



## OPEN ACCESS

## EDITED BY

Guoquan Lu,  
Zhejiang Agriculture and Forestry  
University, China

## REVIEWED BY

Haikuan Feng,  
Beijing Research Center for Information  
Technology in Agriculture, China  
Ben Zhao,  
Chinese Academy of Agricultural Sciences,  
China

## \*CORRESPONDENCE

Ze Zhang

✉ zhangze1227@shzu.edu.cn

Lifu Zhang

✉ zhanglf@radi.ac.cn

<sup>†</sup>These authors have contributed  
equally to this work and share  
first authorship

## SPECIALTY SECTION

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

RECEIVED 06 December 2022

ACCEPTED 14 February 2023

PUBLISHED 01 March 2023

## CITATION

Qin S, Ding Y, Zhou Z, Zhou M, Wang H,  
Xu F, Yao Q, Lv X, Zhang Z and Zhang L  
(2023) Study on the nitrogen content  
estimation model of cotton leaves  
based on “image-spectrum-  
fluorescence” data fusion.  
*Front. Plant Sci.* 14:1117277.  
doi: 10.3389/fpls.2023.1117277

## COPYRIGHT

© 2023 Qin, Ding, Zhou, Zhou, Wang, Xu,  
Yao, Lv, Zhang and Zhang. This is an open-  
access article distributed under the terms of  
the [Creative Commons Attribution License  
\(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or  
reproduction in other forums is permitted,  
provided the original author(s) and the  
copyright owner(s) are credited and that  
the original publication in this journal is  
cited, in accordance with accepted  
academic practice. No use, distribution or  
reproduction is permitted which does not  
comply with these terms.

# Study on the nitrogen content estimation model of cotton leaves based on “image-spectrum-fluorescence” data fusion

Shizhe Qin<sup>1†</sup>, Yiren Ding<sup>1†</sup>, Zexuan Zhou<sup>1</sup>, Meng Zhou<sup>1</sup>,  
Hongyu Wang<sup>1</sup>, Feng Xu<sup>1</sup>, Qiushuang Yao<sup>1</sup>, Xin Lv<sup>1</sup>, Ze Zhang<sup>1\*</sup>  
and Lifu Zhang<sup>2\*</sup>

<sup>1</sup>Xinjiang Production and Construction Crops Oasis Eco-Agriculture Key Laboratory, Shihezi University College of Agriculture, Shihezi, China, <sup>2</sup>Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing, China

**Objective:** Precise monitoring of cotton leaves' nitrogen content is important for increasing yield and reducing fertilizer application. Spectra and images are used to monitor crop nitrogen information. However, the information expressed using nitrogen monitoring based on a single data source is limited and cannot consider the expression of various phenotypic and physiological parameters simultaneously, which can affect the accuracy of inversion. Introducing a multi-source data-fusion mechanism can improve the accuracy and stability of cotton nitrogen content monitoring from the perspective of information complementarity.

**Methods:** Five nitrogen treatments were applied to the test crop, Xinluzao No. 53 cotton, grown indoors. Cotton leaf hyperspectral, chlorophyll fluorescence, and digital image data were collected and screened. A multilevel data-fusion model combining multiple machine learning and stacking integration learning was built from three dimensions: feature-level fusion, decision-level fusion, and hybrid fusion.

**Results:** The determination coefficients ( $R^2$ ) of the feature-level fusion, decision-level fusion, and hybrid-fusion models were 0.752, 0.771, and 0.848, and the root-mean-square errors (RMSE) were 3.806, 3.558, and 2.898, respectively. Compared with the nitrogen estimation models of the three single data sources,  $R^2$  increased by 5.0%, 6.8%, and 14.6%, and the RMSE decreased by 3.2%, 9.5%, and 26.3%, respectively.

**Conclusion:** The multilevel fusion model can improve accuracy to varying degrees, and the accuracy and stability were highest with the hybrid-fusion model; these results provide theoretical and technical support for optimizing an accurate method of monitoring cotton leaf nitrogen content.

## KEYWORDS

cotton, nitrogen, hyperspectral, chlorophyll fluorescence, digital images, data fusion

# 1 Introduction

Cotton is the dominant industry in the Xinjiang region, China's largest production base of high-quality commodity cotton. In 2022, the cotton planting area covered approximately 83% of the farmland in Xinjiang, with over 90% production. Nitrogen is an indispensable life element of cotton crops. It is an important component of protein, nucleic acid, chlorophyll, and some hormones, provides important support for photosynthesis, and significantly impacts crop growth and development (Han et al., 2022; Wang et al., 2022). Nitrogen fertilization is the most used fertilizer type during cotton cultivation and management, with the highest application rate (70–80%) used in drip irrigation. Proper nitrogen fertilizer application can promote cotton growth and increase yield, but unreasonable application reduces fertilizer efficiency and causes fertilizer waste and serious environmental pollution problems (Zhang et al., 2018). Therefore, it is important to understand the nitrogen nutrition of cotton in real-time during the growth process and to develop an efficient and scientific nitrogen monitoring method to improve nitrogen fertilizer efficiency.

Traditional nitrogen monitoring methods are expensive, destructive, and time consuming. With the development of modern information technology, efficient and non-destructive monitoring methods, such as hyperspectral and digital imaging, have become increasingly popular worldwide. For example, researchers screened sensitive bands based on hyperspectral data or combined them with vegetation indices to build and apply nitrogen nutrition inversion models for different crops (Yao et al., 2010; Jin et al., 2013; Zhao et al., 2018; Luo et al., 2020; Zhang et al., 2021); this indicates that crop nitrogen content estimation using hyperspectral techniques is practical and feasible. Chlorophyll fluorescence monitoring differs from hyperspectral monitoring because it provides information directly related to plant physiological and functional properties (Ptushenko et al., 2014). Fluorescence signals can predict the growth status of plants and explain the changes in plant physiological and biochemical processes (Zivcak et al., 2015; Jia, 2016). Nitrogen application rate can notably affect the characteristic changes of chlorophyll fluorescence parameters when used to monitor crop nitrogen content (He et al., 2020), which indicates that fluorescence parameters are ideal indicators of crops' nutritional status and can directly monitor nitrogen content (Feng et al., 2015; Li et al., 2020a). Compared with hyperspectral and chlorophyll fluorescence, digital image technology is simple to operate, visualize, and apply. It includes an interval of interest segmenting, converting image color space (Huang et al., 2020), extracting relevant color and texture features, and building predictive models (Xiao, 2022). In recent years, scholars have achieved good results using this technology to model and invert the nitrogen nutrients of crops such as rice (Saberioon et al., 2013; Wang et al., 2014; Li et al., 2015), corn (Li et al., 2010; Liu and Li, 2010), sugar beet (Moghaddam et al., 2010), and cotton (Zhang et al., 2020; Hong et al., 2022).

The spectra of reflection or absorption can reflect differences in crop composition and structure (Tong et al., 2006). Nitrogen limitation and nitrogen redundancy will cause, among other

symptoms, yellowing or greenness of crop leaves and changes in leaf thickness and water content, resulting in changes in spectral reflectance characteristics. However, it is impossible to obtain the true color based on different color spaces, such as RGB, and shape-related values, such as area features. Nevertheless, the surface color, shape, and other feature changes of crops can be determined using digital images to reflect the nitrogen nutrition status. The changes in chlorophyll fluorescence parameters can directly or indirectly reflect the primary reaction, electron transfer, and CO<sub>2</sub> assimilation process of photosynthesis and affect the energy distribution of crops through nitrogen. This causes dynamic changes in fluorescence, photosynthesis, and heat dissipation that can be used to monitor nitrogen content (Kumar et al., 2014). The three data sources can be used separately for crop nitrogen nutrient inversion. However, the nitrogen information monitoring of single data sources cannot consider the expression of the above phenotype and physiological parameters, and, currently, the monitoring accuracy cannot meet the requirements of precision agriculture. One of the keyways to solve this problem is to effectively integrate the information obtained from "spectrum, fluorescence, and image."

In this study, nitrogen content estimation models that combine multiple machine learning and stacking integrated learning were built from the three levels of feature-level fusion, decision-level fusion, and hybrid fusion using cotton as the research object. The fusion models were compared with the three single data source models. The aim of this study was to improve the accuracy and stability of nitrogen estimation models and provide theoretical and methodological support for accurately monitoring cotton nitrogen content and efficient use of fertilization.

## 2 Materials and methods

### 2.1 Experimental design

#### 2.1.1 Experimental environment

This experiment was performed at the College of Agronomy, Shihezi University, Shihezi City, Xinjiang Uygur Autonomous Region (E 86.06°, N 44.32°) (Figure 1). A laboratory incubator was set to a diurnal temperature of 30°C/26°C on a light/dark cycle of 12 h each, with a light intensity of 20,000 Lux. The potting soil was taken from the surface (0–40 cm) of the agricultural experimental field of Shihezi University (E 86.03°, N44.18°); the soil texture was medium loam, with alkaline hydrolyzable nitrogen, organic matter, available phosphorus, and available potassium contents of 103.8 mg/kg, 18.1 mg/kg, 8.91 mg/kg, and 290.3 mg/kg, respectively, and a field water holding capacity of 24.6%. After air drying and sieving, 3 kg of soil was weighed into flowerpots with a uniform upper diameter of 40 cm and a height of 30 cm.

#### 2.1.2 Experimental treatment

The cotton variety used was Xinluzao No. 53, the main local cultivar. Five nitrogen treatments (0 g/pot (N0), 1.876 g/pot (N1), 5.628 g/pot (N2), 7.504 g/pot (N3), and 9.380 g/pot (N4) of pure

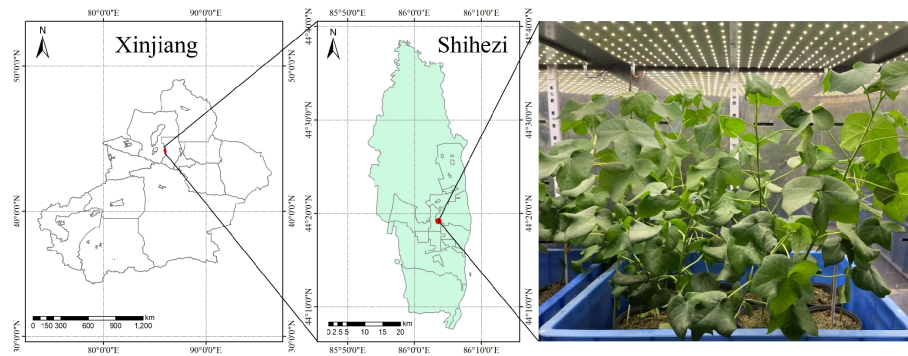


FIGURE 1  
Location and environmental overview of the test area.

nitrogen) were established. Each treatment had nine repetitions, and N2 was the standard nitrogen application treatment. A total of 45 pots were planted, with two cotton plants per pot. Pure phosphorus and pure potassium were applied at 1.876 g/pot and 1.876 g/pot, respectively. The experimental fertilizers used were urea (N 46%), calcium superphosphate ( $P_2O_5$  12%), and potassium chloride ( $K_2O$  60%). All phosphate and potassium fertilizers were applied before sowing; 25% of the nitrogen fertilizers were applied basally, and the remaining 75% were drip-applied with water. Three drip applications were made during the cotton cultivation period. The other management schemes of the pot experiments followed local cotton cultivation management procedures.

## 2.2 Data acquisition and processing

### 2.2.1 Data acquisition

#### 2.2.1.1 Hyperspectral data acquisition

In this experiment, a portable spectral surface spectrometer SR-3500 (Spectral Evolution) was used as the hyperspectral data acquisition instrument, which has a wavelength range of 350–2500 nm at 1 nm intervals. Spectral data were collected at 10:00–14:00 local time on sunny days 60, 80, and 100 days after cotton emergence. Three main cotton stem leaves with uniform growth were selected for each treatment, and non-destructive data

acquisition was conducted using the spectrometer's leaf clamp and built-in light source (Figure 2A). The leaves were labeled with serial numbers. Three positions were measured for each leaf (Figure 2B), and the average data of each of the three measurement points were used as the raw spectral data (R) for that leaf, with whiteboard correction before measuring different leaves; care was taken to avoid leaf veins.

#### 2.2.1.2 Chlorophyll fluorescence data acquisition

The non-destructive sampling of cotton leaf fluorescence parameters was performed using a MultispeQ multifunctional phytometer (Figure 2C), and the measurement times were 10:00–14:00 under light adaptation and 22:00–03:00 under dark adaptation on sunny days. Three positions were measured for each leaf (Figure 2B) and the average was taken as the raw fluorescence data for that leaf.

#### 2.2.1.3 Image data acquisition

A Nikon D5300 digital camera was used to capture cotton leaf images from 10:00–14:00 on sunny days. Cotton leaf images were captured simultaneously as hyperspectral and fluorescence data under light adaptation to avoid errors caused by different acquisition times of data. A white background plate was used as the bottom surface to take photos, and the collection sequence of the blades was consistent with the hyperspectral and fluorescence

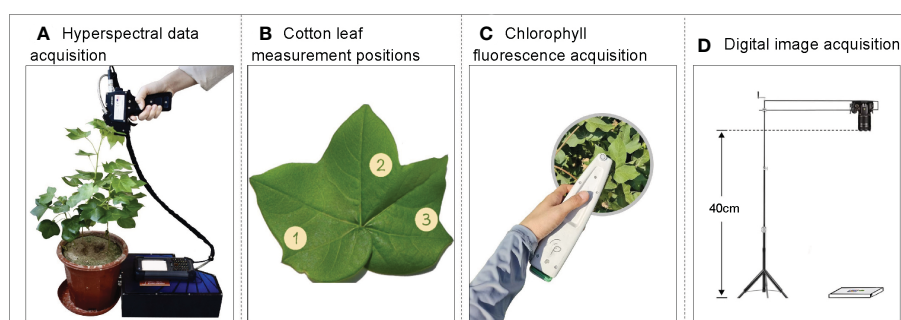


FIGURE 2  
Schematic diagram of data acquisition. (A) Hyperspectral data acquisition, (B) Leaf measurement positions, (C) Chlorophyll fluorescence, and (D) digital image acquisition.

data, which was convenient for subsequent collation and analysis. A 24-color standard color-correction card was placed on the background board for image color correction. When taking photos, the lens and leaf were vertical and kept at a fixed height of 40 cm (Figure 2D). The images were saved in JPEG format.

#### 2.2.1.4 Cotton leaf nitrogen data acquisition

After the spectral, fluorescence, and image data were collected, cotton leaves were harvested, stored in envelopes with ice bags, and brought to the laboratory. The cotton leaves were fixed at 105°C for 30 min, dried at 80°C to a constant weight, crushed, and then heated for digestion using the H<sub>2</sub>SO<sub>4</sub>-H<sub>2</sub>O<sub>2</sub> method. The total nitrogen content was determined using the Kjeldahl method.

## 2.2.2 Data processing

### 2.2.2.1 Hyperspectral data processing method

Data processing was implemented using Python 3.8 programming, and the raw hyperspectral data were first convolutionally smoothed (Savitzky-Golay) to improve the signal-to-noise ratio. Then, the second-order differentiation of the spectral logarithm ([lg(SG)]<sup>2</sup>) was used to eliminate baseline drift and background signals to improve the analytical accuracy (Wang, 2013). Finally, a standard normal variate (SNV) and detrending were used to eliminate spectral errors caused by solid particle size, surface scattering, and light range variations (Jin et al., 2013).

Because there is homogeneity and redundancy in full-band spectral data, two preprocessing methods, SG-SNV-Detrending and [lg(SG)]<sup>2</sup>, were selected to improve the signal-to-noise ratio and analysis accuracy in this study. Pearson's correlation analysis was conducted on the spectra obtained using different pretreatment methods and cotton leaf nitrogen content. The feature bands that significantly correlated at the 0.01 level were selected, and the redundant information was eliminated using the random-frog method to reduce the number of nitrogen-sensitive feature bands while ensuring minimum collinearity among the selected bands.

### 2.2.2.2 Chlorophyll fluorescence data processing method

The variable fluorescence (Fv), PSII maximum photochemical efficiency (Fv'/Fm'), non-photochemical quenching coefficient (NPQ), and photochemical quenching coefficient (qL) were calculated from the parameters of initial fluorescence (F0), maximum fluorescence (Fm), steady-state fluorescence intensity (Fs), initial fluorescence under light adaptation (F0'), maximum fluorescence under light adaptation (Fm'), and variable fluorescence under light adaptation (Fv'). Chlorophyll fluorescence eigenvectors were screened using Pearson correlation analysis for subsequent modeling. The formula for calculating the fluorescence parameters is as follows:

$$Fv = Fm - F0$$

$$Fv' / Fm' = \frac{(Fm' - F0')}{Fm'}$$

$$NPQ = \frac{Fm}{Fm'} - 1$$

$$qL = \frac{(Fm' - Fs)}{(Fm' - F0')} \times \frac{F0'}{Fs}$$

### 2.2.2.3 Image data processing method

Cotton leaf image processing includes color correction, segmentation, and extraction of the color and morphological features of the region of interest. First, the original image (Figure 3A) was corrected using the 24-color-correction card, which reduces the errors caused by solar illumination or the external environment on the color features of the leaf. Then, to eliminate the influence of the color-correction card on cotton leaf segmentation, the maximum outline of the cotton leaf was found to obtain its minimum external rectangle (Figure 3B), the RGB (a color space, composed of red, green, and blue color channels) image was converted into HSV (a color space, composed of hue, saturation, and value channels) format, and the interval of interest to be segmented was detected using the minimum and maximum HSV values. The noise was removed using a closing operation, and the mask image was finally obtained (Figure 3C).

The color features of the cotton leaf images were extracted using Python-OpenCV2. First, the color target image (Figure 3D) was converted into different color spaces to obtain the RGB, HSV, and L\*a\*b\* (a color space composed of a luminance channel and a and b color channels), which are used for color feature extraction of cotton leaves. Finally, nine color feature parameters, including the first-order color moment in the RGB, HSV, and L\*a\*b\* color space of cotton leaves, were obtained.

The size of the leaf area directly determines the intensity of photosynthesis and the final yield of crops (Begonia and Begonia, 2007). Relevant studies have confirmed a significant correlation between crop leaf area and nitrogen content (Zhang, 2019). The morphological features of the cotton leaf area in this study were extracted using the grid method. First, the 24-color-correction card grid image was segmented based on the color feature information (Figure 3E). Because the color-correction card is a customized standard grid of 10 cm × 10 cm, the pseudo-leaf area to be measured can be represented by the number of pixels. The number of pixel points measured on the leaf (pseudo-leaf area) is multiplied by the scaling ratio (K) to calculate the actual area of the leaf, which is used as the morphological feature parameter of the leaf for subsequent modeling. K can be expressed as follows:

$$K = \frac{L}{X}$$

where *L* is the actual area of the color-correction card, which is 100 cm<sup>2</sup>, and *X* is the number of pixels in the pixel coordinate system of the card.

### 2.2.2.4 Nitrogen data processing method

The total nitrogen content of all cotton leaves obtained in 2.2.1 was summarized for processing, and the data marked as NULL due to experimental mis-operation were removed as outliers. Considering the clear difference in data values under different nitrogen application levels, outliers cannot be eliminated using discretization. Therefore, the data under different nitrogen



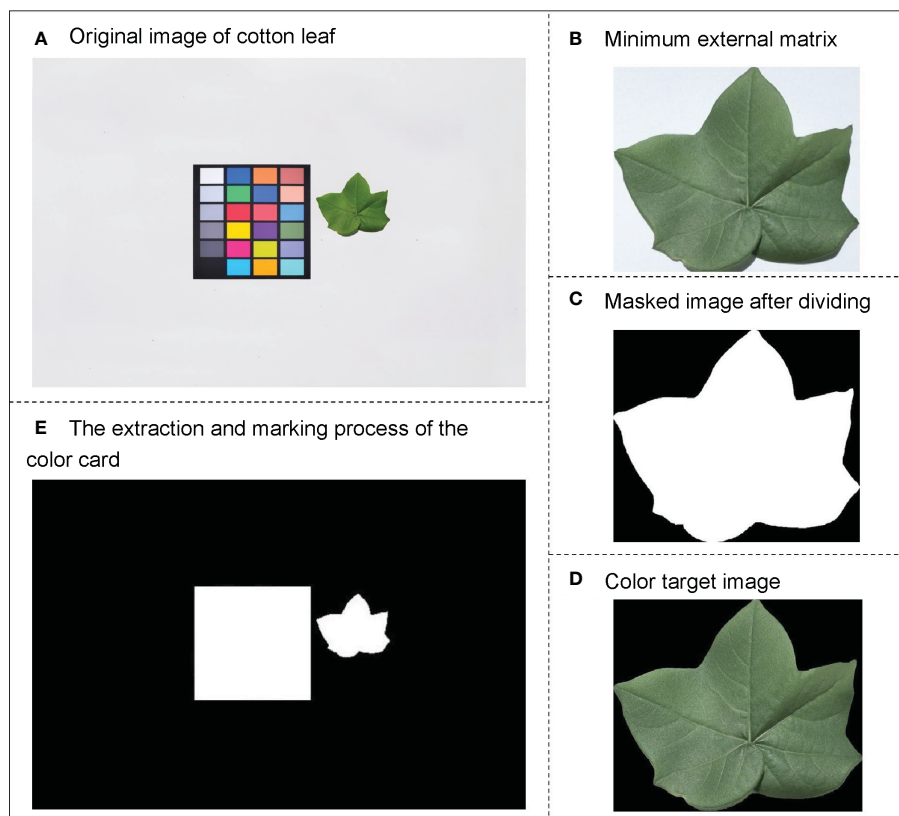


FIGURE 3

Schematic diagram of image data processing. (A) Original image of cotton leaf, (B) Minimum external matrix, (C) Masked image after dividing, (D) Color target image, and (E) Extraction and marking process of color card.

application levels were separated, and the outliers were eliminated again using the criteria of the quartile method: when the nitrogen value was less than the 25th percentile minus 1.5 times the interquartile range (IQR) or greater than the 75th percentile plus 1.5 times the IQR, the data point was eliminated as an outlier.

## 2.3 Model construction and evaluation

Of the 289 groups of cotton leaf data samples, 40, 58, 63, 69, and 59 belonged to the N0, N1, N2, N3, and N4 groups, respectively. The above data samples were randomly divided into training and validation sets using a 7:3 ratio. To evaluate the model's generalization performance, reduce over-fitting, and obtain as much effective information as possible, all machine learning or integrated learning models adopted 5-fold cross-validation.

### 2.3.1 Data-fusion methods

#### 2.3.1.1 Multi-source feature-level fusion method

Feature-level fusion extracts the features from different data sources, such as digital images, chlorophyll fluorescence, and hyperspectral data. The feature vectors extracted from these data sources were connected and substituted into the regression model to obtain inversion results (Figure 4).

The extracted digital image features, chlorophyll fluorescence features, and hyperspectral features were fused using the cascade method. Because integrated learning has strong noise resistance and generalization ability, the stacking integrated learning model was selected for regression prediction. This is a layered integration model; the data set was trained and predicted with the base-learner, and the output value was used as the input value for the next stage of training for the meta-learner to obtain the final forecast result. In this study, the stacking integrated learning algorithm included four learners: random forest (RF), support vector machine (SVM), K-nearest neighbor (KNN), and ridge regression (RR). Subsequently, the appropriate base- and meta-learners were selected to optimize the model fit and generalization ability to obtain the cotton leaf nitrogen inversion results through feature-level fusion.

#### 2.3.1.2 Multi-source decision-level fusion method

Decision-level fusion separately models digital image features, chlorophyll fluorescence features, hyperspectral features, and cotton leaf nitrogen content in three dimensions based on the prediction results to match image and text data (Figure 5).

The features extracted from digital images, chlorophyll fluorescence, and hyperspectral data were used for initial machine learning modeling; RF, KNN, SVM, and RR were used for modeling and evaluation, and the prediction results of the optimal model were

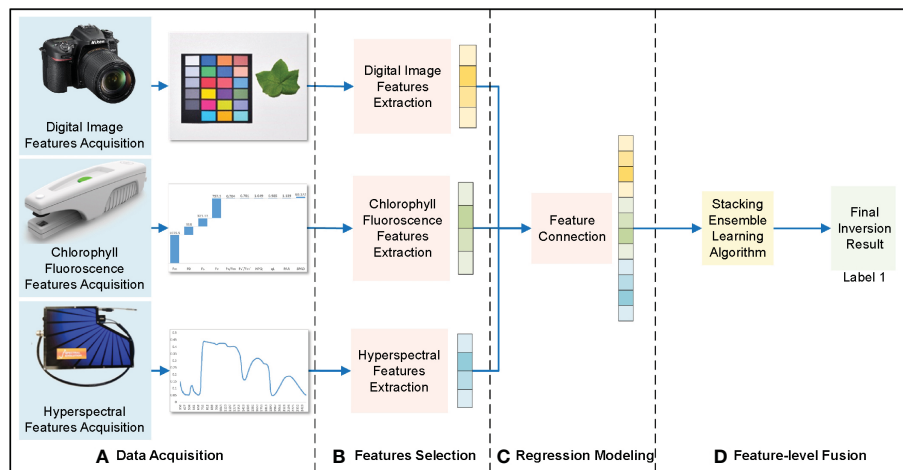


FIGURE 4 Feature-level fusion framework of cotton leaf nitrogen inversion. (A) Data acquisition, (B) Features selection, (C) Regression modeling, (D) Feature-level fusion.

selected. The prediction results of the three data sources were put into the stacking integrated learning algorithm for regression prediction, and the decision-level fusion of cotton leaf nitrogen inversion results was obtained.

### 2.3.1.3 Multi-source hybrid-fusion method

Hybrid fusion is similar to decision-level fusion; the feature-level and decision-level fusion results were fed second into the stacking integrated learning algorithm for regression prediction to obtain hybrid fusion cotton leaf nitrogen inversion results (Figure 6).

### 2.3.2 Model evaluation

The determination coefficient ( $R^2$ ) and root-mean-square error (RMSE) were selected as evaluation indicators of the model.  $R^2$  indicates a deviation between the predicted and actual values; larger

$R^2$  values indicate better fitting models. RMSE analyzes the accuracy of the predicted and the actual values: smaller RMSEs indicate smaller deviations between the predicted and actual values and higher model accuracy. The calculation formulas for  $R^2$  and RMSE are as follows:

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

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

where  $x_i$  is the predicted value,  $y_i$  is the actual value,  $\bar{y}_i$  is the average of the actual values, and  $n$  is the number of samples available for validation.

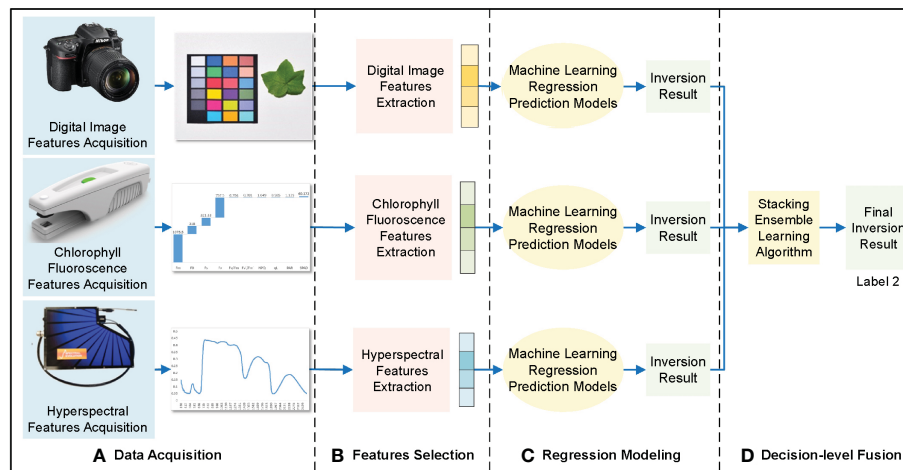


FIGURE 5 Decision-level fusion framework of cotton leaf nitrogen inversion. (A) Data acquisition, (B) Features selection, (C) Regression modeling, (D) Feature-level fusion.

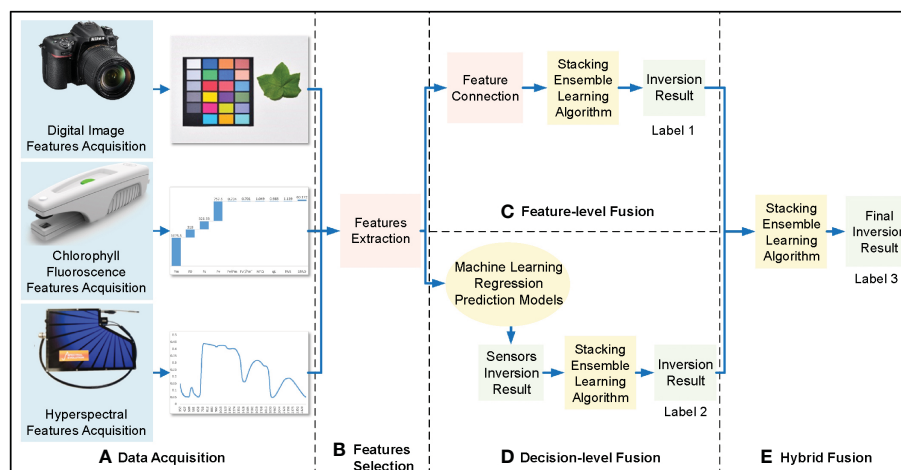


FIGURE 6

Hybrid-fusion framework of cotton leaf nitrogen inversion. (A) Data acquisition, (B) Features selection, (C) Regression modeling, (D) Feature-level fusion.

## 3 Results and analysis

### 3.1 Screening of nitrogen-sensitive feature parameters based on different data sources

#### 3.1.1 Screening of hyperspectral feature parameters

The method in Section 2.2.2.1 was used to obtain the hyperspectral feature parameters. The results showed that there were approximately 10 infrared-sensitive feature bands of nitrogen content selected under the different treatments (Table 1), but with some differences due to the environment and variety. The sensitive feature band of processed spectral information is consistent between SG-SNV-Detrending and R processing, but the  $[\lg(\text{SG})]^n$  processed spectral fluctuation amplitude is larger than that in R processing. This shows that different spectral preprocessing methods affect the selection of sensitive feature bands.

#### 3.1.2 Screening of chlorophyll fluorescence feature parameters

Nitrogen can alter the photosynthetic capacity of crop leaves; fluorescence, as a direct probe for photosynthesis, can be an indicator of crop nitrogen content. In this study, the correlation between chlorophyll fluorescence parameters and cotton leaf nitrogen content was analyzed. The results showed a significant correlation between the chlorophyll fluorescence parameters and cotton leaf nitrogen content (Figure 7). Fm, Fv, Fv/Fm, Fv'/Fm', qL, and NPQ correlated with nitrogen content, and the absolute value

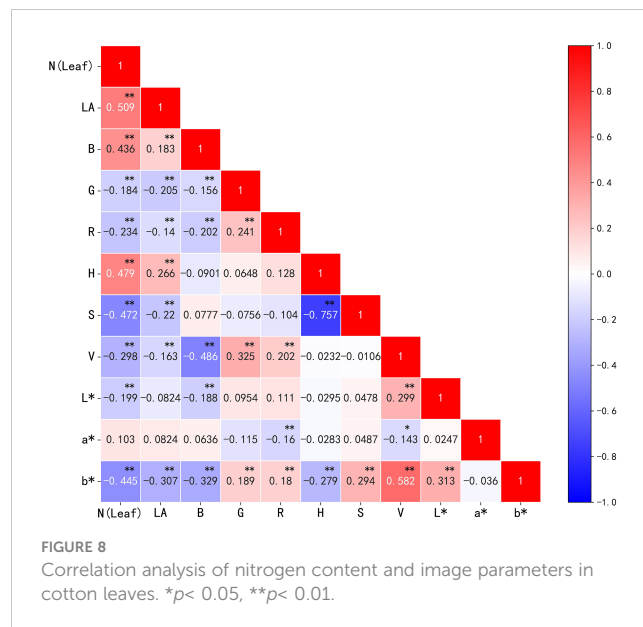
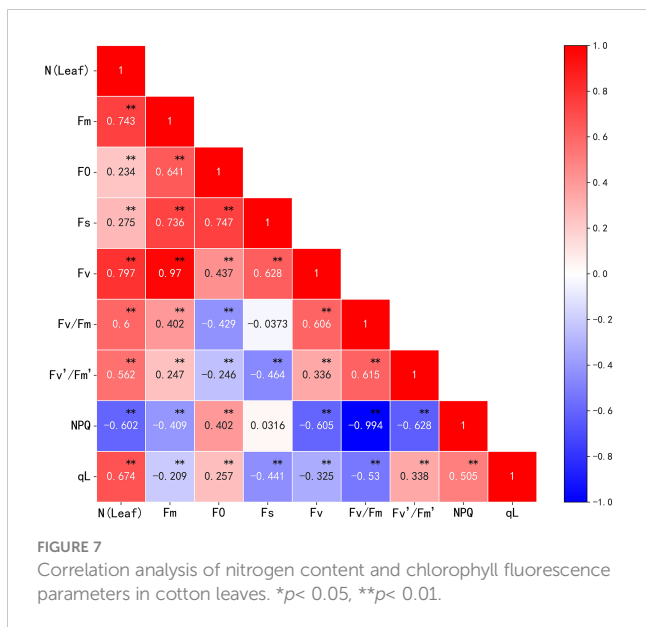
of the correlation coefficient was greater than 0.5. All fluorescence parameters except NPQ correlated positively with cotton leaf nitrogen content. The correlation coefficients between Fm and Fs, Fm and Fv, F0 and Fs, and NPQ and Fv/Fm were 0.736, 0.970, 0.747, and -0.994, respectively. This shows a strong correlation between the corresponding parameters, high information overlap, and some multicollinearity. Therefore, principal component analysis (PCA) was used for linear transformation to remove feature redundancy and noise reduction. The principal component information loads (variance percentages) were 45.507%, 35.048%, 17.780%, 1.181%, 0.438%, 0.036%, 0.011%, and  $2.819 \times 10^{-15}\%$ . This indicates that the first, second, third, and fourth principal components carried more feature information, and the cumulative variance rate reached 99.516%. If the number of principal components continues to increase, the information load will fall below 1%, and the cumulative variance rate will no longer increase significantly. Therefore, the first four principal components were used as chlorophyll fluorescence features for subsequent modeling in this study.

#### 3.1.3 Screening of image feature parameters

The correlations between image parameters (the first-order color moment in the three color spaces and the morphological parameter leaf area (LA)) and the nitrogen content in cotton leaves were analyzed. The results (Figure 8) showed that all parameters except the color parameter  $a^*$  were significantly correlated. LA, B, H, and  $a^*$  correlated positively with nitrogen content, whereas G, R, S, V,  $L^*$ , and  $b^*$  correlated negatively with nitrogen content. Where LA, B, H,

TABLE 1 Screening results of hyperspectral sensitive bands under different treatments.

Pretreatment	Feature Sensitive Bands
R	2212, 2225, 2224, 1253, 2222, 2307, 1704, 2235, 1724, 2213, 1728, 1702, 2337
$[\lg(\text{SG})]^n$	1293, 1174, 1077, 1123, 1650, 1251, 1100, 1244, 1490, 1665
SG-SNV-Detrending	2178, 2213, 2212, 2214, 1725, 2197, 2196, 2195, 1728, 1726



S, and b\* had a highly correlated relationship with nitrogen content,  $R^2$  reached 0.509, 0.436, 0.479, -0.472, and -0.445, respectively, and the correlation between the above parameters was low, indicating that the feature redundancy was low and the collinearity was weak; this means these parameters can be used as the image features for the subsequent modeling processing.

### 3.2 Cotton leaf nitrogen content estimation model based on single data source in fusion framework

#### 3.2.1 Cotton leaf nitrogen content estimation model based on hyperspectral features

The spectral features obtained from Section 3.1.1 were substituted into supervised learning regression algorithms of different principles to model and evaluate the nitrogen content in cotton leaves. We used RF, KNN, SVM, and RR models to avoid duplication of effort due to weak differences between models. RF is a set of multiple decision trees, KNN performs regression using finding the nearest neighbors of the samples, SVM performs regression using the kernel method, and RR is an improved partial least squares regression. From the results of the model evaluation indices (Figure 9), the estimation accuracy of the sensitive feature bands based on R under four model training events was low, with an  $R^2$  value of 0.352, which is lower than that obtained from SG-SNV-Detrending and [lg(SG)] preprocessing. Sensitive feature bands screened using [lg(SG)] and SG-SNV-Detrending pretreatment were substituted into the four models, and  $R^2$  was approximately 0.4–0.7, which can be used for quantitative evaluation of nitrogen content. When constructing the RF model with [lg(SG)] pretreatment, the training set  $R^2$  was 0.965, RMSE was 1.532, and the fit of the model was excellent. However, the validation set  $R^2$  and RMSE were lower than those of the SVM model under the same pretreatment, indicating that its generalization performance is inferior to that of the SVM model. By

comprehensive comparison, the SVM model built with spectral features under the [lg(SG)] pretreatment had an optimal estimation of cotton leaf nitrogen content, with a training set  $R^2$  and RMSE of 0.746 and 3.811, and a validation set  $R^2$  and RMSE of 0.683 and 3.841, respectively.

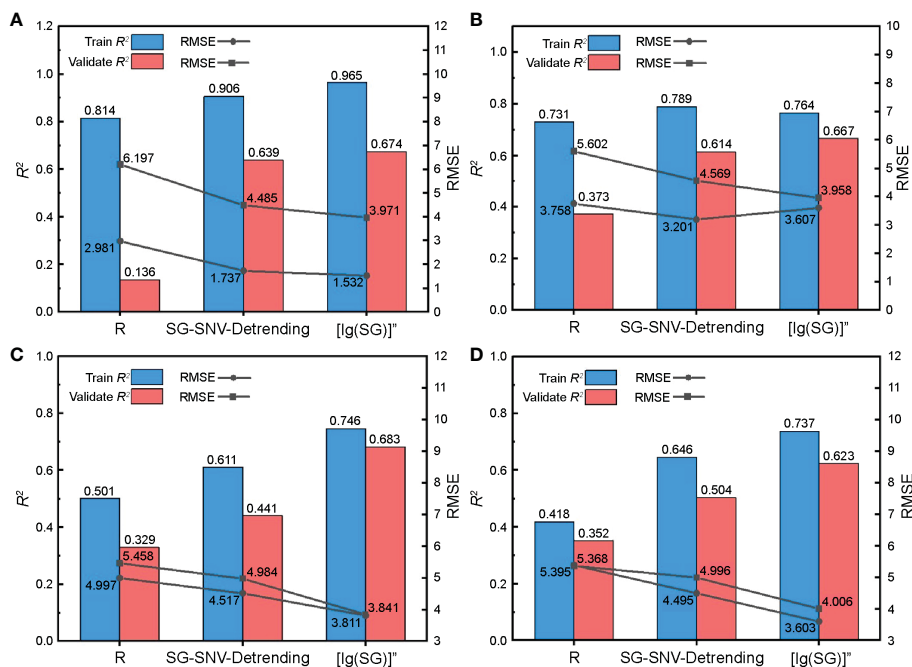
#### 3.2.2 Cotton leaf nitrogen content estimation model based on chlorophyll fluorescence features

The model construction method is consistent with the hyperspectral estimation model in Section 3.2.1; that is, RF, KNN, SVM, and RR machine learning models were used to establish a relationship between the chlorophyll fluorescence features screened in Section 3.1.2 and the nitrogen content of cotton leaves. The results (Figure 10) show that the  $R^2$  of the RF, KNN, and RR training and validation sets were greater than 0.6, and the RMSE was less than 5, indicating that the above three models can estimate nitrogen content. The SVM had the effect of under-fitting. The training set  $R^2$  was 0.457, and the validation set  $R^2$  was 0.525. The model-learning ability was insufficient, and the generalization ability was weak. By comprehensive comparison, the  $R^2$  and RMSE of the estimation model of cotton leaf nitrogen content based on RF were optimal, with a training set  $R^2$  and RMSE of 0.893 and 1.790, and a validation set  $R^2$  and RMSE of 0.702 and 4.086, respectively, which can be used for cotton leaf nitrogen content estimation based on chlorophyll fluorescence features.

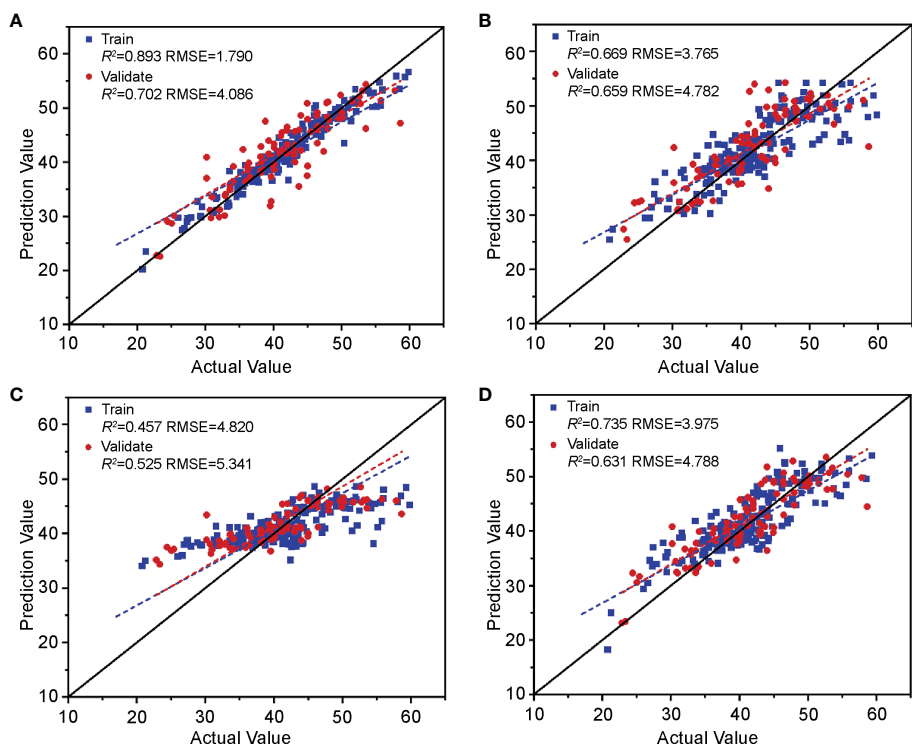
#### 3.2.3 Cotton leaf nitrogen content estimation model based on image features

The above four machine learning models were established based on the image features selected in Section 3.1.3 and the cotton leaf nitrogen content. According to the evaluation results of each model training and validation set (Figure 11), the  $R^2$  of RF, SVM, and RR was approximately 0.7. The RMSE was approximately 3.9, indicating that the RF, SVM, and RR algorithms have excellent

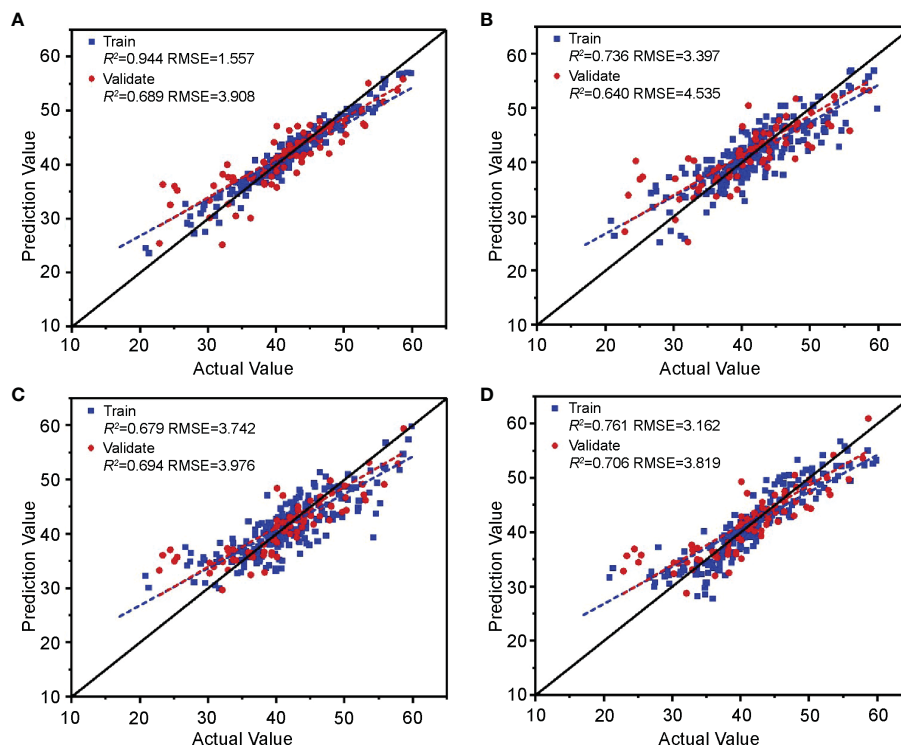




**FIGURE 9** Validation of the cotton leaf nitrogen content estimation model based on hyperspectral features. Validation results of (A) random forest (RF), (B) K-nearest neighbor (KNN), (C) support vector machine (SVM), and (D) ridge regression (RR) models.



**FIGURE 10** Validation of the cotton leaf nitrogen content estimation model based on chlorophyll fluorescence features. Validation results of (A) RF, (B) KNN, (C) SVM, and (D) RR models.



**FIGURE 11**  
Validation of the cotton leaf nitrogen content estimation model based on image features. Validation results of (A) RF, (B) KNN, (C) SVM, and (D) RR models.

accuracy and stability in modeling and can reliably estimate the nitrogen content in cotton leaves, although the  $R^2$  of KNN was 0.64, and the RMSE was 4.535. By comprehensive comparison, the estimation model of cotton leaf nitrogen content based on RR was optimal, with a training set  $R^2$  and RMSE of 0.761 and 3.162, and a validation set  $R^2$  and RMSE of 0.706 and 3.819, respectively, which can be used for cotton leaf nitrogen content estimation based on image features.

### 3.3 Cotton leaf nitrogen content estimation model based on “image-spectrum-fluorescence” data fusions

#### 3.3.1 Cotton leaf nitrogen content estimation model based on feature-level fusion

The feature parameters of hyperspectral, chlorophyll fluorescence, and digital images screened in Section 3.1 were cascaded. In feature-level fusion, after permutation, combination, and screening, the learners in stacking integrated learning obtained the optimal combination using RF and RR as base-learners and SVM as a meta-learner. It can be seen from the evaluation results (Figure 12A) that the training set  $R^2$  of the feature-level fusion model is 0.933 and the RMSE is 1.581, which shows that the goodness of fit of the model to the training set data is excellent, and that it has comprehensively learned the properties of the

training set data. The validation set  $R^2$  was 0.752, and the RMSE was 3.806, indicating that the strong learning ability did not weaken its generalization, and the accuracy remained stable.

#### 3.3.2 Cotton leaf nitrogen content estimation model based on decision-level fusion

The prediction results of the three optimal models constructed using the hyperspectral, chlorophyll fluorescence, and digital image features in Section 3.2 were substituted into the decision-level fusion method for training and validation. Among them, the optimal combination of stacking integrated learning in decision-level fusion included KNN and RR as base-learners and RF as a meta-learner. (Figure 12B) The decision-level fusion model had a training set  $R^2$  and RMSE of 0.997 and 0.349, and a validation set  $R^2$  and RMSE of 0.771 and 3.558, respectively. Although the  $R^2$  value of the decision-level fusion training set is close to 1, the validation set  $R^2$  shows excellent model performance with no significant decline. This can be used for cotton leaf nitrogen content estimation. Combined with the RMSE, the model's accuracy is slightly better than the feature-level fusion model.

#### 3.3.3 Cotton leaf nitrogen content estimation model based on hybrid fusion

The feature-level and decision-level fusion prediction results were substituted into the hybrid fusion for modeling, and training

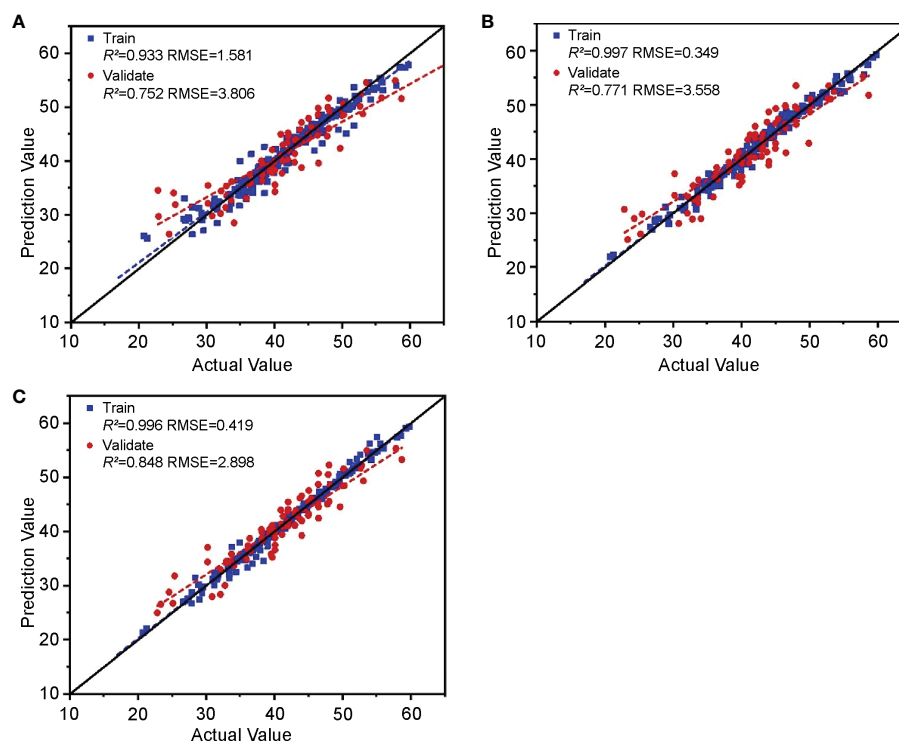


FIGURE 12

Validation of the cotton leaf nitrogen content estimation model based on “image-spectrum-fluorescence” data fusion. Validation results of (A) feature-level fusion, (B) decision-level fusion, and (C) hybrid-fusion models.

validation was conducted. Among them, the optimal combination of stacking integrated learning in hybrid fusion included KNN and RR as base-learners and RF as meta-learner. The results (Figure 12C) show that, in the hybrid-fusion model, the training set  $R^2$  and RMSE were 0.996 and 0.419, and the validation set  $R^2$  and RMSE were 0.848 and 2.898, respectively. Compared with the feature-level and decision-level fusion, the linear fitting relationship between the model’s actual and the predicted values was closer to a 1:1 line. Therefore, the hybrid-fusion model was the best among the cotton leaf nitrogen content estimation models based on “spectrum-fluorescence-image” data fusion.

### 3.4 Models performance comparison test

The algorithm’s advantages enabled us to compare single data source models with multilevel fusion models in the same algorithm dimension by avoiding false improvement in fusion model accuracy. Therefore, the features of the spectrum, chlorophyll fluorescence, and image were substituted into the stacking integrated learning algorithm, and the best combination of learners was selected for training and validation, aiming to unify the algorithm that outputs the final prediction results in the fusion models. The comparison results are shown in Figure 13. Compared with the three single data source models, the feature-level fusion model validation set  $R^2$  improved by 2.8–6.3%, and the RMSE decreased by 4.7–8.2%. The decision-level fusion model validation

set  $R^2$  increased by 2.5–3.6% and the RMSE decreased by 8.8–9.9%. The  $R^2$  of the hybrid-fusion model validation set increased by 12.4–15.9%, and the RMSE decreased by 25.7–26.6%. These results indicate that all multilevel data-fusion methods improved the accuracy of regression fitting of the three single data sources to different degrees. Among them, the accuracy of the feature-level and decision-level models in predicting the outcome, although improved compared with the single data source models, is not significant, as it can combine the prediction results of feature-level fusion and decision-level fusion with the real values for a second fit, greatly improving the evaluation indicators and achieving optimal accuracy and stability of the model. Therefore, the hybrid-fusion model can be applied for cotton leaf nitrogen content monitoring.

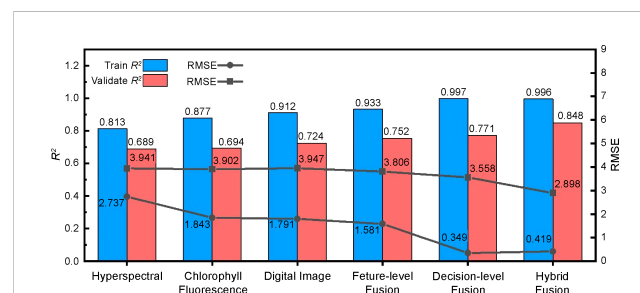


FIGURE 13

Comparison of the precision of the optimal model between multilevel data fusion and three single data sources.

## 4 Discussion

Applying hyperspectral techniques to invert parameters, such as leaf nitrogen content, is one of the main tools for studying the nitrogen nutrient status of cotton. First, preprocessing and data transformation of the original feature spectra are important steps to effectively improve the accuracy of the model. Previous studies have shown that preprocessing of spectra can effectively eliminate noise and baseline drift and highlight the location of spectral feature bands (Arneberg et al., 2007). The accuracy of the models built using relevant studies (Nguyen and Lee, 2006; Li et al., 2016; Wang et al., 2020; Li et al., 2020b) using different forms of spectral transformations is higher than the models built using the original spectra. Feature screening mainly solves the collinearity problem of the spectrum, and wavelengths with less redundant information are selected for direct model construction, which is a simple, fast, and intuitive method easy to implement (Bai et al., 2018; Jia et al., 2019). In this study, the [lg(SG)]<sup>n</sup>-SVM model was the most accurate (training set  $R^2 = 0.746$ , RMSE = 3.811, validation set  $R^2 = 0.683$ , RMSE = 3.841). The original spectra can effectively extract the feature parameters and improve model accuracy using both transformation forms. To prevent overlap between the hyperspectral feature bands obtained from the experiment and the visible region acquired from the digital images, we used the first 10–13 near-infrared region-sensitive bands for modeling to ensure differences between the data sources, reduce the collinearity of each variable of the fusion model, and improve the accuracy and stability, and the selected sensitive bands were similar to those from previous studies (Niu et al., 2000; Xue et al., 2003).

Spectroscopic technology can be used to effectively monitor the nitrogen content of crops. However, spectral reflectance mainly reflects the concentration of plant biochemical components, is not sensitive to crop photosynthetic activity, and cannot directly reveal the photosynthetic physiological state of vegetation (McMurtrey et al., 2003; Ashourloo et al., 2014). Although chlorophyll fluorescence can provide information directly related to plant physiological functions (Smith et al., 2018), it has been widely used for crop growth, development, and nutrition monitoring in recent years. In this study, by analyzing the relationship between chlorophyll fluorescence parameters and nitrogen content in cotton leaves, we found that the correlations between Fm, Fv, Fv/Fm, Fv'/Fm', qL, NPQ, and nitrogen content were stronger than the correlations between fluorescence parameters and nitrogen content of wheat plants established by Feng et al. (2012) and Zivcak et al. (2015). This indicated that different degrees of correlation exist between the fluorescence parameters and nitrogen content for different crops. Because of the strong multicollinearity among individual parameters, PCA was used to reduce the dimensionality, and variables with high information overlap were regrouped to reduce feature redundancy while retaining most of the valid information. The first four principal components were selected to establish a model for estimating the nitrogen content of cotton leaves based on chlorophyll fluorescence characteristics, and the results showed the best RF (training set  $R^2 = 0.893$ , RMSE = 1.790, validation set  $R^2 = 0.702$ , RMSE = 4.086); RF

is an integration of multiple decision trees with high error balance and noise immunity and is widely used by researchers for crop phenotypic indicator monitoring with good modeling results (Han et al., 2019; Lu et al., 2019; Wang et al., 2021).

Digital image technology can identify the subtle differences between the surface layers of samples within the range of visible light. When collecting digital images of cotton leaves, it has been found that the influence of uncertain factors, such as the external environment and light intensity, is likely to lead to inaccurate and inconsistent image colors (You et al., 2020). In this study, the effect of these random errors on image color features was minimized by taking pictures in sufficient sunlight and correcting the standard color card. Huang et al. (2020) and Agarwal and Gupta (2018) proved that the fitting and generalization performance of a model built using combining multiple color spaces was better than that using a single color space. Therefore, in this study, the first moments of RGB, HSV, and L\*a\*b\* color spaces and LA parameters were used to establish correlations with cotton leaf nitrogen content, extract as many of the relevant features of the image as possible from multiple angles, integrate more comprehensive information, and improve the accuracy of the inversion results of the subsequent fusion algorithm. All components of the RGB color space contain brightness information; therefore, they are vulnerable to external conditions; however, the hue and saturation in the HSV color space are separated from the value, and the color channels a\* and b\* in the L\*a\*b\* color space are separated from the luminance, correcting the leaf color deviation in RGB color spaces to some extent. The color-calibration card is secondarily used to standardize the color of all leaves, avoiding large differences in the models' parameters and results due to different external conditions, light, and other uncertain factors. Therefore, the model established in this study is more general. When inverting models of cotton leaf nitrogen content based on digital image features, the evaluation index value of RR was the best among the four machine learning models (training set  $R^2 = 0.761$ , RMSE = 3.162, validation set  $R^2 = 0.706$ , RMSE = 3.819). This shows that although the correlation between the filtered image features is low, there is also a weak linear relationship. The RR loss is unbiased in exchange for numerical stability; therefore, the fitting is more effective. In addition, RF and SVM showed good capability for estimating the nitrogen content based on image features.

Although data-fusion technology has developed rapidly, its application is mainly limited to urban planning, target classification, and recognition (Xing and Meng, 2018; Hoffmann et al., 2019; Zhao et al., 2021). Few studies on the inversion of agricultural parameters have achieved fusion only at the data and feature levels (Zhou et al., 2017; Li, 2018; Zhang, 2020; Jing et al., 2022), and the improvement of model accuracy and stability has been relatively limited. No multimodal and multilevel fusion regression model has been proposed for crop nutrition monitoring. In this study, a multilevel data-fusion model combining multiple machine learning and stacking integrated learning was built to accurately monitor the nitrogen content of cotton crops to obtain as many phenotypic structural, physiological, and biochemical features related to cotton nitrogen as possible through hyperspectral, chlorophyll fluorescence, and digital image data sources.



Out of the three data source features for modeling, feature-level fusion can obtain the highest level of complementary information (training set  $R^2 = 0.933$ , RMSE = 1.581, validation set  $R^2 = 0.752$ , RMSE = 3.806). Decision-level fusion can further train and learn from the results of individual decisions made using each data source (training set  $R^2 = 0.997$ , RMSE = 0.349, validation set  $R^2 = 0.771$ , RMSE = 3.558). Machine learning is adopted for independent modeling of each data source instead of stacking because: (1) the single data source features substituted into the stacking model do not significantly improve the accuracy; (2) the multilevel fusion model must consider running time and efficiency, and the stacking algorithm is used in this study to output the final result of each fusion framework; and (3) combining the three-node parallel-running stacking models with multiple nested stacking models increases the time consumption and the running load of the entire fusion model greatly. Therefore, the machine learning model is used for single data source modeling.

The hybrid-fusion method can integrate the training results of feature-level and decision-level fusion models, has significant anti-noise ability, and significantly improves the regression accuracy of nitrogen content inversion (training set  $R^2 = 0.996$ , RMSE = 0.419, validation set  $R^2 = 0.848$ , RMSE = 2.898). The precision of the fused models is better than the single data source models under the same algorithm, and the precision of the hybrid-fusion model is optimal. This shows the superiority of data fusion using spectral data, chlorophyll fluorescence parameters, and digital images to estimate cotton nitrogen content. However, the consequent limitation is the increase in data acquisition costs. Different data acquisition devices can be integrated to solve this problem and achieve portable, synchronous, and efficient data acquisition. Follow-up research can apply the above methods to monitoring nutrient elements at different scales during different growth periods of crops, providing a reference and ideas for implementing precision agriculture. With changes in the growth period and other factors, the nitrogen nutrition status of cotton leaves will also change, leading to differences in the parameters related to nutritional status extracted from each data source. Therefore, it is necessary to adjust the method of optimizing feature screening and model construction while continuously enriching the experimental data.

## 5 Conclusion

In this study, three data sources—hyperspectral, chlorophyll fluorescence, and digital camera—were used to monitor cotton leaf nitrogen content, screen the features of each data source, and establish multilevel data-fusion models. The following conclusions were drawn:

(1) The original spectral information was pre-treated using different spectral transformation methods. Bands that correlated significantly with the nitrogen content of cotton leaves were selected; each round of processing yielded approximately 10 near-infrared bands as spectral-based features. Correlation analysis between chlorophyll fluorescence parameters and nitrogen content revealed significant collinearity between some parameters, and the first four principal components based on fluorescence were

selected as features using PCA. Meanwhile, the correlation between the extracted image parameters and the nitrogen content was analyzed, and the parameters with strong correlation, low redundancy and low collinearity were selected as the features based on digital images.

(2) The evaluation index values were optimal when the spectral data were processed using  $[\lg(\text{SG})]^p$ -SVM to estimate nitrogen content, the chlorophyll fluorescence features substituted into RF were used to build the nitrogen content estimation model, and the digital image features substituted into RR were used to build the nitrogen content estimation model. The three optimal estimation models of data sources were used for the steps in the decision-level fusion framework.

(3) This study combined the information obtained from the three data sources to build multilevel data-fusion models. Compared with the single data source models under the same algorithm, multilevel data-fusion models had improved accuracy and stability. They can improve the accuracy and stability of cotton leaf nitrogen inversion to varying degrees, and the hybrid-fusion model had the best results. This provides a feasible method and approach precisely monitoring crop nitrogen nutrition.

## Data availability statement

The original contributions presented in the study are included in the article/supplementary material. Further inquiries can be directed to the corresponding authors.

## Author contributions

Conceptualization, SQ, YD, and ZZho. Methodology, SQ. Validation, SQ and QY. Formal analysis, SQ, HW, FX and ZZho. Investigation, SQ and YD. Resources, SQ, YD and MZ. Data curation, SQ and MZ. Writing—original draft preparation, SQ. Writing—review and editing, SQ, LZ and ZZha. Funding acquisition, ZZha, LZ and XL. All authors contributed to the article and approved the submitted version.

## Funding

This research was funded by the National Natural Science Foundation of China (No. 42061058), Key Areas of Research Program of XPCC (No. 2020AB005), “Strong Youth” Scientific and Technological Innovation Backbone Talents Program of XPCC (No. 2022CB002-01).

## Acknowledgments

The authors are grateful to the Xinjiang Production and Construction Crops Oasis Eco-Agriculture Key Laboratory for their contributions to this study. We are also grateful to the

reviewers for their comments and suggestions, which have significantly improved the quality of this paper.

## Conflict of interest

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

## References

- Agarwal, A., and Gupta, S. D. (2018). Assessment of spinach seedling health status and chlorophyll content by multivariate data analysis and multiple linear regression of leaf image features. *Comput. Electron. Agric.* 152, 281–289. doi: 10.1016/j.compag.2018.06.048
- Arneberg, R., Rajalahti, T., Flikka, K., Berven, F. S., Kroksveen, A. C., Berle, M., et al. (2007). Pretreatment of mass spectral profiles: Application to proteomic data. *Anal. Chem.* 79, 7014–7026. doi: 10.1021/ac070946s
- Ashoorloo, D., Mobasheri, M. R., and Huete, A. (2014). Developing two spectral disease indices for detection of wheat leaf rust (*Puccinia triticina*). *Remote Sens.* 6, 4723–4740. doi: 10.3390/rs6064723
- Bai, L. M., Li, F. L., and Chang, Q. R. (2018). Increasing accuracy of hyper-spectral remote sensing for total nitrogen of winter wheat canopy by use of SPA and PLS methods. *J. Plant Nutr. Fertil.* 24, 1178–1184. doi: 10.11674/zwjyf.17478
- Begonia, G. B., and Begonia, M. T. (2007). Plant photosynthetic production as controlled by leaf growth, phenology, and behavior. *Photosynthetica* 45, 321–333. doi: 10.1007/s11099-007-0056-7
- Feng, W., He, L., Zhang, H. Y., Guo, B. B., Zhu, Y. J., Wang, C. Y., et al. (2015). Assessment of plant nitrogen status using chlorophyll fluorescence parameters of the upper leaves in winter wheat. *Eur. J. Agron.* 64, 78–87. doi: 10.1016/j.eja.2014.12.013
- Feng, W., Li, X., Wang, Y., Wang, C., and Guo, T. (2012). Difference of chlorophyll fluorescence parameters in leaves at different positions and its relationship with nitrogen content in winter wheat plant. *Acta Agron. Sin.* 38 (4), 657–664. doi: 10.3724/SP.J.1006.2012.00657
- Han, N., Shi, P. X., Jin, Z., Zhang, Y., Cai, L. F., Lu, Y., et al. (2022). Development of nitrogen nutrient diagnosis technology and its application in peanut. *Liaoning Agric. Sci.* 1, 50–55. doi: 10.3969/j.issn.1002-1728.2022.01.011
- Han, L., Yang, G., Dai, H., Xu, B., Yang, H., Feng, H., et al. (2019). Modeling maize above-ground biomass based on machine learning approaches using UAV remote-sensing data. *Plant Meth.* 15, 10. doi: 10.1186/s13007-019-0394-z
- He, H. F., Yan, C. H., and Wu, N. (2020). Effects of nitrogen application rate on chlorophyll fluorescence characteristics and dry matter accumulation in switchgrass (*Panicum virgatum*) leaves. *Acta Prataculturae Sin.* 29, 141–150. doi: 10.11686/cyxb2020007
- Hoffmann, E. J., Wang, Y., Werner, M., Kang, J., and Zhu, X. X. (2019). Model fusion for building type classification from aerial and street view images. *Remote Sens.* 11, 1259. doi: 10.3390/rs11111259
- Hong, B., Zhang, Z., and Zhang, Q. (2022). The nitrogen content in cotton leaves: Estimation based on digital image. *Chin. Agric. Sci. Bull.* 38, 49–55. doi: 10.11924/j.issn.1000-6850.casb2021-0539
- Huang, F., Gao, S., and Yao, X. (2020). Estimation of winter wheat leaf nitrogen concentration using machine learning algorithm and multi-color space. *J. Nanjing Agric. Univ.* 43, 364–371. doi: 10.7685/j.issn.201903073
- Jia, M. (2016). *Monitoring chlorophyll fluorescence parameters at canopy and leaf scales on hyperspectral reflectance in wheat* (Nanjing: Nanjing Agricultural University).
- Jia, M., Li, W., Wang, K., Zhou, C., Cheng, T., Tian, Y., et al. (2019). A newly developed method to extract the optimal hyperspectral feature for monitoring leaf biomass in wheat. *Comput. Electron. Agric.* 165, 104942. doi: 10.1016/j.compag.2019.104942
- Jin, N., Huang, W., Ren, Y., Luo, J., Wu, Y., Jing, Y., et al. (2013). Hyperspectral identification of cotton verticillium disease severity. *Optik* 124, 2569–2573. doi: 10.1016/j.ijleo.2012.07.026
- Jing, X., Song, X., Zhang, T., Duan, W., and Su, J. (2022). Remote sensing monitoring of wheat stripe rust based on fractional order differential and gaussian process methods. *Spectrosc. Spectral Anal.* 42, 859–865. doi: 10.11975/j.issn.1002-6819.2021.17.016
- Kumar, K. S., Dahms, H. U., Lee, J. S., Kim, H. C., Lee, W. C., and Shin, K. H. (2014). Algal photosynthetic responses to toxic metals and herbicides assessed by chlorophyll a fluorescence. *Ecotoxicol. Environ. Saf.* 104, 51–71. doi: 10.1016/j.ecoenv.2014.01.042
- Li, T. (2018). *Remote sensing monitoring of wheat multi-parameter biomass based on multi-source data* (Ningxia: Ningxia University).
- Li, Y., Chen, D., Walker, C. N., and Angus, J. F. (2010). Estimating the nitrogen status of crops using a digital camera. *Field Crops Res.* 118, 221–227. doi: 10.1016/j.fcr.2010.05.011
- Li, D., Lu, X., and Luo, H. (2020a). Estimation model of nitrogen nutrition in drip irrigation cotton based on chlorophyll fluorescence parameters. *Cotton Sci.* 32, 63–76. doi: 10.11963/1002-7807.ldmzz.20191231
- Li, L., Lu, J., Wang, S., Ma, Y., Wei, Q., Li, X., et al. (2016). Methods for estimating leaf nitrogen concentration of winter oilseed rape (*Brassica napus* L.) using *in situ* leaf spectroscopy. *Ind. Crops Prod.* 91, 194–204. doi: 10.1016/j.indcrop.2016.07.008
- Li, Z. F., Su, J. X., and Fei, C. (2020b). Estimation of chlorophyll content in sugar beet under drip irrigation based on hyperspectral data. *J. Agric. Resour. Environ.* 37, 761–769. doi: 10.13254/j.jare.2019.0388
- Li, L., Zhang, M., Ren, T., Li, X., Cong, R., Wu, L., et al. (2015). Diagnosis of nitrogen nutrition in rice using digital image technology. *J. Plant Nutr. Fertil.* 21 (1), 259–268. doi: 10.11674/zwjyf.2015.0129
- Liu, Y., and Li, Z. (2010). Study on nitrogen nutrient diagnosis in maize using digital image technology. *Maize Sci.* 18, 147–149 + 152. doi: 10.13597/j.cnki.maize.science.2010.04.035
- Lu, N., Zhou, J., Han, Z., Li, D., Cao, Q., Yao, X., et al. (2019). Improved estimation of aboveground biomass in wheat from RGB imagery and point cloud data acquired with a low-cost unmanned aerial vehicle system. *Plant Meth.* 15, 17. doi: 10.1186/s13007-019-0402-3
- Luo, J. J., Yang, H. Y., and Lu, Y. (2020). Classification of nitrogen application levels in rice based on hyperspectral and parameter optimized support vector machine. *Soils Fertil. Sci. China* 5, 250–257. doi: 10.11838/sfsc.1673-6257.19537
- McMurtrey, J. E., Middleton, E. M., Corp, L. A., Campbell, P. K., Butcher, L. M., and Daughtry, C. S. (2003). "Optical reflectance and fluorescence for detecting nitrogen needs in *zea mays* L." in *IGARSS 2003. 2003 IEEE International Geoscience and Remote Sensing Symposium. Proceedings (IEEE Cat. No. 03CH37477)*, (New York, NY: IEEE), Vol. 7. 4602–4604.
- Moghaddam, P. A., Derafshi, M. H., and Shayesteh, M. (2010). A new method in assessing sugar beet leaf nitrogen status through color image processing and artificial neural network. *J. Food Agric. Environ.* 8, 485–489. doi: 10.3168/jds.2010-93-4-1785
- Nguyen, H. T., and Lee, B. W. (2006). Assessment of rice leaf growth and nitrogen status by hyperspectral canopy reflectance and partial least square regression. *Eur. J. Agron.* 24, 349–356. doi: 10.1016/j.eja.2006.01.001
- Niu, Z., Chen, Y., and Sui, H. (2000). Analysis on the mechanism of leaf chemical composition imaging spectrum remote sensing. *J. Remote Sens.* 02, 125–130. doi: 10.11834/jrs.20000208
- Ptushenko, V. V., Ptushenko, O. S., and Tikhonov, A. N. (2014). Chlorophyll fluorescence induction, chlorophyll content, and chromaticity characteristics of leaves as indicators of photosynthetic apparatus senescence in arboreal plants. *Biochem. (Mosc)* 79, 260–272. doi: 10.1134/S0006297914030122
- Saberioon, M. M., Amin, M. S. M., Aimrun, W., Gholizadeh, A., and Anuar, A. A. R. (2013). Assessment of colour indices derived from conventional digital camera for determining nitrogen status in rice plants. *J. Food Agric.* 11, 655–662. doi: 10.1080/03632415.2011.633464
- Smith, W. K., Biederman, J. A., Scott, R. L., Moore, D. J. P., He, M., Kimball, J. S., et al. (2018). Chlorophyll fluorescence better captures seasonal and interannual gross primary productivity dynamics across dryland ecosystems of southwestern north America. *Geophys. Res. Lett.* 45, 748–757. doi: 10.1002/2017GL075922
- Tong, Q., Zhang, B., and Zheng, L. (2006). *Hyperspectral remote sensing – principles, technologies and applications* (Beijing: Higher Education Press).
- Wang, X. (2013). Research and application progress of spectral preprocessing methods in near-infrared analysis. *Sci. Technol. Inf.* 15, 2. doi: 10.3969/j.issn.1672-3791.2013.15.002
- Wang, Y., Li, F., and Wang, W. (2020). Hyper-spectral remote sensing estimation of shoot biomass of winter wheat based on SPA and transformation spectra. *J. Triticeae Crops* 40, 1389–1398. doi: 10.7606/j.issn.1009-1041.2020.11.14

## Publisher's note

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

- Wang, X., Miao, Y., Dong, R., Zha, H., Xia, T., Chen, Z., et al. (2021). Machine learning-based in-season nitrogen status diagnosis and side-dress nitrogen recommendation for corn. *Eur. J. Agron.* 123, 126193. doi: 10.1016/j.eja.2020.126193
- Wang, Y., Wang, D., Shi, P., and Omasa, K. (2014). Estimating rice chlorophyll content and leaf nitrogen concentration with a digital still color camera under natural light. *Plant Meth.* 10, 36. doi: 10.1186/1746-4811-10-36
- Wang, B. F., Yu, Z. Y., Cheng, J. P., Li, Y., Zhang, Z. S., and Yang, X. L. (2022). Research progress on the effect of nitrogen on rice yield and quality formation. *J. Huazhong Agric. Univ.* 41, 76–83. doi: 10.13300/j.cnki.hnlkxb.2022.01.007
- Xiao, S. (2022). Application of deep learning in image recognition of crop diseases and insect pests. *Appl. IC* 39 (1), 240–241. doi: 10.19339/j.issn.1674-2583.2022.01.103
- King, H., and Meng, Y. (2018). Integrating landscape metrics and socioeconomic features for urban functional region classification. *Comput. Environ. Urban Syst.* 72, 134–145. doi: 10.1016/j.compenvurbsys.2018.06.005
- Xue, L., Luo, W., and Cao, W. (2003). Advances in spectral diagnosis of crop water and nitrogen. *J. Remote Sens.* 01, 73–80. doi: 10.11834/jrs.20030113
- Yao, X., Zhu, Y., Tian, Y., Feng, W., and Cao, W. (2010). Exploring hyperspectral bands and estimation indices for leaf nitrogen accumulation in wheat. *Int. J. Appl. Earth Obs. Geoinf.* 12, 89–100. doi: 10.1016/j.jag.2009.11.008
- You, M., Liu, J., Zhang, J., Xv, M., and He, D. (2020). A novel chicken meat quality evaluation method based on color card localization and color correction. *IEEE Access* 8, 170093–170100. doi: 10.1109/ACCESS.2020.2989439
- Zhang, L. (2019). *Research on detection of nitrogen content in apple tree leaves based on digital image processing technology* (Tai'an: Shandong Agricultural University). Master's thesis.
- Zhang, H. (2020). *Research on key phenotypic parameters of crop water and fertilizer based on multi-source image information fusion* (Tianjin: Hebei University of Technology).
- Zhang, Z., Ma, L., and Hong, S. (2020). Research on hyperspectral diagnosis of plant nitrogen nutrient index in drip irrigated cotton fields. *Cotton Sci.* 32, 392–403. doi: 10.11963/1002-7807.zzlx.20200716
- Zhang, W. X., Tong, X. M., and Zhou, T. H. (2021). Remote sensing estimation of cotton leaf nitrogen content based on hyperspectral imaging. *J. Shenyang Agric. Univ.* 52, 586–596. doi: 10.3969/j.issn.1000-1700.2021.05.009
- Zhang, Y., Wang, H., and Lei, Q. (2018). Recommended methods for rational nitrogen application in farmland. *Sci. Agric. Sin.* 51, 11. doi: 10.3864/j.issn.0578-1752.2018.15.009
- Zhao, B., Duan, A., Ata-Ul-Karim, S. T., Liu, Z., Chen, Z., Gong, Z., et al. (2018). Exploring new spectral bands and vegetation indices for estimating nitrogen nutrition index of summer maize. *Eur. J. Agron.* 93, 113–125. doi: 10.1016/j.eja.2017.12.006
- Zhao, W., Li, S., Li, A., and Chen, J. (2021). Deep fusion of hyperspectral images and multi-source remote sensing data for classification with convolutional neural network. *Natl. Remote Sens. Bull.* 25 (7), 1489–1502. doi: 10.11834/jrs.20219117
- Zhou, C. Y., Hua, D. X., Le, J., Yan, Q., Mao, J. D., and Jing, M. (2017). Chlorophyll fluorescence dynamics plant water stress detection method combined with image. *Trans. Chin. Soc. Agric. Mach.* 48 (1), 148–154. doi: 10.6041/j.issn.1000-1298.2017.01.019
- Zivcak, M., Olsovska, K., Slamka, P., Galambosova, J., Rataj, V., Shao, H., et al. (2015). Application of chlorophyll fluorescence performance indices to assess the wheat photosynthetic functions influenced by nitrogen deficiency. *Plant Soil Environ.* 60, 210–215. doi: 10.17221/73/2014-PSE