

Blending Noah, SMOS, and in Situ Soil Moisture Using Multiple Weighting and Sampling Schemes^{clip}

NING ZHANG,^{a,b} STEVEN M. QUIRING,^a AND TRENT W. FORD^c

^a Department of Geography, The Ohio State University, Columbus, Ohio

^b University of California Agriculture and Natural Resources, Davis, California

^c Illinois State Water Survey, University of Illinois at Urbana-Champaign, Urbana, Illinois

(Manuscript received 17 May 2020, in final form 30 April 2021)

ABSTRACT: Soil moisture can be obtained from in situ measurements, satellite observations, and model simulations. This study evaluates the importance of in situ observations in soil moisture blending, and compares different weighting and sampling methods for combining model, satellite, and in situ soil moisture data to generate an accurate and spatially continuous soil moisture product at 4-km resolution. Four different datasets are used: the antecedent precipitation index (API); KAPI, which incorporates in situ soil moisture observations with the API using regression kriging; SMOS L3 soil moisture; and model-simulated soil moisture from the Noah model as part of the North American Land Data Assimilation System (NLDAS). Triple collocation, least squares weighting, and equal weighting are used to generate blended soil moisture products. An enumerated weighting scheme is designed to investigate the impact of different weighting schemes. The sensitivity of the blended soil moisture products to sampling schemes, station density, and data formats (absolute, anomalies, and percentiles) are also investigated. The results reveal that KAPI outperforms API. This indicates that incorporating in situ soil moisture improves the accuracy of the blended soil moisture products. There are no statistically significant ($p > 0.05$) differences between blended soil moisture using triple collocation and equal weighting approaches, and both methods provide suboptimal weighting. Optimal weighting is achieved by assigning larger weights to KAPI and smaller weights to SMOS. Using multiple sources of soil moisture is helpful for reducing uncertainty and improving accuracy, especially when the sampling density is low, or the sampling stations are less representative. These results are consistent regardless of how soil moisture is represented (absolute, anomalies, or percentiles).

KEYWORDS: Soil moisture; Measurements; Remote sensing; Sampling; Surface observations; Model comparison

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 (Abbaspour et al. 2015; Keesstra et al. 2016), agricultural and water resource management (Dobriyal et al. 2012; Pittelkow et al. 2015), runoff and flooding prediction (Brocca et al. 2010; Wanders et al. 2014), and drought monitoring (Dai 2013; Wang et al. 2011).

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 (AMSR-2) (25-km resolution; Imaoka et al. 2010) onboard the *GCOM-W1* satellite, and the Soil Moisture Active Passive (SMAP) satellite (launched 2015; 3-, 9-, and 36-km resolution; Entekhabi et al. 2010). The Advanced Scatterometer onboard the Meteorological Operational (*MetOp-A/B/C*) satellite series (launched 2006, 2012, and 2018 respectively, 25-km resolution) is an active microwave satellite in orbit (Cartwright and Fraser 2021; Wagner et al. 2013). While these sensors have a coarser spatial resolution, they provide 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 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 (Parinussa et al. 2011; Wagner et al. 1999).

^{clip} Supplemental information related to this paper is available at the Journals Online website: <https://doi.org/10.1175/JHM-D-20-0119.s1>.

Corresponding author: Ning Zhang, zhang.7819@osu.edu

Models can provide a valuable source of spatially continuous soil moisture, and, unlike satellites, they can represent root zone moisture conditions. 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 model-simulated soil moisture 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. Ford and Quiring (2019) compared the modeled soil moisture from NLDAS phase 2 (NLDAS-2) and CPC with in situ measurements and found that the 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 with in situ observations. The in situ networks are typically not dense enough to adequately resolve soil moisture variability within each LSM pixel. In addition, there are systematic differences between different soil moisture models, and their reliability 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 (Albergel et al. 2012; Downer and Ogden 2003; Meng and Quiring 2008), but they tend to have large biases in simulating the absolute volumetric water content of the soil (Bi et al. 2016; Xia et al. 2015).

In situ soil moisture measurements from 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 formats (Cosh et al. 2016; Dorigo et al. 2013; Ford and Quiring 2014; Zhang et al. 2017a). Currently, the coordinated in situ soil moisture networks include the International Soil Moisture Network (ISMN) (Dorigo et al. 2011), the North American Soil Moisture Database (NASMD) (Quiring et al. 2016), and the National Soil Moisture Network (Ford et al. 2020, manuscript submitted to *J. Hydrometeor.*). Despite these efforts, the number of stations that measure soil moisture continuously are still very limited globally. In addition, since in situ stations provide a point measurement, this limits their application at larger spatial scales.

In summary, each source of soil moisture data has its strengths and weaknesses. The advantage of in situ observations is that they are the only direct measurement of soil moisture, and thus are often used as a benchmark for models and satellites. The primary disadvantages of in situ observations are that they typically have a sparse spatial density and, as a point measurement, they have limited spatial representativeness. Similarly, model and satellite remote sensing soil moisture have the advantage of representing a larger spatial area and, for the most part, a finer spatial resolution. Of course,

the primary disadvantage of models and satellites is that there can be a large discrepancy between different models and sensors, and the satellites with the highest temporal resolution are confined to near-surface soil moisture and have degraded performance under dense vegetation or in frozen soils. Given the advantages and disadvantages of each source of soil moisture information, they can improve soil moisture monitoring if they are combined.

Multiple methods have been developed to blend in situ, modeled, and remotely sensed soil moisture. Data assimilation is one popular approach to generate soil moisture using observations and models. For example, the SMAP L4 product is based on a data assimilation system that combines SMAP L-band brightness temperature observations with the NASA Catchment land surface model (Reichle et al. 2020). While there are many advantages to data assimilation, it is complicated and computationally expensive to implement. As an alternative, a number of statistical methods, including triple collocation (TC) (Stoffelen 1998) with least squares weighting (LSW) and equal weighting (arithmetic average of parent datasets), have been adopted for blending soil moisture for operational and experimental datasets (Yilmaz et al. 2012; Zeng et al. 2016). One example is the European Space Agency's Climate Initiative for Soil Moisture (ESA CCI SM), which uses a least squares merging scheme with uncertainty characterization to merge multiple sources of soil moisture (Gruber et al. 2019).

The motivation of this study is the gradually enhanced regional and national in situ soil moisture network, which provides the opportunity with other data sources to improve the accuracy and usability of soil moisture monitoring. Blending with in situ soil moisture observations can provide a number of advantages to blending solely based on remote sensing and modeled datasets. For example, remote sensing soil moisture—and soil moisture products generated via data assimilation—typically have a 3–7-day data latency. This latency is acceptable for many applications, such as drought monitoring and monthly or seasonal climate outlooks. However, the longer latency is problematic for applications requiring more rapid updates, including flash flood forecasting and field condition monitoring for agriculture. In addition, soil moisture products based entirely on remote sensing observations do not represent soil moisture conditions in the primary root zone. Although we do not examine root zone soil moisture in this study, the methods are easily applicable for blending root zone soil moisture from in situ and model sources. Last, many blended remote sensing–model datasets are available with superdaily temporal resolution (e.g., every 2–3 days). This is a limitation for two primary reasons. First, most meteorological data, such as temperature, precipitation, and humidity, are updated at daily or subdaily time scales. Daily scale soil moisture better matches these data for hydrological analysis and crop modeling. Second, when standardizing soil moisture using anomalies, percentiles, or similar, the robustness of the standardization is sensitive to the historical soil moisture record length (Ford et al. 2016). This is particularly the case for representing soil moisture extremes, important for drought and flood modeling. Therefore, having a 3-day temporal resolution reduces the historical data record

by 3 times versus a daily resolution product, which can result in less representative standardization.

Despite these advantages, in situ data have not been used as widely as remote sensing and modeled soil moisture in operational or experimental blended datasets. Therefore, there is a need to understand better the advantages and limitations of including in situ soil moisture when blending remote sensing and modeled data. It is important to note that a primary motivation of this study is to test the feasibility of blending in situ, remote sensing, and model soil moisture. This will ultimately complement but not replace the wealth of blended and/or assimilated remote sensing–model soil moisture datasets.

In addition, our understanding of the sensitivity of blended soil moisture data to weighting schemes is lacking. In particular, it is helpful to determine whether the commonly used approaches (e.g., TC, equal weighting) generate optimal weights and how different weighting schemes impact hybrid results. The objective of this study is threefold: 1) to evaluate the effect of incorporating in situ observations into soil moisture blending; 2) to assess the sensitivity to different blending and weighting schemes; and 3) to investigate the impact of sampling schemes, station density, and different data format (e.g., absolute, anomalies, and percentiles) on soil moisture blending.

2. Study area and data

This study is conducted in the south-central region of the United States, covering four states, Texas, Oklahoma, Arkansas, and Louisiana, with a total area of $\sim 1150\,400\text{ km}^2$. The south-central United States is an important agricultural region in the United States, but it is also 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 ($<25\text{ cm}$) to east ($>190\text{ cm}$).

This study uses in situ measurements of soil moisture, satellite-observed soil moisture, model-simulated soil moisture, precipitation, and air temperature. To facilitate comparison, a common period of record from January 2011 to December 2018 was used for all datasets. All the gridded datasets are resampled to a 4-km spatial resolution to match PRISM data, which have the highest spatial resolution among the gridded data used in this study. The final blended soil moisture is also generated at 4-km resolution.

a. In situ soil moisture measurements

The in situ soil moisture data were collected from four sparse networks: the Oklahoma Mesonet (OKM), West Texas Mesonet (WTM), Soil Climate Analysis Network (SCAN), and Climate Reference Network (CRN). Daily soil moisture

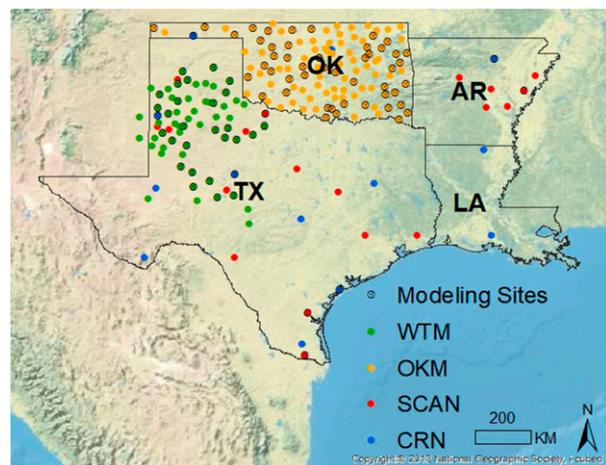


FIG. 1. Study area and location of stations where in situ soil moisture measurements are made. Stations are color coded according to the network they are part of: West Texas Mesonet (WTM; green), Oklahoma Mesonet (OKM; yellow), Soil Climate Analysis Network (SCAN; red), and Climate Reference Network (CRN; blue). A total of 209 stations are presented here, among which 84 stations (indicated by crossed circles) are selected using stratified random sampling for modeling, while the remaining 125 stations were used for out-of-sample validation.

measurements were obtained from the North American Soil Moisture Database (NASMD) in the units of volumetric water content ($\text{m}^3\text{ m}^{-3}$) (Quiring et al. 2016). Since different networks collect data at different time intervals ranging from every 5 min to once per day, for consistency, a single morning measurement (7 a.m. LST) was 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 measurements have passed through a quality assurance and quality control (QAQC) process (Ford and Quiring 2014), and flagged values been removed and filled. The near-surface measurements (5 cm) from a total of 209 stations (Fig. 1) were obtained for this study. The daily in situ soil moisture data are available from 2011 to 2018.

b. SMOS L3 soil moisture

The SMOS satellite was launched in 2009 as part of the ESA Explorer Opportunity science mission. It is dedicated to providing surface (top $\sim 5\text{ cm}$) soil moisture mapping with a target accuracy of $0.04\text{ m}^3\text{ m}^{-3}$ (Kerr et al. 2001). The SMOS Level 3 (L3) soil moisture product used in this study was obtained from the SMOS Data Center (<http://www.catds.fr/sipad>). It is generated at 1-day intervals at a global extent with a spatial resolution of 0.25° . The SMOS L3 products are available from 2010 to the present. Compared with the SMOS level 2 (L2) product, which was produced using a single-orbit algorithm, the SMOS L3 products are produced using the multiorbit algorithm, which enhances the retrievals at the border of the swath (Al Bitar et al. 2017). The SMOS L3 products include both ascending [morning; about 0600 local time (LT)] and

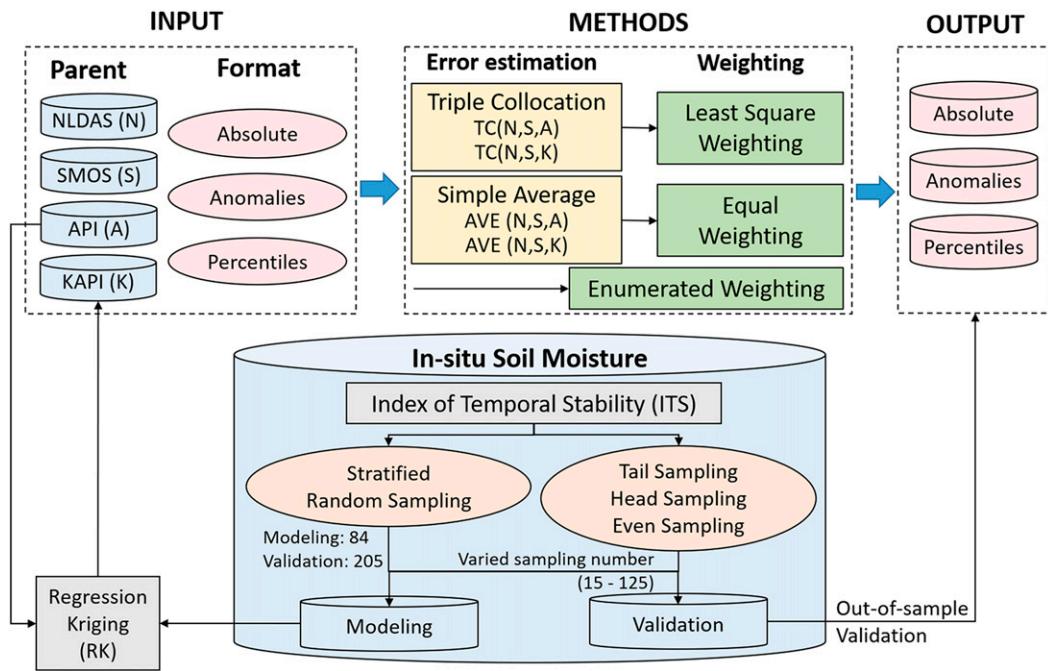


FIG. 2. Schematic diagram showing the methodological framework of this study.

descending (evening; about 1800 LT) overpasses. To be consistent with the daily in situ measurements, which are generated by extracting morning (0700 LT) measurements each day, the ascending SMOS L3 product is used. The nearest neighbor assignment is used to resample SMOS L3 soil moisture from 25 to 4 km to match the spatial resolution of the other datasets.

c. NLDAS-2 Noah soil moisture

This study uses simulated soil moisture from the NLDAS-2 Noah model. The Noah model provides hourly soil moisture fields at 1/8° grid from 1979 to the present. The Noah model has four soil layers, 0–10, 10–40, 40–100, 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 0700 LT is extracted each day to represent the daily soil moisture, and the data from 2011 to 2018 are adopted to match the record length of the in situ data. Finally, the nearest neighbor method is used to resample the simulated soil moisture from 12.5 to 4 km to match the other datasets.

d. PRISM climate data

The PRISM (Parameter-elevation Regressions on Independent Slopes Model) datasets are developed by Oregon State University's PRISM Climate Group ([Daly et al. 2008](#)). PRISM uses surface stations and a weighted regression scheme to generate a daily updated spatial mapping of precipitation and temperature 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

climatological normal (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 the present.

The 4-km daily precipitation from PRISM is used in this study. Since there is a strong coupling between soil moisture and precipitation ([Koster et al. 2004](#)), the latter has been widely used as an important input to estimate the former in various LSMs ([Liang et al. 1996](#); [Liu et al. 2018](#); [Xia et al. 2012](#)). Here, the antecedent precipitation index (API) is used as a soil moisture approximation. The API is calculated solely based on precipitation. The kriged API (KAPI) uses the API and in situ measurements of soil moisture that are incorporated using regression kriging. The API index is introduced in [section 2e\(2\)](#), and regression kriging is introduced in [section 2e\(3\)](#).

e. Data preparation

The data preparation methods are summarized in [Fig. 2](#).

1) ANOMALIES AND PERCENTILES

The volumetric water content of the soil varies as a function of weather conditions, soil characteristics, vegetation, and topography, among other factors, and so it cannot be directly compared between different locations. Therefore, we calculate anomalies and percentiles for all datasets (in situ, SMOS, NLDAS, API, and KAPI) to standardize soil moisture and make them comparable in space and time ([Ford et al. 2015](#); [Zhang et al. 2017a](#)).

Anomalies are calculated by removing the climatological mean from the absolute soil moisture 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 also calculated using an empirical probability distribution function and moving window approach. On each day of the year, all soil moisture measurements that fall within a 31-day window centered on that day are used to construct the empirical probability distribution function. Ford et al. (2016) found that sample sizes between 93 and 186 daily soil moisture observations were required to generate robust percentiles. In our case, all datasets are extracted from 2011 to 2018 (8 years), which provides a large enough sample ($n = 248$) to generate robust percentiles. 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 during the 31-day window centered on the target date.

2) ANTECEDENT PRECIPITATION INDEX

The antecedent precipitation index is a precipitation-based moisture index. It is used to indicate the wetness of a location and serves as an approximation of soil moisture. It has been widely applied in drought monitoring (Crow et al. 2012a), runoff forecasting (Anctil et al. 2004), crop yield prediction (Zhang et al. 2017b), and soil moisture estimation using data assimilation approach (Afshar et al. 2019; Yilmaz and Crow 2013). The API uses the preceding precipitation to estimate the current moisture status, and it is formulated as (Kohler and Linsley 1951)

$$\text{API}(i) = \text{API}(i-1) \times k + \text{PPT}(i), \quad (1)$$

where $\text{API}(i)$ is the API on day i , $\text{PPT}(i)$ is the precipitation occurring on day i , and 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 all 209 stations. The highest correlation ($r = 0.45$) is achieved using $k = 0.92$. Therefore, $k = 0.92$ is used in this study for API calculation.

3) REGRESSION KRIGING

Regression kriging (RK) is one of the most popular and robust spatial interpolation techniques for the digital mapping of soil properties (Keskin and Grunwald 2018). Previous studies demonstrate that RK usually outperforms nongeostatistical methods (Li and Heap 2011; Mishra et al. 2010; Yang et al. 2019), ordinary kriging (Hengl et al. 2004), and cokriging (Eldeiry and Garcia 2010). RK combines a regression between the target variable and one or more auxiliary variables with simple kriging of the regression residuals (Hengl et al. 2007; Odeha et al. 1994). In this study, RK is used to interpolate the point-based in situ soil moisture (target variable) using the API as an auxiliary variable. The RK models can be expressed as two parts (Hengl et al. 2004):

$$\hat{z}(s_0) = \hat{m}(s_0) + \hat{e}(s_0), \quad (2)$$

where $\hat{m}(s_0)$ is the fitted trend. In this study, $\hat{m}(s_0)$ is fitted by a linear model between the auxiliary variable and soil moisture [Eq. (3)]:

$$\hat{m}(s_0) = \hat{\beta}q(s_0), \quad (3)$$

where $\hat{\beta}$ is the estimated model coefficient using generalized least squares, and $q(s_0)$ is the auxiliary variable at the target location s_0 . In this study, the auxiliary variable is the API, represented as absolute values, or anomalies, or percentiles.

The second part of the RK equation is $\hat{e}(s_0)$, which is the interpolated residual. The residual is first calculated from the linear model and then interpolated by simple kriging with an assumed 0 mean:

$$\hat{e}(s_0) = \sum_{i=1}^n \lambda_i e(s_i), \quad (4)$$

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 back to the predicted trend, the final RK predictions (KAPI) are obtained.

4) SITE SELECTION BASED ON THE INDEX OF TEMPORAL STABILITY

In this study, 40% of the stations with soil moisture measurements (84 sites) are used for modeling (crossed circles in Fig. 1), while the remaining 60% of stations (125 sites) are used for the out-of-sample validation. The 84 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

$$\text{ITS}_i = \sqrt{\text{MRD}_i^2 + \text{SDRD}_i^2}, \quad (5)$$

$$\text{MRD}_i = \frac{1}{T} \sum_{j=1}^T \text{RD}_{ij}, \quad (6)$$

$$\text{SDRD}_i = \sqrt{\frac{1}{T-1} \sum_{j=1}^T (\text{RD}_{ij} - \text{MRD}_i)^2}, \quad (7)$$

$$\text{RD}_{ij} = \frac{\theta_{ij} - \bar{\theta}_j}{\bar{\theta}_j}, \quad (8)$$

where θ_{ij} is an individual daily measurement of soil moisture at location $i \in [1, \dots, N]$ and time $j \in [1, \dots, T]$, and $\bar{\theta}_j$ is the spatial average of soil moisture at all locations at time j . The term RD_{ij} is the relative difference of location i at time j , which is introduced by Vachaud et al. (1985). The term 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. The term 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 (Brocca et al. 2012; Cho and Choi 2014; Penna et al. 2013).

A stratified random sampling method is adopted to select 84 modeling sites based on ITS following three steps: 1) calculate and rank the ITS of 209 stations in ascending order, 2) evenly divide the ranked ITS into four groups, and 3) within each group, randomly select 21 sites. The 84 sites are selected in this way to ensure evenly sampled sites across the ITS range, which best mimics the reality that in situ stations have different temporal representativeness.

3. Blending methods

The soil moisture blending schemes used in this study are illustrated in Fig. 2. Triple collocation is used to estimate the error variances of two sets of parent products, including NLDAS, SMOS, and API (N, S, A), and NLDAS, SMOS, and KAPI (N, S, K). The least squares weighting (LSW) is used to merge these products. We also evaluate equal weighting and an enumerated weighting scheme for soil moisture blending to investigate the sensitivity of the blended products to weighting schemes.

a. Triple collocation

Triple collocation is a technique for estimating the error variance of three independent datasets with respect to the unknown truth (Stoffelen 1998). It assumes a linear error model between each product and the unknown truth. The errors from the independent sources are assumed to have a mean of zero, and they are uncorrelated with each other and with the truth. Gruber et al. (2016a) reviewed previous TC analyses of soil moisture and found there are two different notations of TC formula, the difference notation (Scipal et al. 2008; Stoffelen 1998; Yilmaz et al. 2012) and the covariance notation (McColl et al. 2014; Stoffelen 1998). They demonstrated that the two different notations are mathematically identical. In this study, the covariance notation is used because it provides an extra estimate of correlation coefficients (R_i) of the measurement systems with respect to the unknown truth. The covariance notation for TC is based on the variances (σ_i^2) and covariances (σ_{ij}) of the parent products, which can be written as

$$\sigma_i^2 = \beta_i^2 \sigma_t^2 + \sigma_{\epsilon_i}^2, \quad (9)$$

$$\sigma_{ij} = \beta_i \beta_j \sigma_t^2, \quad (10)$$

where $i, j \in (N, S, A)$ or (N, S, K) in this study; and $i \neq j$. Letter N refers to NLDAS, S refers to SMOS, A refers to API, and K refers to KAPI. Here σ_i^2 is the variance of parent product i ; σ_t^2 is the variance of the unknown truth t ; σ_{ij} refers to the covariance between two parent products. Taking the triplets of (N, S, K) as an example, the unscaled error variances of each parent product ($\sigma_{\epsilon_N}^2, \sigma_{\epsilon_S}^2, \sigma_{\epsilon_K}^2$) can be derived by

$$\sigma_{\epsilon_N}^2 = \sigma_N^2 - \frac{\sigma_{NS} \sigma_{NK}}{\sigma_{SK}}, \quad (11)$$

$$\sigma_{\epsilon_S}^2 = \sigma_S^2 - \frac{\sigma_{NS} \sigma_{SK}}{\sigma_{NK}}, \quad (12)$$

$$\sigma_{\epsilon_K}^2 = \sigma_K^2 - \frac{\sigma_{KN} \sigma_{KS}}{\sigma_{NS}}. \quad (13)$$

Note that the covariance notation does not require an a priori rescaling of the datasets, and the error variances obtained from Eqs. (11)–(13) are absolute error variances, which are estimated in their own data space. To facilitate soil moisture blending, they must be converted to a common data space, or to relative error variances. Here, the NLDAS is used as the reference dataset and the rescaling parameters for the other two parent products are computed as (Gruber et al. 2016a; Stoffelen 1998) follows:

$$\beta_N^* = \frac{\beta_N}{\beta_S} = \frac{\sigma_{NK}}{\sigma_{SK}}, \quad (14)$$

$$\beta_K^* = \frac{\beta_N}{\beta_K} = \frac{\sigma_{NS}}{\sigma_{KS}}. \quad (15)$$

The covariance-based TC approach also provides important new information about the performance of the measurement systems, which is the correlation (R_i) between each parent product and the true soil moisture signal. Using the triplets of NLDAS (N), SMOS (S), and in situ measurements (I) as an example, the TC-based correlation for the three parent products can be calculated according to McColl et al. (2014) as

$$R_N = \sqrt{\frac{\sigma_{NS} \sigma_{NI}}{\sigma_N^2 \sigma_{SI}}}, \quad (16)$$

$$R_N = \sqrt{\frac{\sigma_{NS} \sigma_{NI}}{\sigma_N^2 \sigma_{SI}}}, \quad (17)$$

$$R_I = \sqrt{\frac{\sigma_{NI} \sigma_{SI}}{\sigma_I^2 \sigma_{NS}}}. \quad (18)$$

Here, R_i estimates that exceed 1, which may be due to statistical sampling errors (Gruber et al. 2018), are set to 1 (Gruber et al. 2019); only the positive values of the R_i is used, since the parent products are expected to be positively correlated to the unobserved truth.

Two sets of triplets, (N, S, A) and (N, S, K), are evaluated in this study to examine whether incorporating in situ soil moisture (e.g., using KAPI) can improve the accuracy of the blended product over using API, which does not incorporate in situ soil moisture. All parent products are evaluated at the same 125 out-of-sample stations. The stations with less than 100 observations are removed from the TC error estimation (Scipal et al. 2008). A more detailed discussion of the covariance-based TC approach and TC-based correlation coefficient and can be found in McColl et al. (2014) and Gruber et al. (2016a).

b. Least squares weighting

Least squares weighting (LSW) is a commonly used method for blending different soil moisture data sources (Gruber et al.

2019; Yilmaz et al. 2012; Zeng et al. 2016). The least squares framework can be expressed as

$$S_m = w_x S_x + w_y S_y + w_z S_z, \quad (19)$$

where S_m is the merged product; w_x , w_y , and w_z are the relative weights of three parent datasets S_x , S_y , and S_z , respectively. With $w_x + w_y + w_z = 1$, and by minimizing a cost function and the partial derivative of the cost function with respect to w_x and w_y , the optimal estimation of the weights is obtained from

$$w_x = \frac{\sigma_y^2 \sigma_z^2}{\sigma_x^2 \sigma_y^2 + \sigma_x^2 \sigma_z^2 + \sigma_y^2 \sigma_z^2}, \quad (20)$$

$$w_y = \frac{\sigma_x^2 \sigma_z^2}{\sigma_x^2 \sigma_y^2 + \sigma_x^2 \sigma_z^2 + \sigma_y^2 \sigma_z^2}, \quad (21)$$

$$w_z = \frac{\sigma_x^2 \sigma_y^2}{\sigma_x^2 \sigma_y^2 + \sigma_x^2 \sigma_z^2 + \sigma_y^2 \sigma_z^2}, \quad (22)$$

where σ_x^2 , σ_y^2 , and σ_z^2 are the TC-estimated error variance for the two sets of parent datasets [(N, S, A) and (N, S, K)].

c. Enumerated weighting

An enumerated weighing scheme is evaluated in this study to examine whether the TC or equal weighting methods have achieved optimal weights, and to test the sensitivity of hybrid products to weighting schemes. The weight of each parent product is varied over the full range from 0 to 1 with pre-defined intervals. Taking (N, S, K) as an example, as the sum of weights of three parent products should be 1 ($W_N + W_S + W_K = 1$), and the enumerated weighting scheme can be expressed as follows:

$$W_N = 0 - 1 \quad (\Delta W_N = 0.05), \quad (23)$$

$$W_S = \alpha_S (1 - W_N) \quad (\alpha_S = 0 \text{ to } 1, \Delta \alpha_S = 0.05), \quad (24)$$

$$W_K = 1 - W_N - W_S, \quad (25)$$

where ΔW_N and $\Delta \alpha$ denote the fixed intervals (both set to 0.05) of varying weights of NLDAS (W_N) and the proportion of the weight of SMOS (α_S), respectively. With this setting, a total of 441 weighing combinations are generated. (A visual display of the enumerated weight scheme is presented as the background color in Figs. 10 and 11).

d. Goodness of fit

The mean absolute error (MAE) [Eq. (26)] and Pearson correlation coefficient (R_{ei}) [Eq. (27)] between blended products and the out-of-sample in situ measurements are used for the validation and comparison of the blended products:

$$MAE = \frac{1}{n} \sum_{j=1}^n |\theta_o(j) - \theta_e(j)|, \quad (26)$$

$$R_{ei} = \frac{\sigma_{oe}}{\sigma_o \sigma_e}, \quad (27)$$

where n is the number of observations; θ_o and θ_e are observed and estimated soil moisture, respectively; σ_{oe} is the covariance

between observed and estimated soil moisture; and σ_o and σ_e are the standard deviations of observed and estimated soil moisture, respectively. Considering the in situ data have representativeness errors (Crow et al. 2012b; Miralles et al. 2010; Yilmaz et al. 2012), we followed the approach of Gruber et al. (2019) and used Bayesian inference [Eq. (28)] to correct the Pearson correlation (R_{ei}) for the impact of the in situ representativeness error (R_i):

$$R_e = \frac{R_{ei}}{R_i}, \quad (28)$$

where R_e is the corrected correlation between each blended product and the true soil moisture at 4-km scale; R_i is the representativeness error in the in situ measurements in terms of their correlation with true soil moisture at 4-km scale. The value of R_i is calculated by applying TC to the in situ measurements together with NLDAS and SMOS products using Eq. (18).

The Nash–Sutcliffe efficiency (NSE) score [Eq. (29)] is also used to evaluate the impact of sampling schemes and station density on the accuracy of the blended soil moisture products. NSE is a dimensionless indicator of model skill and can be used to assess the products with different units and scales (e.g., absolute soil moisture, anomalies, and percentiles). Its value ranges from $-\infty$ to 1. An NSE of 1 corresponds to perfect skill; an NSE of 0 indicates that the blended product performs the same as using the mean of out-of-sample observations, while a negative NSE ($NSE < 0$) indicates the model prediction is less accurate than the out-of-sample mean.

$$NSE = 1 - \frac{\sum_{i=1}^n [\theta_o(i) - \theta_e(i)]^2}{\sigma_o^2}. \quad (29)$$

4. Results and discussion

a. Comparing parent datasets

Figure 3 presents the spatial pattern of average daily soil moisture from 2011 to 2018 for in situ measurements over 209 stations (Fig. 3a), NLDAS soil moisture (Fig. 3b), the SMOS product (Fig. 3c), KAPI (Fig. 3d), and API (Fig. 3e). In general, all five datasets show a strong moisture gradient from west to east. The SMOS and NLDAS provide stronger soil moisture contrasts and greater spatial variability in soil moisture, while the API and KAPI present a smoother moisture pattern. NLDAS matches well with the spatial pattern of in situ soil moisture variations. SMOS tends to overestimate the dry conditions in the Texas and Oklahoma Panhandles, as well as along the east border of Oklahoma, while KAPI tends to overestimate soil moisture in west Texas.

Figure 4 provides a quantitative comparison of the five datasets in terms of mean (Fig. 4a), standard deviation (STD) (Fig. 4b), and coefficient of variation (CV) (Fig. 4c) over the 125 out-of-sample stations. Here we use absolute soil moisture for the evaluation. Both NLDAS and KAPI have a similar median value to that of the in situ measurements

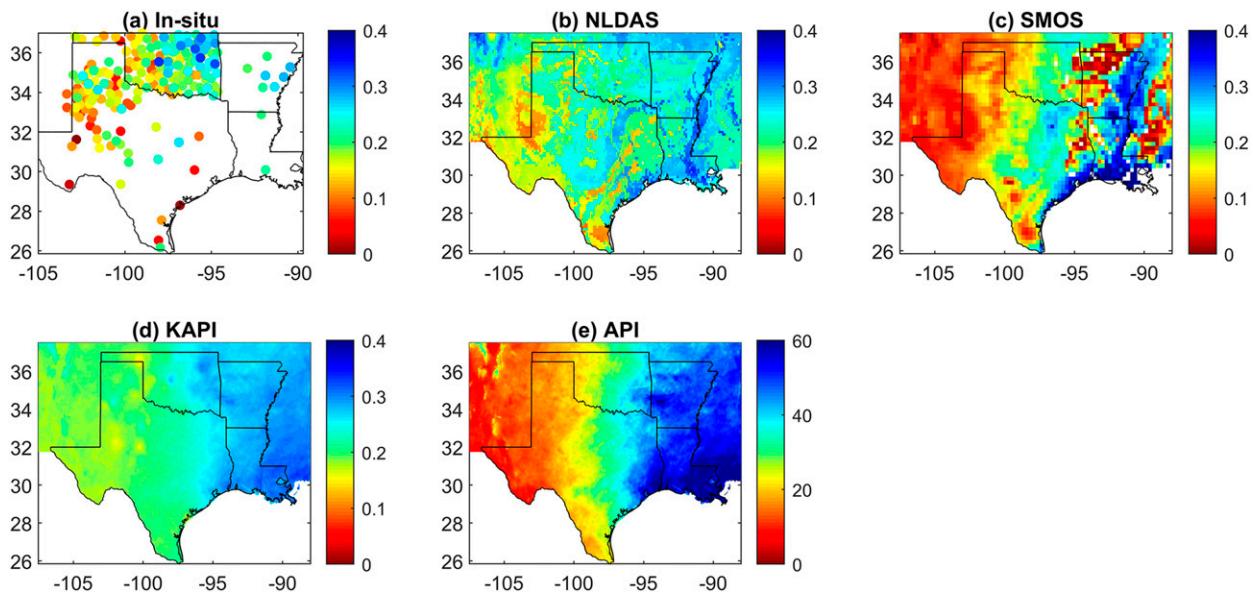


FIG. 3. Mean soil moisture calculated by averaging the daily data from 2011 to 2018 over the study area based on (a) in situ measurements at 209 stations, (b) NLDAS, (c) SMOS, (d) KAPI, and (e) API.

($0.22 \text{ m}^3 \text{ m}^{-3}$ for median), while SMOS presents a significant ($p < 0.05$) underestimation of mean soil moisture ($0.18 \text{ m}^3 \text{ m}^{-3}$) compared with the in situ ones. The negative bias in the SMOS L3 data indicates that this product may overestimate the dryness if used alone without normalization or standardization. NLDAS has the lowest STD ($0.04 \text{ m}^3 \text{ m}^{-3}$ for median), followed by KAPI and in situ data (both are $0.07 \text{ m}^3 \text{ m}^{-3}$), and SMOS has the largest STD ($0.09 \text{ m}^3 \text{ m}^{-3}$). However, it should be noted that the in situ data may be biased toward the KAPI because they are generated using in situ soil moisture.

Due to the different units of API, it is not directly comparable with other datasets in terms of mean and STD, but it can be compared using CV (Fig. 4c). NLDAS has the lowest CV (0.21 for median), followed by KAPI (0.29). In comparison, the in situ data have a CV of 0.31. The higher CV of in situ compared with NLDAS and KAPI may be attributed to it being a point measurement. Both SMOS (0.54) and API (1.02) show significantly ($p < 0.05$) larger CV than the other datasets. The

large CV of SMOS indicates there is a large degree of variability in SMOS soil moisture, which is jointly attributed to its small mean value (Fig. 4a) and large STD (Fig. 4b). The largest CV of API may be explained by its physical discrepancy with soil moisture measures, since API is calculated using precipitation, which does not have an upper limit, whereas soil moisture does. This may result in a wider range of API values corresponding to saturated soil, and thus larger CV of API than other soil moisture data. In summary, Fig. 4 reveals that both NLDAS and KAPI have a similar mean and CV as the in situ measurements, while SMOS tends to underestimate the mean values and has greater dispersion than the in situ data. API has the least agreement with the in situ data.

b. Errors variances from TC

Figure 5 shows the impact of different parent datasets on the error variance (σ_e^2) estimated using the TC approach. Four different sets of soil moisture anomaly products are compared: (N, I, S), (N, I, A), (N, S, A), and (N, S, K). We found that

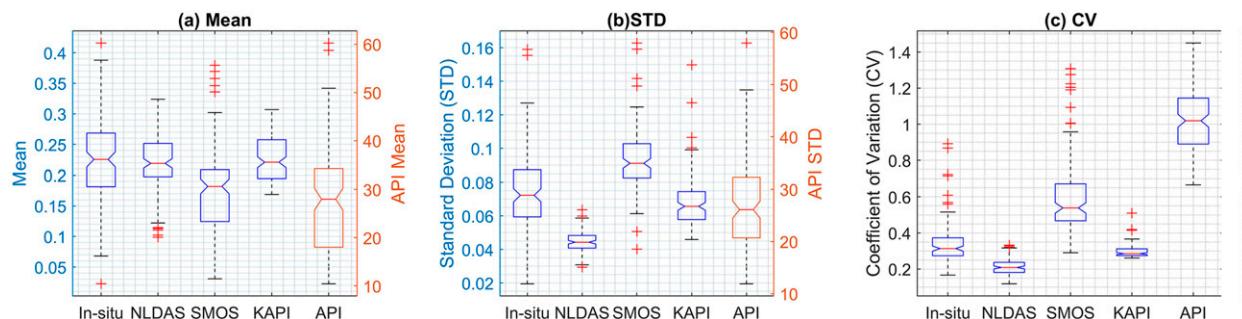


FIG. 4. Comparison of absolute soil moisture at the 125 out-of-sample sites based on in situ measurements, NLDAS, SMOS, KAPI and API: (a) mean, (b) standard deviation (STD), and (c) coefficient of variation (CV).

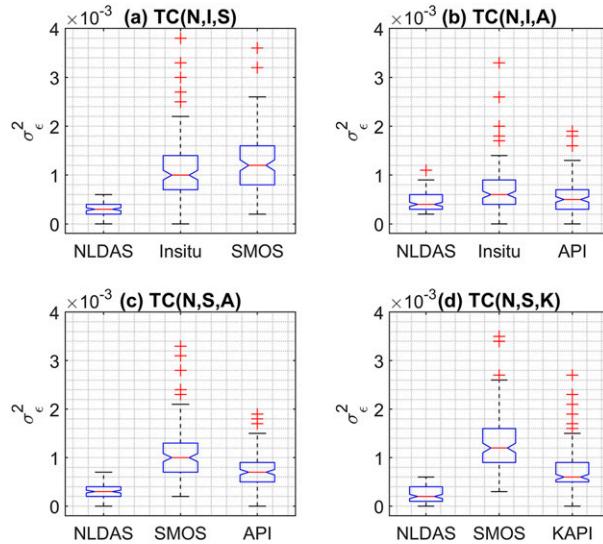


FIG. 5. Triple collocation (TC)-estimated error variances (σ_e^2) for soil moisture anomalies using different combinations of the five datasets [NLDAS (N), in situ (I), SMOS (S), API (A), and KAPI (K)]. The results are shown for the 125 out-of-sample stations.

changing the parent triplets does not change the relative error ranking of the parent datasets, but it does change the absolute values of the σ_e^2 estimated for each product. For example, the σ_e^2 of SMOS is consistently larger than that of NLDAS ($\sigma_{es}^2 > \sigma_{en}^2$), irrespective of the third dataset. The in situ data always have a larger σ_e^2 than NLDAS (Figs. 5a,b). This may be due to the high spatial representativeness errors of in situ measurements (Crow et al. 2012b; Miralles et al. 2010; Yilmaz et al. 2012). To reduce the impact of the representativeness errors in the in situ data, Bayesian inference [Eq. (28)] is adopted to derive the corrected correlation (R_e) between each dataset and the true soil moisture signal. We apply this approach to validate the blended products (section 4d). In general, the relative σ_e^2 ranking of the five datasets is $\sigma_{en}^2 < \sigma_{ek}^2 < \sigma_{ea}^2 < \sigma_{ei}^2 < \sigma_{es}^2$ based on Fig. 5.

c. Weights from LSW

The LSW scheme is adopted with TC-estimated error variance to calculate the weights of the two triplets (N, S, A) and (N, S, K) (Fig. 6). According to LSW theory, the weight of the parent products is inversely proportional to the error variance. Therefore, for the (N, S, A) triplets, NLDAS has the highest weight (0.55 for median), followed by API (0.25), and SMOS has the smallest weights (0.18) (Fig. 6a). Spatially, using the (N, S, A) triplets, the largest weights (>0.8) for NLDAS are clustered in south-central Texas (Fig. 6b), and the highest weights (>0.8) for API are concentrated in the northwest corner and eastern part of the study area (Fig. 6d), while the weights of SMOS are about 0.2 throughout the study area (Fig. 6c). For the (N, S, K) triplets, the largest weights are again given to NLDAS (0.63 for median), followed by KAPI (0.22), and SMOS has the smallest weights (0.12) (Fig. 6e). Spatially, using the (N, S, K) triplets, the

largest weights (>0.8) for NLDAS are located in the western half of the study area (Fig. 6f), while the weights of SMOS are smaller than 0.2 throughout the study area (Fig. 6g). For KAPI, the largest weights are scattered throughout the eastern part of the study area (Fig. 6h). In general, the two triplets (N, S, A) and (N, S, K) present similar weight patterns.

d. Evaluation of blended products

1) COMPARISON OF TC TRIPLETS

In this study, blended soil moisture products are generated from two triplets, (N, S, A) and (N, S, K), using the TC-LSW approach. For comparison, we also employ the simple average (AVE) approach, which gives equal weight to each product. Figure 7 compares the spatial pattern of the four parent products (first row) and the blended products (second row) using TC and AVE. The NLDAS soil moisture has been used as a reference to rescale the other parent products. In Fig. 7, the soil moisture anomalies on 4 July 2014 are used as an example. Distinct differences are observed between the four parent datasets (Figs. 7a–d). For example, scaled SMOS shows much drier conditions (indicated by blue color) in western and southern Texas than NLDAS (Fig. 7b). The scaled API shows wetter condition (indicated by red color) in northern Texas and Oklahoma and in eastern Arkansas than NLDAS (Fig. 7c), while the scaled KAPI has a similar pattern as API (Fig. 7d). In contrast, no apparent visual differences can be observed between all the blended products (second row of Fig. 7), regardless of the triplets or weighting scheme used. Therefore, further quantitative evaluation of the blended results is necessary.

Figure 8 provides a quantitative evaluation of the parent (first column) and blended soil moisture anomalies (last two columns) using the MAE, Pearson correlation (R_{ei}) against in situ measurements, and corrected correlation against the truth (R_e) over the 125 out-of-sample stations. The scaled KAPI presents the lowest MAE, while the scaled SMOS has the largest MAE. Based on an ANOVA test, the scaled SMOS has significantly ($p < 0.05$) higher MAE than other three parent products, and the scaled KAPI presents significantly ($p < 0.05$) lower MAE than API. But the difference between scaled NLDAS and scaled API and KAPI in terms of MAE is not significant ($p > 0.05$) (Fig. 8a). For the blended products, there is no significant ($p > 0.05$) difference in terms of MAE either between the results using TC or AVE, or between using different triplets (N, S, A) or (N, S, K). However, there is a slight improvement in MAE of the (N, S, K) blended results as compared to the parent products (Fig. 8c vs Fig. 8a). The negligible difference between TC and AVE methods agrees with Yilmaz et al. (2012). They found that merged soil moisture generated using a TC-LSW method did not outperform the soil moisture generated using an equally weighted approach.

The evaluation based on R_{ei} (second row in Fig. 8) tells a somewhat different story. Based on an ANOVA test, all the parent products have significantly ($p < 0.05$) different R_{ei} values from each other (Fig. 8d), except the pair between

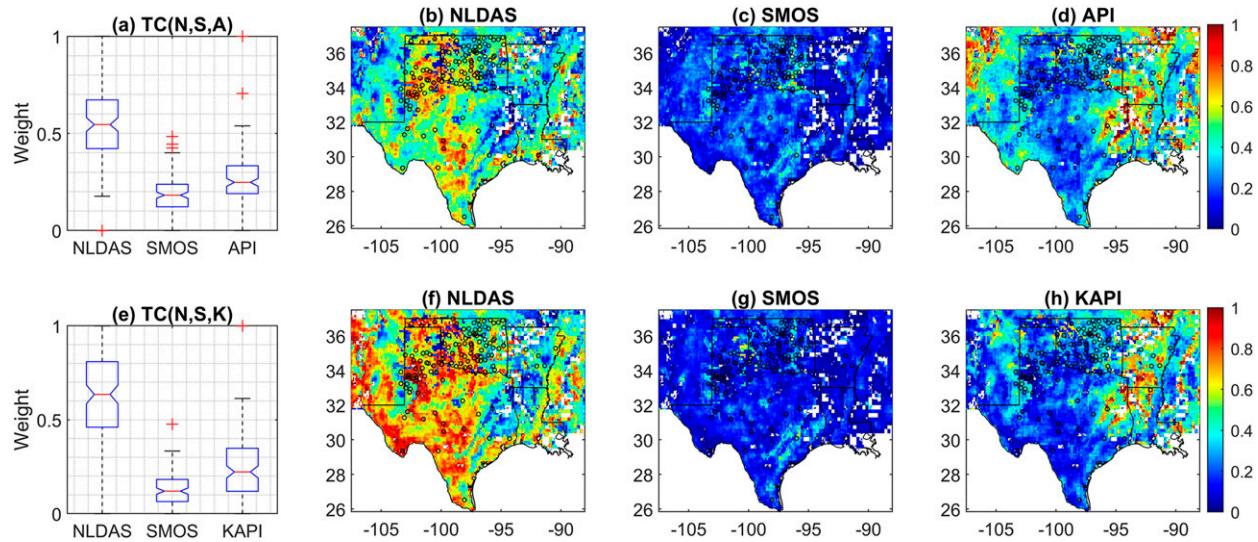


FIG. 6. Comparison of the least squares weighting (LSW)-derived weights based on the TC-estimated error variance using soil moisture anomalies. (a),(e) The weights at 125 out-of-sample stations (black circles in remaining subplots). (b)–(d), (f)–(h) Spatial distribution of weight for each product. Two sets of TC triplets are compared: (N, S, A) in (a)–(d) and (N, S, K) in (e)–(h).

NLDAS and API. The scaled KAPI has the highest R_{ei} (0.78 for median), followed by scaled NLDAS (0.67), scaled API (0.66), and scaled SMOS (0.54). There is no significant ($p > 0.05$) difference in R_{ei} between the blended products using TC or AVE methods. However, a significant ($p < 0.05$) improvement in R_{ei} is observed when using (N, S, K) (median $R_{ei} = 0.75$), instead of (N, S, A) (median $R_{ei} = 0.70$). This indicates that KAPI is superior to API in soil moisture blending, and in situ data can significantly improve the accuracy of the blended soil moisture datasets.

It is worth noting that the parent product KAPI performs best among all parent and blended products, and the insignificant improvement of the blended results over the KAPI parent product is worth exploring. Here, we propose three possible explanations. 1) The validation data are not spatially exhaustive. Although 60% (125) of total stations have been used in the validation, they are still relatively sparse and not evenly distributed in the study area. Figure 1 shows most validation stations are clustered in Oklahoma and west Texas, while relatively few stations are located in south Texas,

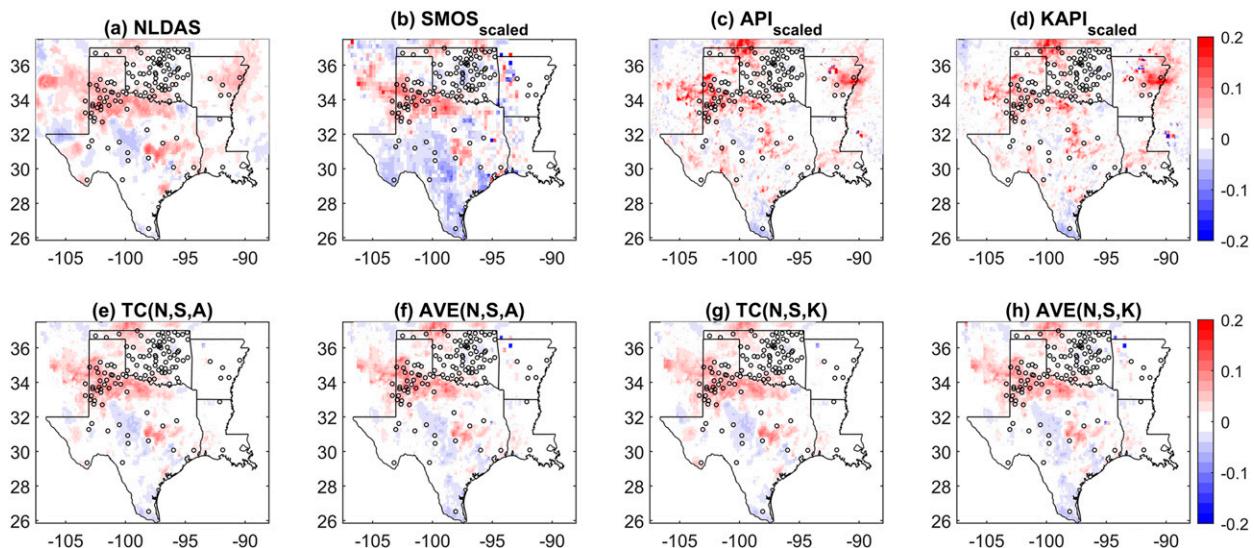


FIG. 7. Example of soil moisture anomalies ($\text{m}^3 \text{m}^{-3}$) on 4 July 2014 for the (a)–(d) parent products and (e)–(h) blended products. “TC” indicates that the blended maps were created using TC errors with LSW weighting; “AVE” indicates that the blended maps were created using equal weighting of the three parent products. The empty circles indicate the locations of the 125 out-of-sample stations. “N” refers to NLDAS, “S” refers to SMOS, “A” refers to API, and “K” refers to KAPI.

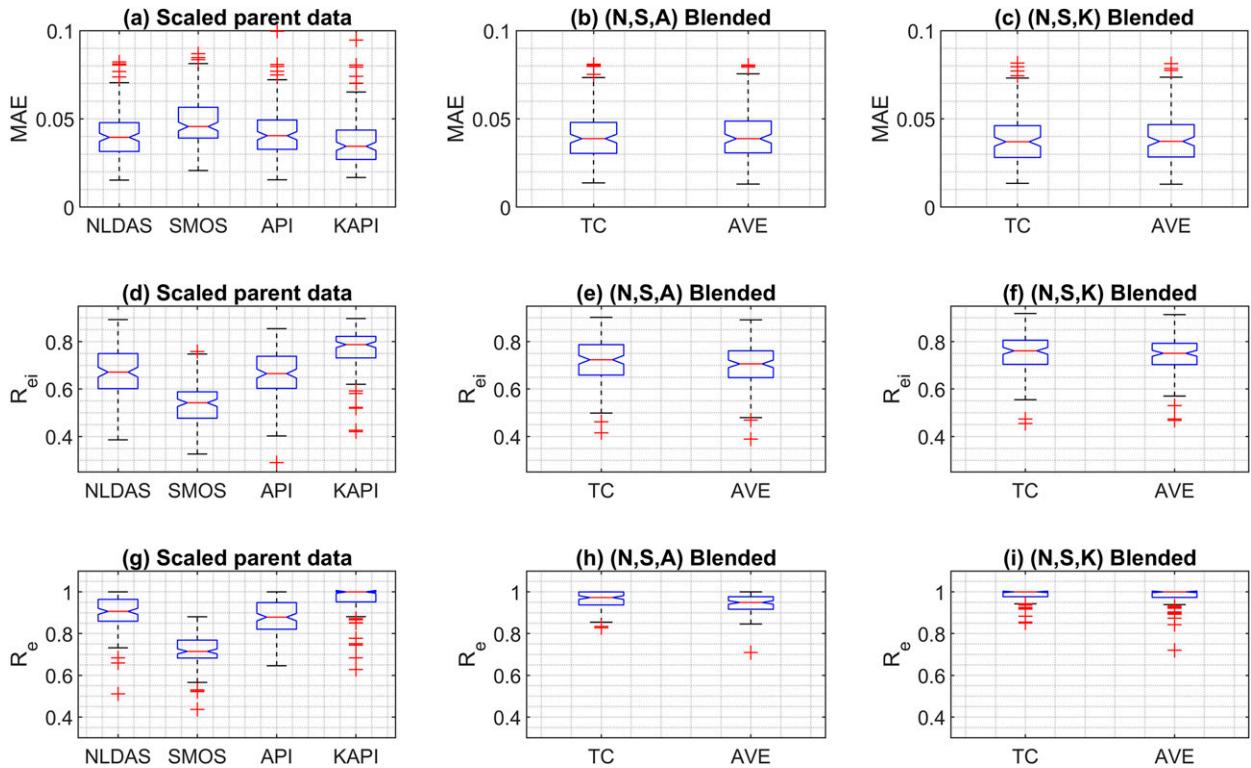


FIG. 8. Evaluation of the (left) parent and (center), (right) blended soil moisture anomalies ($\text{m}^3 \text{m}^{-3}$) using the 125 out-of-sample in situ stations based on (a)–(c) mean absolute error (MAE), (d)–(f) Pearson correlation coefficient (R_{ei}), and (g)–(i) corrected correlation against the truth (R_e). The calculations of MAE, R_{ei} , and R_e are provided in section 3d.

Arkansas, and Louisiana. It is possible that the places where hybrid results showed an improvement over the KAPI are not well captured using only 125 stations. 2) The in situ measurements cannot be considered the “truth” because they are point measurements that may not reflect the soil moisture value at 4-km grid cell. 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. Therefore, the in situ data have representativeness errors (Crow et al. 2012b; Miralles et al. 2010; Yilmaz et al. 2012), which should be removed before the validation. 3) TC is constructed based on several assumptions, such as a linear error model between the parent product and the unknown truth; the errors from the independent sources are assumed to have zero mean and are uncorrelated with each other and with the truth. Any violations of the above assumptions may cause the suboptimal estimation of the error variance.

To reduce the impact of the representativeness error of in situ data, the corrected correlation R_e [Eq. (28)] is used for evaluation. When using R_e , the absolute values of correlation increase substantially for all datasets (Figs. 8g–i), but the overall patterns are the same as those based on R_{ei} . It is interesting to note that after correlation correction, the R_e of scaled KAPI and blended products using (N, S, K) all reach 1, which indicates a perfect correlation with the true soil moisture signal.

2) ERROR SPATIAL PATTERN

The spatial error pattern for the (N, S, K) blended products (red circles) and KAPI (blue circles) are shown in Fig. 9. In places where in situ measurements are dense, especially in Oklahoma and northwest Texas, the MAE of the KAPI (blue circles) is generally similar to or smaller than that of AVE (K, S, N) (red circles). However, in places where in situ measurements are relatively sparse, such as central to southern Texas, northern Arkansas, and Louisiana, the MAE of KAPI is generally larger than that of AVE (N, S, K). This indicates that the blended data are less sensitive to station density and perform better in places where in situ measurements are sparse than the KAPI.

Figure 9 also shows how the ITS values (indicated by the size of the gray dots) compare with the MAE. A site with a smaller ITS value indicates greater temporal stability (i.e., it is a more temporally representative site). There is a positive relationship between MAE and ITS for both the KAPI and the (N, S, K) blends. Sites with higher ITS values (i.e., that are less temporally stable) have higher MAE values. This result agrees with the previous studies that demonstrated that kriging performance declines as the data variation increases (Gotway et al. 1996; Keskin and Grunwald 2018; Li and Heap 2011; Martínez-Cob 1996; Schläpfer and Schmid 1999). This result could also inform site selection for new soil moisture stations. For

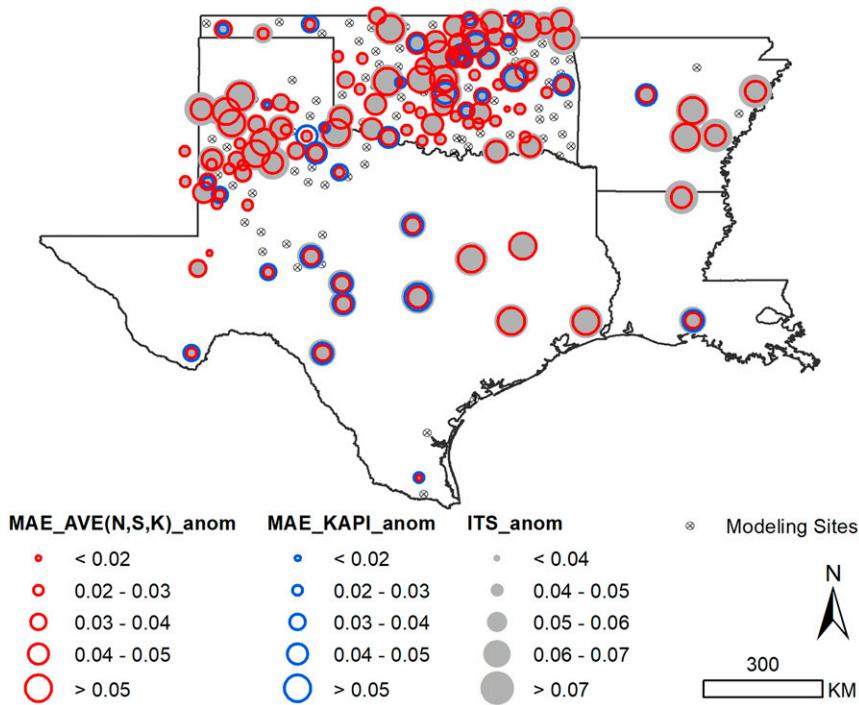


FIG. 9. Spatial distribution of the MAE of soil moisture anomalies ($\text{m}^3 \text{m}^{-3}$) based on the equal weighting of KAPI and NLDAS [AVE(K, N); red circles], KAPI (blue circles), and the index of temporal stability (ITS) (gray circles).

example, the ITS can be calculated using a soil moisture product such as NLDAS, and it can be used to identify locations with a small ITS value if building a new soil moisture network is desired, or to identify locations with a higher ITS if the goal is to address gaps in an existing soil moisture network.

e. Blending scheme sensitivity analysis

This section evaluates how the weighting scheme, sampling scheme, station density, and data formats influence the accuracy of the blended soil moisture products. The goal is to quantify the sensitivity and to identify the optimal blending approach.

1) IMPACT OF WEIGHTING SCHEMES

An enumerated weighing scheme [Eqs. (23)–(25)] is used to evaluate the weights of TC and AVE proposed in section 4d. There are a total of 441 weighting combinations that are considered. Figure 10 shows how the MAE of the blended product varies as a function of the weighting scheme. The soil moisture anomalies are used here as an example. The MAE is sensitive to the weighting scheme, and it ranges from 0.032 to $0.058 \text{ m}^3 \text{m}^{-3}$. The MAE is negatively correlated with the weight of KAPI. The local minimum MAE, denoted by the blue dots in Fig. 10a, tends to be associated with the weighting schemes where SMOS is given little weight. The optimal weights are identified when the minimum MAE is reached. Both local (Fig. 10b) and global (Fig. 10c) minimum MAEs are found with optimal weights of 0.25, 0.71, and 0.04 for NLDAS,

KAPI, and SMOS, respectively. Neither the TC (black triangle) nor AVE (black dot) method reaches optimal weighting; however, they do approach the local minimum MAE and have the potential to reach the optimal weighting. In addition, AVE (N, K), denoted by the red square in Fig. 10a, reaches the local minimum MAE, which is also lower than either the MAE of AVE (N, K, S) or TC (N, K, S). This indicates that SMOS is not an important contributor to the optimal blended products within our study area. However, the conclusion may be different if the study region has relatively few in situ stations and/or is heavily irrigated. In that case, the importance of SMOS may increase since SMOS can capture soil moisture variations due to human activity (e.g., irrigation).

Figure 11 evaluates the impact of weighting schemes on the blended products using Pearson correlation (R_{ei}). The findings are generally consistent with those based on MAE. The R_{ei} ranges from 0.54 to 0.80, depending on the sampling scheme. A positive correlation is observed between R_{ei} and the weight of KAPI. The local (global) optimal weights are 0.25 (0.3) for NLDAS, 0.75 (0.8) for KAPI, and zero for SMOS. The weighting schemes using TC and AVE are similar to the local optimal weights.

2) IMPACT OF SAMPLING SCHEME AND STATION DENSITY

Figure 12 shows the impact of using different sampling schemes and different number of stations on the accuracy of the blended products using Nash–Sutcliffe efficiency (NSE)

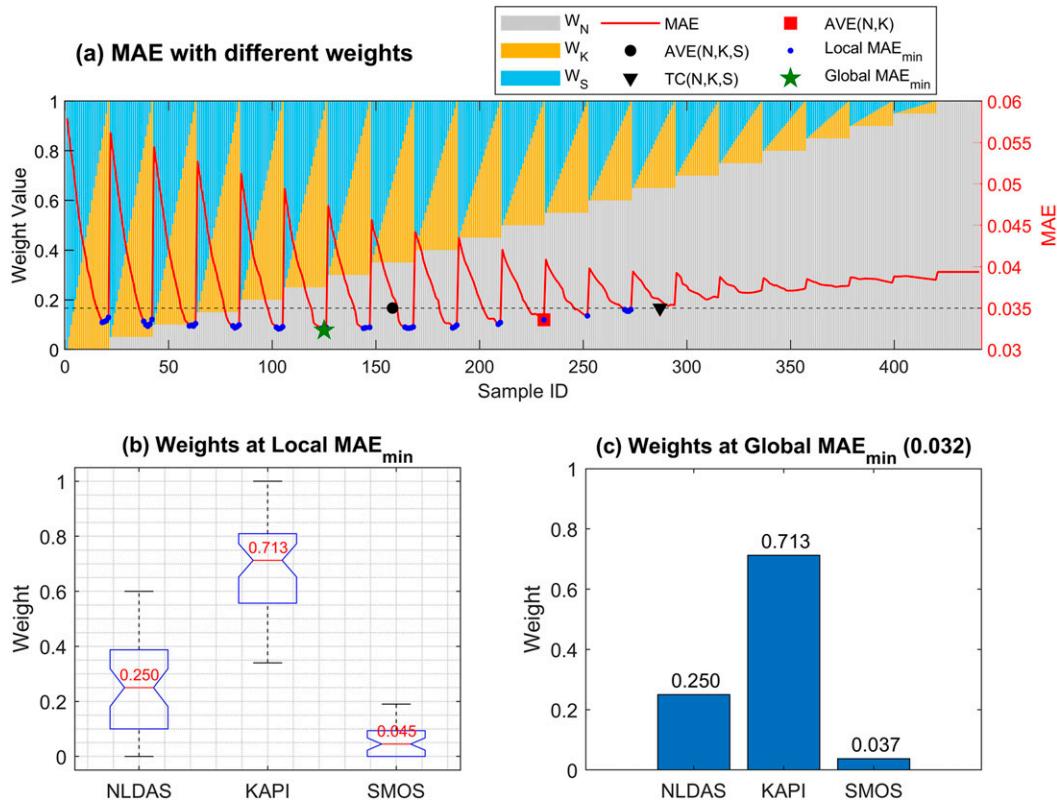


FIG. 10. (a) Mean absolute error (MAE) of the blended soil moisture anomalies ($\text{m}^3 \text{m}^{-3}$) based on an enumerated weighting scheme, (b) the weight of each parent product at the local minimum MAE [blue dots in (a)], and (c) the weight of each parent product at the global minimum MAE [green star in (a)]. Note that in (a) the black dot indicates the MAE of the blended product that is generated from equal weighting of NLDAS, KAPI, and SMOS [AVE(N, K, S)]. The black triangle indicates the MAE of the blended product that is generated from TC-based LSW of NLDAS, KAPI, and SMOS [TC(N, K, S)], and the red square indicates the MAE of the blended product that is generated from equal weighting of NLDAS and KAPI [AVE(N, K)]. The local minimum MAEs are defined as the MAE values at the valley of each MAE cycle and are smaller than the MAE of AVE(N, K, S); the global minimum MAE is the smallest MAE value among all weighting schemes.

and mean absolute error (MAE). MAE and NSE were calculated using soil moisture anomalies at the out-of-sample stations. Three different sampling schemes are evaluated. Head sampling means choosing stations from the head of ascendingly ranked ITS or choosing stations with the lowest ITS. Tail sampling means choosing stations from the tail of ascending ITS or choosing stations with the highest ITS. Even sampling indicates choosing stations evenly from the entire range of ranked ITS values. Head sampling first selects the most stable stations and then gradually adds stations more temporally variable stations, while tail sampling first selects the least stable stations for modeling and then adding more stable stations. In contrast, even sampling selects the stations that are most representative of the population distribution. For each sampling scheme, the number of stations is varied from 15 to 125, and the number of out-of-sample validation stations varies from 194 to 84.

As the number of stations (station density) increases, the NSE (Fig. 12a) also increases for both KAPI and AVE (K, N) using all sampling schemes. When used alone, the KAPI with

the tail sampling scheme has the lowest NSE values. However, when combined with NLDAS, the AVE (K, N) product using tail sampling shows a large improvement. Generally, the variation in the NSE of AVE (K, N) (red shaded area) under different sampling schemes is much smaller than that of KAPI (gray shaded area). This indicates that the blended products, such as AVE (K, N), can reduce the uncertainty in soil moisture estimates, which is especially helpful when there are very few stations or the stations are not representative.

The MAE tends to decrease as the number of stations increases for both KAPI and AVE (K, N) using both the tail and even sampling schemes (Fig. 12b). The greatest reduction in MAE is observed when using AVE (K, N) over KAPI with tail sampling, especially at lower station densities. For example, KAPI has an MAE of $0.06 \text{ m}^3 \text{ m}^{-3}$ using 15 stations with tail sampling, but the MAE drops to $0.0043 \text{ m}^3 \text{ m}^{-3}$ when AVE (K, N) is used. This indicates that when stations are sparsely distributed or less representative, adding an extra source of soil moisture information can greatly improve the accuracy. This also confirms that the blended product (K, S, N) performs

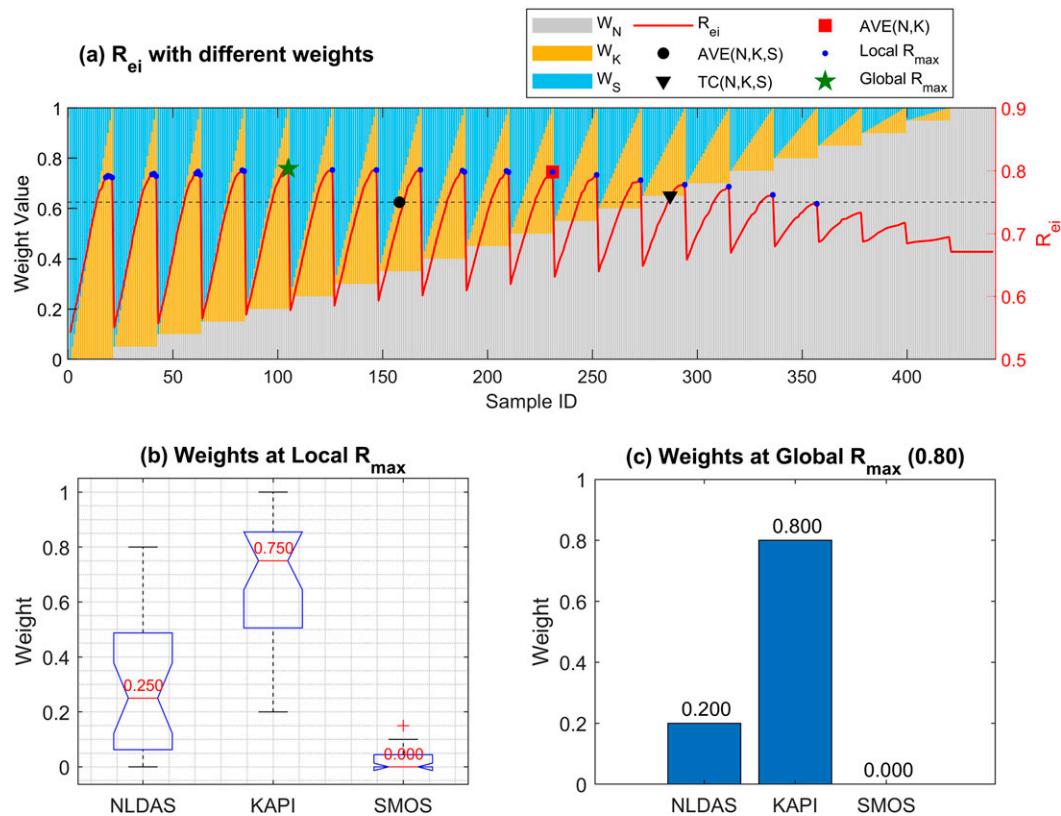


FIG. 11. (a) Pearson correlation coefficient (R_{ei}) of the blended soil moisture anomalies based on an enumerated weighting scheme, (b) the weight of each parent product at the local maximum R_{ei} [blue dots in (a)], and (c) the weight of each parent product at the global maximum correlation [green star in (a)]. In (a), the black dot indicates the Pearson correlation (R_{ei}) of the blended product that is generated from equal weighting of NLDAS, KAPI, and SMOS [AVE(N, K, S)]. The black triangle indicates the Pearson correlation of the blended product that is generated from TC-based LSW of NLDAS, KAPI, and SMOS [TC(N, K, S)] and the red square indicates the Pearson correlation of the blended product that is generated from equal weighting of NLDAS and KAPI [AVE(N,K)]. The local maximum R_{ei} are defined as the R_{ei} values at the peak of each R_{ei} cycle and are larger than the R_{ei} of AVE(N, K, S); the global maximum R_{ei} is the largest R_{ei} value among all weighting schemes.

better than KAPI in areas where in situ stations are sparse. This finding has practical significance for real-world applications because maintaining a dense and representative network of in situ measurements is challenging. The Oklahoma Mesonet is a unique and uncommonly dense network, but in most cases soil moisture stations are sparsely distributed. Therefore, a hybrid soil moisture product is likely to provide the most accurate results.

MAE tends to be higher for both KAPI and AVE (K, N) when head sampling is used (Fig. 12b). This is likely because there is greater heterogeneity in the validation data when the most representative stations are used for training the model. The remaining stations are less representative and have larger temporal variability, which may yield larger errors. In practice, the sampling sites are always a mix of more and less representative sites, which is similar to the even sampling case.

It is also worth noting that increasing station density has a limited impact on the accuracy of the KAPI. For example, the changes in NSE and MAE become more gradual when the

number of stations reaches ~ 55 (Figs. 12a,b). This finding is in agreement with Yuan and Quiring (2017).

3) IMPACT OF SOIL MOISTURE FORMAT

Caution should be taken when conducting error analysis based on absolute soil moisture. This is because the absolute soil moisture can be decomposed into its anomaly and climatological components. It may be necessary to distinguish the error attributed to the anomaly and climatological parts before it is compared with the soil moisture percentiles and anomalies. Although the soil moisture climatology is relatively smooth and thus has limited degrees of freedom, it is still possible that independent products have different climatological errors. For example, Dong et al. (2020) investigated the soil moisture climatological error in multiple data sources and found that the newer L-band remote sensing soil moisture products (e.g., derived from SMOS) have smaller climatological errors than those from older C-/X-band remote sensing products (e.g., AMSR-E) and smaller errors than the four land surface models as well.

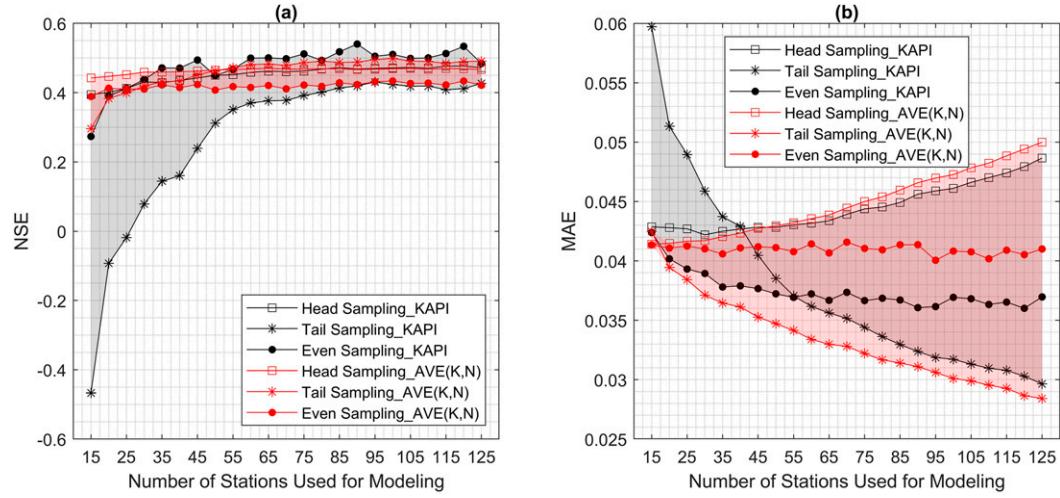


FIG. 12. Evaluation of KAPI and AVE(K, N) using different sampling schemes and different sampling sizes based on (a) Nash-Sutcliffe efficiency (NSE) and (b) mean absolute error (MAE). MAE and NSE were calculated using soil moisture anomalies at the out-of-sample stations. Head sampling (denoted by line with square) means choosing stations from the head of ascendingly ranked ITS or choosing stations with the lowest ITS. Tail sampling (denoted by line with star) means choosing stations from the tail of ascending ITS or choosing stations with the highest ITS. Even sampling (denoted by line with a dot) indicates choosing stations evenly from the entire range of ranked ITS values.

The impact of the soil moisture format (e.g., absolute, anomalies, and percentiles) on the blended results is shown in the supplementary figures. In general, the relative results are consistent regardless of the data format. For example, supplementary Fig. 1 (in the online supplemental material) demonstrates that the ranking of TC-estimated error variance and LSW-derived weights for the parent products are consistent regardless of whether soil moisture is expressed in absolute, anomaly, or percentile format. Supplementary Fig. 2 reveals that the MAE (R_{ei} and R_e) of the blended datasets using absolute soil moisture is higher (lower) than those using anomalies and percentiles. This may be due to the reduced systematic bias using anomalies and percentiles (Ford et al. 2015; Zhang et al. 2017a). However, the performance ranking of the different blended products is consistent regardless of which data format is used. Supplementary Figs. 3 and 4 indicate that although the values of the local/global optimal weights vary with the data format, the relative relationship still holds (e.g., the optimal weights are always larger for KAPI, smaller for NLDAS, and minimal for SMOS).

4) IMPACT OF ERROR CROSS-CORRELATION

The TC analysis requires triplets to have mutually uncorrelated random errors. The modeled, remote sensing, and in situ measurements are widely considered to have mutually independent error structures. However, in our previous analysis (e.g., Figs. 5–8), two modeled datasets (NLDAS soil moisture and API) are used in one TC triplet. To confirm that they have mutually independent errors, we calculated the error cross-correlation between NLDAS soil moisture and API using the extended collocation (EC) method proposed by Gruber et al. (2016b). The EC analysis provides the error

cross-correlation estimation in a least squares framework by generalizing the TC method to an arbitrary number of datasets and relaxing the assumption of zero error cross-correlation for the targeted datasets. Refer to Gruber et al. (2016b) for a detailed description of the method. In this study, four datasets are used in EC analysis, including NLDAS, SMOS, API, and in situ soil moisture, in which the modeled (either NLDAS or API), remote sensing (SMOS), and in situ soil moisture are considered to have zero error cross-correlation, while NLDAS and API is assumed to have nonzero cross-correlation. Using Eqs. (9) and (10) in Gruber et al. (2016b), the cross-correlation errors between NLDAS and API over 125 stations were derived with a median value about 0. Compared with the two examples from Gruber et al. (2016b), one case of high error cross-correlation between two AMSR-E products (median ≈ 0.8), and the other case of negligible cross-correlations between ASCAT and AMSR-E (median ≈ 0.25), the error cross-correlation between NLDAS and API (median ≈ 0.00) is minimal, and they can be considered independent. Therefore, NLDAS soil moisture and API have mutually uncorrelated random errors and can be used as triplets. Therefore, our analysis is valid (Figs. 5–8).

The reason that NLDAS soil moisture is relatively independent of API may be due to the following two reasons. 1) API and NLDAS models are at different levels of complexity. API is used as an approximation of soil moisture and is a simplified statistical model that is based solely on precipitation. In contrast, NLDAS soil moisture is simulated by complex land surface models that use multiple inputs, including precipitation, temperature, radiation, and other surface meteorological data. 2) Although PRISM climatology has been used for the orographic adjustment of the NLDAS precipitation forcing data (Xia et al. 2012), the adjustment is only applied in

mountainous regions (Daly et al. 1994), which is not an issue in our study region (south-central United States).

Finally, the independency between KAPI and validation in situ soil moisture is investigated. Here, each of the 125 validation sites is paired with each of the 84 modeling sites, resulting in a total of 10 500 ($= 125 \times 84$) station pairs. Then the correlation between in situ soil moisture anomalies at each station pair is calculated using all eight years (2011–18) of data (supplementary Fig. 5). Supplementary Fig. 5a shows that the correlations between stations tend to decrease as the distance between stations increases. The quadratic regression line explains approximately 60% of the variance in the correlation between stations. Based on the regression line, the station-paired correlation drops below 0.5 when the distance is larger than 125 km, and the station pairs with distance smaller than 125 km only take about 10% of the total station pairs. This indicates that most (90%) of the station pairs between modeling and validation sites have a distance larger than 125 km and a correlation smaller than 0.5. The correlation of 0.5 is selected according to Fig. 8d, where SMOS presented the lowest correlation with in situ soil moisture with a median value of 0.54. The in situ soil moisture is considered independent of SMOS with a correlation of 0.54, and thus the correlation of 0.5 can be used as a preliminary threshold to determine the dependency between two datasets. If the correlation is smaller than 0.5, the two datasets may be independent. Besides fitting the quadratic line (supplementary Fig. 5a), another approach (histogram) is adopted in supplementary Fig. 5b to determine the distance between station pairs with correlation less than 0.5. The result shows that the correlation of 0.5 falls within the 100–150-km distance bin (supplementary Fig. 5b). The accumulated percentage of station pairs is about 14% ($= 1.6\% + 5.1\% + 7.3\%$) when the correlation larger than 0.5, which indicates that about 86% of the station pairs have a correlation smaller than 0.5.

In summary, supplementary Fig. 5 demonstrates that most (86%–90%) of the station pairs between validation and modeling sites in this study are beyond a distance of 150 km and have a correlation less than 0.5. This indirectly demonstrates that the KAPI, although generated based on the in situ soil moisture at 84 modeling sites, is independent of the in situ soil moisture at the validation sites. Therefore, the in situ soil moisture can be used to validate the KAPI over the 125 validation sites.

5. Conclusions

This study identified the importance of incorporating in situ soil moisture into soil moisture blending, quantified the impact of different weighting schemes, and investigated the impact of sampling schemes, station density, and data format on soil moisture blending. Soil moisture information from multiple sources, including satellite (SMOS L3 SM), model (NLDAS-V2 Noah), and in situ measurements, as well as PRISM precipitation data, are used to generate blended soil moisture products at a 4-km spatial resolution and daily temporal resolution. TC was used to estimate the error variance of the parent products, and LSW was used to generate blended soil

moisture products. An equal weighting approach (AVE) was also compared with the TC-LSW approach. The results indicate that KAPI has the lowest error among all the parent products. The accuracy of the blended products improved when KAPI was incorporated, especially in terms of correlation. This indicates that in situ measurements are valuable for improving the accuracy of blended soil moisture datasets. Likewise, it is reasonable to posit that the inclusion of in situ soil moisture via data assimilation would also improve dataset accuracy. Although the in situ data are not regularly incorporated in SMAP L4, a recent study (Gruber et al. 2018) has incorporated ground-based soil moisture observation via data assimilation for operational agricultural drought monitoring.

The spatial patterns of error in the blended products are related to the station representativeness [i.e., the index of temporal stability (ITS)]. This indicates that ITS can be used to determine the optimal location of new soil moisture stations. Our results agree with previous study from Yee et al. (2016) that temporal stability is helpful to identify representative stations in the setup of common calibration/validation (cal/val) sites.

An enumerated weighting scheme was used to assess the sensitivity of the blended results to different weighting schemes. Our analysis demonstrated that the optimal weights are associated with the largest weights assigned to KAPI, followed by NLDAS, while SMOS receives marginal weight. The weighting schemes used by TC and AVE are both close to the local optimal weighting, and so they are appropriate for operational soil moisture mapping.

The sensitivity of soil moisture blending to sampling schemes, station density and data format were also investigated. Our results demonstrated that using multiple sources of soil moisture helps to reduce the overall uncertainty in the soil moisture estimates, especially when sampling density is low or the sampling stations are less representative. When sampling density is high, sampling schemes tend to have a greater impact on the accuracy of the soil moisture estimates. This finding has implications for real-world applications because achieving a high density of stations that are spatially representative is always challenging.

In contrast, the format of the soil moisture data does not have a major impact on the blending. Generally, our findings are robust and consistent regardless of whether soil moisture is in absolute, anomaly, or percentile format. The performance ranking of the blended products is the same regardless of the data format, and the optimal weights are always larger for KAPI, smaller for NLDAS, and minimal for SMOS.

These findings should be placed in context by acknowledging the study limitations, including the five listed here. 1) The soil moisture products used in this study were all extracted from 0700 LT observations. However, a temporal mismatch may still exist due to the different temporal resolutions of each soil moisture product. Future work can adopt methods to ensure the temporal coherence of different datasets. 2) This study only considered precipitation (API) in the soil moisture kriging. Other variables, such as evapotranspiration, soil texture, vegetation, and topography, may be helpful for soil moisture estimation (Ochsner et al. 2019) and should be considered to improve the kriging accuracy. 3) Geographically

weighted regression kriging (GWRK) (Brunsdon et al. 1996; Fotheringham et al. 2003) considers the spatially nonstationary relationships between the dependent variable and independent variables and weights the regression by their distance to the target point. Therefore, it may be more accurate than RK (Kumar et al. 2012; Yang et al. 2019) and should be explored in future studies. 4) Further study is required to test whether these conclusions are valid in other regions. The proposed approaches may be accurate and directly applied in regions with similar topography and climatology as the south-central United States. In fact, the methodology has been applied operationally to generate national gridded soil moisture percentiles (Zhao et al. 2020). However, caution should be taken when applying these methods in heavily irrigated regions, where human activities strongly control soil moisture variation. 5) Indeed, a primary advantage of blending in situ soil moisture is its representation of root zone conditions. The focus of our next step is to apply the methods developed here for blending root zone soil moisture from in situ and model sources. For example, our previous work (Zhang et al. 2017a) compared different methods to estimate root zone soil moisture using surface measurements and found that exponential filters outperformed the linear regression and artificial neural network. Therefore, exponential filters can be applied to surface remote sensing data to generate RZSM estimates. Together with the model simulated and in situ based root zone soil moisture, the method proposed in this study can be used to generate new root zone soil moisture dataset and analyze the blending sensitivity.

Acknowledgments. 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 NA17OAR4310136). 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. The authors declare that they have no conflict of interest.

Data availability statement. All datasets used in this study are publicly available. The SMOS data can be accessed using the Sipad at www.catds.fr/sipad/, which is a web-based interactive tool produced by Centre Aval de Traitement des Données SMOS (CATDS). 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/>).

REFERENCES

Abbaspour, K. C., E. Rouholahnejad, S. Vaghefi, R. Srinivasan, H. Yang, and B. Kløve, 2015: A continental-scale hydrology and water quality model for Europe: Calibration and uncertainty of a high-resolution large-scale SWAT model. *J. Hydrol.*, **524**, 733–752, <https://doi.org/10.1016/j.jhydrol.2015.03.027>.

Afshar, M. H., M. Yilmaz, and W. Crow, 2019: Impact of rescaling approaches in simple fusion of soil moisture products. *Water Resour. Res.*, **55**, 7804–7825, <https://doi.org/10.1029/2019WR025111>.

Albergel, C., and Coauthors, 2012: Evaluation of remotely sensed and modelled soil moisture products using global ground-based in situ observations. *Remote Sens. Environ.*, **118**, 215–226, <https://doi.org/10.1016/j.rse.2011.11.017>.

Al Bitar, A., and Coauthors, 2017: The global SMOS Level 3 daily soil moisture and brightness temperature maps. *Earth Syst. Sci. Data*, **9**, 293–315, <https://doi.org/10.5194/essd-9-293-2017>.

Anctil, F., C. Michel, C. Perrin, and V. Andréassian, 2004: A soil moisture index as an auxiliary ANN input for stream flow forecasting. *J. Hydrol.*, **286**, 155–167, <https://doi.org/10.1016/j.jhydrol.2003.09.006>.

Bi, H., J. Ma, W. Zheng, and J. Zeng, 2016: Comparison of soil moisture in GLDAS model simulations and in situ observations over the Tibetan Plateau. *J. Geophys. Res.*, **121**, 2658–2678, <https://doi.org/10.1002/2015JD024131>.

Brocca, L., F. Melone, T. Moramarco, W. Wagner, V. Naeimi, Z. Bartalis, and S. Hasenauer, 2010: Improving runoff prediction through the assimilation of the ASCAT soil moisture product. *Hydrol. Earth Syst. Sci.*, **14**, 1881–1893, <https://doi.org/10.5194/hess-14-1881-2010>.

—, T. Tullo, F. Melone, T. Moramarco, and R. Morbidelli, 2012: Catchment scale soil moisture spatial–temporal variability. *J. Hydrol.*, **422–423**, 63–75, <https://doi.org/10.1016/j.jhydrol.2011.12.039>.

Brunsdon, C., A. S. Fotheringham, and M. E. Charlton, 1996: Geographically weighted regression: A method for exploring spatial nonstationarity. *Geogr. Anal.*, **28**, 281–298, <https://doi.org/10.1111/j.1538-4632.1996.tb00936.x>.

Cartwright, J., and A. D. Fraser, 2021: Polar maps of C-band backscatter parameters from the Advanced Scatterometer. *Earth Syst. Sci. Data Discuss.*, <https://doi.org/10.5194/essd-2021-92>.

Chen, F., and Coauthors, 2017: Application of triple collocation in ground-based validation of Soil Moisture Active/Passive (SMAP) level 2 data products. *IEEE J. Selected Top. Appl. Earth Obs. Remote Sens.*, **10**, 489–502, <https://doi.org/10.1109/JSTARS.2016.2569998>.

Chen, Y., K. Yang, J. Qin, L. Zhao, W. Tang, and M. Han, 2013: Evaluation of AMSR-E retrievals and GLDAS simulations against observations of a soil moisture network on the central Tibetan Plateau. *J. Geophys. Res.*, **118**, 4466–4475, <https://doi.org/10.1002/jgrd.50301>.

Cho, E., and M. Choi, 2014: Regional scale spatio-temporal variability of soil moisture and its relationship with meteorological factors over the Korean peninsula. *J. Hydrol.*, **516**, 317–329, <https://doi.org/10.1016/j.jhydrol.2013.12.053>.

Cosh, M. H., and Coauthors, 2016: 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** (4), 1–11, <https://doi.org/10.2136/vzj2015.09.0122>.

Crow, W. T., and M. J. van den Berg, 2010: An improved approach for estimating observation and model error parameters in soil moisture data assimilation. *Water Resour. Res.*, **46**, W12519, <https://doi.org/10.1029/2010WR009402>.

—, S. Kumar, and J. Bolten, 2012a: On the utility of land surface models for agricultural drought monitoring. *Hydrol. Earth Syst. Sci.*, **16**, 3451–3460, <https://doi.org/10.5194/hess-16-3451-2012>.

—, and Coauthors, 2012b: Upscaling sparse ground-based soil moisture observations for the validation of coarse-resolution satellite soil moisture products. *Rev. Geophys.*, **50**, RG2002, <https://doi.org/10.1029/2011RG000372>.

Dai, A., 2013: Increasing drought under global warming in observations and models. *Nat. Climate Change*, **3**, 52–58, <https://doi.org/10.1038/nclimate1633>.

Daly, C., R. P. Neilson, and D. L. Phillips, 1994: A statistical-topographic model for mapping climatological precipitation over mountainous terrain. *J. Appl. Meteor. Climatol.*, **33**, 140–158, [https://doi.org/10.1175/1520-0450\(1994\)033<0140:ASTMFM>2.0.CO;2](https://doi.org/10.1175/1520-0450(1994)033<0140:ASTMFM>2.0.CO;2).

—, and Coauthors, 2008: Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States. *Int. J. Climatol.*, **28**, 2031–2064, <https://doi.org/10.1002/joc.1688>.

Dobriyal, P., A. Qureshi, R. Badola, and S. A. Hussain, 2012: A review of the methods available for estimating soil moisture and its implications for water resource management. *J. Hydrol.*, **458–459**, 110–117, <https://doi.org/10.1016/j.jhydrol.2012.06.021>.

Dong, J., W. T. Crow, and R. Bindlish, 2018: The error structure of the SMAP single and dual channel soil moisture retrievals. *Geophys. Res. Lett.*, **45**, 758–765, <https://doi.org/10.1002/2017GL075656>.

—, and Coauthors, 2020: Comparison of microwave remote sensing and land surface modeling for surface soil moisture climatology estimation. *Remote Sens. Environ.*, **242**, 111756, <https://doi.org/10.1016/j.rse.2020.111756>.

Dorigo, W., and Coauthors, 2011: The International Soil Moisture Network: A data hosting facility for global in situ soil moisture measurements. *Hydrol. Earth Syst. Sci.*, **15**, 1675–1698, <https://doi.org/10.5194/hess-15-1675-2011>.

—, and Coauthors, 2013: Global automated quality control of in situ soil moisture data from the International Soil Moisture Network. *Vadose Zone J.*, **12** (3), 1–21, <https://doi.org/10.2136/vzj2012.0097>.

Downer, C. W., and F. L. Ogden, 2003: Prediction of runoff and soil moistures at the watershed scale: Effects of model complexity and parameter assignment. *Water Resour. Res.*, **39**, 1045, <https://doi.org/10.1029/2002WR001439>.

Eldeiry, A. A., and L. A. Garcia, 2010: Comparison of ordinary kriging, regression kriging, and cokriging techniques to estimate soil salinity using LANDSAT images. *J. Irrig. Drain. Eng.*, **136**, 355–364, [https://doi.org/10.1061/\(ASCE\)IR.1943-4774.0000208](https://doi.org/10.1061/(ASCE)IR.1943-4774.0000208).

Entekhabi, D., and Coauthors, 2010: The Soil Moisture Active Passive (SMAP) mission. *Proc. IEEE*, **98**, 704–716, <https://doi.org/10.1109/JPROC.2010.2043918>.

Ford, T. W., and S. M. Quiring, 2014: Comparison and application of multiple methods for temporal interpolation of daily soil moisture. *Int. J. Climatol.*, **34**, 2604–2621, <https://doi.org/10.1002/joc.3862>.

—, and —, 2019: Comparison of contemporary in situ, model, and satellite remote sensing soil moisture with a focus on drought monitoring. *Water Resour. Res.*, **55**, 1565–1582, <https://doi.org/10.1029/2018WR024039>.

—, A. D. Rapp, and S. M. Quiring, 2015: Does afternoon precipitation occur preferentially over dry or wet soils in Oklahoma? *J. Hydrometeor.*, **16**, 874–888, <https://doi.org/10.1175/JHM-D-14-0005.1>.

—, Q. Wang, and S. M. Quiring, 2016: The observation record length necessary to generate robust soil moisture percentiles. *J. Appl. Meteor. Climatol.*, **55**, 2131–2149, <https://doi.org/10.1175/JAMC-D-16-0143.1>.

Fotheringham, A. S., C. Brunsdon, and M. Charlton, 2003: *Geographically Weighted Regression: The Analysis of Spatially Varying Relationships*. John Wiley & Sons, 282 pp.

Gotway, C. A., R. B. Ferguson, G. W. Hergert, and T. A. Peterson, 1996: Comparison of kriging and inverse-distance methods for mapping soil parameters. *Soil. Sci. Soc. Amer. J.*, **60**, 1237–1247, <https://doi.org/10.2136/sssaj1996.03615995006000040040x>.

Gruber, A., C.-H. Su, S. Zwieback, W. Crow, W. Dorigo, and W. Wagner, 2016a: Recent advances in (soil moisture) triple collocation analysis. *Int. J. Appl. Earth Obs. Geoinf.*, **45**, 200–211, <https://doi.org/10.1016/j.jag.2015.09.002>.

—, —, W. T. Crow, S. Zwieback, W. Dorigo, and W. Wagner, 2016b: Estimating error cross-correlations in soil moisture data sets using extended collocation analysis. *J. Geophys. Res.*, **121**, 1208–1219, <https://doi.org/10.1002/2015JD024027>.

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

—, T. Scanlon, R. van der Schalie, W. Wagner, and W. Dorigo, 2019: Evolution of the ESA CCI soil moisture climate data records and their underlying merging methodology. *Earth Syst. Sci. Data*, **11**, 717–739, <https://doi.org/10.5194/essd-11-717-2019>.

Heggen, R. J., 2001: Normalized antecedent precipitation index. *J. Hydrol. Eng.*, **6**, 377–381, [https://doi.org/10.1061/\(ASCE\)1084-0699\(2001\)6:5\(377\)](https://doi.org/10.1061/(ASCE)1084-0699(2001)6:5(377)).

Hengl, T., G. B. Heuvelink, and A. Stein, 2004: A generic framework for spatial prediction of soil variables based on regression-kriging. *Geoderma*, **120**, 75–93, <https://doi.org/10.1016/j.geoderma.2003.08.018>.

—, —, and D. G. Rossiter, 2007: About regression-kriging: From equations to case studies. *Comput. Geosci.*, **33**, 1301–1315, <https://doi.org/10.1016/j.cageo.2007.05.001>.

Huang, J., H. M. van den Dool, and K. P. Georgarakos, 1996: Analysis of model-calculated soil moisture over the United States (1931–1993) and applications to long-range temperature forecasts. *J. Climate*, **9**, 1350–1362, [https://doi.org/10.1175/1520-0442\(1996\)009<1350:AMCSM>2.0.CO;2](https://doi.org/10.1175/1520-0442(1996)009<1350:AMCSM>2.0.CO;2).

Imaoka, K., M. Kachi, M. Kasahara, N. Ito, K. Nakagawa, and T. Oki, 2010: Instrument performance and calibration of AMSR-E and AMSR2. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.*, **38**, 13–18, http://www.isprs.org/proceedings/XXXVIII/part8/pdf/JTS13_20100322190615.pdf.

Jacobs, J. M., E. C. Hsu, and M. Choi, 2010: Time stability and variability of electronically scanned thinned array radiometer soil moisture during Southern Great Plains hydrology experiments. *Hydrol. Processes*, **24**, 2807–2819, <https://doi.org/10.1002/hyp.7703>.

Keesstra, S., and Coauthors, 2016: Effects of soil management techniques on soil water erosion in apricot orchards. *Sci. Total Environ.*, **551–552**, 357–366, <https://doi.org/10.1016/j.scitotenv.2016.01.182>.

Kerr, Y. H., P. Waldteufel, J.-P. Wigneron, J. Martinuzzi, J. Font, and M. Berger, 2001: Soil moisture retrieval from space: The Soil Moisture and Ocean Salinity (SMOS) mission. *IEEE Trans. Geosci. Remote Sens.*, **39**, 1729–1735, <https://doi.org/10.1109/36.942551>.

Keskin, H., and S. Grunwald, 2018: 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>.

Kohler, M. A., and R. K. Linsley, 1951: *Predicting the Runoff from Storm Rainfall*. U.S. Weather Bureau, Research Paper 34, 10 pp.

Koster, R. D., and Coauthors, 2004: Regions of strong coupling between soil moisture and precipitation. *Science*, **305**, 1138–1140, <https://doi.org/10.1126/science.1100217>.

Kottek, M., J. Grieser, C. Beck, B. Rudolf, and F. Rubel, 2006: World map of the Köppen-Geiger climate classification updated. *Meteor. Z.*, **15**, 259–263, <https://doi.org/10.1127/0941-2948/2006/0130>.

Kumar, S., R. Lal, and D. Liu, 2012: A geographically weighted regression kriging approach for mapping soil organic carbon stock. *Geoderma*, **189–190**, 627–634, <https://doi.org/10.1016/j.geoderma.2012.05.022>.

Li, J., and A. D. Heap, 2011: A review of comparative studies of spatial interpolation methods in environmental sciences: Performance and impact factors. *Ecol. Inform.*, **6**, 228–241, <https://doi.org/10.1016/j.ecoinf.2010.12.003>.

Liang, X., E. F. Wood, and D. P. Lettenmaier, 1996: Surface soil moisture parameterization of the VIC-2L model: Evaluation and modification. *Global Planet. Change*, **13**, 195–206, [https://doi.org/10.1016/0921-8181\(95\)00046-1](https://doi.org/10.1016/0921-8181(95)00046-1).

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

Liu, Q., and Coauthors, 2018: The contributions of precipitation and soil moisture observations to the skill of soil moisture estimates in a land data assimilation system. *J. Hydrometeorol.*, **12**, 750–765, <https://doi.org/10.1175/JHM-D-10-05000.1>.

Martínez-Cob, A., 1996: Multivariate geostatistical analysis of evapotranspiration and precipitation in mountainous terrain. *J. Hydrol.*, **174**, 19–35, [https://doi.org/10.1016/0022-1694\(95\)02755-6](https://doi.org/10.1016/0022-1694(95)02755-6).

McColl, K. A., J. Vogelzang, A. G. Konings, D. Entekhabi, M. Piles, and A. Stoffelen, 2014: Extended triple collocation: Estimating errors and correlation coefficients with respect to an unknown target. *Geophys. Res. Lett.*, **41**, 6229–6236, <https://doi.org/10.1002/2014GL061322>.

Meng, L., and S. M. Quiring, 2008: A comparison of soil moisture models using soil climate analysis network observations. *J. Hydrometeorol.*, **9**, 641–659, <https://doi.org/10.1175/2008JHM916.1>.

Miralles, D. G., W. T. Crow, and M. H. Cosh, 2010: Estimating spatial sampling errors in coarse-scale soil moisture estimates derived from point-scale observations. *J. Hydrometeorol.*, **11**, 1423–1429, <https://doi.org/10.1175/2010JHM1285.1>.

Mishra, U., R. Lal, D. Liu, and M. Van Meirvenne, 2010: Predicting the spatial variation of the soil organic carbon pool at a regional scale. *Soil. Sci. Soc. Amer. J.*, **74**, 906–914, <https://doi.org/10.2136/sssaj2009.0158>.

Mitchell, K. E., and Coauthors, 2004: The multi-institution North American Land Data Assimilation System (NLDAS): Utilizing multiple GCIP products and partners in a continental distributed hydrological modeling system. *J. Geophys. Res.*, **109**, D07S90, <https://doi.org/10.1029/2003JD003823>.

Ochsner, T. E., E. Linde, M. Haffner, and J. Dong, 2019: Mesoscale soil moisture patterns revealed using a sparse in situ network and regression kriging. *Water Resour. Res.*, **55**, 4785–4800, <https://doi.org/10.1029/2018WR024535>.

Odeha, I., A. McBratney, and D. Chittleborough, 1994: Spatial prediction of soil properties from landform attributes derived from a digital elevation model. *Geoderma*, **63**, 197–214, [https://doi.org/10.1016/0016-7061\(94\)90063-9](https://doi.org/10.1016/0016-7061(94)90063-9).

Paloscia, S., S. Pettinato, E. Santi, C. Notarnicola, L. Pasolli, and A. Reppucci, 2013: Soil moisture mapping using Sentinel-1 images: Algorithm and preliminary validation. *Remote Sens. Environ.*, **134**, 234–248, <https://doi.org/10.1016/j.rse.2013.02.027>.

Parinussa, R. M., A. G. Meesters, Y. Y. Liu, W. Dorigo, W. Wagner, and R. A. de Jeu, 2011: Error estimates for near-real-time satellite soil moisture as derived from the land parameter retrieval model. *IEEE Geosci. Remote Sens. Lett.*, **8**, 779–783, <https://doi.org/10.1109/LGRS.2011.2114872>.

Penna, D., L. Brocca, M. Borga, and G. Dalla Fontana, 2013: Soil moisture temporal stability at different depths on two alpine hillslopes during wet and dry periods. *J. Hydrol.*, **477**, 55–71, <https://doi.org/10.1016/j.jhydrol.2012.10.052>.

Pittelkow, C. M., and Coauthors, 2015: Productivity limits and potentials of the principles of conservation agriculture. *Nature*, **517**, 365–368, <https://doi.org/10.1038/nature13809>.

Quiring, S. M., and Coauthors, 2016: The North American Soil Moisture Database: Development and applications. *Bull. Amer. Meteor. Soc.*, **97**, 1441–1459, <https://doi.org/10.1175/BAMS-D-13-00263.1>.

Reichle, R., G. De Lannoy, R. D. Koster, W. T. Crow, J. S. Kimball, and Q. Liu, 2020: SMAP L4 Global 3-hourly 9 km EASE-Grid Surface and Root Zone Soil Moisture Geophysical Data, version 5. SPL4SMGP. NASA National Snow and Ice Data Center Distributed Active Archive Center, accessed 14 June 2021, <https://doi.org/10.5067/9LNYIYOBNBR5>.

Rodell, M., and Coauthors, 2004: The Global Land Data Assimilation System. *Bull. Amer. Meteor. Soc.*, **85**, 381–394, <https://doi.org/10.1175/BAMS-85-3-381>.

Schläpfer, F., and B. Schmid, 1999: Ecosystem effects of biodiversity: A classification of hypotheses and exploration of empirical results. *Ecol. Appl.*, **9**, 893–912, [https://doi.org/10.1890/1051-0761\(1999\)009\[0893:EOBAC\]2.0.CO;2](https://doi.org/10.1890/1051-0761(1999)009[0893:EOBAC]2.0.CO;2).

Scipal, K., T. Holmes, R. De Jeu, V. Naeimi, and W. Wagner, 2008: A possible solution for the problem of estimating the error structure of global soil moisture data sets. *Geophys. Res. Lett.*, **35**, L24403, <https://doi.org/10.1029/2008GL035599>.

Seneviratne, S. I., and Coauthors, 2010: Investigating soil moisture–climate interactions in a changing climate: A review. *Earth Sci. Rev.*, **99**, 125–161, <https://doi.org/10.1016/j.earscirev.2010.02.004>.

Spennemann, P. C., J. A. Rivera, A. C. Saulo, and O. C. Penalba, 2015: A Comparison of GLDAS soil moisture anomalies against standardized precipitation index and multisatellite estimations over South America. *J. Hydrometeorol.*, **16**, 158–171, <https://doi.org/10.1175/JHM-D-13-0190.1>.

Stoffelen, A., 1998: Toward the true near-surface wind speed: Error modeling and calibration using triple collocation. *J. Geophys. Res. Oceans*, **103**, 7755–7766, <https://doi.org/10.1029/97JC03180>.

Tian, L., and S. M. Quiring, 2019: Spatial and temporal patterns of drought in Oklahoma (1901–2014). *Int. J. Climatol.*, **39**, 3365–3378, <https://doi.org/10.1002/joc.6026>.

Vachaud, G., A. Passerat de Silans, P. Balabanis, and M. Vauclin, 1985: Temporal stability of spatially measured soil water probability density function 1. *Soil. Sci. Soc. Amer. J.*, **49**, 822–828, <https://doi.org/10.2136/sssaj1985.03615995004900040006x>.

Wagner, W., G. Lemoine, M. Borgeaud, and H. Rott, 1999: A study of vegetation cover effects on ERS scatterometer data. *IEEE Trans. Geosci. Remote Sens.*, **37**, 938–948, <https://doi.org/10.1109/36.752212>.

—, and Coauthors, 2013: The ASCAT soil moisture product: A review of its specifications, validation results, and emerging applications. *Meteor. Z.*, **22**, 5–33, <https://doi.org/10.1127/0941-2948/2013/0399>.

Wanders, N., D. Karssenberg, A. Roo, S. De Jong, and M. Bierkens, 2014: The suitability of remotely sensed soil moisture for improving operational flood forecasting. *Hydrol. Earth Syst. Sci.*, **18**, 2343–2357, <https://doi.org/10.5194/hess-18-2343-2014>.

Wang, A., D. P. Lettenmaier, and J. Sheffield, 2011: Soil moisture drought in China, 1950–2006. *J. Climate*, **24**, 3257–3271, <https://doi.org/10.1175/2011JCLI3733.1>.

Xia, Y., and Coauthors, 2012: 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. *J. Geophys. Res.*, **117**, 3109, <https://doi.org/10.1029/2011JD016048>.

—, M. B. Ek, Y. Wu, T. Ford, and S. M. Quiring, 2015: Comparison of NLDAS-2 simulated and NASMD observed daily soil moisture. Part I: Comparison and analysis. *J. Hydrometeor.*, **16**, 1962–1980, <https://doi.org/10.1175/JHM-D-14-0096.1>.

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

Yee, M. S., J. P. Walker, A. Monerris, C. Rüdiger, and T. J. Jackson, 2016: On the identification of representative in situ soil moisture monitoring stations for the validation of SMAP soil moisture products in Australia. *J. Hydrol.*, **537**, 367–381, <https://doi.org/10.1016/j.jhydrol.2016.03.060>.

Yilmaz, M. T., and W. T. Crow, 2013: The optimality of potential rescaling approaches in land data assimilation. *J. Hydrometeor.*, **14**, 650–660, <https://doi.org/10.1175/JHM-D-12-052.1>.

—, —, M. Anderson, and C. Hain, 2012: An objective methodology for merging satellite- and model-based soil moisture products. *Water Resour. Res.*, **48**, W11502, <https://doi.org/10.1029/2011WR011682>.

Yuan, S., and S. M. Quiring, 2017: Comparison of three methods of interpolating soil moisture in Oklahoma. *Int. J. Climatol.*, **37**, 987–997, <https://doi.org/10.1002/joc.4754>.

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

Zhang, N., S. Quiring, T. Ochsner, and T. Ford, 2017a: Comparison of three methods for vertical extrapolation of soil moisture in Oklahoma. *Vadose Zone J.*, **16** (10), 1–19, <https://doi.org/10.2136/vzj2017.04.0085>.

—, C. Zhao, S. M. Quiring, and J. Li, 2017b: Winter wheat yield prediction using normalized difference vegetative index and agro-climatic parameters in Oklahoma. *Agron. J.*, **109**, 2700–2713, <https://doi.org/10.2134/agronj2017.03.0133>.

Zhao, C., S. M. Quiring, S. Yuan, D. B. McRoberts, N. Zhang, and Z. Leasor, 2020: Developing and evaluating national soil moisture percentile maps. *Soil. Sci. Soc. Amer. J.*, **84**, 443–460, <https://doi.org/10.1002/saj2.20045>.

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