The International Society of Precision Agriculture presents the

16th International Conference on Precision Agriculture

21–24 July 2024 | Manhattan, Kansas USA

Improving winter wheat plant nitrogen concentration prediction by combining proximal hyperspectral sensing and weather information

Xiaokai Chen[a], Fenling Li[a], Qingrui Chang[a]*, Yuxin Miao[b]**, Kang Yu[c]

a College of Natural Resources and Environment, Northwest A&F University, Yangling 712100, China

b Precision Agriculture Center, Department of Soil, Water and Climate, University of Minnesota, St. Paul., MN, 55108, USA

c Precision Agriculture Lab, Department Life Science Engineering, School of Life Sciences, Technical University of Munich, 85354 Freising, Germany

* Corresponding authors at: College of Natural Resources and Environment, Northwest A&F University, Yangling 712100, China

** Corresponding authors at: Precision Agriculture Center, Department of Soil, Water and Climate, University of Minnesota, St. Paul., MN, 55108, USA

E-mail addresses: changqr@nwsuaf.edu.cn (Q. Chang), ymiao@umn.edu (Y. Miao).

**A paper from the Proceedings of the
16th International Conference on Precision Agriculture
21-24 July 2024
Manhattan, Kansas, United States**

Abstract.

Timely and accurate prediction of winter wheat nitrogen (N) status plays an important role in guiding precision N management. The objectives of this study were to identify potential effective bands and spectral features of proximal hyperspectral sensing data using different preprocessing methods for predicting winter wheat plant N concentration (PNC) with seven machine learning (ML) algorithms and determine the potential to further improve the accuracy of PNC prediction by combining hyperspectral sensing data with weather information. The least absolute shrinkage and selection operator (LASSO) method was applied to identify the effective bands from different preprocessed reflectance (original (OR), first-order derivative (FD), apparent absorption (LOG), and continuum removal (CR)) based on data collected from six site-year field experiments conducted in 2014-2023. The results indicated that effective bands of FD combined with support vector regression (SVR) yield satisfying PNC prediction accuracy, with coefficient of determination (R2), root means square error (RMSE), and residual predictive deviation (RPD) being 0.80, 0.27 and 2.16, respectively. However, when weather information was combined with the proximal hyperspectral sensing data, the accuracy of winter wheat PNC prediction was significantly

improved, with an increase of 0.04-0.46 in R^2, a decrease of 0.03-0.22 in RMSE and an increase of 0.26-1.02 in RPD. Random forest regression (RFR) combining FD and weather information yielded the best PNC predictions (R^2 = 0.85, RMSE = 0.23, and RPD = 2.56). It is concluded that the RFR model combining proximal hyperspectral sensing data and weather information can significantly improve winter wheat PNC prediction and is a promising and practical strategy to predict winter wheat N status. More studies are needed to develop unmanned aerial vehicle or satellite hyperspectral remote sensing-based multi-source data fusion strategies using ML for more efficient monitoring of crop N status over a large area.

Keywords.
Plant nitrogen concentration, Proximal hyperspectral sensing, Weather information, Machine learning

1. Introduction

Wheat (*Triticum aestivum* L.) is one of the most important staple food crops influencing global food security (Jiang et al. 2022). Nitrogen (N) is the most needed nutrient element and plays a vital role in the growth and development of wheat (Lemaire et al. 2008, Liu et al. 2023). However, excessive application of N fertilizer can cause many environmental problems (Padilla et al. 2018, Skiba and Rees 2014), under-application can negatively affect crop photosynthesis and yield (Chlingaryan et al. 2018, Dong et al. 2021b, Miao et al. 2011, Wang et al. 2023). Therefore, N fertilizers should be optimized to meet crop needs without polluting the environment, which requires timely and accurate diagnosis of winter wheat N status.

The traditional biophysical chemical method for determining N concentrations is destructive, time-consuming, laborious and expensive. Proximal and remote sensing techniques have provided new opportunities for non-destructive and accurate monitoring of N status (Chen et al. 2019). A large number of vegetation indices (VIs) have been developed and used to monitor crop N status (Hansen and Schjoerring 2003, Li et al. 2013, 2014, Müller et al. 2008, Yang et al. 2021a). Although acceptable N diagnostic accuracy was achieved, the VIs was generally composed of two or three bands. The band combination with low signal-to-noise ratio will also have a negative impact on the prediction model. At the same time, no VIs has been found to perform consistently to monitor crop N status across different sites and regions, crop varieties and growth stages (Li et al. 2010, Liu et al. 2019, Yang et al. 2023). Some scholars optimized the VIs and used them to monitor N concentration (Li et al. 2010, Yang et al. 2021b, Yang et al. 2023), but the regression models were still calibrated for a specific data set, and the transferability of these models were questionable. It is worth noting that most of the existing VIs with satisfying performances are based on the high correlation between chlorophyll content and N status (Wood et al. 1992). However, the crop growth is a dynamic process of continuous N cycling and turnover (Kattge 2002). With the advance of crop growth and development, especially from vegetative growth to reproductive growth, N will be redistributed to reproductive structures in plants, which will lead to the decrease of the relationship between chlorophyll content and N (Ohyama 2010). Therefore, relying on chlorophyll content for monitoring crop N status can be misleading. It has been pointed out that protein is the main biochemical component containing N in plants (Kokaly and Clark 1999), which may be a better representative (Yasumura et al. 2007). The sensitive band is usually located at a longer wavelength, concentrated in the near infrared and short-wave infrared regions (Berger et al. 2020b). Herrmann et al. (Herrmann et al. 2010) achieved satisfactory results in terms of the ability to predict N content and sensitivity to N content using short-wave infrared reflectance. Therefore, crop N status monitoring may be improved via the introduction of the short-wave infrared region.

Under normal circumstances, obtaining canopy hyperspectral reflectance is inevitably affected by soil background, atmosphere, canopy structure, and illumination changes (Tsai and Philpot 1998). Researchers tried to use first-order derivative (FD) transformation to suppress the impact of illumination changes on canopy reflectance. The results showed that FD transformation spectrum can effectively predict the light-summing ability of crop leaves (Jin et al. 2020); apparent absorption (LOG) and (CR) transformation have also been used to estimate physical and chemical parameters and have achieved satisfactory research results (Huang et al. 2004, LaCapra et al. 1996). It can be seen that it is very necessary to preprocess the original (OR) spectrum using spectral transformation technology. However, there is currently no definite conclusion as to which spectral transformation effect is optimal in PNC prediction of winter wheat.

Machine learning (ML) algorithms are increasingly used to improve crop N status monitoring (Li et al. 2019, Wang et al. 2021, Zha et al. 2020a). Upon integrating proximal hyperspectral sensing data into ML algorithms, the high dimensionality of reflectance data, often referred to as the 'curse of dimensionality' (Berger et al. 2020b, Yang et al. 2023), leads to a reduction in the prediction accuracy of ML algorithms.

Proceedings of the 16th International Conference on Precision Agriculture
21-24 July, 2024, Manhattan, Kansas, United States

2

On the other hand, the narrow bands in hyperspectral data yield a substantial amount of information, resulting in the generation of numerous redundant bands. Directly using all these data in the ML algorithms will result in the risk of overfitting, potentially compromising the interpretability of the regression models. Hence, the essential step of dimensionality reduction in hyperspectral reflectance data becomes crucial to enhance both the prediction performance and interpretability of the ML algorithms(Fu et al. 2021). The least absolute shrinkage and selection operator (LASSO) is a powerful feature identification method, which has been used to process high-dimensional data in linear and nonlinear situations (Cao et al. 2021, Kukreja et al. 2006, Kumar et al. 2019, Ordonez et al. 2018). Song et al. (Song and Wang 2023a) applied LASSO to identify the characteristic variables of leaf photosynthetic parameters (V_{cmax} and J_{max}), evaluating leaf photosynthetic capacity by the random forest regression (RFR) and support vector regression (SVR). The results showed SVR were more accurate. Ordonez et al.(Ordonez et al. 2018) utilized the LASSO method to streamline the sample data dimensionality reduction. From the continuous reflectance spectrum, they identified optimal wavelengths to construct a model for estimating leaf water content. However, the ability of ML algorithms combined with LASSO to identify the effective bands of proximal hyperspectral sensing reflectance to diagnose the N status of winter wheat deserves further exploration.

It is reported that factors such as planting year, region, growth stage, soil type, variety, weather conditions and management conditions will affect the monitoring accuracy of crop growth and N status (Bean et al. 2018, Ruan et al. 2022, Schepers et al. 2004). ML algorithms have the potential to fuse multi-source data improve crop growth and N status monitoring (Grinberg et al. 2020, Küçük et al. 2016, Stas et al. 2016). Some scholars tried to use ML algorithms to combine VIs with different ancillary data to predict crop yield and N status (Ruan et al. 2022). Lu et al. (Lu et al. 2022b) integrated active sensor data with different varieties, environmental variables and different transplanting densities using ML algorithms and found that RFR model with multi-source data fusion greatly improved the diagnostic accuracy of rice N status. Dong et al. (Dong et al. 2021a) proved when the leaf sensor data was combined with environmental and management variables in multiple linear regression (MLR) model, corn N Nutrition Index (NNI) prediction and N status diagnosis was significantly improved. Li et al. (Li et al. 2022b) found the RFR model combining climate and management factors with VIs performed better than the RFR model only using VIs. These studies proved the feasibility of diagnosing crop N status by combining ML algorithms with proximal active multispectral sensors and ancillary. However, studies on the potential of improving crop N status prediction using proximal hyperspectral sensing and multi-source data fusion with ML are still limited.

The objectives of this study were to (1) understand how different types of preprocessed reflectance (OR, FD, LOG, and CR) affect the selection of effective bands of proximal hyperspectral sensing data in monitoring plant N concentration (PNC) with seven ML algorithms; and (2) determine if the introduction of weather information can significantly improve the accuracy of PNC monitoring model.

2. Materials and methods

2.1 Experimental designs

Winter wheat field experiments were conducted in Qian County (108°07'E, 34°38'N), Shaanxi Province, China, from 2014 to 2023 (Figure1). The daily precipitation (mm) and daily mean temperature (℃) during the winter wheat growing stage are illustrated in Fig.2. The plot area varied, with experiments 1 and 2 having 6 m×6 m plots and experiments 3-6 having 9 m×10 m plots. Each experiment consisted of N, P_2O_5, and K_2O as N, P, and K fertilizers, respectively. All fertilizers were applied before planting. Winter wheat in this study was not irrigated and was planted according to the local standard density. Other management practices followed local conventions, with no significant pest problems. Specific details of the six experiments are provided in Table 1.
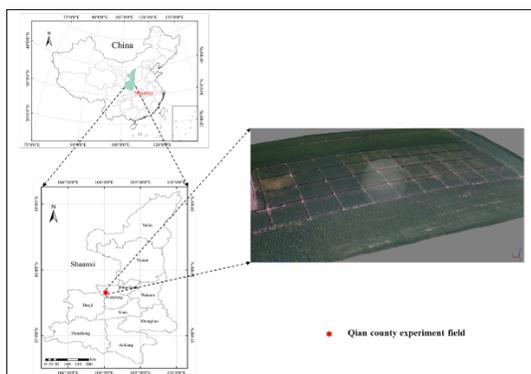

Proceedings of the 16th International Conference on Precision Agriculture
21-24 July, 2024, Manhattan, Kansas, United States

3

Figure 1. Geographic location of Qian County experiment field in this study.

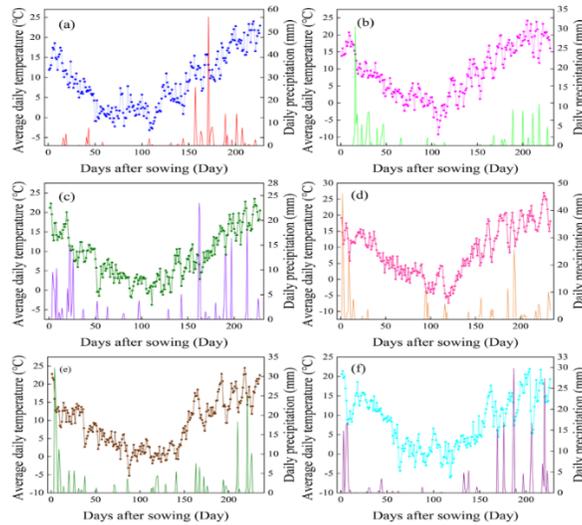Figure 2. Daily precipitation (mm) and Daily mean temperature (℃) during winter wheat growing stage of (a) 2014-2015, (b) 2015-2016, (c) 2016-2017, (d) 2017-2018, (e) 2021-2022, (f) 2022-2023 in this study.

Table 1. Basic information of the field experimental design in this study.

Experiment of year	Location	Cultivar	N rate (Kg/ha)	P rate (Kg/ha)	K rate (Kg/ha)	Irrigation	Sowing date	Sampling stage
Experiment 1 2014-2015	Qian county 108°07'E, 34°38'N	Xiaoyan 22	0, 37.5, 75, 112.5, 150, 187.5	0, 22.5, 45, 67.5, 90, 112.5	0, 15, 30, 45, 60, 75	No	13.Oct.	BBCH23, BBCH39, BBCH56, BBCH65, BBCH75
Experiment 2 2015-2016	Qian county 108°07'E, 34°39'N	Xiaoyan 22	0, 37.5, 75, 112.5, 150, 187.5	0, 22.5, 45, 67.5, 90, 112.5	0, 15, 30, 45, 60, 75	No	9.Oct.	BBCH65, BBCH75
Experiment 3 2016-2017	Qian county 108°07'E, 34°40'N	Xiaoyan 22	0, 30, 60, 90, 120, 150	0, 22.5, 45, 67.5, 90, 112.5	0, 22.5, 45, 67.5, 90, 112.5	No	2.Oct.	BBCH23, BBCH39, BBCH56
Experiment 4 2017-2018	Qian county 108°07'E, 34°41'N	Xiaoyan 22	0, 30, 60, 90, 120, 150	0, 22.5, 45, 67.5, 90, 112.5	0, 22.5, 45, 67.5, 90, 112.5	No	2.Oct.	BBCH23, BBCH39, BBCH65, BBCH75
Experiment 5 2021-2022	Qian county 108°07'E, 34°42'N	Xiaoyan 22	0, 60, 120, 180, 240, 300	0, 30, 60, 90, 120, 150	0, 30, 60, 90, 120, 150	No	1.Oct.	BBCH56, BBCH65
Experiment 6 2022-2023	Qian county 108°07'E, 34°43'N	Xiaoyan 22	0, 45, 90, 135, 180, 225, 270, 315	75	75	No	29.Sep.	BBCH56, BBCH65

2.2 Data collection

2.2.1 Remote sensing data collection

Proximal hyperspectral sensing reflectance measurements were conducted using a handheld spectrometer (HR-1024i) from Spectra Vista Corporation (SVC), Poughkeepsie, USA, in each plot before destructive sampling throughout the growth stages of winter wheat. The SVC device yielded 1024 bands within the wavelength range of 350-2500 nm. Data for the winter wheat canopy spectrum were obtained between 10:00 am and 2:00 pm under clear sky, with no wind or clouds. The SVC sensor lens was positioned vertically downward, approximately 1 m from the vertical height of the canopy, and the field of view angle

**Proceedings of the 16th International Conference on Precision Agriculture
21-24 July, 2024, Manhattan, Kansas, United States**

4

was set at 25°. Calibration of the sensor using a Reference Panels Cases was performed both before and after each spectrum measurement. In the SVC sensor measurement process, two sampling points were evenly collected in each plot, and 10 spectra were repeated for one sampling point. The average value of these repeated spectra was used as the spectral reflectance for the observed sampling point. Furthermore, the spectral reflectance for each plot was determined as the average reflectance of the two sampling points.

2.2.2 Agronomic data collection

Winter wheat plant samples in all experiments were collected following canopy sensing data collection. Simultaneously, global positioning system (GPS) was employed to record the coordinate information of the sampling points. Subsequently, the collected plants were placed in an oven at 105°C for approximately 30 minutes to stop the physiological processes. The drying process continued at 70°C until the weight of the sample remained constant. After drying, the plant samples were ground to fine powders, and the modified Kjeldahl digestion method was utilized to determine the PNC.

2.2.3 Weather information collection

Weather information was introduced to improving monitoring winter wheat N diagnostic in this study. The daily weather data of all six field experiments were obtained from the Qian County Meteorological Station. According to previous studies, growing degree days (GDD), accumulated precipitation (APP) and days (D) from the sowing date to the measurement date were calculated. Considering that the temperature in a period of time may have a more significant impact on crop growth and development, continuously iterative weather data of the 30 days before measurement date were also introduced, including average daily temperature (ADT), average daily minimum temperature (ADT_{min}), average daily maximum temperature (ADT_{max}), and accumulated daily average temperature (ADAT) (Jiang et al. 2022). The formula of GDD was shown as follows, and Tbase = 0℃ according to previous research (Li et al. 2022c):

$$GDD = \sum \left(\frac{(T_{max} + T_{min})}{2} - T_{base} \right) \qquad (1)$$

where T_{max}, T_{min}, and T_{base} are the daily maximum, minimum, and base temperatures, respectively.

The specific statistical descriptions of the weather information considered in this study are provided in Table 2.

Table 2. Weather variables used for machine learning model development.

Year	Growth Stage	GDD(℃)	APP (mm)	D	ADTmax(℃)	ADTmin(℃)	ADAT(℃)	ADT(℃)
	BBCH23	988.55	79.60	167	15.45	0.90	258.10	8.60
	BBCH39	1177.95	161.80	182	16.50	6.65	360.65	12.02
2014-2015	BBCH56	1400.45	185.80	196	18.90	9.10	414.25	13.81
	BBCH65	1650.45	209.00	209	22.80	9.65	507.70	16.92
	BBCH75	1994.95	214.00	226	24.00	15.65	592.35	19.75
2015-2016	BBCH65	1725.05	145.00	213	24.20	13.50	541.35	18.05
	BBCH75	2047.90	162.80	230	24.20	11.80	578.60	19.29
	BBCH23	1159.25	160.50	176	11.80	2.85	226.65	7.56
2016-2017	BBCH39	1422.75	193.80	195	19.05	4.65	344.55	11.49
	BBCH56	1649.45	210.20	209	19.80	9.15	447.50	14.92
	BBCH23	1098.00	190.70	179	20.15	4.35	356.00	11.87
2017-2018	BBCH39	1415.55	244.70	199	20.80	8.35	461.00	15.37
	BBCH65	1768.95	255.60	218	22.95	8.35	518.90	17.30
	BBCH75	2081.35	288.90	233	27.05	14.80	585.05	19.50
2021-2022	BBCH56	1589.90	171.10	207	22.85	9.30	452.75	15.09
	BBCH65	1806.00	192.60	219	24.50	9.30	506.65	16.89
2022-2023	BBCH56	1598.40	196.50	210	21.95	5.35	424.45	14.15
	BBCH65	1863.25	240.50	226	21.95	5.35	471.35	15.71

2.3 Proximal hyperspectral sensing data preprocessing

Data processing of proximal hyperspectral sensors primarily involves smoothing with the Savitzky-Golay

Proceedings of the 16th International Conference on Precision Agriculture
21-24 July, 2024, Manhattan, Kansas, United States

5

(SG) filter, resulting in the original (OR) spectrum. To ensure the validity of proximal hyperspectral sensing data and mitigate the impact of human interference factors and the spectral absorption interval of water during measurement, further processing of the OR spectrum is necessary. In essence, the starting band of 350-400 nm (instability of initial band electromagnetic radiation signal) and the interference in water absorption bands (1360-1420 nm, 1790-1950 nm, 2350-2500 nm) were systematically excluded (Berger et al. 2020a). This exclusion left a total of 1729 bands in the 400-2350 nm range for subsequent studies. Spectral preprocessing in this study included the first-order derivative (FD), apparent absorption (LOG), and continuum removal (CR) techniques. These methods were employed to minimize the noise effects arising from changes in light due to soil background, atmospheric scattering, and alterations in field geometry (Figure 3). The calculated formulas for FD, LOG, and CR are shown as follows:

$$FD_i = \frac{R_{i+1} - R_{i-1}}{2\Delta\lambda} \tag{2}$$

$$LOG_i = \log\left(\frac{1}{R_i}\right) \tag{3}$$

$$CR_i = {R_i}/{R_c} \tag{4}$$

where i refers to the wavelength of band i, FD_i, LOG_i and CR_i refers to FD, LOG and CR reflectance corresponding to wavelength i, respectively. R_c refers to the reflectance of continuum line. $\Delta\lambda$ is the interval between adjacent wavelengths. R_i is the reflectance of the i band.
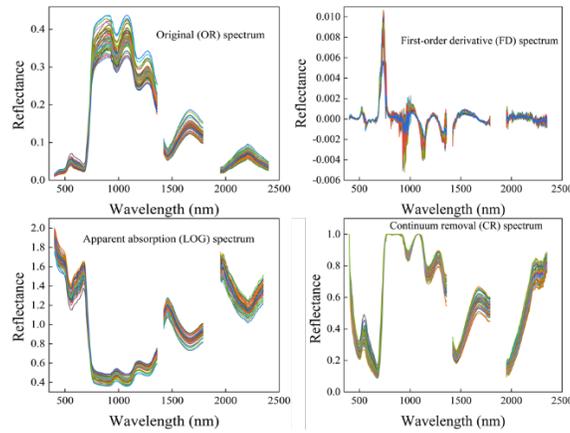

Figure 3. Different forms of proximal hyperspectral sensing data preprocessing.

2.4 Effective bands identification by LASSO

The LASSO method was initially introduced by Tibshirani (Tibshirani 1996). LASSO filters variables by compressing the regression coefficients associated with each variable through a penalty function until the residual sum of squares of the model is minimized. Ultimately, for variables with minimal correlation and uncorrelated variables, the regression coefficient is compressed to zero, leading to the exclusion of the corresponding variables. This process significantly enhances the interpretability of the model. In this study, LASSO was employed to identify the effective bands from various forms of proximal hyperspectral sensing data preprocessing. Ten-fold cross-validation was utilized to determine the optimal set of effective variables. The implementation of LASSO relied on the "glmnet" package in R language.

2.5 Machine learning (ML) algorithms

In this study, ML algorithms were implemented using the caret package within the Rstudio programming environment. The process began by applying the createDataPartition function to meticulously partition the dataset, allocating 75% for the training set (552 data points) and reserving the remaining 25% for the test set (180 data points).

Seven ML algorithms were employed to assess winter wheat PNC, including RFR, SVR, K-Nearest Neighbors Regression (KNNR), Partial Least Squares Regression (PLSR), Gradient Boosting Decision Tree Regression (GBDTR), Elastic Net Regression (ENR), and Decision Tree Regression (DTR). Detailed introductions to three ML algorithms, RFR, SVR, and PLSR, can be found in our previous publication (Chen et al. 2023b). KNNR is utilized for predicting continuous or numerical output values, relying on the proximity of data points within a feature space. Predictions for new data points in KNNR are established by averaging the target values of their K-nearest neighbors within the training dataset. GBDTR is an ensemble learning method combining the predictive power of decision trees with boosting techniques to create highly accurate

Proceedings of the 16th International Conference on Precision Agriculture
21-24 July, 2024, Manhattan, Kansas, United States

6

and robust regression models. ENR is a linear regression technique employed for constructing predictive models, feature selection, and addressing issues in high-dimensional datasets. It harnesses the combined benefits of Ridge Regression and LASSO Regression. DTR builds a regression model in the form of a tree structure, decomposing the dataset into smaller subsets while progressively developing associated decision trees. The result is a tree with decision nodes and leaf nodes. Hyperparameter tuning for the seven ML algorithms involved grid search, followed by ten-fold cross-validation on the calibration set. The optimal parameter or parameter combination was determined based on the highest coefficient of determination (R^2) and lowest root mean square error (RMSE)and mean relative error (MRE) in the calibration set (R^2_{cv}, $RMSE_{cv}$ and MRE_{cv}). These determined hyperparameters were then used to calibrate each prediction model on the entire calibration dataset, and the accuracy was evaluated on an independent test dataset. For specific hyperparameters and range sizes adjusted by each ML algorithm, refer to the research of Ruan et al. (Ruan et al. 2022) and Wang et al. (Wang et al. 2018).

The accuracy of the winter wheat PNC estimation models was evaluated using the R^2, RMSE, and residual prediction deviation (RPD). RPD is calculated as the ratio between the standard deviation (SD) of the test set and RMSE. RPD values can be interpreted as follows: RPD < 1.40 indicates poor prediction performance; 1.40 < RPD < 2.00 suggests rough prediction ability; RPD > 2.00 signifies excellent prediction ability of the model (Chen et al. 2023a). Additionally, the workflow of this study is illustrated in Figure 4.
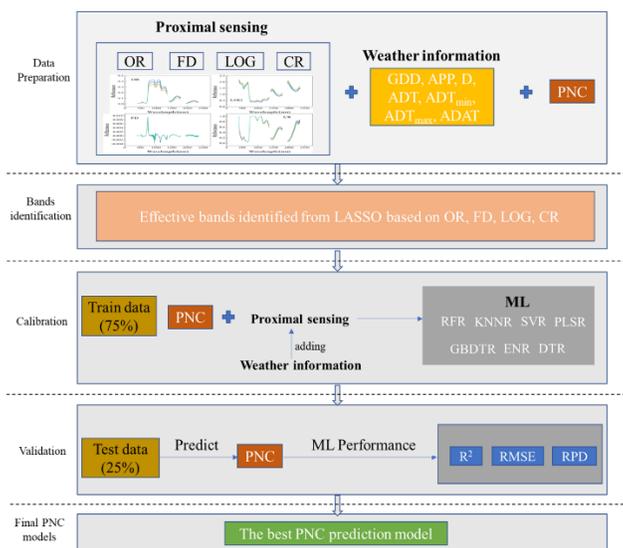

Figure 4. The workflow for predicting winter wheat PNC in this study.

3. Results

3.1 Descriptive statistics

As shown in Table 3, PNC values ranged from 0.47% to 3.25% in the calibration dataset, with an average of 1.46% and a standard deviation of 0.54. In the validation dataset, PNC values ranged from 0.71% to 3.14%, with an average of 1.5%, and a standard deviation of 0.59. Field sampling covered the entire growth period of winter wheat. Consequently, large differences existed between the maximum and minimum PNC values. The distribution range of PNC in the validation set was within the calibration dataset range, affirming the suitability of the dataset division. In summary, the PNC data in both the calibration and validation datasets were well-suited for subsequent studies on evaluating different prediction ML models.

Table 3. Descriptive statistics of winter wheat PNC (%) in this study.

Data sets	Number of samples	Maximum (%)	Minimum (%)	Average	SD
All	732	3.25	0.47	1.46	0.54
Calibration	552	3.25	0.47	1.44	0.53
Validation	180	3.14	0.71	1.5	0.59

3.2 Identification of effective bands by LASSO

Table 4 illustrates the number of effective bands identified through OR, FD, LOG, and CR. The minimum number of effective bands distinguished from OR was 42, covering the visible light range (400-518nm), near-infrared (1037-1154nm), and shortwave infrared (1336-1751nm and 1966-2311nm) regions (Fig. 5).

**Proceedings of the 16th International Conference on Precision Agriculture
21-24 July, 2024, Manhattan, Kansas, United States**

7

In contrast, the maximum number of effective bands identified from FD was 162, located in the visible light range (403-669nm), near-infrared (767-1258nm), and shortwave infrared (1349-2338nm) regions (Fig. 5). LOG yielded 67 effective bands, mainly distributed in the visible light range (400-680nm), near-infrared (1033-1154nm), and shortwave infrared (1336-2337nm) regions (Fig. 5). Additionally, CR identified 79 effective bands, predominantly spanning the visible light range (401-673nm), near-infrared (764-1176nm), and shortwave infrared (1353-2350nm) regions (Figure 5).

Table 4. The number of effective bands under different processed proximal sensing data.

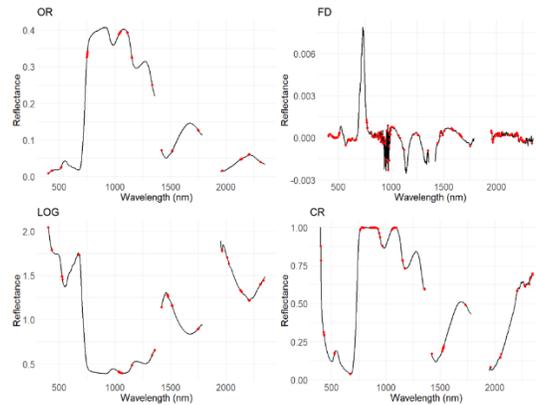| Spectrum transform	Number of full bands	Number of effective bands
OR	1729	42
FD	1729	162
LOG	1729	67
CR	1729	79

Figure 5. The distribution of effective bands for different processed proximal sensing data.

3.4 Estimating PNC with canopy hyperspectral sensing data

When proximal hyperspectral sensing data were used as input variables to estimate winter wheat PNC, the R^2 and RMSE values of various ML algorithms applied to different preprocessed proximal sensing data are presented in Table 5 (ML_{OR}, ML_{FD}, ML_{LOG}, ML_{CR}). For ML_{OR}, the R^2 ranged from 0.33 to 0.63, with RMSE values between 0.36 and 0.49. The $GBDTR_{OR}$ model yielded the best result (R^2 = 0.63, RMSE = 0.36, Figure 6 (a)). In the case of ML_{FD}, the R^2 ranged from 0.43 to 0.80, with RMSE values between 0.27 and 0.46. The SVR_{FD} model performed the best (R^2 = 0.80, RMSE = 0.27, Figure 6 (b)). For ML_{LOG}, the R^2 ranged from 0.34 to 0.74, with RMSE values between 0.30 and 0.48. The ENR_{LOG} model achieved the best performance (R^2 = 0.74, RMSE = 0.34, Figure 6 (c)). Finally, for ML_{CR}, the R^2 ranged from 0.53 to 0.76, with RMSE values between 0.30 and 0.40. The SVR_{CR} model had the best result (R^2 = 0.76, RMSE = 0.30, Figure 6 (d)). In summary, SVR_{FD} demonstrated the best performance in predicting PNC, with an increase of 0.17-0.47 in R^2 and a decrease of 0.09-0.22 in RMSE compared with ML_{OR}.

Table 5. The R2 and the RMSE values of ML algorithms based on different preprocessed proximal hyperspectral sensing data.

Type	OR		FD		LOG		CR	
ML	R^2	RMSE	R^2	RMSE	R^2	RMSE	R^2	RMSE
RFR	0.59	0.40	0.75	0.32	0.55	0.40	0.75	0.31
SVR	0.55	0.47	0.80	0.27	0.62	0.38	0.76	0.30
KNNR	0.47	0.43	0.47	0.44	0.48	0.43	0.72	0.32
GBDTR	0.63	0.36	0.77	0.28	0.63	0.36	0.70	0.32
PLSR	0.43	0.45	0.45	0.44	0.44	0.44	0.61	0.39
ENR	0.61	0.37	0.78	0.28	0.74	0.30	0.75	0.30
DTR	0.33	0.49	0.43	0.46	0.34	0.48	0.53	0.40

Proceedings of the 16th International Conference on Precision Agriculture
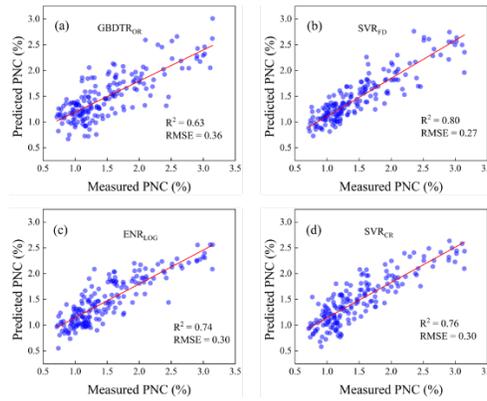**21-24 July, 2024, Manhattan, Kansas, United States**

8

Figure 6. The scatter plots of measured and predicted PNC for different machine learning models using preprocessed proximal hyperspectral sensor data. (The unit of RMSE is %).

3.5 Estimating PNC by combining proximal hyperspectral sensing and weather information

Weather variables were combined with hyperspectral sensing data to construct PNC estimation ML models (ML_{OR+W}, ML_{FD+W}, ML_{LOG+W}, ML_{CR+W}). For the ML_{OR+W} model, the R^2 ranged from 0.76 to 0.83, with RMSE between 0.24 and 0.30. The $GBDTR_{OR+W}$ model produced the most favorable result (R^2 = 0.83, RMSE = 0.24, Figure 7 (a)). Similarly, the ML_{FD+W} model exhibited an R^2 range of 0.76-0.85 and RMSE of 0.23-0.30, with the RFR_{FD+W} model achieving the highest performance (R^2 = 0.85, RMSE = 0.23, Figure 7 (b)). The ML_{LOG+W} model demonstrated an R^2 range of 0.76-0.84 and RMSE of 0.24-0.30. The ENR_{LOG+W} model performed the best (R^2 = 0.84, RMSE = 0.24, Figure 7 (c)). Lastly, the R^2 and RMSE of the ML_{CR+W} model ranged from 0.76-0.85 and 0.23-0.30, respectively. The ENR_{CR+W} model yielded the best result (R^2 = 0.85, RMSE = 0.23, Figure 7 (d)). More information on the performance of different models can be found in Table 6.

Table 6. The R2 and RMSE values of different machine learning models combining proximal hyperspectral sensing and weather information.

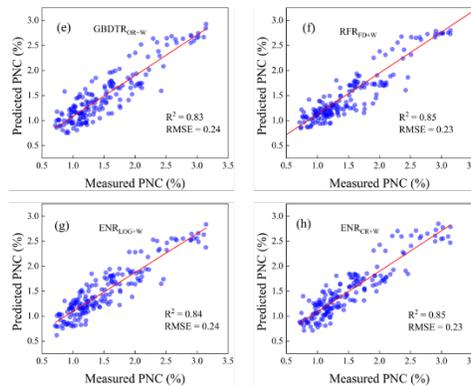| Type	OR+W		FD+W		LOG+W		CR+W	
ML	R^2	RMSE	R^2	RMSE	R^2	RMSE	R^2	RMSE
RFR	0.82	0.25	0.85	0.23	0.81	0.25	0.84	0.24
SVR	0.81	0.26	0.84	0.24	0.79	0.27	0.82	0.25
KNNR	0.81	0.25	0.82	0.25	0.82	0.25	0.83	0.24
GBDTR	0.83	0.24	0.85	0.24	0.82	0.25	0.84	0.24
PLSR	0.76	0.30	0.76	0.30	0.76	0.30	0.76	0.30
ENR	0.82	0.26	0.83	0.25	0.84	0.24	0.85	0.23
DTR	0.79	0.27	0.80	0.26	0.80	0.26	0.80	0.27

Figure 7. The scatter plots of measured and predicted PNC for best performing machine learning models using preprocessed hyperspectral sensing data and weather information (The unit of RMSE is %).

3.6 Comparison of different PNC prediction strategies

In order to further compare the accuracy of the above PNC estimation models, the RPD values of these models were calculated (Figure 8). The ML_{FD} (RPD = 1.30 to 2.13), ML_{LOG} (RPD = 1.23 to 1.94), and ML_{CR}

Proceedings of the 16th International Conference on Precision Agriculture
21-24 July, 2024, Manhattan, Kansas, United States

9

(RPD = 1.46 to 1.97) were superior to ML_{OR} (RPD = 1.22 to 1.65). SVR_{FD} model had the highest RPD (2.16). However, the results showed that the accuracy of ML models combining proximal hyperspectral sensing with weather information were significantly improved (RPD = 1.99 to 2.56) compared with the models using proximal hyperspectral sensing alone (RPD = 1.22 to 2.16), with the accuracy of RFR_{FD+W} model being the highest (RPD = 2.56), followed by ENR_{CR+W} model (RPD = 2.54).
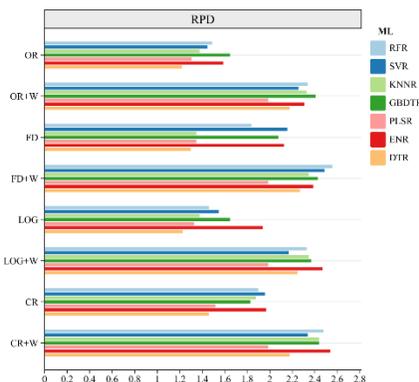

Figure 8. The relative prediction deviation (RPD) values of seven machine learning models combining preprocessed hyperspectral sensing data and weather variables for PNC prediction.

4. Discussion

4.1 The necessity of estimating crop N status with the introduction of the short-wave infrared region

Physiologically, N in plants is used to produce proteins and chlorophyll stored in leaf cells. Notably, chlorophyll represents only a small fraction of leaf N (Homolová et al. 2013), with proteins being the primary N-containing biochemical component. However, most research studies relied on the high correlation between N and chlorophyll content to assess crop N status, predominantly within the visible infrared spectral domain. This approach tends to neglect the redistribution of N concentrated in protein-related structures and non-photosynthetic regions (Berger et al. 2020b), particularly in the short-wave infrared region. To the best of our knowledge, the correlation between N and chlorophyll content in the entire ecosystem is moderately strong, with a correlation coefficient ranging 0.5-0.65 (Homolová et al. 2013). This correlation may stem from the dynamic nature of vegetation growth, characterized by a continuous process of N turnover (Kattge 2002), During this dynamic process, N is transported between organs without relying on chlorophyll. In the context of wheat, N initially binds to wheat leaves during the early stages of growth. As the growth period advances towards the reproduction stage, there is a discernible transfer or redistribution of N from leaves to reproductive structures, such as seeds, ears, or fruits (Ohyama 2010)，Consequently, the correlation between chlorophyll content and N decreases at this stage (Berger et al. 2020a). In contrast, protein emerges as a robust proxy for crop N content, given that rubisco contains up to 50% N in green leaves (Verrelst et al. 2021). Previous studies have highlighted the benefits of using shortwave infrared wavelengths associated with proteins to monitor crop N status (Dunn et al. 2016, Herrmann et al. 2010). Therefore, we propose to focus on protein as the primary N component in leaves and plants and estimate crop N status by the introduction of the short-wave infrared region, capitalizing on the advantages of chlorophyll and protein as effective substitutes for N content.

Hyperspectral sensing technology typically refers to passive sensing capturing subtle characteristics related to physicochemical parameters of crops in continuous narrow bands. However, the utilization of a large number of bands can introduce the curse of dimensionality, giving rise to collinear effects. Additionally, the curse of dimensionality prompts the regression model to intricately train on the data, learning noise in the dataset and consequently reducing the prediction accuracy of the regression model. Therefore, the LASSO method was employed in this study to identify effective bands in full-band proximal hyperspectral data to estimate PNC in winter wheat. Following LASSO processing, the number of estimated effective bands for PNC (Table 3) significantly reduced to 42 (OR), 162 (FD), 67 (LOG), and 79 (CR), based on different preprocessing methods. These results underscore the effectiveness of the LASSO method, which combines feature selection and regularization, mitigating the curse of dimensionality while reducing collinearity and overfitting. Song et al. (Song and Wang 2023b) reached similar conclusions when estimating leaf photosynthetic capacity by combining LASSO with hyperspectral reflectance data. In this study, the effective bands for PNC identified by LASSO are distributed across the visible, near-infrared, and

Proceedings of the 16th International Conference on Precision Agriculture
21-24 July, 2024, Manhattan, Kansas, United States

10

short-wave infrared regions (Figure 5). This further substantiates the necessity of monitoring PNC with the introduction of the short-wave infrared region.

4.2 Combining proximal hyperspectral sensing and weather information to improve winter wheat PNC prediction using ML

Despite employing diverse preprocessing techniques for proximal hyperspectral sensing data and utilizing a combination of ML algorithms for estimating winter wheat PNC, accuracy still needed improvement (Figure 6 and Figure 8). Previous research indicated that the prediction of maize N Nutrition Index (NNI) was improved by combining proximal sensing data with environmental and management information (Dong et al. 2021a). Soil sampling and testing is a time-consuming and labor-intensive process. In contrast, weather information like temperature and precipitation can be easily obtained (Dong et al. 2021a). Extreme temperatures, whether high or low, impose constraints on crop growth. Simultaneously, precipitation plays a crucial role in influencing crop growth, N transport, and the formation of yields. Moreover, changes of temperature and precipitation can significantly influence soil microbial activities and plant root metabolism. For instance, reduced soil water content hampers the capacity of plant roots to absorb and transport N, consequently impacting the overall crop N status (Jiang et al. 2022, Kirschbaum 1995, Naylor et al. 2020). The study of Jiang et al. (Jiang et al. 2022) demonstrated the efficacy of integrating Unmanned Aerial Vehicles (UAVs) multi-spectral remote sensing data with weather information like temperature and precipitation as well as field management data for diagnosing N status of winter wheat at a field scale. To enhance the prediction accuracy of winter wheat PNC, this study explored the inclusion of weather information, such as temperature and precipitation, as supplementary information in the ML algorithms (Figure 7). When weather information was combined with the proximal hyperspectral sensing data, the accuracy of winter wheat PNC prediction was significantly improved, with an increase of 0.04-0.46 in R^2, a decrease of 0.03-0.22 in RMSE and an increase of 0.26-1.02 in RPD.

The integration of remote sensing, environmental, and field management information using ML has become pivotal in predicting both crop yield (Gopal and Bhargavi 2019, Kang et al. 2020, Zhang et al. 2019) and N status (Dong et al. 2021a, Grinberg et al. 2020). In this study, seven ML algorithms were compared for their performance in predicting winter wheat PNC. The RFR model exhibited the highest accuracy when incorporating both proximal hyperspectral sensing and weather information. RFR leverages ensemble learning, a strategy that combines predictions from numerous weak learners to produce a more robust and accurate overall prediction (Genuer et al. 2017). Each decision tree within the ensemble acts as a weak learner, potentially prone to overfitting on specific data subsets. However, the strength of RFR lies in the amalgamation of these trees, compensating for individual weaknesses and thereby improving the collective predictive capability (Feng et al. 2023). Notably, each decision tree in RFR operates as a non-parametric model, effectively capturing irregular data patterns and adapting to intricate relationships. This versatility makes RFR well-suited for handling high-dimensional data and nonlinear associations (Liang et al. 2015). Moreover, RFR is resilient against overfitting. By constructing multiple decision trees and averaging their outcomes, the model mitigates the risk of overly tailoring itself to the training data to some extent. This feature contributes to the model's effectiveness in making accurate predictions on novel, unseen data. In previous studies, many scholars have arrived at similar conclusions when utilizing RFR for the estimation of crop physicochemical parameters (Jiang et al. 2022, Khanal et al. 2018, Zha et al. 2020b).

4.3 Preprocessing methods for proximal hyperspectral sensing

The hyperspectral data of winter wheat canopy inevitably undergo influences from atmospheric conditions, soil background and variations in illumination (Pu 2017). The application of spectral preprocessing techniques enables the partial disentanglement of subtle information within spectra, mitigating background interferences and consequently elevating the prediction precision of crop growth parameters (Kokaly and Clark 1999). Therefore, three different preprocessing techniques, including FD, LOG, and CR, were applied to OS in this study. After the identification of effective bands through LASSO, these preprocessed data were utilized as input variables in conjunction with ML algorithms to estimate winter wheat PNC. As shown in Figure 6 and 8, SVR$_{FD}$ yield satisfactory PNC prediction accuracy compared with ML$_{OR}$. This could be attributed to the fact that FD, compared to other preprocessing methods, has the capability to eliminate various background interferences, emphasize sensitive spectral information, and better suppress the impact of low-frequency components related to illumination variations (Demetriades-Shah et al. 1990, Imanishi et al. 2004, Tsai and Philpot 1998). Previous research has also validated the effectiveness of the first derivative spectra in estimating parameters such as light interception in deciduous forests (Jin et al. 2020, Jin et al. 2022) and chlorophyll content in the canopy of winter wheat (Chen et al. 2023c). Our study affirms the advantages of FD spectral preprocessing techniques in diagnosing winter wheat PNC.

4.4 Challenges and future works

Proceedings of the 16th International Conference on Precision Agriculture
21-24 July, 2024, Manhattan, Kansas, United States

11

In this study, the optimal monitoring model explained 85% of the variation in winter wheat PNC, with an RPD of 2.56, suggesting the potential for further improvement in the model. Future research could explore the integration of more diverse datasets into the PNC prediction model, such as the Shannon Diversity Index (SDI), Average Daily Sunshine Hours (ASH), Sowing Density (SD), N Ratio (NR), and Soil pH (SP), etc. Incorporating such multi-source data can enhance the accuracy and robustness of the ML models, as demonstrated for precision N management (Dong et al. 2021a, Li et al. 2022a, Li et al. 2022b, Lu et al. 2022a). However, for regional-scale crop N status monitoring, the proximal hyperspectral sensing, being point measurements, is still time-consuming and laborious. The UAV remote sensing technology can somewhat overcome these constraints, improving monitoring efficiency and coverage. More studies are needed to develop UAV or satellite hyperspectral remote sensing-based multi-source data fusion strategies using ML for more efficient crop N status monitoring over a large area.

5. Conclusion

This study conducted six site-year field experiments on winter wheat under various N fertilizer treatments. Seven ML algorithms, in conjunction with different preprocessing techniques for proximal hyperspectral sensing data and various weather variables were employed to establish models for predicting winter wheat PNC. The key results indicated that effective bands of FD combined with SVR yield satisfying PNC prediction accuracy ($R^2 = 0.80$, RMSE = 0.27, and RPD = 2.16). When weather information was combined with the proximal hyperspectral sensing data, the accuracy of winter wheat PNC prediction was significantly improved, with an increase of 0.04-0.46 in R^2, a decrease of 0.03-0.22 in RMSE and an increase of 0.26-1.02 in RPD. Random forest combining FD and weather information yielded the best PNC predictions ($R^2 = 0.85$, RMSE = 0.23, and RPD = 2.56). It is concluded that the RFR model combining FD of proximal hyperspectral sensing data and weather information can significantly improve winter wheat PNC prediction compared with using hyperspectral sensing data alone and is a promising and practical strategy to predict winter wheat PNC.

Acknowledgements

References

Bean, G., Kitchen, N., Camberato, J., Ferguson, R., Fernandez, F., Franzen, D., Laboski, C., Nafziger, E., Sawyer, J. and Scharf, P. (2018) Improving an active-optical reflectance sensor algorithm using soil and weather information. Agronomy Journal 110(6), 2541-2551.

Berger, K., Verrelst, J., Feret, J.B., Hank, T., Wocher, M., Mauser, W. and Camps-Valls, G. (2020a) Retrieval of aboveground crop nitrogen content with a hybrid machine learning method. Int J Appl Earth Obs Geoinf 92, 102174.

Berger, K., Verrelst, J., Feret, J.B., Wang, Z., Wocher, M., Strathmann, M., Danner, M., Mauser, W. and Hank, T. (2020b) Crop nitrogen monitoring: Recent progress and principal developments in the context of imaging spectroscopy missions. Remote Sens Environ 242, 111758.

Cao, C., Wang, T., Gao, M., Li, Y., Li, D. and Zhang, H. (2021) Hyperspectral inversion of nitrogen content in maize leaves based on different dimensionality reduction algorithms. Computers and electronics in agriculture 190.

Chen, X., Li, F. and Chang, Q. (2023a) Combination of Continuous Wavelet Transform and Successive Projection Algorithm for the Estimation of Winter Wheat Plant Nitrogen Concentration. Remote Sensing 15(4).

Chen, X., Li, F., Shi, B. and Chang, Q. (2023b) Estimation of Winter Wheat Plant Nitrogen Concentration from UAV Hyperspectral Remote Sensing Combined with Machine Learning Methods. Remote Sensing 15(11).

Chen, X., Li, F., Shi, B., Fan, K., Li, Z. and Chang, Q. (2023c) Estimation of Winter Wheat Canopy Chlorophyll Content Based on Canopy Spectral Transformation and Machine Learning Method. Agronomy 13(3), 783.

Chen, Z., Miao, Y., Lu, J., Zhou, L., Li, Y., Zhang, H., Lou, W., Zhang, Z., Kusnierek, K. and Liu, C. (2019) In-season diagnosis of winter wheat nitrogen status in smallholder farmer fields across a village using unmanned aerial vehicle-based remote sensing. Agronomy 9(10), 619.

Chlingaryan, A., Sukkarieh, S. and Whelan, B. (2018) Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: A review. Computers and electronics in agriculture 151, 61-69.

Demetriades-Shah, T.H., Steven, M.D. and Clark, J.A. (1990) High resolution derivative spectra in remote sensing. Remote sensing of environment 33(1), 55-64.

Proceedings of the 16th International Conference on Precision Agriculture
21-24 July, 2024, Manhattan, Kansas, United States

12

Dong, R., Miao, Y., Wang, X., Chen, Z. and Yuan, F. (2021a) Improving maize nitrogen nutrition index prediction using leaf fluorescence sensor combined with environmental and management variables. Field Crops Research 269.

Dong, R., Miao, Y., Wang, X., Yuan, F. and Kusnierek, K. (2021b) Canopy Fluorescence Sensing for In-Season Maize Nitrogen Status Diagnosis. Remote Sensing 13(24).

Dunn, B.W., Dehaan, R., Schmidtke, L., Dunn, T. and Meder, R. (2016) Using field-derived hyperspectral reflectance measurement to identify the essential wavelengths for predicting nitrogen uptake of rice at panicle initiation. Journal of Near Infrared Spectroscopy 24(5), 473-483.

Feng, Z., Guan, H., Yang, T., He, L., Duan, J., Song, L., Wang, C. and Feng, W. (2023) Estimating the canopy chlorophyll content of winter wheat under nitrogen deficiency and powdery mildew stress using machine learning. Computers and electronics in agriculture 211, 107989.

Fu, Y., Yang, G., Pu, R., Li, Z., Li, H., Xu, X., Song, X., Yang, X. and Zhao, C. (2021) An overview of crop nitrogen status assessment using hyperspectral remote sensing: Current status and perspectives. European Journal of Agronomy 124.

Genuer, R., Poggi, J.-M., Tuleau-Malot, C. and Villa-Vialaneix, N. (2017) Random forests for big data. Big Data Research 9, 28-46.

Gopal, P.M. and Bhargavi, R. (2019) A novel approach for efficient crop yield prediction. Computers and electronics in agriculture 165, 104968.

Grinberg, N.F., Orhobor, O.I. and King, R.D. (2020) An evaluation of machine-learning for predicting phenotype: studies in yeast, rice, and wheat. Machine Learning 109, 251-277.

Hansen, P. and Schjoerring, J. (2003) Reflectance measurement of canopy biomass and nitrogen status in wheat crops using normalized difference vegetation indices and partial least squares regression. Remote sensing of environment 86(4), 542-553.

Herrmann, I., Karnieli, A., Bonfil, D., Cohen, Y. and Alchanatis, V. (2010) SWIR-based spectral indices for assessing nitrogen content in potato fields. International Journal of Remote Sensing 31(19), 5127-5143.

Homolová, L., Malenovský, Z., Clevers, J.G., García-Santos, G. and Schaepman, M.E. (2013) Review of optical-based remote sensing for plant trait mapping. Ecological Complexity 15, 1-16.

Huang, Z., Turner, B.J., Dury, S.J., Wallis, I.R. and Foley, W.J. (2004) Estimating foliage nitrogen concentration from HYMAP data using continuum removal analysis. Remote sensing of environment 93(1-2), 18-29.

Imanishi, J., Sugimoto, K. and Morimoto, Y. (2004) Detecting drought status and LAI of two Quercus species canopies using derivative spectra. Computers and electronics in agriculture 43(2), 109-129.

Jiang, J., Atkinson, P.M., Zhang, J., Lu, R., Zhou, Y., Cao, Q., Tian, Y., Zhu, Y., Cao, W. and Liu, X. (2022) Combining fixed-wing UAV multispectral imagery and machine learning to diagnose winter wheat nitrogen status at the farm scale. European Journal of Agronomy 138.

Jin, J., Arief Pratama, B. and Wang, Q. (2020) Tracing leaf photosynthetic parameters using hyperspectral indices in an alpine deciduous forest. Remote Sensing 12(7), 1124.

Jin, J., Wang, Q. and Song, G. (2022) Selecting informative bands for partial least squares regressions improves their goodness-of-fits to estimate leaf photosynthetic parameters from hyperspectral data. Photosynthesis Research, 1-12.

Kang, Y., Ozdogan, M., Zhu, X., Ye, Z., Hain, C. and Anderson, M. (2020) Comparative assessment of environmental variables and machine learning algorithms for maize yield prediction in the US Midwest. Environmental Research Letters 15(6), 064005.

Kattge, J. (2002) Zur Bedeutung von Stickstoff für den CO2-Düngeeffekt.

Khanal, S., Fulton, J., Klopfenstein, A., Douridas, N. and Shearer, S. (2018) Integration of high resolution remotely sensed data and machine learning techniques for spatial prediction of soil properties and corn yield. Computers and electronics in agriculture 153, 213-225.

Kirschbaum, M.U. (1995) The temperature dependence of soil organic matter decomposition, and the effect of global warming on soil organic C storage. Soil Biology and biochemistry 27(6), 753-760.

Kokaly, R.F. and Clark, R.N. (1999) Spectroscopic determination of leaf biochemistry using band-depth analysis of absorption features and stepwise multiple linear regression. Remote sensing of environment 67(3), 267-287.

Küçük, Ç., Taşkın, G. and Erten, E. (2016) Paddy-rice phenology classification based on machine-learning methods using multitemporal co-polar X-band SAR images. IEEE Journal of selected topics in applied earth observations and remote sensing 9(6), 2509-2519.

Kukreja, S.L., Löfberg, J. and Brenner, M.J. (2006) A least absolute shrinkage and selection operator (LASSO) for nonlinear system identification. IFAC proceedings volumes 39(1), 814-819.

Kumar, S., Attri, S. and Singh, K. (2019) Comparison of Lasso and stepwise regression technique for wheat yield prediction. Journal of Agrometeorology 21(2), 188-192.

LaCapra, V., Melack, J., Gastil, M. and Valeriano, D. (1996) Remote sensing of foliar chemistry of inundated

Proceedings of the 16th International Conference on Precision Agriculture
21-24 July, 2024, Manhattan, Kansas, United States

13

rice with imaging spectrometry. Remote sensing of environment 55(1), 50-58.

Lemaire, G., Jeuffroy, M.-H. and Gastal, F. (2008) Diagnosis tool for plant and crop N status in vegetative stage: Theory and practices for crop N management. European Journal of Agronomy 28(4), 614-624.

Li, D., Miao, Y., Ransom, C.J., Bean, G.M., Kitchen, N.R., Fernández, F.G., Sawyer, J.E., Camberato, J.J., Carter, P.R., Ferguson, R.B., Franzen, D.W., Laboski, C.A.M., Nafziger, E.D. and Shanahan, J.F. (2022a) Corn Nitrogen Nutrition Index Prediction Improved by Integrating Genetic, Environmental, and Management Factors with Active Canopy Sensing Using Machine Learning. Remote Sensing 14(2).

Li, F., Miao, Y., Hennig, S.D., Gnyp, M.L., Chen, X., Jia, L. and Bareth, G. (2010) Evaluating hyperspectral vegetation indices for estimating nitrogen concentration of winter wheat at different growth stages. Precision Agriculture 11, 335-357.

Li, F., Mistele, B., Hu, Y., Chen, X. and Schmidhalter, U. (2013) Comparing hyperspectral index optimization algorithms to estimate aerial N uptake using multi-temporal winter wheat datasets from contrasting climatic and geographic zones in China and Germany. Agricultural and Forest Meteorology 180, 44-57.

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 Journal of Photogrammetry and Remote Sensing 92, 112-123.

Li, F., Wang, L., Liu, J., Wang, Y. and Chang, Q. (2019) Evaluation of Leaf N Concentration in Winter Wheat Based on Discrete Wavelet Transform Analysis. Remote Sensing 11(11).

Li, Y., Miao, Y., Zhang, J., Cammarano, D., Li, S., Liu, X., Tian, Y., Zhu, Y., Cao, W. and Cao, Q. (2022b) Improving Estimation of Winter Wheat Nitrogen Status Using Random Forest by Integrating Multi-Source Data Across Different Agro-Ecological Zones. Frontiers in plant science 13, 890892.

Li, Y., Miao, Y., Zhang, J., Cammarano, D., Li, S., Liu, X., Tian, Y., Zhu, Y., Cao, W. and Cao, Q. (2022c) Improving Estimation of Winter Wheat Nitrogen Status Using Random Forest by Integrating Multi-Source Data Across Different Agro-Ecological Zones. Front Plant Sci 13, 890892.

Liang, L., Di, L., Zhang, L., Deng, M., Qin, Z., Zhao, S. and Lin, H. (2015) Estimation of crop LAI using hyperspectral vegetation indices and a hybrid inversion method. Remote sensing of environment 165, 123-134.

Liu, L., Peng, Z., Zhang, B., Wei, Z., Han, N., Lin, S., Chen, H. and Cai, J. (2019) Canopy nitrogen concentration monitoring techniques of summer corn based on canopy spectral information. Sensors 19(19), 4123.

Liu, S., Bai, X., Zhu, G., Zhang, Y., Li, L., Ren, T. and Lu, J. (2023) Remote estimation of leaf nitrogen concentration in winter oilseed rape across growth stages and seasons by correcting for the canopy structural effect. Remote Sensing of Environment 284.

Lu, J., Dai, E., Miao, Y. and Kusnierek, K. (2022a) Improving active canopy sensor-based in-season rice nitrogen status diagnosis and recommendation using multi-source data fusion with machine learning. Journal of Cleaner Production 380.

Lu, J., Dai, E., Miao, Y. and Kusnierek, K. (2022b) Improving active canopy sensor-based in-season rice nitrogen status diagnosis and recommendation using multi-source data fusion with machine learning. Journal of Cleaner Production 380, 134926.

Miao, Y., Stewart, B.A. and Zhang, F. (2011) Long-term experiments for sustainable nutrient management in China. A review. Agronomy for Sustainable Development 31, 397-414.

Müller, K., Böttcher, U., Meyer-Schatz, F. and Kage, H. (2008) Analysis of vegetation indices derived from hyperspectral reflection measurements for estimating crop canopy parameters of oilseed rape (Brassica napus L.). Biosystems engineering 101(2), 172-182.

Naylor, D., Sadler, N., Bhattacharjee, A., Graham, E.B., Anderton, C.R., McClure, R., Lipton, M., Hofmockel, K.S. and Jansson, J.K. (2020) Soil microbiomes under climate change and implications for carbon cycling. Annual Review of Environment and Resources 45(1), 29-59.

Ohyama, T. (2010) Nitrogen as a major essential element of plants. Nitrogen Assim. Plants 37, 1-17.

Ordonez, C., de la Fuente, M.O., Roca-Pardinas, J. and Rodríguez-Pérez, J.R. (2018) Determining optimum wavelengths for leaf water content estimation from reflectance: a distance correlation approach. Chemometrics and Intelligent Laboratory Systems 173, 41-50.

Padilla, F.M., Gallardo, M., Peña-Fleitas, M.T., De Souza, R. and Thompson, R.B. (2018) Proximal optical sensors for nitrogen management of vegetable crops: A review. Sensors 18(7), 2083.

Pu, R. (2017) Hyperspectral remote sensing: fundamentals and practices, CRC Press.

Ruan, G., Li, X., Yuan, F., Cammarano, D., Ata-Ul-Karim, S.T., Liu, X., Tian, Y., Zhu, Y., Cao, W. and Cao, Q. (2022) Improving wheat yield prediction integrating proximal sensing and weather data with machine learning. Computers and Electronics in Agriculture 195, 106852.

Schepers, A.R., Shanahan, J.F., Liebig, M.A., Schepers, J.S., Johnson, S.H. and Luchiari Jr, A. (2004) Appropriateness of management zones for characterizing spatial variability of soil properties and irrigated corn yields across years. Agronomy Journal 96(1), 195-203.

Proceedings of the 16th International Conference on Precision Agriculture
21-24 July, 2024, Manhattan, Kansas, United States

14

Skiba, U. and Rees, R. (2014) Nitrous oxide, climate change and agriculture. CABI Reviews (2014), 1-7.

Song, G. and Wang, Q. (2023a) Coupling effective variable selection with machine learning techniques for better estimating leaf photosynthetic capacity in a tree species (Fagus crenata Blume) from hyperspectral reflectance. Agricultural and Forest Meteorology 338, 109528.

Song, G. and Wang, Q. (2023b) Coupling effective variable selection with machine learning techniques for better estimating leaf photosynthetic capacity in a tree species (Fagus crenata Blume) from hyperspectral reflectance. Agricultural and Forest Meteorology 338.

Stas, M., Van Orshoven, J., Dong, Q., Heremans, S. and Zhang, B. (2016) A comparison of machine learning algorithms for regional wheat yield prediction using NDVI time series of SPOT-VGT, pp. 1-5, IEEE.

Tibshirani, R. (1996) Regression shrinkage and selection via the lasso. Journal of the Royal Statistical Society Series B: Statistical Methodology 58(1), 267-288.

Tsai, F. and Philpot, W. (1998) Derivative analysis of hyperspectral data. Remote sensing of environment 66(1), 41-51.

Verrelst, J., Rivera-Caicedo, J.P., Reyes-Munoz, P., Morata, M., Amin, E., Tagliabue, G., Panigada, C., Hank, T. and Berger, K. (2021) Mapping landscape canopy nitrogen content from space using PRISMA data. ISPRS J Photogramm Remote Sens 178, 382-395.

Wang, D., Li, R., Liu, T., Liu, S., Sun, C. and Guo, W. (2023) Combining vegetation, color, and texture indices with hyperspectral parameters using machine-learning methods to estimate nitrogen concentration in rice stems and leaves. Field Crops Research 304.

Wang, L., Chang, Q., Yang, J., Zhang, X. and Li, F. (2018) Estimation of paddy rice leaf area index using machine learning methods based on hyperspectral data from multi-year experiments. PLoS One 13(12), e0207624.

Wang, L., Chen, S., Li, D., Wang, C., Jiang, H., Zheng, Q. and Peng, Z. (2021) Estimation of Paddy Rice Nitrogen Content and Accumulation Both at Leaf and Plant Levels from UAV Hyperspectral Imagery. Remote Sensing 13(15).

Wood, C., Reeves, D., Duffield, R. and Edmisten, K. (1992) Field chlorophyll measurements for evaluation of corn nitrogen status. Journal of Plant Nutrition 15(4), 487-500.

Yang, H., Li, F., Hu, Y. and Yu, K. (2021a) Hyperspectral indices optimization algorithms for estimating canopy nitrogen concentration in potato (Solanum tuberosum L.). International Journal of Applied Earth Observation and Geoinformation 102.

Yang, H., Li, F., Hu, Y. and Yu, K. (2021b) Hyperspectral indices optimization algorithms for estimating canopy nitrogen concentration in potato (Solanum tuberosum L.). International Journal of Applied Earth Observation and Geoinformation 102, 102416.

Yang, H., Yin, H., Li, F., Hu, Y. and Yu, K. (2023) Machine learning models fed with optimized spectral indices to advance crop nitrogen monitoring. Field Crops Research 293, 108844.

Yasumura, Y., Hikosaka, K. and Hirose, T. (2007) Nitrogen resorption and protein degradation during leaf senescence in Chenopodium album grown in different light and nitrogen conditions. Functional Plant Biology 34(5), 409-417.

Zha, H., Miao, Y., Wang, T., Li, Y., Zhang, J., Sun, W., Feng, Z. and Kusnierek, K. (2020a) Improving Unmanned Aerial Vehicle Remote Sensing-Based Rice Nitrogen Nutrition Index Prediction with Machine Learning. Remote Sensing 12(2).

Zha, H., Miao, Y., Wang, T., Li, Y., Zhang, J., Sun, W., Feng, Z. and Kusnierek, K. (2020b) Improving unmanned aerial vehicle remote sensing-based rice nitrogen nutrition index prediction with machine learning. Remote Sensing 12(2), 215.

Zhang, L., Zhang, Z., Luo, Y., Cao, J. and Tao, F. (2019) Combining optical, fluorescence, thermal satellite, and environmental data to predict county-level maize yield in China using machine learning approaches. Remote Sensing 12(1), 21.

Proceedings of the 16th International Conference on Precision Agriculture
21-24 July, 2024, Manhattan, Kansas, United States

15