



## OPEN ACCESS

## EDITED BY

S. Qiu,  
Chinese Academy of Agricultural Sciences,  
China

## REVIEWED BY

Jinshun Bai,  
Chinese Academy of Agricultural Sciences  
(CAAS), China  
Jianlin Shen,  
Chinese Academy of Sciences (CAS), China  
Xiaokun Li,  
Huazhong Agricultural University, China

## \*CORRESPONDENCE

Ning Cao  
✉ cao\_ning@jlu.edu.cn

†These authors have contributed equally to  
this work

RECEIVED 14 March 2023

ACCEPTED 13 April 2023

PUBLISHED 25 May 2023

## CITATION

Zhang Y, Wang T, Li Z, Wang T and  
Cao N (2023) Based on machine  
learning algorithms for estimating  
leaf phosphorus concentration of rice  
using optimized spectral indices and  
continuous wavelet transform.  
*Front. Plant Sci.* 14:1185915.  
doi: 10.3389/fpls.2023.1185915

## COPYRIGHT

© 2023 Zhang, Wang, Li, Wang and Cao.  
This is an open-access article distributed  
under the terms of the [Creative Commons  
Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). 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.

# Based on machine learning algorithms for estimating leaf phosphorus concentration of rice using optimized spectral indices and continuous wavelet transform

Yi Zhang<sup>†</sup>, Teng Wang<sup>†</sup>, Zheng Li, Tianli Wang and Ning Cao\*

College of Plant Science, Jilin University, Changchun, China

Remotely estimating leaf phosphorus concentration (LPC) is crucial for fertilization management, crop growth monitoring, and the development of precision agricultural strategy. This study aimed to explore the best prediction model for the LPC of rice (*Oryza sativa* L.) using machine learning algorithms fed with full-band (OR), spectral indices (SIs), and wavelet features. To obtain the LPC and leaf spectra reflectance, the pot experiments with four phosphorus (P) treatments and two rice cultivars were carried out in a greenhouse in 2020–2021. The results indicated that P deficiency increased leaf reflectance in the visible region (350–750 nm) and decreased the reflectance in the near-infrared (NIR, 750–1350 nm) regions compared to the P-sufficient treatment. Difference spectral index (DSI) composed of 1080 nm and 1070 nm showed the best performance for LPC estimation in calibration ( $R^2 = 0.54$ ) and validation ( $R^2 = 0.55$ ). To filter and denoise spectral data effectively, continuous wavelet transform (CWT) of the original spectrum was used to improve the accuracy of prediction. The model based on Mexican Hat (Mexh) wavelet function (1680 nm, Scale 6) demonstrated the best performance with the calibration  $R^2$  of 0.58, validation  $R^2$  of 0.56 and RMSE of  $0.61 \text{ mg g}^{-1}$ . In machine learning, random forest (RF) had the best model accuracy in OR, SIs, CWT, and SIs + CWT compared with other four algorithms. The SIs and CWT coupling with the RF algorithm had the best results of model validation, the  $R^2$  was 0.73 and the RMSE was  $0.50 \text{ mg g}^{-1}$ , followed by CWT ( $R^2 = 0.71$ , RMSE =  $0.51 \text{ mg g}^{-1}$ ), OR ( $R^2 = 0.66$ , RMSE =  $0.60 \text{ mg g}^{-1}$ ), and SIs ( $R^2 = 0.57$ , RMSE =  $0.64 \text{ mg g}^{-1}$ ). Compared with the best performing SIs based on the linear regression models, the RF algorithm combining SIs and CWT improved the prediction of LPC with  $R^2$  increased by 32%. Our results provide a valuable reference for spectral monitoring of rice LPC under different soil P-supplying levels in a large scale.

## KEYWORDS

continuous wavelet transform, leaf phosphorus concentration, machine learning, rice, spectral indices

## 1 Introduction

The fast growth of the global demand for agricultural production is increasing the chemical fertilizer application (Tilman et al., 2011; Mueller et al., 2012; Demay et al., 2023). In intensive cropping systems, phosphorus (P) fertilizer as a nonrenewable resource requires more precise management because of its different effects on yield and the environment (Sharpley and Withers, 1994; Tilman et al., 2002; MacDonald et al., 2011; Townsend and Porder, 2012). However, limiting information for regional soil P fertility status restricts the rational P management strategy development. Globally, imbalance P application within agricultural regions is increasing soil degradation with deficit application, or environmental pollution with an excessive application (Bennett et al., 2001; Carpenter, 2008; MacDonald et al., 2011; Bindraban et al., 2020). The lack of an effective method for non-destructive measurements *in situ* of P limits the holistic understanding of P requirement for crop and soil P-supplying level in a large scale. Therefore, non-destructive measurements are essential for devising precision agricultural policies and the best management practices to optimize the application of P fertilizer to improve grain yield.

As the most promising technology, hyperspectral technology can acquire variation in crop nutrient content timely and nondestructively (Takebe et al., 1990; Hansen and Schjoerring, 2003; Feng et al., 2008; Pimstein et al., 2011). Many studies have documented that leaf or canopy spectral reflectance data can be used to evaluate the nitrogen (N) status of crops, and the N deficiency influences the spectral reflectance of crops in visible region and NIR regions (Daughtry et al., 2000; Zhao et al., 2003; Xue et al., 2004; Zhao et al., 2005; Tian et al., 2014; Zhao et al., 2018). The spectral reflectance of crop leaves is known to be correlated with P status (Milton et al., 1991; Osborne et al., 2002; Yaryura et al., 2009; Pimstein et al., 2011; Mahajan et al., 2017). Generally, P deficiency promoted the visible accumulation of anthocyanin (AnC) (Jiang et al., 2007). AnC is a water-soluble pigment, which shows different colors with the change of soil P availability, and further changes the spectral reflectance of the plant (Viña and Gitelson, 2011). Compared with the spectral study of N, however, studies on crop P content are insufficient. Hence, the development of a leaf phosphorus concentration (LPC) diagnostic model by spectral reflectance technology plays an important role in precision P fertilizer management.

The spectral indices (SIs) are widely used to estimate the P concentration of crops at local, and regional scales (Mahajan et al., 2014; Mahajan et al., 2017). Many studies have shown that the SIs can be used to estimate the P concentration of wheat (Mahajan et al., 2014), litchi (Li et al., 2018a), and rice (Mahajan et al., 2017). However, the literature has shown that the relationship between the P concentration and SIs is still inconsistent. In previous studies, Mahajan et al. (2014) proposed a new normalized difference vegetation index (NDVI) of two band combinations (1080 nm, 1460 nm) for P prediction, and the correlation coefficient ( $R^2$ ) was 0.42. Mahajan et al. (2017) found that NDVI with bands at 1260 nm and 670 nm has a higher prediction accuracy of canopy P status ( $r =$

0.67,  $p < 0.01$ ). Li et al. (2018a) indicated linear regression model constructed by using the ratio of reflectance difference index ( $RRDI_{1465, 1605, 1665}$ ) can well predict leaf P content of litchi ( $R^2_{cv} = 0.95$ ,  $RMSE_{cv} = 0.01$ ), and the selection of sensitive bands and estimation accuracy of LPC were significantly affected by the interrelationship among LPC, pigments, and N. To ensure the performance of SIs, therefore, it is important to select the sensitive bands and suitable algorithms to create the optimized SIs models. To develop optimized SIs and improve the model accuracy of vegetation properties, considering all suitable combinations of the band based on established index formulations are widely used (Mariotto et al., 2013; Rivera et al., 2014; Yang et al., 2021b). However, due to the influence of many factors, such as different crops, growing seasons, and external environment, there is a complex nonlinear relationship between P concentration and spectral characteristics. Thus, it is still unclear whether the SIs can estimate the plant properties with high estimation accuracy (Verrelst et al., 2015; Verrelst et al., 2019). Additionally, to capture accurate and effective spectral information, continuous wavelet analysis (CWA) is becoming a promising tool for estimating biochemical constituent concentrations from leaf reflectance spectra (Cheng et al., 2011). The continuous wavelet transform (CWT) decomposes the leaf reflectance spectra into several scale components, which are composed of wavelet features as a function of wavelength and scale (Cheng et al., 2011; Li et al., 2018b). CWT has been widely used for estimating the leaf water content and nitrogen status, and was proven to be effective and have higher model accuracy compared to SIs (Cheng et al., 2011; Li et al., 2018b; Li et al., 2022).

In recent years, for modeling and analyzing crop growth and vegetation parameters, machine learning has been widely applied (Zhai et al., 2013; Heckmann et al., 2017; Wang et al., 2018; Han et al., 2019). A partial least square regression (PLSR) model was established by Chen et al. (2002) for estimating P concentration in sugarcane leaves, and the  $R^2$  was 0.99. Gao et al. (2019) used the support vector machine (SVM), random forest (RF), and artificial neural network (ANN) algorithms to create models for forage P content estimation, and the SVM model performed best. In addition, the coupling of SIs with machine learning algorithms can improve the accuracy obviously in crop parameter estimation, such as leaf water content (Zhang et al., 2021), and above-ground biomass (Wang et al., 2016; Yang et al., 2021b). The input variables of machine learning can be optimized by using the SIs, such as dimension and multicollinearity reduction (Yang et al., 2021b). However, the previous studies showed the different performances of various models. Therefore, selecting suitable input variables to feed machine learning algorithms is critical for estimating rice LPC.

Previous studies have investigated the full spectrum and feature bands as input variables for machine learning algorithms to estimate the crop LPC. However, limited studies reported the sensitive bands, optimized SIs, and spectral transformation techniques coupling with machine learning algorithms in the estimation of rice LPC. To improve modeling precision and dimension reduction for rice LPC, therefore, there is a need to combine spectral index, wavelet analysis, and machine learning algorithms. In this study, we applied the rice leaf reflectance under

different P application rates and explored the optimal prediction model for LPC by using five machine learning algorithms fed with full-band, spectral indices, and continuous wavelet features. This research aimed to provide a basic reference for LPC spectral monitoring of rice under different soil P-supplying levels in a large scale. The specific objectives were (1) to evaluate the performance of SIs and CWT of original spectrum in estimating rice LPC and (2) to compare the full-band, optimized results of SIs and CWT coupled with five machine learning algorithms in predicting rice LPC.

## 2 Materials and methods

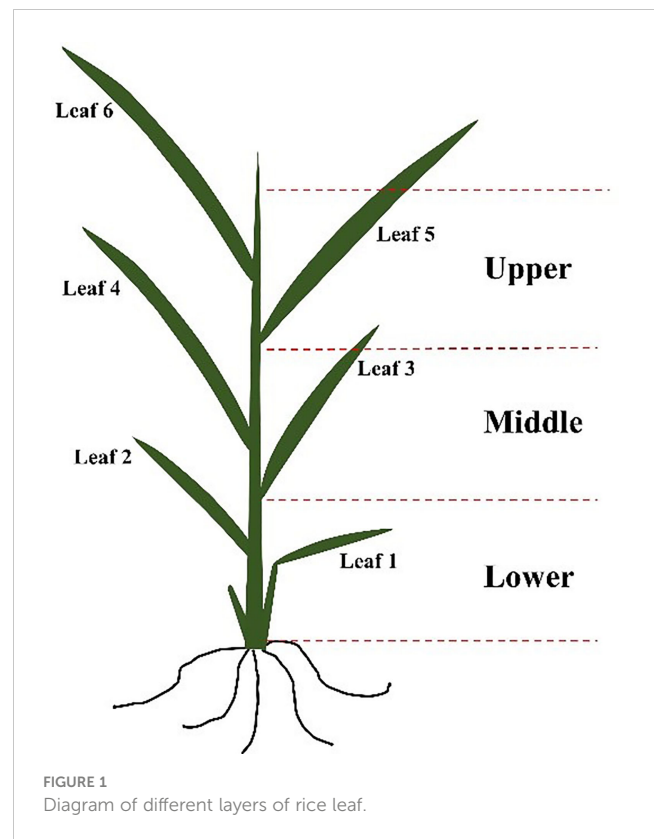
### 2.1 Experimental design and growth conditions

The pot experiments of rice were carried out in the greenhouse of Inner Mongolia Agricultural University (111°42' E, 40°48' N) during 2020-2021 in Hohhot, Inner Mongolia, China. The air temperature and humidity in the greenhouse were maintained at 25-28 °C and 60-70%. The photoperiod was 12h light and 12h dark per day (LD 12:12) in white fluorescent light (about 150  $\mu\text{mol}/\text{m}^2/\text{s}$ ).

Pot experiments with four P treatments, which are 0, 20, 40, and 80 kg  $\text{P}_2\text{O}_5 \text{ ha}^{-1}$ , respectively (P0, P1, P2, and P3), and two rice cultivars (Longjing 31 and Wuyoudao 4) were conducted. The pot size was 40 × 20 × 20 cm. The experiment was a randomized complete block design with ten replicates. Soil pH, organic matter, total N, total P, available N, and available P were 7.8, 17.1 g  $\text{kg}^{-1}$ , 0.67 g  $\text{kg}^{-1}$ , 0.31 g  $\text{kg}^{-1}$ , 29.8 mg  $\text{kg}^{-1}$ , and 8.9 mg  $\text{kg}^{-1}$ , respectively. Phosphorus fertilizer applications such as monopotassium phosphate were performed before sowing.

### 2.2 Spectral data collection

The spectral reflectance of rice leaves in the upper, middle, and lower layers (Figure 1) were measured at the critical stage of P nutrition (tillering stage with six leaves) using a ground object spectrometer PSR+3500 (Spectral Evolution Inc., Lawrence, MA, USA). This instrument records reflectance between 350-2500 nm with a sampling interval of 1 nm and spectral resolution of 3 nm@700 nm, 8 nm@1500 nm, and 6 nm@2100 nm respectively. Output data were composed of the reflectance of 2151 spectral channels. Before measuring, flip the leaf clip and calibrating with the whiteboard in the pistol grip. Put the leaf into the leaf clip during measurement. The observation angle was 90°, the area of view was about 0.5  $\text{cm}^2$  and all spectral measurements were measured between 11:30 a.m. to 2:00 p.m. on clear sunny days (Darvishzadeh et al., 2008; An et al., 2020). Each leaf was measured with three replicates, and the average value was taken as the spectral reflectance of the rice leaf.



### 2.3 Plant sampling and LPC measurements

After spectral data collection, rice leaves in the same layer were collected for measuring leaf dry mass and LPC. All plant samples were oven-dried at 105 °C for 0.5 h and then dried at 75 °C until a constant weight was reached for biomass measurements. After calculating the biomass, the samples were ground to a fine powder (0.25 mm sieve) and the molybdate-blue colorimetric method was used for determining the LPC (mg  $\text{g}^{-1}$ ) of each sample (Murphy and Riley, 1962).

A total number of 456 rice leaf samples were collected during the 2 years of the experiment. The pooled data were divided randomly into an independent calibration dataset (70% of the pooled data, 319 samples) and a validation dataset (30% of the pooled data, 137 samples). The calibration dataset was used to establish the models, and the validation dataset was used to validate the models.

### 2.4 Spectral indices and continuous wavelet transform analysis

#### 2.4.1 Spectral indices (SIs)

A large number of SIs have been created to estimate the nutrition parameters of crops. Especially the two-band SIs including ratio spectral index (RSI), difference spectral index (DSI), and normalized differential spectral index (NDSI) are the most classic SIs algorithms (Jordan, 1969; Rouse et al, 1974; Tucker, 1979). The calculation

formula of these SIs are shown as follows.

$$RSI = \frac{R_{\lambda_1}}{R_{\lambda_2}} \quad (1)$$

$$DSI = R_{\lambda_1} - R_{\lambda_2} \quad (2)$$

$$NDSI = \frac{R_{\lambda_1} - R_{\lambda_2}}{R_{\lambda_1} + R_{\lambda_2}} \quad (3)$$

$R_{\lambda_1}$  and  $R_{\lambda_2}$  represent the reflectance of any two single bands in the range of 350–2500 nm, respectively, and a self-developed code in MATLAB R2021b software (The MathWorks Inc., Massachusetts, USA) was used to select the bands. The relationships between rice LPC and three SIs were analyzed for determining the optimal estimation model of LPC.

## 2.4.2 Continuous wavelet transform (CWT) analysis

CWT is a signal analysis and processing tool which can realize multi-frequency and multi-scale decomposition of spectral information. It decomposes the signal into a series of wavelet functions obtained by the same wavelet basis function. The component in each scale can be directly compared with the input data of spectral reflectivity. At the same time, more valuable spectral information can be obtained (Rivard et al., 2008; Cheng et al., 2011). Usually, choosing the appropriate wavelet function is the primary task of the transform process. In this study, fifteen wavelet functions in MATLAB R2021b were used and ten scales were calculated for each wavelet function. The Mexican Hat (Mexh) wavelet functions smooth the spectral data with the Gaussian function and then calculate the second derivative. It can filter and denoise spectral data effectively (Singh et al., 2013). According to the results of  $R^2$  between wavelet functions and the LPC of rice, the transformation effect based on the Mexh function produced the highest model accuracy. Therefore, Mexh was selected as the basic function of CWT in this study and was realized in MATLAB R2021b.

## 2.5 Machine learning algorithms

### 2.5.1 Partial least squares regression (PLSR)

PLSR is that the eigenvalues are reduced to a small group of unrelated features through a certain operation process, and the least square regression method is performed on these features, which can solve the problems of multi-collinearity between features and feature dimension greater than the sample numbers (Ramadan et al., 2005). In this study, the PLSR program was applied using Python (version 3.7.0, The Python Software Foundation, USA) software, and the parameters were the default settings.

### 2.5.2 Least absolute shrinkage and selection operator (LASSO)

LASSO is a biased estimation algorithm for solving multiple collinear problems (Tibshirani, 2011). Its basic principle is to add L1 regularization constraints to the parameters based on

conventional linear regression, to simplify the refined model and prevent over-fitting of the model. The LASSO program was conducted using Python software, and the selection parameter was set to 'cyclic', which means that the update of the regression coefficient in each iteration is based on the last operation.

### 2.5.3 Random forest (RF)

The RF regression model is based on the decision tree, random attributes are introduced to construct an integrated evaluator (Breiman, 2001). Each decision tree learns independently and predicts independently. The prediction results are determined by averaging over all the trees (Liaw and Wiener, 2002; Hao et al., 2015; Yang et al., 2021b). In this paper, the RF program was applied using Python software, and the parameters were the default settings.

### 2.5.4 Support vector machine (SVM)

SVM is based on the structural risk minimization principle and statistical learning theory, which is suitable for machine learning of small samples (Cortes and Vapnik, 1995). In this study, the kernel function selected when using SVM is the radial basis kernel function (Radial Basis Function), which is suitable for solving partial nonlinear problems. The SVM program was applied using Python software, and the parameters were the default settings in this study.

### 2.5.5 Back propagation artificial neural network (BPANN)

As an artificial intelligence method, BPANN uses an error backpropagation algorithm to obtain the multilayer feedforward neural network (Ramadan et al., 2005). It has a strong nonlinear fitting ability and is widely used. BPANN program was conducted using Python software, and the parameters were the default settings.

The LPC of rice was taken as the dependent variable. The independent variables were the original full band (all 2151 bands ranging from 350–2500 nm, OR), optimized SIs (10 best features), optimized CWT (10 best features), and the combination of SIs and CWT (20 input features, SIs + CWT), respectively. And then the PLSR, LASSO, RF, SVM, and BPANN models were established. A flowchart of the rice LPC estimation model construction is shown in Figure 2.

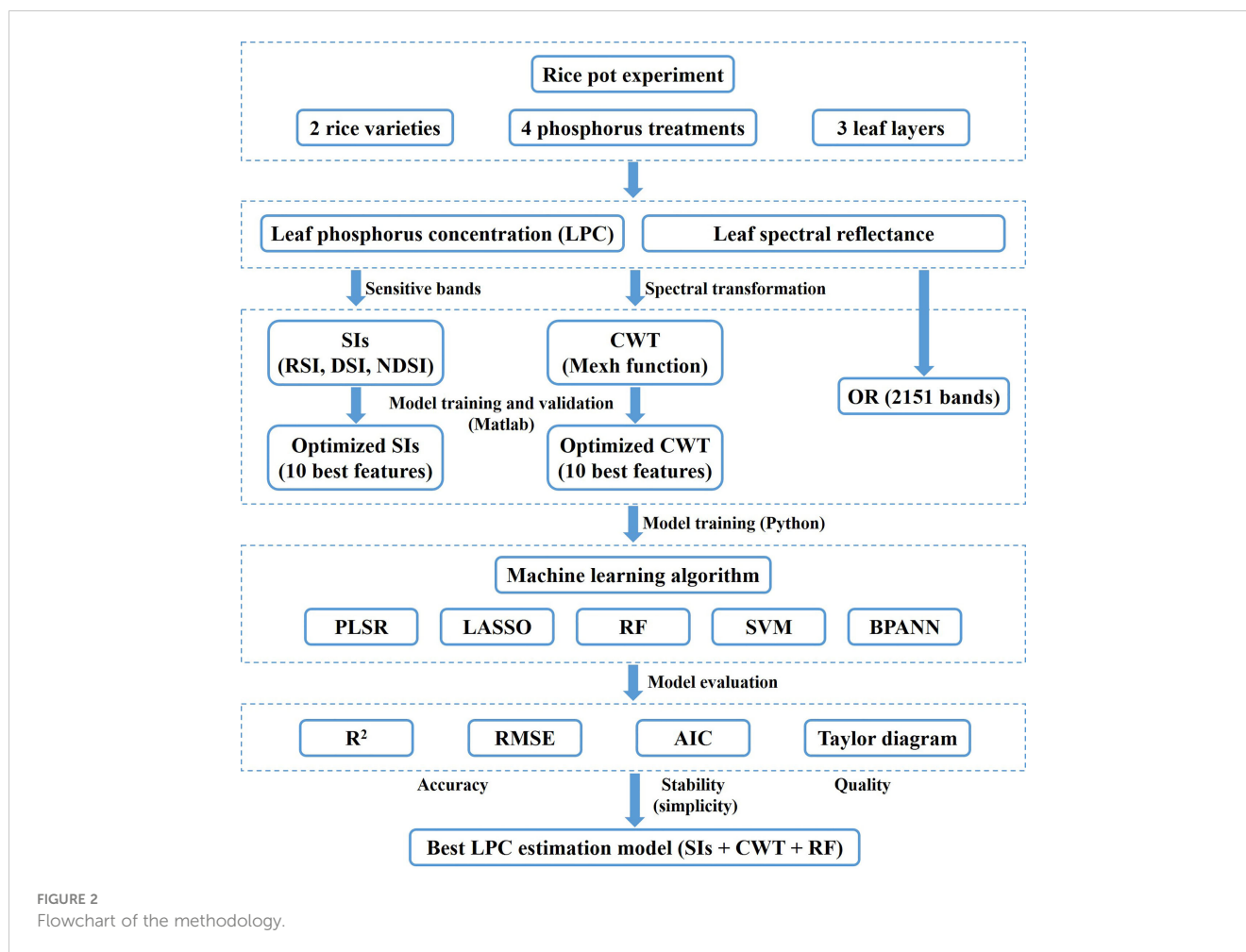
## 2.6 Model accuracy evaluation

The accuracy and simplicity of the model were evaluated by the determination coefficient ( $R^2$ ), root mean square error (RMSE,  $\text{mg g}^{-1}$ ), and Akaike information criterion (AIC). The calculation formula is shown as follows:

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

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{x}_i)^2}{n}} \quad (5)$$

$$AIC = 2k + n \cdot \ln \left( \frac{\sum_{i=1}^n (y_i - \bar{x})^2}{n} \right) \quad (6)$$



where  $\bar{x}$  represents the average of measured values.  $x_i$  and  $y_i$  represent the measured values and predicted values of LPC, respectively.  $n$  is the number of samples, and  $k$  is the number of features. The smaller RMSE with larger  $R^2$  values means better model estimation accuracy. AIC is an index for evaluating the model complexity, and the smaller value means a lower risk of overfitting.

Cross-validation can evaluate the machine learning model skills, which have a lower bias than other methods. The 10-fold coefficient of variation generally attains the lowest mean squared error and variance (Gao et al., 2019). For evaluate the model performance, the coefficient of determination ( $R^2$ ) and root mean squared error (RMSE) of the ten iterations were calculated in this study. Higher  $R^2$  and smaller RMSE indicate that the model has higher accuracy.

Taylor diagram provides a visual framework for the comparative assessment of different model results. The diagram can also be used to quantify the degree of correspondence between the predicted value of the models and the observations. It uses three statistics, the Pearson correlation coefficient, RMSE, and standard deviation (amplitude of variations) between predicted and observed values (Taylor, 2001). In this study, the Taylor diagram was used to evaluate the accuracy of the LPC estimation models based on the machine learning algorithms.

## 2.7 Statistical analysis

A one-way ANOVA was used to compare the means of LPC among different rice varieties, leaf layers, and P treatment based on the least significant difference at a 0.05 level of probability with DSS Statistics.

## 3 Results

### 3.1 Variations in LPC and spectral reflectance

Figure 3A shows the rice LPC in different P fertilizer applications, there was a significant difference among different P treatments. And the variation trend of LPC was  $P3 > P2 > P1 > P0$ . In terms of different leaf layers (Figure 3A), the rice LPC decreased from the upper to the lower layer, and there was no significant difference except for the  $P0$  treatment. The effect of the P application rate on the spectral reflectance of rice leaves in Longjing 31 (LJ31) and Wuyoudao 4 (WYD4) were analyzed, and there was no significant difference between the two rice varieties (Figure 3B).

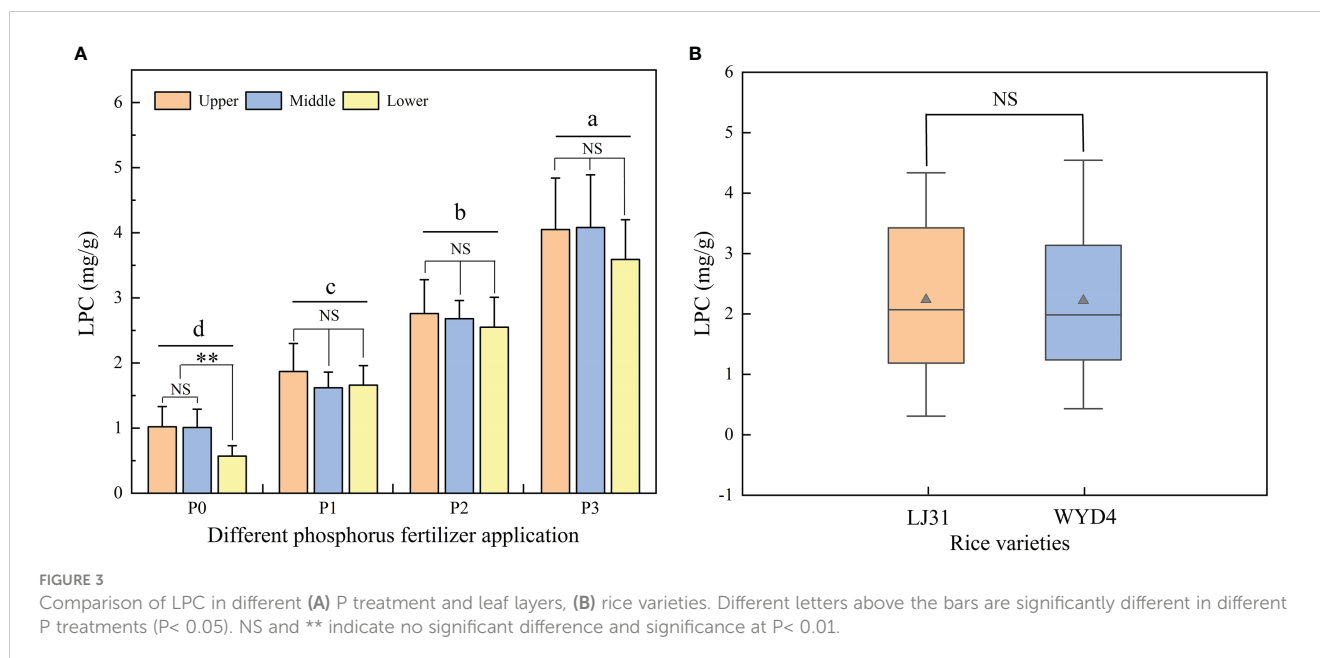
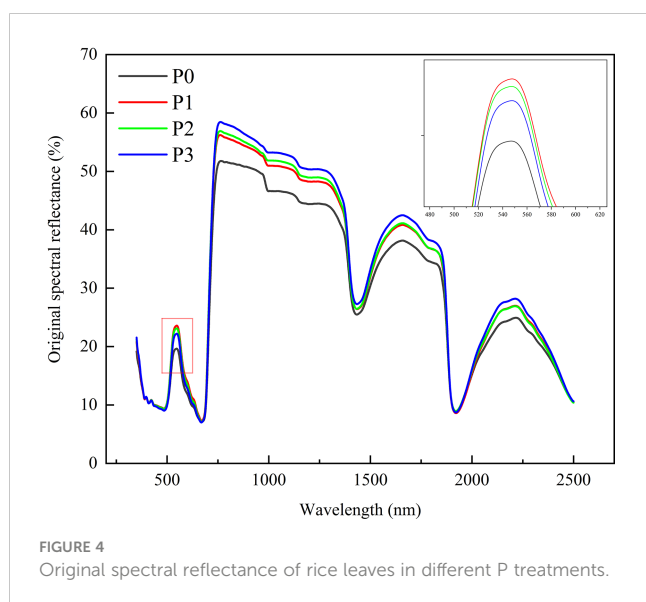


Figure 4 shows the original spectral reflectance of rice leaves in different P treatments in the range of 350–2500 nm. The results showed the P application rate significantly affected the leaf reflectance spectra, and the effects were different in the visible region (350–750 nm) and NIR regions (750–1350 nm). The spectral reflectance of rice leaf was at a low level (25%) in the visible region. The P deficiency mainly increased the leaf reflectance ( $P1 > P2 > P3$ ) at 550 nm. In the NIR regions, in contrast to the visible region, the leaf spectral reflectance was higher, and the P deficiency decreased leaf reflectance ( $P3 > P2 > P1 > P0$ ). Figure 5 shows the original spectral reflectance of rice leaves in different layers. The results showed there was no difference in spectral reflectance between the three layers. Thus, all rice leaf data in



different layers were pooled into one data set, and randomly allocated for model training and testing.

### 3.2 Estimation of rice LPC using spectral indices

To understand the relationships between LPC and RSI, DSI, and NDSI, the contour maps of the determination coefficient ( $R^2$ ) between three SIs and LPC were plotted in Figure 6. As illustrated, the performance of RSI was almost the same as NDSI, and the sensitive regions were mainly located in the NIR regions. The “hot spot” occurred in the area of the combination of 980–1140 nm (horizontal axis) and 960–990 nm (vertical axis). The  $R^2$  for the relationships between LPC and RSI, NDSI in the ranges were higher than 0.4. The sensitive band ranges for DSI were mainly concentrated on 1100–1400 nm (horizontal axis) and 1000–1300 nm (vertical axis). Overall, DSI consisting of 1089 nm and 1070 nm is the best performing spectral index for the estimation of LPC.

Based on the best performing SIs, rice LPC was estimated. The best correlations with LPC were selected to construct the traditional linear regression models (Figure 7). The results showed that the DSI (1089, 1070 nm) had higher  $R^2$  (0.54) in different calibration datasets compared to the RSI (1009, 990 nm) and NDSI (1009, 990 nm). The models were validated by the validation dataset. Relationships between the observed data and the predicted value of LPC by using the three SIs were illustrated in Figure 8. The results showed that the DSI had the best performance with an  $R^2$  of 0.55 and RMSE of  $0.67 \text{ mg g}^{-1}$  compared to RSI and NDSI. Therefore, the changes in LPC caused by different P supply levels can be estimated by optimized spectral index (DSI). However, the estimation accuracy of the linear regression models based on SIs was not high, and the calibration  $R^2$  lower than the validation  $R^2$ . These results showed the SIs models were underfitting and unstable.

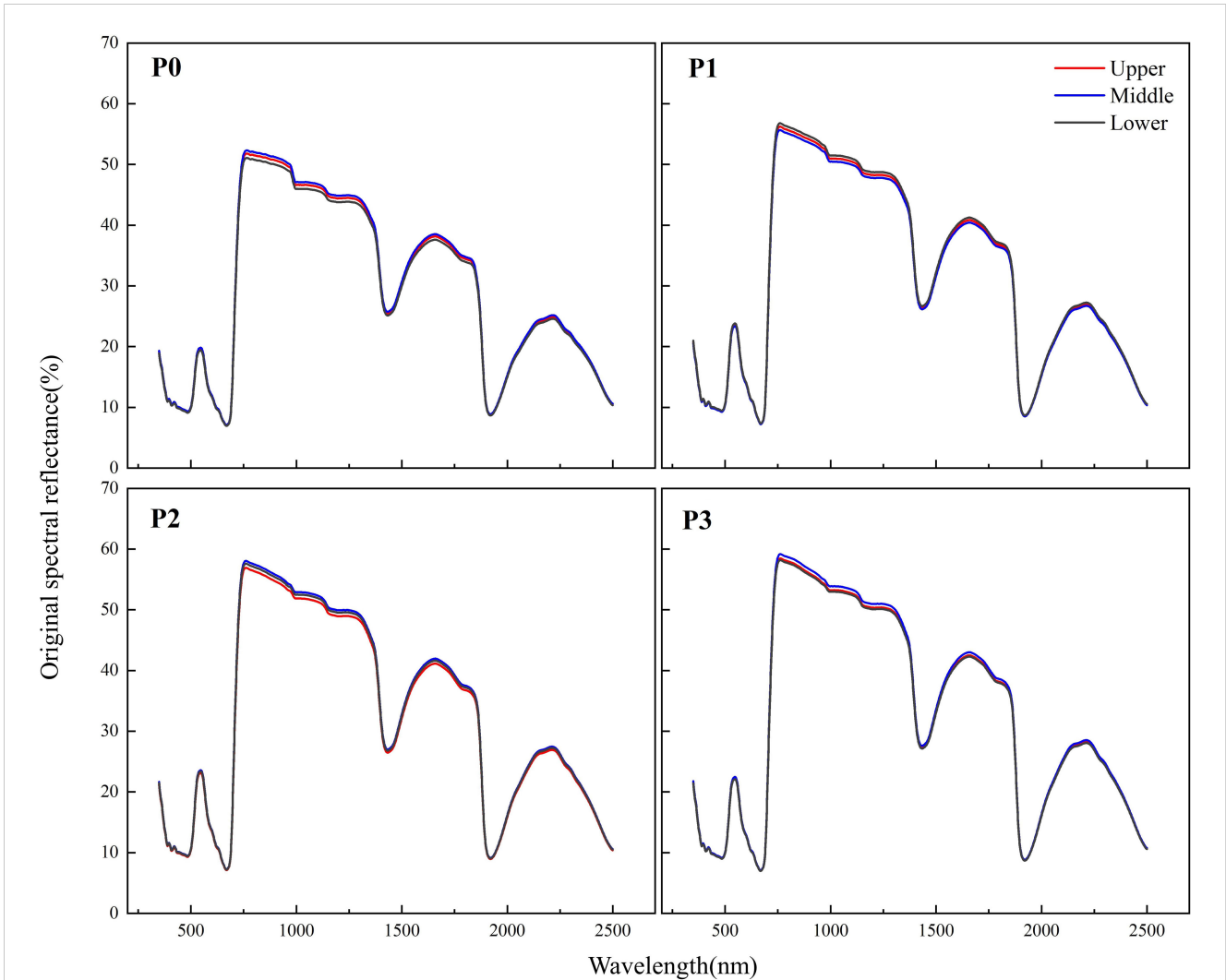


FIGURE 5 Original spectral reflectance of rice leaves in different leaf layers.

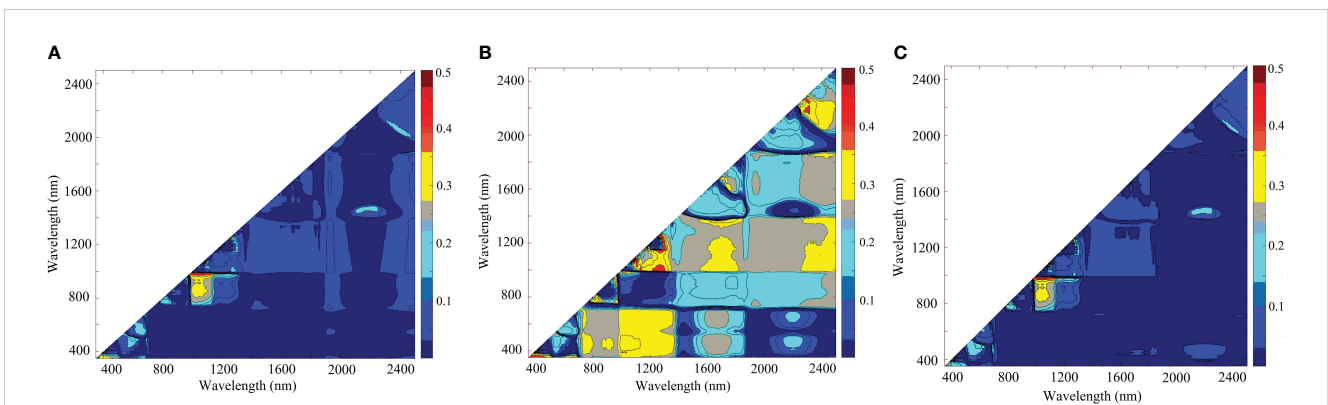


FIGURE 6 Contour maps of the determination coefficient ( $R^2$ ) between LPC and (A) RSI, (B) DSI, and (C) NDSI values.

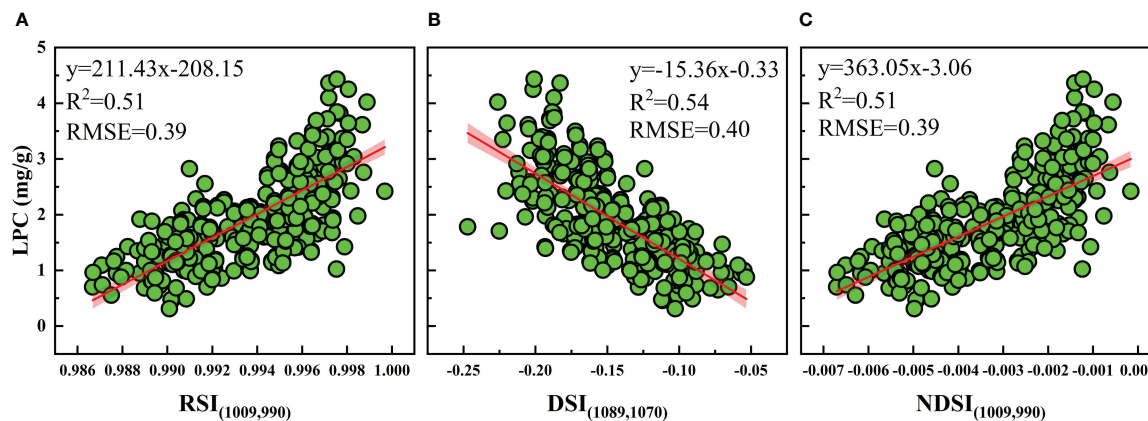


FIGURE 7  
The relationships between LPC and optimized (A) RSI, (B) DSI, and (C) NDSI for the calibration dataset.

### 3.3 Estimation of rice LPC using continuous wavelet transform

Figure 9 shows the relationships between using CWT of reflectance spectra on ten scales based on Mexh function and LPC of rice. Between 400 and 1700 nm, four wavelet features were observed that strongly correlated with the LPC of rice. The feature regions were centered at 400 nm, 1000 nm, 1470 nm, and 1680 nm. An optimal wavelet feature was selected on each scale to construct the LPC estimation model. The wavelet feature at 1680 nm and scale 6 provided the strongest correlation, with calibration  $R^2$  of 0.58, validation  $R^2$  of 0.56, and RMSE of  $0.61 \text{ mg g}^{-1}$  (Table 1). These results represent that the  $R^2$  values are improved by using CWT analysis compare with SIs (validation  $R^2 = 0.55$ ).

### 3.4 Estimation of rice LPC using machine learning algorithms

Figure 10 shows the statistical comparison results between 20 estimation models and the observations. The models constructed

using RF - CWT (point N) and RF - SIs + CWT (point S) were closer to the observation data (point A) on the Taylor diagram, and thus these two models are relatively superior to the other models. And the standard deviation of RF - SIs + CWT was closer to 1, which means the model has the best prediction performance. The accuracy of the 20 models for rice LPC was evaluated with 10-fold cross-validation (Table 2). The result indicates that the RF algorithm fed with the combination of SIs and CWT (RF - SIs + CWT) significantly improved estimation accuracy. In the validation set,  $R^2$  and RMSE were 0.73 and  $0.50 \text{ mg g}^{-1}$ , respectively and the model presents the lowest AIC of -3402.43 (Table 2).

## 4 Discussion

Rice growth is directly affected by soil P-supplying levels (Schachtman et al., 1998; Shen et al., 2011; Jiang et al., 2021). As an important indicator of crop growth, the changes in LPC can be obtained by spectral sensing technology. Previous research has discovered that various crops have varied P spectral response characteristics (Milton et al., 1991; Yaryura et al., 2009;

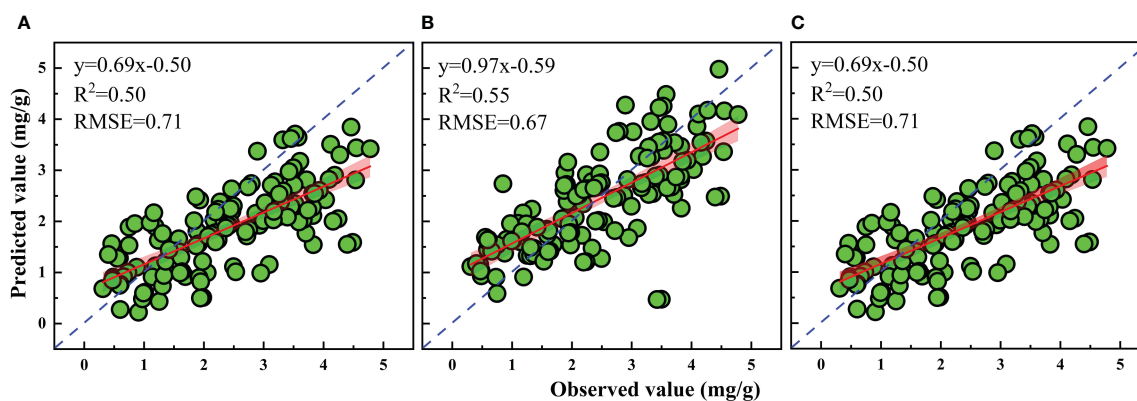


FIGURE 8  
Validation of the estimation models for LPC based on optimized (A) RSI, (B) DSI, and (C) NDSI.



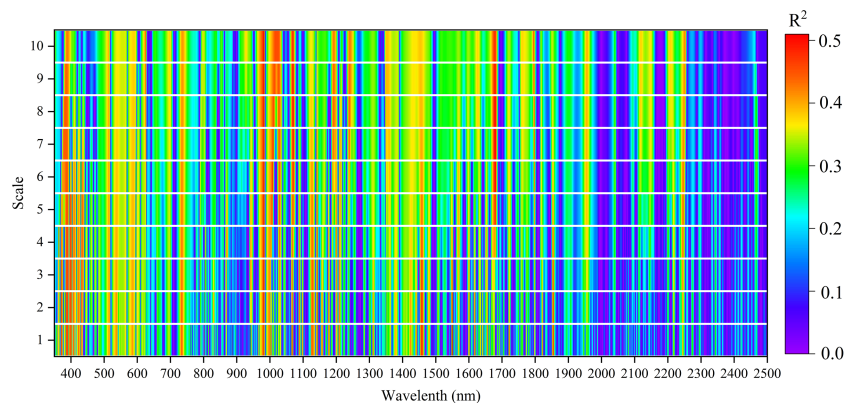


FIGURE 9  
Correlations between CWT and LPC at different transform scales.

Pacumbaba and Beyl, 2011). Our study measured the rice leaves in three layers at the tillering stage. The results showed the rice LPC decreased from the upper to the lower layer, and there was a significant difference between the upper and the lower layer in the P0 treatment. These results demonstrated the P would transfer from old leaves to new leaves when rice is suffered from extreme P deficiency. Previous studies indicated that P remobilization from aging organs to young organs occurred generally during the late vegetative and reproductive growth of plants (Veneklaas et al., 2012; Wang et al., 2021). In this study, the leaf samples were taken at the middle vegetative growth of rice, so there was no significant difference among the three layers under other P treatments. And the P deficiency decreased all rice leaves reflectance in the NIR regions (750-1350 nm), which is similar to the findings of Pacumbaba and Beyl (2011). In addition, many studies have investigated the N nutrition of plants, the sensitive bands of crop N concentration range from 340 nm to 900 nm (Li et al., 2014; Yang et al., 2021a). P concentration of the crop was slightly different from the N, the sensitive bands of crop P concentration were located from the visible region to NIR regions (Osborne et al., 2002;

Yaryura et al., 2009; Ramoelo et al., 2011; Mahajan et al., 2014). In our study, the sensitive bands of LPC were located in the NIR regions (750-1350 nm).

In general, N deficiency increases the leaf reflectance in green and red edge areas, which is due to the decrease of chlorophyll content in leaves (Daughtry et al., 2000; Zhao et al., 2003; Zhao et al., 2005). In P deficiency, one of the characteristic responses of plants is the visible accumulation of anthocyanin (AnC) (Jiang et al., 2007). Existing studies suggested that the AnC spectral feature of plant leaves was peaking around 550 nm in the visible region, and the spectral reflectance of AnC increased sharply near 700nm (Gitelson et al., 2001; Liu et al., 2015; Wang et al., 2020). Moreover, the peak magnitude was closely related to the content of AnC (Gitelson et al., 2001), and also with the increasing of AnC content, the reflectivity of leaves decreased (Liu et al., 2015). The AnC spectral features of plant leaves are similar to our results, which the leaf reflectance decreased with increasing P application rate in the visible region. Therefore, we considered that the spectral reflectance of P is affected by the AnC content of leaves in the visible region. Several studies found that the green (540-560 nm) and red

TABLE 1 Calibration and validation of LPC estimation models based on continuous wavelet function (Mexh).

Feature (wavelength in nm, scale)	Calibration dataset		Validation dataset	
	Fitted equation	R <sup>2</sup>	R <sup>2</sup>	RMSE
Mexh (1550 nm, 1)	$y = 647.56x + 1.80$	0.50	0.48	0.73
Mexh (982 nm, 2)	$y = -36.31x + 2.99$	0.51	0.49	0.71
Mexh (983 nm, 3)	$y = -16.51x + 2.55$	0.50	0.47	0.73
Mexh (982 nm, 4)	$y = -6.70x + 3.14$	0.50	0.50	0.70
Mexh (982 nm, 5)	$y = -4.25x + 3.15$	0.50	0.49	0.71
Mexh (1680 nm, 6)	$y = 22.74x + 0.15$	0.58	0.56	0.61
Mexh (1679 nm, 7)	$y = 15.42x - 0.06$	0.53	0.52	0.66
Mexh (1679 nm, 8)	$y = 11.69x - 0.20$	0.52	0.51	0.69
Mexh (982 nm, 9)	$y = -1.88x + 2.92$	0.51	0.50	0.71
Mexh (982 nm, 10)	$y = -1.72x + 2.80$	0.52	0.52	0.67

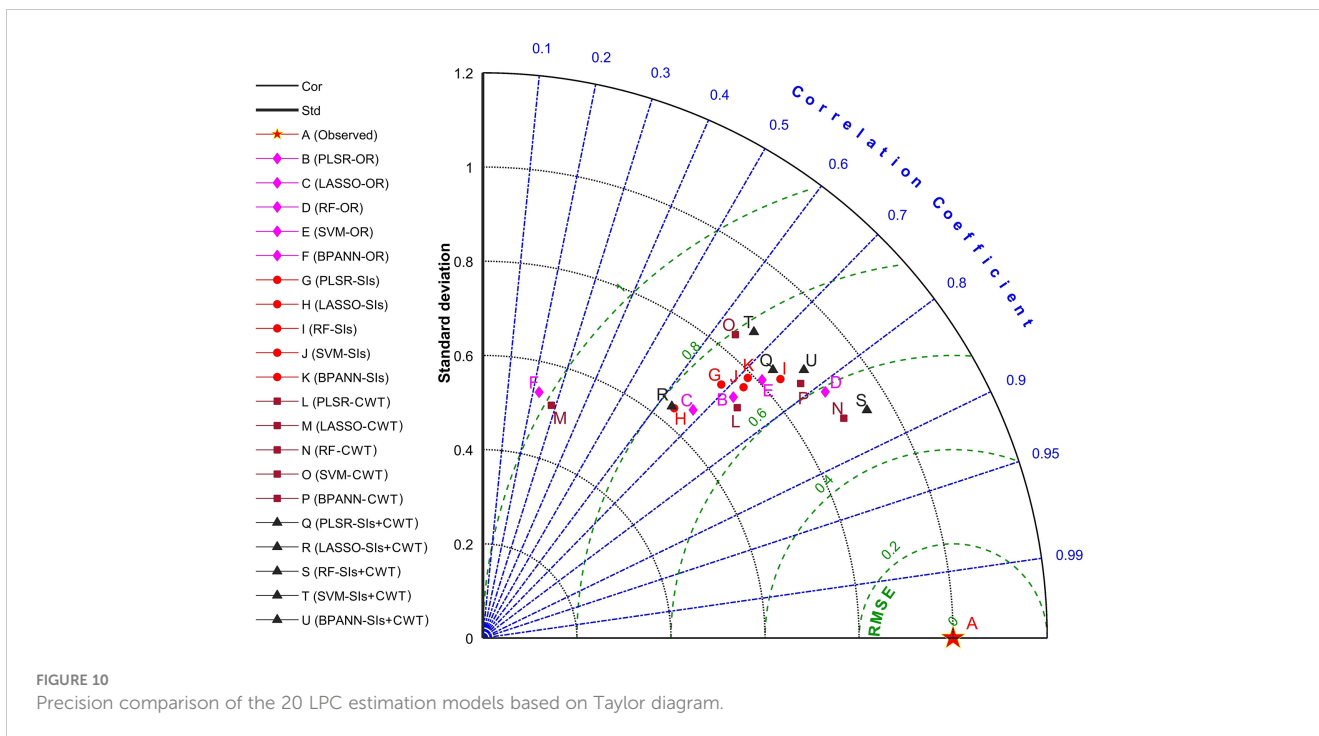


TABLE 2 10-fold cross-validation results of machine learning models.

Variables	No. of bands or features	Models	Calibration dataset		Validation dataset		AIC
			R <sup>2</sup>	R <sup>2</sup>	RMSE		
OR	2151	PLSR	0.49	0.52	0.72	1210.55	
		LASSO	0.47	0.46	0.76	1259.86	
		RF	0.95	0.66	0.60	1044.27	
		SVM	0.52	0.54	0.70	1184.86	
		BPANN	0.48	0.50	0.96	1472.91	
SIs	10	PLSR	0.45	0.47	0.73	-3058.87	
		LASSO	0.43	0.41	0.77	-3010.22	
		RF	0.81	0.57	0.64	-3178.87	
		SVM	0.53	0.52	0.68	-3123.58	
		BPANN	0.48	0.51	0.68	-3143.67	
CWT	10	PLSR	0.59	0.55	0.67	-3137.09	
		LASSO	0.12	0.08	0.95	-2718.64	
		RF	0.97	0.71	0.51	-3385.95	
		SVM	0.50	0.43	0.77	-3010.22	
		BPANN	0.74	0.61	0.62	-3207.83	
SIs + CWT	20	PLSR	0.57	0.54	0.67	-3117.09	
		LASSO	0.42	0.40	0.78	-2978.49	
		RF	0.95	0.73	0.50	-3402.43	
		SVM	0.51	0.44	0.77	-2990.22	
		BPANN	0.73	0.59	0.62	-3187.83	

(640–760 nm) bands were sensitive regions to AnC in plant leaves (Gitelson et al., 2006; Merzlyak et al., 2008; Liu et al., 2015; Wang et al., 2020). In contrast, our results showed the NIR regions (990 nm, 1009 nm, 1070 nm, and 1089 nm) were important to LPC estimation in rice by using SIs. In the optimal CWT, the sensitive bands also were 982 nm, 983 nm, 1550 nm, 1679 nm, and 1680 nm. And according to the feature importance of the RF model (Figure 11), 922 nm, 1134 nm, 983 nm, 923 nm, and 1185 nm were the sensitive bands for rice LPC estimation. The results are similar to the findings of other crops, the NIR was the best sensitive region for P estimation. For example, Ramoelo et al. (2011) indicated that the spectral absorption features used for P estimation of forage were mainly located in the NIR regions. Mahajan et al. (2014) found that the combination of reflectance in NIR and shortwave infrared (SWIR) regions significantly improved the accuracy of P content prediction of wheat. Therefore, NIR regions are more suitable for predicting the LPC of rice at tillering stage.

CWT has significant advantages in effectively obtaining spectral information, denoising, and dimensionality reduction of hyperspectral data (Ebrahimi and Rajaei, 2017; Li et al., 2022). Some previous studies confirmed CWT increased the estimation accuracy of crop leaf nitrogen status in rice, wheat, and summer maize (Li et al., 2018b; Li et al., 2022). Moreover, the Mexh wavelet family is often used as a CWT analysis method. Singh et al. (2013) found that in the quantification of crop leaf pigments, the model obtained by using the Mexh wavelet family has the highest accuracy compared with original spectra and other transformations of spectral reflectance data (Singh et al., 2013). Our study also found that the coefficient of correlation between the spectral data and rice LPC was improved by the CWT (Mexh function) of the original spectral data.

Machine learning methods have also been applied to predict the crop growth information and vegetation parameters, such as leaf water content (Zhang et al., 2021), and above-ground biomass

(Wang et al., 2016; Yang et al., 2021b) to further improve the accuracy of modeling. The estimation accuracy is affected by crop species, vegetation parameters, spectral index, and the type of machine learning algorithm (Chen et al., 2002; Gao et al., 2019). Previous studies showed the different performances of various algorithms. In the current study, PLSR, LASSO, RF, SVM, and BPANN algorithms were used to estimate the rice LPC. The effects of the five machine learning algorithms were different, and the four input variables (OR, SIs, CWT, and SIs + CWT) had a great influence on the estimation effect of the models. The numbers of input features of the models coupled with SIs and CWT were significantly less than that of OR, but the accuracy was improved. The results mean that the dimensionality reduction of input variables is crucial for machine learning algorithms (Yang et al., 2021b). Reducing the dimension can decrease the invalid bands and autocorrelation caused by massive data input, to make the machine learning model more accurate and efficient. In addition, compared with other machine learning algorithms, RF has fewer parameters (Wang et al., 2016). Hence, by incorporating the optimal features of SIs and CWT, the RF model was significantly improved. These results suggest that incorporating suitable input variables could significantly improve model accuracy and robustness. In addition, to determine the stability of the model, independent validation for the RF model was also conducted. The results were similar to the cross-validation results.

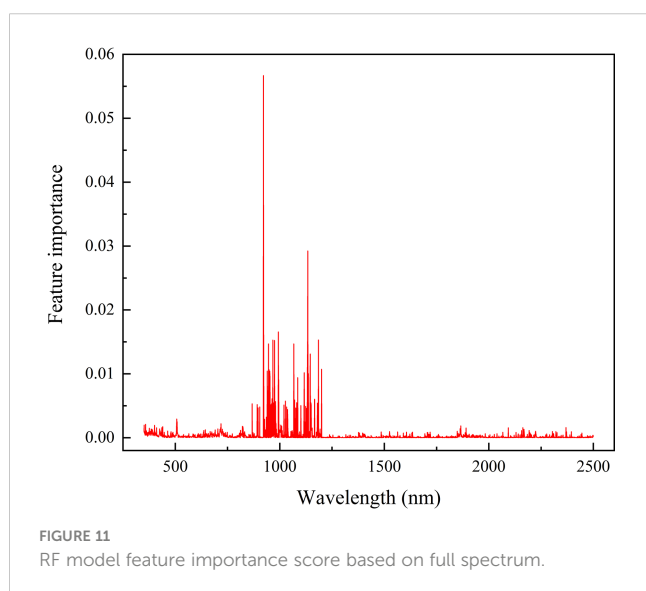
In sum, the combination of spectral index, wavelet analysis, and machine learning algorithms provides an efficient method for improving the estimation accuracy of rice LPC. Our findings may be useful for real time monitoring and diagnosis of rice phosphorus nutrition, and to provide a basic guideline for the best management practices of rice P fertilizer in the future.

## 5 Conclusions

In this study, we integrated SIs and CWT of the original spectrum with machine learning algorithms to offer an optimal prediction model for rice P concentration. The SIs + CWT coupling with the RF model can significantly increase rice LPC estimation accuracy while significantly reducing the number of input variables. The prediction accuracy of LPC with  $R^2$  was increased by 32% compared with the linear regression models. This study provides a new perspective to effectively estimate the P concentration in rice leaves. However, this study only aimed at the tillering stage of potted rice. Hence, in order to improve the applicability and prediction accuracy of the model, more data fusion approaches and new machine learning methods should be considered.

## 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.



## Author contributions

NC and YZ designed the research and supervised the project. TW, TLW, and ZL performed research and analysed data. TW, YZ, and NC wrote and revised the manuscript. All authors contributed to the article and approved the submitted version.

## Funding

This research was funded by the National Natural Science Foundation of China (32272819), National Key Research and Development Program of China (2017YFD0300608-2).

## Acknowledgments

We would like to thank PD Dr. Yuncai Hu, Department of Plant Sciences, Technical University of Munich, for his helpful discussions and English language editing of this manuscript.

## References

- An, G., Xing, M., He, B., Liao, C., Huang, X., Shang, J., et al. (2020). Using machine learning for estimating rice chlorophyll content from *In situ* hyperspectral data. *Remote Sens.* 12 (18). doi: 10.3390/rs12183104
- Bennett, E. M., Carpenter, S. R., and Caraco, N. F. (2001). Human impact on erodable phosphorus and eutrophication: a global perspective: increasing accumulation of phosphorus in soil threatens rivers, lakes, and coastal oceans with eutrophication. *BioScience* 51 (3), 227–234. doi: 10.1641/0006-3568(2001)051[0227:HIOEPA]2.0.CO;2
- Bindraban, P. S., Dimkpa, C. O., and Pandey, R. (2020). Exploring phosphorus fertilizers and fertilization strategies for improved human and environmental health. *Biol. Fertility Soils* 56 (3), 299–317. doi: 10.1007/s00374-019-01430-2
- Breiman, L. (2001). Random forests. *Mach. Learn.* 45 (1), 5–32. doi: 10.1023/A:1010933404324
- Carpenter, S. R. (2008). Phosphorus control is critical to mitigating eutrophication. *Proc. Natl. Acad. Sci.* 105 (32), 11039–11040. doi: 10.1073/pnas.0806112105
- Chen, M., Glaz, B., Gilbert, R. A., Daroub, S. H., Barton, F. E., and Wan, Y. (2002). Near-infrared reflectance spectroscopy analysis of phosphorus in sugarcane leaves. *Agron. J.* 94 (6), 1324–1331. doi: 10.2134/agronj2002.1324
- Cheng, T., Rivard, B., and Sánchez-Azofeifa, A. (2011). Spectroscopic determination of leaf water content using continuous wavelet analysis. *Remote Sens. Environ.* 115 (2), 659–670. doi: 10.1016/j.rse.2010.11.001
- Cortes, C., and Vapnik, V. (1995). Support-vector networks. *Mach. Learn.* 20 (3), 273–297. doi: 10.1023/A:1022627411411
- Darvishzadeh, R., Skidmore, A., Schlerf, M., Atzberger, C., Corsi, F., and Cho, M. (2008). LAI and chlorophyll estimation for a heterogeneous grassland using hyperspectral measurements. *ISPRS J. Photogrammetry Remote Sens.* 63 (4), 409–426. doi: 10.1016/j.isprsjprs.2008.01.001
- Daughtry, C. S. T., Walthall, C. L., Kim, M. S., de Colstoun, E. B., and McMurtrey, J. E. (2000). Estimating corn leaf chlorophyll concentration from leaf and canopy reflectance. *Remote Sens. Environ.* 74 (2), 229–239. doi: 10.1016/S0034-4257(00)00113-9
- Demay, J., Ringeval, B., Pellerin, S., and Nesme, T. (2023). Half of global agricultural soil phosphorus fertility derived from anthropogenic sources. *Nat. Geosci.* 16 (1), 69–74. doi: 10.1038/s41561-022-01092-0
- Ebrahimi, H., and Rajaei, T. (2017). Simulation of groundwater level variations using wavelet combined with neural network, linear regression and support vector machine. *Global Planetary Change* 148, 181–191. doi: 10.1016/j.gloplacha.2016.11.014
- Feng, W., Yao, X., Zhu, Y., Tian, Y. C., and Cao, W. X. (2008). Monitoring leaf nitrogen status with hyperspectral reflectance in wheat. *Eur. J. Agron.* 28 (3), 394–404. doi: 10.1016/j.eja.2007.11.005
- Gao, J., Meng, B., Liang, T., Feng, Q., Ge, J., Yin, J., et al. (2019). Modeling alpine grassland forage phosphorus based on hyperspectral remote sensing and a multi-factor

## 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.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2023.1185915/full#supplementary-material>

- machine learning algorithm in the east of Tibetan plateau, China. *ISPRS J. Photogrammetry Remote Sens.* 147, 104–117. doi: 10.1016/j.isprsjprs.2018.11.015
- Gitelson, A. A., Keydan, G. P., and Merzlyak, M. N. (2006). Three-band model for noninvasive estimation of chlorophyll, carotenoids, and anthocyanin contents in higher plant leaves. *Geophysical Res. Lett.* 33 (11), L11402. doi: 10.1029/2006GL026457
- Gitelson, A. A., Merzlyak, M. N., and Chivkunova, O. B. (2001). Optical properties and nondestructive estimation of anthocyanin content in plant leaves. *Photochem. Photobiol.* 74 (1), 38–45. doi: 10.1562/0031-8655(2001)0740038OPANEO2.0.CO;2
- Han, L., Yang, G., Dai, H., Xu, B., Yang, H., Feng, H., et al. (2019). Modeling maize above-ground biomass based on machine learning approaches using UAV remote-sensing data. *Plant Methods* 15 (1), 10. doi: 10.1186/s13007-019-0394-z
- Hansen, P. M., and Schjoerring, J. K. (2003). Reflectance measurement of canopy biomass and nitrogen status in wheat crops using normalized difference vegetation indices and partial least squares regression. *Remote Sens. Environ.* 86 (4), 542–553. doi: 10.1016/S0034-4257(03)00131-7
- Hao, P., Zhan, Y., Wang, L., Niu, Z., and Shakir, M. (2015). Feature selection of time series MODIS data for early crop classification using random forest: a case study in Kansas, USA. *Remote Sens.* 7 (5), 5347–5369. doi: 10.3390/rs70505347
- Heckmann, D., Schlüter, U., and Weber, A. P. M. (2017). Machine learning techniques for predicting crop photosynthetic capacity from leaf reflectance spectra. *Mol. Plant* 10 (6), 878–890. doi: 10.1016/j.molp.2017.04.009
- Jiang, C., Gao, X., Liao, L., Harberd, N. P., and Fu, X. (2007). Phosphate starvation root architecture and anthocyanin accumulation responses are modulated by the gibberellin-DELLA signaling pathway in arabidopsis. *Plant Physiol.* 145 (4), 1460–1470. doi: 10.1104/pp.107.103788
- Jiang, B., Shen, J., Sun, M., Hu, Y., Jiang, W., Wang, J., et al. (2021). Soil phosphorus availability and rice phosphorus uptake in paddy fields under various agronomic practices. *Pedosphere* 31 (1), 103–115. doi: 10.1016/S1002-0160(20)60053-4
- Jordan, C. F. (1969). Derivation of leaf-area index from quality of light on the forest floor. *Ecology* 50 (4), 663–666. doi: 10.2307/1936256
- Li, L., Geng, S., Lin, D., Su, G., Zhang, Y., Chang, L., et al. (2022). Accurate modeling of vertical leaf nitrogen distribution in summer maize using *in situ* leaf spectroscopy via CWT and PLS-based approaches. *Eur. J. Agron.* 140, 126607. doi: 10.1016/j.eja.2022.126607
- Li, F., Mistele, B., Hu, Y., Chen, X., and Schmidhalter, U. (2014). Optimising three-band spectral indices to assess aerial N concentration, N uptake and aboveground biomass of winter wheat remotely in China and Germany. *ISPRS J. Photogrammetry Remote Sens.* 92, 112–123. doi: 10.1016/j.isprsjprs.2014.03.006
- Li, D., Wang, C., Jiang, H., Peng, Z., Yang, J., Su, Y., et al. (2018a). Monitoring litchi canopy foliar phosphorus content using hyperspectral data. *Comput. Electron. Agric.* 154, 176–186. doi: 10.1016/j.compag.2018.09.007

- Li, D., Wang, X., Zheng, H., Zhou, K., Yao, X., Tian, Y., et al. (2018b). Estimation of area- and mass-based leaf nitrogen contents of wheat and rice crops from water-removed spectra using continuous wavelet analysis. *Plant Methods* 14 (1), 76. doi: 10.1186/s13007-018-0344-1
- Liaw, A., and Wiener, M. (2002). Classification and regression by randomForest. *R News* 2 (3), 18–22.
- Liu, X. Y., Shen, J., Chang, Q. R., Yan, L., Gao, Y. Q., and Xie, F. (2015). Prediction of anthocyanin content in peony leaves based on Visible/Near-infrared spectra. *Trans. Chin. Soc. Agric. Machinery* 46319-324 (9), 342. doi: 10.6041/j.issn.1000-1298.2015.09.047
- MacDonald, G. K., Bennett, E. M., Potter, P. A., and Ramankutty, N. (2011). Agronomic phosphorus imbalances across the world's croplands. *Proc. Natl. Acad. Sci. United States America* 108 (7), 3086–3091. doi: 10.1073/pnas.1010808108
- Mahajan, G. R., Pandey, R. N., Sahoo, R. N., Gupta, V. K., Datta, S. C., and Kumar, D. (2017). Monitoring nitrogen, phosphorus and sulphur in hybrid rice (*Oryza sativa* L.) using hyperspectral remote sensing. *Precis. Agric.* 18 (5), 736–761. doi: 10.1007/s11119-016-9485-2
- Mahajan, G. R., Sahoo, R. N., Pandey, R. N., Gupta, V. K., and Kumar, D. (2014). Using hyperspectral remote sensing techniques to monitor nitrogen, phosphorus, sulphur and potassium in wheat (*Triticum aestivum* L.). *Precis. Agric.* 15 (5), 499–522. doi: 10.1007/s11119-014-9348-7
- Mariotto, I., Thenkabail, P. S., Huete, A., Slonecker, E. T., and Platonov, A. (2013). Hyperspectral versus multispectral crop-productivity modeling and type discrimination for the HypsIRI mission. *Remote Sens. Environ.* 139, 291–305. doi: 10.1016/j.rse.2013.08.002
- Merzlyak, M. N., Chivkunova, O. F., Solovchenko, A. E., Solovchenko, A. F., and Naqvi, K. R. (2008). Light absorption by anthocyanins in juvenile, stressed, and senescing leaves. *J. Exp. Bot.* 59 (14), 3903–3911. doi: 10.1093/jxb/ern230
- Milton, N. M., Eiswerth, B. A., and Ager, C. M. (1991). Effect of phosphorus deficiency on spectral reflectance and morphology of soybean plants. *Remote Sens. Environ.* 36 (2), 121–127. doi: 10.1016/0034-4257(91)90034-4
- Mueller, N. D., Gerber, J. S., Johnston, M., Ray, D. K., Ramankutty, N., and Foley, J. A. (2012). Closing yield gaps through nutrient and water management. *Nature* 490 (7419), 254–257. doi: 10.1038/nature11420
- Murphy, J., and Riley, J. P. (1962). A modified single solution method for the determination of phosphate in natural waters. *Analytica Chimica Acta* 27, 678–681. doi: 10.1016/S0003-2670(00)88444-5
- Osborne, S. L., Schepers, J. S., Francis, D. D., and Schlemmer, M. R. (2002). Detection of phosphorus and nitrogen deficiencies in corn using spectral radiance measurements. *Agron. J.* 94 (6), 1215–1221. doi: 10.2134/agronj2002.1215
- Pacumbaba, R. O., and Beyl, C. A. (2011). Changes in hyperspectral reflectance signatures of lettuce leaves in response to macronutrient deficiencies. *Adv. Space Res.* 48 (1), 32–42. doi: 10.1016/j.asr.2011.02.020
- Pimstein, A., Karnieli, A., Bansal, S. K., and Bonfil, D. J. (2011). Exploring remotely sensed technologies for monitoring wheat potassium and phosphorus using field spectroscopy. *Field Crops Res.* 121 (1), 125–135. doi: 10.1016/j.fcr.2010.12.001
- Ramadan, Z., Hopke, P. K., Johnson, M. J., and Scow, K. M. (2005). Application of PLS and back-propagation neural networks for the estimation of soil properties. *Chemometrics Intellig. Lab. Syst.* 75 (1), 23–30. doi: 10.1016/j.chemolab.2004.04.009
- Ramoelo, A., Skidmore, A. K., Schlerf, M., Mathieu, R., and Heitkönig, I. M. A. (2011). Water-removed spectra increase the retrieval accuracy when estimating savanna grass nitrogen and phosphorus concentrations. *ISPRS J. Photogrammetry Remote Sens.* 66 (4), 408–417. doi: 10.1016/j.isprsjprs.2011.01.008
- Rivard, B., Feng, J., Gallie, A., and Sanchez-Azofeifa, A. (2008). Continuous wavelets for the improved use of spectral libraries and hyperspectral data. *Remote Sens. Environ.* 112 (6), 2850–2862. doi: 10.1016/j.rse.2008.01.016
- Rivera, J. P., Verrelst, J., Delegido, J., Veroustraete, F., and Moreno, J. (2014). On the semi-automatic retrieval of biophysical parameters based on spectral index optimization. *Remote Sens.* 6 (6), 4927–4951. doi: 10.3390/rs6064927
- Rouse, J. W. J., Haas, R. H., Deering, D. W., Schell, J. A., and Harlan, J. C. (1974). "Monitoring The vernal advancement and retrogradation (green wave effect) of natural vegetation," in *Great plains corridor* (Washington, DC, USA: NASA).
- Schachtman, D. P., Reid, R. J., and Ayling, S. M. (1998). Phosphorus uptake by plants: from soil to cell. *Plant Physiol.* 116 (2), 447–453. doi: 10.1104/pp.116.2.447
- Sharpley, A. N., and Withers, P. J. A. (1994). The environmentally-sound management of agricultural phosphorus. *Fertilizer Res.* 39 (2), 133–146. doi: 10.1007/BF00750912
- Shen, J., Yuan, L., Zhang, J., Li, H., Bai, Z., Chen, X., et al. (2011). Phosphorus dynamics: from soil to plant. *Plant Physiol.* 156 (3), 997–1005. doi: 10.1104/pp.111.175232
- Singh, S. K., Hoyos-Villegas, V., Ray, J. D., Smith, J. R., and Fritschi, F. B. (2013). Quantification of leaf pigments in soybean (*Glycine max* (L.) merr.) based on wavelet decomposition of hyperspectral features. *Field Crops Res.* 149, 20–32. doi: 10.1016/j.fcr.2013.04.019
- Takebe, M., Yoneyama, T., Inada, K., and Murakami, T. (1990). Spectral reflectance ratio of rice canopy for estimating crop nitrogen status. *Plant Soil* 122 (2), 295–297. doi: 10.1007/BF02851988
- Taylor, K. E. (2001). Summarizing multiple aspects of model performance in a single diagram. *J. Geophysical Res.* 106 (D7), 7183–7192. doi: 10.1029/2000JD900719
- Tian, Y.-C., Gu, K.-J., Chu, X., Yao, X., Cao, W.-X., and Zhu, Y. (2014). Comparison of different hyperspectral vegetation indices for canopy leaf nitrogen concentration estimation in rice. *Plant Soil* 376 (1), 193–209. doi: 10.1007/s11104-013-1937-0
- Tibshirani, R. (2011). Regression shrinkage and selection via the lasso: a retrospective. *J. R. Stat. Soc. (Statistical Methodology)* 73 (3), 273–282. doi: 10.1111/j.1467-9868.2011.00771.x
- Tilman, D., Balzer, C., Hill, J., and Befort, B. L. (2011). Global food demand and the sustainable intensification of agriculture. *Proc. Natl. Acad. Sci.* 108 (50), 20260–20264. doi: 10.1073/pnas.1116437108
- Tilman, D., Cassman, K. G., Matson, P. A., Naylor, R., and Polasky, S. (2002). Agricultural sustainability and intensive production practices. *Nature* 418 (6898), 671–677. doi: 10.1038/nature01014
- Townsend, A. R., and Porder, S. (2012). Agricultural legacies, food production and its environmental consequences. *Proc. Natl. Acad. Sci. United States America* 109 (16), 5917–5918. doi: 10.1073/pnas.1203766109
- Tucker, C. J. (1979). Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sens. Environ.* 8, 127–150. doi: 10.1016/0034-4257(79)90013-0
- Veneklaas, E. J., Lambers, H., Bragg, J., Finnegan, P. M., Lovelock, C. E., Plaxton, W. C., et al. (2012). Opportunities for improving phosphorus-use efficiency in crop plants. *New Phytol.* 195 (2), 306–320. doi: 10.1111/j.1469-8137.2012.04190.x
- Verrelst, J., Camps-Valls, G., Muñoz-Marí, J., Rivera, J. P., Veroustraete, F., Clevers, J. G. P. W., et al. (2015). Optical remote sensing and the retrieval of terrestrial vegetation bio-geophysical properties – a review. *ISPRS J. Photogrammetry Remote Sens.* 108, 273–290. doi: 10.1016/j.isprsjprs.2015.05.005
- Verrelst, J., Malenovský, Z., van der Tol, C., Camps-Valls, G., Gastellu-Etchegorry, J.-P., Lewis, P., et al. (2019). Quantifying vegetation biophysical variables from imaging spectroscopy data: a review on retrieval methods. *Surveys Geophysics* 40 (3), 589–629. doi: 10.1007/s10712-018-9478-y
- Viña, A., and Gitelson, A. A. (2011). Sensitivity to foliar anthocyanin content of vegetation indices using green reflectance. *IEEE Geosci. Remote Sens. Lett.* 8 (3), 464–468. doi: 10.1109/LGRS.2010.2086430
- Wang, W. D., Chang, Q. R., and Wang, Y. N. (2020). Hyperspectral monitoring of anthocyanins relative content in winter wheat leaves. *J. Triticeae Crops* 40 (6), 754–761. doi: 10.7606/j.issn.1009-1041.2020.06.14
- Wang, J., Chen, Y., Chen, F., Shi, T., and Wu, G. (2018). Wavelet-based coupling of leaf and canopy reflectance spectra to improve the estimation accuracy of foliar nitrogen concentration. *Agric. For. Meteorology* 248, 306–315. doi: 10.1016/j.agrformet.2017.10.017
- Wang, J., Pan, W., Nikiforov, A., King, W., Hong, W., Li, W., et al. (2021). Identification of two glycerophosphodiester phosphodiesterase genes in maize leaf phosphorus remobilization. *Crop J.* 9 (1), 95–108. doi: 10.1016/j.cj.2020.05.004
- Wang, L. a., Zhou, X., Zhu, X., Dong, Z., and Guo, W. (2016). Estimation of biomass in wheat using random forest regression algorithm and remote sensing data. *Crop J.* 4 (3), 212–219. doi: 10.1016/j.cj.2016.01.008
- Xue, L., Cao, W., Luo, W., Dai, T., and Zhu, Y. (2004). Monitoring leaf nitrogen status in rice with canopy spectral reflectance. *Agron. J.* 96 (1), 135–142. doi: 10.2134/agronj2004.1350
- Yang, H., Li, F., Hu, Y., and Yu, K. (2021a). Hyperspectral indices optimization algorithms for estimating canopy nitrogen concentration in potato (*Solanum tuberosum* L.). *Int. J. Appl. Earth Observation Geoinformation* 102, 102416. doi: 10.1016/j.jag.2021.102416
- Yang, H., Li, F., Wang, W., and Yu, K. (2021b). Estimating above-ground biomass of potato using random forest and optimized hyperspectral indices. *Remote Sens.* 13 (12). doi: 10.3390/rs13122339
- Yaryura, P., Cordon, G., Leon, M., Kerber, N., Pucheu, N., Rubio, G., et al. (2009). Effect of phosphorus deficiency on reflectance and chlorophyll fluorescence of cotyledons of oilseed rape (*Brassica napus* L.). *J. Agron. Crop Sci.* 195 (3), 186–196. doi: 10.1111/j.1439-037X.2008.00359.x
- Zhai, Y., Cui, L., Zhou, X., Gao, Y., Fei, T., and Gao, W. (2013). Estimation of nitrogen, phosphorus, and potassium contents in the leaves of different plants using laboratory-based visible and near-infrared reflectance spectroscopy: comparison of partial least-square regression and support vector machine regression methods. *Int. J. Remote Sens.* 34 (7), 2502–2518. doi: 10.1080/01431161.2012.746484
- Zhang, J., Zhang, W., Xiong, S., Song, Z., Tian, W., Shi, L., et al. (2021). Comparison of new hyperspectral index and machine learning models for prediction of winter wheat leaf water content. *Plant Methods* 17 (1), 34. doi: 10.1186/s13007-021-00737-2
- Zhao, B., Duan, A., Ata-Ul-Karim, S. T., Liu, Z., Chen, Z., Gong, Z., et al. (2018). Exploring new spectral bands and vegetation indices for estimating nitrogen nutrition index of summer maize. *Eur. J. Agron.* 93, 113–125. doi: 10.1016/j.eja.2017.12.006
- Zhao, D., Raja Reddy, K., Kakani, V. G., Read, J. J., and Carter, G. A. (2003). Corn (*Zea mays* L.) growth, leaf pigment concentration, photosynthesis and leaf hyperspectral reflectance properties as affected by nitrogen supply. *Plant Soil* 257 (1), 205–218. doi: 10.1023/A:1026233732507
- Zhao, D., Reddy, K. R., Kakani, V. G., and Reddy, V. R. (2005). Nitrogen deficiency effects on plant growth, leaf photosynthesis, and hyperspectral reflectance properties of sorghum. *Eur. J. Agron.* 22 (4), 391–403. doi: 10.1016/j.eja.2004.06.005