Wang L *et al*. (2024)
Notulae Botanicae Horti Agrobotanici Cluj-Napoca
Volume 52, Issue 2, Article number 13728
DOI:10.15835/nbha52213728
Research Article

Diagnostic study of nitrogen nutrition in cotton based on unmanned aerial vehicle RGB images

Lu WANG[a], Qiushuang YAO[b], Ze ZHANG, Shizhe QIN, Hongyu WANG, Feng XU, Mi YANG, Xin LV*, Lulu MA*

*Shihezi University College of Agriculture/ The Key Laboratory of Oasis Eco-agriculture, Xinjiang Production and Construction Group, Shihezi 832003, China; wanglu98@stu.shzu.edu.cn; 20182012108@stu.shzu.edu.cn; zhangze1227@shzu.edu.cn; 20212012024@stu.shzu.edu.cn; wanghongyu@stu.shzu.edu.cn; 20212012017@stu.shzu.edu.cn; yangmi@stu.shzu.edu.cn; lxshz@126.com(*corresponding author); ma_0517@shzu.edu.cn (*corresponding author)*

Abstract

Nitrogen fertilizer levels significantly affect crop growth and development, necessitating precision management. Most studies focus on nitrogen nutrient estimation using vegetation indices and textural features, overlooking the diagnostic potential of color features. Hence, we investigated cotton nitrogen nutrition status using unmanned aerial vehicle (UAV) image features and the nitrogen nutrient index (NNI). Random frog algorithm - and random forest-screened image feature sets significantly correlated with the NNI, which were substituted into four machine learning algorithms for NNI estimation modeling. The composite scores (F) of optimal image feature sets were calculated using the coefficient of variation method for comprehensive cotton nitrogen nutrient diagnosis. Validation of the model for determining the critical nitrogen concentration in cotton yielded a coefficient of determination R^2 = 0.89, root mean square error RMSE = 0.50 g (100 g)$^{-1}$, and mean absolute error MAE = 0.44, demonstrating improved performance. Additionally, our novel NNI estimation model constructed based on the optimal image feature sets exhibited R^2_c = 0.97, $RMSE_c$ = 0.02, MAE_c = 0.02, R^2_v = 0.85, $RMSE_v$ = 0.05, and MAE_v = 0.04. Polynomial fitting of the composite index with NNI indicated that the model was reliable and yielded the following diagnostic criterion: 0.48 < F2 < 0.67 indicated nitrogen overapplication, whereas F2 < 0.48 or F2 > 0.67 indicated nitrogen deficiency. This study demonstrates the superior effectiveness of using UAV RGB image feature sets for NNI estimation and the quick, accurate diagnosis of cotton nitrogen levels, which will help guide nitrogen fertilizer application.

Keywords: composite score; image features; machine learning; nitrogen nutrition index

Abbreviation: AB, adaptive boosting; AK, available potassium; AN, alkali-hydrolyzable nitrogen; AP, available phosphorus; CIVE, color index of vegetation; Con, contrast; Cor, correlation; Dis, dissimilarity; Ent, entropy; ExB, excess blue vegetation index; ExG, excess green vegetation index; ExR, excess red vegetation index; F, comprehensive score; GLCM, gray level co-occurrence matrix; GLI, green leaf index; GRVI, red-green vegetation index; H, hue; Hom, homogeneity; MAE, mean absolute error; MGRVI, modified green red vegetation index; NGRDI, normalized green-red difference index; NNI, nitrogen nutrient index; OM, organic matter; PLSR, partial least squares regression; R^2, coefficient of determination; RFA: random frog algorithm; RF, random forest; RGBVI, red-green-blue vegetation index; RGRI, red green ratio index; RMSE, root mean

Received: 25 Feb 2024. Received in revised form: 06 Apr 2024. Accepted: 29 Apr 2024. Published online: 21 May 2024.
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.

square error; S, saturation; SM, second moment; SVM, support vector machine; UAV, unmanned aerial vehicle; V, value; VAR, variance; VARI, visible atmospherically resistant index.

Introduction

Cotton is one of the largest fiber crops in the world in terms of cultivated area, dominates the global textile industry, and has high economic value (Jans *et al*., 2021; Li *et al*., 2022). China is the world's largest cotton producer, accounting for one-quarter of the world's total annual cotton production (FAOSTAT, 2023). The rise in prices for nitrogen fertilizer, one of the main agricultural fertilizers in China, increases crop production costs and exacerbates the financial burden on agricultural practices (Jin *et al*., 2024). Furthermore, excessive application of nitrogen during cotton production wastes fertilizer resources, and the limited capacity for crop uptake of nitrogen results in soil nitrogen residues (Fenge *et al*., 2023), triggering problems such as increased soil salinization (Tsakmakise *et al*., 2019). Therefore, timely access to the nitrogen nutritional status of cotton is conducive to reducing cotton production costs and improving environmental protection.

Critical nitrogen concentrations have been used to diagnose nitrogen nutrition throughout the crop life cycle (Soratto *et al*., 2022). The nitrogen nutrition index (NNI) is calculated based on measured and critical nitrogen concentrations. Many studies have shown that the NNI can be used to assess the nitrogen nutritional status of crops, and its diagnostic effect is better than that of a single crop agronomic parameter (Shi *et al*., 2021; Yang *et al*., 2016). Although RGB images acquired with unmanned aerial vehicle (UAV)-mounted visible-light cameras contain less spectral information, they have the advantages of high spatial resolution, low cost, and simple data processing, which make it easier to promote their use (Wang *et al*., 2013). Furthermore, combining the vegetation index and a machine learning algorithm can fully utilize UAV RGB image band information (Han *et al*., 2019). Many scholars have already combined spectral and texture information to estimate crop growth parameters (Yang *et al*., 2019; Zheng *et al*., 2020; Li *et al*., 2019), which not only enhances the utilization of spectral features but also mitigates the saturation phenomenon, thus addressing the shortcomings of the inversion of spectral features. Fu *et al*. (2020) found that combining vegetation indices and color features of UAV images to estimate plant nitrogen concentrations in winter wheat was effective. Therefore, the potential of combining vegetation indices, texture features, and color moments of UAV RGB images to estimate the NNI can be explored. Xu *et al*. (2021) used the coefficient of variation method to weigh chlorophyll content, plant water content, plant height, and biomass of winter wheat to construct a comprehensive growth monitoring index, indicating that the comprehensive index performed well in estimating the crop growth parameter. At present, most studies on cotton NNI estimation use multiple image features as model inputs (Fu *et al*., 2020) and rarely combine image features to construct comprehensive indicators. Therefore, in this study, the coefficient of variation method was used to construct comprehensive indicators to explore the feasibility of replacing NNI in nitrogen nutrition diagnosis.

Ultimately, the purpose of this study was to (1) construct a critical nitrogen concentration dilution model to calculate NNI and explore its effect in diagnosing the nitrogen nutrient status of cotton; (2) use random frog algorithm (RFA) and random forest (RF) to screen image feature sets significantly related to the NNI, as well as partial least squares regression (PLSM), support vector machine (SVM), adaptive boosting (AB), and RF modeling, to obtain the optimal image feature set and optimal model for estimating the NNI; and (3) calculate the comprehensive score (F) of the optimal image feature set based on the entropy method and the coefficient of variation method, which will be used to explore its feasibility in nitrogen nutrition diagnosis and provide technical support for real-time and accurate diagnosis of cotton nitrogen nutrition status.

Materials and Methods

Experimental design

This study was conducted in 2022 in Company II of the experimental field of Shihezi University, Xinjiang Uygur Autonomous Region (44°19′N; 85°59′E). The soil texture of the experimental site was loamy, and the basic physicochemical properties were as follows: alkali-hydrolyzable nitrogen (AN) = 60.88 mg kg^{-1}, organic matter (OM) = 19.90 mg kg^{-1}, available potassium (AK) = 134 mg kg^{-1}, and available phosphorus (AP) = 17.95 mg kg^{-1}. The previous crop was cotton. Additional work in 2023 was conducted at the Shihezi Institute of Agricultural Science, Xinjiang Uygur Autonomous Region (44°20′N; 86°3′E), where the soil texture was loamy and the basic physicochemical properties were as follows: AN = 50.90 mg kg^{-1}, OM = 19.20 mg kg^{-1}, AK = 151 mg kg^{-1}, and AP = 18.52 mg kg^{-1}. The previous crop at this site was maize.

The experiment included five nitrogen application levels: N0 (0 kg hm^{-2}), N1 (120 kg hm^{-2}), N2 (240 kg hm^{-2}), N3 (360 kg hm^{-2}), and N4 (480 kg hm^{-2}). Throughout the entire life cycle of cotton, urea (containing 46% nitrogen) and 160 kg hm^{-2} of phosphorus potassium fertilizer (potassium dihydrogen phosphate) are applied. Fertilization is set with 30% base fertilizer and the remaining 70% applied with water. Drip irrigation fertilization is conducted 8 times throughout the entire growth period. The experiment sets different fertilizer application ratios and amounts under different nitrogen gradients, as shown in Table 1. The experimental variety was Xinlu Early 53, the main cultivar in the surrounding area of Shihezi City, Xinjiang, and the planting pattern was "one film, three tubes and six rows," with a plant spacing of 10 cm + 66 cm + 10 cm. Each nitrogen treatment was repeated three times, with a total of 15 plots in a randomized block design and protective rows around the perimeter. Other conditions were arranged according to the local large fields.

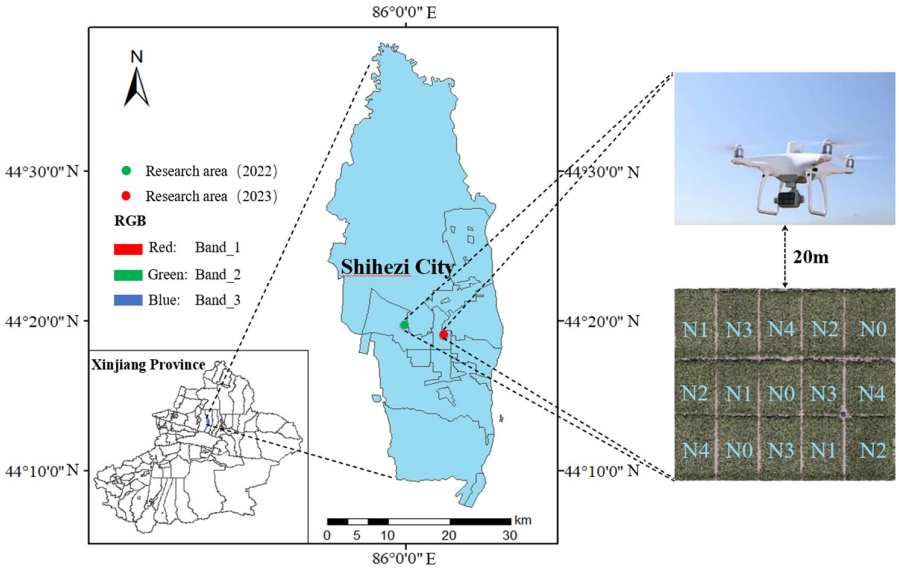

Figure 1. Map of the study area

Acquisition of cotton agronomic parameters

In 2022 (Zhang *et al.,* 2022), samples were collected at the squaring stage (June 21), flowering stage (June 26), full flowering stage (July 3), flowering and boll stage (July 15 and July 25), and full boll stage (August 4). Samples were collected in 2023 at the squaring stage (June 22), flowering stage (July 2), full flowering stage (July 11), flowering and boll stage (July 25 and August 2), full boll stage (August 10), and spitting stage (August 22). Six representative cotton plants were selected from each plot on each sampling day, and the samples were divided into three organ types: stems, leaves, and buds. The samples were deoxidized at 105 °C for 30 min and

baked to constant weight at 80 °C. The dry matter weight of each organ was determined as described previously (Tian *et al.,* 2014).

Table 1. Application date, proportion, and amount of fertilizer in the experimental area

Date		Fertilization ratio	Nitrogen application treatment					Phosphorus and potassium fertilizer
			N0	N1	N2	N3	N4	
2022	June 8th	5%	0	4.20	8.40	12.60	16.80	5.60
	June 16th	11%	0	9.24	18.48	27.72	36.96	12.32
	June 21st	12%	0	10.08	20.16	30.24	40.32	13.44
	June 29th	12%	0	10.08	20.16	30.24	40.32	13.44
	July 5th	16%	0	13.44	26.88	40.32	53.76	17.92
	July 12th	16%	0	13.44	26.88	40.32	53.76	17.92
	July 20th	16%	0	13.44	26.88	40.32	53.76	17.92
	July 28th	12%	0	10.08	20.16	30.24	40.32	13.44
2023	June 13th	5%	0	4.20	8.40	12.60	16.80	5.60
	June 19th	7%	0	5.88	11.76	17.64	23.52	7.84
	June 29th	10%	0	8.40	16.80	25.20	33.60	11.20
	July 8th	15%	0	12.60	25.20	37.80	50.40	16.80
	July 22nd	18%	0	15.12	30.24	45.36	60.48	20.16
	July 30th	18%	0	15.12	30.24	45.36	60.48	20.16
	August 7th	15%	0	12.60	25.20	37.80	50.40	16.80
	August 19th	12%	0	10.08	20.16	30.24	40.32	13.44

UAV data acquisition

A DJI Phantom 4 Advanced UAV (China, DJI) was used to obtain RGB images of the experimental sites at the time of sampling. The flight time for image acquisition was selected under clear, windless, and cloudless weather conditions from 13:00 to 14:00. The UAV platform was equipped with a 20-megapixel image sensor, the lens was set straight down, the camera shutter was set at 1/240 s, and the shooting distance was 20 m above the ground. The front and side overlap properties of the image were set to 75%, and the size of the obtained JPG image was 5472 × 3648 pixels. The obtained original image was imported into DJI Intelligent Graph, and the reflectivity board photo was imported into the radiation correction interface, as well as the corresponding reflectivity. The rebuild function was used to generate an orthophoto, which was exported as a TIFF image.

NNI calculation

Lemaire *et al*. (2008) indicated that the minimum above-ground nitrogen concentration of a plant that maximizes the above-ground dry matter accumulation of cotton is the critical nitrogen concentration. The critical nitrogen concentration dilution curve was modeled as follows:

$$N_c = aW_{max}^{-b} \quad (1)$$

where N_c is the critical nitrogen concentration (g [100 g]$^{-1}$), a is the critical nitrogen concentration of the cotton plant when the aboveground biomass of cotton is 1 t hm^{-2}, W_{max} is the maximum aboveground biomass of cotton, and b is the parameter controlling the slope of the dilution curve of the critical nitrogen concentration.

The NNI refers to the ratio of the actual aboveground nitrogen concentration of crops to the critical nitrogen concentration and is used to determine the nitrogen nutrition status of the crops. The formula for calculating the NNI is as follows:

$$NNI = N_i/N_c \quad (2)$$

where NNI is the nitrogen nutrient index, N_i is the measured nitrogen concentration in the aboveground crop in g (100 g)$^{-1}$, and N_c is the critical nitrogen concentration in the aboveground crop in g (100 g)$^{-1}$. If NNI is greater than 1, it indicates excessive nitrogen application; if NNI is less than 1, it indicates nitrogen deficiency; and if NNI is equal to 1, it indicates the optimal level of nitrogen nutrition (Fabbri et al., 2020).

UAV image processing

Using ENVI software, the region of interest of the UAV RGB image was sketched, and the average values of the red, green, and blue channels of each region of interest were extracted and defined as R, G, and B, respectively. Following normalization, R, G, and B were defined as r, g, and b, which were calculated as r = R/(R + G + B), g = G/(R + G + B), and b = G/(R + G + B). Eleven vegetation indices with good correlation with crop nitrogen nutrient status were selected based on previous studies: green leaf index (GLI), green red vegetation index (GRVI), modified green red vegetation index (MGRVI), excess red vegetation index (ExR), excess blue vegetation index (ExB), excess green vegetation index (ExG), visible atmospherically resistant index (VARI), red-green-blue vegetation index (RGBVI), red green ratio index (RGRI), color index of vegetation (CIVE), and normalized green-red difference index (NGRDI) (Table 2).

Table 2. Vegetation indices used in this study

Vegetation index	Formulation	References
GLI	(2*g-r + b)/(2*g + r + b)	Louhaichi *et al.* (2001)
GRVI	(g-r)/(g + r)	Tucker (1979)
MGRVI	$(g^2-r^2)/(g^2 + r^2)$	Bendig *et al.* (2015)
ExR	1.4*r-g	Meyer and Neto (2008)
ExB	1.4b-g	Qi *et al.* (1994)
ExG	2*g-r-b	Woebbecke *et al.* (1995)
VARI	(g-r)/(g + r-b)	Saberioon and Gholizadeh (2016)
RGBVI	$(g^2-b*r^2)/(g^2 +b*r^2)$	Possoch *et al.* (2016)
RGRI	r/g	Verrelst *et al.* (2008)
CIVE	0.441r−0.811g+0.385b+18.78745	Kataoka *et al.* (2003)
NGRDI	(g − r)/(g + r)	Li *et al.* (2010)

Abbreviations: CIVE, color index of vegetation; ExB, excess blue index; ExG, excess green index; ExR, excess red index; GLI, green leaf index; GRVI, green red vegetation index; MGRVI, modified green red vegetation index; NGRDI, normalized green-red difference index; RGBVI, red-green-blue vegetation index; RGRI, red green ratio index; VARI, visible atmospherically resistant index.

Eight texture features in the region of interest of the RGB image were extracted using ENVI software with a grayscale co-occurrence matrix (GLCM). The gray level-direction co-production matrix were obtained by averaging four values through the four generating directions of θ. The average value was recorded as the feature value. The eight texture features were the mean (Mean), variance (Var), homogeneity (Hom), contrast (Con), dissimilarity (Dis), entropy (Ent), second moment (SM), and correlation (Cor). The terms "_R", "_G", and "_B" were added to texture names to indicate textures based on different GLCM bands (for example, "Mean_R" indicates the mean of the R-band).

The color moments do not need to quantize the color space, and the feature vectors have low dimensions and are mainly concentrated in the low-order moments; therefore, color information can only be expressed by the first-, second-, and third-order moments of the color. With the help of the nine color features of the low-order color moments of the hue (H)-saturation (S)-value (V) color space to express the color information of

the image, the nine components (three color components and two low-order moments on each component) can cover the color moments of the image, no other color features are complex. The formulae for the color moments are as follows (Strickerand Dimai , 1996):

$$\mu_i = \frac{1}{N}\sum_{j=1}^{N} p_{i,j} \quad (3)$$

$$\sigma_i = \left(\frac{1}{N}\sum_{j=1}^{N}(p_{i,j} - \mu_i)^2\right)^{\frac{1}{2}} \quad (4)$$

$$s_i = \left(\frac{1}{N}\sum_{j=1}^{N}(p_{i,j} - \mu_i)^3\right)^{\frac{1}{3}} \quad (5)$$

where $p_{i,j}$ denotes the probability of a pixel with a gray value of j in the i color channel; N is the total number of pixels in the region of interest of each map; μ_i $(1 \le I \le 3)$ denotes the average color in each color channel; and σ_i and s_i denote the variance and skewness of each color channel, respectively.

Modeling method

In this paper, the four machine learning algorithms used for model construction were PLSR, SVM, AB, and RF (Scornet, 2016).

(1) Partial Least Squares Regression (PLSR) combines Principal Component Analysis and Typical Correlation Analysis (Jin et al., 2019) to simplify the characteristics of the independent variables, which can be calculated by regressing the independent variables with multicollinearity. The objective solution is calculated by minimizing the sum of squares of errors to find the best function with this set of data, which makes it easier to identify the system information and noise compared to traditional multiple linear regression models.

(2) Support Vector Machines (SVM) is a powerful and versatile supervised machine learning algorithm that can be used for linear and nonlinear classification, regression, and even anomaly detection. Support Vector Regression (SVR) is proposed in the application of regression problems, and the kernel function linear in SVR is chosen in this study to deal with the nonlinear regression phenomenon of samples.

(3) Adaptive Boosting (AB) algorithm does not screen features, there is no overfitting phenomenon in the training process, and the upper limit of the training error rate decreases gradually with the increase of the number of iterations. In this study, Ridge Regression (RR) is chosen as the base learner of the Adaptive Boosting algorithm, and then multiple weak classifiers are trained based on the same training set, and then multiple weak classifiers are combined into a strong classifier, and the final prediction result is determined by the number of votes.

(4) Random Forest (RF) is a method that scores the importance of each variable and evaluates its role in classification while classifying data, integrates a large number of decision trees together, and utilizes multiple decision trees to train, classify, and predict sample data. In the construction of each decision tree, 2/3 of the training samples are used to train the model and 1/3 are called out-of-bag samples, which are determined in predicting the results by averaging all the trees. The random forest algorithm has been considered as one of the most accurate classification and regression prediction methods. In order to improve the operational efficiency and estimation accuracy, the number of decision trees and the input variables per node can be adjusted.

Model evaluation

In this study, the dataset was divided into modeling and testing sets at a ratio of 7:3, and the coefficient of determination (R^2), root mean square error (RMSE), and mean absolute error (MAE) were used as criteria to evaluate the model performance. When the value of R^2 was closer to 1, the value of RMSE was closer to 0, and the value of MAE was smaller, the model performance was better. These values were calculated as follows:

$$R^2 = \frac{\sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \quad (6)$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2} \quad (7)$$

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}|\overline{y_i} - y_i| \quad (8)$$

where y_i represents the measured value of sample i, $\bar{y}_i$ represents the predicted value of the sample, $\bar{y}$ represents the average value of all the samples, and n is the number of samples.

Results

Construction of critical nitrogen concentration dilution model for cotton

Effects of different nitrogen application levels on aboveground biomass and nitrogen concentration of cotton

As shown in Tables 3 and 4, the aboveground biomass of cotton increased with the advancement of fertility and the increase of N application, and the aboveground biomass of cotton with N application of N3 accumulated the most; the differences between N0 and N1 and N3 were significant, while the differences between N2 and N4 and N3 were not significant.

Table 3. Significance test of cotton aboveground biomass under different nitrogen application levels (2022)

Nitrogen rate	Post emergence time (d)					
	57	62	69	81	91	101
N0	1.95^{d}	3.49^{b}	4.98^{b}	5.86^{c}	6.96^{b}	10.59^{c}
N1	2.29^{c}	3.61^{b}	5.45^{ab}	7.07^{b}	9.20^{a}	11.21^{c}
N2	2.45^{b}	3.74^{b}	6.12^{a}	7.57^{ab}	9.52^{a}	11.44^{c}
N3	2.93^{a}	4.54^{a}	6.13^{a}	7.89^{ab}	10.00^{a}	14.95^{b}
N4	3.00^{a}	4.80^{a}	6.18^{a}	8.00^{a}	10.01^{a}	16.10^{a}

Nitrogen application levels: N0 (0 kg hm^{-2}), N1 (120 kg hm^{-2}), N2 (240 kg hm^{-2}), N3 (360 kg hm^{-2}), and N4 (480 kg hm^{-2}).

Differences in letters between nitrogen application treatments (Duncan test, $p<0.05$).

Table 4. Significance test of cotton aboveground biomass under different nitrogen application levels (2023)

Nitrogen rate	Post emergence time (d)						
	57	67	76	90	98	106	118
N0	1.92^{b}	3.03^{b}	4.67^{c}	7.48^{b}	8.82^{b}	10.02^{b}	12.15^{b}
N1	2.03^{ab}	3.17^{b}	5.14^{c}	9.05^{a}	10.03^{b}	11.89^{b}	12.51^{b}
N2	2.04^{ab}	3.69^{a}	6.17^{b}	9.05^{a}	10.29^{a}	12.76^{a}	15.62^{ab}
N3	2.22^{a}	3.76^{a}	6.69^{ab}	9.28^{a}	10.37^{a}	12.39^{a}	16.90^{a}
N4	2.27^{a}	3.85^{a}	6.78^{a}	9.52^{a}	10.86^{a}	13.00^{a}	16.12^{a}

Nitrogen application levels: N0 (0 kg hm^{-2}), N1 (120 kg hm^{-2}), N2 (240 kg hm^{-2}), N3 (360 kg hm^{-2}), and N4 (480 kg hm^{-2}).

Differences in letters between nitrogen application treatments (Duncan test, $p<0.05$).

As shown in Figure 2, N concentration in cotton showed a decreasing trend during the reproductive process. The differences between N0 and N1 and N3 were significant, while the differences between N2 and N4 and N3 were not significant, indicating that these three N applications did not have a significant effect on the growth and development of cotton.

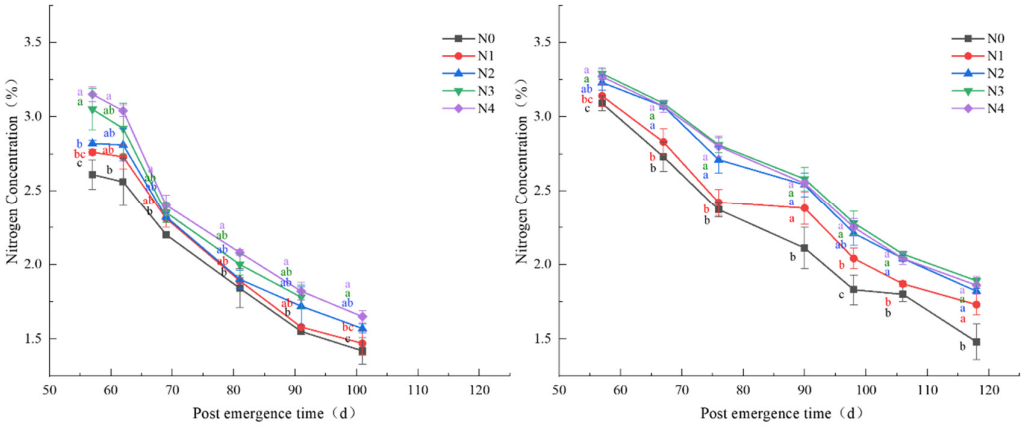

Figure 2. Trends and significance analysis of nitrogen concentration in cotton at different nitrogen application levels (Left: 2022, Right: 2023)
Nitrogen application levels= N0 (0 kg hm^{-2}), N1 (120 kg hm^{-2}), N2 (240 kg hm^{-2}), N3 (360 kg hm^{-2}), and N4 (480 kg hm^{-2})

Construction and verification of critical nitrogen concentration dilution model

During the sampling period, the aboveground biomass of cotton ranged from 1.92 to 16.9 t hm^{-2}. Combined with the critical nitrogen concentration dilution modeling method proposed by Justes et al. (1994), the aboveground biomass of cotton under different nitrogen treatments increased significantly and non-significantly, and nitrogen treatments with significant increases in aboveground biomass were set as nitrogen-limited levels, while those with significant increases in aboveground biomass were set as non-nitrogen-limited levels, which resulted in the critical nitrogen concentration dilution model of cotton as $N_c = 3.96W_{max}^{-0.19}$ (Figure 3), with $R^2 = 0.93$.

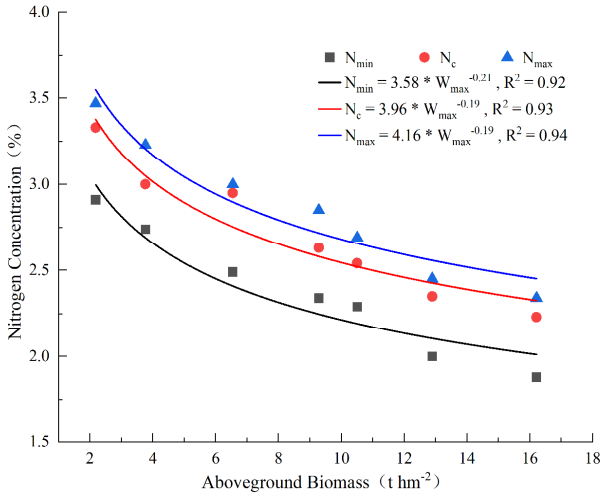

Figure 3. Model of cotton nitrogen concentration dilution curve
N_c, critical nitrogen concentration (g [100 g]$^{-1}$); R^2, coefficient of determination; W_{max}, maximum aboveground biomass of cotton

The data points from 2022 were selected to validate the model, analyze the simulated and measured values, and plot 1:1 histograms (Figure 4). The accuracy of the model was assessed by calculating R^2, RMSE,

and MAE. As shown in Figure 4, R^2 = 0.89, RMSE = 0.50 g (100 g)$^{-1}$, and MAE = 0.44, indicating that the model performance was good and could be used for diagnosing nitrogen nutrition in cotton.

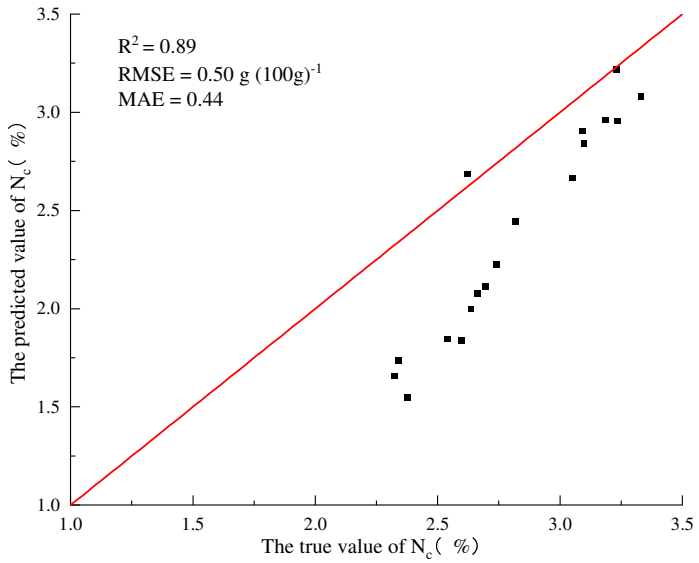

Figure 4. Model validation of the dilution curve for the critical nitrogen concentration in cotton
MAE, mean absolute error; N_c, critical nitrogen concentration (g [100 g]$^{-1}$); R^2, coefficient of determination; RMSE, root mean square error

In this study, NNI was calculated using equation (2) and was used to determine the nitrogen nutritional status of cotton. Table 5 shows the results of descriptive statistical analysis of NNI of cotton in 2022, and Table 6 shows the results of descriptive statistical analysis of NNI of cotton in 2023. The NNI values of cotton in 2022 and 2023 were inferred to be less than 1 at all nitrogen application levels after the full bloom stage, possibly because after entering the boll stage, the synchronization of cotton nutrient growth and reproductive growth, which resulted in a substantial increase in nitrogen demand, and the late stage of fertility, which did not show nitrogen overnutrition because the nitrogen level was not set high enough. The absence of nitrogen overnutrition was because the nitrogen level was not sufficiently high.

Table 5. Statistical analysis of cotton nitrogen nutrient index in 2022

Date (d)	Min	Max	Mean	SD	Variance	Skewness
57	0.710	1.110	0.888	0.073	0.005	0.367
62	0.770	1.190	0.923	0.090	0.008	0.504
69	0.660	1.070	0.824	0.078	0.006	0.734
81	0.520	0.860	0.710	0.069	0.005	-0.412
91	0.540	0.770	0.661	0.058	0.003	0.152
101	0.410	0.80	0.599	0.077	0.006	0.118

SD, standard deviation.

Table 6. Statistical analysis of cotton nitrogen nutrient index in 2023

Date (d)	Min	Max	Mean	SD	Variance	Skewness
57	0.850	1.070	0.957	0.047	0.002	-0.129
67	0.830	1.080	0.960	0.062	0.004	-0.132
76	0.770	1.070	0.932	0.068	0.005	-0.267
90	0.740	1.050	0.910	0.081	0.006	-0.647
98	0.590	0.980	0.831	0.077	0.006	-1.082
106	0.590	0.890	0.769	0.067	0.004	-0.757
118	0.500	0.840	0.717	0.073	0.005	-1.002

SD, standard deviation.

Cotton nitrogen nutrition diagnostic model construction

Cotton nitrogen nutrient diagnosis model based on vegetation index

As shown in Figure 5, 11 vegetation indices extracted from the UAV images were significantly correlated with the NNI during the full fertility period, among which GLI had the best correlation with the NNI (r = 0.56). As shown in Table 7, vegetation indices that were significantly correlated with the NNI and had the top five importance scores were selected using RFA and RF algorithms, among which GLI, ExB, and ExR appeared in the screening results of the two algorithms.

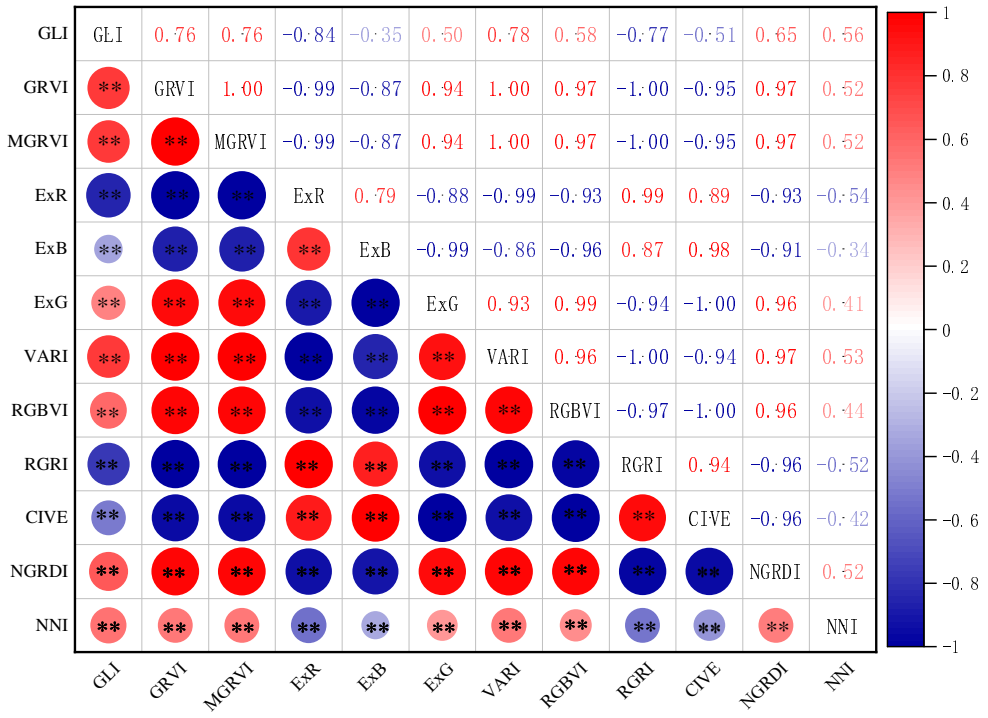

Figure 5. Correlation between vegetation indices and NNI

CIVE, color index of vegetation; ExB, excess blue vegetation index; ExG, excess green vegetation index; ExR, excess red vegetation index; GLI, green leaf index; GRVI, green red vegetation index; MGRVI, modified green red vegetation index; NGRDI, normalized green-red difference index; RGBVI, red-green-blue vegetation index; RGRI, red green ratio index; VARI, visible atmospherically resistant index.

Table 7. Importance of the top five vegetation indices and estimated nitrogen nutrient index for RFA and RF screening

Model	RFA		RF	
	Characteristic	Importance	Characteristic	Importance
Vegetation index	CIVE	0.9697	GLI	0.5267
	ExB	0.9418	ExB	0.1273
	RGRI	0.7243	ExR	0.0816
	GLI	0.1509	ExG	0.0540
	ExR	0.1127	VARI	0.0507

CIVE, color index of vegetation; ExB, excess blue vegetation index; ExG, excess green vegetation index; ExR, excess red vegetation index; GLI, green leaf index; RFA: random frog algorithm; RF, random forest; RGRI, red green ratio index; VARI, visible atmospherically resistant index.

A cotton NNI estimation model was constructed using four machine learning algorithms based on the screened image feature sets for diagnosing nitrogen nutrition in cotton. As shown in Figure 6, the optimal model based on the RFA screening algorithm was RF (modeling set: $R^2 = 0.95$, RMSE = 0.03, MAE = 0.02; test set: $R^2 = 0.70$, RMSE = 0.07, MAE = 0.05), and the optimal model based on the RF screening algorithm was RF (modeling set: $R^2 = 0.95$, RMSE = 0.03, MAE = 0.02; test set: $R^2 = 0.70$, RMSE = 0.07, MAE = 0.05). A comprehensive comparison of the eight-cotton nitrogen nutrient diagnostic models based on vegetation indices revealed that the optimal performance was achieved by the RF model constructed based on the RF screening algorithm.

Cotton nitrogen nutrient diagnosis model based on texture characteristics

As shown in Figure 7, the correlations between 24 texture feature parameters extracted from UAV images and the NNI during the whole fertility period were significantly correlated with NNI, except Ent_R, SM_R, Cor_R, Ent_G, SM_G, Cor_G, Mean_B, Hom_B, Ent_B, SM_B, and Cor_B. Among those that were significantly correlated, the correlations between Con_G and NNI were the best, with an r of -0.63. The top five texture feature parameters that correlated significantly with NNI were selected using the RFA and RF algorithms, among which two algorithms were selected based on the top five importance scores. The top five texture feature parameters with importance scores related to the significance of NNI were selected using the RFA and RF algorithms, and Dis_G appeared in the screening results of two of the algorithms (Table 8).

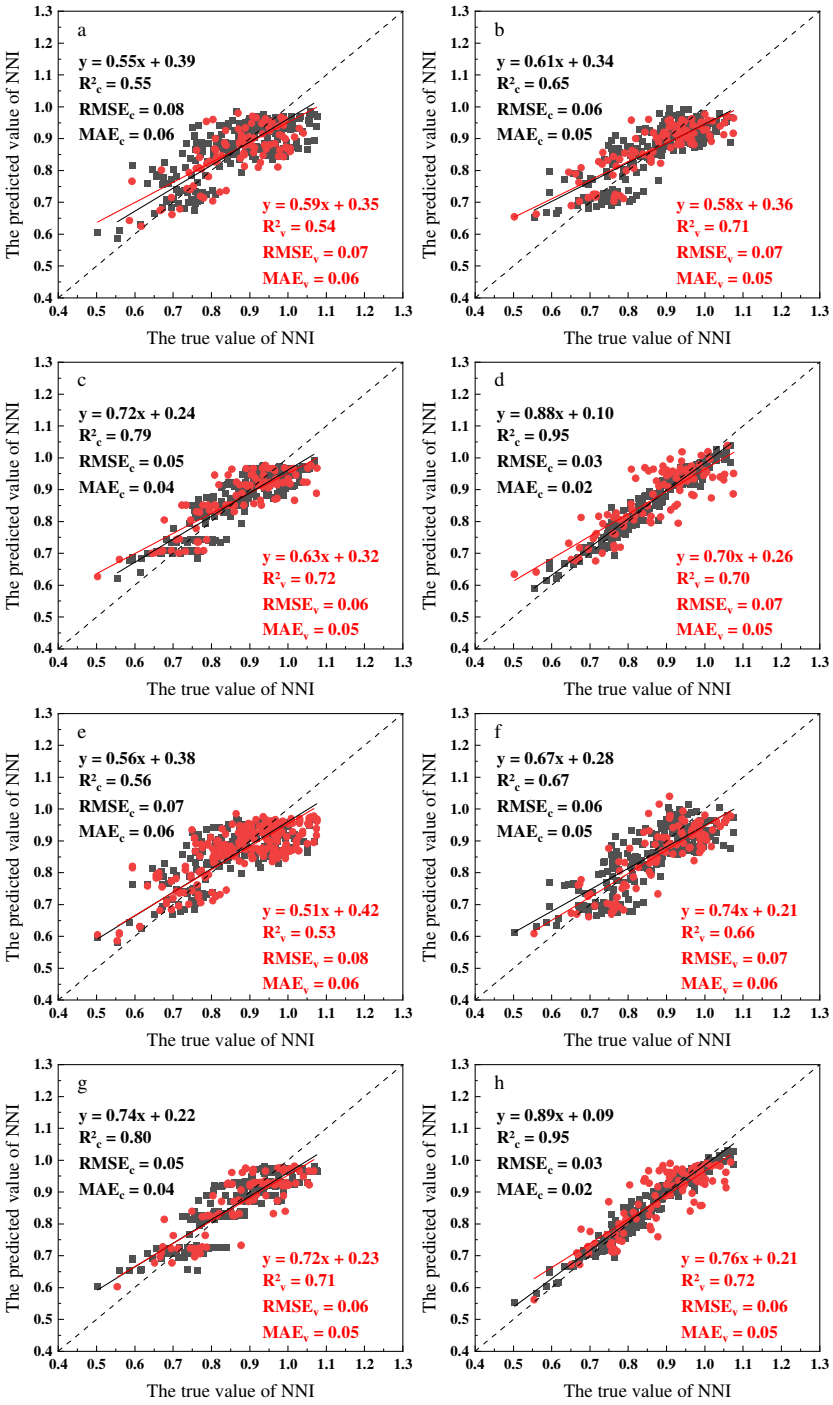

Figure 6. Estimation of the NNI based on vegetation index. RFA: (a) PLSR modeling; (b) SVM modeling; (c) AB modeling; and (d) RF modeling. RF: (e) PLSR modeling; (f) SVM modeling; (g) AB modeling; and (h) RF modeling

AB, adaptive boosting; MAE, mean absolute error; NNI, nitrogen nutrient index; PLSR, partial least squares regression; R^2, coefficient of determination; RFA: random frog algorithm; RF, random forest; RMSE, root mean square error; SVM, support vector machine

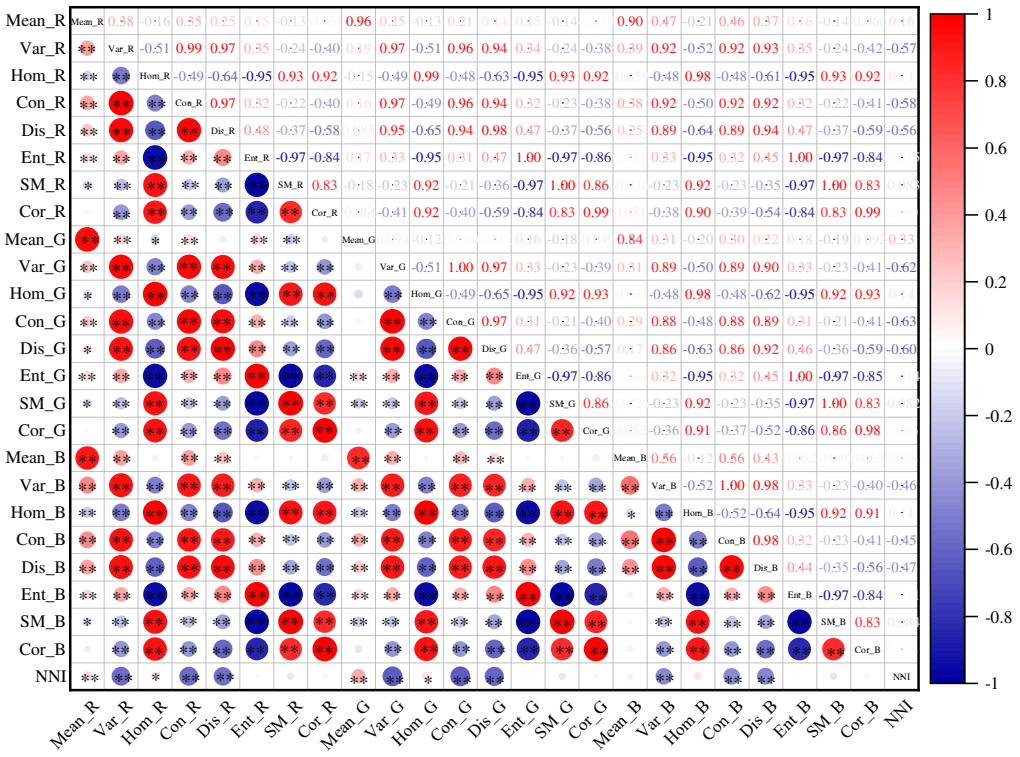

Figure 7. Correlation of texture features with nitrogen nutrient index

B, blue; Con, contrast; Cor, correlation; Dis, dissimilarity; Ent, entropy; G, green; Hom, homogeneity; R, red; SM, second moment; Var, variance

Table 8. Top 5 texture features and importance of estimating nitrogen nutrient index for RFA and RF screening

Model	RFA		RF	
	Characteristic	Importance	Characteristic	Importance
Texture features	Dis_G	0.9969	Mean_G	0.2648
	Hom_R	0.9927	Dis_G	0.2205
	Dis_R	0.9839	Con_G	0.1217
	Hom_G	0.9794	Var_G	0.1124
	Dis_B	0.9769	Mean_R	0.0532

RFA: random frog algorithm; RF, random forest; B, blue; Con, contrast; Dis, dissimilarity; G, green; Hom, homogeneity; R, red; Var, variance.

A cotton NNI estimation model was constructed using four machine learning algorithms based on the screened texture features for the diagnosis of cotton nitrogen nutrition. As shown in Figure 8, the optimal model based on the RFA screening algorithm was RF (modeling set: R^2 = 0.94, RMSE = 0.04, MAE = 0.02; test set: R^2 = 0.56, RMSE = 0.08, MAE = 0.06), and the optimal model based on the RF screening algorithm was RF (modeling set: R^2 = 0.95, RMSE = 0.03, MAE = 0.02; test set: R^2 = 0.63, RMSE = 0.07, MAE = 0.05). A comprehensive comparison of the eight-texture feature-based nitrogen nutrition diagnostic models for cotton revealed that the optimal performance was achieved by the RF model constructed based on the RF screening algorithm.

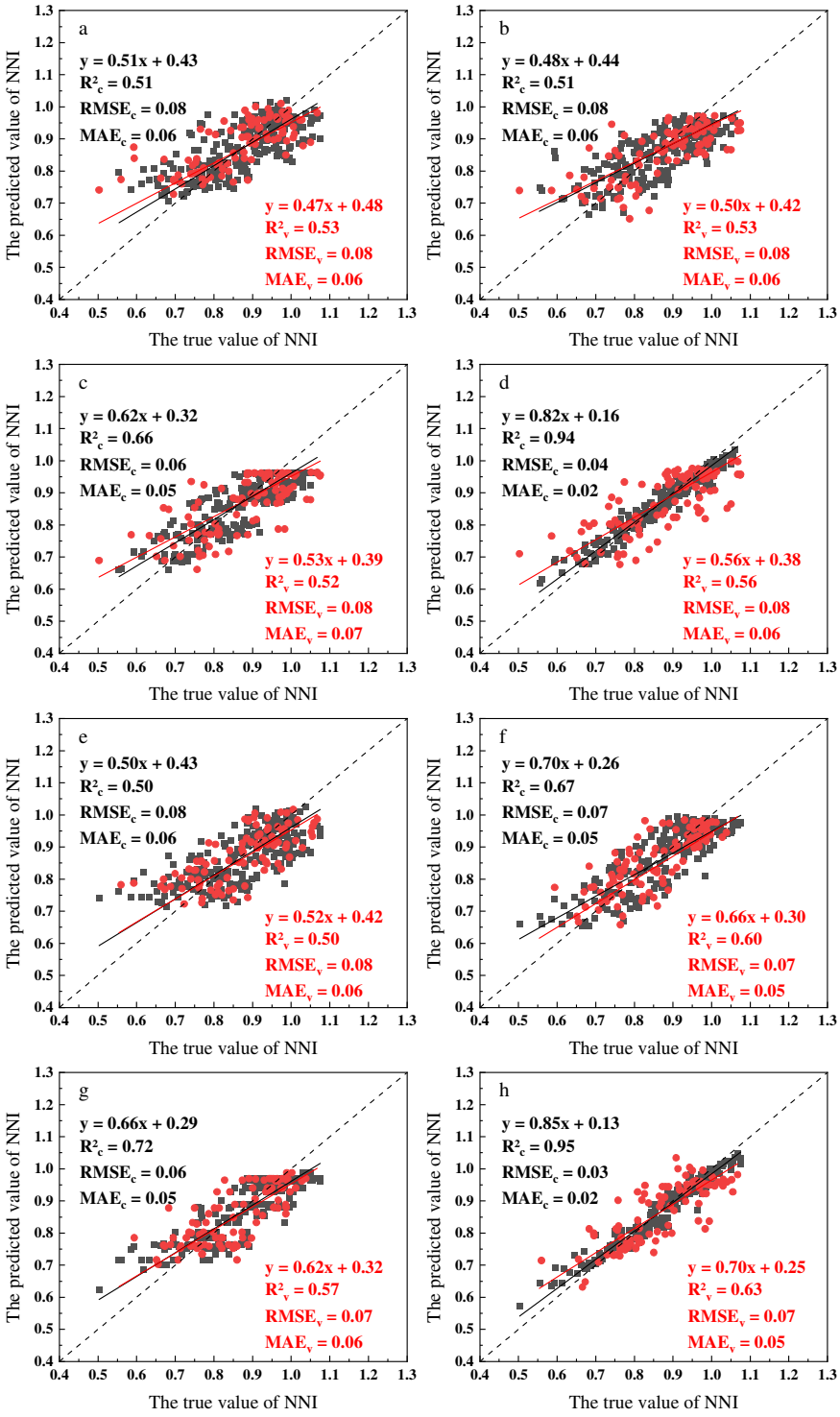

Figure 8. Estimating NNI based on texture features. RFA: (a) PLSR modeling; (b) SVM modeling; (c) AB modeling; and (d) RF modeling. RF: (e) PLSR modeling; (f) SVM modeling; (g) AB modeling; and (h) RF modeling

AB, adaptive boosting; MAE, mean absolute error; NNI, nitrogen nutrient index; PLSR, partial least squares regression; R^2, coefficient of determination; RFA: random frog algorithm; RF, random forest; RMSE, root mean square error; SVM, support vector machine

Cotton nitrogen nutrition diagnosis model based on color moment

The correlations between the nine color moments extracted from the UAV images and the NNI during the entire fertility period are shown in Figure 9, except for H_ske, V_var, and V_ske. All other color moment parameters were significantly correlated with the NNI; among them, H had the best correlation with the NNI, with an r of 0.69. Table 9 displays the top five characteristics selected using the RFA and RF algorithms based on the five importance scores related to the significance of the NNIs. The texture feature parameters were selected using the RFA and RF algorithms, among which H, H_var, S, and V were included in the screening results of the two algorithms.

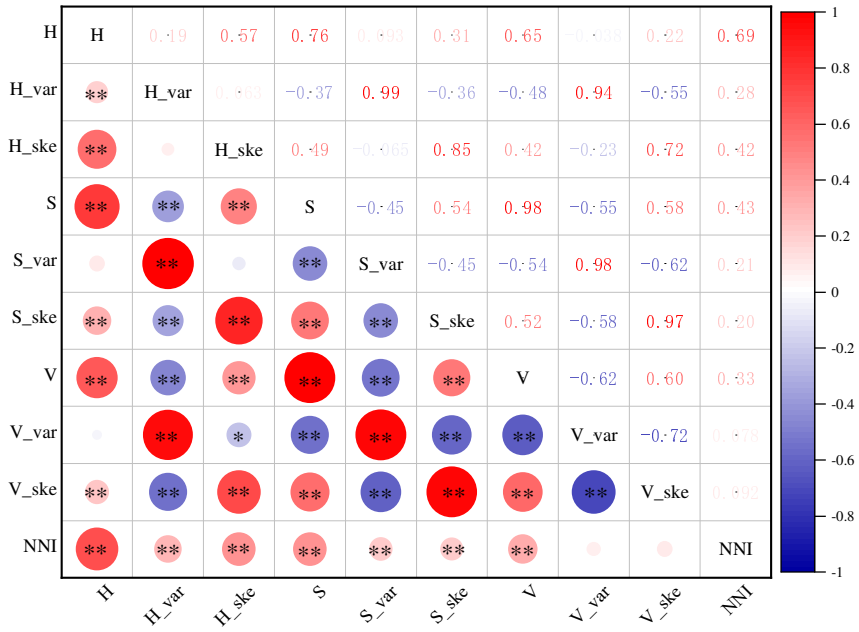

Figure 9. Correlation of color moments with nitrogen nutrient index
H, hue; NNI, nitrogen nutrient index; S, saturation; Ske, skewness; V, value; Var, variance.

Table 9. Importance of the first five color moments and estimated nitrogen nutrient index for RFA and RF screening

Model	RFA		RF	
	Characteristic	Importance	Characteristic	Importance
Color moment	H	0.9999	H	0.6437
	H_var	0.9961	S	0.1523
	S_var	0.8310	S_ske	0.0751
	S	0.1297	V	0.0508
	V	0.0905	H_var	0.0465

RFA: random frog algorithm; RF, random forest; H, hue; NNI, nitrogen nutrient index; S, saturation; Ske, skewness; V, value; Var, variance.

A cotton NNI estimation model was constructed using four machine learning algorithms based on screened color moments to diagnose nitrogen nutrition in cotton. As shown in Figure 10, the optimal model based on the RFA screening algorithm was RF (modeling set: R^2 = 0.97, RMSE = 0.02, MAE = 0.02; test set: R^2 = 0.75, RMSE = 0.06, MAE = 0.04), and the optimal model based on the RF screening algorithm was RF (modeling set: R^2 = 0.97, RMSE = 0.02, MAE = 0.02; test set: R^2 = 0.78, RMSE = 0.05, MAE = 0.04). A comprehensive comparison of the eight color-moment based nitrogen nutrition diagnosis models for cotton

revealed that the optimal performance was achieved with the RF model constructed based on the RF screening algorithm.

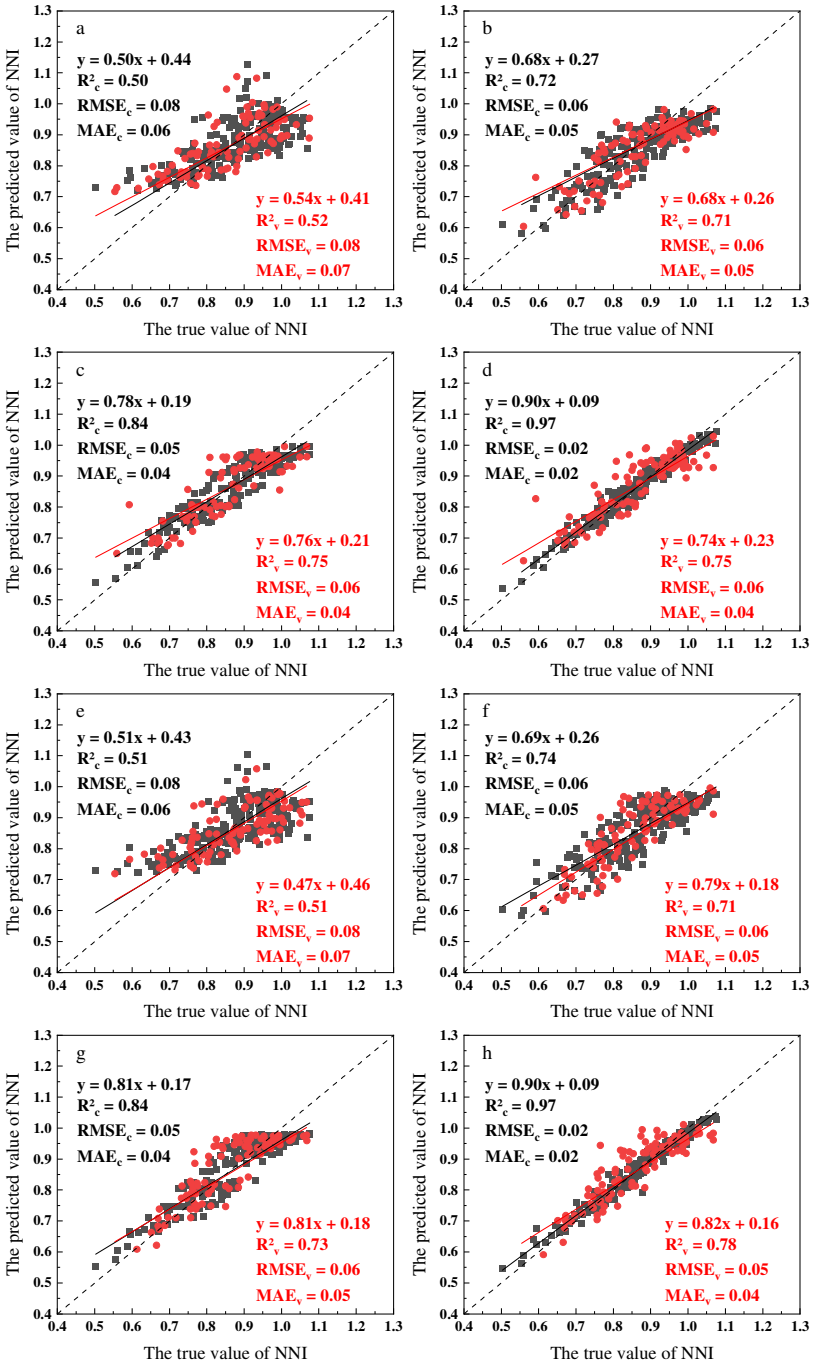

Figure 10. Estimation of the NNI based on color moments. RFA: (a) PLSR modeling; (b) SVM modeling; (c) AB modeling; and (d) RF modeling. RF: (e) PLSR modeling; (f) SVM modeling; (g) AB modeling; and (h) RF modeling

AB, adaptive boosting; MAE, mean absolute error; NNI, nitrogen nutrient index; PLSR, partial least squares regression; R^2, coefficient of determination; RFA: random frog algorithm; RF, random forest; RMSE, root mean square error; SVM, support vector machine

Cotton nitrogen nutrient diagnosis model based on feature fusion

Many studies have shown that the performance of a model built based on feature fusion is better than that of a model with a single class of features as the input (Fu *et al*., 2020). Therefore, in this study, the vegetation index, texture feature, and color moment, which are significantly related to the NNI, were taken as the image feature set, and the top five image feature parameters with importance scores in the image feature set were screened using the RFA and RF algorithms, among which the screening results of the two algorithms all contained Dis_G (Table 10).

Table 10. Top five image features, based on vegetation index, texture feature, and color moment, and importance of the estimated nitrogen nutrient index using RFA and RF screening

Model	RFA		RF	
	Characteristic	Importance	Characteristic	Importance
VIs+TFs+CMs	MGRVI	0.9288	H	0.5612
	VARI	0.8968	S	0.0704
	RGRI	0.8605	GLI	0.0481
	Dis_G	0.8185	ExR	0.0385
	RGBVI	0.7596	Dis_G	0.0314

Dis, dissimilarity; ExR, excess red vegetation index; GLI, green leaf index; G, green; H, hue; MGRVI, modified green red vegetation index; RFA: random frog algorithm; RF, random forest; RGBVI, red-green-blue vegetation index; RGRI, red green ratio index; S, saturation; VARI, visible atmospherically resistant index.

The cotton NNI estimation model was constructed using four machine learning algorithms based on screened image feature parameters for diagnosing nitrogen nutrition in cotton. As shown in Figure 11, the optimal model based on the RFA screening algorithm was RF (modeling set: R^2 = 0.96, RMSE = 0.02, MAE = 0.02; test set: R^2 = 0.81, RMSE = 0.05, MAE = 0.05), and the optimal model based on the RF screening algorithm was RF (modeling set: R^2 = 0.97, RMSE = 0.02, MAE = 0.02; test set: R^2 = 0.85, RMSE = 0.05, MAE = 0.04). The optimal model was the RF model based on the RF screening algorithm.

Comparative analysis of the aforementioned 32 cotton nitrogen nutrient diagnostic models revealed, as depicted in Table 11, the following: the performance of the model constructed with the feature set of RF screening was better than that constructed with the feature set of RFA screening; the performance of the model constructed based on image features was ranked: all features > color moments > vegetation index > texture features; the ranking of modeling algorithms was: RF > AB > SVM > PLSR; and the optimal nitrogen nutrient diagnostic model was the RF model constructed based on all image features of RF screening (modeling set: R^2 = 0.97, RMSE = 0.02, MAE = 0.02; test set: R^2 = 0.85, RMSE = 0.05, MAE = 0.04). Therefore, the image feature set derived from the screening of all image features using RF was considered the optimal image feature set.

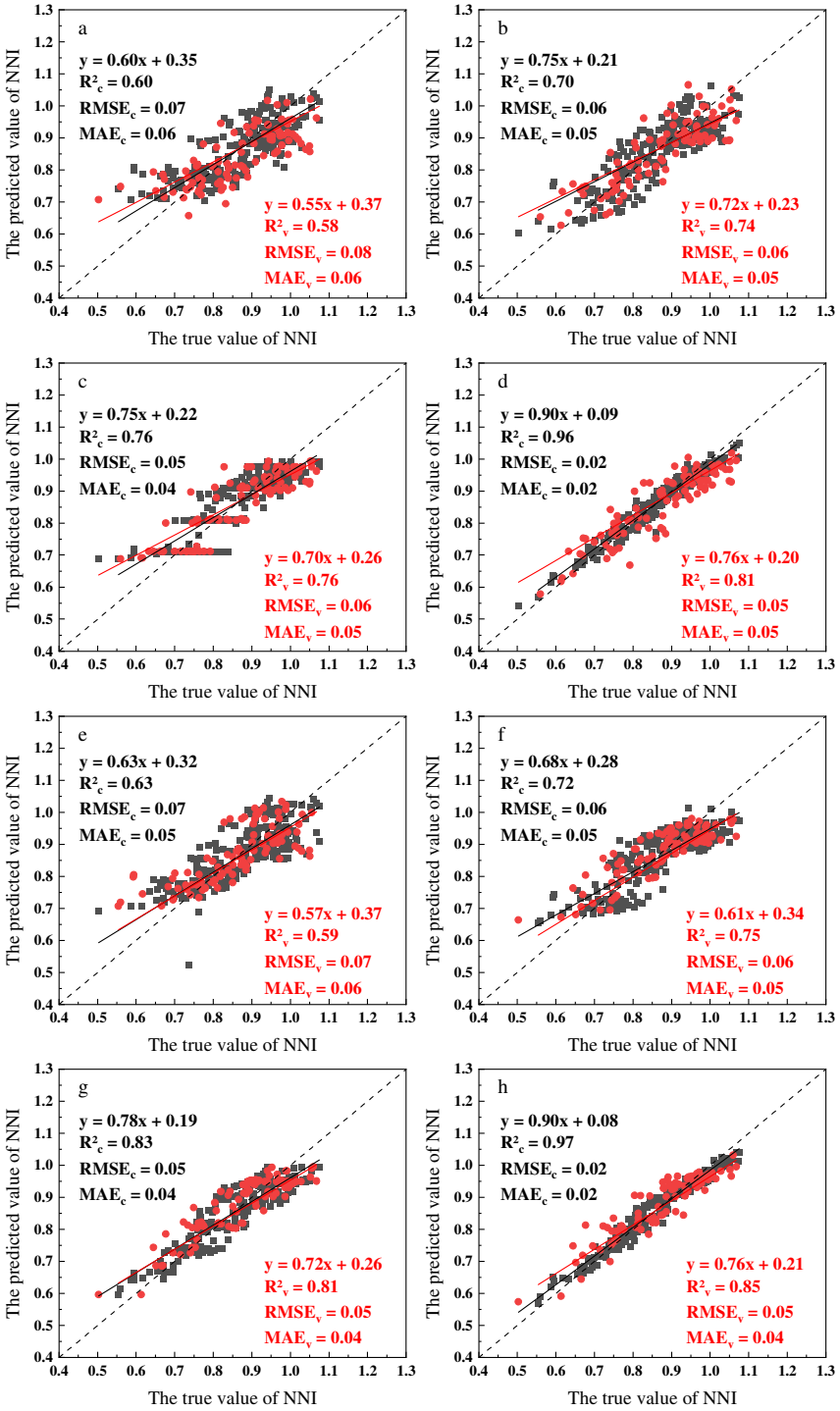

Figure 11. Estimating nitrogen nutrient index based on vegetation index, texture feature, and color moment. RFA: (a) PLSR modeling; (b) SVM modeling; (c) AB modeling; and (d) RF modeling. RF: (e) PLSR modeling; (f) SVM modeling; (g) AB modeling; and (h) RF modeling

AB, adaptive boosting; MAE, mean absolute error; NNI, nitrogen nutrient index; PLSR, partial least squares regression; R^2, coefficient of determination; RFA: random frog algorithm; RF, random forest; RMSE, root mean square error; SVM, support vector machine

Table 11. Indicators associated with the cotton nitrogen nutrient index estimation model

Image features	ML	RFA						RF					
		R^2_c	$RMSE_c$	MAE_c	R^2_v	$RMSE_v$	MAE_v	R^2_c	$RMSE_c$	MAE_c	R^2_v	$RMSE_v$	MAE_v
Vegetation Index	PLSR	0.55	0.08	0.06	0.54	0.07	0.06	0.56	0.07	0.06	0.53	0.08	0.06
	SVM	0.65	0.06	0.05	0.71	0.07	0.05	0.67	0.06	0.05	0.66	0.07	0.06
	AB	0.79	0.05	0.04	0.72	0.06	0.05	0.80	0.05	0.04	0.71	0.06	0.05
	RF	0.95	0.03	0.02	0.70	0.07	0.05	0.95	0.03	0.02	0.72	0.06	0.05
Texture Features	PLSR	0.51	0.08	0.06	0.53	0.08	0.06	0.50	0.08	0.06	0.50	0.08	0.06
	SVM	0.51	0.08	0.06	0.53	0.08	0.06	0.67	0.07	0.05	0.60	0.07	0.05
	AB	0.66	0.06	0.05	0.52	0.08	0.07	0.72	0.06	0.05	0.57	0.07	0.06
	RF	0.94	0.04	0.02	0.56	0.08	0.06	0.95	0.03	0.02	0.63	0.07	0.05
Color Moment	PLSR	0.50	0.08	0.06	0.52	0.08	0.07	0.51	0.08	0.06	0.51	0.08	0.07
	SVM	0.72	0.06	0.05	0.71	0.06	0.05	0.74	0.06	0.05	0.71	0.06	0.05
	AB	0.84	0.05	0.04	0.75	0.06	0.04	0.84	0.05	0.04	0.73	0.06	0.05
	RF	0.97	0.02	0.02	0.75	0.06	0.04	0.97	0.02	0.02	0.78	0.05	0.04
All Features	PLSR	0.60	0.07	0.06	0.58	0.08	0.06	0.63	0.07	0.05	0.59	0.07	0.06
	SVM	0.70	0.06	0.05	0.74	0.06	0.05	0.72	0.06	0.05	0.75	0.06	0.05
	AB	0.76	0.05	0.04	0.76	0.06	0.05	0.83	0.05	0.04	0.81	0.05	0.04
	RF	0.96	0.02	0.02	0.81	0.05	0.05	0.97	0.02	0.02	0.85	0.05	0.04

RFA: random frog algorithm; AB, adaptive boosting; MAE, mean absolute error; ML, machine learning; NNI, nitrogen nutrient index; PLSR, partial least squares regression; R^2, coefficient of determination; RF, random forest; RMSE, root mean square error; SVM, support vector machine.

Diagnostic study of nitrogen nutrition in cotton based on comprehensive indicators of UAV GRB images

Calculate the weights corresponding to the optimal image feature set using the coefficient of variation method and entropy method, respectively (Table 12). The normalized image feature data were multiplied by the corresponding weights to obtain the composite scores of each image feature and then summed to obtain the total score (F), which was used as the comprehensive diagnostic index. Record the comprehensive scores calculated by entropy method and coefficient of variation method as F1 and F2, respectively, and perform polynomial fitting with NNI (Figure 12), and the correlation between the composite index and NNI was obtained with $R^2 = 0.66$, which indicated that the model was reliable. According to the model formula, the values of F2 were 0.48 and 0.67 when the value of NNI was 1. Therefore, when $0.48 < F2 < 0.67$, too much nitrogen was applied to the cotton, whereas when $F2 < 0.48$ or $F2 > 0.67$, the cotton was deficient in nitrogen nutrition, and fertilizer application should be increased.

Table 12. Weights of each image metric

Characteristic parameter		GLI	ExR	Dis_G	H	S
Weight	Entropy method	0.01	0.02	0.13	0.20	0.64
	Coefficient of variation method	0.05	0.07	0.18	0.22	0.49

Dis, dissimilarity; ExR, excess red vegetation index; G, green; GLI, green leaf index; H, hue; S, saturation.

After calculating the weights of the five image features, it can be obtained that GLI and ExR accounted for a relatively small proportion (Table 12), and the weights were recalculated after removing GLI and ExR (Table 13), and then the weights were assigned to calculate the total score as a composite index for polynomial fitting with the NNI (Figure 13), and the model constructed by the entropy method had unchanged R^2, and the model constructed by the coefficient of variation method had reduced R^2 after the removal of GLI and ExR.

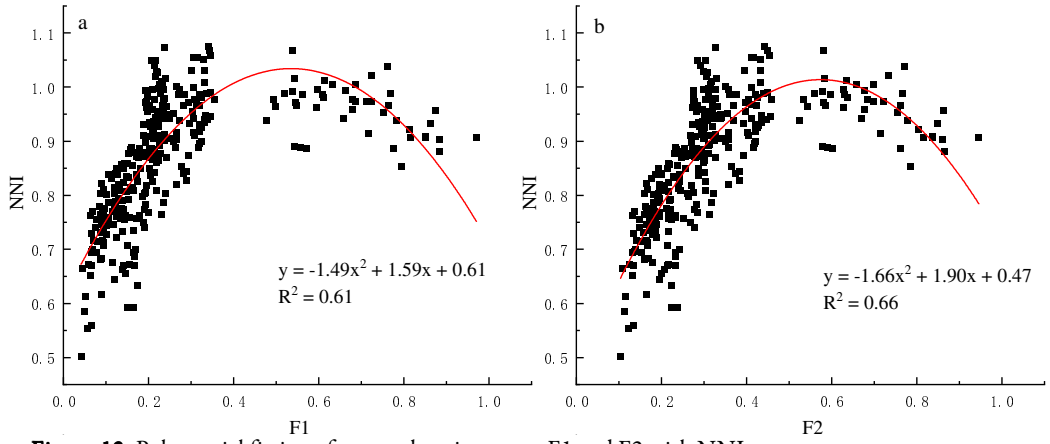

Figure 12. Polynomial fitting of comprehensive scores F1 and F2 with NNI
NNI, nitrogen, nutrient index; R^2, coefficient of determination

Table 13. Weights of each image metric

Characteristic parameter		Dis_G	H	S
Weight	Entropy method	0.13	0.21	0.66
	Coefficient of variation method	0.20	0.25	0.55

GLI, ground level image index; H, hue; S, saturation.

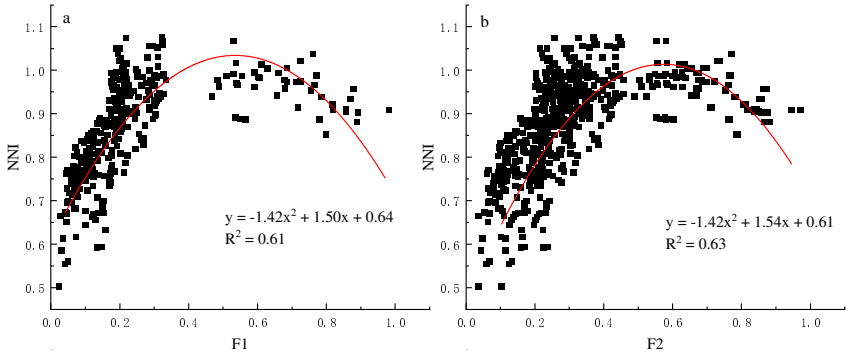

Figure 13. Polynomial fitting of comprehensive scores F1 and F2 with NNI
NNI, nitrogen, nutrient index; R^2, coefficient of determination

Discussion

Nitrogen nutrition diagnostic study for crops based on the NNI

Studies have shown that the NNI calculated with the help of the critical nitrogen concentration dilution model can accurately quantify the nitrogen nutrition level of crops and can be used for the diagnosis of crop nitrogen nutrition (Cheng *et al.*, 2022; Pei *et al.*, 2023), as the NNI value increases with an increase in nitrogen application. This study showed that the overall NNI value decreased with the advancement of cotton growth and development processes, indicating that cotton requires more nitrogen for growth and development in the late reproductive stage, which is consistent with the results of Hou *et al.* (2021). Currently, studies have applied remote sensing technology to estimate crop NNI. For example, Liu *et al.* (2020) constructed a model to estimate the NNI of winter wheat at the flowering stage based on UAV hyperspectral images (RMSE = 0.074). Hui *et al.* (2019) concluded that drone type and flight altitude had no effect on the estimation of the NNI, and

the R^2 values of the NNI estimation model were 0.47-0.86, indicating that drone multispectral images can be used to estimate crop NNI. Qiu *et al*. (2021) used six machine learning algorithms based on UAV RGB images to construct a model for estimating the NNI of rice at key fertility periods, and the results showed that the RF algorithm performed best at all fertility periods, with R^2 values ranging from 0.88 to 0.96 and RMSE values ranging from 0.03 to 0.07. Therefore, the use of UAV image features is feasible for estimating the NNI level of a crop.

Nitrogen nutrition diagnosis in cotton based on UAV image characterization

In this paper, the modeling comparison of four machine learning algorithms found that the NNI estimation model constructed by the RF algorithm was the most effective (modeling set: R^2 = 0.94~0.97, RMSE = 0.02~0.04; test set: R^2 = 0.56~0.85, RMSE = 0.05~0.08), likely because the machine learning modeling process is prone to the overfitting phenomenon, whereas the RF algorithm has stronger robustness and generalization ability (Zheng *et al*., 2018). Crop nitrogen nutrient status is closely related to canopy color, and texture features provide information related to structural features (Colombo *et al*., 2003). Many studies have estimated crop NNIs based on crop canopy spectral, structural, and color information (Maresma *et al*., 2016), but most studies that estimated NNIs based on image features considered only vegetation indexes and texture features; the potential of color moments for estimating the NNI was overlooked, despite the close relationship between the depth of color of crop leaves and nitrogen content (Haider *et al.,* 2021). This study showed that the superior performance ranking of the model constructed on the basis of image features was all features > color moments > vegetation index > texture features, demonstrating that color moments have greater potential in estimating the NNI. The reason for this phenomenon is that the UAV acquires image data 20 meters from the ground, and changes in color features are easier to obtain than changes in texture features, so the performance of vegetation index and color moments in estimating cotton nitrogen nutrient index is better than that of texture features, which is consistent with the results of the study by Pei (2023), the vegetation index is calculated based on the homogenization of the mean of the R, G, and B band, which loses most of the information about the color of the image, and there is a saturation phenomenon that leads to a reduction in estimation accuracy. And there is a saturation phenomenon leading to a reduction in estimation accuracy, while the color moments in H is the hue, S is the saturation, V is the brightness, if the image appears uneven lighting conditions, the impact is only V. The third-order moments can represent the color distribution of the image, and the color information of the acquired image is more complete and more stable (Yang and Sun, 2020). Yang *et al*. (2020) combined the image index and texture features of UAV digital images to construct a "map-spectrum" fusion index and used the partial least squares method to construct an NNI estimation model based on the image index, texture features, and fusion index, respectively. The results of their work showed that the fusion index constructed a model with the highest accuracy (R^2 = 0.6443) and that the fusion index could improve the accuracy of NNI inversion (Yang *et al*., 2020). In this study, the vegetation index, texture features, and color moments were used to estimate the crop NNI, and the results showed that the model fused with multiple data points performed better than the single data model, which was similar to the results of previous studies (Maimaitijiang *et al*., 2020).

Most studies have used evaluation models to assign weights to crop growth indicators to calculate comprehensive growth evaluation indexes, which also provide new ideas for crop nitrogen nutrition diagnosis (Hao *et al*., 2022; Xu *et al*., 2021). Most current studies on estimating the NNI in cotton use multiple image features as model inputs, and few image features are combined to construct a comprehensive index. Therefore, in this paper, using the coefficient of variation method to obtain the total composite score of the optimal image feature set as a composite index, we constructed a NNI estimation model based on the composite index, $y = -1.66x^2 + 1.90x + 0.47$, $R^2 = 0.66$, which is better than the modeling accuracy of the "map-spectrum" estimation model constructed using the partial least square regression algorithm by Yang *et al*. (2020) and has a simpler model structure. Finally, a diagnostic standard for cotton nitrogen nutrition based on the integrated index was

proposed to provide technical support for the rapid and nondestructive diagnosis of crop nitrogen nutrition. The diagnostic criteria were that when 0.48 < F2 < 0.67, cotton was over-applied with nitrogen; when F2 < 0.48 or F2 > 0.67, cotton was deficient in nitrogen.

This study had some limitations. The experimental data in this study came from two years of field experiments with the same cotton variety, and the accuracy of the estimation model was affected by a variety of factors, such as the cotton variety, planting area, and field management program. Therefore, in future studies, field experimental data from different areas should be accumulated over the years to improve the accuracy of the estimation or diagnostic models.

Conclusions

In this study, three feature parameters–vegetation index, texture feature, and color moment–were obtained based on UAV RGB images, and the optimal image feature set was determined using two feature-screening methods: RFA and RF. The NNI estimation model was constructed using four machine learning algorithms–PLSR, SVM, AB, and RF–to diagnose nitrogen nutrition. The optimal feature parameter set was determined by comparing the aforementioned models. The coefficient of variation method was used to calculate F as a comprehensive index, and polynomial fitting was performed using NNI. The nitrogen nutrition status of the cotton plants was determined using the F value.

Two essential implications were drawn from the findings of this study. First, the modeling effect of the image feature set screened by RF was better than RFA, and the performance of the model constructed based on image features can be ranked in terms of superiority: all features > color moments > vegetation index > texture features. Furthermore, the optimal image feature sets are H, S, GLI, ExR, and Dis_G, and the optimal NNI estimation model is the one constructed on the basis of the optimal image feature set using the RF algorithm (modeling set: R^2 = 0.97, RMSE = 0.02, MAE = 0.02; test set: R^2 = 0.85, RMSE = 0.05, MAE = 0.04). These findings demonstrate that the performance of the model constructed by combining multiple image features was superior. Second, the F of the optimal image feature set was calculated by the coefficient of variation method as a composite index, and polynomial fitting with NNI resulted in the model equation: $y = -1.66x^2 + 1.90x + 0.47$, R^2 = 0.66. Furthermore, the model was reliable. The values of F2 were 0.48 and 0.67 when the NNI value was 1. Therefore, when 0.48 < F2 < 0.67, cotton was over-applied with nitrogen, and when F2 < 0.48 or F2 > 0.67, cotton was nitrogen-deficient, indicating that the amount of fertilizer applied should be increased. This study demonstrates that the composite index F2 can effectively diagnose nitrogen nutrition, proposing a new diagnostic standard for crop nitrogen levels.

Authors' Contributions

Conceptualization: LW and QSY; Data curation: LW, QSY, FX and HYW; Funding acquisition; LLM, XL and ZZ; Investigation: LW, SZQ and MY; Supervision: LLM, XL and ZZ; Validation: LW, QSY, SZQ, HYW, FX and MY; Writing - original draft: LW and QSY; Writing - review and editing: LW, QSY, LLM, XL and ZZ. All authors read and approved the final manuscript.

Ethical approval (for researches involving animals or humans)

Not applicable.

Acknowledgements

This study received support from the Youth Science and Technology Innovation Talent Program (CZ006003), National Natural Science Foundation of China (42061058), Crop Intelligent Production Innovation Team (2023TD), Corps Leading Talent Program (2023YZ01), Corps Cotton Breeding AI Big Model Research (2023AA008), Shihezi University High level Talent Research Initiation Project (RCZK202342), and Autonomous Region "Tianchi Talent" Introduction Program (CZ006015).

Conflict of Interests

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

References

Bendig J, Yu K, Aasen H, Bolten A, Bennertz S, Broscheit J, ... Bareth G (2015). Combining UAV-based plant height from crop surface models, visible, and near infrared vegetation indices for biomass monitoring in barley. International Journal of Applied Earth Observation and Geoinformation 39:79-87. *https://doi.org/10.1016/j.jag.2015.02.012*

Cheng M, He J, Wang H, Fan J, Xiang Y, Liu X, ... Zhang F (2022). Establishing critical nitrogen dilution curves based on leaf area index and aboveground biomass for greenhouse cherry tomato: A Bayesian analysis. European Journal of Agronomy 141:126615. *https://doi.org/10.1016/j.eja.2022.126615*

Colombo R, Bellingeri D, Fasolini D, Marino C M (2003). Retrieval of leaf area index in different vegetation types using high resolution satellite data. Remote Sensing of Environment 86(1):120-131. *https://doi.org/10.1016/S0034-4257(03)00094-4*

Fabbri C, Mancini M, dalla Marta A, Orlandini S, Napoli M (2020). Integrating satellite data with a nitrogen nutrition curve for precision top-dress fertilization of durum wheat. European Journal of Agronomy 120:126148. *https://doi.org/10.1016/j.eja.2020.126148*

FAOSTAT (2023). Food and Agriculture Organization Corporate Statistical Database. Retrieved 2023 June 12 from *http://www.fao.org/faostat/*

Feng M, Chen Y, Duan W, Zhu Z, Wang C, Hu Y (2023). Water-energy-carbon emissions nexus analysis of crop production in the Tarim river basin, Northwest China. Journal of Cleaner Production 396:136566. *https://doi.org/10.1016/j.jclepro.2023.136566*

Fu Y, Yang G, Li Z, Song X, Li Z, Xu X, ... Zhao C (2020). Winter wheat nitrogen status estimation using UAV-based RGB imagery and Gaussian processes regression. Remote Sensing 12(22):3778. *https://doi.org/10.3390/rs12223778*

Haider T, Farid MS, Mahmood R, Ilyas A, Khan MH, Haider STA, ... Gul M (2021). A computer-vision-based approach for nitrogen content estimation in plant leaves. Agriculture 11(8):766. *https://doi.org/10.3390/agriculture11080766*

Han L, Yang G, Dai H, Xu B, Yang H, Feng H, ... Yang X (2019). Modeling maize above-ground biomass based on machine learning approaches using UAV remote-sensing data. Plant Methods 15(1):1-19. *https://doi.org/10.1186/s13007-019-0394-z*

Hao K, Fei L, Liu L, Jie F, Peng Y, Liu X, ... Wang X (2022). Comprehensive evaluation on the yield, quality, and water-nitrogen use efficiency of mountain apple under surge-root irrigation in the Loess Plateau based on the improved TOPSIS method. Frontiers in Plant Science 13:853546. *https://doi.org/10.3389/fpls.2022.853546*

Hou X, Xiang Y, Fan J, Zhang F, Hu W, Yan F, ... Li Z (2021). Evaluation of cotton N nutrition status based on critical N dilution curve, N uptake and residual under different drip fertigation regimes in Southern Xinjiang of China. Agricultural Water Management 256:107134. *https://doi.org/10.1016/j.agwat.2021.107134*

Hui Wang, Anders Krogh Mortensen, Peisheng Mao, Birte Boelt, René Gislum (2019). Estimating the nitrogen nutrition index in grass seed crops using a UAV-mounted multispectral camera. International Journal of Remote Sensing 40(7):2467-2482. *https://doi.org/10.1080/01431161.2019.1569783*

Jans Y, von Bloh W, Schaphoff S, Müller C (2021). Global cotton production under climate change–Implications for yield and water consumption. Hydrology and Earth System Sciences 25(4):2027-2044. *https://doi.org/10.5194/hess-25-2027-2021*

Jin J, Wang Q (2019). Evaluation of informative bands used in different PLS regressions for estimating leaf biochemical contents from hyperspectral reflectance. Remote Sensing 11(2):197. *https://doi.org/10.3390/rs11020197*

Jin Z, Li S, Yu F, Xu T (2024). Research on the rice fertiliser decision-making method based on UAV remote sensing data assimilation. Computers and Electronics in Agriculture 216:108508. *https://doi.org/10.1016/j.compag.2023.108508*

Justes E, Mary B, Meynard JM, Machet JM, Thelier-Huché L (1994). Determination of a critical nitrogen dilution curve for winter wheat crops. Annals of botany 74(4):397-407. *https://doi.org/10.1006/anbo.1994.1133*

Kataoka T, Kaneko T, Okamoto H, Hata S (2003). Crop growth estimation system using machine vision. In Proceedings 2003 IEEE/ASME international conference on advanced intelligent mechatronics. Kobe, Japan. IEEE pp 1079-1083.

Lemaire G, Jeuffroy M H, Gastal F (2008). Diagnosis tool for plant and crop N status in vegetative stage: Theory and practices for crop N management. European Journal of Agronomy 28(4):614-624. *https://doi.org/10.1016/j.eja.2008.01.005*

Li N, Li J, Tung S A, Shi X, Hao X, Shi F, ... Luo H (2022). Optimal irrigation amount can increase cotton lint yield by improving canopy structure and microenvironment under non-film deep drip irrigation. Journal of Cleaner Production 360:132156. *https://doi.org/10.1016/j.jclepro.2022.132156*

Li S, Yua F, Ata-UI-Karim S T, Zheng H, Cheng T, Liu X, ... Cao Q (2019). Combining color indices and textures of UAV-based digital imagery for rice LAI estimation. Remote Sensing 11(15):1763. *https://doi.org/10.3390/rs11151763*

Li Y, Chen D, Walker C N, Angus J F (2010). Estimating the nitrogen status of crops using a digital camera. Field Crops Research 118(3):221-227. *https://doi.org/10.1016/j.fcr.2010.05.011*

Liu H, Zhu H, Li Z, Yang G (2020). Quantitative analysis and hyperspectral remote sensing of the nitrogen nutrition index in winter wheat. International Journal of Remote Sensing 41(3):858-881. *https://doi.org/10.1080/01431161.2019.1650984*

Louhaichi M, Borman M M, Johnson D E (2001). Spatially located platform and aerial photography for documentation of grazing impacts on wheat. Geocarto International 16(1):65-70. *https://doi.org/10.1080/10106040108542184*

Maimaitijiang M, Sagan V, Sidike P, Hartling S, Esposito F, Fritschi F B (2020). Soybean yield prediction from UAV using multimodal data fusion and deep learning. Remote Sensing of Environment 237:111599. *https://doi.org/10.1016/j.rse.2019.111599*

Maresma Á, Ariza M, Martínez E, Lloveras J, Martínez-Casasnovas J A (2016). Analysis of vegetation indices to determine nitrogen application and yield prediction in maize (*Zea mays* L.) from a standard UAV service. Remote Sensing 8(12):973. *https://doi.org/10.3390/rs8120973*

Meyer G E, Neto J C (2008). Verification of color vegetation indices for automated crop imaging applications. Computers and Electronics in Agriculture 63(2):282-293. *https://doi.org/10.1016/j.compag.2008.03.009*

Pei Shengzhao (2023). Water and nitrogen monitoring and diagnosis of cotton under membrane drip irrigation in Southern Xinjiang Based on UAV Remote Sensing. MSc Dissertation, North West Agriculture and Forestry University.

Pei SZ, Zeng HL, Dai YL, Bai WQ, Fan JL (2023). Nitrogen nutrition diagnosis for cotton under mulched drip irrigation using unmanned aerial vehicle multispectral images. Journal of Integrative Agriculture 22(8):2536-2552. *https://doi.org/10.1016/j.jia.2023.02.027*

Possoch M, Bieker S, Hoffmeister D, Bolten A, Schellberg J, Bareth G (2016). Multi-temporal crop surface models combined with the RGB vegetation index from UAV-based images for forage monitoring in grassland. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 41:991-998. *https://doi.org/10.5194/isprsarchives-XLI-B1-991-2016*

Qi J, Chehbouni A, Huete A R, Kerr Y H, Sorooshian S (1994). A modified soil adjusted vegetation index. Remote sensing of environment 48(2):119-126. *https://doi.org/10.1016/0034-4257(94)90134-1*

Qiu Z, Ma F, Li Z, Xu X, Ge H, Du C (2021). Estimation of nitrogen nutrition index in rice from UAV RGB images coupled with machine learning algorithms. Computers and Electronics in Agriculture 189:106421. *https://doi.org/10.1016/j.compag.2021.106421*

Saberioon M M, Gholizadeh A (2016). Novel approach for estimating nitrogen content in paddy fields using low altitude remote sensing system. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 41:1011-1015. *https://doi.org/10.5194/isprs-archives-XLI-B1-1011-2016*

Scornet E (2016). Random forests and kernel methods. IEEE Transactions on Information Theory 62(3):1485-1500. *https://doi.org/10.1109/TIT.2016.2514489.*

Shi P, Wang Y, Xu J, Zhao Y, Yang B, Yuan Z, Sun Q (2021). Rice nitrogen nutrition estimation with RGB images and machine learning methods. Computers and Electronics in Agriculture 180:105860. *https://doi.org/10.1016/j.compag.2020.105860*

Soratto RP, Sandaña P, Fernandes FM, Fernandes AM, Makowski D, Ciampitti IA (2022). Establishing a critical nitrogen dilution curve for estimating nitrogen nutrition index of potato crop in tropical environments. Field Crops Research 286:108605. *https://doi.org/10.1016/j.fcr.2022.108605*

Stricker M A, Dimai A (1996). Color indexing with weak spatial constraints. In: Storage and Retrieval for Still Image and Video Databases IV. SPIE, 1996, 2670:29-40.

Sun-zhe YANG, Ai-zhen SUN (2023). Rice nitrogen nutrition diagnosis based on HSV color and LBP texture features. Computer and Modernization (07):86. *https://doi.org/10.3969/j.issn.1006-2475.2023.07.015*

Tian Y C, Gu K J, Chu X, Yao X, Cao W X, Zhu Y (2014). Comparison of different hyperspectral vegetation indices for canopy leaf nitrogen concentration estimation in rice. Plant and Soil 376:193-209. *https://doi.org/10.1007/s11104-013-1937-0*

Tsakmakis I D, Kokkos N P, Gikas G D, Pisinaras V, Hatzigiannakis E, Arampatzis G, Sylaios G K (2019). Evaluation of AquaCrop model simulations of cotton growth under deficit irrigation with an emphasis on root growth and water extraction patterns. Agricultural Water Management 213:419-432. *https://doi.org/10.1016/j.agwat.2018.10.029*

Tucker C J (1979). Red and photographic infrared linear combinations for monitoring vegetation. Remote sensing of Environment 8(2):127-150. *https://doi.org/10.1016/0034-4257(79)90013-0*

Verrelst J, Schaepman M E, Koetz B, Kneubühler M (2008). Angular sensitivity analysis of vegetation indices derived from CHRIS/PROBA data. Remote Sensing of Environment 112(5):2341-2353. *https://doi.org/10.1016/j.rse.2007.11.001*

Wang Y, Wang D, Zhang G, Wang J (2013). Estimating nitrogen status of rice using the image segmentation of GR thresholding method. Field Crops Research 149:33-39. *https://doi.org/10.1016/j.fcr.2013.04.007*

Woebbecke DM, Meyer GE, Von Bargen K, Mortensen DA (1995). Color indices for weed identification under various soil, residue, and lighting conditions. Transactions of the ASAE 38(1):259-269. *https://doi.org/10.13031/2013.27838*

Xu YF, Cheng Q, Wei XP, Yang B, Xia SS (2021). Monitoring of winter wheat growth under UAV using variation coefficient method and optimized neural network. Transactions of the Chinese Society of Agricultural Engineering 37(20):71-80. *https://doi.org/10.11975/j.issn.1002-6819.2021.20.008*

Yang B, Wang M, Sha Z, Wang B, Chen J, Yao X, ... Zhu Y (2019). Evaluation of aboveground nitrogen content of winter wheat using digital imagery of unmanned aerial vehicles. Sensors 19(20):4416. *https://doi.org/10.3390/s19204416*

Yang F, Dai H, Feng H, Yang G, Li Z, Chen Z (2016). Hyperspectral estimation of plant nitrogen content based on Akaike's information criterion. Transactions of the Chinese Society of Agricultural Engineering 32(23):161-167. *https://doi.org/10.11975/j.issn.1002-6819.2016.23.022*

Yang FQ, Feng HK, Xiao TH, Li TC, Guo XQ (2020). Nitrogen nutrition index estimation in winter wheat by UAV spectral information and texture feature fusion. Research of Agricultural Modernization 41(4):718-726. *https://doi.org/10.13872/j.1000-0275.2020.0061*

Zhang T, Wang MF, Zhao Q (2021). Effects of DPC and nitrogen fertilizer through drip irrigation on growth and yield in cotton. Crops 38(4):124-131. *https://doi.org/10.16035/j.issn.1001-7283.2022.04.017*

Zheng H, Li W, Jiang J, Liu Y, Cheng T, Tian Y, ... Yao X (2018). A comparative assessment of different modeling algorithms for estimating leaf nitrogen content in winter wheat using multispectral images from an unmanned aerial vehicle. Remote Sensing 10(12):2026. *https://doi.org/10.3390/rs10122026*

Zheng H, Ma J, Zhou M, Li D, Yao X, Cao W, ... Cheng T (2020). Enhancing the nitrogen signals of rice canopies across critical growth stages through the integration of textural and spectral information from unmanned aerial vehicle (UAV) multispectral imagery. Remote Sensing 12(6):957. *https://doi.org/10.3390/rs12060957*