Downscaling Satellite Soil Moisture using Geomorphometry and Machine Learning

Authors:

Mario Guevara[1] and Rodrigo Vargas[1*]

[1]University of Delaware, Department of Plant and Soil Sciences, Newark, DE, 19716

[*]Corresponding author
E-mail: rvargas@udel.edu

Abstract

Annual soil moisture estimates are useful to characterize trends in the climate system, in the capacity of soils to retain water and for predicting land and atmosphere interactions. The main source of soil moisture spatial information across large areas (e.g., continents) is satellite-based microwave remote sensing. However, satellite soil moisture datasets have coarse spatial resolution (e.g., 25-50 km grids); and large areas from regional-to-global scales have spatial information gaps. We provide an alternative approach to predict soil moisture spatial patterns (and associated uncertainty) with higher spatial resolution across areas where no information is otherwise available. This approach relies on geomorphometry derived terrain parameters and machine learning models to improve the statistical accuracy and the spatial resolution (from 27km to 1km grids) of satellite soil moisture information across the conterminous United States on an annual basis (1991-2016). We derived 15 primary and secondary terrain parameters from a digital elevation model. We trained a machine learning algorithm (i.e., kernel weighted nearest neighbors) for each year. Terrain parameters were used as predictors and annual satellite soil moisture estimates were used to train the models. The explained variance for all models-years was >70% (10-fold cross-validation). The 1km soil moisture grids (compared to the original satellite soil moisture estimates) had higher correlations with field soil moisture observations from the North American Soil Moisture Database (n=668 locations with available data between 1991-2013; 0-5cm depth) than the original product. We conclude that the fusion of geomorphometry methods and satellite soil moisture estimates is useful to increase the spatial resolution and accuracy of satellite-derived soil moisture. This approach can be applied to other satellite-derived soil moisture estimates and regions across the world.

Keywords: Geomorphometry; terrain parameters; machine learning; satellite soil moisture; downscaling; uncertainty.

Introduction

Continuous national to continental scale soil moisture information is increasingly needed to characterize spatial and temporal trends of terrestrial productivity patterns (e.g., production of food, fiber and energy). This is because soil moisture is a key variable regulating hydrological and biogeochemical cycles, and thus studying its spatial-temporal dynamics is crucial for assessing the potential impact of climate change on water resources [1-4]. Currently, the most feasible way to obtain national to continental soil moisture information is using remote sensing. Microwave remote sensing devices deployed on multiple earth observation satellites are able to quantify the dielectric constant of soil surface and retrieve soil moisture estimates [5]. However, there are spatial gaps of satellite-based soil moisture information and its current spatial resolution (> 1km grids) limits its applicability at the ecosystem-to-landscape scales to address the ecological implications of soil moisture dynamics [5-8].

Satellite soil moisture records are an effective indicator for monitoring global soil conditions and forecasting climate impacts on terrestrial ecosystems, because soil moisture estimates are required for assessing feedbacks between water and biogeochemical cycles [9-12]. In addition, accurate soil moisture information is critical to predict terrestrial and atmospheric interactions such as water evapotranspiration or CO_2 emissions from soils [3, 13-15]. However, soil moisture information at spatial resolution of 1x1km pixels or less is not yet available across large areas of the world and the coarse pixel size (>1km pixels) of available satellite soil moisture records is limited for spatial analysis (i.e., hydrological, ecological) at small regional levels (e.g., county- to state). In addition, satellite soil moisture estimates are representative only of the first few 0-5 to 10 cm of top-soil surface [16]. Therefore, comparing multiple sources for satellite soil moisture and field soil moisture estimates is constantly required for precise interpretations of soil moisture spatial patterns [17-19].

There is a pressing need for exploring statistical relationships across different sources of remote sensing information (e.g., topography and soil moisture) and developing alternative soil moisture spatial datasets (i.e., grids) to improve the continental-to-global spatial representation of soil moisture estimates [7]. Spatially explicit soil moisture estimates can be obtained across large areas with a spatial resolution between 25-50 km grids from radar-based microwave platforms deployed across different satellite soil moisture missions [20-21]. The availability of historical soil moisture records of these sources has increased during the last decade with unprecedented levels of temporal resolution (i.e., daily from years 1978-present) at the global scale. However, large areas constantly covered by snow, extremely dry regions or tropical rain forests (where there is a higher content of water above ground) lack of precise soil moisture satellite records due to sensor intrinsic limitations (e.g., saturation or noise) across these environmental conditions [22].

One valuable product that is affected by the aforementioned environmental conditions is the ESA-CCI (European Space Agency Climate Change Initiative) soil moisture product [20-21]. The ESA-CCI mission makes rapidly available long-term soil moisture estimates with daily temporal resolution from the 1978s to date, and it represents the state-of-the-art knowledge tool for assessing long term trends in the climate system. Modeling, validation and calibration frameworks are required for improving the spatial representation of this important dataset, and for predicting soil moisture patterns across areas where no satellite estimates are available.

Currently, there is an increasing availability of fine-gridded information sources and modeling approaches that could be used for increasing the spatial resolution (hereinafter downscaling) of the ESA-CCI satellite soil moisture estimates (e.g., soil moisture predictions across <1x1km grids). Downscaling (and subsequently gap-filling) satellite soil moisture estimates has been the objective of empirical modeling approaches based on sub-grids of soil moisture related information such as soil texture [23]. Other approaches followed environmental correlation methods and generated soil moisture predictions for satellite soil moisture estimates using both data-driven or hypothesis driven models and multiple sub-grids of ancillary information [24-26]. These sub-grids of information usually include vegetation related optical remote sensing imagery, gridded soil information, land cover classes and landforms [27-30]. Most of these approaches have been tested for specific study sites. Other studies have focused on applying a digital soil mapping approach (a reference framework for understanding the spatial distribution of soil variability [31]) and multiple upscaling methods for predicting soil moisture patterns at the continental scale [26, 32]; and an overview of multiple approaches for downscaling satellite soil moisture (e.g., empirically based, physically based) has been previously discussed [33]. Here, we propose that digital terrain analysis (i.e., geomorphometry) can also be applied for empirically downscaling soil moisture satellite-based information across continental-to-global spatial scales.

Geomorphometry is an emergent discipline in earth sciences dedicated to the quantitative analysis of land surface characteristics and topography [34-35]. Topography includes a diversity of hydrologically meaningful terrain parameters (i.e., slope, aspect, curvature) that aim to represent how the landscape physically constrains water inputs (e.g., rainwater, irrigation, overland flow) that reaches the soil surface [35-36]. At the landscape scale, soil moisture is partially controlled by topography related factors (i.e., slope, aspect, curvature) physically constraining soil water inputs and soil hydraulic properties (e.g., soil texture, structure). Based on these geomorphometry principles [35-39], we propose that it is possible to determine which terrain parameters are the strongest predictors of the spatial variability of satellite soil moisture. Statistically coupling the spatial variability of satellite soil moisture with hydrologically meaningful terrain parameters could be an alternative way to improve the spatial resolution and accuracy of satellite soil moisture estimates across the continental scale. This is possible because topography (represented by terrain parameters) directly affects: 1) the angle of the satellite microwave signal at the soil surface; and 2) the overall distribution of water in the landscape.

Topography is a major driver for soil moisture and topography surrogates (e.g., land form or elevation map) have been combined with other variables (e.g., climate, soils, vegetation and land use) for downscaling satellite soil moisture estimates [33]. However, the exclusive use of geomorphometry-derived products for downscaling satellite soil moisture has not yet been explored from national-to-continental scales. This approach is relevant to avoid statistical redundancies and potential spurious correlations when downscaled soil moisture is further used or analyzed with vegetation- or climate-related variables (when these aforementioned variables were used for downscaling of satellite derived soil moisture). In this study, we show the potential of a soil moisture prediction framework purely based on geomorphometry derived products (digital terrain parameters).

Our main objective is to generate a soil moisture prediction framework by coupling satellite soil moisture estimates with hydrologically meaningful terrain parameters as prediction factors. Coupling the complexity of topographic gradients and the multi-temporal nature of satellite soil moisture requires an approach that should account for non-linear relationships. Machine learning approaches could account for non-linearity based on probability and the ability of computer systems to reproduce and 'learn' (i.e., decide the best solution after multiple model realizations) from multiple modeling outputs (i.e., varying model parameters of combinations of training and testing random samples) [40]. Furthermore, machine learning is now a common component of geoscientific research leading the discovery of new knowledge in the earth system [41] including soil greenhouse gas fluxes [8, 42] and soil moisture estimates [43].

We postulate that the data fusion between satellite soil moisture with hydrologically meaningful terrain parameters can enhance the spatial resolution, representativeness and quality (i.e., accuracy) of current coarse satellite soil moisture grids. We focus on the conterminous United States (CONUS) given the large availability of soil moisture records for validating purposes from the North American Soil Moisture Database (NASMD) [44]. This study provides insights for obtaining detailed soil moisture estimates relying on public sources of satellite information and a data-driven framework that could be reproduced and applied across the world. The novelty of this research relies on proposing an alternative approach for obtaining soil moisture gridded measurements across areas where no soil moisture information is available (i.e., from the ESA-CCI) and at a spatial resolution (i.e., 1km) determined by the topographic prediction factors. This approach can be applied to other satellite-derived soil moisture estimates and regions across the world.

Materials and Methods

Our downscaling approach relied on a Digital Elevation Model (DEM), and satellite soil moisture records. Soil moisture information was acquired from the ESA-CCI [20-21]. The development and reliability (i.e., validation) of this remote sensing soil moisture product has been documented by previous studies [20-21, 45]. Our framework includes prediction factors for soil moisture from digital terrain analysis. These terrain predictors were derived across CONUS using 1km grids. Machine learning was used for generating soil moisture predictions (annual, 1991-2016) using as training data the satellite soil moisture estimates provided by the ESA-CCI. Field soil moisture observations from the North American Soil Moisture dataset were used for validating the soil moisture predictions based on digital terrain analysis (Figure 1).

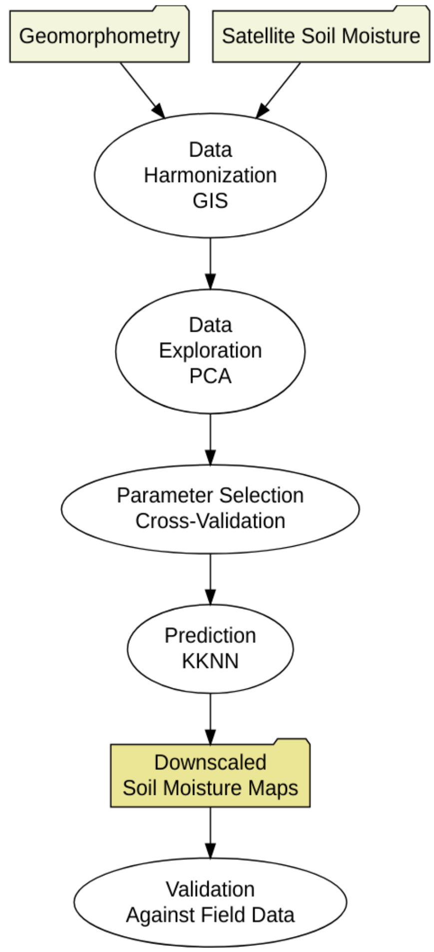

Figure 1: Soil moisture prediction framework. The folders are the inputs and outputs and the ovals are methods for data preparation (data bases harmonization), modeling (for prediction) and validation (for assessing the reliability of soil moisture maps).

Datasets and data preparation

We downscaled the ESA-CCI satellite soil moisture estimates between 1991 and 2016 and validated the downscaled information with field measurements (Supplementary Figure S1). The ESA-CCI soil moisture product has a daily temporal coverage from 1978 to 2016 and a spatial resolution of ~27 km (Supplementary Figure S2). Among several remotely sensed soil moisture products [16, 46-50], we decided to use the ESA-CCI soil moisture product because it covers a larger period of time compared with other satellite soil moisture products (e.g., NASA SMAP). We highlight that satellite soil moisture information is used for training a machine learning model for each year, and independent field soil moisture records area only used for validating the downscaled soil moisture predictions.

For externally validating, we used the NASMD because it has been curated following a strict quality control calibrated for CONUS [44] (Supplementary Figure S1). This data collection effort consists of a harmonized and quality-controlled soil moisture dataset with contributions from over 2000 meteorological stations across CONUS described by Quiring and colleagues [44]. The NASMD also include records of soil moisture registered in the International Soil Moisture Network (ISMN) [44, 51]. The NASMD (unlike the ISMN) provides processed data from each station location in each network [44-19]. We used soil moisture records at 5 cm of depth (n = 5541 daily measurements) from 668 stations with available soil moisture estimates at this depth because radar-based soil moisture estimates are representative for these first few centimeters of topsoil surface [16].

As prediction factors for soil moisture, we calculated hydrologically meaningful terrain parameters for CONUS using information from a radar-based DEM [52-53]. These terrain parameters are quantitative spatial grids representing the topographic variability that directly influence the water distribution across the landscape [35], which supports the physical link between soil moisture and topography. These parameters were the basis for downscaling satellite soil moisture records to 1km grids. This spatial resolution captures the major variability of topographic features across CONUS and is commonly used on large-scale ecosystem studies and soil mapping efforts [53-54].

For the calculation of soil moisture prediction factors, we used automated digital terrain analysis using the System for Automated Geographical Analysis-Geographical Information System (SAGA-GIS) [36]. The automated implementation of SAGA-GIS for Geomorphometry (module for basic terrain analysis) includes a preprocessing stage to remove spurious sinks and reduce the presence of other artifacts in the elevation gridded surface (e.g., false pikes or flat areas). After preprocessing the DEM, 15 hydrologically meaningful terrain parameters were generated for the CONUS from elevation data including primary (i.e., slope, aspect) and secondary parameters (i.e., cross-sectional curvature, longitudinal curvature, analytical hill-shading, convergence index, closed depressions, catchment area, topographic wetness index, length-slope factor, channel network base level, vertical distance to channel network, and valley depth index; Figure 2). The values of these terrain parameters (Supplementary Table S1) were harmonized with the ESA-CCI soil moisture values using as reference the central coordinates of the coarse soil moisture grids (Figure 1 *inputs*; see section 2.3).

Data exploration

We used a principal component analysis (PCA) prior to modeling for data exploration and description of general relationships between soil moisture values and topography (represented by the aforementioned terrain parameters). The purpose was to simplify the dimensionality of the data set to identify the main relationships (between soil moisture and topographic parameters) driving our downscaling framework (Figure 1 *methods*). The PCA was implemented as in previous work [55], based on a reference value representing the 0.95-quantile of the variability obtained by randomly simulating 300 data tables of equivalent size on the basis of a normal distribution. This analysis was applied to the terrain parameters at the locations of the field stations in order to compare the relationship of the first PCA and the values of soil moisture from the ESA-CCI grids and from the field data.

Model building

For this analysis we built a model for each downscaled soil moisture map. We used a machine learning kernel-based model (kernel weighted nearest neighbors, kknn) [56-57] to downscale satellite soil moisture (Figure 1 *methods*). The training dataset for each model/year were the annual values of the ESA-CCI soil moisture product. The kknn model has two main model parameters: the optimum number of neighbors (k) and the optimal kernel function (okf). First, we defined k, which is the number of neighbors to be considered for the prediction. Second, we selected the okf, which is a reference (e.g., triangular, epanechnikov, Gaussian, optimal) for the probability density function of the variable to be predicted. The okf is used to convert distances (i.e., Minkowski distance) into weights used to calculate the k-weighted average. These kknn model parameters (k and okf) were selected by the means of 10-fold cross validation as previously recommended [58]. Cross-validation is a well-known re-sampling technique that divides data into 10 roughly equal subsets. For every possible parameter value (e.g., *k* from 1 to 50 and okf [triangular, epanechnikov, Gaussian, optimal]), 10 different models are generated, each using 90% of the data then being evaluated on the remaining 10%. To predict soil moisture information at 1 km of spatial resolution for each year (between 1991 and 2016), we selected the combination of optimal k and okf that lead to the highest correlation (between observed and predicted data) with the lowest root mean squared error (RMSE) after the cross-validation strategy. Thus, for each year we were able to predict soil moisture across 1x1 km grids (Figure 1 *outputs*).

Validation using field observations across CONUS

Downscaled soil moisture grids were compared against field measurements and we computed the explained variance (r^2) using a linear fit (observed *vs* predicted) for each field soil moisture location. Given the relatively low density and sparse spatial distribution of field data for validating (Supplementary Figure S1), we bootstrapped the independent validation using different sample sizes (from 10 to 100% of data with increments each 10%) to avoid systematic bias associated with the spatial distribution and density of field soil moisture information. We sampled (n = 1000) repeatedly the original and the downscaled soil moisture grids aiming to identify their correlation with the aforementioned validation dataset (i.e., observed vs predicted).

We also computed the spatial structure (spatial autocorrelation) of the explained variance (correlation between geographical distance and variance of r^2 values) for estimating an r^2 map using an interpolation technique known in geostatistics as Ordinary Kriging [59]. Ordinary Kriging is a well-known method for spatial interpolation based on the spatial structure or spatial autocorrelation of the variable of interest (the r^2 values between the field observations and the predicted soil moisture values). The spatial autocorrelation is defined by the relationship between geographical distances and variance of values at a given distance, and it is commonly characterized using variograms. We followed an automated variogram parameterization (the optimal selection for the variogram parameters nugget, sill and range required to perform Ordinary Kriging) proposed in previous work [60].

As implemented in the automap package of R [60], the initial sill is estimated as the mean of the maximum and the median values of the semi-variance. The semi-variance is defined by the variance within multiple distance intervals. For modeling the spatial autocorrelation this algorithm iterates over multiple variogram model parameters selecting the model (e.g., spherical, exponential, Gaussian) that has the smallest residual sum of squares with the sample variogram. The initial range is defined as 0.10 times the diagonal of the bounding box of the data. The initial nugget is defined as the minimum value of the semi-variance. Thus, the parameters used for obtaining a continuous map showing spatial trends in the r^2 were: a Gaussian (normal) model form, a nugget value of 0.06 $m^3\ m^{-3}$, a sill of 0.08 $m^3\ m^{-3}$ and an approximate range of 428.7 km. This map was generated because it could provide insights about overall sources of modeling errors (e.g., environmental similarities in multiple areas showing low or high explained variance) and their spatial distribution. All analyzes were performed in R [61] using public sources of data. A reproducible example (code) for generating the soil moisture predictions in 1km grids is provided as Supplementary Information S1.

Results

The exploratory PCA showed that the first two PCs explained 33% of the total dataset variability (Supplementary Figure S3A), where the first PC explained 18% of total variability and at least five PCs were needed to explain 70% of total variability. The first PC was best correlated with elevation (r=0.82) and with the vertical distance to channel network (r=0.88). Elevation varied negatively with soil moisture, as well as other secondary terrain parameters such as the base level channel network elevation (distance from each pixel to the closer highest point), while the valley depth index varied positively with soil moisture (Supplementary Figure S3B). The relative slope position (indicating the dominance of flat or complex terrain) and the topographic wetness index (which indicates areas where water tends to accumulate) were also correlated with soil moisture across the first 5 PCs. Thus, multiple terrain parameters varied positively and negatively with soil moisture values (Supplementary Information S2).

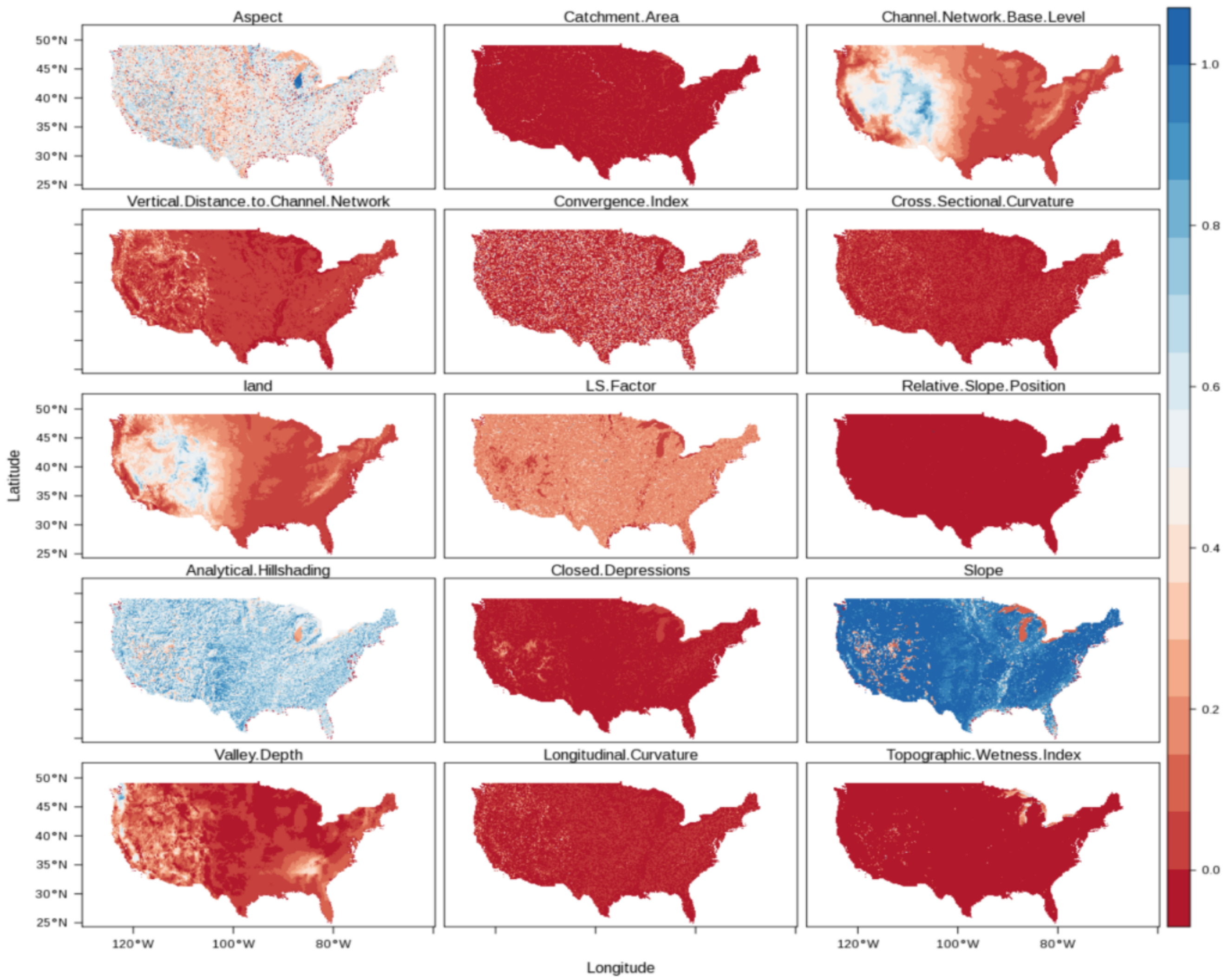

Figure 2. Elevation and hydrologically meaningful terrain parameters at 1x1km of spatial resolution derived using standard the SAGA-GIS basic terrain parameters module. These maps were normalized (between 0-1) and then used as prediction factors to downscale soil moisture across CONUS.

Our framework to predict soil moisture based on topography and remote sensing was able to explain, on average 79±0.1% of the variability of satellite soil moisture information as revealed by the cross-validation strategy. The root mean squared error (RMSE) derived from the cross-validation varied around 0.03 m^3/m^3, while the percentage of explained variance was in all cases above 70% (Table 1).

Table 1. The cross-validation results for each year. This table shows the correlation, root mean squared error (RMSE), the number of training data available (n), the optimal kernel function (okf), and the optimal number of neighbors used for predicting to new data (k).

Model	Year	Correlation	RMSE	n	okf	k
1	1991	0.85	0.03	18058	triangular	18
2	1992	0.89	0.03	18429	triangular	16
3	1993	0.88	0.03	18107	triangular	18
4	1994	0.9	0.03	18367	triangular	16
5	1995	0.88	0.03	18385	triangular	18
6	1996	0.9	0.03	18454	triangular	15
7	1997	0.88	0.03	18428	triangular	15
8	1998	0.88	0.03	18540	triangular	16
9	1999	0.89	0.03	18542	triangular	15
10	2000	0.9	0.03	18547	triangular	15
11	2001	0.9	0.03	18523	triangular	15
12	2002	0.9	0.03	19170	triangular	16
13	2003	0.89	0.03	19132	triangular	16
14	2004	0.89	0.03	18934	triangular	16
15	2005	0.89	0.03	19132	triangular	16
16	2006	0.9	0.03	19131	triangular	16
17	2007	0.88	0.03	19142	triangular	16
18	2008	0.9	0.03	19136	triangular	16
19	2009	0.9	0.03	19142	triangular	16
20	2010	0.88	0.03	19245	triangular	18

21	2011	0.9	0.03	19255	triangular	18
22	2012	0.9	0.03	19252	triangular	16
23	2013	0.89	0.03	19226	triangular	16
24	2014	0.89	0.03	19227	triangular	16
25	2015	0.88	0.03	19231	triangular	16
26	2016	0.88	0.03	19225	triangular	16

By applying the model coefficients to the topographic prediction factors across CONUS, we generated 26 cross-validated maps (for years 1991-2016) of mean annual soil moisture estimates in 1x1km grids (Figure 3). The downscaled product shows a higher level of spatial variability due the increased spatial detail achieved by downscaling soil moisture to 1x1km grids (Supplementary Figure S4). Our predictions reveal a clear bimodal distribution of soil moisture values (e.g., from the east to the west, Figure 4) which is also evident in the original estimate (Supplementary Figure S5). The statistical comparison (squared correlation) between the original product and the downscaled product suggests a high level of agreement showing an r^2 value of 0.72.

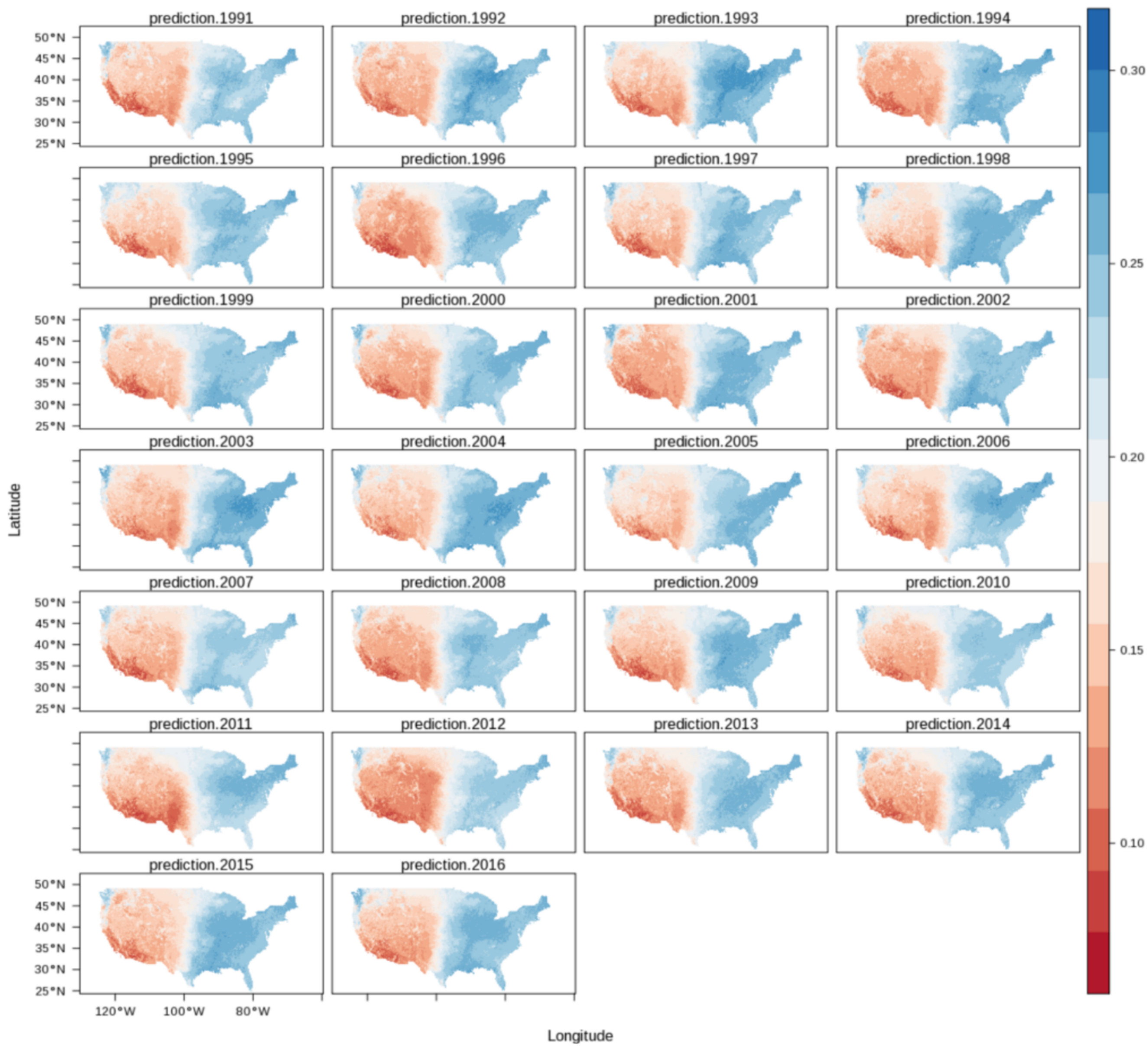

Figure 3. Annual means of soil moisture (1991-2016) downscaled to 1x1km grids across CONUS using terrain parameters as prediction factors.

We provided a visual comparison between the original satellite estimate and the downscaled results including both median and standard deviation values (Figure 4). We also show the uncertainty of the original soil moisture product as reported by its developers and the r^2 map from the validation against field stations. The r^2 map shows the lowest values across the Central Plains of the US and the lower Mississippi basin. The lower values in the r^2 map are consistent with the high uncertainty values of the original satellite estimate (Figure 4).

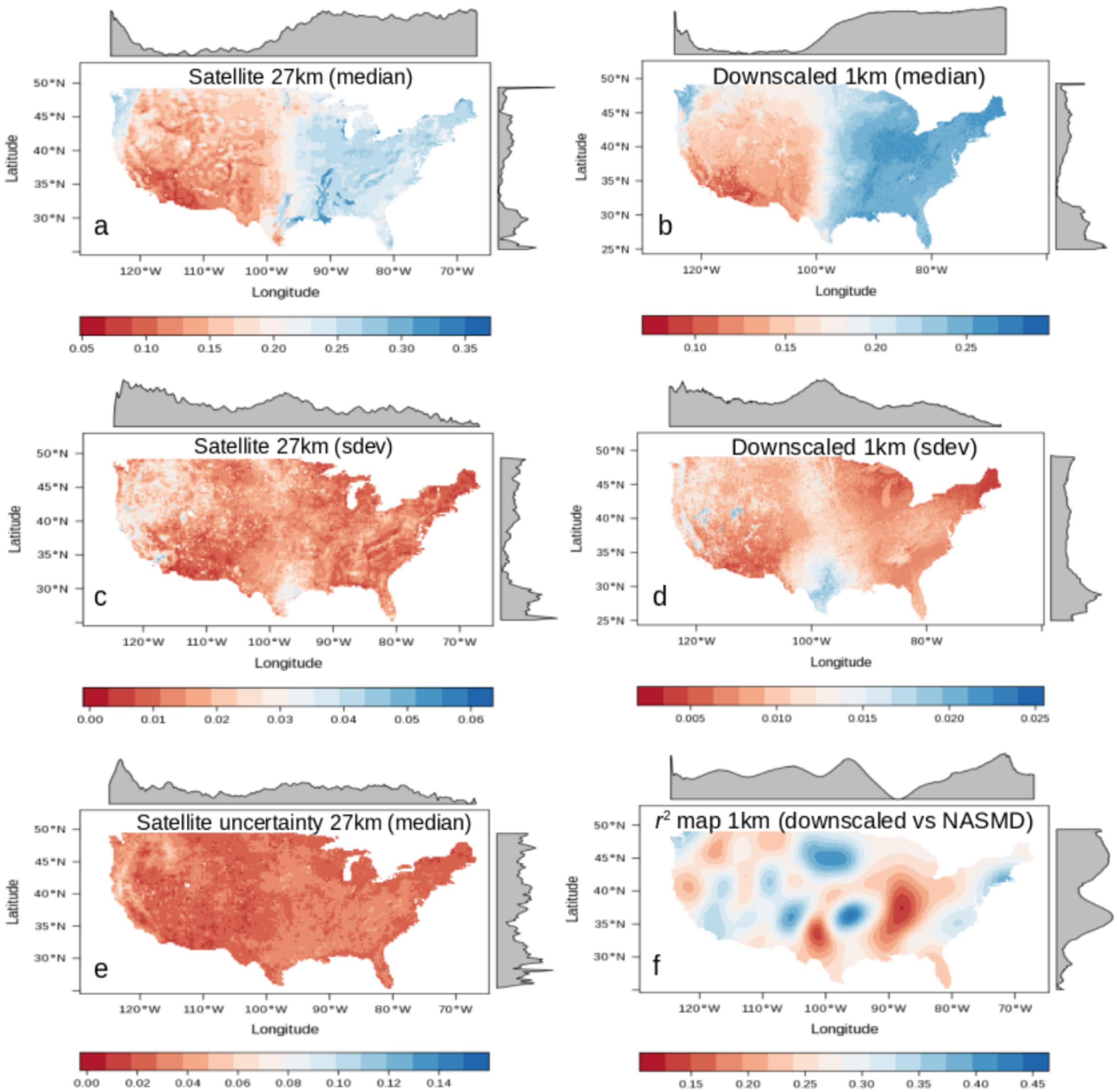

Figure 4. Comparison of the original and the downscaled soil moisture products. Median (a, b) and standard deviation (c, d, sdev) values of satellite soil moisture and downscaled soil moisture values

(1991-2016). We show the uncertainty reported by the ESA-CCI soil moisture (e) and the explained variance map (r^2) between field data and downscaled soil moisture (f).

The r^2 map provided insights about the relationship between soil moisture gridded surfaces and soil moisture field data. Higher r^2 values were found across the east coast, the Northern Plains and water-limited environments across the western states. We found that our soil moisture downscaled output better correlates (nearly 25% improvement) with NASMD field observations when compared to the original soil moisture satellite estimates (Figure 5).

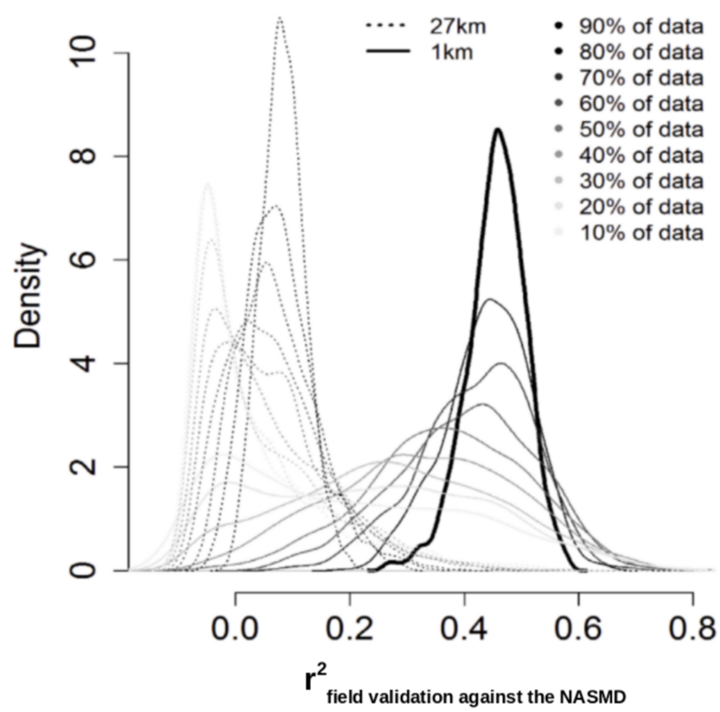

Figure 5. Validation of soil moisture gridded estimates (original 27 and 1km grids)

against NASMD field observations. Dashed line represents the relationship of field stations and soil moisture gridded estimates at 27x27km, while black line represents the relationship between field stations and the downscaled 1x1km soil moisture product. In all cases (all sample sizes), the 1x1km product showed higher r^2 with the NASMD than the ESA-CCI soil moisture estimates.

This improvement was consistent after repeating it using random samples and different sample sizes (from 10 to 90 % of available validation data) from the NASMD field observations (Figure 5). However, there is a sparse distribution of validation data and large areas of CONUS lack of field information for validating/calibrating soil moisture predictions (Figure 6). Considering the quality-controlled records available from the NASMD across CONUS and the coarse scale of the ESA-CCI soil moisture product, our approach suggests an improvement in the spatial resolution (from 27 to 1km grids) of soil moisture estimates while maintaining the integrity of the original satellite values.

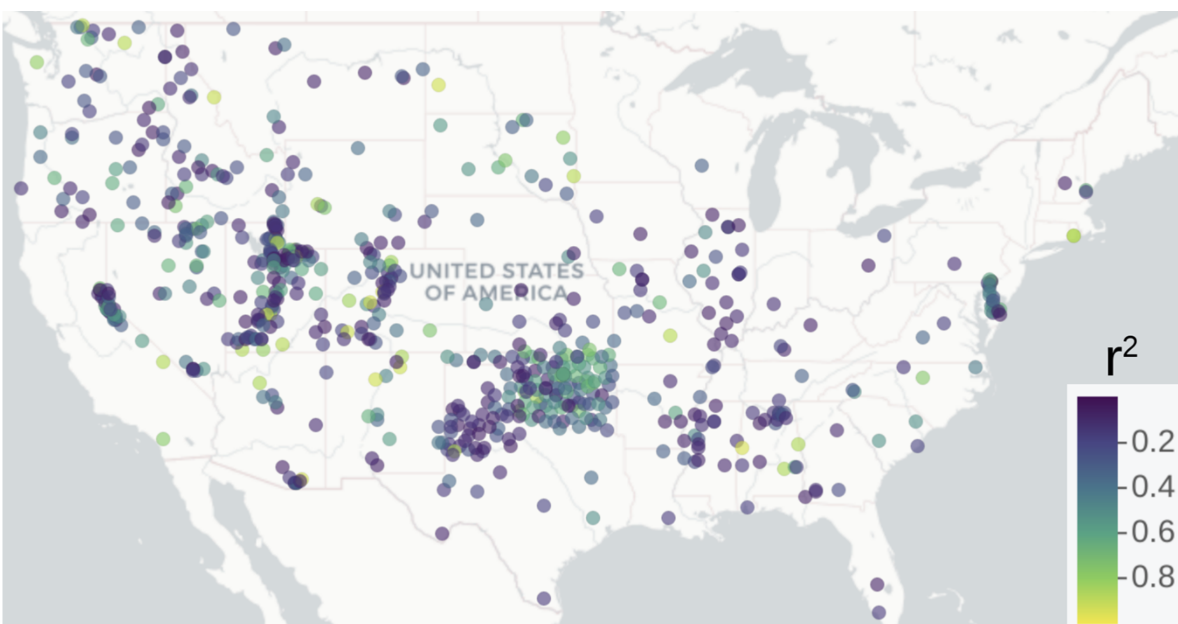

Figure 6 Explained variances computed for each meteorological station of the NASMD and the corresponding pixel of our soil moisture predictions based on geomorphometry.

The original satellite values, the downscaled product and the ISMN dataset showed a similar correlation with the terrain predictors. For example, the first PCA (represented by the distance to channel network and elevation), was negatively correlated with field soil moisture, the satellite original product and our soil moisture predictions. The correlation values were r=-0.17, r=-0.27, and r=-0.28 respectively. These relationships showed a similar pattern in the statistical space (Figure 7).

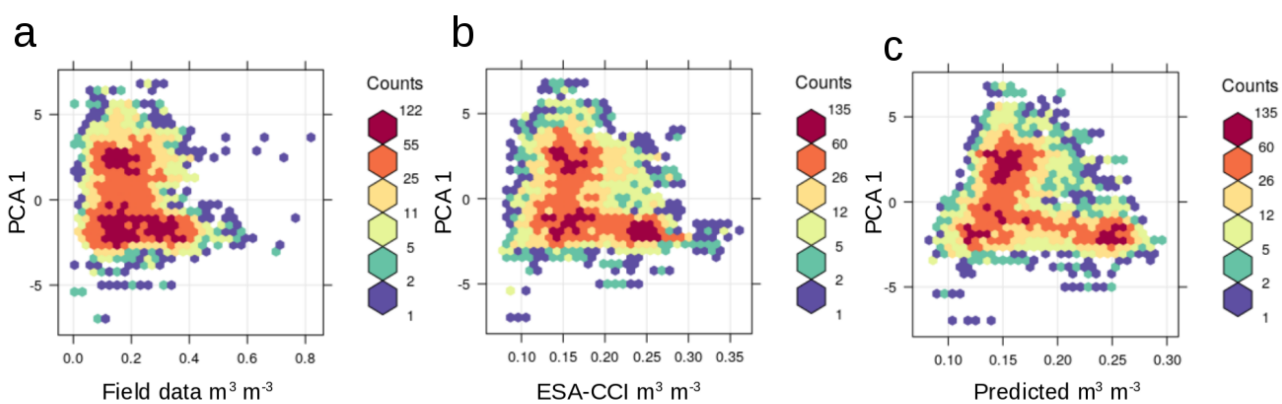

Figure 7 Relationships between the first PC of terrain parameters with soil moisture field data (a), with the ESA-CCI satellite product (b), and with the soil moisture predictions based on terrain parameters (c).

Discussion

Our soil moisture downscaling framework was able to improve the spatial detail of ESA-CCI satellite soil moisture product and its agreement with field soil moisture records from the NASMD. It is well known that topography has a direct influence on the overall water distribution across the landscape [38-39] and in the angle between satellite retrieval and the earth's surface. Thus, we demonstrated how a coarse scale satellite-based soil moisture product (27x27km of spatial resolution), in combination with hydrologically meaningful terrain parameters, can be coupled using machine learning algorithms to generate a fine-gridded and gap-free soil moisture product at the annual scale across CONUS. We found a correlation between field soil moisture estimates and topography that is similar to the correlation between satellite estimates and topography (Figure 7), suggesting that topography can be an effective predictor for direct soil moisture measurements (i.e., from microwave remote sensing). In contrast to previous downscaling efforts using vegetation and climate information [33, 62], we generated 26 annual soil moisture predictions (1991-2016, 1x1 km of spatial resolution) that are independent of ecological data (i.e., vegetation greenness) and climate information, (i.e., precipitation and temperature). This topography-based approach has the advantage that our soil moisture output could be further related to independent datasets of ecological or climate variables [63-64]. Therefore, we provided an alternative (topography-based) approach to predict the satellite soil moisture patterns across finer spatial grids and in areas where no satellite soil moisture is available.

The downscaling process of satellite soil moisture from 27 to 1km grids across CONUS is supported on both internal (Table 1) and independent (Figure 5) validation frameworks to describe modeling performance. Similar results have been found recently for specific study sites [65]. These values showed explained variances >70% and RMSE values considerably below (~0.03 $m^3\ m^{-3}$) the satellite soil moisture mean of 0.22 $m^3\ m^{-3}$, which is suitable for many applications [62], such as the detection of irrigation signals [66]. Our results obtained by the cross-validation strategy and ground validation supports the application of a topography-based model to predict satellite soil moisture estimates (Figure 4).

Our results showed that higher soil moisture values could be found across lower elevations, areas with generally large and gentle slopes mainly across valley bottoms and across catchment areas where water tends to accumulate. This interpretation could explain the short distance in the multivariate analysis of satellite soil moisture estimates to elevation and derived terrain parameters such as the vertical distance (of each pixel) to the nearest channel network, the valley depth index and the topographic wetness index. The multivariate analysis also suggested some degree of statistical redundancy between the topographic prediction factors (Supplementary Information S2) as they were derived from the digital elevation model by the means of geomorphometry [34-39]. For example, we found that the topographic wetness index is highly correlated with the length-slope factor (>0.80%), and this is because they are two secondary parameters that depend on slope [35]. Elevation and slope are respectively required for calculating the valley depth index and the topographic wetness index [36] and these terrain parameters varied closely with soil moisture in the multivariate space (Supplementary Information S2). Thus, understanding the main relationships between topographic prediction factors and soil moisture can be useful for reducing modeling complexity while increasing our capacity to interpret modeling results.

The spatial detail of soil moisture estimates using 1km grids across the continental scale of CONUS is consistent with the variability of soil moisture patterns between the western and eastern United States. While drought scenarios have been recently reported for the western states [67] evidence of precipitation increase has been reported recently in the eastern states [68]. Our soil moisture downscaled estimates (Figure 3) revealed a soil moisture gradient across the Central Plains of CONUS and a clear separation of two major soil moisture data populations (i.e., soil moisture values with a bimodality distribution) from the drier west, to the humid east (Supplementary Figure S5).

The original satellite soil moisture estimates also show this bimodal distribution but with a lesser extent (Supplementary Figure S2). The bimodal distribution of soil moisture could be explained by a negative soil moisture and precipitation feedback in the western CONUS and a positive soil moisture and precipitation feedback in the eastern CONUS [64]. Furthermore, areas with soil moisture bimodality have been recognized across global satellite observations and climate models [69]. We identified areas of low agreement between our soil moisture predictions and field stations (lower r^2 values) across the transitional ecosystems (Figure 4) from drier to humid soil moisture environments (i.e., Central Plains and lower Mississippi basin). It is likely that these transitional areas drive changes in water availability in surface and subsurface hydrological systems [70]. The lower Mississippi basin, specifically the area across the surroundings of the Mississippi delta, is an example of a transitional area experiencing aquifer depletion [71] where both flooding events and droughts tend to occur within shorter distances that are not captured by the original satellite soil moisture information. These are the type areas where we found lower values of agreement (r^2 values) between satellite and ground soil moisture observations. These low correlation values can be also explained by the use of multiple soil moisture networks with different types of sensors and measurement techniques [19]. Also, the imperfections of prediction factors used for soil moisture spatial variability models represent a potential source of uncertainty.

As any downscaling effort dependent on covariates (i.e., terrain parameters), our approach is vulnerable to data quality limitations such as the presence of systematic errors on these covariates. Other errors are derived from input data imperfections and difficulties meeting modeling assumptions. These errors in soil moisture modeling inputs increase the risk of bias and uncertainty propagation to subsequent soil moisture modeling outputs and soil mapping applications [72-74]. For example, elevation data surfaces derived from remote sensing data (such as the global DEM used here) could show artifacts (i.e., false pikes or spurious sinks) due to data saturation or signal noise that can be propagated to final soil moisture predictions [75]. We minimized this issue by using SAGA-GIS [36] as it has adopted methods for preprocessing and perform DEM quality checks [76] before deriving the topographic prediction factors used in this study. Because input covariates could not be fully free of errors, we advocate for reporting information on bias and r^2 values to inform about accuracy (Table 1) as important components for interpreting soil moisture predictions.

Our results suggest that the original coarse scale soil moisture product and the values of soil moisture from the NASMD (Figure 5) are difficult to compare in terms of spatial variability, as is highlighted in previous studies [19]. This is because a satellite soil moisture pixel from the ESA-CCI product provides a value across a larger area (27x27km) than a field measurement at a specific sampling location (defined by geographical coordinates). This scale dependent effect (27x27km vs 1:1 field scale) is reduced (>25%) with soil moisture predictions across finer grids (1km). The downscaled soil moisture maps showed a higher agreement with field soil moisture records from the NASMD (Figure 5), supporting the applicability of this soil moisture product for applications that required higher spatial resolution.

Our soil moisture predictions across 1km grids suggest that topography can be effectively used to improve the spatial detail and accuracy of satellite soil moisture estimates. Several studies have highlighted differences in spatial representativeness between ground-based observations and satellite soil moisture products [73, 77]. Other studies have shown that the spatial representativeness of the ESA-CCI soil moisture compared with field observations is higher from regional-to-continental scales than from ecosystem-to-landscape scales [78-79]. Therefore, large uncertainties of soil moisture spatial patterns (below 1km grids) needs to be resolved for assessing and better understanding the local variability of soil moisture trends. We argue that currently there is an increasing availability of high-quality digital elevation data sources with high levels of spatial resolution (e.g., 1-2 to 30 to 90m grids) across large areas of the world [80-81] that can be used to derive reliable hydrologically meaningful terrain parameters for predicting soil moisture. The relationship of these terrain parameters and field soil moisture (i.e., meteorological stations) is similar to the relationship between terrain parameters and satellite soil moisture gridded estimates (Figure 7).

From a single information source (a remotely sensed DEM), we downscaled satellite records of soil moisture using a framework that theoretically is reproducible across multiple scales. The ultimate goal of reducing the multiple information sources for predicting soil moisture is to reduce the statistical redundancy in further modeling efforts (i.e., land carbon uptake models) and large-scale ecosystem studies (i.e., ecological niche modeling) that combine similar prediction factors for soil moisture (i.e., climate or vegetation indexes). These include models estimating water evapotranspiration trends [82] and process based global carbon models that could also benefit from more accurate and independent soil moisture inputs [74]. To improve the spatial representativeness of satellite soil moisture estimates, the number of studies developing new downscaling approaches based on prediction factors is rapidly expanding [26, 28, 62, 83]. There is a pressing need to solve the current uncertainty of soil moisture estimates to accurately understand how soil moisture is limiting the primary productivity of terrestrial ecosystems [6]. Therefore, our results provide an alternative applicable to continental scales for downscaling satellite soil moisture estimates based on hydrologically meaningful terrain parameters.

The novelty of this approach is that it could be applicable to multiple temporal resolutions (e.g., monthly or daily) as it generates independent models for each period of interest and at multiple spatial scales as the availability of terrain parameters for modeling purposes has increased substantially (i.e., meters) in the last decade. Increasing the temporal resolution of downscaled maps (i.e., from yearly to monthly predictions) is beyond the scope of this study, will increase computational costs, but are theoretically possible following this approach. While monthly or weekly (or even daily soil moisture datasets) are valuable sources for large scale earth system modeling, yearly averages are also valuable for detecting long term trends in the climate-land system. Rather than focusing on temporal variability of soil moisture, our results provide insights for improving the spatial variability and consequently the spatial representation of soil moisture gridded surfaces derived from satellite information.

Conclusion

Recent studies highlight the necessity of detailed soil moisture products to account for soil moisture limitation in terrestrial ecosystems. We developed a geomorphometry-based framework to couple satellite soil moisture records with hydrologically meaningful terrain parameters. We predicted (i.e., downscaled) soil moisture using 1x1km grids across CONUS at a yearly scale from 1991 to 2016. This gap-free soil moisture product improved the spatial detail of the original satellite soil moisture grids and the overall agreement (increased by >20%) of these grids with the NASMD field soil moisture records. Our findings suggest that digital terrain analysis can be applied to elevation data sources to derive hydrologically meaningful terrain parameters and use these parameters predict soil moisture spatial patterns. Our framework is reproducible across the world because it is based on publicly available DEMs, ground and satellite soil moisture data.

Acknowledgments

MG acknowledges a fellowship from CONACyT. RV acknowledges support from the National Science Foundation CIF21 DIBBs (Grant #1724843).

Author Contributions

Conceptualization: Mario Guevara, Rodrigo Vargas

Data curation: Mario Guevara

Formal analysis: Mario Guevara

Writing – original draft: Mario Guevara

Writing – review & editing: Rodrigo Vargas

539 ## References

[1] Greve, P., Gudmundsson, L., Seneviratne, S.I., 2018. Regional scaling of annual mean precipitation and water availability with global temperature change. Earth Syst. Dynam. 9, 227–240. https://doi.org/10.5194/esd-9-227-2018

[2] Seneviratne, S.I., Corti, T., Davin, E.L., Hirschi, M., Jaeger, E.B., Lehner, I., Orlowsky, B., Teuling, A.J., 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

[3] Seneviratne, S.I., Wilhelm, M., Stanelle, T., van den Hurk, B., Hagemann, S., Berg, A., Cheruy, F., Higgins, M.E., Meier, A., Brovkin, V., Claussen, M., Ducharne, A., Dufresne, J.-L., Findell, K.L., Ghattas, J., Lawrence, D.M., Malyshev, S., Rummukainen, M., Smith, B., 2013. Impact of soil moisture-climate feedbacks on CMIP5 projections: First results from the GLACE-CMIP5 experiment: GLACE-CMIP5 EXPERIMENT. Geophysical Research Letters 40, 5212–5217. https://doi.org/10.1002/grl.50956

[4] Western, A.W., Grayson, R.B., Blöschl, G., Willgoose, G.R., McMahon, T.A., 1999. Observed spatial organization of soil moisture and its relation to terrain indices. Water Resour. Res. 35, 797–810. https://doi.org/10.1029/1998WR900065

[5] Dorigo, W., Wagner, W., Albergel, C., Albrecht, F., Balsamo, G., Brocca, L., Chung, D., Ertl, M., Forkel, M., Gruber, A., Haas, E., Hamer, P.D., Hirschi, M., Ikonen, J., de Jeu, R., Kidd, R., Lahoz, W., Liu, Y.Y., Miralles, D., Mistelbauer, T., Nicolai-Shaw, N., Parinussa, R., Pratola, C., Reimer, C., van der Schalie, R., Seneviratne, S.I., Smolander, T., Lecomte, P., 2017. ESA CCI Soil Moisture for improved Earth system understanding: State-of-the art and future directions. Remote Sensing of Environment, Earth Observation of Essential Climate Variables 203, 185–215. https://doi.org/10.1016/j.rse.2017.07.001

[6] Stocker, B. D., Zscheischler, J., Keenan, T. F., Prentice, I. C., Peñuelas, J., & Seneviratne, S. I. 2018. Quantifying soil moisture impacts on light use efficiency across biomes. New Phytol., 218(4), 1430–1449. doi: 10.1111/nph.15123

[7] Brocca, L., Ciabatta, L., Massari, C., Camici, S., & Tarpanelli, A. 2017. Soil Moisture for Hydrological Applications: Open Questions and New Opportunities. Water, 9(2), 140. doi: 10.3390/w9020140

[8] Vargas, R., Sánchez-Cañete, P., Serrano-Ortiz, P., Curiel Yuste, J., Domingo, F., López-Ballesteros, A. and Oyonarte, C., 2018. Hot-moments of soil CO2 efflux in a water-limited grassland. Soil Systems, 2(3), p.47.

[9] Asner, G.P., Alencar, A., 2010. Drought impacts on the Amazon forest: the remote sensing perspective. New phytologist.

[10] Cook, B.D., Corp, L.A., Nelson, R.F., Middleton, E.M., Morton, D.C., McCorkel, J.T., Masek, J.G., Ranson, K.J., Ly, V., Montesano, P.M., 2013. NASA Goddard's LiDAR, Hyperspectral and Thermal (G-LiHT) Airborne Imager. Remote Sensing 5, 4045–4066. https://doi.org/10.3390/rs5084045

[11] Dai, A., 2011. Drought under global warming: a review. Wiley Interdisciplinary Reviews: Climate Change 2, 45–65. https://doi.org/10.1002/wcc.81

[12] Samaniego, L., Thober, S., Kumar, R., Wanders, N., Rakovec, O., Pan, M., Zink, M., Sheffield, J., Wood, E.F., Marx, A., 2018. Anthropogenic warming exacerbates European soil moisture droughts. Nature Climate Change 8, 421–426. https://doi.org/10.1038/s41558-018-0138-5

[13] van der Molen, M.K., Dolman, A.J., Ciais, P., Eglin, T., Gobron, N., Law, B.E., Meir, P., Peters, W., Phillips, O.L., Reichstein, M., Chen, T., Dekker, S.C., Doubková, M., Friedl, M.A.,

Jung, M., van den Hurk, B.J.J.M., de Jeu, R.A.M., Kruijt, B., Ohta, T., Rebel, K.T., Plummer, S., Seneviratne, S.I., Sitch, S., Teuling, A.J., van der Werf, G.R., Wang, G., 2011. Drought and ecosystem carbon cycling. Agricultural and Forest Meteorology 151, 765–773. https://doi.org/10.1016/j.agrformet.2011.01.018

[14] Luo, Y., Ahlström, A., Allison, S.D., Batjes, N.H., Brovkin, V., Carvalhais, N., Chappell, A., Ciais, P., Davidson, E.A., Finzi, A., Georgiou, K., Guenet, B., Hararuk, O., Harden, J.W., He, Y., Hopkins, F., Jiang, L., Koven, C., Jackson, R.B., Jones, C.D., Lara, M.J., Liang, J., McGuire, A.D., Parton, W., Peng, C., Randerson, J.T., Salazar, A., Sierra, C.A., Smith, M.J., Tian, H., Todd-Brown, K.E.O., Torn, M., Groenigen, K.J., Wang, Y.P., West, T.O., Wei, Y., Wieder, W.R., Xia, J., Xu, Xia, Xu, Xiaofeng, Zhou, T., 2016. Toward more realistic projections of soil carbon dynamics by Earth system models. Global Biogeochemical Cycles 30, 40–56. https://doi.org/10.1002/2015GB005239

[15] Walsh, B., Ciais, P., Janssens, I.A., Peñuelas, J., Riahi, K., Rydzak, F., Vuuren, D.P. van, Obersteiner, M., 2017. Pathways for balancing CO2 emissions and sinks. Nature Communications 8, 14856. https://doi.org/10.1038/ncomms14856

[16] Owe, M., Van de Griend, A.A., 1998. Comparison of soil moisture penetration depths for several bare soils at two microwave frequencies and implications for remote sensing. Water Resources Research 34, 2319–2327. https://doi.org/10.1029/98WR01469

[17] Entekhabi, D., Yueh, S., et al., 2014. SMAP handbook—Soil Moisture Active Passive: Mapping Soil Moisture and Freeze/Thaw From Space, Jet Propulsion Lab., California Inst. Technol., Pasadena, Calif.

[18] Singh, R.S., Reager, J.T., Miller, N.L., Famiglietti, J.S., 2015. Toward hyper-resolution land-surface modeling: The effects of fine-scale topography and soil texture on CLM4.0 simulations

over the Southwestern U.S.: Effects of fine-scale resolution on CLM4.0 in Southwest US. Water Resources Research 51, 2648–2667. https://doi.org/10.1002/2014WR015686

[19] Dirmeyer, Paul A., Jiexia Wu, Holly E. Norton, Wouter A. Dorigo, Steven M. Quiring, Trenton W. Ford, Joseph A. Santanello, et al. 2016. "Confronting Weather and Climate Models with Observational Data from Soil Moisture Networks over the United States." Journal of Hydrometeorology 17 (4): 1049–67. https://doi.org/10.1175/jhm-d-15-0196.1.

[20] 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., 2012. Trend-preserving blending of passive and active microwave soil moisture retrievals. Remote Sensing of Environment 123, 280–297. https://doi.org/10.1016/j.rse.2012.03.014

[21] Liu, Y.Y., Parinussa, R.M., Dorigo, W.A., De Jeu, R.A.M., Wagner, W., van Dijk, A.I.J.M., McCabe, M.F., Evans, J.P., 2011. Developing an improved soil moisture dataset by blending passive and active microwave satellite-based retrievals. Hydrology and Earth System Sciences 15, 425–436. https://doi.org/10.5194/hess-15-425-2011

[22] McColl, K.A., Alemohammad, S.H., Akbar, R., Konings, A.G., Yueh, S., Entekhabi, D., 2017. The global distribution and dynamics of surface soil moisture. Nature Geoscience 10, 100–104. https://doi.org/10.1038/ngeo2868

[23] Montzka, C., K. Rötzer, H.R. Bogena, and H. Vereecken. 2018. A new soil moisture downscaling approach for SMAP, SMOS and ASCAT by predicting sub-grid variability. Remote Sens. 10(3):427. doi:10.3390/rs10030427

[24] Afshar, M.H., Yilmaz, M.T., 2017. The added utility of nonlinear methods compared to linear methods in rescaling soil moisture products. Remote Sensing of Environment 196, 224–237. https://doi.org/10.1016/j.rse.2017.05.017

[25] Jin, Y., Ge, Y., Wang, J., Heuvelink, G.B.M., Wang, L., 2018. Geographically Weighted Area-to-Point Regression Kriging for Spatial Downscaling in Remote Sensing. Remote Sensing 10, 579. https://doi.org/10.3390/rs10040579

[26] Kearney, M.R., Maino, J.L., 2018. Can next-generation soil data products improve soil moisture modelling at the continental scale? An assessment using a new microclimate package for the R programming environment. Journal of Hydrology 561, 662–673. https://doi.org/10.1016/j.jhydrol.2018.04.040

[27] Piles, M., Camps, A., Vall-llossera, M., Corbella, I., Panciera, R., Rudiger, C., Kerr, Y.H., Walker, J., 2011. Downscaling SMOS-Derived Soil Moisture Using MODIS Visible/Infrared Data. IEEE Transactions on Geoscience and Remote Sensing 49, 3156–3166. https://doi.org/10.1109/TGRS.2011.2120615

[28] Ranney, K.J., Niemann, J.D., Lehman, B.M., Green, T.R., Jones, A.S., 2015. A method to downscale soil moisture to fine resolutions using topographic, vegetation, and soil data. Advances in Water Resources 76, 81–96. https://doi.org/10.1016/j.advwatres.2014.12.003

[29] Wang, A., Zhang, M., Shi, J., Mu, T., Gong, H., Xie, C., 2012. Space-time analysis on downscaled soil moisture data and parameters of plant growth. Transactions of the Chinese Society of Agricultural Engineering 28, 164–169.

[30] Yu, G., Di, L., Yang, W., 2008. Downscaling of Global Soil Moisture using Auxiliary Data. IEEE, pp. III-230-III–233. https://doi.org/10.1109/IGARSS.2008.4779325

[31] McBratney, A.., Mendonça Santos, M.., Minasny, B., 2003. On digital soil mapping. Geoderma 117, 3–52. https://doi.org/10.1016/S0016-7061(03)00223-4

[32] Bauer-Marschallinger, B., Freeman, V., Cao, S., Paulik, C., Schaufler, S., Stachl, T., ...Wagner, W. (2019). Toward Global Soil Moisture Monitoring With Sentinel-1: Harnessing Assets and

Overcoming Obstacles. IEEE Trans. Geosci. Remote Sens., 57(1), 520–539. doi: 10.1109/TGRS.2018.2858004

[33] Morelo B., Merlin O., Malbeteau Y., Al Bitar A., Cabot F., Stefan V., Kerr Y., Bacon S., Cosh., M., Bindlish R., Jackson T.J. 2016. SMOS disaggregated soil moisture product at 1km resolution: Processor overview and first validation results. Remote Sensing of Environment 180, 361-376 https://doi.org/10.1016/j.rse.2016.02.045

[34] Pike, R.J., Evans, I.S., , T., 2009. Chapter 1 Geomorphometry: A Brief Guide, in: Developments in Soil Science. Elsevier, pp. 3–30.

[35] Wilson, J. P., & Gallant, J. C. (2000). Digital terrain analysis. Terrain analysis: Principles and applications, 6(12), 1–27.

[36] Conrad, O., Bechtel, B., Bock, M., Dietrich, H., Fischer, E., Gerlitz, L., Wehberg, J., Wichmann, V., Böhner, J., 2015. System for Automated Geoscientific Analyses (SAGA) v. 2.1.4. Geoscientific Model Development 8, 1991–2007. https://doi.org/10.5194/gmd-8-1991-2015

[37] Wilson, J.P., 2012. Digital terrain modeling. Geomorphology, Geospatial Technologies and Geomorphological Mapping Proceedings of the 41st Annual Binghamton Geomorphology Symposium 137, 107–121. https://doi.org/10.1016/j.geomorph.2011.03.012

[38] Florinsky, I.V., 2016. Chapter 9 - Influence of Topography on Soil Properties, in: Florinsky, I.V. (Ed.), Digital Terrain Analysis in Soil Science and Geology (Second Edition). Academic Press, pp. 265–270. https://doi.org/10.1016/B978-0-12-804632-6.00009-2

[39] Florinsky, I.V., 2012. The Dokuchaev hypothesis as a basis for predictive digital soil mapping (on the 125th anniversary of its publication). Eurasian Soil Science 45, 445–451. https://doi.org/10.1134/S1064229312040047

[40] Hengl, T., MacMillan, R.A., 2019. Predictive Soil Mapping with R. OpenGeoHub foundation, Wageningen, the Netherlands, 370 pages, www.soilmapper.org, ISBN: 978-0-359-30635-0.

[41] Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., & Prabhat. 2019. Deep learning and process understanding for data-driven Earth system science. Nature, 566(7743), 195. doi: 10.1038/s41586-019-0912-1

[42] Warner, D.L., Guevara, M., Inamdar, S. and Vargas, R., 2019. Upscaling soil-atmosphere CO2 and CH4 fluxes across a topographically complex forested landscape. Agricultural and forest meteorology, 264, pp.80-91.

[43] Coopersmith, E. J., Cosh, M. H., Bell, J. E., & Boyles, R. 2016. Using machine learning to produce near surface soil moisture estimates from deeper in situ records at U.S. Climate Reference Network (USCRN) locations: Analysis and applications to AMSR-E satellite validation. Adv. Water Resour., 98, 122–131. doi: 10.1016/j.advwatres.2016.10.007

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

[45] Dorigo, W., Wagner, W., Albergel, C., Albrecht, F., Balsamo, G., Brocca, L., Chung, D., Ertl, M., Forkel, M., Gruber, A., Haas, E., Hamer, P.D., Hirschi, M., Ikonen, J., de Jeu, R., Kidd, R., Lahoz, W., Liu, Y.Y., Miralles, D., Mistelbauer, T., Nicolai-Shaw, N., Parinussa, R., Pratola, C., Reimer, C., van der Schalie, R., Seneviratne, S.I., Smolander, T., Lecomte, P., 2017. ESA CCI Soil Moisture for improved Earth system understanding: State-of-the art and future directions. Remote Sensing of Environment, Earth Observation of Essential Climate Variables 203, 185–215. https://doi.org/10.1016/j.rse.2017.07.001

[46] Bindlish, R., Jackson, T., Cosh, M., Tianjie Zhao, O'Neill, P., 2015. Global Soil Moisture From the Aquarius/SAC-D Satellite: Description and Initial Assessment. IEEE Geoscience and Remote Sensing Letters 12, 923–927. https://doi.org/10.1109/LGRS.2014.2364151

[47] 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., Kimball, J., Piepmeier, J.R., Koster, R.D., Martin, N., McDonald, K.C., Moghaddam, M., Moran, S., Reichle, R., Shi, J.C., Spencer, M.W., Thurman, S.W., Tsang, L., Van Zyl, J., 2010. The Soil Moisture Active Passive (SMAP) Mission. Proceedings of the IEEE 98, 704–716. https://doi.org/10.1109/JPROC.2010.2043918

[48] Naeimi, V., Paulik, C., Bartsch, A., Wagner, W., Kidd, R., Park, S.-E., Elger, K., Boike, J., 2012. ASCAT Surface State Flag (SSF): Extracting Information on Surface Freeze/Thaw Conditions From Backscatter Data Using an Empirical Threshold-Analysis Algorithm. IEEE Transactions on Geoscience and Remote Sensing 50, 2566–2582. https://doi.org/10.1109/TGRS.2011.2177667

[49] Naeimi, V., Scipal, K., Bartalis, Z., Hasenauer, S., Wagner, W., 2009. An Improved Soil Moisture Retrieval Algorithm for ERS and METOP Scatterometer Observations. IEEE Transactions on Geoscience and Remote Sensing 47, 1999–2013. https://doi.org/10.1109/TGRS.2008.2011617

[50] Wagner, W., Lemoine, G., Rott, H., 1999. A Method for Estimating Soil Moisture from ERS Scatterometer and Soil Data. Remote Sensing of Environment 70, 191–207. https://doi.org/10.1016/S0034-4257(99)00036-X

[51] Dorigo, W.A., Wagner, W., Hohensinn, R., Hahn, S., Paulik, C., Xaver, A., Gruber, A., Drusch, M., Mecklenburg, S., van Oevelen, P., Robock, A., Jackson, T., 2011. The International Soil Moisture Network: a data hosting facility for global in situ soil moisture measurements.

Hydrology and Earth System Sciences 15, 1675–1698. https://doi.org/10.5194/hess-15-1675-2011

[52] Becker, J.J., Sandwell, D.T., Smith, W.H.F., Braud, J., Binder, B., Depner, J., Fabre, D., Factor, J., Ingalls, S., Kim, S.-H., Ladner, R., Marks, K., Nelson, S., Pharaoh, A., Trimmer, R., Von Rosenberg, J., Wallace, G., Weatherall, P., 2009. Global Bathymetry and Elevation Data at 30 Arc Seconds Resolution: SRTM30_PLUS. Marine Geodesy 32, 355–371. https://doi.org/10.1080/01490410903297766

[53] Hengl T, de Jesus JM, MacMillan RA, Batjes NH, Heuvelink GBM, Ribeiro E, et al. SoilGrids1km — Global Soil Information Based on Automated Mapping. Bond-Lamberty B, editor. PLoS ONE. 2014;9: e105992. doi:10.1371/journal.pone.0105992

[54] Tuanmu, M.-N., & Jetz, W. A global 1-km consensus land-cover product for biodiversity and ecosystem modelling. Global Ecol. Biogeogr., 23(9), 1031–1045. 2014. doi: 10.1111/geb.12182

[55] Thuleau S, and Husson F. 2018. FactoInvestigate: Automatic Description of Factorial Analysis. R package version 1.3. https://CRAN.R-project.org/package=FactoInvestigate

[56] Hechenbichler, K., Schliep, K., 2006. Weighted k-nearest-neighbor techniques and ordinal classification, in: Discussion Paper 399, SFB 386.

[57] Hechenbichler, K., Schliep, K., 2004. Weighted k-Nearest-Neighbor Techniques and Ordinal Classification [WWW Document]. URL https://epub.ub.uni-muenchen.de/1769/ (accessed 12.24.16).

[58] Borra, S., Di Ciaccio, A., 2010. Measuring the prediction error. A comparison of cross-validation, bootstrap and covariance penalty methods. Computational Statistics & Data Analysis 54, 2976–2989. https://doi.org/10.1016/j.csda.2010.03.004

[59] Oliver, M. A., & Webster, R. 2014. A tutorial guide to geostatistics: Computing and modelling variograms and kriging. CATENA, 113, 56–69. doi: 10.1016/j.catena.2013.09.006

[60] Hiemstra, P. H., Pebesma, E. J., Twenhöfel, C. J. W., & Heuvelink, G. B. M. 2009. Real-time automatic interpolation of ambient gamma dose rates from the Dutch radioactivity monitoring network. Comput. Geosci., 35(8), 1711–1721. doi: 10.1016/j.cageo.2008.10.011

[61] R Core Team 2018. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

[62] Colliander, A., Fisher, J. B., Halverson, G., Merlin, O., Misra, S., Bindlish, R., ...Yueh, S. 2017. Spatial Downscaling of SMAP Soil Moisture Using MODIS Land Surface Temperature and NDVI During SMAPVEX15. IEEE Geosci. Remote Sens. Lett., 14(11), 2107–2111. doi: 10.1109/LGRS.2017.2753203

[63] He, L., Chen, J. M., Liu, J., Bélair, S., & Luo, X. 2017. Assessment of SMAP soil moisture for global simulation of gross primary production. J. Geophys. Res. Biogeosci., 122(7), 1549–1563. doi: 10.1002/2016JG003603

[64] Tuttle, S. & Salvucci, G. 2016. Empirical evidence of contrasting soil moisture-precipitation feedbacks across the United States. Science **352**, 825–828.

[65] Colliander, A., Jackson, T. J., Bindlish, R., Chan, S., Das, N., Kim, S. B., ...Yueh, S. 2017. Validation of SMAP surface soil moisture products with core validation sites. Remote Sens. Environ., 191, 215–231. doi: 10.1016/j.rse.2017.01.021

[66] Lawston, P. M., Santanello, J. A., & Kumar, S. V. 2017. Irrigation Signals Detected From SMAP Soil Moisture Retrievals. Geophys. Res. Lett., 44(23), 11,860–11,867. doi: 10.1002/2017GL075733

[67] Diffenbaugh, N. S., Swain, D. L., & Touma, D. 2015. Anthropogenic warming has increased drought risk in California. Proc. Natl. Acad. Sci. U.S.A., 112(13), 3931–3936. doi: 10.1073/pnas.1422385112

[68] Easterling, D. R., Kunkel, K. E., Arnold, J. R., Knutson, T., LeGrande, A. N., Leung, L. R., Wehner, M. F. 2017. Precipitation change in the United States. In D. J. Wuebbles, D. W. Fahey, K. A. Hibbard, D. J. Dokken, B. C. Stewart, & T. K. Maycock (Eds.), Climate science special report: Fourth national climate assessment (Vol. I, pp. 207– 230). Washington, DC: U.S. Global Change Research Program

[69] Vilasa L, Miralles DG, de Jeu RAM, Dolman AJ. Global soil moisture bimodality in satellite observations and climate models. Journal of Geophysical Research: Atmospheres. 2017;122: 4299–4311. doi:10.1002/2016jd026099

[70] Dyer, J., & Mercer, A. 2013. Assessment of Spatial Rainfall Variability over the Lower Mississippi River Alluvial Valley. J. Hydrometeorol., 14(6), 1826–1843. doi: 10.2307/24914344

[71] Reba, M. L., Massey, J. H., Adviento-Borbe, M. A., Leslie, D., Yaeger, M. A., Anders, M., & Farris, J. 2017. Aquifer Depletion in the Lower Mississippi River Basin: Challenges and Solutions. Journal of Contemporary Water Research & Education, 162(1), 128–139. doi: 10.1111/j.1936-704X.2017.03264.x

[72] Heuvelink, G.B. M., Millward, A.A., 1999. Error propagation in environmental modelling with GIS. Cartographica 36, 69.

[73] Miralles, D.G., Crow, W.T., Cosh, M.H., 2010. Estimating Spatial Sampling Errors in Coarse-Scale Soil Moisture Estimates Derived from Point-Scale Observations. Journal of Hydrometeorology 11, 1423–1429. https://doi.org/10.1175/2010JHM1285.1

[74] Munguia-Flores, F., Arndt, S., Ganesan, A. L., Murray-Tortarolo, G., & Hornibrook, E. R. C. 2018. Soil Methanotrophy Model (MeMo v1.0): a process-based model to quantify global uptake of atmospheric methane by soil. Geosci. Model Dev., 11(6), 2009–2032. doi: 10.5194/gmd-11-2009-2018

[75] Lindsay J.B, Creed I.F. Removal of artifact depressions from digital elevation models: towards a minimum impact approach. Hydrological Processes. 2005;19: 3113–3126. doi:10.1002/hyp.5835

[76] Planchon, O. & F. Darboux (2001): A fast, simple and versatile algorithm to fill the depressions of digital elevation models. Catena 46: 159-176

[77] Gruber, A., Dorigo, W.A., Zwieback, S., Xaver, A., Wagner, W., 2013. Characterizing Coarse-Scale Representativeness of in situ Soil Moisture Measurements from the International Soil Moisture Network. Vadose Zone Journal 12, 0. https://doi.org/10.2136/vzj2012.0170

[78] Nicolai-Shaw, N., Hirschi, M., Mittelbach, H., Seneviratne, S.I., 2015. Spatial representativeness of soil moisture using in situ, remote sensing, and land reanalysis data: SPATIAL REPRESENTATIVENESS OF SOIL MOISTURE. Journal of Geophysical Research: Atmospheres 120, 9955–9964. https://doi.org/10.1002/2015JD023305

[79] Vargas, R., Sonnentag, O., Abramowitz, G., Carrara, A., Chen, J.M., Ciais, P., Correia, A., Keenan, T.F., Kobayashi, H., Ourcival, J.M. and Papale, D., 2013. Drought influences the accuracy of simulated ecosystem fluxes: a model-data meta-analysis for Mediterranean oak woodlands. Ecosystems, 16(5), pp.749-764.

[80] Nelson, A., Reuter, H.I., Gessler, P., 2009. Chapter 3 DEM Production Methods and Sources, in: Developments in Soil Science. Elsevier, pp. 65–85.

[81] Tadono, T., Ishida, H., Oda, F., Naito, S., Minakawa, K., Iwamoto, H., 2014. Precise Global DEM Generation by ALOS PRISM. ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences II–4, 71–76. https://doi.org/10.5194/isprsannals-II-4-71-2014

[82] Schwingshackl, C., Hirschi, M., Seneviratne, S. I., Schwingshackl, C., Hirschi, M., & Seneviratne, S. I. 2017. Quantifying Spatiotemporal Variations of Soil Moisture Control on Surface Energy Balance and Near-Surface Air Temperature. J. Clim. Retrieved from https://journals.ametsoc.org/doi/full/10.1175/JCLI-D-16-0727.1

[83] Jin, Y., Ge, Y., Wang, J., Heuvelink, G.B.M., Wang, L., 2018. Geographically Weighted Area-to-Point Regression Kriging for Spatial Downscaling in Remote Sensing. Remote Sensing 10, 579. https://doi.org/10.3390/rs10040579