Maize Characteristics Estimation and Classification by Spectral Data under Two Soil Phosphorus Levels

Baiyu Qiao¹, Xiongkui He¹,²,*, Yajia Liu¹,², Hao Zhang¹, Lanting Zhang¹, Limin Liu¹, Alice-Jacqueline Reineke³, Wenxin Liu⁴ and Joachim Müller³

¹ College of Science, China Agricultural University, Beijing 100193, China; b20183100699@cau.edu.cn (B.Q.); liuyajia@cau.edu.cn (Y.L.); SY20203192920@cau.edu.cn (H.Z.); S20203101932@cau.edu.cn (L.Z.); B20193100727@cau.edu.cn (L.L.)
² College of Agricultural Unmanned System, China Agricultural University, Beijing 100193, China
³ Department of Agricultural Engineering, University of Hohenheim, 70593 Stuttgart, Germany; a.reineke@uni-hohenheim.de (A.-J.R.); joachim.mueller@uni-hohenheim.de (J.M.)
⁴ College of Agronomy and Biotechnology, China Agricultural University, Beijing 100193, China; wenxinliu@cau.edu.cn
* Correspondence: xiongkui@cau.edu.cn; Tel.: +86-010-62732446

Abstract: As an essential element, the effect of Phosphorus (P) on plant growth is very significant. In the early growth stage of maize, it has a high sensitivity to the deficiency of phosphorus. The main purpose of this paper is to monitor the maize status under two phosphorus levels in soil by a nondestructive testing method and identify different phosphorus treatments by spectral data. Here, the Analytical Spectral Devices (ASD) spectrometer was used to obtain canopy spectral data of 30 maize inbred lines in two P-level fields, whose reflectance differences were compared and the sensitive bands of P were discovered. Leaf Area Index (LAI) and yield under two P levels were quantitatively analyzed, and the responses of different varieties to P content in soil were observed. In addition, the correlations between 13 vegetation indexes and eight phenotypic parameters were compared under two P levels so as to find out the best vegetation index for maize characteristics estimation. A Back Propagation (BP) neural network was used to evaluate leaf area index and yield, and the corresponding prediction model was established. In order to classify different P levels of soil, the method of support vector machine (SVM) was applied. The results showed that the sensitive bands of P for maize canopy included 763 nm, 815 nm, and 900–1000 nm. P-stress had a significant effect on LAI and yield of most varieties, whose reduction rate reached 41% as a whole. In addition, it was found that the correlations between vegetation indexes and phenotypic parameters were weakened under low-P level. The regression coefficients of 0.75 and 0.5 for the prediction models of LAI and yield were found by combining the spectral data under two P levels. For the P-level identification in soil, the classification accuracy could reach above 86%. These abilities potentially allow for phenotypic parameters prediction of maize plants by spectral data and different phosphorus contents identification with unknown phosphorus fertilizer status.

Keywords: phosphorus; hyperspectral reflectance; maize; LAI; yield

1. Introduction

Maize is one of the most important and strongly expanding agricultural crops worldwide, having the potential for genetic adaptation to a wide climatic range. Global climate change may lead to abiotic stresses (such as low temperature, high temperature, drought, salinity, and so on), thus adversely affecting crop growth and reducing yields. According to the statistics, climate change causes corn yield to vary by approximately 30%, so it constitutes an important food security issue [1]. Additionally, maize is a cereal with a relatively high phosphate demand and a high sensitivity to phosphate-deficiency, particularly in the early growth stage. Hyperspectral detection technology has become a hotspot of crop...
nutrition detection because of its non-destructive, rapid, and real-time characteristics. It is expected to develop into the main technology for real-time diagnosis of crop nutrition, a basis for online information monitoring of crop nutrition and timely fertilization, thereby achieving automated and intelligent management of agricultural production. The basis of crop assessment using remote sensing is the relationship between crop characteristics and canopy spectral characteristics [2]. One of the most important aspects of agricultural remote sensing is to analyze the reflection characteristics of a canopy spectrum and then to estimate the plant nutritional status during crop growth [3]. It becomes possible to quantify the nutritional status, yield, and other parameters of the fields in a non-destructive manner by using hyper-spectrometer, GNSS (global navigation satellite positioning system), and other sensors [4].

At present, many vegetation biophysical properties for remote sensing estimation have been developed. The most widely used algorithm type is the VIs (vegetation indexes), which is the mathematical combination of visible and near-infrared reflection bands. The spectral reflectance of crop canopy in visible and near-infrared regions is mainly affected by the changes in its structure [5]. In the visible region, it interacts with electromagnetic energy; in the near-infrared region, it is affected by the leaf structure [6]. Ben Zhao et al. [7] comprehensively analyzed the canopy spectral reflectance of visible and near-infrared regions for different maize varieties, which provides a reference for the spectral research of this study. Leaf area index and yield, as the most important indexes to evaluate maize growth, are the keys of phenotypic characteristics research [8,9]. Detailed leaf area index maps are of great value to growers, as they can use these maps to optimize decision-making and crop management based on spatial and temporal scales [10]. Traditionally, leaf area index and yield of maize are measured by time-consuming, costly, and destructive methods, which are difficult to apply in the region. Remote sensing monitoring through airborne or ground equipment is a valuable technology to solve this problem [11,12]. Although many studies have tested canopy vegetation characteristics through phenotype testing technology, few studies have explored corn yield assessment [13]. When chemical traits relevant to growth properties (such as water, nutrient content, and leaf area index) have sufficient variability, spectral data are expected to be related to yield [14].

Phosphorus is a component of ADP and ATP, which directly affect almost all the energy required by crops. Therefore, the ability of using remote sensing technology to monitor the status of phosphorus in plants is very important. The main reason for plants to obtain phosphorus is the symbiotic relationship between plant roots and microorganisms [15]. Excessive or insufficient application of phosphorus fertilizer will have a negative impact on crop growth, resulting in a decline in yield and economic losses of farmers [16]. Therefore, it is very necessary to optimize phosphorus fertilizer management, reduce phosphorus fertilizer investment, and improve phosphorus fertilizer utilization efficiency under limited resources, the premise of which is to monitor the phosphorus status of crops timely and accurately. Many researchers at home and abroad have carried out relevant monitoring and research on the state of phosphorus. Zhang et al. [17] used 32P marker technology to quantify the phosphorus absorption, which improved the sustainability of phosphorus availability in poor soil. Malmir et al. [18] used chemometric analysis methods to show that a visible-near-infrared hyperspectral imaging system could predict the content of essential elements such as nitrogen, phosphorus, potassium, and calcium in leaves.

Although many studies have established the models of phosphorus content in leaves or soil, its universality may be limited due to the lack of spectral absorption characteristics related to phosphorus [19,20]. Therefore, it is very important to explore a non-destructive, timely, and rapid monitoring technology for acquiring the phosphorus status [21]. The framework for this paper is divided into several parts, including data acquisition, analysis, and model establishment, as shown in Figure 1. The main purposes of this paper are as follows: 1. Compare the differences of maize canopy spectral characteristics under two P levels in soil, and establish a direct relationship between the spectral characteristics and P content in soil; 2. Predict the phenotypic characteristics by hyperspectral data, so as
to obtain growth parameters easily and quickly in the process of maize growth; and 3. Establish a classification model to obtain the phosphorus status in real-time and then to guide the precise application of phosphorus fertilizers.

Figure 1. The workflow of the main processes adopted for the proposed approach.

2. Materials and Methods

2.1. Study Area and Experimental Design

The experiment was conducted in the summer of 2019–2021 at Shangzhuang experimental station (40°08′12.15″ N, 116°10′44.83″ E, 50.21 m above mean sea level) of China Agricultural University, Haidian District, Beijing, as shown in Figure 2. The climate is temperate humid monsoon, with an average annual temperature of 12.5 °C and an average annual precipitation of 628.9 mm. Maize, wheat, and other crops are planted in the entire study area for scientific research. The soil type is sandy loam.
The test fields were used for long-term experiments of phosphorus gradient. P fertilizers had not been applied to the low-P field since 1985, but 45 kg/ha P$_2$O$_5$ was applied before sowing for the normal-P field every year. Additionally, 240 kg/ha N fertilizer was applied in both trials before planting [16,22]. More detailed information about the fields may be found in Li et al. [16]. Before sowing, the Olsen P in the low-P and normal-P trial were measured by taking nine samples uniformly from the 0–20 cm soil, and the average soil phosphorus content was 1.5 mg/kg and 4.7 mg/kg, respectively. In addition, total nitrogen (N) was determined by the Kjeldahl procedure [23]. Available K (exchangeable-K) and P (Olsen-P) were extracted using NH$_4$OAc and NaHCO$_3$ and determined using a flame photometer and spectrophotometer, respectively [24]. The N and K concentrations were 0.63–0.83 mg/kg and 109.2–147.9 mg/kg for the low-P trial, 0.69–0.77 mg/kg and 135.6–140.3 mg/kg for the normal-P trial, respectively. The concentration of P in soil was the main limiting factor and all other management measures remained the same. A total of 364 maize inbred lines were planted in a randomized block design of three replicates for each of two phosphorus treatments, 30 of which were randomly selected for spectral experiments, numbering 180 plots in all, whose corresponding groups are shown in Table 1. The test plot covered an area of 767.5 m$^2$, with a plant spacing of 0.2 m and a row spacing of 0.5 m. In order to mitigate the border effects, two protection rows were arranged at the edge of the test area. In the full text, the abbreviation of LP represents low phosphorus treatment, and NP represents normal phosphorus treatment.
Table 1. Maize inbred lines and corresponding groups.

| Maize Inbred Line | A Group | Maize Inbred Line | A Group | Maize Inbred Line | A Group |
|------------------|---------|------------------|---------|------------------|---------|
| 150              | A428    | 5237             | A350    | 18-599           | A302    |
| 177              | A242    | 5311             | A272    | 303WX            | A344    |
| 238              | A293    | 7381             | A329    | 384-2            | A320    |
| 268              | A370    | 8902             | A243    | 3H-2             | A444    |
| 501              | A317    | 9642             | A326    | 4F1              | A385    |
| 812              | A236    | 04K5686          | A304    | 7884-4Ht         | A313    |
| 1462             | A399    | 04K5702          | A424    | 835a             | A274    |
| 3411             | A381    | 05W002           | A301    | 835b             | A285    |
| 4019             | A412    | 05WN230          | A324    | 975-12           | A287    |
| 5213             | A387    | 07KS4            | A335    | A619             | A357    |

2.2. Spectral Data and Plant Trait Measurement

In general, the reflectance spectrum measurement of field features includes two steps: one is the measurement of the solar irradiance spectrum, and the other is the measurement of the radiance spectrum in the reflection direction of the target. For the measurements of the solar irradiance spectrum under field conditions, a calibrated diffuse reflection plate can be used as a reference panel, whose radiance spectrum is mainly measured by a feature-spectrometer, so as to calculate the irradiance spectrum of the solar light source reaching the ground surface. The calculation formula is as follows [25]:

\[
Ed(\lambda) = \frac{\pi L_{\text{ref}}(\lambda)}{\rho_{\text{ref}}(\lambda)},
\]

where \(L_{\text{ref}}(\lambda)\) is the radiance of the Lambertian reference plate measured by the spectrometer, and \(\rho_{\text{ref}}(\lambda)\) is the reflectance of the Lambertian reference plate.

According to the definition and formula of directional reflectance, as long as the radiance spectrum reflected by the target and the solar incident signal reflected by the reference plate are measured synchronously, the target spectral reflectance can be calculated, that is:

\[
\rho(t, \theta, \phi, \theta_r, \phi_r, \lambda) = \frac{\pi L_t(\theta_r, \phi_r, \lambda)}{Ed(\theta, \phi, \lambda)} = \frac{L_t(\theta_r, \phi_r, \lambda)}{L_{\text{ref}}(\theta, \phi, \lambda)} \rho_{\text{ref}}(\lambda)
\]

where \(L_t(\theta_r, \phi_r, \lambda)\) is the target radiance spectrum observed in the direction of \((\theta_r, \phi_r)\) by the spectrometer, and \(\rho_t(\theta, \phi, \theta_r, \phi_r, \lambda)\) is the calculated directional reflectance spectrum of field features.

In this experiment, an Analytical Spectral Devices (ASD) Field Spec 4 spectrometer (ASD Inc., Boulder, CO, USA) was used for spectral data acquisition, whose view angle was 25° and wavelength range was 350 nm–2500 nm. It was sunny and windless during all test days. Hyperspectral reflectance measurements of maize were made at the jointing growth stage under clear sky conditions between 10:00 and 14:00 local time. The sensor was distanced 1.0 m above the canopy, which ensured that the view field of spectrometer covered the entire testing canopy. During the measurements, the sensor probe was always pointed vertically downward. Black and baseline reflectance was obtained with a 50 × 50 cm calibration panel before each measurement, whose thickness was 10 mm and material was pressed PTFE. An average of 20 spectra was used for each sample measurement.

At the corresponding spectral reflectance measurement positions, the length and width of all leaves for each sample were measured manually to obtain LAI. At the end of the growing season, the ear length, ear diameter, kernals, 100-kernal weight, ear dry weight, and yield were obtained by destructive samplings in each of the 180 plots. The ears were dried at 65 °C for at least 48 h and weighed to determine the dry weight and actual yield.
2.3. Data Analysis

Based on published literature and the sensitivity of optical and polarization features to LAI and yield, optical spectral vegetation indexes were used to estimate LAI and yield. The leaf area index and yield of crops can be estimated from seven bands of red spectra (600 nm, 603 nm, 636 nm, 639 nm, 642 nm, 666 nm, and 669 nm), four bands of near infrared (738 nm, 741 nm, 744 nm, and 747 nm) and two bands close to the edge of blue and green spectra (498 nm and 501 nm) [2,26–30]. In this paper, 15 vegetation indexes from the spectral data were selected to evaluate the characters of maize, as shown in Table 2. Pearson correlations were used to test the relationship among the selected VIs and phenotypic parameters. A BP neural network (again, with a 70/30 calibration/validation split) was employed to predict LAI and yield, whose accuracy was tested by the regression coefficient R (analyzes conducted in MATLAB 2020b). SVM classification was used to classify spectral data for different phosphorus treatments based on the classification learner toolbox environment in MATLAB 2020b. A 10-fold cross validation method was chosen to protect against overfitting, and an ROC curve was used to check the model accuracy. Sensitivity and specificity of the ROC curve were also applied to evaluate the accuracy of the classification model.

| Vegetation Index                         | Formulation                                                                 | Reference |
|------------------------------------------|-----------------------------------------------------------------------------|-----------|
| Simple Ratio                             | $SR = \frac{\rho_{670}}{\rho_{670} + \rho_{800}}$                         | [31]      |
| Normalized Difference Vegetation Index   | $NDVI = \frac{\rho_{670} - \rho_{700}}{\rho_{670} + \rho_{700}}$           | [32]      |
| Enhanced Vegetation Index                | $EVI = 2.5 \left[ \frac{\rho_{490} + \rho_{610} + \rho_{700} + \rho_{800}}{\rho_{670} + \rho_{700} + \rho_{800} + \rho_{910}} \right]$ | [33]      |
| Soil Adjusted Vegetation Index           | $SAVI = \frac{\rho_{670} - \rho_{700}}{\rho_{670} + \rho_{700}}$          | [32]      |
| Enhanced vegetation Index 2              | $EVI2 = 2.5 \left[ \frac{\rho_{550} + \rho_{670} + \rho_{700} + \rho_{800}}{\rho_{670} + \rho_{700} + \rho_{800} + \rho_{910}} \right]$ | [34]      |
| Normalized Green Red Difference Index    | $NGRDI = \frac{\rho_{670} - \rho_{700}}{\rho_{700} + \rho_{800}}$         | [35]      |
| Optimized Soil Adjusted Vegetation Index | $OSAVI = (1 + 0.16) \left[ \frac{\rho_{670} + \rho_{700} + \rho_{800}}{\rho_{670} + \rho_{700} + \rho_{800}} \right]$ | [36]      |
| Transformed Chlorophyll Absorption Reflectance Index | $TCARI = 3 \left[ \frac{(\rho_{670} - \rho_{700}) - 0.2(\rho_{700} - \rho_{910}) + 0.2(\rho_{910} - \rho_{670})}{\rho_{670} + \rho_{700} + \rho_{800}} \right]$ | [37]      |
| Triangular Greenness Index               | $TGI = -0.5[(670 - 490)(\rho_{700} - \rho_{910}) - (670 - 550)(\rho_{670} - \rho_{280})] - 0.5$ | [32]      |
| Difference Vegetation Index              | $DVI = \frac{\rho_{910}}{\rho_{700} - \rho_{800}}$                         | [31]      |
| Green Normalized Difference Vegetation Index | $GNDVI = \frac{\rho_{910} - \rho_{700}}{\rho_{910} - \rho_{700} + \rho_{800}}$ | [38]      |
| Modified Soil Adjusted Vegetation Index  | $MSAVI = \frac{\rho_{910} - \rho_{700}}{\rho_{910} - \rho_{700} + \rho_{800}}$ | [39]      |

3. Results

3.1. Spectral Characteristics under Different P Levels

The spectral characteristics of plants in the visible band were mainly determined by the concentration of photosynthesis, including chlorophyll a, chlorophyll b, lutein, anthocyanin, and carotenoid. In the near-infrared band, the internal structure of leaves, the size and shape of plant air cavities, and the number of air–water interfaces of plant membranes had a great impact on the intensity of leaf spectral reflectance. Other influencing factors included water content and biochemical concentrations of lignin, cellulose, starch, and protein [40]. Because the time-dependent change in phosphorus would lead to a change in chlorophyll content, the change in leaf reflectivity monitored by spectral sensor could be used as an index to detect the phosphorus status of crops. And due to the comprehensive effects of chlorophyll absorption and mesophyll light dispersion in plant leaves, healthy leaves have a high reflectivity in the red to near-infrared region [30]. Osco et al. [41] used several different machine learning algorithms to predict the hyperspectral nutritional components of citrus leaves and determined the most contributing wavelength or spectral region related to each nutrient element. The measured wavelength range related to phosphorus were 385–411 nm, 438–456 nm, 472–477 nm, 502 nm, 521 nm, 527 nm, and 544–555 nm. However, more band combinations were not involved.

The spectral curves of 30 different maize varieties were divided into two parts, which included the visible light band (400–780 nm) and the near-infrared short-wave region...
(780–1100 nm). The curves of different colors in the following figures represented different maize varieties. As shown in Figure 3 and Figure 5, the field-measured spectra all had typical characteristics of a vegetation spectrum, and these characteristics were similar for different maize varieties and P levels. Specifically, all spectra presented two absorption peaks in the blue-violet and red regions, a strong reflection peak in the green region, and a high reflection platform in the region beyond 680 nm. Within the visible light band, the canopy reflectance of NP was significantly higher than that of LP. There was an obvious change near 763 nm, which can be clearly seen in the first derivative curve of reflectance in Figure 4. In the near-infrared short-wave region, P deficiency could weaken the fluctuation of reflectivity within 815 nm and 900–1000 nm bands, as shown in Figure 5. This result was similar to the conclusion obtained by Osborneel et al. [42], who also demonstrated that the increased plant cells number per leaf area under phosphorus stress had been translated into a significant spectral response in the near-infrared spectrum in maize.

Figure 3. Spectral reflectance curve in the visible light range.
Figure 4. First order spectral reflectance curve in the visible light range.

Figure 5. Spectral reflectance curve in the near-infrared short-wave range.
3.2. Quantitative Analysis of LAI and Yield under Different P Levels

Many biological and physical processes of crops, including interception of light and water (rain and fog), attenuation, transpiration, photosynthesis, and autotrophic respiration of light through the canopy, and carbon and nutrients cycles (such as nitrogen, phosphorus, etc.) can control LAI. It can be seen from Table 3 that the variation range of LAI in the LP plot was 0.5–5.4, and that in the NP plot was 0.4–7.2. The yield of the LP plot was typically 500–2000 kg/hm, and that of the NP plot was typically 1500–3500 kg/hm, with obvious variation trends, which were mainly because the P deficiency made photosynthetic capacity decrease, resulting in the reduction in maize LAI and yield. As early as 2004, Ayala-Silva et al. [43] observed a significant decrease in chlorophyll content in wheat plants grown in an independent phosphorus-deficiency greenhouse. The change in phosphorus content would have a bigger effect on leaf area index than on yield, whose average CVs were approximately 0.46 and 1.12. In other words, the variation range of yield was larger, that is, the CV was also larger. Under the two P levels, there was a significant positive correlation between LAI and yield, and the correlation coefficients were 0.34 and 0.49, respectively.

Table 3. Effect of different phosphorus treatments on leaf area index and yield of maize.

| Sample Size | LAI       | Yield (kg/hm) |
|-------------|-----------|---------------|
| LP          | NP        | LP            | NP            |
| Mean        | 1.66      | 2.85          | 1237.51       | 2098.49       |
| Min         | 0.52      | 0.38          | 0             | 0             |
| Max         | 5.46      | 7.16          | 10,770        | 12,970        |
| Range       | 4.95      | 6.78          | 10,770        | 12,970        |
| SD          | 0.83      | 1.20          | 1533.46       | 2116.76       |
| CV          | 0.50      | 0.42          | 1.24          | 1.00          |
| Rd (%)      | 41.75     | 41.02         |
| r           | 0.34 **   | 0.49 ***      |

Notes: SD represents the standard deviation; Rd represents the relative reduction under low-P stress calculated by (mean (NP)-mean (LP))/mean (NP); CV represents the coefficient of variation; r represents the correlation coefficient, which is related to LAI and yields between low-P (LP) and normal-P treatment (NP). ** p < 0.01, *** p < 0.001.

Regardless of variety, the average yield of maize under P stress decreased by 996.3kg/hm, and the influence rate of P content on LAI and yield was basically the same, which remained at approximately 41%. As expected, this was consistent with the views of Li et al. [16]. However, the sensitivity of different maize varieties to soil phosphorus content was different. That is, different varieties had a different tolerance to phosphorus, which may be relevant to the phosphorus absorption for different genotypes in soil. The development of P-efficient maize varieties that can grow and have a good yield under LP supply should be a major breeding target to enable sustainable agricultural production [44]. It can be seen from Figures 6 and 7 that the leaf area index and yield of most varieties in the NP field were obviously higher than those in the LP field, which meant the P stress was very serious during the growth of maize. Different letters in the figures indicate that the effect of P level on leaf area index and yield was significant and the same letter indicates that the impact was not significant. Groups A242, A243, A370, and A412 were not very sensitive to soil phosphorus content, whose LAIs were basically unchanged. The yields of A274, A320, A326, and A412 were also changed slightly, that is, the change in P content had no more obvious effects on yield for these varieties than the others. These types of varieties should be the focus for breeding.
Figure 6. LAI bars of maize under different varieties and P treatment.

Figure 7. Yield bars of maize under different varieties and P treatment.
Some varieties whose yields had an obvious decline were greatly affected by P-stress, such as A287, A293, A301, A304, A329, A344, and A424. This can be seen intuitively in Figure 7. The yield of some varieties was almost zero, such as A320, A326; however, some varieties, such as A243, A272, A370, and A444, whose yields were higher in the LP field than in the NP field, appeared to show the opposite results. There may be several reasons for these results. First, the pathways of P uptake and acquisition are different for various maize genotypes, which causes different sensitivity to phosphorus. Second, it may be caused by insufficient sample size or repetition. Third, some survey errors in the process of the test, such as environmental factors, sampling, or measurements should be considered. In short, these varieties should be given more attention in future experiments. Moreover, developing soil adaptability maize genotypes is imperative to ensure effective utilization of phosphorus resources. Maize yield increased by approximately 2% per year, which was attributed to the improvement of biological and abiotic stress tolerance, from 50–60% to 75% [45]. In other words, after selecting the varieties sensitive to phosphorus, it would be more conducive for us to cultivate more tolerant varieties and field management.

3.3. Phenotypic Study by Vegetation Indexes

At present, many remote sensing studies related to crop phenotypic parameters were based on the calculation of vegetation index. Lin et al. [46] studied the changeable rules of wheat spectral characteristic parameters with intraday phenology, which was used for the inversion of LAI, and determined the optimal spectral parameters and sub period of LAI. Oklahoma State University developed an instrument named Green Viewer based on canopy multispectral analysis to establish a corresponding diagnostic recommendation model for topdressing of wheat and maize [47]. However, most vegetation indexes are species-specific. Viña et al. [28] evaluated the applicability of different vegetation indexes to remote sensing estimation for multi-temporal LAI of crops with different leaf structure and canopy structure. However, it is still urgent to establish corresponding parameter-characterization models for specific crops.

In this paper, the selected 13 vegetation indexes were used to evaluate eight different phenotypic parameters (including leaf area index, yield, plant height, ear length, ear diameter, kernals, 100-kernal weight, and ear dry weight), and the correlation between them was analyzed. From the correlation coefficient diagrams of Figures 8 and 9, it can be seen that the relativity between most vegetation indexes and phenotypic parameters was significant. Specifically, almost all selected vegetation indexes had a good fitting result with leaf area index and plant height, between which the correlation coefficient could reach above 0.7. Similarly, the selected vegetation indexes could also be used as a good signal for yield prediction due to their significant correlation relationship. However, some correlation values were relatively low or even negative, such as the prediction for 100-kernal weight, which showed a poor result. This indicates that other vegetation indexes need to be found to predict it. It was also obvious that the correlation relationship was weakened due to P stress, which was probably because the P stress disturbed the plant internal structures and made the relationship irregular between them, which was helpful to separate the effects of P on the growth characteristic parameters of maize.

In Figures 8 and 9, LAI represents leaf area index, PH represents plant height, 100KS represents 100-kernel weight, KS represents kernals, EDW represents ear dry weight, EL represents ear length, and ED represents ear diameter.
Figure 8. Correlation coefficient matrix of the NP field.

Figure 9. Correlation coefficient matrix of the LP field.
3.4. LAI and Yield Assessment by BP Neural Network

The ability to accurately estimate maize yield using spectral reflectance before harvest can reduce the time and cost of phenotypic analysis. Integrating crop models and remote sensing data has also become an effective method to monitor crop growth status and yield based on large-area regional data assimilation [27,48]. As early as 2012, Weber et al. [49] used a partial least squares regression model to explain that the proportion of maize genetic variation under drought stress was higher than that under watering conditions. In this paper, a BP neural network was used to evaluate LAI and yield under phosphorus stress.

Based on the above analysis, the relationship of Vis, LAI, and yield in this period was quantitatively analyzed and fitted. The data of the two plots were integrated to predict the LAI and yield of maize, the purpose of which was to facilitate the prediction of the two parameters in the fields with uneven distribution of phosphorus fertilizer, so as to be used for field management and fertilization application in the later stage. Thirteen vegetation indexes were used as input and LAI/yield were used as output for model fitting. Bayesian algorithm in the BP neural network was used to fit the LAI and yield data in the two plots, which included 14 samples in the validation set and 10 hidden neurons. It can be seen in Figures 10 and 11 that the LAI and yield prediction models based on vegetation indexes fit well, and the regression coefficients of the prediction models reach more than 0.75 and 0.5, respectively. In particular, data in the figures represented all measured LAIs or yields and estimated LAIs or yields using the training neural network model; target means the measured LAIs or yields; output means estimated LAIs or yields using the training neural network model; fit represents the fitting curve of target and output; Y = T means an ideal result for which the target equals the output. Many studies have predicted the LAI and yield of maize well. Herrmann et al. [14] explored the yield prediction models for different maize development stages, with the best model found, and the respective R^2 value was 0.73. Due to the differences of P content and test fields, the prediction models had to be verified by more experiments. However, it is undeniable that the results provided a certain reference and basis for our field precision fertilization.

![Figure 10. Establishment of Vis–LAI regression model.](image-url)
3.5. SVM Spectral Classification Models

Due to the difference in P concentration in soil, spectral classification models were established to distinguish different P treatments. Sartin et al. [50] studied a large number of models related to nutrient deficiency types. Based on the difference in chlorophyll coloring, the method of artificial neural network was used to segment the leaf image of P-deficiency cotton and proved the superiority of the neural network method in image segmentation. In this paper, the machine learning algorithm was used to classify spectral data of all testing plots according to the P-level by using MATLAB 2020b. The 10-fold cross validation was used to test the accuracy of the model. All SVM algorithms were selected for training. According to the training results, as shown in Figure 12, the area under the curve (AUC) was 0.92, which indicated a good performance and quality of the classifier. The receiver operating characteristic (ROC) curve showed true positive rate versus false positive rate for the currently selected trained classifier. The marker point (0.15,0.86) on the plot showed the performance of the currently selected classifier. A false positive rate (FPR) of 0.15 indicated that the current classifier assigned 15% of the observations incorrectly to the NP level. A true positive rate (TPR) of 0.86 indicated that the current classifier assigns 86% of the observations correctly to the NP level.

In addition, sensitivity and specificity were also used to characterize the quality of the model. The calculations of these values were as follows:

\[
\text{Sensitivity} = \text{True Positive Rate} \quad (3) \\
\text{Specificity} = 1 - \text{False Positive Rate} \quad (4)
\]

Sensitivity was equaled to the TPR, whose value was 86%, which told us 86% of the NP level was correctly classified by the model. The specificity value was 85%, which told us 85% of the LP level was correctly classified. Overall, the accuracy of the quadratic SVM model was highest, and the recognition accuracy was above 86%, which showed that the model could be better used for P-fertilizer application and management in the field.

Figure 11. Establishment of Vis–yield regression model.
4. Discussion

In this study, the canopy reflectance of different maize inbred lines under two soil phosphorus levels was compared. The main purpose was to establish the direct relationship between the spectral characteristics of the maize canopy and soil phosphorus content. The biggest outcome of this research was that it could easily and quickly obtain the phosphorus status in maize soil, so as to guide precise phosphorus fertilizer application. The measurement of hyperspectral soil phosphorus reflectance provided a portable and low-cost method that could be performed in situ [51]. Because the difference in soil phosphorus content has an impact on the phosphorus content in plant leaves, previous studies mainly focused on the relationship between phosphorus content in maize leaves and the canopy spectrum, so as to reflect the phosphorus content level in soil indirectly.

It has been found that not all maize varieties were sensitive to soil phosphorus content. Additionally, many varieties did not show strong sensitivity to soil phosphorus change that did not cause LAI or yield change. Therefore, in the case of limited phosphorus resources, cultivating and selecting more varieties insensitive to phosphorus fertilizer may be a direction that can be considered. Of course, it should also ensure that the yields of maize will meet the actual production demand.

In addition, more vegetation indices should be explored. In this study, the feasibility of using the existing vegetation indexes to predict leaf area index and yield under different soil phosphorus content was verified, but more vegetation indexes and more growth stages should be studied. Different vegetation indices responded differently to the crop leaf optical properties defined by the leaf chlorophyll, dry mass, and water content [52]. The leaf inclination, solar angle, and sensor bandwidth and calibration all affected the vegetation spectral reflectance [53]. Therefore, more attempts of vegetation indexes combinations need to be made to produce better fitting models. Although there are relatively few studies on the prediction of phosphorus content, some studies have explored new spectral bands and vegetation indices for estimating the nitrogen nutrition index of summer maize, whose results showed that the most sensitive spectral bands were located at 710 nm (red edge...
band) and 512 nm (visible light band), and the optimum spectral vegetation index for estimating nitrogen nutrition index (NNI) was NDSI (R710, R512) [7]. It proved the value of using more vegetation indices for future research.

In this paper, the models between VIs and LAI/yield were built, which verified the applicability of the machine learning method in agricultural vegetation nutrition prediction. Using machine learning for analysis has become a trend of prediction, due to its rapidity, accuracy, and convenience for establishing the prediction models of unknown parameters. Leaf hyperspectral reflectance has been used to estimate nutrient concentrations in plants in narrow bands of the electromagnetic spectrum using partial least squares regression (PLSR) [6]. The application of PLS and specific narrow bands vegetation indices reached significant levels of accuracy in the retrieval of P levels, in comparison to traditional broad band indices [54]. However, the prediction accuracy in this study can be further improved, which may be achieved by increasing the sample size. Moreover, phosphorus stress may affect some internal relations between vegetation indexes and growth parameters, resulting in poor fitting results between them. LAI-VIs relationships are only good for the local area where the model was developed [34].

In the process of spectral detection, very specific models were required for different stages, because different growth stages corresponded to spectral changes in different wavelength ranges [42]. The evaluation models of leaf area index and yield at different growth stages were different. Studies have shown that using vegetation indices as inputs could produce the best results to assess the maize yield in the R4 development stage [14]. Needless to say, the availability of more data, including different soil types and spectral data at different growth stages (not just the jointing stage) is called for to address the limitations of this study and validate its principal findings.

In terms of classification, this study mainly classified the soil phosphorus content according to the reflectance of maize canopies during the growth process, which could quickly and non-destructively monitor the soil phosphorus level and guide the fertilization process. Some studies were based on leaf color scale, but it may be time-consuming. For example, an intelligent oilseed rape nutrition deficiency analysis and diagnosis method was proposed based on HSV color space that quantified non-uniform histograms and combined multiple support vector machine classifiers, which had a high identification rate for the four nutrition deficiency types, including normal, phosphorus deficiency, nitrogen deficiency, and boron deficiency [55]. In addition, an intelligent diagnosis method was proposed for nutrition deficiency based on color characteristics of rapeseed leaves by analyzing and extracting external features of different rapeseed parts, whose results showed that nutrition deficiency diagnosis rate was as high as 87.5%, but only rape leaves (canopy) were studied, without including the actual field cultivated rape [56].

In addition, it should be emphasized that the field used in this paper was specifically used for P-level tests for these years, and the soil phosphorus content measured every year basically remained unchanged. Therefore, the influence of soil heterogeneity on the test conditions was excluded, and only the characteristics of maize canopy spectrum under the conditions of different soil phosphorus content were considered. However, the impact of soil heterogeneity on phosphorus content in actual production fields is not excluded, so more research needs to be completed for different farmland environments.

5. Conclusions

Detecting phosphorous status in soil by non-invasive methods has practical guiding significance for field operation. Therefore, this study focused on maize characteristics estimation and classification by spectral data under two soil phosphorus levels. The results showed that the sensitive bands of P included 763 nm, 815 nm, and 900–1000 nm by comparing the reflectance of maize canopy under two P levels, whose variation in the NP field was higher than that of the LP field. Based on the results of leaf area index and yield under two P levels, most maize varieties showed strong sensitivity to P deficiency and the reduced rate could reach 41%. The Pearson correlation between 13 vegetation indexes and
eight phenotypic parameters showed that the correlation in the NP plot was good, whereas that in the LP plot was slightly poor, that is, P-stress weakened the correlations between them. The best vegetation index was selected for the prediction of maize characteristics. A BP neural network method was shown to feasibly assess LAI and yield by combining the spectral data in the two fields. The regression coefficients of the prediction models were 0.75 and 0.5. The best classification model was established based on all spectral data by using the quadratic SVM to distinguish different P treatments, whose accuracy was more than 86%. This study provides a basis and reference for the application of precision phosphorus fertilizer in maize fields.

Author Contributions: Conceptualization, X.H. and J.M.; methodology, Y.L.; software, B.Q.; validation, B.Q., L.L. and H.Z.; formal analysis, A.J.R.; investigation, L.Z.; resources, L.Z.; data curation, H.Z.; writing—original draft preparation, B.Q.; writing—review and editing, X.H.; visualization, X.H.; supervision, Y.L.; project administration, W.L.; funding acquisition, X.H. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation)-328017493/GRK 2366 (Sino-German International Research Training Group AMAIZE-P) and Research project of Agricultural UAV system (2021AC037).

Data Availability Statement: All relevant data are within the paper.

Acknowledgments: We thank Dongdong Li and Zheng Zhao at China Agricultural University for providing materials in our experiment. We also thank Zhichong Wang and Changling Wang at CCAT of China Agricultural University for suggestions on revisions. We sincerely thank the editors and anonymous reviewers for their constructive comments and suggestions on the manuscript.

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

References

1. Ray, D.K.; Gerber, J.S.; MacDonald, G.K.; West, P.C. Climate variation explains a third of global crop yield variability. *Nat. Commun.* 2015, 6, 5989. [CrossRef] [PubMed]

2. Li, H.; Zhao, C.; Yang, G.; Feng, H. Variations in crop variables within wheat canopies and responses of canopy spectral characteristics and derived vegetation indices to different vertical leaf layers and spikes. *Remote Sens. Environ.* 2015, 169, 358–374. [CrossRef]

3. Li, F.; Miao, Y.; Feng, G.; Yuan, F.; Yue, S.; Gao, X.; Liu, Y.; Liu, B.; Ustin, S.L.; Chen, X. Improving estimation of summer maize nitrogen status with red edge-based spectral vegetation indices. *Field Crops Res.* 2014, 157, 111–123. [CrossRef]

4. Gilliot, J.M.; Michelin, J.; Hadjard, D.; Houot, S. An accurate method for predicting spatial variability of maize yield from UAV-based plant height estimation: A tool for monitoring agronomic field experiments. *Precis. Agric.* 2021, 22, 897–921. [CrossRef]

5. Jin, X.; Li, Z.; Feng, H.; Ren, Z.; Li, S. Deep neural network algorithm for estimating maize biomass based on simulated Sentinel 2A vegetation indices and leaf area index. *CROP J.* 2020, 8, 87–97. [CrossRef]

6. Oliveira, L.F.R.D.; Santana, R.C. Estimation of leaf nutrient concentration from hyperspectral reflectance in Eucalyptus using partial least squares regression. *Sci. Agric.* 2020, 77, e20180409. [CrossRef]

7. Zhao, B.; Duan, A.; Ata-Ul-Karim, S.T.; Liu, Z.; Chen, Z.; Gong, Z.; Zhang, J.; Xiao, J.; Liu, Z.; Qin, A.; et al. Exploring new spectral bands and vegetation indices for estimating nitrogen nutrition index of summer maize. *Eur. J. Agron.* 2018, 93, 113–125. [CrossRef]

8. Curnel, Y.; de Wit, A.J.W.; Duveiller, G.; Defourny, P. Potential performances of remotely sensed LAI assimilation in WOFOST model based on an OSS Experiment. *Agric. For. Meteorol.* 2011, 151, 1843–1855. [CrossRef]

9. Xiao, Z.; Liang, S.; Wang, J.; Jiang, B.; Li, X. Real-time retrieval of Leaf Area Index from MODIS time series data. *Remote Sens. Environ.* 2011, 115, 97–106. [CrossRef]

10. Comba, L.; Biglia, A.; Ricauda Aimonino, D.; Tortia, C.; Mania, E.; Guidoni, S.; Gay, P. Leaf Area Index evaluation in vineyards using 3D point clouds from UAV imagery. *Precis. Agric.* 2019, 21, 881–896. [CrossRef]

11. Quebrago, L.; Perez-Ruiz, M.; Perez-Urrestarazu, L.; Martinez, G.; Egea, G. Linking thermal imaging and soil remote sensing to enhance irrigation management of sugar beet. *Biosyst. Eng.* 2018, 165, 77–87. [CrossRef]

12. Khalil, A.; Comba, L.; Biglia, A.; Ricauda Aimonino, D.; Chiaberge, M.; Gay, P. Comparison of Satellite and UAV-Based Multispectral Imagery for Vineyard Variability Assessment. *Remote Sens.* 2019, 11, 436. [CrossRef]

13. Sankaran, S.; Khot, L.R.; Espinoza, C.Z.; Jarolmasjed, S.; Sathuvali, V.R.; VanDeMark, G.J.; Miklas, P.N.; Carter, A.H.; Pumphrey, M.O.; Knowles, N.R.; et al. Low-altitude, high-resolution aerial imaging systems for row and field crop phenotyping: A review. *Eur. J. Agron.* 2015, 70, 112–123. [CrossRef]
43. Ayala-Silva, T.; Beyl, C.A. Changes in spectral reflectance of wheat leaves in response to specific macronutrient deficiency. *Adv. Space Res.* 2005, 35, 305–317. [CrossRef]

44. Vance, C.P.; Uhde-Stone, C.; Allan, D.L. Phosphorus acquisition and use: Critical adaptations by plants for securing a nonrenewable resource. *New Phytol.* 2003, 157, 423–447. [CrossRef]

45. Araus, J.L.; Serret, M.D.; Edmeades, G.O. Phenotyping maize for adaptation to drought. *Front. Physiol.* 2012, 3, 305. [CrossRef]

46. Lin, Y.; Shen, H.; Tian, Q.; Gu, X. Improving leaf area index retrieval using spectral characteristic parameters and data splitting. *Int. J. Remote Sens.* 2019, 41, 1741–1759. [CrossRef]

47. Peng, Q.; Xu, W. Crop Nutrition and Computer Vision Technology. *Wirel. Pers. Commun.* 2021, 117, 887–899. [CrossRef]

48. Fang, H.; Liang, S.; Hoogenboom, G.; Teasdale, J.; Cavigelli, M. Corn-yield estimation through assimilation of remotely sensed data into the CSM-CERES-Maize model. *Int. J. Remote Sens.* 2008, 29, 3011–3032. [CrossRef]

49. Weber, V.S.; Araus, J.L.; Cairns, J.E.; Sanchez, C.; Melchinger, A.E.; Orsini, E. Prediction of grain yield using reflectance spectra of canopy and leaves in maize plants grown under different water regimes. *Field Crops Res.* 2012, 128, 82–90. [CrossRef]

50. Sartin, M.A.; Da Silva, A.C.; Kappes, C. Image segmentation with artificial neural network for nutrient deficiency in cotton crop. *J. Comput. Sci.* 2014, 10, 1084–1093. [CrossRef]

51. Dhawale, N.M.; Adamchuk, V.; Viscarra, R.; Prasher, S.; Whalen, J.K.; Ismail, A. Predicting Extractable Soil Phosphorus Using Visible/Near-Infrared Hyperspectral. Soil Reflectance Measurements; Paper No. CSBE13-047; The Canadian Society for Bioengineering: Renfrew, ON, Canada, 2013.

52. Jacquemoud, S.; Baret, F. PROSPECT: A model of leaf optical properties spectra. *Remote Sens. Environ.* 1990, 34, 75–91. [CrossRef]

53. Jongshaap, R.E. Sensitivity of a crop growth simulation model to variation in LAI and canopy nitrogen used for run-time calibration. *Ecol. Model.* 2007, 200, 89–98. [CrossRef]

54. Pimstein, A.; Karnieli, A.; Bansal, S.K.; Bonfil, D.J. Exploring remotely sensed technologies for monitoring wheat potassium and phosphorus using field spectroscopy. *Field Crops Res.* 2011, 121, 125–135. [CrossRef]

55. Zhang, K.; Zhang, A.; Li, C. Nutrient deficiency diagnosis method for rape leaves using color histogram on HSV space. *Trans. Chin. Soc. Agric. Eng.* 2016, 32, 179–187.

56. Xu, S.; Lin, W.; Wu, W.; Zhao, H. Nutrient deficiency image diagnose of rapeseed based on color feature. *Chin. J. Oil Crop Sci.* 2015, 37, 576.