Integrating UAV Derived Vegetation and Texture Indices for Estimation of Leaf Nitrogen Concentration in Drip-Irrigated Cotton under Reduced Nitrogen Treatment and Different Plant Densities

Minghua Li
Shihezi University

Yang Liu (ly.0318@163.com)
Shihezi University

Xi Lu
Shihezi University

Jiale Jiang
Shihezi University

Xuehua Ma
Shihezi University

Ming Wen
Gansu Agriculture University

Fuyu Ma
Shihezi University

Research Article

Keywords: Leaf nitrogen concentration, machine learning algorithms, UAV, texture, vegetation index

Posted Date: June 26th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-3074707/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License
Abstract

Background

Accurate assessment of nitrogen (N) status is important for N management and yield improvement. The N status in plant was affected by plant densities and N application rates, while the methods of assessing the N status in drip-irrigated cotton under reduced nitrogen treatment and different plant densities are lacking.

Methods

This study was conducted with four different N treatments (195.5, 299, 402.5, and 506 kg N ha\(^{-1}\)) and three sowing densities (6.9×10\(^4\), 13.8×10\(^4\), and 24×10\(^4\) plants ha\(^{-1}\)) by using a low-cost Unmanned Aerial Vehicle (UAV) system to acquire RGB imagery at 10 m flight altitude at cotton main growth stages. We evaluated the performance of different ground resolutions (1.3-, 2.6-, 5.2-,10.4-, 20.8-, 41.6-, 83.2-, and 166.4-cm-ground-resolution) image textures, vegetation indices (VIs), and their combination for leaf N concentrations (LNC) estimation with four regression methods (stepwise multiple linear regression, SMLR; support vector regression, SVR; extreme learning machine, ELM; random forest, RF).

Results

The results showed that the combination of VIs and texture maintained higher estimation accuracy than using VIs or textures alone. Specifically, the RF regression models had the higher accuracy and stability than SMLR and other two machine learning algorithms. The best accuracy (\(R^2 = 0.87\), RMSE = 3.14g kg\(^{-1}\), rRMSE = 7.00%) was obtained when RF was applied in combination with VIs and texture.

Conclusion

The combination of VIs and textures from UAV images using RF could improve the estimation accuracy of drip-irrigated cotton LNC and may have the potential contribution in rapid and non-destructive nutrition monitoring and diagnosis of other crops or other growth parameters.

1. Introduction

Nitrogen (N) is normally required more than other nutrients for cotton growth, and optimal application rate or ratio of N fertilizers is vital for cotton production and environmental sustainability (Hou et al., 2007). However, an excessive N application leads to greed and late maturity, while insufficient N affects development and yield formation (Bodirsky et al., 2014; Ata-Ul-Karim et al., 2017). Therefore, precise, real-
time, and rapid detection of cotton N nutrient status is beneficial to quantified the N fertilizers management and friendly environment.

Traditional N detection methods are based on direct measurements by chemical analysis, which relies heavily on field measurements and destructive sampling. Those processes making the measurements sufficiently accurate while the destructive process, time-consuming, labor-intensive, and non-reproducible limit their application in large scale and field conditions (Yao et al., 2010). The remote sensing (RS) technology is mainly used in crop N monitoring due to its rapid, non-destructive characteristics (LaCapra et al., 1996; Li et al., 2018), and the unmanned aerial vehicle (UAV)-based RS platforms are used widely in different crops field for data collecting and monitoring due to its low-cost at high Spatio-temporal resolution (Boegh et al., 2002; Tilling et al., 2007; Yao et al., 2014; Blaes et al., 2016).

The RGB, color infrared (CIR), multi-spectral, and high-spectral images derived from different sensors have been obtained and used on various UAV platforms to analysis crop growth conditions (Schut et al., 2018). Especially, low-cost UAV systems with RGB or modified CIR sensors have been widely used on crop N content and biomass estimation (Amaral et al., 2015; Jiang et al., 2019; Lu et al., 2019). However, the N status is strongly influenced by growth stages, and the poor correlation between LNC and RGB image parameters or vegetation indices were observed under the closed canopy or higher plant densities (Prey and Schmidhalter, 2019). Even more, the N signal of crop leaves and plants is hardly captured by canopy spectra at early growth stages, resulting from multiple background factors (eg., water, soil, mulch) and N dilution effects. Previous studies found that the degree of N dilution in winter wheat was more obvious in the late growth stage than that in the earlier growth stage (Zhao et al., 2016; Zhou et al., 2018). Therefore, there is little knowledge about how to enhance the N signal with lower or higher levels of crop canopy, and build up an appropriate model for N-status for whole growth stages.

Texture features can precisely compensate for the shortcomings brought by VIs. A large number of studies have shown that the texture information inherent from UAV imagery shows great potential for crop growth monitoring, especially under dense canopy cover conditions (Yue et al., 2019; Zheng et al., 2020). It has been found that the texture features extracted from UAV images maintained higher accuracy and more sensitivity of remote sensing data to biophysical variables, when comparing to VIs (Dube and Mutanga, 2015). During the early growth period of crops, the leaves often displayed significant color changes due to N nutrients consumed by plant growth and N fertilizer application (Yang et al., 2003). While these color variations between the interior of the crop canopy can be captured by texture indicators (Zheng et al., 2020). Therefore, the texture metrics derived from UAV images would have the potential to better describe the cotton N state changes throughout the growth stage.

The possible and better data fusion depends on the availability of various data sources (Zheng et al., 2019). Data fusion enabled a more complete interpretation of the connection between remote sensing data and crop parameters (Singh et al., 2002; Yang et al., 2003). Research has shown that the fusion of the spectral and texture information derived from UAV images could improve the predictive performance of crop biomass (Zheng et al., 2020). Since VI and texture have own advantages responding to crop
parameters under different growth stages and plant densities, combining their complementary information may help to improve the estimation of crop parameters across the critical growth stages. For examples, the fusion of VIs and texture data approaches have been utilized in estimating aboveground biomass of wheat (Yue et al., 2019), quantification of nitrogen (N) status of rice crop (Zheng et al., 2020), estimating the height of the canopy and aboveground biomass of maize (Lu et al., 2019), etc. The fused data performed better than single data for crop parameter estimation, however, the results found that most of the remote sensing parameters and crop parameters were not simple linear relationships, and the linear models were inadequate in capturing the complex and non-linear relationship between them (Niu et al., 2019). Machine learning regression algorithms have the advantages of dealing with high-dimensional data and nonlinear relationships, which can produce higher accuracy in biomass estimation compared to traditional linear regression techniques (Li et al., 2016; Lu et al., 2019). Among them, random forest (RF) (Breiman, 2001), support vector regression (SVR) (Mountrakis et al., 2011), and extreme learning machine (ELM) (Huang et al., 2006) are widely used for crop parameter estimation. So far as we known, few studies have explored machine learning techniques to estimate the LNC of cotton by combining VI and texture structure information from UAV images. Therefore, whether using machine learning algorithms to estimate drip-irrigated cotton LNC in textures and Vls has a better performance needs to be further explored.

Therefore, the aims of this study were (1) to assess the potential of combining VIs and texture information derived from a low-cost UAV-mounted RGB sensor to enhance the accuracy of estimating LNC in cotton; (2) to assess the estimation performance of three machine learning regression techniques (SVR, ELM, RF) compared to the traditional stepwise multiple linear regression (SMLR) throughout the cotton growth stages under different plant densities and reduced N treatments. The anticipated results will guide how to choose an inexpensive way to perform N estimation and lay the groundwork for the development of non-destructive, rapid monitoring of nitrogen status with UAV for cotton crops.

2. Materials and Methods

2.1 Experimental design

The experiment was conducted in Shihezi University Experiment Base, Shihezi, Xinjiang, China (44°29’ N, 86°1’ E) in 2019 and 2020. The average temperature/rainfall for both years was 20°C/168 mm and 21°C/228 mm during the cotton-growing season, respectively. Four different levels of N treatment (195.5, 299, 402.5, and 506 kg N ha⁻¹) and three levels of densities (6.9×10⁴, 13.8×10⁴, and 24×10⁴ plants ha⁻¹) were conducted using hybrid cotton cultivar Lumianyan24. Treatments were laid out in a split-plot design, with three replications (Fig. 1). The high-density cotton planting pattern was six rows of plants with three irrigation pipes, covered by a 2.05 m plastic film with a plant spacing of 11 cm. The low and medium density planting pattern is one film with three rows and three drip tapes with a plant spacing of 9.5 and 19 cm, respectively (Fig. 2). The area of each plot was 34.2m² (15×2.28), for a total of 36 plots, all other agricultural practices were performed based on local standards.
2.2 Ground Sampling and UAV data acquisition and pre-processing

The ground destructive samplings and UAV campaigns were taken at critical growth stages of cotton (Table 1). After the UAV campaign, three cotton plants were randomly sampled in each plot and divided them into three parts including leaves, stems and reproductive organs. The samples were then oven-dried at 105°C for 30 minutes, followed by drying at 80°C until a constant weight was achieved. Afterall, the dry samples were weighted, ground and passed through a 0.5 mm sieve to evaluate the plant N content. The method of N content assessment was using the micro-Keldjahl (Bremner, 1995).

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Sowing date</th>
<th>Date of UAV flights</th>
<th>Date of field sampling</th>
<th>Growth stage</th>
</tr>
</thead>
<tbody>
<tr>
<td>#1</td>
<td>24th April, 2019</td>
<td>1st July, 2019</td>
<td>1st July, 2019</td>
<td>Full bud stage</td>
</tr>
<tr>
<td></td>
<td>16th July, 2019</td>
<td>16th July, 2019</td>
<td>16th July, 2019</td>
<td>Full flowering stage</td>
</tr>
<tr>
<td></td>
<td>7th August, 2019</td>
<td>7th August, 2019</td>
<td>7th August, 2019</td>
<td>Early full-bolling stage</td>
</tr>
<tr>
<td></td>
<td>16th August, 2019</td>
<td>16th August, 2019</td>
<td>16th August, 2019</td>
<td>Late full-bolling stage</td>
</tr>
<tr>
<td></td>
<td>7th September, 2019</td>
<td>8th September, 2019</td>
<td>8th September, 2019</td>
<td>Boll opening stage</td>
</tr>
<tr>
<td>#2</td>
<td>April 18, 2020</td>
<td>14th June, 2020</td>
<td>15th June, 2020</td>
<td>Full bud stage</td>
</tr>
<tr>
<td></td>
<td>21st August, 2020</td>
<td>21st August, 2020</td>
<td>21st August, 2020</td>
<td>Late full-boll stage</td>
</tr>
<tr>
<td></td>
<td>13th September, 2020</td>
<td>13th September, 2020</td>
<td>13th September, 2020</td>
<td>Boll opening stage</td>
</tr>
</tbody>
</table>

The UAV system was the DJI Mavic Pro series with a four-rotor and a digital camera. Before the preliminary flight, 16 ground control points (GCPs) were established throughout the cotton field experiment site, each marked with a sign, to georeference the UAV images taken at different growth stages, and the Real-Time Kinematic Global Positioning System (RTK-GPS, CHC X900 GNSS) with vertical and horizontal errors within 2 cm and 1 cm was used to acquire the original geographic coordinates.

In our campaign, the UAV with auto-flight mode and predefined operational plan was set to acquire RGB images with about 82% forward and lateral overlap, and the frequency of images acquisition was 1 frame per 5s and the images were saved in JPEG format. Besides, the setting of aperture of camera was f/5, and the flight height was 10 m above ground with the speed of 2 m s⁻¹. All parameters (except the
exposure time) related to camera and UAV were same during the UAV campaign in whole seasons. Each flight campaign was held on a sunny day between 12:00 am and 14:00 local time and approximately 289 images were acquired with a ground resolution of 1.3 cm.

The Agisoft Photoscan 1.2.6 (Agisoft LLC, St. Petersburg, Russia) software was used to process the UAV images to obtain orthophotos. The detailed images processing methodology we used in this study was followed by Lu et al., (2019), and the processing steps and parameter settings can be found in Table 2.

### Table 2
Processing steps with corresponding parameter settings in Agisoft Photoscan software for generation of orthophotos from UAV imagery

<table>
<thead>
<tr>
<th>Task</th>
<th>Parameter setup</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aligning image</td>
<td>Accuracy: high; Pair selection: generic; Key points: 40,000; Tie points: 4000</td>
</tr>
<tr>
<td>Building mesh</td>
<td>Surface type: height field; Source data: dense cloud; Face count: high</td>
</tr>
<tr>
<td>Positioning guided marker</td>
<td>Manual positioning of markers on the even 16 GCPs for all the photos</td>
</tr>
<tr>
<td>Optimizing cameras</td>
<td>Default settings</td>
</tr>
<tr>
<td>Building dense point cloud</td>
<td>Quality: high; Depth filtering: mild</td>
</tr>
<tr>
<td>Building texture</td>
<td>Mapping mode: Generic; Blending mode: Mosaic; Texture size/count: 4096</td>
</tr>
<tr>
<td>Building DEM</td>
<td>Surface: Mesh; Other parameters: default</td>
</tr>
<tr>
<td>Building orthomosaic</td>
<td>Surface: Mesh; Other parameters: default</td>
</tr>
</tbody>
</table>

### 2.3 Selection of image textures

The image textures based on the gray tone spatial dependence matrix (Table 3), using in this study, were defined by Haralick (Haralick and Sabaretnam, 1973). And the calculation formulas of eight image textures were related to all three bands in RGB image using a 3×3 calculation window. The textures in Table 3 are calculated by using
Table 3
Detailed information of textures, calculation window size, and image ground resolution in this study

<table>
<thead>
<tr>
<th>Textures and Abbreviations</th>
<th>Bands</th>
<th>Windows</th>
<th>Ground resolutions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variance (VAR), Entropy (EN), Correlation (COR), Homogeneity (HOM), Second Moment (SE), Dissimilarity (DIS), Contrast (CON), Mean (MEA)</td>
<td>R, G, B</td>
<td>3x3</td>
<td>1.3 cm, 2.6 cm, 5.2 cm, 10.4 cm, 20.8 cm, 41.6 cm, 83.2 cm, 166.4 cm</td>
</tr>
</tbody>
</table>

\begin{align*}
VAR &= \sum_i \sum_j (i-u)^2 p(i,j), \\
HOM &= \sum_i \sum_j \frac{1}{1+(i-j)^2} p(i,j), \\
CON &= \sum_{n=0}^{N_g} n^2 \left[ \sum_i \sum_{j|j-i=n} p(i,j) \right], \\
EN &= -\sum_i \sum_j p(i,j) \log \left( p(i,j) \right), \\
SE &= \sum_i \sum_j \left[ p(i,j) \right]^2, \\
MEA &= \sum_{i=2}^{2N_g} i p_{x+y}(i), \\
COR &= \sum_i \sum_j \frac{(i-j)p(i,j) - u_x u_y}{\sigma_x \sigma_y}, \\
DIS &= \sum_{n=1}^{N_g-1} n \left[ \sum_i \sum_{j|j-i=n} p(i,j) \right].
\end{align*}

Where \( p(i,j) \) is the \((i,j)\) entry in a normalized gray-tone spatial-dependence matrix \( P(i,j)/R_{p_x}(i)=P(i,j)/R \); \( p_x(i) \) is the entry in the marginal-probability matrix obtained by summing the rows of the rows of \( p(i,j) \); \( N_g \) is the number of distinct gray levels in the quantized image; \( p_x(j)=\sum_i p(i,j)p_x(i) \); \( u_x, u_y, \sigma_x, \) and \( \sigma_y \) are the means and standard deviations of \( p_x \) and \( p_y \).

Further details of the calculations can be found in Haralick (Haralick and Sabaretnam, 1973).

### 2.4 Selection of vegetation indices

Ten vegetation indices of the RGB images were used to estimate the LNC in this study (Table 4), and the selected VIs were based on the three bands of the original RGB images.
Table 4 Summary of vegetation indices derived from the aerial orthophotos for the estimation of LNC in cotton

<table>
<thead>
<tr>
<th>Index</th>
<th>Name</th>
<th>Formulation</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>IKAW</td>
<td>Kawashima Index</td>
<td>(IKAW = \frac{R-B}{R+B} )</td>
<td>(Kawashima and Nakatani, 1998)</td>
</tr>
<tr>
<td>RGBVI</td>
<td>Red Green Blue Vegetation Index</td>
<td>(RGBVI = \frac{G^2 - B \cdot R}{G^2 + B \cdot R} )</td>
<td>(Bendig et al., 2015)</td>
</tr>
<tr>
<td>MGRVI</td>
<td>Modified Green Red Vegetation Index</td>
<td>(MGRVI = \frac{G^2 - R^2}{G^2 + R^2} )</td>
<td>(Bendig et al., 2015)</td>
</tr>
<tr>
<td>GLI</td>
<td>Green Leaf Index</td>
<td>(GLI = \frac{2 \cdot g \cdot r \cdot b}{2 \cdot g \cdot r + b} )</td>
<td>(Louhaichi et al., 2001)</td>
</tr>
<tr>
<td>ExGR</td>
<td>Excess Green minus Excess Red</td>
<td>(ExGR = \frac{1.4 \cdot R - G}{G + R + B} )</td>
<td>(Meyer and Neto, 2008)</td>
</tr>
<tr>
<td>GRVI</td>
<td>Green Red Vegetation Index</td>
<td>(GRVI = \frac{G - R}{G + R} )</td>
<td>(Johnson et al., 2001)</td>
</tr>
<tr>
<td>ExG</td>
<td>Excess Green Index</td>
<td>(ExG = \frac{2 \cdot g \cdot r \cdot b}{2 \cdot g \cdot r + b} )</td>
<td>(Woebecke et al., 1995)</td>
</tr>
<tr>
<td>VARI</td>
<td>Visible Atmospherically Resistant Index</td>
<td>(VARI = \frac{g \cdot r \cdot b}{g + r + b} )</td>
<td>(Gitelson et al., 2002)</td>
</tr>
<tr>
<td>GBRI</td>
<td>Green blue ratio index</td>
<td>(GBRI = \frac{G}{B} )</td>
<td>(Gamon and Surfus, 1999)</td>
</tr>
<tr>
<td>GRRI</td>
<td>Green red ratio index</td>
<td>(GRRI = \frac{G}{R} )</td>
<td>(Gamon and Surfus, 1999)</td>
</tr>
</tbody>
</table>

Note: \(R, G, \) and \(B\) represent the digital number of red, green and blue channels, respectively. \(r = R/(R+G+B), g = G/(R+G+B), b = B/(R+G+B)\).

### 2.5 Regression modeling methods

Compared to traditional parametric regression methods, machine learning algorithms would be more suitable for predictive model building for multiple input variables. For building an independent model on LNC estimation using VIs, textures, and their combinations, three machine learning techniques were used and realized in the Rx64 3.4.0 environment software with the caret package (R Development Core Team, 2017).

Random Forest (RF) is a non-linear integrated modeling method with multiple decision trees, and consists of Bootstrap and random subspace methods (Breiman, 2001). Based on this, it shows a better capability to overcome the overfitting problems and to handle small datasets, as well as massive inputs variables (Mutanga et al., 2012; Lu et al., 2019). Mtry and ntree were used in RF to attain the best predictive power, and the ntree was fixed to 1300 and only adjusted mtry to optimize the RF model in this study.

Extreme learning machine (ELM), being a single hidden layer feed-forward neural network, has a faster learning speed, lower training error, and minimum output weight specification than the traditional feed-forward networks (Huang et al., 2006; Huang et al., 2012). It is suitable for real-time training as the
weights of its hidden layer may be generated randomly without any iteration optimization. Moreover, ELM has a better capability to handle complex data and develop regression with multiple highly.

Support vector machine (SVM) is a very powerful machine learning regression approach and based on statistical theory, usually used for pattern recognition and nonlinear regression (Cortes and Vapnik, 1995). SVR has been widely used in previous remote sensing studies to estimate crop parameters (Gao et al., 2018). Due to its ability to handle high-dimensional data and train models with a relatively small number of samples (Mountrakis et al., 2011). In this study, each SVM model automatically estimated a fixed $\sigma$ value for every input variable according to the kernel-based machine learning regression methods provided by the R package "kernlab" (Lin et al., 2017). The Cost was to be further tuned to optimize the predictive performance of the SVM model.

Stepwise multiple linear regression (SMLR) was a commonly used regression method in crop growth parameters (Niu et al., 2019). The variables extracted from the UAV images may be interrelated, the relationships among each variable and the relationships between these variables with the LNC were subjected to simple linear regression as measured with Pearson's correlation coefficient. Moreover, to better assess the performance of three machine learning algorithms methods with traditional regression techniques, the SMLR was used as a reference in this assessment.

2.6 Accuracy assessment

The objective was to build a generalized model across multi-treatment, growth periods, and seasons using various regression techniques, and we pooled 2-years of data with all growth conditions to form a holistic dataset. A generalized model would be more practical than a local model on account of avoiding frequent model calibration for different growth conditions. The dataset was divided into two parts, with 70% of the data used for model calibration and the remaining 30% used for model validation. The coefficient of determination ($R^2$), the root mean square error (RMSE), and the akaike information criterion (AIC) were used to evaluate the accuracy of model calibration. The estimation accuracy is to be evaluated by $R^2$, RMSE, average test prediction accuracy (ATPA), and relative RMSE (rRMSE) of the validation data.

3. Results

3.1 Correlation between LNC, VIs and texture

We analyzed the changes of digital numbers (DN) in different channels with the variation of LNC from the RGB image (Fig. 3a). The DN values in all channels were slightly increased when LNC increased to 4.25%, and then decreased to be flatter as the LNC increased. The DN values in the green channel were higher than those in the blue channel. Further, Fig. 3b showed that the correlation between UAV-derived vegetation indices (based on 1.3-cm-ground-resolution images) with LNC, and the stronger positive correlations were observed in RGBVI, ExGR, ExG, and GLI with Pearson correlation coefficient values above 0.7. Normal R, G, and B were weakly negatively correlated with LNC and their Pearson correlation coefficient values were below −0.5.
For texture-related variables, that most of the image textures in the three channels were strongly correlated with LNC in the five individual growth stages, and the correlation coefficients of the same texture metrics in each channel were close to LNC. During the whole growth stages, only VAR, CON, and DIS had high correlation coefficients with LNC, and they performed consistently under three channels (Fig. 4). Thus, we selected high correlation coefficients (>0.6) variables GLI, ExG, ExGR, GRVI, GBRI, GRRI, MGRVI, and RGBVI variables in the VIs group, and VAR, HOM, CON, DIS in all three channels in the texture group for model developing.

3.2 Comparison of LNC estimation performance among SMLR and machine learning techniques

Table 5 presents the comparison of SMLR, SVR, ELM, and RF for LNC estimation throughout the growth stages of cotton. The results indicate that using VIs alone, RF obtained the optimal calibration ($R^2 = 0.76$, RMSE = 3.72 g kg$^{-1}$, AIC = 414.83) and validation ($R^2 = 0.77$, RMSE = 4.98 g kg$^{-1}$, rRMSE = 8.51%) with performance across the four regression techniques. However, SMLR also showed high accuracy (validation: $R^2 = 0.74$, RMSE = 4.17 g kg$^{-1}$, rRMSE = 12.78%) and was competitive with RF. When using the textures alone, the best performance remained for RF, with a very close performance for SMLR and SVR. In contrast, ELM exhibited lower accuracy than SMLR, and RF achieved the highest level of accuracy among the four regression techniques. Furthermore, this accuracy with RF (Calibration: $R^2 = 0.85$, RMSE = 2.85 g kg$^{-1}$, AIC = 378.59; validation: $R^2 = 0.85$, RMSE = 3.61 g kg$^{-1}$, rRMSE = 7.09%) was even higher than that achieved using the VIs, with $R^2$ increasing by 0.08 for the validation data and RMSE increased by 1.37 g kg$^{-1}$.
### Table 5
Accuracy assessment for the estimation of LNC from vegetation indices, texture and their combinations with SMLR and three machine learning algorithms

<table>
<thead>
<tr>
<th>Input variables</th>
<th>Technique</th>
<th>Calibration (N = 288)</th>
<th>Validation (N = 72)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>R²</td>
<td>RMSE (g/Kg)</td>
</tr>
<tr>
<td>VIs</td>
<td>SMLR</td>
<td>0.72</td>
<td>4.11</td>
</tr>
<tr>
<td></td>
<td>SVR</td>
<td>0.70</td>
<td>4.12</td>
</tr>
<tr>
<td></td>
<td>ELM</td>
<td>0.68</td>
<td>4.25</td>
</tr>
<tr>
<td></td>
<td>RF</td>
<td>0.76</td>
<td>3.72</td>
</tr>
<tr>
<td>Textures</td>
<td>SMLR</td>
<td>0.83</td>
<td>3.21</td>
</tr>
<tr>
<td></td>
<td>SVR</td>
<td>0.82</td>
<td>3.17</td>
</tr>
<tr>
<td></td>
<td>ELM</td>
<td>0.77</td>
<td>3.66</td>
</tr>
<tr>
<td></td>
<td>RF</td>
<td>0.85</td>
<td>2.85</td>
</tr>
<tr>
<td>VIs and Textures</td>
<td>SMLR</td>
<td>0.84</td>
<td>3.11</td>
</tr>
<tr>
<td></td>
<td>SVR</td>
<td>0.83</td>
<td>3.13</td>
</tr>
<tr>
<td></td>
<td>ELM</td>
<td>0.78</td>
<td>3.63</td>
</tr>
<tr>
<td></td>
<td>RF</td>
<td>0.87</td>
<td>2.80</td>
</tr>
</tbody>
</table>

Note: The accuracy metrics were calculated from calibration data and validation data separately. The number in bold for each column represents the maximum R², minimum RMSE, minimum AIC and minimum rRMSE, respectively.

The fusion of VIs and textures further improved for all the regression techniques.

Compared with the traditional method using only VIs, their accuracies were significantly enhanced, the R² of the three regression techniques, SVR, ELM, and RF, increased by 0.14, 0.11, and 0.10, respectively. Consistently, RF produced the most high accuracy in calibration (R² = 0.87, RMSE = 2.80g kg⁻¹, AIC = 378.59) and validation (R² = 0.87, RMSE = 3.14g kg⁻¹, rRMSE = 7.00%). The scatter plots in Fig. 5 also showed that the data points were typically closer to the 1:1 line by their combination data.

### 3.3 The correlation of image textures at different ground resolutions of the image

The Pearson correlation coefficients of the input variables for each channel texture at different ground resolutions images were shown in Fig. 6. The results showed that there was a positive correlation between most textures of images of different ground resolutions. Specifically, a positive correlation existed between the image textures of 1.3-, 2.6-, 5.2-, 10.4-, 20.8-, and 41.6-cm-ground-resolution images,
and the correlation between the texture of the original resolution image and the texture at other resolutions gradually decreases as the resolution decreases, which was most obvious at 41.6 cm. It was shown from another side that the cotton image texture depended on the image ground resolution, therefore, while estimating the LNC, the image texture needed to be calculated from images with different ground resolutions.

3.4 The effect of different ground resolutions images on LNC estimation by using RF

Figure 7 shows the performance of eight different ground resolution images for LNC estimation using only RF. In general, relatively higher ground resolution images yield higher estimation accuracy. Specifically, the ground resolutions above 20.8 cm pixel$^{-1}$ had relatively high estimation accuracy with $R^2$ above 0.77, regardless of using which metric or regression. Secondly, using the same regression method, the accuracy of the combination of VIs and texture was stronger than using VIs or texture alone, while the performance of using texture was stronger than using VIs. The highest accuracy was obtained at 5.2 cm pixel$^{-1}$ with RF by using a combination of VIs and texture. Although the comparable estimation accuracy can also be obtained ground resolution at 10.4 cm, and 20.8 cm pixel$^{-1}$, it was still not as stabilized as the ground resolution at 5.2 cm pixel$^{-1}$ (Fig. 8).

3.5 Performance of three machine learning techniques on LNC estimation at different ground resolutions

The distributions of $R^2$, RMSE, and ATPA for the testing datasets are displayed in Fig. 8. The three machine learning techniques yield similar results to the calibration, with a slightly higher ATPA being observed on the validation data sets. The validation results showed that the regression techniques of RF and SVR had closely approximated $R^2$ and RMSE values for the estimation of LNC. Specifically, The RMSE, $R^2$ and ATPA were close at 1.3-, 2.6-, 5.2-, 10.4-, 20.8-, and 41.6-cm-ground-resolution images. However, the $R^2$, RMSE, and ATPA changed sharply at 41.6 cm ground resolution, and $R^2$ and ATPA decreased faster (except for VIs). Under combined data, The RF with 5.2 cm as the input variable had the highest LNC estimation accuracy, and The SVR with 2.6 cm had the highest LNC estimation accuracy. Compared with the SVR models, the RF models had higher accuracy for LNC estimations. The highest ATPA values for LNC estimations were 90.31%, 94.13%, and 95.42%, respectively, in RF models, which were 4.12%, 5.45%, and 5.73% higher than the corresponding optimal ELM models.

4. Discussion

4.1 The combination of VIs and textures improves the accuracy of LNC estimation

For a long time, the VIs-derived from UAV image was a widely used method for crop LNC estimation, however, its performance still needs improvement in scenarios where only RGB images are accessible (Lu
et al., 2019). There are three reasons for the reduced accuracy achieved by utilizing only VIs from RGB images. The first reason is the lack of near-infrared channels, which hinders the ability to detect variations in vegetation vigor. Secondly, saturate problems in higher canopy cover conditions (Jin et al., 2015). The third reason is the challenge of converting digital number values into reflectance due to the broad spectral ranges of visible channels and imprecise spectral response functions (Zheng et al., 2018). Moreover, the spectrum information utilized in the vegetation indices was primarily obtained from the uppermost layer of leaves in the cotton canopy, which cannot reflect the information of the middle and lower leaves, especially the leaves in the center to late-developing season (Li et al., 2016). Whereas, the texture describes the vertical structure of the canopy (Dube and Mutanga, 2015), which can precisely compensate for the shortcomings caused by VIs.

Research has shown that the accuracy of using textures to estimate N nutrition parameters was better than that using VIs in rice (Zheng et al., 2020). In this study, eight textures with higher correlation with LNC were chosen to estimate LNC (Fig. 4), and the accuracy of the estimated LNC using texture alone was significantly higher than that using VIs alone regardless of the regression method. These results were similar to previous studies, but they only used a simple linear regression method in other crops (Yue et al., 2019). The reason why using textures performs better than using VIs may due to the tonal variation between crop canopy interiors that can be captured by the texture metric.

Moreover, we further found that using the combination of VIs and texture had a significant improvement for LNC estimation than that using VIs or textures alone (Table 5). This may be attributed to the statistical advantage due to the increased number of sources, or the redundancy effect of compensating for noise or defective sources, as well as the superposition of the two advantages. Earlier research has drawn comparable conclusions, but they have utilized a combination of vegetation indices and canopy height metrics to estimate wheat biomass (Lu et al., 2019). Our results illustrated that the fusion of VIs and texture data could be effective in improving N nutrient monitoring. Further, we also tested the multivariate models of LNC on different datasets and the validation results were satisfactory (Table 5 and Fig. 7). However, the suitability of those models may still need to be enhanced by further testing on a wider range of datasets from different geographic sites.

4.2 Comparison of the four regression methods

The stepwise multiple linear regression (SMLR) has been reported to overestimate the vegetation parameters while quantifying them (Grossman et al., 1996). In this study, the accuracy of SMLR was equal to that in SVR, which may result from the moderate number of SMLR input variables in this study (no more than 22). This is consistent with the results of Li et al (Li et al., 2016). However, hundreds or thousands of bands are available in spectral analysis (Jia et al., 2013), so SMLR may not able to handle high-dimensional information. On the other side, machine learning algorithms can handle high-dimensional information with multiple input variables and have been widely used to deal with strong nonlinear relationships between crop biochemical parameters with remotely sensed variables (Gleason and Im, 2012). RF regression has been regarded as one of the popular ensemble learning algorithms that
can combine a large number of regression sub-models and does not respond to noise and over-fitting (Breiman, 2001).

In this study, the performance of RF for LNC estimation using VIs, textures or their combinations were better than that in SMLR, SVR, and ELM (Fig. 5 and Fig. 8), which may be due to the insensitivity of RF to noise and over-fitting. The results in our study were similar to previous studies in that they found the performance of using RF was better for SMLR, SVR, and artificial neural networks (ANN) for wheat and maize biomass estimation (Mutanga et al., 2012; Li et al., 2016). SVR and ELM are both very powerful machine learning regression methods, the biggest advantage of SVR can train a small number of samples (Mountrakis et al., 2011), and ELM does not require much human intervention or any kernel function, being an efficient and fast learning algorithm (Huang et al., 2012). In our study, the accuracy of SVR was consistently moderate while ELM's was the worst, which may be due to the input weights and implied layer thresholds are randomly determined in ELM. The utilization of RF regression was found to be advantageous in achieving high accuracy, as it is a dependable and robust method for handling complex and nonlinear regressions. This assertion has also been supported by previous studies (Wang et al., 2016). The performance of RF regression is still needed to be validated in cotton LNC estimation with datasets from more study areas and cultivars.

4.3 The optimal resolution for LNC estimation

A potential disadvantage of acquiring high-resolution images (with a pixel size of 1.3 cm) is the need to fly the UAV at low altitudes, which can be a significant obstacle to acquiring images efficiently over large areas (Lu et al., 2019). To overcome this issue, one solution could be to use cameras with even higher resolution and fly the UAV at higher altitudes. However, this approach would come with a higher cost and increased weight of the equipment. Instead, relatively lower resolution images can still generate an acceptable accuracy (Zarco-Tejada et al., 2014; Dube and Mutanga, 2015). Our results also confirm that similar accuracy was obtained for ground resolution images at 2.6-, 5.2-, 10.4-, 20.8-, and 41.6-cm pixel$^{-1}$, with $R^2$ above 0.77, compared to a ground resolution of 1.3 cm pixel$^{-1}$ (Fig. 7). This observation is supported by the results of the correlation analysis performed on the texture variables at varying ground resolutions (Fig. 6), where the correlation between the image textures at 1.3-, 2.6-, 5.2-, 10.4-, 20.8-, and 41.6-cm pixel$^{-1}$ was relatively high. However, the performance of ground resolution at 10.4-, 20.8-, and 41.6-cm pixel$^{-1}$ were not stabilized (Fig. 8). Therefore, it was possible to maintain a similar performance among 1.3-5.2-cm pixel$^{-1}$ by adjusting the flight altitude.

In this study, the optimal resolution estimate was 5.2 cm pixel$^{-1}$ image resolution for the LNC estimation (Fig. 7 and Fig. 8). When the resolution was reduced to 5.2 cm pixel$^{-1}$ from the initial resolution of 1.3 cm pixel$^{-1}$ at 10 m, the shape of the cotton canopy underwent only slight changes but remained discernible with ease (Fig. 9). When decreased to a lower resolution, the mixed pixels from the cotton and soil background resulted in the decreased accuracy of LNC estimation. Based on our findings, we suggest utilizing a resolution of 5.2 cm pixel$^{-1}$ for UAV campaigns. This approach would enable us to enhance flight efficiency up to four times with the same UAV without compromising the accuracy of the
estimation, leading to significant savings in UAV flight time. Therefore, a relatively low resolution (5.2 cm pixel\(^{-1}\)) image by increasing the flying height or using a lower resolution camera can still produce an acceptable accuracy.

5. Conclusions

The results of this study indicate that the combination of VIs and texture features enhanced the accuracy of cotton LNC estimation when compared to using VIs or texture features alone. The performance of using the RF algorithm consistently outperformed the other three regression methods (SMLR, SVR, ELM) among the three types of input data. Moreover, we have shown that better accuracy at high resolutions at 1.3 cm pixel\(^{-1}\) using RF techniques can still be obtained at a relatively low resolution of 5.2 cm pixel\(^{-1}\), that the resolution was decreased to one-fourth of the initial orthophotos. Therefore, the combination of VIs and texture information obtained from an inexpensive UAV system using the RF algorithm could have potential in the quick estimation of other parameters for practical applications.

Declarations

Acknowledgements

We thank Dr. Ming Wen, Xi Lu, Jiale Jiang, Xuehua Ma for assistance with data collection and insightful discussion. Thank Professor Yang Liu for his theoretical and technical guidance, and we thank Home for Researchers editorial team (www.home-for-researchers.com) for language editing service.

Author contributions

MGL, YL, and FYM designed the experiments. MGL and YL performed the experiments. MGL, XL, and JLJ analyzed the data. MGL, and YL wrote the manuscript. MGL, MW, and JLJ made the figures. All authors read and approved the final manuscript.

Funding

This work was supported by the National Natural Science Foundation of China under Grant [31860346], the Financial Science and Technology Plan Project of XPCC (2020AB017, 2020AB018), the Financial Science and Technology Plan Project of Shihezi City (2020ZD01), and Shihezi University Scientific Research Cultivation Project for Young Scholars [CXBJ202001].

Availability of data and materials

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Ethics approval and consent to participate
Not applicable.

**Consent for publication**

All authors affirm consent for publication.

**Competing interests**

All authors declare that they have no competing interests.

**References**


Figures
Figure 1

Location of the experimental site and layout of the field plots
Figure 2

Diagram of three density planting patterns for field trials
Figure 3

The changes of digital number values in different channels with LNC and Pearson's correlation coefficients between LNC and individual UAV-derived vegetation indices.

Figure 4

The Pearson correlation coefficient between the different bands of the texture and LNC.
Figure 5

Scatterplots of measured LNC (g kg\(^{-1}\)) in cotton and the LNC estimated with three machine learning techniques.
Figure 6

Pearson's correlation coefficient between input variables of the texture
Figure 7

The performance of LNC estimation from three types of input data
Figure 8

Testing of the three machine learning techniques to estimation of LNC from three types of input data.
Figure 9

RGB image and corresponding B-band DIS and values