AGU ADVANCING EARTH AND SPACE SCIENCE

Water Resources Research

RESEARCH ARTICLE
10.1029/2020WR029249

Regionalizing Root-Zone Soil Moisture Estimates From ESA CCI Soil Water Index Using Machine Learning and Information on Soil, Vegetation, and Climate

Manolis G. Grillakis[1], Aristeidis G. Koutroulis[2], Dimitrios D. Alexakis[1], Christos Polykretis[1], and Ioannis N. Daliakopoulos[3]

[1]Lab of Geophysical-Remote Sensing & Archaeo-Environment, Institute for Mediterranean Studies, Foundation for Research and Technology Hellas, Rethymno, Crete, Greece, [2]School of Environmental Engineering, Technical University of Crete, Chania, Crete, Greece, [3]Department of Agriculture, Hellenic Mediterranean University, Heraklion, Crete, Greece

Key Points:

- We use machine learning to regionalize the calibration of the Soil Water Index based on soil, climate, and vegetation descriptors
- The results are compared to reanalysis data sets, indicating added value to the results of the machine learning calibration

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

M. G. Grillakis,
grillakis@hydrogaia.gr

Citation:

Grillakis, M. G., Koutroulis, A. G., Alexakis, D. D., Polykretis, C., & Daliakopoulos, I. N. (2021). Regionalizing root-zone soil moisture estimates from ESA CCI Soil Water Index using machine learning and information on soil, vegetation, and climate. *Water Resources Research*, *57*, e2020WR029249. https://doi.org/10.1029/2020WR029249

Received 16 NOV 2020
Accepted 31 MAR 2021

Abstract The European Space Agency (ESA), through the Climate Change Initiative (CCI), is currently providing nearly 4 decades of global satellite-observed, fully homogenized soil moisture data for the uppermost 2–5 cm of the soil layer. These data are valuable as they comprise one of the most complete remotely sensed soil moisture data sets available in time and space. One main limitation of the ESA CCI soil moisture data set is the limited soil depth at which the moisture content is represented. In order to address this critical gap, we (a) estimate and calibrate the Soil Water Index using ESA CCI soil moisture against in situ observations from the International Soil Moisture Network and then (b) leverage machine learning techniques and physical soil, climate, and vegetation descriptors at a global scale to regionalize the calibration. We use this calibration to assess the root-zone soil moisture for the period 2001–2018. The results are compared against the European Centre for Medium-Range Weather Forecasts, ERA5 Land, and the Famine Early Warning Systems Network Land Data Assimilation System reanalyses soil moisture data sets, showing a good agreement, mainly over mid latitudes. This work contributes to the exploitation of ESA CCI soil moisture data, while the produced data can support large-scale soil moisture-related studies.

1. Introduction

The water stored in the root-zone layer is one of the most important water resources, as it supplies natural vegetation and crops with the vital water needed. Today, 75% of the total crop area harvested at global scale consists of nonirrigated crops (Portmann et al., 2010), which emphasizes the critical role of soil moisture in global food production and food security. From a natural system perspective, soil moisture is one of the most important components of the hydrological cycle, regulating energy flows and biogeochemical cycles, and hence plays a key role on plant growth (Seneviratne et al., 2010). Lacking systematic in situ measurements of soil moisture, research that focuses on agricultural drought assessment (Bussotti et al., 2014; Grillakis, 2019; Pokhrel et al., 2021; Tramblay et al., 2020) and runoff simulation systems (Bai et al., 2016; Bischiniotis et al., 2019) has been mainly based on model simulations of soil moisture.

Soil moisture varies substantially in space and time rendering its measurement difficult, therefore, satellite remote-sensed soil moisture can be considered as a major global alternative to the limited in situ observations (Lettenmaier et al., 2015). Over the years, an increasing number of synergistic products based on a range of direct and indirect measurements of soil moisture have been developed (W. Dorigo et al., 2017). The European Space Agency (ESA), through the Climate Change Initiative (CCI), is currently providing nearly 4 decades of global satellite-observed, fully homogenized soil moisture (SM) data for the uppermost soil layer, that can be used to analyze the soil moisture climatology at global scale (W. Dorigo et al., 2017; Gruber et al., 2017, 2019). This data set is valuable as it is one of the most complete soil moisture data sets in time and space, currently available. The ESA CCI SM is a homogenized combination of various single-sensor active and passive microwave soil moisture products (W. Dorigo et al., 2017). A major limitation that ESA CCI SM exhibits in comparison to the reanalysis products is the shallow depth for which the soil moisture is assessed (Brocca et al., 2017) (2–5 cm for the combined passive and active CCI SM product), that limits the usability of the data. A second limitation of the ESA CCI SM that is not considered in this study is the uneven fractional coverage of the data both in space and time due to the different satellite products

that are combined in each period. However, the introduction of the ERS 1/2, AMSR-E, ASCAT, and SMAP satellite products in 1990, 2002, 2007, and 2015 respectively, alleviated this limitation, by making significant difference in the fractional coverage of the ESA CCI SM (W. Dorigo et al., 2017). Further, it has been found that the blending/merging algorithms may tamper the long-term trends of the data (Gruber et al., 2019; Liu et al., 2011).

To enable the assessment of the deeper soil layer's moisture, Wagner et al. (1999) established a convolutive transformation of the surface soil moisture estimation, under the assumption of uniform hydraulic conductivity and the absence of evapotranspiration. The established transformation, the Soil Water Index (SWI), utilizes a single parameter that is described as the characteristic time length parameter, T, which parametrizes the hydraulic and evapotranspirative processes that govern the soil moisture evolution. Shallower depths and more permeable soils are being generally associated with smaller T values due to the faster infiltration. In Bauer-Marschallinger et al. (2018), a T value of 1 is correlated to the surface soil moisture, while T value of 5 represented the soil moisture in the layer just below the soil surface. In Grillakis et al. (2016), a T value of 20 was used to reflect root-zone soil moisture. The SWI is usually estimated on normalized satellite soil moisture, hence its outcome is also normalized between 0 (dry) and 1 (wet). In contrast, soil moisture is usually preferred to be expressed in physical units, for example, water volume per soil volume, or kg m^{-2}, however soil moisture anomalies may be more useful than absolute values for specific tasks (Bauer-Marschallinger et al., 2018; Nicolai-Shaw et al., 2015). Furthermore, SWI can also be translated into physical units with the aid of appropriate metadata, that is, scaling the index between the wilting point and the field capacity of the soil.

Terrestrial soil moisture measurements comprise a valuable piece of data that has aided the better understanding of the soil hydrological processes. Soil moisture measurements have been used for the calibration and validation of soil moisture accounting models (Albergel et al., 2008; Grillakis et al., 2016) or to validate satellite-based moisture products (Ceballos et al., 2005; Gruhier et al., 2010; Parrens et al., 2012; Paulik et al., 2014). The International Soil Moisture Network (ISMN) (W. A. Dorigo et al., 2011, 2013) gathers and harmonizes in situ soil moisture measurements from operational networks and validation campaigns worldwide, making them available to the scientific community in an unprecedented database of ground soil moisture observations.

Using point measured information for validating remotely sensed soil moisture product has been subject of extensive research (Albergel et al., 2012; Brocca et al., 2011). Further, calibration procedures have been used to bring remotely sensed soil moisture close to the ground measurements. In their work, Albergel et al. (2008) and Parajka et al. (2009) calibrated the T parameter on in situ soil moisture measurements, showing that a calibrated remote-sensed SWI can simulate soil moisture. Paulik et al. (2014) validated SWI against a large number of ISMN observed time series. Their approach was to use various fixed values of T between 1 and 100 (1, 5, 10, 15, 20, 40, 60, and 100) and then assess the T value that provided the most similar SWI to the ground moisture measurements. Nevertheless, in such comparison approaches, the spatial scale gap between the gridded soil moisture product and the point measurements has been found to be a significant source of errors (Nicolai-Shaw et al., 2015). Crow et al. (2012) mention that even the ground measuring networks with the largest spatial coverage lack the necessary sampling density to describe the horizontal variability in soil moisture fields, which stems from the interactions between soil, vegetation, topography, and climate. From these factors, topography importance decreases while vegetation importance increases with soil depth (Guo et al., 2020). Further studies have provided evidence that soil moisture variability is greatly affected not only by the spatial scale but also by the mean soil moisture state, making the upscaling even more difficult (Famiglietti et al., 2008; Korres et al., 2015). A simple but robust way to tackle with the scale mismatch in the soil moisture measurements is to choose representative location(s) that provide a good agreement to the remote-sensed data (Cosh et al., 2006; Vinnikov et al., 1999).

Root zone is defined as the depth of soil at which the plant roots exploit water, with different types of vegetation exhibiting different rooting depth. It is difficult to define the rooting depth even for a single type of vegetation due to the uneven vertical root distribution. As an indicative example, Raza et al. (2013) analyzed the root biomass distribution of a 11 temperate crops, showing that while half of the roots are located at the uppermost ~15 cm of the soil, the maximum rooting depth exceeds the 100 cm. Further, they indicate that the 61%–78% of the rooting biomass of the various crops were found at the top 30 cm of the soil. Skillful

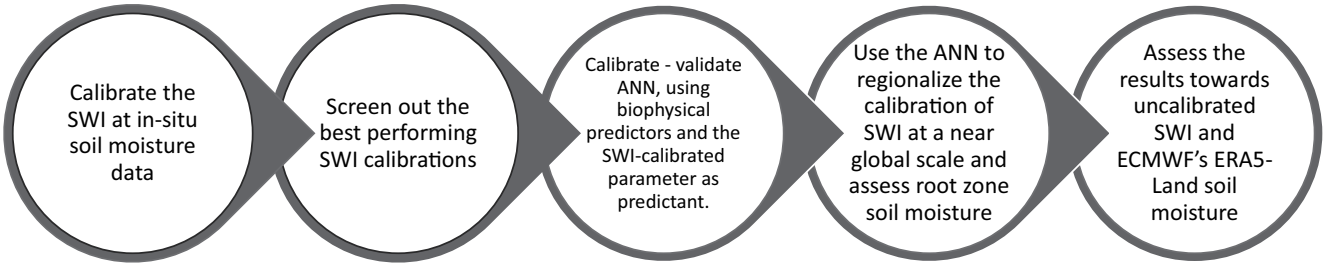

Figure 1. Methodology followed in the present study.

estimates of root-zone moisture can aid to better weather forecasting, flood and drought predictions as well as climate projections (Chen et al., 2014; Dirmeyer et al., 2006).

Machine learning techniques, such as artificial neural networks (ANNs) and support vector machines, have proven useful in simulating the soil water as they can handle noisy data, as well simulate complex nonlinear processes with the aid of the appropriate predictors. Achieng (2019) used machine learning techniques to simulate soil water retention curve of loamy sandy soils. Kumar et al. (2019) used support vector machine-based regression techniques with near-real-time remote-sensed soil moisture to enhance near-real-time satellite-based rainfall estimates for the Ashti watershed in India. Mao et al. (2019) used machine learning to fill-in temporal gaps in high-resolution soil moisture data of SMAP/Sentinel-1. Alexakis et al. (2017) elaborated Sentinel-1, Landsat 8, and machine learning techniques to estimate topsoil soil moisture content. Furthermore, machine learning has been used in the literature to downscale remotely sensed soil moisture products (Srivastava et al., 2013).

Given the technological limitations in the soil moisture remote sensing in deeper rather than surface soil levels, this work aims to explore the machine learning capabilities in combination with ground measurements and biophysical parameters to fill the gap. In situ measurements of ISMN are used to calibrate the SWI index and then use ANNs to regionalize the calibration at a global scale. The calibration is then used to estimate SWI at a root-zone meaningful depth. The result is a global, monthly scale, 0.25° soil moisture data set at a potentially global scale under the restrictions of the ESA CCI SM coverage and the calibration limitations. The result is compared against the ERA5 Land and Famine Early Warning Systems Network (FEWS NET) Land Data Assimilation System (FLDAS) reanalyses soil moisture data.

2. Materials and Methods

The methodology followed can be split into five successive stages. In the first stage, ground, point scale measurements of soil moisture for different regions and depths from ISMN were used to calibrate the value of the T parameter of SWI estimation against ESA CCI SM data. ESA CCI SM data were obtained for each location where ISMN data were available, through nearest neighbor interpolation. Second, the best correlating locations are selected for further processing. From them, the calibrated T values were used as a predictant to train and test an ANN that associates different soil, climate, and land cover characteristics, to obtain an algorithm capable of T value regionalization. Fourth, the ANN was used to regionalize the T value, and then to use it to assess root-zone relative soil moisture based on ESA CCI SM data (rSWI), at a near global scale. Finally, the results of the rSWI were compared against the uncalibrated SWI assessments and the ERA5 Land and FLDAS data. A graphical illustration of the methodology is shown in Figure 1.

2.1. Available Data

2.1.1. Remotely Sensed Soil Moisture

The ESA CCI data (ESA CCI-SOILMOISTURE-L3S-SSMV-COMBINED, version 4.5) were used as the remotely sensed surface soil moisture. This product provides global soil moisture data from about 1979 to 2018 and combines the active scatterometer data and the passive radiometer soil moisture products derived from the AMI-WS, ASCAT, SMMR, SSM/I, TMI, AMSR-E, WindSat, AMSR2, and SMOS satellite instruments.

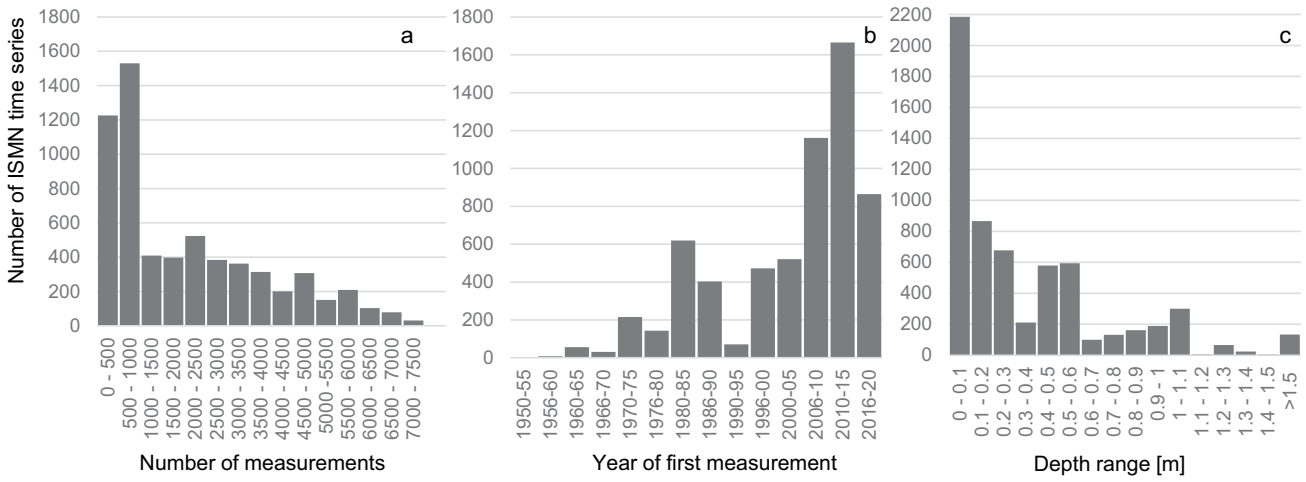

Figure 2. Characteristics of the ISMN data. (a) The distribution of the ISMN time series length in daily measurements, (b) year of the first measurement in each ISMN time series, and (c) the distribution of the measuring depths as defined by the average between the upper and lower measuring depth (intervals on the figure are reported as Lower bound < Depth range ≤ Upper bound). ISMN, International Soil Moisture Network.

Although soil moisture monitoring could be potentially performed by a single-sensor product, individual satellite missions cannot provide continuous monitoring for long (several decades) periods. Hence, the ESA CCI Soil Moisture project was established to aid soil moisture monitoring needs in support of climate research (W. Dorigo et al., 2017). The merging of ESA CCI SM multisensor data is performed on already produced soil moisture products, following a scaling and merging procedure separately for the active and passive products. Finally, the passive and active products are blended to provide the ESA CCI SM. This blending is performed by a location specific weighting of passive and active subproducts, such that the random errors are minimized. The ESA CCI SM data are provided at 0.25° and daily time step, derived by satellite microwave observations between 25 and 50 km^2 and daily or subdaily time steps. The data set is representative for a 2–5-cm depth of the soil (Brocca et al., 2017). Detailed description of the data can be found in W. Dorigo et al. (2017), while the theoretical base of the product can be found in Wagner et al. (2012). Beyond the volumetric soil moisture provided in the combined data set, also ancillary data are provided, such estimates of the volumetric soil moisture error, as well as a quality flags to indicate spurious retrievals.

2.1.2. In Situ Soil Moisture Measurements

Ground measurements from the ISMN (W. Dorigo et al., 2021; W. A. Dorigo et al., 2011, 2013) data were considered for the calibration of T value. The ISMN data set comprises the most comprehensive collection of soil moisture measurements globally. At the time ISMN data were obtained, it included over 6,200 soil moisture data sets, from over 1,700 stations from 49 soil moisture networks. The main drawback of the data offered by ISMN, while harmonized in terms of units, time step, and metadata, is not homogenized over measuring depth and between different sites (W. A. Dorigo et al., 2013). Furthermore, there is a strong coverage bias toward the northern hemisphere and especially Northern America and Eurasia regions (shown later—Figure 4).

Within ISMN data, the gauging duration varies significantly among different stations. The majority of the gauging locations provide less than 1,000 daily measurements but, in some cases, the daily data reach 20 years (Figure 2a). The availability of ISMN data has significantly increased in the recent 2 decades (Figure 2b) since the available technology for soil moisture measurement was improved and turned to be more affordable while also the scientific community better comprehend the importance of soil moisture as a resource. Regarding the available measuring depths (Figure 2c), a large portion of the available data come from soil depths less than (or equal to) 0.1 m. Nevertheless, time series at 0.5 and 1 m depths also constitute a significant fraction of the measurements (Figure 2c).

Table 1
Soil, Vegetation, and Climatic Parameters Used for the Regionalization of T Parameter

	Parameter	Affecting process
1	Available soil water potential in kPa −10 (pF 2.0)	Increased moisture potential at pF 2.0, 2.3, and 2.5 would affect the available (between field capacity and wilting point) water for evapotranspiration
2	Available soil water potential in kPa −20 (pF 2.3)	
3	Available soil water potential in kPa −31.6 (pF 2.5)	
4	Saturated water content (% vol)	Affects the total water capacity of soil
5	Bulk density of fine earth (kg m^{-3})	Higher bulk density negatively affects the soil porosity
6	Clay content (0–2 μm) (mass %)	High clay content links to slower soil water drainage
7	Coarse fragments (% vol)	Coarse fragments are mostly soil moisture inactive. Higher fraction reduces the presence of other soil ingredients
8	Soil organic carbon content in fine earth fraction (g kg^{-1})	Soil organic carbon affects hydraulic conductivity
9	Silt content (2–50 μm) (mass %)	High silt content links to medium soil water drainage
10	Sand content (50–2,000 μm) (mass %)	High sand content links to faster soil water drainage
11	Wilting point water content (% vol)	Affects the total water capacity of soil
12	Absolute depth to bedrock (in cm)	Potential thickness of soil and hence unsaturated zone
13	Average annual precipitation (mm $year^{-1}$)	Most important climate parameters that affect the soil moisture variability
14	Mean annual temperature (°C)	
15	Leaf Area Index (m^2 m^{-2})	Measure of the leaf transpirative surface. Higher LAIs are assumed to correlate to higher transpiration rates and hence to more heavily vegetated regions

Note. Parameters 1–12 are available from ISRIC for 0-, 0.05-, 0.15-, 0.3-, 0.6-, 1-, 2-m depths, and their expected impact to the soil moisture accounting.

2.1.3. Soil, Climatic, and Vegetation Data Sets

The value of the SWI's *T* parameter is related to the soil drainage characteristics, the capillary rise processes that lead to water evaporation, as well as transpiration processes that deplete water for the soil water storage via canopy. In this line, a broad set of physical parameters that characterize parts of the above-mentioned processes were considered. These parameters describe the hydraulic characteristics of the soil and control the water movement within it. In particular, the available soil water capacity is a descriptor of the effective storage of water in the soil (Chandler et al., 2017). The saturated water content describes the storage beyond the field capacity of the soil (O'Geen, 2013). The bulk density is a metric of how densely packed are the soil ingredients which can affect the porosity of the soil (Archer & Smith, 1972). The soil textures (sand/silt/clay fractions) are essential descriptors of the hydraulic transmissivity (O'Geen, 2013). Soil organic carbon is known to affect the hydraulic conductivity of a soil (Nemes et al., 2005). The coarse fragments describe the hydraulically inactive fragment of the soil, but also affect negatively the soil water retention capabilities (Baetens et al., 2009). All these soil parameters were obtained from the International Soil Reference and Information Center (ISRIC)—SoilGrids1km data set (Hengl et al., 2014). Furthermore, in order to consider the plant transpiration-related water extraction from the soil, the Leaf Area Index (LAI) from Copernicus Global Land Service (Smets et al., 2018) was also taken into consideration, in a spatial resolution of 1 km for the period 2015–2019. Finally, climatological characteristics were considered by including mean annual precipitation and temperature as derived from WFDEI Forcing reanalysis data (Weedon et al., 2018). This data set uses ERA-Interim reanalysis, bias adjusted on Global Precipitation Climatology Centre observations. Average annual precipitation and temperature were considered for 1979–2016. An overview of the considered parameters is provided in Table 1.

2.1.4. The ERA5-Land and FLDAS Reanalyses Soil Moisture Products

The ERA5-Land (or E5L) is a reanalysis data set providing high-resolution (~9 km) information of surface variables over several decades between 1981 and near present. The ERA5-Land data is produced by the tiled ECMWF Scheme for Surface Exchanges over Land incorporating land surface hydrology (H-TESSEL) and becomes available through the Copernicus Climate Data Store (Copernicus, 2019). The ERA5-Land is

a rerun of the ERA5 climate reanalysis, employing a series of improvements making it more accurate for all types of land applications, for example, soil moisture estimation. The data used in this study are those produced by the CY45R1 version of the Integrated Forecasting System. The soil moisture in the E5L is simulated in the four levels of the ECMWF surface model, with respective depths: (a) Layer 1: 0–7 cm, (b) Layer 2: 7–28 cm, (c) Layer 3: 28–100 cm, and (d) Layer 4: 100–289 cm. More details about the model can be found in the ECMWF site (ECMWF, 2020).

The FLDAS (McNally, 2018; McNally et al., 2017) is a variant of the NASA Land Information, optimized for agricultural drought assimilation (i.e., soil moisture) in regions with sparse observations, such as Africa, Central America, and Central Asia. The land surface scheme that it uses is *e*, while the meteorological forcings derive from Modern-Era Retrospective analysis for Research and Applications version 2. FLDAS data set provides soil moisture and other variables from 1982 to the near present in a 0.1° resolution. The soil moisture in the FLDAS is simulated in the four levels of the ECMWF surface model, with respective depths: (a) Layer 1: 0–10 cm, (b) Layer 2: 10–40 cm, (c) Layer 3: 40–100 cm, and (d) Layer 4: 100–200 cm.

2.2. Calibration of the SWI Using ISMN In Situ Soil Moisture Data

2.2.1. Soil Water Index Estimation

The convolutive transformation of the surface soil moisture to SWI is described in Equation 1:

$$\mathrm{SWI}(t) = \frac{\sum_i m_s(t_i) e^{(t-t_i)/T}}{\sum_i e^{(t-t_i)/T}} \quad \text{for } t_i \le t \tag{1}$$

where m_s is the surface soil moisture estimate from the remote-sensed soil surface moisture at time t_i and parameter T represents the time scale of soil moisture variation (Ceballos et al., 2005). Details about the recursive SWI estimation from SSM can be found in Wagner et al. (1999) and Marschallinger and Paulik (2019). The implementation of the SWI estimation in this work was performed accordingly to Albergel et al. (2008) and Stroud (1999).

2.2.2. Calibration of *T* Value

The T parameter of SWI was calibrated for a constrained number of available ISMN stations, against the ESA CCI SM data. The selection was based on data availability and ISMN gauging depth criteria, that is, (a) only ISMN time series, points that both ISMN data and ESA CCI SM time series provided at least 365 concurrent daily values were considered from 2001 onwards, (b) only ISMN time series referring to soil depths larger than 0.1 m were considered to avoid a biased calibration towards shallow depths, as 0.1 m soil moisture dynamics are regarded to be well approximated by the surface remote-sensed moisture (Manfreda et al., 2014). The 2001 was selected as the starting year of the analysis due to the much wider availability of the ISMN data (Figure 2b), as well as the expanded coverage of the ESA CCI SM data after the 2001 with the integration of the AMSR-E product (Dorigo et al., 2017). These constrains limited the number of the utilized ISMN time series, from 6,230 to 3,596 (ISMN data locations with more than one soil moisture measurement depths are counted separately). The optimization was performed using unconstrained Nelder–Mead optimization algorithm (Lagarias et al., 1998), such that the coefficient of determination (R^2) between the monthly averaged ISMN measurements and the estimated SWI aggregates is maximized. Further, an optimization tolerance in the T value of 0.5 was used which means that when the solver attempted to take a step that is smaller than 0.5 comparing to the previous estimation, the optimization procedure ended. The monthly aggregates were considered in order to minimize the effect of outliers in either daily ISMN or daily SWI data. Although a range of methods for scaling the SWI to a target measured mean and standard deviation, for example, linear regression (Jackson et al., 2010), linear transformation (Brocca et al., 2010), and CDF matching (Brocca et al., 2011) are available, none of them can significantly alter the correlation coefficient (Paulik et al., 2014). Here, we employed the SWI linear transformation of Paulik et al. (2014). We refer to the best T parameter value obtained, as *T_opt*.

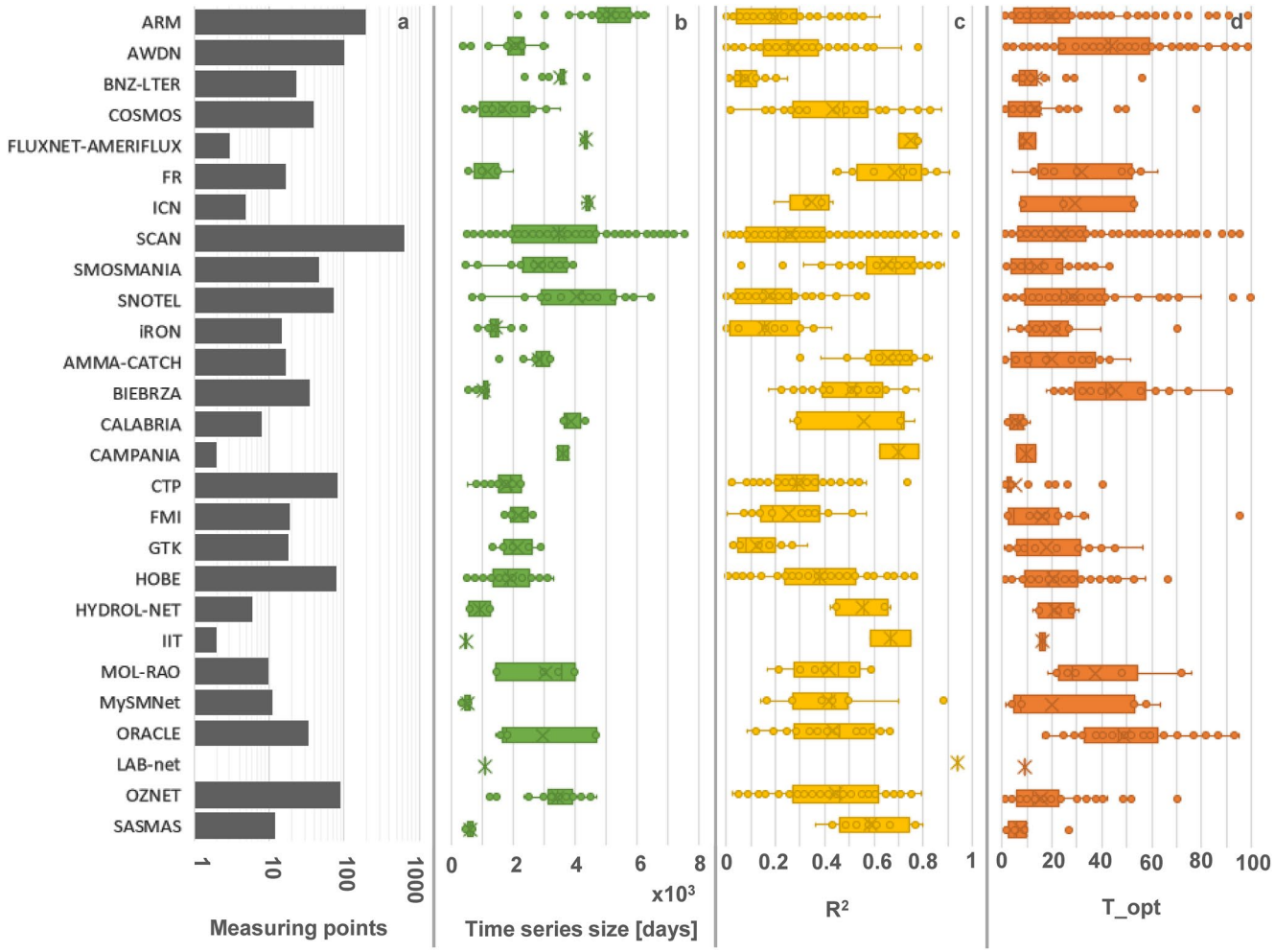

Figure 3. Results of the T calibration along with characteristics of each network time series. In (a), the number of field measurement points per network is shown; (b) shows the size of each network time series (in days). In (c), the attained R^2 is shown for each network and in (d) the optimum T obtained.

2.2.3. Artificial Neural Network Design

A neural network is an adaptive machine learning system that learns by using interconnected nodes or neurons in a layered structure that resembles a human brain. A neural network can learn from data—so it can be trained to recognize patterns, classify data, and perform forecasts. The feed-forward ANN (also known as Multilayered Network) was employed. The network was implemented in Matlab Deep Learning Toolbox (Hudson Beale et al., 2020). The configuration included two hidden layers that use a hyperbolic tangent sigmoid transfer function and a log-sigmoid transfer function, respectively. Gradient descent with momentum and adaptive learning rate was selected as backpropagation training algorithm. The use of log-sigmoid function does not allow for negative outputs (i.e., negative$T$ values) to be assigned. The network equation can be described as in Equation 2.

$$y = tansig\left(W_2\left[logsig\left(W_1X + B_1\right)\right] + B_2\right) \tag{2}$$

where X is the input data deriving from Table 1, while the target y is the optimum T value as it was calibrated on the ISMN data. W_1, W_2, and B_1, B_2 are the weight and bias matrixes for the two layers used.

The ANN training strategy was based on the minimization of the root mean square difference (RMSD) between the T_opt and the T value estimated by the ANN (hereafter referred as T_est). The variables shown

Figure 4. All the International Soil Moisture Network (ISMN) measurement points initially considered from the networks in the Acknowledgments section (all circles) and the 353 points used for the ANN training (red cycles). ANN, artificial neural network.

in Table 1, as well as the depth of the ISMN time series that provided the *T_opt*, were used as predictors. The *T_opt* was used as the target output on which the ANNs were trained. Due to the scale mismatch between the in situ ISMN measurements and the remotely sensed data, the assumption was made that only the ISMN points that provided a decent calibration would represent well the ESA CCI SM grid-cell-level soil moisture. For the purposes of this work, only the *T_opt* that had obtained a coefficient of determination R^2 of at least 0.5 were selected, employing the concept of temporal stability (Cosh et al., 2006; Vinnikov et al., 1999) which assumes that well performing in situ sensors can provide accurate and efficient estimation of large-scale surface soil moisture.

For the ANN optimization, a 20% of the predictors and targets were kept for validation, while also the rest 80% were split into 70% for training (training data set) and 30% as a test data set. The soil-related predictors' depths were selected to be closest to each depth of the ISMN from which each *T_opt* occurred. The available depths of the predictors are shown in Table 1.

2.3. Regionalization— Upscaling

The trained ANN was then used along with the 1-km SoilGrids data, LAI, as well as the downscaled WFDEI climate data, to estimate the *T* value in the respective 1-km resolution in global scale. The estimated global-scale *T* value in 1-km resolution was masked for the regions that the descriptors (Soilgrids, LAI, climate) fall within the range of the calibration points descriptor values. Then, the *T* value was upscaled to the ESA CCI SM resolution by estimating the median of 1 km *T* values within each 0.25° grid cell. Finally, the upscaled *T* was used to estimate the root-zone SWI (rSWI).

3. Results

3.1. SWI *T* Value Calibration

The calibration of *T* value provided diverse results among the 3,596 different ISMN time series. Many of the calibrations (1,967) did not converge to any *T_opt*, resulted to unrealistically high or zero *T_opt* values or did not provide data for the period 2001–2018. We consider that in these cases, the ISMN measurements were not representative of the ESA CCI grid cell measurements and hence the optimization could not converge to a rational result. Here, we present all the calibrations that resulted to *T_opt* between 1 and 100, that correspond to a total number of 1,629 time series. Among them, the R^2 ranged between near 0

Table 2
Information for the 27 In Situ Soil Moisture Measurement Networks

Network	Measurement points	Average time series length (days)	Average R^2	Average T_opt
ARM	195	5,102	0.20	18
AWDN	102	2,070	0.27	40
BNZ-LTER	23	3,490	0.08	15
COSMOS	40	1,667	0.43	17
FLUXNET-AMERIFLUX	3	4,317	0.75	7
FR	17	1,177	0.68	27
ICN	5	4,402	0.35	26
SCAN	638	3,466	0.26	20
SMOSMANIA	46	2,831	0.65	13
SNOTEL	74	4,042	0.18	27
iRON	15	1,440	0.15	24
AMMA-CATCH	17	2,798	0.66	21
BIEBRZA	35	1,043	0.51	43
CALABRIA	8	3,842	0.56	6
CAMPANIA	2	3,594	0.70	10
CTP	81	1,768	0.29	6
FMI	19	2,159	0.26	19
GTK	18	2,135	0.12	23
HOBE	80	1,906	0.38	19
HYDROL-NET	6	899	0.55	19
IIT	2	455	0.67	16
MOL-RAO	10	3,032	0.42	33
MySMNet	11	500	0.42	21
ORACLE	33	2,940	0.43	46
LAB-net	1	1,069	0.94	7
OZNET	89	3,440	0.44	15
SASMAS	12	595	0.59	7

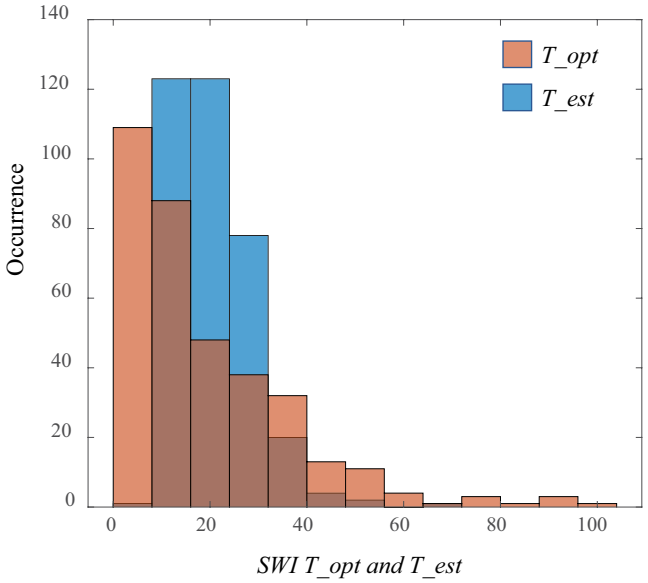

Figure 5. Histogram of the T_opt (blue) and the ANN-obtained T_est (blue). ANN, artificial neural network.

and 0.94. The average R^2 was 0.31 with a standard deviation of 0.27. In Figure 3, the results per network are illustrated. The best mean R^2 was obtained for LAB-net network data; however, this network provides only one time series of measured soil moisture (Table 2). Next best performing networks in the calibration process were FLUXNET-AMERIFLUX, CAMPANIA, FR, IIT, and AMMA-CATCH with R^2 higher than 0.65. The worst performing networks in the calibration process were BNZ-LTER, GTK, iRON, SNOTEL, and ARM with average R^2 less than 0.2. For 353 calibrations, R^2 exceeded 0.5, for 207 and 110 calibrations it exceeded the 0.6 R^2 and 0.7 R^2, respectively.

At a final selection stage, only the *T_opt* that provided $R^2 > 0.5$ were kept for training the ANNs, that is, 353 values of *T_opt*. Figure 4 shows the locations of all the ISMN measurement points initially considered and the 353 finally screened out points from which the *T_opt* was used for the ANN training. It is shown that the 353 points still provide good spatial coverage of the networks located in North America and Alaska, west and central Europe, and Australia. Nevertheless, Eastern Europe and Asia networks are almost entirely unrepresented, as well as already underrepresented South Africa and South America. The average and the median depth of the 353 points was 0.35 and 0.22 m, respectively, denoting that measurement depth is skewed toward the shallow layer.

3.2. Neural Network Results

The ANN was calibrated and validated on the 353 preselected points using the RMSD between *T_opt* and *T_est* as an objective function. The ANN obtained a RMSD of 15.85 and 14.24 for the calibration and validation set, respectively. The distribution of the *T_opt* and the ANN derived *T_est* is shown in Figure 5. It is shown that the ANN estimations do not capture well the low and high *T* values, which explains the RMSD values of the calibration and validation procedure.

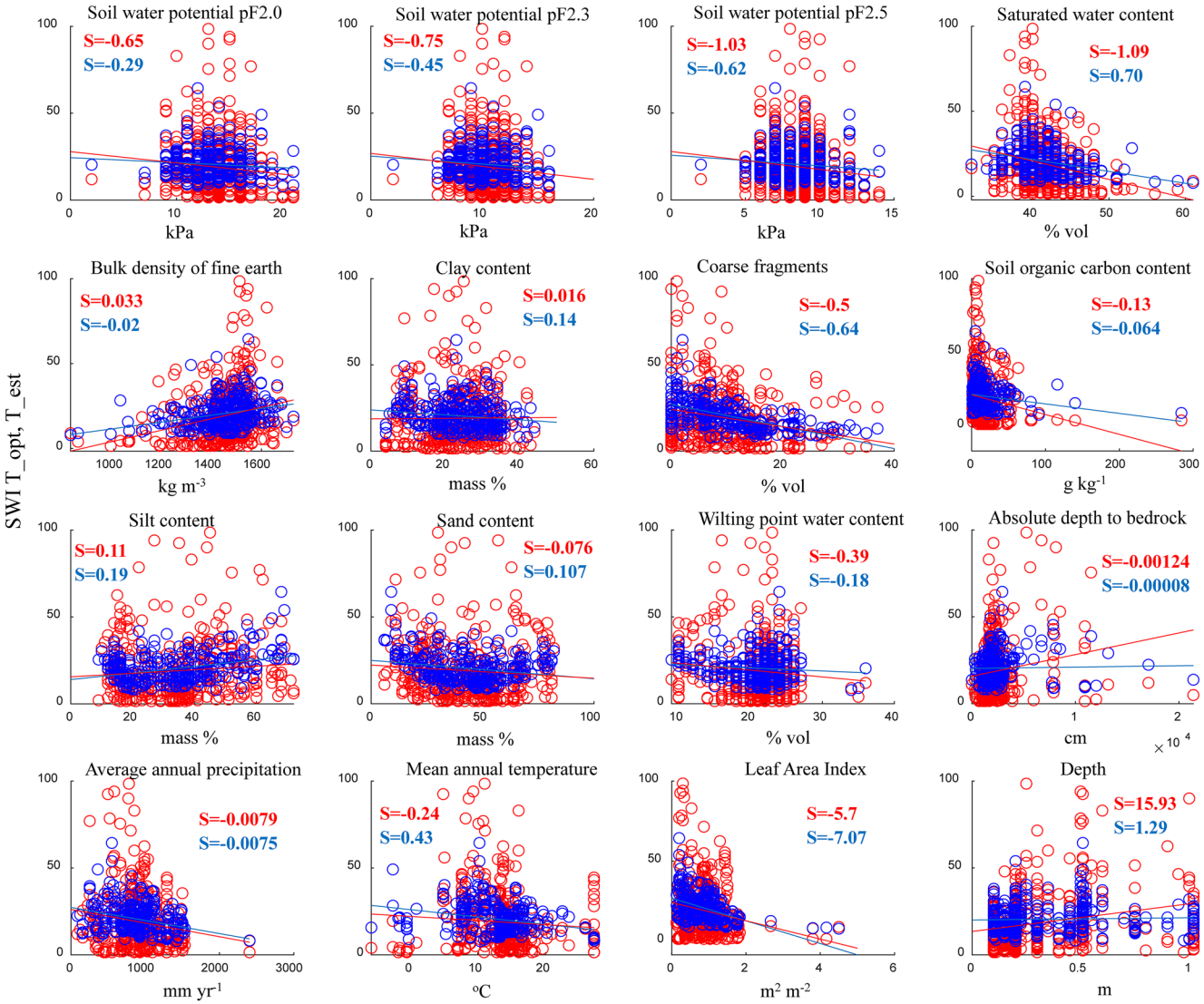

Figure 6. In red, the calibrated *T* on the 353 ISMN points (T_opt) and in blue the estimated *T* parameter at each ISMN points using the ANN (T_est), versus the the values of the 16 physical parameters that were included into the ANN. Respective linear trends are indicated with lines, and their slope *S* are reported. ISMN, International Soil Moisture Network; ANN, artificial neural network.

The results of the ANN estimation for the all the 353 calibration/validation points (*T_est*) are shown in Figure 6 with blue and compared to the *T_opt* (in red). The results show a general agreement in the slope, indicating that the ANN inhered the trend of the observed physical parameters. Among the different parameters, soil organic carbon content and depth to bedrock of Table 1 show the least agreement in the slope, even if they exhibit the same trend sign. Also, other parameters that exhibit smaller but noteworthy difference in slope are the available soil water potential in pF2.0, clay content, wilting point water content, and mean annual temperature.

3.3. Regionalization of the *T_est* Using ANNs

The calibrated ANN was then used to estimate the *T* value. As input to the ANN, the 1 km SoilGrids and LAI data were used, as well as the downscaled WFDEI climate data (nearest neighbor downscaled to the same spatial resolution). For the depth parameter, the value of 0.3 m was used, as SoilGrids data are offered for 0.30 m depth, but also as a representative depth between the mean and the median depth of the 353 points

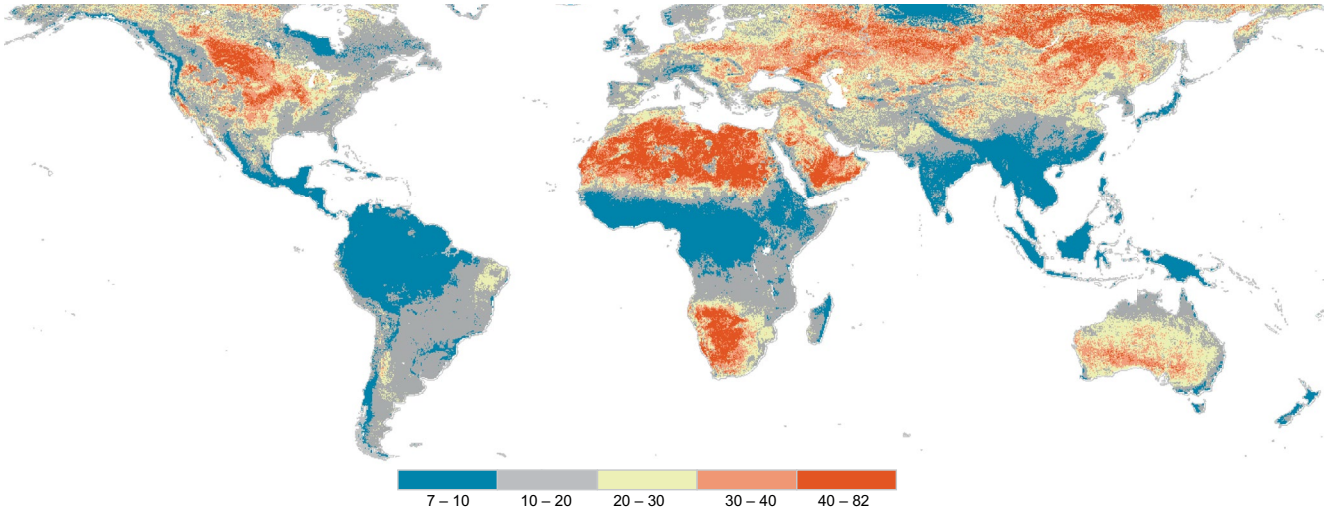

Figure 7. Estimated *T* (T_est) at 1-km spatial resolution.

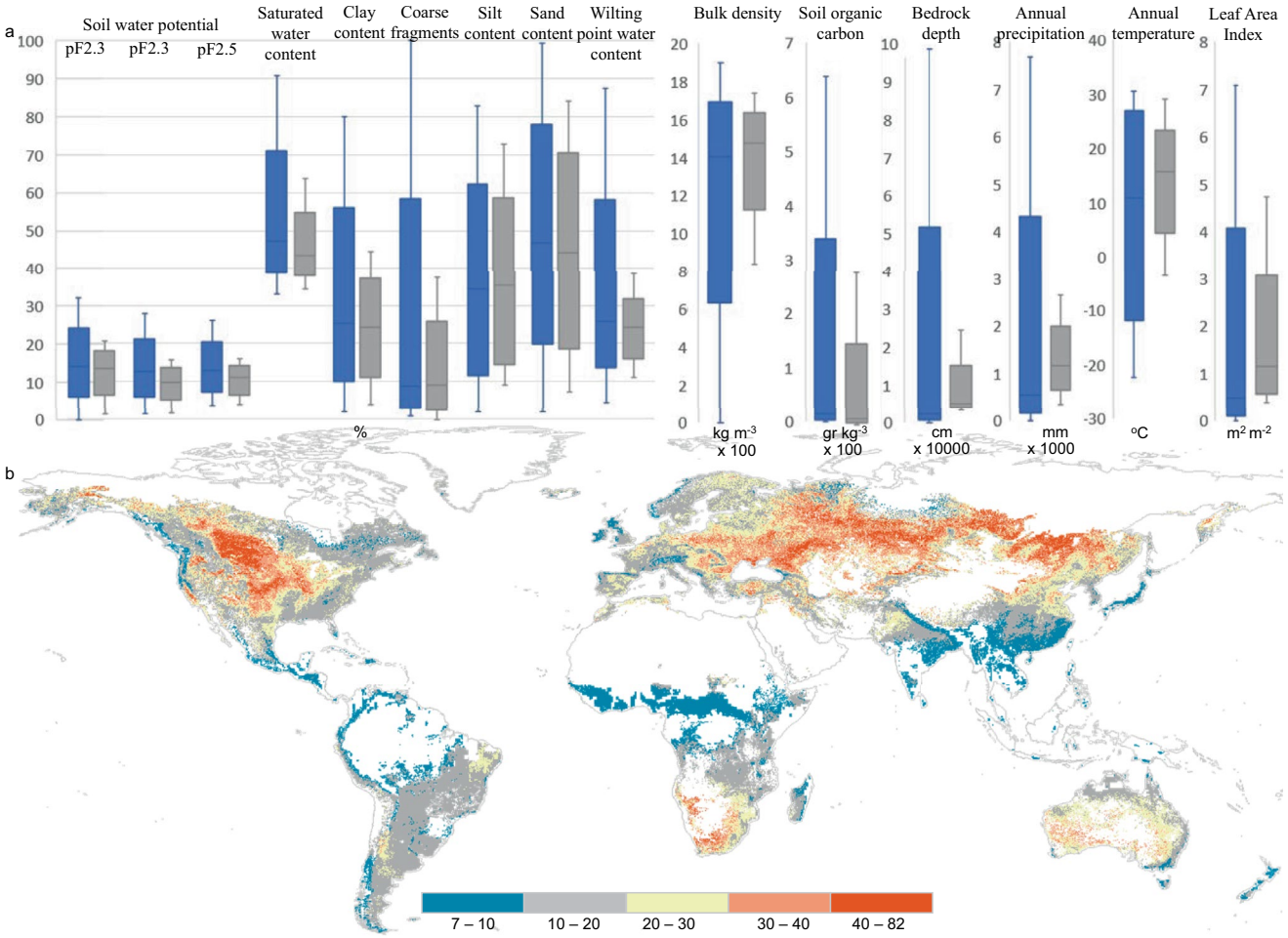

Figure 8. (a) Distribution of the soil, climate, and LAI parameters at global scale (blue) compared to the respective 353 ISMN locations distribution (gray). (b) Same with Figure 7 but for the regions that the soil, climate, and LAI parameters fall within the 353 ISMN locations' range. LAI, Leaf Area Index; ISMN, International Soil Moisture Network.

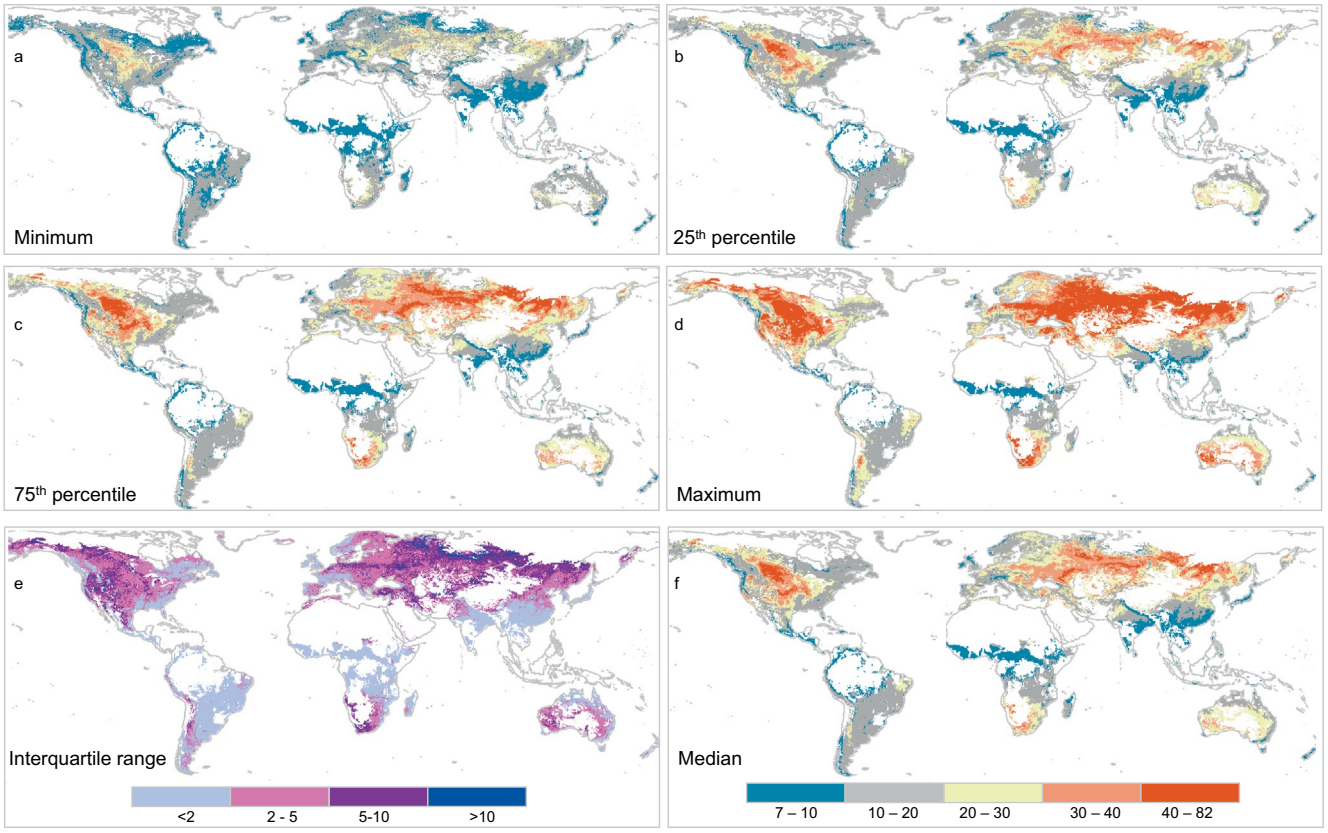

Figure 9. Statistics of T_est within the grid mesh of ESA CCI SM. ESA CCI SM, European Space Agency Climate Change Initiative soil moisture.

used for the calibration–validation (0.35 and 0.22 m, respectively). The result of the ANN application is a global estimation of the *T* value at 1-km spatial resolution (Figure 7).

Nevertheless, the relatively limited number of 353 points that were used for the ANN calibration and validation exhibits inherent spatial coverage limitations and may not be representative for the soil, climate, and LAI data range at the global scale. In Figure 8a, the global distribution of the soil, climate, and LAI parameters is compared against the 353 ISMN locations. In order to avoid the extrapolation of the *T* estimation to regions where the 15 parameters (Table 1 parameters) are outside the limits of the 353 ISMN locations, we masked out those regions. The masked result of Figure 7 is obtained, as shown in Figure 8b. It is shown that 47% of the area is outside the 353 points' parameter range, comprised mainly of tropical rainforests with high LAI values due to dense canopy such as the Indo-Malayan Archipelago and the Amazon and Congo Basins, desert regions such as the Sahara, the Arabian, and the Taklamakan Deserts, as well as cold region in the northern latitudes.

To estimate SWI using the ANN, the upscaling of the *T_est* is required from the spatial scale of ~1 km to the ESA CCI SM 0.25°. Within each 0.25° grid cell, there are 900 *T_est* values estimated by the ANN. In order to upscale and at the same time preserve variability information within each grid cell, the minimum, 25th, median, 75th, and maximum *T_est* for each grid cell were estimated (Figures 9a, 9b, 9f, 9c, and 9d, respectively). The statistics were obtained for each ESA CCI SM grid cell with at least 450 values.

Figure 9 shows that while the minimum and maximum T_est within each grid cell mainly exhibit a noteworthy deviation from the median values, the interquartile range is low and mainly below 10. This indicates that half of the T_est values within each grid cell are mainly within a ±5 range from the median.

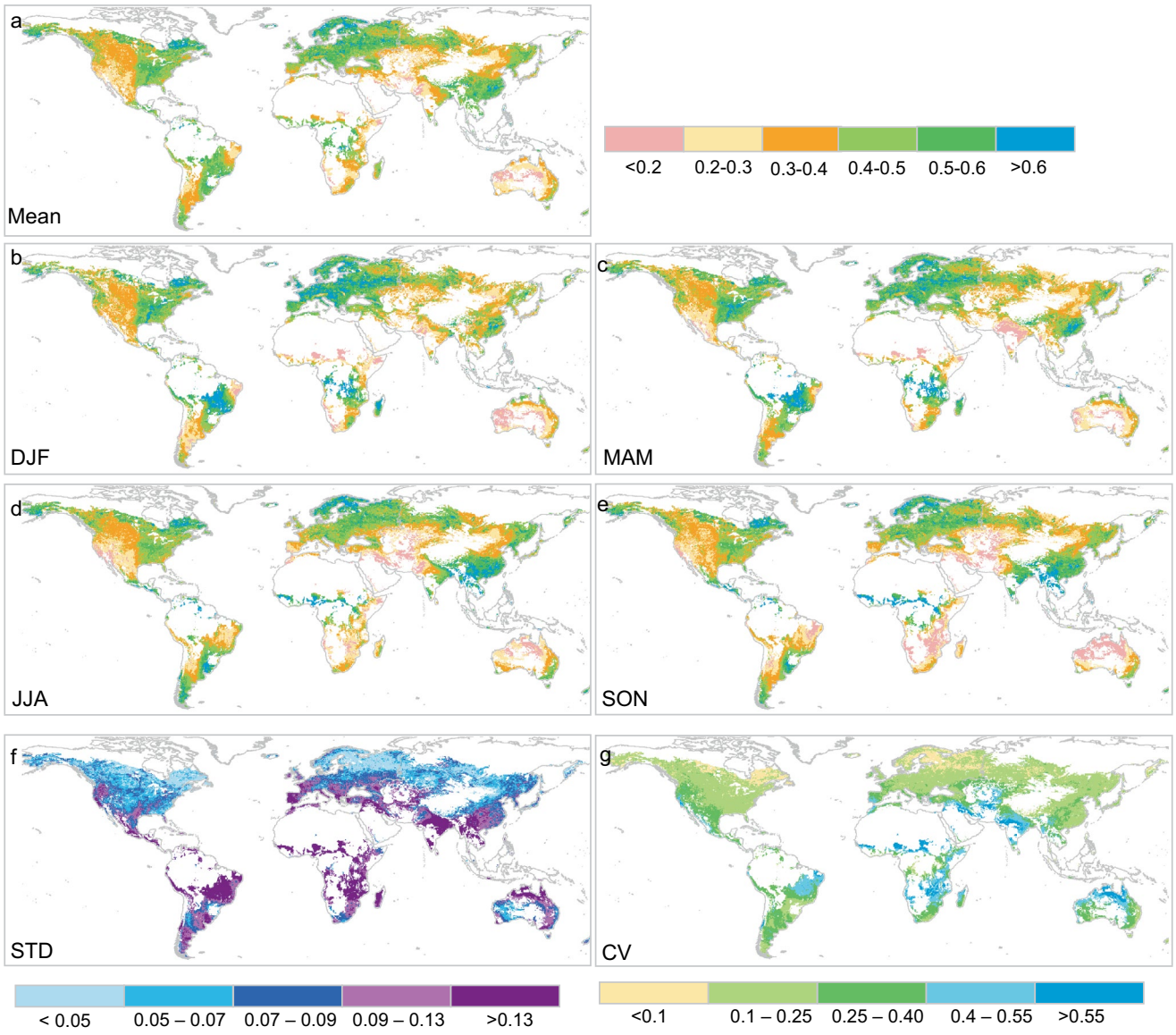

Figure 10. Annual and seasonal mean, standard deviation, and coefficient of variation of the rSWI, between 2001 and 2018. rSWI, root-zone Soil Water Index.

3.4. SWI Estimation

The rSWI was then estimated using the median value of the T_est produced by the ANN and the ESA CCI SM, for the period 2001–2018. The rSWI annual mean, seasonal means, standard deviation, and the coefficient of variation are shown in Figure 10. The annual mean results (Figure 10a) clearly show known climatological patterns, such as the contrast between west and east conterminous United States, the wet gradient from Mediterranean to northern Europe, the west-east Australia contrast, as well as that of north-south China. Furthermore, the seasonal averages (Figures 10b–10e) show the interexchange between different seasons. Climatological aspects are also depicted through the standard deviation of the rSWI, with regions of higher seasonal precipitation variation to exhibit higher values (Figure 10f), as well as higher coefficient of variation (Figure 10g).

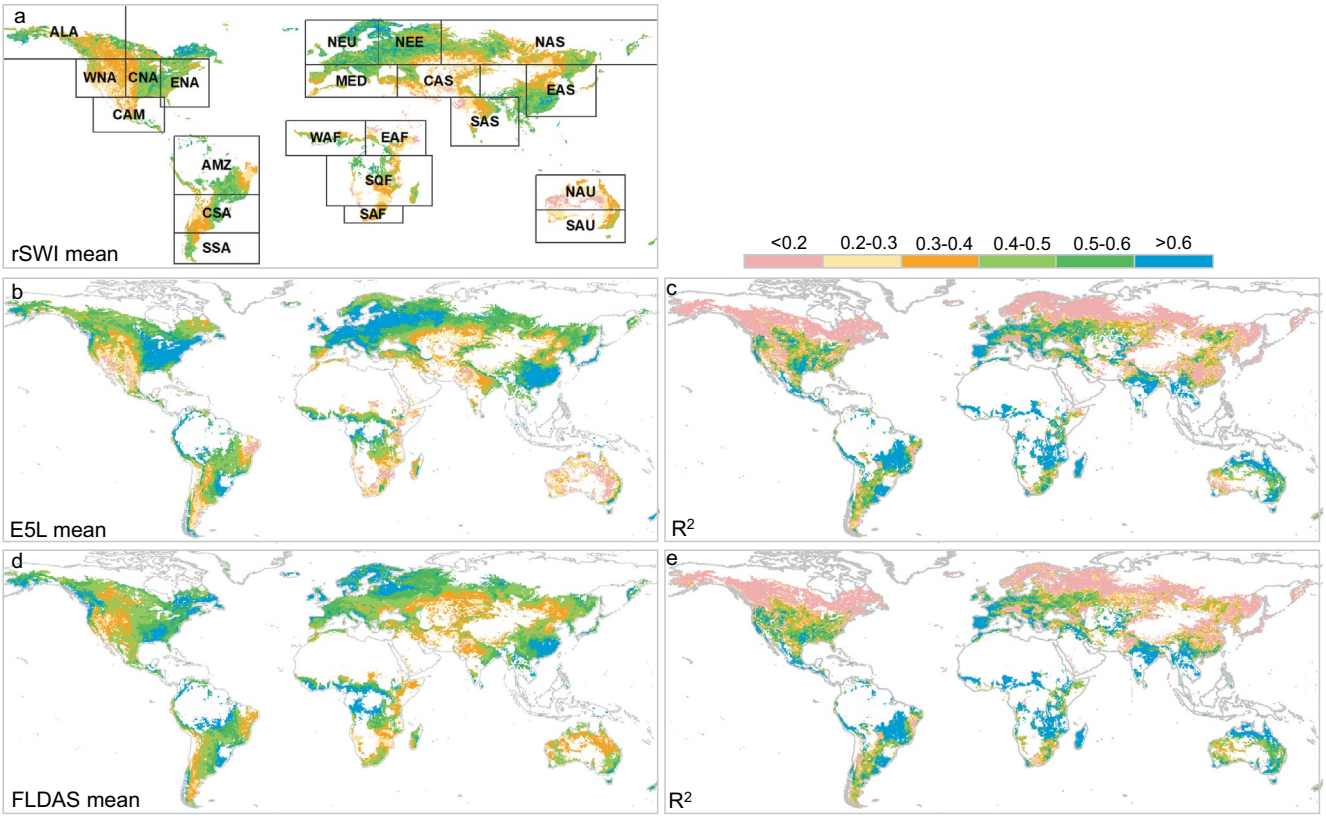

Figure 11. Mean rSWI (a), mean E5L (b), and mean FLDAS (d), for the period 2001–2018. (c, e) The R^2 between the monthly time series of rSWI–E5L and rSWI–FLDAS, for the period 2001–2018. In (a), the different regions are defined according to Giorgi and Bi (2005). rSWI, root-zone Soil Water Index; FLDAS, Famine Early Warning Systems Network Land Data Assimilation System.

3.5. rSWI Evaluation Toward E5L, FLDAS, and the Added Value of the ANN Results

To assess the ability of the estimated rSWI to describe root-zone soil moisture, the ERA 5 Land (E5L) and FLDAS reanalysis data were used. For both data sets, two soil levels were considered. In the case of E5L, the L2 that represents the depth between 7 and 28 cm and L3 that represents 28 and 100 cm were averaged and the outcome was considered as representative for the comparison. For the FLDAS, the two upper layers that represent the average soil moisture of 0–10 cm and 10–40 cm were considered, respectively. In Figure 11, mean rSWI (Figure 11a) is compared to scaled between 0 and 1 E5L and FLDAS soil moisture (Figures 11b and 11d). It is shown that all three data sets exhibit similar patterns, nevertheless, it is also shown that FLDAS but especially E5L exhibits wetter patterns, clearly visible mainly at the eastern contiguous United States, central and Eastern Europe, and East China. Regarding the similarity of the rSWI to the E5L, the R^2 was estimated, showing that for large regions especially in the midlatitudes, the similarity is large with R^2 values ranging over 0.5 (Figures 11c and 11e).

To further assess the regional performance, we make use of the different regions of the globe as defined in Giorgi and Bi (2005), as shown in Figure 11a. The temporal variation of the soil moisture is shown for each region in Figure 12. The rSWI is following the seasonal pattern of E5L and FLDAS with limited regions and periods of disagreement, for example, the first 5 years in the time series in the East Asia (EAS), North Asia (NAS), and Northeast Europe (NEE) as shown in Figure 12. Furthermore, as the regional means refer to normalized soil moisture, the tendency of the E5L and FLDAS to exhibit more extreme (near 0 or near 1) regional means is translated to a more uniform simulated output of E5L comparing to the rSWI results. This is especially pronounced in Alaska (ALA) and North Asia (NAS) regions, where rSWI estimation in the northern hemisphere winter months is not possible due to the extended snow cover. In contrast, the rSWI exhibits a very good similarity in terms of R^2 and RMSD in Mediterranean region (MED), Central America

Table 3
Seasonal R^2 Comparison Between rSWI and E5L, FLDAS Data per Region

		rSWI–E5L				rSWI–FLDAS			
Region		DJF	MAM	JJA	SON	DJF	MAM	JJA	SON
Alaska	ALA	0.24	0.00	0.00	0.06	0.10	0.00	0.01	0.02
Amazon Basin	AMZ	0.38	0.41	0.75	0.79	0.47	0.55	0.73	0.79
Central America	CAM	0.58	0.60	0.81	0.74	0.64	0.45	0.72	0.52
Central Asia	CAS	0.62	0.74	0.89	0.77	0.55	0.56	0.82	0.89
Central North America	CNA	0.69	0.24	0.77	0.65	0.55	0.27	0.69	0.60
Central South America	CSA	0.66	0.71	0.51	0.47	0.55	0.60	0.37	0.29
Eastern Africa	EAF	0.83	0.90	0.54	0.29	0.49	0.90	0.75	0.52
East Asia	EAS	0.23	0.24	0.09	0.33	0.00	0.03	0.54	0.30
Eastern North America	ENA	0.09	0.08	0.33	0.32	0.31	0.07	0.35	0.35
Mediterranean Basin	MED	0.52	0.64	0.84	0.91	0.37	0.65	0.78	0.84
North Asia	NAS	0.09	0.06	0.00	0.10	0.02	0.01	0.27	0.00
North Australia	NAU	0.81	0.85	0.68	0.83	0.84	0.78	0.54	0.38
Northern East Europe	NEE	0.17	0.04	0.24	0.58	0.19	0.04	0.31	0.52
Northern Europe	NEU	0.06	0.03	0.36	0.70	0.01	0.47	0.35	0.74
South Africa	SAF	0.55	0.46	0.75	0.66	0.52	0.34	0.54	0.44
South Asia	SAS	0.25	0.27	0.11	0.44	0.09	0.33	0.35	0.50
South Australia	SAU	0.81	0.65	0.70	0.78	0.73	0.71	0.49	0.79
South Equatorial Africa	SQF	0.64	0.42	0.61	0.64	0.71	0.51	0.44	0.59
Southern South America	SSA	0.31	0.39	0.19	0.48	0.73	0.61	0.14	0.75
Western Africa	WAF	0.67	0.87	0.74	0.81	0.41	0.74	0.80	0.76
Western North America	WNA	0.62	0.49	0.65	0.84	0.77	0.81	0.83	0.90

(CAM), Central Asia (CAS), Central North America (CNA), North Europe (NEU), and North Eastern Europe (NEE) regions. The seasonal correlations and differences are also shown in Table 3.

To assess the added value of the calibrated T using the ANN, the rSWI was estimated using fixed values of T at global scale and compared toward the E5L/FLDAS data. The fixed values were obtained from Paulik et al. (2014) who used the values of 1, 5, 10, 15, 20, 40, 60, and 100. The results were assessed against the E5L/FLDAS, in terms of R^2 and RMSD of the regionally aggregated monthly time series. An overview of the results shown in Figure 13 (and supplementary Tables S1–S4) indicates that the rSWI (as well as the SWI in general) does not agree with the E5L and FLDAS simulation not only in the regions of Alaska and North Asia (ALA, NAS) but also in North and South Australia (NAU, SAU) and West and East Africa (WAF, EAF). This is expected as those regions are dominated by permafrost, glaciers, dense forests, and deserts, in which the SWI approach is not suggested (Marschallinger & Paulik, 2019). In more detail, the results indicate that in terms of R^2, there are regions, where some of the fixed values tested performed better than the ANN calibrated *T_est*, both in terms of R^2 and RMSD, as they were assessed against the E5L data. Nevertheless, the weighted global average results show that none of the fixed T values provided better results against E5L in terms of R^2 and RMSD, while the same results are found comparing to the FLDAS, except T values of 1, 5, and 10 that marginally exhibit a better weighted R^2.

4. Discussion

First, the calibration of SWI's T value using ISMN data shows good results for a number of ISMN locations, with R^2 to range between 0 and >0.9. The results indicated that there are soil moisture measuring networks within ISMN that their measurements are more predictable than others by the SWI estimation on ESA CCI

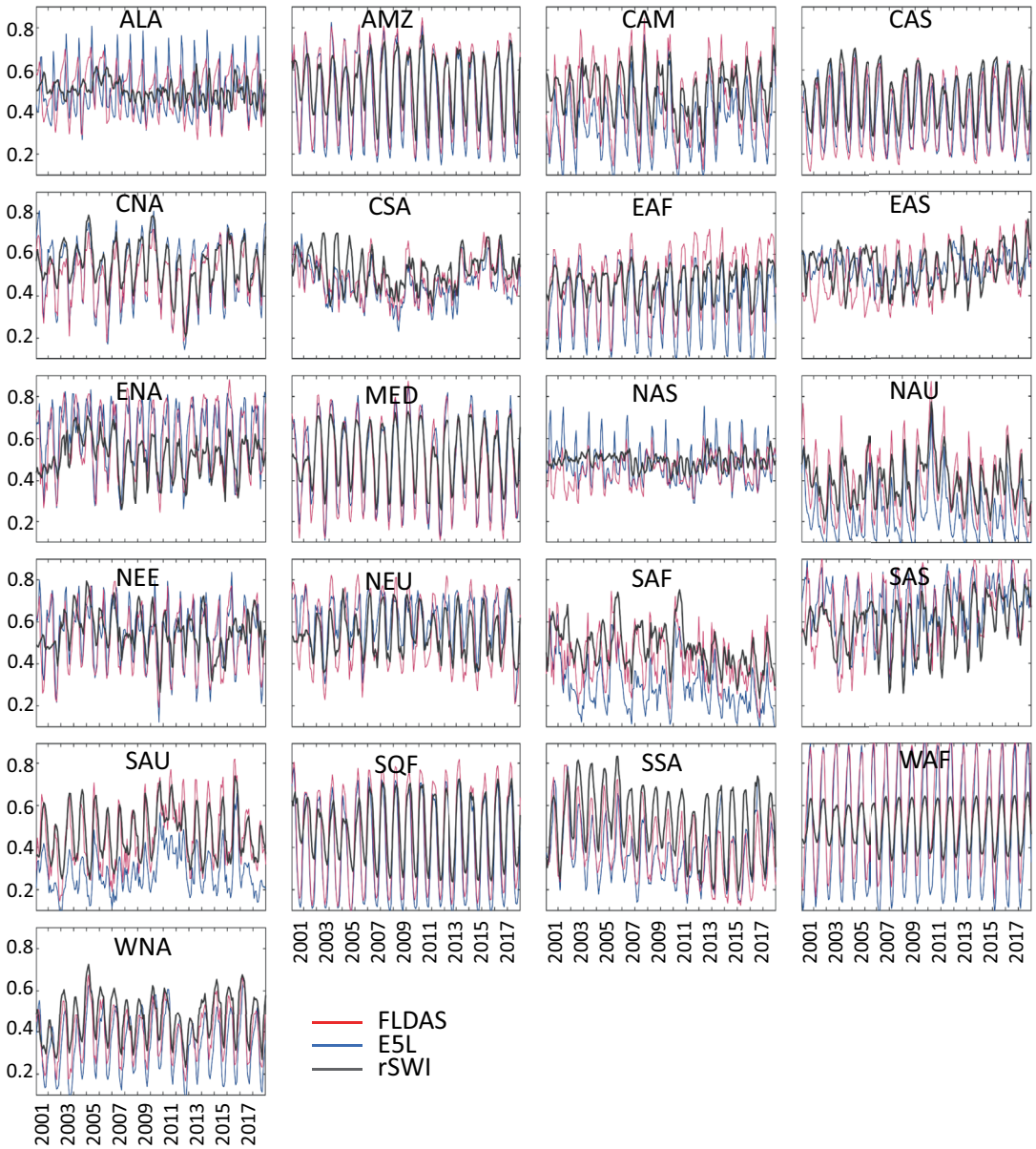

Figure 12. Regional means comparison of the rSWI and the E5L/FLDAS, between 2001 and 2018. rSWI, root-zone Soil Water Index; FLDAS, Famine Early Warning Systems Network Land Data Assimilation System.

SM data. This has been already shown in the study of Paulik et al. (2014), that also found similar results, that is, SWI performs the best with specific ISMN networks' data, such as FLUXNET and ITT, while worse correlation is observed with SNOTEL and ARM. The difference in the correlation performance does not necessarily mean that either product is problematic on its own but that local values are not representative to the remote-sensed information, which was one of the main foci of the study. The variability of the performance has been attributed to the influence of topography, water fraction, noise, and the in situ observation depth (Paulik et al., 2014). Also Paulik et al. (2014) quantified the negative effect of the freezing–thawing of the soil. Here, none of these effects were tested or considered in the calibration of T value, which might increase the performance of specific ISMN measuring locations/points. Further, it is worth mentioning that another potential reason would be the measuring instrumentation methodology that each ISMN uses, for example, neutron probes or permanent time domain reflectometry (TDR). In some cases, the TDRs have been found

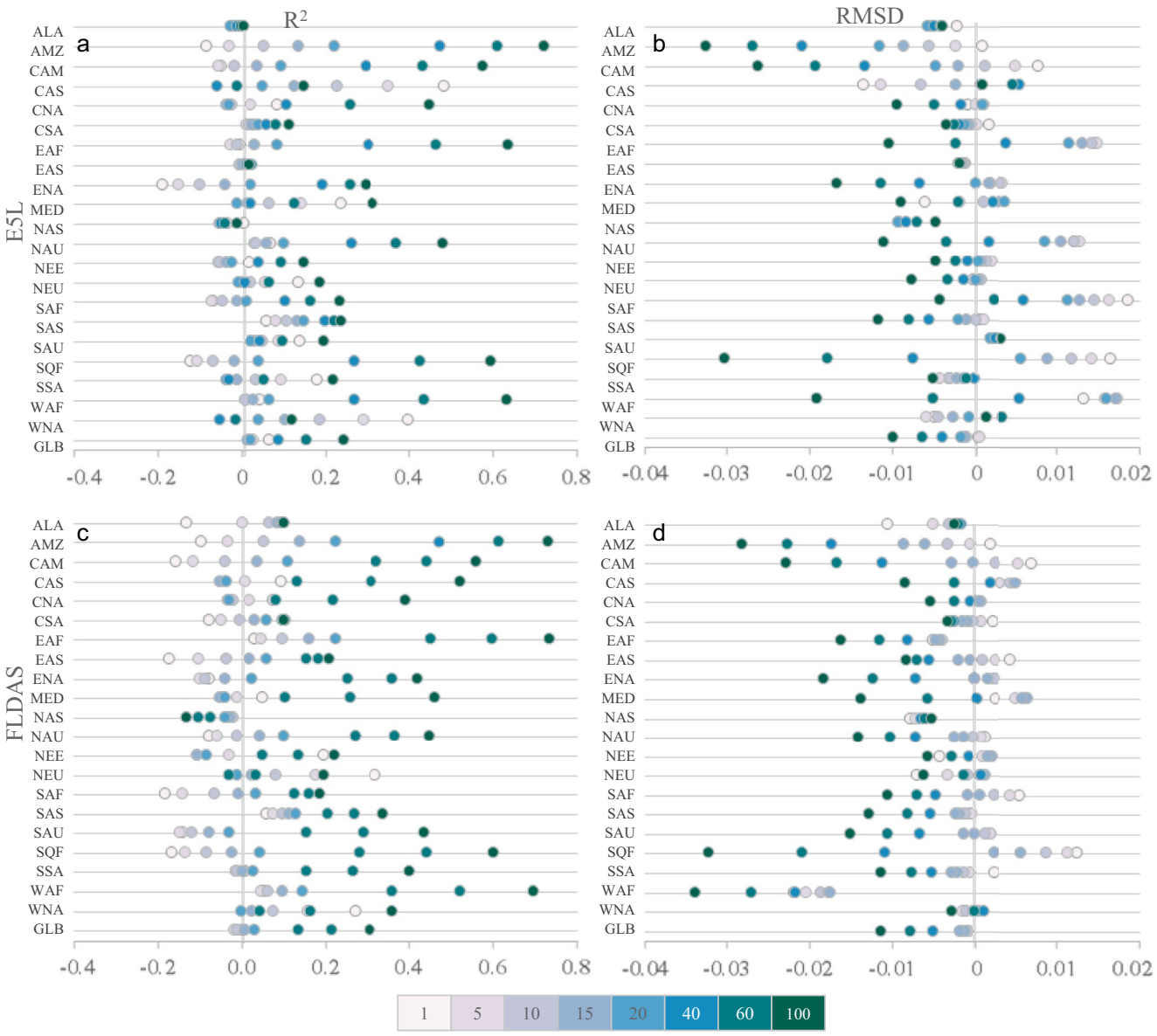

Figure 13. In (a), (c), the R^2 difference between the SWI estimation using fixed values of T and T_est (positive values denote that T_est is better). In (b), (d), the same as in (a) but for the RMSD between the fixed T value assessment and T_est (negative values denote that T_est is better). The color scale indicates the different fixed T values. Both R^2 and RMSD were estimated on the regional monthly time series of rSWI and the E5L/FLDAS, between 2001 and 2018. SWI, Soil Water Index; RMSD, root mean square difference; rSWI, root-zone Soil Water Index; FLDAS, Famine Early Warning Systems Network Land Data Assimilation System.

to be more accurate (Kinama et al., 2010). Hence, in situ measuring methods could add another source of uncertainty that affects the similarity of the SWI produced time series to simulate measured soil moisture.

The trained ANN shows a certain ability to simulate the effect of each physical descriptor in the T values, as shown by the calibration and validation RMSDs, and the slope of the linear regressions of Figure 6. Yet, the high RMSD in the calibration and validation of the ANN is translated as a low ability to simulate the range of T, that is, the higher and lower T values are generally not well simulated by the ANN, as shown in Figures 5 and 6.

The results of rSWI were assessed against E5L and FLDAS reanalysis data. While the comparison revealed useful insights about the ability of the SWI in general and that of the calibrated rSWI to simulate the month-

ly variation soil water, the results should be interpreted with caution. This is due to the large structural discrepancies between the rSWI and the E5L/FLDAS that make the comparison qualitative rather than a real benchmark. First, the E5L and FLDAS data sets provide process-based estimations of soil moisture, produced by the land surface simulations using the driving climate reanalyses variables. On the other hand, rSWI is a convolutive transformation of a remote-sensed observation and while it assimilates the soil moisture decay, it does not by any means integrate the physical processes involved. Furthermore, another important aspect that should be taken into consideration is that the construction of the ESA CCI SM is already including in a statistical sense, information of another reanalysis product, the GLDAS Noah land surface model, against which CDF matching is used to harmonize the climatologies of the individual remote sensing products. This has been found to impact magnitude in the long-term trend of the data (Gruber et al., 2019; Liu et al., 2011). While the trends were not investigated here, this also may add another source of error in the rSWI estimation. Beyond the limitations and the caveats, this study shown for the first time that spatially explicit soil, climate, and vegetation attributes can combinedly aid the SWI root-zone soil moisture estimation.

5. Conclusions

This study is a first attempt to leverage machine learning along with local climate soil and vegetation information to better process the ESA CCI SM data to represent root-zone soil moisture. This was tested through a series of processing and calibration/validations steps, that deems the overall methodology complicated and subject to a series of uncertainties. A major limitation is that we have worked under the assumption that ESA CCI SM data or the resulting SWI can act as a proxy of the ISMN point measuring data. Therefore, here we consider that a good agreement of SWI and ISMN for a limited number of locations is indicative for a much wider region. Another major assumption stems from the representativeness of the ISMN points that are eventually used to train and test the ANN. While there is a wide range in the soil/climate/LAI parameters represented by the 353 points selected for the ANN training and testing, many more effects that remain unexplored may lay well outside the 353 points range. Also, the moderate performance of the ANN model infers that besides the 16 variables used in this study additional predictors may exist. Other limitations stem for the soil, climate, and the LAI data set that was used, that are subject to their own uncertainties. Especially, the SoilGrids soil data set that formed the basis of the predictors in the ANN development (Hengl et al., 2014) indicates that the regression models used to map the soil properties are mainly account for the 20%–50% of observed variability, which indicates that the soil information also consists a main source of uncertainty. Additionally, Hengl et al. (2017) discuss that specific variables such as the coarse fragments and depth to bedrock are represented with high levels of error in the SoilGrids 250 m data set. To this end, given the reliance of the presented analysis on the SoilGrids data set, it may be the largest source of uncertainty.

Besides these limitations, this study yields acceptable results in the root-zone soil moisture, as compared to the E5L and FLDAS reanalyses, especially in the midlatitude regions. We also confirm that SWI can effectively simulate in situ soil moisture measurements for a large number of the ISMN measuring locations. Future research would include additional soil moisture driving parameters such as a more detailed vegetation structure instead of the LAI parameter and the addition of more climate specific parameters instead of the mean annual values. Furthermore, significant impact would be expected by testing different soil data sets, such as the LUCAS topsoil data (Orgiazzi et al., 2018) for the EU region. Finally, further research would consider different types of machine learning models that can also deliver predictor importance such are random forests, as well as different objective function in the calibration of the SWI's T parameter using the ISMN observations.

Data Availability Statement

The FLDAS data used in this study were acquired as part of the mission of NASA's Earth Science Division and archived and distributed by the Goddard Earth Sciences (GES) Data and Information Services Center (DISC). The use of the invaluable ISMN data (W. A. Dorigo et al., 2011, 2013) were used in this study and also each ISMN measuring network separately, that is, ARM (www.arm.gov), AWDN (http://www.hprcc.unl.edu/awdn.php), BNZ–LTER (Van Cleve et al., 2015), COSMOS (Zreda et al., 2012), FLUX-

NET-AMERIFLUX (DAAC, 2013), FR (http://www.inrae.fr/en), ICN (Baldocchi et al., 2001; Hollinger & Isard, 1994), SCAN (AREA & UINTAH, 2010), SMOSMANIA (Albergel et al., 2008; Calvet et al., 2007), SNOTEL (Leavesley et al., 2008), iRON (Osenga et al., 2019), AMMA-CATCH (Cappelaere et al., 2009; Mougin et al., 2009; Pellarin et al., 2009), BIEBRZA (http://www.igik.edu.pl/en), CALABRIA (Auddino et al., 2015; Brocca et al., 2012; W. A. Dorigo et al., 2011), CAMPANIA (Brocca et al., 2011), CTP (Yang et al., 2013), FMI (Rautiainen et al., 2012), GTK (Geological Survey of Finland, www.gtk.fi), HOBE (Bircher et al., 2012), HYDROL–NET (Department of Civil and Environmental Engineering of Perugia, http://www.dica.unipg.it/DICA/ricerca/), IIT (www.iitk.ac.in/), MOL-RAO (Beyrich & Adam, 2007), MySMNet (University Technology Malaysia University Technology Malaysia, www.utm.my), ORACLE (Institut national de recherche en sciences et technologies pour l'environment et l'agriculture, https://bdoh.irstea.fr/ORACLE/), LAB-net (Mattar et al., 2014), OZNET (Rüdiger et al., 2007; Young et al., 2008), UMBRIA (Brocca et al., 2008, 2009, 2012), and SASMAS (Rüdiger et al., 2007; Young et al., 2008).

Acknowledgments
We acknowledge Copernicus Climate Change Service (C3S) and Copernicus Climate Data Store (CDS) for the ERA 5 Land data and more specifically the parameters Volumetric soil water layer 2 and layer 3, from the "ERA5-Land monthly averaged data from 1981 to present" data set, with Publication date 2019-07-12 and http://doi.org/10.24381/cds.68d2bb30. We acknowledge the use of the invaluable ISMN data (W. A. Dorigo et al., 2011, 2013) that were used in this study and also each ISMN measuring network separately, that is, ARM, AWDN, BNZ-LTER, COSMOS, FLUXNET-AMERIFLUX, FR, ICN, SCAN, SMOSMANIA, SNOTEL, iRON, AMMA-CATCH, BIEBRZA, CALABRIA, CAMPANIA, CTP, FMI, GTK, HOBE, HYDROL-NET, IIT, MOL-RAO, MySMNet, ORACLE, LAB-net, OZNET, and SASMAS (see table below).
This research is cofinanced by Greece and the European Union (European Social Fund [ESF]) through the Operational Programme "Human Resources Development, Education and Lifelong Learning" in the context of the project "Reinforcement of Postdoctoral Researchers - 2nd Cycle" (MIS-5033021), implemented by the State Scholarships Foundation (IKY) under the grand agreement no. 2019-050-0503-16972. IND contributed to this research in the context of "DRip Irrigation Precise—DR.I.P.: Development of an Advanced Precision Drip Irrigation System for Tree Crops" (Project Code: T1EDK-03372) which is cofinanced by the European Union and Greek national funds through the Operational Program Competitiveness, Entrepreneurship and Innovation, under the call RESEARCH—CREATE—INNOVATE. A.G.K. acknowledges partial support by the COST Action CA19139: PROCLIAS, supported by COST (European Cooperation in Science and Technology).

References

Achieng, K. O. (2019). Modelling of soil moisture retention curve using machine learning techniques: Artificial and deep neural networks vs support vector regression models. *Computers & Geosciences*, *133*, 104320. https://doi.org/10.1016/j.cageo.2019.104320

Albergel, C., de Rosnay, P., Gruhier, C., Muñoz-Sabater, J., Hasenauer, S., Isaksen, L., et al. (2012). Evaluation of remotely sensed and modelled soil moisture products using global ground-based in situ observations. *Remote Sensing of Environment*, *118*, 215–226. https://doi.org/10.1016/j.rse.2011.11.017

Albergel, C., Rüdiger, C., Pellarin, T., Calvet, J.-C., Fritz, N., Froissard, F., et al. (2008). From near-surface to root-zone soil moisture using an exponential filter: An assessment of the method based on in-situ observations and model simulations. *Hydrology and Earth System Sciences*, *12*, 1323–1337. https://doi.org/10.5194/hess-12-1323-2008

Alexakis, D., Mexis, F.-D., Vozinaki, A.-E., Daliakopoulos, I., & Tsanis, I. (2017). Soil moisture content estimation based on Sentinel-1 and auxiliary earth observation products. A hydrological approach. *Sensors*, *17*, 1455. https://doi.org/10.3390/s17061455

Archer, J. R., & Smith, P. D. (1972). The relation between bulk density, available water capacity, and air capacity of soils. *Journal of Soil Science*, *23*, 475–480. https://doi.org/10.1111/j.1365-2389.1972.tb01678.x

AREA, U., & UINTAH, P. O. F. (2010). United States Department of Agriculture Natural Resources Conservation Service.

Auddino, M., Dominici, R., & Viscomi, A. (2015). Evaluation of yield sediment in the Sfalassà Fiumara (south-western, Calabria) by using Gavrilović method in GIS enviroment. *Rendiconti Online Della Società Geologica Italiana*, *33*, 3–7.

Baetens, J. M., Verbist, K., Cornelis, W. M., Gabriels, D., & Soto, G. (2009). On the influence of coarse fragments on soil water retention. *Water Resources Research*, *45*, W07408. https://doi.org/10.1029/2008WR007402

Bai, P., Liu, X., Yang, T., Liang, K., & Liu, C. (2016). Evaluation of streamflow simulation results of land surface models in GLDAS on the Tibetan plateau. *Journal of Geophysical Research: Atmospheres*, *121*, 12180–12197. https://doi.org/10.1002/2016JD025501

Baldocchi, D., Falge, E., Gu, L., Olson, R., Hollinger, D., Running, S., et al. (2001). FLUXNET: A new tool to study the temporal and spatial variability of ecosystem-scale carbon dioxide, water vapor, and energy flux densities. *Bulletin of the American Meteorological Society*, *82*, 2415–2434. https://doi.org/10.1175/1520-0477(2001)082<2415:FANTTS>2.3.CO;2

Bauer-Marschallinger, B., Paulik, C., Hochstöger, S., Mistelbauer, T., Modanesi, S., Ciabatta, L., et al. (2018). Soil moisture from fusion of scatterometer and SAR: Closing the scale gap with temporal filtering. *Remote Sensing*, *10*, 1030. https://doi.org/10.3390/rs10071030

Beyrich, F., & Adam, W. K. (2007). *Site and data report for the Lindenberg reference site in CEOP-Phase I*. Offenbach am Main, Germany: Selbstverlag des Deutschen Wetterdienstes.

Bircher, S., Skou, N., Jensen, K. H., Walker, J. P., Rasmussen, L., & Verhoest, N. (2012). A soil moisture and temperature network for SMOS validation in Western Denmark. *Hydrology and Earth System Sciences*, *16*, 1445–1463.

Bischiniotis, K., van den Hurk, B., Zsoter, E., Coughlan de Perez, E., Grillakis, M., & Aerts, J. C. J. H. (2019). Evaluation of a global ensemble flood prediction system in Peru. *Hydrological Sciences Journal*, *64*, 1171–1189. https://doi.org/10.1080/02626667.2019.1617868

Brocca, L., Ciabatta, L., Massari, C., Camici, S., & Tarpanelli, A. (2017). Soil moisture for hydrological applications: Open questions and new opportunities. *Water*, *9*, 140. https://doi.org/10.3390/w9020140

Brocca, L., Hasenauer, S., Lacava, T., Melone, F., Moramarco, T., Wagner, W., et al. (2011). Soil moisture estimation through ASCAT and AMSR-E sensors: An intercomparison and validation study across Europe. *Remote Sensing of Environment*, *115*, 3390–3408. https://doi.org/10.1016/j.rse.2011.08.003

Brocca, L., Melone, F., & Moramarco, T. (2008). On the estimation of antecedent wetness conditions in rainfall-runoff modelling. *Hydrological Processes*, *22*, 629–642. https://doi.org/10.1002/hyp.6629

Brocca, L., Melone, F., Moramarco, T., & Morbidelli, R. (2009). Antecedent wetness conditions based on ERS scatterometer data. *Journal of Hydrology*, *364*, 73–87. https://doi.org/10.1016/j.jhydrol.2008.10.007

Brocca, L., Melone, F., Moramarco, T., Wagner, W., Naeimi, V., Bartalis, Z., & Hasenauer, S. (2010). Improving runoff prediction through the assimilation of the ASCAT soil moisture product. *Hydrology and Earth System Sciences*, *14*, 1881–1893. https://doi.org/10.5194/hess-14-1881-2010

Brocca, L., Moramarco, T., Melone, F., Wagner, W., Hasenauer, S., & Hahn, S. (2012). Assimilation of surface- and root-zone ASCAT soil moisture products into rainfall–runoff modeling. *IEEE Transactions on Geoscience and Remote Sensing*, *50*, 2542–2555. https://doi.org/10.1109/TGRS.2011.2177468

Bussotti, F., Ferrini, F., Pollastrini, M., & Fini, A. (2014). The challenge of Mediterranean sclerophyllous vegetation under climate change: From acclimation to adaptation. *Environmental and Experimental Botany*, *103*, 80–98. https://doi.org/10.1016/j.envexpbot.2013.09.013

Calvet, J. C., Fritz, N., Froissard, F., Suquia, D., Petitpa, A., & Piguet, B. (2007). In situ soil moisture observations for the CAL/VAL of SMOS: The SMOSMANIA network. In *International Geoscience and Remote Sensing Symposium (IGARSS)* (pp. 1196–1199). https://doi.org/10.1109/IGARSS.2007.4423019

Cappelaere, B., Descroix, L., Lebel, T., Boulain, N., Ramier, D., Laurent, J.-P., et al. (2009). The AMMA-CATCH experiment in the cultivated Sahelian area of south-west Niger—Investigating water cycle response to a fluctuating climate and changing environment. *Journal of Hydrology*, *375*, 34–51. https://doi.org/10.1016/j.jhydrol.2009.06.021

Ceballos, A., Scipal, K., Wagner, W., & Martínez-Fernández, J. (2005). Validation of ERS scatterometer-derived soil moisture data in the central part of the Duero Basin, Spain. *Hydrological Processes*, *19*, 1549–1566. https://doi.org/10.1002/hyp.5585

Chandler, D. G., Seyfried, M. S., McNamara, J. P., & Hwang, K. (2017). Inference of soil hydrologic parameters from electronic soil moisture records. *Frontiers of Earth Science*, *5*, 25. https://doi.org/10.3389/feart.2017.00025

Chen, M., Willgoose, G. R., & Saco, P. M. (2014). Spatial prediction of temporal soil moisture dynamics using HYDRUS-1D. *Hydrological Processes*, *28*, 171–185. https://doi.org/10.1002/hyp.9518

Copernicus. (2019). *ERA5-Land monthly averaged data from 1981 to present dataset*. https://doi.org/10.24381/cds.68d2bb30

Cosh, M. H., Jackson, T. J., Starks, P., & Heathman, G. (2006). Temporal stability of surface soil moisture in the Little Washita River watershed and its applications in satellite soil moisture product validation. *Journal of Hydrology*, *323*, 168–177. https://doi.org/10.1016/j.jhydrol.2005.08.020

Crow, W. T., Berg, A. A., Cosh, M. H., Loew, A., Mohanty, B. P., Panciera, R., et al. (2012). Upscaling sparse ground-based soil moisture observations for the validation of coarse-resolution satellite soil moisture products. *Reviews of Geophysics*, *50*, RG2002. https://doi.org/10.1029/2011RG000372

DAAC, O. (2013). *Fluxnet web page*. Oak Ridge, TN:Oak Ridge National Laboratory Distributed Active Archive Center.

Dirmeyer, P. A., Gao, X., Zhao, M., Guo, Z., Oki, T., & Hanasaki, N. (2006). GSWP-2: Multimodel analysis and implications for our perception of the land surface. *Bulletin of the American Meteorological Society*, *87*, 1381–1398. https://doi.org/10.1175/BAMS-87-10-1381

Dorigo, W., Himmelbauer, I., Aberer, D., Schremmer, L., Petrakovic, I., Zappa, L., et al. (2021). The International Soil Moisture Network: Serving Earth system science for over a decade. *Hydrology and Earth System Sciences Discussions*, *2021*, 1–83. https://doi.org/10.5194/hess-2021-2

Dorigo, W., Wagner, W., Albergel, C., Albrecht, F., Balsamo, G., Brocca, L., et al. (2017). ESA CCI Soil Moisture for improved Earth system understanding: State-of-the art and future directions. *Remote Sensing of Environment*, *203*, 185–215.

Dorigo, W. A., Wagner, W., Hohensinn, R., Hahn, S., Paulik, C., Xaver, A., et al. (2011). The International Soil Moisture Network: A data hosting facility for global in situ soil moisture measurements. *Hydrology and Earth System Sciences*, *15*, 1675–1698. https://doi.org/10.5194/hess-15-1675-2011

Dorigo, W. A., Xaver, A., Vreugdenhil, M., Gruber, A., Hegyiová, A., Sanchis-Dufau, A. D., et al. (2013). Global automated quality control of in situ soil moisture data from the International Soil Moisture Network. *Vadose Zone Journal*, *12*, 1–21. https://doi.org/10.2136/vzj2012.0097

ECMWF. (2020). ECMWF [WWW document]. Retrieved from https://confluence.ecmwf.int/rest/ecmwfjsd/1.0/exporthtml/pagedisplay/CKB/ERA5-Land%3A+data+documentation (accessed November 11, 2020).

Famiglietti, J. S., Ryu, D., Berg, A. A., Rodell, M., & Jackson, T. J. (2008). Field observations of soil moisture variability across scales. *Water Resources Research*, *44*, W01423. https://doi.org/10.1029/2006WR005804

Giorgi, F., & Bi, X. (2005). Updated regional precipitation and temperature changes for the 21st century from ensembles of recent AOGCM simulations. *Geophysical Research Letters*, *32*, L21715. https://doi.org/10.1029/2005GL024288

Grillakis, M. G. (2019). Increase in severe and extreme soil moisture droughts for Europe under climate change. *The Science of the Total Environment*, *660*, 1245–1255. https://doi.org/10.1016/j.scitotenv.2019.01.001

Grillakis, M. G., Koutroulis, A. G., Komma, J., Tsanis, I. K., Wagner, W., & Blöschl, G. (2016). Initial soil moisture effects on flash flood generation—A comparison between basins of contrasting hydro-climatic conditions. *Journal of Hydrology*, *541*, 206–217. https://doi.org/10.1016/j.jhydrol.2016.03.007

Gruber, A., Dorigo, W. A., Crow, W., & Wagner, W. (2017). Triple collocation-based merging of satellite soil moisture retrievals. *IEEE Transactions on Geoscience and Remote Sensing*, *55*, 6780–6792. https://doi.org/10.1109/TGRS.2017.2734070

Gruber, A., Scanlon, T., van der Schalie, R., Wagner, W., & Dorigo, W. (2019). Evolution of the ESA CCI Soil Moisture climate data records and their underlying merging methodology. *Earth System Science Data*, *11*, 717–739. https://doi.org/10.5194/essd-11-717-2019

Gruhier, C., de Rosnay, P., Hasenauer, S., Holmes, T., de Jeu, R., Kerr, Y., et al. (2010). Soil moisture active and passive microwave products: Intercomparison and evaluation over a Sahelian site. *Hydrology and Earth System Sciences*, *14*, 141–156. https://doi.org/10.5194/hess-14-141-2010

Guo, X., Fu, Q., Hang, Y., Lu, H., Gao, F., & Si, J. (2020). Spatial variability of soil moisture in relation to land use types and topographic features on hillslopes in the black soil (mollisols) area of northeast China. *Sustainability*, *12*, 3552. https://doi.org/10.3390/SU12093552

Hengl, T., de Jesus, J. M., MacMillan, R. A., Batjes, N. H., Heuvelink, G. B. M., Ribeiro, E., et al. (2014). SoilGrids1km—Global soil information based on automated mapping. *PLoS One*, *9*, e105992. https://doi.org/10.1371/journal.pone.0105992

Hengl, T., Mendes de Jesus, J., Heuvelink, G. B. M., Ruiperez Gonzalez, M., Kilibarda, M., Blagotić, A., et al. (2017). SoilGrids250m: Global gridded soil information based on machine learning. *PLoS One*, *12*, e0169748. https://doi.org/10.1371/journal.pone.0169748

Hollinger, S. E., & Isard, S. A. (1994). A soil moisture climatology of Illinois. *Journal of Climate*, *7*, 822–833.

Hudson Beale, M., Hagan, M., & Demuth, H. (2020). *Deep learning toolbox (TM) reference, MATLAB*. 1 Apple Hill Drive Natick, MA 01760-2098.

Jackson, T. J., Cosh, M. H., Bindlish, R., Starks, P. J., Bosch, D. D., Seyfried, M., et al. (2010). Validation of advanced microwave scanning radiometer soil moisture products. *IEEE Transactions on Geoscience and Remote Sensing*, *48*, 4256–4272. https://doi.org/10.1109/TGRS.2010.2051035

Kinama, J. M., Stigter, C. J., Ong, C. K., Ng'ang'a, J. K., & Gichuki, F. N. (2010). *Comparing soil moisture by TDR and neutron probe at 30 cm depth and crop yields in a maize and cowpea agroforestry and grass strip farming system semi-arid Kenya*.

Korres, W., Reichenau, T. G., Fiener, P., Koyama, C. N., Bogena, H. R., Cornelissen, T., et al. (2015). Spatio-temporal soil moisture patterns—A meta-analysis using plot to catchment scale data. *Journal of Hydrology*, *520*, 326–341. https://doi.org/10.1016/j.jhydrol.2014.11.042

Kumar, A., Ramsankaran, R., Brocca, L., & Munoz-Arriola, F. (2019). A machine learning approach for improving near-real-time satellite-based rainfall estimates by integrating soil moisture. *Remote Sensing*, *11*, 2221. https://doi.org/10.3390/rs11192221

Lagarias, J. C., Reeds, J. A., Wright, M. H., & Wright, P. E. (1998). Convergence properties of the Nelder–Mead simplex method in low dimensions. *SIAM Journal on Optimization*, *9*, 112–147. https://doi.org/10.1137/S1052623496303470

Leavesley, G. H., David, O., Garen, D. C., Lea, J., Marron, J. K., Pagano, T. C., et al. (2008). *A modeling framework for improved agricultural water supply forecasting*. AGUFM 2008, C21A-0497.

Lettenmaier, D. P., Alsdorf, D., Dozier, J., Huffman, G. J., Pan, M., & Wood, E. F. (2015). Inroads of remote sensing into hydrologic science during the WRR era. *Water Resources Research*, *51*, 7309–7342. https://doi.org/10.1002/2015WR017616

Liu, Y. Y., Parinussa, R. M., Dorigo, W. A., De Jeu, R. A. M., Wagner, W., van Dijk, A. I. J. M., et al. (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

Manfreda, S., Brocca, L., Moramarco, T., Melone, F., & Sheffield, J. (2014). A physically based approach for the estimation of root-zone soil moisture from surface measurements. *Hydrology and Earth System Sciences*, *18*, 1199–1212. https://doi.org/10.5194/hess-18-1199-2014

Mao, H., Kathuria, D., Duffield, N., & Mohanty, B. P. (2019). Gap filling of high-resolution soil moisture for SMAP/Sentinel-1: A two-layer machine learning-based framework. *Water Resources Research*, *55*, 6986–7009. https://doi.org/10.1029/2019WR024902

Marschallinger, B., & Paulik, C. (2019). *Copernicus global land operations "vegetation and energy" "CGLOPS-1" algorithm theoretical basis document—Soil Water Index collection 1 km—Version 1—Issue I1.20*.

Mattar, C., Santamaría-Artigas, A., Durán-Alarcón, C., Olivera-Guerra, L., & Fuster, R. (2014). LAB-net the first Chilean soil moisture network for remote sensing applications. In *Proceedings IV Recent Advances in Quantitative Remote Sensing* (pp. 22–26).

McNally, A. (2018). *FLDAS Noah Land Surface Model L4 global monthly 0.1 × 0.1 degree (MERRA-2 and CHIRPS)*. https://doi.org/10.5067/5NHC22T9375G

McNally, A., Arsenault, K., Kumar, S., Shukla, S., Peterson, P., Wang, S., et al. (2017). A land data assimilation system for sub-Saharan Africa food and water security applications. *Scientific Data*, *4*, 1–19. https://doi.org/10.1038/sdata.2017.12

Mougin, E., Hiernaux, P., Kergoat, L., Grippa, M., de Rosnay, P., Timouk, F., et al. (2009). The AMMA-CATCH Gourma observatory site in Mali: Relating climatic variations to changes in vegetation, surface hydrology, fluxes and natural resources. *Journal of Hydrology*, *375*, 14–33. https://doi.org/10.1016/j.jhydrol.2009.06.045

Nemes, A., Rawls, W. J., & Pachepsky, Y. A. (2005). Influence of organic matter on the estimation of saturated hydraulic conductivity. *Soil Science Society of America Journal*, *69*, 1330–1337. https://doi.org/10.2136/sssaj2004.0055

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. *Journal of Geophysical Research: Atmospheres*, *120*, 9955–9964. https://doi.org/10.1002/2015JD023305

O'Geen, A. T. (2013). *Soil water dynamics [WWW document]*. Nature Education Knowledge. Retrieved from https://www.nature.com/scitable/knowledge/library/soil-water-dynamics-103089121/ (accessed October 12, 2020).

Orgiazzi, A., Ballabio, C., Panagos, P., Jones, A., & Fernández-Ugalde, O. (2018). LUCAS soil, the largest expandable soil dataset for Europe: A review. *European Journal of Soil Science*, *69*, 140–153. https://doi.org/10.1111/ejss.12499

Osenga, E. C., Arnott, J. C., Endsley, K. A., & Katzenberger, J. W. (2019). Bioclimatic and soil moisture monitoring across elevation in a mountain watershed: Opportunities for research and resource management. *Water Resources Research*, *55*, 2493–2503. https://doi.org/10.1029/2018WR023653

Parajka, E. C., Naeimi, J. C., Blöschl, K. A., & Komma, J. W. (2009). Matching ERS scatterometer based soil moisture patterns with simulations of a conceptual dual layer hydrologic model over Austria. *Hydrology and Earth System Sciences*, *13*, 259–271. https://doi.org/10.5194/hess-13-259-2009

Parrens, M., Zakharova, E., Lafont, S., Calvet, J.-C., Kerr, Y., Wagner, W., & Wigneron, J.-P. (2012). Comparing soil moisture retrievals from SMOS and ASCAT over France. *Hydrology and Earth System Sciences*, *16*, 423–440. https://doi.org/10.5194/hess-16-423-2012

Paulik, C., Dorigo, W., Wagner, W., & Kidd, R. (2014). Validation of the ASCAT Soil Water Index using in situ data from the International Soil Moisture Network. *International Journal of Applied Earth Observation and Geoinformation*, *30*, 1–8. https://doi.org/10.1016/j.jag.2014.01.007

Pellarin, T., Laurent, J. P., Cappelaere, B., Decharme, B., Descroix, L., & Ramier, D. (2009). Hydrological modelling and associated microwave emission of a semi-arid region in South-western Niger. *Journal of Hydrology*, *375*, 262–272. https://doi.org/10.1016/j.jhydrol.2008.12.003

Pokhrel, Y., Felfelani, J. P., Satoh, Y., Boulange, J., Burek, P., Gädeke, A., et al. (2021). Global terrestrial water storage and drought severity under climate change. *Nature Climate Change*, *11*, 226–233. https://doi.org/10.1038/s41558-020-00972-w

Portmann, F. T., Siebert, S., & Döll, P. (2010). MIRCA2000—Global monthly irrigated and rainfed crop areas around the year 2000: A new high-resolution data set for agricultural and hydrological modeling. *Global Biogeochemical Cycles*, *24*, GB1011. https://doi.org/10.1029/2008GB003435

Rautiainen, K., Lemmetyinen, J., Pulliainen, J., Vehvilainen, J., Drusch, M., Kontu, A., et al. (2012). L-band radiometer observations of soil processes in boreal and subarctic environments. *IEEE Transactions on Geoscience and Remote Sensing*, *50*, 1483–1497. https://doi.org/10.1109/TGRS.2011.2167755

Raza, A., Friedel, J. K., Moghaddam, A., Ardakani, M. R., Loiskandl, W., Himmelbauer, M., & Bodner, G. (2013). Modeling growth of different Lucerne cultivars and their effect on soil water dynamics. *Agricultural Water Management*, *119*, 100–110. https://doi.org/10.1016/j.agwat.2012.12.006

Rüdiger, C., Hancock, J. K., Hemakumara, H. M., Jacobs, M. R., Kalma, J. D., Martinez, C., et al. (2007). Goulburn River experimental catchment data set. *Water Resources Research*, *43*, W10403. https://doi.org/10.1029/2006WR005837

Seneviratne, S. I., Corti, T., Davin, E. L., Hirschi, M., Jaeger, E. B., Lehner, I., et al. (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

Smets, B., Jacobs, T., Verger, E. L. (2018). *Gio Global Land Component—Lot I "operation of the global land component" product user manual*.

Srivastava, P. K., Han, D., Ramirez, M. R., & Islam, T. (2013). Machine learning techniques for downscaling SMOS satellite soil moisture using MODIS land surface temperature for hydrological application. *Water Resources Management*, *27*, 3127–3144. https://doi.org/10.1007/s11269-013-0337-9

Stroud, P. D. (1999). *A recursive exponential filter for time-sensitive data* (Tech. Rep. LAUR-99-5573 131). Washington, DC: Los Alamos National Laboratory.

Tramblay, Y., Koutroulis, A., Samaniego, L., Vicente-Serrano, S. M., Volaire, F., Boone, A., et al. (2020). Challenges for drought assessment in the Mediterranean region under future climate scenarios. *Earth-Science Reviews*, *210*, 103348. https://doi.org/10.1016/j.earscirev.2020.103348

Van Cleve, K., Chapin, F. S. S., Ruess, R. W. (2015). Bonanza Creek Long Term Ecological Research Project Climate Database.

Vinnikov, K. Y., Robock, A., Qiu, S., & Entin, J. K. (1999). Optimal design of surface networks for observation of soil moisture. *Journal of Geophysical Research*, *104*, 19743–19749.

Wagner, W., Dorigo, W., De Jeu, R., Fernandez, J. K., Benveniste, J., Haas, E., & Ertl, M. (2012). Fusion of active and passive microwave observations to create an essential climate variable data record on soil moisture. In *ISPRS annals of the photogrammetry, remote sensing and spatial information sciences* (pp. 315–321). Göttingen, Germany: Copernicus GmbH. https://doi.org/10.5194/isprsannals-I-7-315-2012

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

Weedon, G., Balsamo, G., Bellouin, N., Gomes, S., Best, M., & Viterbo, P. (2018). *The WFDEI Meteorological Forcing Data. Research Data Archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory [WWW Document]*. https://doi.org/10.5065/486N-8109

Yang, K., Qin, J., Zhao, L., Chen, Y., Tang, W., Han, M., et al. (2013). A multiscale soil moisture and freeze–thaw monitoring network on the third pole. *Bulletin of the American Meteorological Society*, *94*, 1907–1916.

Young, R., Walker, J., Yeoh, N., Smith, A., Ellett, K., Merlin, O., et al. (2008). *Soil moisture and meteorological observations from the Murrumbidgee catchment*. Melbourne, Australia: Department of Civil and Environmental Engineering, University of Melbourne.

Zreda, M., Shuttleworth, W. J., Zeng, X., Zweck, C., Desilets, D., Franz, T., et al. (2012). COSMOS: The COsmic-ray Soil Moisture Observing System. *Hydrology and Earth System Sciences Discussions*, *9*, 4505–4551. https://doi.org/10.5194/hessd-9-4505-2012