iForest
Biogeosciences and Forestry

Research Article
doi: 10.3832/ifor4758-018
vol. 19, pp. 186-194

Application of machine learning models and Euclidean distance to predict soil spatial properties

Najmeh Ansari [(1)],
Mostafa Moradi [(1)],
Ruhollah Taghizadeh-Mehrjardi [(2)]

Accurate estimation of the spatial properties of forest soil is essential for sustainable land management. This research aimed to map key soil properties in temperate forests using hybrid machine learning models. The integration of predictive and geostatistical techniques has gained prominence in soil science, with hybrid approaches enhancing prediction accuracy. Precise soil maps serve as a foundational resource for effective soil management, motivating the development of more accurate and cost-efficient mapping methods. We used hybrid machine-learning techniques that incorporated Euclidean distance (Dis), remote sensing (RS) data, and digital elevation models (DEM) to improve spatial predictions. We hypothesized that integrating Euclidean distance data into predictive models would boost the accuracy of soil property maps. Model performance was evaluated using root-mean-square error (RMSE), coefficient of determination (R^2), mean absolute error (MAE), and concordance correlation coefficient (CCC). The hybrid RF+GA+Dis model offered the highest accuracy for predicting soil organic carbon (RMSE = 1.810, R^2 = 0.763, MAE = 1.883, CCC = 0.803), with similar improvements observed for soil nitrogen, organic carbon stock, bulk density, calcium carbonate, and soil texture. Hybrid models consistently demonstrated a superior performance over individual machine learning methods such as Random Forest (RF) and Genetic Algorithm (GA). Incorporating ancillary data, especially from DEM and RS, substantially ameliorated prediction accuracy for soil physicochemical properties. Among the tested models, those using a broader range of covariates demonstrated a superior performance, with RF+GA+Dis outperforming RF+Dis. These findings confirm that integrating machine learning with comprehensive spatial data is a cost-effective and reliable approach for generating high-resolution soil maps, facilitating precision land management and informed decision-making. This highlights the value of hybrid modeling and diverse covariates in advancing soil property prediction.

Keywords: Machine Learning, Euclidean Distance, Forest Soil, Ancillary Data

Introduction

Studies on forest soil physicochemical properties are essential because they reflect the influence of plant communities, root systems, and litterfall on soil characteristics (Veen et al. 2019, Moradi et al. 2017, Forogh Nasab et al. 2020). Further, plant communities have different effects on soil microorganisms (Li et al. 2025), which, in turn, influence the soil properties, including organic matter decomposition and nutrient availability (Veen et al. 2019). Meanwhile, forest soils contain a considerable amount of carbon and organic matter. Thus, it is essential to understand soil properties and identify the impact of management on forest resources (Moradi et al. 2022, Behmanesh et al. 2024, Klein-Raufhake et al. 2024).

Different ecosystems, each characterized by unique conditions and plant species, contribute to variations in soil properties. Accordingly, researchers strive to understand these properties and the factors that drive their changes (Setälä et al. 2016). Nevertheless, due to the complexity and wide range of influencing factors, examination of soil physicochemical properties is often time-consuming and expensive. As such, there is a growing demand for efficient techniques to estimate the spatial variability of soil properties.

The importance of soil spatial properties lies in their sensitivity to factors such as bedrock composition (Davatgar et al. 2012), land use (Wan et al. 2019), human activities (Moradi Behbahani et al. 2017), and runoff (Hashemi Rad et al. 2018). Nevertheless, spatial properties of plant species could also be derived from soil nutrient properties (Wan et al. 2019, John et al. 2007). Hence, given the interaction between plants and soil, soil spatial property studies could identify potential lands for afforestation (Wan et al. 2019) and provide information about the presence or absence of plant species.

In current soil science research, geostatistics and soil map production constitute crit-

(1) Department of Forestry, Faculty of Natural Resources, Behbahan Khatam Alanbia University of Technology, Behbahan, Khuzestan (Iran); (2) Faculty of Agriculture and Natural Resources, Ardakan University, Yazd (Iran)

@ Mostafa Moradi (moradi4@gmail.com)

Received: Nov 14, 2024 - Accepted: Oct 25, 2025

Citation: Ansari N, Moradi M, Taghizadeh-Mehrjardi R (2026). Application of machine learning models and Euclidean distance to predict soil spatial properties. iForest 19: 186-194. - doi: 10.3832/ifor4758-018 [online 2026-06-02]

Communicated by: Lorenzo MW Rossi

ical aspects that not only serve as fundamental requirements for monitoring soil physicochemical properties but also help identify optimal soil sampling locations (Ogunwole et al. 2014). Machine learning has gained substantial traction across various scientific fields (Wadoux et al. 2020). This popularity stems from the machine learning ability to generate soil property prediction maps with greater accuracy than traditional geostatistical methods, such as ordinary kriging. Further, combining machine learning with ordinary kriging yields even more accurate results (Tziachris et al. 2019). Nowadays, machine learning methods are widely regarded as promising for generating accurate soil maps worldwide. They have provided accurate maps of soil organic carbon (Parvizi & Fatehi 2025), soil particle maps (Kassai et al. 2025), and different soil types (Manteghi et al. 2024).

Among hybrid geostatistical models, regression kriging is particularly noteworthy, as it merges soil property data with DEM and remote sensing information (Cianfrani et al. 2018). Regression kriging and kriging artificial neural networks are widely employed in hybrid geostatistical modeling (Cheng et al. 2004, Hengl et al. 2007). Regression kriging, in particular, excels in estimating soil organic carbon and nitrogen in large-scale, homogenous ecosystems (Bangroo et al. 2020). RF can accurately predict soil organic carbon quantity (Parvizi & Fatehi 2025). In addition, machine learning is an effective technique, not only for ascertaining soil properties and degradation (Parajuli et al. 2025), but also for studying soil nutrient recovery in plantations (Pandey et al. 2025).

The application of machine learning models in soil science is most evident in digital soil mapping and the analysis of infrared spectral data, enabling more precise predictions of soil types and properties (Padarian et al. 2020). Machine learning methods consistently outperform traditional soil mapping approaches (Zolfaghari Nia et al. 2023). This combination of machine learning and geostatistics marks a significant contribution to advancing soil science, offering a promising pathway to improved soil monitoring and management. Despite advances in machine learning algorithms and software packages, there is no single best general algorithm for mapping soil properties. Accordingly, researchers continue to develop improved, compatible, and cost-effective algorithms.

The combination of predictive and geostatistical methods has become increasingly popular in soil science, with hybrid methods proving effective in improving the accuracy of soil property predictions (Song et al. 2017, Chen et al. 2019, Yang et al. 2019). In light of these developments, digital soil mapping (DSM) has evolved to enhance accuracy and refine methodologies. Accurate soil maps provide a fundamental framework for soil management, prompting researchers to seek more precise and cost-effective mapping techniques. This study aims to generate accurate soil maps by applying hybrid machine learning methods that integrate distance data (particularly Euclidean distance) with remote sensing and digital elevation model (DEM) data. Through integrating these diverse data sources with machine learning models, this approach aims to enhance the selection of optimal methods for generating soil maps with reduced error (Tziachris et al. 2019). We hypothesize that incorporating Euclidean distance data into predictive models improves the spatial mapping of soil properties. In addition, combining Euclidean distance, remote sensing, and DEM data with machine learning models is suggested as the most effective approach for predicting soil properties in temperate forests.

Materials and methods

Study site

The study site was located in Kheyrud Experimental Forest at the University of Tehran, Mazandaran province, northern Iran. This forest covers 8,000 ha and is located at a latitude of 36° 27′ N to 36° 40′ N and a longitude of 51° 32′ E to 51° 43′ E (Fig. 1). The mean annual temperature is 8.6 °C, and the annual precipitation is 1380 mm. Two uneven-aged parcels (236 and 327) dominated by *Fagus orientalis* were chosen in the present study. Other species, including *Carpinus betulus* and *Acer velutinum*, were recorded at the studied sites (Sefidi & Mohadjer 2010). This site was selected because it is similar to North American and East Asian temperate forests but has lower levels of human disturbance and exploitation (Sagheb Talebi et al. 2013).

Soil sampling and analysis

Soil sampling was performed using a regular grid design with a spacing of 100 × 50 m. A total of 126 soil samples were collected from the 0-10 cm mineral soil layer, following removal of surface litter and woody debris. The samples were air-dried and subsequently prepared for physicochemical analysis. The contents of soil nitrogen (Bremner & Mulvaney 1982), potassium (Merwin & Peech 1951), phosphorus (Olsen et al. 1954), and soil organic carbon (Walkley & Black 1951) were determined. In addition, a cylindrical metal core sampler (4.8 cm diameter by 10.1 cm high) was used to measure soil bulk density (Page-Dumroese et al. 1999), and soil texture was also determined using a hydrometer method (Holliday 1990).

Ancillary data

Landsat-8 (30 m resolution) and Sentinel-2 (10, 20, and 60 m resolution) satellite images were utilized as ancillary data in this study, providing valuable spectral information across multiple bands that help detect variations in vegetation cover, moisture content, and mineral composition, all being crucial indicators of soil properties (Tab. 1). Sentinel-2 (10, 20, and 60 m resolution) was used to leverage the full spectral range and enhance the model performance. Prior to analysis, all bands were resampled to a resolution of 30 m using bilinear interpolation to ensure spatial consistency with our soil sampling grid. This approach aligns with current practices in digital soil mapping, where spectral richness often outweighs small variations in native spatial resolution, especially at local-to-regional scales (Kebonye et al. 2024, Wang et al. 2024). The use of multiple bands with varying resolutions enables the inclusion of

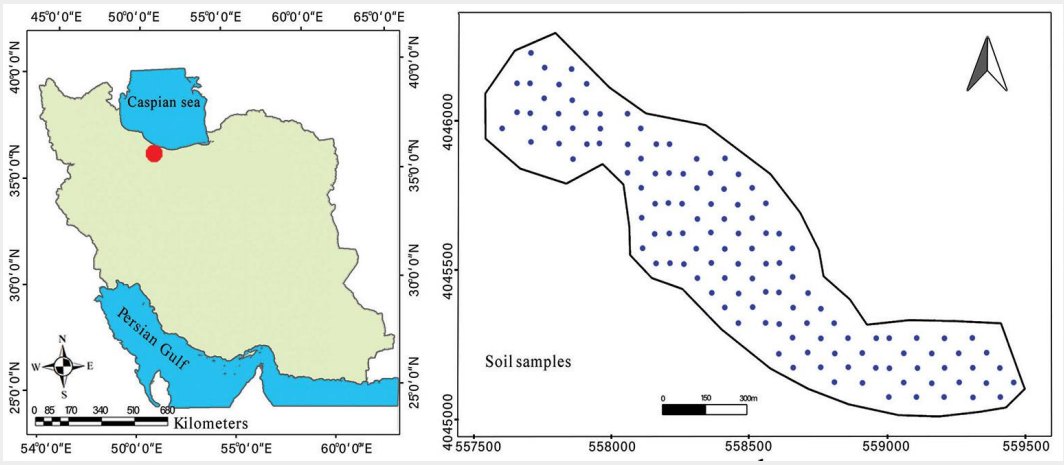

Fig. 1 - Location of the study site (left panel, in red) and spatial distribution of sampling points (right panel, blue dots).

valuable information, such as vegetation indices, soil brightness, and moisture content, which are key factors influencing soil properties. In addition to these spectral datasets, DEM data, including slope and Euclidean distance, were incorporated. Slope data are necessary for understanding soil erosion and deposition patterns, as well as hydrological processes, which significantly influence soil nutrient distribution. Euclidean distance can ameliorate the soil property prediction (Sun et al. 2017); thus, it was used to measure the proximity of each soil sample to specific landscape features, such as streams, roads, and vegetation zones, often correlating with soil property variations owing to the influence of water flow, anthropogenic activities, and ecological transitions.

We employed a 30 m-resolution DEM derived from the Shuttle Radar Topography Mission (SRTM) because our target soil maps were generated at 30 m resolution. Various modeling scenarios were developed and tested to optimize soil property prediction. The modeling scenarios were designed to systematically assess the contribution of different covariate sets to model performance and to test the hypothesis that including spatial metrics, such as Euclidean distance, improves prediction accuracy. In digital soil mapping (DSM), scenario-based testing is a recognized approach to exploring the relative importance of different covariate groups and their interactions (Nabiollahi et al. 2019). The introduction of Euclidean distance in the latter scenarios was informed by studies showing that it accounts for spatial relationships between soil samples and landscape features, often exhibiting strong spatial autocorrelation with soil properties (Behrens et al. 2018).

The first scenario used a random forest (RF) model in conjunction with ancillary data (satellite images and DEMs). The second scenario introduced a hybrid model that integrates random forest with a genetic algorithm (GA) and ancillary data (satellite images and DEM). The third scenario used a RF with extended ancillary data, incorporating satellite images, DEMs, and Euclidean distances. The final scenario employed a hybrid model that combined RF, GA, and extended ancillary data (satellite images, DEM, and Euclidean distance). By incorporating Euclidean distance data, the third and fourth scenarios aimed to capture the spatial variability associated with proximity to landscape features. After ascertaining the results across these four scenarios, the best-performing model was selected to generate the soil property maps.

Tab. 1 - Ancillary data used for soil properties estimation.

Terrain attributes (Name)	Terrain attributes (Name)
Analytical hill-shading	Sentinel 2.6
Aspect	Sentinel 2.7
Catchment Area	Sentinel 2.8
Catchment Slope	Sentinel 2.9
Channel Network Base Level	Sentinel 2.10
Channel Network Distance	NDVI. S
Convergence Index	EVI. S
Cross-Sectional Curvature	TVI. S
Generalized Surface	SAVI. S
Gradient	LSWI. S
Longitudinal Curvature	Brightness Index S
LS-factor	Clay Index S
Maximum Curvature	Salinity Index S
Minimum Curvature	Carbonate Index S
Modified Catchment Area	Gypsum Index S
Morphometric Features	Landsat 8.1
MRRTF	Landsat 8.2
MRVBF	Landsat 8.3
Negative Openness	Landsat 8.4
Plan Curvature	Landsat 8.5
Positive Openness	Landsat 8.6
Profile Curvature	NDVI. L
Relative Slope Position	EVI.L
Slope	SAVI.L
Topographic Wetness Index	NDMI.L
Topographic Wetness Index 1	COSRI. L
Total Catchment Area	Brightness Index L
Valley Depth	Clay Index L
Wind Effect	Salinity Index L
Sentinel 2.1	Carbonate Index L
Sentinel 2.2	Gypsum Index L
Sentinel 2.3	-
Sentinel 2.4	-
Sentinel 2.5	-

Machine learning models

The genetic algorithm is an optimization technique based on natural selection and genetic principles, inspired by the biological process of evolution and demonstrating Darwin's theory of survival (Michalewicz & Schoenauer 1996). The integration of GA with geographic information systems (c) considerably enhances data analysis (Mansor et al. 2012). By selecting the most relevant variables, GA improves predictive model performance (Xie et al. 2015). This technique has been widely applied in soil classification and property mapping (Xie et al. 2015, Tziachris et al. 2019) because of its robustness and ability to avoid local minima (Emadi et al. 2020).

The RF model, on the other hand, employs an ensemble of classification and regression trees to improve predictive accuracy and reduce overfitting (Breiman 2001). As a tree-based model, it identifies the most informative variable at each node split, making it resilient to outliers. This robustness, speed, and ease of use make random forest a preferred method in predictive modeling (Cutler et al. 2007).

Despite the performance ofmachine learning models in predicting soil properties, we need more alternatives to generate more accurate maps (Zeraatpisheh et al. 2019). Hence, in addition to standalone RF and GA models, this study also utilized two hybrid machine learning methods: (i) Random Forest combined with Distance (RF+Dis): Incorporates Euclidean distance data to improve spatial variability prediction; (ii) Random Forest combined with Genetic Algorithm and Distance (RF+GA+Dis): Integrates GA for variable selection and Euclidean distance data to optimize the prediction of soil properties.

The inclusion of Euclidean distance is a recent method that causes spatial autocorrelation, thus generating more accurate digital soil maps (Behrens et al. 2018). Hence, the hybrid models used in the present study leverage the strengths of both GA and RF to improve predictive accuracy, par-

ticularly for mapping spatial variations in soil properties.

The performance of each studied model was judged by RSME, MAE, R^2, CCC, and error. These parameters are defined as follows:

$$RMSE = \sqrt{\sum_{i=1}^{n} (Pi - Oi)^2} \quad (1)$$

$$MAE = \frac{1}{n} \sum_{i=1}^{n} |Pi - Oi| \quad (2)$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (Oi - p)^2}{\sum_{i=1}^{n} (Oi - \hat{O})^2} \quad (3)$$

$$CCC = \frac{2 r \sigma_o \sigma_p}{\sigma_o^2 + \sigma_p^2 + [O' - P']^2} \quad (4)$$

where n represents the number of samples; Oi and Pi denote observed and predicted soil variables, respectively; O' and P' show the means for the observed and predicted soil variables, respectively; variances of observed and predicted values are represented by σ_o and σ_p.

K-fold cross-validation techniques (10-fold cross-validation) were applied for model validation (Kebonye et al. 2024). Cross-validation is a robust method for evaluating model performance, ensuring that all available samples are used for both training and testing. In *K*-fold cross-validation, the dataset is randomly split into *K* equally sized subsets, or folds. Along each iteration, one fold is used as the test set, while the remaining *K*-1 folds are utilized as the training set. This process is repeated *K* times, ensuring that each fold is utilized as the test set exactly once. The performance metrics from each *K* iteration are then averaged to provide a comprehensive assessment of the model's accuracy. Using all data points for training and testing, this method ensures that multiple samples and scenarios are ascertained to lower the risk of overfitting and provide a more accurate measure of generalization. Cross-validation is essential for mitigating overfitting since the model is trained and tested on different subsets of the data, helping identify models that may perform well on the training set but poorly on unseen data. It also provides a more reliable estimate of model performance than a simple train-test split, by averaging the results over multiple iterations. It also aids in optimal model selection by comparing the performance across different models or parameter configurations. Overall, *K*-fold cross-validation enhances the reliability and generalizability of predictive models, ensuring that the selected models accurately forecast soil properties across diverse landscapes.

Further, we employed 2500 trees (ntree = 2500), with the *mtry* value optimized *via* cross-validation (typically set to the square root of the number of predictors) in the RF model. A population size of 50, crossover probability of 0.8, mutation rate of 0.01, and a maximum of 100 generations were applied in GA. Since the RF model is not sensitive to variable scaling, and categorical variables were not included, no standardization was applied. Nevertheless, all continuous covariates were checked for outliers and normalized if extreme values were present, to avoid bias in interpreting feature importance.

The entire data analysis and modeling was performed using the R programming language (version 4.2.2). We used the "randomForest" package for RF modeling and the "GA" package to implement the genetic algorithm. Data preprocessing and visualization were performed using packages such as "caret", "raster", and "ggplot2".

Results

Soil properties overview

Our results revealed that the mean soil nitrogen content at the study site was 0.31% (Tab. 2), whereas the mean soil organic carbon (SOC) was 7.7%. The average soil calcium carbonate ($CaCO_2$) content was 1%, and the mean soil bulk density was 1.38 g cm^{-3}, with values ranging from 1.16 to 1.66 g cm^{-3} (Tab. 2). The mean percentages of soil clay, silt, and sand were 24.8%, 31.6%, and 43.5%, respectively (Tab. 2). The mean soil carbon stock (SCS) was 107.5 tons per hectare, ranging from 22.6 to 203 t ha^{-1} (Tab. 2).

Tab. 2 - The summary of soil physicochemical properties in the studied site (SOCS: soil organic carbon stock).

Soil properties	Minimum	Maximum	Mean	Std. Deviation
Organic carbon (%)	1.74	15.40	7.70	2.53
Nitrogen (%)	0.14	0.70	0.31	0.09
$CaCO_3$ (%)	0.20	2.30	1.09	0.39
Silt (%)	8.00	49.44	31.6	7.47
Sand (%)	16.96	84.16	43.55	12.89
Clay (%)	6.56	45.60	24.82	7.73
Bulk density (g cm^{-3})	1.16	1.66	1.38	0.07
SOCS (ton ha^{-1})	22.62	234.08	107.49	37.55

Machine learning models

Tab. 2 outlines the performance results of different machine learning algorithms for soil property prediction. Model performance was evaluated using RMSE, MAE, R^2, CCC, and other error metrics. Given lower error values and higher R^2, our findings indicate that the hybrid model RF+GA+Dis generated the most accurate estimates of soil physicochemical properties (Tab. 3). In contrast, the non-hybrid models revealed higher error levels and lower coefficients of determination (Tab. 3). Although two hybrid models were tested in this study, the results showed that the RF+Dis model was less accurate than the RF+GA+Dis model in estimating soil properties. Nevertheless, the RF+Dis model outperformed non-hybrid models (Tab. 3).

Hybrid model of DEM+Rs

Applying the hybrid model, incorporating DEM and RS data, yielded insightful results for forecasting various soil properties. For SOC, the wind effect was the most significant covariate in lowering estimation errors. The wind effect, often modeled using indices such as wind fetch or wind exposure, is essential for determining soil erosion and deposition patterns, thereby influencing the spatial distribution of organic matter. The brightness index (S), a measure of soil reflectance, and Sentinel 2.6, a specific spectral band from the Sentinel-2 satellite, also played significant roles in enhancing SOC predictions (Fig. 2A). The brightness index helps distinguish soil types based on surface reflectance, while Sentinel 2.6, which is sensitive to soil mineralogy, aids in capturing variations in soil organic content. On the contrary, Sentinel 2.8 and NDVI.S (Normalized Difference Vegetation Index for Sentinel) contributed the least to mitigating the estimation error. This suggests that vegetation cover, as represented by NDVI, and certain spectral bands are less effective in SOC prediction within this hybrid model.

In soil nitrogen (N) prediction, the wind effect remained the most influential variable given its impact on soil nutrient redistribution through erosion and deposition (Fig. 2A). Normalized Difference Moisture Index for Landsat (NDMI.L) and aspect, which represents the directional slope of the terrain, had minimal influence on nitrogen estimation. The channel network base level (CNBL), representing the hierarchical position of streams, and aspect significantly ameliorated the estimation of soil $CaCO_2$ by reflecting landscape drainage and aspect-induced microclimatic variations (Fig. 2A). On the other hand, the modified catchment area (MCA), indicating drainage areas, and profile curvature, measuring slope convexity and concavity, contributed little to error reduction. Sentinel 2.6 and

Tab. 3 - Performances of different machine learning models in estimating soil physicochemical properties (Soil prop.). (SOC): soil organic carbon; (N): nitrogen; (BD): bulk density; (SOCS): soil organic carbon stock; (RF): Random Forest; (GA): Genetic Algorithm; (DIS): Euclidean distance.

Soil prop.	ML models	RMSE	MAE	R^2	CCC	%Error	Soil prop.	ML models	RMSE	MAE	R^2	CCC	%Error
SOC (%)	RF	2.488	2.055	0.304	0.344	32.299	Sand (%)	RF	13.017	10.970	0.440	0.480	29.886
	RF+GA	2.432	2.035	0.507	0.547	31.570		RF+GA	12.644	10.818	0.511	0.551	29.029
	RF+DIS	2.124	1.921	0.616	0.656	27.565		RF+DIS	10.292	8.922	0.578	0.618	23.629
	RF+DIS+GA	1.810	1.883	0.763	0.803	23.497		RF+DIS+GA	8.962	6.190	0.691	0.731	20.574
N (%)	RF	0.080	0.065	0.329	0.369	25.627	Clay (%)	RF	8.365	7.102	0.395	0.435	33.699
	RF+GA	0.078	0.065	0.527	0.567	24.966		RF+GA	8.290	7.113	0.524	0.564	33.399
	RF+DIS	0.071	0.054	0.605	0.645	22.926		RF+DIS	6.912	6.007	0.557	0.597	27.845
	RF+DIS+GA	0.068	0.047	0.711	0.751	21.787		RF+DIS+GA	5.849	5.393	0.680	0.720	23.566
$CaCO_2$ (%)	RF	0.366	0.309	0.384	0.424	33.388	BD (gr cm^{-3})	RF	0.078	0.065	0.372	0.412	5.641
	RF+GA	0.358	0.307	0.474	0.514	32.607		RF+GA	0.073	0.062	0.434	0.474	5.289
	RF+DIS	0.314	0.293	0.579	0.619	28.662		RF+DIS	0.070	0.064	0.579	0.619	5.076
	RF+DIS+GA	0.250	0.243	0.602	0.642	22.758		RF+DIS+GA	0.069	0.063	0.651	0.691	4.961
Silt (%)	RF	7.197	5.990	0.327	0.367	22.765	SOCS (ton ha^{-1})	RF	37.197	30.425	0.507	0.547	34.603
	RF+GA	6.813	5.910	0.479	0.519	21.551		RF+GA	34.841	29.293	0.449	0.489	32.411
	RF+DIS	6.300	5.521	0.518	0.558	19.928		RF+DIS	31.940	28.853	0.622	0.662	29.712
	RF+DIS+GA	5.175	5.170	0.648	0.688	16.369		RF+DIS+GA	24.848	24.898	0.641	0.681	23.115

Sentinel 2.9 were identified as the most effective covariates for predicting the soil silt content (Fig. 2A). In contrast, Landsat 8.3 and profile curvature significantly lowered the prediction error for soil sand and clay, as Landsat 8.3 captures soil texture characteristics, whereas profile curvature reflects the erosional and depositional processes affecting soil texture. For soil bulk density estimation, Landsat 8.5 and NDVI.S proved highly effective in reducing prediction error, suggesting that vegetation indices and spectral bands are strong predictors of bulk density. Wind effect and CNBL were the most critical covariates in estimating soil carbon stock, given their influence on nutrient accumulation and landscape drainage patterns (Fig. 2B).

Hybrid model of DEM+Rs+Dis

Integration of the distance variable into the hybrid model further improved predictions of soil properties. For soil organic matter estimation, the wind effect remained the dominant variable, underscoring the significance of erosion and deposition patterns in organic matter distribution. The channel network base level (CNBL) and Sentinel 2.3, a spectral band that captures soil and vegetation characteristics, also significantly contributed to SOC estimation (Fig. 3A). CNBL reflects drainage patterns, while Sentinel 2.3 helps distinguish soil and

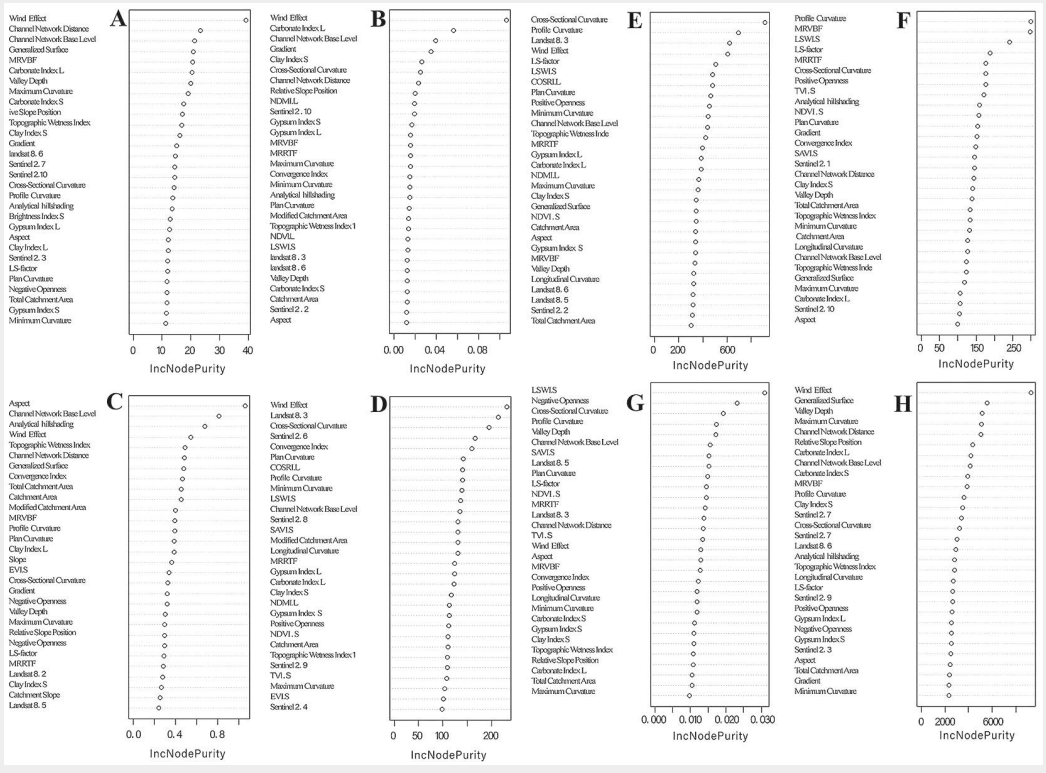

Fig. 2 - Variables for soil physiochemical properties in the hybrid model DEM+RS (A: soil organic carbon; B: soil nitrogen; C: soil calcium carbonate; D: soil silt; E: soil sand; F: soil clay; G: bulk density; H: soil carbon stock).

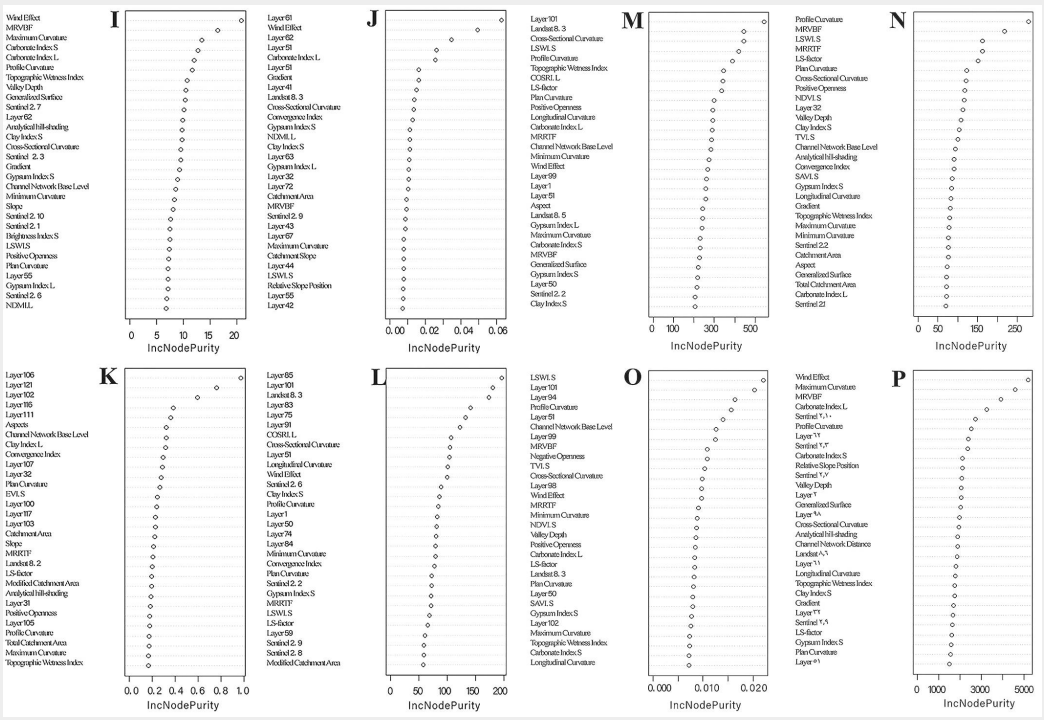

Fig. 3 - Variables for soil physiochemical properties in the hybrid model DEM+RS+Dis (I: soil organic carbon; J: soil nitrogen; K: soil calcium carbonate; L: soil silt; M: soil sand; N: soil clay; O: bulk density; P: soil carbon stock).

vegetation types, which are closely related to SOC levels.

In predicting soil nitrogen, layer 61, layer 56 (spectral bands from remote sensing data), and the wind effect were the best predictors (Fig. 3A). Layers 61 and 56 capture specific spectral signatures associated with the nitrogen content. For soil calcium carbonate estimation, layers 106, 102, and CNBL were the most influential variables (Fig. 3A). Layers 106 and 102 capture mineralogical signatures correlating with $CaCO_2$, whereas CNBL represents landscape drainage, which affects $CaCO_2$ accumulation.

For soil silt estimation, layers 65, 91, and 83 were the most impactful in reducing error (Fig. 3A). These layers, which represent specific spectral bands, are sensitive to soil textural properties. Soil sand estimation benefited most from layer 1, Landsat 8.1, and profile curvature (Fig. 3B). Layer 1 captures soil surface characteristics, Landsat 8.1 provides texture-specific spectral information, and profile curvature reflects erosion and deposition processes that influence soil texture.

In soil clay estimation, Sentinel 2.4, Topographic Wetness Index 1 (TWI1), and CROSRI.L (Cross-Slope Roughness Index for Landsat) were the most effective variables (Fig. 3B). Sentinel 2.4 identifies clay minerals, TWI1 evaluates the soil moisture content, and CROSRI.L reflects surface roughness owing to clay dispersion.

For soil bulk density, negative openness, Landsat 8.1, LSWIS (Landsat Soil Wetness Index Sentinel), and layer 62 were identified as the most important covariates (Fig. 3B). Negative openness measures landscape concavity, affecting soil compaction, while LSWIS and layer 62 capture soil moisture and texture characteristics. For soil carbon stock estimation, layer 62 and the wind effect emerged as the most crucial variables (Fig. 3B). Layer 62 presents spec-

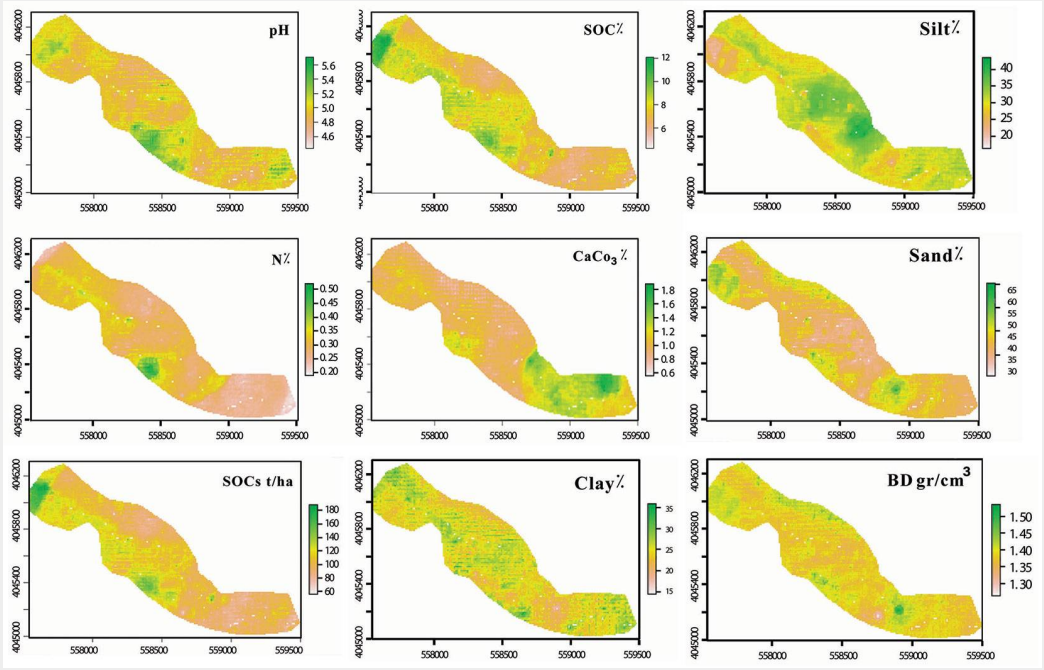

Fig. 4 - Soil spatial property maps using a hybrid model of DEM+RS+Dis (Soil pH, organic carbon, nitrogen, and calcium carbonate; soil silt, sand, clay; bulk density; and soil carbon stock).

tral data associated with the soil organic content, and the wind effect influences the distribution of soil organic carbon through erosion and deposition processes.

Soil physicochemical spatial estimation

Given the superior predictive performance of the hybrid model incorporating DEM, RS, and Dis variables, comprehensive soil property maps were generated using this approach. The SOC map (Fig. 4A) indicates that the highest organic matter values are found in the southern and southwestern parts of the study area, reaching approximately 12%. In contrast, the northern, central, and eastern regions exhibited the lowest SOC levels, with values below 6%. High SOC values typically represent fertile soils rich in nutrients with better water retention, which can promote healthy plant growth. Conversely, low SOC levels can result in poor soil structure, reduced nutrient availability, and increased susceptibility to erosion.

The soil nitrogen map (Fig. 4A) shows that the eastern and northern parts of the study area have the lowest nitrogen values, around 0.2%, whereas the highest levels, approximately 0.6%, are concentrated in the south-central region. High nitrogen concentrations are often linked to greater soil fertility and productivity, supporting vigorous plant growth. Conversely, excessively low nitrogen levels can lead to nutrient deficiencies, resulting in stunted growth and reduced crop yields.

Soil $CaCO_2$ content was highest in the eastern part of the study area, exceeding 1.8% (Fig. 4A). Elevated $CaCO_2$ levels can improve soil structure and buffer acidity, boosting overall soil health. However, excessively high levels may reflect calcareous soils, which can inhibit nutrient availability for certain crops. The soil carbon stock (SCS) map (Fig. 4B) reveals that the southwestern region of the site has had the highest values, reaching up to 180 t ha^{-1}. High carbon stock suggests soils with abundant organic matter, thereby improving fertility, water retention, and climate change mitigation. In contrast, low carbon stock values indicate soils that are less capable of supporting sustainable agricultural practices and may require organic matter amendments.

The soil physical properties map (Fig. 4B) indicates that the central region of the study area contains the highest silt content, whereas the sand content is lowest in the same area. High silt content can improve soil moisture retention and nutrient-holding capacity, though excessive amounts can lead to poor drainage. In contrast, high sand content promotes good drainage but often causes poor nutrient retention. The highest clay content was found across the eastern, western, and central regions of the study area. Soils with high clay content typically have excellent water and nutrient retention but may also experience poor drainage and compaction.

The soil bulk density map (Fig. 4B) reveals that bulk density does not follow a specific spatial pattern, with no extensive, homogeneous areas with similar values. High bulk density generally corresponds to compacted soils, which can impede root growth, reduce water infiltration, and limit soil aeration. Conversely, low bulk density typically signifies well-structured soils that are more conducive to plant growth.

These comprehensive soil property maps, generated using the hybrid model of DEM+RS+Dis, provide critical insights into the spatial distribution of key soil attributes and serve as a valuable tool for precision agriculture, land management, and sustainable environmental practices.

Discussion

In the present study, different machine learning models and hybrid models were employed to provide spatial maps of soil properties in temperate forests. Our results revealed that the hybrid model RF+GA+Dis achieved the best performance in estimating soil properties, as evidenced by a higher coefficient of determination and lower error. Given the high accuracy of machine learning methods, scientists are increasingly applying them to generate soil property maps (Song et al. 2017, Szabó et al. 2019, John et al. 2020, Emadi et al. 2020, Bouslihim et al. 2021, Zolfaghari Nia et al. 2023). Hybrid models have been widely used to reduce feature dimensionality and overall costs. It seems that even a hybrid model can provide such conditions, but adding the Euclidean distance to the hybrid models can offer more accurate maps. This is because Euclidean distance can introduce spatial autocorrelation into the input data (Kim et al. 2023, Behrens et al. 2018).

Comparing non-hybrid machine learning models (RF and AG) with hybrid models (RF+GA+Dis and RF+Dis), non-hybrid models produced higher errors and lower coefficients of determination. This indicates the importance of using hybrid models to generate soil property maps. This finding is consistent with other studies highlighting the importance of the hybrid model for generating more accurate soil maps (Chen et al. 2019, Tziachris et al. 2019). The need to generate highly accurate maps has long been recognized to ameliorate soil spatial property maps, and in the present study, using Euclidean distance yields more accurate maps, which is in accordance with the findings of Kim et al. (2023).

Studies demonstrate that machine learning can generate highly accurate soil spatial properties maps (John et al. 2020, Bouslihim et al. 2021, Zolfaghari Nia et al. 2023). However, our results indicated that hybrid models can produce far better outcomes. In the present study, even hybrid models with more parameters generated more accurate maps than the hybrid model with fewer parameters. The hybrid model RF+Dis yielded less accurate results than RF+GA+Dis, but provided better results when compared to the RF and GA as separate models.

Nowadays, ancillary data are considered a significant factor in enhancing the accuracy of soil spatial maps and have been used in many studies (Filippi et al. 2018, Chatterjee et al. 2020). Further, other studies have highlighted its importance in achieving better results (Marchetti et al. 2008, Piccini et al. 2020). In the present study, we used DEM and RS data as ancillary variables to predict soil properties. The best ancillary variable for predicting each soil property differed between the hybrid models. However, wind effect was the most influential ancillary variable for predicting soil properties in both hybrid machine learning models. Ancillary data contributed to improvements in soil spatial mapping, aligning with findings from other studies (Dai et al. 2014, Filippi et al. 2018). Meanwhile, using RS and DEM data as ancillary inputs revealed that these datasets can yield different results in forecasting soil properties. Each ancillary variable may behave differently when predicting various soil properties, which is consistent with the findings of Mousavi et al. (2020).

In our study, soil property mapping was performed using the hybrid model RF+GA+Dis. According to the maps, the highest levels of soil nitrogen and organic matter were recorded in the western and southwestern parts of the study site, while the lowest levels were observed in the eastern part. This variation is attributed to differences in tree species composition. Pure *Fagus orientalis* stands are found in the eastern part, whereas mixed stands of *Fagus orientalis* and *Carpinus betulus* are present in the western part of the site. The differences in soil nitrogen and organic matter are a consequence of these varying tree species. This finding is consistent with the results of Habashi (2015), who reported lower organic matter in pure *Fagus orientalis* stands than in mixed stands of *Fagus orientalis* and *Carpinus betulus*. The hybrid model, when combined with appropriate ancillary data, generated accurate maps, even distinguishing among different forest stand types. Our results indicate that a hybrid machine learning model, combined with Euclidean distance, can accurately predict soil properties even in temperate forests, aligning with the findings of Behrens et al. (2018).

Conclusion

Overall, hybrid models offered significantly higher accuracy than standalone machine learning models such as RF and GA. The inclusion of ancillary data, particularly from DEM and RS, markedly boosted the predictive accuracy of soil physico-chemical properties. Among the hybrid models, those incorporating a broader range of parameters demonstrated a superior performance. Specifically, the hybrid model combining RF, GA, and Dis variables outperformed the RF+Dis model. Further, the

RF+Dis hybrid model yielded better results than the individual models. Integrating machine learning models with ancillary data from DEM and RS provided a cost-effective solution for generating high-accuracy soil spatial property maps, enabling precise and sustainable land management practices. The results highlight the value of combining complementary machine learning approaches with a comprehensive set of covariates to optimize soil property predictions, paving the way for improved soil management strategies and better-informed decisions.

References

Bangroo SA, Najar GR, Achin E, Truong PN (2020). Application of predictor variables in spatial quantification of soil organic carbon and total nitrogen using regression kriging in the North Kashmir forest Himalayas. Catena 193: 104632. - doi: 10.1016/j.catena.2020.104632

Behmanesh S, Moradi M, Pourrezaei J, Basiri R (2024). Does road construction have beneficial effects on vegetation biodiversity and tree regeneration in arid woodlands? Land Degradation and Development 35 (7): 2508-2517. - doi: 10.1002/ldr.5076

Behrens T, Schmidt K, Rossel RV, Gries P, Scholten T, MacMillan RA (2018). Spatial modelling with Euclidean distance fields and machine learning. European Journal of Soil Science 69 (5): 757-770. - doi: 10.1111/ejss.12687

Bouslihim Y, Rochdi A, El Amrani Paaza N (2021). Machine learning approaches for the prediction of soil aggregate stability. Heliyon 7 (3): e06480. - doi: 10.1016/j.heliyon.2021.e06480

Breiman L (2001). Random forests. Machine Learning 45: 5-32. - doi: 10.1023/A:1010933404324

Bremner JM, Mulvaney CS (1982). Nitrogen total. In: "1. Methods of soil analysis. Part 2. Chemical and microbiological properties" (Page JL ed). Agronomy Society of America and Soil Science Society of America, Madison, WI, USA, pp. 595-624. - doi: 10.2134/agronmonogr9.2.2ed.c31

Chatterjee S, Huang J, Hartemink AE (2020). Establishing an empirical model for surface soil moisture retrieval at the U.S. climate reference network using Sentinel-1 backscatter and ancillary data. Remote Sensing 12: 1242. - doi: 10.3390/rs12081242

Chen L, Ren C, Li L, Wang Y, Zhang B, Wang Z, Li L (2019). A comparative assessment of geostatistical, machine learning, and hybrid approaches for mapping topsoil organic carbon content. ISPRS International Journal of Geo-Information 8: 174. - doi: 10.3390/ijgi8040174

Cheng XF, Shi XZ, Yu DS, Pan XZ, Wang HJ, Sun WX (2004). Using GIS spatial distribution to predict soil organic carbon in subtropical China. Pedosphere 14: 425-431. [online] URL: https://www.researchgate.net/publication/279578125

Cianfrani C, Buri A, Verrecchia E, Guisan A (2018). Generalizing soil properties in geographic space: approaches used and ways forward. PLoS One 13 (12): e0208823. - doi: 10.1371/journal.pone.0208823

Cutler DR, Edwards TC, Beard KH, Cutler A, Hess KT, Gibson J, Lawler JJ (2007). Random forests for classification in ecology. Ecology 88 (11): 2783-2792. - doi: 10.1890/07-0539.1

Dai FQ, Zhou QG, Lv ZQ, Wang XM, Liu GC (2014). Spatial prediction of soil organic matter content integrating artificial neural network and ordinary kriging in Tibetan Plateau. Ecological Indicator 45: 184-194. - doi: 10.1016/j.ecolind.2014.04.003

Davatgar N, Neishabouri M, Sepaskhah A (2012). Delineation of site-specific nutrient management zones for a paddy cultivated area based on soil fertility using fuzzy clustering. Geoderma 173: 111-118. - doi: 10.1016/j.geoderma.2011.12.005

Emadi M, Taghizadeh-Mehrjardi R, Cherati A, Danesh M, Mosavi A, Scholten T (2020). Predicting and mapping of soil organic carbon using machine learning algorithms in Northern Iran. Remote sensing 12 (14): 2234. - doi: 10.3390/rs12142234

Filippi P, Cattle SR, Bishop T, Jones EJ, Minasny B (2018). Combining ancillary soil data with Vis-NIR spectra to improve predictions of organic and inorganic carbon content of soils. MethodsX 5: 551-560. - doi: 10.1016/j.mex.2018.05.019

Forogh Nasab M, Moradi M, Moradi G, Taghizade-Mehrjardi R (2020). Topsoil carbon stock and soil physicochemical properties in riparian forests and agricultural lands of southwestern Iran. Eurasian Soil Science 53 (10): 1389-1395. - doi: 10.1134/S1064229320100075

Habashi H (2015). Microbial respiration and microbial biomass C relationship with soil organic matter in different types of mixed beech forest. Forest Research and Development 1 (2): 135-144.

Hashemi Rad S, Kiani F, Meftah Helghi M, Hematzadeh Y (2018). Effect of land use on physical and chemical parameters of soil and sediment in Qarnave and Yelcheshme watersheds, Golestan Province, Iran. Environmental Resources Research 2: 89-102. - doi: 10.22069/ijerr.2018.9866.1116

Hengl T, Heuvelink GBM, Rossiter DG (2007). About regression-kriging: from equations to case studies. Computers and Geosciences 33: 1301-1315. - doi: 10.1016/j.cageo.2007.05.001

Holliday VT (1990). Methods of soil analysis, Part 1, physical and mineralogical methods (2nd edn). A. Klute, Ed., 1986, American Society of Agronomy, Agronomy Monographs 9 (1), Madison, Wisconsin, 1188. Geoarchaeology 5 (1): 87-89. - doi: 10.1002/gea.3340050110

John K, Isong IA, Kebonye NM, Ayito EO, Agyeman PC, Afu SM (2020). Using machine learning algorithms to estimate soil organic carbon variability with environmental variables and soil nutrient indicators in an alluvial soil. Land 9: 487. - doi: 10.3390/land9120487

John R, Dalling JW, Harms KE, Yavitt JB, Stallard RF, Mirabello M, Hubbell SP, Valencia R, Navarrete H, Vallejo M, Foster RB (2007). Soil nutrients influence spatial distributions of tropical tree species. Proceedings of the National Academy of Sciences USA 104: 864-869. - doi: 10.1073/pnas.0604666104

Kassai P, Kocsis M, Szatmári G, Makó A, Mészáros J, Laborczi A, Magyar Z, Takacs K, Szabó B (2025). Large-scale mapping of soil particle size distribution using legacy data and machine learning-based pedotransfer functions. Geoderma 454: 117178. - doi: 10.1016/j.geoderma.2025.117178

Kebonye NM, Taghizadeh-Mehrjardi R, John K, Agyeman PC, Kakhani N, Seletlo Z, Motlhetlhi L, Moyo B, Scholten T (2024). Spatial scale drives pedodiversity-elevation relationship in Botswana. Geomatica 76 (2): 100037. - doi: 10.1016/j.geomat.2024.100037

Kim HJ, Mawuntu KBA, Park TW, Kim HS, Park JY, Jeong YS (2023). Spatial Autocorrelation Incorporated Machine Learning Model for Geotechnical Subsurface Modeling. Applied Sciences 13 (7): 4497. - doi: 10.3390/app13074497

Klein-Raufhake T, Hölzel N, Schaper JJ, Elmer M, Fornfeist M, Linnemann B, Meyer M, Neuenkamp L, Rentemeister K, Santora L, Jens W, Hamer U (2024). Disentangling the impact of forest management intensity components on soil biological processes. Global Change Biology 31 (1): e70018. - doi: 10.1111/gcb.70018

Li H, Meng Z, Ren X, Han Y (2025). Variations in the physicochemical properties of soil, enzyme activities, and the characteristics of bacterial communities within algal biocrusts and subsoils across different plant communities. Global Ecology and Conservation 58: e03455. - doi: 10.1016/j.gecco.2025.e03455

Mansor SB, Pormanafi S, Mahmud ARB, Pirasteh S (2012). Optimization of land use suitability for agriculture using integrated geospatial models and genetic algorithms. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences 2: 234-299. - doi: 10.5194/isprsannals-I-2-229-2012

Manteghi S, Moravej K, Mousavi SR, Delavar MA, Mastinu A (2024). Digital soil mapping for soil types using machine learning approaches at the landscape scale in the arid regions of Iran. Advances in Space Research 74 (1): 1-16. - doi: 10.1016/j.asr.2024.04.042

Marchetti A, Piccini C, Francaviglia R, Santucci S, Chiuchiarelli I (2008). Evaluation of soil organic matter content in Teramo province, Central Italy. Advances in GeoEcology 39: 345-356. [online] URL: http://www.cabidigitallibrary.org/doi/full/10.5555/20093074215

Merwin HD, Peech M (1951). Exchangeability of soil potassium in the sand, silt, and clay fractions as influenced by the nature of the complementary exchangeable cation 1. Soil Science Society of America Journal 15 (C): 125-128. - doi: 10.2136/sssaj1951.036159950015000C0026x

Michalewicz Z, Schoenauer M (1996). Evolutionary algorithms for constrained parameter optimization problems. Evolutionary Computation 4 (1): 1-32. - doi: 10.1162/evco.1996.4.1.1

Moradi Behbahani S, Moradi M, Basiri R, Mirzaei J (2017). Sand mining disturbances and their effects on the diversity of arbuscular mycorrhizal fungi in a riparian forest of Iran. Journal of Arid Land 9 (6): 837-849. - doi: 10.1007/s40333-017-0028-0

Moradi M, Imani F, Naji HR, Moradi Behbahani S, Ahmadi MT (2017). Variation in soil carbon stock and nutrient content in sand dunes after afforestation by *Prosopis juliflora* in the Khuzestan province (Iran). iForest 10 (3): 585-589. - doi: 10.3832/ifor2137-010

Moradi M, Jorfi MR, Basiri R, Yusef Naanaei S, Heydari M (2022). Beneficial effects of livestock exclusion on tree regeneration, understory plant diversity, and soil properties in semiarid

forests in Iran. Land Degradation and Development 33 (2): 324-332. - doi: 10.1002/ldr.4154
Mousavi A, Shahabzi F, Oustan S, Jafarzadeh A, Minasny B (2020). Application of two data mining techniques for mapping the spatial distribution of soil organic carbon (case study: east shore of Urmia Lake). Water and Soil 34 (3): 689-705. - doi: 10.22067/jsw.v34i3.84154
Nabiollahi K, Eskandari Sh Taghizadeh-Mehrjardi R, Kerry R, Triantafilis J (2019). Assessing soil organic carbon stocks under land-use change scenarios using random forest models. Carbon Management 10 (1): 63-77. - doi: 10.1080/17583004.2018.1553434
Ogunwole JO, Obidike EO, Timm LC, Odunze AC, Gabriels DM (2014). Assessment of spatial distribution of selected soil properties geospatial statistical tools. Communications in Soil Science and Plant Analysis 45 (16): 2182-2200. - doi: 10.1080/00103624.2014.912288
Olsen SR, Cole CV, Watanabe FS, Dean LA (1954). Estimation of available phosphorus in soils by extraction with sodium bicarbonate. USDA, Washington, DC, USA, pp. 1-19. [online] URL: http://books.google.com/books?id=d-oaM80x5agC
Padarian J, Minasny B, Mcbratney AB (2020). Machine learning and soil sciences: a review aided by machine learning tools. Soil 6: 35-52. - doi: 10.5194/soil-6-35-2020
Page-Dumroese DS, Brown RE, Jurgensen MF, Mroz GD (1999). Comparison of methods for determining bulk densities of Rocky Forest soils. Soil Science Society of America Journal 63 (2): 379. - doi: 10.2136/sssaj1999.03615995006300020016x
Pandey M, Mishra A, Swamy SL, Anderson JT, Thakur TK (2025). Machine learning-based monitoring of land cover and reclamation plantations on coal-mined landscape using Sentinel 2 data. Environmental and Sustainability Indicators 25 (25): 100585. - doi: 10.1016/j.indic.2025.100585
Parajuli M, Cristan R, Daniel MJ, Rijal A, Mitchell D, McDonald T, Gallagher T (2025). Integrating high-resolution imagery, deep learning, and GIS to estimate soil erosion following timber harvesting. Remote Sensing Applications: Society and Environment 37 (37): 101498. - doi: 10.1016/j.rsase.2025.101498
Parvizi Y, Fatehi S (2025). Geospatial digital mapping of soil organic carbon using machine learning and geostatistical methods in different land uses. Scientific Report 15: 4449. - doi: 10.1038/s41598-025-88062-9
Piccini C, Francaviglia R, Marchetti A (2020). Predicted maps for soil organic matter evaluation: the case of Abruzzo Region (Italy). Land 9: 349. - doi: 10.3390/land9100349
Sagheb Talebi K, Sajedi T, Pourhashemi M (2013). Forests of Iran. Plant and vegetation. Springer Nature, Netherlands, pp. 152. - doi: 10.1007/978-94-007-7371-4
Sefidi K, Mohadjer MRM (2010). Characteristics of coarse woody debris in successional stages of natural beech (*Fagus orientalis*) forests of Northern Iran. Journal of Forest Science 56 (1): 7-17. - doi: 10.17221/113/2008-jfs
Setälä HM, Francini G, Allen JA, Hui N, Jumpponen A, Kotze DJ (2016). Vegetation type and age drive changes in soil properties, nitrogen, and carbon sequestration in urban parks under cold climate. Frontiers in Ecology and Evolution 4: 93. - doi: 10.3389/fevo.2016.00093
Song YQ, Yang LA, Li B, Hu YM, Wang AL, Zhou W, Cui XS, Liu YL (2017). Spatial prediction of soil organic matter using a hybrid geostatistical model of an extreme learning machine and ordinary kriging. Sustainability 9 (5): 754. - doi: 10.3390/su9050754
Sun W, Zhang X, Zou B, Wu T (2017). Exploring the potential of spectral classification in estimation of soil contaminant elements. Remote Sensing 9 (6): 632. - doi: 10.3390/rs9060632
Szabó B, Szatmári G, Takács K, Laborczi A, Makó A, Rajkai K, Pásztor L (2019). Mapping soil hydraulic properties using random-forest-based pedotransfer functions and geostatistics. Hydrology and Earth System Sciences 23 (6): 2615-2635. - doi: 10.5194/hess-23-2615-2019
Tziachris P, Aschonitis V, Chatzistathis T, Papadopoulou M (2019). Assessment of spatial hybrid methods for predicting soil organic matter using DEM derivatives and soil parameters. Catena 174: 206-216. - doi: 10.1016/j.catena.2018.11.010
Veen GF, Fry EL, Hooven FC, Kardol P, Morriën E, De Long JR (2019). The role of plant litter in driving plant-soil feedbacks. Frontiers in Environmental Science 7: 168. - doi: 10.3389/fenvs.2019.00168
Wadoux A, Minasny B, McBratney AB (2020). Machine learning for digital soil mapping: applications, challenges and suggested solutions. Earth Science Reviews 210: 103359. - doi: 10.31223/osf.io/8eq6s
Walkley A, Black IA (1951). An examination of the Degtjareff method for determining organic carbon in soils: effect of variations in digestion conditions and of inorganic soil constituents. Soil Science 63: 251-263. - doi: 10.1097/00010694-194704000-00001
Wan JZ, Yu JH, Yin GJ, Song ZM, Wei DX, Wang CJ (2019). Effects of soil properties on the spatial distribution of forest vegetation across China. Global Ecology and Conservation 18: e00635. - doi: 10.1016/j.gecco.2019.e00635
Wang N, Chen S, Huang J, Frappart F, Taghizadeh R, Zhang X, Wigneron JP, Xue J, Xiao Y, Peng J, Shi Z (2024). Global soil salinity estimation at 10 m using multi-source remote sensing. Journal of Remote Sensing 4: 130. - doi: 10.34133/remotesensing.0130
Xie X, Zheng Y, Li Y (2015). Genetic Algorithm and its performance analysis for scheduling a single crane. Discrete Dynamics in Nature and Society 2015: 18436. - doi: 10.1155/2015/618436
Yang H, Miao N, Li S, Ma R, Liao Z, Wang W, Sun H (2019). Relationship between stand characteristics and soil properties of two typical forest plantations in the mountainous area of Western Sichuan, China. Journal of Mountain Science 16 (8): 1816-1832. - doi: 10.1007/s11629-018-5265-y
Zeraatpisheh M, Ayoubi S, Jafari A, Tajik S, Finke P (2019). Digital mapping of soil properties using multiple machine learning in a semi-arid region, central Iran. Geoderma 338: 445-452. - doi: 10.1016/j.geoderma.2018.09.006
Zolfaghari Nia M, Moradi M, Moradi G, Taghizadeh-Mehrjardi R (2023). Machine learning models for prediction of soil properties in the riparian forests. Land 12 (1): 32. - doi: 10.3390/land12010032