agriculture

MDPI

Article

Potato Late Blight Severity and Epidemic Period Prediction Based on Vis/NIR Spectroscopy

Bingru Hou [1], Yaohua Hu [2], Peng Zhang [1] and Lixia Hou [1,*]

[1] College of Mechanical and Electronic Engineering, Northwest A&F University, Xianyang 712100, China; houbr@nwafu.edu.cn (B.H.); zhangpe185@nwafu.edu.cn (P.Z.)
[2] College of Optical, Mechanical and Electrical Engineering, Zhejiang A&F University, Hangzhou 311300, China; huyaohua@zafu.edu.cn
* Correspondence: houlixia@nwafu.edu.cn

Abstract: Late blight caused by *Phytophthora infestans* is a destructive disease in potato production, which can lead to crop failure in severe cases. This study combined visible/near-infrared (Vis/NIR) spectroscopy with machine learning (ML) and chemometric methods for rapid detection of potato late blight. The determination of disease severity was accomplished by two methods directly or indirectly based on differences in reflectance. One approach was to utilize ML algorithms to build a model that directly reflects the relationship between disease level and spectral reflectance. Another method was to first use partial least squares to construct a predictive model of internal physicochemical values, such as relative chlorophyll content (SPAD) and peroxidase (POD) activity, and then use an ML model to classify disease levels based on the predicted values. The classification accuracy based on these two methods could reach up to 99 and 95%, respectively. The changes in physicochemical values during the development of disease were further investigated. Regression models for fitting changes in SPAD value and POD activity were developed based on temperature and incubation time, with determination coefficients of 0.961 and 0.997, respectively. The prediction of epidemic period was realized by combining regression and classification models based on physicochemical values with an accuracy of 88.5%. It is demonstrated that rapid non-destructive determination of physicochemical values based on Vis/NIR spectroscopy for potato late blight detection is feasible. Furthermore, it is possible to guide the control of disease throughout the epidemic period.

Keywords: disease classification; machine learning; partial least squares; physicochemical values; potato

Citation: Hou, B.; Hu, Y.; Zhang, P.; Hou, L. Potato Late Blight Severity and Epidemic Period Prediction Based on Vis/NIR Spectroscopy. *Agriculture* **2022**, *12*, 897. https://doi.org/10.3390/agriculture12070897

Academic Editors: Ahmed Mustafa Rady and Ewa Ropelewska

Received: 2 June 2022
Accepted: 16 June 2022
Published: 21 June 2022

1. Introduction

Potato is a multi-purpose product and the fourth largest food crop in the world after wheat, rice and maize [1,2]. Potato occupies an essential position in China's economic development due to its high yield, adaptability and wide distribution. Yet, potato yields are severely constrained by pests and diseases. Late blight is a worldwide potato disease that causes the most severe yield losses [3]. It can affect all parts of the plant, including leaves, stems and tubers, and is a huge threat to potato production [4]. The prevalence of late blight is influenced by environmental factors. It can destroy an entire field in a matter of days with high humidity and certain temperature conditions [5]. Thus, it is important to achieve early detection of the disease and prediction of the epidemic period for the control of potato late blight.

The detection of plant diseases has relied on visual observations by agronomists or laboratory examinations in the past. However, these traditional evaluation methods are time-consuming and labor-intensive [6]. Recently, spectroscopic methods have overcome the drawbacks of traditional methods and have been widely used in the field

for non-destructive detection of crop diseases due to advantages of rapid analysis, non-destructiveness and high sensitivity [7,8]. In response to disease stress, the internal physiological structure and biochemical properties of plants are altered [9]. A certain extent of leaf yellowing, wilting and withering will be exhibited in appearance, along with a change in spectral features [10]. Diseases can be detected at the plant, canopy and larger scale of crop based on the differences in spectral features [11].

A number of sensitive wavelength bands are more obvious in response to plant disease infection, and the spectral reflectance of healthy and disease-infested plants differs significantly at these sensitive regions. Typically, most studies use chemometrics or machine learning methods to build models of spectral information on the disease level, which define the severity of disease directly based on changes in spectral reflectance. There are a number of relevant studies on potato disease detection. Duarte-Carvajalino et al. [12] and Rodríguez et al. [13] used several machine learning methods to train multi-spectral data of the canopy to predict the severity of potato late blight. Huang et al. [14] proposed a method for rapid non-destructive detection of hollow heart of potato based on semi-transmission hyperspectral imaging. It first identified eight spectral feature variables using a feature wavelength selection algorithm, and the support vector machine (SVM) model based on these variables had a recognition rate of 94.6% for potato hollow disease. After that, the hyperparameters of the SVM model were optimized with the artificial fish swarm algorithm (AFSA), and the accuracy of the optimized model could reach 100%. Using the SVM model, Griffel et al. [15] evaluated the feasibility of spectral reflectance of five different wavelength bands in the Vis and NIR regions for distinguishing whether plants were infected with potato Y virus. The spectral data in the ranges of 500–900 and 720–1300 nm were found to be preferable for the qualitative determination of potato Y virus, with the highest accuracy of 89.9%. Atherton et al. [16] used partial least squares (PLS) and combined principal component analysis (PCA) and spectral change analysis to determine potato early blight at two different growth stages. PCA successfully distinguished between plants with more severe disease and those with less severe disease, while spectral change analysis provided the seven wavelengths most sensitive to early blight infection. Garhwal et al. [17] analyzed the hyperspectral data by multivariate statistical methods to obtain 34 bands sensitive to potato zebra chip disease, and a partial least squares discriminant analysis (PLS-DA) model based on this optimal set of bands achieved an accuracy of 89%. Han et al. [18] developed a PLS determination model for potato black heart disease based on Vis/NIR spectra, and successfully simplified the model by combining PCA and spectral morphological features to use a combination of six wavelengths. The total accuracy of the simplified model could reach 96.7%. From these studies, it is clear that the method to detect disease directly based on reflectance can achieve the desired effect. It is necessary to first obtain proper feature variables to simplify the model, so dimensional reduction in the spectral data is essential. Since this method only considers the correlation between spectral reflectance and disease level, it can only be used for the quantification of disease severity.

Actually, plant physiological and biochemical properties are constantly changing with the development of diseases. The chlorophyll content is an important indicator to characterize the health status of plants [19,20]. The SPAD value is usually measured as a substitute for chlorophyll content. Moreover, peroxidase (POD) used to evaluate plant stress resistance is constantly changing during plant disease and directly related to the severity of plant disease [21]. Hence, some studies have also attempted to discern the degree of crop infection based on changes in the internal physicochemical values of the leaves. Ban et al. [22] developed a SPAD estimation model based on hyperspectral vegetation index for detecting apple mosaic virus infection. Li and Hu [23] first developed a POD activity prediction model based on hyperspectral data, a kinetic model was then established to fit the changes in POD activity, and finally the severity of potato late blight was evaluated based on the changes. Nevertheless, most of these studies focused on the development of high-performance predictive models for the rapid determination of physicochemical values. In fact, the models are also based on the full band or the characteristic band to achieve the

prediction of the physicochemical values. For example, Kang et al. [24] established a model for estimating SPAD values of rice leaves under rice blast disease stress. Kong et al. [25] rapidly determined POD activity in tomato leaves infected with botrytis cinerea based on hyperspectroscopy. Cheng et al. [26] applied Vis/NIR hyperspectra to detect POD activity in cucumber leaves with early bacterial angular leaf-spot disease stress. These studies had not further analyzed the development of physicochemical values over time. The physicochemical values change continuously with time, and can be applied to evaluate the disease severity. Therefore, the changes in physicochemical values can be used to explore the development and to forecast the epidemic period of the disease.

The main objectives of this study were: (1) to investigate the changes in spectral reflectance and the response of physicochemical values to incubation time and environment conditions during the stages of disease; (2) to establish a classification model based on reflectance and physicochemical values to quantify the disease severity by combining pre-processing, dimensional reduction and classification methods; (3) to predict the epidemic period based on the response of physicochemical values through time.

2. Materials and Methods

2.1. Leaf Samples

Zhongshu No. 18 seed potato, which is a late blight moderately susceptible variety was planted during February 2021 in the farmland of Hanzhong, Shaanxi. A number of leaves in good health and uniform size were collected from potato plants at a late stage of growth (8–10 weeks after emergence) and inoculated with Phytophthora infestans using the ex vivo inoculation method. The concentration of the sporangial suspension used for inoculation was controlled at about 4×10^4/mL. After inoculation, the container in which the leaves were located was sealed and moisturized with cling film. Finally, the leaves were placed in an RGX-250B artificial climate incubator (Kuntian Laboratory Instruments Co., Ltd., Shanghai, China) at a temperature of (17 ± 1) °C for 24 h to ensure successful infestation. The successfully infested leaves were randomly divided into three groups and put into an artificial climate chamber at 15, 20 and 25 °C with 70% humidity for 6 days. The temperature gradients were set from 15 to 25 °C according to the optimum temperature (16~24 °C) for the prevalence of potato late blight. The purpose is to imitate different environmental conditions for disease infestation in order to investigate the effect of temperature on the development of late blight. The severely diseased leaves at the end of the cultivation period had dried out or decayed. As a result, these leaves were not available for spectral collection, so a total of 270 valid leaf samples were eventually used for subsequent experimental studies.

2.2. Data Acquisition and Division

2.2.1. Spectra Collection

A USB4000 Vis/NIR fiber optic spectrometer (Ocean Optics, Inc., Dunedin, FL, USA) in the wavelength range of 350–1000 nm combined with a PRB-D0-IS30 portable reflectance probe (Yunxiang Optoelectronics Technology Co., Ltd., Wuxi, Jiangsu, China) was used to collect spectra from the leaves. The diseased area was selected as the region of interest (ROI) on the leaves. The probe was then used to irradiate the surface of the ROI. During this process, it is important to ensure that the leaves and the probe fit perfectly. Once the spectral curve had stabilized, the spectral reflectance data were saved. The measurement was repeated three times, and then the average value was calculated and taken as the spectral reflectance of the leaf sample.

2.2.2. Measurements of Physicochemical Values

This section includes the measurement of SPAD value and POD activity of the leaves. The measurement of peroxidase activity destroys the leaves, which affects the subsequent experiments. Therefore, we selected a batch of leaves in good health and of uniform size for the experimental work, and the leaves were different for each day of the experiment.

Measurements of chlorophyll content were performed using a SPAD-502 chlorophyll meter (Konica Minolta, Inc., Tokyo, Japan). The determination of peroxidase activity was achieved by using a spectrophotometric method. The process was as follows. The ROI of the leaves was cut off, weighed and ground thoroughly, and the crude extract of peroxidase was obtained by adding the prepared reagents and centrifuged. The absorbance value was then measured by a YMB-7A enzyme standard meter (YENDA Instruments & Equipment Co., Ltd., Xiamen, Fujian, China). The absorbance value at 470 nm was recorded every 1 min three times. The change in absorbance value by 0.01 per minute was defined as one enzyme activity unit. The POD activity was calculated according to Equation (1).

$$\text{POD activity} = \frac{\Delta A_{470} \times V_t}{FW \times V_s \times t \times 0.01} \quad (1)$$

where ΔA_{470} is the amount of change in absorbance at 470 nm; V_t is the total volume of enzyme solution (mL); FW is the fresh weight of the sample (g); V_s is the volume of enzyme solution taken at the time of determination (mL); and t is the time of enzyme action (min).

2.2.3. Disease Classification

The National Standard of the People's Republic of China GB/T 17980-34-2000 stipulates that there are 6 levels of late blight according to the ratio of the disease spot area to the total leaf area, i.e., level 0: 0%; level 1: <5%; level 3: 6 to 10%; level 5: 11 to 20%; level 7: 21 to 50%; and level 9: >50%. The artificial climate chamber was completely enclosed and although the humidity was set at around 70%, the leaves were in a more humid environment. Compared to the field environment, the rate of disease onset was faster and the spread of the spots was consequently faster. Some leaves could be classified as level 3 after only 1 day of cultivation. Thus, this study made slight adjustments to the above grading criteria in relation to the basic symptoms of leaf disease [27], as shown in Table 1. Potato leaves with varying infection levels are shown in Figure 1.

Table 1. Criteria for classification of potato late blight.

Disease Level	Ratio (%)	Symptoms
Healthy (level 0)	0	No disease spots
Mild disease (level 1)	<10	Small, light to dark green, irregularly shaped spots
Moderate disease (level 2)	10~50	Large dark brown lesions on the leaf surface
Severe disease (level 3)	>50	White mold on the leaf surface when the environment is wet, or the whole leaf dries and shrinks when it is dry

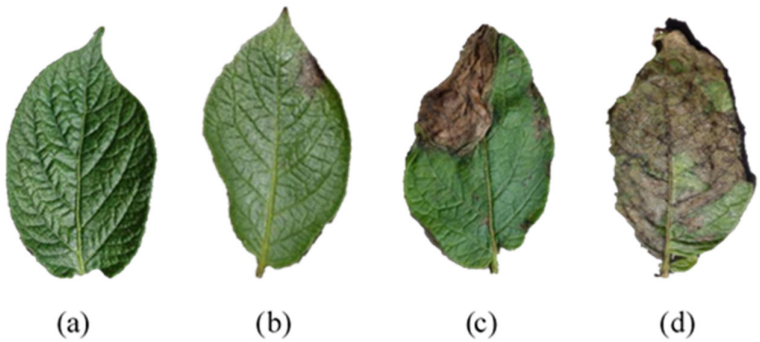

Figure 1. Potato leaves with varying infection levels: (**a**) level 0; (**b**) level 1; (**c**) level 2; (**d**) level 3.

2.2.4. Datasets

Each leaf sample had its own disease level, POD activity, SPAD value, infection time and environmental temperature. This study used the Kennard–Stone (KS) algorithm to divide the data into a calibration set and a prediction set in a ratio of 2:1. The calibration set and the prediction set were used to train the model and to test the generalization ability of the model, respectively. The KS algorithm ensures the uniformity of sample distribution in space [28]. Table 2 shows the statistical results of the physicochemical values for the two sample sets.

Table 2. Statistical results of physicochemical values of the sample set.

Metric	Sets	Number	Range	Mean ± SD
POD activity /U (g min)$^{-1}$	Calibration	180	17.23~138.43	73.27 ± 39.35
	Prediction	90	17.33~138.03	71.63 ± 38.23
SPAD value	Calibration	180	22.20~39.90	32.49 ± 3.74
	Prediction	90	26.90~39.50	32.38 ± 3.32

2.3. Pre-treatment of Spectral Data

The data acquired by the spectrometer includes not only information about the composition of the sample, but also extraneous information such as instrument noise, stray light, background information and baseline drift [29], so the spectral data should first be pre-processed to establish a stable and reliable mathematical model. A number of common methods were adopted for pre-processing. Baseline was used for baseline correction to reduce the interference caused by baseline drift. Mean centering (MC) is a data enhancement transformation method that reinforces the differences between data to improve the representativeness of the model. Moving average smoothing (MA), Savitzky–Golay smoothing (SG) and median filter are all data smoothing approaches that diminish the random errors introduced by the instruments. Multiplicative scatter correction (MSC) and standard normal variate (SNV) can eliminate the errors resulting from the uneven distribution of sample particle size during the spectral data acquisition.

2.4. Dimensional Reduction in Spectral Data

Spectral data are highly dimensional and complex. The large amount of data not only increases the complexity of the modeling, but the large amount of redundant and collinear data it contains can have an impact on the accuracy of the modeling [30–32]. Consequently, the dimensionality reduction process is crucial. Dimensional reduction involves two means of feature extraction and feature selection. Feature extraction is the mapping of the original feature space to a new lower dimensional space, where the feature values are changed. Feature selection is the selection of a subset of features from the original features, with no change to the original feature.

2.4.1. Feature Extraction

Linear discriminant analysis (LDA) and neighborhood components analysis (NCA) utilized in this study are widely used methods for feature extraction. LDA ensures that the intra-class variance of each class is small and the inter-class mean disparity is large. NCA learns the optimal linear transformation matrix for data dimensional reduction by continuously raising the accuracy of the KNN algorithm based on the Mahalanobis distance measurement [33]. They are both supervised learning methods, so they must be combined with ML methods to attain reflectance-based disease severity classification.

2.4.2. Feature Selection

To select effective wavelengths to predict physicochemical values, the performance of competitive adaptive reweighted sampling (CARS), random frog (Frog), uninformative variable elimination (UVE) and random forest (RF) algorithms were contrasted.

CARS is able to select the wavelengths in the PLS model with large correlation coefficients by using adaptive reweighted sampling and an exponentially decreasing function, and uses cross-validation to progressively select the subset with the minimum RMSE [34,35]. The variables in this subset are the selected wavelengths. However, Monte Carlo sampling caused the results to be inconsistent from run to run, so the CARS algorithm was run 100 times. Random frog evaluates the importance by calculating the probability of being selected for each variable [36]. The algorithm was also proposed based on Monte Carlo random sampling. The average probability of selection over 10 runs was therefore taken as the evaluation criterion. UVE adds the random noise matrix to the spectral matrix. The PLS model is established by cross-validation and the regression coefficient matrix is obtained. UVE analyzes the stability of wavelength variables based on the coefficient [37]. Random forest was performed to assess the importance of variables based on the mean decrease in impurity [38].

Except for the CARS algorithm, the remaining three algorithms only reflected the importance of wavelengths to the model and could not explicitly provide the set of selected wavelengths. Generally, a threshold is set as a criterion, and then features with importance higher than the threshold are retained. Obviously, the threshold value affects the number of filtered features and thus the prediction capability. Therefore, the threshold cannot be set randomly and artificially, especially when the features do not differ much in importance. To tackle this problem, a genetic algorithm was applied to optimize the number of features and maximize the prediction performance.

2.5. Partial Least Squares Regression

PLS is a widely used multivariate statistical method for chemometric analysis. It allows for severe collinearity among variables and the number of samples can be less than the number of variables [39]. However, PLS is sensitive to outliers, so outlier detection is carried out using Monte Carlo sampling prior to modeling [40]. Outlier samples were tested based on POD activity in developing a prediction model of physicochemical values using PLS because of the large error in POD activity under the same experimental conditions. The samples with high mean or standard deviation are considered as outliers, which need to be removed for the robustness of the model. The essential aspect to be considered in establishing a PLS model is the determination of the number of latent variables (LVs). This was tested through multiple modeling. The LVs were gradually increased from 1, and a 10-fold cross-validation was performed to build the model. The LVs corresponding to the minimum $RMSE_p$ were adopted as the final number of variables involved in the model [41]. The coefficient of determination of the prediction set ($R_p{}^2$) (Equation (2)) and the root mean square error ($RMSE_P$) (Equation (3)) are generally used as criteria for evaluating the performance of PLS models.

$$RMSE_p = \sqrt{\frac{1}{N}\sum_{1}^{N}(y_i - \hat{y}_i)^2} \tag{2}$$

$$R_p{}^2 = 1 - \frac{\sum_{1}^{N}(y_i - \hat{y}_i)^2}{\sum_{1}^{N}(y_i - \overline{y})^2} \tag{3}$$

where N is the total number of samples in the prediction set; y_i is the true value of metrics; $\hat{y}_i$ is the predicted value of metrics; and $\overline{y}$ is the mean of metrics.

2.6. Machine Learning Algorithms

In this study, ML algorithms were used to explore the response of physicochemical values with time series and environmental conditions during disease development stages, in addition to classifying disease severity based on reflectance and physicochemical values.

Therefore, some common methods were chosen for modeling, such as support vector machines (SVM), k-nearest neighbor (KNN) and decision trees (DT). The methods can be applied to both classification and regression. All ML methods used in this study were implemented in the Scikit-learn (sklearn) library based on Python.

The discrimination of disease level based on spectral reflectance can be regarded as a multi-classification problem. The number of samples in each class was not exactly equal. In addition to using the accuracy (ACC) for measuring the classification performance of the model on global samples, weighted precision (P) (Equation (4)) and weighted recall (R) (Equation (5)) were needed for further evaluation.

$$P = \sum_{1}^{n} W_i P_i \tag{4}$$

$$R = \sum_{1}^{n} W_i R_i \tag{5}$$

$$P_i = \frac{TP_i}{TP_i + FP_i} \tag{6}$$

$$R_i = \frac{TP_i}{TP_i + FN_i} \tag{7}$$

where n is the total number of classes; W_i, P_i and R_i are the corresponding weights, precision and recall for each class, respectively; TP_i, FP_i and FN_i are the number of true positive, false positive and false negative samples corresponding to each class, respectively. In fact, a regression model can be fitted to the changes in the physicochemical values to reflect their response with cultivation time and temperature. The $R_p{}^2$ and the $RMSE_P$ were also adopted to evaluate the performance of the regression model.

2.7. Double Factor Variance Analysis

A two-factor ANOVA was performed to decompose the differences in the physicochemical values and to infer whether the temperature or time factors significantly influenced the changes in the physicochemical values. The temperature factor consists of 3 levels: 15, 20 and 25 °C. There are 7 levels of time, namely 0, 1, 2, 3, 4, 5 and 6 d.

3. Results and Discussion

3.1. Response of Spectral Reflectance and Physicochemical Values at Different Disease Stages

3.1.1. Analysis of Spectral Curve Characteristics

Spectral data in the 500~850 nm band were selected for analysis. The average spectra of the leaves at four disease levels are shown in Figure 2. The solid line in Figure 2 is the average reflectance curve; the corresponding highlighted part is the error band of the reflectance curve. As can be seen in Figure 2, the spectral curves have a significant reflection peak at 550 nm and maintain a high reflectance in the NIR region. The reflectance spectral features at 550 nm are dominated by chlorophyll content, while those in the NIR band are determined by a combination of leaf cell structure and canopy structure. With the severity of the disease, the chlorophyll content in the leaves declined, the reflection of green light became weaker, the cell walls of the leaves were destroyed [42], the intercellular gaps became larger, and the multiple reflections were weakened, so the reflectance at 550 nm and in the NIR spectral band both decreased.

3.1.2. Analysis of the Changes in Physicochemical Values

Figure 3 shows the trends of POD activity (Figure 3a) and SPAD value (Figure 3b) with cultivation time at different temperature conditions. It can be seen that when the leaves were successfully infested, referring to the first day of inoculation, there was a slight decrease in POD activity and a slight increase in SPAD value. Thereafter, as time passed, the POD activity showed a significant upward trend and the SPAD value was obviously

decreasing. Chlorophyll content tends to decrease with disease stress, which leads to a drop in SPAD values. The change in POD activity is related to the defense system in the plant. The decline in POD activity at the beginning of the disease is due to the disruption of the metabolic balance in the plant by the pathogen, and the addition afterward is due to the production of large amounts of POD enzymes in the plant to remove the reactive oxygen damage caused by the pathogen [43,44]. In addition, the rate of changes in physicochemical values at different temperature conditions was not the same.

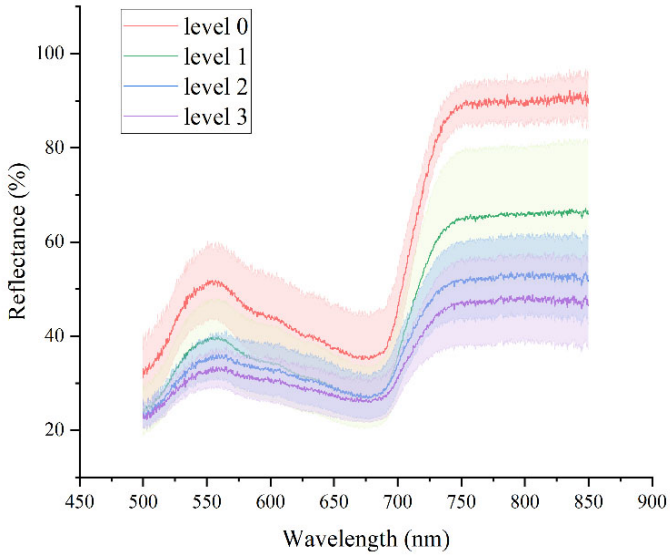

Figure 2. The average spectra of leaves at four disease levels.

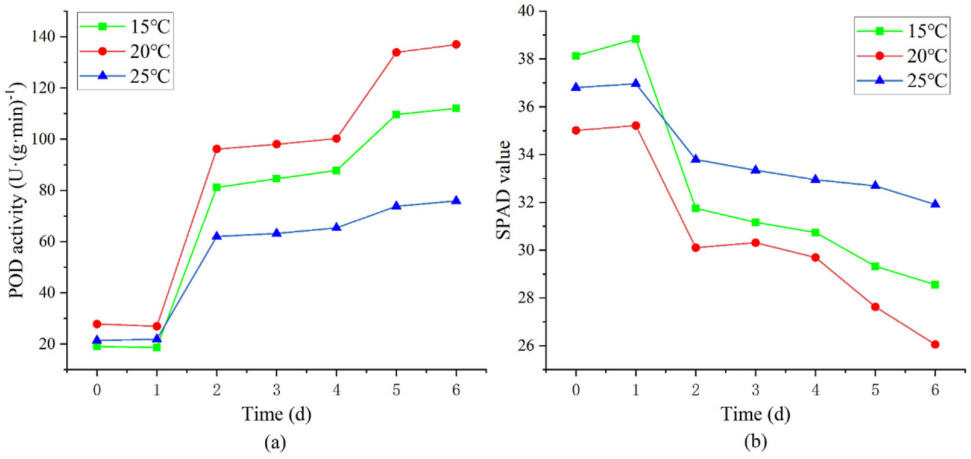

Figure 3. Trends of physicochemical values with time at different temperature conditions: (**a**) POD activity; (**b**) SPAD value.

Further, a two-factor ANOVA was performed on the two physicochemical values and the results are shown in Figure 4. The results indicate that the different temperature treatments caused significant differences ($p < 0.05$) between the two physicochemical values, and the rate of change in physicochemical values was the fastest at 20 °C and the slowest at 25 °C. Significant differences were found between the physicochemical values of infected and healthy leaves at all times except the first day. This demonstrates that temperature and time are two important factors that influence the changes in physicochemical values.

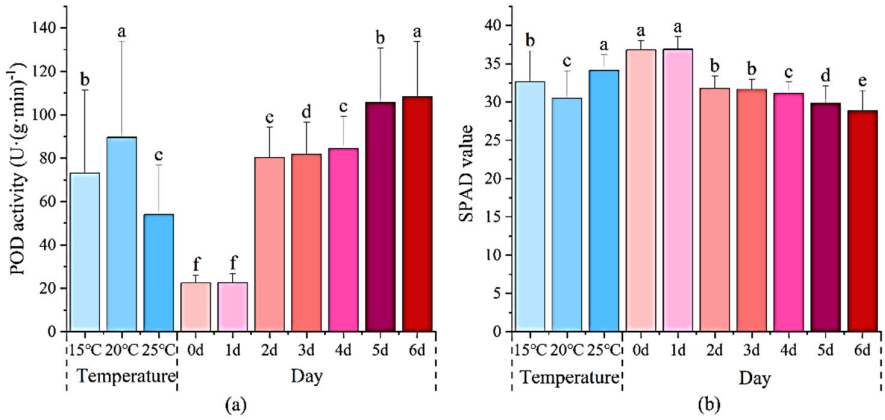

Figure 4. Results of two-factor ANOVA for physicochemical values: (**a**) POD activity; (**b**) SPAD value. Note: The letters a–f in the figure represent the results of the differences between the data. The same letters indicate non-significant differences between data, while different letters mean significant differences between data.

3.2. Optimal Pre-Processing Method

3.2.1. Pre-Processing Method for Spectral Reflectance-Based Classification Models

Figure 5 compares the performance of multiple pre-processing methods combined with SVM, KNN and DT algorithms for disease-degree classification. The performance was assessed by the classification accuracy. The pre-processing methods that could improve the classification performance varied for different ML algorithms. The MC and MF methods slightly improved the accuracy of the SVM model. However, MC caused the classification ability of the remaining two models to be severely degraded. MA and SG enhanced the performance of the KNN model. Baseline, MA, SG and MF all led to higher accuracy of the DT model. Comparing all of the pre-processing methods, MA and MF both promoted the performance of two ML models, MSC and SNV were worse than no treatment, especially MSC. Since MA reduced the performance of the SVM model, so MF was eventually utilized to combine the ML method for classification. In addition, the same pre-processing method combined with multiple classification models exhibited significant differences in the accuracy. It is shown that the choice of classification algorithm is more important than the pre-processing method.

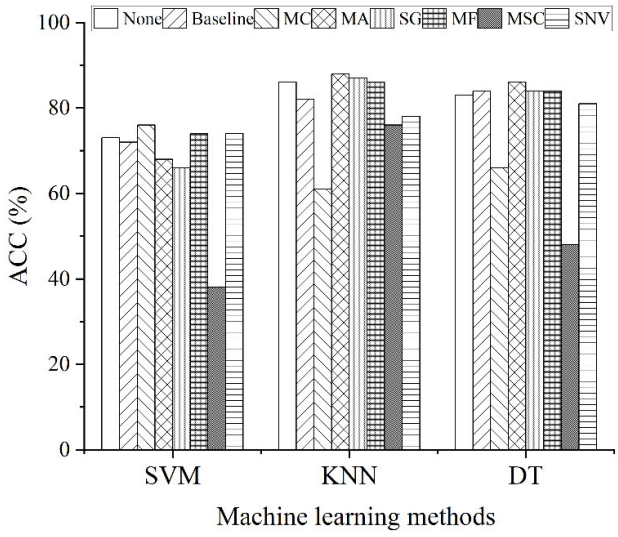

Figure 5. Comparison of the classification performance of different pre-processing methods combined with machine learning models.

3.2.2. Pre-Processing Methods for Predictive Models of Physicochemical Values

Table 3 shows the effects of the prediction models for POD activity and SPAD values based on the raw and pre-treated spectra. Among them, the number of latent variables differed due to the size and distribution of the physicochemical values and the pre-treatment methods. As can be seen in Table 3, the pre-treatment methods did not all improve the performance of the models. The baseline, MSC and SNV pre-treatment methods all reduced the predictive ability of the model for both physicochemical values. The models treated by MC had the same results in predicting the calibration set, but became slightly less effective for the prediction set. For the POD activity prediction model, only the MF method improved the performance. MA and SG improved the performance of the SPAD value prediction model, but the fit of MA to the calibration set became worse. Therefore, the spectral data were processed using MF and SG before building the POD activity and SPAD value prediction models, respectively.

Table 3. Performance of models based on different pre-treatment methods for predicting physicochemical values.

Methods	POD Activity					SPAD Value				
	LVs	R_c^2	R_p^2	$RMSE_c$	$RMSE_p$	LVs	R_c^2	R_p^2	$RMSE_c$	$RMSE_p$
None	10	0.995	0.965	2.662	7.186	10	0.996	0.978	0.252	0.542
Baseline	10	0.995	0.957	2.776	7.973	11	0.997	0.973	0.209	0.583
MC	10	0.995	0.965	2.662	7.184	10	0.996	0.978	0.252	0.506
MA	12	0.997	0.962	2.178	7.465	11	0.995	0.980	0.261	0.520
SG	12	0.997	0.963	2.247	7.361	12	0.997	0.983	0.194	0.480
MF	11	0.997	0.968	2.159	6.929	11	0.997	0.977	0.201	0.550
MSC	9	0.951	0.893	8.689	13.067	9	0.961	0.873	0.753	1.291
SNV	10	0.975	0.901	6.269	12.352	9	0.972	0.883	0.635	1.173

Note: LVs refers to the number of latent variables; R_c^2 and R_p^2 refer to the coefficients of determination of the calibration set and the prediction set, respectively; $RMSE_c$ and $RMSE_p$ refer to the root mean square error of the calibration set and the prediction set, respectively.

3.3. Effective Spectral Features

3.3.1. Wavelength Extraction

Figure 6 compares the performance of the feature extraction methods combined with the SVM, KNN and DT algorithms for disease-degree classification. The dimensional reduction process greatly improves the classification ability of the model, and the accuracy increases to 99%. From the results, it can be seen that the classification accuracy of all three ML models built on the data extracted by NCA was higher than that of the models based on the LDA-extracted data. The effect of extraction using LDA was better than that of NCA. In addition, there were no major differences in the performance of the classification models after parameter optimization. This illustrates that the option of dimensional reduction, rather than classification algorithm, should be considered first when building classification models based on spectral reflectance.

3.3.2. Wavelength Selection

The performance of the models based on different feature selection methods for predicting physicochemical values is shown in Table 4. For the POD activity prediction model, all feature selection methods enhanced the performance of the model. However, the models built after wavelength selection using UVE and RF had a slightly decreasing predictive capability for SPAD values. This was because the number of wavelengths was dropped by more than 86% after CARS and random frog selection, which effectively eliminated redundant data and retained effective wavelengths. UVE and RF preserved approximately 56 and 28% of the spectral data, respectively, and a secondary selection based on this should be considered.

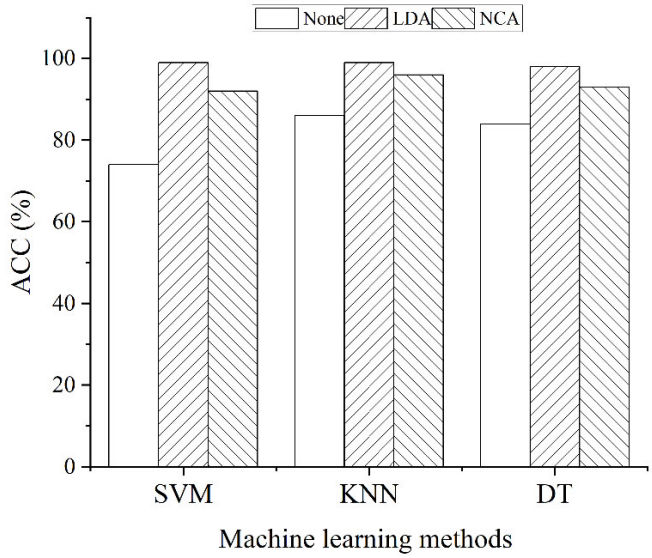

Figure 6. Comparison of the classification performance of different feature extraction methods combined with machine learning models.

Table 4. Performance of models based on different feature selection methods to predict physicochemical values.

Methods	POD Activity					SPAD Value				
	N	R_c^2	R_p^2	$RMSE_c$	$RMSE_p$	N	R_c^2	R_p^2	$RMSE_c$	$RMSE_p$
None	1793	0.997	0.968	2.159	6.929	1793	0.997	0.983	0.194	0.480
CARS	140	0.997	0.989	2.319	3.943	131	0.997	0.995	0.226	0.254
Frog	227	0.999	0.995	1.497	2.811	248	0.996	0.995	0.246	0.254
UVE	857	0.997	0.970	2.153	6.734	999	0.999	0.984	0.127	0.542
RF	479	0.998	0.972	1.642	6.581	507	0.997	0.972	0.202	0.577

Note: N (number) refers to the number of selected wavelengths.

The POD activity prediction model established by spectral data filtered by random frog was the most effective, with 87.3% reduction in the number of wavelengths and 59.4% decrease in $RMSE_p$. Figure 7 shows the average probability of selection for each wavelength obtained by random frog, and the wavelengths selected by an iterative search using GA correspond to a probability threshold of 0.15. The SPAD value is best predicted using CARS. The number of wavelengths was decreased by 92.7% and the $RMSE_p$ was reduced by 45.9%. Figure 8 shows the results of the CARS runs. Figure 8a indicates that the number of selected variables decreases gradually as the number of runs increases, and the number of runs was roughly exponentially related to the number of variables. It can be seen from Figure 8b that the RMSECV reached the minimum value after 21 sampling runs. The corresponding selected wavelengths distribution is shown in Figure 9.

It is observed from Figure 9 that the effective wavelengths related to the two physicochemical values were distributed in both the Vis and NIR region. It was mainly concentrated in red (622~780 nm) and green (492~577 nm) in the visible region. In the red region, the number of effective wavelengths associated with the two physicochemical values accounted for 43.6 and 38.1% of the total number, respectively. Of these, 50% of the effective wavelengths in the red region came from the red edge region. The number of effective wavelengths in green region represented 22.5 and 29.0% of the total number, respectively. Surprisingly, these wavelengths (green, red, red-edge, and near-infrared regions) are similarly more sensitive to late blight infestation. There are significant differences in reflectance between

healthy and infected plants in these regions [45,46]. This confirms the feasibility of assessing the level of late blight based on physicochemical values.

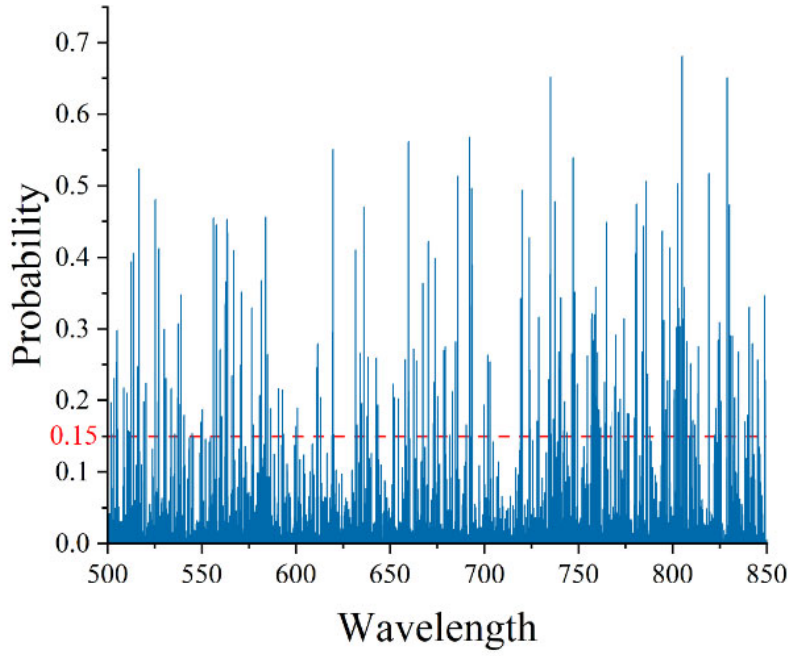

Figure 7. The average selection probability of each wavelength obtained by random frog. Note: The threshold value of the selected probability is 0.15, and wavelengths with probability higher than this threshold will be retained.

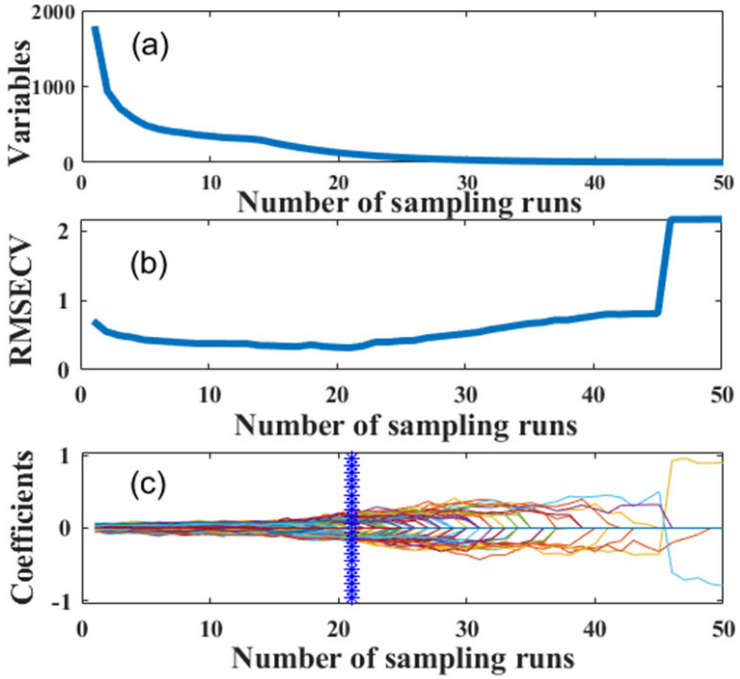

Figure 8. Results of CARS runs: (**a**) the trend of the variables number in the model with the number of sampling runs; (**b**) the trend of the RMSECV with the number of sampling runs; (**c**) the trend of correlation coefficient of each variable in the model with the number of sampling runs.

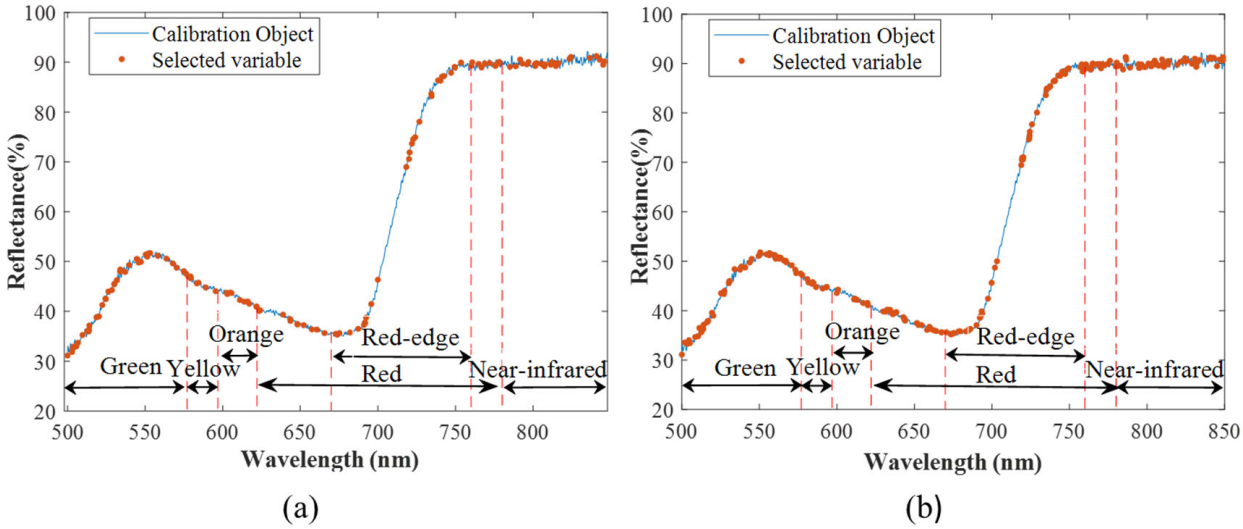

Figure 9. Distribution of sensitive wavelengths associated with physicochemical values: (**a**) POD activity; (**b**) SPAD value.

3.4. Classification Performance

The performance of the classification models based on reflectance and physicochemical values built by combining pre-processing, dimensional reduction and classification methods is shown in Table 5. The process of predicting the physicochemical values using PLS generated bias, which was incorporated into the classification model thereby reducing the accuracy. Thus, this study also uses gradient boosted decision trees (GBDT) for classification, which is one of the ensemble methods that can integrate multiple weak classifiers to improve accuracy. The accuracy of the model built using GBDT was improved to 95%. Table 6 shows the statistics of disease levels of all leaf samples at different temperatures and cultivation periods. As can be seen from Table 6, the rate of disease development was approximately the same under different temperature conditions. The disease developed into level 1 one day after incubation, then became level 2 in 4–5 days and turned into level 3 during the 6th day. It can be seen from Figure 3 that the two physicochemical values of the leaves at the fourth to fifth day of cultivation were apparently different from those at other periods, while the physicochemical values of healthy and level 1 diseased leaves were relatively close. Therefore, the highest recognition rate was achieved for leaves with level 2 disease and the worst for healthy leaves when using GBDT for classification based on physicochemical values. Healthy and level 1 diseased leaves were easily misclassified. The classification performance of the reflectance-based model was better, but both models had the potential to underestimate the severity of disease.

Table 5. The performance of the classification models based on reflectance and physicochemical values built by combining pre-processing, dimensional reduction and classification methods.

Model		**Reflectance-Based**				**Physicochemical Values-Based**			
						POD		**SPAD**	
Method	Pretreatment	MF				MF		SG	
	Reduction	LDA				Frog		CARS	
	Classification	SVM				GBDT			
		level 0	level 1	level 2	level 3	level 0	level 1	level 2	level 3
Metric	P(%)	100	98	100	100	89	96	99	95
	R(%)	100	100	100	92	89	97	97	95
	ACC(%)	99				95			

Table 6. The statistics of disease levels of all leaf samples at different temperatures and cultivation periods.

Temp (°C)	Time (d)	Level	Temp (°C)	Time (d)	Level	Temp (°C)	Time (d)	Level
15	0	0	20	0	0	25	0	0
15	1	1	20	1	1	25	1	1
15	2	1	20	2	1	25	2	1
15	3	1	20	3	1	25	3	1
15	4	1	20	4	2	25	4	2
15	5	2	20	5	2	25	5	2
15	6	3	20	6	3	25	6	3

3.5. Epidemic Period Prediction Based on Physicochemical Values

Temperature and time remarkably affected the changes in physicochemical values during disease development, which in turn indicate disease levels. Thus, the prediction of the epidemic period can be based on the physicochemical values. This study aims to predict when the next stage of the disease is reached. For example, a sample at 15 °C for 2 days is predicted to develop into a level 2 disease after 3 days. Regression models based on temperature and time could be developed to fit the changes in physicochemical values. The performance of the established regression models is shown in Table 7, which shows that the fitting effect of using DT and KNN was comparable and slightly better than SVM. It is possible to predict the epidemic period by combining the regression and classification model based on physicochemical values. Figure 10 shows a confusion matrix to indicate the results of the prediction of the epidemic period. The accuracy is 88.5%. The classification model misclassified the disease level of 12 samples, resulting in a lower accuracy.

Table 7. Performance of regression models fitting the changes in physicochemical values.

Methods	POD Activity				SPAD Value			
	R_c^2	R_p^2	$RMSE_c$	$RMSE_p$	R_c^2	R_p^2	$RMSE_c$	$RMSE_p$
DT	0.998	0.997	1.764	2.286	0.963	0.961	0.693	0.704
KNN	0.998	0.997	1.764	2.287	0.963	0.961	0.693	0.704
SVM	0.997	0.996	1.850	2.382	0.959	0.958	0.725	0.732

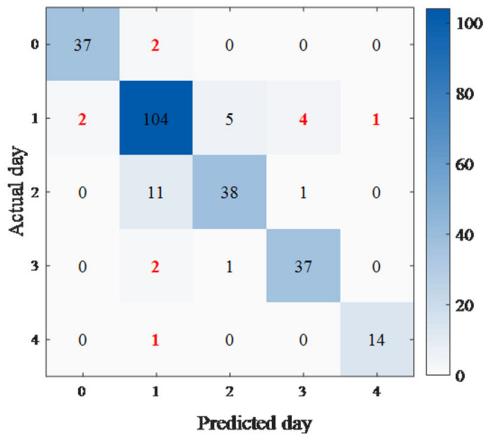

Figure 10. The confusion matrix for the prediction results of the epidemic period. Note: The samples marked in red in the figure are those with misclassified disease levels.

4. Conclusions

This study systematically combined pre-processing, dimensional reduction and classification methods to determine the severity of disease based on spectral reflectance and physicochemical values, and further explored the feasibility of predicting the epidemic period based on changes in physicochemical values. The main conclusions include:

(1) The spectral reflectance and physicochemical values varied with the development of disease. The mean reflectance decreased, the POD activity decreased slightly and then increased significantly, and the SPAD value rose marginally and then declined continuously. It demonstrates that the changes in reflectance and physicochemical values can reflect the disease level.

(2) For the reflectance-based classification model, it is most essential to choose the dimensional reduction method, followed by the classification method. In this case the established MF–LDA–SVM classification model had the best classification performance with an accuracy of 99%.

(3) A high-performing prediction model is a prerequisite for classifying the severity of disease based on physicochemical values. MF for pre-treatment combined with Frog for wavelength selection improved the predictability of POD activity. SG for pre-treatment combined with CARS for wavelength selection led to much better prediction of SPAD values. The grading using GBDT based on the predicted physicochemical values was 95% accurate. In addition, there were significant differences in reflectance at sensitive wavelengths extremely relevant to both physicochemical values under different levels of disease stress. This indicates the feasibility of rapid non-destructive determination of physicochemical values based on Vis/NIR spectroscopy for potato late blight disease classification.

(4) Temperature and time are important factors that influence the changes in physicochemical values. The Rp^2 of the fitted regression models based on these two factors for POD activity and SPAD value were 0.997 and 0.961, respectively. The final prediction of epidemic period was achieved by combining regression and classification models based on physicochemical values with an accuracy of 88.5%.

The potato leaves in this study were infected through human intervention. The disease had a rapid onset rate. Therefore, the use of spray inoculation to infect potato plants could be considered in future work. The development of late blight disease interacting with temperature and humidity should be investigated to provide a theoretical guide for the early detection of potato late blight disease in the field.

Author Contributions: Data curation, methodology, project administration, writing—review and editing, B.H.; funding acquisition and project administration, Y.H.; data curation, P.Z.; project administration, L.H. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the National Natural Science Foundation of China (31971787, 32171894).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare that they have no conflict of interest concerning this research.

References

1. Boivin, M.; Bourdeau, N.; Barnabé, S.; Desgagné-Penix, I. Sprout Suppressive Molecules Effective on Potato (*Solanum tuberosum*) Tubers during Storage: A Review. *Am. J. Potato Res.* **2020**, *97*, 451–463. [CrossRef]
2. Zhang, H.; Xu, F.; Wu, Y.; Hu, H.-H.; Dai, X.-F. Progress of potato staple food research and industry development in China. *J. Integr. Agric.* **2017**, *16*, 2924–2932. [CrossRef]
3. Majeed, A.; Muhammad, Z.; Ullah, Z.; Ullah, R.; Ahmad, H. Late Blight of Potato (*Phytophthora infestans*) I: Fungicides Application and Associated Challenges. *Turk. J. Agric. Food Sci. Technol.* **2017**, *5*, 261–266. [CrossRef]
4. Perez, W.; Forbes, G.A. *Potato Late Blight: Technical Manual*; International Potato Center: Lima, Peru, 2010; Available online: http://www.cipotato.org/publications/pdf/005446.pdf (accessed on 28 April 2016).
5. Campos, H.; Ortiz, O. *The Potato Crop: Its Agricultural, Nutritional and Social Contribution to Humkanhkind*; Springer: New York, NY, USA, 2020. [CrossRef]

6. Khaled, A.Y.; Abd Aziz, S.; Bejo, S.K.; Nawi, N.M.; Seman, I.A.; Onwude, D.I. Early detection of diseases in plant tissue using spectroscopy—Applications and limitations. *Appl. Spectrosc. Rev.* **2017**, *53*, 36–64. [CrossRef]
7. Ali, M.M.; Bachik, N.A.; Muhadi, N.; Yusof, T.N.T.; Gomes, C. Non-destructive techniques of detecting plant diseases: A review. *Physiol. Mol. Plant Pathol.* **2019**, *108*, 101426. [CrossRef]
8. Zahir, S.A.D.M.; Omar, A.F.; Jamlos, M.F.; Azmi, M.A.M.; Muncan, J. A review of visible and near-infrared (Vis-NIR) spectroscopy application in plant stress detection. *Sens. Actuators A Phys.* **2022**, *338*, 113468. [CrossRef]
9. Xie, Y.P.; Chen, F.N.; Zhang, J.C.; Zhou, B.; Wu, K.H. Study on monitoring of common diseases of crops based on hyper-spectral technology. *Spectrosc. Spectr. Anal.* **2018**, *38*, 2233–2240. [CrossRef]
10. Mora-Romero, G.A.; Félix-Gastélum, R.; Bomberger, R.A.; Romero-Urías, C.; Tanaka, K. Common potato disease symptoms: Ambiguity of symptom-based identification of causal pathogens and value of on-site molecular diagnostics. *J. Gen. Plant Pathol.* **2022**, *88*, 89–104. [CrossRef]
11. Zhang, D.R.; Fang, H.; He, Y. Research of crop disease based on visible/near infrared spectral image technology: A review. *Spectrosc. Spectr. Anal.* **2019**, *39*, 1748–1756. [CrossRef]
12. Duarte-Carvajalino, J.M.; Alzate, D.F.; Ramirez, A.A.; Santa-Sepulveda, J.D.; Fajardo-Rojas, A.E.; Soto-Suárez, M. Evaluating Late Blight Severity in Potato Crops Using Unmanned Aerial Vehicles and Machine Learning Algorithms. *Remote Sens.* **2018**, *10*, 1513. [CrossRef]
13. Rodríguez, J.; Lizarazo, I.; Prieto, F.; Angulo-Morales, V. Assessment of potato late blight from UAV-based multispectral imagery. *Comput. Electron. Agric.* **2021**, *184*, 1513. [CrossRef]
14. Huang, T.; Li, X.-Y.; Xu, M.-L.; Jin, R.; Ku, J.; Xu, S.-M.; Wu, Z.-Z.; Huang, T.; Li, X.Y.; Xu, M.L.; et al. Non-destructive detection research for hollow heart of potato based on semi-transmission hyperspectral imaging and svm. *Spectrosc. Spectr. Anal.* **2015**, *35*, 198–202. [CrossRef]
15. Griffel, L.; Delparte, D.; Edwards, J. Using Support Vector Machines classification to differentiate spectral signatures of potato plants infected with Potato Virus Y. Comput. *Electron. Agric.* **2018**, *153*, 318–324. [CrossRef]
16. Atherton, D.; Watson, D.G.; Zhang, M.; Qin, Z.; Liu, X. Hyperspectral Spectroscopy for Detection of Early Blight (*Alternaria solani*) Disease in Potato (*Solanum tuberosum*) Plants at Two Different Growth Stages. In Proceedings of the ASABE Annual International Meeting, New Orleans, LA, USA, 26–29 July 2015. [CrossRef]
17. Garhwal, A.S.; Pullanagari, R.R.; Li, M.; Reis, M.M.; Archer, R. Hyperspectral imaging for identification of Zebra Chip disease in potatoes. *Biosyst. Eng.* **2020**, *197*, 306–317. [CrossRef]
18. Han, Y.F.; Cheng-xu, L.; Yuan, Y.W.; Yang, B.N.; Zhao, Q.L.; Cao, Y.F.; Yin, X.Q. PLS-Discriminant Analysis on Potato Blackheart Disease Based on VIS-NIR Transmission Spectroscopy. *Spectrosc. Spectr. Anal.* **2021**, *41*, 1213–1219. [CrossRef]
19. Agarwal, A.; Gupta, S.D. Assessment of spinach seedling health status and chlorophyll content by multivariate data analysis and multiple linear regression of leaf image features. *Comput. Electron. Agric.* **2018**, *152*, 281–289. [CrossRef]
20. Rakib, M.; Borhan, A.; Jawahir, A. The relationship between SPAD chlorophyll and disease severity index in Ganoderma-infected oil palm seedlings. *J. Bangladesh Agric. Univ.* **2019**, *17*, 355–358. [CrossRef]
21. Minaeva, O.M.; Akimova, E.E.; Tereshchenko, N.N.; Zyubanova, T.I.; Apenysheva, M.V.; Kravets, A.V. Effect of Pseudomonas Bacteria on Peroxidase Activity in Wheat Plants when Infected with Bipolaris sorokiniana. *Russ. J. Plant Physiol.* **2018**, *65*, 717–725. [CrossRef]
22. Ban, S.; Tian, M.; Chang, Q. Estimating the severity of apple mosaic disease with hyperspectral images. *Int. J. Agric. Biol. Eng.* **2019**, *12*, 148–153. [CrossRef]
23. Li, Q.; Hu, Y. Kinetic models of peroxidase activity in potato leaves infected with late blight based on hyperspectral data. *Int. J. Agric. Biol. Eng.* **2019**, *12*, 160–165. [CrossRef]
24. Kang, L.; Gao, R.; Kong, Q.M.; Jia, Y.J.; Shi, Y.B.; Su, Z.B. Estimation of SPAD VaIue of rice Ieaves based on hyperspectraI image (Agarwal and Dutta Gupta). *J. Northeast Agric. Univ.* **2020**, *51*, 89–96. [CrossRef]
25. Kong, W.; Liu, F.; Zhang, C.; Bao, Y.; Yu, J.; He, Y. Fast detection of peroxidase (POD) activity in tomato leaves which infected with *Botrytis cinerea* using hyperspectral imaging. *Spectrochim. Acta Part A Mol. Biomol. Spectrosc.* **2014**, *118*, 498–502. [CrossRef] [PubMed]
26. Cheng, F.; Zhao, Y.R.; Yu, K.Q.; Lou, B.G.; He, Y. Application of Hyper-Spectra for Detecting Peroxidase Content in Cucumber Leaves with Early Disease Stress. *Spectrosc. Spectr. Anal.* **2017**, *37*, 1861–1865. [CrossRef]
27. Sha, W.; Razukas, A. Potato cultivars susceptibility to potato late blight (*Phytopthtora infestans*). *Appl. Ecol. Environ. Res.* **2008**, *6*, 95–106. [CrossRef]
28. Nawar, S.; Mouazen, A.M. Optimal sample selection for measurement of soil organic carbon using on-line vis-NIR spectroscopy. *Comput. Electron. Agric.* **2018**, *151*, 469–477. [CrossRef]
29. Pasquini, C. Near infrared spectroscopy: A mature analytical technique with new perspectives—A review. *Anal. Chim. Acta* **2018**, *1026*, 8–36. [CrossRef]
30. Khulal, U.; Zhao, J.; Hu, W.; Chen, Q. Nondestructive quantifying total volatile basic nitrogen (TVB-N) content in chicken using hyperspectral imaging (HSI) technique combined with different data dimension reduction algorithms. *Food Chem.* **2016**, *197 Pt B*, 1191–1199. [CrossRef]
31. Qi, Z.; Wu, X.; Yang, Y.; Wu, B.; Fu, H. Discrimination of the Red Jujube Varieties Using a Portable NIR Spectrometer and Fuzzy Improved Linear Discriminant Analysis. *Foods* **2022**, *11*, 763. [CrossRef]

32. Zhu, H.; Chu, B.; Zhang, C.; Liu, F.; Jiang, L.; He, Y. Hyperspectral Imaging for Presymptomatic Detection of Tobacco Disease with Successive Projections Algorithm and Machine-learning Classifiers. *Sci. Rep.* **2017**, *7*, 4125. [CrossRef]
33. Zhang, L.; Rao, Z.; Ji, H. NIR Hyperspectral Imaging Technology Combined with Multivariate Methods to Study the Residues of Different Concentrations of Omethoate on Wheat Grain Surface. *Sensors* **2019**, *19*, 3147. [CrossRef]
34. Chen, J.; Li, G. Prediction of moisture content of wood using Modified Random Frog and Vis-NIR hyperspectral imaging. *Infrared Phys. Technol.* **2020**, *105*, 103225. [CrossRef]
35. Li, J.; Wu, J.; Lin, J.; Li, C.; Lu, H.; Lin, C. Nondestructive Identification of Litchi Downy Blight at Different Stages Based on Spectroscopy Analysis. *Agriculture* **2022**, *12*, 402. [CrossRef]
36. Ashie, A.; Lei, H.; Han, B.; Xiong, M.; Yan, H. Fast determination of three components in milk thistle extract with a hand-held NIR spectrometer and chemometrics tools. *Infrared Phys. Technol.* **2021**, *113*, 103629. [CrossRef]
37. Wang, Y.; Guo, W.; Zhu, X.; Liu, Q. Effect of homogenisation on detection of milk protein content based on NIR diffuse reflectance spectroscopy. *Int. J. Food Sci. Technol.* **2019**, *54*, 387–395. [CrossRef]
38. Bhadra, S.; Sagan, V.; Maimaitijiang, M.; Maimaitiyiming, M.; Newcomb, M.; Shakoor, N.; Mockler, T. Quantifying Leaf Chlorophyll Concentration of Sorghum from Hyperspectral Data Using Derivative Calculus and Machine Learning. *Remote Sens.* **2020**, *12*, 2082. [CrossRef]
39. Zhang, D.; Yang, Y.; Chen, G.; Tian, X.; Wang, Z.; Fan, S.; Xin, Z. Nondestructive evaluation of soluble solids content in tomato with different stage by using Vis/NIR technology and multivariate algorithms. *Spectrochim. Acta Part A Mol. Biomol. Spectrosc.* **2021**, *248*, 119139. [CrossRef]
40. Yang, X.; Liu, G.; He, J.; Kang, N.; Yuan, R.; Fan, N. Determination of sugar content in Lingwu jujube by NIR–hyperspectral imaging. *J. Food Sci.* **2021**, *86*, 1201–1214. [CrossRef]
41. Chu, X.; Miao, P.; Zhang, K.; Wei, H.; Fu, H.; Liu, H.; Jiang, H.; Ma, Z. Green Banana Maturity Classification and Quality Evaluation Using Hyperspectral Imaging. *Agriculture* **2022**, *12*, 530. [CrossRef]
42. Dutta, S.; Singh, S.K.; Panigrahy, S. Assessment of Late Blight Induced Diseased Potato Crops: A Case Study for West Bengal District Using Temporal AWiFS and MODIS Data. *J. Indian Soc. Remote Sens.* **2013**, *42*, 353–361. [CrossRef]
43. El_Komy, M.H.; Saleh, A.A.; Ibrahim, Y.E.; Molan, Y.Y. Early production of reactive oxygen species coupled with an efficient antioxidant system play a role in potato resistance to late blight. *Trop. Plant Pathol.* **2020**, *45*, 44–55. [CrossRef]
44. Maksimov, I.V.; Sorokan', A.V.; Chereoanova, E.A.; Surina, O.B.; Troshina, N.B.; Yarullina, L.G. Effects of salicylic and jasmonic acids on the components of pro/antioxidant system in potato plants infected with late blight. *Russ. J. Plant Physiol.* **2011**, *58*, 299–306. [CrossRef]
45. Kundu, R.; Dutta, D.; Nanda, M.K.; Chakrabarty, A. Near Real Time Monitoring of Potato Late Blight Disease Severity using Field Based Hyperspectral Observation. *Smart Agric. Technol.* **2021**, *1*, 100019. [CrossRef]
46. Ray, S.S.; Jain, N.; Arora, R.K.; Chavan, S.; Panigrahy, S. Utility of Hyperspectral Data for Potato Late Blight Disease Detection. *J. Indian Soc. Remote Sens.* **2011**, *39*, 161–169. [CrossRef]