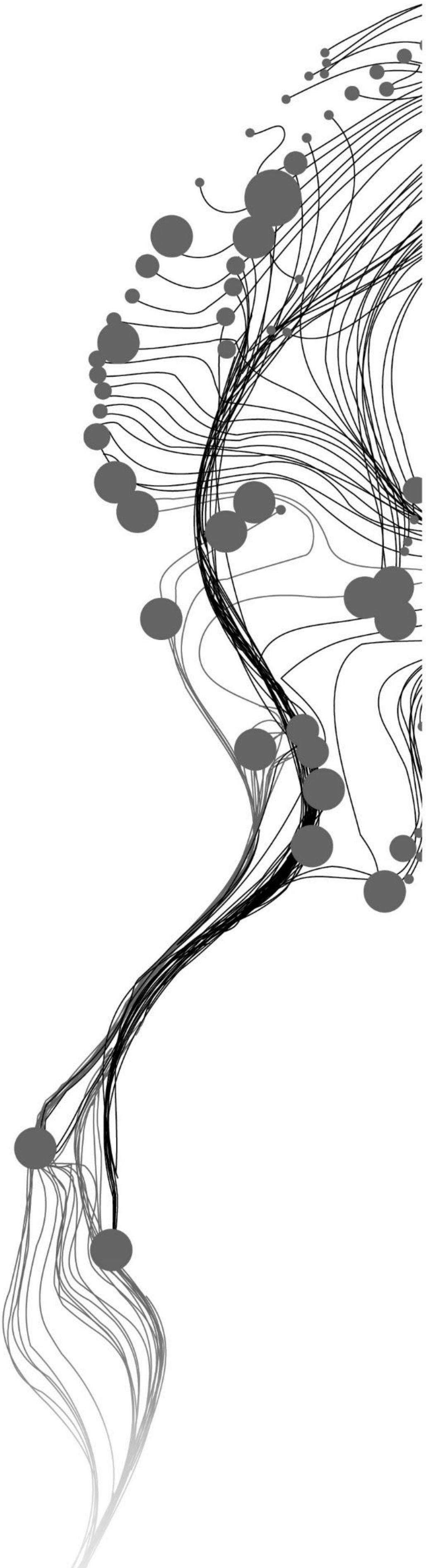
ASSESSING THE VARIATION OF NITROGEN AND CARBON IN HEALTHY AND LODGED WHEAT CROPS USING SENTINEL-2 AND FIELD SPECTROSCOPY MEASUREMENTS

OYETOUN OLATORERA ALONGE
JUNE 2024

SUPERVISORS:
Dr R. Darvishzadeh

Prof. A.D. Nelson



ASSESSING THE VARIATION OF NITROGEN AND CARBON IN HEALTHY AND LODGED WHEAT CROPS USING SENTINEL-2 AND FIELD SPECTROSCOPY MEASUREMENTS

OYETOUN OLATORERA ALONGE
Enschede, The Netherlands, JUNE 2024

Thesis submitted to the Faculty of Geo-Information Science and Earth Observation of the University of Twente in partial fulfilment of the requirements for the degree of Master of Science in Geo-information Science and Earth Observation.
Specialization: Natural Resources Management

SUPERVISORS:
Dr R. Darvishzadeh
Prof. A.D. Nelson

THESIS ASSESSMENT BOARD:
Dr.ir. T.A. Groen (Chair)
Prof. A. Ramoelo (External Examiner)

Abstract

Estimation of foliar nitrogen and carbon content in crops is critical for assessing the nutrient cycle and optimal nutrient management, to improve crop health monitoring, and enhance yield predictions. This study investigates the estimation and retrieval of nitrogen and carbon content in healthy and lodged wheat using field hyperspectral remote sensing data taken during the field campaign between 30 April and May 18, 2023, and Sentinel-2 data of 23 April 2023 was obtained from the European Space Agency's (ESA) Copernicus Open Access Hub, covering our study area in Bonifiche Ferraresi farm, Jolanda di Savoia, Ferrara, Italy.

Wheat field coordinates in lodged and healthy fields, together with leaf samples, were collected and wet chemistry analysis was performed to measure nitrogen and carbon content in leaf samples. Widely used vegetation indices known for the estimation of nitrogen and carbon were used to estimate these biochemical parameters. Pearsons's correlation coefficient test was used to determine important bands for the estimation and retrieval of carbon and nitrogen in both lodged and healthy wheat plots. The significant bands were further analysed using continuum removal and band depth analysis, trimming down the number of significant bands to a lower number by selecting the most important bands, for discriminating lodged and healthy wheat samples. Partial Least Squares Regression (PLSR) was used further for the estimation of carbon and nitrogen both at the field hyperspectral and at Sentinel-2 levels.

The result of spectral analysis demonstrated significant differences in spectral properties between healthy and lodged wheat, with lodged wheat showing higher reflectance across visible, near infrared (NIR), and shortwave infrared regions. Among the vegetation indices used, narrowband NDVI was found to provide better predictive accuracy for foliar carbon and nitrogen content in lodged and healthy wheat, with lodged biochemical parameters having higher prediction values than healthy wheat's. Using the selected bands from the band depth analysis, in the PLSR model, the result showed the highest R^2 CV value for carbon in lodged wheat (0.41), followed closely by nitrogen (0.39), nitrogen and carbon in healthy wheat had R^2 value of 0.32 and 0.31 respectively, using venetian blind validation, when randomized validation was used all the biochemical parameters in both healthy and lodged wheat experienced higher prediction value, with healthy wheat having the highest prediction of 94% and 91% for nitrogen and carbon respectively. Sentinel-2 data for the PLSR model, the highest R^2 CV value was recorded by lodged nitrogen (0.60), followed by healthy carbon (44), healthy nitrogen (0.42) and then lodged carbon (0.37) using venetian blind validation, when randomized validation was used, the same trend with increase prediction was noted, but the highest prediction was from lodged nitrogen (0.80), followed by lodged carbon (0.78), the healthy wheat biochemicals have same prediction value of 0.72.

The results showed that lodged wheat has higher reflectance throughout the spectrum due to change in the structure of the canopy and plant stress. This leaded to better prediction of the biochemical parameters in lodged wheat, although lodging changed the absorption band location for nitrogen and carbon content, when compared with healthy wheat. Hyperspectral data proved superior in predicting these biochemical parameters, particularly in lodged wheat, due to its ability to capture subtle spectral variations. In contrast, Sentinel-2 data, despite its broader spectral bands, also provided valuable insights but showed lower predictive accuracy compared to hyperspectral data.

The findings emphasize the importance of selecting appropriate spectral bands and vegetation indices suited for the specific crop condition, and that validation methods play a crucial role in model accuracy. Finally, it highlights the superior performance of hyperspectral data over Sentinel-2 data for precise agricultural monitoring and management.

Acknowledgements

The completion of this work is owed to God Almighty who deserves my immeasurable appreciation and profound gratitude. It is all by His grace I live to attain this academic level. I have benefited immensely from mentorship, friendship, intellectual insight, assistance, and kindness of people so many that this space would not permit me to mention. Please if your name is not mentioned here be assured of my eternal gratitude and thanks.

My gratitude goes to my supervisors, Dr. Roshanak Darvishzadeh, and Prof. Dr. Andy Nelson for supervising this work, for their tolerant, patience, constructive feedback, concern, and encouragement which guided me in this research. I am grateful to you both for impacting me positively in my academic journey.

I appreciate the chairman of thesis assessment board Dr.ir. T.A. Groen for your suggestion and comments every time you assess this work, and this has helped me to improve the output quality of this research. I am grateful to Kathrin Zweers-Peter who put me through all the laboratory process to be done which is an important part of my data. To Dr Çelik, Furkan, thank you for the times you've attended to my codes even during weekends, to Dr. Duren, Iris van thank you for your mentorship, Willem Nieuwenhuis, I appreciate your technical help. Theresa Boogaard-Burke and Jorike Siemerink, thank you so much for your care, help and for everything.

My appreciation goes Prof. A.J. Omojola of the University of Ibadan, for approving my study leave, God bless you sir and to NUFFIC for the OKP scholarship, which has given me this wonderful opportunity of being a graduate of the great Faculty of ITC, University of Twente, Netherlands.

My utmost appreciation goes to my husband, Dr. A.J. Alonge and our children, Delight, Desire and Divine for their support, prayer, understanding, love and standing by me during my study. I would like to express my profound gratitude to my mother, Mrs C.O. Adeyeye, and siblings and their perfect halves: Adedoyin & Jite, Dr. Olaonipekun & Fani Oluwaseyifunmi, Olariike & Tosin, and Oladimeji for their prayers and support. Dr Sola Owolabi; thank you for the encouragement, to my dad Prof. E.T Jaiyeoba, so much love.

I wish to also express my boundless appreciation to my course mates at the NRM specialisation Anchal, Emmanuel, Pauline, Charlie, Kwasi, Padma, Lawrence, Prince, Alma, and others. To my friends Tionge Munthali, Kaka Mosun & Vinaida Sno, thanks for everything. To my brother and sisters from another mother: Kelvin, Tendai, Tatenda and Stella, thank you so much. To my family here in the Netherlands, Patrick & Winnie Eneche, Segun & Kemi Adeoye, David & Isabel Oshinaya, Mummy Queen, Pastor Donald and Pastor Nelson and others, love you all.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. CHAPTER 1: **INTRODUCTION**

1.1 Wheat and crop lodging

Food security is a paramount concern of our time, posing complex challenges that extend far beyond the simple question of food availability, affecting 9% of the world's population. (Ghosh et al, 2022). The prevalence of global hunger has exhibited an upward trend since 2014, despite increased food production over the past three decades. In 2019, a staggering two billion individuals, constituting approximately 25.9% of the global population, encountered hunger or faced inadequate and insufficient access to nourishing food (Raj et al., 2022). According to Grote et al. (2020), 820 million people are under-nourished in terms of energy intake, and 1.3 billion people suffer from micronutrient deficiencies. Grains occupy a significant position within the global food supply, providing essential nutrients, including proteins, carbohydrates, vitamins, and minerals. Wheat, rice, maize, oats, and barley are versatile grains that can be used to create a wide range of foods and recipes, they are the foundation of basic foods and are widely cultivated worldwide (Albahri et al., 2023). Among these, wheat is an important cereal crop, and it is extensively cultivated. It has high nutritional value and has a variety of uses. It supports agricultural economies, adds to global food security, has cultural significance and is crucial for maintaining human populations (Shewry & hey, 2015; FAO, 2020). The use of wheat is rising globally, even in countries where the environment is not ideal for growing wheat. The estimated world production for wheat is 790,752 (1000 MT) for the year 2024 (USDA, 2024) Figure 1 shows the world wheat production distribution, and despite significant progress in global food production, millions still contend with food insecurity, a situation where individuals and communities lack consistent access to safe, nutritious, and culturally acceptable food that meets their dietary needs for an active and healthy life (Raphela & Pillay, 2021). Production of wheat is susceptible to biotic and abiotic stresses like climate change (Rahman et al., 2017), pests and diseases (Figueroa et al., 2017. Nigus et al., 2022, and lodging is one of the challenges in wheat production (Wu et al., 2022: Chauhan et al., 2019, 2020, 2021, Li et al., 2023).

Crop lodging is prevalent in field crops and is characterized by the bending or fracturing plant stems or stalks (Chauhan et al., 2019, 2020, 2021; Shrestha et al., 2019; Wu et al., 2022). This phenomenon mostly affects cereal crops, including but not limited to wheat, maize (corn), barley, and rice. The adverse impact on plant stability leads to the crop collapsing or tilting, resulting in reduced agricultural output (Chauhan et al., 2019; Shrestha et al., 2019). When plants become lodged, their ability to absorb nutrients, carryout photosynthesis efficiently, and transport nutrients is hampered. Chauhan et al., (2020) also recognized that lodging typically occurs during the later part of the booting stage, because of the grains' weight. The booting stage in wheat represents the onset of the reproductive phase, indicating the shift from the vegetative phase to the reproductive phase. Lodging leads to disease vulnerability, decreased agricultural output, hindered mechanical harvesting and storage, increased post-harvest losses, and further strain on food supplies.

There are two types of crop lodging: root lodging and stem lodging. Root lodging is commonly caused by insufficient root development, shallow planting, imbalanced soil fertility, root damage from pests, the impact of strong winds and rain. However, stem lodging occurs later in the crop's growth, it is primarily associated with unfavourable weather conditions, such as powerful winds, heavy rainfall, or severe storms. These conditions exert physical stress on the plants, making them more prone to bending or breaking (Chauhan et al., 2020; Wu et al., 2022). Aside from environmental conditions, various factors can contribute to stem lodging, including pests and diseases, plant variety traits, soil variables, nutrient uptake, topographical features of the farmland, excessive plant height, and excessive nitrogen application, among others.

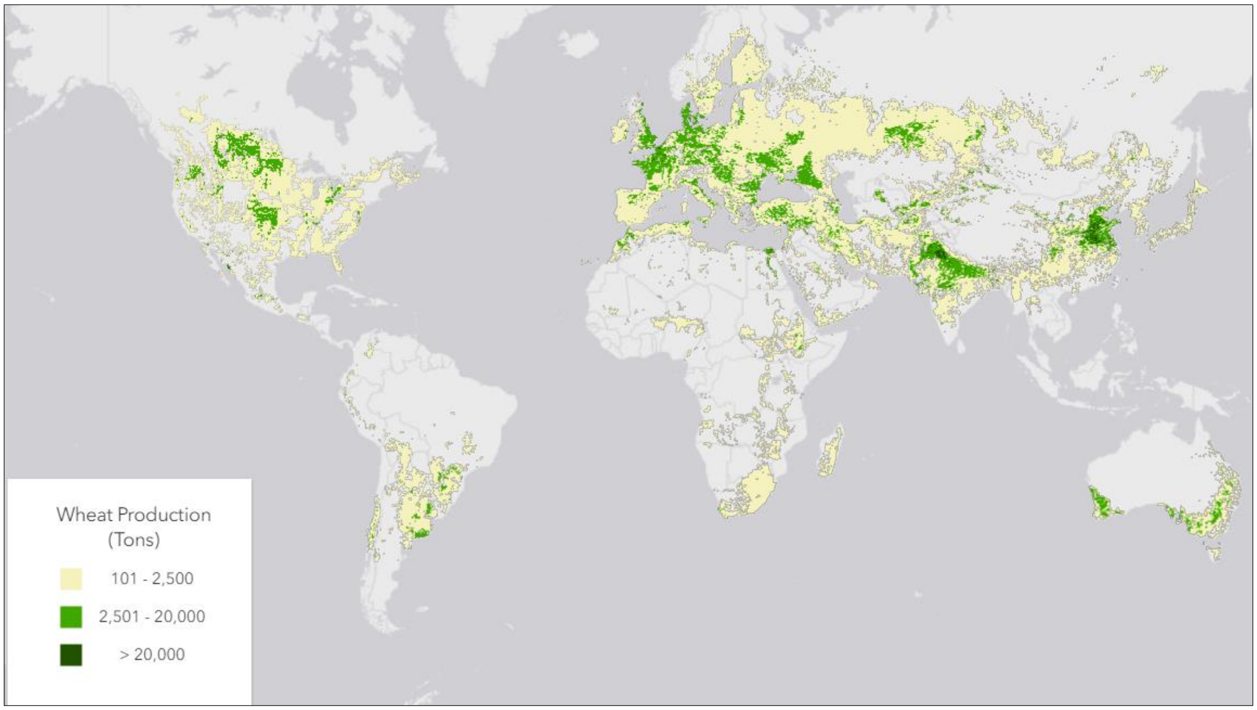


Figure 1: Wheat production in tonnes. (Sourced: USDA Wheat Explorer)

1.2 Role of carbon and nitrogen.

Carbon and nitrogen are important nutrients for plant growth, influencing structural integrity and lodging resistance. Carbon, obtained through photosynthesis, is important to produce cellulose and lignin, which strengthen plant cell walls and enhance stress tolerance. Insufficient carbon reduces cell wall strength and plant rigidity, increasing vulnerability to environmental stresses like wind and heavy rainfall (Li et al., 2023). Nitrogen, absorbed from the soil, is essential for synthesizing amino acids, proteins, and chlorophyll. However, excessive nitrogen can reduce stem quality and increase lodging risk by accelerating stem growth beyond the support capacity of stems and roots. Research indicates that high nitrogen levels decrease stem resistance to lodging and disrupt the allocation of photosynthetic carbon, impacting yield formation and stem quality (Wu et al., 2023).

The interaction between carbon and nitrogen cycles is complex, affecting plant physiology, growth, and ecosystem dynamics. Elevated nitrogen levels can enhance carbon sequestration but may also disturb the carbon-nitrogen balance, reducing photosynthetic rates and overall growth (Sinto et al.,2022). High nitrogen fertilization increases lodging risk in crops like rice and Italian ryegrass, impacting yield and stem strength. Appropriate nitrogen management strategies can mitigate lodging without reducing yield, emphasizing the need for balanced nutrient management (Svečnjak et al., 2020; Chauhan et al., 2020; Mizuta et al., 2023).

Understanding the coordination of carbon and nitrogen is crucial for optimizing plant health and productivity, as imbalances can significantly affect plant physiology and growth (Saiz-Fernández et al., 2017; Bicharanloo et al., 2021).

1.3 Remote sensing to estimate plant carbon and nitrogen.

Various remote sensing methodologies and approaches have been used to assess plants' nitrogen and carbon composition. Method involving hyperspectral imaging and vegetation indices to track the nutritional status of plants and estimate the biophysical content of crops has been found to be effective (Ma et al., 2022). Additionally, crop biophysical and biochemical properties have been estimated by using multispectral and hyperspectral imaging techniques using satellite, aircraft, and drone platforms (Sharifi, 2020). The use of spectral reflectance data extracted from multispectral imagery has proven to be successful in predicting nitrogen and carbon content in citrus canopy (Liu et al., 2016). Various models, including partial least squares regression (PLSR), multiple linear regression (MLR), and support vector machine (SVM), are used to estimate the nitrogen and carbon content in plants (Liu et al., 2016).

Using remote sensing data to estimate plant carbon and nitrogen levels has both benefits and drawbacks (Silva et al., 2008). Multispectral remote sensing offers a broad assessment of vegetation status, but it lacks the precise spectral data required for accurate content estimation due to its restricted spectral resolution. Hyperspectral remote sensing is well-suited for conducting thorough examinations of certain absorption characteristics associated with carbon and nitrogen molecules. It enables the creation of accurate spectral indices for quantitative analysis (Lu et al., 2020; Cotrozzi et al., 2018).

While remote sensing can be used to assess plants' nitrogen and carbon content, it comes with certain obstacles and limitations. One obstacle arises from the requirement to precisely calibrate and validate remote sensing models for specific geographic areas. This is because various vegetation types and environmental factors can influence the correlation between the spectral reflectance of nitrogen and carbon content (Ozdogan et al., 2010). The spectral resolution of remote sensing data is limited, which means it may not be able to detect slight differences in nitrogen and carbon content. This limitation is particularly evident when using broadband images (Wu et al., 2023). The resolution of remote sensing data in space and time may not always match the scale and timing of nitrogen and carbon changes in plant canopies. Additionally, factors that can cause confusion, such as background factors and canopy structure, can introduce uncertainties when estimating nitrogen and carbon content (Machwitz et al., 2021). Notwithstanding these difficulties, remote sensing presents a comprehensive and non-invasive method to assess the nitrogen and carbon levels in plants, offering vital information for enhancing nitrogen and carbon management in agriculture (Kumar et al., 2019).

Remote sensing has been used to detect and characterise crop lodging, offering various methodologies and technologies to enhance agricultural monitoring and management. UAV imagery has been used to identify lodging in wheat (Zhao et al., 2020; Yu et al., 2022), highlighting the capability of high-resolution UAV images to capture detailed spatial information critical for assessing the extent and severity of lodging. Additionally, Chauhan et al. (2021) developed an innovative remote sensing technique using Synthetic Aperture Radar (SAR) data to map the vulnerability of wheat to lodging, with SAR's ability to penetrate cloud cover and operate independently of weather conditions being particularly valuable for consistent monitoring over large areas. Beyond these examples, a broad spectrum of research explores different remote sensing technologies for crop lodging detection, including multispectral and hyperspectral imaging (Lowe et al., 2017), which detect changes in crop canopy structure and reflectance characteristics associated with lodging. LiDAR has also been used to assess crop height and biomass, providing precise measurements of crop displacement and canopy deformation (Li et al., 2017), making LiDAR an important remote sensing method that can be used in detection of lodging. The integration of machine learning algorithms with remote sensing data further improves the accuracy and efficiency of lodging detection and characterization (Zhang et al., 2020). Hyperspectral remote sensing, this technology offers a more detailed spectral resolution compared to multispectral imaging, enabling the detection of subtle biochemical and physiological changes in crops (Pascucci et al., 2020), and it can be effective in identifying early signs of stress and lodging before they become visually apparent, thereby providing a proactive approach to crop management and intervention strategies.

Hyperspectral imaging spectroscopy is a remote sensing technique that captures detailed spectral information across a wide range of wavelengths, compared with multispectral imaging, with a few bands, hyperspectral data offers a continuous spectrum (Prabira et al., 2022). Hyperspectral data offers detailed spectral information in diverse applications such as agriculture and environmental monitoring (Darvishzadeh et al, 2006; 2010; Pascucci et al., 2020). Numerous studies have utilized hyperspectral data to evaluate various biophysical parameters, including biomass (Wang et al., 2017), leaf area index (Darvishzadeh et al, 2008), vegetation cover (Darvishzadeh et al, 2007; 2009; Chanchí et al., 2023), canopy height (Miraki & Sohrabi, 2022), soil moisture (Jiang et al., 2022), and biochemical parameters such as chlorophyll content (Inoue et al., 2016), pathogen or plant stress detection (Gold et al., 2020; Galieni et al., 2020), carotenoids (Falcioni et al.,2023), mineral content (Wang et al., 2022), nutrient content (Zhang et al., 2023), plant water content (Zhang & Zhou, 2019), and lignin/cellulose (Feng et al., 2018). Preprocessing methods, such as noise removal and atmospheric correction, play a crucial role in hyperspectral measurements to enhance the precision of information extraction (Li et al., 2021).

Other remotely sensed techniques are also important, for example Sentinel-2 mission, operated by the European Space Agency, consists of twin satellites (Sentinel-2A and Sentinel-2B) equipped with multispectral sensors. These satellites provide valuable data for agricultural research. Sentinel-2 offers frequent revisit times (every 5 days) and high spatial resolution (10–20 meters). The spectral bands of Sentinel-2 exhibit sensitivity to many vegetative features, rendering it a powerful tool for evaluating biochemical factors in crop plants (Panwar et al., 2020; Segarra et al., 2020; Mahathi et al.,2023). The use of Sentinel-2 data in differentiating between healthy and lodged wheat, as well as in promptly identifying the occurrence and intensity of lodging, has been demonstrated (Chauhan et al., 2020). In addition, Sentinel-2 data can be integrated with other data sources, such as Sentinel-1 (Chauhan et al., 2020), drone imagery (Li et al., 2021), hyperspectral measurement (Mielke et al., 2014; Transon et al., 2017), used in abiotic and biotic stress detection (Pereira-Pires et al., 2021) to leverage the distinct capabilities of various sensors in estimating diverse plant properties. Combining remotely sensed data with ground-based measurements ensures accurate results and strengthens the reliability of agricultural assessments (Farbo et al., 2022).

1.4 Vegetation indices

Vegetation indices are important for assessing various aspects of plant health. The Normalized Difference Vegetation Index (NDVI) leverages the absorption of visible light and reflection of near-infrared light by healthy vegetation to provide a measure of vegetation health, with values ranging from -1 to 1. NDVI is instrumental in evaluating vegetation health, monitoring drought, and estimating biomass due to its robustness in indicating photosynthetic activity (Rouse et al., 1974; Zapata et al., 2023). The Enhanced Vegetation Index (EVI) improves upon NDVI by reducing atmospheric influences and enhancing sensitivity in high biomass regions, calculated using an advanced formula incorporating blue band. EVI is particularly effective in densely vegetated areas, facilitating accurate monitoring of plant vigour and biomass production (Huete et al., 2002). The Cellulose Absorption Index (CAI) estimates cellulose concentration by measuring the absorption characteristics of cellulose (Daughtry, 2001), which is a major component of plant cell walls, and aiding in crop health assessment, disease stress detection, and fertilizer optimization in precision agriculture. Similarly, the Lignin Cellulose Absorption Index (LCAI) provides insights into both lignin and cellulose content, further refining the understanding of plant structural composition and health (Daughtry et al., 2005). The Normalized Difference Lignin Index (NDLI) focuses on nitrogen and lignin content (Serrano et al., 2002), which is inversely related to forage quality, and is useful for assessing plant quality. The Plant Senescence Reflectance Index (PSRI) indicates the onset of plant senescence by measuring changes in leaf pigment concentrations, crucial for managing harvest timing and crop rotation strategies (Merzlyak et al., 1999). Lastly, the Ratio Vegetation Index (RVI) uses the simple ratio of near infrared to red band reflectance, offering an alternative method for evaluating vegetation density and health (Pearson & Miller, 1972). Together, these indices provide comprehensive tools for foliar carbon and nitrogen content monitoring in precision agriculture, giving a clearer view in vegetation health monitoring.

1.5 PLSR in remote sensing

In remote sensing, various statistical and machine learning techniques are employed to analyze and interpret complex datasets, each with distinct advantages and limitations. Partial Least Squares Regression (PLSR) is particularly notable for handling multicollinearity effectively and reducing dimensionality by extracting latent variables, making it well-suited for datasets with many highly correlated spectral bands (Jin & Wang, 2019). This makes PLSR advantageous for predicting vegetation indices, soil properties, and water quality. In contrast, Multiple Linear Regression (MLR) is simpler and computationally efficient but struggles with multicollinearity and high-dimensional data. Support Vector Machines (SVM) and Random Forest (RF) offer robust handling of non-linear relationships and high-dimensional spaces, with RF providing insights into feature importance and SVM being less prone to overfitting, though both can be computationally intensive. Neural Networks (NN) excel in modeling complex, non-linear relationships and are highly flexible but require large datasets and significant computational resources, often lacking interpretability. PLSR is frequently chosen for its ability to manage multicollinearity and dimensionality reduction in spectral data, essential for remote sensing applications, though the choice of technique should ultimately depend on the specific task requirements, data characteristics, and available computational resources (Cheng & Sun, 2017; Shen et al., 2020; Liu et al., 2017).

The above literature review demonstrates that, there has been no research conducted on the assessment of nitrogen and carbon contents in healthy and lodged crops using hyperspectral data. Additionally, no studies have yet utilized Sentinel-2 data to monitor the relationship between crop carbon and nitrogen content in lodged and healthy wheat. This relationship is crucial for the assessment of plant productivity, plant health, quality of grains, and ensuring food security. The outcome of this study is anticipated to enhance precision agriculture methods by offering well-informed remedies for crop lodging resulting from excessive nitrogen levels and inadequate carbon levels in wheat plants. This will enable farmers to develop more effective management plans to prevent lodging and subsequent reductions in crop yield.

1.6 Problem statement

Lodging presents a significant obstacle for the agriculture industry, for it leads to substantial reductions in both yield, crop quality, optimal crop management and yield prediction, as it alters the structural and spectral properties of plants. Crop lodging primarily occurs in cereal crops, including wheat, maize, barley, and rice. The lodging of crops is influenced by various factors, with nitrogen and carbon levels playing crucial roles in the strength and stability of plant stems. Excessive nitrogen usage has been identified as a major contributor to the heightened risk of lodging, whereas the presence of carbon is crucial to produce structural elements that enhance plant stiffness and ability to withstand environmental pressures.

Previous studies highlight the complex relationship between nitrogen, carbon, and lodging in crops and that an uneven distribution of nitrogen fertilizer can decrease the ability of plant stems to withstand stress, hence compromising their structural strength (Li et al., 2023). Moreover, the harmonization and equilibrium of carbon and nitrogen concentrations in plants are crucial for attaining maximum growth, advancement, ability to withstand stress and yield. This underscores the importance of comprehending the different types of plant nutrient interactions, especially when it comes to lodging in wheat crops, thereby helping in plant resilience, lodging resistance and yield prediction.

Despite the use of remote sensing techniques in different research for the evaluation of nitrogen or carbon content in plants, there remains a notable gap in the exploration of remote sensing data for assessing the relationship between nitrogen and carbon contents in healthy and lodged wheat canopies, to overcome the challenges that lies in accurately monitoring and managing nutrient levels in lodged wheat to mitigate economic losses, enhance yield prediction and optimize crop health. Therefore, the problem addressed in this study is how to effectively utilize hyperspectral and Sentinel-2 data to accurately estimate and retrieve nitrogen and carbon content in lodged and healthy wheat, thereby improving crop health monitoring and reducing the adverse effects of lodging on wheat yield.

1.7 Aim and objectives of the study.

The objective of the study is to examine the relationships between lodging occurrence and severity, and variations in nitrogen and carbon content in wheat using remote sensing data. To evaluate the influence of lodging on nitrogen and carbon variations, hyperspectral measurements and Sentinel-2 data will be used in this study.

1.7.1. Specific objectives

1. To assess the relationships between spectral reflectance and nitrogen and carbon content in lodged and healthy (un-lodged) wheat canopy, using the field hyperspectral measurements.
2. To analyse the relationships between lodging and nitrogen and carbon variabilities in wheat canopy, obtained using the Sentinel-2 data (S2 data).

1.7.2 Research questions and hypothesis

1a. What is the relationship between lodging and nitrogen content in wheat canopy estimated using the field hyperspectral measurements?

Ho: There are no specific hyperspectral bands that accurately estimate nitrogen content within the lodged wheat canopy.

Ha: There are specific hyperspectral bands that accurately estimate nitrogen content within the lodged wheat canopy, showing a significant relationship with changes in nitrogen levels due to lodging.

1b. What is the relationship between lodging and carbon content in wheat canopy estimated using the field hyperspectral measurements?

Ho: *Ho:* There are no specific hyperspectral bands that accurately estimate carbon content within the lodged wheat canopy.

Ha: There are specific hyperspectral bands that accurately estimate carbon content within the lodged wheat canopy, showing a significant relationship with changes in carbon levels due to lodging.

2a. To what extent does lodging affect the retrieval of nitrogen in wheat canopies using S2 data?

Ho: There is no significant impact of lodging on the retrieval of nitrogen in wheat using Sentinel-2 data. The spectral bands identified for nitrogen can retrieve nitrogen despite lodging in wheat.

Ha: Lodging significantly affects the retrieval of nitrogen in wheat using Sentinel-2 data. The spectral bands identified for nitrogen cannot retrieve nitrogen in lodged wheat.

2b. To what extent does lodging affect the retrieval of carbon in wheat canopies using S2 data?

Ho: There is no significant impact of lodging on the retrieval of carbon in wheat canopies using Sentinel-2 data. The spectral bands identified for carbon can retrieve carbon despite lodging in wheat.

Ha: Lodging significantly affects the retrieval of carbon in wheat canopies using Sentinel-2 data. The spectral bands identified for carbon cannot retrieve carbon in lodged wheat.

2. CHAPTER 2: METHODOLOGY

2.1 Study area

The research site is situated at the Bonifiche Ferraresi farm, Jolanda di Savoia, Ferrara, Italy (44°52'59"N, 11°58'48"E). Bonifiche Ferraresi is a major agri-food enterprise in Italy, covering over 6500 hectares (Figure 2). The primary agricultural produce includes durum wheat (*Triticum durum*), soft wheat (*Triticum aestivum*), rice (*Oryza sativa*), corn (*Zea mays*), barley (*Hordeum vulgare*), soybean (*Glycine max*), and potatoes (*Solanum tuberosum),* with several other horticultural and medicinal species. These crops are commonly cultivated in a sequential rotation in successive years. The primary crop is wheat (field sizes range from 2.38 to 84.86 hectares) predominantly grown in clayey and silty soils. The climate of the region is mild and temperate.

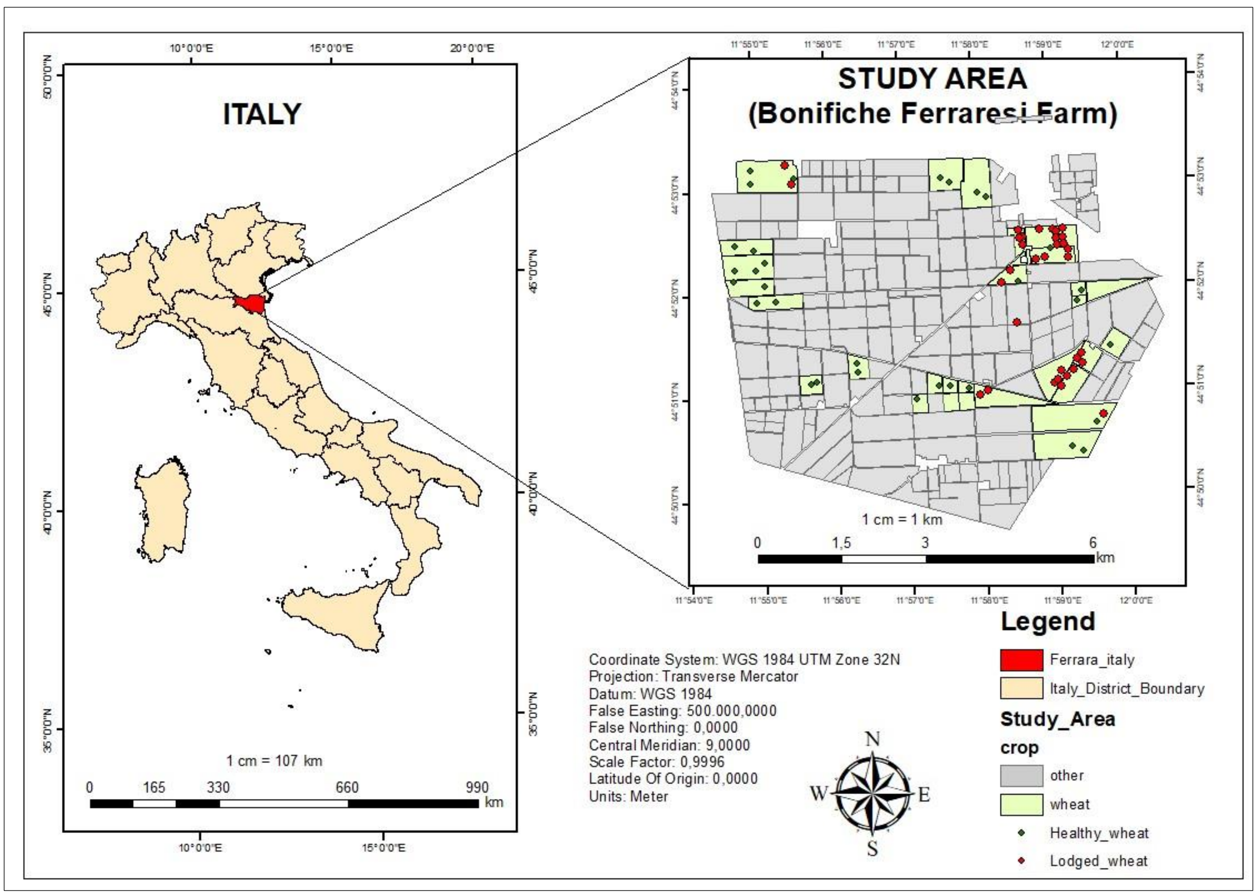


Figure 2: Location of the study area - Bonifiche Ferraresi farm in Jolanda di Savoia, within the Ferrara region of Italy (Zoom-in). (Source: Author).

2.2 Flowchart of method

Carbon and nitrogen content estimation and retrieval performed in this research followed detailed steps shown in the methodological flowchart (Figure 3) and steps taken to achieve the set objectives according to the type of data used.

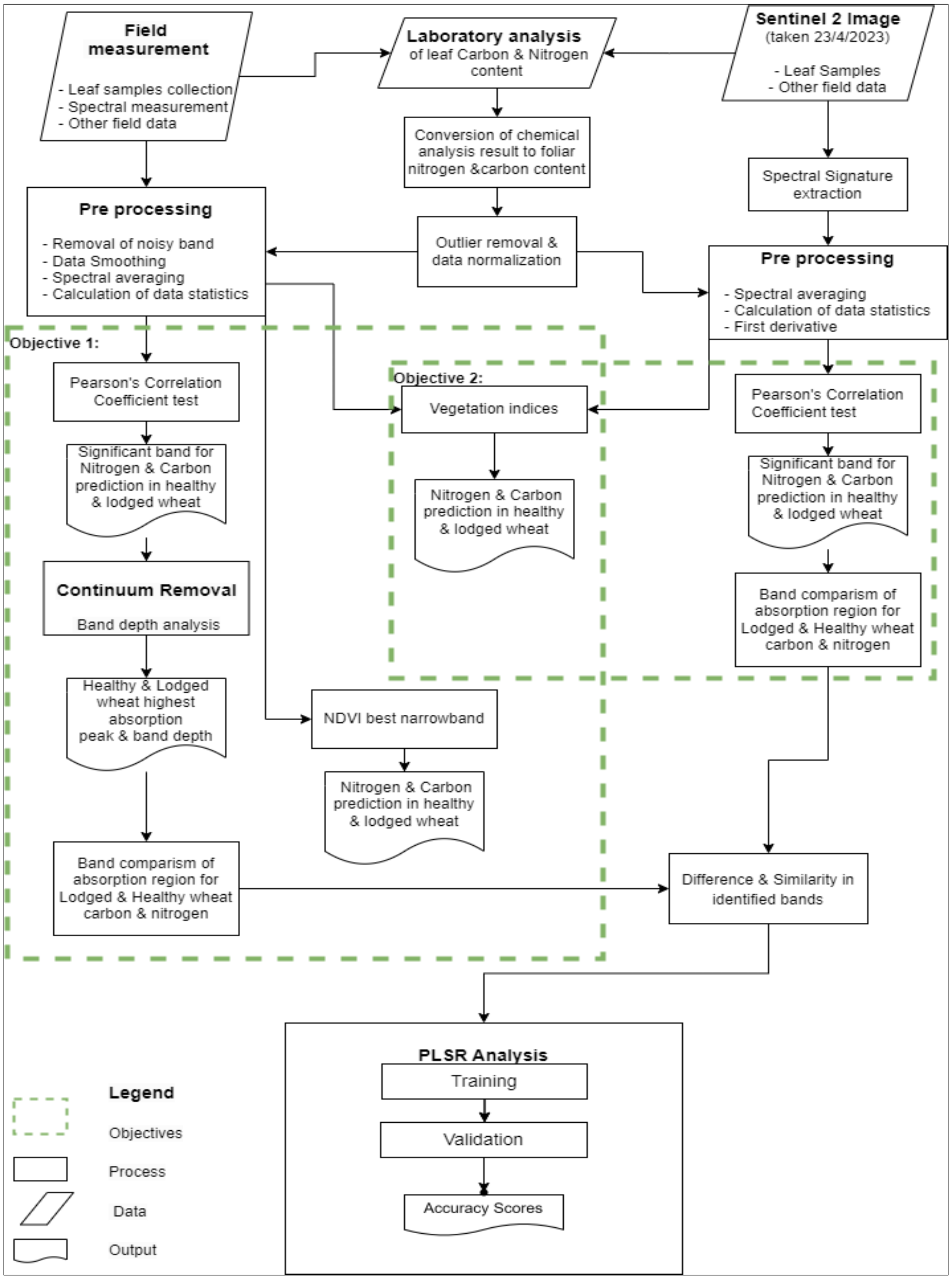


Figure 3: Methodological flowchart of this research.

2.3 Data collection

2.3.1 Sampling technique

Stratified random sampling was used for sampling at three levels (Figure 3): sampling at plot level (90m × 90m), subplot level (15m × 15m), and microplot level (1m × 1m). A total of 65 plots were sampled, out of which 33 were lodged wheat plots and 32 were healthy wheat plots (Table 1). Crop biophysical parameters were collected during the field campaign held between 30 April and May 18, 2023. The parameters that were measured in the field include slant height (cm), lodge area (%), SPAD, plant density, leaf area index (LAI), vertical lodge height (cm), fresh biomass (t/ha), leaf area (cm^2), canopy height (cm), cover percentage (%), shoot numbers, point of line failure (cm) and tillers. Slant height (cm), vertical lodged height was used to assess lodging occurrence and severity in the field. Fresh wheat plant samples from one plant within the micro plot were collected, wrapped with wet paper, and kept in a Ziplock bag to preserve the sample for subsequent leaf area and fresh weight measurements.

Table 1: The number of sample plots at three different levels.

	Plot level	Subplot level	Microplot level
Lodged wheat	33	165	498
Healthy wheat	32	157	470
Total	**65**	**322**	**968**

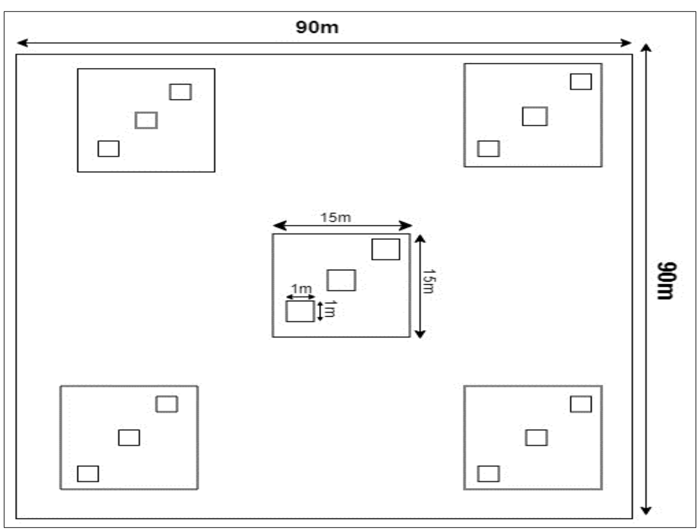


Figure 4: The Diagram of Stratified Random Sampling

2.2.2 *Data selection*

The samples used for the hyperspectral part of this study were at microplot level, which was initially 968 samples but were reduced to 163 samples and at subplot level from 322 samples to 305 samples (Table 1 & 2) due to several constraints. These constraints included the high cost of laboratory analysis, the unavailability of some spectral measurements at the microplot level, and practical issues such as missing field values, incomplete leaf samples needed for laboratory analysis, incorrect labelling, and spoilage of samples before drying due to adverse weather and distance to the laboratory. Consequently, the sample size was significantly reduced, impacting the overall data quality and quantity. For the Sentinel-2 part of the study, samples were collected at the subplot level. However, some subplots were excluded from the analysis due to incorrect location details that fell outside the study area. The total amount of sample data used in the study is summarized in Table 2.

Table 2: The number of healthy and lodged wheat used at hyperspectral, and Sentinel-2 level.

	Subplot level	Microplot level
Lodged wheat	159	93
Healthy wheat	146	70
Total	**305**	**163**

2.2.3 *Field spectral measurement*

An ASD FieldSpec-3 Full-Range (350–2500 nm) spectrometer (Analytical Spectral Devices, Boulder, CO, USA) was used to collect canopy hyperspectral measurements in 74 healthy and 115 lodged micro-plots and 22 healthy and 40 lodged subplots. The spectroradiometer was calibrated against the reference white panel before taking the canopy spectral measurements in each micro plot. Five canopy spectral measurements were taken at each micro-plot and ten readings were recorded to minimise external noise for each measurement.

2.2.4 *Laboratory analysis*

The fresh samples were transferred into labelled paper envelops and were dried in an oven set to 65 degrees for 48 hours in ITC laboratory. The dried samples were ground with coffee grinders and the powder was kept in labelled envelopes. The envelopes were kept in airtight containers until further process commenced. A Perkin Elmer CHN Analyzer-2400 Series was used to quantify nitrogen and carbon content in the powdered samples of wheat crops. A detailed description of the methodology used by this instrument is given by Steve (2015). In brief, to prepare samples, clean all weighing utensils and use tweezers to handle tin capsules. Weigh an empty capsule, tare it using another empty capsule on the Perkin Elmer balance, then add the sample. Weigh again, adjust as needed, and note the weight, which should be between 1.700mg to 1.800mg. Clean the preparation area between samples. When the capsule has the correct sample amount, flatten, and fold it into a C shape, re-weigh, record details, and store in a clean tray, avoiding cross-contamination.

Figure 5: The sorting process of the sample before it is weighed and placed in the CHN Analyzer-2400 Series (behind the author, at the left side).

Before starting the machine good lighting in ensured in the laboratory and the gases (hydrogen, helium, and compressed air) needed for the procedure is checked to be supplied to the machine. The instrument was set on 'STANDBY'; operational mode by pressing 'PARAMETERS', turning off Gas Saver, and verifying combustion, reduction, and detector oven temperatures (925°C, 640°C, 82.6°C, respectively) are within limits. If correct, I proceed with purging gas lines to remove diffused N_2 using specified durations for Helium and Oxygen. While purging, the PC was started, and I log in to EA 2400 Data Manager, and followed on-screen instructions, entering 'PARAMETERS', and confirming values as prompted. This setup ensures the instrument is ready for use. To save results, a folder named with the analysis date (e.g., 20240130) was created. Excel file is opened and filled in the necessary information in "Logbook CHN." The oxygen valve is then switched off ('PARAMETERS' > 20). A blank sample is run ('SINGLE RUN' > 1 BK), during which the sample drop area is clean. After running helium-only blanks, the oxygen valve was switched back on and repeat the blank run.

The machine is condition using the acetanilide standard. For sample analysis, 'SINGLE RUN' or 'AUTO RUN' modes are used, entering sample ID and weight. After analysis, the system is switched to Gas Saver Mode and data output is saved by clicking "Sync to Storage."

2.2.5 Conversion of nitrogen and carbon content to area-based measurements

Nitrogen and carbon content that was measured in the laboratory, was in percentage grams of the dry weight of the leaf sample. This is multiplied by leaf mass area in grams per centimeter squared to convert to foliar nitrogen content at the plot level. The same was done with carbon content using the formula.

$$\text{foliar nitrogen}(g/cm^2) = \text{leaf nitrogen } (\%) * \text{LMA}(g/cm^2)$$

Where LMA is leaf mass per area.

2.2.6 Data preprocessing for hyperspectral measurements

The raw field hyperspectral data collected with the spectrometer on the field undergo a series of essential steps to enhance data quality and prepare it for subsequent analysis; removal of the noisy band affected by water and atmospheric region; 350nm and 399 nm, 1330nm -1406nm, 1810nm – 1952nm and above 2351nm were also removed. A total of 764 bands were removed while the remaining 1736 band were used in the analysis. The 100-reflectance data collected each microplot were averaged; to reduce the measurement noise and this represent a sample. Savitzky Golay filter (Savitzky and Golay, 1964) was used to smooth the data with a frame size of 9 (1^{st} degree polynomial) was used, this data was used for further analysis. The reflectance data for healthy and lodged wheat are shown in Figure 5. Healthy and lodged wheat spectral was used to calculate the mean, minimum and maximum reflectance, and their correspondent 1^{st} derivative (Figure 5) which are important for feature differentiation in the field hyperspectral data for wheat.

2.2.7 Sentinel-2 reflectance acquisition

Sentinel-2 satellite imagery - atmospherically and geometrically corrected - covering the study area and was taken on 23^{rd} April 2023 was used to assess the wheat fields for lodged and healthy crops. The image was used during the extraction of the mean reflectance of lodged and healthy wheat on ENVI Classic 5.7 + IDL 8.9 using NRS toolbox. The data were then used for further analysis. The data were explored to show the difference in the spectral signature of the lodged and healthy wheat considering the maximum, mean, and minimum reflectance. Then the 1^{st} derivative was also calculated to show the difference between them and this help to show the likely bands that can be used to differentiate between the lodged and healthy wheat.

2.2.8 Outliers removal and normalization of data

Median Absolute Deviation (MAD) method was used to remove outlies in the data. The MAD is less affected by extreme values (leys et al., 2013). This further reduced the used data at the subplot level, healthy wheat are 109 samples, lodged were 105 samples and at microplot level, healthy wheat are 50 samples, lodged were 82 samples. To ensure that the data was standardized and comparable across different scales, z-score normalization to the carbon and nitrogen content was applied both in field hyperspectral and Sentinel-2 measurements. This process involved standardizing the input vector by subtracting the mean and dividing by the standard deviation of the data. For example, the normalization function used for lodged carbon was defined as 'normal_lc =

@(carbon) (carbon - mean(carbon)) / std (carbon)'. This transformation centered the data around zero and scaled it to have a unit variance, ensuring that each feature contributed equally to the analysis. The normalized carbon levels were then used in subsequent analyses to improve the robustness and performance of the predictive models.

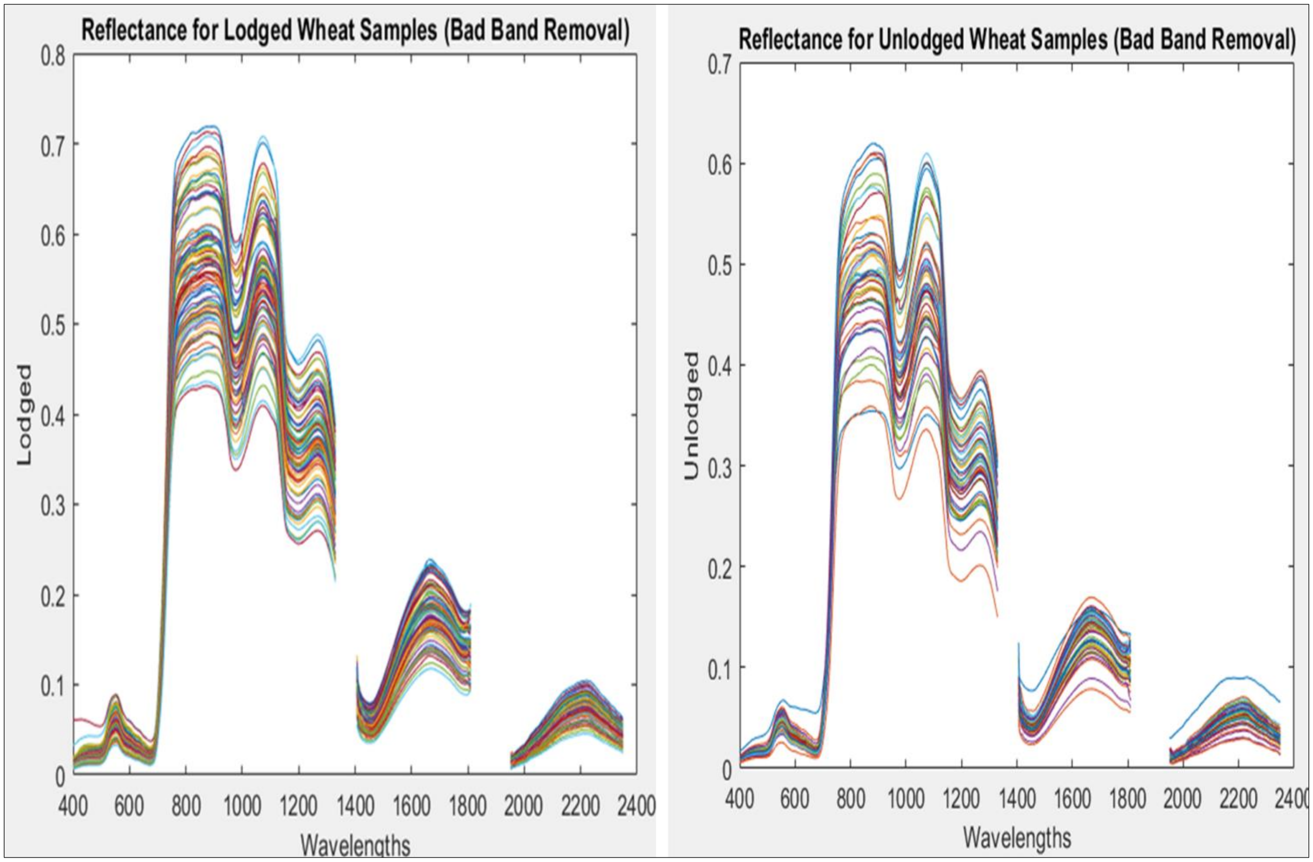


Figure 6: (a) Mean reflectance of lodged-left wheat (n=82), (b) Mean reflectance of healthy wheat – right (n=50). The image was taken after all the preprocessing steps.

2.3 Data analysis

Pearsons's correlation coefficient test was used to analyse the healthy and lodged wheat mean reflectance, to determine if there is a significant difference in the correlation. The p-value threshold was set at 0.05 and any correlation that has a p-value greater than the set threshold was considered not significant. This help to identify the spectral bands or regions that can estimate carbon or nitrogen content in both healthy and lodged wheat.

Continuum removal was applied to the hyperspectral data to analyze specific absorption features in the last step that is related to nitrogen and carbon content in wheat. This technique normalizes the reflectance spectrum by fitting and removing a continuum line, thereby highlighting the absorption bands more clearly, in doing so, there is a better correlation of the spectral features with biochemical properties of the plants. To isolate specific absorption features for analysis, the reflectance spectra were divided by the continuum at each wavelength, resulting in the continuum-removed spectra (RCR = R/C). This transformation normalizes the reflectance spectra, with values ranging from 0 to 1, where the first and last points become 1, highlighting absorption features. Band depth (BD) was calculated by subtracting the continuum-removed reflectance (R') from 1.

Vegetation indices (VI) is a valuable tool for non-destructive and quick assessments of plant health, biophysical properties, and physiological characteristics. Vegetation indices such as the Normalized Difference Vegetation Index (NDVI) and Enhanced Vegetation Index (EVI), are commonly used to generate a prediction model for canopy nitrogen content (Zhonglin et al.,2022). In this research, VI was used in monitoring carbon and nitrogen variation in wheat leaf, using the reflectance of the field hyperspectral and Sentinel-2 data. Narrowband indices were calculated using the spectral information from the hyperspectral data, finding the best band combinations that can be used to predict carbon and nitrogen content using vegetation indices. The indices for nitrogen content estimation are normalised difference vegetation index (NDVI), enhanced vegetation index (EVI), and ratio vegetation index (RVI), while the indices for carbon content estimation are the cellulose absorption index (CAI), plant senescence reflectance index (PSRI), Lignin Cellulose Absorption Index (LCAI), and normalized difference lignin index (NDLI)

Table 3: The vegetation indices used in this study.

Vegetation indices	**Formula**	**Purpose**
Normalised Difference Vegetation Index (NDVI)	$NDVI = \frac{NIR - Red}{NIR + Red}$	NDVI is associated with healthy vegetation, and it is an indication of nitrogen content. Nitrogen shows absorption in the red and the near infrared bands (Rouse et al., 1974).
Enhanced Vegetation Index (EVI)	$\text{EVI} = G * \frac{(\rho_{NIR} - \rho_{red})}{(\rho_{NIR} + C1 * \rho_{red} - C2 * \rho_{blue} + L)}$	EVI is sensitive to high vegetation cover, less saturation, reduced atmospheric and contamination due to canopy background (Huete et al., 2002).
Plant Senescence Reflectance Index (PSRI)	$\frac{678\text{nm} - 500\text{nm}}{750\text{nm}}$	PSRI is designed to detect plant senescence and stress by identifying changes in leaf pigments, specifically the decrease in chlorophyll and the increase in carotenoids. It is useful for understanding the physiological status of vegetation (Merzlyak et al., 1999).
Cellulose Absorption Index (CAI)	CAI = 0,5 (2020nm+2220) – 2100nm	Cellulose is an indication of carbon content in plant and CAI is sensitive

		to cellulose content (Daughtry, 2001)
Normalized Difference Lignin Index (NDLI)	$NDLI = \frac{\log\left(\frac{1}{\rho_{1754}}\right) - \log\left(\frac{1}{\rho_{1680}}\right)}{\log\left(\frac{1}{\rho_{1754}}\right) + \log\left(\frac{1}{\rho_{1680}}\right)}$	NDLI is an index to assess Nitrogen and lignin, which can be used to predict carbon content in plant (Serrano et al., 2002).
Ratio Vegetation Index (RVI)	RED/NIR	RVI is a simple index used in remote sensing to measure the relative amount of green vegetation. It is calculated as the ratio of the reflectance in the near-infrared (NIR) region to the reflectance in the red region. The RVI is sensitive to vegetation density and biomass (Pearson & Miller, 1972)
Lignin Cellulose Absorption Index (LCAI)	$\mathrm{LCAI} = \frac{R_{2220} - R_{2100}}{R_{2220} + R_{2100}}$	LCAI is a vegetation index specifically designed to assess lignin and cellulose content in plant material. Lignin and cellulose are two of the most abundant organic compounds in the biosphere and are significant carbon reservoirs in terrestrial ecosystems. It is used in monitoring vegetation health, maturity, and biochemical composition (Daughtry et al., 2005)

Optimal hyperspectral band combinations for estimating nitrogen and carbon content were assessed through the calculation of narrowband indices. The coefficient of determination between the calculated narrowband and carbon or nitrogen content was determined for each pair, resulting in an R^2 matrix. The corresponding wavelengths associated with the highest R^2 values were then considered as the best band combination.

To analyze the distribution and variability of carbon and nitrogen content in the lodged and healthy wheat, box plots was used. These plots provided a clear visual representation of the central tendency, spread, and potential outliers in the data, facilitating a comprehensive understanding of the nutrient status across the wheat samples. The box plots were generated using Matlab, ensuring accurate depiction of quartiles and interquartile ranges, thereby highlighting the variations of nitrogen and carbon content within the dataset.

In this study, we compared two cross-validation methods, venetian blinds, and randomized validation, to evaluate their impact on the retrieval accuracy of nitrogen content in wheat canopies using Sentinel-2 data. The venetian blinds validation method involves systematically partitioning the data into equally spaced segments, ensuring that each subset maintains the original sequence of the data. This method provides consistent and repeatable results, which is critical for reproducibility in sequential data analysis. In contrast, the randomized validation method involves shuffling the data before dividing it into training and test sets, aiming to minimize sampling bias and ensure that each subset is representative of the entire dataset. While randomized validation is typically preferred for its statistical robustness and ability to mitigate bias, it inherently introduces variability between runs due to the random shuffling process.

Partial Least Squares Regression (PLSR) had been used to develop a robust model for the data analysis. PLSR establish the relationship between biochemical concentrations and reflectance spectra (Darvishzadeh et al. 2008c). PLSR combines the features of principal component analysis (PCA) and multiple regression, compressing many variables into a few latent variables (PLS factors). This approach incorporates the response variable information during decomposition, effectively reducing overfitting. In this study, spectra data and normalised data of carbon and nitrogen content were analyzed using PLS Toolbox on Matlab. The lodged and healthy wheat had 82 and 50 samples respectively at the field level while at the satellite level, lodged and healthy wheat had 105 and 109 samples respectively. The response variable was mean-centred and PLSR model was developed using the SIMPLS algorithm, with the number of latent variables determined through cross-validation. For validation, Venetian blinds cross-validation method was used with varied numbers of splits and a blind thickness of 1. In this approach, the dataset was systematically divided into segments, ensuring that every sample was used for validation exactly once while maintaining the spatial structure of the data. This method helps in reducing the potential bias and variance, providing a more reliable estimate of the model's performance. Model validation was conducted through venetian blinds, where the coefficient of determination for cross-validated (R^2 CV), cross validation bias and root mean square error cross validation (RMSECV) between predicted and observed values were calculated to evaluate model predictive performance.

CHAPTER 3: RESULTS

3. Results

Hyperspectral data statistics are shown in Table 4; the detailed information on the healthy dataset, showing that nitrogen ranges between 2.95 and 6.75 g/m² and 69,22 to 242,07 g/m² for carbon, while lodged nitrogen and carbon ranges between 2.82 to 10.38 g/m² and 61.97 to 263.48 g/m² respectively.

Table 4: Statistical summary of nitrogen and carbon in healthy and lodged wheat at the field level (hyperspectral)

Hyp_Variables	Unit	Minimum	Maximum	Mean	Standard deviation
Healthy Nitrogen	g/m²	2,95	6,75	5,23	1,02
Lodged Nitrogen	g/m²	2,82	10,38	6,15	1,95
Healthy Carbon	g/m²	69,22	242,07	149,41	37,28
Lodged Carbon	g/m²	61,97	263,48	141,11	50,66

Sentinel-2 data statistics were shown in Table 5 below; the detailed information on the healthy dataset, showing that nitrogen ranges between 2.65 and 8.57 g/m² and 104.58 to 269.77 g/m² for carbon, while lodged nitrogen and carbon ranges between 1.98 g/m² to 12.89 g/m² and 87.38 to 295.10 g/m² respectively.

Table 5: Statistical summary of nitrogen and carbon in healthy and lodged wheat at satellite level (Sentinel-2).

S2_Variables	Unit	Minimum	Maximum	Mean	Standard deviation
Healthy Nitrogen	g/m²	2,65	8,57	5,82	1,29
Lodged Nitrogen	g/m²	1,98	12,89	7,73	2,11
Healthy Carbon	g/m²	104,58	269,77	176,34	35,66
Lodged Carbon	g/m²	87,38	295,10	181,55	57,26

3.1 Mean spectral reflectance of lodged and healthy wheat.

In the mean reflectance of lodged and healthy wheat (Figure 7) there are distinct differences in the spectral signatures at the field level (hyperspectral - left) and satellite level (Sentinel-2- right), the different in reflectance was much due to broad band nature of the Sentinel-2 data and lots of bands which were present in the field hyperspectral data, which were missing in Sentinel-2, leading to missing information in some regions at the satellite level spectra (Figure 7 - right).

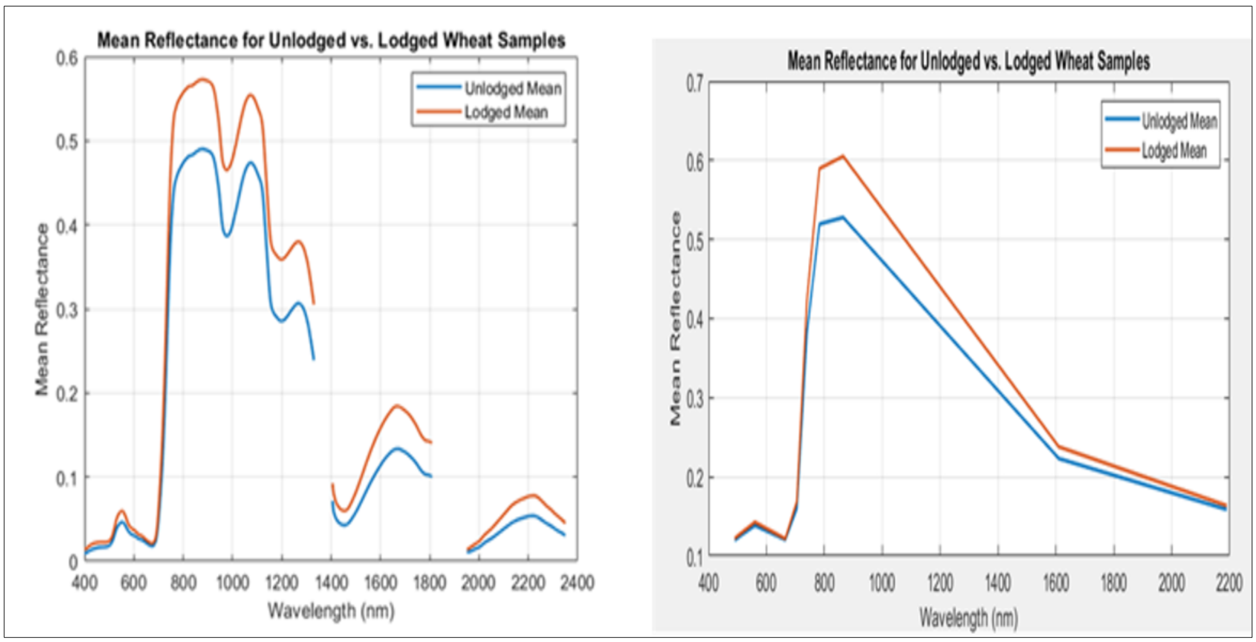


Figure 7: Average spectral signature of lodged and healthy wheat, on the left is the field hyperspectral measurement (lodged n=82, healthy n=50) and on the right is the satellite spectra obtained from Sentinel-2 imagery imagery of 23rd of April 2023 (lodged n=105, healthy n=109) in Bonifiche Ferraresi farm, Jolanda di Savoia, Ferrara, Italy.

Minimum, and maximum reflectance values was calculated at hyperspectral (Figure 8) and sentinel-2 (Figure 9) level, to enhance the differentiation between lodged and healthy wheat. Analyzing the mean, minimum, and maximum reflectance values allow for a comprehensive characterization of the spectral properties of both lodged and healthy wheat, helping to identify subtle differences that may not be apparent when only considering the mean reflectance. Minimum and maximum reflectance values provide insights into the variability within the wheat samples, essential for understanding the range of spectral responses and identifying any outliers or anomalies indicative of lodging. The first derivative of the reflectance values enhances the sensitivity of the analysis to changes in spectral reflectance, helping to identify inflection points and subtle shifts in the spectral curve critical for distinguishing between lodged and healthy wheat.

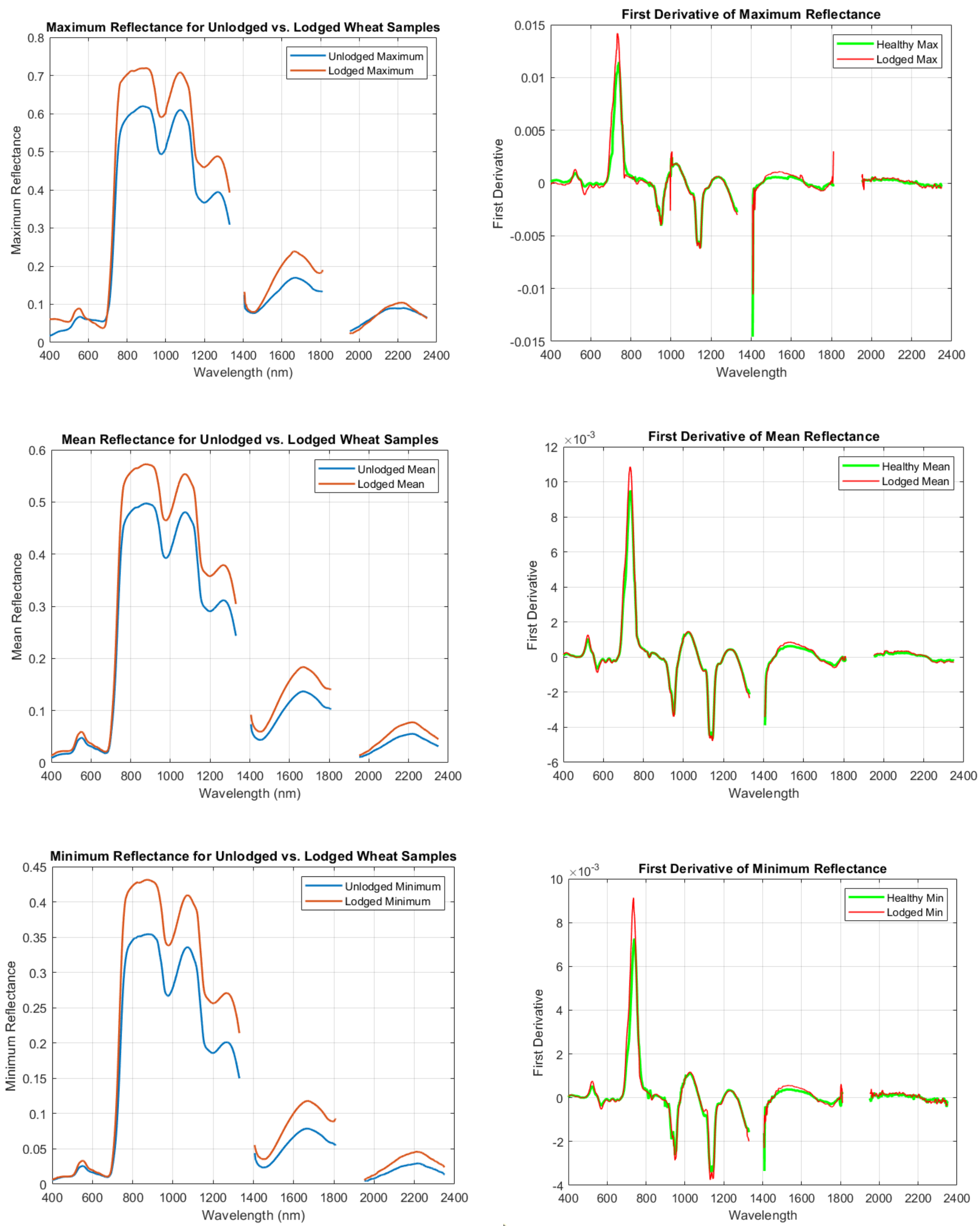


Figure 8: Using field hyperspectral (a)Maximum reflectance of lodged vs healthy wheat and the corresponding 1st derivative is shown beside it. (b)Mean reflectance of lodged vs healthy wheat and the corresponding 1st derivative is shown beside it (c) Minimum reflectance of lodged vs healthy wheat and the corresponding 1st derivative is shown beside it.

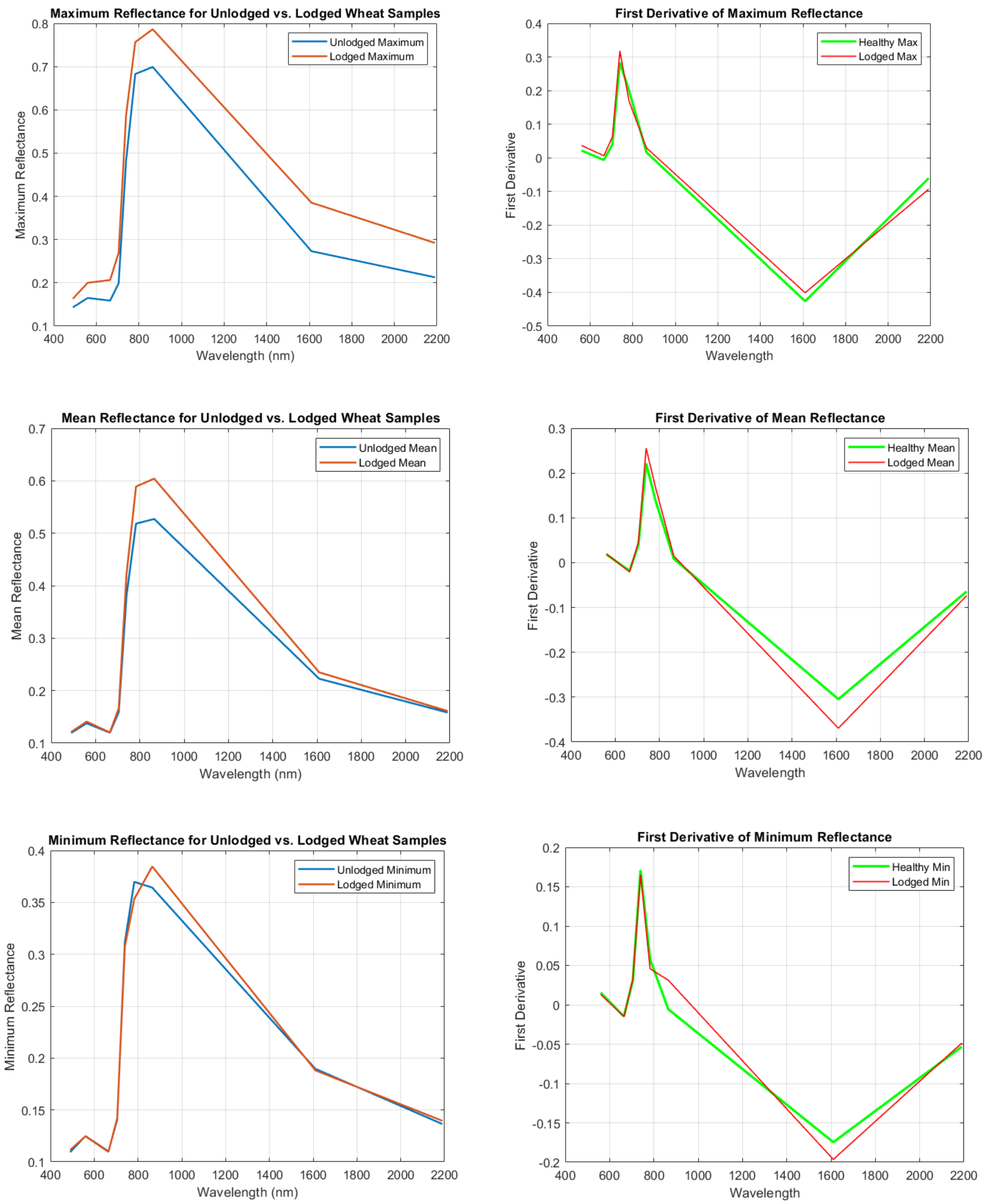


Figure 9: Using Sentinel-2 (a)Maximum reflectance of lodged vs healthy wheat and the corresponding 1st derivative is shown beside it. (b)Mean reflectance of lodged vs healthy wheat and the corresponding 1st derivative is shown beside it (c) Minimum reflectance of lodged vs healthy wheat and the corresponding 1st derivative is shown beside it.

3.2 Pearson’s correlation coefficient result

After conducting Pearson’s correlation coefficient to test the significance of difference in the lodged and healthy at a 0.05 confidence level on the field hyperspectral data across the entire spectral range, several spectral regions were identified where the spectral bands showed significant differences in carbon and nitrogen content of lodged wheat (Figure 10), while that of healthy wheat do not show any statistical correlation. These significant regions are presented in Table 6, from these spectral regions, specific absorption features that were used for further investigation in this study were identified. Sentinel-2 Pearson’s correlation result showed significant only in healthy wheat nitrogen, Figure 11 showed this, other parameters were correlated but not significant.

Table 6: Significant spectral regions at field hyperspectral level (left) which is only noticed in the lodged wheat spectra and at Sentinel-2 level (right) only in healthy wheat, no significant band is seen in healthy carbon and lodged carbon and nitrogen i.e. p-value < 0.05.

Nitrogen wavelength	Carbon wavelength
range (nm)	Range(nm)
1411-1558	1410-1538
1951-2350	1951-2229
	2266-2350

Nitrogen wavelength
Range (nm)
740
783
865
1610

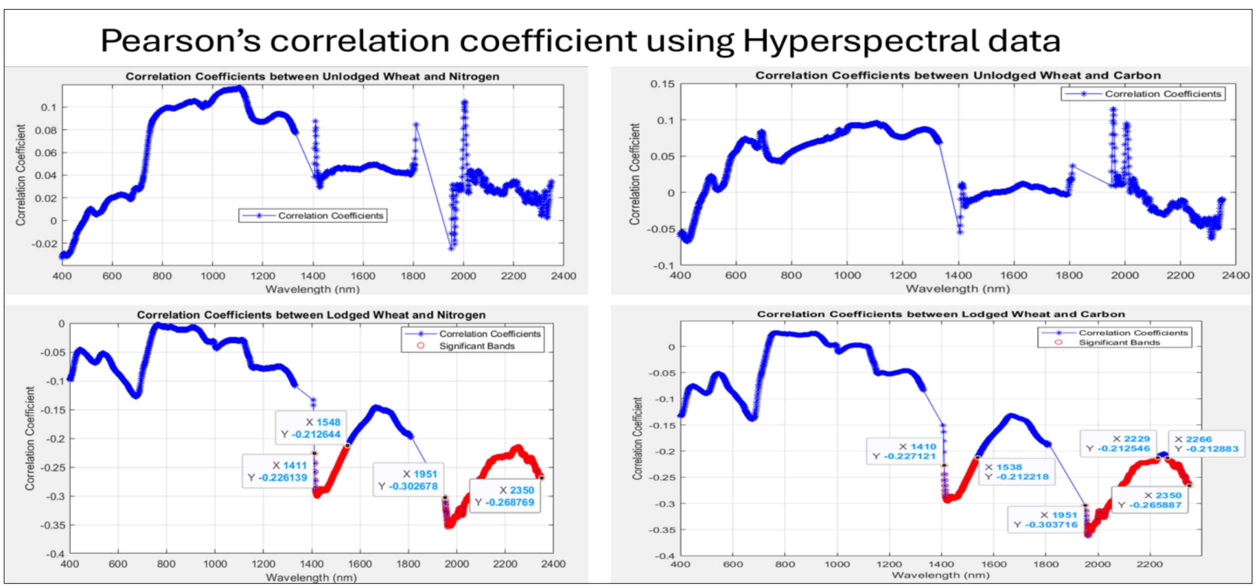


Figure 10: Correlation coefficient of healthy wheat nitrogen (top left) content and carbon (top right), lodged wheat nitrogen (bottom left) content and carbon (bottom right) content in field hyperspectral data, this Figure shows significance in some bands (in red colour) in the lodged wheat, that can be used in estimate nitrogen and carbon content. Healthy wheat shows correlation, but they are not significant.

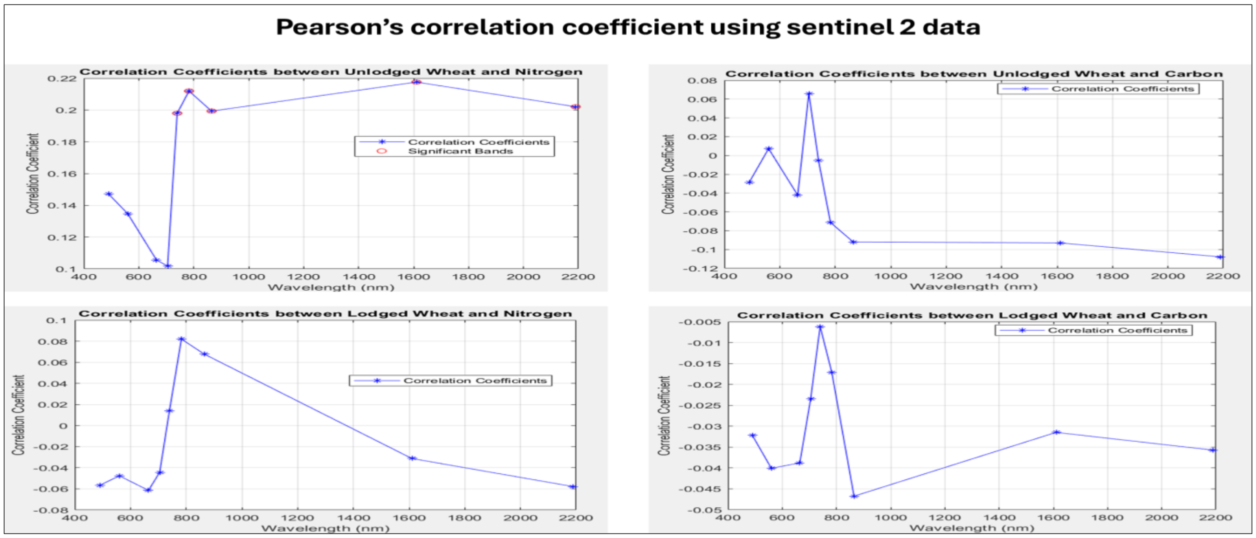


Figure 11: Correlation coefficient of healthy and lodged wheat nitrogen and carbon using Sentinel-2 data; healthy nitrogen content (top left) shows significant band in red colour, healthy carbon (Top right), lodged wheat nitrogen (bottom left) and carbon (bottom right) content do not have significant band that can be used for these variables' retrieval.

3.3 Absorption features and continuum removal

The absorption features and continuum removal (Figure 12) steps goes together, as the continuum removal was done to determine where the absorption features are located and then band depth analysis was done to determine the deference between the lodged and healthy wheat spectra. After the result of the Pearsons correlation showing that healthy wheat does not show any significant band, it was decided that the whole spectral will be used in this step to really consider which location have high band depth difference. Considering the two tables, three of the absorption peaks at field level is close to the located peaks at the plot level.

Table 7: (a-left) absorption peak band location using field hyperspectral data and (b-right) absorption peak band location using Sentinel-2 data.

Bands with highest absorption peak (nm)	Band depth	Band depth
	Lodged	Healthy
672	0.43	0.37
714	0.78	0.74
976 & 977	0.82	0.80
1164 & 1165	0.76	0.73
1463	0.55	0.52

Bands with highest absorption peak (nm)	Band depth	Band depth
	Lodged	Healthy
665	0.35	0.38
705	0.39	0.41
1610	0.90	0.97

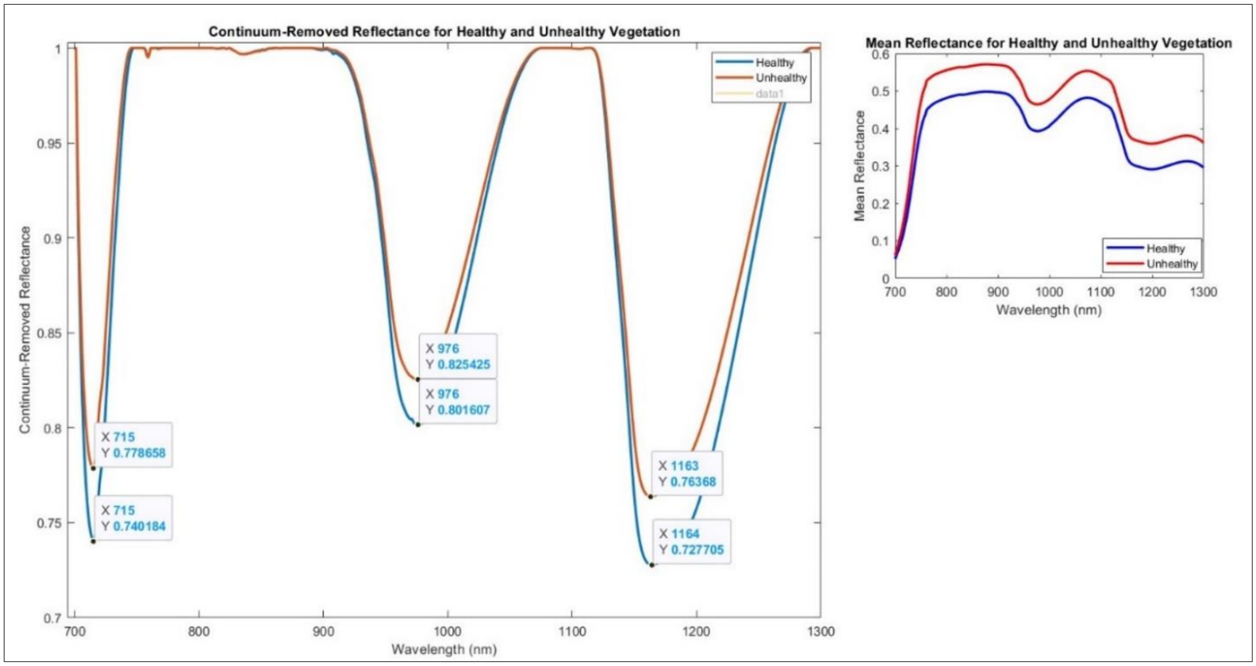


Figure 12: Continuum removal reflectance of hyperspectral data.

3.5 Vegetation indices

The coefficient of determination (R^2) for various vegetation indices used to analyze the nitrogen content in wheat are used to know the accuracy and effectiveness of each index in predicting or correlating with the foliar nitrogen and carbon content in healthy and lodged wheat. Higher R^2 values indicate stronger relationships between the vegetation indices and studied parameters, suggesting that the index is a better predictor. Normalized Difference Vegetation Index (NDVI) achieved the highest R^2 (0.18), followed by Cellulose Absorption Index (CAI) and Enhanced Vegetation Index (EVI) both having same R^2 (0.03), specifically for estimating healthy nitrogen content using Sentinel-2 data but its performance in lodged nitrogen estimation was too low in the indices used for this study. Additionally, hyperspectral data proved more effective than Sentinel-2 in estimating nitrogen content in lodged wheat than in healthy wheat. Among the indices, the Lignin Cellulose Absorption Index (LCAI) and the Plant Senescence Reflectance Index (PSRI) showed better performance (0.10) when compared with other indices used in estimating nitrogen content in lodged wheat. However, the results also indicate that hyperspectral data generally underperforms in estimating healthy nitrogen content but performed well in lodged nitrogen in most of the indices used, see Appendix (B5 and C5).

Estimation of carbon content using hyperspectral data shows that the Lignin Cellulose Absorption Index (LCAI) and the Cellulose Absorption Index (CAI) are the most effective indices for estimating carbon content 0.13 and 0.04 respectively in lodged vegetation. Conversely, the Normalized Difference Lignin Index (NDLI) and CAI provide more accurate estimates for carbon content in healthy wheat 0.06 and 0.03 respectively compared to other indices. When using Sentinel-2 data, the retrieval rates for both healthy and lodged carbon content were generally low. However, the Plant Senescence Reflectance Index (PSRI) proved to be the most effective, achieving an R^2 value barely above 0.1 which is recorded in healthy carbon.

This low R^2 values recorded in our study indicate that the vegetation indices used in this study are not well-suited for estimating carbon and nitrogen content in both healthy and lodged vegetation using the available data. The limitations of these indices and the data highlight the need for more robust, multi-faceted approaches that consider the complexity of wheat biochemical parameters estimation. So, we took a step further in by using narrowband NDVI in prediction of carbon and nitrogen content in wheat.

3.6 Best narrowband

The study used best narrowband NDVI (Normalized Difference Vegetation Index) to identify two spectral bands for predicting carbon and nitrogen content in both lodged and healthy conditions. As described in the methodology, the selected bands were occasionally too close in spectral range, potentially reducing the reliability of predictions. To address this, a sorting process was implemented to identify alternative bands that, while possibly yielding slightly lower R^2 values, could provide more stable estimations of the variables. Table 8 below lists the bands initially chosen by the model along with their corresponding R^2 values. It also includes the sorted bands that have been deemed more reliable for accurately estimating the variables.

Table 8: Best narrowband NDVI analysis in lodged and healthy wheat using field hyperspectral. The sorted bands are selected band that are more likely to give a more reliable result.

Variables	Model Band (nm)	R^2	Sorted Band (nm)	R^2
Lodged wheat nitrogen	1463 & 1469	0.29	1460 & 1470	0.27
Healthy wheat nitrogen	2330 & 2332	0.15	2003 & 2337	0.14
Lodged wheat carbon	1462 & 1470	0.32	1457 & 1471	0.31
Healthy wheat carbon	401 & 403	0.25	2085 & 2095	0.21

3.7 Carbon and nitrogen estimation

Healthy wheat nitrogen content shows a median value is 5.4 g/m², with a tight interquartile range (IQR) suggesting less variability in nitrogen content levels. The data points are quite concentrated. Lodged wheat reveals a higher median nitrogen content is 5.9 g/m² and has a wider IQR, indicating more variability, suggesting that lodged wheat nitrogen content is more varied when compared with healthy nitrogen content. In healthy wheat carbon content, the median carbon content is 127 g/m² with a narrow IQR, implying consistent carbon levels among healthy carbon, while lodged carbon median is 136 g/m², but with a slightly wider IQR, suggesting slightly more variability in carbon content in lodged wheat compared to healthy wheat (Figure 13). Lodged wheat exhibits higher nitrogen content and more variability in nitrogen and carbon content compared to healthy wheat.

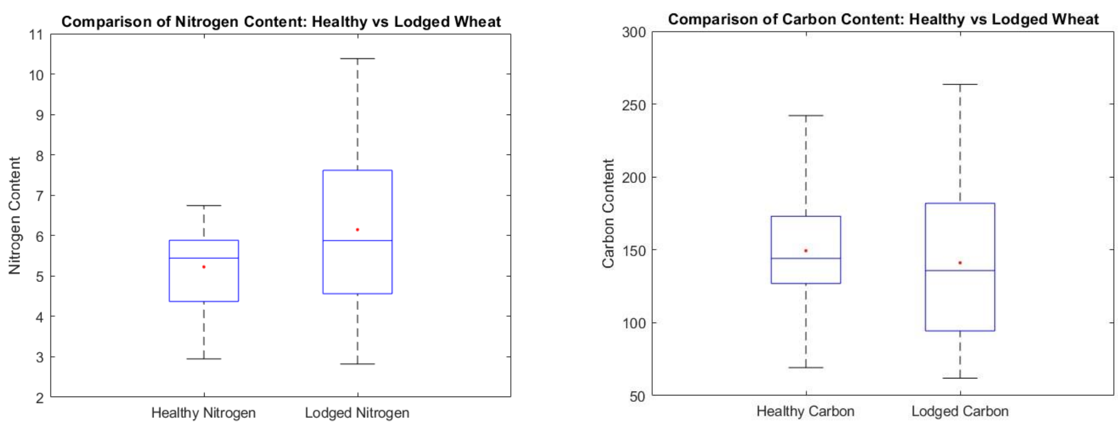
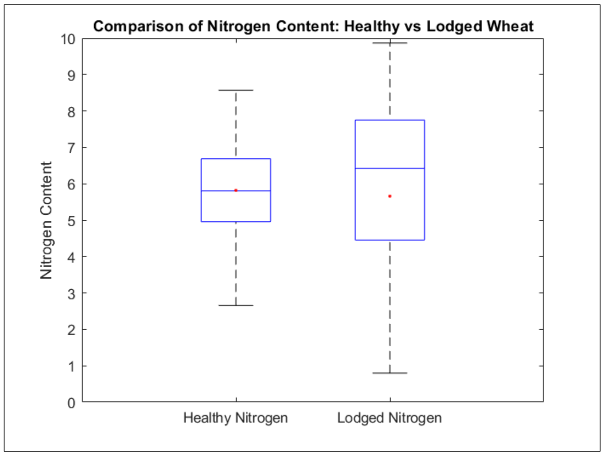


Figure 13:Boxplot comparison of lodged and healthy wheat nitrogen (left) and carbon (right) content using hyperspectral data; note that nitrogen and carbon content unit are in g/m².

In the boxplot in Figure 14, healthy nitrogen using Sentinel-2 data shows a median nitrogen content of 5.8 g/m² with a narrow interquartile range (IQR), indicating low variability and a consistent nitrogen content level in healthy wheat. Lodged nitrogen has a slightly higher median nitrogen content of 6.4 g/m², but the range is more compressed compared to the healthy wheat. suggests that lodging may affect nitrogen retrieval or estimation. Healthy wheat shows a median carbon content of 1.8 g/m², with a compact IQR. This reflects uniformity in carbon content among the healthy samples, lodged wheat shows a lower median carbon content, around 1.7 g/m², with a broader IQR than the healthy samples. This indicates a greater variability and lower carbon levels in lodged wheat, which could reflect stress or damage. Lodged wheat exhibits higher nitrogen content but lower carbon content, variations between healthy and lodged wheat in terms of carbon content is notable and it is suggesting significant impacts of lodging on carbon content.



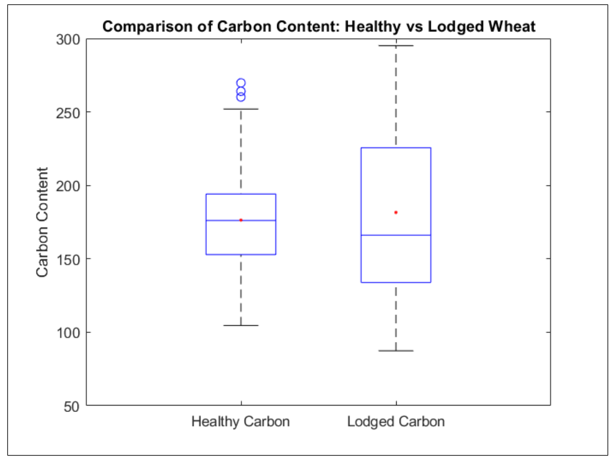


Figure 14: Boxplot comparison of lodged and healthy wheat nitrogen (left) and carbon (right) content using Sentinel-2 data; note that carbon and nitrogen content unit are in g/m^2.

3.8 Partial least square regression analysis

3.8.1 Estimation of carbon and nitrogen content using field hyperspectral data

The Partial Least Squares Regression (PLSR) models were evaluated using both venetian blinds cross-validation and randomized validation methods to assess their predictive performance on hyperspectral data for lodged and healthy carbon and nitrogen content estimation. The significant bands were determined during band depth analysis and all the identified bands (17 bands) were used in healthy wheat because the Pearson's correlation did not show significant correlation throughout the spectrum, while in the lodged wheat category, the bands identified after band depth analysis, that is in the region that was significant in Pearson's correlation (6 bands) were used. The PLSR prediction of lodged carbon and nitrogen where higher than that of healthy carbon and nitrogen when venetian blinds validation is used while in randomized validation, healthy nitrogen, and carbon prediction are higher than lodged nitrogen and carbon estimate. The RMSE-CV of lodged carbon has the highest of 1.007 while healthy nitrogen has the lowest with the value of 1.004 using venetian blinds validation but in the randomized validation, healthy nitrogen has the highest RMSE-CV of 1.012 followed by healthy carbon with the value of 1.010, while lodged carbon has the lowest value of 1.005. Detailed information is in Table 9, comparing R^2 CV of both validation methods, the randomization method consistently has a higher result in all biochemical components, indicating that the randomized validation method suggesting better predictive performance and model robustness compared to the venetian blinds' validation method.

Table 9: Performance of PLSR prediction at field hyperspectral level, using venetian and randomized validation method and their results. Preprocessing method used in the PLSR for the reflectance data is 1st Derivative (order: 2, window: 9 pt)

Parameters	Venetian blinds validation			Randomized validation		
	R^2 CV	RMSE-CV	CV-Bias	R^2 CV	RMSE-CV	CV-Bias
Lodged wheat carbon	0.41	1.007	-0.035	0.71	1.005	-0.034
Healthy wheat carbon	0.31	1.005	0.021	0.91	1.010	0.021
Lodged wheat nitrogen	0.39	1.005	0.031	0.73	1.006	-0.031
Healthy wheat nitrogen	0.32	1.004	0.020	0.94	1.012	0.020

3.8.2 Retrieval of carbon and nitrogen content using field Sentinel-2 data

The significant bands that have the highest absorption peaks (665nm, 705nm and 1610nm) in the band depth analysis were used to estimate nitrogen and carbon retrieval in both lodged and healthy wheat using Sentinel-2 data. To assess the predictive performance of Sentinel-2 data (Table 10), PLSR models were evaluated using both venetian blinds cross-validation and randomized validation methods. The retrieval prediction of lodged nitrogen was the higher with R^2 CV of 0.60 and lodged carbon was the lowest with 0.37 R^2 CV, healthy carbon and nitrogen were 0.44 and 0.42 respectively when venetian blinds validation is used. While randomized validation for healthy nitrogen, and carbon R^2 CV have the same value of 0.72, which is lower than lodged nitrogen and carbon estimate of 0.80 and 0.78 respectively. The RMSE-CV of lodged nitrogen has the highest of 1.007 while lodged carbon was the lowest and healthy carbon and nitrogen shared the same value of 1.004 using venetian blinds validation but in the randomized validation, lodged carbon and nitrogen sheared same and highest RMSE-CV of 1.006, while healthy nitrogen has the lowest value of 1.004. Detailed information is in Table 10, comparing R^2 CV of both validation method, it is seen that randomized consistently have a higher result in all biochemical components, just as shown in the predictive model for hyperspectral data above.

At Sentinel-2 level, it was only healthy nitrogen that have bands that were significant in retrieval after running a Pearson's correlation (740nm, 783nm, 865nm, 1610nm, 2190nm), others like heathy carbon, lodged nitrogen and carbon were correlated but the correlation is not significant. So, the bands with significant correlation in healthy nitrogen were also used for a PLSR analysis to compare it performance, with the bands from band depth analysis, detailed result is in Table 11.

Table 10: Performance of PLSR prediction at Sentinel-2 level, using venetian and randomized validation method and their results.

Parameters	Venetian blinds validation			Randomized validation		
	R^2 CV	RMSE-CV	CV-Bias	R^2 CV	RMSE-CV	CV-Bias
Lodged wheat carbon	0.37	1.003	-0.004	0.78	1.006	-0.005
Healthy wheat carbon	0.44	1.004	-0.014	0.72	1.005	-0.014
Lodged wheat nitrogen	0.60	1.007	-0.006	0.80	1.006	-0.006
Healthy wheat nitrogen	0.42	1.004	0.019	0.72	1.004	0.019

Table 11: Comparing healthy nitrogen retrieval using Pearson's correlation significant bands (PCSB) and recommended band from the band depth analysis (BDA) t Sentinel-2 level used in PLSR prediction, using venetian and randomized validation method.

Variables_S2 data	Venetian blinds validation			Randomized Validation		
	R^2 CV	RMSE-CV	CV-Bias	R^2 CV	RMSE-CV	CV-Bias
Healthy nitrogen (PCSB)	0.40	1.004	0.025	0.71	1.004	0.025
Healthy nitrogen (BDA)	0.42	1.004	0.019	0.72	1.004	0.019

CHAPTER 4: Discussion

4. Discussion

The relationship between carbon and nitrogen in healthy and lodged wheat reveals critical insights into their biochemical interactions under varying physiological conditions (Zhang et al., 2017a & 2017b). Understanding the relationship is vital for optimizing crop health, enhancing yield, enabling precision agriculture, reducing economic losses, and fostering agricultural research and development. The primary objective of this study was to evaluate the effectiveness of hyperspectral data for estimating nitrogen and carbon content in both healthy and lodged wheat and if Sentinel-2 data is effective in retrieval of nitrogen and carbon content in both healthy and lodged wheat. To discriminate between healthy and lodged wheat, spectral signature was used and the relationship between heathy and lodged wheat and carbon and nitrogen were determined by considering the significant bands for carbon and nitrogen using Pearson's correlation coefficient test. This study found that lodging significantly alters the spectral properties of wheat and impacts the estimation of these nutrients, consequently, the models developed were able to predict nitrogen and carbon content more accurately in lodged wheat than in healthy wheat, particularly because the reflectance changes made lodged wheat more distinguishable in hyperspectral analysis. In healthy wheat, the uniformity and stability in the plant structure result in consistent reflectance patterns, with specific spectral bands around 2003nm and 2337nm being crucial for nitrogen prediction and 2085nm & 2095nm to be crucial for carbon prediction, these regions had been identified to be important as nitrogen and carbon absorption bands (Curran, 1989; Mutaga et al., 2004). On the contrary, lodged wheat exhibits increased reflectance across various wavelengths due to structural changes and stress responses, making bands around 1460nm and 1470 nm more relevant for nitrogen estimation and 1457nm and 1471nm to be more relevant for carbon estimation, the band region noted for nitrogen in lodged wheat in this study shows that this region noted for carbon content (Curran, 1989; Curran et al., 2001). This difference showed in the important bands in both lodge and healthy wheat highlights how lodging alters the spectral properties of wheat, making certain wavelengths informative depending on the plant's condition.

4.1 Spectral reflectance of lodged and healthy wheat

The comparative analysis of reflectance properties between healthy and lodged wheat reveals significant spectral differences, with lodged wheat consistently showing higher reflectance (Chauhan et al., 2019) across visible (400-700 nm), near-infrared (700-1400 nm), and shortwave infrared (1400-2400 nm) regions. This increased reflectance in lodged wheat is likely due to structural changes (altered leaf structure and increasing leaf area exposed to the sensor) and orientation of the leaves and stems after lodging, loss of chlorophyll or changes in leaf angle (expose more leaf area to incident light), and reduced water content (Sun et al., 2021). The first derivative of the reflectance spectra further differentiates healthy from lodged wheat by highlighting specific wavelengths associated with chlorophyll absorption (around 680 nm and 730 nm) and water absorption (around 1450 nm and 1950 nm). These spectral characteristics were used to effectively distinguish between healthy and lodged wheat using both hyperspectral data and Sentinel-2 data, differentiation of the lodged and healthy wheat will support better crop management and yield optimization by allowing early detection of lodging and associated stress.

4.2 Spectral region analysis

This research revealed that lodging significantly impacts the spectral properties of wheat, which is more noted when hyperspectral data was used. Using Pearson's correlation coefficient, we identified specific wavelengths closely linked to carbon and nitrogen content in lodged wheat, while healthy wheat showed no significant correlations, indicating more uniform reflectance patterns. While using Sentinel-2 data, healthy nitrogen showed significant, but others have correlation that is not significant. The identification of specific absorption features through continuum removal and band depth analysis confirmed their utility in highlighting significant spectral regions. Continuum removal helps in isolating the absorption features, while band depth analysis quantifies the extent of these features, providing a clearer understanding of the spectral properties (Mutanga et al., 2004; Zhao & Lifu, 2013). Continuum removal and band depth analysis is not common in Sentinel-2 studies, but it was done so we can compare the results output of these bands after building our model.

4.3 Vegetation indices and best narrowband

The result of our analysis showed that using Lignin Cellulose Absorption Index (LCAI) and Plant Senescence Reflectance Index (PSRI) vegetation index predicts lodged nitrogen content better than healthy nitrogen, although the prediction of both indices is low with R^2 value of about 0.10 and 0.03 respectively, while in healthy nitrogen the best prediction is by using NDLI with just 0.02 predictability, using hyperspectral data. For carbon LCAI predicts lodged carbon content better and NDLI predict healthy carbon the best, with the value of 0.13 and 0.06 respectively. At Sentinel-2 level, the best predictive vegetation index for healthy nitrogen was NDVI with predicts value of 0.18, for lodged nitrogen content, the value was less than 0.02. Predicting healthy carbon content was difficult when compared with healthy nitrogen, the best is PSRI with the prediction value barely over 0.01. For lodged carbon and nitrogen prediction using Sentinel-2 data, the prediction was difficult, and the best prediction was for lodged nitrogen with 0.01 using NDVI and RVI. These findings suggest that different vegetation indices are better suited for specific conditions and highlight the importance of selecting appropriate indices based on the health status of the crop, as different indices offer varying levels of predictability for nitrogen and carbon content with in healthy and lodged wheat. Sentinel-2 data showed lower predictive accuracy for both healthy and lodged wheat compared to hyperspectral data, as indicated by lower R^2 values. For instance, the best predictive value for healthy nitrogen using Sentinel-2 data was achieved with the NDVI index at 0.18, while for lodged nitrogen content, the value was less than 0.02. This contrast suggests that S2's broader spectral bands are less sensitive to the subtle spectral differences needed for accurate biochemical parameter estimation compared to the narrower bands used in hyperspectral imaging.

The best narrow band result showed a higher predictive model for lodged and healthy wheat carbon content when compared with the normal vegetation indices. Some of the bands selected by the model were too close together and using closely spaced spectral bands for predictive model can lead to redundancy, over-fitting, noise sensitivity, and computational inefficiency, leading to unreliable result, selecting a broader range of bands can improve model performance and generalization. The R^2 value for best narrowband were showed in Table 8 and the prediction of lodged wheat for carbon and nitrogen performed better than that of healthy wheat.

4.4 Nitrogen and carbon content estimation

The box-plot analysis revealed distinct differences in nitrogen and carbon content between healthy and lodged wheat. Healthy wheat exhibited a narrower inter-quartile range (IQR) for nitrogen content, suggesting lower variability and more consistent nitrogen levels, but lodged wheat showed higher median nitrogen content and greater variability, indicating that nitrogen significantly affects wheat lodging. The same trend is seen in carbon content, with healthy wheat having a narrow IQR and lodged wheat exhibiting higher median values and greater variability. Higher nitrogen and carbon content can contribute to lodging by promoting rapid and heavy growth without necessarily strengthening the plant's structure adequately. Therefore, it's crucial to manage nutrient levels carefully to ensure optimal plant health and minimize the risk of lodging.

4.5 PLSR model

Estimating carbon and nitrogen content using hyperspectral data was achieved by using bands identify from band depth analysis and the highest accuracy recorded using R^2 CV was found in lodged carbon followed closely by lodged nitrogen, with the value of 0.41 and 0.39 respectively, using venetian blind validation. Healthy nitrogen and carbon with the value of 0.94 and 0.91 respectively was the highest prediction when randomized validation was used. At Sentinel-2 level, the retrieval of carbon and nitrogen in lodged and healthy wheat was predicted better in lodged wheat carbon in both model validation and nitrogen and carbon have more accurate results in lodged wheat when randomized validation was used. Both models; using hyperspectral and Sentinel-2 data reveals that model accuracy for predicting nitrogen and carbon content in wheat varies significantly, and this depends on the validation method used. The inconsistency in model results noted in randomized validation at each run of the model suggests the need for a more robust evaluation approach. This inconsistency may highlight the importance of carefully selecting validation techniques to ensure a comprehensive and accurate assessment of our model's performance. Venetian blinds validation systematically covers the entire dataset, making it effective in capturing patterns and variances associated with lodged wheat, while randomized validation ensures a more generalized model. Considering the information on RMSEcv on Table 9, it should be noted that randomized validation has higher error values when compared to venetian blinds in healthy carbon and nitrogen, and in lodged nitrogen, but the bias in both validation is the same. At Sentinel-2 level, RMSEcv on Table 10, showed that randomized validation has higher error values when compared to venetian blinds in healthy carbon and nitrogen, and in lodged carbon, and bias is also a bit higher in randomized validation in both healthy and lodged carbon content while they have same RMSEcv value in healthy nitrogen. Comparing healthy nitrogen retrieval using Pearson's correlation significant bands (PCSB) and important bands from the band depth analysis (BDA) at Sentinel-2 level, results of both validation method showed a better retrieval prediction when band depth analysis bands are used for the prediction.

This study shows an overall result that hyperspectral remote sensing can estimate carbon and nitrogen content in lodged wheat using our models, provided the bands identified to have high absorption peak during the process of band analysis is used. This is evident from the higher variability in nitrogen and carbon content in lodged wheat in our boxplot analysis and significant difference shown by the Pearson's correlation coefficient affirm that remote sensing techniques can be effectively used in carbon and nitrogen estimation and management even in lodged areas within a field. Although there was no significant band identified in lodged wheat using Sentinel-2 data, our model was still able to retrieve carbon and nitrogen in lodged wheat.

The use of two different models in this study also reveal the importance of comparing different validation for accurately predicting the more stable conditions of wheat biochemical content. Using both methods allows for a robust evaluation, ensuring that the model is reliable across different conditions and providing valuable understandings for wheat health monitoring and nutrient management. The variability showed in randomized validation can be a disadvantage when exact reproducibility of results is required, as observed in our repeated analyses where the randomized validation produced slightly different results each time. On the other hand, the venetian blinds method yielded consistent results across multiple runs, making it a more reliable choice for studies where reproducibility and consistency are paramount. Therefore, in the context of our research, where the goal is to achieve reproducible and comparable results with the same dataset and methodology, the Venetian blinds validation method is more suitable. This ensures that any observed differences in model performance are due to actual changes in the data or methodology, rather than variability introduced by the validation process itself.

CHAPTER 5: CONCLUSIONS

This research confirmed the effectiveness of hyperspectral in estimating nitrogen and carbon content in wheat, highlighting the spectral differences between healthy and lodged wheat plays a key role in retrieval and estimation of carbon and nitrogen content. Lodged wheat consistently exhibited higher reflectance and greater variability in nitrogen and carbon content compared to healthy wheat. The analysis result indicated that the spectral properties of lodged wheat significantly differed from those of healthy wheat, making it more responsive to hyperspectral analysis. The structural changes and stress responses associated with lodging increased reflectance across various wavelengths, thereby enhancing the model's ability to predict nitrogen and carbon content in lodged wheat.

The key finding in this research is that lodging affected the estimation of biochemical parameters by making certain spectral features more prominent, which in turn improved the accuracy of predictions for lodged wheat. The structural and physiological changes due to lodging made the nitrogen and carbon content more detectable using hyperspectral imaging. Specific spectral bands were identified as crucial for nitrogen and carbon prediction in both healthy and lodged wheat. For healthy wheat, bands around 2003 nm and 2337 nm were important for nitrogen, and bands around 2085 nm and 2095 nm for carbon. For lodged wheat, bands around 1460 nm and 1470 nm were relevant for nitrogen, and bands around 1457 nm and 1471 nm for carbon.

In the Sentinel-2 nitrogen and carbon content retrieval from wheat, it was confirmed that Sentinel-2 is effective in retrieving nitrogen and carbon content in both lodged and healthy wheat. This study identified specific wavelengths significantly associated with nutrient content estimation (1463nm, 2001nm, 2270nm, 2276nm, 2309nm, and 2312nm) in lodged wheat, and the whole 17 bands derived from the continuum removal (check appendix B10) was used in healthy wheat; using hyperspectral data and for nutrient content retrieval (665nm, 705nm, and 1610nm) in both lodged and healthy wheat using Sentinel-2 data, enhancing the precision of remote sensing applications in agricultural monitoring.

Vegetation indices, particularly the Lignin Cellulose Absorption Index (LCAI) and Plant Senescence Reflectance Index (PSRI), were better predictors for lodged nitrogen content, while the Normalized Difference Lignin Index (NDLI) was more suitable for healthy carbon content using hyperspectral data. The use of narrowband NDVI provided higher predictive accuracy for both conditions, with lodged carbon having the highest R^2 followed by lodged nitrogen. Sentinel-2 data showed lower predictive accuracy compared to hyperspectral data due to its broader spectral bands, which are less sensitive to subtle spectral differences.

PLSR models further validated the findings, showing that model accuracy varies significantly with the validation method used. venetian blinds validation proved more effective for lodged wheat, capturing the complexities associated with lodging. On the contrary, randomized validation provided a more generalized model, better suited for healthy wheat conditions.

The implications for precision agriculture in our study area are that by selecting appropriate spectral bands and validation methods, it is possible to enhance the accuracy of nitrogen and carbon content estimation and retrieval predictions, having a comprehensive assessment of model performance. Venetian blinds validation provided more consistent and reliable results compared to randomized validation, which exhibited higher variability. The variability in nitrogen and carbon content between healthy and lodged wheat, as revealed by box-plot analysis, emphasizes the need for careful nutrient management to minimize lodging risks. Higher median nitrogen and carbon content in lodged wheat suggest that excessive nutrient levels can promote rapid growth, potentially leading to lodging.

The implication of this study is that the ability to accurately estimate nitrogen and carbon content in wheat through hyperspectral remote sensing supports more precise crop monitoring and management. This can lead to optimized nutrient application, reducing waste and improving yield. Early detection of lodging and associated nutrient stress can facilitate timely interventions, optimize fertilization practices, and improve yield and crop quality. Using of multiple validation methods ensures robust model performance, providing reliable tools for agricultural researchers and practitioners to monitor and manage crop health under varying conditions.

This study emphasizes hyperspectral remote sensing is a powerful tool for estimating biochemical parameters in wheat, particularly in distinguishing and managing the effects of lodging and that it is importance of integrating advanced remote sensing techniques with robust validation methods to achieve reliable and actionable insights for agricultural productivity and sustainability, supports optimal nutrient management, improves crop health monitoring, enhances yield predictions, and facilitates precision agriculture practices. Additionally, it contributes to environmental sustainability and economic benefits for farmers. Retrieval of these nutrients can be achieved more accurately, enabling farmers to adopt more efficient and sustainable farming practices.

LIST OF REFERENCES

Bicharanloo, B., Cavagnaro, T. R., Keitel, C., & Dijkstra, F. A. (2021). Nitrogen Fertilisation Increases Specific Root Respiration in Ectomycorrhizal but Not in Arbuscular Mycorrhizal Plants: A Meta-Analysis. In *Frontiers in Plant Science* (Vol. 12). Frontiers Media S.A. https://doi.org/10.3389/fpls.2021.711720

Chanchí Golondrino, G. E., Ospina Alarcón, M. A., & Saba, M. (2023). Vegetation Identification in Hyperspectral Images Using Distance/Correlation Metrics. *Atmosphere*, *14*(7). https://doi.org/10.3390/atmos14071148

Chauhan, S., Darvishzadeh, R., Boschetti, M., & Nelson, A. (2020a). Discriminant analysis for lodging severity classification in wheat using RADARSAT-2 and Sentinel-1 data. *ISPRS Journal of Photogrammetry and Remote Sensing*, *164*, 138–151. https://doi.org/10.1016/j.isprsjprs.2020.04.012

Chauhan, S., Darvishzadeh, R., Boschetti, M., & Nelson, A. (2020b). Estimation of crop angle of inclination for lodged wheat using multi-sensor SAR data. *Remote Sensing of Environment*, *236*. https://doi.org/10.1016/j.rse.2019.111488

Chauhan, S., Darvishzadeh, R., Boschetti, M., & Nelson, A. (2020c). Estimation of crop angle of inclination for lodged wheat using multi-sensor SAR data. *Remote Sensing of Environment*, *236*. https://doi.org/10.1016/j.rse.2019.111488

Chauhan, S., Darvishzadeh, R., Boschetti, M., & Nelson, A. (2020d). Estimation of crop angle of inclination for lodged wheat using multi-sensor SAR data. *Remote Sensing of Environment*, *236*. https://doi.org/10.1016/j.rse.2019.111488

Chauhan, S., Darvishzadeh, R., Boschetti, M., Pepe, M., & Nelson, A. (2019). Remote sensing-based crop lodging assessment: Current status and perspectives. In *ISPRS Journal of Photogrammetry and Remote Sensing* (Vol. 151, pp. 124–140). Elsevier B.V. https://doi.org/10.1016/j.isprsjprs.2019.03.005

Chauhan, S., Darvishzadeh, R., van Delden, S. H., Boschetti, M., & Nelson, A. (2021). Mapping of wheat lodging susceptibility with synthetic aperture radar data. *Remote Sensing of Environment*, *259*. https://doi.org/10.1016/j.rse.2021.112427

Cheng, JH., Sun, DW (2017). Partial Least Squares Regression (PLSR) Applied to NIR and HSI Spectral Data Modeling to Predict Chemical Properties of Fish Muscle. Food Eng Rev 9, 36–49. https://doi.org/10.1007/s12393-016-9147-1

Cotrozzi, L., Townsend, P. A., Pellegrini, E., Nali, C., & Couture, J. J. (2018). Reflectance spectroscopy: a novel approach to better understand and monitor the impact of air pollution on Mediterranean plants. *Environmental Science and Pollution Research*, *25*(9), 8249–8267. https://doi.org/10.1007/s11356-017-9568-2

Curran, P. J. (1989). Remote sensing of foliar chemistry. Remote Sensing of Environment, 30(3), 271–278. https://doi.org/10.1016/0034-4257(89)90069-2

Curran, P. J., Dungan, J. L., & Peterson, D. L. (2001). Estimating the foliar biochemical concentration of leaves with reflectance spectrometry: Testing the Kokaly and Clark methodologies. Remote Sensing of Environment, 76(3), 349–359. https://doi.org/10.1016/S0034-4257(01)00182-1

Darvishzadeh, R. (n.d.). *Hyperspectral remote sensing of vegetation parameters using statistical and physical models*.

Darvishzadeh, R., Atzberger, C., & Skidmore, A. K. (n.d.). *HYPERSPECTRAL VEGETATION INDICES FOR ESTIMATION OF LEAF AREA INDEX*.

Darvishzadeh, R., Atzberger, C., Skidmore, A., & Schlerf, M. (n.d.). *RETRIEVAL OF VEGETATION BIOCHEMICALS USING A RADIATIVE TRANSFER MODEL AND HYPERSPECTRAL DATA*.

Daughtry, C. (2001). "Discriminating Crop Residues from Soil by Short-Wave Infrared Reflectance." *Agronomy Journal* 93: 125-131.

Daughtry, C., E. Hunt, Jr., P. Doraiswamy, and J. McMurtrey III (2005). "Remote Sensing the Spatial Distribution of Crop Residues." Agronomy Journal 97: 864-871.

Falcioni, R., Antunes, W. C., Demattê, J. A. M., & Nanni, M. R. (2023). A Novel Method for Estimating Chlorophyll and Carotenoid Concentrations in Leaves: A Two Hyperspectral Sensor Approach. *Sensors*, *23*(8). https://doi.org/10.3390/s23083843

Farbo, A., Meloni, R., Blandino, M., Sarvia, F., Reyneri, A., Borgogno-Mondino, E. (2022). Spectral Measures from Sentinel-2 Imagery vs Ground-Based Data from Rapidscan© Sensor: Performances on Winter Wheat. In: Borgogno-Mondino, E., Zamperlin, P. (eds) Geomatics for Green and Digital Transition. ASITA 2022. Communications in Computer and Information Science, vol 1651. Springer, Cham. https://doi.org/10.1007/978-3-031-17439-1_15

Feng, X., Yu, C., Liu, X., Chen, Y., Zhen, H., Sheng, K., & He, Y. (2018). Nondestructive and rapid determination of lignocellulose components of biofuel pellet using online hyperspectral imaging system. *Biotechnology for Biofuels*, *11*(1). https://doi.org/10.1186/s13068-018-1090-3

Galieni, A., D'Ascenzo, N., Stagnari, F., Pagnani, G., Xie, Q., & Pisante, M. (2021). Past and Future of Plant Stress Detection: An Overview from Remote Sensing to Positron Emission Tomography. In *Frontiers in Plant Science* (Vol. 11). Frontiers Media S.A. https://doi.org/10.3389/fpls.2020.609155

Gold, K. M., Townsend, P. A., Chlus, A., Herrmann, I., Couture, J. J., Larson, E. R., & Gevens, A. J. (2020). Hyperspectral measurements enable pre-symptomatic detection and differentiation of contrasting physiological effects of late blight and early blight in potato. *Remote Sensing*, *12*(2). https://doi.org/10.3390/rs12020286

Huete, A., Didan,K., Miura, T., Rodriguez,E. P., Gao, X., Ferreira, L. G. (2002). Overview of the radiometric and biophysical performance of the MODIS vegetation indices. Remote Sensing of Environment 83, 195-213. doi:10.1016/S0034-4257(02)00096-2.

Inoue, Y., Guérif, M., Baret, F., Skidmore, A., Gitelson, A., Schlerf, M., Darvishzadeh, R., & Olioso, A. (2016). Simple and robust methods for remote sensing of canopy chlorophyll content: a comparative analysis of hyperspectral data for different types of vegetation. *Plant Cell and Environment*, *39*(12), 2609–2623. https://doi.org/10.1111/pce.12815

Jiang, X., Luo, S., Ye, Q., Li, X., & Jiao, W. (2022). Hyperspectral Estimates of Soil Moisture Content Incorporating Harmonic Indicators and Machine Learning. *Agriculture (Switzerland)*, *12*(8). https://doi.org/10.3390/agriculture12081188

Jin, J.; Wang, Q (2019). Evaluation of Informative Bands Used in Different PLS Regressions for Estimating Leaf Biochemical Contents from Hyperspectral Reflectance. Remote Sens. 11, 197. https://doi.org/10.3390/rs11020197

Kumar, P., Rani, M., Chandra, P., Haroon, P., Bhagwan, S., & Chaudhary, S. (n.d.). *Applications and Challenges of Geospatial Technology Potential and Future Trends*.

Li, C., Chang, Y., Luo, Y., Li, W., Jin, M., Wang, Y., Cui, H., Sun, S., Li, Y., & Wang, Z. (2023). Nitrogen regulates stem lodging resistance by breaking the balance of photosynthetic carbon allocation in wheat. *Field Crops Research*, *296*. https://doi.org/10.1016/j.fcr.2023.108908

Li, G., Han, W., Huang, S., Ma, W., Ma, Q., & Cui, X. (2021). Extraction of sunflower lodging information based on uav multi-spectral remote sensing and deep learning. *Remote Sensing*, *13*(14). https://doi.org/10.3390/rs13142721

Li, W. qian, HAN, M. ming, PANG, D. wei, CHEN, J., WANG, Y. yuan, DONG, H. he, CHANG, Y. lan, JIN, M., LUO, Y. li, LI, Y., & WANG, Z. lin. (2022). Characteristics of lodging resistance of high-yield winter wheat as affected by nitrogen rate and irrigation managements. *Journal of Integrative Agriculture*, *21*(5), 1290–1309. https://doi.org/10.1016/S2095-3119(20)63566-3

Li, Y. H., Tan, X., Zhang, W., Jiao, Q. Bin, Xu, Y. X., Li, H., Zou, Y. B., Yang, L., & Fang, Y. P. (2021). Research and Application of Several Key Techniques in Hyperspectral Image Preprocessing. *Frontiers in Plant Science*, *12*. https://doi.org/10.3389/fpls.2021.627865

Li, Y.; Guo, Y.; Gong, L.; Liu, C. (2023). Harvesting Route Detection and Crop Height Estimation Methods for Lodged Farmland Based on AdaBoost. Agriculture 13, 1700. https://doi.org/10.3390/agriculture13091700

Liu, L.; Ji, M.; Buchroithner, M (2017). Combining Partial Least Squares and the Gradient-Boosting Method for Soil Property Retrieval Using Visible Near-Infrared Shortwave Infrared Spectra. Remote Sens. 9, 1299. https://doi.org/10.3390/rs9121299

Liu, Q., Luo, L., & Zheng, L. (2018). Lignins: Biosynthesis and biological functions in plants. In *International Journal of Molecular Sciences* (Vol. 19, Issue 2). MDPI AG. https://doi.org/10.3390/ijms19020335

Liu X, Lyu Q, He S., Yi S., Hu D., Wang Z., Xie R., Zheng Y., Deng L., (2016). Estimation of carbon and nitrogen contents in citrus canopy by low-altitude remote sensing. International Journal of Agricultural and Biological Engineering, 149-157, 9(5)

Liu, Q., Yin, C., Li, X., He, C., Ding, Z., & Du, X. (2022). Lodging resistance of rice plants studied from the perspective of culm mechanical properties, carbon framework, free volume, and chemical composition. *Scientific Reports*, *12*(1). https://doi.org/10.1038/s41598-022-24714-4

Lowe, A., Harrison, N. & French, A.P (2017). Hyperspectral image analysis techniques for the detection and classification of the early onset of plant disease and stress. Plant Methods 13, 80. https://doi.org/10.1186/s13007-017-0233-z

Lu, B., Dao, P. D., Liu, J., He, Y., & Shang, J. (2020). Recent advances of hyperspectral imaging technology and applications in agriculture. In *Remote Sensing* (Vol. 12, Issue 16). MDPI AG. https://doi.org/10.3390/RS12162659

Ma, C., Zhai, L., Li, C., & Wang, Y. (2022). Hyperspectral Estimation of Nitrogen Content in Different Leaf Positions of Wheat Using Machine Learning Models. *Applied Sciences (Switzerland)*, *12*(15). https://doi.org/10.3390/app12157427

Machwitz, M., Pieruschka, R., Berger, K., Schlerf, M., Aasen, H., Fahrner, S., Jiménez-Berni, J., Baret, F., & Rascher, U. (2021). Bridging the Gap Between Remote Sensing and Plant Phenotyping—Challenges and Opportunities for the Next Generation of Sustainable Agriculture. *Frontiers in Plant Science*, *12*. https://doi.org/10.3389/fpls.2021.749374

Mielke, C., Bösche, N. K., Rogass, C., Segl, K., Gauert, C., & Kaufmann, H. (n.d.). *POTENTIAL APPLICATIONS OF THE SENTINEL-2 MULTISPECTRAL SENSOR AND THE ENMAP HYPERSPECTRAL SENSOR IN MINERAL EXPLORATION*. https://doi.org/10.12760/01-2014-2-07

Miraki, M., & Sohrabi, H. (2022). Using canopy height model derived from UAV imagery as an auxiliary for spectral data to estimate the canopy cover of mixed broadleaf forests. *Environmental Monitoring and Assessment*, *194*(1). https://doi.org/10.1007/s10661-021-09695-7

Mizuta, K., Araki, H., & Takahashi, T. (2023). Relationship between canopy coverage at the initiation of stem elongation and lodging in wheat. *European Journal of Agronomy*, *148*. https://doi.org/10.1016/j.eja.2023.126855

Mutanga, O., Skidmore, A. K., & Prins, H. H. T. (2004). Predicting in situ pasture quality in the Kruger National Park, South Africa, using continuum-removed absorption features. Remote Sensing of Environment, 89(3), 393–408. https://doi.org/10.1016/J.RSE.2003.11.001

Ozdogan, M., Yang, Y., Allez, G., & Cervantes, C. (2010). Remote sensing of irrigated agriculture: Opportunities and challenges. In *Remote Sensing* (Vol. 2, Issue 9, pp. 2274–2304). https://doi.org/10.3390/rs2092274

Pascucci, S., Pignatti, S., Casa, R., Darvishzadeh, R., & Huang, W. (2020). Special issue "hyperspectral remote sensing of agriculture and vegetation." In *Remote Sensing* (Vol. 12, Issue 21, pp. 1–7). MDPI AG. https://doi.org/10.3390/rs12213665

Pearson, R. L.; Miller, L. D. (1972). Remote mapping of standing crop biomass for estimation of the productivity of the short-grass Prairie, Pawnee National Grasslands, Colorado. Proceedings of the Eighth International Symposium on Remote Sensing of Environment. Willow Run Laboratories, Environmental Research Institute of Michigan. Pg1357-1381.

Pereira-Pires, J.E., Mora, A., Aubard, V., Silva, J.M.N., Fonseca, J.M. (2021). Assessment of Sentinel-2 Spectral Features to Estimate Forest Height with the New GEDI Data. In: Camarinha-Matos, L.M., Ferreira, P., Brito, G. (eds) Technological Innovation for Applied AI Systems. DoCEIS 2021. IFIP Advances in Information and Communication Technology, vol 626. Springer, Cham. https://doi.org/10.1007/978-3-030-78288-7_12

Rouse, J.W., Haas, R.H., Schell, J.A. and Deering, D.W. (1974) Monitoring Vegetation Systems in the Great Plains with ERTS. Third ERTS-1 Symposium NASA, NASA SP-351, Washington DC, 309-317.

Saiz-Fernández, I., De Diego, N., Brzobohatý, B., Muñoz-Rueda, A., & Lacuesta, M. (2017). The imbalance between C and N metabolism during high nitrate supply inhibits photosynthesis and overall growth in maize (Zea mays L.). *Plant Physiology and Biochemistry*, *120*, 213–222. https://doi.org/10.1016/j.plaphy.2017.10.006

Serrano, L., Peñuelas, J., & Ustin, S. L. (2002). Remote sensing of nitrogen and lignin in Mediterranean vegetation from AVIRIS data. Remote Sensing of Environment, 81(2-3), 355–364. doi:10.1016/s0034-4257(02)00011-1

Sharifi, A. (2020). Remotely sensed vegetation indices for crop nutrition mapping. *Journal of the Science of Food and Agriculture*, *100*(14), 5191–5196. https://doi.org/10.1002/jsfa.10568

Shen, L.; Gao, M.; Yan, J.; Li, Z.-L.; Leng, P.; Yang, Q.; Duan, S.-B (2020). Hyperspectral Estimation of Soil Organic Matter Content using Different Spectral Preprocessing Techniques and PLSR Method. Remote Sens. 12, 1206. https://doi.org/10.3390/rs12071206

Shrestha, S., Rebecca, M., Laza, C., Mendez, K. V, Bhosale, S., & Dingkuhn, M. (2019). *Title: 1 The Blaster: A Methodology to Induce Rice Lodging at Plot Scale to Study Lodging Resistance 2 3*.

Sinto, A., Sathee, L., Singh, D., Jha, S. K., Adavi, S. B., Kumar, R. R., Chinnusamy, V., & Singh, M. P. (2022). Elevated CO2 and Nitrogen dose affect grain ionome, grain morphology and associated gene expression in wheat (Triticum aestivum L.). *Indian Journal of Genetics and Plant Breeding*, *82*(2), 143–152. https://doi.org/10.31742/IJGPB.82.2.2

Steve Veysey. (2015). CHNS Instructions. *Iowa State University, Chemical Instrumentation Facility*, *Revised 05/29/2015 S.V.*, 1–16. https://www.cif.iastate.edu/sites/default/files/uploads/Other_Inst/CHN/CHNS%20Instructions.pdf

Svečnjak, Z., Jareš, D., Kovačević, M., Barić, K., Ivanušić, T., Maćešić, D., & Dujmović Purgar, D. (2020). Changes in seed yield and quality resulting from lodging in italian ryegrass crop. *Journal of Central European Agriculture*, *21*(1), 51–55. https://doi.org/10.5513/JCEA01/21.1.2748

Transon, J., d'Andrimont, R., Maugnard, A., & Defourny, P. (2018). Survey of hyperspectral Earth Observation applications from space in the Sentinel-2 context. In *Remote Sensing* (Vol. 10, Issue 2). MDPI AG. https://doi.org/10.3390/rs10020157

USDA, Foreign Agricultural Service, U.S. Department of Agriculture. Wheat 2023, World Production. Wheat Explorer. International Production Assessment Division. https://ipad.fas.usda.gov/cropexplorer/cropview/commodityView.aspx?cropid=0410000 Accessed Online: 24/02/2024.

Wang, C., Nie, S., Xi, X., Luo, S., & Sun, X. (2017). Estimating the biomass of maize with hyperspectral and LiDAR data. *Remote Sensing*, *9*(1). https://doi.org/10.3390/rs9010011

Wang, S., Zhou, K., Wang, J., & Zhao, J. (2022). Identifying and Mapping Alteration Minerals Using HySpex Airborne Hyperspectral Data and Random Forest Algorithm. *Frontiers in Earth Science*, *10*. https://doi.org/10.3389/feart.2022.871529

Wu, B., Zhang, M., Zeng, H., Tian, F., Potgieter, A. B., Qin, X., Yan, N., Chang, S., Zhao, Y., Dong, Q., Boken, V., Plotnikov, D., Guo, H., Wu, F., Zhao, H., Deronde, B., Tits, L., & Loupian, E. (2023). Challenges and opportunities in remote sensing-based crop monitoring: a review. In *National Science Review* (Vol. 10, Issue 4). Oxford University Press. https://doi.org/10.1093/nsr/nwac290

Wu, D. H., Chen, C. T., Yang, M. Der, Wu, Y. C., Lin, C. Y., Lai, M. H., & Yang, C. Y. (2022). Controlling the lodging risk of rice based on a plant height dynamic model. *Botanical Studies*, *63*(1). https://doi.org/10.1186/s40529-022-00356-7

Wu, M., Jiang, H., Wei, Z., Li, W., Gao, K., Wang, D., Wei, X., Tian, P., Cui, J., Di, Y., Wu, Z., & Yang, M. (2023). Influence of Nitrogen Application Rate on Stem Lodging Resistance Rice under Dry Cultivation. *Agronomy*, *13*(2). https://doi.org/10.3390/agronomy13020426

Wu, W., & Ma, B. L. (2022). Understanding the trade–off between lodging resistance and seed yield and developing some non–destructive methods for predicting crop lodging risk in canola production. *Field Crops Research*, *288*. https://doi.org/10.1016/j.fcr.2022.108691

Yu, J., Cheng, T., Cai, N., Lin, F., Zhou, X. G., Du, S., Zhang, D., Zhang, G., & Liang, D. (2022). Wheat lodging extraction using Improved_Unet network. *Frontiers in Plant Science*, *13*. https://doi.org/10.3389/fpls.2022.1009835

Zapata, F., Smith, E., Kreinovich, V., Phuong, N.H. (2023). Why Normalized Difference Vegetation Index (NDVI)?. In: Phuong, N.H., Kreinovich, V. (eds) Biomedical and Other Applications of Soft Computing. Studies in Computational Intelligence, vol 1045. Springer, Cham. https://doi.org/10.1007/978-3-031-08580-2_9

Zhang, F., & Zhou, G. (2019). Estimation of vegetation water content using hyperspectral vegetation indices: A comparison of crop water indicators in response to water stress treatments for summer maize. *BMC Ecology*, *19*(1). https://doi.org/10.1186/s12898-019-0233-0

Zhang, M., Wang, H., Yi, Y., Ding, J., Zhu, M., Li, C., Guo, W., Feng, C., & Zhu, X. (2017). Effect of nitrogen levels and nitrogen ratios on lodging resistance and yield potential of winter wheat (Triticum aestivum L.). *PLoS ONE*, *12*(11). https://doi.org/10.1371/journal.pone.0187543

Zhang, W., Wu, L., Ding, Y., Yao, X., Wu, X., Weng, F., Li, G., Liu, Z., Tang, S., Ding, C., & Wang, S. (2017). Nitrogen fertilizer application affects lodging resistance by altering secondary cell wall synthesis in japonica rice (Oryza sativa). *Journal of Plant Research*, *130*(5), 859–871. https://doi.org/10.1007/s10265-017-0943-3

Zhang, Y., Xiao, J., Yan, K., Lu, X., Li, W., Tian, H., Wang, L., Deng, J., & Lan, Y. (2023). Advances and Developments in Monitoring and Inversion of the Biochemical Information of Crop Nutrients Based on

Hyperspectral Technology. In *Agronomy* (Vol. 13, Issue 8). Multidisciplinary Digital Publishing Institute (MDPI). https://doi.org/10.3390/agronomy13082163

Zhang, Z.; Flores, P.; Igathinathane, C.; L. Naik, D.; Kiran, R.; Ransom, J.K. (2020). Wheat Lodging Detection from UAS Imagery Using Machine Learning Algorithms. Remote Sens. 12, 1838. https://doi.org/10.3390/rs12111838

Zhao, H., & Lifu, Z., (2013). A new method of continuum removal for mixing spectral analysis. 34th Asian Conference on Remote Sensing 2013, ACRS 2013. 2. 1304-1310.

Zhao, B., Li, J., Baenziger, P. S., Belamkar, V., Ge, Y., Zhang, J., & Shi, Y. (2020). Automatic wheat lodging detection and mapping in aerial imagery to support high-throughput phenotyping and in-season crop management. *Agronomy*, *10*(11). https://doi.org/10.3390/agronomy10111762

Zhonglin, Wang., Junxu, Chen., Jiawei, Zhang., Xian-he, Tan., Muhammad, Ali, Raza., Jun, Ma., Yan, Zhu., Feng, Yang., Wenyu, Yang. (2022). Assessing canopy nitrogen and carbon content in maize by canopy spectral reflectance and uninformative variable elimination. Crop Journal, doi: 10.1016/j.cj.2021.12.005

Appendix

A. Field and laboratory information

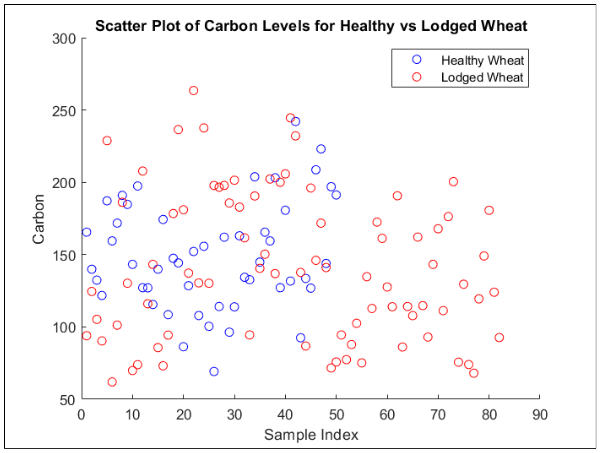
1. Pictures taken on the field during the field campaign in May 2023

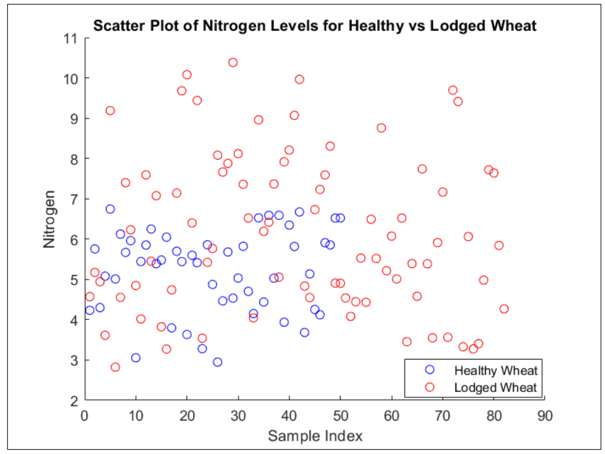
2. The wheat samples taken on the field were kept and taken to the laboratory for further processing.

3. Samples were dried in the oven (top), the Perkin Elmer CHN Analyzer-2400 Series used for nitrogen and carbon analysis (bottom).

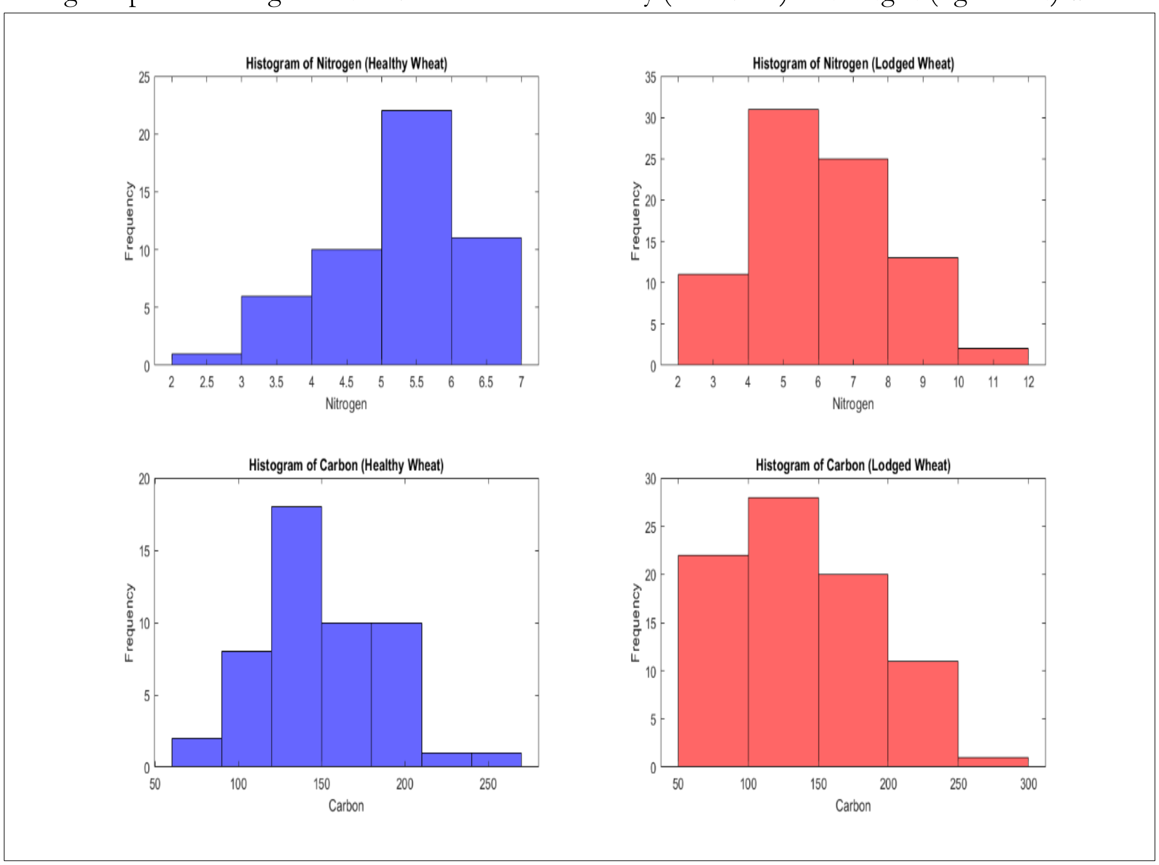
B. Hyperspectral information

1. Scattered plot of carbon (left) and nitrogen (right) content in healthy and lodged wheat

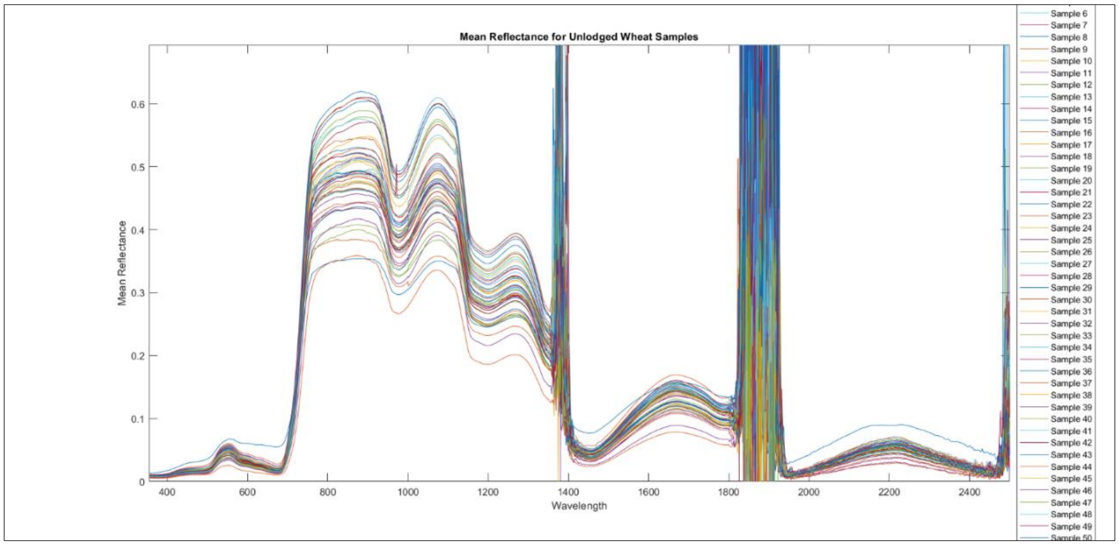




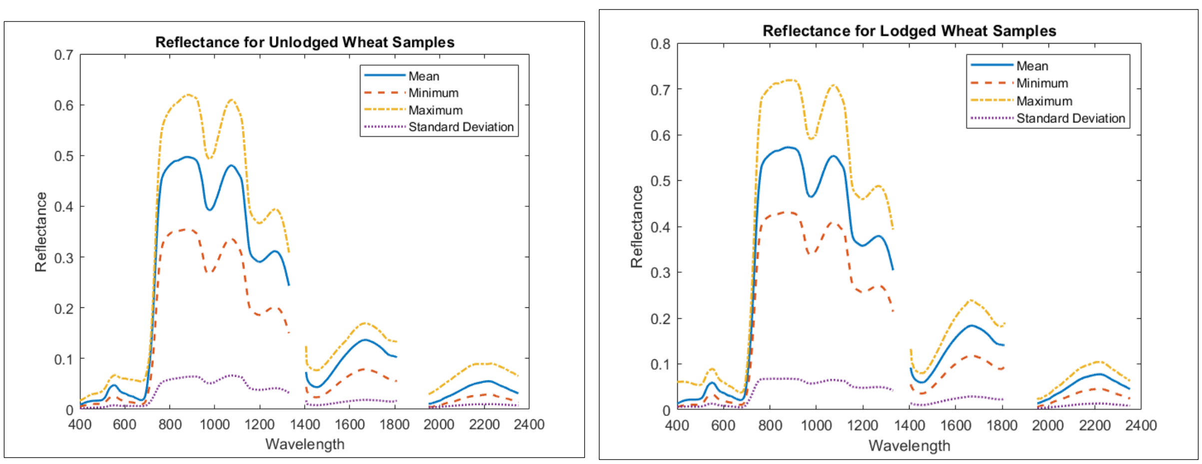
2. Histogram plot of nitrogen and carbon content in healthy (left - blue) and lodged (right - red) wheat.



3. The raw spectral reflectance of healthy wheat samples beforeremoval of noisy bands



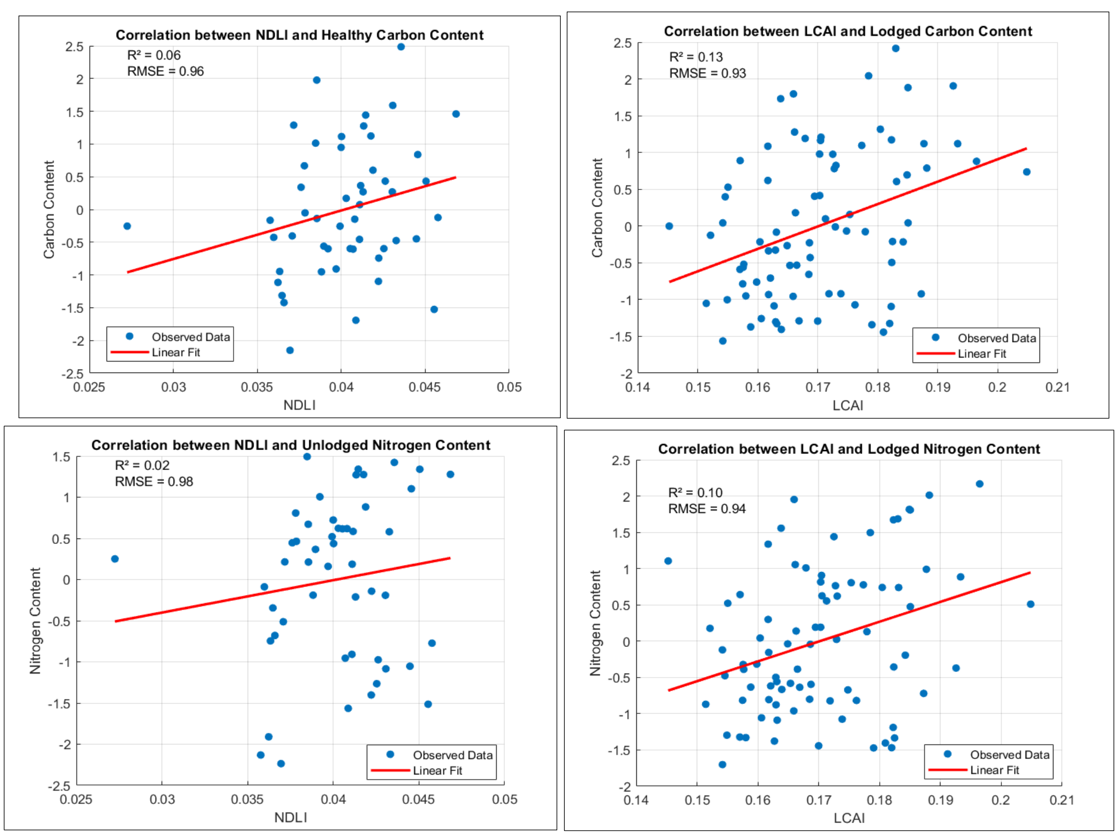
4. Average mean, minimum, maximum, and standard deviation of lodged (right) and healthy (left) wheat samples.



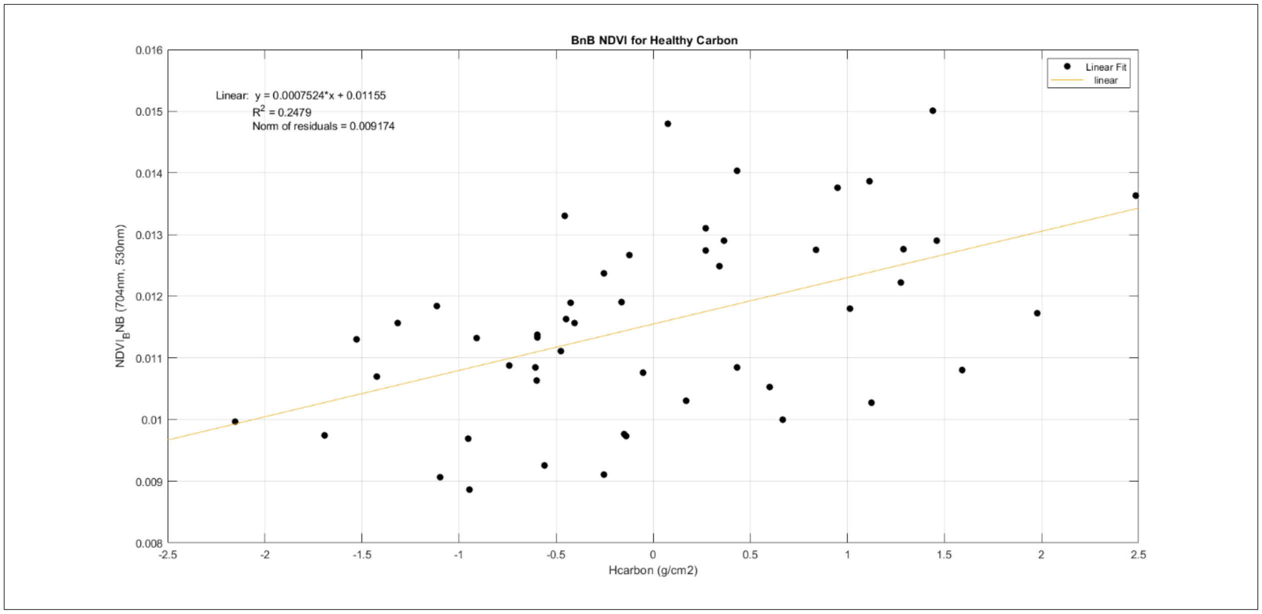
5. The result of vegetation indices used in estimation and retrieval of carbon and nitrogen using Sentinel-2 data

Vegetation indices	Lodged carbon	Healthy carbon	Lodged nitrogen	Healthy nitrogen
EVI	0.0006	0.0054	0.0016	0.0121
NDVI	0.00001	0.0039	0.0054	0.00042
PSRI	0.0022	0.0120	0.028	0.0051
RVI	0.0016	0.0055	0.0198	0.0016
CAI	0.00053	0.0029	0.0194	0.0011
LCAI	0.00051	0.00046	0.1026	0.0029
NDLI	0.00054	0.0021	0.013	0.018

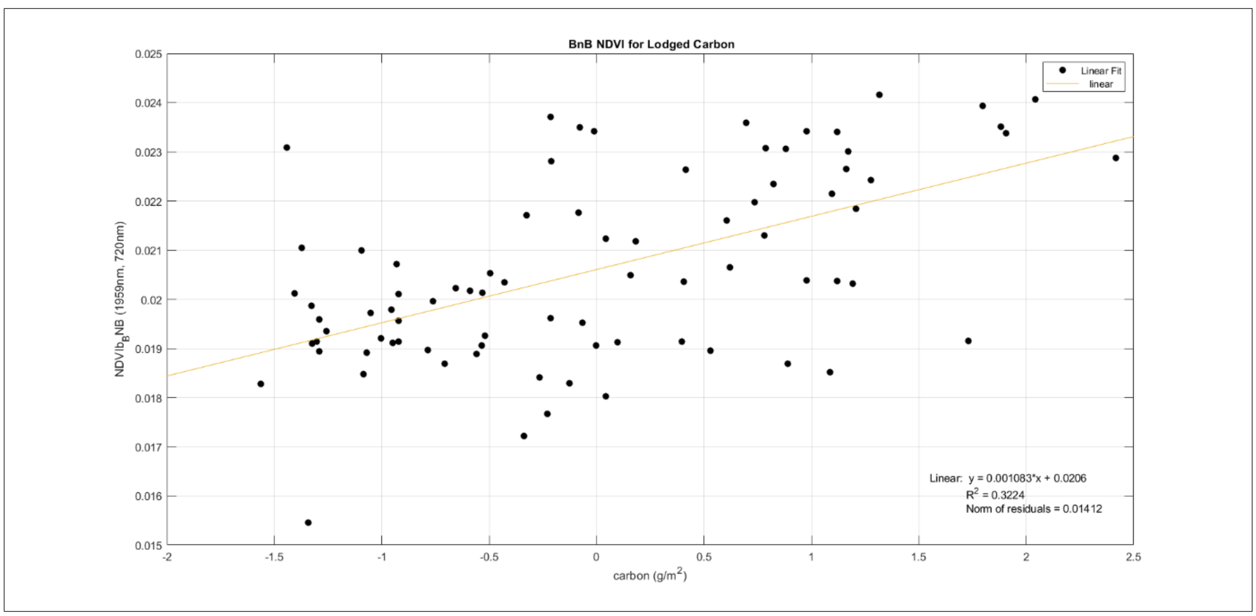
6. Vegetation indices used in this study that showed the best correlation result when used in prediction of lodged or healthy carbon and nitrogen content



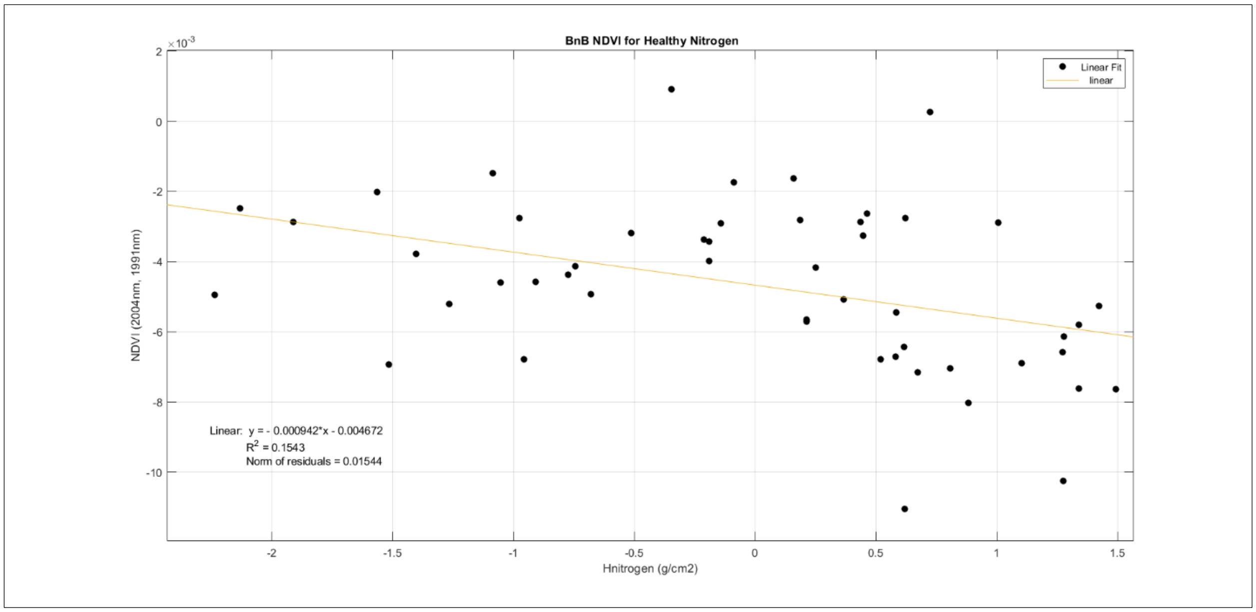
7. Best narrowband correlation graph for healthy carbon



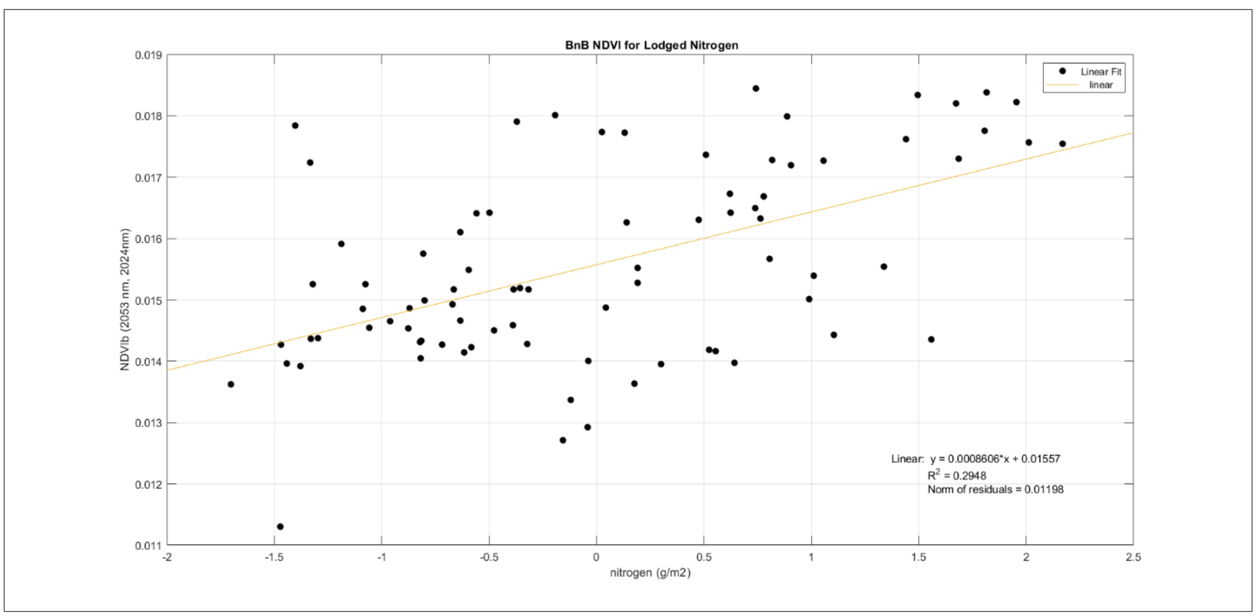
8. Best narrowband correlation graph for lodged carbon



9. Best narrowband correlation graph for healthy nitrogen



10. Best narrowband correlation graph for lodged nitrogen



11. Continuum removal bands of importance in both lodged and healthy wheat throughout the spectrum

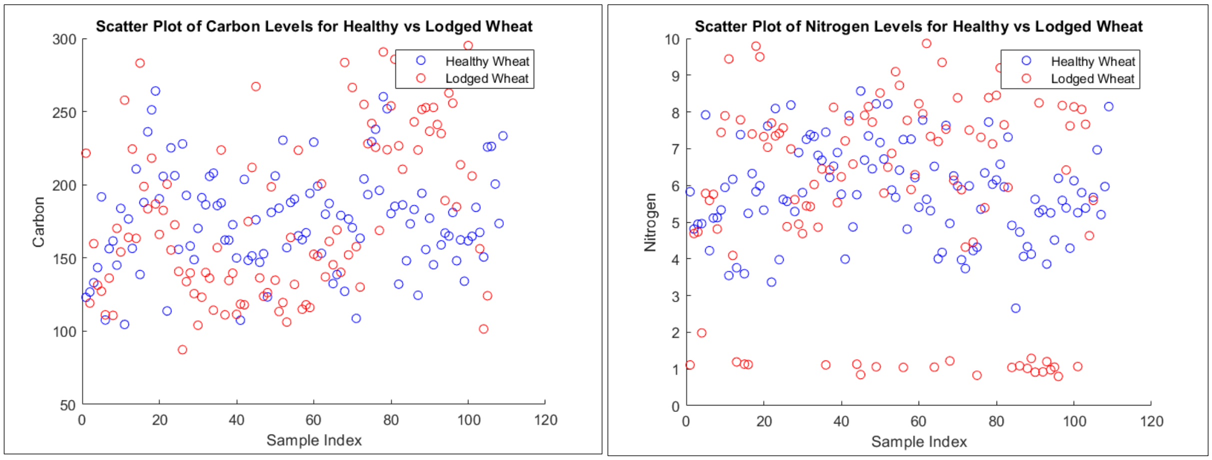
Region	Healthy / Lodged	Values	Band location (nm)	Difference
Visible	Healthy-Lodged	0.532743-0.570673	497	0.039
	Healthy-Lodged	0.372014-0.429055	672	0.057
NIR	Healthy-Lodged	0.744453-0.781867	714	0.038
	Healthy-Lodged	0.801403- 0.824614	977 & 976	0.023
	Healthy-Lodged	0.726994-0.76317	1165 & 1164	0.036
SWIR I	Healthy-Lodged	0.999716-0.999865	1320 & 1319	0.000149
SWIR II	Healthy-Lodged	0.524604 - 0.550526	1463	0.026
	Healthy-Lodged	0.952874-0.950651	1754 & 1775	0.002
SWIR III	Healthy-Lodged	0.830165-0.836237	2001 & 2001	0.006
	Healthy-Lodged	0.987304 & 0.989961	2270 & 2276	0.003
	Healthy-Lodged	0.984283 & 0.987887	2312 & 2309	0.004

12. Most important bands from feature selection

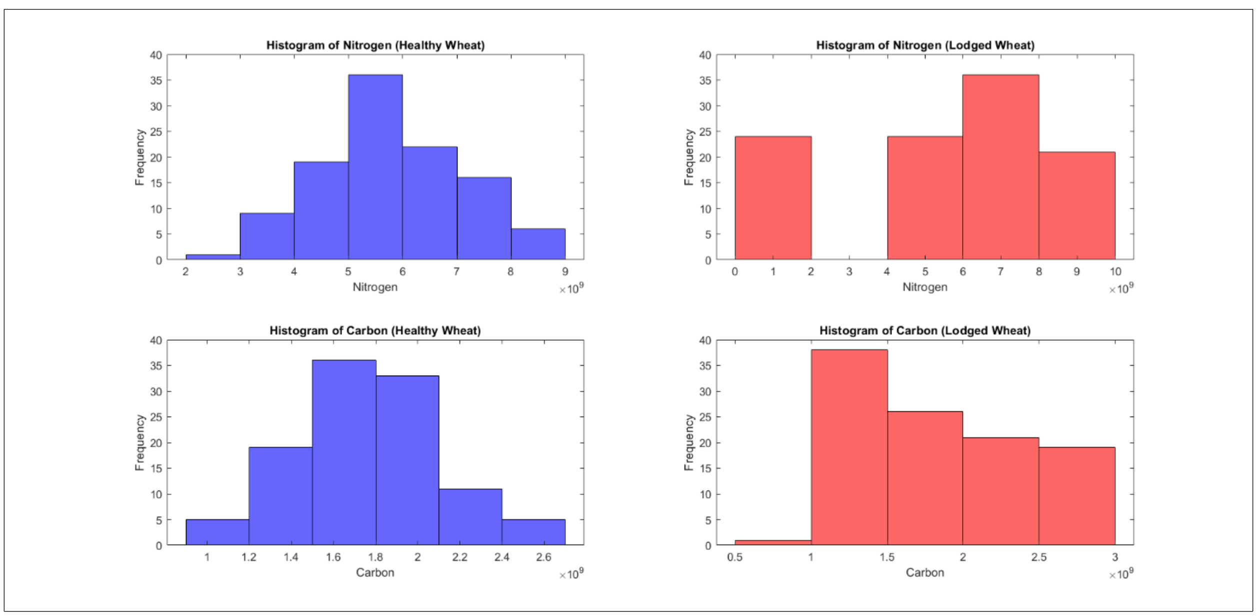
Most important bands	Healthy wheat carbon	Lodged wheat carbon	Healthy wheat nitrogen	Lodged wheat nitrogen
1	705	1347	682	1347
2	704	1346	680	1346
3	703	1345	681	1345
4	706	1348	683	1348
5	702	1344	679	1344
6	701	1343	684	1343
7	707	1349	678	1349
8	700	1342	677	1342
9	699	1350	685	1350
10	697	1341	676	1341
11	698	1351	686	1351
12	696	1340	675	1352
13	708	1352	674	1353
14	695	1353	687	1340
15	694	1354	673	1354
16	709	1339	688	1355
17	693	1355	672	1356
18	692	1356	689	1357
19	710	1357	690	1387
20	691	1358	671	1386

C. Sentinel-2 information

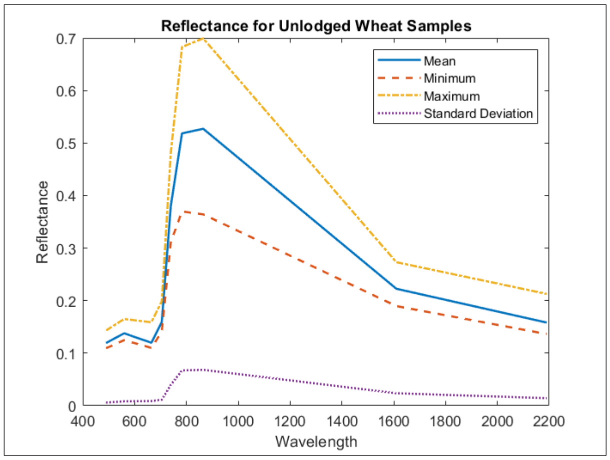
1. Scattered plot of carbon (left) and nitrogen (right) content in healthy and lodged wheat

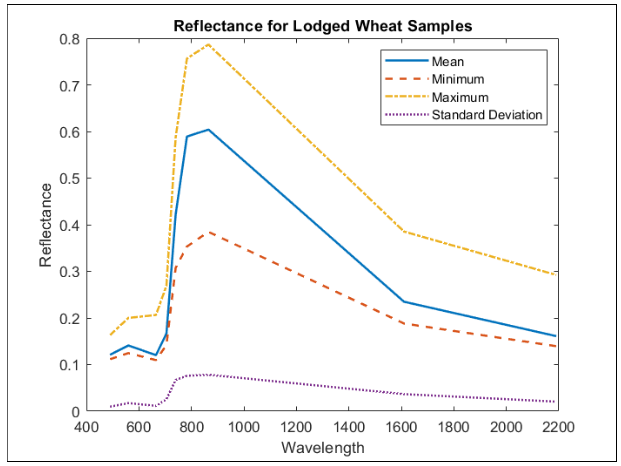


2. Histogram plot of nitrogen and carbon content in healthy (left - blue) and lodged (right - red) wheat.



3. Average mean, minimum, maximum, and standard deviation of lodged (right) and healthy (left) wheat samples.

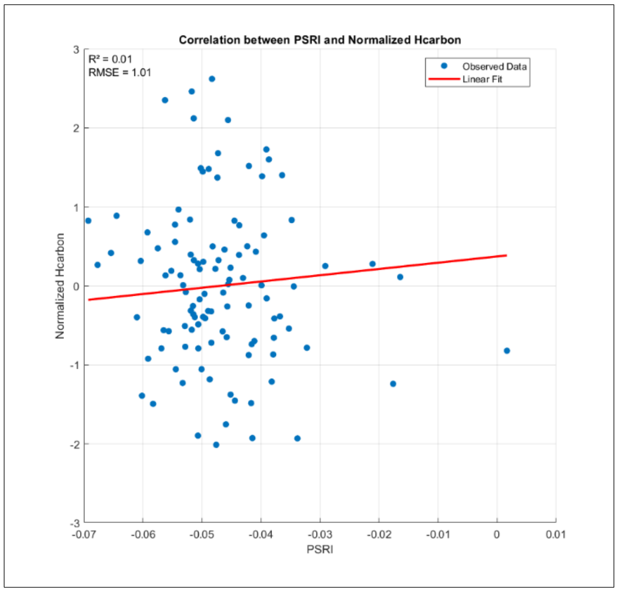


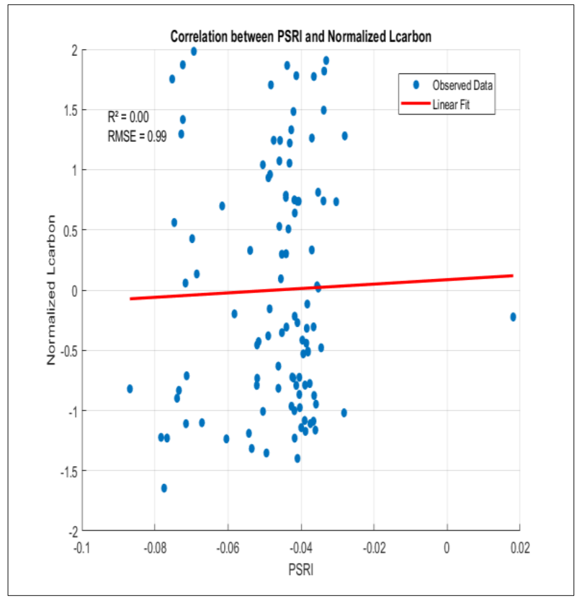


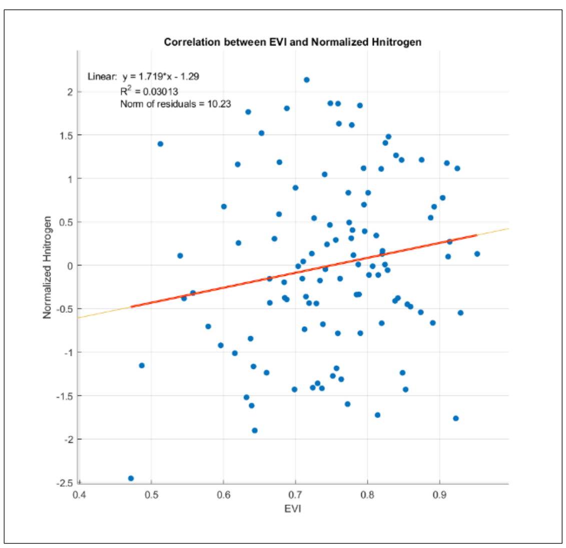
4. Vegetation indices used in this study that showed the best correlation result when used in prediction of lodged or healthy carbon and nitrogen content

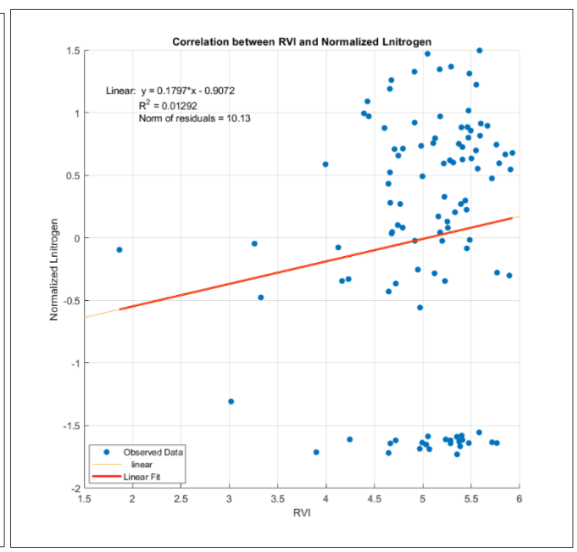
Vegetation indices	Lodged carbon	Healthy carbon	Lodged nitrogen	Healthy nitrogen
EVI	0.0041	0.0080	0.0016	0.0121
NDVI	0.0256	0.0011	0.0054	0.00042
PSRI	0.0071	0.0030	0.028	0.0051
RVI	0.0049	0.0071	0.0198	0.0016
CAI	0.0383	0.0257	0.0194	0.0011
LCAI	0.1277	0.00075	0.1026	0.0029
NDLI	0.0282	0.0625	0.013	0.018

5. Vegetation indices used in this study that showed the best correlation result when used in prediction of lodged or healthy carbon and nitrogen content


Correlation between PSRI and Normalized Hcarbon
R^2 = 0.01
RMSE = 1.01
Observed Data
Linear Fit
Normalized Hcarbon
PSRI


Correlation between PSRI and Normalized Lcarbon
R^2 = 0.00
RMSE = 0.99
Observed Data
Linear Fit
Normalized Lcarbon
PSRI


Correlation between EVI and Normalized Hnitrogen
Linear: y = 1.719*x - 1.29
R^2 = 0.03013
Norm of residuals = 10.23
Normalized Hnitrogen
EVI


Correlation between RVI and Normalized Lnitrogen
Linear: y = 0.1797*x - 0.9072
R^2 = 0.01292
Norm of residuals = 10.13
Observed Data
linear
Linear Fit
Normalized Lnitrogen
RVI

Data management plan

Section 1: Organisation and documentation

Name of the main folder(s):	MSC_THESIS
Name of secondary and tertiary folders (if applicable):	- HYPERSPECTRAL SPECTROSCOPY -2023 -SLANT HEIGHT (cm) -2023 -LODGE AREA (%) -2023 -SPAD -2023 -PLANT DENSITY -2023 -LEAVE AREA INDEX (LAI) -2023 -VERTICAL LODGE HEIGHT (cm) -2023 -FRESH BIOMASS (T/HA) -2023 -CANOPY HEIGHT (cm) -2023 -COVER PERCENTAGE (%) -2023 -SHOOT NUMBERS -2023 -POINT OF LINE FAILURE (cm) -2023 -TILLERS -2023 - SENTINEL 2 -2023 - ITALY SHAPE FILE -2023
Version control strategy	THE SPATIO-TEMPORAL ANALYSIS WAS SPLIT INTO DIFFERENT WHICH ACCORDING TO THE DIFFERENT VERSIONS CREATED, WILL HAVE THE FOLLOWING NAMES: NAME OF FOLDER ANALYSIS_NAME OF ORIGIN DATASET (SENTINEL 2) _V1…... THE MINORS VERSIONS WILL BE NAMED AS V1.1, V1.2…...
Metadata standards used:	ISO 191155-1
Readme file contents:	This M.SC_THESIS_DATASET NAME readme.txt file was generated on 20240631 by NAME. Oyetoun Olatorera, Alonge --------------------------------- GENERAL INFORMATION

Title of Dataset: M.SC_THESIS_DATASET
Author Information (OYETOUN M-GEO WO, ITC, Hallenweg 8, 7522 NH Enschede)

Principal Investigator:
Alonge, O.O. (Oyetoun, Student M-GEO-WO)
<o.o.alonge@student.utwente.nl>;

Date of data collection (multiple datasets,
From 2023.09.05-2024.05.18)

Geographic location of data collection: ITALY, GCS_WGS_1984
< 44°52'59"N, 11°58'48"E>

DATA & FILE OVERVIEW

File list:
M.SC_THESIS_DATASET
- HYPERSPECTRAL SPECTROSCOPY
    -2023
-SLANT HEIGHT (CM)
    -2023
-LODGE AREA (%)
    -2023
-SPAD
    -2023
-PLANT DENSITY
    -2023
-LEAVE AREA INDEX (LAI)
    -2023
-VERTICAL LODGE HEIGHT (CM)
    -2023
-FRESH BIOMASS (T/HA)
    -2023
-CANOPY HEIGHT (CM)
    -2023
-COVER PERCENTAGE (%)
    -2023
-SHOOT NUMBERS
    -2023
-POINT OF LINE FAILURE (CM)
    -2023
-TILLERS
    -2023
-SENTINEL 2
    -2023
-ITALY SHAPE FILE
    -2023

	Usage These datasets are free to use for research and non-commercial purposes. If you use these datasets in your work, please cite the original source of the data. **Sources** The datasets were collected from various sources, including government agencies, non-governmental organizations, and research institutions. The specific sources for each dataset are listed in the file headers. **Contact** If you have any questions or comments about these datasets, please contact the above-mentioned email address. I welcome any feedback or suggestions for improving the datasets or this repository.

Section 2 | Storage and sharing of research data.

Data	Storage Location	Back-up location and frequency	Strategy to prevent unauthorized access to data during research	Plan for pseudonymization or anonymization of data (if applicable)
MASTER FILES	To ensure data security and accessibility, the work will be stored on the local disk of my personal computer. A folder on the ITC cloud was also used for storage. However, I have decided to store the work on my local disks to protect the data from any unauthorized access by third parties.	The backup files will be securely stored in two locations: on Google Drive and on a hard disk. To ensure the safety and integrity of the data, a copy of the analysis will be saved at every stage of the work. This approach will help to minimize the risk of data loss or corruption and enable easy retrieval of previous versions if needed.	I will be responsible for working on my computer, which means that I will prevent unauthorized access to data during the research by securing my PC.	N/A
Copy of the Data	I will keep a copy of the data on my computer systems, as well as on Microsoft Teams and Google Drive for backup purpose.	To ensure that the project updates are safely backed up, a weekly backup will be conducted and stored in Google Drive	N/A	N/A