual parent products. However, our study evaluated different parent datasets, and the TC-based weights did vary with the input datasets (Fig. 4a-d). Since the K-API was found the most accurate among the three parent products, and Yilmaz et al. (2012) did not use this product, our results are not directly comparable. But both our study and Yilmaz et al. (2012) utilized NLDAS Noah, and both studies found that the hybrid datasets that use NLDAS have smaller errors (MAEs and RMSEs).

https://doi.org/10.5194/hess-2019-549
Preprint. Discussion started: 18 November 2019

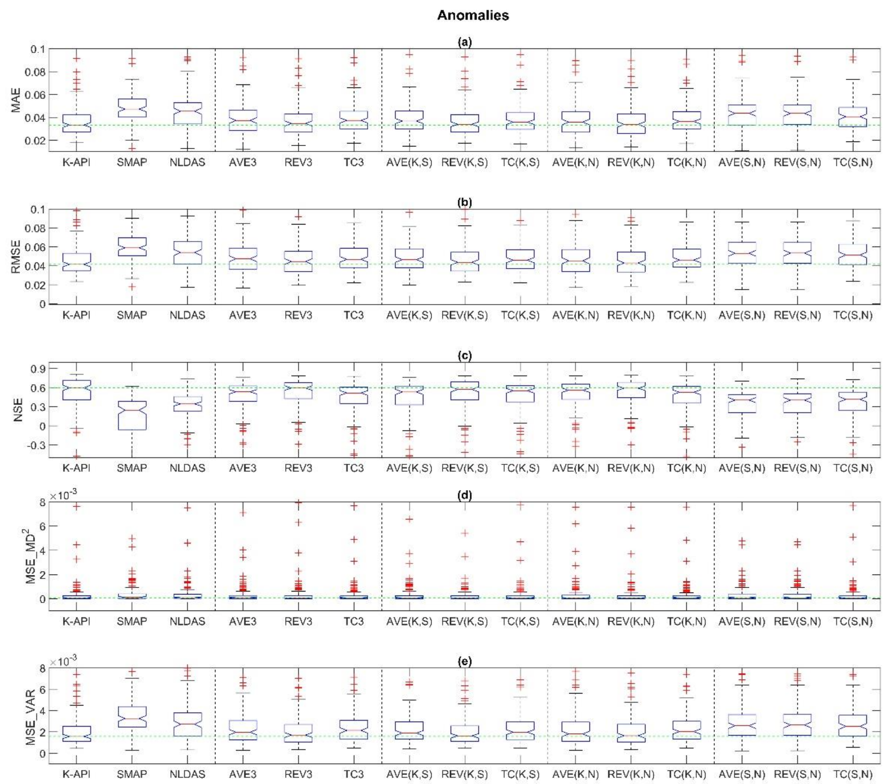

Fig. 8 Comparison of parent and hybrid products of soil moisture anomalies using different blending methods (simple average (AVE), REV- and TC-based) on (a) MAE, (b) RMSE, (c) NSE, (d) MSE_MD^2 and (e) MSE_VAR. The green line indicates the median error of K-API among 127 out-of-sample stations. AVE3, REV3 and TC3 respectively indicate the hybrid results using all three parent products based on simple average, REV and TC analysis.

The lack of significant improvement of the merged datasets verses the parent products may be attributed to (1) sub-optimal weights because neither TC and REV consider temporal variations in errors. Both TC and REV give only one error estimation at each location for the entire period. In reality, the error in each parent product likely varies both spatially (from location to location) and temporally (from day to day). Thus, the temporally fixed error estimation that is provided by TC and REV is likely not optimal. (2) The in-situ measurements cannot be considered the “truth”

https://doi.org/10.5194/hess-2019-549
Preprint. Discussion started: 18 November 2019

Hydrology and Earth System Sciences
Discussions

because they are point measurements that may not reflect the soil moisture value for each 4 km grid cell. In addition, the use of in situ measurements as truth may also be biased towards the K-API. As we found in Fig. 4, the in-situ soil moisture have large spatial representativeness errors. Even for the densest in-situ network used in this study, such as the Oklahoma Mesonet, there is only one station within each 4-km grid cell. Considering K-API is generated using in-situ soil moisture, the error patterns of K-API may follow closely with that of in-situ data, which is also confirmed by Fig. 4. This bias may result in smaller errors of K-API when evaluated using in-situ data. (3) The validation data are not spatially exhaustive. Although 60% (127) of total stations have been used in the validation, they are still relatively sparse and not evenly distributed in the study area. Fig. 1 shows most validation stations are clustered in Oklahoma and west Texas, while few stations are located in south Texas, Arkansas and Louisianan. It is possible that the places where hybrid results showed an improvement over the parent product (K-API) are not well captured using only 127 stations.

When comparing the results from the various blending methods, there is no statistically significant difference between the merged datasets using AVE, REV or TC, even though the REV-weighted datasets perform slightly better (slightly lower MAE/ RMSE, and slightly higher NSE) than other two methods. This indicates that the more complicated blending methods (LSW using TC and REV estimates) are not necessarily superior to the simple average. This result agrees with the findings from Yilmaz et al. (2012) that the merged soil moisture anomalies using LSW and TC estimates did not outperform the equally-weighted results.

Considering the different blending methods correspond to different weighting schemes (Fig. 7 and Fig. S5), this result suggests two possible conclusions: (1) if the optimal weighting (either TC- or REV-based) has been achieved, then the weighting scheme does not have a

https://doi.org/10.5194/hess-2019-549
Preprint. Discussion started: 18 November 2019

significant impact on the merged results; (2) if the optimal weight has not been achieved, then there is still an optimal weighting scheme to be identified that can significantly reduce the errors. The evaluation of hybrid datasets using the two parent datasets (Fig. 8) suggests the first conclusion (weighting scheme does not have a significant impact on the merged results) is most likely. According to Fig. 6, the weights calculated using AVE, REV and TC-based methods have covered all the possible weighting schemes of two datasets, including equal weighting (AVE) and two cases of unbalanced weighing (the product given larger weights by TC will be given smaller weights by REV). Still, no significant differences are observed when different weighting schemes are applied. In this case, the simple average (equal weighting) is recommended for soil moisture blending, as the more complicated weighting schemes do not outperform this approach.

It is also found that the combination of SMAP and NLDAS (e.g., AVE(S,N), REV(S,N) and TC(S,N)) result in a statistically significant increase in MAE values, while combining K-API with either SMAP or NLDAS has similar accuracy as the merged datasets using all three datasets. This indicates (1) incorporating three datasets may not be necessary to generate the most accurate soil moisture product and, (2) in-situ measurement is valuable for improving the accuracy of blended soil moisture datasets. K-API is the only dataset that incorporates the in-situ measurements, and it has the lowest error among all parent datasets (Fig. 8). These results are consistent when RMSE (Fig. 8b) or NSE (Fig. 8c) as considered instead of MAE.

The impact of data format on hybrid results is examined by comparing Fig. 8 with Fig. S6 to S8. It is found the MAE of hybrid datasets using anomalies (about 0.035 m^3m^{-3}) is lower than that of the absolute datasets (about 0.055 m^3m^{-3}). The hybrid datasets using anomalies and percentiles also have higher NSE (around 0.6) values than that of (around 0.3) absolute and RK-gridded datasets. The improved performance of the anomaly and percentile datasets are mainly due to the removal of bias error (MSE_MD^2). Using Eq.S2 to Eq. S4 from the Supplementary Text, the MSE can be decomposed to differences in the mean or bias (MSE_MD^2) and differences in the variance (MSE_VAR). It is found that the bias for both the soil moisture anomalies (Fig. 8d) and soil moisture percentiles (Fig. S6d) are close to zero. Therefore, most of their error is due to differences in variance (Fig. 8e and Fig. S6e). This is reasonable since both anomalies and percentiles are methods for standardizing the datasets and they are useful for removing the systematic bias between different data sets (Ford et al., 2015;Zhang et al., 2017a). In contrast, the errors of absolute soil moisture (Fig. S7d and S7e) and RK-gridded absolute soil moisture (Fig. S8d and S8e) have similar proportions of error that are due to bias (0.02 m^3m^{-3}) and variance (0.02 m^3m^{-3}). This indicates when using soil moisture data in absolute formats, the bias-related errors are present in the final datasets.

Fig. 9 shows maps of soil moisture anomalies on March 31, 2015 for each of the parent datasets (a-c) and the merged datasets based on using three products (d-f) and two products (g-i). There are distinct differences between the three parent datasets. The K-API (Fig. 9a) has a smoother pattern than the other two datasets. In addition, the in-situ anomalies (with blue dots indicate positive anomalies and red dots indicate negative anomalies) seem to match better with that of K-API. However, the differences between the maps become less distinguishable after blending the three datasets using AVE (Fig. 9d), REV- (Fig. 9e) and TC-based LSW (Fig. 9f). There is also no dramatic change of spatial patterns when changing the number of input datasets from three (Fig. 9d) to two (Fig. 9g-i) using simple average, which agrees with the results from Fig. 8.

https://doi.org/10.5194/hess-2019-549
Preprint. Discussion started: 18 November 2019

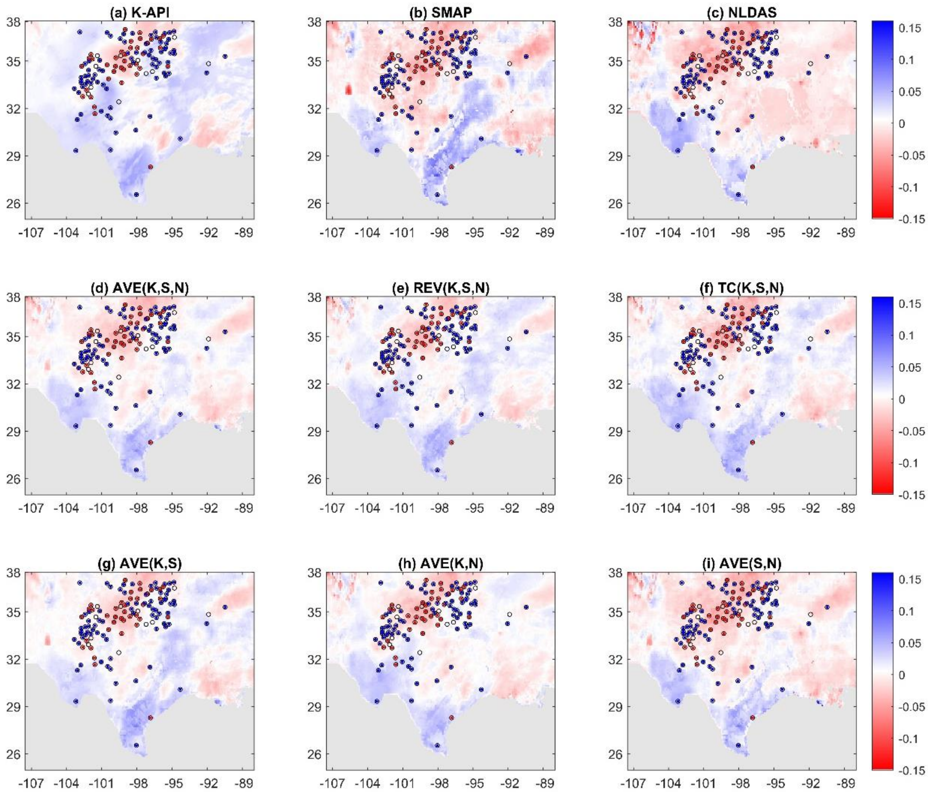

Fig. 9 Maps of parent and hybrid products of soil moisture anomalies on March 31, 2015. (a)-(c) present the three parent products (K-API, SMAP and NLDAS); (d)-(f) respectively represent the hybrid product from 3 parent products using simple average, least square weighting using REV estimated errors and TC estimated errors; (g)-(i) represent the hybrid products using simple average of two parent products. The red dots represent the in-situ stations with negative anomalies, the blue dots present the in-situ stations with positive anomalies, and the empty circles present the in-situ stations with no measurement on that day.

Finally, the spatial patterns of errors for K-API and AVE(K,N) are shown in Fig. 10. By overlaying the ITS, a consistent pattern is observed between the spatial distribution of MAE and ITS (Fig. 10). That is, sites that are less temporally representative sites have higher MAE values than those that are more temporally representative. This agrees with previous finding that the

https://doi.org/10.5194/hess-2019-549
Preprint. Discussion started: 18 November 2019

kriging performance declines as the data variation increases (Schläpfer and Schmid, 1999;Martínez-Cob, 1996;Li and Heap, 2011;Keskin and Grunwald, 2018). Gotway et al. (1996) found that the performance of both inverse distance weighting and ordinary kriging declines as CV increases when mapping soil properties. Based on a review of more than 50 spatial interpolation studies, Li and Heap (2011) found that data variation has a significant impact on the performance of spatial interpolation methods. Generally, accuracy decreases as CV increases. Keskin and Grunwald (2018) also found an inverse relationship between the accuracy of RK models and the variation of soil properties based on a review of more than 70 studies of RK model.

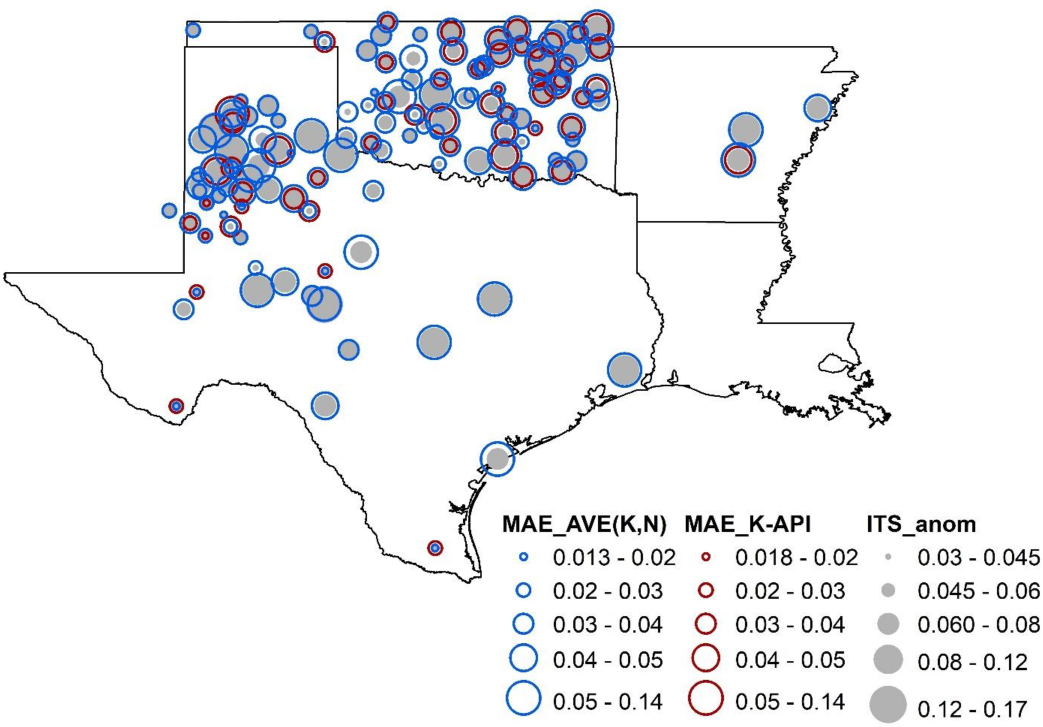

Fig. 10 The spatial distribution of MAE of AVE(K,N), K-API and ITS index using anomalies soil moisture.

https://doi.org/10.5194/hess-2019-549
Preprint. Discussion started: 18 November 2019

The correlation between temporal stability and error is further demonstrated in Fig. 11. The results show that the MAE has a higher correlation with ITS than REV or CV. This finding is consistent for all data formats (absolute, anomalies and percentiles). As the variability in the data increases, the predictive accuracy of RK decreases.

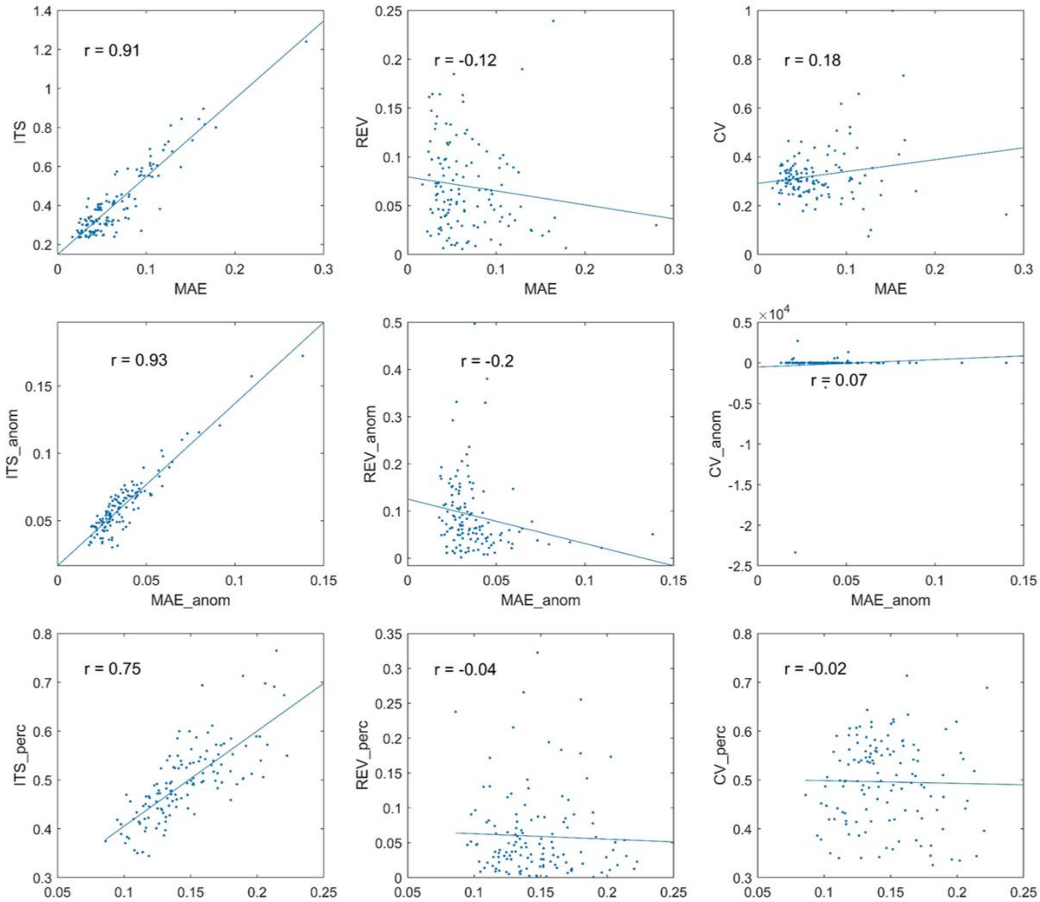

Fig. 11 Relationship between MAE and three indices (ITS, REV and CV) using absolute soil moisture (first row), anomalies (second row) and percentiles (third row).

Another interesting finding from Fig. 10 is that the MAEs of K-API (red circles) are generally smaller than that of AVE(K,N) (blue circles), especially in Oklahoma and northwest Texas. However, in places where in-situ measurements are sparse, such as the central to the south

https://doi.org/10.5194/hess-2019-549
Preprint. Discussion started: 18 November 2019

Hydrology and Earth System Sciences Discussions
Open Access
EGU

of Texas and east Arkansas, the AVE(K,N) has similar and sometimes smaller MAE than K-API. This indicates when in-situ measurements are sparse, using additional sources of soil moisture information (such that from NLDAS) may help to increase the accuracy. To further confirm this point, the assessment of K-API and AVE(K,N) with varying numbers of sampling points and different sampling schemes are provided in Fig. 12 using (a) NSE and (b) MAE, respectively. Three different sampling schemes are compared, including tail sampling (choosing sampling points from the tail of ascending ITS), head sampling (choosing sampling points from the head of ascending ITS) and even sampling (choosing sampling points evenly from the ascending ITS). Since ITS is an indicator of the temporal stability, a more ‘‘representative’’ site is characterized by a lower ITS. Therefore, the tail sampling scheme selects the least “representative” sites for RK, head sampling selects the most “representative” sites for RK, while even sampling selects the sites most close to the population distribution for RK.

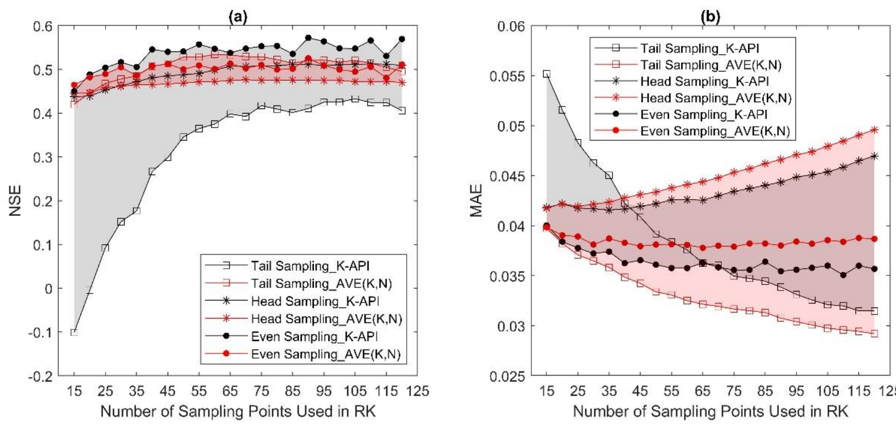

Fig. 12 Assessment of K-API and AVE(K,N) under different sampling scheme and sampling points in terms of (a) NSE and (b) MAE. Both the NSE and MAE are calculated using the out-of-sample stations over the entire study period. “Tail Sampling” indicate choosing sampling points from the tail of ascending ITS (denoted by line with square), “Head Sampling” indicates choosing sampling points from the head of ascending ITS (denoted by line with asterisk), while “Even

https://doi.org/10.5194/hess-2019-549
Preprint. Discussion started: 18 November 2019

Sampling" indicates choosing sampling points evenly from the ascending ITS (denoted by line with dot).

A positive NSE trend is observed as sample density increases (Fig. 12a). The most significant improvement is observed for K-API using tail sampling (black line with squares). When used alone, the K-API under tail sampling shows the lowest NSE values among all datasets using all sampling schemes. However, when combined with the NLDAS, or the AVE(K,N) product (red line with squares) using tail sampling shows comparable NSE values with other product and other sampling schemes. Generally, the NSE variation of AVE(K,N) (red shaded area) under different sampling schemes is much smaller than that of K-API product (grey shaded area). This indicates that the hybrid product AVE(K,N) can reduce uncertainties and it is especially helpful for reducing the errors caused by using too few stations or using unrepresentative stations, as compared with the K-API product.

A decreasing trend in MAE is observed as sample density increases using both the tail and even sampling schemes, which is consistent with the NSE results (Fig. 12b). Although the AVE(K,N) has a larger MAE than K-API when using even sampling, the differences are not statistically significant ($p>0.05$) based on ANOVA. However, when using tail sampling, the AVE(K,N) shows a statistically significant ($p<0.05$) improvement over K-API, especially at lower sample densities. The MAE of K-API is 0.055 m^3m^{-3} using 15 stations, but it drops to only 0.004 m^3m^{-3} when AVE(K,N) is used. This indicates when sampling sites are less representative and sparsely distributed, adding an extra source of soil moisture information (e.g., the product of AVE(K,N)) significantly improves the accuracy. This finding has practical significance for real-world applications, where achieving dense and representative sampling is always challenging. The Oklahoma Mesonet is a unique and uncommonly densely network, in most cases, soil moisture

https://doi.org/10.5194/hess-2019-549
Preprint. Discussion started: 18 November 2019

stations are sparsely distributed. In these cases, the hybrid product, AVE(K,N), may perform better than K-API.

It is also worth noting that the sampling scheme has a larger impact on RK results than the sample density. Although increasing the station density generally improves the accuracy, the improvement gradually decreases and it levels off when the number of stations is >50 (Fig. 12a and b. This agrees well with previous findings (Yuan and Quiring, 2017). In contrast, the change in the sampling scheme may yield a completely different trend of MAE. As shown in Fig. 12(b), an increasing trend of MAE is observed for both K-API and AVE(K,N) when head sampling is adopted. This increasing trend may due to the higher degree of heterogeneity in validation data of head sampling. Considering the head sampling selects the most representative sites for RK modeling, while the remaining sites are less representative and have larger temporal variability, which may yield larger errors. Thus, the head sampling should be avoided for RK modeling, and the even sampling scheme or the bootstrapping random sampling may be more reasonable. In reality, the sampling sites are always a mix of more and less representative sites.

In summary, both increasing sample density and adding an extra source of soil moisture data can improve the accuracy, especially when the station representativeness and station density are low (Fig. 12). Increasing the station density helps to capture the spatial variation of the target variable, while using an additional source of soil moisture tends to lead to a more substantial improvement in accuracy.

5. Conclusions

This work is the first study that compared multiple methods (REV-, TC- and RK-based LSW and simple average) and considered multiple data formats (absolute, anomalies and percentiles) for soil moisture blending from multiple sources, including satellite (SMAP L4-SM),

https://doi.org/10.5194/hess-2019-549
Preprint. Discussion started: 18 November 2019

model (NLDAS-V2 Noah), and in-situ measurement. All soil moisture datasets are generated at 4-km spatial resolution and updated daily. The results indicate that the SMAP data have a large negative bias (-18%) compared with in-situ measurements (Fig. 3), thus it should be used with caution without standardization, especially for drought monitoring. Both the absolute and relative errors from TC vary with the input datasets (Fig. 4). In contrast, REV provides an absolute measurement error, but in a relative (ratio) format. Generally, the TC-estimated error variance tend to show the opposite pattern as REV. That is, the soil moisture products that have a low error variance based on REV, tend to have a larger error variance using TC. The RK-estimated error variances are similar for different datasets (Fig. 4 and Fig. 6).

The hybrid results are not sensitive to the weighting scheme that is used. There were no statistically significant differences between the hybrid datasets when using different weighting methods (TC, REV, RK). There is also no significant advantage to using more complicated weighting (LSW) over the simple average (AVE). The merged products from two datasets (with one fixed as K-API) are found to have comparable accuracy with merged products using three datasets. This indicates that in-situ measurements are valuable for improving the accuracy of blended soil moisture datasets. In terms of different data formats (absolute, anomalies, percentiles or RK-gridded soil moisture), the NSE for anomalies and percentiles (0.60) is higher than that of absolute soil moisture (0.25) mainly due to reduced biases when using anomalies and percentiles (Supplementary. Fig. 6-8). However, the relative errors (or error ranking) is independent of the data format used. The errors from RK are highly correlated with ITS (Fig. 10 and Fig. 11). This indicates that the predictive capability of RK decreases as the heterogeneity increases.

Both K-API and AVE(K,N) are recommended as optimal soil moisture datasets. Considering the PRISM dates back to 1895 and NLDAS dates back to 1979, a long-term soil

https://doi.org/10.5194/hess-2019-549
Preprint. Discussion started: 18 November 2019

moisture record can be generated by adopting the methods in this study. It is also found K-API can be used alone if the station density is high (>50 stations in our case). However, when the station density is low (<50 stations) and the stations are not representative, the hybrid product (AVE(K,N)) has significantly better performance. Increasing station density helps to capture the spatial variation of the target variable, while using an extra source of soil moisture information may help to reduce the overall uncertainties (Fig. 12). This has significant practical implications for real-world applications because achieving a high density of stations that are spatially representative is always challenging.

Finally, there are some limitations of this study, such as: (1) the soil moisture products used in this study were extracted with time slot covering 7 am. However, a temporal mismatch between different soil moisture products may still exist due to their different temporal resolutions. Future work can adopt the daily average or other methods to ensure the temporal coherence of different datasets. (2) this study only considered precipitation (API) in soil moisture kriging, while in future studies other variables, such as soil properties, land cover and topography (DEM), may be helpful for soil moisture estimation (Ochsner et al., 2019) and should be considered to improve the accuracy using RK. (3) Geographically weighted regression kriging (GWRK) (Brunsdon et al., 1996;Fotheringham et al., 2003) considers the spatially non-stationarity relationships between dependent variable and independent variables and weights the regression points by their distance to the target point. Therefore, it may be more accurate than RK (Yang et al., 2019;Kumar et al., 2012), and should be explored in future studies for estimating soil moisture. (4) Further study is required to test whether these conclusions are valid in other parts of the world.

Data Availability

All datasets used in this study are publicly available. The SMAP data can be accessed through National Snow and Ice Data Center (http://nsidc.org/data/smap). The NLDAS-V2 Noah soil moisture products can be accessed through NASA's Earth Observing System Data and Information System (EOSDIS) (https://disc.gsfc.nasa.gov/). The in-situ soil moisture measurements can be accessed through the National Soil Moisture Network (http://www.nationalsoilmoisture.com/).

Author contribution

NZ designed and carried out the study under supervision of SQ. NZ prepared the original manuscript and all the co-authors contributed scientifically by providing editing, comments and suggestions.

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

Acknowledgement

This work was financially supported by the NOAA Modeling, Analysis, Predictions and Projections (MAPP) “Developing National Soil Moisture Products to Improve Drought Monitoring” project (Grant number: NA17OAR4310136).

References

Abbaspour, K. C., Rouholahnejad, E., Vaghefi, S., Srinivasan, R., Yang, H., and Kløve, B.: A continental-scale hydrology and water quality model for Europe: Calibration and uncertainty of a high-resolution large-scale SWAT model, Journal of Hydrology, 524, 733-752, 2015.

Albergel, C., de Rosnay, P., Gruhier, C., Muñoz-Sabater, J., Hasenauer, S., Isaksen, L., Kerr, Y., and Wagner, W.: Evaluation of remotely sensed and modelled soil moisture products using global ground-based in situ observations, Remote Sensing of Environment, 118, 215-226, https://doi.org/10.1016/j.rse.2011.11.017, 2012.

Anctil, F., Michel, C., Perrin, C., and Andréassian, V.: A soil moisture index as an auxiliary ANN input for stream flow forecasting, Journal of Hydrology, 286, 155-167, https://doi.org/10.1016/j.jhydrol.2003.09.006, 2004.

https://doi.org/10.5194/hess-2019-549
Preprint. Discussion started: 18 November 2019

Bi, H., Ma, J., Zheng, W., and Zeng, J.: Comparison of soil moisture in GLDAS model simulations and in situ observations over the Tibetan Plateau, Journal of Geophysical Research: Atmospheres, 121, 2658-2678, 2016.

Brocca, L., Melone, F., Moramarco, T., Wagner, W., Naeimi, V., Bartalis, Z., and Hasenauer, S.: Improving runoff prediction through the assimilation of the ASCAT soil moisture product, Hydrology and Earth System Sciences, 14, 1881-1893, 2010.

Brocca, L., Tullo, T., Melone, F., Moramarco, T., and Morbidelli, R.: Catchment scale soil moisture spatial–temporal variability, Journal of hydrology, 422, 63-75, 2012.

Brunsdon, C., Fotheringham, A. S., and Charlton, M. E.: Geographically weighted regression: a method for exploring spatial nonstationarity, Geographical analysis, 28, 281-298, 1996.

Caires, S., and Sterl, A.: Validation of ocean wind and wave data using triple collocation, Journal of Geophysical Research: Oceans, 108, 2003.

Chen, F., Crow, W. T., Colliander, A., Cosh, M. H., Jackson, T. J., Bindlish, R., Reichle, R. H., Chan, S. K., Bosch, D. D., and Starks, P. J.: Application of triple collocation in ground-based validation of Soil Moisture Active/Passive (SMAP) level 2 data products, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 10, 489-502, 2017.

Chen, Y., Yang, K., Qin, J., Zhao, L., Tang, W., and Han, M.: Evaluation of AMSR-E retrievals and GLDAS simulations against observations of a soil moisture network on the central Tibetan Plateau, Journal of Geophysical Research: Atmospheres, 118, 4466-4475, 10.1002/jgrd.50301, 2013.

Cho, E., and Choi, M.: Regional scale spatio-temporal variability of soil moisture and its relationship with meteorological factors over the Korean peninsula, Journal of Hydrology, 516, 317-329, 2014.

Cosh, M. H., Ochsner, T. E., McKee, L., Dong, J. N., Basara, J. B., Evett, S. R., Hatch, C. E., Small, E. E., Steele-Dunne, S. C., Zreda, M., and Sayde, C.: The Soil Moisture Active Passive Marena, Oklahoma, In Situ Sensor Testbed (SMAP-MOISST): Testbed Design and Evaluation of In Situ Sensors, Vadose Zone J., 15, 11, 10.2136/vzj2015.09.0122, 2016.

Crow, W., and Van den Berg, M.: An improved approach for estimating observation and model error parameters in soil moisture data assimilation, Water Resources Research, 46, W12519, 10.1029/2010WR009402, 2010.

Crow, W., Kumar, S., and Bolten, J.: On the utility of land surface models for agricultural drought monitoring, Hydrology and Earth System Sciences, 16, 3451-3460, 2012a.

Crow, W. T., Berg, A. A., Cosh, M. H., Loew, A., Mohanty, B. P., Panciera, R., de Rosnay, P., Ryu, D., and Walker, J. P.: Upscaling sparse ground ‐ based soil moisture observations for the validation of coarse ‐ resolution satellite soil moisture products, Reviews of Geophysics, 50, 2012b.

Dai, A.: Increasing drought under global warming in observations and models, Nature Climate Change, 3, 52, 2013.

Daly, C., Halbleib, M., Smith, J. I., Gibson, W. P., Doggett, M. K., Taylor, G. H., Curtis, J., and Pasteris, P. P.: Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States, International Journal of Climatology: a Journal of the Royal Meteorological Society, 28, 2031-2064, 2008.

Delworth, T. L., and Manabe, S.: The influence of potential evaporation on the variabilities of simulated soil wetness and climate, Journal of Climate, 1, 523-547, 1988.

Dirmeyer, P. A., Wu, J., Norton, H. E., Dorigo, W. A., Quiring, S. M., Ford, T. W., Santanello Jr, J. A., Bosilovich, M. G., Ek, M. B., and Koster, R. D.: Confronting weather and climate

https://doi.org/10.5194/hess-2019-549
Preprint. Discussion started: 18 November 2019

models with observational data from soil moisture networks over the United States, Journal of hydrometeorology, 17, 1049-1067, 2016.

Dobriyal, P., Qureshi, A., Badola, R., and Hussain, S. A.: A review of the methods available for estimating soil moisture and its implications for water resource management, Journal of Hydrology, 458, 110-117, 2012.

Dong, J., Crow, W. T., and Bindlish, R.: The Error Structure of the SMAP Single and Dual Channel Soil Moisture Retrievals, Geophysical Research Letters, 45, 758-765, 2018.

Dorigo, W., Scipal, K., Parinussa, R. M., Liu, Y. Y., Wagner, W., De Jeu, R. A., and Naeimi, V.: Error characterisation of global active and passive microwave soil moisture data sets, 2010.

Dorigo, W., Wagner, W., Hohensinn, R., Hahn, S., Paulik, C., Xaver, A., Gruber, A., Drusch, M., Mecklenburg, S., and Oevelen, P. v.: The International Soil Moisture Network: a data hosting facility for global in situ soil moisture measurements, Hydrology and Earth System Sciences, 15, 1675-1698, 2011.

Dorigo, W., Xaver, A., Vreugdenhil, M., Gruber, A., Hegyiova, 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 J., 12, 2013.

Downer, C. W., and Ogden, F. L.: Prediction of runoff and soil moistures at the watershed scale: Effects of model complexity and parameter assignment, Water Resources Research, 39, 10.1029/2002wr001439, 2003.

Eldeiry, A. A., and Garcia, L. A.: Comparison of ordinary kriging, regression kriging, and cokriging techniques to estimate soil salinity using LANDSAT images, J. Irrig. Drainage Eng-ASCE, 136, 355-364, 2010.

Entekhabi, D., Njoku, E. G., O'Neill, P. E., Kellogg, K. H., Crow, W. T., Edelstein, W. N., Entin, J. K., Goodman, S. D., Jackson, T. J., and Johnson, J.: The soil moisture active passive (SMAP) mission, Proceedings of the IEEE, 98, 704-716, 2010.

Fang, H., Wei, S., Jiang, C., and Scipal, K.: Theoretical uncertainty analysis of global MODIS, CYCLOPES, and GLOBCARBON LAI products using a triple collocation method, Remote Sensing of Environment, 124, 610-621, 2012.

Ford, T. W., and Quiring, S. M.: Comparison and application of multiple methods for temporal interpolation of daily soil moisture, International Journal of Climatology, 34, 2604-2621, 2014.

Ford, T. W., Rapp, A. D., and Quiring, S. M.: Does afternoon precipitation occur preferentially over dry or wet soils in Oklahoma?, Journal of Hydrometeorology, 16, 874-888, 2015.

Ford, T. W., Wang, Q., and Quiring, S. M.: The observation record length necessary to generate robust soil moisture percentiles, Journal of Applied Meteorology and Climatology, 55, 2131-2149, 2016.

Ford, T. W., and Quiring, S. M.: Comparison of contemporary in situ, model, and satellite remote sensing soil moisture with a focus on drought monitoring, Water Resources Research, 55, 1565-1582, 2019.

Fotheringham, A. S., Brunsdon, C., and Charlton, M.: Geographically weighted regression: the analysis of spatially varying relationships, John Wiley & Sons, 2003.

Gotway, C. A., Ferguson, R. B., Hergert, G. W., and Peterson, T. A.: Comparison of kriging and inverse-distance methods for mapping soil parameters, Soil Sci. Soc. Am. J., 60, 1237-1247, 1996.

https://doi.org/10.5194/hess-2019-549
Preprint. Discussion started: 18 November 2019

Gruber, A., Dorigo, W. A., Zwieback, S., Xaver, A., and Wagner, W.: Characterizing coarse-scale representativeness of in situ soil moisture measurements from the International Soil Moisture Network, Vadose Zone J., 12, 2013.

Gruber, A., Su, C.-H., Zwieback, S., Crow, W., Dorigo, W., and Wagner, W.: Recent advances in (soil moisture) triple collocation analysis, International Journal of Applied Earth Observation and Geoinformation, 45, 200-211, 2016.

Heggen, R. J.: Normalized antecedent precipitation index, Journal of hydrologic Engineering, 6, 377-381, 2001.

Hengl, T., Heuvelink, G. B., and Stein, A.: A generic framework for spatial prediction of soil variables based on regression-kriging, Geoderma, 120, 75-93, 2004.

Hengl, T., Heuvelink, G. B., and Rossiter, D. G.: About regression-kriging: from equations to case studies, Computers & geosciences, 33, 1301-1315, 2007.

Hirschi, M., Seneviratne, S. I., Alexandrov, V., Boberg, F., Boroneant, C., Christensen, O. B., Formayer, H., Orlowsky, B., and Stepanek, P.: Observational evidence for soil-moisture impact on hot extremes in southeastern Europe, Nature Geoscience, 4, 17, 2011.

Huang, J., van den Dool, H. M., and Georgarakos, K. P.: Analysis of model-calculated soil moisture over the United States (1931–1993) and applications to long-range temperature forecasts, Journal of Climate, 9, 1350-1362, 1996.

Imaoka, K., Kachi, M., Kasahara, M., Ito, N., Nakagawa, K., and Oki, T.: Instrument performance and calibration of AMSR-E and AMSR2, International Archives of the Photogrammetry, Remote Sensing and Spatial Information Science, 38, 13-18, 2010.

Jacobs, J. M., Hsu, E. C., and Choi, M.: Time stability and variability of electronically scanned thinned array radiometer soil moisture during Southern Great Plains hydrology experiments, Hydrological processes, 24, 2807-2819, 2010.

Keesstra, S., Pereira, P., Novara, A., Brevik, E. C., Azorin-Molina, C., Parras-Alcántara, L., Jordán, A., and Cerdà, A.: Effects of soil management techniques on soil water erosion in apricot orchards, Science of the Total Environment, 551, 357-366, 2016.

Kerr, Y. H., Waldteufel, P., Wigneron, J.-P., Martinuzzi, J., Font, J., and Berger, M.: Soil moisture retrieval from space: The Soil Moisture and Ocean Salinity (SMOS) mission, IEEE transactions on Geoscience and remote sensing, 39, 1729-1735, 2001.

Keskin, H., and Grunwald, S.: Regression kriging as a workhorse in the digital soil mapper's toolbox, Geoderma, 326, 22-41, https://doi.org/10.1016/j.geoderma.2018.04.004, 2018.

Kohler, M. A., and Linsley, R. K.: Predicting the runoff from storm rainfall, in: Res. Paper,, US Weather Bureau, Washington, DC, 1951.

Koster, R. D., Dirmeyer, P. A., Guo, Z., Bonan, G., Chan, E., Cox, P., Gordon, C., Kanae, S., Kowalczyk, E., and Lawrence, D.: Regions of strong coupling between soil moisture and precipitation, Science, 305, 1138-1140, 2004.

Kottek, M., Grieser, J., Beck, C., Rudolf, B., and Rubel, F.: World map of the Köppen-Geiger climate classification updated, Meteorologische Zeitschrift, 15, 259-263, 2006.

Kumar, S., Lal, R., and Liu, D.: A geographically weighted regression kriging approach for mapping soil organic carbon stock, Geoderma, 189, 627-634, 2012.

Li, J., and Heap, A. D.: A review of comparative studies of spatial interpolation methods in environmental sciences: Performance and impact factors, Ecological Informatics, 6, 228-241, https://doi.org/10.1016/j.ecoinf.2010.12.003, 2011.

Liang, X., Wood, E. F., and Lettenmaier, D. P.: Surface soil moisture parameterization of the VIC-2L model: Evaluation and modification, Global and Planetary Change, 13, 195-206, 1996.

https://doi.org/10.5194/hess-2019-549
Preprint. Discussion started: 18 November 2019

Hydrology and Earth System Sciences Discussions

Lievens, H., and Verhoest, N. E. C.: Spatial and temporal soil moisture estimation from RADARSAT-2 imagery over Flevoland, The Netherlands, Journal of Hydrology, 456-457, 44-56, https://doi.org/10.1016/j.jhydrol.2012.06.013, 2012.

Liu, Q., Reichle, R. H., Bindlish, R., Cosh, M. H., Crow, W. T., de Jeu, R., De Lannoy, G. J., Huffman, G. J., and Jackson, T. J.: The contributions of precipitation and soil moisture observations to the skill of soil moisture estimates in a land data assimilation system, Journal of Hydrometeorology, 19, 2018.

Martínez-Cob, A.: Multivariate geostatistical analysis of evapotranspiration and precipitation in mountainous terrain, Journal of Hydrology, 174, 19-35, 1996.

McColl, K. A., Vogelzang, J., Konings, A. G., Entekhabi, D., Piles, M., and Stoffelen, A.: Extended triple collocation: Estimating errors and correlation coefficients with respect to an unknown target, Geophysical Research Letters, 41, 6229-6236, 2014.

Meng, L., and Quiring, S. M.: A Comparison of Soil Moisture Models Using Soil Climate Analysis Network Observations, Journal of Hydrometeorology, 9, 641-659, 10.1175/2008jhm916.1, 2008.

Miralles, D. G., Crow, W. T., and Cosh, M. H.: Estimating spatial sampling errors in coarse-scale soil moisture estimates derived from point-scale observations, Journal of Hydrometeorology, 11, 1423-1429, 2010.

Mishra, U., Lal, R., Liu, D., and Van Meirvenne, M.: Predicting the spatial variation of the soil organic carbon pool at a regional scale, Soil Sci. Soc. Am. J., 74, 906-914, 2010.

Mitchell, K. E., Lohmann, D., Houser, P. R., Wood, E. F., Schaake, J. C., Robock, A., Cosgrove, B. A., Sheffield, J., Duan, Q., and Luo, L.: The multi‐institution North American Land Data Assimilation System (NLDAS): Utilizing multiple GCIP products and partners in a continental distributed hydrological modeling system, Journal of Geophysical Research: Atmospheres, 109, 2004.

Mittelbach, H., and Seneviratne, S. I.: A new perspective on the spatio-temporal variability of soil moisture: temporal dynamics versus time-invariant contributions, Hydrology and Earth System Sciences, 16, 2169-2179, 2012.

Ochsner, T. E., Linde, E., Haffner, M., and Dong, J.: Mesoscale Soil Moisture Patterns Revealed Using a Sparse In Situ Network and Regression Kriging, Water Resources Research, 55, 4785-4800, 10.1029/2018wr024535, 2019.

Odeha, I., McBratney, A., and Chittleborough, D.: Spatial prediction of soil properties from landform attributes derived from a digital elevation model, Geoderma, 63, 197-214, 1994.

Paloscia, S., Pettinato, S., Santi, E., Notarnicola, C., Pasolli, L., and Reppucci, A.: Soil moisture mapping using Sentinel-1 images: Algorithm and preliminary validation, Remote Sensing of Environment, 134, 234-248, 2013.

Parinussa, R. M., Meesters, A. G., Liu, Y. Y., Dorigo, W., Wagner, W., and de Jeu, R. A.: 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, 2011.

Penna, D., Brocca, L., Borga, M., and Dalla Fontana, G.: Soil moisture temporal stability at different depths on two alpine hillslopes during wet and dry periods, Journal of Hydrology, 477, 55-71, 2013.

Pittelkow, C. M., Liang, X., Linquist, B. A., Van Groenigen, K. J., Lee, J., Lundy, M. E., Van Gestel, N., Six, J., Venterea, R. T., and van Kessel, C.: Productivity limits and potentials of the principles of conservation agriculture, Nature, 517, 365, 2015.

Quiring, S. M., Ford, T. W., Wang, J. K., Khong, A., Harris, E., Lindgren, T., Goldberg, D. W., and Li, Z.: The North American Soil Moisture Database: Development and Applications, Bulletin of the American Meteorological Society, 97, 1441-1459, 10.1175/bams-d-13-00263.1, 2016.

Reichle, R., Lannoy, G. D., Koster, R. D., Crow, W. T., Kimball, J. S., and Liu, Q.: SMAP L4 Global 3-hourly 9 km EASE-Grid Surface and Root Zone Soil Moisture Geophysical Data, Version 4. SPL4SMGP, Boulder, Colorado USA. NASA National Snow and Ice Data Center Distributed Active Archive Center, https://doi.org/10.5067/KPJNN2GI1DQR, 2018.

Reichle, R. H., De Lannoy, G. J. M., 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., González-Zamora, Á., Holifield Collins, C. D., 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, 10.1175/JHM-D-17-0063.1, 2017.

Robock, A., Vinnikov, K. Y., Schlosser, C. A., Speranskaya, N. A., and Xue, Y.: Use of Midlatitude Soil Moisture and Meteorological Observations to Validate Soil Moisture Simulations with Biosphere and Bucket Models, Journal of Climate, 8, 15-35, 10.1175/1520-0442(1995)008<0015:uomsma>2.0.co;2, 1995.

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, 10.1175/bams-85-3-381, 2004.

Roebeling, R., Wolters, E., Meirink, J., and Leijnse, H.: Triple collocation of summer precipitation retrievals from SEVIRI over Europe with gridded rain gauge and weather radar data, Journal of Hydrometeorology, 13, 1552-1566, 2012.

Schläpfer, F., and Schmid, B.: Ecosystem effects of biodiversity: a classification of hypotheses and exploration of empirical results, Ecological Applications, 9, 893-912, 1999.

Scipal, K., Holmes, T., De Jeu, R., Naeimi, V., and Wagner, W.: A possible solution for the problem of estimating the error structure of global soil moisture data sets, Geophysical Research Letters, 35, 2008.

Seneviratne, S. I., Corti, T., Davin, E. L., Hirschi, M., Jaeger, E. B., Lehner, I., Orlowsky, B., and Teuling, A. J.: Investigating soil moisture–climate interactions in a changing climate: A review, Earth-Science Reviews, 99, 125-161, 2010.

Spennemann, P. C., Rivera, J. A., Saulo, A. C., and Penalba, O. C.: A Comparison of GLDAS Soil Moisture Anomalies against Standardized Precipitation Index and Multisatellite Estimations over South America, Journal of Hydrometeorology, 16, 158-171, 10.1175/jhm-d-13-0190.1, 2015.

Stoffelen, A.: Toward the true near‐surface wind speed: Error modeling and calibration using triple collocation, Journal of Geophysical Research: Oceans, 103, 7755-7766, 1998.

Su, C. H., Ryu, D., Crow, W. T., and Western, A. W.: Beyond triple collocation: Applications to soil moisture monitoring, Journal of Geophysical Research: Atmospheres, 119, 6419-6439, 2014.

https://doi.org/10.5194/hess-2019-549
Preprint. Discussion started: 18 November 2019

Takagi, K., and Lin, H. S.: Changing controls of soil moisture spatial organization in the Shale Hills Catchment, Geoderma, 173-174, 289-302, https://doi.org/10.1016/j.geoderma.2011.11.003, 2012.

Tian, L., and Quiring, S. M.: Spatial and temporal patterns of drought in Oklahoma (1901–2014), International Journal of Climatology, 39, 3365-3378, 2019.

Vachaud, G., Passerat de Silans, A., Balabanis, P., and Vauclin, M.: Temporal Stability of Spatially Measured Soil Water Probability Density Function 1, Soil Sci. Soc. Am. J., 49, 822-828, 1985.

Vinnikov, K. Y., and Yeserkepova, I.: Soil moisture: Empirical data and model results, Journal of Climate, 4, 66-79, 1991.

Vinnikov, K. Y., Robock, A., Speranskaya, N. A., and Schlosser, C. A.: Scales of temporal and spatial variability of midlatitude soil moisture, Journal of Geophysical Research: Atmospheres, 101, 7163-7174, 1996.

Wagner, W., Lemoine, G., Borgeaud, M., and Rott, H.: A study of vegetation cover effects on ERS scatterometer data, IEEE Transactions on Geoscience and Remote Sensing, 37, 938-948, 1999.

Wagner, W., Hahn, S., Kidd, R., Melzer, T., Bartalis, Z., Hasenauer, S., Figa-Saldaña, J., de Rosnay, P., Jann, A., and Schneider, S.: The ASCAT soil moisture product: A review of its specifications, validation results, and emerging applications, Meteorologische Zeitschrift, 22, 5-33, 2013.

Wanders, N., Karssenberg, D., Roo, A. d., De Jong, S., and Bierkens, M.: The suitability of remotely sensed soil moisture for improving operational flood forecasting, Hydrology and Earth System Sciences, 18, 2343-2357, 2014.

Wang, A., Lettenmaier, D. P., and Sheffield, J.: Soil moisture drought in China, 1950–2006, Journal of Climate, 24, 3257-3271, 2011.

Wang, T., Liu, Q., Franz, T. E., Li, R., Lang, Y., and Fiebrich, C. A.: Spatial patterns of soil moisture from two regional monitoring networks in the United States, Journal of Hydrology, 552, 578-585, 2017.

Xia, Y., Mitchell, K., Ek, M., Sheffield, J., Cosgrove, B., Wood, E., Luo, L., Alonge, C., Wei, H., and Meng, J.: Continental‐scale water and energy flux analysis and validation for the North American Land Data Assimilation System project phase 2 (NLDAS‐2): 1. Intercomparison and application of model products, Journal of Geophysical Research: Atmospheres, 117, 2012.

Xia, Y., Sheffield, J., Ek, M. B., Dong, J., Chaney, N., Wei, H., Meng, J., and Wood, E. F.: Evaluation of multi-model simulated soil moisture in NLDAS-2, Journal of Hydrology, 512, 107-125, http://dx.doi.org/10.1016/j.jhydrol.2014.02.027, 2014.

Xia, Y., Ek, M. B., Wu, Y., Ford, T., and Quiring, S. M.: Comparison of NLDAS-2 simulated and NASMD observed daily soil moisture. Part I: Comparison and analysis, Journal of Hydrometeorology, 16, 1962-1980, 2015.

Yang, S.-H., Liu, F., Song, X.-D., Lu, Y.-Y., Li, D.-C., Zhao, Y.-G., and Zhang, G.-L.: Mapping topsoil electrical conductivity by a mixed geographically weighted regression kriging: A case study in the Heihe River Basin, northwest China, Ecological Indicators, 102, 252-264, https://doi.org/10.1016/j.ecolind.2019.02.038, 2019.

Yao, X., Fu, B., Lü, Y., Sun, F., Wang, S., and Liu, M.: Comparison of four spatial interpolation methods for estimating soil moisture in a complex terrain catchment, PloS one, 8, e54660, 2013.

https://doi.org/10.5194/hess-2019-549
Preprint. Discussion started: 18 November 2019

Hydrology and Earth System Sciences
Discussions

Yilmaz, M., Crow, W., Anderson, M., and Hain, C.: An objective methodology for merging satellite‐and model‐based soil moisture products, Water Resources Research, 48, 2012.

Yuan, S., and Quiring, S. M.: Comparison of three methods of interpolating soil moisture in Oklahoma, International Journal of Climatology, 37, 987-997, 2017.

Zeng, Y., Su, Z., van der Velde, R., Wang, L., Xu, K., Wang, X., and Wen, J.: Blending satellite observed, model simulated, and in situ measured soil moisture over Tibetan Plateau, Remote sensing, 8, 268, 2016.

Zhang, N., Quiring, S., Ochsner, T., and Ford, T.: Comparison of Three Methods for Vertical Extrapolation of Soil Moisture in Oklahoma, Vadose Zone J., 16, 2017a.

Zhang, N., Zhao, C., Quiring, S. M., and Li, J.: Winter Wheat Yield Prediction Using Normalized Difference Vegetative Index and Agro-Climatic Parameters in Oklahoma, Agronomy Journal, 109, 2700-2713, 2017b.

Zhao, Y., Peth, S., Wang, X., Lin, H., and Horn, R.: Controls of surface soil moisture spatial patterns and their temporal stability in a semi‐arid steppe, Hydrological Processes, 24, 2507-2519, 2010.