Detecting and mapping forest nutrient deficiencies: Eucalyptus variety *(Eucalyptus grandis x and Eucalyptus urophylla)* trees in KwaZulu-Natal, South Africa

by
Leeth Singh

January 2022

UNIVERSITY OF KWAZULU-NATAL
INYUVESI YAKWAZULU-NATALI

Detecting and mapping forest nutrient deficiencies: Eucalyptus variety ***(Eucalyptus grandis x and Eucalyptus urophylla)*** trees in KwaZulu-Natal, South Africa

by

LEETH SINGH

Submitted in fulfilment of the academic requirements of

Doctor of Philosophy

in College of Agriculture, Engineering & Science

School of Agricultural, Earth and Environmental Sciences

University of KwaZulu-Natal

Pietermaritzburg

South Africa

January 2022

PREFACE

The candidate completed the research in this thesis while based in the Discipline of Geography, School of Agricultural, Earth and Environmental Sciences of the College of Agriculture, Engineering and Science, University of KwaZulu-Natal Pietermaritzburg, South Africa. The research was financially supported by the National Research Foundation South Africa.

The submission of this thesis has not been to another university and the results reported are solely investigations by the candidate.

— [redacted] —

Signed: Professor Onisimo Mutanga

Date: January 2022

DECLARATION 1: PLAGIARISM

Leeth Singh declares that:

(i) the research reported in this dissertation, except where otherwise indicated or acknowledged, is my original work.

(ii) this dissertation has not been submitted in whole or in part for any degree or examination to any other university.

(iii) this dissertation does not contain other persons' data, pictures, graphs, or other information unless expressly acknowledged as being sourced from other persons.

(iv) this dissertation does not contain other persons' writing unless expressly acknowledged as being sourced from other researchers. Where other written sources have been quoted, then:

   a. their words have been re-written, but the general information attributed to them has been referenced.

   b. where their exact words have been used, their writing has been placed inside quotation marks, and referenced.

(v) where I have used material for which publications followed, I have indicated in detail my role in the work.

(vi) this dissertation is primarily a collection of material, prepared by me, published as journal articles, or presented as a poster and oral presentations at conferences. In some cases, additional material has been included.

(vii) this dissertation does not contain text, graphics or tables copied and pasted from the Internet unless expressly acknowledged, and the source is detailed in the dissertation and the References sections.

[redacted]

Signed: Leeth Singh

Date: January 2022

DECLARATION 2: PUBLICATIONS

Asterisks (*) indicate the primary and corresponding author for each paper:

1. Singh L*, Mutanga O, Mafongoya P, Peerbhay KY, Crous J 2022. Hyperspectral remote sensing for foliar nutrient detection in forestry: A near-infrared perspective. *Remote Sensing Applications: Society and Environment* 25 (2022), 1-12.
2. Singh L*, Mutanga O, Mafongoya P, Peerbhay KY, Dovey S 2021. Detecting nutrient deficiencies in *Eucalyptus grandis x and Eucalyptus urophylla* trees using hyperspectral remote sensing and random forest. *South African Journal of Geomatics* 10 (2), 207–222.
3. Singh L*, Mutanga O, Mafongoya P, Peerbhay KY, Ismail R (In preparation). A rapid diagnostic tool for detecting tree growth nutrients using infrared spectroscopy and vertical canopy positioning.
4. Singh L*, Mutanga O, Mafongoya P, Peerbhay KY, Dovey S (In preparation). Comparing unmanned aerial vehicle and PlanetScope imagery for classifying nutrient management regimes in commercial forestry plantations using a deep learning artificial neural network.
5. Singh L*, Mutanga O, Mafongoya P, Peerbhay KY, Ismail R (In preparation). Predicting forestry health indicators using high-resolution UAS imagery and a deep learning artificial neural network in South Africa.

—[redacted]———

Signed: Leeth Singh

Date: January 2022

ACKNOWLEDGEMENTS

First and foremost, I would like to thank God for granting me the knowledge, health, and support to read for a PhD. I would like to also thank my parents for their constant advice, tender love and strength that allowed me to pursue my academic endeavours. To my mom, thank you for assisting me on campus with all the administrative tasks. To my dad, thanks for motivating me to take on this journey and for being proud of me.

I would like to thank my supervisory team: Professor Onisimo Mutanga, Professor Paramu Mafongoya and Dr Kabir Peerbhay, who groomed my interests and provided invaluable guidance and support. To Prof Onisimo Mutanga, thanks for your guidance and patients with me. Also, Prof Paramu Mafongoya, thanks for your financial assistance, continued assistance, and motivation. And finally, to Dr Kabir Peerbhay, you played an integral part throughout my PhD. Thanks for your inspiration, time, and expertise during it all. I am grateful for your determination and support over my years of study and belief in my capabilities.

A tribute to Dr Riyad Ismail: I know we will all remember him even though he is no longer with us. I still see Dr Riyad Ismail as an integral role player in academic endeavours in the field of remote sensing. He inspired me to study remote sensing further and always encouraged me to work harder. I dedicate this thesis to him.

I cannot forget special people like my fellow friends and colleagues, Dr Romano Lottering, Shorba Harkhu, Mrs Shani Ramroop, Mr Donavan DeVos, Dr Khatab Abdalla, Mr Agostinho, Mrs Precious Nciliba, Ms Karusha Naidoo and all my student buddies throughout my journey as a PhD student. Most importantly, to a very extraordinary human being in my life, my wife Karen, I extend my gratitude for your steadfast love, understanding and support of my research endeavours. Thank you for pushing all the way until the end.

Finally, I would like to thank my funders as this work is based on the research supported wholly by the Department of Science and Technology (DST) and the National Research Foundation of South Africa (NFR) (Grant Numbers: 86893 & 114898). I would like to thank the Institute of Commercial Forestry Research (ICFR), specifically Dr Richard Burgendorf, for their field and laboratory work technical support. I would like to thank senior scientists Mr Nkosinathi Kaptein and Dr Steven Dovey from the ICFR and Sappi Forests. Their consultation is gratefully acknowledged for understanding the materials and methods utilized in this study. I would like to thank the reviewers for their comments that enhanced this study.

TABLE OF CONTENTS

Page

ABSTRACT

Nutrient deficiencies in commercial forestry environments stunt plant growth and reduce survival, resulting in a loss of time, resources, and trees that can become more susceptible to a host of infections. Ineffective and inefficient nutrient screening methods could lead to the release of unhealthy trees for in-field planting, wasting functional space and inevitably impeding forest production. Therefore, the early detection and continuous monitoring of nutrient deficiencies are essential to support management decisions for an effective nutrient management regime. This study aimed to develop and explore innovative detection techniques to map nutrient deficiencies in commercial forest plantations.

In the first part of this thesis, the focus is on reviewing existing literature on mapping nutrient deficiencies using operational near-infrared (NIR), remotely sensed data. The review provides a synopsis of the application of near-infrared spectroscopy (NIRS) for detecting foliar nutrients, focusing on the best spectral noise removal methods, data pre-processes, and statistical models. The primary outcomes suggest that NIRS provide reasonably accurate results by utilising carefully selected pre-processing data methods and statistical models that reduce spectral noise.

In the second part of this study, the focus was on developing a forest nursery experiment to test the capability of remote sensing to detect macronutrient and micronutrient deficiencies rapidly using a non-destructive approach. This study entailed creating a pot trial experiment to acquire full-waveform hyperspectral data (350nm-2500nm) from 135 young trees in a controlled forestry nursery environment. This study quantified nitrogen (N), phosphorus (P), potassium (K), calcium (Ca), magnesium (Mg), sodium (Na), manganese (Mn), iron (Fe), copper (Cu), zinc (Zn), and boron (B) in a commonly planted commercial hybrid variety. Utilizing the robustness of the random forest (RF) algorithm, N and P produced R^2 of 0.95 and 0.89, respectively, and for micronutrients such as Mn and Cu produced R^2's of 0.90 and 0.86, respectively. This study identified the most effective regions (red-edge, NIR, visible (VIS) and short-wave infrared-2 (SWIR-2) for detecting macronutrients and micronutrients regions in this study.

These positive results prompted the need to understand the distribution of nutrient content across the four vertical canopy positions (VCP) (Quartiles 1-4) and develop a rapid diagnostic tool for accurate nutrient detection using an in-field handheld NIRS device. As a result, quartiles two and four were the best positions to take a measurement for detecting

macronutrients, and micronutrients when sampling using NIRS and the partial least squares (PLS) algorithm.

In the final section of this study, the focus was on upscaling the findings from the first part of the thesis. In the final section the focus was to test the capabilities of unmanned aerial system (UAS) imagery using a very high resolution Micasense sensor and satellite imagery (PlanetScope) in conjunction with an ANN to classify four nutrient regimes in live standing forestry compartment. Both images successfully classified the four nutrient regimes with an overall accuracy (OA) above 80%, with Kappa coefficient (KHAT) above 75 using four hidden layers and 30 epochs. Chapter 5 found that the UAS imagery performed slightly better than the satellite imagery; however, they were both seamlessly accurate. Lastly, the outcome of chapter 5 was tested in chapter 6. Hence, in chapter 6 tested very high-resolution UAS imagery to predict macronutrients and micronutrients using a deep learning ANN in a compartment forest. Variable importance measures provided helpful information in the prediction model. The utilization of the NIR and red-edge wavebands highly contributed to the prediction model with R^2's for various nutrients ranging between 0.14 and 0.75.

Overall, this study advocates for the potential use of advanced remote sensing technology to detect and map nutrient deficiencies in commercial forestry environments, at nursery and compartment levels. The results from this study provide an alternative nutrient screening framework for the commercial forestry industry that require quality planting material for long- and short-term resource sustainability on a large scale.

LIST OF TABLES

Table **Page**

No table of figures entries found.

LIST OF FIGURES

Figure **Page**

LIST OF ACRONYMS

%N	Percentage Of Nitrogen
2RF	Rehabilitation
AI	Artificial Intelligence
AIC	Akaike Information Criterion
AIS	Airborne Imaging Spectrometer
ANN	Artificial Neural Network
AOI	Area Of Interest
ASD	Analytical Spectrometer Device
ASL	Above Sea Level
AVIRIS	Airborne Visible/Infrared Imaging Spectrometer
B	Boron
BD	Band Depth
BDR	Band Depth Ratio
BRDF	Bidirectional Reflectance Distribution Function
BRF	Bidirectional Reflectance Factor
C	Carbon
Ca	Calcium
CAO	Carnegie Airborne Observatory
CARS	Competitive Adaptive Reweighted Sampling
CART	Classification And Regression Trees
C-H	Carbon-Hydrogen
Chl	Chlorophyll
CIg	Chlorophyll Index Green
CNS	Carbon Nanostructures
COP	Conference Of the Parties
CRDR	Continuum-Removed Derivative Reflectance
Cu	Copper
CVA	Cross-Validation Accuracy
DA	Discriminative Analysis
DER	Derivative
DST	Department Of Science and Technology
DT	Decision Tree
EWT	Equivalent Water Thickness
FAO	Food And Agriculture Organization
Fe	Iron
FERT	Nutrient Replacement
FOS	Forest Observation System
FOV	Field Of View
FSF	Field Spectroscopy Facility
FT-IR	Fourier Transform Infrared
FT-NIR	Fourier Transform-Near Infrared
GCP	Ground Control Point
GNDVI	Green Normalized Difference Vegetation Index

GPP	Gross Primary Production
GPS	Global Positioning System
GSD	Ground Sample Distance
HIRIS	High-Resolution Imaging Spectrometer
ICFR	Institute Of Commercial Forestry Research
IPP	Inflexion Point Position
IR	Infrared
ISI	International Scientific Indexing
K	Potassium
KHAT	Kappa Coefficient
KNN	K-Nearest Neighbour
KNP	Kruger National Park
LAI	Leaf Area Index
LC	Lack Of Correlation
LDM	Leaf Dry Mass
LIDAR	Light Detection and Ranging
LL	Lambda-Lambda
LMA	Leaf Mass Per Area
LNC	Leaf Nitrogen Concentration
LOOCV	Leave One - Out Cross Validation
LR	Linear Regression
LS-SVM	Least-Squares Support Vector Machines
MAE	Mean Absolute Error
MAP	Mean Annual Precipitation
MAT	Mean Annual Temperature
MAX	Maximum
MCCV	Monte Carlo Cross-Validation
Mg	Magnesium
MIN	Minimum
MLP	Multi-Layer Perceptron
Mn	Manganese
MNDI	Normalized Difference Vegetation Index
MNF	Minimum Noise Fraction
MSC	Multiplicative Scatter Correction
MSE	Mean Squared Error
N	Nitrogen
Na	Sodium
NBDI	Normalised Band Depth Index
NDRE	Normalized Difference Red Edge
NDVI	Normalized Difference Vegetation Index
NERC	Natural Environment Research Council
NGRDI	Normalized Green-Red Difference Index
N-H	Nitrogen-Hydrogen
NIR	Near-Infrared
NIRS	Near-Infrared Spectroscopy
NIRV	Near-Infrared Variation

NO3	Nitrate
NPK	Nitrogen Phosphorus Potassium
NRF	National Research Foundation of South Africa
NU	Nonunity Slope
OA	Overall Accuracy
O-H	Oxygen-Hydrogen
OOB	Out-Of-Bag
ORF	Oblique Random Forest
P	Phosphorus
PCA	Principal Component Analysis
PLS	Partial Least Squares
PLS-DA	Partial Least Squares-Discriminative Analysis
PLSR	Partial Least Squares Regression
PPM	Parts Per Million
REDD+	Reducing Emissions from Deforestation and Forest Degra
REM	Nutrient Removal
reNDVI	Normalized Difference Red-Edge
R2	Coefficient Of Determination
RET	Residue Retention
RF	Random Forest
RGB	Red-Green-Blue
RMS	Root Means Square
RMSE	Root Mean Square Errors
RMSEC	Root Mean Square Error of Calibration
RMSECAL	Calculated Root Mean Square Error
RMSECV	Root Mean Square Error Cross Validation
RMSEP	Root Mean Square Error of Prediction
RPD	Residual Prediction Deviation
RPDCC	Ratio Of Prediction to The Deviation of Cross-Validation
RR	Ridge Regression
S/N	Signal-To-Noise
SAS	Statistical Analysis System
SAVI	Soil Adjusted Vegetation Index
SB	Squared Bias
SCOPE	Soil Canopy Observation, Photochemistry and Energy
SD	Standard Deviation
SEC	Standard Error of Calibration
SEP	Standard Error of Prevision
SG	Savitzky-Golay
SIF	Solar-Induced Fluorescence
SNR	Signal-To-Noise Ratio
SNV	Standard Normal Variate
SSC	Solid Soluble Content
SVM	Support Vector Machines
SVMP	Support Vector Machine-Polynomial

SVMR	Support Vector Machines Regression
SVM-RFE	Support Vector Machine-Recursive Feature Elimination
SWIR-1	Short-Wave Infrared-1
SWIR-2	Short-Wave Infrared-2
TA	Training Accuracy
UAV	Unmanned Aerial Vehicle
UTM	Universal Transverse Mercator
UV	Ultraviolet
VCP	Vertical Canopy Positioning
VI	Variable Importance
VIS	Visible
VIS-NIR	Visible–Near-Infrared
VN	Vector Normalization
Zn	Zinc

CHAPTER 1: General introduction

1.1 Introduction

The global commercial forestry industry is amongst the most critical contributors towards economic growth, generating approximately 600 billion US dollars per year from forest products and supports combatting the effects of food security; climate change; greenhouse gas reduction and creating green jobs (Boadi *et al.*, 2014; FAO, 2016; Peña-Lévano *et al.*, 2019). Hence, producing high quality and quantity trees is imperative for sustaining the industry's growth and solving many auxiliary global impacts (Cipullo *et al.*, 2019). The most significant factors affecting tree production and health are water and nutrients such as N, P and K (Porras-Soriano *et al.*, 2009; Silva & Uchida, 2000). However, determining these factors become a challenge over a large scale and across many heterogenous environments. Hence, research efforts need to focus on acquiring advanced technology for rapidly assessing forest health parameters to maximize productivity and improve health in the future (Porras-Soriano *et al.*, 2009).

Recently, the inception of remote sensing technology as a valuable resource and tool for detecting and mapping discrete foliar biochemicals over large scale has made many breakthroughs (Watt *et al.*, 2019). More specifically, previous research shows that both hyperspectral and multispectral remote sensing capabilities can detect and map foliar biochemicals such as Nitrogen, Phosphorous, Potassium (NPK) with relatively good accuracy (Xulu *et al.*, 2019). This provides beneficial solutions for overcoming many commercial forestry challenges such as accurately estimating yield and forest health parameters (Rubilar *et al.*, 2018). Furthermore, the planting of nutrient deficient trees decrease the economic performance of many forestry companies. Therefore, a rapid approach to monitoring the nutrient status of trees would benefit the commercial forestry industry by saving time, and economic resources within the commercial forestry value chain.

Commercial forestry nurseries are high-throughput environments, producing hundreds to thousands of saplings for in-field planting, using standard nutrient regimes that focus on initial plant rooting and growth. Unproductive trees eventually increase pressure on the economic feasibility of forestry systems and nutrient cycles of the land. Hence, the optimal supply of nutrients plays a critical role in plant survival, growth, and overall productivity (Silva & Uchida, 2000). Conventional nutrient assessments are performed using wet chemistry laboratory-based techniques (Köhl *et al.*, 2006). Whilst conventional nutrient assessments provide highly accurate results; they become inefficient and ineffective when the demand for high-quality timber is imperative for the survival of many commercial forestry companies today. Hence, trees need to reach optimal levels of nutrient retention for their survival and rapid growth. However, maintaining sufficient levels of nutrient retention becomes problematic at a large scale, especially when traditional nutrient assessments are labour intensive and timely. Furthermore, the timeous assessment of nutrients is critical for avoiding nutrient deficiency or nutrient toxicity (Dordas, 2008). As a result, nutrient-deficient or toxic nutrient trees become susceptible to pests and diseases, which is usually only detected after in-field planting. This has a devasting effect on the performance of many commercial forestry companies that value time as an indicator of growth.

Advancements in remote sensing to successfully detect and map key foliar biochemical information, mainly nutrient information, has been demonstrated across various applications using various statistical modelling techniques and imaging platforms. These applications have shown to be more efficient and effective in generating nutrient assessments by reducing the time and labour of acquiring a sample over a large area. Hyperspectral and multispectral imaging are two different types of remote sensing which differ in their spectral resolution. Generally, hyperspectral data contain narrower wavebands (high spectral resolution) whereas multispectral data contain wider wavebands (low spectral resolution) within the electromagnetic spectrum (Bioucas-Dias *et al.*, 2013). Currently, there are three systems/platforms namely handheld/proximal, aerial and satellite which hyperspectral and multispectral data can be acquired. A handheld/proximal system is a small portable machine for example a analytical spectrometer device used mainly for laboratory experiments, an aerial system are aircrafts and UAS's for example an aeroplane or drone used for larger scale (AOI's >5 ha) experiments and a satellite system are orbital satellites for example NASA's Landsat imagery used for large to global scale experiments (Bioucas-Dias *et al.*, 2013). Many studies have explored the capabilities of hyperspectral data, which produces contiguous waveband information for the detection of discrete nutrient information such as NPK (Abdel-Rahman *et al.*, 2017; Amirruddin *et al.*, 2020; Ansari *et al.*, 2016; Axelsson *et al.*, 2013; Eshkabilov *et al.*, 2021; Knyazikhin *et al.*, 2013; Li *et al.*, 2018; Mahajan *et al.*, 2014; Meacham-Hensold *et al.*, 2019; Mutanga *et al.*, 2004b; Oliveira & Santana, 2020; Peterson *et al.*, 1988; Pullanagari *et al.*, 2018; Singh *et al.*, 2017a; Wang *et al.*, 2017; Yu *et al.*, 2020). In addition, multispectral studies have shown benefits for large scale mapping, especially in many heterogeneous environments (Cai *et al.*, 2019; Chemura *et al.*, 2018; L. Chen *et al.*, 2019; Dash *et al.*, 2018; Gara *et al.*, 2018; Kokaly & Clark, 1999; Osco *et al.*, 2020a; Singh *et al.*, 2017b; Walshe *et al.*, 2020). Advancements in remote sensing technologies allow for more accurate macronutrient and micronutrient deficiencies using hyperspectral proximal and multispectral imaging techniques combined with highly sophisticated algorithmic modelling techniques. A combination of advanced remote sensing platforms and data modelling techniques make it possible to detect and map discrete nutrient information using non-destructive approaches.

Currently, the South African forestry industry is highly sanctioned by climate and economic policies (adoption of sustainable development goals). Increased restrictions on access to arable land and water put pressure on the sustainability of the commercial forestry industry (McEwan & Steenkamp, 2014). However, many commercial forestry organizations still use past knowledge and practices to inform current day decision making. Hence, the survival of the South African forestry industry hinges on the development of novel, innovative, and modernized solutions to improve residue management, site preparation, establishment, coppice management, weed control and fire protection activities using science and technology (McEwan & Steenkamp, 2014). Reducing costs on fertilizer use and identifying nutrient deficient trees play a crucial role in the sustainability of the commercial forestry industry. Furthermore, improvement in the efficacy of remote sensing technologies can provide many foresters, forestry operation managers, forestry organizations executives, laboratory technicians, and forestry researchers with readily available nutrient information for proactive decision making. Decision-makers can use this information to inform the industry's plant health and economic performance, such as forest mensuration, yield estimates, cost efficiency, productivity, and returns on investment within the forestry industry. Hence, the adoption of remote sensing technologies has the capabilities of providing forestry industries with timely nutrient assessments to improve plant survival.

1.2 Aims and objectives

This research aims to investigate the potential use of remote sensing to accurately detect and map foliar nutrient deficiencies occurring within commercial forestry environments in KwaZulu-Natal, South Africa.

The objectives of the research are:

1. To provide a synopsis of the application of remote sensing for detecting foliar nutrients.

2. To investigate the ability of remote sensing to rapidly detect nutrient deficiencies of saplings in a nursery environment using hyperspectral data.

3. To investigate the influence of VCP in improving the detection accuracy of nutrients of saplings in a nursery environment using NIRS.

4. To compare the classification accuracy of very-high-resolution UAS imagery and high-resolution PlanetScope imagery in four field nutrient management regimes using a deep learning ANN.

5. To provide a framework for predicating nutrients in a compartment forest using very-high-resolution imagery and a deep learning ANN.

1.3 Outline of thesis

This thesis is presented as a set of research papers to address each objective outlined in section 1.2 above. Each research paper is primarily self-contained with the following items: an introduction, materials and methods, results and discussion, and a conclusion. Two chapters (2 & 3) have been published as research papers in International Scientific Indexing (ISI) journals, while the remaining three chapters (4, 5 & 6) are in preparation. Including the introduction and synthesis, this thesis consists of seven chapters. Figure 1.1 shows a graphical interpretation of the thesis. The thesis is divided into two parts:

In the first part, this thesis reviews existing literature on mapping nutrient deficiencies. The literature review chapter provides a synopsis of the application of NIRS for detecting foliar nutrients, focusing on spectral noise, pre-processing data methods, and statistical models. The primary outcomes suggest that by carefully selecting pre-processing data methods and statistical models that reduce spectral noise, NIRS provide reasonably accurate results. The critical findings from the literature review were translated into research papers in chapters 3 and 4 of this thesis. These chapters have a specific focus on investigating the capabilities of remote sensing to rapidly detect nutrient deficiencies of saplings in a nursery environment using hyperspectral and NIR data. A pot trial experiment was implemented to induce nutrient deficiencies in a nursery environment for N, P, K, Ca, Mg, Na, Mn, Fe, Cu, Zn, and B in young commercially planted forest variety. Pre-processing techniques were used to understand the distribution of nutrient content across the four VCPs.

In the second part of the thesis, the capabilities of airborne and satellite remote sensing were tested for detecting and classifying nutrient information in a heterogeneous compartment forest environment using multispectral data. These chapters have a specific focus on investigating the capabilities of remote sensing to rapidly detect nutrient deficiencies of saplings in a field environment using very high-resolution multispectral data. These chapters tested the use of UAS data and satellite imagery in conjunction with deep learning approaches to predict and classify nutrient concentrations. Very-high-resolution 12-bit multispectral Red Edge-MX Micasense and PS2-PlanetScope imagery provided closely matched waveband competencies and spatial resolutions of 8cm and 3.7m, respectively. These chapters compared the capabilities of both platforms to predict and classify four nutrient management regimes (residue retention (RET), nutrient removal (REM), nutrient replacement (FERT), and rehabilitation (2RF)) in a commercial compartment forest. These chapters used a deep learning (ANN) approach as the optimal strategy for detecting and classifying nutrient data.

Finally, a synthesis is presented in chapter 7. The synthesis is a formal summary of all the findings and conclusions from the preceding chapters.

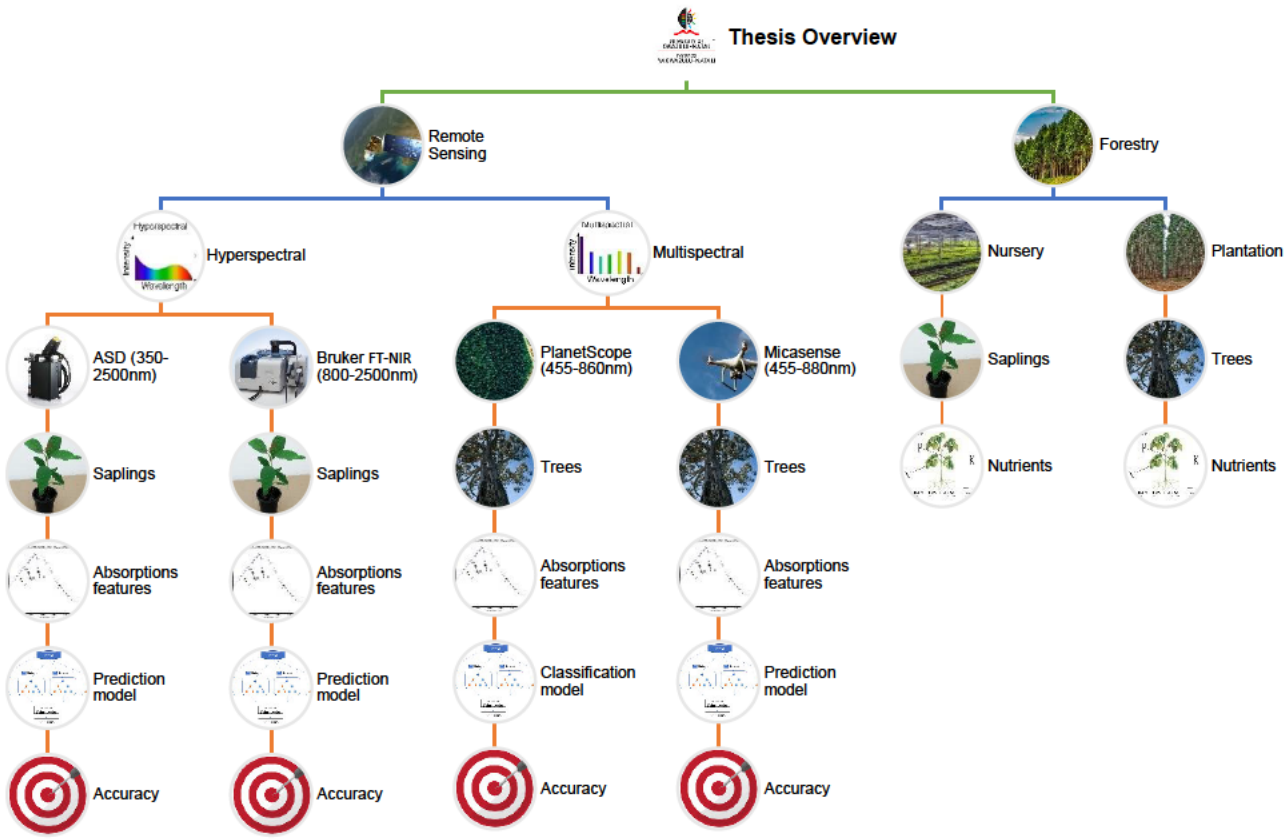


Figure 1. 1: A graphical interpretation of the thesis showing the methodological flow of thesis.

CHAPTER 2: Literature review: Remote sensing for foliar nutrient detection in forestry: A near-infrared perspective

A synopsis of the application of hyperspectral remote sensing for detecting foliar nutrients

This chapter is based on:

Singh L*, Mutanga O, Mafongoya P, Peerbhay KY, Crous J 2022. Hyperspectral remote sensing for foliar nutrient detection in Forestry: A near-infrared perspective. *Remote Sensing Applications: Society and Environment* 25 (2022), 1-12.

Abstract

Over the past decade, hyperspectral remote sensing as a rapid, non-destructive technique for vegetation assessment has considerably contributed to the efficacy of remote sensing. This review paper provides a synopsis of the application of hyperspectral remote sensing for detecting foliar nutrients. The focus was to review spectral noise, pre-processing data methods, and NIR technology statistical models. This chapter used an integrative approach to critically analyse a decade (2010-2020) of research. The primary outcomes suggest that NIR technology provides reasonably accurate results by utilising strategically selected data pre-processing methods and statistical models that reduce spectral noise. Sample sizes, latent variables, and leaf water content were the main factors determining successful outcomes. The constraints presented motivate future research to understand the effects of epicuticle wax and trichomes on leaf optical properties.

Keywords: hyperspectral, spectral noise, data pre-processing, statistical model, chemical compound, variable selection, foliar nutrients

2.1 Introduction

The intensive management of forest plantations has significantly evolved to meet the global supply and demand for forest products (Rubilar *et al.*, 2018). Advances in our understanding of silviculture practices combined with the progression of information technologies have revolutionized the forestry industry (Rubilar *et al.*, 2018). Trees are given a strategic supply of nutrients at the nursery level to build up nutrient reserves for subsequent in-field planting for maximizing rooting and shoot production (Timmer, 1997). However, traditional nutrient assessment methods become inadequate to meet high demands when assessing large sample sizes. Operationally, the intensification of labour to meet this demand further exacerbates costs, deeming the process forfeited (Payn *et al.*, 1999). An alternative is to reduce sample sizes and extrapolate nutrient assessments, saving initial costs. However, not all trees are homogeneous; hence nutrient concentrations are irregular. An ineffective nutrient regime could leave plants exposed to the intrusion of pests and diseases, and overfertilization leading to toxicity (Payn *et al.*, 1999).

Nutrient deficiencies plague plant production, significantly reducing industrial plantation output and productivity. Accurately diagnosing nutrient deficiencies at early stages of growth will substantially increase in-field survival on highly valued plantations (Mee *et al.*, 2017). The timely assessment of plant nutrient status during the early stages of growth is critical for maintaining optimal nutrient levels to maximize shoot production and rooting (Turner & Lambert, 2017). Furthermore, 'hidden' nutrient deficiencies cannot be visually observed and interpreted with a human eye alone which could flaw diagnosis, disrupting remedial action for affected plants (Mee *et al.*, 2017). The current advances in analytical techniques make the detection of hidden nutrient deficiencies more possible. The introduction of remotely sensed data, specifically NIRS, can provide almost instantaneous results for large-scale precision silviculture practice (Rubilar *et al.*, 2018; Watt *et al.*, 2019). NIRS is a type of high-energy vibrational spectroscopy which sensors in the wavelength range of 750-2500 nm (Pasquini, 2018). However, the impact of spectral noise in highly dimensional data continues to implicate the development of efficient nutrient assessments. A recent review paper by Watt *et al.* (2019) concludes that early research has demonstrated the potential of remotely sensed data, particularly the NIR region, for diagnosing nutrient deficiencies; however, little research has used these models for this purpose.

Three main objectives outline this review. The three main objectives provide the reader with (1) hyperspectral remote sensing (NIRS) and its relation to plant physiological traits, (2) the impact of spectral noise in high dimensional data and (3) strategic data pre-processing and statistical modelling techniques. Also, this chapter presented variable selection methods and accuracy assessment used by most studies in this review. To accomplish the objectives in this review, this chapter critically examined a decade of research that used hyperspectral remote sensing (NIRS) to detect foliar nutrient deficiencies in a forestry environment. This chapter unpacked peer-reviewed scientific research articles between 2010 and 2020 using google scholars advanced search configurations in the 'incognito mode. Using the 'incognito mode' prevents skewing the data to previous searched articles outside the scope of this article. This chapter used the following search configurations: (1) 'with all the words' = "near-infrared" & "forestry" & "tree" & "foliar" & "spectroscopy" & "nutrient" and "deficiency" & "remote sensing"; (2) 'where my words occur' = 'anywhere in the article'; and (3) 'return articles dated between' = '2010'-'2020'. The search results gathered approximately 61 results of journal articles on 31st March 2020. The structure of this review takes an integrative form similar to that of Varhola *et al.* (2010).

2.2 Remote sensing and understanding the physiological basis of tree characteristics

In the 1700s, Pierre Bouguer coined the Beer-Lambert law, which provided researchers with a physiological basis to understand the physical link between chemical compounds and electromagnetic radiation. In the 1800s, William Herschel, a musician and amateur astronomer, discovered "infrared (IR) radiation", laying the foundation for modern NIRS (Ring, 2000). William Herschel studied the relationship between VIS and invisible light rays and found simple ways of examining absorption and reflection (Ring, 2000). This relationship established the way scientists experimented with IR radiation to discover a physical link to chemical compounds in the late 1900s (Curran, 1989; Dixit & Ram, 1985; Elvidge, 1990; Peterson *et al.*, 1988; Sasaki *et al.*, 1984; Wessman *et al.*, 1989; Weyer, 1985). More specifically, the works of Karl Norris and Lois Weyer played a crucial role in developing a NIRS application for organic (molecular) sampling (Weyer, 1985; Williams & Norris, 1987).

2.2.1 Past trends

Karl Norris and Lois Weyer explored detecting organic compounds in a molecule during energy transitions (thermal radiation). These energy transitions occur during molecular bonding from the electrostatic force of attraction between oppositely charged ions (Weyer, 1985; Williams & Norris, 1987). The authors used the magnitude of these energy transitions (excitation) to distinguish between different organic compounds known as vibrational states.

Weyer (1985) described three bond vibrations, namely: (1) carbon-hydrogen (C-H); (2) nitrogen-hydrogen (N-H); and (3) oxygen-hydrogen (O-H) (Table 2.1). For example, the excitation of the C-H bond links to aliphatic, aromatic, olefins, and oxygenated compounds. Table 2.1 provides a list of chemical compounds associated with each bond vibration. Researchers then aimed to understand which regions of the electromagnetic spectrum had known absorption features for detection of specific chemical compounds, for example, N, P and K (Cael *et al.*, 1975; Curran, 1989; Elvidge, 1990; Sasaki *et al.*, 1984). Their findings consolidate many studies using NIRS today. The introduction of NIR spectrometer devices revolutionized many commercial industries, scientific institutions, and organizations, mainly for its ability to rapidly assess large quantities of data (Weyer, 1985).

Sasaki *et al.* (1984) developed methods for uniquely determining spectral curves to estimate xylene isomers and dyes in VIS and IR absorption regions. Dixit and Ram (1985) discovered that derivatives enhanced smaller peaks encapsulated by more massive peaks and the separation of overlapping wavebands. Huete (1986) suggested using a factor-analytical inversion model, which encouraged the decomposition of spectral mixtures into Eigenspectral and Eigenvector matrices. Peterson *et al.* (1988), Card *et al.* (1988), and Curran (1989) similarly investigated the use of absorption wavebands in the VIS and NIR region to predict foliar chemical concentrations in plant material.

Table 2. 1: A list of bond vibrations with their chemical compounds

Electron transition/bond vibration	Chemical compound	Absorption band
C-H	Aliphatic	1100 nm, 1250 nm, 1300 nm, 1450 nm, 1600 nm, 1800 nm, 2000 nm, 2400 nm.
	Aromatic	1143 nm, 1420-1450 nm, 1685 nm, 2150 nm, 2460 nm.
	Olefins	1180 nm, 1620 nm, 1680 nm, 2100 nm, 2150 nm.
	Oxygenated	1650 nm, 2200 nm, 2210 nm, 2150 nm, 2250 nm.
N-H	Amines	1450 nm, 1500 nm, 1530 nm, 1960 nm, 1990 nm, 2000 nm, 2020 nm.
	Amides	1500 nm, 2000 nm.
O-H	Alcohols and phenols	1000 nm, 1400-1440 nm, 2000 nm.
	Water	960 nm, 1200 nm, 1440 nm, 1450 nm, 1660 nm, 1940 nm, 1960 nm, 2000 nm.
	Silanols	1385 nm, 1900 nm, 2220 nm.

	Acids and hydroperoxides	1450 nm, 1505 nm, 1900 nm, 2070 nm, 2100 nm.

*The list of absorption bands was found in: (Curran, 1989; Dixit & Ram, 1985; Peterson *et al.*, 1988; Weyer, 1985)

Peterson *et al.* (1988) investigated the use of remote sensing for estimating biochemical content of forest leaves and canopies using a Perkin-Elmer Model 360 laboratory spectrophotometer (400-2400 nm) and an airborne imaging spectrometer (AIS) satellite image (1100-2400 nm) on three heterogeneous sites in Alaska, Wisconsin, and Oregon in North America. The authors found significant correlations with the short-wave IR region and biochemical content, specifically N and lignin, produced standard prediction errors comparable to wet chemical laboratory assessments. Similarly, Card *et al.* (1988) predicted leaf chemistry using VIS and NIRS of dry, ground leaf material from deciduous and conifer tree species in Alaska, Wisconsin, and California in North America. The authors analysed seven chemical compounds (sugar, starch, protein, cellulose, total chlorophyll (Chl), lignin, and total N using reflectance spectra acquired from a Perkin-Elmer Model 330 laboratory spectrophotometer (400-2446 nm) and stepwise regression. The authors found the VIS and NIR regions to have high correlations with the chemical compounds analysed in their study using a stepwise regression. However, insufficient sample sizes did not allow prediction for all chemicals, and they suggest the implementation of techniques for reducing instrument error.

Curran (1989) provided forty-two absorption features in the VIS (380-700 nm) and NIR (800-2500 nm) spectral regions that are related to foliar chemical concentrations (lignin, cellulose, sugar, starch, and water). The authors located foliar wavebands using computer models such as stepwise multiple regression and deconvolution processes; and AIS I and II equipped with 124 NIR wavebands; airborne VIS/IR imaging spectrometer (AVIRIS) equipped with 209 VIS and NIR wavebands; and high-resolution imaging spectrometer (HIRIS) equipped with 192 VIS and NIR wavebands. Nonetheless, this study provided three accounts of criticism of using a stepwise regression: (1) overfitting of wavebands during modelling; (2) multi-collinearity of chemicals; and (3) waveband omissions. Subsequently, the use of more strategic portions of the electromagnetic spectrum became apparent. Hence, the strategic advancements of NIR spectrometers and various remote sensing devices are continually developing to enhance their functionality, utility, and capability.

2.2.2 Current trends

Presently, there is a plethora of research that investigated the use of remote sensing for canopy chemistry and planted functional traits (Asner & Martin, 2009; Asner *et al.*, 2011; Au *et al.*, 2020; Girard *et al.*, 2020; Knyazikhin *et al.*, 2013; Lepine *et al.*, 2016; Martin *et al.*, 2018; Rodrigues *et al.*, 2020; Shi *et al.*, 2019; Stein *et al.*, 2014; Ustin, 2013; van der Meer, 2018; van der Tol *et al.*, 2019; Watt *et al.*, 2019; Windley & Foley, 2015; Zeng *et al.*, 2019; Zhang *et al.*, 2020). These studies consider remote sensing an alternative to wet chemistry assessments as it effectively reduces the time for adaptive management practices. Figure 2.1 shows a schematic workflow diagram for detecting foliar biochemicals using hyperspectral data. Many studies used hyperspectral data as the basis for their investigations. For example, earlier studies such as Zhao *et al.* (2005) explored the capabilities of hyperspectral reflectance (350 to 2500 nm) properties to determine the effects of N deficiency on sorghum growth. The authors found linear correlations with reflectance ratios of R_{405}/R_{715} (R^2 = 0.68) and R_{1075}/R_{735} (R^2 = 0.64) for Leaf N and Chl concentrations, respectively.

Zhang *et al.* (2013) investigated the potential of VIS and NIR hyperspectral imaging systems (380 to 1030 nm) for determining N, P, K in oilseed rape leaves using partial least squares regression (PLSR) and least-squares support vector machines (LS-SVM). The authors revealed that hyperspectral imaging is a promising technique for detecting macronutrients, with the PLSR and LS-SVM models predicting R^2 accuracies above 0.70. Axelsson *et al.* (2013) explored the possibilities of retrieving N, P, K, Ca, Mg, and Na in mangroves of the Berau Delta, Indonesia, using hyperspectral data (450 to 2490 nm). Their model successfully detected N with an R^2 of 0.67; however, P, K, Ca, Mg, and Na revealed slightly discouraging results (Axelsson *et al.*, 2013). Similarly, Mahajan *et al.* (2014) detected N, P, K and Sulphur in wheat (*Triticum aestivum* L.) using hyperspectral imaging (350 to 2500 nm) and eight vegetation indices (VIs). Their study reported lower R^2s of <0.42 using VIs; however, a combination of the short-wave infrared (SWIR), NIR and the VIS region was more effective in monitoring plant nutrient status.

A later study by Osco *et al.* (2020a) presented a framework based on a host of machine learning algorithms (k-Nearest Neighbour (kNN), Lasso Regression, Ridge Regression, SVM, ANN, Decision Tree (DT), and Random Forest (RF)) to predict a full range of macronutrients and micronutrients (N, P, K, Mg, S, Cu, Fe, Mn, and Zn) using a handheld hyperspectral spectrometer device (380 to 1020 nm). The authors assessed the training data using Cross-Validation and Leave-One-Out and used the Relief-F metric of the algorithms for the prediction. Using a host of algorithms, Osco *et al.* (2020a) produced higher R^2 predictions of >0.72 for all macro-and micronutrients compared to Mahajan *et al.* (2014) and (Zhang *et al.*, 2013). Finally, Eshkabilov *et al.* (2021) successfully found optimal waveband regions between 506–601 nm and 634–701 nm for detecting discrete nutrient content variables (nitrate (NO_3–), Ca^{2+}, K^+, solid soluble content (SSC), pH, and total Chl) using hyperspectral images (400 to 1000 nm) of the freshly cut lettuce leaves. The authors produced R^2s between 0.78 and 0.99 using (PLSR) and principal component analysis (PCA) techniques. With improvements in remote sensing technology and research knowledge, more studies have found correlations with specific regions of the electromagnetic region.

Many studies found the NIR region of the electromagnetic spectrum capable and reliable for further investigation. For example, Windley and Foley (2015) measured foliar concentrations of total N, in vitro dry matter digestibility, and available N of a multi-species dataset of New Zealand trees. The authors found NIRS robust for measuring nutritional traits with R^2's ranging from 0.83-0.99 using modified-PLSR. Zeng *et al.* (2019) found the NIR spectral region resilient against soil background contamination, allowing for the robust calculation of Solar-induced Chl fluorescence (SIF). The authors estimated the fraction of total emitted NIR SIF (760 nm) photons that escape the canopy by combining the NIR reflectance of vegetation (NIR_V) and the fraction of absorbed photosynthetically active radiation (fPAR) using a Soil Canopy Observation, Photochemistry and Energy (SCOPE) model. Their NIR_V based approach could explain variations in the escape ratio with an R^2 of 0.91, and an RMSE of 1.48% across various simulations where canopy structure, soil brightness, and sun-sensor-canopy geometry are varied.

Au *et al.* (2020) used PLSR to model the relationship between NIR spectra and the foliar concentration of two ecologically critical chemical traits, available N, and total formylated phloroglucinol compounds, using a FOSS-NIR System 6500 (400-2498 nm) of *Eucalyptus* leaves. However, their study proposed using different cross-validation techniques for model fitting and selection for testing the variation in large chemical and spectral datasets of 80 *Eucalyptus* species in eastern and southern Australia. The author's main findings were: (1) geographic location influenced the predictability of N, (2) prediction error increased when assessing samples from different locations in Australia, (3) prediction accuracy of the available N model differed little whether 300 or up to 987 calibration samples and (4) merely relying on spectral variation (assessed by Mahalanobis distance) may misinform researchers into how many reference values are required.

Furthermore, Rodrigues *et al.* (2020) evaluated the use of visible–near-infrared (VIS-NIR) spectroscopy for predicting the production of leaf dry mass (LDM), as well as macronutrients and micronutrients contents of soybean leaves grown under limestone-mining coproducts using an analytical spectrometer device (ASD) FieldSpec 3 spectroradiometer (350-2500 nm) in Tietê, São Paulo, Brazil. As a result, the authors obtained $R^2p > 0.50$ and $RPD_p > 1.50$ for the variables LDM, P, K, Mg, S, and Zn using PLSR. Also, the authors found the following waveband regions 380- 400 nm, 500-530 nm, 600-690 nm, and 700-750 nm important in their prediction model. The latest NIRS technologies provide forestry stakeholders with a rapid and non-destructive approach to assess tree health (Cipullo *et al.*, 2019). Subsequently, NIRS technology in large forestry nurseries can significantly competitively advantage.

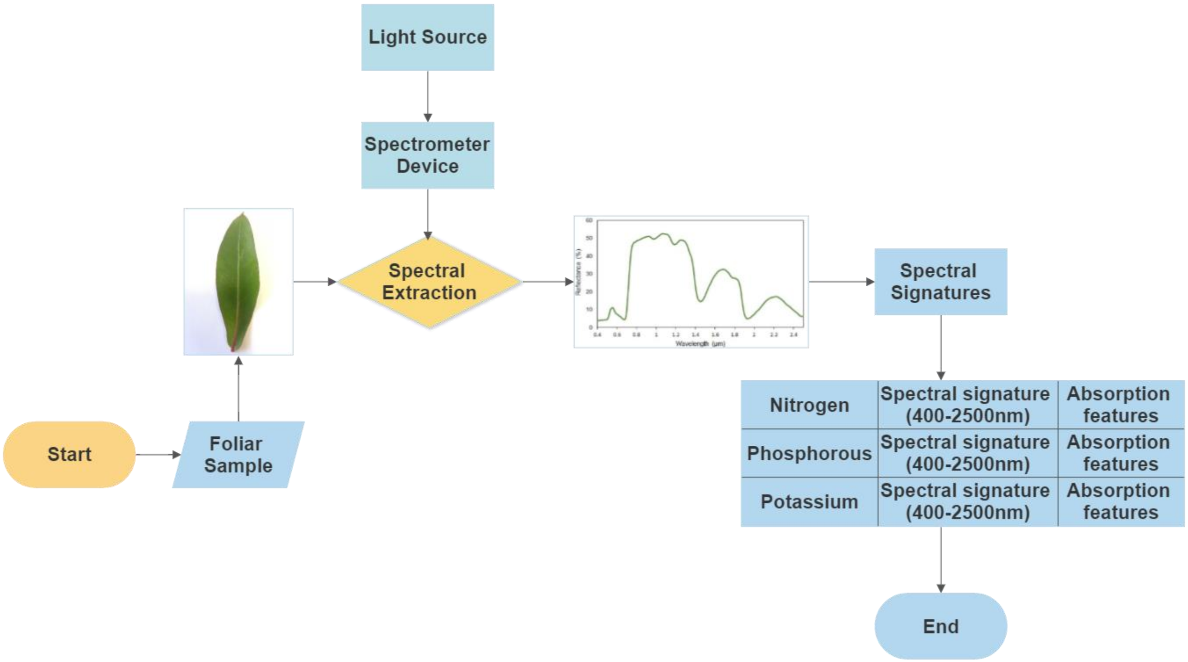


Figure 2. 1: Schematic flowchart showing the process of nutrient detection using a hyperspectral spectrometer device.

2.3 Leaf nutrient distribution

Early research has shown that trees relocate nutrients throughout the canopy leaves as a conservation mechanism; therefore, the sampling position of acquiring a representative sample is an integral part of more accurately determining nutrient content (Gara *et al.*, 2018). Furthermore, studies have reported inconsistent spectral values to sample on the adaxial (top) surface compared to the abaxial (bottom) surface of the same leaf (Lu & Lu, 2015; Warburton *et al.*, 2014). For example, Warburton *et al.* (2014) measured relative water content, leaf water potential and stomatal conductance in *Eucalyptus grandis* leaves using Thermo-Scientific microPhazir NIR spectrometer (1600-2400 nm) and PLSR in a controlled environment facility in Australia. The authors acquired spectral reflectance data from the adaxial and abaxial leaf surfaces and the upper and lower leaves in the stem. As a result, R^2's using cross-validation were R^2_{CV} = 0.85 for relative water content, R^2_{CV} = 0.74 for leaf water potential and R^2_{CV} = 0.80 for stomatal conductance. Similarly, Lu and Lu (2015) estimated leaf Chl content using ASD FieldSpec 3 portable spectrophotometer (350-2500 nm) and vegetation indices in northeast China. The authors acquired reflectance spectra from both adaxial and abaxial leaf surfaces of white poplar (*Populus alba)* and Siberian Elm (*Ulmus pumila* var. *pendula*.). As a result, spectral reflectance values were higher on the abaxial surface than adaxial surfaces in the VIS wavelengths (400-700 nm), whereas the authors found the opposite for the NIR wavelengths (700-1000 nm) for both plant species.

2.4 Spectranomics

Additionally, "Spectranomics" is a newly developing concept that explores the relationship between plant canopy species and their functional traits to their spectral-optical properties (Asner & Martin, 2009). Asner and Martin (2009) combined chemical (N, P, Chl-a, Chl-b) and spectral remote sensing (400nm-2500nm) perspectives to facilitate canopy diversity mapping. Asner *et al.* (2011) further developed this concept by examining leaf hemispherical reflectance and transmittance spectra, along with a 21-chemical portfolio, in 6136 humid tropical forest canopies. They developed up-scaling methods using a combination of canopy radiative transfer, PLSR and high-frequency noise modelling techniques using a spectral range of 400nm-2500nm. Similarly, Stein *et al.* (2014) aimed to determine the relationship between spectral reflectance (350nm–2500 nm) and foliar nutrient concentration (e.g. N, P, K, Calcium (Ca), Magnesium (Mg)) in *loblolly* pine, and to investigate the role of geographic scale in model accuracy. The authors found that localized *loblolly* pine nutrient studies are less likely to produce successful models than studies across a large geographic region. McManus *et al.* (2016) discovered the link between foliar reflectance spectra (350nm–2500 nm) and the phylogenetic composition of tropical canopy tree communities using nine biochemical traits that relate to a wide range of leaf functions. Whilst Martin *et al.* (2018) tested the concept of the foliar trait (Leaf mass per area (LMA)) retrieval and chemical data (P, Ca, K, Mg, B, Fe) using imaging spectroscopy (Carnegie Airborne Observatory (CAO)) data (350nm–2510 nm) constrained with simultaneous light detection and ranging (LiDAR) measurements.

2.5 Spectral noise & high dimensional datasets

This section delves into (a) the challenges of dealing with spectral noise in hyperspectral NIR data and (b) strategic methodologies for reducing spectral noise. Also, (c) explored the effects of moisture content and epicuticle wax on extracting a representative sample.

2.5.1 Challenges of spectral noise

Demetriades-Shah *et al.* (1990) define spectral noise as a signal of interest accompanied by background noise and other unwanted signals. Spectral noise is expressed as the Signal-to-noise (S/N) ratio between the wanted signal and the unwanted background noise as:

$$SNR = \frac{P_{signal}}{P_{noise}}$$

The suspended particles cause scattering and increase absorption by lengthening the path of the analytical beam through the sample (Demetriades-Shah *et al.*, 1990). In addition, spectral noise may occur when the sensor malfunctions; the sensor is affected by environmental constituents or the Bidirectional Reflectance Distribution Function (BDRF). A study by Knyazikhin *et al.* (2013) further exemplifies quantifying the retrieval of any biochemical information from spectral electromagnetic data is subject to leaf and canopy bidirectional reflectance factor (BRF). A conference paper by Ustin (2013) agrees with Knyazikhin *et al.* (2013) and states that future research should address these problems by quantifying the physical interactions. Hyperspectral sensors produce higher spectral noise than multispectral sensors through acquiring highly discrete spectral information (Agjee *et al.*, 2018). As a result, large continuums of data become damaged or lost (Peerbhay *et al.*, 2013).

In practice, Lepine *et al.* (2016) tested the influence of spectral resolution, spatial resolution and sensor fidelity on relationships between observed patterns of foliar percentage Nitrogen (%N) and canopy reflectance. Their study revealed almost no reduction in the strength of relationships between reflectance and %N when using coarser bandwidths from AVIRIS imagery, but instead saw declines with increasing spatial resolution and loss of sensor fidelity. Signal processing is a continuously developing field with new signal denoising techniques available across many fields such as photogrammetry, bioinformatics and remote sensing (Koziol *et al.*, 2018). The most standard denoising techniques are the Savitzky-Golay (SG) or Fourier-filtering, and the more advanced approaches are PCA and the Minimum Noise Fraction (MNF) (Koziol *et al.*, 2018).

2.5.2 Strategic denoising methodologies

As remote sensing scientists, the practice is encouraged to undertake best practice data acquisition methods to reduce noise whilst acquiring a representative sample. Hence, most of the studies in this review have used data pre-processing techniques to reduce spectral noise and normalize spectral reflectance values. For example, Asner *et al.* (2011) investigated the impact of high-frequency noise (sensor and residual artefacts noise following atmospheric correction) on PLSR predictions. They applied noise using data from AVIRIS imagery taken over tropical forests as a noise source for very SWIR simulations. The authors found that noise negatively affects PLSR results varying degrees depending on wavelength range and chemical constituent. Zhai *et al.* (2013) estimated N, P, and K contents in the leaves of different plants using laboratory-based VIS and NIRS using a FieldSpec Pro portable spectroradiometer (350-2500 nm) in Jiangsu Province, China. The authors compared regression models PLSR and SVM regression methods for estimating the N, P, and K content present in leaves of diverse plants. As a result, the support vector machines regression (SVMR) method accounted for more than 90% of N, P, and K variation compared to PLSR, which accounted for 59.1%, 50.9%, and 50.6% of the variation using, respectively.

Similarly, Amirruddin *et al.* (2017) quantified N status on various ages (maturity classes) of *Tenera* oil palm stands using a Geophysical and Environmental Research Corporation 1500 model spectroradiometer (350-1050 nm) in Malacca, Malaysia. The authors compared machine learning algorithms: Discriminative Analysis (DA) feature selection and Support Vector Machine-Recursive Feature Elimination (SVM-RFE) to determine the best spectral wavebands needed for quantifying N status. As a result, their study found that DA outperformed SVM in all maturity classes of *Tenera* oil palms. Furthermore, the authors developed spectral signatures that illustrate 'deficient N' and 'optimum N' levels using the electromagnetic spectrum (Amirruddin et al., 2017).

Koziol *et al.* (2018) investigated the spectral denoising efficiency and signal distortion properties of several spectral noise removal techniques such as Fourier transform, Mean Filter, Weighted Mean Filter, Gauss Filter, Median Filter, spatial Wavelets and Deep Neural Networks. The authors also tested spatial noise removal techniques such as SG, Fourier transforms, PCA, MNF, and spectral wavelets, using high-definition Fourier transform infrared (FT-IR) data (3900 cm^{-1} to 900 cm^{-1}) as an input. As a result, their study showed that multivariate based techniques of PCA and MNF outperformed any other spatial and spectral denoising method (Koziol *et al.*, 2018). Agjee *et al.* (2018) evaluated the influence of simulated spectral noise on RF and oblique random forest (oRF) classification performance. The authors used two node-splitting models (ridge regression (RR) and support vector machines (SVM)) to discriminate healthy and infested vegetation using hyperspectral data (350-2500 nm).

Advancements in remote sensing technology will produce more accurate sensors, enabling more precise acquisitions of remotely sensed data. Currently, there is no coherent framework for noise removal. However, the viability of using deep learning ANN to remove noise still needs further testing and assessment across many different IR spectrometers and materials (Koziol *et al.*, 2018). Hence, future research should emphasize comparing numerous case studies and scenarios to provide suitable noise removal frameworks for analysis.

2.5.3 Impact of moisture content & epicuticle wax

Another common problem when acquiring spectra is the influence of moisture content (aquaphotomics) within the leaf. Previous research has shown that water is ubiquitous in biological samples, and its effect on chemical compounds change the intensity and shifts absorption wavebands which have been a long term challenge (Kokaly & Clark, 1999; Pasquini, 2018). Kokaly and Clark (1999) illustrate the effect of moisture content on spectra obtained from a leaf when dry and exposed to 10% moisture, with 25% soil background effects and 50 m residual atmosphere. As a result, a lower spectral curve is produced, which is an inaccurate account of the actual chemical concentration of the plant. The presence of leaf glaucousness (epicuticle wax) & trichomes (presence of hairs) have a considerable impact on leaf reflectance values (Holmes & Keiller, 2002; Vanderbilt & Grant, 1985). The two studies have hypothesized that the amount of light specularly reflected by a leaf depends on plant species and is related to the canopy's physiological status and development stage (Vanderbilt & Grant, 1985). Most studies have used the ultraviolet (UV) and VIS spectral regions; an opportunity still exists to understand the effects of cuticle wax & trichomes more closely when sensed using the NIR region of the electromagnetic spectrum.

2.6 NIR data pre-processing methods & statistical modelling

This section was divided into four parts: (a) sample strategy, (b) choosing a pre-processing data method, (c) choosing an appropriate statistical model and (d) variable selection. Many remote sensing research has explored innovative data pre-processing methods to analyse reflectance data (Zhai *et al.*, 2013). The utility of spectral data pre-processing methods is an essential component for deriving a representative spectral sample. Essentially, applying a pre-processing data method has many benefits, such as reducing spectral data dimensionality, spectral noise, data redundancy and impurity, especially when employing high dimensional and multivariate data.

2.6.1 Sampling strategy

Firstly, deriving a suitable sampling strategy is the primary step to acquiring accurate results using NIR scanning systems. Sampling strategies should assimilate steps that reduce background noise and enhance the integrity of data for modelling (Atkinson & Curran, 1995; Zhu *et al.*, 2019). Samples obtained for modelling should contain high variability of the target site, providing a more stable model less vulnerable to outlier scenarios (Atkinson & Curran, 1995; Au *et al.*, 2020; Zhu *et al.*, 2019). Table 2.2 below shows a variety of studies with different sampling strategies derived based on their application. Furthermore, using a reliable wet chemistry assessment method is integral to validating the data before modelling. Therefore, researchers should emphasize the accurate execution of wet chemistry assessments.

Table 2. 2: A comparative table of various research studies who detected foliar nutrients

Author	Chemical	Plant	Sample size	Instrument	Scanning window (nm)	Wet chemistry
Ulissi *et al.* (2011)	N	Tomato leaves	15	a (portable) single-channel spectrophotometer	400-800	FlowSys, Systea, Italy
Zhai *et al.* (2013)	N, P, K	Various plants (rice, corn, sesame, soybean, tea, grass, arbour, and shrub)	95	FieldSpec Pro Portable spectroradiometer	350-2500	N: Kjeldahl method P & K: Mo–Sb colourimetry and a corning flame photometer
Afandi *et al.* (2016)	N	Rice crop	48	handheld spectroradiometer	700-1075	Kjeldahl method
Lequeue *et al.* (2016)	N	Tomato leaf powder	216	Fourier Transform-IR imaging Microscope (Hyperion 3000, Bruker Optics, Ettlingen, Germany)	350-2500	Elemental analyzer (Thermo Finnigan, San Jose, CA, USA)

Amirruddin *et al.* (2017)	N	Immature palm tree	150	1500 model spectroradiometer	350-1050	TruMec Series CNS Carbon/ Nitrogen/Sulfur analyzer instrument
Masemola and Cho (2019)	N	*Eucalyptus* trees	53	ASD spectroradiometer	350-2500	Leco FP528 N analyser
(Murguzur *et al.*, 2019)	N, P, Carbon (C)	Vascular plants	N: 552 P: 291 C: 424	FieldSpec 3, ASD Inc., Boulder, Colorado	350-2500	Colourimetric method and the CNS elemental analyzer
Guo *et al.* (2019)	N	Rubber trees (Hevea brasiliensis)	200	FieldSpec 3, ASD Inc., Boulder, Colorado	350-2500	Indophenol blue colourimetry Method and the continuous flow analyzer
Au *et al.* (2020)	Available N (N_A)	*Eucalyptus* trees	3662	Foss-NIR Systems 6500;	400-2498	N/A
(Oliveira & Santana, 2020)	N, P, K, S Ca, Mg, Mn, B, Zn, Cu, Fe	*Eucalyptus* clones	1350	CI-710 mini-spectrometer	400-900	Spectrophotometry: Ca, Mg, S, Zn, Fe, Mn; Colorimetry: P; Flame photometry: K; Kjeldahl: N

2.6.2 Pre-processing

Secondly, an important step is choosing the most appropriate spectral data pre-processing method. Section 3. b of this review highlights some essential strategies to eliminate spectral noise from hyperspectral data. The statistical model's success depends on the pre-processing data method (Schmitt *et al.*, 2014). There are many variations in spectral data pre-processing methods such as signal derivatives, vector normalization (VN), or multiplicative scatter correction (MSC). Table 2.3 below shows studies with different scanning windows, pre-processing data methods and data splitting methods used on different leaf material and scanning systems.

Table 2. 3: A comparative table of various research studies detailing pre-processing data methods for foliar nutrients

Author	Nutrient	Software	Pre-processor	Data split (%)
Ulissi *et al.* (2011)	Chlorophyll (Chl) N	MATLAB V7.0	Stavisky-Golay, Multiple Scatter Correction, Orthogonal Signal correction	85/15
Zhai *et al.* (2013)	N P K	PLSR: ParLeS v3.0 SVMR: LIBSVM toolbox	Absorbance transformation (log(1/Ref)) First derivative Light scatter and baseline correction (MSC and standard normal variate (SNV)) Detrending Wavelet Median filter Data enhancement (normalization and mean centre) First derivative transformation	70/30
Afandi *et al.* (2016)	N	N/A	None	70/30
Lequeue *et al.* (2016)	N C	Unscrambler® X software version 10.3	Smoothing (Stavisky-Golay algorithm) First derivatives	
Amirruddin *et al.* (2017)	N	DA: Statistical Analysis System (SAS) 9.4 software SVM: WEKA 3.6.9 software	DA feature selection (STEPDISC) SVM- Recursive Feature Elimination (SVM-RFE)	70/30
Masemola and Cho (2019)	N	MATLAB toolbox	Toolbox Field Spectroscopy Facility (FSF) Post Processing Toolbox First derivative Applied process MSC SNV SG smoothing convolution Embedded pre-processes Mean centring ('mean') Auto-scaling MSC plus mean centring MSC plus autoscaling	Randomized
(Murguzur *et al.*, 2019)	N P C	R-PLS package	centring, scaling, SNV,	85/15

	smoothing based on moving averages, baseline corrections and 1st and second-order SG derivatives	

2.6.3 Statistical modelling

Thirdly, a statistical model should be selected based primarily on the application of the research, used for either classification or prediction of the data used in the study. Table 2.4 provides a detailed description of studies that mainly used statistical prediction models for foliar analysis. An important step is to calibrate and test the models. Splitting data into training and test data is essential for model calibration and validation. The following studies demonstrate the application of using different splits (Curran, 1989; Curran *et al.*, 2001; Donkin *et al.*, 1993b; Guo *et al.*, 2010; Mutowo *et al.*, 2018; Pasquini, 2018). Earlier studies used standard multivariate statistical algorithms such as PLS (Menesatti *et al.*, 2010; Ulissi *et al.*, 2011) and PLSR (Zhai *et al.*, 2013). For instance, Menesatti *et al.* (2010) used PLS to make chemical determinations on citrus tree leaves detecting N, P, K, Ca, Mg, Fe, Zn, and Mn using the Vis-NIR region (310-1100 nm). Their study found relatively high correlations with R^2's ranging from 0.88 for Mg and 0.48 for P. Alternatively, PLSR is an effective method of estimating the nutrient content of plants (Zhai *et al.*, 2013). However, when Zhai *et al.* (2013) compared the model performance of PLSR and SVMR in detecting N, P, K using Vis-NIR (1000-2500 nm), they found that the PLSR model produced satisfactory results with R^2's of 0.59, 0.51, 0.51 for N, P, K, respectively. The SVMR model outperformed PLSR; as a result, the SVMR model accounted for more than 90% of the variation. Most of the studies in table 2.4 have focused on measuring N as it relates to the general health of the plant species. However, not much research has measured the entire range of macronutrients and micronutrients.

2.6.4 Variable selection

Lastly, recent studies successfully implemented variable selection. Variable selection enables scientists to test the ability of a spectral waveband to detect a feature accurately. For example, variable selection algorithms frequently detect water absorption features. Considering the extreme case of spectral data pre-processing, the method of variable selection aims to eliminate variables not contributing to improving the model's overall performance (Pasquini, 2018). Pasquini (2018) reviews and lists many different variable selection methods for improving model performance. Studies have generally shown decent to significantly accurate results when using variable selection to predict chemical properties of foliar material. Overall, variable selection seeks to: simplify models for more straightforward interpretation, shorter training times, avoiding problems of dimensionality and overfitting. For example, Mutanga *et al.* (2004b) applied the 'continuum removal on absorption features' concept to predict macronutrients N, P, K, Ca and Mg using a GER 3700 spectroradiometer (350–2500 nm) in a savanna grassland in Kruger National Park (KNP), South Africa. The authors tested four variables for estimating canopy concentrations N, P, K, Ca, and Mg: (i) continuum-removed derivative reflectance (CRDR), (ii) band depth (BD), (iii) band depth ratio (BDR) and (iv) normalized band depth index (NBDI) using Stepwise linear regression. As a result, their study produced the highest using CRDR data, which yielded R^2 values of 0.70, 0.80, 0.64, 0.50 and 0.68 with root mean square errors (RMSE) of 0.01, 0.004, 0.03, 0.01 and 0.004 for N, P, K, Ca, and Mg, respectively.

The results of their study justify the use of pre-processing data methods for successfully estimating nutrient content in dry foliar samples (Zhai *et al.*, 2013). Furthermore, the differences in prediction accuracy between PLSR and SVMR show the importance of selecting a suitable algorithm.

Table 2. 4: A comparative table of various research studies detailing statistical analysis and variable selection methods for foliar nutrient analysis.

Author	**Nutrient**	**Algorithm**	**Model calibration parameters**	**Model validation parameters**	**Latent variables**	**Results**
Ulissi *et al.* (2011)	Chlorophyll (Chl) N	Partial least squares regression (PLSR)	R^2	R^2	11	R^2 = 0.94
			RMSE	SEP		
			The standard error of prevision (SEP)	RMSE		SEP = 0.35
			Root mean square error in calibration (RMSEC)	Squared bias (SB)		

			Root mean square error in validation (RMSECV)	Nonunity slope (NU)		RMSE = 0.41
				Lack of correlation (LC)		
Zhai *et al.* (2013)	N P K	PLSR	R^2_{CV}	R^2_V	6	R^2_V = 0.66
			$RMSE_{CV}$	$RMSE_V$		$RMSE_V$ = 0.577
			Leave one - out cross validation (LOOCV)	Residual prediction deviation (RPD)		RPD = 1.75
			Akaike information criterion (AIC)			
			K-fold cross-validation using RMSE			
Afandi *et al.* (2016)	N	Artificial Neural Network (ANN)	RMSE	RMSE try-error method	11	RMSE = 0.32
			3-fold cross-validation			
Lequeue *et al.* (2016)	N C	Partial Least Squares (PLS) regression	the standard error of calibration (SEC)	R^2_C	Selected by the software	R^2_C = 0.9
			the standard error of cross-validation (SECcv)	the ratio of prediction to the deviation of cross-validation (RPD_{CV})		RPDcv = >3

			determination coefficient of calibration (R^2_C)			
			the ratio of prediction to the deviation of the calibration (RPD_C)	RPD_C		R^2_p = > 0.90
Amirruddin *et al.* (2017)	N	Discriminative analysis (DA) for classification	Training accuracy (TA %)	TA (%)	DA: 8-30 wavebands	DA: TA (%) = 87.95 CVA (%) = 87.90
		Support vector machine (SVM)	Cross-validation accuracy (CVA %)	CVA(%)	SVM: 6 wavebands	SVM: TA (%) = 94.13 CVA (%) = 81.74 (Averaged across immature, young mature and prime mature leaves)
Masemola and Cho (2019)	N	Competitive adaptive reweighted sampling (CARS) Monte Carlo Cross-Validation	R^2_{CAL} $RMSE_{CAL}$	R^2_P $RMSE_P$	6	sMC-PLS: R^2_P = 0.76 $RMSE_P$ = 0.30 PLS-CARS: R^2_P = 0.74 $RMSE_P$ = 0.25

		(MCCV)-CARS Significant multivariate correlation (sMC)-PLS				<u>PLS-MCCV-CARS:</u> R^2_P = 0.76 $RMSE_P$ = 0.14 (Obtained from fresh leaf spectra)
Murguzur *et al.* (2019)	N P C	PLS	Internal 10-fold cross-validation	R^2 RMSE CV RMSEP	N/A	<u>N:</u> R^2 = 0.94 RMSEP = 0.20 <u>P:</u> R^2 = 0.76 RMSEP = 0.05 <u>C:</u> R^2 = 0.82 RMSEP = 1.16

2.7 Summary & Discussion

This review examined a decade (2010 – 2020) of research to provide a synopsis of the past and present techniques for detecting foliar nutrients using NIRS. The best practice of this technology will provide high throughput commercial industry with a rapid and cost-effective alternative to assessing the nutrient status of their plants. Hence, future research should support implementing a NIRS system with a standardized approach to sample preparation, pre-processing, and statistical modelling.

2.7.1 Section overview

An essential part of this review was to list and compare the pre-processing data methods used from the latest research studies. It is important to note that NIR spectrometers generally produce a large amount of noise towards the end of the spectrum. There was no standard data pre-processing method from the studies presented in this review that could deal with such noise. However, most studies preferred to use the SG smoothing as the primary data pre-processing method. Furthermore, most studies used the first derivative transformation to reduce background S/N in NIR data (Lequeue *et al.*, 2016; Masemola & Cho, 2019; Murguzur *et al.*, 2019; Zhai *et al.*, 2013). To overcome the challenge of S/N problems caused by light scattering, studies here employed mainly two methods: 1. MSC and 2. SNV.

Furthermore, pre-processors such as MSC, SNV and SG into high dimensional hyperspectral data have improved prediction accuracy compared to untransformed data (Ustin & Jacquemoud, 2020). For example, Zhai *et al.* (2013) successfully employed MSC and SNV combined with the wavelet detrending method to correct light scattering variation and baseline of N, P, and K content present in leaves of diverse plants using laboratory-based VIS and near-infrared (Vis-NIR) reflectance spectroscopy. However, SNV predicted N with the highest accuracy for all the leaf spectral datasets (Masemola & Cho, 2019). More commonly, most studies listed in this review used wavelet detrending as a successful method for reducing spectral noise. Furthermore, 'mean centring', and most studies also employed 'auto-scaling; however, these two methods are typically embedded and automated in the software used. This chapter has stressed the importance of reducing spectral noise. The reduction in spectral noise has shown to significantly improve results in the studies by (Agjee *et al.*, 2018; Koziol *et al.*, 2018; Peerbhay *et al.*, 2013).

Finally, following data cleaning for noise and obscurities using pre-processing data methods, statistical models can be produced for either prediction or classification. An important part is selecting the most suitable algorithm (statistical model). For regression, most studies have used the PLSR algorithm for predicting foliar nutrients in vascular plants, tomato leaves, grasses, and various other shrubs (Cho *et al.*, 2007; Meuret *et al.*, 1993; Oliveira & Santana, 2020; Peng *et al.*, 2019). Many studies in this review found much higher correlations when using PLSR than other algorithms for predicting foliar nutrients (Abdel-Rahman *et al.*, 2017; Murguzur *et al.*, 2019; Singh *et al.*, 2015). For example, Murguzur *et al.* (2019); and Ulissi *et al.* (2011) successfully predicted (R^2 => 0.90) N levels using the PLSR algorithm. However, some studies found SVMR performed better in estimating N, P and K as SVMR have built-in noise and overfitting removal mechanisms (Amirruddin et al., 2017; Zhai et al., 2013).

For accuracy assessment, most studies used the $R^{2,}$ and root means square error (RMSE) to test the predictive ability of the models (Cho *et al.*, 2007; Mutowo *et al.*, 2018; Pasquini, 2018; Zhai *et al.*, 2013). Most studies preferred to use the root means square error cross-validation

($RMSE_{CV}$), as well as the ratio of prediction (RPD) for the goodness of fit and SEC for calibration error. These studies: Afandi *et al.* (2016), Murguzur *et al.* (2019), and Zhai *et al.* (2013) the K-fold cross-validation technique as a resampling procedure for further calibration of their models.

2.7.2 Latest research

The latest research conducted (2010-2020) shows a trend towards NIR technology in various applications and strategies. Most studies used NIR configured spectrometer devices in this review, while a few studies strategically selected the NIR region using full-spectrum hyperspectral data (350-2500 nm) for foliar analysis (Lequeue *et al.*, 2016; Masemola & Cho, 2019; Murguzur *et al.*, 2019). N is the most common plant health chemical parameter to monitor (Ustin, 2013; Windley & Foley, 2015). Hence, most of the examples in this review investigated estimating N levels within plant leaf material, whereas very few studies investigated other macronutrients (P, K, Ca, Mg, Na) and micronutrient (Mn, Fe, Cu, Zn, B). The sample sizes differed significantly from 15 to >1000 samples per study. It is important to note; these samples represented the total number of reference samples and not the spectral samples. However, researchers found that smaller sample sizes did not significantly affect the prediction results than studies with bigger sample sizes. The NIR instrumentation used by most studies were partly handheld devices and bench devices. The wet chemistry analysis performed differed across the laboratories; as a result, most nutrients did not show any significant pattern besides N using the conventional method called the 'Kjeldahl method' as the preferred method.

2.8 Recommendations

The findings of this review had a specific focus on the latest data pre-processing methods and statistical models for forest foliar nutrient assessment. This review highlighted the challenges and opportunities before model development. The leaf reflectance values affect spectral noise, moisture content; epicuticle wax; and adaxial and abaxial sampling. With this said, selecting the best data pre-processing method and statistical model is application-specific. It is vital to remain relevant with the latest research in this evolving research domain. The influence of artificial intelligence (AI) and better computing power will exceedingly enhance many of the pre-processing data methods and statistical models mentioned in this review. Essentially, the methodologies gathered in this literature review will be tested in the preceding chapters of this thesis.

Recently, studies have shown that pre-processing data methods could significantly improve results. Furthermore, collinearity and high dimensional data cause overfitting problems, especially when using sizeable contiguous data sets. Hence, before modelling, an integral part of data analysis is to employ pre-processing and variable selection methods. The studies presented in this review were limited by mainly: small sample size (Afandi *et al.*, 2016; Lequeue *et al.*, 2016); the number of latent variables used in the model (Zhai *et al.*, 2013); and leaf water content (Masemola & Cho, 2019). This review found that glaucousness and trichomes influenced spectral reflectance. Furthermore, the impact was species-dependent and related to the plant's physiological status.

This thesis recommends future research to investigate the utilization of MSC, SNV and SG as pre-processing data methods combined with PLSR and SVMR as statistical models. Future studies should investigate these statistical models using multiple validation parameters against the data produced from NIR technology and wet chemistry to reduce spectral noise. Furthermore, studies have shown that the plant's age, seasonality, and temperature affect epicuticle wax and trichome production; leaf reflectance will vary across these elements. Hence, an opportunity exists in understanding the impact of epicuticle wax and trichomes, moisture content, and the effects of sampling the adaxial and abaxial leaf surfaces across age, seasonality and temperature and heterogeneous trees with the reduction of spectral noise. The information gathered, and lessons learnt in this chapter are important for developing the strategies for detecting nutrient deficiencies in the upcoming chapters of this thesis.

CHAPTER 3: Investigating the ability of remote sensing to rapidly detect nutrient deficiencies of saplings in a nursery environment using hyperspectral data

To investigate the ability of remote sensing to rapidly detect nutrient deficiencies of saplings in a nursery environment using hyperspectral data.

This chapter was based on:

Singh L*, Mutanga O, Mafongoya P, Peerbhay KY, Dovey S 2021. Detecting nutrient deficiencies in *Eucalyptus grandis* trees using hyperspectral remote sensing and random forest. *South African Journal of Geomatics* 10 (2), 207–222.

Abstract

Nutrient deficiencies in commercial forest trees often lead to stunted growth and reduced chances of field survival, resulting in a loss of time productivity and trees that can become more susceptible to a host of infections. While conventional foliar analytical methods provide accurate results, they are not time and cost-effective in a high productivity environment. This study aims to test the capability of remote sensing to detect macronutrient and micronutrient deficiencies rapidly in juvenile trees. This study acquired full-waveform handheld/proximal hyperspectral data (350-2500nm) from 135 young trees planted in individual pots in a controlled forestry nursery environment. This study quantified N, P, K, Ca, Mg, Na, Mn, Fe, Cu, Zn, and B in young commercially planted forest variety. This study identified the most critical wavebands for detecting nutrient deficiencies using built-in RF variable importance (VI) measures. The RF algorithm's robustness significantly reduced the dataset's noise whilst producing promising results for certain macronutrients such as P and N (0.95 and 0.89, respectively) and micronutrients such as Mn and Cu (0.90 and 0.86, respectively). This study identified the red-edge, NIR, VIS and SWIR-2 regions of the electromagnetic spectrum as the most effective regions for detecting macronutrients and micronutrients in this study. This study recommends testing the use of strategic portions of the electromagnetic spectrum for reducing noise and enabling faster computing time, such as portable NIR technology.

Keywords: hyperspectral, forestry, random forest, nitrogen, foliar nutrients

3.1 Introduction

Nutrient deficient trees present a challenge to the commercial forestry sector. Underproductive trees put pressure on economic systems and nutrient cycles. The adequate supply of macronutrients and micronutrients play a crucial role in supporting plant development, positively influencing forest productivity (Silva & Uchida, 2000). N, P, K, Fe, B and Zn are essential macronutrients for several plant physiological functions such as photosynthesis, enzymatic reactions, respiration, ribonucleic acid formation, tryptophan synthesis, maintaining genetic information, root development, stomatal regulation, protection against oxidative damage and creating amino acids (Silva & Uchida, 2000). However, macronutrient and micronutrient deficiencies often result in stunted growth, chlorosis, reduced protein content, weak stem production and can cause early maturity in some plants (Silva & Uchida, 2000).

Traditionally, scientists obtained foliar nutrient information using destructive sampling methods such as wet chemistry analysis, which involve ground-based periodic surveys and tedious laboratory work that is costly and time-consuming (Pullanagari *et al.*, 2016). However, researchers have made little progress using indirect spectral methods (Oliveira *et al.*, 2017). Remote sensing offers a rapid, non-destructive, and effective approach for detecting key nutrient levels in forest trees. Handheld/proximal hyperspectral data can benefit high productivity environments, such as in younger plants in forest nurseries. The detection of foliar nutrients occurs through specific absorption features within the electromagnetic spectrum. Earlier research explains the physiological link between foliar nutrient content and remote sensing, e.g., (Curran, 1989; Dixit & Ram, 1985; Elvidge, 1990). More specifically, hyperspectral systems (350–2500nm) capture detailed spectral information; however, they are often sensitive to the influence of spectral noise, which negatively impacts classification approaches (Agjee *et al.*, 2018). Spectral noise can significantly impact the quality of the data acquired; hence, classification approaches' performance will deteriorate (Agjee *et al.*, 2018).

For example, Oliveira *et al.* (2017) successfully estimated the N content of 25-month *Eucalyptus* trees and compared a wide range of variable importance (VI) results. As a result, the authors obtained the best R^2's of 0.97 using inflexion point position (IPP), normalized difference red-edge (reNDVI) and modified red-edge normalized difference vegetation index (mNDI) in the 400–900nm range. A later study by Oliveira and Santana (2020) estimated the full range of macronutrients and micronutrients: N, P, K, S, Ca, Mg, Mn, B, Zn, Cu, and Fe in *Eucalyptus* clones using the NIR region (400-900 nm) and PLSR. As a result, the authors predicted all nutrients using the coefficient of determination of cross-validation (R_{CV}^2), with the lowest and highest estimate was Mg (0.22 R_{CV}^2) and N (0.95 R_{CV}^2), respectively. The authors found that PLSR and variable selection methods increased the accuracy of nutrient concentration estimates and suggested future studies to use wavelength ranges above 900 nm. Osco *et al.* (2020b) tested machine learning algorithms: kNN, lasso regression, RR, SVM, ANN, DT, and RF using a proximal hyperspectral sensor (380- 1020nm) to predict nutrient content on a Valencia-orange orchard. The authors obtained high predictions (R^2) above 0.73 for all algorithms and found that RF was the most suitable algorithm.

Many studies have predicted nutrient concentrations using handheld/proximal hyperspectral data and a vast array of computational algorithms (Abdel-Rahman *et al.*, 2017; Ferwerda *et al.*, 2005; Oliveira & Santana, 2020; Wang *et al.*, 2018). However, few studies used a full-waveform hyperspectral proximal sensor and RF to predict macronutrient and micronutrient deficiency. Therefore, this study aimed to predict macronutrients and micronutrients in *Eucalyptus* hybrid trees using full-waveform handheld/proximal hyperspectral data (350-2500nm) and the RF algorithm. Furthermore, to our knowledge, no studies have identified the most critical wavebands for detecting nutrient deficiencies in younger trees within a nursery setting. The outcomes of this study will promote the use of remote sensing scanning systems for rapid diagnosis of macronutrient and micronutrient deficiencies in high productivity commercial forestry environments.

3.2 Material and Methods

3.2.1 Study area

This research experiment was conducted under a controlled nursery environment at the ICFR nursery in Pietermaritzburg, KwaZulu Natal, South Africa (29°37'40.20"S and 30°24'13.63"E). This study examined a *Eucalyptus* hybrid (*Eucalyptus grandis x Eucalyptus urophylla*). The *Eucalyptus* genus is a hardwood perennial native to Australia (Myburg *et al.*, 2014). Commercial forestry industries commonly grow *Eucalyptus* trees for their fast growth and superior wood properties (Myburg *et al.*, 2014). Hence, more than 100 countries across six continents (>20 million ha) grow *Eucalyptus* trees as a timber resource (Myburg *et al.*, 2014). The hybrid species *Eucalyptus grandis* and *Eucalyptus urophylla* used in this study are native to Newcastle, New South Wales to Bundaberg in Queensland and the Indonesian Archipelago Timor, respectively (Pajares, 2015; Pinto *et al.*, 2014). The shape of hybrid *Eucalyptus grandis x Eucalyptus urophylla* leaves was lanceolate with the adaxial side dark green and the abaxial slightly paler than the adaxial side.

3.2.2 Experimental design

A pot trial experiment was conducted to develop more explicit diagnostic indicators and measures of changes in soil nutrient status (Figure 3.1). 135 hybrid seeds *Eucalyptus grandis x Eucalyptus urophylla* were obtained from a commercial plantation seed orchard in KwaZulu Natal, Midlands, South Africa, in June 2014. To minimize the effect of a microclimate, pots were randomly arranged in the designated nursery environment. The pots were under an open-sided plastic cover to exclude rainfall and kept under natural sunlight. This study's soil type was Inanda soil, the predominant soil type in the Midlands, South Africa (Mucina & Rutherford, 2006). The soil texture was silty clay (56% sand & silt: 44% clay). Distilled water was added automatically via drip irrigation to maintain optimal soil moisture conditions. Drippers had an output rate of 2.2L per hour programmed to water twice a day for three minutes. Fertilizer was added once per week over four weeks and then left to acclimatize for another four weeks (Table 3.2). The canopy characteristics of the leaf material sampled in this study were at the sapling stage (Juvenile) of growth with leaf area was 6cm to 10cm long and 2cm to 3cm wide. The saplings grew to a height of 30cm to 60cm with a canopy width of approximately 30cm to 40cm. The root characteristics of the saplings had an elongated rooting structure. The hybrid *Eucalyptus* is designed with an extensive tap-root rooting system to anchor the trees and horizontal roots that keep the trees upright when planted in the field (Dye, 1996).

Figure 3. 1: *Eucalyptus grandis x Eucalyptus urophylla* trees planted in pots and an adaxial leaf representation.

The experiment continued until each nutrient reached its depletion threshold determined through foliar and growth diagnosis (Table 3.1). Suggested depletion thresholds were determined by Reuter and Robinson (1997). This strategy enabled calibration of extractable soil nutrient levels with tree growth and foliar nutrient diagnostics to improve laboratory soil data interpretation (Table 3.2).

Table 3. 1: Fertilizer compounds used to exclude specific nutrients in pot experiment treatments

Nutrient	Fertilizer compound added	Compound formula
N	Urea	CH_4N_2O
P	Sodium Dihydrogen Orthophosphate dihydrate	$NaH_2PO_4.H_2O$
H	Potassium Chloride	KCl
Ca	Calcium Chloride	$CaCl_2.2H_2O$
Mg	Magnesium Chloride hexahydrate	$MgCl_2.6H_2O$
Na	Sodium Sulphate anhydrous	Na_2SO_4
Micro-nutrients	Micro-Nutrient Mix	Zn, Cu, Fe, B, Mn, Mo
	Zinc Chloride	$ZnCl_2$

Table 3. 2: Pot trial experimental design

Treatments	Nutrient
Exclusion of macronutrients	-N, -P, -K, -Ca, -Mg, -Na
Exclusion of micronutrients	-Mn, -Fe, -Cu, -Zn, -B (combined)
Control 1	No fertilizer
Control 2	Full Fertilizer

3.2.3 Spectral measurements

Spectral reflectance measurements were taken on the 2nd of February 2017, using a handheld field ASD (FieldSpec® three spectrometers) synchronously with foliar sampling. The ASD measures at a sampling interval range of 1.4nm for 350-1000nm and 2nm for 1000-2500nm. Reflectance measurements were taken 1m above the pot, using the fibre optic cable set at 25 degrees field of view (FOV), pointed at the nadir position. A white reference panel treated with a barium sulphate of known reflectivity to calibrate the sensor was used every ten minutes (Spectralon Labsphere, Inc., Sutton, New Hampshire). Ten measurements of each plant (per pot) were acquired to derive the representative reflectance spectra for each pot for a total of 135 pots. The ASD instrument operator was positioned as far away from the area under observation to reduce interferences caused by anthropogenic shadow and reflection (Figure

3.2). Each pot was placed in direct sunlight on a cloudless day between 10:00h and 14:00h Central African Time.

Furthermore, pots were placed on a stable black platform in an open area to minimize bidirectional reflectance distribution function (BRDF) effects and background scattering. After averaging the spectra, all spectra were converted from radiance to reflectance using ViewSpec Pro software (ASD Inc., Boulder, Colorado, version 6.0.11). All spectra were radiometric and atmospheric corrected to reduce noise using the Natural Environment Research Council (NERC) field of spectroscopy templates (NERC, Undated).

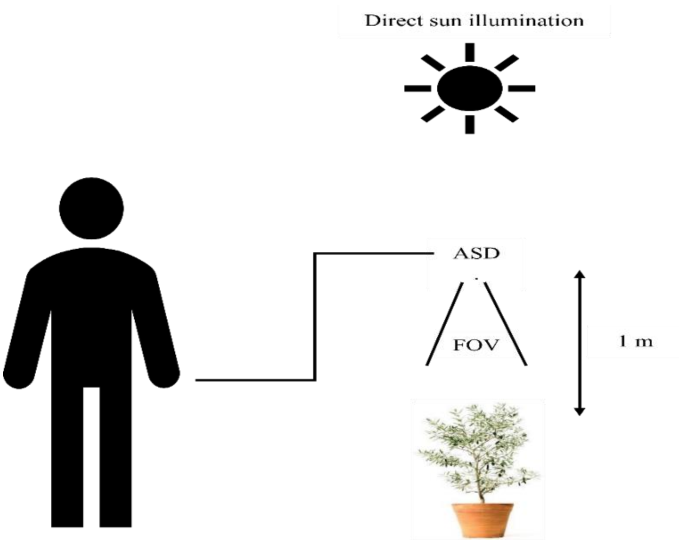


Figure 3. 2: Schematic drawing of ASD measurements procedure.

3.2.4 Wet chemistry

Foliar samples were taken from fully expanded leaves from the top third of all 25 trees in the sample plots for foliar diagnostics. All samples were dried, weighed and analysed for physio-chemical properties and expressed by leaf concentration (%/dry weight). All laboratory tests were executed using standard measures as described by (Donkin *et al.*, 1993a, 1993b).

3.2.5 Reference data t-test

In this study foliar chemistry was measured before and after nutrient depletion. Nutrient depletion thresholds were defined in Reuter and Robinson (1997). The foliar chemistry results were used to test the significant difference in means ($\bar{x}$) of foliar macronutrient and micronutrient at both high and low levels of nutrient content using a paired τ-test for nutrient deficient plants. As a result, the paired τ-test provides a score of the significance (ρ-value) by calculating the difference in the $\bar{x}$ of two groups data (Ruxton, 2006). The ρ-value was used to calculate the statistical hypothesis test results, assuming the null hypothesis is correct (Meng, 1994). If the ρ-value is $<= 0.05$, the two datasets are significantly different, and if the ρ-value is > 0.05, the result is insignificant; hence the trees did not deplete the targeted nutrient. The τ-test was calculated using the following formula metrics:

$$t = \frac{m}{s/\sqrt{n}}$$

Where τ is the paired τ-test, m and s are the mean and the standard deviation of the difference in samples, respectively. N is the size of the sample.

3.2.6 Random Forest

This study used RF for regression based on the classification and regression trees (CART) to predict (Breiman, 2001). The RF ensemble was implemented using the "randomForest" package R in R statistical software (R Development Core Team). Compared to several other machine learning algorithms, the RF algorithm delivers the most consistent results, especially when using high dimensional handheld/proximal hyperspectral data for predicting foliar nutrient data (Amirruddin *et al.*, 2020). The RF algorithm produces decision trees by drawing a subset of training samples through a replacement method known as "bagging". The bagging process refers to selecting the same sample several times while the remaining samples can remain unselected. During the training process, approximately two-thirds of the samples from the training set are used as "in-bag” samples, while the remaining one third "out-of-bag" (OOB) samples are used as an internal cross-validation technique to test the RF models' performance (Belgiu & Drăguţ, 2016). The error produced through the cross-validation technique is known as the OOB error. The RF algorithm splits each node by a user-defined parameter called the m_{try} function whilst each decision tree is independently produced without pruning (Breiman, 2001). Another user-defined parameter is the n_{tree} which grows the forest whilst the algorithm generates trees which have high variance and low bias (Breiman, 2001). The model grows trees until the final classification is taken by the average class assignment probabilities using the arithmetic mean. The RF algorithm evaluates the model using the final classification and produces a new unlabelled data input against all the decision trees, and each tree votes for class membership. Here, the membership class that receives the maximum votes is the final model selected (Belgiu & Drăguţ, 2016). The model was parametrized using the m_{try} and n_{tree} functions. RF recognizes that classification accuracy is sensitive to m_{try} than the n_{tree} function. R sets the default n_{tree} value at 500, whereby errors stabilize as the R package grows trees (Belgiu & Drăguţ, 2016). The n_{tree} value can be optimized to find the best detection accuracies and lowest error rates (Mutanga *et al.*, 2012). After numerous runs through the RF algorithm, an n_{tree} value of 750 provided more accurate results with the data used in this study. The m_{try} and n_{tree} functions were set to 46 and 750, respectively.

3.2.7 Variable importance

Determining the most critical variables is essential for improving model optimisation, simplification and robustness, especially when dealing with high dimensional data in this study (Liaw & Wiener, 2002). RF has three measures of variable importance. VI bases the first measure on the number of times a candidate variable is selected in the model. The second measure of importance is based on the Gini impurity when a variable is chosen to split a node, as Breiman (2001) proposed. Finally, the third measure is the permutation of a variable as an ensemble of VI (Breiman, 2001). In this study, the third measure (permutation of variables) of variable importance, and the mean squared error (MSE) in percentage was chosen for determining VI. VI was determined for each nutrient to generate a coherent account of the relevant variables used in the prediction models for nutrient deficiency.

3.2.8 Accuracy assessment

The final dataset was split into training (70%) and test (30%) (Breiman, 2017). The R^2 was used for prediction and RMSE to assess the RF algorithm's performance in determining nutrient deficiency. The R^2 was calculated using predicted and observed values, values closer to one, predict better results. The RMSE was calculated using predicted and observed values, values closer to zero, predict better results. Higher R^2 and lower RMSE values indicate a reliable model.

3.3 Results

3.3.1 Descriptive statistics and reference t-test

Table 3.3 summarizes descriptive statistics related to *Eucalyptus grandis x Eucalyptus urophylla* foliar macronutrient and micronutrient at low and high levels. A random subset (10 pots) was taken for each nutrient on the repetitive measured pots before and after inducing deficiency to test if nutrient deficiency had occurred. The paired τ-test revealed that there was a significant difference ($\rho \leq 0.05$) in $\bar{x}$ whereby the average $\rho - value = 0.0067$.

Table 3. 3: Descriptive statistics of *Eucalyptus grandis x Eucalyptus urophylla* foliar macronutrient and micronutrient content.

Nutrient	High				Low			
	Min	**Max**	**Median**	**SD**	**Min**	**Max**	**Median**	**SD**
Macronutrients (%)								
N	0.59	3.67	2.13	0.94	0.83	3.16	1.995	0.50
P	0.002	0.41	0.206	0.08	0.015	0.50	0.2575	0.06
K	0.02	2.45	1.235	0.39	0.26	1.02	0.64	0.11
Ca	0.31	2.88	1.595	0.49	0.39	1.61	1	0.25
Mg	0.003	0.88	0.441	0.11	0.21	0.59	0.4	0.08
Na	0.08	0.32	0.2	0.06	0.05	0.21	0.13	0.05
Micronutrients (ppm)								
Mn	29.96	5 137.42	2583.7	968.84	837.85	8 417.55	4627.7	1 620.46
Fe	23.11	669.57	346.34	69.98	43.04	214.85	128.95	32.92
Cu	3.42	10.88	7.15	1.44	5.51	17.77	11.64	2.57
Zn	3.44	47.17	25.305	9.24	13.19	64.64	38.915	10.73
B	23.79	166.49	95.14	39.65	5.70	98.53	52.115	22.15

3.3.2 Detecting macronutrient N, P, K, Ca, Mg, Na, using RF

An important step was applied to determine the nutrient content of each tree. Furthermore, it was essential to test trees' prediction accuracy with combined low and high nutrient concentrations. The RF model ($n_{tree} = 750$ and $m_{try} = 46$) successfully predicted the macronutrient of trees with low and high nutrient concentration levels. Figure 3.3 shows predicted versus observed distribution plots for all macronutrients. As a result, the RF algorithm successfully predicted macronutrients N, P, K, Ca, Mg, and Na with R^2's of 0.88, 0.95, 0.73, 0.88, 0.70, 0.75 and RMSE's of 0.15, 0.13, 0.14, 0.10, 0.06, 0.04, respectively (Figure 3.3).

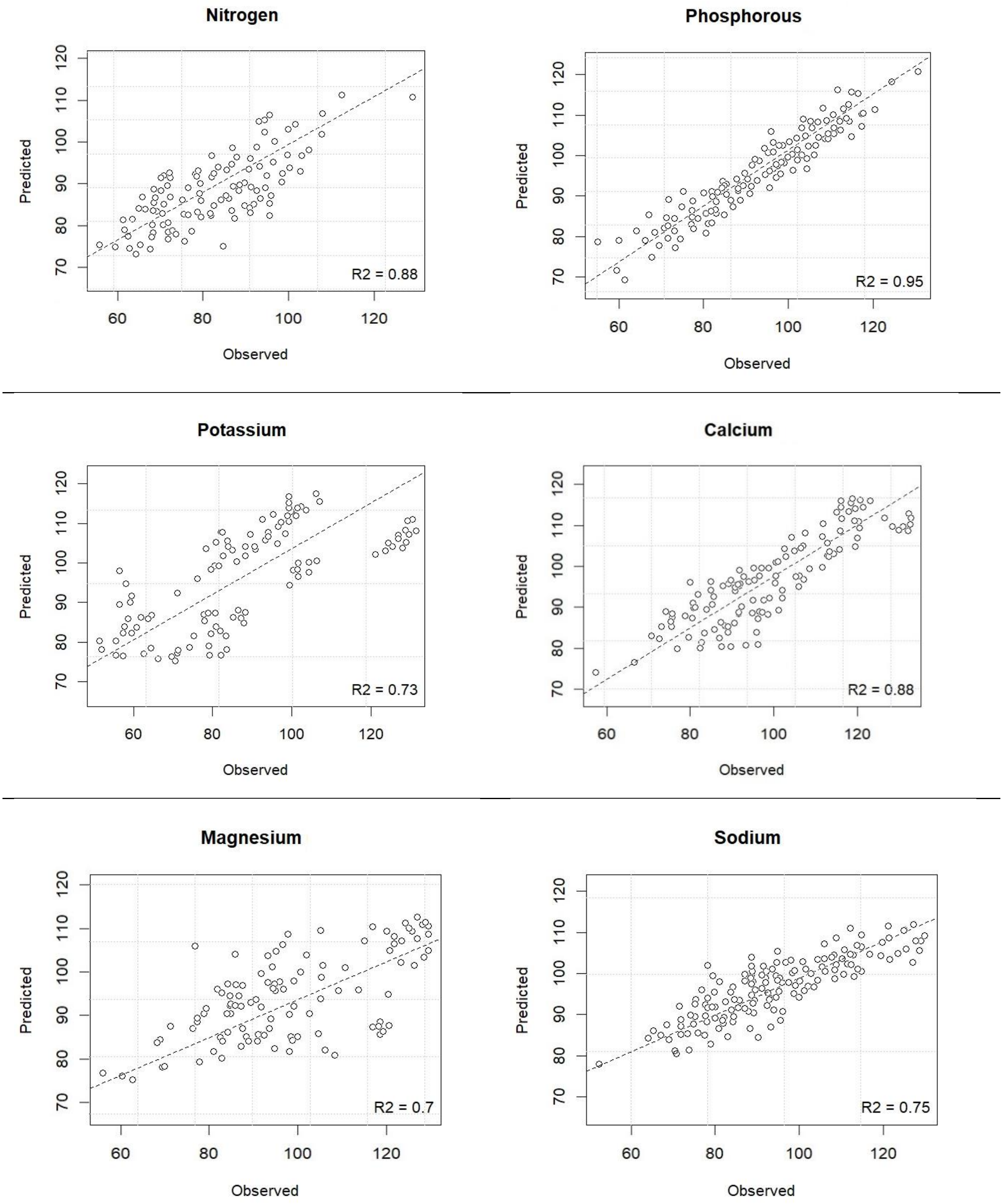


Figure 3. 3: The one-to-one relationship between predicted versus observed macronutrients: N, P, K, Ca, Mg, Na using RF.

3.3.3 Detecting micronutrients Fe, Mn, Cu, Zn, B using RF

To predict micronutrients, the same n_{tree} and m_{try} input values were used in the macronutrient RF prediction model. Figure 3.4 shows predicted versus observed distribution plots for all micronutrients. As a result, the RF algorithm successfully predicted low and high levels of micronutrients concentrations Fe, Mn, Cu, Zn, and B with R^2's of 0.79, 0.90, 0.86, 0.69, 0.66 and RMSE's of 11.86, 745.01, 1.70, 15.09, 2.54, respectively (Figure 3.4).

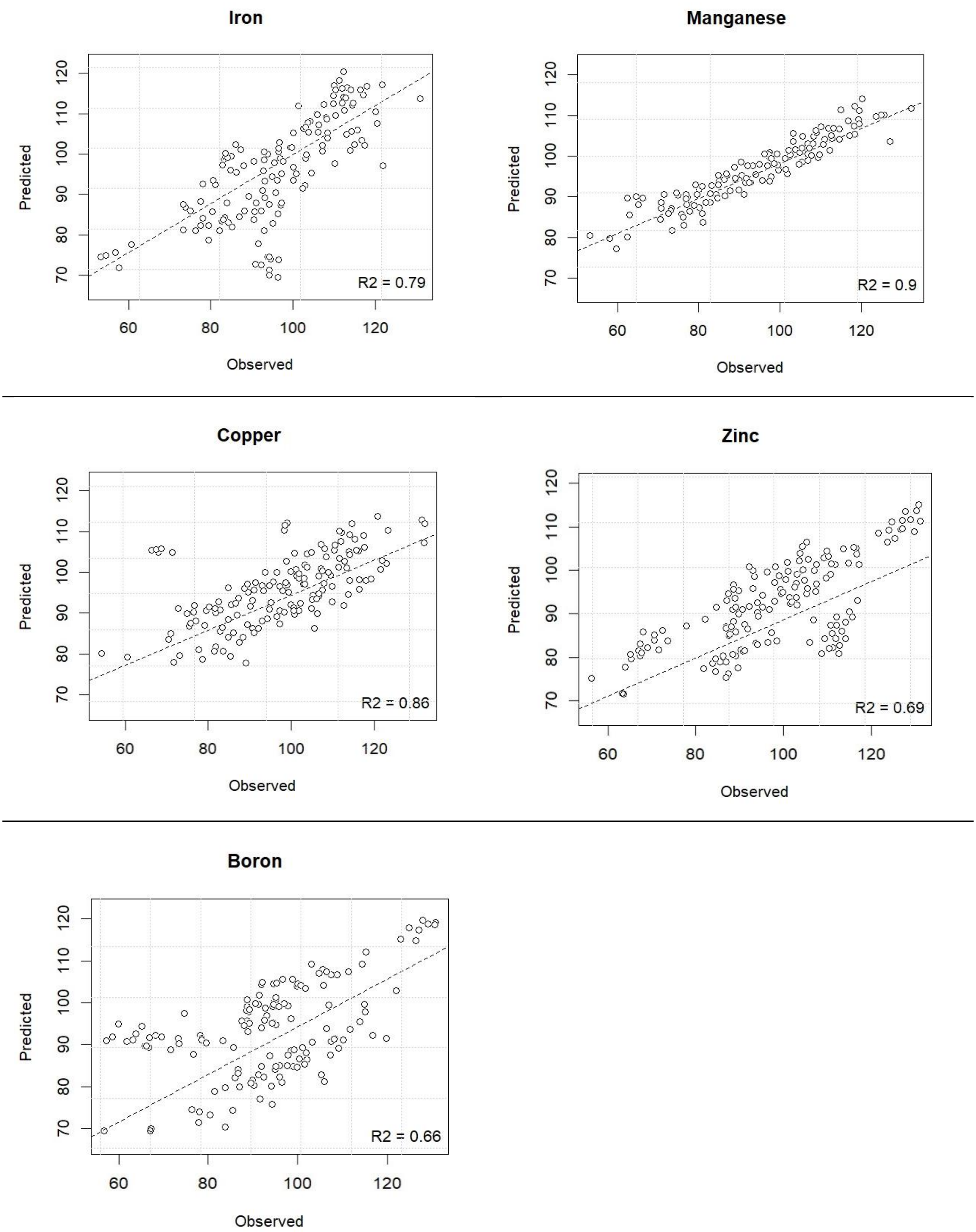


Figure 3. 4: The one-to-one relationship between predicted versus observed for micronutrients: Fe, Mn, Cu, Zn, and B using RF.

3.3.4 Variable importance of macronutrients

Figure 3.5 shows the most effective wavebands for detecting macronutrients at low and high concentrations used in the prediction model. The RF ensemble measures VI to obtain the most important wavebands. This study considered the top ten most important wavebands for each macronutrient. The SWIR region was partitioned into SWIR-1 (1300nm to 1900nm) and SWIR-2 (1900nm to 2500nm). For N content and Na concentrations, the most important wavebands are (1949nm, 675nm, 722nm, 892nm, 908nm, 918nm, 930nm, 947nm, 950nm, 955nm) and (901nm, 931nm, 955nm, 1131nm, 1371nm, 1420nm, 1960nm, 902nm, 907nm, 908nm) found in the NIR and red edge regions of the spectrum, respectively. For P, Ca, and Mg concentrations, the important wavebands are (429nm, 353nm, 360nm, 367nm, 407nm, 413nm, 438nm, 444nm, 452nm, 453nm), (357nm, 383nm, 354nm, 355nm, 362nm, 389nm, 408nm, 422nm, 507nm, 672nm) and (1295nm, 1302nm, 1980nm, 1296nm, 429nm, 442nm, 444nm, 449nm, 450nm, 453nm) found in the VIS region, respectively. Whilst, for K concentration detection, the most important wavebands are (1912nm, 389nm, 350nm, 390nm, 2045nm, 2124nm, 2139nm, 2494nm, 406nm, 2098nm) found in the VIS (40 %) and SWIR-2 region (60 %), respectively (Figure 3.5).

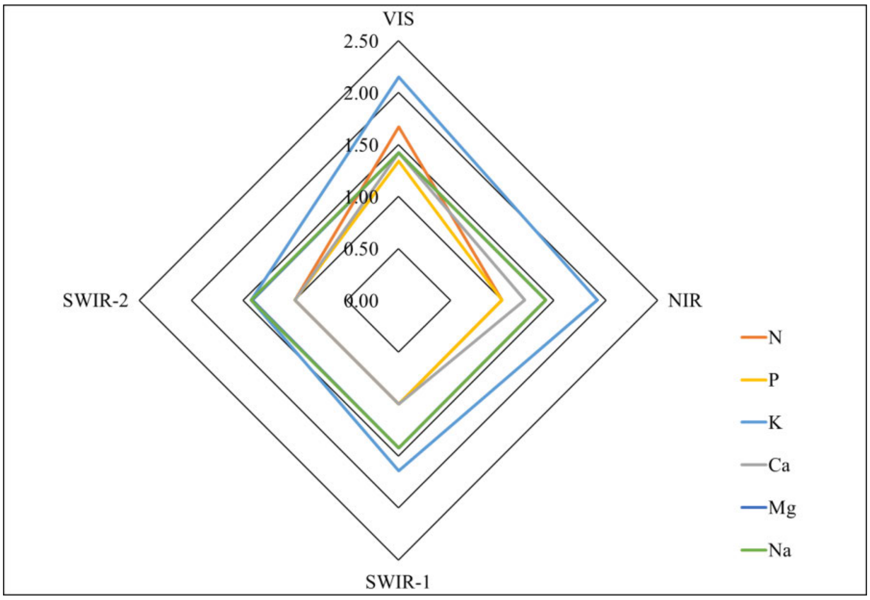


Figure 3. 5: Radar plot showing important spectral regions for detecting deficient macronutrients: N, P, K, Ca, Mg, and Na. The electromagnetic regions are illustrated in the following colours: VIS (yellow), NIR (green), SWIR-1 (blue) and SWIR-2 (red).

3.3.5 Variable importance of micronutrients

Figure 3.6 shows the most effective wavebands for detecting micronutrients at the low and high concentrations used in the prediction model. Fe, Zn and B produced correlations with the NIR region. The most critical wavebands for detecting Fe, Zn and B in this study were (1123nm, 1128nm, 894nm, 907nm, 919nm, 922nm, 941nm, 958nm, 961nm, 969nm), (788nm, 1186nm, 949nm, 1564nm, 445nm, 469nm, 636nm, 714nm, 742nm, 744nm), (784nm, 896nm, 904nm, 1263nm, 897nm, 1069nm, 1073nm, 1170nm, 1307nm, 1318nm), respectively. Mn and Cu showed strong correlations with the SWIR-2 region. The most critical wavebands for detecting Mn and Cu in this study were (2496nm, 1909nm, 2418nm, 2495nm, 2451nm, 1800nm, 1877nm, 1222nm, 1260nm, 1291nm) and (1906nm, 1921nm, 1914nm, 1923nm, 1922nm, 1909nm, 1929nm, 2483nm, 405nm, 1225nm) found in the SWIR-2 region, respectively (Figure 3.6).

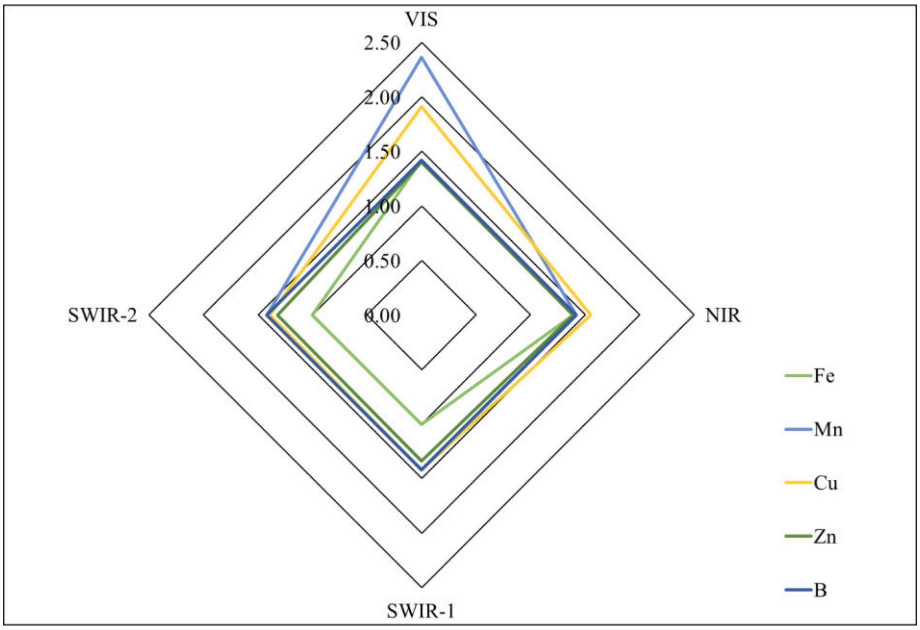


Figure 3. 6: Radar plot showing important spectral regions for detecting deficient micronutrients: Fe, Mn, Cu, Zn, and B. The electromagnetic regions are illustrated in the following colours: VIS (yellow), NIR (green), SWIR-1 (blue) and SWIR-2 (red).

3.4 Discussion

In commercial forestry, the supply of nutrients from root to shoot is vital for plant growth and forest productivity. Quantifying nutrient-deficient trees in a compartment remains unworkable and could prove challenging when dealing with many younger plants in the nursery. Hence new methods are needed that can be adaptable early, either at the nursery before planting, to provide rapid detections or out in the field (Quentin *et al.*, 2017). The early detection of potential nutrient depletion at the nursery level could guide the optimisation of future forest management practices and lead to a more robust approach to nutrient measurement and assessment (Garcia *et al.*, 2018). A recent study by Acevedo *et al.* (2020) suggests that the use of nutrient loading at the nursery level will improve seedling nutritional status, morphological attributes, and the growth of new roots. The authors suggest modelling growth responses to improve their understanding of physiological processes further.

This study successfully determined foliar macronutrients in *E. grandis x E. urophylla* at low and high nutrient concentration levels using handheld/proximal hyperspectral data and RF. The prediction results were high throughout all macronutrients, vital for improving remote sensing efficacy, particularly in deficient samples. The prediction results (R^2) of the most limiting growth nutrients N, P, K in this study explain the findings of previous studies (Adams et al., 2000; Axelsson et al., 2013; Özyiğit & Bilgen, 2013), which detected similar R^2 accuracies in foliar wheat and grass samples. The use of the RF algorithm significantly improved detection accuracy when compared to previous studies that used VI's and PLS to detect N, P, K and Na in wheat samples (Mahajan et al., 2014; Oliveira et al., 2017; Pimstein et al., 2011). For example, P predicted considerably better than previous studies (Mahajan et al., 2014; Özyiğit & Bilgen, 2013) with R^2's above 0.90, while previous studies produced R^2's below 0.50. The RF algorithm has built-in parameter fine-tuning, which permits the optimisation of the n_{tree} function, providing more robust results than generalized VIs used in previous studies. RMSE values remained low; however, N samples were found to be considerably higher than the rest but still permissible. Our findings confirm that foliar micronutrients can be detected in E. grandis x E. urophylla at both deficient and not deficient levels using handheld/proximal hyperspectral data and RF. While many studies generally predict macronutrients N, P, K, this study predicted a wide range of micronutrients. However, this study predicted micronutrients: Ca, Mg, Na, Mg at low and high concentrations. This study could predict micronutrients better with prediction accuracies from 0.66 to 0.90. While Ca produced the highest R^2 among Ca, Mg and Na. While other studies used VI's (Adams et al., 2000; Oliveira et al., 2017; Özyiğit & Bilgen, 2013), this study achieved higher accuracies when using the RF algorithm. This achievement validates and promotes the robustness and effectiveness of the RF ensemble to discriminate each micronutrient, especially when using high dimensional data.

An important step was determining which waveband regions correlate with the deficient nutrients. To our knowledge, our study is the first to examine the critical wavebands for detecting macronutrient and micronutrient deficiencies in foliar tree material. Hence, this study results could not be directly compared to previous studies, mainly examining N deficiency in heterogeneous environments (Blackmer *et al.*, 1996; Goel *et al.*, 2003). However, the VI results correlated well with other corresponding regions of the electromagnetic spectrum associated with general reflectance markers in foliar material. For example, this study found that most macronutrient deficiencies correlate with wavebands in the VIS (P, Ca, Mg) and NIR (N, Na) regions. Blackmer *et al.* (1996) found similar correlations in the VIS region when examining N deficiency in corn using a portable spectroradiometer (350-1100nm). In the absence of the latest instrumentation, their study could not examine the NIR edge and SWIR regions, which were essential regions for determining N deficiency in our study.

According to Goel *et al.* (2003), wavebands 498nm and 671nm in the VIS region correlated with N stress in foliar corn material. Similarly, waveband 675nm closely related to their study in this study (Figure 3.5). However, higher correlations were found in this study for N deficiencies in the red-edge and NIR region. Like N, most micronutrient deficiencies correlated with the NIR (Fe, Zn, B) regions related to Liew *et al.* (2008). Also, Mn and Cu deficiencies were more closely related to the SWIR-2 region. Obtaining lower RMSE values than higher R^2 values were crucial for generating more robust models for each nutrient during modelling. Lower RMSE values improve the technology's efficacy and provide confidence to the user (forester, nursery manager or technical staff), particularly during the system's implementation into commercial forestry nurseries. The RF algorithm helped provide VI measures, essential for identifying the most critical wavebands when using high dimensional data. Deriving reference paired τ-test results formed an essential component of this study for deciphering between nutrient-deficient trees and not nutrient deficient. The results from the reference paired τ-test showed a significant difference ($\rho \leq 0.05$) in samples with deficient and not deficient for all macronutrients and micronutrients.

This study provides a framework for proactive decision-making about the nutrient health status of a tree. Nurseries could use this method for quality control and risk assessment purposes. Rapid spectroscopy is cost-effective, time-efficient and requires fewer resources for the chemical processing of samples. Furthermore, future studies should upscale this assessment to a live standing compartment. This study will help foresters, land managers, and commercial timber industries rapidly assess each tree's health status within a compartment. Upscaling to the hyperspectral satellite data would be beneficial; however, problems of resolution (e.g., spectral) may hinder the detection of macronutrient and micronutrients deficiency in *Eucalyptus Grandis* trees.

3.5 Summary and Conclusion

To our knowledge, this is the first study that has explored remote sensing of a full range of tree macronutrients (N, P, K, Ca, Mg, Na) deficiencies and micronutrients (Fe, Mn, Cu, Zn, B) deficiencies, using full-waveform handheld/proximal hyperspectral data (350-2500nm) and the RF algorithm. From this study, this study concludes that:

- The study successfully predicted N, P, K, Ca, Mg, S, Fe, Mn, Cu, Zn, and B in E. grandis x E. urophylla using handheld/proximal hyperspectral data and RF analysis.
- Variable importance results predicted wavebands for detecting nutrient deficiencies in E. grandis x E. urophylla.
- The results improve the efficacy of using remote sensing methods for nutrient analysis in a high productivity forestry nursery environment.

This study was an important first step for detecting nutrient deficiencies at a micro-level (nursery environment). Future studies should use this study as a framework for rapid plant nutrient analysis in commercial forestry nurseries. Future research could upscale the results from this study from nursery to field level as well as investigate detecting the distribution of key nutrients within forest trees using remote sensing imagery. Hence, the upcoming chapter will attempt to detect nutrient deficiencies at macro-scale (field environment). Understanding the capabilities of remote sensing to detect nutrient deficiencies at both scales are important for implementation of the technology commercially.

CHAPTER 4: Impacts of vertical canopy positioning in the detection of nutrients in hybrid *Eucalyptus* trees using near-infrared datasets

Impacts of vertical canopy positioning in the detection of nutrients in hybrid Eucalyptus trees using near-infrared datasets

This chapter was based on:

Singh L*, Mutanga O, Mafongoya P, Peerbhay KY, Ismail R (Preparation). A Rapid Diagnostic Tool for Detecting Tree Growth Nutrients Using Near-Infrared Spectroscopy and Vertical Canopy Positioning.

Abstract

Ineffective nutrient screening technologies could lead to the release of unhealthy trees for in-field planting, wasting functional space and time and inevitably impeding production. In this study, macronutrients: N, P, K, Ca, Na, Mg, and micronutrients: Fe, B, Cu, Mn, Zn of 135 *Eucalyptus grandis* saplings were measured using NIRS and the vertical canopy gradient technique. *Eucalyptus grandis* seeds were planted in two-litre plastic pots filled with topsoil and placed them in a controlled nursery environment. Non-destructive samples were acquired at four VCPs using a handheld NIR spectrometer device. Leaf samples were picked and scanned on both adaxial and abaxial sides to provide a representative sample. Pre-processing techniques: PCA and SG smoothing were applied before implementing the PLS regression algorithm to understand the distribution of nutrient content across the four VCPs. The combination of NIRS and VCP successfully determined nutrient concentrations. Overall, the findings of this work provide an alternative screening framework for commercial forestry nurseries that require quality planting material for long- and short-term resource sustainability.

Keywords: Remote sensing, forest biochemistry, infrared image sensors, vegetation mapping, vertical canopy positioning

4.1 Introduction

Nutrient availability plays a significant role in plant survival and growth, with an optimal supply of critical nutrients vital for increasing plant stem volume and density and overall productivity. N, P, and K are among the most critical macronutrients known for efficient plant growth. For instance, N is the major component of Chl and amino acids used for photosynthesis and protein synthesis, effective for plant physiological functions and development (Richardson *et al.*, 2009; Vessey, 2003). Several other macronutrients are responsible for overall organism function (Khan & Lee, 2013; Martinelli *et al.*, 2000; Nguyen *et al.*, 2015) and physiological processes in plants (Bahar *et al.*, 2018; Garcia & Zimmermann, 2014). Although plants require an abundant supply of macronutrients for overall growth, they also require adequate micronutrients such as B, Fe, Cu, Mn, molybdenum (Mo), and Zn (Ma *et al.*, 2012; Rengel, 2007). Predominantly, the concentration of within-tree macronutrients and micronutrients are in the leaves and branches (Martinelli *et al.*, 2000). Therefore, an important step is understanding the distribution of nutrients across a tree's vertical canopy gradient (Fig. 2).

The distribution of nutrients across the vertical canopy gradient of a tree is complex and variable, challenging most estimation processes (Gara *et al.*, 2018; Gara *et al.*, 2019; Zhao *et al.*, 2016). Forest canopies are spatially heterogeneous environments, and the relationships between the chemical content, light interception, canopy structure and patterns of photosynthesis are poorly understood (Ellsworth & Reich, 1993; Mutowo *et al.*, 2019; Wang & Li, 2013). Previous studies only took measurements at a single VCP which did not represent nutrient distribution across the VCP. Furthermore, research has shown that the tree's re-translocate nutrients throughout the tree canopy as a conservation strategy (Fife *et al.*, 2008). Hence, assessing the VCP variation to chemical content can provide valuable insight into the partitioning of nutrient resources within forest trees (Ellsworth & Reich, 1993).

In the 1970s, researchers began investigating the interaction of electromagnetic radiation as a proxy for detecting foliar biochemicals and their distribution within forest canopies (Aber, 1979a, 1979b; Hansen *et al.*, 1987; Reich *et al.*, 1990). The principles of Beer's law provided researchers with a physiological basis to understand the physical link between organic substances and electromagnetic radiation. Henceforth, the interactions of light with organic compounds triggered energy transitions during partial intermolecular bonding known as vibrational states (Weyer, 1985). Weyer (1985) describes three bond vibrations, namely: (1) C-H; (2) N-H; and (3) O-H absorption. Weyer (1985) also experimented with instrumentation, analysis techniques, and remote sensing applications for detecting organic substances. As a result, a plethora of research findings expose specific regions of the electromagnetic spectrum where organic compounds absorption occurs (Curran, 1989; Dixit & Ram, 1985; Elvidge, 1990; Sasaki *et al.*, 1984; Wessman *et al.*, 1989). Previous researchers used mathematical models combined with radiometers to find the best sampling points through the VCP to understand the distribution of foliar biochemicals within forest canopies. For example, Ellsworth and Reich (1993) found that N content was distributed throughout the forest canopy in a spatially patterned way in response to height using a portable integrating radiometer (400 nm to 700 nm). However, primitive technological innovations limited early research whereby most studies typically estimated C, N, leaf area index (LAI), foliage, and canopy gas exchange using point-based approaches and tools such as cameras and point quadrats. Furthermore, not many studies estimated macronutrients and micronutrients that are important for tree growth and the ecological functions of the forest.

The latest technological innovations offer platforms for measuring nutrient content through the vertical canopy gradient using non-destructive approaches (Gara *et al.*, 2018). Gara *et al.* (2018) found that variations in leaf mass accounted for most of the variables useful for canopy-level scaling relationships and the partitioning of nutrient resources. Typically, laboratories analyse nutrient content using wet chemistry techniques from extractable soil and foliar dry matter (Chase *et al.*, 2016). For example, a common wet chemistry technique for determining nitrogen content is the Kjeldahl method that requires fresh weight samples to be dried, grinded, and sieved into dry weight samples (e.g., leaves) before chemical analysis (Bremner, 1960). In summary, foliar dry matter is placed into a digestion flask for distillation and titration using chemical solutions and results are derived from the residual weight (pre-weight - post-weight).Although these methods produce accurate results, they often become tedious, prove costly over large scale sampling operations and do not account for the variation of nutrient content in live standing trees (Downes *et al.*, 1997). Furthermore, there are limited expertise and resources to process, analyse and interpret such information. The latest advancements in remote sensing technologies can offer a more practical, faster, and provide a broader area coverage for forestry opportunities in estimating and monitoring foliar nutrient content in trees (Mutowo *et al.*, 2018; Pasquini, 2018; Quentin *et al.*, 2017).

For example, Li *et al.* (2018) successfully estimated leaf nitrogen concentration (LNC) in the upper, middle and lower layers of oilseed rape (*Brassica napus L.)* and obtained R^2 results ranging between 0.83 to 0.90 using handheld hyperspectral data (350 nm to 2500 nm), PLS, lambda-lambda R^2 (LL R^2) and SVM models. Gara *et al.* (2018) found that leaf traits and leaf reflectance co-vary across the vertical canopy profile. The authors successfully estimated foliar N, Chl, C, and equivalent water thickness (EWT) between the upper, middle and lower parts of a tree using handheld hyperspectral data and Partial least squares-discriminative analysis (PLS-DA) (350 nm to 2500 nm). Meacham-Hensold *et al.* (2019) successfully predicted %N in a high-throughput tobacco (*Nicotiana tabacum*) environment using PLSR. The authors predicted an R^2 of 0.83 for the %N. However, only a few studies have estimated foliar nutrients at various VCPs using NIRS and PLSR. Recently, NIRS technologies have also shown success in estimating foliar nutrients.

NIRS provides a non-destructive approach for nutrient sampling in a high throughput nursery environment. Non-destructive sampling approaches reduce the risk of epidemiological exposure in young plants (Muñoz-Huerta *et al.*, 2013). Furthermore, non-destructive NIRS allows multiple nutrients to be analysed simultaneously from a single scan; NIRS reduces data dimensionality by using a smaller portion of the electromagnetic spectrum (Murguzur *et al.*, 2019). Previous research omits analytical processes once calibration models are in place, allowing for a further reduction in dimensional data reduces processing costs up to 80% (Murguzur *et al.*, 2019). Previous studies have found leaf optical properties closely related to foliar chemistry in the NIR region (Asner, 1998; Ferreira *et al.*, 2018; Mutanga *et al.*, 2004b; Van Deventer *et al.*, 2015; Wallis *et al.*, 2019; Yu *et al.*, 2020). For example, Ferreira *et al.* (2018) determined carbon (C), N, extractives, acid-soluble lignin, Klason insoluble lignin, and holo-cellulose in *Eucalyptus* harvest residues using handheld NIR sensor (1100 nm to 2500 nm) combined with PLS-DA. Asner *et al.* (2011) determined a range of leaf chemical traits including N, P, K, Ca, Mg, Zn, B, Fe, Mn in humid tropical forests using airborne VIS to NIR (400 nm to 1050 nm) combined with PLSR.

In summary, nutrients are active components in the development and growth of forest trees. However, methods used to determine their nutrient content or deficiency have proven challenging. This study, therefore, aims to accurately detect valuable macronutrients and micronutrients required for optimal tree growth over a vertical canopy gradient using NIR technology. This study will expand on Gara et al. (2018) research as a basis for vertical canopy research and take a step further by introducing an additional canopy position. Trees relocate nutrients throughout the canopy as a conservation mechanism; therefore, an additional sample position combined with using an NIR sensor and the PLSR algorithm would be a promising approach to more accurately determining nutrient content. Future earth exploration satellite sensors will offer an opportunity for scientists to seamlessly integrate satellite imagery with handheld NIR spectrometer devices (Adão *et al.*, 2017). Hence, NIR data will be readily provided as an effective alternative to acquiring nutrient information over a large area, rapidly and more accurately.

4.2 Methods and materials

4.2.1 Study area and species description

This research experiment was conducted under a controlled nursery environment at the ICFR nursery in Pietermaritzburg, KwaZulu Natal, South Africa (29° 37’40.20 “S and 30° 24’13.63 “E) (Figure 4.1). Figure 4.1 shows the study area location at the ICFR in Pietermaritzburg, KwaZulu Natal, South Africa with nearby important cities. This study examined a *Eucalyptus* hybrid (*Eucalyptus grandis x Eucalyptus urophylla).* The *Eucalyptus* genus is a hardwood perennial native to Australia (Myburg *et al.*, 2014). Commercial forestry industries commonly grow *Eucalyptus* trees for their fast growth and superior wood properties (Myburg *et al.*, 2014). Hence, more than 100 countries across six continents (>20 million ha) grow *Eucalyptus* trees as a timber resource (Myburg *et al.*, 2014). More specifically, the hybrid species *Eucalyptus grandis* and *Eucalyptus urophylla* used in this study originated in Newcastle, New South Wales, to Bundaberg in Queensland and the Indonesian Archipelago and Timor (Myburg *et al.,* 2014; Pinto *et al.*, 2014). The shape of hybrid *Eucalyptus grandis x Eucalyptus urophylla* leaves was lanceolate with the adaxial side dark green and the abaxial slightly paler than the adaxial side.

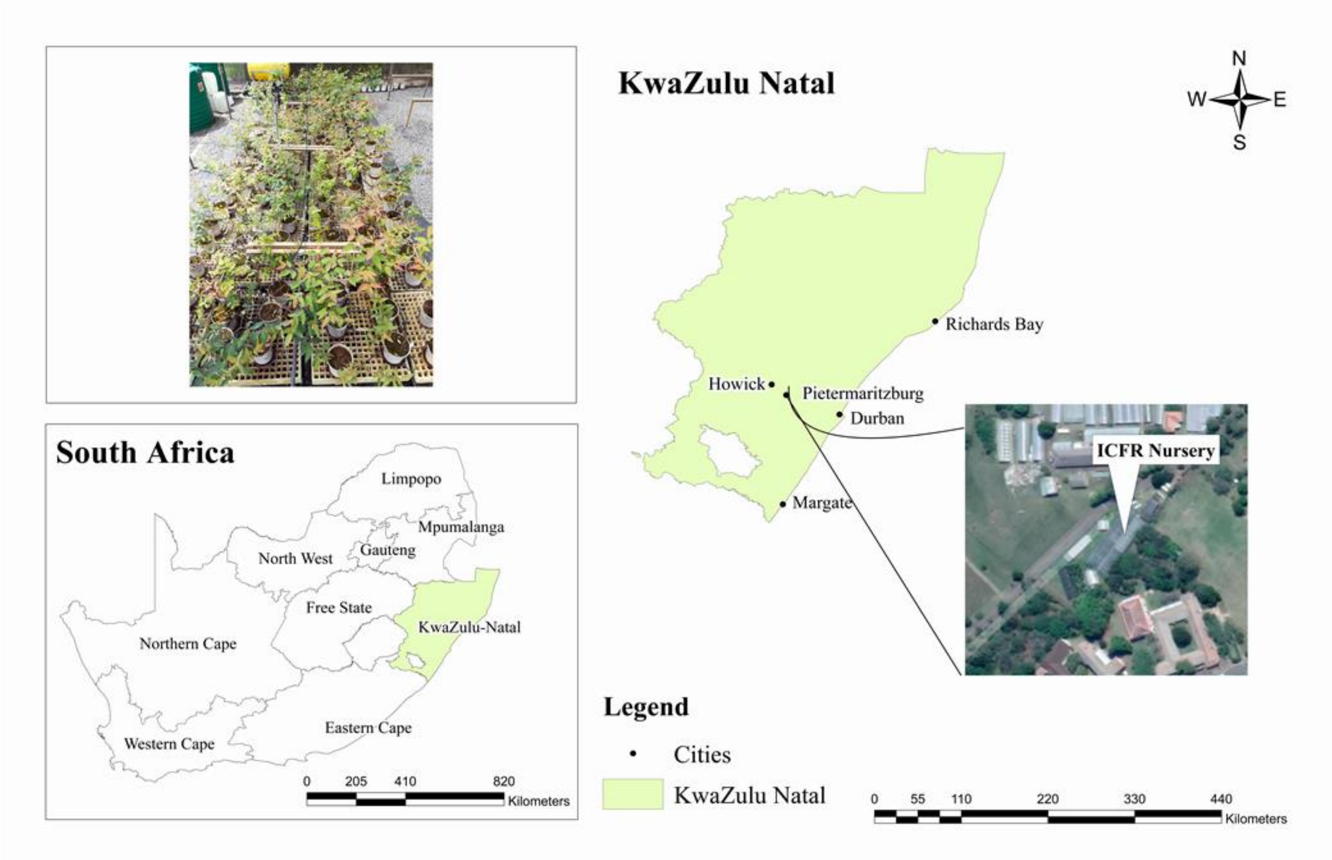


Figure 4. 1: The location of the ICFR nursery study site in KwaZulu Natal, South Africa and pot trial experiment within the nursery environment.

4.2.2 Experimental design

A pot experiment was conducted to develop more explicit diagnostic indicators and measures of changes in soil nutrient status. 135 mature hybrid seeds *Eucalyptus grandis x Eucalyptus urophylla* obtained from a commercial plantation seed orchard in KwaZulu Natal, Midlands, South Africa, in June 2014. To minimize the effect of a microclimate, pots were randomly arranged in the designated nursery environment under natural sunlight. The pots were under an open-sided plastic cover to exclude rainfall at the nursery. The soil type used in this study was Inanda soil (Meyer *et al.*, 1983), and the soil texture was silty clay (56% sand & silt; 44% clay). Distilled water was added automatically via drip irrigation to maintain optimal soil moisture conditions. Fertilizer was added once per week over four weeks and then left to acclimatize for another four weeks (Gara *et al.*, 2018). The canopy characteristics of the leaf material sampled in this study were at the sapling stage of growth with a height of 30 cm to 60 cm. The size of the leaf area was 6 cm to 10 cm long and 2 cm to 3 cm wide. The canopy width was approximately 30 cm to 40 cm. The tree species used in this study has an elongated root structure inside the vase. The height of each sapling grew to approximately 45 cm. The root physiognomies had an elongated rooting structure (Louro *et al.*, 1999). The hybrid *Eucalyptus* has an extensive tap-root rooting system designed to anchor the trees and horizontal roots that keep the trees upright when planted in the field (Dye, 1996).

4.2.3 NIRS leaf spectral measurements

For each leaf sample measurement, approximately three (3) grams of fresh leaves were randomly sampled from all four canopy layers (Q1-Q4) for each sapling (Figure 4. 3). After carefully discriminating four canopy layers, leaves were picked from the outer and interior canopy in all directions (Gara *et al.*, 2018). Four canopy layers were determined based on the height of the sapling along the stem. The picked leaf samples from four canopy layers were packaged in Ziploc bags, stored in a portable cooler, and transported them to the ICFR laboratory, Pietermaritzburg, South Africa, to perform all laboratory measurements within two hours of leaf picking. Figure 4.2 shows reflectance and first derivative NIR spectral signatures of a randomized leaf sample in this study. NIR spectra were acquired using the Fourier Transform-near infrared (FT-NIR) spectrometer device (Model: Multi-Purpose Analyzer, Bruker Optik GmbH, Ettlingen, Germany) at wavelengths from 9000 to 4000 cm-1 (800 to 2500 nm) at 4 cm-1 sampling intervals (2074 spectral wavebands) (Figure 4.2).

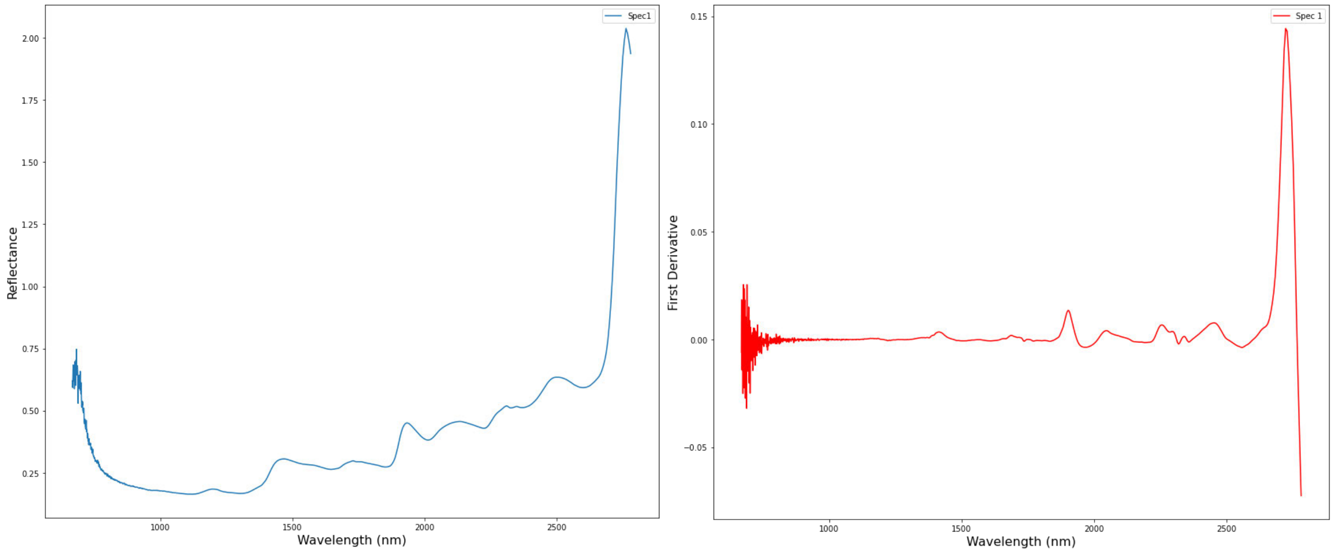


Figure 4. 2: Reflectance and first derivative spectra plot of an N spectral measurement.

The spectrometer device was calibrated using OPUS software using the Opus Manual. The S/N ratio root means square (RMS) was 1.513029, and the S/N peak-to-peak (PP) was 0.142029. To minimize BRDF effects and atmospheric perturbations, spectral measurements were acquired on clear days between 10:00 and 14:00 h central African time (Gara *et al.*, 2018). A 60 cm fibre optic cable attached to a pistol grip and an external trigger was used to collect each leaf sample's spectra to minimize spectral noise. Two scans were performed on both the adaxial and abaxial sides to provide a representative sample. A spectralon white reference panel coated with a barium sulphate of known reflectivity was used to calibrate the sensor. Calibration was performed continuously after every 15 to 20 measurements to offset any changes to the room environment. Scans of the leaf veins, midrib, and the petiole were avoided to reduce background scattering and noise (Gara *et al.*, 2018). Figure 4.3 illustrates how each vertical canopy position was determined across a sapling profile. In total, 1080 spectral reflectance measurements were collected (135 plants x 4 canopy layers x 2 sides) (Figure 4.3). To ensure the ratio of direct to diffuse incoming solar radiation was constant, the measurement setup maintained a steady and consistent view angle when measuring each sampling in an open area (Darvishzadeh *et al.*, 2008).

4.2.4 Wet chemistry analysis

Foliar samples were taken from fully expanded leaves from the top third of all 25 trees in the sample plots for foliar diagnostics. All samples were dried, weighed and analysed for physio-chemical properties and expressed by leaf concentration (%/dry weight). All laboratory tests were executed using standard measures as described by (Donkin *et al.*, 1993a, 1993b).

4.3 Statistical analysis

4.3.1 Spectra evaluation & noise removal

Before chemometric analysis, all spectral reflectance data were analysed for noise and averaged per leaf canopy position and pot. All spectra were transformed from reflectance to the first derivative (Tsai & Philpot, 1998). A moving second-order polynomial *Savitsky* filter was applied with a window size of 5 to all spectra to reduce noise in the dataset (Savitzky & Golay, 1964; Tsai & Philpot, 1998).

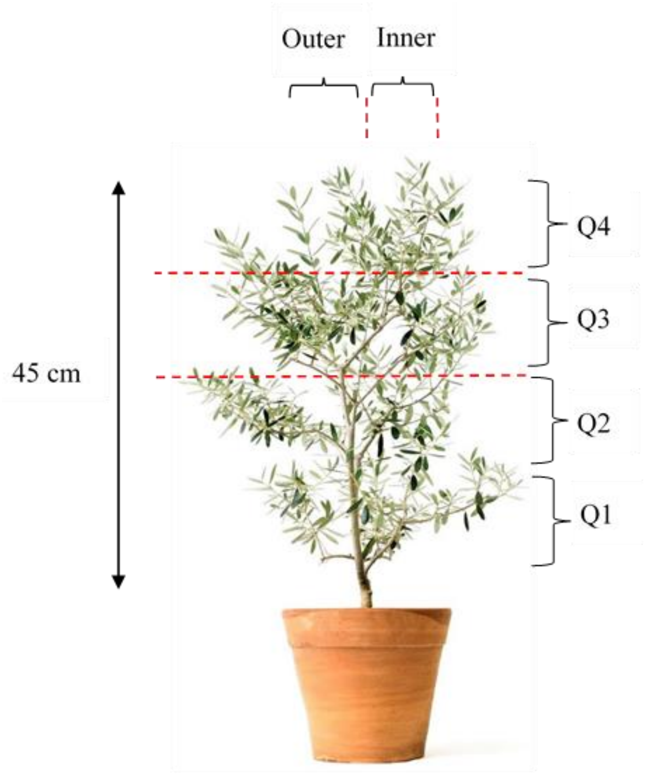


Figure 4. 3: Illustration of the four canopy layers (Q1, Q2, Q3, Q4) with outer and inner tree segmentations.

4.3.2. Partial least squares regression

Partial least squares regression (PLSR) was the preferred model approach due to the collinearity among input variables as the number of input variables is large, relative to the number of observations (Martin *et al.*, 2018). A large amount of research has suggested using PLS analysis as a benchmark for detecting chemometric techniques (Hansen & Schjoerring, 2003; Norgaard *et al.*, 2000). Moreover, previous research successfully predicted plant nutrient data combined with NIR spectral measurements using PLSR (Abdel-Rahman *et al.*, 2017; Gara *et al.*, 2018; Ge *et al.*, 2019; Meacham-Hensold *et al.*, 2019). A PLSR transforms spectral data from its original form to eigenvectors and calculates the covariance between the response and predictor variables (Wang *et al.*, 2019). Kiala *et al.* (2016) explain PLSR in a three-step process. In the first step, decomposition of the independent variables and the response variable. The second step is the prediction or expression of *Y-values* using *X-values* which as the regression coefficient and an error matrix of the relationship between *Y-values* and *X-values*. In the third step, the algorithm uses the predicted *Y-values* to produce a predictive model of the subsequent response variable (Kiala *et al.*, 2016). PLSR is an extension of the multiple linear regression model, which combines both PCA and multiple linear regression (Abdi, 2010).

PLSR was implemented in RStudio statistical software (Team, 2020) using the "caret" and "prospectr" libraries. In simple terms, the goal of PLSR is to predict a set of dependent variables from a set of independent variables (Abdi, 2010). The principle behind PLS is to search for a set of latent variables (Kembhavi *et al.*, 2011; Zhao *et al.*, 2013).

4.3.3 Accuracy assessment

The final dataset was split for each treatment into 70 % training and 30 % test data (Gara *et al.*, 2018). The R^2 was used to indicate the percentage of variation within the data. However, in this study the adjusted R^2 predictions was preferred as it reduces the number of predictors in the model, which helps overcome problems of overfitting (Huang et al., 2019; Ward et al., 2019). The adjusted R^2 formula is:

$$\mathrm{R}^2_{\mathrm{adjusted}} = 1 - \left[\frac{(1-\mathrm{R}^2)(\mathrm{n}-1)}{\mathrm{n}-\mathrm{k}-1}\right]$$

where R^2 is the sample R-square, k is the number of predictors, and *n* is the total sample size. The RMSE and the mean absolute error (MAE) was used to calculate the amount of variation between predicted and observed values, values closer to 0, predicted better results (Wang *et al.*, 2019). The RMSE formula is:

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

where y

$$RMSE[\%\} = \frac{RMSE}{\bar{Y}} \times 100$$

4.4 Results

4.4.1 Leaf canopy positions

An essential step in this study is to predict which canopy level is best to measure. In this study, four vertical canopy levels (Q1, Q2, Q3, Q4) were measured to ensure total coverage of the entire sapling. Figure 4.4 shows a line graph and bar graph of each biochemical correlation with each position. In the line graph the Y-axis shows the R^2 result, and the X-axis shows the nutrient. Each line in the graph represents a vertical canopy position. The vertical spread of the line data shows the variability of the results across each nutrient. For example, the orange colour line is vertical canopy position (2) which highly correlates with the nutrient sodium. In the bar graph the Y-axis shows the R^2 result, and the X-axis shows the nutrient. Each bar in the graph represents the vertical canopy position. The colour at the top of the bar shows the vertical canopy position with the highest correlation for a nutrient. The bar graph does not show the spread of data but more importantly the highest correlation result for each nutrient. For example, the orange block on top of the red block shows that sodium predicts the highest in vertical canopy position (2). Overall, Q2 best predicted the macronutrients: N, P, K, Ca, Mg, and Na. The highest correlations for N and Ca were found in Q3 with an R^2 of 0.88 and 0.40, respectively. The highest correlations for P, Mg, and Na were found in Q2 with R^2's of 0.17, 0.59, and 0.82, respectively. K predicted highest in Q1 with an R^2 of 0.54.

Table 4.2 shows the R^2, MAE and RMSE results of each vertical canopy position and for each biochemical. Overall, Q4 best-predicted micronutrients: Fe, Mn, Cu, Zn and B. The highest correlations for Fe, Mn and B were found in Q3 and Q4 with R^2's of 0.31, 0.29, and 0.58, respectively. Cu and Zn predicted highest in Q1 with R^2's of 0.74 and 0.64, respectively (Figure 4.4).

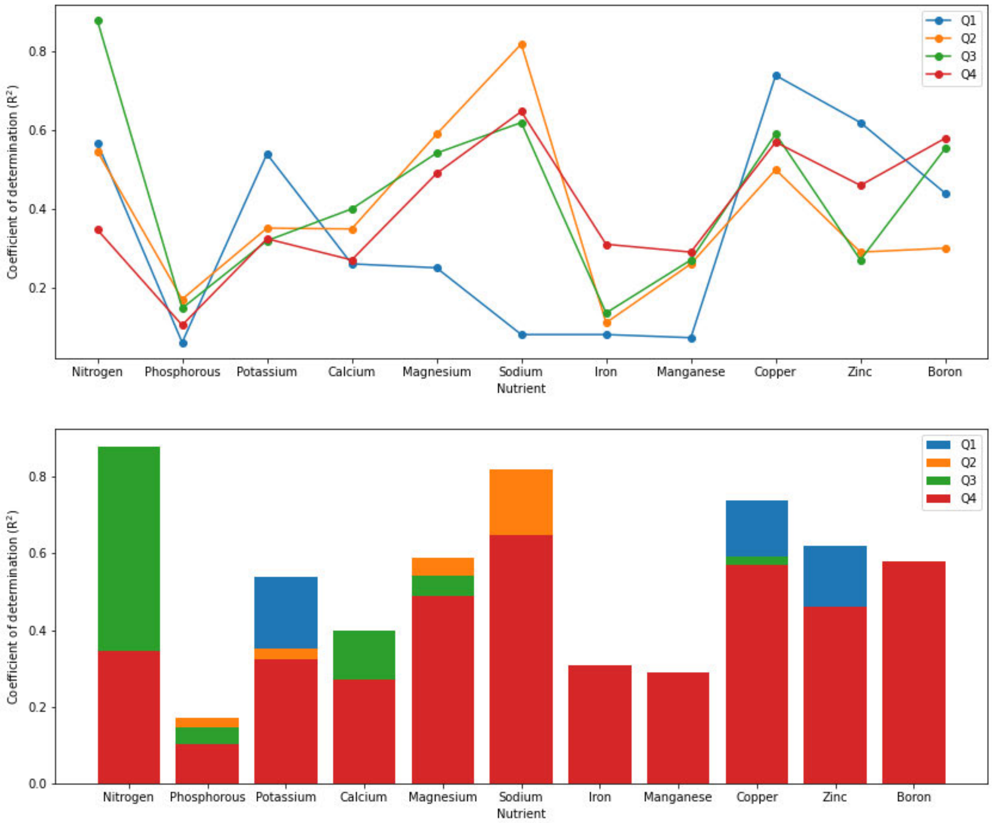


Figure 4. 4: Line and bar graph showing each biochemical correlation at each canopy leaf position. Q1 to Q4 represents each vertical canopy leaf position.

4.4.2 Macronutrients and micronutrients predictions using PLSR

Table 4.1 shows R^2 results for all nutrients with first and second derivatives, where the first derivative shows the direction, a point is moving at any point on the curve whilst the second derivative shows how the direction is changing based on the slope of the tangent. To strengthen the results, it was important to show MAE values combined with predicted versus observed graphs. MAE values show how measured values differ from actual real values. Figure 4.5 and figure 4.6 show the relationship between predicted versus observed values using a 95% confidence interval for macronutrients and micronutrients. The 95% confidence interval shows the mean estimate for each nutrient and the estimate of the range in variation for a preceding test. Overall, the accuracies for the detection of macronutrients outperformed micronutrients. Prediction results produced R^2's within a range of 0.16 to 0.92 and 0.16 to 0.87 for micronutrients and macronutrients using a first derivative transformation, respectively. Also, RMSE values ranged between 0.02 % to 0.19 % and 1.01 ppm to 1289 ppm for macronutrients and micronutrients, respectively (Table 4.1). Prediction results produced R^2's within a range of 0.21 to 0.92 and 0.23 to 0.87 for micronutrients and macronutrients using a second derivative transformation, respectively. Also, RMSE values ranged between 0.02 % to 0.41 % and 1.14 ppm to 577.4 ppm for macronutrients and micronutrients, respectively (Table 4.1)(Figure 4.5).

Furthermore, the majority (82 %) of the biochemicals (N, K, Ca, S, Fe, B, Mn, Zn) performed better using a smoothing window equal to 3 and segment break equal to 5. The minority (18 %) of the biochemicals, specifically P and Mg, performed better using a smoothing window equal to 11, 13 and segment break equal to 13, 15, respectively (Table 4.1). This study predicted N, S, Cu, and B with an R^2 above 0.80 (Figure 4.6).

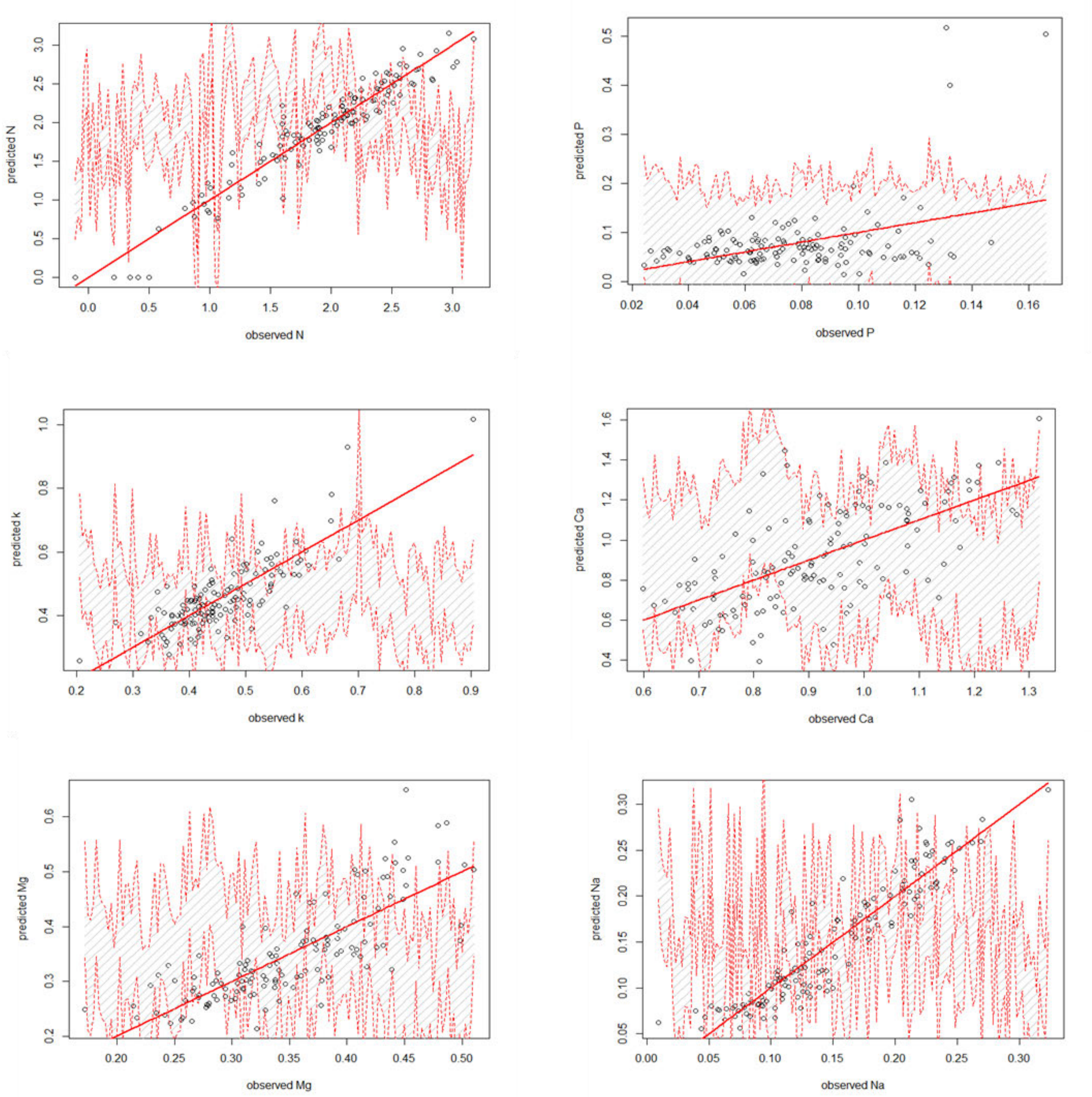


Figure 4. 5: One-to-one relationship (g/kg) with a 95% confidence interval between predicted and observed macronutrients: N, P, K, Ca, Mg, and Na using PLSR with all spectra averaged for each nutrient.

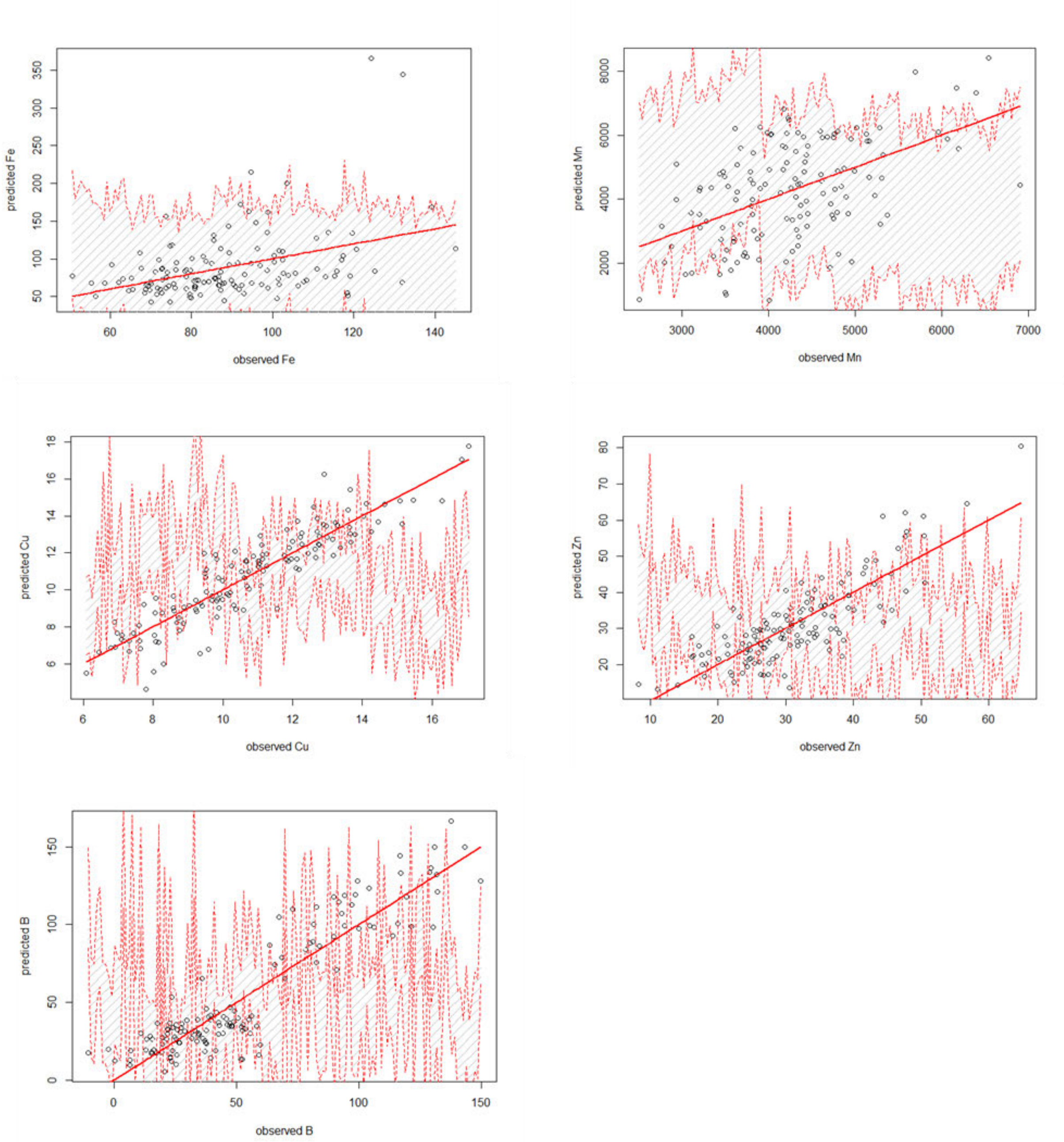


Figure 4. 6: One-to-one relationship (g/kg) with a 95% confidence between predicted and observed micronutrients: Fe, Mn, Cu, Zn, and B using PLSR with all spectra averaged for each nutrient.

Table 4. 1: R^2, MAE and RMSE's of all nutrients averaged

	Biochemical	DER	*R2*	MAE	RMSE	RMSE %
Macronutrients	Nitrogen	1st	0.91	0.38	0.53	19
		2nd	0.57	0.37	0.42	41
	Phosphorous	1st	0.16	0.04	0.06	6
		2nd	0.21	0.04	0.06	6
	Potassium	1st	0.66	0.07	0.09	6
		2nd	0.71	0.08	0.06	6
	Calcium	1st	0.40	0.17	0.21	19
		2nd	0.37	0.17	0.20	20
	Magnesium	1st	0.63	0.04	0.06	5
		2nd	0.67	0.05	0.05	5
	Sodium	1st	0.86	0.03	0.04	2
		2nd	0.86	0.04	0.02	2
Micronutrients	Iron	1st	0.16	27.63	39.97	4163
		2nd	0.23	28.13	39.80	3980
	Manganese	1st	0.26	1261.16	1465.50	128900
		2nd	0.30	1304.01	1306.00	57740
	Copper	1st	0.83	1.46	1.91	101
		2nd	0.84	1.45	1.03	114
	Zinc	1st	0.67	7.77	9.70	691
		2nd	0.68	7.62	6.42	642
	Boron	1st	0.84	26.62	33.78	1740
		2nd	0.75	25.19	20.20	1629

R^2, MAE, and RMSE of first and second derivatives for each biochemical with all spectra averaged. *DER = derivative.

Table 4. 2: R^2, MAE, and RMSE at the different leaf positions

	Biochemical		Q1	Q2	Q3	Q4
Macronutrients	Nitrogen	*R2*	0.57	0.55	0.88*	0.35
		MAE	0.40	0.40	0.43	0.42
		RMSE	0.55	0.56	0 56	0.58
	Phosphorous	*R2*	0.06	0.17*	0 15	0.10
		MAE	0.04	0.04	0.04	0.03
		RMSE	0.06	0.06	0.06	0.06
	Potassium	*R2*	0.54*	0.35	0 32	0.32
		MAE	0.08	0.07	0.07	0.07
		RMSE	0.10	0.09	0 10	0.10
	Calcium	*R2*	0.26	0.35	0.40*	0.27
		MAE	0.19	0.17	0 18	0.18
		RMSE	0.23	0.22	0 22	0.23
	Magnesium	*R2*	0.25	0.59*	0 54	0.49
		MAE	0.06	0.05	0.49	0.05
		RMSE	0.08	0.06	0.63	0.07
	Sodium	*R2*	0.08	0.82*	0.62	0.65
		MAE	0.06	0.04	0.04	0.04
		RMSE	0.06	0.05	0.05	0.05
Micronutrients	Iron	*R2*	0.08	0.11	0 14	0.31*
		MAE	28.78	27.49	26.65	25.87
		RMSE	41.10	39.93	39.22	36.94
	Manganese	*R2*	0.07	0.26	0 27	0.29*
		MAE	1308.51	1236.18	1257.79	1219.55
		RMSE	1543.08	1471.29	1478 29	1429.19
	Copper	*R2*	0.74*	0.50	0 59	0.57
		MAE	1.80	1.67	1.65	1.64
		RMSE	2.27	2.14	2 10	2.12
	Zinc	*R2*	0.62*	0.29	0 27	0.46
		MAE	8.46	7.65	8.02	7.28
		RMSE	10.99	10.21	10.23	9.46
	Boron	*R2*	0.44	0.30	0 55	0.58*
		MAE	31.47	28.49	25.84	25.75
		RMSE	38.51	35.14	32.62	32.56

The asterisk (*) symbol Indicates the best performing canopy position for each biochemical. R^2, MAE and RMSE at the different leaf positions (Q1 to Q4) for each biochemical using the best smoothing window and segment break in the first derivative.

4.4.3 Variable importance within each leaf canopy position

Assessing the VI of which wavelengths best discriminate each nutrient within each leaf canopy position was also essential in this study. Table 4.3 explicitly shows the most important wavelengths (nm) for detecting each nutrient (N, P, K, Ca, Mg, Na, Fe, Mn, Cu, Zn, B) within in each vertical canopy position (Q1, Q2, Q3, Q4). Variable importance of each vertical canopy level was assessed using the "varImp" function within the "caret" package. This study predicted most nutrients within the SWIR region. Only N, P, Mn could be more discernible within the NIR region of the electromagnetic spectrum (Table 4.3).

Table 4. 3: Variable important wavelength for each nutrient within each canopy position

	Biochemical	Canopy Position	Variable important wavelengths (nm)									
Macronutrients	Nitrogen	Q1	1672	1671	1673	1669	1674	1668	1675	1667	1676	1666
		Q2	1671	1669	1672	1668	1673	1667	1674	1666	1675	1665
		Q3	863	864	863	864	904	1651	1650	1649	903	864
		Q4	1667	1668	1666	1669	1665	1671	1664	1672	1191	1191
	Phosphorous	Q1	2224	2222	2116	2115	2118	2113	2111	2120	2225	2110
		Q2	2321	2323	2319	2325	2327	2317	1834	1832	1835	1831
		Q3	1224	1265	1275	1090	1276	1091	1265	1223	1266	1224
		Q4	1224	1265	1275	1090	1276	1091	1265	1223	1266	1224
	Potassium	Q1	2225	2224	2222	2227	1653	1652	1655	2220	1651	1656
		Q2	2321	2319	1778	1777	2323	1681	1682	1779	1776	1684
		Q3	1680	1679	1678	1681	1677	1773	1775	1772	1776	1676
		Q4	1681	1680	1682	1679	1684	1678	2247	2245	2249	2251
	Calcium	Q1	1689	1690	1688	1691	1687	1692	1686	1693	1695	1685
		Q2	2266	2268	2264	2262	2270	2260	2258	2256	2251	2253
		Q3	940	940	939	940	961	2195	941	2197	2193	937
		Q4	2193	2192	2195	2190	2197	2188	2199	2186	911	2201
	Magnesium	Q1	2210	2212	2208	2214	2207	2205	1692	1693	2203	1691
		Q2	1688	1687	1689	1686	1690	1685	1684	1691	1682	1692
		Q3	1685	1686	1684	1687	1682	1688	1681	1689	1680	1793
		Q4	1689	1688	1690	1687	1691	1686	1692	1685	1684	1693
	Sodium	Q1	2224	2222	2220	2225	2218	2106	2108	2104	2110	2103
		Q2	1712	1714	1715	1711	1716	1710	1717	1709	1718	845
		Q3	1678	1677	1679	1676	1680	1681	1675	1682	1674	1684
		Q4	1680	1679	1681	1678	1682	1677	1684	1676	1685	1675
Micronutrients	Iron	Q1	2225	2224	2227	2222	2229	2372	1291	1763	1764	1765
		Q2	1769	1770	1767	1771	1772	1773	1766	1775	1776	1765
		Q3	1675	1674	1676	1677	1673	1678	1672	1679	1680	1671
		Q4	1679	1678	1680	1677	1681	1676	1682	1684	1675	1685
	Manganese	Q1	931	931	950	950	932	941	1117	1117	999	1283
		Q2	1265	1265	1222	1222	1264	1257	931	1263	1266	1246
		Q3	1246	1257	1246	1245	1205	1205	1257	1263	1258	1222
		Q4	1246	1246	1245	1247	1248	985	1245	1223	1265	1205
	Copper	Q1	963	999	964	1000	999	987	988	963	987	966
		Q2	1668	1667	1669	1666	1671	1665	1672	1664	1673	1663
		Q3	1664	1663	1665	1662	1666	1667	1668	1671	1669	1661
		Q4	1675	1674	1673	1676	1672	1677	1671	1669	1678	1665
	Zinc	Q1	1651	1652	1650	1653	1649	1655	1648	1647	1656	1646
		Q2	1783	1784	1830	1782	1831	1828	1786	1832	1827	1834
		Q3	1783	1784	1782	1786	1781	1787	1794	1795	1788	1793
		Q4	1684	1682	1685	1681	2245	2243	1686	2247	2241	2249
	Boron	Q1	2224	2225	2222	1652	1651	1650	1653	2220	1649	842
		Q2	1682	1684	1681	1685	1680	2319	1679	1686	2321	1771
		Q3	1678	1677	1679	1676	1680	1681	1675	1682	1674	1684
		Q4	1678	1679	1677	1680	1681	1676	1682	1675	1684	1674

Table 4.4 shows the calculated means and standard deviations (SD) of R^2's and the distribution of data around the mean R^2 for each biochemical. The standard deviation was significant in explaining the variance when measuring each canopy position. The lowest SD was 0.04 R^2's for Cu, and the highest SD was 0.16 R^2's for N, whereas the average mean and SD was 0.39 and 0.10 for all biochemicals combined, respectively (Table 4.4).

Table 4. 4: Mean R^2 and standard deviation of leaf canopy positioning for each biochemical

	Biochemical		Q1	Q2	Q3	Q4	Mean R^2	SD R^2
Macronutrients	Nitrogen	*R2*	0.57	0.55	0.88*	0.35	0.49	0.12
		MAE	0.40	0.40	0.43	0.42	0.41	0.02
		RMSE	0.55	0.56	0.56	0.58	0.56	0.01
	Phosphorous	*R2*	0.06	0.17*	0.15	0.10	0.10	0.04
		MAE	0.04	0.04	0.04	0.03	0.04	0.00
		RMSE	0.06	0.06	0.06	0.06	0.06	0.00
	Potassium	*R2*	0.54*	0.35	0.32	0.32	0.33	0.02
		MAE	0.08	0.07	0.07	0.07	0.07	0.00
		RMSE	0.10	0.09	0.10	0.10	0.10	0.00
	Calcium	*R2*	0.26	0.35	0.40*	0.27	0.29	0.05
		MAE	0.19	0.17	0.18	0.18	0.18	0.01
		RMSE	0.23	0.22	0.22	0.23	0.22	0.01
	Magnesium	*R2*	0.25	0.59*	0.54	0.49	0.43	0.16
		MAE	0.06	0.05	0.49	0.05	0.16	0.22
		RMSE	0.08	0.06	0.63	0.07	0.21	0.28
	Sodium	*R2*	0.08	0.82*	0.62	0.65	0.45	0.32
		MAE	0.06	0.04	0.04	0.04	0.04	0.01
		RMSE	0.06	0.05	0.05	0.05	0.05	0.01
Micronutrients	Iron	*R2*	0.08	0.11	0.14	0.31*	0.11	0.03
		MAE	28.78	27.49	26.65	25.87	27.20	1.25
		RMSE	41.10	39.93	39.22	36.94	39.30	1.75
	Manganese	*R2*	0.07	0.26	0.27	0.29*	0.20	0.11
		MAE	1308.51	1236.18	1257.79	1219.55	1255.51	38.65
		RMSE	1543.08	1471.29	1478.29	1429.19	1480.46	47.04
	Copper	*R2*	0.74*	0.50	0.59	0.57	0.55	0.05
		MAE	1.80	1.67	1.65	1.64	1.69	0.07
		RMSE	2.27	2.14	2.10	2.12	2.16	0.08
	Zinc	*R2*	0.62*	0.29	0.27	0.46	0.34	0.10
		MAE	8.46	7.65	8.02	7.28	7.85	0.50
		RMSE	10.99	10.21	10.23	9.46	10.22	0.62
	Boron	*R2*	0.44	0.30	0.55	0.58*	0.43	0.13
		MAE	31.47	28.49	25.84	25.75	27.89	2.71
		RMSE	38.51	35.14	32.62	32.56	34.71	2.81

4.5 DISCUSSION

This study detected foliar nutrients using NIRS and VCP using PLSR in 135 *Eucalyptus grandis* saplings. The study included an additional position of measurement (Q4), which enabled further investigation into which part of the sapling denoted the best place to determine nutrient content using a non-destructive approach. Understanding the concentration and distribution of active leaf biochemicals within vegetation canopies significantly influences gross primary production (GPP) and is empirical for developing accurate sapling nutrient prediction models (Alton & North, 2007). NIR technology offers forestry companies a robust, simplified, and operational alternative for retrieving specific biochemical information to inform sapling breeding programs and mill pulping requirements.

The objective of this study was to investigate the best sampling position to measure macronutrients and micronutrients at different canopy positions by segmenting saplings into four quartiles using NIRS. The idea was to select and inform forest nursery managers on the best sampling position, which will reduce sampling time and still provide an accurate account of the nutrient content in the entire sapling. This study used a similar method reported by Gara *et al.* (2018) and adapted it by adding a fourth quartile which further segmented the sapling. However, Gara *et al.* (2018) studied the shifting of wavelengths from the lower to upper portions of the tree and measured N, Chl, specific leaf area, C, and effective water thickness. As a result, their study used significance ρ-value at 0.1, 0.05, and 0.01 to compare the lower to upper regions. Whereas this study specifically observed the trees most effective macronutrients and micronutrients between all leaf canopy positions. The finding in this study shows that macronutrients and micronutrients are better predicted in Q2 and Q4, respectively.

However, when observing the overall predictions of each nutrient, P produced significantly low results within NPK compared to other studies using PLSR with all spectra averaged. Commonly, NPK has high prediction accuracies when compared to earlier research. Similar studies such as Abdel-Rahman *et al.* (2017) also predicted NPK using PLSR, which produced significantly higher P content results than this study. Nonetheless, the results in this study are comparable to other studies that determined valuable nutrients in young trees using remotely sensed data (Mutowo *et al.*, 2018, 2019). For instance, Mutowo *et al.* (2019) achieved an R^2 of > 0.90 for N with a 2-canopy level (top and medium) using Sentinel-2 imagery (443 nm to 2190 nm) and weighted means (RMSE) approach. Asner *et al.* (2011) achieved an R^2 of 0.59 for N and 0.51 for P. Menesatti *et al.* (2010) measured N, P, K, Ca, Mg, Fe, Zn, and Mn in citrus leaves using a VI-NIR spectrometer (310 to 1100 nm) and PLS for prediction. As a result, like this study, their study also found significantly low P accuracies during testing; however, their study achieved their highest accuracy for K compared to our study.

A potential limitation for accurately sampling at leaf level studies understands the process of nutrient re-translocation at the species level. Nutrient re-translocation is a process when a plant removes nutrients from plant tissue into the perennial part of the plant before senescence as a conservation mechanism (Fife *et al.*, 2008). As a result, lower nutrient content levels could be in the leaves at the acquisition time. For example, Fife *et al.* (2008) found significant differences in the amount of nutrient content in the leaves compared to the stem or trunk after taking measurements at 12 months and 22 months of age in *Acacia mearnsii* De Wild. (a N-fixing species); *Eucalyptus globulus; E. fraxinoides* H. Deane Maiden; *E. grandis* W. Hillex Maiden; *Pinus radiata*; and *Casuarina glauca* Sieber ex Spreng species.

However, this study provides evidence to sample at specific locations to find differences in nutrient content. A constraint of this study is the lack of variability in environmental conditions that would be of impact in a more natural environment. An essential step for future research would be to investigate the chemical distribution throughout the canopy of live standing trees using a methodology like this study. Detecting foliar nutrients within a forest environment can offer fast and reliable results to forestry and agricultural sectors. Future work can use such methods developed in this study for estimating wood density properties and other key traits such as cellulose and lining that would be instrumental towards the manufacturing value chain of the forestry business. The generalization of this study’s findings supports analyses of live standing trees. However, this will require future studies to understand the re-translocation of plant biochemistry within the bigger context and consider all external environmental impacts, which may influence nutrient content. Also, future studies could upscale this approach to fully grown trees in the field to examine the extent of the VCP method for the detection of critical nutrients and to understand the impact of taking measurements from the inner and outer portions of the trees. The adoption of coherent sampling methods and analytical procedures for biomonitoring to better understand plant and leaf processes is critical for the sustainability of many forestry industries worldwide (Loppi *et al.*, 1997). The success of this study promotes the use of NIR spectroscopy for determining the nutrient status of saplings. Furthermore, providing a good platform for forestry industries to rapidly detect the nutrient status of their younger plants while ensuring optimal growth of their future plantations.

4.6 Conclusion

In conclusion, the study detected and predicted foliar nutrients in *Eucalyptus grandis* saplings under nursery conditions. The study made the following findings:

- The findings of this study suggest that spectral measurements are best taken from canopy position levels Q2 and Q4 for macronutrients and micronutrients using NIRS, respectively.
- Not many studies have examined the full range of macronutrients (N, P, K, Ca, Mg, Na) and micronutrients (Fe, Mn, Cu, Zn, B) nutrient concentrations across four VCP's using NIR spectroscopy combined with an effective PLS algorithm.

Overall, the study found no distinct VCP better predicted all nutrients in this study. However, this study offers guidance to the optimal tree position suitable for producing the best detection of each macronutrient and micronutrient. The next step would be upscaling the methodologies developed in this study onto larger areas within the forest plantation using higher resolution airborne sensors to accurately detect contemporary nutrient management regimes. Lastly, finding the optimal position to take a representable sample are important for understanding the distribution of nutrient deficiencies throughout a tree canopy. In the upcoming chapter we explore airborne sensors to test the capabilities of airborne sensors and the detection accuracy of nutrients on a broader scale.

CHAPTER 5: Comparing the classification accuracy of ultra-high-resolution UAS imagery and very-high-resolution PlanetScope imagery in four nutrient management regimes using a deep learning artificial neural network

To compare the classification accuracy of ultra-high-resolution UAS imagery and very-high-resolution PlanetScope imagery in four nutrient management regimes using a deep learning artificial neural network.

This chapter was based on:

Singh L*, Mutanga O, Mafongoya P, Peerbhay KY, Dovey S (Preparation). Comparing unmanned aerial vehicle and PlanetScope imagery for classifying nutrient management regimes in commercial forestry plantations using a deep learning artificial neural network.

Abstract

Improvements in spatial resolutions of remotely sensed data can improve crop health status mapping. Coarser-resolution imagery conceivably confounds nutrient assessments. Proximal sensors provide incredibly high spectral resolutions; however, they are not practical for assessing large compartment forest canopies. This study aims to evaluate the capabilities of ultra-high-resolution UAS imagery and very-high-resolution PlanetScope imagery combined with a deep learning ANN to classify four nutrient management regimes in a *Eucalyptus* compartment forest. Using a confusion matrix both pixel-based image sources successfully classified the four nutrient management regimes with an OA above 80%, KHAT above 75% using four hidden layers and 30 epochs. UAS imagery performed better than PlanetScope imagery; however, the results show the potential of very-high-resolution PlanetScope imagery closely matching the results of ultra-high-resolution UAS imagery. Advancements in remote sensing and machine learning provide resourceful in improving the effectiveness and efficiency of nutrient assessments.

Keywords: PlanetScope, deep learning, forestry, foliar biochemicals

5.1 Introduction

UAS technology and very-high-resolution satellite imagery have shown potential to assist in maximising global forest plantations returns (Dash *et al.*, 2018; Krishna, 2018; Taddese *et al.*, 2020). Recently, research has focused significantly on developing remote sensing methods using very-high-resolution imagery for forest management (Barbedo, 2019; Mutanga *et al.*, 2016). UAS's and very-high-resolution satellite imagery can provide precise, timely, cost-effective, and versatile services for crop management (Chen *et al.*, 2019; Krishna, 2018). Advanced UAS photogrammetric sensors such as the Red Edge-MX Micasense camera provide a host of crop management services such as precision agriculture, crop health, irrigation monitoring and detecting pests and diseases are essential information for making an informed decision (Berger *et al.*, 2020; Fareed & Rehman, 2020; Krishna, 2018; Oliveira *et al.*, 2020). The Micasense camera provides centre meter pixel accuracy, designed for precision agricultural applications (Babaeian *et al.*, 2021). The Micasense camera is small and lighter for UAS drone assembly and is generally cheaper than actual hyperspectral cameras (Suomalainen *et al.*, 2021). To date, not many studies have explored the capabilities of new-generation UAS technology and very-high-resolution satellite imagery to classify nutrient management regimes in commercial forestry compartments. Intelligent nutrient management regimes provide a valuable competitive edge for commercial forestry growers. Furthermore, practical nutrient assessments provide essential information for forest mensuration, forest health and estimating yield potential.

Globally, forests play an integral role in climatic stabilization and sequestering C in biomass and soil (Yousefpour *et al.*, 2018). Research suggests that forest management is a primary driver of biodiversity, climate change, ecosystem health and ecosystem services (Kahl & Bauhus, 2014; Yousefpour *et al.*, 2018). Problems exist when commercial forests become nutrient deficient; effectively, time is lost, resulting in no return on investment for commercial forestry companies. Hence, the supply of adequate nutrients to trees directly affects the income potential of forestry industries. In this regard, a crop's response to standard but essential inputs such as nutrients (fertilizer), water and other amendments varies immensely (Krishna, 2018). UAS's provide ultra-high-resolution imagery with centimetre accuracy, enabling more discrete information for subsequent analysis. UAS's have restrictions to local area applications (<10 km^2), whilst very-high-resolution satellite imagery provides extensive area coverage of the area of interest (AOI)(Johansen *et al.*, 2020). The standard procedure is to validate UAS imagery with ground data enhancing the reliability of the data collected. UAS's and very high-resolution satellite imagery offers a wide range of services to farmers and reduce drudgery (Krishna, 2018). UAS's and very-high-resolution satellite imagery offers non-destructive assessments which disturb soil neither its biotic factors nor the physicochemical properties of the targeted AOI (Krishna, 2018).

Furthermore, the lack of suitable plantation nutrient management regimes for maintaining balanced soil nutrients results in a decline in tree productivity, dystrophic soils, and the interruption of natural nutrient cycling processes (Crous *et al.*, 2007). Long-term changes in soil nutrient levels have been challenging to detect in previous studies directly attributable to management induced losses (Krishna, 2018). More specifically, nutrient-related productivity decline in plantation forests can function the accuracy of sampling methods that require calibration with plantation forest soils. Detecting nutrient-related productivity decline in plantation forests can be challenging in complex plantation environments. Traditional sampling techniques for evaluating different nutrient management regimes have to wait until harvest; hence the early detection of nutrient stress cannot be tracked (Cai *et al.*, 2019). It is important to understand the cost implications of employing these systems operationally in a practical environment. The cost of these platforms and associated sensors are rapidly decreasing while their capabilities and sophistication are constantly improving (Dash *et al.*, 2018). There are many factors to consider when comparing the cost-benefit of UAS versus satellite imagery such as optical quality (low, medium, high), minimum order area, radiometric resolution, spatial resolution, data volume per hectare, revisit time, and price per hectare. (Sozzi *et al.*, 2021). However, satellite imagery is characterized by a lower break-even point in hectares compared to UAS imagery. Furthermore, satellite imagery most commonly requires a minimum order compared to UAS image that is flown based on the parameters of the AOI.

Previous studies have demonstrated the use of ultra-high-resolution imagery (>2m) on airborne UAS platforms for mapping biophysical and biochemical properties of crops using mainly vegetation spectral indices (Cai *et al.*, 2019; Dash *et al.*, 2018; Lu *et al.*, 2018; Lussem *et al.*, 2019; Osco *et al.*, 2020a) and machine learning (Gracia-Romero *et al.*, 2020; Han & Watchareeruetai, 2019; Zha *et al.*, 2020). However, there are substantial costs associated with acquiring centimetre accurate UAS imagery, which generally covers a smaller area than satellite imagery. Therefore, satellite imagery still plays a vital role in precision agriculture mapping applications by mapping at larger spatial scales, technically not achievable for UASs (Johansen *et al.*, 2020). Later studies have revealed that UAS data can complement satellite imagery (Dash *et al.*, 2018). For example, Johansen *et al.* (2020) tested UAS and Worldview-3 imagery capabilities to map the condition of macadamia tree crowns using the RF classifier.

As a result, Worldview-3 imagery underperformed compared to UAS imagery; however, the authors correctly classified five condition categories with out of bag accuracy above 98.5%. Dash *et al.* (2018) tested the sensitivity of multispectral imagery collected from time-series UAS and satellite imagery to detect herbicide-induced stress in a carefully controlled experiment carried out in a mature Pinus radiata D. Don plantation in New Zealand. The authors compared the performance of spectral indices Normalized Difference Vegetation Index (NDVI), green normalized difference vegetation index (GNDVI), and the red edge normalized difference vegetation index (RENDVI). The authors found that UAS imagery provides superior sensitivity to physiological stress than satellite imagery.

In summary, many studies have mapped and monitored the nutrient status of crops such as sunflower, cotton, wheat, grass, turfgrass, corn, potato, sugar beet, rice, soybean, and canola using multispectral, hyperspectral and red-green-blue (RGB) UAS imagery (Barbedo, 2019; Dash et al., 2018; Gracia-Romero et al., 2020; Jay et al., 2019; Lussem et al., 2019). Furthermore, the majority of the studies mentioned above predicted N concentration, P and K using spectral vegetation indices such as normalized difference red edge (NDRE), NDVI, normalized green-red difference index (NGRDI), PLSR, RF and SVMR (Barbedo, 2019). However, few studies have compared the performance of ultra-high-resolution UAS's and very-high-resolution PlanetScope imagery for classifying nutrient management regimes combined with a deep learning ANN in a commercial forestry compartment. Hence, the aim of this is to evaluate the capabilities of ultra-high-resolution UAS imagery and very-high-resolution PlanetScope imagery combined with a deep learning ANN to classify four nutrient management regimes in a Eucalyptus compartment forest. This study will provide exploratory clarity to the trade-offs of two very-high-resolution remote sensing platforms for optimizing forest management practices. Climate Smart Forestry recognizes synergies among climate change mitigation and other forest benefits, essentially optimizing forest management practices to contribute to climate change mitigation to contribute to the ambitious COP 21 goals (Yousefpour et al., 2018).

5.2 Materials and methods

5.2.1 Study area

The study area is located at the Clan Sappi plantation in KwaZulu-Natal, South Africa (29°20'6" S, 30°27'13" E)(Figure 5.1). The location of the nutrient management regime in the mist belt grassland bioregion of the KwaZulu-Natal, Midlands, occupies 1.4ha. The compartment consisted of 3-4-year-old *Eucalyptus grandis x Eucalyptus urophylla* tree species variety, *Eucalyptus grandis* (flooded gum or rose gum) is native to New South Wales, and *Eucalyptus urophylla* (Timor white gum, Timor Mountain gum, Popo or Ampupu) is native to the Indonesian Archipelago and Timor. The mean annual precipitation (MAP) range for the study area is 1000-1100 mm, and the mean annual temperature (MAT) is 18.1 degrees Celsius, situated at 788 m above sea level (ASL) (Mucina & Rutherford, 2006). The study area lithology consists of shale and dolerite, and the soil type is Inanda which has a silty clay soil texture (Mucina & Rutherford, 2006). The soil texture comprises of 56% sand and 44% of clay. The study area was preferred due to anthropogenic influences, a controlled environment, and typical site characteristics for compartment planting of the tree species.

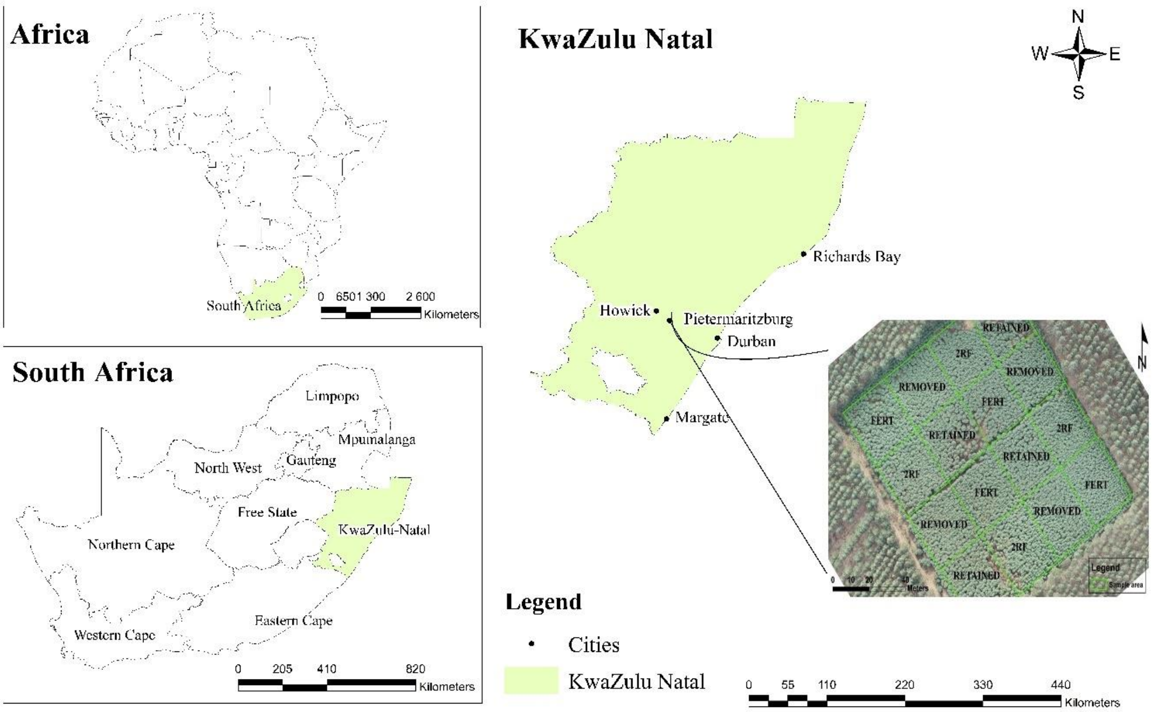


Figure 5. 1: The location of the study area in a plantation forest of KwaZulu Natal, Midlands, South Africa, and the nutrient management regime formation.

5.2.2 Experimental design & field data

The nutrient management regime trial comprised four treatments with four replicates arranged in a Latin square design. Trees were planted at 1m x 1m espacement (i.e., 10 000 sph), with plots of 10 x 15 trees (750 m^2) and growth measurement plots of 5 x 5 trees centrally located within each plot (Figure 5.1). Figure 5.1 illustrates the polygons depicting each nutrient management regime. A series of 16 ground control points (GCPs) were established at the centre of each polygon represented in Figure 5.1 within the study area. Each GCP was located using a Trimble GeoExplorer 6000 series GeoXH global positioning system (GPS). The GPS has a centimetre-level accuracy of less than 10 cm, which improves accuracy when locating each GCP (Trimble Navigation Ltd., Sunnyvale, CA, USA). The GCP's were used to accurately geo-rectify the UAS imagery (Dash *et al.*, 2018). The four treatments consist of:

1. Residue retention (RET): All biomass (stem, branches, and foliage) is retained and dispersed over the plot (to simulate full nutrient retention).

2. Nutrient removal (REM): All biomass is removed with no fertilizer application.

3. Nutrient replacement (FERT): All biomass is removed, followed by fertilizer addition to replace lost nutrients. The annual nutrient loss was calculated and replaced with granular fertilizer at around 110% of the removed elements; dispersed over the whole-plot area. Fertilizer addition commenced after the first harvest (first coppice crop) to simulate intensive nutrient removal with management amelioration.

4. Rehabilitation (2RF): Same as treatment 2, but when soil depletion is confirmed, this treatment will test the ability of fertilizer to ameliorate soil fertility following re-establishment of all plots to standard operational espacement and to recover naturally or through fertilizer addition. Thus, the first three treatments were implemented for the study's first phase.

5.2.3 UAS imagery & pre-processing

UAS imagery was collected under cloudless (0%) conditions between 10:00 and 14:00 h central African time using DJI Phantom 4 Pro quadcopter in October 2018. UAS imagery was collected using a Red Edge-MX Micasense narrowband multispectral camera (Micasense Inc., Seattle, WA, USA). The Red Edge-MX Micasense camera has a pixel size of 2 cm and a pixel depth of 12-bits (Lussem *et al.*, 2019). The Red Edge-MX Micasense camera has the following waveband setting (blue = 455–495 nm, green = 540–580 nm, red = 658–678 nm, red edge = 707–727 nm and NIR = 800–880 nm). The DJI Phantom 4 Pro was flown at an altitude of approximately 60 m above the AOI, resulting in a ground sample distance (GSD) of 8 cm (Dash *et al.*, 2018). The image was acquired as a snapshot of the AOI with a focal length lens of 5.5 mm and a FOV of 151.47 degrees. The camera was housed in a gimbal to ensure the nadir orientation of the camera during data collection. The sensor was calibrated using a reference panel (white reference) coated with a barium sulphate of known reflectivity. Sensor calibration took place directly before and after each flight to offset any change in the sun's atmospheric condition and irradiance (Singh *et al.*, 2017a).

The UAS image was processed using (Structure-from-Motion software PhotoScan v1.4) (AgiSoft LLC, St. Petersburg, Russia) in conjunction with the GCPs speculated in the experimental design of this study (Lussem *et al.*, 2019). Image pixels were extracted from the polygons (Figure 5.1) using the "Extract by Mask tool" in the "Spatial Analyst" toolbox in ArcGIS 10.7. Each nutrient management regime represented various pixels based on the UAS image resolution.

5.2.4 PlanetScope imagery & pre-processing

The acquisition of PlanetScope imagery coincided with the time of acquiring the UAS imagery. PlanetScope multispectral imagery was obtained under cloudless conditions in October 2018 by the supplier. PlanetScope satellite imagery consists of a constellation of individual small Dove satellites, namely a CubeSat 3U form factor. The constellation has approximately 130 satellites that image the entire earth on a daily temporal resolution. The PS2 instrument onboard the Dove satellite collects images with a frame size of 24 x 8 km and a bit depth of 12-bits. PlanetScope imagery comprises four wavebands operating over the 455-860 nm wavelength range with a spatial resolution of 3.7 m. Specific wavelength ranges are 455-515 nm, 500-590 nm, 590-670 nm, and 780-860 nm for the blue, green, red, and NIR.

PlanetScope product level 3A imagery was obtained which is radiometric, and sensor corrected by the vendor. ENVI 5.2 image processing software (L3Harris Technologies, 2020) was used to test and further pre-processed the PlanetScope imagery for any other irregularities. PlanetScope product level 3A imagery is orthorectified and projected to Universal Transverse Mercator (UTM) by the vendor, and the WGS-84 Geodetic System was used, similarly for the UAS imagery above. Similarly, pixels were extracted from the polygons (Figure 5.1) using the "Extract by Mask tool" in the "Spatial Analyst" toolbox in ArcGIS 10.7. Each nutrient management regime represented various pixels based on the PlanetScope image resolution.

Table 5. 1: UAS and PlanetScope configurations and specifications

Platform name	Platform type	Sensor type	Radiometric resolution	Spatial resolution	Wavebands	Spectral resolution (nm)
DJI Phantom 4	UAS		12-bit	8 cm/px @ 120 m	Blue	455–495
					Green	540–580

		Red Edge-MX Micasense			Red	658–678
					Red-edge	707–727
					NIR	800–880
PlanetScope	Satellite	4 band	12-bit	3.7 m/px	Blue	455-515
					Green	500-590
					Red	590-670
					NIR	780-860

5.2.5 Deep learning artificial neural network (ANN) algorithm

In general, an ANN stems from artificial intelligence, which is the ability of a sophisticated computer machine to perform human intelligence tasks (Atkinson & Tatnall, 1997; Berg & Nyström, 2018; Reichstein *et al.*, 2019). A deep learning ANN is based on constructing increasingly sophisticated hierarchical architectures using two or more hidden layers with multi-layer neurons (Reichstein *et al.*, 2019). Multi-layer neurons can cycle or loop information between different neurons, creating powerful pattern recognition abilities to learn intricate multivariate data patterns (Atkinson & Tatnall, 1997; Mutanga & Skidmore, 2004; Reichstein *et al.*, 2019). Previous studies used various deep learning ANN models such as multi-layer perceptron (MLP), radial basis function, and backpropagation in forest modelling of remotely sensed data (Atkinson & Tatnall, 1997; Liu *et al.*, 2013; Omer *et al.*, 2017; Wang *et al.*, 2009).

Figure 5.2 represents the network structure of an ANN; whereby hidden layers bridge the gap between input layers and output layers. Artificial neurons receive a set of weighted inputs to produce an output through an activation function (Atkinson & Tatnall, 1997)(Figure 5.2). An activation function is liable for converting the summed weighted input from the node into the node's activation or output to help the network learn intricate patterns in the data. Several types of activation functions such as sigmoid, rectified linear units, or hyperbolic tangents exist within the neurons of each hidden layer (Berg & Nyström, 2018). Combined with the activation function, each neuron in the ANN is assigned a bias, including the output neurons and excluding the input neurons, whilst the connections between neurons in subsequent layers are represented by matrices of weights (Berg & Nyström, 2018). The weighted input for a deep learning ANN is defined as:

$$z_j^l = \sum_k w_{jk}^l \, \sigma_{l-1}\left(z_k^{l-1}\right) + b_j^l$$

where the sum of all inputs is taken to the neuron *j* in Layer *l*, which is the number of neurons, the deep learning ANN naturally defines a recursion in terms of previous weighted inputs through the ANN. The calculation that terminates any recursion is defined as:

$$\sigma_0\left(z_j^0\right) = \mathcal{Y}_j^0 = x_j$$

More specifically, in figure 5.2, Layer 0 is the input layer which is consists of imagery (UAS and PlanetScope) and the response variable (nutrient management regimes), and Layer L is the output layer. Layer l – 1 and Layer l represent the hidden layers; in this figure 5.2, there are two hidden layers, whereas, in this study, four hidden layers were used (Berg & Nyström, 2018).

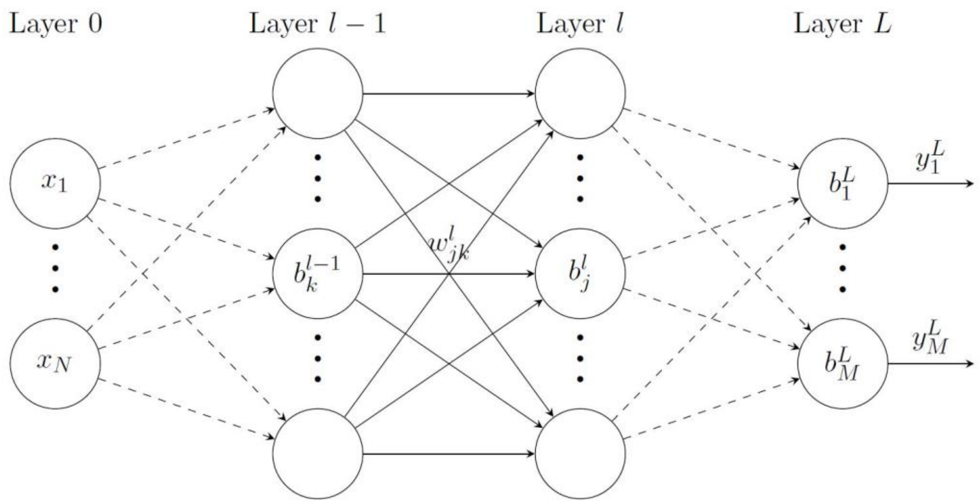


Figure 5. 2: A fully connected feedforward ANN (Berg & Nyström, 2018).

The feedforward algorithm for computing the output is defined as:

$$y^L = \sigma_L(z^L)$$

$$Z^L = w^L\sigma_L(z^{L-1}) + b^L$$

$$z^{L-1} = w^{L-1}\sigma_{L-2}(z^{L-2}) + b^{L-1}$$

$$\vdots$$

$$z^1 = w^1 x + b^1$$

The backpropagation algorithm for calculating the gradients of the cost function is defined as:

$$\delta_j^{\mathcal{L}} = \frac{\partial \mathcal{C}}{\partial \mathcal{Y}_j^{\mathcal{L}}}\, \sigma_{\mathcal{L}}'\left(\mathcal{Z}_j^{\mathcal{L}}\right), \frac{\partial \mathcal{C}}{\partial w_{jk}^l} = \mathcal{Y}_k^{l-1}\delta_j^l,$$

$\delta_j^l = \sum_k w_{kj}^{l+1}\ \delta_k^{l+1}\ \sigma'_{\mathcal{L}}(z_j^l), \frac{\partial C}{\partial b_j^l} = \delta_k^l.$

The deep learning ANN was executed in Rapid Miner studio software (version 7.3). Rapid Miner provides an integrated tool for neural network analysis that supports all the machine learning process steps, including data preparation, results in visualisation, validation, and optimisation (Alsaqer & Sasi, 2017; Kanmani & Jayapradha, 2017). This study set the number of neurons to 1000 and epochs to 30, with four hidden layers to increase the ANN's depth for a deep learning approach.

5.2.6 Accuracy assessment

The extracted pixels for each image (PlanetScope and UAS) were randomly split into training (60%) and test (40%) datasets. A confusion matrix was used to evaluate the accuracy assessment or network performance (Cömert & Kocamaz, 2016). The confusion matrix produces an OA and KHAT, which indicates the percentage of correctly classified pixels and the effectiveness of the overall classification (Congalton & Green, 2019; Tu *et al.*, 2018). KHAT values range between 0 to 1 in percentage; values closer to 1 predict the best results. The confusion matrix measures the user and producer accuracy embedded in its system. The user accuracy corresponds to an error of commission (inclusion), and producer accuracy corresponds to an error of omission (exclusion)(Atkinson & Tatnall, 1997). The confusion matrix was calculated by comparing the ground truth data (plantation biochemicals) with the in-depth learning ANN classification results. OA and KHAT equations are as follows:

$$OA = \frac{\sum_{i=1}^{T} Xii}{n}$$

$$Kappa = \frac{n * \sum_{i=1}^{T} Xii - \sum_{i=1}^{T}\ \sum_{i=1}^{T} Xij}{n^2 - \sum_{i=1}^{T}\ \sum_{i=1}^{T} Xij}$$

T is the number of classes; Xii represents the correctly classified pixels in class i; Xij represents the incorrectly classified pixels in class i; and n represents the number of pixels participating in the classification (Tu *et al.*, 2018).

5.3 Results

5.3.1 Basic soil physio-chemical properties

Table 5.2 shows the basic assessment of soil physio-chemical properties (attribute) was conducted at 0-20 cm soil depth in the study area. These properties are essential attributes of ancillary data, potentially used to enhance this study's nutrient regime classification accuracy.

Table 5. 2: General study area information and basic soil physio-chemical properties of the 0 - 20 cm soil depth for the nutrient management regime trial.

Attribute	Study area
Sand & Silt %	56
Clay %	44
Texture	Silty Clay
pH (KCl)	3.55
pH (H_2O)	3.89
Exchangeable acidity (cmol$_c$ kg^{-1})	7.16
Organic carbon (WB) %	8.35
N %	0.51
C : N	16.4
P (ppm)	6.51
K^+ (cmol$_c$ kg^{-1})	0.20
Ca^{2+} (cmol$_c$ kg^{-1})	0.60
Mg^{2+} (cmol$_c$ kg^{-1})	0.27
Na^+ (cmol$_c$ kg^{-1})	0.06
ECEC (cmol$_c$ kg^{-1})	8.29
Base saturation (%)	13.6

5.3.2 Deep learning ANN using UAS imagery

As a result, very-high-resolution 6cm spatial resolution UAS imagery pixels produced an overall classification accuracy of 87.62%, a KHAT statistic value of 0.83, and an error rate of 12.38%. Table 5.3 & 5.4 is a confusion matrix showing the performance of a deep learning ANN and the user and producer accuracy which is calculated by dividing the number of correctly classified pixels in each nutrient regime by the total number of pixels in the corresponding column. The combination of UAS imagery and a deep learning ANN produced excellent classification accuracy for each nutrient management regime, with the user and producer accuracy ranging from 75% to 100% and 82% to 94%, respectively (Table 5.3). The classification of nutrient removed plot pixels produced the best results with user and producer accuracies of 100% and 82%, respectively. Similarly, the retained plot pixels' classification produced better results than other plots with a user and producer accuracy of 94% and 88% for nutrient retained plots, respectively (Table 5.3). However, the classification of rehabilitated plots pixels produced lower results with user and producer accuracies of 75% and 86% in this study, respectively.

Deep learning model scoring history

Figure 5.3 shows the scoring history of the UAS image model that was trained for 30 epochs using four hidden layers and 200 neurons. The choice of limiting the epochs to 30 was made based on the empirical observation that the process converged well within 30 epochs (Figure 5.3). Figure 5.3 shows that most "learning" occurred during the first ten epochs.

Table 5. 3: Confusion matrix based on the deep learning ANN and the UAS imagery five waveband pixels. *Bold values indicate the number of correctly classified pixels.

	Removed	Rehabilitated	Fertilized	Retained	Row total	User's accuracy (%)
Removed	**4100**				4100	100
Rehabilitated	900	**4300**	200	300	5700	75
Fertilized		500	**4700**	300	5500	85
Retained		200	100	**4600**	4900	94
Column total	5000	5000	5000	5200	**20200**	
Producer's accuracy (%)	82	86	94	88		

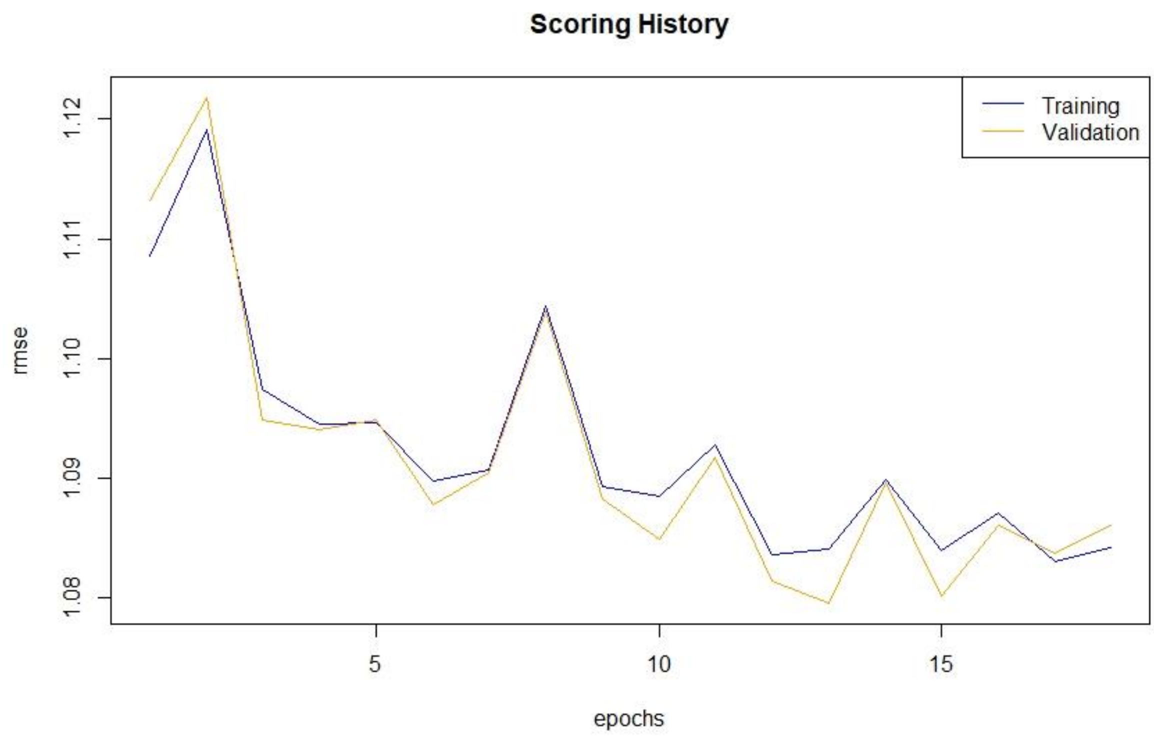


Figure 5. 3: The scoring history of the deep learning ANN using UAS imagery. * Y-axis shows classification error (RMSE), and the x-axis shows the number of epochs. The series lines show training (yellow) and validation (purple) converging after 30 epochs.

5.3.3 Deep learning ANN using PlanetScope imagery

As a result, very-high-resolution 3m spatial resolution PlanetScope imagery pixels produced an overall classification accuracy of 81.50%, a KHAT statistic value of 0.75 and an error rate of 18.50%. The combination of PlanetScope satellite imagery and a deep learning ANN produced excellent classification accuracies for each nutrient management regime, with user and producer accuracies ranging from 73% to 100% and 76% to 90%, respectively (Table 5.4). Similarly, to the UAS imagery, the PlanetScope satellite image best-classified nutrient removed plot pixels with user and producer accuracies of 100% and 76%, respectively. The classification of nutrient fertilized and retained plot pixels produced the same results with user and producer accuracies of 80% and 80%, respectively. The rehabilitated plot pixels classification performed well with 73% and 90% user and producer accuracy, respectively (Table 5.4).

Deep learning model scoring history

We replicated the same parameters in the UAS image (Figure 5.4). Figure 5.4 shows the scoring history for the PlanetScope image model that was trained for 30 epochs using four hidden layers and 200 neurons. In Figure 5.4 the most “learning” occurred during the first ten epochs.

Table 5. 4: Confusion matrix based on the deep learning ANN and the PlanetScope imagery four waveband pixels. *Bold values indicate the number of correctly classified pixels.

	Removed	Rehabilitated	Fertilized	Retained	Row total	User’s accuracy (%)
Removed	**3800**				3800	100
Rehabilitated	1200	**4500**	500		6200	73
Fertilized			**4000**	1000	5000	80
Retained		500	500	**4000**	5000	80
Column total	5000	5000	5000	5000	**20000**	
Producer’s accuracy (%)	76	90	80	80		

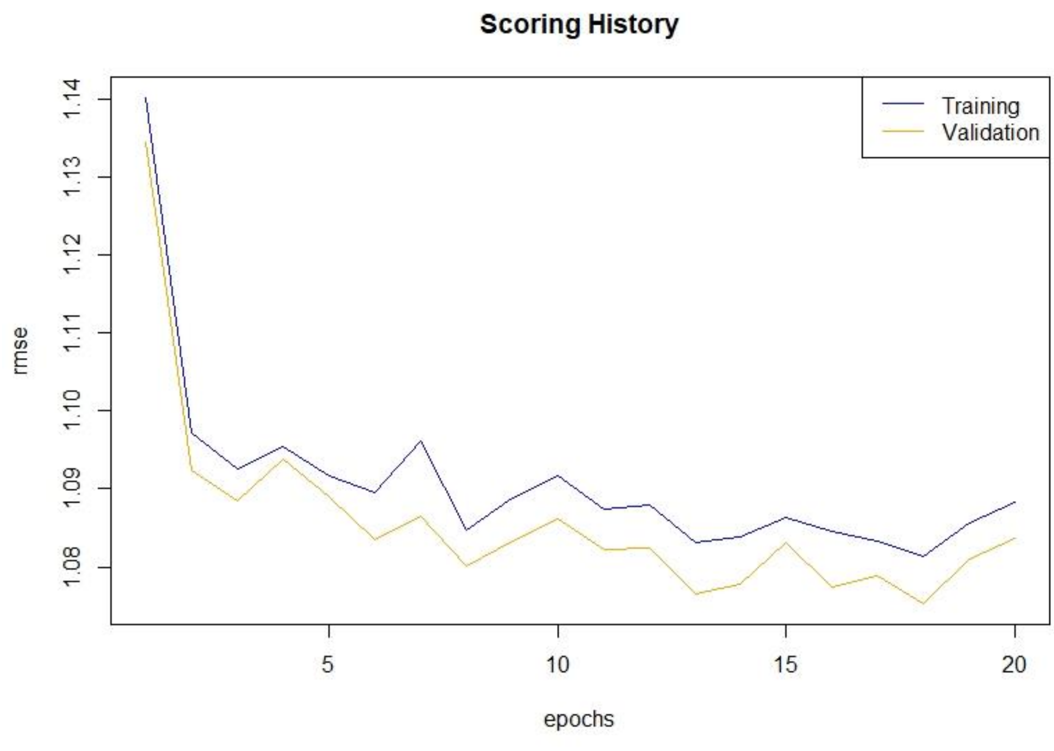


Figure 5. 4: The scoring history of the deep learning ANN using PlanetScope imagery. * Y-axis shows classification error (RMSE), and the x-axis shows the number of epochs. The series lines show training (yellow) and validation (purple) converging after 30 epochs.

5.3.4 Comparing UAS and PlanetScope imagery using a deep learning ANN

Very-high-resolution UAS and PlanetScope imagery pixels accurately classified nutrient management regimes with overall classification accuracy> 80%. Table 5.5 compares the overall classification accuracy between using UAS and PlanetScope imagery using overall accuracy, KHAT and error rate. Ultimately, the UAS imagery performed better than PlanetScope imagery. When comparing the differences in actual values between the UAS and PlanetScope imagery, there is a 6.12%, 0.08% and a 6.12% difference in the OA, KHAT and error rate, respectively (Table 5.5).

Table 5. 5: Overall classification performance of UAS and PlanetScope imagery pixels using a deep learning ANN.

Imagery	Overall accuracy (%)	KHAT (%)	Error rate (%)
UAS	87.62	0.83	12.38
PlanetScope	81.50	0.75	18.50

Furthermore, table 5.6 compared the user and producer accuracy of very high-resolution UAS and PlanetScope imagery to classify each nutrient management regime. Figure 5.5 illustrates the classification of all nutrient management regimes in the study area using UAS imagery. As a result, the UAS image pixels produced higher user accuracy (>=75%) when compared to PlanetScope image pixels for the removed, fertilized and retained plots. However, the rehabilitated plot produced a lower user accuracy for the UAS image pixels (75%) than the PlanetScope image pixels (86%). The producer accuracy of the UAS image pixels was higher (>=82%) for rehabilitated, fertilized, and retained plots compared to the PlanetScope image pixels. However, the removed plot produced a lower producer accuracy for the UAS image pixels (82%) than the PlanetScope image pixels (100%) (Table 5.6).

Table 5. 6: Comparing UAS and PlanetScope image pixel user and producer accuracies across all nutrient management regimes.

	Removed	Rehabilitated	Fertilized	Retained
UAS				
User's accuracy (%)	100	75	85	94
Producer's accuracy (%)	82	86	94	88
PlanetScope				
User's accuracy (%)	76	90	80	80
Producer's accuracy (%)	100	73	80	80

*Producer accuracy = omission error and user accuracy = commission error.

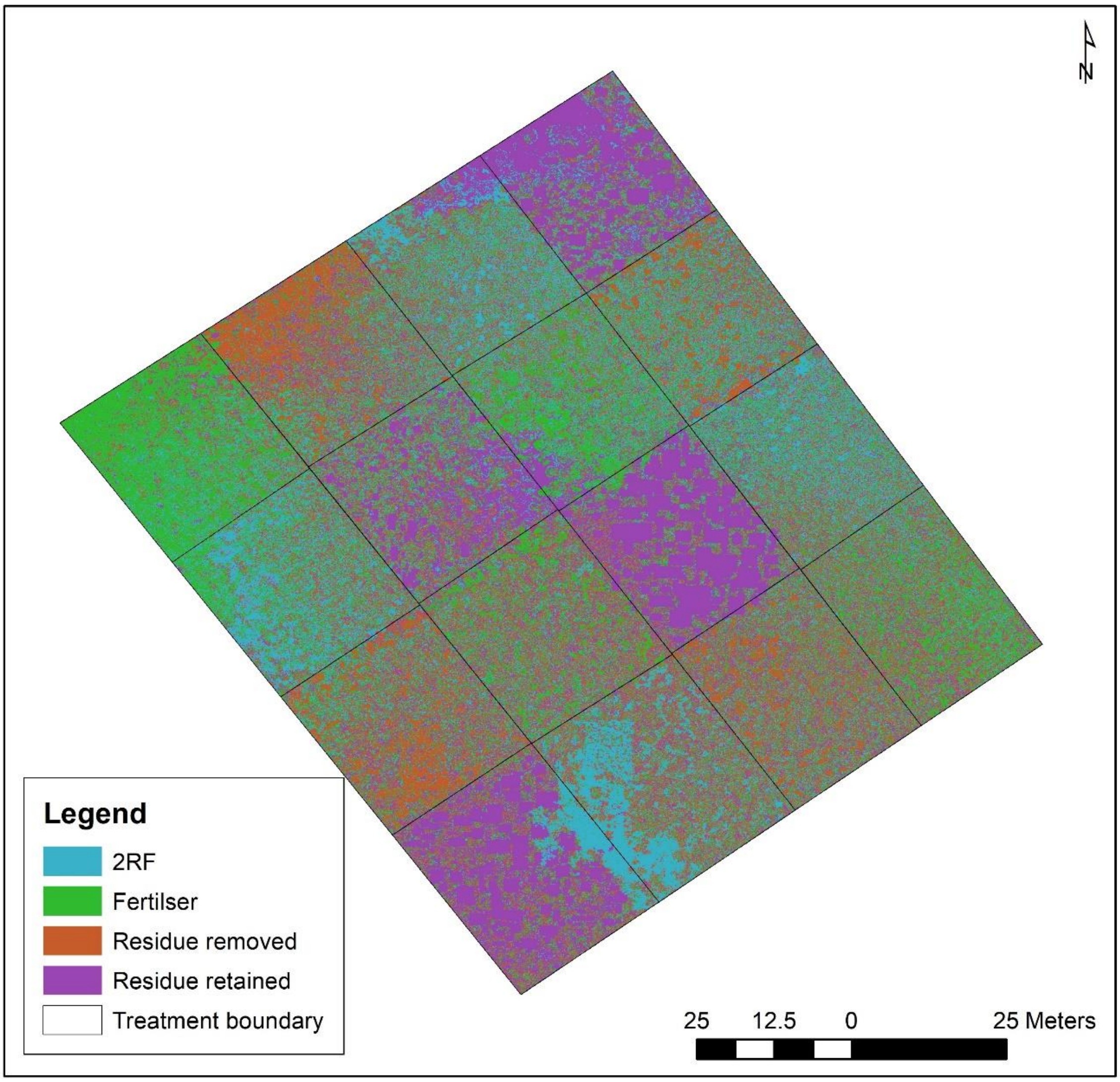


Figure 5. 5: The classification of nutrient management regimes using UAS imagery. *Each nutrient management regime is indicated within each plot on the map (please refer to experimental design and field data section of this study for information).

5.4 Discussion

This study has shown a deep learning ANN's potential to accurately classify critical commercial nutrient management regimes using ultra-high-resolution UAS imagery and very-high-resolution satellite imagery in KwaZulu-Natal, South Africa. A deep learning ANN combined with very-high-resolution UAS imagery provides an ideal framework for classifying several nutrient management regimes in a commercial forestry compartment. Additionally, this study has shown that very high-resolution satellite imagery can classify several nutrient management regimes.

5.4.1 Classification using UAS imagery

The deep learning ANN successfully classified all nutrient management regimes with a lower spectral resolution and high spatial resolution UAS sensor (Figure 5.5). The UAS sensor produced an OA above 87% and a KHAT value of 0.83% for all nutrient management regimes. A multi-perceptron ANN was more effective and efficient, reducing computational time while producing a low error rate of 12.38%. To date, not many studies have used a similar experimental design; hence the results cannot be compared to many other studies. A similar study by Escalante *et al.* (2019) reported an OA of 83% when estimating N fertilization in Barley using UAS imagery and a deep learning approach. Their study used a 3-waveband RGB image, whereas this study used a 5-waveband image, including a NIR and red-edge waveband which improved accuracy and predictive capabilities. The use of the Micasense camera improves predictions in this study. Although the Micasense camera provides exceptional performance for the objectives obtained in this study, it is more expensive than commonly used RGB cameras. However, for the level of detail required in this study, the Micasense camera provided the relevant performance.

5.4.2 Classification using PlanetScope imagery

The deep learning ANN successfully classified all nutrient management regimes using high spatial resolution PlanetScope satellite imagery. To date, this is the first study to use a combination of PlanetScope imagery and an ANN to detect nutrient concentrations in a forestry environment. Similar studies using high-resolution imagery from different sources combined with a deep learning approach. The results obtained in this study are comparable to Pereira *et al.* (2022), who found UAS imagery superior to PlanetScope and Sentinel-2 imagery. Our findings concur that the accuracy of the predictions is consistent with the area of interest's area size (AOI's). Satellite imagery is more suitable for larger areas than UAS imagery. Hence, UAS imagery can provide higher levels of spatial accuracy when compared to satellite imagery (Watt *et al.*, 2019). PlanetScope imagery's performance was exceptional compared to previous studies (Berger *et al.*, 2020; Dash *et al.*, 2018; Johansen *et al.*, 2020; Lussem *et al.*, 2019). The overall results indicate the usefulness and capabilities of 4-band PlanetScope satellite imagery in detecting discrete nutrients in a compartment forest.

5.4.3 Comparing classification using UAS and PlanetScope imagery

The objective of this study was to compare the utility of ultra-high-resolution centimetre accuracy UAS and very-high-resolution PlanetScope imagery for classifying each nutrient management regime in a compartment forest. This study refers to ultra-high-resolution imagery with less than 1m spatial resolution, and very-high-resolution imagery is > 1m and < 10m (Watt *et al.*, 2019). These platforms were selected based on their similarities in spectral, radiometric, and very-high spatial resolutions. Table 5.1 shows the configurations of each platform which are almost identical for two vastly different imaging platforms. The UAS imagery performed better than satellite imagery; however, its margin surpassed the PlanetScope imagery was small. However, limitations exist amongst both platforms, which include the study design. UAS deliver exceptional high resolutions, although there are limitations such as gusty wind conditions, flight restrictions, battery capacity, and a skilled UAS pilot (Pereira *et al.*, 2022). When compared to satellite platforms, these limitations do not exist. However, satellite platforms do not offer the same versatility as UAS platforms, such as timely collection, administrative acquisition processes, zoning into the AOI, and image format issues. Another limitation exists in the affordability of both platforms. UAS imagery RGB generally costs less than satellite imagery (multispectral or hyperspectral) due to the complexity, investment, and technology requirements for building and launching a satellite. This study was essential to show that the latest very-high-resolution satellite imagery can be utilized for mapping nutrients with highly accurate precision in discrete areas where UAS's are generally restricted (Johansen *et al.*, 2020). This study shows that UAS imagery is better at predicting nutrient concentrations in heterogeneous environments when compared to satellite imagery. Hence, the detection of discrete foliar information requires higher spatial resolutions combined with NIR and red edge wavebands. Furthermore, the battery life of a UAS platform a limitation for studies that require a series of images for monitoring change detection. However, in this study, only a snapshot was required.

5.5 Conclusion

Overall, this study's results indicate that ultra-high-resolution UAS imagery and very-high-resolution satellite imagery can classify nutrient management regimes in a commercial plantation forest in KwaZulu Natal, South Africa. The critical point was that PlanetScope results were closely matched with the UAS imagery, which supports satellite imagery for mapping discrete land use biochemicals using a deep learning ANN. This chapter provided important feedback for understanding the effectiveness of airborne (UAS imagery) and satellite imagery for detecting nutrient deficiencies at plantation level. As a result, the information gathered here will be used to improve the detection accuracies of nutrient deficient trees in the preceding chapter.

CHAPTER 6: Improving the prediction of nutrient content in a compartment forest using very-high-resolution imagery and a deep learning artificial neural network

To provide a framework for predicting nutrients in a compartment forest using very-high-resolution imagery and a deep learning artificial neural network.

This chapter was based on:

Singh L*, Mutanga O, Mafongoya P, Peerbhay KY, Ismail R (Preparation). Predicting forestry health indicators using high-resolution UAS imagery and a deep learning artificial neural network in South Africa.

Abstract

Accurately predicting forest quality variables provides stakeholders with valuable information on forest condition, assessment, and effective management. Very high-resolution spatial resolution, remotely sensed datasets provide resources for evaluating finite ecological forest health indicators. Therefore, this study aims to evaluate how a well-calibrated broad-band multispectral camera (Micasense RedEdge-M), captured using an UAS, can predict tree macronutrients and micronutrients using a deep learning ANN in a heterogenous *Eucalyptus* compartment forest. As a result, this study found the NIR and red-edge wavebands to be the most important contributors to predicting macronutrients and micronutrients with successful predictions resulting in R^2's ranging between 0.45 and 0.75, with RMSE's below 0.08. Exploring heterogeneous environments tests the capability of remotely sensed data to detect subtle changes in discrete biochemical information. Future studies should further understand the contribution of climatic variability as a component of ancillary data to create a more holistic construct for predicting foliar nutrients.

Keywords: UAS, deep learning, forestry, foliar biochemicals

6.1 Introduction

Precise estimates of forest quality variables are needed to support global climate change policy initiatives such as Reducing Emissions from Deforestation and forest Degradation (REDD+) (Gomes *et al.*, 2010) and the Forest Observation System (FOS) (Schepaschenko *et al.*, 2019). Forest nutrient availability is a dominant driver for C retention (De Vries, 2014). In a paper, Fernández-Martínez *et al.* (2014) describe nutrient-rich forests with a higher rate of net ecosystem production and lower ecosystem respiration which increases C allocation in the woody tissues and fungal root symbionts and exudates. These nutrient-rich forests have higher retention of NPK, whereas nutrient-poor forests show symptoms of nutrient deficiencies, including a reduction in leaf area and lower foliar concentrations (Fernández-Martínez *et al.*, 2014; Watt *et al.*, 2019). NPK is associated with metabolically active proteins, including RuBisCo, directly linked to plant productivity (Cavender-Bares *et al.*, 2020; Silva & Uchida, 2000). Hence, NPK are essential contributors to assessing plant health and rates of C retention in forest environments (Fernández-Martínez *et al.*, 2014; Silva & Uchida, 2000; Watt *et al.*, 2019). More specifically, estimating nutrient quality biochemicals are critical for monitoring forest growth and improving decision-support systems for specific agronomic practices (Guo *et al.*, 2020). However, the precise measurement of these nutrient quality biochemicals at the compartment level becomes challenging for silviculture practice when using conventional nutrient assessment methods (Watt *et al.*, 2019).

Conventional nutrient assessment methods are time-consuming, labour intensive, expensive, and generally not consistent for extrapolating a single measurement over a large (<10 km^2) sample area (Pullanagari *et al.*, 2016; Watt *et al.*, 2019). Moreover, these assessments are generally only carried out at the end of a growing season, which provides little information regarding the temporal effects of climate, and other environmental constituents have on nutrient retention (Pullanagari *et al.*, 2016). Very high spatial resolution multispectral satellite data have contributed immensely to determining nutrient quality biochemicals across homogeneous and heterogeneous environments. Hence, foliar nutrient information require very high spatial resolutions for successful detection accuracies (Watt *et al.*, 2019). Recent developments in remote sensing offer an opportunity to such highly detailed mapping demands made by the forestry sector. Very-high-resolution UAS's imagery offer significant potential for improving the precision accuracy of mapping biochemical content of vegetation (Cai *et al.*, 2019; Chemura *et al.*, 2018). UAS's provide incomparable resolution and data densities than alternative platforms however they are limited in flight range of up to 10 km^2 (Watt *et al.*, 2019). To date, a few studies have investigated the potential of very-high-resolution multispectral UAS imagery for accurately predicting NPK in commercial plantation forests. Multispectral UAS imagery provides a more affordable option with strategic band configurations that elevate issues of overfitting and spectral noise when compared to hyperspectral imaging (Mutanga *et al.*, 2004a; Mzinyane *et al.*, 2016). Hence, the focus of research should be on developing remote sensing methods for effectively and efficiently assessing the quality of commercial forestry plantations using a cost-effective approach for sampling at compartment level (Fassnacht *et al.*, 2016; Galidaki *et al.*, 2017; Köhl *et al.*, 2006; Watt *et al.*, 2019). Accurate nutrient assessments using remote sensing enable forest managers to make informed decisions on the health of their plantations yield potential and overall productivity (Watt *et al.*, 2019). In addition, further research on this issue will improve the efficacy of remote sensing, providing a cost-effective, timely, non-destructive alternative to traditional nutrient assessments (Dash *et al.*, 2018).

Emergent literature presents the potential use of UAS's as an alternative platform to satellite and airborne platforms given their low-cost operation in environmental monitoring, higher spatial resolutions, and high flexibility in image acquisition (Dash *et al.*, 2018; Mutanga *et al.*, 2016). Combining very high spatial resolution UAS imagery and broad-band spectral resolutions can provide an ideal framework for detecting foliar biochemicals at the compartment level (Lussem *et al.*, 2019). Furthermore, a combination of very-high-resolution UAS imagery and a deep learning artificial intelligence machine learning approach has improved biochemical information detection and accuracy. For example, Cai *et al.* (2019) tested the consistency of CubeSat-based (455nm-860nm) chlorophyll index green (CIg) against UAS-based CIg for N stress in different N management practices of cornfield trials in Champaign County, Illinois. Their study showed that CubeSat-based CIg produced high correlations with UAS-based CIg (correlation above 0.9). Montgomery *et al.* (2020) measured the crop nutrient status of flue-cured tobacco using a combination of UAS imagery (20-megapixel RGB camera and a spatial resolution of 0.05m), canopy structure and multiple linear regression model in Wilson, North Carolina. The authors used a low-cost UAS equipped with consumer-grade RGB cameras (multi-view stereo images) in a 0.5 ha field. The authors successfully measured NPK and B. The most robust relationships produced adjusted R^2's of 0.81 and 0.41 for N and B, respectively. Their study showed the positive influence of canopy structure and spectral reflectance when measuring the crop nutrient status of tobacco trees. Osco *et al.* (2020a) predicted leaf N in maize crops using UAS imagery (SenseFly Parrot Sequoia multispectral sensor (550nm-790nm)) and machine learning models in Brazil. The authors successfully predicted leaf N using a series of spectral vegetation indices NDVI, NDRE, green normalized difference vegetation (GNDVI), and the soil adjusted vegetation index (SAVI)) and machine learning models with variations (REPTree (REPT), Random Forest (RF), kNN, Support Vector Machine-Polynomial (SVMP), linear regression (LR)). Their study produced the highest R^2 of 0.91 for leaf N concentration using the RF machine learning algorithm using a GSD of 0.10m.

Similarly, Costa *et al.* (2021) determined leaf nutrient concentrations in citrus trees using multispectral UAS imagery (Micasense (465nm-1000nm)) and machine learning in Polk County, Florida, USA. As a result, the estimation model successfully mapped two citrus tree varieties (Hamlin and Valencia) using a gradient boosting regression tree. The authors achieved high precision for macronutrients (N, P, K, Mg, Ca, and sulphur (S)) with an average error of 9% and 17% and moderate precision for micronutrients (Zn, B, Mn, Fe, Cu) with an average error of 16% and 30% for Hamlin and Valencia citrus trials, respectively.

Previous studies predicted mainly crop health status using vegetation indices and indirect measures such as the chlorophyll index green combined with mainly linear regression modelling techniques in the USA and Chile (Cai *et al.*, 2019; Costa *et al.*, 2021; Dash *et al.*, 2018; Kattenborn *et al.*, 2019; Montgomery *et al.*, 2020). This study aims to provide a framework for detecting macronutrients and micronutrient content using very-high-resolution (<10 cm) multispectral imagery and a deep learning ANN in a compartment forest in the KwaZulu-Natal Midlands, South Africa.

6.2 Material and methods

6.2.1 Study area/map

This study was conducted at the Clan Sappi plantation located in the Midlands, South Africa (29°20’6” S, 30°27’13” E)(Figure 6.1). The study area is 1.4 hectares consisting of a nutrient management regime trial situated in the mist belt grassland bioregion of the KwaZulu-Natal, Midlands. The species variety planted was 3-4-year-old *Eucalyptus grandis x Eucalyptus urophylla*. *Eucalyptus grandis* (flooded gum or rose gum) is native to New South Wales, and *Eucalyptus urophylla* (Timor white gum, Timor Mountain gum, popo or ampupu) is native to the Indonesian Archipelago and Timor. The study area is situated at 788m ASL receives a MAP range of 1000-1100 mm and a MAT of 18.1 degrees Celsius (Mucina & Rutherford, 2006). The lithology is shale and dolerite, consisting of a silty clay soil texture comprising of 56% sand and silt and 44% of clay.

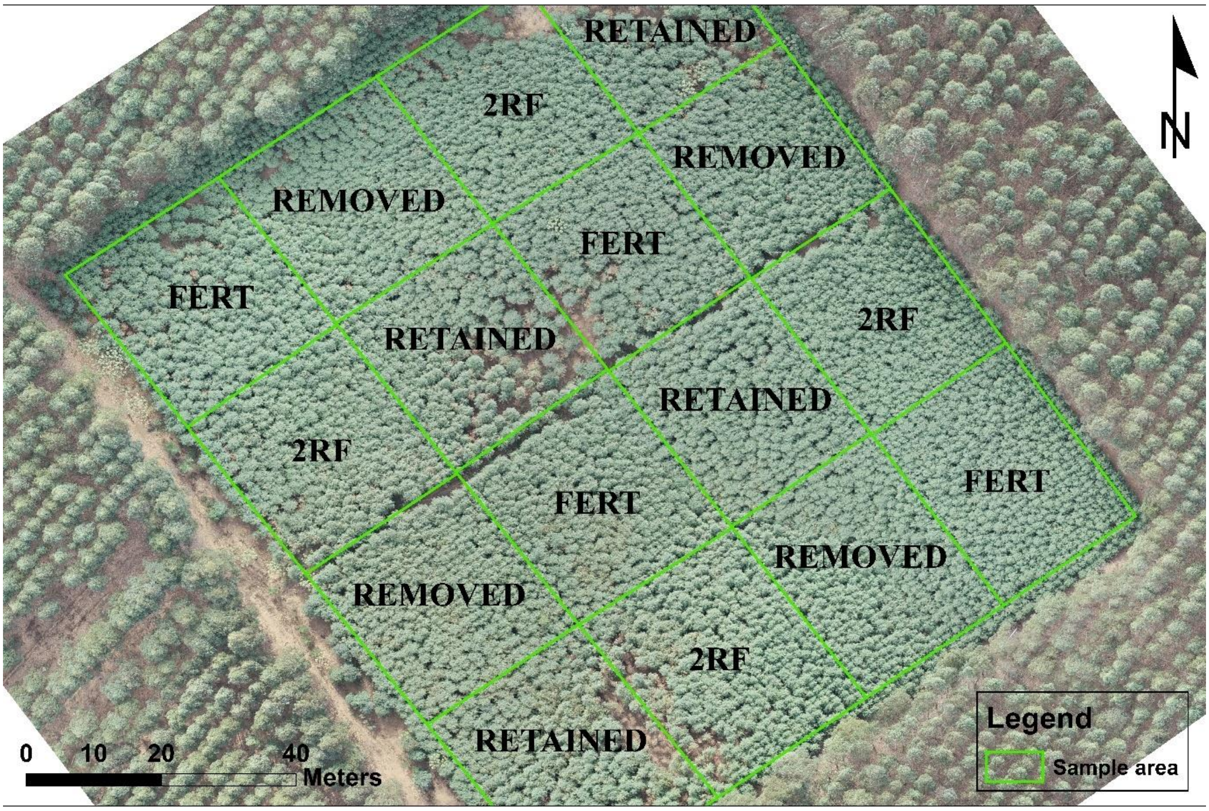


Figure 6. 1: The location of the study area in a commercial forest of KwaZulu Natal, Midlands, South Africa, and the nutrient management regime formation.

6.2.2 Experimental design & field layout

This section is similar to section 5.2.2 of this thesis. A nutrient management regime trial comprised four treatments with four replicates arranged in a Latin square design. A series of 16 GCPs were established at the centre of each polygon represented in Figure 6.1 within the study area. Each GCP was located using a Trimble GeoExplorer 6000 series GeoXH GPS.

6.2.3 UAS imagery & pre-processing

The UAS image specifications and pre-processing techniques is similar to section 5.2.3 of this thesis. UAS imagery was acquired under cloudless (0%) conditions between 10:00 and 14:00 h central African time using DJI Phantom 4 Pro quadcopter in October 2018. The DJI Phantom 4 Pro was flown at an altitude of approximately 60 m above the AOI, resulting in a GSD of 8 cm (Dash *et al.*, 2018). The image was acquired as a snapshot of the AOI with a focal length lens of 5.5 mm and a FOV of 151.47 degrees. The sensor was calibrated using a reference panel (white reference) coated with a barium sulphate of known reflectivity. Sensor calibration occurred directly before and after each flight to offset any change in the sun's atmospheric condition and irradiance (Singh *et al.*, 2017a). The UAS image was processed using (Structure-from-Motion software PhotoScan v1.4) (AgiSoft LLC, St. Petersburg, Russia) in conjunction with the GCPs speculated in the experimental design of this study (Lussem *et al.*, 2019). Each nutrient management regime represented various pixels based on the UAS image resolution.

6.2.4 Deep learning artificial neural network (ANN) algorithm

A deep learning ANN architecture constructs increasingly sophisticated hierarchical architectures using multiple hidden layers with multi-layer neurons (Reichstein *et al.*, 2019). Multi-layer neurons can cycle or loop information between different neurons, creating powerful pattern recognition capabilities to learn intricate multivariate data patterns (Atkinson & Tatnall, 1997; Mutanga & Skidmore, 2004; Reichstein *et al.*, 2019). Figure 6.2 represents the network structure of an ANN; whereby hidden layers bridge the gap between input layers and output layers. Artificial neurons receive a set of weighted inputs to produce an output through an activation function (Atkinson & Tatnall, 1997) (Figure 6.2). An activation function is responsible for transforming the summed weighted input from the node into the node's activation or output to help the network learn intricate patterns in the data. Several types of activation functions such as sigmoid, rectified linear units, or hyperbolic tangents exist within the neurons of each hidden layer (Berg & Nyström, 2018). Combined with the activation function, each neuron in the ANN is assigned a bias, including the output neurons and excluding the input neurons, whilst the connections between neurons in subsequent layers include matrices of weights (Berg & Nyström, 2018). The weighted input for a deep learning ANN is:

$$z_j^l = \sum_k w_{jk}^l \, \sigma_{l-1}\left(z_k^{l-1}\right) + b_j^l$$

The sum of all inputs is taken to the neuron j in Layer l, the number of neurons. The deep learning ANN naturally defines a recursion in previous weighted inputs through the ANN. The calculation that terminates any recursion is:

$$\sigma_0\left(z_j^0\right) = \mathcal{Y}_j^0 = x_j$$

More specifically, in figure 6.2, Layer 0 is the input layer which is consists of imagery (UAS and PlanetScope) and the response variable (nutrient management regimes), and Layer L is the output layer. Layer l – 1 and Layer l represent the hidden layers; in this figure 6.2, there are two hidden layers, whereas, in this study, four hidden layers were used (Berg & Nyström, 2018).

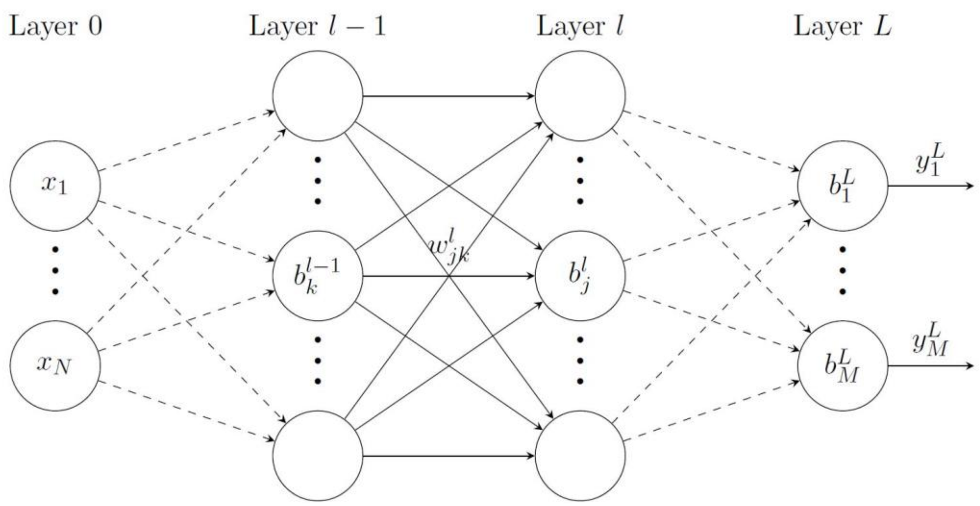


Figure 6. 2: A fully connected feedforward ANN (Berg & Nyström, 2018).

The deep learning ANN was executed in Rapid Miner studio software (version 7.3). Rapid Miner provides an integrated tool for neural network analysis that supports all the machine learning process steps, including data preparation, results in visualization, validation, and optimization (Alsaqer & Sasi, 2017; Kanmani & Jayapradha, 2017). This study set the number of neurons to 1000 and epochs to 30, with four hidden layers to increase the ANN's depth for a deep learning approach.

6.2.5 Accuracy assessment

The final dataset was split into training (70%) and test (30%) data (Breiman, 2017). Similar to section 3.2.8, R^2 was used for prediction and RMSE. Higher R^2 and lower RMSE values indicate a reliable model.

6.3 Results

6.3.1 Predicting of macronutrients and micronutrients

Tables 6.1 and 6.2 summarize the bare soil physiochemical properties and prediction statistics. This study successfully predicted macronutrients and micronutrients using very-high-resolution UAS imagery and a deep learning ANN. As a result, R^2's predictions ranged between 0.45 and 0.75, with RMSE values ranging between 0.01 and 38.04 for both macronutrients and micronutrients. P produced the highest R^2 value of 0.75, whilst B produced the lowest R^2 value of 0.54 (Table 6.2). The UAS image model was trained for 30 epochs using four hidden layers and 200 neurons for both macronutrients and micronutrients. The choice of limiting the epochs to 30 was made based on the empirical observation that the process converged well within 30 epochs.

Table 6. 1: General study area information and bare soil physio-chemical properties of the 0 - 20 cm soil depth for the nutrient management regime trial in the Midlands, South Africa.

Attribute	Study area
Sand & Silt %	56
Clay %	44
Texture	Silty Clay
pH (KCl)	3.55
pH (H_2O)	3.89
Exchangeable acidity ($cmol_c\ kg^{-1}$)	7.16
Organic carbon (WB) %	8.35
N %	0.51
C : N	16.4
P (ppm)	6.51
K^+ ($cmol_c\ kg^{-1}$)	0.20
Ca^{2+} ($cmol_c\ kg^{-1}$)	0.60
Mg^{2+} ($cmol_c\ kg^{-1}$)	0.27
Na^+ ($cmol_c\ kg^{-1}$)	0.06
ECEC ($cmol_c\ kg^{-1}$)	8.29
Base saturation (%)	13.6

Table 6. 2: Summary of predictive statistics and results using UAS imagery and a deep learning ANN.

Biochemical	Min	Max	Mean	R^2	RMSE
Nitrogen	47	60	53	0.71	6.72
Phosphorous	0.09	0.21	0.15	0.75	0.02
Potassium	0.63	1.04	0.81	0.74	0.07
Calcium	1.31	1.81	1.57	0.68	0.12
Magnesium	0.27	0.60	0.44	0.63	0.06
Sodium	0.14	0.22	0.17	0.63	0.01
Iron	251	412	338	0.45	38.04

Manganese	0.56	0.87	0.67	0.49	0.05
Copper	8.29	11.9	10.36	0.56	0.44
Zinc	15.5	22.62	19.01	0.54	3.08
Boron	22.51	38.79	30.63	0.54	4.98

Figure 6.3 and 6.4 displays the scatterplots used to show the variability of the data points over the mean in this study. K produced the highest R^2 of 0.74, whilst Fe produced the lowest R^2 of 0.45 (Figure 6.3 and 6.4). This study produced a mean R^2 prediction of 0.57 for all macronutrients and micronutrients. RMSE values were consistent with all R^2 predictions, with the highest RMSE of 4.98 and lowest RMSE of 0.02 for B and K, respectively (Table 6.2).

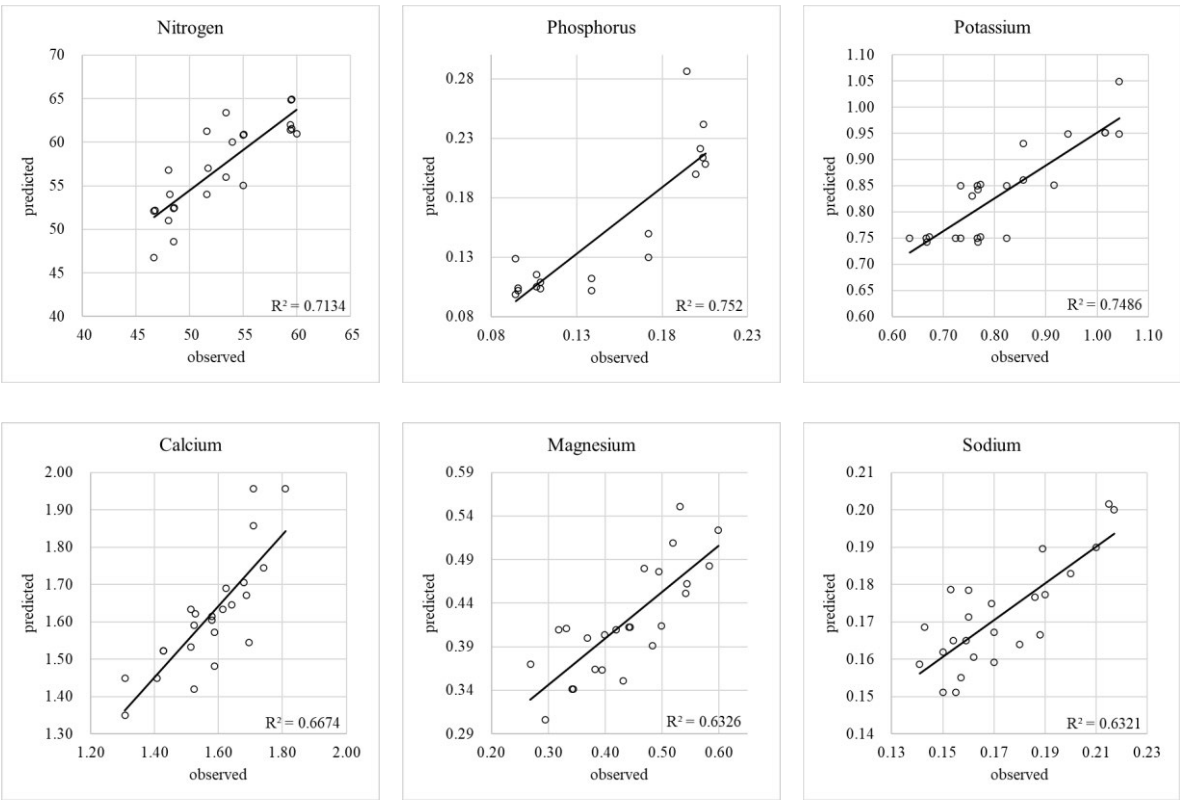


Figure 6. 3: Scatterplots showing the one-to-one relationship between predicted versus observed values of macronutrients using a deep learning ANN.

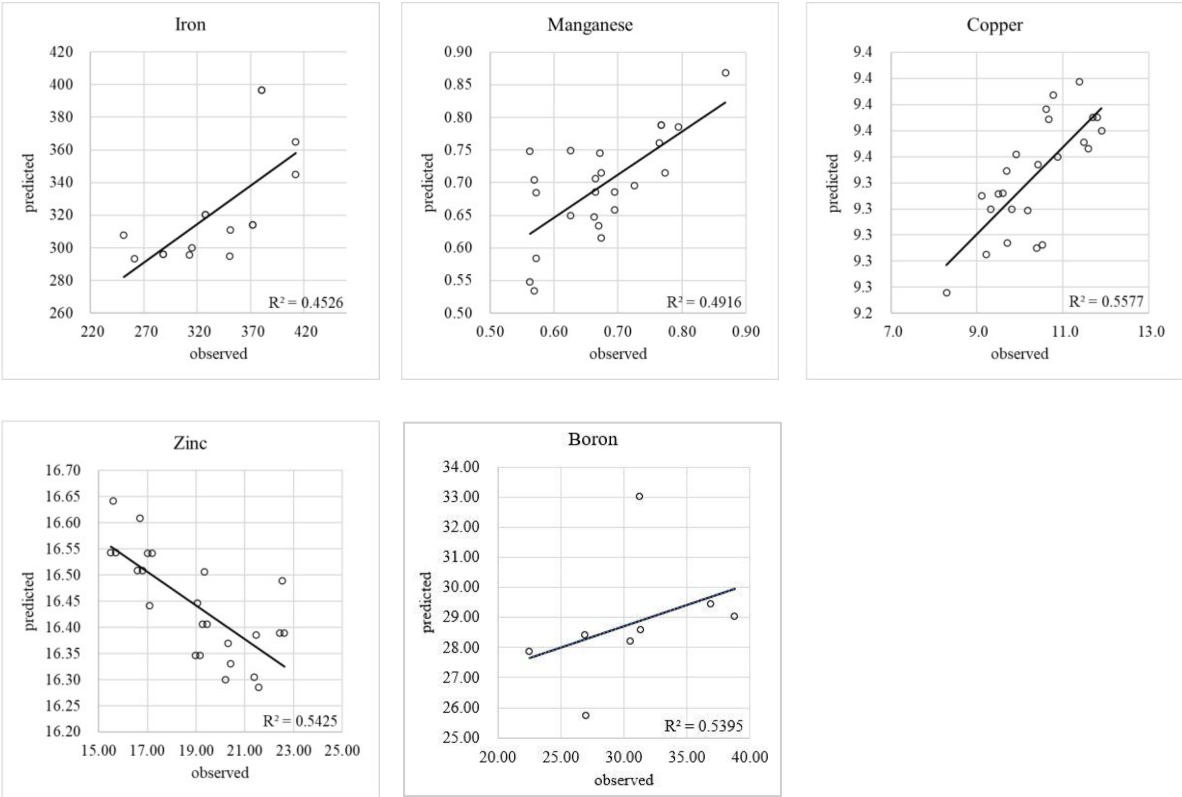
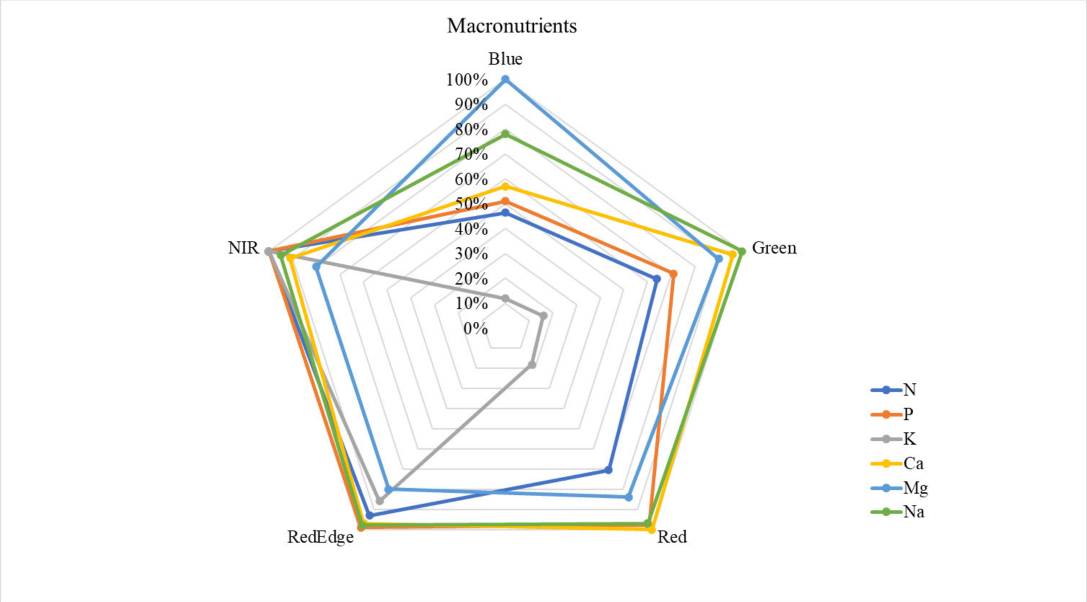


Figure 6. 4: Scatterplots showing the one-to-one relationship between predicted versus observed values of micronutrients using a deep learning ANN.

6.3.2 Variable importance

An important step was determining which wavebands onboard the UAS camera sensor were more valuable in the prediction model for predicting macronutrients and micronutrients in this study. Figure 6.5 illustrates a radar plot of variable importance measures for micronutrients and macronutrients, the colours represent a nutrient and the scale (0-100%) shows the percentage of importance to an electromagnetic region. The VI percentages were averaged across all macronutrients and micronutrients for each waveband (i.e., red, green, blue, NIR and red-edge). The NIR and red-edge wavebands produced the highest correlations with an average NIR waveband of 94% and a red-edge waveband of 92% across all macronutrients in this study. In contrast, the blue waveband produced the lowest correlation of 57% (Figure 6.5). The red-edge and red wavebands produced the highest correlations with an average red-edge waveband of 84% and a red waveband of 68% across all macronutrients in this study. In contrast, the NIR waveband produced the lowest correlation of 63% (Figure 6.5). Overall, the NIR and red-edge wavebands produced the highest correlations for predicting both macronutrients and micronutrients in this study (Figure 6.5). Figure 6.6 represents prediction maps of three macronutrients (NPK) and three micronutrients (Cu, Fe, Zn) concentrations using five-waveband UAS (Micasense) imager, higher concentrations are green in colour and lower concentrations are red in colour of the nutrients.



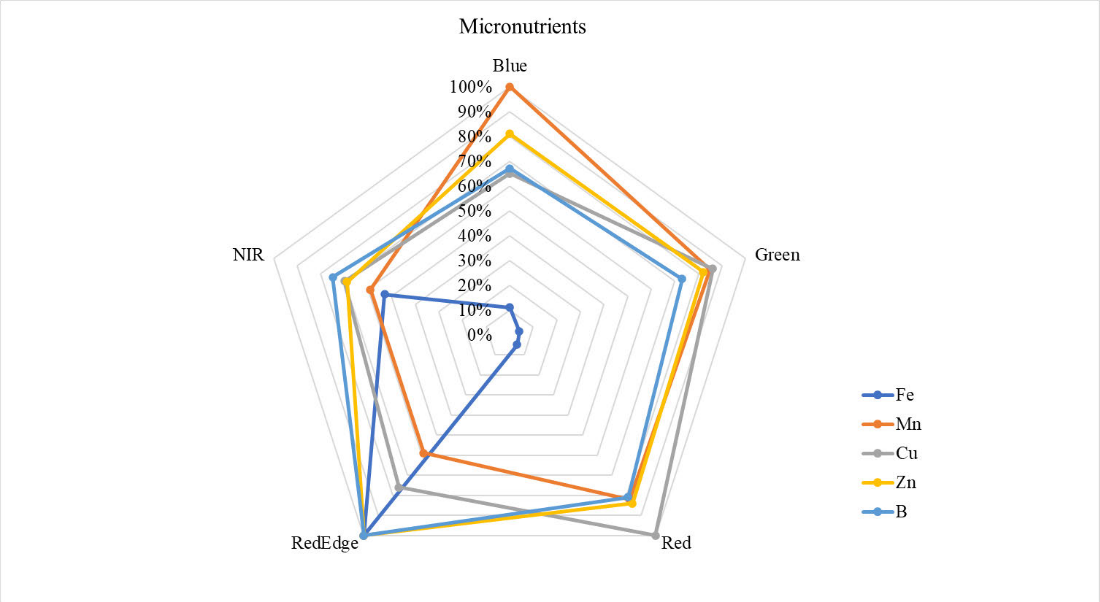


Figure 6. 5: Radar plots showing variable importance of the most essential wavebands used to predict macronutrients and micronutrients.

Macronutrients	Micronutrients

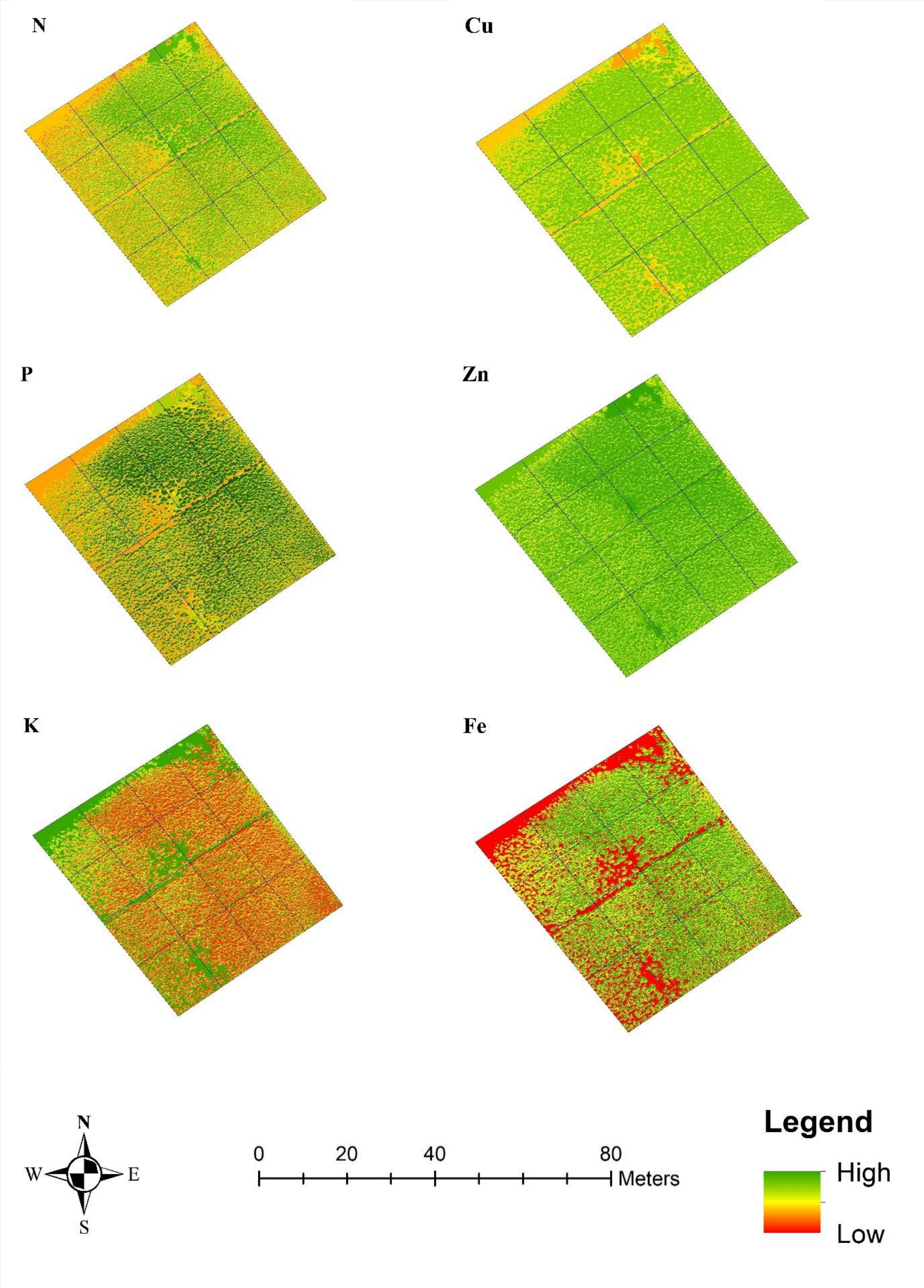

N
Cu
P
Zn
K
Fe
N
W
E
S
0
20
40
80
Meters
Legend
High
Low

Figure 6. 6: Prediction maps of three macronutrients (NPK) and three micronutrients (Cu, Fe, Zn) concentrations using five-waveband UAS (Micasense) imagery in KwaZulu Natal, Midlands, South Africa.

6.4 Discussion

Commercial afforestation can produce effective climate change mitigation strategies such as offsetting anthropogenic C, stabilizing climate risks, contributing several essential ecosystem services, and providing a safety net to meet basic human needs (De Vries, 2014; Forster *et al.*, 2021; Gamfeldt *et al.*, 2013; Luyssaert *et al.*, 2008). A challenge is maintaining the health of extensive commercial forests, as trees require an optimal supply of nutrients for growth (Dash *et al.*, 2018). Conventional methods provide accurate assessments; however, reproducing these assessments over large forestry compartments becomes ineffective and inefficient. The development of innovative methods that are more effective and efficient for large scale forest health assessment is critical for enhancing tree growth, reducing exposure to pests and disease, and ultimately improving yield (Köhl *et al.*, 2006). This study assesses the capability of the Micasense sensor onboard a UAS to detect and map essential forest health macronutrients and micronutrients effectively. Additionally, this study implemented VI measures to show the essential wavebands needed to detect macronutrients and micronutrients accurately. This study presents a reliable framework for detecting and monitoring forest health nutrient indicators of commercially grown *Eucalyptus grandis* variety in a compartment forest in KwaZulu-Natal, South Africa.

6.4.1 Predicting macronutrients and micronutrients using UAS imagery and a deep learning approach

The results show that very-high-resolution UAS imagery combined with a deep learning ANN successfully detected and mapped macronutrients and micronutrients within a commercial forestry environment. These results justify previous research recommendations for using very-high-resolution imagery and a deep learning artificial intelligence approach to map discrete foliar biochemicals in a compartment forest. This study site was preferred due to the homogeneity of commercial forestry conditions. Homogeneity allowed for consistency when sampling, which appropriates the objectives of this study. The results from this study were compared to Cai *et al.* (2019), who produced similar results. Cai *et al.* (2019) found that UAS-based imagery provided more spatial details when compared to CubeSat-based imagery. Implementing very high spatial resolution imagery was essential in our study for detecting the full range of macronutrient and micronutrient information. Another study by Montgomery *et al.* (2020) produced similar results to our study even when implementing ancillary data into their regression models, such as canopy structural data and water flow information. This study did not implement ancillary data into the modelling phase. Without ancillary data, this study produced similar results due to the robustness of using a deep learning approach combined with high-quality imagery. A few nutrients produced lower predictions results, such as Iron (0.45). The cause of this is unknown; however previous studies have related lower prediction accuracies to the mobility (translocation) of nutrient concentration within the plant physiology (Montgomery *et al.*, 2020; Silva & Uchida, 2000). This study was constrained to a snapshot of the study area; hence future studies could estimate nutrients over a more extended temporal resolution to understand the translocation of nutrients within the tree.

The prediction maps show the concentrations of nutrients within each different nutrient regime. Higher concentrations are green in colour, whereas lower concentrations are depicted red. Unfortunately, there is no distinct pattern showing the variation of nutrient concentration across each nutrient management regime. The problem could be caused by the pixel density of the image or the lower spectral resolution of the Micasense sensor. The distribution of macronutrient and micronutrient concentrations seems higher (greener in colour) in the eastern regions of the maps (figure 6.6). The abnormal distribution could be due to water absorption or water accumulation in the eastern part of the compartment.

6.4.2 Variable importance

An essential step in this study was to understand which wavebands contributed to each nutrient's prediction model. Overall, VI results show that the NIR and red-edge wavebands best predict macronutrients, with average VI of 94% and 92% for the NIR and red-edge wavebands, respectively. In comparison, the red and red-edge wavebands were the best in predicting micronutrients, with average VI of 70% and 87% for the red and red-edge waveband, respectively (Figure 6.5). The red-edge waveband was a vital descriptor in this study, previous studies that suggested using the red-edge waveband (Mutanga *et al.*, 2004b; Mutowo *et al.*, 2018; Rodrigues *et al.*, 2020; Zhang *et al.*, 2020). The blue waveband produced the lowest correlation of 57% for macronutrients whilst the blue, green and NIR wavebands produced the lowest results for micronutrients; however, they were still significantly higher. These results are not surprising as previous studies found similar waveband performances in this study (Mutowo *et al.*, 2018; Shi *et al.*, 2019; Türker-Kaya & Huck, 2017). VI results can improve the efficacy of the basic spectral waveband setting for different types of acquisition purposes.

For example, engineers can prioritize fitting spectral sensors with the red-edge waveband for remote sensing of nutrient information than using lower priority blue waveband settings. These results enable forest managers and critical forestry stakeholders to make investment decisions on the type of remote sensing technology to acquire for specific field surveys.

6.5 Conclusion

Overall, this chapter aimed to improve the detection accuracy of nutrient deficient trees using the best possible practices gathered from the previous chapters in this thesis.

The conclusions of this study are as follows:

- To date, the NIR and red-edge wavebands are high contributors to the prediction model of macronutrients and micronutrients.
- A combination of very high-resolution imagery (<10cm GSD) and deep learning ANN can accurately predict full range macronutrients and micronutrients.

The challenges of the forestry industry should guide future studies. Future studies should enhance conventional practice by improving detection accuracy. Enhancing the practice and effectiveness of remote sensing in a high-throughput environment will improve the efficacy of remote sensing. As a result, this chapter successfully detected nutrient deficiencies with a reasonable accuracy for in-field implementation at commercial plantation level.

CHAPTER 7: Detecting and mapping forest nutrient deficiencies: *Eucalyptus grandis x and Eucalyptus urophylla* trees in KwaZulu-Natal, South Africa: A synthesis

7.1 Introduction

The effects of nutrient deficiencies are often amplified when the productivity of valuable resources is at risk, which wastes functional space, time and inevitably impedes production. More importantly, ineffective nutrient screening technologies result in untimely in-field planting of trees before reaching optimal levels. Conventional nutrient screening regimes use destructive sampling methods, which involve periodic ground-based surveys and tedious laboratory assays that are costly and time-consuming (Pullanagari *et al.*, 2016). Hence, accurately quantifying nutrient deficient trees at a compartment level and within extensive commercial forestry remains unworkable, especially when dealing with many samples (Quentin *et al.*, 2017). The timely detection of nutrient depletion at the nursery level could optimize forest management practices for in-field planting (Garcia *et al.*, 2018). Hence, new methods must provide rapid detection capacities, reduce labour, and maximize time. Therefore, detecting and mapping nutrient deficiencies in commercial forestry are critical to maximising productivity, especially within a limited forest production area. In this context, the utility of remote sensing has offered great potential to enhance traditional nutrient detection methods in a high productivity environment.

However, the use of remotely sensed data for identifying nutrient concentrations has proven difficult, owing to the lack of appropriate spectral and spatial resolutions. Furthermore, the efficacy of the technology has not been widely adopted; hence a lack of understanding of the principles of remote sensing has led researchers to focus on mainly indirect estimates such as NDVI using traditional statistical data pre-processing methods. Also, conventional remote sensing mainly used hyperspectral spectrometer devices which caused overfitting and multicollinearity problems, thus selecting essential wavebands for detection was limited. Lastly, earlier statistical methods provided minimal accounts of the data acquired from remote sensing devices; hence scientists could not determine definitive relationships between remote sensing and agricultural health indicators. Most importantly, the challenges mentioned above weakened the efficacy of remote sensing for detecting nutrients in a forestry environment. Therefore, the aim of this thesis was to develop an alternative nutrient screening framework for the commercial forestry industry using the latest advancements in remote sensing and chemometric data analysis techniques. The benefits produced in this thesis present a new paradigm for the commercial forestry industry that requires quality planting material for long- and short-term resource sustainability.

In this thesis, the objectives were to (1) provide a synopsis of the application of remote sensing for detecting foliar nutrients, (2) investigate the ability of remote sensing to rapidly detect nutrient deficiencies of saplings in a nursery environment using hyperspectral data, (3) investigate the influence of VCP in improving the detection accuracy of nutrients of saplings in a nursery environment using NIR data, (4) compare the classification accuracy of ultra-high-resolution UAS imagery and very-high-resolution PlanetScope imagery in four nutrient management regimes using a deep learning ANN, and (5) provide a framework for the detection of nutrients in a compartment forest using very-high-resolution imagery and a deep learning ANN. In a nutshell, Chapter 2 provides an introduction and overall background for the thesis conceptual design, which address objective 1. the synthesis for chapters 3 and 4 (chapter 7.3) was combined, which specifically focus on a nursery experiment using handheld analytical devices to test the findings of chapter 2, which address objectives 2 and 3. Similarly, chapters 5 and 6 (chapter 7.4) was combined which provide an upscale approach with a specific focus on airborne and spaceborne remote sensing technology for large-scale mapping in a compartment forest, these chapters address objectives 4 and 5.

7.2 Summary of results/findings

Research objective 1 was achieved in chapter 2. The main purpose of chapter 2 was to provide evidence regarding the scientific background and underpinning for remote sensing of nutrient biochemicals. The latest scientific literature showed a trend towards NIR technology in various applications and strategies used to detect foliar nutrients. The main outcomes of this chapter suggest that NIR technology provides reasonably accurate results when utilising strategically selected data pre-processing methods and statistical models that reduce spectral noise. An important element is understanding the impact of sample sizes, latent variables, and leaf water content were the main factors determining successful outcomes. NIR studies were reported to gather more accurate detection accuracies when compared to other portions of the electromagnetic spectrum (Elvidge, 1990; Kokaly & Clark, 1999; Peñuelas & Filella, 1998). Furthermore, the scientific underpinnings of the physiological basis for remote sensing of foliar biochemicals was discussed. More specifically, the physical laws of how remote sensing of foliar biochemicals is made possible using a non-destructive approach (Curran, 1989; Elvidge, 1990; Fourty *et al.*, 1996; Wessman *et al.*, 1989). The potential of using hyperspectral NIR remote sensing for foliar nutrient detection in a forestry environment. Hence, this chapter critically analysed a decade (2010-2020) of research using an integrative approach with a specific focus on NIR technology. In this chapter most research articles followed a similar approach to detecting nutrients in foliar samples (Figure 2.1). In summary, the importance of the NIR region of the electromagnetic spectrum formed the basis for the type of remote sensing used in this chapter. A reflection on future research motivates understanding the effects of epicuticle wax and trichomes on leaf optical properties. Overall, this chapter formed the foundation of the preceding chapters in this thesis.

Research objective 2 was achieved chapter 3. Chapter 3 tested the capability of remote sensing to detect nutrient deficiencies in a *Eucalyptus* variety using the RF algorithm. This chapter successfully detected full range macronutrients (N, P, K, Ca, Mg, Na) and micronutrients (Mn, Fe, Cu, Zn, B) using full-waveform hyperspectral data (350-2500 nm). The robustness of the RF algorithm produced promising results for certain macronutrients such as P and N (0.95 and 0.89, respectively) and micronutrients such as Mn and Cu (0.90 and 0.86, respectively). This chapter found the prediction results (R^2) of the most limiting growth nutrients N, P, K in this chapter explain the findings of previous studies (Adams *et al.*, 2000; Axelsson *et al.*, 2013; Özyiğit & Bilgen, 2013). Chapter 3 identified the red-edge, NIR, VIS and SWIR-2 regions of the electromagnetic spectrum as the most critical wavebands for detecting nutrient deficiencies using built-in RF measures importance (Figure 3.5). The outcomes of chapter 3 were used as a basis for chapter 4 of this thesis.

Research objective 3 was achieved in chapters 4. Chapter 4 intended to enhance the sampling strategy in chapter 3. The goal of chapter 4 was to find the most effective sampling position to enhance spectroscopic sampling procedures in a fully operational nursery environment for detecting nutrient deficiencies before infield planting. Furthermore, chapter 4 attempted to find the best position for sampling across the vertical canopy gradient of a sapling and to understand the translocation of nutrients throughout the sapling. Chapter 4 built on the study by Gara *et al.* (2018) who used a vertical canopy gradient model for sampling. Chapter 4 findings show the best place to take a representative sample is from VCP Q2 and Q4 for enhancing the detection of macronutrients and micronutrients using NIRS.

Research objective 4 was achieved in chapters 5. Chapter 5 compared the accuracies of very high-resolution imagery namely the Micasense sensor onboard a UAS platform and the PlanetScope imagery onboard a satellite platform (Table 5.1). This chapter successfully classified nutrients using a deep learning ANN for both platforms. As a result, the Micasense 5 waveband imagery performed slightly better than the satellite imagery with an overall classification accuracy of 87.62%, a KHAT statistic value of 0.83, and an error rate of 12.38% (Figure 5.5). Whereas the PlanetScope satellite imagery produced an overall classification accuracy of 81.50%, a KHAT statistic value of 0.75 and an error rate of 18.50%. This chapter concluded that high spectral resolution UAS imagery performed better than most studies who used RGB cameras. Furthermore, the UAS platform could capture more detail due to its higher spatial resolution (6 cm) when compared to the PlanetScope satellite imagery.

Research objective 5 was achieved in chapters 6. Chapter 6 was designed from the most successful findings of chapter 5. Hence, the best performing imagery in Chapter 6 was used to provide a framework for detecting macronutrients and micronutrients combined was implemented with using a deep learning ANN. Chapter 6 developed VI measures to identify the most important wavebands in the 5 waveband Micasense UAS image (red, green, blue, NIR, red-edge) which highly influenced the prediction model (Figure 6.5). Chapter 6 successfully predicted all macronutrients and micronutrients and developed VI measures that can be used in a specialized remote sensing framework. As a result, this chapter successfully predicted macronutrients and micronutrients with R^2's ranging between 0.54 and 0.75 with RMSE's below 0.08.

7.3 Research gaps

This thesis addressed the research gaps specifically pertaining to the remote sensing of nutrient biochemicals. Furthermore, this thesis uses a systematic approach by testing different methodologies across many different platforms. Chapters 3 and 4 were based on a nutrient-dependent nursery experiment to understand the physiological basis for remote sensing of foliar biochemicals. The foundation of the nursery experiments was to implement a controlled environment with known experimental inputs and outputs. Two different analytical spectrometer devices were tested, namely the ASD (chapter 3) and Bruker Fourier Transform-NIR (FT-NIR) spectrometer devices (chapter 4). The outcomes of chapter 3 suggest that future studies test the capabilities of using strategic portions (NIR region) of the electromagnetic spectrum to reduce spectral noise in the dataset and enable faster computing time. Chapter 4 improved the findings of Gara *et al.* (2018) by enhancing their sampling strategy by adding a fourth quartile which further segmented the sapling. Furthermore, this chapter added one more VCP (Q4) (Figure 4.3) and sampled both the adaxial and abaxial sides of leaf material which many studies have not addressed. Chapter 5 and 6 aimed to upscale chapters 3 and 4 from proximal (handheld) sensors and a nursery environment to airborne and satellite sensors in a compartment forest environment. The upscaling approach enables for transferability between scientific experiments and operationally using the science in a field environment. Furthermore, chapters 5 and 6 differed in the statistical approach: chapter 5 aimed to classify macronutrients and micronutrients whilst chapter 6 aimed to predict macronutrients and micronutrients. An important step was to understand the performance of classification accuracy before making infield predictions. The overall findings suggested an alternative screening framework for commercial forestry nurseries requiring quality planting material for long-term and short-term resource sustainability.

7.4 Strengths and limitations of the methodology

This thesis has many strengths and limitations in the methodologies used for its practical and operational use of remote sensing technology. In this section, the strengths and limitations are summarized described in more detail. In this thesis, strengths include sampling at large quantities, rapid monitoring of nutrient status, application of pre-processing data methods, reducing time and labour of acquiring a sample, and employing methods of variable selection. This thesis demonstrates the ability of remote sensing to sampling large scale compartment forest in a heterogeneous environment. This thesis describes the use of remote sensing for rapid detection and monitoring of the nutrient status of trees. The thesis demonstrates the use of pre-processing data methods for reducing spectral data dimensionality, spectral noise, data redundancy and impurity, especially when employing high dimensional and multivariate data. The thesis demonstrates that the technology is more efficient and effective in generating nutrient assessments by reducing the time and labour of acquiring a sample over a large area. Lastly, the thesis employs variable selection methods that enable scientists to test the ability of a spectral waveband to detect a feature more accurately in an operational environment.

In this thesis, limitations include the cost-benefit of the technology, physiological functions of trees, spectral noise, the influence of moisture content and trichomes, and image resolution accuracies. A limitation of technology exists in the affordability of the remote sensing platforms. UAS imagery RGB generally costs less than satellite imagery (multispectral or hyperspectral) due to the complexity, investment, and technology requirements for building and launching a satellite. UAS deliver exceptional high resolutions, although there are limitations such as gusty wind conditions, flight restrictions, battery capacity, and employing a skilled UAS pilot. However, satellite platforms do not offer the same versatility as UAS platforms, such as timely collection, administrative acquisition processes, zoning into the AOI, and image format issues. A limitation of the technology is to better understand relationship/link between the process of nutrient re-translocation at the species level and remote sensing. The physiological functions of the tree and the distribution of nutrients across the tree need to be further understood. The issue of spectral noise for the scope of requirements in the remote sensing of nutrient variables. Hyperspectral remote sensing is the more preferred as it produces narrower bands than multispectral imaging, however hyperspectral imagery is highly influenced by the presence of spectral noise. Furthermore, the influence of moisture content and trichomes and the interaction of electromagnetic waves on the leaves needs to be further understood when acquiring a representative sample.

The adoption of the technology for operational use in a commercial forestry industry presents a strength and limitation. The strength is that commercial forestry industries have already adopted the technology as it will improve their screening performance for large scale sampling before infield planting. To date, there are many commercial nurseries that are researching ways of improving their remote sensing frameworks for more accurate and efficient use of the technology. However, there are many institutions that have not transitioned fully from wet chemistry to remote sensing screening due to issues with the efficacy of technology and ideological reasoning when comparing to wet chemistry sampling methodologies. Hence, further research is required to improve the efficacy of the technology and create confidence in the use of the technology for decision makers to make more informed decisions.

7.5 Conclusions

This thesis aimed to investigate the potential use of remote sensing to accurately detect and map foliar nutrient deficiencies occurring within commercial forestry environments in KwaZulu-Natal, South Africa. The research undertaken in this thesis has demonstrated the capability of remotely sensed technologies to detect and map foliar nutrient deficiencies by comparing different imaging platforms and study areas. The main conclusions were based on the subsequent observations presented in this thesis:

1. Data cleaning for noise and obscurities using pre-processing data methods, statistical models can be produced for either prediction or classification. An important part is selecting the most suitable algorithm (statistical model). The influence of AI and better computing power will exceedingly enhance many of the pre-processing data methods and statistical models mentioned in this review.

2. The study successfully predicted N, P, K, Ca, Mg, S, Fe, Mn, Cu, Zn, and B in *E. grandis x E. urophylla* using hyperspectral data and RF analysis. The RF reduced the dataset's noise whilst producing competent results for certain macronutrients such as P

and N (0.95 and 0.89, respectively) and micronutrients such as Mn and Cu (0.90 and 0.86, respectively). The red-edge and NIR portions of the electromagnetic spectrum wavebands for detecting both macronutrients and micronutrients in *E. grandis x E. urophylla*.

3. The best positions to collect spectral measurements along the vertical canopy gradient are from VCP: Q2 and Q4 in this chapter. Overall, the study found no distinct VCP that better predicted all nutrients in this chapter.

4. Centimetre accurate ultra-high-resolution UAS imagery has proven more effective than PlanetScope satellite imagery in classifying macronutrients and micronutrients in commercial plantation forestry in KwaZulu Natal, South Africa. PlanetScope satellite imagery closely matched the results of the UAS image, which is valuable for mapping discrete land use biochemicals using a deep learning ANN.

5. The utilization of the NIR and red-edge wavebands highly contributed to the prediction model. A combination of very high-resolution imagery (<10cm GSD) and deep learning ANN can accurately predict full range macronutrients and micronutrients in compartment forest. Optimizing sensors with pre-set waveband configurations will enhance the speed and detection rates of these discrete foliar biochemicals.

7.6 Recommendations and future research

The future of foliar nutrient deficiency detection and mapping lies in further understanding the influences of plant's age, seasonality, and temperature that affect epicuticle wax and trichome production. Future research should investigate these influences by including this information into the mapping framework for in-depth analysis. Hence, an opportunity exists in understanding the impact of epicuticle wax and trichomes, moisture content, and the effects of sampling the adaxial and abaxial leaf surfaces across age, seasonality, temperature, and heterogeneous trees. While such variables would improve the detection accuracies of foliar nutrients and facilitate proper management decisions, monitoring areas at risk of nutrient depletion should be investigated.

Research should also focus on developing chemometric models that are plant variety-specific as a "plug and play" framework for detecting nutrient deficiencies rapidly. In this regard, forest managers should regularly implement detection and mapping techniques to provide temporal information related to the extent, distribution, and rate of nutrient depletion. The availability of affordable hyperspectral data sources would make high spatial resolution satellite data more accessible for nutrient monitoring applications in the forestry sector.

To further optimize the results of this thesis, future research should investigate efficient methods of reducing spectral noise in high spatial data, especially for foliar biochemicals which require narrower-band configurations. Effective spectral noise reduction strategies should enable efficient detection of nutrient stressed plants in a high-throughput environment. Hence, the continuous testing of new remote sensing sensors and the use of artificial intelligent algorithms will optimize the detection accuracies found in this thesis.

Finally, while this thesis focuses on detecting and mapping nutrient deficiencies within a commercial forestry environment, future research may consider diagnosing heterogeneous environments. In context, the automated detection techniques developed would be valuable to other agricultural and cropping sectors owing to the widespread and diverse ranges of nutrient deficiencies.

References

Abdel-Rahman, E. M., Mutanga, O., Odindi, J., Adam, E., Odindo, A., & Ismail, R. (2017). Estimating Swiss chard foliar macro-and micronutrient concentrations under different irrigation water sources using ground-based hyperspectral data and four partial least squares (PLS)-based (PLS1, PLS2, SPLS1 and SPLS2) regression algorithms. *Computers and Electronics in Agriculture, 132,* 21-33. doi: https://doi.org/10.1016/j.compag.2016.11.008

Abdi, H. (2010). Partial least squares regression and projection on latent structure regression (PLS Regression). *Wiley interdisciplinary reviews: computational statistics, 2*(1), 97-106

Aber, J. (1979a). Foliage-height profiles and succession in northern hardwood forests. *Ecology, 60*(1), 18-23

Aber, J. (1979b). A method for estimating foliage-height profiles in broad-leaved forests. *The Journal of Ecology*, 35-40

Acevedo, M., Rubilar, R., Dumroese, R. K., Ovalle, J. F., Sandoval, S., & Chassin-Trubert, R. (2020). Nitrogen loading of Eucalyptus globulus seedlings: nutritional dynamics and influence on morphology and root growth potential. *New Forests*, 1-16. doi: https://doi.org/10.1007/s11056-020-09778-2

Adams, M. L., Norvell, W. A., Philpot, W. D., & Peverly, J. H. (2000). Spectral detection of micronutrient deficiency in 'Bragg'soybean. *Agronomy Journal*, *92*(2), 261-268. doi: (https://doi.org/10.2134/agronj2000.922261x)

Adão, T., Hruška, J., Pádua, L., Bessa, J., Peres, E., Morais, R., & Sousa, J. (2017). Hyperspectral imaging: A review on UAV-based sensors, data processing and applications for agriculture and forestry. *Remote Sensing, 9*(11), 1110

Afandi, S. D., Herdiyeni, Y., Prasetyo, L. B., Hasbi, W., Arai, K., & Okumura, H. (2016). Nitrogen content estimation of rice crop based on near-infrared (NIR) reflectance using artificial neural network (ANN). *Procedia Environmental Sciences, 33*, 63-69. doi: http://10.1016/j.proenv.2016.03.057

Agjee, N. e. H., Mutanga, O., Peerbhay, K., & Ismail, R. (2018). The impact of simulated spectral noise on random forest and oblique random forest classification performance. *Journal of Spectroscopy, 2018*. doi: https://doi.org/10.1155/2018/8316918

Alsaqer, A. F., & Sasi, S. (2017). Movie review summarization and sentiment analysis using rapidminer *2017 International Conference on Networks & Advances in Computational Technologies (NetACT)* (pp. 329-335): IEEE.

Alton, P. B., & North, P. (2007). Interpreting shallow, vertical nitrogen profiles in tree crowns: A three-dimensional, radiative-transfer simulation accounting for diffuse sunlight. *Agricultural and forest meteorology, 145*(1-2), 110-124

Amirruddin, A. D., Muharam, F. M., Ismail, M. H., Ismail, M. F., Tan, N. P., & Karam, D. S. (2020). Hyperspectral remote sensing for assessment of chlorophyll sufficiency levels in mature oil palm (Elaeis guineensis) based on frond numbers: Analysis of decision tree and random forest. *Computers and Electronics in Agriculture*, *169*, 105221. doi: https://doi.org/10.1016/j.compag.2020.105221

Amirruddin, A. D., Muharam, F. M., & Mazlan, N. (2017). Assessing leaf scale measurement for nitrogen content of oil palm: performance of discriminant analysis and Support Vector Machine classifiers. *International Journal of Remote Sensing, 38*(23), 7260-7280. doi: https://doi.org/10.1080/01431161.2017.1372862

Ansari, M. S., Young, K. R., & Nicolas, M. E. (2016). Determining Wavelength for Nitrogen and Phosphorus Nutrients Through Hyperspectral Remote Sensing in Wheat (Triticum aestivum L.) Plant. *International Journal of Bio-Resource & Stress Management, 7*. doi: (https://doi.org/10.5958/0976-4038.2016.00125.1)

Asner, G. P. (1998). Biophysical and biochemical sources of variability in canopy reflectance. *Remote sensing of Environment, 64*(3), 234-253. doi: https://doi.org/10.1016/S0034-4257(98)00014-5

Asner, G. P., & Martin, R. E. (2009). Airborne spectranomics: mapping canopy chemical and taxonomic diversity in tropical forests. *Frontiers in Ecology and the Environment, 7*(5), 269-276

Asner, G. P., Martin, R. E., Knapp, D. E., Tupayachi, R., Anderson, C., Carranza, L., . . . Weiss, P. (2011). Spectroscopy of canopy chemicals in humid tropical forests. *Remote Sensing of Environment, 115*(12), 3587-3598

Atkinson, P. M., & Curran, P. J. (1995). Defining an optimal size of support for remote sensing investigations. *IEEE Transactions on Geoscience and Remote Sensing, 33*(3), 768-776. doi: http://10.1109/36.387592

Atkinson, P. M., & Tatnall, A. R. (1997). Introduction neural networks in remote sensing. *International Journal of remote sensing, 18*(4), 699-709

Au, J., Youngentob, K. N., Foley, W. J., Moore, B. D., & Fearn, T. (2020). Sample selection, calibration and validation of models developed from a large dataset of near-infrared spectra of tree leaves. *Journal of Near-Infrared Spectroscopy*, 0967033520902536. doi: https://doi.org/10.1177/0967033520902536

Axelsson, C., Skidmore, A. K., Schlerf, M., Fauzi, A., & Verhoef, W. (2013). Hyperspectral analysis of mangrove foliar chemistry using PLSR and support vector regression. *International Journal of Remote Sensing, 34*(5), 1724-1743. doi: (https://doi.org/10.1080/01431161.2012.725958)

Babaeian, E., Paheding, S., Siddique, N., Devabhaktuni, V. K., & Tuller, M. (2021). Estimation of root zone soil moisture from ground and remotely sensed soil information with multisensor data fusion and automated machine learning. *Remote Sensing of Environment, 260*, 112434

Bahar, N. H., Gauthier, P. P., O'Sullivan, O. S., Brereton, T., Evans, J. R., & Atkin, O. K. (2018). Phosphorus deficiency alters scaling relationships between leaf gas exchange and associated traits in a wide range of contrasting Eucalyptus species. *Functional Plant Biology, 45*(8), 813-826

Barbedo, J. G. A. (2019). A review on the use of unmanned aerial vehicles and imaging sensors for monitoring and assessing plant stresses. *Drones, 3*(2), 40

Belgiu, M., & Drăguţ, L. (2016). Random forest in remote sensing: A review of applications and future directions. *ISPRS Journal of Photogrammetry and Remote Sensing, 114*, 24-31. doi: (https://doi.org/10.1016/j.isprsjprs.2016.01.011)

Berg, J., & Nyström, K. (2018). A unified deep artificial neural network approach to partial differential equations in complex geometries. *Neurocomputing, 317*, 28-41

Berger, K., Verrelst, J., Féret, J.-B., Wang, Z., Wocher, M., Strathmann, M., . . . Hank, T. (2020). Crop nitrogen monitoring: Recent progress and principal developments in the context of imaging spectroscopy missions. *Remote Sensing of Environment, 242*, 111758

Bioucas-Dias, J. M., Plaza, A., Camps-Valls, G., Scheunders, P., Nasrabadi, N., & Chanussot, J. (2013). Hyperspectral remote sensing data analysis and future challenges. *IEEE Geoscience and remote sensing magazine, 1*(2), 6-36

Blackmer, T. M., Schepers, J. S., Varvel, G. E., & Walter-Shea, E. A. (1996). Nitrogen deficiency detection using reflected shortwave radiation from irrigated corn canopies. *Agronomy Journal, 88*(1), 1-5. doi: (https://digitalcommons.unl.edu/agronomyfacpub/338)

Boadi, S., Baah-Acheamfour, M., Ulzen-Appiah, F., & Murtaza Jamro, G. (2014). Nontimber forest product yield and income from Thaumatococcus daniellii under a mixed tree plantation system in Ghana. *International Journal of Forestry Research, 2014*. doi: (http://dx.doi.org/10.1155/2014/524863)

Breiman, L. (2001). Random forests. *Machine Learning, 45*(1), 5-32. doi: (https://doi.org/10.1023/A:1010933404324)

Breiman, L. (2017). *Classification and regression trees*: Routledge.

Bremner, J. (1960). Determination of nitrogen in soil by the Kjeldahl method. *The Journal of Agricultural Science, 55*(1), 11-33

Cael, J., Gardner, K., Koenig, J., & Blackwell, J. (1975). Infrared and Raman spectroscopy of carbohydrates. Paper V. Normal coordinate analysis of cellulose I. *The Journal of Chemical Physics, 62*(3), 1145-1153. doi: https://doi.org/10.1063/1.430558

Cai, Y., Guan, K., Nafziger, E., Chowdhary, G., Peng, B., Jin, Z., . . . Wang, S. (2019). Detecting in-season crop nitrogen stress of corn for field trials using UAV-and CubeSat-based multispectral sensing. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 12*(12), 5153-5166

Card, D. H., Peterson, D. L., Matson, P. A., & Aber, J. D. (1988). Prediction of leaf chemistry by the use of visible and near-infrared reflectance spectroscopy. *Remote Sensing of Environment, 26*(2), 123-147. doi: https://doi.org/10.1016/0034-4257(88)90092-2

Cavender-Bares, J., Gamon, J. A., & Townsend, P. A. (2020). *Remote Sensing of Plant Biodiversity*: Springer Nature.

Chase, C. W., Kimsey, M. J., Shaw, T. M., & Coleman, M. D. (2016). The response of light, water, and nutrient availability to pre-commercial thinning in dry inland Douglas-fir forests. *Forest Ecology and Management, 363*, 98-109

Chemura, A., Mutanga, O., Odindi, J., & Kutywayo, D. (2018). Mapping spatial variability of foliar nitrogen in coffee (Coffea arabica L.) plantations with multispectral Sentinel-2 MSI data. *ISPRS Journal of Photogrammetry and Remote Sensing, 138*, 1-11. doi: https://doi.org/10.1016/j.isprsjprs.2018.02.004

Chen, Jin, Y., & Brown, P. (2019). An enhanced bloom index for quantifying floral phenology using multi-scale remote sensing observations. *ISPRS Journal of Photogrammetry and Remote Sensing, 156*, 108-120

Chen, L., Wang, Y., Ren, C., Zhang, B., & Wang, Z. (2019). Assessment of multi-wavelength SAR and multispectral instrument data for forest aboveground biomass mapping using random forest kriging. *Forest Ecology and Management, 447*, 12-25

Cho, M. A., Skidmore, A., Corsi, F., Van Wieren, S. E., & Sobhan, I. (2007). Estimation of green grass/herb biomass from airborne hyperspectral imagery using spectral indices and partial least squares regression. *International journal of applied Earth observation and geoinformation, 9*(4), 414-424. doi: https://doi.org/10.1016/j.jag.2007.02.001

Cipullo, S., Nawar, S., Mouazen, A., Campo-Moreno, P., & Coulon, F. (2019). Predicting bioavailability change of complex chemical mixtures in contaminated soils using visible and near-infrared spectroscopy and random forest regression. *Scientific reports, 9*(1), 1-12. doi: https://doi.org/10.1038/s41598-019-41161-w

Cömert, Z., & Kocamaz, A. F. (2016). A study based on gray level co-occurrence matrix and neural network community for determination of hypoxic fetuses *International Artificial Intelligence and Data Processing Symposium (IDAP)* (pp. 569-573).

Congalton, R. G., & Green, K. (2019). *Assessing the accuracy of remotely sensed data: principles and practices*: CRC press.

Costa, L., Kunwar, S., Ampatzidis, Y., & Albrecht, U. (2021). Determining leaf nutrient concentrations in citrus trees using UAV imagery and machine learning. *Precision Agriculture*, 1-22

Crous, J., Morris, A., & Scholes, M. (2007). The significance of residual phosphorus and potassium fertiliser in countering yield decline in a fourth rotation of Pinus patula in Swaziland. *Southern Hemisphere Forestry Journal, 69*(1), 1-8

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

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

Darvishzadeh, R., Skidmore, A., Schlerf, M., & Atzberger, C. (2008). Inversion of a radiative transfer model for estimating vegetation LAI and chlorophyll in a heterogeneous grassland. *Remote Sensing of Environment, 112*(5), 2592-2604

Dash, J., Pearse, G., & Watt, M. (2018). UAV multispectral imagery can complement satellite data for monitoring forest health. *Remote Sensing, 10*(8), 1216

De Vries, W. (2014). Nutrients trigger carbon storage. *Nature Climate Change, 4*(6), 425-426

Demetriades-Shah, T. H., Steven, M. D., & Clark, J. A. (1990). High resolution derivative spectra in remote sensing. *Remote Sensing of Environment, 33*(1), 55-64. doi: https://doi.org/10.1016/0034-4257(90)90055-Q

Dixit, L., & Ram, S. (1985). Quantitative analysis by derivative electronic spectroscopy. *Applied Spectroscopy Reviews, 21*(4), 311-418. doi: https://doi.org/10.1080/05704928508060434

Donkin, M., Pearce, J., & Chetty, P. (1993a). Methods for routine plant analysis in the ICFR laboratories. *ICFR Bulletin Series, 8*, 93

Donkin, M., Pearce, J., & Chetty, P. (1993b). Methods for routine plant analysis in the ICFR laboratories. *, 6*(93), 8

Dordas, C. (2008). Role of nutrients in controlling plant diseases in sustainable agriculture. A review. *Agronomy for sustainable development, 28*(1), 33-46

Downes, G. M., Hudson, I., Raymond, C. A., Dean, G. H., Michell, A. J., Schimleck, L. R., . . . Muneri, A. (1997). *Sampling plantation eucalypts for wood and fibre properties*: CSIRO publishing.

Dye, P. (1996). Response of Eucalyptus grandis trees to soil water deficits. *Tree Physiology, 16*(1-2), 233-238. doi: https://doi.org/10.1093/treephys/16.1-2.233

Ellsworth, D., & Reich, P. (1993). Canopy structure and vertical patterns of photosynthesis and related leaf traits in a deciduous forest. *Oecologia, 96*(2), 169-178

Elvidge, C. D. (1990). Visible and near-infrared reflectance characteristics of dry plant materials. *Remote Sensing, 11*(10), 1775-1795. doi: https://doi.org/10.1080/01431169008955129

Escalante, H., Rodríguez-Sánchez, S., Jiménez-Lizárraga, M., Morales-Reyes, A., De La Calleja, J., & Vazquez, R. (2019). Barley yield and fertilization analysis from UAV imagery: a deep learning approach. *International Journal of Remote Sensing, 40*(7), 2493-2516

Eshkabilov, S., Lee, A., Sun, X., Lee, C. W., & Simsek, H. (2021). Hyperspectral imaging techniques for rapid detection of nutrient content of hydroponically grown lettuce cultivars. *Computers and Electronics in Agriculture, 181*, 105968

FAO. (2016). *FAO-KRC Regional cross-sectoral Policy Dialogue on Forests for Food Security*. http://www.fao.org/3/a-bl730e.pdf.

Fareed, N., & Rehman, K. (2020). Integration of Remote Sensing and GIS to Extract Plantation Rows from A Drone-Based Image Point Cloud Digital Surface Model. *ISPRS International Journal of Geo-Information, 9*(3), 151

Fassnacht, F. E., Latifi, H., Stereńczak, K., Modzelewska, A., Lefsky, M., Waser, L. T., . . . Ghosh, A. (2016). Review of studies on tree species classification from remotely sensed data. *Remote Sensing of Environment, 186*, 64-87

Fernández-Martínez, M., Vicca, S., Janssens, I. A., Sardans, J., Luyssaert, S., Campioli, M., . . . Obersteiner, M. (2014). Nutrient availability as the key regulator of global forest carbon balance. *Nature Climate Change, 4*(6), 471-476. doi: https://doi.org/10.1038/nclimate2177

Ferreira, G. W., Roque, J. V., Soares, E. M., da Silva, I. R., da Silva, E. F., Vasconcelos, A. d. A., & Teófilo, R. F. (2018). Temporal decomposition sampling and chemical characterization of eucalyptus harvest residues using NIR spectroscopy and chemometric methods. *Talanta*

Ferwerda, J. G., Skidmore, A. K., & Mutanga, O. (2005). Nitrogen detection with hyperspectral normalized ratio indices across multiple plant species. *International Journal of Remote Sensing, 26*(18), 4083-4095. doi: (https://doi.org/10.1080/01431160500181044)

Fife, D., Nambiar, E., & Saur, E. (2008). Retranslocation of foliar nutrients in evergreen tree species planted in a Mediterranean environment. *Tree physiology, 28*(2), 187-196

Forster, E. J., Healey, J. R., Dymond, C., & Styles, D. (2021). Commercial afforestation can deliver effective climate change mitigation under multiple decarbonisation pathways. *Nature Communications, 12*(1), 1-12

Fourty, T., Baret, F., Jacquemoud, S., Schmuck, G., & Verdebout, J. (1996). Leaf optical properties with explicit description of its biochemical composition: direct and inverse problems. *Remote sensing of Environment, 56*(2), 104-117

Galidaki, G., Zianis, D., Gitas, I., Radoglou, K., Karathanassi, V., Tsakiri–Strati, M., . . . Mallinis, G. (2017). Vegetation biomass estimation with remote sensing: focus on forest and other wooded land over the Mediterranean ecosystem. *International Journal of Remote Sensing, 38*(7), 1940-1966

Gamfeldt, L., Snäll, T., Bagchi, R., Jonsson, M., Gustafsson, L., Kjellander, P., . . . Philipson, C. D. (2013). Higher levels of multiple ecosystem services are found in forests with more tree species. *Nature communications*, *4*(1), 1-8
Gara, T. W., Darvishzadeh, R., Skidmore, A. K., & Wang, T. (2018). Impact of vertical canopy position on leaf spectral properties and traits across multiple species. *Remote sensing*, *10*(2), 346. doi: https://doi.org/10.3390/rs10020346
Gara, T. W., Darvishzadeh, R., Skidmore, A. K., Wang, T., & Heurich, M. (2019). Evaluating the performance of PROSPECT in the retrieval of leaf traits across canopy throughout the growing season. *International Journal of Applied Earth Observation and Geoinformation*, *83*, 101919
Garcia, K., & Zimmermann, S. D. (2014). The role of mycorrhizal associations in plant potassium nutrition. *Frontiers in plant science*, *5*, 337
Garcia, W., Amann, T., & Hartmann, J. (2018). Increasing biomass demand enlarges negative forest nutrient budget areas in wood export regions. *Scientific Reports*, *8*(1), 1-7. doi: https://doi.org/10.1038/s41598-018-22728-5
Ge, Y., Atefi, A., Zhang, H., Miao, C., Ramamurthy, R. K., Sigmon, B., . . . Schnable, J. C. (2019). High-throughput analysis of leaf physiological and chemical traits with VIS–NIR–SWIR spectroscopy: a case study with a maize diversity panel. *Plant methods*, *15*(1), 66
Girard, A., Schweiger, A. K., Carteron, A., Kalacska, M., & Laliberté, E. (2020). Foliar Spectra and Traits of Bog Plants across Nitrogen Deposition Gradients. *Remote Sensing*, *12*(15), 2448. doi: https://doi.org/10.3390/rs12152448
Goel, P. K., Prasher, S. O., Landry, J., Patel, R. M., Bonnell, R., Viau, A. A., & Miller, J. (2003). Potential of airborne hyperspectral remote sensing to detect nitrogen deficiency and weed infestation in corn. *Computers and Electronics in Agriculture*, *38*(2), 99-124. doi: (https://doi.org/10.1016/S0168-1699(02)00138-2)
Gomes, R., Bone, S., Cunha, M., Nahur, A. C., Moreira, P. F., Meneses-Filho, L., . . . Moutinho, P. (2010). Exploring the bottom-up generation of REDD+ policy by forest-dependent peoples. *Nature*, *2*, 836
Gracia-Romero, A., Kefauver, S. C., Vergara-Díaz, O., Hamadziripi, E., Zaman-Allah, M. A., Thierfelder, C., . . . Araus, J. L. (2020). Leaf versus whole-canopy remote sensing methodologies for crop monitoring under conservation agriculture: a case of study with maize in Zimbabwe. *Scientific reports*, *10*(1), 1-17
Guo, P.-T., Li, M.-F., Luo, W., & Cha, Z.-Z. (2019). Estimation of foliar nitrogen of rubber trees using hyperspectral reflectance with feature bands. *Infrared Physics & Technology*, *102*, 103021. doi: https://doi.org/10.1016/j.infrared.2019.103021
Guo, T.-t., Zhang, B., Guo, L., Li, D.-s., Wu, Y., Wu, J.-j., & Zhao, L. (2010). Classification of plant leaves by near-infrared spectroscopy using ANN and wavelet *2010 Second International Workshop on Education Technology and Computer Science* (Vol. 2, pp. 20-23): IEEE.
Guo, Y., Yin, G., Sun, H., Wang, H., Chen, S., Senthilnath, J., . . . Fu, Y. (2020). Scaling effects on chlorophyll content estimations with RGB camera mounted on a UAV platform using machine-learning methods. *Sensors*, *20*(18), 5130
Han, K. A. M., & Watchareeruetai, U. (2019). Classification of nutrient deficiency in black gram using deep convolutional neural networks *2019 16th International Joint Conference on Computer Science and Software Engineering (JCSSE)* (pp. 277-282): IEEE.
Hansen, & Schjoerring, J. (2003). Reflectance measurement of canopy biomass and nitrogen status in wheat crops using normalized difference vegetation indices and partial least squares regression. *Remote sensing of environment*, *86*(4), 542-553
Hansen, P., McRoberts, R., Isebrands, J., & Dixon, R. (1987). Determining CO2 exchange rate as a function of photosynthetic photon flux density. *Photosynthetica*, *21*, 98-101
Holmes, G., & Keiller, D. (2002). Effects of pubescence and waxes on the reflectance of leaves in the ultraviolet and photosynthetic wavebands: a comparison of a range of species. *Plant, Cell & Environment*, *25*(1), 85-93. doi: https://doi.org/10.1046/j.1365-3040.2002.00779.x

Huang, C.-Y., Wei, H.-L., Rau, J.-Y., & Jhan, J.-P. (2019). Use of principal components of UAV-acquired narrow-band multispectral imagery to map the diverse low stature vegetation fAPAR. *GIScience & remote sensing, 56*(4), 605-623
Huete, A. R. (1986). Separation of soil-plant spectral mixtures by factor analysis. *Remote Sensing of Environment, 19*(3), 237-251. doi: https://doi.org/10.1016/0034-4257(86)90055-6
Jay, S., Baret, F., Dutartre, D., Malatesta, G., Héno, S., Comar, A., . . . Maupas, F. (2019). Exploiting the centimeter resolution of UAV multispectral imagery to improve remote-sensing estimates of canopy structure and biochemistry in sugar beet crops. *Remote Sensing of Environment, 231*, 110898
Johansen, K., Duan, Q., Tu, Y.-H., Searle, C., Wu, D., Phinn, S., . . . McCabe, M. F. (2020). Mapping the condition of macadamia tree crops using multi-spectral uav and worldview-3 imagery. *ISPRS Journal of Photogrammetry and Remote Sensing, 165*, 28-40
Kahl, T., & Bauhus, J. (2014). An index of forest management intensity based on assessment of harvested tree volume, tree species composition and dead wood origin. *Nature Conservation, 7*, 15
Kanmani, W., & Jayapradha, B. (2017). Prediction of Default Customer in Banking Sector using Artificial Neural Network. *International Journal on Recent and Innovation Trends in Computing and Communication, 5*(7), 293-296
Kattenborn, T., Lopatin, J., Förster, M., Braun, A. C., & Fassnacht, F. E. (2019). UAV data as alternative to field sampling to map woody invasive species based on combined Sentinel-1 and Sentinel-2 data. *Remote Sensing of Environment, 227*, 61-73
Kembhavi, A., Harwood, D., & Davis, L. S. (2011). Vehicle detection using partial least squares. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 33*(6), 1250-1265
Khan, A. L., & Lee, I.-J. (2013). Endophytic Penicillium funiculosum LHL06 secretes gibberellin that reprograms Glycine max L. growth during copper stress. *BMC plant biology, 13*(1), 86
Kiala, Z., Odindi, J., Mutanga, O., & Peerbhay, K. (2016). Comparison of partial least squares and support vector regressions for predicting leaf area index on a tropical grassland using hyperspectral data. *Journal of Applied Remote Sensing, 10*(3), 036015. doi: (https://doi.org/10.1117/1.JRS.10.036015)
Knyazikhin, Y., Schull, M. A., Stenberg, P., Mõttus, M., Rautiainen, M., Yang, Y., . . . Lewis, P. (2013). Hyperspectral remote sensing of foliar nitrogen content. *Proceedings of the National Academy of Sciences, 110*(3), E185-E192
Köhl, M., Magnussen, S. S., & Marchetti, M. (2006). *Sampling methods, remote sensing and GIS multiresource forest inventory*: Springer Science & Business Media.
Kokaly, R. F., & Clark, R. N. (1999). Spectroscopic determination of leaf biochemistry using band-depth analysis of absorption features and stepwise multiple linear regression. *Remote sensing of environment, 67*(3), 267-287. doi: https://doi.org/10.1016/S0034-4257(98)00084-4
Koziol, P., Raczkowska, M. K., Skibinska, J., Urbaniak-Wasik, S., Paluszkiewicz, C., Kwiatek, W., & Wrobel, T. P. (2018). Comparison of spectral and spatial denoising techniques in the context of High Definition FT-IR imaging hyperspectral data. *Scientific reports, 8*(1), 1-11. doi: https://doi.org/10.1038/s41598-018-32713-7
Krishna, K. R. (2018). *Agricultural drones: a peaceful pursuit*: Taylor & Francis.
L3Harris Technologies, I. (2020). ENVI 5.2: environment for visualizing images. Florida, United States: 1988-2020 Harris Geospatial Solutions, Inc.
Lepine, L. C., Ollinger, S. V., Ouimette, A. P., & Martin, M. E. (2016). Examining spectral reflectance features related to foliar nitrogen in forests: Implications for broad-scale nitrogen mapping. *Remote Sensing of Environment, 173*, 174-186
Lequeue, G., Draye, X., & Baeten, V. (2016). Determination by near-infrared microscopy of the nitrogen and carbon content of tomato (Solanum lycopersicum L.) leaf powder. *Scientific reports, 6*, 33183. doi: https://doi.org/10.1038/srep33183
Li, L., Jákli, B., Lu, P., Ren, T., Ming, J., Liu, S., . . . Lu, J. (2018). Assessing leaf nitrogen concentration of winter oilseed rape with canopy hyperspectral technique considering a non-uniform vertical nitrogen distribution. *Industrial Crops and Products, 116*, 1-14

Liaw, A., & Wiener, M. (2002). Classification and regression by randomForest. *R News, 2*(3), 18-22
Liew, O. W., Chong, P. C. J., Li, B., & Asundi, A. K. (2008). Signature optical cues: emerging technologies for monitoring plant health. *Sensors*, *8*(5), 3205-3239. doi: (https://doi.org/10.3390/s8053205)
Liu, M., Wang, M., Wang, J., & Li, D. (2013). Comparison of random forest, support vector machine and back propagation neural network for electronic tongue data classification: Application to the recognition of orange beverage and Chinese vinegar. *Sensors and Actuators B: Chemical, 177*, 970-980
Loppi, S., Nelli, L., Ancora, S., & Bargagli, R. (1997). Passive monitoring of trace elements by means of tree leaves, epiphytic lichens and bark substrate. *Environmental Monitoring and Assessment*, *45*(1), 81-88
Louro, R. P., Dos Santos, A. V., & Machado, R. D. (1999). Ultrastructure of Eucalyptus grandis× Eucalyptus urophylla. I. Shoots cultivated in vitro in multiplication and elongation-rooting media. *International journal of plant sciences*, *160*(2), 217-227
Lu, B., He, Y., & Liu, H. (2018). Mapping vegetation biophysical and biochemical properties using unmanned aerial vehicles-acquired imagery. *International Journal of Remote Sensing*, *39*(15-16), 5265-5287
Lu, X., & Lu, S. (2015). Effects of adaxial and abaxial surface on the estimation of leaf chlorophyll content using hyperspectral vegetation indices. *International Journal of Remote Sensing*, *36*(5), 1447-1469. doi: https://doi.org/10.1080/01431161.2015.1012277
Lussem, U., Bolten, A., Menne, J., Gnyp, M., & Bareth, G. (2019). Ultra-high spatial resolution UAV-based imagery to predict biomass in temperate grasslands. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci*, 443-447
Luyssaert, S., Schulze, E.-D., Börner, A., Knohl, A., Hessenmöller, D., Law, B. E., . . . Grace, J. (2008). Old-growth forests as global carbon sinks. *Nature*, *455*(7210), 213-215
Ma, T., Wu, W., & Wang, Y. (2012). Transcriptome analysis of rice root responses to potassium deficiency. *BMC Plant Biology, 12*(1), 161
Mahajan, G., Sahoo, R., Pandey, R., Gupta, V., & Kumar, D. (2014). Using hyperspectral remote sensing techniques to monitor nitrogen, phosphorus, sulphur and potassium in wheat (Triticum aestivum L.). *Precision Agriculture, 15*(5), 499-522. doi: (https://doi.org/10.1007/s11119-014-9348-7)
Martin, R. E., Chadwick, K. D., Brodrick, P. G., Carranza-Jimenez, L., Vaughn, N. R., & Asner, G. P. (2018). An Approach for Foliar Trait Retrieval from Airborne Imaging Spectroscopy of Tropical Forests. *Remote Sensing*, *10*(2), 199
Martinelli, L. A., Almeida, S., Brown, I., Moreira, M., Victoria, R., Filoso, S., . . . Thomas, W. (2000). Variation in Nutrient Distribution and Potential Nutrient Losses by Selective Logging in a Humid Tropical Forest of Rondonia, Brazil 1. *Biotropica*, *32*(4a), 597-613
Masemola, C., & Cho, M. A. (2019). Estimating leaf nitrogen concentration from similarities in fresh and dry leaf spectral bands using a model population analysis framework. *International Journal of Remote Sensing*, *40*(17), 6841-6860. doi: https://doi.org/10.1080/01431161.2019.1597300
McEwan, A., & Steenkamp, J. (2014). Silviculture modernization in the South African forestry industry *Proceedings of the second international congress of silviculture, 26th–29th Novembe, Florence. Florence: Accademia Italiana di Scienze Forestali*.
McManus, K. M., Asner, G. P., Martin, R. E., Dexter, K. G., Kress, W. J., & Field, C. B. (2016). Phylogenetic structure of foliar spectral traits in tropical forest canopies. *Remote Sensing*, *8*(3), 196
Meacham-Hensold, K., Montes, C. M., Wu, J., Guan, K., Fu, P., Ainsworth, E. A., . . . Raines, C. (2019). High-throughput field phenotyping using hyperspectral reflectance and partial least squares regression (PLSR) reveals genetic modifications to photosynthetic capacity. *Remote sensing of environment*, *231*, 111176
Mee, C., Balasundram, S. K., & Hanif, A. H. M. (2017). Detecting and monitoring plant nutrient stress using remote sensing approaches: A review. *Asian J. Plant Sci*, *16*, 1-8. doi: http://10.3923/ajps.2017.1.8

Menesatti, P., Antonucci, F., Pallottino, F., Roccuzzo, G., Allegra, M., Stagno, F., & Intrigliolo, F. (2010). Estimation of plant nutritional status by Vis–NIR spectrophotometric analysis on orange leaves [Citrus sinensis (L) Osbeck cv Tarocco]. *Biosystems Engineering*, *105*(4), 448-454. doi: https://doi.org/10.1016/j.biosystemseng.2010.01.003
Meng, X.-L. (1994). Posterior predictive P-values. *The Annals of Statistics, 22*(3), 1142-1160. doi: http://doi:10.1214/aos/1176325622
Meuret, M., Dardenne, P., Biston, R., & Poty, O. (1993). The use of NIR in predicting nutritive value of Mediterranean tree and shrub foliage. *Journal of Near-Infrared Spectroscopy, 1*(1), 45-54
Meyer, J., Wood, R., McIntyre, R., & Leibbrandt, N. (1983). Classifying soils of the South African sugar industry on the basis of their nitrogen mineralizing capacities and organic matter contents *Proc S Afr Sug Technol Ass* (Vol. 57, pp. 151-158): Citeseer.
Montgomery, K., Henry, J. B., Vann, M. C., Whipker, B. E., Huseth, A. S., & Mitasova, H. (2020). Measures of canopy structure from low-cost UAS for monitoring crop nutrient status. *Drones, 4*(3), 36
Mucina, L., & Rutherford, M. (2006). The vegetation of South Africa, Lesotho and Swaziland. Strelitzia 19.,(South African National Biodiversity Institute: Pretoria, South Africa). *Memoirs of the Botanical Survey of South Africa*
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, 13*(8), 10823-10843
Murguzur, F. J. A., Bison, M., Smis, A., Böhner, H., Struyf, E., Meire, P., & Bråthen, K. A. (2019). Towards a global arctic-alpine model for Near-infrared reflectance spectroscopy (NIRS) predictions of foliar nitrogen, phosphorus and carbon content. *Scientific reports*, *9*(1), 1-10. doi: https://doi.org/10.1038/s41598-019-44558-9
Mutanga, Dube, T., & Ahmed, F. (2016). Progress in remote sensing: vegetation monitoring in South Africa. *South African Geographical Journal, 98*(3), 461-471. doi: (https://doi.org/10.1080/03736245.2016.1208586)
Mutanga, & Skidmore, A. (2004). Integrating imaging spectroscopy and neural networks to map grass quality in the Kruger National Park, South Africa. *Remote sensing of environment*, *90*(1), 104-115
Mutanga, Skidmore, A., & Prins, H. T. (2004a). Discriminating sodium concentration in a mixed grass species environment of the Kruger National Park using field spectrometry. *International journal of remote sensing, 25*(20), 4191-4201. doi: https://doi.org/10.1080/01431160410001720207
Mutanga, Skidmore, A. K., & Prins, H. (2004b). Predicting in situ pasture quality in the Kruger National Park, South Africa, using continuum-removed absorption features. *Remote sensing of Environment*, *89*(3), 393-408. doi: https://doi.org/10.1016/j.rse.2003.11.001
Mutanga, O., Adam, E., & Cho, M. A. (2012). High density biomass estimation for wetland vegetation using WorldView-2 imagery and random forest regression algorithm. *International Journal of Applied Earth Observation and Geoinformation, 18*, 399-406. doi: (https://doi.org/10.1016/j.jag.2012.03.012)
Mutowo, G., Mutanga, O., & Masocha, M. (2018). Evaluating the Applications of the Near-Infrared Region in Mapping Foliar N in the Miombo Woodlands. *Remote Sensing, 10*(4), 505. doi: https://doi.org/10.3390/rs10040505
Mutowo, G., Mutanga, O., & Masocha, M. (2019). Including shaded leaves in a sample affects the accuracy of remotely estimating foliar nitrogen. *GIScience & Remote Sensing*, 1-14
Myburg, A. A., Grattapaglia, D., Tuskan, G. A., Hellsten, U., Hayes, R. D., Grimwood, J., . . . Bauer, D. (2014). The genome of Eucalyptus grandis. *Nature, 510*(7505), 356-362. doi: https://doi.org/10.1038/nature13308
Mzinyane, T., Van Aardt, J., Ahmed, F., & Gebreslasie, M. (2016). Predicting soil nitrogen content using narrow-band indices from Eucalyptus grandis canopies as proxy. *South African Journal of Plant and Soil, 33*(1), 23-31
NERC. (Undated). *Excel Post Processing Templates and User Guides*. Retrieved from https://fsf.nerc.ac.uk/resources/post-processing/Excel.shtml

Nguyen, G., Rothstein, S., Spangenberg, G., & Kant, S. (2015). Role of microRNAs involved in plant response to nitrogen and phosphorous limiting conditions. *Frontiers in plant science*, *6*, 629

Norgaard, L., Saudland, A., Wagner, J., Nielsen, J. P., Munck, L., & Engelsen, S. B. (2000). Interval partial least-squares regression (iPLS): a comparative chemometric study with an example from near-infrared spectroscopy. *Applied Spectroscopy*, *54*(3), 413-419

Oliveira, L. F. R. d., Oliveira, M. L. R. d., Gomes, F. S., & Santana, R. C. (2017). Estimating foliar nitrogen in Eucalyptus using vegetation indexes. *Scientia Agricola*, *74*(2), 142-147. doi: (http://dx.doi.org/10.1590/1678-992x-2015-0477)

Oliveira, L. F. R. d., & Santana, R. C. (2020). Estimation of leaf nutrient concentration from hyperspectral reflectance in Eucalyptus using partial least squares regression. *Scientia Agricola*, *77*(6). doi: https://doi.org/10.1590/1678-992x-2018-0409

Oliveira, R. A., Näsi, R., Niemeläinen, O., Nyholm, L., Alhonoja, K., Kaivosoja, J., . . . Markelin, L. (2020). Machine learning estimators for the quantity and quality of grass swards used for silage production using drone-based imaging spectrometry and photogrammetry. *Remote Sensing of Environment*, *246*, 111830

Omer, G., Mutanga, O., Abdel-Rahman, E. M., Peerbhay, K., & Adam, E. (2017). Mapping leaf nitrogen and carbon concentrations of intact and fragmented indigenous forest ecosystems using empirical modeling techniques and WorldView-2 data. *Isprs Journal of Photogrammetry and Remote Sensing*, *131*, 26-39. doi: 10.1016/j.isprsjprs.2017.07.005

Osco, L., Junior, J., Ramos, A., Furuya, D., Santana, D., Teodoro, L., . . . Junior, C. (2020a). Leaf nitrogen concentration and plant height prediction for maize using UAV-based multispectral imagery and machine learning techniques. *Remote Sensing*, *12*(19), 3237

Osco, L., Ramos, A., & Faita, M. (2020b). A Machine Learning Framework to Predict Nutrient Content in Valencia-Orange Leaf Hyperspectral Measurements. *Remote Sensing*, *12*(6), 906. doi: https://doi.org/10.3390/rs12060906

Özyiğit, Y., & Bilgen, M. (2013). Use of spectral reflectance values for determining nitrogen, phosphorus, and potassium contents of rangeland plants. *Journal of Agricultural Science and Technology*, *15*(7), 1537-1545

Pajares, G. (2015). Overview and current status of remote sensing applications based on unmanned aerial vehicles (UAVs). *Photogrammetric Engineering & Remote Sensing*, *81*(4), 281-330. doi: https://doi.org/10.14358/PERS.81.4.281

Pasquini, C. (2018). Near-Infrared Spectroscopy: a mature analytical technique with new perspectives–A review. *Analytica chimica acta*. doi: https://doi.org/10.1016/j.aca.2018.04.004

Payn, T. W., Hill, R. B., HÖck, B. K., Skinner, M. F., Thorn, A. J., & Rijkse, W. C. (1999). Potential for the use of GIS and spatial analysis techniques as tools for monitoring changes in forest productivity and nutrition, a New Zealand example. *Forest Ecology and Management*, *122*(1-2), 187-196. doi: https://doi.org/10.1016/S0378-1127(99)00041-9

Peerbhay, K. Y., Mutanga, O., & Ismail, R. (2013). Commercial tree species discrimination using airborne AISA Eagle hyperspectral imagery and partial least squares discriminant analysis (PLS-DA) in KwaZulu–Natal, South Africa. *ISPRS Journal of Photogrammetry and Remote Sensing*, *79*, 19-28. doi: https://doi.org/10.1016/j.isprsjprs.2013.01.013

Peña-Lévano, L. M., Taheripour, F., & Tyner, W. E. (2019). Climate change interactions with agriculture, forestry sequestration, and food security. *Environmental and Resource Economics*, *74*(2), 653-675

Peng, Z., Guan, L., Liao, Y., & Lian, S. (2019). Estimating total leaf chlorophyll content of gannan navel orange leaves using hyperspectral data based on partial least squares regression. *IEEE Access*, *7*, 155540-155551. doi: http://10.1109/ACCESS.2019.2949866

Peñuelas, J., & Filella, I. (1998). Visible and near-infrared reflectance techniques for diagnosing plant physiological status. *Trends in plant science*, *3*(4), 151-156

Pereira, F. d. S., de Lima, J., Freitas, R., Dos Reis, A., do Amaral, L., Figueiredo, G., . . . Magalhães, P. (2022). Nitrogen variability assessment of pasture fields under an

integrated crop-livestock system using UAV, PlanetScope, and Sentinel-2 data. *Computers and Electronics in Agriculture, 193*, 106645
Peterson, D. L., Aber, J. D., Matson, P. A., Card, D. H., Swanberg, N., Wessman, C., & Spanner, M. (1988). Remote sensing of forest canopy and leaf biochemical contents. *Remote Sensing of Environment, 24*(1), 85-108. doi: https://doi.org/10.1016/0034-4257(88)90007-7
Pimstein, A., Karnieli, A., Bansal, S. K., & Bonfil, D. J. (2011). Exploring remotely sensed technologies for monitoring wheat potassium and phosphorus using field spectroscopy. *Field Crops Research, 121*(1), 125-135. doi: (https://doi.org/10.1016/j.fcr.2010.12.001)
Pinto, D. S., Resende, R. T., Mesquita, A., Rosado, A. M., & Cruz, C. D. (2014). Early selection in tests for growth traits of Eucalyptus urophylla clones test. *Scientia Forestalis, 42*(102), 251-257. doi: http://www.ipef.br/publicacoes/scientia/a
Porras-Soriano, A., Soriano-Martín, M. L., Porras-Piedra, A., & Azcón, R. (2009). Arbuscular mycorrhizal fungi increased growth, nutrient uptake and tolerance to salinity in olive trees under nursery conditions. *Journal of plant physiology, 166*(13), 1350-1359
Pullanagari, R., Kereszturi, G., & Yule, I. (2016). Mapping of macro and micro nutrients of mixed pastures using airborne AisaFENIX hyperspectral imagery. *ISPRS Journal of Photogrammetry and Remote Sensing, 117*, 1-10
Pullanagari, R., Kereszturi, G., Yule, I., & Irwin, M. (2018). Determining Uncertainty Prediction Map of Copper Concentration in Pasture from Hyperspectral Data Using Qunatile Regression Forest *IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium* (pp. 3809-3811): IEEE.
Quentin, A. G., Rodemann, T., Doutreleau, M. F., Moreau, M., & Davies, N. W. (2017). Application of near-infrared spectroscopy for estimation of non-structural carbohydrates in foliar samples of Eucalyptus globulus Labilladiere. *Tree Physiology, 37*(1), 131-141. doi: (https://doi.org/10.1093/treephys/tpw083
Reich, P., Ellsworth, D., Kloeppel, B., Fownes, J., & Gower, S. (1990). Vertical variation in canopy structure and CO2 exchange of oak-maple forests: influence of ozone, nitrogen, and other factors on simulated canopy carbon gain. *Tree Physiology, 7*(1-2-3-4), 329-345
Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., & Carvalhais, N. (2019). Deep learning and process understanding for data-driven Earth system science. *Nature, 566*(7743), 195-204
Rengel, Z. (2007). Cycling of micronutrients in terrestrial ecosystems *Nutrient cycling in terrestrial ecosystems* (pp. 93-121): Springer.
Reuter, D., & Robinson, J. B. (1997). *Plant Analysis: An Interpretation Manual* (Second ed.): CSIRO publishing.
Richardson, A. E., Barea, J.-M., McNeill, A. M., & Prigent-Combaret, C. (2009). Acquisition of phosphorus and nitrogen in the rhizosphere and plant growth promotion by microorganisms. *Plant and soil, 321*(1-2), 305-339
Ring, E. (2000). The discovery of infrared radiation in 1800. *The Imaging Science Journal, 48*(1), 1-8. doi: https://doi.org/10.1080/13682199.2000.11784339
Rodrigues, M., Nanni, M. R., Cezar, E., dos Santos, G. L. A. A., Reis, A. S., de Oliveira, K. M., & de Oliveira, R. B. (2020). Vis–NIR spectroscopy: from leaf dry mass production estimate to the prediction of macro-and micronutrients in soybean crops. *Journal of Applied Remote Sensing, 14*(4), 044505. doi: https://doi.org/10.1117/1.JRS.14.044505
Rubilar, R. A., Allen, H. L., Fox, T. R., Cook, R. L., Albaugh, T. J., & Campoe, O. C. (2018). Advances in silviculture of intensively managed plantations. *Current Forestry Reports, 4*(1), 23-34. doi: https://doi.org/10.1007/s40725-018-0072-9
Ruxton, G. D. (2006). The unequal variance t-test is an underused alternative to Student's t-test and the Mann–Whitney U test. *Behavioral Ecology, 17*(4), 688-690. doi: (https://doi.org/10.1093/beheco/ark016)
Sasaki, K., Kawata, S., & Minami, S. (1984). Estimation of component spectral curves from unknown mixture spectra. *Applied optics, 23*(12), 1955-1959. doi: https://doi.org/10.1364/AO.23.001955
Savitzky, A., & Golay, M. J. (1964). Smoothing and differentiation of data by simplified least squares procedures. *Analytical chemistry, 36*(8), 1627-1639

Schepaschenko, D., Chave, J., Phillips, O. L., Lewis, S. L., Davies, S. J., Réjou-Méchain, M., . . . Herault, B. (2019). The Forest Observation System, building a global reference dataset for remote sensing of forest biomass. *Scientific data*, *6*(1), 1-11

Schmitt, S., Garrigues, S., & de la Guardia, M. (2014). Determination of the mineral composition of foods by infrared spectroscopy: A review of a green alternative. *Critical reviews in analytical chemistry*, *44*(2), 186-197. doi: https://doi.org/10.1080/10408347.2013.835695

Shi, H., Lei, Y., Prates, L. L., & Yu, P. (2019). Evaluation of near-infrared (NIR) and Fourier transform mid-infrared (ATR-FT/MIR) spectroscopy techniques combined with chemometrics for the determination of crude protein and intestinal protein digestibility of wheat. *Food chemistry*, *272*, 507-513. doi: https://doi.org/10.1016/j.foodchem.2018.08.075

Silva, J. A., & Uchida, R. (2000). Plant Nutrient Management in Hawaii's Soils, Approaches for Tropical and Subtropical Agriculture. *Plant Nutrient Management in Hawaii's Soils*, 31-55

Singh, A., Serbin, S. P., McNeil, B. E., Kingdon, C. C., & Townsend, P. A. (2015). Imaging spectroscopy algorithms for mapping canopy foliar chemical and morphological traits and their uncertainties. *Ecological Applications*, *25*(8), 2180-2197. doi: https://doi.org/10.1890/14-2098.1

Singh, L., Mutanga, O., Mafongoya, P., & Peerbhay, K. (2017a). Remote sensing of key grassland nutrients using hyperspectral techniques in KwaZulu-Natal, South Africa. *Journal of Applied Remote Sensing*, *11*(3), 036005

Singh, L., Mutanga, O., Mafongoya, P., & Peerbhay, K. Y. (2017b). Multispectral mapping of key grassland nutrients in KwaZulu-Natal, South Africa. *Journal of Spatial Science*, 1-18

Sozzi, M., Kayad, A., Gobbo, S., Cogato, A., Sartori, L., & Marinello, F. (2021). Economic comparison of satellite, plane and uav-acquired NDVI images for site-specific nitrogen application: Observations from Italy. *Agronomy*, *11*(11), 2098

Stein, B. R., Thomas, V. A., Lorentz, L. J., & Strahm, B. D. (2014). Predicting macronutrient concentrations from loblolly pine leaf reflectance across local and regional scales. *GIScience & Remote Sensing*, *51*(3), 269-287

Suomalainen, J., Oliveira, R. A., Hakala, T., Koivumäki, N., Markelin, L., Näsi, R., & Honkavaara, E. (2021). Direct reflectance transformation methodology for drone-based hyperspectral imaging. *Remote Sensing of Environment*, *266*, 112691

Taddese, H., Asrat, Z., Burud, I., Gobakken, T., Ørka, H. O., Dick, Ø. B., & Næsset, E. (2020). Use of Remotely Sensed Data to Enhance Estimation of Aboveground Biomass for the Dry Afromontane Forest in South-Central Ethiopia. *Remote Sensing*, *12*(20), 3335

Team, R. (2020). RStudio: Integrated Development for R (Version Version 1.2.5042). RStudio, Inc., Boston, MA. Retrieved from http://www.rstudio.com/

Timmer, V. (1997). Exponential nutrient loading: a new fertilization technique to improve seedling performance on competitive sites. *New Forests*, *13*(1-3), 279-299. doi: https://doi.org/10.1023/A:1006502830067

Tsai, F., & Philpot, W. (1998). Derivative analysis of hyperspectral data. *Remote Sensing of Environment*, *66*(1), 41-51

Tu, Y., Bian, M., Wan, Y., & Fei, T. (2018). Tea cultivar classification and biochemical parameter estimation from hyperspectral imagery obtained by UAV. *PeerJ*, *6*, e4858

Türker-Kaya, S., & Huck, C. (2017). A review of mid-infrared and near-infrared imaging: principles, concepts and applications in plant tissue analysis. *Molecules*, *22*(1), 168

Turner, J., & Lambert, M. (2017). Analysis of Foliage Phosphorus Requirements of Radiata Pine Plantations. *Communications in Soil Science and Plant Analysis*, *48*(18), 2218-2229. doi: (https://doi.org/10.1080/00103624.2017.1409355)

Ulissi, V., Antonucci, F., Benincasa, P., Farneselli, M., Tosti, G., Guiducci, M., . . . Pari, L. (2011). Nitrogen concentration estimation in tomato leaves by VIS-NIR non-destructive spectroscopy. *Sensors*, *11*(6), 6411-6424. doi: (https://doi.org/10.3390/s110606411)

Ustin, S. L. (2013). Remote sensing of canopy chemistry. *Proceedings of the national academy of sciences*, *110*(3), 804-805

Ustin, S. L., & Jacquemoud, S. (2020). How the Optical Properties of Leaves Modify the Absorption and Scattering of Energy and Enhance Leaf Functionality *Remote Sensing of Plant Biodiversity* (pp. 349-384): Springer, Cham.

van der Meer, F. (2018). Near-infrared laboratory spectroscopy of mineral chemistry: A review. *International journal of applied earth observation and geoinformation, 65*, 71-78. doi: https://doi.org/10.1016/j.jag.2017.10.004

van der Tol, C., Vilfan, N., Dauwe, D., Cendrero-Mateo, M. P., & Yang, P. (2019). The scattering and re-absorption of red and near-infrared chlorophyll fluorescence in the models Fluspect and SCOPE. *Remote sensing of environment, 232*, 111292. doi: https://doi.org/10.1016/j.rse.2019.111292

Van Deventer, H., Cho, M., Mutanga, O., & Ramoelo, A. (2015). Capability of models to predict leaf N and P across four seasons for six sub-tropical forest evergreen trees. *ISPRS Journal of Photogrammetry and Remote Sensing, 101*, 209-220

Vanderbilt, V. C., & Grant, L. (1985). Plant canopy specular reflectance model. *IEEE Transactions on Geoscience and Remote Sensing*(5), 722-730. doi: http://0196-2892/85/0900-0722

Varhola, A., Coops, N. C., Weiler, M., & Moore, R. D. (2010). Forest canopy effects on snow accumulation and ablation: An integrative review of empirical results. *Journal of Hydrology, 392*(3-4), 219-233. doi: https://doi.org/10.1016/j.jhydrol.2010.08.009

Vessey, J. K. (2003). Plant growth promoting rhizobacteria as biofertilizers. *Plant and soil, 255*(2), 571-586

Wallis, C. I., Homeier, J., Peña, J., Brandl, R., Farwig, N., & Bendix, J. (2019). Modeling tropical montane forest biomass, productivity and canopy traits with multispectral remote sensing data. *Remote sensing of environment, 225*, 77-92

Walshe, D., McInerney, D., Van De Kerchove, R., Goyens, C., Balaji, P., & Byrne, K. A. (2020). Detecting nutrient deficiency in spruce forests using multispectral satellite imagery. *International Journal of Applied Earth Observation and Geoinformation, 86*, 101975

Wang, Townsend, P. A., Schweiger, A. K., Couture, J. J., Singh, A., Hobbie, S. E., & Cavender-Bares, J. (2019). Mapping foliar functional traits and their uncertainties across three years in a grassland experiment. *Remote sensing of environment, 221*, 405-416

Wang, Q., & Li, P. (2013). Canopy vertical heterogeneity plays a critical role in reflectance simulation. *Agricultural and forest meteorology, 169*, 111-121

Wang, Y., Wang, F., Huang, J., Wang, X., & Liu, Z. (2009). Validation of artificial neural network techniques in the estimation of nitrogen concentration in rape using canopy hyperspectral reflectance data. *International Journal of Remote Sensing, 30*(17), 4493-4505

Wang, Z., Skidmore, A. K., Darvishzadeh, R., & Wang, T. (2018). Mapping forest canopy nitrogen content by inversion of coupled leaf-canopy radiative transfer models from airborne hyperspectral imagery. *Agricultural and Forest Meteorology, 253*, 247-260. doi: (https://doi.org/10.1016/j.agrformet.2018.02.010)

Wang, Z., Skidmore, A. K., Wang, T., Darvishzadeh, R., Heiden, U., Heurich, M., . . . Hearne, J. (2017). Canopy foliar nitrogen retrieved from airborne hyperspectral imagery by correcting for canopy structure effects. *International Journal of Applied Earth Observation and Geoinformation, 54*, 84-94

Warburton, P., Brawner, J., & Meder, R. (2014). *Handheld near-infared spectroscopy for the prediction of Leaf physiological status in tree seedlings* (Report No. 1751-6552). Retrieved from Journal of Near-Infrared Spectroscopy.

Ward, K. J., Chabrillat, S., Neumann, C., & Foerster, S. (2019). A remote sensing adapted approach for soil organic carbon prediction based on the spectrally clustered LUCAS soil database. *Geoderma, 353*, 297-307

Watt, M. S., Pearse, G. D., Dash, J. P., Melia, N., & Leonardo, E. M. C. (2019). Application of remote sensing technologies to identify impacts of nutritional deficiencies on forests. *ISPRS Journal of Photogrammetry and Remote Sensing, 149*, 226-241

Wessman, C. A., Aber, J. D., & Peterson, D. L. (1989). An evaluation of imaging spectrometry for estimating forest canopy chemistry. *International Journal of Remote Sensing, 10*(8), 1293-1316. doi: https://doi.org/10.1080/01431168908903969

Weyer, L. G. (1985). Near-infrared spectroscopy of organic substances. *Applied Spectroscopy Reviews, 21*(1-2), 1-43. doi: https://doi.org/10.1080/05704928508060427

Williams, P., & Norris, K. (1987). *Near-infrared technology in the agricultural and food industries*: American Association of Cereal Chemists, Inc.

Windley, H. R., & Foley, W. J. (2015). Landscape-scale analysis of nutritional traits of New Zealand tree foliage using near-infrared spectroscopy. *Forest Ecology and Management, 357*, 161-170

Xulu, S., Gebreslasie, M. T., & Peerbhay, K. Y. (2019). Remote sensing of forest health and vitality: a South African perspective. *Southern Forests: a Journal of Forest Science, 81*(2), 91-102

Yousefpour, R., Augustynczik, A. L. D., Reyer, C. P., Lasch-Born, P., Suckow, F., & Hanewinkel, M. (2018). Realizing mitigation efficiency of European commercial forests by climate smart forestry. *Scientific reports, 8*(1), 1-11

Yu, P., Zhang, M., Ziyan, X., Yang, T., Yali, S., Zhou, T., . . . Lin, Y. (2020). Estimation of leaf nutrition status in degraded vegetation based on field survey and hyperspectral data. *Scientific Reports (Nature Publisher Group), 10*(1)

Zeng, Y., Badgley, G., Dechant, B., Ryu, Y., Chen, M., & Berry, J. A. (2019). A practical approach for estimating the escape ratio of near-infrared solar-induced chlorophyll fluorescence. *Remote Sensing of Environment, 232*, 111209. doi: https://doi.org/10.1016/j.rse.2019.05.028

Zha, H., Miao, Y., Wang, T., Li, Y., Zhang, J., Sun, W., . . . Kusnierek, K. (2020). Improving unmanned aerial vehicle remote sensing-based rice nitrogen nutrition index prediction with machine learning. *Remote Sensing, 12*(2), 215

Zhai, Y., Cui, L., Zhou, X., Gao, Y., Fei, T., & Gao, W. (2013). Estimation of nitrogen, phosphorus, and potassium contents in the leaves of different plants using laboratory-based visible and near-infrared reflectance spectroscopy: comparison of partial least-square regression and support vector machine regression methods. *International journal of remote sensing, 34*(7), 2502-2518. doi: (https://doi.org/10.1080/01431161.2012.746484)

Zhang, S., Li, J., Wang, S., Huang, Y., Li, Y., Chen, Y., & Fei, T. (2020). Repaid Identification and Prediction of Cadmium–Lead Cross-Stress of Different Stress Levels in Rice Canopy Based on Visible and Near-Infrared Spectroscopy. *Remote Sensing, 12*(3), 469. doi: https://doi.org/10.3390/rs12030469

Zhang, X., Liu, F., He, Y., & Gong, X. (2013). Detecting macronutrients content and distribution in oilseed rape leaves based on hyperspectral imaging. *Biosystems engineering, 115*(1), 56-65

Zhao, Caiafa, C. F., Mandic, D. P., Chao, Z. C., Nagasaka, Y., Fujii, N., . . . Cichocki, A. (2013). Higher order partial least squares (HOPLS): a generalized multilinear regression method. *IEEE transactions on pattern analysis and machine intelligence, 35*(7), 1660-1673

Zhao, C., Li, H., Li, P., Yang, G., Gu, X., & Lan, Y. (2016). Effect of vertical distribution of crop structure and biochemical parameters of winter wheat on canopy reflectance characteristics and spectral indices. *IEEE Transactions on Geoscience and Remote Sensing, 55*(1), 236-247

Zhao, D., Reddy, K. R., Kakani, V. G., & Reddy, V. (2005). Nitrogen deficiency effects on plant growth, leaf photosynthesis, and hyperspectral reflectance properties of sorghum. *European journal of agronomy, 22*(4), 391-403

Zhu, X., Ma, L., Li, C., Tang, L., Zhao, Y., & Zhang, H. (2019). Sampling strategy and product validation over nonuniform surface-based on TEM and CGM upscaling: A case study on LAI. *Nonlinear Optics and Applications XI, 11026*, 110261I. doi: https://doi.org/10.1117/12.2525310