Reconstruction of root zone soil moisture according to the data from passive microwave radiometer and machine learning in the arid steppe region of Southern Western Siberia

Andrei A. Bondarovich	Altai State University, 61 Lenin Ave., Barnaul, 656049, Russia
Egor Yu. Mordvin	Altai State University, 61 Lenin Ave., Barnaul, 656049, Russia
Nikita M. Pochyomin	Altai State University, 61 Lenin Ave., Barnaul, 656049, Russia
Anatoly A. Lagutin	Altai State University, 61 Lenin Ave., Barnaul, 656049, Russia

Our study focuses on reconstruction root zone soil moisture (RZSM) in the Kulunda plain, a representative dry steppe area in southern Western Siberia, using remote sensing data (RSD) and machine learning techniques. We employed modern machine learning methods with soil surface layer moisture data from the AMSR2 passive microwave radiometer as the primary predictor. Additionally, we incorporated data from local meteorological and soil hydrological stations, as well as gravity lysimeter data for 2015–2017. This choice of predictors was based on the extensive time series of continuous observations and the availability of selected meteorological parameters. Among the machine learning models we evaluated, Random Forest (RF) and Extreme Gradient Boosting (XGW) yielded the best results, achieving statistical metrics of R-squared (R2) values of 0.96 and 0.94, respectively, with corresponding root mean square error (RMSE) values of 0.34 and 0.41.

Acta Biologica Sibirica 9: 805–829 (2023) doi: 10.5281/zenodo.10061576

Corresponding author: Andrei A. Bondarovich (a9130262571@gmail.com)

Academic editor: A. Matsyura | Received 20 September 2023 | Accepted 9 October 2023 | Published 4 November 2023

http://zoobank.org/63CE11DD-EBD8-4A10-A3F0-0043B0BFA1F1

Citation: Bondarovich AA, Mordvin EYu, Pochyomin NM, Lagutin AA (2023) Reconstruction of root zone soil moisture according to the data from passive microwave radiometer and machine learning in the arid steppe region of Southern Western Siberia. Acta Biologica Sibirica 9: 805–829. https://doi.org/10.5281/zenodo.10061576

Keywords

Soil moisture, root-zone, machine learning, model, Kulunda plain, Western Siberia

Introduction

Soil moisture (SM) profoundly influences water, energy, greenhouse gases, and climate cycles (Seneviratne et al. 2010; Dong et al. 2016; Mao et al. 2020; Luo et al. 2020; Humphrey et al. 2021). Given the increasing scarcity of freshwater resources, the concept of green water, encompassing atmospheric precipitation, evaporation, and root-zone soil moisture (RZSM), has gained significance (Lai et al. 2022; Wang-Erlandsson et al. 2022; Huang et al. 2023). Defining RZSM, particularly the soil horizons involved, is essential. In this study, we focus on RZSM depths between 25 and 35 cm.

Estimating RZSM is particularly crucial in arid regions, which make up 40% of the world's land surface and are prone to soil degradation (FAO, 2000). This degradation affects 34% of all agricultural lands globally, with 733 million ha in drylands. These regions also experience high levels of "water stress" (35-70%). In light of climate and land use changes, investigating greenhouse gas emissions, soil moisture use efficiency, evapotranspiration, drought, and crop yields is essential for arid areas (Changming et al. 2002; Robinson et al. 2008; Bodner et al. 2015, Fisher et al. 2017; Zhang et al. 2018; Lal 2020; Liu et al. 2022; Afshar et al. 2022; Chai et al. 2023; Yue et al. 2023; Kamran et al. 2023; Zhang et al. 2023). Many of the above challenges are related to the introduction of conservation agriculture (CA), among which No-Tillage (hereafter NT) and Mini-Tillage (hereafter MT) are the best known (Kassam et al. 2019).

Various approaches for monitoring surface soil moisture (SSM) and RZSM include ground observations, remote sensing inversion, and data assimilation methods (Margulis et al. 1997; Walker et al. 2001; Drusch 2007; Albergel et al. 2008; Draper et al. 2009; Wyatt et al. 2021; Qin et al. 2022; Li et al. 2022; Travova 2022; Shahi et al. 2023). While direct soil moisture measurements are accurate, they have limitations (Seneviratne et al. 2010; Babaeian et al. 2019; Li et al. 2022). Groundbased observations encompass labor-intensive gravity measurements, handheld devices, and automatic observation stations (Robinson et al. 2008; Seneviratne et al. 2010; Zhang et al. 2016; Dong et al. 2016; Li, Zhang 2021; Hodges et al. 2023).

The world's largest platform for collecting and publishing soil moisture measurements from more than 2800 sensors worldwide is the International Soil Moisture Network (hereafter ISMN, Dorigo et al. 2021).

Remote sensing provides continuous spatial coverage, primarily for SSM observations (Remote ... 2013; Babaeian et al. 2019; Travova 2022). Passive microwave radiometers, such as AMSR-E and AMSR-2 (AMSR2), hold promise for regionalscale SSM estimation with high temporal coverage (Wigneron et al. 2003; Travova 2022). However, extracting RZSM from remote observations is non-trivial due to the nonlinear relationship between SSM and RZSM (Babaeian et al. 2019, Cheng et al. 2022). Various methods have attempted to link satellite SSM to RZSM, including multiple regression, exponential filters, data assimilation, and artificial intelligence, AI (Mahmood et al. 2012; Pal et al. 2016; Baldwin et al. 2017; Li 2021; Stefan et al. 2021; Tian et al. 2022; Yu et al. 2022; Celik et al. 2022; Yinglan et al. 2023).

While data assimilation is efficient, it is computationally expensive. Exponential filters have limitations, especially during drought and with increasing soil depth. Machine learning (ML) algorithms, such as Random Forest (RF), support vector machine (SVM), k Nearest Neighbors (kNN), and Boosting, offer promise for modeling soil moisture at different depths.

In Russia, only 2% of the land has optimal moisture conditions, with over 78% of arable land concentrated in arid zones (State... 2018). The Kulunda plain, in southern Western Siberia, is a prime example. Conservation technologies like No-Till and Mini-Till are being introduced to combat soil degradation, accompanied by a monitoring network studying meteorological and soil hydrological parameters (Kulunda 2020).

Our study's objective is to reconstruct RZSM at a 30 cm depth using satellite SSM data, machine learning models, and additional regressors like air temperature, humidity, and atmospheric precipitation from local monitoring networks and Roshydromet's open data.

Materials and methods

The continental climate of the Kulunda steppe is characterized by a long, cold, lowsnow winter and a short but hot summer. The average annual temperature is about 0 °C, the average temperature in January is -19 °C, the absolute minimum is -47 °C, the average temperature in July is +19 °C, and the absolute maximum is +40 °C. The frost-free period lasts 112–120 days a year from May 15-25 to September 10-15. Annual precipitation is about 250-450 mm and about 200 mm in April–October. The region represents the eastern edge of the Dfb Köppen climate (snow/humidity/ warm summer) and is quite humid compared to the drainless areas further south in the Republic of Kazakhstan (Bsk Köppen). The soil cover of the dry steppes of Kulunda is represented by chestnut, meadow-chestnut, meadow soils, solonts, and solonchaks of different degree of hydromorphism. Chestnut soils differ considerably in granulometric composition as a result of the ancient limnic and aeolian genesis of the territory. Loamy loam (15–19 % clay, 11–20 % silt, 65–70 % sand) prevails, the humus content (2–4 %) is relatively high (Balykin et al. 2016). Of the total steppe area, only 2.8 million ha can be attributed to fallow lands and natural grasslands (Wesche et al 2020). Taking into account climatic changes, especially the increasing aridity found in western Siberia (Kharlamova 2020), the adaptation of conservation tillage (CA) technologies that focus on greenhouse gas sequestration and improving the water regime is important here (Kulunda 2020).

We used data from weather station (hereafter WS), two lysimetric station (hereinafter Lys 1 and Lys 1) to study the water regime under different farming technologies, in natural and agricultural cenoses and the data from two soil hydrological stations to study "No-Tillage" (hereinafter NT) and Conventional Tillage (hereinafter CT) in the dry steppe subzone since 2013. Our model plots was in the farm "Partner" located in the central part of the Kulunda steppe (village Poluyamki, Mikhailovsky district, Altai Krai, Russia) – see Fig. 1.

Figure 1. *Location map of the stations for measuring meteorological and soil hydrological parameters in the dry steppe subzone of southern western Siberia (Kulunda steppe, Altai Krai, Poluyamki village). Google Earth substrate. WS and lysimetric station coordinates: E79°42.786' N52°03.959'; CT: N52°04.180 E79°54.014; NT: N52°04.128 E79°54.006.*

WS is equipped with pyranometer CMP3 (Kipp&Sonnen, Netherlands): solar radiation (R), W/m^2 and multisensor WXT 536 (Vaisala, Finland): temperature (T)°C and of the air humidity (φ) %, liquid precipitation (P) mm, air pressure (p) hPa, wind speed (V) m/s and direction (degrees) – see Fig. 2. Soil stations measure soil moisture (%) and temperature (°C) using Hydra Probe (Stevens, USA) TDR sensors and osmotic pressure (pF) at depths of 30 cm, 60 cm and 120 cm in automatic mode (Fig. 3).

The common methods among dielectric classes are time domain reflectometry (TDR), frequency domain reflectometry (FDR), amplitude domain reflectometry (ADR), time domain transmission (TDT), and electrical capacitance volume tomography (ECVT). Sensors that measure SM using Time Domain Reflectometry (TDR) principle are the most accurate (error 1-2 %) (Mukhlisin et al. 2021). A study of the calibration accuracy of eight commercially available (TDR) sensors showed (we highlight only the two most used) that higher accuracy was demonstrated by the Theta Probe product (Delta-T Devices, UK) compared to the Hydra Probe sensor (Stevens, USA) (Vaz et al. 2013). In another field experiment in two types of soils, these sensors also showed different accuracy but were evaluated as the most suitable (even without calibration to local conditions) for organizing a monitoring network for SM content (Jabro et al. 2018).

Figure 2. *Meteorological station in Poluyamki village (photo by Bondarovich).*

Figure 3. *Installation of the "Hydra Probe" (Stevens, USA) of the soil hydrological station (photo by Bondarovich).*

During June to August 2013, a containerized weighable lysimeter station (Polyethylene PE-HD; manufacturer ‘UGT Muencheberg’, BRD) was installed at the ‘Partner’ test farm of the KULUNDA project in Poluyamki. The soilmonoliths (surface area of 1m^2, 2 m depth) extracted from both an arable land site (furthermore Lys 1) and a fallow site which was plowed once in the 1950s, but since then covered with natural steppe vegetation (Bunch-grass steppes, dominant species – *Stipapennata*) (furthermore Lys 2 Figs 4, 5). The lysimetric station performs actual evapotranspiration (ETa) and soil solution measurements using precision scales, as well as using Watermark TDR sensors (Spectrum Technologies, USA), moisture (%) and temperature (°C) at depths of 30, 50 and 120 cm (Balykin et al. 2016; Bondarovich et al. 2020).

The soil in the test sector is identified as dark chestnut soils (WRB classification – Haplic, FAO classification – Kastanozem) (Balykin et al. 2016; Mizgirev et al. 2020). In Lys 1 the following crop rotation in 2013 – *Triticum aestivum* L., in 2014 – *Písum*, in 2015 – *Triticum aestivum* L., in 2016–2018 – fallow (Fig. 6).

To characterize the moisture content of vegetation periods at the time of field experiments, we give the hydrothermal coefficient (HTC) of Selyaninov (Hydrothermal Coefficient 2023). We see that this period 2013-2018 can be characterized as "quite wet" with values ranging from 0.9 to 1.3 (see Table 1). Consequently, the use of these data in the training sample of the model does not make it possible to predict the subsequent dry growing seasons of 2019–2020, which were also recorded in the neighboring Republic of Kazakhstan (Abdrakhmanova and Filippova 2022).

A more detailed description of the monitoring network, test soil profiles, results on soil temperature and moisture data processing, and evapotranspiration have been published in a number of papers (Balykin et al. 2016; Bondarovich et al. 2020; Meissner et al. 2020; Bondarovich et al. 2021). In the field experiment, due to soil tillage (in CT variant up to 24 cm, in NT variant up to 5 cm), it was not possible to install sensors for SSM measurement (1-5 cm). Based on the objectives of the experiment at Lys, identical horizons were selected for observation, and SSM was also not observed. Thus, continuous SSM data are not available. To obtain SSM data within the test farm, express measurements were made from 5 to 100 cm, every 10 cm, but these were periodic and of little use for statistics. In the context of this study, we were only interested in ground-based RZSM data at 30 cm depth from Lys 1 and Lys 2, CT and NT, and a set of WS meteorological parameters for model training.

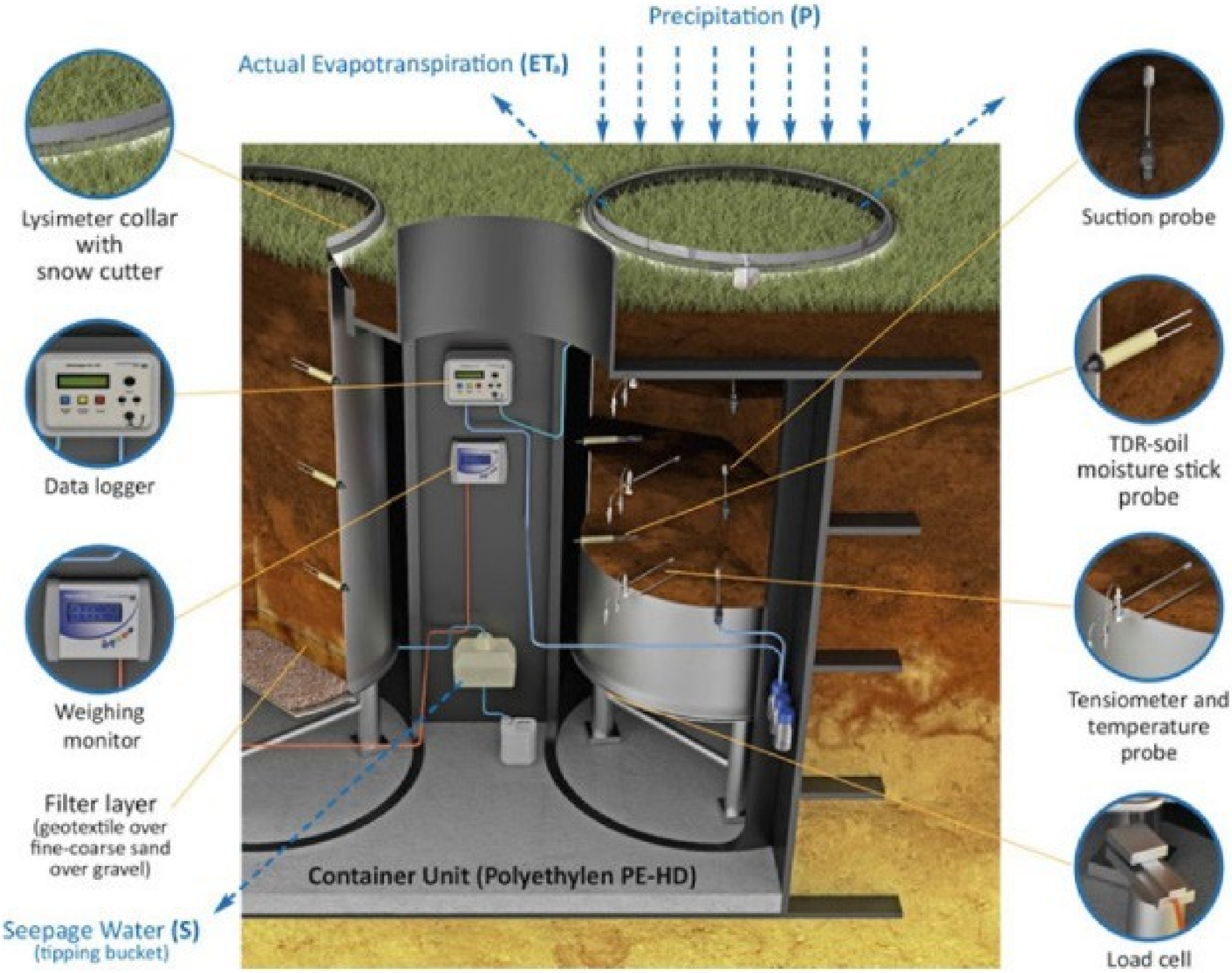

Figure 4. *Scheme of the two-section gravity lysimeter (Balykin et al. 2016).*

Figure 5. *Above-ground part of the two-section (Lys 1 – left section, Lys 2 – right) gravity lysimeter in the village of Poluyamki (photo by Bondarovich).*

Figure 6. *View of monoliths, Lys 1, arable mechanical fallow (left), Lys 2, steppe vegetation, Stipa pennata (right), Poluyamki village (photo by A.A. Bondarovich).*

The SSM information was obtained from the AMSR2 passive microwave radiometer on board the

GCOM-W1 (Global Change Observation Mission 1 Water) satellite, which was launched by the Japan Aerospace Exploration Agency (JAXA) in May 2012. According to the orbit configuration of GCOM-W1, AMSR2 performs both upward (13:30 local time) and downward (01:30 local time) imaging of the Earth's surface. The spatial resolution of the product AMSR2 SMC (Soil Moisture Content up to a 5 cm depth) is 0.1 degrees (~ 10 km), and the data are presented on hourly, daily, and monthly time scales. The survey results are freely available and distributed through the JAXA electronic resource (https://gcm-w1.jaxa.jp). There are various works dedicated to assessing the reliability of satellite SSM data (instruments: AMSR2, ASCAT, SMOS, and SMAP) and the results of their assimilation into soil models, such as: AMSR2 JAXA algorithm, Land Parameter Retrieval Model, SMAP and ERA5 reanalysis. Meanwhile, AMSR2 estimates show higher precision in arid areas (Chen et al. 2012; Kim et al. 2015; Zhang et al. 2017; Liu et al. 2019; Ma et al. 2019; Kaihotsu et al. 2019; Li et al. 2022; Yi et al. 2023; Min et al. 2023).

Year	Transition through +5/ start/end of vegetation period	Transition through +5/ start/end of vegetation period	Vegetation period (days)	Sum of active temperatures (through +10, °C)	Sum of precipitation for the vegetation period (mm)	Selyaninov HTC
2013	01.05[1] /3.10	21.05/27.09	156[1]	2262[1]	295.9[1]	0.9[1]
2014	29.03/08.10	30.04/18.09	193	2350	260	1.1
2015	11.04/11.10	18.04/21.09	183	2785	257	0.9
2016	01.04/04.10	15.04/01.10	186	2791	340	1.2
2017	11.04/23.09	16.04/22.09	165	2691	338	1.3
2018	12.04/14.10	21.04/08.10	185	2450	313	1.3

Table 1. *Dates of transition of stable mean daily air temperatures through 5 °C, 10 °C, sum of active temperatures and Selyaninov HTC in 2013–2018 according to WS "Poluyamki" data*

Note: [1] – data from 01.05.2013.

Result

A critical step in recovering RZSM information from satellite observations is selecting the optimal set of regressors. These regressors can be readily obtained from local monitoring networks or open databases. Given the limited global operational system for ground-based SM observations, parameters from the surface layer of the atmosphere, such as air temperature, relative humidity, precipitation, and evapotranspiration, become the primary sources of information for accurate RZSM forecasting (Travova 2022). It's worth noting the importance of including predictors like topography and physicochemical soil properties (Celik et al. 2022), as well as various climatic index parameters to enhance the description of nonlinear relationships (Chang et al. 2023).

As previously mentioned, machine learning techniques have established themselves as effective tools for addressing these problems. In the following sections, wepresent results for several of the outlined approaches, evaluating the quality of the models using widely accepted metrics such as Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and R-squared (R2).

In our current study, we analyzed a set of meteorological parameters derived from data spanning 2013 to 2020 from CT, NT, Lys 1, Lys 2, and WS stations. Two essential points to consider are: a) WS and Lys 1/Lys 2 stations are located within a 5-meter radius, while CT and NT stations are approximately 13,000 meters away from WS and Lys 1/Lys 2 (see Fig. 1); and b) Lys sensors tend to overestimate RZSM readings at a depth of 30 cm by roughly 50% compared to CT and NT data. This discrepancy is due to the steel vessel walls of the Lys monoliths, which prevent horizontal runoff and consequently increase moisture in the confined space (Bondarovich et al. 2021). We also note that CT and NT data were excluded from the subsequent analysis, because the observations of

Lys 1 and Lys 2 are, in our opinion, of higher quality (fewer omissions and artifacts in the data), and the time series is the longest. Following an initial analysis of the observation vector, we selected the following values for training our regression models:

L1: Soil moisture at a 30 cm depth (in %) from lysimetric observations (Lys 1). This sample represents soil cultivated using deep-loosing methods typical for agrocenoses in dry and arid steppes on chestnut soils.

L2: Soil moisture at a 30 cm depth (in %) from lysimetric observations (Lys 2). This sample reflects undisturbed areas of the dry steppe.

T, f, p, P: Observations from the automatic weather station, including temperature, humidity, and air pressure at 2 meters above the soil surface at the locations of lysimeters L1 and L2, as well as daily precipitation (P) in millimeters obtained via a precipitation gauge.

SMC5: Information on soil moisture in the surface soil layer up to a 5 cm depth from the AMSR-2/GCOM-W1 radiometer.

Additionally, we considered several features related to data recording intervals and seasonal variations in different parameters. These observations had varying temporal resolutions, for instance, AMSR2 data were available at 12-hour intervals, while the weather station and lysimeters provided hourly measurements. Thus, during data preparation, the 2015–2017 series were normalized to daily averages The transition to daily average values is due to the fact that the process of moisture change in deeper soil layers occurs with a time delay and less intensity compared to the surface layer. The position of non-linear dependence between surface and depth moisture content is now almost axiomatic in monitoring and modeling of soil moisture regime. (Biniak-Pieróg 2017; Li, Zhang 2021).

Out of all the data spanning 2013-2020, we obtained continuous observation series (Lys 1, Lys 2, T, f, p, P) only for the growing seasons of 2015, 2016, and 2017. In 2016, a fallow (mechanical fallow) experiment was conducted on the Lys 1 monolith to measure evapotranspiration, ensuring that crop transpiration did not affect RZSM. To properly utilize data from both monoliths, we corrected the 2016 Lys 1 data for the transpiration effect. This correction utilized observations from Lys 1 (L1) and Lys 2 (L2) in 2015 and 2017 to calculate the relative difference (Cdif) between Lys 2 and Lys 1 readings, defined as:

$Cdif = \mu\ (L1(2015) – L2(2015)\ e\ L1(2015) \cap L2(2015)) + \mu\ (L1(2015) – L2(2015)$

$e\ L1(2015)\ U\ L2(2015))\ /\ 2.$

Subsequently, we applied this correction to the L1(2016) observations using the following expression:

$L1(2016)\_cor = L1(2016) + Cdif * (L1(2016) – L2(2016)).$

In the next section, we present the seasonal RZSM dynamics for 2016 based on L1 and L2, along with the transpiration-corrected RZSM observations for Lys 1 2016 (L1(2016)_cor) (see Fig. 7). This figure illustrates how incorporating transpiration effects in the soil-atmosphere system influences the temporal dynamics of soil moisture.

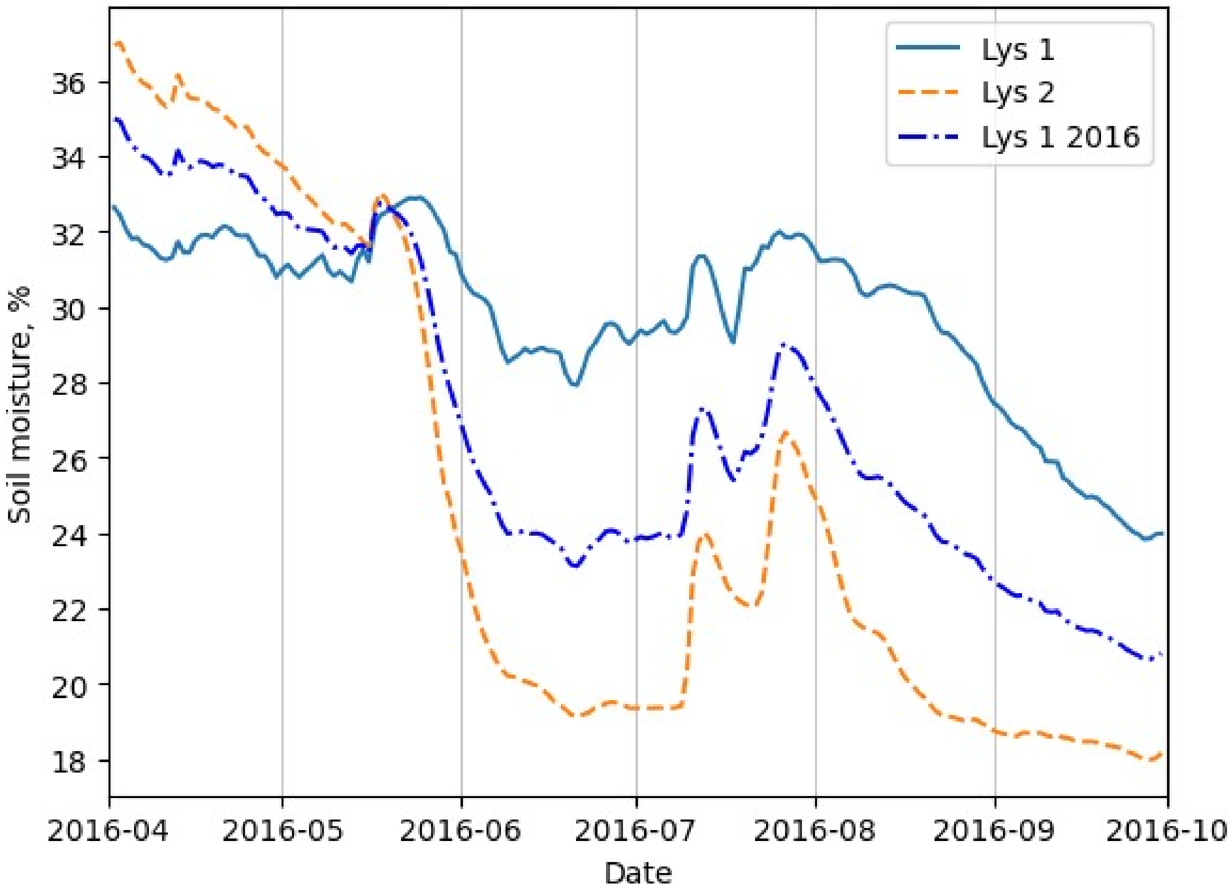

Figure 7. *Seasonal variation of the RZSM (soil moisture content at a 30 cm depth) from the Lys 1 and Lys 2 data and transpiration-corrected RZSM observations for Lys 1 (L1(2016)_ cor).*

A crucial aspect of modeling is identifying the degree of statistical intercorrelation among the regressors. While some authors often apply correlation analysis for this purpose (Grjibowski 2008), it's worth questioning its applicability due to certain limitations, particularly the assumption of linearity in trait relationships (David 1938). The relationship between soil moisture at different horizons during the growing season is typically nonlinear (Li et al. 2021). Our constructed correlation matrix (Pearson R2) reveals that the relationship between meteorological parameters and RZSM values from Lys 1 and Lys 2 data is indeed nonlinear in nature (see Fig. 8a). In addition to the correlation coefficient, we've also explored other methods, such as Mutual Information Connectivity (MI) (Mutual... 2023), which has unveiled significant interrelations between lysimeter readings, meteorological parameters, and SMC5 (see Fig. 8a, b). The work at this stage was performed with different types of geospatial data, where Lys 1 and Lys 2 are point observations and AMSR2 results are integral over a 10x10 km area. However, most of the pixel used covers arable land, so the results obtained characterize the water regime of agrocenoses.

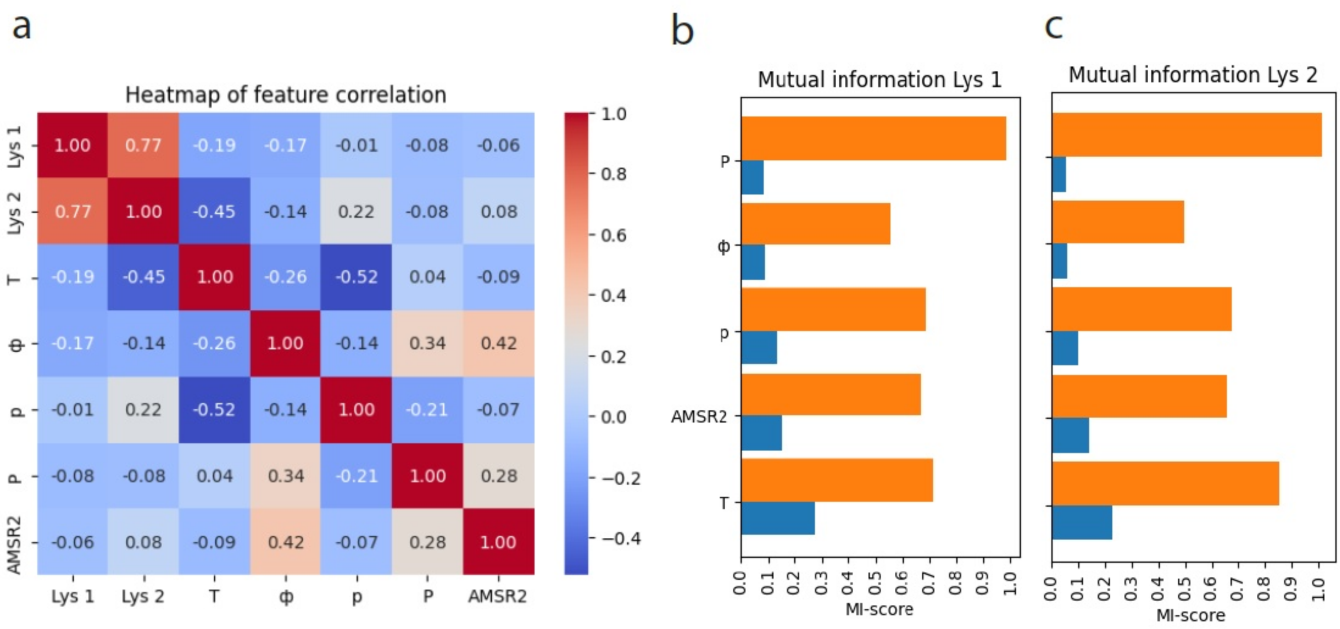

Figure 8. *Thermal correlation matrix (Pearson coefficient): a) between RZSM (%) at 30 cm depth according to both monoliths Lys 1, Lys 2, meteorological parameters: T (C°) – air temperature, φ (%) – air humidity, p (hPa) – atmospheric pressure, P (mm) – precipitation according to WS WXT 536 , SMC5 (%) at 1-5 cm depth by AMSR2; (b) MI between meteorological parameters, SMC5 (from AMSR2 data) and Lys 1; (c) MI between meteorological parameters, SMC5 (from AMSR2 data) and Lys 2. Note: in b and c, the blue color is the smoothing one day, the orange color is the smoothing with a window of 20 days.*

In previous studies (Bondarovich et al. 2020; 2021), it was established that the dynamics of soil moisture and temperature at a 30 cm depth are indeed correlated with atmospheric parameters, but this relationship exhibits inertia, as previously observed patterns dictate. Moreover, the amplitude of changes in parameters Lys 1 and Lys 2 during the growing season was notably less significant compared to meteorological parameters. Consequently, we employed a temporal smoothing procedure for all parameters except Lys 1 and Lys 2. Empirically, we determined the optimal smoothing window size by assessing the relationship between RZSM and the selected regressors using the Mutual Information (MI) method. The best results, within the range of 5–30 days, were achieved with a 20-day window size (refer to Fig. 8c). The outcomes of the smoothing process are depicted in Fig. 9. Notably, the seasonal variations in the dataset now closely resemble the variations in the Lys 1 and Lys 2 series (see Fig. 9).

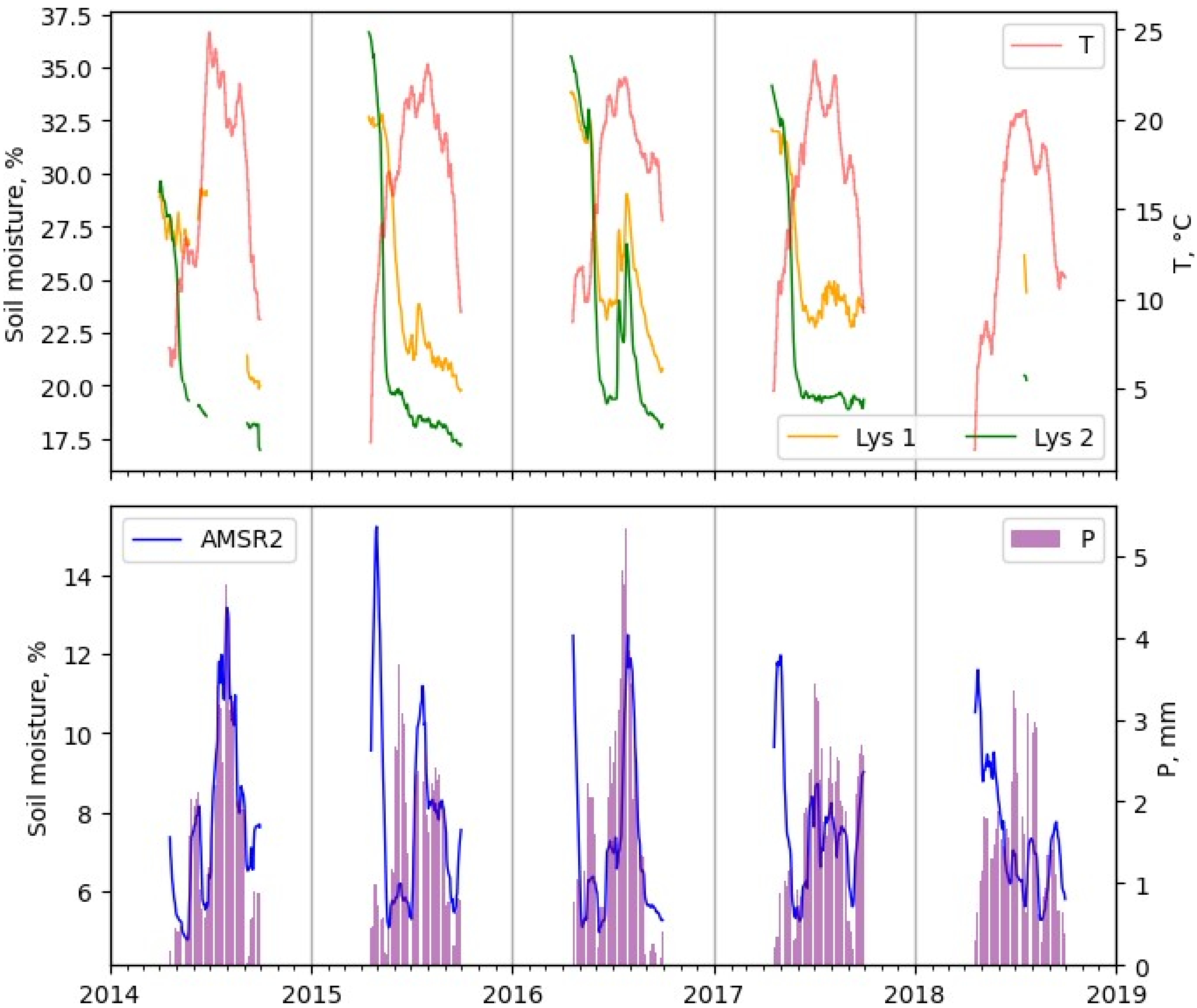

Figure 9. *Seasonal course of RZSM in % from Lys 1 and Lys 2 data and air temperature by WS (top figure), as well as SMC5 in % (from AMSR2 data) and precipitation by WS in 2014- 2018 after the smoothing procedure (window of 20 days).*

Additionally, we introduced two supplementary parameters, namely, the cumulative sum of active temperatures (Com_T) exceeding 10°C and the ordinal number of calendar months (Month). The inclusion of these parameters aimed to capture the temporal dynamics and seasonality of the modeled process, thereby reducing noise in the forecast data, especially towards the end of the growing season.

Following preprocessing, the selected regressors were as follows: T (air temperature), f (air humidity), p (atmospheric pressure at 2 m level), P (sum of daily precipitation), SMC5 (soil moisture at 5 cm from AMSR2 satellite data), Month (month number), and Com_T (cumulative sum of active temperatures). The final dataset consisted of 474 values covering the period from 2015 to 2017. Observations were limited to the growing season, spanning from April 15 to September 30.

According to the customary practice, we split the training dataset into training and test subsets in an 80/20% proportion (Kisekka et al. 2022). Model training was performed separately for the Lys 1 and Lys 2 series. We used default hyperparameter values for most methods (default scikit-learn library for Python), except for K-nearest Neighbors (KNN), where the number of nearest neighbors was set to 20. Notably, the applied approaches did not employ autoregressive principles. The metrics obtained during model training are summarized in Table 2.

Metrics for Lys 1 model			
Metric	KNN	RF	XGB
R2	0.883	0.984	0.981
MAE	1.013	0.287	0.310
MSE	1.614	0.215	0.267
RMSE	1.271	0.464	0.517
Metrics for Lys 2 model			
Metric	KNN	RF	XGB
R2	0.937	0.996	0.994
MAE	0.901	0.195	0.268
MSE	1.736	0.113	0.168
RMSE	1.318	0.337	0.410

Table 2. *Quality and statistical metrics of RZSM prediction models Lys 1, Lys 2*

These results demonstrate that the outlined approaches yield high-quality predictions in the test dataset, with prediction errors falling within the range of 0.2– 0.3%. A visual representation of the metric quality is presented in the scatterplot (refer to Fig. 10).

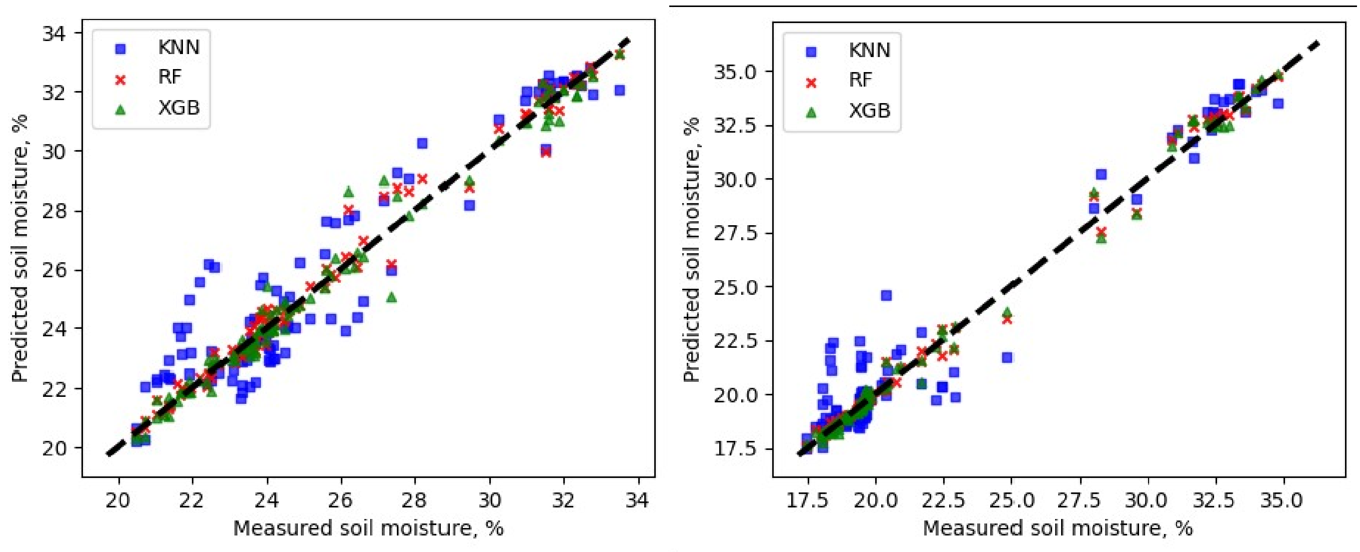

Figure 10. *Scatter plot diagram of the inhabited by roots of the soil layer moisture content in 2015–2017. obtained from observations from the test dataset (Lys 1 in the left figure, Lys 2 on the right figure).*

We also observed the seasonal variations in soil moisture at a 30 cm depth from 2015 to 2017, as predicted from the test sample data (Fig. 11). It's worth noting that the obtained series align with the seasonal distribution typical of the dry steppes of Eurasia (Banzragch et al. 2010; Kukharuk et al. 2017).

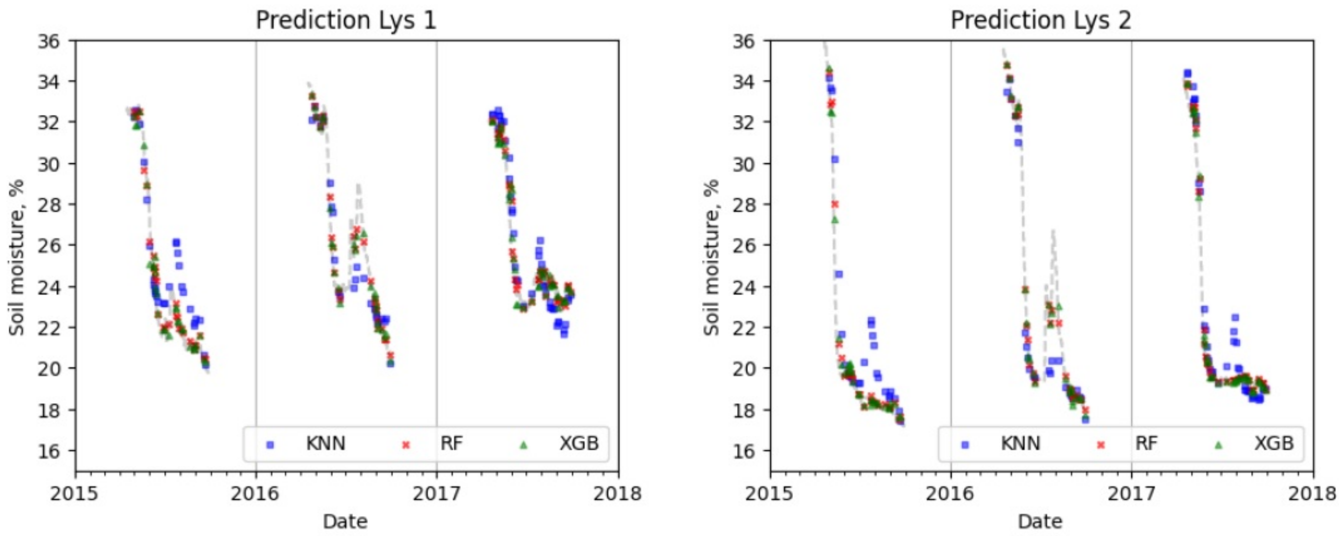

Figure 11. *Seasonal behavior of soil moisture at 30 cm depth in 2015–2017 predicted from the test sample data.*

In the final stage of the study, we predicted soil moisture behavior at a 30 cm depth for the years 2014 and 2018, as only for this period the initial meteorological data were representative. The prediction results were generated using K-nearest Neighbors (KNN), Random Forest (RF), and Extreme Gradient Boosting (XGB) models (Fig. 12). Notably, these methods, in general, replicate the seasonal trends of the desired value.

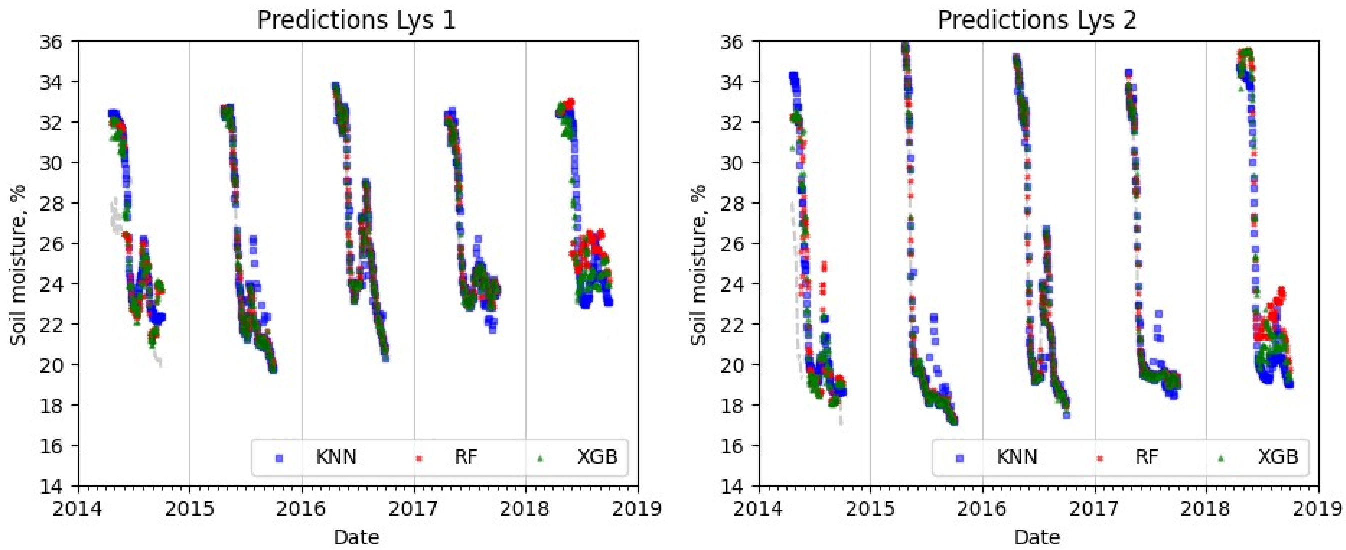

Figure 12. *Predicted seasonal course of RZSM (30 cm) in 2013-2018, predicted from the data from the full dataset (test and trane) Lys 1, Lys 2.*

The RF and XGB methods demonstrated the highest quality in reproducing the seasonal patterns of soil moisture at a 30 cm depth. These models have a proven track record of successful applications in predicting RZSM in arid regions (Adab et al. 2020, Carranza et al. 2021, Acharya et al. 2021, Zhang et al. 2022, Wang et al. 2022, He et al. 2022, Kisekka et al. 2022, Wang, Fu 2023). Some authors have also highlighted the superior performance of the XGB model for forecasting tasks using both satellite and ground data (Nguyen et al. 2022).

However, it's important to note that the outlined approaches excel in a space bounded by the extremes of the training dataset, and they may struggle with extrapolation in cases with limited data, a situation known as the cold start model training problem (Zhiyuan et al. 2023).

In summary, our study showcases the successful application of machine learning models, specifically K-nearest Neighbors (KNN), Random Forest (RF), and Extreme Gradient Boosting (XGB), in predicting root zone soil moisture (RZSM) at a 30 cm depth. These models demonstrate high predictive accuracy and reproduce the seasonal dynamics of soil moisture, especially in dry steppe regions like the Kulunda plain. While RF and XGB exhibit better performance, it's crucial to consider data limitations when applying these models, particularly in situations with sparse or limited datasets. Nonetheless, our findings provide valuable insights into RZSM forecasting, which is vital for addressing water resource challenges in arid regions.

To validate the models we conducted a comparison between the predicted estimates, Selyaninov's HTC, and the standardized precipitation index (SPI), recognized by the World Meteorological Organization as a key tool for monitoring mete- orological droughts (WMO, 2016).

These comparisons reveal a reasonable agreement between monthly mean RZSM from the XGB model and the Selyaninov HTC and SPI indices. This alignment underscores the reliability of the predictive estimates (Table 3). Due to the lack of alternative ground-based sources of RZSM information, this approach was proposed to verify the model results. We also estimated the correlation coefficient between the reconstructed RZSM and HTC, which is approximately equal to 0.2. Obviously, the above approach can be conditionally used to verify the quality of the previously obtained model, in view of the implicit mechanism linking HTC and RZSM. The solution of this problem requires additional research.

In general, the highest-quality models were obtained using the Lys 2 steppe vegetation monolith data. This result can be explained to the fact that various products, such as AMSR2 and SMAP, consistently demonstrate better performance in extracting soil moisture from vegetated areas (Zhang et al. 2017). In the regional context, the prediction results for arable land (Lys 1) take on greater significance, as it forms the foundation for the dry steppe subzone of Altai Krai. Notably, this is where we observe the most significant prediction errors, which may be attributed to factors including the retrieval of multiple 2016 values and the unique aspects of moisture dynamics associated with crops and agricultural practices. Future research should consider these factors to enhance the accuracy of soil moisture predictions at reference sites.

Furthermore, it is important to highlight that our proposed models were solely trained on new data, without accounting for past experience and autoregressive principles. The incorporation of historical data and the consideration of temporal dependencies hold the potential for further improving the predictive capabilities of these models (Žliobaitė et al. 2015).

Additionally, it is crucial to acknowledge that this study was conducted with a relatively small dataset, which could limit the generalizability of the results. This limitation has been addressed in separate publications (Hošovský et al. 2023). Expanding the dataset and conducting studies in diverse geographical regions can serve to validate and refine these findings.

Year	Month	HTC	SPI	XGB Lys 1*	XGB Lys 2*
2014	4	0.5	-1.4	355.2	367.5
	5	1.1	0.4	971.8	948.9
	6	0.5	-1.1	755.2	614.5
	7	1.8	1.2	717.9	609.9
	8	0.6	0.2	723.7	598.4
	9	0.6	-0.4	681.8	564.6
2015	4	1.2	-0.5	359.0	387.5
	5	0.5	-0.6	963.3	763.2
	6	0.9	1.0	698.8	579.9
	7	1.2	0.1	696.4	569.2
	8	0.9	1.5	660.0	563.4
	9	0.5	-0.4	616.2	527.2
2016	4	0.9	0.4	353.8	375.9
	5	1.1	0.6	972.0	958.2
	6	0.9	0.6	726.9	603.4
	7	1.9	1.4	807.8	695.1
	8	0.5	-0.9	775.9	645.4
	9	0.2	-0.9	647.7	553.3
2017	4	0.8	-0.7	359.7	374.2
	5	0.6	-0.4	937.4	858.0
	6	0.8	0.4	713.9	588.9
	7	1.5	0.5	737.1	602.0
	8	0.7	0.0	749.5	602.0
	9	1.6	2.0	704.0	576.3
2018	4	1.8	0.1	354.2	394.4
	5	1.8	1.1	986.6	1074.8
	6	1.1	1.5	819.4	711.8
	7	1.2	-0.2	777.4	667.9
	8	0.6	-0.8	779.8	647.6
	9	0.5	-0.5	745.0	621.3

Figure 13. *Table 3. Comparison of RZSM predictions with Selyaninov HTC and SPI*

Note: * – summary soil moisture of month (mm).

Conclusion

This research is focused on developing a method to reconstruction soil moisture content at the root layer level using meteorological data and AMSR2 satellite microwave radiometer information. Our approach leverages advanced machine learning techniques, including KNN, XGB, and RF. Here are the key findings:

By employing a 20-day moving average window to preprocess the input data, we achieved a twofold increase in the correlation between selected predictors and root zone soil moisture (RZSM).

The models we developed consistently demonstrated high-quality predictions, with prediction errors falling within the narrow range of 0.2–0.3 %. Notably, the predicted seasonal variations in soil moisture at a 30 cm depth from 2015 to 2017 closely matched the area's actual distribution characteristics.

When it comes to replicating the seasonal RZSM patterns at a 30 cm depth in 2014 and 2018, the RF and XGB methods outperformed others. Comparing these predictive results with the monthly moisture availability index of HTC and SPI revealed a satisfactory level of agreement.

These findings underscore the effectiveness of our methodology in forecasting soil moisture levels and its potential utility in various applications.

Acknowledgements

The work was supported by the Ministry of Science and Higher Education of the Russian Federation (State Assignment for scientific research carried out at Altai State University, project FZMW-2023-0007).

References

Abdrakhmanova AO, Filippova AV (2022) Degree of drought and duration of the interphase period of potato development according to planting time in Western Kazakhstan. Izvestiya Orenburg State Agrarian University 4 (96): 50–55.

Acharya U, Daigh ALM, Oduor PG (2021) Machine Learning for Predicting Field Soil Moisture Using Soil, Crop, and Nearby Weather Station Data in the Red River Valley of the North. Soil Systems 5: 57. https://doi.org/10.3390/soilsystems5040057

Adab H, Morbidelli R, Saltalippi C, Moradian M, Ghalhari GAF (2020) Machine learning to estimate surface soil moisture from remote sensing data. Water 12(11): 3223. https://doi.org/10.3390/w12113223

Afshar MH, Bulut B, Duzenli E, Amjad M, Yilmaz MT (2022) Global spatiotemporal consistency between meteorological and soil moisture drought indices. Agricultural and Forest Meteorology 316: 108848. https://doi.org/10.1016/j.agrformet.2022.108848

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

Babaeian E, Sadeghi M, Jones SB, Montzka C, Vereecken H, Tuller M (2019) Ground, proximal, and satellite remote sensing of soil moisture. Reviews of Geophysics 57: 530–616. https://doi.org/10.1029/2018RG000618

Baldwin D, Manfreda S, Keller K, Smithwick EAH (2017) Predicting root zone soil moisture with soil properties and satellite near-surface moisture data across the conterminous United States. Journal of Hydrology 546: 393–404. https://doi.org/10.1016/j.jhydrol.2017.01.020

Balykin D, Puzanov A, Stephan E, Meissner R (2016) Using the innovative Lysimeter technology in the German-Russian research project KULUNDA. In: Mueller L, Sheudshen AK, Eulenstein F (Eds) Novel methods for monitoring and managing land and water resources Siberia. Springer International Publishing, Cham, 387–399. https://doi.org/10.1007/978-3-319-24409-9_16

Biniak-Pieróg M (2017) Monitoring of atmospheric precipitation and soil moisture as basis for the estimation of effective supply of soil profile with water. Wydawnictwo Uniwersytetu Przyrodniczego we Wrocławiu / Uniwersytet Przyrodniczy we Wrocławiu, Wrocław, 198 pp. [In Polish]

Bodner G, Nakhforoosh A, Kaul HP (2015) Management of crop water under drought: a review.

Agronomy for Sustainable Development 35: 401–442. https://doi.org/10.1007/s13593-015-0283-4

Bondarovich AA, Scherbinin V, Ponkina EV, Matsyura A, Puzanov A, Stephan E, Balykin D, Rupp H, Meissner R (2020) Soil Moisture and Evapotranspiration In: KULUNDA: Climate Smart Agriculture, Innovations in Landscape Research. Springer Nature, Switzerland AG. https://doi.org/10.1007/978-3-030-15927-6_12

Bondarovich AA, Matsyura AV, Schmidt G (2021) The Application of Digital Monitoring in Studying Soil Water Holding Capacity and Implications for Conservation Agriculture in the South-Western Siberia. IOP Conference Series: Earth and Environmental Science 670: 012056. https://doi.org/10.1088/1755-1315/670/1/012056

Brakhasi F, Walker JP, Ye N, Wu X, Shen X, Yeo IY, Boopathi N, Kim E, Kerr Y, Jackson T (2023) Towards soil moisture profile estition in the root zone using L- and P-band radiometer observations: A coherent modelling approach. Science of Remote Sensing 7:100079. https://doi.org/10.1016/j.srs.2023.100079

Coleen C, Corjan N, Michiel P, Martine van der P (2021) Root zone soil moisture estimation with Random Forest. Journal of Hydrology 593: 125840. https://doi.org/10.1016/j.jhydrol.2020.125840

Celik MF, Isik MS, Yuzugullu O, Fajraoui N, Erten E (2014) Soil Moisture Prediction from Remote Sensing Images Coupled with Climate, Soil Texture and Topography via Deep Learning. Remote Sensing 14: 5584. https://doi.org/10.3390/rs14215584

Chai Y, Liu H, Yu Y, Yang Q, Zhang X, Zhao W, Guo L, Yetemen O (2023) Strategies of Parameter Optimization and Soil Moisture Sensor Deployment for Accurate Estimation of Evapotranspiration Through a Data-driven Method. Agricultural and Forest Meteorology 331: 109354. https://doi.org/10.1016/j.agrformet.2023.109354

Chang Y, Yi W, Chen J, Liu X, Meng W, Fan Zh, Zhang R, Hai Ch (2023) Quantifying the Coupled Effect between Soil Moisture and Climate in the Desert Steppe Environment of Inner Mongolia, China. Water 15: 1150. https://doi.org/10.3390/w15061150

Chen Xiu-zhi, Chen Shui-sen, Zhong Ruo-fei, Su Yong-xian, Liao Ji-shan, Dan Li, Liu-sheng Han, Yong Li, Xia Li (2012) A semi-empirical inversion model for assessing surface soil moisture using AMSR-E brightness temperatures. Journal of Hydrology 456–457: 1–11. https://doi.org/10.1016/j.jhydrol.2012.05.022

Cheng M, Li B, Jiao X, Huang X, Fan H, Lin R, Liu K (2022) Using multimodal remote sensing data to estimate regional-scale soil moisture content: A case study of Beijing, China. Agricultural Water Management 260: 107298. https://doi.org/10.1016/j.agwat.2021.107298

David FN (1938) Tables of the Ordinates and Probability Integral of the Distribution of the Correlation Coefficient in Small Samples. Cambridge University, Cambridge. Dong Wang, Yu Liu, Gao-Lin Wu, Lu-Ming Ding, Zheng Yang, Hong-Min Hao (2015) Effect of rest-grazing management on soil water and carbon storage in an arid grassland (China). Journal of Hydrology 527: 754–760. https://doi.org/10.1016/j.jhydrol.2015.05.036

Dorigo W, Himmelbauer I, Aberer D, Schremmer L, Petrakovic I, Zappa L, Preimesberger W, Xaver A, Annor F, Ardö J, Baldocchi D, Bitelli M, Blöschl G, Bogena H, Brocca L, Calvet J-C, Camarero JJ, Capello G, Choi M, Cosh MC, van de Giesen N, Hajdu I, Ikonen J, Jensen KH, Kanniah KD, de Kat I, Kirchengast G, Rai PK, Kyrouac J, Larson K, Liu S, Loew A, Moghaddam M, Fernández JM, Bader CM, Morbidelli R, Musial JP, Osenga E, Palecki MA, Pellarin T, Petropoulos GP, Pfeil I, Powers J, Robock A, Rüdiger C, Rummel U, Strobel M, Su Z, Sullivan R, Tagesson T, Varlagin A, Vreugdenhil M, Walker J, Wen J, Wenger F, Wigneron JP, Woods M, Yang K, Zeng Y, Zhang X, Zreda M, Dietrich

S, Gruber A, van Oevelen P, Wagner W, Scipal K, Drusch M, Sabia R (2021) The International Soil Moisture Network: serving Earth system science for over a decade. Hydrology and Earth System Sciences 25: 5749–5804. https://doi.org/10.5194/hess-25-5749-2021

Vaz CM, Jones S, Meding M, Tuller M (2013) Evaluation of standard calibration functions for eight electromagnetic soil moisture sensors. Vadose Zone Journal 12(2). https://doi.org/10.2136/vzj2012.0160

Draper CS, Mahfouf J-F, Walker JP (2009) An EKF assimilation of AMSR-E soil moisture into the ISBA land surface scheme. Journal of Geophysical Research 114: D20104. https://doi.org/10.1029/2008JD011650

Drusch M (2007) Initializing numerical weather prediction models with satellite-derived surface soil moisture: Data assimilation experiments with ECMWF's Integrated Forecast System and the TMI soil moisture data set. Journal of Geophysical Research 112: D03102. https://doi.org/10.1029/2006JD007478

FAO (2000) Land resource potential and constraints at regional and country levels. World Soil Resources Report 90. Rome FAO (2021)The State of the World's Land and Water Resources for Food and Agriculture–Systems at breaking point. Synthesis report 2021. Rome. https://doi.org/10.4060/cb7654en

Fisher JB, Melton F, Middleton E, Hain C, Anderson M, Allen R, McCabe MF, Hook S, Baldocchi D, Townsend PA, Kilic A, Tu K, Miralles DD, Perret J, Lagouarde J-P, Waliser D, Purdy AJ, French A, Schimel D, Famiglietti JS, Stephens G, Wood EF (2017) The future of evapotranspiration: Global requirements for ecosystem functioning, carbon and climate feedbacks, agricultural management, and water resources. Water Resources Research 53(4): 2618–2626. https://doi.org/10.1002/2016WR020175

Ghorbani MA, Zadeh HA, Isazadeh M, Terzi O (2016) A comparative study of artificial neural network (MLP, RBF) and support vector machine models for river flow prediction. Environmental Earth Sciences 75: 476. https://doi.org/10.1007/s12665-015-5096-x

Global Climate Observing System (2010) Implementation plan for the global observing system for climate in support of the UNFCCC. Technical Report WMO/TD, 1244:23. Available from: https://www.wmo.int/pages/prog/gcos/Publications/gcos-92_GIP_ES.pdf/.

Grjibovski AM (2008) Correlation analysis. Human Ecology 9: 50–59. He B, Jia B, Zhao Y, Wang X, Wei M, Dietzel R (2022) Estimate soil moisture of maize by combining support vector machine and chaotic whale optimization algorithm. Agricultural Water Management 267: 107618. https://doi.org/10.1016/j.agwat.2022.107618

Hodges B, Tagert ML, Paz JO, Meng Q (2023) Assessing in-field soil moisture variability in the active root zone using granular matrix sensors. Agricultural Water Management 282: 108268. https://doi.org/10.1016/j.agwat.2023.108268

Hollmann R, Merchant CJ, Saunders R, Downy C, Buchwitz M, Cazenave A, Chuvieco E, Defourny P, de Leeuw G, Forsberg R, Holzer-Popp T, Paul F, Sandven S, Sathyendranath S, van Roozendael M, Wagner W (2013) The ESA climate change initiative: Satellite data records for essential climate variables. Bulletin of the American Meteorological Society 94(10): 1541–1552. https://doi.org/10.1175/BAMS-D-11-00254.1

Zhiyuan W, Ning L, Guodong L, Xinyu L, Yue W, Lin Z (2023) Meta-Learning-Based Spatial-Temporal Adaption for Coldstart Air Pollution Prediction. International Journal of Intelligent Systems 2023: 373455. https://doi.org/10.1155/2023/3734557

Huang S, Zhang X, Wang C, Chen N (2023) Two-step fusion method for generating 1 km seamless multi-layer soil moisture with high accuracy in the Qinghai-Tibet plateau. ISPRS Journal of Photogrammetry and Remote Sensing 197: 346–363. https://doi.org/10.1016/j.isprsjprs.2023.02.009

Humphrey VW, Berg AM, Ciais PG, Pierre G, Martin J, Reichstein M, Seneviratne SI, Frankenberg Ch (2021) Soil moisture-atmosphere feedback dominates land carbon uptake variability. Nature 59: 65–69. https://doi.org/10.1038/s41586-021-03325-5

Imaoka K, Kachi M, Kasahara M, Ito N, Nakagawa K, Oki T (2010) Instrument performance and calibration of AMSR-E and AMSR2. International archives of the photogrammetry, remote sensing and spatial information science 38(8): 13–18.

Jabro JD, Stevens WB, Iversen WM (2018) Field performance of three real-time moisture sensors in sandy loam and clay loam soils. Archives of Agronomy and Soil Science 64(7): 930–938. https://doi.org/10.1080/03650340.2017.1393528

Kaihotsu I, Asanuma J, Aida K, Oyunbaatar D (2019) Evaluation of the AMSR2 L2 soil moisture product of JAXA on the Mongolian Plateau over seven years (2012-2018). SN Applied Sciences 1: 1477. https://doi.org/10.1007/s42452-019-1488-y

Kamran M, Yan Zh, Ahmad I, Jia Q, Ghani MU, Chen X, Chang Sh, Li T, Siddique KHM, Fahad S, Hou F (2023) Assessment of greenhouse gases emissions, global warming potential and net ecosystem economic benefits from wheat field with reduced irrigation and nitrogen management in an arid region of China. Agriculture, Ecosystems & Environment 341: 108197. https://doi.org/10.1016/j.agee.2022.108197

Kang J, Li X, Jin R, Ge Y, Wang J, Wang J (2014) Hybrid Optimal Design of the Eco-Hydrological Wireless Sensor Network in the Middle Reach of the Heihe River Basin. China Sensors 14(10): 19095–19114. https://doi.org/10.3390/s141019095

Karamvand A, Hosseini SA, Sharafati A (2022) SMAP products for prediction of surface soil moisture by ELM network model and agricultural drought index. Acta Geophysica 71: 1845–1856. https://doi.org/10.1007/s11600-022-00973-7

Kassam A, Friedrich T, Derpsch R (2019) Global spread of conservation agriculture. International Journal of Environmental Studies 76(1): 29–51. https://doi.org/10.1080/00207233.2018.1494927

Kharlamova NF (2020) Climatic Variability of the Kulunda Steppe Belt. In: KULUNDA: Climate Smart Agriculture, Innovations in Landscape Research. Springer Nature. Switzerland AG. https://doi.org/10.1007/978-3-030-15927-6_3

Kim S, Liu YiY, Johnson FM, Parinussa RM, Sharma A (2015) A global comparison of alternate AMSR2 soil moisture products: Why do they differ? Remote Sensing of Environment 161: 43–62. https://doi.org/10.1016/j.rse.2015.02.002

Kisekka I, Peddinti SR, Kustas WP et al (2022) Spatial-temporal modeling of root zone soil moisture dynamics in a vineyard using machine learning and remote sensing. Irrig Sci, 40:761-777. https://doi.org/10.1007/s00271-022-00775-1

Kuharuk NS, Smirnova LG, Narozhnaya AG, Chendev YuG, Glazunov GP (2017) Dynamics of soil moisture conservation areas of forest-steppe on the background of interdecadal climate variability. Scientific Bulletins of Belgorod State University. Seria Natural Sciences 25(274): 79–90.

KULUNDA (2020) Climate Smart Agriculture, South Siberian Agro-steppe as Pioneering Region for

Sustainable Land Use, Innovations in Landscape Research. Springer Nature, Switzerland AG. https://doi.org/10.1007/978-3-030-15927-6

Ladányi Z, Barta K, Blanka V, Pálffy B (2021) Assessing Available Water Content of Sandy Soils to Support Drought Monitoring and Agricultural Water Management. Water Resources Management 35: 869–880. https://doi.org/10.1007/s11269-020-02747-6

Lafond GP, Walley F, May WE, Holzapfel CB (2011) Long term impact of No-Till on soil properties and crop productivity on the Canadian prairies. Soil & Tillage Research 117: 110–123. https://doi.org/10.1016/j.still.2011.09.006

Lai EN, Wang-Erlandsson L, Virkki V, Porkka M, van der Ent RJ (2022) Root zone soil moisture in over 25 % of global land permanently beyond pre-industrial variability as early as 2050. EGUsphere [preprint]. https://doi.org/10.5194/egusphere-2022-971

Lal R (2020) Managing soils for negative feedback to climate change and positive impact on food and nutritional security. Soil Science and Plant Nutrition 66(1): 1–9. https://doi.org/10.1080/00380768.2020.1718548

Li B, Wang C, Gu X, Zhou X, Ma M, Li L, Feng Zh, Ding T, Li X, Jiang T, Li X, Zheng X (2022) Accuracy calibration and evaluation of capacitance-based soil moisture sensors for a variety of soil properties. Agricultural Water Management 273: 107913. https://doi.org/10.1016/j.agwat.2022.107913

Li J, Zhang L (2021) Comparison of Four Methods for Vertical Extrapolation of Soil Moisture Contents from Surface to Deep Layers in an Alpine Area. Sustainability 13(16): 8862. https://doi.org/10.3390/su13168862

Li P, Zha Y, Shi L, Michael TC-H, Zhang Y, Zeng W (2020) Comparison of the use of a physical-based model with data assimilation and machine learning methods for simulating soil water dynamics. Journal of Hydrology 584: 124692. https://doi.org/10.1016/j.jhydrol.2020.124692

Liao L, Zhang L, Bengtsson L (2008) Soil moisture variation and water consumption of spring wheat and their effects on crop yield under drip irrigation. Irrigation and Drainage Systems 22: 253–270. https://doi.org/10.1007/s10795-008-9055-5

Liu C, Zhang X, Zhang Y (2002) Determination of daily evaporation and evapotranspira- tion of winter wheat and maize by large-scale weighing lysimeter and micro-lysimeter. Agricultural and Forest Meteorology 111(2): 109–120. https://doi.org/10.1016/S0168-1923(02)00015-1

Liu Z, Fu YH, Shi X, Lock TR, Kallenbach RL, Yuan Zh (2022) Soil moisture determines the effects of climate warming on spring phenology in grasslands. Agricultural and Forest Meteorology 323: 109039. https://doi.org/10.1016/j.agrformet.2022.109039

Liu E, Zhu Y, Lü H, Horton R, Gou Q, Wang X, Ding Zh, Xu H, Pan Y (2023) Estimation and Assessment of the Root Zone Soil Moisture from Near-Surface Measurements over Huai River Basin. Atmosphere 14: 124. https://doi.org/10.3390/atmos14010124

Liu J, Chai L, Lu Z, Liu S, Qu Y, Geng D, Song Y, Guan Y, Guo Z, Wang J, Zhu Z (2019) Evaluation of SMAP, SMOS-IC, FY3B, JAXA, and LPRM Soil Moisture Products over the Qinghai-Tibet Plateau and Its Surrounding Areas. Remote Sensing 11: 792. https://doi.org/10.3390/rs11070792

Luo Q, Yang K, Chen Y, Zhou X (2020) Method development for estimating soil organic carbon content in an alpine region using soil moisture data. Science China Earth Sciences 63: 591–601. https://doi.org/10.1007/s11430-019-9554-8

Ma H, Zeng J, Chen N, Zhang X, Cosh MH, Wang W (2019) Satellite surface soil moisture from SMAP, SMOS, AMSR2 and ESA CCI: A comprehensive assessment using global ground-based observations. Remote Sensing of Environment 231: 111215. https://doi.org/10.1016/j.rse.2019.111215

Mahmood R, Littell A, Hubbard KG, You J (2012) Observed data-based assessment of relationships among soil moisture at various depths, precipitation, and temperature. Applied Geography 34: 255–264. https://doi.org/10.1016/j.apgeog.2011.11.009

Mao Na, Li Wenrao, Wei Xiaorong, Shao Ming'an (2020) Soil moisture- and texture-dependent effects of soil deposition on evaporation and carbon emission. Soil and Tillage Research 204: 104703. https://doi.org/10.1016/j.still.2020.104703

Margulis SA McLaughlin D, Entekhabi D, Dunne S (2022) Land data assimilation and estimation of soil moisture using measurements from the Southern Great Plains 1997 Field Experiment. Water Resources Research 38(12): 1299. https://doi.org/10.1029/2001WR001114

Meissner R, Rupp H, Haselow L (2020) Use of lysimeters for monitoring soil water balance parameters and nutrient leaching. Climate Change and Soil Interactions: 171–205. https://doi.org/10.1016/B978-0-12-818032-7.00007-2

Min X, Li D, Shangguan YL, Tian S, Shi Z (2023) Characterizing the accuracy of satellite-based products to detect soil moisture at the global scale. Geoderma 432: 116388. https://doi.org/10.1016/j.geoderma.2023.116388

Mizgirev A, Zierdt M, Illiger P, Kudryavtsev AE, Bondarovich AA, Stephan E, Schmidt G (2020) Soils of the Central and Western Steppes of the Altai Krai. In: KULUNDA: Climate Smart Agriculture, Innovations in Landscape Research. Springer Nature. Switzerland AG. https://doi.org/10.1007/978-3-030-15927-6_5

Nandintsetseg B, Shinoda M, Kimura R, Ibaraki Y (2010) Relationship between Soil Moisture and Vegetation Activity in the Mongolian Steppe. Scientific online letters on the atmosphere. SOLA 6: 029–032. https://doi.org/10.2151/sola.2010-008

Nguyen TT, Ngo HH, Guo WS, Chang SW, Nguyen DD, Nguyen CT, Zhang J, Liang S, Bui XT, Hoang NB (2022) A low-cost approach for soil moisture prediction using multisensor data and machine learning algorithm. Science of The Total Environment 833: 155066. https://doi.org/10.1016/j.scitotenv.2022.155066

Pal M, Maity R, Dey S (2016) Statistical Modelling of Vertical Soil Moisture Profile: Coupling of Memory and Forcing. Water Resour Manage 30: 1973–1986. https://doi.org/10.1007/s11269-016-1263-4

Qin J, Tian J, Yang K, Lu H, Li X, Yao L, Shi J (2022) Bias correction of satellite soil moisture through data assimilation. Journal of Hydrology 610: 127947. https://doi.org/10.1016/j.jhydrol.2022.127947

Remote Sensing of Energy Fluxes and Soil Moisture Content (2017) Petropoulos GP (Ed.) CRC Press (Taylor & Francis), Boca Raton.

Robinson DA, Campbell CS, Hopmans JW, Hornbuckle BK, Jones SB, Knight R, Ogden F, Selker J, Wendroth O (2008) Soil Moisture Measurement for Ecological and Hydrological Watershed-Scale Observatories: A Review. Vadose Zone Journal 7(1): 358–389. https://doi.org/10.2136/vzj2007.0143

Robock A, Vinnikov KYa, Srinivasan G, Entin JK, Hollinger SE, Speranskaya NA, Liu S, Namkhai A

(2000) The global soil moisture data bank. Bulletin of the American Meteorological Society 81(6): 1281–1299. https://doi.org/10.1175/1520-0477(2000)081<1281:TGSMDB>2.3.CO;2

Selyaninov's hydrothermal coefficient. Meteorological Dictionary. Available from: http://meteorologist.ru/gidrotermicheskiy-koeffitsient-selyaninova.html/Accessed on September 26, 2023.

Seneviratne SI, Corti T, Davin EL, Hirschi M, Jaeger EB, Lehner I, Orlowsky B, Teuling AJ (2010) Investigating soil moisture-climate interactions in a changing climate: A review. Earth-Science Reviews 99(3–4): 125–161. https://doi.org/10.1016/j.earscirev.2010.02.004

Shahi AP, Rai PK, Islam R, Mishra VN (2023) Remote sensing data extraction and inversion techniques: A review. In: Abhay KS, Shani T (Eds) Earth Observation, Atmospheric Remote Sensing. Elsevier: 85–104. https://doi.org/10.1016/B978-0-323-99262-6.00021-3

Stefan V-G, Indrio G, Escorihuela M-J, Quintana-Seguí P, Villar JM (2021) High-Resolution SMAP-Derived Root-Zone Soil Moisture Using an Exponential Filter Model Calibrated per Land Cover Type. Remote Sensing 13: 1112. https://doi.org/10.3390/rs13061112

Tian J, Zhang Y, Guo J, Zhang X, Ma N, Wei H, Tang Z (2022) Predicting root zone soil moisture using observations at 2121 sites across China. Science of The Total Environment 847: 157425. https://doi.org/10.1016/j.scitotenv.2022.157425

Walker JP, Willgoose GR, Kalma JD (2001) One-dimensional soil moisture profile retrieval by assimilation of near-surface observations: a comparison of retrieval algorithms. Advances in Water Resources 24(6): 631–650. https://doi.org/10.1016/S0309-1708(00)00043-9

Wang S, Fu G (2023) Modelling soil moisture using climate data and normalized difference vegetation index based on nine algorithms in alpine grasslands. Frontiers in Environmental Science 11: 1130448. https://doi.org/10.3389/fenvs.2023.1130448

Wang JJ, Wu F, Shang JL, Zhou Q, Ahmad I, Zhou GS (2022) Saline soil moisture mapping using Sentinel-1A synthetic aperture radar data and machine learning algorithms in humid region of China's East Coast. Catena 213: 106189. https://doi.org/10.1016/j.catena.2022.106189

Wang-Erlandsson L, Tobian A, van der Ent RJ, Fetzer I, te Wierik S (2022) A planetary boundary for green water. Nature Reviews Earth & Environment 3: 380–392. https://doi.org/10.1038/s43017-022-00287-8

Wesche K, Korolyuk A, Lashchinsky N, Silantyeva MM, Rosche C, Hensen I (2020) The Kulunda Steppe as Part of the Eurasian Steppe Belt. In: KULUNDA: Climate Smart Agriculture, Innovations in Landscape Research. Springer Nature. Switzerland AG. https://doi.org/10.1007/978-3-030-15927-6_2

Wigneron J-P, Calvet J-C, Pellarin T, Van de Griend AA, Berger M, Ferrazzoli P (2023) Retrieving near-surface soil moisture from microwave radiometric observations: current status and future plans. Remote Sensing of Environment 85(4): 489–506. https://doi.org/10.1016/S0034-4257(03)00051-8

World Meteorological Organization (WMO) and Global Water Partnership (GWP) (2016) Handbook of Drought Indicators and Indices (Svoboda M, Fuchs BA). Integrated Drought Management Programme (IDMP), Integrated Drought Management Tools and Guidelines Series 2. Geneva.

Wyatt BM, Ochsner TE, Zou CB (2021) Estimating root zone soil moisture across diverse land cover types by integrating in-situ and remotely sensed data. Agricultural and Forest Meteorology 307:

108471. https://doi.org/10.1016/j.agrformet.2021.108471

Yi C, Li X, Zeng J, Fan L, Xie Z, Gao L, Xing Z, Ma H, Boudah A, Zhou H, Zhou W, Sheng Y, Dong T, Wigneron J-P (2023) Assessment of five SMAP soil moisture products using ISMN ground-based measurements over varied environmental conditions. Journal of Hydrology 619: 129325. https://doi.org/10.1016/j.jhydrol.2023.129325

Yinglan A, Guoqiang W, Peng H, Xiaoying L, Baolin X, Qingqing F (2022) Root-zone soil moisture estimation based on remote sensing data and deep learning. Environmental Research 212(B): 113278. https://doi.org/10.1016/j.envres.2022.113278

Yu J, Wu Y, Xu L, Peng J, Chen G, Shen X, Lan R, Zhao C, Zhangzhong L (2022) Evaluating the Hydrus-1D Model Optimized by Remote Sensing Data for Soil Moisture Simulations in the Maize Root Zone. Remote Sensing 14(23): 6079. https://doi.org/10.3390/rs14236079

Yue K, Fornara DA., Heděnec P, Wu Q, Peng Y, Peng X, Ni X, Wu F, Peñuelas J (2023) No tillage decreases GHG emissions with no crop yield tradeoff at the global scale. Soil and Tillage Research 228: 105643. https://doi.org/10.1016/j.still.2023.105643

Yu-Lin L, Jian-Yuan C, Tong-Hui Z, Ha-Lin Z (2003) Measurement of evapotranspiration of irrigated spring wheat and maize in a semi-arid region of north China. Agricultural Water Management 61(1): 1–12. https://doi.org/10.1016/S0378-3774(02)00177-4

Zeng J, Li Z, Chen Q, Bi H, Qiu J, Zou P (2015) Evaluation of remotely sensed and reanalysis soil moisture products over the Tibetan Plateau using in-situ observations. Remote Sensing of Environment 163: 91–110. https://doi.org/10.1016/j.rse.2015.03.008

Zhang G, Zhang Y, Zhao D, Liu S, Wen X, Han J, Liao Y (2023) Quantifying the impacts of agricultural management practices on the water use efficiency for sustainable production in the Loess Plateau region: A meta-analysis. Field Crops Research 291: 108787. https://doi.org/10.1016/j.fcr.2022.108787

Zhang U, Wang S, Wang H, Wang R, Wang X, Li J (2018) Crop yield and soil properties of dryland winter wheat-spring maize rotation in response to 10-year fertilization and conservation tillage practices on the Loess Plateau. Field Crops Research 225: 170–179. https://doi.org/10.1016/j.fcr.2018.07.003

Zhang X, Zhang T, Zhou P, Shao Y, Gao S (2017) Validation Analysis of SMAP and AMSR2 Soil Moisture Products over the United States Using Ground-Based Measurements. Remote Sensing 9(2): 104. https://doi.org/10.3390/rs9020104

Zhang YW, Shangguan ZP (2016) The coupling interaction of soil water and organic carbon storage in the long vegetation restoration on the Loess Plateau. Ecological Engineering 91: 574–581. https://doi.org/10.1016/j.ecoleng.2016.03.033

Zhang YY, Chen YB, Chen LF, Xu SC, Sun HZ (2022) A machine learning-based approach for generating high-resolution soil moisture from SMAP products. Geocarto International 37(27): 1–22. https://doi.org/10.1080/10106049.2022.2105406

Žliobaitė I, Bifet A, Read J et al (2015) Evaluation methods and decision theory for classification of streaming data with temporal dependence. Machine Learning 98: 455–482. https://doi.org/10.1007/s10994-014-5441-4

Zouzou Y, Citakoglu H (2023) General and regional cross-station assessment of machine learning models for estimating reference evapotranspiration. Acta Geophysica 71: 927– 947.

https://doi.org/10.1007/s11600-022-00939-9