International Journal of
Agriculture, Environment and Food Sciences
e-ISSN: 2618-5946 https://dergipark.org.tr/jaefs

DOI: https://doi.org/10.31015/2025.4.14.r Int. J. Agric. Environ. Food Sci. 2025; 9 (4): 1150-1161

Prediction of Soil Physicochemical and Biochemical Attributes under Different Land Uses through VNIRS–Based PLSR Models

Süreyya Betül RUFAİOĞLU[1] , Fatma KAPLAN[2] , Ali Volkan BİLGİLİ[3]

[1,2,3]Harran University, Faculty of Agriculture, Department of Soil Science and Plant Nutrition, Sanliurfa, Türkiye.

Article History
Received: October 9, 2025
Accepted: December 6, 2025
Published Online: December 11, 2025
Final Version: December 26, 2025

Article Info
Type: Research Article
Subject: Agricultural Systems Analysis and Modelling

Corresponding Author
Süreyya Betül Rufaioğlu
sureyyarufaioglu@harran.edu.tr

Author ORCID
[1]https://orcid.org/0009-0006-0225-7629
[2]https://orcid.org/0000-0002-4873-3997
[3]https://orcid.org/0000-0002-4727-8283

Available at
https://dergipark.org.tr/jaefs/issue/99001/1800262

DergiPark AKADEMİK

Abstract

This study aimed to evaluate the effects of different land use types (Melissa officinalis, cotton, pistachio, and uncultivated) on the physicochemical and biochemical properties of soils developed on the same parent material under semi-arid conditions, and to assess the potential of Visible–Near Infrared Spectroscopy (VNIRS) for predicting these soil attributes. The soils in the study area are formed on limestone-derived colluvial–alluvial deposits characteristic of the Harran soil series, classified as Vertic Calciorthids (Soil Taxonomy) and Calcic Vertisols (WRB). Laboratory analyses included soil texture, pH, electrical conductivity (EC), calcium carbonate, organic matter (OM), water retention parameters, and enzyme activities (β-glucosidase, dehydrogenase, alkaline phosphatase). Spectral reflectance data in the 350–2500 nm range were used to develop Partial Least Squares Regression (PLSR) models for soil property estimation. The models demonstrated good calibration performance for EC (R^2 = 0.93), OM (R^2 = 0.49), and dehydrogenase activity (R^2 = 0.93), while validation accuracy remained modest (R^2 = 0.46, 0.43, and 0.75, respectively), reflecting the limitations of the small sample size. Texture-related parameters (sand, silt, clay) showed limited predictive accuracy (R^2 = 0.10). Distinct absorption bands at 1400, 1900, and 2200 nm were associated with soil moisture and clay minerals. Although Melissa-cultivated soils tended to show higher organic matter and enzyme activity, these differences should be interpreted cautiously due to the limited number of samples, representing only preliminary indications rather than generalizable trends. Overall, the findings suggest that VNIRS has potential as a rapid and cost-effective approach for characterizing soil biochemical indicators and supporting sustainable land management in semi-arid regions, but further studies with larger datasets are needed to confirm its predictive reliability.

Keywords: Land use, Soil enzyme activity, VNIRS, PLSR, Soil biochemical properties

Cite this article as: Rufaioglu, S.B., Kaplan, F., Bilgili, A.V. (2025). Prediction of Soil Physicochemical and Biochemical Attributes under Different Land Uses through VNIRS–Based PLSR Models. International Journal of Agriculture, Environment and Food Sciences, 9 (4): 1150-1161. https://doi.org/10.31015/2025.4.14.r

INTRODUCTION

Soil health is a fundamental determinant of ecosystem productivity and sustainability, regulating plant growth, carbon sequestration, and key soil functions (Lehmann & Kleber, 2015; Gao et al., 2025). Enzyme activities, widely used as indicators of soil health, play essential roles in the decomposition of organic matter, nutrient mobilization, and microbial processes (Schimel & Weintraub, 2003; Sinsabaugh & Follstad Shah, 2012). As core drivers of soil biological functioning, enzyme activities provide sensitive early-warning signals of degradation and shifts in nutrient cycling. Land use practices strongly influence these biochemical processes by altering soil organic matter dynamics, microbial communities, and nutrient cycling (Bolan et al., 2011; Ren et al., 2025). In semi-arid regions, where soil systems are particularly vulnerable, land use changes often intensify soil degradation and disrupt enzyme activity patterns (Wang et al., 2025). Understanding how land use regulates enzyme-mediated soil functions is therefore crucial for developing sustainable management strategies.

Visible–Near Infrared Spectroscopy (VNIRS) has emerged as a rapid, non-destructive, and cost-effective tool for assessing soil properties (Bilgili et al., 2010; Kaplan & Bilgili, 2024; Çullu et al., 2024; Álvarez et al., 2025; Singha et al., 2023). Unlike conventional laboratory analyses which are time-consuming, reagent-intensive, and labor-demanding VNIRS enables fast, repeatable measurements with minimal sample preparation, making it highly suitable for large-scale soil

monitoring. Previous studies have demonstrated its ability to estimate key physicochemical attributes and overall soil quality (Kinoshita et al., 2012). Recent findings further indicate that VNIRS can successfully predict biological parameters such as enzyme activities and microbial biomass under arid conditions (Hosseini et al., 2024) and can reliably model diverse soil attributes including SOC, SOM, pH, EC, macro-nutrients, and texture fractions using PLSR-based approaches (Vibhute et al., 2018). Despite these advances, the ability of VNIRS to quantify enzyme activities across contrasting land use types in semi-arid regions remains poorly understood, representing a significant gap in the literature.

Land use intensity and vegetation type are known to influence SOC levels, nutrient cycling, and enzyme responses (Ma et al., 2023; Raiesi & Beheshti, 2015). Continuous cultivation often reduces microbial biomass and enzymatic activity (Rayegani et al., 2016), whereas forest and pasture ecosystems tend to support higher biological functioning due to improved organic matter inputs and microbial diversity (Wang & Huang, 2020; Singh et al., 2007). Moreover, shifts in organic residue quality, rhizosphere effects, and alterations in microbial community composition directly regulate the activity of hydrolytic and oxidative enzymes. Interactions between nutrient inputs and climate-driven moisture conditions further complicate soil biochemical responses (Ren et al., 2025). These complexities underscore the need for rapid, scalable approaches to monitor soil function under varying land management and climatic regimes (Gao et al., 2025). VNIRS offers a particularly promising solution for such assessments because enzyme activities often correlate with organic matter chemistry and moisture-sensitive absorption bands detectable in the VNIR region.

The novelty of this study lies in combining VNIRS with detailed biochemical assessments to evaluate enzyme activity responses to different land use types in a semi-arid environment. Specifically, we investigate whether VNIRS can accurately estimate soil enzyme activities (β-glucosidase, alkaline phosphatase, dehydrogenase) and key physical–chemical properties across land use systems using spectral reflectance in the 350–700 nm and 350–2500 nm ranges. To the best of our knowledge, this is the first study to systematically compare the performance of VNIRS–PLSR models for enzyme activities across multiple land use types in a semi-arid region. By integrating spectral data with laboratory analyses, this study provides new evidence on the potential of VNIRS to serve as a rapid and reliable alternative to conventional methods for soil monitoring and precision land management.

MATERIALS AND METHODS

Locations and Soil Sampling

The study was conducted within the Haliliye district of Sanliurfa province, near Göbeklitepe, between latitudes 37°12'37"–37°13'05" N and longitudes 38°59'59"–38°59'05" E. The elevation of the study area ranges from 577 to 624 meters and exhibits diverse topographic features and vegetation characteristics. Although the soil samples examined originate from the same parent material limestone-derived colluvial–alluvial deposits characteristic of the Harran soil series, classified as Vertic Calciorthids (Soil Taxonomy) and Calcic Vertisols (WRB.,2022) the landforms differ across the sites. Accordingly, sampling was carried out across four distinct land use areas: pistachio (363 m²), Melissa *officinalis* (72 m²), cotton (99 m²), and uncultivated land (14 m²) (Figure 1). A total of 16 soil samples were collected, with four samples taken from each land use type, using a random sampling method on January 3, 2024, from a depth of 0–30 cm. This sampling design enabled a detailed evaluation of the effects of different land use practices on the physical and chemical properties of the soils. Analyses conducted on the collected samples enabled a detailed evaluation of the effects of different land use practices on the physical and chemical properties of the soils. In this study, each of the four land use types was represented by four independent composite samples (n = 4 per land use type), resulting in a total of 16 samples. Each composite sample consisted of five subsamples collected within a 5–10 m radius and thoroughly homogenized to ensure representativeness. Although the total sample size is limited, this design follows VNIRS calibration studies conducted on shallow soil horizons in semi-arid regions, where spectral homogeneity is prioritized over bulk sample volume (Stenberg et al., 2010). All laboratory measurements (physical, chemical, and enzymatic) were performed in triplicate, and mean values were used for calibration to increase analytical robustness and reduce measurement error.

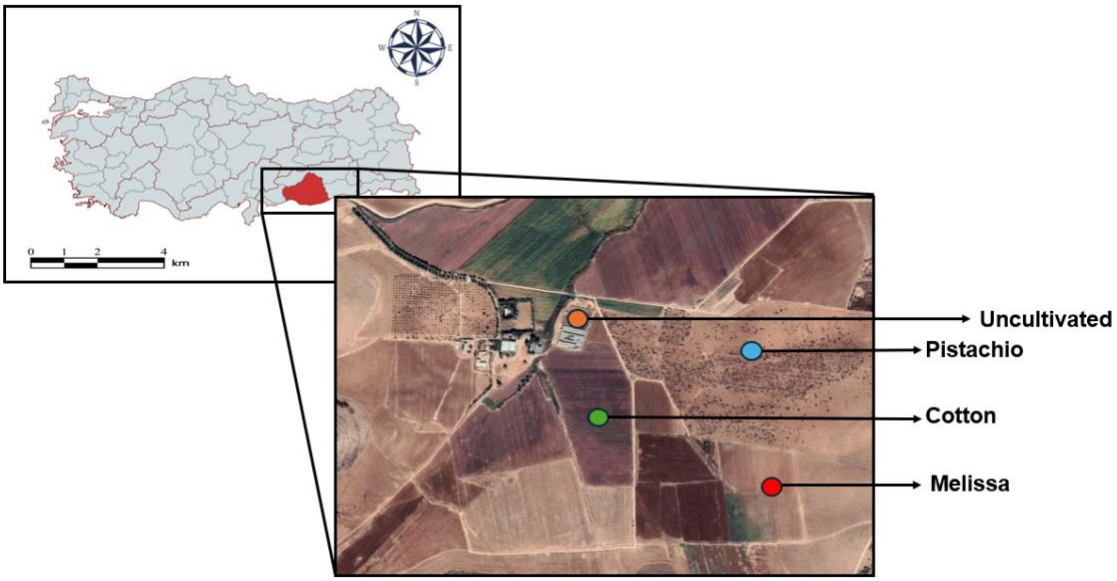

Figure 1. The location of the study area and the distribution of the types of land use

The region is located within a semi-arid climate zone, with annual temperature and precipitation values exhibiting certain fluctuations. In January, the average temperature was recorded as 8.5°C, with values ranging between 2.3°C and 16.2°C. During the same period, total precipitation was 71.4 mm, and the average relative humidity was measured at 74.5%. (Table 1).

Table 1. Climate information for the year and month of soil sampling (MGM, 2024).

Month	Max.temp. (°C)	Min.temp. (°c)	Averagetemp. (°C)	Rainfall (mm)	Humidity (%)
January	16.2	2.3	8.5	71.4	74.5

Physical and chemical analysis of samples

After sampling, the soil samples were air-dried and sieved through a 2 mm mesh to prepare them for laboratory analyses. Standard procedures were employed to determine the physical and chemical properties of the soil. Soil texture (clay, silt, sand) was measured using the hydrometer method developed by Bouyoucos (1951). Soil pH and electrical conductivity (EC) were determined in saturated paste extracts following the procedure described by McLean (1982). Soil organic matter (SOM) content was quantified using the modified Walkley–Black method (Nelson & Sommers, 1982), while calcium carbonate ($CaCO_3$) content was measured using a Scheibler-type calcimeter (Tüzüner, 1990). Water retention properties, including field capacity (FC), wilting point (WP), available water capacity (AWC), and water-holding capacity (WHC), were assessed using the pressure plate apparatus as outlined by Klute (1986). Plant-available phosphorus (Available P) was analyzed using the Olsen sodium bicarbonate extraction method (Olsen, 1954). Aggregate stability (AS) was evaluated using the rainfall simulator–based wet stability method (Kemper & Rosenau, 1986; Gugino et al., 2009).

Enzyme analysis in soil

Soil enzyme activities were quantified using established colorimetric procedures widely applied in soil biology (Tabatabai, 1982; Schinner et al., 1996; Eivazi et al., 2003), with all assays performed in triplicate. β-glucosidase (BG) activity was determined following Eivazi & Tabatabai (1988), using p-nitrophenyl-β-D-glucopyranoside (25 mM) as the substrate. One gram of soil was incubated with 4 mL modified universal buffer (pH 6.0) at 37°C for 1 h, and the reaction was terminated using 0.5 M $CaCl_2$ and 0.1 M THAM–NaOH before measuring absorbance at 410 nm; results were expressed as mg pNP g^{-1} soil h^{-1}. Alkaline phosphatase (ALP) activity was quantified according to Tabatabai & Bremner (1969) using p-nitrophenyl phosphate (25 mM) in MUB at pH 11, under incubation and detection conditions similar to the BG assay and reported as µg pNP g^{-1} soil h^{-1}. Acid phosphatase was initially evaluated but excluded due to inconsistent responses, and therefore only ALP results are presented. Dehydrogenase (DHA) activity was measured following Casida et al. (1964) and Hopkins et al. (1996) by incubating 5 g soil with 5 mL of 0.5% TTC at 37°C for 24 h under anaerobic conditions, extracting the resulting triphenyl formazan (TPF) with methanol, and measuring absorbance at 485 nm; results were expressed as µg TPF g^{-1} soil 24 h^{-1}. These fully detailed protocols ensure reproducibility and follow internationally accepted standards for assessing soil microbial activity in semi-arid soils (Schinner et al., 1996; Acosta-Martínez & Tabatabai, 2000).

Spectral Analysis of Soil Samples by Visible Near Infrared Spectroscopy

The spectral data of the soil samples were collected at the Spectroradiometer and EM38 Laboratories of the Department of Soil Science and Plant Nutrition at Harran University. Prior to measurement, the soil samples were air-dried, sieved through a 2 mm mesh, and placed into 4 cm-diameter optical petri dishes. A Tungsten Quartz Halogen lamp was used as the light source, and the reflected light from each sample was transmitted to a computer through a fiber-optic cable. VNIR reflectance spectra within the 350–2500 nm wavelength range were obtained using an ASD FieldSpecPro III spectroradiometer with a 1 nm spectral resolution, resulting in 2151 spectral bands per sample. Before scanning, the instrument was calibrated using a white Spectralon reference panel, and recalibration was performed at regular intervals (every 10 samples) to minimize reflectance drift, following established VNIRS quality-control protocols (Stenberg et al., 2010; Ben-Dor et al., 1999). Each soil sample was scanned three times, with the petri dish rotated by 120° between scans to reduce directional reflectance effects, and the mean spectrum was used for further analysis. Instrument warm-up and lamp stabilization procedures were completed before measurements to ensure spectral consistency.

Prior to chemometric modeling, spectral data were preprocessed using commonly applied signal-enhancement techniques. A Savitzky–Golay (SG) smoothing filter (second-order polynomial, 15-point window) was applied to reduce high-frequency noise, and the final reflectance values were calculated by normalizing the soil reflectance measurements against the white reference panel (Equation 1). All preprocessing steps adhered to established protocols for soil VNIRS applications (Stenberg et al., 2010; Ben-Dor et al., 1999).

$$R = \frac{soil\ reflection}{White\ spektralon\ reflectance\ (reference)} \qquad \text{Equation 1}$$

Partial Least Square Regression (PLSR)

Prior to model construction, all spectral predictor variables (X matrix) were standardized using z-score normalization to ensure equal weighting across the 350–2500 nm wavelength range. Soil biochemical and physicochemical parameters (Y matrix) were also standardized to minimize scale-related bias and improve model stability. PLSR was used to model the relationship between the spectral data and measured soil properties through a multivariate linear calibration framework

(Esbensen & Geladi, 2010). Full cross-validation was performed to evaluate model reliability. In this procedure, each sample is sequentially omitted, the model is recalibrated using the remaining samples, and the omitted sample is predicted before being reinserted into the dataset. This iterative approach allowed for both robust model assessment and the identification of potential overfitting. The optimal number of latent factors for each model was determined according to the lowest cross-validation error, following standard chemometric recommendations (Milos & Bensa, 2018). This systematic validation ensured that each PLSR model achieved maximum predictive performance while avoiding unnecessary model complexity.

Performance Evaluation Metrics

PLSR modeling was conducted using Unscrambler X (Version 10.5, Camo Analytics), a widely used chemometric software for multivariate calibration. The software was employed to perform spectral preprocessing (where applicable), latent factor optimization, and full cross-validation procedures due to its robustness and proven performance in soil spectroscopy applications.

Complementary statistical analyses including descriptive statistics, Pearson correlation coefficients, and data visualization were performed using Python 3.10. The SciPy (1.10), NumPy (1.26), and pandas (2.0) libraries were used for numerical operations and statistical computation, while Matplotlib (3.8) and Seaborn (0.13) were used to generate all exploratory and correlation-based plots. The accuracy of predictions obtained using the Partial Least Squares Regression (PLSR) method was assessed using metrics such as the coefficient of determination (R^2) (Equation 2) and the root mean square error (RMSE). The R^2 value reflects how closely the data conform to the regression line, whereas RMSE indicates the dispersion of validation dataset samples around the 1:1 regression line and the average prediction error. RMSE is computed as the square root of the mean of the squared differences between observed and predicted values (Equation 3). In the equation, n represents the number of samples in the validation dataset, and Y_pred and Y_lab correspond to the predicted and actual measured values, respectively.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2}{\sum_{i=1}^{n}(Y_i - \bar{Y}_i)^2} \qquad \text{Equation 2}$$

$$RMSE = \sqrt{\frac{\sum_{i-1}^{n}(Yestimated - Ylab.observed)^2}{n}} \qquad \text{Equation 3}$$

RESULT

Soil parameters were predicted through calibration models developed between the parameters and their corresponding reflectance values using the Partial Least Squares Regression (PLSR) method. A cross-validation approach was applied to assess the accuracy of the predictions. This method systematically tests the predictive capability of the model by dividing the dataset into training and testing subsets. The model is developed using the training data and then its performance is evaluated on the test data. This process plays a critical role in determining the reliability and validity of the model under real-world conditions. Cross-validation also helps reduce the risk of overfitting and improves the generalizability of the model. The cross-validation results are presented with the coefficient of determination (R^2) and the root mean square error of prediction (RMSEP) for various soil parameters (Tables 2 and 3). These tables illustrate the relationship between the predicted values obtained from the PLSR model and the laboratory measurements, highlighting the parameters for which the model performs best. The findings also provide a basis for evaluating which wavelength range is more suitable for predicting specific soil parameters. Average spectral reflectance curves of soils under different land use types (Melissa, cotton, pistachio, and fallow) measured across the 350–700 nm (left) and 350–2500 nm (right) wavelength ranges. Reflectance intensity increases toward the near-infrared region, with distinct absorption features observed near 1400, 1900, and 2200 nm corresponding to soil moisture and clay mineral bands (Figure.2).

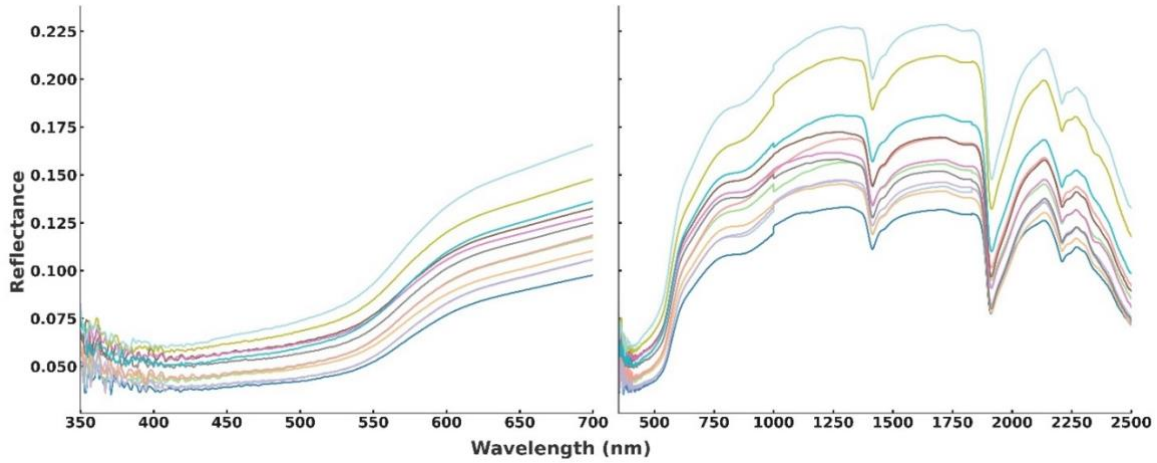

Figure 2. Reflectance spectra of soil samples measured in the 350–2500 nm wavelength range in the visible and near-infrared (VNIR) regions. The left panel illustrates the visible range (350–700 nm), where reflectance gradually increases with wavelength. The right panel presents the near-infrared range (700–2500 nm).

Descriptive statistics of the measured soil physicochemical and biochemical properties are presented in Table 2. The results revealed a moderate variation among soil parameters across different land use types. Lime content ranged between 24.46% and 31.30%, reflecting the calcareous nature of the soils in the study area. SOM values were relatively low (1.18–1.81%), consistent with semi-arid conditions where organic residue input is limited. EC varied from 292.67 to 365.67 µS cm^{-1}, indicating non-saline to slightly saline conditions, while pH showed minimal variation (7.31–7.36), suggesting a slightly alkaline character typical of calcareous soils. Texture analysis indicated that clay was the dominant fraction (52.67–66%), followed by sand (15.33–24%) and silt (18.67–23.33%), classifying the soils predominantly as clayey. Among water-related properties, AWC ranged between 26.05% and 38.47%, while the WP and WHC exhibited higher variability, reflecting the influence of organic matter and texture on soil water retention. AS values (44.69–62.83%) demonstrated moderate structural stability, likely influenced by differences in vegetation cover and management practices. Regarding biochemical parameters, DHA varied from 11.54 to 19.17 µg TPF g^{-1} soil, BG from 4.72 to 10.98 mg pNP g^{-1} soil, and ALP from 9.21 to 16.37 µg g^{-1} soil, indicating active microbial processes, particularly in cultivated soils. Available P content ranged from 5.81 to 9.41 mg kg^{-1}, showing moderate nutrient availability.

The Shapiro–Wilk normality test results ($p > 0.05$ for most variables) indicated that the data generally followed a normal distribution, except for silt and WHC, which showed slightly skewed distributions. Overall, the descriptive statistics reflect the heterogeneity of soil physical, chemical, and biochemical properties under different land use conditions in the semi-arid environment.

Table.2 Descriptive statistics of soil physicochemical and biochemical properties, including minimum, maximum, mean, standard deviation (SD), skewness, kurtosis, and Shapiro–Wilk test p-values.

Indicator	Min	Max	Mean	SD	Skewness	Kurtosis	Shapiro–Wilk p
Lime	24.457	31.3	27.197	2.903	0.752	-0.871	0.356
SOM	1.178	1.81	1.568	0.273	-0.833	-0.861	0.353
EC	292.667	365.667	323.333	35.793	0.238	-1.685	0.291
pH	7.313	7.363	7.338	0.02	0	-1.018	0.848
Sand	15.333	24	19.333	3.569	0.316	-1	0.798
Clay	52.667	66	59.833	5.64	-0.268	-1.196	0.955
Silt	18.667	23.333	20.833	2.517	0.035	-1.954	0.084
AWC	26.05	38.472	32.595	5.559	-0.139	-1.532	0.802
WP	13.33	24.718	19.959	4.785	-0.635	-0.926	0.560
WHC	5.581	16.853	12.637	4.891	-0.887	-0.822	0.253
AS (%)	44.693	62.833	54.284	9.555	-0.039	-1.942	0.126
DHA	11.537	19.167	14.696	3.689	0.288	-1.619	0.331
ALP	9.213	16.377	12.465	3.057	0.317	-1.222	0.922
β-glucosidase	4.724	10.982	7.958	2.677	-0.112	-1.313	0.977
Available P (mg kg^{-1})	5.812	9.413	7.54	1.588	0.114	-1.477	0.876

The Pearson correlation analysis between spectral reflectance values and soil physicochemical and biochemical properties across the 350–2500 nm wavelength range is presented in Figure 3. The correlation patterns revealed clear spectral sensitivity for most soil parameters, with distinct wavelength regions showing either strong positive or negative relationships. Positive correlations were particularly evident for lime, SOM, EC, and enzyme-related parameters, especially within the visible (400–700 nm) and near-infrared (700–1300 nm) regions. SOM and DHA) exhibited the strongest positive correlations in the NIR region, suggesting that spectral responses in this range are primarily influenced by organic carbon and associated microbial processes. Conversely, pH and available P displayed negative correlations, reflecting their limited optical sensitivity and indirect relationship with reflectance intensity. Among physical parameters, sand and AS demonstrated positive trends with increasing wavelength, while clay and WP showed negative correlations, indicating the higher absorption capacity of clay-rich soils. Distinct absorption features near 1400, 1900, and 2200 nm were consistently associated with soil water and clay mineral bands, emphasizing the influence of moisture and texture on spectral behavior. Overall, the correlation analysis identified specific wavelength regions most responsive to soil organic matter and enzyme activity variations, confirming the potential of VNIRS data to capture biochemical processes and guiding the selection of sensitive bands for subsequent PLSR modeling.

In this modeling study, the Partial Least Squares Regression (PLSR) method was applied using spectral data within the 350–2500 nm wavelength range to evaluate the calibration (RMSEC, R^2) and validation (RMSEP, R^2) performance for various soil physical, chemical, and biochemical properties (Table 3, Figure 3). High calibration accuracy was achieved for key chemical properties such as SOM, EC, and pH. For EC, the calibration R^2 value was 0.93 with an RMSEC of 9.69 µS/cm, while the validation R^2 dropped to 0.46 and the RMSEP was calculated as 30.68 µS/cm. The validation R^2 values for SOM and pH were determined as 0.43 and 0.40, respectively. For soil texture components, lower validation performance was observed, with R^2 values of 0.42 for clay, 0.31 for sand, and 0.10 for silt. Meanwhile, for water-related parameters such as PAW, a strong calibration fit ($R^2 = 0.93$) was achieved, but a decrease in validation accuracy was noted ($R^2 = 0.26$). Among the biochemical properties, high calibration R^2 values were recorded for DHA and BG enzyme activities, at 0.93 and 0.72, respectively. In the validation phase, these values dropped to 0.75 and 0.66. For parameters such as ALP and Available P, relatively low R^2 values were observed.

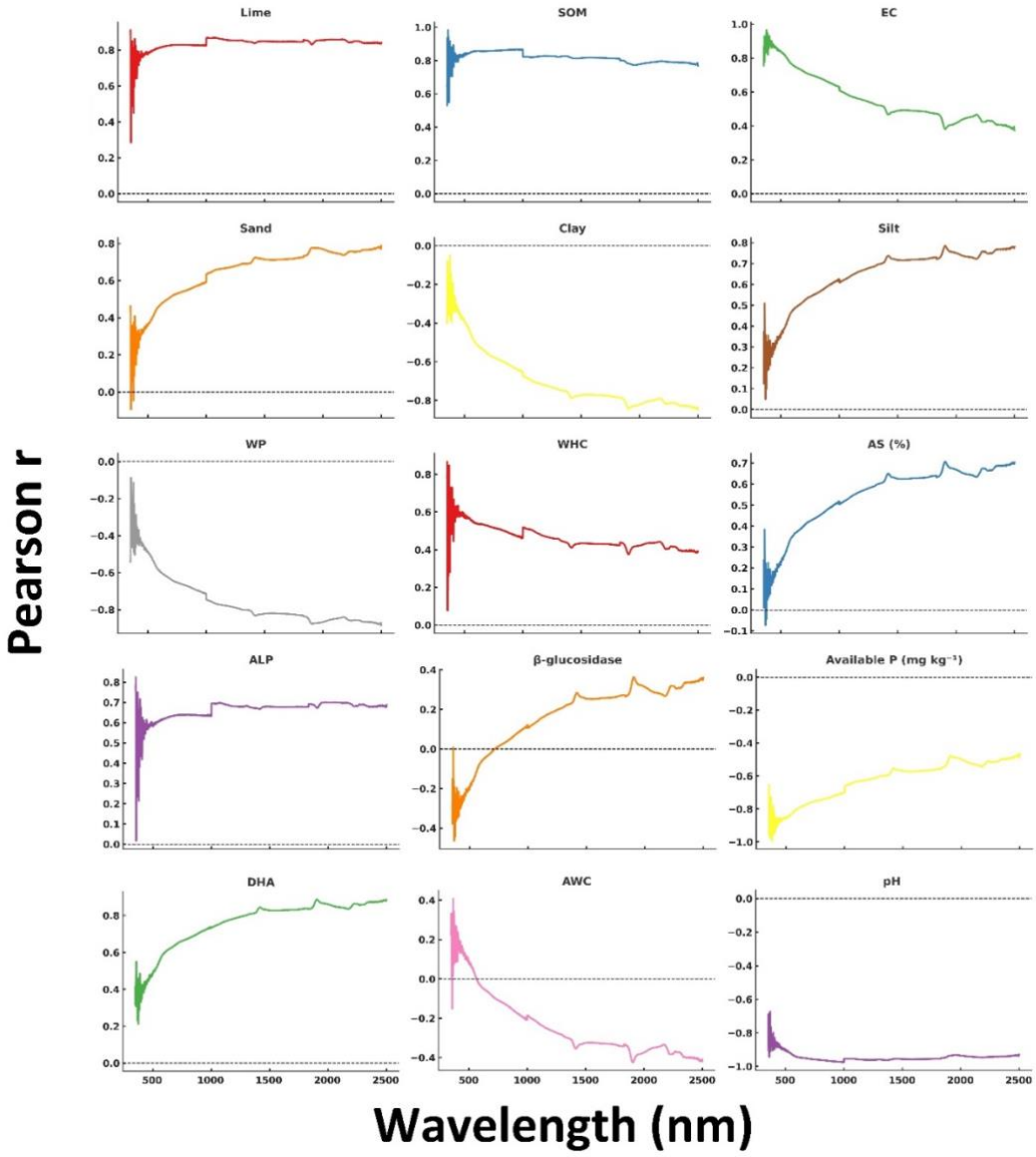

Figure 3. Pearson correlation coefficients (r) between soil physicochemical and biochemical properties.

Table 3. Calibration and validation statistics for soil physicochemical and biochemical properties predicted using spectral reflectance data (350-2500 nm). Metrics include root mean square error of calibration (RMSEC), root mean square error of prediction (RMSEP), and coefficient of determination (R^2) for both calibration and validation datasets.

	Parameters		***Calibration***		***Validation***		
		Unit	**RMSEC**ǂ	**R^2**	**RMSEP**ɪ		**R^2**
	Sand	%	2.69	0,46	3.31	0.31	
	$CaCO_3$	%	0.69	0.95	2.11	0.66	
	SOM	%	0.23	0.49	0.27	0.43	
	EC	µS/cm	9.69	0.93	30.68	0.46	
	pH		0.015	0.46	0.018	0.40	
	Clay	%	3.28	0.70	4.97	0.42	
	Silt	%	2.67	0.25	3.20	0.10	
	AWC	%	2.23	0.82	4.26	0.46	
	WP	%	3.95	0.35	4.59	0.27	
	WHC	%	1.38	0.93	5.06	0.26	
	DHA	µg/g	0.89	0.93	1.86	0.75	
	BG	mg/g	1.48	0.72	1.81	0.66	
	ALP	µg/g	3.88	0.11	4.60		NA
	Available P	mg/kg	2.33	0.35	3.00	0.10	
	AS	%	2.5	0.95	7.51	0.62	

‡Root Mean square error of calibration, ‡Root Mean square error of prediction.

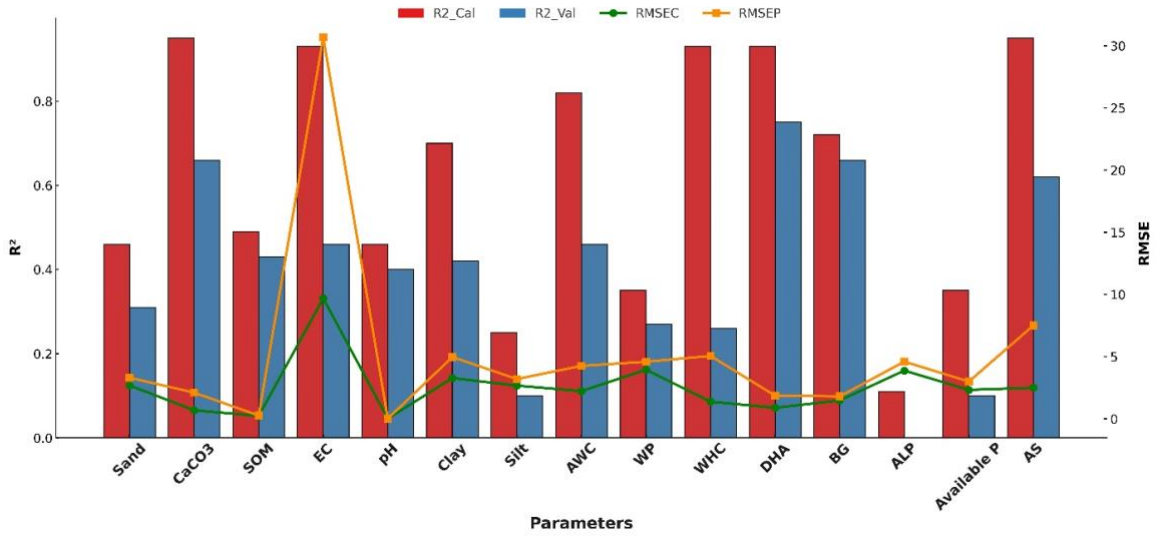

Figure 3. Calibration and validation performance of spectral models for predicting soil physicochemical and biochemical properties (350-2500 nm). Bars represent the coefficient of determination (R^2) for calibration (red) and validation (blue), while lines represent the root mean square error of calibration (RMSEC, green) and root mean square error of prediction (RMSEP, orange).

Following the PLSR modeling performed within the 350–2500 nm spectral range, a second analysis was conducted in the 700–2500 nm range to evaluate how model accuracy changes within a narrower spectral band. This analysis is particularly important to reveal the effect of spectral bandwidth on model performance. The modeling results within the 700–2500 nm range generally demonstrated acceptable calibration performance for certain soil properties, while a noticeable decrease in accuracy was observed during the validation phase. For instance, the validation R^2 values for OM and pH were recorded as 0.43 and 0.40, respectively similar to those previously obtained from the wider spectral range. However, for electrical conductivity (EC), model accuracy dropped significantly in validation, with an R^2 of only 0.09. Likewise, the prediction of soil texture components showed limited accuracy: sand, clay, and silt had validation R^2 values of 0.32, 0.34, and 0.10, respectively. In terms of soil water retention parameters, while calibration R^2 values were relatively high for PAW, FC, and WP for example, PAW: $R^2 = 0.93$ these values decreased significantly during validation (PAW: $R^2 = 0.26$). For biochemical parameters, DHA and BG enzyme activities were among the most successfully predicted variables in this spectral range as well. The validation R^2 was calculated as 0.69 for DHA and 0.60 for BG. In contrast, parameters such as ALP and Available P showed poor prediction performance (Table 4, Figure 4).

In conclusion, the PLSR analyses performed within the 700–2500 nm range exhibited more limited predictive accuracy compared to the broader 350–2500 nm range. This decline was particularly evident in physical soil properties; however, certain biochemical parameters, especially microbial activity indicators like DHA still showed acceptable levels of accuracy. These findings emphasize the critical role of wavelength selection in the success of soil spectral modeling.

Table 4. Calibration and validation statistics for soil physicochemical and biochemical properties predicted using spectral reflectance data (700-2500 nm). Metrics include root mean square error of calibration (RMSEC), root mean square error of prediction (RMSEP), and coefficient of determination (R^2) for both calibration and validation datasets.

Parameters		***Calibration***		***Validation***	
	Unit	**RMSEC‡**	**R^2**	**RMSEP‡**	**R^2**
Sand	%	2.66	0.47	3.29	0.32
$CaCO_3$	%	2.34	0.51	3.14	0.26
SOM	%	0.23	0.49	0.27	0.43
EC	µS/cm	30.78	0.36	40.11	0.09
pH		0.015	0.46	0.018	0.40
Clay	%	4.07	0.54	5.3	0.34
Silt	%	2.68	0.25	3.2	0.10
AWC	%	3.91	0.46	5.11	0.23
WP	%	3.97	0.35	4.61	0.26
WHC	%	5.15	0.10	5.95	NA
DHA	µg/g	1.58	0.78	2.06	0.69
BG	mg/g	1.57	0.69	1.95	0.60
ALP	µg/g	3.88	0.10	4.6	NA
Available P	mg/kg	2.37	0.33	3.07	0.066
AS	%	0.4232	0,99	10.0125	0.34

‡Root Mean square error of calibration, ‡Root Mean square error of prediction.

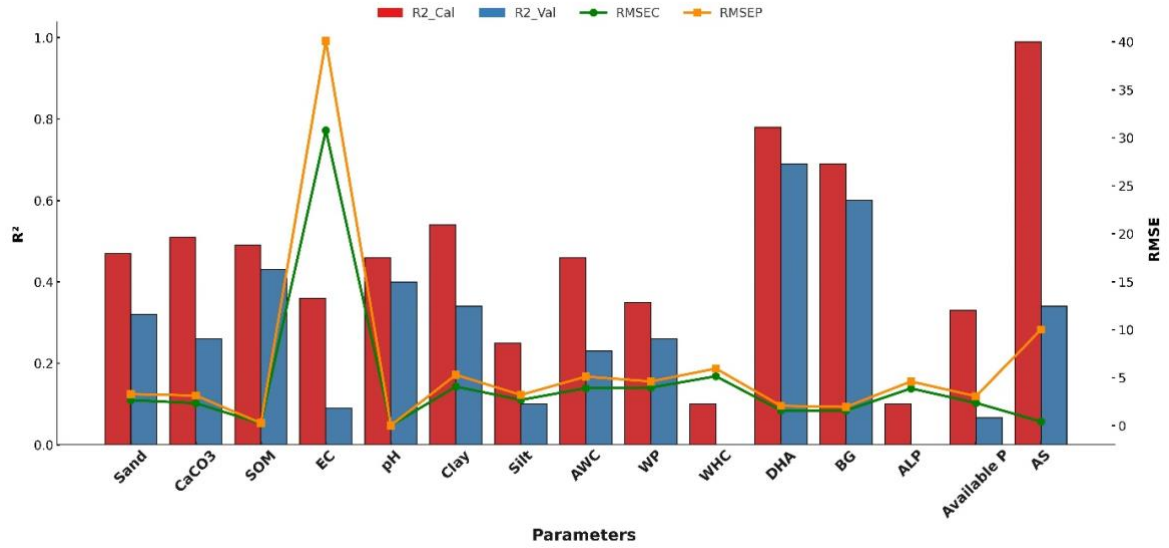

Figure 4. Calibration and validation performance of spectral models for predicting soil physicochemical and biochemical properties (700-2500 nm). Bars represent the coefficient of determination (R^2) for calibration (red) and validation (blue), while lines represent the root mean square error of calibration (RMSEC, green) and root mean square error of prediction (RMSEP, orange).

The average spectral reflectance profiles of soils under different land use types are presented in (Figure 5.) These spectral curves, obtained for the 350–700 nm (a) and 350–2500 nm (b) wavelength ranges, reveal the spectral responses of four distinct land use categories (Melissa, cotton, pistachio, and uncultivated). The most noticeable separation among the curves occurred within the 600–1200 nm range, where the reflectance values tended to diverge more clearly. In the near-infrared region, variations were also observed around the absorption features near 1400, 1900, and 2200 nm, which are commonly associated with water and clay mineral absorptions. Rather than indicating definitive distinctions among land use types, these patterns suggest preliminary spectral differences that warrant further investigation with a larger dataset.

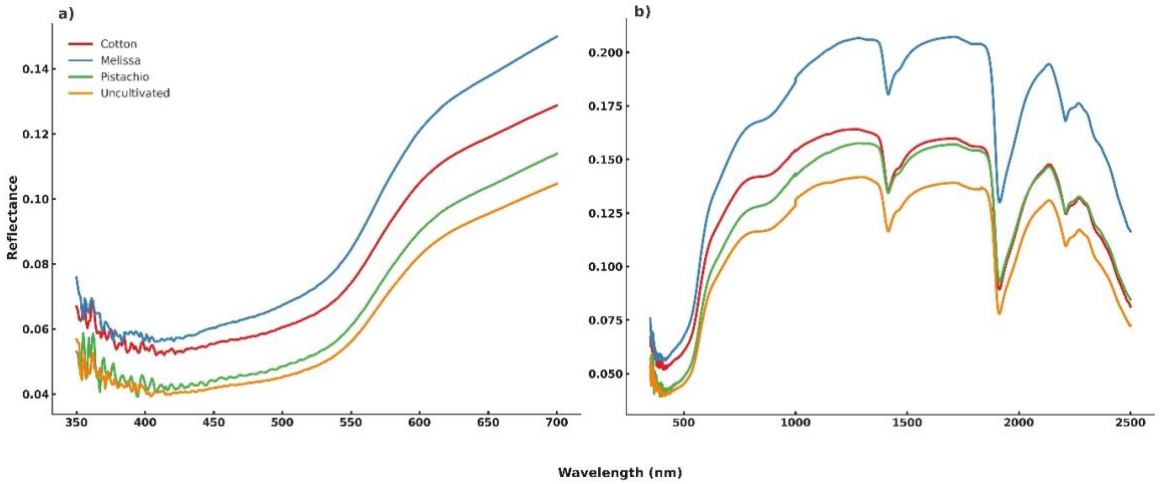

Figure 5. Average spectral reflectance curves of soil samples collected from different land use types (uncultivated land, cotton, pistachio, and Melissa) are presented. The graphs separately show a) the 350–700 nm and b) the 350–2500 nm wavelength ranges.

The combined correlation patterns between soil physicochemical and biochemical parameters and spectral reflectance values across the 350–2500 nm wavelength range are shown in Figure 6. Overall, distinct wavelength-dependent variations were observed, indicating that different soil attributes respond uniquely to specific spectral regions. Positive correlations dominated the visible and near-infrared regions (400–1300 nm) for most chemical and biochemical parameters, particularly SOM, EC, DHA, and BG activities. These parameters exhibited correlation coefficients (r) exceeding 0.6 in the NIR region, suggesting that the reflectance behavior in this range is largely governed by organic compounds, moisture, and microbially active constituents. Lime and AS also showed strong positive associations with reflectance intensity, reflecting their influence on soil structure and spectral brightness. In contrast, variables such as pH, Available phosphorus P, and clay

content demonstrated negative correlations, especially around the 1400 nm and 1900 nm absorption features, where water and hydroxyl groups in clay minerals strongly absorb incoming radiation. Water retention parameters, including AWC, WP, and WHC, showed variable correlations across the NIR–SWIR regions, indicating their complex relationship with both moisture and organic matter content. The pronounced absorption bands at approximately 1400, 1900, and 2200 nm correspond to molecular vibrations of H–O–H, O–H, and Al–OH bonds, indicating the spectral influence of soil moisture and clay minerals. However, given the limited sample size, these observations should be interpreted cautiously. While the spectral patterns suggest the potential of VNIRS to detect chemical and mineralogical variations in soils, a larger dataset would be required to draw definitive conclusions regarding its sensitivity and reliability for prediction and modeling applications.

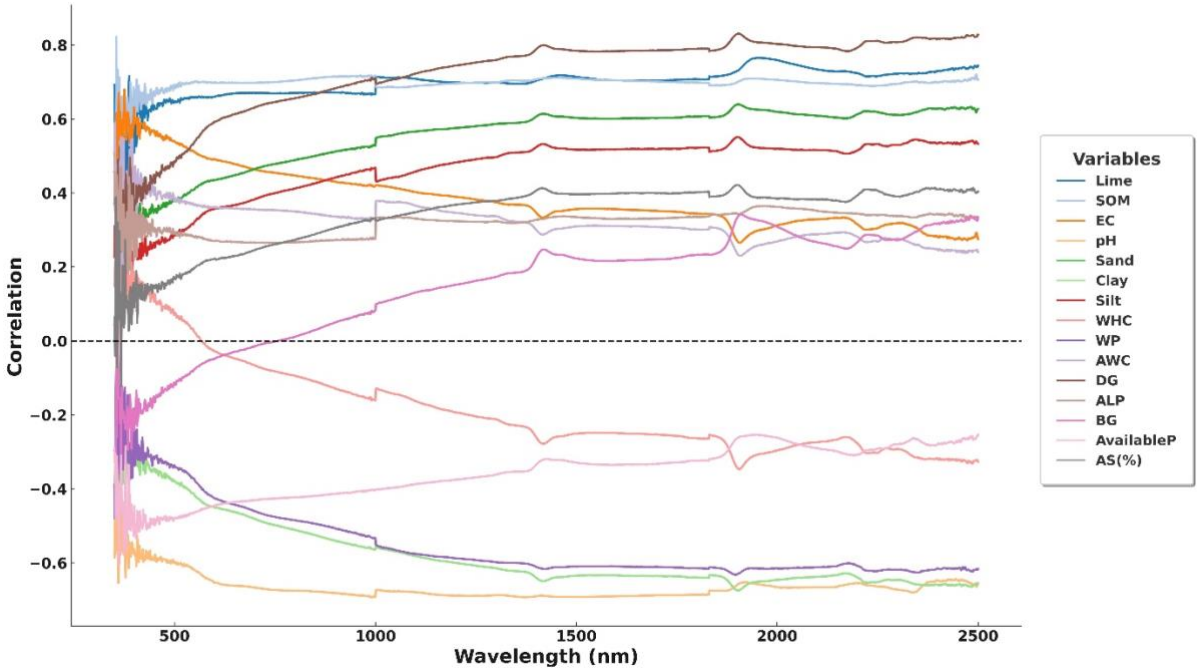

Figure 6. Correlation between soil physicochemical and biochemical properties and spectral reflectance across the 350–2500 nm wavelength range.

DISCUSSION

This study investigated the effects of different land use types on the physical, chemical, and biochemical properties of soils under semi-arid conditions and evaluated the potential of VNIRS spectral data for predicting these properties. The findings revealed that spectral methods demonstrated promising preliminary predictive performance particularly for biochemical and certain chemical properties. The high calibration accuracy obtained for enzyme activities such as dehydrogenase and β-glucosidase indicates that biochemical processes associated with microbial activity and organic matter are reflected in the spectral signals. This is consistent with previous studies reporting that organic matter composition and microbial processes leave distinct signatures in VNIRS spectra (Wang et al., 2025; Stenberg et al., 2010; Rossel & Webster, 2012). Model performance was significantly lower for physical soil properties. Textural components such as sand, clay, and silt exhibited low validation accuracy due to their weak spectral representation and their complex interactions with surface conditions, mineralogical composition, and soil moisture. This finding aligns with the literature indicating that textural attributes have weak spectral signatures and that VNIRS is limited in predicting physical soil parameters (Ben-Dor, 2008; Tsolis & Barouchas, (2023); Stoner & Baumgardner, 1981). Lobell and Asner (2002) further emphasized that soil reflectance is more sensitive to moisture and mineralogical properties, whereas particle size contributes only marginally to the spectral signal. The distinct absorption bands observed around 1400, 1900, and 2200 nm correspond to OH bond vibrations of water and Al–OH combination bands of clay minerals (Clark et al., 1990; Hunt, 1977). The reflectance variations observed in these regions, particularly in melissa-cultivated soils where organic matter and microbial activity were higher, are consistent with the literature. Regarding chemical properties, VNIRS-based models showed moderate predictive performance for variables such as organic matter, calcium carbonate, and to a certain extent electrical conductivity. This is in agreement with previous studies reporting stronger spectral representation of chemical soil parameters in the VNIR region (Reeves et al., 2002; Gomez et al., 2008; Gozukara et al., 2025). Nevertheless, the low validation accuracy for complex chemical attributes such as EC indicates that salinity is influenced by multiple interacting factors including moisture, mineralogy, and surface conditions that complicate its spectral expression (Farifteh et al., 2010). Overall, the findings of this study demonstrate that VNIRS has considerable potential to detect biochemical processes and certain chemical soil properties. However, limitations such as the small sample size, the restricted representation of land use categories, and the weak spectral expression of physical soil attributes constrain model performance. Despite these limitations, the advantages of VNIRS including rapid measurement, low cost, environmental friendliness, and the ability to evaluate multiple soil properties simultaneously make it a valuable tool for both research and practical soil management applications (Stenberg et al., 2010; Ben-Dor, 1999).

CONCLUSION

This study demonstrated that land use type has a significant influence on the physicochemical and biochemical properties of soils developed under semi-arid conditions. Soils cultivated with *Melissa officinalis* exhibited higher organic matter content and enzyme activities (particularly dehydrogenase and β-glucosidase), indicating improved microbial activity and soil health compared to cotton, pistachio, and fallow lands. These results confirm that vegetation type and management practices play a crucial role in maintaining soil biological functionality in fragile dryland ecosystems. The application of Visible–Near Infrared Spectroscopy (VNIRS) combined with Partial Least Squares Regression (PLSR) modeling provided promising results for the estimation of soil biochemical indicators. The models achieved strong calibration performance for key variables such as electrical conductivity, organic matter, and dehydrogenase activity, while validation accuracy was moderate. However, prediction capability for physical parameters (e.g., texture fractions) remained limited, suggesting that these properties are less spectrally active and may require complementary approaches. This outcome is consistent with the generally weak direct spectral expression of particle-size fractions and their dependence on mineralogical composition and microaggregate structure, which reduces their detectability by VNIRS. Distinct absorption bands at 1400, 1900, and 2200 nm associated with soil water and clay mineral vibrations were consistently detected across land use types, confirming the sensitivity of VNIRS to moisture-related features. These findings highlight the method's potential as a rapid, cost-effective, and environmentally friendly alternative to conventional laboratory analyses for soil monitoring. While the study provides valuable preliminary evidence, several limitations should be acknowledged. The relatively small sample size, limited replication across land use categories, and the inherently complex spectral behavior of enzyme activities constrain the robustness of the predictive models. These factors may also contribute to the moderate validation performance observed for several soil properties. Therefore, the results should be interpreted as an initial assessment rather than definitive predictive models.

Despite these constraints, the study contributes to the growing body of work exploring the feasibility of VNIRS for biochemical soil assessment under semi-arid conditions. The inclusion of enzyme activities as prediction targets represents a meaningful step toward expanding VNIRS applications, although further research is required to strengthen model performance and confirm reproducibility.

Future research should focus on (i) expanding the sample size and spatial coverage to enhance model robustness, (ii) integrating advanced spectral preprocessing and machine learning algorithms (e.g., Random Forest, ANN, or hybrid models) to improve predictive accuracy, and (iii) developing portable VNIRS-based field platforms and drone-integrated systems for large-scale soil health monitoring. Such advancements would improve the scalability and reliability of spectral soil assessment, supporting sustainable soil management strategies and precision agriculture practices, particularly in semi-arid regions vulnerable to degradation and nutrient loss.

Compliance with Ethical Standards

Peer Review

This article has been reviewed by independent experts in the field using a rigorous double-blind peer review process.

Conflict of Interest

The authors declare no conflicts of interest.

Author Contributions

All authors contributed equally to the study design, data collection, analysis, and manuscript preparation.

REFERENCES

Acosta-Martínez, V., & Tabatabai, M. A. (2000). Enzyme activities in a limed agricultural soil. *Biology and Fertility of soils*, *31*(1), 85-91.

Álvarez, V. E., Arias-Rios, J. A., Guidalevich, V., Marchelli, P., Tittonell, P. A., & El Mujtar, V. A. (2025). Using near-infrared spectroscopy as a cost-effective method to characterise soil and leaf properties in native forest. *Geoderma Regional*, 40, e00948.

Ben-Dor, E., Irons, J. R., & Epema, G. F. (1999). Soil reflectance. *Remote sensing for the earth sciences: Manual of remote sensing*, *3*(3), 111-188.

Ben-Dor, E., Taylor, R. G., Hill, J., Demattê, J. A. M., Whiting, M. L., Chabrillat, S., & Sommer, S. (2008). Imaging spectrometry for soil applications. *Advances in agronomy*, *97*, 321-392.

Bilgili, A.V., Vas Es, H.M., Akbas, F., Durak, A., Hively, W.D., 2010. Visible-near infrared reflectance spectroscopy for assessment of soil properties in a semi-arid area of Turkey. J. Arid Environ. 74, 229–238. https://doi.org/10.1016/j.jaridenv.2009.08.011.

Bolan, N. S., Adriano, D. C., Kunhikrishnan, A., James, T., McDowell, R. & Senesi, N. (2011). Dissolved organic matter: Biogeochemistry, dynamics, and environmental significance in soils. *Advances in Agronomy*, 110, 1–75.

Bouyoucos G.J. (1951). A Recalibration of The Hydrometer Method for Making Mechanical Analysis of Soils. Agronomy Journal 43: 434-438.

Clark, R. N., Swayze, G. A., Singer, R. B., & Pollack, J. B. (1990). High-resolution reflectance spectra of Mars in the 2.3-μm region: Evidence for the mineral scapolite. *Journal of Geophysical Research: Solid Earth*, *95*(B9), 14463-14480.

Çullu, M. A., Şeker, H., Gozukara, G., Günal, H. & Bilgili, A. V. (2024). Rapid characterization of soil horizons for different soil series utilizing Vis-NIR spectral information. Geoderma Regional, 38, e00853.

Deng S, Popova I (2011) Carbohydrate hydrolases. In: Dick RP (ed) SSSA book series. American Society of Agronomy, Crop Science Society of America, and Soil Science Society of America, pp 185–209. https://doi.org/10.2136/sssabookser9.c9

Eivazi, F., Bayan, M. R., & Schmidt, K. (2003). Select soil enzyme activities in the historic Sanborn Field as affected by long-term cropping systems. *Communications in Soil Science and Plant Analysis*, *34*(15-16), 2259-2275.

Esbensen, K.H., Geladi, P., 2010. Principles of Proper Validation: Use and Abuse of Re-Sampling for Validation. Special Issue Article. J. Chemometrics, 2010;24: 168–187.

Farifteh, J., Tolpekin, V., Van Der Meer, F., & Sukchan, S. (2010). Salinity modelling by inverted Gaussian parameters of soil reflectance spectra. *International Journal of Remote Sensing*, *31*(12), 3195-3210.

Gao, M., Hu, W., Li, M., Wang, S., & Chu, L. (2025). Network analysis was effective in establishing the soil quality index and differentiated among changes in land-use type. *Soil and Tillage Research*, *246*, 106352.

Gao, M., Hu, W., Li, M., Wang, S., & Chu, L. (2025). Network analysis was effective in establishing the soil quality index and differentiated among changes in land-use type. *Soil and Tillage Research*, *246*, 106352.

Gomez, C., Rossel, R. A. V., & McBratney, A. B. (2008). Soil organic carbon prediction by hyperspectral remote sensing and field vis-NIR spectroscopy: An Australian case study. *Geoderma*, *146*(3-4), 403-411.

Gözükara, İ. E. & Demattê, J. A. M. (2025). Prediction accuracy of pXRF, MIR, and Vis-NIR spectra for soil properties: A comparative study. *Soil Science Society of America Journal*, 89(3), 251–264. https://doi.org/10.1002/saj2.20755

Gugino, B.K, Idowu, O.J, Schindelbeck, R.R., Van Es H.M., Wolfe, D.W., Moebius-Clune, B.N., Thies J.E., and Abawi G.S. 2009. Cornell Soil Health Assessment Training Manual, Second Edition,Cornell University, Geneva, New York.

Hopkins, D. W., Alef, K., & Nannipieri, P. (1996). Methods in Applied Soil Microbiology and Biochemistry. *Journal of Applied Ecology*, *33*(1), 178.

Hosseini, E., Zarei, M., Moosavi, A. A., Ghasemi-Fasaei, R., Baghernejad, M., & Mozaffari, H. (2024). Feasibility of Vis-NIR spectroscopy approach to predict soil biological attributes in arid land soils. *Plos one*, *19*(9), e0311122.

Hunt, G. R. (1977). Spectral signatures of particulate minerals in the visible and near infrared. *Geophysics*, *42*(3), 501-513.

Kaplan, F., Bilgili, A.V. (2024). Characterization of Some Properties of Soils Formed on Basalt Parent Material Using Spectroradiometric and Geostatistical Techniques. ISPEC Journal of Agricultural Sciences, 8(3): 583-603. DOI: https://doi.org/10.5281/zenodo.12603883.

Kemper, W.D., Rosenau, R.C. 1986. Aggregate Stability and Size Distribution, In A. Klute et al., Methods of Soil Analysis, part 1, Physical and Minerological Methods, pp.425-442, 2d ed. Agronomy Monograph 9, Soil Science Society of America, Madison.

Klute, A., 1986. Water Retonbon Labaratory Methots in a Klute (ed.) Methots of Soil Analysis. Part I, Physical and Minerological Properties, Argon, No 9, Amer. Soc. Of Agronomy, Inc, Madison, Wisconsin, USA.

Lehmann, J. & Kleber, M. (2015). The contentious nature of soil organic matter. *Nature*, 528(7580), 60–68.

Lobell, D. B., & Asner, G. P. (2002). Moisture effects on soil reflectance. *Soil Science Society of America Journal*, *66*(3), 722-727.

Ma, X., Wang, J., Li, Y., & Zhang, H. (2023). Land-use change and its effects on soil quality: A review. *Environmental Reviews*, 31(2), 119–138.

McLean EO (1982). Soil pH and lime requirement. In: Page AL, Miller RH, Keeney DR, editors. Methods of Soil Analysis, Part 2. Madison, WI, USA: ASA and SSSA, pp. 199–223.

MGM. (2024). Meteorological Observation Data – Sanliurfa Central Station. General Directorate of Meteorology, Ankara. Retrieved from https://www.mgm.gov.tr

Milos, B. & Bensa, A., 2018. Estimation of Organic Carbon and Calcium Carbonates in Agricultural Soils by Vis-NIR Spectroscopy. Poljoprivreda, 24 (1): 45-51.

Nelson, D. W., & Sommers, L. E. (1982). Total carbon, organic carbon, and organic matter. *Methods of soil analysis: Part 2 chemical and microbiological properties*, *9*, 539-579.

Olsen, S. R. (1954). *Estimation of available phosphorus in soils by extraction with sodium bicarbonate* (No. 939). US Department of Agriculture.

Piao, S., Friedlingstein, P., Ciais, P., Zhou, L. & Chen, A. (2009). The carbon balance of terrestrial ecosystems in China. *Nature*, 458(7241), 1009–1013.

Raiesi, F., & Beheshti, A. (2015). Soil quality response to land-use change: A comparative study of agricultural and natural ecosystems. *Applied Soil Ecology*, 95, 73–85.

Rayegani, B., Barati, S., Sohrabi, T. A., & Sonboli, B. (2016). Remotely sensed data capacities to assess soil degradation. *The Egyptian Journal of Remote Sensing and Space Science*, *19*(2), 207-222.

Reeves, J. B., McCarty, G. W., & Meisinger, J. J. (1999). Near infrared reflectance spectroscopy for the analysis of agricultural soils. *Journal of Near Infrared Spectroscopy*, *7*(3), 179-193.

Ren, T., Smreczak, B., Ukalska-Jaruga, A., Li, X., Hassan, W., & Cai, A. (2025). Differential impacts of nitrogen addition on soil dissolved organic carbon in humid and non-humid regions: A global meta-analysis. *Journal of Environmental Management*, 377, 124744.

Ren, Z., Zhang, Y., & Li, F. (2025). Impact of irrigation and fertilization practices on soil salinity in cotton fields. *Agronomy*, 14(4), 1150. https://doi.org/10.3390/agronomy14041150

Rossel, R. V., & Webster, R. (2012). Predicting soil properties from the Australian soil visible–near infrared spectroscopic database. *European Journal of Soil Science*, *63*(6), 848-860.

Schimel, J. P., & Weintraub, M. N. (2003). The implications of exoenzyme activity on microbial carbon and nitrogen limitation in soil: a theoretical model. *Soil Biology and Biochemistry*, *35*(4), 549-563.

Schinner F, Öhlinger R, Kandeler E, Margesin R (eds) (1996) Methods in soil biology. Springer, Berlin. https://doi.org/10.1007/978-3-642-60966-4

Singh, J. S., Pandey, V. C., & Singh, D. P. (2007). Efficient soil carbon sequestration in agroecosystems. *Environmental Management*, 40(3), 506–518.

Sinsabaugh, R. L., & Follstad Shah, J. J. (2012). Ecoenzymatic stoichiometry and ecological theory. *Annual Review of Ecology, Evolution, and Systematics*, 43, 313–343.

Stenberg, B., Rossel, R. A. V., Mouazen, A. M., & Wetterlind, J. (2010). Visible and near infrared spectroscopy in soil science. *Advances in agronomy*, *107*, 163-215.

Stoner, E. R., & Baumgardner, M. F. (1981). Characteristic variations in reflectance of surface soils. *Soil Science Society of America Journal*, *45*(6), 1161-1165.

Tabatabai MA (1982) Methods of soil analysis part 2. Chemical and microbiological properties. In: Soil Enzymes, 2nd edn. American Society of Agronomy, Soil Science Society of America-Madison, pp 903–947

Tabatabai MA, Bremner JM (1969) Use of p-nitrophenyl phosphate for assay of soil phosphatase activity. Soil Biol Biochem 1:301–307

Tsolis, V., & Barouchas, P. (2023). Biochar as soil amendment: The effect of biochar on soil properties using VIS-NIR diffuse reflectance spectroscopy, biochar aging and soil microbiology—A review. *Land*, *12*(8), 1580.

Tüzüner, A., 1990. Toprak ve Su Analiz Laboratuvarları El Kitabı. T.C. Tarım ve Köy İşleri Bakanlığı, Köy Hizmetleri Genel Müdürlüğü, s: 61-73, Ankara.

Wang, G., Tang, X., Zhang, Q., Li, B., & Li, M. (2025). The Relationship Between Soil Organic Matter Composition and Soil Enzymes Activities in Various Land Use Types in the Upper Watershed of Danjiangkou Reservoir in China. *Land Degradation & Development*, *36*(8), 2557-2570.

Wang, S., & Huang, W. (2020). Soil microbial diversity and function in response to land-use change. *Applied Soil Ecology*, 155, 103670.

WRB, World Reference Base for Soil Resources. International Soil Classification System for Naming Soils and Creating Legends for Soil Maps, 4th ed. (IUSS, Vienna, 2022). https://wrb.isric.org/files/WRB_fourth_edition_2022-12-18.pdf