Lin X *et al.* (2023)
Notulae Botanicae Horti Agrobotanici Cluj-Napoca
Volume 51, Issue 1, Article number 13059
DOI:10.15835/nbha51113059
Research Article

A cotton leaf nitrogen monitoring model based on spectral-fluorescence data fusion

Xiu LIN[1], Faduman HA[2], Lulu MA[1], Xiangyu CHEN[1], Yiru MA[1], Aiqun CHEN[3], Zhenan HOU[1]*, Ze ZHANG[1]*

[1]*Shihezi University, College of Agriculture, The Key Laboratory of Oasis Eco-agriculture, Xinjiang Production and Construction Group, Shihezi 832003, China; malulu@stu.shzu.edu.cn; 1849819772@qq.com; chenxinagyu@stu.shzu.edu.cn; mayiru@stu.shzu.edu.cn; 1742329237@qq.com; zhangze1227@shzu.edu.cn (*corresponding author)*
[2]*Xinjiang Agricultural Vocational and Technical College, Changji, 831100, China; 1442100913@qq.com*
[3]*Kekedala Agricultural Technology Popularization Station of the Fourth Division of Xinjiang Production and Construction Corps, Kekedala, 835213, China; malulu_shzu@163.com*

Abstract

In the present study, hyperspectral imaging and remote sensing of fluorescence were integrated to monitor the nitrogen content in leaves of drip-irrigated cotton at different growth periods in northern Xinjiang, China. Based on the spectrum and chlorophyll fluorescence parameters of nitrogen content in cotton leaves of different growth periods obtained through the shuffled frog-leaping algorithm (SFLA), the successive projection algorithm (SPA), grey relational analysis (GRA), and competitive adaptive reweighted sampling (CARS), a monitoring model of nitrogen content in cotton leaves was established via on hyperspectral imaging, chlorophyll fluorescence parameters, and spectral-fluorescence data fusion. The results showed that: (1) there were significant positive correlations between the chlorophyll fluorescence parameters Fv'/Fm', Fv/Fm, Yield, Fm, NPQ, and the nitrogen content at each growth period. (2) The effectiveness of chlorophyll fluorescence parameters in inversion of nitrogen content was the highest at the budding period and the blooming period, and the coefficients of determination (R^2) of the validation sets were 0.745 and 0.709, respectively. (3) In the monitoring model for cotton leaf nitrogen in the blooming period that was established based on the decision-level algorithm and spectral-fluorescence data fusion, the R^2 value of the training set reached 0.961, and that of the validation set was 0.828. In conclusion, the findings of this study suggest that the feature-level fusion and decision-level fusion algorithms of spectral-fluorescence data can effectively improve the accuracy and reliability of cotton leaf nitrogen monitoring.

Keywords: characteristic parameter; chlorophyll fluorescence parameters; cotton leaf nitrogen; hyperspectral imaging; monitoring model

Introduction

Cotton, as a major economic crop of China, is a strategic material (Xu *et al.*, 2021). Nitrogen fertilizer is the most widely used fertilizer in cotton cultivation; it plays an essential role in cotton growth and affects

Received: 27 Dec 2022. Received in revised form: 09 Feb 2023. Accepted: 14 Feb 2023. Published online: 29 Mar 2023.
From Volume 49, Issue 1, 2021, Notulae Botanicae Horti Agrobotanici Cluj-Napoca journal uses article numbers in place of the traditional method of continuous pagination through the volume. The journal will continue to appear quarterly, as before, with four annual numbers.

both yield and quality (Wang *et al*., 2015; Li *et al*., 2016; Liu *et al*., 2021; Zhong *et al*., 2021). Studies have shown that proper application of nitrogen fertilizer can effectively increase the yield of crops, yet excessive application can lead to inconsistent growth, late flowering, low crop yield and quality (Barraclough *et al*., 2010; Jay *et al*., 2017). Moreover, excess usage also wastes resources and potentially damages the environment (Serrano *et al*., 2015; Halihash, 2016). Hence, it is important to achieve accurate and efficient monitoring of nitrogen demand during cotton growth, thereby determining nitrogen deficiency (Suo, 2016). An efficient nitrogen content monitoring method is necessary for formulating precise and efficient nitrogen fertilizer application plans.

Conventional crop nutrition monitoring methods have the limitations of being slow, time-consuming, and labor-intensive (Li *et al*., 2016; Fei *et al*., 2021) and can damage the crop leaves. Hence, such methods are not suitable for long-term repeated crop monitoring and do not meet the requirements of precision agriculture for high reliability and rapid deployment (He, 2016). Individual monitoring methods such as those based on soil plant analysis development (SPAD) values to predict the nutritional status of crops have the problems of poor stability and low degree of fitting (Shi *et al*., 2018). Image features-based crop nitrogen monitoring methods involve destructive sampling processes and use heavy instruments that are not easy to carry (Yang *et al*., 2020). Therefore, there is an urgent need for innovative technologies that can overcome the limitations of conventional monitoring methods and quickly and accurately predict the nutritional demands of crops. In recent years, development of hyperspectral remote sensing and chlorophyll fluorescence technologies has provided solutions to these problems (Li *et al*., 2021; Xu *et al*., 2021). Hyperspectral remote sensing technology has the characteristics of high resolution and strong continuity between bands; it is able to obtain information that cannot be detected on the surface (Zdoan *et al*., 2021). Compared with the hyperspectral signal, the chlorophyll fluorescence parameters are more closely related to the physiological and ecological changes and photosynthesis of crops. Existing studies have shown that fluorescence signals can be detected before the chlorophyll content decreases, and hence can be used to predict the growth state of crops and explain the underlying physiological and biochemical changes, thereby achieving crop growth status monitoring (Zivcak *et al*., 2014; Feng *et al*., 2015), particularly regarding environmental stress. Therefore, it is important to integrate the hyperspectral remote sensing with chlorophyll fluorescence monitoring technology, thereby achieving rapid and accurate crop nitrogen content monitoring and facilitating efficient application of nitrogen fertilizer.

In this study, the spectral and fluorescence parameters of cotton leaves at different growth periods were determined, and these were then used in combination with cotton nitrogen content parameters to construct an estimation model on the basis of spectral-fluorescence data fusion. It is expected that the model can provide technical support for high-efficiency utilization of nitrogen fertilizer and thereby improving cotton yield.

Materials and Methods

Study area and study design

The study area was located in the second plot of Shihezi University's experimental field (85°59'E, 44°19'N), as shown in Figure 1. The soil in the study area is mostly loam; the average annual precipitation is 125.9–207.7 mm, with a large temperature difference between daytime and nighttime. The previous crop was also cotton. The soil was a loam that contained organic matter at 19.12 g/kg, total nitrogen 12.76 mg/kg, available phosphorus (P) 20.68 mg / kg, and available potassium (K) 165.11 mg/kg at the depth of 0-20 cm.

The cotton variety used in this study was 'Xinluzao 53'. A randomized block design was used. The width of the study area was 7.05 m; the length of the plot was 25 m, and the area of the plot was 176.25 m^2. The cotton planting pattern was 66 cm + 10 cm, with one film mulch, three pipes, and six rows. Six nitrogen application levels were used, 0, 120, 240, 360, 480, and 278 kg/ha, denoted as N0, N1, N2, N3, N4, and NC, respectively. At each level there were three replicates, resulting in a total of 18 test plots. Protective rows were set between each plot, as shown in Figure 1. The plots were irrigated according to local standards, and the remaining management was carried out in accordance with local high-yield planting requirements.

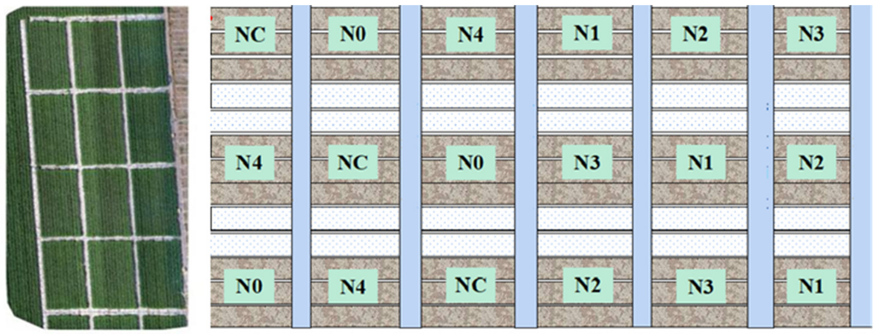

Figure 1. Study area plot layout map

Data collection and analysis

Spectral data of cotton leaves

A portable ground object spectroradiometer (SR-3500, Spectral Evolution, USA) was used to collect hyperspectral reflectance data of cotton leaves and the parameters are shown in Table 1. The spectral range of the spectroradiometer was 350–2500 nm, and the interval was 1 nm. Each leaf was sampled at three different points, with three repetitions at each point. The average of the three measurements was taken as the spectral value of a point, and the average value of three points on a leaf was taken as the spectral datum of the leaf. Three plants were selected from each plot for spectral data collection, and the spectroradiometer was calibrated each time a new plant was measured.

Table 1. Technical parameters of SR-3500 portable full spectrum ground object spectrometer

Technical indicators	Parameter	Technical indicators	Parameter
Spectral range	350-2500 nm	Field of view	25°; Integrating sphere
Spectral resolution	3.5 nm (350-1000 nm)	Spectral bandwidth	1.5 nm (350-1000 nm)
	10 nm (1000-1900 nm)		3.8 nm (1000-1900 nm)
	7 nm (1900-2500 nm)		2.5 nm (1900-2500 nm)

Chlorophyll fluorescence data of cotton leaves

The fluorescence parameters of living cotton leaves were measured using MultispeQ (PhotosynQ, USA). The measurements were carried out on clear and cloudless days at Beijing time (UTC+8) 12:00–16:00 (light adaptation) and 00:00–05:00 (dark adaptation). Three cotton plants were measured in each plot. The parameters of maximum fluorescence yield of leaves in the light (Fm'), photochemical quenching coefficient(qP), non-photochemical quenching coefficient (NPQ), maximum fluorescence yield of leaves in the dark (Fm), Yield, and optimal/maximal photochemical efficiency of PS IIin the dark. (Fv/Fm) were used for analysis.

Fm' is the maximum fluorescence under light, which is the fluorescence intensity when all PSII centers are open in the photo-adapted state; Fm is the fluorescence yield when the PS II reaction center is completely closed. It can reflect the electron transfer through PS II; Fv/Fm reflects the maximum light energy conversion efficiency in PS II system; qP reflects the proportion of open reaction centers in total PS II; NPQ reflects the change of heat dissipation. Yield is the PS II actual photochemical quantum efficiency, which reflects the primary light energy capture efficiency of the PS reaction center with partial closure.

Determination of nitrogen content in cotton leaves

Cotton leaf samples were collected sequentially in the budding period, flowering period, blooming period, full boll period, and boll opening period. Specifically, the leaf on the third main stem (L3) was collected, and the nitrogen content of the leaves was determined by the Kjeldahl method.

Screening of characteristic parameters

Characteristic spectral parameters

The hyperspectral data included 2150 bands. Due to homogeneity and redundancy in the spectral data, using full-band data can affect the accuracy of monitoring (Hansen *et al*., 2003; Feng *et al*., 2015). Therefore, it is necessary to reduce the data dimension and redundancy and only include sensitive bands. In this study, the shuffled frog-leaping algorithm (SFLA), the successive projection algorithm (SPA), grey relational analysis (GRA), and competitive adaptive reweighted sampling (CARS) were used to determine the characteristic bands with strong correlations with cotton leaf nitrogen content.

Shuffled frog-leaping algorithm (SFLA) is a new heuristic population evolution algorithm with high efficiency and excellent global search ability. The basic principle of the hybrid leapfrog algorithm is described. Aiming at the problem that the individual spatial position changes greatly before and after the update operation caused by the local update strategy of the algorithm, which reduces the convergence speed, an improved leapfrog algorithm based on the threshold selection strategy is proposed.

Successive projection algorithm (SPA) is a forward feature variable selection method. SPA uses vector projection analysis to project the wavelength onto other wavelengths, compare the size of the projection vector, take the wavelength with the largest projection vector as the wavelength to be selected, and then select the final characteristic wavelength based on the correction model. SPA selects the combination of variables with the least redundant information and the least collinearity.

Grey relational analysis (GRA) is to quantitatively characterize the degree of correlation between multiple factors, thus revealing the main characteristics of the grey system. Correlation analysis is the basis of grey system analysis and prediction.

Competitive adaptive reweighted sampling (CARS) is a feature variable selection method combining Monte Carlo sampling and the partial least squares (PLS) model regression coefficients. In the CARS, the points with larger absolute weight of the regression coefficient in the PLS model are retained as a new subset each time through adaptive weighted sampling (ARS), and the points with smaller weight are removed. Then the PLS model is established based on the new subset. After multiple calculations, the wavelength in the subset with the smallest root mean square error (RMSECV) of the PLS model is selected as the characteristic wavelength.

Characteristic fluorescence parameters

SPSS 24.0 software was used to analyze the correlation between chlorophyll fluorescence parameters and cotton leaf nitrogen content, and Duncan's test was used to determine the significance of differences between samples. Then, correlation and regression between the chlorophyll fluorescence parameters and the nitrogen content were performed, and the parameters with significant correlations were selected.

Model construction and validation

Data fusion methods

In this paper, two fusion algorithms of feature level and decision level are used to fuse the characteristic band and fluorescence characteristic parameters, and the monitoring model of cotton leaf nitrogen content is established. The two fusion algorithms are implemented in Python 3.9.

Feature-level fusion is carried out at the feature level, that is, extracting features from different data sources, fusing them before training the final classifier, and fusing the extracted spectral features and chlorophyll fluorescence feature parameters using the cascade method.

Decision-level fusion uses spectral characteristics, chlorophyll fluorescence parameters and cotton leaf nitrogen content to model separately in two dimensions to make decisions on the results. The algorithm first inputs the extracted characteristic parameter values into the full connection layer and the classification layer for the establishment and prediction evaluation of the initial estimation model, and carries out machine learning on the prediction results under the two states, and obtains the final results of the nitrogen content of cotton leaves.

Cotton leaf nitrogen content model

In order to solve the collinearity problem in hyperspectral data, four machine learning algorithms, i.e., random forest regression (RFR), support vector regression (SVR), feature-level fusion, and decision-level fusion, were used to construct four cotton leaf nitrogen monitoring models based on spectral-fluorescence data fusion.

Model validation

In this study, 54 data samples were obtained in each growth period, and these were divided into a training set and a validation set with a 2:1 ratio, i.e., 36 samples in the training set and 18 samples in the validation set (Ma *et al*., 2022). The coefficient of determination (R^2), the relative root mean square error (rRMSE), and the root mean square error (RMSE) were used to evaluate the accuracy of each model. A larger R^2 value indicates better fitting performance of the model, and smaller rRMSE and RMSE values indicate higher accuracy of the model. The calculation equations are as follows:

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

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i-x_i)^2} \quad (2)$$

$$rRMSE = \frac{RMSE}{\bar{x}} \times 100\% \quad (3)$$

where i is the i-th data point; xi represents the measured fluorescence or spectral value of the i-th data point; yi denotes the model predicted value of fluorescence or spectral data of the i-th data point. $\bar{x}$ is the average measured fluorescence or spectral value, and $\bar{y}$ is the predicted fluorescence or spectral value.

Results

Characteristic spectral and fluorescence parameters

Dimensionality reduction of the hyperspectral data of the L3 leaf at five different growth periods was carried out using SFLA, SPA, CARS, and GRA, and the characteristic spectral bands were obtained (Table 2). The reflectance and nitrogen content of each characteristic band were significant at the 0.01 level, i.e., the bands were significantly correlated with cotton leaf nitrogen content. The results showed that the characteristic bands obtained by different algorithms were adjacent in both the visible light and near-infrared ranges. Hence, all four algorithms achieved dimensionality reduction of the spectral data to a certain extent.

Table 2. Characteristic spectral parameters in different growth periods

Growth period	Method	Number of characteristic bands	Characteristic bands
Budding period	SFLA	10	713, 715, 568, 718, 567, 571, 705, 699, 700, 709
	SPA	15	547, 573, 548, 574, 546, 572, 544, 549, 719, 575, 718, 543, 570, 545, 708
	CARS	16	553, 565, 567, 570, 571, 572, 704, 706, 707, 708, 709, 710, 712, 714, 720, 721
	GRA	15	704, 713, 546, 701, 545, 542, 571, 702, 707, 706, 543, 711, 544, 565, 700
Flowering period	SFLA	10	2035, 2090, 1466, 2047, 1540, 2089, 2046, 2005, 1510, 2091
	SPA	15	2076, 2247, 2091, 2395, 2043, 2033, 2237, 2381, 2377, 2233, 2057, 2333, 2371, 1994, 1452
	CARS	19	1461, 1521, 1530, 1898, 1999, 2006, 2012, 2225, 2329, 2343, 2348, 2357, 2373, 2382, 2386, 2390, 2394, 2397, 2400
	GRA	15	1452, 1456, 1461, 1464, 1519, 1469, 1473, 1477, 1482, 1486, 2096, 2081, 2085, 1491, 1453
Blooming period	SFLA	10	1935, 1937, 2008, 957, 1936, 2339, 2366, 2362, 2363, 2231
	SPA	15	2368, 451, 406, 444, 412, 479, 1901, 1916, 450, 405, 1907, 410, 1392, 1881, 418
	CARS	9	966, 967, 968, 971, 973, 976, 978, 994, 1134, 1135, 1136, 1137, 1138, 1139, 1140, 1409, 1410, 1411
	GRA	15	402, 449, 405, 428, 403, 421, 408, 450, 435, 451, 440, 400, 439, 417, 432
Full boll period	SFLA	10	517, 632, 515, 591, 631, 692, 685, 691, 647, 506
	SPA	15	685, 692, 682, 633, 706, 507, 570, 718, 545, 540, 523, 501, 504, 727, 652
	CARS	13	596, 602, 604, 607, 642, 648, 650, 705, 706, 707, 708, 711, 727
	GRA	15	537, 521, 516, 545, 510, 532, 550, 524, 525, 502, 546, 527, 529, 501, 542
Boll opening period	SFLA	10	517, 632, 515, 591, 631, 692, 685, 691, 647, 506
	SPA	15	685, 692, 682, 633, 706, 507, 570, 718, 545, 540, 523, 501, 504, 727, 652
	CARS	9	596, 602, 604, 607, 642, 648, 650, 705, 706, 707, 708, 711, 727
	GRA	15	537, 521, 516, 545, 510, 532, 550, 524, 525, 502, 546, 527, 529, 501, 542

Duncan's test was used to compare the differences between samples, and 13 characteristic chlorophyll fluorescence parameters highly correlated with cotton leaf nitrogen were obtained. The results showed that there were differences in the characteristic chlorophyll fluorescence parameters in L3 leaves at different growth periods. A range of 8–11 characteristic fluorescence parameters were significant (Table 3). In particular, 11 parameters were significant in the blooming period. Five fluorescence parameters, Fv'/Fm', Fv/Fm, Yield, Fm, and NPQ, were significant in all five growth periods.

Table 3. Characteristic chlorophyll fluorescence parameters in different growth periods

Growth period	Characteristic parameter										
Budding period	**Fv'/Fm'**	**Fv/Fm**	**Yield**	qP	qN	**Fm**	**NPQ**	Fv/F0	Rfd		
	0.69**	0.66**	0.82**	0.79**	0.87**	0.69**	0.75**	0.86**	0.86**		
Flowering period	**Fv'/Fm'**	**Fv/Fm**	**Yield**	qN	**Fm**	**NPQ**	Fv/F0	Fm/F0	Rfd		
	0.58**	0.68**	0.65**	0.78**	0.60**	0.68**	0.76**	0.74**	0.75**		
Blooming period	**Fv'/Fm'**	**Fv/Fm**	**Yield**	qP	qN	Fm'	**Fm**	**NPQ**	Fv/F0	Fm/F0	Rfd
	0.84**	0.83**	0.82**	0.84**	0.69**	0.76**	0.76**	0.84**	0.59**	0.61**	0.87**
Full boll period	**Fv'/Fm'**	**Fv/Fm**	**Yield**	qP	F0'	Fm'	**Fm**	**NPQ**	Fs	Fm/F0	
	0.74**	0.53**	0.67**	0.64**	0.62**	0.76**	0.80**	0.65**	0.58**	0.50**	
Boll opening period	**Fv'/Fm'**	**Fv/Fm**	**Yield**	qP	**Fm**	**NPQ**	F0'	Fv/F0	Fm/F0	Rfd	
	0.63**	0.67**	0.59**	0.68**	0.69**	0.66**	0.65**	0.74**	0.42**	0.63**	

Cotton leaf nitrogen monitoring model based on characteristic hyperspectral parameters

<u>RFR-based cotton leaf nitrogen monitoring model</u>

Based on the characteristic bands listed in Table 2, an RFR cotton leaf nitrogen monitoring model was created, and the effectiveness of the model was verified (Figure 2). The RFR-GRA model had the highest accuracy in the budding period. The R^2 of the training set reached 0.812; the RMSE was 1.54; the R^2 of the validation set was 0.757, and the RMSE was 1.29. The linear fitting showed a close to 1:1 ratio between the measured values and the predicted values. The RFR-CARS model in the full boll period had the lowest accuracy. The R^2 of the training set was 0.619; the RMSE was 1.60, and the R^2 of the validation set was only 0.020, with an RMSE of 1.84. There was an over-fitting phenomenon that led to a poor generalization ability of the model, and hence affected the model performance.

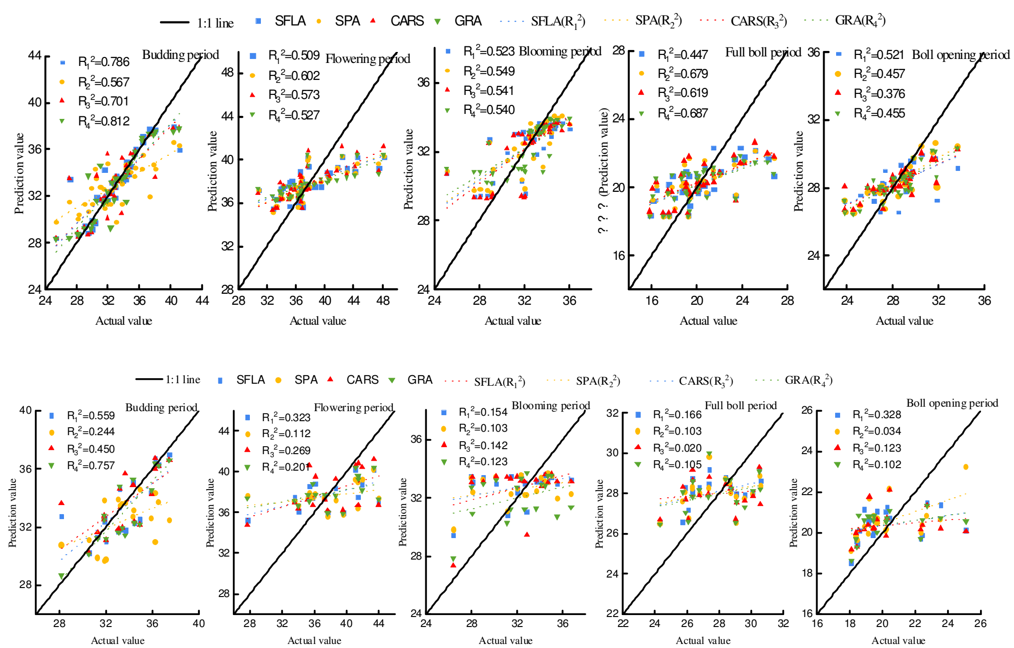

Figure 2. RFR-based cotton nitrogen monitoring models

<u>SVR-based cotton leaf nitrogen monitoring model</u>

The accuracy of the SVR-based cotton leaf nitrogen monitoring models varied among different growth periods, and the effectiveness of the model in the budding period was significantly better than that in other growth periods (Figure 3). Specifically, the SVR-GRA algorithm had the strongest modeling and inversion ability in the budding period; the R^2 values of the training set and the validation set reached 0.852 and 0.810, with RMSE of 1.80 and 1.20, respectively. In contrast, the performance of the models was poor in the flowering period, and there was over-fitting in all four models, indicating that the results could not be used for cotton leaf nitrogen monitoring. In summary, among all the SVR-based models, the GRA model showed the best effect for the characteristic bands.

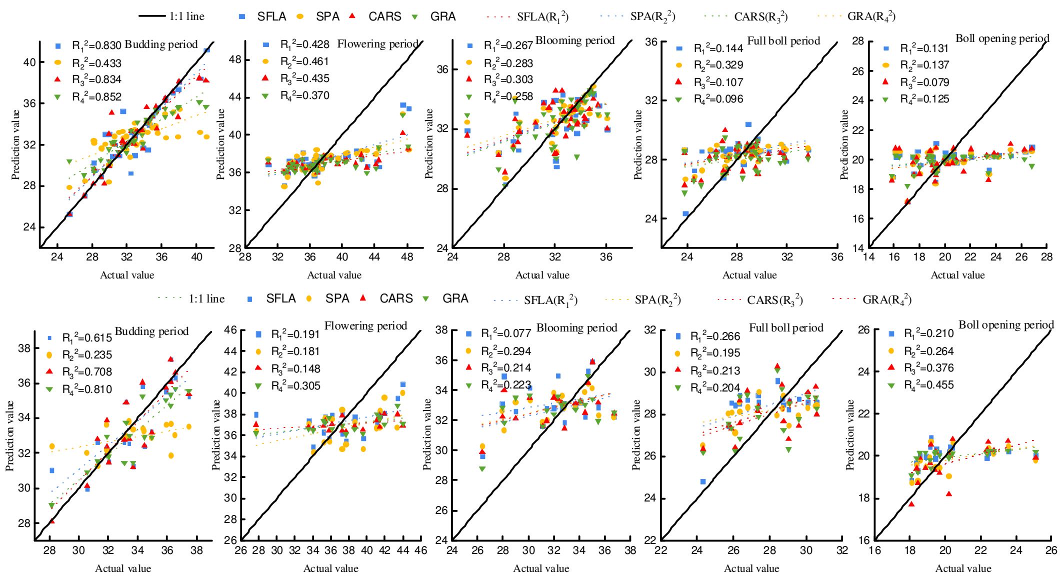

Figure 3. SVR-based cotton nitrogen monitoring models

Cotton leaf nitrogen monitoring model based on characteristic chlorophyll fluorescence parameters

<u>RFR-based cotton leaf nitrogen monitoring model</u>

Taking the characteristic chlorophyll fluorescence parameters obtained by the correlation analysis as the input variables and the cotton leaf nitrogen as the output variable, an L3 cotton leaf nitrogen model was constructed based on the RFR algorithm. There were 54 data points divided into a training set and a validation set with a ratio of 2:1. The results showed that the RFR models established based on characteristic fluorescence parameters had a good ability to estimate cotton leaf nitrogen content (Figure 4). Specifically, the monitoring model in the flowering period demonstrated the best performance, with R^2 above 0.83 in both training and validation sets, and RMSE less than 1.5. The model performance in the flowering period was slightly inferior to that in the other growth periods. The error was small in the training set but larger in the validation set, indicating an over-fitting problem that affected the generalization ability of the model.

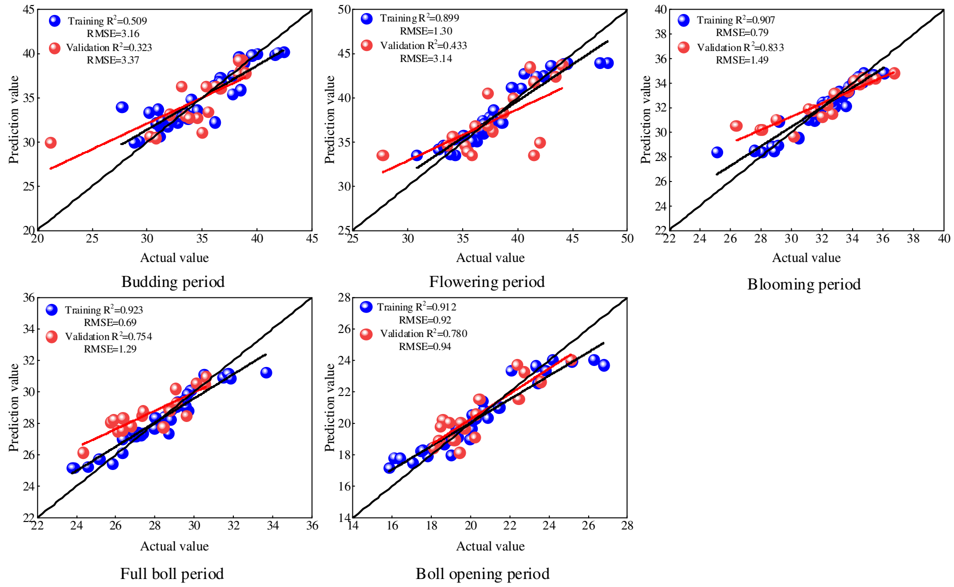

Figure 4. RFR-based cotton nitrogen monitoring models (blue: training set, red: validation set)

<u>SVR-based cotton leaf nitrogen monitoring model</u>

The results of the monitoring models created based on the SVR algorithm showed differences in the estimation of nitrogen content in different growth periods. The model in the blooming period had the highest accuracy. The RMSE values of the training and validation sets were 1.03 and 1.85, respectively, and the R^2 was above 0.7. Moreover, the model in the flowering period had low accuracy. Compared to the R^2 of the model in the blooming period, the R^2 decreased by 13.95% and 119.50% in the training and validation sets, while the RMSE increased by 67.41% and 45.10%, respectively. Therefore, on the basis of the above results, the blooming period and the budding period are the best times to estimate cotton leaf nitrogen based on the chlorophyll fluorescence parameters.

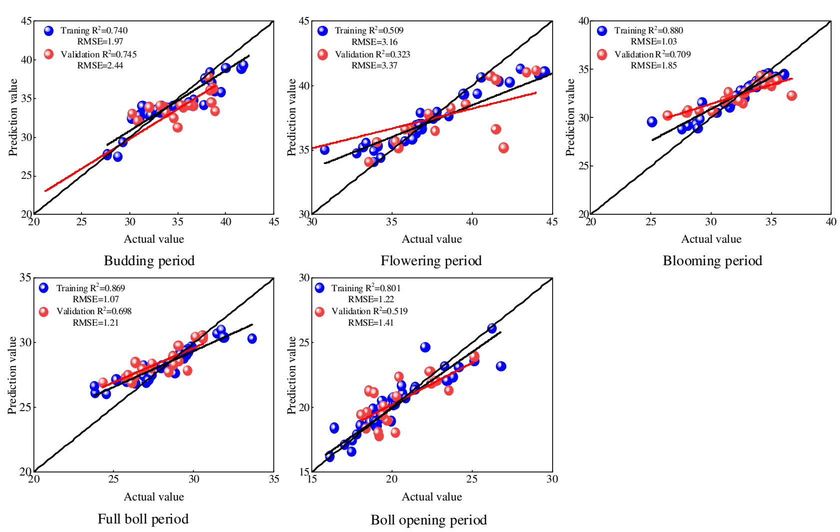

Figure 5. SVR-based cotton nitrogen monitoring models (blue: training set, red: validation set)

Cotton leaf nitrogen monitoring model based on spectral-fluorescence data fusion

Cotton leaf nitrogen monitoring model based on feature-level fusion

L3 cotton leaf nitrogen models at different growth periods were constructed based on the RFR and SVR algorithms. The inversion results are shown in Figure 6. The RFR-CARS model in the blooming period had the best performance. In the training set, R^2 = 0.800, rRMSE (%) = 2.48%, and RMSE = 0.80, while in the validation set, R^2 = 0.871, rRMSE (%) = 4.45%, and RMSE = 1.43. The accuracy was higher than for the other models in the same growth period. The RFR-SPA model in the budding period had the lowest accuracy. In the training set R^2 = 0.340, rRMSE (%) =7.45%, and RMSE = 2.46; in the validation set, R^2 = 0.341, rRMSE (%) = 6.52%, and RMSE = 2.19. Compared with the models created based solely on spectral and fluorescence parameters, the accuracy of the fusion models was greatly improved. Compared with the best model based on the characteristic spectral parameters, the R^2 in the training set and validation set was increased by 2.93% and 5.79%, respectively, and the RMSE was reduced by 75.31% and 50.60%. Compared with the best model based on the characteristic fluorescence parameters, the R^2 of the training set and validation set increased by −1.72% and 5.91%, respectively, and the RMSE decreased by −14.46% and 1.21%.

Comparing the models in different growth periods, the models based on the characteristic bands obtained by SFLA and CARS had the highest accuracy, followed by GRA and SPA. Except for the boll opening period, the best models in all growth periods (i.e., budding period, flowering period, blooming period, and full boll period) were RFR-based models, suggesting that the RFR algorithm had a higher generalization ability and inversion accuracy, and better stability compared to other methods. Hence, it is suitable for estimation of cotton leaf nitrogen content.

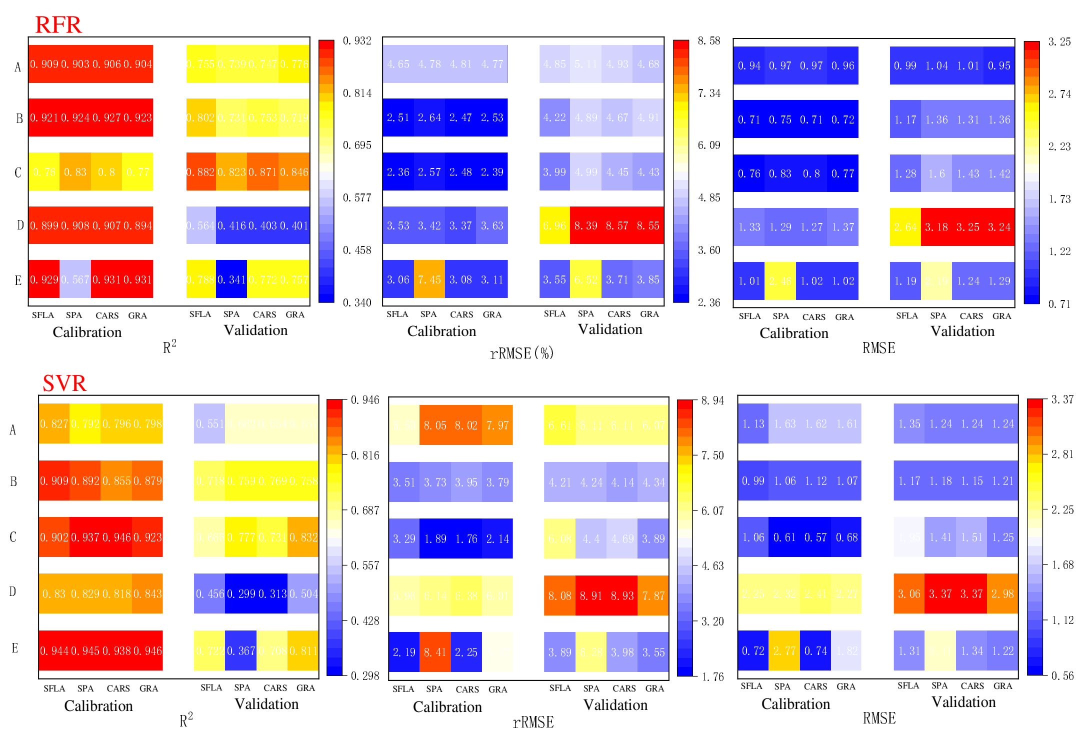

Figure 6. Cotton leaf nitrogen monitoring models based on feature-level fusion (A:Boll opening period; B:Full boll period; C:Blooming period; D:Flowering period; E:Budding period)

Cotton leaf nitrogen monitoring model based on decision-level fusion

Decision modeling was carried out according to the results based on spectral and fluorescence parameters. The fitting results are shown in Figure 7. For the RFR model in the blooming period, the R^2 of the training set reached 0.961, significantly higher than that of the models created in previous sections, and the RMSE was greatly reduced. The R^2 was further increased by 5.7% from the model of feature-level fusion. In terms of the estimation ability of the models for nitrogen content, RFR-CARS had the best estimation ability in the blooming period (training set: R^2 = 0.945, RMSE = 0.58, validation set: R^2 = 0.921, RMSE =1.60). RFR-SPA had a poor estimation ability ((training set: R^2 = 0.749, RMSE = 1.76). In conclusion, based on the progressive increase of the R^2 value from 0.269 (spectral data), 0.833 (fluorescence parameter), 0.907 (feature set fusion) to 0.921 (decision-level fusion), the decision modeling exhibited outstanding performance in inversion of cotton leaf nitrogen content.

Similarly, the accuracy of the monitoring models based on the SVR algorithm showed that except for the SVR-SPA in the budding period, all models had an R^2 value of over 0.870 and an RMSE below 1.30. The model in the blooming period had the best accuracy (R^2 = 0.945, RMSE = 0.57). The values of R^2 were 253.93%, 7.39%, and 4.78% higher than those of the models based on spectral parameters, fluorescence parameters, and feature-level fusion, and the RMSE values were 75.83%, 50.49%, and 51.89% lower. The validation results showed that SVR-GRA was the best in the budding period (R^2 = 0.808 and RMSE = 1.30). Compared with the model based on spectral parameters and feature-level fusion, the model accuracy did not improve significantly. When compared with the model based on the fluorescence parameters, the R^2 was improved by 12.85%. However, for the SVR-GRA model in the flowering period that had the lowest decision modeling accuracy, the R^2 was 21.97% and 15.17% higher than that based on spectral and fluorescence parameters, and the RMSE was reduced by 11.29% and 20.77%, respectively. However, the R^2 of the feature-level fusion model was 35.48% higher than that of the decision fusion model, indicating that decision-level fusion was unsuitable for models in all growth periods.

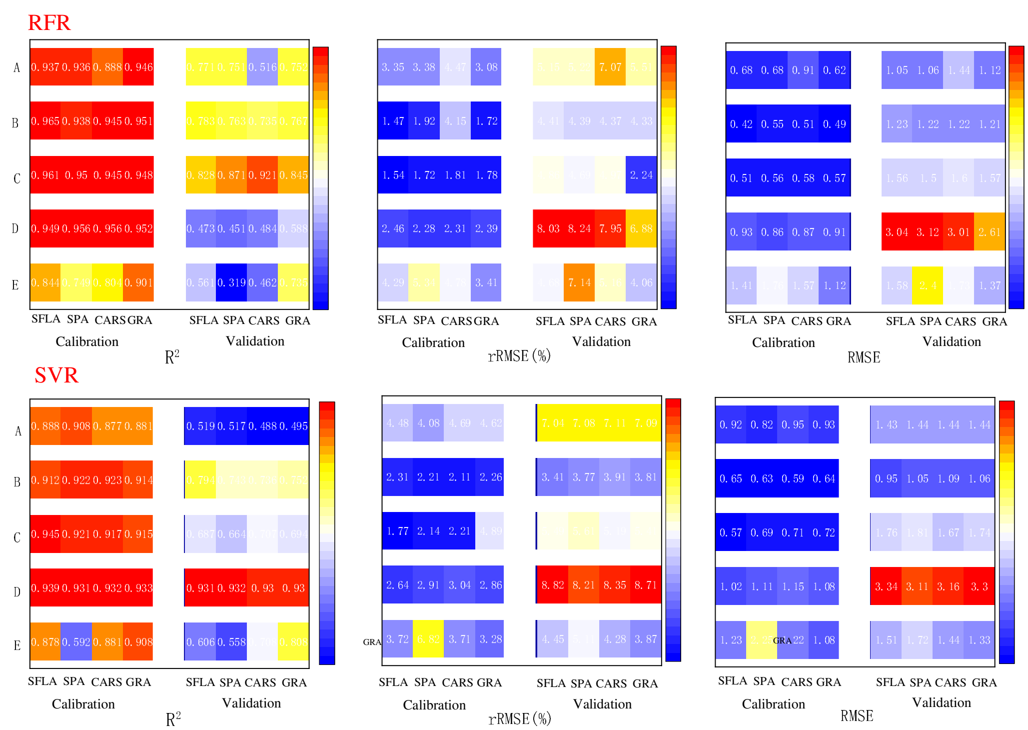

Figure 7. Parameters of the cotton leaf nitrogen monitoring models based on feature-level fusion (A:Boll opening period; B:Full boll period; C:Blooming period; D:Flowering period; E:Budding period)

Discussion

Due to redundancy, how to extract characteristic spectral bands has always been a key research topic in crop remote sensing monitoring (Blackburn, 1998; Hansen *et al*., 2003; Zhu *et al*., 2007). In conventional multispectral data, researchers usually perform regression on individual bands one by one in order to obtain characteristic bands that are sensitive to crop physiological and biochemical characteristics (Niu *et al*., 2000; Ma *et al*., 2020). However, for hyperspectral information, it is necessary to use algorithms to effectively and accurately extract characteristic information that is highly correlated with the target. In this study, four methods, i.e., SFLA, SPA, CARS, and GRA, were used to screen the characteristic bands, thereby achieving rapid extraction and analysis of spectral information. The selected characteristic spectral bands can effectively improve the accuracy of cotton leaf nitrogen monitoring models. However, when different machine learning algorithms are used to screen characteristic bands, some of the effective spectral information may be excluded. Moreover, bands with low correlation with nitrogen content may be selected. Although we analyzed the correlations between the bands with cotton leaf nitrogen content before screening (significance at the 0.01 level), their physical significance remains to be explored. Two machine learning algorithms were used in this study to construct L3 cotton leaf nitrogen monitoring models. Among these, SVR-GRA in the budding period had the best performance. However, the accuracy for the flowering period and full boll period was lower than the canopy spectral models employed in previous studies (Wen *et al*., 2016; Ma *et al*., 2020), a result that may be due to differences in the data collection environment. The human and meteorological factors were more complicated in this study, and these may have affected the model accuracy.

In this study, the fluorescence parameters Fv/Fm, Fv'/Fm', Fv/F0 and Yield of L3 cotton leaves showed decreasing trends with the growth period, consistent with the results of existing studies (Zhang *et al*., 2003; Li *et al*., 2020; Ding *et al*., 2020). Under different levels of nitrogen application, the fluorescence parameters Fv/Fm, Fv'/Fm', and Fv/F0 were greater in the high nitrogen application groups (N3, N4) and lesser in the low nitrogen application groups. The results suggest that with insufficient supply of nitrogen, cotton leaves are prone to photoinhibition, and increasing nitrogen fertilizer application can effectively enhance the photosynthesis of cotton leaves and reduce photoinhibition. Similar results have been reported for corn and wheat (Khamis *et al*., 1990; Zhang *et al*., 2003). In this study, by analyzing the relationship between the fluorescence parameters and cotton nitrogen content at different leaf positions, we found that there were differences in the model accuracy for L3 cotton leaves at different growth periods. Research has shown that stronger photosynthesis is associated with older leaf age at the top and higher reliability of leaf nitrogen content inversion by different fluorescence parameters (Dong *et al*., 2000). Therefore, the L3 leaf can be used as a functional leaf for cotton nitrogen monitoring. In addition, the models in this study were different from the linear models constructed by Chen *et al*. (2007) and Ma et al. 2006), suggesting that the correlation between fluorescence parameters and nitrogen content is different for different crops, and that crop nitrogen monitoring based on different fluorescence parameters is feasible.

To monitor the severity of wheat leaf rust, Bai *et al*. (2020) took fused spectral index and canopy SIF as independent variables to build a monitoring model. The results showed that regardless of the modeling method, the model accuracy based on data fusion was better than when the data were used alone. Jing *et al*. (2019) used random forest to build a wheat rust monitoring model based on SIF-spectral data fusion and found that compared with single data type modeling, the R^2 of collaborative modeling was increased by 4%, and the RMSE decreased by 22%, indicating that data fusion effectively enhanced the monitoring accuracy of the model. In this study, the fluorescence-spectral data were fused to build a cotton leaf nitrogen monitoring model. The model accuracy was greatly improved, and the accuracy of decision-level fusion was higher than that of feature-level fusion, consistent with the findings of previous studies.

Although the models based on feature-level fusion and decision-level fusion had better accuracy for the L3 leaf in all growth periods, the accuracy in some growth periods was still inferior to that of the models based on spectrum and fluorescence data alone. For example, SVR-GRA had the best performance in the budding period. However, most fusion models achieved better results than the models based on a single data source, and the stability and accuracy for all leaf positions in the whole growth period were much higher than with other modeling methods. The application of the model in the estimation of vegetation physicochemical parameters needs to be further analyzed.

Conclusions

In this study, hyperspectral reflectance and chlorophyll fluorescence data were integrated, and based on different characteristic bands screening methods, RFR, SVR, feature-level fusion, and decision-level fusion were used to estimate the nitrogen content of L3 leaves at different growth periods. The main findings are as follows:

(1) In the cotton leaf nitrogen monitoring models based on hyperspectral characteristic bands, the accuracy and stability of SVR-GRA was the best (R^2 = 0.852, RMSE = 1.20).

(2) The fluorescence parameters Fv/Fm, Fv'/Fm', Fv/F0, and Yield under different nitrogen applications gradually decreased with the progression of the growth period. Based on the characteristic fluorescence parameters, cotton leaf nitrogen monitoring models were constructed, and the accuracy in the blooming period was the highest, with R^2 in both the training set and the validation set above 0.85, suggesting good estimation performance.

(3) Among the cotton leaf nitrogen monitoring models based on spectral-fluorescence data fusion, the RFR-CAR model based on feature-level fusion had the best performance in the blooming period; in the training set, R^2 = 0.800, rRMSE (%) = 2.48%, and RMSE = 0.80; in the validation set, R^2 = 0.871, rRMSE (%) = 4.45%, and RMSE = 1.43. The results of decision modeling showed that the RFR-CARS model had the best inversion ability for the L3 leaf nitrogen content in the blooming period, with R^2 in the validation and training set both above 0.920, indicating that feature-level fusion and decision-level fusion significantly improved the accuracy and stability of the models.

Authors' Contributions

XL were responsible for determination of some indexes and the writing of this manuscript. LLM and XYC performed the experiments and collected all data sets. AQC, FH and YRM were responsible for the mapping and typesetting. ZAH were responsible for the revision. ZZ, the corresponding author, were responsible for the revision and quality control of the paper. All authors read and approved the final manuscript.

Ethical approval (for researches involving animals or humans)

Not applicable.

Acknowledgements

This work was supported by the National Natural Science Foundation of China, China (Grant No. 42061058), Technological Innovation Talent Plan of XPCC, China (Grant No. 2022CB002-01) and Corps financial science and technology plan of XPCC, China (Grant No. 2020AB005).

Conflict of Interests

The authors declare that there are no conflicts of interest related to this article.

References

Bai ZP (2020). Remote sensing monitoring of wheat stripe rust based on reflectance spectroscopy and SIF. Xi'an: Xi'an University of Science and Technology, (in Chinese)

Barraclough PB, Howarth JR, Jones J, Lopez-Bellido R, Parmar S, Shepherd CE, Hawkesford MJ (2010). Nitrogen efficiency of wheat: Genotypic and environmental variation and prospects for improvement. European Journal of Agronomy 33(1):1-11. *https://doi.org/10.1016/j.eja.2010.01.005*

Blackburn GA (1998). Spectral indices for estimating photosynthetic pigment concentrations: A test using senescent tree leaves. International Journal of Remote Sensing 19(4):657-675. *https://doi.org/10.1080/014311698215919*

Chen B, Zheng H, Luo G, Chen C, Bao A, Liu T, Chen X (2020). Estimating nitrogen content in cotton canopy using hyperspectral imaging. Journal of Irrigation and Drainage 39(03):35-41. (in Chinese)

Chen YS (2007). Quantitative study on the effects of nitrogen on post-flowering leaf photosynthesis and chlorophyll fluorescence parameters of cucumber in greenhouse. Nanjing: Nanjing Agricultural University. (in Chinese)

Ding YR, Li DM, Ma LL (2020). Study on inversion model of cotton chlorophyll fluorescence parameters and cotton growth index under drip irrigation. Agricultural Research in the Arid Areas 38(06):234-242. *https://doi.org/10.7606/j.issn.1000-7601.2020.06.31*

Dong HZ, Li WJ, Tang W, Li ZH (2000). Photosynthetic characters of field grown cotton leaves. Shandong Agricultural Sciences 6(7-9):15. (in Chinese) *https://doi.org/10.3969/j.issn.1001-4942.2000.06.002*

Feng W, He L, Zhang HY, Guo BB, Zhu YJ, Wang CY, Guo TC (2015). Assessment of plant nitrogen status using chlorophyll fluorescence parameters of the upper leaves in winter wheat. European Journal of Agronomy 64:78-87. *https://doi.org/10.1016/j.eja.2014.12.013*

Halihashi Y, Li QJ, Zhang Y (2019). Effects of organic manure nitrogen replacing chemical fertilizer nitrogen on cotton nutrient uptake, nitrogen utilization efficiency and yield. Soil and Fertilizer Sciences in China 3:137-142. (in Chinese). *https://doi.org/10.11838/sfsc.1673-6257.18267*

Hansen PM, Schjoerring JK(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. *https://doi.org/10.1016/S0034-4257(03)00131-7*

He T (2016). Hyperspectral remote sensing model for maize nitrogen monitoring. Shenyang: Shenyang Agricultural University. (in Chinese)

Hongyun YANG, Jianjun LUO, Aizhen S, Ying WAN, Wenlong YI (2020). Study on estimation model of total nitrogen content in rice leaves based on image characteristics. Acta Agriculturae Zhejiangensis 32:2232-2243. (in Chinese) *https://doi.org/10.3969/j.issn.1004-1524.2020.12.15*

Jay S, Maupas F, Bendoula R, Gorretta N (2017). Retrieving LAI, chlorophyll and nitrogen contents in sugar beet crops from multi-angular optical remote sensing: Comparison of vegetation indices and PROSAIL inversion for field phenotyping. Field Crops Research 210:33-46. *https://doi.org/10.1016/j.fcr.2017.05.005*

Jing X, Bai Z, Gao Y, Liu LY (2019). Wheat stripe rust monitoring by random forest algorithm combined with SIF and reflectance spectrum. Transactions of the Chinese Society of Agricultural Engineering 35(13):154-161. (in Chinese) *https://doi.org/10.11975/j.issn.1002-6819.2019.13.017*

Khamis S, Lamaze T, Lemoine Y, Foyer C (1990). Adaptation of the photosynthetic apparatus in maize leaves as a result of nitrogen limitation. Plant Physiology 94(3):1436-1443.1436-1443.

Li DM, Lv X, Luo HH (2020). A monitoring model of nitrogen nutrition of drip irrigation cotton based on chlorophyll fluorescence parameters. Cotton Science 32(01):63-76. (in Chinese) *https://doi.org/10.11963/1002-7807.ldmzz.20191231*

Li F, Li D, Elsayed S, Hu Y, Schmidhalter U (2021). Using optimized three-band spectral indices to assess canopy N uptake in corn and wheat. Science Direct. European Journal of Agronomy 127:126-286. *https://doi.org/10.1016/j.eja.2021.126286*

Li JH, Wang F, Li JW, Zou RB, Liao GP (2016). Multifractal methods for rapeseed nitrogen nutrition qualitative diagnosis modeling. Journal of Biological Mathematics 9:285-297. *https://doi.org/10.1142/S1793524516500649*

Li MF, Peng WY, Li HT (2016). Effects of the combined application of nitrogen and boron on the yield, petiole annulus and nutrient uptake of cotton. Soil and Fertilizer Sciences in China 4:97-102. (in Chinese) *https://doi.org/10.11838/sfsc.20160416*

Liu L, Chen YH, Lei ZX (2021). A study on high-quality development of the cotton industry of Xinjiang [J]. Macroeconomic Management 10:77-83. (in Chinese)

Lyu T, Shen J, Ma J, Ma P, Yang Z, Dai Z, ... Li M (2021). Effects of slow-release nitrogen fertilizer combined with urease inhibitor on photosynthetic characteristics of cotton. China Cotton 48(10):1-7. (in Chinese) *https://doi.org/10.11963/cc20210128*

Ma JF (2006). Nitrogen content monitoring in wheat and rice based on leaf chlorophyll fluorescence parameters. Nanjing: Nanjing Agricultural University. (in Chinese).

Ma L, Chen X, Zhang Q, Lin J, Yin C, Ma Y, ... Lv X (2022). Estimation of nitrogen content based on the hyperspectral vegetation indexes of interannual and multi-temporal in cotton. Agronomy 12(6):1319. *https://doi.org/10.3390/agronomy12061319*

Niu Z, Chen YH, Sui HZ, Zhang QY, Zhao CJ (2000). Mechanism analysis of leaf biochemical concentration by high spectral remote sensing. Journal of Remote Sensing 4(2):125-129. (in Chinese) *https://doi.org/10.3321/j.issn:1007-4619.2000.02.008*

Serrano L, Filella I, Penuelas J (2000). Remote sensing of biomass and yield of winter wheat under different nitrogen supplies. Crop Science 40:723-731. *https://doi.org/10.2135/cropsci2000.403723x*

Shi XH, Cai HJ (2018). Estimation of nitrogen nutrition index of greenhouse tomato under different water and nitrogen fertilizer treatments based on leaf SPAD. Transactions of the Chinese Society of Agricultural Engineering 34:116-126. (in Chinese) *https://doi.org/10.11975/j.issn.1002-6819.2018.17.016*

Suo JY (2016). Evaluation of Nitrogen Application Efficiency and Nutritional Status of Cotton in Xinjiang. Urumqi: Xinjiang Agricultural University. (in Chinese)

Wang H, Guo Z, Shi Y, Zhang Y, Yu Z (2015). Impact of tillage practices on nitrogen accumulation and translocation in wheat and soil nitrate-nitrogen leaching in drylands. Soil and Tillage Research 153:20-27. *https://doi.org/10.1016/j.still.2015.03.006*

Wen PF (2016). A spectral estimation model for cotton nitrogen and a diagnosis system. Xinjiang: Shihezi University. (in Chinese)

Xu HC, Yao B, Wang Q, Chen TT, Zhu TZ, He H, ... Wu LQ (2021). Determination of suitable band width for estimating rice nitrogen nutrition index based on leaf reflectance spectra. Scientia Agricultura Sinica 54:4525-4539. (in Chinese) *https://doi.org/10.3864/j.issn.0578-1752.2021.21.004*

Xu W, Chen P, Zhan Y, Chen S, Zhang L, Lan Y (2021). Cotton yield estimation model based on machine learning using time series UAV remote sensing data. International Journal of Applied Earth Observation and Geoinformation 104:102511. *https://doi.org/10.1016/j.jag.2021.102511*

Zdoan G, Lin X, Sun DW (2021). Rapid and noninvasive sensory analyses of food products by hyperspectral imaging: Recent application developments. Trends in Food Science & Technology 111:151-165. *https://doi.org/10.1016/j.tifs.2021.02.044*

Zhang L, Shangguan Z, Mao M, Yu G (2003). Effects of long-term application of nitrogen fertilizer on leaf chlorophyll fluorescence of upland winter wheat. Chinese Journal of Applied Ecology (05):695-698. (in Chinese)

Zhang WF, Gou L, Wang ZL (2003). Effect of nitrogen on chlorophyll fluorescence of leaves of high-yielding cotton in Xinjiang. Scientia Agricultura Sinica (08): 893-898. (in Chinese)

Zhong KG, Li N, Hua K (2021). Analysis and forecast of influencing factors of cotton yield in Xinjiang based on grey theory. Agriculture and Technology 41(21):1-3. (in Chinese) *https://doi.org/10.19754/j.nyyjs.20211115001*

Zhu Y, Tian Y, Yao X, Liu X, Cao W (2007). Analysis of common canopy reflectance spectra for indicating leaf nitrogen concentrations in wheat and rice. Plant Production Science 10(4):400-411. *https://doi.org/10.1626/pps.10.400*

Živčák M, Olšovská K, Slamka P, Galambošová J, Rataj V, Shao HB, Brestič M (2014). Application of chlorophyll fluorescence performance indices to assess the wheat photosynthetic functions influenced by nitrogen deficiency. Plant Soil and Environment 60(5):210-215. *https://doi.org/10.2135/cropsci2013.08.0551*