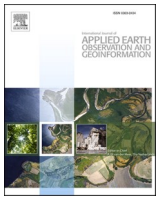
International Journal of Applied Earth Observations and Geoinformation 104 (2021) 102584

Contents lists available at ScienceDirect

International Journal of Applied Earth Observations and Geoinformation

journal homepage: www.elsevier.com/locate/jag



Leaf nitrogen content estimation using top-of-canopy airborne hyperspectral data

Rahul Raj [a,b,c,*], Jeffrey P. Walker [b], Rohit Pingale [c], Balaji Naik Banoth [d], Adinarayana Jagarlapudi [e]

[a] IITB-Monash Research Academy, India
[b] Monash University, Melbourne, Australia
[c] Indian Institute of Technology, Bombay, India
[d] Professor Jayashankar Telangana State Agricultural University, India
[e] Indian Institute of Technology, Bombay, India

ARTICLE INFO

Keywords:
Hyperspectral sensing
Leaf nitrogen content
Drone
Precision agriculture
CHNS

ABSTRACT

Remote estimation of leaf nitrogen content is a critical requirement for precision farm management. Precise knowledge of nitrogen distribution in the crop enables farmers to decide the fertilisation amount required at specific locations on the farm. Importantly, nitrogen related molecules in plants are transported using water molecules, and water molecules surround the amide bonds (a plant protein created from nitrogen). Consequently, the nitrogen in various crop parts loses its activity in the absence of sufficient water molecules. The association of water molecules around plant proteins makes the optical remote estimation of plant nitrogen challenging as nitrogen and water molecules simultaneously affect the reflectance data. Moreover, the coarse spatial resolution of satellite data and sparse canopy coverage at early growth stages of the crop make it challenging to estimate leaf-level nitrogen contents. Accordingly, this research developed a leaf nitrogen content estimation model using drone-based top-of-canopy 400–1000 nm pure pixel hyperspectral images collected from a maize research farm treated with different water and nitrogen levels. Leaf level spectral signatures were also collected using a field spectroradiometer and used to identify indices more sensitive to nitrogen than water. The leaves were also destructively sampled for obtaining ground truth leaf water and nitrogen content. Red-edge region bands of electromagnetic spectra were identified to be sensitive to leaf nitrogen content. A synthetic data was created using maximum and minimum values of these indices and crop growth stage information, which was further used for training a gradient-boosting machine model to estimate leaf nitrogen content from drone-based hyperspectral images. The estimated leaf nitrogen content values from drone observations were critically analysed with respect to leaf water content values. For water-stressed areas, the model gave an R^2 and RMSE of 0.63 and 2.74 mg/g, respectively. However, the model did not perform adequately for well irrigated areas, having an R^2 and RMSE of 0.26 and 4.54 mg/g, respectively.

1. Introduction

Plants can uptake soil nitrogen as nitrate and ammonia and utilise it for plant growth (Ohyama, 2010). Thus for agriculture purposes, fertiliser in the form of ammonia (NH_3) or ammonium nitrate (NH_4NO_3) is supplied to the farm to provide sufficient nitrogen to the soil (Mason, 1977; Craig and Wollum, 1982; Gezgin and Bayrakll, 1995; van Grinsven et al., 2015). Moreover, nitrogen is a mobile macronutrient in plants that changes its content temporally (Kutman et al., 2011; Masclaux-Daubresse et al. 2010), tending to move from old leaves to new/fresh leaves in order to increase their biomass (Masclaux-Daubresse et al. 2010). Nitrogen is also a critical nutrient for biomass creation in grain-producing plants, with plant nitrogen concentration decreasing as dry biomass of the canopy increases (Chen et al., 2010). Interestingly, this trend is similar to the change in leaf water content over progressive crop growth stages (Raj et al., 2021).

Approximately, 30–50% of nitrogen in green leaves is in the form of D-ribulose 1–5-diphosphate carboxylase (RuBisCO) (Kokaly, 2001),

* Corresponding author at: Patna, Bihar, India.
E-mail addresses: Rahul.Rahulraj@monash.edu, rahul.raj@iitbombay.org (R. Raj), Jeff.Walker@monash.edu (J.P. Walker), Rohitpingale103@iitb.ac.in (R. Pingale), Adi@iitb.ac.in (A. Jagarlapudi).

https://doi.org/10.1016/j.jag.2021.102584
Received 3 September 2021; Received in revised form 9 October 2021; Accepted 10 October 2021

being the key photosynthetic enzyme/protein in green leaves (Gutteridge and Gatenby, 1995; Andersson and Backlund, 2008). Other functional involvement of nitrogen in plants is in the form of different proteins (Schrader, 1984; Howitt and Udvardi, 2000). Maintaining the nitrogen level in plants above a critical value is very important because nitrogen is used to form biomass with the progress in growth stages (Leghari et al., 2016). This critical value represents the minimum nitrogen concentration to be maintained in the plant so as to maximize crop growth (Blumenthal et al., 2008). However, using a high amount of nitrogen in the farm is not only expensive but leads to water pollution, as the irrigated water is wasted through runoff and leaching (Knox et al., 2012; Elrashidi et al., 2005), meaning that optimal use of fertilisers should be applied on the farm.

Interestingly, water is also a major contributor to the protein's three-dimensional structure (Franks, 1988), and protein controls the structure of its surrounding water, known as the hydration of protein (Franks, 1988). This protein hydration is critical, as its biological activity reduces in the absence of hydrating water (Chaplin, 2006). Consequently, the association of water and protein makes it difficult to estimate the precise amount of nitrogen in a crop and more so remotely.

The remote estimation of canopy/leaf nitrogen content essentially depends on the vibrational properties of the amide bonds of the plant proteins (Kokaly, 2001; Damodaran, 2008). Nitrogen shows higher light absorption characteristics in ultraviolet bands (Ogawa et al., 1964) and short-wave infrared bands (Widlowski et al., 2015). The various stretching, bending, and torsion in the different amide bonds present in proteins have shown absorption characteristics at wavelengths longer than 3000 nm (Haris and Chapman, 1994). Table 1 shows the light absorption wavelengths for amide bonds in proteins (Haris and Chapman, 1994). Moreover, researchers have also found a few wavelengths in the 400–2500 nm region (515, 520, 525, 550, 575, 743, 1116, 2173, and 2359 nm) of the electromagnetic spectrum correlated to nitrogen content (Thenkabail et al., 2016). However, apart from 2173 and 2359 nm wavelengths (Kokaly, 2001), none of the other bands have known causation (small absorption features) for this correlation.

There is no known nitrogen/protein absorption wavelength available between the 400–1000 nm region of the electromagnetic spectrum, which has a light absorption sensitivity to the nitrogen content in the leaves. Thus, the only logical way that the 400–1000 nm reflectance data can be used for nitrogen estimation is through empirical proxy relationships with parameters such as greenness (Hansen and Schjoerring, 2003) and chlorophyll of the crop (Curran et al., 1992; Haboudane et al., 2002), or some empirical relations between vegetation indices and vegetation nitrogen content (Reyniers et al., 2006; Hansen and Schjoerring, 2003; Chen et al. 2010).

Some studies have shown a correlation between red-edge-based indices and nitrogen content of the crop. The double peak canopy nitrogen index (DCNI) is an example of a red-edge-based index where 720, 700, and 670 nm wavelengths were used (Chen et al. 2010). The DCNI formula, along with other indices, is shown in Table 2. Similarly, Yao et al. (2010) used a spectroradiometer with 400–2500 nm range to collect canopy level reflectance from a wheat crop. Yao et al. (2010) studied various indices, and the involvement of 720, 725, and 736 nm wavelengths showed the usefulness of the red-edge region for nitrogen estimation. In another study, Feng et al. (2008) also used red-edge region wavelengths to estimate wheat crop leaf nitrogen. Feng et al. (2008) recommended spectroradiometer-based reflectance signatures to create red-edge position index (Cho and Skidmore, 2006) and mND705 (normalised difference index at 705 nm) (Sims and Gamon, 2002) indices for leaf nitrogen content estimation.

Stroppiana et al. (2006) experimented with data from a index, called the optimal normalized difference index (NDI_{opt}), to be more sensitive than the traditional normalized difference vegetation index (NDVI) to changes in the plant nitrogen concentration while also being less affected by crop biophysical properties. Du et al. (2016) estimated nitrogen content in rice leaves using 32-band active hyperspectral sensing. The central wavelength of these bands was between 500 and 910 nm, with four wavelengths in the red-edge region. The authors used a total of 32 bands for making a machine learning model, with a maximum R^2 of 0.75 obtained. In another study, Fan et al. (2019) used canopy level spectroradiometer-based spectral signatures from a maize crop to estimate leaf nitrogen content. Partial least square regression analysis was carried out on the collected data resulting in R^2 of 0.77. Similarly, Tan et al. (2018) collected temporal spectroradiometer-based hyperspectral reflectance spectra from an experimental wheat crop canopy. The study was focused on a statistical analysis of available methods of leaf nitrogen estimation, with an index named NREAI found to give the highest R^2 of 0.97, with chlorophyll used as a proxy indicator for estimating leaf nitrogen. Tian et al. (2014) also used spectroradiometer-based canopy (including background) reflectance spectra of a rice crop and found the simple ratio of 553 and 537 nm bands more reliable for leaf nitrogen content estimation under various cultivation conditions.

In a drone-based study, Liu et al. (2017) used spectral signatures (450–950 nm) of a wheat crop at different growth stages. Field-based spectroradiometer readings were also collected simultaneously. A few bands, including from the red-edge region, were selected using correlation analysis. The Back Propagation (BP) neural network and multifactor statistical regression method were implemented on the selected bands for the leaf nitrogen content model training and evaluation. The model gave an R^2 between 0.85 and 0.96 for different growth stages of the crop. Similarly, Liang et al. (2018) used ground-based spectroradiometer and aircraft-based hyperspectral data from an experimental winter wheat farm. However, this was mixed-pixel data as the spatial resolution of the aircraft data was 3 m. First-derivative indices at 520 nm and 715 nm were found to produce an R^2 of 0.75. The authors recommended to use less than 30 nm bandwidth for leaf nitrogen content estimation. Tian et al. (2011) created two-band and three-band hyperspectral indices for estimating paddy canopy-level leaf nitrogen concentration. The data was collected from ground, airborne (AVIRIS), and spaceborne (Hyperion satellite) platforms. The newly identified two-band index - $\frac{R533}{R565}$, and three-band index $\frac{R705}{R717+R491}$ resulted in an R^2 less than 0.76. The use of the green-colour wavelength region shows that the higher correlation was due to the difference in green colour among different nitrogen treatment plots.

In very few studies, leaf water and nitrogen contents were studied together using hyperspectral data. Strachan et al. (2002) used canopy-level 350–1000 nm hyperspectral data to demonstrate the maize development under nitrogen and water stress conditions. Canonical discriminate analysis was used to classify different nitrogen rate canopies. The authors suggested to carry out more research to understand the dynamics of nitrogen estimation under various water stress conditions. In another study, Feng et al. (2016) developed a water resistance nitrogen index (WRNI) and tested it on a winter wheat crop. Canopy level reflectance spectra were collected (~400–1000 nm), and plants were destructively sampled for estimating leaf water content and leaf nitrogen content. WRNI was calculated using the ratio of normalised

Table 1
Infrared wavelength absorption of amide bonds in protein (adapted from Haris and Chapman, 1994).

Wavelength (nm)	Chemical bond origin
3030	—NH stretching
3225	N—H stretching
5917–6250	C═O stretching, C-N stretching, N—H bending
6349–6757	C-N stretching, N—H bending
7686–8137	C-N stretching, C═O stretching, N—H bending, O═C-N bending
13030–16000	O═C-N bending
12500–15630	—NH bending
16500–18620	C═O bending
50,000	C—N torsion

Table 2
Indices found to be used for leaf nitrogen content estimation in literature.

Index Formula	short form	Full form	Reference
$\frac{(1+0.45)(R_{800}^2+1)}{R_{670}+0.45}$	Vi_{opt}	Optimal vegetation index	Reyniers et al., 2006
$\frac{R_{573}-R_{440}}{R_{573}+R_{440}}$	$NDVI_{g-b}$	Green-blue normalised difference vegetation index	Hansen and Schjoerring, 2003
$\frac{\frac{R_{720}-R_{700}}{R_{700}-R_{670}}}{R_{720}-R_{670}+0.03}$	$DCNI$	Double peak canopy nitrogen index	Chen et al. 2010
$\frac{R_{450}}{R_{550}}$	$BGI2$	Blue Green Index 2	Chen et al. 2010
$\frac{MCARI}{MTVI2}$	Combined Index	$MCARI = [(R_{700}-R_{670}-0.2)(R_{700}-R_{550})]\left(\frac{R_{700}}{R_{670}}\right)$ $MTVI2 = 1.5^{*}\frac{(1.2^{*}(R_{800}-R_{550})-2.5^{*}(R_{670}-R_{550}))}{\sqrt{(2^{*}R_{880}+1)^2-(6^{*}R_{800}-5^{*}R_{670}^{\frac{1}{2}})-0.5}}$	Eitel et al., 2007Daughtry et al., 2000Haboudane et al., 2004
$3^{*}\left[\begin{matrix}(R_{700}-R_{670}-0.2)\\ ^{*}(R_{700}-R_{550})\end{matrix}\right]\left(\frac{R_{700}}{R_{670}}\right)$	TCARI	Transformed chlorophyll absorption in reflectance index	Haboudane et al., 2002

difference red-edge (Fitzgerald et al., 2006) and a floating-position water band index (Strachan et al., 2002). The WRNI gave R^2 between 0.79 and 0.85. Corti et al. (2017) conducted a pot experiment under different water and nitrogen treatments. Spectroradiometer-based 400–1000 nm spectral signatures were collected and used to estimate crop biophysical and biochemical properties, including leaf water and nitrogen content. All of the crop parameters were estimated using various indices. However, no analysis was presented to understand the effect of water and nitrogen variables on their estimation.

From the above literature, it was found that very little work is available on pure-pixel analysis for remote estimation of leaf nitrogen content with respect to the change in leaf water content. Considering the association of water molecules with plant protein, the bands or indices for estimating leaf nitrogen content from visible to near-infrared region (400–1000 nm) are expected to be affected by the water molecules around plant proteins. None of the available studies have checked if their models are resulting in high correlation with leaf nitrogen simply due to their sensitivity to vegetation water content. Moreover, the indices available in the literature were created and tested for mixed-pixel data, and there has been no equivalent study with pure pixels. Accordingly, narrow-band pure-pixel hyperspectral data has been used in this study to estimate leaf nitrogen content, and the results analysed with respect to change in leaf water content.

2. Materials and methods

In this study, leaf-level narrow-band and pure pixel spectral signatures were collected using a drone-based hyperspectral imager and hand-held spectroradiometer. Both data were recorded around the same time in the 400–1000 nm range. The leaves from which spectroradiometer-based spectral signatures were collected were subsequently plucked and the leaf total nitrogen content were obtained using the Dumas method based CHNS instrumental analyser (Dhaliwal et al., 2014). The collected data were then used to identify the bands and indices more sensitive to change in leaf nitrogen content (LNC) than leaf water content (LWC). The maximum and minimum values of indices, growth stage information, were then used to create synthetic linear data to train a gradient boosting machine learning algorithm (GBM). Drone-based hyperspectral data were used to evaluate the model, and the results critically analysed with respect to LWC information. The framework of this research is shown in Fig. 1.

2.1. Site description and data acquisition

This research was conducted during the post-monsoon (*Rabi*) season of 2018–19, in a research farm located at 17°19′28″N and 78°23′56E, as shown in Fig. 2. A 'Cargil 900 m gold' maize crop variety was used for this study. Three different irrigation levels and nitrogen treatments were used, combinations of which resulted in nine unique treatment plots.

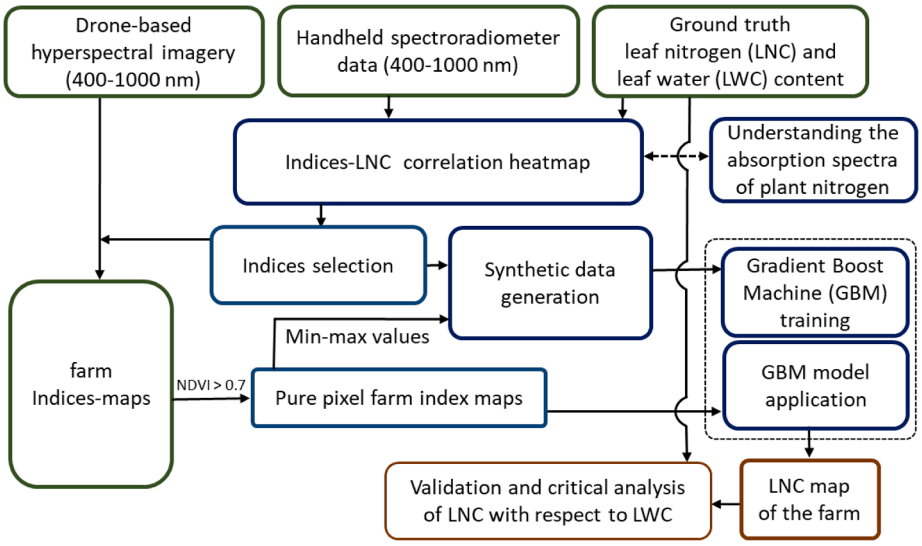


Fig. 1. The structure of the leaf nitrogen content estimation model used in this research.

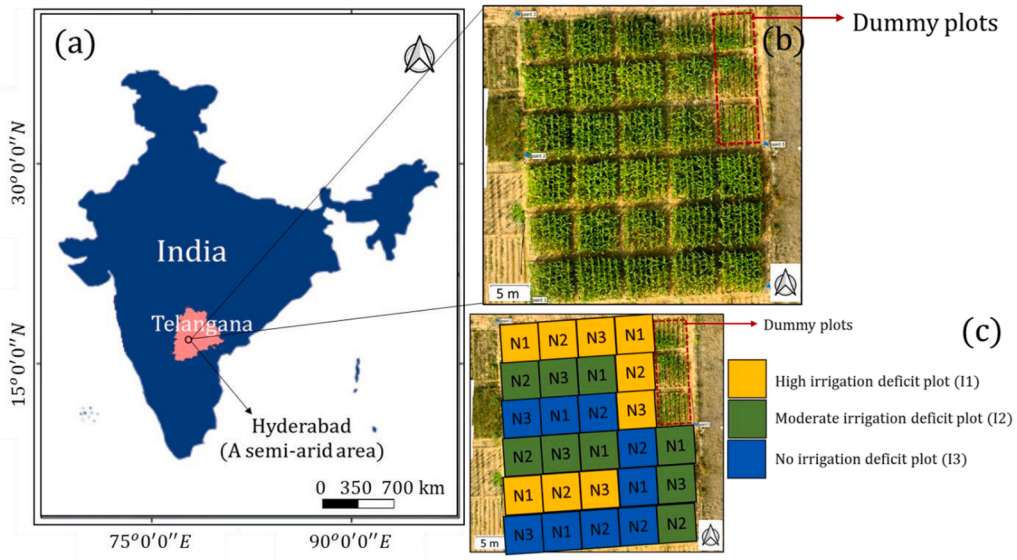


Fig. 2. (a) Map of India with highlighted study area; (b) The research farm top-of-canopy image; (c) The layout of the research farm indicating the treatments given to each plot. I1, I2, I3 and N1, N2, N3 represent low, moderate and high doses of water and nitrogen treatments, respectively.

These nine treatments were replicated thrice to create a total of 27 plots. Fig. 2(c) shows the layout of these treatments, where I1, I2, and I3 represent low, moderate, and high irrigation plots, while N1, N2, and N3 represent low, moderate, and high doses of nitrogen treatments, respectively. A detailed explanation of the site can be found in Raj et al. (2021).

Drone-based hyperspectral data were collected using a Bayspec OCI-FHR (400–1000 nm; 241 bands; 2.4 nm FWHM) push-broom hyperspectral imager. A DJI Matrice 600 hexacopter was used to fly the hyperspectral camera at a height of 50 m above the research farm. As the imager was of the line-scanner (push-broom) type with frame rate of 50 fps, the speed of the drone was fixed at 2 m/sec. As shown in Fig. 3, the imager was fitted inside a gimbal attached to a DJI Matrice 600 hexacopter drone. The gimbal is one of the crucial parts of the system, keeping the imager stable while the drone is flying, minimising vibrations that can affect *ortho*-distortions, artifacts, missing spatial lines, and out-of-focus issues (Oehlschläger et al., 2018).

The data was collected at six-leaf, pre-tasselling, tasselling, silking, dough, and maturity crop growth stages on sunny days with clear sky conditions. Around the same time as the drone-based data collection, an SVC GER1500 spectroradiometer (400–1000 nm; 381 bands; 2.1 nm FWHM) was used to collect the leaf spectra of a single leaf from each treatment plot. To minimise the angle effect on the reflectance spectra, spectroradiometer was operated as per the suggestions given by He et al. (2016). Destructive sampling of the same leaf was then carried out, and the leaves were packed into pre-weighed airtight zip bags. Analytical balance with 10th of a milligram accuracy was used for weighing purposes. Later, individual leaf-filled zip-bags were again weighed, and the subtraction of before and after weights of zip-bags used as the fresh weight of the leaves. The leaves were then cut into small pieces, put into a pre-weighed aluminum foil vessel, and weighed again. The leaf-filled aluminum foil vessel was then put into an oven at 60 °C for around 72 h to get completely dry samples. The dried samples in the aluminum foil vessels were weighed again, and the dry weight of the leaves obtained. The fresh weight and dry weight of the leaf samples were used to obtain the leaf water content using

$$LWC = \frac{Freshweight - Dryweight}{Freshweight} \quad (1)$$

The dry leaves were then ground separately and passed through a sieve to get finely powdered samples. The powdered samples were then fed into a ThermoFisher Scientific FlashSmart Elemental Analyzer (CHNS instrument) to obtain the nitrogen content in each leaf sample. A total of 272 leaf samples were analysed using the CHNS instrument. The CHNS analyser works on the Dumas method in which flash combustion is done for instantaneous oxidation of the sample. Later, a chromatographic column and thermal conductivity detector were used to separate and detect the combustion products. The CHNS analyser was calibrated using k-factor analysis of 2.5-Bis (5-*tert*-butyl-benzoxazol-2-yl) thiophene (also known as BBOT). Leaf powder samples in a quantity of 3–4 mg were used for the analysis. Each leaf sample was replicated twice, and every 50th sample was replicated thrice in the CHNS analyser to check the consistency of the results. Average values of these replications were assigned as ground-truth leaf nitrogen content. The CHNS-based LNC was considered actual LNC. Fig. 4 shows the step-by-step approach used, starting from leaf collection to making its fine powder.

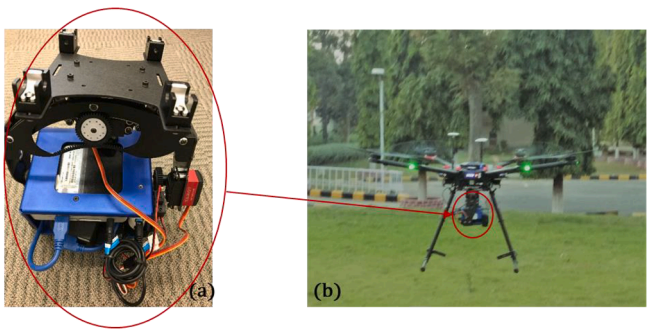


Fig. 3. (a) Hyperspectral Imager installed inside a gimbal and (b) the complete gimble setup installed on a DJI Matrice 600 hexacopter drone.

2.2. Preprocessing of hyperspectral data

The collected spectroradiometer data had a high-frequency noise associated with it. A Savitsky-Golay filter was used to smooth the data by removing high-frequency noise. Fig. 5 shows the raw and smooth spectra of a leaf. The hyperspectral sensor on the UAV was a line scanner type from Bayspec, requiring the raw data to be preprocessed in their cube creator software. The software does the radiometric correction and orthorectification and gives a hyperspectral cube. However, the pixel radiometric values contained high-frequency noise that needed to be corrected using the Savitsky-Golay filter, similar to the

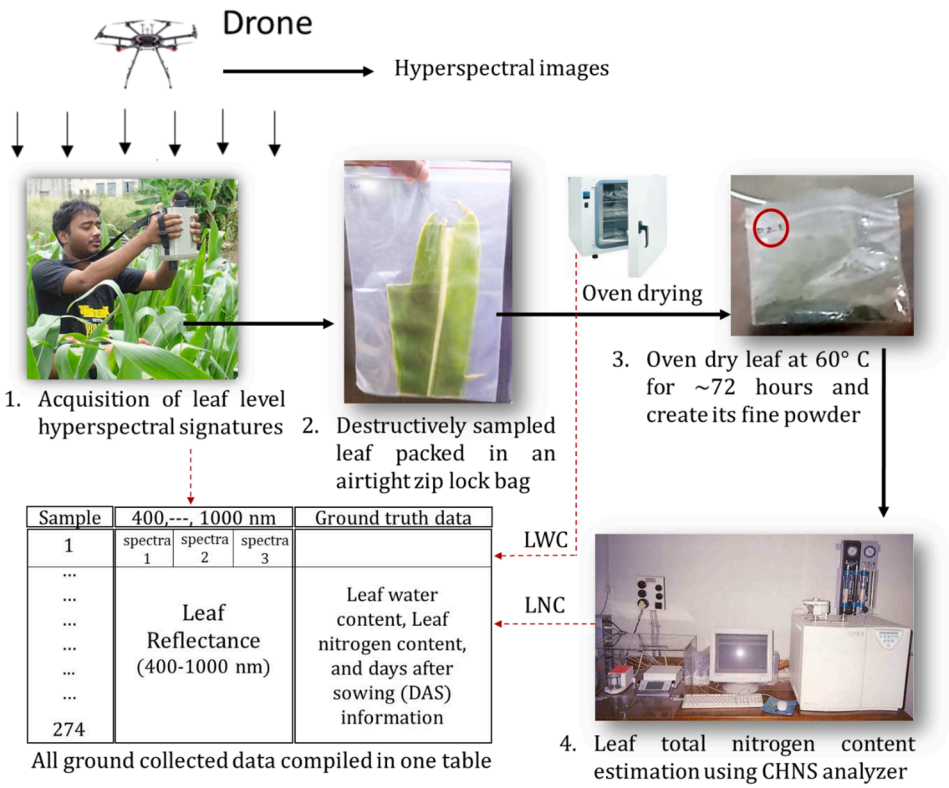


Fig. 4. Process indicating spectroradiometer and drone-based leaf-level hyperspectral data collection. Further, the leaves were destructively sampled, oven-dried, ground, and used in the CHNS analyser. The obtained leaf water (LWC) and nitrogen (LNC) content were coupled with respective leaf hyperspedctral signatures and saved as csv files.

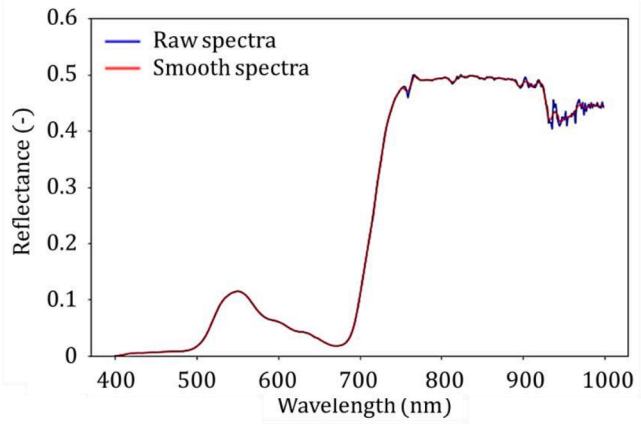


Fig. 5. Raw spectra having high-frequency noise (majorly after 900 nm), and its smooth version after applying a Savitzky -Golay filter.

spectroradiometer data correction. The hyperspectral data also had some missing lines in some of the stitched tiles. These missing lines were corrected by replacing the missing values with band-wise average values of adjacent pixels.

2.3. Index creation for leaf nitrogen content

The indices identified in this research were created using the narrow-band, pure pixel leaf-level hyperspectral signatures obtained from the spectroradiometer data. The 381 bands of the sapectroradiometer were used to create 72,390 ($^{381}C_2$) unique two-band normalised difference indices using

$$Normalised\,difference\,index\,(NDI) = \frac{Reflectance\,at\,band\,i - Reflectance\,at\,band\,j}{Reflectance\,at\,band\,i + Reflectance\,at\,band\,j} \quad (2)$$

where *i*and *j* are spectroradiometer bands. The correlation coefficient between each of the 72,390 indices and LNC was obtained and presented as a heatmap in Fig. 6(a). Interestingly, highly correlated areas in the index-LNC correlation heatmap were found to be the same as highly correlated areas in the index-LWC correlation heatmap published in Raj et al. (2021). This restricted use of the index-LNC correlation heatmap for finding nitrogen-sensitive indices independently from information on water content as the same indices were also highly sensitive to LWC. To find the indices more correlated with LNC than LWC, a correlation difference heatmap between the LNC and LWC heatmap was created. The difference heatmap is shown in Fig. 6(b). After comparing the correlation heatmaps and the difference heatmap, four indices were selected for further analysis. The identified wavelengths for these indices are indicated in Fig. 6 and listed in Table 3. Out of these four indices, only the RedEdge1 index showed spatial variability on the drone-based image. Thus, the RedEdge1 index along with the DCNI index from literature was selected for further analysis.

2.4. Leaf nitrogen content estimation model

Drone-based hyperspectral band images were used to create farm-index maps for the newly identified index - RedEdge1 (Table 3) and the DCNI index (Chen et al., 2010). The spatial resolution of the farm map was around 1 cm resulting in most canopy pixels being a pure pixel. The purity of the vegetation pixels resulted in higher narrow-band NDVI values of each vegetation pixel when compared to NDVI values of background pixels. Pixels having an NDVI value less than 0.7 were either non-vegetative or mixed pixel at the leaf edges. These lower NDVI pixels

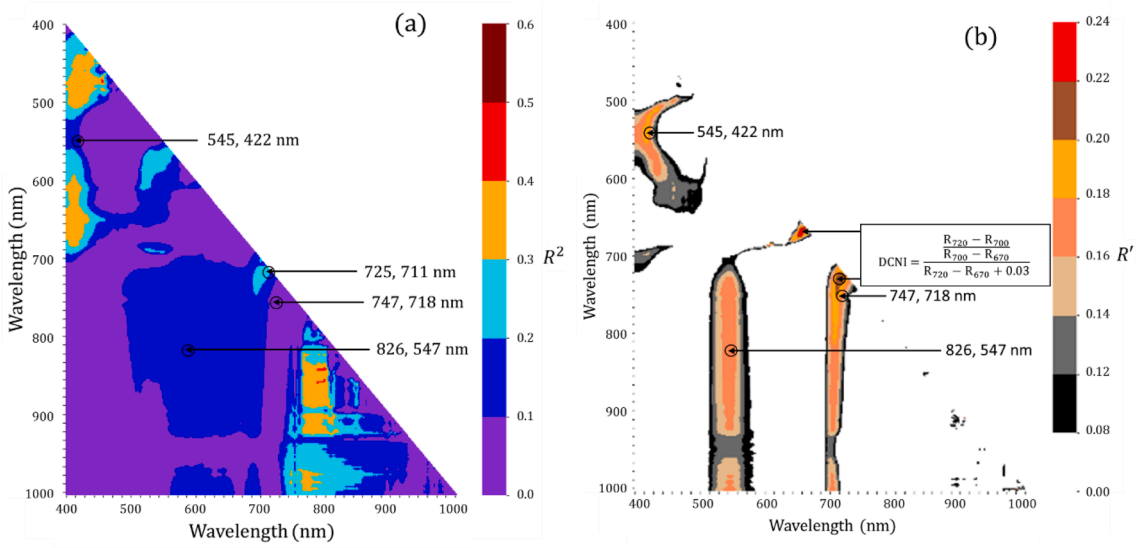


Fig. 6. (a) Heatmap of R^2 between narrow-band normalized difference indices and leaf nitrogen content. (b) The heat map of LNC-LWC correlation coefficient difference R' showing only those indices having a superior correlation with LNC compared to LWC. Indices in the white part of heatmap are correlated more with LWC than LNC.

Table 3
Pure pixel, narrow-band indices identified in this research for LNC estimation.

Index Formula	Short from	Range	Usability
$\frac{R_{545} - R_{422}}{R_{545} + R_{422}}$	$GBslope_{Index}$	0–0.35	Spectroradiometer data only
$\frac{R_{826} - R_{547}}{R_{826} + R_{547}}$	$GreenNIR_{Index}$	0.2–0.8	
$\frac{R_{747} - R_{718}}{R_{747} + R_{718}}$	$RedEdge2_{Index}$	0.1–0.5	
$\frac{R_{725} - R_{711}}{R_{725} + R_{711}}$	$RedEdge1_{Index}$	0.2–0.4	Spectroradiometer and drone data

were therefore assigned a null value in the map. A similar background removal approach was used in Raj et al. (2021).

Synthetic data was created for model training, using the maximum and minimum values of individual pure pixel vegetation index maps. As RedEdge1 and DCNI were both positively correlated with LNC, the minimum and maximum values were assigned to the minimum and maximum ground-truth LNC, respectively. A straight line was interpolated between the extreme values and 1000 random Gaussian distributed points generated within 10% of the interpolated value as shown in Fig. 7. Along with the synthetic index data, a decreasing trend of LNC with progressive growth stages was modeled using days after sowing (DAS) information. The DAS-LNC data was created using Gaussian noise around a third-order polynomial fit to the median values of ground truth data. Use of synthetic data for model training will reduce the model dependence on collected data resulting in improved model repeatability.

A GBM model for LNC estimation was trained on 20,000 sampled points from the synthetic data. The GBM model was chosen for its promising performance in Raj et al. (2021). Moreover, the property of

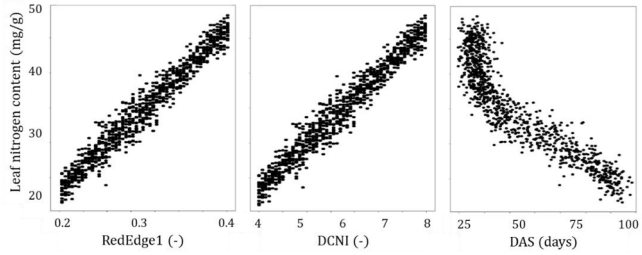


Fig. 7. The synthetic data for RedEdge1, DCNI, and DAS relation with LNC.

the GBM model to convert weak learners to strong learners by performing sequential improvement of decision trees (Friedman, 2001) made it a suitable choice for LNC estimation. Optimal hyperparameters of the model were obtained using a GridSearchCV algorithm (Zhao et al., 2020). The best parameters were obtained with a learning rate of 0.395, minimum sample split of 12, and number of estimators of 1,900. The LNC obtained from the GBM model was termed as estimated LNC.

3. Results

The temporal ground-truth LNC for various irrigation and nitrogen treatment plots are shown in Fig. 8. LNC was found to be decreasing with progress in growth stages from 6-leaf to maturity in all the treatments. The plants were found to be responsive to nitrogen application as LNC increased temporarily, especially for those plots treated with limited irrigation. The effect of different soil nitrogen treatments on actual leaf nitrogen content was found to be reduced for sufficiently irrigated plots.

Various indices identified from the literature (Table 2) have been tested to estimate LNC. These indices were created to estimate canopy nitrogen content and, to date, mainly tested on canopy-level low spatial resolution data. The performance of these indices on pure-pixel narrow-band data was poor. Apart from DCNI, no other index available in literature could give an R^2 greater than 0.05. The RedEdge1 index (identified in this research) and DCNI (identified from literature) farm index maps were created. Average index values for each treatment were used to create scatter plots with respective CHNS-based LNC values. As shown in Fig. 9, it was found that the RedEdge1 index and DCNI were correlated to LNC with an R^2 of 0.27 and 0.20, respectively.

The GBM model estimated temporal and spatial distribution of crop LNC is shown in Fig. 10. Comparison of these maps with LWC showed that, in general, LNC was high in the areas where LWC was also high. Although this is true, as LNC and LWC have a high correlation (R^2 of 0.7) as shown in Fig. 11, the efficiency of the LNC model could be checked by analysing the LNC distribution in the same LWC area. The analysis of water-stressed and non-water stressed plots was undertaken separately to see the effect of different LWC on the LNC model. Moreover, the correlation between estimated LNC and CHNS-based LNC was made. The maps were analysed for the 6-leaf and pre-tasseling stages, where the water-stressed plots gave an R^2 of 0.63 and RMSE of 2.74 mg/g, but the plots with higher LWC gave an R^2 of 0.26 and RMSE of 4.54 mg/g. This shows that the LNC model can identify nitrogen stress areas from

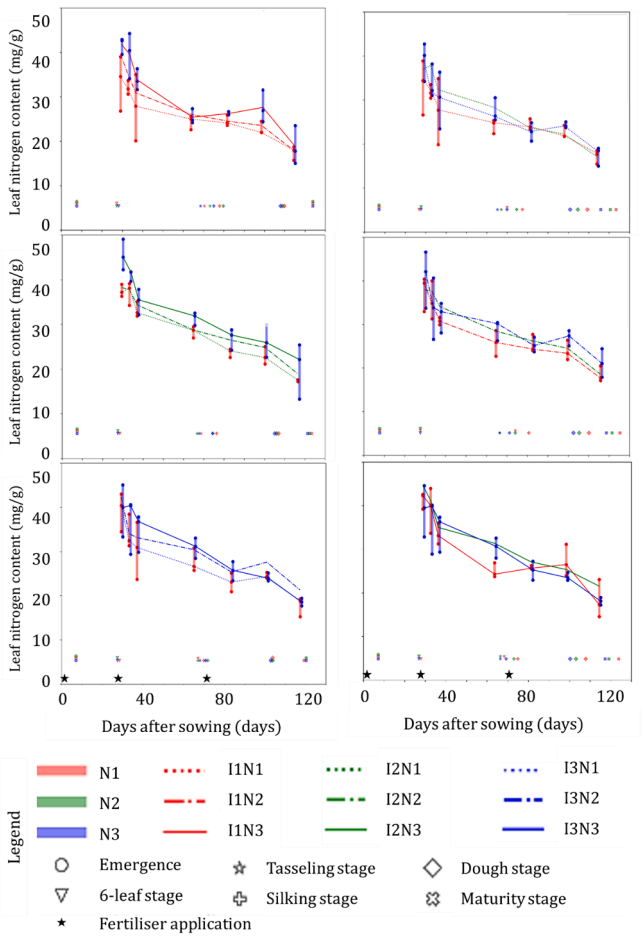


Fig. 8. Temporal CHNS-based LNC values for different irrigation and nitrogen treatment plots.

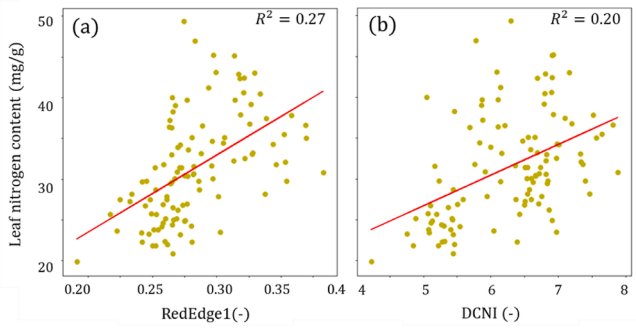


Fig. 9. Scatter plot between ground truth LNC and plot-wise averaged index value of (a) RedEdge1 and (b) DCNI index.

the water-stressed plots, but the model could not perform well for regions where no water stress was present. This is a limitation of the model, suggesting that the LNC model should only be applied once the LWC model has been used to classify low and high LWC areas, as suggested in Raj et al. (2021). The scatter plots between estimated LNC and CHNS-based LNC for different conditions are shown in Fig. 12, with the model giving an R^2 of 0.33 and RMSE of 5.35 mg/g when tested on 6-leaf and pre-tasseling stage data with no discrimination of water-stressed regions. However, the model accuracy increased to an R^2 of 0.63 and RMSE of 2.74 mg/g when applied to only the water-stressed regions.

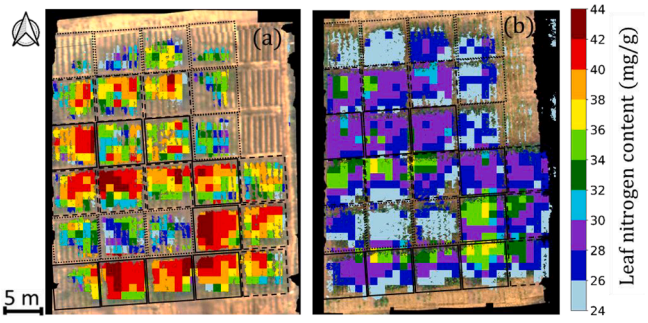


Fig. 10. Colour-coded farm leaf nitrogen map obtained from the trained GBM showing nitrogen content in plant leaves on (a) 6-leaf stage and (b) pre-tasseling stage.

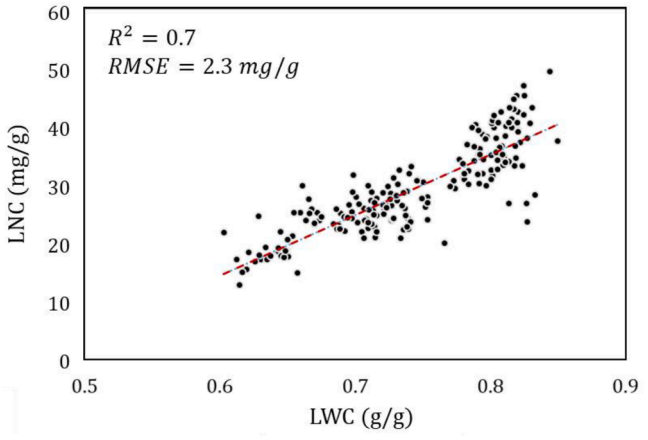


Fig. 11. Scatter plot between LWC and LNC indicating a strong dependence of LNC on LWC.

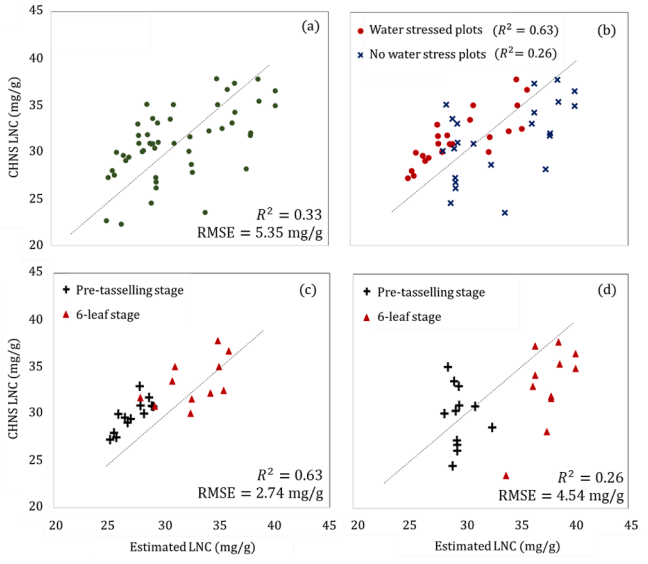


Fig. 12. (a) Scatter plot between estimated and CHNS-derived LNC values for all plots; (b) Water stress classification-based scatter plot for all plots; (c) Growth-stage based scatter plot for water-stressed plots only; (d) Growth-stage based scatter plot for non-water stressed plots only.

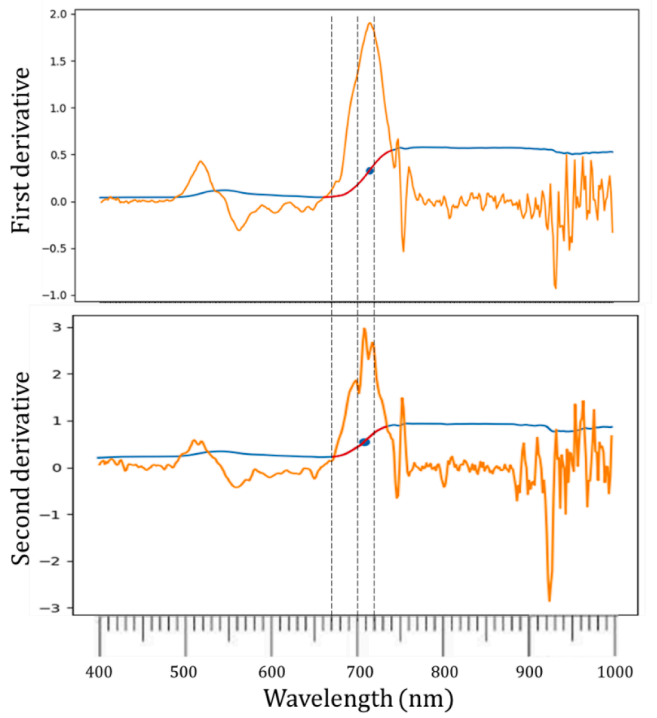


Fig. 13. Orange lines show derivatives of typical leaf reflectance spectra. The two peaks can be observed in the second derivative spectra around 700 nm and 720 nm. The blue line spectra is the leaf reflectance spectra. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

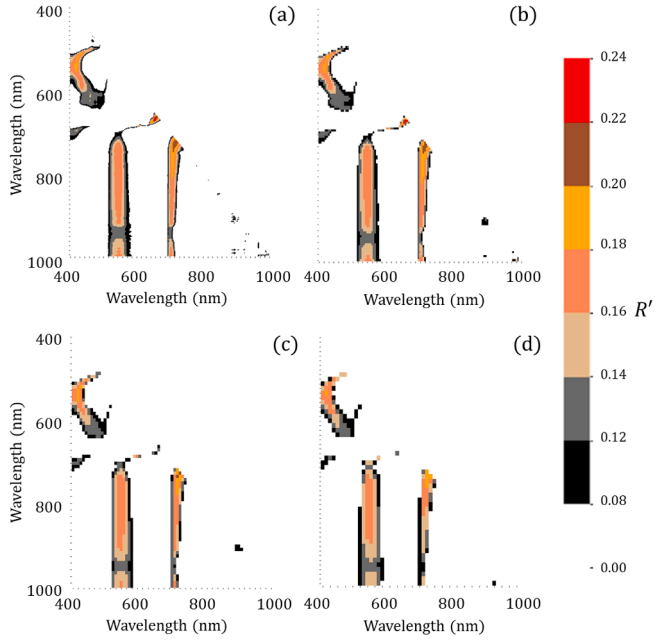


Fig. 14. Comparison of correlation coefficient difference heat map showing only those indices which show superior correlation with LNC compared to LWC. The bandwidths used to create these leaf nitrogen content heatmaps are as follows: (a) 2 nm, (b) 5 nm, (c) 8 nm, (d) 11 nm.

4. Discussion

The indices used in this research – DCNI and RedEdge1 – were created using bands from the red edge region of the electromagnetic spectrum. Chen et al. (2010), who introduced the DCNI index, presented a detailed analysis of the sensitivity of nitrogen concentration to the relative height changes in the peaks of derivative spectra of the leaf spectral signature. In this research, those peaks were found in the red-edge zone of the spectra around 700 nm and 720 nm, as depicted in Fig. 13. Importantly, Chen et al. (2010) found that the nitrogen concentration in the leaves was highly correlated with the relative height changes of those peaks, which can be estimated using the ratio of the average heights of the two peaks. However, Chen et al. (2010) did not give any scientific reasoning for this high correlation. Chen et al. (2010) also added the 670 nm wavelength in the DCNI index to reduce the effects of LAI on the index, which may be helpful for mixed pixel data, but not for pure-pixel data. The LNC model presented in this research was for pure vegetation pixels, which enabled the RedEdge1 index to give better results than DCNI without adding other factors in the index to remove the effects of crop biophysical properties.

The literature-based indices presented in Table 2 did not show high correlation with leaf nitrogen content. This may be because literature-based indices, until now, have been mostly validated for canopy level nitrogen content which are influenced by the biomass and other biophysical properties of the crop available in each pixel area. However, in this research the indices were tested for pure-pixel leaf level total nitrogen content which is independent of any crop biophysical properties.

One crucial observation made in this research is that the identified indices tended to lose sensitivity to LNC estimation as the bandwidth broadened. Fig. 14 shows the comparative correlation difference heatmap between LNC and LWC created with different bandwidth data. The analysis was undertaken on bandwidth correlation difference heatmaps created using a 2 nm to 11 nm bandwidth dataset. The indices performed similarly until 5 nm bandwidth data, with the correlation reducing drastically as the bandwidth further broadened. Thus, this research suggests that for distinguishing LNC and LWC, data should be collected with sensors having the central wavelengths given in Table 3 with a less than 5 nm bandwidth.

5. Conclusion

There are no known amide bond absorption bands available in the 400–1000 nm electromagnetic range which can be directly used for leaf nitrogen content estimation. However, the red edge of the electromagnetic spectrum was found to be sensitive to crop leaf nitrogen content. The peaks in the derivative spectra of red-edge between 700 and 725 nm (bandwidth less than 5 nm) have been found more sensitive to leaf nitrogen content than leaf water content. Moreover, it was found necessary to differentiate between water-stressed and non-stressed areas as the proposed LNC model performed better for water-stressed crops. Importantly, recommendations of this LNC study are that sensors for mapping LNC should be based on narrow bands (bandwidth less than 5 nm) centered at and around 700 and 725 nm, and any future research for LNC estimation should also consider the sensitivity of leaf water content on their model.

CRediT authorship contribution statement

Rahul Raj: Conceptualization, Methodology, Formal analysis, Writing – review & editing, Visualization, Validation. **Jeffrey P. Walker:** Conceptualization, Writing – review & editing, Investigation, Resources, Supervision. **Rohit Pingale:** Data curation. **Balaji Naik Banoth:** Conceptualization, Validation, Resources. **Adinarayana Jagarlapudi:** Investigation, Resources, Writing – review & editing, Supervision, Project administration, Funding acquisition.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence

the work reported in this paper.

Acknowledgments

BharatRohan company has helped selflessly in flying the drone for hyperspectral data collection. Authors also thank Priyankar Chand, Milan Malhotra, and Bhupii Sharma for their help in preparing samples for CHNS analysis.

References

Andersson, I., Backlund, A., 2008. Structure and function of Rubisco. Plant Physiol. Biochem. 46 (3), 275–291.

Blumenthal, J.M., Baltensperger, D.D., Cassman, K.G., Mason, S.C. and Pavlista, A.D., 2008. Importance and effect of nitrogen on crop quality and health. In Nitrogen in the Environment (pp. 51-70). Academic Press.

Chaplin, M., 2006. Do we underestimate the importance of water in cell biology? Nat. Rev. Mol. Cell Biol. 7 (11), 861–866.

Chen, P., Haboudane, D., Tremblay, N., Wang, J., Vigneault, P., Li, B., 2010. New spectral indicator assessing the efficiency of crop nitrogen treatment in corn and wheat. Remote Sens. Environ. 114 (9), 1987–1997.

Cho, M.A., Skidmore, A.K., 2006. A new technique for extracting the red edge position from hyperspectral data: The linear extrapolation method. Remote Sens. Environ. 101 (2), 181–193.

Corti, M., Gallina, P.M., Cavalli, D., Cabassi, G., 2017. Hyperspectral imaging of spinach canopy under combined water and nitrogen stress to estimate biomass, water, and nitrogen content. Biosyst. Eng. 158, 38–50.

Craig, J.R., Wollum, A.G., 1982. Ammonia volatilization and soil nitrogen changes after urea and ammonium nitrate fertilization of Pinus taeda L. Soil Sci Soc Am J. 46 (2), 409–414.

Curran, P.J., Dungan, J.L., Macler, B.A., Plummer, S.E., Peterson, D.L., 1992. Reflectance spectroscopy of fresh whole leaves for the estimation of chemical concentration. Remote Sens. Environ. 39 (2), 153–166.

Damodaran, S., 2008. Amino acids, peptides and proteins. Fennema's food chemistry 4, 425–439.

Daughtry, C.S.T., Walthall, C.L., Kim, M.S., De Colstoun, E.B., McMurtrey Iii, J.E., 2000. Estimating corn leaf chlorophyll concentration from leaf and canopy reflectance. Remote Sens. Environ. 74 (2), 229–239.

Dhaliwal, G.S., Gupta, N., Kukal, S.S. and Meetpal-Singh, 2014. Standardization of automated Vario EL III CHNS analyzer for total carbon and nitrogen determination in plants. Commun Soil Sci Plant Anal, 45(10), pp.1316-1324.

Du, L., Gong, W., Shi, S., Yang, J., Sun, J., Zhu, B., Song, S., 2016. Estimation of rice leaf nitrogen contents based on hyperspectral LIDAR. Int J Appl Earth Obs Geoinf 44, 136–143.

Eitel, J.U.H., Long, D.S., Gessler, P.E., Smith, A.M.S., 2007. Using in-situ measurements to evaluate the new RapidEye™ satellite series for prediction of wheat nitrogen status. Int. J. Remote Sens. 28 (18), 4183–4190.

Elrashidi, M.A., Mays, M.D., Peaslee, S.D., Hooper, D.G., 2005. A technique to estimate nitrate–nitrogen loss by runoff and leaching for agricultural land, Lancaster County. Nebraska. Commun Soil Sci Plant Anal 35 (17–18), 2593–2615.

Fan, L., Zhao, J., Xu, X., Liang, D., Yang, G., Feng, H., Yang, H., Wang, Y., Chen, G., Wei, P., 2019. Hyperspectral-based estimation of leaf nitrogen content in corn using optimal selection of multiple spectral variables. Sensors 19 (13), 2898.

Feng, W., Yao, X., Zhu, Y., Tian, Y.C., Cao, W.X., 2008. Monitoring leaf nitrogen status with hyperspectral reflectance in wheat. Eur J Agron 28 (3), 394–404.

Feng, W., Zhang, H.Y., Zhang, Y.S., Qi, S.L., Heng, Y.R., Guo, B.B., Ma, D.Y., Guo, T.C., 2016. Remote detection of canopy leaf nitrogen concentration in winter wheat by using water resistance vegetation indices from in-situ hyperspectral data. Field Crops Res. 198, 238–246.

Fitzgerald, G.J., Rodriguez, D., Christensen, L.K., Belford, R., Sadras, V.O., Clarke, T.R., 2006. Spectral and thermal sensing for nitrogen and water status in rainfed and irrigated wheat environments. Precis. Agric. 7 (4), 233–248.

Franks, Felix, 1988. Protein hydration. In: Franks, F. (Ed.), Characterization of Proteins. Biological Methods, 1st. Humana Press Inc., pp. 127–154

Friedman, Jerome H., 2001. Greedy function approximation: a gradient boosting machine. Ann. Stat. 29, 1189–1232, 5. https://www.jstor.org/stable/2699986.

Gezgin, S., Bayraklı, F., 1995. Ammonia volatilization from ammonium sulphate, ammonium nitrate, and urea surface applied to winter wheat on a calcareous soil. J. Plant Nutr. 18 (11), 2483–2494.

Gutteridge, S., Gatenby, A.A., 1995. Rubisco synthesis, assembly, mechanism, and regulation. Plant Cell 7 (7), 809.

Haboudane, D., Miller, J.R., Pattey, E., Zarco-Tejada, P.J., Strachan, I.B., 2004. Hyperspectral vegetation indices and novel algorithms for predicting green LAI of crop canopies: Modeling and validation in the context of precision agriculture. Remote Sens. Environ. 90 (3), 337–352.

Haboudane, D., Miller, J.R., Tremblay, N., Zarco-Tejada, P.J., Dextraze, L., 2002. Integrated narrow-band vegetation indices for prediction of crop chlorophyll content for application to precision agriculture. Remote Sens. Environ. 81 (2–3), 416–426.

Hansen, P.M., Schjoerring, J.K., 2003. Reflectance measurement of canopy biomass and nitrogen status in wheat crops using normalized difference vegetation indices and partial least squares regression. Remote Sens. Environ. 86 (4), 542–553.

Haris, Parvez I., Chapman, Dennis, 1994. Analysis of polypeptide and protein structures using Fourier transform infrared spectroscopy. Microscopy, Optical Spectroscopy, and Macroscopic Techniques 183–202.

He, L., Song, X., Feng, W., Guo, B.B., Zhang, Y.S., Wang, Y.H., Wang, C.Y., Guo, T.C., 2016. Improved remote sensing of leaf nitrogen concentration in winter wheat using multi-angular hyperspectral data. Remote Sens. Environ. 174, 122–133.

Howitt, S.M., Udvardi, M.K., 2000. Structure, function and regulation of ammonium transporters in plants. Biochim Biophys Acta 1465 (1–2), 152–170.

Knox, J.W., Kay, M.G., Weatherhead, E.K., 2012. Water regulation, crop production, and agricultural water management - Understanding farmer perspectives on irrigation efficiency. Agric. Water Manag. 108, 3–8.

Kokaly, R.F., 2001. Investigating a physical basis for spectroscopic estimates of leaf nitrogen concentration. Remote Sens. Environ. 75 (2), 153–161.

Kutman, U.B., Yildiz, B., Cakmak, I., 2011. Effect of nitrogen on uptake, remobilization and partitioning of zinc and iron throughout the development of durum wheat. Plant Soil. 342 (1), 149–164.

Leghari, S.J., Wahocho, N.A., Laghari, G.M., HafeezLaghari, A., MustafaBhabhan, G., HussainTalpur, K., Bhutto, T.A., Wahocho, S.A. and Lashari, A.A., 2016. Role of nitrogen for plant growth and development: A review. Adv. Environ. Biol., 10(9), pp.209-219.

Liang, L., Di, L., Huang, T., Wang, J., Lin, L., Wang, L., Yang, M., 2018. Estimation of leaf nitrogen content in wheat using new hyperspectral indices and a random forest regression algorithm. Remote Sens. 10 (12), 1940.

Liu, H., Zhu, H., Wang, P., 2017. Quantitative modelling for leaf nitrogen content of winter wheat using UAV-based hyperspectral data. Int. J. Remote Sens. 38 (8–10), 2117–2134.

Masclaux-Daubresse, C., Daniel-Vedele, F., Dechorgnat, J., Chardon, F., Gaufichon, L., Suzuki, A., 2010. Nitrogen uptake, assimilation and remobilization in plants: challenges for sustainable and productive agriculture. Ann. Bot. 105 (7), 1141–1157.

Mason, M.G., 1977. Sources of nitrogen for cereals: urea, ammonium nitrate or sulphate of ammonia?. J Dept Agr S Australia, Western Australia, Series 4, 18(1), pp.15-19.

Oehlschläger, J., Schmidhalter, U. and Noack, P.O., 2018, September. UAV-based hyperspectral sensing for yield prediction in winter barley. In 2018 9th Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS) (pp. 1-4). IEEE.

Ogawa, M., Tanaka, Y., Jursa, A.S., et al., 1964. Isotope shift of the nitrogen absorption bands in the vacuum ultraviolet region. Can. J. Phys. https://doi.org/10.1139/p64-157.

Ohyama, T., 2010. Nitrogen as a major essential element of plants. Nitrogen assimilation in plants 37 (2), 2–17.

Raj, R., Walker, J.P., Vinod, V., Pingale, R., Naik, B., Jagarlapudi, A., 2021. Leaf water content estimation using top-of-canopy airborne hyperspectral data. Int J Appl Earth Obs Geoinf 102, 102393.

Reyniers, M., Walvoort, D.J., De Baardemaaker, J., 2006. A linear model to predict with a multi-spectral radiometer the amount of nitrogen in winter wheat. Int. J. Remote Sens. 27 (19), 4159–4179.

Sims, D.A., Gamon, J.A., 2002. Relationships between leaf pigment content and spectral reflectance across a wide range of species, leaf structures and developmental stages. Remote Sens. Environ. 81 (2–3), 337–354.

Schrader, L.E., 1984. Function and Transformations of Nitrogen in Higher Plants. Nitrogen in crop production 55–65.

Strachan, I.B., Pattey, E., Boisvert, J.B., 2002. Impact of nitrogen and environmental conditions on corn as detected by hyperspectral reflectance. Remote Sens. Environ. 80 (2), 213–224.

Stroppiana, D., Boschetti, M., Confalonieri, R., Bocchi, S., Brivio, P.A., 2006. Evaluation of LAI-2000 for leaf area index monitoring in paddy rice. Field Crops Res. 99 (2–3), 167–170.

Tan, C., Du, Y., Zhou, J., Wang, D., Luo, M., Zhang, Y., Guo, W., 2018. Analysis of different hyperspectral variables for diagnosing leaf nitrogen accumulation in wheat. Front. Plant Sci. 9, 674.

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 Soil 376 (1), 193–209.

Tian, Y.C., Yao, X., Yang, J., Cao, W.X., Hannaway, D.B., Zhu, Y., 2011. Assessing newly developed and published vegetation indices for estimating rice leaf nitrogen concentration with ground-and space-based hyperspectral reflectance.Field Crops. Res. 120 (2), 299–310.

Thenkabail, P.S. and Lyon, J.G. eds., 2016. Hyperspectral remote sensing of vegetation. CRC press.

van Grinsven, H.J., Bouwman, L., Cassman, K.G., van Es, H.M., McCrackin, M.L., Beusen, A.H., 2015. Losses of ammonia and nitrate from agriculture and their effect on nitrogen recovery in the European Union and the United States between 1900 and 2050. J. Environ. Qual. 44 (2), 356–367.

Widlowski, J.L., Mio, C., Disney, M., Adams, J., Andredakis, I., Atzberger, C., Brennan, J., Busetto, L., Chelle, M., Ceccherini, G. and Colombo, R., 2015. The fourth phase of the radiative transfer model intercomparison (RAMI) exercise: Actual canopy scenarios and conformity testing. Remote Sens. Environ., 169, pp.418-437.

Yao, X., Zhu, Y., Tian, Y., Feng, W., Cao, W., 2010. Exploring hyperspectral bands and estimation indices for leaf nitrogen accumulation in wheat. Int J Appl Earth Obs Geoinf 12 (2), 89–100.

Zhao, S., Mao, X., Lin, H., Yin, H., Xu, P., 2020. Machine Learning Prediction for 50 Anti-Cancer Food Molecules from 968 Anti-Cancer Drugs. Int. J. Intell. Sci 10 (1), 1–8.