

# Water Resources Research

## RESEARCH ARTICLE

10.1029/2018WR024039

### Key Points:

- A framework is applied to validate in situ, model, and satellite soil moisture data for drought monitoring
- SCAN in situ stations exhibit high relative error variance, compared with other in situ networks
- NLDAS-2 models and SMAP L3 remote sensing products outperform other products tested

### Supporting Information:

- Supporting Information S1

### Correspondence to:

T. W. Ford,  
twford@siu.edu

### Citation:

Ford, T. W., & Quiring, S. M. (2019). Comparison of contemporary in situ, model, and satellite remote sensing soil moisture with a focus on drought monitoring. *Water Resources Research*, 55, 1565–1582. <https://doi.org/10.1029/2018WR024039>

Received 10 SEP 2018

Accepted 30 JAN 2019

Accepted article online 5 FEB 2019

Published online 25 FEB 2019

## Comparison of Contemporary In Situ, Model, and Satellite Remote Sensing Soil Moisture With a Focus on Drought Monitoring

Trent W. Ford<sup>1</sup>  and Steven M. Quiring<sup>2</sup> 

<sup>1</sup>Department of Geography and Environmental Resources, Southern Illinois University at Carbondale, Carbondale, IL, USA, <sup>2</sup>Atmospheric Sciences Program, Department of Geography, The Ohio State University, Columbus, OH, USA

**Abstract** Soil moisture is a key drought indicator; however, current in situ soil moisture infrastructure is inadequate for large-scale drought monitoring. One initiative of the ongoing National Soil Moisture Network program is the development of a near real-time drought monitoring product that integrates in situ, model, and satellite remote sensing data. Data integration from diverse sources requires large-scale validation prior to integration. This study develops a framework for assessing the fidelity of in situ, model, and satellite soil moisture data sets. Here we evaluate data from over 100 in situ monitoring stations that are part of nine monitoring networks; North American Land Data Assimilation System Phase 2 and Climate Prediction Center land surface models; and Soil Moisture Active-Passive, Soil Moisture and Ocean Salinity, and European Space Agency-Climate Change Initiative (ESA-CCI) satellite products. The results indicate the majority of in situ stations exhibit low error variance and are spatially representative; however, some networks and individual stations exhibit anomalously high error variance or are sited in a way that make them not spatially representative of a larger area. Overall, North American Land Data Assimilation System Phase 2 is the modeled product that consistently performed best, and Soil Moisture Active-Passive L3 is the remotely sensed product that consistently performed the best. They were able to both capture in situ soil moisture variability and provide an accurate depiction of drought conditions. The methods and verification framework applied in this study can be used to evaluate any soil moisture data set in any region of the world.

**Plain Language Summary** Soil moisture is an important indicator of drought; however, there are very few monitoring stations that directly measure soil moisture across the contiguous United States. High-quality model and/or satellite remote sensing soil moisture estimates can help fill in the gaps of in-ground soil moisture measurements. However, integration of soil moisture data from diverse sources requires assessment of data quality. This study applies a series of methods for evaluating the integrity of in-ground, model, and satellite soil moisture data sets across the United States. The results indicate that the majority of in-ground sensors tested exhibit high data quality, although some exhibit high measurement error and/or low spatial representativeness. The results also indicate that the land surface models that are part of the North American Land Data Assimilation System Phase 2 and the Soil Moisture Active-Passive remotely sensed soil moisture products best match in-ground soil moisture measurement variability. Additionally, North American Land Data Assimilation System Phase 2 and Soil Moisture Active-Passive data sets accurately depict drought occurrence. The methods applied here can be used to evaluate the quality of soil moisture data sets in any region of the world.

## 1. Introduction

Drought is one of the most destructive and costly natural disasters, resulting in diminished agricultural production (Hatfield et al., 2011), reduced water resources (van Dijk et al., 2013), and deadly heat waves (Hoerling et al., 2013; Schubert et al., 2014). In response, drought strategies are being developed worldwide, with a particular focus on comprehensive drought monitoring (Hayes et al., 2011). As part of drought monitoring strategies, soil moisture is often used as an indicator of moisture stress and agricultural drought (Quiring & Papkyriakou, 2003). This is because root zone soil moisture in vegetated regions has a significant influence on evapotranspiration rates (Miralles et al., 2014). Soil moisture persistence or “memory” can affect atmospheric conditions on subseasonal to seasonal timescales (Liu et al., 2014; Whan et al., 2015), thereby providing a critical source of information for drought monitoring and prediction on such timescales.

Because of its importance for drought monitoring and prediction, numerous operational and experimental soil moisture data sets are available for drought applications. The majority of these products are based on model-simulated soil moisture. For example, the University of Washington Experimental Surface Water Monitor (Wood, 2008), the NOAA Climate Prediction Center (Fan & van den Dool, 2004), the North American Land Data Assimilation System Phase 2 (NLDAS-2; Xia et al., 2012), and the Princeton Drought Monitoring and Forecasting project (Sheffield et al., 2014) all provide model-simulated soil moisture to support drought monitoring applications. The U.S. Drought Monitor (USDM; Svoboda et al., 2002) uses some of these soil moisture data sets to help generate weekly, nationwide drought maps. However, there are limitations to using model-derived soil moisture since each land-surface model has biases, and model performance varies significantly from region to region and model to model (Kumar et al., 2017; Xia et al., 2015a). Therefore, one approach is to integrate in situ, satellite-, and model-derived soil moisture into a comprehensive drought monitoring product that leverages the advantages of all three data sources. During the National Soil Moisture Network (NSMN) workshop in June 2016, near real time, national soil moisture data sets that integrate in situ, satellite-, and model-derived soil moisture were identified as the highest priority by the workshop participants (McNutt et al., 2016). There was a broad consensus among the 52 workshop participants who represented a variety of federal and state agencies, universities, and the private sector that a high-resolution gridded soil moisture product that leverages multiple in situ networks, satellite platforms, and land surface models is needed. Their sentiment reflects how critical in situ soil moisture observations are for national drought monitoring and prediction, as well as for calibrating and validating the suite of model-based soil moisture-drought products. This need has existed for years, and it has been repeatedly identified as a high priority, but, to date, it has not been addressed in a comprehensive way.

Previous studies have demonstrated that the inclusion of in situ soil moisture information improves drought monitoring and drought prediction systems (Ford et al., 2015). However, these studies are limited in scope, either spatially or temporally, because there has been relatively little effort devoted to assembling and homogenizing in situ soil moisture measurements for national-scale drought monitoring. As a result of the ongoing NSMN project and other previous initiatives, the foundation has been laid for a high-resolution, near real time gridded soil moisture product that exploits multiple in situ networks, satellite platforms, and land surface models. An important precursor to the development of this type of product is a comprehensive, national-scale assessment of in situ, satellite, and model soil moisture data fidelity. This paper employs a variety of soil moisture comparison/verification methods to develop a comprehensive soil moisture validation framework, with a focus on drought monitoring. The goal is to identify the in situ, satellite-, and model-derived soil moisture data sets that are best suited for developing a national soil moisture product.

## 2. Data

### 2.1. In Situ Soil Moisture

Daily in situ soil moisture observations from 100 stations that are part of nine monitoring networks (Table 1; Figure 1) were downloaded from the North American Soil Moisture Database ([www.nationalsoilmoisture.com](http://www.nationalsoilmoisture.com)). These networks included the Delaware Environmental Observing System (DEOS; <http://www.deos.udel.edu/data/>), Enviro-weather (formerly Michigan Automated Weather Network; <https://mawn.geo.msu.edu/>), the NOAA Hydrometeorological Testbed (NOAA; <https://hmt.noaa.gov>), the Oklahoma Mesonet (<https://www.mesonet.org/>), the USDA-Natural Resources Conservation Service Soil Climate Analysis Network (SCAN; <https://www.wcc.nrcd.usda.gov/scan/>), the USDA-Natural Resources Conservation Service SNOwpack TELemetry (SNOTel; <https://www.wcc.nrcs.usda.gov/snow/>), the Soil Moisture Sensing Controller and Optimal Estimator (SoilScape; <https://www.soilscape.usc.edu/>), and the West Texas Mesonet (WTX Mesonet, <http://www.depts.ttu.edu/nwi/research/facilities/wtm/index.php>).

DEOS includes a series of stations throughout Delaware, southeast Pennsylvania, and northwest Maryland (Legates et al., 2005). A variety of meteorological and agricultural variables are observed at DEOS stations, including soil moisture at 5-cm depth. DEOS employs Campbell 616L, time domain reflectometer, sensors to estimate  $\theta$  at 5-min intervals. DEOS provides daily average  $\theta$  at most stations within the network. We obtain data from 26 DEOS stations that continually report daily average  $\theta$  over the time period 2014 to 2017 (Figure 1). The 26 stations span an area between 38.5°N and 40.2°N latitude by 76°W and 75.2°W longitude, a density of one station per 492 km<sup>2</sup>.

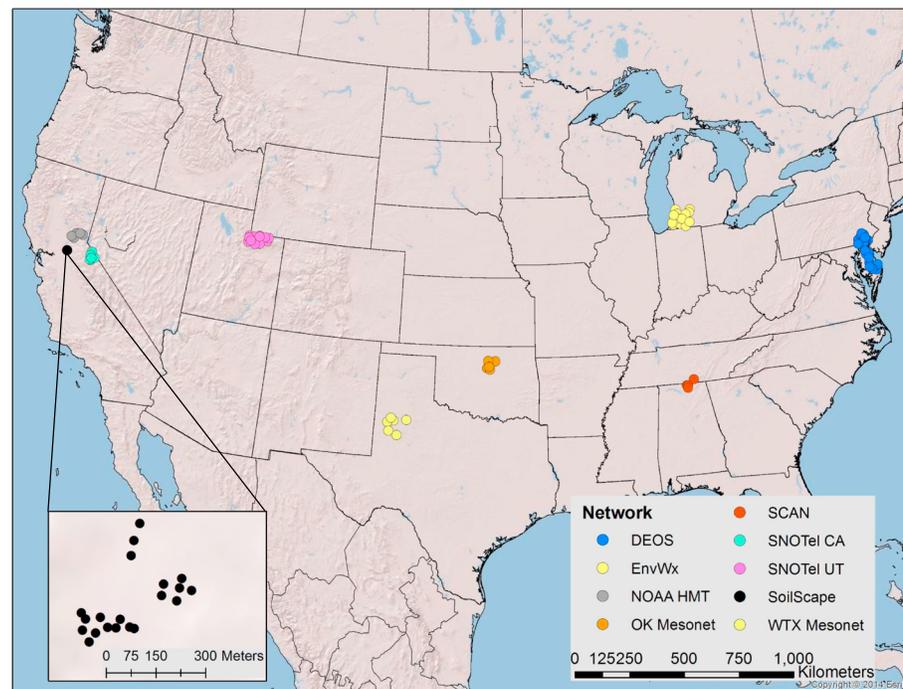
**Table 1**  
*Details of In Situ, Model, and Satellite Soil Moisture Data Used in the Study*

In situ network	Location	# of stations	Years	Sensor	Sensor depths (cm)	Reference
Delaware Environmental Observation System (DEOS)	Delaware, Pennsylvania	26	4	CS 616	5	Legates et al. (2005)
Enviro-weather (EnvWx)	Michigan	16	10	CS 616	15, 45	Andresen et al. (2011)
NOAA Hydrometeorological Testbed (NOAA HMT)	California	3	10	CS 616	10, 15	Zamora et al. (2011)
Oklahoma Mesonet (OK Mesonet)	Oklahoma	5	19	CS 229 L	5, 25, 60	McPherson et al. (2007)
Soil Climate Analysis Network (SCAN)	Alabama	6	14	Stevens Hydraprobe	5, 10, 20, 50, 100	Schaefer et al. (2007)
SNOWpack TELEmetry (SNOTel)	California, Utah	21	13	Stevens Hydraprobe	5, 20, 50	Schaefer and Paetzold (2001)
Soil Moisture Sensing Controller and Optimal Estimator (SoilScape)	California	17	5	Decagon EC-5	5, 20, 50	Moghaddan et al. (2010)
West Texas Mesonet (WTX Mesonet)	Texas	6	16	CS 616	5, 20, 60, 75	Schroeder et al. (2005)
Model/satellite	VERSION	Horizontal resolution	Years		Depths (cm)	Reference(s)
CPC one-layer (model)		1°	10		0–160	Huang et al. (1996), van den Dool et al. (2003)
NLDAS-2 Noah (model)	2.8	0.125°	19		0–10, 10–40, 40–100	Chen et al. (1996), Xia et al. (2012)
NLDAS-2 Mosaic (model)		0.125°	19		0–10, 10–40, 40–100	Koster and Suarez (1992)
NLDAS-2 VIC (model)	4.0.3	0.125°	19		0–10, 10–40, 40–100	Liang et al. (1994)
SMAP L4 (satellite/model)	SML4SMGP	9 km	3		0–5, 0–100	Reichle et al. (2018)
SMAP L3 (satellite)	V5	36 km	3		~ 0–5	O'Neill et al. (2018)
SMOS L3 (satellite)	CATDS-PDC 1 day global	40 km	8		~ 0–5	Al Bitar et al. (2017)
ESA Soil Moisture CCI (ECV)	Combined 04.2	0.25°	19		~ 0–5	Liu et al. (2012), Dorigo et al. (2017), Gruber et al. (2017)

Enviro-weather (EnvWx) is a network of meteorological and agricultural environmental monitoring networks across Michigan and eastern Wisconsin (Andresen et al., 2011). Campbell 616L sensors are used to estimate  $\theta$  at hourly intervals at most EnvWx stations at approximately 15- and 45-cm depths. We obtained hourly  $\theta$  from 16 EnvWx stations that continually report data over the time period 2008 to 2017 (Figure 1). We averaged the hourly  $\theta$  estimates to daily resolution to match the other networks, models, and satellites. EnvWx stations were selected from the southwest corner of the network because station density is highest in that region. The 16 stations span an area between 41.8°N and 42.6°N latitude and 86.4°W and 85.3°W longitude, a density of one station per 494 km<sup>2</sup>.

NOAA HMT is a network of meteorological and hydrological monitoring stations that observe multiple variables in numerous river basins and locations across the United States (Zamora et al., 2011). We obtained hourly  $\theta$  from three NOAA HMT stations within the American River Basin in eastern California (Figure 1). CS616 sensors estimate  $\theta$  at 10- and 15-cm depths at each of these stations, and each station continually reported data over the time period 2008 to 2017. We averaged hourly  $\theta$  estimates at each of the three NOAA HMT stations to daily resolution. The stations span an area between 39°N and 39.3°N latitude by 120.9°W and 120.4°W longitude, a density of one station per 237 km<sup>2</sup>.

The OK Mesonet is a network of meteorological and agricultural monitoring stations across the state of Oklahoma. CS229L heat dissipation sensors, sited at 5-, 25-, and 60-cm depths, are used to measure soil matric potential—from which  $\theta$  is estimated—at 15-min intervals at nearly all OK Mesonet stations (McPherson et al., 2007). The focus of this analysis was to develop and test in situ, model, and satellite



**Figure 1.** In situ station map. Inset shows distribution of high density SoilScape stations.

soil moisture data validation across a wide variety of climates, soil types, sensors, and networks. Therefore, we obtained daily averaged  $\theta$  from five OK Mesonet stations that are relatively densely sited in the north-central portion of Oklahoma (Figure 1). The five stations continuously report data over the time period 1999 to 2017 and span an area between 36°N and 36.4°N latitude by 97.2°W and 96.8°W longitude, a density of one station per 320 km<sup>2</sup>.

SCAN contains over 200 stations in all 50 states continuously monitoring soil moisture, as well as a variety of meteorological and hydrological variables (Schaefer et al., 2007). SCAN uses Stevens Hydraprobe dielectric reflectometers, sited at 5-, 10-, 20-, 50-, and 100-cm depths, to estimate  $\theta$  at hourly intervals. We obtained daily averaged  $\theta$  from six SCAN stations located in northern Alabama and southern Tennessee (Figure 1). These SCAN stations are densely sited—relative to the rest of the network—and have continuously reported soil moisture over the time period 2004 to 2017. The six stations span an area between 34.8°N and 35.1°N latitude by 86.6°W and 86.9°W longitude, a density of one station per 176 km<sup>2</sup>.

The SNOTel network is comprised of over 300 stations that monitor various meteorological and hydrological conditions across the western United States (Schaefer & Paetzold, 2001). Like SCAN, Stevens Hydraprobes are used to estimate hourly  $\theta$  at 5, 20, and 50 cm at SNOTel stations. We obtained daily averaged  $\theta$  from 6 SNOTel stations in eastern California, as well as 15 SNOTel stations in northeastern Utah (Figure 1). Validation and comparison analyses are carried out separately for California SNOTel stations (SNOTel CA) and Utah SNOTel stations (SNOTel UT). All 20 SNOTel stations have continuously reported data over the time period 2005 to 2017. The California SNOTel stations span an area between 38.3°N and 38.7°N latitude by 119.6°W and 119.4°W longitude (153 km<sup>2</sup> per station), while the Utah stations span an area between 40.5°N and 41°N latitude by 111.1°W and 109.9°W longitude (308 km<sup>2</sup> per station).

SoilScape is a high-density network of soil moisture nodes. Multiple soil moisture sensors are clustered in groups within areas <1 km<sup>2</sup> at several locations across the western United States (Moghaddan et al., 2010). The network uses Decagon EC-5 capacitance probes to estimate 20-min  $\theta$  at 5-, 20-, and 50-cm depths. We obtained 20-min  $\theta$  from the series of stations within the Tonzi, California, node in northern California, 17 stations total (Figure 1), all reporting data over the time period 2013 to 2017. We averaged the 20-min  $\theta$  data to daily resolution. All 17 of these stations are located within a 1-km<sup>2</sup> area, representing the densest network analyzed in this study.

The WTX Mesonet consists of over 100 stations across West Texas and New Mexico that monitor meteorological and soil conditions (Schroeder et al., 2005). WTX Mesonet uses CS616 sensors to monitor hourly  $\theta$  at 5-, 20-, 60-, and 75-cm depths. We obtained daily averaged  $\theta$  from six stations in West Texas (Figure 1), all continuously reporting data over the time period 2002 to 2017. The six stations span an area between 32.9°N and 33.7°N latitude by 102.4°W and 101.4°W longitude, a density of 1272 km<sup>2</sup> per station.

In situ stations within each network were carefully selected based on (1) covering a diversity of climate regions, (2) record length, and (3) measurement continuity. All stations were screened using a quality control process (Quiring et al., 2016), and stations and/or networks that do not consistently report soil moisture or those that have not reported data since the beginning of 2018 were not considered. Although the record length varied by network (Table 1), all stations within each network reported data over the entire time period, eliminating the potential confounding effects of stations coming and going. All in situ data were acquired in units of volumetric water content ( $\theta$ , m<sup>3</sup>/m<sup>3</sup>). A general overview of each in situ network is included in Table 1.

## 2.2. Modeled Soil Moisture

Modeled soil moisture has increasingly become an invaluable asset for historical climate analysis and climate monitoring and forecasting (Koster et al., 2010; Xia et al., 2014), in no small part due to the multitude of high-quality, modeled soil moisture products available. Many of these products, however, are produced with too long of a latency to be effectively used for the near real time (i.e., daily to subweekly) soil moisture and drought monitoring that is the primary application of data validation/comparison efforts in this paper. Therefore models assessed in this study are selected foremost based on data availability/latency and data record length. Three of the land surface models that are used in this study—Noah, Mosaic, and VIC—are part of the NLDAS-2. Each model simulates hourly soil moisture at multiple depths. Here we use 0- to 10-cm, 10- to 40-cm, and 40- to 100-cm layers. The spatial resolution of NLDAS-2 is 1/8° (Table 1), and NLDAS-2 model output is available from 1979 to the present. All 39 years of NLDAS-2 soil moisture were used to convert the  $\theta$  into soil moisture percentiles (see section 3.2). Hourly  $\theta$  fields from each model were averaged to daily resolution to match the temporal resolution of the in situ and satellite data sets.

The NLDAS-2 models were selected for evaluation in this study because (1) all three models have been previously shown to have high fidelity (Xia et al., 2015a, 2015b), (2) model-simulated soil moisture data are available since 1979 and this covers the entire period of record for the in situ and satellite data, and (3) NLDAS-2 soil moisture have a relatively short latency (~4 days; Ek et al., 2017). Data latency is a crucial consideration for this study and future work because real-time soil moisture is required for drought identification. Additionally, NLDAS-2 soil moisture data are available as  $\theta$  or total column water (mm), which allows us to generate soil moisture anomalies or percentiles using a variety of methods. Also pertinent for our model selection is the upcoming NLDAS-2.5 product, which will become operational in early 2019 and will have a 0-day latency.

In addition to the NLDAS-2, we will also evaluate daily model-simulated soil moisture from the NOAA Climate Prediction Center (CPC)'s one-layer “leaky bucket” hydrological model (van den Dool et al., 2003). The CPC model computes total (0–160 cm) column moisture at a 1° resolution across the contiguous United States. Soil moisture data are available from 2008 to the present. The CPC model was selected for this study because it currently informs both CPC climate forecasts and the USDM and data are available with an approximately 1-day latency.

The subset of models evaluated in this study obviously represents only a small fraction of available modeled soil moisture products. Because it is not feasible to evaluate all existing models, we selected NLDAS-2 and CPC models because they are commonly used for near real time climate monitoring and forecasting and are currently used as input to the USDM. Therefore, large-scale assessment of these systems provides important information for the providers and users of modeled soil moisture.

## 2.3. Satellite Soil Moisture

Satellite-based microwave remote sensing soil moisture fields were obtained from the Soil Moisture Active-Passive (SMAP L3; Entekhabi et al., 2010), the Soil Moisture and Ocean Salinity (SMOS L3; Kerr et al., 2010), and the European Space Agency Program on Global Monitoring of Essential Climate Variables (ECVs; Dorigo et al., 2017; Gruber et al., 2017; Liu et al., 2012) soil moisture data sets (Table 1). SMAP and SMOS

were selected because they have been previously shown to have high fidelity (Colliander et al., 2017; Jackson et al., 2012). The SMAP L3 products, with a horizontal resolution of 36 km, have an approximately 50-hr latency, while the SMOS L3 products, with a horizontal resolution of 40 km, have a 7-day latency. The primary limitation of SMAP and SMOS is their relatively short data records (~3 and ~9 years, respectively). Therefore, we will also evaluate the ECV data set in this study. ECV is a merged active-passive product that has a relatively long data record (1992–2016) and a 0.25° horizontal resolution. The primary limitation of the ECV product is its data latency, as of July 2018, the 2017 data were still not available. However, the ECV data set will soon be produced with a 10 day latency via the Copernicus Climate Change service (<https://climate.copernicus.eu/>). SMAP L3, SMOS L3, and ECV all provide daily soil moisture fields for analysis.

This study will also evaluate the SMAP L4 surface and root zone soil moisture products (Reichle et al., 2017). These data sets are produced by assimilating SMAP L-band brightness temperature observations into the NASA Catchment land surface model. The resulting 3-hr soil moisture data are available at 0- to 5-cm (SMAP L4 0–5) and 0- to 100-cm (SMAP L4 0–100) layers at 9-km spatial resolution. The 3-hr  $\theta$  fields in both SMAP L4 data sets were averaged to daily resolution. The SMAP L4 products have a higher spatial resolution than SMAP L3 (9 km compared to 36 km). In addition, the data latency of SMAP L4 is only ~3 days. Finally, it is important to note that none of the model or satellite soil moisture products use any of the in situ observations for their product calibration; therefore, the comparison is not confounded by data dependency.

### 3. Methods

#### 3.1. Soil Moisture Data Processing

To compare in situ soil moisture to the NLDAS-2, SMAP L4, and CPC models, we vertically interpolated the in situ  $\theta$  data to the model levels, following the procedure used by Dirmeyer et al. (2016). Sensors are assumed to represent a layer whose top is exactly halfway between it and the next shallowest sensor and whose bottom is exactly between it and the next deepest sensor. For example, SNOTel stations have sensors placed at 5, 20, and 50 cm; therefore, the 5-cm sensor represents a “layer” spanning the surface down to 12.5 cm (i.e., halfway between 5- and 20-cm sensors), while the 20-cm sensors represents a “layer” spanning 12.5 to 35 cm, and the 50-cm sensor a “layer” spanning 35 cm to the deepest model layer (100 cm for the Noah model). The final interpolated value for the model layer is the weighted average of all of the observation “layers” that overlap the model layer. In the case of SNOTel stations, the 0- to 10-cm Noah layer is directly compared to the 5-cm soil moisture, as the 5-cm sensor represents a layer entirely spanning the 0- to 10-cm Noah layer. The 10- to 40-cm Noah layer is compared to a weighted combination of all three SNOTel sensors, with sensor weights representing the fraction of the model layer they contain (i.e., 2.5 cm of 30-cm layer = 8% weight for 5-cm sensor). A schematic of the weighting procedure is shown in the supporting information (Figure S1).

The layer-weighted in situ time series were used for comparison with the land-surface models. This weighting procedure was done separately for each model and in situ network pair comparison, as sensor depth is inconsistent between monitoring networks (Table 1). The weighting procedure has been successfully implemented in previous model comparison/validation efforts (see Dirmeyer et al., 2016) and provides a fair comparison as the models evaluated simulate continuous layers instead of discrete points in the soil column. The shallowest in situ sensor was compared to SMAP L3, SMOS L3, and ECV data sets. Each in situ-model/satellite comparison was completed over a fixed period, determined by the shorter of the two records. This way, all in situ stations in each network were included over the entire comparison, eliminating potential confounding effects of stations coming online and off-line. Prior to comparison with model/satellite data, daily in situ  $\theta$  was converted to anomalies ( $\theta_a$  m<sup>3</sup>/m<sup>3</sup>) by subtracting the climatological mean of a 15-day moving window surrounding the calendar day. For example,  $\theta_a$  at a given station on 20 July 2010 is  $\theta$  with the mean  $\theta$  of 13 to 27 July of every year in the data record. Using anomalies instead of volumetric water content accounts for data set differences in absolute soil moisture and seasonal variability that affects soil moisture across most of the United States. Daily  $\theta_a$  at individual stations was then averaged over all stations within each model or satellite grid cell. The average number of in situ stations contained within single grid cells of each satellite or model data set is reported in Table 2.

**Table 2**  
Average Number of In Situ Stations in Each Network Contained Within a Single Grid Cell of Each Satellite or Model Data Set

	CPC	NLDAS-2	SMAP L4	SMAP L3	SMOS L3	ECV
DEOS	4.8	1.04	1	2.67	1.72	1.41
EnvWx	2.75	1	1	1.22	1	1
NOAA HMT	2	1	1	1.33	1.33	1
OK Mesonet	5	1	1	1.25	1.25	1
SNOTel CA	1.67	1	1	1.67	1.67	1.25
SNOTel UT	7.5	1.15	1.07	2.14	1.5	1.25
SoilScape	15	15	7.5	15	15	15
WTX Mesonet	2	1	1	1.2	1	1

The grid cell-averaged in situ  $\theta_a$  were then compared to the model and satellite  $\theta_a$ , computed using the same anomaly calculation method. Averaging across stations within each grid cell reduces, but does not eliminate, the scale mismatch issues associated with comparing point measurements with model/satellite data.

### 3.2. Soil Moisture Percentiles

Because the focus of this study is on utilizing soil moisture data for drought monitoring, we defined drought as a  $\theta$  percentile of  $\leq 20$ . This follows the USDM classification scheme, where soil moisture drier than the 20th percentile is considered “moderate drought.” Daily  $\theta$  percentiles were computed independently for each in situ network, model, and satellite data set. Soil moisture standardization, particularly for monitoring extreme events like drought, must account for the seasonal cycle of soil moisture that is common throughout regions in the midlatitudes. However, computing soil moisture percentiles separately by calendar month or day of the year can result in the same percentile value for dramatically different soil moisture conditions. This introduces false variability where the difference between (i.e.,) the 20th and 80th percentiles of  $\theta$  is not physically meaningful. In addition, computing percentiles requires enough data points to generate a stable distribution, particularly when examining the distribution tails. Our solution to account for all of these confounding factors is to first standardize daily  $\theta$  using the climatological mean of each calendar day. This accounts for the seasonal cycle of soil moisture variability during that time of the year. Then we compute percentiles of those daily soil moisture anomalies, using the entire data record, not separately by calendar month or day. Including the entire daily time series of  $\theta$  anomalies when computing percentiles both reduces the chances of introducing false variability and ensures a sufficient data record length as to generate stable distributions.

In practice, we first averaged daily in situ  $\theta$  values across all stations within a model or satellite grid cell. We then standardized these grid cell-averaged time series by subtracting the climatological mean of each calendar day. Daily percentiles were then computed from the empirical cumulative distribution function of the entire data record for each grid cell-averaged in situ time series.

### 3.3. Evaluation Methods

#### 3.3.1. Methods for In Situ Validation

Most soil moisture validation studies regard in situ measurements as the benchmark. However, the accuracy, representativeness, and overall quality of in situ soil moisture observations across the United States are not equal. Disparities in station siting, sensor types and sensor installation, and site maintenance are just a few of the many potential sources of error in the in situ measurements. Although quality control procedures can help to flag and remove some suspicious measurements, it cannot account for all siting-, installation-, or maintenance-induced errors. Therefore, it is imperative to first assess the fidelity of *in situ* data prior to comparing these data to model or satellite soil moisture.

We use two approaches for evaluating the quality of the in situ soil moisture measurements. The first approach, which is adopted from Dirmeyer et al. (2016), estimates the random in situ observational error. Soil moisture varies as a first-order Markov process, such that the autocorrelation of daily soil moisture at an in situ station decreases as lag  $\tau$  increases:

$$r(\tau) = \exp(-\lambda\tau), \quad (1)$$

where  $1/\lambda$  is the timescale. Robock et al. (1995) showed that for actual observed data, a linear best fit of  $\ln(r)$  versus  $\tau$  for a range of lags does not cross zero at a value of  $r = 1$ , but instead at a correlation  $r < 1$ , and this displacement of the autocorrelation at  $\tau = 0$  is an indicator of measurement error such that

$$r(\tau = 0) = 1 - a. \quad (2)$$

Dirmeyer et al. (2016) show that the displacement of the autocorrelation ( $a$ ) is directly related to the ratio of soil moisture variance contributed by random measurement error ( $\delta^2$ ) to the real variance of the daily soil moisture time series ( $\sigma^2$ ), such that

$$\delta^2 / \sigma^2 = (a / (1 - a)). \quad (3)$$

Using this method, we can estimate the ratio of error variance to real soil moisture variance, with higher ratios representing a larger proportion of measurement error, and therefore lesser data quality.

The second approach for in situ soil moisture verification evaluates the spatial representativeness of the in situ measurements using triple collocation (TC). Specifically, we adopt the formulation of Gruber et al. (2013) to evaluate the ability of stations within each in situ monitoring network to represent soil moisture conditions over an area approximating the size of a model or satellite grid cell (e.g., ~50 km). A typical error model for TC is

$$i = \alpha_i + \beta_i\theta + \epsilon_i, \quad (4)$$

where  $i \in [X, Y, Z]$  are three spatially and temporally collocated soil moisture data sets;  $\theta$  is the unknown, true soil moisture;  $\alpha_i$  and  $\beta_i$  are systematic gains of data set  $i$  with respect to  $\theta$ ; and  $\epsilon_i$  is the additive zero-mean random noise. We select the in situ data as the reference and rescale the SMOS L3 and Noah land surface model data into the in situ data space:

$$\theta_{,s}^*(t) = \theta_R + \sqrt{\frac{\text{var}(\theta_R)}{\text{var}(\theta_s) [\theta_s(t) - \bar{\theta}_s]}}, \quad (5)$$

where  $\theta_s(t)$  is the soil moisture anomaly at time  $t$ ,  $\theta_R$  is the reference data (in situ) soil moisture anomaly, and  $\theta_{,s}^*$  is the rescaled anomaly at time  $t$ . Soil moisture anomalies are computed by subtracting the climatological monthly mean from each daily observation in each calendar month. We use SMOS L3 and Noah to represent satellite- and model-derived soil moisture, respectively, because these two data sets are independent of each other and have at least 8 years of data that can be used from the TC. Averaged random errors ( $e_i$ ) of each data set (in situ, Noah, and SMOS) are calculated using the rescaled measurements from equation (2) such that

$$e_x = \langle (\theta_x - \theta_y)(\theta_x - \theta_z) \rangle, \quad (6)$$

where  $e_x$  is the averaged random error of data set  $x$ , and  $\theta_x$ ,  $\theta_y$ , and  $\theta_z$  are rescaled measurements from data sets  $x$ ,  $y$ , and  $z$ , respectively. Stations that are identified through TC as having less random error, more accurately capture coarse-scale soil moisture variability. Therefore, these stations are more valuable for national-scale drought monitoring than those that only represent highly localized soil moisture conditions.

### 3.3.2. Methods for Model and Satellite Comparison

After identifying which in situ monitoring networks are of sufficient quality for comparison with model/satellite data sets, we use a number of approaches for the comparison. First, we assess differences in daily  $\theta$  variability (i.e., standard deviation,  $\sigma$ ) between the soil moisture data sets, differences in absolute  $\theta$  values (i.e., bias), and shared variance via the coefficient of variation ( $R^2$ ). Additionally, we assess the drought monitoring capability of the various model and satellite products. The questions we are seeking to answer, with regard to model/satellite drought monitoring efficacy, are (1) how well does the model/satellite correctly identify drought and (2) how well does the model/satellite discern between drought and nondrought conditions. We implement accuracy ( $A$ ) and probability of detection ( $POD$ ) forecast verification metrics to answer these questions. Both metrics are based on a contingency, with droughts identified

as binary events (i.e., 1 = drought, and 0 = no drought). Calculation of the metrics is according to equations (7) and (8), such that

$$A = \frac{\text{hits} + \text{correct negatives}}{\text{total}}, \quad (7)$$

$$POD = \frac{\text{hits}}{\text{hits} + \text{misses}}. \quad (8)$$

$A$  represents the fraction of total predictions that were correct and range from 1 (perfect prediction) to 0, and  $POD$  measures the fraction of observed (in situ) droughts that were correctly depicted as model/satellite droughts, ranging from 0 to 1, with 1 representing a perfect score.

In addition, we adopt the accuracy metric to evaluate the frequency with which shared in situ model or in situ satellite droughts are depicted in the same USDM drought category. This metric ( $A_{cat}$ ) is calculated as

$$A_{cat} = \frac{1}{N} \sum_{i=1}^K n(F_i O_i), \quad (9)$$

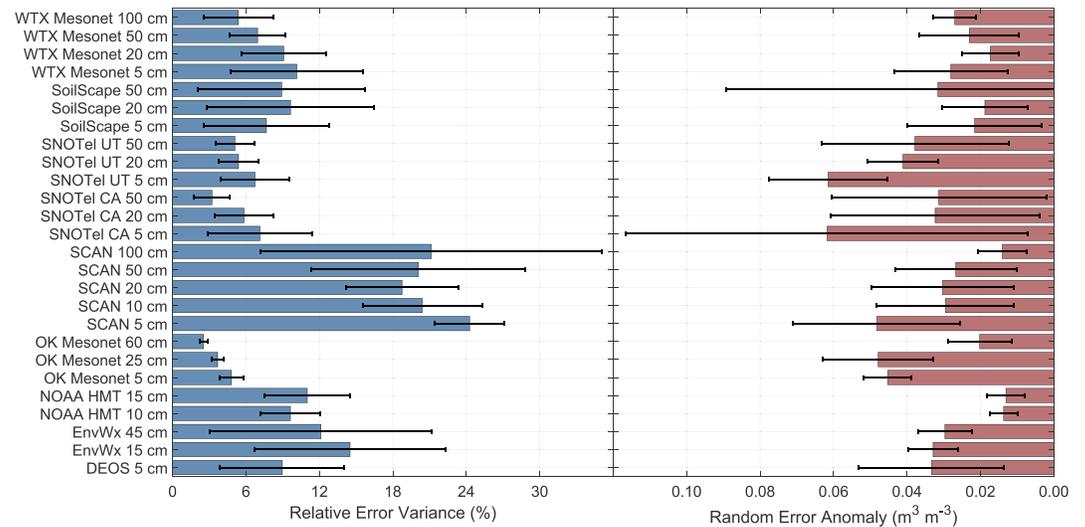
where  $N$  is the total number of shared in situ-model/satellite droughts and  $n(F_i O_i)$  is the number of shared in situ-model/satellite droughts that are both depicted in intensity category  $i$ . We adopt the USDM drought intensity categories of D1 to D4 (Svoboda et al., 2002). Accuracy values range from 0 to 1, with 1 being a perfect score, indicating all shared droughts were depicted in the same USDM intensity category. The primary limitation of the accuracy metric is that it is strongly influenced by the most common category, while we are most interested in how well the data sets match each other with regard to drought intensity.

## 4. Results

### 4.1. In Situ Validation

The fidelity of daily  $\theta$  at each in situ monitoring station was evaluated prior to model/satellite comparison. The relative error variance indicates the relative proportion of variability from measurement error to real soil moisture variability. Relative error variance (%) was computed at each station and averaged by network and depth (Figure 2), with error bars representing the range of individual station relative error variance values for each network. SNOTel, OK Mesonet, and WTX Mesonet stations exhibit the lowest relative error variance, with network-averaged values  $\leq 10\%$ ; meaning 10% or less of the overall variability in daily  $\theta$  is attributed to measurement error.

Assessment of relative error variance at the station level via a two way analysis of variance showed no significant ( $\alpha = 0.05$ ) difference by measurement depth, but that significant differences did exist between measurement networks. Further analysis revealed that these results were entirely due to the SCAN stations, which exhibited a network-averaged relative error variance exceeding 20%. For perspective, the SCAN network average relative error variance exceeded the maximum of all but six stations from the other networks. We computed the relative error variance from daily soil moisture observations from 63 other SCAN stations across the contiguous United States over the same 14-year period and found relative error variance was elevated—relative to the other networks assessed in this study—in these additional SCAN stations. This somewhat agrees with the results of Dirmeyer et al. (2016) who show SCAN stations, on average, exhibit relatively high relative error variance. The relative error variance is related to the displacement of the autocorrelation of daily soil moisture, and therefore the daily variability of soil moisture at a site. This is reflected in a comparison of daily time series of soil moisture from two nearby stations in Reese Center, Texas, one part of SCAN and the other part of the West Texas Mesonet network. Daily soil moisture from the 5-cm sensors at the two stations, located  $\sim 1$  km apart, exhibit considerable differences in daily variability, with the SCAN station exhibiting a coefficient of variation (0.62) twice that of the nearby West Texas Mesonet station (0.32). Enhanced daily variability, particularly when soils are drier than the climatological mean condition, at the SCAN site (relative to the nearby West Texas Mesonet site) decreases autocorrelation, thereby increasing  $a$  (equations (2) and (3)) and the relative error variance. This comparison, along with the consistency of



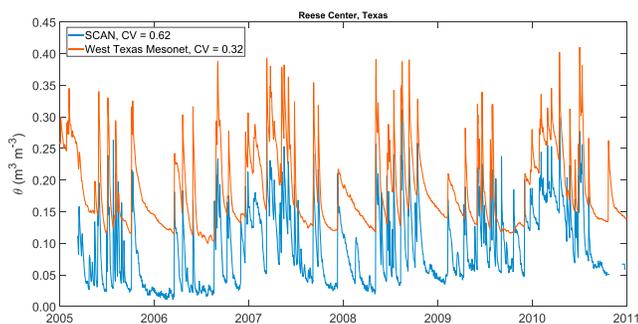
**Figure 2.** In situ soil moisture (left) relative error variance and (right) random anomaly error. Errors are calculated from daily volumetric water content and are averaged over each in situ network.

elevated relative error variance at numerous, diversely sited SCAN stations suggests the relative error variance at a single site is strongly determined by the climate, land cover, or soil texture characteristics of that site but instead is determined by the variability of daily soil moisture observations, irrespective of whether those variations are real or artificial. It is unclear whether this difference in variability between two nearby sites (Figure 3) is related to differences in sensor type; however, the same Stevens Hydraprobes used at the SCAN sites are also employed at the SNOTel sites in California and Utah, stations that exhibited low relative error variance.

Because there was no apparent relationship between sensor measurement depth and the relative error variance of the corresponding soil moisture time series, we assessed the quality of each individual sensor by comparing its relative error variance to that of all sensors, irrespective of measurement depth. The resultant distribution (Figure S2a) shows distinct bimodality, with modes of approximately 7% and 20% relative error variance. Based on the distribution we adopt a threshold relative error variance value of 15%, beyond which sensors are not considered for subsequent comparison with model or satellite data sets. We only use a station for comparison if none of its sensors exceeds 15% relative error variance, because all stations within a network must include observations at the same depths to facilitate fair, consistent comparison. Entire stations are omitted from the comparison analysis if all of the sensors exceed the 15% limit; this resulted in the omission of two DEOS stations, five EnvWx stations, and two SoilScape

stations. Only three SCAN sensors fell below the 15% relative error variance limit, and we therefore omitted all six SCAN stations from subsequent model/satellite comparison. It is important to note that while the SCAN stations exhibited relatively high error variance, one should not immediately consider the data of lesser quality. Indeed, SCAN is considered one of the highest quality soil moisture monitoring networks and is frequently used for model and satellite validation (e.g., Al Bitar et al., 2012; Njoku et al., 2003; Pan et al., 2016; Xia et al., 2014). However, given the results presented here, corroborating those of Dirmeyer et al. (2016), we suggest further inquiry into the causes of enhanced variability and relative error variance of SCAN soil moisture.

TC analysis was used, as a supplement to relative error variance, to evaluate the quality of in situ soil moisture measurements from monitoring networks. Specifically, TC is used to estimate the spatial representativeness of each in situ monitoring station, determined by the random error. Random errors were computed for each in situ station and then



**Figure 3.** Time series of daily 5-cm volumetric water content ( $\text{m}^3/\text{m}^3$ ) from two in situ stations near Reese Center, Texas. The SCAN station is shown in blue, and the West Texas Mesonet station in red. The coefficient of determination ( $\text{CV}, \sigma/\mu$ ) is reported for both time series.

averaged by network and depth (Figure 2). NOAA HMT, WTX Mesonet, and SoilScape networks exhibited the lowest random errors, suggesting in situ stations part of these networks best captured larger-scale patterns in the variability of  $\theta$ . Despite exhibiting low relative error variance, SNOTel stations showed among the highest random errors. This is not necessarily surprising, given SNOTel stations are sited for snowpack measurements in the western United States and therefore are often located in frequently snow-covered environments with complex topography. It is expected that a single in situ station may have difficulty capturing soil moisture variability across 50- to 100-km spatial scales in these locations. With that being said, only one station exhibited random errors exceeding  $0.10 \text{ (m}^3/\text{m}^3)$ . Our results are to those of Gruber et al. (2013) and suggests the in situ sites selected for this study are spatially representative. The only station with random errors  $>0.10$  is the SNOTel Poison Flat site in California, which—according to Jordan Clayton of the U.S. Department of Agriculture-Natural Resources Conservation Service—is in a lowland area that is always wetter than the surrounding region and is therefore “only representative of the alluvial area immediately adjacent to the stream” (personal communication, 18 December 2018). The distribution of TC random error, with all depths grouped together, is leptokurtic and exhibits strong, positive skew similar to a Rayleigh distribution (Figure S2b). Because all sensors, except for those of the SNOTel Poison Flat site, span a tight TC random error range, we only omitted the Poison Flat station from our subsequent model/satellite comparison.

Unlike relative error variance, the TC results exhibit statistically significant differences in both measurement depth and measurement network, based on a one-way analysis of variance. Deeper sensors exhibited lower random errors and thus were more spatially representative. This is most likely because atmospheric processes that influence daily soil moisture variability and spatial heterogeneity are modulated, making deeper-layer soil moisture more spatially homogenous and representative of larger areas surrounding the sensor. The significant random error differences by measurement network seem to be driven by topography and elevation, as there is a significant, negative correlation between station elevation and spatial representativeness. Similar to relative error variance, TC results at the station or network levels show no noticeable (or significant) relationship with either sensor type or soil texture.

Taken together, the relative error variance and TC random anomaly error provide a comprehensive framework for evaluating both the overall quality and spatial representativeness of in situ soil moisture across disparate measurement networks, sensors, climates, soils, and installation and quality control procedures. As a result of in situ validation using these procedures, at least one sensor at 16 sites (two DEOS, five EnvWx, six SCAN, one SNOTel, and two SoilSCAPE) were flagged, and so these sites were not used in subsequent model/satellite comparison.

#### 4.2. Model and Satellite Comparison

There are clear differences in  $\theta_a$  variability ( $\text{m}^3/\text{m}^3$ ) between in situ and model/satellite data sets (Figure 4), although these differences vary by in situ network and model/satellite product. Overall, model and satellite data sets tend to underestimate daily  $\theta_a$  variability (Figure 4, shaded in red), although these underestimates rarely exceed  $0.03 \text{ (m}^3/\text{m}^3)$ . Model/satellite overestimates of  $\theta_a$  (Figure 4, shaded in blue), although less prevalent, are larger in magnitude and appear systematic for certain in situ networks and model/satellite data sets. Specifically, daily  $\theta_a$  variability is larger in every model/satellite data set as compared to NOAA HMT sites, although further comparison finds that NOAA HMT site daily  $\theta_a$  variability is consistently one-half to one third that of all other stations, including nearby SNOTel CA and SoilScape sites. This suggests site-level processes reducing daily  $\theta_a$  variability at NOAA HMT sites are not affecting larger-scale soil moisture data sets. Otherwise, differences in  $\theta_a$  variability were unrelated to the in situ station relative error variance and random anomaly error, as well as the density of stations in a single model/satellite grid cell.

When variability differences are averaged over all networks (Figure 4, last column) the VIC model 40- to 100-cm data exhibited the largest deviation from in situ observations, overestimating  $\theta_a$  variability at all networks located in arid or semiarid environments while capturing well  $\theta_a$  variability in the two humid environment networks (DEOS and EnvWx). Interestingly, this issue is not apparent in the Noah or Mosaic models, nor in the near-surface VIC layers. This suggests that the VIC model may include some subsurface processes that result in enhanced deep-soil moisture variability that is not consistent with the in situ observations.

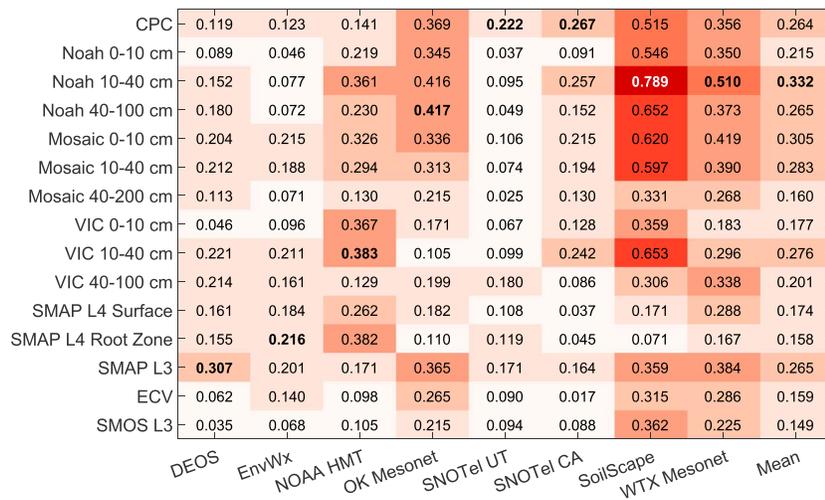
CPC	<b>0.000</b>	0.004	<b>0.084</b>	0.036	0.013	<b>0.043</b>	0.032	0.011	0.028
Noah 0-10 cm	-0.006	-0.005	0.024	-0.030	-0.008	-0.007	-0.004	-0.023	-0.007
Noah 10-40 cm	-0.010	-0.001	0.016	-0.026	-0.017	-0.006	-0.005	-0.030	-0.010
Noah 40-100 cm	-0.021	-0.012	<b>0.011</b>	-0.016	-0.035	-0.005	-0.004	-0.022	-0.013
Mosaic 0-10 cm	-0.010	-0.012	0.039	-0.005	-0.015	<b>0.003</b>	<b>0.003</b>	-0.019	-0.002
Mosaic 10-40 cm	-0.012	-0.009	0.035	-0.013	-0.005	0.011	-0.004	-0.018	-0.002
Mosaic 40-200 cm	-0.015	-0.010	0.031	-0.007	-0.040	-0.019	-0.018	-0.044	-0.015
VIC 0-10 cm	-0.010	-0.022	0.020	-0.039	<b>0.060</b>	-0.027	-0.009	-0.031	-0.007
VIC 10-40 cm	-0.021	-0.027	0.021	-0.006	<b>-0.002</b>	-0.015	-0.009	-0.033	-0.012
VIC 40-100 cm	-0.004	<b>0.000</b>	<b>0.100</b>	0.008	0.009	0.048	0.048	0.048	0.032
SMAP L4 0-5 cm	-0.012	-0.008	0.017	-0.023	-0.027	-0.018	<b>-0.003</b>	-0.023	-0.012
SMAP L4 0-100 cm	-0.017	0.002	0.017	<b>-0.002</b>	-0.025	-0.006	-0.017	-0.025	-0.009
SMAP L3	0.002	0.002	0.032	-0.015	-0.021	-0.019	0.006	<b>0.004</b>	<b>-0.001</b>
ECV	-0.004	-0.012	0.042	-0.024	-0.022	-0.019	<b>-0.003</b>	-0.020	-0.008
SMOS L3	<b>0.082</b>	0.031	0.045	0.011	-0.015	-0.018	<b>0.051</b>	0.017	0.026
	DEOS	EnVWx	NOAA HMT	OK Mesonet	SNOTel UT	SNOTel CA	SoilScape	WTX Mesonet	Mean

**Figure 4.** Model/satellite-in situ differences in  $\theta$  variability ( $\text{m}^3/\text{m}^3$ ). Positive differences (negative differences) represent more (less) model/satellite variability with respect to in situ data and are shaded in blue (red). Differences are averaged over each in situ network; the far right column shows the overall model/satellite product difference averaged over all networks. The smallest differences (i.e., closest to 0) for each in situ network are bolded.

Although there is no consistent relationship between the model/satellite and in situ differences in  $\theta_a$  variability and the spatial resolution of the model/satellite data set, the two products with the coarsest resolutions—CPC and SMOS L3—exhibited the largest differences, with the exception of the VIC 0- to 100-cm layer. Additionally, the CPC and SMOS L3 products exhibited the highest network-to-network variability in  $\theta_a$  variability differences (Figure S2). The SMAP L3 product exhibited the overall lowest  $\theta_a$  variability differences ( $-0.001 \text{ m}^3/\text{m}^3$ ), followed by the Mosaic 0- to 10-cm and 10- to 40-cm layers ( $-0.002 \text{ m}^3/\text{m}^3$ ), while the SMAP L4 surface and root zone products exhibited the lowest overall network-to-network variability in  $\theta_a$  variability differences (Figure S2).

The amount of shared variance between the in situ and model/satellite  $\theta_a$  data sets was assessed using the coefficient of determination ( $R^2$ ). Overall, and with the exception of SMAP L3, the model data sets exhibited higher shared variance with in situ observations than the satellite-informed products (Figure 5). Specifically, the CPC model, Noah 0- to 10-cm and 40- to 100-cm, Mosaic 0- to 10-cm and 10- to 40-cm, and the VIC 10- to 40-cm layers exhibited  $R^2$  values exceeding 0.25 (Figure 5, last column). SMAP L3 outperformed SMOS L3, ECV, and SMAP L4 data sets with average  $R^2$  values exceeding 0.25. This is a particularly interesting result for in situ networks that have sensors throughout the root zone, such as West Texas Mesonet (i.e., sensors at 5, 20, 60, and 75 cm), because the SMAP L3 product, which only represents soil moisture only the top 5–10 cm, outperformed the SMAP L4 product, which represents soil moisture in the 0- to 100-cm soil column. When averaged over all three NLDAS-2 models, shared variance was highest at the 10- to 40-cm layer. The SMAP L3 product also exhibited the lowest internetwork variability of shared variance (Figure S3a), suggesting consistent performance across diverse monitoring networks/regions.

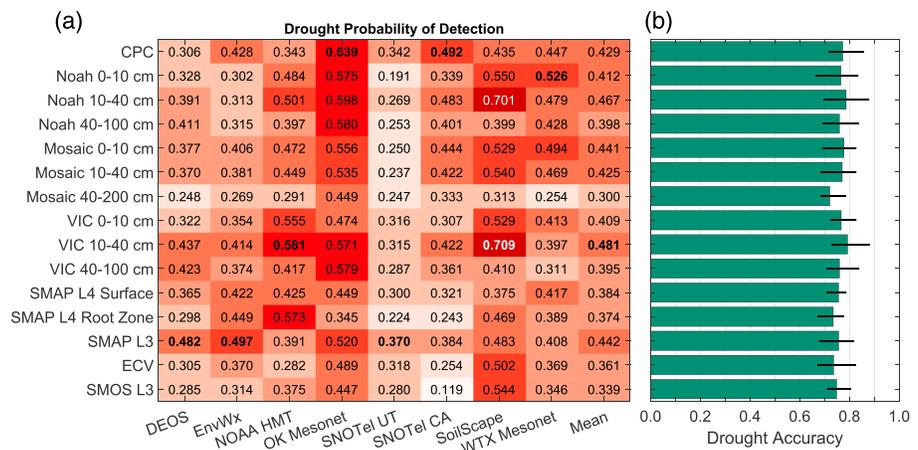
It is clear from Figure 5 that the model and satellite products overall exhibit the largest shared variance with SoilScape stations, most likely attributed to the high density of the SoilScape stations or nodes. Station density seems to be more important to the models than to the satellite-informed products, as the models exhibit considerably more variability in shared variance between in situ networks than the satellite data sets (Figure S3b). It is important to note, though, that there is a large gap in station density between SoilScape (13.75 stations per grid cell) and the next densest network, SNOTel Utah (2.44 stations per grid cell). Therefore, it would be difficult to assess, based on our data, if the relationship between model/satellite-in situ soil moisture shared variance and in situ station density is purely linear or only becomes important when station density exceeds some threshold. Clearly, other factors are contributing to the patterns in Figure 5, as  $R^2$  values are generally very low between model/satellite products and SNOTel Utah stations.



**Figure 5.** Shared variance ( $R^2$ ) between model/satellite and in situ  $\theta_a$  ( $m^3/m^3$ ). Values are averaged over each in situ network; the far right column shows the overall model/satellite product shared variance averaged over all networks. The largest values (i.e., closest to 1) for each in situ network are bolded.

### 4.3. Drought Monitoring

Because the overarching goal of this project is to improve near real time drought monitoring, we have focused the model/satellite comparison on the dry end of the soil moisture distribution. We identify drought conditions as any day in which the daily  $\theta_a \leq 20$ th percentile and evaluate the ability of model/satellite products to both identify drought conditions—as depicted by in situ observations—and discern between drought and nondrought conditions. The latter of these evaluations is accomplished via the accuracy metric ( $A$ ), which measures the fraction of correct identifications (drought and nondrought) to total predictions. Drought/nondrought accuracy varied very little between model and satellite products (Figure 6a), with grid cell-level accuracy values ranging from 68% to 89%. This consistent accuracy is most likely due to the high frequency of nondrought conditions, which, by definition, occur on 80% of days. With that being said, high drought accuracy results suggest each of the model/satellite products can discern between drought and nondrought conditions for the majority of days tested.



**Figure 6.** Comparison of drought occurrence between in situ and model/satellite data sets. Panel (a) shows the probability of drought detection of each model/satellite product, averaged over all grid cells for each in situ network. Panel (b) shows the overall accuracy of drought detection for each model/satellite product, averaged over all in situ networks. Error bars in panel b represent the range of accuracy values across all in situ sensors.

CPC	0.105	0.128	0.109	0.163	0.088	0.132	0.277	0.159	0.145
Noah 0-10 cm	0.122	0.101	0.197	0.202	0.073	0.119	0.302	0.183	0.162
Noah 10-40 cm	0.162	0.109	0.139	0.233	0.075	0.179	0.522	0.148	0.196
Noah 40-100 cm	0.140	0.104	0.102	0.194	0.083	0.154	0.273	0.126	0.147
Mosaic 0-10 cm	0.126	0.134	0.145	0.189	0.086	0.137	0.212	0.172	0.150
Mosaic 10-40 cm	0.126	0.118	0.103	0.189	0.076	0.130	0.284	0.152	0.147
Mosaic 40-200 cm	0.090	0.107	0.066	0.157	0.085	0.086	0.169	0.094	0.107
VIC 0-10 cm	0.107	0.123	0.206	0.149	0.094	0.106	0.230	0.143	0.145
VIC 10-40 cm	0.141	0.138	0.187	0.181	0.097	0.119	0.356	0.136	0.169
VIC 40-100 cm	0.147	0.155	0.119	0.211	0.115	0.101	0.259	0.079	0.148
SMAP L4 0-5 cm	0.091	0.099	0.123	0.117	0.087	0.060	0.118	0.091	0.098
SMAP L4 0-100 cm	0.085	0.103	0.032	0.128	0.071	0.061	0.112	0.091	0.085
SMAP L3	0.073	0.053	0.086	0.151	0.075	0.087	0.095	0.108	0.091
ECV	0.103	0.095	0.060	0.105	0.050	0.018	0.187	0.089	0.088
SMOS L3	0.047	0.029	0.036	0.080	0.020	0.008	0.090	0.052	0.045
	DEOS	EnVWx	NOAA HMT	OK Mesonet	SNOTel UT	SNOTel CA	SoilScape	WTX Mesonet	Mean

**Figure 7.** Accuracy of model/satellite soil moisture drought severity compared to that from in situ observations. Drought severity is determined using the percentile categories of the U.S. Drought Monitor. Values shown are averaged over each in situ network; the far right column shows the overall model/satellite product accuracy averaged over all networks.

The ability of models/satellites to identify drought conditions is more explicitly evaluated by the probability of detection ( $POD$ ) metric, which measures the fraction of observed (in situ) droughts that were correctly depicted as model/satellite droughts. Considerably more variability exists in drought  $POD$  between in situ networks and model/satellite products (Figure 6b). In general, the model products outperformed the satellite-informed products, with the VIC 10- to 40-cm layer data exhibiting the highest  $POD$  (0.481). For perspective, a  $POD$  of 0.481 reveals that 48% of in situ droughts were also depicted as droughts by the VIC model. Consistent with previous results, the SMAP L3 product outperformed all other satellite-informed data sets, and most model products as well ( $POD = 0.442$ ). At the other end, the Mosaic 40- to 200-cm data exhibited the lowest  $POD$  (0.300) by a considerable margin, perhaps due in part to the fact that none of the in situ networks have used sensors below 75 cm. The majority of the 160-cm Mosaic layer is deeper than the deepest in situ sensors, and less representative of drought as depicted by those sensors. To this end, the 10- to 40-cm VIC and Noah layers exhibited higher  $POD$  than the 0- to 10-cm layers, and although the Mosaic 0- to 10-cm layer outperformed its 10- to 40-cm layer, both exhibited  $POD$  values exceed 0.40. Similar to the variability and shared variance results, model/satellite products most accurately identified drought conditions at SoilScape stations. This is not surprising given their high density of in situ measurements at these locations. Also consistent with previous results is the relatively poor model/satellite performance at SNOTel stations.

Evaluation of the model/satellite data sets was accomplished using  $A$  and  $POD$ , and it is based on drought as a binary variable, drought or no drought. This analysis is complemented by using the accuracy ( $A_{cat}$ ) metric to assess how often drought events of a certain intensity as defined by a model/satellite data set corresponds with drought of the same intensity in the in situ data. We adopt the four-category classification scheme of the USDM for this analysis. The mean  $A_{cat}$  values are calculated for each model/satellite-in situ combination (Figure 7). When evaluating identification of both the occurrence and severity of drought, a noticeable difference appears between model and satellite products. All of the model data sets exhibit a higher average  $A_{cat}$  (i.e., better drought severity depiction) than any of the satellite products (Figure 7), which, with the exception of SMAP L3, is consistent with other analyses conducted in this study. Further analysis revealed SMAP L3 percentiles and drought severity categories exhibited larger variability than the 5-cm in situ observations, likely due to the short SMAP data record, relative to NLDAS-2 and CPC. Despite SMAP L3 having the best match with the daily variability of in situ  $\theta_a$ , the SMAP distribution of daily  $\theta_a$  values is considerably smaller than that of the NLDAS-2 and CPC models. The data record length is imperative for the stability of soil moisture percentiles based on that record (i.e., Ford et al., 2016), and therefore, the shorter SMAP data record enhances daily percentile variability and reduces the effectiveness of the SMAP product for identifying drought severity based on soil moisture percentiles. Concurrently, the data record length of in situ

observations impacts the stability of in situ soil moisture percentiles as well. This potentially explains why nearly all model and satellite data sets identified drought occurrence and severity with high skill at Oklahoma Mesonet stations, relative to other networks, as the Mesonet has the longest data record of all networks (19 years).

We did not find the horizontal resolution of model/satellite products, in situ sensor depth, or site-specific characteristics such as vegetation density, soil texture, and sensor type significant explanatory factors for model/satellite drought occurrence and severity accuracy. Although we expect these factors to influence the ability of the model, satellite, or in situ data to effectively represent soil moisture, they do not appear to play a noteworthy role in influencing the results presented here.

## 5. Discussion and Conclusions

Soil moisture is a key variable for drought monitoring and prediction; however, there are relatively few in situ soil moisture stations. Data collection efforts like the International Soil Moisture Network (Dorigo et al., 2011) and the NSMN (McNutt et al., 2016) are reasons to be optimistic about the future of in situ soil moisture monitoring. However, as new in situ networks come online and existing networks improve and/or expand their monitoring infrastructure, it is important to understand the overall quality and representativeness of the in situ observations, particularly if these measurements are to be used for model or satellite calibration and validation. The in situ networks assessed here are among the most widely used for such purposes. Our coupled relative error variance and triple collocation in situ data validation demonstrate that the measurements from these networks are generally of high quality. However, variability exists both within and between networks. The six SCAN stations evaluated exhibited unusually high relative error variance, suggesting a considerable fraction of the daily soil moisture variability at these stations could be attributed to measurement error. Further evaluation revealed this phenomenon was common among SCAN stations, even when compared to nearly colocated stations from other networks. This elevated error variance in the SCAN data cannot be attributed to station location, sensor type, monitoring depth, or spatial representativeness. Therefore, further site-specific investigations are needed to identify the cause. This is especially important because SCAN data are widely used for validating (and in some cases, calibrating) model and satellite products. SNOTel stations are found to exhibit relatively low spatial representativeness. Although, with the exception of the Poison Flat station in California, all SNOTel stations remained within an acceptable range of random anomaly errors. One advantage of in situ data validation based on relative error variance and triple collocation is that it can be completed for each individual sensor, and therefore, it can be used to identify poorly performing or damaged sensors at otherwise well-sited stations.

Comparison of model/satellite soil moisture products with in situ observations revealed a number of noteworthy findings. The two products with the coarsest horizontal resolution, SMOS L3 and CPC, exhibited the largest deviation from in situ daily soil moisture anomaly variability, as well as the highest network-to-network deviation variability. On the other hand, a consistent relationship between product horizontal resolution and overall performance was not found. This is partly attributable to the relatively small sample size of in situ stations and model/satellite products examined. Overall, satellite-informed products, with the exception of SMAP L3, exhibited less shared variance with in situ soil moisture anomalies than the NLDAS-2 and CPC models. SMAP L3 performed similarly well to the models and exhibited the lowest internetwork variability for all validation metrics. Model performance, particularly NLDAS-2 model performance, increased dramatically when compared to the high-density SoilScape stations, relative to other in situ networks. In contrast, SMAP L3 performed consistently at most networks and was less influenced by the network's station density. Interestingly, SMAP L3 outperformed SMAP L4 surface and root zone products at nearly all networks, even those with sensors placed throughout the root zone. SMAP L3 also consistently outperformed SMOS L3 and ECV data sets.

Despite the strong performance of NLDAS-2 models, discrepancies were apparent between the modeled and in situ data. Mo et al. (2012) found that uncertainties in soil moisture from the NLDAS are mostly attributed to precipitation forcing, although model structure and parameter errors are also important sources of soil moisture uncertainty (Xia et al., 2014). The former of these issues is mitigated through direct assimilation of soil moisture measurements, namely, satellite-based observations (Kumar et al., 2014; Reichle et al., 2017). Based on the results presented here, assimilation of SMAP soil moisture may also help increase

consistency of NLDAS-2 soil moisture accuracy across diverse sites. It is also important to note that future phases of NLDAS and the NCEP Unified Land Data Assimilation System (NULDAS) will contain operational data assimilation of soil moisture and snow. Problems with NLDAS model soil hydraulic parameters and model structure should be addressed, given that improved parameterizations result in more realistic soil moisture dynamics (Kishné et al., 2017).

Drought occurrence at in situ stations were most accurately identified by the CPC and NLDAS-2 models, as well as by the SMAP L3 product. With that being said, no model or satellite data set was able to capture more than 48% of in situ drought events, when averaged over all networks. Not surprisingly, the model and satellite products were less accurate in identifying drought severity, based on USDM percentiles, than identifying drought occurrence. This decrease in accuracy was largest for the SMAP L3 product, and further analysis revealed this was at least partly attributable to the relatively short SMAP data record, which decreased percentile stability and reduced drought severity accuracy. Among the NLDAS-2 models, the 10- to 40-cm layer most consistently identified in situ drought events. This is not surprising, given the importance of root zone soil moisture for vegetation productivity.

Overall, we found the large majority of in situ stations assessed have high fidelity and can be used with confidence for model/satellite data validation. Our results showed that the NLDAS-2 models and SMAP L3 data were able to capture daily  $\theta_a$  variability, had moderate-to-high degree of shared variance with the in situ observations, and were able to accurately identify drought occurrence. The CPC model did not perform as well as the NLDAS-2 models or SMAP L3, as it tended to overestimate  $\theta_a$  variability. However, the CPC model was able to identify drought conditions with similar accuracy as NLDAS-2 models and SMAP L3. SMAP L3 was able to outperform SMAP L4 (surface and root zone), ECV, and SMOS L3 products based on nearly all metrics of evaluation. These results are consistent with prior studies that have shown SMAP outperform other satellite-based soil moisture products (Chen et al., 2018; Kim et al., 2018; Montzka et al., 2017). SMAP retrievals are based on L-band passive microwave observations, which have been shown to be more sensitive to soil moisture than X-band or C-band sensors. SMOS L-band retrievals have been shown to be significantly affected by radio frequency interference (Kumar et al., 2018), which partially explains the superior performance of SMAP versus SMOS.

The metrics used in this study provide a framework for evaluation of in situ, model, and satellite soil moisture fidelity and can be applied across diverse geographic regions to understand the extent and source of soil moisture data uncertainty. Future work will apply this framework to all of the NSMN in situ stations (roughly 1,000+ total) and to as many operational land surface models and satellite remote sensing products as can be obtained.

#### Acknowledgments

In situ observations as part of the NOAA HMT, OK Mesonet, SCAN, SNOTel, and WTX Mesonet are available at the National Soil Moisture Network website (<https://nationalsoilmoisture.com>). DEOS observations are available at the DEOS website ([http://www.deos.udel.edu/data/agirrigation\\_retrieval.php](http://www.deos.udel.edu/data/agirrigation_retrieval.php)), EnvWx observations are available at the Enviroweather website (<https://enviroweather.msu.edu/index.php>), and SoilScape observations are available at the SoilSCAPE website ([http://soilscape.usc.edu/bootstrap/sites\\_and\\_data.html](http://soilscape.usc.edu/bootstrap/sites_and_data.html)). CPC model soil moisture is available at the NCEP website ([ftp://ftp.cpc.ncep.noaa.gov/GIS/USDM\\_Products/soil/](ftp://ftp.cpc.ncep.noaa.gov/GIS/USDM_Products/soil/)), NLDAS-2 model soil moisture is available at the GSFC website ([https://disc.gsfc.nasa.gov/datasets?keywords=NLDAS,SMAP\\_L4](https://disc.gsfc.nasa.gov/datasets?keywords=NLDAS,SMAP_L4)) and SMAP L3 data are available at the NSIDC website (<https://nsidc.org/data/smap/smap-data.html>), SMOS L3 data are available at the CATDS website (<http://www.catds.fr/Products/Available-products-from-CPDC>), and ECV data are available at the ESA CCI Soil Moisture website (<http://www.esa-soilmoisture-cci.org/>). This work was supported by NOAA Grant NA17OAR4310148.

#### References

- Al Bitar, A., Leroux, D., Kerr, Y. H., Merlin, O., Richaume, P., Sahoo, A., & Wood, E. F. (2012). Evaluation of SMOS soil moisture products over continental US using the SCAN/SNOTEL network. *IEEE Transactions on Geoscience and Remote Sensing*, *50*(5), 1572–1586.
- Al Bitar, A., Mialon, A., Kerr, Y. H., Cabot, F., Richaume, P., Jacqueline, E., et al. (2017). The global SMOS level 3 daily soil moisture and brightness temperature maps. *Earth System Science Data*, *9*(1), 293–315. <https://doi.org/10.5194/essd-9-293-2017>
- Andresen, J., Olsen, L., Aichele, T., Bishop, B., Brown, J., Landis, J., et al. (2011). Enviro-weather: A weather-based pest and crop management information system for Michigan. Proc. 7th International Integrated Pest Management Symposium, Memphis, TN.
- Chen, F., Crow, W. T., Bindlish, R., Colliander, A., Burgin, M. S., Asanuma, J., & Aida, K. (2018). Global-scale evaluation of SMAP, SMOS and ASCAT soil moisture products using triple collocation. *Remote Sensing of Environment*, *214*(1), 1–13. <https://doi.org/10.1016/j.rse.2018.05.008>
- Chen, F., Mitchell, K., Schaake, J., Xue, Y., Pan, H. L., Koren, V., et al. (1996). Modeling of land surface evaporation by four schemes and comparison with FIFE observations. *Journal of Geophysical Research*, *101*(D3), 7251–7268. <https://doi.org/10.1029/95JD02165>
- Colliander, A., Jackson, T. J., Bindlish, R., Chan, S., Das, N., Kim, S. B., et al. (2017). Validation of SMAP surface soil moisture products with core validation sites. *Remote Sensing of Environment*, *191*, 215–231. <https://doi.org/10.1016/j.rse.2017.01.021>
- Dirmeyer, P. A., Wu, J., Norton, H. E., Dorigo, W. A., Quiring, S. M., Ford, T. W., et al. (2016). Confronting weather and climate models with observational data from soil moisture networks over the United States. *Journal of Hydrometeorology*, *17*(4), 1049–1067. <https://doi.org/10.1175/JHM-D-15-0196.1>
- Dorigo, W. A., Wagner, W., Albergel, C., Albrecht, F., Balsamo, G., Brocca, L., et al. (2017). ESA CCI soil moisture for improved Earth system understanding: State-of-the-art and future directions. *Remote Sensing of Environment*, *203*, 185–215.
- Dorigo, W. A., Wagner, W., Hohensinn, R., Hahn, S., Paulik, C., Xaver, A., et al. (2011). The International Soil Moisture Network: A data hosting facility for global in situ soil moisture measurements. *Hydrology and Earth System Sciences*, *15*(5), 1675–1698. <https://doi.org/10.5194/hess-15-1675-2011>
- Ek, M. B., Peters-Lidard, C. D., Xia, Y., Mocko, D. M., Meng, J., Kumar, S. V., et al. (2017). Next phase of the NCEP Unified Land Data Assimilation System (NULDAS): Vision, requirements, and implementation. NLDAS White Paper (17 pp.). Retrieved from [http://www.emc.ncep.noaa.gov/mmb/nldas/White\\_Paper\\_for\\_Next\\_Phase\\_LDAS\\_final.pdf](http://www.emc.ncep.noaa.gov/mmb/nldas/White_Paper_for_Next_Phase_LDAS_final.pdf)

- Entekhabi, D., Njoku, E. G., O'Neill, P. E., Kellogg, K. H., Crow, W. T., Edelstein, W. N., et al. (2010). The soil moisture active passive (SMAP) mission. *Proceedings of the IEEE*, 98(5), 704–716. <https://doi.org/10.1109/JPROC.2010.2043918>
- Fan, Y., & van den Dool, H. (2004). Climate Prediction Center global monthly soil moisture data set at 0.5° resolution for 1948 to present. *Journal of Geophysical Research*, 109, D10102. <https://doi.org/10.1029/2003JD004345>
- Ford, T. W., McRoberts, D. B., Quiring, S. M., & Hall, R. E. (2015). On the utility of in situ soil moisture observations for flash drought early warning in Oklahoma, USA. *Geophysical Research Letters*, 42, 9790–9798. <https://doi.org/10.1002/2015GL066600>
- Ford, T. W., Wang, Q., & Quiring, S. M. (2016). The observation record length necessary to generate robust soil moisture percentiles. *Journal of Applied Meteorology and Climatology*, 55(10), 2131–2149. <https://doi.org/10.1175/JAMC-D-16-0143.1>
- Gruber, A., Dorigo, W., Crow, W., & Wagner, W. (2017). Triple collocation-based merging of satellite soil moisture retrievals. *IEEE Transactions on Geoscience and Remote Sensing*, 55(12), 6780–6792. <https://doi.org/10.1109/TGRS.2017.2734070>
- Gruber, A., Dorigo, W. A., Zwieback, S., Xaver, A., & Wagner, W. (2013). Characterizing coarse-scale representativeness of in situ soil moisture measurements from the International Soil Moisture Network. *Vadose Zone Journal*, 12(2). <https://doi.org/10.2136/vzj2012.0170>
- Hatfield, J. L., Boote, K. J., Kimball, B. A., Ziska, L. H., Izaurralde, R. C., Ort, D., et al. (2011). Climate impacts on agriculture: Implications for crop production. *Agronomy Journal*, 103(2), 351–370. <https://doi.org/10.2134/agronj2010.0303>
- Hayes, M., Svoboda, M., Wall, N., & Widhalm, M. (2011). The Lincoln declaration on drought indices: Universal meteorological drought index recommended. *Bulletin of the American Meteorological Society*, 92(4), 485–488. <https://doi.org/10.1175/2010BAMS103.1>
- Hoerling, M., Kumar, A., Dole, R., Nielsen-Gammon, J. W., Eischeid, J., Perlwitz, J., et al. (2013). Anatomy of an extreme event. *Journal of Climate*, 26(9), 2811–2832. <https://doi.org/10.1175/JCLI-D-12-0027.1>
- Huang, J., van den Dool, H. M., & Georgarakos, K. P. (1996). Analysis of model-calculated soil moisture over the United States (1931–1993) and applications to long-range temperature forecasts. *Journal of Climate*, 9(6), 1350–1362. <https://doi.org/10.1175/1520-0442>
- Jackson, T. J., Bindlish, R., Cosh, M. H., Zhao, T., Starks, P. J., Bosch, D. D., et al. (2012). Validation of Soil Moisture and Ocean Salinity (SMOS) soil moisture over watershed networks in the US. *IEEE Transactions on Geoscience and Remote Sensing*, 50(5), 1530–1543. <https://doi.org/10.1109/TGRS.2011.2168533>
- Kerr, Y. H., Waldteufel, P., Wigneron, J. P., Delwart, S., Cabot, F., Boutin, J., et al. (2010). The SMOS mission: New tool for monitoring key elements of the global water cycle. *Proceedings of the IEEE*, 98(5), 666–687. <https://doi.org/10.1109/JPROC.2010.2043032>
- Kim, H., Parinussa, R., Konings, A. G., Wagner, W., Cosh, M. H., Lakshmi, V., et al. (2018). Global-scale assessment and combination of SMAP and ASCAT (active) and AMSR2 (passive) soil moisture products. *Remote Sensing of Environment*, 204, 260–275. <https://doi.org/10.1016/j.rse.2017.10.026>
- Kishné, A. S., Yimam, Y. T., Morgan, C. L. S., & Dornblaser, B. C. (2017). Evaluation and improvement of the default soil hydraulic parameters for the Noah Land Surface Model. *Geoderma*, 285, 247–259. <https://doi.org/10.1016/j.geoderma.2016.09.022>
- Koster, R. D., Mahanama, S. P. P., Yamada, T. J., Balsamo, G., Berg, A. A., Boissier, M., et al. (2010). Contribution of land surface initialization to subseasonal forecast skill: First results from a multi-model experiment. *Geophysical Research Letters*, 37, L02402. <https://doi.org/10.1029/2009GL041677>
- Koster, R. D., & Suarez, M. J. (1992). Modeling the land surface boundary in climate models as a composite of independent vegetation stands. *Journal of Geophysical Research*, 97(D3), 2697–2715. <https://doi.org/10.1029/91JD01696>
- Kumar, S. V., Dirmeyer, P. A., Peters-Lidard, C. D., Bindlish, R., & Bolten, J. (2018). Information theoretic evaluation of satellite soil moisture retrievals. *Remote Sensing of Environment*, 204, 392–400. <https://doi.org/10.1016/j.rse.2017.10.016>
- Kumar, S. V., Peters-Lidard, C. D., Mocko, D., Reichle, R., Liu, Y., Arsenault, K. R., et al. (2014). Assimilation of remotely sensed soil moisture and snow depth retrievals for drought estimation. *Journal of Hydrometeorology*, 15(6), 2446–2469. <https://doi.org/10.1175/JHM-D-13-0132.1>
- Kumar, S. V., Wang, S., Mocko, D. M., & Peters-Lidard, C. D. (2017). Similarity assessment of land surface model outputs in the North American Land Data Assimilation (NLDAS). *Water Resources Research*, 53, 8941–8965. <https://doi.org/10.1002/2017WR020635>
- Legates, D. R., Leathers, D. J., DeLiberty, T. L., Quelch, G. E., Brinson, K., Butke, J., et al. (2005). The Delaware Environmental Observing System. 21st International Conference on Interactive Information Processing Systems (IIPS) for Meteorology, Oceanography, and Hydrology.
- Liang, X., Lettenmaier, D. P., Wood, E. F., & Burges, S. J. (1994). A simply hydrologically based model of land surface water and energy fluxes for general circulation models. *Journal of Geophysical Research*, 99(D7), 14,415–14,428. <https://doi.org/10.1029/94JD00483>
- Liu, D., Wang, G., Mei, R., Yu, Z., & Yu, M. (2014). Impact of initial soil moisture anomalies on climate mean and extremes over Asia. *Journal of Geophysical Research: Atmospheres*, 119, 529–545. <https://doi.org/10.1002/2013JD020890>
- Liu, Y. Y., Dorigo, W. A., Parinussa, R. M., de Jeu, R. A. M., Wagner, W., McCabe, M. F., et al. (2012). Trend-preserving blending of passive and active microwave soil moisture retrievals. *Remote Sensing of Environment*, 123, 280–297. <https://doi.org/10.1016/j.rse.2012.03.014>
- McNutt, C., Strobel, M., Lucido, J., & Quiring, S. (2016). National Soil Moisture Network Workshop 2016: Progress made, future directions. <https://www.drought.gov/drought/sites/drought.gov.drought/files/2016%20SMMW%20Boulder%20FINAL%2012232016.pdf>
- McPherson, R. A., Fiebrich, C. A., Crawford, K. C., Kilby, J. R., Grimsley, D. L., Martinez, J. E., et al. (2007). Statewide monitoring of the mesoscale environment: A technical update on the Oklahoma Mesonet. *Journal of Atmospheric and Oceanic Technology*, 24(3), 301–321. <https://doi.org/10.1175/JETCH1976.1>
- Miralles, D. G., Teuling, A. J., van Heerwaarden, C. C., & de Arellano, J. V. G. (2014). Mega-heatwave temperatures due to combined soil desiccation and atmospheric heat accumulation. *Nature Geoscience*, 7(5), 345–349. <https://doi.org/10.1038/ngeo2141>
- Mo, K. C., Chen, L. C., Shukla, S., Bohn, T. J., & Lettenmaier, D. P. (2012). Uncertainties in North American land data assimilation systems over the contiguous United States. *Journal of Hydrometeorology*, 13(3), 996–1009. <https://doi.org/10.1175/JHM-D-11-0132.1>
- Moghaddan, M., Burgin, M., Castillo, A., Entekhabi, D., Goykhman, Y., Kakhbod, A., et al. (2010). Soil moisture sensing controller and optimal estimator (SoilSCAPE): First deployment of the wireless sensor network and latest progress on soil moisture satellite retrieval validation strategies. Earth Science Technology Forum (ESTF10).
- Montzka, C., Bogen, H. R., Zreda, M., Moneris, A., Morrison, R., Muddu, S., & Vereecken, H. (2017). Validation of spaceborne and modeled surface soil moisture products with Cosmic-Ray Neutron probes. *Remote Sensing*, 9(2), 103. <https://doi.org/10.3390/rs9020103>
- Njoku, E. G., Jackson, T. J., Venkat, L., Chan, T. K., & Nghiem, S. V. (2003). Soil moisture retrieval from AMSR-E. *IEEE Transactions on Geoscience and Remote Sensing*, 41(2), 215–229. <https://doi.org/10.1109/TGRS.2002.808243>
- O'Neill, P. E., Chan, S., Njoku, E. G., Jackson, T., & Bindlish, R. (2018). SMAP L3 Radiometer Global Daily 36 km EASE-Grid Soil Moisture, Version 5. Boulder, Colorado, USA. NASA National Snow and Ice Data Center Distributed Active Archive Center, accessed January, 2018.

- Pan, M., Cai, X., Chaney, N. W., Entekhabi, D., & Wood, E. F. (2016). An initial assessment of SMAP soil moisture retrievals using high-resolution model simulations and in situ observations. *Geophysical Research Letters*, *43*, 9662–9668. <https://doi.org/10.1002/2016GL069964>
- Quiring, S. M., Ford, T. W., Wang, J. K., Khong, A., Harris, E., Lindgren, T., et al. (2016). The North American soil moisture database: Development and applications. *Bulletin of the American Meteorological Society*, *97*(8), 1441–1459. <https://doi.org/10.1175/BAMS-D-13-0026.1>
- Quiring, S. M., & Papkyriakou, T. N. (2003). An evaluation of agricultural drought indices for the Canadian prairies. *Agricultural and Forest Meteorology*, *118*(1–2), 49–62. [https://doi.org/10.1016/S0168-1923\(03\)00072-8](https://doi.org/10.1016/S0168-1923(03)00072-8)
- Reichle, R., de Lannoy, G., Koster, R. D., Crow, W. T., Kimball, J. S., & Liu, Q. (2018). SMAP L4 Global 3-hourly 9 km EASE-Grid Surface and Root Zone Soil Moisture Geophysical Data, Version 4. Boulder, CO. NASA National Snow and Ice Data Center Distributed Active Archive Center, accessed January, 2018.
- Reichle, R. H., de Lannoy, G. J. M., Liu, Q., Ardizzone, J. V., Colliander, A., Conaty, A., et al. (2017). Assessment of the SMAP Level-4 surface and root-zone soil moisture product using in situ measurements. *Journal of Hydrometeorology*, *18*(10), 2621–2645. <https://doi.org/10.1175/JHM-D-17-0063.1>
- Robock, A., Vinnikov, K. Y., Schlosser, C. A., Speranskaya, N., & Xue, Y. (1995). Use of midlatitude soil moisture and meteorological observations to validate soil moisture simulations with biosphere and bucket models. *Journal of Climate*, *8*(1), 15–35.
- Schaefer, G. L., Cosh, M. H., & Jackson, T. J. (2007). The USDA natural resources conservation service soil climate analysis network (SCAN). *Journal of Atmospheric and Oceanic Technology*, *24*(12), 2073–2077. <https://doi.org/10.1175/2007JTECHA930.1>
- Schaefer, G. L., & Paetzold, R. F. (2001). SNOTEL (SNOWpack TELelemetry) and SCAN (soil climate analysis network). Proc. Intl. Workshop on Automated Weather Stations for Appl. In Ar. And Water Resour. Mgmt.
- Schroeder, J. L., Burgett, W. S., Haynie, K. B., Sonmez, I., Skwira, G. D., Doggett, A. L., & Lipe, J. W. (2005). The West Texas mesonet: A technical overview. *Journal of Atmospheric and Oceanic Technology*, *22*(2), 211–222. <https://doi.org/10.1175/JTECH-1690.1>
- Schubert, S. D., Wang, H., Koster, R. D., Suarez, M. J., & Groisman, P. (2014). Northern Eurasian heat waves and droughts. *Journal of Climate*, *27*(9), 3169–3207. <https://doi.org/10.1175/JCLI-D-13-0060.1>
- Sheffield, J., Wood, E. F., Chaney, N., Guan, K., Sadri, S., Yuan, X., et al. (2014). A drought monitoring and forecasting system for sub-Saharan African water resources and food security. *Bulletin of the American Meteorological Society*, *95*(6), 861–882. <https://doi.org/10.1175/BAMS-D-12-00124.1>
- Svoboda, M., LeComte, D., Hayes, M., Heim, R., Gleason, K., Angel, J., et al. (2002). The drought monitor. *Bulletin of the American Meteorological Society*, *83*(8), 1181–1190. <https://doi.org/10.1175/1520-0477-83.8.1181>
- van den Dool, H., Huang, J., & Fan, Y. (2003). Performance and analysis of the constructed analogue method applied to US soil moisture over 1981–2001. *Journal of Geophysical Research*, *108*(D16), 8617. <https://doi.org/10.1029/2002JD003114>
- van Dijk, A. I. J. M., Beck, H. E., Crosbie, R. S., de Jeu, R. A. M., Liu, Y. Y., Podger, G. M., et al. (2013). The millennium drought in southeast Australia (2001–2009): Natural and human causes and implications for water resources, ecosystems, economy, and society. *Water Resources Research*, *49*, 1040–1057. <https://doi.org/10.1002/wrcr.20123>
- Whan, K., Zscheischler, J., Orth, R., Shongwe, M., Rahimi, M., Asare, E. O., & Seneviratne, S. I. (2015). Impact of soil moisture on extreme maximum temperatures in Europe. *Weather and Climate Extremes*, *9*, 57–67. <https://doi.org/10.1016/j.wace.2015.05.001>
- Wood, A. W. (2008). The University of Washington Surface Water Monitor: An experimental platform for national hydrologic assessment and prediction. In 22nd Conference on Hydrology. Retrieved from <https://ams.confex.com/ams/pdfpapers/134844.pdf>
- Xia, Y., Ek, M. B., Wu, Y., Ford, T., & Quiring, S. M. (2015a). Comparison of NLDAS-2 simulated and NASMD observed daily soil moisture. Part I: Comparison and analysis. *Journal of Hydrometeorology*, *16*(5), 1962–1980. <https://doi.org/10.1175/JHM-D-14-0096.1>
- Xia, Y., Ek, M. B., Wu, Y., Ford, T., & Quiring, S. M. (2015b). Comparison of NLDAS-2 simulated and NASMD observed daily soil moisture. Part II: Impact of soil texture classification and vegetation type mismatches. *Journal of Hydrometeorology*, *16*(5), 1981–2000. <https://doi.org/10.1175/JHM-D-14-0097.1>
- Xia, Y., Mitchell, K., Ek, M., Sheffield, J., Cosgrove, B., Wood, E., et al. (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. *Journal of Geophysical Research*, *117*, D03109. <https://doi.org/10.1029/2011JD016048>
- Xia, Y., Sheffield, J., Ek, M. B., Dong, J., Chaney, N., Wei, H., et al. (2014). Evaluation of multi-model simulated soil moisture in NLDAS-2. *Journal of Hydrology*, *512*, 107–125. <https://doi.org/10.1016/j.jhydrol.2014.02.027>
- Zamora, R. J., Ralph, F. M., Clark, E., & Scheider, T. (2011). The NOAA Hydrometeorological testbed soil moisture observing networks: Design, instrumentation, and preliminary results. *Journal of Atmospheric and Oceanic Technology*, *28*(9), 1129–1140. <https://doi.org/10.1175/2010JTECHA1465.1>