Assessment of Nutrients in Crop Canopy to Estimate Nutrients in Yield Using Ground Spectral Measurements
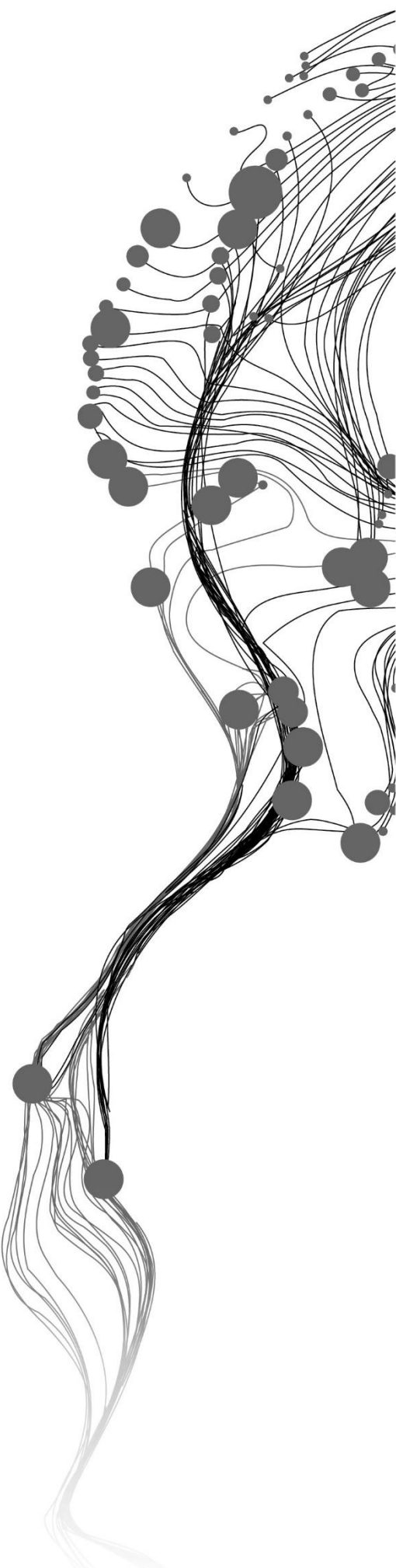
PAULINE KENDI MEME
July, 2024

SUPERVISORS:
Dr. Michael Marshall
Dr. Mariana Belgiu

Assessment of Nutrients in Crop Canopy to Estimate Nutrients in Yield Using Ground Spectral Measurements

PAULINE KENDI MEME
Enschede, The Netherlands, July 2024

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

SUPERVISORS:
Dr. Michael Marshall
Dr. Mariana Belgiu

THESIS ASSESSMENT BOARD:
Prof. A.D. Nelson (Chair)
Dr. Ir. S. Salama ((External Examiner, University of Twente - WRS)
Drs. R.G. Nijmeijer (Procedural advisor)

ABSTRACT

Assessing nutrient concentrations in crop canopies as it develops is crucial for estimating nutrient concentration in the final grain yield. This study investigates the use of ground spectral measurements to estimate eight macro- and micro-nutrient concentrations in wheat canopy across two growth stages (i.e., vegetative and reproductive) and grains. These nutrients are Calcium (Ca), Magnesium (Mg), Iron (Fe), Nitrogen (N), Zinc (Zn), Sulphur (S), Potassium (K), and Phosphorus (P). Field data were collected from 10 experimental sampling units in Italy, including biophysical and biochemical properties, ground spectral measurements, and laboratory-measured nutrient concentrations. Partial Least Squares Regression (PLSR) was applied to estimate nutrient concentrations in the canopy using ground spectral data.

Important spectral regions were identified using Variable Importance in Projection (VIP) scores and PLSR loadings, with Visible and Near Infrared being effective for nutrients associated with chlorophyll content and plant growth, such as N and P, while Shortwave Infrared was useful for nutrients associated with structural and protein components, such as Ca, Mg and S. The PLSR models explained the variances in nutrient concentrations with R^2 values ranging from 0.2 to 0.67 for the vegetative stage and 0.2 to 0.87 for the reproductive stage, showing moderate predictive capabilities ($R^2 > 0.6$) for most nutrients. The PLSR estimated nutrient concentrations in wheat canopy were compared with laboratory-measured nutrients concentration, reporting Root Mean Square Error (RMSE) values ranging from 0.2 to 1441.41 mg/kg for N and K, respectively. Nutrient concentrations in grains were estimated using canopy nutrient estimates from the PLSR models, resulting in R^2 values between 0.2 and 0.81 and Relative RMSE (RRMSE) values from 0.04% to 0.12% for Ca and Fe, respectively.

These results demonstrate the potential of using ground spectral measurements and PLSR to estimate nutrient concentrations in wheat canopy and grains, providing a non-destructive and efficient approach for assessing crop nutritional quality. Results on spectral response for various nutrients estimations are specifically important for future development or fine-tuning of remote sensing tools focused on nutrient estimations, which farmers can then use as decision support tools for crop nutritional quality management strategies. While the PLSR model predictions were promising, their accuracies can further be improved by validating these findings across different crops, environmental conditions and increasing number of samples. Additionally, comparative use of other models to enhance prediction accuracy and practical applications of ground spectral data for nutrient estimations to support farmers as a decision-making support tool is recommended. As demonstrated by the findings, integrating ground spectral measurements in estimation of macro- and micro-nutrients concentrations at key growth stages can contribute to informed timely decisions on optimizing agricultural practices to enhance the nutritional quality of crops and ensure food security in the face of changing environmental conditions.

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

AGB	Above ground biomass
ALIA	Absorbed leaf infrared region
Ca	Calcium
CCC	Canopy chlorophyl content
CNC	Canopy nitrogen content
EnMAP	Environmental Mapping and Analysis Program
ESU	Experimental Sampling Unit
Fe	Iron
GPC	Grain protein content
HNBs	Hyperspectral Narrow Bands
K	Potassium
LAI	Leaf area index
LCC	Leaf chlorophyll content
LMA	Leaf mass per area
LNC	Leaf nitrogen content
LWC	Leaf water content
Mg	Magnesium
MSAVI2	Modified Soil-Adjusted Vegetation Index
NDRE	Normalized Difference Red Edge
NIR	Near infra-red
OSAVI	Optimized Soil-Adjusted Vegetation Index
P	Phosphorus
PLS	Partial least squares
PLSR	Partial Least Squares Regression
PRISMA	PRecursore IperSpettrale della Missione Applicativa
R^2	Coefficient of Determination
REP	Red Edge Position
RF	Randon Forest
RMSE	Root Mean Square Error
RMSEP	Root Mean Square Error of Prediction
S	Sulphur
SAVI	Soil-Adjusted Vegetation Index
SVM	Support Vector Machine
SWIR	Shortwave infra-red
TSS	Total Soluble Solids
VIP	Variable Importance in Projection
Zn	Zinc

1. INTRODUCTION

1.1. Background

The global agricultural sector is facing the problem of decreasing nutrient concentrations in crops, a challenge further intensified by climate change implications on food and nutritional security (Myers et al., 2017, 2014). This decrease in crop nutrients concentrations could potentially amplify the prevalence of micronutrient deficiency, also referred to as "hidden hunger". This form of malnutrition currently affects many people in the world, as highlighted by Etienne et al., (2018) and Gashu et al., (2021). The inadequacy of micronutrients can trigger severe health complications like stunted growth in children, cognitive disorders, and weakened immune systems. Mitigating this issue is vital to attaining the United Nations' Sustainable Development Goal of "Zero Hunger" by 2030, aligning specifically with Indicator 2.1.1: "Prevalence of undernourishment" which aims to end hunger and ensure access by all people, especially the poor and those in vulnerable situations, to safe, nutritious and sufficient food all year round, and Indicator 2.2.2: "Prevalence of malnutrition" among children under 5 years of age (United Nations, n.d.).

Conventional methods for evaluating nutrient concentrations in crops involve field sample collection during the crop harvesting period followed by laboratory analysis (i.e., wet chemical analysis) to determine macro- and micro-nutrient compositions, as outlined in Caporaso et al., (2018). This approach, however, is laborious and financially strenuous, particularly for less developed nations. The advances in technology and ongoing studies in Earth Observation have increased interest in utilizing remote sensing imagery for various applications including using spectral vegetation indices to estimate leaf and canopy properties (Benami et al., 2021). Also, hyperspectral imagery and near-infrared spectroscopy is becoming increasingly useful to reduce the challenges of traditional methods of measuring key crop traits at leaf, canopy and ecosystem scales (Meacham-Hensold et al., 2019). Hyperspectral imaging (HSI) that combines imaging and spectroscopy, due to its capability to capture a wide spectrum of wavelengths, provides detailed data essential for assessing crop nutritional quality (Ma et al., 2022). Similarly, ground spectral measurements offers a non-destructive method for assessing plant biophysical and biochemical properties, essential in evaluating crop nutritional quality (Pandey et al., 2017).

Remote sensing is a powerful and versatile technology that enables the collection of information about Earth's surface and atmosphere at various scales (Wang et al., 2020). It has been widely used in multiple fields such as agriculture, forestry and environmental studies. Remote sensing has proven to be advantageous over traditional ground measurements techniques due to its convenience, efficiency, and cost-effectiveness, particularly over large areas and across time (Ni et al., 2018). The integration of remote sensing data with field-based information has been recognized as necessary, highlighting the importance of combining these technologies for making informed farming decision (Xie et al., 2013). Additionally, remote sensing has been used in

assessing and monitoring nutrient status in crops and their overall health to improve crop production by providing valuable information on nutrient stress (Adhikary et al., 2022). For example, it has been applied in in-season crop health monitoring like evaluation of soil moisture, hence contribute to efficient water use in agriculture (Singh et al., 2023). In general, remote sensing technology has and is still being used to effectively provide data at large spatial resolutions and high temporal resolutions to facilitate studying and addressing crop production related challenges (Adão et al., 2017; Benediktsson and Wu, 2021).

As aforementioned, remote sensing is an essential tool for evaluating nutritional content in crops. By retrieving biophysical crop parameters from remotely-sensed data, researchers can integrate remotely sensed data source directly into models used for estimation of nutritional quality (Sharifi, 2020). This approach has been recognized for its rapid, convenient, and non-invasive, providing a quicker and more reliable methods for evaluating the nutrient levels in crops than conventional methods (Huang et al., 2015; Sharifi, 2020).

Remote sensing methods are useful for evaluating crop health and productivity based on physiological, and biochemical properties like nutrient deficiencies and water access, and thus provided an avenue for identifying different and monitoring different types of crops in a spatial and temporal manner (Nguyen et al., 2019; Qian et al., 2019). Integration of proximal crop reflectance measurements with remote sensing technologies have been utilized to estimate the nitrogen nutrition index in crops highlighting the applicability of remote sensing in crop nutritional assessment as shown by Souza et al. (2020).

1.2. Problem statement

An accurate estimation and understanding of nutrients concentrations in crops at different crop growth stages is essential to ensure nutritious crop yields, Yet there are no extensive studies that track the dynamics of nutrients concentrations in crops throughout the development stages. Additionally, the relationship between nutrient estimates across various growth stages and their final concentrations in grains is not yet fully established. Also, the conventional methods for nutrient measurements are invasive, expensive and time-consuming. Although using hyperspectral reflectance data, such as ground spectral measurements, provide a promising alternative to the conventional method of collecting field samples and subsequently carrying out laboratory analysis to estimate the nutrient concentrations in crops, there remains an important research gap.

Specifically, it is unclear which spectral regions are the most important for accurate estimation of specific nutrients concentrations in crops at different growth stages. This study is the first of its kind to assess canopy nutrient dynamics using ground spectral measurements, providing insights into how nutrients are absorbed, utilized, and redistributed within the wheat plant during its development. By tracing the fate of nutrients from the canopy to the grains, this research offers a novel approach to understanding nutrient translocation and accumulation, which is critical for optimizing fertilization strategies and improving crop yield and quality.

The ground spectral measurements have inherent uncertainties that requires assessment, quantification and reduction to improve their applicability in estimating nutrient concentration in crops such that nutrient estimates closely match the actual measurements. Addressing these uncertainties is essential for improving the reliability of nutrient predictions based on the spectral data and contributing to timely crop management interventions.

1.3. Study relevance

The significance of this study lies in improving nutrient estimation in crop canopy by using ground spectral measurements, thereby reducing uncertainities in existing data on nutrient concentration in crop yields. By doing so, it identifies which nutrients can be estimated, identifies the best crop growth stages for those nutrients predictions, and quantifies the contribution of each nutrient estimate at specific growth stages to the final nutrient concentration in grains at maturity. As such, this study highlights on mechanisms of nutrient transport, using biophysical and biochemical properties captured by ground spectroscopy.

1.4. Objectives and research questions

1.4.1. Objectives

The main objective is to assess nutrient concentration in wheat canopy and grains using ground-spectral measurements

Specific objectives are:
  i. To estimate Calcium (Ca), Iron (Fe), Magnesium (Mg), Potasium (K), Phosphorus (P), Suplhur (S), and Zinc (Zn) concentration in the canopy at main wheat growth stages (vegetative and reproductive) using ground spectral measurements.
  ii. To estimate the nutrient concentration in wheat grain based on nutrient estimates from the main growth stages.
  iii. To compare the estimated nutrient concentration in the canopy and grains with the measured nutrients in the laboratory.

1.4.2. Research questions

To contribute to filling gaps in the literature, this research addresses these points:
  i. What are most important spectral bands that can be used in estimating nutrient concentrations in wheat at main growth stages?
  ii. What crop growth stage is most important in estimating nutrient concentration in grains?

1.4.3. Hypotheses
  i. We expect the spectral characteristics by red-edge bands during the reproductive growth stage to be the best predictors not only of nitrogen but also of overall nutrients concentrations in the canopy and subsequent grains.
  ii. We expect reproductive growth stage to show the highest impact on the nutrient concentrations in the grain due to the enhanced photosynthesis and grain filling processes during this stage, leading to increased nutrient accumulation in the grains.

1.5.   Structure of thesis

This thesis is structured as follows. The first chapter introduces the study, outlining the background, problem statement, relevance, objectives, research questions, and hypotheses. Second chapter provides a literature review on the link between climate change and crop nutritional quality, nutrient estimation methods, and the role of remote sensing and machine learning techniques. The third chapter describes the methodology, including the sampling design, data processing, and nutrient estimation techniques adopted for the study. The forth chapter presents the results and discussion, focusing on exploring the observed field data, PLSR model performance, and comparisons of estimated and observed nutrient concentrations. Finally, the last chapter concludes with key findings and recommendations.

2. LITERATURE REVIEW

2.1. Link between climate change and crop nutritional quality

Climate change has been identified as a crucial factor that affects the nutrition concentration in crops (Soares et al., 2019) . For instance, increased levels of carbon dioxide (CO_2), resulting from climate change, have been found to increase the carbon to nitrogen ratio in crops, which in turn reduces the grain protein content (Porter and Semenov, 2005). Similarly, changes in temperature and rainfall patterns can cause modifications in crop phenology, nutrient absorption, and metabolic activities, ultimately affecting the nutritional composition of crops (Marcos-barbero et al., 2021). For example, higher temperatures can accelerate plant growth and shorten the maturation period, leading to lower nutrient accumulation in crops. Moreover, shifts in rainfall patterns can affect the availability of water, resulting in decreased nutrient uptake and nutrient dilution as shown in Tao et al. (2017). These changes directly impact human needs and food security, emphasizing the importance of understanding the effects of climate change on crop nutritional quality as shown in studies by Asseng et al. (2019). As much as relationship between climate change factors and their impact on crop yields have been studied, not much attention has been on their impact on crop nutritional quality (Ahmed et al., 2019).

Wheat crop is no exception to effects of climate change. Past studies have shown that climate change stressors such as increased CO_2 levels, heat, drought, and salinity stress can alter wheat grain quality e.g., grain weight, nutrient content, fiber, protein composition, starch granules ,and free amino acid composition, among others and subsequently affecting their nutritional value (Ali et al., 2017). Genotypic variability in wheat under changing climatic conditions can lead to shifts in metabolic proteins, potentially influencing the nutritional value of wheat grains (Dwivedi et al., 2017). Chronic exposure to limited water and ozone has also been found to influence the nutritional quality of wheat (Soares et al., 2019). These findings highlight the complex interplay between climate change and crop nutritional quality, underscoring the need for sustainable agricultural practices focused on key crop growth stages to ensure nutritious crops for human health.

2.2. Estimation methods of nutrient concentrations in crops

The effective estimation of nutrient concentration in crops is essential to ensure optimal crop production and achieving food security as described by Lynch (2019). There are couple of methods that can be used, each with their inherent characteristics as described in the below subsections.

2.2.1. Field-based measurements

Field-based measurement is the most commonly used method to estimate nutrient concentration in crops. It involves field-based plant sampling and subsequent wet-chemical analysis in the laboratory. Some of the studies that have used this method include Cockson et al. (2019),

Maillard et al. (2015) and Roth et al. (1989). Roth et al. (1989) conducted plant tissue experiments to estimate nitrogen fertilizer requirements of winter wheat, emphasizing the use of specific tests to measure stem nitrate concentration, whole-plant Kjeldahl-N concentration, and crop N uptake between growth stages. Maillard et al. (2015) focused on leaf mineral nutrient remobilization during leaf senescence and modulation by nutrient deficiency. Leaf senescence represents a critical stage where nutrients are remobilized from aging leaves to other plant parts, such as developing seeds or storage organs, to ensure optimal nutrient utilization and plant growth. Nutrient remobilization during senescence is a tightly regulated process involving various molecular mechanisms, including autophagy, a cellular process that plays a pivotal role in recycling nutrients from senescing tissues to developing organs, such as seeds or storage compartments (Guiboileau et al., 2012). Similarly, Cockson et al. (2019) investigated the impact of soil test levels and fertilization with phosphorus and potassium of field crop tissues concentration, emphasizing the relevance of leaf tissue nutrient analysis. Moreover, Smith et al. (2019) researched the impact on field drought conditions on the nutrient concentration of leaf and seed tissue in common bean, providing insights into the influence of environmental stress on crop nutrient levels. Garcia and Grusak, (2015) focused on mineral accumulation in vegetative tissue during seed development and demonstrated the relevance of leaf tissue nutrient analysis in understanding nutrient dynamic in crops. These studies highlight that field-based methods are commonly used. However, the field-based methods are labour-intensive, crop invasive and thus reducing potential crop yield, expensive in terms of expertise and sophisticated equipment requirements, and neither suitable for large-scale nor real-time monitoring as described in study by Kalaji et al. (2018).

2.2.2. Remote sensing methodology

With the advancement in remote sensing technologies over the past decades, there are nowadays availability of sensors and imaging systems that provide hyperspectral data at large spatial coverages and high temporal resolution. For example, the Italian Space Agency developed and still manages a medium-resolution hyperspectral satellite known as PRecursore IperSpettrale della Missione Applicativa (PRISMA) that has been in operation since March 2019 to assist assessment of spectral signatures of different materials. The hyperspectral data is available at 30m spatial resolution and can capture images across 250 Hyperspectral Narrow Bands (HNBs) within a spectral range of 400-2500nm (Su et al., 2020). Environmental Mapping and Analysis Program (EnMAP) is another hyperspectral satellite mission by German Space Agency that provides over 240 continuous spectral bands in the wavelength range between 420 and 2450 nm with a ground resolution of 30 m×30 m and provides valuable geochemical, biochemical, and biophysical data for environmental monitoring and analysis. Another source of hyperspectral data is field spectroscopy that provides detailed ground spectral measurements as compared to satellite based information.

These hyperspectral data sources are progressively being applied to estimate nutrient concentration in various crops (Miphokasap et al., 2012). Visible or Near infrared reflectance spectroscopy has been successfully used for predicting nitrogen concentration at the canopy level in field crops (Jie et al., 2014). Hyperspectral data has been utilized to distinguish nitrogen, phosphorus and potassium deficiencies in crops, showcasing the potential of canopy

hyperspectral reflectance data for nutrient stress analysis (Liu et al., 2020). Additionally, studies have focused on use of drones-based canopy reflectance data to detect potassium deficiency in canola crop (Severtson et al., 2016). Similarly, remote sensing has been applied to obtain reflectance which can be used to derive vegetation indices under varying nitrogen concentration rates at diverse growth stages of rice (Din et al., 2019).

Studies have attempted to link nutrient concentration in crops during their development to that in grains, however the use of remote sensing is limited. For example, Peleg et al. (2009) analysed the concentration of various minerals in wheat grains focusing on their genetics and physiological characteristics. Similarly, Wang et al. (2021) explored how different manipulations of source (production of assimilates through photosynthesis) and sink (storage of assimilates in grains) can affect the accumulation of micronutrients and macronutrients in wheat grains. These studies did not apply remote sensing. Research, however, indicates that predicting nutrient content in crops through spectroscopy is feasible by determining spectral plant status using visible and near-infrared spectral responses from plant canopies as demonstrated in Bagheri et al. (2012). In the study, they employed various vegetative indices e.g., SAVI (Soil-Adjusted Vegetation Index), OSAVI (Optimized Soil-Adjusted Vegetation Index), and MSAVI2 (Modified Soil-Adjusted Vegetation Index) and identified that wavelengths between 630-860 nm were most suitable for diagnosing nitrogen content. Additionally, Zhao et al. (2019) focused on the use of Sentinel-2A satellite data to monitor and predict grain protein content in winter wheat thus facilitating nutrient management and harvest timing .The study successfully correlated Sentinel-2A vegetation indices with wheat nitrogen parameters and grain protein content. The prediction models used in the study demonstrated high accuracy, especially for plant nitrogen accumulation. Validation using multi-year and regional ground data confirmed reliable prediction and inversion results with Sentinel-2A imagery.

Literature shows that use of hyperspectral data for estimating nutrient concentration in crops are non-destructive alternative, and can offer timely assessment of nutritional concentration in crops as it develops, through analysis of canopy reflectance. However, remote sensing technologies also has its limitations in that they have inherent uncertainties because of random and systematic errors may originate from sensor calibration uncertainties, resampling or retrieval errors, inaccurate measurements and data collection processes, amongst others, that affect accurate estimation of nutrients in crops. Successful use of hyperspectral data also requires careful selection and extraction of specific predictive bands to represent accurate crop growth information (at a certain growth stage) for nutrient estimation which in most cases is unknown. For this purpose, various models including machine learning are often integrated with remote sensing data. For example, in the study by Tan et al. (2020), the Partial Least Squares (PLS) regression model was employed to predict the grain protein content (GPC) of wheat using remote sensing data.

2.3. Remote sensing on nutritional quality

2.3.1. Bio-physical and bio-chemical properties

Biophysical properties in crops refers to the physical characteristics and processes of plants including traits such as leaf area index (LAI), canopy structures, and spectral reflectance, which are essential for understanding the growth, development, and overall health of crops (Danson et al., 2003; Vitale et al., 2016; Wali et al., 2020). These properties are essential for estimating crop parameters, monitoring growth status, and predicting yield, making them valuable for agricultural production (Bahrami et al., 2021; Tomíček et al., 2021). On the other hand, bio-chemical properties are the chemical composition and characteristics of plants encompassing aspects such as bioactive compounds, enzyme activities and nutrient content, which directly influence the nutritional quality, physiological processes of crops(Dong et al., 2021).

The nutritional quality of crops is closely linked to their bio-physical and bio-chemical processes and traits as they play critical role in determining their productivity and nutritional value (Cooper et al., 2020). Additionally, the bio-chemical composition of crops, including the presence of bioactive and antioxidants, directly influence their nutritional quality (Keutgen et al., 2019). Some of the factors that have been found to affect the bio-physical and bio-chemical properties of crops include the soil properties and other human induced applications like fertilizer applications, that ultimately affect the composition of nutrients in crops (Arnó et al., 2012; García-Gaytán et al., 2018). Also, the genetic potential of crops and their growing conditions have been found to influence their bio-chemical and nutritional composition (Uprety et al., 2010).

Understanding the closely link between the bio-physical and bio-chemical properties of crops is essential for optimizing agricultural practices to enhance the nutritional quality of crops and ensure food security in the face of changing environmental conditions (Reyes et al., 2021;Challinor et al., 2010). For this purpose, the use of remote sensing applications are being explored to provide insights into crop nutritional status(Din et al., 2017; Thorp et al., 2012; Yang et al., 2011). Also, comparing field-measured spectral data with models liked PROSAIL. PROSAIL, a radiative transfer model, can aid in determining biochemical properties like leaf chlorophyll content by simulating leaf reflectance spectra based on known biochemical and structural leaf parameters as indicated by Berger et al. (2018a) and Zhao et al. (2004). By utilizing PROSAIL, the spectral responses of leaves to different chlorophyll concentrations can be analysed, enabling the estimation of chlorophyll content non-destructively and accurately. As such, this model facilitates the interpretation of hyperspectral data to derive chlorophyll-related vegetation indices, providing valuable insights into plant health and nutrient concentrations. Models such as PLSR have also been developed and used to predict photosynthetic and biochemical properties in wheat canopies (Robles-Zazueta et al., 2022). Also as shown in Liu et al. (2015) crop traits retrieval can give more detailed information for better understanding of wheat biophysical properties.

2.3.2. Spectral bands

Spectral bands are essential in assessing nutrient concentration in crop canopies. By analyzing the reflectance patterns of different spectral bands, specific nutrient deficiencies or imbalances in crop canopies can be identified (Hatfield et al., 2008). For example, hyperspectral data has enabled the estimation of canopy nitrogen concentration to demonstrate their applicability in timely determination of nutrient status in crops as shown in the study by Hatfield et al. (2008).

An attempt to demonstrate the potential use of field spectroscopy in estimating nutrient quality was made in Adão et al. (2017) where crop chlorophyll content was assessed. In a study by Hou et al. (2022), the reflectance spectral features at 550nm and the near-infrared (NIR) band was shown to be crucial in determining chlorophyll content and leaf cell structure in crops, emphasizing the importance of specific spectral bands in assessing nutrient quality. In a study by Liu et al. (2020), hyperspectral data was used to enable them differentiate N, P and K deficiencies in winter oilseed, emphasizing the significance of spectral bands in discriminating different nutrient stress conditions. Zhou et al. (2016) found airborne hyperspectral reflectance measurements to be effective in remotely estimating canopy nitrogen content in winter wheat, which highlights the potential of hyperspectral bands in assessing nutrient levels. A study by Bagheri et al. (2012) identified that wavelengths between 630-860 nm were most suitable for diagnosing nitrogen content.

Literature shows that some studies have identified several important spectral bands for assessing nutrients in crops. For example, in their study, Mejía-Correal et al. (2023) explored a range of wavelengths from visible to shortwave infra-red (SWIR) for estimating Total Soluble Solids (TSS) in grapes and found the tested range of wavelengths to be effective for TSS estimation. Pereira-Obaya et al. (2023) identified distinct wavelengths such as 560, 680, 1400, and 1935 nm, as effective for assessing plant health in chestnut trees. Similarly, Stagnari et al. (2023) highlighted the relevance of the blue-cyan and early NIR spectral regions in determining nitrate content in spinach.

Field studies have shown that vegetation indices like Red Edge Position (REP), calculated from narrow spectral bands in the red-edge region (around 700 ±40 nm) correlates strongly with nitrogen content (González-Piqueras et al., 2017). This is because reflectance in the red-edge region is sensitive to crop canopy chlorophyll, which is closely related to nitrogen status (Clevers and Gitelson, 2013). Red-edge bands have also shown importance in estimating other elements essential like chlorophyll content and photosynthetically active radiation for crop health apart from nitrogen as described in Mashiane et al. (2023). In a study by Prananto et al. (2020) it was found that NIR spectroscopy can detect macro-nutrients such as N, P, and S directly because they are major constituents of NIR-sensitive organic compounds. On the other hand, macro-nutrients and micro-nutrients that exist mostly in inorganic forms such as Ca, Mg, and K are detectable through association with organic compounds and indirect correlations with organic compounds.

Weiss et al.(2020) conducted a research exploring assessment of N on wheat leaves using the red-edge position (700 nm) demonstrating the ability of the technique to estimate nitrogen content remotely. A study by Santos et al. (2013) explored the detection of P and K in wheat grains using hyperspectral imaging pinpointed certain spectral bands within the visible and near-infrared spectrums as particularly responsive to variations in the macro-and micro-nutrients. Prananto et al. (2020) explores the potential of NIR bands ((700–2500 nm) to estimate both macro- and micronutrient contents in plant leaves, indicating effective prediction of nutrient content, but it primarily targeted leaf tissues rather than grains. Their study found that both macro- (N, P, K, S, Ca, Mg) and micronutrients (Fe, Zn, Mn, Cu) in plant leaves could be accurately predicted within the NIR bands. Macronutrients, which are more reliably detected due to their presence in NIR-sensitive organic compounds, showed high prediction accuracies. For instance, N had a median calibration R^2 of 0.98 and validation R^2 of 0.94, P had a median calibration R^2 of 0.91 and

validation R^2 of 0.62, and K had a median calibration R^2 of 0.94 and validation R^2 of 0.64. In contrast, micronutrients exhibited lower prediction accuracies, such as Fe with a median calibration R^2 of 0.81 and validation R^2 of 0.72, and Zn with a median calibration R^2 of 0.85 and validation R^2 of 0.62. These results highlight the greater reliability of NIR spectroscopy in detecting macronutrients compared to micronutrients. They attributed this to the fact that micronutrients and some macronutrients, primarily existing in inorganic forms, are identified through their association with organic compounds and indirect correlations.

A couple of indices have been explored for estimating crop. For instance, indices based on red-edge, like the Normalized Difference Red Edge (NDRE), have been recognized as successful methods for assessing nutrients in crops using spectral analysis (Prey et al., 2018). .NIR spectroscopy has also been utilized to rapidly and accurately assess various nutritional components of wheat, such as grain protein concentration, chlorophyll content (Jiao et al., 2022; McGRATH et al., 1997). Schirrmann et al. (2016) developed multispectral indices to assess N, P, and K concentrations in wheat canopies, emphasizing the potential of theses indices for integrated nutrient management. These examples illustrate the importance of selecting precise spectral bands for accurate nutrient analysis.

2.4. Integration of machine learning in crop nutrient estimation

The use of machine learning techniques to estimate crop nutrients and their nutritional quality has been explored using various methods, including remote sensing and hyperspectral canopy sensing (Cummings et al., 2021; de Oliveira et al., 2021; Stanton et al., 2017; Yu et al., 2018). Integrating ground spectral measurements with machine learning approaches, like Partial Least Squares Regression (PLSR), allows for analysis of spectral data, facilitating accurate and efficient assessment of nutritional quality (Furbank et al., 2021).

Application of machine learning techniques and spectral indices from un-manned vehicles (UAV) have been explored for estimating chlorophyll content in crops (Narmilan et al., 2022). By integrating machine learning and NIR spectroscopy, Sun et al. (2023) was able to retrieve canopy chlorophyll content and leaf area index, while Li et al. (2022) was able to estimate nitrogen nutrition index in crops. Some other applications showing integration of machine learning in assessing crop nutritional quality include that by Narmilan et al. (2022) in predicting canopy chlorophyll content in sugarcane; and crops. Muñoz-Huerta et al. (2014) incorporated machine learning algorithms with NIR canopy reflectance to determine crop N requirements. The use of machine learning to estimate nutrients concentrations was reported as successful due to its ability to provide accurate, efficient, and non-invasive monitoring of crop health and nutritional needs.

The application of machine learning has some limitations that should be considered for their successful use in estimating crop nutrients. For example, the performance of machine learning techniques depends on spectral data used for training, which can affect the accuracy of retrieving crop variables from spectral reflectance data (Revill et al., 2020). Additionally, the effectiveness of spectral imaging techniques can be influenced by external factors such as lighting conditions and the presence of noise, which may impact the accuracy of the results (Li et al., 2020). The integration of spectroscopy and machine learning to evaluate nutrient levels requires some level of expertise for successful implementation and optimization (Li et al., 2020). Some of the

machine learning techniques that can be used for estimating crop nutrient quality include PLSR, random forest (RF) and support vector machine (SVM), among others.

2.5.   Partial Least Squares Regression

2.5.1.   Description and its application

PLSR is a statistical technique extensively applied across multiple disciplines, including agriculture, to assess and estimate nutrient content in crops (Peng et al., 2019). This technique is useful when dealing with a large number of predictors, in this case independent variables, and limited sample size. Unlike conventional regression techniques, PLSR creates latent variables or components. These components are designed to capture the most variance in both the predictors and explanatory variables by finding the linear combinations that best describes their relationship.

In literature, PLSR is recognized as a powerful calibration tool in chemometrics projecting data into a reduced-dimensional space formed through orthogonal latent components, maximizing the covariance between spectral data and chemical concentration matrices, as noted by Vohland and Emmerling (2011). It has been utilized in multiple agricultural applications, for instance, estimating leaf nutrient contents in crops (De Silva et al., 2023), estimate above ground biomass in winter wheat (Yue et al., 2018), or predicting soyabean yields under varied water applications (Crusiol et al., 2021). The PLSR technique has also been used to determine leaf nutrient levels in Eucalyptus through leaf hyperspectral reflectance as explored by de Oliveira and Santana (2020), and to evaluate soil phosphorous concentrations in rice paddies, as demonstrated by Kawamura et al. (2019). PLSR has also been utilized to predict different macro- and micro-nutrients e.g., carbon, nitrogen, and phosphorous concentrations in crops across the globe, for example its application in China by Xu et al. (2017).

The PLSR technique has been recognised for its potential in estimating bio-physical and bio-chemical properties in crops. It has shown superior performance over conventional techniques like multiple linear regression in estimating a wide range of vegetation bio-physical and bio-chemical properties (Yue et al., 2018). Studies have documented its successful use to accurately estimate leaf area index (LAI), above ground biomass (AGB), nitrogen levels and chlorophyll content when integrated with hyperspectral remote sensing data (Berger et al., 2018b; Darvishzadeh et al., 2008; Maqbool et al., 2012; Muñoz-Huerta et al., 2013; Robles-Zazueta et al., 2022; Thorp et al., 2012).

Some additional uses of PLSR is in estimation of various crop leaf water content, by incorporating remote sensing properties like backscattering coefficient or brightness temperatures as highlighted by Elmetwalli et al.(2021) and Vereecken et al. (2012). For their calibration and validation in the context of assessing crop characteristics, samples from cover crops fields are usually used (Xia et al., 2021). These studies and examples demonstrate the unique opportunity of integrating hyperspectral remote sensing with PLSR that can provide non-invasive and timely alternative method for accurate estimation of crop bio-physical and bio-chemical traits towards sustainable crop production, also echoed in Naik et al. (2020).

2.5.2.   How PLSR works

As mentioned above, PLSR is a statistical method used to model complex data structures where the predictor variables are many and highly collinear. It addresses issues in multiple linear

regression (MLR) see Eqn. (2.1), particularly where the predictor matrix X might be singular due to more variables than samples or due to multicollinearity among the variables.

$$Y = X\beta + \varepsilon \tag{2.1}$$

where β is the regression coefficient matrix and ε is the error.

The PLSR overcomes the singular matrix in MLR (X^T) by transforming X into a new space following Eqn. (2.2).

$$X = TP + \varepsilon \tag{2.2}$$

where T are the scores, P the loadings, and ε the residuals.

Therefore, instead of regressing Y on X, PLSR regresses Y on the scores T, which involves calculating the regression coefficients (B) following Eqn. (2.3).

$$B = P(T^TT)^{-1}T^TY \tag{2.3}$$

The PLSR algorithm then iteratively calculates weights for X and Y to maximize the covariance between transformed X (scores t) and Y as shown in Eqn. (2.4) and (2.5).

$$t = Xw \tag{2.4}$$

$$u = Yq \tag{2.5}$$

where w and q are weight vectors for X and Y, respectively. The scores t are often normalized by their norm.

After calculating scores and loadings, PLSR deflates the X and Y matrices, removing the information explained by the extracted components following Eqn. (2.6) and (2.7). The process is repeated until a specified number of components are extracted or until the deflation yields no further information. After every iteration , vectors w,t,p and q are saved as columns in matrices W,T,P and Q, respectively

$$E_{n+1} = E_n - tp^T \tag{2.6}$$

$$F_{n+1} = F_n - tq^T \tag{2.7}$$

where E and F are initialised as X and Y, respectively.

Using the scores T, the final model coefficients are calculated to map back to the original predictor space following Eqn (2.8). R is a transformation matrix linking weights and loadings to align with the original variable space as shown in Eqn (2.9).

$$B = R(T^TT)^{-1}T^TY \tag{2.8}$$

$$R = W(P^TW)^{-1} \tag{2.9}$$

The PLSR components, loadings, and scores play pivotal roles in capturing and interpreting the variance in data.

2.5.3. Number of components and scores

Components, also known as latent variables, are derived from the data to maximize the covariance between predictor variables X and response variables Y as highlighted by Rosipal and Krämer (2006). The number of components used is a critical parameter and is typically determined via cross-validation to prevent overfitting. Detailed description of PLSR algorithm can be found in Mevik and Wehrens (2007).

Scores represent the projection of the original data onto the space defined by the PLSR components, showing how much information each component captures from original data according to McIntosh and Lobaugh (2004).

2.5.4. Loadings

Loadings indicate the weights assigned to the original variables in creating the latent variables (components) that maximize the covariance between predictor variables X and response variables Y. As such, the loading factor indicate how each element in X contributes to explaining Y within each component (El-Hendawy et al., 2019). This aids in interpreting the relationships between X and Y for nutrient estimation and assessment in crops. They help in understanding which variables are most influential in defining the components, thereby shedding light on the underlying patterns within the data.

Together, these elements help in reducing dimensionality, managing multicollinearity, and enhancing the interpretability of complex datasets, making PLSR particularly valuable in fields like chemometrics and spectroscopy.

3.   METHODOLOGY AND DATA

3.1.   Study area

3.1.1.   Geographical location, topography and climate

The study was conducted within a region in Italy as shown in Figure 1. It spans the geographical range of 44°87'0"N to 44°89'0"N and 11°95'0"E to 11°97'0"E. The area lies on the eastern edge of the Po River plain, one of the most fertile and agriculturally productive regions in Italy. The soils in the area are predominantly alluvial soils formed from sediment deposits of the Po River and its tributaries as highlighted by Giuseppe et al. (2014). These soils compositions range from sandy loam to silty clay loam, providing optimal drainage and water retention beneficial for wheat cultivation and are nutrient-rich due to periodic floodwater enrichment depositing silt and organic matter.

The study area is characterized by temperate climate with distinct summer and winter seasons. It experiences daily temperatures that range between 3.9 to 23.1°C during the mild winter and humid summer, respectively as stated by Poggi et al. (2022). In winter season, the mean daily temperature can fall up to -4°C according to Mercuri et al. (2013). The area receives monthly rainfall of between 41.1 to 54.6 mm during winter and summer periods, respectively according to Belgiu et al. (2023).
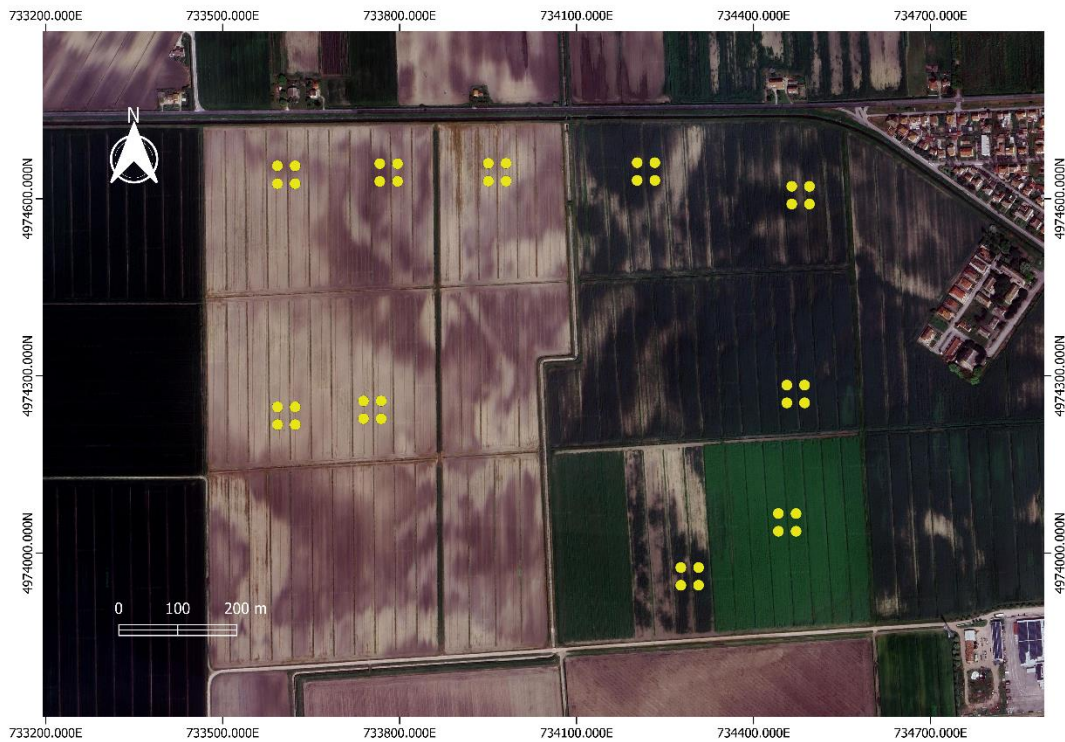
Figure 1: Study area showing the points (in yellow) where wheat samples were collected i.e., experimental sampling units.

3.1.2. Selected crop of interest

In this study, wheat (*Triticum aestivum L.*) was selected as the crop of interest due to its paramount importance in global food security as a staple cereal alongside rice and maize as noted by Shrawat and Armstrong (2018). According to Tagliabue et al. (2022), wheat is one of the main crops grown in the area besides barley, corn, sugar beet, alfalfa, soybean, rice, and medicinal plant. According to the Food and Agriculture Organization (FAO), wheat is the second most-produced cereal worldwide, following maize, with a production of over 731 million tonnes in 2019. As a predominant staple crop in temperate zones, it is a primary food source, providing about 20% of the daily protein and food energy for 4.5 billion people, underscoring its economic and nutritional importance as highlighted in Shiferaw et al. (2013).

Wheat crop is adaptable to extensive cultivation and long-term storage allowing it to be used to support approximately 35% of the global population despite challenges from pests and diseases. In the present study area, the wheat crop is planted and harvested in March and July, respectively. The predominant varieties of wheat cultivated in the area include augusto, bologna, bramante, cesare, giorgione, odisseo, orobel, santorin, sypassion, and titolavio (Belgiu et al., 2023).

The selection of macro-nutrients (Ca, Mg, K, S, P, and N) and micro-nutrients (Fe and Zn) is crucial due to their essential roles in plant growth and human health. These nutrients play vital roles such as bone health (Ca, Mg), oxygen transport (Fe), amino acid synthesis (N), immune function (Zn), and energy production (P). Deficiencies in these nutrients can lead to health issues like anemia, weakened immunity, bone disorders, and impaired growth.

3.2.   Sampling design for field measurements collection

For collection of data, a field campaign was conducted during the summer of 2023. To facilitate this, a total of ten experimental sample units (ESU) for each wheat growth stages were designed. Two growth stages, consistent with the present study objectives were considered: vegetative, reproductive.

The ESUs were of 60 m x 60 m in size. Every ESU consisted of four plots of 15 m x 15 m dimension as shown in Figure 2. Within each plot, five locations were selected from where wheat canopy spectral measurements together with bio-chemical and bio-physical parameters were taken. The bio-physical and bio-chemical parameters collected were leaf area index (LAI), leaf water content (LWC), leaf mass per area (LMA), average leaf inclination angle (ALIA), leaf chlorophyll content (LCC), canopy chlorophyl content (CCC), leaf nitrogen content (LNC) and canopy nitrogen content (CNC). In summary, there were 10 ESUs containing four plots each to result in 40 samples being collected for every growth stage.
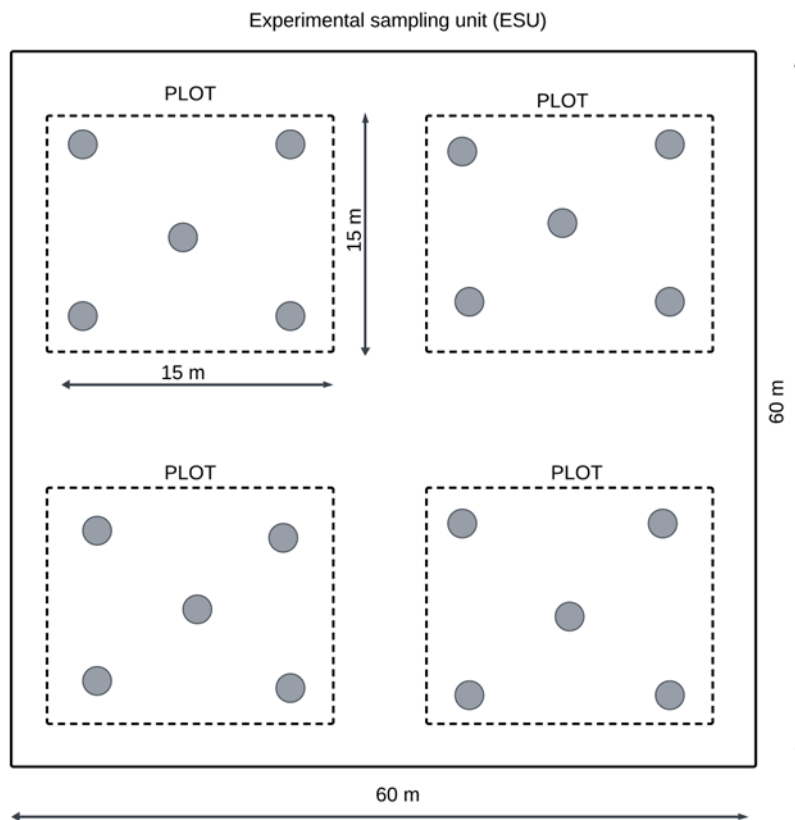
Figure 2: Schematic representation of the experimental sampling units (ESUs).

To obtain a single spectral measurement for a plot in each of the growth stages, each of the spectral measurements from the five points were averaged. The fresh whole plant samples were separated into leaf, stalk and grain compartments. Thereafter, fresh and dry weight of each of the compartments were obtained in the field before being taken to the laboratory for nutrient

analysis. These spectra and nutrient measurements were provided by the supervision team in comma separated values (CSV) format.

3.3.   Data pre-liminary processing

Although the ground spectra measurements and the actual nutrient measurements were initially deemed as "ready for analysis", further scrutiny to verify their completeness and accuracy were incorporated. This involved handling of missing data points within the time series, removing duplicates and carrying out cleaning procedures on the datasets to enhance data quality. Additionally, checks were implemented to ensure the input files, specifically column names that represent nutrients being analysed and formats, were correctly structured to meet the input requirements of the PLSR model. This pre-processing was essential to ensure the quality of the datasets for analysis and subsequent study findings.

The ground spectral measurements have inherent uncertainties that requires assessment, quantification and reduction to improve their applicability in estimating nutrient concentration in crops such that nutrient estimates closely match the actual measurements. In this study, the ground spectral measurements underwent several preprocessing steps to ensure data quality. First, the data were checked for completeness and duplicates. Thereafter, a logarithmic transformation was applied to stabilize variance due to anomalies resulting from spectra peaks, followed by smoothing with a spline filter. Atmospheric windows (1338-1449, 1793-1993, and 2356-2500 nm) were omitted from the analysis to exclude water bands and other interferences following Belgiu et al. (2023).

3.4.   Estimation of nutrient concentration in canopy

Figure 3 illustrates the methodology adopted for addressing the research questions and achieve the objectives. To address the first objective and related question, PLSR detailed in section 2.5 was used to analyse the full spectral response at main wheat growth stages (vegetative and reproductive) to the nutrient concentrations in the canopy. The decomposition of a matrix of ground-spectral measurements into linear combinations using partial least squares (PLS) was utilized to explain variations in nutrient concentrations.
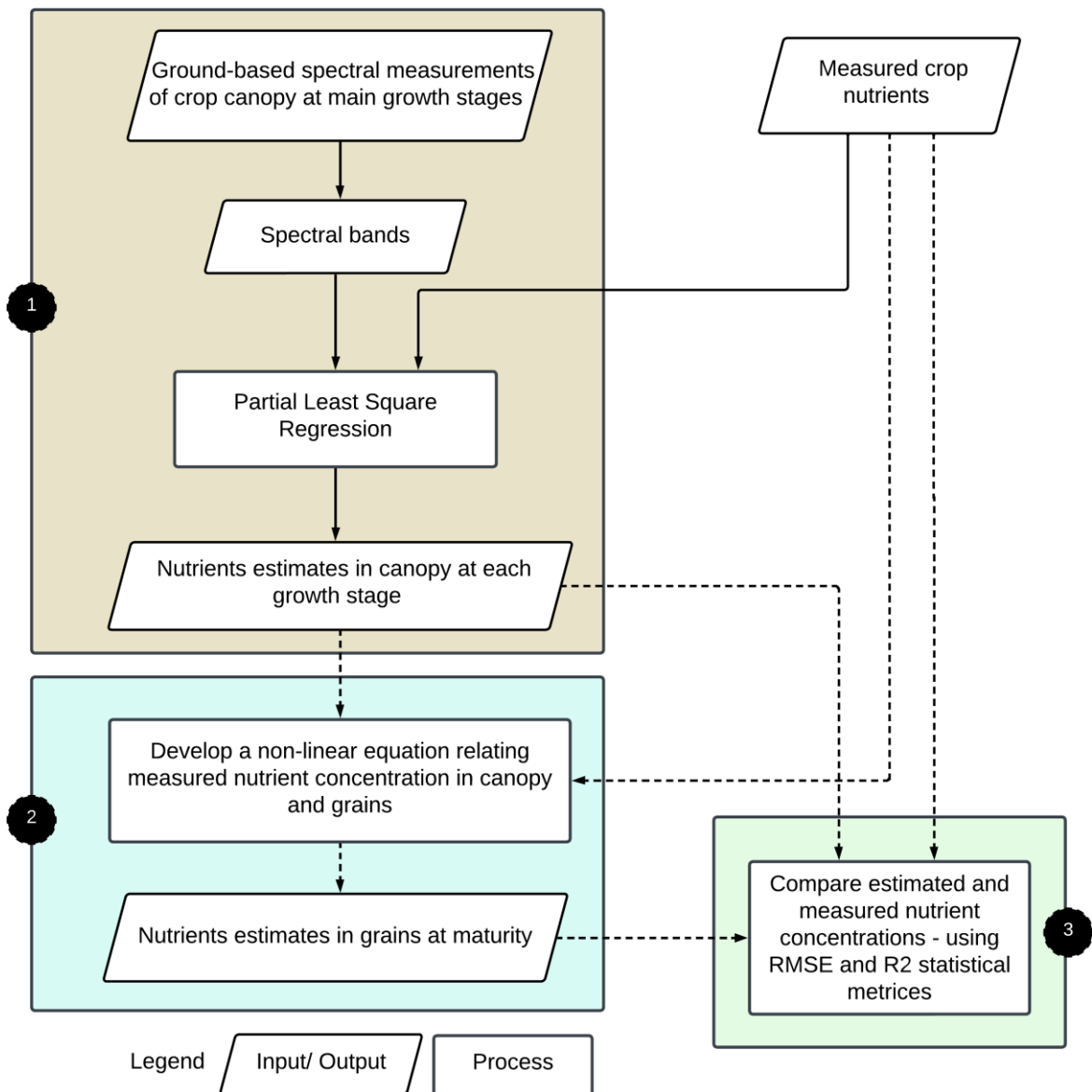
Figure 3: Flow chart describing the methodology. The numbers 1, 2, 3 represent where each objective is being addressed.

The objective of PLS was to predict response variable or Y based on explanatory variables or X, while describing the common structure inherent to both variables. As a regression technique, it facilitates the discovery of underlying factors i.e., linear composites of X that optimally model Y by following Equation (3.1) as detailed in section

$$Y = X\beta + \cdots \ldots + TP^T + \varepsilon \qquad (3.1)$$

where β is the slope of a standard linear regression model, ε is the error term, TP^T represents the transformation of X into a series of correlation coefficients or scores (T) and their corresponding loadings P^T. The first component explains the most variation in Y, followed by second, third etc.

The PLSR analysis focused on the loadings to understand the strength and direction of estimated nutrient concentrations among important regions of the spectrum, including visible, NIR, SWIR wavelength. The scores were used to estimate nutrient concentration in canopy at a given stage.

The PLSR was applied due to its robustness in handling collinear and high-dimensional spectral data (Ryan and Ali, 2016; Wold et al., 2001) . The analysis aimed to develop a predictive model that accurately estimates nutrients concentration based on the spectral characteristics of the canopy. In this context, spectra measurements constituted the explanatory variables X, while the estimated nutrient concentrations acted as the Y variable. The measured nutrients concentration in canopy at various growth stages served as the reference. The PLSR algorithm implemented in R environment using "caret" package was used for analysis.

To assess the relationship between ground spectral measurements and nutrient content in wheat crops, biplots of the first two PLSR components were used. Biplots display both the scores and the loadings simultaneously to facilitate interpretation of both the spectral data and nutrient concentrations. In the interpretation of the boxplots, the position (in terms of which quadrant) and direction of the loadings(in terms of right or left) indicate how each spectral variable contributes to the components, while the sample scores reveal the distribution and clustering of observations. This method facilitates understanding of the multivariate relationships, identifying key spectral bands associated with nutrient concentrations to facilitate assessing the ability of the PLSR models to discriminate between different nutrient concentrations in the wheat samples. Additionally, the Variable Importance in Projection (VIP) plots that shows a measure of a variable's importance in modelling Y using X were used to identify the most important bands for estimating various nutrients. In order to include most important variables in the models, a lower VIP threshold of 0.1 was applied as opposed to 1 that is commonly used in most studies (Doesburg et al., 2019; Fan et al., 2016).

3.5. Estimation of nutrient concentration in grains at maturity

The second objective focused on estimating nutrient concentrations in wheat grain at maturity across 40 plots. This required three key inputs for each nutrient: (1) laboratory-measured nutrient concentrations in the canopy during vegetative and reproductive stages (2) laboratory-measured nutrient concentration in grains at maturity; and (3) nutrient estimates in the canopy for these two growth stages, as were estimated using PLSR in the first objective.

To address this objective, the initial step involved establishing a statistical relationship between the nutrient concentrations measured in the grain at maturity and those measured in the canopy at the two growth stages. To test and establish the presence of linear or non-linear relationships (see example in equation (3.2) and (3.3)) between nutrients measured in the canopy and grains, the "lmtest' package in R was used.

$$Y = aX_1 + bX_2 + C \qquad (3.2)$$

$$Y = aX_1^2 + b\sqrt{X_2} + C \qquad (3.3)$$

where, Y is the nutrient concentration in the grain; X_1, and X are the nutrient concentrations at the vegetative and reproductive stages, respectively; C is a constant; and a and b are coefficients.

The coefficients obtained from this statistical relationship (a, and b in the examples) indicated how nutrient concentrations at each growth stage contributed to the final nutrient concentration in the grain. This analysis aimed to reveal which growth stages have most impact on the nutrient concentrations in the grain.

Finally, this established statistical relationship was used to predict nutrient concentrations in the grain. Here, the nutrient estimates obtained from the PLSR for the specified growth stages served as input variables.

3.6.   Comparison of estimated and laboratory-measured nutrient concentrations

To compare the estimated and laboratory-measured nutrient concentrations, statistical metrics including the Root Mean Square Error (RMSE) and the Coefficient of Determination (R^2) were used. RMSE provides a measure of the differences between the nutrient concentration values estimated by the PLSR (in case of ground-spectral measurements in objective 1) or established regression equation (in case of nutrients concentration for two growth stages in objective 2) and their corresponding values measured in the laboratory, to quantify the prediction error. Its formula is shown in equation (3.4).

$$RMSE = \sqrt{\frac{1}{n}\sum_{i-1}^{n}(y_i - \widehat{y_i})^2} \tag{3.4}$$

where n is the number of observations, y_i is the actual value of the dependent variable; and $\widehat{y_i}$ is the predicted value of the dependent variable.

The RMSE values range from -∞ to ∞. Lower RMSE values would indicate a closer match between obtained estimates and laboratory measurements, thus affirming the reliability (or predictive power) of using ground spectral measurements for nutrient estimations.

Relative RMSE that is calculated by dividing the RMSE by the mean of the actual values, then multiplying by 100 to express it as a percentage, to normalize the error in relation to the magnitude of the actual observed data was also used. This is due to high variability in RMSE values and also to allow easy comparison across different nutrients.

The R^2 value was used to measure the proportion of variance in the laboratory-measured nutrient concentrations that is predictable from the estimated nutrient concentrations. It ranges from 0 to 1 and its equation is shown in equation (3.5). An R^2 value closer to 1 suggests a high degree of correlation, supporting the study hypothesis of a high relationship between estimated and laboratory measured nutrient concentrations.

$$R^2 = 1 - \frac{\sum_{i-1}^{n}(y_i - \widehat{y_i})^2}{\sum_{i-1}^{n}(y_i - \bar{y})^2} \tag{3.5}$$

where $\bar{y}$ is the mean value of the actual values of the dependent variable.

These analyses were conducted for the different growth stages (when focusing on nutrient concentration in canopy) and at maturity when focusing on grains to identify specific patterns and variations in nutrient dynamics.

4. RESULTS AND DISCUSSION

4.1. Field data summary

4.1.1. Nutrient concentration in canopy and grains measured in the laboratory

The results on the nutrient composition at different growth stages (i.e., vegetative and reproductive) and for grains is shown in Figure 4.
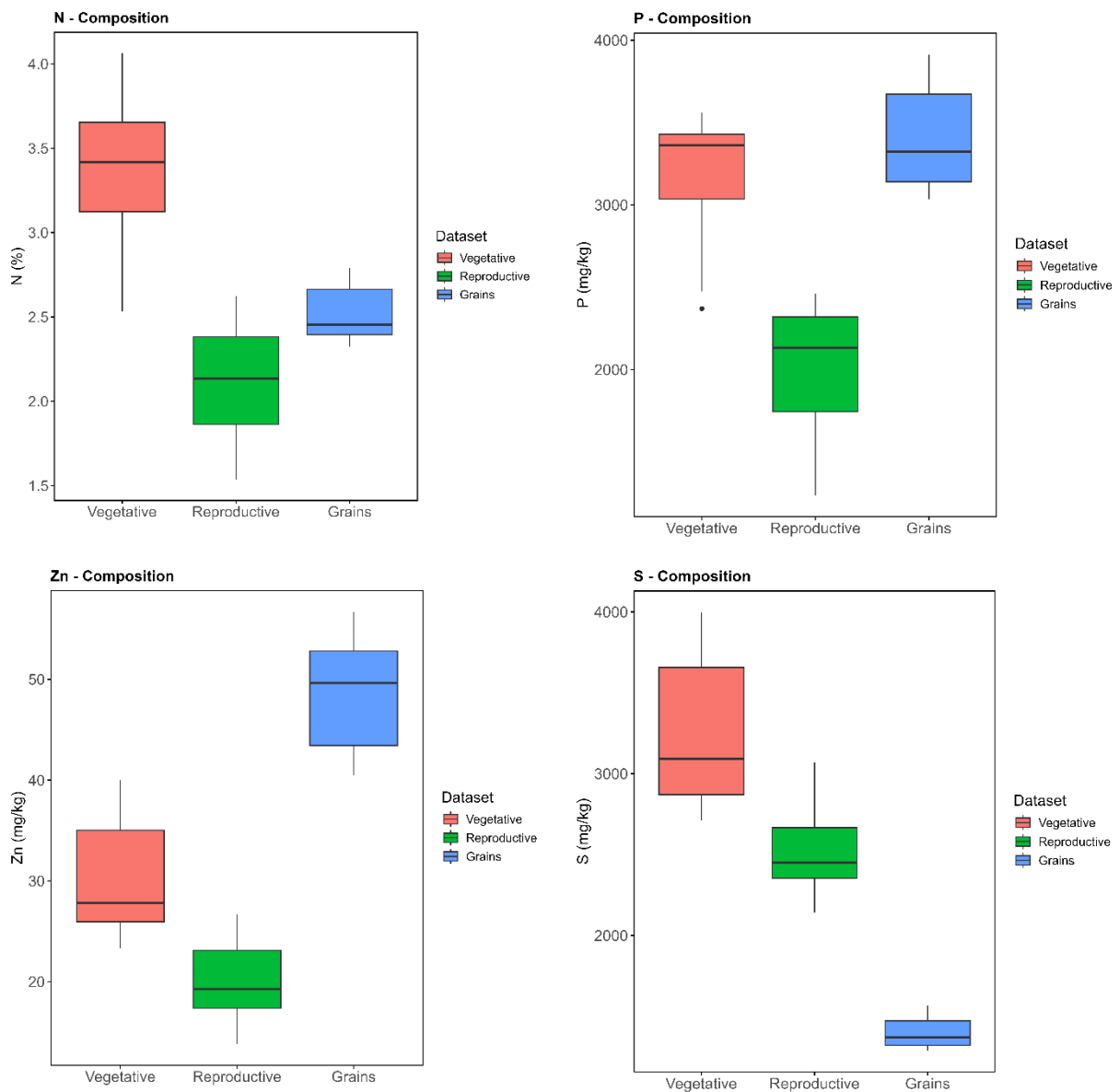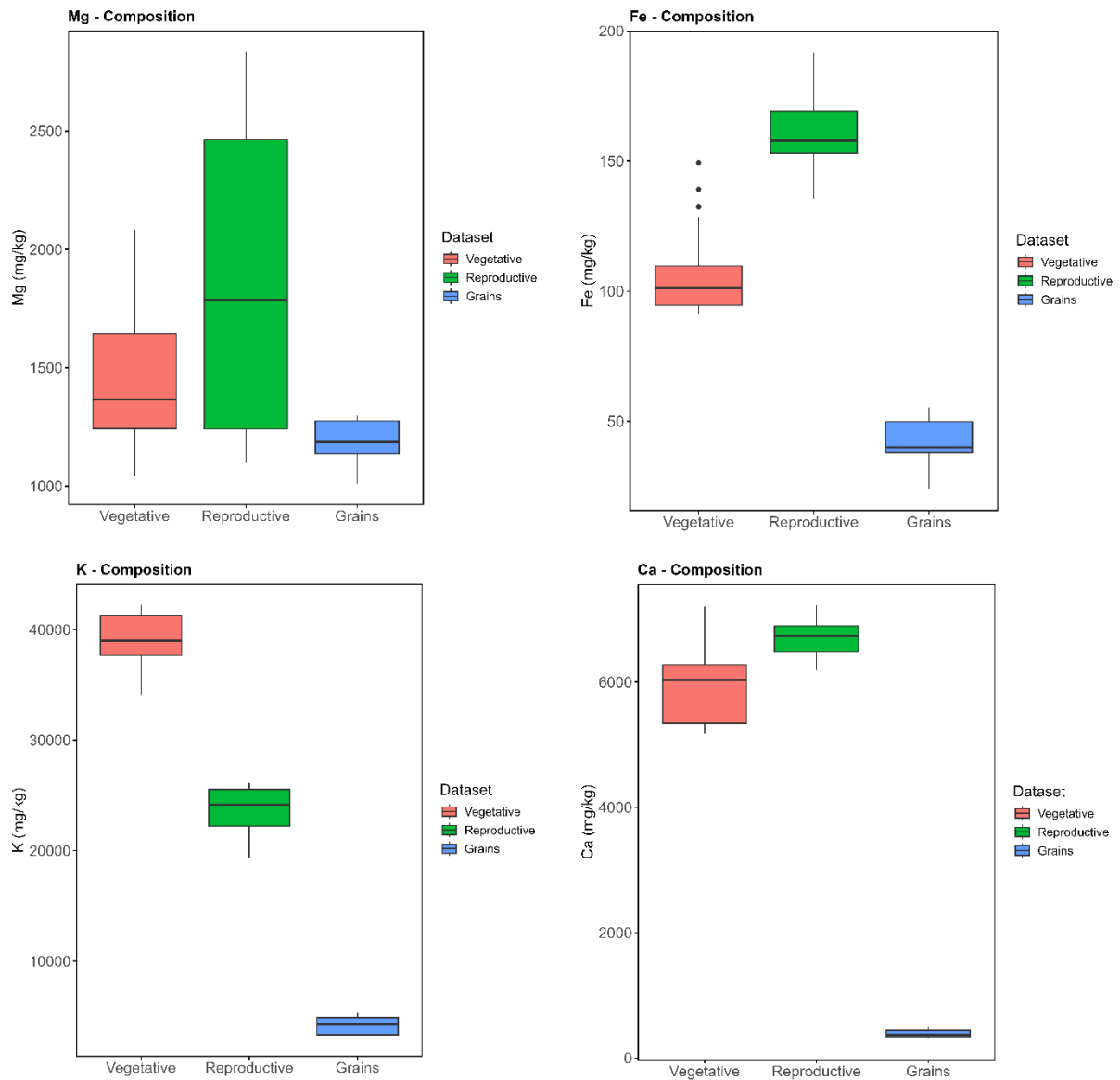
Figure 4: Box whisker plots representing the nutrient composition of wheat canopy (vegetative and reproductive stages) and grains. A single box-whisker represents a series of 10 ESU plots in the study area. The black horizontal line inside the box represents the median (50th percentile), the top and bottom boundary of the box represents the 25th and 75th percentiles, while the whiskers indicate the extreme values (5% and 95%) excluding outliers. The black dots show the outliers.

As shown in terms of median values of the boxplots, a comparative insights indicate that highest concentrations of most nutrients (with exception of Mg, Fe and Ca) is shown during the vegetative growth stages. This suggest that a potential high nutrient uptake during this stage of wheat development. This nutrient concentration decreases at reproductive stages, but still higher than that in the grains in some cases. Further, the grains exhibits the lowest nutrient concentrations for most nutrients. This suggests that a lower concentration of nutrients is translocated into the grains. These patterns highlight how nutrient uptake and accumulation vary across different growth stages, with initial stages (i.e., vegetative stage) generally having higher nutrient concentrations and a lower concentration of nutrients being mobilized or remobilized into the grains. In some instances, P and Zn concentrations in grains were higher than in the

canopy during the early growth phase. This may be due to the mobilization of P and Zn from the canopy to the grains during the later growth phase.

In general, each stage shows distinct ranges and median values for the different nutrients, indicating how nutrient concentrations vary across different growth stages of the wheat crop. Outliers suggest variability in nutrient content, which could be due to several factors including soil composition, fertilization, environmental conditions and potential measurement errors.

4.1.2. Correlation between biophysical or biochemical properties with nutrients

In Figure 5, heatmaps are presented to illustrate the relationships between various biophysical or biochemical properties and nutrient concentrations in wheat crop during its development stages. The figures show how specific bio-physical and bio-chemical properties correlate with essential micro- and macro-nutrients. Positive correlations, denoted in shades of red and orange, suggest that as one property increases, the corresponding nutrient concentration also tends to rise. Conversely, negative correlations, highlighted in blue and purple, indicate an inverse relationship.
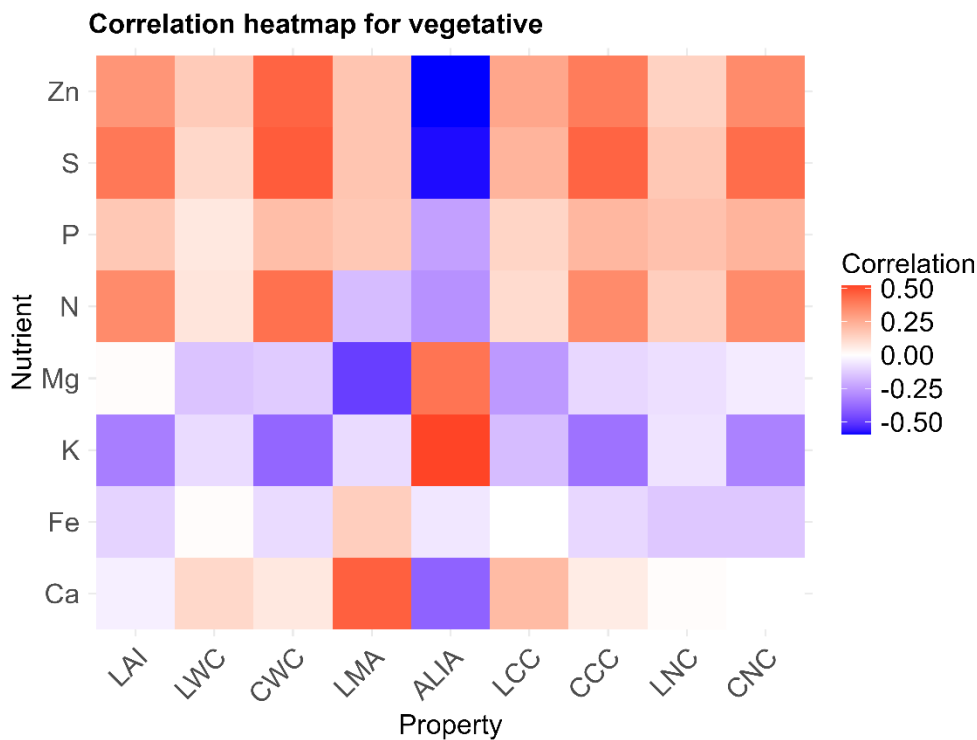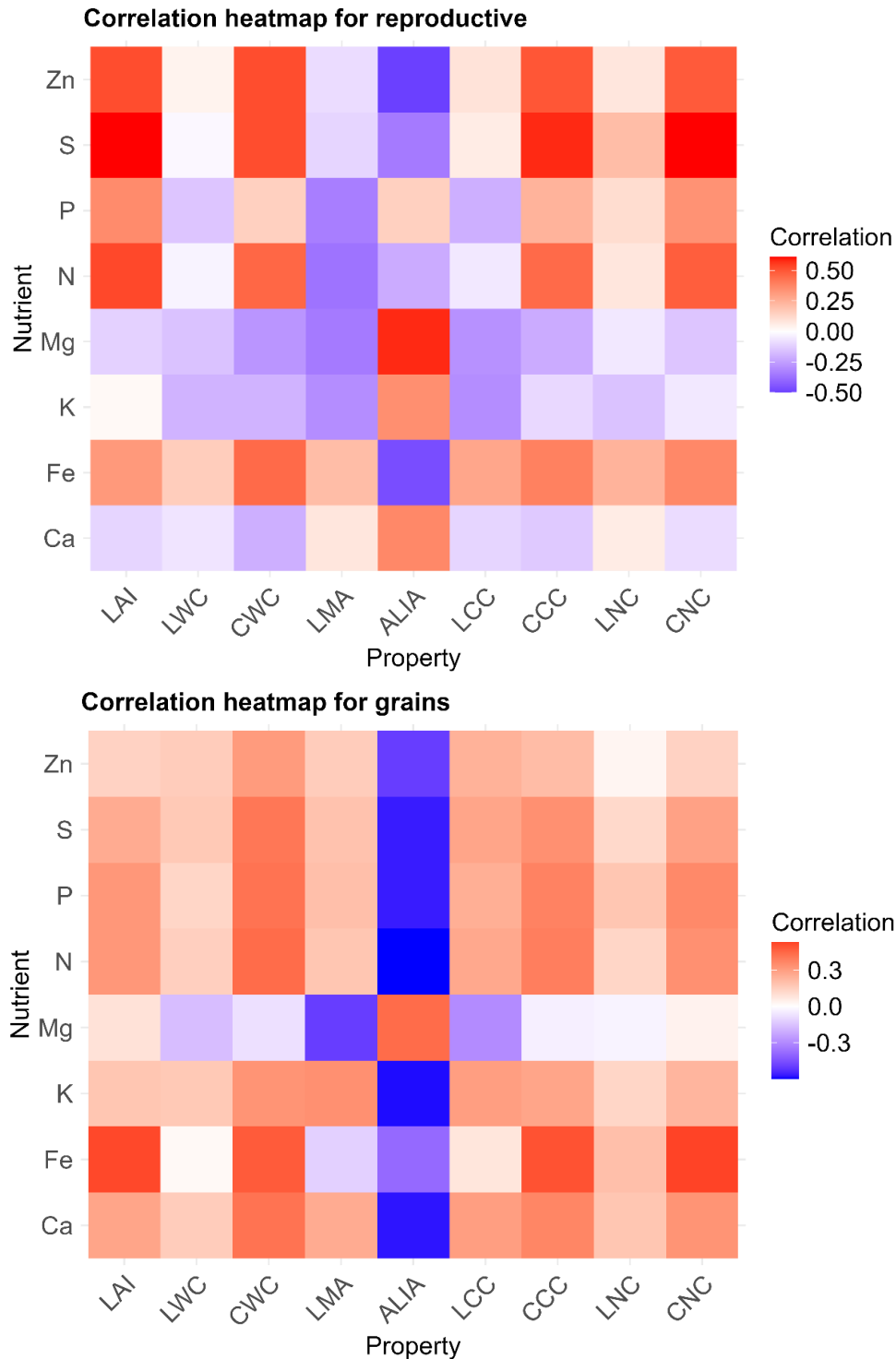
Figure 5: Correlation heatmaps illustrating the relationships between various biophysical or biochemical properties and nutrient concentrations in wheat crop in the canopy (vegetative stage at the top and reproductive stage at the middle) and grains at the bottom. The y-axis represents the nutrients and the x-axis represents the properties leaf area index (LAI), leaf water content (LWS), canopy water content (CWC), leaf mass area (LMA), average lea inclination angle (ALIA), leaf chlorophyll content (LCC), canopy chlorophyll content (CCC), leaf nitrogen content (LNC) and canopy nitrogen content (CNC). Positive correlations are depicted in shades of red and orange, indicating a direct relationship, while negative correlations are shown in shades of blue and purple, indicating an inverse relationship. The intensity of the colour corresponds to the strength of the correlation.

Results of the heatmaps reveal a consistent pattern of strong positive correlations between properties such as LAI, LWC, and CCC with nutrients like S, N, and K. This highlights their potential role in enhancing nutrient uptake and utilization in wheat crops. For example, the increased LAI suggests a larger photosynthetic area, which can enhance the plant's capacity to absorb and utilize nutrients more effectively, thus improving nutrient uptake and utilization in crops.

The negative correlations observed in the heatmaps, particularly between properties like LMA and nutrients such as P, Mg, and Fe, suggest potential trade-offs in plant physiology. These trade-offs may reflect the wheat crop's adaptive strategies to optimize growth under varying environmental conditions. For instance, the strong negative correlation between LMA and Mg in the heatmaps implies that plants with higher leaf mass per area might exhibit reduced magnesium concentrations, possibly due to altered metabolic demands or nutrient allocation strategies as described by Wright et al. (2005). Similarly, the negative correlation between LMA and Fe highlights the intricate interactions between different nutrients and the need for a balanced approach in nutrient management to avoid deficiencies that could impair plant health and productivity. Ensuring optimal amount of one nutrient without negatively impacting others is essential for maintaining overall plant growth.

The integration of canopy nutrient data and leaf traits can offer an understanding of nutrient dynamics within crop canopies and grains. Various bio-physical and bio-chemical properties that affect nutrient uptake and allocation strategies, can potentially impact nutrient concentrations in different plant parts also noted by Bertheloot et al. (2008). For instance, higher LAI can enhance photosynthesis, thereby increasing the uptake and assimilation of these essential nutrients, which can improve plant growth and resilience. Such an understanding can aid in optimizing the use of fertilizers and water resources, contributing to improved nutritional quality and yields towards meeting the global demand for food more efficiently.

4.1.3. Results on ground spectral measurements pre-processing

The preprocessing of the ground spectral measurements effectively stabilized variance and reduced noise, as seen in the smoother spectral features in Figure 6. For brevity only a sample at one of the sample plots is shown but the pattern is replicated for other spectral measurements. Removal of atmospheric windows ensured the exclusion of water bands, resulting in clearer and more reliable spectral data for subsequent PLSR analysis. These steps improved the quality of the ground spectral measurements for subsequent spectral analysis by mitigating variance instability and atmospheric interference.
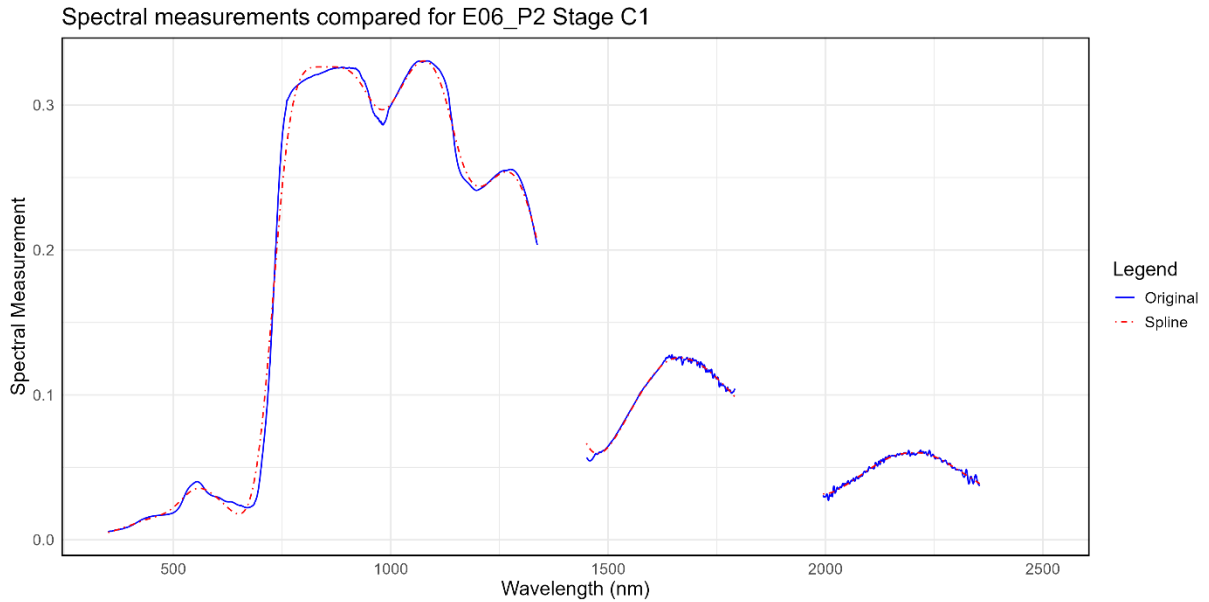
Figure 6: Pre-processed ground spectral measurements showing original (blue solid line) and smoothed spectra (red dashed line) after transformation and applying spline filtering and ESU09 and plot 1 for vegetative stage.

4.2. Estimation of nutrient concentration using Partial Least Squares Regression and crop canopy spectra measurements

In this section, the results of PLSR analysis carried out to estimate nutrient concentrations from crop canopy spectra are presented. The results are structured to present various PLSR outputs, its performance and key findings as demonstrated in the following sub-sections.

4.2.1. Performance of the PLSR model

The performance of the PLSR model for each of the nutrients across the crop canopy are summarised in Table 1. Results showed that the PLSR ground spectra models had variable performance across different nutrients and growth stages. S and Zn had the highest R^2 values, indicating the model explains 67% of their variances, making them the best-performing models among the nutrients for the vegetative growth stage. In contrast, Fe has the lowest R^2 value at 20%, indicating the limited ability of the model to explain its variance. For the other nutrients for this vegetative growth stage, the models explained 56, 20, 56, 49 and 54% of the variability in Ca, K, Mg, and N.

During the reproductive stage, Mg showed the highest R^2 of 0.87, indicating that the model explained a large proportion of variance for this nutrient. Ca had the lowest R^2 value, suggesting that the model may not be as effective in estimating this nutrient as it could only on average explain up to 20% variability in Ca concentration. In terms of performance across the growth stages, PLSR models generally accounted for most of the variabilities in nutrients during the reproductive growth stages.

Table 1: Summary of the PLSR decomposition of ground spectra measurements alongside model performance at a given wheat growth stage for each nutrient. Ncomp is the number of components for each model (maximum = 5). R^2 and RMSEP are unitless while the units for RMSE and MAE are same as those for response variable (nutrients concentrations i.e. mg kg^{-1} for all nutrients except N that is in %)

Stage	Nutrient	R^2	RMSE	RMSEP	MAE	Ncomp
Vegetative (C1)	Ca	0,56	382,06	400,08	305,86	5
	Fe	0,20	12,70	13,12	10,0	1
	K	0,56	1300,83	1329,66	1047,3	2
	Mg	0,49	199,00	205,01	156,1	3
	S	0,67	279,63	284,97	235,2	5
	P	0,21	275,55	283,14	226,3	2
	Zn	0,67	3,45	3,53	2,8	4
	N	0,54	0,28	0,29	0,2	3
Reproductive (C2)	Ca	0,20	278,17	280,96	231,8	3
	Fe	0,74	8,06	8,54	6,4	5
	K	0,61	1441,41	1484,64	1190,6	5
	Mg	0,87	257,29	262,72	220,0	5
	S	0,71	160,32	164,98	127,9	5
	P	0,62	258,21	261,91	212,6	5
	Zn	0,51	2,97	3,03	2,5	3
	N	0,42	0,25	0,25	0,2	5

The results of optimal number of components determined based on cross-validation (see Table 2) and detailed in Appendix B, showed that nutrients such as Fe and P used fewer number of components (1 and 2, respectively) as compared to Ca and S for the vegetative stage to suggest that the models were simple. The results show that the model was complex mostly for the nutrient concentrations during the reproductive stage (i.e., greatest number of components). These findings suggest that to avoid over-fitting while maintaining predictive power of the models, a tailored approach for each nutrient is needed during the vegetative stage.

Table 2: Number of component and their percentage explained variances

Stage	Nutrient	Ncomp	Comp.1	Comp.2	Comp.3	Comp.4	Comp.5
Vegetative (C1)	Ca	5	61,31	25,83	11,39	0,83	0,33
	Fe	1	56,96				
	K	2	59,85	35,25			
	Mg	3	61,89	16,68	19,95		
	S	5	56,96	37,57	3,98	0,85	0,27
	P	2	31,95	15,75			
	Zn	4	58,42	36,59	3,09	1,31	
	N	3	47,12	47,68	3,73		
Reproductive (C2)	Ca	3	80,23	11,21	7,05		
	Fe	5	68,4	27,05	3,05	0,58	0,48
	K	5	24,82	69,32	4,36	0,74	0,27

Mg	5	77,63	17,28	3,6	0,52	0,57
S	5	78,16	17,61	2,23	0,87	0,37
P	5	75,89	16,39	6,22	0,34	0,46
Zn	3	79,73	15,82	2,94		
N	5	80,1	15,56	0,95	2,26	0,75

Table 2 shows the results of variance explained by each of the components. For most nutrients, the first few components explained the majority of the variance. For instance in Ca during the vegetative growth stage, most variance is captured by the first two components (61.31% and 25.83%) though up to 5 components were used. Similarly, for Zn during the reproductive growth stage, most variance is captured by the first two components (80.1% and 15.56%) as demonstrated by scree plots in Figure 7 and Appendix A. This suggests that while more components can be added to marginally improve the model, the bulk of the explanatory power is concentrated in the first few components. It is noted that more number of components generally lead to better performance but can also risk overfitting.
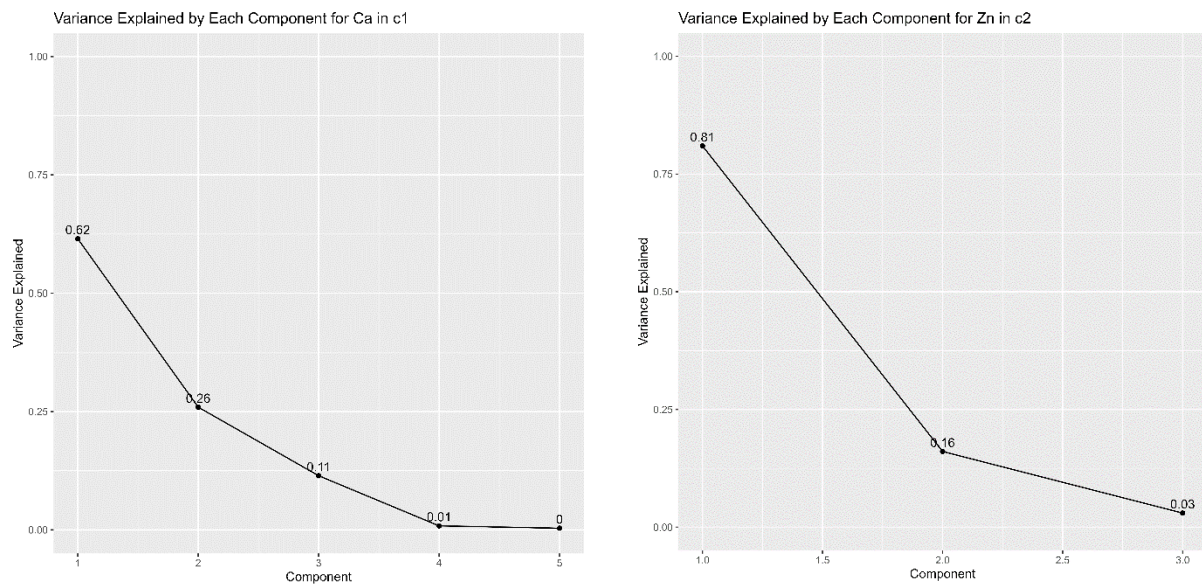
Figure 7: Scree plots of percentage variance explained by each component of Ca and Zn for vegetative and reproductive growth stages, respectively.

For predictive accuracy, the model errors (RMSE and MAE) were the lowest for N and Zn during both the vegetative and reproductive growth stages, indicating better performance of the PLSR model for estimating these nutrients. Conversely, the models for K demonstrated lower accuracy, indicated by relatively high model errors in both stages. This weaker model performance might be attributed to higher variability or measurement scale in K nutrient concentration. It is noted that the PLSR models only had access to limited number of samples for training and validation. Future studies to improve estimation of K by using ground spectral measurements could potentially be explored in the future through increasing number of samples and further data processing, outlier removal, or PLSR model simplification by reducing the number of components. However, these steps were not undertaken in the present study.

Overall, these findings suggest that the PLSR model was more effective in estimating and explaining variability in nutrients such as S and Zn during the vegetative stage, and showed better performance for magnesium Mg and Fe during the reproductive stage, a finding that is also consistent with study by Belgiu et al. (2023) who found models to be effective at early growth stages. Conversely, the PLSR model performed poorly in explaining the variances of nutrients like Fe and P in the vegetative phase, and exhibited weaker performance for Ca and N in the reproductive stage. Additionally, the model errors revealed that K had the highest error in both stages, indicating the need for further model refinement.

4.2.2. Identifying bands significantly correlated with nutrients concentration

4.2.2.1. Analysis based on correlation of first two components using biplots

The correlation between the first two components for the PLSR models at the vegetative and reproductive stages are shown in Figure 8 with further results included in Appendix D. In the vegetative stage, the biplots indicate strong correlations between certain ground spectra measurements and various nutrients. For N, wavelengths 1919, 1242, and 1464 show strong positive correlations with principal component 1, while 350 shows a negative correlation. Vectors 1019 and 1242 in the upper right quadrant are highly correlated and strongly influence both components 1 and 2. In contrast, wavelengths 1464, 796, 1687, and 1090 in the lower right quadrant are correlated and highly influence component 1, opposing vector 350 and the samples in the upper left quadrant. Vectors pointing outward indicate positive correlations with other elements, with clusters along these vectors representing samples with high N values.

Further results for the vegetative stage show that for Ca, wavelengths 2356, 573, 2133, and 1910 are positively correlated with principal component 1 (PC1), while wavelength 350 is negatively correlated. Vectors 2356, 573, and 2133 in the upper right quadrant are highly correlated and significantly influence both components 1 and 2, as indicated by their distance from the origin. Samples in this quadrant are closely related and show opposite tendencies to those in the lower left quadrant on component 2. In contrast,, wavelengths 1910, 1687, 796, 1464, 1242, and 1019 in the lower right quadrant are correlated and strongly influence component 1. They oppose the vector 350 and the samples in the upper left quadrant. Vectors pointing right indicate positive correlations with other elements in that direction, with clusters of points along these vectors representing samples with high Ca values. Similarly, for Fe, there is minimal correlation as the vectors are close to the origin. K shows a positive correlation with principal component 2 (PC2) and a negative one with PC1 for variables 2356, 573, and 2133, while 350 is positively correlated with PC1. Mg has wavelengths like 796, 1242, and 573 showing a positive correlation with PC2 and negative with PC1, with 350 indicating a strong positive correlation with PC1. S exhibits positive correlations with PC1 and PC2 for variables 1019, 1242, and 1464, and negative correlation for 350. P indicates strong negative correlations with PC1 and PC2 for variables 2356, 573, and 2133, with positive correlation for 350. Zn presents strong positive correlations with PC1 for 1019, 1242, and 1464, and negative for 350.
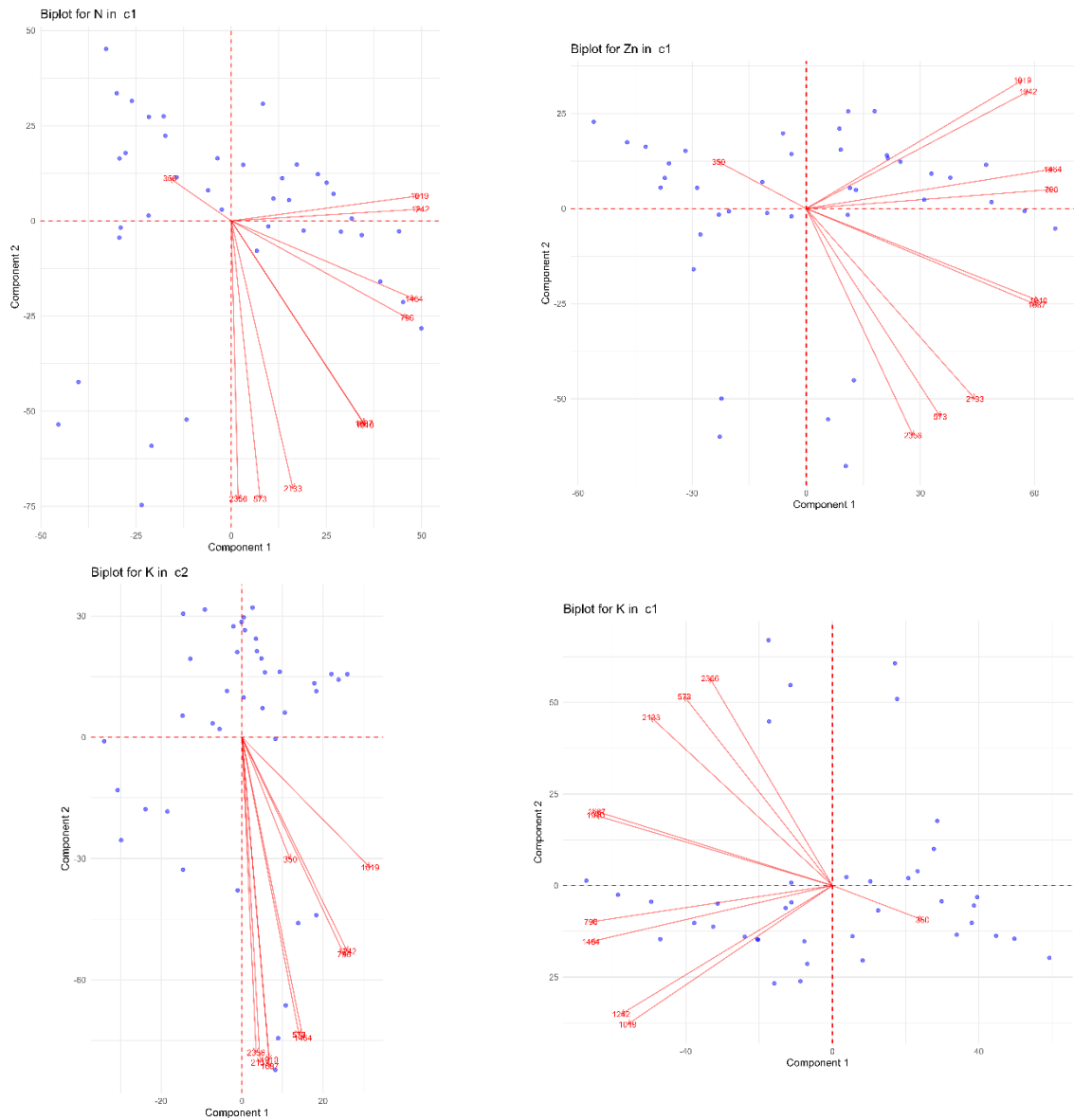
Figure 8: Biplots showing the correlation between the component 1 (PC1) and component 2 (PC2) for the models in vegetative and reproductive stages.

In the reproductive stage, Ca shows positive correlations with PC2 and negative with PC1 for variables 1019 and 796, with 573 also significantly contributing. For Fe, wavelengths 2133, 2356, and 1464 are positively correlated with PC1. K has strong positive correlations with PC1 and negative with PC2 for 1019 and 1242, while 2356, 1910, 2133, and 1687 significantly contribute to PC1, contrasting with the lesser contribution of 350. Mg in this stage shows positive correlations with PC2 and negative with PC1 for 1019, 573, and 1242.

The findings from the correlation between the first PC1 and PC2 highlights the most important ground spectra measurements that influence nutrient concentration in wheat crop, with varying patterns across vegetative and reproductive stages. Strong positive or negative correlations indicate which variables are key contributors, while vectors close to the origin suggest minimal

impact. As such an insight is obtained that can aid in identifying crucial spectral features for efficient nutrient monitoring and management.

4.2.2.2.   Analysis based on contribution of spectral regions using VIP plots

Results of revealing the most important spectral bands for estimating various nutrients are summarised using VIP plots, for example Figure 9 and further results included in Appendix E, which highlight that certain spectral bands are crucial for predicting the concentrations of specific nutrients. Analysis during the vegetative stage area that for Mg, the critical bands are 350 - 450 nm, 450 - 650 nm, and 2100 - 2400 nm. Estimation of K is mainly influenced by bands at 700 - 1350 nm, 1450 - 1800 nm, 2050 - 2400 nm, and 400 - 600 nm. Most contributions for Ca are from bands at 350 - 750 nm, 950 - 2400 nm, and 1150 - 1650 nm. The estimation of S is most influenced by bands at 2250 - 2400 nm, 1650 - 2050 nm, 700 - 950 nm, and 350 - 600 nm. N, Zn and Fe had lower VIP scores than the applied VIP threshold of 0.1 to indicate that none of the bands contributed to explaining the variance in these nutrients concentrations during the vegetative stage. This implies that alternative methods or additional data may be required to accurately determine the key spectral bands for predicting the content of Fe, N, and Zn.

The contributions of the spectral regions in this vegetative stage indicate that different nutrients are best estimated by specific spectral bands, reflecting their unique absorption and reflection characteristics. Mg is strongly associated with both ultra-violet rays (UV) – visible (VIS) and shortwave infrared 2 (SWIR2) regions, suggesting these wavelengths capture critical information about Mg-related processes in vegetation. The most important bands influencing K estimation span a broad range from VIS to SWIR2, highlighting its widespread interaction across different spectral regions. Most important bands for Ca in VIS and shortwave infrared (SWIR1) - SWIR2 suggest these regions are crucial for detecting its presence and concentration. S shows a strong influence from SWIR2, SWIR1-SWIR2, Red edge-SWIR1, and UV-VIS regions, indicating these bands are vital for capturing S-related spectral signatures. The detection of S is attributable to vibrational properties of proteins and amino-acids which are detectable in SWIR due to water absorption that is linked to Oxygen-hydrogen bonds.
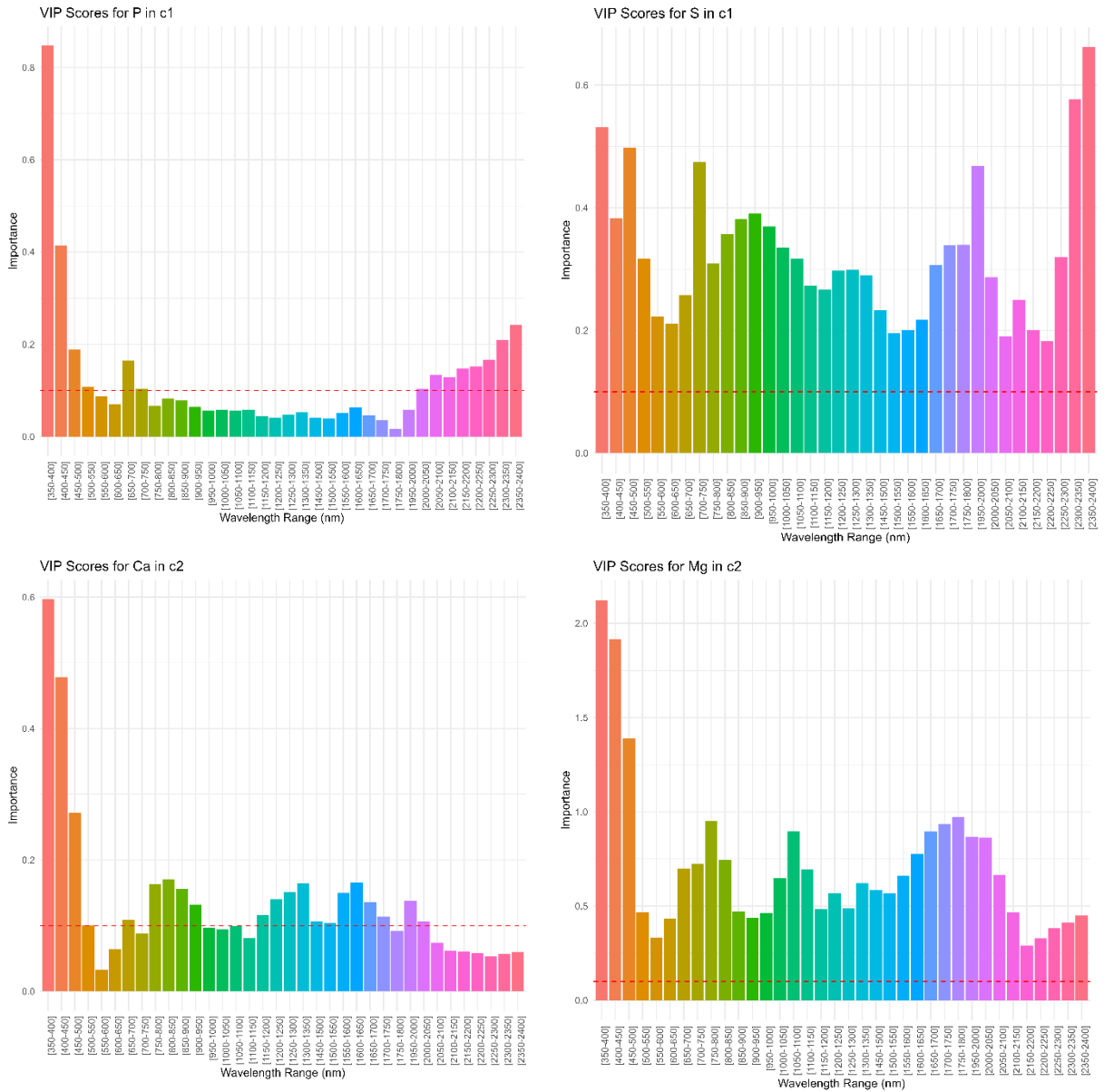
Figure 9: Variable Importance in Projection (VIP) plots for P and S nutrients during the vegetative stage (c1) and Ca and Mg during the reproductive stage.

For the reproductive stage, the analysis of spectral data identified the critical bands for estimating various nutrients as follows. For P, the important bands are 1000 - 1150 nm, 350 - 500 nm, 700 - 850 nm, and 1550 - 2150 nm. This suggest that most of the P's absorption and reflectance activities is across SWIR1, UV - VIS, Red edge, and SWIR2 regions, indicating that these wavelengths capture essential aspects of P's role in reproductive processes. Mg is mostly influenced by bands at 350 - 500 nm, 650 - 850 nm, 1000 - 1150 nm, and 1600 - 2150 nm. These influential bands for Mg spanning UV - VIS, Red edge - NIR, and SWIR1-SWIR2 highlights its extensive interaction across various spectral regions, which may be linked to its role in chlorophyll function and energy transfer. For Ca, the key bands are 350 - 500 nm and 750 - 950 nm which are within the UV - VIS and NIR regions, respectively, to suggest that these wavelengths are crucial for detecting its presence and concentration during the reproductive stage.

In general, during the vegetative stage, bands in the UV - VIS, VIS, NIR, SWIR1, and SWIR2 regions are crucial for nutrients like Mg, K and S while in the reproductive stage, the critical bands for P, Mg, Ca are found in the UV - VIS, Red edge, NIR, SWIR1, and SWIR2 regions. During photosynthesis, K is instrumental in regulating crop water potential when carbon is exchanged with atmosphere, which explains its detectability at the NIR or SWIR2 regions associated with water absorption features. Fe and Mg that are classified as metallic nutrients bond electrostatically to carbon compounds e.g., cellulose and lignin as highlighted in Belgiu et al. (2023), which results in absorption at longer wavelengths in the SWIR. Further, these nutrients are involved in chlorophyl biosynthesis to cause strong absorption characteristic due to Chlorophyll. Accessory pigments like carotenoids and anthocyanins are present in the visible spectrum to affect the nutrients detections at the early growth stages. UV is also useful for detecting nutrients associated with proteins and aromatic amino acids due to its influence on the N-H bonds. These findings support the hypothesis that red-edge bands during the reproductive stage are significant predictors of overall nutrient concentrations, due to their sensitivity to chlorophyll. An understanding of these critical bands and their contributions provides better insights into crop development and the nutrition quality during key stages and thus improved crop monitoring towards curbing nutrients deficiencies and food security.

4.2.3. Results of PLSR loadings

Results of the relative contributions of ground spectral features to nutrients estimation for wheat at vegetative and reproductive stages are shown in Figure 10 and Figure 11, respectively. In the context of this study, the PLSR loadings facilitated showing the strength and direction of the response among important regions of the spectrum (VIS – 400 to 680 nm, NIR – 730 to1000 nm, SWIR1 – 1000 to 1700 nm, and SWIR2 – 1700 to 2500 nm).

4.2.3.1. Important spectral bands for predicting nutrients

During the vegetative stage, the analysis of PLSR loadings identified important spectral regions for various nutrients in wheat canopy. For Fe, positive loadings spanned 379 nm to 2355 nm, with a peak at 2164 nm in the mid-infrared region, indicating a correlation between higher reflectance and higher Fe content, possibly due to water absorption features. Ca estimation showed positive loadings from 381 nm to 2355 nm, with a notable peak at 1523 nm in the NIR region, highlighting the importance of the plant water content and internal leaf structure. The broad range that includes VIS, NIR and SWIR regions suggest that Ca is influenced by multiple features. The positive loadings across these regions can also be attributed to various plant physiological processes. For example, in the visible range, reflectance is influenced by pigments like chlorophyll, which can indirectly reflect Ca concentrations. The NIR region is sensitive to the internal structure of the leaves and water content, which are crucial for understanding plant health and nutrient status. The SWIR region often relates to water absorption and other biochemical constituents of the plant (e.g., proteins, amino acids, cellulose lignin) as highlighted in Marshall et al. (2022). Negative loadings, indicating higher reflectance (or lower absorption) correlating with lower Ca content, were found from 350 nm to 380 nm, with the most pronounced negative loading at 350 nm at the UV-visible edge. This may reflect stress conditions

or lower pigment concentrations associated with reduced Ca concentration. This could be due to various factors affecting the plant's ability to absorb or utilize Ca effectively within this region.

K estimation identified positive loadings from 350 nm to 391 nm, with a peak at 350 nm at the UV-visible edge, suggesting healthier, pigment-rich plants that reflects more light. Mg content correlated with positive loadings from 350 nm to 380 nm, with a peak at 350 nm, indicating healthier plant tissues, while negative loadings in the NIR region suggested less dense or stressed tissues. P estimation showed positive loadings from 350 nm to 2355 nm, with a peak at 839 nm, indicating healthy internal canopy structures that reflects more NIR light. Negative loadings were identified from 350 nm to 2355 nm, with a peak at 652 nm, suggesting higher reflectance in the visible range correlates with lower P content, reflecting potentially less chlorophyl and poor plant stress. S content was linked to positive loadings from 381 nm to 2355 nm, with a peak at 1323 nm in the NIR region while negative loadings from 350 nm to 380 nm were observed for S, with a peak at 350 nm, suggest higher reflectance in the UV-visible edge correlates with lower S content to indicate stress or lower pigment concentration. N estimation showed positive loadings from 381 nm to 2355 nm, with a peak at 971 nm in the NIR region, associated with healthier plant tissues. Similarly, negative loadings from 350 nm to 380 nm, with a peak at 650 nm were observed for N to indicate higher reflectance in the visible range correlates with lower N content, possibly due to reduced chlorophyll concentrations. Zn content was associated with positive loadings from 381 nm to 2355 nm, with a peak at 1337 nm in the NIR region, reflecting healthier plant tissues.

In the reproductive stage, Zn estimation showed positive loadings from 350 nm to 2355 nm, with a peak at 1681 nm in the NIR region, indicating healthier plant tissues that reflects more NIR light possibly due to presence of water content in the plant tissues. No significant negative loadings were identified to indicate that there are no specific wavelengths where higher reflectance is associated with lower Zn content, suggesting a generally positive relationship between reflectance and Zn content. P content showed positive loadings across 350 nm to 2355 nm, with a peak at 2264 nm in the SWIR region, highlighting the importance of P-related compound absorption features. S content was linked to positive loadings from 350 nm to 2355 nm, with a peak at 1535 nm in the NIR region. Mg showed no positive loadings but had negative loadings from 350 nm to 2355 nm, with a peak at 1553 nm in the NIR region, indicating less dense or stressed tissues. K content showed positive loadings from 350 nm to 2355 nm, with a peak at 809 nm in the NIR region, reflecting healthier plant structures, while negative loadings at 1994 nm in the SWIR region suggested a weak inverse relationship. Fe estimation showed positive loadings from 350 nm to 2355 nm, with a peak at 2046 nm in the SWIR region, indicating healthier plant tissues. Ca content showed no significant positive loadings but had negative loadings from 350 nm to 2355 nm, with a peak at 1648 nm in the NIR region, indicating less dense or stressed tissues. N estimation showed positive loadings from 350 nm to 2355 nm, with a peak at 1687 nm in the NIR region, reflecting healthier plant structures.
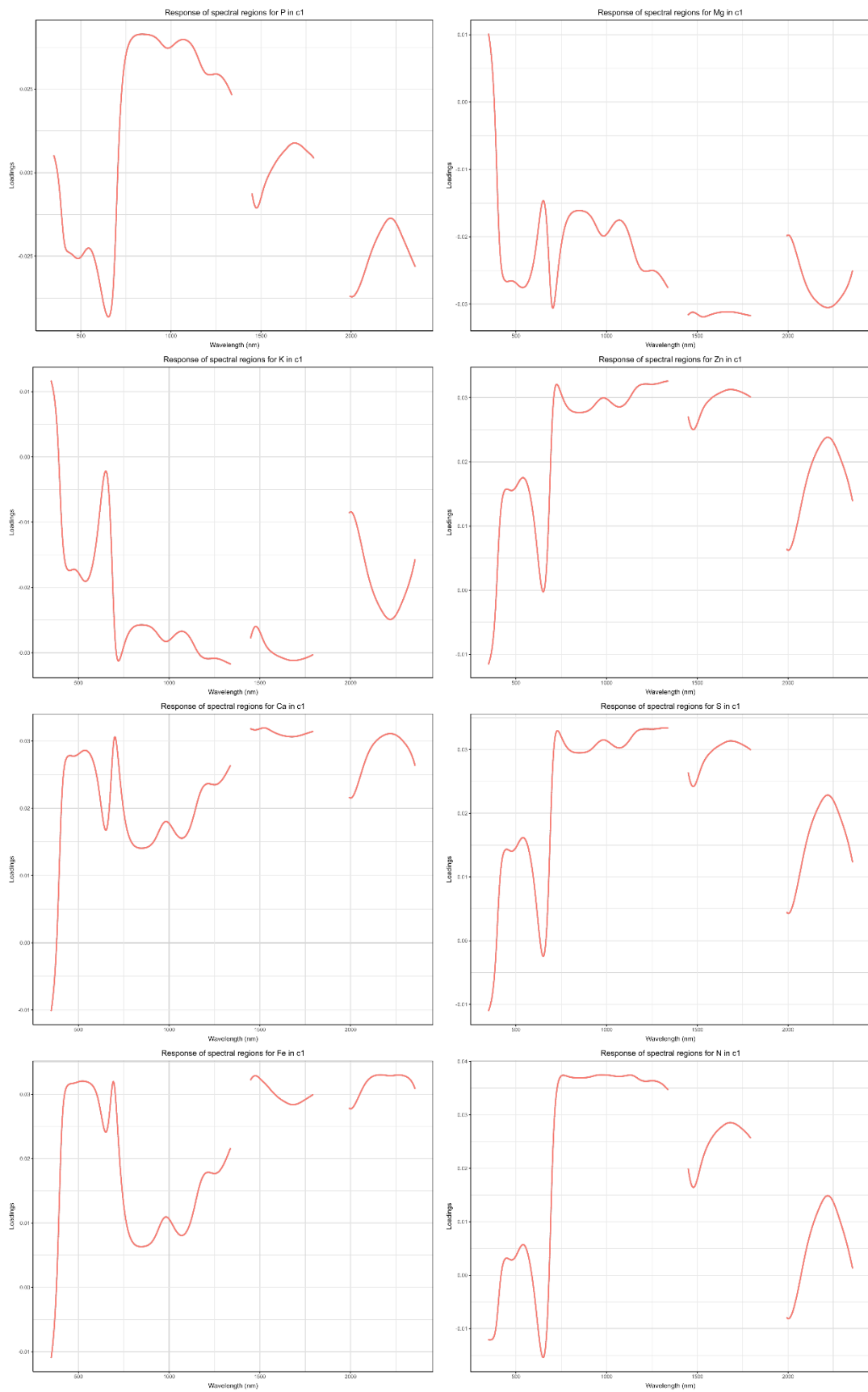
Figure 10: Response of spectral regions (as defined by PLSR loadings) to eight macro- and micro-nutrients in wheat canopy at vegetative growth stage.
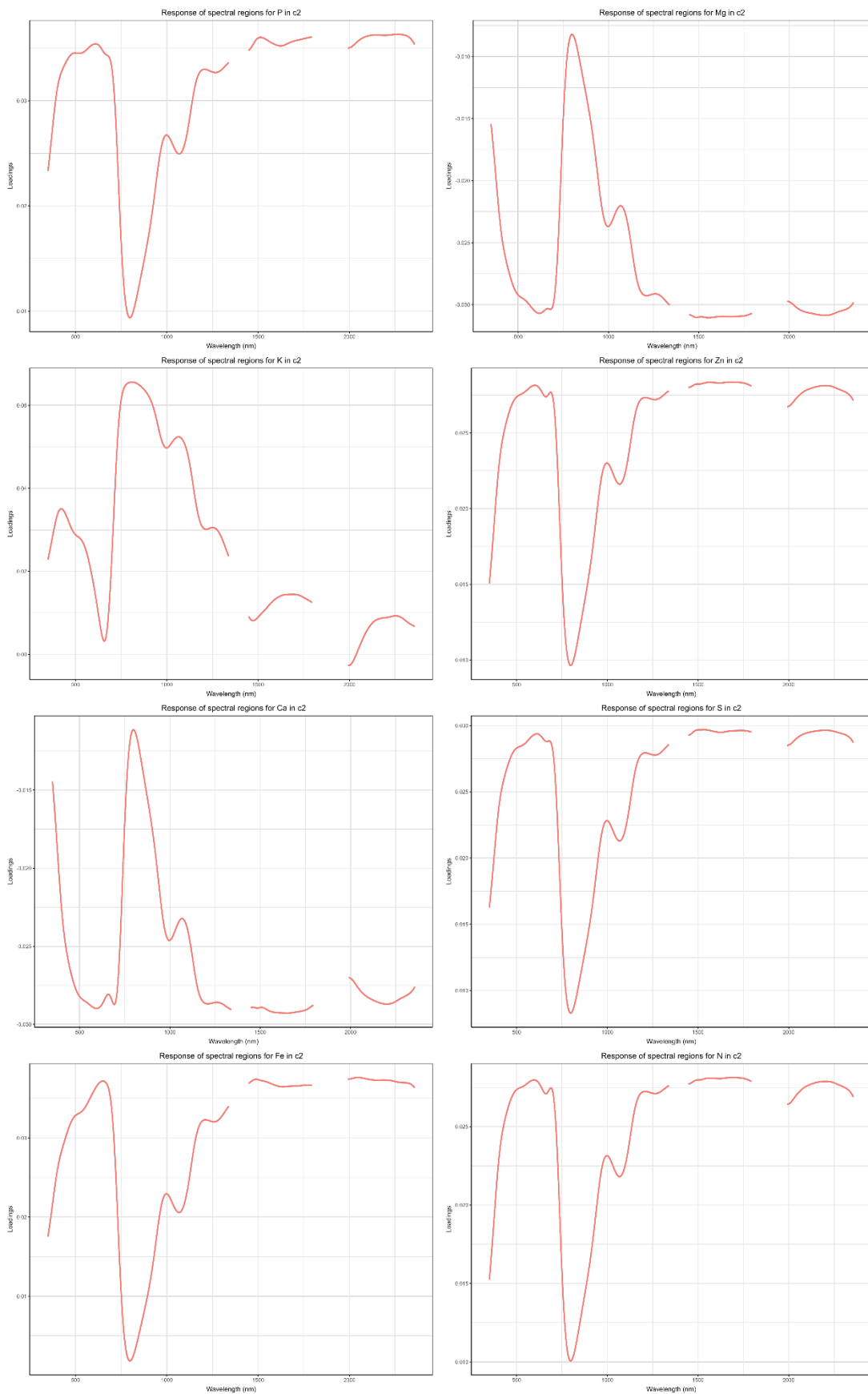
Figure 11: Response of spectral regions (as defined by PLSR loadings) to eight macro- and micro-nutrients in wheat canopy at reproductive growth stage.

4.2.3.2. Differences in spectral response between growth stages

During the vegetative stage, both positive and negative loadings influenced nutrient estimations across a broad range of wavelengths, including UV-visible, NIR, and SWIR regions. Positive loadings often indicated healthier, pigment-rich plants with robust internal structures, while negative loadings suggested stress conditions and reduced pigment concentrations. For example, Mg and K showed positive loadings at the UV-visible edge, indicating healthier plants, while negative loadings in the NIR region suggested stress or lower nutrient content. In contrast, the reproductive stage was characterized by a dominance of positive loadings, particularly in the NIR and SWIR regions, indicating healthier plant structures and a generally positive relationship between reflectance and nutrient content. Negative loadings were less pronounced, highlighting the stability of nutrient content during this stage. For instance, Fe and Zn showed strong positive loadings in the NIR and SWIR regions, reflecting healthier plant tissues with sufficient nutrient levels.

The differences in spectral response between growth stages can be attributed to plant physiology and nutrient dynamics. During the vegetative stage, plants focus on growth and development, involving a complex interplay of pigments, water content, and structural integrity. This results in a wider range of significant spectral responses as the plant undergoes active growth processes. The presence of both positive and negative loadings reflects the variability in nutrient content due to fluctuating environmental conditions and physiological changes. In the reproductive stage, the plant's focus shifts to seed production and maturation, leading to more consistent spectral responses dominated by positive loadings. This reflects the plant's overall health and nutrient status with less variability compared to the vegetative stage. The stability in nutrient content during this stage is crucial for successful reproduction and seed development, resulting in fewer significant negative loadings.

Environmental factors such as changes in water availability, light intensity, and temperature might also have influence the spectral responses. During the vegetative stage, the plant's active growth and adaptation to these conditions cause variability in spectral reflectance. In the reproductive stage, the plant's physiological processes stabilize, resulting in more consistent and positive spectral responses.

In general, these spectral response analysis provides an understanding of the key wavelengths and regions that can help in developing targeted remote sensing tools for various nutrients estimations and monitoring. Farmers can use such the developed remote sensing tool tuned to these specific wavelengths to detect and manage various nutrients deficiencies more effectively. Furthermore, future models for nutrient estimation can be improved by focusing on these identified wavelengths to enhance their accuracies and reliability.

4.2.4. Comparison of estimated and observed nutrients concentration

Results comparing the estimates and observed nutrient concentrations in the canopy are summarised in Table 3. The metrics for nutrient estimation models reveal varying levels of performance across different nutrients. The R^2 values indicate that the models for Ca (0.62), K

(0.55), S (0.71), Zn (0.71), and N (0.52) have moderate to strong correlations between estimated and actual observed nutrient concentration, explaining a substantial portion of the variance during the vegetative stage. However, the models for Fe (0.17), Mg (0.48), and P (0.18) exhibit weaker correlations, suggesting less reliable predictions for these nutrients. The metrics for nutrient estimation models during the reproductive stage reveal notable variations in model performance as well. High R^2 values for Fe (0.81), K (0.73), Mg (0.9), S (0.8), and P (0.7) suggest strong correlations between estimated and actual nutrient values, indicating these models effectively capture the variance in nutrient content. However, the lower R^2 values for Ca (0.2), Zn (0.53), and N (0.53) suggest weaker models for these nutrients, indicating less reliable predictions.

Evaluating the RMSE values for vegetative stage, it is found that the models for Zn (2.98) and N (0.26) exhibited very low error magnitudes, indicating accurate prediction capabilities. In contrast, nutrients like K (1265.2) and Ca (314.68) have higher RMSE values, pointing to larger model prediction errors. This differences highlights the varying degrees of accuracies across the PLSR models, with some nutrients requiring further refinement to improve nutrient estimations based on ground spectra measurements. RMSE values during the reproductive stage provide further insights into the models prediction accuracy. Nutrients like Fe (6.32), Zn (2.67), and N (0.2) exhibit low RMSE values, indicating high precision in predictions. In contrast, higher RMSE values for K (1112.97) and Ca (261.44) indicate larger prediction errors, highlighting areas where model refinement is needed to improve accuracy.

The table also show RRMSE values that provide a relative measure of prediction accuracy. Lower RRMSE values during vegetative stage for Ca (0.05), K (0.03), S (0.07), P (0.08), and N (0.08) suggest that these models are relatively accurate compared to the actual values. On the other hand, higher RRMSE values for Fe (0.12), Mg (0.13), and Zn (0.1) indicate less accurate predictions, emphasizing the need for model improvement. Overall, while the models for certain nutrients perform well, others like Fe, Mg, and P need enhancement to achieve better prediction reliability and accuracy. Low RRMSE values during reproductive stage for Ca (0.04), Fe (0.04), K (0.05), S (0.05), and N (0.09) indicate relatively accurate predictions compared to the actual values. However, higher RRMSE values for Mg (0.11), P (0.11), and Zn (0.13) suggest these models are less accurate, emphasizing the need for further model improvements.

Table 3: Comparison of estimated and observed nutrients concentration for vegetative and reproductive stages in terms of R^2, RMSE and relative RMSE. R^2 is unitless; RMSE units are same as those for response variable (i.e. mg kg^{-1} for all nutrients except N that is in %) and RRMSE is in %.

Stage	Nutrient	R2	RMSE	RRMSE
Vegetative	Ca	0,62	314,68	0,05
	Fe	0,17	12,12	0,12
	K	0,55	1265,2	0,03
	Mg	0,48	186,27	0,13
	S	0,71	230,57	0,07
	P	0,18	260,78	0,08
	Zn	0,71	2,98	0,1
	N	0,52	0,26	0,08
Reproductive	Ca	0,2	261,44	0,04
	Fe	0,81	6,32	0,04
	K	0,73	1112,97	0,05
	Mg	0,9	207,85	0,11
	S	0,8	121,19	0,05
	P	0,7	211,89	0,11
	Zn	0,53	2,67	0,13
	N	0,53	0,2	0,09

During the vegetative stage, nutrients like Ca, K, S, and Zn exhibit moderate to strong R^2 values, indicating reliable models, while Fe, Mg, and P show weaker correlations. In contrast, the reproductive stage shows improved model performance for Fe, K, Mg, and S with high R^2 values, suggesting stronger models, while Ca and Zn remain less reliable. RMSE values across both stages highlight that nutrients like Zn and N still had low error magnitudes, indicating high precision. However, high RMSE values for K and Ca persist in both stages, indicating consistent areas for improvement. RRMSE values reinforce these observations, with generally lower values during the reproductive stage for key nutrients, indicating relatively more accurate predictions, though Mg and Zn still require enhancement. Overall, the reproductive stage models generally show improved accuracy and reliability compared to the vegetative stage, especially for Fe, K, Mg, and S. Scatterplots of estimated versus observed nutrients concentrations for both stages can further illustrate these differences in model performance as shown in Figure 12 and Figure 13 for the best performing models.
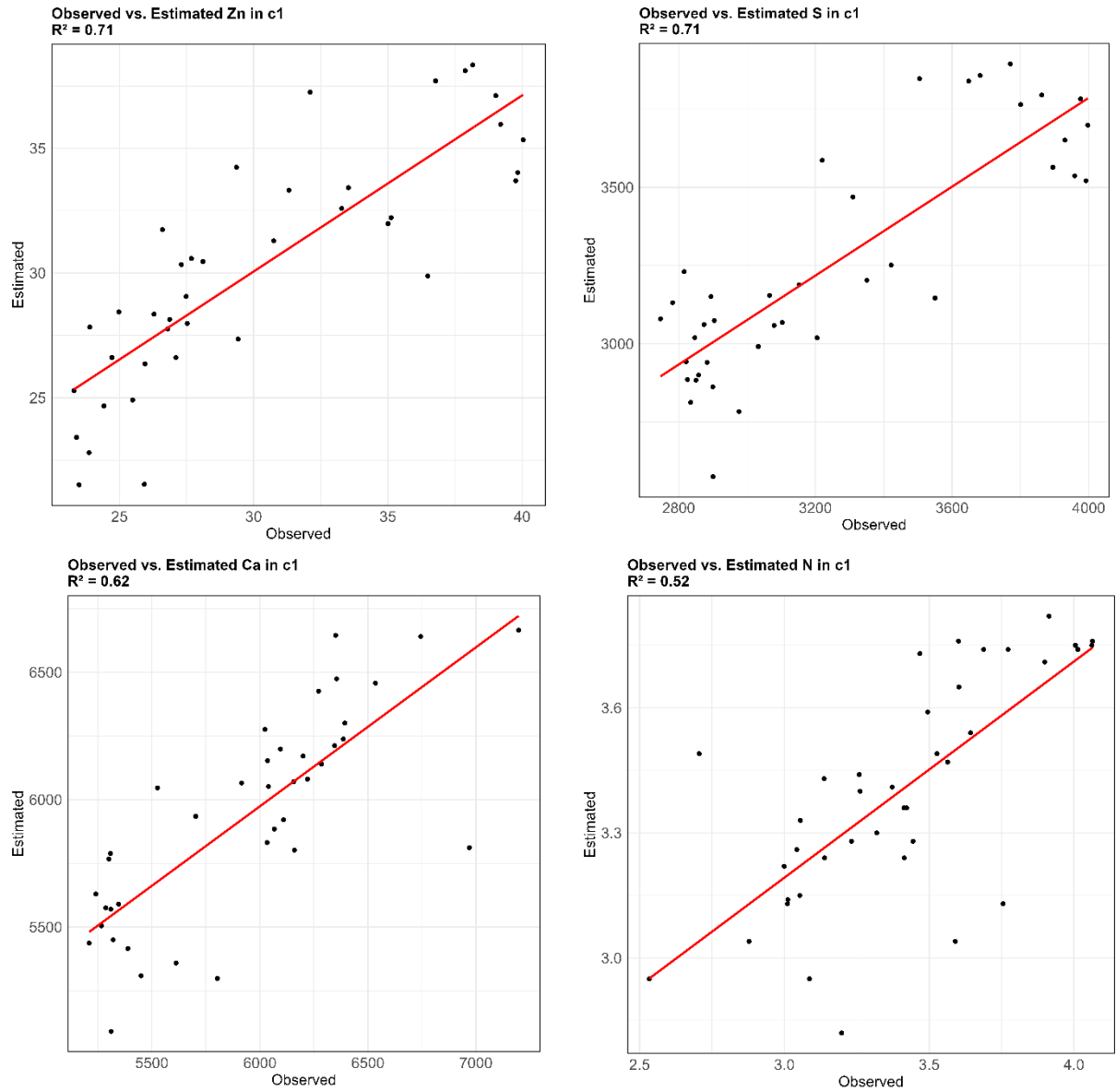
Figure 12: Scatterplots comparing the estimated and observed nutrients in wheat canopy at the vegetative stage.
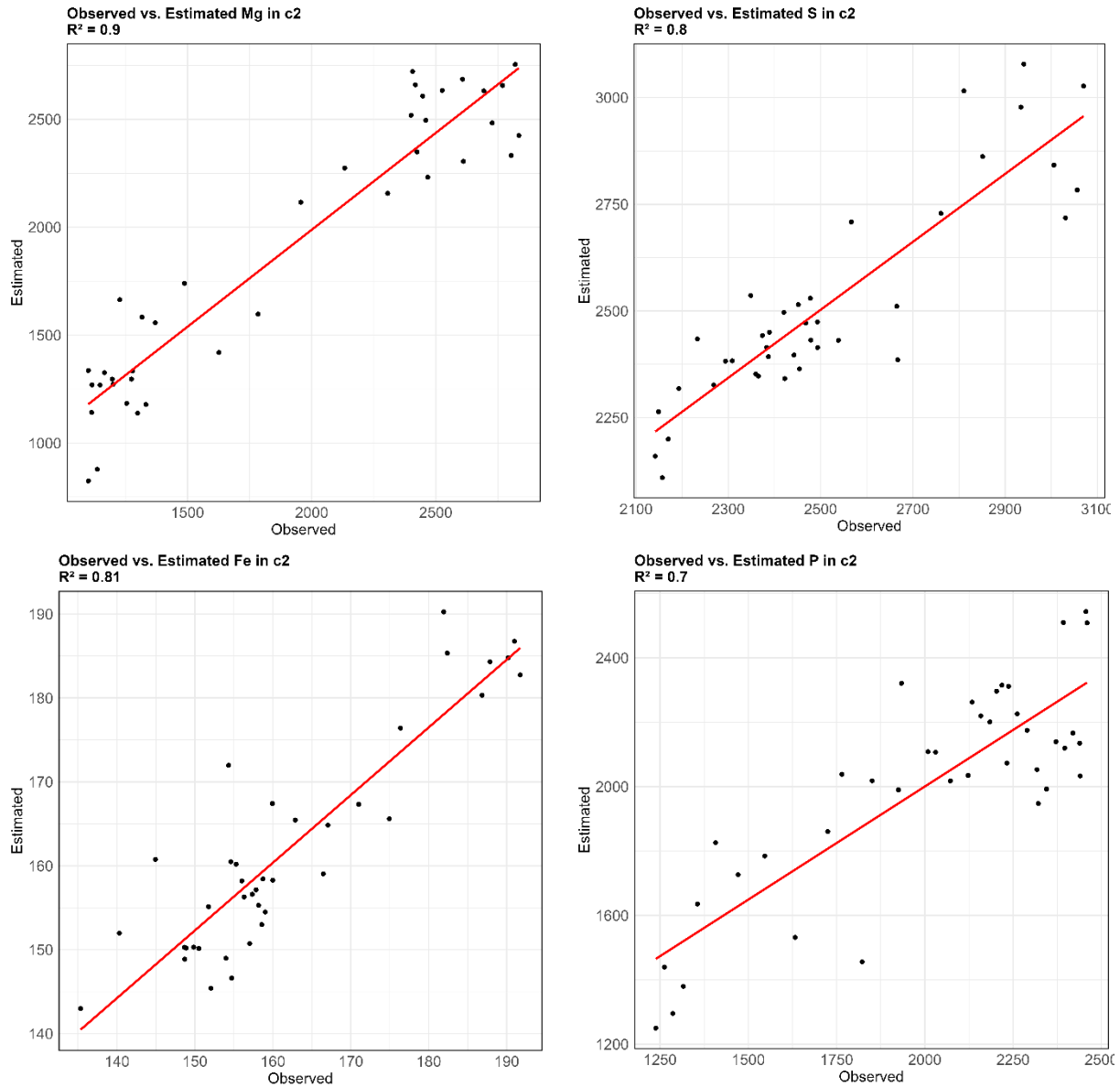
Figure 13: Scatterplots comparing the estimated and laboratory-measured (observed) nutrients in wheat canopy at the reproductive stage.

4.3. Estimation of nutrient concentration in grains using estimated nutrients across the canopy

The assessment of nutrient contributions from the vegetative and reproductive stages as presented in Figure 14 shows varied impacts on nutrient concentrations in wheat grains. Notably, Fe exhibits a strong positive influence during the reproductive stage (0.34) compared to a negative influence during the vegetative growth stage (-0.25), indicating enhanced accumulation of Fe due to reproductive growth processes. In contrast, nutrients like Ca, K, and Mg have lower or negative contributions during the reproductive stage, suggesting that these nutrients are less influenced by the reproductive stage and are instead more impacted during the vegetative stage.

For nutrients such as P, Zn, and N, the vegetative stage shows a higher positive contribution, with coefficients of 0.94, 0.92, and 0.31, respectively, compared to negative or minimal

contributions during the reproductive stage. This indicates that nutrient accumulation for these elements is more critical during the vegetative stage. These results on vegetative stage are important from a decision-making perspective, because farm managers have more time to intervene during the crop development in terms of improving crop nutritional quality.

Overall, while the hypothesis expect reproductive growth stage to show the highest impact on the nutrient concentrations in the grain due to the enhanced photosynthesis and grain filling processes during this stage, leading to increased nutrient accumulation in the grains, the results suggest that the vegetative growth stage plays a more crucial role in the accumulation of most nutrients, with the exception of Fe, where the reproductive stage is crucial.
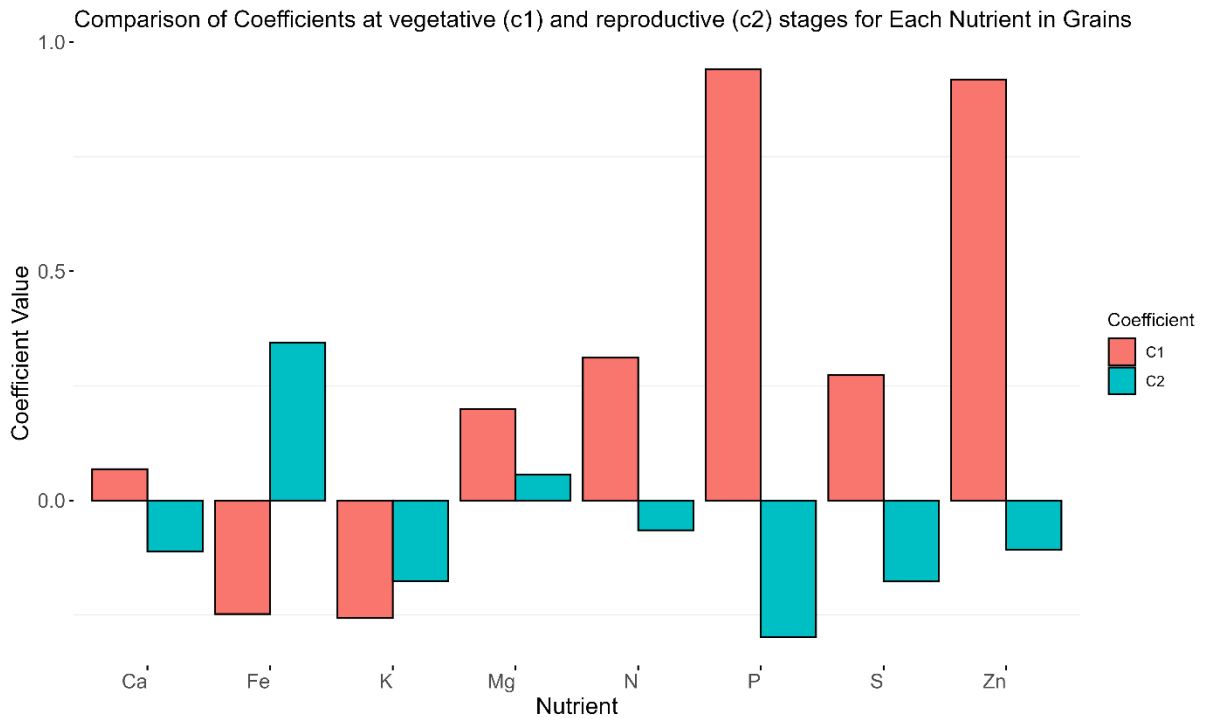

Figure 14: Composition of observed nutrients for vegetative and reproductive stages to that in the final grains.

The comparison of estimated and observed nutrient concentrations in wheat grains is provided in Table 4 and visually represented in Figure 15. High R^2 values for nutrients like K (0.70), Mg (0.75), and Fe (0.61) indicate that the non-linear models used to relate nutrients in canopy and that in the grains explain most portion of the variance, demonstrating strong predictive capabilities. However, the low R^2 values for P (0.20) and N (0.30) suggest that the non-linear models perform poorly, explaining little variance and indicating weak predictive power.

The RMSE values further illustrate the precision of these models in estimating nutrients in grains using that from the vegetative an reproductive stages. Nutrients such as Fe (5.57), Zn (3.45), and N (0.13) have low RMSE values, indicating high precision in their predictions. However, higher RMSE values for K (419.62) and P (255.20) suggest these models struggle with accuracy, resulting in larger average estimation errors. The RRMSE values align with these findings, where Mg (3.74) and S (4.48) exhibit relatively low RRMSE, indicating accurate estimations relative to actual

nutrient measurements, whereas Fe (13.11) and Ca (11.87) show high RRMSE, pointing to a higher deviations from actual nutrient concentrations.

Table 4: Comparison of estimated and laboratory-measured nutrients concentration in wheat grains in terms of R-squared, RMSE and relative RMSE. R2 is dimensionless, The units for RMSE are in mg kg^{-1} except for Nitrogen that are in %, and RRMSE is in %.

Nutrient	R^2	RMSE	RRMSE
Ca	0,48	46,21	11,87
Fe	0,61	5,57	13,11
K	0,70	419,62	9,98
Mg	0,75	44,48	3,74
S	0,55	62,58	4,48
P	0,20	255,20	7,46
Zn	0,59	3,45	7,12
N	0,30	0,13	5,25
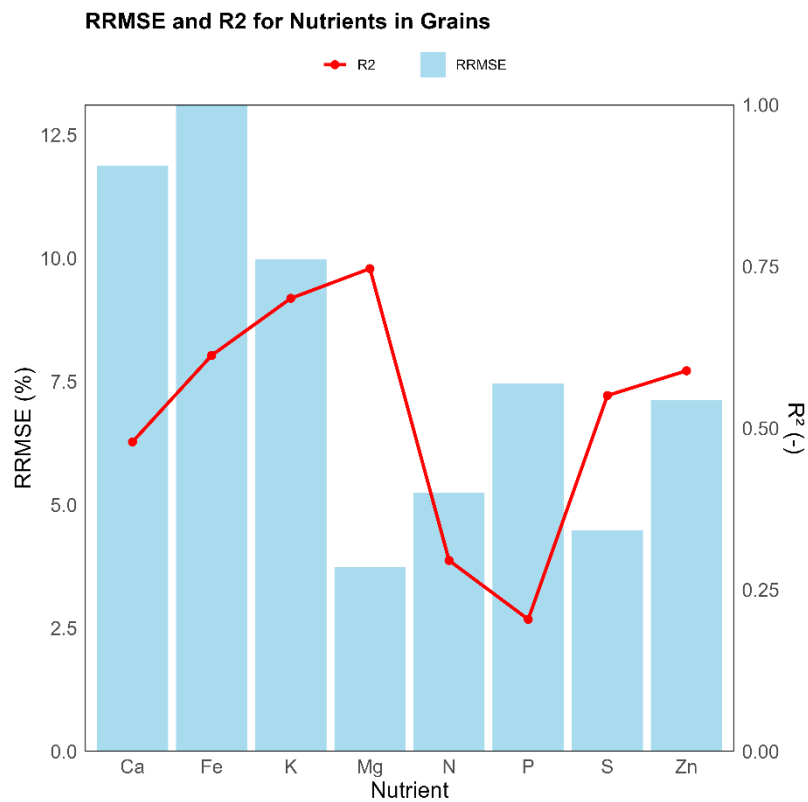
Figure 15: Performance metrics in terms of R² and relative RMSE for estimated nutrients in grains when compared to their laboratory-measured estimates.

In general, the models for K, Mg, and Fe show the best overall performance, with high explanatory power and reasonable error margins, making them reliable for estimating these nutrients concentration in wheat grains using canopy-based nutrients estimates from PLSR models. In contrast, the models for P and N require substantial improvement due to their low

predictive power. The models for Ca, S, and Zn demonstrate moderate effectiveness, suggesting their potential use but also indicating room for enhancement. Although the model show high predictive power for K, it however exhibits high model error that needs further refinement. Scatterplots of predicted versus actual nutrient concentrations in Figure 16 further illustrate these results, highlighting areas of underestimation or overestimation to support the needs for specific model refinement for improved accuracy.
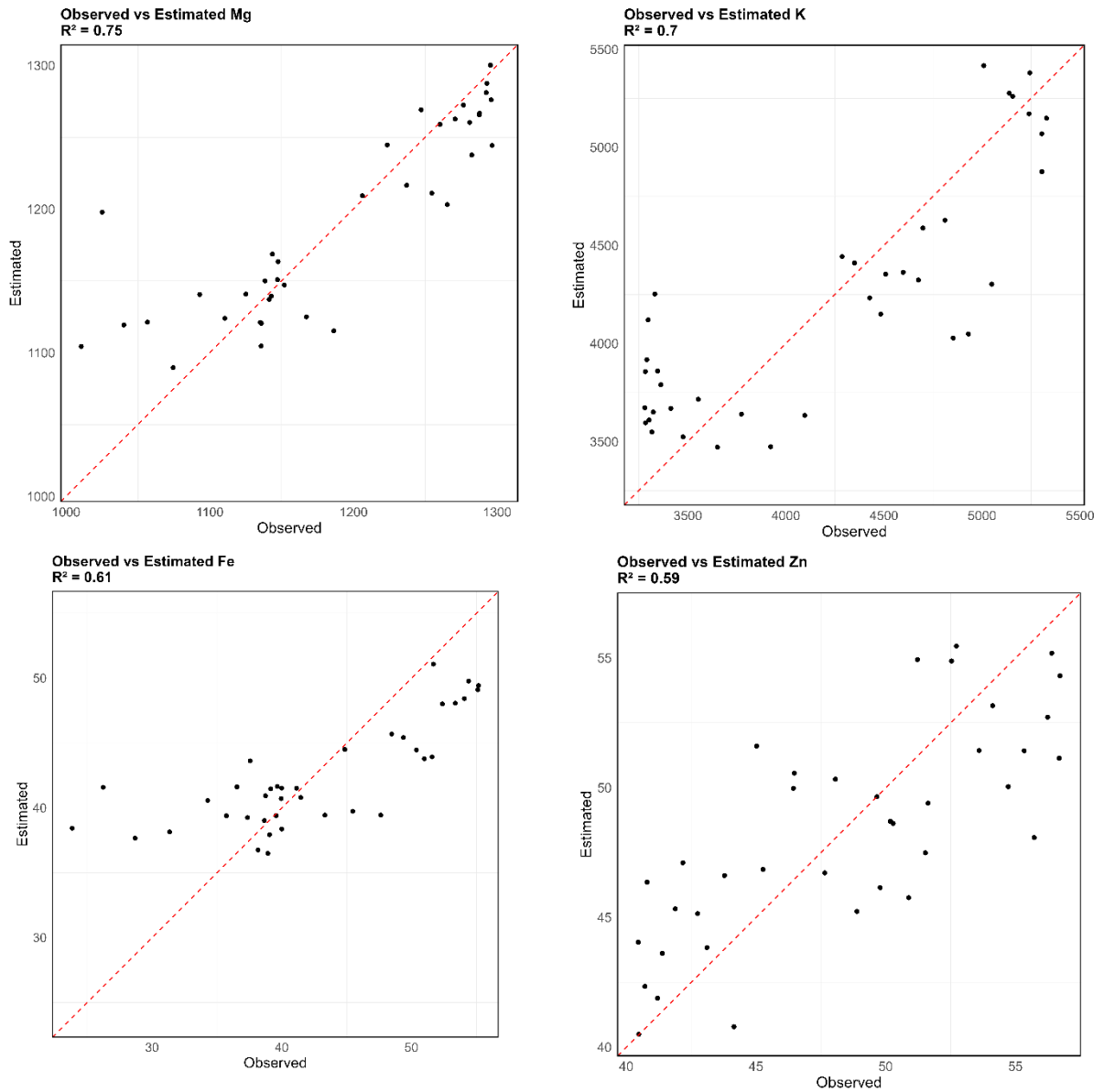

Figure 16: Scatterplots comparing the estimated and laboratory-measured (observed) nutrients in wheat grains for best performing models.

5. CONCLUSION AND RECOMMENDATION

5.1. Conclusion

This study assessed the potential of using ground spectral measurements to estimate nutrient concentrations in wheat canopies at different growth stages (vegetative and reproductive) and in grains at maturity. The macro- and micro-nutrients of focus were Calcium (Ca), Irom (Fe), Magnesium (Mg), Potasium (K), Phosphorus (P), Suplhur (S), and Zinc (Zn). The findings indicate that Partial Least Squares Regression (PLSR) models can moderately predict nutrient concentrations with varying degrees of accuracy across different nutrients. The key conclusions drawn from the study include:

a) **Nutrient concentration dynamics** – The nutrient concentrations in the canopy were generally higher during the vegetative stage compared to the reproductive stage, with the lowest concentrations observed in the grains. This finding highlights higher nutrient uptake during the early growth stages and subsequent translocation or reduction by the time of grain maturity.

b) **PLSR model performance** – The PLSR models exhibited varying performance, with R^2 values ranging from 0.2 to 0.67 for the vegetative stage and 0.2 to 0.87 for the reproductive stage. Nutrients such as Sulphur (S) and Magnesium (Mg) showed higher predictive accuracy, especially during the vegetative and reproductive stages, respectively.

c) **Spectral band importance –** Most important spectral bands for nutrient estimation were identified in the visible (VIS) and near-infrared (NIR) regions, with common peaks around 550 nm (green region) and 700-800 nm (NIR). These bands were crucial for predicting nutrient concentrations in wheat canopies.

d) **Nutrient estimation in grains –** The nutrient estimates in the grains using canopy nutrient concentrations estimated from PLSR models showed moderate predictive capabilities, with R^2 values between 0.2 and 0.9. This indicates a potential for ground spectral measurements in predicting final grain nutrient content, but with some limitations.

5.2. Recommendation

Based on the findings, further research on improving the accuracy and applicability of nutrient estimation using ground spectral measurements is encouraged. Some of these recommendations include:

a) **Further model refinement and increasing number of sizes –** While the PLSR model predictions were promising, their accuracies can further be improved by increasing the number of samples available for model training and validation as opposed to limited

samples that were used. Also, additional data preprocessing steps, such as outlier removal and noise reduction, to improve model accuracy for nutrients with lower predictive performance, such as Potassium (K) is recommended.

b) **Extending the study to other crops, geo-graphical regions and sample variability – ** additional work is needed to validate the findings across different crops and environmental conditions to generalize the applicability of ground spectral measurements in estimating nutrient concentrations especially within the low income countries affected by micro-nutrients deficiencies. As such the robustness and reliability of the models will be ensured. Also, focus on key stages particularly the vegetative stage, where nutrient uptake is highest to inform targeted fertilization strategies to optimize nutrient quality in crop yields. Lastly, although the sample sizes used in the study were from different agricultural farms, further studies can explore data with higher variabilities in terms of soil properties and different ecosystems.

c) **Exploration of additional models – ** in the current study, only PLSR was used. A further work to compare the models predictive power when using ground spectral measurements for nutrients estimations e.g., by using other machine learning models, such as Random Forest (RF) and Support Vector Machine (SVM), is recommended to potentially improve nutrient estimation accuracy and reduce model errors.

LIST OF REFERENCES

Adão, T., Hruška, J., Pádua, L., Bessa, J., Peres, E., Morais, R., Sousa, J.J., 2017. Hyperspectral Imaging: A Review on UAV-Based Sensors, Data Processing and Applications for Agriculture and Forestry. Remote Sensing 2017, Vol. 9, Page 1110 9, 1110. https://doi.org/10.3390/RS9111110

Adhikary, S., Biswas, B., Naskar, M.K., Mukherjee, B., Singh, A.P., Atta, K., Adhikary, S., Biswas, B., Naskar, M.K., Mukherjee, B., Singh, A.P., Atta, K., 2022. Remote Sensing for Agricultural Applications. Arid Environment - Perspectives, Challenges and Management. https://doi.org/10.5772/INTECHOPEN.106876

Ahmed, S., Griffin, T.S., Kraner, D., Schaffner, M.K., Sharma, D., Hazel, M., Leitch, A.R., Orians, C.M., Han, W., Stepp, J.R., Robbat, A., Matyas, C., Long, C., Xue, D., Houser, R.F., Cash, S.B., 2019. Environmental Factors Variably Impact Tea Secondary Metabolites in the Context of Climate Change. Front Plant Sci 10, 458137. https://doi.org/10.3389/FPLS.2019.00939/BIBTEX

Ali, S., Liu, Y., Ishaq, M., Shah, T., Abdullah, Ilyas, A., Din, I.U., 2017. Climate Change and Its Impact on the Yield of Major Food Crops: Evidence from Pakistan. Foods 2017, Vol. 6, Page 39 6, 39. https://doi.org/10.3390/FOODS6060039

Arnó, J., Rosell, J.R., Blanco, R., Ramos, M.C., Martínez-Casasnovas, J.A., 2012. Spatial variability in grape yield and quality influenced by soil and crop nutrition characteristics. Precis Agric 13, 393–410. https://doi.org/10.1007/S11119-011-9254-1/METRICS

Asseng, S., Martre, P., Maiorano, A., Rötter, R.P., O'Leary, G.J., Fitzgerald, G.J., Girousse, C., Motzo, R., Giunta, F., Babar, M.A., Reynolds, M.P., Kheir, A.M.S., Thorburn, P.J., Waha, K., Ruane, A.C., Aggarwal, P.K., Ahmed, M., Balkovič, J., Basso, B., Biernath, C., Bindi, M., Cammarano, D., Challinor, A.J., De Sanctis, G., Dumont, B., Eyshi Rezaei, E., Fereres, E., Ferrise, R., Garcia-Vila, M., Gayler, S., Gao, Y., Horan, H., Hoogenboom, G., Izaurralde, R.C., Jabloun, M., Jones, C.D., Kassie, B.T., Kersebaum, K.C., Klein, C., Koehler, A.K., Liu, B., Minoli, S., Montesino San Martin, M., Müller, C., Naresh Kumar, S., Nendel, C., Olesen, J.E., Palosuo, T., Porter, J.R., Priesack, E., Ripoche, D., Semenov, M.A., Stöckle, C., Stratonovitch, P., Streck, T., Supit, I., Tao, F., Van der Velde, M., Wallach, D., Wang, E., Webber, H., Wolf, J., Xiao, L., Zhang, Z., Zhao, Z., Zhu, Y., Ewert, F., 2019. Climate change impact and adaptation for wheat protein. Glob Chang Biol 25, 155–173. https://doi.org/10.1111/GCB.14481

Bagheri, N., Ahmadi, H., Alavipanah, S., Omid, M., 2012. Soil-line vegetation indices for corn nitrogen content prediction. Int Agrophys 26, 103–108. https://doi.org/10.2478/v10247-012-0016-8

Bahrami, H., Homayouni, S., Safari, A., Mirzaei, S., Mahdianpari, M., Reisi-Gahrouei, O., 2021. Deep Learning-Based Estimation of Crop Biophysical Parameters Using Multi-Source and Multi-Temporal Remote Sensing Observations. Agronomy 2021, Vol. 11, Page 1363 11, 1363. https://doi.org/10.3390/AGRONOMY11071363

Belgiu, M., Marshall, M., Boschetti, M., Pepe, M., Stein, A., Nelson, A., 2023. PRISMA and Sentinel-2 spectral response to the nutrient composition of grains. Remote Sens Environ 292, 113567. https://doi.org/10.1016/J.RSE.2023.113567

Benami, E., Jin, Z., Carter, M.R., Ghosh, A., Hijmans, R.J., Hobbs, A., Kenduiywo, B., Lobell, D.B., 2021. Uniting remote sensing, crop modelling and economics for agricultural risk management. Nature Reviews Earth & Environment 2021 2:2 2, 140–159. https://doi.org/10.1038/s43017-020-00122-y

Benediktsson, J.A., Wu, Z., 2021. Distributed Computing for Remotely Sensed Data Processing. Proceedings of the IEEE 109, 1278–1281. https://doi.org/10.1109/JPROC.2021.3094335

Berger, K., Atzberger, C., Danner, M., D'Urso, G., Mauser, W., Vuolo, F., Hank, T., 2018a. Evaluation of the PROSAIL Model Capabilities for Future Hyperspectral Model Environments: A Review Study. Remote Sensing 2018, Vol. 10, Page 85 10, 85. https://doi.org/10.3390/RS10010085

Berger, K., Atzberger, C., Danner, M., D'Urso, G., Mauser, W., Vuolo, F., Hank, T., 2018b. Evaluation of the PROSAIL Model Capabilities for Future Hyperspectral Model Environments: A Review Study. Remote Sensing 2018, Vol. 10, Page 85 10, 85. https://doi.org/10.3390/RS10010085

Bertheloot, J., Martre, P., Andrieu, B., 2008. Dynamics of Light and Nitrogen Distribution during Grain Filling within Wheat Canopy. Plant Physiol 148, 1707–1720. https://doi.org/10.1104/pp.108.124156

Caporaso, N., Whitworth, M.B., Fisk, I.D., 2018. Protein content prediction in single wheat kernels using hyperspectral imaging. Food Chem 240, 32–42. https://doi.org/10.1016/j.foodchem.2017.07.048

Challinor, A.J., Simelton, E.S., Fraser, E.D.G., Hemming, D., Collins, M., 2010. Increased crop failure due to climate change: assessing adaptation options using models and socio-economic data for wheat in China. Environmental Research Letters 5, 034012. https://doi.org/10.1088/1748-9326/5/3/034012

Clevers, J.G.P.W., Gitelson, A.A., 2013. Remote estimation of crop and grass chlorophyll and nitrogen content using red-edge bands on Sentinel-2 and -3. International Journal of Applied Earth Observation and Geoinformation 23, 344–351. https://doi.org/10.1016/J.JAG.2012.10.008

Cockson, P., Landis, H., Smith, T., Hicks, K., Whipker, B.E., 2019. Characterization of Nutrient Disorders of Cannabis sativa. Applied Sciences 2019, Vol. 9, Page 4432 9, 4432. https://doi.org/10.3390/APP9204432

Cooper, M., Tang, T., Gho, C., Hart, T., Hammer, G., Messina, C., 2020. Integrating genetic gain and gap analysis to predict improvements in crop productivity. Crop Sci 60, 582–604. https://doi.org/10.1002/CSC2.20109

Crusiol, L.G.T., Nanni, M.R., Furlanetto, R.H., Sibaldelli, R.N.R., Cezar, E., Sun, L., Foloni, J.S.S., Mertz-Henning, L.M., Nepomuceno, A.L., Neumaier, N., Farias, J.R.B., 2021. Yield Prediction in Soybean Crop Grown under Different Levels of Water Availability Using Reflectance Spectroscopy and Partial Least Squares Regression. Remote Sensing 2021, Vol. 13, Page 977 13, 977. https://doi.org/10.3390/RS13050977

Cummings, C., Miao, Y., Paiao, G.D., Kang, S., Fernández, F.G., 2021. Corn Nitrogen Status Diagnosis with an Innovative Multi-Parameter Crop Circle Phenom Sensing System. Remote Sensing 2021, Vol. 13, Page 401 13, 401. https://doi.org/10.3390/RS13030401

Danson, F.M., Rowland, C.S., Baret, F., 2003. Training a neural network with a canopy reflectance model to estimate crop leaf area index. Int J Remote Sens 24, 4891–4905. https://doi.org/10.1080/0143116031000070319

Darvishzadeh, R., Skidmore, A., Schlerf, M., Atzberger, C., Corsi, F., Cho, M., 2008. LAI and chlorophyll estimation for a heterogeneous grassland using hyperspectral measurements. ISPRS Journal of Photogrammetry and Remote Sensing 63, 409–426. https://doi.org/10.1016/J.ISPRSJPRS.2008.01.001

de Oliveira, A.F., Serra, S., Ligios, V., Satta, D., Nieddu, G., 2021. Assessing the Effects of Vineyard Soil Management on Downy and Powdery Mildew Development. Horticulturae 2021, Vol. 7, Page 209 7, 209. https://doi.org/10.3390/HORTICULTURAE7080209

de Oliveira, L.F.R., Santana, R.C., 2020. Estimation of leaf nutrient concentration from hyperspectral reflectance in Eucalyptus using partial least squares regression. Sci Agric 77, e20180409. https://doi.org/10.1590/1678-992X-2018-0409

De Silva, A.L., Trueman, S.J., Kämper, W., Wallace, H.M., Nichols, J., Hosseini Bai, S., 2023. Hyperspectral Imaging of Adaxial and Abaxial Leaf Surfaces as a Predictor of Macadamia Crop Nutrition. Plants 12, 558. https://doi.org/10.3390/PLANTS12030558/S1

de souza, R., Peña-Fleitas, T.M., Thompson, R.B., Gallardo, M., Padilla, F.M., 2020. Assessing Performance of Vegetation Indices to Estimate Nitrogen Nutrition Index in Pepper. Remote Sensing 2020, Vol. 12, Page 763 12, 763. https://doi.org/10.3390/RS12050763

Di Giuseppe, D., Bianchini, G., Faccini, B., Coltorti, M., 2014. Combination of wavelength dispersive X-ray fluorescence analysis and multivariate statistic for alluvial soils classification: a case study from the Padanian Plain (Northern Italy). X-Ray Spectrometry 43, 165–174. https://doi.org/10.1002/XRS.2535

Din, M., Ming, J., Hussain, S., Ata-Ul-Karim, S.T., Rashid, M., Tahir, M.N., Hua, S., Wang, S., 2019. Estimation of dynamic canopy variables using hyperspectral derived vegetation indices under varying N rates at diverse phenological stages of rice. Front Plant Sci 9, 353371. https://doi.org/10.3389/FPLS.2018.01883/BIBTEX

Din, M., Zheng, W., Rashid, M., Wang, S., Shi, Z., 2017. Evaluating hyperspectral vegetation indices for leaf area index estimation of Oryza sativa L. at diverse phenological stages. Front Plant Sci 8, 237162. https://doi.org/10.3389/FPLS.2017.00820/BIBTEX

Doesburg, F., Smit, J.M., Paans, W., Onrust, M., Nijsten, M.W., Dieperink, W., 2019. Use of infrared thermography in the detection of superficial phlebitis in adult intensive care unit patients: A prospective single-center observational study. PLoS One 14, e0213754. https://doi.org/10.1371/journal.pone.0213754

Dong, H., Fan, S., Sun, H., Chen, C., Wang, A., Jiang, L., Ma, D., 2021. Rhizosphere-Associated Microbiomes of Rice (Oryza sativa L.) Under the Effect of Increased Nitrogen Fertilization. Front Microbiol 12, 730506. https://doi.org/10.3389/FMICB.2021.730506/BIBTEX

dos Santos, C.A.T., Lopo, M., Páscoa, R.N.M.J., Lopes, J.A., 2013. A Review on the Applications of Portable Near-Infrared Spectrometers in the Agro-Food Industry. Appl Spectrosc 67, 1215–1233. https://doi.org/10.1366/13-07228

El-Hendawy, S., Al-Suhaibani, N., Alotaibi, M., Hassan, W., Elsayed, S., Tahir, M.U., Mohamed, A.I., Schmidhalter, U., 2019. Estimating growth and photosynthetic properties of wheat grown in simulated saline field conditions using hyperspectral reflectance sensing and multivariate analysis. Scientific Reports 2019 9:1 9, 1–15. https://doi.org/10.1038/s41598-019-52802-5

Elmetwalli, A.H., Tyler, A.N., Moghanm, F.S., Alamri, S.A.M., Eid, E.M., Elsayed, S., 2021. Integration of radiometric ground-based data and high-resolution quickbird imagery with multivariate modeling to estimate maize traits in the nile delta of Egypt. Sensors 21, 3915. https://doi.org/10.3390/S21113915/S1

Etienne, P., Diquelou, S., Prudent, M., Salon, C., Maillard, A., Ourry, A., 2018. Macro and micronutrient storage in plants and their remobilization when facing scarcity: The case of drought. Agriculture (Switzerland). https://doi.org/10.3390/agriculture8010014

Fan, X., Tang, S., Li, G., Zhou, X., 2016. Non-Invasive Detection of Protein Content in Several Types of Plant Feed Materials Using a Hybrid Near Infrared Spectroscopy Model. PLoS One 11, e0163145. https://doi.org/10.1371/journal.pone.0163145

Furbank, R.T., Silva-Perez, V., Evans, J.R., Condon, A.G., Estavillo, G.M., He, W., Newman, S., Poiré, R., Hall, A., He, Z., 2021. Wheat physiology predictor: predicting physiological traits in wheat from hyperspectral reflectance measurements using deep learning. Plant Methods 17, 1–15. https://doi.org/10.1186/S13007-021-00806-6/FIGURES/10

Garcia, C.B., Grusak, M.A., 2015. Mineral accumulation in vegetative and reproductive tissues during seed development in Medicago truncatula. Front Plant Sci 6, 155788. https://doi.org/10.3389/FPLS.2015.00622/BIBTEX

García-Gaytán, V., Hernández-Mendoza, F., Coria-Téllez, A.V., García-Morales, S., Sánchez-Rodríguez, E., Rojas-Abarca, L., Daneshvar, H., 2018. Fertigation: Nutrition, Stimulation and Bioprotection of the Root in High Performance. Plants 2018, Vol. 7, Page 88 7, 88. https://doi.org/10.3390/PLANTS7040088

Gashu, D., Nalivata, P.C., Amede, T., Ander, E.L., Bailey, E.H., Botoman, L., Chagumaira, C., Gameda, S., Haefele, S.M., Hailu, K., Joy, E.J.M., Kalimbira, A.A., Kumssa, D.B., Lark, R.M., Ligowe, I.S., McGrath, S.P., Milne, A.E., Mossa, A.W., Munthali, M., Towett, E.K., Walsh, M.G., Wilson, L., Young, S.D., Broadley, M.R., 2021. The nutritional quality of cereals varies geospatially in Ethiopia and Malawi. Nature 594, 71–76. https://doi.org/10.1038/s41586-021-03559-3

González-Piqueras, J., Lopez-Corcoles, H., Sánchez, S., Villodre, J., Bodas, V., Campos, I., Osann, A., Calera, A., 2017. Monitoring crop N status by using red edge-based indices. Advances in Animal Biosciences 8, 338–342. https://doi.org/10.1017/S2040470017000243

Guiboileau, A., Yoshimoto, K., Soulay, F., Bataillé, M., Avice, J., Masclaux-Daubresse, C., 2012. Autophagy machinery controls nitrogen remobilization at the whole-plant level under both limiting and ample nitrate conditions in Arabidopsis. New Phytologist 194, 732–740. https://doi.org/10.1111/j.1469-8137.2012.04084.x

Hatfield, J.L., Gitelson, A.A., Schepers, J.S., Walthall, C.L., 2008. Application of Spectral Remote Sensing for Agronomic Decisions. Agron J 100, S-117. https://doi.org/10.2134/AGRONJ2006.0370C

Hou, B., Hu, Y., Zhang, P., Hou, L., 2022. Potato Late Blight Severity and Epidemic Period Prediction Based on Vis/NIR Spectroscopy. Agriculture 2022, Vol. 12, Page 897 12, 897. https://doi.org/10.3390/AGRICULTURE12070897

Huang, J., Ma, H., Su, W., Zhang, X., Huang, Y., Fan, J., Wu, W., 2015. Jointly Assimilating MODIS LAI and et Products into the SWAP Model for Winter Wheat Yield Estimation. IEEE J Sel Top Appl Earth Obs Remote Sens 8, 4060–4071. https://doi.org/10.1109/JSTARS.2015.2403135

Jiao, Q., Sun, Q., Zhang, B., Huang, W., Ye, H., Zhang, Z., Zhang, X., Qian, B., 2022. A random forest algorithm for retrieving canopy chlorophyll content of wheat and soybean trained with prosail simulations using adjusted average leaf angle. Remote Sens (Basel) 14, 98. https://doi.org/10.3390/RS14010098/S1

Jie, W., Zhao, H.B., Shen, C.W., Chen, Q.W., Dong, C.X., Xu, Y.C., 2014. Determination of Nitrogen Concentration in Fresh Pear Leaves by Visible/Near-Infrared Reflectance Spectroscopy. Agron J 106, 1867–1872. https://doi.org/10.2134/AGRONJ13.0303

Kalaji, H.M., Bąba, W., Gediga, K., Goltsev, V., Samborska, I.A., Cetner, M.D., Dimitrova, S., Piszcz, U., Bielecki, K., Karmowska, K., Dankov, K., Kompała-Bąba, A., 2018. Chlorophyll fluorescence as a tool for nutrient status identification in rapeseed plants. Photosynth Res 136, 329–343. https://doi.org/10.1007/s11120-017-0467-7

Kawamura, K., Tsujimoto, Y., Nishigaki, T., Andriamananjara, A., Rabenarivo, M., Asai, H., Rakotoson, T., Razafimbelo, T., 2019. Laboratory Visible and Near-Infrared Spectroscopy with Genetic Algorithm-Based Partial Least Squares Regression for Assessing the Soil Phosphorus Content of Upland and Lowland Rice Fields in Madagascar. Remote Sensing 2019, Vol. 11, Page 506 11, 506. https://doi.org/10.3390/RS11050506

Keutgen, A.J., Wszelaczyńska, E., Pobereżny, J., Przewodowska, A., Przewodowski, W., Milczarek, D., Tatarowska, B., Flis, B., Keutgen, N., 2019. Antioxidant properties of potato tubers (Solanum

tuberosum L.) as a consequence of genetic potential and growing conditions. PLoS One 14, e0222976. https://doi.org/10.1371/JOURNAL.PONE.0222976

Li, D., Miao, Y., Ransom, C.J., Bean, G. Mac, Kitchen, N.R., Fernández, F.G., Sawyer, J.E., Camberato, J.J., Carter, P.R., Ferguson, R.B., Franzen, D.W., Laboski, C.A.M., Nafziger, E.D., Shanahan, J.F., 2022. Corn Nitrogen Nutrition Index Prediction Improved by Integrating Genetic, Environmental, and Management Factors with Active Canopy Sensing Using Machine Learning. Remote Sensing 2022, Vol. 14, Page 394 14, 394. https://doi.org/10.3390/RS14020394

Li, D., Zhang, P., Chen, T., Qin, W., 2020. Recent Development and Challenges in Spectroscopy and Machine Vision Technologies for Crop Nitrogen Diagnosis: A Review. Remote Sensing 2020, Vol. 12, Page 2578 12, 2578. https://doi.org/10.3390/RS12162578

Liu, S., Peng, Y., Du, W., Le, Y., Li, L., 2015. Remote Estimation of Leaf and Canopy Water Content in Winter Wheat with Different Vertical Distribution of Water-Related Properties. Remote Sensing 2015, Vol. 7, Pages 4626-4650 7, 4626–4650. https://doi.org/10.3390/RS70404626

Liu, S., Yang, X., Guan, Q., Lu, Z., Lu, J., 2020. An Ensemble Modeling Framework for Distinguishing Nitrogen, Phosphorous and Potassium Deficiencies in Winter Oilseed Rape (Brassica napus L.) Using Hyperspectral Data. Remote Sensing 2020, Vol. 12, Page 4060 12, 4060. https://doi.org/10.3390/RS12244060

Lynch, J.P., 2019. Root phenotypes for improved nutrient capture: an underexploited opportunity for global agriculture. New Phytologist 223, 548–564. https://doi.org/10.1111/NPH.15738

Ma, J., Zheng, B., He, Y., 2022. Applications of a Hyperspectral Imaging System Used to Estimate Wheat Grain Protein: A Review. Front Plant Sci 13, 837200. https://doi.org/10.3389/fpls.2022.837200

Maillard, A., Diquélou, S., Billard, V., Laîné, P., Garnica, M., Prudent, M., Garcia-Mina, J.M., Yvin, J.C., Ourry, A., 2015. Leaf mineral nutrient remobilization during leaf senescence and modulation by nutrient deficiency. Front Plant Sci 6, 1–15. https://doi.org/10.3389/FPLS.2015.00317/ABSTRACT

Maqbool, R., Percival, D.C., Adl, M.S., Zaman, Q.U., Buszard, D., 2012. In situ estimation of foliar nitrogen in wild blueberry using reflectance spectra. Canadian Journal of Plant Science 92, 1155–1161. https://doi.org/10.4141/CJPS2011-203/ASSET/IMAGES/LARGE/CJPS2011-203F5.JPEG

Marcos-barbero, E.L., Pérez, P., Martínez-carrasco, R., Arellano, J.B., Morcuende, R., 2021. Genotypic variability on grain yield and grain nutritional quality characteristics of wheat grown under elevated co2 and high temperature. Plants 10, 1043. https://doi.org/10.3390/PLANTS10061043/S1

Mashiane, K., Adelabu, S., Ramoelo, A., 2023. Comparative Analysis of Single Bands, Vegetation Indices, and Their Combination in Predicting Grass Species Nitrogen in a Protected Mountainous Area. Applied Sciences 2023, Vol. 13, Page 7960 13, 7960. https://doi.org/10.3390/APP13137960

McGRATH, V.B., BLAKENEY, A.B., BATTEN, G.D., 1997. Fructan to nitrogen ratio as an indicator of nutrient stress in wheat crops. New Phytologist 136, 145–152. https://doi.org/10.1111/J.1469-8137.1997.TB04741.X

McIntosh, A.R., Lobaugh, N.J., 2004. Partial least squares analysis of neuroimaging data: applications and advances. Neuroimage 23, S250–S263. https://doi.org/10.1016/J.NEUROIMAGE.2004.07.020

Meacham-Hensold, K., Montes, C.M., Wu, J., Guan, K., Fu, P., Ainsworth, E.A., Pederson, T., Moore, C.E., Brown, K.L., Raines, C., Bernacchi, C.J., 2019. High-throughput field phenotyping using hyperspectral reflectance and partial least squares regression (PLSR) reveals genetic modifications to photosynthetic capacity. Remote Sens Environ 231, 111176. https://doi.org/10.1016/J.RSE.2019.04.029

Mejía-Correal, K.B., Marcelo, V., Sanz-Ablanedo, E., Rodríguez-Pérez, J.R., 2023. Total Soluble Solids in Grape Must Estimation Using VIS-NIR-SWIR Reflectance Measured in Fresh Berries. Agronomy 13, 2275. https://doi.org/10.3390/AGRONOMY13092275/S1

Mercuri, A.M., Torri, P., Casini, E., Olmi, L., 2013. Climate warming and the decline of Taxus airborne pollen in urban pollen rain (Emilia Romagna, northern Italy). Plant Biol 15, 70–82. https://doi.org/10.1111/J.1438-8677.2012.00624.X

Mevik, B.-H., Wehrens, R., 2007. Journal of Statistical Software The pls Package: Principal Component and Partial Least Squares Regression in R.

Miphokasap, P., Honda, K., Vaiphasa, C., Souris, M., Nagai, M., 2012. Estimating Canopy Nitrogen Concentration in Sugarcane Using Field Imaging Spectroscopy. Remote Sensing 2012, Vol. 4, Pages 1651-1670 4, 1651–1670. https://doi.org/10.3390/RS4061651

Muñoz-Huerta, R.F., Guevara-Gonzalez, R.G., Contreras-Medina, L.M., Torres-Pacheco, I., Prado-Olivarez, J., Ocampo-Velazquez, R. V., 2013. A Review of Methods for Sensing the Nitrogen Status in Plants: Advantages, Disadvantages and Recent Advances. Sensors 2013, Vol. 13, Pages 10823-10843 13, 10823–10843. https://doi.org/10.3390/S130810823

Muñoz-Huerta, R.F., Ortiz-Melendez, A. de J., Guevara-Gonzalez, R.G., Torres-Pacheco, I., Herrera-Ruiz, G., Contreras-Medina, L.M., Prado-Olivarez, J., Ocampo-Velazquez, R. V., 2014. An Analysis of Electrical Impedance Measurements Applied for Plant N Status Estimation in Lettuce (Lactuca sativa). Sensors 2014, Vol. 14, Pages 11492-11503 14, 11492–11503. https://doi.org/10.3390/S140711492

Myers, S.S., Smith, M.R., Guth, S., Golden, C.D., Vaitla, B., Mueller, N.D., Dangour, A.D., Huybers, P., 2017. Climate Change and Global Food Systems: Potential Impacts on Food Security and Undernutrition. Annu Rev Public Health 38, 259–277. https://doi.org/10.1146/annurev-publhealth-031816-044356

Myers, S.S., Zanobetti, A., Kloog, I., Huybers, P., Leakey, A.D.B., Bloom, A.J., Carlisle, E., Dietterich, L.H., Fitzgerald, G., Hasegawa, T., Holbrook, N.M., Nelson, R.L., Ottman, M.J., Raboy, V., Sakai, H., Sartor, K.A., Schwartz, J., Seneweera, S., Tausz, M., Usui, Y., 2014. Increasing CO2 threatens human nutrition. Nature 510, 139–142. https://doi.org/10.1038/nature13179

Naik, B.B., Naveen, H.R., Sreenivas, G., Choudary, K.K., Devkumar, D., Adinarayana, J., 2020. Identification of Water and Nitrogen Stress Indicative Spectral Bands Using Hyperspectral Remote Sensing in Maize During Post-Monsoon Season. Journal of the Indian Society of Remote Sensing 48, 1787–1795. https://doi.org/10.1007/S12524-020-01200-W/FIGURES/8

Narmilan, A., Gonzalez, F., Salgadoe, A.S.A., Kumarasiri, U.W.L.M., Weerasinghe, H.A.S., Kulasekara, B.R., 2022. Predicting Canopy Chlorophyll Content in Sugarcane Crops Using Machine Learning Algorithms and Spectral Vegetation Indices Derived from UAV Multispectral Imagery. Remote Sensing 2022, Vol. 14, Page 1140 14, 1140. https://doi.org/10.3390/RS14051140

Nguyen, V.C., Jeong, S., Ko, J., Ng, C.T., Yeom, J., 2019. Mathematical Integration of Remotely-Sensed Information into a Crop Modelling Process for Mapping Crop Productivity. Remote Sensing 2019, Vol. 11, Page 2131 11, 2131. https://doi.org/10.3390/RS11182131

Ni, B., Wang, Daqing, Deng, Z., Xu, H., Wang, Donghao, Jiang, X., 2018. Review on the Groundwater Potential Evaluation Based on Remote Sensing Technology. IOP Conf Ser Mater Sci Eng 394, 052038. https://doi.org/10.1088/1757-899X/394/5/052038

Pandey, P., Ge, Y., Stoerger, V., Schnable, J.C., 2017. High throughput in vivo analysis of plant leaf chemical properties using hyperspectral imaging. Front Plant Sci 8. https://doi.org/10.3389/fpls.2017.01348

Peleg, Z., Cakmak, I., Ozturk, L., Yazici, A., Jun, Y., Budak, H., Korol, A.B., Fahima, T., Saranga, Y., 2009. Quantitative trait loci conferring grain mineral nutrient concentrations in durum wheat × wild emmer wheat RIL population. Theoretical and Applied Genetics 119, 353–369. https://doi.org/10.1007/S00122-009-1044-Z/METRICS

Peng, Y., Zhao, L., Hu, Y., Wang, G., Wang, L., Liu, Z., 2019. Prediction of Soil Nutrient Contents Using Visible and Near-Infrared Reflectance Spectroscopy. ISPRS International Journal of Geo-Information 2019, Vol. 8, Page 437 8, 437. https://doi.org/10.3390/IJGI8100437

Pereira-Obaya, D., Castedo-Dorado, F., Sanz-Ablanedo, E., Mejía-Correal, K.B., Rodríguez-Pérez, J.R., 2023. Leaf Trait Hyperspectral Characterization of Castanea sativa Miller Affected by Dryocosmus kuriphilus Yasumatsu. Agronomy 2023, Vol. 13, Page 923 13, 923. https://doi.org/10.3390/AGRONOMY13030923

Poggi, G.M., Aloisi, I., Corneti, S., Esposito, E., Naldi, M., Fiori, J., Piana, S., Ventura, F., 2022. Climate change effects on bread wheat phenology and grain quality: A case study in the north of Italy. Front Plant Sci 13, 936991. https://doi.org/10.3389/FPLS.2022.936991/BIBTEX

Porter, J.R., Semenov, M.A., 2005. Crop responses to climatic variation. Philosophical Transactions of the Royal Society B: Biological Sciences 360, 2021–2035. https://doi.org/10.1098/RSTB.2005.1752

Prananto, J.A., Minasny, B., Weaver, T., 2020. Near infrared (NIR) spectroscopy as a rapid and cost-effective method for nutrient analysis of plant leaf tissues, in: Advances in Agronomy. Academic Press, pp. 1–49. https://doi.org/10.1016/bs.agron.2020.06.001

Prey, L., von Bloh, M., Schmidhalter, U., 2018. Evaluating RGB Imaging and Multispectral Active and Hyperspectral Passive Sensing for Assessing Early Plant Vigor in Winter Wheat. Sensors 2018, Vol. 18, Page 2931 18, 2931. https://doi.org/10.3390/S18092931

Qian, Y., Yang, Z., Di, L., Rahman, M.S., Tan, Z., Xue, L., Gao, F., Yu, E.G., Zhang, X., 2019. Crop Growth Condition Assessment at County Scale Based on Heat-Aligned Growth Stages. Remote Sensing 2019, Vol. 11, Page 2439 11, 2439. https://doi.org/10.3390/RS11202439

Revill, A., Florence, A., Macarthur, A., Hoad, S., Rees, R., Williams, M., 2020. Quantifying Uncertainty and Bridging the Scaling Gap in the Retrieval of Leaf Area Index by Coupling Sentinel-2 and UAV Observations. Remote Sensing 2020, Vol. 12, Page 1843 12, 1843. https://doi.org/10.3390/RS12111843

Reyes, F., Gosme, M., Wolz, K.J., Lecomte, I., Dupraz, C., 2021. Alley cropping mitigates the impacts of climate change on a wheat crop in a mediterranean environment: A biophysical model-based assessment. Agriculture (Switzerland) 11, 356. https://doi.org/10.3390/AGRICULTURE11040356/S1

Robles-Zazueta, C.A., Pinto, F., Molero, G., Foulkes, M.J., Reynolds, M.P., Murchie, E.H., 2022. Prediction of Photosynthetic, Biophysical, and Biochemical Traits in Wheat Canopies to Reduce the Phenotyping Bottleneck. Front Plant Sci 13, 828451. https://doi.org/10.3389/fpls.2022.828451

Rosipal, R., Krämer, N., 2006. Overview and recent advances in partial least squares. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) 3940 LNCS, 34–51. https://doi.org/10.1007/11752790_2/COVER

Roth, G.W., Fox, R.H., Marshall, H.G., 1989. Plant Tissue Tests for Predicting Nitrogen Fertilizer Requirements of Winter Wheat. Agron J 81, 502–507. https://doi.org/10.2134/AGRONJ1989.00021962008100030022X

Ryan, K., Ali, K., 2016. Application of a partial least-squares regression model to retrieve chlorophyll-a concentrations in coastal waters using hyper-spectral data. Ocean Science Journal 51, 209–221. https://doi.org/10.1007/s12601-016-0018-8

Schirrmann, M., Giebel, A., Gleiniger, F., Pflanz, M., Lentschke, J., Dammer, K.H., 2016. Monitoring Agronomic Parameters of Winter Wheat Crops with Low-Cost UAV Imagery. Remote Sensing 2016, Vol. 8, Page 706 8, 706. https://doi.org/10.3390/RS8090706

Severtson, D., Callow, N., Flower, K., Neuhaus, A., Olejnik, M., Nansen, C., 2016. Unmanned aerial vehicle canopy reflectance data detects potassium deficiency and green peach aphid susceptibility in canola. Precis Agric 17, 659–677. https://doi.org/10.1007/S11119-016-9442-0/FIGURES/7

Sharifi, A., 2020. Remotely sensed vegetation indices for crop nutrition mapping. J Sci Food Agric 100, 5191–5196. https://doi.org/10.1002/JSFA.10568

Shiferaw, B., Smale, M., Braun, H.J., Duveiller, E., Reynolds, M., Muricho, G., 2013. Crops that feed the world 10. Past successes and future challenges to the role played by wheat in global food security. Food Secur 5, 291–317. https://doi.org/10.1007/S12571-013-0263-Y/TABLES/9

Shrawat, A.K., Armstrong, C.L., 2018. Development and Application of Genetic Engineering for Wheat Improvement. CRC Crit Rev Plant Sci 37, 335–421. https://doi.org/10.1080/07352689.2018.1514718

Singh, M., Karada, M.S., Rai, R.K., Pratap, D., Agnihotri, D., Singh, A.K., Singh, B.K., 2023. A Review on Remote Sensing as a Tool for Irrigation Monitoring and Management. International Journal of Environment and Climate Change 13, 203–211. https://doi.org/10.9734/IJECC/2023/V13I61817

Smith, M.R., Veneklaas, E., Polania, J., Rao, I.M., Beebe, S.E., Merchant, A., 2019. Field drought conditions impact yield but not nutritional quality of the seed in common bean (Phaseolus vulgaris L.). PLoS One 14, e0217099. https://doi.org/10.1371/JOURNAL.PONE.0217099

Soares, J.C., Santos, C.S., Carvalho, S.M.P., Pintado, M.M., Vasconcelos, M.W., 2019. Preserving the nutritional quality of crop plants under a changing climate: importance and strategies. Plant and Soil 2019 443:1 443, 1–26. https://doi.org/10.1007/S11104-019-04229-0

Stagnari, F., Polilli, W., Campanelli, G., Platani, C., Trasmundi, F., Scortichini, G., Galieni, A., 2023. Nitrate Content Assessment in Spinach: Exploring the Potential of Spectral Reflectance in Open Field Experiments. Agronomy 2023, Vol. 13, Page 193 13, 193. https://doi.org/10.3390/AGRONOMY13010193

Stanton, C., Starek, M.J., Elliott, N., Brewer, M., Maeda, M.M., Chu, T., 2017. Unmanned aircraft system-derived crop height and normalized difference vegetation index metrics for sorghum yield and aphid stress assessment. J Appl Remote Sens 11, 026035. https://doi.org/10.1117/1.jrs.11.026035

Su, P., Tarkoma, S., Pellikka, P.K.E., 2020. Band Ranking via Extended Coefficient of Variation for Hyperspectral Band Selection. Remote Sensing 2020, Vol. 12, Page 3319 12, 3319. https://doi.org/10.3390/RS12203319

Sun, Q., Jiao, Q., Chen, X., Xing, H., Huang, W., Zhang, B., 2023. Machine Learning Algorithms for the Retrieval of Canopy Chlorophyll Content and Leaf Area Index of Crops Using the PROSAIL-D Model with the Adjusted Average Leaf Angle. Remote Sensing 2023, Vol. 15, Page 2264 15, 2264. https://doi.org/10.3390/RS15092264

Tagliabue, G., Boschetti, M., Bramati, G., Candiani, G., Colombo, R., Nutini, F., Pompilio, L., Rivera-Caicedo, J.P., Rossi, M., Rossini, M., Verrelst, J., Panigada, C., 2022. Hybrid retrieval of crop traits from multi-temporal PRISMA hyperspectral imagery. ISPRS Journal of Photogrammetry and Remote Sensing 187, 362–377. https://doi.org/10.1016/J.ISPRSJPRS.2022.03.014

Tan, C., Zhou, X., Zhang, P., Wang, Z., Wang, D., Guo, W., Yun, F., 2020. Predicting grain protein content of field-grown winter wheat with satellite images and partial least square algorithm. PLoS One 15, e0228500. https://doi.org/10.1371/JOURNAL.PONE.0228500

Tao, W., Wu, J., Wang, Q., 2017. Mathematical model of sediment and solute transport along slope land in different rainfall pattern conditions. Scientific Reports 2017 7:1 7, 1–11. https://doi.org/10.1038/srep44082

Thorp, K.R., Wang, G., West, A.L., Moran, M.S., Bronson, K.F., White, J.W., Mon, J., 2012. Estimating crop biophysical properties from remote sensing data by inverting linked radiative transfer and ecophysiological models. Remote Sens Environ 124, 224–233. https://doi.org/10.1016/J.RSE.2012.05.013

Tomíček, J., Mišurec, J., Lukeš, P., 2021. Prototyping a Generic Algorithm for Crop Parameter Retrieval across the Season Using Radiative Transfer Model Inversion and Sentinel-2 Satellite Observations. Remote Sensing 2021, Vol. 13, Page 3659 13, 3659. https://doi.org/10.3390/RS13183659

Uprety, D.C., Sen, S., Dwivedi, N., 2010. Rising atmospheric carbon dioxide on grain quality in crop plants. Physiology and Molecular Biology of Plants 16, 215–227. https://doi.org/10.1007/S12298-010-0029-3/METRICS

Vereecken, H., Weihermüller, L., Jonard, F., Montzka, C., 2012. Characterization of Crop Canopies and Water Stress Related Phenomena using Microwave Remote Sensing Methods: A Review. Vadose Zone Journal 11, vzj2011.0138ra. https://doi.org/10.2136/VZJ2011.0138RA

Vitale, L., Di Tommasi, P., D'Urso, G., Magliulo, V., 2016. The response of ecosystem carbon fluxes to LAI and environmental drivers in a maize crop grown in two contrasting seasons. Int J Biometeorol 60, 411–420. https://doi.org/10.1007/S00484-015-1038-2/METRICS

Vohland, M., Emmerling, C., 2011. Determination of total soil organic C and hot water-extractable C from VIS-NIR soil reflectance with partial least squares regression and spectral feature selection techniques. Eur J Soil Sci 62, 598–606. https://doi.org/10.1111/J.1365-2389.2011.01369.X

Wali, E., Tasumi, M., Moriyama, M., 2020. Combination of Linear Regression Lines to Understand the Response of Sentinel-1 Dual Polarization SAR Data with Crop Phenology—Case Study in Miyazaki, Japan. Remote Sensing 2020, Vol. 12, Page 189 12, 189. https://doi.org/10.3390/RS12010189

Wang, H.J., Zhou, Y., Wang, S.X., Wang, F.T., Zhao, Q., 2020. APPLICATION OF REMOTE SENSING BIG DATA FOR RAPID RESPOND TO LANDSLIDE LAKE DISASTER MONITOR. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences XLII-3-W10, 37–43. https://doi.org/10.5194/ISPRS-ARCHIVES-XLII-3-W10-37-2020

Wang, L., Xia, H., Li, X., Qiao, Y., Xue, Yanhui, Jiang, X., Yan, W., Liu, Y., Xue, Yanfang, Kong, L., 2021. Source–sink manipulation affects accumulation of zinc and other nutrient elements in wheat grains. Plants 10, 1032. https://doi.org/10.3390/PLANTS10051032/S1

Weiss, M., Jacob, F., Duveiller, G., 2020. Remote sensing for agricultural applications: A meta-review. Remote Sens Environ 236, 111402. https://doi.org/10.1016/j.rse.2019.111402

Wold, S., Sjostrom, M., Eriksson, L., Sweden°°, S., 2001. PLS-regression: a basic tool of chemometrics. Chemometrics and Intelligent Laboratory Systems 58, 109–130.

Wright, I.J., Reich, P.B., Cornelissen, J.H.C., Falster, D.S., Groom, P.K., Hikosaka, K., Lee, W., Lusk, C.H., Niinemets, Ü., Oleksyn, J., Osada, N., Poorter, H., Warton, D.I., Westoby, M., 2005. Modulation of leaf economic traits and trait relationships by climate. Global Ecology and Biogeography 14, 411–421. https://doi.org/10.1111/j.1466-822x.2005.00172.x

Xia, Y., Guan, K., Copenhaver, K., Wander, M., 2021. Estimating cover crop biomass nitrogen credits with Sentinel-2 imagery and sites covariates. Agron J 113, 1084–1101. https://doi.org/10.1002/AGJ2.20525

Xie, W., Xue, Y., Zhai, L., Sang, H., 2013. DATA FUSION TECHNOLOGY OF MULTI-PLATFORM EARTH OBSERVATION ON AGRICULTURE. The International Archives of the

Photogrammetry, Remote Sensing and Spatial Information Sciences XL-7-W1, 189–192. https://doi.org/10.5194/ISPRSARCHIVES-XL-7-W1-189-2013

Xu, S., Zhao, Y., Shi, X., Wang, M., 2017. Rapid Determination of Carbon, Nitrogen, and Phosphorus Contents of Field Crops in China Using Visible and Near-Infrared Reflectance Spectroscopy. Crop Sci 57, 475–489. https://doi.org/10.2135/CROPSCI2016.03.0194

Yang, X.H., Huang, J.F., Wu, Y.P., Wang, J.W., Wang, P., Wang, X.M., Huete, A.R., 2011. Estimating biophysical parameters of rice with remote sensing data using support vector machines. Sci China Life Sci 54, 272–281. https://doi.org/10.1007/S11427-011-4135-4/METRICS

Yu, K., Anderegg, J., Mikaberidze, A., Karisto, P., Mascher, F., McDonald, B.A., Walter, A., Hund, A., 2018. Hyperspectral canopy sensing of wheat septoria tritici blotch disease. Front Plant Sci 9, 326021. https://doi.org/10.3389/FPLS.2018.01195/BIBTEX

Yue, J., Feng, H., Yang, G., Li, Z., 2018. A Comparison of Regression Techniques for Estimation of Above-Ground Winter Wheat Biomass Using Near-Surface Spectroscopy. Remote Sensing 2018, Vol. 10, Page 66 10, 66. https://doi.org/10.3390/RS10010066

Zhao, C., Zhou, Q., Wang, J., Huang, W., 2004. Spectral Indices Redefined in Detecting Nitrogen Availability for Wheat Canopy. Commun Soil Sci Plant Anal 35, 853–864. https://doi.org/10.1081/CSS-120030362

Zhao, H., Song, X., Yang, G., Li, Z., Zhang, D., Feng, H., 2019. Monitoring of Nitrogen and Grain Protein Content in Winter Wheat Based on Sentinel-2A Data. Remote Sensing 2019, Vol. 11, Page 1724 11, 1724. https://doi.org/10.3390/RS11141724

Zhou, X., Huang, W., Kong, W., Ye, H., Luo, J., Chen, P., 2016. Remote estimation of canopy nitrogen content in winter wheat using airborne hyperspectral reflectance measurements. Advances in Space Research 58, 1627–1637. https://doi.org/10.1016/J.ASR.2016.06.034
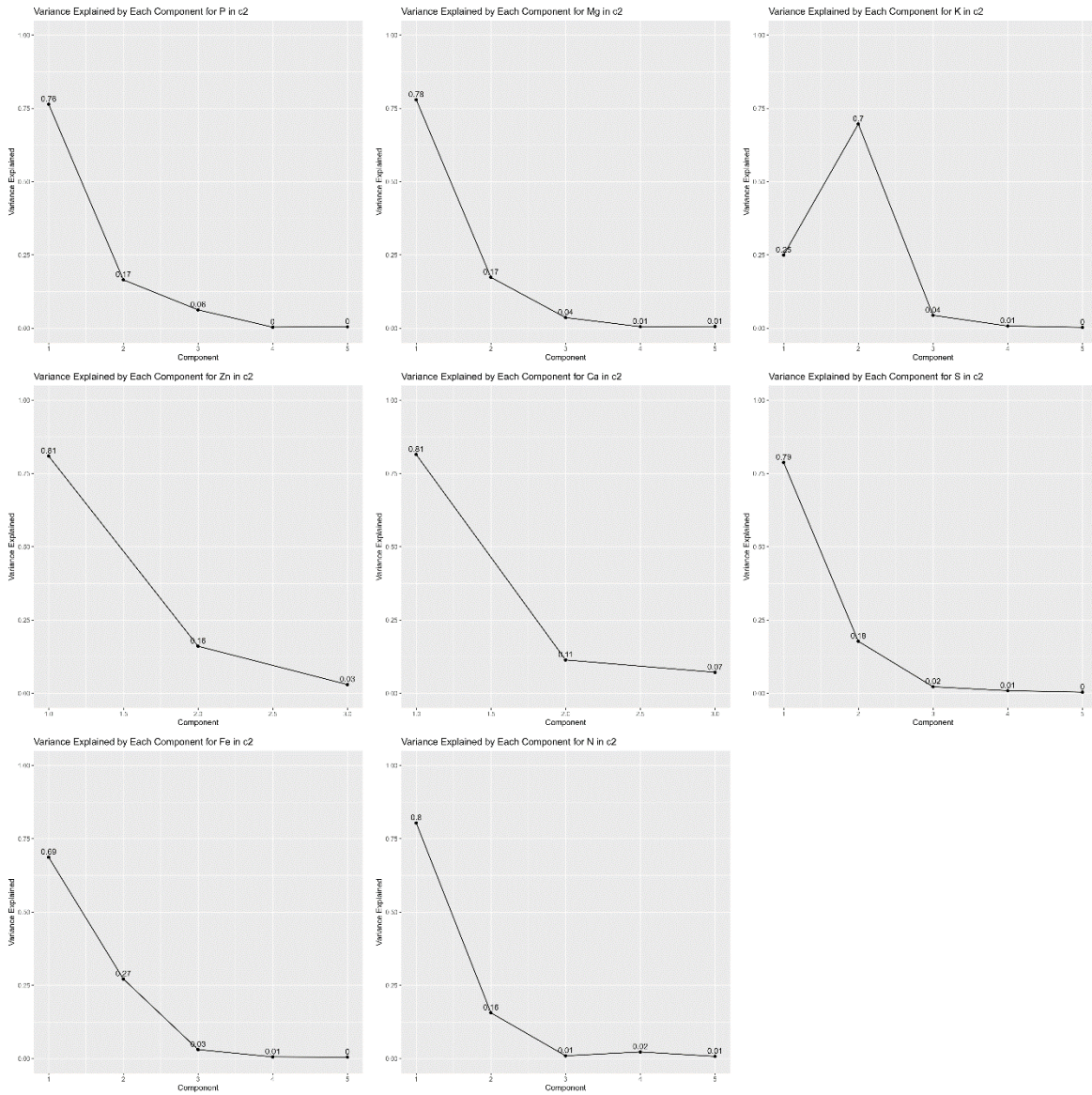
APPENDICES

Appendix A: Percentage of explained variances by each PLSR component for each nutrient at vegetative and reproductive stages

(a) Explained variances by each component at vegetative stage
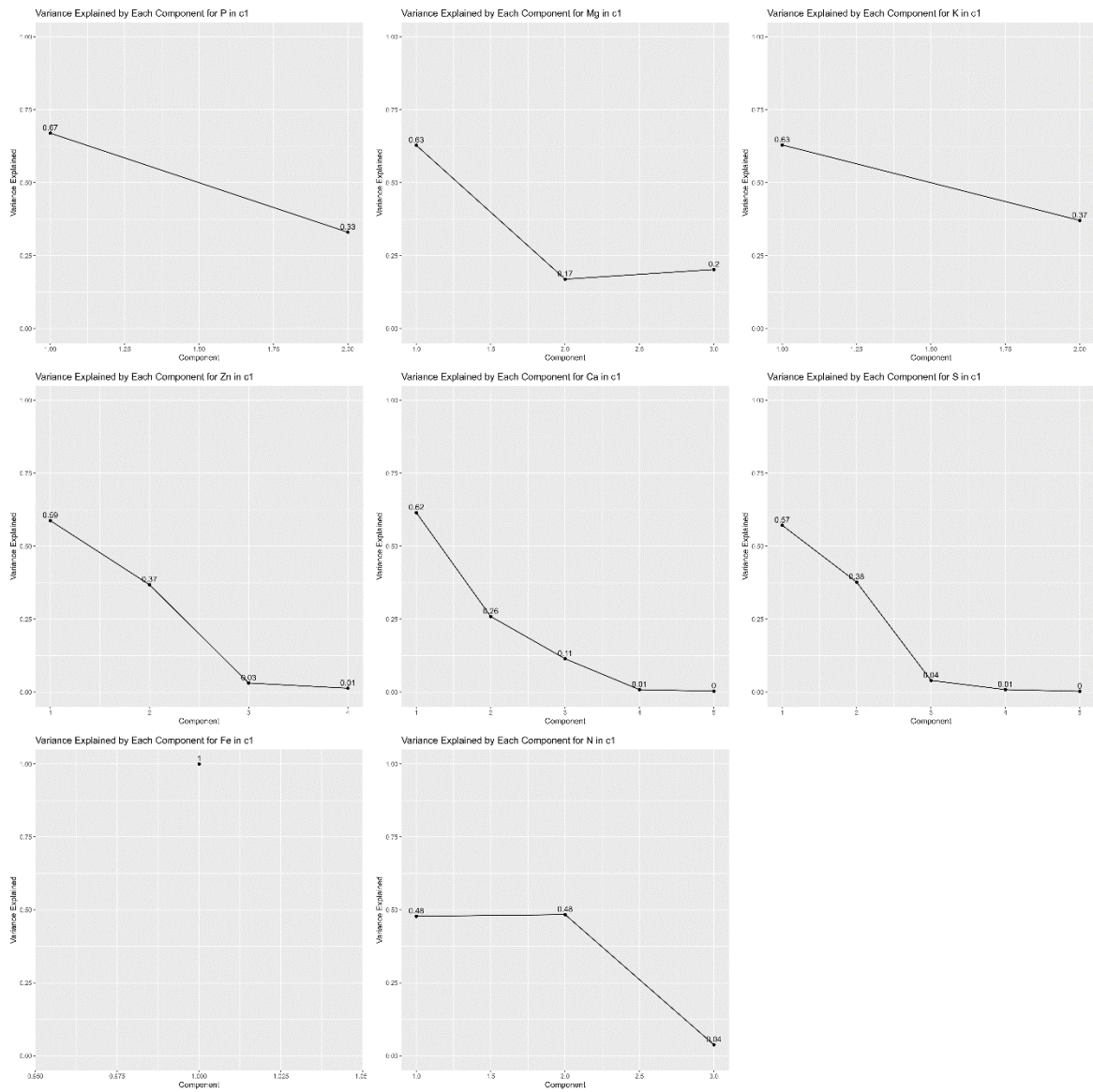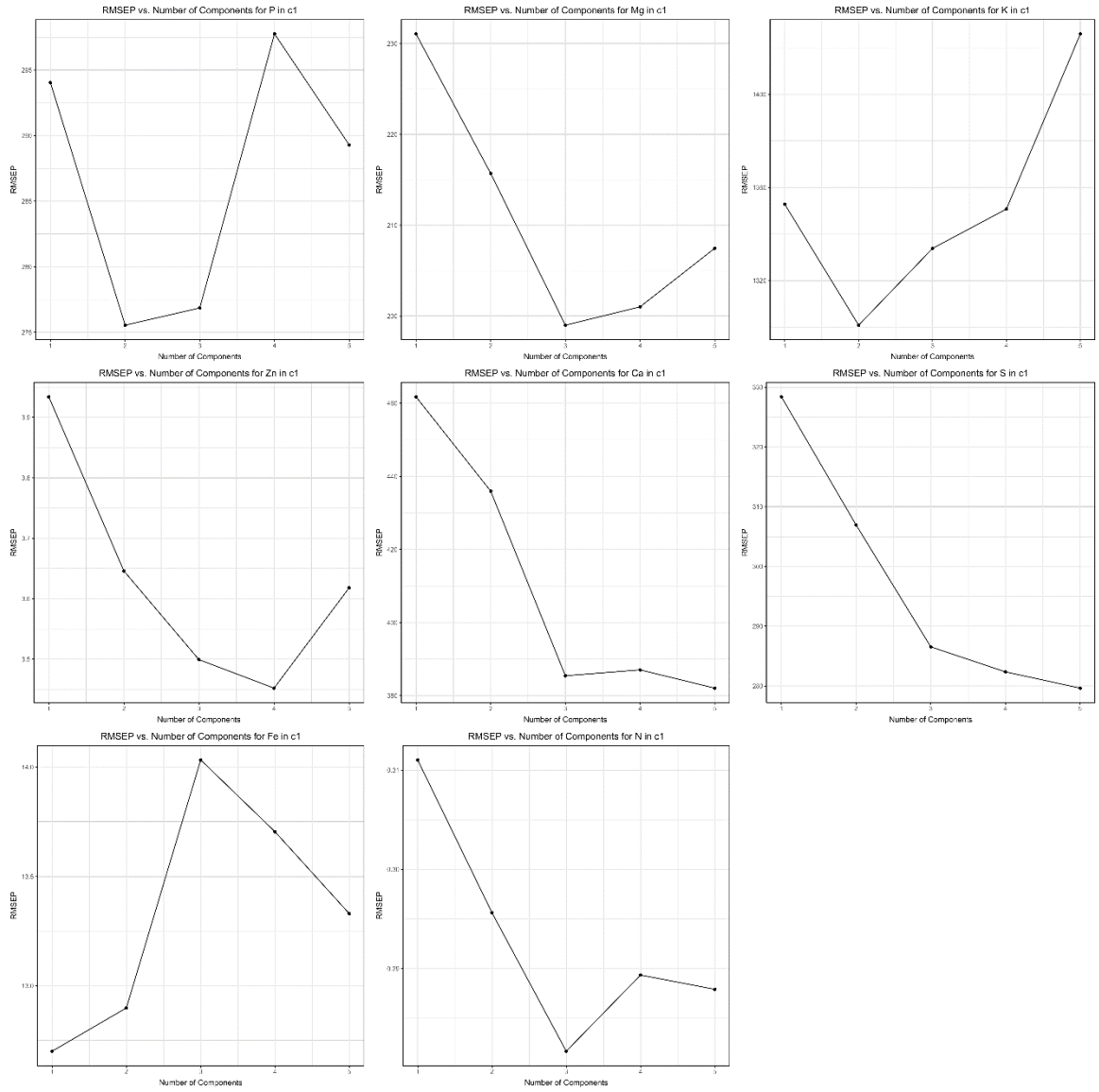
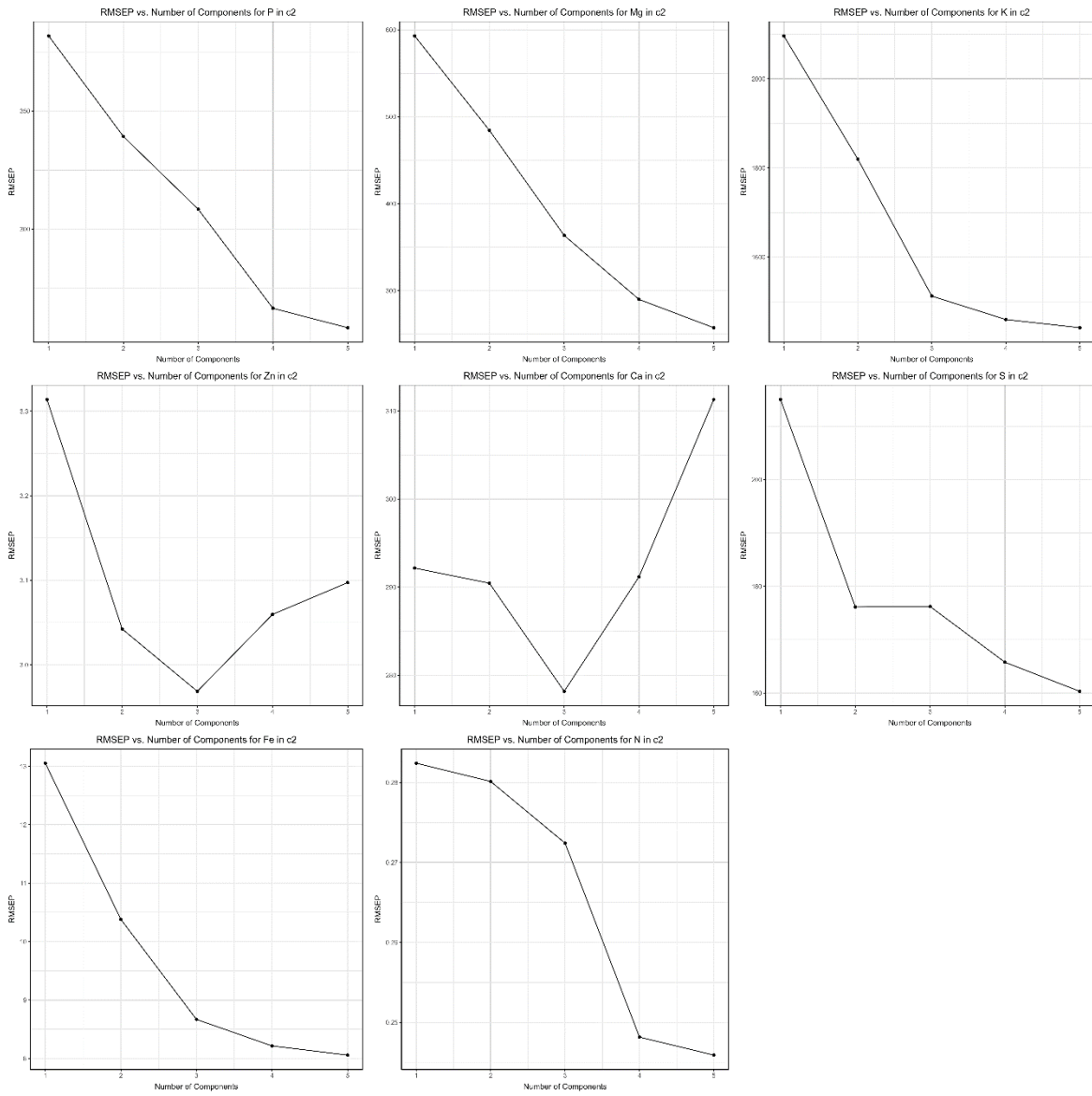*(b) Explained variances each component at reproductive stage*

Appendix B: Cross validation results for determining optimal PLSR number of components for different models at vegetative and reproductive stages
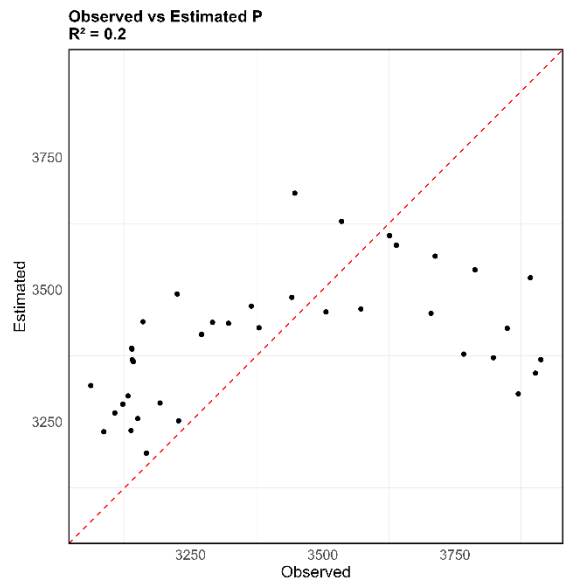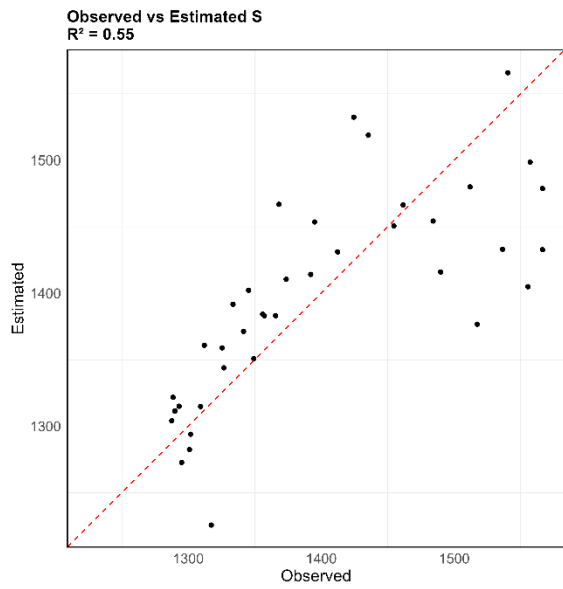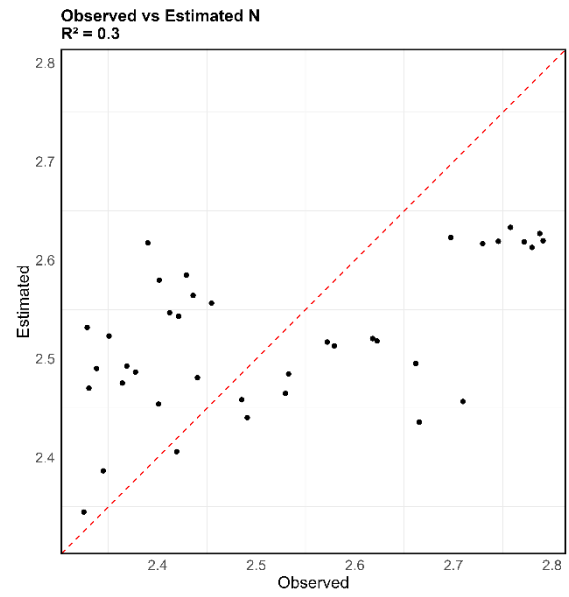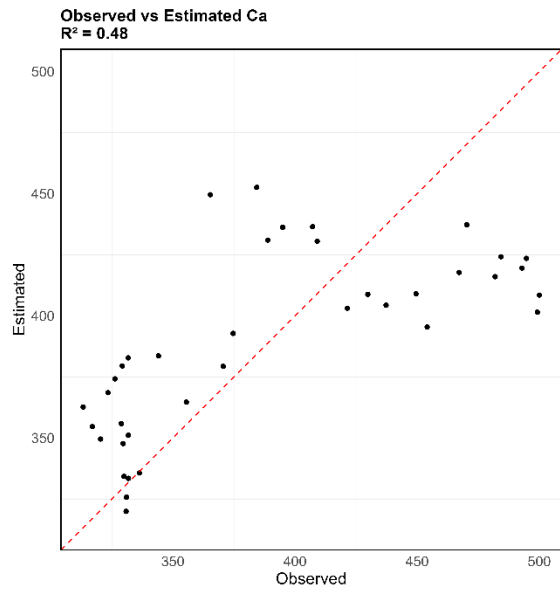
(a) Cross-validation results - vegetative stage

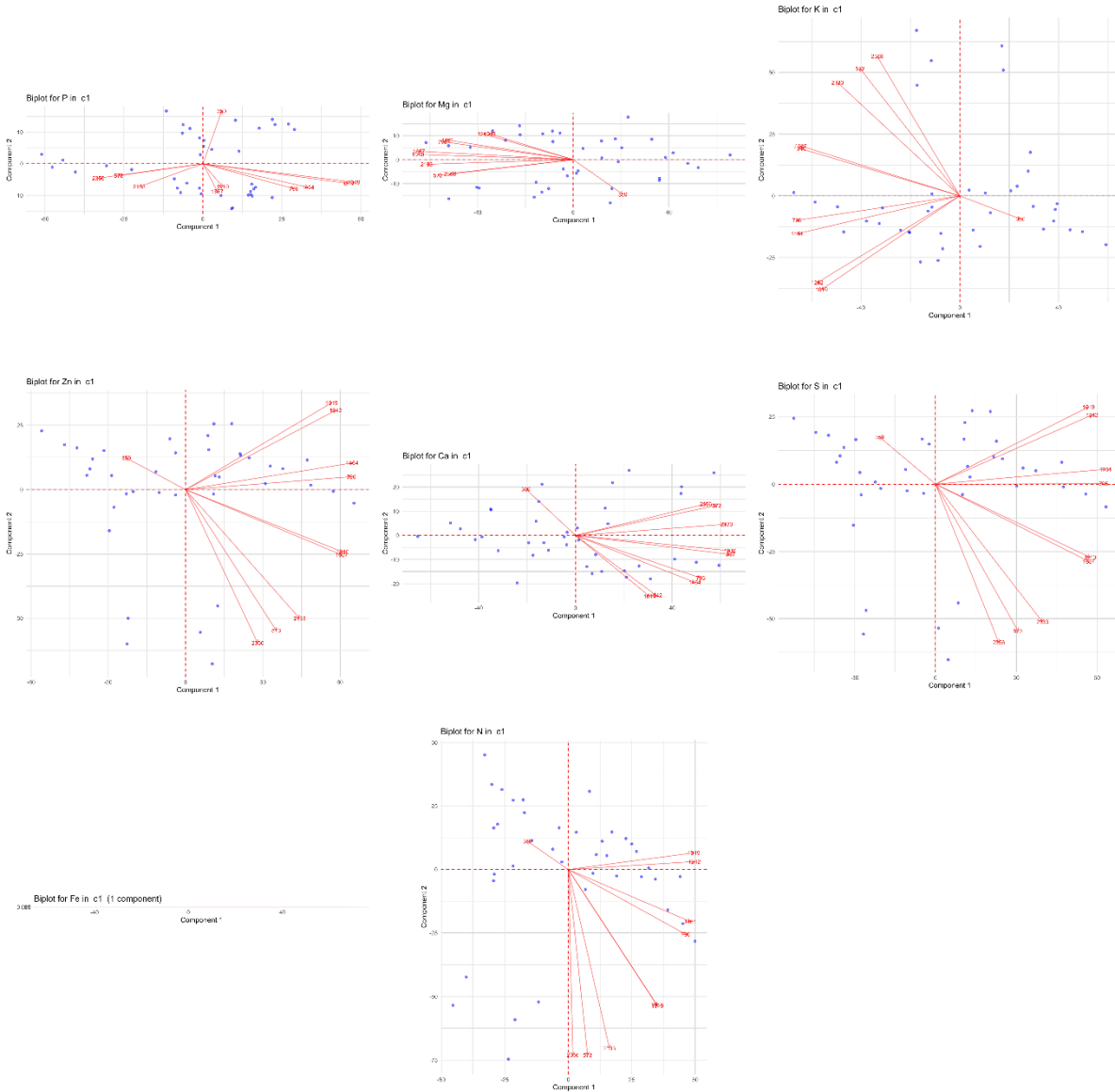## *(b) Cross-validation results – reproductive stage*

Appendix C: Scatterplots showing comparison of poorly estimated nutrients in grains compared to observed nutrients

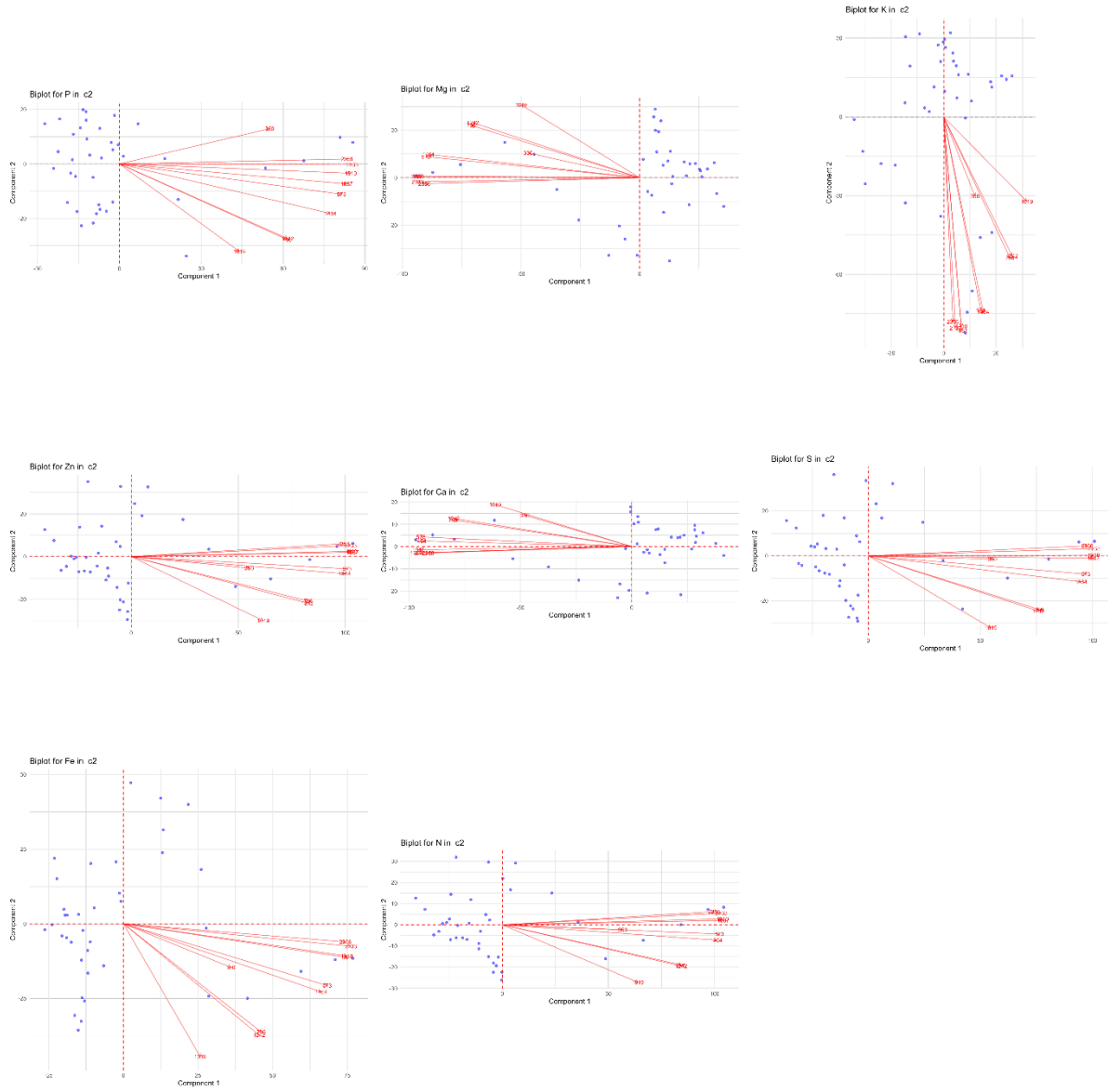## Appendix D: Biplots of first two PLSR components at the vegetative and reproductive stages for each nutrient
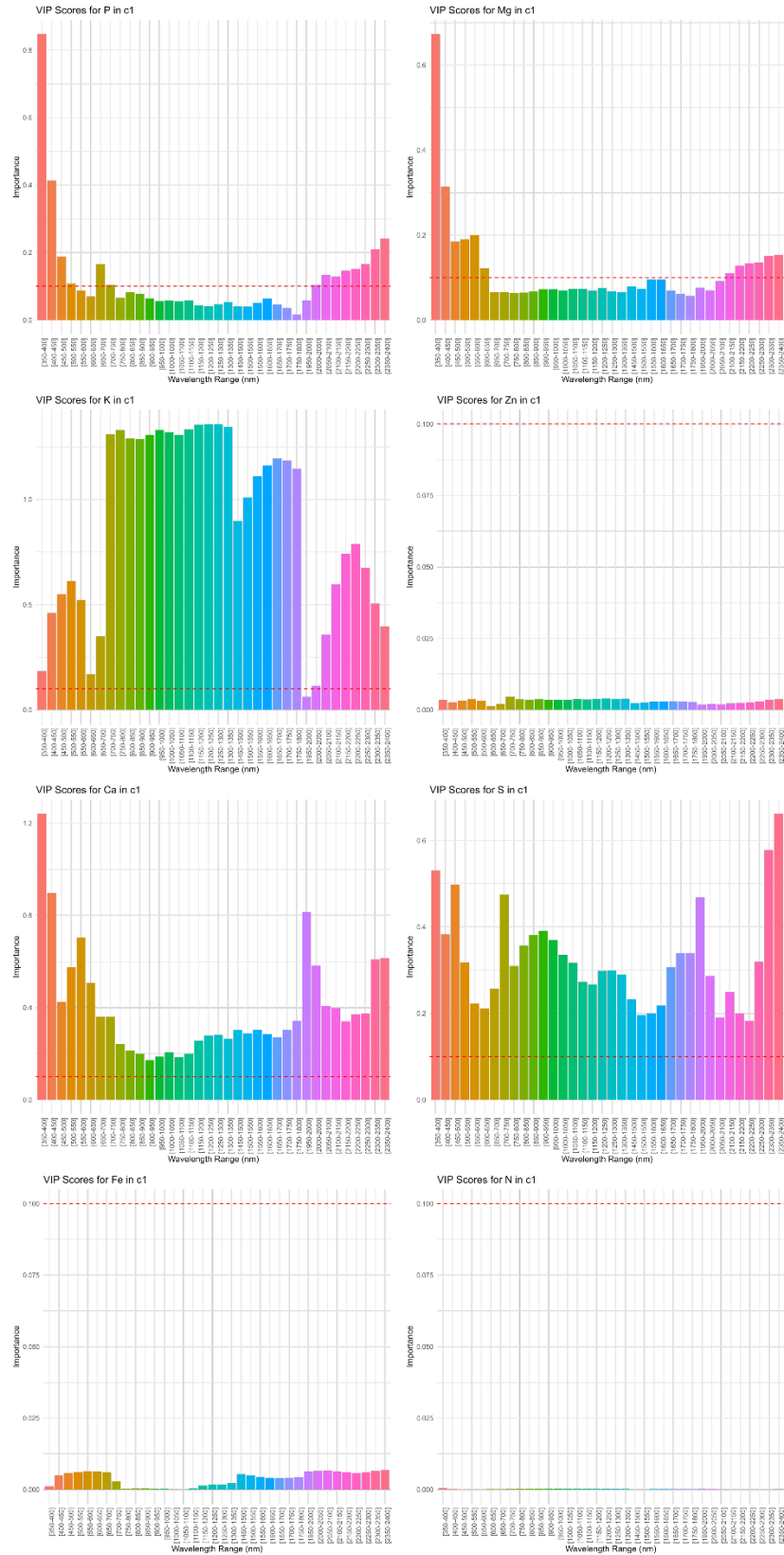
(a) Biplots of PC1 and PC2 at vegetative stage
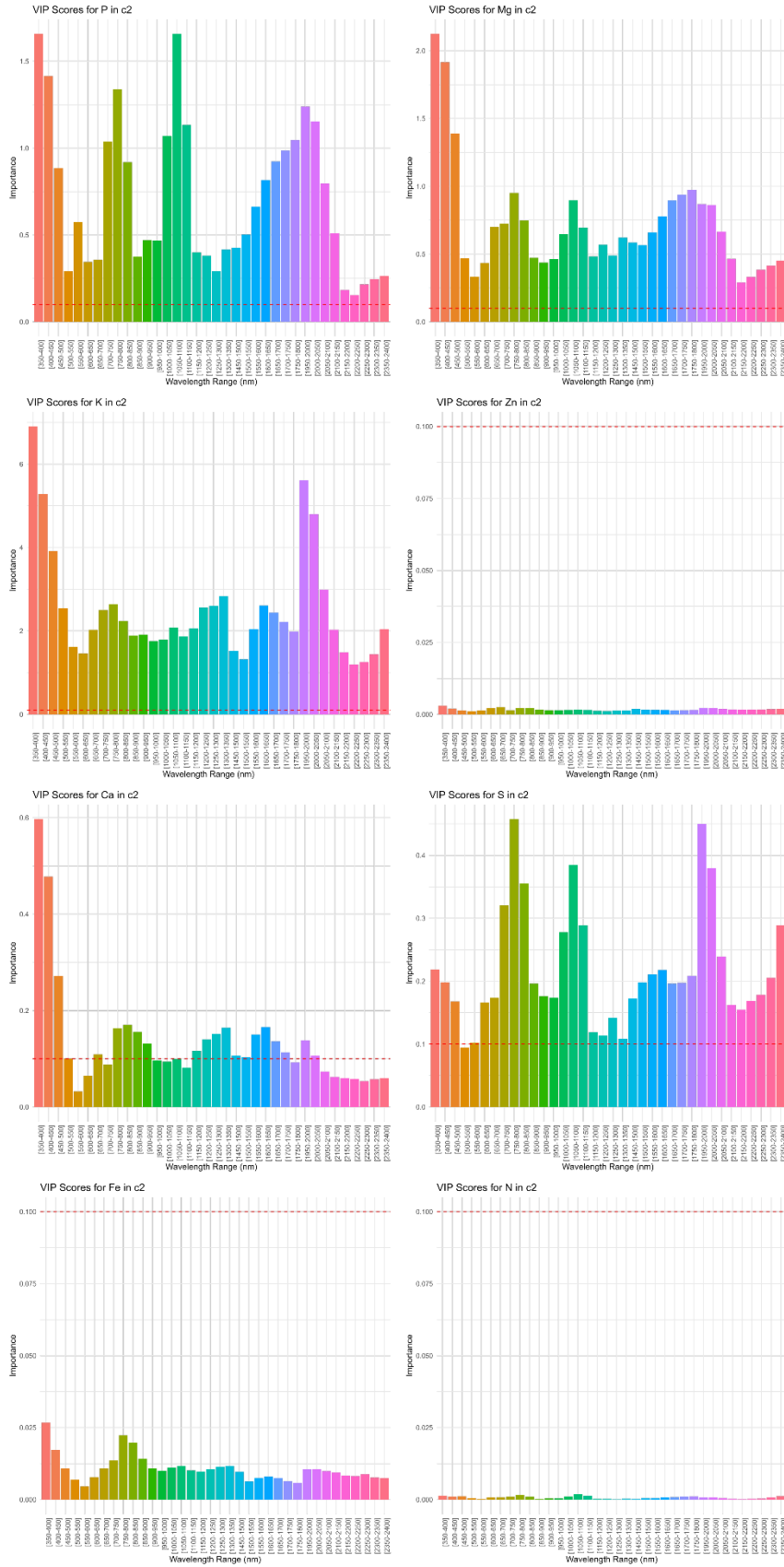
(b) Biplots of PC1 and PC2 at reproductive stage

Appendix E: VIP plots for vegetative and reproductive stages

(a) VIP during vegetative stage

(b) VIP for reproductive stage

Appendix F: Ethical considerations, risks and contingencies

No ethical issues are present as the data collection was conducted within the EO4Nutri project, with the Bonifiche Ferraresi farm, a project partner, providing consent for data collection

Appendix G: Data management plan

Data Collection	The data used in the study were collected from 40 Experimental Sampling Units in Italy, including biophysical and biochemical properties, ground spectral measurements, and laboratory-measured nutrient concentrations during summer 2023 campaign. MSc supervisors provided pre-modified PLSR R script.
Data Format	Ground spectra and laboratory nutrient measurements were provided in comma-separated values (CSV) format.
Data Pre-processing	Data quality checks included completeness verification, duplicate removal, and handling missing data. Further processing involved spline filter smoothing to prepare ready inputs for Partial Least Squares Regression models. Some spectral bands were omitted to exclude interferences.
Data Storage	A structured collection of data, including research output and R scripts are submitted to the University of Twente through the link \\ad.utwente.nl\itc\Archive\CourseData\Upload\s3009211 with appropriate documentation and metadata sheet to allow anyone else capable of repeating my research be able to reproduce and confirm my results and draw the same conclusions.
Data Access and Sharing	- Access to the data will be controlled and granted to relevant persons(s) upon approval by the University of Twente. - Data will be made available upon request under a data-sharing agreement ensuring proper citation and acknowledgment (where necessary).
Data Preservation	Long-term preservation will be ensured by storing data in recognized digital repositories with appropriate metadata to facilitate future use and discovery.
Responsibilities	Data management responsibilities lie with me as the student and supervisory team, ensuring adherence to data management policies and procedures.