Estimation of Canopy Nitrogen Density of Lodging Maize via UAV-based Hyperspectral Images

Qian Sun a, b, Liping Chen a, c, Sen Zhang b, Xiaohe Gu b,*, Jingping Zhou b, Limin Gu d, Wenchao Zhen e

a College of Information and Electrical Engineering, China Agricultural University, Beijing 100083, China
b Research Center of Information Technology, Beijing Academy of Agriculture and Forestry Sciences, Beijing 100097, China
c Research Center of Intelligent Equipment, Beijing Academy of Agriculture and Forestry Sciences, Beijing 100097, China
d Key Laboratory of North China Water-saving Agriculture, Ministry of Agriculture and Rural Affairs, Baoding Hebei 071001, China
e College of Agronomy, Hebei Agricultural University, Baoding Hebei 071001, China

* Correspondence: 405174959@qq.com; guxh@nercita.org.cn

Abstract: Rapid and non-destructive monitoring the temporal dynamic changes of canopy nitrogen status of lodging maize is of great significance to explore the influence of lodging on plant nutrient transport and yield loss of maize. The purpose of this study was to use UAV-based hyperspectral imaging technology to monitor the temporal changes of canopy nitrogen density (CND) of lodging maize in plot-scale. Through the combination of physiological index (leaf nitrogen content) and canopy structural index (canopy height) of maize population, the CND was constructed to indicate the severity of lodging maize. Firstly, three feature selection algorithms were selected to reduce the dimensionality of hyperspectral data, including mutual information (MI), recursive feature elimination (RFE) and SelectFromModel (SFM). Then, the spectral-based CND estimation model was constructed by using canopy spectral information and multi-temporal observation data of CND, and nine inversion models were constructed based on different dimensionality reduction and regression methods. Finally,
the accuracy of each CND model was tested by the determination coefficient ($R^2$) and root mean square error (RMSE). The main results were as follows: (1) The CND of non-lodging maize was the lowest, and the more serious the lodging, the higher the maize CND. It was because CND was highly related to the canopy height of the maize population. (2) With the same modeling algorithms (such as linear regression, LR), RFE algorithm (VT stage, $R^2=0.47$, RMSE =26.94 mg/m3) was better than MI ($R^2=0.34$, RMSE =30.03 mg/m3) and SFM ($R^2=0.42$, RMSE =28.10 mg/m3) algorithm. (3) In the same dimensionality reduction method, the accuracy of RFR (R3 stage, MI-RFR, $R^2=0.92$, RMSE =15.27 mg/m3) algorithm was higher than that of ABR (MI-ABR, $R^2=0.89$, RMSE =17.72 mg/m3) and LR (MI-LR, $R^2=0.53$, RMSE =36.55 mg/m3) algorithms. (4) Among the 9 inversion models, the fitting effect of RFE-RFR algorithm was the best and the model was the most stable (VT stage, training set: $R^2=0.92$, RMSE =10.60 mg/m3, testing set: $R^2=0.92$, RMSE =9.84 mg/m3; R3 stage, training set: $R^2=0.93$, RMSE =13.83 mg/m3, testing set: $R^2=0.92$, RMSE =24.42 mg/m3). Therefore, the algorithm combining feature selection with regression model effectively reduced the dimension of hyperspectral data and improved the estimation accuracy and computational efficiency of lodging maize CND. This study provided a reference for efficient and non-destructive estimation of lodging maize CND using UAV-based hyperspectral images.

Keywords: lodging maize, unmanned aerial vehicle (UAV), hyperspectral images, canopy nitrogen density (CND), feature selection.

1 Introduction

Maize is one of the three major crops in the world, which plays a vital role in global crop production and security (Herrmann et al., 2020; Mi et al., 2011). Due to the influence of climate warming, extreme weather occurs frequently around the world, which has a serious influence on agricultural production. In recent years, the extremely heavy precipitation caused by typhoons showed an increasing trend and was mainly concentrated in summer. At this time, it was in the prosperous period of maize growth, coupled with its structural traits (such as tall plants, thin stalks and weak lodging resistance), maize was prone to lodging (Abedon et al., 1999; Hondroyianni et al., 2000). Lodging stress not only decreased photosynthetic rate and grain quality (Remison & Akinleye, 1978), but also decreased yield (Setter et al., 1997) and hindered mechanized harvest in large area (Yinghui et al., 2005). The period and severity of lodging had different influences on the maize growth (Mi et
Lodging occurred in the early stage (V10-V12), middle stage (V13-V15) and late stage (V17-R1) of maize, the yield decreased by 2-6%, 5-15% and 13-31% compared with the control group (Carter & Hudelson, 1988). Lodging inhibited the transport of plant nutrient and influenced crop growth, result in the yield loss of crop. Therefore, the rapid accurate diagnosis and timing change analysis of nitrogen status in maize after lodging was helpful to explore the influences of lodging on plant nutrient transport and yield loss.

The traditional lodging measurement method was that agricultural technicians got into the field to investigate the number, type and angle of lodging plants. Traditional methods were based on visual assessment, which depended on the observer's skills, experience and was very time-consuming (Bock et al., 2010; Singh et al., 2019). The remote sensing provides a new technology for crop lodging monitoring. The spectral detection technology was fast, real-time, non-destructive and convenient (Atzberger, 2013). UAV-based hyperspectral has high spatial and spectral resolution, which is helpful to obtain subtle changes of ground objects (Faïcal et al., 2017; Sahoo et al., 2015). It is an important means to obtain crop lodging and growth information on a small scale, especially in crop lodging emergency monitoring. Lodging led to the decrease of canopy height, more leaves overlapping each other, which changed the light conditions and photosynthetic capacity of plant components (mainly stalk and leaf). The growth of lodging crops was stressed, the roots and stalks were damaged, which inhibited the transshipment and transportation of nutrients, resulting in the yield loss of crop (Setter et al., 1997).

Lodging crops inhibited leaf photosynthesis, and the damage to the root and stalk of the plant influenced the nitrogen transport between soil and plant, resulting in the reduction of nitrogen. Leaf nitrogen is closely related to chlorophyll, and is one of the most important elements in chloroplast and Rubisco enzyme. Plant nitrogen content is directly related to crop senescence and photosynthesis intensity (Wang et al., 2021). There are narrow bandwidth and continuous wavebands in hyperspectral data, showing subtle changes in crop spectral features, which is beneficial to the meticulous monitoring of nitrogen in lodging crops. Hyperspectral images were used
to accurately estimate key crop traits (nitrogen, chlorophyll and photosynthetic capacity) at the leaf and canopy level, and to assess the influences of nitrogen deficiency on crop yield (Wang et al., 2021). Many studies on crop nitrogen were carried out using hyperspectral data, including leaf nitrogen content (LNC) (Ma et al., 2022; Raj et al., 2021), leaf nitrogen accumulation (LNA) (Guo et al., 2021), plant nitrogen concentration (PNC) (Li et al., 2010), nitrogen nutrition index (NNI) (Liu et al., 2020) and so on. The LNC of maize was estimated using airborne hyperspectral data at the top of the canopy (Raj et al., 2021). As the acquisition of spectral data was often based on the canopy level, the crop nitrogen estimation model was also based on the canopy level. When the canopy leaves were taken as a whole object, most studies were conducted on the upper leaves, which ignored the vertical heterogeneity of the crop canopy. Some studies explored the vertical distribution of nitrogen in crop canopy at different growth stages, and developed an effective method to estimate nitrogen in each leaf layer and total nitrogen in canopy by determining the correlation between different leaf layers and nitrogen status (He et al., 2019; Li et al., 2022). Because the canopy leaf cover was influenced by different growth stages, the LNC in the canopy was not constant from the top to the bottom of the leaf layer and varied with the growth stage. Therefore, the canopy nitrogen content (CNC) (Wang et al., 2022) and canopy nitrogen density (CND) (Li et al., 2019; Ma et al., 2022; Zhao et al., 2012) were proposed and explored. The CNC and canopy carbon content (CCC) of maize were evaluated by hyperspectral remote sensing data (Wang et al., 2022). Multiple lookup tables (Multi-LUT) were used to estimate CND of winter wheat based on field and UAV hyperspectral data (Li et al., 2019). It was found that CND was more sensitive than LNC to the change of canopy spectral features (Zhao et al., 2012). The research progress of remote sensing monitoring of canopy nitrogen in rice and wheat was reviewed (Zheng et al., 2022). The above studies on crop CND were aimed at crop with upright growth, and CND was the total amount of nitrogen per unit area (Li et al., 2019). For normal crop populations, the canopy structure was generally considered to be uniform. However, the canopy height of lodging maize population was lower than that of non-lodging maize population, and the leaves of lodging maize were tilted and covered with
each other, which led to great changes in canopy structure. The spatial distribution of maize leaves after lodging
changed in three-dimensional space, which led to the change of the spatial distribution of nitrogen. In this paper,
CND was defined as the total nitrogen content per unit volume to indicate the physiological status of maize
population after lodging. Lodging caused the change of maize population structure, which affected the
photosynthetic efficiency and nutrient transport efficiency of leaves (Berry & Spink, 2012). Therefore, it is
possible to estimate the nitrogen status of lodging maize by UAV-based hyperspectral technology. This study
tries to establish a quantitative relationship between hyperspectral information and CND of lodging maize
to achieve rapid and non-destructive estimation of nitrogen status of lodging maize.

There are hundreds of bands and high spectral resolution in hyperspectral data, which provides rich spectral
features to monitor lodging crop accurately and timely (Faiçal et al., 2017). However, hyperspectral data are
redundant and multicollinearity (Fang et al., 2019). In order to deal with the large amount of information in
hyperspectral data and to avoid the problems caused by a large amount of data, it is necessary to reduce the
computational complexity and alleviate the dimension of hyperspectral data. Dimensionality reduction is an
important preprocessing step in hyperspectral data analysis, which reduces computational complexity and
improves statistical morbidity by discarding redundant features. The feature selection method allows extraction
from the data with the most relevant information in the waveband, which retains the key information and physical
meaning of the hyperspectral data. In this paper, mutual information (MI), recursive feature elimination (RFE)
and SelectFromModel (SFM) algorithms were used to select the optimal waveband combination sensitive to
nitrogen to quantitatively realize the CND evaluation of lodging maize.

After lodging, canopy structure was changed greatly, roots and stalks of maize plant were damaged, which
resulted in the inhibition of photosynthesis and nitrogen transport (Khan et al., 2018; Robertson et al., 2017;
Shah et al., 2017). The images of lodging maize were obtained by UAV-based hyperspectral sensor in this study.
Through the fusion of spatial and spectral information, the CND of lodging maize was estimated and the severity
of lodging maize was evaluated on the plot scale. The objective of this study was to propose a CND estimation method of lodging maize based on the change of canopy spectrum after lodging. The specific objectives include: (i) analyzing the change of maize CND with different lodging types; (ii) comparing the wavebands screening performance of MI, RFE and SFM algorithms in CND estimation; (iii) using linear regression (LR), random forest regression (RFR) and AdaBoost regression (ABR) algorithms to construct the CND hyperspectral estimating model of lodging maize.

2 Material and Methods

2.1 Study area and Experiment design

2.1.1 General situation of the study area

Maize lodging control experiments were conducted in two sites (Fig. 1). The site 1 located in Zhaoxian Experimental Base (37°50′ N, 114°49′ E) of Agricultural and Forestry Research Institute of Shijiazhuang, Hebei Province, which is in North China Plain and with flat terrain and an average elevation of 45 m. It is the temperate continental monsoon climate. The average annual temperature, precipitation and annual sunshine are 12.30 °C, 464 mm and 2800 h. There is light loam cinnamon soil, the content of organic matter and total nitrogen are 14.30 g/kg, 1.20 g/kg. The site 2 located in Xinxiang Comprehensive Experimental Base (35°18′ N, 113°53′ E) of Chinese Academy of Agricultural Sciences in Xinxiang, Henan Province. It is the temperate continental monsoon climate. The average annual temperature, precipitation are 14.10 °C, 573.40 mm, the annual sunshine is 2407 h, and the average elevation is 73.20 m. There is fluvo-aquic soil, the content of organic matter and total nitrogen are 13.28 g/kg, 1.19 g/kg.
2.1.2 Experiment design

Maize lodging experiments were carried out in Zhao County of Hebei Province and Xinxiang County of Henan Province in 2020 and 2021 respectively. In 2020, the maize variety was *Zhengdan 958*, the sowing date was June 15, and the planting density was 69000 plants/ha. In 2021, the maize variety was *Junzhong 707*, the sowing date was June 9, and the planting density was 67500 plants/ha. Fertilization, spraying and others were carried out in the conventional way of field management. The maize lodging experiment was carried out by artificially controlled lodging in the field. The field was irrigated one day before lodging to make the field capacity saturated, which was convenient for maize lodging. Lodging treatment was carried out by experimenters with many years of field experience to ensure that lodging types were consistent with the real lodging situation.

When the plant is laid flat on the ground completely from above the root, the root lodging (RL) occurs in maize.
When the internodes of the stalks are bent and not broken, the stalk buckling (SB) occurs in maize. SB1 refers to the bending of the stalks at the 2-3 internodes above the root, while SB2 refers to the bending of the internodes near the maize ear. When the plant inclines to form an angle with the vertical line of the ground, the stalk inclining (SI) occurs. The tilt angle of the maize plant was set to 60° (SI60°), 45° (SI45°), and 30° (SI30°). There were six types of lodging maize and non-lodging (CK) maize were designed in the experiment (Fig. 2b), which were three repetitions. The plot size was 3.50 m × 3.50 m with an interval of 1.20 m (Fig. 2a). In addition, the lodging period was in the late growth stage, at tassel stage (VT) (2–3 d before silking) and milk stage (R3) (15–20 d after pollination).

Fig. 2 (a) Design of maize lodging experiment and (b) schematic diagram of lodging type of maize.
2.2 Data acquisition

2.2.1 UAV-based hyperspectral images

The UAV-based hyperspectral flight was carried out at the experimental base at 12-13:00 pm when the sky was clear and cloudless and the sun intensity was stable. At the same time, the ground agronomic traits of maize were collected. A six-rotor S900 electric UAV (Fig. 3a) (SZ DJI Technology Co., Ltd., Shenzhen, China) was equipped with a UHD185 hyperspectral imager (Cubert GmbH, Ulm, Germany) (Fig. 3b) to obtain the images of the experiment area (Fig. 3c). The length of single arm of the UAV was 386 mm, net weight was 4 kg, load mass was 6 kg, and flight duration was 15–20 min. The UHD185 imager (Fig. 4b) is a full-frame, non-scanning, real-time imaging airborne hyperspectral system that acquired full-color photos (.jpg) and hyperspectral cube photos (.cub) with 125 channels at 450–950 nm. The UHD185 was calibrated using a standard whiteboard before flight. The flight altitude of the UAV was set at 30 m, and the heading and lateral overlap were set at 80% and the image acquisition interval was 1 ms. The resolution of the original images obtained was 1000 pixels × 1000 pixels.
Fig. 3 (a) UAV-based platform, (b) UHD185 sensor and its parameters, (c) hyperspectral imagery of the experimental area.

2.2.2 Agronomic traits of maize

The acquisition of maize agronomic traits was synchronized with the acquisition of UAV-based hyperspectral images. The lodging date and collection date of maize agronomic traits at VT stage and R3 stage in 2020 and 2021 were shown in Table 1.

Canopy height (CH): In every plot, three representative plants were randomly selected to measure their height which was the vertical distance between the top of the plant and the ground. The average value of three maize plants was taken as the canopy height of the maize population.

Dry weight (DW): A representative whole plant was selected in every plot and brought back to the laboratory for stalk and leaf separation. The fresh weight of the leaves and stalks was weighed, then put them into the marked paper bags and placed them in the oven. The leaves and stalks of the maize were dried at 105 °C for 30
min (de-enzyme, to stop the reaction and the role of enzymes in the plant), and then dried at 85 °C (lasting for about 48 h) and weighed.

Leaf nitrogen content (LNC): The dried maize leaves were crushed and filtered through a 0.20 mm sieve. The dried sample was weighed as 0.15g (accurate to 0.0001g) and then digested with H₂SO₄-H₂O₂. The LNC of maize leaf was measured by flow analyzer (Auto Analyser-III, Bran Luebbe, Germany).

<table>
<thead>
<tr>
<th>Year</th>
<th>Growth stage</th>
<th>Lodging date</th>
<th>Data acquisition date</th>
</tr>
</thead>
<tbody>
<tr>
<td>2020</td>
<td>VT</td>
<td>Aug 20</td>
<td>Aug 27, Sep 6</td>
</tr>
<tr>
<td></td>
<td>R3</td>
<td>Sep 1</td>
<td>Sep 6</td>
</tr>
<tr>
<td>2021</td>
<td>VT</td>
<td>July 30</td>
<td>July 31, Aug 6, Aug 14, Aug 20, Aug 27</td>
</tr>
<tr>
<td></td>
<td>R3</td>
<td>Aug 13</td>
<td>Aug 14, Aug 20, Aug 27</td>
</tr>
</tbody>
</table>

2.2.3 Canopy nitrogen density (CND)

The canopy height of lodging maize was lower than that of non-lodging maize (Fig. 4). After lodging, the normal distribution of leaves and stalks was destroyed. The leaves and stalks of maize were tilted and covered each other after lodging, which greatly changed the physical structure and morphology of the plant (Fig.4). In three dimensions, it was ideally assumed that the maize population was a cuboid. After lodging, the shape of the cuboid changed into a parallelepiped, and the volume of the parallelepiped also changed (Fig.4). After lodging, not only the nitrogen content in the leaves changed, but also the spatial structure and distribution of nitrogen in the maize canopy changed accordingly. Therefore, the CND was constructed to indicate the changes of population structure and nutritional activity of maize after lodging in this paper. The CND was the sum of leaf nitrogen content per unit volume of maize population after lodging. The change of CND was closely related to canopy height and leaf nitrogen content. The calculation formula is as follows:
\[ \text{CND (mg/m}^3) = \frac{\text{LNC} \times \text{DW} \times \frac{n}{S \times CH}}{1} \]  

where, \( \text{LNC} \) was leaf nitrogen content (mg/g), \( \text{DW} \) was the leaf dry weight of sampled maize (one plant) (g), \( n \) is the number of maize plants in a plot, \( S \) is the plot area (m\(^2\)), \( CH \) was the canopy height (m).

![Fig. 4 Simulation diagram of lodging maize](image)

2.3 Methods and models

2.3.1 Waveband selection

The hyperspectral wavebands are many and narrow, and there is a strong correlation among multiple continuous wavebands, which leads to data redundancy and increases the complexity of the inversion model. There were 125 wavebands in UAV-based hyperspectral data of maize canopy collected in this study. The relationship between maize CND and canopy reflectance under lodging stress was established. The "dimension disaster" of hyperspectral data must be solved to improve the computational efficiency of CND model and reduce data redundancy. The main information was extracted from the original data by waveband (feature) selection method, which reduced the complexity of the inversion model without losing too much original spectral information. Therefore, mutual information (MI), recursive feature elimination (RFE) and SelectFromModel (SFM) algorithms were used to select the optimal spectral combination from maize canopy reflectance (Fig. 5).

MI is a filtering feature selection algorithm. The mode of MI is based on mutual information to select features, which is a measure of the interdependence between features and tags. MI detects both linear and non-linear relationships and is strongly related with the concept of entropy (Pascoal et al., 2017). Generally, the more MI
between dependent variables (targets) and independent variables (features), the higher the interdependence of
tags on this feature. On the contrary, the MI between dependent variables (targets) and independent variables
(.features) is 0, it means that the tags and feature are independent. MI finds features that have a potential
relationship with targets because the model learns the relationship between features related to targets. In this
paper, MI algorithm was used to select several wavebands from the canopy hyperspectrum of lodging maize,
which not only reduced the number of wavebands, but also reduced the complexity of maize CND model.
RFE is an encapsulated feature selection algorithm. The model was built repeatedly, the optimal features were
selected, the selected features were put aside, and the remaining features were repeated until all features were
traversed. Features ranking refers to the order of features are removed and the optimal combination of features
is selected according to their importance (Guyon et al., 2002). The optimal wavebands were selected by RFE
method, which not only ranked the wavebands according to their importance to maize CND, but also maintained
the interpretability of the original spectral data. In this paper, linear regression (LR), random forest regression
(RFR) and AdaBoost regression (ABR) algorithms were selected as the estimators in RFE.
SFM is an embedded feature selection algorithm. SFM is a meta-transformer, which is used to select features or
reduce dimensions according to the importance weight to improve the accuracy score of the estimator or its
performance on ultra-high-dimensional datasets (Krupinova et al., 2022). SFM is used with any estimator which
has the attribute of feature importance. The weight coefficient of features is obtained by using an estimator for
training, and then the importance of each feature is sorted according to the weight coefficient. In this paper, SFM
was used to automatically select the wavebands subset that were important to the maize CND in the training of
the estimator. The estimators used with SFM were linear regression (LR), random forest regression (RFR) and
AdaBoost regression (ABR) algorithms.

2.3.2 Regression algorithm

In this paper, the sensitive wavebands of maize CND were screened based on MI, RFE and SFM algorithms to
reduce the dimension of canopy hyperspectrum and remove redundancy data. Combined with LR, RFR and ABR algorithm, nine inversion models were constructed, namely MI-LR, MI-RFR, MI-ABR, RFE-LR, RFE-RFR, RFE-ABR, SFM-LR, SFM-RFR, SFM-ABR (Fig.5). The relationship between canopy hyperspectral information and maize CND was established to realize the accurate inversion of nitrogen of lodging maize.

Fig.5 Process diagram, feature selection and regression algorithm were combined to establish CND model.

2.3.3 Model evaluation

The maize CND model was evaluated by two indexes: stability and estimation ability. The stability was tested by the determination coefficient ($R^2$). The estimation ability was tested by the root mean square error ($RMSE$). The $R^2$ was closer to 1, the stability of the CND model was higher. The smaller the $RMSE$, the smaller the deviation between the measured CND value and the estimated CND value. The calculation formula of $R^2$ and $RMSE$ as follows (Chicco, Warrens, & Jurman, 2021):

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

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{n}} \quad \#(3)$$
where, $y_i$ and $\bar{y}$ were the measured CND and the mean of measured CND, $y'_i$ and $\bar{y}'$ were the estimated CND and the mean of estimated CND, $i$ was the $ith$ sample, $n$ was the number of samples.

3 Results

3.1 Maize CND with different lodging types

With the increase of the days after lodging (DAL), the maize CND of the same lodging type showed a downward trend at the VT stage and R3 stage (Fig. 6). It was because the nitrogen content of maize decreased with the increase of DAL. After lodging, the canopy structure collapsed and maize leaves and stalks blocked each other, resulting in the decrease of light interception ability of canopy and continued to influence the nutrient uptake of plants. In addition, with the recovery growth of maize plant, the canopy height became higher than that at the beginning of lodging. Lodging maize has the ability of self-recovery, which is shown in the upward growth of maize plants, that is, the increase of maize canopy height, resulting in the decrease of the maize CND.

Among the maize CND with different lodging types at VT stage, the CND of SB1 type was the highest on DAL 1, 14 and 28, and the CND of RL type was the highest on DAL 7 and 21 (Fig. 6a). Overall, the CND of RL and SB1 types were higher than other lodging types, the CK was the lowest. At R3 stage, the CND of SB1 type was the highest on DAL 1 and 7, the SI30° type was the lowest, and the CND of SB2 type was the highest on DAL 14 (Fig. 6b). The change of maize CND with different lodging types were consistent at VT stage and R3 stage.

In general, the CND of RL and SB types were higher than that of SI type. The main reason was that the canopy height of maize caused by RL and SB type was lower than that of SI type, and the lamination degree of maize leaves of RL and SB type was also higher than that of SI type.
The collected data was divided into training set and testing set, and the CND values of lodging maize at VT stage and R3 stage was calculated (Table 2). The statistical indexes mainly included maximum value (Max), minimum value (Min), mean value (M), standard deviation (SD) and coefficient of variation (CV). The difference of CND of lodging maize in the same growth stage was small between the two years. The CND of lodging maize at R3 stage was higher than the CND of lodging maize at VT stage.

Table 2 Statistical information of maize CND (mg/m³)

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Year</th>
<th>Growth stage</th>
<th>No.</th>
<th>Max</th>
<th>Min</th>
<th>M</th>
<th>SD</th>
<th>CV (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training set</td>
<td>2021</td>
<td>VT</td>
<td>105</td>
<td>231.21</td>
<td>50.60</td>
<td>120.94</td>
<td>36.95</td>
<td>30.55</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R3</td>
<td>63</td>
<td>347.64</td>
<td>61.33</td>
<td>130.84</td>
<td>53.59</td>
<td>40.96</td>
</tr>
<tr>
<td>Testing set</td>
<td>2020</td>
<td>VT</td>
<td>42</td>
<td>219.49</td>
<td>62.61</td>
<td>118.31</td>
<td>34.06</td>
<td>28.79</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R3</td>
<td>21</td>
<td>406.81</td>
<td>73.31</td>
<td>191.05</td>
<td>85.44</td>
<td>44.72</td>
</tr>
</tbody>
</table>

3.2 Sensitive wavebands selection

Generally, most of the sensitive wavebands selected by the nine CND inversion models were in the red edge and near infrared band at VT stage (Fig.7). At R3 stage, the sensitive wavebands selected by MI-LR, MI-RFR, MI-
ABR and RFE-LR algorithms were in the visible band (Fig. 8a, b, d, g), while most of the sensitive bands selected by RFE-RFR, RFE-ABR, SFM-LR, SFM-RFR and SFM-ABR algorithms were in the red edge and near infrared bands (Fig. 8c, e, f, h, i). The results showed that the relationship between CND and the reflectance of red edge and near infrared bands was important, while the red edge and near infrared spectrum was mainly affected by the canopy composition and nutritional activity of maize. The canopy spectrum of non-lodging maize mainly came from the upper regular leaves; after lodging, the canopy structure collapsed, and the canopy spectrum mainly came from maize stalks and cascading leaves. The multi-layer leaves of maize canopy provided multiple scattering and transmission, and the red edge and near infrared reflectance changed with the number of leaf layers in the canopy, so the maize canopy spectrum of different lodging types was more sensitive to red edge and near infrared bands.

Fig. 7 The selected wavebands at VT stage, (a), (d), (g) were MI algorithm, (b), (e), (h) were RFE algorithm,
Fig. 8 The selected wavebands at R3 stage, (a), (d), (g) were MI algorithm, (b), (e), (h) were RFE algorithm, (c), (f), (i) were SFM algorithm.

3.3 The results of maize CND models

The sensitive wavebands were screened from the canopy spectrum of maize population by MI, RFE, SFM combined with LR, RFR and ABR algorithms, and the CND model of lodging maize was constructed (Fig. 9).

For the same regression algorithm, the accuracy of maize CND model constructed by the sensitive waveband selected by RFE algorithm was higher than that of MI and SFM algorithms at VT stage and R3 stage. At VT stage, the $R^2$ of CND model constructed by RFE-LR was 27.58% and 10.02% higher than that of MI-LR and SFM-LR. The $R^2$ of CND model constructed by RFE-RFR was 1.57% and 0.83% higher than that of MI-RFR.
The $R^2$ of CND model constructed by RFE-ABR was 16.80% and 2.54% higher than that of MI-ABR and SFM-ABR (Fig. 9a). At R3 stage, the $R^2$ of CND model constructed by RFE-LR was 29.93% and 5.65% higher than that of MI-LR and SFM-LR. The $R^2$ of CND model constructed by RFE-RFR was 1.56% and 1.22% higher than that of MI-RFR and SFM-RFR. The $R^2$ of CND model constructed by RFE-ABR was 2.03% and 1.33% higher than that of MI-ABR and SFM-ABR (Fig. 9b).

For the same feature selection algorithm, the accuracy of maize CND model constructed by RFR algorithm was higher than that of LR and ABR algorithms at VT stage and R3 stage. The accuracy of CND model of LR algorithm was the lowest, and $R^2$ was the smallest and RMSE was the largest. For maize lodging at VT stage, the $R^2$ of CND model constructed by MI-RFR was 62.43% and 28.05% higher than that of MI-LR and MI-ABR. The $R^2$ of CND model constructed by RFE-RFR was 48.94% and 14.88% higher than that of RFE-LR and RFE-ABR. The $R^2$ of CND model constructed by SFM-RFR was 53.67% and 16.35% higher than that of SFM-LR and SFM-ABR. For maize lodging at R3 stage, the $R^2$ of CND model constructed by MI-RFR was 41.79% and 3.06% higher than that of MI-LR and MI-ABR. The $R^2$ of CND model constructed by RFE-RFR was 18.22% and 2.60% higher than that of RFE-LR and RFE-ABR. The $R^2$ of CND model constructed by SFM-RFR was 21.89% and 2.70% higher than that of SFM-LR and SFM-ABR.

The CND model constructed by RFE-RFR algorithm (for example, at R3 stage, training set: $R^2=0.93$, RMSE =13.83mg/m³; testing set: $R^2=0.92$, RMSE =24.42mg/m³) was more accurate and more stable than RFE-LR algorithm (R3 stage, training set: $R^2=0.76$, RMSE =26.07mg/m³; testing set: $R^2=0.88$, RMSE =29.43mg/m³) and RFE-ABR algorithm (R3 stage, training set: $R^2=0.91$, RMSE =16.15mg/m³; testing set: $R^2=0.98$, RMSE =11.60mg/m³). There was little difference in accuracy between the training set and the testing set of the CND model constructed by RFE-RFR.
Fig. 9 The $R^2$ and RMSE of maize CND models with different algorithms in training and testing sets at (a) VT stage and (b) R3 stage. The subscript tr represented the training set and subscript te represented the testing set. Through the above analysis, the accuracy of RFE-RFR algorithm was high and the prediction results were stable at VT stage and R3 stage. In order to validate the CND estimation model of lodging maize, the measured and estimated CND values were compared and analyzed at VT stage (Fig. 10a) and R3 stage (Fig. 10b). The measured and estimated CND values of training set and testing set were gathered near line 1:1, and the $R^2$ was more than 0.90, indicating that the estimation accuracy of maize CND under lodging stress was high, and the CND model was stable.

Fig. 10 Comparison between measured CND and estimated CND between the training set and testing set, the subscript tr represented the training set and subscript te represented the testing set.
3.4 Spatial distribution of the maize CND

According to the above results, the CND values of lodging maize at VT stage and R3 stage were calculated by RFE-RFR algorithm, and the spatial distribution of maize CND with different lodging types on UAV-based hyperspectral images was shown in Fig. 11. Even in the same lodging type, the spatial variability of nitrogen status in maize canopy existed in every plot. The changes of CND in each plot were explained by the growth status of individual maize, such as plant height and nitrogen content of single maize. The lodging types of maize at VT stage and R3 stage were RL, SB1, SB2, SI60°, SI45°, SI30° and CK from left to right. It was obvious that the maize CND of RL type was the highest, followed by SB1 type, and the CK was the lowest. Among the SI types, the CND of SI60° was the highest, followed by SI45°. The results showed that the more serious the lodging stress, the higher the maize CND. Because CH was the main factor affecting CND value, the more serious the lodging stress, the lower the maize CH and the higher the total nitrogen content per unit volume. The results further showed that based on the hyperspectral data of UAV, the RFE-RFR algorithm could improve the accuracy of CND estimation of lodging maize, and it had a good potential in the estimation of CND of lodging maize.

Fig. 11 Maize CND map estimated at VT and R3 stage using UAV-based hyperspectral data on Sep 6, 2020
4 Discussion

4.1 UAV-based hyperspectral estimation of lodging maize CND

Timely and accurate assessment of crop lodging severity is of great practical significance to post-disaster production management and rapid settlement of insurance claims (Kendall et al., 2017; Tian et al., 2018). The hyperspectral sensor has the advantages of small, narrow band, high spectral resolution, which distinguishes the subtle differences in the spectra of ground objects (Faiçal et al., 2017). Imaging hyperspectral sensor combines image technology with spectral technology, which reflects not only the two-dimensional image information of the target, but also the spectral dimension information (Deng et al., 2018; Gu et al., 2021). The canopy spectrum of crops was a comprehensive reflection of population structure and nutritional activity, and the CND constructed in this paper reflected the information of both. Considering the change of nitrogen content in volume, the CND was introduced to indicate the nitrogen changes of lodging maize, which reflected the comprehensive changes of nitrogen in the canopy horizontal and vertical direction. The nitrogen content of the maize after lodging directly influenced the nutrients of maize plant, and it was also related to the extent of recovery after lodging and the final yield of maize. Therefore, the CND combined with CH not only reflected the nitrogen status of maize population, but also indicated the growth and late recovery of lodging maize.

Strong winds and continuous precipitation brought by extreme weather easily caused crops to tilt and even lodging. Maize lodging generally occurs in the VT stage and R3 stage, and there are mainly three types of root lodging (RL), stalk buckling (SB) and stalk inclining (SI). In this paper, different lodging types of maize were simulated by manual control method (Fig. 2b, Fig. 3c). The most direct apparent changes of maize under lodging stress were the decrease of vertical height of canopy and the disorder of directional arrangement of leaves (Fig. 3c), which led to great changes in canopy structure (Fig. 4). Lodging led to more leaves covering each other, which directly reduced the light area of leaves and influenced the photosynthesis of crops. The photosynthesis of crops was inhibited and the physiological of leaves was weakened, which led to the decrease of nitrogen.
content in leaves and influenced the growth status of crops (Berry & Spink, 2012). The spatial resolution of the UAV-based hyperspectral images obtained was centimeter in this paper, which completely distinguished the stalk and leaf of the maize plant (Fig. 3c). The spectral resolution of the UHD185 hyperspectral sensor was 4 nm, which provided rich spectral information and could better identify the detailed information (nitrogen) of maize plants. In the application of hyperspectral data in agricultural remote sensing, crop growth status and yield can be easily predicted by accurately obtaining agronomic traits such as moisture and nitrogen content (Ma et al., 2022; Raj et al., 2021; S. Wang et al., 2021). Lodging leads to changes in canopy structure, damage to roots and stalks, and decrease in physiological activity of crops. A CND model was constructed through the spatial analysis of crop lodging, which described not only the physiological features of crop population, but also the structural features of lodging population. The change of maize CND, the relationship between CND and canopy spectrum with different lodging types were analyzed. With the increase of DAL, the maize CND with the same lodging type decreased in both VT and R3 stages (Fig. 6). On one hand, the nitrogen content of maize decreased with the increase of DAL. On the other hand, with the recovery growth of maize plant, the increase of maize canopy height, which decreased the maize CND. In general, at VT stage and R3 stage, the maize CND of RL and SB types were higher than that of SI type, and the CK was the lowest (Fig. 6a, b). The reason was that the CH of RL and SB type was lower than that of SI type, and the lamination degree of maize leaves of RL and SB type was also higher than that of SI type. UAV-based hyperspectral system has the advantages of large amount of spectral information and high accuracy of data acquisition, which provides an accurate database and a simple and efficient solution for large-area crop lodging stress monitoring (Ampatzidis et al., 2020). The results showed that it was feasible to estimate the maize CND under lodging stress using UAV-based hyperspectral system. This paper improved and deepened the theory of crop lodging monitoring by hyperspectral remote sensing, and laid a foundation for timely and accurate monitoring of maize lodging severity.
4.2 Performance of lodging maize CND model

The UAV-based hyperspectral sensors obtained 125 narrow wavebands in this study, which brought challenges to data mining in establishing a reliable estimation model (Khan et al., 2015). Since the CND estimation model constructed by all wavebands was prone to over-fitting, and the prediction ability and interpretability of the model were poor (Thor et al., 2017). Therefore, mining effective spectral information was very important to estimate the CND of lodging maize. Dimensionality reduction of hyperspectral data was widely regarded as an effective means to improve the accuracy and interpretability of model estimation. The wavebands sensitive to lodging were selected according to certain criteria to retain the features of the original hyperspectrum (Fu et al., 2022). For specific applications, due to the different ranking criteria and search strategies of waveband sensitivity, the wavebands selected by different feature selection methods were not consistent. In this study, MI, RFE and SFM algorithms were compared to select the sensitive wavebands to CND of lodging maize. Based on the above feature selection algorithm, combined with regression algorithms, nine joint algorithms were proposed to reduce the dimensionality of hyperspectral data and accurately evaluate the CND of lodging maize.

Crop CND was estimated by multiple methods based on hyperspectral data from many aspects. Based on the lookup table (LUT) method to estimate the CND of winter wheat with multiple growth stages, the highest accuracy of the model was 0.77 (Li et al., 2019). But there were many input parameters of the model, and the calculation time and LUT size were significantly increased. The highest accuracy of estimating winter wheat CND by green normalized difference vegetation index (GNDVI) was 0.83 (Zhao et al., 2012). However, the spectral index consisted of special wavebands, which might ignore the wavebands that were more sensitive to crop CND. The stable CND model of cotton at flowering and boll stage was established by multiple linear regression \( R^2 = 0.53-0.67 \) (Ma et al., 2022), but the accuracy of the model needed to be improved. For the CND of lodging maize in this study, the comparison of feature selection algorithms showed that RFE algorithm was better than MI and SFM algorithms. The principle of RFE is to fit a given regression model, sort the wavebands
according to their importance, discard the least important wavebands, and re-fit the regression model for the
remaining wavebands. This process is repeated until a certain number of waveband subsets are retained. In three
regression algorithms, the performance of RFR was better than LR and ABR. RFR belongs to the Bagging class
of ensemble algorithms, which forms a strong model by training several weak models. The prediction results of
RFR are obtained by averaging the results of all internal binary decision in order to use several different sub-
models to increase the robustness and stability of the final model (that is, to reduce the variance). The accurate
evaluation of lodging maize CND was realized through data reduction of RFE-RFR algorithm, and achieved
good results (Fig. 9, Fig. 10). Its implementation was simple and efficient. Using only 10 wavebands, the
estimation accuracy of lodging maize CND reached more than 0.90 (Fig. 10). The selected sensitive wavebands
provided a reference for the multispectral sensors, and the large-scale monitoring of maize lodging could be
realized in the future.

Although some good results were achieved in this study, there were still other experiments and analyses to be
supplemented and improved. The lodging experiment was only conducted on maize crop, without considering
the lodging stress of other crops. In the future research, lodging experiments were conducted on a variety of
crops to analyze the changes of physiological status of different crops, in order to verify the universality of the
CND estimation model.

5 Conclusion

Focusing on the key issue of monitoring crop lodging by remote sensing, for the purpose of CND estimation of
maize with different lodging types, the internal relationship between canopy hyperspectral features and nitrogen
of maize was explored by UAV-based hyperspectral technology. The higher the lodging severity of maize, the
greater the maize CND. With the longer the days after lodging, the maize CND tended to decrease. Combining
MI, RFE and SFM (feature selection algorithm) with LR, RFR and ABR (regression algorithm), the
hyperspectral wavebands sensitive to maize CND were selected and the CND estimation model was constructed.
For the feature selection algorithm, RFE algorithm was better than MI and SFM algorithm. The $R^2$ of the CND model constructed by RFE-RFR was 1.57% and 0.83% higher than that of MI-RFR and SFM-RFR at VT stage.

The $R^2$ of CND model constructed by RFE-RFR was 1.56% and 1.22% higher than that of MI-RFR and SFM-RFR at R3 stage. For the regression algorithm, RFR algorithm was better than LR and ABR algorithm, and there were high accuracy and good stability of the CND model constructed by RFR algorithm. The $R^2$ of the CND model constructed by RFE-RFR was 48.94% and 14.88% higher than that of RFE-LR and RFE-ABR at VT stage. The $R^2$ of the CND model constructed by RFE-RFR was 18.22% and 2.60% higher than that of RFE-LR and RFE-ABR at R3 stage. The accuracy of the CND model constructed by RFE-RFR algorithm in the testing set ($R^2=0.92$, $RMSE=9.84\text{mg/m}^3$ at VT stage; $R^2=0.92$, $RMSE=24.42\text{mg/m}^3$ at R3 stage) was close to that of the training set ($R^2=0.92$, $RMSE=10.60\text{mg/m}^3$ at VT stage; $R^2=0.93$, $RMSE=13.83\text{mg/m}^3$ at R3 stage), indicating that the CND model was stable. Therefore, the combination of feature selection and regression algorithm to reduce the dimension of hyperspectral data could achieve good performance in CND estimation of lodging maize.

**Funding**

This work was supported by the National Natural Science Foundation of China [42271319] and the Key Research and Development Projects in Hebei Province [31327001D].

**References**


This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4364605


This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4364605