



## Article

# Winter Wheat SPAD Value Inversion Based on Multiple Pretreatment Methods

Lanzhi Shen <sup>1,2,3</sup>, Maofang Gao <sup>1,\*</sup> , Jingwen Yan <sup>3</sup>, Qizhi Wang <sup>1</sup> and Hua Shen <sup>4</sup>

<sup>1</sup> Institute of Agricultural Resources and Regional Planning, Chinese Academy of Agricultural Sciences/Key Laboratory of Agricultural Remote Sensing, Ministry of Agriculture and Rural Affairs, Beijing 100081, China

<sup>2</sup> 21st Century Space Technology Application Co., Ltd., Beijing 100096, China

<sup>3</sup> Key Laboratory of Digital Signal and Image Processing of Guangdong Province, Shantou University, Shantou 515063, China

<sup>4</sup> School of Emergency Management, Institute of Disaster Prevention, Sanhe 065201, China

\* Correspondence: gaomaofang@caas.cn

**Abstract:** SPAD value was measured by a portable chlorophyll instrument, which can reflect the relative chlorophyll content of vegetation well. Chlorophyll is an important organic chemical substance in plants that acquires and transmits energy during photosynthesis. The continuous spectral curve of winter wheat can be obtained rapidly in a specific band range by using hyperspectral remote sensing technology to estimate the SPAD value of winter wheat, which is of great significance to the growth monitoring and yield estimation research of winter wheat. In this study, with winter wheat as the research object, the spectral data and corresponding SPAD value in different growth stages were used as the data source, 20 kinds of data preprocessing spectra and sensitive spectral indices set the data as model input values, the partial least square regression (PLSR) model was established to estimate the SPAD value, and the model estimation results of different model input values at different growth stages were compared in detail. The results showed that the set of sensitive spectral indices selected in this study as input values can effectively improve the accuracy and stability of the PLSR model. In addition, the effects of 20 spectral data pretreatment methods on the estimation results of the SPAD value were compared and analyzed in different growth stages. It was found that the spectral data pretreated by the combination of wavelet packet denoising, first-order derivative transformation and principal component analysis can improve the accuracy and stability of PLSR model, and it is suitable for all growth stages. The results also showed that the estimation model is highly sensitive to the standard deviation of the SPAD value ( $STD_{chl}$ ) in sample sets. When the standard deviation is greater than 5.5 SPAD, the larger the  $STD_{chl}$  is, the higher the model estimation accuracy is, and the more stable the model is. At this time, the model estimation accuracy is higher ( $R^2_V$  is greater than 0.5, ratio of performance to deviation is greater than 1.4), which can meet the estimation requirements of the SPAD value.

**Keywords:** hyperspectral remote sensing; partial least squares regression; chlorophyll content; SPAD value



**Citation:** Shen, L.; Gao, M.; Yan, J.; Wang, Q.; Shen, H. Winter Wheat SPAD Value Inversion Based on Multiple Pretreatment Methods. *Remote Sens.* **2022**, *14*, 4660. <https://doi.org/10.3390/rs14184660>

Academic Editor: Yongguang Zhang

Received: 6 August 2022

Accepted: 14 September 2022

Published: 18 September 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Chlorophyll is the main pigment in plants that aids in photosynthesis and plays a great role in the acquisition and transmission of energy in plants [1,2]. As an important index to evaluate the healthy growth of vegetation, it is important to obtain the chlorophyll content quickly and nondestructively [3,4]. A portable chlorophyll instrument is used to measure a relative chlorophyll content, which has real-time, rapid, and non-destructive advantages [5].

At present, hyperspectral remote sensing technology has been widely studied and applied in agriculture, including spectral feature analysis of crop leaves or canopy, monitoring and inversion of some physiological and biochemical parameters in crops, crop

identification and classification, pest monitoring, etc. [6–8]. This technology can quickly obtain the continuous spectral curve of crops, and its small difference can reflect different characteristics such as crop growth. The visible range band (400–780 nm) of the spectrum is the strong absorption band of pigment [9]. The change of visible spectrum waveform is closely related to the content of chlorophyll in leaves. Jin et al. analyzed the spectrum of rice leaves and the corresponding chlorophyll content and found that the sensitive bands related to the chlorophyll content of rice leaves were 450–686 and 750–780 nm [10]. Lu et al. analyzed the correlation between the spectrum of cherry leaves and chlorophyll content after making various changes [11]. Using stepwise linear regression analysis, it was found that the variable  $\log(1/R_{741})$  was highly correlated with the chlorophyll of cherry leaves. Zhang et al. conducted a variety of typical changes in the original winter wheat spectrum and found that the spectral change data based on the second derivative  $R''$  can accurately estimate the chlorophyll content under low temperature stress [12]. In the range of red light, the research of red edge parameters is also the focus of many researchers. Bonham Carter et al. defined the red edge position as the wavelength value corresponding to the first-order differential maximum in the red band and then analyzed its relationship with leaf chlorophyll content in detail [13]. Yang Jie et al. compared the reflectance of any two bands in the spectrum of rice leaves from 350 to 2500 nm and calculated its normalization index [14]. They proposed that the band ratio  $SR(R_{724}, R_{709})$  and normalization index  $ND(R_{780}, R_{709})$  are the two spectral indices most closely related to the chlorophyll content of rice leaves, and the determination coefficients are above 0.9. To extract sensitive bands, Wang et al. employed the elastic net approach to minimize the dimensionality of hyperspectral data and then retrieved the chlorophyll content of winter wheat [15]. Qiu et al. applied the SPAD value and spectral techniques to study the change of nitrogen during rape growth [16].

Because the spectral data obtained by the instrument has a certain noise, most researchers will preprocess the spectral data in the estimation of chlorophyll content of winter wheat [10–14]. However, single or two or three kinds of preprocessed spectral data or some spectral indices are generally used as a model input to estimate winter wheat chlorophyll content, and different preprocessing methods are used in different studies [11,12,14]. Few studies compare the effects of these common preprocessing or spectral indices on the results of a winter wheat leaf green content estimation model. Therefore, it is necessary to select some common preprocessed spectral data and select some spectral indices with high correlation with chlorophyll content as input values, and compare and analyze the model accuracy under various input values. In addition, for the remote sensing monitoring of winter wheat chlorophyll content, most researchers perform research based on the sample data of a certain growth period [17], which lacks universality for all growth periods. Therefore, it is necessary to consider the whole growth period of winter wheat and find a chlorophyll estimation model suitable for the whole growth period.

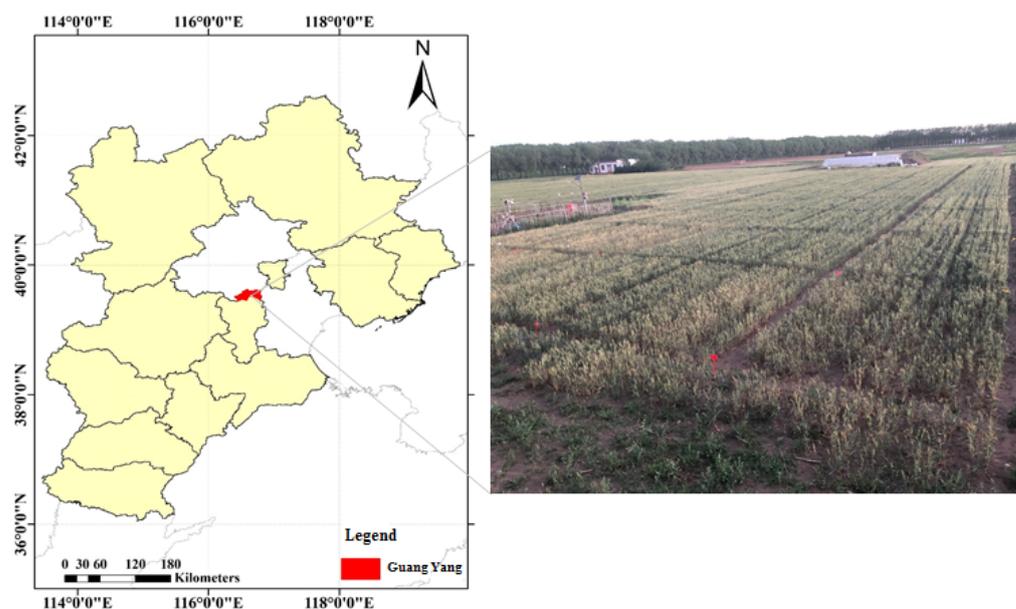
Therefore, this study will collect the spectral data of Winter Wheat in multiple growth stages (from jointing stage to maturity stage), take different combinations of preprocessed spectral data and spectral indices data [18] as input values, respectively, and model to estimate the SPAD value. In addition, PLSR [19,20] will be used to compare and analyze the estimation results of different models and explore the influencing factors of model accuracy and stability. It has certain application value for estimating the SPAD value in different growth stages and provides a theoretical basis and data preprocessing reference for estimating the chlorophyll content of vegetation from spaceborne hyperspectral satellite images.

## 2. Experiments and Methods

### 2.1. Study Area

The study area is located in the international agricultural high tech Industrial Park of Chinese Academy of Agricultural Sciences in Langfang City, Hebei Province. The geographical coordinates are  $116^{\circ}35'34''E$ – $116^{\circ}35'36''E$  and  $39^{\circ}35'50''N$ – $39^{\circ}35'53''N$  (Figure 1). The average

annual temperature is 13.9 °C, and the average annual precipitation is 393.4 mm, belonging to the warm temperate continental monsoon climate. The total area of the test area is 1200 m<sup>2</sup> (20 m wide and 60 m long). The soil texture is mainly loamy, and the soil types are mainly brown soil, moisture soil and cinnamon soil, with a unit weight of 1.5 g/cm<sup>3</sup>. The winter wheat variety selected in the experiment is Jingdong 12, and the fertilization level is consistent with the field fertilization level in this area (pure nitrogen (N) 256.5 kg·hm<sup>-2</sup>, phosphorus pentoxide (P<sub>2</sub>O<sub>5</sub>) 240 kg·hm<sup>-2</sup>, all input at one time during sowing).



**Figure 1.** Overview of the study area.

## 2.2. Spectral and SPAD Value Data Acquisition

The data used in the study mainly include the spectral data of winter wheat leaves at various growth stages and the corresponding SPAD value.

The spectral data of winter wheat leaves were measured by portable ground object spectrometer ASD FieldSpec4. The wavelength range of the spectrometer is 350–2500 nm, the spectral sampling interval is 1.4 nm (350–1000 nm), 2 nm (1001–2500 nm) and the spectral resolution is 3 nm (700 nm), 8 nm (1400, 2100 nm). The spectra of winter wheat leaves were measured from 10:30 to 13:30 Beijing time when the weather was sunny, cloudless or light cumulus, windless or wind speed lower than 1 m/s. Before measurement, the optical fiber probe is vertically aligned with the standard reference plate (whiteboard) for calibration. The winter wheat spectrum measurement is re-calibrated with the whiteboard every 15 min to prevent the influence caused by the drift of the sensor response system and the change of the solar incidence angle [21]. During the measurement, the measurer was dressed in dark color, facing the vertical direction of the sun (perpendicular to the main plane direction), holding a pistol, and the optical fiber probe was vertically pointing downward at the no shadow in the front leaf of the top leaf of winter wheat. The spectral measurement position of winter wheat was selected at the place with less stem veins and more mesophyll. The vertical distance between the optical fiber probe and the blade is 1 cm. The spectral curves of each leaf sample are measured repeatedly, and 5 spectral curves are collected, and the average value of the 5 spectral curves is taken as the final spectral reflectance of the sample. The spectrum of green vegetation is mainly affected by pigment content in the range of 400–900 nm, and it will show obvious peak and valley characteristics, which is a representative band range distinguished from non-green vegetation [22]. The study only makes correlation analysis on the spectral data in the range of 400–900 nm. The SPAD value was measured by portable chlorophyll meter SPAD-502 and replaced by a SPAD value. After the leaf spectrum is measured, a portable chlorophyll meter is used to evenly measure five values at the spectral measurement points

immediately and the average value is calculated as the SPAD value sample. Each growth stage was randomly sampled in the experimental area.

Spectral data and SPAD data were collected at the jointing stage, heading stage, flowering stage, grain filling stage, milk ripening stage and mature stage. The data of the whole growth period include the data of six single-growth periods. The set aside method is used to divide the training set and the verification set, in which 4/5 of the samples are used for training, and the remaining 1/5 are reserved for verification. In order to make the types of samples in the training set and the verification set as balanced as possible, the samples in each growth period are sorted from small to large according to the SPAD value, and then, one sample from every four is taken into the verification set according to the approximate ratio of 4:1 between the training set and the verification set, and the remaining samples are put into the training set. The data set size and SPAD value information of each growth period are shown in Table 1.

**Table 1.** The introduction of winter wheat sample sets on each growth stages.

| Growth Stages       | Data Collection Data | Sample Set       | Number of Samples | SPAD Value |         |         |                    |
|---------------------|----------------------|------------------|-------------------|------------|---------|---------|--------------------|
|                     |                      |                  |                   | Maximum    | Minimum | Average | Standard Deviation |
| Jointing stage      | 26 April 2019        | Training set     | 132               | 55.5       | 25.2    | 43.645  | 5.632              |
|                     |                      | Validation set   | 32                | 56.8       | 35.5    | 46.350  | 5.481              |
|                     |                      | Total sample set | 164               | 56.8       | 25.2    | 44.173  | 5.722              |
| Heading stage       | 9 May 2019           | Training set     | 156               | 67.2       | 45.9    | 54.341  | 3.511              |
|                     |                      | Validation set   | 44                | 68.9       | 49.7    | 55.823  | 3.652              |
|                     |                      | Total sample set | 200               | 68.9       | 45.9    | 54.667  | 3.604              |
| Flowering stage     | 13 May 2019          | Training set     | 78                | 60.7       | 46.4    | 53.304  | 2.652              |
|                     |                      | Validation set   | 20                | 65.1       | 51.7    | 56.585  | 3.196              |
|                     |                      | Total sample set | 98                | 65.1       | 46.4    | 53.973  | 3.087              |
| Filling stage       | 22 May 2019          | Training set     | 156               | 70.1       | 46.6    | 55.112  | 3.187              |
|                     |                      | Validation set   | 39                | 61.4       | 43.8    | 54.754  | 3.133              |
|                     |                      | Total sample set | 195               | 70.1       | 43.8    | 55.041  | 3.180              |
| Milk ripening stage | 28 May 2019          | Training set     | 156               | 65.3       | 27.0    | 50.796  | 6.558              |
|                     |                      | Validation set   | 44                | 62         | 17.3    | 49.584  | 8.125              |
|                     |                      | Total sample set | 200               | 65.3       | 17.3    | 50.529  | 6.969              |
| Maturity stage      | 31 May 2019          | Training set     | 156               | 63.6       | 3.2     | 43.013  | 12.580             |
|                     |                      | Validation set   | 39                | 58         | 2.5     | 42.236  | 13.099             |
|                     |                      | Total sample set | 195               | 63.6       | 2.5     | 42.857  | 12.722             |
| Whole growth stage  |                      | Training set     | 842               | 76.8       | 17.3    | 51.702  | 6.432              |
|                     |                      | Validation set   | 210               | 63.6       | 2.5     | 43.578  | 12.592             |
|                     |                      | Total sample set | 1052              | 76.8       | 2.5     | 50.079  | 8.694              |

### 2.3. Spectral Index Collection

Affected by pigments, plant leaves have strong absorption characteristics in visible light band, especially in red band, and strong reflection characteristics in near-infrared band, which is the physical basis of remote sensing monitoring of green plants [23]. Through the reflectivity value of each band of visible light and near infrared, different vegetation indices can be obtained by different combinations of derivative values. 62 spectral indices related to the inversion of plant chlorophyll content in the published literature were collected and sorted out (Table 2). By analyzing the correlation between these spectral indices and the SPAD value, the sensitive spectral indices set will be selected to predict the SPAD value of winter wheat.

**Table 2.** Spectral indices used to estimate the vegetation physiological and ecological parameters.

| Number | Spectral Indices        | Formula   | Reference | Number | Spectral Indices               | Formula  | Reference |
|--------|-------------------------|---|-----------|--------|--------------------------------|--|-----------|
| V1     | NDVI                    | $\frac{R_{800} - R_{670}}{R_{800} + R_{670}}$     | [24]      | V2     | CI                             | $\frac{R_{750}}{(R_{700} - R_{710}) - 1}$                | [25]      |
| V3     | GI                      | $\frac{R_{554}}{R_{677}}$                         | [26]      | V4     | SIPI (680)                     | $\frac{R_{800} - R_{455}}{R_{800} - R_{680}}$            | [27]      |
| V5     | VI(700)                 | $\frac{R_{700} - R_{670}}{R_{700} + R_{670}}$     | [28]      | V6     | RDVI                           | $\frac{(R_{800} - R_{670})}{\sqrt{(R_{800} + R_{670})}}$ | [29]      |
| V7     | SR                      | $\frac{R_{800}}{R_{680}}$                         | [30]      | V8     | SR2                            | $\frac{R_{700}}{R_{670}}$                                | [31]      |
| V9     | MNDVI8                  | $\frac{(R_{755} - R_{730})}{(R_{755} + R_{730})}$ | [32]      | V10    | DPI                            | $\frac{D_{688} * D_{710}}{D_{697}^2}$                    | [33]      |
| V11    | D2-C                    | $\frac{D_{705}}{D_{722}}$                         | [33]      | V12    | D1-CF                          | $\frac{D_{730}}{D_{706}}$                                | [33]      |
| V13    | NDVI2-LS,C <sub>a</sub> | $\frac{R_{750} - R_{705}}{R_{750} + R_{705}}$     | [34]      | V14    | Gitelson-LS,C <sub>Total</sub> | $\frac{1}{R_{700}}$                                      | [35]      |
| V15    | Datt4-LS,C              | $\frac{R_{754}}{R_{704}}$                         | [36]      | V16    | Datt1-LS,C <sub>b</sub>        | $\frac{R_{672}}{R_{550}}$                                | [37]      |
| V17    | SIPI                    | $\frac{R_{800} - R_{445}}{R_{800} - R_{680}}$     | [27]      | V18    | SR5                            | $\frac{R_{752}}{R_{690}}$                                | [38]      |
| V19    | SR3-LS,S                | $\frac{R_{440}}{R_{690}}$                         | [39]      | V20    | Vogelmann-LS,C                 | $\frac{R_{740}}{R_{720}}$                                | [40]      |
| V21    | Boochs1                 | $D_{703}$   | [41]      | V22    | Boochs2                        | $D_{720}$  | [41]      |
| V23    | SOFDR (625–795)         | $\sum_{625}^{795} D$                              | [42]      | V24    | CI                             | $\frac{R_{800} - R_{550}}{R_{800}}$                      | [43]      |
| V25    | MSR                     | $\frac{R_{800} - R_{445}}{R_{680} - R_{445}}$     | [44]      | V26    | REP                            | $Max(D_{680}, D_{760})$                                  | [45]      |
| V27    | BGI                     | $\frac{R_{450}}{R_{550}}$                         | [46]      | V28    | SIPI (705)                     | $\frac{R_{800} - R_{455}}{R_{800} + R_{705}}$            | [47]      |
| V29    | ZM                      | $\frac{R_{750}}{R_{710}}$                         | [46]      | V30    | Datt-LS,C                      | $\frac{R_{850} - R_{710}}{R_{850} - R_{680}}$            | [36]      |
| V31    | SR6                     | $\frac{R_{750}}{R_{550}}$                         | [38]      | V32    | MNDVI1                         | $\frac{R_{755} - R_{745}}{R_{755} + R_{745}}$            | [32]      |
| V33    | D <sub>690_red</sub>    | $D_{690}$   | [47]      | V34    | NDVI3                          | $\frac{R_{682} - R_{553}}{R_{682} + R_{553}}$            | [48]      |
| V35    | Maccioni-LS,C           | $\frac{R_{780} - R_{710}}{R_{780} - R_{680}}$     | [49]      | V36    | MTCI                           | $\frac{R_{754} - R_{709}}{R_{709} - R_{681}}$            | [50]      |
| V37    | NPCI                    | $\frac{R_{680} - R_{430}}{R_{680} + R_{430}}$     | [27]      | V38    | GNDVI-C <sub>a</sub>           | $\frac{R_{800} - R_{550}}{R_{800} + R_{550}}$            | [51]      |
| V39    | Datt3-LS,C              | $\frac{R_{850}}{R_{710}}$                         | [36]      | V40    | Datt5-LS,C <sub>a,total</sub>  | $\frac{R_{672}}{R_{550} * R_{708}}$                      | [36]      |
| V41    | Datt2-LS                | $\frac{R_{860}}{R_{550} * R_{708}}$               | [37]      | V42    | SR4                            | $\frac{R_{750}}{R_{700}}$                                | [38]      |
| V43    | SR1-LS,C <sub>a</sub>   | $\frac{R_{675}}{R_{700}}$                         | [52]      | V44    | SRPI                           | $\frac{R_{430}}{R_{680}}$                                | [27]      |

Table 2. Cont.

| Number | Spectral Indices | Formula                       | Reference   | Number | Spectral Indices | Formula   | Reference |
|--------|------------------|-------------------------------|---|--------|------------------|---|-----------|
| V45    | Vogelmann2-LS,C  | $\frac{R_{715}}{R_{705}}$     | [40]  | V46    | Vogelmann3-LS,C  | $\frac{R_{734} - R_{747}}{R_{715} + R_{726}}$       | [40]      |
| V47    | SOFDR (680–780)  | $\sum_{680}^{780} D$          | [45]  | V48    | Carter-I         | $\frac{R_{695}}{R_{760}}$                           | [53]      |
| V49    | DVI              | $R_{800} - R_{670}$           | [30]  | V50    | PSSRC            | $\frac{R_{800}}{R_{680}} + \frac{R_{800}}{R_{635}}$ | [54]      |
| V51    | Gitelson-RG      | $\frac{R_{800}}{R_{550}} - 1$ | [55]  | V52    | BI               | $R_{800} + R_{670} + \frac{R_{550}}{\sqrt{3}}$      | [43]      |
| V53    | TCARI            |                               | $3 * \left( (R_{700} - R_{670}) - 0.2 * (R_{700} - R_{550}) * \left( \frac{R_{700}}{R_{670}} \right) \right)$                               |        |                  |   | [56]      |
| V54    | MCARI1           |                               | $1.2 * [2.5 * (R_{800} - R_{670}) - 1.3 * (R_{800} - R_{550})]$   |        |                  |   | [57]      |
| V55    | MCARI            |                               | $\frac{[(R_{700} - R_{670}) - 0.2 * (R_{700} - R_{550})]}{\left( \frac{R_{700}}{R_{670}} \right)}$  |        |                  |   | [58]      |
| V56    | TVI              |                               | $0.5 * [120 * (R_{750} - R_{550}) - 200 * (R_{670} - R_{550})]$   |        |                  |   | [59]      |
| V57    | MCARI2           |                               | $\frac{1.5 * [2.5 * (R_{800} - R_{670})] - 1.3 * (R_{800} - R_{550})}{\sqrt{(2 * R_{800} + 1) - (6 * R_{800} - 5 * \sqrt{R_{670}}) - 0.5}}$ |        |                  |   | [57]      |
| V58    | EVI              |                               | $2.5 * \left( \frac{R_{800} - R_{670}}{R_{800} - (6 * R_{670}) - (7.5 * R_{475}) + 1} \right)$  |        |                  |   | [60]      |
| V59    | TCARI2-Wu        |                               | $3 * \left( (R_{750} - R_{705}) - 0.2 * (R_{750} - R_{550}) * \left( \frac{R_{750}}{R_{705}} \right) \right)$                               |        |                  |   | [61]      |
| V60    | DD               |                               | $(R_{749} - R_{720}) - (R_{701} - R_{672})$   |        |                  |   | [62]      |
| V61    | MND (705)        |                               | $\frac{(R_{750} - R_{705})}{(R_{750} + R_{705} - 2R_{445})}$  |        |                  |   | [44]      |
| V62    | MNDVI            |                               | $\frac{(R_{800} - R_{680})}{(R_{800} + R_{680} - 2R_{445})}$  |        |                  |   | [36]      |

#### 2.4. Spectral Preprocessing Methods

Denosing, data form transformation and dimension reduction single preprocessing were applied to the original winter wheat spectral data in turn (Table 3). Two different denosing approaches are no denosing (ND) and wavelet packet denosing (WPD) [63–65]. Five different mathematical forms of transformation are the original spectral reflectance data  $R$ , reciprocal  $1/R$ ,  $\log(R)$ , first derivative  $R'$  and reciprocal first derivative  $(1/R)'$  [66]. Two different dimensionality reduction approaches are no dimension reduction (NDR) and principle component analysis dimension reduction (PCADR) [67–70].

Table 3. Combination pretreatment.

| Preprocessing            | Name                          |
|--------------------------|-------------------------------|
| Denosing                 | ND, WPD                       |
| Data form transformation | $R, R', (1/R)', 1/R, \log(R)$ |
| dimension reduction      | NDR, PCADR                    |

Denosing, data form transformation, and dimension reduction single preprocessing are combined to form a total of 20 combined preprocessing methods. The model's input value is various combinations of preprocessed spectral data, and the model estimation is established by PLSR.

### 2.5. Partial Least Squares Regression

The link between vegetation spectrum and vegetation chlorophyll concentration can be studied using partial least squares regression (PLSR). This method can estimate the dependent variable through the linear combination of independent variables [71]. It is a multiple independent variables to multiple dependent variables linear regression modeling technique. When the sample size is small and there are multiple correlations between variables, the PLSR method has more advantages than the traditional regression model [71]. Due to the possible strong correlation between independent variables, it is easy to lead to overfitting of the problem when solving the linear regression problem of multiple independent variables to multiple dependent variables. However, the PLSR technique will identify several linear independent new variables to substitute the previous independent variables, allowing the diversity between independent variables to be maximized and the overfitting problem to be alleviated.

In this paper, the preprocessed spectral data and the sensitive vegetation indices with high correlation with winter wheat chlorophyll are used as the input values of the model, and the PLSR method is used to estimate and analyze the SPAD value of winter wheat.

### 2.6. Model Performance Evaluation Indices

The root mean square error of the training set ( $RMSE_T$ ), the root mean square error of the validation set ( $RMSE_V$ ), the determination coefficient of the training set  $R^2_T$ , the determination coefficient of the verification set  $R^2_V$ , and the ratio of performance to deviation (RPD) were all used to carefully consider the model's effectiveness in the research.

The determination coefficient  $R^2$  is the second power of the correlation coefficient  $R$ . The calculation formula of the correlation coefficient  $R$  is shown in Formula (1).

$$R_i = \frac{\text{cov}(x, y)}{\sqrt{D(x)}\sqrt{D(y)}} = \frac{\sum_{n=1}^N (x_{ni} - \bar{x}_i)(y_n - \bar{y})}{\sqrt{\sum_{n=1}^N (x_{ni} - \bar{x}_i)^2} \sqrt{\sum_{n=1}^N (y_n - \bar{y})^2}} \quad (1)$$

In Formula (1), the  $R_i$  is the correlation coefficient between the winter wheat leaf spectral reflectance in band  $i$  and the winter wheat leaf SPAD value,  $N$  is the total number of samples,  $x_{ni}$  is the spectral reflectance of winter wheat leaves in band  $i$  of the  $n$ th sample,  $\bar{x}_i$  is the average spectral reflectance of winter wheat leaves in band  $i$  of all samples,  $y_n$  is the SPAD value of winter wheat leaves in the  $n$ th sample, and  $\bar{y}$  is the average SPAD value of winter wheat leaves in all samples. The value of determination coefficient is equal to  $R^2$ .

The calculation formula of root mean square error ( $RMSE$ ) is shown in Formula (2).

$$RMSE = \sqrt{\frac{\sum_{n=1}^N (y_{nt} - y_{np})^2}{N}} \quad (2)$$

In Formula (2),  $RMSE$  is root mean square error,  $N$  is the total number of samples,  $y_{nt}$  is the measured value of winter wheat leaf SPAD value,  $y_{np}$  is the estimated value of winter wheat leaf SPAD value.

The calculation formula of RPD is shown in Formula (3).

$$RPD = SD/RMSE_V \quad (3)$$

In Formula (3), RPD is the ratio of performance to deviation, SD is the standard deviation of SPAD value of winter wheat leaves in the validation set.

Generally speaking, the higher the value of the  $R^2$ , the higher the correlation between the predicted value and the real value. The RMSE value should be as low as possible. The stronger the model's stability, the closer the  $RMSE_T$  and  $RMSE_V$  are. RPD can be categorized into three categories: The model is useful for accurately predicting the SPAD value of winter wheat when RPD is equal to or greater than 2.0. The model's reliability can be enhanced by fine-tuning when the RPD is below 2.0. The model is unstable when RPD is equal or falls below 1.4 [72].

### 3. Result

#### 3.1. Sensitive Spectral Indices Screening

By analyzing the correlation between these spectral indices and SPAD value, the sensitive spectral indices set was selected to predict the SPAD value of winter wheat.

Figure 2 shows the changes of 62 spectral indices and correlation coefficient of leaf SPAD value in jointing, heading, flowering, filling, milk maturity, maturity and whole growth stage of winter wheat. Each vertical line in the figure corresponds to a spectral index. The seven points on the vertical line represent six single growth periods and whole growth periods. The length of the vertical line is the maximum and minimum difference between spectral index and correlation coefficient of chlorophyll concentration in six single growth periods and whole growth periods. The association between each spectral index and SPAD value is more or less affected by different growth periods, as seen in the figure. The association between spectral index V11 and chlorophyll concentration at heading stage is  $-0.11594$ , but the association between spectral index V11 and leaf SPAD value at maturity stage is as high as  $-0.919$ , which fluctuates greatly. We found that the vertical line length corresponding to each spectral index ranges from 0.1 to 1.1. The sensitive spectral indices selected by comprehensive consideration can reflect the differences of different growth periods, and it is applicable to all growth periods. It is based on the spectral indices corresponding to the maximum value of positive correlation coefficient and the minimum value of negative correlation coefficient between each growth period and SPAD value, to screen the sensitive spectral indices.

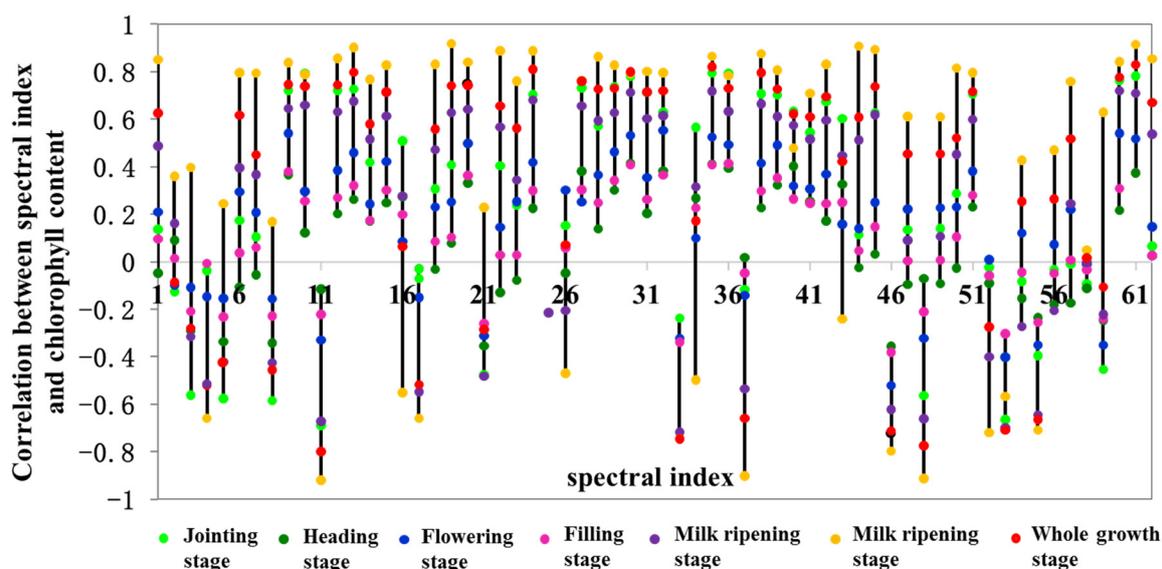


Figure 2. The association between spectral indices and SPAD value at different growth stages.

Table 4 lists the spectral indices and correlation coefficients corresponding to the maximum positive correlation coefficient and the minimum negative correlation coefficient of the SPAD value in different growth periods. The maximum positive and negative correla-

tion coefficients at jointing, milk maturity, maturity and whole growth stages are higher, above  $\pm 0.7$ , but the maximum positive and negative correlation coefficients at heading and filling are slightly smaller, basically below  $\pm 0.5$ . According to Table 3, 11 spectral indices V7, V8, V13, V24, V25, V26, V27, V35, V42, V49 and V50 are finally selected as sensitive spectral indices. The bands involved in these sensitive spectral indices are basically in the range of 670–755 nm, which are concentrated near the red edge, and some spectral indices also involve the 550 nm green peak band. This shows that the SPAD value has a great correlation with the red edge information and green peak information. The hyperspectral differences of winter wheat leaves at various growth stages are also mainly reflected in the vicinity around the green peak and red edge.

**Table 4.** Sensitive spectral indices.

| Growth Stages       | Highest Positive Correlation Spectral Indices | Highest Positive Correlation Coefficient | Lowest Negative Correlation Spectral Indices | Lowest Negative Correlation Coefficient |
|---------------------|---|--|--|---|
| Jointing stage      | V7  | 0.795                                    | V26  | −0.742                                  |
| Heading stage       | V24   | 0.415                                    | V42  | −0.400                                  |
| Flowering stage     | V25   | 0.553                                    | V35  | −0.521                                  |
| Filling stage       | V27   | 0.413                                    | V35  | −0.380                                  |
| Milk ripening stage | V49   | 0.720                                    | V26  | −0.717                                  |
| Maturity stage      | V13   | 0.919                                    | V8   | −0.919                                  |
| Whole growth stage  | V50   | 0.830                                    | V8   | −0.799                                  |

### 3.2. Estimation of SPAD Value Based on Sensitive Spectral Indices Set

The value range of each sensitive spectral index is quite different. In the first place, the normalized spectral indices are employed as the input value of the PLSR model to complete the subsequent SPAD value estimation. Table 5 shows the estimated findings of PLSR chlorophyll concentration for each growth period. Figures 3–9 demonstrate a scatter diagram of the estimation results of the SPAD value in various growth phases, in which figure (a) is the estimation results of training set and figure (b) is the estimation results of the validation set. The results showed that the estimation results of the heading stage, flowering stage and filling stage were poor, with RPD value less than 1.4 and  $R^2_V$  value less than 0.1. The estimation results of the jointing stage and milk ripening stage were better, RPD value was greater than 1.5 and  $R^2_V$  value was greater than 0.6. The estimation results of the mature stage and whole growth stage are the best, RPD value is basically above 2, and especially in the mature stage, the  $R^2_V$  value was up to 0.974.

**Table 5.** Estimation results of chlorophyll concentration in various growth stages with spectral indices set as input value.

| Growth Stages       | RMSE <sub>T</sub> (SPAD) | RMSE <sub>V</sub> (SPAD) | RPD   | R <sup>2</sup> <sub>T</sub> | R <sup>2</sup> <sub>V</sub> |
|---------------------|--------------------------|--------------------------|-------|-----------------------------|-----------------------------|
| Jointing stage      | 3.204                    | 3.473                    | 1.578 | 0.687                       | 0.706                       |
| Heading stage       | 2.906                    | 3.894                    | 0.938 | 0.319                       | 0.010                       |
| Flowering stage     | 2.215                    | 3.609                    | 0.885 | 0.310                       | 0.002                       |
| Filling stage       | 2.646                    | 3.214                    | 0.975 | 0.315                       | 0.063                       |
| Milk ripening stage | 4.422                    | 5.136                    | 1.582 | 0.552                       | 0.634                       |
| Maturity stage      | 4.517                    | 3.624                    | 3.615 | 0.882                       | 0.974                       |
| Whole growth stage  | 3.894                    | 6.399                    | 1.956 | 0.634                       | 0.792                       |

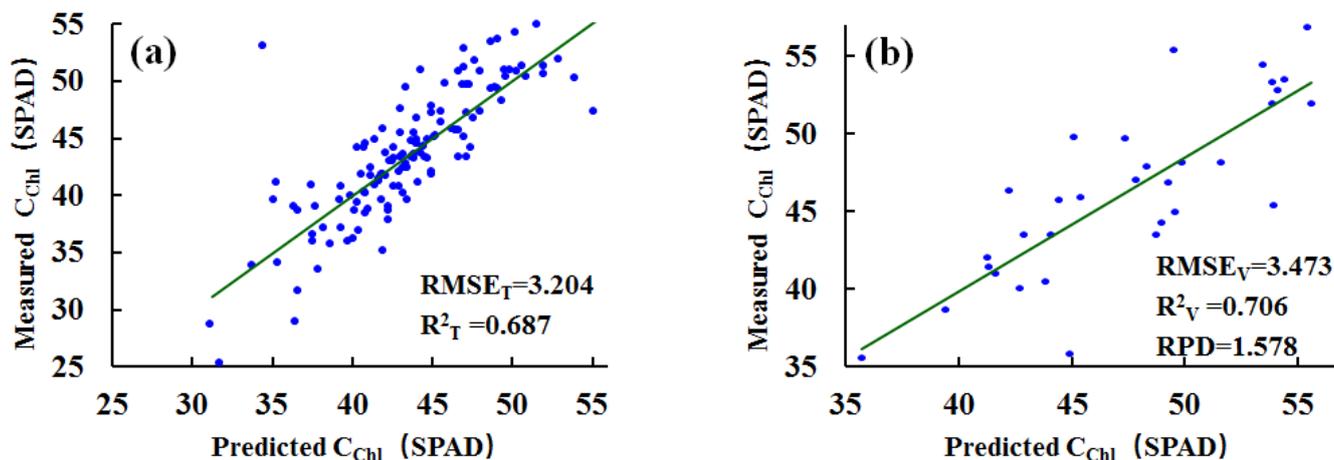


Figure 3. The estimated results of SPAD value for (a) training set and (b) validation set at jointing stage.

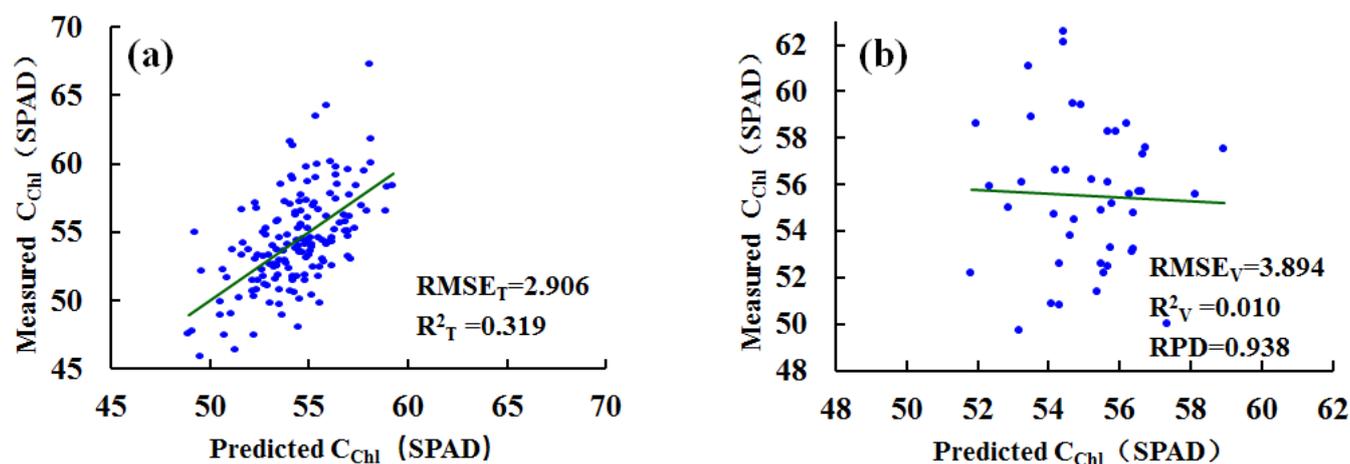


Figure 4. The estimated results of SPAD value for (a) training set and (b) validation set at heading stage.

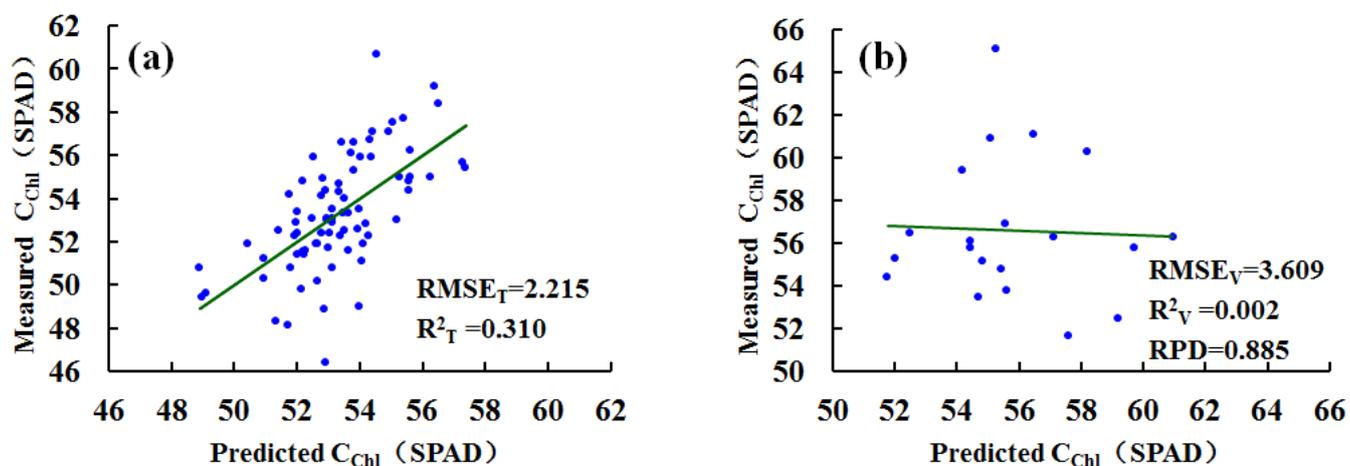


Figure 5. The estimated results of SPAD value for (a) training set and (b) validation set at flowering stage.

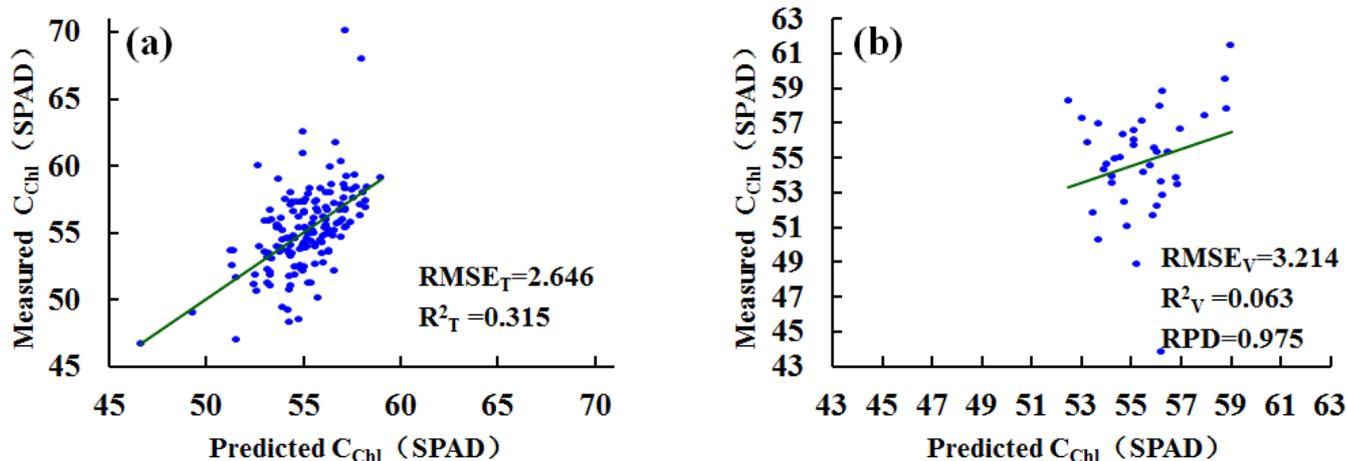


Figure 6. The estimated results of SPAD value for (a) training set and (b) validation set at filling stage.

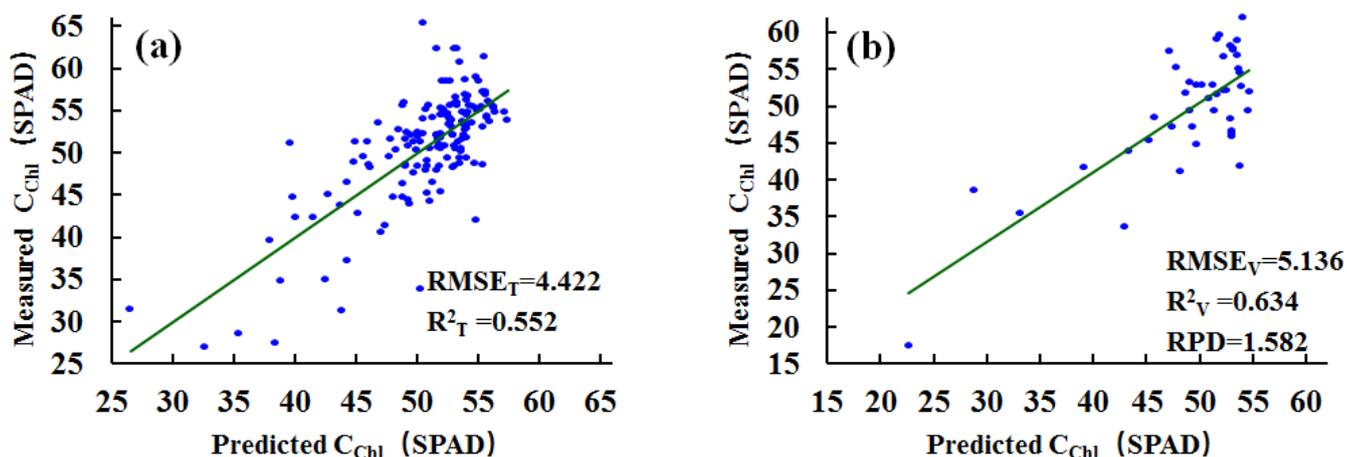


Figure 7. The estimated results of SPAD value for (a) training set and (b) validation set at milk stage.

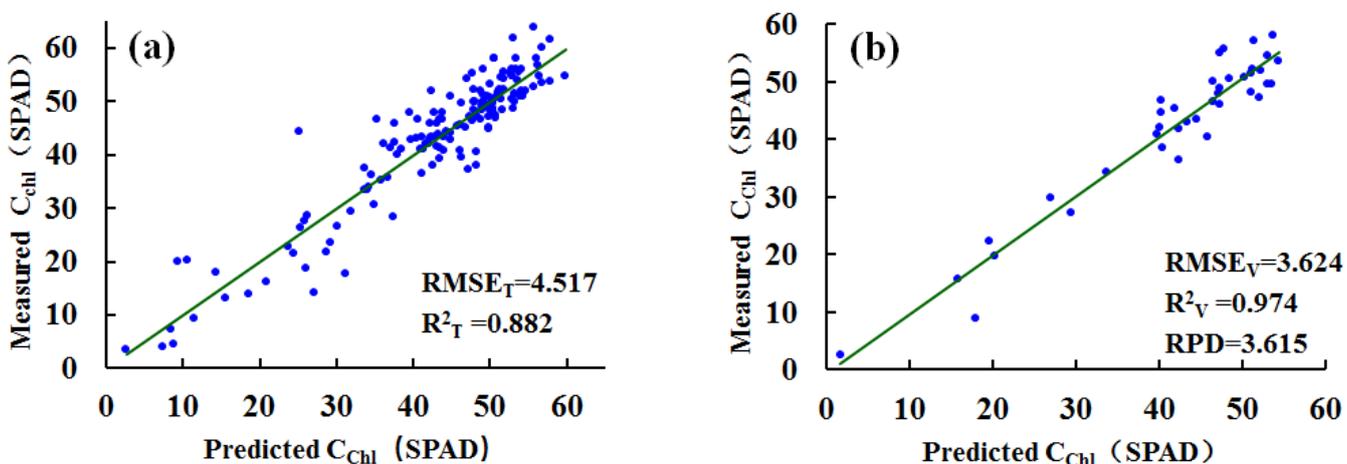
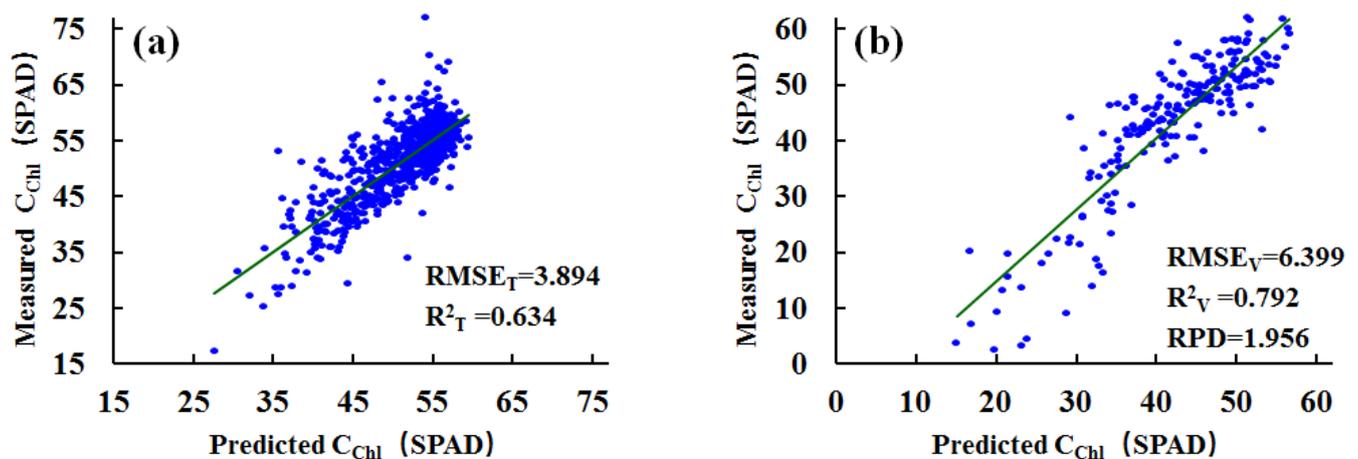


Figure 8. The estimated results of SPAD value for (a) training set and (b) validation set at mature stage.



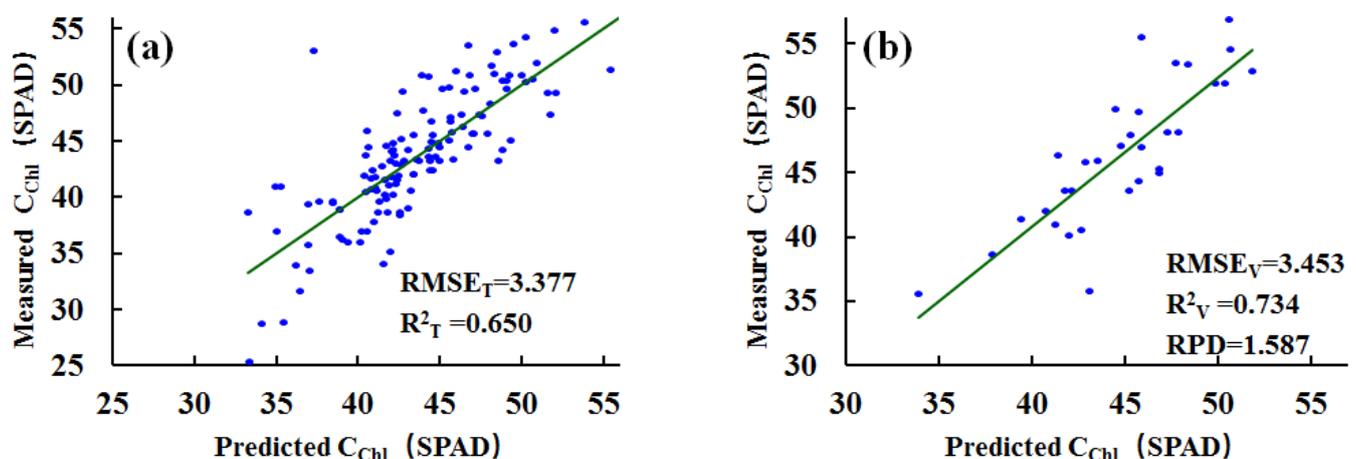
**Figure 9.** The estimated results of SPAD value for (a) training set and (b) validation set at whole growth stage.

### 3.3. Estimation of SPAD Value Based on Different Pretreatment Approaches

There are great differences between the estimated results of the preprocessed spectral data model in different growth stages. Table 6 lists the best model's estimation findings under 20 kinds of combined pretreatments in different growth periods. Figures 10–16 are scatter plots of the estimation results of the preprocessing optimal model in six single growth stages, as well as the entire growth stage.  $R^2_V$  growth value and RDP growth value in Table 6 refer to the growth values of  $R^2_V$  and RDP, respectively, compared with the model estimation results of the original spectral data.

**Table 6.** The optimal estimation results under the data preprocessing of each growth stages.

| Growth Stages | Pretreatment Method | RMSE <sub>T</sub> (SPAD) | RMSE <sub>V</sub> (SPAD) | RPD   | R <sup>2</sup> <sub>T</sub> | R <sup>2</sup> <sub>V</sub> | R <sup>2</sup> <sub>V</sub> Growth Value | RPD Growth Value |
|---------------|---------------------|--------------------------|--------------------------|-------|-----------------------------|-----------------------------|--|------------------|
| Jointing      | WPD-(1/R)'-PCADR    | 3.377                    | 3.453                    | 1.587 | 0.650                       | 0.734                       | 0.318                                    | 0.578            |
| Heading       | WPD-1/R-PCADR       | 2.616                    | 3.532                    | 1.034 | 0.450                       | 0.180                       | 0.178                                    | 0.405            |
| Flowering     | WPD-(1/R)'-PCADR    | 0.545                    | 3.948                    | 0.809 | 0.983                       | 0.092                       | 0.071                                    | 0.220            |
| Filling       | ND-(1/R)'-PCADR     | 2.777                    | 2.616                    | 1.198 | 0.244                       | 0.340                       | 0.330                                    | 0.225            |
| Milk ripening | WPD-(1/R)'-PCADR    | 4.881                    | 6.266                    | 1.297 | 0.452                       | 0.512                       | 0.005                                    | 0.101            |
| Maturity      | WPD-R'-PCADR        | 4.143                    | 3.734                    | 3.508 | 0.903                       | 0.970                       | 0.109                                    | 1.170            |
| Whole growth  | WPD-R'-PCADR        | 2.844                    | 6.107                    | 2.062 | 0.806                       | 0.785                       | 0.044                                    | 0.180            |



**Figure 10.** The estimated results of SPAD value for (a) training set and (b) validation set based on pretreatment WPD-(1/R)'-PCADR at jointing stage.

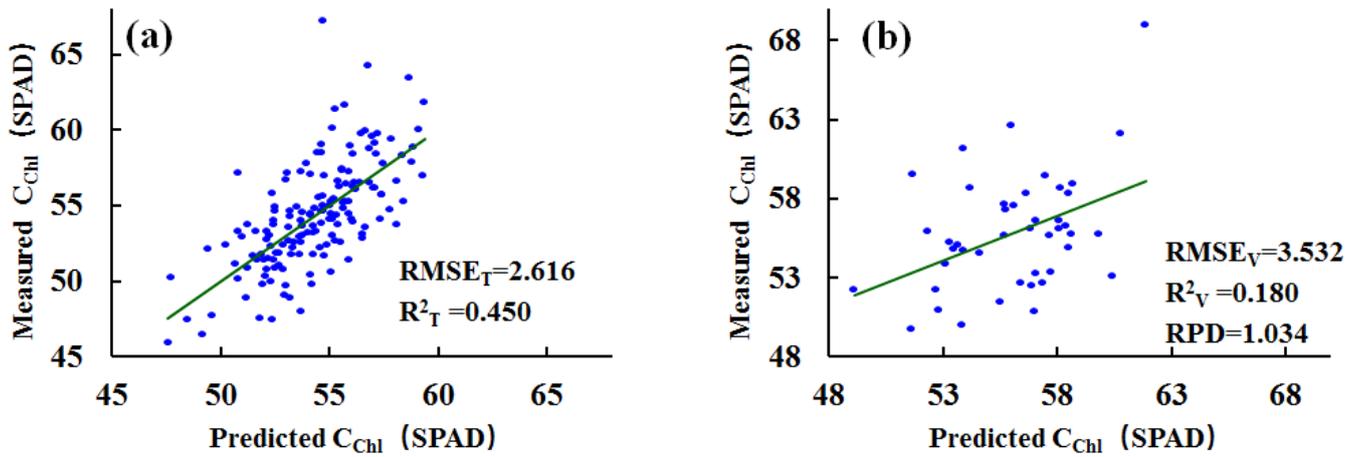


Figure 11. The estimated results of SPAD value for (a) training set and (b) validation set based on pretreatment WPD-1/R-PCADR at heading stage.

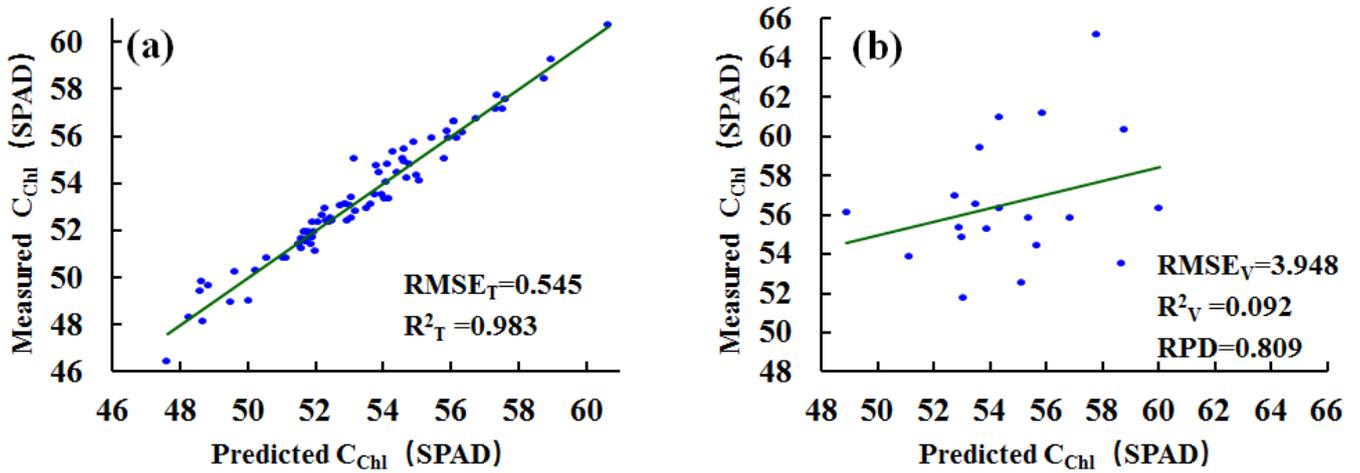


Figure 12. The estimated results of SPAD value for (a) training set and (b) validation set based on pretreatment WPD-(1/R)-PCADR at flowering stage.

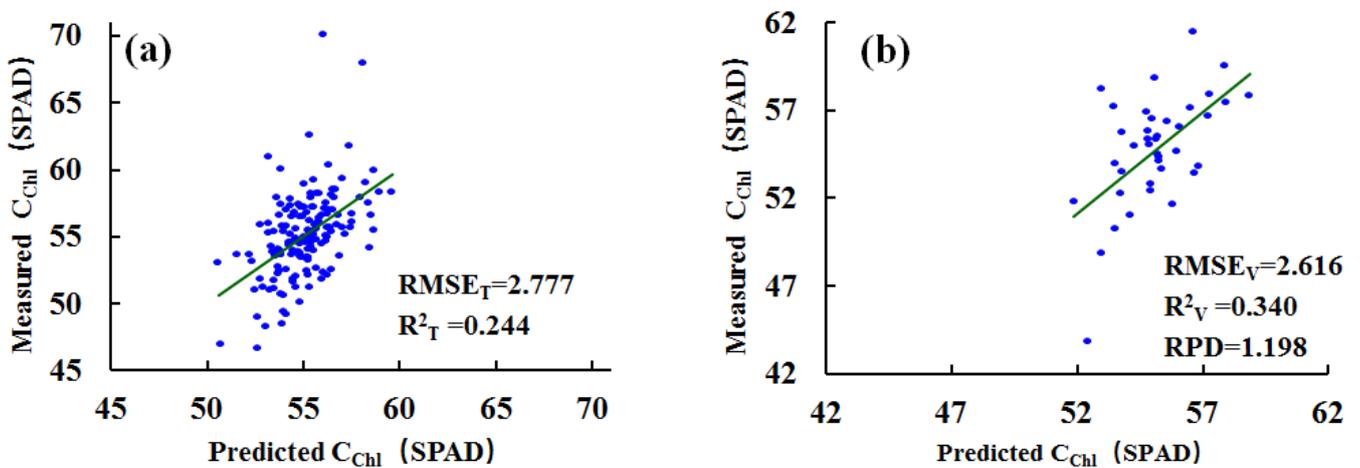


Figure 13. The estimated results of SPAD value for (a) training set and (b) validation set based on pretreatment ND-(1/R)-PCADR at filling stage.

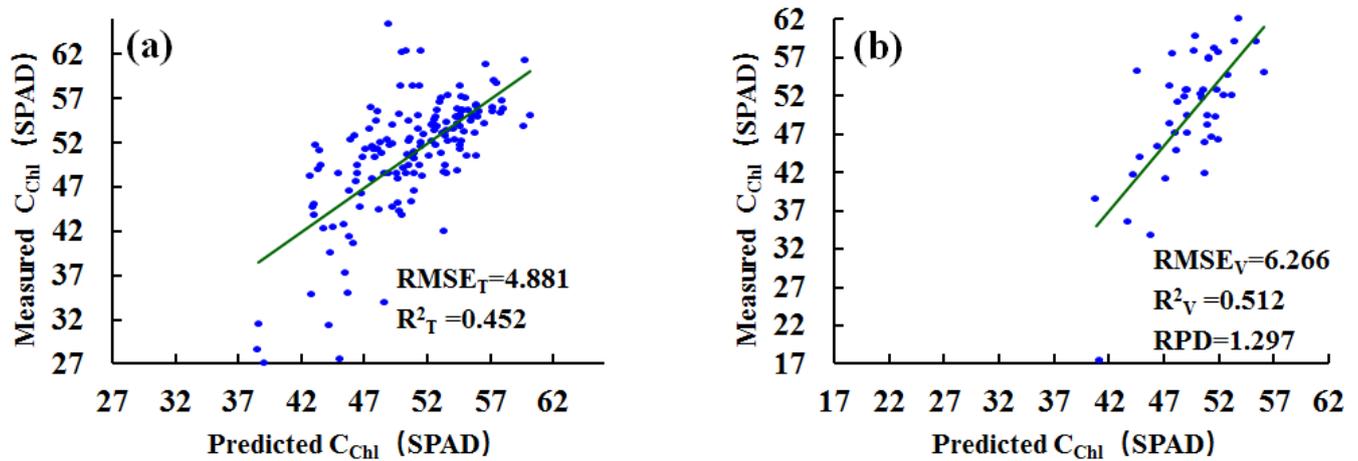


Figure 14. The estimated results of SPAD value for (a) training set and (b) validation set based on pretreatment WPD-(1/R)-PCADR at milk stage.

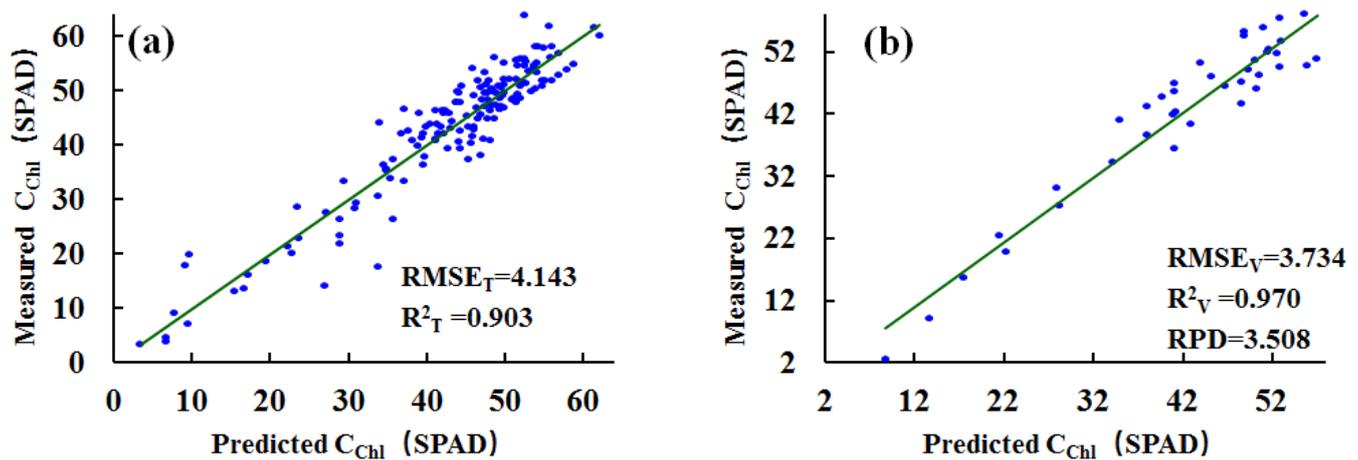


Figure 15. The estimated results of SPAD value for (a) training set and (b) validation set based on pretreatment WPD-R'-PCADR at mature stage.

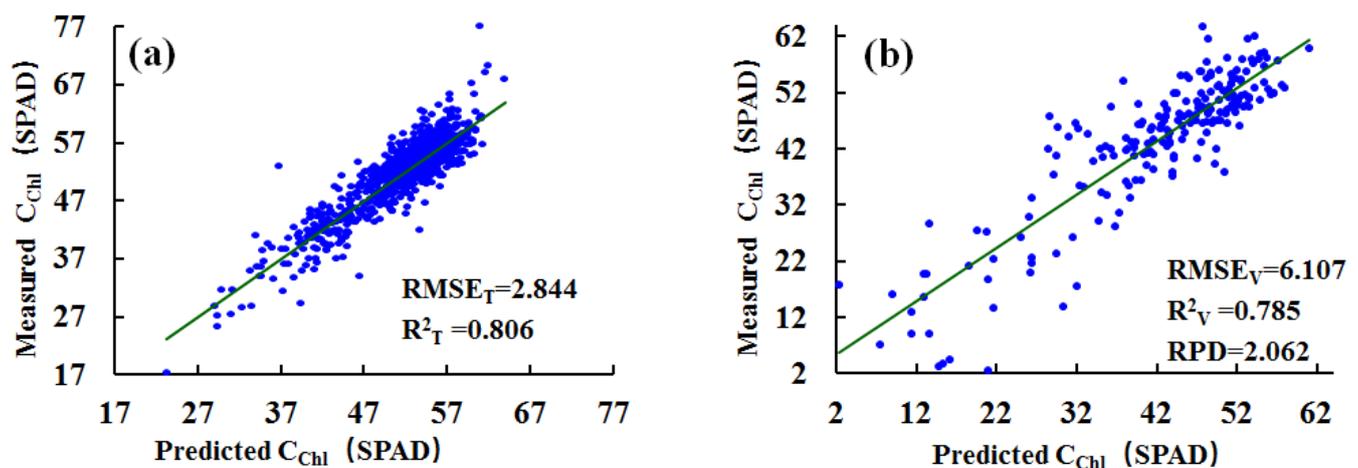


Figure 16. The estimated results of SPAD value for (a) training set and (b) validation set based on pretreatment WPD-R'-PCADR at whole growth stage.

Table 6 and Figures 10–16 demonstrate that the chlorophyll estimation results of the optimal model in different growth periods are basically obtained by preprocessing WPD-R'-PCADR or WPD-(1/R)-PCADR, and the  $R^2_V$  value can be up to 0.970, indicating that the combination of WPD denoising, PCA dimensionality reduction and first derivative mathematical form transformation can well increase the PLSR SPAD value estimate model's accuracy. This is because the first derivative transformation can highlight useful information, denoising can decrease the noise in the original spectral information, and dimensionality reduction can better select useful information and avoid overfitting in model training due to data redundancy. In addition, the chart also shows that the RPD value in the maturity stage and the whole growth stage is greater than 2, and the model  $R^2_V$  value is above 0.75; especially in the mature stage, the  $R^2_V$  value was up to 0.97, which can be a good estimate of the SPAD value of winter wheat. The RPD value of jointing stage is greater than 1.4, which can also better estimate the SPAD value of winter wheat. However, the estimation results of the SPAD value in the heading stage, flowering stage, filling stage and milk maturity stage were extremely poor,  $R^2_V$  was less than 0.4, and the corresponding RPD value was less than 1.4.

### 3.4. Comparison of SPAD Value Estimation Results of Different Model Inputs

Taking the sensitive spectral indices or the preprocessed spectral data as the model input value, the estimation results have some similarities, which are mainly reflected in different growth periods.

Firstly, the estimation accuracy of the mature stage and whole growth stage is very high, RPD value is greater than 2, and  $R^2_V$  can be up to 0.97 in the mature stage. The estimation accuracy of the jointing stage and milk maturity stage is relatively high, RPD value is basically greater than 1.4, and the accuracy  $R^2_V$  value is also greater than 0.5. The model's accuracy in predicting the stages of heading, flowering, and filling are relatively poor,  $R^2_V$  value is less than 0.4, and the estimation results are obviously underfitted (Tables 5 and 6). The estimation accuracy of different growth periods varies widely, and one of the influencing factors is the original sample set. The number of samples collected in every growth period is similar; however, the  $STD_{chl}$  in the sample set is substantially different (Table 1). From the turning green to jointing and then to the heading stage, the temperature increases continuously, the winter wheat leaves grow continuously, and the top leaf size also increases continuously. During this period, the chlorophyll content increases sharply. From filling to the mature stage, the grain in the wheat ear takes shape, its reproductive growth is dominant, and the vegetative organs such as leaves gradually stop growing and begin to wither and fall off. Coupled with the damage of high temperature to leaf cells, the chlorophyll content in the leaves decreases significantly. Therefore, the sample's SPAD value measured in these two periods has a large range, and the  $STD_{chl}$  in the sample set is also large. The standard deviation of the SPAD value in the sample set at the jointing stage, milk maturity stage, maturity stage and whole growth stage is greater than 5.5. From heading stage to filling stage, the number and size of leaves basically did not change, and the chlorophyll content increased slowly. Therefore, the range of SPAD value in the sample set measured during the growth period is small, and the  $STD_{chl}$  in the sample set is also small. The standard deviation of the SPAD value in the heading, flowering and filling stage is less than 4.

Secondly, using preprocessed spectral data or sensitive spectral indices set as the model input value can significantly boost the precision and stability of the PLSR-based SPAD value estimating model, and when the  $STD_{chl}$  of all samples in the model sample set is greater than 5.5 (Tables 1, 5 and 6), the estimation accuracy is higher (the estimation accuracy is greater than 0.5), and the model is stable (RPD is greater than 1.4). Moreover, when  $STD_{chl}$  in the sample set is greater than 5.5, the larger the SPAD value in the sample set, the more accurate the model's estimation. Most important of all, the  $STD_{chl}$  in the sample set and the estimation accuracy of the model from high to low are the mature stage, whole growth stage, jointing stage and milk mature stage. It can be seen that the

SPAD value estimation model based on PLSR is particularly responsive to the input sample set's  $STD_{chl}$  in addition to the high requirements for the spectral input value of the model (appropriate spectral data preprocessing or appropriate spectral indices set). In addition, there are great differences in the estimation results of single growth period data. If only the data of a single growth period are used to build a model to estimate the SPAD value, the universality of the estimation results is too weak. It is suggested to collect the data of mixed growth period (such as the whole growth period in this paper) for relevant estimation research, and the stability of the model is stronger.

The estimation results of input values of different models also have some differences. When the  $STD_{chl}$  in the input value sample set is greater than 6, taking the sensitive spectral indices as the input value is more accurate than taking the preprocessed spectral data. The calculation formula of spectral indices involves the mathematical form transformation of spectral data. Taking spectral indices as the model's input value also has a certain function of data compression and dimensionality reduction. Therefore, the spectral indices set selected in this study can significantly aid in improving the model's accuracy and reliability.

### 3.5. Estimation of SPAD Value with Different Chlorophyll Standard Deviations

For the sake of further confirming the estimation model's sensitivity to the  $STD_{chl}$  in the input sample set, we estimated the SPAD value in different standard deviation sample sets. Within this investigation, the SPAD value in all samples ranges from 2.5 to 76.8. Taking 5 as the gradient and making statistics in the range of 0 to 80 (Figure 17), it is found that the SPAD value of most samples is concentrated in the range of 40 to 60. Based on the principle that the sample set has as many standard deviations (STD) of different SPAD values as possible, the sample set with different SPAD value ranges is set. The range of SPAD value in the sample set ranges from 5 to 80. The spectral data of each sample set pretreated by WPD-(1/R)-PCADR are used as the input value of the model, and the model estimation accuracy of different SPAD value ranges was analyzed using the PLSR SPAD value estimation model (Table 7).

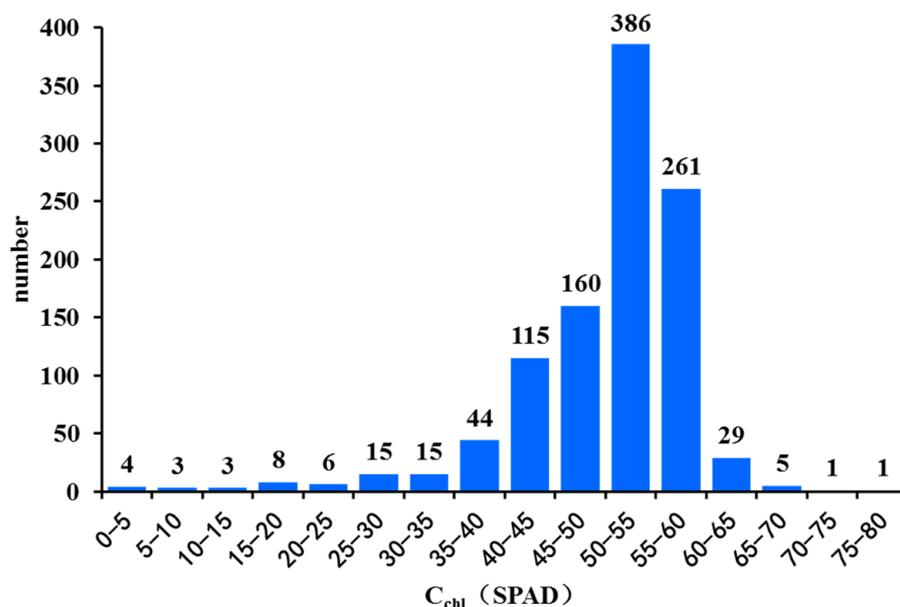
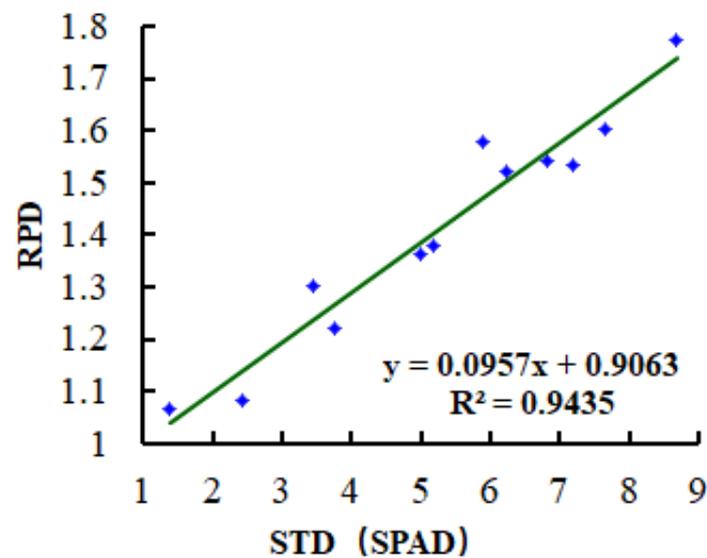


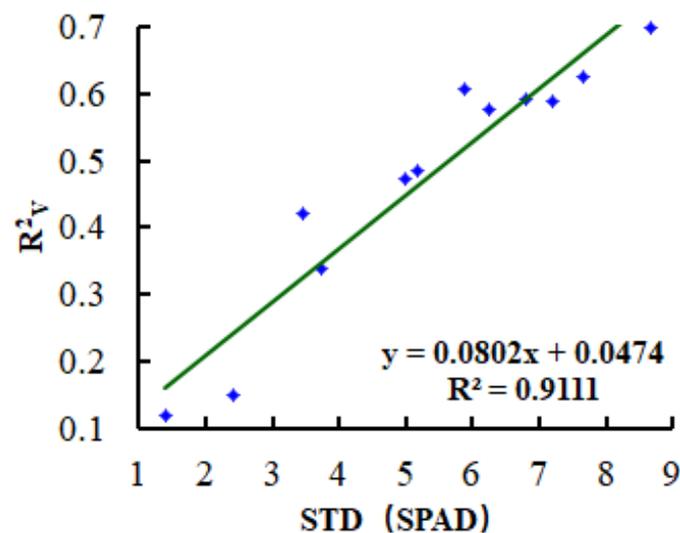
Figure 17. SPAD value distribution.

**Table 7.** The estimation results of different SPAD value ranges.

| SPAD Value Range | STD (SPAD) | RMSE <sub>T</sub> (SPAD) | RMSE <sub>V</sub> (SPAD) | RPD   | R <sup>2</sup> <sub>T</sub> | R <sup>2</sup> <sub>V</sub> |
|------------------|------------|--------------------------|--------------------------|-------|-----------------------------|-----------------------------|
| 50–55            | 1.393      | 1.299                    | 1.305                    | 1.065 | 0.133                       | 0.121                       |
| 50–60            | 2.424      | 2.159                    | 2.265                    | 1.083 | 0.203                       | 0.151                       |
| 45–60            | 3.462      | 2.641                    | 2.653                    | 1.301 | 0.420                       | 0.422                       |
| 45–65            | 3.748      | 3.018                    | 3.070                    | 1.223 | 0.352                       | 0.338                       |
| 40–65            | 4.981      | 3.271                    | 3.657                    | 1.362 | 0.57                        | 0.472                       |
| 40–80            | 5.173      | 3.644                    | 3.699                    | 1.378 | 0.509                       | 0.485                       |
| 35–80            | 5.878      | 3.664                    | 3.694                    | 1.578 | 0.614                       | 0.607                       |
| 30–80            | 6.242      | 4.096                    | 4.065                    | 1.523 | 0.573                       | 0.576                       |
| 25–80            | 6.803      | 4.469                    | 4.365                    | 1.544 | 0.572                       | 0.592                       |
| 15–55            | 7.197      | 4.500                    | 4.565                    | 1.534 | 0.616                       | 0.589                       |
| 15–80            | 7.664      | 3.689                    | 4.748                    | 1.603 | 0.771                       | 0.626                       |
| 00–80            | 8.680      | 4.161                    | 4.841                    | 1.774 | 0.773                       | 0.697                       |

The research shows that when the SPAD value range in the sample set is small (50~55, 50~60), the standard deviation of the sample set is also small, and the corresponding model R<sup>2</sup><sub>V</sub> value is also poor (R<sup>2</sup><sub>V</sub> is less than 0.2). As the range of SPAD values in the sample set gradually increases, the standard deviation also increases, and the corresponding model estimation accuracy (R<sup>2</sup><sub>V</sub>) basically shows a gradually increasing trend, and the more stable the model estimation result is (RPD). Figure 18 shows the variation law of RPD with the STD<sub>chl</sub> in the sample set, and Figure 19 shows the variation law of R<sup>2</sup><sub>V</sub> with the STD<sub>chl</sub> in the sample set. It can be seen from the two figures that the STD<sub>chl</sub> in the sample set is highly correlated with the PLSR SPAD value estimation model's accuracy and reliability. When the STD<sub>chl</sub> in the sample set is larger than 5.5, RPD is greater than 1.4 and R<sup>2</sup><sub>V</sub> is higher than 0.5, the model is stable, and its accuracy is high. At this time, the PLSR model could be used to predict the SPAD value with high accuracy.

**Figure 18.** The correlation diagram of model RPD and STD of sample SPAD values.



**Figure 19.** The correlation diagram of model  $R^2_v$  and standard deviation of sample SPAD values.

#### 4. Discussion

##### 4.1. Pretreatment Methods

In this study, common spectral preprocessing methods were compared, and it was found that using both sensitive spectral indices and preprocessed spectral data as model input values could increase the model's estimation accuracy and reliability to some extent. This also illustrates the importance of spectral data preprocessing. In particular, it can be seen from Table 6 that, no matter in which growth stages, the model inversion accuracy was highest when the pretreatment method was WPD-R'-PCADR or WPD-(1/R)-PCADR. When WPD and PCADR are used combined, the benefits of the derivative technique become clear. The derivative approach, according to Oldham et al. and Li et al., is not only a useful tool for studying spectral data, but it also aids in the resolution of multicollinearity issues [73,74]. As a result, it can be utilized to improve the analysis's sensitivity. It also aids in the reduction of noise to some extent. The first derivative (FD) provides the reflectance spectrum's slope, whereas the second derivative (SD) provides the reflectance spectrum's slope change [74]. Although SD allows more absorption peaks to be separated, it also introduces noise and may lead to errors. Through SD and FD changes, the spectrum is severely deformed and leads to sharp peaks. For the pretreatment of spectral data, the fractional derivative can reduce the range of spectral variation to a certain extent and retain the form properties of the original spectrum as far as possible, which is more useful than the integer derivative (FD, SD). Fractional derivatives can really be utilized to include further valuable data coming from remote sensing by extending the order to non-integers, which could also add more details to the spectrum than integer derivatives. Only FD was employed to measure SPAD values in this research. Other higher-order integer and fractional derivatives will be used to preprocess spectral data in the future.

##### 4.2. Influence of Sample Set on Model Accuracy

The results of this paper show that (Figures 18 and 19) the  $STD_{chl}$  in the model input sample set is highly sensitive to the accuracy of the PLSR model in evaluating the SPAD value. This has to do with the characteristics of the PLSR method itself. PLSR is a machine learning algorithm, which has high requirements for the model input sample set. It needs the number of samples to be as large and diverse as possible. When the sample set is too small or the data range is too monotonous, it is easy to lead to the phenomenon of underfitting and overfitting, and the generalization ability of the model is weak. In this study, the  $STD_{chl}$  in the sample set is too small, which is the performance of poor sample diversity. The SPAD values in the heading stage, flowering stage and milk ripening stage are relatively concentrated due to their own physiological characteristics. The standard

deviation of the SPAD value in the corresponding sample set is less than 4, and the accuracy of the corresponding model is also low. However, there are not enough data to prove whether the accuracy tends to be stable when the standard deviation is large to a certain value. More growth period data need to be collected for further verification.

Table 7 reveals that when the  $STD_{chl}$  in the sample set is greater than 5.5, the larger the  $STD_{chl}$  in the sample set, the higher the estimation accuracy of the model. Due to the small number of samples collected from the research, whether the threshold of 5.5 can be applied to other scenes needs to be further verified. In addition, based on the biological characteristics of winter wheat itself, the content of chlorophyll in leaves has a certain range. The  $STD_{chl}$  in the sample set we collected will always reach an upper limit. At this time, the relationship between the  $STD_{chl}$  and the accuracy of the model also needs to be further explored and verified. In following research, we will conduct deeper exploration, research and verification on the relationship between the sample set and model accuracy.

#### 4.3. Universality of Data

Taking winter wheat as the research object and taking the hyperspectral data and corresponding SPAD value in different growth periods as the data source, this paper establishes a model by using the PLSR method to estimate the SPAD value and analyzes the estimation results in detail. The research has achieved some results, but the scope of the study area is small. When the variety of winter wheat changes or the growing environment changes, the research results need to be further analyzed and verified. In addition, it is relatively difficult to obtain data at the leaf scale of winter wheat, because it is susceptible to weather (when there are clouds and winds, etc.), and data collectors need certain professional experience. At present, in the quantitative remote sensing's field, combining ground non-imaging spectral data with UAV imaging spectral data and satellite image multi-spectral data to realize multi-level and multi-scale crop growth monitoring is an unavoidable tendency. Wang et al. combined ground hyperspectral data with GF1 satellite multispectral images to achieve an inversion of chlorophyll concentration geographic variation in the winter wheat canopy [75]. In the subsequent study, we attempted to expand the research scope and conduct satellite-scale estimation of the winter wheat SPAD value, so as to analyze and find efficient and high-precision estimation methods for winter wheat chlorophyll that are suitable for more conditions.

### 5. Conclusions

Depending on the spectral data and SPAD value in various growth stages, taking the sensitive spectral indices and preprocessed spectral data as the model input values, the model was constructed by the PLSR method to estimate the SPAD value, and the estimated results were compared and analyzed in detail. The main conclusions are as follows:

- (1) The 11 spectral indices (V13, V15, V16, V17, V25, V39, V41, V42, V43, V47 and V57) selected in this paper can be used as the input values of the model, which could boost the model's precision and reliability while calculating the SPAD value.
- (2) Compared with the original spectral data and preprocessed spectral data as the model's input value, especially spectral data preprocessed by WPD-(1/R)-PCADR or WPD-R'-PCADR, it can increase the accuracy of the SPAD value estimation model and enhance the model's stability.
- (3) When the  $STD_{chl}$  in the sample set is less than 4, the estimation results of the SPAD value are prone to underfitting; when the  $STD_{chl}$  in the sample set is greater than 5.5, the greater the  $STD_{chl}$  in the sample set, the higher the model's estimating accuracy. In this case, the advantage of using sensitive spectral indices and preprocessing spectral dataset as model input values to increase the estimation model's precision and stability is obvious. In addition, when the  $STD_{chl}$  in the sample set was greater than 6, the model with the sensitive spectral indices as the model input had a greater estimation accuracy than the model with the pretreatment spectral dataset.

- (4) When using the spectral data and corresponding SPAD value data of a single growth period as the data source, the estimation results were not representative when using the PLSR method to predict the SPAD value, and the universality of data related to other growth periods was poor. Modeling using data from the whole growth period can improve the universality ability and stability of the model. It is recommended that spectral data for the whole fertility period be used as the data source and that the  $STD_{chl}$  corresponding to the data source should be as large as possible (standard deviation of SPAD values > 5.5).

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

**Funding:** This research was supported by Central Public-interest Scientific Institution Basal Research Fund (grant number Y2022GH05), Fundamental Research Funds for Central Non-profit Scientific Institution (grant number 1610132021024) and Shantou University Team Building Project for Innovative and Strong University (grant number 2017KCXTD015).

**Acknowledgments:** The data of this field experiment were obtained with the assistance of students Ya Gao, Qiang Li, Shilei Li, Renjun Li, Xiaolin Zhu and Jingjing Yang from the Institute of agricultural resources and regional planning, Chinese Academy of Agricultural Sciences. We would like to express our great thanks to them.

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

## References

- Guo, Y.K.; Xu, M.; Zhang, X.J.; Liu, Y.L. Hyperspectral inversion of chlorophyll content combined with PRO-4SAIL and BP neural network. *Bull. Surv. Mapp.* **2020**, *516*, 24–27.
- Luo, Q.F. Research progress of chlorophyll and development and utilization of chlorophyll copper Na. *For. Prod. Chem. Commun.* **1995**, *3*, 32–33.
- Xiao, J. *Research on Plant Chlorophyll Inversion Based on Geometric Spectral Integration*; Wuhan University: Wuhan, China, 2018.
- Qian, B.; Huang, W.; Huichun, Y. Inversion of winter wheat chlorophyll contents based on improved algorithms for red edge position. *Nongye Gongcheng Xuebao Trans. Chin. Soc. Agric. Eng.* **2021**, *36*, 162–170.
- Jiang, L.F.; Shi, F.C.; Wang, H.T.; Zu, Y.G. Application tryout of chlorophyll meter SPAD-502. *J. Ecol.* **2005**, *12*, 1543–1548.
- Tong, Q.X.; Zhang, B.; Zhang, L.F. Current progress of hyperspectral remote sensing in China. *J. Remote Sens.* **2016**, *20*, 689–707.
- Wang, Y.Y.; Chen, Y.H.; Li, J.; Huang, W.J. Two new red border parameters indicating the severity of winter wheat stripe rust. *J. Remote Sens.* **2007**, *6*, 875–881.
- Li, D.; Chen, J.M.; Zhang, X.; Yan, Y.; Zhu, J.; Zheng, H.B.; Zhou, K.; Yao, X.; Tian, Y.C.; Zhu, Y.; et al. Improved estimation of leaf chlorophyll content of row crops from canopy reflectance spectra through minimizing canopy structural effects and optimizing off-noon observation time. *Remote Sens. Environ.* **2020**, *248*, 111985. [[CrossRef](#)]
- Wang, B.Z.; You, Y.M.; Lu, M.Z.; Zhao, Z.Y. *Physical World Tour*; Beijing Institute of Technology Press: Beijing, China, 2015; Volume 12, p. 101.
- Jin, Z.Y.; Tian, Q.J.; Hui, F.M. Study on the Relationship between chlorophyll concentration and spectral reflectance in rice. *Remote Sens. Technol. Appl.* **2003**, *18*, 134–137.
- Lu, X.; Peng, H. Predicting Cherry Leaf Chlorophyll Concentrations Based on Foliar Reflectance Spectra Variables. *J. Indian Soc. Remote Sens.* **2015**, *43*, 109–120. [[CrossRef](#)]
- Zhang, X.R.; Feng, M.C.; Yang, W.D.; Wang, C.; Guo, X.L.; Shi, C.C. Using spectral transformation processes to estimate chlorophyll content of winter wheat under low temperature stress. *Chin. J. Eco-Agric.* **2017**, *25*, 1351–1359.
- Bonham-Carter, G.F. Numerical procedures and computer program for fitting an inverted gaussian model to vegetation reflectance data. *Comput. Geosci.* **1988**, *14*, 339–356. [[CrossRef](#)]
- Yang, J.; Tian, Y.C.; Yao, X. Hyperspectral estimation model of chlorophyll content in upper leaves of rice. *Acta Ecol. Sin.* **2009**, *29*, 6561–6571.
- Wang, T.L.; Gao, M.F.; Cao, C.L.; You, J.; Zhang, X.W.; Shen, L.Z. Winter wheat chlorophyll content retrieval based on machine learning using in situ hyperspectral data. *Comput. Electron. Agric.* **2022**, *193*, 106728. [[CrossRef](#)]
- Qiu, Z.J.; Song, H.Y.; He, Y.; Fang, H. Variation rules of the nitrogen content of the rape growth stage using SPAD and special. *Trans. CSAE* **2007**, *23*, 150–154.
- Jin, Q. *Extraction of Chlorophyll Hyperspectral Characteristics and Comprehensive Evaluation of Estimation Model in Winter Wheat Under Low Temperature Stress*; Shanxi Agricultural University: Jinzhong, China, 2017.

18. Srivastava, P.K.; Gupta, M.; Singh, U.; Prasad, R.; Pandey, P.C.; Raghubanshi, A.S.; Petropoulos, G.P. Sensitivity analysis of artificial neural network for chlorophyll prediction using hyperspectral data. *Environ. Dev. Sustain.* **2021**, *23*, 5504–5519. [[CrossRef](#)]
19. Shen, L.Z.; Gao, M.F.; Yan, J.W.; Li, Z.L.; Duan, S.B. Hyperspectral estimation of soil organic matter content using different spectral preprocessing techniques and PLSR method. *Remote Sens.* **2020**, *12*, 1206. [[CrossRef](#)]
20. Atzberger, C.; Gu'erif, M.; Baret, F.; Werner, W. Comparative analysis of three chemometric techniques for the spectroradiometric assessment of canopy chlorophyll content in winter wheat. *Comput. Electron. Agric.* **2010**, *73*, 165–173. [[CrossRef](#)]
21. Zha, Y.S.; Lu, Z.L.; Zhang, X. Design of Micro Solar Sensor for Cubic Star. *Chin. J. Sens. Actuators* **2018**, *31*, 31–35.
22. Wang, H. *Study on Growth Metabolism and Canopy Spectral Characteristics of MAIZE Under Different Fertilization Systems*; Shandong Agricultural University: Taian, China, 2015.
23. He, Y.; Feng, L. *Fundamentals of Geospatial Informatics*; Zhejiang University Press: Hangzhou, China, 2010.
24. Tucker, C.J. Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sens. Environ.* **1979**, *8*, 127–150. [[CrossRef](#)]
25. Gitelson, A.A.; Vina, A.; Ciganda, V. Remote estimation of canopy chlorophyll content in crops. *Geophys. Res. Lett.* **2005**, *32*, L08403. [[CrossRef](#)]
26. Zarco-Tejada, P.J.; Berjon, A.; Lopezlozano, R.; Miller, J.; Martin, P.; Cachorro, V.; Gonzalez, M.; Defrutos, A. Assessing vineyard condition with hyperspectral indices: Leaf and canopy reflectance simulation in a row-structured discontinuous canopy. *Remote Sens. Environ.* **2005**, *99*, 271–287. [[CrossRef](#)]
27. Peñuelas, J.; Baret, F.; Filella, I. Semi-empirical indices to assess carotenoids—Chlorophyll a ratio from leaf spectral reflectance. *Photosynthetica* **1995**, *31*, 221–230.
28. Gitelson, A.A.; Kaufman, Y.J.; Stark, R.; Rundquist, D. Novel algorithms for remote estimation of vegetation fraction. *Remote Sens. Environ.* **2002**, *80*, 76–87. [[CrossRef](#)]
29. Roujean, J.L.; Breon, F.M. Estimating PAR absorbed by vegetation from bidirectional reflectance measurements. *Remote Sens. Environ.* **1995**, *51*, 375–384. [[CrossRef](#)]
30. Jordan, C.F. Derivation of Leaf-Area Index from Quality of Light on the Forest Floor. *Ecology* **1969**, *50*, 663–666. [[CrossRef](#)]
31. McMurtrey, J.E.; Chappelle, E.W.; Kim, M.S.; Meisinger, J.J.; Corp, L.A. Distinguishing nitrogen fertilization levels in field corn (*Zea mays* L.) with actively induced fluorescence and passive reflectance measurements. *Remote Sens. Environ.* **1994**, *47*, 36–44. [[CrossRef](#)]
32. Mutanga, O.; Skidmore, A.K. Narrow band vegetation indices overcome the saturation problem in biomass estimation. *Int. J. Remote Sens.* **2004**, *25*, 3999–4014. [[CrossRef](#)]
33. Zarco-Tejada, P.J.; Pushnik, J.C.; Dobrowski, S.; Ustin, S.L. Steady-state chlorophyll a fluorescence detection from canopy derivative reflectance and double-peak red-edge effects. *Remote Sens. Environ.* **2003**, *84*, 283–294. [[CrossRef](#)]
34. Gitelson, A.; Merzlyak, M.N. Quantitative estimation of chlorophyll-a using reflectance spectra: Experiments with autumn chestnut and maple leaves. *J. Photochem. Photobiol. B Biol.* **1994**, *22*, 247–252. [[CrossRef](#)]
35. Gitelson, A.A.; Buschmann, C.; Lichtenthaler, H.K. The chlorophyll fluorescence ratio F735\_F700 as an accurate measure of the chlorophyll content in plants. *Remote Sens. Environ.* **1999**, *69*, 296–303. [[CrossRef](#)]
36. Datt, B. Visible/near infrared reflectance and chlorophyll content in Eucalyptus leaves. *Int. J. Remote Sens.* **2010**, *20*, 2741–2759. [[CrossRef](#)]
37. Datt, B. Remote sensing of chlorophyll a, chlorophyll b, chlorophyll a + b, and total carotenoid content in eucalyptus leaves. *Remote Sens. Environ.* **1998**, *66*, 111–121. [[CrossRef](#)]
38. Gitelson, A.A.; Merzlyak, M.N. Remote estimation of chlorophyll content in higher plant leaves. *Int. J. Remote Sens.* **1997**, *18*, 2691–2697. [[CrossRef](#)]
39. Lichtenthaler, H.K.; Lang, M.; Sowinska, M.; Heisel, F.; Miehé, J.A. Detection of Vegetation Stress Via a New High Resolution Fluorescence Imaging System. *J. Plant Physiol.* **1996**, *148*, 599–612. [[CrossRef](#)]
40. Vogelmann, J.E.; Rock, B.N. Moss, D.M. Red edge spectral measurements from sugar maple leaves. *Int. J. Remote Sens.* **1993**, *14*, 1563–1575. [[CrossRef](#)]
41. Boochs, F.; Kupfer, G.; Dockter, K.; KÜHbauch, W. Shape of the red edge as vitality indicator for plants. *Int. J. Remote Sens.* **1990**, *11*, 1741–1753. [[CrossRef](#)]
42. Elvidge, C.; Chen, Z. Comparison of broad-band and narrow-band red and near-infrared vegetation indices. *Remote Sens. Environ.* **1995**, *54*, 38–48. [[CrossRef](#)]
43. Liu, J.G.; Moore, J.M. Hue image RGB colour composition. A simple technique to suppress shadow and enhance spectral signature. *Int. J. Remote Sens.* **1990**, *11*, 1521–1530. [[CrossRef](#)]
44. Sims, D.A.; Gamon, J.A. Relationships between leaf pigment content and spectral reflectance across a wide range of species, leaf structures and developmental stages. *Remote Sens. Environ.* **2002**, *81*, 337–354. [[CrossRef](#)]
45. Filella, I.; Penuelas, J. The red edge position and shape as indicators of plant chlorophyll content, biomass and hydric status. *Int. J. Remote Sens.* **1994**, *15*, 1459–1470. [[CrossRef](#)]
46. Zarco-Tejada, P.J.; Miller, J.R.; Noland, T.L.; Mohammed, G.H.; Sampson, P.H. Scaling-Up and Model Inversion Methods with Narrowband Optical Indices for Chlorophyll Content Estimation in Closed Forest Canopies with Hyperspectral Data. *IEEE Trans. Geosci. Remote Sens.* **2001**, *39*, 1491–1507. [[CrossRef](#)]

47. Croft, H.; Chen, J.M.; Zhang, Y. The applicability of empirical vegetation indices for determining leaf chlorophyll content over different leaf and canopy structures. *Ecol. Complex.* **2014**, *17*, 119–130. [[CrossRef](#)]
48. Gandia, S.; Fernández, G.; García, J.C.; Moreno, J. Retrieval of vegetation biophysical variables from CHRISPROBA data in the SPARC campaign. In Proceedings of the Proba Workshop, ESA/ESRIN, Frascati, Italy, 28–30 January 2004.
49. Maccioni, A.; Agati, G.; Mazzinghi, P. New vegetation indices for remote measurement of chlorophylls based on leaf directional reflectance spectra. *J. Photochem. Photobiol. B Biol.* **2001**, *61*, 52–61. [[CrossRef](#)]
50. Dash, J.; Curran, P.J. The MERIS terrestrial chlorophyll index. *Int. J. Remote Sens.* **2004**, *25*, 5403–5413. [[CrossRef](#)]
51. Gitelson, A.A.; Kaufman, Y.J.; Merzlyak, M.N. Use of a green channel in remote sensing of global vegetation from EOS-MODIS. *Remote Sens. Environ.* **1996**, *58*, 289–298. [[CrossRef](#)]
52. Chappelle, E.W.; Kim, M.S.; Iii, M.M. Ratio analysis of reflectance spectra (RARS): An algorithm for the remote estimation of the concentrations of chlorophyll A, chlorophyll B, and carotenoids in soybean leaves. *Remote Sens. Environ.* **1992**, *39*, 239–247. [[CrossRef](#)]
53. Carter, G.A. Ratios of leaf reflectances in narrow wavebands as indicators of plant stress. *Int. J. Remote Sens.* **1994**, *15*, 697–703. [[CrossRef](#)]
54. Blackburn, G.A. Quantifying chlorophylls and carotenoids at leaf and canopy scales—An evaluation of some hyperspectral approaches. *Remote Sens. Environ.* **1998**, *66*, 273–285. [[CrossRef](#)]
55. Gitelson, A.A.; Gritz, Y.; Merzlyak, M.N. Relationships between leaf chlorophyll content and spectral reflectance and algorithms for non-destructive chlorophyll assessment in higher plant leaves. *J. Plant Physiol.* **2003**, *160*, 271–282. [[CrossRef](#)]
56. Haboudane, D.; Miller, J.R.; Tremblay, N.; Zarco-Tejada, P.J.; Dextraze, L. Integrated narrow-band vegetation indices for prediction of crop chlorophyll content for application to precision agriculture. *Remote Sens. Environ.* **2002**, *81*, 416–426. [[CrossRef](#)]
57. Haboudane, D.; Miller, J.R.; Pattey, E.; Tejada, P.J.; Strachan, I.B. Hyperspectral vegetation indices and novel algorithms for predicting green LAI of crop canopies: Modeling and validation in the context of precision agriculture. *Remote Sens. Environ.* **2004**, *90*, 337–352. [[CrossRef](#)]
58. Daughtry, C.S.T.; Walthall, C.L.; Kim, M.S.; Colstoun, E.B.; McMurtrey, J.E. Estimating Corn Leaf Chlorophyll Concentration from Leaf and Canopy Reflectance. *Remote Sens. Environ.* **2000**, *74*, 229–239. [[CrossRef](#)]
59. Broge, N.H.; Leblanc, E. Comparing prediction power and stability of broadband and hyperspectral vegetation indices for estimation of green leaf area index and canopy chlorophyll density. *Remote Sens. Environ.* **2001**, *76*, 156–172. [[CrossRef](#)]
60. Huete, A.R.; Liu, H.Q.; Batchily, K.; Leeuwen, W. A comparison of vegetation indices global set of TM images for EOS-MODIS. *Remote Sens. Environ.* **1997**, *59*, 440–451. [[CrossRef](#)]
61. Wu, C.; Niu, Z.; Tang, Q.; Huang, W. Estimating chlorophyll content from hyperspectral vegetation indices: Modeling and validation. *Agric. For. Meteorol.* **2008**, *148*, 1230–1241. [[CrossRef](#)]
62. Maire, G.L.; François, C.; Dufrêne, E. Towards universal broad leaf chlorophyll indices using PROSPECT simulated database and hyperspectral reflectance measurements. *Remote Sens. Environ.* **2004**, *89*, 1–28. [[CrossRef](#)]
63. Kong, L.J. *Matlab Wavelet Analysis Super Learning Manual*; The People's Posts and Telecommunications Press: Beijing, China, 2014; pp. 300–301.
64. Virmani, J.; Kumar, V.; Kalar, N. SVM-Based Characterization of Liver Ultrasound Images Using Wavelet Packet Texture Descriptors. *J. Digit. Imaging* **2012**, *26*, 530–543. [[CrossRef](#)]
65. Han, Z.Y.; Lun, S.X.; Wang, J. *Research on Robust Feature Extraction and Visualization of Speech Signals*; Northeastern University Press: Shenyang, China, 2012.
66. Shen, L.Z.; Gao, M.F.; Yan, J.W.; Yao, Y.M. Estimation model of soil organic matter based on SVR and PLSR. *Agric. Inf. China* **2019**, *31*, 58–71.
67. Pham, Q.T.; Liou, N.S. The development of on-line surface defect detection system for jujubes based on hyperspectral images. *Comput. Electron. Agric.* **2022**, *194*, 106743. [[CrossRef](#)]
68. Cao, C.L.; Wang, T.L.; Gao, M.F.; Li, Y.; Li, D.D.; Zhang, H.J. Hyperspectral inversion of nitrogen content in maize leaves based on different dimensionality reduction algorithms. *Comput. Electron. Agric.* **2021**, *190*, 106461. [[CrossRef](#)]
69. Juan, P.R.; Jochem, V.; Jordi, M. Hyperspectral dimensionality reduction for biophysical variable statistical retrieval. *ISPRS J. Photogramm. Remote Sens.* **2017**, *132*, 88–101.
70. Pearson, K. On lines and planes of closest fit to systems of points in space. *Phil. Mag.* **1901**, *2*, 559–572. [[CrossRef](#)]
71. Shi, L.; Li, Z.L.; Cui, G.M. A Model of Silicon Content in Hot Metal of Blast Furnace Based on Partial Least Square Regression. *J. Inn. Mong. Univ. Nat. Sci. Ed.* **2010**, *41*, 427–430.
72. Chang, C.W.; Laird, D.A.; Mausbach, M.J.; Hurburgh, C.R. Near infrared reflectance spectroscopy-Principal components regression analysis of soil properties. *Soil Sci. Soc. Am. J.* **2001**, *65*, 480–490. [[CrossRef](#)]
73. Oldham, K.B.; Spanier, J. The fractional calculus. *Math. Gaz.* **1974**, *56*, 396–400.
74. Li, B.; Xie, W. Adaptive fractional differential approach and its application to medical image enhancement. *Comput. Electr. Eng.* **2015**, *45*, 324–335. [[CrossRef](#)]
75. Wang, T.T. *Remote Sensing Inversion of Winter Wheat Chlorophyll Based on Hyperspectral and GF-1 Satellite Images*; Northwest A&F University: Shaanxi, China, 2019.