https://doi.org/10.5194/egusphere-2025-896
Preprint. Discussion started: 7 April 2025

Identifying Important Features for Downscaling Soil Moisture to 1-km in the Contiguous United States

Eshita A. Eva[1*], Zack Leasor[2], Iliyana Dobreva[1], Steven M. Quiring[1]

[1]Department of Geography, The Ohio State University, Columbus, OH, USA
[2]Climate Science in the School of Natural Resources, University of Missouri, Columbia, MO

**Corresponding author address:* Eshita A. Eva, Department of Geography, The Ohio State University, 1145 Derby Hall, 154 N Oval Mall, Columbus, OH 43210
E-mail: evaeshita22@gmail.com

https://doi.org/10.5194/egusphere-2025-896
Preprint. Discussion started: 7 April 2025

Abstract

Soil moisture is a fundamental state variable in climatology, meteorology, and hydrology. Many of the available soil moisture products have a coarse spatial resolution that is not useful for agricultural applications. This study used Random Forest to identify which features are most helpful for accurately downscaling soil moisture to 1-km resolution. Fourteen features were considered: precipitation, antecedent precipitation index, maximum daily air temperature, minimum daily air temperature, mean daily air temperature, diurnal temperature range, dew point temperature, elevation, slope, aspect, normalized difference vegetation index, leaf area index (LAI), soil texture, and land use/land cover. The analysis of variable importance was repeated using two different sources of soil moisture data (e.g., satellite-derived soil moisture from NASA's Soil Moisture Active Passive (SMAP) and model-derived soil moisture from the North American Land Assimilation System (NLDAS)) and ~~two different ways of representing soil saturation (e.g., volumetric water content (VWC) and percentiles)~~. We found that dew point temperature is the most important variable for downscaling SMAP percentiles (0.18), NLDAS VWC (0.27), and NLDAS percentiles (0.17) over CONUS, while elevation is the most important variable for downscaling SMAP VWC (0.28). Dew point temperature is crucial for downscaling in most regions of the United States, except in the South and WestNorthCentral, where elevation is the most important feature. The accuracy of the downscaling varies by region. In the South, SMAP VWC and NLDAS VWC downscaling are relatively accurate, both have mean absolute errors of ~0.07. The MAE values in the South region are 0.196 for SMAP percentiles and 0.175 for NLDAS percentiles.

Keywords: Soil moisture, downscaling, SMAP, NLDAS, Random Forest

https://doi.org/10.5194/egusphere-2025-896
Preprint. Discussion started: 7 April 2025

1. Introduction

Soil moisture is an important state variable in the climate system. Soil moisture is essential for examining, monitoring, and forecasting drought and streamflow (Abbaspour et al. 2015; Keesstra et al. 2016), hydrological parameter distribution (Carrijo et al. 2019; Dobriyal et al. 2012; Pittelkow et al. 2015), flash flood management (Brocca et al. 2014; Wanders et al. 2014), drought (Ruosteenoja et al. 2020; Wagner et al. 2013) and weather forecasting (Seneviratne et al. 2010; Zhang et al. 2019). Soil moisture is also important for agriculture. It influences crop yield, irrigation planning, disease outbreaks, pest control, and determining optimal management practices, including irrigation (Dobriyal et al. 2012; Pittelkow et al. 2015).

Soil moisture is a small constituent of the total available freshwater (0.0015%) and the global water cycle (0.05%) (Robinson et al. 2008). However, soil moisture is a critical element of land-atmosphere interactions because it plays a significant role in modulating exchanges of water between the land and atmosphere through evapotranspiration and transpiration (Ford et al. 2016; Zhang et al. 2019). Processes such as evapotranspiration, runoff, infiltration, and groundwater recharge are influenced by the soil moisture and so it also has a significant part to play in the hydrologic cycle (Vereecken et al. 2008).

Although soil moisture is useful for many purposes, national soil moisture data are not readily available at high resolution (e.g., kilometer scale). For example, satellite soil moisture products are limited to 1-km spatial resolution and model-derived products are limited to 12.5-km, therefore often it is not sufficient for field-scale applications in the agricultural sector. To make these datasets applicable to the finer scale, downscaling is necessary. It is possible to conduct soil moisture downscaling in different ways. To assess soil moisture heterogeneity at a finer scale, one approach is to combine ancillary data with low-resolution soil moisture estimates. This has been implemented in several studies to obtain soil moisture at high spatial resolution (Alemohammad et al. 2018; Srivastava et al. 2013). The spatial heterogeneity of soil moisture is influenced by numerous factors, such as topography, precipitation, temperature, soil texture, and vegetation. The purpose of this study is to identify which variables are most important for downscaling soil moisture to 1-km spatial resolution. We downscaled SMAP and NLDAS using Random Forest. We compared the importance of the ancillary variables for two different data sources and two different methods for representing soil moisture (e.g., percentiles and volumetric water content). To the best of our knowledge, this is the first study to evaluate feature importance over CONUS using both satellite-derived and modeled soil moisture.

2. Data

2.1 Soil moisture data

2.1.1 In-situ soil moisture

In-situ datasets are useful for calibration and validation of the SMAP and NLDAS soil moisture. In-situ measurements of soil moisture were used to validate the results of the downscaling approach. The in-situ measurement of soil moisture is based on sensors installed in various national and state networks. The networks used in this study include: the U.S. Climate Reference Network (CRN), Delaware Environmental Observing System (DEOS), North Carolina Environment and Climate Observing Network (ECONet), Illinois Climate Network (ICN), Kansas Mesonet (KS Mesonet, New Jersey Weather and Climate Network (NJWCN), NOAA Hydrometeorological Testbed (NOAA), New York Mesonet (NY Mesonet), Oklahoma Mesonet (OK Mesonet), Soil Climate Analysis Network (SCAN), South Dakota Mesonet (SD Mesonet), Snowpack Telemetry (SNOTEL), Texas Soil Observation Network (TxSon), Georgia Automated Environmental Monitoring Network (GA AEMN), and West Texas Mesonet (WTX Mesonet). A

https://doi.org/10.5194/egusphere-2025-896
Preprint. Discussion started: 7 April 2025

total of 1542 in-situ stations were used for this study. The number of stations varies from region to region greatly. For example, the South has a lot of stations, while the North~~W~~est has few stations. In-situ data from 2001 to 2021 were used for validating the downscaled soil moisture.

2.1.2 SMAP soil moisture

Remote sensing data provides gridded products instead of point-based soil moisture measurements and improves spatial coverage. Ford and Quiring (2019) examined the remote sensing soil moisture datasets provided by SMAP (SMAP L3 and SMAP L4), SMOS, and ESA-CCI and they concluded that SMAP Level 3 products consistently performed best among the four datasets. Though SMAP L3 products perform better than other remote sensing products, it has a significant number of missing values and dates. For that reason, SMAP L4 data are used in this study. The SMAP L4 ~~product is a~~ model-derived value-added products of surface and root zone soil moisture that support key SMAP applications. SMAP produces Level 4 data products that combine surface observations with a land surface model using a data assimilation system to simulate root zone soil moisture. The SMAP L4 data are available from 2015 to the present and they provide surface soil moisture (0-5 cm) and root zone soil moisture (0-100 cm). This study used SMAP data from 2015 to 2021.

2.1.3 NLDAS soil moisture

Land surface models use climate inputs and parameterizations of the environment to simulate soil moisture based on model-derived equations and assumptions. The model-derived soil moisture products can be achieved consistently at the global scale because the model can be operated seamlessly. However, model-simulated soil moisture has limitations because there is a nonlinear relationship between climate parameters and soil water content. In addition, soils and

https://doi.org/10.5194/egusphere-2025-896
Preprint. Discussion started: 7 April 2025

vegetation are spatially heterogeneous and this is not accurately captured by models (Xia et al. 2015). Ford and Quiring (2019) examined the NLDAS-2 and found that it performed better than the CPC-modeled soil moisture when validated against in-situ measurements. Therefore, this study uses NLDAS-2 model-simulated soil moisture. The soil moisture data are provided hourly at 1/8° spatial resolution by the NLDAS-2 Noah model. The data are available at four soil layers: 0-10 cm, 10-40 cm, 40-100 cm, and 100-200 cm. To align the temporal granularity with in-situ and satellite data, we averaged the hourly data to daily. NLDAS-2 data from 2001 to 2021 were used in this study. The downscaling was done every day from 2001 to 2021 within the CONUS to generate a 1-km soil moisture from NLDAS-2.

2.2 Ancillary variables

This study evaluated a total of 14 features: precipitation, antecedent precipitation index, maximum daily air temperature, minimum daily air temperature, mean daily air temperature, diurnal temperature range, dew point temperature, elevation, slope, aspect, Normalized Difference Vegetation Index, Leaf Area Index, soil texture, and land use/land cover.

Soil texture (sand, silt, and clay percentage) came from the gridded soil survey geographic database (SSURGO) provided by Natural Resources Conservation Services (NRCS) (https://www.nrcs.usda.gov/). It has a spatial resolution of 10-30 m. Precipitation, mean temperature, maximum temperature, minimum temperature, and dew point temperature were obtained from Parameter-elevation Regressions on the Independent Slopes Model (PRISM) lab (https://prism.oregonstate.edu/). The spatial resolution of PRISM is 4-km. The Antecedent Precipitation Index (API) was calculated from the precipitation data for each day. The antecedent precipitation index (API) formula is,

https://doi.org/10.5194/egusphere-2025-896
Preprint. Discussion started: 7 April 2025

$$API = P_i * N * 0.98$$

where: P_i: daily precipitation i days before a storm, N: total number of days to consider before the storm, K: a recession constant that is less than 1.0 (here, a value of 0.98 was used).

The diurnal temperature range (DTR) was calculated based on the difference between the maximum to minimum temperature. Elevation data were extracted from GTOP30 provided by the United States Geological Survey (USGS), and the spatial resolution is 1-km. Slope and aspect were derived from the elevation data. The Normalized Difference Vegetation Index (NDVI) and Leaf Area Index (LAI) are from the Advanced Very High-Resolution Radiometer (AVHRR) satellite. The spatial resolution of the AVHRR satellite is 9-km. Land use data were collected from the National Land Cover Dataset (NLCD). The native spatial resolution of NLCD is 30-m. Table 1 provides a description of all of the ancillary variables.

Table 1: Spatial and temporal resolution of variables evaluated for downscaling soil moisture

https://doi.org/10.5194/egusphere-2025-896
Preprint. Discussion started: 7 April 2025

Variable Name	Temporal Domain	Temporal Resolution	Spatial Resolution	Source
Elevation	2018	N/A	1-km	USGS EROS Archive - Digital Elevation - Global 30 Arc-Second Elevation (GTOPO30)
Slope	2018	N/A	1-km	USGS EROS Archive - Digital Elevation - Global 30 Arc-Second Elevation (GTOPO30)
Aspect	2018	N/A	1-km	USGS EROS Archive - Digital Elevation - Global 30 Arc-Second Elevation (GTOPO30)
Soil Texture	2021	N/A	10 to 30-m	Gridded Soil Survey Geographic (gSSURGO) Database
Land Use	2021	N/A	30-m	National Land Cover Dataset (NLCD)
Precipitation	2001 to 2021	Daily	4-km	Parameter-elevation Relationships on Independent Slopes Model (PRISM)
Antecedent Precipitation Index (API)	2001 to 2021	Daily	4-km	Parameter-elevation Relationships on Independent Slopes Model (PRISM)
Mean Temperature	2001 to 2021	Daily	4-km	Parameter-elevation Relationships on Independent Slopes Model (PRISM)
Maximum Temperature	2001 to 2021	Daily	4-km	Parameter-elevation Relationships on Independent Slopes Model (PRISM)
Minimum Temperature	2001 to 2021	Daily	4-km	Parameter-elevation Relationships on Independent Slopes Model (PRISM)
Dew Point Temperature	2001 to 2021	Daily	4-km	Parameter-elevation Relationships on Independent Slopes Model (PRISM)
~~Normalized Difference Vegetation Index (NDVI)~~	~~2001 to 2021~~	~~Daily~~	~~4-km~~	~~Parameter-elevation Relationships on Independent Slopes Model (PRISM)~~
Normalized Difference Vegetation Index (NDVI)	2001 to 2021	Daily	9-km	Advanced very-high-resolution radiometer (AVHRR)
Leaf Area Index (LAI)	2001 to 2021	Daily	9-km	Advanced very-high-resolution radiometer (AVHRR)

2.3 Soil Moisture Standardization

Volumetric water content (VWC) varies according to weather conditions, soil properties, vegetation cover, topographic features, and various other elements. Thus, it is not possible to directly compare the VWC from different locations since a fine-textured soil will almost always

https://doi.org/10.5194/egusphere-2025-896
Preprint. Discussion started: 7 April 2025

have a higher VWC than a coarse-textured soil. Therefore, VWC is commonly standardized using an approach such as soil moisture percentiles (Ford et al. 2016; Zhang et al. 2019). In this paper, we converted the soil moisture datasets to percentiles using the following approach. First, the VWC data were converted to VWC anomalies (m^3 m^{-3}) by subtracting the climatological mean from the daily VWC value (Crow et al. 2012). A moving window approach was used to calculate the climatological mean. Following Chen et al. (2019), a 31-day moving window was used surrounding the target day, and all years of data from the period of record from that window were used to calculate the mean. Next, daily percentiles were calculated using the empirical cumulative distribution function based on a 31-day moving approach that used the entire period of record. According to Ford et al. (2016), at least 6 years of daily data are required to create stable and robust percentiles. In our case, soil moisture percentiles were calculated based on 7 years of SMAP data and 21 years of NLDAS data.

3. Methods

The NLDAS and SMAP soil moisture data were downscaled using Random Forest. This algorithm was applied separately to the SMAP soil moisture (9-km) and the NLDAS soil moisture (12.5-km) to generate a 1-km soil moisture dataset over CONUS. Figure 1 provides a summary of the methods used to identify the most important features for downscaling soil moisture. The first step was to retrieve from different sources and aggregate the ancillary variables by applying nearest-neighbor interpolation. Then an RF model was generated for downscaling SMAP and NLDAS soil moisture products. The third step was to summarize feature importance. Each step is described in detail below.

https://doi.org/10.5194/egusphere-2025-896
Preprint. Discussion started: 7 April 2025

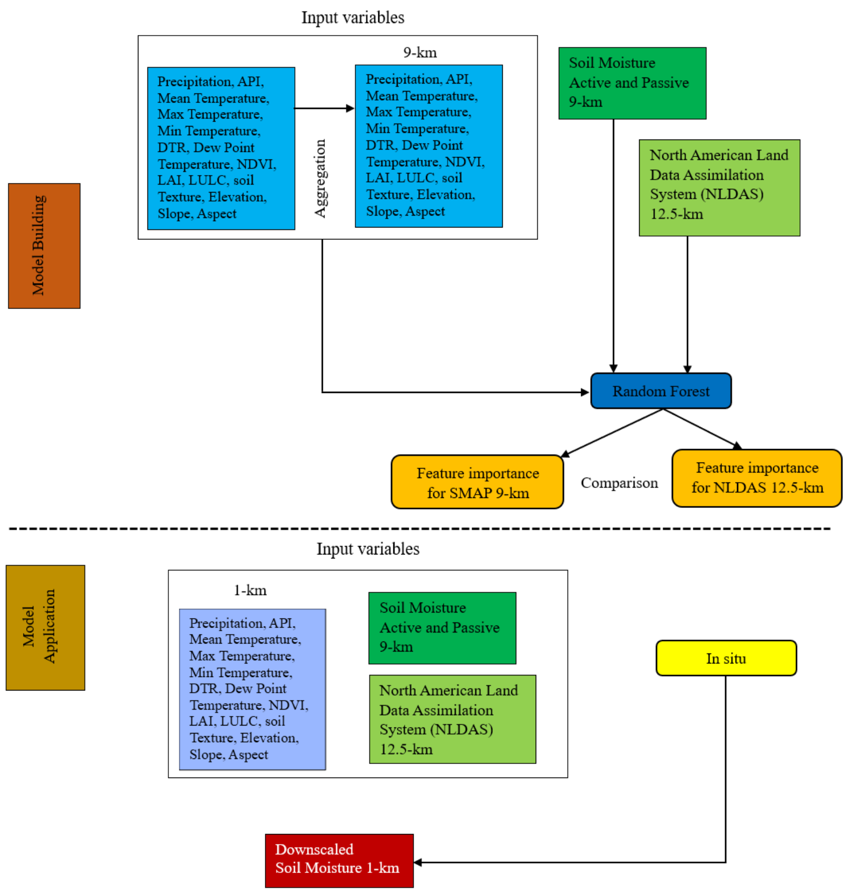

Figure 1: Schematic representation of the methodology used in this study

3.1 Nearest neighbor interpolation

All ancillary data are first aggregated to a resolution of 9-km and 12.5-km to construct the downscaling models. Nearest neighbor interpolation is used to resample the ancillary data (Table 1). Nearest neighbor is a method of spatial resampling that calculates the value of a new pixel based on the value of its nearest neighbor. This method was selected because it is computationally efficient and previous soil moisture downscaling studies have applied the nearest neighbor method (Alemohammad et al. 2018; Li et al. 2016; Liu et al. 2020, 2011). This

https://doi.org/10.5194/egusphere-2025-896
Preprint. Discussion started: 7 April 2025

method was applied twice, once to generate features at the same resolution as SMAP (9-km) and to generate features at the same resolution as NLDAS (12.5-km).

3.2 Downscaling soil moisture with random forest

Soil moisture downscaling to 1-km from the native SMAP (9-km) and NLDAS (12.5-km) resolutions was done using Random Forest. Random Forest (RF) was first developed by (Breiman 2001) and then later advanced by (Cutler et al. 2012). RF is an integrated learning algorithm that can be used for both classification of a categorical response and regression for continuous variables (Li et al. 2013; Lu et al. 2014). The advantages of RF are that it is very fast to train and predict, it only depends on a few tuning parameters, and it can be used for high-dimensional problems, outlier detection, and visualization (Cutler et al. 2012). The regression model in RF was used for downscaling soil moisture in this study.

The dataset was randomly split into 80% training and 20% testing for model evaluation. To optimize the model, we performed randomized search cross-validation over a set of hyperparameters, including the number of trees (n_estimators): {20, 40, 60, …, 200}, maximum depth of trees (max_depth): {10, 20, 30, …, 110, None}, maximum number of features per split (max_features): {'auto', 'sqrt'}, minimum number of samples per split (min_samples_split): {2, 5, 10}, minimum number of samples per leaf (min_samples_leaf): {1, 2, 4}, and bootstrap sampling method: {True, False}. The hyperparameter tuning is the same for each region.

The RF model for SMAP was generated using data from 2015 to 2021, and the RF model for NLDAS was generated using data from 2001 to 2021. A 3-fold cross-validation was used for cross-validation, each fold randomly selected 80% of the data for training and the remaining 20% was used for validation. The RF model with the highest accuracy was selected based on the

https://doi.org/10.5194/egusphere-2025-896
Preprint. Discussion started: 7 April 2025

cross-validation. This model was used to downscale the soil moisture to 1-km resolution and this downscaled soil moisture was validated using the in-situ measurements. This process was repeated separately for SMAP and NLDAS.

3.3 Model evaluation

The correlation coefficient (R), coefficient of determination (R^2), mean absolute error (MAE), mean squared error (MSE), and root mean square error (RMSE) were used for validating the downscaled soil moisture data. R and R^2 measure the goodness-of-fit of the model. MAE, MSE and RMSE measure the error between the actual and downscaled soil moisture values.

$$\mathrm{R} = \frac{\mathrm{Cov}\,(\mathrm{SM_o}, \mathrm{SM_p})}{\sqrt{\mathrm{Var}[\mathrm{SM_o}]\mathrm{Var}[\mathrm{SM_p}]}}$$

$$R^2 = 1 - \frac{\sum_{\mathrm{i=1}}^{\mathrm{n}}[(\mathrm{SM_p} - \mathrm{SM_{mp}})(\mathrm{SM_o} - \mathrm{SM_{mo}})]^2}{\sum_{\mathrm{i=1}}^{\mathrm{n}}(\mathrm{SM_p} - \mathrm{SM_{mp}})^2 \sum_{\mathrm{i=1}}^{\mathrm{n}}(\mathrm{SM_o} - \mathrm{SM_{mo}})^2}$$

$$\mathrm{MAE} = \frac{1}{\mathrm{n}} \sum_{\mathrm{i=1}}^{\mathrm{n}} |\mathrm{SM_o} - \mathrm{SM_p}|$$

$$MSE = \frac{1}{\mathrm{n}} \sum_{\mathrm{i=1}}^{\mathrm{n}} (\mathrm{SM_o} - \mathrm{SM_p})^2$$

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{\mathrm{i=1}}^{\mathrm{n}}(\mathrm{SM_o} - \mathrm{SM_p})^2}{\mathrm{n}}}$$

Where n is the number of observations, SM_o, SM_p, SM_{mo}, SM_{po} represent the observed soil moisture value, predicted soil moisture value, mean observed soil moisture value, and mean predicted soil moisture value.

https://doi.org/10.5194/egusphere-2025-896
Preprint. Discussion started: 7 April 2025

4 Results

4.1 Feature importance for SMAP downscaling

There are two measures that can be used to quantify the predictive power of individual features in RF Regression: (1) an increased mean squared error (IncMSE) and, (2) an increase in the number of nodes purified (IncNudePurity). IncMSE measures how an individual feature changes when it is permuted at random. IncMSE measures the degree to which the accuracy of Random Forest decreases when a feature is removed. Fourteen features were used to downscale the coarse resolution of soil moisture products to 1-km resolution. The IncMSE was calculated for each feature in each of the 9 regions to determine feature importance.

At the national (CONUS) scale, elevation is the most influential variable and dew point temperature is the second most important feature (Figure 2). Maximum temperature is also ranked as an important variable that improves the accuracy of downscaling of SMAP VWC over CONUS. The least important features at the national scale are slope, aspect, and LAI.

In addition to evaluating feature importance for each region, the analysis was also replicated by evaluating feature importance for different regions of CONUS. Figure 3 shows the feature importance for SMAP VWC by region. The dew point temperature is the most important feature in all regions, except the WestNorthCentral and South regions. Elevation is the most important variable in these two regions. API is the second-most important variable in the NorthWest, West, and SouthWest regions. Maximum, minimum, and mean temperature also play a significant role in downscaling soil moisture. LAI is a feature that consistently is ranked as a relatively unimportant feature for downscaling soil moisture (ranks from 0.1 to 0.2 across the 9 regions).

This study evaluates both VWC and soil moisture percentiles. Therefore, we can compare the relative feature importance for downscaling soil moisture percentiles. At the CONUS scale, dewpoint temperature is the most influential variable (Figure 4). This is followed by API and elevation. The results at the CONUS scale are relatively consistent with the regional results. Additionally, precipitation, maximum, mean temperatures, and DRT are all relatively important variables for downscaling soil moisture percentiles at the CONUS scale (rank ranges from 3 to 6).

Similar to the results for SMAP VWC, dew point temperatures are the most influential variable in all regions except the South (Figure 5). Elevation is the most influential feature in the South. The second most important feature is either API or elevation. API is the second-most influential variable in the NorthWest, SouthWest, and EastNorthCentral, NorthEast, SouthEast, and Central regions. While elevation is the second most important feature of WestNorthCentral regions. LAI, aspect, slope, NDVI, and LULC are the least important features for downscaling soil moisture percentiles.

https://doi.org/10.5194/egusphere-2025-896
Preprint. Discussion started: 7 April 2025

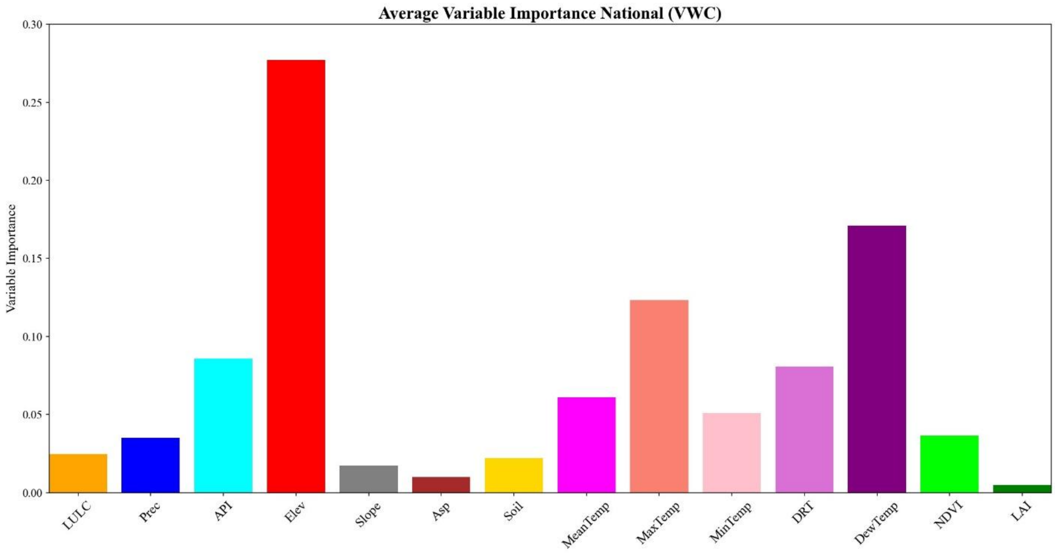

Figure 2: Feature importance (IncMSE%) for CONUS for SMAP VWC

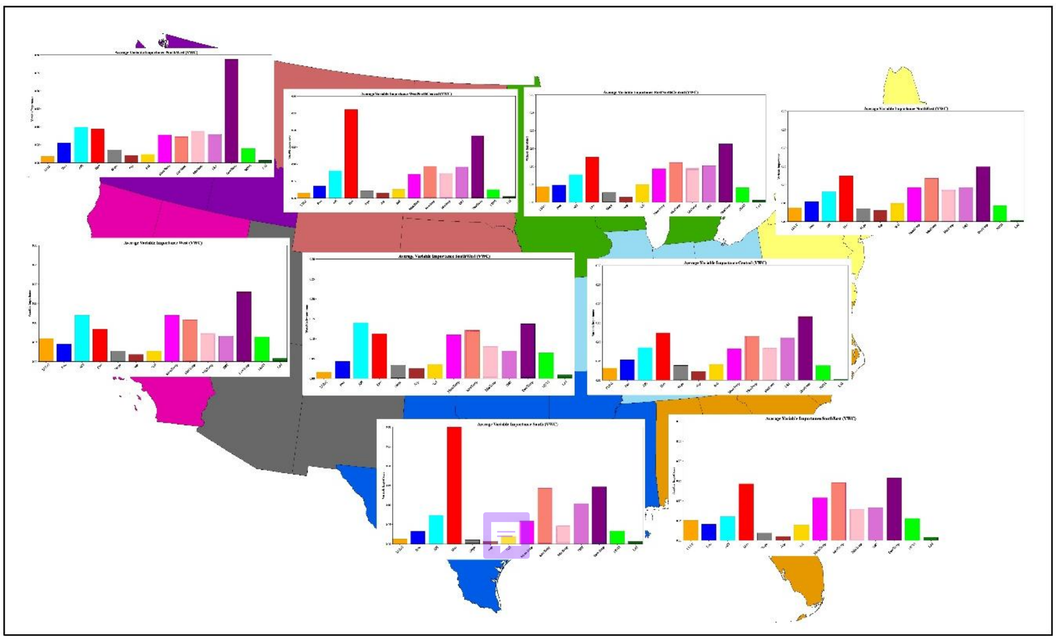

Figure 3: Features importance (IncMSE%) by region for SMAP VWC

https://doi.org/10.5194/egusphere-2025-896
Preprint. Discussion started: 7 April 2025

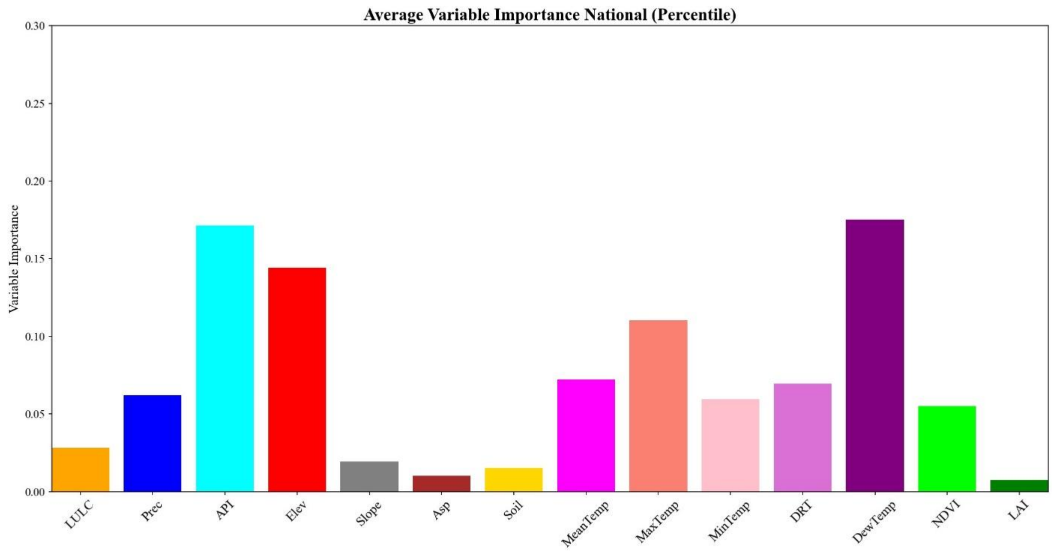

Figure 4: Feature importance (IncMSE%) for CONUS for SMAP percentile

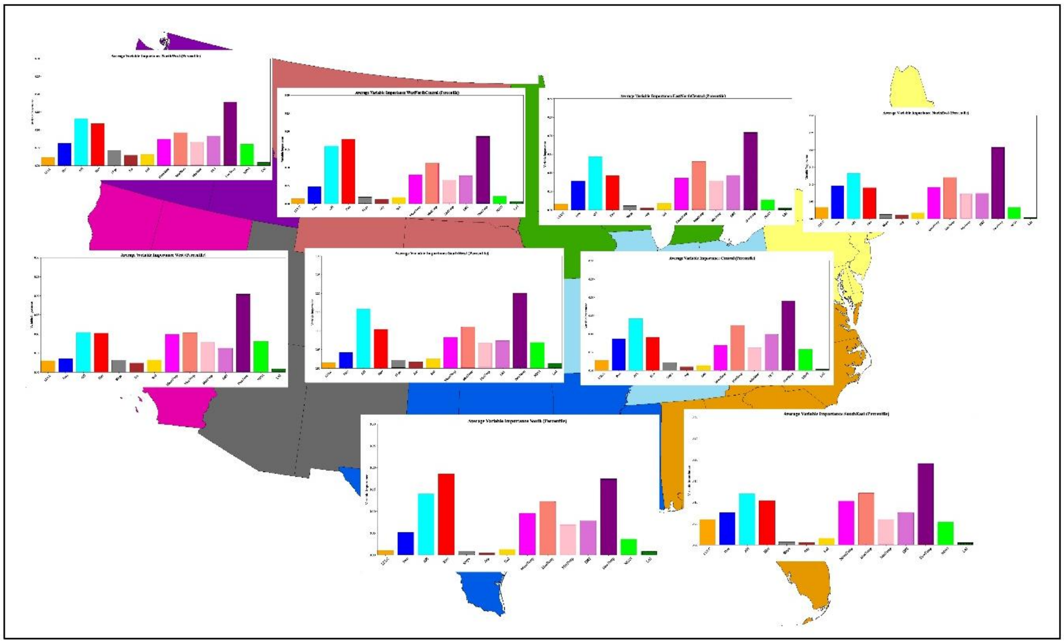

Figure 5: Feature importance (IncMSE%) by region for SMAP percentiles

https://doi.org/10.5194/egusphere-2025-896
Preprint. Discussion started: 7 April 2025

4.2 Feature importance for NLDAS downscaling

At the CONUS scale, dew point temperature is the most important feature (Figure 6). Elevation and maximum temperature have almost the same influence and are ranked as the second and third most important features on the CONUS scale. Though elevation was the highest influential variable for SMAP VWC, elevation is ranked second here. However, as SMAP VWC, slope, aspect, LAI ranked the lowest.

Figure 7 shows the feature importance for downscaling NLDAS VWC in each region. Dew point temperature is the most important feature in all regions. Elevation is the second-highest important feature in all regions, except the NorthWest, West, SouthWest, and EastNorthCentral regions. The second most important variable for the NorthWest, SouthWest, and EastNorthCentral region is maximum temperature; for the West is mean temperature. Maximum and mean temperature are also influential features for downscaling NLDAS VWC. They typically rank as each region's 3rd or 4th most important features. While LAI and aspect are generally the least important features. These results are quite consistent with the SMAP feature importance.

At the CONUS scale, dewpoint temperature is again the most influential feature (Figure 8). This is followed by API and elevation. Maximum and mean temperatures are also influential features at the CONUS scale. The results of the feature importance for NLDAS soil moisture percentiles are consistent with NLDAS VWC. Dew point temperature is the most important feature in all regions (Figure 9). API is the second most important feature in all regions, except the NorthWest, West, WestNorthCentral. In these regions, elevation is the second most important feature. LAI and aspect consistently have the lowest feature importance for downscaling NLDAS soil moisture percentiles.

https://doi.org/10.5194/egusphere-2025-896
Preprint. Discussion started: 7 April 2025

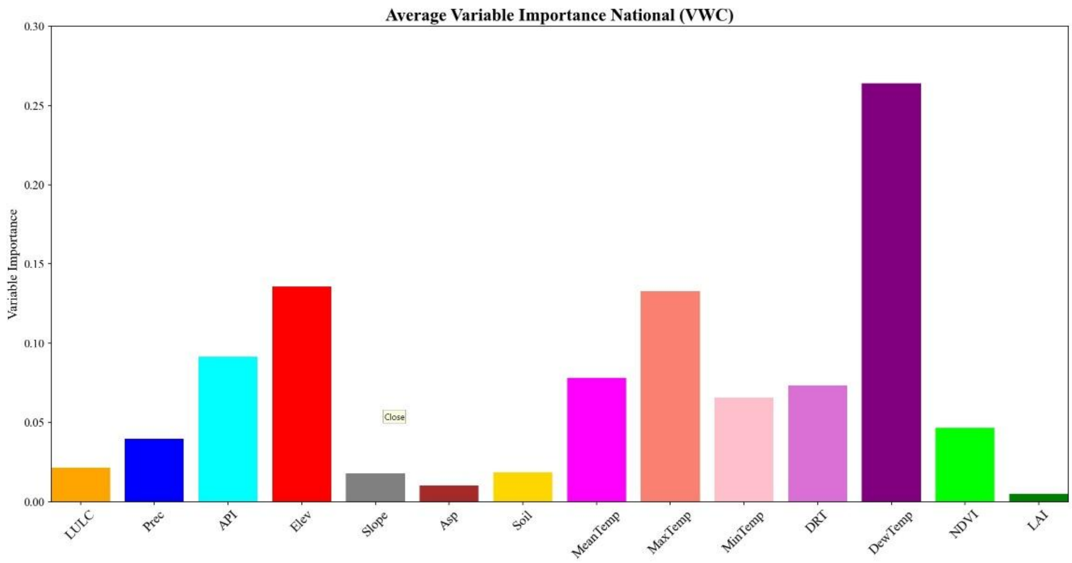

Figure 6: Feature importance (IncMSE%) for CONUS for NLDAS VWC

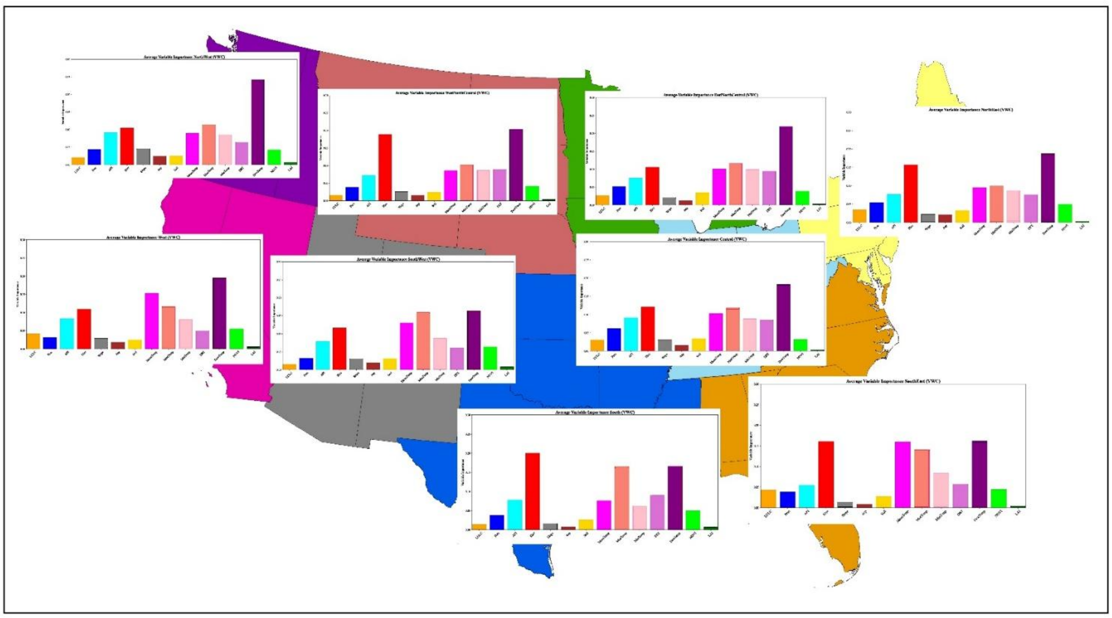

Figure 7: Feature importance (IncMSE%) by region for NLDAS VWC

https://doi.org/10.5194/egusphere-2025-896
Preprint. Discussion started: 7 April 2025

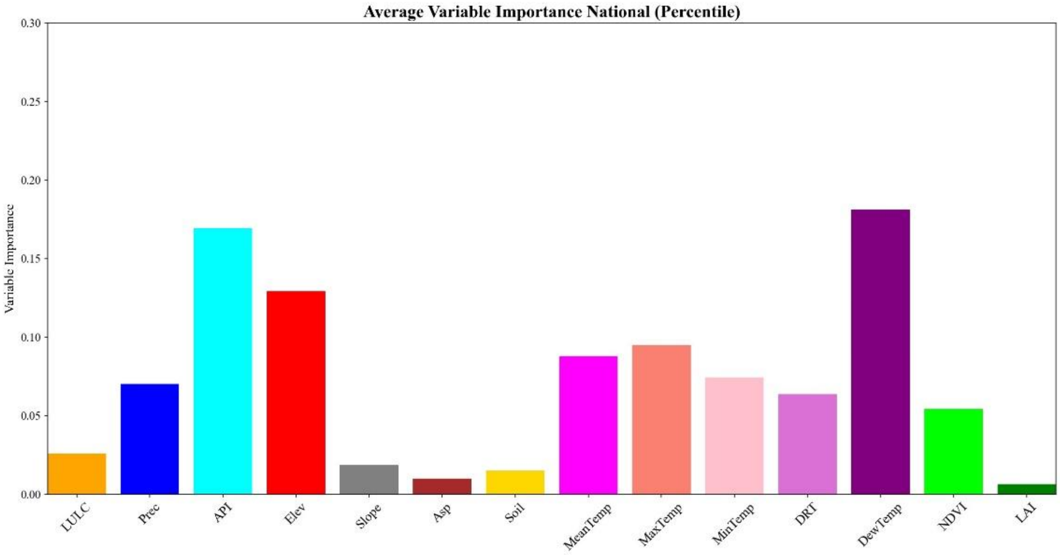

Figure 8: Feature importance (IncMSE%) for CONUS for NLDAS percentiles

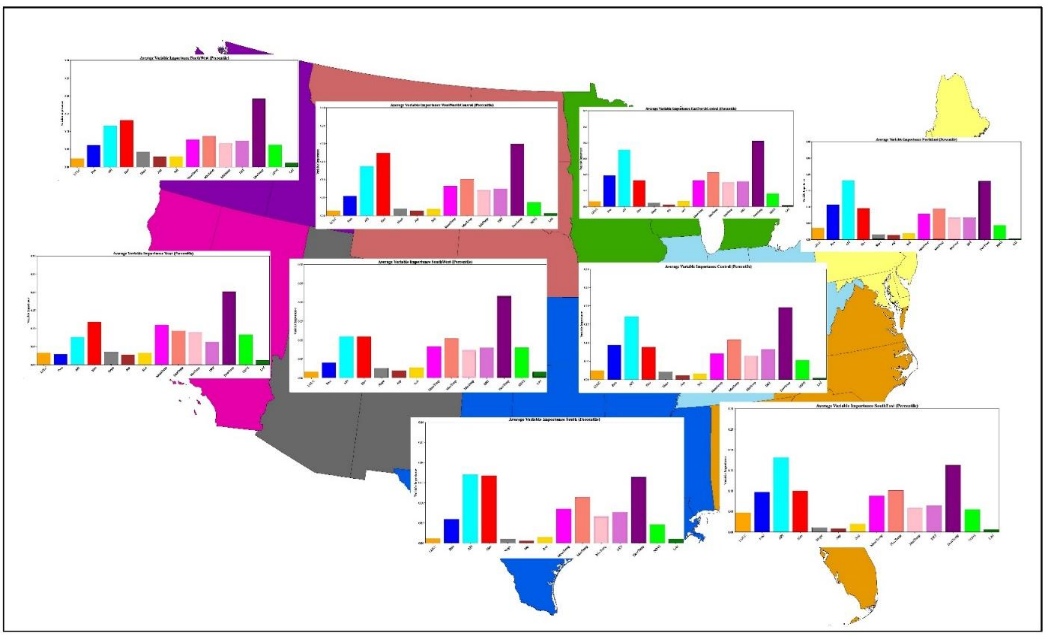

Figure 9: Feature importance (IncMSE%) by region for NLDAS percentile

https://doi.org/10.5194/egusphere-2025-896
Preprint. Discussion started: 7 April 2025

4.3 Partial dependence plot (PDP) of variables

Figure 10 represents the relationship between predicted soil moisture and the other variables (dew point temperature, elevation, API, and maximum temperature).

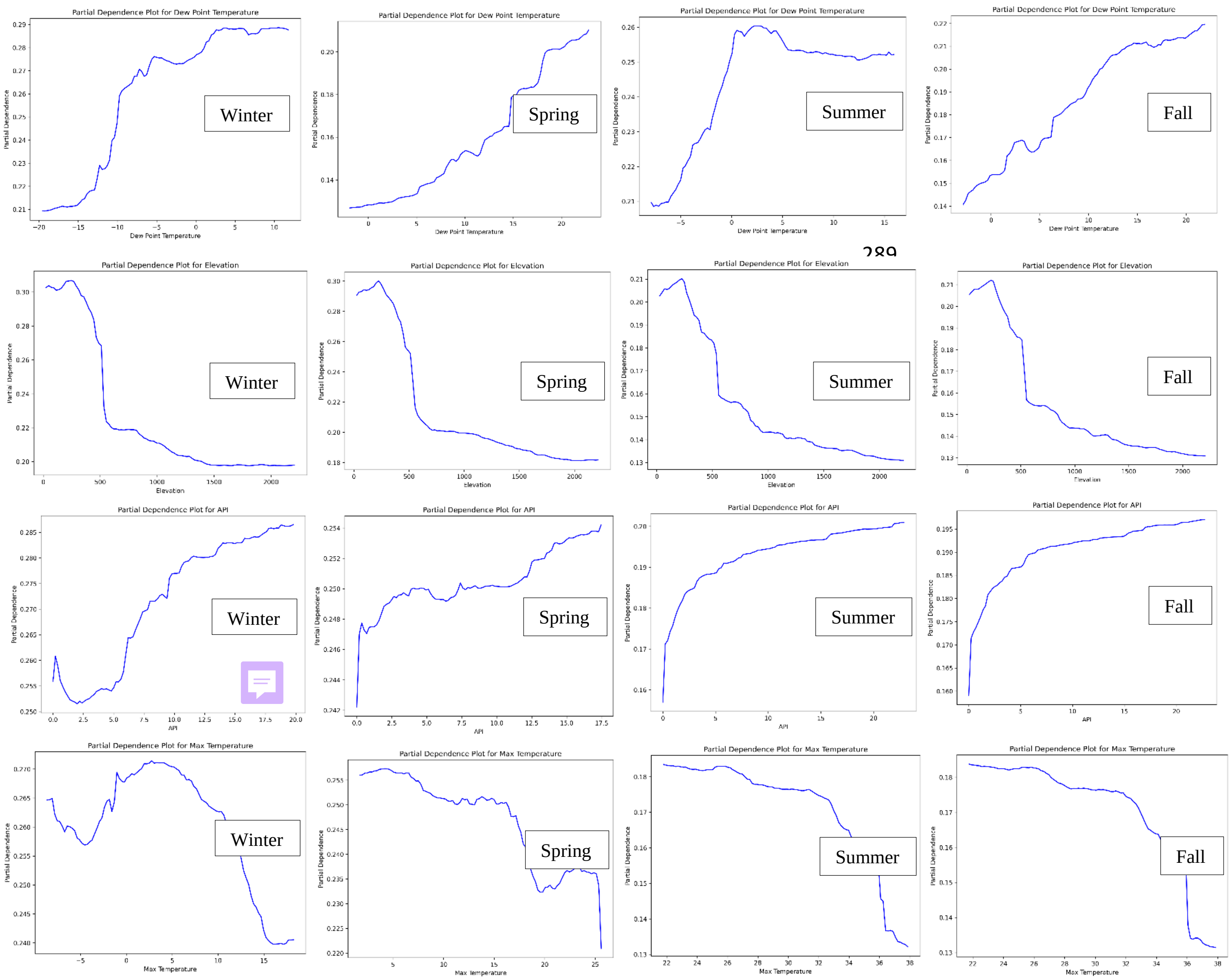

Figure 10: Partial dependence plot for the four most important variables

https://doi.org/10.5194/egusphere-2025-896
Preprint. Discussion started: 7 April 2025

The higher dew point temperature is related to the high soil moisture values in all seasons. The lower elevation seems to have high soil moisture values in four seasons which is reasonable because in the high elevation, the water content in the soil is low. If the API value is increasing, the value of water content in the soil will be high and vice versa. On the other hand, with the increasing maximum temperature, the soil moisture will be decreasing.

4.3 Accuracy of VWC downscaling

Table 2 shows the accuracy of the VWC downscaling in each region based on the validation against in situ measurements. As shown in the table, the R-value is < 0.5 and R^2 < 0.25 for NLDAS downscaling. The ~~average~~ absolute difference between the SMAP downscaling and observed value (MAE) ranges from 0.074 to 0.092. The Central region has a relatively low MAE of 0.074, indicating a small average absolute difference between the predicted and actual values. The highest value of MSE is 0.014 corresponding to SouthEast and 0.013 to EastNorthCentral and the lowest value is 0.008 corresponding to South. The ubRMSE for the South is 0.082 which is the lowest and for the SouthEast is 0.109 which is the highest. Although the R^2 values are relatively low, the errors (MAE and ubRMSE) are also low. This indicates that the downscaling is skillful.

Table 2: Accuracy of 1-km downscaled soil moisture (VWC) by region for SMAP and NLDAS

	R		**R2**		**MAE**		**MSE**		**RMSE**		**ubRMSE**		**Bias**	
	SMAP	*NLDAS*	*SMAP*	*NLDAS*	*SMAP*	*NLDAS*	*SMAP*	*NLDAS*	*SMAP*	*NLDAS*	*SMAP*	*NLDAS*	*SMAP*	*NLDAS*
Central	0.451	0.352	0.204	0.146	0.074	0.086	0.009	0.011	0.093	0.104	0.089	0.091	-0.008	-0.021

https://doi.org/10.5194/egusphere-2025-896
Preprint. Discussion started: 7 April 2025

EastNorthCentral	0.403	0.494	0.157	0.248	0.092	0.072	0.013	0.010	0.115	0.097	0.093	0.087	0.053	-0.017
NorthEast	0.323	0.337	0.104	0.131	0.075	0.084	0.009	0.010	0.095	0.100	0.089	0.085	0.001	-0.017
NorthWest	0.268	0.298	0.065	0.090	0.080	0.085	0.009	0.010	0.097	0.100	0.092	0.089	0.022	0.014
South	0.568	0.551	0.323	0.304	0.072	0.069	0.008	0.007	0.091	0.086	0.082	0.080	-0.033	-0.020
SouthEast	0.419	0.622	0.176	0.386	0.094	0.084	0.014	0.010	0.116	0.101	0.109	0.090	0.034	0.010
SouthWest	0.364	0.622	0.132	0.073	0.078	0.091	0.010	0.012	0.098	0.108	0.093	0.097	-0.020	-0.004
West	0.308	0.270	0.099	0.081	0.075	0.084	0.009	0.010	0.094	0.099	0.092	0.090	0.016	0.025
WestNorthCentral	0.520	0.494	0.270	0.248	0.073	0.080	0.008	0.010	0.091	0.097	0.085	0.087	-0.011	0.004
National	0.514	0.509	0.192	0.259	0.082	0.083	0.010	0.010	0.102	0.100	0.100	0.095	-0.015	-0.011

4.4 Accuracy of percentile downscaling

Table 3 provides the evaluation metrics of soil moisture downscaling for percentiles by region. The South region has the lowest MAE of 0.195, indicating the smallest average absolute difference between predicted and actual values among all the regions. The MSE is also the lowest which is 0.061 for this region, indicating a higher average squared difference. The RMSE of 0.248 is also the lowest among the regions, indicating the smallest average absolute difference. On the other hand, the NorthWest and SouthWest regions are characterized by the highest MAE of 0.223, which indicates a greater average absolute difference between the predicted and actual values compared to other regions. As indicated by the MSE value of ~0.076, there is a greater average squared difference between the two regions. RMSE of ~0.275 indicates

Table 3: Accuracy of 1-km downscaled soil moisture (percentiles) by region for SMAP and NLDAS

	R		**R^2**		**MAE**		**MSE**		**RMSE**		**ubRMSE**		**Bias**	
	SMAP	*NLDAS*	*SMAP*	*NLDAS*	*SMAP*	*NLDAS*	*SMAP*	*NLDAS*	*SMAP*	*NLDAS*	*SMAP*	*NLDAS*	*SMAP*	*NLDAS*
Central	0.423	0.536	0.187	0.297	0.200	0.177	0.065	0.050	0.256	0.224	0.222	0.198	-0.055	-0.016

https://doi.org/10.5194/egusphere-2025-896
Preprint. Discussion started: 7 April 2025

EastNorthCentral	0.411	0.429	0.198	0.304	0.207	0.195	0.065	0.056	0.256	0.236	0.216	0.189	-0.026	-0.045
NorthEast	0.269	0.307	0.085	0.120	0.207	0.211	0.066	0.068	0.257	0.261	0.226	0.220	-0.048	-0.034
NorthWest	0.229	0.194	0.050	0.048	0.223	0.237	0.076	0.084	0.276	0.291	0.249	0.252	0.0276	-0.058
South	0.442	0.488	0.195	0.238	0.195	0.175	0.061	0.049	0.248	0.223	0.224	0.208	-0.068	-0.019
SouthEast	0.394	0.437	0.156	0.195	0.207	0.189	0.066	0.055	0.258	0.236	0.227	0.212	-0.081	-0.027
SouthWest	0.229	0.229	0.053	0.058	0.223	0.228	0.077	0.079	0.278	0.281	0.256	0.253	0.000	-0.042
West	0.238	0.180	0.062	0.047	0.216	0.226	0.072	0.076	0.268	0.276	0.245	0.245	-0.015	-0.025
WestNorthCentral	0.363	0.367	0.141	0.144	0.215	0.218	0.073	0.074	0.270	0.272	0.253	0.244	-0.021	-0.054
National	0.425	0.370	0.029	0.138	0.217	0.226	0.073	0.080	0.270	0.284	0.260	0.262	-0.033	-0.080

a greater average absolute difference among the regions. The lowest MAE value is 0.175 in the South region. The lowest MAE value indicates the closer alignment between actual and predicted values. The average squared difference between the expected and actual values ranges from 0.050 to 0.086. Similar patterns can be observed between MSE and RMSE values. The lowest value of RMSE is 0.223 indicating good predictions for South regions.

5 Discussion

5.1 Feature importance

A complex interaction between pedologic, topographic, vegetative, and meteorological factors is responsible for the extensive horizontal variability in surface soil moisture fields (Mohanty and Skaggs 2001). It is difficult to isolate and measure these factors, however, understanding their magnitude is critical for determining soil moisture upscaling strategies (Crow et al. 2012).

Topography is ~~another~~ crucial factor that influences the spatial organization of soil moisture at various scales. The topographic attributes such as slope, aspect affect the movement of water through gravity-driven processes, such as runoff and infiltration, leading to variations in

https://doi.org/10.5194/egusphere-2025-896
Preprint. Discussion started: 7 April 2025

soil moisture across the landscape (Crow et al. 2012). This research ~~also complements~~ that topography is an important factor because elevation is the most influential factor and second highest factor relatively in SMAP VWC and NLDAS VWC downscaling over CONUS. In addition, elevation ranks third for percentile downscaling for both data sources. Understanding the role of topography is essential for accurately characterizing and predicting soil moisture patterns. Charpentier and Groffman (1992) suggested that variations in topography play a significant role in shaping the spatial distribution of soil moisture, with more diverse topographic features leading to increased variability in soil moisture. This study shows the South and WestNorthCentral regions have the influences of slope regardless of data sources and units of measurement. Since the WestNorthCentral region includes portions of the Rocky Mountains, and the mountainous area has high variability in elevation. It has been demonstrated that relative slope position significantly impacts determining soil moisture variation. When considering different time scales, a straightforward averaging of soil moisture values across a slope is likely to produce errors (Mohanty et al. 2000; Jacobs 2004). Here, this research shows differences because slope has a relatively low impact on the downscaling for both data sources and both units of measurement (VWC and percentile). In addition, the aspect (direction of slope) also has less impact on the downscaling in all the regions and over CONUS too. The effects of slope on soil moisture are more localized and dependent on other factors rather than being significant. It is important to mention that slope influences soil moisture in various ways, including its length and direction (aspect). Although these effects are generally less pronounced than those that elevation has on the local microclimate and hydrology, they can be observed (Bennie et al. 2008).

Among these meteorological factors, precipitation is considered the single most important forcing factor for soil moisture content and its distribution (Crow et al. 2012). This

https://doi.org/10.5194/egusphere-2025-896
Preprint. Discussion started: 7 April 2025

research agrees with Crow et al. (2012) because the precipitation index (API) ranks second in terms of SMAP percentile downscaling over the CONUS. Furthermore, API is also a dominant factor for the South, WestNorthCentral, NorthWest, SouthWest regions. Sivapalan et al. (1987) and Famiglietti et al. (1999) found the mean soil moisture trend changed with the precipitation gradient. The soil moisture storage, drainage, and water budget for different climate patterns changed with the intensity of rainfall (Kim et al. 1997; Salvucci 2001). To date, there has not been any research that has shown that dew point temperature is useful for downscaling soil. The present study found that dew point temperature is the most important feature for downscaling soil moisture for both products and both measurements, except SMAP VWC over the CONUS. Although elevation has the highest impact on SMAP VWC downscaling over the CONUS, dew point temperature has the highest influence over all the regions except the South, and WestNorthCentral regions both for SMAP VWC and percentile downscaling. For NLDAS VWC and percentiles, only the South region is different.

There are a number of ways that the dew point temperature can influence soil moisture. First, when the air temperature cools to the dew point temperature water vapor will condense, forming dew on the soil surface and vegetation (Monteith, 1957, Jacobs et al. 2000). Second, dew point temperature can also affect the transpiration rates of plants (Ambrose et al. 2009). Soil evaporation and plant transpiration will increase as a function of the vapor pressure gradient. Therefore, dew point temperature fluctuations can influence the wetting and drying cycles of the soil (Zhou et al. 2008).

Precipitation, mean, maximum, and minimum temperatures also have an impact on soil moisture. Higher temperatures tend to be associated with an increased vapor pressure deficit between the soil surface and the air, and this leads to a higher evaporation rate from the soil.

https://doi.org/10.5194/egusphere-2025-896
Preprint. Discussion started: 7 April 2025

Higher temperatures can cause the soil to dry out more rapidly, particularly from the surface of the soil (Balugani et al. 2014). They can also cause vegetation to utilize more soil water and so they can cause the depletion of soil moisture (Oren et al. 1999).

The distribution of soil moisture is significantly affected by soil heterogeneity, due to variations in soil properties such as texture, organic matter content, porosity, and structure. As a result of these variations, soil moisture can significantly differ over small spatial distances, influencing local hydrological processes. The color of the soil can also influence the rate of evaporative drying for bare or lightly vegetated soils by affecting the albedo. Soil hydraulic conductivity affects processes such as water infiltration, redistribution, and drainage, which in turn impact the distribution of soil moisture across a given area (Moore et al. 1988; Kim and Barros 2002). Although previous research concluded that soil type has a significant impact on soil moisture, this research found that soil type has minimal impact on downscaling soil moisture. In addition, whether over the CONUS or for all the regions, soil type is one of the least influential variables. The reason could be that other factors like elevation, prevailing climate, and precipitation patterns have a greater impact than the soil texture.

Previous work has concluded that variations in soil moisture are more strongly influenced by vegetation than by soil and topography (Crow et al. 2012). Our results do not support this conclusion (at least for 1-km soil moisture) since the vegetation variables (NDVI and LAI) were less important than topography. Soil moisture is spatially variable as a result of various processes that involve water uptake, transpiration, and the surface energy budget. The influence of land cover characteristics must be understood to accurately model and predict soil moisture dynamics. Land cover affected soil moisture distribution significantly at the satellite footprint during the NAFE'05 field campaign in Australia (Panciera et al. 2008). The regional distribution of soil

https://doi.org/10.5194/egusphere-2025-896
Preprint. Discussion started: 7 April 2025

moisture was noted to be impacted by variations in land cover (Cosh et al. 2004; Joshi and Mohanty 2010). However, the results of our study indicated that land cover was not an important feature for downscaling soil moisture to 1-km resolution. It had minimal impact on the downscaling regardless of the data sources or soil moisture measurement. These results were consistent over the CONUS and all regions.

Drawing broad conclusions regarding the impact of soil, topography, land cover, vegetation, and meteorological forcing can be challenging because feature important is strongly influenced by the spatial scale of analysis and the region of interest (Crow et al. 2012). Our study results demonstrate that there can be substantial region variations in feature importance and that feature importance is also a function of the data source, unit of measurement and spatial resolution.

Regions may differ in terms of climatic, geographic, and soil characteristics, which can have a significant impact on the relationship between the input features and soil moisture. Regions with limited or incomplete data may be difficult to predict accurately. Our analysis demonstrates that other than elevation, temperature is an important feature in the South. This region includes Texas, Louisiana, Mississippi, Arkansas, Kansas, and Oklahoma. It is characterized by a warm and humid climate. Soil moisture is influenced by precipitation and topography. Periodic droughts have a strong influence on soil water content (Bond et al. 2008; Engle et al. 2008).

The northeastern region includes New York, Pennsylvania, Connecticut, Delaware, Maine, Maryland, Massachusetts, New Hampshire, New Jersey, Vermont, and Rhode Island. It has a temperate climate with distinct seasonal patterns. Seasonal variations in precipitation and snowmelt can significantly impact soil moisture patterns (Daly et al. 2008). Our results

https://doi.org/10.5194/egusphere-2025-896
Preprint. Discussion started: 7 April 2025

demonstrate that API, dew point temperature, and temperature influence soil moisture. The northwestern region includes Idaho, Oregon, and Washington. It has a diverse climate ranging from maritime to continental. The EastNorthCentral includes Iowa, Michigan, Minnesota, and Wisconsin. It has a humid continental climate with distinct seasonal changes. The region is dominated by the Great Lakes. Soil moisture depends on temperature and precipitation patterns (Kunkel et al. 2013). Our results demonstrate that API and temperature have a significant impact on soil moisture.

The WestNorthCentral includes Montana, Nebraska, North Dakota, South Dakota, and Wyoming. It is characterized by a continental climate with hot summers and cold winters. Snowmelt and spring rains can increase soil moisture levels (Kunkel et al. 2013). Temperature and dew point temperature are important features for accurately downscaling soil moisture. The West includes California and Nevada. It has a diverse range of climates, from Mediterranean to arid and semi-arid. This region with mountainous terrain may experience higher soil moisture levels due to increased precipitation (Nielsen et al. 2024). The SouthWest includes Arizona, Colorado, New Mexico, and Utah. It has an arid to semi-arid climate. However, soil moisture is lower in this region (Huxman et al. 2004) and API has an impact on the soil moisture. Moreover, topographic features such as mountains and valleys can cause localized variations in soil moisture (Seager et al. 2007), which is evident in the present study.

5.3 Model evaluation

Abbaszadeh et al. (2019) downscaled SMAP 9-km (SMAP level 3) soil moisture to 1-km over the CONUS using ensemble learning methods and they found overall the ubRMSE between downscaled and in-situ observations met the SMAP accuracy requirement of 0.04. The downscaled soil moisture data showed a strong correlation (R=0.325 to 0.997, average 0.715)

https://doi.org/10.5194/egusphere-2025-896
Preprint. Discussion started: 7 April 2025

and high accuracy (ubRMSE=0.010 to 0.141 m³/m³, average 0.041 m³/m³) compared to the in-situ soil moisture measurements (Xu et al. 2022). Our study shows that the ubRMSE value for SMAP VWC is 0.100 m³/m³. The reason could be the different methods of downscaling because our analysis used RF, while Xu et al. (2022) used an ensemble approach. In addition, Xu et al. (2022) used SMAP level 3, while this analysis used SMAP level 4.

Sun and Cui (2021) downscaled VWC in the central United States using support vector machine (SVM). The correlation between their downscaled soil moisture and the in-situ measurements ranged from 0.174 to 0.754, and their RMSE ranged from 0.063 to 0.101. In our study, the correlation between the downscaled soil moisture and in-situ measurements in the central region was 0.452 for SMAP VWC and 0.352 for NLDAS VWC. The RMSE was 0.093 for SMAP VWC and 0.104 for NLDAS VWC. Therefore, our results are relatively comparable in terms of correlation and error.

Guevara and Vargas (2019) downscaled the soil moisture from 27-km to 1-km by training the kernel-weighted nearest neighbors across the conterminous United States. They showed that the downscaled soil moisture had an R^2 value of 0.46. Our results generally have lower correlations for all the regions, except the South. It is possible that this is because Guevara and Vargas (2019) used different methods of downscaling such kernel-weighted nearest neighbors, while our approach was based on. There is also the difference of spatial native resolutions of soil moisture products that were used for downscaling.

Liu et al. (2020) downscaled soil moisture using six methods over different regions including Oklahoma. The R^2 value for their Oklahoma region ranged from 0.287 to 0.714, and the MAE ranged from 0.027 to 0.055. The MAE in the South region in our study ranged from 0.069 to 0.195. A similar study in Oklahoma state was done by Jiang and Cotton (2004) applying

https://doi.org/10.5194/egusphere-2025-896
Preprint. Discussion started: 7 April 2025

the artificial neural network (ANN). Their RMSE was <0.04. Xu et al. (2021) downscaled SMAP 36-km soil moisture to 3-km and 1-km by applying the convolution neural network in Oklahoma. Xu et al. (2021) had a correlation of 0.659 and ubRMSE of 0.052 for their 1-km downscaled soil moisture using SMAP level 3 products from 1^{st} January 2018 to 30^{th} December 2018. While these studies had lower errors than our analysis, this is likely because they focused on a single state, while our region encompasses multiple states. In addition, they downscaled only one year soil moisture, but we downscaled daily soil moisture from 2015 to 2021 and evaluated performance using this entire period. In addition, Xu et al. (2021) utilized a convolutional neural network (CNN) and we used RF for downscaling.

Warner et al. (2021) downscaled VWC in Delaware to 100-m resolution using kernel k-nearest neighbor (KKNN) based on SSM estimates from the European Space Agency's Climate Change soil moisture products. They had an MAE of 0.048 (Warner et al. 2021). In comparison, our MAE in the NorthEast region, which includes Delaware, ranged from 0.076 to 0.082. Warner et al. (2021) likely had greater accuracy because they their study area is a single state (Delaware), while our NorthEast region includes Delaware, Connecticut, Maine, Maryland, Massachusetts, New Hampshire, New York, Pennsylvania, Rhode Island, and Vermont. In addition, Warner et al. (2021) downscaled different soil moisture data than our study.

Overall, the soil moisture downscaling is more accurate in the South, Central, and NorthEast regions, as they have lower error than in other regions. The downscaling was not as accurate in the NorthWest, SouthWest, West, and WestNorthCentral regions.

6 Conclusions

https://doi.org/10.5194/egusphere-2025-896
Preprint. Discussion started: 7 April 2025

This study evaluated and compared feature importance for downscaling satellite and model-derived soil moisture products across different regions. Random forest was applied to predict the soil moisture products (SMAP and NLDAS) to 1-km resolution over CONUS and in-situ data were used to validate the model. The conclusions are:

1) Results indicated that dew point temperature is the most important feature and elevation is the second most important feature for downscaling soil moisture in the United States. In general, the atmospheric features (e.g., temperature) have more impact on the downscaling than vegetation features. Vegetation features such as NDVI and LAI, as well as topographic features such as slope and aspect were not important for downscaling soil moisture to 1-km resolution.
2) Based on the accuracy metrics, downscaling VWC is more accurate than soil moisture percentiles.
3) Downscaled SMAP and NLDAS VWC was most accurate in the South and WestNorthCentral regions.

Our results can be used to improve feature selection for soil moisture downscaling. However, it is likely that the optimal features for downscaling soil moisture are strongly dependent on the spatial resolution of the analysis and the climatic, topographic and edaphic characteristics of the study region.

Future research can advance this work in the following ways. First, we can increase the validity of the accuracy assessment by using more in-situ data to validate the downscaling. This study compares the 1-km downscaled soil moisture to a single station. A better approach would be to use a dense network of stations to upscale the in-situ data to match the resolution of the soil moisture products. Second, given the substantial regional variations in performance, future work, would benefit from including a more comprehensive set of features that account for soil moisture dynamics in each region. Third, the method of upscaling the features can be improved. Last, this study only used a single method for downscaling. Using other downscaling approaches may result in increased accuracy and provide additional insights.

Author Contributions:

EAE-Conceptualization, methodology, formal analysis, validation, visualization, writing-original draft, **ZL-** Data acquisition, validation **ID-** Data acquisition, validation **SQ**-Funding acquisition, supervision, writing-review, editing.

Funding: This research was funded by the United States Department of Agriculture (USDA) under Grant AQD-109872.

Acknowledgments: The authors thank the anonymous reviewers for providing such valuable comments.

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

https://doi.org/10.5194/egusphere-2025-896
Preprint. Discussion started: 7 April 2025

References

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

Abbaszadeh, P., H. Moradkhani, and X. Zhan, 2019: Downscaling SMAP Radiometer Soil Moisture Over the CONUS Using an Ensemble Learning Method. *Water Resources Research*, **55**, 324–344, https://doi.org/10.1029/2018WR023354.

Alemohammad, S. H., J. Kolassa, C. Prigent, F. Aires, and P. Gentine, 2018: Global downscaling of remotely sensed soil moisture using neural networks. *Hydrology and Earth System Sciences*, **22**, 5341–5356, https://doi.org/10.5194/hess-22-5341-2018.

Ambrose, A. R., S. C. Sillett, and T. E. Dawson, 2009: Effects of tree height on branch hydraulics, leaf structure and gas exchange in California redwoods. *Plant, Cell & Environment*, **32**, 743–757, https://doi.org/10.1111/j.1365-3040.2009.01950.x.

Baldwin, D., K. J. Naithani, and H. Lin, 2017: Combined soil-terrain stratification for characterizing catchment-scale soil moisture variation. *Geoderma*, **285**, 260–269, https://doi.org/10.1016/j.geoderma.2016.09.031.

Balugani, E., M. Lubczynski, and K. Metselaar, 2014: A framework for sourcing of evaporation between saturated and unsaturated zone in bare soil condition. *Hydrological Sciences Journal*, **61**, 150527103244004, https://doi.org/10.1080/02626667.2014.966718.

Bennie, J., B. Huntley, A. Wiltshire, M. Hill, and R. Baxter, 2008: Slope, aspect and climate: Spatially explicit and implicit models of topographic microclimate in chalk grassland. *Ecological Modelling*, **216**, 47–59, https://doi.org/10.1016/j.ecolmodel.2008.04.010.

Bond, B. J., F. C. Meinzer, and J. R. Brooks, 2008: How Trees Influence the Hydrological Cycle in Forest Ecosystems. *Hydroecology and Ecohydrology*, P.J. Wood, D.M. Hannah, and J.P. Sadler, Eds., Wiley, 7–35.

Breiman, L., 2001: Random Forests. *Machine Learning*, **45**, 5–32, https://doi.org/10.1023/A:1010933404324.

Brocca, L., and Coauthors, 2014: Soil as a natural rain gauge: Estimating global rainfall from satellite soil moisture data. *Journal of Geophysical Research: Atmospheres*, **119**, 5128–5141, https://doi.org/10.1002/2014JD021489.

Carrijo, D. R., C. Li, S. J. Parikh, and B. A. Linquist, 2019: Irrigation management for arsenic mitigation in rice grain: Timing and severity of a single soil drying. *Science of The Total Environment*, **649**, 300–307, https://doi.org/10.1016/j.scitotenv.2018.08.216.

Charpentier, M. A., and P. M. Groffman, 1992: Soil moisture variability within remote sensing pixels. *Journal of Geophysical Research: Atmospheres*, **97**, 18987–18995, https://doi.org/10.1029/92JD00882.

Chen, S., D. She, L. Zhang, M. Guo, and X. Liu, 2019: Spatial Downscaling Methods of Soil Moisture Based on Multisource Remote Sensing Data and Its Application. *Water*, **11**, 1401, https://doi.org/10.3390/w11071401.

Cosh, M., T. Jackson, and R. Bindlish, 2004: Watershed scale temporal and spatial stability of soil moisture and its role in validating satellite estimates. *Remote Sensing of Environment*, **92**, 427–435, https://doi.org/10.1016/j.rse.2004.02.016.

Crow, W. T., and Coauthors, 2012: Upscaling sparse ground-based soil moisture observations for the validation of coarse-resolution satellite soil moisture products. *Reviews of Geophysics*, **50**, https://doi.org/10.1029/2011RG000372.

Cutler, A., D. R. Cutler, and J. R. Stevens, 2012: Random Forests. *Ensemble Machine Learning: Methods and Applications*, C. Zhang and Y. Ma, Eds., Springer US, 157–175.

Daly, C., M. Halbleib, J. I. Smith, W. P. Gibson, M. K. Doggett, G. H. Taylor, J. Curtis, and P. P. Pasteris, 2008: Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States. *International Journal of Climatology*, **28**, 2031–2064, https://doi.org/10.1002/joc.1688.

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

Engle, D., B. Coppedge, and S. Fuhlendorf, 2008: From the Dust Bowl to the Green Glacier: Human Activity and Environmental Change in Great Plains Grasslands. Vol. 196 of, 253–271.

Famiglietti, J. S., and Coauthors, 1999: Ground-based investigation of soil moisture variability within remote sensing footprints During the Southern Great Plains 1997 (SGP97) Hydrology Experiment. *Water Resources Research*, **35**, 1839–1851, https://doi.org/10.1029/1999WR900047.

Fang, B., V. Lakshmi, R. Bindlish, and T. J. Jackson, 2018: Downscaling of SMAP Soil Moisture Using Land Surface Temperature and Vegetation Data. *Vadose Zone Journal*, **17**, 170198, https://doi.org/10.2136/vzj2017.11.0198.

Ford, T. W., Q. Wang, and S. M. Quiring, 2016: The Observation Record Length Necessary to Generate Robust Soil Moisture Percentiles. *Journal of Applied Meteorology and Climatology*, **55**, 2131–2149, https://doi.org/10.1175/JAMC-D-16-0143.1.

Guevara, M., and R. Vargas, 2019: Downscaling satellite soil moisture using geomorphometry and machine learning. *PLoS ONE*, **14**, e0219639, https://doi.org/10.1371/journal.pone.0219639.

Huxman, T. E., and Coauthors, 2004: Precipitation pulses and carbon fluxes in semiarid and arid ecosystems. *Oecologia*, **141**, 254–268, https://doi.org/10.1007/s00442-004-1682-4.

I. D. Moore, G. J. Burch, and D. H. Mackenzie, 1988: Topographic Effects on the Distribution of Surface Soil Water and the Location of Ephemeral Gullies. *Transactions of the ASAE*, **31**, 1098–1107, https://doi.org/10.13031/2013.30829.

Im, J., S. Park, J. Rhee, J. Baik, and M. Choi, 2016: Downscaling of AMSR-E soil moisture with MODIS products using machine learning approaches. *Environ Earth Sci*, **75**, 1120, https://doi.org/10.1007/s12665-016-5917-6.

Jacobs, A. F. G., B. G. Heusinkveld, and S. M. Berkowicz, 2000: Dew measurements along a longitudinal sand dune transect, Negev Desert, Israel. *Int J Biometeorol*, **43**, 184–190, https://doi.org/10.1007/s004840050007.

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

Jiang, H., and W. Cotton, 2004: Soil moisture estimation using an artificial neural network: A feasibility study. *Canadian Journal of Remote Sensing*, **30**, https://doi.org/10.5589/m04-041.

Joshi, C., and B. P. Mohanty, 2010: Physical controls of near-surface soil moisture across varying spatial scales in an agricultural landscape during SMEX02. *Water Resources Research*, **46**, 2010WR009152, https://doi.org/10.1029/2010WR009152.

Keesstra, S. D., and Coauthors, 2016: The significance of soils and soil science towards realization of the United Nations Sustainable Development Goals. *SOIL*, **2**, 111–128, https://doi.org/10.5194/soil-2-111-2016.

Kim, C. P., J. N. M. Stricker, and R. A. Feddes, 1997: Impact of soil heterogeneity on the water budget of the unsaturated zone. *Water Resources Research*, **33**, 991–999, https://doi.org/10.1029/97WR00364.

Kim, G., and A. Barros, 2002: Downscaling of remotely sensed soil moisture with a modified fractal interpolation method using contraction mapping and ancillary data. *Remote Sensing of Environment*, **83**, 400–413, https://doi.org/10.1016/S0034-4257(02)00044-5.

Kim, M., J. Im, H. Han, J. Kim, S. Lee, M. Shin, and H.-C. Kim, 2015: Landfast sea ice monitoring using multisensor fusion in the Antarctic. *GIScience & Remote Sensing*, **52**, 239–256, https://doi.org/10.1080/15481603.2015.1026050.

Krueger, E. S., T. E. Ochsner, and S. M. Quiring, 2019: Development and Evaluation of Soil Moisture-Based Indices for Agricultural Drought Monitoring. *Agronomy Journal*, **111**, 1392–1406, https://doi.org/10.2134/agronj2018.09.0558.

Kunkel, K. E., L. E. Stevens, S. E. Stevens, and L. Sun, Regional Climate Trends and Scenarios for the U.S. National Climate Assessment Part 4. Climate of the U.S. Great Plains.

Li, M., J. Im, and C. Beier, 2013: Machine learning approaches for forest classification and change analysis using multi-temporal Landsat TM images over Huntington Wildlife Forest. *GIScience & Remote Sensing*, **50**, 361–384, https://doi.org/10.1080/15481603.2013.819161.

Li, X., W. Chen, X. Cheng, and L. Wang, 2016: A Comparison of Machine Learning Algorithms for Mapping of Complex Surface-Mined and Agricultural Landscapes Using ZiYuan-3 Stereo Satellite Imagery. *Remote Sensing*, **8**, 514, https://doi.org/10.3390/rs8060514.

https://doi.org/10.5194/egusphere-2025-896
Preprint. Discussion started: 7 April 2025

Liu, J., F. Rahmani, K. Lawson, and C. Shen, 2022: A multiscale deep learning model for soil moisture integrating satellite and in-situ data. *Geophysical Research Letters*, **n/a**, e2021GL096847, https://doi.org/10.1029/2021GL096847.

Liu, Q., and Coauthors, 2011: The Contributions of Precipitation and Soil Moisture Observations to the Skill of Soil Moisture Estimates in a Land Data Assimilation System. *Journal of Hydrometeorology*, **12**, 750–765.

Liu, Y., W. Jing, Q. Wang, and X. Xia, 2020: Generating high-resolution daily soil moisture by using spatial downscaling techniques: a comparison of six machine learning algorithms. *Advances in Water Resources*, **141**, 103601, https://doi.org/10.1016/j.advwatres.2020.103601.

Lu, Z., J. Im, J. Rhee, and M. Hodgson, 2014: Building type classification using spatial and landscape attributes derived from LiDAR remote sensing data. *Landscape and Urban Planning*, **130**, 134–148, https://doi.org/10.1016/j.landurbplan.2014.07.005.

Mohanty, B. P., and T. H. Skaggs, 2001: Spatio-temporal evolution and time-stable characteristics of soil moisture within remote sensing footprints with varying soil, slope, and vegetation. *Advances in Water Resources*, **24**, 1051–1067, https://doi.org/10.1016/S0309-1708(01)00034-3.

Mohanty, B. P., T. H. Skaggs, and J. S. Famiglietti, 2000: Analysis and mapping of field-scale soil moisture variability using high-resolution, ground-based data during the Southern Great Plains 1997 (SGP97) Hydrology Experiment. *Water Resources Research*, **36**, 1023–1031, https://doi.org/10.1029/1999WR900360.

Nielsen, M., B. I. Cook, K. Marvel, M. Ting, and J. E. Smerdon, 2024: The Changing Influence of Precipitation on Soil Moisture Drought With Warming in the Mediterranean and Western North America. *Earth's Future*, **12**, e2023EF003987, https://doi.org/10.1029/2023EF003987.

Oren, R., J. S. Sperry, G. G. Katul, D. E. Pataki, B. E. Ewers, N. Phillips, and K. V. R. Schäfer, 1999: Survey and synthesis of intra- and interspecific variation in stomatal sensitivity to vapour pressure deficit. *Plant, Cell & Environment*, **22**, 1515–1526, https://doi.org/10.1046/j.1365-3040.1999.00513.x.

Panciera, R., J. P. Walker, J. D. Kalma, E. J. Kim, J. M. Hacker, O. Merlin, M. Berger, and N. Skou, 2008: The NAFE'05/CoSMOS Data Set: Toward SMOS Soil Moisture Retrieval, Downscaling, and Assimilation. *IEEE Trans. Geosci. Remote Sensing*, **46**, 736–745, https://doi.org/10.1109/TGRS.2007.915403.

Park, S., J. Im, E. Jang, and J. Rhee, 2016: Drought assessment and monitoring through blending of multi-sensor indices using machine learning approaches for different climate regions. *Agricultural and Forest Meteorology*, **216**, 157–169, https://doi.org/10.1016/j.agrformet.2015.10.011.

Peng, J., A. Loew, O. Merlin, and N. E. C. Verhoest, 2017: A review of spatial downscaling of satellite remotely sensed soil moisture: Downscale Satellite-Based Soil Moisture. *Rev. Geophys.*, **55**, 341–366, https://doi.org/10.1002/2016RG000543.

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

https://doi.org/10.5194/egusphere-2025-896
Preprint. Discussion started: 7 April 2025

Rhee, J., S. Park, and Z. Lu, 2014: Relationship between land cover patterns and surface temperature in urban areas. *GIScience & Remote Sensing*, **51**, 521–536, https://doi.org/10.1080/15481603.2014.964455.

Robinson, D. A., and Coauthors, 2008: Soil Moisture Measurement for Ecological and Hydrological Watershed-Scale Observatories: A Review. *Vadose Zone Journal*, **7**, 358–389, https://doi.org/10.2136/vzj2007.0143.

Ruosteenoja, K., T. Markkanen, and J. Räisänen, 2020: Thermal seasons in northern Europe in projected future climate. *International Journal of Climatology*, **40**, 4444–4462, https://doi.org/10.1002/joc.6466.

Salvucci, G., 2001: Estimating the moisture dependence of root zone water loss using conditionally averaged precipitation. *Water Resources Research - WATER RESOUR RES*, **37**, 1357–1366, https://doi.org/10.1029/2000WR900336.

Seager, R., and Coauthors, 2007: Model Projections of an Imminent Transition to a More Arid Climate in Southwestern North America. *Science*, **316**, 1181–1184, https://doi.org/10.1126/science.1139601.

——, A. Tzanova, and J. Nakamura, 2009: Drought in the Southeastern United States: Causes, Variability over the Last Millennium, and the Potential for Future Hydroclimate Change. https://doi.org/10.1175/2009JCLI2683.1.

Seneviratne, S. I., T. Corti, E. L. Davin, M. Hirschi, E. B. Jaeger, I. Lehner, B. Orlowsky, and A. J. Teuling, 2010: Investigating soil moisture–climate interactions in a changing climate: A review. *Earth-Science Reviews*, **99**, 125–161, https://doi.org/10.1016/j.earscirev.2010.02.004.

Sivapalan, M., K. Beven, and E. F. Wood, 1987: On hydrologic similarity: 2. A scaled model of storm runoff production. *Water Resources Research*, **23**, 2266–2278, https://doi.org/10.1029/WR023i012p02266.

Srivastava, P. K., D. Han, M. R. Ramirez, and T. Islam, 2013: Machine Learning Techniques for Downscaling SMOS Satellite Soil Moisture Using MODIS Land Surface Temperature for Hydrological Application. *Water Resour Manage*, **27**, 3127–3144, https://doi.org/10.1007/s11269-013-0337-9.

Sun, H., and Y. Cui, 2021: Evaluating Downscaling Factors of Microwave Satellite Soil Moisture Based on Machine Learning Method. *Remote Sensing*, **13**, 133, https://doi.org/10.3390/rs13010133.

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

Wagner, W., and Coauthors, 2013: The ASCAT Soil Moisture Product: A Review of its Specifications, Validation Results, and Emerging Applications. *Meteorologische Zeitschrift*, **22**, 5–33, https://doi.org/10.1127/0941-2948/2013/0399.

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

https://doi.org/10.5194/egusphere-2025-896
Preprint. Discussion started: 7 April 2025

Warner, D., M. Guevara, J. Callahan, and R. Vargas, 2021: Downscaling satellite soil moisture for landscape applications: A case study in Delaware, USA. *Journal of Hydrology: Regional Studies*, **38**, 100946, https://doi.org/10.1016/j.ejrh.2021.100946.

Xu, M., N. Yao, H. Yang, J. Xu, A. Hu, L. Goncalves, and G. Liu, 2022: Downscaling SMAP soil moisture using a Wide & Deep learning method over the Continental United States. *Journal of Hydrology*, **609**, 127784, https://doi.org/10.1016/j.jhydrol.2022.127784.

Xu, W., Z. Zhang, Z. Long, and Q. Qin, 2021: Downscaling SMAP Soil Moisture Products With Convolutional Neural Network. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, **14**, 4051–4062, https://doi.org/10.1109/JSTARS.2021.3069774.

Zhang, N., S. M. Quiring, and T. W. Ford, 2019: *Blending SMAP, Noah and In Situ Soil Moisture Using Multiple Methods*. Vadose Zone Hydrology/Modelling approaches,.

Zhao, W., N. Sánchez, H. Lu, and A. Li, 2018: A spatial downscaling approach for the SMAP passive surface soil moisture product using random forest regression. *Journal of Hydrology*, **563**, 1009–1024, https://doi.org/10.1016/j.jhydrol.2018.06.081.

Zhou, X., H. S. Lin, and E. A. White, 2008: Surface soil hydraulic properties in four soil series under different land uses and their temporal changes. *CATENA*, **73**, 180–188, https://doi.org/10.1016/j.catena.2007.09.009.

Dew - Monteith - 1957 - Quarterly Journal of the Royal Meteorological Society - Wiley Online Library. https://rmets.onlinelibrary.wiley.com/doi/10.1002/qj.49708335706 (Accessed June 6, 2024).