

## Article

# Dissection of Hyperspectral Reflectance to Estimate Photosynthetic Characteristics in Upland Cotton (*Gossypium hirsutum* L.) under Different Nitrogen Fertilizer Application Based on Machine Learning Algorithms

Peng Han <sup>1</sup>, Yaping Zhai <sup>1</sup>, Wenhong Liu <sup>1</sup>, Hairong Lin <sup>1</sup>, Qiushuang An <sup>1</sup>, Qi Zhang <sup>1</sup>, Shugen Ding <sup>1</sup>, Dawei Zhang <sup>2,\*</sup>, Zhenyuan Pan <sup>1,\*</sup> and Xinhui Nie <sup>1,\*</sup>

<sup>1</sup> Key Laboratory of Oasis Ecology Agricultural of Xinjiang Production and Construction Corps, Agricultural College, Shihezi University, Shihezi 832003, China

<sup>2</sup> Research Institute of Economic Crops, Xinjiang Academy of Agricultural Sciences, Urumqi 830091, China

\* Correspondence: zbzdw012@126.com (D.Z.); panzhenyuandawood@163.com (Z.P.); xjnxh2004130@126.com (X.N.)



**Citation:** Han, P.; Zhai, Y.; Liu, W.; Lin, H.; An, Q.; Zhang, Q.; Ding, S.; Zhang, D.; Pan, Z.; Nie, X. Dissection of Hyperspectral Reflectance to Estimate Photosynthetic Characteristics in Upland Cotton (*Gossypium hirsutum* L.) under Different Nitrogen Fertilizer Application Based on Machine Learning Algorithms. *Plants* **2023**, *12*, 455. <https://doi.org/10.3390/plants12030455>

Academic Editors: Michael Gomez Selvaraj, Sindhuja Sankaran and Alvaro Sanz-Saez

Received: 31 August 2022

Revised: 16 December 2022

Accepted: 13 January 2023

Published: 19 January 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/>).

**Abstract:** Hyperspectral technology has enabled rapid and efficient nitrogen monitoring in crops. However, most approaches involve direct monitoring of nitrogen content or physiological and biochemical indicators directly related to nitrogen, which cannot reflect the overall plant nutritional status. Two important photosynthetic traits, the fraction of absorbed photosynthetically active radiation (FAPAR) and the net photosynthetic rate (Pn), were previously shown to respond positively to nitrogen changes. Here, Pn and FAPAR were used for correlation analysis with hyperspectral data to establish a relationship between nitrogen status and hyperspectral characteristics through photosynthetic traits. Using principal component and band autocorrelation analyses of the original spectral reflectance, two band positions (350–450 and 600–750 nm) sensitive to nitrogen changes were obtained. The performances of four machine learning algorithm models based on six forms of hyperspectral transformations showed that the light gradient boosting machine (LightGBM) model based on the hyperspectral first derivative could better invert the Pn of function-leaves in cotton, and the random forest (RF) model based on hyperspectral first derivative could better invert the FAPAR of the cotton canopy. These results provide advanced metrics for non-destructive tracking of cotton nitrogen status, which can be used to diagnose nitrogen nutrition and cotton growth status in large farms.

**Keywords:** cotton; nitrogen monitoring; absorbed photosynthetically active radiation; net photosynthetic rate; remote sensing monitoring; machine learning algorithms models

## 1. Introduction

Nitrogen is an important component of the structural and vital substances of crops, is an essential element for crop metabolism, and affects the growth state of crops [1–3]. Nitrogen deficiency leads to a decrease in leaf thickness [4], closure of leaf stomata [5], degradation of photosynthetic enzymes and thylakoids [6], decrease in chlorophyll content [7,8], and obstruction of light-energy capture, electron transfer, and carboxylation rate processes in leaves [9]. Simultaneously, the growth and development of nitrogen-deficient plants are hindered; plants are short and thin and the leaf area is reduced, which reduces the utilization efficiency of photosynthetically active radiation (PAR) in the crop canopy [10]. Excessive nitrogen application leads to the close arrangement of leaf mesophyll cells [11], which is not conducive to the entry of external CO<sub>2</sub> into the cells, resulting in a decrease in the photosynthetic rate and even toxic effects on crops [12]. In addition, excessive nitrogen application can lead to excessive leaf numbers and dense canopies, which reduces canopy

absorption and light energy utilization in crops [13,14]. The fraction of absorbed photosynthetically active radiation (FAPAR) and the net photosynthetic rate (Pn) are effective tools to assess the response of crops to nitrogen changes at the canopy and leaf scales, respectively [15–21]. It is of great significance to monitor the growth status of crops to determine their nitrogen nutrition status through changes in crop photosynthetic capacity. Traditional methods rely on in-field gas-exchange systems [22] or optical instrument measurement [23,24]. Such approaches can provide precise photosynthetic information, but are costly, time-consuming, and difficult to accomplish, especially in large cotton-growing areas. Therefore, pre-symptomatic or pre-visual detection of plant physiological changes is urgently required to avoid severe damage.

Remote sensing monitoring, as the current mainstream mode of detection, has the unique advantages of being rapid, non-destructive, and multidimensional, and is widely used in the agricultural production process [25–28]. The spectral characteristics of the plant canopy reflect various types of state information during plant growth [7,29,30]. Research has indicated that the spectral characteristics of plant canopy reflections are closely related to several biochemical and biophysical properties, such as pigment concentrations [31,32], plant vigor [33,34], water status [35,36], and nutritional status [37,38]. Different nitrogen states in cotton cause reflectance changes in multiple bands of the spectrum, such as 550–700 nm [39], 705–715 nm [40], and 1325–1575 nm [41]. Hyperspectral remote sensing plays an important role in plant nitrogen nutrition monitoring as a cutting-edge technology. This has been reported for many plant species, and the hyperspectral method has been used to provide rapid, non-destructive estimates of nitrogen status in most major crops, including rice [42,43], maize [44,45], wheat [26,46], and others [27,41,47]. Many studies have confirmed that the photosynthetic ability of crops, which is an important indicator of crop growth status, can be used to indirectly assess growth status [5,20,48]. At the same time, researchers have also constructed hyperspectral estimation models of the photosynthetic characteristics of different crops. Zhou et al. [36] used hyperspectral remote sensing to monitor the photosynthetic characteristics of citrus Pn under water stress. Jin et al. [49] used the hyperspectral vegetation index to evaluate the maximum carboxylation and electron transfer rates of alpine deciduous forests. There are some examples of monitoring plant photosynthetic characteristics using hyperspectral applications [7,50,51]. However, the current research on the nitrogen nutrition status of cotton primarily uses plant or leaf nitrogen content, dry matter, and photosynthetic pigments as direct monitoring indicators, and does not take into account changes in cotton physiology and morphology. Therefore, it is meaningful to study how to use remote sensing to monitor cotton photosynthetic parameters to assess the nitrogen nutrition status effectively and non-destructively.

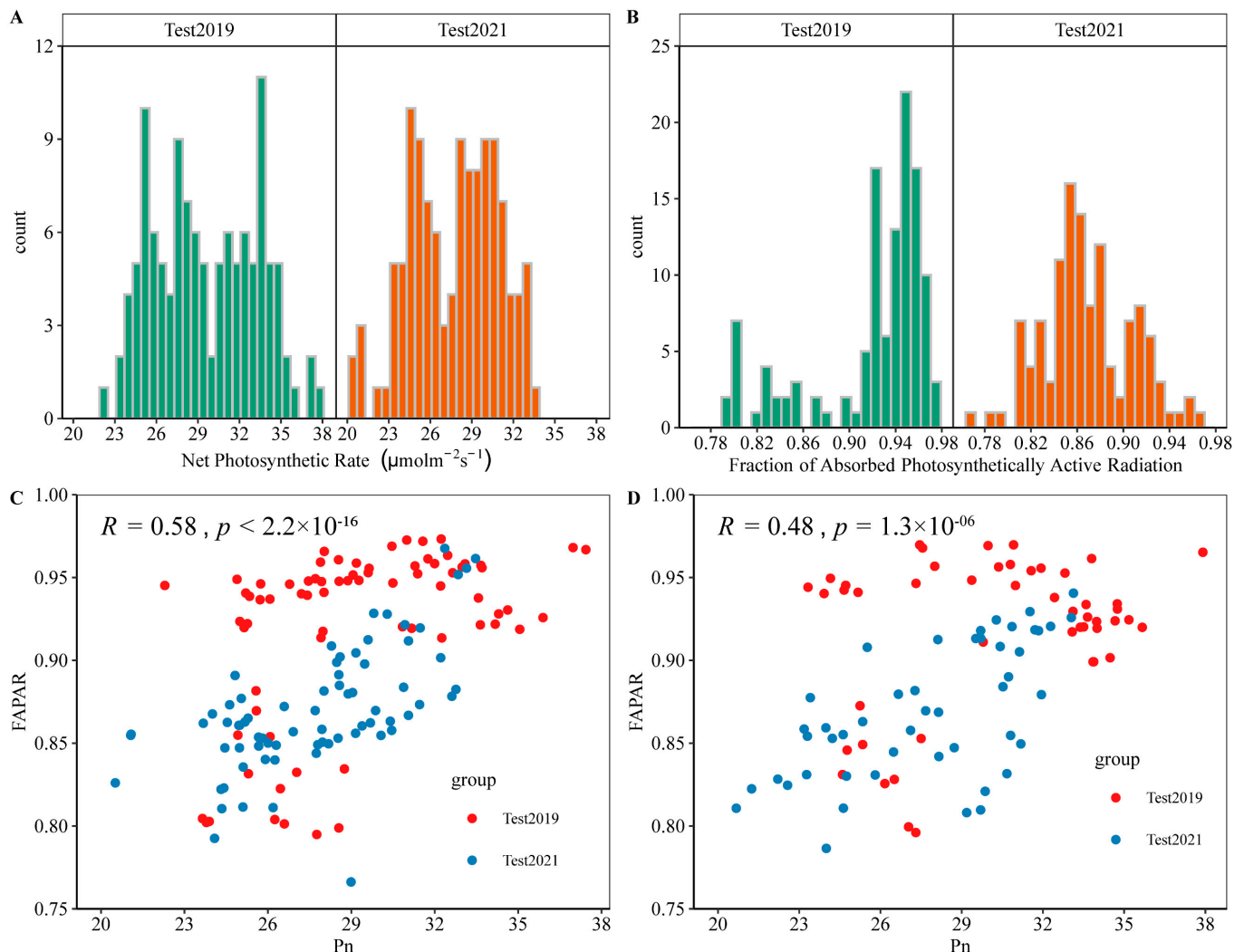
The aim of this study was to construct a remote sensing estimation model for photosynthetic characteristic parameters and population photosynthetically active radiation in cotton leaves under different nitrogen levels. During the key growth period of cotton, the photosynthetic parameters of cotton leaves, population photosynthetically active radiation, and canopy spectral information were obtained. By mining spectral information, the optimal modeling method and hyperspectral variables were used to construct an estimation model of the photosynthetic characteristic parameters of cotton leaves and the population's photosynthetically active radiation. This study provides a theoretical basis for efficient and non-destructive monitoring of the nitrogen nutrition status of cotton and provides scientific guidance for specifying scientific and reasonable fertilization programs.

## 2. Results

### 2.1. Data Distribution of Cotton Photosynthetic Characteristics Parameters

To obtain a dataset with variations, the spectral reflectance, Pn, and FAPAR of the cotton canopy for two cotton varieties were obtained at five growth stages under four nitrogen fertilizer applications across 2 years. A total of 240 data pairs were collected, with 30 pairs for each factor level. In all experiments, Pn and FAPAR were in the range of 20.51–37.91  $\mu\text{mol CO}_2 \cdot \text{m}^{-2} \cdot \text{s}^{-1}$  and 0.7949–0.9925, respectively (Figure 1A,B). The dataset

for subsequent modeling was divided into training ( $n = 180$ ) and validation ( $n = 60$ ). The training and validation datasets were balanced to prevent bias in regression and metrics. Spearman's correlation analysis showed that there was a significant positive correlation between Pn and FAPAR (Figure 1C,D).

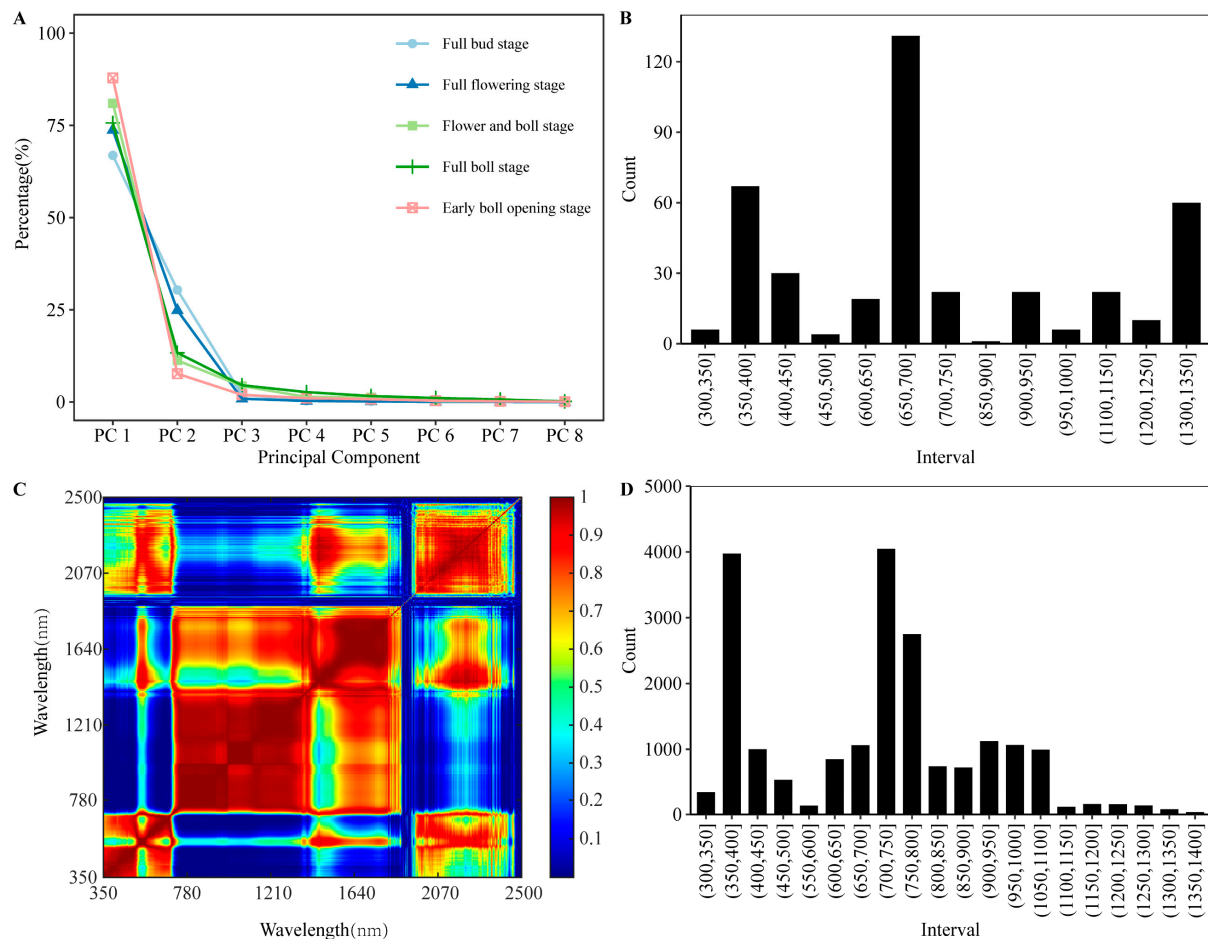


**Figure 1.** Data distributions of net photosynthetic rate (Pn) and fraction of absorbed photosynthetically active radiation (FAPAR) in cotton. (A) Histograms of Pn distribution in all experiments. (B) Histograms of FAPAR distribution in all experiments. (C) Correlation plots of Pn and FAPAR contents in training dataset. (D) Correlation plots of Pn and FAPAR contents in validation datasets. Statistical results for correlation are shown.

## 2.2. Preprocessing and Dimensionality Reduction for Hyperspectral Data

Dimension reduction and preprocessing are important methods for ensuring the reliability and efficiency of a model. Principal component analysis showed that the first eight principal components contributed to 99% of the variability across the reflectance spectrum (Figure 2A). The top ten loading factors, as the main contributors to these principal components, were statistically analyzed. The results showed that the three regions of the hyperspectral characteristics of cotton affected by the nitrogen application level were 350–450 nm (accounting for 18.25% of the total bands), 650–700 nm (32.75%), and 1300–1350 nm (15.00%) (Figure 2B). The hyperspectral data of cotton at different stages of growth were then analyzed for band autocorrelation (Figures 2C and S1). We selected 2000 combinations of bands with the lowest correlation; the frequency of each

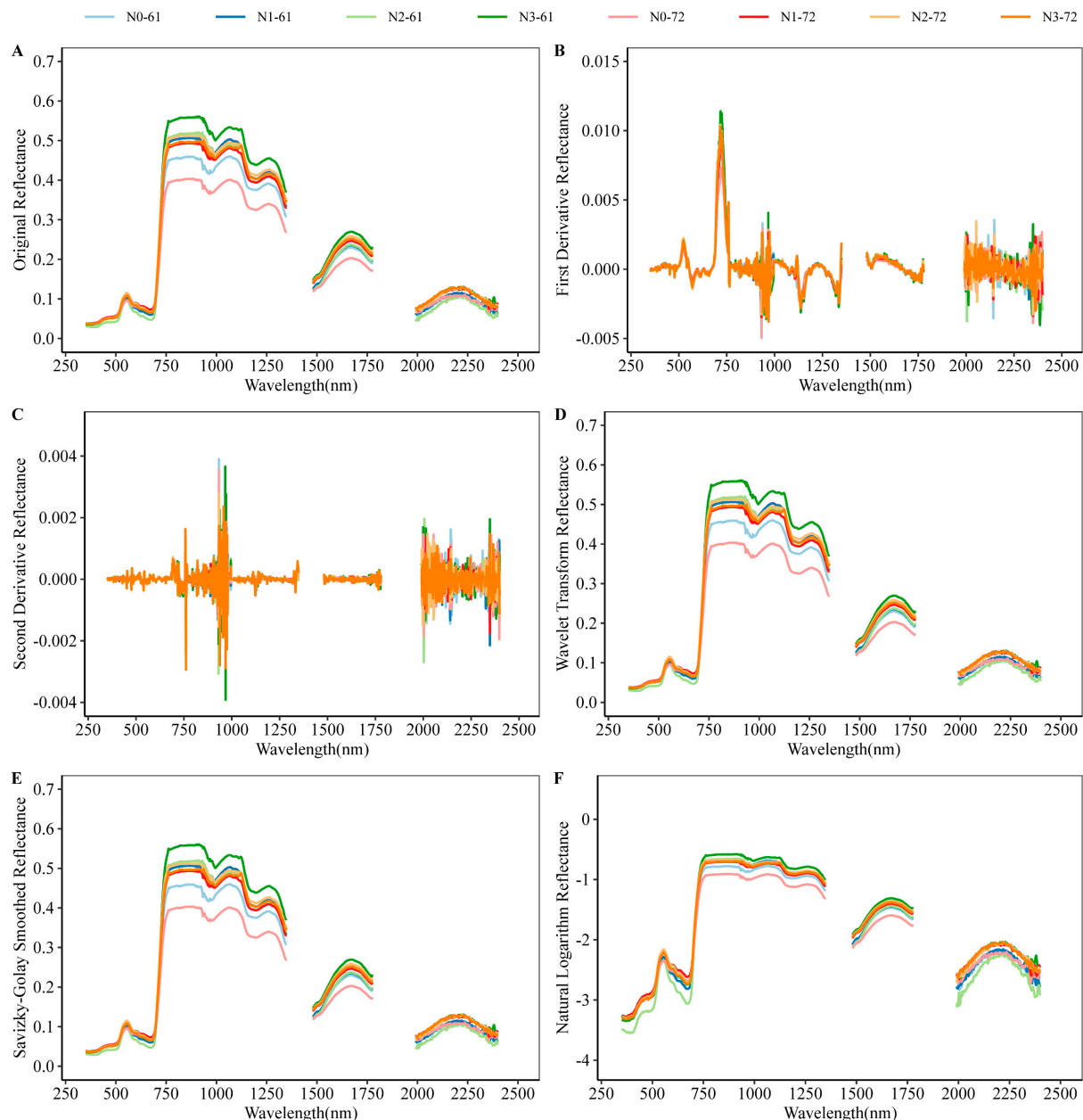
band was counted, and the hyperspectral bands with rich information were in the ranges of 350–450 nm (25.29%), 600–750 nm (39.94%), and 750–1000 nm (23.62%) (Figure 2D). Previous studies have shown that 350–450 nm and 600–650 nm are strong absorption bands of chlorophyll and carotenoids, and 700–750 nm is related to the content of chlorophyll (a + b) per unit area of plants [52]. Therefore, the high spectral reflectance of 350–450 and 600–750 nm was considered the best band range for model construction.



**Figure 2.** Dimensionality reduction performed using principal component analysis and band autocorrelation analysis. (A) Contribution of each principal component to total variation in cotton hyperspectral. (B) Frequency distribution of bands selected by the first eight principal components in each band range. (C) Decision coefficient of autocorrelation of spectral reflectance in different bands of cotton canopy at full bud stage. (D) Distribution of autocorrelation selected bands in each band range.

By comparing the basic soil fertility of the two-year experimental area, the total nitrogen content and organic matter content of the experimental area in 2019 were lower than those in 2020 (Table S1). To avoid insignificant spectral differences between different treatments due to the high base fertility of the soil, the spectral reflectance data of the 2019 experiment (Test2019) were chosen for analysis. The 1350–1480 and 1780–1990 nm wavelengths were removed because the water vapor absorption peak has a large effect on the reflectivity of the cap. The variation trends of the canopy hyperspectral reflectance of the two cotton varieties in different growth periods were similar (Figures 3A and S2–S5), but there were significant differences in the canopy spectral reflectance in different wavelength ranges. There is a large difference in the reflectivity between the visible and near-infrared bands. Five preprocessing methods, namely, first-derivative reflectance (FD), second-derivative reflectivity (SD), wavelet transform reflectivity (WT), Savitzky–Golay smoothed

reflectivity (SG), and natural logarithm reflectivity (NL), were applied to the original spectral reflectance (OR) data to build a regression model. Compared with the original spectral reflectance, the derivative-type spectra reduced the intra-class variance by removing noise from the original spectral data (Figures 3B,C and S2–S5). After preprocessing the spectral data using two smoothing methods, small variations in reflectance were eliminated across the wavelength range, making it more uniform (Figures 3D,E and S2–S5). With the natural logarithmic change, subtle differences in the original spectral curve became more pronounced (Figures 3F and S2–S5).

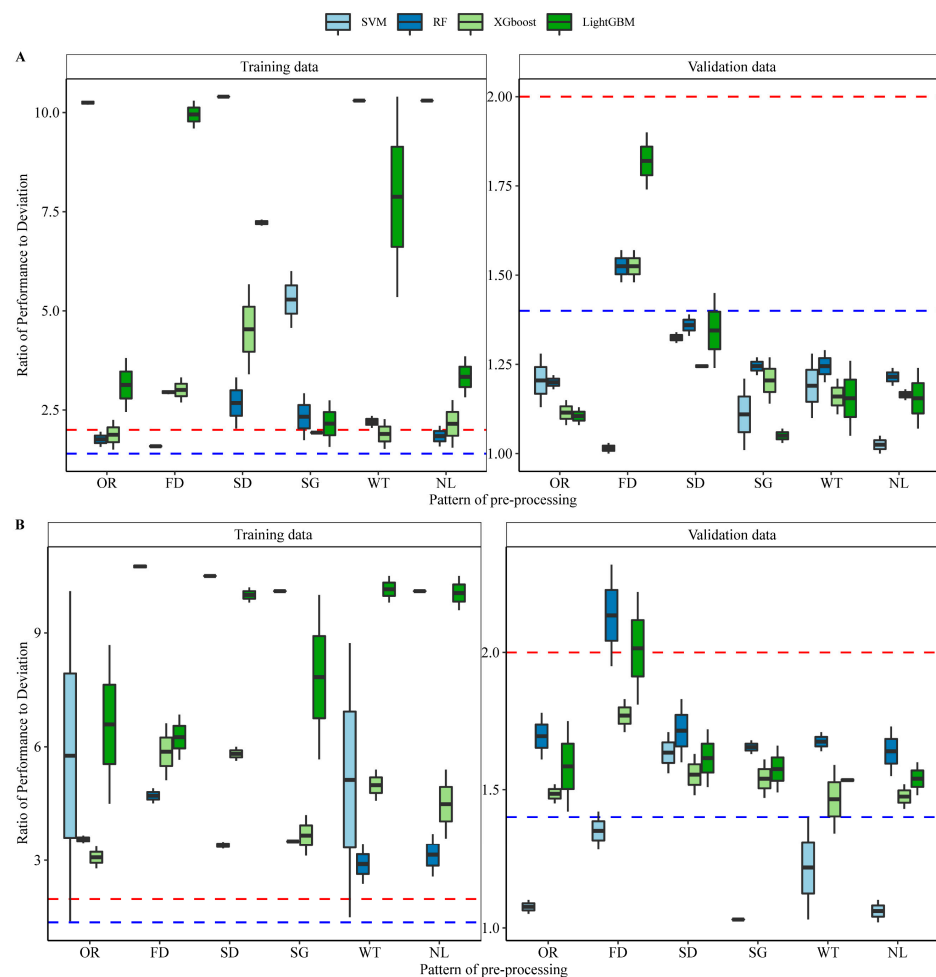


**Figure 3.** Original and preprocessed spectral data of two cotton varieties under different nitrogen treatments at full bud stage in Test2019. (A) Spectral original reflectance; (B) first-derivative reflectivity; (C) second-derivative reflectivity; (D) wavelet transform reflectivity; (E) Savitzky–Golay smoothed reflectivity; (F) natural logarithm reflectivity.

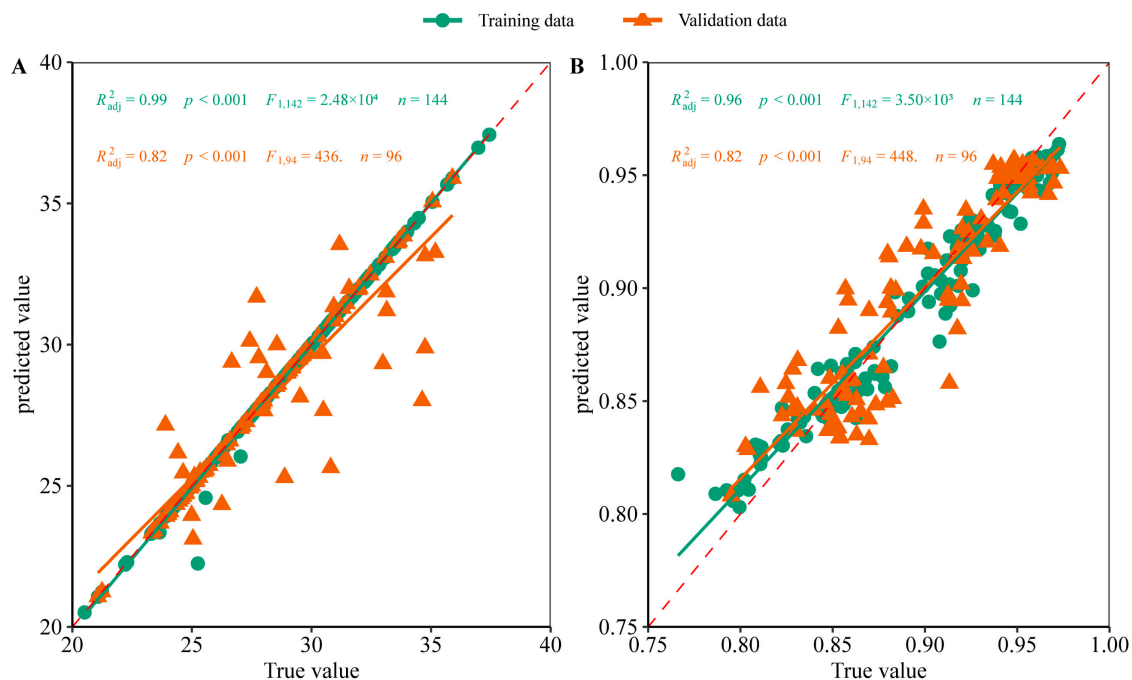


### 2.3. Machine Regression Models of Pn and FAPAR Based on Characteristic Bands

To simplify the model and reduce over-fitting, the reflectance wavelengths of 350–450 and 600–750 nm reflectance were used as characteristic bands to construct the model. Correlation estimation models of cotton leaf Pn and canopy FAPAR were constructed using four different machine learning algorithms: Support vector machine (SVM), random forest (RF), extreme gradient boosting (XGBoost), and light gradient boosting machine (LightGBM). Model performance was evaluated using the ratio of performance to deviation (RPD) values and robustness to multiple repetitions. The machine learning models indicated high performance and robustness when first-derivative reflectivity and second-derivative reflectivity were applied as hyperspectral data in the training and validation datasets. For Pn modeling, RF, XGBoost, and LightGBM models indicated high performance and robustness when FD was applied as hyperspectral data both in the training data and validation data, as most RPD values were greater than 1.4, representing an acceptable prediction level (Figures 4A and 5A). For FAPAR modeling, RF and LightGBM models based on FD also demonstrated high performance and robustness in training and validation data, with RPD values above 2.0, representing an accurate prediction level (Figures 4B and 5B). These results were supported by the coefficient of determination ( $R^2$ ), root mean square error (RMSE), and mean absolute percent error (MAPE), which are indicators of model performance (Figures S6–S8).



**Figure 4.** Model performance and robustness for each preprocessing of reflectance. (A) Machine learning models for Pn. (B) machine learning models for the fraction of absorbed photosynthetically active radiation (FAPAR). The ratio of performance to deviation (RPD) was applied to evaluate the accuracy of each model. Figures are plots of the RPD values in each repeat. Blue and red lines indicate RPD values of 1.4 and 2.0, respectively, as accuracy thresholds.



**Figure 5.** Measured vs. predicted values after applying machine learning models to predict net photosynthetic rate (Pn) and fraction of absorbed photosynthetically active radiation (FAPAR). (A) Pn—light gradient boosting machine (LightGBM); (B) FAPAR—random forest (RF). The red dotted line is the 1:1 line.

### 3. Discussion

Nitrogen management is problematic in large-scale cotton cultivation. However, laboratory methods for measuring nitrogen content and photosynthetic characteristics remain complex experimental techniques that are not suitable for large-scale and real-time monitoring [22–24]. Therefore, novel techniques are required to efficiently assess nitrogen nutrient status and monitor crop photosynthetic physiology. The nitrogen nutrition status of plants is not only reflected in the changes in nitrogen content in plants, but also in changes in plant growth status, including changes in the internal physiology and external morphology of plants, which can be reflected in plants' spectral signature [29,50,53]. Although using remote sensing to assess plant responses to environmental changes is a very active research topic, most studies to date have focused on nitrogen content or hormone content related to nitrogen stress [54,55]. At present, there are few studies on the correlation between cotton photosynthetic characteristics and hyperspectral data for estimating the nitrogen nutrition status of cotton. Therefore, the current study aimed to address these questions by using the difference in hyperspectral reflectance caused by different nitrogen application conditions, using a machine learning algorithm to accurately estimate cotton Pn and FAPAR, and implementing non-destructive estimation of cotton nitrogen nutrition status.

Hyperspectral techniques can obtain detailed spectral information from plants and can be used to closely monitor plant growth as well as physiological and biochemical properties [41,50,53,56]. Nevertheless, the use of full hyperspectral bands may cause information redundancy and interference [57,58]. In this study, principal component analysis and band autocorrelation based on the original spectral reflectance jointly screened the characteristic bands. Two highly correlated band regions (350–450 and 600–750 nm) were co-selected, located in the visible bands. Previous research has shown that the light depth in these wavelength ranges is involved in the photosynthetic process of plants [59,60], and the reflectance can reflect photomorphogenesis in and out of plants [61,62].

At the same time, it is necessary to use preprocessing methods to improve the level of hyperspectral modeling. Preprocessing of original spectral data was performed using

the first derivative, second derivative, wavelet transform smoothing, Savitzky–Golay smoothing, and natural logarithm transformation to reduce the effect of background noise. These preprocessing methods, based on the full-band original spectral dataset, retained all the information of the original spectral data. Our research shows that the accuracy and stability of the prediction model were improved by using suitable preprocessing and transformation and that the machine learning model built using derivative-type spectral data was better than other preprocessing methods. Peng et al. believed that when the distance between the sensor and the collected crops was small, the low-order differential could effectively remove spectral background noise and improve the inversion ability of the model [63]. Xia et al. also reported that the application of derivative spectroscopy improved the accuracy of the model [64].

Based on a powerful and flexible algorithmic framework, machine learning has the potential to analyze hyperspectral reflectance data, thereby avoiding the waste of hyperspectral information in traditional research [30,65,66]. In this study, most FAPAR models were more robust than Pn models. The RPD values of the predicted data for most FAPAR models were greater than 1.4, which is an acceptable prediction level. This may be related to the collection form of hyperspectral data. Canopy-level hyperspectral information has a stronger predictive ability for FAPAR; its predictive ability for Pn requires improvement.

The machine learning methods LightGBM, RF, and XGBoost showed high performance and robustness for both Pn and FAPAR. The excellent performance of LightGBM may be attributed to its unique leaf growth strategy. RF is more robust for input variables and outliers. XGBoost adds a regular term to control the complexity of the model and reduce the possibility of overfitting. Therefore, machine learning is a powerful tool for estimating crop-related indicators using hyperspectral technology. In recent years, machine learning models have been increasingly used in agricultural remote sensing [41,65–67]. However, in previous studies, a single band or vegetation index was often used to remotely estimate crop growth indicators, resulting in a large waste of hyperspectral data and unstable model predictions [55,68,69]. In this study, we used a larger range of characteristic bands to predict the cotton nitrogen nutrient status. This will help improve the difficulty of monitoring large areas of cotton. This study contributes to the establishment of non-destructive monitoring management strategies and accurate modeling methods.

## 4. Materials and Methods

### 4.1. Site Characteristics

The experiments were conducted from May to September 2019 (Test2019) and 2021 (Test2021), respectively, in two experimental fields with different basal fertilities (Table S1) at the field experiment station of the College of Agriculture, Shihezi University, Shihezi, Xinjiang Uygur Autonomous Region, China (43°26′–45°20′ N, 84°58′–86°24′ E). Shihezi is a hinterland city in Xinjiang with a typical temperate continental climate. The annual average temperature, annual average relative humidity, precipitation, and annual average frost-free period are approximately 7 °C, 55%, 180 mm, and 170 days, respectively. The rainfall and temperature during the cotton growing season and the time period of the experiment (from April to October) are shown in Figure S9.

### 4.2. Experimental Design

The experiment was conducted using a split-plot experimental design of randomized groups with three replicates, which included treatments with four amounts of nitrogen fertilizer as follows: (i) No nitrogen application (N0), without N fertilization application; (ii) less nitrogen application (N1), N fertilizer applied at 150 kg N ha<sup>−1</sup>; (iii) common nitrogen application (N2), N fertilizer applied at 300 kg N ha<sup>−1</sup>; and (iv) excessive nitrogen application (N3), nitrogen fertilizer applied at 450 kg N ha<sup>−1</sup>. The sub-area treatments with cotton of two different genotypes were as follows: (i) Xinluzao61 and (ii) Xinluzao72.

Except for nitrogen fertilizer, the other fertilizers, phosphate (P) and potassium (K), were the same for all treatments: 105 kg P<sub>2</sub>O<sub>5</sub> ha<sup>−1</sup> and 75 kg K<sub>2</sub>O ha<sup>−1</sup>. The application



rates were recommended by local agricultural technicians and were sufficient to meet the requirements of cotton. The fertilizer types used in this study were urea (46.4% N), potassium dihydrogen phosphate (52% P<sub>2</sub>O<sub>5</sub>, 34% K<sub>2</sub>O), and potassium sulfate (52% K<sub>2</sub>O). All fertilizers were applied in multiple drops of water (Tables S2 and S3).

Two cotton varieties were selected (Xinluzao 61 and Xinluzao 72); both have excellent variety characteristics and are the major cultivars in Xinjiang (Table S4). A plastic film mulch and drip irrigation were used in the experiments. Before sowing, the experimental plots were covered with 2.28 m-wide sheets of transparent plastic film. Drip irrigation lines were installed beneath the plastic film. The cultivars were sown through holes in a plastic film mulch. A wide-narrow row-spacing (66 + 10 cm) configuration was used in the experimental plots, with a planting density of 174,000 plants ha<sup>-1</sup>. Weeds, pests, and diseases were controlled using chemicals via local methods. No obvious weed, pest, or disease problems occurred during the cotton-growing season.

#### 4.3. Photosynthetic Measurements

Following the standard method [22], cotton reached peak photosynthesis at 11:30–13:30 Beijing time on a day with sunny and stable weather. The net photosynthetic rate (P<sub>n</sub>, μmol CO<sub>2</sub>·m<sup>-2</sup>·s<sup>-1</sup>) of cotton key functional leaves was measured on clear and cloudless days during five reproductive periods using an LI-6400XT portable photosynthesis instrument (LI-COR Bioscience Inc., Lincoln, NE, USA). The data were collected on the 55th, 70th, 85th, 100th, and 115th days after cotton emergence, involving the full bud stage, full flowering stage, flower and boll stage, full boll stage, and early boll-opening stage of cotton. The effective radiation of the light source was PAR 1800 μmol·m<sup>-2</sup>·s<sup>-1</sup>, the leaf chamber used was 2 × 3 cm, the gas flow rate was 500 mmol·s<sup>-1</sup>, and the CO<sub>2</sub> injector was 400 μmol·mol<sup>-1</sup>. Ten plants from each treatment group were centrally selected for each measurement period.

PAR was synchronously sampled with the leaf net photosynthetic rate and measured in the time domain. PAR was measured using an LI-250A light meter with an LI-190SA quantum sensor (LI-COR Bioscience, Inc., Lincoln, NE, USA). According to previous methods [23,24], the measured solar radiation included incoming PAR above the canopy (PAR<sub>t</sub>) and the three types of PAR transmitted through the canopy to the ground (PAR<sub>s</sub>, PAR<sub>n</sub>, and PAR<sub>w</sub>). PAR<sub>t</sub> was measured using a quantum sensor placed horizontally 0.5 m above the cotton canopy surface pointing toward the sky; PAR<sub>s</sub> were measured with line quantum sensors placed approximately 5 cm above the ground, pointing upward; PAR<sub>w</sub> and PAR<sub>n</sub> were measured with line quantum sensors placed approximately 5 cm above the ground in wide and narrow rows of cotton, respectively, pointing upward. Each type of PAR was obtained using the mean value of the three different measurement directions. The average values of cotton PAR<sub>s</sub>, PAR<sub>w</sub>, and PAR<sub>n</sub> were taken as the cotton-bottom PAR (PAR<sub>b</sub>). The fraction of absorbed photosynthetically active radiation (FAPAR) was determined using Equation (1).

$$\text{FAPAR} = \frac{(\text{PAR}_t - \text{PAR}_b)}{\text{PAR}_t} \quad (1)$$

#### 4.4. Hyperspectral Measurement Processing

For synchronization with the measurement of photosynthesis, the spectral radiance of the cotton canopy was measured with a full-range hyperspectral SR-3500 portable ground object spectrometer (Spectral Evolution, Lawrence, KS, USA) under natural light on a sunny day in the field. The instrument collected data in the 350–2500 nm spectral range, with a resampled spectral resolution of 3 nm before 1000 nm, 8 nm between 1000 and 1500 nm, and 6 nm after 1500 nm.

For cotton plant canopies where photosynthesis was measured, the spectral reflectance was measured 50 cm above the cotton canopy using a hyperspectral one-handed remote-measuring device. The acquired reflectance spectra could be affected by several environmental factors. Whiteboard calibration was used to eliminate or minimize these side

effects [70]. Before each measurement, the instrument was optimized once, and a standard whiteboard calibration was carried out. The reflectance of the cotton canopy was computed as the ratio of leaf radiances relative to the radiance from the white reference panel. More than five measurements at each location and averaged to produce a spectral reflectance. Equation (2) was used for cotton canopy reflectance calculation.

$$R_{ci} = \frac{R_{mi}}{R_{wi}} \quad (2)$$

where  $R_{ci}$  is the corrected reflectance of the cotton canopy,  $R_{wi}$  is the reflectance of the whiteboard, and  $R_{mi}$  is the measured reflectance of the cotton canopy.

Owing to background noise, the signal-to-noise ratio of the original spectral reflectance was low. Smoothing, derivative processing, and mathematical transformation are considered to be important for improving the signal-to-noise ratio [71,72]. To be more conducive to the recognition of original spectral data, five types of preprocessing methods, including first derivative, second derivative, wavelet transform smoothing, Savitzky–Golay smoothing, and natural logarithm transformation, were adopted to eliminate data noise and highlight the change law of reflectance with wavelength.

#### 4.5. Filtering Characteristic Bands

Using the splice correction function that comes with the instrument, reflectance data at 1-nm steps were obtained across the entire wavelength domain from 400 to 2500 nm. Hyperspectral data contain information on many bands and there is large data redundancy and collinearity, which can affect model building [57,58]. Thus, it is necessary to reduce data redundancy by reducing data dimensionality. To be more conducive to the screening of the hyperspectral feature band range, the dimensionality reduction methods of principal component analysis and band autocorrelation were used to select the feature band.

Principal component analysis can analyze the main factors that cause differences in a large number of variables, and then achieve the effect of data dimensionality reduction [73,74]. By screening the bands corresponding to the top ten weight coefficients of each principal component that provided more than 99% of the spectral variation interpretation, the optimal band for the spectral information of the cotton canopy was selected.

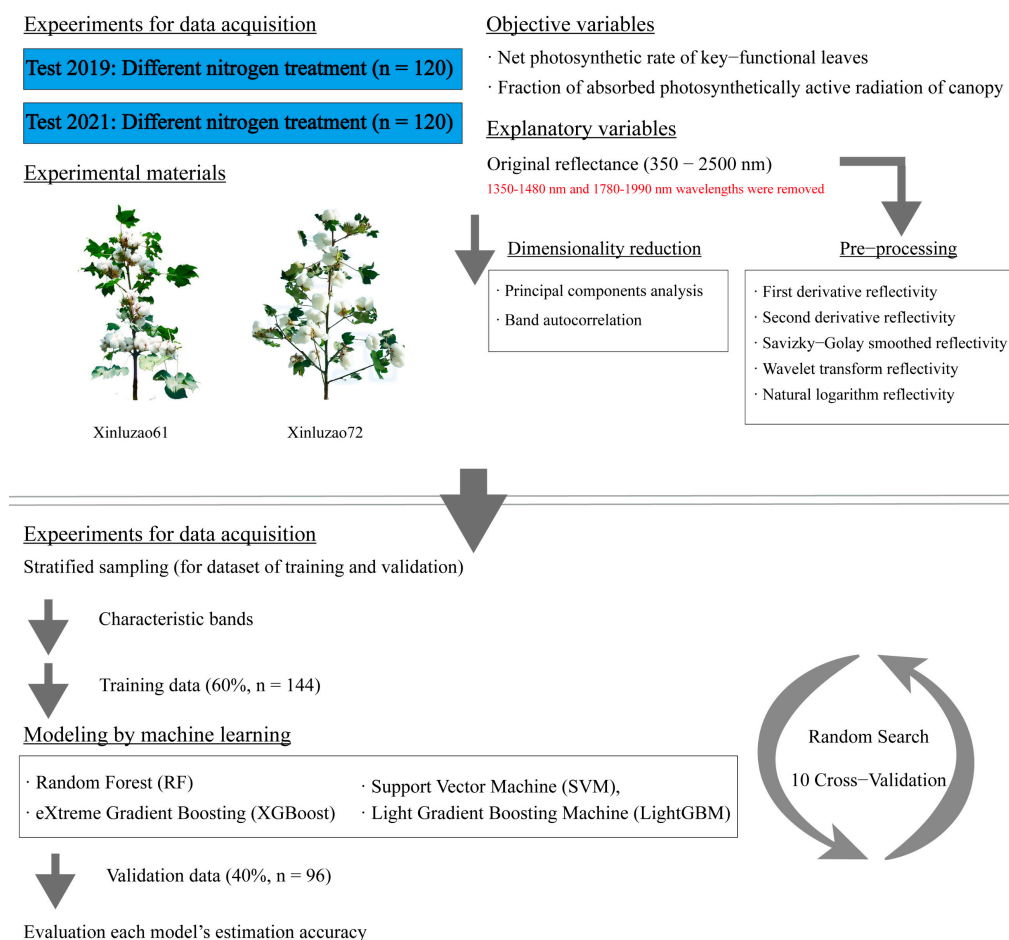
Band autocorrelation can reduce spectral content redundancy through the correlation relationship between the two bands [75]. The smaller the correlation coefficient, the lower the information redundancy between the two bands. Using the spectral reflectance in the 350–2500 nm band, the  $r^2$  correlation coefficient was calculated using a pairwise combination, and all the combinations constituted a  $2151 \times 2151$   $r^2$  matrix. By selecting the band combinations corresponding to 2000 smaller values in the  $r^2$  matrix of different growth periods, the occurrences of different bands were counted, and the band range with rich hyperspectral information was determined according to the band distribution.

However, various methods have been devised to select the range of characteristic bands, each with its own advantages and limitations. Therefore, in this study, we consider the common part of the characteristic bands screened by the two methods as the characteristic band range.

#### 4.6. Model Construction

The process of generating regression models is described in a traditional manner [76], with slight modifications. For modeling, the data from Test 2019 and Test 2021 were divided into two groups by stratified sampling (Figure 6): The training dataset (60%) and the validation dataset (40%).

Based on the feature bands screened by band autocorrelation analysis and principal component analysis, correlation estimation models of cotton leaf Pn and canopy FAPAR were constructed using four different machine learning algorithms, including SVM, RF, XG-Boost, and LightGBM with spectral OR, FD, SD, WT, SG, and NL as independent variables.



**Figure 6.** Experiment and modeling designs in this study.

SVM can better minimize the structural risk and select the optimal solution between the accuracy of the given data and the complexity of the function to obtain the best model promotion ability. The SVM model has the advantage of solving the problems of the small sample size and nonlinear, high-dimensional data [77].

The decision tree is a common machine learning algorithm that is easy to understand and interpret. RF builds a bagging ensemble with a decision tree as the base learner and further introduces random feature selection in the training process of the decision tree, which has great advantages in processing high-dimensional data [77]. XGBoost is a large-scale parallel boosting tree tool that has the advantages of rapid speed, good effect, and the ability to handle large-scale data. Its regularization, parallel structure, and high flexibility have introduced the ability to improve the prediction model [78]. LightGBM is a new member of the boosting set model that adopts the strategy of leaf-wise tree growth to construct decision trees and features ultrafast efficiency in coping with large datasets [79]. LightGBM has shown potential as an efficient model method for various agricultural targets using reflectance data such as seed germination [80], lignin content [81], variety classification [82], and crop breeding [83].

#### 4.7. Data Analysis and Charting

The preprocessing of the five types of hyperspectral data was performed using Origin 2022b (OriginLab, Northampton, MA, USA) and Microsoft Excel 2019 (Microsoft Corporation, Redmond, WA, USA). The SG smoothing window size was set to 5 [84]. We set the wavelet type to DB6, the expansion mode to periodic, and the truncation rate to 20% in the WT smoothing [85].

Four machine learning algorithms were performed using the “mlr3” package (V. 0.13.3) in R 4.1.3 [86]. This package integrates various machine learning algorithm packages to achieve a unified and neat machine learning process-oriented operation. To optimize these machine learning algorithms, we used ten-fold cross-validation and a random search to tune the hyperparameters. Information about the hyperparameters of these machine learning algorithms is presented in Table S5. All figures were created using the R package “ggplot2” V. 3.3.5 [87] and the extension packages. Statistical analyses were conducted using the R statistical programming language V. 4.1.3.

#### 4.8. Model Testing and Evaluation

The prediction metrics used to evaluate the abovementioned algorithms were  $R^2$ , RMSE, and MAPE, which were calculated from Equations (3)–(5).

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

$$R^2 = 1 - \frac{\sum_{i=1}^n \sum (y_i - \hat{y}_i)^2}{\sum_{i=1}^n \sum (y_i - \bar{y})^2} \quad (4)$$

$$\text{MAPE} = \frac{1}{n} \sum_{i=1}^n \left| \frac{\hat{y}_i - y_i}{y_i} \right| \times 100\% \quad (5)$$

where  $\hat{y}_i$  is the predicted Pn or FAPAR value,  $y_i$  is the measured Pn or FAPAR value,  $\bar{y}$  is the mean value of the measured Pn or FAPAR, and  $n$  is the number of validations.

In the field of spectroscopy, the RPD has been used as the standard method to report the quality of a model [37,88–91]. The quality of the prediction model was interpreted according to three classes of RPD as follows:  $\text{RPD} > 2$  indicates accurate model predictions,  $\text{RPD}$  of 1.4–2 indicates reasonably acceptable predictions, and  $\text{RPD} < 1.4$  indicates poor model predictions. RPD was conducted using “yardstick” (R package).

## 5. Conclusions

In this study, correlation estimation models between hyperspectral data and photosynthetic characteristic parameters (Pn and FAPAR) were established using machine learning algorithms under different nitrogen conditions. The results showed that hyperspectral reflectance data and machine learning methods have excellent potential for estimating leaf Pn and canopy FAPAR in cotton. Among them, the LightGBM model (selected super-parameters were the shrinkage rate: 0.053, number of boosting iterations: 479, maximum number of leaves in one tree: 22, limit the maximum depth for tree model: 5, and minimal number of data in one leaf: 18), based on the hyperspectral first derivative, can better invert the Pn of cotton leaves. The RF model (selected super-parameters were the number of trees: 87; number of variables used to split the nodes: 68; minimum number of unique cases in a terminal node: 3; and maximum depth at which a tree should be grown: 37), based on the hyperspectral first derivative, can better invert the FAPAR of the cotton canopy. These results provide advanced metrics for non-destructive tracking of crop nitrogen status. In the future, these techniques can be used to diagnose nitrogen nutrition and the growth status of crops on large farms in real time.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/plants12030455/s1>, Figure S1: Autocorrelation of spectral reflectance in different bands of cotton canopy at different growth-stages. Note: (A) is full flowering stage; (B) is flower and boll stage; (C) is full boll stage; (D) is early boll opening stage. Figure S2: Original and preprocessed spectral data of two cotton varieties under different nitrogen treatments at full flowering stage in Test2019. Notes: The order of pictures is the same as in Figure 3. Figure S3: Original and preprocessed spectral data of two cotton varieties under different nitrogen treatments at flower and boll stage in Test2019. Notes: The order of pictures is the same as in Figure 3. Figure S4: Original

and preprocessed spectral data of two cotton varieties under different nitrogen treatments at full boll stage e in Test2019. Notes: The order of pictures is the same as in Figure 3. Figure S5: Original and preprocessed spectral data of two cotton varieties under different nitrogen treatments at early boll opening stage in Test2019. Notes: The order of pictures is the same as in Figure 3. Figure S6: The coefficient of determination ( $R^2$ ) for each regression model and pre-processing of reflectance. Notes: (A) The machine learning models for Pn. (B) The machine learning models for FAPAR. Figure S7: The root mean squared error (RMSE) for each regression model and pre-processing of reflectance. Notes: (A) The machine learning models for Pn. (B) The machine learning models for FAPAR. Figure S8: The mean absolute percent error (MAPE) for each regression model and pre-processing of reflectance. Notes: (A) The machine learning models for Pn. (B) The machine learning models for FAPAR. Figure S9: The temperature and precipitation changes of Shihezi, Xinjiang in 2019 and 2021 from sowing to harvesting of cotton. Notes: Weather data from April to October of 2019 (A,B) and 2021 (C,D) were collected from <https://www.worldweatheronline.com> (1 August 2022). Table S1: The soil type and fertility of the two experimental fields. Table S2: The summarization of fertilization with irrigation during the whole growth period in 2019. Table S3: The summarization of fertilization with irrigation during the whole growth period in 2021. Table S4: Characteristics of two cotton varieties. Table S5: Tuning hyperparameters in each machine learning algorithm.

**Author Contributions:** X.N. and D.Z. designed the experiment; X.N., D.Z., Z.P., W.L. and H.L. revised and guided the content of the article; P.H., Y.Z., Q.A., Q.Z. and S.D. conducted the experiment; P.H. and Y.Z. conducted the data analysis; P.H. finished the writing of the manuscript. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by the Special Fund for Major Science and Technology in Xinjiang Uygur Autonomous Region, China (No. 2021A02001–4), the Key Programs for Science and Technology Development of Shihezi city, Xinjiang production and Construction Corps, China (No. 2022NY01), and the High-Level Talent Scientific Research Start-UP project of shihezi university, Shihezi city, Xinjiang production and Construction Corps, China (No. RCZK202013).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The raw data supporting the conclusions of this article will be made available upon request to the corresponding authors.

**Acknowledgments:** We thank Tongyu Hou of Shihezi University and Jingxin Zhang of Harbin University of Commerce for their insight and guidance in carrying out the research.

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

## References

1. Leghari, S.J.; Wahocho, N.A.; Laghari, G.M.; HafeezLaghari, A.; MustafaBhabhan, G.; HussainTalpur, K.; Bhutto, T.A.; Wahocho, S.A.; Lashari, A.A. Role of nitrogen for plant growth and development: A review. *Adv. Environ. Biol.* **2016**, *10*, 209–219.
2. Hirel, B.; Bertin, P.; Quilleré, I.; Bourdoncle, W.; Attagnant, C.; Delley, C.; Gouy, A.; Cadiou, S.; Retailliau, C.; Falque, M.; et al. Towards a Better Understanding of the Genetic and Physiological Basis for Nitrogen Use Efficiency in Maize. *Plant Physiol.* **2001**, *125*, 1258–1270. [[CrossRef](#)] [[PubMed](#)]
3. Novoa, R.; Loomis, R.S. Nitrogen and plant production. *Plant Soil* **1981**, *58*, 177–204. [[CrossRef](#)]
4. Hou, W.; Tränkner, M.; Lu, J.; Yan, J.; Huang, S.; Ren, T.; Cong, R.; Li, X. Diagnosis of Nitrogen Nutrition in Rice Leaves Influenced by Potassium Levels. *Front. Plant Sci.* **2020**, *11*, 165. [[CrossRef](#)]
5. Tewari, R.K.; Kumar, P.; Sharma, P.N. Oxidative Stress and Antioxidant Responses in Young Leaves of Mulberry Plants Grown Under Nitrogen, Phosphorus or Potassium Deficiency. *J. Integr. Plant Biol.* **2007**, *49*, 313–322. [[CrossRef](#)]
6. Mu, X.; Chen, Y. The physiological response of photosynthesis to nitrogen deficiency. *Plant Physiol. Biochem.* **2021**, *158*, 76–82. [[CrossRef](#)] [[PubMed](#)]
7. Zhao, D.; Reddy, K.R.; Kakani, V.G.; Reddy, V.R. Nitrogen deficiency effects on plant growth, leaf photosynthesis, and hyperspectral reflectance properties of sorghum. *Eur. J. Agron.* **2005**, *22*, 391–403. [[CrossRef](#)]
8. Haripriya Anand, M.; Byju, G. Chlorophyll meter and leaf colour chart to estimate chlorophyll content, leaf colour, and yield of cassava. *Photosynthetica* **2008**, *46*, 511–516. [[CrossRef](#)]
9. Xiong, D.; Chen, J.; Yu, T.; Gao, W.; Ling, X.; Li, Y.; Peng, S.; Huang, J. SPAD-based leaf nitrogen estimation is impacted by environmental factors and crop leaf characteristics. *Sci. Rep.* **2015**, *5*, 13389. [[CrossRef](#)]



10. Massignam, A.; Chapman, S.; Hammer, G.; Fukai, S. Effects of nitrogen supply on canopy development of maize and sunflower. *Crop Pasture Sci.* **2012**, *62*, 1045–1055. [\[CrossRef\]](#)
11. Gutiérrez-Rodríguez, E.; Lieth, H.J.; Jernstedt, J.A.; Labavitch, J.M.; Suslow, T.V.; Cantwell, M.I. Texture, composition and anatomy of spinach leaves in relation to nitrogen fertilization. *J. Sci. Food Agric.* **2013**, *93*, 227–237. [\[CrossRef\]](#) [\[PubMed\]](#)
12. Albornoz, F. Crop responses to nitrogen overfertilization: A review. *Sci. Hortic.* **2016**, *205*, 79–83. [\[CrossRef\]](#)
13. Luo, C.; Guo, Z.; Xiao, J.; Dong, K.; Dong, Y. Effects of Applied Ratio of Nitrogen on the Light Environment in the Canopy and Growth, Development and Yield of Wheat When Intercropped. *Front. Plant Sci.* **2021**, *12*, 719850. [\[CrossRef\]](#) [\[PubMed\]](#)
14. Tang, L.; Zhu, X.C.; Cao, M.Y.; Cao, W.X.; Zhu, Y. Relationships of rice canopy PAR interception and light use efficiency to grain yield. *Ying Yong Sheng Tai Xue Bao* **2012**, *23*, 1269–1276. [\[PubMed\]](#)
15. Dreccer, M.; Schapendonk, A.; Slafer, G.; Rabbinge, R. Comparative response of wheat and oilseed rape to nitrogen supply: Absorption and utilisation efficiency of radiation and nitrogen during the reproductive stages determining yield. *Plant Soil* **2000**, *220*, 189–205. [\[CrossRef\]](#)
16. Gitelson, A.A.; Peng, Y.; Huemmrich, K.F. Relationship between fraction of radiation absorbed by photosynthesizing maize and soybean canopies and NDVI from remotely sensed data taken at close range and from MODIS 250 m resolution data. *Remote Sens. Environ.* **2014**, *147*, 108–120. [\[CrossRef\]](#)
17. Serrano, L.; Filella, I.; Penuelas, J. Remote sensing of biomass and yield of winter wheat under different nitrogen supplies. *Crop Sci.* **2000**, *40*, 723–731. [\[CrossRef\]](#)
18. Milroy, S.P.; Bange, M.P. Nitrogen and light responses of cotton photosynthesis and implications for crop growth. *Crop Sci.* **2003**, *43*, 904–913. [\[CrossRef\]](#)
19. Yang, X.; Geng, J.; Huo, X.; Lei, S.; Lang, Y.; Li, H.; Liu, Q. Effects of different nitrogen fertilizer types and rates on cotton leaf senescence, yield and soil inorganic nitrogen. *Arch. Agron. Soil Sci.* **2021**, *67*, 1507–1520. [\[CrossRef\]](#)
20. Guo, H.X.; Liu, W.Q.; Shi, Y. Effects of different nitrogen forms on photosynthetic rate and the chlorophyll fluorescence induction kinetics of flue-cured tobacco. *Photosynthetica* **2006**, *44*, 140–142. [\[CrossRef\]](#)
21. Ghosh, A.; Dass, A.; Krishnan, P.; Kaur, R.; Rana, K. Assessment of photosynthetically active radiation, photosynthetic rate, biomass and yield of two maize varieties under varied planting dates and nitrogen application. *J. Environ. Biol.* **2017**, *38*, 683. [\[CrossRef\]](#)
22. Long, S.P.; Bernacchi, C.J. Gas exchange measurements, what can they tell us about the underlying limitations to photosynthesis? Procedures and sources of error. *J. Exp. Bot.* **2003**, *54*, 2393–2401. [\[CrossRef\]](#) [\[PubMed\]](#)
23. Lv, Z.; Meng, R.; Man, J.; Zeng, L.; Wang, M.; Xu, B.; Gao, R.; Sun, R.; Zhao, F. Modeling of winter wheat fAPAR by integrating Unmanned Aircraft Vehicle-based optical, structural and thermal measurement. *Int. J. Appl. Earth Obs. Geoinf.* **2021**, *102*, 102407. [\[CrossRef\]](#)
24. Viña, A.; Gitelson, A.A. New developments in the remote estimation of the fraction of absorbed photosynthetically active radiation in crops. *Geophys. Res. Lett.* **2005**, *32*, L17403. [\[CrossRef\]](#)
25. Atzberger, C. Advances in remote sensing of agriculture: Context description, existing operational monitoring systems and major information needs. *Remote Sens.* **2013**, *5*, 949–981. [\[CrossRef\]](#)
26. Shoshany, M.; Goldshleger, N.; Chudnovsky, A. Monitoring of agricultural soil degradation by remote-sensing methods: A review. *Int. J. Remote Sens.* **2013**, *34*, 6152–6181. [\[CrossRef\]](#)
27. Shanmugapriya, P.; Rathika, S.; Ramesh, T.; Janaki, P. Applications of remote sensing in agriculture—A Review. *Int. J. Curr. Microbiol. Appl. Sci.* **2019**, *8*, 2270–2283. [\[CrossRef\]](#)
28. Abdullridha, J.; Ampatzidis, Y.; Ehsani, R.; de Castro, A.I. Evaluating the performance of spectral features and multivariate analysis tools to detect laurel wilt disease and nutritional deficiency in avocado. *Comput. Electron. Agric.* **2018**, *155*, 203–211. [\[CrossRef\]](#)
29. Grzybowski, M.; Wijewardane, N.K.; Atefi, A.; Ge, Y.; Schnable, J.C. Hyperspectral reflectance-based phenotyping for quantitative genetics in crops: Progress and challenges. *Plant Commun.* **2021**, *2*, 100209. [\[CrossRef\]](#)
30. Virnodkar, S.S.; Pachghare, V.K.; Patil, V.; Jha, S.K. Remote sensing and machine learning for crop water stress determination in various crops: A critical review. *Precis. Agric.* **2020**, *21*, 1121–1155. [\[CrossRef\]](#)
31. Kwon, Y.S.; Pyo, J.; Kwon, Y.-H.; Duan, H.; Cho, K.H.; Park, Y. Drone-based hyperspectral remote sensing of cyanobacteria using vertical cumulative pigment concentration in a deep reservoir. *Remote Sens. Environ.* **2020**, *236*, 111517. [\[CrossRef\]](#)
32. Blackburn, G.A. Wavelet decomposition of hyperspectral data: A novel approach to quantifying pigment concentrations in vegetation. *Int. J. Remote Sens.* **2007**, *28*, 2831–2855. [\[CrossRef\]](#)
33. Bandyopadhyay, D.; Bhavsar, D.; Pandey, K.; Gupta, S.; Roy, A. Red Edge Index as an Indicator of Vegetation Growth and Vigor Using Hyperspectral Remote Sensing Data. *Proc. Natl. Acad. Sci. India Sect. A Phys. Sci.* **2017**, *87*, 879–888. [\[CrossRef\]](#)
34. Prey, L.; Von Bloh, M.; Schmidhalter, U. Evaluating RGB Imaging and Multispectral Active and Hyperspectral Passive Sensing for Assessing Early Plant Vigor in Winter Wheat. *Sensors* **2018**, *18*, 2931. [\[CrossRef\]](#) [\[PubMed\]](#)
35. Pôças, I.; Rodrigues, A.; Gonçalves, S.; Costa, P.M.; Gonçalves, I.; Pereira, L.S.; Cunha, M. Predicting Grapevine Water Status Based on Hyperspectral Reflectance Vegetation Indices. *Remote Sens.* **2015**, *7*, 16460–16479. [\[CrossRef\]](#)
36. Zhou, J.-J.; Zhang, Y.-H.; Han, Z.-M.; Liu, X.-Y.; Jian, Y.-F.; Hu, C.-G.; Dian, Y.-Y. Evaluating the performance of hyperspectral leaf reflectance to detect water stress and estimation of photosynthetic capacities. *Remote Sens.* **2021**, *13*, 2160. [\[CrossRef\]](#)

37. Han, N.; Zhang, B.; Liu, Y.; Peng, Z.; Zhou, Q.; Wei, Z. Rapid Diagnosis of Nitrogen Nutrition Status in Summer Maize over Its Life Cycle by a Multi-Index Synergy Model Using Ground Hyperspectral and UAV Multispectral Sensor Data. *Atmosphere* **2022**, *13*, 122. [\[CrossRef\]](#)
38. Lelong, C.; Lanore, M.; Caliman, J.-P. Evaluation of hyperspectral remote sensing relevance to estimate oil palm trees nutrition status remote sensing. In Proceedings of the International Symposium on Recent Advances in Quantitative Remote Sensing, Valence, Spain, 25–29 September 2006; pp. 147–152.
39. de Oliveira, M.R.R.; Queiroz, T.R.G.; dos Santos Teixeira, A.; Moreira, L.C.J.; de Oliveira Leão, R.A. Reflectance spectrometry applied to the analysis of nitrogen and potassium deficiency in cotton. *RCA* **2020**, *51*, e20196705. [\[CrossRef\]](#)
40. Marang, I.J.; Filippi, P.; Weaver, T.B.; Evans, B.J.; Whelan, B.M.; Bishop, T.F.A.; Murad, M.O.F.; Al-Shammari, D.; Roth, G. Machine Learning Optimised Hyperspectral Remote Sensing Retrieves Cotton Nitrogen Status. *Remote Sens.* **2021**, *13*, 1428. [\[CrossRef\]](#)
41. Yamashita, H.; Sonobe, R.; Hirono, Y.; Morita, A.; Ikka, T. Dissection of hyperspectral reflectance to estimate nitrogen and chlorophyll contents in tea leaves based on machine learning algorithms. *Sci. Rep.* **2020**, *10*, 17360. [\[CrossRef\]](#)
42. Tan, K.; Wang, S.; Song, Y.; Liu, Y.; Gong, Z. Estimating nitrogen status of rice canopy using hyperspectral reflectance combined with BPSO–SVR in cold region. *Chemom. Intell. Lab. Syst.* **2018**, *172*, 68–79. [\[CrossRef\]](#)
43. Yi, Q.-X.; Huang, J.-F.; Wang, F.-M.; Wang, X.-Z.; Liu, Z.-Y. Monitoring Rice Nitrogen Status Using Hyperspectral Reflectance and Artificial Neural Network. *Environ. Sci. Technol.* **2007**, *41*, 6770–6775. [\[CrossRef\]](#) [\[PubMed\]](#)
44. Goel, P.K.; Prasher, S.O.; Patel, R.M.; Landry, J.A.; Bonnell, R.B.; Viau, A.A. Classification of hyperspectral data by decision trees and artificial neural networks to identify weed stress and nitrogen status of corn. *Comput. Electron. Agric.* **2003**, *39*, 67–93. [\[CrossRef\]](#)
45. Cilia, C.; Panigada, C.; Rossini, M.; Meroni, M.; Busetto, L.; Amaducci, S.; Boschetti, M.; Picchi, V.; Colombo, R. Nitrogen Status Assessment for Variable Rate Fertilization in Maize through Hyperspectral Imagery. *Remote Sens.* **2014**, *6*, 6549–6565. [\[CrossRef\]](#)
46. Thorp, K.R.; Wang, G.; Bronson, K.F.; Badaruddin, M.; Mon, J. Hyperspectral data mining to identify relevant canopy spectral features for estimating durum wheat growth, nitrogen status, and grain yield. *Comput. Electron. Agric.* **2017**, *136*, 1–12. [\[CrossRef\]](#)
47. Ling, B.; Goodin, D.G.; Raynor, E.J.; Joern, A. Hyperspectral analysis of leaf pigments and nutritional elements in tallgrass prairie vegetation. *Front. Plant Sci.* **2019**, *10*, 142. [\[CrossRef\]](#)
48. Li, X.-Y.; Liu, G.-S.; Yang, Y.-F.; Zhao, C.-H.; Yu, Q.-W.; Song, S.-X. Relationship between hyperspectral parameters and physiological and biochemical indexes of flue-cured tobacco leaves. *Agric. Sci. China* **2007**, *6*, 665–672. [\[CrossRef\]](#)
49. Jin, J.; Arief Pratama, B.; Wang, Q. Tracing Leaf Photosynthetic Parameters Using Hyperspectral Indices in an Alpine Deciduous Forest. *Remote Sens.* **2020**, *12*, 1124. [\[CrossRef\]](#)
50. Inoue, Y.; Peñuelas, J.; Miyata, A.; Mano, M. Normalized difference spectral indices for estimating photosynthetic efficiency and capacity at a canopy scale derived from hyperspectral and CO<sub>2</sub> flux measurements in rice. *Remote Sens. Environ.* **2008**, *112*, 156–172. [\[CrossRef\]](#)
51. Clevers, J.G.; Kooistra, L. Using hyperspectral remote sensing data for retrieving canopy chlorophyll and nitrogen content. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2011**, *5*, 574–583. [\[CrossRef\]](#)
52. Hennessy, A.; Clarke, K.; Lewis, M. Hyperspectral Classification of Plants: A Review of Waveband Selection Generalisability. *Remote Sens.* **2020**, *12*, 113. [\[CrossRef\]](#)
53. Zhao, D.; Raja Reddy, K.; Kakani, V.G.; Read, J.J.; Carter, G.A. Corn (*Zea mays* L.) growth, leaf pigment concentration, photosynthesis and leaf hyperspectral reflectance properties as affected by nitrogen supply. *Plant Soil* **2003**, *257*, 205–218. [\[CrossRef\]](#)
54. Yi, Q.; Jiapaer, G.; Chen, J.; Bao, A.; Wang, F. Different units of measurement of carotenoids estimation in cotton using hyperspectral indices and partial least square regression. *ISPRS J. Photogramm. Remote Sens.* **2014**, *91*, 72–84. [\[CrossRef\]](#)
55. Yin, C.; Lin, J.; Ma, L.; Zhang, Z.; Hou, T.; Zhang, L.; Lv, X. Study on the Quantitative Relationship Among Canopy Hyperspectral Reflectance, Vegetation Index and Cotton Leaf Nitrogen Content. *J. Indian Soc. Remote Sens.* **2021**, *49*, 1787–1799. [\[CrossRef\]](#)
56. Meacham-Hensold, K.; Fu, P.; Wu, J.; Serbin, S.; Montes, C.M.; Ainsworth, E.; Guan, K.; Dracup, E.; Pederson, T.; Driever, S.; et al. Plot-level rapid screening for photosynthetic parameters using proximal hyperspectral imaging. *J. Exp. Bot.* **2020**, *71*, 2312–2328. [\[CrossRef\]](#) [\[PubMed\]](#)
57. Cao, C.; Wang, T.; Gao, M.; Li, Y.; Li, D.; Zhang, H. Hyperspectral inversion of nitrogen content in maize leaves based on different dimensionality reduction algorithms. *Comput. Electron. Agric.* **2021**, *190*, 106461. [\[CrossRef\]](#)
58. Mojaradi, B.; Abrishami-Moghaddam, H.; Zoj, M.J.V.; Duin, R.P.W. Dimensionality Reduction of Hyperspectral Data via Spectral Feature Extraction. *IEEE Trans. Geosci. Remote Sens.* **2009**, *47*, 2091–2105. [\[CrossRef\]](#)
59. Yoder, B.J.; Waring, R.H. The normalized difference vegetation index of small Douglas-fir canopies with varying chlorophyll concentrations. *Remote Sens. Environ.* **1994**, *49*, 81–91. [\[CrossRef\]](#)
60. Peñuelas, J.; Gamon, J.A.; Fredeen, A.L.; Merino, J.; Field, C.B. Reflectance indices associated with physiological changes in nitrogen- and water-limited sunflower leaves. *Remote Sens. Environ.* **1994**, *48*, 135–146. [\[CrossRef\]](#)
61. Vogelmann, T.C. Plant tissue optics. *Annu. Rev. Plant Biol.* **1993**, *44*, 231–251. [\[CrossRef\]](#)
62. Carvalho, R.F.; Takaki, M.; Azevedo, R.A. Plant pigments: The many faces of light perception. *Acta Physiol. Plant.* **2011**, *33*, 241–248. [\[CrossRef\]](#)
63. Peng, Y.; Zhu, X.; Xiong, J.; Yu, R.; Liu, T.; Jiang, Y.; Yang, G. Estimation of Nitrogen Content on Apple Tree Canopy through Red-Edge Parameters from Fractional-Order Differential Operators using Hyperspectral Reflectance. *J. Indian Soc. Remote Sens.* **2021**, *49*, 377–392. [\[CrossRef\]](#)

64. Yao, X.; Huang, Y.; Shang, G.; Zhou, C.; Cheng, T.; Tian, Y.; Cao, W.; Zhu, Y. Evaluation of Six Algorithms to Monitor Wheat Leaf Nitrogen Concentration. *Remote Sens.* **2015**, *7*, 14939–14966. [\[CrossRef\]](#)
65. Kamilaris, A.; Prenafeta-Boldú, F.X. Deep learning in agriculture: A survey. *Comput. Electron. Agric.* **2018**, *147*, 70–90. [\[CrossRef\]](#)
66. Liakos, K.G.; Busato, P.; Moshou, D.; Pearson, S.; Bochtis, D. Machine learning in agriculture: A review. *Sensors* **2018**, *18*, 2674. [\[CrossRef\]](#) [\[PubMed\]](#)
67. Fu, P.; Meacham-Hensold, K.; Guan, K.; Bernacchi, C.J. Hyperspectral leaf reflectance as proxy for photosynthetic capacities: An ensemble approach based on multiple machine learning algorithms. *Front. Plant Sci.* **2019**, *10*, 730. [\[CrossRef\]](#)
68. Semeraro, T.; Mastroleo, G.; Pomes, A.; Luvisi, A.; Gissi, E.; Aretano, R. Modelling fuzzy combination of remote sensing vegetation index for durum wheat crop analysis. *Comput. Electron. Agric.* **2019**, *156*, 684–692. [\[CrossRef\]](#)
69. Yi, Q.-X.; Bao, A.-M.; Wang, Q.; Zhao, J. Estimation of leaf water content in cotton by means of hyperspectral indices. *Comput. Electron. Agric.* **2013**, *90*, 144–151. [\[CrossRef\]](#)
70. Yan, L.; Xiong, C.; Qu, H.; Liu, C.; Chen, W.; Zheng, L. Non-destructive determination and visualisation of insoluble and soluble dietary fibre contents in fresh-cut celeries during storage periods using hyperspectral imaging technique. *Food Chem.* **2017**, *228*, 249–256. [\[CrossRef\]](#) [\[PubMed\]](#)
71. Shen, L.; Gao, M.; Yan, J.; Li, Z.-L.; Leng, P.; Yang, Q.; Duan, S.-B. Hyperspectral Estimation of Soil Organic Matter Content using Different Spectral Preprocessing Techniques and PLSR Method. *Remote Sens.* **2020**, *12*, 1206. [\[CrossRef\]](#)
72. Qiao, X.-X.; Wang, C.; Feng, M.-C.; Yang, W.-D.; Ding, G.-W.; Sun, H.; Liang, Z.-Y.; Shi, C.-C. Hyperspectral estimation of soil organic matter based on different spectral preprocessing techniques. *Spectrosc. Lett.* **2017**, *50*, 156–163. [\[CrossRef\]](#)
73. Zhu, S.; Zhou, L.; Gao, P.; Bao, Y.; He, Y.; Feng, L. Near-Infrared Hyperspectral Imaging Combined with Deep Learning to Identify Cotton Seed Varieties. *Molecules* **2019**, *24*, 3268. [\[CrossRef\]](#) [\[PubMed\]](#)
74. Bajcsy, P.; Groves, P. Methodology for Hyperspectral Band Selection. *Photogramm. Eng. Remote Sens.* **2004**, *70*, 793–802. [\[CrossRef\]](#)
75. Sun, K.; Geng, X.; Ji, L. An efficient unsupervised band selection method based on an autocorrelation matrix for a hyperspectral image. *Int. J. Remote Sens.* **2014**, *35*, 7458–7476. [\[CrossRef\]](#)
76. Sonobe, R.; Hirono, Y.; Oi, A. Non-Destructive Detection of Tea Leaf Chlorophyll Content Using Hyperspectral Reflectance and Machine Learning Algorithms. *Plants* **2020**, *9*, 368. [\[CrossRef\]](#) [\[PubMed\]](#)
77. Yuan, H.; Yang, G.; Li, C.; Wang, Y.; Liu, J.; Yu, H.; Feng, H.; Xu, B.; Zhao, X.; Yang, X. Retrieving Soybean Leaf Area Index from Unmanned Aerial Vehicle Hyperspectral Remote Sensing: Analysis of RF, ANN, and SVM Regression Models. *Remote Sens.* **2017**, *9*, 309. [\[CrossRef\]](#)
78. Yang, X.; Yang, R.; Ye, Y.; Yuan, Z.; Wang, D.; Hua, K. Winter wheat SPAD estimation from UAV hyperspectral data using cluster-regression methods. *Int. J. Appl. Earth Obs. Geoinf.* **2021**, *105*, 102618. [\[CrossRef\]](#)
79. Luo, X.; Xu, L.; Huang, P.; Wang, Y.; Liu, J.; Hu, Y.; Wang, P.; Kang, Z. Nondestructive Testing Model of Tea Polyphenols Based on Hyperspectral Technology Combined with Chemometric Methods. *Agriculture* **2021**, *11*, 673. [\[CrossRef\]](#)
80. Zhou, S.; Sun, L.; Xing, W.; Feng, G.; Ji, Y.; Yang, J.; Liu, S. Hyperspectral imaging of beet seed germination prediction. *Infrared Phys. Technol.* **2020**, *108*, 103363. [\[CrossRef\]](#)
81. Gao, W.; Zhou, L.; Liu, S.; Guan, Y.; Gao, H.; Hui, B. Machine learning prediction of lignin content in poplar with Raman spectroscopy. *Bioresour. Technol.* **2022**, *348*, 126812. [\[CrossRef\]](#)
82. Ge, X.; Sun, J.; Lu, B.; Chen, Q.; Xun, W.; Jin, Y. Classification of oolong tea varieties based on hyperspectral imaging technology and BOSS-LightGBM model. *J. Food Process Eng.* **2019**, *42*, e13289. [\[CrossRef\]](#)
83. Yan, J.; Xu, Y.; Cheng, Q.; Jiang, S.; Wang, Q.; Xiao, Y.; Ma, C.; Yan, J.; Wang, X. LightGBM: Accelerated genomically designed crop breeding through ensemble learning. *Genome Biol.* **2021**, *22*, 271. [\[CrossRef\]](#) [\[PubMed\]](#)
84. Huang, Y.; Li, J.; Yang, R.; Wang, F.; Li, Y.; Zhang, S.; Wan, F.; Qiao, X.; Qian, W. Hyperspectral Imaging for Identification of an Invasive Plant *Mikania micrantha* Kunth. *Front. Plant Sci.* **2021**, *12*, 626516. [\[CrossRef\]](#)
85. Lai, Y.; Ni, Y.; Kokot, S. Discrimination of Rhizoma *Corydalis* from two sources by near-infrared spectroscopy supported by the wavelet transform and least-squares support vector machine methods. *Vib. Spectrosc.* **2011**, *56*, 154–160. [\[CrossRef\]](#)
86. Lang, M.; Binder, M.; Richter, J.; Schratz, P.; Pfisterer, F.; Coors, S.; Au, Q.; Casalicchio, G.; Kotthoff, L.; Bischl, B. mlr3: A modern object-oriented machine learning framework in R. *JOSS* **2019**, *4*, 1903. [\[CrossRef\]](#)
87. Valero-Mora, P.M. ggplot2: Elegant Graphics for Data Analysis. *J. Stat. Soft.* **2010**, *35*, 160–167. [\[CrossRef\]](#)
88. Chang, C.-W.; Laird, D.A.; Mausbach, M.J.; Hurburgh, C.R. Near-Infrared Reflectance Spectroscopy–Principal Components Regression Analyses of Soil Properties. *Soil Sci. Soc. Am. J.* **2001**, *65*, 480–490. [\[CrossRef\]](#)
89. Razakamanarivo, R.H.; Grinand, C.; Razafindrakoto, M.A.; Bernoux, M.; Albrecht, A. Mapping organic carbon stocks in eucalyptus plantations of the central highlands of Madagascar: A multiple regression approach. *Geoderma* **2011**, *162*, 335–346. [\[CrossRef\]](#)

90. Du, C.; Zhou, J.; Wang, H.; Chen, X.; Zhu, A.; Zhang, J. Determination of soil properties using Fourier transform mid-infrared photoacoustic spectroscopy. *Vib. Spectrosc.* **2009**, *49*, 32–37. [[CrossRef](#)]
91. Hucklesby, D.P.; Brown, C.M.; Howell, S.E.; Hageman, R.H. Late Spring Applications of Nitrogen for Efficient Utilization and Enhanced Production of Grain and Grain Protein of Wheat1. *Agron. J.* **1971**, *63*, 274–276. [[CrossRef](#)]

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