



Article

The Prediction Model of Total Nitrogen Content in Leaves of Korla Fragrant Pear Was Established Based on Near Infrared Spectroscopy

Mingyang Yu ^{1,2,†}, Xinlu Bai ^{3,4,†} , Jianping Bao ^{1,2}, Zengheng Wang ^{1,2}, Zhihui Tang ⁵, Qiangqing Zheng ⁶ and Jinhu Zhi ^{3,4,*}

¹ National and Local Joint Engineering Laboratory of High Efficiency and High-Quality Cultivation and Deep Processing Technology of Characteristic Fruit Trees in Southern Xinjiang, Alar 843300, China; 10757231040@taru.edu.cn (M.Y.); baobao-xinjiang@126.com (J.B.); 10757213041@stumail.taru.edu.cn (Z.W.)

² College of Horticulture and Forestry Science, Tarim University, Alar 843300, China

³ College of Agriculture, Tarim University, Alar 843300, China; bxl0112@163.com

⁴ The Research Center of Oasis Agricultural Resources and Environment in Southern Xinjiang, Tarim University, Alar 843300, China

⁵ Institute of Mechanical Equipment, Xinjiang Academy of Agricultural Sciences, Shihezi 832000, China; xjtzh0701@sina.com

⁶ Institute of Forestry and Horticulture, Xinjiang Academy of Agricultural Sciences, Shihezi 832000, China; zhengqq369@163.com

* Correspondence: zjhzky@163.com

† These authors contributed equally to this work.

Abstract: In order to efficiently detect total nitrogen content in Korla fragrant pear leaves, near-infrared spectroscopy technology was utilized to develop a detection model. The collected spectra underwent various preprocessing techniques including first-order derivative, second-order derivative, Savitzky–Golay + second-order derivative, multivariate scattering correction, multivariate scattering correction + first-order derivative, and standard normal variable transformation + second-order derivative. A competitive adaptive reweighted sampling algorithm was employed to extract characteristic wavelengths, and a prediction model for the total nitrogen content of fragrant pear leaves was established by combining the random forest algorithm, genetic algorithm-based random forest algorithm, radial basis neural network algorithm, and extreme learning machine algorithm. The study found that spectral preprocessing of SNV + SD along with the radial basis neural network algorithm yielded better predictions for total nitrogen content of fragrant pear leaves. The validation set results showed an R^2 of 0.8547, RMSE of 0.291%, and RPD of 2.699. Therefore, the SNV + SD + CARS + RBF algorithm combination model proved to offer optimal comprehensive performance in predicting the total nitrogen content of fragrant pear leaves.

Keywords: Korla fragrant pear; leaves; total nitrogen content; near infrared spectroscopy



Citation: Yu, M.; Bai, X.; Bao, J.; Wang, Z.; Tang, Z.; Zheng, Q.; Zhi, J. The Prediction Model of Total Nitrogen Content in Leaves of Korla Fragrant Pear Was Established Based on Near Infrared Spectroscopy. *Agronomy* **2024**, *14*, 1284. <https://doi.org/10.3390/agronomy14061284>

Academic Editors: Ana Elisa Rato and Julio Nogales-Bueno

Received: 26 April 2024

Revised: 5 June 2024

Accepted: 8 June 2024

Published: 14 June 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Leaves are the most sensitive part of a tree to changes in nutritional status and play a crucial role in synthesizing photosynthetic products in fruit trees. The mineral elements found in leaves act as a key indicator of the nutritional well-being of fruit trees. The presence or absence of essential mineral nutrients directly affects the photosynthetic capacity of leaves, thereby influencing the overall growth, development, and fruit production of fruit trees [1,2]. Traditional methods for evaluating the nutrient content of fragrant pear trees involve complex chemical procedures such as drying, grinding, and treating collected leaves with concentrated sulfuric acid and hydrogen peroxide. Although these methods provide relatively accurate results, they are time-consuming, labor-intensive, sample-destructive, and potentially hazardous to the operator's health [3,4]. Hence, the adoption

of near-infrared spectroscopy technology is essential for creating an efficient, precise, and non-destructive detection technique to monitor nutrient levels in fragrant pear leaves and anticipate nitrogen deficiency in fragrant pear trees.

Near-infrared spectroscopy (NIRS) is increasingly recognized for its objectivity, speed, and cost-effectiveness [5]. It is known for its ease of use, rapid results, and environmentally friendly nature, enabling the identification of multiple components within a single substance. NIRS is also valued for its high sensitivity, resolution, and scanning speed, making it a preferred tool for qualitative and quantitative analysis in various fields, including fruit and vegetable origin identification [6], component analysis [7], and assessment of pesticide and microbial contamination levels [8,9]. Researchers like Petit et al. [10] have successfully utilized near-infrared reflectance spectroscopy and arctic–alpine calibration models to evaluate the chemical properties of individual leaves. Their study demonstrated accurate measurement of nitrogen ($R^2 = 0.88$, RMSE = 0.824%), phosphorus ($R^2 = 0.65$, RMSE = 0.081%), carbon ($R^2 = 0.78$, RMSE = 2.199%), and silicon content in silicon-rich growth forms ($R^2 = 0.67$, RMSE = 0.677%). This highlights the potential of near-infrared spectroscopy for analyzing chemical properties in plant leaves. Therefore, this research aims to detect total nitrogen content in fragrant pear leaves using near-infrared spectroscopy in combination with chemical determination methods to create a predictive model.

In order to develop precise regression models, researchers commonly preprocess spectral data to remove background effects, compensate for additive or multiplicative influences [11], and reduce systematic errors [12]. Dyah and other scholars [13] utilized FTIR spectroscopy to identify pepper plants infected by pepper yellow leaf curl virus (PYLCV). They applied various preprocessing techniques such as baseline correction, normalization (standard normal variables, vectors, and min-max), and denoising (Savitzky–Golay smoothing, first and second order derivatives) before inputting the data into MLPNN, SVM, and LDA3 algorithms. The study found that the first-order derivative achieved the highest classification accuracy of 100% across all three models. Preprocessing spectral data did not result in significant loss by truncating the spectral range. Multivariate scattering correction (MSC) addresses scattering effects from uneven particle distribution and size, while standard normal variable transformation (SNV) mitigates solid particle size, surface scattering, and optical path variations in diffuse reflectance spectra. The first derivative (FD) and second derivative (SD) effectively remove baseline interference, enhancing resolution and sensitivity. The Savitsky–Golay smoothing (SG) method successfully eliminates noise. This study combines MSC with FD to reduce diffuse reflection and baseline influence, SG with SD to eliminate noise and baseline influence, and SNV with SD to reduce dispersion effects and baseline influence, resulting in three preprocessing combinations: MSC + FD, SG + SD, and SNV + SD. The study evaluates six preprocessing methods: MSC, FD, SD, MSC + FD, SG + SD, and SNV + SD, all of which enhance original spectral information and model accuracy.

Large spectral data can pose a significant challenge for models, potentially leading to reduced efficiency and the risk of overfitting. Zhao et al. [14] applied the competitive adaptive weighted sampling (CARS) algorithm to select canopy dust removal features and developed a canopy dust retention estimation model using random forest regression (RFR). Through the analysis of airborne hyperspectral images, a dust distribution map was created. The results indicated that the 2DCOS-CARS-RFR model outperformed other approaches, demonstrating higher accuracy rates (with RMSE and RPD values of 0.820 g/m², 3.910 g/m², and 2.357 in the R^2 validation set, respectively). Guo et al. [15] employed near-infrared transmission spectroscopy in combination with the partial least squares (PLS) algorithm to accurately identify the water core and soluble solid content of 663 ‘Fuji’ apple samples from Xinjiang (159), Shanxi (252), and Shandong Province (252). Various variable selection methods such as synergy interval (SI), successive projections algorithm (SPA), genetic algorithm (GA), and competitive adaptive reweighted sampling (CARS) were utilized to identify characteristic variables. The study illustrates that using the competitive adaptive reweighted sampling (CARS) variable selection method along with

the PLS model produces the most effective results in quantitatively determining the water core and soluble solids content of apple samples. Predictive correlation coefficient (R_p), root-mean-square error of prediction (RMSEP), and residual predictive deviation (RPD) were employed to evaluate the model performance. The CARS-PLS models exhibited superior prediction capabilities using spectra in the range of 600–1000 nm, with R_p , RMSEP, and RPD values of 0.9562, 1.340%, and 3.720 for apple water core degree; 0.9808, 0.327 oBx, and 4.845 for apple SSC. The study underscores the importance of extracting characteristic bands from spectral data, emphasizing the value of reducing data processing time.

Zhang et al. [16] employed terahertz spectral data in conjunction with radial basis function neural network (RBFNN) and back-propagation neural network (BPNN) algorithms to develop a model for detecting tomato nitrogen content based on terahertz spectra under nitrogen stress conditions. The researchers gathered 80 samples from 65-day-old tomato plants, selecting 20 leaf samples for each of the 4 nitrogen stress gradients. These samples underwent scanning using a terahertz time domain spectroscopy measurement system to acquire spectral information, with 10 sampling points scanned per sample and the data averaged for analysis. The BPNN model yielded corrected root-mean-square error (RMSEC) and predicted root-mean-square error (RMSEP) values of 0.1722% and 0.1843%, respectively, along with coefficients of determination for the calibration set (R_c^2) and prediction set (R_p^2) at 0.8447 and 0.8375. Conversely, the RBFNN model demonstrated RMSEC and RMSEP values of 0.1322% and 0.1855% respectively, with R_c^2 and R_p^2 values of 0.8714 and 0.8463. The RBFNN model exhibited slightly superior accuracy compared to the BPNN model. Despite the availability of various techniques for element detection in fruit trees using spectral data, the application of neural network algorithms for predictive models remains limited. This research offers valuable insights into modeling algorithms for detecting nitrogen content in fruit tree leaves.

This study investigates the total nitrogen content of Korla fragrant pear leaves using chemical methods for measurement and near-infrared spectra for data collection. The dataset included 180 leaf samples, each with 1557 wave numbers from near-infrared spectroscopy scans as the characteristic attribute and 1 total nitrogen content as the target attribute. Various preprocessing techniques were applied to the spectra, and the CARS algorithm was used to extract important wave number information. By combining four different prediction algorithms, the study established an optimal model for predicting total nitrogen content in fragrant pear leaves. This model enables rapid and non-destructive determination of nutritional elements in fragrant pear leaves, offering valuable insights for the scientific cultivation of fragrant pear trees.

2. Material and Methods

2.1. Materials and Equipment

A Fourier near-infrared spectrometer (WI, USA) was utilized to scan Korla fragrant pear leaves and collect leaf spectral information. The instrument had a resolution of 8 cm^{-1} and a gain of $2\times$. The instrument's built-in background served as a reference, and the average spectra was generated from 32 scans. Samples were sourced from Korla fragrant pears grown by the 15th Company of the 9th Regiment in Alar City, Xinjiang Production and Construction Corps (Urumqi, China). Three nitrogen application levels were established based on local fruit farmers' practices: 326.08 g/plant (N1), 652.17 g/plant (N2), 978.26 g/plant (N3); along with 7 fertilization periods as follows: T1 (bud germinating stage of fruit tree), T2 (young fruit expansion stage), T3 (rapid fruit expansion stage), T4 (bud germinating stage of fruit tree and young fruit expansion stages), T5 (bud germinating stage of fruit tree and rapid fruit expansion stages), T6 (young and rapid fruit expansion stages), T7 (bud germinating stage of fruit tree, young fruit expansion, and rapid fruit expansion stages). There are a total of 21 treatments, each consisting of 3 trees, resulting in a total of 63 trees. A total of 189 leaf samples were collected from these 63 trees. The pear trees in Liyuan were spaced at $1.5\text{ m} \times 4\text{ m}$, and were 6 years old. Mature leaves from the middle and lower sections of the current-year branches on the outer edges of the canopy of

each test tree were gathered during the fragrant pear fruit setting period (23 April 2022), fruit expansion period (11 July 2022), and fruit maturity period (10 September 2022). A single leaf from each cardinal direction (east, south, west, north) was carefully collected, labeled, and stored in a ziplock bag in a 4 °C freezer until spectral scanning and total nitrogen content analysis were conducted, as illustrated in Figure 1 for data acquisition.

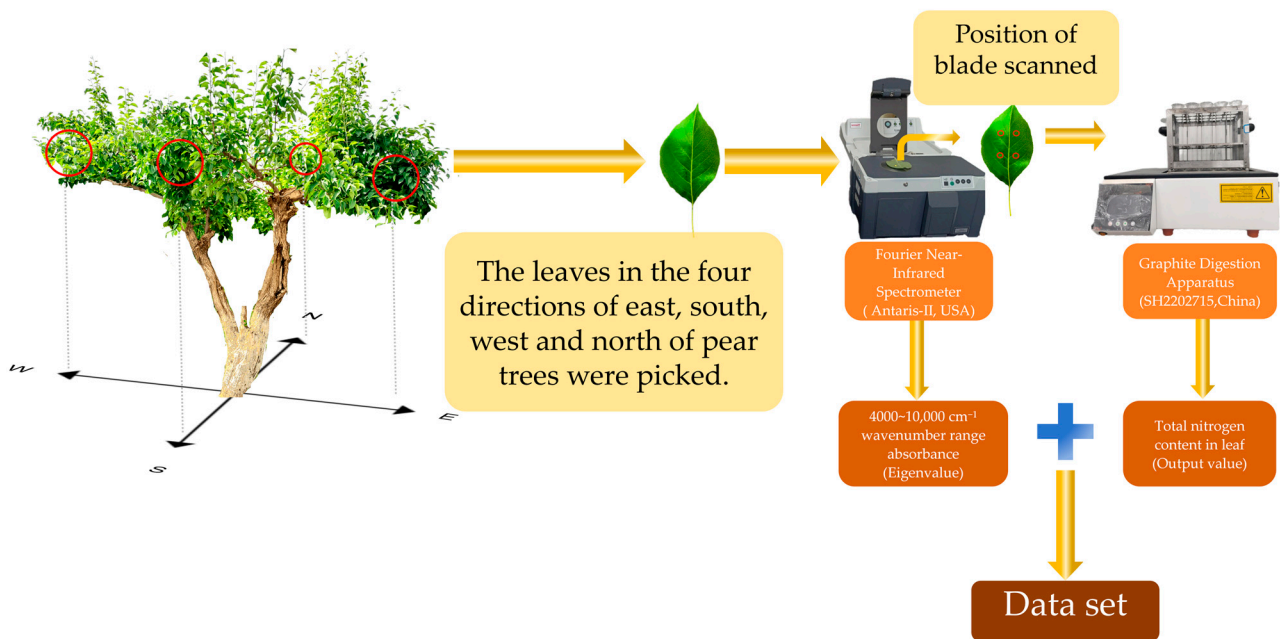


Figure 1. Data set acquisition flow chart. Note: The red circle on the tree is the sampling position, and the red circle on the leaf is the scanning spectral position.

2.2. Acquisition of the Original Spectra of Pear Leaves

Before collecting spectra, the test sample was removed from the 4 °C freezer and allowed to equilibrate in the room where the instrument was located for 12 h. The average storage and room temperature were both maintained at 24 °C to ensure consistent spectral collection conditions for all samples. Prior to collecting spectra, the instrument was powered on and allowed to warm up for 30 min before conducting diffuse reflection correction. The fragrant pear leaves were thoroughly cleaned and secured in place. When measuring the leaves, the leaf veins were used as reference boundaries. A total of 4 areas were selected on both the upper and lower ends of the leaves, assigned with different Arabic numerals and representing the spectral curves collected in each area with different colors for distinction. Each area was scanned 4 times, with an acquisition range of 4000–10,000 cm^{-1} , a resolution of 8 cm^{-1} , a gain of 2 $\times$, and 32 scans. This process yielded 16 spectral curves for a leaf sample, and the average of these curves was utilized as the final absorbance A value of the leaf sample. Subsequently, the data were exported for further analysis and processing to develop a prediction model.

2.3. Determination of Total Nitrogen in Pear Leaves

After collecting spectral data, the leaf samples were then sent back to the laboratory for cleaning in the prescribed order: tap water $\rightarrow$ 0.1% detergent solution $\rightarrow$ tap water $\rightarrow$ distilled water. The entire washing process did not exceed 2 min. Once washed, excess water on the leaf surface was quickly absorbed, followed by drying in a blast-drying oven at 105 °C for 20 min, then further drying at 80 °C until a constant weight was achieved. The dried samples were then be crushed using a stainless steel crusher, passed through a 60-mesh nylon sieve, and stored in a ziplock bag for determination of the total nitrogen content. Boiling the fragrant pear leaf sample with $\text{H}_2\text{SO}_4\text{-H}_2\text{O}_2$ yielded a test liquid. Ammonium in the test liquid was reacted with Nessler's reagent under alkaline conditions (pH = 11)

to form an orange complex according to the reaction formula. The reaction formula is as follows:



Substances that can cause turbidity in the measurement solution include ions such as Ca^{2+} , Mg^{2+} , Fe^{3+} , S^{2-} , ketones, and alcohols. In plant analysis, interference is primarily caused by Ca^{2+} and Mg^{2+} ions, and sodium tartrate can be added to mask these interferences. The digestion liquid was diluted to 100 mL with water, and the filtrate (or the supernatant that has been clarified) was taken for the determination of nitrogen (N) and other elements. In total 5 mL of the test solution mentioned above was taken and transferred to a 50 mL volumetric flask, where 2 mL of a 100 g/L sodium tartrate solution was added, mixed well, then potassium hydroxide (KOH) solution was added (amount determined by using phenolphthalein as an indicator to neutralize the acid in the solution). Water up to 40 mL was added, then 2.5 mL of Nessler's reagent, and diluted with water to the volume and mixed well. After 30 min, the color was measured using a spectrophotometer at a wavelength of 425 nm [17]. A blank test was conducted simultaneously with the sample measurement to correct for any reagent errors. To prepare the standard curve, 0.25, 5.00, 7.50, 10.00, and 12.50 mL of 10 $\mu\text{g}/\text{mL}$ nitrogen ($\text{NH}_4^+\text{-N}$) standard solutions were pipetted into six 50 mL volumetric flasks. The same color development steps were followed as for sample determination. The concentrations of this standard series were 0.5, 1.0, 1.5, and 2.0, 2.5 $\mu\text{g}/\text{mL}$ nitrogen ($\text{NH}_4^+\text{-N}$), respectively, and the colorimetry was performed at a wavelength of 420 nm. The zero point of the instrument was adjusted after using the blank digestion solution to develop color. The calculation is as follows:

$$\text{N}(\%) = \rho \cdot V \cdot t_s \times 10^{-4} / m \quad (3)$$

In the formula, ρ represents the mass concentration ($\mu\text{g} \cdot \text{mL}^{-1}$) of the chromogenic liquid N ($\text{NH}_4^+\text{-N}$) obtained from the standard curve; V represents the volume of the chromogenic liquid (mL); t_s represents the fractionation multiple, and the digestion liquid is constant. Volume (mL) divided by absorbed cooking liquid volume (mL); m represents sample mass (g).

2.4. Spectral Pretreatment Method

During the process of near-infrared spectra collection and sample chemical value analysis, external interference from instruments and man-made sources, such as high-frequency machine noise, light scattering, and test samples, can impact the precision and accuracy of spectral analysis and sample quality. Preprocessing of near-infrared spectra can help reduce background interference and address overlapping peaks [18], thus improving accuracy. In this experiment, six methods were used for preprocessing spectra: multivariate scattering correction (MSC), first derivative (FD), second derivative (SD), multivariate scattering correction + first derivative (MSC + FD), Savitzky–Golay + second derivative (SG + FD), and standard normal variable transformation (SNV) + second derivative (SNV + SD).

2.5. Modeling Methodology

The sample set partitioning based on joint x-y distance (SPXY) algorithm was used to split the dataset into a calibration set (training set) and a validation set (test set) at a ratio of 3:1. This allows the algorithmic model to learn from the available data, enabling it to predict or classify unknown data. Given the extensive amount of data in this experiment, despite preprocessing the spectral information and removing outliers, there was still a possibility of noise and outliers impacting the model's establishment. The random forest (RF) algorithm is known for its robustness in handling noisy data and outliers, although it may face challenges such as overfitting and local optimal solutions. To address these issues, the genetic algorithm-based random forest (GA-RF) algorithm and the extreme learning

machine (ELM) algorithm were employed in this study. The focus of this experiment was on detecting the total nitrogen content of plant leaves, a task that requires considering the growth and development of plants. Due to the limitations of ordinary linear functions in this context, the radial basis function (RBF) algorithm, known for its non-linear function approximation capabilities, was utilized. The four algorithms utilized in the training process were all classified under machine learning. The RF algorithm essentially belongs to the integration algorithm in machine learning, an algorithm based on the idea of bagging, which is essentially an integration algorithm; the ELM algorithm is a common machine learning algorithm, which strictly belongs to neural networks; the GA-RF algorithm belongs to optimization algorithms in machine learning; and the RBF algorithm belongs to deep learning algorithms in machine learning. The calibration set data were employed to train the algorithm model, while the validation set data were utilized to assess the accuracy of the algorithm model. A brief overview of the aforementioned machine learning methods is provided below:

Random Forest: (1) Randomly select samples and features: Each decision tree is trained using 135 randomly selected samples from the sample set, using the spectral information of these samples as input; (2) Build a decision tree: 100 decision trees are trained based on the selected samples and features. Each decision tree is constructed by recursively dividing the dataset; (3) Prediction: The test samples are inputted into all constructed decision trees, and the final output result is obtained through voting or averaging.

Genetic optimization of the random forest algorithm involves initializing parameters such as 100 decision trees, each with a maximum depth of 15 and a minimum number of 5 samples per node. The algorithm then uses a genetic algorithm to optimize these parameters. This process includes generating a population of individuals, evaluating them based on a fitness function, and selecting individuals with higher fitness. New individuals are created through crossover and mutation operations, with probabilities of 0.7 and 0.01, respectively, and added to the population. This cycle of evaluation, selection, and generation continues for 20 iterations. Finally, the optimized parameters are used to train the random forest model and make predictions on the test set.

Radial basis neural network algorithm: (1) Determine the structure of the RBF network, including the input layer, hidden layer and output layer. The hidden layer is set to 2 layers using the Gaussian radial basis function formula as follows, and the radial basis function expansion speed is set to 100. The formula is:

$$K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2) \quad (4)$$

Among them, x_i and x_j represent sample points, while γ is a parameter that determines the level of curvature of the function curve. A higher value of γ results in a steeper function curve, whereas a lower value of γ leads to a flatter function curve; (2) The parameters of the radial basis function are initialized, including the center determined through the K-means algorithm and the width selected through cross-validation; (3) The hidden layer output, representing the distance between the input sample and the radial basis function, is calculated using the Euclidean distance method. The formula is as follows:

$$\|x - c_j\| = \sqrt{(x - c_j)^T (x - c_j)} \quad (5)$$

The center vector of hidden layer neuron j , denoted as c_j , is calculated based on the input vector x and the specific hidden layer neuron j .

(4) Perform linear regression on the hidden layer output using the least squares method; (5) Update the network parameters, including the center and width of the radial basis function, as well as the weight and bias of the output layer, through the back propagation algorithm; (6) Repeat steps 3–5 until convergence conditions are met, train the model with the set parameters, and make predictions on the test set.

The extreme learning machine algorithm involves two main steps: (1). Randomly generating weights and biases between the input layer and hidden layer during initialization; (2). Inputting 135 sample sets into the network, and calculating the hidden layer output matrix H with 50 nodes. The Sigmoid activation function is used in this process. The formula is as follows:

$$f(x) = \frac{1}{1 + e^{-x}} \quad (6)$$

This function can effectively scale any real number to the $(0, 1)$ interval. As the input x increases, the output approaches 1, while as x decreases, the output approaches 0.

(3) The hidden layer output matrix H and label vector Y are inputted into the linear regression model to obtain the parameter vector LW , calculated as $LW = \text{pinv}(H') \times t_{\text{train}}$. Here, $\text{pinv}(H')$ represents the Moore–Penrose inverse of matrix H , and t_{train} is the target variable vector in the training dataset. Subsequently, the model is tested using the test set to calculate the test error.

2.6. Data Processing

Chemical values and spectral averages were calculated using Excel 2016, with the spectral information of 185 leaves being matched one-to-one with the chemical values. The total nitrogen content image was created using Origin 2021. MATLAB R2021b software was utilized to preprocess the spectral information (employing algorithms such as MSC, SNV, FD, SD, MSC + FD, SG + SD), extract feature bands (using the CARS algorithm), and construct models (using RF, GA-RF, RBF, ELM algorithms).

3. Results

3.1. Analysis of Total Nitrogen Content in Pear Leaves

Figure 2 illustrates the total nitrogen content of fragrant pear samples collected at various growth stages and nitrogen fertilizer levels. The findings reveal a range of total nitrogen content, spanning from 1.27% to 4.22%. The average nitrogen content for all samples was calculated at 2.64%, with a standard deviation of 0.82%. The observed variability in nitrogen content among samples receiving different nitrogen application levels offers valuable information for the development of models.

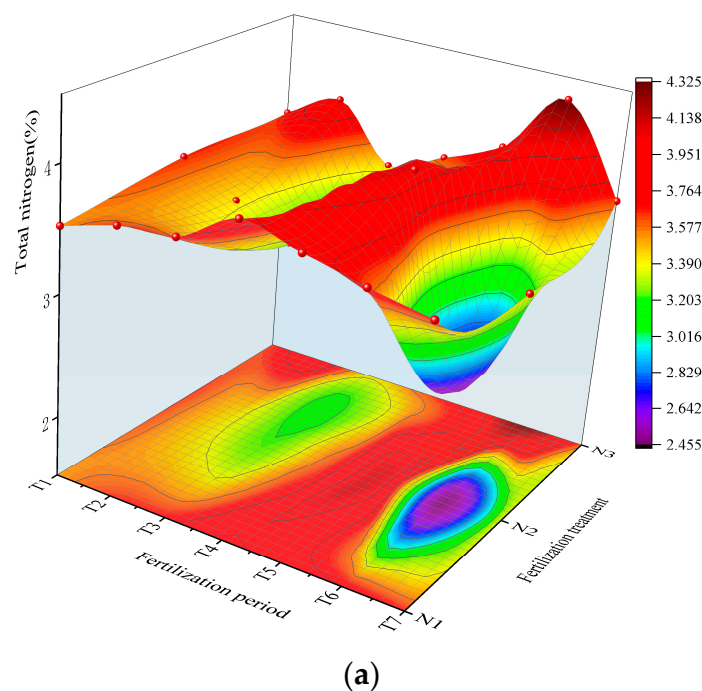


Figure 2. Cont.

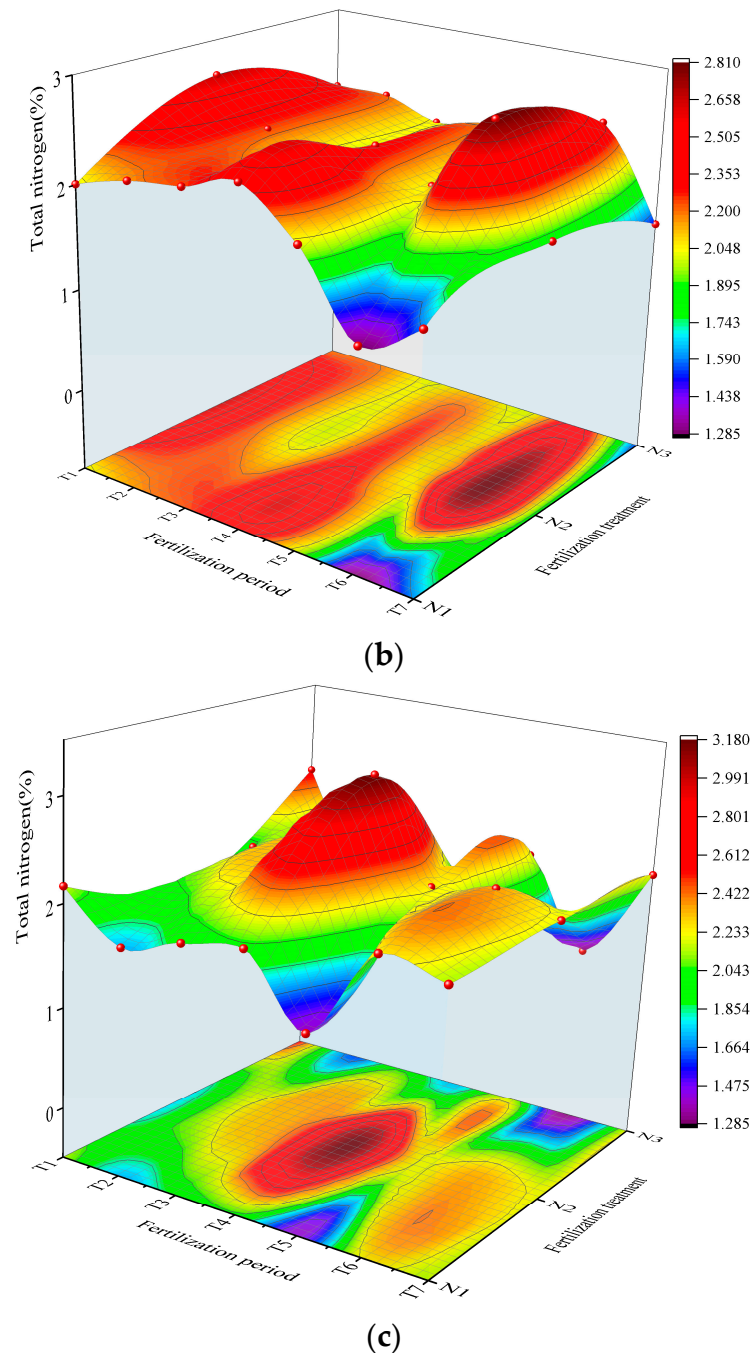


Figure 2. Leaf total nitrogen content at different growth stages under different nitrogen application levels, (a) is the leaf total nitrogen content sampled at the fruit setting stage; (b) is the total nitrogen content of leaves sampled during fruit expansion; (c) is the total nitrogen content of leaves sampled at fruit maturity stage.

3.2. Analysis of Near Infrared Spectroscopy of Pear Leaves

Figure 3 illustrates the full-band original spectra of a fragrant pear leaf sample using a near-infrared spectrometer. The test involved scanning the complete near-infrared spectra to validate the integrity and accuracy of the leaf spectra wavelength, spanning from 4000 to 10,000 cm^{-1} . The spectra exhibited clear absorption peaks, primarily linked to the vibration and absorption of O-H, N-H, C-H, and S-H bonds in molecules [19]. Prominent absorption peaks were detected at 7085.19 cm^{-1} and 8816.95 cm^{-1} , with a slight variation around the 8400.43 cm^{-1} region. A significant decline in the spectra was noted at 9542 cm^{-1} . Although the shapes of the original absorption spectra from different samples with varying nitrogen

application levels were similar, there were slight differences in absorbance levels for each spectral line, which could be advantageous for modeling purposes.

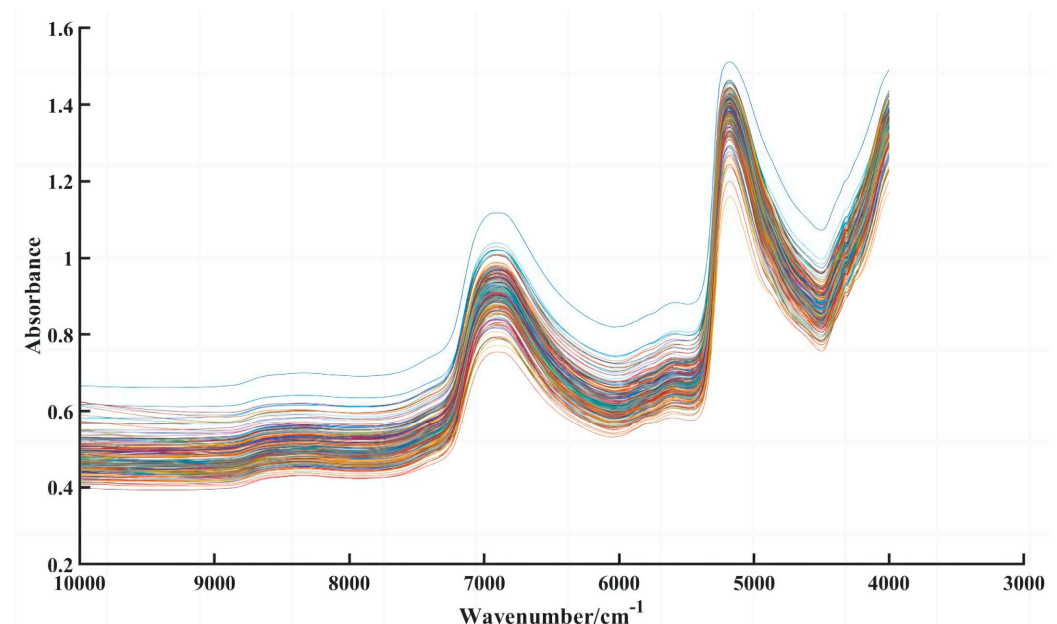


Figure 3. Primitive spectra of pear leaves.

3.3. Removal of Abnormal Samples

The spectra of fragrant pear leaves may contain abnormal components that could interfere with subsequent model predictions. To enhance the stability and accuracy of predicting the total nitrogen content in fragrant pear leaves, this study utilized the Mahalanobis distance method [20] to remove outliers during data processing. Initially, there were 185 sample data points in the experiment, and 5 outliers were identified and removed using the Mahalanobis distance method. Additionally, extreme values within the same type of data were also eliminated to improve the modeling effect, resulting in 180 sample data points.

$$MD_{(K)} = \sqrt{(X_K - \mu)^T \Sigma^{-1} (X_K - \mu)} \quad (7)$$

The formula for $MD_{(k)}$ involves calculating the Mahalanobis distance of various spectral curves in the k-band using the difference matrix X_k , the mean vector μ , and the inverse of the covariance matrix Σ^{-1} .

3.4. Division of Sample Set Results

In the experiment, the 180 remaining fragrant pear leaf samples from the previous chapters were utilized. The samples were partitioned into a calibration set and a validation set in a 3:1 ratio using the sample set partitioning based on joint xy distance (SPXY) method [21]. This ratio resulted in 135 samples in the calibration set (training set) and 45 samples in the validation set (test set). The fragrant pear leaf spectra were then categorized based on the SPXY algorithm. Analysis of Table 1 reveals that the maximum total nitrogen content of fragrant pear leaves in the calibration set and validation set were 4.22% and 3.87%, respectively, while the minimum values were 1.27% and 1.60%, respectively. The average total nitrogen content of fragrant pear leaves across all samples in this study was 2.64%, with averages of 2.63% and 2.45% for the calibration set and validation set, respectively. This distribution range was deemed suitable for modeling purposes, ensuring that the model remained free from overfitting.

Table 1. Sample total nitrogen content statistics and sample set division results.

Sample	Number of Sample	Maximum Value (%)	Minimum Value (%)	Mean Value (%)	Standard Deviation
Overall	180	4.22	1.27	2.64	0.82
Sample calibration set	135	4.22	1.27	2.63	0.83
Sample validation set	45	3.87	1.60	2.45	0.67

3.5. Comparison of Spectral Preprocessing and Extraction of Characteristic Wavelengths after Eliminating Outliers

3.5.1. Comparison of Spectral Preprocessing after Eliminating Outliers

In this study, the original spectral information was used to identify and eliminate abnormal data through the Mahalanobis distance method. Subsequently, six different combinations of preprocessing methods (MSC, FD, SD, MSC + FD, SG + SD, and SNV + SD) were employed to preprocess the spectra. Figure 4 illustrates the comparison of spectral images after applying the six preprocessing methods with Figure 3 representing the original spectral image. The results demonstrated that the spectral image following MSC preprocessing appeared smoother than the original, showing reduced noise bands. Additionally, the peak amplitudes after preprocessing were significantly higher than in the original spectra, highlighting a distinct absorption peak in the wave number range of 6000~7800 cm⁻¹. Specifically, a second frequency doubling related to the protein-CONH2 group was visible in the 5640~5880 cm⁻¹ region, while the 4220~5640 cm⁻¹ range exhibited a combined frequency associated with the protein -CH, -CH₂, and -CH₃ groups. Although this particular region was not clearly discernible in the original spectra, it became more pronounced in the images involving first-order derivative and second-order derivative preprocessing (FD, SD, MSC + FD, SG + SD, and SNV + SD). These observations suggest that preprocessed images enhance the analysis of spectral information.

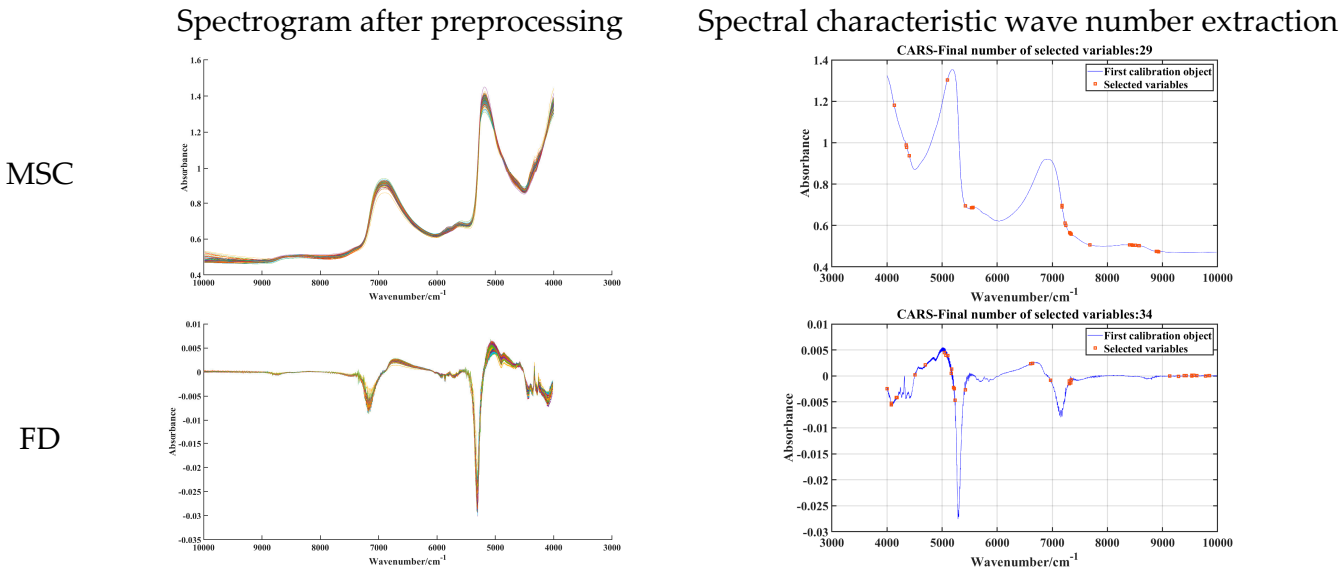


Figure 4. Cont.

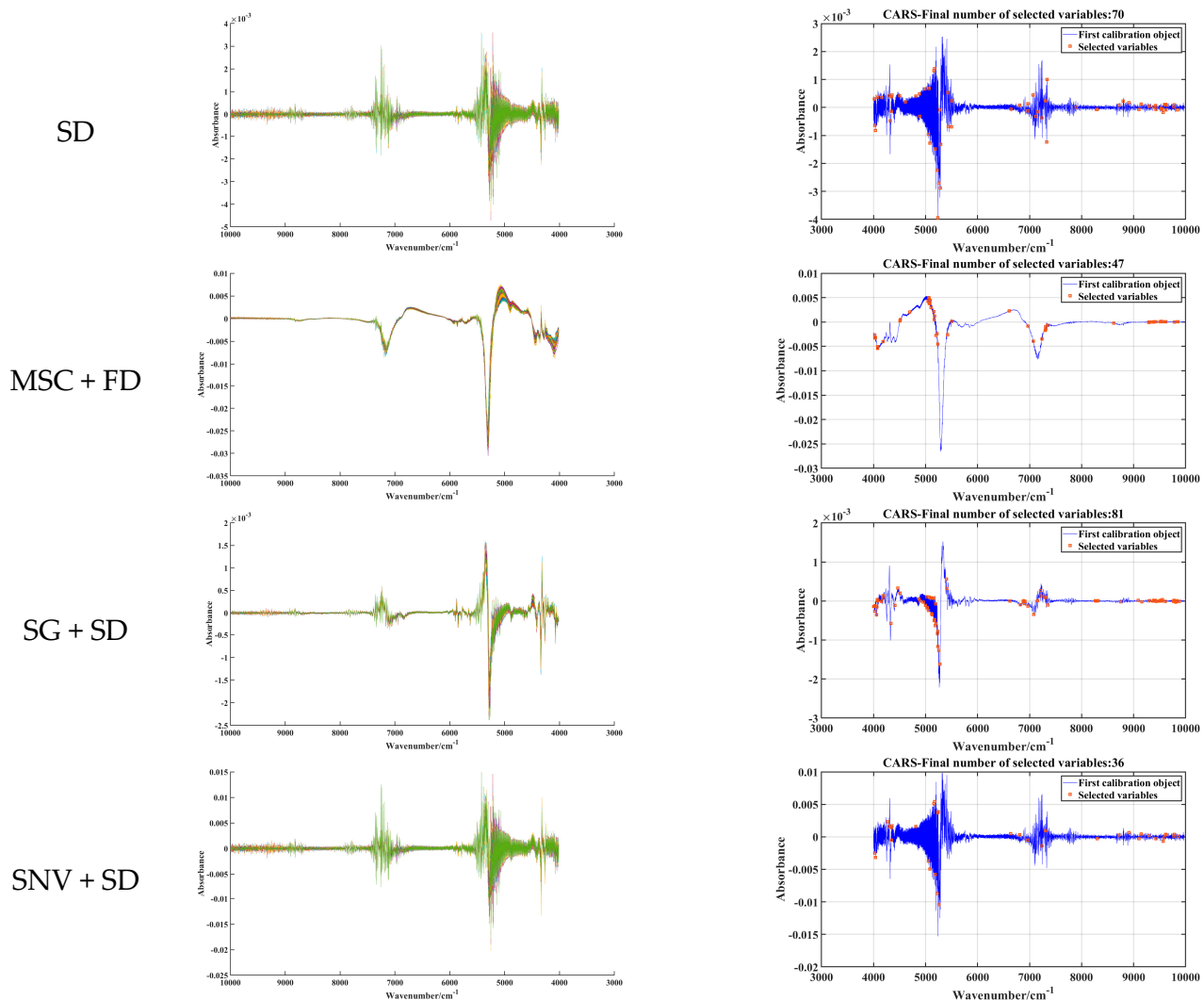


Figure 4. Six kinds of pre-treatment spectra and their characteristic wavenumber extraction positions.

3.5.2. Characteristic Wavelength Extraction

Competitive adaptive reweighted sampling (CARS) is an efficient algorithm utilized for spectral feature band selection. This method adjusts the selection probability of each band through Monte Carlo sampling and exponential decay functions to identify the optimal band combination that enhances modeling performance [22,23]. The results of characteristic wavelength selection using the CARS algorithm are depicted in Figure 5a–c, illustrating the changing trends of variables, root-mean-square error of interactive validation (RMSECV) value, and regression coefficient value, the presence of the star-pendant line in Figure 5c indicates that the root—mean—square error is minimized at that location, respectively. The experiment comprised 100 sampling times. Figure 5a shows a rapid decrease in modeling variables at the start of sampling, transitioning from rough to fine selection. Figure 5b demonstrates the relationship between sampling times and RMSECV, with the minimum value of 0.0878 observed at 73 sampling times. Figure 5c showcases the change in regression coefficient values for each variable at different sampling times, with the vertical line indicating the location of minimal root-mean-square error. The algorithm extracted 29 characteristic wave numbers from the 1557 wave numbers of the MSC preprocessing spectra, 34 from the FD preprocessing spectra, and 70 from the SD preprocessing spectra. Additionally, 47 characteristic wave numbers were extracted from the MSC + FD preprocessing spectra, 81 from the SG + SD preprocessing spectra, and 36

from the SNV + SD preprocessing spectra (Figure 4). Modeling with characteristic wave numbers can enhance model stability and reduce modeling time.

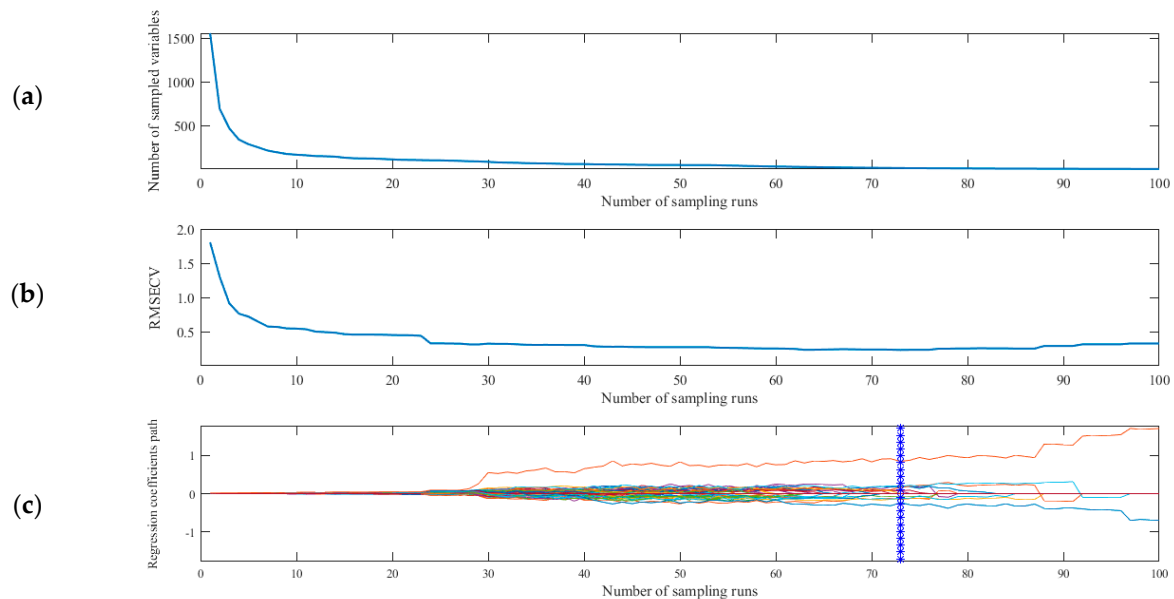


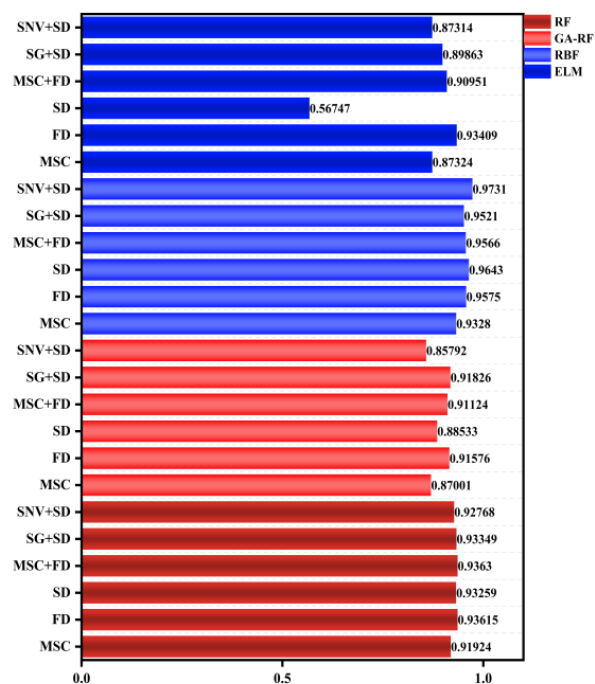
Figure 5. The principle of extracting variables by CARS operation. (a) The relationship between modeling variables and the number of shipments; (b) the relationship between sampling times and RMSECV; (c) the trend of regression coefficient value change.

3.6. Determination and Validation of the Model for the Total Nitrogen Content of Kuril Balsam Pear

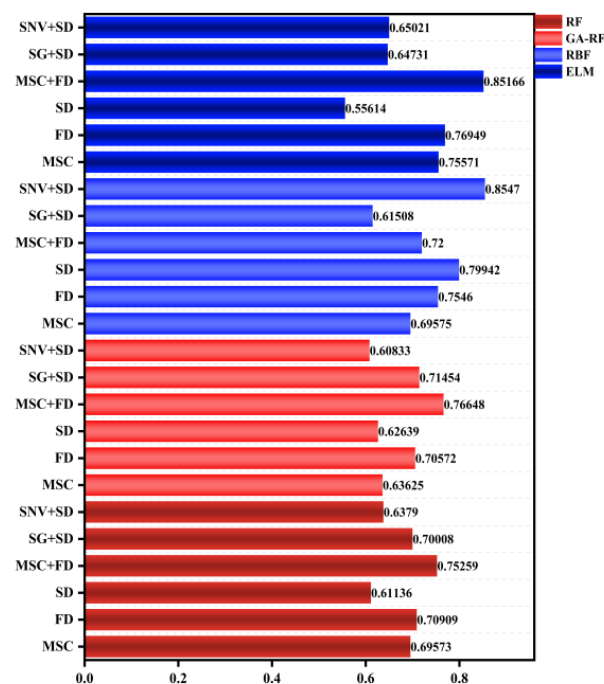
3.6.1. Determination of Combined Models

In order to determine the most appropriate prediction model for the nitrogen content of Korla fragrant pear leaves, this study utilized the random forest algorithm (RF), genetic algorithm-based random forest algorithm (GA-RF), radial basis neural network algorithm (RBF), and extreme learning machine algorithm (ELM). The spectral information was preprocessed using six algorithms (MSC, FD, SD, MSC + FD, SG + SD, SNV + SD) and combined in various ways with the four main algorithms. A total of 24 model combinations were tested, including MSC + CARS + RF, SNV + CARS + RF, FD + CARS + RF, SD + CARS + RF, MSC + FD + CARS + RF, SG + SD + CARS + RF, MSC + CARS + GA-RF, SNV + CARS + GA-RF, FD + CARS + GA-RF, SD + CARS + GA-RF, MSC + FD + CARS + GA-RF, SG + SD + CARS + GA-RF, MSC + CARS + RBF, SNV + CARS + RBF, FD + CARS + RBF, SD + CARS + RBF, MSC + FD + CARS + RBF, SG + SD + CARS + RBF, MSC + CARS + ELM, SNV + CARS + ELM, FD + CARS + ELM, SD + CARS + ELM, MSC + FD + CARS + ELM, SG + SD + CARS + ELM. The study utilized 135 calibration sets from the SPXY algorithm as the dataset for model establishment, with an additional 35 verification sets for model validation. The modeling and verification results are presented in Figure 5. During model establishment, only the R^2 values of the RF and RBF algorithms exceeded 0.9. However, upon comparison with the verification set, only the R^2 values of SNV + SD for the RBF algorithm and MSC + FD for the ELM algorithm surpassed 0.8 (Figure 6a,b). The R^2 value of SNV + SD for the RBF algorithm slightly outperformed MSC + FD for the ELM algorithm. The SNV + SD + CARS + RBF combination demonstrated the lowest RMSE at 0.0911%. No significant differences were observed in the validation set (Figure 6c,d). Figure 6e,f illustrate the residual prediction deviation (RPD) for the calibration and validation sets, with the SNV + SD + CARS + RBF combination yielding higher values compared to other algorithm combinations, reaching 4.5127 and 2.699, respectively. A higher R^2 and RPD generally indicate a smaller RMSE and better model prediction [24]. In conclusion, the RBF algorithm proves more suitable for predicting the nitrogen content of fragrant pear leaves.

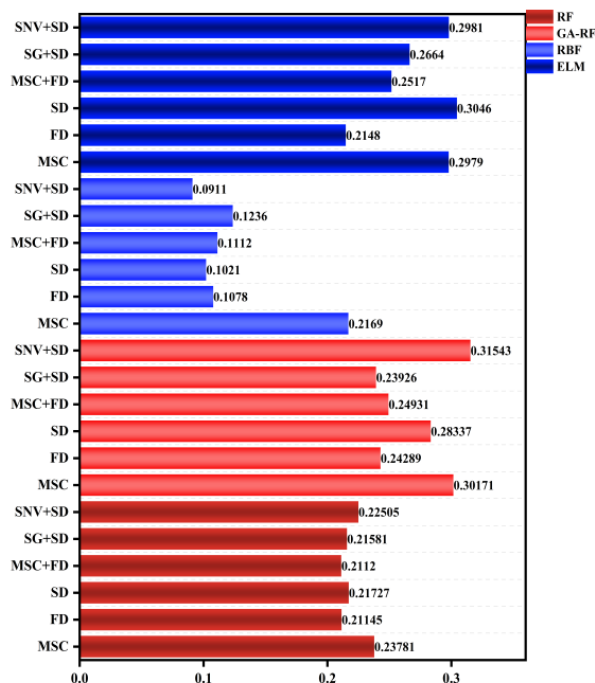
Thus, the SNV + SD + CARS + RBF algorithm combination serves as a more reasonable prediction model for the total nitrogen content of Korla fragrant pear leaves.



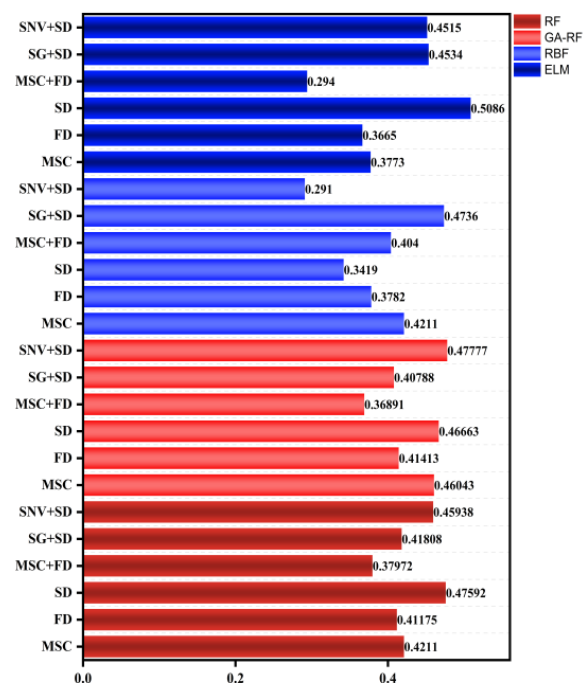
(a)



(b)

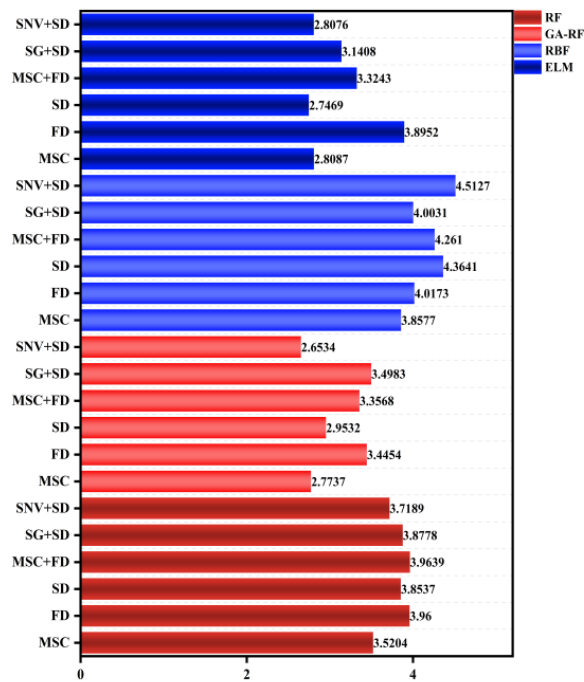


(c)

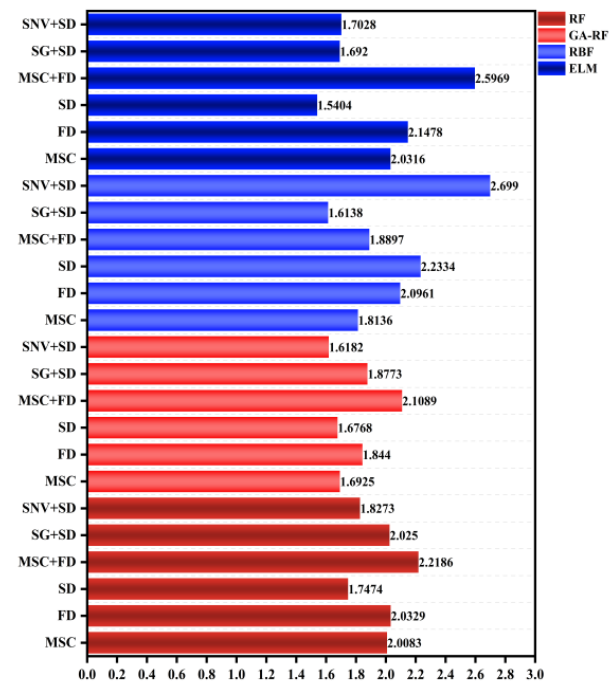


(d)

Figure 6. Cont.



(e)

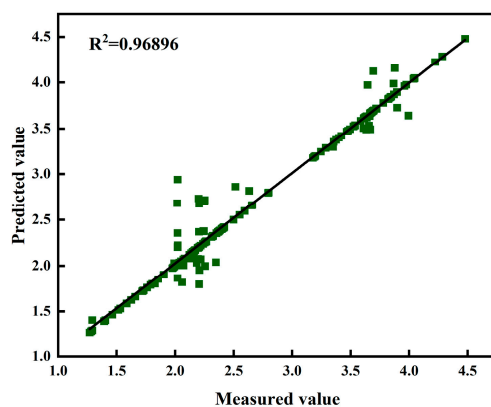


(f)

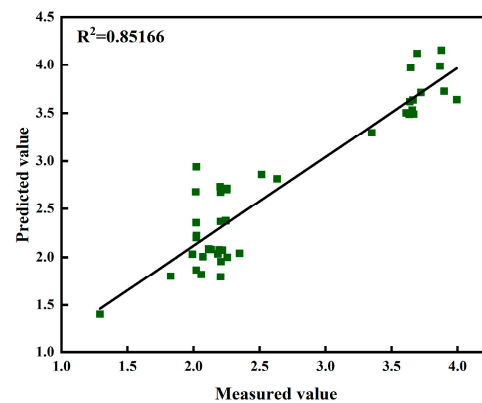
Figure 6. (a) Deterministic coefficients for each model calibration set; (b) deterministic coefficients for each model validation set; (c) root-mean-square error for each model calibration set; (d) Root-mean-square error for each model validation set; (e) Residual prediction bias for each model calibration set; (f) Residual prediction bias for each model validation set.

3.6.2. Validation of the Combined Model of SNV + SD + CARS + RBF Algorithm

The R^2 of the fit coefficient for measured and predicted values of all samples in the validation set of the model constructed using a combination of SNV, SD, CARS, and RBF algorithms was 0.96896 (Figure 7a). Additionally, the R^2 of the fit coefficient for measured and predicted values in the validation set was 0.85166 (Figure 7b), with an absolute deviation lower than 0.6 (Figure 7c). These results suggest that an R^2 value greater than 0.90 indicates excellent prediction accuracy, while values between 0.81 and 0.90 are considered good, 0.66 to 0.80 are also good, and 0.50 to 0.65 are deemed unsatisfactory [25]. Therefore, it can be concluded that the prediction model utilizing SNV + SD + CARS + RBF algorithms is effective for predicting nitrogen content in Kuril balsam pear leaves.



(a)



(b)

Figure 7. Cont.

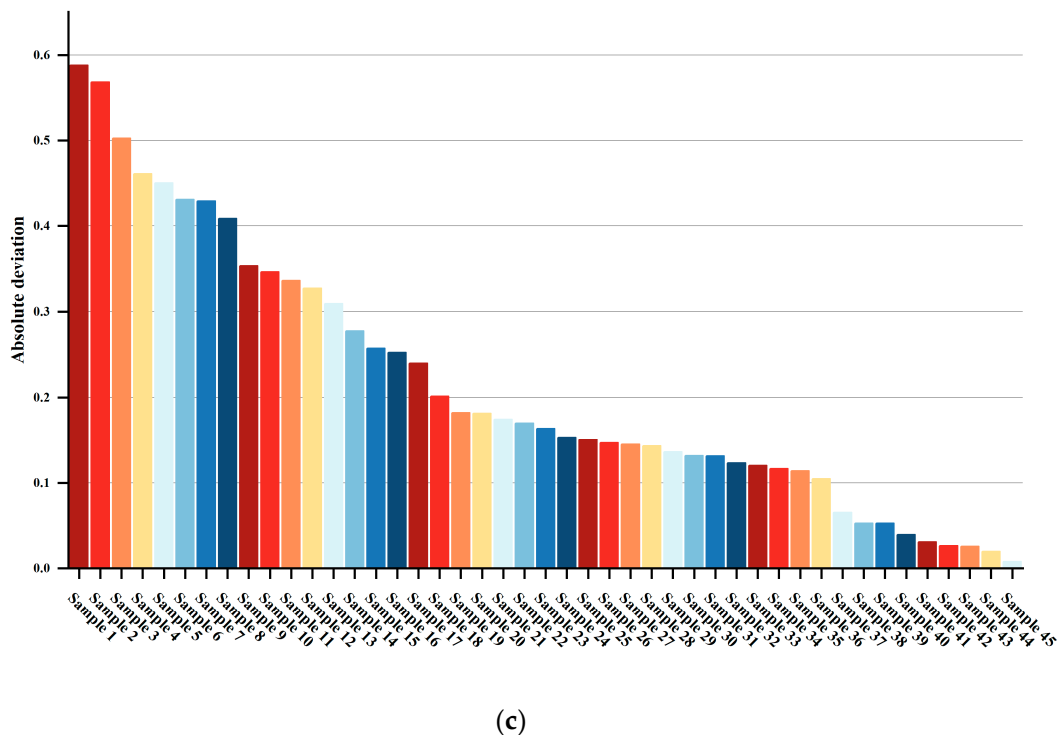


Figure 7. (a) The fitting results of the measured values and predicted values of all samples; (b) the fitting results of the measured values and predicted values of the validation set samples; (c) the absolute deviation between the measured value and the predicted value of the validation set sample.

4. Discussion

The search for a fast, non-destructive, and accurate detection method is crucial in diagnosing the nutritional status of fruit trees. Near-infrared non-destructive testing technology offers the benefits of speed and accuracy and is widely utilized in various fields such as agriculture and pharmaceuticals, including the examination of fruits like pears, walnuts, and apples [26–29]. Mishra et al. [30] employed Vis-NIR spectroscopy technology to assess nitrogen (N) and potassium (K) concentrations in bell pepper leaves. Their findings indicated that VisNIR spectroscopy showed promise in predicting N and K levels in pepper leaves, with root-mean-square errors (RMSEP) of 0.28% and 0.44%, respectively. By selecting specific wavelengths, the model's predictive performance for N and K is enhanced compared to PLS regression. Through wavelength selection, the RMSEP for N and K is reduced by 19% and 15%, respectively, in comparison to PLS regression. This study supports the development of protocols for non-destructive prediction of key plant chemical components like K and N without the need for wet chemical analysis. However, the study did not address the removal of outlier samples, which led to a decrease in the model's prediction performance. In a separate study by Lin et al. [31], hyperspectral reflectance of *Populus euphratica* and other trees in certain regions was investigated. The Mahalanobis distance method was utilized to identify significant differences between the original and transformed spectra of the trees, and a stepwise discrimination method was employed to evaluate the effectiveness of the selected difference bands. The results demonstrated that the Mahalanobis distance method serves as an efficient feature band extraction technique, enhancing the accuracy of the preprocessing method for recognition.

The spectral signal in the sample is often affected by stray light, noise, baseline drift, and other factors, which can significantly impact the final qualitative and quantitative analysis results. Therefore, it is crucial to mitigate the influence of these interference factors on the spectral signal before proceeding with modeling. To effectively address these issues, enhancing the signal-to-noise ratio, and establishing a more dependable model, preprocessing of the collected spectral information is necessary. Sonobe et al. [32]

utilized a combination of five preprocessing methods: first derivative reflectance (FDR), continuum-removed spectra (CR), de-trending (DT), multiplicative scatter correction (MSC), and standard normal variate (SNV) to reduce noise in wasabi leaf spectral data for vegetation characterization. The study demonstrated that preprocessing techniques can yield highly accurate estimates. When extracting spectral features, a genetic algorithm (GA) was employed, which may face challenges such as premature convergence and getting stuck in local optima [33]. To address this issue, parameters were set in this experiment so that the CARS algorithm ran 1000 times, allowing each wavelength to accumulate selected frequencies for feature extraction from the spectral information. Due to variations in the preprocessed spectral image, the characteristic wave number position of each process was altered to some degree. Therefore, in this experiment, each preprocessed spectral information was fed into the CARS algorithm to determine the characteristic wave number position.

Effective management of pear trees involves making decisions on the timing and rates of nitrogen application [34]. The nitrogen content in the leaves fluctuates as the fruit matures, with the lowest average value typically observed during the fruit maturity stage (Figure 2c). Pear trees, being woody plants, rely on two primary nitrogen sources for their nutritional needs: nitrogen reabsorption from storage and nitrogen uptake from the roots [35]. This study focuses on optimizing nitrogen application timing and quantity by local fruit farmers to enhance the growth and development of fragrant pear trees and fruits. It also emphasizes the importance of non-destructive nitrogen status diagnosis for timely fertilization to enhance fruit quality and yield. In contrast to previous research highlighting the limitations of leaf nitrogen determination through spectral measurement [35,36], this study introduces the RBF (the RBF algorithm, as the only feed-forward neural network, effectively overcomes the local minimum problem, is easy to train, and exhibits fast learning convergence speed). Research has demonstrated its capability to approximate any continuous nonlinear network with high precision [37] algorithms to develop a more accurate model. Among the 24 modeling methods tested, the model created by combining SNV + SD + CARS + RBF (the validation set R^2 was 0.8547, with RMSE = 0.291% and RPD = 2.699) outperformed other algorithm combinations, providing a reliable leaf total nitrogen detection model. This not only enhances work efficiency but also offers valuable insights for pear nitrogen supplementation. Future research could explore alternative algorithms like adaptive boosting and gradient boost decision tree, as well as different data partitioning methods to avoid overfitting.

5. Conclusions

In this study, we aimed to study whether the combination of near-infrared spectroscopy and chemometric methods can be used as a technology to detect the total nitrogen content of Korla fragrant pear leaves. In this study, we effectively utilized the CARS algorithm to extract spectral features from fragrant pear leaf samples, enhancing the model's detection capability while reducing computation time. The RBF algorithm effectively extracted features from the characteristic spectra, demonstrating strong performance in our study. The determination coefficient R^2 of the calibration set was 0.9731, and for the verification set, it was 0.8547. This method proved to be suitable for predicting the total nitrogen content of fragrant pear leaves through a predictive algorithm combination model of SNV + SD + CARS + RBF. Utilizing near-infrared spectroscopy technology for rapid data collection and the nonlinear function approximation capability of the RBF algorithm, we can now explore the development of a rapid detection system to accurately determine the total nitrogen content of fragrant pear trees at different stages of fruit growth. Near-infrared spectroscopy technology simplifies sample processing and efficiently gathers elemental information through spectra, while deep learning excels in interpreting spectral data. Our study offers insights for future applications in Korla fragrant pear production and tree cultivation. Moving forward, research should explore spectra linked to the growth of

Korla fragrant pear trees using different fertilizers, considering the constraints of spectra collection and fertilization in our current study.

Author Contributions: Z.W. and M.Y. contributed to this work by designing the study, obtaining data, performing statistical analyses, writing the manuscript and interpreting the data. J.B. contributed to the revision and review of the article. Z.T., X.B., Q.Z. and J.Z. participated in the conception and design of the study, interpreted the data, and reviewed and edited the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Bintuan science and technology program (2021AA005) and the President's Fund Innovation Team Project of Tarim University (TDZKCX202301).

Data Availability Statement: The original contributions presented in the study are included in the article, further inquiries can be directed to the corresponding author.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

MSC	Multiple Scattering Correction
FD	First Order Derivative
SD	Second Order Derivative
MSC + FD	Multiple Scattering Correction + First Order Derivatives
SG + SD	Savitzky–Golay + Second-Order Derivatives
SNV + SD	Standard Normal Variable Transformation + Second Order Derivative
SPXY	Sample set Partitioning based on joint X-Y distance
CARS	Competitive Adaptive Reweighted Sampling
RF	Random Forest
GA-RF	Genetic Algorithm-based Random Forest
RFB	Radial Basis Function Neural Network
ELM	Extreme Learning Machines
RMSECV	Root-Mean-Square Error of Cross-Validation
RMSE	Root-Mean-Square Error
RDP	Relative Percent Difference

References

1. Buwalda, J.G.; Meekings, J.S. Seasonal accumulation of mineral nutrients in leaves and fruit of Japanese pear (*Pyrus serotina* Rehd.). *Sci Hortic.* **1990**, *41*, 209–222. [\[CrossRef\]](#)
2. Wang, Z.H.; Liu, R.; Fu, L.; Tao, S.T.; Bao, J.P. Effects of Orchard Grass on Soil Fertility and Nutritional Status of Fruit Trees in Korla Fragrant Pear Orchard. *Horticulturae* **2023**, *9*, 903. [\[CrossRef\]](#)
3. Lu, X.Q.; Yuan, M.; He, N.; Zhao, S.; Zhu, H.; Liu, W. Rapid detection of lycopene content in watermelon by colorimeter. *Chin. Melon Veg.* **2021**, *34*, 41–45.
4. Ma, X.T.; Luo, H.P.; Gao, F.; Wang, C.X. Research and application of near infrared spectroscopy in apple detection. *J. Food Saf. Qual. Insp.* **2022**, *13*, 4219–4227.
5. Rahman, A.; Wang, S.; Yan, J.; Xu, H. Intact Macadamianut quality assessment using near-infrared spectroscopy and multivariate analysis. *J. Food Compos. Anal.* **2021**, *102*, 104033. [\[CrossRef\]](#)
6. Chen, B.; Zhang, M.; Chen, H.; Mujumdar, A.S.; Guo, Z. Progress in smart labels for rapid quality detection of fruit and vegetables: A review. *Postharvest Biol. Technol.* **2023**, *198*, 112261. [\[CrossRef\]](#)
7. Badaró, A.T.; Garcia-Martin, J.F.; del Carmen López-Barrera, M.; Barbin, D.F.; Alvarez-Mateos, P. Determination of pectin content in orange peels by near-infrared hyperspectral imaging. *Food Chem.* **2020**, *323*, 126861. [\[CrossRef\]](#)
8. Yu, G.W.; Ma, B.X.; Chen, J.C.; Dang, F.M.; Li, X.Z.; Li, C.; Wang, G. Vis/NIR spectra discriminant of pesticide residues on the Hamimel on surface by gadf and multi-scaleCNN. *Spectrosc. Spect. Anal.* **2022**, *41*, 3701–3707.
9. Sun, J.; Ge, X.; Wu, X.; Dai, C.; Yang, N. Identification of pesticide residues in lettuce leaves based on near infrared transmission spectroscopy. *J. Food Process Eng.* **2018**, *41*, e12816. [\[CrossRef\]](#)
10. Petit Bon, M.; Böhner, H.; Kaino, S.; Moe, T.; Brathen, K.A. One leaf for all: Chemical traits of single leaves measured at the leaf surface using near-infrared reflectance spectroscopy. *Methods Ecol. Evol.* **2020**, *11*, 1061–1071. [\[CrossRef\]](#)
11. Maleki, M.R.; Mouazen, A.M.; Ramon, H.; De Baerdemaeker, J. Multiplicative Scatter Correction during On-line Measurement with Near Infrared Spectroscopy. *Biosyst. Eng.* **2007**, *96*, 427–433. [\[CrossRef\]](#)
12. Tsai, F.; Philpot, W. Derivative Analysis of Hyperspectral Data. *Remote Sens. Environ.* **1998**, *66*, 41–51. [\[CrossRef\]](#)

13. Agustika, D.K.; Mercuriani, I.; Purnomo, C.W.; Hartono, S.; Triyana, K.; Iliescu, D.D.; Leeson, M.S. Fourier transform infrared spectrum pre-processing technique selection for detecting PYLCV-infected chilli plants. *Spectrochim. Acta Part. A Mol. Biomol. Spectrosc.* **2022**, *278*, 121339. [[CrossRef](#)] [[PubMed](#)]
14. Zhao, Y.; Lei, S.; Yang, X.; Gong, C.; Wang, C.; Cheng, W.; Li, H.; She, C. Study on Spectral Response and Estimation of Grassland Plants Dust Retention Based on Hyperspectral Data. *Remote Sens.* **2020**, *12*, 2019. [[CrossRef](#)]
15. Guo, Z.; Wang, M.; Agyekum, A.A.; Wu, J.; Chen, Q.; Zuo, M.; El-Seedi, H.R.; Tao, F.; Shi, J.; Ouyang, Q.; et al. Quantitative detection of apple watercore and soluble solids content by near infrared transmittance spectroscopy. *J. Food Eng.* **2020**, *279*, 109955. [[CrossRef](#)]
16. Zhang, X.; Duan, C.; Wang, Y.; Gao, H.; Hu, L.; Wang, X. Research on a nondestructive model for the detection of the nitrogen content of tomato. *Front. Plant Sci.* **2023**, *13*, 1093671. [[CrossRef](#)] [[PubMed](#)]
17. Bao, S.D. *Soil Agrochemical Analysis*, 3rd ed.; China Agriculture Press: Beijing, China, 2000; pp. 30–108.
18. Wang, J.J.; Liu, H.; Ren, G.X. Establishment of near-infrared model of pea quality traits and analysis of regional differences. *J. Plant Genet. Resour.* **2014**, *15*, 779–787+801.
19. Luo, W.; Huan, S.; Fu, H.; Wen, G.; Cheng, H.; Zhou, J.; Wu, H.; Shen, G.; Yu, R. Preliminary study on the application of near infrared spectroscopy and patternrecognition method stoclassify different types of apple samples. *Food Chem.* **2011**, *128*, 555–561. [[CrossRef](#)] [[PubMed](#)]
20. Li, Z.; Wang, J.Q.; Li, Z.Q.; Zou, D.Q.; Zhang, X.G.; Han, J.W.; Liu, W.Y. Hyperspectral monitoring and modeling of cotton canopy nitrogen. *J. Jiangxi Agric. Univ.* **2023**, *45*, 1273–1284.
21. Galvão, R.K.H.; Araujo, M.C.U.; José, G.E.; Pontes, M.J.C.; Silva, E.C.; Saldanha, T.C.B. A method for calibration and validation subset partitioning. *Talanta* **2005**, *67*, 736–740. [[CrossRef](#)]
22. Chen, H.; Lin, Z.; Tan, C. Fast quantitative detection of sesame oil adulteration by near-infrared spectroscopy and chemometric models. *Vib. Spectrosc.* **2018**, *99*, 178–183. [[CrossRef](#)]
23. Huan, K.W.; Chen, X.A.; Song, X.Y.; Dong, W. Variable selection in near-infrared spectra: Application to quantitative non-destructive determination of protein content in wheat. *Infrared Phys. Technol.* **2021**, *119*, 103937. [[CrossRef](#)]
24. Li, M.; Han, D.H.; Liu, W. Non-destructive measurement of soluble solids content of three melon cultivars using portable visible/near infrared spectroscopy. *Biosyst. Eng.* **2019**, *188*, 31–39. [[CrossRef](#)]
25. Wang, M. Study of Optimal Model for Nondestructive Detection of Kiwifruit, Peach and Pear Quality Characteristic by NIR Spectroscopy. Master's Thesis, Northwest A & F University, Yangling, China, 2013.
26. Liu, Y.D.; Liao, J.; Li, B.; Jiang, X.; Zhu, M.; Yao, J.; Wang, Q. Robustness of the global model of soluble solids in Gong pear based on near infrared spectroscopy. *Spectrosc. Spectr. Anal.* **2022**, *42*, 2781–2787.
27. Liu, Y.D.; Hu, X.; Zhu, M.W.; Yao, J.; Li, B.; Liao, J.; Zhang, J.; Jing, H. The influence of the parameters of the near-infrared on-line detection device on the applicability of the apple sugar model. *J. S. China Agric. Univ.* **2022**, *43*, 108–114.
28. Luo, L.Q.; Wang, T.; Liu, G.Q.; Zhao, W.G.; Zhang, R.; Yu, J.; Lu, B.; Chen, T.C. Detection model of soluble protein content in walnut kernel based on near infrared spectroscopy. *J. Fruit. Trees* **2023**, *40*, 1750–1761.
29. Moros, J.; Armenta, S.; Garrigues, S.; de la Guardia, M. Near infrared determination of Diuron in pesticide formulations. *Anal. Chim. Acta* **2005**, *543*, 124–129. [[CrossRef](#)]
30. Mishra, P.; Herrmann, I.; Angileri, M. Improved prediction of potassium and nitrogen in dried bell pepper leaves with visible and near-infrared spectroscopy utilising wavelength selection techniques. *Talanta* **2021**, *225*, 121971. [[CrossRef](#)] [[PubMed](#)]
31. Lin, H.J.; Zhang, H.F.; Gao, Y.Q.; Li, X.; Yang, F.; Zhou, Y.F. Mahalanobis Distance Based Hyperspectral Characteristic Discrimination of Leaves of Different Desert Tree Species. *Guang Pu Xue Yu Guang Pu Fen Xi* **2014**, *34*, 3358–3362.
32. Sonobe, R.; Yamashita, H.; Mihara, H.; Morita, A.; Ikka, T. Hyperspectral reflectance sensing for quantifying leaf chlorophyll content in wasabi leaves using spectral pre-processing techniques and machine learning algorithms. *Int. J. Remote Sens.* **2021**, *42*, 1311–1329. [[CrossRef](#)]
33. Li, J.; Lu, W.; Fan, Y. Groundwater Pollution Sources Identification Based on Hybrid Homotopy-Genetic Algorithm and Simulation Optimization. *Environ. Eng. Sci.* **2021**, *38*, 777–788. [[CrossRef](#)]
34. Wang, J.; Shen, C.; Liu, N.; Jin, X.; Fan, X.; Dong, C.; Xu, Y. Non-Destructive Evaluation of the Leaf Nitrogen Concentration by In-Field Visible/Near-Infrared Spectroscopy in Pear Orchards. *Sensors* **2017**, *17*, 538. [[CrossRef](#)] [[PubMed](#)]
35. Ulissi, V.; Antonucci, F.; Benincasa, P. Nitrogen content estimation in tomato leaves by VIS-NIR non-destructive spectroscopy. *Sensors* **2011**, *11*, 6411–6424. [[CrossRef](#)] [[PubMed](#)]
36. Graeff, S.; Claupein, W. Quantifying nitrogen status of corn (*Zea mays* L.) in the field by reflectance measurements. *Eur. J. Agron.* **2003**, *19*, 611–618. [[CrossRef](#)]
37. Hu, Y.; You, J.J.; Liu, J.N.K.; He, T. An eigenvector based center selection for fast training scheme of RBFNN. *Inf. Sci.* **2018**, *428*, 62–75. [[CrossRef](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.