Article

Application of a Fusion Attention Mechanism-Based Model Combining Bidirectional Gated Recurrent Units and Recurrent Neural Networks in Soil Nutrient Content Estimation

Huan Wang, Lixin Zhang * and **Jiawei Zhao**

School of Mechanical and Electrical Engineering, Shihezi University, Shihezi 832003, China; WangH_Cn@126.com (H.W.); zhaojiawei_cn@126.com (J.Z.)
* Correspondence: lixin_zhang_cn@126.com

Abstract: Accurately estimating soil nutrient content, including soil organic matter (OM), nitrogen (N), phosphorus (P), and potassium (K) levels, is crucial for optimizing agricultural practices and ensuring sustainable crop production. This paper proposes a model based on a fusion attention mechanism that combines bidirectional gated recurrent units (BiGRU) and recurrent neural networks (RNN) to estimate soil nutrient content. The proposed model integrates the fused attention mechanism with BiGRU and RNN to enhance the accuracy and effectiveness of soil nutrient prediction. The fused attention mechanism captures key features in the input data, while the BiGRU architecture captures both forward and backward contextual information, enabling the model to capture long-term dependencies in the data. The results demonstrate that the proposed Att-BiGRU-RNN model outperforms other constructed models, exhibiting a higher prediction accuracy and robustness. The model shows good estimation capabilities for soil OM, N, P, and K with estimation accuracies (R^2) of 0.959, 0.907, 0.921, and 0.914, respectively. The application of this model in soil nutrient estimation has the potential to optimize fertilizer management, enhance soil fertility, and ultimately improve crop yield. Further research can explore the applicability of this model in precision agriculture and sustainable soil management practices, benefiting the agricultural sector and contributing to food security and environmental sustainability.

Keywords: soil nutrient; hyperspectral; attention mechanism; deep learning; feature extraction

1. Introduction

In the face of the current global food crisis, the urgency of implementing precision agriculture practices to enhance agricultural operations is increasing [1,2]. The growing population and evolving dietary preferences are exerting immense pressure on our ability to sustainably produce an adequate food supply [3]. To address these challenges, precision agriculture has emerged as a promising approach that emphasizes the need for location-specific and data-driven agricultural technologies [4,5].

Precision agriculture involves the application of advanced technologies and data analysis to optimize resource allocation, improve crop productivity, and minimize environmental impacts [6,7]. Soil nutrient levels, including soil organic matter (OM), nitrogen (N), phosphorus (P), and potassium (K), play a critical role in supporting crop health and ensuring optimal agricultural productivity [8,9]. These nutrients are vital for various biochemical and physiological processes in plants, affecting their growth, development, and overall yield [10,11]. Soil organic matter provides a source of carbon and nutrients, enhances soil structure and water-holding capacity, and facilitates nutrient cycling [12,13]. Nitrogen is crucial for protein synthesis and overall plant growth [14,15]. Phosphorus is essential for energy transfer and storage, root development, and flowering [16,17]. Potassium regulates water movement, enhances disease resistance, and contributes to overall plant vigor [18,19]. Accurate estimation of soil nutrient levels is a crucial component of precision agriculture

as it provides valuable insights into the availability and distribution of essential nutrients such as soil OM, N, P, and K [20,21].

By precisely assessing nutrient levels in the soil, farmers can adjust their fertilizer management strategies to ensure crops receive the right nutrients at the right time and in the right quantity [22,23]. This targeted approach minimizes nutrient waste, reduces environmental pollution, and maximizes crop yield and quality [24,25]. Additionally, it enables farmers to optimize resource utilization and make informed decisions, thereby enhancing the profitability and long-term sustainability of farms [26]. Accurate estimation of soil nutrient levels is vital for effective nutrient management and sustainable agricultural practices [27,28]. However, traditional soil nutrient analysis methods are often time-consuming, labor-intensive, and costly, making them unsuitable for large-scale monitoring and real-time decision-making [29,30]. These methods involve chemical extraction and analysis to determine nutrient concentrations in the soil [31,32]. While these methods ensure the reliability of soil nutrient detection, they are time-consuming and have stringent personnel and equipment requirements [33]. Moreover, they provide localized information and do not capture the spatial variability of soil nutrient levels in the field [34].

Another challenge is the complexity of the relationship between information obtained from hyperspectral data and soil nutrient content [35,36]. Hyperspectral data contain numerous spectral bands, often with redundant and unrelated information, making it challenging to extract meaningful features for accurate nutrient estimation [37]. Existing models and methods struggle to effectively capture the complex correlations and non-linear relationships present in hyperspectral data, resulting in suboptimal prediction accuracy [38,39]. Furthermore, the importance of different soil nutrients, such as soil OM, N, P, and K, varies depending on the crop and soil type. Developing a comprehensive model that can accurately estimate multiple soil nutrients simultaneously is a challenging task.

In recent years, there has been growing interest in applying machine learning and data-driven methods to soil nutrient estimation [40,41]. Machine-learning methods can directly map the non-linear relationships between input data and corresponding soil property information without the need for building degradation models and computing correlation functions, effectively reducing uncertainty in the data and having advantages in handling high-dimensional data and extracting non-linear patterns in prediction [42]. Machine-learning techniques like support vector regression (SVR) have been widely used for soil nutrient estimation, leveraging their ability to effectively fit large and complex datasets [43]. SVR models are trained based on spectral reflectance values and corresponding soil nutrient measurements to establish regression relationships [44]. However, these models often struggle to capture the non-linear relationships present in hyperspectral data [45].

Other machine-learning techniques such as decision trees [46], principal component analysis (PCA) [47], and random forests [48] have also been used for soil nutrient estimation. These models utilize the power of statistical learning to extract meaningful relationships between spectral features and soil nutrient content [49]. While these methods have shown promising results, they may face challenges in handling high-dimensional hyperspectral data and capturing complex interactions [50]. The reason behind this lies in the fact that convolutional neural networks (CNNs), support vector machines (SVMs), and traditional feedforward neural networks are unable to fully explore the information in soil nutrient data and handle temporal sequences [51]. This limitation results in less accurate predictions from network structures.

To address the issue of underexplored information in soil nutrient data and the inability of convolutional neural networks to handle temporal sequences, this paper proposes a temporal-spatial attention mechanism (TSAM) combined with bidirectional gating recursive units (Bi-GRU) network model for soil nutrient content prediction, with a specific focus on soil OM, N, P, and K. This approach has the potential to revolutionize soil nutrient estimation by providing more accurate and efficient predictions, enabling farmers and land managers to make data-driven decisions and optimize nutrient management practices. By improving the estimation of soil nutrient content, we can enhance agricultural productivity,

reduce input costs, minimize environmental impacts, and contribute to sustainable and responsible agricultural practices.

This research has made significant contributions to the field of agricultural soil nutrient estimation. The main contributions are as follows:

1. Development of New Models: This study proposed and developed five different hybrid models for estimating soil nutrient content, namely FDT-MLR, FDT-SVR, FDT-CNN, BiGRU-RNN, and Att-BiGRU RNN models. Each model combines different techniques and architectures to leverage the strengths of various methods. The FDT-MRR model is included in the model development. The FDT-SVR model combines feature extraction using frequency domain transformation (FDT) with support vector regression (SVR). The FDT-NN model combines FDT with convolutional neural networks (CNN). The BiGRU RNN model includes bidirectional gating recursive units (BiGRU) and RNN. Additionally, the Att-BiGRU-RNN model combines attention mechanisms with BiGRU and RNN. These models are designed to capture the complex relationships and correlations between hyperspectral data and soil nutrient content, thereby improving prediction accuracy.

2. Model Validation and Comparison: Rigorous validation and comparison of the proposed hybrid models were conducted to assess their performance in soil nutrient estimation. Real-world hyperspectral data and corresponding soil nutrient measurements were used to train and test the models. Performance metrics such as the mean absolute error, root mean square error, and determination coefficient were used to evaluate the accuracy and robustness of each model. Through this comparative analysis, the study provides valuable insights into the strengths and limitations of each approach, enabling practitioners to choose the most suitable method based on their specific needs.

The contributions of this research go beyond the development and evaluation of hybrid models. By advancing the field of soil nutrient estimation, this study promotes informed decision-making in precision agriculture and supports sustainable soil management practices. Accurate estimation of soil nutrient content allows farmers to optimize fertilization strategies, minimize nutrient waste, and enhance crop productivity. It also contributes to reducing environmental pollution associated with the excessive use of fertilizers. Furthermore, this research underscores the potential of advanced techniques such as deep learning and attention mechanisms in addressing the challenges of soil nutrient estimation. By demonstrating the effectiveness of these technologies in capturing complex relationships in hyperspectral data, the study lays the foundation for further advancements in this field. In conclusion, this research makes contributions by developing new soil nutrient estimation hybrid models, validating their performance, and providing in-depth insights into their strengths and limitations. These contributions have the potential to strengthen precision agriculture practices, increase crop yields, promote sustainable soil management, benefiting farmers, land managers, and the environment as a whole.

2. Proposed Methods

2.1. Soil Sample Collection

Dingbian County, located in the northwest corner of Shaanxi Province and the westernmost tip of Yulin City, serves as the research area for this study (Figure 1). It is positioned between $107°15'$ and $108°22'$ E longitude and between $36°49'$ and $37°53'$ N latitude, covering a land area of 6920 km^2. The region experiences a temperate semi-arid continental monsoon climate, characterized by distinct seasonal variations. The climate in Dingbian County exhibits several notable features. Springs are windy, summers are dry, autumns are overcast and rainy, and winters are severe. The region experiences ample sunshine, with an annual average of 2856.2 h. The average annual temperature stands at 7.9 °C, and the average annual rainfall is approximately 320.1 mm. The frost-free period lasts around 125 days on average, with an absolute frost-free period of 121 days. The duration of sunshine throughout the year amounts to 2743.3 h. Moreover, the region receives a total solar radiation thermal energy of 137.37 kcal/cm^2/year, which adequately meets the light energy requirements of crops. The annual average accumulated temperature exceeds

3566 °C, ensuring that temperatures remain above freezing point (≥0 °C). This information provides an overview of the geographical and climatic characteristics of Dingbian County, which are essential factors for understanding the research context and interpreting the study findings.
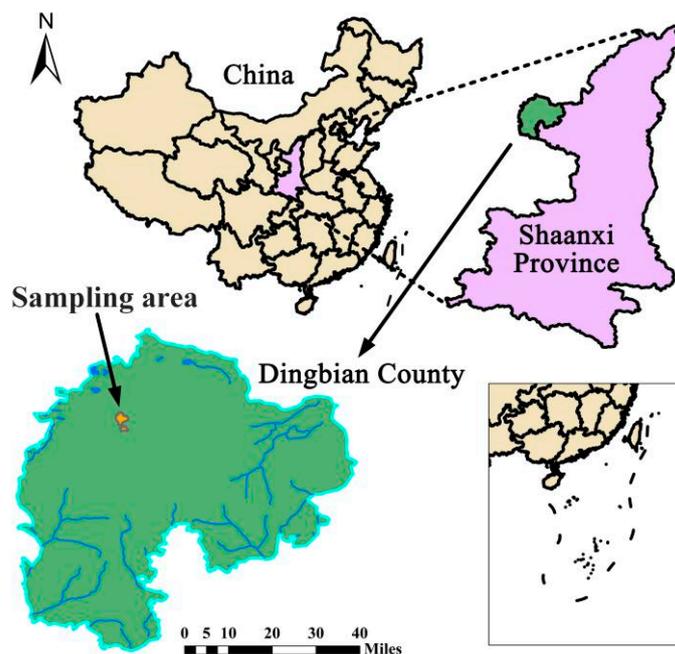
Figure 1. The study area is located in Dingbian County, Shaanxi Province, China.

From 5 to 10 August 2022, a field investigation and sampling were conducted in the research area. Sampling points were uniformly distributed, and GPS locations and corresponding environmental information were recorded. For the study, soil samples were collected from a depth of 0 to 20 cm. A total of 298 soil samples were collected in the research area. After collection, the soil samples were carefully sealed in bags to ensure their preservation and prevent any contamination or loss of soil properties during transportation and storage. This meticulous sample handling process aimed to maintain the integrity and representativeness of the soil samples for subsequent analysis and evaluation.

Initially, the soil moisture content was determined using the drying method. This involved subjecting the soil samples to controlled heating to remove moisture and obtain accurate measurements of moisture content. After moisture determination, the soil samples underwent a series of processing steps. First, they were air-dried to remove any excess moisture. Next, the dried soil samples were carefully crushed to break down any aggregates and ensure uniformity. Foreign materials such as rocks, roots, or debris were meticulously removed to obtain a clean soil sample. Throughout this process, care was taken to retain the integrity of the soil samples, ensuring that they remained intact for subsequent analysis. These processed soil samples were then ready for further examination and evaluation of their properties. This air-dried soil was then mixed thoroughly and used for further analysis, including high-resolution spectroscopy and the determination of soil OM, N, P, and K content.

2.2. Determination of Soil OM, N, P, and K

The analysis of soil nutrient content, including soil OM, N, P, and K, was carried out using standard laboratory methods. The specific procedures for each nutrient determination are outlined below:

The determination of soil OM involves the loss-on-ignition method. A subsample of the soil sample is weighed and placed in a pre-weighed crucible. The crucible with the soil sample is then heated in a furnace at a high temperature (typically around 550 °C) for

several hours to burn off the OM. After the ignition process, the crucible is allowed to cool in a desiccator and then reweighed. The weight loss corresponds to the OM content of the soil sample.

The determination of soil nitrogen content is typically performed using the Kjeldahl method. In this method, a portion of the soil sample is digested with a mixture of concentrated sulfuric acid and a catalyst, such as potassium sulfate. The digestion process breaks down the organic nitrogen compounds present in the soil into ammonium ions. The ammonium ions are then converted into ammonium sulfate through a series of chemical reactions. The ammonium sulfate solution is titrated with a standardized solution of sulfuric acid, and the nitrogen content is calculated based on the volume of acid required for neutralization.

Soil phosphorus content is determined using the Olsen or Bray extraction methods. These methods involve extracting the available phosphorus from the soil sample using specific chemical solutions. To determine the phosphorus concentration in the soil samples, a colorimetric method was employed. The extract obtained from the soil sample was mixed with a colorimetric reagent, resulting in the formation of a colored compound. The intensity of the color was measured using a spectrophotometer, which quantitatively assessed the phosphorus content. To establish a quantitative relationship, a calibration curve was generated using standard phosphorus solutions with known concentrations. The absorbance of the sample was then compared to the calibration curve to determine the phosphorus concentration. This colorimetric approach provided a reliable and accurate measurement of phosphorus levels in the soil samples, allowing for further analysis and interpretation of their nutrient content.

The determination of soil potassium content typically involves the extraction of exchangeable potassium using ammonium acetate or another suitable extractant. The extract is then analyzed using flame photometry or atomic absorption spectroscopy. These methods measure the intensity of the emitted or absorbed light by potassium ions and provide a quantitative measure of the potassium concentration in the soil sample.

By employing these standard laboratory methods, the soil nutrient content, including soil OM, N, P, and K, can be accurately determined. The measured nutrient values serve as the ground truth data for training and validating the soil nutrient estimation models in this research. The results are shown in Table 1. The 298 soil samples were arranged in ascending order of element concentrations, and a concentration gradient method was used to partition them into a validation set and a training set, ensuring that both sets represented the concentration distribution of the entire dataset. A total of 75 samples were designated as the validation set, while the remaining 223 samples constituted the training set. The training and validation sets, after partitioning, maintained consistency with the original dataset in terms of minimum values, maximum values, means, standard deviations, and coefficients of variation for each attribute. The validation set exhibited an increase in the average values of soil attributes and a decrease in standard deviations. This suggests that the validation set contains higher and more concentrated concentrations of organic matter. Since the validation set includes samples with higher concentrations, it is used to assess the model's performance within this particular concentration range.

Table 1. Statistics of nutrient contents in soil samples.

Soil Nutrients	Number of Samples	Minimum $(g \cdot kg^{-1})$	Maximum $(g \cdot kg^{-1})$	Mean $(g \cdot kg^{-1})$	Standard Deviation $(g \cdot kg^{-1})$
OM	298	2.457	4.112	2.983	0.487
N	298	0.723	0.848	0.776	0.225
P	298	0.505	0.776	0.611	0.275
K	298	1.862	2.472	2.210	0.262

2.3. Spectroscopy and Pretreatment

The ground spectrometer ASD Fieldspec 4 from Analytical Spectral Devices in the United States is used to measure soil hyperspectral data. The model number of this spectrometer measures a spectral range of 350 to 2500 nm [52]. During the field measurement, we selected clear and cloudless noon to conduct, with the sensor probe vertically downward and approximately 80 cm away from the soil surface, to obtain the field undisturbed soil spectrum (situ spectrum) [53]. During the indoor spectral measurement, the soil sample was prepared by filling and flattening the processed air-dried soil in a blood container with a diameter of 10 cm and a depth of 2 cm. The soil surface was leveled using a straight edge to ensure uniformity [54]. A standard DC tungsten quartz halogen lamp served as the light source for the measurement. The lamp was placed inside a black box, providing controlled lighting conditions. The light was incident on the soil sample at an angle of 45°, with a distance of 30 cm between the light source and the soil sample. To capture the spectral information, a spectrometer was used, positioned at a distance of 15 cm from the soil sample. The spectrometer measured the intensity of light reflected from the soil sample within a field of view angle of 25° [55,56]. Figure 2 depicts the experimental setup, illustrating the described configuration of the spectral measurement setup. This setup allowed for consistent and standardized measurement of the spectral properties of the soil sample, facilitating the analysis and interpretation of the collected spectral data. We obtained an indoor dry spectrum of the soil and performed timely whiteboard calibration during the measurement process. To eliminate spectral interference information, edge bands with excessive noise (350–399 nm and 2451–2500 nm) were removed, with a sampling interval of 1 nm.
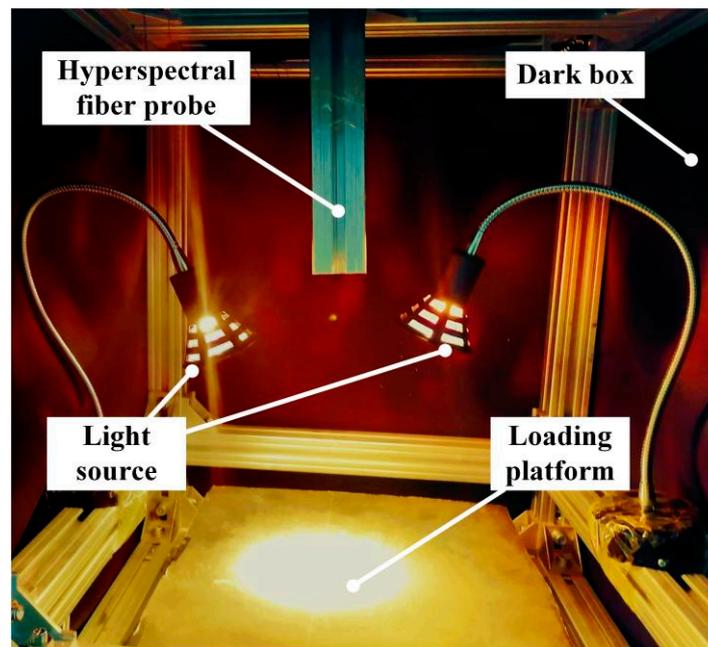

Figure 2. Laboratory data acquisition device.

3. Model Overview

3.1. External Parameter Orthogonalization (EPO)

EPO projects all the spectra onto a space orthogonal to the influencing factor to be removed [57], thereby achieving the purpose of removing this influencing factor. The field spectrum X(n × m). It can be expressed in matrix form as follows:

$$X = XP + XQ + R \tag{1}$$

P(n × m) represents a projection matrix of useful spectral information (soil properties); Q(n × m) represents a projection matrix of useless spectral information (additional environmental factors); R(n × m) is the residual matrix.

The objective of EPO (extraction of useful spectrum) is to extract the valuable spectrum P from the given spectrum X by subtracting the irrelevant information represented by Q. In this process, the identity matrix I is utilized to compute the difference P = X(I − Q). The resulting spectrum P captures the essential and meaningful features for further analysis and interpretation. For the spectrum of saline soil, removing the water factor, the nutrients spectrum can be regarded as a useful spectrum P, and the water spectrum can be regarded as a useless spectrum Q. The water spectral projection Q can be obtained by singular value decomposition of the difference spectrum Dn × m of *Situ-spectra* and *Dry-spectra*, as the useful spectral information in the difference spectrum Dn × m is mainly reflected by soil moisture.

Based on the above principles, this paper uses the EPO idea to study the effect of soil moisture removal based on the difference spectrum Dn × m of *Situ-spectra* and *Dry-spectra*, and obtains an EPO corrected spectrum. The relevant calculations are implemented in MATLAB R2019b.

3.2. Fractional Order Differential

Integer-order differential spectroscopy is the most widely used spectral curve transformation method in soil spectral analysis. Integer-order differential spectroscopy has many advantages, but ignoring differential spectral gradient information limits the accuracy of the estimation model. Fractional differentiation is a method that can refine spectral information and improve modeling accuracy, essentially the slope of arbitrary spectral reflectance. There are three commonly used differential definition methods: *Geunwald-Letnikov*, *Riemann-Liouville*, and *Caputo* [58]. In this paper, the most widely used G-L method is selected to mine soil spectral details. This article selects the most widely used G-L method to mine soil spectral fine information. In MATLAB R2019b, the differential transformation of soil spectral reflectance in the order of 0.2 is achieved. The G-L method expression is [59]:

$$
\frac{d^a f(x)}{dx^a} \approx f(x) + (-a)f(x-1) +
$$
$$
\frac{(-a)(-a+1)}{2} f(x-1) + \cdots \tag{2}
$$
$$
\frac{\Gamma(-a+1)}{n!\Gamma(-a+1)} f(x-n)
$$

In this context, a is an arbitrary order, $f(x)$ is the spectrum, x is the corresponding wavelength, and n is the difference between the upper and lower limits in the differential transformation. It determines the range of differentiation applied to the soil spectral reflectance data. Γ-Gamma is a function.

3.3. Att-BiGRU-RNN Model

The encoder-decoder model, initially developed for natural language processing, has found extensive applications in recent years, particularly in time series forecasting, establishing itself as a pivotal sequence-to-sequence model. Given the intrinsic sequential nature of spectral information across various bands in hyperspectral data, this paper introduces an attention-based bi-directional gated recurrent unit combined with a recurrent neural network model (Att-BiGRU-RNN) for the precise estimation of soil nutrient content. The architecture of this model is depicted in Figure 3, which provides a schematic diagram illustrating the key components, including the encoder and decoder sections. Detailed information regarding each layer's input, output, and parameters can be found in Table 2.
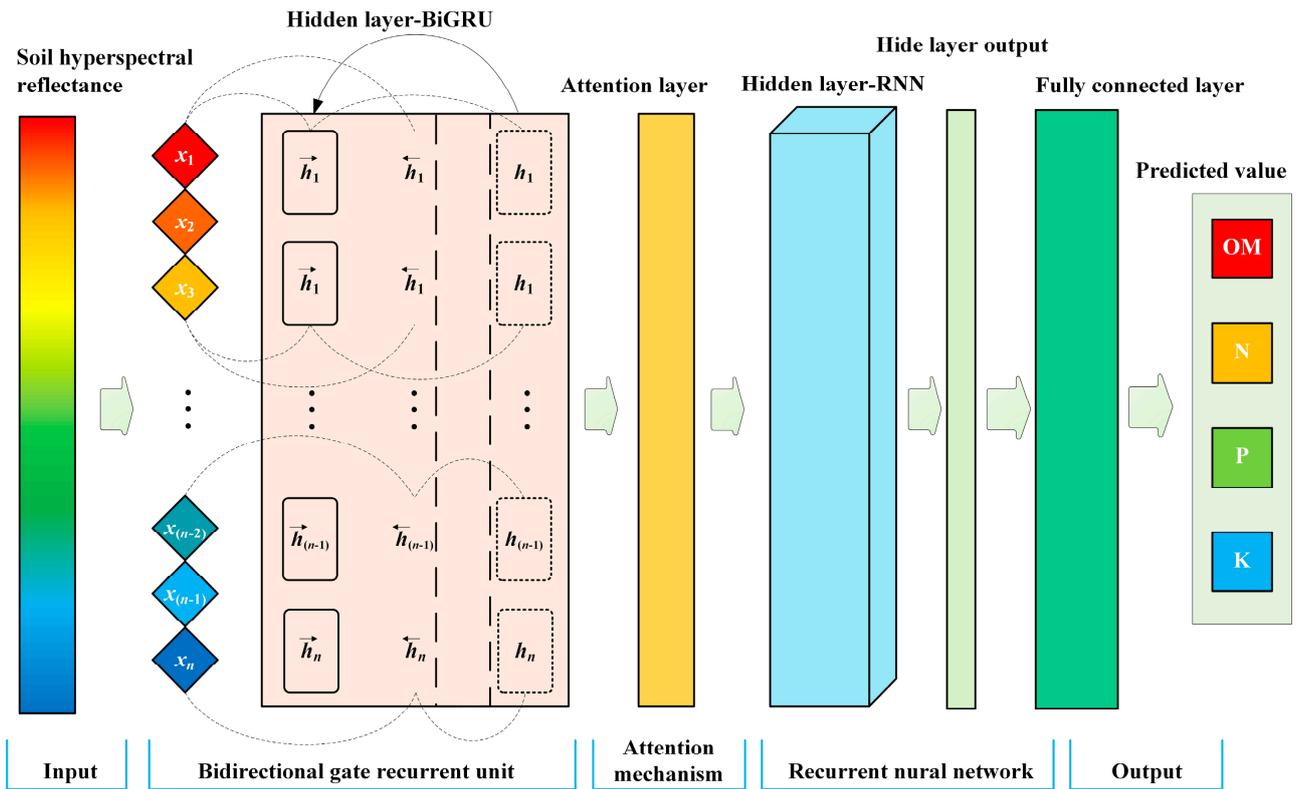
Figure 3. Schematic diagram of the Att-BiGRU-RNN model.

Table 2. Att-BiGRU-RNN network structure parameters.

Layer Type		Input Size	Hidden Size	Bidirectional Processing	Activation Function	Output Size
Att-BiGRU Layer		2151	128	Bidirectional	Tanh	2151
RNN Layer	First Layer	2151	64	NO	Tanh	2151
	Second Layer	64	64	NO	Tanh	64
Fully Connected Layer		64	64	-	ReLU	4
Output Layer		4	-	-	Softmax	4

The encoding module employs bi-directional gated recurrent units (BiGRU) to extract feature information from the spectral sequence X, yielding feature spectral sequence h_t. The decoding module incorporates a fusion attention mechanism to obtain dynamic weights for different spectral feature sequences, thereby performing a weighted average of the encoder's hidden states h_t across various bands. This results in a dynamically weighted spectral feature sequence that is subsequently processed by the decoder's RNN hidden layer, ultimately producing the final feature sequence S. Finally, the softmax function is applied to derive the predicted labels Y.

Bidirectional gated recurrent unit (BiGRU) is an RNN variant with gated structures, including update gate z and reset gate r. These gate structures allow for selective information transmission within the hidden layer, addressing the vanishing gradient problem in RNNs and overcoming short-term memory limitations. The computational process is as follows:

$$
\begin{aligned}
z_i &= \sigma(W_z x_i + U_z h_{t'} + b_z) \\
r_t &= \sigma(W_r x_t + U_z h_{t'} + b_r) \\
\tilde{h}_t &= \tanh(W_h x_t + U_t(r_t \odot h_{t'}) + b_h) \\
\overleftrightarrow{h}_t &= (1 - z_t) \odot h_{t'} + z_t \odot \tilde{h}_t
\end{aligned}
\tag{3}
$$

In this context, *W*, *U*, and *b* are the parameter matrices and biases for the update gate and reset gate in GRU, $W \in R^{k \times n}$, $U \in R^{k \times k}$, $b \in R^k$. x_i is the spectral information of the *t*-th wavelength or band. $\odot$ represents element-wise multiplication. h_t is the previous hidden state information for the *t*-th band, used for computing the forward hidden output (h_{t-1}) during forward pass and backward hidden output (h_{t+1}) during backward pass. $\tilde{h}_t$ represents the hidden state update for the GRU unit. $\overleftrightarrow{h}_t$ is the hidden state output information for the *t*-th band, computed during the forward pass for the forward hidden output and during the backward pass for the backward hidden output. z_t is the update gate vector. r_t is the reset gate vector. σ is the sigmoid activation function. tanh is the hyperbolic tangent activation function.

In this paper, the encoding module employs BiGRU to extract spectral sequence features, consisting of two GRUs that propagate in opposite directions, providing bidirectional spectral sequence information to the output layer simultaneously. Based on Equation (3), the forward and backward hidden state outputs $\overrightarrow{h}_t$ and $\overleftarrow{h}_t$ are calculated separately for the input sequence, and their outputs are concatenated to obtain the final feature spectral sequence.

In the decoding module, a fusion attention mechanism is designed within an RNN structure. The essence of attention is to focus on specific parts of the model and assign higher weights to them, emphasizing the contribution of important spectral bands to the classification model. Figure 4 illustrates the attention mechanism structure. Due to the differences in spectral features for soil nutrient content across different bands, the attention mechanism assigns a weight a_i to each spectral feature, adaptively allocating higher weights to important feature bands, thus enhancing their contribution to the classification model.
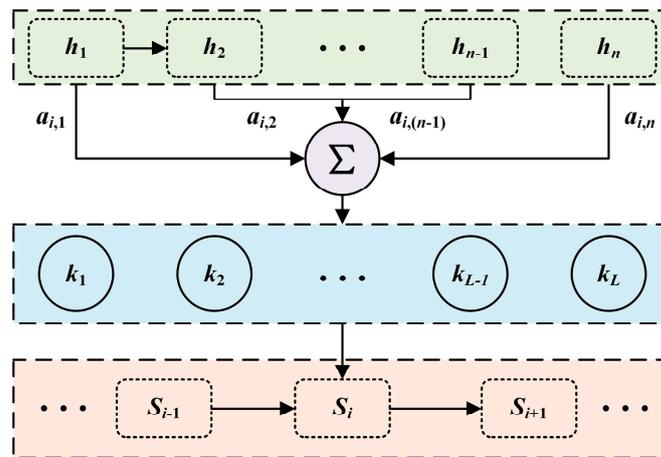

Figure 4. Attention mechanism structure diagram.

The calculation process for obtaining the dynamically weighted fused feature vector, decoding module output (s_t), and final predicted labels (Y) is as follows:

$$e_{t,\,i} = v^{\mathrm{T}} \tan \mathrm{h}(U_1 h_i + U_2 s_{t-1} + b_t)$$

$$a_{t,\,i} = \frac{\exp(e_{t,\,i})}{\sum_{i=1}^{n} \exp(e_{t,\,i})}$$

$$\mathrm{k}_t = \sum_{i=1}^{n} a_{t,\,i} h_i$$

$$y_m = f(\mathrm{s}_t)$$

(4)

In this context, $e_{t,\,i}$ represents the relevance between the *t*-th band and the *i*-th band in the attention mechanism. v, U_1, U_2 are the weight coefficient matrices in the attention mechanism, $v \in \mathrm{R}^m$, $U_1 \in \mathrm{R}^{m \times k}$, $U_2 \in \mathrm{R}^{m \times m}$. b_t is the bias coefficient. $a_{t,\,i}$ is the attention allocation coefficient for the *i*-th band in the encoder to the *t*-th band in the decoder. *n* is the number of spectral bands. *f* is the softmax classifier function. y_m is the predicted labels.

s_t is the output of the decoding module obtained after passing through the hidden layer of the decoding module.

The experimental model in this paper was run on an environment consisting of an Intel Core i5-12490 3.0 GHz 12-core processor running Windows 11, an NVIDIA RTX A2000 (6 GB) GPU, and CUDA v11.0 for training and testing. Table 2 provides the network structure parameters for the Att-BiGRU-RNN model. While keeping the network structure parameters fixed, the study used the open-source AutoML toolkit NNI (neural network intelligence) to search for optimal values of variable parameters. The Adam optimizer was selected as the gradient update rule.

3.4. Convolutional Neural Network

The following description outlines a model based on reference [55] for soil property prediction. The model architecture comprises a three-layer CNN combined with a fully connected neural network. We have chosen this architecture as a point of reference for comparison with our own "Att-BiGRU-RNN Model". The CNN includes three convolutional layers and two pooling layers, followed by three fully connected layers.

To assess the dissimilarity between predicted and actual values, we have opted for the cross-entropy function as the objective function, commonly used in classification tasks. To bridge the gap between the output of the convolutional layers and the input of the fully connected layer, we have introduced a fully connected layer named F4. F4's role is to flatten the output of the third pooling layer (S3) into a one-dimensional sequence, aligning with the structure of the reference model.

Various activation functions are applied to different layers of the model. The rectified linear unit (ReLU) function is employed in the convolutional layers to introduce non-linearity and capture intricate patterns and features within the data. For the fully connected layer, the softmax function is used to transform the output into a probability distribution across different classes, a common practice in multi-class classification tasks.

The CNN model is illustrated in Figure 5, where C1, C2, and C3 represent convolutional layers, and S1, S2, and S3 represent pooling layers. F4 corresponds to the fully connected layer, and F5 represents the dropout layer.
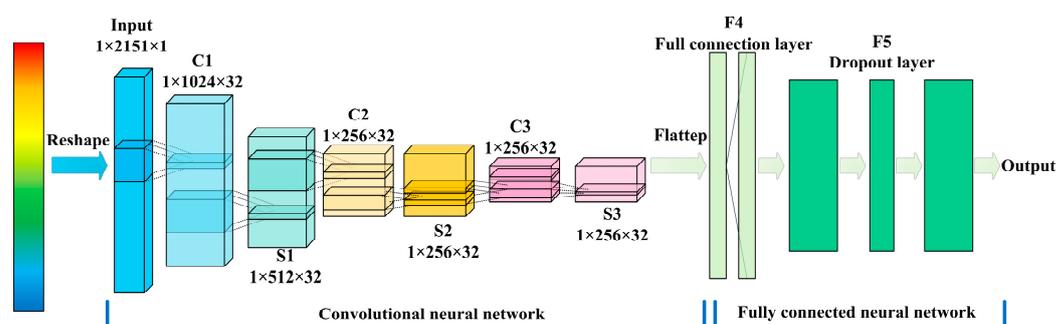

Figure 5. Convolutional neural network model structure.

To mitigate overfitting, two methods are employed in this study. Firstly, the dropout method is utilized, where neurons are randomly dropped out with a specified probability during the training phase. This technique reduces interdependence between hidden nodes and enhances the model's generalization ability. Secondly, the cross-validation method is adopted. A validation set is set aside to evaluate the performance and metrics of the model trained on the training set. Given the limited amount of data, K-fold cross-validation is considered a more suitable approach, with K greater than 2, to ensure a reliable assessment of model performance. This is especially critical as high-level neural networks typically involve a substantial number of parameters to train, and inadequate fold numbers in cross-validation can result in unstable and unreliable results, particularly when dealing with limited data.

The primary purpose of this model is to compare its performance with our own "Att-BiGRU-RNN Model". The model was trained and tested in the following runtime environment: Intel Core i5-12490 3.0 GHz 12-core processor running Windows 11, NVIDIA RTX A2000 (6G) GPU, and CUDA v11.0. The programming language used includes OpenCV 3.3, Keras 2.0, TensorFlow 2.0, and Python 3.6.

3.5. Multivariate Linear Regression Model

In this section, a multivariate linear regression (MLR) model is introduced as a comparison to the Att-BiGRU-RNN model in terms of predictive performance. The MLR model is based on the classical least squares method and aims to address multiple objectives simultaneously. It establishes a linear relationship between a dependent variable and multiple independent variables. By assuming a linear equation, the MLR model captures the relationship between these variables. This model serves as a benchmark for comparison with the Att-BiGRU-RNN model, allowing for an evaluation of their respective predictive capabilities.

3.6. Support Vector Machine Model

In this study, comparative experiments were conducted using two different models: SVR and CNN. SVR is a well-known classification model that aims to find a hyperplane capable of accurately dividing training data and maximizing the geometric interval. To enhance the performance and reduce prediction errors, the radial basis function (RBF) kernel is employed as the kernel function in the SVR model. The RBF kernel is known for its ability to reduce model complexity and improve prediction accuracy. In this study, the SVR model incorporates the RBF kernel with optimal values selected for the moderating factors. The moderating factors include C, epsilon (ε), and gamma (γ). The optimal values chosen for these parameters are C = 1, epsilon (ε) = 0.1, and gamma (γ) = 0.1. These parameter selections were made to ensure the SVR model achieves the best possible performance in terms of accuracy and generalization. By leveraging the RBF kernel and optimizing the moderating factors, the SVR model aims to provide more accurate and reliable predictions for the given soil property estimation task. These values were determined through experimentation and tuning to achieve the best performance. Additionally, by constructing a model based on fractional differentiation, the prediction accuracy of the SVR model in certain frequency bands is increased, the researchers also constructed a fractional order differentiation-based model known as FDT-SVR. This approach incorporates the fractional order differentiation method to enhance the model's accuracy in predicting soil nutrients content within specific bands. By comparing the performance of SVR and CNN models, the researchers aim to evaluate the suitability of these models for predicting soil properties and identify the model that provides the most accurate and reliable predictions in their study. The running environment of this model was trained and tested on the Intel Core i5-12490 3.0 GHz 12 core processor on Windows 11, NVIDIA RTX A2000 (6G) GPU, and CUDA v11.0. The programming language uses OpenCV 3.3, Keras 2.0, and Python 3.6.

3.7. Model Evaluation Criteria

The evaluation of the estimation model adopts the determination coefficient (R^2) and root mean square error (RMSE). The evaluation formulas for R^2 and RMSE are as follows:

$$R^2 = 1 - \frac{\sum\limits_{i=1}^{n} (y_i - Y_i)^2}{\sum\limits_{i=1}^{n} (y_i - \bar{y})^2} \tag{5}$$

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - Y_i)^2} \tag{6}$$

We evaluate the optimal model using the precision–recall (P-R) curve, as follows:

$$\text{Precision} = \frac{TP}{TP + FP} \tag{7}$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{8}$$

In the formula, precision represents the proportion of correctly predicted positive soil samples among all samples predicted as positive, measuring the accuracy of the model. Recall represents the proportion of correctly predicted positive soil samples among all true positive samples, measuring the model's ability to identify positive instances. *TP* represents the number of soil spectral detections that match the actual situation, *FP* represents the number of soil spectral detections that do not match the actual situation, *TN* represents the number of undetected soil spectra that are not relevant to the actual situation, while *FN* represents the number of undetected soil spectra that are relevant to the situation at hand. The P-R curve is a curve drawn with recall as the horizontal axis and precision as the vertical axis.

4. Results and Discussion

4.1. Removal of Soil Moisture Factor Based on EPO

Once the soil moisture factor based on the EPO is eliminated, the average value curve of the indoor air-dried soil spectrum is calculated for each type of sample. This is depicted in Figure 6. The overall trend of the average spectral curve shape of various samples is consistent, especially in the visible light region. After 750 nm, the spectral reflectance difference between samples increases, with a significant increase in spectral reflectance as nutrient content increases.
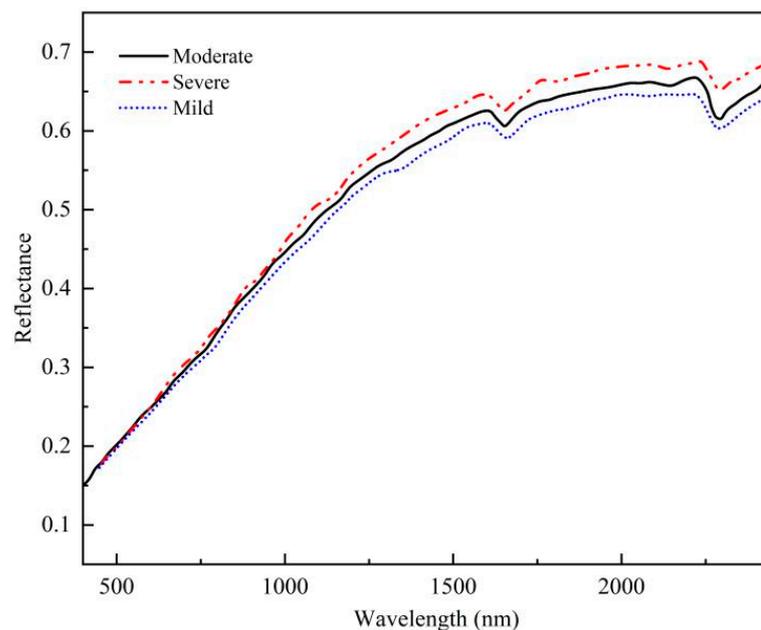

Figure 6. Average value curve of indoor air-dried soil spectrum for each type of sample.

Following the removal of the soil moisture factor through the orthogonalization of external parameters, correlation analysis was performed on the fractional derivative of nutrient content and spectral reflectance of each type of sample, as shown in Figure 7. Notably, the data dimensionality remained unchanged at "2151*1", maintaining a consistent representation throughout the analysis.
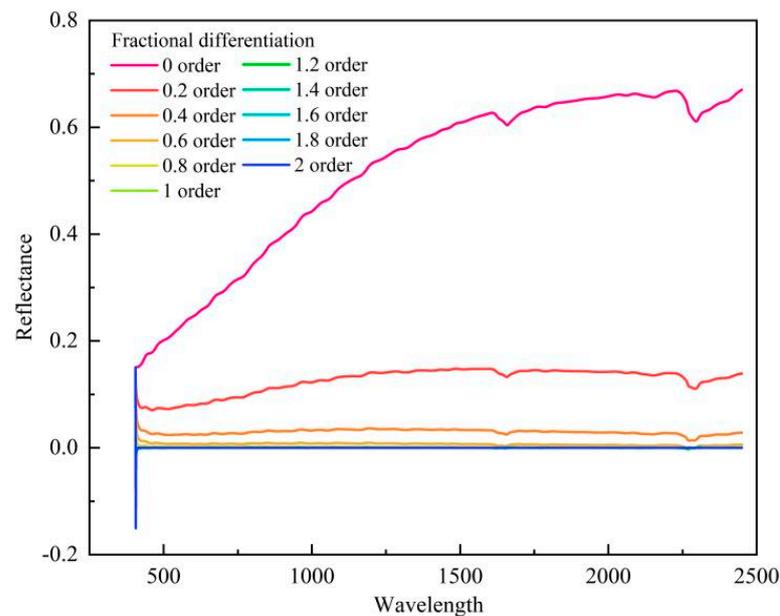
Figure 7. Correlation analysis of fractional differential between nutrient content and soil sample spectral reflectance.

We conducted correlation analysis between EPO and soil nutrients and water content, and analyze the changes in correlation coefficients before and after correction, as shown in Figure 8.
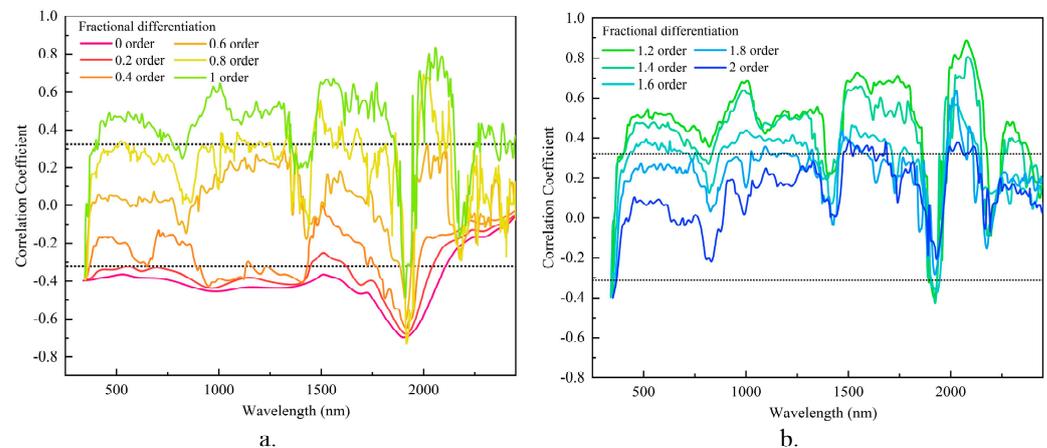
Figure 8. Correlation analysis results of 0–2 order soil nutrients and water content.

It can be seen that the overall waveform of the correlation between the spectrum and nutrients before and after EPO correction is consistent, and the change in the correlation coefficient value after correction is small, with only a few bands showing significant fluctuations. Compared to before correction, the overall waveform of the correlation between spectrum and moisture after EPO correction has significantly improved. After correction, the values of the correlation coefficients have been significantly reduced, with only a few bands showing significant fluctuations.

This indicates that the EPO method effectively reduces the correlation between spectra and soil moisture in most spectral ranges, and can remove the effect of soil moisture. At the same time, the correlation between spectra and soil nutrients was also weakened in local bands.

The estimation effect of soil nutrients after removal of the soil moisture factor based on EPO is shown in Table 3. Comparing the results before and after applying EPO correction, the R^2 values for the training set and validation set increased by 0.15 and 0.33, respectively.

This significant improvement in accuracy demonstrates the effectiveness of EPO correction in soil salt modeling using field undisturbed soil spectra. It enables more precise and stable quantitative analysis of soil salt content. This method can be employed to mitigate the influence of moisture and enhance the accuracy of nutrient estimation in undisturbed field soil samples. By doing so, it enables rapid, in situ, and accurate estimation of soil nutrients.

Table 3. Estimation effect of soil nutrients after removal of the soil moisture factor based on EPO.

Input Spectra	Training Set		Validation Set	
	R^2	RMSE	R^2	RMSE
Situ Spectra	0.61	0.59	0.42	0.56
EPO Spectra	0.76	0.58	0.75	0.54

4.2. Correlation Analysis between Fractional Differential Spectrum and Nutrient Content

Figure 8a illustrates the correlation coefficient distribution between the 0~2-order fractional differential spectrum and soil nutrient content. The dotted line on the parallel abscissa axis indicates the critical line denoting a highly significant correlation level. Figure 8b displays the absolute values of the maximum correlation coefficients associated with each fractional order.

From Figure 8a, it is evident that as the fractional order increases, the correlation coefficient curve exhibits a gradually increasing trend of volatility. Overall, the number of bands passing the highly significant correlation level test decreases, with only a very significant negative correlation observed in the 0–0.6 order range. The original spectrum (0-order) from external parameter orthogonalization closely resembles the correlation coefficient curve of the 0.2-order differential spectrum. The bands that exhibit the highest significant correlation are primarily concentrated within the 450–2150 nm range, with maximum absolute correlation coefficient values of 0.701 and 0.712 occurring at 1907 nm and 1908 nm, respectively. The correlation coefficient curves of the 0.4–1 order differential spectra follow a similar pattern, with gradually increasing positive correlations and occasional curve valleys. The bands that pass the highly significant correlation level test are still predominantly found within the 1700–2300 nm range. The absolute values of the correlation coefficients for the 0.4, 0.6, 0.8, and 1 order differential spectra appear at 1922 nm, 1911 nm, 1916 nm, and 2055 nm, respectively, with corresponding values of 0.680, 0.718, 0.734, and 0.864. The maximum absolute correlation coefficient value is observed at 2074 nm, reaching 0.912.

In summary, the maximum absolute correlation coefficient values for integer-order spectral transformation (orders 0, 1, and 2) of external parameter orthogonalization are 0.701, 0.865, and 0.422, respectively. The maximum absolute correlation coefficient values for fractional order differential spectral correlation coefficients are 0.712, 0.680, 0.718, 0.734, 0.912, 0.839, 0.671, and 0.704, respectively. It is evident that, except for the first-order differential spectrum, the absolute correlation coefficient values increase after applying fractional differential spectral transformation compared to the original spectrum (0-order) and 2-order differential spectrum. The highest value of 0.912 is observed at the 1.2-order, indicating an increase of 0.210, 0.046, and 0.489 compared to the integer-order spectra. These findings demonstrate that fractional differential spectroscopy has the ability to uncover latent information in the spectrum, enhance the sensitivity of spectral information to soil nutrient content, and improve the accuracy of estimation models.

4.3. Determination of Sensitive Waveband

To leverage the benefits of orthogonal fractional differential spectral transformation of external parameters, this study ranked the correlation coefficients' absolute values between each fractional differential and soil nutrient content from highest to lowest. The top 10 fractional differential spectral bands associated with soil nutrient content were selected, and a matrix diagram displaying the absolute correlation coefficient values between these selected bands and soil nutrient content was generated. These results are presented in

Figure 9. Additionally, Figure 10 illustrates the corresponding fractional order for the maximum absolute correlation coefficient value. The selected wavebands include the first-order bands at 2030 nm, 2032 nm, 2055 nm, and 2093 nm, the 1.2-order bands at 1992 nm, 2016 nm, 2054 nm, 2074 nm, and 2094 nm, and the 1.4-order band at 2071 nm. The absolute correlation coefficient values range from 0.803 to 0.912, all surpassing the highly significant correlation level test. Notably, most of the selected wavebands are located near the water absorption peak within the 1700–2300 nm range. This indicates that the spectrum of undisturbed soil in the field is considerably influenced by external factors such as water content and measurement conditions. The spectral reflectance in this band range demonstrates a close relationship with soil nutrient content, suggesting the presence of rich information that can facilitate the accurate estimation of soil nutrient content.
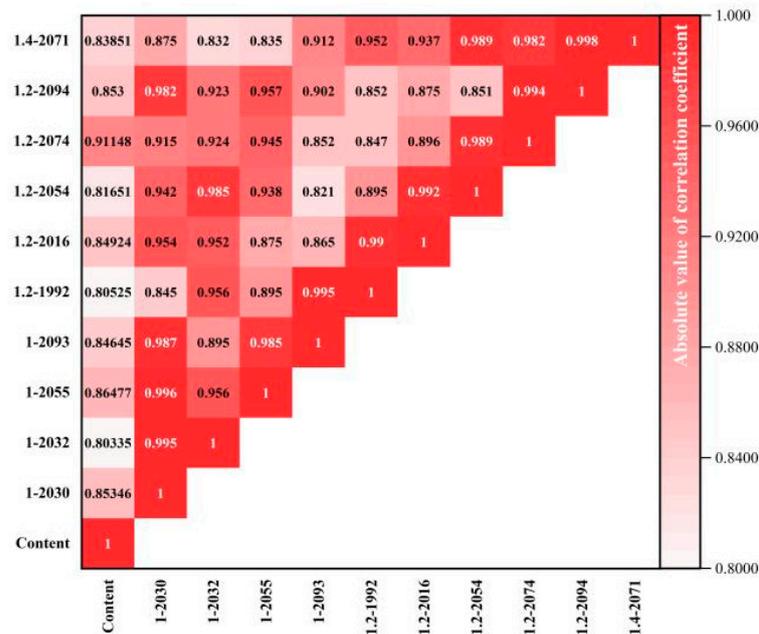

Figure 9. Matrix diagram of the top 10 fractional differential and soil nutrient content correlation coefficients.
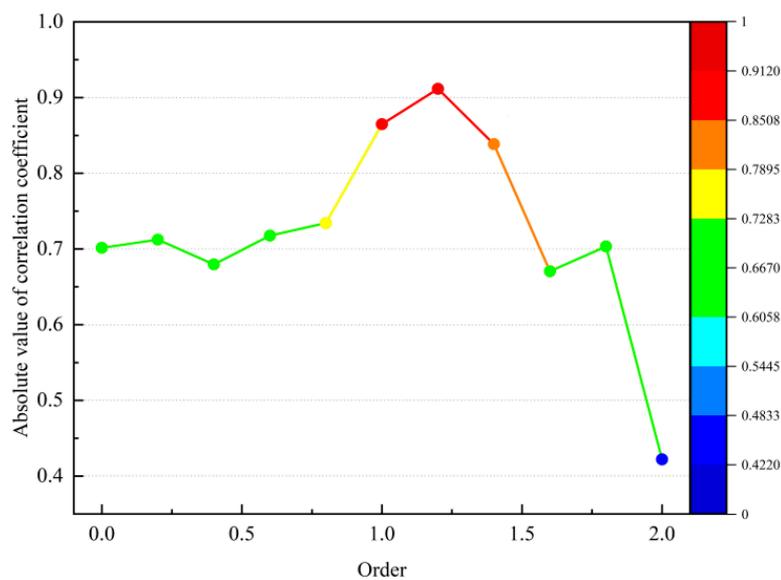

Figure 10. Order corresponding to the maximum absolute correlation coefficient.

4.4. Analysis of Model Prediction Results

The predictive accuracy of the Att-BiGRU-RNN model was evaluated by calculating the R-squared (R^2) and root mean square error (RMSE) for the FDT-MLR, FDT-CNN, FDT-SVR, and BiGRU-RNN models on both the training and validation sets. The corresponding results are presented in Table 4. From the table, it is evident that the high-resolution spectroscopy prediction accuracy for soil organic matter (OM) is consistently higher than that for nitrogen, phosphorus, and potassium elements, irrespective of the training or validation set.

Table 4. Evaluate the performance of each model on the training and validation sets before and after optimization.

Soil Nutrients	Predictive Model	Training Set		Validation Set	
		$R_T{}^2$	$\mathbf{RMSE}_T$	$R_V{}^2$	$\mathbf{RMSE}_V$
OM	BiGRU-RNN	0.931	0.136	0.893	0.166
	Att-BiGRU-RNN	0.962	0.120	0.959	0.118
	FDT-MLR	0.753	0.169	0.727	0.173
	FDT-CNN	0.823	0.188	0.806	0.177
	FDT-SVR	0.697	0.215	0.688	0.191
N	BiGRU-RNN	0.908	0.034	0.818	0.034
	Att-BiGRU-RNN	0.972	0.032	0.907	0.033
	FDT-MLR	0.768	0.039	0.734	0.039
	FDT-CNN	0.840	0.040	0.817	0.024
	FDT-SVR	0.772	0.042	0.726	0.029
P	BiGRU-RNN	0.901	0.045	0.896	0.034
	Att-BiGRU-RNN	0.988	0.037	0.921	0.025
	FDT-MLR	0.799	0.049	0.762	0.039
	FDT-CNN	0.809	0.059	0.781	0.024
	FDT-SVR	0.660	0.067	0.651	0.029
K	BiGRU-RNN	0.929	0.034	0.899	0.039
	Att-BiGRU-RNN	0.959	0.032	0.914	0.025
	FDT-MLR	0.805	0.039	0.798	0.024
	FDT-CNN	0.868	0.040	0.857	0.032
	FDT-SVR	0.858	0.042	0.813	0.031

After optimization with attention mechanisms, the BiGRU-RNN model exhibited a significant improvement in accuracy. In terms of the validation set results, the most significant change was observed in the $R_v{}^2$ value for nitrogen, which increased from 0.818 to 0.907. The largest change in RMSE was observed for soil OM, which decreased from 0.166 g·kg^{-1} to 0.118 g·kg^{-1}.

Among these three elements, soil organic matter exhibited the highest estimation accuracy, with a coefficient of determination (R^2) of 0.959 and an RMSE of 0.118 g·kg^{-1}. Compared to the pre-optimized model, the BiGRU-RNN model showed an improvement of 7.39% in accuracy and 28.92% in reliability. The estimation accuracies for nitrogen, phosphorus, and potassium elements were 0.907, 0.921, and 0.914, respectively.

Based on the evaluation using Taylor diagrams, Figure 11 evaluates the efficiency of different models in predicting soil properties. Among the models considered, the Att-BiGRU-RNN model consistently demonstrated the best performance across multiple soil parameters.
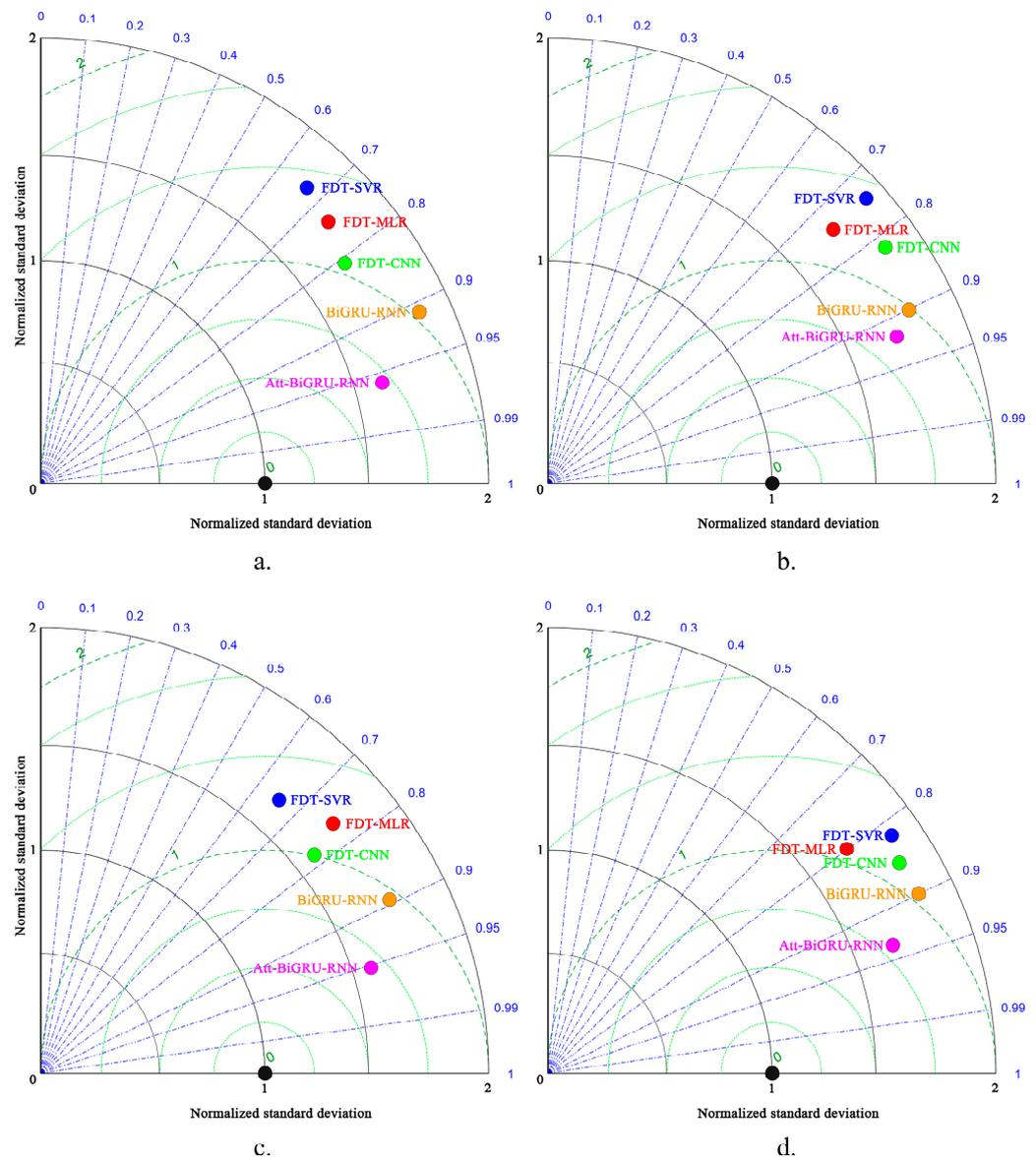
Figure 11. Taylor plots of prediction effects of different models on soil nutrient elements. (**a**) OM prediction effect; (**b**) N element prediction effect; (**c**) P element prediction effect; (**d**) K element prediction effect.

For the prediction of soil OM, the Att-BiGRU-RNN model exhibited a high correlation coefficient of 0.959, indicating a strong relationship between predicted and observed values. Additionally, it achieved a low normalized standard deviation of 0.523, implying minimal variability in the predictions. In contrast, the SVR model showed a lower correlation coefficient of 0.783 and a normalized standard deviation of 1.0, indicating a poorer performance in capturing the variation in soil organic matter.

Regarding N estimation, the Att-BiGRU-RNN model demonstrated a high correlation coefficient of 0.907, indicating strong predictive capability and a reliable relationship between predicted and observed values. Moreover, it exhibited a low normalized standard deviation of 0.407, indicating relatively low variability in the predicted values. In contrast, the SVR model exhibited a lower correlation coefficient of 0.726 and a higher normalized standard deviation of 0.992, suggesting a poorer performance in capturing the variations in nitrogen levels.

Similarly, in the estimation of P levels, the Att-BiGRU-RNN model outperformed the SVR model with a higher correlation coefficient of 0.927, indicating a strong relationship

between predicted and observed values. Additionally, it exhibited a lower normalized standard deviation of 0.651, suggesting more consistent and accurate predictions. In contrast, the SVR model had a lower correlation coefficient of 0.651 and a higher normalized standard deviation of 0.985, indicating a limited accuracy and higher variability in phosphorus estimation.

On the other hand, the SVR model exhibited a lower correlation coefficient of 0.742 and a higher normalized standard deviation of 1.0, indicating a poorer performance in predicting potassium (K) levels. In contrast, the Att-BiGRU-RNN model demonstrated a superior performance with a higher correlation coefficient of 0.923, indicating a strong relationship between predicted and observed values. Additionally, it exhibited a lower normalized standard deviation of 0.619, suggesting more consistent predictions for potassium levels.

Overall, based on the evaluation using Taylor diagrams, the Att-BiGRU-RNN model consistently exhibited the best performance in predicting soil OM, N, P, and K. Its high correlation coefficients and low normalized standard deviations suggest its effectiveness in accurately capturing the variations in these soil properties.

The boxplots presented in Figure 12 provide a comprehensive comparison of the different models used for predicting soil properties, specifically soil organic matter (OM), nitrogen (N), phosphorus (P), and potassium (K). From the boxplots, it is evident that the Att-BiGRU-RNN model exhibits the closest similarity between the median of the predicted values and the observed values for these soil properties. This indicates that the Att-BiGRU-RNN model is capable of producing predictions that closely align with the actual measurements. While some variations are observed in the lower quartile (Q25) and the range of the data (maximum and minimum values) among the models, the Att-BiGRU-RNN model consistently outperforms the other models in terms of prediction accuracy. The differences in these statistical variables suggest that the Att-BiGRU-RNN model provides more reliable and precise predictions compared to the other models. On the other hand, the FDT-MLR and FDT-SVR models exhibit significant variations in the statistical variables for soil OM, N, P, and K. This indicates that these models have limitations in accurately capturing the complexities and variations present in the soil properties. Overall, the results emphasize the superior performance of the Att-BiGRU-RNN model in generating predictions that closely match the observed values for soil OM, N, P, and K. In contrast, the FDT-MLR and FDT-SVR models exhibit a less favorable performance, indicating their limitations in achieving the same level of accuracy as the Att-BiGRU-RNN model.

In conclusion, the performance of the hybrid Att-BiGRU-RNN model surpasses that of the standalone BiGRU-RNN model, CNN, and SVR models. This is because the BiGRU-RNN model exhibits a good estimation accuracy but lacks robustness in terms of distribution state and probability density fitting, while the Att-BiGRU-RNN model possesses better global search capabilities. The results of this study clearly demonstrate that the hybrid Att-BiGRU-RNN model developed in this research can serve as an effective and viable approach for accurately predicting soil OM, N, P, and K.

Figure 13 depicts the performance of the optimized Att-BiGRU-RNN model in terms of accuracy and recall. Different model parameter choices can meet different estimation requirements. Using hyperspectral technology for soil nutrient detection can identify the nutrient distribution in a certain area on a large scale and perform statistical mapping of soil nutrient distribution areas.

From the P-R (Precision–Recall) curves, it can be observed that the Att-BiGRU-RNN model exhibits the largest area under the curve for identifying soil OM content, indicating its superior robustness. It can maintain a good recall rate even in high-precision identification tasks. On the other hand, for the identification of N content, the P-R curve for the Att-BiGRU-RNN model has the smallest area, suggesting that a trade-off needs to be made between prediction accuracy and recall rate in practical operations. For large-scale identification of N element distribution areas, a higher recall rate may be preferred, while for precision-demanding identification tasks, a higher prediction accuracy can be prioritized.
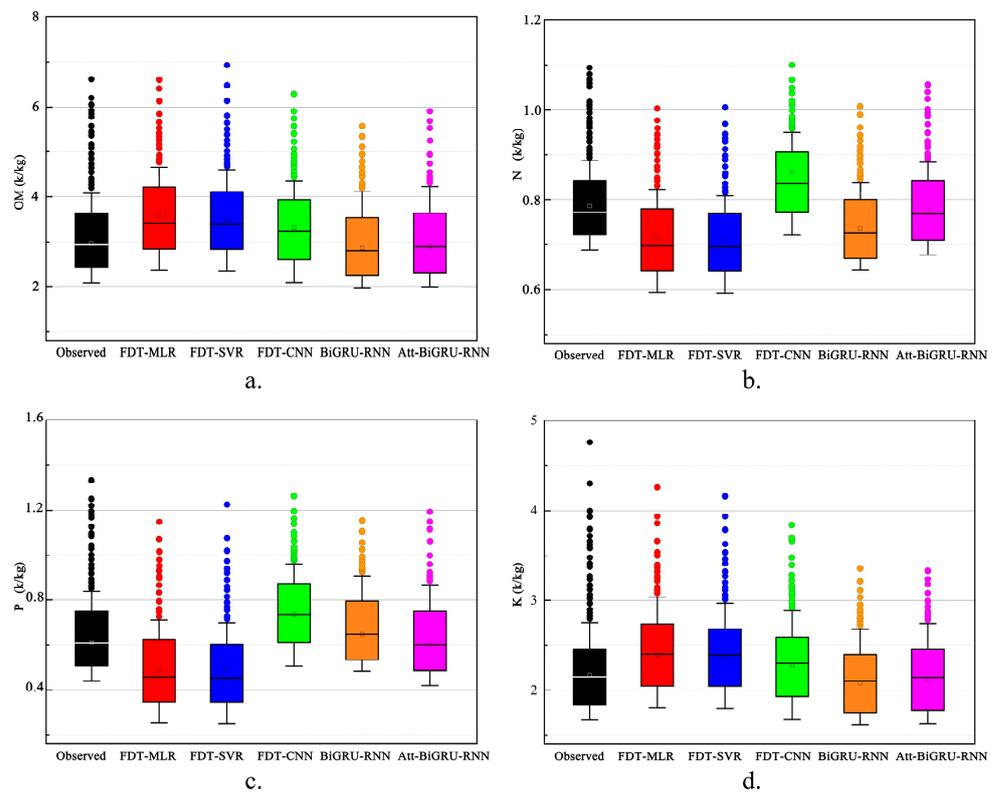
Figure 12. Boxplots of prediction effects of different models on soil nutrient elements. (**a**) OM prediction effect; (**b**) N element prediction effect; (**c**) P element prediction effect; (**d**) K element prediction effect.
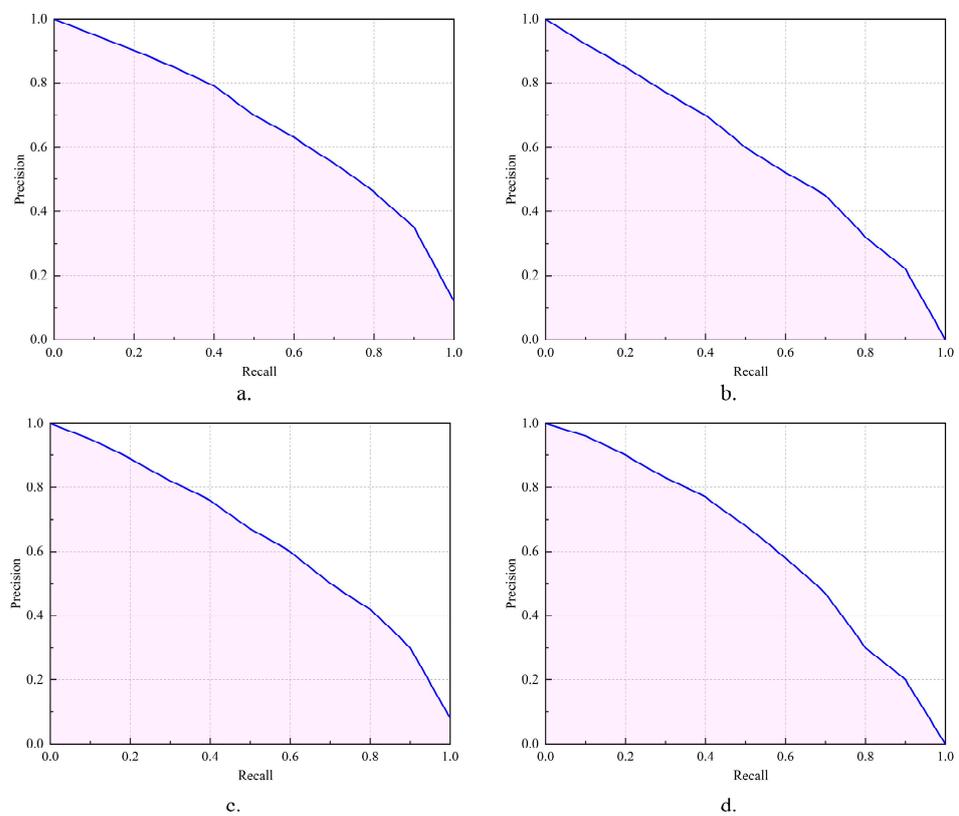
Figure 13. P-R curve of the Att-BiGRU-RNN model for soil nutrient content. (**a**) P-R curve of OM; (**b**) P-R curve of N element; (**c**) P-R curve of P element; (**d**) P-R curve of K element.

5. Conclusions

The BiGRU-RNN model demonstrates a strong estimation accuracy, yet it faces challenges in capturing the distribution state and fitting probability density effectively. In contrast, the hybrid Att-BiGRU-RNN model leverages the advantages of the BiGRU-RNN model while incorporating the attention mechanism, resulting in enhanced global search capabilities and an improved overall performance. These findings from our study underscore that the optimized Att-BiGRU-RNN model offers a promising alternative for precise predictions of soil OM, N, P, and K.

Furthermore, we employed Taylor diagrams to evaluate the efficiency of our models. The Att-BiGRU-RNN model exhibited the highest correlation coefficients and the lowest normalized standard deviations for the prediction of OM, N, P, and K, indicating its superior performance. In contrast, the SVR model displayed lower correlation coefficients and higher normalized standard deviations, indicative of a comparatively poorer performance. The Att-BiGRU-RNN model consistently demonstrated the best predictive accuracy among the models we assessed.

To further illustrate the model comparisons, we utilized boxplots to visualize the similarities and differences. The Att-BiGRU-RNN model consistently exhibited the closest resemblance between the median of predicted values and the observed values for OM, N, P, and K. While there were variations in the lower quartile and data range among the models, the Att-BiGRU-RNN model consistently outperformed the other models in terms of prediction accuracy. Conversely, the FDT-MLR and FDT-SVR models displayed larger variations in statistical variables, implying their comparatively less favorable performance.

In summary, our study highlights the superior performance of the hybrid Att-BiGRU-RNN model when compared to other models, positioning it as a promising choice for predicting soil nutrient elements. The fusion of BiGRU-RNN with the attention mechanism significantly enhances estimation accuracy and provides enhanced global search capabilities. This study emphasizes the potential of the Att-BiGRU-RNN model in large-scale identification of areas affected by drought or severe salinization and in statistical mapping of soil nutrient distribution areas.

For future research directions, we recognize the importance of optimizing model parameters to address specific estimation requirements. Additionally, exploring the application of hyperspectral techniques in soil nutrient detection holds promise for advancing agricultural management and environmental monitoring, and we aim to further investigate this avenue.

References

1. Aslan, M.F.; Durdu, A.; Sabanci, K.; Ropelewska, E.; Gültekin, S.S. A Comprehensive Survey of the Recent Studies with UAV for Precision Agriculture in Open Fields and Greenhouses. *Appl. Sci.* **2022**, *12*, 1047. [CrossRef]
2. Boursianis, A.D.; Papadopoulou, M.S.; Diamantoulakis, P.; Liopa-Tsakalidi, A.; Barouchas, P.; Salahas, G.; Wan, S.; Goudos, S.K. Internet of Things (Iot) and Agricultural Unmanned Aerial Vehicles (UAVs) in Smart Farming: A Comprehensive Review. *Internet Things* **2022**, *18*, 100187. [CrossRef]
3. Filho, W.L.; Nagy, G.J.; Setti, A.F.F.; Sharifi, A.; Donkor, F.K.; Batista, K.; Djekic, I. Handling the Impacts of Climate Change on Soil Biodiversity. *Sci. Total Environ.* **2023**, *869*, 161671. [CrossRef] [PubMed]
4. Khanna, A.; Kaur, S. An Empirical Analysis on Adoption of Precision Agricultural Techniques among Farmers of Punjab for Efficient Land Administration. *Land Use Policy* **2023**, *126*, 106533. [CrossRef]
5. Liakos, K.G.; Busato, P.; Moshou, D.; Pearson, S.; Bochtis, D. Machine Learning in Agriculture: A Review. *Sensors* **2018**, *18*, 2674. [CrossRef] [PubMed]
6. Liu, Y.; Ma, X.Y.; Shu, L.; Hancke, G.P.; Abu-Mahfouz, A.M. From Industry 4.0 to Agriculture 4.0: Current Status, Enabling Technologies, and Research Challenges. *IEEE Trans. Ind. Inform.* **2021**, *17*, 4322–4334. [CrossRef]
7. Khanal, S.; Fulton, J.; Shearer, S. An Overview of Current and Potential Applications of Thermal Remote Sensing in Precision Agriculture. *Comput. Electron. Agric.* **2017**, *139*, 22–32. [CrossRef]
8. Hossain, M.Z.; Bahar, M.M.; Sarkar, B.; Donne, S.W.; Ok, Y.S.; Palansooriya, K.N.; Kirkham, M.B.; Chowdhury, S.; Bolan, N. Biochar and Its Importance on Nutrient Dynamics in Soil and Plant. *Biochar* **2020**, *2*, 379–420. [CrossRef]
9. Hengl, T.; Leenaars, J.G.B.; Shepherd, K.D.; Walsh, M.G.; Heuvelink, G.B.M.; Mamo, T.; Tilahun, H.; Berkhout, E.; Cooper, M.; Fegraus, E.; et al. Soil Nutrient Maps of Sub-Saharan Africa: Assessment of Soil Nutrient Content at 250 M Spatial Resolution Using Machine Learning. *Nutr. Cycling Agroecosyst.* **2017**, *109*, 77–102. [CrossRef]
10. Kirkby, C.A.; Richardson, A.E.; Wade, L.J.; Passioura, J.B.; Batten, G.D.; Blanchard, C.; Kirkegaard, J.A. Nutrient Availability Limits Carbon Sequestration in Arable Soils. *Soil Biol. Biochem.* **2014**, *68*, 402–409. [CrossRef]
11. Nkebiwe, P.M.; Weinmann, M.; Bar-Tal, A.; Muller, T. Fertilizer Placement to Improve Crop Nutrient Acquisition and Yield: A Review and Meta-Analysis. *Field Crops Res.* **2016**, *196*, 389–401. [CrossRef]
12. Cotrufo, M.F.; Soong, J.L.; Horton, A.J.; Campbell, E.E.; Haddix, M.L.; Wall, D.H.; Parton, A.J. Formation of Soil Organic Matter Via Biochemical and Physical Pathways of Litter Mass Loss. *Nat. Geosci.* **2015**, *8*, 776–779. [CrossRef]
13. Witzgall, K.; Vidal, A.; Schubert, D.I.; Hoeschen, C.; Schweizer, S.A.; Buegger, F.; Pouteau, V.; Chenu, C.; Mueller, C.W. Particulate Organic Matter as a Functional Soil Component for Persistent Soil Organic Carbon. *Nat. Commun.* **2021**, *12*, 4115. [CrossRef] [PubMed]
14. Li, Z.L.; Tian, D.S.; Wang, B.X.; Wang, J.S.; Wang, S.; Chen, H.Y.H. Microbes Drive Global Soil Nitrogen Mineralization and Availability. *Glob. Chang. Biol.* **2019**, *25*, 1078–1088. [CrossRef] [PubMed]
15. Oswald, R.; Behrendt, T.; Ermel, M.; Wu, D.; Su, H.; Cheng, Y. Hono Emissions from Soil Bacteria as a Major Source of Atmospheric Reactive Nitrogen. *Science* **2013**, *341*, 1233–1235. [CrossRef] [PubMed]
16. Turner, B.L.; Lambers, H.; Condron, L.M.; Cramer, M.D.; Leake, J.R.; Richardson, A.E. Soil Microbial Biomass and the Fate of Phosphorus During Long-Term Ecosystem Development. *Plant Soil* **2013**, *367*, 225–234. [CrossRef]
17. Wu, X.J.; Rensing, C.; Han, D.F.; Xiao, K.Q.; Dai, Y.X.; Tang, Z.X. Genome-Resolved Metagenomics Reveals Distinct Phosphorus Acquisition Strategies between Soil Microbiomes. *mSystems* **2022**, *7*, e01107-21. [CrossRef] [PubMed]
18. Etesami, H.; Emami, S.; Alikhani, H.A. Potassium Solubilizing Bacteria (KSB): Mechanisms, Promotion of Plant Growth, and Future Prospects—A Review. *J. Soil Sci. Plant Nutr.* **2017**, *17*, 897–911. [CrossRef]
19. Soumare, A.; Sarr, D.; Diedhiou, A.G. Potassium Sources, Microorganisms and Plant Nutrition: Challenges and Future Research Directions. *Pedosphere* **2023**, *33*, 105–115. [CrossRef]
20. Zheng, L.H.; Li, M.Z.; Pan, L.; Sun, R.Y.; Tang, N. Computation of Spectra in Estimation of Soil Nutrients. In Proceedings of the 4th International Symposium on Intelligent Information Technology in Agriculture (ISIITA), Beijing, China, 26–29 October 2007.
21. Peng, Y.P.; Wang, L.; Zhao, L.; Liu, Z.H.; Lin, C.J.; Hu, Y.M. Estimation of Soil Nutrient Content Using Hyperspectral Data. *Agriculture* **2021**, *11*, 1129. [CrossRef]
22. Peng, Y.P.; Zhao, L.; Hu, Y.M.; Wang, G.X.; Wang, L.; Liu, Z.H. Prediction of Soil Nutrient Contents Using Visible and near-Infrared Reflectance Spectroscopy. *ISPRS Int. J. Geo-Inf.* **2019**, *8*, 437. [CrossRef]
23. Wang, H.; Zhang, L.X.; Zhao, J.W.; Hu, X.; Ma, X. Application of Hyperspectral Technology Combined with Bat Algorithm-Adaboost Model in Field Soil Nutrient Prediction. *IEEE Access* **2022**, *10*, 100286–100299. [CrossRef]
24. Martins, K.V.; Dourado-Neto, D.; Reichardt, K.; Favarin, J.L.; Sartori, F.F.; Felisberto, G.; Mello, S.C. Maize Dry Matter Production and Macronutrient Extraction Model as a New Approach for Fertilizer Rate Estimation. *An. Acad. Bras. Cienc.* **2017**, *89*, 705–716. [CrossRef] [PubMed]
25. Singh, S.; Kasana, S.S. Quantitative Estimation of Soil Properties Using Hybrid Features and Rnn Variants. *Chemosphere* **2022**, *287*, 131889. [CrossRef] [PubMed]
26. An, X.F.; Li, M.Z.; Zheng, L.H.; Liu, Y.M.; Sun, H. A Portable Soil Nitrogen Detector Based on Nirs. *Precis. Agric.* **2014**, *15*, 3–16. [CrossRef]
27. Ma, L.; Li, A.Q.; Yua, H.L.; Chen, G.F. Hyperspectral Remote Sensing Estimation of Soil Nutrients in the Black Soil Region Based on Computer Vision Model. *ScienceAsia* **2022**, *48*, 287–293. [CrossRef]

28. Xie, S.G.; Li, Y.H.; Wang, X.; Liu, Z.X.; Ma, K.L.; Ding, L.W. Research on Estimation Models of the Spectral Characteristics of Soil Organic Matter Based on the Soil Particle Size. *Spectrochim. Acta A Mol. Biomol. Spectrosc.* **2021**, *260*, 119963. [CrossRef]

29. Bhat, R.; Sujatha, S.; Jose, C.T. Assessing Soil Fertility of a Laterite Soil in Relation to Yield of Arecanut (*Areca catechu* L.) in Humid Tropics of India. *Geoderma* **2012**, *189*, 91–97. [CrossRef]

30. Liu, R.M.; Xu, F.; Yu, W.W.; Shi, J.H.; Zhang, P.P.; Shen, Z.Y. Analysis of Field-Scale Spatial Correlations and Variations of Soil Nutrients Using Geostatistics. *Environ. Monit. Assess.* **2016**, *188*, 126. [CrossRef]

31. Lin, J.H.; Wang, M.H.; Zhang, M.; Zhang, Y.N.; Chen, L. Electrochemical Sensors for Soil Nutrient Detection: Opportunity and Challenge. In Proceedings of the 1st International Conference on Computer and Computing Technologies in Agriculture (CCTA 2007), Wuyishan, China, 18–20 August 2007.

32. Pohl, L.; Maurischat, P.; Schiedung, M.; Weber, T.K.D.; Winkler, P. Soil Science Challenges-an Interdisciplinary Overview of Current and Future Topics. *J. Plant Nutr. Soil Sci.* **2022**, *185*, 691–693. [CrossRef]

33. Smaling, E.M.A.; Fresco, L.O.; de Jager, A. Classifying, Monitoring and Improving Soil Nutrient Stacks and Flows in African Agriculture. *Ambio* **1996**, *25*, 492–496.

34. Ayamba, B.E.; Abaidoo, R.C.; Opoku, A.; Ewusi-Mensah, N. Mechanisms for Nutrient Interactions from Organic Amendments and Mineral Fertilizer Inputs under Cropping Systems: A Review. *PeerJ* **2023**, *11*, e15135. [CrossRef] [PubMed]

35. Kaur, G.; Das, K.; Hazra, J. Soil Nutrients Prediction Using Remote Sensing Data in Western India: An Evaluation of Machine Learning Models. In Proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Waikoloa, HI, USA, 26 September–2 October 2020.

36. Wu, C.Y.; Chen, Y.F.; Hong, X.J.; Liu, Z.L.; Peng, C.H. Evaluating Soil Nutrients of Dacrydium Pectinatum in China Using Machine Learning Techniques. *For. Ecosyst.* **2020**, *7*, 30. [CrossRef]

37. Lin, L.X.; Gao, Z.Q.; Liu, X.X. Estimation of Soil Total Nitrogen Using the Synthetic Color Learning Machine (Sclm) Method and Hyperspectral Data. *Geoderma* **2020**, *380*, 114664. [CrossRef]

38. Chlingaryan, A.; Sukkarieh, S.; Whelan, B. Machine Learning Approaches for Crop Yield Prediction and Nitrogen Status Estimation in Precision Agriculture: A Review. *Comput. Electron. Agric.* **2018**, *151*, 61–69. [CrossRef]

39. Singh, S.; Kasana, S.S. Estimation of Soil Properties from the Eu Spectral Library Using Long Short-Term Memory Networks. *Geoderma Reg.* **2019**, *18*, e00233. [CrossRef]

40. Andrade, R.; Silva, S.H.G.; Benedet, L.; Mancini, M.; Lima, G.J.; Nascimento, K.; Amaral, F.H.C.; Silva, D.R.G.; Ottoni, M.V.; Carneiro, M.A.C.; et al. Proximal Sensing Provides Clean, Fast, and Accurate Quality Control of Organic and Mineral Fertilizers. *Environ. Res.* **2023**, *236*, 116753. [CrossRef]

41. Yamac, S.S.; Negis, H.; Seker, C.; Memon, A.M.; Kurtulus, B.; Todorovic, M. Saturated Hydraulic Conductivity Estimation Using Artificial Intelligence Techniques: A Case Study for Calcareous Alluvial Soils in a Semi-Arid Region. *Water* **2022**, *14*, 3875. [CrossRef]

42. Niu, Z.; Shi, L.; Qiao, H.B.; Xu, X.; Wang, W.W.; Ma, X.M. Construction of a Hyperspectral Estimation Model for Total Nitrogen Content in Shajiang Black Soil. *J. Plant Nutr. Soil Sci.* **2023**, *186*, 196–208. [CrossRef]

43. Garcia-Nieto, P.J.; Garcia-Gonzalo, E.; Fernandez, J.R.A.; Muniz, C.D. Predictive Modelling of Eutrophication in the Pozn De La Dolores Lake (Northern Spain) by Using an Evolutionary Support Vector Machines Approach. *J. Math. Biol.* **2018**, *76*, 817–840. [CrossRef]

44. Li, H.; Leng, W.; Zhou, Y.; Chen, F.; Xiu, Z.; Yang, D. Evaluation Models for Soil Nutrient Based on Support Vector Machine and Artificial Neural Networks. *Sci. World J.* **2014**, *2014*, 478569. [CrossRef] [PubMed]

45. Yao, H.B.; Tian, L.; Kaleita, A. Hyperspectral Image Feature Extraction and Classification for Soil Nutrient Mapping. In Proceedings of the 4th European Conference on Precision Agriculture, Berlin, Germany, 15 June 2003.

46. Liu, H.; Zhang, X.; Cao, L.; Li, D.X. A Study on Evaluation of Farmland Fertility Levels Based on Optimization of the Decision Tree Algorithm. In Proceedings of the 2nd International Conference on Computer Science and Network Technology (ICCSNT), Changchun, China, 29–31 December 2012.

47. Hu, H.L.; Wu, J.D.; Wang, X.D.; Huang, K.K. Relation Model of Soil Nutrients and Tobacco Quality Based on Principal Component and Regression Analysis. In Proceedings of the 2nd International Symposium on Computer, Communication, Control and Automation (3CA), Singapore, 1–2 December 2013.

48. Jonathan, L.; Adeline, F.; Andyne, L.; Celine, P.; Hugues, C. Prediction of Forest Nutrient and Moisture Regimes from Understory Vegetation with Random Forest Classification Models. *Ecol. Indic.* **2022**, *144*, 109446. [CrossRef]

49. Trontelj ml., J.; Chambers, O. Machine Learning Strategy for Soil Nutrients Prediction Using Spectroscopic Method. *Sensors* **2021**, *21*, 4208. [CrossRef] [PubMed]

50. Uwiragiye, Y.; Ngaba, M.Y.; Zhao, M.Z.; Elrys, A.S.; Heuvelink, G.B.M.; Zhou, J.B. Modelling and Mapping Soil Nutrient Depletion in Humid Highlands of East Africa Using Ensemble Machine Learning: A Case Study from Rwanda. *Catena* **2022**, *217*, 106499. [CrossRef]

51. Reashma, S.; Pillai, A.S. Edaphic Factors and Crop Growth Using Machine Learning–a Review. In Proceedings of the International Conference on Intelligent Sustainable Systems (ICISS), Palladam, India, 7–8 December 2017.

52. Wang, D.C.; Zhang, D.Y.; Zhao, J.L.; Li, C.J.; Zhu, D.Z.; Huang, W.J.; Li, Y.F.; Yang, X.D. Using Extraction of Red Edge Position to Validate Consistency of Hyperspectral Imaging and Non-Imaging Data. *Spectrosc. Spect. Anal.* **2011**, *31*, 2450–2454.

53. Tian, A.H.; Zhao, J.S.; Tang, B.H.; Zhu, D.M.; Fu, C.B.; Xiong, H.G. Study on the Pretreatment of Soil Hyperspectral and Na$^+$ Ion Data under Different Degrees of Human Activity Stress by Fractional-Order Derivatives. *Remote Sens.* **2021**, *13*, 3974. [CrossRef]

54. Dworak, V.; Mahns, B.; Selbeck, J.; Gebbers, R.; Weltzien, C. Hyperspectral Imaging Tera Hertz System for Soil Analysis: Initial Results. *Sensors* **2020**, *20*, 5660. [CrossRef]

55. Wang, H.; Zhang, L.X.; Zhao, J.W.; Hu, X.; Ma, X. Application of Hyperspectral Technology Combined with Genetic Algorithm to Optimize Convolution Long- and Short-Memory Hybrid Neural Network Model in Soil Moisture and Organic Matter. *Appl. Sci.* **2022**, *12*, 10333. [CrossRef]

56. Ge, X.Y.; Wang, J.Z.; Ding, J.L.; Cao, X.Y.; Zhang, Z.P.; Liu, J.; Li, X.H. Combining Uav-Based Hyperspectral Imagery and Machine Learning Algorithms for Soil Moisture Content Monitoring. *PeerJ* **2019**, *7*, e6926. [CrossRef]

57. Hans, G.; Allison, B. Temperature and Moisture Insensitive Prediction of Biomass Calorific Value from near Infrared Spectra Using External Parameter Orthogonalization. *J. Near Infrared Spectrosc.* **2019**, *27*, 259–269. [CrossRef]

58. Almeida, R.; Bastos, N.R.O.; Monteiro, M.T.T. Modeling Some Real Phenomena by Fractional Differential Equations. *Math. Methods Appl. Sci.* **2016**, *39*, 4846–4855. [CrossRef]

59. Chen, Q.; Liu, J.P.; Tang, Z.H.; Li, J.Q.; Wu, M. Detection and Extraction of Image Edge Curves and Detailed Features Using Fractional Differentiation. *Acta Electron. Sin.* **2013**, *41*, 1873–1880.