Trends in Analytical Chemistry 185 (2025) 118142

Contents lists available at ScienceDirect

Trends in Analytical Chemistry

journal homepage: www.elsevier.com/locate/trac

Infrared spectroscopy for soil NPK estimation: Advances, challenges, and future directions in predictive modelling

Shengchang Huai [a,b], Qingyue Zhang [a], Yuwen Jin [a], Weijia Yu [c], Jeroen Meersmans [b], Shichao Wang [d], Gilles Colinet [b,**], Changai Lu [a,*]

[a] Institute of Agricultural Resources and Regional Planning, Chinese Academy of Agricultural Sciences, Beijing, China
[b] Gembloux Agro-Bio Tech, University of Liege, Passage des deportes 2, 5030, Gembloux, Belgium
[c] Institute of Crop Germplasm Resources, Shandong Academy of Agricultural Sciences, Jinan, Shandong Province, China
[d] Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Shijiazhuang, Hebei, China

ARTICLE INFO

Keywords:
Soil nutrients
Infrared spectroscopy
NPK estimation
Machine learning
Predictive models
Interference factors

ABSTRACT

Soil nitrogen (N), phosphorus (P), and potassium (K) are vital nutrients that underpin plant health and agricultural productivity. Precise and rapid monitoring of soil NPK levels is crucial for promoting sustainable farming practices. Infrared spectroscopy (IR) has emerged as a promising non-invasive technology for real-time soil nutrient assessment. However, current predictive models face challenges in achieving high precision and robustness due to the inherent heterogeneity of soil and variations in environmental conditions. This review critically examines the latest advancements in infrared spectroscopy for NPK estimation, emphasizing innovations in data acquisition, preprocessing, variable selection, and modelling techniques. The effects of soil heterogeneity and environmental factors on predictive accuracy are thoroughly evaluated, alongside advanced strategies proposed to address these limitations. Particular focus is placed on the role of soil components in influencing spectroscopy-based NPK estimation models, emphasizing the need for future research to refine characteristic NPK spectroscopy band selection. Furthermore, the development of correction sub-models that account for the interference patterns of soil components is recommended to improve model accuracy and stability. Infrared spectroscopy holds great promise for precision agriculture by enabling real-time soil nutrient management, thereby contributing to global food security.

1. Introduction

Soil nutrient management plays a critical role in global food security and sustainable agricultural development [1]. As climate change and soil degradation intensify, global agriculture is facing unprecedented challenges [2,3]. According to the United Nations Food and Agriculture Organization (FAO), approximately 33 % of the world's farmland is moderately to highly degraded [4]. Accurate and regular monitoring of soil nutrients is essential to ensure high crop productivity, reduce environmental pollution, and optimize fertilizer usage. However, traditional soil nutrient assessment methods are often time-consuming, labor-intensive, and reliant on harmful chemicals, highlighting the urgent need for rapid, efficient, and eco-friendly alternatives [5,6].

Infrared spectroscopy (IR) has emerged as a powerful, non-destructive tool for soil nutrient analysis, offering real-time data without the need for complex sample preparation. Its application in soil monitoring aligns with the FAO's call for innovative and sustainable solutions to global soil health challenges [7]. Infrared spectroscopy identifies key soil components by detecting absorption bands corresponding to the vibrational modes of bonds like C–C, C–O, C–H, and N–H, and analyzing their distinct spectral signatures [8–10]. By establishing robust regression models that correlate spectral data with specific soil nutrient indices, IR spectroscopy has enabled the quantitative analysis of soil components [11]. Global research trends (Fig. S1) indicate growing interest and collaboration across institutions, underscoring its importance. However, challenges persist in addressing soil heterogeneity and enhancing the accuracy and adaptability of models across different regions and scales.

The spatial and temporal variability of soil NPK (nitrogen (N), phosphorus (P), and potassium (K)) contents during crop growth

* Corresponding author.
** Corresponding author.
E-mail addresses: Jeroen.Meersmans@uliege.be (J. Meersmans), luchangai@caas.cn (C. Lu).

https://doi.org/10.1016/j.trac.2025.118142
Received 5 November 2024; Received in revised form 19 December 2024; Accepted 9 January 2025
Available online 30 January 2025
0165-9936/

underscores the necessity for rapid monitoring to optimize fertilization and enhance precision agriculture [12]. However, most research on infrared spectroscopy for soil nutrient assessment has concentrated on organic matter and carbon due to the technique's high sensitivity to organic molecular vibrations, yielding highly accurate models (Fig. S2). The successive application of various modelling methods has advanced the high-resolution spectral quantification of soil organic matter [11,13,14]. While numerous studies have applied infrared spectroscopy to predict soil NPK contents, a universally accepted model has yet to emerge. Recent research has increasingly focused on developing NPK estimation models, driven by progress in variable selection techniques, machine learning algorithms, and the mitigation of spectral interference factors like water absorption (Fig. 1). Discrepancies in estimation accuracy arise from differences in preprocessing, modelling techniques, and soil properties. In particular, the weak absorption features of phosphorus and potassium in the visible/near-infrared range (VNIR) highlight the need for more accurate predictive models [15–17].

This review aims to (1) evaluate the methods for data acquisition, spectral data preprocessing, variable selection, and modelling in infrared spectroscopy for soil NPK estimation, (2) critically assess the performance and limitations of current NPK prediction models, and (3) identify key factors influencing feature extraction and model accuracy. Furthermore, novel strategies proposed to enhance spectral feature selection and model development to improve NPK prediction accuracy, supporting the advancement of precision agriculture.

2. Spectral data acquisition

Infrared spectroscopy data acquisition relies on two main techniques: proximal sensing and remote sensing (Fig. 2). Proximal sensing can be performed in two environments: laboratory spectroscopy for controlled settings and field in-situ spectroscopy for real-time, on-site measurements. Remote sensing primarily involves airborne and satellite-based methods, typically using image spectroscopy to cover larger areas (Table 1).

Laboratory infrared spectroscopy is widely used to determine soil properties, increasingly replacing traditional physicochemical methods [19,20]. Soil samples are typically preprocessed through steps like removing debris, drying, grinding, and sieving to a standard 2 mm particle size, which is sufficient for reliable spectral analysis [21]. Although finer particles can improve reflectance, further reduction in size beyond 2 mm shows minimal improvement in prediction accuracy. The measurement process involves irradiating the soil samples with a halogen lamp, followed by whiteboard correction to ensure consistent calibration [22,23]. Recent studies have established standard preprocessing protocols, enhancing the precision and repeatability of soil nutrient analysis using infrared spectroscopy [24].

Field in-situ spectroscopy provides the advantage of real-time, on-site data collection, making it highly suitable for large-scale soil monitoring and precision agriculture [8]. Unlike laboratory-based methods, it captures soil properties in their natural state, reducing the need for sample preparation [25]. However, it is more susceptible to environmental variables like soil moisture, surface roughness, and atmospheric conditions, which can introduce noise into the data [26,27]. To address these challenges, advanced techniques such as external parameter orthogonalization (EPO) have been developed to minimize moisture-related interference, significantly improving the reliability of in-situ measurements [26].

Airborne and satellite-based remote sensing technologies provide wide coverage, frequent monitoring, and cost-effective solutions for large-scale soil analysis [28]. Traditional multispectral methods, such as Multispectral Scanner (MSS) and Thematic Mapper (TM), often face challenges in accurately predicting soil NPK content due to lower spectral resolution and the influence of atmospheric conditions and vegetation [29–31]. However, advances in hyperspectral remote sensing, with improved spectral resolution and signal-to-noise ratios, now allow for more precise detection of soil nutrient variations [32,33]. These technologies are increasingly applied to assess crop growth and

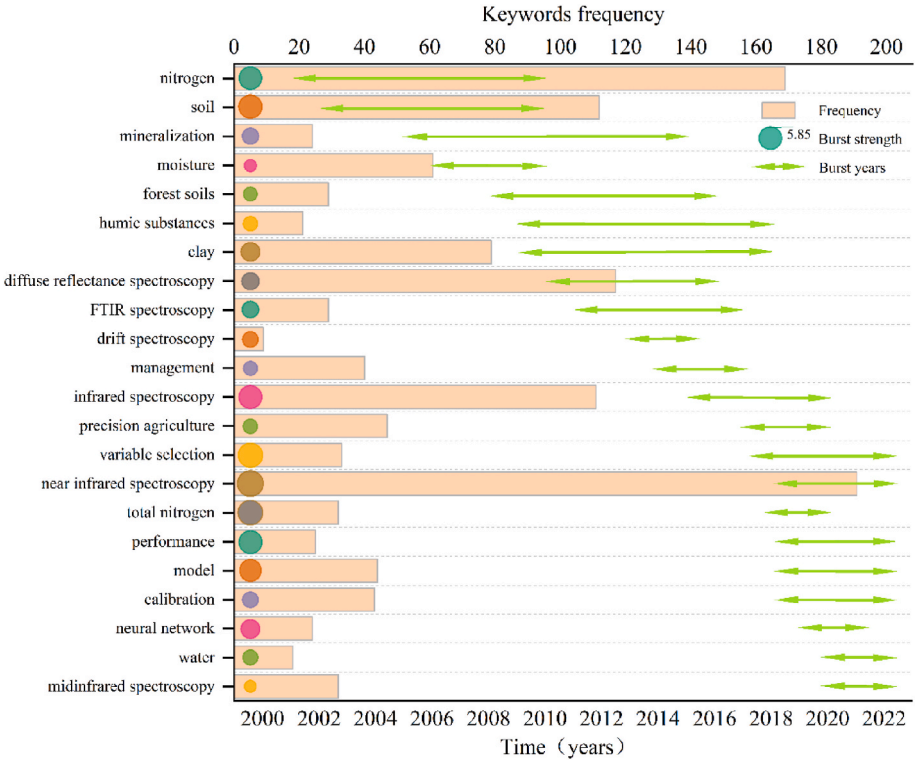

Fig. 1. Burst analysis of keywords on the applications of infrared spectroscopy in soil NPK content during 1995–2023. Note: Publication literatures were collected from the Web of Science from January 1, 2001, to March 29, 2023. The query formula is "TS= (infrared spectroscopy AND (soil nitrogen OR soil phosphorus OR soil potassium))". Burst analysis of keywords were created using citespace-6.2.2 software.

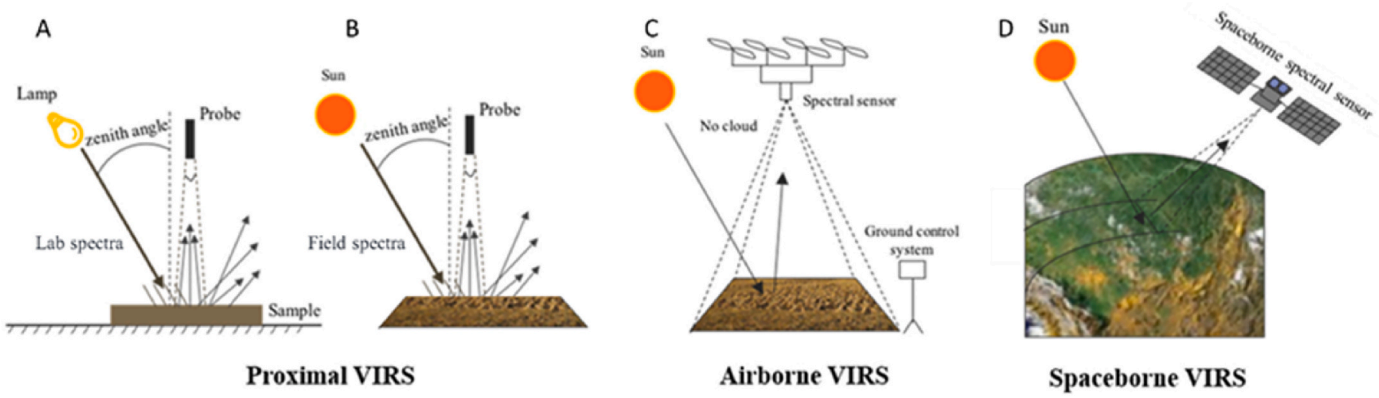

Fig. 2. Schematic diagram of the spectroscopy acquisition equipment setup and spectroscopy acquisition (modified from Xiyue Jia, 2021 [18]). Note: VIRS refers to visible and infrared reflectance spectroscopy.

Table 1
Spectroscopy data acquisition methods for soil NPK estimation.

Method	Advantages	Limitations	Applications
Proximal Sensing (Laboratory)	High accuracy due to controlled environment	Time-consuming, sample preparation; small-scale	Detailed research analysis
Proximal Sensing (In-situ)	Real-time; on-site data; no need for sample preparation	Affected by moisture and surface conditions; less precision than lab methods	Large-scale field monitoring
Remote Sensing (Airborne)	Wide area coverage; frequent monitoring; cost-effective for large-scale	Lower resolution; affected by atmospheric	Regional soil nutrient mapping
Remote Sensing (Satellite)	Global data collection; long-term monitoring	Lower signal-to-noise ratio; interference from vegetation and atmospheric factors	Global agricultural management

estimate soil NPK levels through vegetation indices and spectral analysis. For instance, Naveen et al. predicted soil labile nitrogen content by establishing a relationship between tidal vegetation reflectance from Hyperion remote sensing images and soil nitrogen content [34].

3. Spectroscopy data preprocessing, feature band selection, and modelling methods

The spectra collected by spectrometers contain both sample information and noise from background and stray light. Variations in spectral response across different energy bands result in non-smooth curves, while environmental factors like temperature and illumination can cause baseline shifts and increased scattering [35]. Therefore, spectral data undergo preprocessing to eliminate noise, optimize the range, and remove the effects of extraneous variables [36]. Furthermore, to address redundancy among spectroscopy variables, characteristic bands related to soil NPK content are selected, enhancing model predictive accuracy [37]. Ultimately, appropriate modelling techniques and evaluation systems are employed to develop an accurate soil NPK spectroscopy analysis model.

3.1. Spectral data preprocessing

Spectral data preprocessing is essential for enhancing the accuracy of soil NPK predictions. Common denoising and smoothing methods include Moving Average (MA), Savitzky-Golay (SG) convolution smoothing, Wavelet Transform (WT), Gaussian Smoothing (GS), and Median Filtering (MV) (Table 2). The MA method applies a sliding window to calculate averages across the data set, but an optimal window size must balance noise reduction and retention of key spectral features [10] (Fig. 3). SG smoothing, which uses polynomial least squares fitting, provides more accurate data preservation by emphasizing central data points [38]. WT offers both temporal and spectral analysis, effectively separating noise from meaningful signals [39]. GS, through convolution with a Gaussian kernel, reduces normally distributed noise, while

Table 2
A comparative review of spectral preprocessing techniques for precise soil NPK estimation.

Preprocessing	Description	Advantages	Challenges	Application Scenario
Moving Average (MA)	Applies a sliding window to compute the average	Simple and reducing random noise.	Requires balance in window size to avoid oversmoothing.	Simple signals with significant noise.
Savitzky-Golay (SG)	Uses polynomial fitting to smooth data	Retains signal details, especially sharp features.	Sensitive to polynomial order and window size	Complex spectral data that require shape preservation.
Wavelet Transform (WT)	Decomposes signal in time/frequency domains	Handles both localized and broad spectral features.	More complex to implement and interpret.	Best for multi-scale noise handling and signal analysis.
Gaussian Smoothing (GS)	Convolves with a Gaussian kernel to reduce noise.	Smooths noise while preserving overall data trends.	May blur sharp transitions in the signal.	Suitable for spectra with evenly distributed noise.
Median Filtering (MV)	Replaces each point with the median of its neighbors	Effective at removing spike noise and preserving edges.	Less effective for continuous noise and small variations.	Best for signals with impulsive spike noise.

Formula

$MA: y_{i,smooth} = \frac{1}{2k+1}\sum_{i=-k}^{k} y_{i+j}$

$SG: y_{i,smooth} = \sum_{i=-k}^{k} c_j \cdot y_{i+j} c_j$: The coefficients of the Savitzky-Golay filter

$WT: W(a,b) = \frac{1}{\sqrt{a}}\int_{-\infty}^{\infty} y(t)\psi^*\left(\frac{t-b}{a}\right)dta/b$: Scale/Translation parameter; ψ^*: Complex conjugate

$GS: y_{i,smooth} = \sum_{j=-k}^{k} G(j)\cdot y_{i+j} G(j)$ is the Gaussian function

$MA: y_{i,smooth} = median\{y_{i-k},\ldots,y_i,\ldots,y_{i+k}\}$

Note: y_i represents original data value at position; $y_{i,smooth}$ defines smoothed data value at position; k represents half the window width.

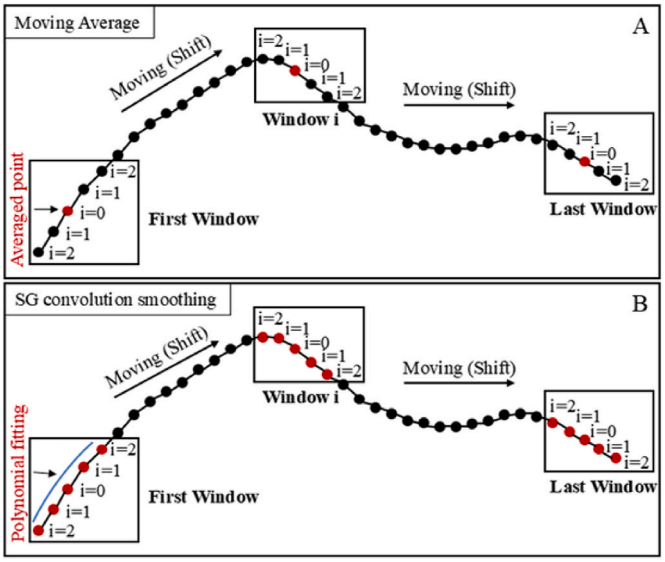

Fig. 3. Illustration of moving average and Savitzky-Golay smoothing methods.

Median Filtering preserves edges by replacing values with neighborhood medians [40].

Multivariate scatter correction (MSC) and standard normal variate (SNV) transformations are crucial for augmenting spectroscopy by elucidating complex features such as steps, peaks, and valleys, which correlate significantly with soil nutrient content while minimizing environmental and instrumental influences [41] (Fig. 4B and C). MSC employs regression analysis between measured and reference spectra, using derived regression coefficients (slope and intercept) to correct the measured data and eliminate scattering effects due to irregular soil particle sizes and distribution. In contrast, SNV standardizes the sample by calculating the standard deviation across all variables and normalizing accordingly. Research suggests that MSC and SNV typically exhibit linear correlation and yield similar results [42].

Selecting the appropriate preprocessing technique is crucial and depends on soil characteristics, environmental conditions, and specific spectral data features. SG smoothing combined with MSC has proven highly effective in predicting soil nutrients [43,44]. Additionally, WT has demonstrated significant advantages for improving nitrogen prediction accuracy [45,46]. By choosing the most suitable methods, the shape and characteristic features of the spectral curve can be preserved, leading to more reliable predictions.

3.2. *Spectral data transformation and feature bands selection*

Spectral transformation techniques, such as derivative processing, are extensively employed to enhance spectral information by removing baseline shifts and improving the identification of key absorption features (Fig. 4D and E). First-order derivatives effectively correct baseline variations and minimize linear background interference, making them particularly advantageous for predicting soil nutrients with low concentrations or weak absorption peaks [47,48]. However, derivative processing inherently retains noise while attenuating the signal, potentially reducing the signal-to-noise ratio. Higher-order derivatives (e.g., third or fourth) can improve spectral resolution but significantly amplify noise, complicating data interpretation and reducing reliability.

After transforming the spectral data, feature band selection is critical to improving model accuracy and efficiency by eliminating irrelevant spectral bands. Common methods include correlation coefficient analysis, stepwise regression, competitive adaptive reweighted sampling (CARS), uninformative variable elimination (UVE), and genetic algorithms (GA). Each method has strengths depending on the complexity of the dataset and the relationships between variables.

Variable selection methods vary significantly in their effectiveness. CARS and UVE are especially useful for simplifying models by removing redundant information, while GA offer a robust approach for selecting optimal wavelengths [49–51]. These methods help improve model performance by focusing on the most informative spectral bands, reducing model complexity, and enhancing predictive power.

3.3. *Quantitative modelling methods for soil nitrogen, phosphorus, and potassium*

Quantitative modelling methods establish mathematical relationships between infrared spectral data and soil NPK content. Commonly used techniques include multivariate linear regression (MLR), partial least squares regression (PLSR), and principal component regression (PCR). Nonlinear methods, such as support vector machines (SVM), random forests (RF), artificial neural networks (ANN), and Gaussian process regression (GPR), are gaining popularity due to their ability to model complex relationships in spectral data.

PLSR, a widely used linear approach, integrates principal component analysis to decompose spectral data and select the most informative components for regression [52]. It has shown consistent performance for predicting soil NPK levels. Nonlinear methods, such as SVM and RF, leverage kernel functions to handle non-linearities and can provide higher accuracy, especially in heterogeneous soil environments [53]. ANN, inspired by biological neural networks, excels in capturing complex patterns and has been increasingly applied in NPK modelling [54].

Recent advances in deep learning, particularly in backpropagation neural networks (BPNN) and convolutional neural networks (CNN), have enhanced model accuracy for soil nutrient prediction. BPNN optimizes prediction by adjusting network weights through gradient descent, while CNN captures spatial relationships in spectral data [55]. These methods often outperform simpler techniques due to their ability to capture nonlinear relationships. However, they typically require larger datasets and precautions to avoid misleading results [56]. For small and simple datasets, simpler models are advisable, as they enhance interpretability by focusing on practical and meaningful predictor variables rather than relying solely on model capabilities to "choose the best variables [57].

4. Accuracy and applicability of different soil nitrogen, phosphorus, and potassium spectroscopy models

Progress has been made in developing infrared spectroscopy models that, with improved feature extraction, preprocessing, and robust regression, better predict soil NPK concentrations. Models for total and available nitrogen show high accuracy due to clearer absorption bands and strong spectral correlations. However, modelling phosphorus and potassium remains challenging due to weaker absorption features, leading to lower prediction accuracy. Advancing feature band selection and regression techniques is key to improving these models. Future research should focus on refining these methods to enhance prediction reliability and robustness, particularly for phosphorus and potassium.

4.1. *Soil nitrogen spectroscopy model*

Numerous studies have demonstrated the potential of using infrared spectroscopy to analyze total nitrogen (TN) content in soil (Table 3). PLSR combined with near-infrared (NIR) or mid-infrared (MIR) spectroscopy is widely used, showing reliable accuracy across different soil types and conditions [58–64]. With a higher number of spectral bands, MIR typically outperforms NIR in predicting TN due to its enhanced sensitivity to nitrogen-related molecular vibrations [65,66]. However, in certain cases, NIR has shown comparable accuracy, particularly when paired with advanced regression models such as SVM and ANN [66–70].

Recent studies have shown that nonlinear methods, including SVM and ANN, provide more accurate predictions than linear methods like

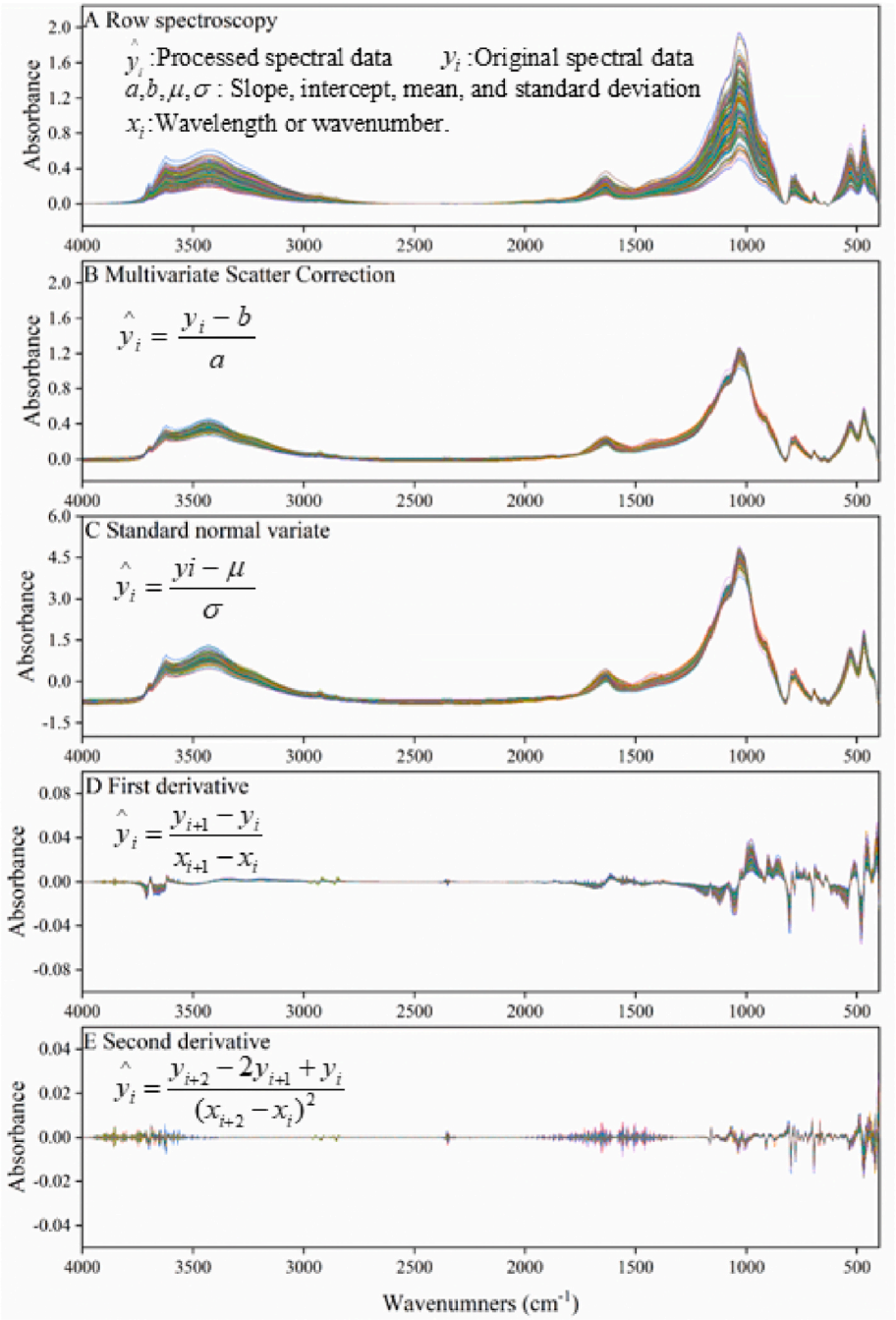

Fig. 4. Changes in spectral curves after preprocessing of mid-infrared spectroscopy. A: raw spectra, B: multiplicative scatter correction, C: standard normal variate, D: First derivative, E: Second derivative.

PLSR, particularly in heterogeneous soil conditions. These models capture complex nonlinear relationships in spectral data, leading to improved TN prediction performance. For example, Shao et al. found that combining infrared spectroscopy with SVM models resulted in higher predictive accuracy for TN compared to traditional PLSR models [91].

In addition to TN, infrared spectroscopy has been applied to estimate soil available nitrogen (AN). Viscarra Rossel et al. summarized studies on soil AN spectroscopy conducted between 1986 and 2002, reporting that the validation R^2 for predicting AN using laboratory-acquired spectra ranged from 0.72 to 0.92 [93]. Typically, the accuracy of in-situ spectroscopy for predicting AN is significantly lower than that of laboratory spectroscopy [70]. Shao et al. using 280 soil samples from Zhejiang, China, established an available nitrogen model by applying a least squares support vector machine (LS-SVM) to NIR, MIR, and NIR-MIR data [91]. The study revealed that SVM combined with NIR data led to high predictive accuracy for AN. Additionally, Paz-Kagan et al. utilized NIR and PLSR models to predict the concentrations of ammonium nitrogen ($N–NH_4^+$) ($R^2 = 0.83$ and ratio of performance to deviation (RPD) = 2.43) and nitrate nitrogen ($N–NO_3^-$) ($R^2 = 0.81$ and RPD = 2.68), yielding satisfactory results.

4.2. Soil phosphorus spectroscopy model

Research on soil total phosphorus (TP) modelling using infrared spectroscopy has faced challenges due to the complex chemical forms of phosphorus in soil and its weak absorption features (Table 4). Unlike nitrogen, phosphorus is present in both organic and inorganic forms,

Table 3
Summary of soil nitrogen content estimations using VNIR and MIR spectroscopy. Soil properties assessed include total nitrogen (TN, g kg^{-1}), available nitrogen (AN, mg kg^{-1}), ammonium nitrogen ($N–NH_4^+$, mg kg^{-1}), and nitrate nitrogen (NO_3^-N, mg kg^{-1}). Prediction quality is categorized as: A (well predicted, $R^2 > 0.8$), B (acceptable prediction, $0.6 < R^2 \leq 0.8$) and C (poor prediction, $R^2 \leq 0.6$). RPD refers to the ratio of performance to deviation. The R^2 and RPD values presented are validation R^2 and validation RPD. MBL refers to memory-based learning. M-PLSR refers to Modified-PLSR.

Predicted properties	Regression technique (R^2; RPD)	IR technique	Measure	Samples	Concentration range	References
TN	PLSR (0.86; -)[B]	NIR	Lab	165	0.06-0.08	[58]
TN	PLSR (0.80; 2.77)[B]	NIR	Lab	400	0.08–12.56	[71]
TN	PLSR (0.93; -)[A]	NIR	Lab	60	–	[59]
TN	PLSR (0.98; 7.56)[A]	NIR	Lab	431	0.47-3.10	[60]
TN	PLSR (0.79; 2.21)[A]	NIR	Lab	308	0.7–12.2	[72]
TN	PLSR (0.74; 2.0)[B]	NIR	Lab	70	0.6-3.5	[73]
TN	PLSR (0.77; 2.00)[B]	NIR	Lab	360	0.74-4.67	[74]
TN	PLSR (0.87; 2.60)[A]	NIR	In suit	130	0.7-2.4	[61]
TN	PLSR (0.86; 2.66)[A]	NIR	In suit	131	0.18-10.12	[62]
TN	PLSR (0.69; 1.78)[A]	NIR	In suit	130	–	[75]
TN	PLSR (0.72; 1.89)[B]	VNIR	Lab	62	0.43-2.18	[76]
TN	PLSR (0.89; 3.0)[B]	VNIR	Lab	122	1.5–5.5	[63]
TN	PLSR (0.86; 1.96)[A]	MIR	Lab	291	–	[77]
TN	PLSR (0.95; -)[A]	MIR	Lab	180	10.2–126.0	[64]
TN	PLSR (0.35; 1.30)[C]	MIR	Lab	68	0.5–20.6	[78]
TN	PLSR (0.85; 4.12)[A]	NIR	Lab	544	0.3-2.2	[65]
	PLSR (0.93; 4.39)[A]	MIR				
TN	PLSR (0.95; 4.12)[A]	NIR	Lab	217	0.6-2.8	[79]
	PLSR (0.96; 4.39)[A]	MIR				
TN	PLSR (0.54; -)[C]	NIR	Lab		–	[15]
	PLSR (0.30; -)[C]	MIR				
TN	PLSR (0.73; 1.89)[B]	VNIR	Lab	458	0–3.8	[80]
	PLSR (0.71; 1.82)[B]	MIR				
TN	PLSR (0.75; -)[B]	NIR	Lab	2825	0.05–7.50	[81]
	PLSR (0.74; -)[B]	MIR				
	PLSR (0.75; -)[B]	NIR–MIR				
TN	M-PLSR (0.89; 3.0)[A]	VNIR	In suit	201	0.04-0.46	[82]
TN	ANN (0.86; 2.54)[A]	NIR	Lab	151	0.8-2.8	[67]
TN	RF (0.86; -)[A]	MIR	Lab	700	0.1–10.0	[83]
TN	PLSR (0.73; -)[B]	VNIR	Lab	701	0.10-3.62	[66]
	SVM (0.75; -)[B]					
	PLSR (0.90; -)[A]	MIR				
	SVM (0.86; -)[A]					
TN	PLSR (0.86; -)[A]	MIR	Lab	20153	0–6.91	[69]
	ANN (0.97; -)[A]					
TN	PLSR (0.87; 2.81)[A]	VNIR	Lab	183	0.3–3.7	[70]
	PLSR (0.86; 2.68)[A]		In suit			
	SVM (0.88; 3.05)[A]		In suit			
TN	PLSR (0.92; 3.56)[B]	VNIR	Lab	105	0.6-3.9	[84]
	MLR (0.96; 5.16)[C]					
TN	PLSR (0.71; 1.82)[B]	MIR	Lab	458	0–3.8	[85]
	GPR (0.85; 2.52)[A]					
TN	PLSR (0.82; 2.35)[A]	VNIR	Lab	148	0.39-4.25	[86]
	PCR (0.78; 2.13)[B]					
	BPNN (0.83; 2.44)[A]					
	SVM (0.86; 2.69)[A]					
TN	PLSR (0.95; -)[A]	MIR	Lab	54211	0–419	[87]
	MBL (0.97; -)[A]					
AN	PLSR (0.69; 1.60)[B]	NIR	In suit	130	1.5–15.4	[61]
AN	BPNN (0.90; -)[A]	VNIR	Lab	144	74.5–181.2	[88]
AN	PLSR (0.44; 1.32)[C]	MIR	In suit	300	25.6–133.5	[89]
AN	PLSR (0.71; -)[B]	VNIR	Lab	100	14.04–60.41	[90]
AN	PLSR (0.86; 2.49)[A]	VNIR	Lab	183	15.8–295.0	[70]
	PLSR (0.76; 1.91)[B]		In suit			
	SVM (0.76; 1.91)[B]		In suit			
AN	PLSR (0.85; -)[A]	NIR	Lab	280	32.26–208.30	[91]
	SVM (0.90; -)[A]					
	PLSR (0.83; -)[A]	MIR				
	SVM (0.87; -)[A]					
	PLSR (0.84; -)[A]	NIR–MIR				
	SVM (0.88; -)[A]					
$N–NH_4^+$	PLSR (0.83; 2.43)[A]	VNIR	Lab	220	–	[92]
$N–NH_4^+$	PLSR (0.92; 3.40)[A]	VNIR	Lab	105	0.45–110.99	[84]
	MLR (0.92; 3.49)[A]					
$N–NO_3^-$	PLSR (0.81; 2.68)[A]	VNIR	Lab	220	–	[92]
$N–NO_3^-$	PLSR (0.73; 1.89)[B]	VNIR	Lab	105	1.32–126.56	[84]
	MLR (0.53; 1.37)[C]					

Table 4
Summary of soil phosphorus content estimations using VNIR and MIR spectroscopy. Soil properties assessed include total phosphorus (TP, mg kg^{-1}), available phosphorus (P-avl, mg kg^{-1}). Prediction quality is categorized as: A (well predicted, $R^2 > 0.8$), B (acceptable prediction, $0.6 < R^2 \leq 0.8$) and C (poor prediction, $R^2 \leq 0.6$). RPD refers to the ratio of performance to deviation. The R^2 and RPD values presented are validation R^2 and validation RPD. M-PLSR refers to Modified-PLSR. The units in references 15 and 98 are $cmol_c\ dm^{-3}$ and kg ha^{-1}, respectively.

Predicted properties	Regression technique (R^2; RPD)	IR technique	Measure	Samples	Concentration range	References
TP	PLSR (0.26; -)[C]	VNIR	Lab	36	211–1232	[97]
TP	PLSR (0.61; 1.60)[B]	VNIR	Lab	146	10–7490	[95]
TP	PLSR (0.69; 2.82)[B]	VNIR	Lab	448	0.01–2870	[96]
TP	PLSR (0.07; 0.93)[C]	NIR	In suit	216	–	[75]
TP	PLSR (0.36; 1.26)[C]	VNIR	Lab	450	7.00–591.96	[94]
	RF (0.25; 1.16)[C]					
	PLSR (0.39; 1.29)[C]	MIR				
	RF (0.20; 1.13)[C]					
TP	PCR (0.69; 1.78)[B]	VNIR	Lab	148	220–1100	[86]
	PLSR (0.72; 1.90)[B]					
	BPNN (0.74; 1.95)[B]					
	SVM (0.76; 2.03)[B]					
TP	PLSR (0.55; -)[C]	VNIR	Lab	379	130–913	[98]
	Cubist (0.65; -)[B]					
	SVM (0.85; -)[C]					
	XGBoost (0.77; -)[C]					
AP	PLSR (0.26;-)[C]	VNIR	Lab	798	12–1660	[99]
AP	PLSR (0.56; 1.51)[C]	VNIR	Lab	146	0.06–124.89	[95]
AP	PLSR (0.69; 1.80)[B]	VNIR	In suit	175	19.90–121.91	[100]
AP	PLSR (0.25; 1.1)[C]	NIR	Lab	148	0–69.6	[101]
AP	PLSR (0.69; 2.2)[B]	NIR	Lab	70	6–212	[73]
AP	PLSR (0.95; 4.53)[A]	NIR	Lab	431	0–64	[60]
AP	PLSR (0.72; 1.80)[B]	NIR	In suit	130	–	[61]
AP	PLSR (0.10; 1.02)[C]	NIR	In suit	216	–	[75]
AP	PLSR (0.48; 0.50)[C]	MIR	Lab	291	–	[77]
AP	PLSR (0.50; 0.81)[C]	MIR	Lab	1456	3.15–33.20	[102]
AP	PLSR (0.07; -)[C]	MIR	Lab	186	50–900	[103]
AP	PLSR (0.43; 1.31)[C]	MIR	In suit	300	60.0–289.0	[89]
AP	PLSR (0.62; -)[B]	MIR	Lab	54211	0–224	[87]
AP	PLSR (0.05; 0.01)[C]	VNIR	Lab	198	2.0–117.0	[104]
	PLSR (0.23; 1.9)[C]	MIR				
AP	PLSR (0.90; -)[A]	NIR	Lab	100	4.0–540.0	[15]
	PLSR (0.84; -)[A]	MIR				
AP	PCR (0.80; -)[B]	NIR	In suit	108	13.2–248.1	[105]
AP	ANN (0.81; -)[A]	VNIR	Lab	41	–	[106]
AP	BPNN (0.81; 2.23)[A]	VNIR	Lab	153	–	[107]
AP	BPNN (0.82; -)[A]	VNIR	Lab	144	6.0–304.7	[88]
AP	M-PLSR (0.70; 1.80)[B]	VNIR	In suit	201	17.5–177.5	[82]
AP	SVM (0.80; 2.27)[A]	VNIR	Lab	235	0.47-46.14	[108]
AP	SVM (0.24; -)[C]	VNIR	Lab	1259	–	[109]
	SVM (0.38; -)[C]	MIR				
AP	MLR (0.52; -)[C]	NIR	Lab	118	0.42–161.00	[110]
AP	PLSR (0.39; 1.17)[C]	VNIR	Lab	105	2.15–337.52	[84]
	MLR (0.48; 1.33)[C]					
AP	PLSR (0.14; 1.08)[C]	MIR	Lab	20000	0–595	[69]
	ANN (0.16; 1.09)[C]					
AP	PLSR (0.82; -)[A]	NIR	Lab	280	25.26–343.50	[91]
	SVM (0.83; -)[A]					
	PLSR (0.85; -)[A]	MIR				
	SVM (0.88; -)[A]					
	PLSR (0.83; -)[A]	NIR–MIR				
	SVM (0.85; -)[A]					
AP	PLSR (0.29; 1.17)[C]	VNIR	Lab	183	0.7–108.0	[70]
	PLSR (0.43; 1.33)[C]		In suit			
	SVM (0.36; 1.27)[C]					

complicating its spectral detection [94]. PLSR combined with VNIR spectroscopy has been used to achieve moderate success in predicting TP content, but the accuracy remains suboptimal due to the poor correlation between phosphorus's spectral characteristics and its concentration in the soil [15,60,86,95,96].

Machine learning models, such as SVM and ANN, show promise in improving the predictive power of phosphorus models by capturing nonlinear relationships in the data [68,91,107,108]. However, the generalization ability of these models still needs improvement, particularly when dealing with diverse soil types and varying phosphorus availability [75,77,95,101–104]. Furthermore, MIR spectroscopy, which provides more detailed molecular information, has shown potential for enhancing phosphorus predictions but requires further research to optimize its application. The introduction of feature selection methods, such as CARS and UVE, has contributed to better identification of phosphorus-related spectral bands. However, the challenge of extracting reliable spectral information from weak phosphorus absorption signals remains a key area for future research [68,100,106].

4.3. Soil potassium spectroscopy model

PLSR has also been applied to determine the total potassium content in soil (Table 5). The weak and dispersed infrared absorption features of total soil potassium pose challenges for PLSR models in extracting sufficient information for accurate predictions [75,111]. Despite the inherent advantage of neural network models in managing complex

Table 5
Summary of soil potassium content estimations using VNIR and MIR spectroscopy. Soil properties assessed include total potassium (TK, g kg^{-1}), available potassium (K-avl, mg kg^{-1}) and exchangeable potassium (K-exch, mg kg^{-1}). Prediction quality is categorized as: A (well predicted, $R^2 > 0.8$), B (acceptable prediction, $0.6 < R^2 \leq 0.8$) and C (poor prediction, $R^2 \leq 0.6$). RPD refers to the ratio of performance to deviation. The R^2 and RPD values presented are validation R^2 and validation RPD. M-PLSR refers to Modified-PLSR. The units in references 79, 93, 87, 74, 75, and 116 are cmolc kg^{-1}. The units in references 60 and 98 are $cmol_c$ dm^{-3} and kg ha^{-1}, respectively.

Predicted properties	Regression technique (R^2; RPD)	IR technique	Measure	Samples	Concentration range	References
TK	PLSR (0.72; 1.68)[B]	VNIR	Lab	146	2.7–46.0	[95]
TK	PLSR (0.39; 1.26)[C]	NIR	In suit	216	–	[75]
TK	PLSR (0.12; -)[C]	VNIR	In suit	134	–	[111]
TK	M-PLSR (0.72; -)[B]	VNIR	Lab	317	0.13-1.40	[113]
TK	PCR (0.54; 1.78)[C]	VNIR	Lab	148	4.97-29.94	[86]
	PLSR (0.58; 1.90)[C]					
	BPNN (0.65; 1.95)[B]					
	SVM (0.63; 2.03)[B]					
TK	PCR (0.29; 1.19)[C]	VNIR	Lab	178	–	[112]
	PLSR (0.34; 1.23)[C]					
	LS-SVM (0.22; 1.14)[C]					
	BPNN (0.45; 1.34)[C]					
AK	PLSR (0.34;-)[C]	VNIR	Lab	798	40–2050	[99]
AK	PLSR (0.79; 2.27)[B]	NIR	Lab	431	0.05-1.60	[60]
AK	PLSR (0.01; 1.02)[C]	NIR	In suit	216	–	[75]
AK	PLSR (0.34; -)[C]	MIR	Lab	186	–	[103]
AK	RF (0.51; -)[C]	MIR	Lab	700	0.1–90.0	[83]
AK	PCR (0.6; -)[B]	NIR	Lab	106	114–849	[105]
AK	ANN (0.8; -)[A]	VNIR	Lab	41	–	[106]
AK	BPNN (0.94; -)[A]	VNIR	Lab	144	80.2–866.5	[68]
AK	PLSR (0.07; 0.77)[C]	VNIR	Lab	183	32.5–105.0	[70]
	PLSR (0.03; 0.89)[C]		In suit			
	SVM (0.14; 0.91)[C]					
AK	PLSR (0.71; 1.59)[B]	VNIR	Lab	105	25.89–6600.24	[84]
	MLR (0.95; 4.26)[C]					
AK	PLSR (0.80; -)[A]	NIR	Lab	280	41.52–343.5	[91]
	SVM (0.83; -)[A]					
	PLSR (0.85; -)[A]	MIR				
	SVM (0.89; -)[A]					
	PLSR (0.83; -)[A]	NIR–MIR				
	SVM (0.85; -)[A]					
K-exch	PLSR(0.62;-)[C]	VNIR	Lab	36	0–3.1	[97]
K-exch	PLSR (0.70; 1.5)[B]	NIR	Lab	70	210–890	[73]
K-exch	PLSR(0.28; 1.1)[C]	NIR	Lab	148	0–1.4	[101]
K-exch	PLSR (0.18; 1.09)[C]	MIR	In suit	300	35–360	[89]
K-exch	PLSR (0.33; -)[C]	MIR	Lab	183	–	[103]
K-exch	PLSR (0.45; 1.32)[C]	VNIR	Lab	458	0.1–4.2	[80]
	PLSR (0.39; 1.27)[C]	MIR				
K-exch	PLSR (0.47; -)[C]	NIR	Lab	2825	0.04-1.78	[81]
	PLSR (0.54; -)[C]	MIR				
	PLSR (0.59; -)[C]	NIR–MIR				
K-exch	MPLS (0.90; 3.2)[A]	VNIR	In suit	201	31–684	[82]
K-exch	PCR (0.55; -)[C]	VNIR	Lab	149	16.2–1757.2	[114]
K-exch	BPNN (0.82; -)[B]	VNIR	Lab	168	70–780	[115]
K-exch	SVM (0.32; 1.22)[C]	VNIR	Lab	233	14–438	[108]
K-exch	SVM (0.21; -)[C]	VNIR	Lab	1259	–	[109]
	SVM (0.25; -)[C]	MIR				
K-exch	MLR (0.48; -)[C]	NIR	Lab	118	–	[110]
K-exch	PLSR (0.63; 1.44)[B]	VNIR	Lab	168	26–93	[116]
	BPNN (0.68; 1.52)[B]					
K-exch	PLSR (0.50; -)[C]	MIR	Lab	54211	0–1150	[87]
	MBL (0.72; -)[B]					
K-exch	PLSR (0.64; -)[B]	MIR	Lab	40000	0–32.33	[117]
	CNN (0.79; -)[B]					

nonlinear relationships and their demonstrated improvements over PLSR and PCR models in predicting soil total potassium content, their accuracy remains suboptimal and requires further enhancement to achieve satisfactory levels [86,112].

Whether using near-infrared or mid-infrared spectroscopy, the accuracy of PLSR models for predicting soil available potassium, including exchangeable potassium, remains generally low [70,80,81,101,103,97,99]. The introduction of neural network models has significantly improved the prediction accuracy for soil AK and K-exch, with the R^2 value for soil available potassium exceeding 0.8 [68,106]. SVM models demonstrate high prediction accuracy for soil available potassium but show lower accuracy for exchangeable potassium [91,108,109].

5. Factors influencing NPK spectroscopy model accuracy and directions for future research

Soil spectral reflectance is influenced by measurement conditions and intrinsic properties like moisture, organic matter, and minerals (Fig. 5). At specific wavelengths, soil spectral characteristics is tied to molecular vibrations and electronic transitions, directly impacting NPK quantification [118]. Accurately identifying and correcting for interference factors, such as moisture and surface roughness, is crucial for enhancing the precision and stability of NPK spectral models.

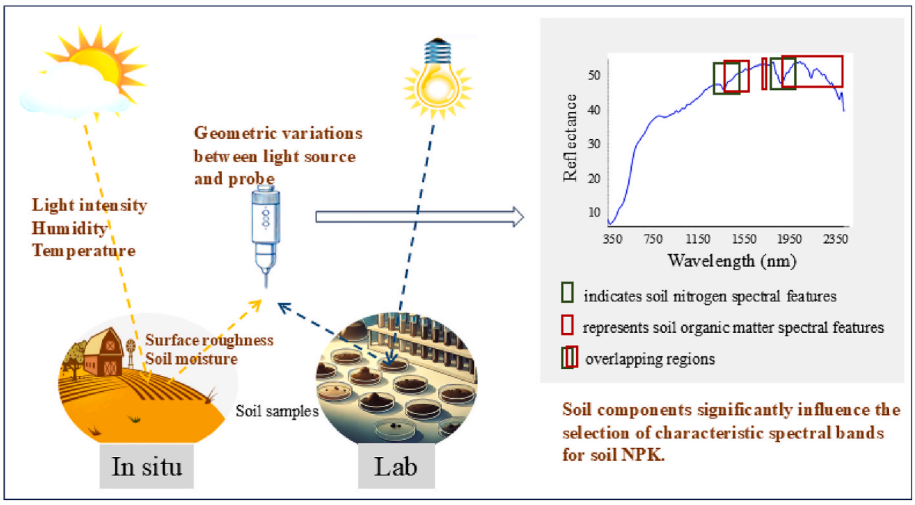

Fig. 5. Factors influencing the accuracy of soil NPK spectral models under in situ and laboratory conditions.

5.1. Spectroscopy measurements

Laboratory factors affecting analysis mainly include geometric conditions like light sources, probe-to-soil distance, and angles. Increased distance between soil samples and sensors significantly affects sensor output due to external light [119]. In situ measurements are primarily influenced by the relative movement between the soil surface and the spectroscopy probe. Studies indicate that after baseline correction, this relative movement is the main source of error, rather than the variation in distance between the soil surface and the probe [120]. In addition, in situ spectroscopy is influenced by environmental factors such as light, humidity, and soil surface properties (e.g., roughness, clods, slope) [121]. Soil moisture also significantly affects in situ measurements. Rossel et al. compared spectra of air-dried, ground, and sieved soil samples in the lab with untreated field samples, identifying key differences at 1400 nm and 1900 nm [122].

Methods like EPO, spectral transformation (DS), and orthogonal signal correction (OSC) are used to eliminate environmental influences such as soil moisture and surface characteristics. EPO, introduced by Roger et al. for removing temperature effects in fruit sugar content prediction, projects spectra onto a space orthogonal to the influencing factors [123]. This approach has since been applied to soil moisture and roughness. DS analyses the difference spectra between field and lab measurements to establish a correlation and remove environmental factors [27]. Despite receiving limited attention, OSC demonstrates potential advantages over conventional methods, as it decouples laboratory sample requirements from field locations and effectively eliminates extraneous variability, such as environmental noise and instrumental artifacts, from the data [124].

5.2. Feature band selection

Since soil spectra are the result of the comprehensive inversion of soil components, the selection of soil NPK spectroscopy features are influenced by other soil parameters. Therefore, it is essential to identify the key factors that interfere with NPK feature extraction and understand their patterns. Mitigating the impact of interference factors remains a key focus for future research aimed at improving the accuracy of prediction models.

5.2.1. Soil minerals

Soil minerals constitute more than 90 % of dry soil, and the parent material of the soil determines the fundamental shape (skeleton) of the soil reflectance spectroscopy [125]. Clay minerals (e.g., montmorillonite, kaolinite, illite) exhibit distinct infrared absorption bands. These mineral-specific bands can overlap or interfere with the characteristic absorption bands of NPK, reducing quantification accuracy [126]. For instance, the O–H group in clay minerals displays prominent absorption peaks in the near and mid-infrared regions, which may coincide with phosphate absorption peaks, potentially complicating the prediction of phosphorus content [127,128]. The particle size, morphology, and surface properties of soil minerals can affect their infrared spectral performance. For example, finer particles have larger specific surface areas, leading to stronger scattering effects, which may weaken or enhance absorption intensity at specific wavelengths, thereby impacting the spectral identification of NPK characteristics [129]. Iron oxides can significantly affect the soil's reflective and absorptive spectra, leading to characteristic absorption peaks in specific bands, such as the visible region (500–700 nm) and the near-infrared region (900–1100 nm) [130,131]. Soils with higher iron oxide content typically exhibit lower reflectance [132]. The strong absorptive capacity of iron oxides can interfere with the spectral signals of NPK, causing their characteristic absorption peaks to be obscured or shifted. This interference may reduce the accuracy of spectral models in predicting soil nutrient content. Additionally, carbonates (such as calcite and dolomite) exhibit specific absorption features in the NIR and MIR spectral regions, which can overlap or interfere with the absorption peaks of NPK elements. This overlap can hinder the accurate identification and quantification of the characteristic absorption peaks of NPK in soil spectra [133].

5.2.2. Soil organic matter

Soil spectral reflectance is significantly negatively correlated with soil organic matter content [134]. The removal of soil organic matter can enhance the soil reflectance within the 400–2500 nm range [163]. Ben-Dor et al. demonstrates that the decrease in reflectance is particularly significant in soils with high organic matter content [135]. Studies indicate that organic matter exhibits absorption peaks between 1100 nm and 2500 nm [135,136], which overlap with the characteristic absorption regions of soil NPK. This overlap can obscure the spectral signals of NPK, making them difficult to distinguish and thus reducing the accuracy of predictive models. Furthermore, organic matter may cause shifts or distortions in NPK absorption peaks due to potential physical or chemical interactions between organic matter and NPK elements [135]. For instance, acidic functional groups in organic matter might react with soil phosphates, altering the chemical environment of phosphorus and causing slight shifts in its spectral absorption peaks [137,138]. The presence of organic matter complicates the effective modelling of soil NPK spectral signals using simple linear regression, necessitating the use of more advanced multivariate analysis methods, such as SVM.

5.2.3. Mitigation of interference effects

Soil spectra encapsulate the intricate interactions among various soil components, with the spectral characteristics of NPK being significantly influenced by minerals, organic matter, and iron oxides [8]. Enhancing the accuracy of NPK prediction models necessitates the mitigation of interference from these non-target components [139]. Although spectral preprocessing can reduce some interference, its effectiveness is limited. A more promising strategy involves excluding spectral bands significantly affected by other soil components (such as clay minerals, iron oxides, and organic matter), while selecting those closely related to NPK. In this context, the selection of an appropriate calibration set plays a vital role in mitigating interference [140,141]. A diverse and representative calibration set can ensure that the model effectively distinguishes between target NPK features and interfering factors, such as overlapping absorption bands of organic matter or minerals. A well-designed calibration set not only enhances the robustness of correction models but also improves the reliability of spectral band selection by elucidating key patterns in spectral data [142,143]. For instance, by including samples spanning a wide range of organic matter and clay contents, such a calibration set enables models to better accommodate spectral variability induced by these factors. Additionally, understanding the mechanisms of interference and integrating correction factors for components such as organic matter and clay into models can further enhance accuracy.

Future research should prioritize the following: 1) systematic identification of interference mechanisms and their corresponding spectral ranges to minimize overlap with target features, 2) development of advanced correction sub-models to address these confounding factors, and 3) establishment of standardized guidelines for calibration set construction to optimize its effectiveness in mitigating interference. These advancements will collectively improve the precision and reliability of NPK prediction models, ultimately supporting the development of precision agriculture and sustainable soil management practices.

5.3. Future modelling strategy

The inherent complexity of soil components often introduces nonlinearities in spectral data due to significant interference among these components. Machine learning techniques, particularly ANN, have gained attention due to their robust nonlinear mapping capabilities. Unlike traditional regression models, ANN operate through a data-driven approach, identifying patterns without requiring explicit theoretical assumptions about variable independence [144]. While this capability is advantageous, it also increases the risk of overfitting—a phenomenon where models excel on training data but perform poorly on validation or unseen datasets [145]. Overfitting is commonly caused by limited dataset size, excessive model complexity, and noise within spectral data, leading to high performance indicators (e.g., R^2) during calibration but reduced prediction accuracy during validation or practical application.

To address overfitting, several strategies have proven effective. Beyond careful selection of representative calibration datasets and optimized feature band selection, regularization techniques like Ridge and LASSO Regression control model complexity by introducing penalty terms, reducing the risk of overfitting [146]. Furthermore, k-fold cross-validation, when paired with independent test sets, not only provides a reliable evaluation of model performance but also enhances its generalization across diverse datasets [147].

Future research encourages the integration of statistical and machine learning models to develop hybrid approaches that balance data interpretability with strong generalization while addressing critical challenges like overfitting and interference. For instance, combining neural networks with chemometric methods—such as principal component regression neural networks (PCR-ANN) and partial least squares regression neural networks (PLSR-ANN)—has demonstrated potential to enhance model robustness, resistance to interference, and accuracy in infrared spectroscopy analysis [148,149]. Additionally, localized regression methods tailored to specific soil properties have proven effective in addressing both heterogeneity and nonlinear relationships [150,151]. By constructing calibration subsets based on criteria such as soil texture, organic matter content, or mineral composition, these methods improve model precision and applicability in diverse soil environments. For instance, selecting samples with properties closely matching the estimation set enables models to accommodate large concentration ranges and significant sample variability, thereby enhancing both accuracy and robustness in practical applications [151].

6. Conclusion

The application of infrared spectroscopy for soil NPK estimation has shown significant promise in advancing precision agriculture through its potential for real-time, non-destructive soil nutrient assessment. This review highlights the substantial progress made in enhancing data acquisition techniques, refining spectral preprocessing methods, and optimizing modelling approaches, particularly with the integration of machine learning techniques. Despite these advancements, challenges remain, particularly in accounting for soil heterogeneity and environmental variability, which continue to impact model robustness and accuracy.

Future research must focus on refining spectral feature selection processes and developing more sophisticated correction sub-models to address the interference caused by soil components such as moisture, organic matter, and mineral content. Machine learning models, particularly deep learning algorithms, offer new avenues for improving the predictive accuracy of NPK levels, but they require larger, more diverse datasets to avoid overfitting and enhance generalizability across different soil types and conditions.

To fully unlock the potential of infrared spectroscopy in soil nutrient management, further efforts are needed to improve phosphorus and potassium prediction models, which remain less accurate due to their weak absorption features. Collaborative research, leveraging the synergy between hyperspectral data and advanced predictive algorithms, could propel the field forward, making soil NPK infrared spectroscopy a critical tool in addressing global food security challenges and promoting sustainable agricultural practices.

CRediT authorship contribution statement

Shengchang Huai: Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Investigation, Conceptualization. **Qingyue Zhang:** Writing – review & editing, Writing – original draft, Visualization, Methodology. **Yuwen Jin:** Writing – review & editing, Writing – original draft, Visualization, Validation. **Weijia Yu:** Writing – original draft, Visualization, Validation, Methodology. **Jeroen Meersmans:** Writing – review & editing, Visualization, Validation. **Shichao Wang:** Writing – review & editing, Visualization, Validation, Supervision. **Gilles Colinet:** Writing – original draft, Visualization, Validation, Methodology, Conceptualization. **Changai Lu:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Project administration, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Changai Lu reports financial support was provided by Ministry of Science and Technology of the People's Republic of China.

Acknowledgements

This study was supported by Mechanism of Soil Spectral Response in Complex Environments and Field Validation (CAAS-CSAL-202402) and the National Key Research and Development Program of China (2022YFD1500100).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.trac.2025.118142.

Data availability

Data will be made available on request.

References

[1] M. Cavallito, FAO: "Soil protection is the first response to the global food crisis," Re Soil Foundation. https://resoilfoundation.org/en/agricultural-industry/fao-soil-nutrients-food-crisis/, 2022. (Accessed 11 September 2024).
[2] R.V. Rojas, M. Achouri, J. Maroulis, L. Caon, Healthy soils: a prerequisite for sustainable food security, Environ. Earth Sci. 75 (180) (2016), https://doi.org/10.1007/s12665-015-5099-7 s12665-015-5099–7.
[3] L. Qiao, X. Wang, P. Smith, J. Fan, Y. Lu, B. Emmett, R. Li, S. Dorling, H. Chen, S. Liu, T.G. Benton, Y. Wang, Y. Ma, R. Jiang, F. Zhang, S. Piao, C. Müller, H. Yang, Y. Hao, W. Li, M. Fan, Soil quality both increases crop production and improves resilience to climate change, Nat. Clim. Change 12 (2022) 574–580, https://doi.org/10.1038/s41558-022-01376-8.
[4] Food and Agriculture Organization of the United Nations, in: The Future of Food and Agriculture: Trends and Challenges, Food and Agriculture Organization of the United Nations, Rome, 2017.
[5] P. Musa, H. Sugeru, E.P. Wibowo, Wireless sensor networks for precision agriculture: a review of NPK sensor implementations, Sensors 24 (2023) 51, https://doi.org/10.3390/s24010051.
[6] M. Zeraatpisheh, S. Ayoubi, A. Jafari, S. Tajik, P. Finke, Digital mapping of soil properties using multiple machine learning in a semi-arid region, central Iran, Geoderma 338 (2019) 445–452, https://doi.org/10.1016/j.geoderma.2018.09.006.
[7] L. Montanarella, The global soil partnership, IOP Conf. Ser. Earth Environ. Sci. 25 (2015) 012001, https://doi.org/10.1088/1755-1315/25/1/012001.
[8] M. Nocita, A. Stevens, B. van Wesemael, M. Aitkenhead, M. Bachmann, B. Barthès, E. Ben Dor, D.J. Brown, M. Clairotte, A. Csorba, P. Dardenne, J.A. M. Demattê, V. Genot, C. Guerrero, M. Knadel, L. Montanarella, C. Noon, L. Ramirez-Lopez, J. Robertson, H. Sakai, J.M. Soriano-Disla, K.D. Shepherd, B. Stenberg, E.K. Towett, R. Vargas, J. Wetterlind, Chapter four - soil spectroscopy: an alternative to wet chemistry for soil monitoring, in: D.L. Sparks (Ed.), Advances in Agronomy, Academic Press, 2015, pp. 139–159, https://doi.org/10.1016/bs.agron.2015.02.002.
[9] R.N. Clark, T.V.V. King, M. Klejwa, G.A. Swayze, N. Vergo, High spectral resolution reflectance spectroscopy of minerals, J. Geophys. Res. Solid Earth 95 (1990) 12653–12680, https://doi.org/10.1029/JB095iB08p12653.
[10] A. Gholizadeh, L. Borůvka, M.M. Saberioon, J. Kozák, R. Vašát, K. Němeček, Comparing different data preprocessing methods for monitoring soil heavy metals based on soil spectral features, Soil Water Res. 10 (2015) 218–227, https://doi.org/10.17221/113/2015-SWR.
[11] B. Stenberg, R.A. Viscarra Rossel, A.M. Mouazen, J. Wetterlind, Visible and near infrared spectroscopy in soil science, in: Advances in Agronomy, Elsevier, 2010, pp. 163–215, https://doi.org/10.1016/S0065-2113(10)07005-7.
[12] R.A.V. Rossel, J. Bouma, Soil sensing: a new paradigm for agriculture, Agric. Syst. 148 (2016) 71–74, https://doi.org/10.1016/j.agsy.2016.07.001.
[13] I. Barra, S.M. Haefele, R. Sakrabani, F. Kebede, Soil spectroscopy with the use of chemometrics, machine learning and pre-processing techniques in soil diagnosis: recent advances-A review, Trac. Trends Anal. Chem. 135 (2021) 116166, https://doi.org/10.1016/j.trac.2020.116166.
[14] V. Sharma, R. Chauhan, R. Kumar, Spectral characteristics of organic soil matter: a comprehensive review, Microchem. J. 171 (2021) 106836, https://doi.org/10.1016/j.microc.2021.106836.
[15] M.A.N. Coutinho, F. de O. Alari, M.M.C. Ferreira, L.R. do Amaral, Influence of soil sample preparation on the quantification of NPK content via spectroscopy, Geoderma 338 (2019) 401–409, https://doi.org/10.1016/j.geoderma.2018.12.021.
[16] M. Gandariasbeitia, G. Besga, I. Albizu, S. Larregla, S. Mendarte, Prediction of chemical and biological variables of soil in grazing areas with visible- and near-infrared spectroscopy, Geoderma 305 (2017) 228–235, https://doi.org/10.1016/j.geoderma.2017.05.045.
[17] J.M. Soriano-Disla, L.J. Janik, R.A.V. Rossel, L.M. Macdonald, M.J. McLaughlin, The performance of visible, near-, and mid-infrared reflectance spectroscopy for prediction of soil physical, chemical, and biological properties, Appl. Spectrosc. Rev. 49 (2014) 139–186, https://doi.org/10.1080/05704928.2013.811081.
[18] X. Jia, D. O'Connor, Z. Shi, D. Hou, VIRS based detection in combination with machine learning for mapping soil pollution, Environ. Pollut. 268 (2021) 115845, https://doi.org/10.1016/j.envpol.2020.115845.
[19] A.J. Margenot, F.J. Calderón, S.J. Parikh, Limitations and potential of spectral subtractions in fourier-transform infrared spectroscopy of soil samples, Soil Science Soc of Amer J 80 (2016) 10–26, https://doi.org/10.2136/sssaj2015.06.0228.
[20] J.J. Roberts, D. Cozzolino, Wet or dry? The effect of sample characteristics on the determination of soil properties by near infrared spectroscopy, TrAC, Trends Anal. Chem. 83 (2016) 25–30, https://doi.org/10.1016/j.trac.2016.08.002.
[21] C. Nduwamungu, N. Ziadi, G.F. Tremblay, L.-É. Parent, Near-infrared reflectance spectroscopy prediction of soil properties: effects of sample cups and preparation, Soil Sci. Soc. Am. J. 73 (2009) 1896–1903, https://doi.org/10.2136/sssaj2008.0213.
[22] Y. Hong, M.A. Munnaf, A. Guerrero, S. Chen, Y. Liu, Z. Shi, A.M. Mouazen, Fusion of visible-to-near-infrared and mid-infrared spectroscopy to estimate soil organic carbon, Soil Tillage Res. 217 (2022) 105284, https://doi.org/10.1016/j.still.2021.105284.
[23] A.M. Mouazen, J. De Baerdemaeker, H. Ramon, Towards development of on-line soil moisture content sensor using a fibre-type NIR spectrophotometer, Soil Tillage Res. 80 (2005) 171–183, https://doi.org/10.1016/j.still.2004.03.022.
[24] E. Ben Dor, C. Ong, I.C. Lau, Reflectance measurements of soils in the laboratory: standards and protocols, Geoderma 245–246 (2015) 112–124, https://doi.org/10.1016/j.geoderma.2015.01.002.
[25] N.M. Dhawale, V.I. Adamchuk, S.O. Prasher, R.A. Viscarra Rossel, A.A. Ismail, J. Kaur, Proximal soil sensing of soil texture and organic matter with a prototype portable mid-infrared spectrometer, Eur. J. Soil Sci. 66 (2015) 661–669, https://doi.org/10.1111/ejss.12265.
[26] B. Minasny, A.B. McBratney, V. Bellon-Maurel, J.-M. Roger, A. Gobrecht, L. Ferrand, S. Joalland, Removing the effect of soil moisture from NIR diffuse reflectance spectra for the prediction of soil organic carbon, Geoderma 167–168 (2011) 118–124, https://doi.org/10.1016/j.geoderma.2011.09.008.
[27] W. Ji, R.A. Viscarra Rossel, Z. Shi, Accounting for the effects of water and the environment on proximally sensed vis–NIR soil spectra and their calibrations, Eur. J. Soil Sci. 66 (2015) 555–565, https://doi.org/10.1111/ejss.12239.
[28] Q. Guan, R. Zhao, F. Wang, N. Pan, L. Yang, N. Song, C. Xu, J. Lin, Prediction of heavy metals in soils of an arid area based on multi-spectral data, J. Environ. Manag. 243 (2019) 137–143, https://doi.org/10.1016/j.jenvman.2019.04.109.
[29] E. Ben-Dor, S. Chabrillat, J.A.M. Demattê, G.R. Taylor, J. Hill, M.L. Whiting, S. Sommer, Using Imaging Spectroscopy to study soil properties, Rem. Sens. Environ. 113 (2009) S38–S55, https://doi.org/10.1016/j.rse.2008.09.019.
[30] C. Hbirkou, S. Pätzold, A.-K. Mahlein, G. Welp, Airborne hyperspectral imaging of spatial soil organic carbon heterogeneity at the field-scale, Geoderma 175–176 (2012) 21–28, https://doi.org/10.1016/j.geoderma.2012.01.017.
[31] G. Vivone, Multispectral and hyperspectral image fusion in remote sensing: a survey, Inf. Fusion 89 (2023) 405–417, https://doi.org/10.1016/j.inffus.2022.08.032.
[32] J.L. Campbell, J.A. Maxwell, S.M. Andrushenko, S.M. Taylor, B.N. Jones, W. Brown-Bury, A GUPIX-based approach to interpreting the PIXE-plus-XRF spectra from the Mars Exploration Rovers: I. Homogeneous standards, Nucl. Instrum. Methods Phys. Res. Sect. B Beam Interact. Mater. Atoms 269 (2011) 57–68, https://doi.org/10.1016/j.nimb.2010.10.004.
[33] P. Shi, F. Castaldi, B. van Wesemael, K. Van Oost, Large-Scale, High-resolution map** of soil aggregate stability in croplands using APEX hyperspectral imagery, Rem. Sens. 12 (2020) 666, https://doi.org/10.3390/rs12040666.
[34] N.J.P. Anne, A.H. Abd-Elrahman, D.B. Lewis, N.A. Hewitt, Modeling soil parameters using hyperspectral image reflectance in subtropical coastal wetlands, Int. J. Appl. Earth Obs. Geoinf. 33 (2014) 47–56, https://doi.org/10.1016/j.jag.2014.04.007.
[35] K.M. Yusof, S. Isaak, N.C.A. Rashid, N.H. Ngajikin, NPK detection spectroscopy ON NON-agriculture soil, Jurnal Teknologi 78 (2016), https://doi.org/10.11113/jt.v78.8382.
[36] Seema, A.K. Ghosh, B.S. Das, N. Reddy, Application of VIS-NIR spectroscopy for estimation of soil organic carbon using different spectral preprocessing techniques and multivariate methods in the middle Indo-Gangetic plains of India, Geoderma Regional 23 (2020) e00349, https://doi.org/10.1016/j.geodrs.2020.e00349.
[37] Y.-H. Yun, H.-D. Li, B.-C. Deng, D.-S. Cao, An overview of variable selection methods in multivariate analysis of near-infrared spectra, TrAC, Trends Anal. Chem. 113 (2019) 102–115, https://doi.org/10.1016/j.trac.2019.01.018.
[38] S.R. Delwiche, J.B. Reeves, A graphical method to evaluate spectral preprocessing in multivariate regression calibrations: example with Savitzky-Golay filters and partial least squares regression, Appl. Spectrosc. 64 (2010) 73–82, https://doi.org/10.1366/000370210790572007.
[39] H. Yang, B. Kuang, A. Mouazen, Selection of spectral preprocessing methods for soil texture classification, Adv. Mater. Res. 181–182 (2011) 416–421. https://doi.org/10.4028/www.scientific.net/AMR.181-182.416.
[40] D.R. Newman, J.M.H. Cockburn, L. Drăguț, J.B. Lindsay, Local scale optimization of geomorphometric land surface parameters using scale-standardized Gaussian scale-space, Comput. Geosci. 165 (2022) 105144, https://doi.org/10.1016/j.cageo.2022.105144.
[41] G. Tranter, B. Minasny, A.B. McBratney, R.A.V. Rossel, B.W. Murphy, Comparing spectral soil inference systems and mid-infrared spectroscopic predictions of soil moisture retention, Soil Sci. Soc. Am. J. 72 (2008) 1394–1400, https://doi.org/10.2136/sssaj2007.0188.

[42] F. Berg, Å. Rinnan, Chapter 5. Calibration transfer methods, in: Infrared Spectroscopy for Food Quality Analysis and Control, 2009, pp. 105–118, https://doi.org/10.1016/B978-0-12-374136-3.00005-5.
[43] C. Huazhou, J. University, Guangzhou, Guangdong, China, Combination optimization of multiple scatter correction and Savitzky-Golay smoothing modes applied to the near infrared spectroscopy analysis of soil organic matter, Comput. Appl. Chem. 28 (2011) 518–522, https://doi.org/10.1007/s12274-011-0112-2.
[44] M. Guolin, D. Jianli, Z. Zipeng, Soil organic matter content estimation based on soil covariate and VIS-NIR spectroscopy, laser optoelectron, Prog. 57 (2020) 192801, https://doi.org/10.3788/LOP57.192801.
[45] Z. Sun, Y. Zhang, J. Li, W. Zhou, Spectroscopic determination of soil organic carbon and total nitrogen content in pasture soils, Commun. Soil Sci. Plant Anal. 45 (2014) 1037–1048, https://doi.org/10.1080/00103624.2014.883628.
[46] H. Yang, Y. Qian, F. Yang, J. Li, W. Ju, Using wavelet Transform of hyperspectral reflectance data for extracting spectral features of soil organic carbon and nitrogen, Soil Sci. 177 (2012) 674–681, https://doi.org/10.1097/SS.0b013e3182792bcc.
[47] A.K. Patel, J.K. Ghosh, Soil fertility status assessment using hyperspectral remote sensing, in: C.M. Neale, A. Maltese (Eds.), Remote Sensing for Agriculture, Ecosystems, and Hydrology XXI, SPIE, 2019, p. 14, https://doi.org/10.1117/12.2533115. Strasbourg, France.
[48] T.W. Tong Jiajun, Improvement of NIR model by fractional order Savitzky-Golay derivation (FOSGD) coupled with wavelength selection, Chemometr. Intell. Lab. Syst. 143 (2015). https://www.zhangqiaokeyan.com/journal-foreign-detail/070401270731.html (accessed May 6, 2024).
[49] G. Smith, Step away from stepwise, J. Big Data 5 (2018) 32, https://doi.org/10.1186/s40537-018-0143-6.
[50] G. Tang, Y. Huang, K. Tian, X. Song, H. Yan, J. Hu, Y. Xiong, S. Min, A new spectral variable selection pattern using competitive adaptive reweighted sampling combined with successive projections algorithm, Analyst 139 (2014) 4894–4902, https://doi.org/10.1039/C4AN00837E.
[51] D. Wu, X. Chen, X. Zhu, X. Guan, G. Wu, Uninformative variable elimination for improvement of successive projections algorithm on spectral multivariable selection with different calibration algorithms for the rapid and non-destructive determination of protein content in dried laver, Anal. Methods 3 (2011) 1790–1796, https://doi.org/10.1039/C1AY05075C.
[52] V. Centner, 3.21 - multivariate approaches: UVE-PLS, in: S.D. Brown, R. Tauler, B. Walczak (Eds.), Comprehensive Chemometrics, Elsevier, Oxford, 2009, pp. 609–618, https://doi.org/10.1016/B978-044452701-1.00031-4.
[53] M. Reza Keyvanpour, M.B. Shirzad, Chapter 14 - machine learning techniques for agricultural image recognition, in: M.A. Khan, R. Khan, M.A. Ansari (Eds.), Application of Machine Learning in Agriculture, Academic Press, 2022, pp. 283–305, https://doi.org/10.1016/B978-0-323-90550-3.00011-4.
[54] B. Yegnanarayana, Artificial Neural Networks, PHI Learning Pvt. Ltd., 2009.
[55] K. Han, Y. Wang, A review of artificial neural network techniques for environmental issues prediction, J. Therm. Anal. Calorim. 145 (2021) 2191–2207, https://doi.org/10.1007/s10973-021-10748-9.
[56] J. Heaton, Ian Goodfellow, Yoshua Bengio, Aaron Courville, Deep learning, Genet. Program. Evolvable Mach. 19 (2018) 305–307, https://doi.org/10.1007/s10710-017-9314-z.
[57] F.E. Harrell, Regression Modeling Strategies: with Applications to Linear Models, Logistic and Ordinal Regression, and Survival Analysis, Springer International Publishing, Cham, 2015, https://doi.org/10.1007/978-3-319-19425-7.
[58] Y. He, M. Huang, A. García, A. Hernández, H. Song, Prediction of soil macronutrients content using near-infrared spectroscopy, Comput. Electron. Agric. 58 (2007) 144–153, https://doi.org/10.1016/j.compag.2007.03.011.
[59] Y. Zhang, M. Li, L. Zheng, Y. Zhao, X. Pei, Soil nitrogen content forecasting based on real-time NIR spectroscopy, Comput. Electron. Agric. 124 (2016) 29–36, https://doi.org/10.1016/j.compag.2016.03.016.
[60] J.B. Carra, M. Fabris, J. Dieckow, O.R. Brito, P.R.S. Vendrame, L. Macedo Dos Santos Tonial, Near-infrared spectroscopy coupled with chemometrics tools: a rapid and non-destructive alternative on soil evaluation, Commun. Soil Sci. Plant Anal. 50 (2019) 421–434, https://doi.org/10.1080/00103624.2019.1566465.
[61] M. Kodaira, S. Shibusawa, Using a mobile real-time soil visible-near infrared sensor for high resolution soil property mapping, Geoderma 199 (2013) 64–79, https://doi.org/10.1016/j.geoderma.2012.09.007.
[62] B.H. Kusumo, C.B. Hedley, M.J. Hedley, A. Hueni, M.P. Tuohy, G.C. Arnold, The use of diffuse reflectance spectroscopy for in situ carbon and nitrogen analysis of pastoral soils, Aust. J. Soil Res. 46 (2008) 623–635, https://doi.org/10.1071/SR08118.
[63] J. Wetterlind, B. Stenberg, M. Söderström, Increased sample point density in farm soil mapping by local calibration of visible and near infrared prediction models, Geoderma 156 (2010) 152–160, https://doi.org/10.1016/j.geoderma.2010.02.012.
[64] D.B. Madhavan, M. Kitching, D.S. Mendham, C.J. Weston, T.G. Baker, Mid-infrared spectroscopy for rapid assessment of soil properties after land use change from pastures to Eucalyptus globulus plantations, J. Environ. Manag. 175 (2016) 67–75, https://doi.org/10.1016/j.jenvman.2016.03.032.
[65] G.W. McCarty, J.B. Reeves, Comparison of near infrared and mid infrared diffuse reflectance spectroscopy for field-scale measurement of soil fertility parameters, Soil Sci. 171 (2006) 94–102, https://doi.org/10.1097/01.ss.0000187377.84391.54.
[66] U.J. dos Santos, J.A. de Melo Dematte, R.S.C. Menezes, A.C. Dotto, C.C. B. Guimarães, B.J.R. Alves, D.C. Primo, E.V. Sampaio, Predicting carbon and nitrogen by visible near-infrared (Vis-NIR) and mid-infrared (MIR) spectroscopy in soils of Northeast Brazil, Geoderma Regional 23 (2020) e00333, https://doi.org/10.1016/j.geodrs.2020.e00333.
[67] S. Alomar, S.A. Mireei, A. Hemmat, A.A. Masoumi, H. Khademi, Comparison of Vis/SWNIR and NIR spectrometers combined with different multivariate techniques for estimating soil fertility parameters of calcareous topsoil in an arid climate, Biosyst. Eng. 201 (2021) 50–66, https://doi.org/10.1016/j.biosystemseng.2020.11.007.
[68] Q. Wu, Y. Yang, Z. Xu, Y. Jin, Y. Guo, C. Lao, Applying local neural network and visible/near-infrared spectroscopy to estimating available nitrogen, phosphorus and potassium in soil, Guang Pu Xue Yu Guang Pu Fen Xi = Guang Pu (2014). https://www.semanticscholar.org/paper/%5BApplying-local-neural-network-and-visible-to-and-Wu-Yang/e241daacf28f99fea7ae8e32e3bf0567e76a9d70. (Accessed 1 July 2024).
[69] N.K. Wijewardane, Y. Ge, S. Wills, Z. Libohova, Predicting physical and chemical properties of US soils with a mid-infrared reflectance spectral library, Soil Sci. Soc. Am. J. 82 (2018) 722–731, https://doi.org/10.2136/sssaj2017.10.0361.
[70] J. Wenjun, S. Zhou, H. Jingyi, L. Shuo, In situ measurement of some soil properties in paddy soil using visible and near-infrared spectroscopy, PLoS One 9 (2014) e105708, https://doi.org/10.1371/journal.pone.0105708.
[71] R. Reda, T. Saffaj, B. Ilham, O. Saidi, K. Issam, L. Brahim, E.M. El Hadrami, A comparative study between a new method and other machine learning algorithms for soil organic carbon and total nitrogen prediction using near infrared spectroscopy, Chemometr. Intell. Lab. Syst. 195 (2019) 103873, https://doi.org/10.1016/j.chemolab.2019.103873.
[72] B. Kuang, A.M. Mouazen, Calibration of visible and near infrared spectroscopy for soil analysis at the field scale on three European farms, Eur. J. Soil Sci. 62 (2011) 629–636, https://doi.org/10.1111/j.1365-2389.2011.01358.x.
[73] D. Cozzolino, W.U. Cynkar, R.G. Dambergs, N. Shah, P. Smith, In situ measurement of soil chemical composition by near-infrared spectroscopy: a tool toward sustainable vineyard management, Commun. Soil Sci. Plant Anal. 44 (2013) 1610–1619, https://doi.org/10.1080/00103624.2013.768263.
[74] B.G. Barthès, E. Kouakoua, M. Clairotte, J. Lallemand, L. Chapuis-Lardy, M. Rabenarivo, S. Roussel, Performance comparison between a miniaturized and a conventional near infrared reflectance (NIR) spectrometer for characterizing soil carbon and nitrogen, Geoderma 338 (2019) 422–429, https://doi.org/10.1016/j.geoderma.2018.12.031.
[75] M. Schirrmann, R. Gebbers, E. Kramer, Performance of automated near-infrared reflectance spectrometry for continuous in situ mapping of soil fertility at field scale, Vadose Zone J. 12 (2013), https://doi.org/10.2136/vzj2012.0199 vzj2012.0199.
[76] Y. Liu, Q. Jiang, T. Shi, T. Fei, J. Wang, G. Liu, Y. Chen, Prediction of total nitrogen in cropland soil at different levels of soil moisture with Vis/NIR spectroscopy, Acta Agric. Scand. Sect. B Soil Plant Sci 64 (2014) 267–281, https://doi.org/10.1080/09064710.2014.906644.
[77] I. Djuuna, L. Abbott, C. Russell, Determination and prediction of some soil properties using partial least square (PLS) calibration and mid-infra red (MIR) spectroscopy analysis, J. Tropical Soil. 16 (2013) 93–98, https://doi.org/10.5400/jts.2011.v16i2.93-98.
[78] H. Sepahvand, R. Mirzaeitalarposhti, K. Beiranvand, M. Feizian, T. Müeller, Prediction of soil carbon levels in calcareous soils of Iran by mid-infrared reflectance spectroscopy, Environ. Pollut. Bioavailabil. 31 (2019) 9–17, https://doi.org/10.1080/09542299.2018.1549961.
[79] H.T. Xie, X.M. Yang, C.F. Drury, J.Y. Yang, X.D. Zhang, Predicting soil organic carbon and total nitrogen using mid- and near-infrared spectra for Brookston clay loam soil in Southwestern Ontario, Canada, Can. J. Soil Sci. 91 (2011) 53–63, https://doi.org/10.4141/cjss10029.
[80] J.M. Soriano-Disla, L.J. Janik, D.J. Allen, M.J. McLaughlin, Evaluation of the performance of portable visible-infrared instruments for the prediction of soil properties, Biosyst. Eng. 161 (2017) 24–36, https://doi.org/10.1016/j.biosystemseng.2017.06.017.
[81] J.-M. Johnson, E. Vandamme, K. Senthilkumar, A. Sila, K.D. Shepherd, K. Saito, Near-infrared, mid-infrared or combined diffuse reflectance spectroscopy for assessing soil fertility in rice fields in sub-Saharan Africa, Geoderma 354 (2019) 113840, https://doi.org/10.1016/j.geoderma.2019.06.043.
[82] J.-P. Gras, B.G. Barthès, B. Mahaut, S. Trupin, Best practices for obtaining and processing field visible and near infrared (VNIR) spectra of topsoils, Geoderma (2014) 214–215, https://doi.org/10.1016/j.geoderma.2013.09.021, 126–134.
[83] E.K. Towett, K.D. Shepherd, A. Sila, E. Aynekulu, G. Cadisch, Mid-infrared and total X-ray fluorescence spectroscopy complementarity for assessment of soil properties, Soil Sci. Soc. Am. J. 79 (2015) 1375–1385, https://doi.org/10.2136/sssaj2014.11.0458.
[84] X. Yu, Q. Liu, Y. Wang, X. Liu, X. Liu, Evaluation of MLSR and PLSR for estimating soil element contents using visible/near-infrared spectroscopy in apple orchards on the Jiaodong peninsula, Catena 137 (2016) 340–349, https://doi.org/10.1016/j.catena.2015.09.024.
[85] R. Martínez-España, A. Bueno-Crespo, J. Soto, L.J. Janik, J.M. Soriano-Disla, Developing an intelligent system for the prediction of soil properties with a portable mid-infrared instrument, Biosyst. Eng. 177 (2019) 101–108, https://doi.org/10.1016/j.biosystemseng.2018.09.013.
[86] S. Xu, Y. Zhao, M. Wang, X. Shi, Comparison of multivariate methods for estimating selected soil properties from intact soil cores of paddy fields by Vis–NIR spectroscopy, Geoderma 310 (2018) 29–43, https://doi.org/10.1016/j.geoderma.2017.09.013.
[87] J. Sanderman, K. Savage, S.R.S. Dangal, Mid-infrared spectroscopy for prediction of soil health indicators in the United States, Soil Sci. Soc. Am. J. 84 (2020) 251–261, https://doi.org/10.1002/saj2.20009.

[88] Q. Wu, Y. Yang, Z. Xu, Y. Jin, Y. Guo, C. Lao, Applying local neural network and visible/near-infrared spectroscopy to estimating available nitrogen, phosphorus and potassium in soil, Guang Pu Xue Yu Guang Pu Fen Xi = Guang Pu 34 (2014) 2102–2105, https://doi.org/10.3964/j.issn.1000-0593(2014)08-2102-04.
[89] W. Ji, V.I. Adamchuk, A. Biswas, N.M. Dhawale, B. Sudarsan, Y. Zhang, R. A. Viscarra Rossel, Z. Shi, Assessment of soil properties *in situ* using a prototype portable MIR spectrometer in two agricultural fields, Biosyst. Eng. 152 (2016) 14–27, https://doi.org/10.1016/j.biosystemseng.2016.06.005.
[90] E.S. Mohamed, A.A.E. Baroudy, T. El-beshbeshy, M. Emam, A.A. Belal, A. Elfadaly, A.A. Aldosari, A.M. Ali, R. Lasaponara, Vis-NIR spectroscopy and satellite landsat-8 OLI data to map soil nutrients in arid conditions: a case study of the northwest coast of Egypt, Rem. Sens. 12 (2020) 3716, https://doi.org/10.3390/rs12223716.
[91] Y. Shao, Y. He, Nitrogen, phosphorus, and potassium prediction in soils, using infrared spectroscopy, Soil Res. 49 (2011) 166–172, https://doi.org/10.1071/SR10098.
[92] T. Paz-Kagan, M. Shachak, E. Zaady, A. Karnieli, A spectral soil quality index (SSQI) for characterizing soil function in areas of changed land use, Geoderma 230–231 (2014) 171–184, https://doi.org/10.1016/j.geoderma.2014.04.003.
[93] R.A. Viscarra Rossel, D.J.J. Walvoort, A.B. McBratney, L.J. Janik, J.O. Skjemstad, Visible, near infrared, mid infrared or combined diffuse reflectance spectroscopy for simultaneous assessment of various soil properties, Geoderma 131 (2006) 59–75, https://doi.org/10.1016/j.geoderma.2005.03.007.
[94] J. Yin, Z. Shi, B. Li, F. Sun, T. Miao, Z. Shi, S. Chen, M. Yang, W. Ji, Prediction of soil properties in a field in typical black soil areas using in situ MIR spectra and its comparison with vis-NIR spectra, Rem. Sens. 15 (2023) 2053, https://doi.org/10.3390/rs15082053.
[95] D. Xu, W. Ma, S. Chen, Q. Jiang, K. He, Z. Shi, Assessment of important soil properties related to Chinese Soil Taxonomy based on vis–NIR reflectance spectroscopy, Comput. Electron. Agric. 144 (2018) 1–8, https://doi.org/10.1016/j.compag.2017.11.029.
[96] X.-Y. Hu, Application of visible/near-infrared spectra in modeling of soil total phosphorus, Pedosphere 23 (2013) 417–421, https://doi.org/10.1016/S1002-0160(13)60034-X.
[97] R. Recena, V.M. Fernández-Cabanás, A. Delgado, Soil fertility assessment by Vis-NIR spectroscopy: predicting soil functioning rather than availability indices, Geoderma 337 (2019) 368–374, https://doi.org/10.1016/j.geoderma.2018.09.049.
[98] S. Xu, Y. Zhao, Y. Wang, Optimizing machine learning models for predicting soil pH and total P in intact soil profiles with visible and near-infrared reflectance (VNIR) spectroscopy, Comput. Electron. Agric. 218 (2024) 108643, https://doi.org/10.1016/j.compag.2024.108643.
[99] M.-C. Marmette, V. Adamchuk, J. Nault, S. Tabatabai, R. Cocciardi, Comparison of the Performance of Two Vis-NIR Spectrometers in the Prediction of Various Soil Properties, 2018.
[100] A.M. Mouazen, M.R. Maleki, J. De Baerdemaeker, H. Ramon, On-line measurement of some selected soil properties using a VIS–NIR sensor, Soil Tillage Res. 93 (2007) 13–27, https://doi.org/10.1016/j.still.2006.03.009.
[101] P.R.S. Vendrame, R.L. Marchão, D. Brunet, T. Becquer, The potential of NIR spectroscopy to predict soil texture and mineralogy in Cerrado Latosols, Eur. J. Soil Sci. 63 (2012) 743–753, https://doi.org/10.1111/j.1365-2389.2012.01483.x.
[102] F. Ma, C.W. Du, J.M. Zhou, Y.Z. Shen, Investigation of soil properties using different techniques of mid-infrared spectroscopy, Eur. J. Soil Sci. 70 (2019) 96–106, https://doi.org/10.1111/ejss.12741.
[103] L.J. Janik, R.H. Merry, J.O. Skjemstad, Can mid infrared diffuse reflectance analysis replace soil extractions? Aust. J. Exp. Agric. 38 (1998) 681–696, https://doi.org/10.1071/EA97144.
[104] S.R. Araújo, M. Söderström, J. Eriksson, C. Isendahl, P. Stenborg, JoséA. M. Demattê, Determining soil properties in Amazonian Dark Earths by reflectance spectroscopy, Geoderma 237–238 (2015) 308–317, https://doi.org/10.1016/j.geoderma.2014.09.014.
[105] C.D. Christy, Real-time measurement of soil attributes using on-the-go near infrared reflectance spectroscopy, Comput. Electron. Agric. 61 (2008) 10–19, https://doi.org/10.1016/j.compag.2007.02.010.
[106] K.W. Daniel, N.K. Tripathi, K. Honda, Artificial neural network analysis of laboratory and in situ spectra for the estimation of macronutrients in soils of Lop Buri (Thailand), Soil Res. 41 (2003) 47–59, https://doi.org/10.1071/sr02027.
[107] H. Qi, T. Paz-Kagan, A. Karnieli, X. Jin, S. Li, Evaluating calibration methods for predicting soil available nutrients using hyperspectral VNIR data, Soil Tillage Res. 175 (2018) 267–275, https://doi.org/10.1016/j.still.2017.09.006.
[108] M.C. Sarathjith, B.S. Das, S.P. Wani, K.L. Sahrawat, A. Gupta, Comparison of data mining approaches for estimating soil nutrient contents using diffuse reflectance spectroscopy, Curr. Sci. 110 (2015) 1031, https://doi.org/10.18520/cs/v110/i6/1031-1037.
[109] F.S. Terra, J.A.M. Demattê, R.A. Viscarra Rossel, Spectral libraries for quantitative analyses of tropical Brazilian soils: comparing vis–NIR and mid-IR reflectance data, Geoderma 255–256 (2015) 81–93, https://doi.org/10.1016/j.geoderma.2015.04.017.
[110] A. Gholizade, M.A.M. Soom, M.M. Saberioon, L. Borůvka, Visible and near infrared reflectance spectroscopy to determine chemical properties of paddy soils, J. Food Agric. Environ. 11 (2013) 859–866.
[111] K. Metzger, F. Liebisch, J.M. Herrera, T. Guillaume, F. Walder, L. Bragazza, The use of visible and near-infrared spectroscopy for in-situ characterization of agricultural soil fertility: a proposition of best practice by comparing scanning positions and spectrometers, Soil Use Manag. 40 (2024) e12952, https://doi.org/10.1111/sum.12952.
[112] X.-Y. Li, P.-P. Fan, Y. Liu, G.-L. Hou, Q. Wang, M.-R. Lv, Prediction results of different modeling methods in soil nutrient concentrations based on spectral technology, J. Appl. Spectrosc. 86 (2019) 765–770, https://doi.org/10.1007/s10812-019-00891-5.
[113] D. Cozzolino, A. Morón, The potential of near-infrared reflectance spectroscopy to analyse soil chemical and physical characteristics, J. Agric. Sci. 140 (2003) 65–71, https://doi.org/10.1017/S0021859602002836.
[114] C.-W. Chang, D.A. Laird, M.J. Mausbach Jr., C.R. Hurburgh, Near-infrared reflectance spectroscopy - principal components regression analyses of soil properties, Soil Sci. Soc. Am. J. 65 (2001) 480–490, https://doi.org/10.2136/sssaj2001.652480x.
[115] A.M. Mouazen, B. Kuang, J. De Baerdemaeker, H. Ramon, Comparison among principal component, partial least squares and back propagation neural network analyses for accuracy of measurement of selected soil properties with visible and near infrared spectroscopy, Geoderma 158 (2010) 23–31, https://doi.org/10.1016/j.geoderma.2010.03.001.
[116] L. Xuemei, L. Jianshe, Using short wave visible–near infrared reflectance spectroscopy to predict soil properties and content, Spectrosc. Lett. 47 (2014) 729–739, https://doi.org/10.1080/00387010.2013.840315.
[117] F. Albinet, Y. Peng, T. Eguchi, E. Smolders, G. Dercon, Prediction of exchangeable potassium in soil through mid-infrared spectroscopy and deep learning: from prediction to explainability, Artificial Intellig. Agricult. 6 (2022) 230–241, https://doi.org/10.1016/j.aiia.2022.10.001.
[118] S. Chabrillat, E. Ben-Dor, J. Cierniewski, C. Gomez, T. Schmid, B. Van Wesemael, Imaging spectroscopy for soil mapping and monitoring, Surv. Geophys. 40 (2019) 361–399, https://doi.org/10.1007/s10712-019-09524-0.
[119] A.K. Salman, S.E. Aldulaimy, H.J. Mohammed, Y.M. Abed, Performance of soil moisture sensors in gypsiferous and salt-affected soils, Biosyst. Eng. 209 (2021) 200–209, https://doi.org/10.1016/j.biosystemseng.2021.07.006.
[120] K.A. Sudduth, J.W. Hummel, Portable, near-infrared spectrophotometer for rapid soil analysis, Transact. ASAE 36 (1993) 185–193, https://doi.org/10.13031/2013.28329.
[121] T. Angelopoulou, A. Balafoutis, G. Zalidis, D. Bochtis, From laboratory to proximal sensing spectroscopy for soil organic carbon estimation—a review, Sustainability 12 (2020) 443, https://doi.org/10.3390/su12020443.
[122] R.A. Viscarra Rossel, S.R. Cattle, A. Ortega, Y. Fouad, In situ measurements of soil colour, mineral composition and clay content by vis–NIR spectroscopy, Geoderma 150 (2009) 253–266, https://doi.org/10.1016/j.geoderma.2009.01.025.
[123] J.-M. Roger, F. Chauchard, V. Bellon-Maurel, EPO–PLS external parameter orthogonalisation of PLS application to temperature-independent measurement of sugar content of intact fruits, Chemometr. Intell. Lab. Syst. 66 (2003) 191–204, https://doi.org/10.1016/S0169-7439(03)00051-0.
[124] N.A. Woody, R.N. Feudale, A.J. Myles, S.D. Brown, Transfer of multivariate calibrations between four near-infrared spectrometers using orthogonal signal correction, Anal. Chem. 76 (2004) 2595–2600, https://doi.org/10.1021/ac035382g.
[125] W. de S. Mendes, J.A.M. Demattê, B.R. Bonfatti, M.E.B. Resende, L.R. Campos, A. C.S. da Costa, A novel framework to estimate soil mineralogy using soil spectroscopy, Appl. Geochem. 127 (2021) 104909, https://doi.org/10.1016/j.apgeochem.2021.104909.
[126] J. Madejová, FTIR techniques in clay mineral studies, Vib. Spectrosc. 31 (2003) 1–10, https://doi.org/10.1016/S0924-2031(02)00065-6.
[127] P. Fan, X. Li, H. Qiu, G.-L. Hou, Spectral analysis of total phosphorus in soils based on its diagnostic reflectance spectra, Result. Chem. 3 (2021) 100145, https://doi.org/10.1016/j.rechem.2021.100145.
[128] J.D. Russell, H.W. Van der Marel, H. Beutelspacher, Atlas of Infrared Spectroscopy of Clay Minerals and Their Admixtures, vol. 12, Elsevier, Amsterdam, 1977, pp. 279–280, https://doi.org/10.1180/claymin.1977.012.3.11, 1976. viii + 396 pp. £34.35, Clay Miner.
[129] M.J. Wilson (Ed.), Clay Mineralogy: Spectroscopic and Chemical Determinative Methods, Springer Netherlands, Dordrecht, 1994, https://doi.org/10.1007/978-94-011-0727-3.
[130] E.R. Stoner, M. Baumgardner, Physiochemical, site, and bidirectional reflectance factor characteristics of uniformly moist soils [Brazil, Spain and the United States of America], https://www.semanticscholar.org/paper/Physiochemical%2C-site%2C-and-bidirectional-reflectance-Stoner-Baumgardner/44888e7ad24df66f9515b4cf9e3b6319592d55d5, 1980. (Accessed 30 August 2024).
[131] G. Hunt, Visible and near-infrared spectra of minerals and rocks : I silicate minerals. https://www.semanticscholar.org/paper/Visible-and-near-infrared-spectra-of-minerals-and-%3A-Hunt/d6d90ca316b10199236f28926df4f29b8e01aa74, 1970. (Accessed 30 August 2024).
[132] T. Coleman, O. Montgomery, Soil moisture, organic matter, and iron content effect on the spectral characteristics of selected vertisols and alfisols in Alabama, Photogramm. Eng. Rem. Sens. (1987). https://www.semanticscholar.org/paper/Soil-moisture%2C-organic-matter%2C-and-iron-content-on-Coleman-Montgomery/a8ba93dcf5af32583c7b8e9d509cfad0cb451137. (Accessed 30 August 2024).
[133] R. Mirzaeitalarposhti, M.S. Demyan, F. Rasche, G. Cadisch, T. Müller, Overcoming carbonate interference on labile soil organic matter peaks for midDRIFTS analysis, Soil Biol. Biochem. 99 (2016) 150–157, https://doi.org/10.1016/j.soilbio.2016.05.010.
[134] T. He, J. Wang, Z. Lin, Y. Cheng, Spectral features of soil organic matter, Geo Spatial Inf. Sci. 12 (2009) 33–40, https://doi.org/10.1007/s11806-009-0160-x.
[135] E. Ben-Dor, Y. Inbar, Y. Chen, The reflectance spectra of organic matter in the visible near-infrared and short wave infrared region (400–2500 nm) during a

controlled decomposition process, Rem. Sens. Environ. 61 (1997) 1–15, https://doi.org/10.1016/S0034-4257(96)00120-4.
[136] H. Liu, W. Yu, X. Zhang, Q. Ma, H. Zhou, Z. Jiang, Study on the main influencing factors of black soil spectral characteristics, Spectrosc. Spectr. Anal. 29 (2009) 3019–3022, https://doi.org/10.3964/j.issn.1000-0593(2009)11-3019-04.
[137] A. Mihoub, M.D. Bouhoun, A. Naeem, M.L. Saker, Low-molecular weight organic acids improve plant availability of phosphorus in different textured calcareous soils, Arch. Agron Soil Sci. 63 (2017) 1023–1034, https://doi.org/10.1080/03650340.2016.1249477.
[138] K. Jindo, Y. Audette, F.L. Olivares, L.P. Canellas, D.S. Smith, R.P. Voroney, Biotic and abiotic effects of soil organic matter on the phytoavailable phosphorus in soils: a review, Chem. Biol. Technol. Agric. 10 (2023) 29, https://doi.org/10.1186/s40538-023-00401-y.
[139] S. Taghdis, M.H. Farpoor, M. Mahmoodabadi, Pedological assessments along an arid and semi-arid transect using soil spectral behavior analysis, Catena 214 (2022) 106288, https://doi.org/10.1016/j.catena.2022.106288.
[140] W. Ng, B. Minasny, W. de S. Mendes, J.A.M. Demattê, The influence of training sample size on the accuracy of deep learning models for the prediction of soil properties with near-infrared spectroscopy data, SOIL 6 (2020) 565–578, https://doi.org/10.5194/soil-6-565-2020.
[141] Y. Liu, Y. Liu, Y. Chen, Y. Zhang, T. Shi, J. Wang, Y. Hong, T. Fei, Y. Zhang, The influence of spectral pretreatment on the selection of representative calibration samples for soil organic matter estimation using vis-NIR reflectance spectroscopy, Rem. Sens. 11 (2019) 450, https://doi.org/10.3390/rs11040450.
[142] B. Stenberg, R.A. Viscarra Rossel, A.M. Mouazen, J. Wetterlind, Chapter five - visible and near infrared spectroscopy in soil science, in: D.L. Sparks (Ed.), Advances in Agronomy, Academic Press, 2010, pp. 163–215, https://doi.org/10.1016/S0065-2113(10)07005-7.
[143] E. Ben-Dor, S. Chabrillat, J.A.M. Demattê, G.R. Taylor, J. Hill, M.L. Whiting, S. Sommer, Using Imaging Spectroscopy to study soil properties, Rem. Sens. Environ. 113 (2009) S38–S55, https://doi.org/10.1016/j.rse.2008.09.019.
[144] U. Kamath, J. Liu, Introduction to interpretability and explainability, in: U. Kamath, J. Liu (Eds.), Explainable Artificial Intelligence: an Introduction to Interpretable Machine Learning, Springer International Publishing, Cham, 2021, pp. 1–26, https://doi.org/10.1007/978-3-030-83356-5_1.
[145] W. Fu, W.S. Hopkins, Applying machine learning to vibrational spectroscopy, J. Phys. Chem. A 122 (2018) 167–171, https://doi.org/10.1021/acs.jpca.7b10303.
[146] J. Friedman, T. Hastie, R. Tibshirani, Regularization paths for generalized linear models via coordinate descent, J. Stat. Software 33 (2010) 1–22.
[147] T. Hastie, R. Tibshirani, J. Friedman, The Elements of Statistical Learning, Springer, New York, NY, 2009, https://doi.org/10.1007/978-0-387-84858-7.
[148] L.-N. Li, X.-F. Liu, F. Yang, W.-M. Xu, J.-Y. Wang, R. Shu, A review of artificial neural network based chemometrics applied in laser-induced breakdown spectroscopy analysis, Spectrochim. Acta B Atom Spectrosc. 180 (2021) 106183, https://doi.org/10.1016/j.sab.2021.106183.
[149] Y. Xiong, C. McCarthy, J. Humpal, C. Percy, Near-infrared spectroscopy and deep neural networks for early common root rot detection in wheat from multi-season trials, Agron. J. 116 (2024) 2370–2390, https://doi.org/10.1002/agj2.21648.
[150] S. Liu, H. Shen, S. Chen, X. Zhao, A. Biswas, X. Jia, Z. Shi, J. Fang, Estimating forest soil organic carbon content using vis-NIR spectroscopy: implications for large-scale soil carbon spectroscopic assessment, Geoderma 348 (2019) 37–44, https://doi.org/10.1016/j.geoderma.2019.04.003.
[151] F. Gogé, R. Joffre, C. Jolivet, I. Ross, L. Ranjard, Optimization criteria in sample selection step of local regression for quantitative analysis of large soil NIRS database, Chemometr. Intell. Lab. Syst. 110 (2012) 168–176, https://doi.org/10.1016/j.chemolab.2011.11.003.