



Article

Machine Learning Downscaling of SoilMERGE in the United States Southern Great Plains

Kenneth Tobin ^{1,*}, Aaron Sanchez ¹, Daniela Esparza ¹, Miguel Garcia ¹, Deepak Ganta ² and Marvin Bennett ²

¹ Center for Earth and Environmental Studies, Texas A&M International University, Laredo, TX 78041, USA; aaron.sanchez@tamiu.edu (A.S.); daniela.esparza@tamiu.edu (D.E.); miguel.garcia01@tamiu.edu (M.G.)

² School of Engineering, Texas A&M International University, Laredo, TX 78041, USA; deepak.ganta@tamiu.edu (D.G.); mbennett@tamiu.edu (M.B.)

* Correspondence: ktobin@tamiu.edu; Tel.: +1-956-326-2417

Abstract: SoilMERGE (SMERGE) is a root-zone soil moisture (RZSM) product that covers the entire continental United States and spans 1978 to 2019. Machine learning techniques, Random Forest (RF), eXtreme Gradient Boosting (XGBoost), and Gradient Boost (GBoost) downscaled SMERGE to spatial resolutions straddling the field scale domain (100 to 3000 m). Study area was northern Oklahoma and southern Kansas. The coarse resolution of SMERGE (0.125 degree) limits this product's utility. To validate downscaled results *in situ* data from four sources were used that included: United States Department of Energy Atmospheric Radiation Measurement (ARM) observatory, United States Climate Reference Network (USCRN), Soil Climate Analysis Network (SCAN), and Soil moisture Sensing Controller and optimal Estimator (SoilSCAPE). In addition, RZSM retrievals from NASA's Airborne Microwave Observatory of Subcanopy and Surface (AirMOSS) campaign provided a nearly spatially continuous comparison. Three periods were examined: era 1 (2016 to 2019), era 2 (2012 to 2015), and era 3 (2003 to 2007). During eras 1 and 2, RF outperformed XGBoost and GBoost, whereas during era 3 no model dominated. Performance was better during eras 1 and 2 as opposed to the pre-L band era 3. Improvements across all eras, regions, and models realized from downscaling included an increase in correlation from 0.03 to 0.42 and a decrease in *ubRMSE* from -0.0005 to -0.0118 m³/m³. This study demonstrates the feasibility of SMERGE downscaling opening the prospect for the development of a long-term RZSM dataset at a more desirable field-scale resolution with the potential to support diverse hydrometeorological and agricultural applications.



Citation: Tobin, K.; Sanchez, A.; Esparza, D.; Garcia, M.; Ganta, D.; Bennett, M. Machine Learning Downscaling of SoilMERGE in the United States Southern Great Plains. *Remote Sens.* **2023**, *15*, 5120. <https://doi.org/10.3390/rs15215120>

Academic Editor: Dusan Gleich

Received: 6 October 2023

Revised: 18 October 2023

Accepted: 21 October 2023

Published: 26 October 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: rootzone soil moisture; random forest; downscaling

1. Introduction

Satellite-derived soil moisture products, such as Soil Moisture Active Passive (SMAP) and Soil Moisture and Ocean Salinity (SMOS), have revolutionized hydrological and agricultural studies by providing estimates of surface soil moisture (SM) worldwide [1,2]. However, their use is hindered by their coarse spatial resolution (0.25–0.50 degrees) and their lack of ability to penetrate below the surface skin or the top 5 cm layer. While products like the European Space Agency Climate Change Initiative (ESA-CCI) have blended satellite retrievals to enhance observation frequency [3], there remains a pressing need for higher spatial resolution data to support hydrological and ecological applications [4] and accurate drought monitoring [5]. To bridge this gap, advanced machine learning (ML) techniques have emerged as promising tools for downscaling satellite-based soil moisture data. ML offers advantages in handling large and noisy datasets from dynamic and non-linear systems [6–10]. In recent years, studies have successfully employed ML algorithms such as Random Forest (RF), gradient boost decision tree, and eXtreme Gradient Boosting (XGBoost) to enhance spatial resolution and accuracy of surface soil moisture estimates [11–16]. Notable improvements have included increased correlations and decreased unbiased root mean square error [11,14–16]. While surface soil moisture retrievals

are valuable, the scientific community increasingly recognizes the importance of deeper root-zone soil moisture (RZSM) data, as it directly influences agricultural productivity and groundwater interactions. However, directly sensing RZSM remains a challenge. Efforts like NASA's SigNals of Opportunity: P-band Investigation (SNOPI) plan to use the penetrating P-band to retrieve RZSM, but at present, RZSM is often inferred from surface measurements. Two common approaches used to estimate RZSM include the Ensemble Kalman Filter [17] and Exponential Filter [18,19], which, while useful, are not as robust as having direct retrievals of soil moisture.

In this context, the SoilMERGE (SMERGE) product stands as a notable endeavor, covering the continental United States and spanning multiple decades (1979 to 2019) [20]. Like other products, SMERGE faces spatial resolution constraints (0.125 degrees) and is available at a daily time step. SMERGE provides an overall estimate of RZSM between 0 to 40 cm and is based on the fusion of NLDAS Noah-2 land surface model output with surface satellite retrievals from the European Space Agency (ESA) Climate Change Initiative (CCI). In this study, we address the following critical questions related to the downscaling of SMERGE: (1) What is the most optimal ML technique and (2) what downscaling resolution provides the most robust results? To answer these questions, we explore three ML approaches and compare downscaled SMERGE with diverse datasets, including *in situ* measurements and airborne radar estimates of RZSM from the Marena Oklahoma Soil Moisture Active Passive In Situ Testbed (MOISST) site associated with NASA's Airborne Microwave Observatory of Subcanopy and Surface (AirMOSS) campaign [21]. The analysis spans three distinct eras (2016 to 2019, 2012 to 2015, and 2003 to 2007) and aims to achieve finer spatial resolution for SMERGE, thereby enabling more accurate RZSM estimation, which can be used to support diverse applications.

2. Study Areas

Figure 1 provides an overview of this study's focus areas within north-central Oklahoma and south-central Kansas. This is a location where SMERGE exhibited robust performance [20] making it an ideal candidate to explore downscaling. Rectangular areas in Figure 1 reflect zones where SMERGE was downscaled. During era 1 (2016 to 2019), two regions (in red) associated with the United States Department of Energy Atmospheric Radiation Measurement (ARM) observatory were examined (ARM_1_1, ARM_1_2). The naming convention specifies Network_Era_Region where ARM_1_1 represents ARM era 1, region 1. Data from MOISST and Soil moisture Sensing Controller and oPtimal Estimator (SoilSCAPE) comprise era 2 observations (2012 to 2015), which are indicated in blue. Excessive missing data precluded the use of ARM during era 2. During era 3 (2003 to 2007; in black), ARM was divided into four distinct zones (ARM_3_1, ARM_3_2, ARM_3_3, and ARM_3_4). Table 1 indicates the ARM stations utilized in each region. The location of United States Climate Reference Network (USCRN; Stillwater 2W, Stillwater 5WNW) and Soil Climate Analysis Network (SCAN; Abrams) sensors are also indicated in Figure 1.

Table 1. United States Department of Energy Atmospheric Radiation Measurement (ARM) sites by era/region.

Era_Region	<i>In Situ</i> Station
Era 1, Region 1	Anthony, Ashton, Bryon, Lamont-CF1, Maple City, Medford, Newkirk, Pawhuska
Era 1, Region 2	Marshall, Morrison, Omega, Ringwood, Tyron, Waukomis
Era 3, Region 1	Hillsboro, Towanda
Era 3, Region 2	Ashton, Byron, Lamont-CF1
Era 3, Region 3	Elk Falls, Pawhuska, Tyro
Era 3, Region 4	El Reno, Meeker

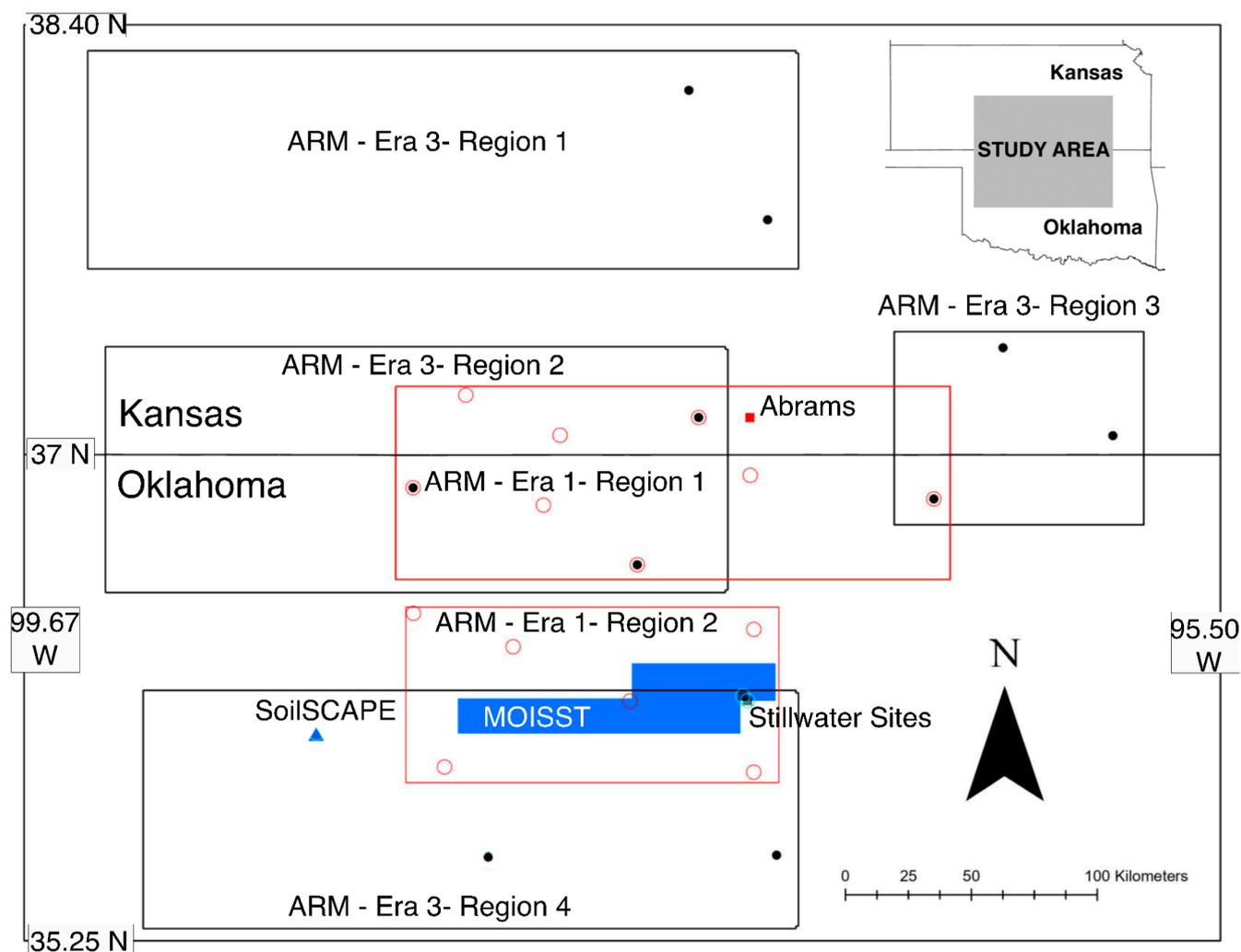


Figure 1. Locality map with era 1 sites in red, era 2 in blue, and era 3 in black. Upper-right inset map shows the study area location in Kansas and Oklahoma. Locations of ARM era 1 and era 3 *in situ* sites are indicated by red open circles and black circles, respectively. The Marena Oklahoma Soil Moisture Active Passive In Situ Testbed (MOISST) site is a solid blue rectangle and the Soil moisture Sensing Controller and oPtimal Estimator (SoilSCAPE) site is indicated by a blue triangle. Other *in situ* data from United States Climate Reference Network (USCRN) Stillwater sites are squares with a blue halo (eras 1 and 2) and the Soil Climate Analysis Network (SCAN) Abrams site with a red square (era 1).

Table 2 provides an overview of the physical characteristics of these focus areas. Clay, silt, and sand values exhibit great variability in all areas. On average, all areas have a similar soil texture with near-equal clay, silt, and sand values representing an overall loamy texture. Elevations generally define a moderate relief within the regions. ARM_1_1, ARM_1_2, ARM_3_3, and MOISST have average elevations of less than 400 m, whereas ARM_3_1, ARM_3_2, ARM_3_4, and SoilSCAPE have higher elevations. In terms of land use and land cover (LULC), herbaceous and cultivated crops dominate all examined areas except for ARM_3_3 (Table 2), which, instead of cultivated crops, has significant hay/pasture and deciduous forest.

Table 3 describes the meteorological conditions within the study areas. Values are generally consistent. Lower spring and fall temperatures were recorded during era 2, as represented by MOISST and SoilSCAPE, when compared against eras 1 and 3. A pronounced drying trend from east to west is present in the study area. This is reflected by the higher warm season (April to October) precipitation values present in ARM_3_3, which

is further east than the other ARM era 3 regions. Additionally, during era 2, MOISST had higher precipitation values than SoilSCAPE, which is located to the west.

Table 2. Physical characteristics of examined regions.

Network_Era_Region	Clay (%)	Silt (%)	Sand (%)	Dominant Land Cover	Secondary Land Cover	Elevation (m)
ARM_1_1	7–50 (30)	1–67 (46)	4–90 (24)	Cultivated Crops	Herbaceous	244–449 (347)
ARM_1_2	8–51 (27)	1–64 (37)	6–90 (36)	Herbaceous	Cultivated Crops	251–443 (332)
MOISST	5–42 (24)	4–65 (37)	13–90 (40)	Herbaceous	Cultivated Crops	267–377 (322)
SoilSCAPE	13–23 (19)	22–58 (45)	19–65 (35)	Herbaceous	None	520–535 (523)
ARM_3_1	3–48 (27)	2–64 (36)	3–95 (37)	Cultivated Crops	Herbaceous	371–694 (517)
ARM_3_2	1–54 (27)	1–67 (40)	7–98 (34)	Herbaceous	Cultivated Crops	280–677 (427)
ARM_3_3	17–57 (34)	18–63 (44)	3–62 (22)	Herbaceous	Hay/PastureDeciduous Forest	202–475 (287)
ARM_3_4	7–50 (24)	1–65 (35)	6–90 (41)	Herbaceous	Cultivated Crops	247–666 (413)

Range of values for soil texture and elevation (%) by region are given with average values indicated in parentheses.

Table 3. Meteorological characteristics of study areas during warm season (April to October).

Network_Era_Region	April Mean Temp (°C)	July Mean Temp (°C)	October Mean Temp (°C)	Warm Season Precipitation (mm)
ARM_1_1	16.8	23.4	19.0	114.7
ARM_1_2	17.3	24.5	19.2	98.0
MOISST	12.9	26.3	15.3	94.6
SoilSCAPE	15.3	27.0	15.8	77.0
ARM_3_1	17.3	24.2	17.8	85.3
ARM_3_2	19.7	25.2	18.2	83.3
ARM_3_3	18.2	25.1	18.4	113.9
ARM_3_4	19.2	25.7	19.6	86.8

3. Methodology

This study's methodologies were implemented in five steps that include: (1) Gathering data; (2) selecting dates to use for validation; (3) executing machine learning downscaling; (4) evaluation of each model run using objective metrics; and (5) performance comparison between different models and spatial resolutions.

3.1. Data Gathering

RZSM data used in this study includes SMERGE version 2.0 (Table 4). The downscaling approach described below incorporated independent variables (Table 4) that are both static (soil texture, elevation, aspect, slope) and dynamic (Normalized Vegetation Difference Index—NDVI, albedo, Leaf Area Index—LAI, surface temperature) following the approach of [13]. Note that a one-month lag was used for NDVI and LAI, consistent with the approach of [20]. The pre-processing of the above datasets was conducted in ArcGIS Pro, where values were extracted into points with a 30 m resolution. The ArcGIS AggregatePoints function was used to aggregate points into grid files at the different resolutions examined in this study.

For evaluation, both *in situ* and AirMOSS data were used. *In situ* data from the ARM, SCAN, SoilSCAPE, and USCRN networks were gathered from the International Soil Moisture Network portal [22], and hourly results were converted into a daily overall estimate of soil moisture, between 0 to 40 cm using a proportional weighting scheme. For example, a common configuration at ARM was to have sensors at 5 cm, 15 cm, 25 cm, and 35 cm, resulting in a corresponding weight for each sensor of 18.75%, 31.25%, 25.00%, and 25.00%, respectively. AirMOSS L2/3 RZSM retrievals using the University of Southern California algorithm were applied at MOISST as described by [21].

Table 4. Data sources used for SoilMERGE (SMERGE) downscaling.

Data Source	Description and Download URL (Accessed on 30 June 2023)
Static Variables	
Elevation	USGS Elevation Products (3DEP), 1/3 arc-sec DEM: TNM Download v2 (nationalmap.gov)
Soil Texture	Gridded National Soil Survey Geographic Database (gNATSGO), the ratio of sand, silt, and clay (Spatial Resolution = 30 m): https://www.nrcs.usda.gov/resources/data-and-reports/gridded-national-soil-survey-geographic-database-gnatsgo
Dynamic Variables	
SMERGE	Smerge-Noah-CCI root zone soil moisture 0-40 cm L4 daily 0.125 × 0.125 degree V2.0 (SMERGE_RZSM0_40CM): https://www.tamui.edu/cees/smerge/data.shtml
Albedo	MCD15A3H v061 MODIS/Terra+Aqua Leaf Area Index/FPAR 4-Day L4 Global 500 m SIN Grid: https://lpdaac.usgs.gov/products/mcd43a3v006/
LAI	MCD15A3H v061 MODIS/Terra+Aqua Leaf Area Index/FPAR 4-Day L4 Global 500 m SIN Grid: https://lpdaac.usgs.gov/products/mcd15a3hv061/
NDVI	Temporally Smoothed Weekly AQUA Collect 6 (C6) Moderate Resolution Imaging Spectroradiometer (MODIS) Normalized Difference Vegetation Index (NDVI) at 250 m: Remote Sensing Phenology CONUS 250 m Smoothed NDVI (usgs.gov)
Temperature	Daily mean temperature, calculated as (tmax + tmin)/2 (Spatial Resolution = 4 km): https://ftp.prism.oregonstate.edu/daily/tmean/

3.2. Date Selection and Data Organization

This study followed the approach of [13], which indicated that spatial variation is more important than temporal variation. With a focus on the warm season (April to October) dates were selected that were separated from each other by at least seven days. *In situ* data were also screened using the methods of [22], and dates with anomalous spikes and plateaus were rejected. For all *in situ* datasets, if there were fewer than ten available dates with valid measurements, then that site was not utilized. Also, if a site had multiple sensors, the sensor with the highest correlation with baseline or default version of SMERGE was used for analysis. *In situ* datasets with a correlation < 0.5 with default SMERGE were rejected.

At ARM, era 1 (2016 to 2019) and era 3 (2003 to 2007) were examined. Era 1 was divided into two regions, and era 3 consisted of four regions (Figure 1). At the ARM site, a total of 47 and 52 dates were examined during eras 1 and 3, respectively (Table 5). At the SoilSCAPE site, *in situ* estimates of RZSM were within a field scale cluster of sensors or nodes that spanned from 2012 to 2015 (era 2). This study focused on a small site near Canton, Oklahoma between latitude 36.000 to 36.003°N and longitude 98.628 to 98.633°W (Figure 1). Daily soil moisture measurements were obtained at depths of 4 cm, 13 cm, and 30 or 40 cm. These values were weighted to determine an overall estimate of RZSM in the top 40 cm. Seventy-nine dates were used for SMERGE downscaling (Table 5) at SoilSCAPE. Note that *in situ* data are missing for some of the selected dates. This was performed intentionally to provide more data to support ML analysis. Basically, the dataset used at SoilSCAPE was inflated temporally to compensate for this site's small spatial footprint.

The AirMOSS project (2012 to 2015; era 2) directly sensed RZSM estimates using the P-band radar, which operates at a low frequency (420–440 MHz). RSZM is estimated up to a depth of 40 cm. This study focuses on the MOISST site. This site is unique in that it affords a continuous comparison with the downscaled SMERGE estimate. This is unlike the discrete *in situ* measurements from ARM and SoilSCAPE. Observations focused on the 13 SMERGE grids that had greater than 75% coverage of AirMOSS retrievals [23] (Figure 2a) and span latitude 36.000 to 36.250°N and longitude 97.000 to 98.125°W (Figure 1). The 22 dates with acceptable AirMOSS retrievals from MOISST are listed in Table 5.

Table 5. Dates with acceptable values for comparison.

Network_Era	Dates
ARM_1	20160401, 20160421, 20160505, 20160524, 20160715, 20160729, 20160812, 20160826, 20160909, 20160923, 20161007, 20161021, 20170401, 20170415, 20170429, 20170513, 20170527, 20170610, 20170624, 20170708, 20170722, 20170805, 20170819, 20170902, 20170916, 20170930, 20171014, 20171028, 20180401, 20180415, 20180429, 20180513, 20180527, 20180610, 20180624, 20180708, 20180722, 20180805, 20180819, 20180902, 20180916, 20180930, 20181014, 20181028, 20190401, 20190415, 20190429
ARM_3	20030401, 20030415, 20030429, 20030513, 20030527, 20030610, 20030624, 20030731, 20030814, 20030828, 20030910, 20031023, 20040401, 20040415, 20040430, 20040514, 20040528, 20040617, 20040701, 20040715, 20040729, 20040812, 20040826, 20040909, 20040923, 20041007, 20041021, 20050401, 20050415, 20050429, 20050513, 20050527, 20050610, 20050628, 20050713, 20050727, 20050822, 20050905, 20050920, 20051004, 20051018, 20060401, 20060415, 20060707, 20060806, 20060826, 20060909, 20060923, 20061007, 20061021, 20070620, 20070813
MOISST_2	20121024, 20121027, 20121030, 20130617, 20130716, 20130719, 20130723, 20130927, 20140416, 20140418, 20140424, 20140708, 20140711, 20140715, 20141014, 20141017, 20141021, 20150416, 20150420, 20150807, 20150811, 20150814
SoilSCAPE_2	20120421, 20120601, 20120608, 20120615, 20120628, 20120705, 20120712, 20120719, 20120726, 20120802, 20120817, 20120824, 20120831, 20120907, 20120922, 20120929, 20121006, 20121013, 20121020, 20121030, 20121106, 20121113, 20121120, 20121126, 20130407, 20130414, 2013042, 20130428, 20130505, 20130512, 20130519, 20130530, 20130606, 20130613, 20130626, 20130703, 20130710, 20130717, 20130724, 2013080, 20130810, 20131006, 20131013, 20131020, 20131110, 20131117, 20131124, 20141027, 20150401, 20150408, 20150415, 20150422, 20150429, 20150506, 20150513, 20150520, 20150527, 20150603, 20150610, 20150617, 20150624, 20150701, 20150708, 20150715, 20150722, 20150729, 20150805, 20150812, 20150819, 20150826, 20150902, 20150909, 20150916, 20150923, 20150930, 20151007, 20151014, 20151021, 20151028

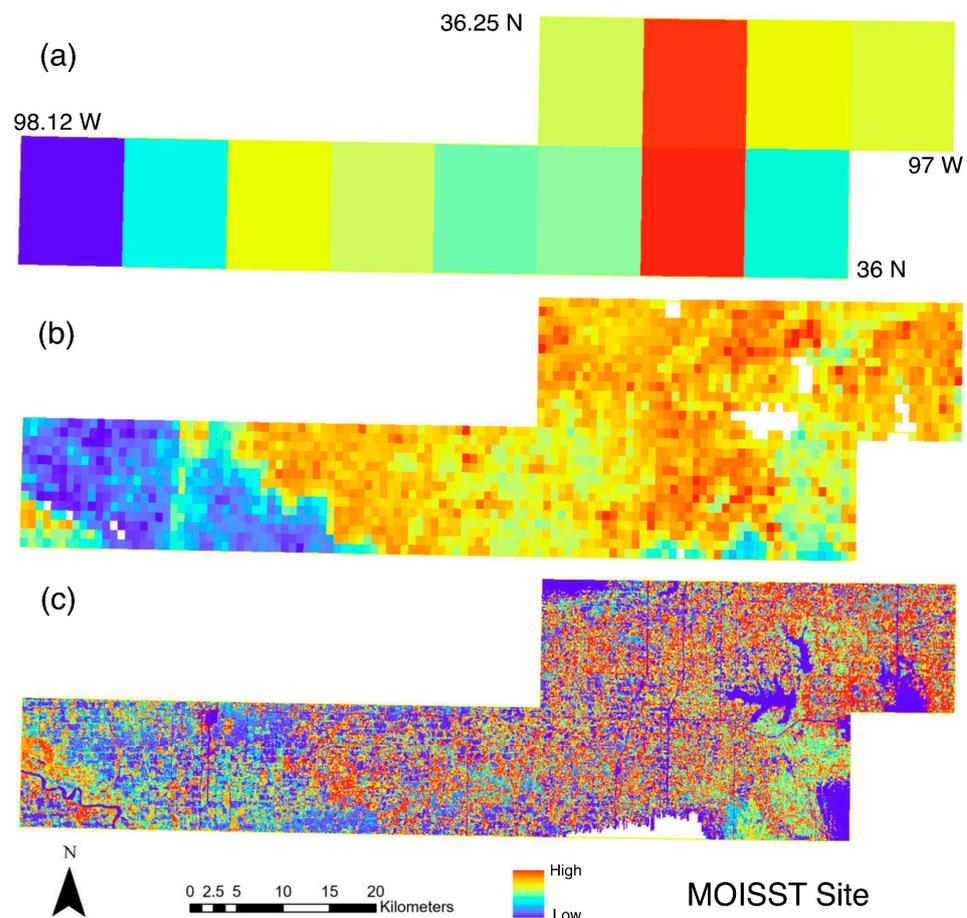


Figure 2. Comparison of (a) Default SMERGE, (b) Downscaled SMERGE (1000 m resolution), and (c) AirMOSS (90 m resolution) for the MOISST site on 24 October 2012. Bounding coordinates are provided in (a).

During eras 1 and 2, in situ data from the Stillwater 2W and 5 WNW stations from USCRN and Abrams from SCAN were used for additional evaluation (Figure 1). Daily soil moisture measurements were obtained at depths of 5 cm, 10 cm, 20 cm, and 50 cm and were used to estimate RZSM between the surface and 40 cm depth using a proportional weighting scheme.

3.3. Machine Learning Implementation

Random Forest (RF) served as our baseline machine learning algorithm, which has been used to down-scale soil moisture estimates in many studies e.g., [12,13,24–27]. We also executed eXtreme Gradient Boosting (XGBoost) [12] and Gradient Boost (GBoost) [12]. The dataset was randomly split (70% training and 30% testing) using the SKLearning `train_test_split` function, and rows that corresponded to ARM and SoilSCAPE *in situ* data sites were removed from the training set and inserted into the testing set for later data verification purposes. Experimentation between 400 to 1400 m for ARM, 400 to 3000 m for MOISST, and 30 to 100 m for SoilSCAPE was executed to determine the optimum spatial resolution for each era/region. The average independent values within grids at varying resolutions were obtained with zonal statistics in ArcGIS Pro. To implement the down-scaling of SMERGE, we used TensorFlow Decision Forests' Random Forest implantation, Distributed (Deep) Machine Learning Community (DMLC)'s XGBoost, and Sklearing's Gradient Boost (GBoost); all of these were set to run as regressors. The hypertuning parameters/settings used for these ML models are specified in Table 6. Hypertuning for three models was conducted via iteration method, where a range of tuning values were given and ideal parameters were found over hundreds of runs. For the TensorFlow Random Forest model, the tuner helper function was included in the iterations.

Table 6. Machine learning algorithm settings/hypertuning parameters.

Random Forest (RF)
<ul style="list-style-type: none"> • Tuner = forest.tuner.RandomSearch(num_trials = 135, use_predefined_hps = True) • Winner_take_all = True • Categorical_algorithm = 'CART' • Honest = True • Honest_fixed_separation = True • Honest_ratio_leaf_examples = 0.75 • Bootstrap_size_ratio = 1.05 • Adapt_bootstrap_size_ratio_for_maximum_training_duration = True • Keep_non_leaf_label_distribution = False • Max_depth = 9
eXtreme Gradient Boosting (XGBoost)
<ul style="list-style-type: none"> • N_estimators = 500 • Max_depth = 10 • Tree_method = 'hist'
Gradient Boost (GBoost)
<ul style="list-style-type: none"> • N_estimators = 175 • Max_depth = 10 • Min_samples_split = 4 • Learning_rate = 0.3 • Loss = squared_error

Independent variable sensitivity was examined with metrics customized for each model, and the results of this analysis are summarized in Table 7. Two approaches were used to gauge model sensitivity because of the different model structures. RF sensitivity was gauged using TensorFlow's Inverse Mean Minimum Depth (IMMD). The mean IMMD for

each independent variable was calculated by tracking the depth of the first occurrence of a feature in each tree in the forest. The depth was averaged for every feature, with the inverse of it taken, resulting in higher IMMD values, reflecting greater sensitivity. For XGBoost and GBoost, independent variable sensitivity was evaluated using the interpretability model tool SHapley Additive exPlanations (SHAP). The reason we opted not to use SHAP for RF is that this tool is not compatible with the TensorFlow Decision Forest library. SHAP explains the individual variable contributions to the predictions made by the model. The higher the SHAP value, the higher the importance of that feature towards the given prediction.

Table 7. Model sensitivity results.

Network_Era _Region	Model Type	Date	Albedo	Clay	Aspect	Temp	Elev	NDVI	Lai	Sand	Silt	Slope
ARM_1_1	RF	H	H	H	M	H	M	L	M	M	L	L
ARM_1_1	XGBoost	H	L	M	M	H	M	L	L	H	M	M
ARM_1_1	GBoost	H	L	M	M	H	M	L	L	H	M	M
ARM_1_2	RF	H	H	M	M	H	M	L	M	L	H	L
ARM_1_2	XGBoost	H	L	M	M	H	M	L	L	H	H	M
ARM_1_2	GBoost	H	L	M	M	H	M	L	L	H	H	M
MOISST	RF	H	H	H	M	M	L	M	L	M	M	L
MOISST	XGBoost	H	M	H	L	H	M	L	L	M	M	M
MOISST	GBoost	H	M	H	M	H	M	L	L	M	M	L
SoilSCAPE	RF	H	H	L	M	M	M	H	H	L	L	M
SoilSCAPE	XGBoost	H	H	M	M	H	M	M	H	L	L	L
SoilSCAPE	GBoost	H	H	M	M	H	M	M	H	L	L	M
ARM_3_1	RF	H	H	H	M	L	H	L	M	L	M	L
ARM_3_1	XGBoost	H	L	M	M	M	H	L	L	M	H	L
ARM_3_1	GBoost	H	L	M	M	M	H	L	L	M	H	M
ARM_3_2	RF	H	H	H	M	L	H	M	L	M	L	L
ARM_3_2	XGBoost	H	L	M	M	M	H	L	L	H	H	M
ARM_3_2	GBoost	H	L	M	M	H	H	L	L	H	M	M
ARM_3_3	RF	H	H	H	M	M	H	M	M	L	L	L
ARM_3_3	XGBoost	H	L	H	M	H	H	L	L	M	M	M
ARM_3_3	GBoost	H	L	M	M	H	H	L	L	M	M	M

H represents high sensitivity, M is medium sensitivity, and L is low sensitivity.

Because of the different tools used to evaluate sensitivity, a direct numerical comparison between IMMD and SHAP was not feasible. Therefore, sensitivity was based on relative ranking. For a given model and spatial resolution, the most sensitive variable was assigned 1, and subsequent variables were assigned a number based on their relative ranking, with the least sensitive variable being 11. Variables that ranked between 1 and 4 were deemed highly sensitive. Rankings between 5 to 8 were designed as moderately sensitive, and those variables ranked greater than 8 were considered to have a low sensitivity. Table 7 shows the average rankings between different spatial resolutions ran for a given model with the sensitivity indicated as high (H), medium (M), or low (L).

3.4. Model Evaluation

A comparison of default and downscaled SMERGE at different spatial resolutions was made against *in situ* data from the ARM and SoilSCAPE sites and SCAN and USCRN sensors. In addition, downscaled SMERGE was evaluated against AirMOSS data from the MOISST site. Standard evaluation metrics (correlation, r ; unbiased root-squared error, $ubRMSE$ [m^3/m^3]) were utilized in these comparisons. Delta r and Delta $ubRMSE$ were used to compare relative performance between default and downscaled versions of SMERGE and are defined as follows:

$$\text{Delta } r = \text{Downscaled SMERGE } r - \text{Default SMERGE } r \quad (1)$$

$$\text{Delta } ubRMSE = \text{Downscaled SMERGE } ubRMSE - \text{Default SMERGE } ubRMSE \quad (2)$$

where Downscaled SMERGE is the value obtained from ML and Default SMERGE is the value obtained from the original 12.5 km resolution product. Delta values represent a specific value from a model type and spatial resolution combination (e.g., RF, 400 m) executed within a given network/era/region (e.g., ARM_1_1). Delta r and Delta $ubRMSE$ values provided insight into whether the downscaled model results exhibited improvement (or degradation) compared with the Default SMERGE product. Improvement for r is defined as having Downscaled SMERGE $>$ Default SMERGE and for $ubRMSE$ Default SMERGE $>$ Downscaled SMERGE.

3.5. Comparisons within a Region/Era

To facilitate comparison across the different models and spatial resolutions examined, an objective metric was derived that combined correlation and $ubRMSE$ metrics. The objective metric varies between zero and one and is defined as indicated below:

$$\text{If Delta } r \leq 0 \text{ Then } a = 0 \quad (3)$$

$$\text{If Delta } r > 0 \text{ Then } a = [\text{Delta } r / \text{Maximum Delta } r] \times 0.5 \quad (4)$$

$$\text{If Delta } ubRMSE \geq 0 \text{ Then } b = 0 \quad (5)$$

$$\text{If Delta } ubRMSE < 0 \text{ Then } b = [\text{Delta } ubRMSE / \text{Minimum Delta } ubRMSE] \times 0.5 \quad (6)$$

$$\text{Objective Metric} = a + b \quad (7)$$

where maximum Delta r and minimum Delta $ubRMSE$ are the maximum and minimum values respectively obtained within an era/region, where a is the component of the objective metric derived from Delta r , and b is the component of the objective metric derived from Delta $ubRMSE$. The objective metric value ranges between zero and one. A value of one indicates that a model type and spatial resolution combination had the best possible r and $ubRMSE$ values within an era/region. Zero indicates that downscaling yielded no improvement based on both r and $ubRMSE$ metrics. Models with an objective metric between 0 to 0.8 were described as slightly improved. Objective metrics between 0.8 to 1.0 were designated as high performing. If a downscaled model had improvement in only one metric, then this model was deemed as non-improved. This represents a model that recorded improvement in r but not $ubRMSE$ or vice versa.

4. Results

Model sensitivity results are summarized in Table 7. The date and aspect independent variables generally have consistently high and moderate sensitivity, respectively. Elevation mostly exhibits moderate sensitivity, except for ARM era 3, where this variable has a high sensitivity. Albedo and NDVI have a higher sensitivity for RF compared to XGBoost and GBoost models. For albedo, the exception is in SoilSCAPE, which has a high sensitivity for all models. Also, during ARM era 1 and ARM_3_1, NDVI had a uniformly low sensitivity. LAI also generally had a higher sensitivity for RF compared with XGBoost and GBoost models. In SoilSCAPE, LAI sensitivity was uniformly high, and in MOISST and ARM_3_2, low sensitivity was recorded for all models. Conversely, temperature exhibited a lower sensitivity for RF models compared with XGBoost and GBoost, but during ARM era 1, a uniformly high sensitivity was noted for all models. Within ARM, slope, sand, and silt mostly had a higher sensitivity for XGBoost and GBoost models than that recorded by RF. In MOISST and SoilSCAPE, slope had low to moderate sensitivity without a consistent pattern

between downscaling models, whereas sand and silt had uniform moderate sensitivity within MOISST and low sensitivity in SoilSCAPE. Finally, clay does not have a consistent relationship among models and generally has a moderate to high sensitivity. Model sensitivity was also examined as a function of spatial resolution. In general, sensitivity was consistent as a function of spatial resolution. The exception is the RF model at the MOISST site, where silt exhibited a marked increase in sensitivity at coarser spatial resolutions.

Figure 3 indicates the number of models with a specified performance from eras 1 and 2. A total of 12 models were executed from each ARM region (four using each ML approach). At ARM, during era 1, 63% of the executed models resulted in improved performance (Figure 3a,b). Of these, 17% were high performing and included all three ML approaches. MOISST and SoilSCAPE (era 2) outperformed ARM (era 1), which had a total of 18 and 6 models executed, respectively. For MOISST, 100% of the downscaling attempts yielded improvements compared to Default SMERGE (Figure 3c). At MOISST, 44% of the models were high performing with a slight preference for RF over the other model types. Figure 2b shows that the spatial distribution of Downscaled SMERGE is similar to AirMOSS data (Figure 2c) from MOISST. At SoilSCAPE the same was true except for a single attempt, resulting in an 83% improvement rate (Table 8). However, only one RF model (17%) from SoilSCAPE was high performing. Figure 4 summarizes the results from ARM era 3, with only 37% of downscaling attempts yielding improvements. In ARM_3_4 not a single model yielded an improvement, compared with Default SMERGE. ARM_3_1, ARM_3_2, and ARM_3_3 had success rates of 42%, 50%, and 58%, respectively. Overall, high performing models from ARM era 3 was only 10%.

Table 8. SoilSCAPE results.

Model Type	Resolution m	Default SMERGE r	Downscaled SMERGE r	Default SMERGE ubRMSE	Downscaled SMERGE ubRMSE	Objective Metric
RF	100	0.4805	0.5217	0.1127	0.1122	0.9061
RF	30	0.4662	0.5169	0.1210	0.1207	0.7738
XGBoost	100	0.4805	0.4837	0.1127	0.1126	0.1200
XGBoost	30	0.4662	0.4665	0.1210	0.1210	0.0188
GBoost	100	0.4805	0.4809	0.1127	0.1127	0.0137
GBoost	30	0.4662	0.4662	0.1210	0.1210	0

Best model results are indicated in bold.

Figure 5 and Table A1 illustrates ARM era 1 results as a function of spatial resolution. In ARM_1_1, RF outperformed the other ML approaches. Optimal performance for RF was recorded at 400 m with declining objective metrics obtained at coarser spatial resolutions. XGBoost and GBoost yielded inconsistent results, and these approaches recorded non-improvement at 700 and 1000 m. When compared against the SCAN Abrams site, none of the models yielded an improvement. Conversely, all RF models were non-improved in ARM_1_2. XGBoost and GBoost exhibited erratic behavior as a function of spatial resolution with best results obtained by XGBoost at a 400m resolution. Interestingly, comparison with USCRN data from the Stillwater sites within ARM_1_2 yielded the opposite with improvement realized with only RF and not XGBoost and GBoost models.

Figure 6 and Table A2 depict the results from MOISST (Era 2). RF recorded the highest objective metric at a 400 m spatial resolution. Examination of an *in situ* sensor from the USCRN Stillwater 5 WNW site also had RF outperforming the other methods. RF exhibited a different trend in terms of performance as a function spatial resolution compared against XGBoost and GBoost. Maximum objective metrics for RF were noted at 400 and 3000 m with lessened performance at the resolutions between 700 to 2000 m. Conversely, XGBoost and GBoost had maximum objective metrics around 1400 to 2000 m with lower values at other resolutions. SoilSCAPE (Table 8), also from era 2, recorded a similar preference for RF. It is noteworthy to indicate that the optimal spatial resolution for all era 1 and 2 regions was ≤ 700 m.

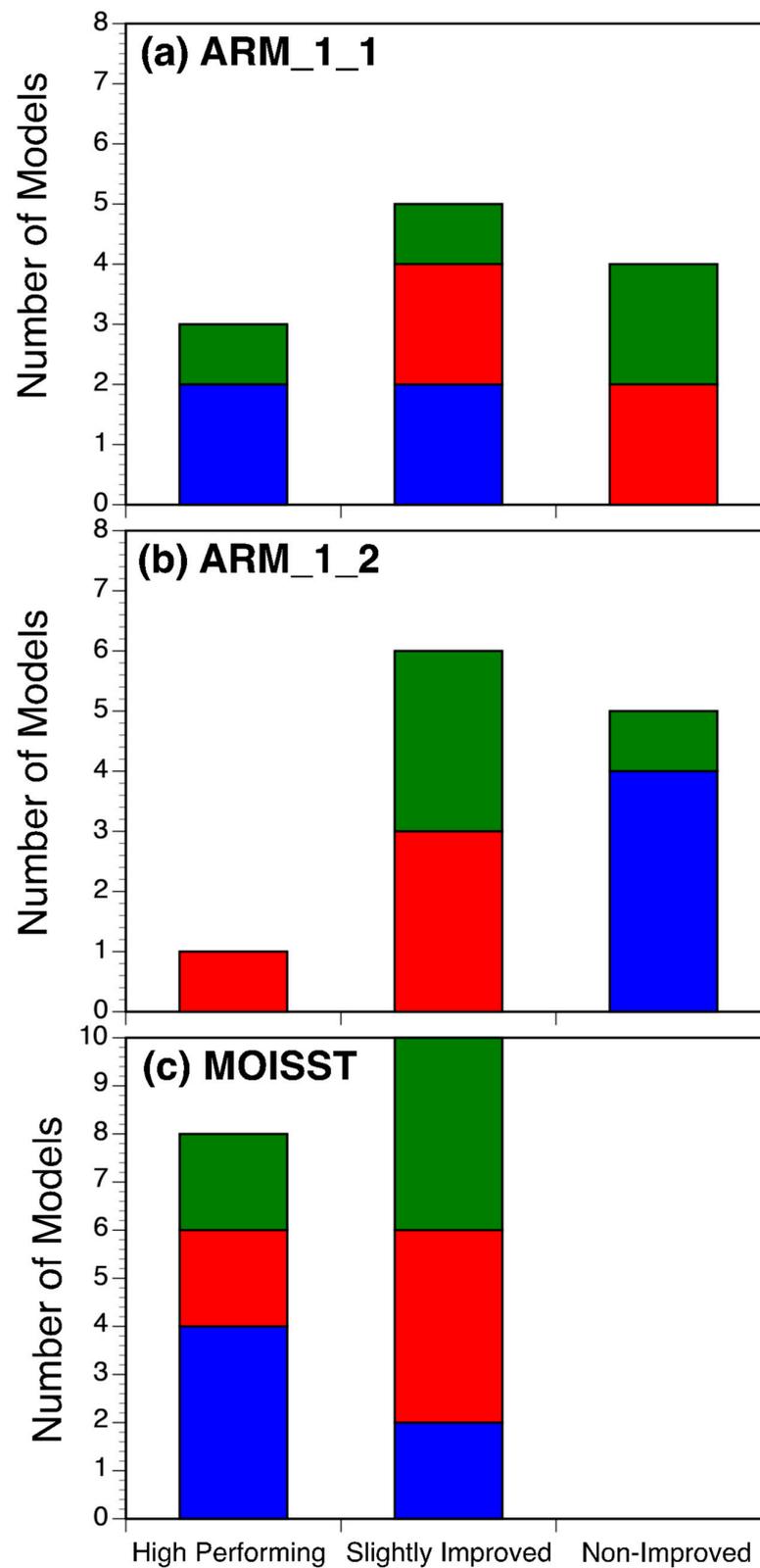


Figure 3. Model performance in eras 1 and 2 regions examined with (a) ARM_1_1, (b) ARM_1_2, and (c) MOISST. Random Forest (RF) is blue, eXtreme Gradient Boosting (XGBoost) is red, and Gradient Boost (GBoost) is green.

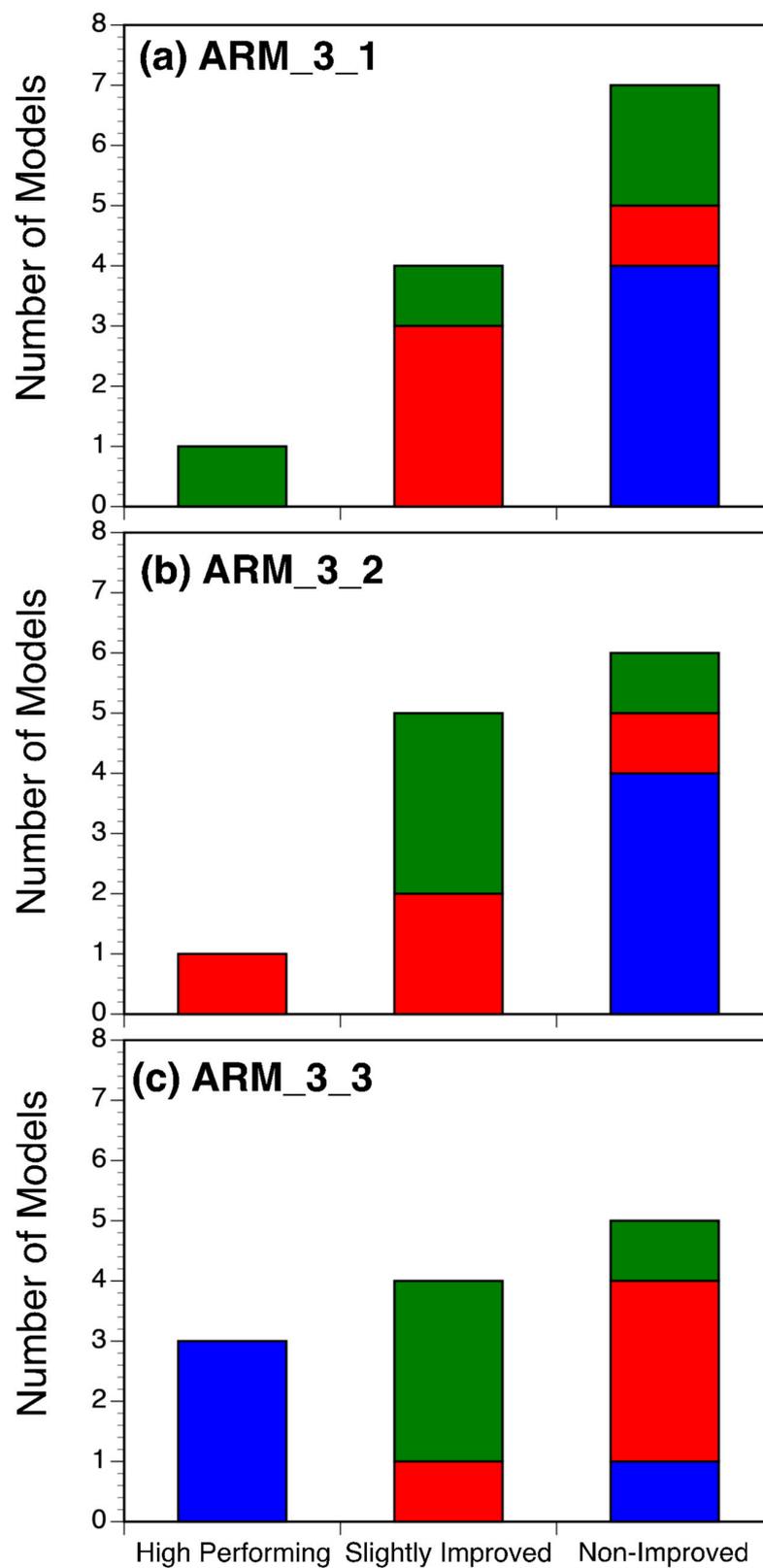


Figure 4. Model performance in era 3 regions examined with (a) ARM_3_1, (b) ARM_3_2, and (c) ARM_3_3. RF is blue, XGBoost is red, and GBoost is green.

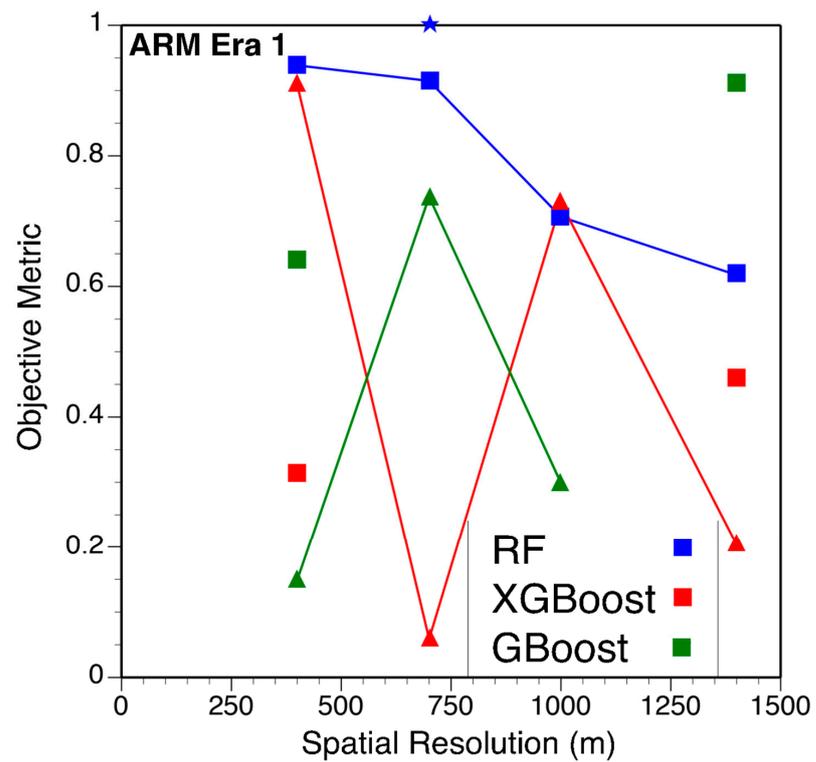


Figure 5. Objective metric as a function of spatial resolution for ARM Era 1. ARM_1_1 represented by squares and ARM_1_2 by triangles. *In situ* comparison with the USCRN Stillwater sites is indicated as a star. Only objective metrics for improved models were plotted. Blue indicates RF, red indicates XGBoost, and green indicates GBoost (also see legend).

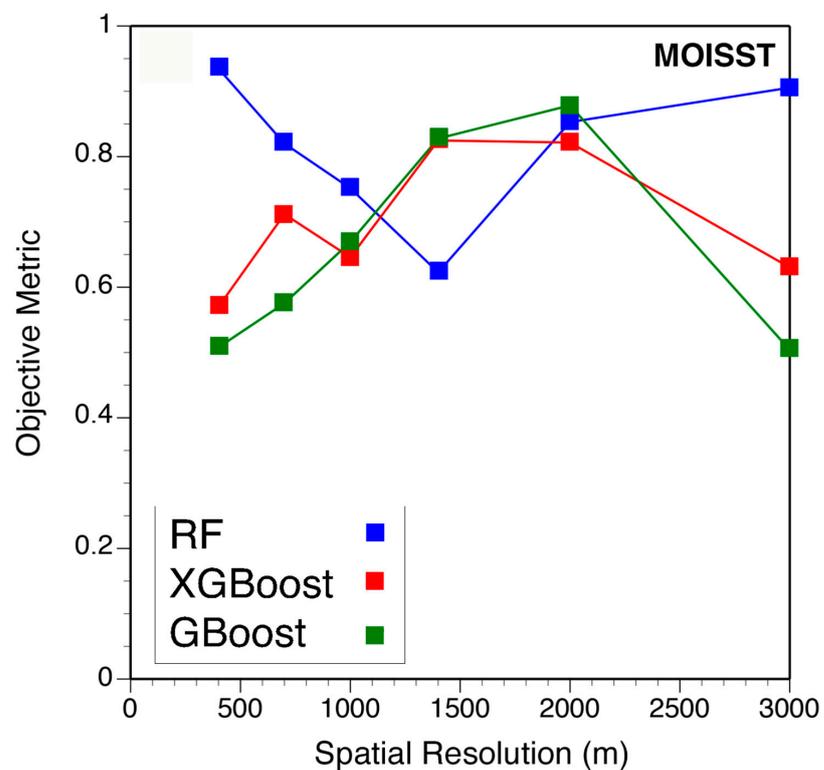


Figure 6. Objective metric as a function of spatial resolution for the MOISST. Colors are as indicated in Figure 5 (also see legend).

During ARM era 3, improvements in downscaled results were noted in ARM_3_1, ARM_3_2, and ARM_3_3 (Figure 7; Table A3) and the model that yielded the highest objective metric varied. In ARM_3_1, GBoost had the highest objective metric at a spatial resolution of 700 m and in ARM_3_2 XGBoost, at a 1400 m resolution, yielded the best results. Note that ARM_3_1 and ARM_3_2 had only one high performing model per region (Figure 4). ARM_3_3 differed in that RF yielded three high performing models, with the highest objective metric at 1000 m. For ARM_3_4, no downscaling results yielded an improvement over Default SMERGE. In general, for era 3 the optimal spatial resolution was ≥ 700 m. Also, in all three regions the general trend was for the objective metric to increase as spatial resolutions becomes coarser.

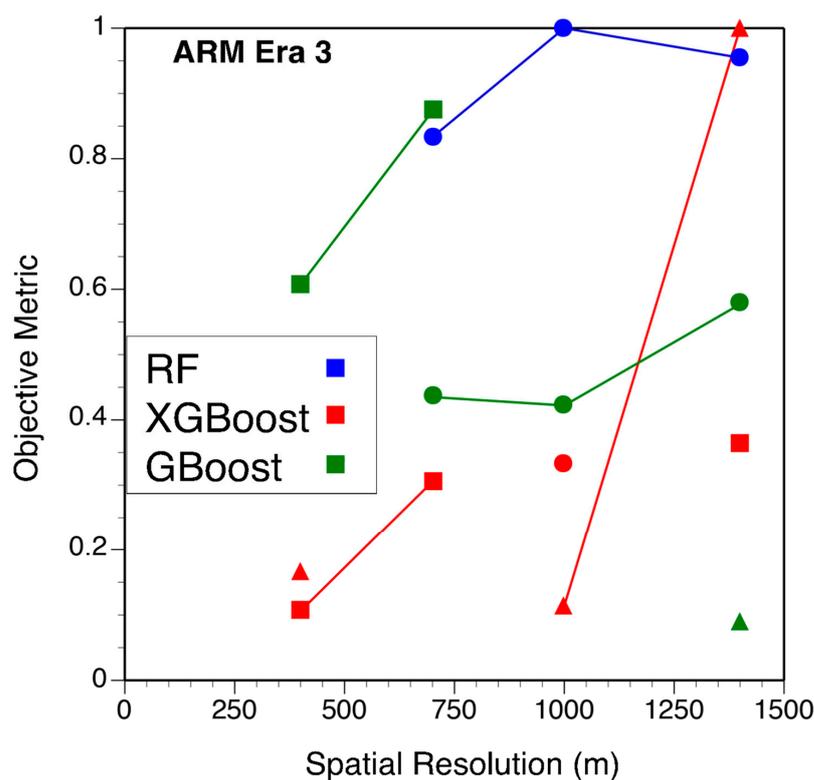


Figure 7. Objective metric as a function of spatial resolution for the ARM Era 3. Colors are as indicated in Figure 5 (also see legend). ARM_3_1 represented by squares, ARM_3_2 by triangles, and ARM_3_3 by circles.

5. Discussion

Overall, improvement in downscaled SMERGE across all eras, regions, and models ranged from 0.03 to 0.42 for Delta r and -0.0005 to -0.0118 m^3/m^3 for Delta $ubRMSE$ (Tables 8 and A1, Tables A2 and A3). These results are comparable to results from previous ML downscaling efforts. Ref. [15] downscaled SMAP on the Iberian Peninsula using RF, yielding an increase in correlation of 0.31 and a decrease in $ubRMSE$ of 0.026 m^3/m^3 compared to SMAP at its native resolution. Reference [14], from a study in China, yielded an increase of 0.1 to 0.2 for correlation and a decrease of 0.01 m^3/m^3 for $ubRMSE$ compared with original land surface model data. Reference [11] also noted an improvement in correlation by 0.1 for RF generated soil moisture, compared with the default ESA-CCI product. [16] used RF to develop the 1 km resolution ChinaCropSM estimate of soil moisture that had a correlation value of 0.93 compared to a 0.35 for ESA-CCI. $ubRMSE$ also recorded a dramatic improvement, from 0.093 to 0.033 m^3/m^3 in this product.

There are clear differences in performance between eras facilitated by comparison of ARM data between eras 1 and 3. During era 1, more than half (15 out of 24) of the downscaling models yielded improvements (Figure 3a,b). Conversely, during era 3 only

slightly more than a third of the models (16 out of 48) were improved (Figure 4). A similar trend was noted for high-performing models. For era 1, 17% of downscaled models were high performing compared with the only 10% high performing models for era 3. During era 1, deeper penetrating L-band retrievals, with increased accuracy, were available from the SMOS [2] and SMAP [28] missions and included in the ESA-CCI product that forms the backbone of SMERGE. Another consideration is the completeness of the independent variables and validation datasets. Missing albedo and LAI data is present during era 3 (Table 9). Note that spatial averaging minimized the impact of these missing datasets at coarser resolutions (1000 to 1400 m). Also, albedo and LAI, have a higher sensitivity within RF models compared with XGBoost and GBoost (Table 7). This could explain the relative underperformance of RF during ARM era 3 compared with other models. RF had a success rate of 19% during era 3. XGBoost and GBoost had higher success rates of 44% and 50%, respectively. Another dimension to examine is the increasing interpolation within the SMERGE product during the earlier eras (Table 9). The ESA-CCI product had some missing daily data in the product that was estimated by interpolation within SMERGE, see [17]. During eras 1 and 2 the degree of interpolation is relatively low (1 to 17%) unlike era 3 (24 to 36%). Interpolation can produce uncertainties within SMERGE estimates that get propagated to the downscaled version, possibly contributing to the poor performance of SMERGE downscaling during era 3. Finally, the spatial representativeness of the *in situ* data is different between eras 1 and 3 within the ARM regions (Figure 1). The coverage of *in situ* stations with acceptable data (correlation with Default SMERGE > 0.5) is greater in era 1. Interestingly, the region with the best *in situ* coverage is the small region ARM_3_3, which recorded the best performance of all ARM regions during era 3.

As notable as the above ARM comparisons are, they are still based on sparse *in situ* data, which [29] indicated can be problematic when providing validation for coarse-resolution satellite-based soil moisture products. The spatial variability within a grid, even for a downscaled product, may not be represented by a point *in situ* measurement. A spatial mismatch exists between a grid mean and sparse *in situ* sampling that can increase uncertainty and produce spurious errors within the downscaled product. This sampling issue is negated for the era 2 sites where MOISST RZSM retrievals were collected over a continuous extent during the AirMOSS campaign. Additionally, SoilSCAPE is a small site, less than a square kilometer, with a cluster of 21 sensors. As such, the validation at these sites can be considered a best-case scenario as reflected by an improvement rate of 100% and 83% for MOISST and SoilSCAPE, respectively. Of particular note are the improvements seen at the 100 m resolution at SoilSCAPE, suggesting that under ideal circumstances downscaling to field scale resolutions is feasible.

Spatial resolution trends are different between eras and models. RF at ARM_1_1 and MOISST had a maximum objective metric at 400 m with declining performance at coarser spatial resolutions out to 1400 m (Figures 5 and 6). Conversely, the highest objective metric for XGBoost and GBoost models at MOISST was at 1400 to 2000 m (Figure 6). During ARM era 3, all models recorded increasing performance with coarser spatial resolutions (Figure 7). During this era, incomplete albedo and LAI likely hampered model execution at finer spatial resolutions (400 to 700 m).

This study has provided valuable insights for the future development of a regional downscaled version of SMERGE. This has implications in that SMERGE is a long duration RZSM product that provides a retrospective estimate of this variable unlike SMOS and SMAP that are limited temporally (post-2010). Therefore, this work lays the groundwork for the development of a long-term field-scale estimate that can support diverse user communities. The new downscaled product will focus on the United States Southern Great Plains where the Default SMERGE performance was the best [20]. This work is not intended to provide a comprehensive validation. Instead, its goal was to determine possible temporal coverage, spatial resolution, and ML model to be used to develop the downscaled product. Additional ranked correlation analyses will be applied to fully validate downscaled SMERGE. Inclusion of these techniques here is beyond this work's scope.

Table 9. Reasons for incompleteness in datasets.

Network_Era_Region or Site	Resolution m	Percent Complete	Data Used for Training	Incomplete MOISST/In Situ Data	Missing Albedo/LAI	Percentage SMERGE Interpolated
ARM_1_1	1400	100%	-	0%	0%	10%
ARM_1_1	1000	100%	-	0%	0%	11%
ARM_1_1	700	100%	-	0%	0%	16%
ARM_1_1	400	100%	-	0%	0%	17%
SCAN_1_Abrams	1000	34%	66%	0%	0%	-
SCAN_1_Abrams	700	32%	68%	0%	0%	-
ARM_1_2	1400	100%	-	0%	0%	9%
ARM_1_2	1000	100%	-	0%	0%	9%
ARM_1_2	700	100%	-	0%	0%	11%
ARM_1_2	400	100%	-	0%	0%	14%
USCRN_1_Stillwater Sites	1400	29%	71%	0%	0%	-
USCRN_1_Stillwater Sites	1000	25%	75%	0%	0%	-
USCRN_1_Stillwater Sites	700	31%	69%	0%	0%	-
AirMOSS_2_MOISST	3000	82%	-	18%	0%	1%
AirMOSS_2_MOISST	2000	84%	-	16%	0%	1%
AirMOSS_2_MOISST	1400	83%	-	17%	0%	1%
AirMOSS_2_MOISST	1000	82%	-	18%	0%	1%
AirMOSS_2_MOISST	700	83%	-	17%	0%	1%
AirMOSS_2_MOISST	400	81%	-	19%	0%	1%
USCRN_2_Stillwater 5WNW	700	46%	54%	0%	0%	-
SoilSCAPE_2	100	65%	-	0%	0%	10%
SoilSCAPE_2	30	72%	-	0%	0%	10%
ARM_3_1	1400	100%	-	0%	0%	24%
ARM_3_1	1000	93%	-	0%	7%	24%
ARM_3_1	700	75%	-	0%	25%	25%
ARM_3_1	400	63%	-	0%	37%	25%
ARM_3_2	1400	100%	-	0%	0%	29%
ARM_3_2	1000	92%	-	0%	8%	29%
ARM_3_2	700	77%	-	0%	23%	29%
ARM_3_2	400	77%	-	0%	23%	29%
ARM_3_3	1400	92%	-	7%	1%	36%
ARM_3_3	1000	92%	-	7%	1%	36%
ARM_3_3	700	91%	-	7%	2%	36%
ARM_3_3	400	77%	-	6%	17%	36%

6. Conclusions

This study successfully downscaled SMERGE, focusing on the warm season (April to October) in the Southern Great Plains (Oklahoma and Kansas). More robust SMERGE downscaling results were yielded during eras 1 (2016 to 2019) and 2 (2012 to 2015), where RF produced optimal results at ≤ 700 m. Improvements in the downscaled SMERGE at the 100 m resolution at SoilSCAPE were particularly noteworthy. These results suggest that SMERGE can be successfully downscaled to the field scale with the advent of L-band microwave retrievals after 2010. However, some caution regarding this conclusion is warranted, given the small area of the SoilSCAPE site. During era 3, downscaling efforts were less successful for several reasons. Optimum model and spatial resolution were less consistent across the ARM era 3 regions, but in general, they exceeded 700 m. Results from this study straddles existing (Sentinel 1 & 2) and planned NASA ISRO Synthetic Aperture Radar (NISAR) mission's capabilities. In addition, downscaled RZSM from long duration products like SMERGE can support more robust hydrologic and ecologic modeling and drought monitoring than surface satellite SM estimates.

Author Contributions: Conceptualization, K.T.; methodology, K.T., A.S., D.E. and M.G.; software, A.S. and D.G.; validation, K.T. and A.S.; formal analysis, K.T.; investigation, K.T.; resources, K.T.; data curation, K.T. and A.S.; writing—original draft preparation, K.T.; writing—review and editing, M.B.; visualization, K.T.; supervision, K.T. and A.S.; project administration, K.T.; funding acquisition, K.T. and D.G. All authors have read and agreed to the published version of the manuscript.

Funding: The authors acknowledge support from the United States Department of Energy Research Development and Partnership Pilot (RDPP, award number DE-SC0023067). Support from NASA Climate Indicator and Data Products for Future National Climate Assessments program through award # NNX16AH30G and NSF Geoscience Equipment (Award Number 1636769) is also gratefully acknowledged.

Data Availability Statement: The data presented in this study are available upon request from the corresponding author.

Acknowledgments: The Texas Advanced Computing Center (TACC; <http://www.tacc.utexas.edu>) at The University of Texas at Austin also provided computational resources that have contributed to the research results reported within this paper. The assistance of Franco Zamora (TAMU ARC Writing Consultant) is also greatly appreciated.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Table A1. ARM Era 1 results.

Model Type	Region	Spatial Resolution m	Default SMERGE r	Downscaled SMERGE r	Default SMERGE ubRMSE	Downscaled SMERGE ubRMSE
RF	1	1400	0.5714	0.6907	0.0710	0.0678
RF	1	1000	0.5714	0.6870	0.0710	0.0666
RF	1	700	0.5714	0.7409	0.0710	0.0660
RF	1	400	0.5714	0.7318	0.0710	0.0654
XGBoost	1	1400	0.5714	0.6550	0.0710	0.0685
XGBoost	1	1000	0.5714	0.5243	0.0710	0.0701
XGBoost	1	700	0.5714	0.5249	0.0710	0.0724
XGBoost	1	400	0.5714	0.6298	0.0710	0.0693
GBoost	1	1400	0.5714	0.7100	0.0710	0.0650
GBoost	1	1000	0.5714	0.4850	0.0710	0.0720
GBoost	1	700	0.5714	0.4489	0.0710	0.0736
GBoost	1	400	0.5714	0.6477	0.0710	0.0660
RF	2	1400	0.6037	0.6554	0.0970	0.1049
RF	2	1000	0.6044	0.6966	0.0970	0.1024
RF	2	700	0.6034	0.7145	0.0970	0.1010
RF	2	400	0.6017	0.6979	0.0972	0.1007
XGBoost	2	1400	0.6037	0.6649	0.0970	0.0970
XGBoost	2	1000	0.6044	0.7567	0.0970	0.0929
XGBoost	2	700	0.6034	0.6128	0.0970	0.0965
XGBoost	2	400	0.6017	0.7511	0.0972	0.0897
GBoost	2	1400	0.6037	0.4444	0.0970	0.1071
GBoost	2	1000	0.6044	0.6349	0.0970	0.0934
GBoost	2	700	0.6034	0.6757	0.0970	0.0881
GBoost	2	400	0.6017	0.6181	0.0972	0.0954
In Situ Comparison from USCRN Stillwater Sites						
RF	2	1400	0.6324	0.3496	0.0712	0.0800
RF	2	1000	0.7502	0.5777	0.0684	0.0757
RF	2	700	0.6417	0.6756	0.0681	0.0673
XGBoost	2	1400	0.6324	0.5975	0.0712	0.0724
XGBoost	2	1000	0.7502	0.6692	0.0684	0.0681
XGBoost	2	700	0.6417	0.4898	0.0681	0.0717
GBoost	2	1400	0.6324	0.4663	0.0712	0.0752
GBoost	2	1000	0.7502	0.6692	0.0684	0.0757
GBoost	2	700	0.6417	0.4713	0.0681	0.0716

Best model results are indicated in bold.

Table A2. MOISST results.

Model Type	Resolution m	Default SMERGE r	Downscaled SMERGE r	Default SMERGE ubRMSE	Downscaled SMERGE ubRMSE
RF	3000	0.4806	0.8193	0.0572	0.0508
RF	2000	0.4375	0.7608	0.0627	0.0568
RF	1400	0.4330	0.7161	0.0656	0.0620
RF	1000	0.4113	0.7099	0.0698	0.0648
RF	700	0.3800	0.7304	0.0755	0.0704
RF	400	0.3321	0.7497	0.0860	0.0805
XGBoost	3000	0.4806	0.8092	0.0572	0.0542
XGBoost	2000	0.4375	0.8082	0.0627	0.0579
XGBoost	1400	0.4330	0.8337	0.0656	0.0612
XGBoost	1000	0.4113	0.7745	0.0698	0.0671
XGBoost	700	0.3800	0.7461	0.0755	0.0720
XGBoost	400	0.3321	0.7239	0.0860	0.0847
GBoost	3000	0.4806	0.8072	0.0572	0.0557
GBoost	2000	0.4375	0.7997	0.0627	0.0571
GBoost	1400	0.4330	0.8460	0.0656	0.0614
GBoost	1000	0.4113	0.7865	0.0698	0.0670
GBoost	700	0.3800	0.7552	0.0755	0.0739
GBoost	400	0.3321	0.7314	0.0860	0.0856
In Situ Comparison from USCRN Stillwater 5WNW					
RF	700	0.5650	0.6189	0.0360	0.0347
XGBoost	700	0.5650	0.2316	0.0360	0.0460
GBoost	700	0.5650	0.5774	0.0360	0.0368

Best model results are indicated in bold.

Table A3. ARM Era 3 results.

Model Type	Region	Resolution m	Default SMERGE r	Downscaled SMERGE r	Default SMERGE ubRMSE	Downscaled SMERGE ubRMSE
RF	1	1400	0.6109	0.4994	0.0304	0.0207
RF	1	1000	0.6087	0.5889	0.0312	0.0194
RF	1	700	0.6235	0.5781	0.0293	0.0201
RF	1	400	0.6463	0.6299	0.0288	0.0197
XGBoost	1	1400	0.6109	0.6374	0.0304	0.0264
XGBoost	1	1000	0.6087	0.5476	0.0312	0.0287
XGBoost	1	700	0.6235	0.6406	0.0293	0.0251
XGBoost	1	400	0.6463	0.6479	0.0288	0.0265
GBoost	1	1400	0.6109	0.2918	0.0304	0.0381
GBoost	1	1000	0.6087	0.5956	0.0312	0.0344
GBoost	1	700	0.6235	0.6922	0.0293	0.0205
GBoost	1	400	0.6463	0.7016	0.0288	0.0239
RF	2	1400	0.7032	0.7255	0.0271	0.0274
RF	2	1000	0.7115	0.6617	0.0265	0.0289
RF	2	700	0.7076	0.4798	0.0277	0.0344
RF	2	400	0.7097	0.5728	0.0266	0.0310
XGBoost	2	1400	0.7032	0.8047	0.0271	0.0246
XGBoost	2	1000	0.7115	0.7215	0.0265	0.0262
XGBoost	2	700	0.7076	0.6519	0.0277	0.0299
XGBoost	2	400	0.7097	0.7212	0.0266	0.0261
GBoost	2	1400	0.7032	0.7180	0.0271	0.0270
GBoost	2	1000	0.7115	0.4374	0.0265	0.0374
GBoost	2	700	0.7076	0.6381	0.0277	0.0380
GBoost	2	400	0.7097	0.2011	0.0266	0.0550
RF	3	1400	0.4895	0.7029	0.0348	0.0255
RF	3	1000	0.4895	0.7120	0.0348	0.0251
RF	3	700	0.4880	0.6605	0.0350	0.0263
RF	3	400	0.5302	0.4529	0.0346	0.0332
XGBoost	3	1400	0.4895	0.4879	0.0348	0.0333
XGBoost	3	1000	0.4895	0.5606	0.0348	0.0314
XGBoost	3	700	0.4880	0.4467	0.0350	0.0349
XGBoost	3	400	0.5302	0.2835	0.0346	0.0382
GBoost	3	1400	0.4895	0.5945	0.0348	0.0282
GBoost	3	1000	0.4895	0.6025	0.0348	0.0315
GBoost	3	700	0.4880	0.6069	0.0350	0.0317
GBoost	3	400	0.5302	0.4645	0.0346	0.0357

Best model results are indicated in bold.

References

- Entekhabi, D.; Njoku, E.G.; O'Neill, P.E.; Kellogg, K.H.; Crow, W.T.; Edelstein, W.N.; Entin, J.K.; Goodman, S.D.; Jackson, T.J.; Johnson, J.; et al. The Soil Moisture Active Passive (SMAP) mission. *Proc. IEEE* **2010**, *98*, 704–716. [[CrossRef](#)]
- Kerr, Y.H.; Waldteufel, P.; Wigneron, J.P.; Maerinuzzi, J.M.; Font, J.; Berger, M. Soil moisture retrieval from space: The Soil Moisture and Ocean Salinity (SMOS) mission. *IEEE Trans. Geosci. Remote Sens.* **2001**, *39*, 1729–1735. [[CrossRef](#)]
- Liu, Y.Y.; Dorigo, W.A.; Parinussa, R.M.; de Jeu, R.A.M.; Wagner, W.; McCabe, M.F.; Evans, J.P.; van Dijk, A.I.J.M. Trend-preserving blending of passive and active microwave soil moisture retrievals. *Remote Sens. Environ.* **2012**, *123*, 280–297. [[CrossRef](#)]
- O'donnell, M.S.; Manier, D.J. Spatial Estimates of Soil Moisture for Understanding Ecological Potential and Risk: A Case Study for Arid and Semi-Arid Ecosystems. *Land* **2022**, *11*, 1856. [[CrossRef](#)]
- Bhardwaj, J.; Kuleshov, Y.; Chua, Z.-W.; Watkins, A.B.; Choy, S.; Sun, Q. Evaluating Satellite Soil Moisture Datasets for Drought Monitoring in Australia and the South-West Pacific. *Remote Sens.* **2022**, *14*, 3971. [[CrossRef](#)]
- Peng, J.; Loew, A.; Merlin, O.; Verhoest, N.E.C. A review of spatial downscaling remotely sensed soil moisture: Downscale satellite-based soil moisture. *Rev. Geophys.* **2017**, *55*, 341–366. [[CrossRef](#)]
- Sabaghy, S.; Walker, J.P.; Renzullo, L.J.; Jackson, T.J. Spatially enhanced passive microwave derived soil moisture: Capabilities and opportunities. *Remote Sens. Environ.* **2018**, *209*, 551–580. [[CrossRef](#)]
- Srivastava, P.; Han, D.; Ramirez, M.R.; Islan, T. Machine learning techniques for downscaling SMOS satellite soil moisture using MODIS land surface temperature for hydrologic applications. *Water Resour. Manag.* **2013**, *27*, 3127–3144. [[CrossRef](#)]
- Im, J.; Park, S.; Rhee, J.; Balk, J.; Choi, M. Downscaling of AMSR-E soil moisture with MODIS products using machine learning approaches. *Environ. Earth Sci.* **2016**, *75*, 1120. [[CrossRef](#)]
- Liu, Y.; Yang, Y.; Jing, W.; Yue, X. Comparison of different machine learning approaches for monthly satellite-based soil moisture downscaling over Northeast China. *Remote Sens.* **2018**, *10*, 31. [[CrossRef](#)]
- Zhang, L.; Zeng, Y.; Zhuang, R.; Szabó, B.; Manfreda, S.; Han, Q.; Su, Z. In Situ Observation-Constrained Global Surface Soil Moisture Using Random Forest Model. *Remote Sens.* **2022**, *13*, 4893. [[CrossRef](#)]
- Liu, Y.; Xia, X.; Yao, L.; Jing, W.; Zhou, C.; Huang, W.; Li, Y.; Yang, J. Downscaling satellite retrieved soil moisture using Regression tree-based machine learning algorithms over Southwest France. *Earth Space Sci.* **2020**, *7*, e2020EA001267. [[CrossRef](#)]
- Zappa, L.; Forkel, M.; Xaver, A.; Dorigo, W. Deriving field scale soil moisture from satellite observations and ground measurements in a hilly agricultural region. *Remote Sens.* **2019**, *11*, 2596. [[CrossRef](#)]
- Abowarda, A.S.; Bai, L.; Zhang, C.; Long, D.; Li, X.; Huang, Q.; Sun, Z. Generating surface soil moisture at 30 m spatial resolution using both data fusion and machine learning toward better water resources management at the field scale. *Remote Sens. Environ.* **2021**, *255*, 112301. [[CrossRef](#)]
- Zhao, W.; Sánchez, N.; Lu, H.; Li, A. A spatial downscaling approach for the SMAP passive surface soil moisture product using random forest regression. *J. Hydrol.* **2018**, *563*, 1009–1024. [[CrossRef](#)]
- Cheng, F.; Zhang, Z.; Zhuang, H.; Han, J.; Luo, Y.; Cao, J.; Zhang, L.; Zhang, J.; Xu, J.; Tao, F. ChinaCropSM1 km: A fine 1 km daily soil moisture dataset for dryland wheat and maize across China during 1993–2018. *Earth Syst. Sci. Data* **2023**, *15*, 395–409. [[CrossRef](#)]
- Reichle, R.H.; Crow, W.T.; Koster, R.D.; Sharif, H.O.; Mahanama, S.P.P. Contribution of soil moisture retrievals to land data assimilation products. *Geophys. Res. Lett.* **2008**, *35*, L01404. [[CrossRef](#)]
- Wagner, W.; Lemoine, G.; Rott, H. A method for estimating soil moisture from ERS scatterometer and soil data. *Remote Sens. Environ.* **1999**, *70*, 191–207. [[CrossRef](#)]
- Albergel, C.; Ruediger, C.; Pellarin, T.; Calvet, J.-C.; Fritz, N.; Froissard, F.; Suquia, D.; Petitpa, A.; Pignatelli, B.; Martin, E. From near-surface to root-zone soil moisture using an exponential filter: An assessment of the method based on in-situ observations and model simulations. *Hydrol. Earth Syst. Sci.* **2008**, *12*, 1323–1337. [[CrossRef](#)]
- Tobin, K.J.; Crow, W.T.; Dong, J.; Bennett, M.E. Validation of a new soil moisture product Soil MERGE or SMERGE. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2019**, *12*, 3351–3365. [[CrossRef](#)]
- Tabatabaenejad, A.; Burgin, M.; Duan, X.; Moghaddam, M. P-Band radar retrieval of subsurface soil moisture profile as a second-order polynomial: First AirMOSS results. *IEEE Trans. Geosci. Remote Sens.* **2015**, *53*, 645–658. [[CrossRef](#)]
- Dorigo, W.A.; Xavier, A.; Vreugdenhil, M.; Gruber, A.; Hegyiová, A.; Sanchis-Dufau, A.D.; Zamojski, D.; Cordes, C.; Wagner, W.; Drusch, M. Global automated quality control of in situ soil moisture data from the International Soil Moisture Network. *Vadose Zone J.* **2013**, *12*, 1–21. [[CrossRef](#)]
- Tobin, K.J.; Crow, W.T.; Bennett, M.E. Root zone soil moisture comparisons: AirMOSS, SMERGE, and SMAP. *IEEE Geosci. Remote Sens. Lett.* **2022**, *19*, 1–5. [[CrossRef](#)]
- Jing, W.; Zhang, P.; Zhao, X. Reconstructing monthly ECV global soil moisture with an improved spatial resolution. *Water Resour. Manag.* **2018**, *32*, 2523–2537. [[CrossRef](#)]
- Yan, R.; Bai, R. A new approach for soil moisture downscaling in the presence of seasonal difference. *Remote Sens.* **2020**, *12*, 2818. [[CrossRef](#)]
- Kovačević, J.; Cvijetinović, Ž.; Stančić, N.; Brodić, N.; Mihajlović, D. New downscaling approach using ESA CCI SM products for obtaining high resolution surface soil moisture. *Remote Sens.* **2020**, *12*, 1119. [[CrossRef](#)]
- Xu, Y.; Wang, L.; Ma, Z.; Li, B.; Bartels, R.; Liu, C.; Zhang, X.; Dong, J. Spatially explicit model for statistical downscaling of satellite passive microwave soil moisture. *IEEE Trans. Geosci. Remote Sens.* **2020**, *58*, 1182–1191. [[CrossRef](#)]

28. Reichle, R.H.; De Lannoy, G.J.M.; Liu, Q.; Ardizzone, J.V.; Colliander, A.; Conaty, A.; Crow, W.; Jackson, T.J.; Jones, L.A.; Kimball, J.S.; et al. Assessment of the SMAP Level-4 surface and root-zone soil moisture product using in situ measurements. *J. Hydrometeorol.* **2017**, *18*, 2621–2645. [[CrossRef](#)]
29. Crow, W.T.; Berg, A.A.; Cosh, M.H.; Loew, A.; Mohanty, B.P.; Panciera, R.; de Rosnay, P.; Ryu, D.; Walker, J.P. Upscaling sparse ground-based soil moisture observations for the validation of coarse-resolution satellite soil moisture products. *Rev. Geophys.* **2012**, *50*, RG2002. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.