



Statistical Exploration of SENTINEL-1 Data, Terrain Parameters, and *in-situ* Data for Estimating the Near-Surface Soil Moisture in a Mediterranean Agroecosystem

Sarah Schönbrodt-Stitt^{1*}, Nima Ahmadian¹, Markus Kurtenbach¹, Christopher Conrad², Nunzio Romano³, Heye R. Bogena⁴, Harry Vereecken⁴ and Paolo Nasta³

¹ Department of Remote Sensing, Institute of Geography and Geology, University of Würzburg, Würzburg, Germany,

² Department of Geoecology, Institute of Geosciences and Geography, University of Halle-Wittenberg, Halle/Saale, Germany,

³ Department of Agricultural Sciences, Division of Agricultural, Forest and Biosystems Engineering, University of Naples

Federico II, Naples, Italy, ⁴ Agrosphere Institute (Institute of Bio- and Geosciences), Forschungszentrum Jülich, Jülich, Germany

OPEN ACCESS

Edited by:

Luca Brocca,
National Research Council (CNR), Italy

Reviewed by:

Angelica Tarpanelli,
National Research Council (CNR), Italy
Rogier Van Der Velde,
University of Twente, Netherlands

*Correspondence:

Sarah Schönbrodt-Stitt
sarah.schoenbrodt-stitt@
uni-wuerzburg.de

Specialty section:

This article was submitted to
Water and Hydrocomplexity,
a section of the journal
Frontiers in Water

Received: 19 January 2021

Accepted: 18 June 2021

Published: 14 July 2021

Citation:

Schönbrodt-Stitt S, Ahmadian N,
Kurtenbach M, Conrad C, Romano N,
Bogena HR, Vereecken H and
Nasta P (2021) Statistical Exploration
of SENTINEL-1 Data, Terrain
Parameters, and *in-situ* Data for
Estimating the Near-Surface Soil
Moisture in a Mediterranean
Agroecosystem.
Front. Water 3:655837.
doi: 10.3389/frwa.2021.655837

Reliable near-surface soil moisture (θ) information is crucial for supporting risk assessment of future water usage, particularly considering the vulnerability of agroforestry systems of Mediterranean environments to climate change. We propose a simple empirical model by integrating dual-polarimetric Sentinel-1 (S1) Synthetic Aperture Radar (SAR) C-band single-look complex data and topographic information together with *in-situ* measurements of θ into a random forest (RF) regression approach (10-fold cross-validation). Firstly, we compare two RF models' estimation performances using either 43 SAR parameters ($\theta_{\text{Nov}}^{\text{SAR}}$) or the combination of 43 SAR and 10 terrain parameters ($\theta_{\text{Nov}}^{\text{SAR+Terrain}}$). Secondly, we analyze the essential parameters in estimating and mapping θ for S1 overpasses twice a day (at 5 a.m. and 5 p.m.) in a high spatiotemporal (17 × 17 m; 6 days) resolution. The developed site-specific calibration-dependent model was tested for a short period in November 2018 in a field-scale agroforestry environment belonging to the "Alento" hydrological observatory in southern Italy. Our results show that the combined SAR + terrain model slightly outperforms the SAR-based model ($\theta_{\text{Nov}}^{\text{SAR+Terrain}}$ with 0.025 and 0.020 m³ m⁻³, and 89% compared to $\theta_{\text{Nov}}^{\text{SAR}}$ with 0.028 and 0.022 m³ m⁻³, and 86% in terms of RMSE, MAE, and R²). The higher explanatory power for $\theta_{\text{Nov}}^{\text{SAR+Terrain}}$ is assessed with time-variant SAR phase information-dependent elements of the C2 covariance and Kennaugh matrix (i.e., K1, K6, and K1S) and with local (e.g., altitude above channel network) and compound topographic attributes (e.g., wetness index). Our proposed methodological approach constitutes a simple empirical model aiming at estimating θ for rapid surveys with high accuracy. It emphasizes potentials for further improvement (e.g., higher spatiotemporal coverage of ground-truthing) by identifying differences of SAR measurements between S1 overpasses in the morning and afternoon.

Keywords: near-surface soil moisture, Sentinel-1 single-look complex data, SAR backscatters, terrain parameters, Alento hydrological observatory, Mediterranean environment

INTRODUCTION

Globally, experts from diverse applications such as climate, hydrological, and crop yield modeling agree on the importance of the near-surface soil moisture (θ) for improving risk assessment and land and water resources management (Babaeian et al., 2019). It is mainly due to θ acting as a state variable for controlling land-atmosphere feedback and thereby controlling ecosystem dynamics (e.g., Seneviratne et al., 2010; Dorigo et al., 2017).

At the plot to the field-scale, the direct thermo-gravimetric method provides reliable *in-situ* measurements of θ . Alternatively, indirect methods through the use of portable devices or stationary sensors (e.g., Time Domain Reflectometry) measure the soil dielectric properties from which the soil moisture is inferred through semi-empirical equations (e.g., Rowlandson et al., 2018; Babaeian et al., 2019). However, their applications are rather unfeasible for large-scale operations due to excessive demand for time (e.g., maintenance) and costs (e.g., labor) for surveying (e.g., Santi et al., 2016). Moreover, the significance of ground-based methods is limited to the represented plot settings, and the risk of generating non-representative patterns of θ by spatial regionalization prevails (Chen et al., 2017).

Spaceborne remote sensing (RS) provides spatially explicit information as satellites sense the same ground trace in regular time intervals, allowing for continuous monitoring (Babaeian et al., 2019). Estimates of θ are retrieved from different sensors measuring optical and thermal spectra (e.g., Rahimzadeh-Bajgirani et al., 2013; Zhang and Zhou, 2016), by passive and active microwave sensors (e.g., Schmugge and Jackson, 1997; Das and Paul, 2015), or the synergistic use of different sensor types such as using radar and optical data from Sentinel-2, Landsat, and MODIS (e.g., Attarzadeh et al., 2018; Ayehu et al., 2020; Foucras et al., 2020; Han et al., 2020; Ma et al., 2020). Synthetic Aperture Radars (SAR) are among the most effective and flexible active microwave sensor systems (e.g., Wang and Qu, 2009; Santi et al., 2016) due to their ability to penetrate the near-surface soil layer up to a depth of 5 cm (i.e., for C-band), which in turn enables to observe θ by directly relating the microwave scattering and emission to the water content of the focused object (e.g., Paloscia et al., 2013; Santi et al., 2016; Mohanty et al., 2017; Babaeian et al., 2019).

High spatiotemporal SAR-based retrieval of θ predominantly focusses on bare land surfaces (e.g., Datta et al., 2020), sparsely vegetated areas such as grassland and meadows (e.g., Xu et al., 2020), or agricultural farmland when the vegetation cover is low such as covered by crop residues (e.g., Ayehu et al., 2020). This is mainly due to the unique challenge resulting from the SAR backscatter's high sensitivity to surface characteristics (e.g., roughness) and vegetation properties (e.g., vegetation canopy architecture). From a physical perspective, this is an inherent source of uncertainty in deriving θ (Alemohammad et al., 2019), which can be widely disentangled using backscatter models of different complexity. Numerical theoretical models such as the Integral Equation Model (Baghdadi et al., 2006) and the Kirchhoff Approximation (Gu et al., 2019), and semi-empirical

backscatter models such as the Oh model (Oh et al., 1992) and the Water Cloud Model (e.g., Baghdadi et al., 2017) relate the sensor-specific microwave scattering to the target at ground (e.g., bare soil) by means of physical-based descriptions (e.g., Paloscia et al., 2013). They intend to be not site-specific and universally applicable (Paloscia et al., 2013). However, these physical surface scattering models are not valid for vegetation surfaces, hence should not be applied for conditions as present in dense and woody vegetation cover areas (Petropoulos et al., 2015; MirMazloumi and Sahebi, 2016).

The models above mainly require complex radiative transfer algorithmic handling to solve the physical equation systems (e.g., Das and Paul, 2015), thus potentially reducing feasibility for practitioners as stated by Ezzahar et al. (2019). Empirical backscatter models present a simplification by investigating the interaction of microwaves with site-specific surface characteristics but are limited to the area under investigation and require the input of high-quality reference data, e.g., surface roughness (Barrett et al., 2009).

Data-driven machine learning (ML) approaches were investigated for the operational and systematic mapping and monitoring of θ (e.g., Santi et al., 2016; Adeyemi et al., 2018; Cai et al., 2019; Efremova et al., 2019; Datta et al., 2020). Provided the availability of data sufficiently characterizing the system for which θ is to be estimated, they present an adequate alternative (Ezzahar et al., 2019) as they expand the concept of the above backscatter models and allow integration of additional environmental covariates (parameters) beyond those defined for physical modeling.

ML approaches effectively handle large spatial datasets and break down complexity by extracting representative parameters (Karpatne et al., 2019). They can account for the non-linearity of relationships, which is mainly present in RS-based θ retrieval in vegetated and heterogeneous areas (e.g., Chakrabarti et al., 2015). Random forest regression and support vector machine applied on pixels (Hajdu et al., 2018; Datta et al., 2020) or image objects (Ezzahar et al., 2019) and *K*-Nearest Neighbors for non-parametric regression analysis (Datta et al., 2020) are reported to lead to predictions of increasing accuracy, with even reported less required data compared to mechanistic models (Adeyemi et al., 2018), and enhanced readability of results (Attarzadeh et al., 2018; Ezzahar et al., 2019). ML approaches are increasingly used to estimate θ in agricultural watersheds across humid to arid climate regimes and at different scales (e.g., Quesney et al., 2000; Hachani et al., 2019). They are reported to improve hydrological modeling in wetlands (Dabrowska-Zielinska et al., 2018) and to support the monitoring of water tables dynamics (Asmuß et al., 2019), and the estimation of crop water stress (El-Shirbeny and Abutaleb, 2017), and the water use efficiency (Efremova et al., 2019).

Commonly, studies employ the backscatter intensity Sigma-naught (σ°) either in single (VV or VH) polarization (e.g., Quesney et al., 2000; Attarzadeh et al., 2018; Xu et al., 2020) or dual (VV + VH) polarization (e.g., Dabrowska-Zielinska et al., 2018; Efremova et al., 2019; Datta et al., 2020) from SAR amplitude information only. The topographic effect on radar backscatter, i.e., foreshortening, layover, and shadowing

(Chen and Wang, 2008), is considered in the preprocessing of SAR data such as by terrain flattening and orthorectification to remove radiometric distortions (Small, 2011) and to correct distortions induced by the sensor inclination and topography, respectively (Filippini, 2019). Less attention, however, is paid to the integration of relevant topographic parameters to retrieve spatial patterns of θ in ML approaches.

Contador et al. (2006) applied Artificial Neural Networks for modeling θ based on topography in a Mediterranean watershed in Spain. They stress the strength of topographic variables in predicting area-wide θ as they contribute to explaining spatial information spanning from the very vicinity of a point (e.g., aspect and slope) reflecting the local position to the entire terrain. Notably, compound, topographic attributes such as the contributing area, the flow accumulation, the wetness index, and the topographic roughness allow for integrating the spatial coherence and geographical organization between the different terrain positions (Contador et al., 2006). Subsequently, terrain features enable approximating of spatial patterns θ from an eco-hydrological perspective (e.g., Western et al., 2002; Robinson et al., 2008).

Guevara and Vargas (2019) successfully employed topographic information in an ML approach to downscale RS-based soil moisture products over the conterminous USA and labeled several parameters as “hydrologically meaningful.” In this way, it was possible to adequately address the landscape’s capability to physically constrain water inputs (e.g., rain and irrigation water, overland flow) and support the link between spatial variability of soil moisture θ and topography (Guevara and Vargas, 2019).

While the number of studies on Sentinel-1 C-band SAR-based estimation of θ , either by using physical backscatter models or ML approaches, is steadily increasing, and the possibilities for a more accurate analysis under vegetation (e.g., cropland) are explored, application examples from agroforestry sites are not known. This is noteworthy as those sites hold a multifunctional role for a sustainable development model for rural areas (Santoro et al., 2020). Particularly across European Mediterranean regions that is home to one of the global hotspots of agroforestry sites (Gauquelin et al., 2018), they should provide water supply for multiple uses (e.g., domestic water use, irrigation, ecosystem functions) during the dry season (e.g., García-Ruiz et al., 2011). Therefore, θ is crucial for creating scenario-based projections for supporting specific risk assessments (e.g., Nasta et al., 2020).

Our study aims to estimate and map a reasonable short time series of θ in a Mediterranean dry sub-humid, field-scale agroforestry ecosystem in southern Italy using an ML approach. Our objective is to develop a site-specific simplified empirical calibration model that provides the capability to handle SAR backscatters and scarce *in-situ* information of topsoil θ for a short observation period in November 2018. For calibrating θ from SAR measurements, we aim to employ dual-polarization phase and amplitude information from Sentinel-1 SAR single-look complex (SLC) data to retrieve additional parameters beyond the recent scope of physical and semi-empirical models. To account for the effect of the topography on spatial patterns of θ , we further test the integration of a set of

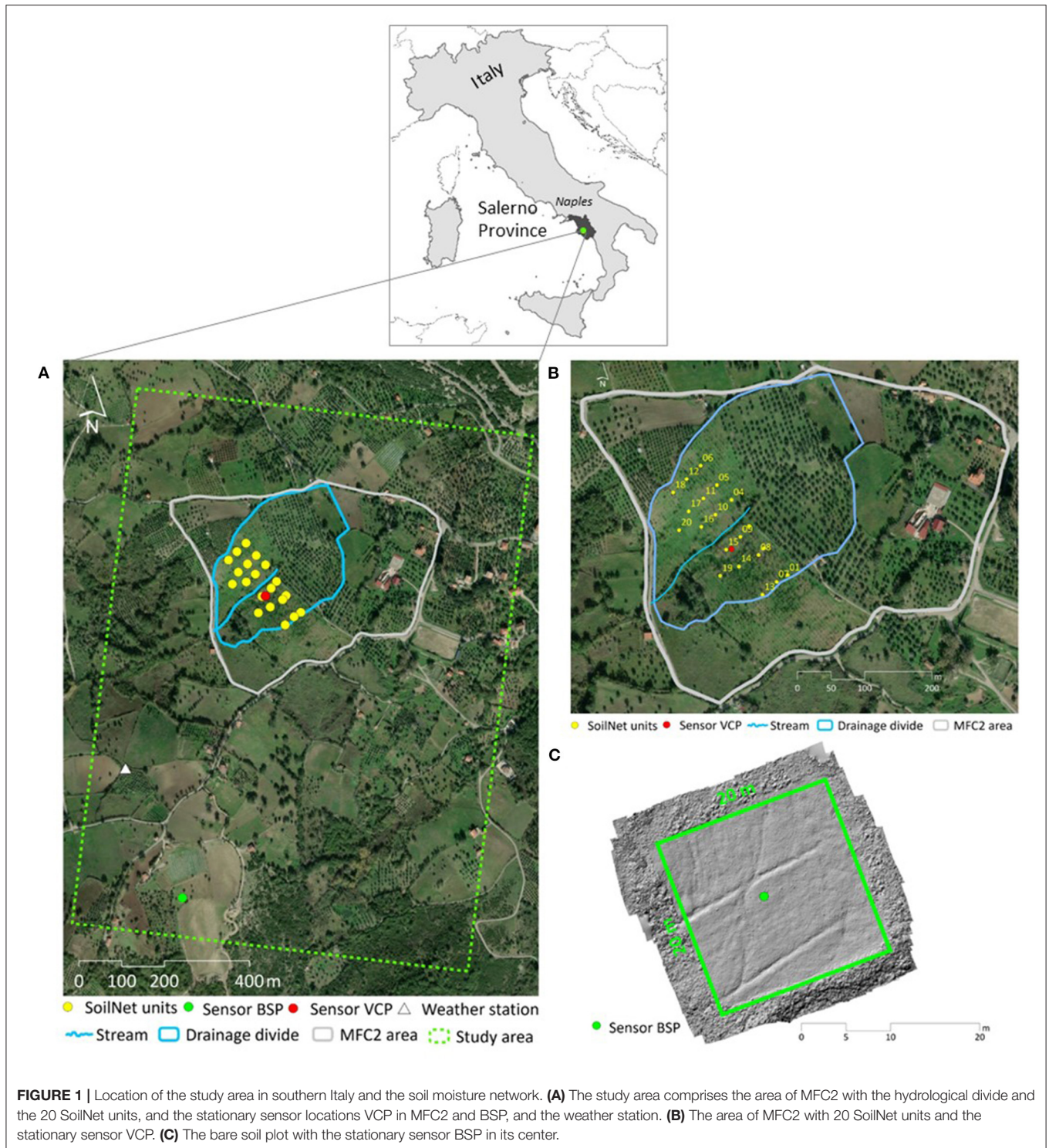
so-called ‘hydrological meaningful’ parameters. For evaluating its estimation performance, we compare the *in-situ* θ with estimated θ information when building the model (i) with SAR-based parameters only and (ii) with the combination of SAR and terrain parameters. Furthermore, we aim at identifying the most relevant prediction parameters and their temporal behavior (e.g., variability) to enhance understanding of their sensitivity.

STUDY AREA

Our study focuses on a field-scale, predominantly agroforestry area in the headwater zone of the Upper Alento River Catchment in southern Italy (Figure 1; Nasta et al., 2017). It is represented by the rectangular tile (green dashed line) with a size of 1.18 km² in Figure 1A, which comprises the experimental monitoring site MFC2 (Figure 1B) with a size of about 30 ha (drainage divide with an area of 8 ha is represented by the solid blue line; Romano et al., 2018), and its immediate surroundings. The climate is Mediterranean dry sub-humid with hot-dry summer and mild-rainy winter. The average annual rainfall measured at the meteorological station (Figure 1A) amounts to 1,229.3 mm, with 68% (834.9 mm) occurring from October to March. With an average monthly rainfall of 152.2 mm, November is the wettest month (Nasta et al., 2017). The potential evapotranspiration averages 629 mm per year (Nasta et al., 2020).

Our *in-situ* measurements of θ spatially concentrated over 20 designated locations (Figure 1B) that belong to the wireless sensor network (SoilNet; Bogena et al., 2010) installed in the drainage divide of MFC2 in 2016 (Nasta et al., 2020). We further installed two stationary sensors to include the real-field condition in terms of vegetation cover and land-use close (station name VCP; vegetation cover plot represented by the red dot in Figure 1A) and bare soil conditions without any vegetation cover (station name BSP; bare soil plot represented by the green dot in Figure 1A). Later is located about 750 m from the VCP sensor and is represented by a 20 × 20 m square with the stationary sensor located in its center. For the monitoring period, the grass was frequently cut for maintaining the bare soil conditions, as exemplarily shown for March 25, 2019 (Figure 1C) using a hillshade derived from a UAV-based DEM (pixel size ~ 6 mm).

The physical-geographic settings of the moderately steep, south to west facing survey and sensor locations, with mainly silty clayey loam to clayey topsoils (Table 1), reflect the general situation of the study area (Nasta et al., 2019). Vegetation at the 21 observational points within the drainage divide of MFC2 (Figure 1B) is predominantly characterized by the occurrence of cherry (*Prunus* sp.) and walnut (*Juglans* sp.) trees growing over a 5 m regular plantation grid with the presence of dense herbaceous vegetation. At the time of investigations (Nov. 2018), all trees showed an indication of progressive abscission. Additionally, olive trees (*Olea europaea*) with heights between 3 and 5 m grow over a regular plantation grid in proximity to the observational points (Figure 1B). Furthermore, smaller plantations (e.g., olives and vineyards) and agricultural fields allocate in the southern study area (Figure 1A). In Nov. 2018, the majority of those fields lay fallow.



MATERIALS AND METHODS

A calibration procedure is necessary to convert SAR-based measurements into θ values at the ground. Therefore, we based our *in-situ* sampling framework on a measurement sequence per satellite overpasses of Sentinel-1. The constellation of the two

polar-orbiting satellites (Sentinel-1 A/B; after this, S1A and S1B) allows for a short revisit time of 6 days over Europe. In Nov. 2018, we organized eight sampling campaigns in four satellite overpass days (c.f., section *In-situ* Near-Surface Soil Moisture Measurements). Gridded information derived from the S1 SAR products and topographic data (c.f., section Data Processing and

TABLE 1 | Geographical positions (WGS 84/UTM zone 33N), topographic characteristics as derived from a DEM with a cell size of 5 m, and topsoil (0–5 cm) properties of the 20 SoilNet units (01–20) and the stationary sensor locations VCP and BSP.

Plot	X	Y	Topographic properties				Soil texture			Soil classification
			Altitude (m a.s.l.)	Slope (°)	Aspect (°)	Exposition	Sand (%)	Silt (%)	Clay (%)	
01	515685	4468257	434.9	9.0	205	SW	18.57	46.23	35.20	Silty clay loam
02	515653	4468295	436.4	9.4	237	SW	28.20	36.85	34.95	Clay loam
03	515626	4468325	434.9	8.6	240	SW	21.90	37.98	40.12	Clay
04	515600	4468365	432.2	17.8	267	W	15.00	33.80	51.20	Clay
05	515581	4468389	425.8	13.3	285	W	12.67	45.37	41.96	Silty clay
06	515556	4468419	421.0	6.2	229	SW	15.45	39.25	45.30	Clay
07	515669	4468244	432.0	7.4	194	S	34.68	40.06	25.26	Loam
08	515637	4468283	433.6	7.5	241	SW	33.77	36.36	29.87	Clay loam
09	515611	4468313	432.3	8.8	243	SW	22.02	41.50	36.48	Clay loam
10	515578	4468346	427.6	11.5	277	W	38.83	22.65	38.52	Clay loam
11	515557	4468368	423.3	9.1	286	W	15.73	37.51	46.76	Clay
12	515533	4468399	419.7	3.5	261	W	16.62	48.18	35.20	Silty clay loam
13	515646	4468225	429.7	5.3	163	S	24.20	33.64	42.16	Clay
14	515614	4468264	430.4	5.4	279	W	20.00	47.87	32.13	Clay loam
15	515588	4468294	428.4	8.3	254	W	35.10	32.98	31.92	Clay loam
16	515578	4468346	427.6	11.5	277	W	20.05	33.19	46.76	Clay
17	515558	4468369	423.3	9.1	286	W	22.28	39.20	38.52	Clay loam
18	515533	4468399	419.7	3.5	261	W	15.30	48.22	36.48	Silty clay loam
19	515585	4468253	427.4	7.7	278	W	19.02	48.35	32.63	Silty clay loam
20	515521	4468322	417.4	10.3	267	W	19.07	50.08	30.85	Silty clay loam
VCP	515602	4468297	430.6	7.9	242	SW	22.02	41.50	36.48	Clay loam
BSP	515404	4467589	351.9	17.1	172	S	19.70	25.16	55.14	Clay

S, south; SW, southwest; W, west.

Dataset Construction) and *in-situ* measurements of θ built up the dataset for the spatial estimation of θ (c.f., section Near-Surface Soil Moisture Estimation Using Machine Learning).

In-situ Near-Surface Soil Moisture Measurements

Throughout the observation period, the sampling campaigns were repeatedly conducted over the above exact 20 locations (Figure 1B) to enable fast surveying at the designated and easy-to-recognize flagged locations without using a GPS device. Due to logistic reasons (i.e., darkness during the overpass at 5 a.m.), we set the start and end times of surveying to ~ 30 min before and after the start and end times of the S1 satellite overpasses (Table 2). We used a portable, low-cost frequency Stevens HydraProbe soil sensor (Stevens Water Monitoring Systems, Inc.) with waveguide lengths of 5.7 cm and inserted it vertically into the topsoil for measuring point-scale θ . The rugged HydraProbe uses the principle of coaxial impedance dielectric reflectometry (Bellingham, 2019) and infers θ in units of water fraction (wfv) with an accuracy of ± 0.03 wfv $m^3 m^{-3}$ for fine-textured soils and a precision of ± 0.003 wfv $m^3 m^{-3}$ (Kammerer et al., 2014).

At the stationary sensor locations VCP and BSP (Figures 1B,C), further two vertically inserted HydraProbes, each connected to a GeoPrecision datalogger

TABLE 2 | Dates with start and end time of overpasses of the satellites Sentinel-1 A/B (S1A/B) during the observation period in Nov. 2018.

Date of overpass	Start/end time of overpass GMT/UTC + 1h	Satellite platform
10 Nov 2018	05:04:00–05:04:27 a.m./04:56:14–04:56:42 p.m.	S1A_d124/S1B_a44
16 Nov 2018	05:03:18–05:03:45 a.m./04:56:58–04:57:26 p.m.	S1B_d124/S1A_a44
22 Nov 2018	05:03:59–05:04:26 a.m./04:56:14–04:56:42 p.m.	S1A_d124/S1B_a44
28 Nov 2018	05:03:11–05:03:38 a.m./04:56:57–04:57:25 p.m.	S1B_d124/S1A_a44

a/d, ascending/descending satellite track. 44/124, relative orbit numbers at the start time of the data track.

(www.geo-precision.com) and following the exact technical specification as for the portable probe, continuously recorded θ in time-steps of 1 min. We considered those values recorded precisely at the time of satellite overpass (Table 2) for further analyses. Eventually, we compiled a dataset of 176 *in-situ* measurements (i.e., measurements recorded during eight overpasses at 22 position) of θ according to the proposed calibration procedures.

Data Processing and Dataset Construction

The SAR backscatter that is composed of a pair of symbols denoting the polarization—in our study VV (vertical transmit and vertical receive) and VH (vertical transmit and horizontal receive)—is determined by soil physical properties (e.g., dielectric properties and particle size distribution), topography, and vegetation (plant and canopy structure, vegetation water content). Physical or semi-empirical models attempt to disentangle these properties from the SAR measurements (intensity and phase) that, in turn, are affected by the frequency, polarization, and incidence angle (El Hajj et al., 2018) interacting with the surface properties (Wang and Qu, 2009; Das and Paul, 2015). In our modeling approach, we use such parameters that describe the SAR backscatters and statistically identify the relevant parameters to predict θ .

SAR-Based Parameters

We processed S1A and S1B C-band (5.405 GHz) Level-1 single-look complex (SLC) dual-polarization (dual-pol) SAR data using the Sentinel-1 Toolbox in the Sentinel Application Platform (SNAP). The SAR images were retrieved with a constant and identical Interferometric Wide (IW) sub-swath mode from ascending and descending tracking modes (Table 2). The local incidence angles of the S1A and S1B satellites over our study area vary between near and far range for IW2 mode with values of 36.2–41.9° and 30.6–46.1°, respectively. The look direction (i.e., antenna pointing) for the S1A/B satellites is constantly right, concerning the flight direction. The azimuth steering angle in IW2 mode is $\pm 0.6^\circ$.

We first split the swath into desirable sub-swath. Essential preprocessing steps included geocoding of images to add the orbit file with the satellite's exact position and velocity information. Data was radiometrically calibrated to retrieve complex numbers from the digital image pixels and backscattering coefficients (i.e., Beta-naught β° , Sigma-naught σ° , Gamma-naught γ° , and terrain-flattened Gamma-naught $TF\gamma^\circ$). The deburst function (Moreira et al., 2013) was applied afterward to remove border noise between the sub-swaths, thus creating continuous images. The polarimetric matrix generation function was used to create the dual-pol covariance matrix, according to Nielsen et al. (2017). Afterward, the multi-looking approach was applied to reducing the speckle noise, thus enhancing the radiometric accuracy, the image quality, and consequently, the interpretability of the SAR image (Moreira et al., 2013). Speckle noise was reduced by applying refined Lee speckle filtering (7×7 window) according to Yommy et al. (2015). Due to the topographical effect and the tilt of the satellite sensor, some distortion takes place in SAR images. Therefore, the Range Doppler terrain correction method (Filipponi, 2019) was applied to correct these distortions, and georeferencing was carried out based on (WGS 84, UTM Zone 33N).

The resulting gridded SAR parameters had a spatial resolution of 17×17 m and contained both polarizations (cross-pol VH and co-pol VV). To avoid the effects of the different local incidence angles on the backscatter, hence to enable the radiometric comparison of multi-temporal SAR images, all images were normalized to a constant incidence angle of 35°

using the quadratic cosine correction based on the theoretical model of Lambert's law and assuming Lambertian surface as typical for moderate vegetation (Mladenova et al., 2013). In addition, we computed mathematical band combinations (MBC) of both polarization in terms of addition (VH+VV), subtraction (VH-VV, VV-VH), ratio (VH/VV, VV/VH), and multiplication (VH*VV). These MBCs were reported to counteract radiometric instability and vegetation moisture variations introduced by using original polarization (e.g., Omar et al., 2017; Ahmadian et al., 2019) and hence are likely able to improve the performance of our empirical model. The definition of Kenneough's matrix and its elements for dual-pol S1 data was used based on the calculation by Schmitt et al. (2015) and Ullmann et al. (2017). SVH and SVV refer to the complex signal of cross-pol and co-pol channels, respectively (Table 3). $|SVV|^2$ and $|SVH|^2$ indicate the intensity of VV and VH channels, whereas Re and Im hold the real and the imaginary part of the inter-channel correlation and * denotes the complex conjugate (Table 3). The Kenneough elements can be divided into two intensity-only channels K0 and K1, and the two channels K5 and K6 containing the real (Re) and the imaginary (Im) part, respectively (Ullmann et al., 2017). Since the coefficients of the Kenneough matrix elements for the intensities differ slightly in the literature (e.g., Schmitt et al., 2015; Ullmann et al., 2017), the second version was used according to Schmitt et al. (2015) and is called K0S and K1S, respectively (Table 3).

Terrain Parameters

We further incorporated continuous, primary local topographic attributes (PLTA) and secondary compound topographic attributes (SCTA) as commonly used environmental proxy-drivers in θ modeling studies (e.g., Western et al., 2002; Beaudette et al., 2013). All attributes were derived using the System for Automated Geoscientific Analyses (Conrad et al., 2015). Due to the predominant processing of SAR parameters, we met the need for the exact cell sizes by re-sampling (i.e., bilinear interpolation) the DEM with an original cell size of 5 m (Nasta et al., 2017) toward the joint spatial resolution of 17×17 m. The gridded PLTAs (Table 4) are the altitude above the channel network (AAC), aspect, elevation, slope angle, and plan and profile curvature. Derived SCTAs (Table 4) comprised flow accumulation, SAGA wetness index (SWI), and topographic roughness index (TRI).

In total, we derived 43 SAR and 10 steady-state terrain parameters as spatial gridded datasets. As the SAR measurements are time-variant, they were calculated for each S1 overpass, resulting in 433 single spatial grids.

Near-Surface Soil Moisture Estimation Using Machine Learning

Our ML approach used is the random forests (RF) regression analysis. RF is an ensemble of randomized Classification and Regression Trees (CART), e.g., Breiman (2001). We applied the "caret" R package for regression analysis (Kuhn et al., 2020) for the R Project for Statistical Computing and employed our dataset of 176 *in-situ* measurements (c.f., section *In-situ* Near-Surface Soil Moisture Measurements).

TABLE 3 | List of 43 SAR-based parameters as used as predictors in $\theta_{\text{Nov}}^{\text{SAR}}$ and $\theta_{\text{Nov}}^{\text{SAR+Terrain}}$.

Parameter grid name	Parameter (<i>polarization</i>); unit	Purpose
B_VH	β° (VH)	SAR backscatter referring to the radar brightness in the slant range plane (Small, 2011)
B_VV	β° (VV)	
B_VH_div_VV	β° (VH)/ β° (VV)	
B_VH_min_VV	β° (VH) – β° (VV)	
B_VH_multi_VV	β° (VH) * β° (VV)	
B_VH_plus_VV	β° (VH) + β° in (VV)	
B_VV_div_VH	β° (VV)/ β° (VH)	
B_VV_min_VH	β° (VV) – β° (VH)	
G_VH	γ° (VH)	SAR backscattering coefficients in the plane perpendicular to the line of sight from sensor to an ellipsoidal model of the earth surface (Small, 2011)
G_VV	γ° (VV)	
G_VH_div_VV	γ° (VH)/ γ° (VV)	
G_VH_min_VV	γ° (VH) – γ° (VV)	
G_VH_multi_VV	γ° (VH) * γ° (VV)	
G_VH_plus_VV	γ° (VH) + γ° (VV)	
G_VV_div_VH	γ° (VV)/ γ° (VH)	
G_VV_min_VH	γ° (VV) – γ° (VH)	
S_VH	σ° (VH)	SAR backscattering coefficients normalized to ground area which locally tangent to an ellipsoidal model of earth surface (Small, 2011)
S_VV	σ° (VV)	
S_VH_div_VV	σ° (VH) / σ° (VV)	
S_VH_min_VV	σ° (VH) – σ° (VV)	
S_VH_multi_VV	σ° (VH) * σ° (VV)	
S_VH_plus_VV	σ° (VH) + σ° (VV)	
S_VV_div_VH	σ° (VV)/ σ° (VH)	
S_VV_min_VH	σ° (VV) – σ° (VH)	
TFG_VH	TF γ° (VH)	Terrain flattened (TF) SAR backscattering coefficient accounting for the topographic variations (Small, 2011)
TFG_VV	TF γ° (VV)	
TFG_VH_div_VV	TF γ° (VH)/TF γ° (VV)	
TFG_VH_min_VV	TF γ° (VH) – TF γ° (VV)	
TFG_VH_multi_VV	TF γ° (VH) * TF γ° (VV)	
TFG_VH_plus_VV	TF γ° (VH) + TF γ° (VV)	
TFG_VV_div_VH	TF γ° (VV)/TF γ° (VH)	
TFG_VV_min_VH	TF γ° (VV) – TF γ° (VH)	
C11	1st element of polarimetric C2 matrix; -	The 1st and 4th elements (i.e., C11 and C22) refer to ensemble averaging of first (co-pol VV) and second channel (cross-pol VH), which represents the modulus (amplitude). The 2nd and 3rd elements (i.e., C12imag and C12real) contain the information about the real and imagery part of the channels (Erten, 2012; Nielsen et al., 2017; Mandal et al., 2019)
C12imag	2nd element of polarimetric C2 matrix; -	
C12real	3rd element of polarimetric C2 matrix; -	
C22	4th element of polarimetric C2 matrix; -	
K0	Kennaugh element K0 = $0.5 (SVV ^2 + 2 SVH ^2)$;	Total intensity (sum) can be found in K0 and K0S. The difference between co- and cross-pol intensity appears in K1 (and K1S). The elements K5 and K6 relates to real (Re) and imagery (Im) information and consequently hold useful information for deterministic—usually man-made—targets exclusively (Schmitt et al., 2015; Ullmann et al., 2017)
K0S	Kennaugh element K0S = $ SVV ^2 + SVH ^2$; -	
K1	Kennaugh element K1 = $0.5 (SVV ^2 - 2 SVH ^2)$;	
K1S	Kennaugh element K1S = $ SVV ^2 - SVH ^2$; -	
K5	Kennaugh element K5 = $\text{Re} (SVV SVH^*)$; -	
K6	Kennaugh element K6 = $-\text{Im} (SVV SVH^*)$; -	
ProjLocIncAngle	Projected local incidence angle	Local incidence angle projected into the range plane; refers to the angle between the incoming radiation and the projected surface normal into the range plane (SNAP)

VV, vertical transmit/vertical receive; VH, vertical transmit/horizontal receive; β° , Beta-naught; σ° , Sigma-naught; γ° , Gamma-naught.

TABLE 4 | Ten terrain parameters as used as predictors in $\theta_{\text{Nov}}^{\text{SAR+Terrain}}$.

Parameter grid name	Type of terrain parameter; Parameter; unit	Purpose
AAC	PLTA; Altitude above channel network; m a.s.l.	Vertical distance of the original elevation to the interpolated channel network base level elevation; proxy on the relative soil gravitational potentials, thus local draining potentials (Nobre et al., 2011)
Aspect_d	PLTA; Aspect; °	Key attribute in describing the landscape's spatial heterogeneity affecting the surface energy balance and as such the soil water retention capacity and water availability (Geroy et al., 2011)
Aspect_r	PLTA; Aspect; rad	
Elevation	PLTA; Altitude; m a.s.l.	Key attribute in describing the physical landscape properties and their spatial patterns (Beaudette et al., 2013)
FlowAcc	SCTA; Flow accumulation	Cumulative count of the pixel number naturally draining toward the outlet, thus a proxy on the effects of the depth and velocity of flow; in the present study integrating the multiple flow direction based on the maximum downslope gradient (Top-Down processing; Qin et al., 2011)
PlanCurvature	PLTA; Plan curvature; rad m ⁻¹	Curvature of a contour line formed by intersecting a horizontal plane with the surface; proxy on the convergence or divergence of water during downslope flow (Amatulli et al., 2018)
ProfileCurvature	PLTA; Profile curvature; rad m ⁻¹	Curvature of the surface in the direction of the steepest slope affecting the acceleration or deceleration of surface water and indicating local changes of inflow velocity (Bogaart and Troch, 2006; Contador et al., 2006)
Slope	PLTA; Slope angle; °	Key attribute in describing the landscape's spatial heterogeneity and describing catchment-related hydrological processes (e.g., Bogaart and Troch, 2006)
SWI	SCTA; SAGA Wetness Index; -	Measure on the spatial scale effect of the topography on hydrological processes and identifier in flow paths; based on the modified catchment area for predicting a realistic terrain-related soil moisture potential (Böhner and Selige, 2006)
TRI	SCTA; Topographic Roughness Index; -	Descriptor of the terrain profile indicating the surface heterogeneity of the landscape (Riley et al., 1999)

PLTA, primary, local topographic attribute; SCTA, secondary compound topographic attribute.

The prediction datasets included (i) SAR parameters only (after this $\theta_{\text{Nov}}^{\text{SAR}}$) and (ii) the combination of SAR and terrain parameters (after this $\theta_{\text{Nov}}^{\text{SAR+Terrain}}$). Since 70–30% splitting for training/testing did not prove successful because of the limited number of point-observations, we decided for k -fold cross-validation with $k = 10$, splitting the randomly shuffled data into 10 complementary subsets. From those 10 subsets, nine subsets were used to build and fit the model and one for testing the model. This step was repeated until each subset was used for testing once. The total error rate of the model was then calculated as the average of the individual error rates of the $k = 10$ individual runs. For both RF models, we kept the tuning parameters n_{tree} (number of trees) and m_{try} (number of variables randomly sampled at each split) as default values ($n_{\text{tree}} = 500$, $m_{\text{try}} = \text{no. of variables}/3$), which we considered as appropriate as RF is insensitive to over-adaptation. We chose the coefficient of determination (R^2), the root-mean-square error (RMSE), and the corresponding mean absolute error (MAE) as estimation performance indicators.

Additionally, we computed the predictor importance for $\theta_{\text{Nov}}^{\text{SAR}}$ and $\theta_{\text{Nov}}^{\text{SAR+Terrain}}$. It presents the evaluated, scaled relationship between each predictor and the estimates. We used the percentage increase in mean squared error (%IncMSE) as the importance measure for the regression models. It refers to the MSE of predictions, which is calculated by comparing the difference in the influence on model accuracy between a parameter and a randomly permuted version of it at each split in

a tree when it is selected (Breiman, 2001). Subsequent importance analysis provides a ranking of predictors and information of how much the model accuracy decreases if we leave out a particular parameter in estimating θ . A one-way ANOVA test was conducted on the averaged values of the topmost ranking parameters as revealed by $\theta_{\text{Nov}}^{\text{SAR}}$ to analyze potential differences between the SAR measurements in the morning (5 a.m.) and afternoon (5 p.m.) of each overpass day (Table 2). We assume this step to better identify those parameters sensitive to within-day variations of θ .

RESULTS

In-situ Soil Moisture

In-situ θ values from all eight sampling campaigns ($n = 176$) spanned from $0.14 \text{ m}^3 \text{ m}^{-3}$, measured at position 18 on Nov. 16 (5 a.m.) to $\sim 0.62 \text{ m}^3 \text{ m}^{-3}$ at the position BSP on Nov. 22 (5 p.m.). With an average of $\sim 0.34 \text{ m}^3 \text{ m}^{-3}$ and a standard deviation (SD) of almost $0.08 \text{ m}^3 \text{ m}^{-3}$, the total measured θ values follow a distinct positive, right-tailed distribution (Table 5).

On a daily basis, average *in-situ* θ values ranged between $\sim 0.30 \text{ m}^3 \text{ m}^{-3}$ in the morning (5 a.m.) and afternoon (5 p.m.) of Nov. 16 and $\sim 0.42 \text{ m}^3 \text{ m}^{-3}$ (Nov. 22, 5 p.m.; Table 5). The highest θ values ($0.48\text{--}0.62 \text{ m}^3 \text{ m}^{-3}$) were recorded without exception at the sensor location BSP. For all sampling campaigns, θ measured at the BSP location (Figure 1C) was twice as high as at 21 locations within the MFC2 area (Figure 1B). On average,

TABLE 5 | Descriptive statistics of the *in-situ* soil moisture ($\text{m}^3 \text{m}^{-3}$) over the entire 22 locations at the time of satellite overpasses.

Statistics	Total	10-Nov. 2018		16-Nov. 2018		22-Nov. 2018		28-Nov. 2018	
		5 a.m.	5 p.m.	5 a.m.	5 p.m.	5 a.m.	5 p.m.	5 a.m.	5 p.m.
Number of observation (<i>n</i>)	176	22	22	22	22	22	22	22	22
Min	0.140	0.290	0.280	0.140	0.200	0.330	0.320	0.338	0.319
Max	0.617	0.571	0.524	0.488	0.479	0.612	0.617	0.596	0.596
1st quartile	0.330	0.358	0.358	0.260	0.251	0.400	0.372	0.393	0.373
Median	0.390	0.395	0.400	0.295	0.305	0.410	0.410	0.410	0.400
3rd quartile	0.410	0.410	0.410	0.308	0.340	0.428	0.460	0.410	0.410
Average	0.377	0.385	0.382	0.297	0.304	0.412	0.423	0.409	0.403
SD	0.075	0.057	0.051	0.068	0.070	0.056	0.064	0.049	0.057
CV %	20.00	14.80	13.40	22.90	23.20	13.70	15.20	12.00	14.20
Skewness*	0.090	1.052	0.155	0.592	0.433	1.711	0.951	2.258	1.605
<i>p</i> -value	< 0.001	0.004	0.005	0.043	0.494	0.001	0.111	< 0.001	0.003

*Pearson's coefficient of skewness. Normality of distribution was tested using the Shapiro-Wilk test with $\alpha = 0.05$.

the difference amounts to 140%. For all sampling campaigns, the variation of measured θ values compared to the averages was distinct, with the lowest coefficient of variation (CV) of 12% in the morning of Nov. 28 and the highest CV of 23.2% in the afternoon of Nov. 16 (Table 5). The development of *in-situ* θ indicated spatiotemporal variations, hence allows us to distinguish between dry and moist situations.

Performance of the RF Models Over the Observational Points

Initial θ estimation of the RF models (c.f., sections Near-Surface Soil Moisture Estimation Using Machine Learning) over the 22 sampling and sensor locations returned very high overall coefficients of determination with $R^2 = 0.86$ for $\theta_{\text{Nov}}^{\text{SAR}}$ and $R^2 = 0.89$ for $\theta_{\text{Nov}}^{\text{SAR+Terrain}}$, even though minor systematic errors in the form of overestimations of low θ ($< 0.30 \text{ m}^3 \text{ m}^{-3}$) were observed in both models (Figure 2). Compared to $\theta_{\text{Nov}}^{\text{SAR}}$, the model $\theta_{\text{Nov}}^{\text{SAR+Terrain}}$ can reduce this systematic error slightly. This becomes visible in the corresponding values of RSME = $0.025 \text{ m}^3 \text{ m}^{-3}$ and MAE = $0.020 \text{ m}^3 \text{ m}^{-3}$ indicating a higher agreement between observed and predicted θ values when including topographic information (Figure 2B).

Additional support in understanding the temporal pattern of θ is gained with the provision of the predictive capability of $\theta_{\text{Nov}}^{\text{SAR}}$ and $\theta_{\text{Nov}}^{\text{SAR+Terrain}}$ for the single S1 overpasses in the morning (5 a.m.) and in the afternoon (5 p.m.; Figure 3). The model $\theta_{\text{Nov}}^{\text{SAR}}$ generally returned high values of R^2 spanning from 0.71 (Nov. 16, 5 p.m.) to 0.84 (Nov. 22, 5 p.m.). Except for Nov. 16, the explained variances are slightly higher in the afternoon ($R^2 = 0.79$ for Nov. 10, $R^2 = 0.84$ for Nov. 22) than in the morning ($R^2 = 0.72$ for Nov. 10, $R^2 = 0.82$ for Nov. 22). For Nov. 28, R^2 nearly equals for both S1A/B overpasses ($R^2 \sim 0.79$ at 5 a.m./5 p.m.). The lowest RMSE values of $\sim 0.023 \text{ m}^3 \text{ m}^{-3}$ indicate the highest agreement between observed and estimated at Nov. 10 (5 p.m.) and Nov. 28 (5 a.m.), respectively. In contrast, the lowest agreements (RMSE $\sim 0.034 \text{ m}^3 \text{ m}^{-3}$ and $\sim 0.038 \text{ m}^3 \text{ m}^{-3}$, respectively) were found for Nov. 16 (Figure 3). The same trend of prediction capability applies for

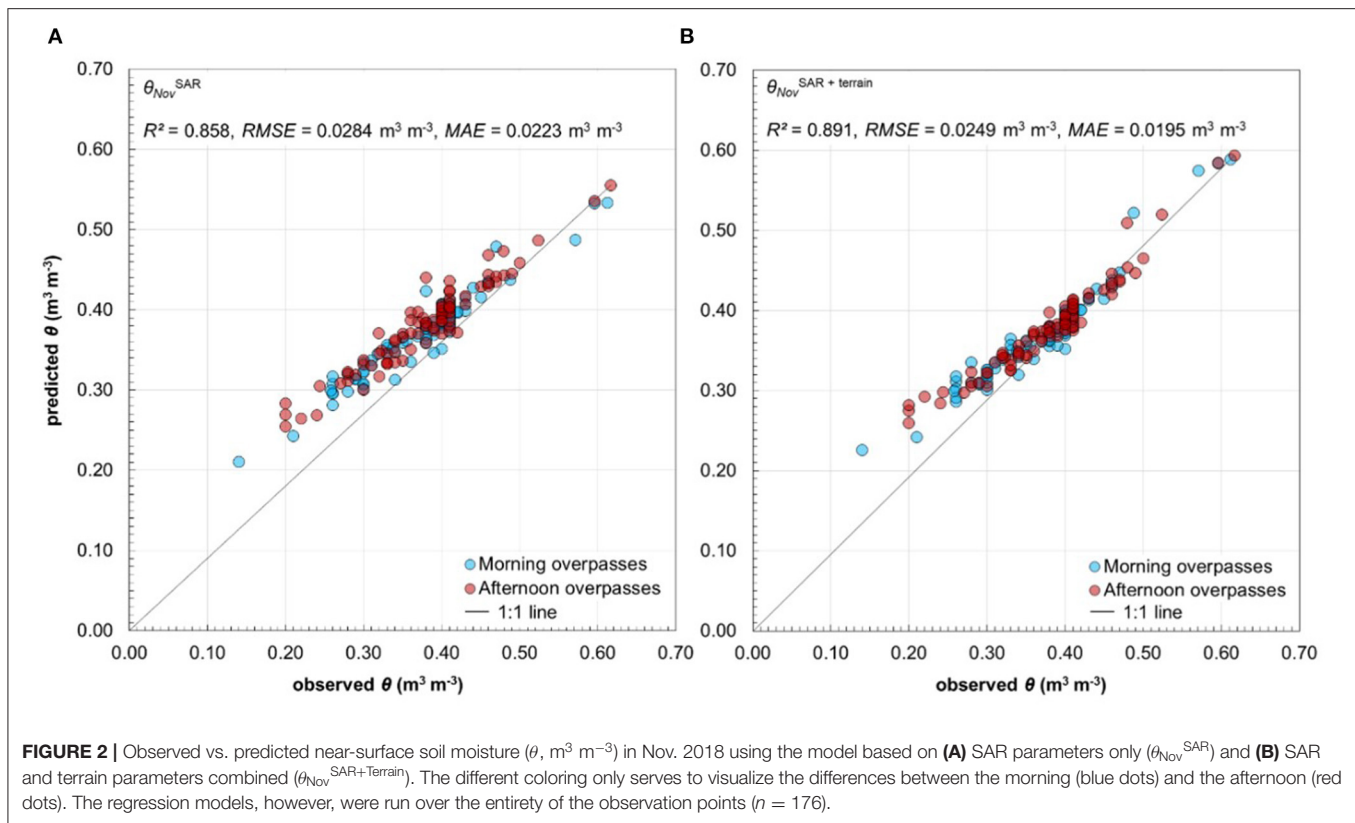
$\theta_{\text{Nov}}^{\text{SAR+Terrain}}$ too. Hence, the lowest values of R^2 were found for Nov. 16 ($R^2 = 0.71$ at 5 a.m., $R^2 = 0.69$ at 5 p.m.). The highest R^2 values from ~ 0.88 to 0.90 were revealed for all other S1 overpasses (Figure 3). Despite for Nov. 16, the incorporation of terrain parameters thus increases the prediction accuracy by the averaged factor of 1.1 and results in a higher agreement between observed and estimated θ (RMSE from $0.016 \text{ m}^3 \text{ m}^{-3}$ to $0.021 \text{ m}^3 \text{ m}^{-3}$; Figure 3). Again, Nov. 16 remains excluded from this observation.

Time Series of the Near-Surface Soil Moisture

Spatial estimates of θ from $\theta_{\text{Nov}}^{\text{SAR}}$ and $\theta_{\text{Nov}}^{\text{SAR+Terrain}}$ for the eight S1 overpasses significantly differ with $p < 0.001$ as revealed by ANOVA testing (Figure 3).

The time series derived from $\theta_{\text{Nov}}^{\text{SAR}}$ show distinct, scattered patterns (pixel size of 17 m) of θ . At the beginning of our observation period, the spatial estimates average to $0.380 \text{ m}^3 \text{ m}^{-3}$ (CV 9.6%) and $0.394 \text{ m}^3 \text{ m}^{-3}$ (CV 12.6%) on Nov. 10 (at 5 a.m. and 5 p.m., respectively). For Nov. 16, the modeling resulted in a slightly lower spatial average of $0.376 \text{ m}^3 \text{ m}^{-3}$ (CV 9.7%) on Nov. 16 (at 5 a.m.; Figure 3). Six days later (Nov. 22), spatial patterns of estimated θ change to moist conditions at 5 a.m. (average = $0.402 \text{ m}^3 \text{ m}^{-3}$, CV 11.9%) and at 5 p.m. (average = $0.419 \text{ m}^3 \text{ m}^{-3}$, CV 13.3%). On Nov. 28, the θ situation nearly returns to that at the start day with similar spatial patterns and average spatial estimates for the morning and afternoon, respectively (Figures 3, 4). The spatial patterns depict a distinct preference of dry and wet pixels so that the area of MFC2 (Figures 1A,B) appears as distinctly drier compared to the very north-eastern and southern part of the study area (Figure 4). The spatial development of θ takes the form of narrow wet bands stretching toward the central study area.

Even though θ estimates obtained from $\theta_{\text{Nov}}^{\text{SAR+Terrain}}$ temporally evolve comparable to those from $\theta_{\text{Nov}}^{\text{SAR}}$ (Figure 4), the predicted spatial patterns characterize as less scattered but rather smoothed (Figure 5). For the single S1 overpasses, spatial estimates average from 0.377 (Nov. 16, 5 a.m.) to $0.396 \text{ m}^3 \text{ m}^{-3}$



(Nov. 22, 5 p.m.) for $\theta_{\text{Nov}}^{\text{SAR+Terrain}}$ (**Figure 3**). The coefficient of variations depicted (CV 12.0% for Nov. 22, 5 p.m. to CV 13.9% for Nov. 28, 5 a.m.) point at slightly higher dispersion of spatial estimates around the average when predicted with $\theta_{\text{Nov}}^{\text{SAR+Terrain}}$ as compared to predicting with SAR parameters only (**Figure 3**).

Again remarkable is the constantly drier situation in the central study area, represented by the area of MFC2, and supplementing to this, the constantly wetter condition in the southern study area with few very wet spots ($\sim 0.51\text{--}0.56 \text{ m}^3 \text{m}^{-3}$) surrounding the sensor location BSP (**Figure 5**). A narrow north-south to the west-east stretching wet band ($\sim 0.44\text{--}0.48 \text{ m}^3 \text{m}^{-3}$) in the very northern study area constantly occurs throughout the entire time series (**Figure 5**).

However, the influence of the terrain parameters on the $\theta_{\text{Nov}}^{\text{SAR+Terrain}}$ model becomes visible when comparing spatial predictions with that of $\theta_{\text{Nov}}^{\text{SAR}}$ (**Figures 4, 5**). While the θ pattern returned by $\theta_{\text{Nov}}^{\text{SAR+Terrain}}$ follows a clear topographical trend and shows the same patterns over the entire period under study, $\theta_{\text{Nov}}^{\text{SAR}}$ lacks these constants.

Importance of SAR and Terrain Parameters in Near-Surface Soil Moisture Estimation

The importance of SAR and terrain parameters incorporated in the ensemble models $\theta_{\text{Nov}}^{\text{SAR}}$ and $\theta_{\text{Nov}}^{\text{SAR+Terrain}}$ (c.f., section Near-Surface Soil Moisture Estimation Using Machine Learning) are sorted decreasingly from top to bottom in **Figure 6**. For $\theta_{\text{Nov}}^{\text{SAR}}$, the topmost 10 positions (%IncMSE > 47) are

predominantly occupied by four out of six Kennaugh elements (K1, K6, K1S, and K0S) and three out of four elements of the C2 matrix (C22, C12imag, and C11). This can be due to information hold by these variables, e.g., the total intensity in K0S and the difference between co-pol and cross-pol intensity in K1. Thus, it is likely that changing the dielectric constant of the target (i.e., soil and vegetation cover) due to the alignment of water dipoles directly affects the intensities and amplitudes (i.e., C11 and C22). The VV co-pol SAR backscatters $\text{TF}\gamma^\circ$ and β° , and the mathematical band combination (MBC) of $\gamma^\circ\text{VV} + \gamma^\circ\text{VH}$ (**Table 3**) further rank amongst the 10 topmost positions (**Figure 6A**). Up to position 20 (%IncMSE > 31), mainly MBCs of the SAR backscatters β° , γ° , $\text{TF}\gamma^\circ$, and σ° either with VH+VV or VH*VV rank with steadily decreasing parameter importance (**Figure 6A**). Below ranking position 20 (%IncMSE < 30), the VH cross-pol SAR backscatters β° , γ° , and σ° , as well as the remaining 16 MBCs of the SAR backscatters (i.e., VV-VH or VH-VV), and two elements of the C2 (C12real) and Kennaugh (K5) matrix indicate minor prediction importance.

The incorporation of terrain parameters into the regression modeling ($\theta_{\text{Nov}}^{\text{SAR+Terrain}}$) clearly emphasizes the topography as most crucial in spatially estimating θ (**Figure 6B**). The 10 topmost ranking positions (%IncMSE > 47) are occupied with seven out of 10 terrain parameters. They mainly characterize as PLTA (i.e., AAC, elevation, slope, aspect in degree, and profile curvature). Among the topmost ranking parameters, the wetness and topographic roughness indices additionally point at the topographic heterogeneity as necessary. Plan curvature

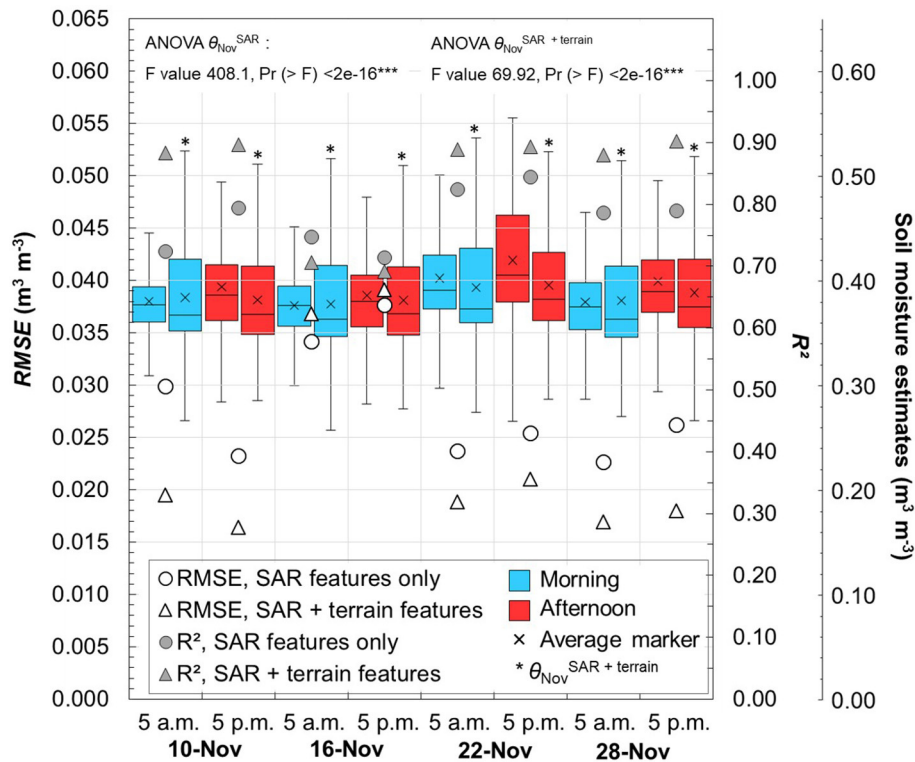


FIGURE 3 | Estimation performance indicators R^2 (right y axis) and RMSE (left y axis) for the models θ_{Nov}^{SAR} and $\theta_{Nov}^{SAR+Terrain}$ to approximate the *in-situ* soil moisture (θ) at the eight Sentinel-1 overpasses in Nov. 2018. The box plots show the distributions of the predicted spatial estimates (second y axis at the right) of θ in the morning (blue color) and in the afternoon (red color) for θ_{Nov}^{SAR} and $\theta_{Nov}^{SAR+Terrain}$ (indicated by *).

(%IncMSE \sim 38) and flow accumulation (%IncMSE \sim 4) show moderate to hardly any predictive strength. Among the SAR parameters, the modulus of the polarization channels C11 (%IncMSE \sim 63) and C22 (%IncMSE \sim 49) and the Kennaugh element K1 (%IncMSE \sim 44) are of highest prediction importance in $\theta_{Nov}^{SAR+Terrain}$. Below ranking position 10 (%IncMSE $<$ 40), further C2 and Kennaugh matrix elements and the SAR backscatters and their MCBs rank in nearly the same order as revealed for θ_{Nov}^{SAR} (Figures 6A,B).

DISCUSSION

Spatiotemporal Dynamics of the Near-Surface Soil Moisture

We consider the observed temporal pattern of near-surface soil moisture (θ) to follow the hydrometeorological conditions in the study area. As shown by Nasta et al. (2019), the absence of rainfall and hence, concurrent evapotranspiration losses characterize the first days in November 2018 before starting our observation period and the days between Nov. 10 and Nov. 16. Thus, a likely effect of drying up of θ resulting in a distinct dry spatial pattern in the morning and afternoon of both days is supposed. This can be depicted from both models, the SAR-only-based RF (θ_{Nov}^{SAR}) and the SAR+terrain-based RF ($\theta_{Nov}^{SAR+Terrain}$). Two comparatively large rainfall events between Nov. 18 and Nov. 22 amounting to a cumulative rainfall of \sim 79 mm (Nasta et al., 2019)

are discussed to have triggered field wetting with an effect on the spatiotemporal patterns of θ .

When looking at the spatial pattern of the model outputs (Figures 4, 5), an all-time higher spatial θ becomes apparent for the direct surrounding of the sensor location BSP, referring to the bare soil plot (Figure 1C). Compared to the predominantly silty clay loamy and clay loamy topsoil at MFC2 (Table 1), the bare soil plot is the richest in clay (55% clay). Only plot 04 presents a frequently small wet spot in the time series derived from $\theta_{Nov}^{SAR+Terrain}$ as to be seen from the mapped estimates in the center of the hydrological divide (Figure 5), which is interestingly not apparent from the mapping results when using the SAR parameters as only predictor ensemble (Figure 4). With 51.2%, the clay content at plot 04 is nearly comparable to that at BSP. Specifically, we assume the higher clay content at both positions and the differences in topsoil textures in general (Table 1) to explain the spatial patterns of θ . Numerous studies confirm soils with finer particle sizes (i.e., clay) to retain more soil water, thus influencing the residual water content, independently from short-term changes in rainfall patterns (e.g., Amanabadi et al., 2018; Jiménez-de-Santiago et al., 2019).

Although the time series of θ shows a distinct response to rainfall events (Nasta et al., 2019) and patterns appear plausible, the estimation of θ is, to a certain extent, questionable. Probably, the vegetation is constraining the spatial estimation. The reason might be two-fold. On the one hand, the vegetation considerably

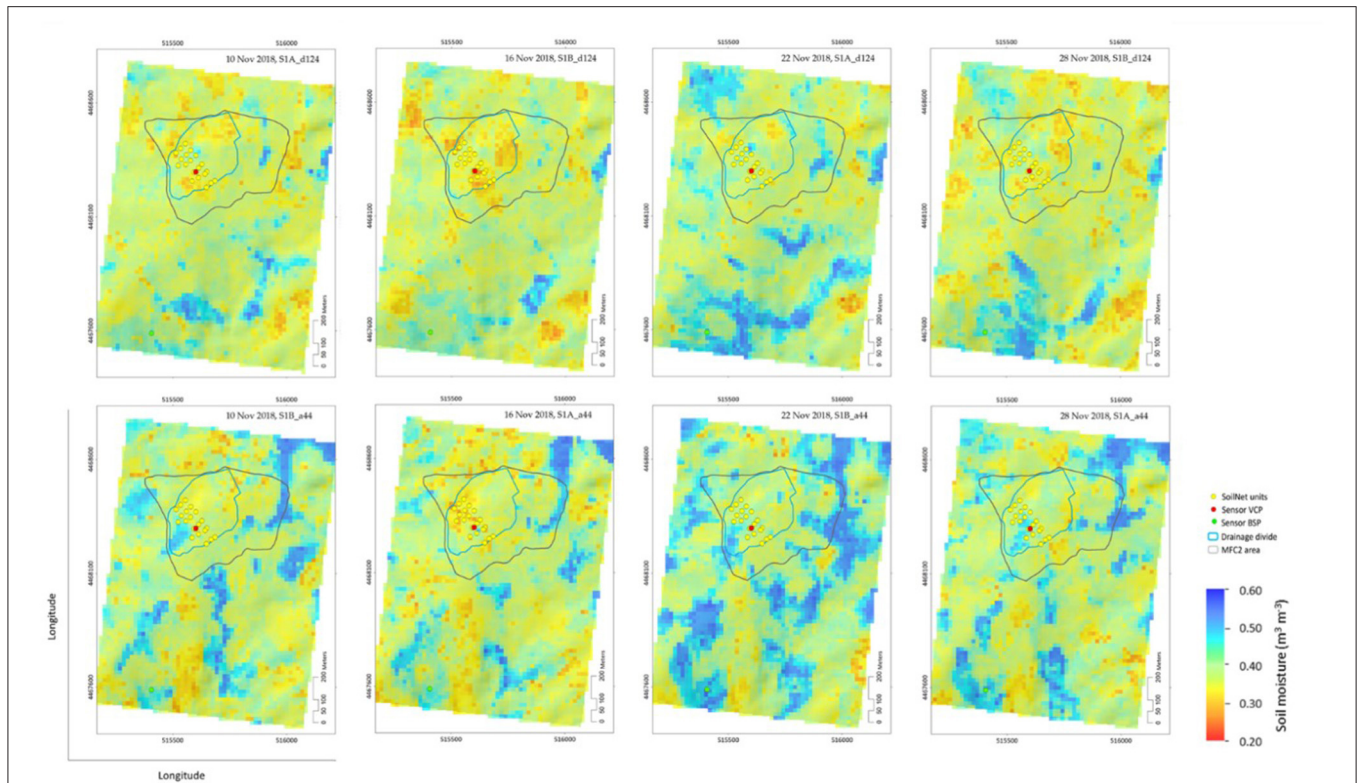


FIGURE 4 | Spatiotemporal patterns of predicted near-surface soil moisture (θ , $\text{m}^3 \text{m}^{-3}$) in the morning (upper row) and in the afternoon based on SAR parameters only ($\theta_{\text{Nov}}^{\text{SAR}}$). The pixel size is $17 \times 17 \text{ m}$. White pixels refer to no data pixels.

governs θ by interception and evaporation, and root water uptake (e.g., Guderle and Hildebrandt, 2015). Yet, this effect cannot be quantified as the survey of such information (i.e., root activity) was considered not feasible in this study. However, we suppose this effect is not distinctly decisive for the present study as Guderle and Hildebrandt (2015) discuss θ data at least from measured recordings to already include information from evapotranspiration and root water uptake patterns. This might be even more relevant as the mainly two-layer vegetation with trees and grass cover (c.f., section Study Area) paused in terms of vegetation activity, and the generally smaller trees at MFC2 were characterized by ongoing abscission at the time of surveying, respectively. However, the vegetation cover poses another uncertainty in our study; The fact that C-band signals are largely influenced by vegetation cover may attenuate the soil contribute to the signal (Quesney et al., 2000). It is also possible that the signals penetrate through vegetation but cannot penetrate the soil as they would do without vegetation.

Additionally, the higher soil moisture content is reported to reduce the penetration depths of the SAR beam (Gorrah et al., 2014). Thus, under- and overestimations in the spatial patterns of θ may be possible, even though the direction of their response to rainfall is considered plausible. The L-band (compared to shorter C-band wavelength) is often recommended to minimize the influence of vegetation on SAR backscattering in estimating θ (e.g., Barrett et al., 2009; El Hajj et al., 2018;

Zribi et al., 2019). However, this comes mainly at the cost of data availability and accessibility, data expenses (e.g., Sinha et al., 2015), and the reduced spatial resolution (e.g., Babaeian et al., 2019). Depending on the area of interest and research objective in terms of feasibility, a suitable wavelength must be selected.

Importance of SAR Backscatters and the Topography

We chose our terrain parameters (Table 4) for incorporation into the modeling framework because of their capability to reflect the site-specific soil hydraulic behavior. Therefore, the incorporation of the “hydrologically meaningful” terrain parameters (Guevara and Vargas, 2019) is in line with studies that employed geomorphometry in empirical approaches, such as for downscaling RS-based soil moisture estimates (e.g., Guevara and Vargas, 2019; Mascaro et al., 2019; Zappa et al., 2019). For instance, the SAGA wetness index (SWI) simulates hydrological processes based upon the topography. It is thus capable of reflecting the spatial distribution of soil moisture (Böhner and Selige, 2006), a fact that is regarded as the reasons for the explanatory strength of the SWI in our spatiotemporal modeling of θ (Figure 6A).

Though we did not quantify the effect of the soil as meaningful properties such as texture, organic matter content, bulk density, and hydraulic conductivity (e.g., Jiménez-de-Santiago et al., 2019) were not available as gridded datasets, we ascribe our

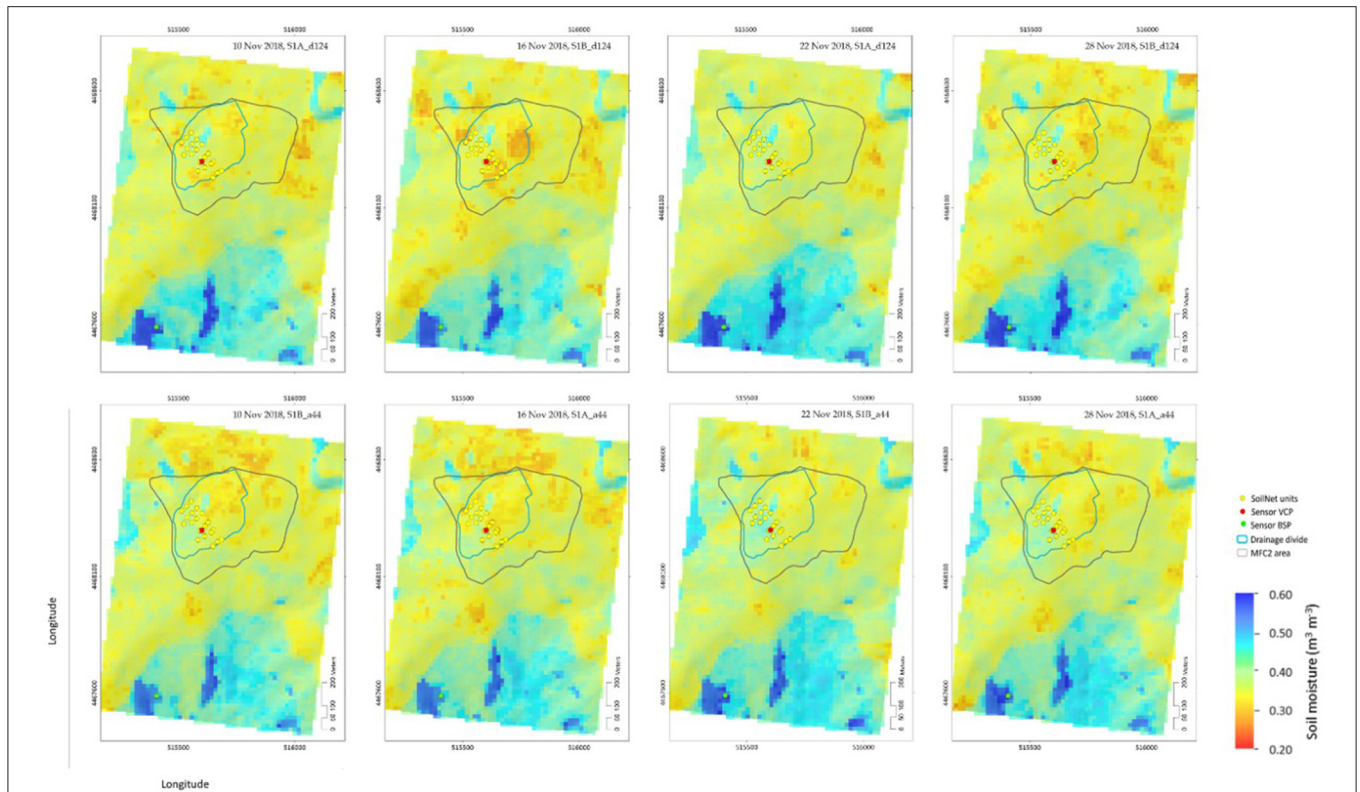


FIGURE 5 | Spatiotemporal patterns of predicted near-surface soil moisture (θ , $\text{m}^3 \text{m}^{-3}$) in the morning (upper row) and in the afternoon (lower row) based on the combination of SAR and terrain parameters ($\theta_{\text{Nov}}^{\text{SAR+Terrain}}$). The pixel size is $17 \times 17 \text{ m}$. White pixels refer to no data pixels.

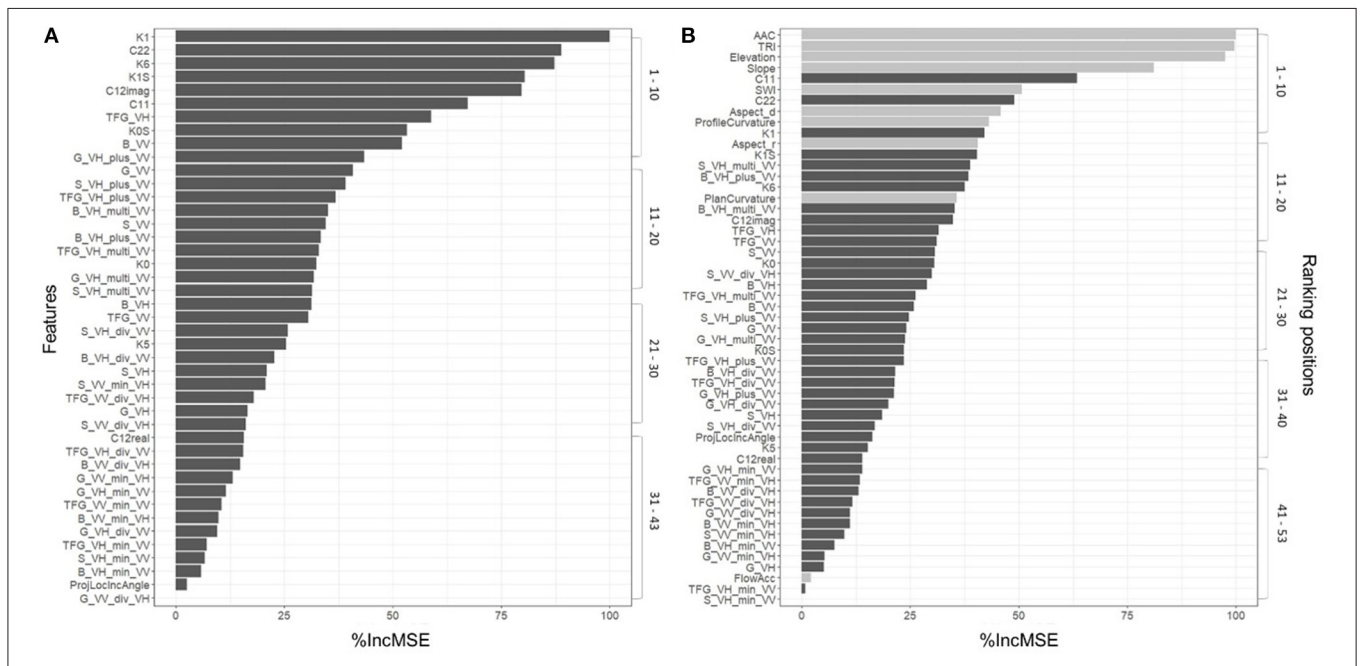


FIGURE 6 | Importance of SAR (dark gray) and terrain (light gray) parameters in the ensemble models built from (A) 43 SAR parameters for $\theta_{\text{Nov}}^{\text{SAR}}$ and (B) 53 SAR and terrain parameters for $\theta_{\text{Nov}}^{\text{SAR+Terrain}}$.

$\theta_{\text{Nov}}^{\text{SAR+Terrain}}$ approach high capability to approximating more realistic patterns of θ compared to $\theta_{\text{Nov}}^{\text{SAR}}$. We mainly assume this to result from the incorporation of “vague proxies” (Behrens and Viscarra Rossel, 2020) in the form of our DEM derivatives (Table 4). They are primarily discussed to drive spatial soil property variations and therefore are commonly applied in soil-landscape modeling and digital soil mapping (e.g., Mulder et al., 2011; Zhang et al., 2017). Besides this proxy effect, the topography holds a distinct direct effect; Due to its location in a topographic depression, i.e., characterized by a larger specific contributing area (Qin et al., 2011), the southern-most study area that is represented by the plot location BSP, is found to be generally wetter compared to the higher allocated and less inclined area of MFC2 (Figure 5). Against this background, we assume our choice to incorporate topographic data as crucial baseline information toward the spatial mapping of θ .

In terms of the employed SAR parameters, the intensity-based elements K1, K1S, and K6 and C22, C12imag and C11 from the dual-polarized Kennaugh matrix and covariance matrix, respectively (c.f., section SAR-Based Parameters) were identified most crucial, in both approaches $\theta_{\text{Nov}}^{\text{SARand}}$ and $\theta_{\text{Nov}}^{\text{SAR+Terrain}}$ (Figure 6). To facilitate the discussion of SAR parameters, we tested those topmost important SAR parameters for their temporal significance during the individual S1 satellite overpasses (i.e., at 5 a.m./p.m.) using one-way ANOVA. This is as we assume that they act sensitively to the study area’s hydrometeorological conditions (c.f., section Spatiotemporal Dynamics of the Near-Surface Soil Moisture). Figure 7 provides information on their time-variant distributions in the RF model $\theta_{\text{Nov}}^{\text{SAR}}$. Averaged values of the Kennaugh elements K1 and K1S range from 0.009 to 0.012 and from 0.033 to 0.041, respectively. On average, they were measured lowest on Nov. 10 (5 a.m.) and highest on Nov. 22 (5 p.m.). Element K6 was measured lowest (average of -0.001) on Nov. 16 (5 p.m.) and highest (up to an average of 0.001) at 5 a.m. on Nov. 16, 22, and 28. Averaged values of the C2 elements vary from 0.045 to 0.058 (C22), from -0.001 to 0.001 (C12imag), and from 0.012 to 0.017 (C11), respectively. Except for C12imag, the lowest and highest averages were again revealed for Nov. 16, 5 a.m. and Nov. 22, 5 a.m./5 p.m., respectively (Figure 7). We not only observe significant differences ($p < 0.001$ mainly) in the individual elements between morning and afternoon shots of the same day, except for C11 on Nov. 22 (Figure 7). We also observe a qualitative relationship between those days assigned to drier conditions (Nov. 10 and Nov. 16.) and to wetter conditions (Nov. 22), respectively (c.f., section Spatiotemporal Dynamics of the Near-Surface Soil Moisture).

Although empirical models need a significant number of experimental measurements to establish a useful empirical relationship for inversion of soil parameter from backscattering observations (Oh et al., 1992; Walker et al., 2004), and therefore are site-specific and not recommended applicable for monitoring special events or trends beyond the domain of calibration, they are reported to yield accurate results (Wigneron et al., 2003). For this reason, we developed an empirical ML model using an RF 10-fold cross-validation to estimate θ . Even if there is a high correlation between the predictor parameters observed (Table 6), such as between the second versioned Kennaugh elements K0 and

K0S (Schmitt et al., 2015), their inclusion was regarded essential to evaluate their effectiveness and importance in terms of their capability for empirically estimating θ . This consideration is supported by Schmitt and Brisco (2013), who state the Kennaugh elements, though derived from dual-co-polarized imaging modes (HH and VV), to distinctly contribute to the robustness and reliability in change detection sufficient for wetland monitoring.

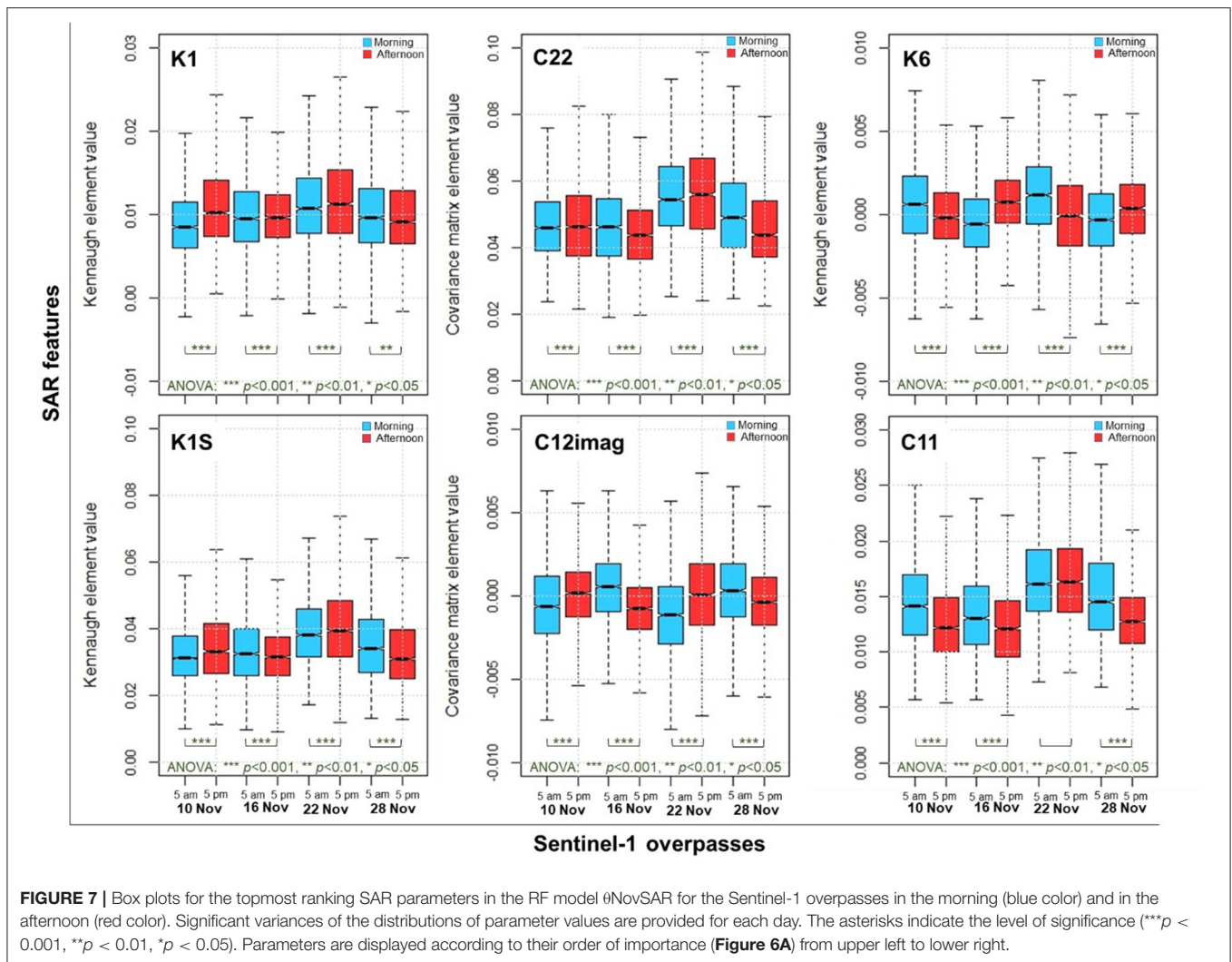
From Liao et al. (2018), who tested the performance of different polarimetric parameter sets from RADARSAT2 quad-pool data for cropland classification in humid southwestern Ontario, we see that specific elements of the coherency matrix performed best and largely contributed to an overall classification accuracy of 90%. Yet, only a little information is available in such context of θ estimation, with no available information for agroforestry systems in the Mediterranean. Hence, we assume our approach in which we implemented a similar S1 SAR-based predictor set for regression analysis and identified elements of the Kennaugh and covariance matrices to be essential as currently best feasible to retrieving θ in a high spatiotemporal resolution. However, against this background, we aim to omit such highly correlated parameters with respect to their parameter importance in future research to gain optimum parameters.

Transferability and Constraints of the Modeling Framework

Our modeling framework comprises a standardized survey scheme for measuring point-scale θ in accordance with S1 satellite overpasses and RF-based spatial regression analysis to derive a reasonable short time series of θ in November 2018 in a Mediterranean agroforestry environment.

At this point, it should be briefly mentioned that both the measuring instrument (Stevens HydraProbe) and the timing of the field observations in tune with the overpasses of S1 are directly comparable to several recently published studies (Ayehu et al., 2020; Datta et al., 2020; Han et al., 2020; Ma et al., 2020), and thus are regarded as reliable from a technical point of view.

Our framework considers the spatial heterogeneity of the terrain and time-variant backscatter and polarization variables from S1 SLC SAR C-band products (Tables 3, 4) and improves our knowledge of the spatiotemporal dynamics of θ and controlling parameters. Yet, our approach is capable of improvements, most importantly the increase in data availability in terms of higher spatial coverage and a more extended period to optimize the modeling scheme (e.g., testing on unseen data) and to enable more generalized statements, e.g., accurate differentiation of the meteorological forcing (i.e., rainfall) and its response on θ . Of course, this would require more calibration data at the 22 sensor locations considered (Figures 1B,C) and beyond to gather more spatially distributed information. Later is particularly true as zones outside MFC2 show remarkable differences between the SAR-only-based and the SAR+terrain-based RF models and between the SAR measurements for the S1 overpasses in the morning and afternoon. We remind that our empirical calibration models were trained over 22 specific locations at the precise time of eight satellite overpasses for a short observation period of 4 days only. As we trained



our site-specific RF model with *in-situ* data representing the wettest period in the year, we must assume that for the wet-to-dry transition zone, as well as for the dry period and dry-to-wet transition zone (Nasta et al., 2020), it will come up against its limits of what is meaningful. Thus, a final statement on the transferability of the model(s) is only qualitatively feasible at present when considering the same physio-geographic and hydrological settings (e.g., soil texture, slope, vegetation cover) and incidence angle (e.g., Walker et al., 2004). In this context, we consider our approach as best affordable at the given time.

We are aware that both soil moisture and the SAR backscatter are influenced by more factors than addressed in our study. For instance, the soil surface roughness was not considered, although it is reported a major limiting factor for active microwave soil moisture RS (e.g., Sahebi et al., 2002; Schuler et al., 2002; Walker et al., 2004; Wang et al., 2016; El-Shirbeny and Abutaleb, 2017). As no tillage and storm events, commonly known to largely influence the backscatter (e.g., Baghdadi et al., 2008; Dalla Rosa et al., 2012), were noticed and recorded, respectively, during our

short observation period (i.e., 18 days), we assumed that the surface roughness did not vary and ignored it, particularly as its field survey was considered not feasible.

Furthermore, the occurrence of clouds caught by the close-by mountain ridge (Romano et al., 2018) did not allow for embedding optical (e.g., Sentinel-2) vegetation indices as proxies on the moisture contained in the vegetation and the vegetation activity, such as the leaf area index, the (green) normalized difference vegetation and normalized difference water index, and the fractional vegetation cover (e.g., Attarzadeh et al., 2018; Han et al., 2020; Xu et al., 2020). Here, our model framework offers an open structure toward their gridded integration to account for the potential attenuation of the SAR signal caused by vegetation and enhanced representation of the landscape.

However, we consider our RF approach ($\theta_{Nov}^{SAR+Terrain}$ in particular) to compensate for numerous effects and return plausible results despite the neglected factors. This encourages us to continue this kind of empirical assessment by augmenting the *in-situ* sampling for capturing the entire annual course of hydrometeorological and vegetation conditions. Our achieved

TABLE 6 | Spearman's correlation coefficients ρ ($\alpha = 0.05$), of the topmost important SAR parameters represented by each three elements of the covariance and Kenough matrices.

Element	K1	C22	K6	K1S	C12imag	C11
K1	1.00	0.83	-0.09	0.93	0.09	0.24
C22	0.83	1.00	-0.14	0.97	0.14	0.70
K6	-0.09	-0.14	1.00	-0.12	-1.00	-0.05
K1S	0.93	0.97	-0.12	1.00	0.12	0.53
C12imag	0.09	0.14	-1.00	0.12	1.00	0.05
C11	0.24	0.70	-0.05	0.53	0.05	1.00

The structure of elements follows their parameter importance (from left to right) in the model θ_{Nov}^{SAR} .

results show that although the SAR backscatter is composed of different factors, our site-specific calibration-dependent empirical model is capable of disentangling such effects on the backscatter. Thus, our methodological approach is promising for using and optimizing simplified empirical models over universal and complex physically-based models to spatiotemporally estimate θ . This is particularly encouraging as practitioners might be interested in feasible (e.g., fast surveys for ground-truthing and set-up of predictor sets composed of meaningful parameters) and reliable approaches of high accuracy (Attarzadeh et al., 2018), but at the same time have a limited understanding of the complex theoretical physical properties (Ezzahar et al., 2019).

CONCLUSIONS AND OUTLOOK

We present a simplified, site-specific calibration model to explore the potentials of Sentinel-1 A/B dual-pol SAR-based phase and amplitude information and *in-situ* near-surface soil moisture (θ) measured at times of satellite overpasses for the empirical inversion of SAR measurements to gridded estimates of θ . The addition of hydrological meaningful terrain parameters supports capturing the spatial variability of θ by linking the topography at the immediate, surveyed locations to their spatial coherence at the field-scale. We thus not only spatially approximate θ from a technical perspective in terms of physical-numerical interactions of SAR backscatters with surface properties, but rather from a combined approach allowing an eco-hydrological perspective on the spatial patterns of θ . Beyond the predominantly statistical approach, we developed a short-term reasonable time series of θ for a Mediterranean agroforestry ecosystem in southern Italy for November 2018.

Although our proposed model gained high accuracy, we anticipate that its general information value and significance can be enhanced by incorporating more observational data and environmental covariates such as soil (e.g., soil texture and structure) and vegetation properties (e.g., leaf area index and canopy structure), and land use to enable for a more holistic approach. Here, our proposed ML model framework offers an open structure for integrating relevant gridded data. We consider our framework to provide impulses for improvements to address current constraints (i.e., a limited number of *in-situ* data in space

and time). Future research will focus on increasing the number of *in-situ* observations at times of satellite overpasses over the investigated locations and beyond to optimize the modeling and testing, hence enhancing the mapping accuracy. Subsequently, more robust statements on the spatial heterogeneity and its effect on θ pattern by less over- and underestimations are envisaged. By identifying relevant SAR measurements and topographic information to efficiently parameterize the empirical calibration model to the situation on-site, we attempted to reduce the predictor space toward an optimum of sensitive parameters. The reduction of the computational efforts (e.g., processing time and data storage capacity) is concluded to be of specific relevance for practitioners interested in reliable and feasible approaches to monitoring θ .

Eventually, we consider coupling our model framework and results to hydrological models to support the assessment of surface water fluxes and water balance components, e.g., in groundwater recharge modeling for deriving scenarios of water usage in the Mediterranean agroforestry ecosystems.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

SS-S, PN, and CC: manuscript design. SS-S: statistical analysis, design of figures and tables, and manuscript editing. PN and NR: regional and hydrological-hydrogeological expertise. MK: *in-situ* data survey. MK and PN: VCP and BSP sensor maintenance. SS-S, CC, PN, and NR: study design. SS-S, NA, and MK: data processing. HB and HV: provision of SoilNet monitoring infrastructure. All authors contributed to the article and approved the submitted version.

FUNDING

The Bavarian Research Alliance (grant no. BayIntAn_UWUE_2019_95) and the Italian Ministry of Education, University and Research MiUR (grant no. 2010JHF437) funded the research. The Helmholtz Association and the German Federal Ministry of Education and Research (BMBF) provided funding for the TERENO sensor network infrastructure. Financial support for the maintenance activities was provided by the MiUR-PRIN project. This publication was supported by the Open Access Publication Fund of the University of Würzburg.

ACKNOWLEDGMENTS

We acknowledge Luigi Giordano from Agriturismo Tre Morene in Monteforte Cilento (Salerno Province) and Benedetto Sica from the University of Naples for supporting the installation of the bare soil plot and for regularly cutting the grass cover.

REFERENCES

- Adeyemi, O., Grove, I., Peets, S., Domun, Y., and Norton, T. (2018). Dynamic neural network modelling of soil moisture content for predictive irrigation scheduling. *Sensors* 18:3408. doi: 10.3390/s18103408
- Ahmadian, N., Ullmann, T., Verelst, J., Borg, E., Zöllitz, R., and Conrad, C. (2019). Biomass assessment of agricultural crops using multi-temporal dual-polarimetric TerraSAR-X data. *PFG J. Photogramm. Remote Sens. Geoinf. Sci.* 87, 159–175. doi: 10.1007/s41064-019-00076-x
- Alemohammad, S. H., Jagdhuber, T., Moghaddam, M., and Entekhabi, D. (2019). Soil and vegetation scattering contributions in L-band and P-band polarimetric SAR observations. *IEEE Trans. Geosci. Remote Sens.* 57, 8417–8429. doi: 10.1109/TGRS.2019.2920995
- Amanabadi, S., Mohammadi, M. H., Masihabadi, M. H., and Vazirinia, M. (2018). Predicting continuous form of soil-water characteristics curve from limited particle size distribution data. *Water SA* 44, 428–435. doi: 10.4314/wsa.v44i3.10
- Amatulli, G., Domisch, S., Tuanmu, M.-N., Parmentier, B., Ranipeta, A., Malczyk, J., et al. (2018). A suite of global, cross-scale topographic variables for environmental and biodiversity modeling. *Sci. Data* 5:180040. doi: 10.1038/sdata.2018.40
- Asmuß, T., Bechtold, M., and Tiemeyer, B. (2019). On the potential of Sentinel-1 for high resolution monitoring of water table dynamics in grasslands on organic soils. *Remote Sens.* 11:1659. doi: 10.3390/rs11141659
- Attarzadeh, R., Amini, J., Notarnicola, C., and Greifeneder, F. (2018). Synergetic use of Sentinel-1 and Sentinel-2 data for soil moisture mapping at plot scale. *Remote Sens.* 10:1285. doi: 10.3390/rs10081285
- Ayehu, G., Tadesse, T., Gessesse, B., Yigrem, Y., and Melesse, A. M. (2020). Combined use of Sentinel-1 SAR and Landsat sensors for residual soil moisture retrieval over agricultural fields in the Upper Blue Nile Basin, Ethiopia. *Sensors* 20:3282. doi: 10.3390/s20113282
- Babaeian, E., Sadeghi, M., Jones, S. B., Montzka, C., Vereecken, H., and Tuller, M. (2019). Ground, proximal, and satellite remote sensing of soil moisture. *Rev. Geophys.* 57, 530–616. doi: 10.1029/2018RG000618
- Baghdadi, N., Cerdan, O., Zribi, M., Auzet, A. V., Darboux, F., and El Hajj, M. (2008). Operational performance of current synthetic aperture radar sensors in mapping soil surface characteristics in agricultural environments: application to hydrological and erosion modelling. *Hydrol. Process.* 22, 9–20. doi: 10.1002/hyp.6609
- Baghdadi, N., El Hajj, M., Zribi, M., and Bousbih, S. (2017). Calibration of the Water Cloud Model at C-band for winter crop fields and grasslands. *Remote Sens.* 9:969. doi: 10.3390/rs9090969
- Baghdadi, N., Holah, N., and Zribi, M. (2006). Calibration of the Integral Equation Model for SAR data in C-band and HH and VV polarizations. *Int. J. Remote Sens.* 27, 805–816. doi: 10.1080/01431160500212278
- Barrett, B., Dwyer, E., and Whelan, P. (2009). Soil moisture retrieval from active spaceborne microwave observations: an evaluation of current techniques. *Remote Sens.* 1, 210–242. doi: 10.3390/rs1030210
- Beaudette, D. E., Dahlgren, R. A., and O'Geen, A. T. (2013). Terrain-shape indices for modeling soil moisture dynamics. *Soil Sci. Soc. Am. J.* 77, 1696–1710. doi: 10.2136/sssaj2013.02.0048
- Behrens, T., and Viscarra Rossel, R. A. (2020). On the interpretability of predictors in spatial data science: the information horizon. *Sci. Rep.* 10:16737. doi: 10.1038/s41598-020-73773-y
- Bellingham, K. (2019). *Soil Geomorphology. A Pedological Guide for Soil Sensor Applications*. Stevens Appl. Note. Portland, OR: Stevens Water Monitoring Systems, Inc., 25.
- Bogaart, P. W., and Troch, P. A. (2006). Curvature distribution within hillslopes and catchments and its effect on the hydrological response. *Hydrol. Earth Syst. Sci.* 10, 925–936. doi: 10.5194/hess-10-925-2006
- Bogena, H. R., Herbst, M., Huisman, J. A., Rosenbaum, U., Weuthen, A., and Vereecken, H. (2010). Potential of wireless sensor networks for measuring soil water content variability. *Vadose Zo. J.* 9, 1002–1013. doi: 10.2136/vzj2009.0173
- Böhner, J., and Selige, T. (2006). "Spatial prediction of soil attributes using terrain analysis and climate regionalisation," in *SAGA-Analyses and Modelling Applications*, eds J. Böhner, K. R. McCloy, and J. Strobl (Göttingen: Göttinger Geographische Abhandlungen), 13–28.
- Breiman, L. (2001). Random forest. *Mach. Learn.* 45, 5–32. doi: 10.1023/A:1010933404324
- Cai, Y., Zheng, W., Zhang, X., Zhangzhong, L., and Xue, X. (2019). Research on soil moisture prediction model based on deep learning. *PLoS ONE* 14:e0214508. doi: 10.1371/journal.pone.0214508
- Chakrabarti, S., Bongiovanni, T., Judge, J., Nagarajan, K., and Principe, J. C. (2015). Downscaling satellite-based soil moisture in heterogeneous regions using high-resolution remote sensing products and information theory: a synthetic study. *IEEE Trans. Geosci. Remote Sens.* 53, 85–101. doi: 10.1109/TGRS.2014.2318699
- Chen, H., Fan, L., Wu, W., and Liu, H.-B. (2017). Comparison of spatial interpolation methods for soil moisture and its application for monitoring drought. *Environ. Monit. Assess.* 189:525. doi: 10.1007/s10661-017-6244-4
- Chen, Z., and Wang, J. (2008). A new method for minimizing topographic effects on RADARSAT-1 images: an application in mapping human settlements in the mountainous Three Gorges Area, China. *Can. J. Remote Sens.* 34, 13–25. doi: 10.5589/m08-005
- Conrad, C., Bechtel, B., Bock, M., Dietrich, H., Fischer, E., Gerlitz, L., et al. (2015). System for automated geoscientific analyses (SAGA) v. 2.1.4. *Geosci. Model Dev.* 8, 2271–2312. doi: 10.5194/gmdd-8-2271-2015-supplement
- Contador, J. F. L., Maneta, M., and Schnabel, S. (2006). Prediction of near-surface soil moisture at large scale by digital terrain modeling and neural networks. *Environ. Monit. Assess.* 121, 213–232. doi: 10.1007/s10661-005-9116-2
- Dabrowska-Zielinska, K., Musial, J., Malinska, A., Budzynska, M., Gurdak, R., Kiryla, W., et al. (2018). Soil moisture in the Biebrza wetlands retrieved from Sentinel-1 imagery. *Remote Sens.* 10:1979. doi: 10.3390/rs10121979
- Dalla Rosa, J., Cooper, M., Darboux, F., and Medeiros, J. C. (2012). Soil roughness evolution in different tillage systems under simulated rainfall using a semivariogram-based index. *Soil Tillage Res.* 124, 226–232. doi: 10.1016/j.still.2012.06.001
- Das, K., and Paul, P. K. (2015). Present status of soil moisture estimation by microwave remote sensing. *Cogent Geosci.* 1:1084669. doi: 10.1080/23312041.2015.1084669
- Datta, S., Das, P., Dutta, D., and Giri, R., Kr. (2020). Estimation of surface moisture content using Sentinel-1 C-band SAR data through machine learning models. *J. Indian Soc. Remote Sens.* 49, 887–896. doi: 10.1007/s12524-020-01261-x
- Dorigo, W., Wagner, W., Albergel, C., Albrecht, F., Balsamo, C., Brocca, L., et al. (2017). ESA CCI soil moisture for improved earth system understanding: state-of-the-art and future directions. *Remote Sens. Environ.* 203, 185–215. doi: 10.1016/j.rse.2017.07.001
- Efremova, N., Zausaev, D., and Antipov, G. (2019). Prediction of soil moisture content based on satellite data and Sequence-to-Sequence networks. *arXiv:1907.03697v1* 5.
- El Hajj, M., Baghdadi, N., Bazzi, H., and Zribi, M. (2018). Penetration analysis of SAR signals in the C and L bands for wheat, maize, and grasslands. *Remote Sens.* 11:31. doi: 10.3390/rs11010031
- El-Shirbeny, M. A., and Abutaleb, K. (2017). Sentinel-1 radar data assessment to estimate crop water stress. *World J. Eng. Technol.* 05, 47–55. doi: 10.4236/wjet.2017.52B006
- Erten, E. (2012). The performance analysis based on SAR sample covariance matrix. *Sensors* 12, 2766–2786. doi: 10.3390/s120302766
- Ezzahar, J., Ouadi, N., Zribi, M., Elfarkh, J., Aouade, G., Khabba, S., et al. (2019). Evaluation of backscattering models and support vector machine for the retrieval of bare soil moisture from Sentinel-1 data. *Remote Sens.* 12:72. doi: 10.3390/rs12010072
- Filippini, F. (2019). Sentinel-1 GRD preprocessing workflow. *Proceedings* 18:11. doi: 10.3390/ECRS-3-06201
- Foucras, M., Zribi, M., Albergel, C., Baghdadi, N., Calvet, J.-C., and Pellarin, T. (2020). Estimating 500-m resolution soil moisture using Sentinel-1 and optical data synergy. *Water* 12:866. doi: 10.3390/w12030866
- García-Ruiz, J. M., López-Moreno, J. I., Vicente-Serrano, S. M., Lasanta-Martínez, T., and Beguería, S. (2011). Mediterranean water resources in a global change scenario. *Earth Sci. Rev.* 105, 121–139. doi: 10.1016/j.earscirev.2011.01.006
- Gauquelin, T., Michon, G., Joffre, R., Duponnois, R., Génin, D., Fady, B., et al. (2018). Mediterranean forests, land use and climate change: a social-ecological perspective. *Reg. Environ. Change* 18, 623–636. doi: 10.1007/s10113-016-0994-3
- Geroy, I. J., Gribb, M. M., Marshall, H. P., Chandler, D. G., Benner, S. G., and McNamara, J. P. (2011). Aspect influences on soil water retention and storage. *Hydrol. Process.* 25, 3836–3842. doi: 10.1002/hyp.8281
- Gorrab, A., Zribi, M., Baghdadi, N., Lili-Chabaane, Z., and Mougnot, B. (2014). "Multi-frequency analysis of soil moisture vertical heterogeneity effect on radar

- backscatter,” in *2014 1st International Conference on Advanced Technologies for Signal and Image Processing (ATSIP)* (Sousse), 379–384.
- Gu, W., Xu, H., and Tsang, L. (2019). A numerical Kirchhoff simulator for GNSS-R land applications. *Prog. Electromagn. Res.* 164, 119–133. doi: 10.2528/PIER18121803
- Guderle, M., and Hildebrandt, A. (2015). Using measured soil water contents to estimate evapotranspiration and root water uptake profiles—a comparative study. *Hydrol. Earth Syst. Sci.* 19, 409–425. doi: 10.5194/hess-19-409-2015
- Guevara, M., and Vargas, R. (2019). Downscaling satellite soil moisture using geomorphometry and machine learning. *PLoS ONE* 14:e0219639. doi: 10.1371/journal.pone.0219639
- Hachani, A., Ouassar, M., Paloscia, S., Santi, E., and Pettinato, S. (2019). Soil moisture retrieval from Sentinel-1 acquisitions in an arid environment in Tunisia: application of Artificial Neural Networks techniques. *Int. J. Remote Sens.* 40, 9159–9180. doi: 10.1080/01431161.2019.1629503
- Hajdu, I., Yule, I., and Dehghan-Shear, M. H. (2018). “Modelling of near-surface soil moisture using machine learning and multi-temporal Sentinel 1 images in New Zealand,” in *IGARSS 2018–2018 IEEE International Geoscience and Remote Sensing Symposium* (Valencia), 1422–1425.
- Han, Y., Bai, X., Shao, W., and Wang, J. (2020). Retrieval of soil moisture by integrating Sentinel-1A and MODIS data over agricultural fields. *Water* 12:1726. doi: 10.3390/w12061726
- Jiménez-de-Santiago, D. E., Lidón, A., and Bosch-Serra, A. D. (2019). Soil water dynamics in a rainfed Mediterranean agricultural system. *Water* 11:799. doi: 10.3390/w11040799
- Kammerer, G., Nolz, R., Rodny, M., and Loiskandl, W. (2014). Performance of Hydra Probe and MPS-1 soil water sensors in topsoil tested in lab and field. *J. Water Resour. Prot.* 6, 1207–1219. doi: 10.4236/jwarp.2014.613110
- Karpatne, A., Ebert-Uphoff, I., Ravela, S., Babaie, H. A., and Kumar, V. (2019). Machine learning for the geosciences: challenges and opportunities. *IEEE Trans. Knowl. Data Eng.* 31, 1544–1554. doi: 10.1109/TKDE.2018.2861006
- Kuhn, M., Wing, J., Weston, S., Mayer, Z., R Core Team, Benesty, M., et al. (2020). *Package “caret”: Classification and Regression Training*. Version 6.0–86. Available online at: <https://CRAN.Rproject.org/package=caret>
- Liao, C., Wang, J., Huang, X., and Shang, J. (2018). Contribution of minimum noise fraction transformation of multi-temporal RADARSAT-2 polarimetric SAR data to cropland classification. *Can. J. Remote Sens.* 44, 15–231. doi: 10.1080/07038992.2018.1481737
- Ma, C., Li, X., and McCabe, M. F. (2020). Retrieval of high-resolution soil moisture through combination of Sentinel-1 and Sentinel-2 data. *Remote Sens.* 12:2303. doi: 10.3390/rs12142303
- Mandal, D., Vaka, D. S., Bhogapurapu, N. R., Vanama, V. S. K., Kumar, V., Rao, Y. S., et al. (2019). Sentinel-1 SLC preprocessing workflow for polarimetric applications: a generic practice for generating dual-pol covariance matrix elements in SNAP S-1 Toolbox. *Preprints* 2019110393. doi: 10.20944/preprints201911.0393.v1
- Mascaro, G., Ko, A., and Vivoni, E. R. (2019). Closing the loop of satellite soil moisture estimation via scale invariance of hydrologic simulations. *Sci. Rep.* 9:16123. doi: 10.1038/s41598-019-52650-3
- MirMazloumi, S., and Sahebi, M. R. (2016). Assessment of different backscattering models for bare soil surface parameters estimation from SAR data in band C, L, and P. *Eur. J. Remote Sens.* 49, 261–278. doi: 10.5721/EuJRS20164915
- Mladenova, I. E., Jackson, T. J., Bindlish, R., and Hensley, S. (2013). Incidence angle normalization of radar backscatter data. *IEEE Trans. Geosci. Remote Sens.* 51, 1791–1804. doi: 10.1109/TGRS.2012.2205264
- Mohanty, B. P., Cosh, M. H., Lakshmi, V., and Montzka, C. (2017). Soil moisture remote sensing: state-of-the-science. *Vadose Zone J.* 16, 1–9. doi: 10.2136/vzj2016.10.0105
- Moreira, A., Prats-Iraola, P., Younis, M., Krieger, G., Hajnsek, I., and Papathanassiou, K. P. (2013). A tutorial on synthetic aperture radar. *IEEE Geosci. Remote Sens. Mag.* 1, 6–43. doi: 10.1109/MGRS.2013.2248301
- Mulder, V. L., de Bruin, S., Schaepman, M. E., and Mayr, T. R. (2011). The use of remote sensing in soil and terrain mapping—a review. *Geoderma* 162, 1–19. doi: 10.1016/j.geoderma.2010.12.018
- Nasta, P., Bogena, H. R., Sica, B., Weuthen, A., Vereecken, H., and Romano, N. (2020). Integrating invasive and non-invasive monitoring sensors to detect field-scale soil hydrological behavior. *Front. Water* 2:26. doi: 10.3389/frwa.2020.00026
- Nasta, P., Palladino, M., Ursino, N., Saracino, A., Sommella, A., and Romano, N. (2017). Assessing long-term impact of land-use change on hydrological ecosystem functions in a Mediterranean upland agro-forestry catchment. *Sci. Total Environ.* 605–606, 1070–1082. doi: 10.1016/j.scitotenv.2017.06.008
- Nasta, P., Schönbrodt-Stitt, S., Kurtenbach, M., Ahmadian, N., Bogena, H. R., Vereecken, H., et al. (2019). “Integrating ground-based and remote sensing-based monitoring of near-surface soil moisture in a Mediterranean environment,” in *2019 IEEE International Workshop on Metrology for Agriculture and Forestry (MetroAgriFor)* (Portici), 274–279.
- Nielsen, A. A., Canty, M. J., Skriver, H., and Conradsen, K. (2017). “Change detection in multi-temporal dual polarization Sentinel-1 data,” in *2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)* (Fort Worth, TX), 3901–3908.
- Nobre, A. D., Cuartas, L. A., Hodnett, M., Rennó, C. D., Rodrigues, G., Silveira, A., et al. (2011). Height above the nearest drainage—a hydrologically relevant new terrain model. *J. Hydrol.* 404, 13–29. doi: 10.1016/j.jhydrol.2011.03.051
- Oh, Y., Sarabandi, K., and Ulaby, F. T. (1992). An empirical model and an inversion technique for radar scattering from bare soil surfaces. *IEEE Trans. Geosci. Remote Sens.* 30, 370–381. doi: 10.1109/36.134086
- Omar, H., Mismam, M. A., and Kassim, A. R. (2017). Synergetic of PALSAR-2 and Sentinel-1A SAR polarimetry for retrieving aboveground biomass in Dipterocarp Forest of Malaysia. *Appl. Sci.* 7:675. doi: 10.3390/app7070675
- Paloscia, S., Pettinato, S., Santi, E., Notarnicola, C., Pasolli, L., and Reppucci, A. (2013). Soil moisture mapping using Sentinel-1 images: algorithm and preliminary validation. *Remote Sens. Environ.* 134, 234–248. doi: 10.1016/j.rse.2013.02.027
- Petropoulos, G. P., Ireland, G., and Barrett, B. (2015). Surface soil moisture retrievals from remote sensing: Current status, products and future trends. *Phys. Chem. Earth* 83–84, 36–56. doi: 10.1016/j.pce.2015.02.009
- Qin, C.-Z., Zhu, A. X., Pei, T., Li, B. L., Scholten, T., Behrens, T., et al. (2011). An approach to computing topographic wetness index based on maximum downslope gradient. *Precis. Agric.* 12, 32–43. doi: 10.1007/s11119-009-9152-y
- Quesney, A., Le Hégarat-Masclé, S., Taconet, O., Vidal-Madjar, D., Wigneron, J. P., Loumagne, C., et al. (2000). Estimation of watershed soil moisture index from ERS/SAR data. *Remote Sens. Environ.* 72, 290–303. doi: 10.1016/S0034-4257(99)00102-9
- Rahimzadeh-Bajgirani, P., Berg, A. A., Champagne, C., and Omasa, K. (2013). Estimation of soil moisture using optical/thermal infrared remote sensing in the Canadian Prairies. *ISPRS J. Photogramm. Remote Sens.* 83, 94–103. doi: 10.1016/j.isprsjprs.2013.06.004
- Riley, S. J., DeGloria, S. D., and Elliot, R. (1999). A terrain ruggedness index that quantifies topographic heterogeneity. *Interm. J. Sci.* 5, 23–27.
- Robinson, D. A., Campbell, C. S., Hopmans, J. W., Hornbuckle, B. K., Jones, S. B., Knight, R., et al. (2008). Soil moisture measurement for ecological and hydrological watershed-scale observatories: a review. *Vadose Zone J.* 7, 358–389. doi: 10.2136/vzj2007.0143
- Romano, N., Nasta, P., Bogena, H., De Vita, P., Stellato, L., and Vereecken, H. (2018). Monitoring hydrological processes for land and water resources management in a Mediterranean ecosystem: the Alento River catchment observatory. *Vadose Zone J.* 17:180042. doi: 10.2136/vzj2018.03.0042
- Rowlandson, T. L., Berg, A. A., Bullock, P. R., Hanis-Gervais, K., Ojo, E. R., Cosh, M. H., et al. (2018). Temporal transferability of soil moisture calibration equations. *J. Hydrol.* 56, 349–358. doi: 10.1016/j.jhydrol.2017.11.023
- Sahebi, M. R., Angles, J., and Bonn, F. (2002). A comparison of multi-polarization and multi-angular approaches for estimating bare soil surface roughness from spaceborne radar data. *Can. J. Remote Sens.* 28, 641–652. doi: 10.5589/m02-060
- Santi, E., Paloscia, S., Pettinato, S., and Fontanelli, G. (2016). Application of artificial neural networks for the soil moisture retrieval from active and passive microwave spaceborne sensors. *Int. J. Appl. Earth Obs. Geoinf.* 48, 61–73. doi: 10.1016/j.jag.2015.08.002
- Santoro, A., Venturi, M., Bertani, R., and Agnoletti, M. (2020). A review of the role of forests and agroforestry systems in the FAO Globally Important Agricultural Heritage Systems (GIAHS) programme. *Forests* 11:860. doi: 10.3390/f11080860
- Schmitt, A., and Brisco, B. (2013). Wetland monitoring using the Curvelet-based change detection method on polarimetric SAR imagery. *Water* 5, 1036–1051. doi: 10.3390/w5031036

- Schmitt, A., Wendleder, A., and Hinz, S. (2015). The Kennaugh element framework for multi-scale, multi-polarized, multi-temporal and multi-frequency SAR image preparation. *ISPRS J. Photogramm. Remote Sens.* 102, 122–139. doi: 10.1016/j.isprsjprs.2015.01.007
- Schmugge, T., and Jackson, T. (1997). “Passive microwave remote sensing of soil moisture,” in *Land Surface Processes in Hydrology*, eds S. Sorooshian, H. V. Gupta, and J. C. Rodda (Berlin, Heidelberg: Springer Berlin Heidelberg), 239–262.
- Schuler, D. L., Lee, J.-S., Kasilingam, D., and Nesti, G. (2002). Surface roughness and slope measurements using polarimetric SAR data. *IEEE Trans. Geosci. Remote Sens.* 40, 687–698. doi: 10.1109/TGRS.2002.1000328
- 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 Sci. Rev.* 99, 125–161. doi: 10.1016/j.earscirev.2010.02.004
- Sinha, S., Jegannathan, C., Sharma, L. K., and Nathawat, M. S. (2015). A review of radar remote sensing for biomass estimation. *Int. J. Environ. Sci. Technol.* 12, 1779–1792. doi: 10.1007/s13762-015-0750-0
- Small, D. (2011). Flattening Gamma: radiometric terrain correction for SAR imagery. *IEEE Trans. Geosci. Remote Sens.* 49, 3081–3093. doi: 10.1109/TGRS.2011.2120616
- Ullmann, T., Banks, S. N., Schmitt, A., and Jagdhuber, T. (2017). Scattering characteristics of X-, C- and L-band PolSAR data examined for the Tundra environment of the Tuktoyaktuk Peninsula, Canada. *Appl. Sci.* 7:595. doi: 10.3390/app7060595
- Walker, J. P., Houser, P. R., and Willgoose, G. R. (2004). Active microwave remote sensing for soil moisture measurement: a field evaluation using ERS-2. *Hydrol. Process.* 18, 1975–1997. doi: 10.1002/hyp.1343
- Wang, H., Magagi, R., Goita, K., Jagdhuber, T., and Hajnsek, I. (2016). Evaluation of simplified polarimetric decomposition for soil moisture retrieval over vegetated agricultural fields. *Remote Sens.* 8:142. doi: 10.3390/rs8020142
- Wang, L., and Qu, J. J. (2009). Satellite remote sensing applications for surface soil moisture monitoring: a review. *Front. Earth Sci. China* 3, 237–247. doi: 10.1007/s11707-009-0023-7
- Western, A. W., Grayson, R. B., and Blöschl, G. (2002). Scaling of soil moisture: a hydrologic perspective. *Annu. Rev. Earth Planet. Sci.* 30, 149–180. doi: 10.1146/annurev.earth.30.091201.140434
- Wigneron, J.-P., Calvet, J.-C., Pellarin, T., Van de Griend, A., Berger, M., and Ferrazzoli, P. (2003). Retrieving near-surface soil moisture from microwave radiometric observations: current status and future plans. *Remote Sens. Environ.* 85, 489–506. doi: 10.1016/S0034-4257(03)00051-8
- Xu, C., Qu, J. J., Hao, X., and Wu, D. (2020). Monitoring surface soil moisture content over the vegetated area by integrating optical and SAR satellite observations in the permafrost region of the Tibetan Plateau. *Remote Sens.* 12, 183. doi: 10.3390/rs12010183
- Yommy, A. S., Liu, R., and Wu, A. S. (2015). “SAR image despeckling using Refined Lee Filter,” in *2015 7th International Conference on Intelligent Human-Machine Systems and Cybernetics (Hangzhou)*, 260–265.
- Zappa, L., Forkel, M., Xaver, A., and Dorigo, W. (2019). Deriving field scale soil moisture from satellite observations and ground measurements in a hilly agricultural region. *Remote Sens.* 11:2596. doi: 10.3390/rs11222596
- Zhang, D., and Zhou, G. (2016). Estimation of soil moisture from optical and thermal remote sensing: a review. *Sensors* 16:1308. doi: 10.3390/s16081308
- Zhang, G., Liu, F., and Song, X. (2017). Recent progress and future prospect of digital soil mapping: a review. *J. Integr. Agric.* 16, 2871–2885. doi: 10.1016/S2095-3119(17)61762-3
- Zribi, M., Muddu, S., Bousbih, S., Al Bitar, A., Tomer, S. K., Baghdadi, N., et al. (2019). Analysis of L-band SAR data for soil moisture estimations over agricultural areas in the tropics. *Remote Sens.* 11:1122. doi: 10.3390/rs11091122

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Schönbrodt-Stitt, Ahmadian, Kurtenbach, Conrad, Romano, Bogena, Vereecken and Nasta. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.