


Article

Spectral Quantitative Analysis and Research of Fusarium Head Blight Infection Degree in Wheat Canopy Visible Areas

Yanyu Chen ¹, Xiaochan Wang ^{1,*} , Xiaolei Zhang ¹, Ye Sun ², Haiyan Sun ³, Dezhi Wang ¹ and Xin Xu ¹

¹ College of Engineering, Nanjing Agricultural University, Nanjing 210031, China

² College of Food Science and Light Industry, Nanjing Tech University, Nanjing 210009, China

³ Institute of Plant Protection, Jiangsu Academy of Agricultural Sciences, Nanjing 210014, China

* Correspondence: wangxiaochan@njau.edu.cn

Abstract: Obtaining complete and consistent spectral images of wheat ears in the visible areas of in situ wheat canopies poses a significant challenge due to the varying growth posture of wheat. Nevertheless, detecting the presence and degree of wheat Fusarium head blight (FHB) in situ is critical for formulating measures that ensure stable grain production and supply while promoting green development in agriculture. In this study, a spectral quantitative analysis model was developed to evaluate the infection degree of FHB in an in situ wheat canopy's visible areas. To achieve this, a spectral acquisition method was used to evaluate the infection degree of FHB in a wheat canopy's visible areas. Hyperspectral images were utilized to obtain spectral data from healthy and mildly, moderately, and severely infected wheat ear canopies. The spectral data were preprocessed, and characteristic wavelengths were extracted using twelve types of spectral preprocessing methods and four types of characteristic wavelength extraction methods. Subsequently, sixty-five spectral quantitative prediction models for the infection degree of FHB in the in situ wheat canopy visible areas were established using the PLSR method, based on the original spectral data, preprocessed spectral data, original spectral characteristic wavelengths extracted data, and preprocessed spectral characteristic wavelengths extracted data. Comparative analysis of the models indicated that the MMS + CARS + PLSR model exhibited the best prediction effect and could serve as the spectral quantitative analysis model for the evaluation of the infection degree of FHB in an in situ wheat canopy's visible areas. The model extracted thirty-five characteristic wavelengths, with a modeling set coefficient of determination (R^2) of 0.9490 and a root-mean-square error (RMSE) of 0.2384. The testing set of the coefficient of determination (R^2) was 0.9312, with a root-mean-square error (RMSE) of 0.2588. The model can facilitate the spectral quantitative analysis of the infection degree of FHB in the in situ wheat canopy visible areas, thereby aiding in the implementation of China's targeted poverty alleviation and agricultural power strategy.

Keywords: in situ wheat canopy; fusarium head blight; spectrum; infection degree; visible areas



Citation: Chen, Y.; Wang, X.; Zhang, X.; Sun, Y.; Sun, H.; Wang, D.; Xu, X. Spectral Quantitative Analysis and Research of Fusarium Head Blight Infection Degree in Wheat Canopy Visible Areas. *Agronomy* **2023**, *13*, 933. <https://doi.org/10.3390/agronomy13030933>

Academic Editor: Mario Cunha

Received: 6 March 2023

Revised: 15 March 2023

Accepted: 20 March 2023

Published: 21 March 2023



Copyright: © 2023 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

Wheat is a crucial food crop, providing sustenance for approximately 35~40% of the world's population [1,2]. China, as the largest wheat-producing country globally, boasts perennial planting areas and total output that accounts for over 20% of the country's grain crops [3,4]. Wheat Fusarium head blight (FHB), an irreversible disease mainly caused by Fusarium species, such as *Fusarium graminearum* Schwabe [5], significantly impacts crop yields. The disease may cause a reduction of 10~20% in yield, while severe cases may lead to a reduction of 80~90% or even no production at all [6,7]. However, obtaining accurate information on wheat FHB in situ is challenging due to the complexity and diversity of the Chinese farmland environment.

Currently, there exist several methods for detecting wheat FHB, including artificial field detection [8], biological detection [9], machine vision detection [10,11], and spectral

imaging detection [12–14]. However, artificial field detection and biological detection methods are often time-consuming, laborious, and inefficient. Machine vision detection techniques can only obtain superficial image information [15]. In contrast, spectral imaging detection methods can provide not only superficial information but also internal information, making it an effective tool for disease detection [16–19].

Several studies have been conducted to identify and diagnose FHB in wheat using spectral data. Huang employed Fisher linear discriminant analysis (FLDA) and support vector machine (SVM) to establish disease severity identification models based on spectral data obtained from ex situ wheat ears under a black background [20]. PSO-SVM algorithm was also used to fuse spectral, texture, and color features to build a diagnosis model for FHB infection degree [21], achieving a validation set accuracy of 92%. Additionally, a PSO-SVM FHB detection model was developed by combining wavelet features, achieving an overall accuracy of 93.5% [22]. Zhang collected spectral data from infected wheat ears at non-original locations and developed a DCNN FHB recognition model based on fusion features, which achieved an R^2 of 0.97 [23]. Moreover, a new FHB index (FDI) was established based on spectral information ranging from 384 to 1050 nm, which showed a better ability to distinguish the FHB infection degree than the other 16 existing spectral indices [24]. Zhang also obtained hyperspectral microscopic image data of infected ex situ wheat ears, and established an FHB classification index, achieving good results [25]. Jin used black cloth to obtain pure hyperspectral images of wheat ears and constructed a 2D-CNN-BidGRU hybrid neural network model to identify and distinguish healthy and infected wheat ears [26]. Furthermore, Jin collected hyperspectral images of wheat ears under the same conditions and used a convolutional neural network to classify and model the pixel features of FHB, achieving a rapid determination of the disease in hyperspectral images [27]. Bauriegel collected spectral data under a black background and conducted early identification of wheat FHB [28]. Liu Linyi analyzed the spectral characteristics SIPI and TVI with the highest sensitivity to canopy FHB, using the least-square method to construct a canopy wheat FHB monitoring model from ex situ spectral data obtained under a black background, achieving better results [29]. Han also obtained spectral data from ex situ wheat ears under a black background, achieving a recognition accuracy of 86.2% for the testing set of the model constructed using the MSC + XLW + SVM algorithm [30]. It is worth noting that ex situ spectral information can avoid interference from other factors and obtain better modeling results than in situ spectral information, which is affected by the soil background, crop varieties, canopy structure, and other factors that require further exploration and improvement. Żelazny collected spectral data from in situ wheat and explored the effect of the sequential support vector machine algorithm on FHB identification [31]. Whetton also obtained in situ spectral data in the laboratory and the field, and preliminary FHB detection was realized through the combination of image and spectrum analysis, supporting follow-up research [32,33].

Limited studies have been conducted on the in situ spectral imaging detection of FHB in wheat, with most research focusing on ex situ detection. This study aims to address this gap by exploring the detection of FHB in the in situ wheat canopy using hyperspectral imaging. In this process, the wheat ear growth posture varies significantly due to the influence of ear weight. In the horizontal plane, the central axis of wheat ears is seldom parallel to the plumb line, and most ears exhibit a certain inclination angle. As such, obtaining hyperspectral images of wheat ears with consistent integrity poses a challenge. Statistical analysis shows that the wheat ear inclination angle fluctuates within a range of $-15^\circ \sim 15^\circ$, but the spectral fluctuation within this range is minor, enabling the true reflection of wheat FHB to some extent. By aligning the hyperspectral camera lens plane parallel to the horizontal plane, only hyperspectral information from the visible areas of wheat ears in the canopy can be obtained. Research indicates that the number of wheat spikelets in the visible areas of each wheat ear accounts for approximately 50% of the total spikelets. Given the irreversible nature of FHB disease, the early detection and prompt implementation of corresponding measures are critical for minimizing crop losses.

Therefore, obtaining spectral information from 50% of wheat ears in the visible areas of the canopy can reflect and predict the degree and trend of wheat FHB infection to a certain extent. This information can provide a valuable reference for the formulation of measures aimed at promoting stable grain production and supply, encouraging green development in agriculture, and facilitating the implementation of the precise poverty alleviation and agricultural power strategies of the country.

2. Materials and Methods

2.1. Experimental Design

The experiment was conducted at the outdoor experimental area of the College of Engineering, Nanjing Agricultural University, spanning from September 2021 to June 2022. The test sample selected was Huaimai 33, the leading wheat variety in Jiangsu Province [34]. This wheat variety is susceptible to FHB, thus, making it conducive to the infection of *Fusarium graminearum* Schwabe and the outbreak of FHB in the later stages. The test soil was chosen to have a constant nutrient distribution ratio and sterilization. Hole sowing in the soil trough was utilized as the cultivation method. Four groups of soil troughs measuring 5000×400 mm were set up, with two groups in the control and experimental groups. Each hole was planted with three to five seeds, and the hole spacing was set at 100 mm. After sowing, the soil was covered with a film to preserve moisture. Watering was regularly performed to ensure the healthy and balanced growth of wheat. The *Fusarium graminearum* Schwabe spore suspension was evenly sprayed on the wheat ear when the wheat was flowering by more than 50%. The wheat ear was then moisturized for 48 h to ensure the outbreak of FHB [35]. Spectral data were obtained and analyzed seven days after the inoculation of *Fusarium graminearum* Schwabe from 10:00 to 14:00 under sunny, windless, and cloudless weather conditions with slight fluctuations in light intensity [36]. The experimental procedures are illustrated in Figure 1.

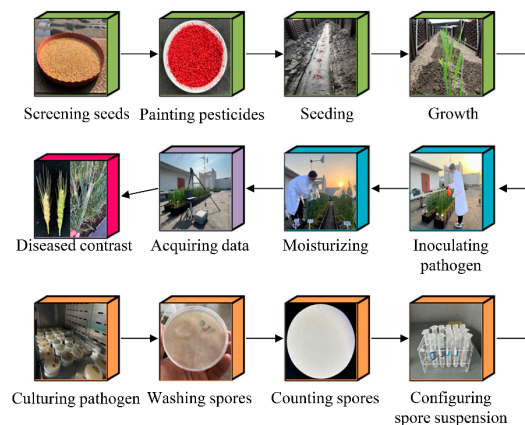


Figure 1. Schematic of experiment procedures.

2.2. Data Acquisition

2.2.1. Spectral Data Acquisition

The in situ hyperspectral imaging acquisition system comprised a SOC710-VP portable hyperspectral imager (manufactured by Surface Optics Company in the San Diego, CA, USA), a tripod, a power supply, a computer with hyperspectral data acquisition software (Hyper Scanner), a test bench, soil troughs, and wheat samples, as shown in Figure 2. This system can collect wavelengths ranging from 375 to 1050 nm, with a spectral resolution of 2.10 nm and 128 wavelengths. The system can capture hyperspectral images of the wheat canopy visible areas, where the wheat ear inclination angle ranges from -15° to 15° .

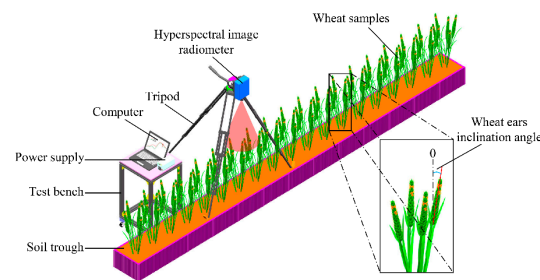


Figure 2. In situ hyperspectral imaging acquisition system.

During the in situ hyperspectral imaging acquisition process, noise was inevitably generated by the relevant components of the system and environmental fluctuations. To minimize the effect of noise on the spectral data, the Spectral Radiance Analysis Toolkit was utilized to correct the original data using a gray plate [37]. This enabled the acquisition of corrected hyperspectral images of the wheat canopy visible areas, which were further analyzed in subsequent steps.

RGB images are created by synthesizing red light (approximately 620~750 nm), green light (approximately 495~570 nm), and blue light (approximately 450~495 nm), and do not correspond to a specific wavelength range. In order to better visualize and differentiate between healthy and infected wheat ear, as shown in Figure 3a, the present study employed ENVI 5.6 software to synthesize RGB images using three wavelengths, namely, 648.68 nm, 549.33 nm, and 451.61 nm, from hyperspectral images. The Build 3D Cube tool was utilized to visualize the overall three-dimensional hyperspectral imaging data, and the resulting 3D maps effectively represented the healthy and infected wheat ears. The selection of region of interest (ROI) was accomplished using the ROI Tool region of interest selection tool, which extracted circular ROIs with a dimension of 5×5 pixels for healthy and infected wheat ears. The average reflection spectrum of the pixels within the ROIs was calculated to generate the reflection spectrum curve of the healthy and infected wheat ears [38]. Figure 3b depicts the reflection spectrum curves obtained through the aforementioned process.

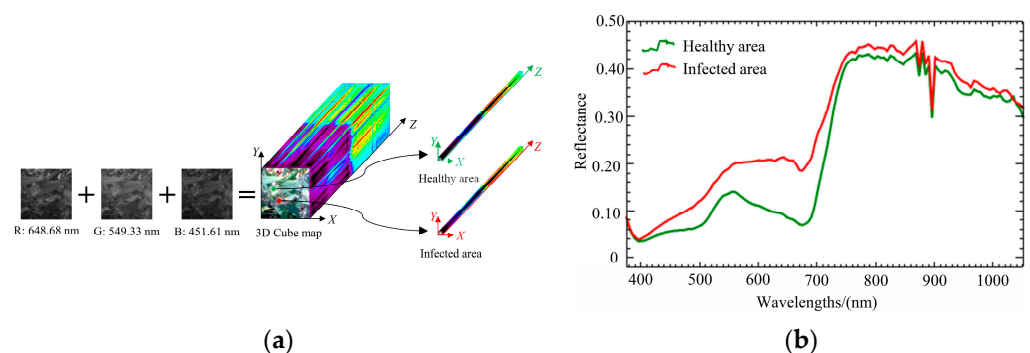


Figure 3. Schematic of spectral data acquisition. (a) Region of interest division and spectral data extraction; (b) average reflectance spectral curves of healthy and infected area of wheat ears.

The spectral data obtained in this study were subject to variability due to equipment limitations, environmental factors, and manual selection of the region of interest. As a result, spectral data with large differential wavelengths and significant noise at both ends (375~400 nm and 1000~1050 nm) were eliminated manually to ensure the reliability of the data. Ultimately, 451 sets of spectral data with 114 wavelengths within the range of 400~1000 nm were obtained as the sample data for subsequent processing and analysis.

2.2.2. Infection Degree Determination

During the collection of hyperspectral images of the wheat canopy visible areas, an RGB camera was used to capture clear RGB images simultaneously. The obtained RGB

images were utilized to determine the degree of FHB infection in wheat based on the ratio of infected wheat spikelets to the total number of spikelets within the wheat canopy's visible areas [39]. Figure 4 illustrates the determination of the FHB infection degree through the aforementioned process.

According to the Chinese national standard of reference [39], the calculation formula of the infection degree of wheat FHB in the canopy's visible areas is as follows:

$$D_{HBI} = \frac{N_{HBI}}{N_{ALL}} \times 100\%, \quad (1)$$

where D_{HBI} represents the infection degree (%) of wheat FHB in the canopy visible areas; N_{HBI} represents the number of infected wheat spikelets in the canopy visible areas; and N_{ALL} represents the total number of infected wheat spikelets in the canopy visible areas.

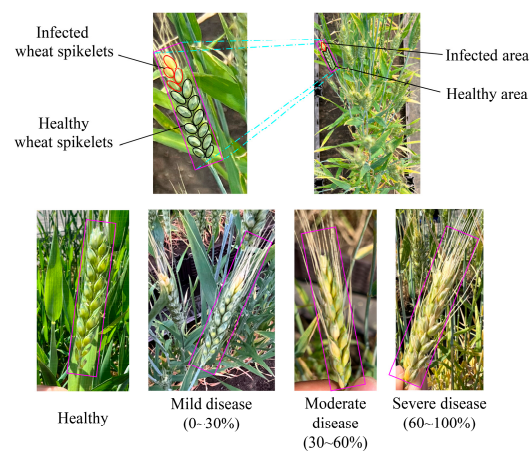


Figure 4. Schematic of wheat infection degree.

2.3. Data Processing and Modeling Evaluation Method

2.3.1. Preprocessing Spectral Data

The spectral data collected from in situ visible areas of wheat infected with FHB contained a significant amount of noise and other external interferences due to equipment noise and environmental factors. Thus, preprocessing the spectral data was necessary to remove irrelevant information [40].

In this study, to investigate the effects of different preprocessing methods on spectral data, twelve conventional spectral preprocessing methods were employed [41–43]. These methods included first derivative (FD), second derivative (SD), detrend correction (DC), continuous wavelet transform (CWT), multiplicative scatter correction (MSC), standard normal variate (SNV), moving average smoothing (MAS), Savitzky–Golay convolution smoothing (SGCS), mean centering (MC), standard scaling (SS), max–min scaling (MMS), and vector normalization (VN). The 451 sets of spectral data obtained comprised 114 wavelengths with a range of 400~1000 nm.

2.3.2. Characteristic Wavelengths Extraction

Hyperspectral data contain a wealth of information; however, they also suffer from problems of information redundancy and multicollinearity, which can reduce modeling efficiency and accuracy while increasing computational complexity [44,45]. To address these issues, various characteristic wavelength extraction methods were employed in this study, including the successive projections algorithm (SPA), competitive adaptive reweighted sampling (CARS), genetic algorithm (GA), and uninformative variables elimination (UVE) [46,47]. These methods were used to reduce the dimensionality of original and preprocessed spectral data, thereby simplifying the data and improving the efficiency of modeling calculations.

The successive projections algorithm (SPA) is a forward characteristic variable selection method [48,49]. This algorithm projects the wavelength onto other wavelengths, and then compares the projection vector size. The wavelength with the most significant projection vector is selected as the wavelength to be used. Finally, based on the correction model, the final characteristic wavelength is selected. Using SPA, the wavelength combination with the least redundant information and the least collinearity can filter out redundant and invalid information, reduce the collinearity between variables, and improve modeling efficiency [50].

Competitive adaptive reweighted sampling (CARS) is a method for selecting characteristic variables that combines Monte Carlo sampling and PLS model regression coefficients [51,52]. It employs the exponentially decreasing function (EDF) and adaptive reweighted sampling (ARS) to select individuals with larger absolute weights of regression coefficients in the PLSR model as a subset of wavelength variables each time. By conducting multiple calculations, multiple wavelength variable subsets are obtained. The characteristic wavelength is selected from the subset with the smallest root-mean-square error of cross-validation (RMSECV) of the PLSR model.

The genetic algorithm (GA) is a metaheuristic search method that mimics the principles of natural evolution and genetics [53–55]. Its core concept of “survival of the fittest” emulates the biological phenomenon of natural selection. The GA algorithm is implemented by iteratively optimizing the population’s individual structure through selection, crossover, and mutation processes based on genetic mechanisms. In this process, unfavorable variables are eliminated, and favorable ones are preserved. An appropriate fitness function evaluates the individual’s advantages and disadvantages. Finally, the characteristic wavelength is extracted through the GA algorithm.

Uninformative variables elimination (UVE) is a wavelength-selection method that utilizes the regression coefficient of the partial least-squares regression (PLSR) model [56–58]. The UVE method involves adding random noise variables to the sample variables and eliminating the uninformative wavelength variables through leave-one-out cross-validation. This approach retains the variables that are useful for model prediction, thereby reducing the number of variables and the model’s complexity.

2.3.3. Regression Model Establishment and Evaluation

Spectral data preprocessing and feature wavelength extraction are essential steps in spectral analysis, as they significantly impact the subsequent analysis and modeling results. Data preprocessing eliminates interference, corrects data, and enhances data accuracy and reliability, while feature-wavelength extraction selects representative wavelengths, reduces data dimensionality, enhances interpretability, and reduces noise interference. Spectral data modeling is critical in evaluating the effectiveness of spectral data preprocessing and feature-wavelength extraction. The optimal modeling combination can be obtained by assessing the modeling and prediction sets’ R^2 and RMSE. This study employed the PLSR method to model and analyze the original spectral data, preprocessed spectral data, original spectral characteristic wavelengths extracted data, and preprocessed spectral characteristic wavelengths extracted data, constructing sixty-five model combinations. Through comparative analysis, the optimal model for quantitative analysis of the FHB infection degree in the in situ wheat canopy visible areas was determined based on PLSR modeling.

The partial least-squares regression (PLSR) algorithm is a prominent multivariate statistical data analysis technique commonly employed in spectral data analysis [59–61]. The modeling procedure blends the advantages of principal component analysis, canonical correlation analysis, and linear regression analysis. It facilitates simultaneous regression modeling, data structure simplification, and correlation analysis between two sets of variables. PLSR efficiently extracts more comprehensive and profound data information, which leads to the derivation of a more appropriate regression model.

In this study, the prediction performance and the optimal prediction model of preprocessing methods, the characteristic wavelength extraction method, and the PLSR method

on the infection degree of FHB in the in situ wheat canopy visible areas were investigated. The original spectral data, preprocessed spectral data, original spectral characteristic wavelengths extracted data, and preprocessed spectral characteristic wavelengths extracted data were randomly divided into modeling (315 samples) and testing (136 samples) sets in a 7:3 ratio. A PLSR model was used to establish a spectral quantitative prediction model for the infection degree of FHB in the in situ wheat canopy visible areas. The best prediction model was determined based on the minimum root-mean-square error (*RMSE*) and the maximum coefficient of determination (R^2) of the modeling and testing sets. The formulas for calculating *RMSE* and R^2 are as follows:

$$R^2 = 1 - \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - \bar{y})^2}, \quad (2)$$

$$RMSE = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (y_i - \hat{y}_i)^2}, \quad (3)$$

where N represents the sample size; y_i represents the actual value of the i sample; \hat{y}_i represents the predicted value of sample i ; \bar{y} represents the average of the actual values of all samples; R^2 represents the coefficient of determination (the greater R^2 is, the higher the correlation is); and *RMSE* is the root-mean-square error (the smaller the *RMSE*, the better the prediction effect of the model).

3. Result and Discussion

3.1. Original Spectral Data

A total of 451 sets of original spectra (Figure 5) and average spectra (Figure 6) were acquired from healthy and infected wheat ears (with varying levels of infection: mild, moderate, and severe) in visible areas of the canopy, covering a wavelength range of 400~1000 nm. The spectral curves of healthy and infected wheat ears exhibit different characteristics due to variations in the contents of chlorophyll, carotene, and water [62].

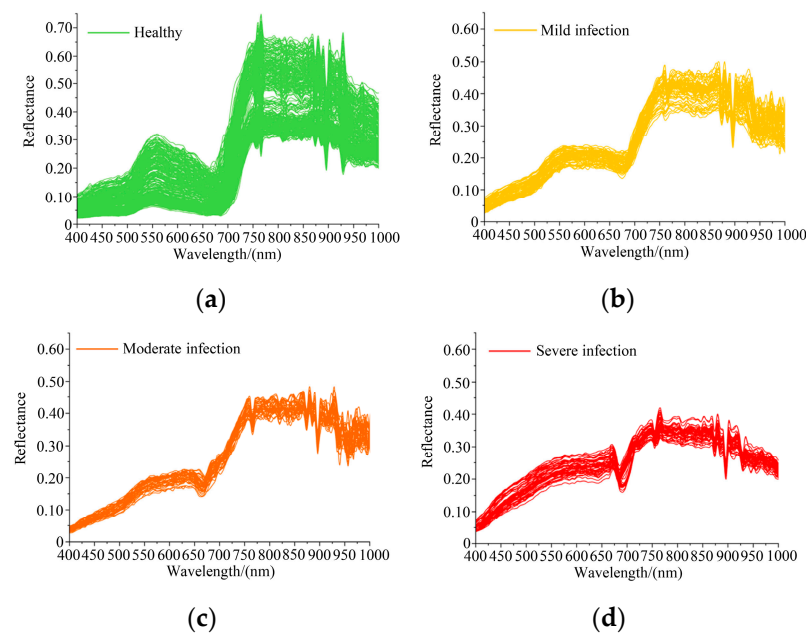


Figure 5. Original spectral curves of healthy and infected wheat ears. (a) Healthy wheat ears; (b) mildly infected wheat ears; (c) moderately infected wheat ears; (d) severely infected wheat ears.

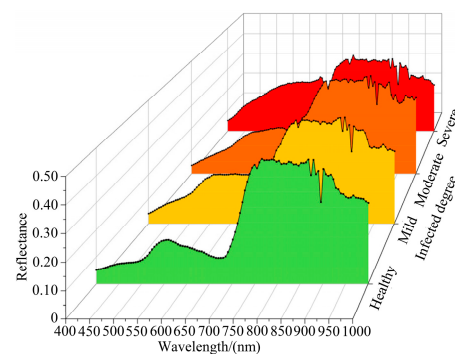


Figure 6. Average spectral curves of healthy and infected wheat ears.

Within the range of 450~700 nm, the spectral reflectance of healthy wheat ears exhibited a reduction when compared to infected wheat ears. Moreover, the reflectance displayed green peaks and red valleys. As the infection degree increased, the green peaks and red valleys gradually diminished. The spectral reflectance exhibited a rapid increase within the range of 700~750 nm, which was attributed to the high reflection platform's transition band. The platform was formed by the strong absorption of the chlorophyll red band and multiple scattering of the near-infrared band, and it is commonly referred to as the "red edge" [63]. The position of the red edge varies depending on the biomass, pigment content, and growth of the healthy wheat ears. In particular, when these ears have a high biomass, a high pigment content, and vigorous growth, the red edge moves to the longwave direction, which is also known as a "redshift." Conversely, the red edge of the infected wheat ears shifted towards the shortwave direction, commonly referred to as a "blueshift." Within the wavelength range of 750~870 nm, the spectral reflectance of healthy wheat ears was observed to be higher compared to infected wheat ears, and the spectral reflectance exhibited a decreasing trend with an increase in the degree of infection.

Clear differences could be observed between the spectral curves of healthy and infected wheat ears, characterized by evident peak–valley and high–low variations [62]. Notably, the spectral curve of severely infected wheat ears exhibited considerable changes with different degrees of infection. Conversely, the spectral curve trends of mildly and moderately infected wheat ears were comparable, potentially affecting the modeling's prediction accuracy.

3.2. Preprocessed Spectral Data

The preprocessing results are shown in Figure 7. Baseline correction techniques are essential in spectral analysis to eliminate unwanted variations in the spectra caused by systematic or instrumental factors. Among the widely used baseline correction methods are derivative processing, detrending correction, and continuous wavelet transform. The first derivative (FD) (Figure 7a) is a method that eliminates baseline shifts in the spectrum, while the second derivative (SD) (Figure 7b) eliminates baseline drift. Derivative processing eliminates the interference of other backgrounds and improves the spectral resolution and sensitivity, but introduces noise, which reduces the signal-to-noise ratio. The detrend correction (DC) (Figure 7c) is another method that eliminates baseline drift and improves the spectral resolution. On the other hand, continuous wavelet transform (CWT) (Figure 7d) amplifies the local spectral information and achieves smooth noise filtering, background removal, data compression, and multi-scale overlapping-signal-refinement analysis.

Scattering correction is a crucial preprocessing step in spectroscopy to improve the accuracy of measurements by reducing the effects of scattering. Multiplicative scatter correction (MSC) (Figure 7e) and standard normal variate (SNV) (Figure 7f) are commonly used techniques for this purpose. MSC scales the spectrum by dividing it by a reference spectrum, while SNV normalizes the data by subtracting the mean and dividing by the standard deviation. Both techniques weaken the influence of scattering on the original spectrum, enhancing the spectral features of interest.

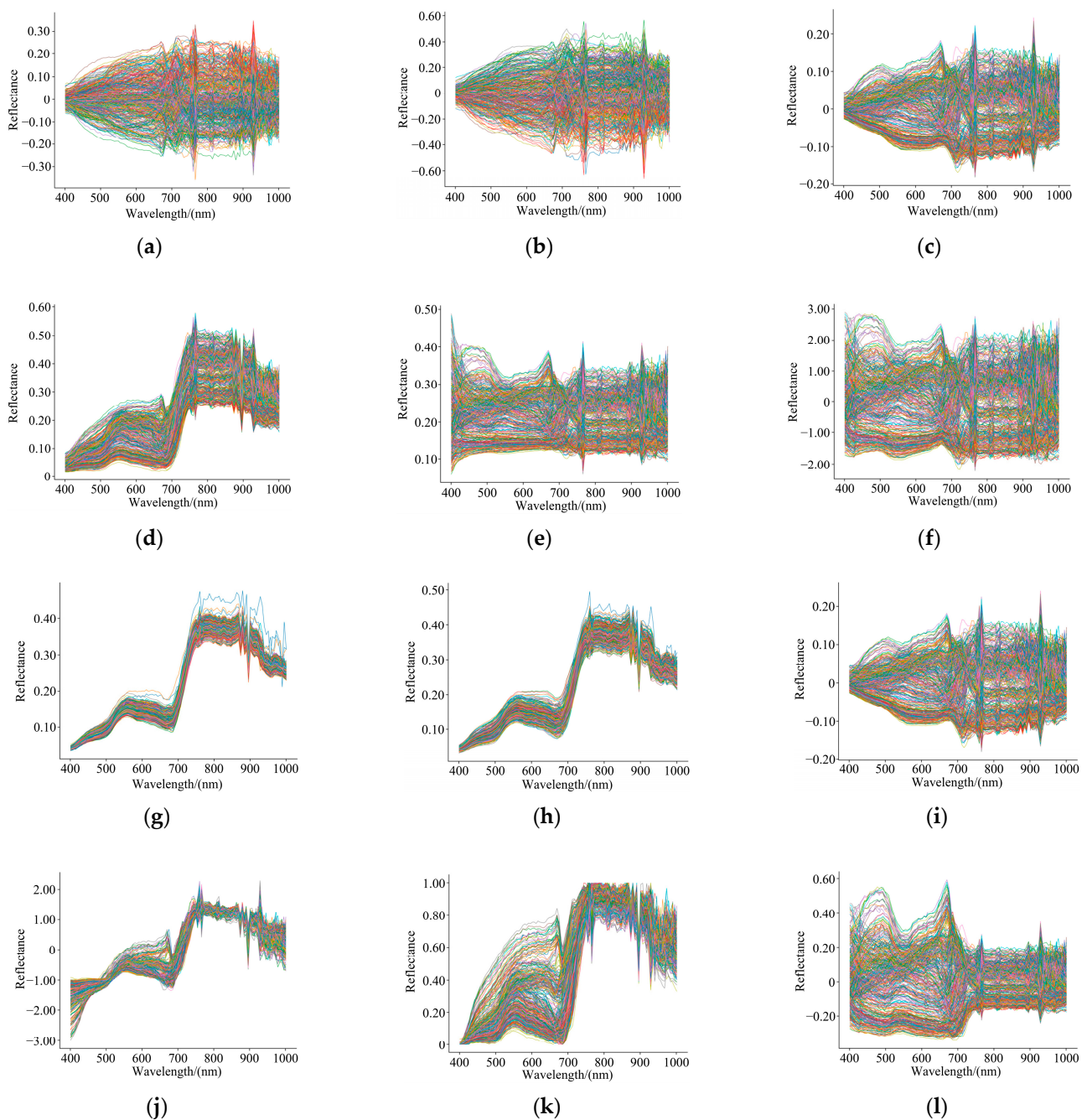


Figure 7. Preprocessed spectral curves of healthy and diseased wheat ears. (a) First derivative, FD; (b) second derivative, SD; (c) detrend correction, DC; (d) continuous wavelet transform, CWT; (e) multiplicative scatter correction, MSC; (f) standard normal variate, SNV; (g) moving average smoothing, MAS; (h) Savitzky–Golay convolution smoothing, SGCS; (i) mean centering, MC; (j) standard scaling, SS; (k) max–min scaling, MMS; (l) vector normalization, VN.

Smoothing techniques, such as moving average smoothing (MAS) (Figure 7g) and Savitzky–Golay convolution smoothing (SGCS) (Figure 7h), are commonly used in spectral data analysis. MAS involves calculating the average value of a specified number of adjacent data points to smooth out fluctuations in the signal. SGCS uses a convolution kernel to fit a polynomial curve to the spectral data and smooths the curve using weighted averaging. Both techniques improve the signal-to-noise ratio and reduce the influence of random noise in the original spectrum. These techniques are important preprocessing steps for spectral data analysis, which can enhance the accuracy and reliability of spectral measurements.

Scaling is a critical preprocessing step in data analysis to eliminate the adverse effects of scale differences and improve model robustness and predictive ability. Four commonly used scaling methods are mean centering (MC) (Figure 7i), standard scaling (SS) (Figure 7j), max–min scaling (MMS) (Figure 7k), and vector normalization (VN) (Figure 7l). These techniques adjust the scale of the data to ensure that each feature contributes equally to the analysis, improving the accuracy and interpretability of the results. The choice of scaling method depends on the specific characteristics of the data and the analysis goals.

3.3. Characteristic Wavelength Extraction Data

The extraction of informative features from high-dimensional spectral data is a crucial step in data analysis. Characteristic extraction algorithms, such as SPA, CARS, GA, and UVE, can effectively reduce the dimensionality of the data while preserving essential information. This paper employs these methods to extract characteristic wavelengths from the original and preprocessed spectral data. It is important to note that different data preprocessing and feature extraction methods may yield distinct characteristic wavelengths that require further modeling and evaluation. Therefore, the selection of an appropriate method should be based on the specific characteristics of the data and the research objectives.

3.3.1. SPA Characteristic Wavelength Extraction

The results of characteristic wavelength extraction using the SPA method are presented in Table 1. The number of extracted characteristic wavelengths was primarily distributed between 24 and 29. It was observed that the SPA method extracted the same 29 characteristic wavelengths from both the original spectral data and the data preprocessed with MSC, SNV, MC, and VN, indicating that these four preprocessing methods did not significantly affect the characteristic wavelengths extracted by the SPA method. Furthermore, the DC, CWT, and MMS preprocessing methods extracted 29, 24, and 26 characteristic wavelengths, respectively. Many of these wavelengths were shared with those extracted from the original spectral data, suggesting that these preprocessing methods had little impact on the SPA method’s characteristic wavelength extraction. In contrast, the SS preprocessing method resulted in the extraction of only 14 characteristic wavelengths, which may provide some advantages over other preprocessing methods. Finally, it was observed that the FD, SD, MAS, and SGCS preprocessing methods only extracted two characteristic wavelengths, which may lead to the loss of valuable information.

Table 1. Characteristic wavelengths extracted using SPA under different data processing methods.

Data Processing Method	Number of Wavelengths	Corresponding Characteristic Wavelength/(nm)
SPA	29	410.95, 431.24, 436.33, 472.04, 549.33, 632.89, 675.10, 680.40, 701.63, 717.61, 722.94, 728.28, 738.97, 744.32, 776.52, 787.29, 798.08, 814.30, 819.71, 825.13, 835.98, 841.42, 846.85, 857.74, 863.19, 885.04, 895.99, 901.48, 945.50
FD + SPA	2	669.81, 934.47
SD + SPA	2	675.10, 951.02
DC + SPA	29	431.24, 436.33, 472.04, 549.33, 632.89, 675.10, 680.40, 701.63, 717.61, 722.94, 728.28, 738.97, 744.32, 776.52, 787.29, 798.08, 803.48, 814.30, 819.71, 825.13, 835.98, 841.42, 846.85, 857.74, 863.19, 885.04, 895.99, 901.48, 945.50
CWT + SPA	24	410.95, 426.16, 431.24, 461.82, 559.71, 648.68, 659.24, 669.81, 701.63, 722.94, 728.28, 760.40, 814.30, 819.71, 857.74, 863.19, 868.65, 885.04, 895.99, 901.48, 912.45, 945.50, 967.62, 984.26
MSC + SPA	29	410.95, 431.24, 436.33, 472.04, 549.33, 632.89, 675.10, 680.40, 701.63, 717.61, 722.94, 728.28, 738.97, 744.32, 776.52, 787.29, 798.08, 814.30, 819.71, 825.13, 835.98, 841.42, 846.85, 857.74, 863.19, 885.04, 895.99, 901.48, 945.50

Table 1. Cont.

Data Processing Method	Number of Wavelengths	Corresponding Characteristic Wavelength/(nm)
SNV + SPA	29	410.95, 431.24, 436.33, 472.04, 549.33, 632.89, 675.10, 680.40, 701.63, 717.61, 722.94, 728.28, 738.97, 744.32, 776.52, 787.29, 798.08, 814.30, 819.71, 825.13, 835.98, 841.42, 846.85, 857.74, 863.19, 885.04, 895.99, 901.48, 945.50
MAS + SPA	2	664.52, 760.40
SGCS + SPA	2	669.81, 760.40
MC + SPA	29	410.95, 431.24, 436.33, 472.04, 549.33, 632.89, 675.10, 680.40, 701.63, 717.61, 722.94, 728.28, 738.97, 744.32, 776.52, 787.29, 798.08, 814.30, 819.71, 825.13, 835.98, 841.42, 846.85, 857.74, 863.19, 885.04, 895.99, 901.48, 945.50
SS + SPA	14	436.33, 502.84, 538.96, 685.70, 755.03, 776.52, 787.29, 819.71, 825.13, 863.19, 890.52, 945.50, 967.62, 989.81
MMS + SPA	26	416.01, 421.09, 477.16, 549.33, 675.10, 712.28, 722.94, 733.62, 744.32, 760.40, 776.52, 781.90, 787.29, 814.30, 819.71, 825.13, 841.42, 846.85, 857.74, 863.19, 885.04, 895.99, 928.96, 945.50, 967.62, 989.81
VN + SPA	29	410.95, 431.24, 436.33, 472.04, 549.33, 632.89, 675.10, 680.40, 701.63, 717.61, 722.94, 728.28, 738.97, 744.32, 776.52, 787.29, 798.08, 814.30, 819.71, 825.13, 835.98, 841.42, 846.85, 857.74, 863.19, 885.04, 895.99, 901.48, 945.50

3.3.2. CARS Characteristic Wavelength Extraction

Table 2 shows that the CARS algorithm extracted a range of 32–42 characteristic wavelengths. Notably, CARS extracted several identical characteristic wavelengths from the original and preprocessed spectral data. Nevertheless, there were some disparities between the two data sets. The SGCS preprocessed data yielded the minimum of 11 characteristic wavelengths, while MSC preprocessed data yielded the highest number of 42 characteristic wavelengths. The efficacy of CARS in extracting characteristic wavelengths under different data preprocessing methods requires further modeling and analysis.

Table 2. Characteristic wavelengths extracted using CARS under different data processing methods.

Data Processing Method	Number of Wavelengths	Corresponding Characteristic Wavelength/(nm)
CARS	34	416.01, 421.09, 431.24, 436.33, 441.41, 446.51, 466.93, 472.04, 482.29, 497.69, 523.45, 544.14, 559.71, 575.32, 580.53, 585.74, 596.19, 601.42, 606.65, 638.15, 675.10, 680.40, 701.63, 717.61, 722.94, 738.97, 776.52, 781.90, 787.29, 819.71, 825.13, 857.74, 939.98, 945.50
FD + CARS	19	410.95, 421.09, 472.04, 502.84, 513.14, 523.45, 564.91, 596.19, 601.42, 632.89, 659.24, 669.81, 696.32, 701.63, 738.97, 803.48, 819.71, 825.13, 841.42
SD + CARS	12	669.81, 738.97, 472.04, 596.19, 601.42, 653.96, 659.24, 669.81, 792.68, 803.48, 819.71, 841.42
DC + CARS	32	410.95, 421.09, 436.33, 441.41, 446.51, 466.93, 472.04, 482.29, 492.55, 497.69, 502.84, 523.45, 549.33, 559.71, 585.74, 601.42, 606.65, 643.41, 675.10, 680.40, 701.63, 717.61, 722.94, 728.28, 781.90, 787.29, 819.71, 841.42, 857.74, 874.11, 939.98, 945.50
CWT + CARS	28	410.95, 421.09, 446.51, 466.93, 472.04, 482.29, 492.55, 497.69, 523.45, 533.79, 564.91, 585.74, 601.42, 606.65, 669.81, 675.10, 701.63, 712.28, 722.94, 749.67, 755.03, 760.40, 787.29, 819.71, 825.13, 857.74, 939.98, 945.50
MSC + CARS	42	405.89, 421.09, 466.93, 472.04, 482.29, 492.55, 497.69, 523.45, 544.14, 559.71, 580.53, 596.19, 601.42, 606.65, 622.38, 627.63, 632.89, 643.41, 664.52, 669.81, 675.10, 680.40, 685.70, 691.01, 696.32, 701.63, 717.61, 722.94, 738.97, 765.77, 771.14, 776.52, 781.90, 787.29, 814.30, 819.71, 846.85, 857.74, 863.19, 874.11, 934.47, 945.50

Table 2. Cont.

Data Processing Method	Number of Wavelengths	Corresponding Characteristic Wavelength/(nm)
SNV + CARS	38	416.01, 421.09, 436.33, 446.51, 466.93, 472.04, 482.29, 492.55, 497.69, 523.45, 533.79, 549.33, 559.71, 564.91, 580.53, 585.74, 590.96, 601.42, 606.65, 627.63, 638.15, 643.41, 664.52, 669.81, 675.10, 680.40, 701.63, 717.61, 722.94, 738.97, 771.14, 776.52, 781.90, 787.29, 819.71, 857.74, 939.98, 945.50
MAS + CARS	15	482.29, 507.98, 559.71, 596.19, 632.89, 691.01, 696.32, 701.63, 776.52, 808.89, 819.71, 852.30, 868.65, 874.11, 895.99
SGCS + CARS	11	441.41, 451.61, 461.82, 492.55, 497.69, 554.52, 596.19, 622.38, 798.08, 803.48, 906.96
MC + CARS	26	410.95, 421.09, 436.33, 446.51, 472.04, 482.29, 497.69, 523.45, 549.33, 580.53, 601.42, 606.65, 638.15, 675.10, 680.40, 701.63, 717.61, 722.94, 765.77, 776.52, 781.90, 787.29, 819.71, 825.13, 830.55, 857.74
SS + CARS	34	410.95, 421.09, 436.33, 446.51, 461.82, 466.93, 472.04, 482.29, 497.69, 518.29, 549.33, 559.71, 575.32, 601.42, 606.65, 611.89, 627.63, 638.15, 643.41, 664.52, 675.10, 701.63, 717.61, 722.94, 728.28, 781.90, 787.29, 808.89, 819.71, 857.74, 895.99, 939.98, 945.50, 951.02
MMS + CARS	35	410.95, 421.09, 436.33, 461.82, 466.93, 472.04, 482.29, 497.69, 518.29, 523.45, 554.52, 559.71, 601.42, 606.65, 617.13, 669.81, 675.10, 701.63, 706.95, 717.61, 722.94, 728.28, 760.40, 765.77, 776.52, 781.90, 808.89, 819.71, 825.13, 857.74, 885.04, 895.99, 939.98, 945.50, 956.55
VN + CARS	36	405.89, 421.09, 436.33, 441.41, 446.51, 466.93, 472.04, 482.29, 497.69, 523.45, 544.14, 549.33, 559.71, 580.53, 585.74, 601.42, 606.65, 627.63, 638.15, 648.68, 664.52, 669.81, 675.10, 680.40, 701.63, 717.61, 722.94, 776.52, 781.90, 787.29, 808.89, 819.71, 825.13, 857.74, 939.98, 945.50

3.3.3. GA Characteristic Wavelength Extraction

Table 3 presents the distribution of characteristic wavelengths extracted using the GA algorithm, which was mainly concentrated between 36 and 42 wavelengths. The characteristic wavelengths extracted using GA from the original and pre-processed spectral data were comparable to those extracted using CARS. Although some identical wavelengths were extracted using GA under different data preprocessing methods, there were still some differences in the overall comparison. The SGCS preprocessed data exhibited a minimum of 12 characteristic wavelengths, while MSC and VN preprocessed data yielded the highest number of 42 characteristic wavelengths. Further modeling and analysis are necessary to assess the superiority of GA in extracting characteristic wavelengths under different data processing methods.

Table 3. Characteristic wavelengths extracted using GA under different data processing methods.

Data Processing Method	Number of Wavelengths	Corresponding Characteristic Wavelength/(nm)
GA	36	400.83, 421.09, 431.24, 436.33, 472.04, 523.45, 559.71, 580.53, 601.42, 617.13, 622.38, 627.63, 643.41, 648.68, 653.96, 675.10, 701.63, 728.28, 755.03, 760.40, 771.14, 776.52, 787.29, 814.30, 819.71, 825.13, 857.74, 890.52, 939.98, 945.50, 951.02, 956.55, 973.16, 978.71, 984.26, 995.37
FD + GA	22	410.95, 421.09, 436.33, 441.41, 446.51, 461.82, 513.14, 523.45, 559.71, 564.91, 570.11, 653.96, 659.24, 669.81, 760.40, 765.77, 819.71, 945.50, 956.55, 967.62, 984.26, 995.37
SD + GA	31	405.89, 421.09, 477.16, 492.55, 513.14, 523.45, 533.79, 544.14, 549.33, 559.71, 564.91, 585.74, 590.96, 596.19, 611.89, 653.96, 659.24, 738.97, 744.32, 771.14, 798.08, 819.71, 846.85, 852.30, 879.57, 917.95, 951.02, 956.55, 978.71, 984.26, 995.37
DC + GA	40	400.83, 421.09, 431.24, 472.04, 523.45, 544.14, 559.71, 580.53, 601.42, 606.65, 611.89, 617.13, 622.38, 627.63, 653.96, 664.52, 669.81, 675.10, 701.63, 712.28, 717.61, 722.94, 733.62, 755.03, 760.40, 776.52, 787.29, 814.30, 819.71, 825.13, 857.74, 890.52, 928.96, 939.98, 945.50, 956.55, 973.16, 978.71, 984.26, 995.37

Table 3. Cont.

Data Processing Method	Number of Wavelengths	Corresponding Characteristic Wavelength/(nm)
CWT + GA	38	410.95, 421.09, 426.16, 431.24, 441.41, 472.04, 482.29, 601.42, 617.13, 622.38, 653.96, 659.24, 664.52, 669.81, 675.10, 685.70, 701.63, 717.61, 722.94, 728.28, 744.32, 749.67, 760.40, 776.52, 787.29, 814.30, 819.71, 825.13, 857.74, 890.52, 901.48, 939.98, 945.50, 956.55, 973.16, 978.71, 984.26, 995.37
MSC + GA	42	405.89, 421.09, 431.24, 472.04, 523.45, 559.71, 580.53, 601.42, 606.65, 617.13, 622.38, 632.89, 653.96, 659.24, 664.52, 669.81, 675.10, 680.40, 685.70, 701.63, 712.28, 717.61, 722.94, 738.97, 755.03, 760.40, 771.14, 776.52, 787.29, 808.89, 814.30, 819.71, 825.13, 857.74, 890.52, 939.98, 945.50, 956.55, 973.16, 978.71, 984.26, 995.37
SNV + GA	39	405.89, 421.09, 431.24, 436.33, 472.04, 482.29, 523.45, 559.71, 580.53, 601.42, 617.13, 622.38, 653.96, 669.81, 675.10, 680.40, 701.63, 712.28, 717.61, 722.94, 738.97, 744.32, 755.03, 760.40, 776.52, 787.29, 798.08, 814.30, 819.71, 825.13, 846.85, 857.74, 939.98, 945.50, 956.55, 973.16, 978.71, 984.26, 995.37
MAS + GA	13	400.83, 431.24, 482.29, 533.79, 564.91, 575.32, 643.41, 733.62, 765.77, 776.52, 945.50, 951.02, 1000.94
SGCS + GA	12	400.83, 426.16, 451.61, 482.29, 680.40, 803.48, 825.13, 874.11, 885.04, 895.99, 906.96, 951.02
MC + GA	40	400.83, 421.09, 431.24, 472.04, 523.45, 538.96, 559.71, 564.91, 580.53, 601.42, 617.13, 622.38, 653.96, 664.52, 669.81, 675.10, 680.40, 691.01, 701.63, 717.61, 722.94, 738.97, 755.03, 760.40, 771.14, 776.52, 787.29, 808.89, 814.30, 819.71, 825.13, 857.74, 890.52, 939.98, 945.50, 956.55, 973.16, 978.71, 984.26, 995.37
SS + GA	37	436.33, 466.93, 472.04, 482.29, 523.45, 544.14, 549.33, 559.71, 575.32, 638.15, 643.41, 675.10, 685.70, 701.63, 717.61, 722.94, 728.28, 738.97, 744.32, 765.77, 771.14, 776.52, 803.48, 808.89, 819.71, 835.98, 846.85, 852.30, 863.19, 874.11, 923.45, 945.50, 956.55, 962.08, 978.71, 984.26, 989.81
MMS + GA	36	405.89, 421.09, 472.04, 482.29, 513.14, 518.29, 559.71, 580.53, 590.96, 601.42, 606.65, 617.13, 632.89, 638.15, 643.41, 675.10, 717.61, 722.94, 728.28, 755.03, 760.40, 771.14, 776.52, 798.08, 808.89, 814.30, 819.71, 825.13, 857.74, 890.52, 939.98, 945.50, 956.55, 973.16, 978.71, 984.26
VN + GA	42	400.83, 421.09, 431.24, 472.04, 523.45, 538.96, 549.33, 559.71, 564.91, 580.53, 601.42, 617.13, 622.38, 627.63, 653.96, 669.81, 675.10, 680.40, 691.01, 701.63, 717.61, 722.94, 728.28, 738.97, 755.03, 760.40, 771.14, 776.52, 787.29, 814.30, 819.71, 825.13, 852.30, 857.74, 890.52, 939.98, 945.50, 956.55, 973.16, 978.71, 984.26, 995.37

3.3.4. UVE Characteristic Wavelength Extraction

Table 4 shows that the UVE algorithm extracted a large number of characteristic wavelengths, which mainly ranged from 55 to 74 wavelengths. Interestingly, the original spectral data and SNV preprocessed spectral data yielded the same 69 characteristic wavelengths. In contrast, FD, SD, DC, CWT, MSC, MC, SS, MMS, and VN preprocessing spectral data extracted 56, 37, 56, 44, 55, 69, 65, 74, and 69 characteristic wavelengths, respectively. Furthermore, most of these preprocessing methods had a negligible effect on the characteristic wavelengths extracted using UVE. Specifically, they extracted more identical wavelengths with the characteristic wavelengths extracted from the original spectral data. However, only 17 and 7 characteristic wavelengths were extracted from MAS and SGCS preprocessed spectral data, respectively. These findings imply that these two preprocessing methods may lead to the loss of significant information.

Table 4. Characteristic wavelengths extracted using UVE under different data processing methods.

Data Processing Method	Number of Wavelengths	Corresponding Characteristic Wavelength/(nm)
UVE	69	405.89, 410.95, 416.01, 421.09, 426.16, 431.24, 441.41, 451.61, 456.71, 461.82, 466.93, 472.04, 477.16, 482.29, 487.42, 492.55, 497.69, 502.84, 507.98, 513.14, 528.62, 533.79, 538.96, 544.14, 549.33, 554.52, 559.71, 564.91, 570.11, 575.32, 580.53, 585.74, 590.96, 596.19, 601.42, 606.65, 611.89, 643.41, 648.68, 653.96, 659.24, 664.52, 669.81, 675.10, 680.40, 722.94, 733.62, 738.97, 744.32, 749.67, 760.40, 765.77, 771.14, 776.52, 781.90, 798.08, 808.89, 814.30, 819.71, 825.13, 841.42, 846.85, 857.74, 863.19, 934.47, 945.50, 956.55, 984.26, 995.37
FD + UVE	56	421.09, 426.16, 441.41, 451.61, 456.71, 461.82, 466.93, 472.04, 477.16, 482.29, 487.42, 492.55, 497.69, 502.84, 507.98, 513.14, 528.62, 533.79, 538.96, 544.14, 549.33, 554.52, 559.71, 564.91, 570.11, 575.32, 580.53, 585.74, 590.96, 601.42, 611.89, 617.13, 632.89, 643.41, 648.68, 653.96, 669.81, 733.62, 738.97, 744.32, 749.67, 760.40, 771.14, 798.08, 803.48, 808.89, 819.71, 825.13, 846.85, 863.19, 885.04, 945.50, 951.02, 956.55, 984.26, 995.37
SD + UVE	37	410.95, 421.09, 461.82, 472.04, 477.16, 482.29, 487.42, 492.55, 497.69, 502.84, 507.98, 513.14, 533.79, 538.96, 544.14, 549.33, 554.52, 559.71, 564.91, 570.11, 575.32, 601.42, 611.89, 738.97, 744.32, 749.67, 760.40, 771.14, 803.48, 808.89, 819.71, 825.13, 846.85, 951.02, 956.55, 978.71, 995.37
DC + UVE	56	410.95, 416.01, 421.09, 426.16, 431.24, 451.61, 456.71, 461.82, 466.93, 472.04, 477.16, 482.29, 487.42, 492.55, 497.69, 502.84, 507.98, 513.14, 533.79, 538.96, 544.14, 549.33, 554.52, 559.71, 564.91, 570.11, 575.32, 580.53, 585.74, 590.96, 601.42, 611.89, 648.68, 653.96, 659.24, 664.52, 669.81, 675.10, 680.40, 733.62, 738.97, 744.32, 749.67, 760.40, 765.77, 771.14, 776.52, 814.30, 819.71, 825.13, 846.85, 857.74, 863.19, 934.47, 945.50, 956.55
CWT + UVE	44	410.95, 416.01, 421.09, 426.16, 451.61, 456.71, 461.82, 466.93, 472.04, 477.16, 482.29, 487.42, 492.55, 497.69, 502.84, 507.98, 513.14, 533.79, 538.96, 544.14, 549.33, 554.52, 559.71, 564.91, 570.11, 575.32, 580.53, 585.74, 590.96, 601.42, 653.96, 659.24, 664.52, 669.81, 675.10, 738.97, 749.67, 760.40, 776.52, 819.71, 825.13, 857.74, 945.50, 956.55
MSC + UVE	55	410.95, 416.01, 421.09, 426.16, 431.24, 451.61, 456.71, 461.82, 466.93, 472.04, 477.16, 482.29, 487.42, 492.55, 497.69, 502.84, 507.98, 513.14, 533.79, 538.96, 544.14, 549.33, 554.52, 559.71, 564.91, 570.11, 575.32, 580.53, 585.74, 590.96, 601.42, 611.89, 653.96, 659.24, 664.52, 669.81, 675.10, 680.40, 733.62, 738.97, 744.32, 749.67, 760.40, 765.77, 771.14, 776.52, 814.30, 819.71, 825.13, 846.85, 857.74, 863.19, 934.47, 945.50, 956.55
SNV + UVE	69	405.89, 410.95, 416.01, 421.09, 426.16, 431.24, 441.41, 451.61, 456.71, 461.82, 466.93, 472.04, 477.16, 482.29, 487.42, 492.55, 497.69, 502.84, 507.98, 513.14, 528.62, 533.79, 538.96, 544.14, 549.33, 554.52, 559.71, 564.91, 570.11, 575.32, 580.53, 585.74, 590.96, 596.19, 601.42, 606.65, 611.89, 643.41, 648.68, 653.96, 659.24, 664.52, 669.81, 675.10, 680.40, 722.94, 733.62, 738.97, 744.32, 749.67, 760.40, 765.77, 771.14, 776.52, 781.90, 798.08, 808.89, 814.30, 819.71, 825.13, 841.42, 846.85, 857.74, 863.19, 934.47, 945.50, 956.55, 984.26, 995.37
MAS + UVE	17	482.29, 487.42, 492.55, 549.33, 554.52, 559.71, 564.91, 570.11, 575.32, 580.53, 585.74, 601.42, 606.65, 765.77, 776.52, 945.50, 984.26
SGCS + UVE	7	400.83, 461.82, 477.16, 482.29, 487.42, 507.98, 895.99
MC + UVE	69	405.89, 410.95, 416.01, 421.09, 426.16, 431.24, 441.41, 451.61, 456.71, 461.82, 466.93, 472.04, 477.16, 482.29, 487.42, 492.55, 497.69, 502.84, 507.98, 513.14, 528.62, 533.79, 538.96, 544.14, 549.33, 554.52, 559.71, 564.91, 570.11, 575.32, 580.53, 585.74, 590.96, 596.19, 601.42, 606.65, 611.89, 643.41, 648.68, 653.96, 659.24, 664.52, 669.81, 675.10, 680.40, 722.94, 733.62, 738.97, 744.32, 749.67, 760.40, 765.77, 771.14, 776.52, 781.90, 798.08, 808.89, 814.30, 819.71, 825.13, 841.42, 846.85, 857.74, 863.19, 934.47, 945.50, 956.55, 984.26, 995.37
SS + UVE	65	400.83, 405.89, 410.95, 416.01, 421.09, 426.16, 431.24, 436.33, 441.41, 446.51, 451.61, 456.71, 461.82, 466.93, 472.04, 482.29, 487.42, 492.55, 497.69, 502.84, 507.98, 533.79, 538.96, 544.14, 549.33, 554.52, 559.71, 564.91, 570.11, 606.65, 611.89, 617.13, 622.38, 627.63, 632.89, 638.15, 643.41, 648.68, 653.96, 659.24, 664.52, 669.81, 675.10, 680.40, 685.70, 722.94, 733.62, 738.97, 749.67, 755.03, 760.40, 771.14, 776.52, 787.29, 808.89, 819.71, 825.13, 846.85, 857.74, 863.19, 895.99, 934.47, 945.50, 956.55, 984.26

Table 4. Cont.

Data Processing Method	Number of Wavelengths	Corresponding Characteristic Wavelength/(nm)
MMS + UVE	74	400.83, 410.95, 416.01, 421.09, 426.16, 441.41, 446.51, 451.61, 456.71, 461.82, 466.93, 472.04, 477.16, 482.29, 487.42, 492.55, 497.69, 502.84, 507.98, 513.14, 518.29, 528.62, 533.79, 538.96, 544.14, 549.33, 554.52, 559.71, 564.91, 570.11, 575.32, 580.53, 585.74, 617.13, 622.38, 627.63, 632.89, 638.15, 643.41, 648.68, 653.96, 659.24, 664.52, 669.81, 675.10, 680.40, 685.70, 701.63, 706.95, 712.28, 717.61, 728.28, 733.62, 738.97, 744.32, 749.67, 765.77, 771.14, 776.52, 798.08, 814.30, 819.71, 825.13, 830.55, 846.85, 857.74, 863.19, 890.52, 923.45, 934.47, 945.50, 956.55, 984.26, 995.37
VN + UVE	69	405.89, 410.95, 416.01, 421.09, 426.16, 431.24, 441.41, 451.61, 456.71, 461.82, 466.93, 472.04, 477.16, 482.29, 487.42, 492.55, 497.69, 502.84, 507.98, 513.14, 528.62, 533.79, 538.96, 544.14, 549.33, 554.52, 559.71, 564.91, 570.11, 575.32, 580.53, 585.74, 590.96, 596.19, 601.42, 606.65, 611.89, 643.41, 648.68, 653.96, 659.24, 664.52, 669.81, 675.10, 680.40, 722.94, 733.62, 738.97, 744.32, 749.67, 760.40, 765.77, 771.14, 776.52, 781.90, 798.08, 808.89, 814.30, 819.71, 825.13, 841.42, 846.85, 857.74, 863.19, 934.47, 945.50, 956.55, 984.26, 995.37

3.4. Results and Analysis of Regression Modeling

3.4.1. PLSR Modeling of Original Spectral Data

In this study, PLSR modeling was conducted on the original spectral data without any treatment, and the obtained results are presented in Table 5. The modeling set R^2 of the PLSR model was 0.9441, and the corresponding $RMSE$ was 0.2442. The testing set R^2 was 0.9215, and the corresponding $RMSE$ was 0.2676, which indicates that the model had good predictive ability. However, the original spectral data contained a large number of wavelengths, specifically 114, which could increase the computational complexity and reduce the data processing efficiency if modeled directly.

Table 5. PLSR modeling results of original spectral data.

Model	Number of Wavelengths	Modeling Set		Testing Set	
		Coefficient of Determination (R^2)	Root-Mean-Square Error ($RMSE$)	Coefficient of Determination (R^2)	Root-Mean-Square Error ($RMSE$)
PLSR	114	0.9441	0.2442	0.9215	0.2676

3.4.2. PLSR Modeling of Preprocessed Spectral Data

The results of the PLSR modeling of preprocessed spectral data are presented in Table 6. The impact of MSC, SNV, MC, and VN preprocessing methods on the PLSR modeling was found to be insignificant compared to that of the original spectral data. This indicated that these four pre-processing methods did not affect the PLSR modeling in this study. On the other hand, the R^2 and $RMSE$ values of the PLSR modeling set and testing set of SS and MMS preprocessed spectral data were slightly better than those of the original spectral data, suggesting that these two preprocessing methods enhanced the PLSR modeling effect to a certain extent. However, the PLSR modeling set and testing set of DC and CWT preprocessed spectral data showed slightly worse R^2 and $RMSE$ values than those of the original spectral data, indicating that these two preprocessing methods reduced the PLSR modeling effect in this study. Furthermore, the PLSR modeling set and testing set of FD, SD, MAS, and SGCS preprocessed spectral data demonstrated inferior R^2 and $RMSE$ values, indicating that these four preprocessing methods were unsuitable for direct PLSR modeling in this study. It should be noted that, although direct PLSR modeling of preprocessed spectral data can improve the prediction accuracy to a certain extent, the preprocessed spectral data contained 114 wavelengths and required complex data calculation, which affects processing efficiency.

Table 6. PLSR modeling results of preprocessed spectral data.

Model	Number of Wavelengths	Modeling Set		Testing Set	
		Coefficient of Determination (R^2)	Root-Mean-Square Error (RMSE)	Coefficient of Determination (R^2)	Root-Mean-Square Error (RMSE)
FD + PLSR	114	0.4868	0.4250	0.4036	0.4452
SD + PLSR	114	0.1001	0.4965	0.1203	0.4731
DC + PLSR	114	0.9428	0.2456	0.9203	0.2686
CWT + PLSR	114	0.9418	0.2467	0.9117	0.2755
MSC + PLSR	114	0.9441	0.2442	0.9215	0.2676
SNV + PLSR	114	0.9441	0.2442	0.9215	0.2676
MAS + PLSR	114	0.0275	0.4988	0.0376	0.5006
SGCS + PLSR	114	0.0661	0.4938	0.0457	0.4995
MC + PLSR	114	0.9441	0.2442	0.9215	0.2676
SS + PLSR	114	0.9440	0.2443	0.9304	0.2596
MMS + PLSR	114	0.9450	0.2433	0.9292	0.2607
VN + PLSR	114	0.9441	0.2442	0.9215	0.2676

3.4.3. PLSR Modeling of Original Spectral Characteristic Wavelengths Extracted Data

The results of PLSR modeling of the original spectral characteristic wavelength extraction data are presented in Table 7. The SPA method was employed to extract 29 characteristic wavelengths, which effectively reduced the dimensionality of the original spectral data. However, the R^2 and RMSE values obtained for both the SPA + PLSR modeling set and the testing set were inferior to those of the PLSR model. This outcome suggests that the SPA method might have caused the loss of some critical spectral information during data dimensionality reduction. On the other hand, the CARS and GA methods identified 34 and 36 characteristic wavelengths, respectively. The R^2 and RMSE values obtained for the CARS + PLSR and GA + PLSR modeling sets and testing sets were superior to those of the PLSR model, indicating that these methods were able to remove some noise information and optimize the critical spectral information while performing data dimensionality reduction. The UVE method, however, extracted 69 characteristic wavelengths. While the R^2 and RMSE values of the modeling set of the UVE + PLSR model were slightly inferior to those of the PLSR model, the R^2 and RMSE values of the testing set of the UVE + PLSR model were slightly better than the PLSR model. Nonetheless, there was a significant difference in the modeling effect when compared to the CARS + PLSR and GA + PLSR models.

Table 7. PLSR modeling results of original spectral characteristic wavelengths extracted data.

Model	Number of Wavelengths	Modeling Set		Testing Set	
		Coefficient of Determination (R^2)	Root-Mean-Square Error (RMSE)	Coefficient of Determination (R^2)	Root-Mean-Square Error (RMSE)
SPA + PLSR	29	0.9312	0.2573	0.9092	0.2774
CARS + PLSR	34	0.9551	0.2312	0.9304	0.2596
GA + PLSR	36	0.9449	0.2434	0.9323	0.2578
UVE + PLSR	69	0.9395	0.2491	0.9236	0.2657

3.4.4. PLSR Modeling of Preprocessed Spectral Characteristic Wavelengths Extracted Data

Table 8 displays the results of PLSR modeling for preprocessed spectral characteristic wavelength extraction data. The spectral data were preprocessed using four methods, namely, FD, SD, MAS, and SGCS, to extract characteristic wavelengths using SPA, CARS, GA, and UVE, respectively. The R^2 and RMSE values obtained for the PLSR modeling set and testing set indicated that these four preprocessing methods were unsuitable for PLSR modeling of preprocessed spectral characteristic wavelength extraction data in this study.

For the other eight types of preprocessed spectral data, SPA, CARS, GA, and UVE were employed to extract characteristic wavelengths. The PLSR modeling set and testing set for extracting characteristic wavelength data yielded relatively better R^2 and RMSE values, which could be further compared and analyzed.

Table 8. PLSR modeling results of preprocessed spectral characteristic wavelengths extracted data.

Model	Number of Wavelengths	Modeling Set		Testing Set	
		Coefficient of Determination (R^2)	Root-Mean-Square Error (RMSE)	Coefficient of Determination (R^2)	Root-Mean-Square Error (RMSE)
FD + SPA + PLSR	2	0.3855	0.4446	0.4139	0.4432
FD + CARS + PLSR	19	0.5201	0.4179	0.4501	0.4362
FD + GA + PLSR	22	0.4929	0.4237	0.4445	0.4373
FD + UVE + PLSR	56	0.4836	0.4256	0.4431	0.4376
SD + SPA + PLSR	2	0.1011	0.4964	0.1391	0.4705
SD + CARS + PLSR	12	0.1872	0.4840	0.1383	0.4707
SD + GA + PLSR	31	0.1580	0.4883	0.0964	0.4763
SD + UVE + PLSR	37	0.1463	0.4900	0.0685	0.4799
DC + SPA + PLSR	29	0.9313	0.2572	0.9093	0.2773
DC + CARS + PLSR	32	0.9521	0.2350	0.9312	0.2588
DC + GA + PLSR	40	0.9445	0.2438	0.9305	0.2595
DC + UVE + PLSR	56	0.9363	0.2524	0.9245	0.2649
CWT + SPA + PLSR	24	0.9290	0.2593	0.9102	0.2767
CWT + CARS + PLSR	28	0.9468	0.2413	0.9258	0.2638
CWT + GA + PLSR	38	0.9421	0.2464	0.9312	0.2589
CWT + UVE + PLSR	44	0.9284	0.2598	0.9141	0.2736
MSC + SPA + PLSR	29	0.9312	0.2573	0.9092	0.2774
MSC + CARS + PLSR	42	0.9537	0.2329	0.9284	0.2614
MSC + GA + PLSR	42	0.9455	0.2427	0.9326	0.2575
MSC + UVE + PLSR	55	0.9378	0.2509	0.9265	0.2632
SNV + SPA + PLSR	29	0.9312	0.2573	0.9092	0.2774
SNV + CARS + PLSR	38	0.9546	0.2318	0.9342	0.2559
SNV + GA + PLSR	39	0.9446	0.2437	0.9369	0.2533
SNV + UVE + PLSR	69	0.9395	0.2491	0.9236	0.2657
MAS + SPA + PLSR	2	0.0339	0.4980	0.0577	0.4980
MAS + CARS + PLSR	15	0.0263	0.4990	0.0376	0.5006
MAS + GA + PLSR	13	0.0144	0.5005	0.0096	0.5042
MAS + UVE + PLSR	17	0.0487	0.4961	0.0348	0.5009
SGCS + SPA + PLSR	2	0.0568	0.4950	0.0623	0.4973
SGCS + CARS + PLSR	11	0.0966	0.4897	0.0223	0.5026
SGCS + GA + PLSR	12	0.0699	0.4933	0.0443	0.4997
SGCS + UVE + PLSR	7	0.0735	0.4928	0.0458	0.4995
MC + SPA + PLSR	29	0.9312	0.2573	0.9092	0.2774
MC + CARS + PLSR	26	0.9483	0.2395	0.9160	0.2721
MC + GA + PLSR	40	0.9473	0.2407	0.9325	0.2576
MC + UVE + PLSR	69	0.9395	0.2491	0.9236	0.2657
SS + SPA + PLSR	14	0.9223	0.2652	0.9251	0.2644
SS + CARS + PLSR	34	0.9571	0.2286	0.9342	0.2560
SS + GA + PLSR	37	0.9515	0.2357	0.9354	0.2548
SS + UVE + PLSR	65	0.9560	0.2300	0.9170	0.2713
MMS + SPA + PLSR	26	0.9345	0.2541	0.9147	0.2731
MMS + CARS + PLSR	35	0.9493	0.2384	0.9376	0.2526
MMS + GA + PLSR	36	0.9439	0.2445	0.9324	0.2577
MMS + UVE + PLSR	74	0.9323	0.2562	0.9247	0.2647
VN + SPA + PLSR	29	0.9312	0.2573	0.9092	0.2774
VN + CARS + PLSR	36	0.9569	0.2289	0.9264	0.2632
VN + GA + PLSR	42	0.9452	0.2430	0.9342	0.2560
VN + UVE + PLSR	69	0.9395	0.2491	0.9236	0.2657

For the same preprocessing method, CARS and GA characteristic wavelength extraction data for ePLSR modeling yielded better results than SPA and UVE. Among the different preprocessing techniques, DC + CARS + PLSR showed the most favorable outcome for the DC method, with an R^2 value of 0.9312 and an $RMSE$ of 0.2588. Similarly, for the CWT method, CWT + GA + PLSR proved to be the most effective, with an R^2 value of 0.9312 and an $RMSE$ of 0.2589. For the MSC method, MSC + GA + PLSR yielded the highest R^2 value of 0.9326 and the lowest $RMSE$ of 0.2575. Likewise, for the SNV method, SNV + GA + PLSR gave the best results, with an R^2 value of 0.9369 and an $RMSE$ of 0.2533. For the MC method, MC + GA + PLSR resulted in an R^2 value of 0.9325 and an $RMSE$ of 0.2576. SS + GA + PLSR led to the best results for the SS method, with an R^2 value of 0.9354 and an $RMSE$ of 0.2548. Additionally, MMS + CARS + PLSR demonstrated the most favorable results for the MMS method, with an R^2 value of 0.9376 and an $RMSE$ of 0.2526. Finally, for the VN method, VN + GA + PLSR exhibited the highest R^2 value of 0.9342 and an $RMSE$ of 0.2560.

For the same feature wavelength extraction method, it has been observed that PLSR modeling of spectral data preprocessed using scale scaling and scattering correction methods can yield better results compared to baseline correction and smoothing preprocessing methods. Among the different characteristic wavelength extraction methods, SS + SPA + PLSR exhibited the highest efficacy for SPA, with an R^2 value of 0.9251 and an $RMSE$ value of 0.2644. For CARS, MMS + CARS + PLSR yielded the best performance, with an R^2 value of 0.9376 and an $RMSE$ value of 0.2526. GA exhibited the best results with SNV + GA + PLSR, with an R^2 value of 0.9369 and an $RMSE$ value of 0.2533. For UVE, MSC + UVE + PLSR delivered the best results, with an R^2 value of 0.9265 and an $RMSE$ value of 0.2632.

Based on the evaluation criteria of the R^2 and $RMSE$ of the testing set, a comprehensive comparison of the different preprocessing methods was conducted, and it was found that the MMS + CARS + PLSR, SNV + GA + PLSR, and SS + GA + PLSR models demonstrated superior performance. These models achieved an R^2 value of more than 0.9350 and controlled the $RMSE$ below 0.2550. A comparison of the predicted and true values for the testing set of these three models is illustrated in Figure 8, which indicated that the prediction effect of these models was satisfactory, with no significant deviation between the predicted and true values. However, in terms of computational complexity and efficiency, the GA method was more intricate than the CARS method. Therefore, MMS + CARS + PLSR was selected as the optimal model for spectral quantitative analysis of the infection degree of FHB in the wheat canopy's visible area. The model extracted thirty-five characteristic wavelengths, and the R^2 and $RMSE$ of the modeling set were 0.9493 and 0.2384, respectively. The R^2 and $RMSE$ of the testing set were 0.9312 and 0.2588, respectively. The proposed model enabled a spectral quantitative analysis of the infection degree of FHB in the in situ wheat canopy's visible area, providing useful insights for national food security and green agricultural development.

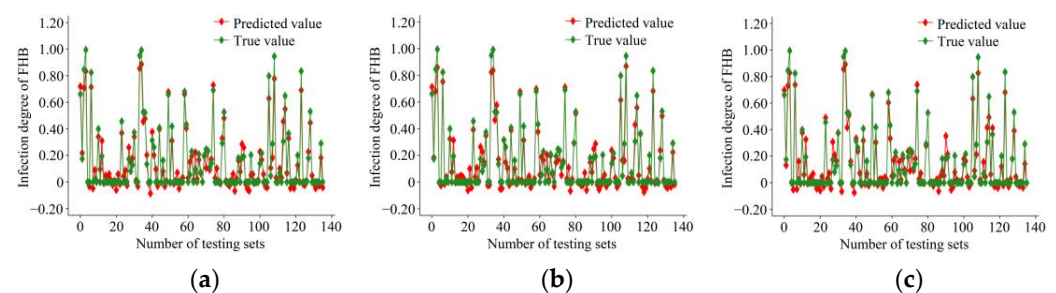


Figure 8. Comparison of the prediction effects of three better models. (a) Prediction effect of MMS + CARS + PLSR model; (b) prediction effect of SNV + GA + PLSR model; (c) prediction effect of SS + GA + PLSR model.

3.5. Some Discussions

This study utilized hyperspectral imaging technology to develop a method for spectrally quantifying the degree of Fusarium head blight (FHB) infection in an in situ wheat canopy's visible areas. The variation in wheat ear growth posture posed a significant challenge in obtaining complete and consistent spectral images in the in situ canopy visible areas. Nonetheless, the study found that acquiring 50% hyperspectral information of wheat ears in the in situ canopy's visible areas could reflect and predict the degree and trend of wheat FHB infection to a certain extent. To this end, the study established sixty-five different spectral quantitative prediction models using combinations of twelve pretreatment methods, four characteristic wavelength extraction methods, and PLSR modeling methods, of which the MMS + CARS + PLSR model was the most effective. The study's findings provide valuable insights into utilizing hyperspectral imaging technology for detecting and evaluating FHB infection in the visible areas of a wheat canopy and highlight the challenges of obtaining spectral images of wheat ears in the visible areas of the canopy with good integrity and consistency. The findings of this study could be helpful for farmers and agricultural policymakers to take necessary measures to prevent FHB infection and ensure stable grain production, ultimately contributing to the green development of agriculture.

However, there are some limitations in this study to be considered. A limited amount of research has been conducted on using in situ spectral imaging to detect Fusarium head blight (FHB) in wheat [18,31–33]. Instead, most of the research in this area has focused on ex situ detection methods [20–30]. In order to ensure the originality and authenticity of wheat growth posture and spectral data, in situ spectral data of wheat grown in the experimental soil tank were collected. Compared with the ex situ data acquisition, the experimental conditions of in situ data acquisition are difficult to control, and data acquisition is complex. Therefore, only 451 sets of spectral data were obtained in this study, and the number of samples was small. The reliability of the model needs to be further verified and improved. In further research, we will continue to explore the method of in situ spectral imaging to detect wheat FHB and strengthen the sample size collection to improve the model's reliability. At the same time, this study explored the spectral quantitative prediction model of the infection degree of FHB in the visible areas of an in situ wheat canopy based on PLSR regression. The modeling effects based on the neural network, support vector machine, multiple linear regression, and other methods must be further explored and analyzed in subsequent studies.

In addition, the experimental soil box can reasonably simulate the in situ growth state of wheat, but there are specific gaps and limitations compared with the field environment. Collecting the corresponding spectral data in the field environment is a complex and challenging task. As a field experiment is affected by various environmental factors, the complexity of the experimental data is increased. Therefore, when collecting spectral data in a field environment, we must be more careful in selecting the appropriate spectral instrument and collection time. Moreover, different collection methods should be selected for vegetation characteristics, and the influence of environmental factors on data should be considered. Finally, we need to use excellent data processing and analysis methods to process the collected data. At present, the spectral quantitative prediction of wheat FHB still faces many challenges and difficulties. More researchers are needed to actively participate in this field and jointly promote the progress of this work.

4. Conclusions

In this study, a method for spectral quantitative analysis of the in situ FHB infection degree in the visible areas of the wheat canopy was developed using hyperspectral imaging technology and a detection and evaluation method for FHB. Through data acquisition experiments and model comparison analysis, the following conclusions were drawn:

(1) The growth posture of wheat ears varies greatly due to the influence of their weight, and most have an inclination angle of (-15° ~ 15°), resulting in only 50% of the spectral images

of wheat ears in the canopy's visible areas being obtained. Obtaining spectral images of wheat ears in the canopy's visible areas with good integrity and consistency is challenging.

(2) The irreversible disease of wheat FHB will cause severe losses once infected, and obtaining 50% hyperspectral information of wheat ears in the canopy visible areas can reflect and predict the degree and trend of wheat FHB infection to a certain extent. This method can be used to determine the spectral acquisition method of FHB and the evaluation method of infection degree in the wheat canopy's visible areas.

(3) This study established sixty-five different combinations of spectral quantitative prediction models based on twelve pretreatment methods, four characteristic wavelength extraction methods, and PLSR modeling methods. The MMS + CARS + PLSR model was determined as the optimal spectral quantitative analysis model for the infection degree of FHB in the wheat canopy visible areas. The model extracted thirty-five characteristic wavelengths; the modeling set of the R^2 was 0.9490, the RMSE was 0.2384, the testing set of the R^2 was 0.9312, and the RMSE was 0.2588. These findings may provide a constructive reference for the spectral detection of the infection degree of FHB in the visible areas of in situ wheat canopies. Furthermore, such insights could guide the country to formulate measures to ensure stable grain production and promote the green development of agriculture.

Author Contributions: Y.C.: writing—review and editing, conceptualization, methodology, formal analysis, investigation, data curation, visualization, software. X.W.: funding acquisition, conceptualization, methodology, project administration, supervision. X.Z.: data curation, software, methodology, validation. Y.S.: data curation, software, methodology. H.S.: data curation, methodology. D.W.: data curation, software. X.X.: data curation, software. All authors have read and agreed to the published version of the manuscript.

Funding: This work was funded by the National Key R&D Program (Sino-British Cooperation)—Research on on-line detection technology and sensor of farmland soil, crop chlorophyll, diseases and pests (2019YFE0125201).

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- Hou, S.; Dang, H.; Huang, T.; Huang, Q.; Li, C.; Li, X.; Sun, Y.; Chu, H.; Qiu, W.; Liu, J.; et al. Targeting High Nutrient Efficiency to Reduce Fertilizer Input in Wheat Production of China. *Field Crops Res.* **2023**, *292*, 108809. [\[CrossRef\]](#)
- Zhang, Z.; Li, Z.; Chen, Y.; Zhang, L.; Tao, F. Improving Regional Wheat Yields Estimations by Multi-Step-Assimilating of a Crop Model with Multi-Source Data. *Agric. For. Meteorol.* **2020**, *290*, 107993. [\[CrossRef\]](#)
- Jing-Song, S.; Guang-Sheng, Z.; Xing-Hua, S. Climatic Suitability of the Distribution of the Winter Wheat Cultivation Zone in China. *Eur. J. Agron.* **2012**, *43*, 77–86. [\[CrossRef\]](#)
- Xu, X.; He, P.; Chuan, L.; Liu, X.; Liu, Y.; Zhang, J.; Huang, X.; Qiu, S.; Zhao, S.; Zhou, W. Regional Distribution of Wheat Yield and Chemical Fertilizer Requirements in China. *J. Integr. Agric.* **2021**, *20*, 2772–2780. [\[CrossRef\]](#)
- Wang, H.; Sun, S.; Ge, W.; Zhao, L.; Hou, B.; Wang, K.; Lyu, Z.; Chen, L.; Xu, S.; Guo, J.; et al. Horizontal Gene Transfer of Fhb7 from Fungus Underlies Fusarium Head Blight Resistance in Wheat. *Science* **2020**, *368*, eaba5435. [\[CrossRef\]](#) [\[PubMed\]](#)
- Mang, H.; Wang, Y.; Gao, Y.; He, Y.; Jiang, P.; Wu, L.; Zhang, X. Review and prospect on the breeding for the resistance to Fusarium Head Blight in wheat. *Sci. Agric. Sin.* **2022**, *55*, 837–855.
- Huang, C.; Jiang, Y.; Li, C. Occurrence, yield loss and dynamics of wheat disease and insect pests in China from 1987 to 2018. *Plant Prot.* **2020**, *46*, 186–193.
- NY/T 1443.4-2007; Rules for Resistance Evaluation of Wheat to Diseases and Insect Pests Part4: Rule for Resistance Evaluation of Wheat to Wheat Scab. Agricultural Ministry of the People's Republic of China: Beijing, China; Standards Press of China: Beijing, China, 2017.
- Abbas, A.; Yli-Mattila, T. Biocontrol of Fusarium Graminearum, a Causal Agent of Fusarium Head Blight of Wheat, and Deoxynivalenol Accumulation: From In Vitro to In Planta. *Toxins* **2022**, *14*, 299. [\[CrossRef\]](#)
- Wang, D. Grading Disease Severity and Developing Diagnosis System of Wheat Fusarium Head Blight Based on Computer Vision Technology. Master's Thesis, Anhui University, Anhui, China, 2020.
- Bao, W.; Yang, X.; Liang, D.; Hu, G.; Yang, X. Lightweight Convolutional Neural Network Model for Field Wheat Ear Disease Identification. *Comput. Electron. Agric.* **2021**, *189*, 106367. [\[CrossRef\]](#)

12. Ding, J.; Liang, K.; Han, D.; Xu, J.; Shen, M. Detection of Vomiting Toxin Content in Wheat Scab Seeds by Near-Infrared Hyperspectral Based on ICO-SPA Feature Extraction. *J. Triticeae Crops* **2019**, *39*, 867–876.
13. Jiang, X.; Zhang, B.; Zhao, T.; Xiong, C.; Shen, F.; Hei, X.; Liu, Q.; Zhou, H.; Liu, X. Screening of DON Contamination in Wheat Based on Visible/Near Infrared Spectroscopy. *Spectrosc. Spectr. Anal.* **2019**, *39*, 3904–3909.
14. Zhang, N.; Yang, G.; Zhao, C.; Zhang, J.; Yang, X.; Pan, Y.; Huang, W.; Xu, B.; Li, M.; Zhu, X.; et al. Progress and prospects of hyperspectral remote sensing technology for crop diseases and pests. *Natl. Remote Sens. Bull.* **2021**, *25*, 403–422.
15. Qingqing, H.; Ling, J.; Zhenghua, Z.; Shu, J.; Chen, G.; Wei, M.; Wenxi, L.; Tao, L.; Bin, L.; Changwei, T. A Lightweight Model for Wheat Ear Fusarium Head Blight Detection Based on RGB Images. *Remote Sens.* **2022**, *14*, 3481.
16. da Conceição, R.R.P.; Simeone, M.L.F.; Queiroz, V.A.V.; de Medeiros, E.P.; de Araújo, J.B.; Coutinho, W.M.; da Silva, D.D.; de Araújo Miguel, R.; de Paula Lana, U.G.; de Resende Stoianoff, M.A. Application of Near-Infrared Hyperspectral (NIR) Images Combined with Multivariate Image Analysis in the Differentiation of Two Mycotoxicogenic Fusarium Species Associated with Maize. *Food Chem.* **2020**, *344*, 128615. [[CrossRef](#)]
17. Bai, X.; Yu, J.; Fu, Z.; Zhang, L.; Li, X. Application of Spectral Imaging Technology for Detecting Crop Disease Information: A Review. *Spectrosc. Spectr. Anal.* **2020**, *40*, 350–355.
18. Baraa, A.M.; Krishnaswamy, R.A.; Whetton, R.L.; Damien, V.; Damien, E.; Philippe, V.; Mouazen, A.M. Detection of Fusarium Head Blight in Wheat under Field Conditions Using a Hyperspectral Camera and Machine Learning. *Comput. Electron. Agric.* **2022**, *203*, 107456.
19. Lv, Y.; Lv, W.; Han, K.; Tao, W.; Zheng, L.; Weng, S.; Huang, L. Determination of Wheat Kernels Damaged by Fusarium Head Blight Using Monochromatic Images of Effective Wavelengths from Hyperspectral Imaging Coupled with an Architecture Self-Search Deep Network. *Food Control* **2022**, *135*, 108819.
20. Huang, L.; Wu, Z.; Huang, W.; Ma, H.; Zhao, J. Identification of Fusarium Head Blight in Winter Wheat Ears Based on Fisher's Linear Discriminant Analysis and a Support Vector Machine. *Appl. Sci.* **2019**, *9*, 3894. [[CrossRef](#)]
21. Huang, L.; Li, T.; Ding, C.; Zhao, J.; Zhang, D.; Yang, G. Diagnosis of the Severity of Fusarium Head Blight of Wheat Ears on the Basis of Image and Spectral Feature Fusion. *Sensors* **2020**, *20*, 2887. [[CrossRef](#)]
22. Huang, L.; Wu, K.; Huang, W.; Dong, Y.; Ma, H.; Liu, Y.; Liu, L. Detection of Fusarium Head Blight in Wheat Ears Using Continuous Wavelet Analysis and PSO-SVM. *Agriculture* **2021**, *11*, 998. [[CrossRef](#)]
23. Zhang, D.-Y.; Chen, G.; Yin, X.; Hu, R.-J.; Gu, C.-Y.; Pan, Z.-G.; Zhou, X.-G.; Chen, Y. Integrating Spectral and Image Data to Detect Fusarium Head Blight of Wheat. *Comput. Electron. Agric.* **2020**, *175*, 105588. [[CrossRef](#)]
24. Zhang, D.; Wang, Q.; Lin, F.; Yin, X.; Gu, C.; Qiao, H. Development and Evaluation of a New Spectral Disease Index to Detect Wheat Fusarium Head Blight Using Hyperspectral Imaging. *Sensors* **2020**, *20*, 2260. [[CrossRef](#)] [[PubMed](#)]
25. Zhang, N.; Pan, Y.; Feng, H.; Zhao, X.; Yang, X.; Ding, C.; Yang, G. Development of Fusarium Head Blight Classification Index Using Hyperspectral Microscopy Images of Winter Wheat Spikelets. *Biosyst. Eng.* **2019**, *186*, 83–99. [[CrossRef](#)]
26. Jin, X.; Jie, L.; Wang, S.; Qi, H.J.; Li, S.W. Classifying Wheat Hyperspectral Pixels of Healthy Heads and Fusarium Head Blight Disease Using a Deep Neural Network in the Wild Field. *Remote Sens.* **2018**, *10*, 395. [[CrossRef](#)]
27. Jin, X.; Lu, J.; Fu, Y.; Wang, S.; Xu, G.; Li, S. A classification method for hyperspectral imaging of Fusarium head blight disease symptom based on deep convolutional neural network. *Acta Agric. Zhejiangensis* **2019**, *31*, 315–325.
28. Bauriegel, E.; Giebel, A.; Geyer, M.; Schmidt, U.; Herppich, W.B. Early Detection of Fusarium Infection in Wheat Using Hyperspectral Imaging. *Comput. Electron. Agric.* **2011**, *75*, 304–312. [[CrossRef](#)]
29. Liu, L. Research on the Methods of Wheat Fusarium Head Blight and Powdery Mildew Monitoring Using Remote Sensing Technology at Different Scales. Ph.D. Thesis, University of Chinese Academy of Sciences, Beijing, China, 2020.
30. Han, D. Research on Detection Method of Wheat Ear Scab Infection Based on Hyperspectral Imaging Technology. Master's Thesis, Nanjing Agricultural University, Nanjing, China, 2020.
31. Želazny, W.R.; Chrpová, J.; Hamouz, P. Fusarium Head Blight Detection from Spectral Measurements in a Field Phenotyping Setting—A Pre-Registered Study. *Biosyst. Eng.* **2021**, *211*, 97–113. [[CrossRef](#)]
32. Whetton, R.L.; Hassall, K.L.; Waine, T.W.; Mouazen, A.M. Hyperspectral Measurements of Yellow Rust and Fusarium Head Blight in Cereal Crops: Part 1: Laboratory Study. *Biosyst. Eng.* **2018**, *166*, 101–115. [[CrossRef](#)]
33. Whetton, R.L.; Waine, T.W.; Mouazen, A.M. Hyperspectral Measurements of Yellow Rust and Fusarium Head Blight in Cereal Crops: Part 2: On-Line Field Measurement. *Biosyst. Eng.* **2018**, *167*, 144–158. [[CrossRef](#)]
34. Cao, S.; Yang, H.; Leng, S.; Shu, Y.; Deng, Y.; Sun, H.; Zhang, A.; Zhang, Y.; Li, W.; Chen, H. Identification of the Resistance to Fusarium Head Blight and Toxin Accumulation of Major Wheat Cultivars in Jiangsu Province. *J. Triticeae Crops* **2022**, *42*, 958–968.
35. Arata, G.J.; Martínez, M.; Elguezabal, C.; Rojas, D.; Cristos, D.; Dinolfo, M.I.; Arata, A.F. Effects of Sowing Date, Nitrogen Fertilization, and Fusarium Graminearum in an Argentinean Bread Wheat: Integrated Analysis of Disease Parameters, Mycotoxin Contamination, Grain Quality, and Seed Deterioration. *J. Food Compos. Anal.* **2022**, *107*, 104364. [[CrossRef](#)]
36. Feng, S.; Zhao, D.; Guan, Q.; Li, J.; Liu, Z.; Jin, Z.; Li, G.; Xu, T. A Deep Convolutional Neural Network-Based Wavelength Selection Method for Spectral Characteristics of Rice Blast Disease. *Comput. Electron. Agric.* **2022**, *199*, 107199. [[CrossRef](#)]
37. Gai, Z.; Sun, L.; Bai, H.; Li, X.; Wang, J.; Bai, S. Convolutional Neural Network for Apple Bruise Detection Based on Hyperspectral. *Spectrochim. Acta Part A Mol. Biomol. Spectrosc.* **2022**, *279*, 121432. [[CrossRef](#)]
38. Kang, L.; Yuan, J.; Gao, R.; Kong, Q.; Jia, Y.; Su, Z. Early Detection and Identification of Rice Blast Based on Hyperspectral Image. *Spectrosc. Spectr. Anal.* **2021**, *41*, 898–902.

39. GB/T15796-2011; Rules for Monitoring and Forecast of the Wheat Head Blight. Agricultural Ministry of the People's Republic of China: Beijing, China; Standards Press of China: Beijing, China, 2011.
40. Khodabakhshian, R.; Bayati, M.R.; Emadi, B. Adulteration Detection of Sudan Red and Metanil Yellow in Turmeric Powder by NIR Spectroscopy and Chemometrics: The Role of Preprocessing Methods in Analysis. *Vib. Spectrosc.* **2022**, *120*, 103372. [\[CrossRef\]](#)
41. Yao, H.; Lewis, D. CHAPTER 2—Spectral Preprocessing and Calibration Techniques. In *Hyperspectral Imaging for Food Quality Analysis and Control*; Sun, D.-W., Ed.; Academic Press: San Diego, CA, USA, 2010; pp. 45–78, ISBN 978-0-12-374753-2.
42. Jiao, Y.; Li, Z.; Chen, X.; Fei, S. Preprocessing Methods for Near-Infrared Spectrum Calibration. *J. Chemom.* **2020**, *34*, e3306. [\[CrossRef\]](#)
43. Bian, X. Spectral Preprocessing Methods. In *Chemometric Methods in Analytical Spectroscopy Technology*; Chu, X., Huang, Y., Yun, Y.-H., Bian, X., Eds.; Springer: Singapore, 2022; pp. 111–168, ISBN 978-981-19162-5-0.
44. Shen, S.S. Optimal Band Selection and Utility Evaluation for Spectral Systems. In *Hyperspectral Data Exploitation*; John Wiley & Sons, Ltd.: Hoboken, NJ, USA, 2007; pp. 227–243, ISBN 978-0-470-12462-8.
45. Chang, C.-I. Hyperspectral Information Compression. In *Hyperspectral Data Processing*; John Wiley & Sons, Ltd.: Hoboken, NJ, USA, 2013; pp. 541–544, ISBN 978-1-118-26978-7.
46. Bajcsy, P.; Groves, P. Methodology for Hyperspectral Band Selection. *Photogramm. Eng. Remote Sens.* **2004**, *70*, 793–802. [\[CrossRef\]](#)
47. Sonobe, R.; Hirono, Y. Applying Variable Selection Methods and Preprocessing Techniques to Hyperspectral Reflectance Data to Estimate Tea Cultivar Chlorophyll Content. *Remote Sens.* **2023**, *15*, 19. [\[CrossRef\]](#)
48. Araújo, M.C.U.; Saldanha, T.C.B.; Galvão, R.K.H.; Yoneyama, T.; Chame, H.C.; Visani, V. The Successive Projections Algorithm for Variable Selection in Spectroscopic Multicomponent Analysis. *Chemom. Intell. Lab. Syst.* **2001**, *57*, 65–73. [\[CrossRef\]](#)
49. Soares, S.F.C.; Gomes, A.A.; Araujo, M.C.U.; Filho, A.R.G.; Galvão, R.K.H. The Successive Projections Algorithm. *TrAC Trends Anal. Chem.* **2013**, *42*, 84–98. [\[CrossRef\]](#)
50. Pang, L.; Wang, L.; Yuan, P.; Yan, L.; Xiao, J. Rapid Seed Viability Prediction of Sophora Japonica by Improved Successive Projection Algorithm and Hyperspectral Imaging. *Infrared Phys. Technol.* **2022**, *123*, 104143. [\[CrossRef\]](#)
51. Li, H.; Liang, Y.; Xu, Q.; Cao, D. Key Wavelengths Screening Using Competitive Adaptive Reweighted Sampling Method for Multivariate Calibration. *Anal. Chim. Acta* **2009**, *648*, 77–84. [\[CrossRef\]](#) [\[PubMed\]](#)
52. Kumar, K. Competitive Adaptive Reweighted Sampling Assisted Partial Least Square Analysis of Excitation-Emission Matrix Fluorescence Spectroscopic Data Sets of Certain Polycyclic Aromatic Hydrocarbons. *Spectrochim. Acta Part A Mol. Biomol. Spectrosc.* **2021**, *244*, 118874. [\[CrossRef\]](#)
53. Reeves, C.R. *Genetic Algorithms*; Liu, L., Özsu, M.T., Eds.; Springer: Boston, MA, USA, 2009; pp. 1224–1227.
54. Kramer, O. *Genetic Algorithm Essentials*; Springer International Publishing: Cham, Switzerland, 2017; Volume 679.
55. Caredda, M.; Mara, A.; Ciulu, M.; Floris, I.; Pilo, M.I.; Spano, N.; Sanna, G. Use of Genetic Algorithms in the Wavelength Selection of FT-MIR Spectra to Classify Unifloral Honeys from Sardinia. *Food Control* **2023**, *146*, 109559. [\[CrossRef\]](#)
56. Centner, V.; Massart, D.L.; de Noord, O.E.; de Jong, S.; Vandeginste, B.M.; Sterna, C. Elimination of Uninformative Variables for Multivariate Calibration. *Anal. Chem.* **1996**, *68*, 3851–3858. [\[CrossRef\]](#) [\[PubMed\]](#)
57. Han, Q.-J.; Wu, H.-L.; Cai, C.-B.; Xu, L.; Yu, R.-Q. An Ensemble of Monte Carlo Uninformative Variable Elimination for Wavelength Selection. *Anal. Chim. Acta* **2008**, *612*, 121–125. [\[CrossRef\]](#) [\[PubMed\]](#)
58. Song, X.; Huang, Y.; Tian, K.; Min, S. Near Infrared Spectral Variable Optimization by Final Complexity Adapted Models Combined with Uninformative Variables Elimination—a Validation Study. *Optik* **2020**, *203*, 164019. [\[CrossRef\]](#)
59. Geladi, P.; Kowalski, B.R. Partial Least-Squares Regression: A Tutorial. *Anal. Chim. Acta* **1986**, *185*, 1–17. [\[CrossRef\]](#)
60. Mehmood, T.; Liland, K.H.; Snipen, L.; Sæbø, S. A Review of Variable Selection Methods in Partial Least Squares Regression. *Chemom. Intell. Lab. Syst.* **2012**, *118*, 62–69. [\[CrossRef\]](#)
61. Xie, Z.; Feng, X.; Chen, X. Partial Least Trimmed Squares Regression. *Chemom. Intell. Lab. Syst.* **2022**, *221*, 104486. [\[CrossRef\]](#)
62. Chu, X.; Yuan, H.; Lu, W. Progress and Application of Spectral Data Pretreatment and Wavelength Selection Methods in NIR Analytical Technique. *Prog. Chem.* **2004**, *16*, 528–542.
63. Sun, W.; Wang, R.; Gao, R.; Li, Q.; Wu, H.; Feng, L. Crop Disease Recognition Based on Visible Spectrum and Improved Attention Module. *Spectrosc. Spectr. Anal.* **2022**, *42*, 1572–1580.

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.