Comparison of the Prediction of Winter Wheat Leaf Water Content by Using New Hyperspectral Index and Machine Learning Models

Juanjuan Zhang
Henan Agricultural University https://orcid.org/0000-0001-8460-1601
Wen Zhang
Henan Agricultural University
Shuping Xiong
Henan Agricultural University
Zhaoxiang Song
Adelphi University
Wenzhong Tian
Luoyang of Agricultural and Foresty
Lei Shi
Henan Agricultural University
Xinming Ma (xinmingma@126.com)
Henan Agricultural University

Research

Keywords: winter wheat, leaf water content, spectral index, characteristic band, modeling method, inversion model

Posted Date: November 2nd, 2020

DOI: https://doi.org/10.21203/rs.3.rs-96551/v1

Version of Record: A version of this preprint was published at Plant Methods on March 31st, 2021. See the published version at https://doi.org/10.1186/s13007-021-00737-2.

Abstract

In this study, hyperspectral technology was used to establish the winter wheat leaf water content inversion model to provide technical reference for winter wheat precision irrigation. In a field experiment, seven different wheat varieties for different irrigation times were treated during two consecutive years. The data onto canopy spectral reflectance and leaf water content (LWC) of winter wheat were collected. Five different modeling methods, Spectral index, partial least squares (PLSR), random forest (RF), extreme random tree (ERT) and k-nearest neighbor (KNN) were used to construct LWC estimation models. The results showed that the canopy spectral reflectance was directly proportional to the irrigation times, especially in the near infrared band. As for LWC, the prediction effect of the newly differential spectral index DVI (R1185, R1308) is better than the existing spectral index, and R^2 are 0.78. Because of the large amount of hyperspectral data. The correlation coefficient method (CA) and loading weight (x-Lw) are used to select the water characteristic bands from the full band. The results show that the accuracy of the model based on the characteristic band is not significantly lower than that of the full band. Among these models, the ERT- x-Lw model performs best (R^2 and RMSE of 0.88 and 1.81; 0.84 and 1.62 for calibration and validation, respectively). In addition, the accuracy of LWC estimation model constructed by ERT-x-Lw was better than that of DVI (R1185, R1307). The results provide technical reference and basis for crop water monitoring and diagnosis under similar production conditions.

1. Backgroud

Wheat is the main food crop in North China. Due to the imbalance between precipitation and water demand during the growing period, reasonable irrigation has become a necessary condition for high yield of wheat (Wang et al., 2017). As an important component of plant canopy structure and other biochemical changes, leaf water content is an important indicator to reflect crop water status and indirectly reflect the profit and loss of soil water (José R et al., 2018; William T et al., 2000). Therefore, leaf water content can be used as an important reference index for irrigation decision (Wang et al., 2001). Hyperspectral remote sensing technology has the advantages of fast, economic, and non-destructive. It can be used to monitor. the growth of crops by obtaining the reflectance information on plants. Thus, it is of great significance for precision irrigation and water-saving irrigation to construct the diagnosis model of water status by hyperspectral remote sensing technology

At present, hyperspectral remote sensing technology has been widely used in crop water monitoring scenarios. In some studies, the relationship between wheat leaf water content and hyperspectral was analyzed, and the spectral index was used to estimate leaf water content. For example, the leaf water content of rice, peanut, soybean, and wheat can be well predicted by WI (R900 / R970) / NDVI (R900-R680) / (R900 + 680) (Inoue, 1993). Zhao et al. (2016) study showed that there was a significant correlation between leaf water content and Normalized Difference Water Index (NDWI), Simple Ratio (SR), and Shortwave Infrared Perpendicular Water Stress Index (SPSI). Rapaport et al. (2016) developed Water Balance Index (WABI, (R1500–R531/R1500 + R531)) for monitoring the plant water status in grapevine under field conditions. In recent years, some machine learning methods have been used in the modeling

and analysis of crop growth information and water index of wheat. Zhang et al. (2018) used reflectance data in the range of 859–1640 nm and partial least squares (PLSR), artificial neural network (ANN) and support vector machine (SVN) algorithms to construct models for estimating wheat leaf water content and equivalent water thickness content, respectively, and compared the prediction accuracy with the vegetation index model as the reference model. Bas B et al. (2017) established PLSR, multiple linear regression, SVN and random forest (RF) prediction models to calculate relative water content of wheat through measuring the hyperspectral and relative water content change data of ten different wheat varieties in different time periods indoors. When the full band data are used for modeling, there are some issues such as interference of redundant band information and long operation time caused by large amount of data, and have a certain impact on the accuracy of the model (Chen et al., 2017). Therefore, it is very important to select sensitive bands related to agronomic parameters. Sun et al. (2019) utilized continuous wavelet transform (CWT) to decompose and transform the canopy spectra under different irrigation treatments, and the PLSR model constructed in 2400 nm, 1596 nm and 2397 nm bands can effectively estimate wheat equivalent water thickness (EWT). Bas B et al. (2019) found that the full band was reduced to 14 by the loading weight method, which did not reduce the accuracy of the model. In addition to the above modeling methods, modeling methods such as random forest, extreme random tree and k-nearest neighbor algorithm are also used in the estimation of biomass, nitrogen content, leaf area and SPAD (Yue et al., 2016; George et al., 2018; Ri-Xian et al., 2016). However, these methods are rarely used in the estimation of wheat leaf water content.

Considering the leaf canopy reflectance of ten wheat varieties under different irrigation treatments from 2018 to 2020, five different models was established separately, to provide a reference for hyperspectral monitoring of winter wheat leaf moisture under similar production conditions in the future. The aim of this study were to: (i) analyze the effect of different irrigation times on wheat leaf water content and spectral reflectance; (ii) evaluate the performance of newly spectral indices of leaf water in comparison to existing spectral indices; (iii) evaluate and compare different model algorithms for leaf water status.

2. Experimental Materials And Methods

2.1. Experiment setup

Experiment 1: The experiment was conducted in the field of Luoyang Academy of Agricultural and Forestry Sciences (E:112.47 °, N:34.66° from 2018 to 2019). The experimental variables included different irrigation period treatments and different cultivars of winter wheat. All experiments consisted of a randomized complete block design and the total experiment area is 13 m^2 (2.6 m × 5 m) with two replicates. The soil type is loamy cinnamon soil (Organic Matter: 14.4 $g.kg^{-1}$, N:1.83 $g.kg^{-1}$, P_2O_5:24.6 $mg.kg^{-1}$, K:126.9 $mg.kg^{-1}$). Additional details regarding the experimental design and sample collection time are shown in Table 1.

Table 1
Experiment treatment and sampling periods of the experiments

Exp.no.	Year	Cultivars	Irrigation treatments	Irrigation date	Sampling time and date
		Luomai No.27 Luomai No.34 ZhengmaiNo.16 ZhongyuNo.1211 ZhengmaiNo.136 Zhoumai No.18 Zhengmai No.22	w0	Irrigation during bottom stage/October 18,2018	March 15 (5 days before jointing water)
			w1	Irrigation during bottom stage/October 18,2018	March 30 (10 days after jointing water)
Exp.1	2018–2019			Irrigation during jointing stage/March 20,2019	April 18 (heading stage)
			w2	Irrigation during bottom stage/October 18,2018	April 25 (5 days before filling water)
				Irrigation during jointing stage/March 20,2019	May 10 (10 days after filling water)
				Irrigation during grain filling stager/May 1,2019	May 20 (20 days after flowering)
			W0	Irrigation during bottom stage/October 22,2019	March 13 (5 days before jointing water)
			W1	Irrigation during bottom stage/October 22,2019	March 28 (10 days after jointing water)

Exp.no.	Year	Cultivars	Irrigation treatments	Irrigation date	Sampling time and date
Exp.2	2019–2020	Same as above		Irrigation during jointing stage/March 18,2020	April 15 (heading stage)
			W2	Irrigation during bottom stage/October 22,2019	April 23 (5 days before filling water)
				Irrigation during jointing stage/March 18,2020	May 8 (10 days after grouting)
				Irrigation during grain filling stager/April 25,2020	May 18 (20 days after flowering)

Experiment 2: The location of experiment is the same as experiment 1 (E:112.47 °, N:34.66° from 2019 to 2020). The soil type in this area is loamy cinnamon soil (Organic Matter: 13.2 $g.kg^{-1}$, N:1.05 $g.kg^{-1}$, P_2O_5:18.6 $mg.kg^{-1}$, K:116.3 $mg.kg^{-1}$). The area of the sub region is 10.4 m^2 (2.6 m × 4 m). As for other experimental design, cultivation management measures and sampling period are the same as experiment 1.

2.2. Determination method and Index

2.2.1. Measuring plant hyperspectral data

Plant canopy spectral measurements were obtained with ASD Field Spec Pro spectrometer (Analytical Spectral Devices, Boulder, CO, USA) at 1 m above the canopy (canopy height was approximately 70–80 cm when wheat was mature) with a 25 field of view. Reflectance values with 1 nm spectral resolution in the 350–2500 nm range were collected. To minimize the effects caused by sky and field conditions, spectral measurements were obtained from 10 sites in each plot and averaged into a single spectral sample. For each experiment, measurements were obtained on several different dates that reflected the major growth stages of wheat.

2.2.2. Determination of leaf water

After the canopy spectrum was measured, the wheat plants at corresponding points were collected and all the leaves were extracted. The moisture content of wheat leaves was determined by drying method. The fresh weight of the leaves was weighed with an analytical balance (accuracy of 0.01 g). The leaves

were put into bags and sterilized at 105 ℃ for 30 min. After that, it was dried at 80 ℃ until the weight remained constant weight, and the dry weight of the leaves was weighed at the end. The LWC was calculated using the following formula:

$LWC = (m_f - m_d)/ m_f \times 100\%$ (1)

LWC is the water content of leaf, %; m_f is the fresh weight of leaves, g; m_d is the dry weight of leaf, g.

2.3. Data analysis and Utilization

During the two years of the experiment, a total of 252 wheat samples were collected by the researchers. The test data in 2018–2019 are used as modeling samples (n = 126) and the test data from 2019–2020 are used as validation samples (n = 126). The statistical characteristics of leaf water content in each sample set are shown in Table 2.

Table 2
Statistics characteristics of Wheat Leaf Water Content

Sample sets	Experiment year	Number of samples	Max.	Mini.	Mean	Standard deviation	Variation coefficient
Total	2018–2020	252	88.61	55.87	78.82	4.76	6.02
modeling	2018–2019	126	88.61	55.87	78.80	5.71	7.25
validation	2019–2020	126	83.75	67.54	78.84	3.92	4.98

2.3.1. Characteristic band screening

The hyperspectral band of 350–2500 nm was collected in the experiment (Due to atmospheric noise of hyperspectral spectrometer at 1350–1400 nm, 1800–1950 nm, 2450–2500 nm, the spectral bands in this range are removed), a total of 1901 dimensional spectral bands. If all the spectral bands are used as the input of the model, it will cause "dimension disaster". Therefore, the correlation coefficient method and the load factor method are used to screen the characteristic bands.

(1) Correlation coefficient (CA)

This method determines the characteristic band according to the correlation coefficient between the spectral band and the parameters. The correlation between leaf water content and canopy reflectance under different irrigation treatments was analyzed. The characteristic wavelength was determined by selecting the maximum absolute value of correlation coefficient and the position of wave crest and trough.

(2) x-Loading weight (x-Lw)

The loading weight based on PLSR model can clarify the influence proportion of different dependent variables on the total independent variables, which is of great significance for the rapid screening of characteristic bands. In this study, the peak and trough are selected as the characteristic band.

Based on the feature bands selected by the above two methods as dependent variables, the inversion models of leaf water content of winter wheat were constructed by PLSR, RF, ERT and KNN respectively.

2.3.2. modeling method

(1) Spectral index

Combined with the existing research, it has been proposed that the spectral index related to crop water status has been selected by researchers, as shown in Table 3, and the relationship with leaf water content has also been analyzed. In order to obtain better spectral parameters, the normalized vegetation index (NDVI), the ratio vegetation index (RVI) and the difference vegetation index (DVI) were calculated in the range of 350–2500 nm, which were shown in following formula, and the relationship between them and leaf water content was analyzed, so as to determine the optimal spectral estimation of leaf water content parameters were measured.

$$NDVI(\lambda 1, \lambda 2) = (R\lambda 1 - R\lambda 2)/(R\lambda 1 + R\lambda 2) \quad (2)$$

$$RVI(\lambda 1, \lambda 2) = R\lambda 1/R\lambda 2 \quad (3)$$

$$DVI(\lambda 1, \lambda 2) = R\lambda 1 - R\lambda 2 \quad (4)$$

R λ 1 and R λ 2 are reflectance of any two bands in the range of 350–2500 nm, respectively.

Table 3
The calculation method and reference of spectral index related to crop moisture status.

Spectral index	Definition or equation	Reference
Ratio Index	R1650/R2220	Elvidge et al. (1985)
Normalized differential water index, NDWI	(R820-1240)/(R820+1240)	Gao et al. (1996)
Moisture stress index, MSI	R1600/R820	Hunt et al. (1989)
Maximum water difference index, MDWI	(Rmax1500-1750) -(Rmin1500-1750)/(Rmax1500-1750) +(Rmin1500-1750)	Eitel et al. (2006)
Hyperspectral normalized difference vegetation index , hNDVI	(R900-R680)/(R900-R680)	Rouse et al. (1974)
Water index, WI	R900/R970	Peñuelas et al. (1993)
Simple ratio water index, SRWI	R820/R1200	Zarco-Tejada et al. (2001)
Normalized Difference Infrared Index, NDII	(R820-R1649)/(R820+R1679)	Hardisky et al. (1983)
WI/hNDVI	(R900/R970)/[(R900-R680)/(R900-R680)]	Zhang et al. (2018)
FD730-955	FD730/FD955	Liang et al. (2013)

(2) Partial Least-Squares Regression (PLSR)

Partial least squares regression (PLSR) can effectively establish regression model under the condition of serious multiple correlation of independent variables. In the process of modeling, principal component analysis can be used to judge whether the independent variables can significantly improve the prediction ability. Therefore, this method can explain the multiple autocorrelations among multiple hyperspectral features (Shao J et al., 1993). In this study, PLSR model was carried out in R version 3.5.3 using “pls” package, and "Loo" cross validation was used to determine the number of components.

(3) Random Forest (RF)

Random Forest can establish the relationship between multiple independent variables and a dependent variable, improve the prediction accuracy of the model through many classification trees, and the sample

data can be fully utilized (Victor et al., 2014). To apply this technique, we used python 2.7. With the n tree = 500 used in this study.

(4) Extreme Random Trees (ERT)

Extreme random tree is a top-down method, which is very similar to random forest, but it is different from random forest in two points: firstly, it does not adopt bootstrap sampling replacement strategy, but directly uses the original training samples, in order to reduce the deviation; secondly, it gets the bifurcation value completely randomly, so as to realize the decision tree to bifurcation. The result is smaller and more stable than the random forest (Uddin et al. ,2015). To apply this technique, we used python 2.7.

(5) k-nearest neighbor (KNN)

k-nearest neighbor was proposed by Cover and Hart. It is a classification algorithm based on the proximity of similar samples in the pattern space (Cover T et al., 1967) Euclidean distance is used to measure the similarity between samples, and the larger the distance, the less similarity. In this study, The K-nearest neighbor algorithm was performed in python 2.7. Cross validation method was used to determine K value, K = 3.

2.4. Model validation

The coefficient of determination (R^2), root mean square error (RMSE) and prediction bias (PRD) were used to evaluate the accuracy of the model. The larger R^2 is, and the smaller RMSE is, which indicates that the model has good prediction accuracy. RPD > 1.4, indicating that the prediction ability of the model is acceptable, and the model can be applied.

3. Results

3.1. Effects of irrigation times on leaf water content (LWC) and canopy spectral reflectance of Wheat

In order to explain the effects of different irrigation times on LWC and canopy spectrum, the experiment 1 data were taken as an example. It can be seen from Fig. 1 that with the advance of growth, the LWC of all varieties increased first and then decreased. It reached the highest10 days after jointing water (Mar.30). At the early wheat growth stage, there was little difference in LWC among different irrigation times. However, it decreased rapidly in the late growth period, and the difference increased gradually. The LWC content in w0 was the lowest, which was significantly lower than that in w2 and w1. The results showed that filling water could delay the senescence of leaves and delay the green retention of leaves.

Taking Luomai No.27 as an example, under the same irrigation times, the canopy reflectance of wheat was slightly lower than that of 10 days after jointing water (Mar.30). However, canopy reflectance at 10 days after jointing water was higher than that at 10 days after filling water. Therefore, the canopy reflectance first increased and then decreased with the growth process. With the increase of irrigation times, the canopy reflectance in the visible light region (350–750 nm) did not change significantly. The main reason is that the growth of crops is accelerated after getting enough water after irrigation. Leaf area index, biomass and canopy spectral reflectance increased. It is more obvious after filling water, that is w2 > w1 > w0. Under different irrigation times, the variation of canopy spectral reflectance was basically the same among different wheat varieties.

3.2. Correlation between LWC and spectral index of wheat leaves

Under different irrigation treatments, the correlation between LWC and existing water related vegetation indices was analyzed, such as Ratio Index, NDWI, MSI, MDVI, hNDVI, NDII, WI, SRWI, WI / hNDVI and $FD_{730-955}$. In order to find the best spectral index to estimate the water content of wheat leaves. Based on the leaf spectral reflectance, the relationship between LWC and NDVI, RVI and DVI was analyzed by the researchers based on the samples tested in 2018–2019, and the sensitive band range with larger R^2 value was determined. In NDVI, RVI and DVI bands, some "hot spots" with high correlation coefficient between LWC and RSI, NDS and DVI were identified (Fig. 3). According to Fig. 3, the highest R^2 value was extracted from the hot spot area. NDVI, RVI, and DVI consisting of 1185 nm and 1307 nm perform best for LWC.

Based on the correlation comparison between 13 spectral indices and LWC (Fig. 4). In the modeling set, the spectral indices with R^2 higher than 0.6 are MSI, NDWI, hNDVI, WI, NDII and $FD_{730-955}$ in order; for verification, the data from experiment 2 are calculated, with R^2 ranging from 0.38 to 0.78. The results show that the proposed DVI (R1185, R1307) performs best through high R^2 and low RMSE. The best linear equation of LWC predicted by DVI value is shown in Fig. 5 (R^2 = 0.85, RMSE = 2.25.00; R^2 = 0.78, RMSE = 1.95). Thus the newly proposed indices can be used for accurate estimation of changes in RWC caused by irrigation times in wheat.

3.3. Extraction of characteristic bands of LWC

The correlation between LWC and original spectral reflectance (350–2500 nm) under different irrigation times was shown in Fig. 6. The results showed that the correlation coefficient of LWC ranged from − 0.83 to 0.87, and the maximum positive correlation and negative correlation were 0.86 (618 nm) and − 0.83 (769 nm), respectively. By selecting the maximum absolute value of correlation coefficient and the peak and trough, the optimal wavelengths are 505 nm, 551 nm, 681 nm, 747–831 nm, 989 nm, 1101 nm, 1158 nm, 1445 nm, 1716 nm, 1782 nm, 1978 nm, 2000 nm, 2007 nm, 2038 nm, 2242 nm and 2394 nm.

According to the contribution rate of PLSR model and RMSEP, the number of principal components is determined. When the principal component was 3, RMSEP was 2.34% and explained 96.12% variance (88.86%, 6.03% and 1.23% were explained respectively), as shown in Fig. 7A. Therefore, the best

characteristic band is determined by the peak and trough of the loading weight value of the three principal components (Fig. 7B). According to the above analysis, the optimal band is: 588 nm, 663 nm, 674 nm, 680 nm, 700 nm, 763 nm, 777 nm, 783 nm, 808 nm, 816 nm, 970 nm, 977 nm, 984 nm, 1070 nm, 1072 nm, 1156 nm, 1205 nm, 1246 nm, 1264 nm, 1402 nm, 1445 nm, 1456 nm, 1660 nm, 1678 nm, 1957 nm, 1702 nm, 2221 nm and 2252 nm.

3.4. Comparison of LWC inversion models constructed by different methods

The independent variables were 1901 bands (full band in 350–2500 nm), 100 bands and 28 bands screened by correlation coefficient (CA) and x-Loading weight (x-Lw), respectively, and LWC was the dependent variable. After the PLSR, KNN, RF and extreme ERT models were established (Table 4). Then, compared with the full band model, the accuracy of the model is not significantly reduced with the feature bands extracted from CA and x-Lw as dependent variables. However, the input variables are reduced, which improves the efficiency of the model. Modeling and validation of R^2, RMSE and PRD were considered.

Table 4
Result of modeling sets and validation sets with different modeling methods

Modeling method	Feature band screening method	Number of modeled bands	Calibration			Validation		
			R2	RMSE	RPD	R2	RMSE	RPD
PLSR	Full band	1901	0.87	2.11	2.73	0.82	1.78	2.09
	CA	100	0.86	2.10	2.52	0.84	1.61	2.43
	x-LW	28	0.86	2.61	1.96	0.84	1.73	2.04
RF	Full band	1901	0.88	2.80	1.26	0.83	1.69	2.31
	CA	100	0.90	2.49	1.63	0.81	1.75	2.15
	x-LW	28	0.88	1.57	3.10	0.80	1.86	1.92
ERF	Full band	1901	0.87	1.36	2.18	0.85	1.52	2.30
	CA	100	0.86	1.95	2.48	0.82	1.76	2.34
	x-LW	28	0.88	1.46	3.37	0.84	1.62	2.39
KNN	Full band	1901	0.82	2.10	1.61	0.83	1.61	2.20
	CA	100	0.85	2.00	2.25	0.80	1.79	2.16
	x-LW	28	0.84	2.04	1.80	0.80	1.74	1.83

The performance of the model is as follows: PLSR-CA > ERF-CA > RF-CA > KNN-CA. However, for the 28 characteristic bands selected by x-Lw method, the performance of the model is as follows: ERF-x-Lw > PLSR-x-Lw > RF-x-Lw > KNN-x-Lw. Compared with CA method, the dependent variable is reduced by 98.63%

by x-Lw method, which significantly improves the modeling efficiency. Compared with 12 models constructed by different methods, ERT-x-Lw model showed higher R^2 value and lower RMSE. In the modeling set, $R^{2,}$ RMSE, and RPD were 0.88, 1.46, and 3.37, respectively; in the validation set, R^2, RMSE, and RPD were 0.84, 1.62, and 2.39, respectively (Fig. 8).

4. Discussion

The growth and development of wheat can be directly affected by water deficit. As an important indicator of wheat growth, leaf water content can be monitored by hyperspectral technology. In this study, the spectral reflectance of wheat canopy increased first and then decreased with the advance of growth period. The canopy reflectance was lower before jointing water, and increased 10 days after jointing water, but decreased significantly 10 days after grouting water. The reason may be due to the rapid growth of wheat biomass and leaf area at jointing stage. In the jointing stage, timely irrigation can increase the water absorption of the plant, resulting in the increase of leaf water content, further accelerate the growth of wheat plants, and finally increase the canopy reflectance. However, at the late stage of grain filling, the indexes such as plant withering and yellow abscission, leaf area and leaf water content gradually decreased, which led to the decrease of canopy reflectance (Yang et al., 2019; Guo et al., 2016).

In addition, different irrigation times had significant effect on canopy spectral reflectance. There was no significant difference between w0 treatment and w1, w2 treatment before and after jointing water. Especially in the near infrared region (750–1350 nm), canopy reflectance increased significantly with the increase of irrigation times after jointing water, which was due to the increase of wheat plant height, chlorophyll content and net photosynthetic rate. However, the canopy reflectance of w2 was significantly higher than that of w1 and w0. This indicates that the leaf senescence and photosynthesis time after flowering can be delayed by irrigation at the filling stage; however, without irrigation, wheat plants grow short, the leaves turn yellow and wither ahead of time, lower leaf water content and cell structure changes, which eventually leads to wheat yield reduction (Guo et al., 2016).

The researchers believe that the change of canopy reflectance is caused by the change of LWC. Therefore, the model is constructed by the characteristic response band of leaf water content, which can be used to diagnose and retrieve leaf water content. Due to the difference of experimental conditions in different studies, different bands are selected in spectral index. For example, for the ratio spectral index, the bands at 1300 nm and 1200 nm were selected to predict the water content of wheat leaves, with R^2 of 0.63 (Jiang et al., 2019). In addition, 1391 nm and 1830 nm bands were selected to predict wheat leaf water status (Das et al., 2017). DVI (R185, R1307) was constructed in this study, which has high accuracy. It has consistency and similarity with the waveband, which is selected by Jiang et al. (2019) and Wu et al. (2009). Furthermore, these two bands are selected in this paper, which are in the water

sensitive near-infrared band (Ranjan R et al., 2017). It is superior to the water spectral parameters constructed by previous studies.

Due to inclusion of too many latent variables led to over-fitting (Ecarnot et al., 2013). To improve the modeling accuracy, machine learning and other methods have also been used to model and analyze the moisture index of wheat. Several researchers in the past, based on the grey correlation analysis method, the spectral index with high correlation of leaf water content was selected. These spectral indices were used as independent variables in PLSR and BP neural network models to predict wheat leaf water content, with R^2 of 0.72 and 0.8, respectively. (Jin et al., 2013; Ranjan R et al., 2017). In this study, correlation coefficient method (CA) and x-loading weight (x-Lw) are used to select characteristic bands. Compared with CA method, x-Lw method reduces the number of bands by 72%, which may be due to the concentration of sensitive bands extracted by CA method and the smaller adjacent interval (Zhang Jue et al., 2019). 28 characteristic bands related to leaf water content were selected by x-Lw. Among them, the wavelengths of 663 nm, 674 nm, 680 nm and 700 nm are located in the "red edge" range of the visible light region, which can indirectly diagnose the water status of wheat due to the high reflection of crop leaf structure (Ullah et al., 2014); 1156 nm, 1205 nm, 1246 nm and 1264 nm are related to leaf and canopy cell structure (Gopal Krishna et al., 2019); 1402 nm, 1445 nm, 1456 nm and 1957 nm are related to water absorption band(Wang et al., 20009). This is basically consistent with the water related bands selected by Gopal et al. (2019) Compared with the four modeling methods, the model was constructed through x-Lw-ERT, which makes the best prediction effect of leaf water content. The determination coefficient R^2 and prediction determination coefficient R^2 were 0.88 and 0.84 respectively, and RMSE were 1.46 and 1.62, respectively, which were higher than those of PLSR, RF and KNN models. The reason for this is that ERT has better generalization ability and more stable performance (Randal S. et al., 2017). However, KNN breaks the continuous characteristics of the band because it learns and predicts according to the distance features between different samples (Wu et al.,2017). It is suggested that extreme random tree (ERT) may be a reliable modeling method to improve the modeling accuracy, machine learning and other methods have also been used to model and analyze the moisture index of wheat.

5. Conclusion

Based on the field experiment of different irrigation treatments in two years, the effect of different irrigation times on canopy reflectance spectrum was studied. Five different modeling were compared. The results showed that irrigation in jointing + irrigation in filling period could increase leaf water content, leaf area and biomass, delay plant senescence and increase canopy reflectance. The new model was constructed by DVI (R1185, R1307), which can be used to estimate the LWC of winter wheat. The accuracy of the Extreme random tree model based on the x-loading weight method is better than the former models, and the decision coefficient of modeling and prediction is 0.84 respectively. Thus, both models can be used to estimate the water content of wheat leaves effectively.

Declarations

Ethics approval and consent to participate:

Not applicable

Consent for publication:

This manuscript has not been published elsewhere and is not under consideration by another journal.

Availability of data and materials:

The processed data required to reproduce these findings cannot be shared at this time as the data also forms part of an ongoing study.

Competing interests:

This manuscript is no conflicts of interest to declare.

Funding:

National key RESEARCH and development programs (2016YFD0300609), Key projects of Science and Technology of Henan Province (192102110012) and Modern Agriculture (wheat) Industrial Technology System Project of Henan Province (S2016-01-G04).

Authors' contributions:

Wen Zhang & Zhaoxiang Song: Conceptualization, Methodology, Data curation: Application of statistical, mathematical, computational, or other formal techniques to analyze or synthesize study data, Visualization, Writing- Original draft preparation. Juanjuan Zhang& Xinming Ma Writing - Review & Editing, Funding acquisition, Conceptualization.

Lei Shi &Shuping Xiong: Project administration

Wenzhong Tian: Provision of study materials, laboratory samples, instrumentation

Acknowledgements:

Acknowledgements: This research was supported by National key RESEARCH and development programs (2016YFD0300609), Key projects of Science and Technology of Henan Province

(192102110012) and Modern Agriculture (wheat) Industrial Technology System Project of Henan Province (S2016-01-G04). We are also most grateful to Mr. Zhaoxiao Song and Dr.Juanjuan Zhang.

References

1. Cover T, Hart P. Nearest neighbor pattern classification. IEEE Transactions on Information. 1967;13(1):21–7. https://doi.org/10.1007/978-3-540-74976-9_25.
2. Zhang C, Liu J, Shang J, Cai H. (2018). Capability of crop water content for revealing variability of winter wheat grain yield and soil moisture under limited irrigation. Sci Total Environ, 631–2. https://doi.org/10.1016/j.scitotenv.2018.03.004.
3. Christopher D, Elvidge, Ronald JP, Lyon. Estimation of the vegetation contribution to the 1.65/2.22 micrometer ratio in airborne Thematic-Mapper imagery of the Virginia Range. Nevada International Journal of Remote Sensing. 1985;6(1):75–88. https://dx.doi.org/10.1080/01431168508948425.
4. Gao B-C. (1996). NDWI—a normalized difference water index for remote sensing of vegetation liquid water from space. Remote Sens. Environ, 58, 322–331. http://proceedings.spiedigitallibrary.org/ on 05/12/2015.
5. George P, Sinha K. (2018). Evaluation of the Use of Hyperspectral Vegetation Indices for Estimating Mangrove Leaf Area Index in Middle Andaman Island, India. Remote Sensing Letters, 9(11), https://doi.org/10.1080/2150704X.2018.1508910.
6. Gopal Krishna RN, Sahoo P, Singh, et al. (2019). Comparison of various modelling approaches for water deficit stress monitoring in rice crop through hyperspectral remote sensing. Agric Water Manag, 213. https://doi.org/10.1016/j.agwat.2018.08.029.
7. Guo rui, et al. (2016). Growth metabolism of wheat under drought stress at the jointing-booting stage. Chinese Journal of Plant Ecology, 40(12), 1319–27. http://www.plant-ecology.com/EN/10.17521/cjpe.2016.0107.
8. Hardisky MA, Klemas V, Smart RM. The influence of soil salinity, growth form, and leaf moisture on the spectral radiance of Spartina alterniflora canopies. Photogram Engineer Remote Sensing. 1983;49:77–83.
9. Hunt ER Jr, Rock BN. Detection of changes in leaf water content using near and middle-infrared reflectance. Remote Sens Environ. 1989;30(1):43–54. http://dx.doi.org/10.1016/0034-4257(89)90046-1.
10. Inoue Y, Morinaga S, et al. Non-destructive estimation of water status of intact crop leaves based on spectral reflectance measurements. Japanese Journal of Crop Science. 1993;62:462–9.
11. Jan UH, Eitel, et al. Suitability of existing and novel spectral indices to remotely detect water stress in Populus spp. Forest Ecology Management. 2006;229(1):170–82. https://doi.org/10.1016/j.foreco.2006.03.027.
12. Jiang Jinbao H, Wenjiang, et al. (2010). Using Canopy Hyperspectral Ratio Index to Retrieve Relative Water Content of Wheat Under Yellow Rust Stress. Spectroscopy Spectral Analysis, 30(07), 1939–43.

https://doi.org/10.3964/j.issn.1000-0593(2010)07-1939-05.

13. Jin X, Xu X, et al. (2013). Estimation of Leaf Water Content in Winter Wheat Using GreyRelational Analysis-Partial Least Squares Modeling with Hyperspectral Data. Agron J,105(5). https://doi.org/10.2134/agronj2013.0088.
14. José R, Rodríguez-Pérez C, Ordóñez, Ana B, González-Fernández, et al. (2018). Leaf water content estimation by functional linear regression of field spectroscopy data. Biosys Eng,165, 36–46. https://doi.org/10.1016/j.biosystemseng.2017.08.017.
15. Liang liang, Zhang lianpeng, et al. Estimating Canopy Leaf Water Content in Wheat Based on Derivative Spectral. Scientia Agricultura Sinica. 2013;46(1):18–29. http://doi:org/10.3864/j.issn.0578-1752.2013.01.003.
16. Olson Randal S, Cava W, La M, Zairah, et al. (2018). Data-driven advice for applying machine learning to bioinformatics problems. Pacific Symposium on Biocomputing, 23, 192–203. www.World scientific. Comby 80.82.77.83 on 12/20/17.
17. Peñuelas J, Piñol J, Ogaya R, Filella I. Estimation of plant water concentration by the reflectance water index WI (R900/R970). Int J Remote Sensing. 1997;18:2869–75. http://dx.doi.org/10.1080/014311697217396.
18. Rapaport T, Hochberg U, Cochavi A, et al. (2017). The potential of the spectral 'water balance index' (WABI) for crop irrigation scheduling. New Phytol, 216(3), 741–57. https://www.sci-hub.tf/10.1111/nph.14718.
19. Ranjan R, Sahoo RN, Chopra UK, et al. (2017). Assessment of Water Status in Wheat (Triticum aestivum L.) Using Ground Based Hyperspectral Reflectance. Proceedings of the National Academy of ences, India Section B: Biological ences, 87(2), 377–388. https://doi.org/10.1007/s40011-015-0618-6.
20. Ri-Xian C, Ya-Dong L, Jin-Dong FU. Estimation of Winter Wheat Leaf Nitrogen Accumulation using Machine Learning Algorithm and Visible Spectral. Spectroscopy Spectral Analysis. 2016;36(6):1837–42. https://dx.doi.org/10.3964/j.issn.1000-0593(2016)06-1837-06.
21. Rouse JW, Haas RH, Schell JA, Deering DW. (1979). Monitoring vegetation systems in the Great Plains with ERTS. In: Freden, S.C., Mercanti, E.P., Becker, M, editors, Third Earth Resources Technology Satellite–1 Syposium. Volume I: Technical Presentations. NASA SP-351, NASA, Washington, D.C, pp. 309–317. https://dx.doi.org/citeulike-article-id:12009708.
22. Uddin MT, Azher MU. (2015). Human Activity Recognition from Wearable Sensors using Extremely Randomized Trees. In Proceedings of the IEEE International Conference on Dhaka, Dhaka, Bangladesh, 21–23. https://doi.org/10.1109/ICEEICT.2015.7307384.
23. Saleem Ullah, Andrew K, et al. Retrieval of leaf water content spanning the visible to thermal infrared spectra. ISPRSJ Photogramm Remote Sens. 2014;93:56–64. https://doi.org/10.1016/j.isprsjprs.2014.04.005.
24. Shao J. Linear model selection by cross-validation. J Am Stat Assoc. 1993;88:486–94. https://doi.org/10.1080/01621459.1993.10476299.

25. Sun Qian GU, XiaoHe SUN, Lin WANG, Miao, et al. Spectral Response Analysis of Canopy Water Content of Winter Wheat Under Different Irrigation Conditions. Scientia Agricultura Sinica. 2019;52(14):2425–35. 10.11869/j.issn.100-8551.2014.01.0309.
26. Victor R-G, Mendes MP, Garcia-Soldado MJ, et al. (2014). Predictive modeling of groundwater nitrate pollution using Random Forest and multisource variables related to intrinsic and specific vulnerability: A case study in an agricultural setting (Southern Spain). Sci Total Environ, 476–7. https://doi.org/10.1016/j.scitotenv.2014.01.001.
27. Wang Jihua Z, Chunjiang, et al. (2001). Effect of Soil Water Content on the Wheat Leaf Water Content and the Physiological Function. Journal of Triticeae Crops, (4), 42–27.
28. William T, Pockman JS, Sperry. Vulnerability to xylem cavitation and the distribution of Sonoran Desert vegetation. American Journal Botany. 2000;87(9):1287–99. https://doi.org/10.2307/2656722.
29. Wu C, Niu Z, Tang Q, et al. Predicting vegetation water content in wheat using normalized difference water indices derived from ground measurements. J Plant Res. 2009;122(3):317–26. https://doi.org/10.1007/s10265-009-0215-y.
30. Xueyan Wu, Shuihua W, Zhang Yudon. (2017). Survey on theory and application of k-Nearest-Neighbors algorithm. Computer Engineering Applications,53(21), 1–7. http://cea.ceaj.org/CN/Y2017/V53/I21/1.
31. Yang chen, et al. Hyperspectral monitoring of aboveground dry biomass of winter wheat under different irrigation treatments. Chin J Ecol. 2019;38(06):1767–73. http://www.cje.net.cn/CN/Y2019/V38/I6/1767.
32. Yue Jibo Y, Guijun, Feng Haikuan. Comparative of remote sensing estimation models of winter wheat biomass based on random forest algorithm. Transactions of the Chinese Society of Agricultural Engineering. 2016;32(18):175–82.
33. Zarco-Tejada PJ, Ustin SL. (2001). Modeling canopy water content for carbon estimates from MODIS data at land EOS validation sites. Proceedings of the IEEE 2001 International Geoscience and Remote Sensing Symposium, 342–344. https://doi.org/10.1109/IGARSS.2001.976152.
34. Shiqin, Wang, et al. Relationship between land-use and sources and fate of nitrate in groundwater in a typical recharge area. of the North China Plain. 2017;609:607–20. https://doi.org/10.1016/j.scitotenv.2017.07.176.
35. Zhao S, Wang Q, Yao Y, Du S, Zhang C, Li J, Zhao J. Estimating and validating wheat leaf water content with three MODIS spectral indexes: A case study in Ning xia Plain, China. Journal of Agricultural Science Technology. 2016;18:387–98.
36. Zhang Jue T, Haiqing Z, Zhiyu Z, Lina Z, Jing, Li Fei. Moisture content dectection in silage maize raw material based on hyperspectrum and improved discrete particle swarm. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE). 2019;35(1):285–93.

Figures

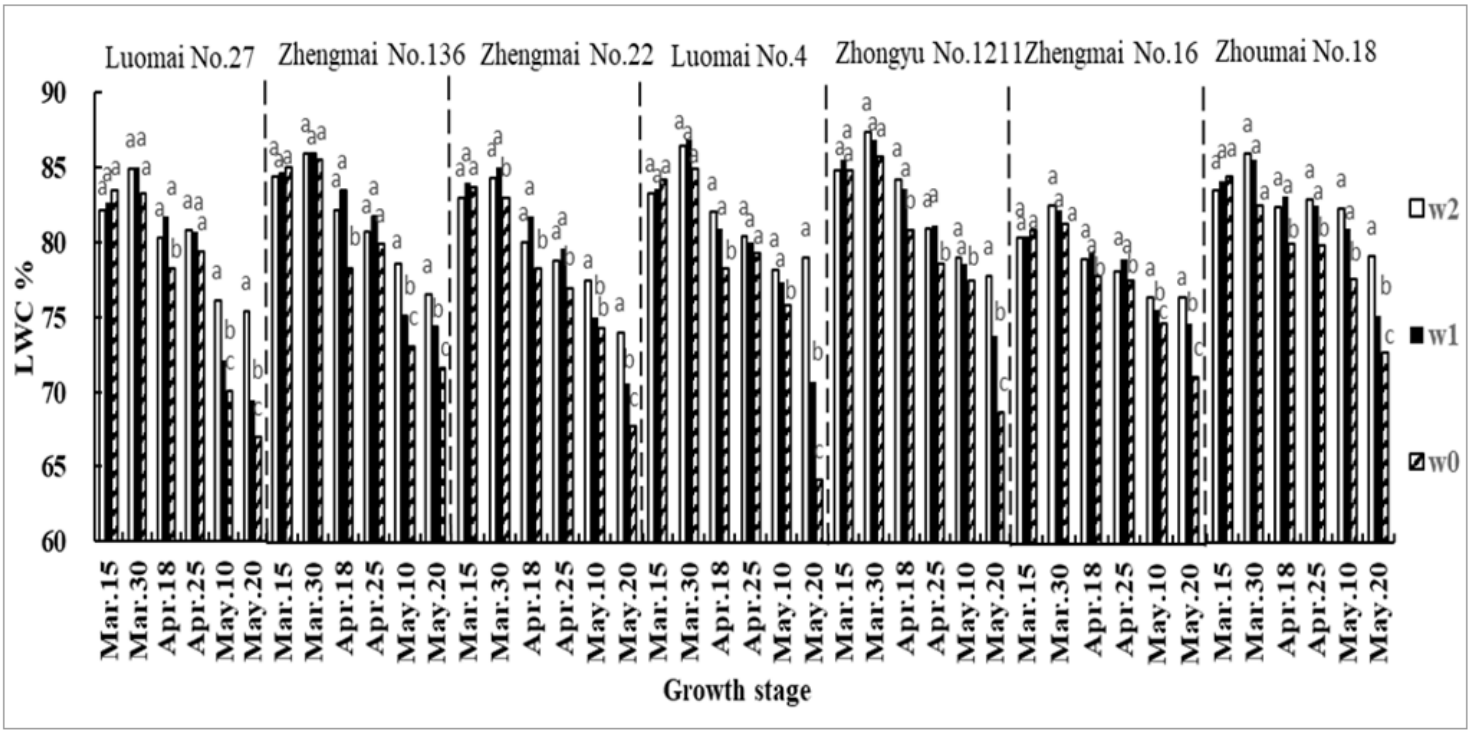

Figure 1

Effects of different irrigation times on water content of wheat leaves Note: the letters on the column indicate the significant difference among different treatments in the same period (P < 0.05)

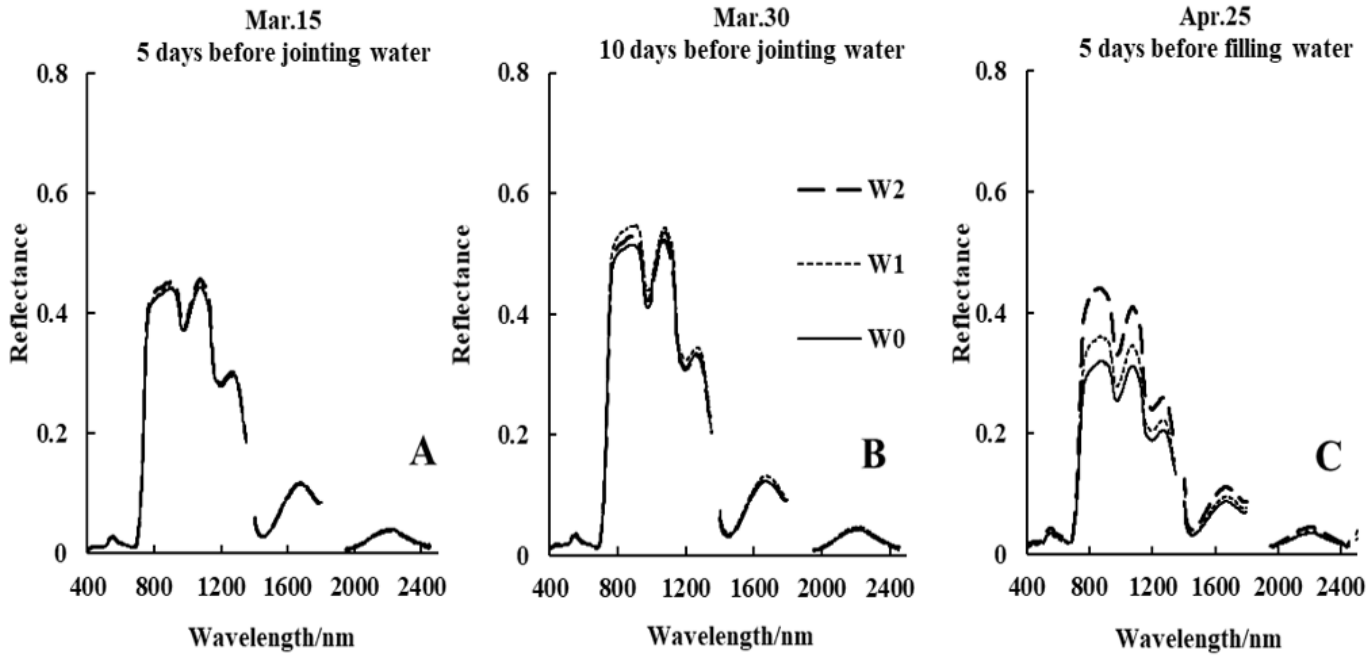

Figure 2

Effects of different irrigation times on canopy spectra before jointing water(A),10 days after jointing water (B)and grouting water(C)

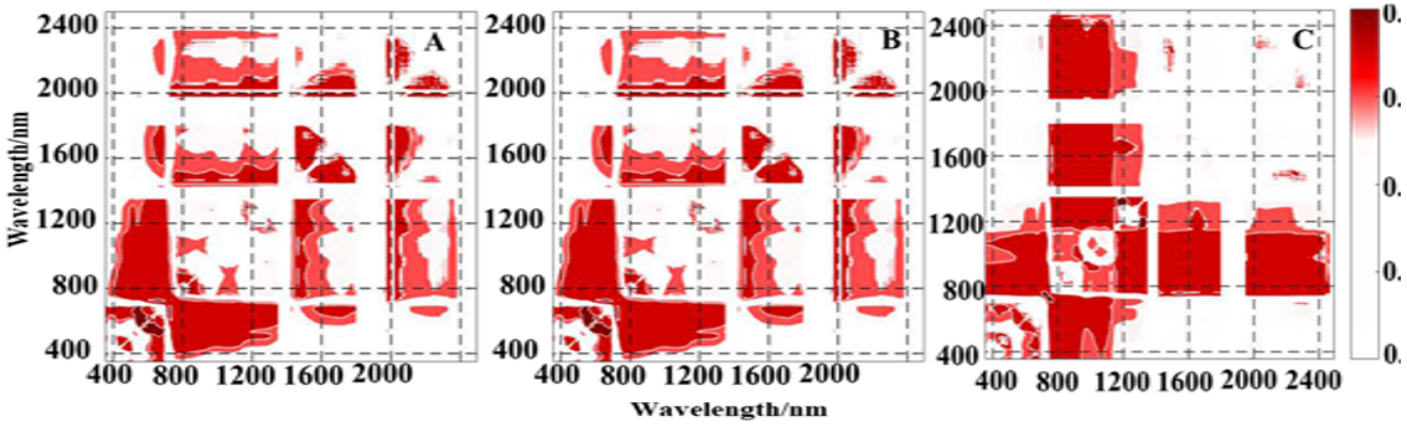

Figure 3

The contour maps of the coefficient of determination between LWC and RVI(A), NDVI(B), and DVI(C)values based on normalized band depth index

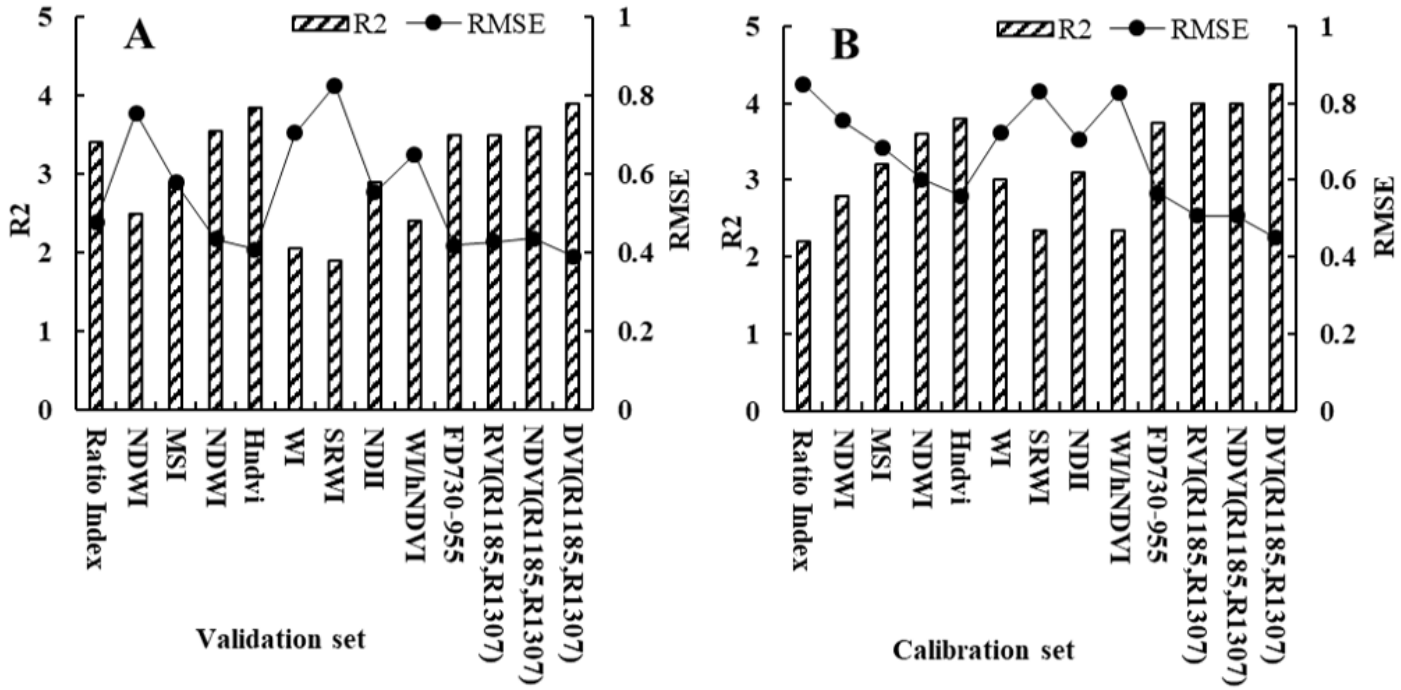

Figure 4

Correlation between spectral index and LWC in calibration set(A) and verification machine(B)

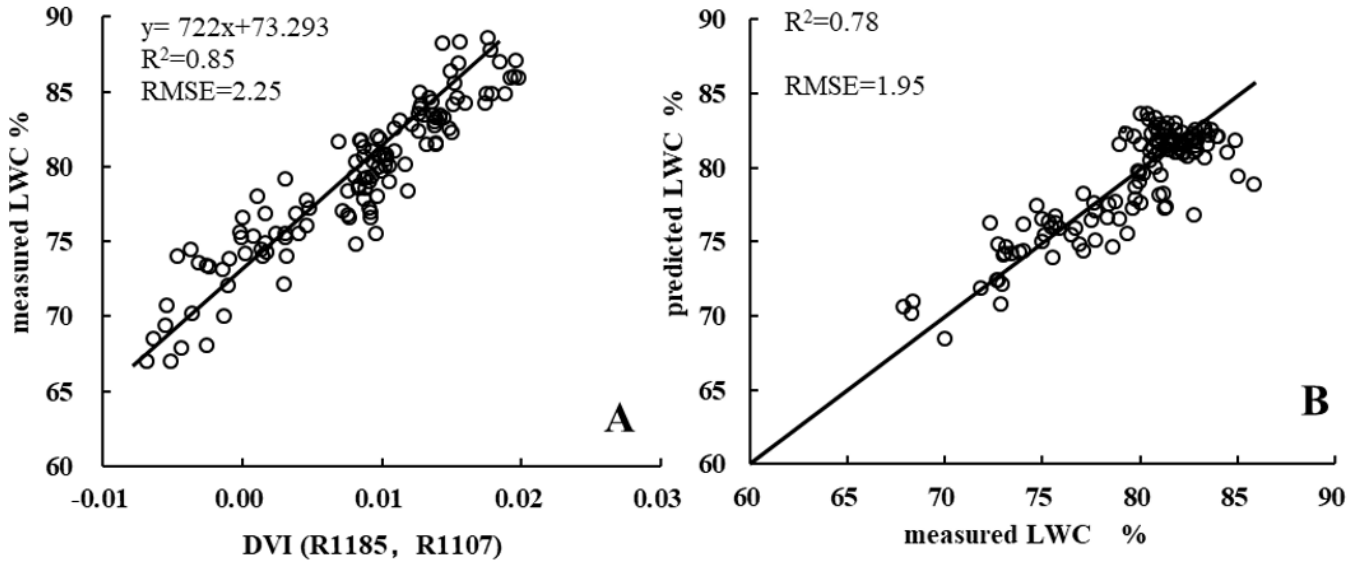

Figure 5

Quadratic regression model based on DVI (R1185, R1107) to predict wheat LWC (A calibration set, B verification machine)

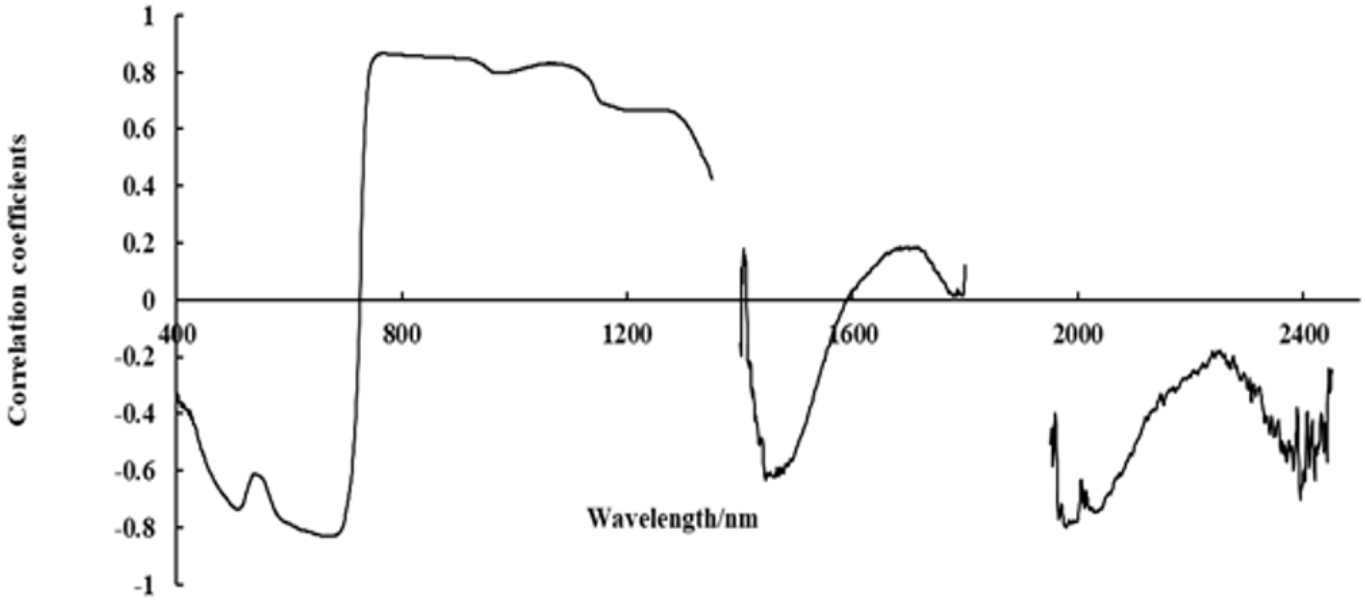

Figure 6

Correlation between canopy reflectance and leaf water content under different irrigation times

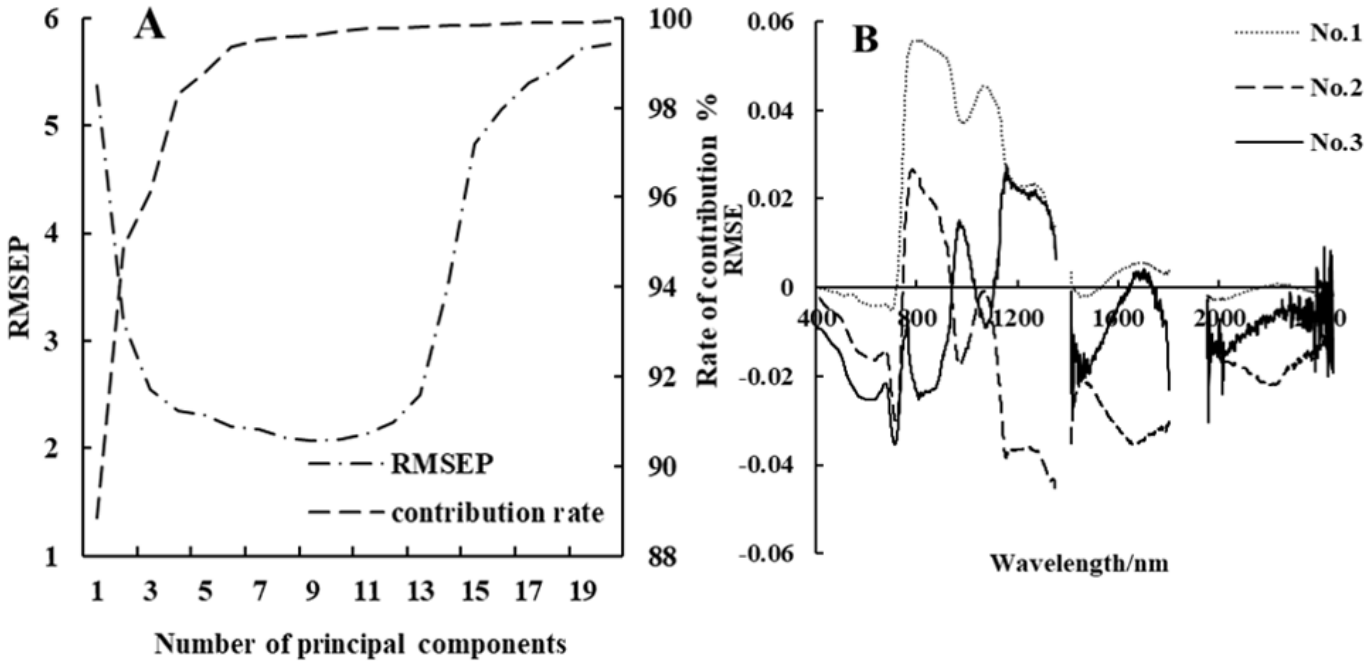

Figure 7

Principal component contribution rate(A) and load value calculated(B) by PLSR regression

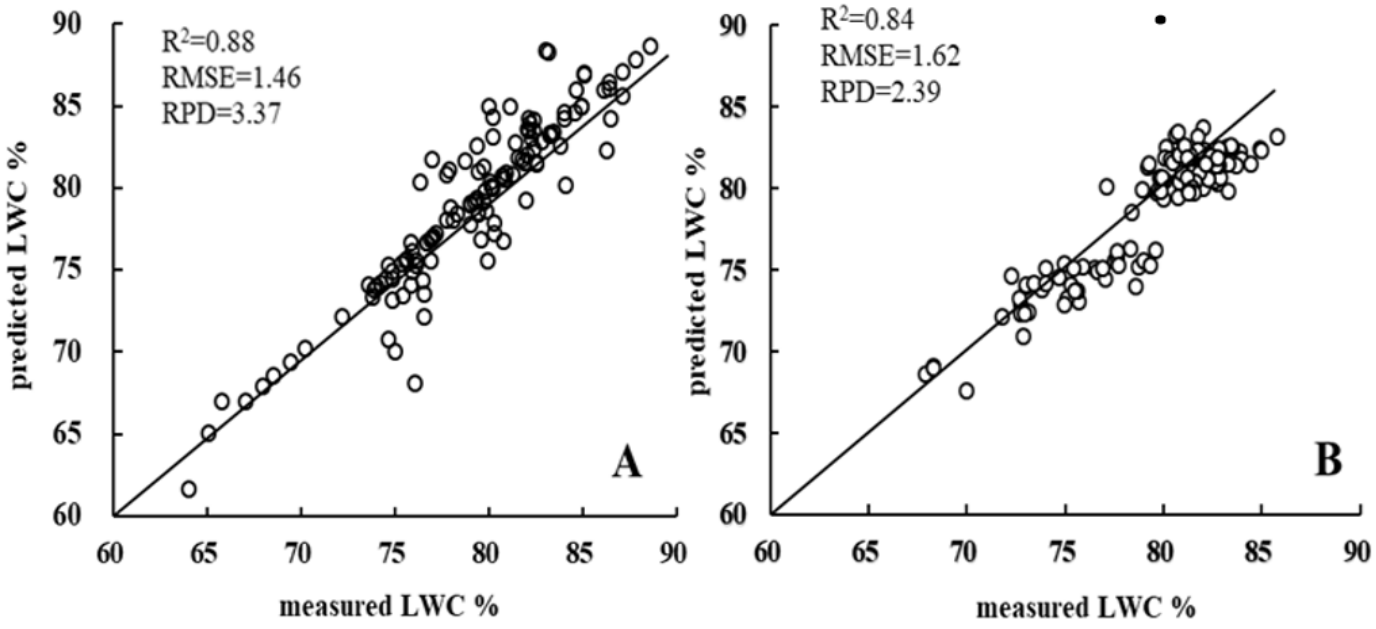

Figure 8

Predicted versus observed LWC values for ERT-x-Lw regression models based on (A)calibration set and (B) verification machine