



Potential of Multivariate Statistical Technique Based on the Effective Spectra Bands to Estimate the Plant Water Content of Wheat Under Different Irrigation Regimes

Hui Sun^{1,2}, Meichen Feng¹, Lujie Xiao¹, Wude Yang^{1*}, Guangwei Ding^{3*}, Chao Wang¹, Xueqin Jia¹, Gaihong Wu¹ and Song Zhang¹

OPEN ACCESS

Edited by:

Bangyou Zheng,
Commonwealth Scientific
and Industrial Research Organisation
(CSIRO), Australia

Reviewed by:

Salah Elsayed Mohamed Elsayed,
University of Sadat City, Egypt
Bappa Das,
Central Coastal Agricultural Research
Institute (ICAR), India

*Correspondence:

Wude Yang
sxauywd@126.com
Guangwei Ding
guangwei.ding@northern.edu

Specialty section:

This article was submitted to
Technical Advances in Plant Science,
a section of the journal
Frontiers in Plant Science

Received: 20 November 2020

Accepted: 19 January 2021

Published: 26 February 2021

Citation:

Sun H, Feng M, Xiao L, Yang W,
Ding G, Wang C, Jia X, Wu G and
Zhang S (2021) Potential
of Multivariate Statistical Technique
Based on the Effective Spectra Bands
to Estimate the Plant Water Content
of Wheat Under Different Irrigation
Regimes.
Front. Plant Sci. 12:631573.
doi: 10.3389/fpls.2021.631573

¹ Institute of Dry Farming Engineering, Shanxi Agricultural University, Taigu, China, ² College of Resource and Environment, Shanxi Agricultural University, Taigu, China, ³ Department of Chemistry, Northern State University, Aberdeen, SD, United States

Real-time, nondestructive, and accurate estimation of plant water status is important to the precision irrigation of winter wheat. The objective of this study was to develop a method to estimate plant water content (PWC) by using canopy spectral proximal sensing data. Two experiments under different water stresses were conducted in 2014–2015 and 2015–2016. The PWC and canopy reflectance of winter wheat were collected at different growth stages (the jointing, booting, heading, flowering, and filling stages in 2015 and the jointing, booting, flowering, and filling stages in 2016). The performance of different spectral transformation approaches was further compared. Based on the optimal pretreatment, partial least squares regression (PLSR) and four combination methods [i.e., PLSR-stepwise regression (SR), PLSR-successive projections algorithm (SPA), PLSR-random frog (RF), and PLSR-uninformative variables elimination (UVE)] were used to extract the sensitive bands of PWC. The results showed that all transformed spectra were closely correlated to PWC. The PLSR models based on the first derivative transformation method exhibited the best performance (coefficient of determination in calibration, $R^2_C = 0.96$; root mean square error in calibration, $RMSE_C = 20.49\%$; ratio of performance to interquartile distance in calibration, $RPIQ_C = 9.19$; and coefficient of determination in validation, $R^2_V = 0.86$; root mean square error in validation, $RMSE_V = 46.27\%$; ratio of performance to interquartile distance in validation, $RPIQ_V = 4.34$). Among the combination models, the PLSR model established with the sensitive bands from PLSR-RF demonstrated a good performance for calibration and validation ($R^2_C = 0.99$, $RMSE_C = 11.53\%$, and $RPIQ_C = 16.34$; and $R^2_V = 0.84$, $RMSE_V = 44.40\%$, and $RPIQ_V = 4.52$, respectively). This study provides a theoretical basis and a reference for estimating PWC of winter wheat by using canopy spectral proximal sensing data.

Keywords: band selection, canopy reflectance, PLSR, transformation method, water content

INTRODUCTION

Climate change has increased the frequency and intensity of drought events. In most part of Asia, drought has been recorded to intensify during the last decades (Miyan, 2015), which raises numerous challenges in agriculture. Water is one of the important factors in plant growth and yield formation. With the increase of water deficit, crops, such as wheat, appear to be gradually wilted, and the rate of growth slows down, resulting in 20–40% production loss in cereal (Daryanto et al., 2017). Thus, accurate and real-time estimation of plant water status could help the planters to adjust the irrigation management effectively and efficiently.

The conventional way to measure plant water content is based on destructive field sampling and laboratory analysis, which is always time consuming and labor demanding. In contrast, spectral reflectance is considered as a rapid and nondestructive technique taken into practical application (Cho et al., 2007). Plant water condition directly influences cell turgor and internal space of tissue resulting in the changes of leaf structure. It causes the absorption, transmission, and reflection of light in the leaves changing the value of canopy reflection eventually. The near-infrared and short-infrared spectral regions are sensitive to the water content of plant leaf and canopy (Ihuoma and Madramootoo, 2017). Based on the theory, researchers proposed many vegetation indices to monitor water content, such as land surface water index (LSWI) (Xiao et al., 2005), simple ratio water index (SRWI) (Zarco-Tejada et al., 2003), normalized difference water index (NDWI) (Gao, 1996), and so on. Furthermore, plant water content affects chlorophyll content and canopy size. For this reason, previous findings indicated that bands in the visible region can be used indirectly to assess plant water status (Das et al., 2017; El-Hendawy et al., 2019a).

Hyperspectral remote sensing, having numerous continuous narrow bands and providing crucial information, has shown great potential for the accurate retrieval of plant parameters (Clevers et al., 2010). As for plant water status, there were approaches based on narrow band vegetation indices (Winterhalter et al., 2011; Wang et al., 2015; Zhang and Zhou, 2015; Fang et al., 2017) and spectral absorption features (Tian et al., 2001; Mutanga et al., 2005; Cheng et al., 2011). However, the canopy spectrum can generate comprehensive information which not only extracts the expression of the target variation but also other factors (e.g., soil background, noise of instrument) (Demetriades-Shah et al., 1990). In order to reduce useless information and improve the signal-to-noise ratio, the reflectance data are properly processed or transformed before the actual analysis. Numerous studies have indicated that different pretreatment methods are used to predict the water status of plant, including the first derivative (Liang et al., 2013), continuum removal (CR)

(González-Fernández et al., 2015), and normalization (Sun et al., 2015). Moreover, the hyperspectral data contains a large number of bands (e.g., FieldSpec 3.0 Spectrometer has more than 1,900 bands), resulting in redundancy and multicollinearity. In the modeling process, it also makes the number of samples much smaller than the number of independent variables (spectral wavelengths) used in the spectral analysis (Atzberger et al., 2010). Partial least square regression (PLSR) is an effective method to solve these problems. PLSR, which combines principal component analysis and multiple linear regression, can easily process the data matrix and solve the correlation between independent variables. It has been used to estimate vegetation water content (Li et al., 2008; Mirzaie et al., 2014). However, the PLSR model with full spectrum is complicated in practical applications. In order to remove the irrelevant information, reduce the number of input variables, simplify the complexity of the model, and improve the interpretability of the model, it is necessary to select the sensitive wavelengths. There are already many variable selection methods. Zou et al. (2010) made a review about the variable selection methods in near-infrared spectroscopy. Some variable selection methods were used in plant water status, such as Bipls-SPA (backward interval PLS in combination with successive projection algorithm) (Zhang et al., 2012), PLSR-SR (Das et al., 2017), random frog (Chen and Li, 2020), and so on. These models established with sensitive bands also had good prediction ability.

As an important step in quantitative spectral analysis, data preprocessing had a significant influence on improving PLSR model performance. PLSR is proved to be a useful approach in selecting sensitive wavelengths (Sharabian et al., 2014; Wang et al., 2017; Zeng et al., 2018) and quantitating plant parameters. Considering the advantages of pretreatment, more sensitive bands would be selected from the spectra with optimal pretreatment. To date, few studies compared multivariate statistical methods in sensitive band selection based on the optimal pretreatment method. In this study, nine common transformation methods in combination with PLSR were initiated and compared to select the suitable transformation method of canopy reflectance from field spectrometer data for estimating plant water content in winter wheat. Meanwhile, in order to reduce the input variables of the PLSR model, sensitive wavelengths were selected by PLSR and four combined methods [i.e., PLSR in combination with stepwise regression (SR), successive projections algorithm (SPA), random frog (RF), and uninformative variables elimination (UVE)].

MATERIALS AND METHODS

Site Description and Experimental Design

The experiment was conducted from 2014 to 2016 at the experiment station of Shanxi Agricultural University (E112°34'19.96", N37°25'19.81"), Shanxi Province (P. R. China). The experimental site has a temperate continental climate condition. The average annual temperature is 9.8°C, the annual

Abbreviation: C, calibration; CV, cross-validation; FC, field capacity; LV, latent variable; MSC, multiplicative scatter correction; PLSR, partial least squares regression; PWC, plant water content; R^2 , coefficient of determination; RF, random frog; RMSE, root mean square error; RPIQ, ratio of performance to interquartile distance; SD, standard deviation; SPA, successive projections algorithm; SR, stepwise regression; UVE, uninformative variables elimination; V, validation; VIP, variable importance in projection.

frost-free season is around 175 days, and the annual precipitation is around 450 mm.

Experiments were conducted in a water-experiment pool. The refilled soil in the pool is classified as Calcareous Cinnamon soil (Alfisols in US taxonomy) with 9.60 g/kg organic matter, 57.75 g/kg available nitrogen, 22.10 mg/kg available phosphate, and 185.48 mg/kg available potassium. The density of the soil is 1.36 g/cm³. The field capacity (FC) is 24.14%.

Jinnong 190 and Chang 4738 were planted in 2014–2015, while Chang 4738 and Zhongmai 175 were studied in 2015–2016. The experiments were set up in a randomized complete block design with three replications. There were five upper limits of irrigation for each 10-day period included in 2014–2015: W₁ (80% of the FC), W₂ (60% of the FC), W₃ (45% of the FC), W₄ (35% of the FC), and W₅ (less than 30% of the FC). There were five irrigation regimes included in 2015–2016: I₁ (four irrigations at jointing, booting, flowering, and filling stage), I₂ (three irrigations at jointing, booting, and filling stage), I₃ (two irrigations at jointing and flowering stage), I₄ (two irrigations at jointing and filling stage), and I₅ (without irrigation). The upper limit in 2015–2016 was 80% of the FC. The volume of water was controlled by the water meter. The plot area was 6 m². The critical growth stages of wheat were selected based on the information available from the previous studies (Zadoks et al., 1974). For all treatments, the fertilizer was applied as basal dose with pure nitrogen (urea) 150 kg/ha, P₂O₅ 150 kg/ha, and K₂O 150 kg/ha.

Canopy Reflectance Measurement

The canopy spectral reflectance was captured from jointing stage to filling stage with a FieldSpec 3.0 Spectrometer [Analytical Spectral Devices (ASD), Boulder, CO, United States]. The sample dates are listed in **Table 1**. The spectral range of the device is 350–2,500 nm, with a sampling interval of 1.4 nm and spectral resolution of 3 nm between 350 and 1,000 nm and a sampling interval of 2 nm and spectral resolution of 10 nm between 1,000 and 2,500 nm. The measurements were conducted under clear sky conditions during 10:00–14:00. The sensor with a field of view of 25° was held at a height of 1 m above the canopy vertically to cover a sensing area of the wheat canopy (~44.4 cm in diameter). In order to reduce the random error, spectral measurements were determined at three sites in each plot. Ten reflectance curves per site were averaged. A 40-cm² BaSO₄ calibration panel was used for calibrating the baseline reflectance prior to each measurement.

Plant Water Content Measurement

Samples were collected consistent with the spectral measurements, i.e., 20 plants were clipped at ground level in each plot and immediately put into valve bags avoiding water

loss. Then, the fresh plant was weighed. Samples were dried in an oven at 105°C for half an hour, then dried at 80°C to constant weight and reweighed. Plant water content (PWC) was calculated as Equation (1):

$$PWC (\%) = \frac{FW - DW}{DW} \times 100\% \quad (1)$$

Where FW is the fresh weight of wheat (g), and DW is the dry weight of wheat (g).

Spectral Transformation Methods

In order to get rid of the instrument noise, a moving Savitzky–Golay filter (Savitzky and Golay, 1964) with a window width of 5 nm and a polynomial of second degree was applied. The spectrum from 400 to 2,450 nm was selected. Due to the strong influence of the water vapor absorption peak on the canopy reflectance, wavelengths around 1,400 (1,350–1,400 nm) and 1,900 nm (1,800–1,950 nm) were removed.

According to previous studies, nine transformation methods were conducted on the canopy reflectance to improve the accuracy and select the best method in evaluating the PWC of winter wheat. The transformation methods are listed in **Table 2**, and the raw spectrum and the transformed spectra are shown in **Figure 1**.

Partial Least Squares Regression

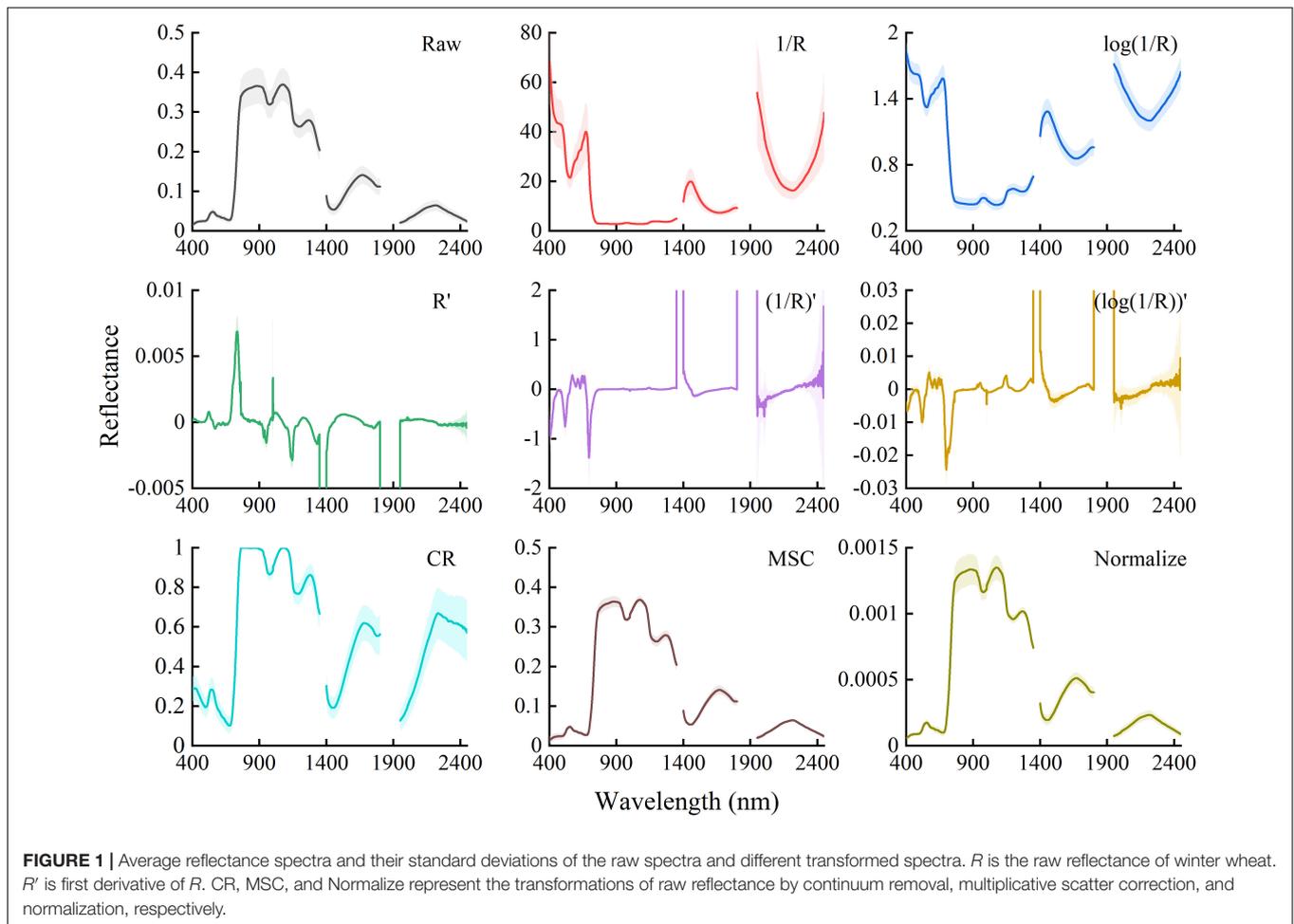
Partial least squares regression is one of the multivariate statistical analysis methods, which is a powerful tool in chemometrics and other fields (Wold et al., 1983). PLSR can handle data with high dimensionality and multicollinearity by reducing extensive collinear variables to noncorrelated factors (latent variables, LVs) (Wold, 1966). Then, the estimation model was established with LVs as the independent variables. In addition, the model is a linear regression model by projecting the independent variables and dependent variables (observable variables) to a new space. The method summarizes the spatial change information of independent variables, which can interpret the dependent variable as much as possible. In order to avoid overfitting,

TABLE 2 | Transformation methods used in this study.

No.	Transformation	Abbreviation	References
1	Raw reflectance	R	
2	The reciprocal reflectance	$1/R$	
3	Logarithm of the reciprocal reflectance	$\text{Log}(1/R)$	Grossman et al. (1996)
4	First derivative reflectance	R'	Demetriades-Shah et al. (1990)
5	First derivative of the reciprocal reflectance	$(1/R)'$	
6	First derivative of logarithm of the reciprocal reflectance	$(\text{Log}(1/R))'$	
7	Continuum removal	CR	Clark and Roush (1984)
8	Multiplicative scatter correction	MSC	Liang et al. (2010)
9	Normalization	Normalize	Yu et al. (1999)

TABLE 1 | Field reflectance spectra measurement dates in 2015 and 2016.

Year	Date 1	Date 2	Date 3	Date 4	Date 5
2015	13 April	24 April	3 May	13 May	22 May
2016	13 April	22 April	–	13 May	21 May



models were validated by leave-one-out cross-validation, and the $RMSE_{CV}$ was a basic criterion of determining the number of the latent variables. In general, the optimal latent variable number of the PLSR model was determined by the minimum RMSE of cross-validation, which was the inflection point or the location where the curve of $RMSE_{CV}$ became smooth (Kohonen et al., 2009).

Selection of Sensitive Bands

The aim of selecting sensitive bands is to extract the informative bands, simplify the complexity of the model, and construct a stable PWC estimation model. Based on a PLSR model, regression coefficients (called B -coefficients) and the variable importance in the projection (VIP) were commonly used as variable selection methods. In this study, five selection methods were used. In these methods, the first step was to selected the band ranges by the PLSR regression coefficients (B -coefficients) and the variable importance in the PLSR projection (VIP). B -coefficients represent the importance of each band in predicting the dependent variable. VIP delineates the relative importance of each band in the PLSR model for predicting PWC. The VIP method selects variables by calculating the VIP score for each variable and excludes all the variables with VIP score below

threshold 1 (Maestre, 2004). The VIP value for band j was calculated by the equation (Tran et al., 2014):

$$VIP_j = \sqrt{p \times \frac{\sum_{m=1}^M w_{mj}^2 SS(b_m \cdot t_m)}{\sum_{m=1}^M SS(b_m \cdot t_m)}} \quad (2)$$

Where p is the number of bands, M is the number of selected LVs, w_{mj} is the corresponding loading weight (the linear combination of the independent variables that define the LVs) of the j -th band for the m -th latent variable, and $SS(b_m \cdot t_m)$ is the explained sum of squared Y (PWC) by the PLSR model with m -th latent variable.

More effective bands could be selected through combining the B -coefficients and the VIP value in band selection (Chong and Jun, 2005). However, the combination method may select some bands with false positive (Tran et al., 2014). Thus, it is necessary to combine with other band selection methods. Four other variable selection methods were studied.

Stepwise Regression

Stepwise regression is a common method of selecting variables. The aim of the regression is to screen for significant bands by establishing the relation between the dependent variables and the independent variable. In this study, bidirectional elimination that

combined forward and backward elimination was used. It is an iterative process starting with no band. In each step, the band that contributes the most in predicting the dependent variable will be added to the model for a criterion of P value ($P < 0.01$). Then, the P value of all bands in the model will be tested, and bands whose P value is above the certain threshold (0.05) are removed. The process will repeat until all significant bands are in the regression model in estimating the dependent variable.

Successive Projections Algorithm

Successive projections algorithm is a method to solve the collinearity problem among variables (Araújo et al., 2001). It used a simple projection operation to obtain subsets of variables with minimal collinearity and is a forward algorithm for band selection. A new selected variable had the maximum projection value on the orthogonal subspace of the previous selected variable. The sensitive bands and the number of sensitive bands were determined based on the smallest root mean square error (RMSE) of cross-validation of multiple linear regression model.

Random Frog

Random frog is an efficient method for band selection in recent years. It was based on the framework of invertible-jump Markov Monte Carlo. In this method, PLSR was used as a modeling method. According to definite criteria, the variable subsets are updated constantly. When the number of iterations is reached, the selected frequency of each variable was calculated as the basis for eliminating redundant information (Yun et al., 2013). It requires a threshold of probability which is usually set empirically. Thus, a forward variable selection was used. According to frequency, bands were ranked from the largest to the smallest. In each iteration, the top-ranked band was added into the PLSR model. With the leave-one-out cross-validation, the number of bands was decided when the minimum RMSE was reached.

Uninformative Variables Elimination

Uninformative variables elimination is proposed to eliminate uninformative variable (Centner et al., 1996). Bands were evaluated with the stability coefficients of the regression coefficient (B -coefficient). A same size matrix of noise was jointed with the spectra matrix. The PLSR prediction model was established by the new independent variable matrix. Having leave-one-out validation, the B -coefficient matrix with sample in row and variable (band and noise) in column was obtained. Then, the stability coefficients were calculated with the average value divided by the standard deviation value of each variable B -coefficient. Compared with noise variable, the band with lower absolute stability coefficient would be removed.

Performance Evaluation of the Developed Models

The experiment data collected from 2014 to 2015 ($n = 150$) and from 2015 to 2016 ($n = 120$) were used for model calibration (C) and validation (V), respectively. The PLSR models for estimating the PWC were evaluated based on the coefficient of determination (R^2), RMSE, and ratio of performance to

interquartile distance (RPIQ). The calculation formulae are listed as Equations 3–5:

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \tilde{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (3)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (\tilde{y}_i - y_i)^2}{nc}} \quad (4)$$

$$RPIQ = \frac{Q3 - Q1}{RMSE} \quad (5)$$

Where y_i , \tilde{y}_i are the measured PWC and predicted PWC of the i -th sample in different datasets; \bar{y} is the mean value of PWC in the calibration and validation sets; and Q3 and Q1 are third quartile and first quartile of the datasets, respectively. According to the value of RPIQ, four categories of models are defined: RPIQ between 2.02 and 2.70 illustrates a poor model where only high and low values are distinguishable; RPIQ between 2.70 and 3.37 demonstrates a model where quantitative predictions are possible; RPIQ between 3.37 and 4.05 exhibits a good, quantitative model; and RPIQ > 4.05 implies an excellent model (Ludwig et al., 2017). Generally, a robust model should have a higher R^2 and RPIQ and lower RMSE.

Data Processing Software

Stepwise regression was analyzed in SPSS 19.0 (SPSS Inc., Chicago, United States). PLSR models were performed with PLS Toolbox version 6.20 (Eigenvector Research, Inc., Wenatchee, WA, United States) that ran under MATLAB version R2010a. Other band selection methods were performed with the same MATLAB environment. Graphs were prepared with Origin 9.0 (Microcal, United States).

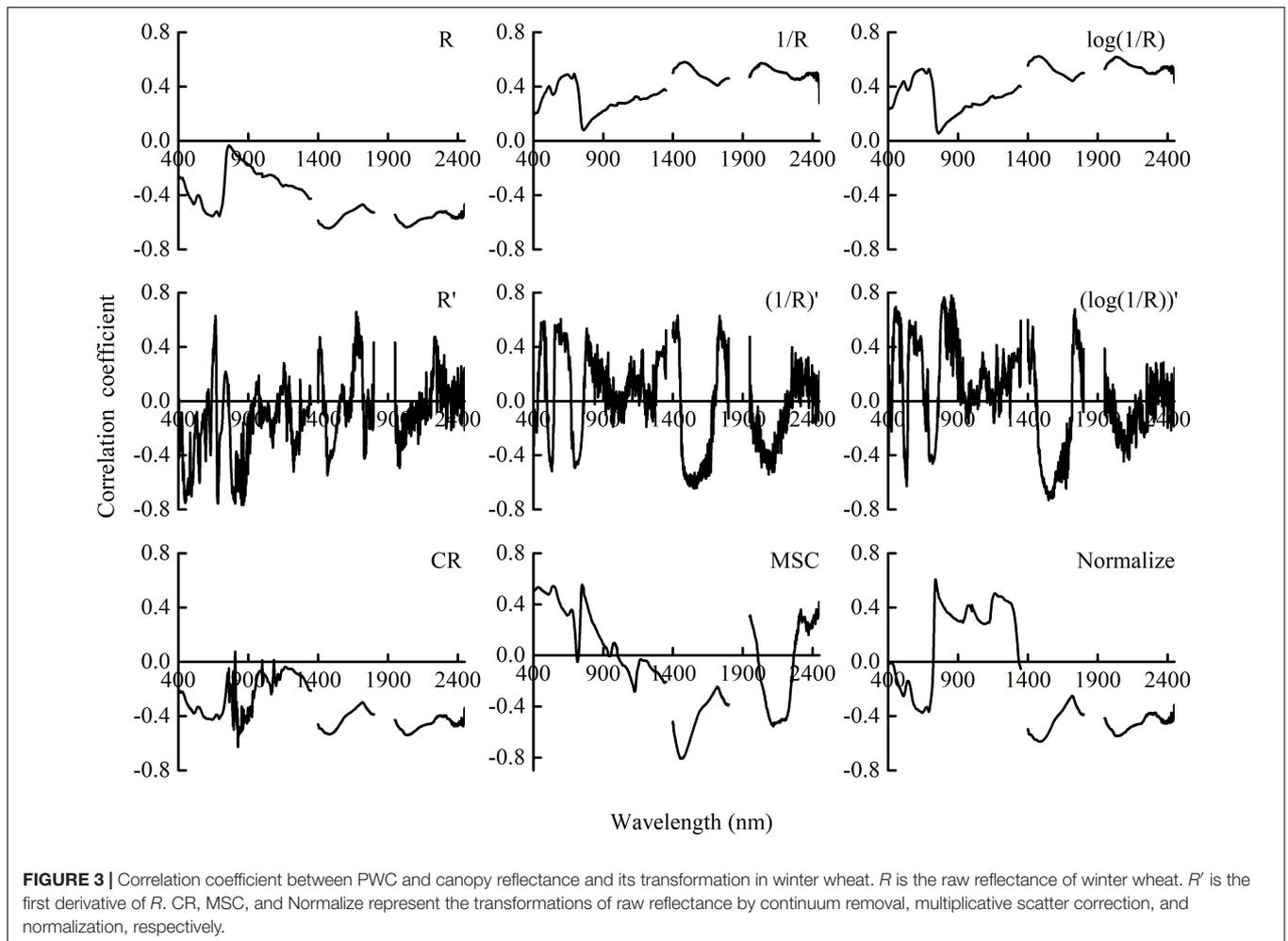
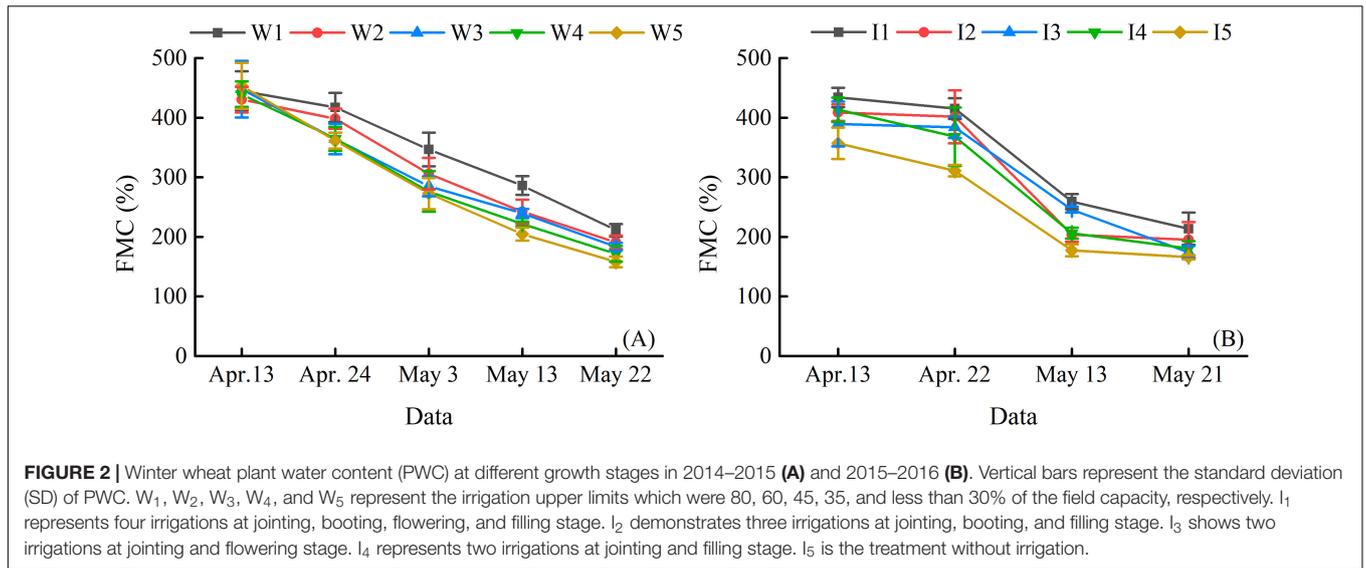
RESULTS

Variation of Winter Wheat Plant Water Content Between Water Treatments

Figure 2 shows the PWC at the main growth stages in the 2014–2015 and 2015–2016 growing seasons. Winter wheat PWC tended to decrease throughout the entire growing seasons of two consecutive years, and the minimum PWC was obtained on May 22, 2015, and May 21, 2016, respectively. The PWC was increased with the increase of irrigation ceiling (2014–2015) and frequency (2015–2016).

Correlation of PWC and Canopy Transformed Reflectance

The correlation coefficients between PWC and canopy reflectance with different transformation methods are shown in Figure 3. Correlations were negative in the range of the bands implemented in this study. Compared with raw canopy reflectance, curve with normalized transformation exhibited the same pattern and kept the basic features. $1/R$ and $\log(1/R)$ showed an opposite curve pattern compared with raw reflectance. The transformation



methods with derivation can obviously enhance the correlation in the visible regions to raw reflectance, up to 0.76 for R' . The correlation coefficients for the treatment of CR significantly

decreased. Multiplicative scatter correction (MSC) coefficients did exhibit the same pattern to R . The exception was that MSC was positive in the visible region.

PLSR Model for Estimating PWC Based on the Full Spectrum

The RMSE trend of the PLSR model leave-one-out cross-validation based on different spectral transformation models is shown in **Figure 4**. Except for CR, the number of LVs for different transformation methods was selected by the minimum $RMSE_{CV}$.

To quantify the PWC of winter wheat, the data from 2014 to 2015 through different pretreatments were used to establish the PLSR models. The data from 2015 to 2016 were used as an external validation to the PLSR models. The comparative analysis of the prediction accuracy of PLSR models based on different spectral transformations is listed in **Table 3** and **Figure 5**. The results showed that the PLSR model based on the full spectrum was suitable for estimating the PWC with R^2_C values above 0.87, except for the CR. The models based on the reciprocal method and the derivative method did show a higher R^2_C ($R^2 > 0.90$) and $RPIQ_C$ ($RPIQ > 6.0$) and lower $RMSE_C$ ($RMSE < 31\%$). In addition, the models with derivative method (PLSR-4, PLSR-5, and PLSR-6) exhibited less LVs than the other treatments (except for CR). On the other hand, the performances of PLSR-5 and PLSR-6 in validation were poorer than PLSR-4, illustrating that the model was optimal. Therefore, among all the spectral transformation methods, the first derivative reflectance was superior in estimating the PWC in winter wheat.

PLSR Model for Estimating PWC Based on the Sensitive Bands

Selection of the Sensitive Bands

In order to simplify the model and avoid overfitting, sensitive bands were selected based on the first derivative transformed spectrum. The B -coefficients and VIP values derived from the PLSR-4 model indicated the importance of different bands in predicting PWC (**Figure 6**). The VIP and B -coefficients did show similar trends. First, bands where the VIP value was above 1 (the threshold of VIP) or the absolute values of B -coefficients were greater than the criterion (the mean of absolute value of B -coefficients) were selected. Due to the redundancy and

inefficiency of the band, four methods (SR, SPA, RF, and UVE) were then used to further select the sensitive bands. The results of the method of using VIP and B -coefficient values of PLSR to select bands (denoted as B+VIP) and the four band selection methods are shown in **Figure 6**. B+VIP extracted the largest range of sensitive bands. In addition, other methods effectively reduced the number of sensitive bands, especially SR. The locations of the selected bands using different methods were similar, and they were located in the spectral regions centered at about 680, 860, 980, 1,285, 1,580, 1,660, 1,980, 2,184, 2,250, 2,350, and 2,430 nm. Most of the sensitive bands were distributed in the near-infrared and short-wave infrared regions.

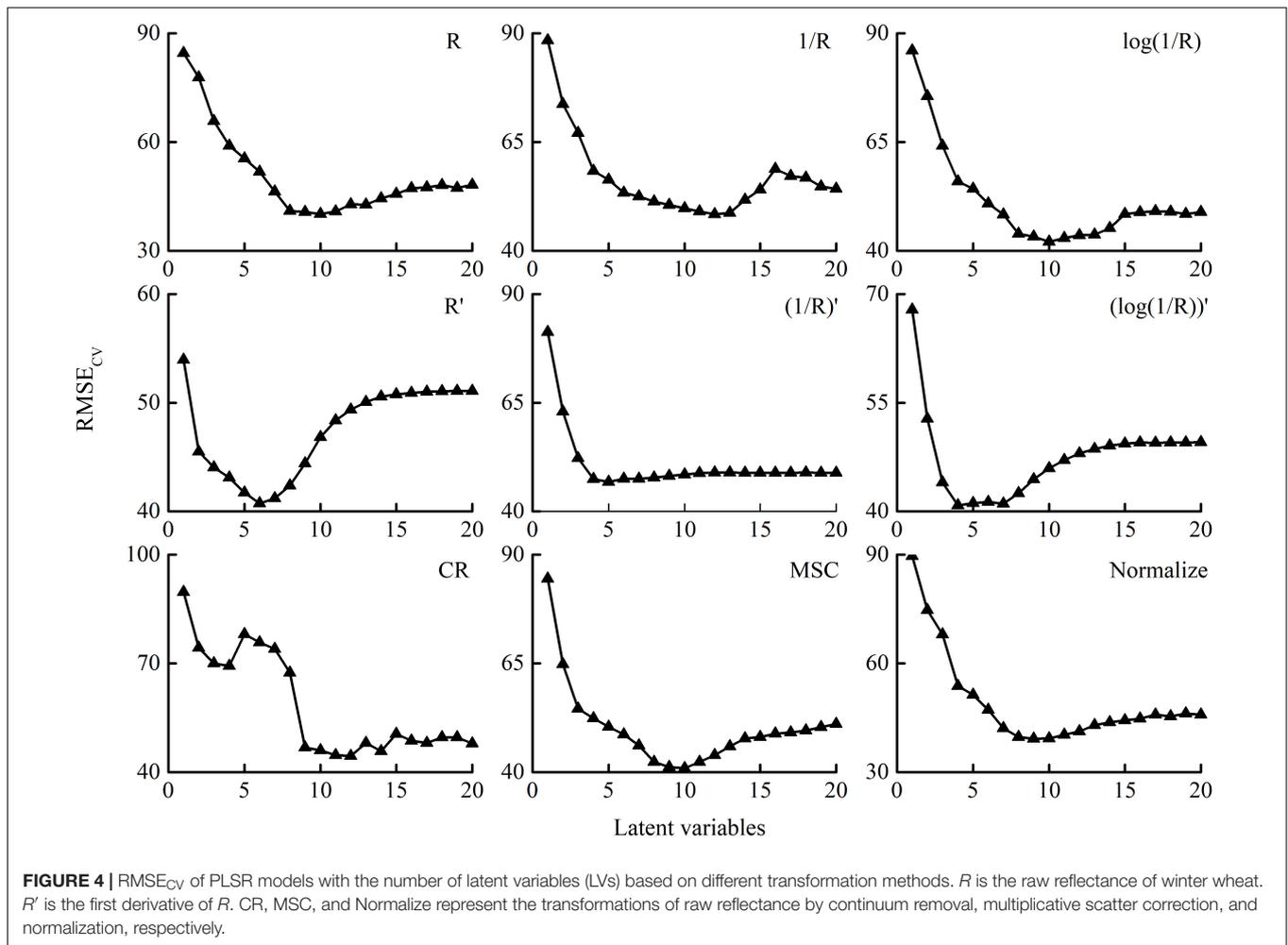
PLSR Model Based on Sensitive Bands

Using the sensitive bands selected in the *Selection of the Sensitive Bands* section, new PLSR models were established for estimating the PWC of winter wheat (**Table 4**), and the relationships between the measured value and the predicted value in model validation are shown in **Figure 7**. The models with sensitive bands demonstrated robust performance in model calibration ($R^2_C > 0.80$, $RMSE_C < 45\%$, and $RPIQ_C > 4.2$). Compared with the method (B+VIP of the PLSR model), the number of variables selected by the combined variable selection methods was reduced by at least 75%. Except for B+VIP-RF, the model accuracies were slightly reduced. Take B+VIP-SR for example, only 21 bands were selected to build the new model. The model accuracies were reduced in calibration and validation with R^2 differences of 0.02 and 0.07 and RMSE differences of 4.22 and 15.3%, respectively. Among different variable selection methods, 244 sensitive bands were extracted through B+VIP-RF, and the model established by these sensitive bands performed best during model calibration and validation ($R^2_C = 0.99$, $RMSE_C = 11.53\%$, $RPIQ_C = 16.34$; $R^2_V = 0.84$, $RMSE_V = 44.40\%$, $RPIQ_V = 4.52$). Considering the number of variables and model performance, the combination of PLSR and RF (B+VIP-RF) was optimal. According to $RPIQ_V$, the new model had good accuracy for PWC. In **Figure 7**, the data

TABLE 3 | Performances of PLSR models for PWC based on different spectral transformation methods.

No.	Spectral transformations	Calibration dataset (2014–2015)					Validation dataset (2015–2016)		
		LVs	$RMSE_{CV}$	R_C^2	$RMSE_C$	$RPIQ_C$	R_V^2	$RMSE_V$	$RPIQ_V$
1	R	8	40.99	0.88	35.35	5.33	0.78	75.03	2.68
2	$1/R$	12	48.38	0.91	30.78	6.12	0.82	47.81	4.20
3	$\log(1/R)$	10	42.10	0.90	31.97	5.89	0.83	58.63	3.43
4	R'	6	40.73	0.96	20.49	9.19	0.86	46.27	4.34
5	$(1/R)'$	4	47.43	0.91	30.87	6.10	0.73	52.43	3.83
6	$(\log(1/R))'$	4	40.84	0.91	29.95	6.29	0.81	50.54	3.98
7	CR	4	69.27	0.65	59.42	3.17	0.53	79.29	2.53
8	MSC	9	41.17	0.89	33.18	5.68	0.82	71.06	2.83
9	Normalize	8	39.67	0.88	34.70	5.43	0.83	56.64	3.55

Latent variable indicates the number of latent variables selected by the model. R_C^2 and R_V^2 represent the coefficient of determination of model calibration and validation, respectively. $RMSE_C$, $RMSE_{CV}$, and $RMSE_V$ represent the root mean square error of calibration, cross-validation (leave-one-out) and validation, respectively. $RPIQ_C$ and $RPIQ_V$ represent the ratio of performance to interquartile distance of the calibration and validation models, respectively. R is the raw reflectance of winter wheat. R' is the first derivative of R . CR, MSC, and Normalize represent the transformations of raw reflectance by continuum removal, multiplicative scatter correction, and normalization, respectively.



points were close to the 1:1 line, indicating good estimation in different ranges of PWC.

DISCUSSION

Methods for Spectral Proximal Sensing Data Preprocessing in Estimating Water Content

In the estimation with hyperspectral remote sensing, the disturbing factors (e.g., soil background and instrument noise) had a negative influence. By using suitable preprocessing and transformation, the accuracy and stability of the prediction model were improved. In this study, nine transformation methods were analyzed and compared. Compared with raw reflectance (*R*), the correlation between CR and PWC significantly decreased (Figure 3). The poor correlation with PWC was associated with the result of the transformation. The location and reflection feature of spectral absorption of variables is emphasized by CR transformation, and the values are normalized (Clark and Roush, 1984; Mutanga et al., 2005). However, the absorption features of parameters, in addition to water content, are also

enhanced. It may result in the pattern of RMSE_{CV} fluctuated with the increase in the number of LVs (Figure 4). On the contrary, other transformation methods increased the correlation to some extent and did show good performance ($R^2_C > 0.87$, RMSE_C < 36.00%, and RPIQ_C > 5.30) in model calibration. Consistent with Zhang et al. (2012), derivative transformations reduced the complexity of the model with less LVs but showed better correlation and good performance in water content estimation (Figure 4 and Table 3). The full-spectrum model with the first derivative transformation was the optimum (Table 3), indicating its advantages in estimating plant water content with canopy reflectance. Derivative transformation resolved overlapping spectra and emphasized the weak but meaningful peak (Shibayama et al., 1993). In this study, soil may contribute more to the canopy reflectance of plant under severe water stress. The first derivative transformation could suppress the spectral response of the soil background by converting it into a constant (Demetriades-Shah et al., 1990), and it could eliminate the effects of canopy architecture difference under various water conditions. In addition, our study elucidated that the correlation analysis can be used as an elementary selection method for spectral transformations.

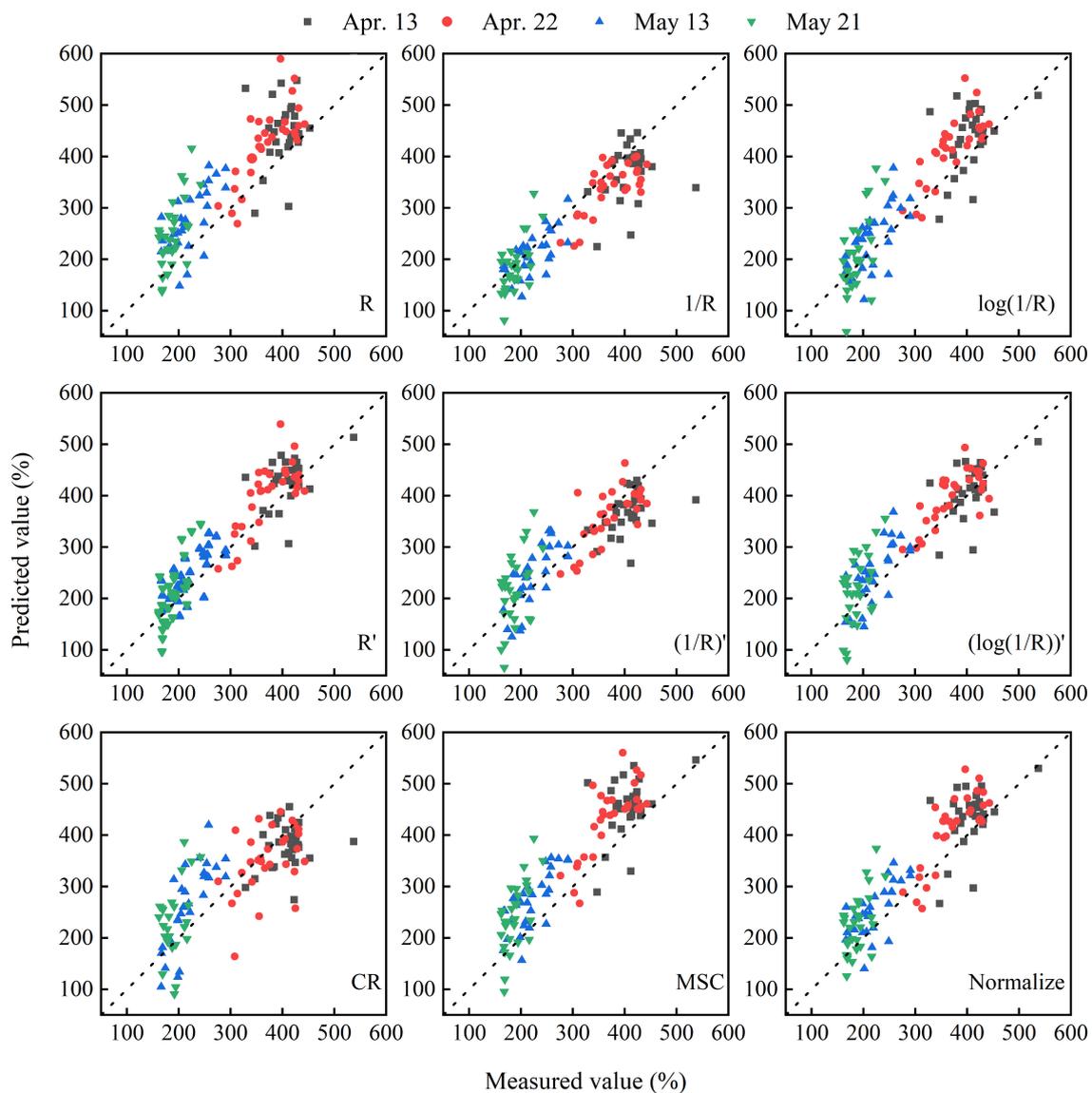


FIGURE 5 | The correlation between measured and predicted values in estimating PWC in winter wheat based on different transformation methods. The dotted line represents the 1:1 line. R is the raw reflectance of winter wheat. R' is the first derivative of R . CR, MSC, and Normalize represent the transformations of raw reflectance by continuum removal, multiplicative scatter correction, and normalization, respectively.

Approaches to Sensitive Band Selection

Five variable selection methods were used to extract the sensitive bands for PWC estimation. The numbers of input variables (sensitive bands) of the estimation models were significantly reduced, and B+VIP-SR had the most effects. The number of sensitive bands was reduced to only 21. Compared with the full spectrum, the calibration model with sensitive bands selected by B+VIP had higher estimation accuracy and better stability. This indicated that variable selection could reduce model complexity and improve model performance due to the removal of irrelevant and interference variables (Krishna et al., 2019). The B -coefficient and VIP value of the PLSR model reflected the weight of variables in predicting PWC, and they can be used as the basis for band

selection (Sharabian et al., 2014; Wang et al., 2017). The criterion of the two parameters used in this study is effective to make it a good variable selection method. However, the number of selected bands is still very large, and further band extraction is needed. Thus, this method can be used as an initial band selection method.

On this basis, the other four methods were used to further reduce the number of sensitive bands. Comparing four combined band selection methods, the new calibration models demonstrated similar results, with high R^2 and RPIQ and low RMSE. This indicated that the four methods used in this study were effective. Among the four methods, SPA showed the lowest R^2 and the highest RMSE in model calibration, followed by

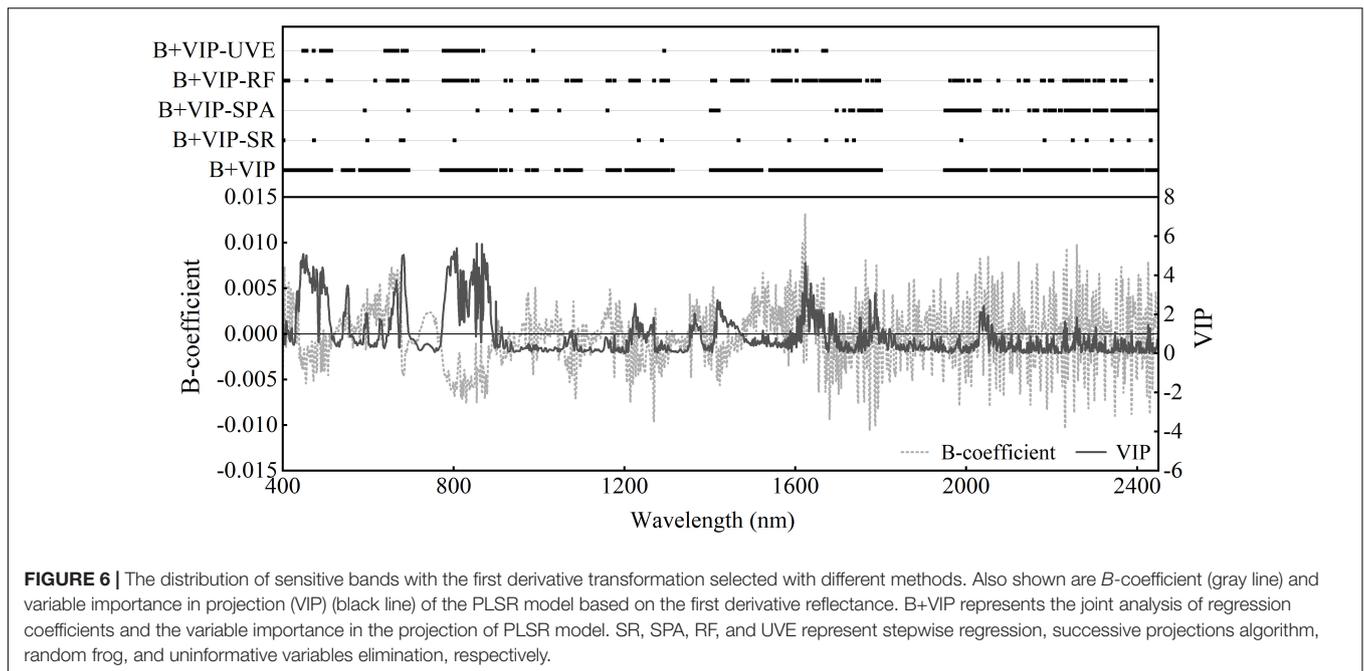


FIGURE 6 | The distribution of sensitive bands with the first derivative transformation selected with different methods. Also shown are B -coefficient (gray line) and variable importance in projection (VIP) (black line) of the PLSR model based on the first derivative reflectance. B+VIP represents the joint analysis of regression coefficients and the variable importance in the projection of PLSR model. SR, SPA, RF, and UVE represent stepwise regression, successive projections algorithm, random frog, and uninformative variables elimination, respectively.

TABLE 4 | The performance of PLSR models based on the first derivative spectrum sensitive bands.

Method	Number	Calibration dataset (2014–2015)			Validation dataset (2015–2016)		
		R_c^2	RMSE _C	RPIQ _C	R_v^2	RMSE _V	RPIQ _V
B+VIP	987	0.97	17.59	10.71	0.85	44.93	4.47
B+VIP-SR	21	0.95	21.81	8.64	0.78	60.23	3.34
B+VIP-SPA	145	0.81	44.67	4.22	0.62	63.89	3.14
B+VIP-RF	244	0.99	11.53	16.34	0.84	44.40	4.52
B+VIP-UVE	115	0.88	34.81	5.41	0.85	60.74	3.31

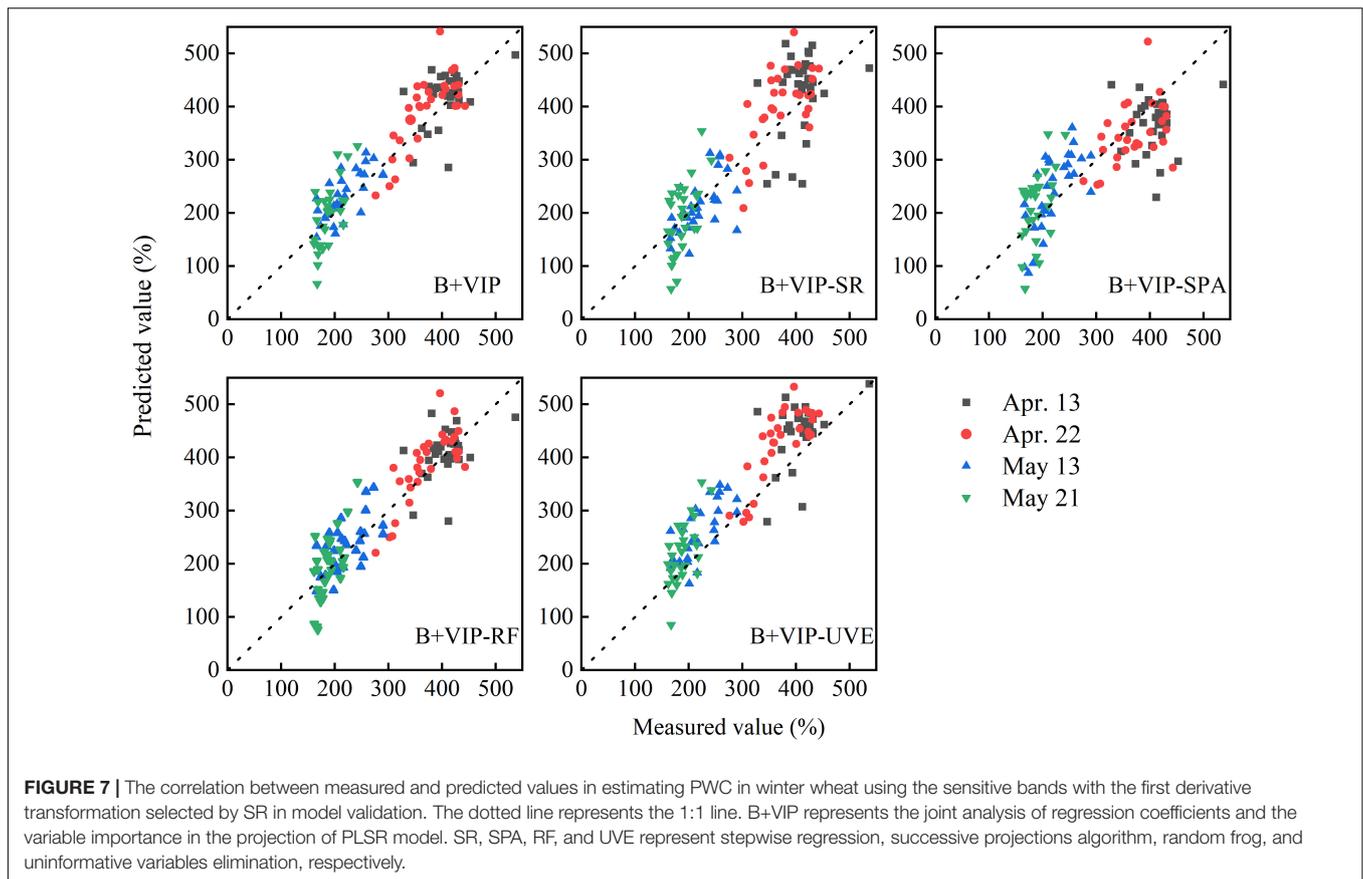
B+VIP represents the joint analysis of regression coefficients and the variable importance in the projection of the PLSR model. SR, SPA, RF, and UVE represent stepwise regression, successive projections algorithm, random frog, and uninformative variables elimination, respectively. R_c^2 and R_v^2 represent the coefficient of determination of model calibration and validation, respectively. RMSE_C, RMSE_{Cv}, and RMSE_V represent the root mean square error of calibration, cross-validation (leave-one-out) and validation, respectively. RPIQ_C and RPIQ_V represent the ratio of performance to interquartile distance of the calibration and validation models, respectively.

UVE. It may be explained by the fact that SPA had selected the bands with minimum collinearity while ignoring some important bands and containing the noise information (Soares et al., 2013). Although SPA was performed on the basis of the PLSR (B+VIP) method, the noisy bands may still be selected. It indicated that B+VIP cannot remove all noisy bands. As for UVE, a previous study had proved that it can effectively extract sensitive bands (Yang et al., 2015). It determined invalid bands by comparing the stability coefficient of spectrum and noise. On the basis of PLSR, some effective bands with lower stability coefficient may be deleted, which resulted in the decreasing of model accuracy. It indicated that the combination of PLSR and UVE is not appropriate. With respect to the other methods, they showed

good capabilities in band extraction. SR played a vital role in variable selection in a small number of samples, contributing to more reliable results. The result of PLSR-SR was similar to that of previous studies (Li et al., 2017; Wang et al., 2017; Xie et al., 2020), which reported that PLSR-SR could extract the effective sensitive bands and build the estimation model with acceptable accuracy. In addition, since only 21 bands were selected, PLSR-SR has the potential for developing the new instrument. RF can effectively extract sensitive bands (Table 4), which was proven by studies (Chen and Li, 2020). With 244 bands, PLSR-RF showed better or similar accuracy than PLSR in model calibration and validation, respectively. It indicated the advantages of RF in band extraction. Furthermore, the model accuracy has difference between the validation dataset and calibration dataset. In this study, two experimental datasets of different varieties and water stress treatments were used for model calibration and validation. Although the selected bands have good performance in model calibration, changes in cultivars and cultivation environment may cause the displacement of sensitive bands, resulting in some changes in the model validation. Considering the number of bands and the performance of model calibration and validation, PLSR-RF was the optimal method.

Sensitive Bands for Plant Water Content Estimation

The sensitive bands selected through five variable selection methods did exhibit similar spectral regions with the central band being 680, 860, 980, 1,285, 1,580, 1,660, 1,980, 2,184, 2,250, 2,350, and 2,430 nm (Figure 6). It indicated that these central bands were useful in the estimation of winter wheat PWC. These sensitive bands were distributed in each spectral range studied. The new models based on the sensitive bands exhibited good estimation power with the RPIQ of model



calibration and validation above 3.1. It is consistent with previous studies, showing that the combination of visible, near-infrared, and short-wave infrared bands certified robust performance (Zhang et al., 2013; Bendig et al., 2015; Kusnierek and Korsaeht, 2015). Using the method of B+VIP-RF, the estimation model had the best performance and the bands were located around 400–410, 500, 645–690, 770–850, 930, 980–995, 1,100, 1,165, 1,220, 1,290, 1,450–1,480, 1,550–1,590, 1,610–1,750, 1,960–1,990, 2,200, and 2,230–2,360 nm. It had been proved that the near-infrared and short-wave infrared spectral regions (900–2,500 nm) were directly affected by water (Gao and Goetz, 1994; Yebra et al., 2013). Majority of the sensitive bands were in the short-wave infrared region which may provide more information. Specifically, 980–995, 1,220, 1,450–1,480, 1,750, and 2,230–2,360 nm were near the major water absorption bands at 970, 1,200, 1,450, 1,750, and 2,250 nm (Danson et al., 1992; Shibayama et al., 1993; Pu et al., 2003; Clevers et al., 2008). The region of 1,150–1,260 nm was proved to be one of the optimal bands for ground-based remote sensing of vegetation water content (Sims and Gamon, 2003). The 2,230–2,360-nm bands were located in the region of 2,080–2,350 nm which can be used for water content estimation, but it was strongly affected by the soil reflectance when the cover is less than 100% (Ripple, 1986). Concurrently, many selected sensitive bands were indirectly relevant to the plant water content. The selected bands located in the visible region were related to the absorption of chlorophyll content in plants (Yebra et al., 2013; Steidle Neto et al., 2017). Band valleys

around 500 and 680 nm did show a good correlation with the chlorophyll pigments representing the color characteristics (Yu et al., 2014). This can be explained by that water content exhibited a significant correlation with chlorophyll content and other pigments (Jaleel et al., 2009; Keyvan, 2010; Din et al., 2011). Some bands were located in the near-infrared region, which reflected the internal structure of the cells. It is mainly because plant water status directly influences cell turgor and internal space of the tissue resulting in the change of internal scattering. Some bands were selected in the specific region which were associated with the absorption of protein for the N–H asymmetry stretch and amide II at 1,980 nm and C–H stretch at 2,240 nm (Curran, 1989; Sandra et al., 2000). Presumably, it is because of the close relationship between water status and protein.

Estimation of Plant Water Content With Multivariate Statistical Methods

As a powerful multivariate statistical method, PLSR models had good performance in estimating the PWC of winter wheat (Tables 3, 4). In model validation, the data distributions from different growth stages were inconsistent (Figures 5, 7), indicating that model performance was affected by plant growth stages. The results were in accordance with previous studies (Rischbeck et al., 2016; Wang et al., 2017). Plant development was related to plant internal structure, substances, and function, which would affect the location of sensitive bands in different growth stages. As the growth period is promoted, the data

points of majority models were getting closer to the 1:1 line, and the RMSE of the measured and predicted PWC decreased (data not shown), with the minimum RMSE appearing on May 21. It indicated that the filling stages performed better than the other stages. In addition, the model performance was not only affected by the plant growth but also by the cultivars and experimental treatment (El-Hendawy et al., 2019a,b). In this study, the influence of these factors on multivariate statistics was not considered, and further research would be conducted in the future.

CONCLUSION

In this study, the results demonstrated that the canopy spectral proximal sensing data could be used for estimating the PWC of winter wheat in an accurate and nondestructive way. Among different spectral transformation methods, the first derivative transformation had closer correlation with PWC, and the PLSR-4 model was more effective in calibration and validation, indicating that the first derivative transformation was the optimal method for processing the canopy spectral data in estimating PWC. The sensitive bands extracted by the combination of PLSR and RF obtained the best performance. The verification accuracy of the model slightly decreased compared with the full-spectrum model, which is acceptable. The comprehensive method (PLSR-RF) performed best in extracting the sensitive bands and building a simple and stable model. The findings of

this study provide technical support for large-scale monitoring of plant water status by using canopy spectral proximal sensing data in field production.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

HS, MF, and WY conceived the research. HS, MF, LX, and WY designed the experiments. HS, LX, CW, XJ, GW, and SZ carried out the field measurements. HS, MF, and CW provided technical support. GD reviewed and edited the manuscript. All authors have read and agreed to the published version of the manuscript.

FUNDING

This work was supported by the National Natural Science Foundation of China (Grant No. 31871571), Applied Basic Research Project of Shanxi Province (Grant No. 201801D221299), and Science and Technique Innovation Project of Shanxi Agricultural University (Grant No. 2020BQ32).

REFERENCES

- Araújo, M. C. U., Saldanha, T. C. B., Galvão, R. K. H., Yoneyama, T., Chame, H. C., and Visani, V. (2001). The successive projections algorithm for variable selection in spectroscopic multicomponent analysis. *Chemometr. Intell. Lab. Syst. 57*, 65–73. doi: 10.1016/S0169-7439(01)00119-8
- Atzberger, C., Guérif, M., Baret, F., and Werner, W. (2010). Comparative analysis of three chemometric techniques for the spectroradiometric assessment of canopy chlorophyll content in winter wheat. *Comput. Electron. Agric. 73*, 165–173. doi: 10.1016/j.compag.2010.05.006
- Bendig, J., Yu, K., Aasen, H., Bolten, A., Bennertz, S., Broscheit, J., et al. (2015). Combining UAV-based plant height from crop surface models, visible, and near infrared vegetation indices for biomass monitoring in barley. *Int. J. Appl. Earth Obs. Geoinf. 39*, 79–87. doi: 10.1016/j.jag.2015.02.012
- Centner, V., Massart, D. L., De Noord, O. E., And, S. D. J., Vandeginste, B. G. M., and Sterna, C. (1996). Elimination of uninformative variables for multivariate calibration. *Anal. Chem. 68*, 3851–3858. doi: 10.1021/ac960321m
- Chen, J., and Li, G. (2020). Prediction of moisture content of wood using Modified Random Frog and Vis-NIR hyperspectral imaging. *Infrared Phys. Technol. 105*:103225. doi: 10.1016/j.infrared.2020.103225
- Cheng, T., Rivard, B., and Sánchez-Azofeifa, A. (2011). Spectroscopic determination of leaf water content using continuous wavelet analysis. *Remote Sens. Environ. 115*, 659–670. doi: 10.1016/j.rse.2010.11.001
- Cho, M. A., Skidmore, A., Corsi, F., Wieren, S. E. V., and Sobhan, I. (2007). Estimation of green grass/herb biomass from airborne hyperspectral imagery using spectral indices and partial least squares regression. *Int. J. Appl. Earth Obs. Geoinf. 9*, 414–424. doi: 10.1016/j.jag.2007.02.001
- Chong, I. G., and Jun, C. H. (2005). Performance of some variable selection methods when multicollinearity is present. *Chemometr. Intell. Lab. Syst. 78*, 103–112. doi: 10.1016/j.chemolab.2004.12.011
- Clark, R. N., and Roush, T. L. (1984). Reflectance spectroscopy: quantitative analysis techniques for remote sensing applications. *J. Geophys. Res. Solid Earth 89*(B7), 6329–6340. doi: 10.1029/JB089iB07p06329
- Clevers, J. G. P. W., Kooistra, L., and Schaepman, M. E. (2008). Using spectral information from the NIR water absorption features for the retrieval of canopy water content. *Int. J. Appl. Earth Obs. Geoinf. 10*, 388–397. doi: 10.1016/j.jag.2008.03.003
- Clevers, J. G. P. W., Kooistra, L., and Schaepman, M. E. (2010). Estimating canopy water content using hyperspectral remote sensing data. *Int. J. Appl. Earth Obs. Geoinf. 12*, 119–125. doi: 10.1016/j.jag.2010.01.007
- Curran, P. J. (1989). Remote sensing of foliar chemistry. *Remote Sens. Environ. 30*, 271–278. doi: 10.1016/0034-4257(89)90069-2
- Danson, F., Steven, M., Malthus, T., and Clark, J. (1992). High-spectral resolution data for determining leaf water content. *Int. J. Remote Sens. 13*, 461–470. doi: 10.1080/01431169208904049
- Daryanto, S., Wang, L., and Pierre-André, J. (2017). Global synthesis of drought effects on cereal, legume, tuber and root crops production: a review. *Agric. Water Manage. 179*, 18–33. doi: 10.1016/j.agwat.2016.04.022
- Das, B., Sahoo, R. N., Pargal, S., Krishna, G., Verma, R., Chinnusamy, V., et al. (2017). Comparison of different uni- and multi-variate techniques for monitoring leaf water status as an indicator of water-deficit stress in wheat through spectroscopy. *Biosyst. Eng. 160*, 69–83. doi: 10.1016/j.biosystemseng.2017.05.007
- Demetriades-Shah, T. H., Steven, M. D., and Clark, J. A. (1990). High resolution derivative spectra in remote sensing. *Remote Sens. Environ. 33*, 55–64. doi: 10.1016/00344257(90)90055-Q
- Din, J., Khan, S. U., Ali, I., and Gurmani, A. R. (2011). Physiological and agronomic response of canola varieties to drought stress. *J. Animal Plant Sci. 21*, 78–82. doi: 10.4081/ijas.2011.e11
- El-Hendawy, S. E., Alotaibi, M., Al-Suhaibani, N., Al-Gaadi, K., Hassan, W., Dewir, Y. H., et al. (2019a). Comparative performance of spectral reflectance indices and multivariate modeling for assessing agronomic parameters in advanced spring wheat lines under two contrasting irrigation regimes. *Front. Plant Sci. 10*:1537. doi: 10.3389/fpls.2019.01537
- El-Hendawy, S. E., Al-Suhaibani, N. A., Elsayed, S., Hassan, W. M., Dewir, Y. H., Refay, Y., et al. (2019b). Potential of the existing and novel spectral reflectance

- indices for estimating the leaf water status and grain yield of spring wheat exposed to different irrigation rates. *Agric. Water Manage.* 217, 356–373. doi: 10.1016/j.agwat.2019.03.006
- Fang, M., Ju, W., Zhan, W., Cheng, T., Qiu, F., and Wang, J. (2017). A new spectral similarity water index for the estimation of leaf water content from hyperspectral data of leaves. *Remote Sens. Environ.* 196, 13–27. doi: 10.1016/j.rse.2017.04.029
- Gao, B., and Goetz, A. F. H. (1994). Extraction of dry leaf spectral features from reflectance spectra of green vegetation. *Remote Sens. Environ.* 47, 369–374. doi: 10.1016/0034-4257(94)90104-X
- Gao, B. C. (1996). NDWI—a normalized difference water index for remote sensing of vegetation liquid water from space. *Remote Sens. Environ.* 58, 257–266. doi: 10.1016/S0034-4257(96)00067-3
- González-Fernández, A. B., Rodríguez-Pérez, J. R., Marabel, M., and Álvarez-Taboada, F. (2015). Spectroscopic estimation of leaf water content in commercial vineyards using continuum removal and partial least squares regression. *Sci. Hortic.* 188, 15–22. doi: 10.1016/j.scienta.2015.03.012
- Grossman, Y. L., Ustin, S. L., Jacquemoud, S., Sanderson, E. W., Schmuck, G., and Verdebout, J. (1996). Critique of stepwise multiple linear regression for the extraction of leaf biochemistry information from leaf reflectance data. *Remote Sens. Environ.* 56, 182–193. doi: 10.1016/0034-4257(95)00235-9
- Ithuma, S. O., and Madramootoo, C. A. (2017). Recent advances in crop water stress detection. *Comput. Electron. Agric.* 141, 267–275. doi: 10.1016/j.compag.2017.07.026
- Jaleel, C. A., Manivannan, P., Wahid, A., Farooq, M., Al-Juburi, H. J., Somasundaram, R., et al. (2009). Drought stress in plants: a review on morphological characteristics and pigments composition. *Int. J. Agric. Biol.* 11, 100–105. doi: 10.3763/ijas.2009.0459
- Keyvan, S. (2010). The effects of drought stress on yield, relative water content, proline, soluble carbohydrates and chlorophyll of bread wheat cultivars. *J. Animal Plant Sci.* 8, 1051–1060.
- Kohonen, J., Reinikainen, S. P., Aaljoki, K., and Höskuldsson, A. (2009). Non-linear PLS approach in score space. *Chemometr. Intell. Lab. Syst.* 97, 159–163. doi: 10.1016/j.chemolab.2009.03.010
- Krishna, G., Sahoo, R. N., Singh, P., Patra, H., Bajpai, V., Das, B., et al. (2019). Application of thermal imaging and hyperspectral remote sensing for crop water deficit stress monitoring. *Geocarto Int.* 18, 1–18. doi: 10.1080/10106049.2019.1618922
- Kusnierek, K., and Korsath, A. (2015). Simultaneous identification of spring wheat nitrogen and water status using visible and near infrared spectra and powered partial least squares regression. *Comput. Electron. Agric.* 117, 200–213. doi: 10.1016/j.compag.2015.08.001
- Li, G., Wang, C., Feng, M., Yang, W., Li, F., and Feng, R. (2017). Hyperspectral prediction of leaf area index of winter wheat in irrigated and rainfed fields. *PLoS One* 12:e0183338. doi: 10.1371/journal.pone.0183338
- Li, L., Cheng, Y. B., Ustin, S., Hu, X. T., and Riaño, D. (2008). Retrieval of vegetation equivalent water thickness from reflectance using genetic algorithm (GA)-partial least squares (PLS) regression. *Adv. Space Res.* 41, 1755–1763. doi: 10.1016/j.asr.2008.02.015
- Liang, L., Zhang, L., Lin, H., Li, C., and Yang, M. (2013). Estimating canopy leaf water content in wheat based on derivative spectra. *Sci. Agric. Sinica* 46, 18–29. doi: 10.3864/j.issn.0578-1752.2013.01.003
- Liang, X., Ji, H., Wang, P., Rao, Z., and Shen, B. (2010). Study of building quantitative analysis model for chlorophyll in winter wheat with reflective spectrum using MSC-ANN algorithm. *Spectrosc. Spectr. Anal.* 30, 188–191. doi: 10.3964/j.issn.1000-0593201001-0188-04
- Ludwig, B., Vormstein, S., Niebuhr, J., Heinze, S., Marschner, B., and Vohland, M. (2017). Estimation accuracies of near infrared spectroscopy for general soil properties and enzyme activities for two forest sites along three transects. *Geoderma* 288, 37–46. doi: 10.1016/j.geoderma.2016.10.022
- Maestre, F. T. (2004). On the importance of patch attributes, environmental factors and past human impacts as determinants of perennial plant species richness and diversity in Mediterranean semiarid steppes. *Divers. Distrib.* 10, 21–29. doi: 10.1111/j.1472-4642.2004.00057.x
- Mirzaie, M., Darvishzadeh, R., Shakiba, A., Matkan, A., Atzberger, C., and Skidmore, A. (2014). Comparative analysis of different uni- and multi-variate methods for estimation of vegetation water content using hyper-spectral measurements. *Int. J. Appl. Earth Obs. Geoinf.* 26, 1–11. doi: 10.1016/j.jag.2013.04.004
- Miyani, M. A. (2015). Droughts in Asian least developed countries: vulnerability and sustainability. *Weather Clim. Extrem.* 7, 8–23. doi: 10.1016/j.wace.2014.06.003
- Mutanga, O., Skidmore, A. K., Kumar, L., and Ferwerda, J. (2005). Estimating tropical pasture quality at canopy level using band depth analysis with continuum removal in the visible domain. *Int. J. Remote Sens.* 26, 1093–1108. doi: 10.1080/01431160512331326738
- Pu, R., Ge, S., Kelly, N., and Gong, P. (2003). Spectral absorption features as indicators of water status in coast live oak (*Quercus agrifolia*) leaves. *Int. J. Remote Sens.* 24, 1799–1810. doi: 10.1080/01431160210155965
- Ripple, W. (1986). Spectral reflectance relationships to leaf water stress. *Photogramm. Eng. Remote Sensing* 52, 1669–1675.
- Rischbeck, P., Elsayed, S., Mistehele, B., Barmeier, G., Heil, K., and Schmidhalter, U. (2016). Data fusion of spectral, thermal and canopy height parameters for improved yield prediction of drought stressed spring barley. *Eur. J. Agron.* 78, 44–59. doi: 10.1016/j.eja.2016.04.013
- Sandra, E. K., Franklin, E., Barton, I. I., and William, R. W. (2000). Predicting protein content by near infrared reflectance spectroscopy in diverse cereal food products. *J. Near Infr. Spectrosc.* 8, 35–43. doi: 10.1255/jnirs.262
- Savitzky, A., and Golay, M. J. (1964). Smoothing and differentiation of data by simplified least squares procedures. *Anal. Chem.* 36, 1627–1639. doi: 10.1021/ac60214a047
- Sharabian, V. R., Noguchi, N., and Ishi, K. (2014). Significant wavelengths for prediction of winter wheat growth status and grain yield using multivariate analysis. *Eng. Agric. Environ. Food* 7, 14–21. doi: 10.1016/j.eaef.2013.12.003
- Shibayama, M., Takahashi, W., Morinaga, S., and Akiyama, T. (1993). Canopy water deficit detection in paddy rice using a high resolution field spectroradiometer. *Remote Sens. Environ.* 45, 117–126. doi: 10.1016/0034-4257(93)90036-W
- Sims, D. A., and Gamon, J. A. (2003). Estimation of vegetation water content and photosynthetic tissue area from spectral reflectance: a comparison of indices based on liquid water and chlorophyll absorption features. *Remote Sens. Environ.* 84, 526–537. doi: 10.1016/S0034-4257(02)00151-7
- Soares, S. F. C., Gomes, A. A., Araujo, M. C. U., Filho, A. R. G., and Galvão, R. K. H. (2013). The successive projections algorithm. *TrAC Trends Anal. Chem.* 42, 84–98. doi: 10.1016/j.trac.2012.09.006
- Steidle Neto, A. J., Lopes, D. C., Pinto, F. A. C., and Zolnier, S. (2017). Vis/NIR spectroscopy and chemometrics for non-destructive estimation of water and chlorophyll status in sunflower leaves. *Biosyst. Eng.* 155, 124–133. doi: 10.1016/j.biosystemseng.2016.12.008
- Sun, Y., Wang, Y., and Huang, J. (2015). “Estimating canopy water content of wetland vegetation using hyperspectral and multispectral remote sensing data,” in *Proceedings of the SPIE Remote Sensing*, (Bellingham, WA: International Society for Optics and Photonics).
- Tian, Q., Tong, Q., Pu, R., Guo, X., and Zhao, C. (2001). Spectroscopic determination of wheat water status using 1650–1850 nm spectral absorption features. *Int. J. Remote Sensing* 22, 2329–2338. doi: 10.1080/01431160118199
- Tran, T. N., Afanador, N. L., Buydens, L. M. C., and Blanchet, L. (2014). Interpretation of variable importance in partial least squares with significance multivariate correlation (sMC). *Chemometr. Intell. Lab. Syst.* 138, 153–160. doi: 10.1016/j.chemolab.2014.08.005
- Wang, C., Feng, M., Yang, W., Ding, G., Xiao, L., Li, G., et al. (2017). Extraction of sensitive bands for monitoring the winter wheat (*Triticum aestivum*) growth status and yields based on the spectral reflectance. *PLoS One* 12:e0167679. doi: 10.1371/journal.pone.0167679
- Wang, X., Zhao, C., Guo, N., Li, Y., Jian, S., and Yu, K. (2015). Determining the canopy water stress for spring wheat using canopy hyperspectral reflectance data in loess plateau semiarid regions. *Spectrosc. Lett.* 48, 492–498. doi: 10.1080/00387010.2014.909495
- Winterhalter, L., Mistehele, B., Jampatong, S., and Schmidhalter, U. (2011). High throughput phenotyping of canopy water mass and canopy temperature in well-watered and drought stressed tropical maize hybrids in the vegetative stage. *Eur. J. Agron.* 35, 22–32. doi: 10.1016/j.eja.2011.03.004
- Wold, H. (1966). Estimation of principal components and related models by iterative least squares. *Multivariate Anal.* 1, 391–420.

- Wold, S., Martens, H., and Wold, H. (1983). "The multivariate calibration problem in chemistry solved by the PLS method," in *Matrix Pencils*, eds B. Kågström and A. Ruhe (Berlin: Springer), 286–293. doi: 10.1007/BFb0062108
- Xiao, X., Boles, S., Liu, J., Zhuang, D., Frolking, S., Li, C., et al. (2005). Mapping paddy rice agriculture in southern China using multi-temporal MODIS images. *Remote Sens. Environ.* 95, 480–492. doi: 10.1016/j.rse.2004.12.009
- Xie, Y., Feng, M., Wang, C., Yang, W., Sun, H., Yang, C., et al. (2020). Hyperspectral monitor on chlorophyll density in winter wheat under water stress. *Agron. J.* 112, 3667–3676. doi: 10.1002/agj2.20306
- Yang, B., Chen, J., Chen, L., Cao, W., Yao, X., and Zhu, Y. (2015). Estimation model of wheat canopy nitrogen content based on sensitive bands. *Trans. Chinese Soc. Agric. Eng.* 31, 176–182.
- Yebra, M., Dennison, P. E., Chuvieco, E., Riaño, D., Zylstra, P., Hunt, E. R., et al. (2013). A global review of remote sensing of live fuel moisture content for fire danger assessment: moving towards operational products. *Remote Sens. Environ.* 136, 455–468. doi: 10.1016/j.rse.2013.05.029
- Yu, B., Ostland, M., Gong, P., and Pu, R. (1999). Penalized discriminant analysis of in situ hyperspectral data for conifer species recognition. *IEEE Trans. Geosci. Remote Sensing* 37, 2569–2577. doi: 10.1109/36.789651
- Yu, K., Zhao, Y., Liu, Z., Li, X., Liu, F., and He, Y. (2014). Application of visible and near-infrared hyperspectral imaging for detection of defective features in loquat. *Food Bioprocess Technol.* 7, 3077–3087. doi: 10.1007/s11947-014-1357-z
- Yun, Y., Li, H., Wood, L. R. E., Fan, W., Wang, J., Cao, D., et al. (2013). An efficient method of wavelength interval selection based on random frog for multivariate spectral calibration. *Spectrochim. Acta A Mol. Biomol. Spectrosc.* 111, 31–36. doi: 10.1016/j.saa.2013.03.083
- Zadoks, J. C., Chang, T. T., and Konzak, C. F. (1974). A decimal code for the growth stages of cereals. *Weed Res.* 14, 415–421. doi: 10.1111/j.1365-3180.1974.tb01084.x
- Zarco-Tejada, P. J., Rueda, C. A., and Ustin, S. L. (2003). Water content estimation in vegetation with MODIS reflectance data and model inversion methods. *Remote Sens. Environ.* 85, 109–124. doi: 10.1016/S0034-4257(02)00197-9
- Zeng, W., Xu, C., Gang, Z., Wu, J., and Huang, J. (2018). Estimation of sunflower seed yield using partial least squares regression and artificial neural network models. *Pedosphere* 28, 764–774. doi: 10.1016/S1002-0160(17)60336-9
- Zhang, F., and Zhou, G. (2015). Estimation of canopy water content by means of hyperspectral indices based on drought stress gradient experiments of maize in the north plain China. *Remote Sensing* 7, 15203–15223. doi: 10.3390/rs71115203
- Zhang, N., Hong, Y., Qin, Q., and Liu, L. (2013). VSDD: a visible and shortwave infrared drought index for monitoring soil and vegetation moisture based on optical remote sensing. *Int. J. Remote Sensing* 34, 4585–4609. doi: 10.1080/01431161.2013.779046
- Zhang, Q., Li, Q., and Zhang, G. (2012). Rapid determination of leaf water content using VIS/NIR spectroscopy analysis with wavelength selection. *J. Spectrosc.* 27, 93–105. doi: 10.1155/2012/276795
- Zou, X., Zhao, J., Povey, M. J., Holmes, M., and Mao, H. (2010). Variables selection methods in near-infrared spectroscopy. *Anal. Chim. Acta* 667, 14–32. doi: 10.1016/j.aca.2010.03.048

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Sun, Feng, Xiao, Yang, Ding, Wang, Jia, Wu and Zhang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.