

**Estimating mineral content of indigenous browse species using
laboratory spectroscopy and Sentinel-2 imagery**

by

Michael Jabulani Gama

Submitted in partial fulfilment of the requirements for the degree

Master of Science

in

Forest Management and the Environment

In the Department of Plant and Soil Sciences

Faculty of Natural and Agricultural Sciences

University of Pretoria

Pretoria

Supervisor: Prof M.A. Cho

Co-supervisor: Prof P.W. Chirwa

July 2018

DECLARATION

I, Michael Jabulani Gama declare that the thesis entitled “estimating mineral content of indigenous browse species using laboratory spectroscopy and Sentinel-2 imagery”, which I hereby submit for the degree of Master of Science at the University of Pretoria, is my own work and has not been submitted previously by me for a degree at this or any other tertiary institution.

SIGNATURE:

DATE:

ACKNOWLEDGEMENTS

The author would like to thank the following for their contribution and support towards the completion of this research project:

Professor Moses Azong Cho and Professor Paxie Chirwa for stimulating ideas on the subject matter, reviewing the content and their constant guidance;

Professor Wyne Twine for collaboration and support during field campaigns at the Wits Rural Facility (WRF);

Professor Coert Geldenhuys for his inputs in conceptualising the research idea;

Mrs Cecilia Masemola for assistance with data handling and analysis;

My entire family; and special thanks to Samke, Sphe and Mpilo for their patience and understanding during the course of my study;

Sincere gratitude goes to the following institutions:

The South African Forestry Company (SAFCOL) for funding my studies;

The Council for Scientific and Industrial Research (CSIR) and

The University of the Witwatersrand, for allowing me access to their facilities.

ABSTRACT

Understanding variation in the foliar nutrient among indigenous species of the bushveld is crucial for rural livelihoods, in particular the integration of trees into agroecosystems. The study explored nutrient composition of common browse species with regard to nitrogen (N), phosphorus (P), potassium (K) and calcium (Ca) using leaf spectra (400–2500nm) and chemical data collected from nine bushveld species, along with partial least squares (PLS) analysis. The work further explored the relationship between canopy reflectance of Sentinel-2 image and foliar properties of the identified species.

Spectroscopic analysis reveals useful information about nitrogen at leaf and canopy scales whereas modelling reflectance using satellite image did not yield satisfactory results. At the leaf level, nitrogen was highly correlated with leaf spectral reflectance ($R^2=0.72$, $p<0.05$) for winter and ($R^2=0.88$, $p<0.05$) for summer. The coefficient of determination for the relationships between leaf reflectance and concentrations of phosphorus, potassium and calcium were low. Modelling the relationship using Sentinel-2 data also showed higher correlations ($R^2=0.44$, $p<0.05$) for nitrogen compared with the other nutrients investigated.

Table of contents

CHAPTER 1: INTRODUCTION	1
1.1 Background	1
1.2 Science-based agroforestry	2
1.3 Hyperspectral reflectance of leaf material	3
1.4 Monitoring leaf nutrients using satellite data	4
1.4 Problem statement	4
1.5 Objectives of the study	5
1.6 Research questions	6
CHAPTER 2: LITERATURE REVIEW	7
2.1 Productivity in agroecosystems	7
2.2 Spectral analysis	8
2.3 Monitoring bio-resources using remote sensing	16
CHAPTER 3: MATERIALS AND METHODS	18
3.1 General description of the study area	18
3.2 Data collection	19
3.3 Chemical analysis	20
3.4 Spectral measurements	22
3.5 Satellite data	24
CHAPTER 4: RESULTS AND DISCUSSION	28
4.1 Foliar nutrient variations	28
4.2. Nutrient concentration and spectral reflectance of dried leaf samples	35
4.3. Modelling nutrient concentrations using Sentinel-2 images	40
CHAPTER 5: CONCLUSION	43
REFERENCES	45

List of tables

Table 2.1: Overview of wavelength ranges and materials used in the papers reviewed	9
Table 2.2: Steps in the multivariate model-construction	13
Table 3.1: Tree species investigated in the study	19
Table 3.2: List of Sentinel-2 images used in the study	25
Table 3.3: Description of Sentinel-2 bands used in the study	27
Table 4.1: Summary of the chemical data recorded	28
Table 4.2: Pairwise comparisons for nitrogen in the wet season	30
Table 4.3: Pairwise comparisons for nitrogen in the dry season	30
Table 4.4: Pairwise comparisons for nitrogen in both wet and dry seasons	31
Table 4.5: Performance of the models on the estimation of N, P, K & Ca	34
Table 4.6: Performance of models for Sentinel-2 data	40

List of figures

Figure 1.1: Workflow outlining the processes followed in the study	6
Figure 2.1: FieldSpec® 3 Spectroradiometer	11
Figure 2.2: Multivariate analysis - classification of chemometrics methods	12
Figure 3.1: Map of the study area	18
Figure 3.2: The Spectro Genesis ICP-OES: SPECTRO Analytical Instruments	21
Figure 3.3: Collecting leaf spectra using FieldSpec® 3 spectroradiometer	22
Figure 3.4: Reflectance spectra of dried ground leaf samples	23
Figure 3.5: Workflow of iCOR atmospheric correction	26
Figure 4.1: Folia nutrient variation by species	29
Figure 4.2: Foliar nutrient variation by group	33
Figure 4.3: Foliar nutrient variation between dry and wet seasons	34
Figure 4.4: Cross-validated RMSEP curves for chemical vs hyperspectral data: winter	36
Figure 4.5: Cross-validated predictions for chemical vs hyperspectral data: winter	37
Figure 4.6: Cross-validated RMSEP curves for chemical vs hyperspectral data: summer	38
Figure 4.7: Cross-validated predictions for chemical vs hyperspectral data: summer	39
Figure 4.8: Cross-validated RMSEP curves for Sentinel-2 data	41
Figure 4.9: Cross-validated predictions for Sentinel-2 data	42

Appendices

Appendix A01: Mean foliar nutrient composition for tree species	56
Appendix A02: Statistical analysis for chemical data	58
Appendix A03: Q-Q Plots for N, P, K & Ca	61
Appendix A04: The continuum-removed spectra and spectral analysis	62
Appendix A05: Spectral bands and indices used from Sentinel-2 image	64

LIST OF ABBREVIATIONS

ASD	:	analytical spectral devices
ANN	:	artificial neural network
ESA	:	European Space Agency
EO	:	earth observation
GPS	:	global positioning system
MLR	:	multiple linear regressions
MSI	:	multispectral instrument
NIR	:	near infrared
PCR	:	principal component analysis
PLSR	:	partial least squares regression
RE	:	red-edge
RF	:	random forest
RMSEP:		root mean square error of prediction
SWIR	:	shortwave infrared
Vis	:	visible light
WRF	:	Wits Rural Facility

CHAPTER 1

INTRODUCTION

1.1 Background

Trees are important natural assets, which play vital economic and environmental roles, especially with regard to crop and livestock productivity. They constitute fodder for livestock and game in the arid and semi-arid zones (Chepape *et al.*, 2011). Fodder trees and shrubs contain appreciable amounts of nutrients that are deficient in other feed resources such as grasses especially during dry seasons (Lukhele & van Ryssen, 2000). The integration of trees in production landscapes is known as agroforestry; and it bestows a wide range of ecosystem benefits to improve the quality of life of rural households (Mbow *et al.*, 2014).

Agroforestry is recognised as a viable option for optimising land productivity and reducing pressure on the indigenous forests (Mukolwe, 1999). Due to the persistent threat of food shortages, projected climate change and rising prices of agricultural inputs, agroecology has experienced special interest from research and development communities (Akinnifesi *et al.*, 2010; Chirwa *et al.*, 2015). It is considered a cost-effective and ecologically sound approach to enhance food security particularly in the era of adverse environmental conditions (Mbow *et al.*, 2014). Considerable work has focussed on integration of indigenous trees into agroecosystems to enhance productivity, improve income opportunities and address conservation related issues (Kwesiga *et al.*, 2003; Garrity *et al.*, 2006; Kalaba *et al.*, 2010; Chirwa *et al.*, 2015).

Indigenous trees demonstrate the potential to supplement exotic species due to the minimal impact they have on the environment (Mukolwe, 1999; Lukhele & van Ryssen, 2000; Tegegne, 2008). Possible applications include establishment of indigenous multipurpose trees on agro-ecological systems and on sensitive sites such as riparian zones, water-stressed catchments and land cleared of alien plants to improve land management and protect biodiversity (Everson *et al.*, 2011). While the potential benefits of agroforestry are well documented, particular aspects need to be adapted to suit the areas where the trees are introduced (Akinnifesi *et al.*, 2010).

Scientific basis for actions is critical to enhance the conservation and sustainable use of organic resources and their contributions to human well-being. Nassoro *et al.* (2015) assert that the use of browse trees generally as supplements to ruminant feeding is limited by lack of information on the nutritive potential. This necessitates the need for efficient methods to investigate their composition for purposes of conservation and utilisation in agricultural production systems (Chirwa *et al.*, 2015). Among the research imperatives is the need to domesticate new tree species to provide a wide range of products and services (Garrity *et al.*, 2006).

1.2 Science-based agroforestry

Palm *et al.* (2001) assert that despite the critical services that organic inputs provide to agricultural productivity, the use of organic materials for fodder and soil fertility management has been based primarily on trial and error. The resource quality of a multipurpose plant varies with the plant species, plant parts and their maturity. Therefore, it is essential that these qualities are investigated. Plant materials are classified by taxonomic family, genus, and species and whether they are able to nodulate and fix nitrogen or not (Palm *et al.*, 2001).

Leaf nitrogen is one of the key factors limiting agricultural production and ecosystem functioning. It is a good indicator of photosynthetic capacity of plants and measure of rangeland quality (Ramoelo & Cho, 2018). Nitrogen is related to protein and biomass and it plays a crucial role in understanding the feeding patterns and distribution of browsers. The concept of cultivating and domesticating high value indigenous tree species features prominently in the research and development agenda of most developing countries. Efforts made by countries such as the Democratic Republic of Congo to test methods of agroforestry using native trees are helping to take pressure off the continent's national parks (World Agroforestry Centre, 2015).

The nutrient composition of some browse species has been investigated using conventional wet techniques (e.g. Lukhele & Ryssen, 2000). Methods based on tissue analysis have been widely applied due to their reliability. However, they are costly because they require use of reagents, time-consuming and destructive (Muñoz-Huerta *et al.*, 2013). A significant barrier to plant mineral analysis in general is the price versus the perceived value by farmers (Rossa *et al.*, 2015). For instance, the delay in processing time with wet analysis can result in significant losses on

crops, should any corrective action be necessary (van Maarschalkerweerd & Husted, 2015).

To overcome these limitations, efforts have been directed to the study and development of spectroscopy in the visible to shortwave infrared (Vis-SWIR) which is fast, non-destructive and less costly (Huang *et al.*, 2009). The composition of an organic matrix can be estimated by reflectance spectra due to organic functional groups or the association between minerals and the organic functional groups (García-Sánchez *et al.*, 2017). Reflectance spectroscopy offers rapid plant mineral analysis with instant results (Nomngongo *et al.*, 2017). Remote sensing techniques have also been proven to be valuable for in-field plant nutrient estimation and distribution patterns on landscapes (Shi *et al.*, 2015).

1.3 Hyperspectral reflectance of leaf material

Absorption of electromagnetic radiation in the near infrared (NIR) and other parts of the electromagnetic spectrum reflects the molecular composition of a sample, and is routinely used for fast analysis in science and industry (Rossa *et al.*, 2015). Spectral features within the shortwave and near-infrared region have been shown to correlate with the chemical composition of organic materials such as dried and ground plant leaves (Riley & Canaves, 2002). The technique is widely accepted as a tool for analysis in fields such as the food and agricultural industries (İlknur ŞEN, 2003).

Near infrared (NIR) equipment can record spectra for solid and liquid samples with no pre-treatment, implement continuous methodologies, provide spectra quickly and predict physical and chemical parameters from the spectrum (Rossa *et al.*, 2015). Spectrometers used are essentially identical with those employed in other regions of the electromagnetic spectrum. Equipment can incorporate a variety of devices, depending on the characteristics of the sample and the particular analytical conditions (Blanco & Villarroya, 2002).

Quantification of macronutrients and micronutrients by NIR spectroscopy in soils has been demonstrated by various studies (e.g. İlknur ŞEN, 2003; Awiti *et al.*, 2007). It is reported that measurement of soil N demonstrated good results whilst calibrations for P and K in soil were less successful (Ward *et al.*, 2011). The potential of NIR technique for rapid analysis of N content in cotton leaves (Riley & Cánaves, 2002),

and macro- and micronutrients in sugarcane (Yarce & Rojas, 2012) and in yerba mate indicate that the technique is useful (Rossa *et al.*, 2015).

1.4 Monitoring leaf nutrients using satellite data

Monitoring foliar nutrients using traditional methods of leaf harvesting and transportation to laboratories for analysis implies a number of difficulties. Species of interest are sometimes inaccessible, because of dense overgrowth or they could be located in swamps (van Deventer *et al.*, 2015). Remote sensing techniques complement ground based monitoring systems. They essentially involve the ability to detect and characterise unique patterns of nutrient phenology across species, seasons and regions (Cho *et al.*, 2012).

In spite of the advances made in mapping foliar nutrients using remote sensing, the relationship between foliar nutrient concentration and spectral reflectance across species, season and ecosystems remains poorly understood (Ramoelo *et al.*, 2015). The relationship between foliar nutrients and spectral information of multipurpose trees in relation to seasonal variation is of significance in resource characterisation (van Deventer *et al.*, 2015). Due to their fine spatial resolution and reasonable number of spectral bands, new sensors such as RapidEye, WorldView-2 and Sentinel-2 are a trade-off in terms of the benefits offered by multispectral and hyperspectral remotely sensed data. Their spectral bands are configured within unique portions of the electromagnetic spectrum such as the red-edge (Dhau *et al.*, 2017).

The recent availability of time series of Sentinel-2 imagery at fine spatial resolution (10m) and high temporal frequency (every 5 days with S-2A and S-2B) represents a significant step in the use of Earth Observation (EO) data for the monitoring of forest resources (Simonetti *et al.*, 2017). Sentinel-2 imagery is freely available since 2015 through the Copernicus programme of the European Commission. This study therefore tested this method; and the results henceforth could offer an opportunity for monitoring and mapping tree species that are alternative sources of nitrogen in agroecosystems.

1.4 Problem statement

The potential of leaf spectroscopy as a robust technique for simultaneous prediction of leaf biochemical properties is well acknowledged. However, the challenge is to move the application of laboratory spectroscopy toward a diagnostic screening tool that can aid the development of reliable spectral case definition to characterize fodder quality of various plant species for agricultural and environmental management. The lack of suitable reference material of high quality to correlate the nutrient concentrations to fast spectral measurements is an eminent setback.

Although canopy-level hyperspectral measurements have been linked to plant chemistry, the relationships are frequently lower in precision and accuracy compared to those of leaf-level studies. It remains unclear how well leaf properties can be retrieved from spectral reflectance data acquired from Sentinel-2. While laboratory soil and foliar nutrient analysis is common practice, similar studies have not been conducted to estimate foliar nutrients of trees in the subtropical or bushveld regions using Sentinel-2 data. The multilevel approach can enhance rapid characterization of organic resources at leaf, canopy and landscape scales.

1.5 Objectives of the study

Overall aim of the study was to investigate the utility of spectral data at the laboratory level and spaceborne platform in predicting the leaf nutrients: nitrogen (N), phosphorus (P), potassium (K) and calcium (Ca) of leguminous and non-legumes species of the Lowveld ecosystem of South Africa.

The specific objectives of the study were to:

1. compare leaf nitrogen levels amongst the broad-leaved leguminous, narrow-leaved leguminous and non-leguminous plants of the savanna ecosystem;
2. evaluate laboratory spectroscopy on dry samples in estimating the concentration of nitrogen, phosphorus, potassium and calcium in leaves of identified browse species between dry and wet seasons;
3. investigate the relationship between foliar nutrients (N, P, K and Ca) and spectral reflectance at spaceborne level using Sentinel-2 data.

1.6 Research questions

The questions to be investigated were:

1. What are the patterns of leaf nitrogen variation among the broad-leaved leguminous, fine-leaved leguminous and non-leguminous plants of the Lowveld savanna between dry and wet seasons?
2. Can the quantities of nitrogen, phosphorus, potassium and calcium of the trees be accurately predicted using hyperspectral reflectance of dry samples?
3. Can Sentinel-2 satellite image be used to estimate foliar nitrogen, phosphorus, potassium and calcium in the selected browse species of the savanna in South Africa?

The conceptual framework of the study is outlined in Figure 1.1.

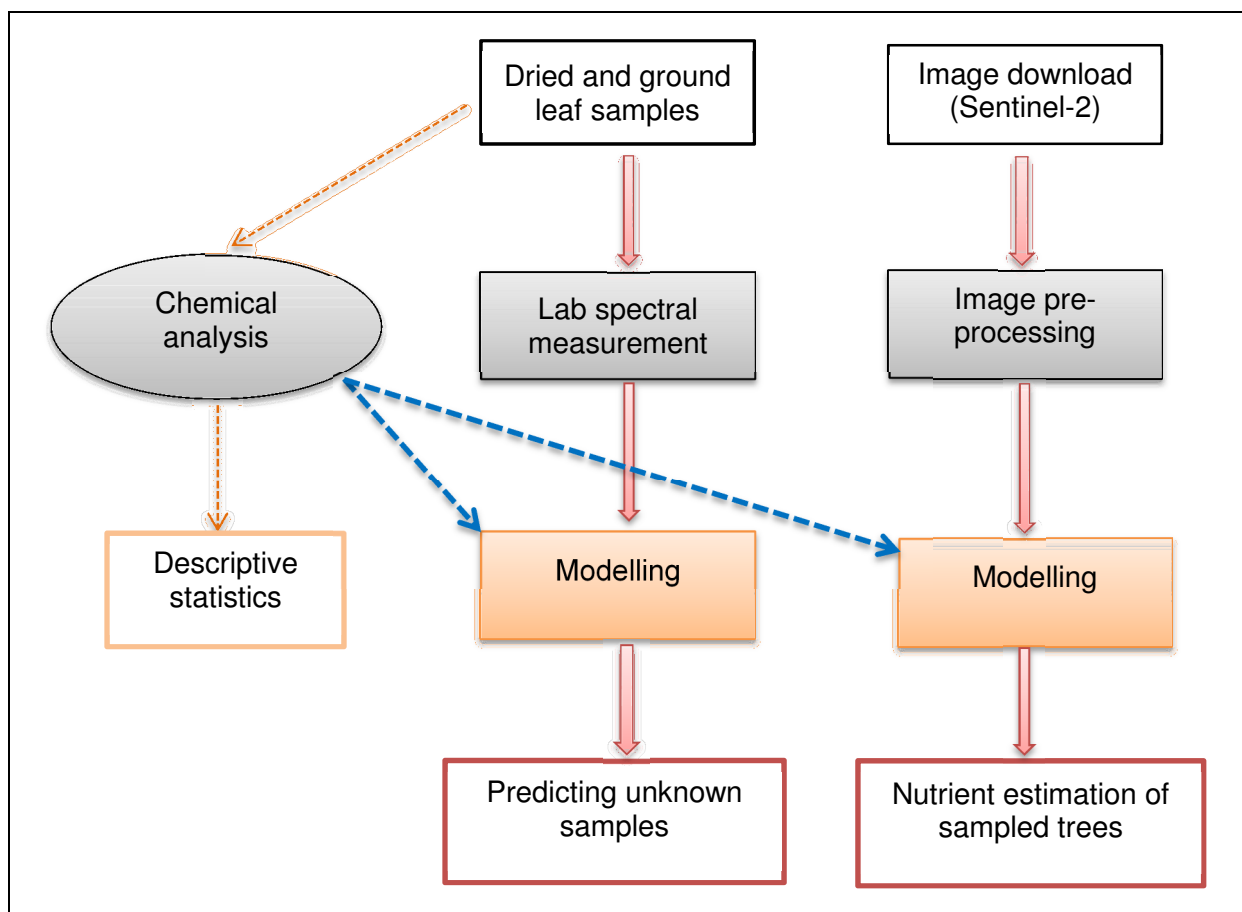


Figure 1.1: Workflow outlining the processes followed in the study.

CHAPTER 2

LITERATURE REVIEW

2.1 Productivity in agroecosystems

Indigenous browse species are an important source of nutrients for livestock and game in southern Africa. They remain green for longer, into the dry season compared with grasses (Lukhele and van Ryssen, 2003). Some browse plants have high crude protein content - up to about 25% on a dry matter basis (Chepape *et al.*, 2011). Garrity *et al.* (2006) stress the need for a predictive understanding for the management of multipurpose trees in tropical agroecosystems.

Numerous studies have since reported nutrient contents, resource quality, decomposition, and nutrient release patterns for a variety of plant species from tropical agroecosystems (Palm *et al.*, 2001). The quality of organic resource depends on the plant species and the stage of maturity of the leaves. It is therefore essential that quality indicators are identified for the various plants (Akinnifesi *et al.*, 2010). Nair (1993) highlighted that nitrogen is generally considered a good indicator of feed quality as litter that is high in nitrogen is decomposed rapidly compared to woody residues and other lignified materials. Other than taxonomic classification, plants material varies qualitatively according to the plant parts (Palm *et al.*, 2001).

The comparison between nutrient compositions of plants of different families has been explored. Palm *et al.* (2001) revealed that plants in Leguminosae and Asteraceae families contained significantly higher amounts of leaf N compared with plants in other families, which contained less than 3% N (about 18.75% crude protein). This is consistent with the findings by Lukhele and van Ryssen (2003), who reported that many browse species found in the bushveld region contained less than 17% crude protein. El hassan *et al.* (2000) analysed multipurpose leguminous trees and concluded that they could be good protein supplements, provided they are degraded in the digestive tract of the animal and are non-toxic.

Palm *et al.* (2001) further highlighted that in general there is variability in nutrient content for plant species - within and between locations. The authors assert that if there is little variability in values for a particular species, it may not be necessary to analyse those materials repeatedly in the future. However, if there is considerable

variability then correlations might be established between resource quality parameters and climate, soil or genotypic provenances. According to Lukhele and van Ryssen (2003), a seasonal effect was observed in *Combretum apiculatum*, where the crude protein concentration in winter was significantly lower than that of other seasons, while variations in mineral content between seasons was not significant for other browse species investigated.

Plant tissue analysis has been used as a reference technique to estimate plant nutrients in relation to recent fast analytical methods (Muñoz-Huerta *et al.*, 2013). The combustion method for total nitrogen determination proposed by Dumas (1831) is one of the reference methods for the determination of the total nitrogen in an organic matrix. The sample is combusted at high temperature in an oxygen atmosphere. Via subsequent oxidation and reduction tubes, nitrogen is quantitatively converted to nitrogen gas (N₂). Other volatile combustion products are either trapped or separated. A thermal conductivity detector measures nitrogen gas and the results are given as percentage or milligrams of nitrogen, which may be converted into crude protein by using conversion factors (Muñoz-Huerta *et al.*, 2013).

2.2 Spectral analysis

2.2.1 Theory and application

Hyperspectral reflectance of leaf material is a fast and non-destructive technique that is used to provide multi-constituent analysis (García-Sánchez *et al.*, 2017). The technique covers wavelength range closer to the mid-infrared and broadens up to the visible region (Nomgongo *et al.*, 2017). The basis for Vis-SWIR spectroscopy is molecular vibrations. When a sample is irradiated, electromagnetic radiation is absorbed selectively according to the specific vibration frequencies of the molecules present (Givens & Deaville, 1999). Incoming radiation with a frequency corresponding to that of the molecular vibrations is absorbed, and the remaining is either reflected or transmitted and this gives rise to a spectrum which yields information about the sample's molecular composition (Stuart, 2004).

Bonds commonly found in biological systems such as C-H, O-H and N-H bonds are infrared active (Yarce & Rojas, 2012), whereas molecules such as Cl₂ or H₂C=CH₂ are considered as infrared inactive (İlknur ŞEN, 2003; Ward *et al.*, 2011). The

frequency of molecular vibrations depends on the strength of chemical bonds and the mass of each atom involved. The hydrogen atom is the lightest thus exhibits the largest vibrations and the greatest deviations from harmonic behaviour; therefore the main bands typically observed in the NIR region correspond to bonds containing hydrogen and other light atoms (Blanco & Villarroya, 2002; İlknur ŞEN, 2003).

Overtone and combination bands of hydrogen-bearing functional groups (C-H, O-H, and N-H) dominate the NIR spectra of plant material due to the light mass of the hydrogen atom. These show broad and overlapping bands which are less suited for structural studies but offer some advantages for quantitative analysis (García-Sánchez *et al.*, 2017). The absorption features of certain biochemicals obtained through NIR spectroscopy of dried and ground leaf specimens were outlined by Curran (1989).

Table 2.1: Overview of wavelength ranges and materials used in the papers reviewed

Wavelength range (nm)	Material(s) evaluated	Nutrients	Reference
830–2500	Soil	N, P, K	İlknur ŞEN (2003)
400–2500	Canopy (fresh leaves)	N, P, Chlorophyll a and b, carotenoids, anthocyanins, leaf water and SLA	Asner & Martin (2008)
830–2500	Dry, ground grasses	N, P, K	Ward <i>et al.</i> (2011)
400–2500	Dry, ground wheat and rice straw	K, Ca, Mg, Fe	Huang <i>et al.</i> (2009)
830–2500	Dry, ground sugarcane leaves	N, P, K, Ca, Mg, Cu, Zn, Mn, Fe	Yarce & Rojas (2012)
1000–2500	Fresh barley leaves	Cu	van Maarschalkerweerd <i>et al.</i> (2015)
1000-2500	Dry, ground yerba mate leaves	N, C, P, K, Ca, Mg, Na, Fe, Mn, Cu, Zn	Rossa <i>et al.</i> (2015)

It was also determined at early stages in chronological research that NIRS can be used to analyse macro and micronutrients where a consistent correlation between a mineral and a spectroscopically active compound exists (Murray, 1986; van Maarschalkerweerd *et al.*, 2015). The mineral may directly form part of the compound or be essential in its biosynthesis (de Aldana *et al.*, 1995). It has been

reported, for instance, that magnesium is associated with chlorophyll; and potassium is associated with organic acids, especially malate (Givens & Deaville, 1999, Yarce & Rojas, 2012). The wavelength ranges and materials investigated in the papers reviewed are shown in Table 2.1. The technique can make multiple measurements within a few seconds and has high sensitivity relative to other infrared techniques (İlknur ŞEN, 2003).

The indirect nature of spectroscopic methods for plant mineral analysis makes strict validation of the analytical concentration range and specificity crucial (van Maarschalkerweerd & Husted, 2015). At very low concentrations, a linear relationship between nutrient concentration and spectroscopic data may fade due to induction of side reactions and biotic or abiotic stresses can influence the spectra (Huang *et al.*, 2009).

The relationship between concentration of a specific component and absorbed energy is further complicated by overlapping of spectral bands from different constituents present in the sample (Stuart, 2004). Since there is no mathematical law to describe the interaction of radiation with a scattering medium containing a heterogeneous distribution of absorbing species, NIR spectroscopy is largely an empirical secondary technique requiring calibration using samples of known composition determined using standard methods (Givens & Deaville, 1999).

The FieldSpec® 3 (Analytical Spectral Devices Inc., Boulder, Colorado, USA) shown in Figure 2.1, is a reflectance spectrophotometer operating in the diffuse reflectance mode (Awiti *et al.*, 2007). The data obtained is typically broad, extensively overlapped bands of spectra in the range 350-2500nm. The device provides state-of-the-art, real-time performance and creates powerful information that helps to improve, simplify and streamline research, and production processes for a multitude of material measurement solutions (Awiti *et al.*, 2007). NIR spectroscopy requires chemometrics to extract as much relevant information as possible (Wold *et al.*, 2001).

In most investigations, NIR spectra are measured on dried and ground plant material to ensure homogeneity and to avoid interference from water (Huang *et al.*, 2009).

Kumar *et al.* (2006) highlighted that spectral analysis of fresh leaves is more problematic than that of dried leaves wherein the dominant effect of water absorption



Figure 2.1: FieldSpec® 3 Spectroradiometer (ASD Inc., USA)

(<https://www.asdi.com/about-us/news/asd-introduces-fieldspec-3-spectroradiometer>

[accessed: 23 Feb 2017]).

largely masks the signatures of the biochemical components over the 1850nm-2000 nm region (Givens & Deaville, 1999). The cell structure of fresh plant material scatters radiation as it passes through multiple air and water surfaces, a phenomenon that may obscure the subtle biochemical absorption features (Kumar *et al.*, 2006). Variations in particle size can account for up to 90% of the variance in NIR spectra and can substantially influence the predicted values (Huang *et al.*, 2009).

2.2.2 Multivariate analysis

The indirect correlation between NIR spectra and nutrient concentrations means that great caution needs to be taken during method development (van Maarschalkerweerd & Husted, 2015). The process should minimise interference from the spectra of strongly overlapping constituents and from light scatter variations by combining, spectral data from many different wavelengths (Givens & Deaville, 1999). The next step is to derive a mathematical model relating the spectral data to the reference or chemical data, a process called chemometrics (İlknur ŞEN, 2003). The main steps to be followed in constructing a model are shown in Table 2.2.

A number of multivariate analysis methods are available including multiple stepwise linear regression (MSLR), partial least squares regression (PLSR), modified partial least squares (MPLS) and principal component regression (PCR), as outlined in Figure 2.2 (Roussel *et al.*, 2014).

Multivariate analysis methods have the basic form:

$$Y = a + b_1X_1 + b_2X_2 + \dots + b_nX_n \quad (\text{Equation 1})$$

where Y is the component to be calibrated; a the intercept; X_1, X_2, \dots, X_n are independent spectral variables; and b_1, b_2, \dots, b_n are regression coefficients. In calibration, a set of X s and known Y s are used to derive the regression coefficients, whereas in validation and subsequent prediction, a set of X s and the derived regression coefficients are used to predict an unknown Y (Wold *et al.*, 2001; İlknur ŞEN, 2003).

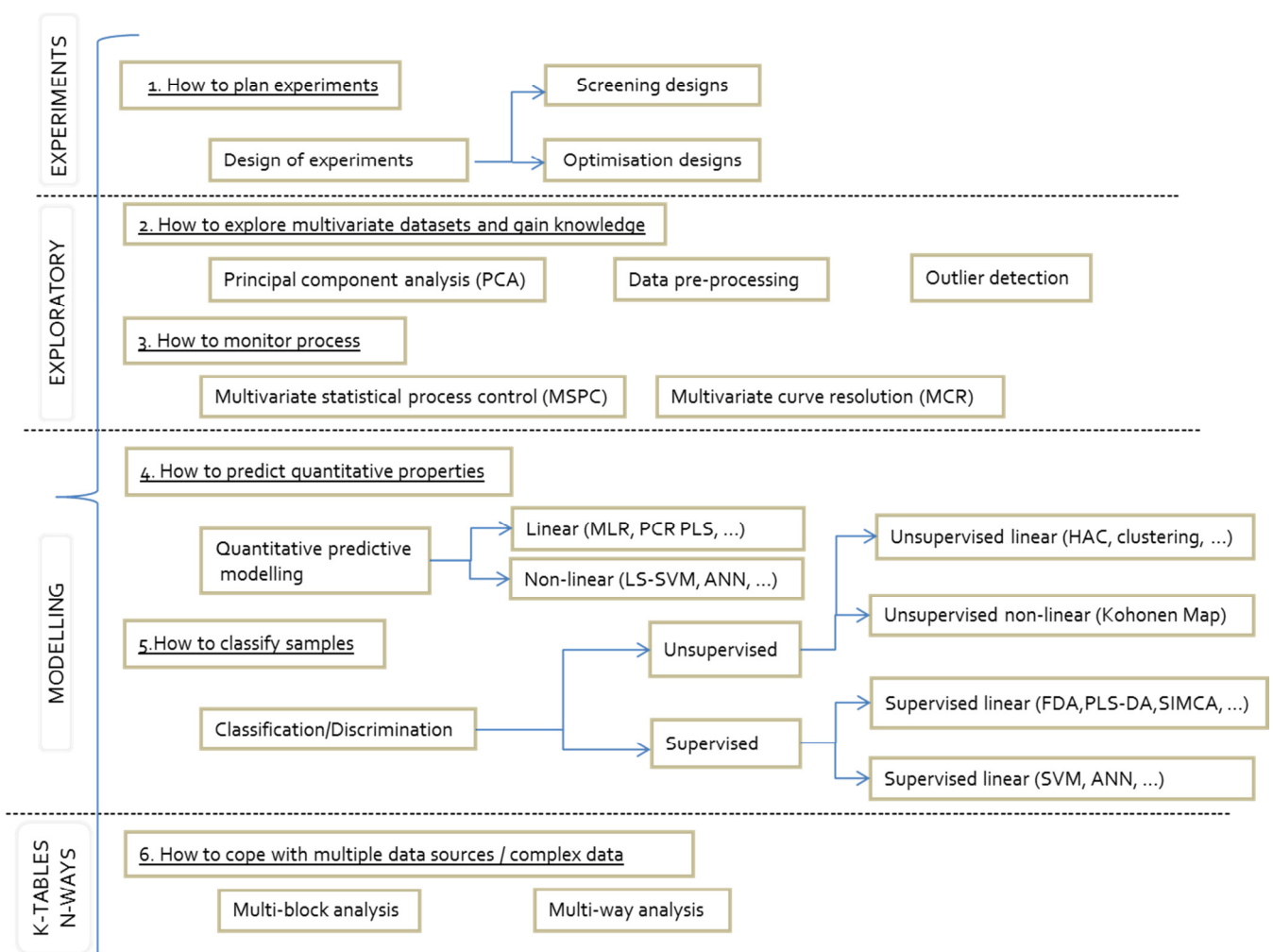


Figure 2.2: Multivariate analysis - classification of chemometrics methods (adapted from Roussel *et al.*, 2014).

Table 2.2: Steps in the multivariate model-construction

Step	Purpose
1. Choosing the calibration samples	To select a set of samples representative of the whole population
2. Determining the target parameter using the reference method	To determine the value of the measured property in an accurate, precise manner. The quality of the value dictates that of the calibration model
3. Recording of NIR spectra	To obtain physico–chemical information in a reproducible manner
4. Pre-treatment of spectra	To reduce unwanted contributions (such as shifts and scatter) to the spectra
5. Constructing the model	To establish the spectrum–property relationship using multivariate methods
6. Validating the model	To ensure that the model accurately predicts the property of interest in samples not subjected to the calibration process
7. Predicting unknown samples	To predict rapidly the property of interest in new, unknown samples

(Blanco & Villarroya, 2002).

2.2.3 Pre-treatment of spectra

Before the analysis, the X - and Y -variables are often transformed to make their distributions fairly symmetrical to minimize the contributions of irrelevant information into spectra (Barnes *et al.*, 1989; Wold *et al.*, 2001). This enables development of simple but more robust models (Givens & Deaville, 1999).

Pre-treatment methods for NIR spectra include: normalization; derivatives (usually first or second) (Rossa *et al.*, 2015); the multiplicative scatter correction (MSC); the standard normal variate (SNV); de-trending (DT); or a combination thereof (Blanco & Villarroya, 2002). Baker *et al.* (1994) showed that in silage, the effects of different particle sizes could be overcome by the use of the standard normal variate-detrending (SNV-D) transformation of Barnes *et al.* (1989).

Absorptions in a spectrum have two components: continuum and individual features. The continuum is the "background absorption" onto which other absorption features are superimposed which may be due to the wing of a larger absorption feature (Clark, 1999). The continuum removal technique was introduced in an effort to reduce background effects and normalise the spectral reflectance of water

absorption features (Clark & Roush, 1984). It can be viewed as an albedo normalization technique which is based on the computation of the continuum of a given spectrum (Kokaly & Clark, 1999).

The continuum is calculated by fitting a convex hull over the top of the spectrum, utilising straight line segments that connect local spectra maxima (Clark & Roush, 1984; Kokaly & Clark, 1999). The continuum is then removed by dividing the reflectance value for each point in the absorption features by the reflectance level of the continuum line (convex hull) at the corresponding wavelength (Mutanga & Ismail, 2010). According to Mevik & Wehrens (2007), the technique is an effective method to highlight absorption features of minerals.

2.2.4 Partial least-squares regression

Multivariate regression methods such as principal component regression (PCR) and partial least squares (PLS) enjoy large popularity in various fields in the natural sciences (Asner & Martin, 2008). PLS regression for mineral concentrations in plant tissue makes use of the correlation between the mineral and a biochemical such as pigment (de Aldana *et al.*, 1995). The technique has been designed to confront the situations where there are many, predictor variables, and relatively few samples. PLS regression relates two data matrices (X, Y) through a linear multivariate model, and is widely used in reflectance spectroscopy data analysis and chemometrics (Mevik & Wehrens, 2007).

Principal component regression (PCR) combines principal component analysis (PCA) and multiple linear regression (MLR) (Wold *et al.*, 2001). The independent variables are first decomposed into orthogonal principal components using the nonlinear iterative partial least squares algorithm and full cross validation of the calibration set. The maximum number of principal components is then defined according to the minimum value of the root-mean-square error of the cross validation (Mevik & Wehrens, 2007). In the final step, the chosen principal components are used to calibrate the MLR models. The advantage of PCR over the normal MLR is that the principal components are uncorrelated and the noise is filtered out (Nomngongo *et al.*, 2017).

PLS regression developed from multiple linear regression (MLR). The method decomposes both predictor and response variables into new components (Wold *et al.*, 2001). It is used to construct predictive models when there are many predictor variables that are possibly correlated. It is ideal for analysing data with multiple, collinear, variables and simultaneously model several response variables (Mevik & Wehrens, 2007). The PLS approach is advantageous because it uses the continuous spectrum as a single measurement rather than a band-by-band analysis (Asner & Martin, 2008).

With PLS regression, the calibration set is divided into several groups for the cross-validation. Cross-validation is essential in order to select the optimal number of factors and to avoid over fitting. Each group is then validated using a calibration developed on the other samples. Validation errors generated are combined into a root mean square error of cross-validation RMSECV. It has been reported that the RMSECV is the best single estimate for the prediction capability of the equation and that this statistic is similar to the average root mean square error of prediction (RMSEP) from 10 randomly-chosen prediction sets (González-Martín *et al.*, 2015).

The coefficient of determination (R^2) establishes a correlation between the analytical data obtained in the laboratory and those predicted by the calibration equations for each of the components analysed (Rossa *et al.*, 2015). A high R^2 value (e.g. 0.75 to 1.0) indicates a strong correlation, while with some calibrations an R^2 value less than 0.75 may be useful for monitoring purposes (Garcia Sanchez *et al.*, 2017). Following calibration the cross validation error called the root mean square error of cross validation (RMSECV) or root mean square error of prediction (RMSEP) is obtained, which should be taken into account most closely when evaluating the calibration. It is calculated based on the number of samples in the set and the differences between the estimated values and those obtained by standard methods of analysis. The following formula is used:

$$\text{RMSEP} = \sqrt{\frac{1}{N} \cdot \sum (x_i - y_i)^2} \quad (\text{Equation 2})$$

where $(x_i - y_i)$ is the difference between the measured value by laboratory methods and the predicted value by PLS model, and N is the total number of samples in the test (Naes *et al.*, 2002).

In “leave one out” cross validation (LOOCV), the response variable(s) are estimated from the fitted model based on the remaining samples. For example, if there are 20 samples, the response variable of each sample will be predicted by fitted model from the 19 remaining samples iteratively to determine the performance of the model. The merit of the cross-validation is the capability to detect outliers and provide unbiased assessment of the prediction error (Mevik & Wehrens, 2007).

2.3 Monitoring bio-resources using remote sensing

Remote sensing has been a valuable source of information over the course of the past few decades in mapping and monitoring forests. It provides a cost-effective tool to help forest managers better understand forest characteristics, such as forest area, locations, and species, even down to the level of characterizing individual trees (Adam *et al.*, 2012). The technology provides a synoptic view of the canopy and an opportunity to evaluate plant communities using light reflectance (Hatfield *et al.*, 2008). Understanding the spatial distribution of environmental parameters can allow for changes in management practices in human dominated ecosystems.

New space-borne sensors such as RapidEye, WorldView-2 and Sentinel-2 designed with red-edge bands provide an opportunity for assessing leaf nutrients at regional scale (Ramoelo *et al.*, 2015). The red-edge based vegetation indices of Sentinel-2 provide an opportunity for better estimates of leaf biochemicals and biomass (Mutanga *et al.*, 2012). The lack of red-edge bands in older satellite multispectral sensors hindered the production of landscape maps of leaf N and biomass. Sentinel-2, with spectral bands comparable to those of RapidEye and WorldView-2 was recently launched by the European Space Agency (Ramoelo *et al.*, 2015).

Earlier work by Adam *et al.* (2012) demonstrated that the new red-edge band in the RapidEye sensor has the potential for classifying tree species in semi-arid environments when integrated with other standard bands. The recent availability of time series of Sentinel-2 imagery at fine spatial resolution (10m) and high temporal frequency (every 5 days with S-2A and S-2B) represents a significant “scale-step” in the use of Earth Observation (EO) data for the monitoring of forest resources (Simonetti *et al.*, 2017).

The Sentinel-2 MultiSpectral Instrument (MSI) is designed to measure the earth-reflected radiance in 13 spectral bands, spanning from the Visible and Near Infrared (VNIR) to the Short Wave Infrared (SWIR). These include: four 4 bands at 10m special resolution namely: blue (490nm), green (560nm), red (665nm) and near infrared (842nm); six (6) bands at 20m special resolution which include four 4 narrow bands for vegetation characterisation (705nm, 740nm, 783nm and 865nm) and 2 larger SWIR bands (1610nm and 2190nm) for applications such as snow / ice / cloud detection or vegetation moisture stress assessment (ESA, 2015).

Sentinel-2A data has spectral bands that are much similar to Landsat 8 (except for the thermal bands of the Landsat 8 Thermal Infrared Sensor); and together they provide a core capacity on which a viable set of globally consistent services in the forestry domain could be based (Adan, 2017). This sets the stage for a number of innovative and challenging applications, and to the redesign of monitoring systems for more accurate monitoring of forest degradation (Simonetti *et al.*, 2017).

The extraction of leaf N and biomass using remote sensing has been limited to the use of statistical analysis, the most common techniques being simple linear and stepwise multiple regression (Ramoelo *et al.*, 2012). The latter suffers from overfitting and multi-collinearity (Curran, 1989; Kokaly *et al.*, 2009). Machine learning techniques such as random forest (RF) and artificial neural network (ANN) were found to be robust, and circumvents overfitting and multi-collinearity problem (Skidmore *et al.*, 2010; Mutanga *et al.*, 2012). RF and ANN were used more for classification than for regression analysis (Adam *et al.*, 2012). In regression, RF was found to be more robust than other parametric regression techniques (Mutanga *et al.*, 2012).

CHAPTER 3

MATERIALS AND METHODS

3.1 General description of the study area

The study was conducted at the Wits Rural Facility (24 °33'S; 31 °1'E), a research and development station (Figure 3.1) situated in the central Lowveld savanna in Limpopo Province, South Africa. The soil type is predominantly shallow sandy lithosols. The vegetation is typically lowland savannah woodlands, dominated by members of the Combretaceae (particularly *Combretum collinum*, *Combretum hereroense*, *Combretum apiculatum* and *Terminalia sericea*); and Mimosoideae (such as *Vachellia gerrardii*, *V. nilotica*, *V. swazica* and *Dichrostachys cinerea*). The mean annual rainfall is 665±123mm, received mostly during the summer months between October and April. The mean annual temperature is approximately 22°C and frost is rare (Shackleton, 2001).

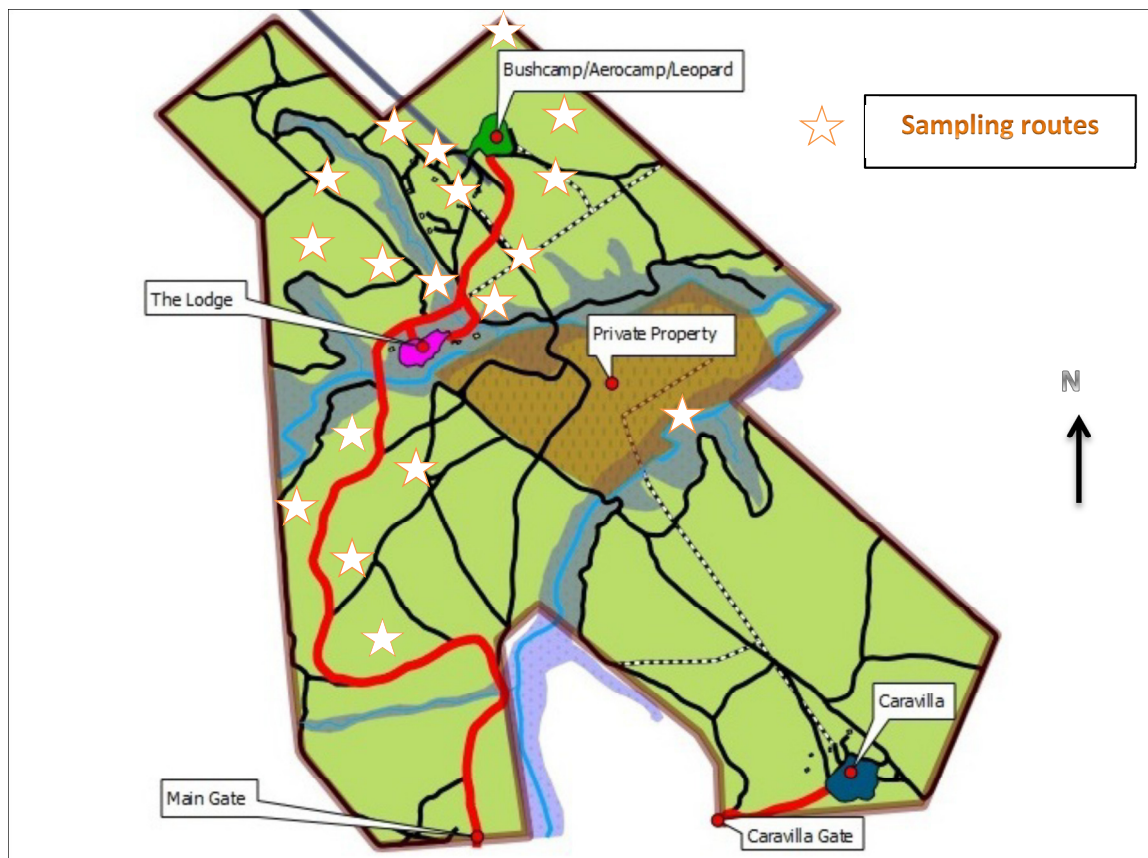


Figure 3.1: The study area

3.2 Data collection

3.2.1 Plant species investigated in the study

The dataset comprised of nine plant species mainly from the Fabaceae (or Leguminosae) family that are also common browse species for livestock and game in the Lowveld (Chepape *et al.*, 2011). A preliminary visit to the study area was conducted to identify tree species of interest, which were validated by ecologists at the Wits Rural Facility and a botanist at the University of Pretoria. The species comprised of three (3) broad-leaved leguminous trees, three (3) fine-leaved leguminous trees and three (3) non-legume species, hereafter referred to as “groups” or “type” (Table 3.1).

Table 3.1: Tree species investigated in the study

Species	Family	Type	Foliage
<i>Bauhinia galpinii</i>	Fabaceae	Broad-leaved, leguminous	Deciduous
<i>Philenoptera violacea</i>	Fabaceae	Broad-leaved, leguminous	Deciduous
<i>Schotia brachypetala</i>	Fabaceae	Broad-leaved, leguminous	Evergreen
<i>Peltophorum africanum</i>	Fabaceae	Narrow-leaved, leguminous	Evergreen
<i>Dichrostachys cinerea</i>	Fabaceae	Narrow-leaved, leguminous	Evergreen
<i>Vachellia gerrardii</i>	Fabaceae	Narrow-leaved, leguminous	Evergreen
<i>Combretum apiculatum</i>	Combretaceae	Broad-leaved, non-leguminous	Deciduous
<i>Terminalia sericea</i>	Combretaceae	Broad-leaved, non-leguminous	Deciduous
<i>Euclea natalensis</i>	Ebenaceae	Broad-leaved, non-leguminous	Evergreen

(Van Wyk & Van Wyk, 2013; <http://pza.sanbi.org/> [accessed: 22 September, 2016]).

3.2.2 Leaf sampling and handling

Leaf samples were collected on 22 August 2016 (dry season) and on 13 January 2017 (wet season), hereafter referred to as “winter” and “summer” respectively. About 100g of leaf material was collected from each tree; wherein five trees were

targeted per species. Leaf samples were, as much as possible, collected from different positions of the crown (top, middle, bottom, south, east, etc.). Samples were put in paper bags and labelled. Sampling points (waypoints) were recorded using Garmin *eTrex30*, global positioning system (GPS) instrument. In the laboratory, the samples were oven-dried at 60°C for 48 hours, and then ground and sieved through a 1mm screen (Campbell & Plank, 1997). The samples were placed in polythene bottles and stored in a cabinet, pending chemical and spectral analyses (Lukhele & van Ryssen, 2003).

3.3 Chemical analysis

3.3.1 Wet chemistry

Foliar nitrogen was determined following the Dumas dry oxidation combustion method (Dumas, 1831). The finely milled sample was used directly on a Carlo Erba NA 1500 Carbon/Nitrogen/Sulphur Analyzer, using approximately 10mg sample, weighed into a tin foil container for each determination. The instrument uses gas chromatography (GC) to separate the gases and yield nitrogen in the form of nitrogen gas (N₂) (Jimenez & Ladha, 1993).

The sample is ignited at high temperature (about 980°C) in oxygen on a chrome oxide catalyst to produce carbon dioxide, nitrogen gas and oxides of nitrogen. The gases produced pass through silvered cobalt oxide (to remove oxides of sulphur and halogens) and a column of copper (540°C), which reduces the oxides of nitrogen (NO_x) to (N₂) and also removes excess O₂. After removal of water vapour by a trap of anhydrous magnesium perchlorate, N₂ and CO₂ gases are finally separated by gas chromatography (GC) using a helium carrier gas and detected by a thermal conductivity detector (Jimenez & Ladha, 1993).

The instrument is calibrated against a certified standard of a pure organic compound of known composition. The compound chosen for our calibration standard is phenylalanine, an amino acid, which contains 8.48% N and 65.4% C. The software - PeakNet was used, which allows non-linear calibrations of peak height or peak area while smoothing out of high frequency noise and flexibility in peak integration (<http://www.dionex-france.com/library/manuals/software/34941-10.pdf> [accessed: 20 April, 2017]).

Potassium, calcium and phosphorus were determined by optical emission spectroscopy using the Spectro Genesis inductively coupled plasma optical emission spectrometer (ICP-OES) (Figure 3.2). Approximately 0.3g of each sample was digested using the Kjeldahl wet oxidation process (Kovacs *et al.*, 1996). This method uses nitric acid and hydrogen peroxide to digest plant samples; the elements are then determined by optical emission spectroscopy with inductively coupled plasma excitation.

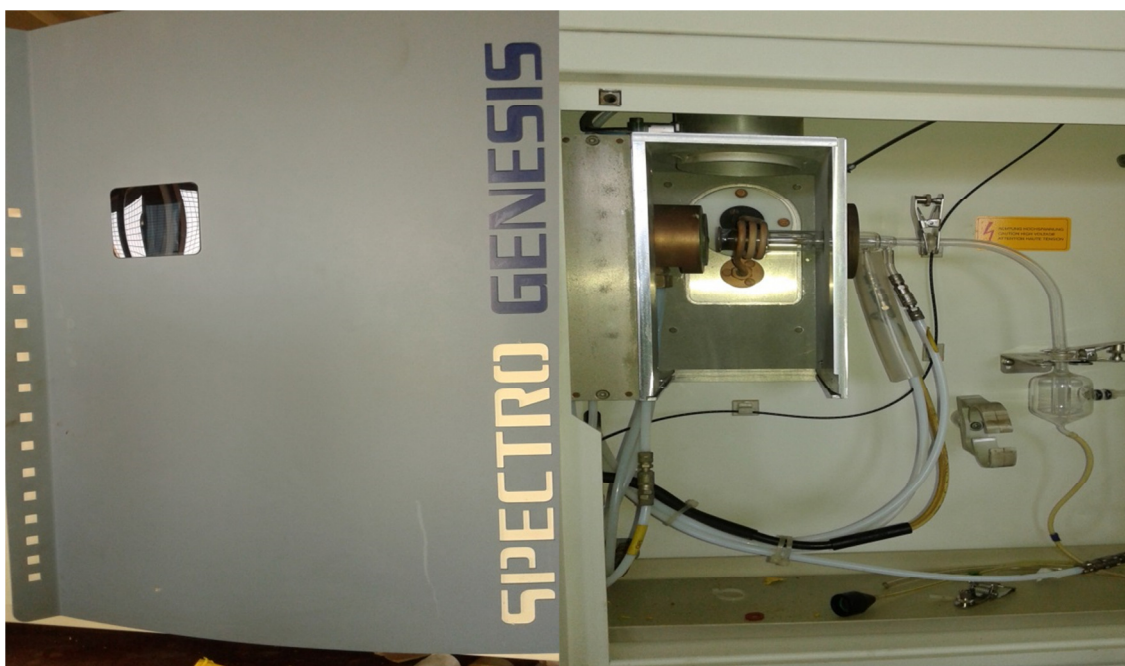


Figure 3.2: The Spectro Genesis ICP-OES: SPECTRO Analytical Instruments (<https://www.spectro.com/products/icp-oes-aes-spectrometers/genesis-icp-analysis> [accessed: 06 April 2017]).

3.3.2 Statistical data analysis

Descriptive statistics on foliar nutrient content were computed using ANalysis Of VAriance (ANOVA) programmed into the R statistical software. Boxplots were generated, followed by Q-Q plots and the Shapiro-Wilk test to test normality of data. Two-way ANOVA was used to test differences in nitrogen and the Tukey's post hoc tests were performed to obtain all pairwise comparisons (Mangiafico, 2015). The non-parametric Kruskal-Wallis test was computed to check statistical differences for phosphorus, potassium and calcium, followed by pairwise comparisons using Wilcoxon rank sum test (<http://www.sthda.com/english/wiki/kruskal-wallis-test-in-r> [accessed: 27 July 2018]).

3.4 Spectral measurements

3.4.1 Recording of spectra

The spectral reflectance of dried and ground foliage was measured using FieldSpec[®] 3 spectroradiometer (Analytical Spectral Devices Inc., Boulder, Colorado, USA) with a fibre optic probe (Figure 3.3). The instrument has a sampling interval of 1 nm for the 350 nm to 2500 nm spectral region with a total number of 2151 data points per spectrum (Awiti *et al.*, 2007).



Figure 3.3: Using FieldSpec[®] 3 spectroradiometer to collect spectra of dried and ground leaf samples.

The radiance spectra were normalized against a 99% white reference to produce relative reflectance spectra for each measurement. Ground leaf samples were arranged to fill in the field of view of the contact probe. Samples were placed on a spectrally black surface to minimize the background spectral noise. Five spectral measurements were taken on different parts of the sample and were averaged to obtain representative spectra. Figure 3.4 shows the spectra collected from the dried and ground leaf samples.

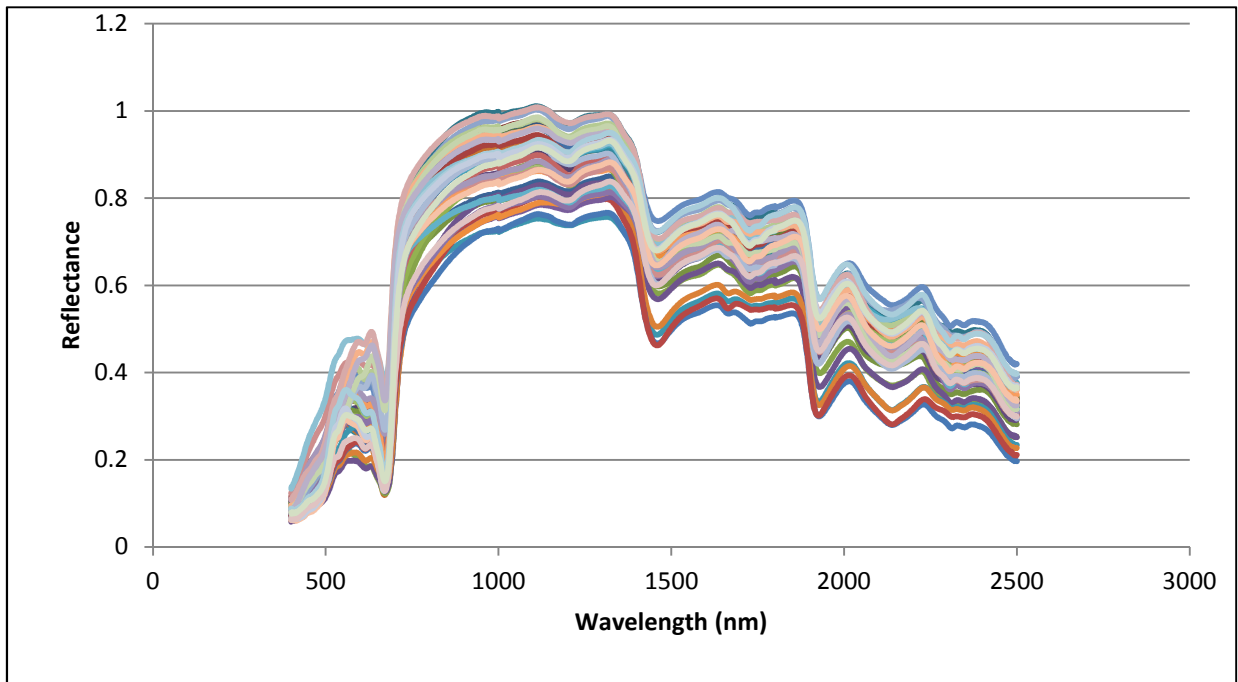


Figure 3.4: Reflectance spectra of dried and ground leaf samples.

3.4.2 Data pre-treatment

Due to some noise in the spectra between 350-399 nm, this interval was excluded from the analyses (Awiti *et al.*, 2007). The raw data were subjected to continuum removal using the `prospectr` package of the R software (Stevens & Ramirez-Lopez, 2015). The treatment is necessary to allow minimization of the scattering effect. Curran *et al.* (2001) asserts that estimating foliar biochemical concentration was more accurate when continuum-removed and band-normalised spectra were used rather than other pre-treatment methods (e.g. the first derivative of spectra - FDS).

Removing the continuum standardises isolated absorption features for comparison purposes (Clark & Roush, 1984; Clark, 1999). The continuum-removed spectra are shown in Appendix A04. The continuum-removed spectra were then used to develop a model through the method of chemometrics using R statistical programming language (Garcia & Filmoser, 2017).

3.4.3 Constructing the models

Partial least squares regression (PLSR) was used to model relationship between the chemical and spectral measurements of samples. Forty-five (45) samples were used

to build the models for winter and summer seasons respectively. For all samples, the chemical data determined by the reference methods were added to corresponding NIR spectra. The following packages programmed into the R software were used analyse spectral data: `pls`, `chemometrics`, `ChemometricsWithR`, `prospectr` and `inspectr`. Partial-least squares regression was used to develop models.

Validation of the model was done using 'leave one out' cross validation (LOOCV) because of the small sample sizes (Mevik & Wehrens, 2007). The root mean square error (RMSE) and the coefficient of determination (R^2) were considered as statistic measures of precision and accuracy (Garcia & Filmoser, 2017). A typical way of fitting a PLSR model was as follows:

```
NleafChem <- pls(Nitrogen ~ crALLSPECTRA, ncomp = 10, data = leafTrain,
validation = "LOO").
```

This fits a model with 10 components, and includes leave-one-out (LOO) cross-validated predictions. We then get an overview of the fit and validation results with the summary method as follows: `summary(NleafChem)`.

Different regions of the spectra were explored to assess performance of the model. The script for fitting the PLSR model(s) is shown in Appendix A03. The optimum number of components used each model was determined using the local minimum of RMSE of the model developed using 10 components (Mevik & Wehrens, 2007). Statistical accuracy and precision metrics such as the coefficient of determination (R^2), root mean square error (RMSE) and p values were used (Ramoelo & Cho, 2018).

3.5 Satellite data

3.5.1 Image acquisition and pre-processing

Sentinel-2 images covering the study area were downloaded from <http://glovis.usgs.gov/> [retrieved: 28 June, 2017], corresponding to the dates of leaf sampling (see Table 3.2). Atmospheric correction was performed using Image Correction Plugin for Snap Toolbox Software (iCOR), previously known as OPERA (Sterckx *et al.*, 2015). The iCOR plugin for the Sentinel-2 Snap Toolbox is scene and sensor generic atmospheric correction algorithm that can handle land and water

targets and is adaptable with minimal efforts to other hyper- or multi-spectral sensors (VITO, 2018).

All input data required for the atmospheric correction are derived from the image itself or delivered through pre-calculated look-up-tables (VITO, 2017). Through the use of a single atmospheric correction implementation, discontinuities in the reflectance between land and the highly dynamic water areas are reduced (Ramoelo & Cho, 2018).

Table 3.2: List of Sentinel-2 images used in the study

Product ID	Date
S2A_OPER_MSI_L1C_TL_MPS__20160812T112604_A005952_T36JUT_N02.04	2016-08-12
S2A_MSIL1C_20170119T074231_N0204_R092_T36JUT_20170119T075734	2017-01-19

(<https://glovis.usgs.gov/app> [accessed: 28 June 2017]).

3.5.2 Point extraction

In order to extract multispectral data, the images were pre-processed in which digital numbers were converted to surface reflectance. Spectral reflectance of georeferenced ground points (GPS coordinates of the sampled trees) were extracted from the image. Trees with larger canopies were purposively sampled (Ramoelo *et al.*, 2014); and their chemical, hyperspectral (laboratory spectroscopy), multispectral spectral (Sentinel-2) as well as indices were used in generating the model. Due to the fewer number of trees with large canopies, the dataset comprised of data recorded in both dry and wet seasons.

3.5.3 Spectral bands and indices

This study utilised eight Sentinel-2 MSI bands that are critical for characterisation of vegetation (Skidmore *et al.*, 2010; Ramoelo *et al.*, 2014; Ramoelo & Cho, 2018). The bands are centred at 490nm, 560nm, 665nm, 705nm, 740nm, 783nm, 842nm and 865nm (Table 3.3). The workflow of iCOR atmospheric correction algorithm is outlined in Figure 3.5.

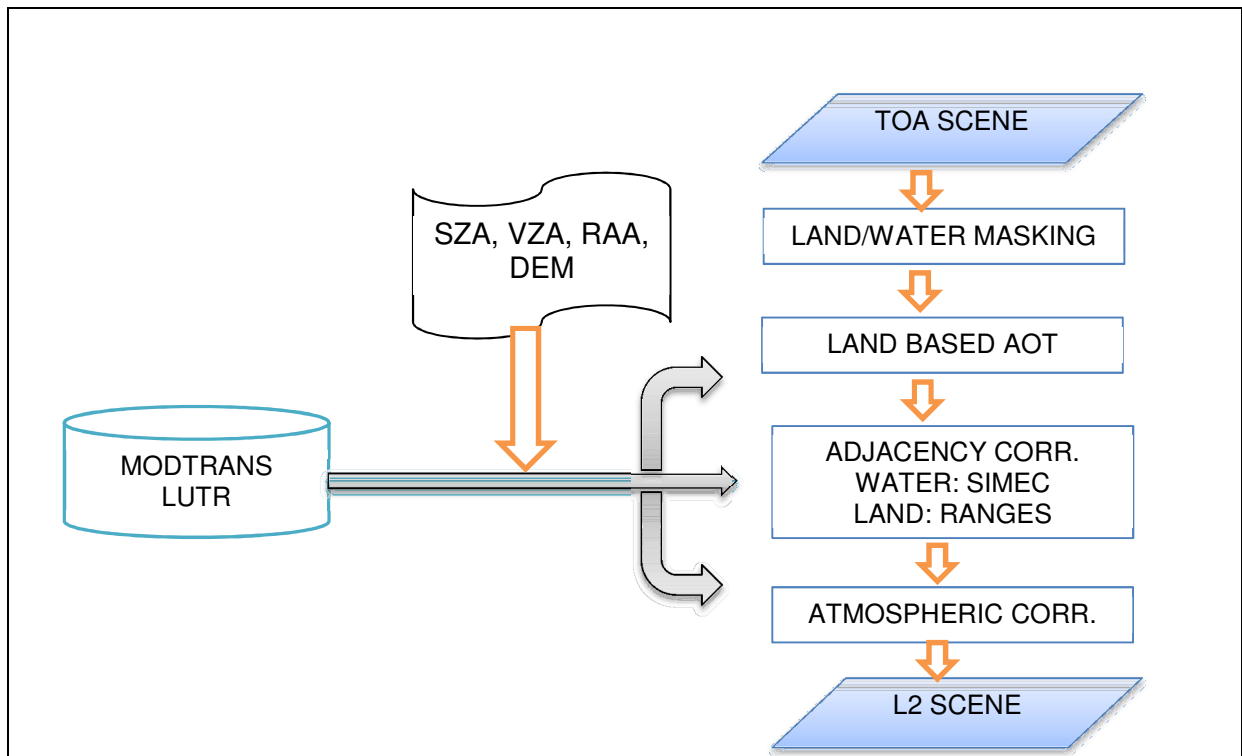


Figure 3.5: Workflow of iCOR atmospheric correction algorithm. (Adapted from: https://blog.vito.be/remotesensing/icor_available [accessed: 12 Jan 2018]).

LUT = Look-up-Table, SZA = Solar Zenith Angle, VZA = View Zenith Angle, RAA = Relative Azimuth Angle, DEM = Digital Elevation Model, TOA = Top-of-Atmosphere, AOT = Aerosol Optical Thickness, SIMEC = Similarity Environment Correction, L2 = Level 2 Atmospherically Corrected.

Conventional and red-edge based vegetation indices such as modified simple ratio mSR705 $(B06 - B01) / (B05 - B01)$; normalized difference vegetation index - NDVI $(B08 - B04) / (B08 + B04)$; red-edge normalized difference vegetation index - RE NDVI $(B08 - B06) / (B08 + B06)$; NDVI-Green $(B03 * (B08 - B04) / (B08 + B04))$, were computed (Ramoelo *et al.*, 2014; <http://www.sentinel-hub.com/eotaxonomy/indices> [accessed: 26 October, 2017]). A total of eighteen bands and vegetation indices were used as an input into the model (Appendix A04).

Table 3.3: Description of Sentinel-2 bands used in the study

Spectral Band	Central wavelength (nm)	Band width (nm)	Resolution
Band 2 – blue	490nm	65nm	10m
Band 3 – green	560nm	35nm	10m
Band 4 – red	665nm	30nm	10m
Band 5 – vegetation red edge	705nm	15nm	20m
Band 6 – vegetation red edge	740nm	15 nm	20m
Band 7 – vegetation red edge	783nm	20nm	20m
Band 8 – NIR	842nm	115nm	10m
Band 8b – vegetation red edge	865nm	20nm	20m

ESA (2015); retrieved from http://www.gdal.org/frmt_sentinel2.html [15 September, 2017]).

3.5.4 Model construction

Partial least squares regression (PLSR) was used to model relationships between leaf chemical composition and multispectral data (indices and bands) from Sentinel-2 (Ramoelo *et al.*, 2015). Models were computed using the pls package in R programming language and were validated using 'leave one out' cross validation (LOOCV) because of the small sample sizes (Mevik & Wehrens, 2007). The root mean square error (RMSE) and the coefficient of determination (R^2) were considered as statistic measures of precision and accuracy (Garcia & Filmoser, 2017).

CHAPTER 4

RESULTS AND DISCUSSION

4.1 Foliar nutrient variations

The summary of the results of chemical analysis is presented in Table 4.1. Nitrogen concentration ranged from 0.7 to 3.5; phosphorus from 0.06 to 0.23; potassium from 0.33 to 1.67; and calcium ranged from 0.30 to 3.27 g/100g.

Table 4.1: Summary of the chemical data recorded

	Nitrogen (g/100g)	Phosphorus (g/100g)	Potassium (g/100g)	Calcium (g/100g)
Highest value	3.540	0.2300	1.6700	3.270
Mean	2.063	0.1168	0.7674	1.207
Lowest value	0.710	0.0600	0.3300	0.300

Figure 4.1 shows the boxplots for the variation in nutrient levels among the species in both dry and wet seasons. The highest concentration of nitrogen was recorded in *Philenoptera violacea* specimen during summer; and the lowest was in *Terminalia sericea* plant during winter. Phosphorus concentration was highest in *Dichrostachys cinerea* during summer, while the lowest was recorded in *Terminalia sericea* during winter. Potassium was highest in *Philenoptera violacea* during summer, the lowest was in *Combretum apiculatum* during winter. The highest amount of calcium was recorded in *Dichrostachys cinerea* leaves during winter, while the lowest was recorded in *Combretum apiculatum* during summer.

The Shapiro–Wilk test showed that leaf nitrogen was normally distributed, while phosphorus, potassium and calcium were not. *Philenoptera violacea* contained significantly higher ($p < 0.05$) nitrogen than most species during the summer season, with the exception of *Vachellia gerrardii* and *Dichrostachys cinerea* (Table 4.2). In the same period, nitrogen concentration was significantly higher ($p < 0.05$) in both *Vachellia gerrardii* and *Dichrostachys cinerea* compared with *Terminalia sericea* and *Euclea natalensis*. The results show that during winter there was less variability in nitrogen (CV=40.5%) concentration among the species compared to summer (CV=51.4%), corroborating the findings by Ramoelo *et al.* (2014).

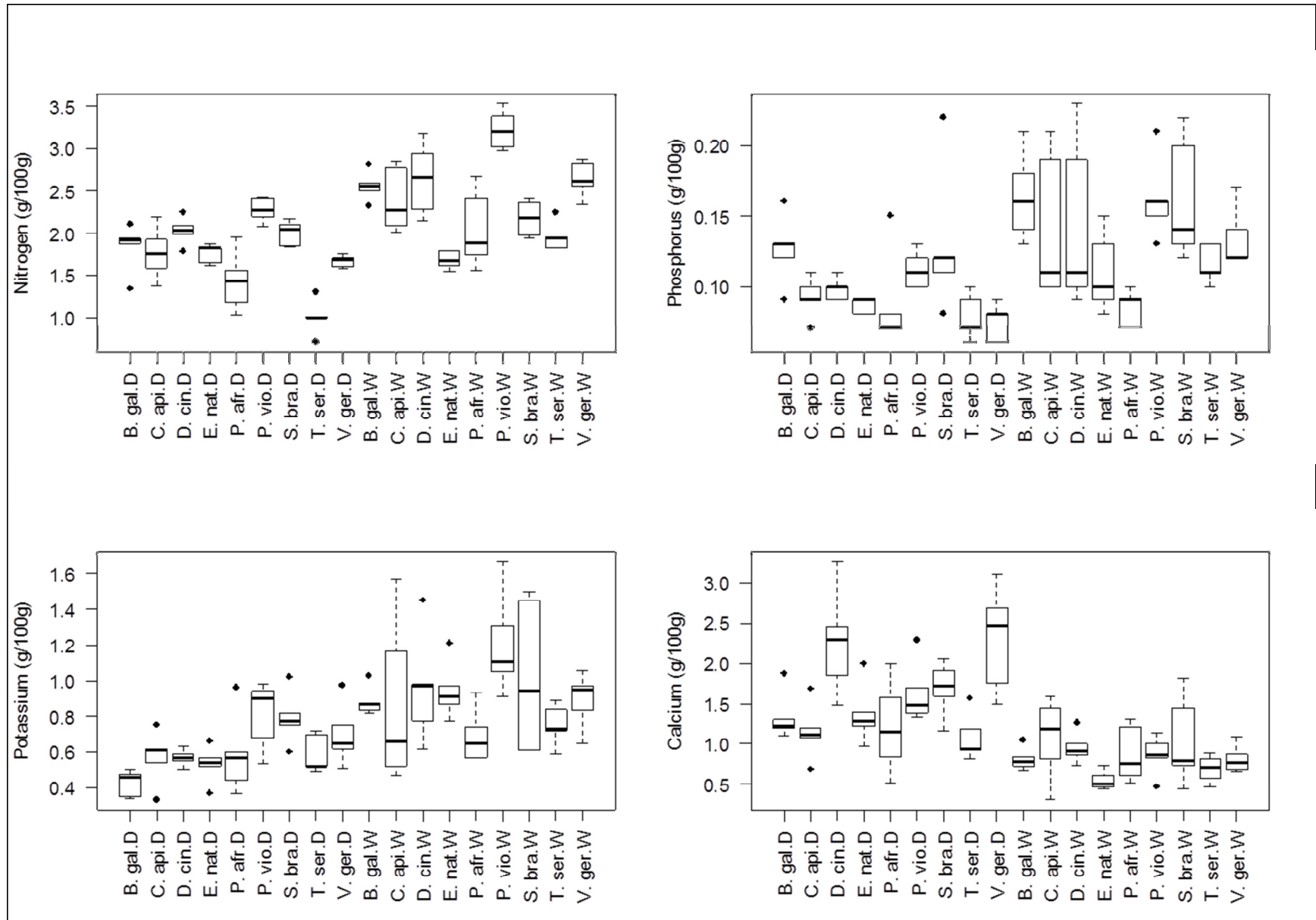


Figure 4.1: Foliar nutrient variation by species in the dry (D) and wet (W) seasons

Table 4.2: Pairwise comparisons for nitrogen in the wet season

	<i>B. gal</i>	<i>P. vio</i>	<i>S. bra</i>	<i>P. afr</i>	<i>D. cin</i>	<i>V. ger</i>	<i>C. api</i>	<i>T. ser</i>	<i>E. nat</i>
<i>B. gal</i>									
<i>P. vio</i>	*								
<i>S. bra</i>		**							
<i>P. afr</i>		**							
<i>D. cin</i>									
<i>V. ger</i>									
<i>C. api</i>		**							
<i>T. ser</i>	*	**			**	**			
<i>E. nat</i>		**			**	**	**		

Significance codes: '*' p < 0.05
 '***' p < 0.01

Table 4.3: Pairwise comparisons for nitrogen in the dry season

	<i>B. gal</i>	<i>P. vio</i>	<i>S. bra</i>	<i>P. afr</i>	<i>D. cin</i>	<i>V. ger</i>	<i>C. api</i>	<i>T. ser</i>	<i>E. nat</i>
<i>B. gal</i>									
<i>P. vio</i>									
<i>S. bra</i>									
<i>P. afr</i>		**							
<i>D. cin</i>									
<i>V. ger</i>		*							
<i>C. api</i>									
<i>T. ser</i>	**	**	**			*	**		
<i>E. nat</i>								**	

Significance codes: '*' p < 0.05
 '***' p < 0.01

Table 4.4: Pairwise comparisons for nitrogen in both wet and dry seasons

	<i>B. gal</i>	<i>P. vio</i>	<i>S. bra</i>	<i>P. afr</i>	<i>D. cin</i>	<i>V. ger</i>	<i>C. api</i>	<i>T. ser</i>	<i>E. nat</i>
<i>B. gal</i>									
<i>P. vio</i>	**								
<i>S. bra</i>		**							
<i>P. afr</i>	*	**							
<i>D. cin</i>				**					
<i>V. ger</i>		**							
<i>C. api</i>		**							
<i>T. ser</i>	**	**	**		**	**	**		
<i>E. nat</i>	*	**			**				

Significance codes: '*' p < 0.05
 '**' p < 0.01

The results of statistical tests for mean nitrogen in both dry and wet seasons are shown in Table 4.4. *Philenoptera violacea* contain significantly higher ($p < 0.05$) levels of nitrogen compared with other species except with *Dichrostachys cinerea*. The findings corroborate the work by du Toit (2003) in the Lowveld savanna, which showed foliar nitrogen concentrations in excess of 3% for *Philenoptera violacea*, well in excess of foliar nitrogen in neighbouring non-leguminous plants. The non-leguminous species *Terminalia sericea* contained significantly lower ($p < 0.05$) amounts of nitrogen than other species except *Peltophorum africanum* and *Euclea natalensis*.

The results of foliar nitrogen variation between groups are shown in Figure 4.2. The concentration was significantly lower ($p < 0.05$) in non-leguminous plants than in leguminous plants during summer, while the difference in nitrogen between the two legume groups was not significant. During winter, nitrogen concentration was significantly lower ($p < 0.05$) in non-legumes than in broad-leaved legumes while the difference in mean nitrogen between narrow-leaved legumes and non-legumes was not significant.

The nitrogen content was highest in two broad-leaved legumes *Philenoptera violacea* and *Bauhinia galpinii*. The results suggest that leguminous plants are better sources of nitrogen (crude protein) than non-leguminous plants, which is consistent with the findings by El hassan *et al.* (2000) and Lukhele & van Ryssen (2003). The studies reported that many non-leguminous species found in the bushveld region contained less than 17% crude protein, while most of the leguminous trees were found to be good protein sources.

Leaf nitrogen is related to protein (Clifton *et al.*, 1994). The high protein content in leguminous plants is due to the action of nitrogen-fixing bacteria that inhabit root nodules and have symbiosis relationship with leguminous plants (Hungria & Franco. 1993; Jacobs *et al.*, 2007). Nonetheless the potential of each legume species needs to be explored as there is significant variability among them. For instance *Peltophorum africanum*, a narrow-leaved leguminous plant was found to be among the lowest in nitrogen concentration.

In terms of phosphorus concentrations in the different groups, broad-leaved legumes contained significantly higher ($p < 0.05$) phosphorus than narrow-leaved legumes and non-leguminous species while the difference in between narrow-leaved legumes and the non-legumes was not significant (Figure 4.2). Differences in potassium and calcium between the three groups were not significant. Seasonal effect was significant ($p < 0.05$) for all the four nutrients. Concentrations of nitrogen, phosphorus and potassium were significantly higher ($p < 0.05$) in summer than in winter whereas calcium was significantly higher ($p < 0.05$) in winter than in summer (Figure 4.3).

Green plants generally have higher nutrient concentrations due to active metabolic activities (e.g. photosynthesis) portraying vigour and health of vegetation. During dry periods the nutrients are generally translocated from the leaves to the roots (Majeke *et al.*, 2008). The interaction effect between the “groups” and “seasons” was highly significant ($p < 0.05$) for nitrogen, owing to changes soil conditions (temperature and humidity) on the activity of nitrogen-fixing bacteria (Jacobs *et al.*, 2003).

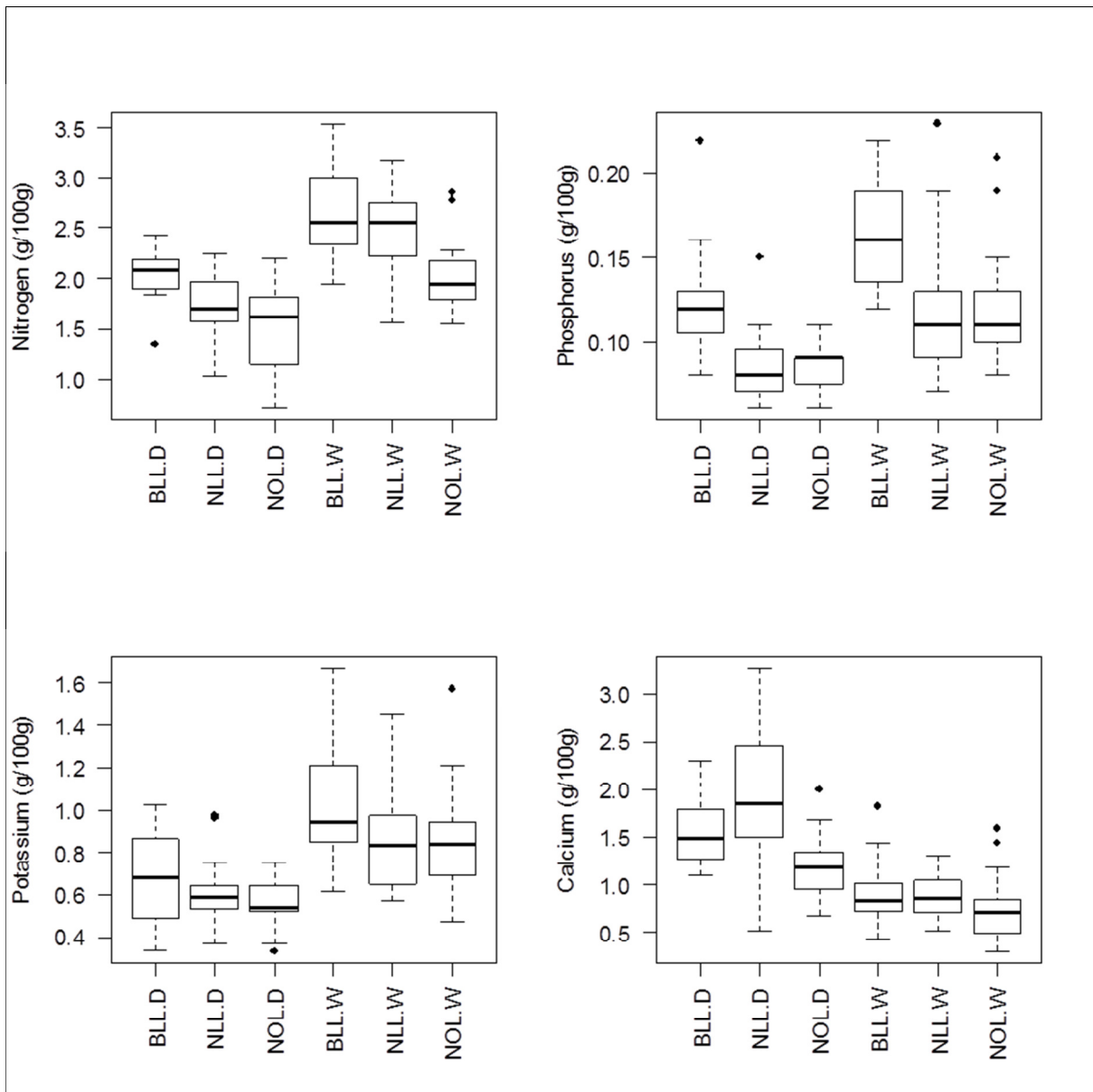


Figure 4.2: Foliar nutrient variation by group across the two seasons.

BLL.D = Broad leaved legume – dry season; NLL.D = Narrow leaved legume – dry season; NOL.D = Non-legume – dry season; BLL.W = Broad leaved legume – wet season; NLL.W = Narrow leaved legume – wet season; NOL.W = Non legume – wet season.

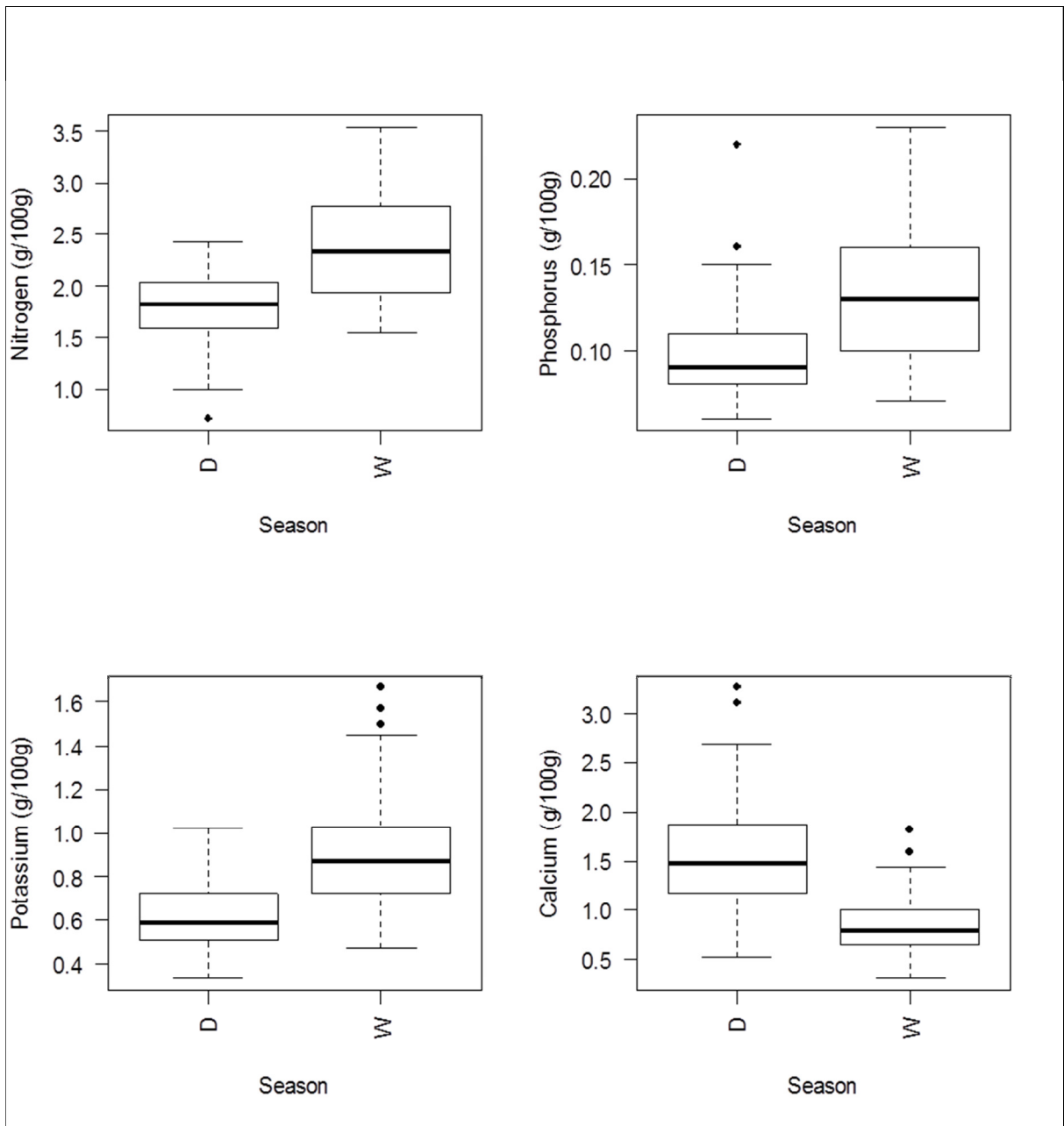


Figure 4.3: Foliar nutrient concentration between dry and wet seasons.

4.2. Nutrient concentration and spectral reflectance of dried leaf samples.

Table 4.5 is a summary of the performance of the statistical models developed to relate foliar nutrients measured by reference methods to hyperspectral data obtained from laboratory spectroscopy. Figures 4.4 to 4.7 show the graphical output of the models. There was a significant relationship between foliar nitrogen concentration and hyperspectral reflectance of ground leaf samples both during summer ($R^2=0.88$, $p<0.05$) and winter seasons ($R^2=0.76$, $p<0.05$). The higher coefficient of determination in summer can be attributed to the higher variability in nitrogen content observed in chemicals analysis and as highlighted under section 4.1 above. During summer plants have more vigour and high levels of photosynthetic pigments such as chlorophyll (Ramoelo *et al.*, 2014). This study corroborates what was achieved by Ramoelo *et al.* (2014).

The relationship between foliar phosphorus, potassium and calcium concentrations and hyperspectral reflectance of ground leaf samples was generally low. The lowest correlations were observed with potassium ($R^2=0.08$) during winter and calcium ($R^2=0.01$) during summer as shown in Table 4.5..

Table 4.5: Performance of the models on the estimation of foliar nutrients using spectrometer data

Season	<i>n</i>	Nutrient	RMSEP	R^2	<i>p</i> values < 0.05
Winter	45	N	0.21	0.76	Yes
		P	0.02	0.49	Yes
		K	0.17	0.08	Yes
		Ca	0.45	0.43	Yes
Summer	45	N	0.18	0.88	Yes
		P	0.03	0.41	Yes
		K	0.26	0.20	Yes
		Ca	0.32	0.01	Yes

The results demonstrate that spectral reflectance of dried and ground leaf samples coupled with partial least squares regression can be applied as a fast analytical technique to evaluate nitrogen content of species of interest to be integrated in small-scale agricultural systems as organic inputs. Work by Galvez-Sola *et al.* (2015)

on citrus leaves of lemon, mandarin, orange, and grapefruit found high accuracy regarding the estimation of nitrogen ($R^2 = 0.96$) and as well as acceptable estimates for potassium, iron and zinc. In yerba mate plants (*Ilex paraguariensis*), the prediction was good for phosphorus and copper but not for potassium and calcium (Rossa *et al.*, 2015). These data show that the leaf spectral response depends on the species studied, so for each species it is necessary to make the appropriate calibrations.

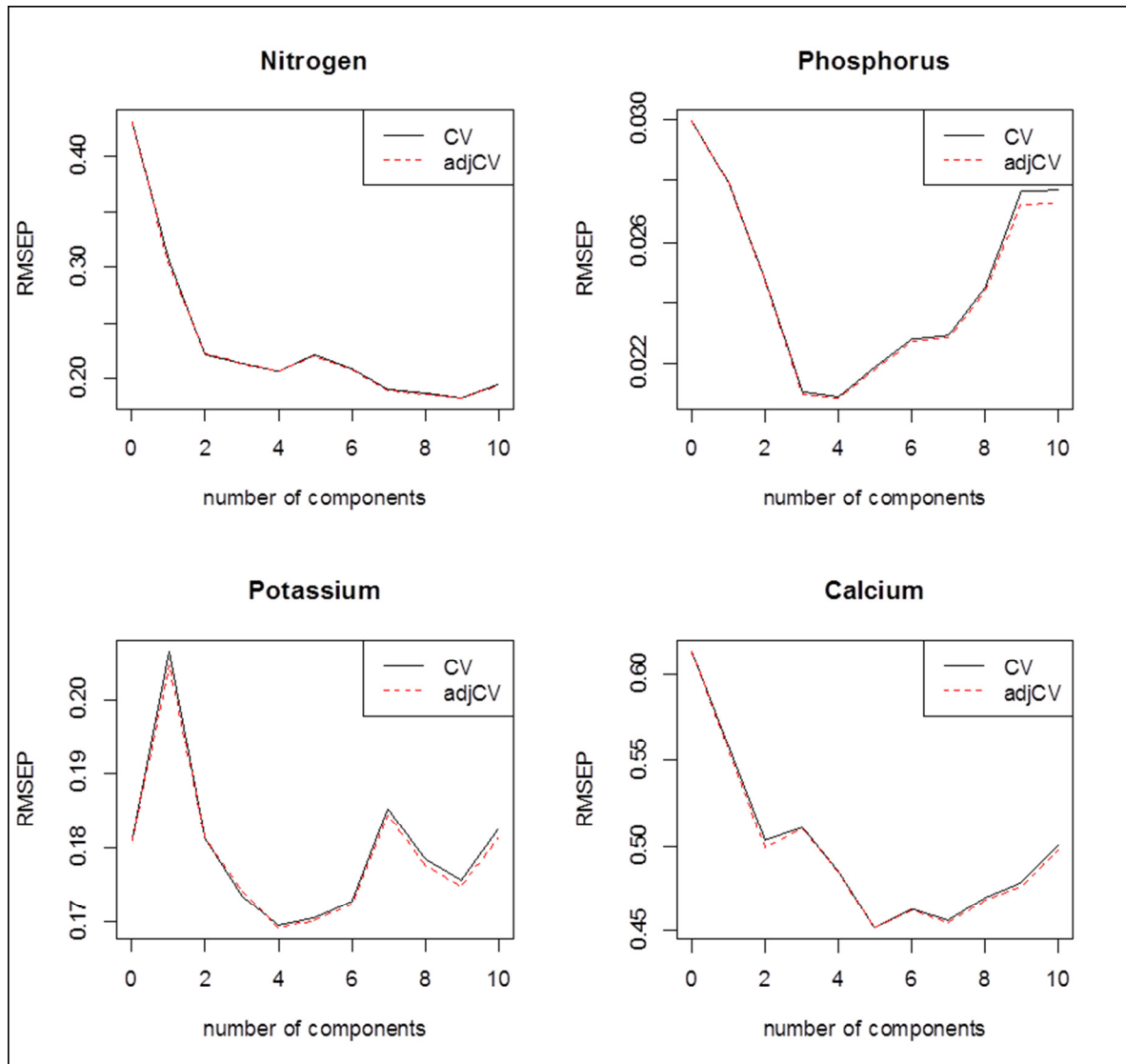


Figure 4.4: Cross-validated RMSEP curves for hyperspectral vs chemical: winter

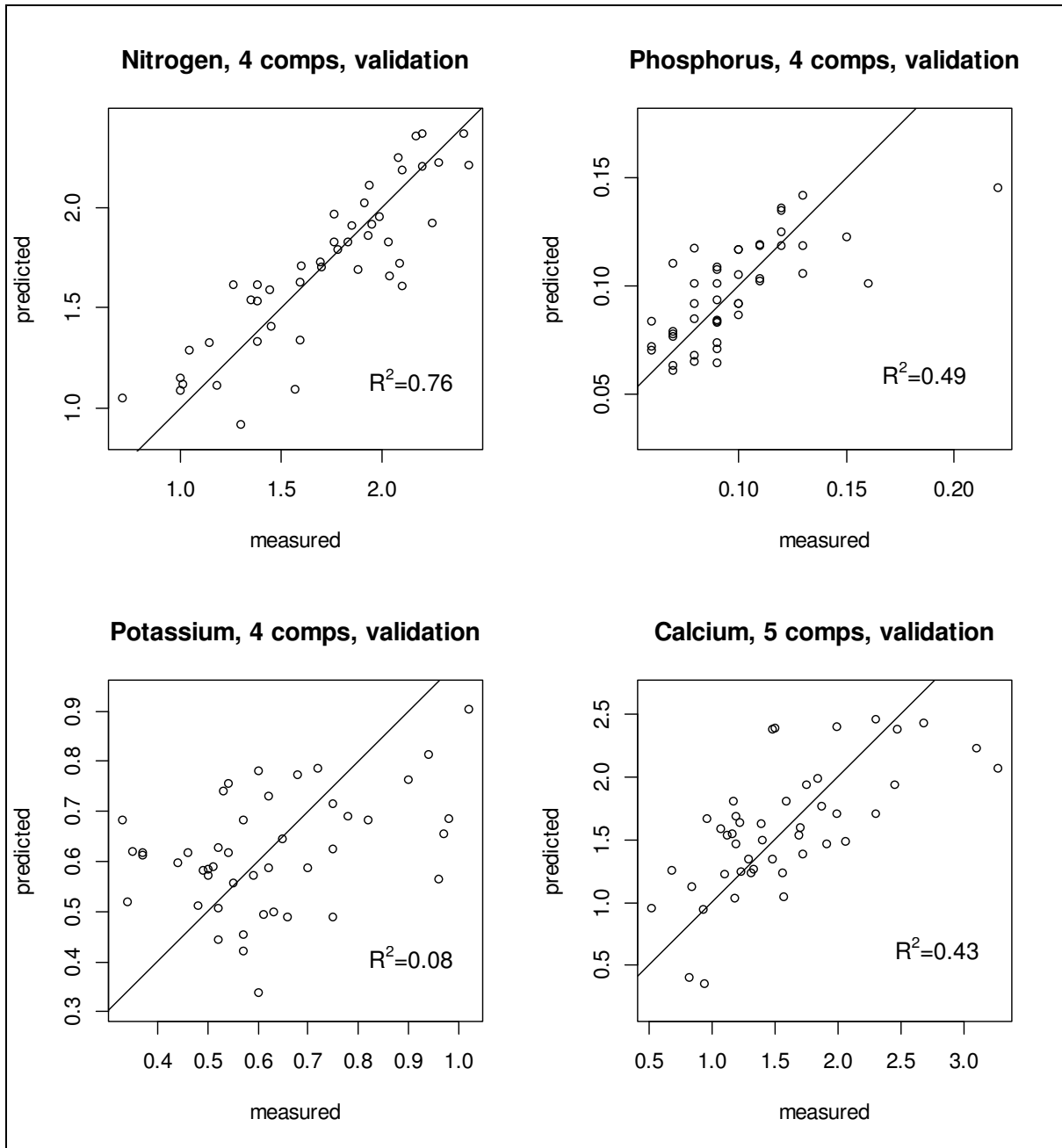


Figure 4.5: Cross-validated predictions for hyperspectral vs chemical: winter.

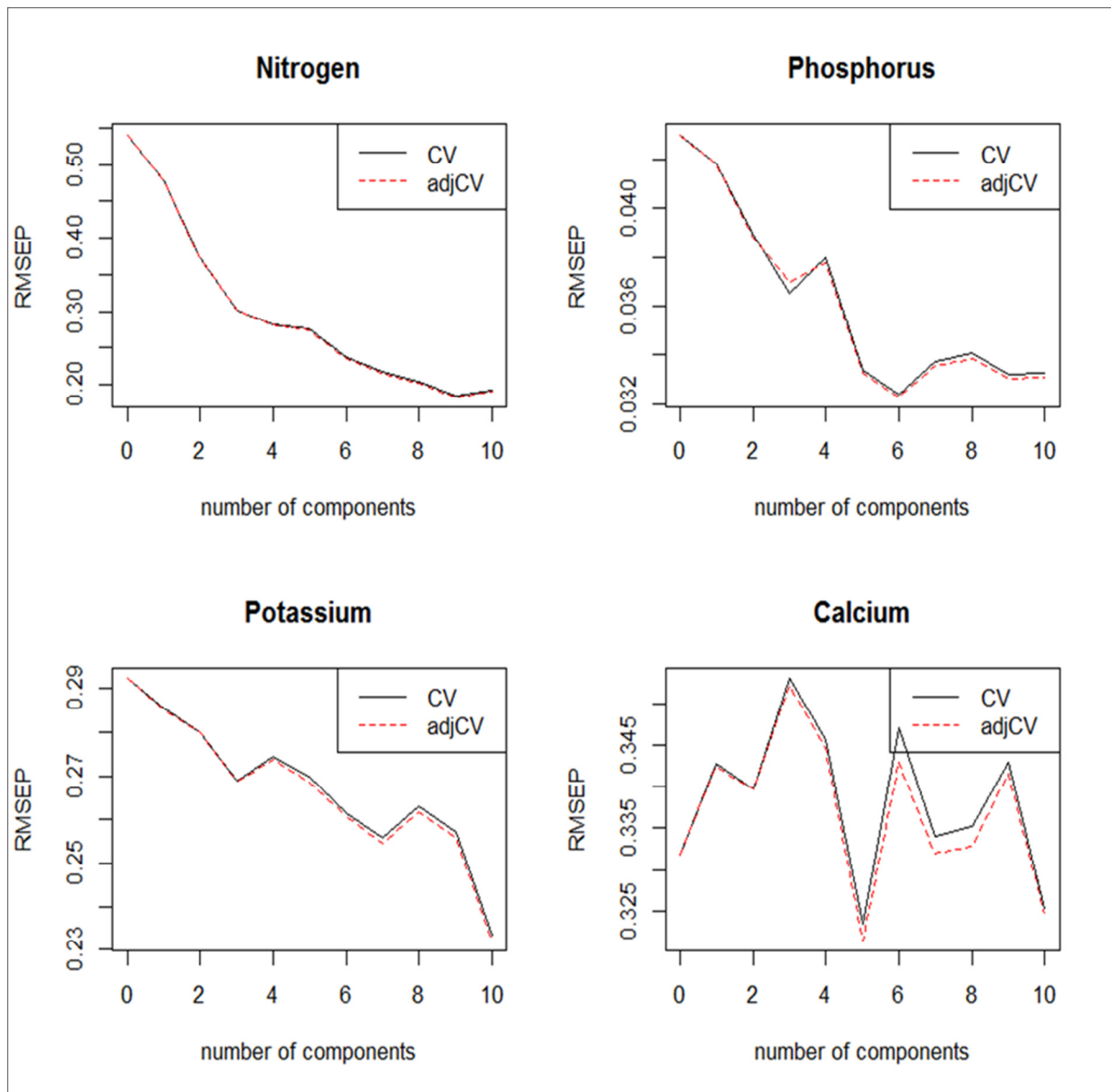


Figure 4.6: Cross-validated RMSEP curves for hyperspectral vs chemical: summer.

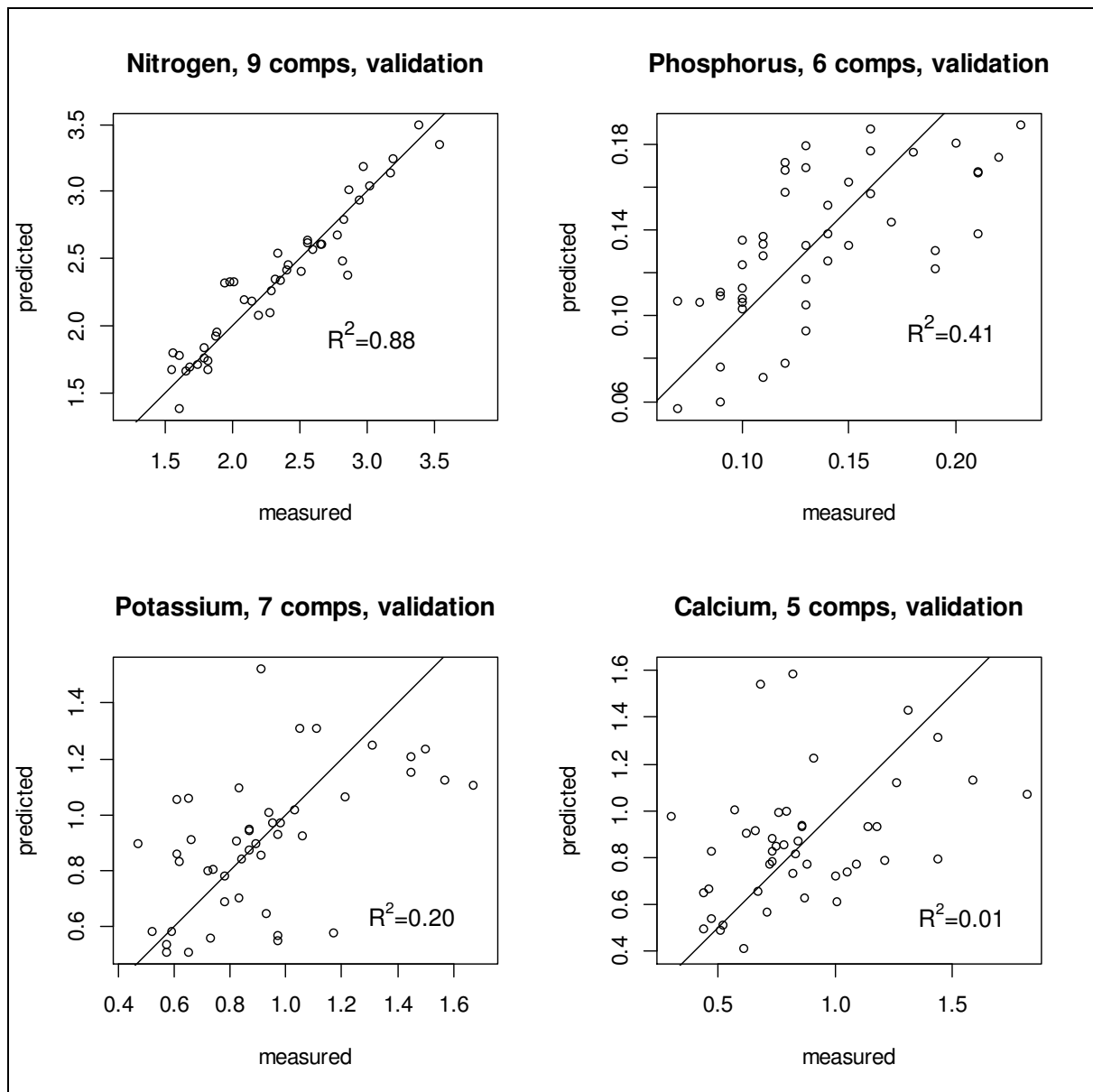


Figure 4.7: Cross-validated predictions for hyperspectral vs chemical: summer

4.3. Modelling nutrient concentrations using Sentinel-2 images

Table 4.6 shows the results of the models relating chemical composition to spectral reflectance of Sentinel-2. Graphical outputs of the model are shown in Figures 4.8 & 4.9.

Table 4.6: Performance of models for Sentinel-2 data

Season	<i>n</i>	Nutrient	RMSE	R ²	<i>p</i> values < 0.05
Combined dataset	19	N	0.22	0.44	Yes
	19	P	0.04	0.04	Yes
	19	K	0.27	0.23	Yes
	19	Ca	0.41	0.25	Yes

The prediction accuracy for nitrogen was higher compared to that of potassium, phosphorus and calcium. Phenology is a major factor of this outcome, given the fact that most of the vegetation indices particularly RE-based indices depend on the vegetation vigour and greenness. There was significant correlation ($R^2=0.44$, $p<0.05$) between foliar nitrogen composition and the predictions made by the PLS model using Sentinel-2 data. As outlined by Garcia Sanchez *et al.* (2017), R^2 values that are less than 0.75, although not acceptable, they may be useful for monitoring purposes. It is assumed that the models could perform better with improved sampling techniques and the use of larger datasets.

Estimation of leaf nutrients using Sentinel-2 data yielded low performances, which could be attributed to low dimensionality of the data. PLS regression performs better when exploratory variables are in their tens to hundreds and even thousands, especially using spectroscopy or hyperspectral data (Wold *et al.*, 2001). Ramoelo *et al.* (2015) demonstrated that machine learning techniques, such as random forest improved the estimation of leaf nitrogen by 49%, and are quite robust if well parameterized. Nonetheless various studies demonstrated that Sentinel-2 can be used to accurately estimate leaf N concentrations (Ramoelo & Cho, 2018).

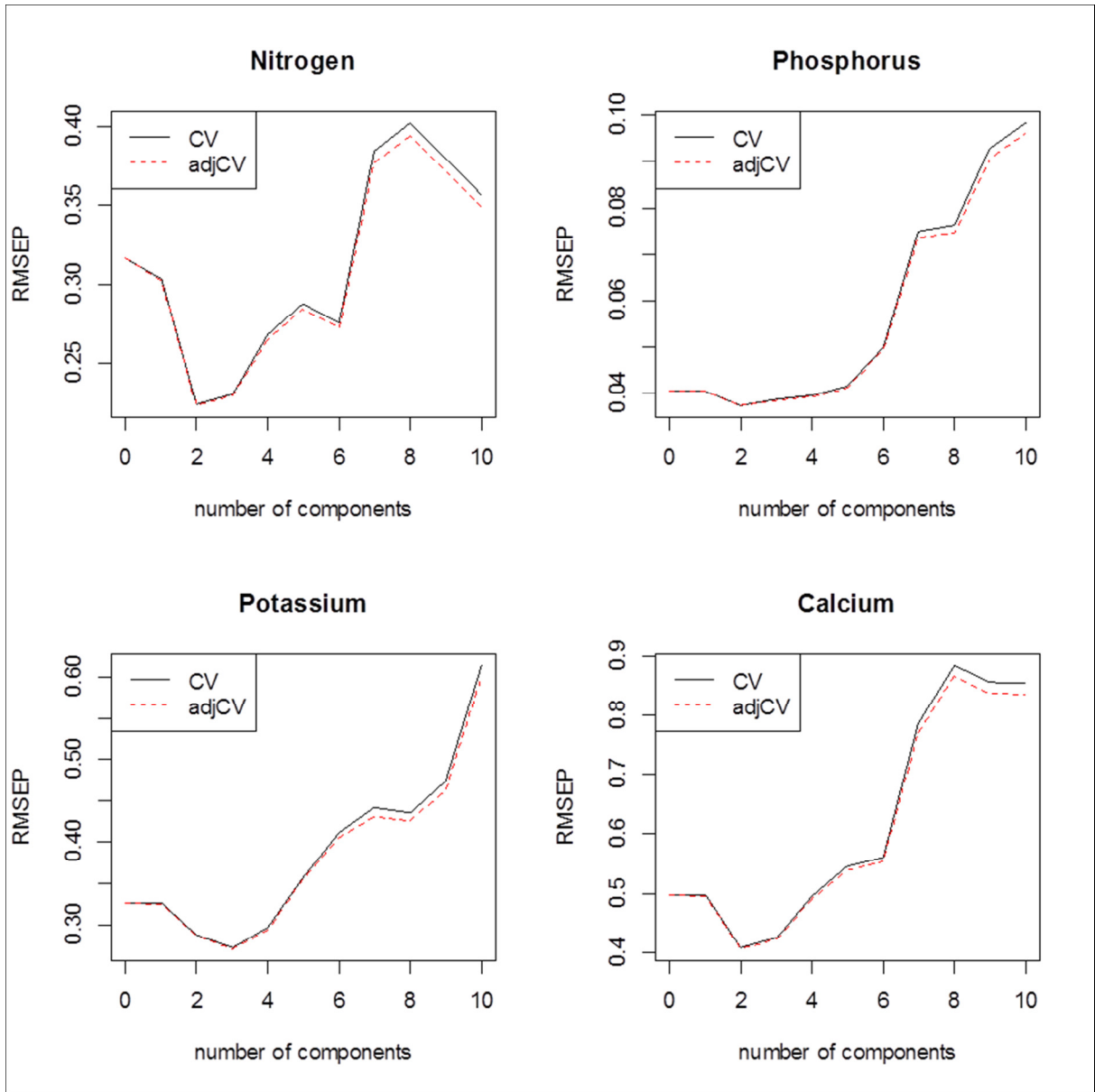


Figure 4.8: Cross-validated RMSEP curves for Sentinel-2 data

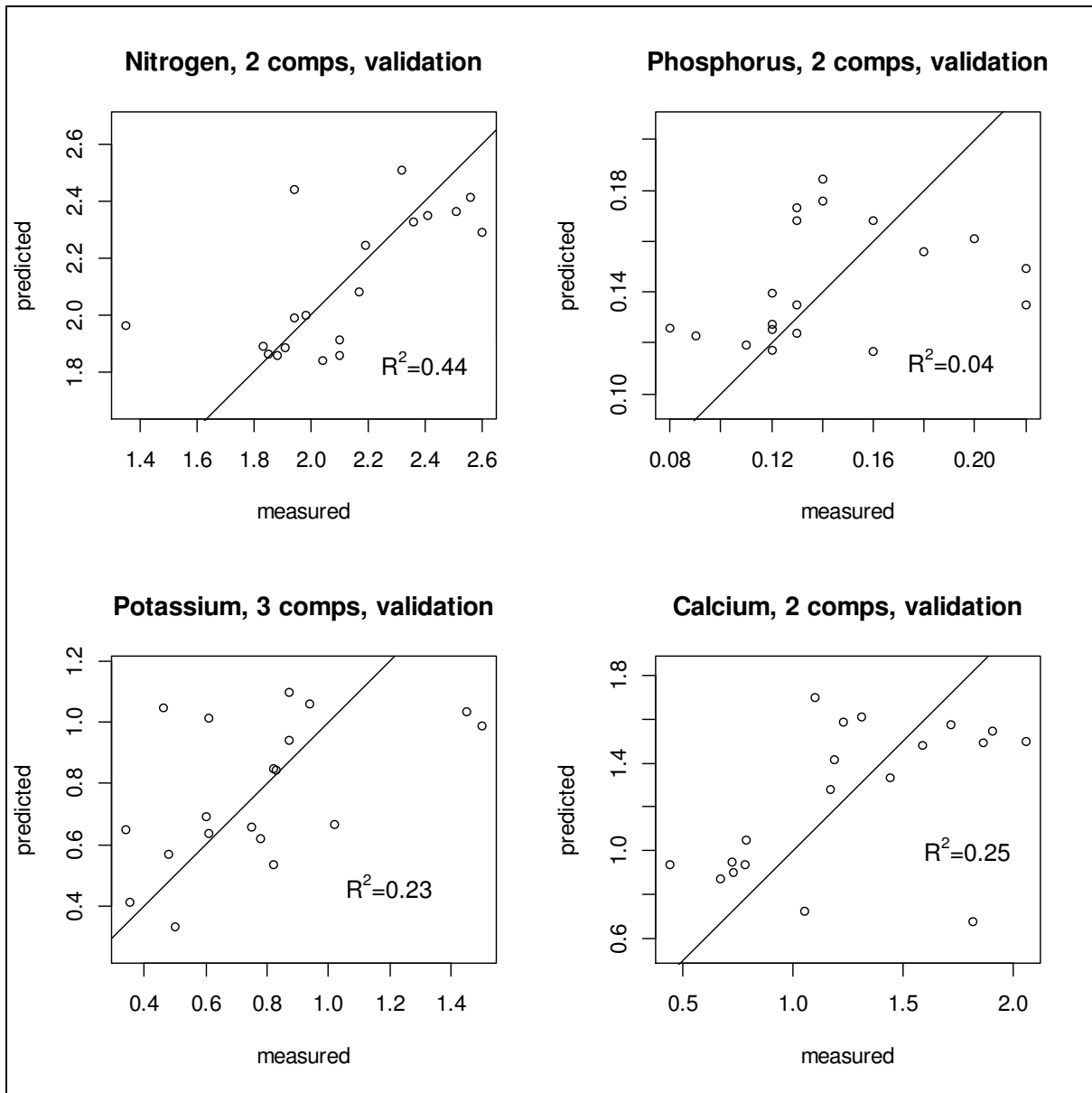


Figure 4.9: Cross-validated predictions for Sentinel-2 data

CHAPTER 5

CONCLUSION

Spatial and taxonomic variation in leaf chemistry is recognized as important both for the functional role that trees play in various ecosystems as well as their response to environmental change. The study combined field campaigns, chemical analysis, laboratory spectroscopy on dry leaves as well as multispectral reflectance of vegetation to identify opportunities for scaling “leaf spectral to chemical” relationships to canopies. The analysis was done in the context of the foliar nutrients, as elements for evaluating organic resources for small scale agricultural production systems and nitrogen in particular as the key indicator for resource quality.

Science based on imaging spectroscopy and techniques has been driven by the assumption that improved identification of particular spectral features leads to better estimation of foliar biochemicals. Season specific analysis showed that wet season models performed better than the dry season one. Various literature indicate that estimation models depend on elements being analysed and the type of plant species. Efforts for the selection of ideal combinations of trees, shrubs and crops that will benefit each other and the environment; and improve income of the small-scale farmer need to be up-scaled.

This study demonstrated a potential for hyperspectral data to estimate leaf nitrogen using multivariate techniques. Ultimately, it affirms the vast utility of laboratory spectroscopy in evaluating and understanding nutrient value of trees and shrubs for integration into livestock and crop production systems. Remote chemical detection on agroecosystems is crucial for understanding resource quality and sustainable utilisation thereof. Nonetheless foliar nutrient estimation using Sentinel-2 data did not show great potential for the estimation of leaf nitrogen, phosphorus, potassium and calcium.

Since livestock production is already an important component of many smallholder farming systems, scientists need to develop innovative ways to capitalise on the use of fast analytical methods to enhance resource management and food security. Research needs to address ways to overcome the potentially negative impacts of loss of biodiversity. The global importance of the ecosystem services provided by

trees (e.g. water and nutrient cycling, fodder, energy production, erosion control, carbon sequestration, biological diversity) has not been fully recognised (Morris, 2011). The study has contributed to understanding measurements between the different components of a multifunctional landscape.

REFERENCES

- Adam, E.M., Mutanga, O., Rugege, D. & Ismail, R. (2012). Discriminating the papyrus vegetation (*Cyperus papyrus* L.) and its co-existent species using random forest and hyperspectral data resampled to HYMAP. *International Journal of Remote Sensing*, 33: 552-569.
- Adan, M.A. (2017). Integrating Sentinel-2 derived indices and terrestrial laser scanner to estimate above-ground biomass/carbon in the Ayer Hitam tropical forest, Malaysia. MSc Thesis, Faculty of Geo-information Science, University of Twente, Enschede, The Netherlands.
- Akinnifesi, F.K., Ajayi, O.C., Sileshi, G., Chirwa, P.W. & Chianu, J. (2010). Fertiliser trees for sustainable food security in the maize-based production systems of East and Southern Africa: a review. *Agronomy for Sustainable Development*, 30: 615–629.
- Analytical Spectral Devices Inc. (1997). *FieldSpec user's guide*. Analytical Spectral Devices Inc., Boulder CO.
- Asner, G.P. & Martin, E.R. (2008). Spectral and chemical analysis of tropical forests: Scaling from leaf to canopy levels. *Remote Sensing of Environment*, 112: 3958–3970.
- Awiti, A. O., Walsh, M. G., Shepherd, K. D. & Kinyamario, J. (2007). Soil condition classification using infrared spectroscopy: A proposition for assessment of soil condition along a tropical forest-cropland chronosequence. *Geoderma*, doi:10.1016/j.geoderma.2007.08.021.
- Baker, C. W., Givens, D. I. & Deaville, E. R. (1994). Prediction of organic matter digestibility *in vivo* of grass silage by near infrared reflectance spectroscopy: effect of calibration method, residual moisture and particle size. *Animal Feed Science and Technology*, 50: 17-26.
- Barnes, R. J., Dhanoa, M. S. & Lister, S. J. (1989). Standard normal variate transformation and de-trending of near infrared diffuse reflectance spectra. *Applied Spectroscopy*, 43: 772-777.

- Blanco, M. & Villarroya, I. (2002). NIR spectroscopy: a rapid-response analytical tool. *Trends in Analytical Chemistry*, 21 (4).
- Campbell, C.R. & Plank, C.O. (1997). Preparation of plant tissue for laboratory analysis, pg37. In *Handbook of Reference Methods for Plant Analysis*. Kalra, Y.P. (eds). CRC Press.
- Chepape, R.M., Mbatha, K.R. & Luseba, D. (2011). Local use and knowledge validation of fodder trees and shrubs browsed by livestock in Bushbuckridge area, South Africa. *Livestock Research for Rural Development*, 23 (6).
- Chirwa, P.W., Larwanou, M., Syampungani, S. & Babalola, F.D. (2015). Management and restoration practices in degraded landscapes of Eastern Africa and requirements for upscaling. *International Forestry Review*, 17 (s3), 4.
- Cho, M.A., Mathieu, R., Asner, G.P., Naidoo, L., van Aardt, J., Ramoelo, A., Debba, P., Wessels, K., Main, R., Smit, I.P.J. & Erasmus, B. (2012). Mapping tree species composition in South African savannas using an integrated airborne spectral and LiDAR system. *Remote Sensing of Environment*, 125: 214–226.
- Clark, R.N. & Roush, T.L. (1984). Reflectance spectroscopy: quantitative analysis techniques for remote sensing applications. *Journal of Geophysical Research*, 89(B7): 6329–6340.
- Clark, R. N. (1999). Spectroscopy of rocks and minerals, and principles of spectroscopy, In: *Manual of Remote Sensing, Volume 3: Remote Sensing for the Earth Sciences*, p3-58. (Rencz, A.N. Ed.) John Wiley and Sons, New York.
- Clifton, K.E., Bradbury, J.W., Vehrencamp, S.L. (1994). The fine-scale mapping of grassland protein densities. *Grass Forage Science*, 49: 1-8.
- 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 the Environment*, 76: 349-359.
- Dhau, I., Adam, E., Mutanga, O., Ayisi, K., Mohamed, E. Rahman, A., Odindi, J. & Masocha, M. (2017). Testing the capability of spectral resolution of the new

multispectral sensors on detecting the severity of grey leaf spot disease in maize crop. *Geocarto International*. doi: 10.1080/10106049.2017.1343391.

De Aldana, B. R. V., Criado, B. G., Ciudad, A. G. & Corona, M. E. P. (1995). Estimation of mineral-content in natural grasslands by near-infrared reflectance spectroscopy. *Communications in Soil Science and Plant Analysis*, 26: 1383–1396.

du Toit, J.T., Rogers, K.H. & (2003). *The Kruger Experience: Ecology and Management of Savanna Heterogeneity*, Island Press, Washington.

Dumas, J.B.A. (1831). Procédes de l'analyse organique. *Annals of Chemistry and of Physics*, 247: 198–213.

El hassan, S.M., Lahlou K. A., Newbold, C.J. & Wallace, R.J. (2000). Chemical composition and degradation characteristics of foliage of some African multipurpose trees. *Animal Feed Science and Technology*, 86: 27-37.

European Space Agency (ESA). (2015). SENTINEL-2 User Handbook. https://earth.esa.int/documents/247904/685211/Sentinel-2_User_Handbook [accessed: 14 September, 2017].

Everson, C.S., Dye, P.J., Gush, M.B. & Everson, T.M. (2011). Water use of grasslands, agroforestry systems and indigenous forests. *Water SA*, 37 (5).

García-Sánchez, F., Galvez-Sola, L., Martínez-Nicolás, J.J., Muelas-Domingo, R. & Nieves, M. (2017). Using near-infrared spectroscopy in agricultural systems, In: *Developments in Near-Infrared Spectroscopy*, 98-127.

<http://www.intechopen.com/books/developments-in-nearinfrared-spectroscopy>

[accessed: 16 November, 2016]. doi: 10.5772/67236.

Garcia, H. & Filzmoser, P. (2017). Multivariate Statistical Analysis using the R package chemometrics. <http://cran.r-project.org> [accessed: 20 September, 2017]).

Garrity, D., Okono, A., Grayson, M. & Parrott, S. (2006). *World Agroforestry into the Future*. Nairobi: World Agroforestry Centre.

Givens, D. I. & Deaville, E. R. (1999). The current and future role of near infrared reflectance spectroscopy in animal nutrition: a review. *Australian Journal of Agricultural Research*, 50: 1131-1145.

- González-Martín, M.I., Escuredo, O., Revilla, I., Vivar-Quintana, A.M., Coello, A.C., Riocerezo, C.P. & Moncada, G.W. (2015). Determination of the mineral composition and toxic element contents of propolis by near infrared spectroscopy. *Sensors*, 15: 27854-27868.
- Hatfield, J.L., Gitelson, A.A. & Schepers, J.S. (2008). Application of spectral remote sensing for agronomic decisions. *Agronomy Journal Supplement*, 117-131.
- Huang, C. J., Han, L. J., Yang, Z. L. & Liu, M. (2009). Exploring the use of near infrared reflectance spectroscopy to predict minerals in straw. *Fuel* 88: 163–168.
- Hungria, M. & Franco, A.A. (1993). Effects of high temperature on nodulation and nitrogen fixation. *Phaseolus vulgaris* L. *Plant and Soil*, 149: 95-102.
- İlknur ŞEN. (2003). Spectroscopic determination of major nutrients (N, P, K) of soil. Masters' Dissertation. İzmir Institute of Technology. İzmir, Turkey.
- International Centre for Research on Agroforestry (ICRAF) (1997). *Annual Report*, 1996. Majestic printing works limited. Nairobi, Kenya.
- Jacobs, M.S., Pettit, N.E., Naiman, R.J. (2007). Nitrogen fixation by the savanna tree *Philenoptera violacea* (Klotzsch) Schrire (Apple leaf) of different ages in a semi-arid riparian landscape. *South African Journal of Botany*, 73: 163–167.
- Kalaba, K.F., Chirwa, P.W., Syampungani, S. and Ajayi, C.O. (2010). Contribution of agroforestry to biodiversity and livelihoods improvement in rural communities of Southern African regions. In: *Tropical Rainforests and Agroforests under Global Change*. Tschardtke *et al.* (eds.). Environmental Science and Engineering, 461–476. DOI 10.1007/978-3-642-00493-3_22.
- Kjeldahl, J. (1883) Neue methode zur bestimmung des stickstoffs in organischen körper. *Zeitschrift für Analytische Chemie*, 22: 366–382.
- 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: 267-287.

- Kwesiga, F., Akinnifesi, F.K., Mafongoya, P.L., McDermott, M.H. & Agumy, A. (2003). Agroforestry research and development in southern Africa during the 1990s: Review and challenges ahead. *Agroforestry Systems*, 59: 173–186.
- Kovacs, B., Gyori, Z., Prokisch, J., Loch, J. & Daniel, P. (1996). *Communications in Soil and Plant Analysis*, 27 (5-8): 1177 – 1198.
- Kumar, L., Schmidt, K.S., Dury, S. & Skidmore, A.K. (2001). Imaging spectroscopy and vegetation science, In: *Image spectroscopy*, pp. 111-154. (Van Der Meer, F.D., De Jong, S.M., Eds.), Kluwer Academic Publishers, Dordrecht.
- Lukhele, M.S. & van Ryssen, J.B.J. (2003). The chemical composition and potential nutritive value of the foliage of four subtropical tree species in southern Africa for ruminants. *South African Journal of Animal Science*, 33 (2).
- Majeke, B., van Aardt, J. & Cho, M.A. (2008). Imaging spectroscopy of foliar biochemistry in forestry environments. *Southern Forests*, 70(3): 275–285.
- Mbow, C., Van Noordwijk, M., Luedeling, E., Neufeldt, H., Minang P. A. and Kowero, G. (2014). Agroforestry solutions to address food security and climate change challenges in Africa. *Current Opinion in Environmental Sustainability*, 6, 61–67.
- Mevik, B. & Wehrens, R. (2007). The pls Package: Principal component and partial least squares regression in R. *Journal of Statistical Software*, 18 (2).
- Morris, C.D. (2011). Rangeland management for sustainable conservation of natural resources, In: *Grassland Productivity and Ecosystem Services*. (Lemaire G, Hodgson J & Chabbi, A. Eds). CABI Publishing, Wallingford, UK.
- Mangiafico, S.S. (2015). An R Companion for the Handbook of Biological Statistics, version 1.3.2. <http://rcompanion.org/documents/RCompanionBioStatistics.pdf>. [accessed: 15 April, 2018].
- Mukolwe, M.O. (1999). The potential of agroforestry in the conservation of high value indigenous trees: a case study of Umzimvubu District, Eastern Cape. Masters' Thesis. School of Environment and Development, University of Natal. Pietermaritzburg.

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: 10823-10843.

Murray, I. (1986). Near infrared reflectance analysis of forages, In: *Recent advances in animal nutrition*, pp. 141- 156. (Haresign, W. & Cole, A. Eds). Butterworths, London.

Mutanga, O. & Ismail R. (2010). Variation in foliar water content and hyperspectral reflectance of *Sirex noctilio* infested *Pinus patula* trees. *Southern Forests*, 72(1): 1–7.

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.

Naes, T., Isacksson, T., Fearn, T. & Davies, T. (2002). *A user-friendly guide to multivariate calibration and classification*. NIR Publications, Chichester.

Nair, P.K.R. (1993). *An introduction to agroforestry*. Kluwer Academic Publishers. Dordrecht, The Netherlands.

Nassoro, Z., Rubanza, C.D.K. & Kimaro, A.A. (2015). Evaluation of nutritive value of browse tree fodder species in semi-arid districts of Tanzania. *Journal of Food, Agriculture and Environment* 13 (3&4): 113–120.

Nomngongo, P.N., Munonde, T.S., Mpupa, A. & Biata, N.R. (2017). Near-infrared spectroscopy combined with multivariate tools for analysis of trace metals in environmental matrices, In: *Developments in Near-Infrared Spectroscopy*, pp131-141. Cranfield University, United Kingdom. (Kyprianidis, K. Eds)

Palm, C.A., Gachengo, C.N., Delve, R.J., Cadisch, G. and Giller, N.E. (2001). Organic inputs for soil fertility management in tropical agroecosystems: application of an organic resource database. *Agriculture, Ecosystems and Environment*, 83: 27–42.

Ramoelo, A., Skidmore, A.K., Cho, M.A., Schlerf, M., Mathieu, R. & Heitkönig, I.M.A. (2012). Regional estimation of savanna grass nitrogen using the red-edge band of the spaceborne RapidEye sensor. *International Journal of Applied Earth Observation and Geoinformation*, 19: 151-162.

Ramoelo, A., Cho, M.A., Madonsela, S., Mathieu, R., van der Korchove, R., Kaszta, Z. & Wolf, E. (2014). A potential to monitor nutrients as an indicator of rangeland quality using space borne remote sensing. Eighth International Symposium of the Digital Earth (ISDE8), IOP Publishing. doi:10.1088/1755-1315/18/1/012094.

Ramoelo, A., Cho, M.A., Mathieu, R., Madonsela, S., van der Kerchove, R., Kaszta, Z. & E. Wolff. (2015). Monitoring grass nutrients and biomass as indicators of rangeland quality and quantity using random forest modelling and WorldView-2 data. *International Journal of Applied Earth Observation and Geoinformation*, 43: 43–54.

Ramoelo, A. & Cho, M.A. (2018). Explaining leaf nitrogen distribution in a semi-arid environment predicted on Sentinel-2 imagery using a field spectroscopy derived model. *Remote Sensing*, 10, 269. doi:10.3390/rs10020269.

Riley, M.L. & Canaves, L.C. (2002). FT-NIR Spectroscopic analysis of nitrogen in cotton leaves. *Biological Systems Engineering: Papers and Publications*. Paper 350.

Rossa, Ü.B., Angelo, A.C., Nisgoski, S., Westphalen, D.J., Frizon, C.N.T. & Hoffmann-Ribani, R. (2015). Application of the NIR method to determine nutrients in yerba mate (*Ilex paraguariensis* A. St.-Hill) leaves. *Communications in Soil Science and Plant Analysis*, 46 (18): 2323-2331.

Shackleton, C.M. (2001). Managing regrowth of an indigenous savannah tree species (*Terminalia sericea*) for fuelwood: the influence of stump dimensions and post-harvest coppice pruning. *Biomass and Bioenergy*, 20: 261–270.

Shi, T., Wang, J., Liu, H. & Wu, G. (2015). Estimating leaf nitrogen concentration in heterogeneous crop plants from hyperspectral reflectance, *International Journal of Remote Sensing*, 36:18, 4652-4667. doi: 10.1080/01431161.2015.1088676

Simonetti, D., Marelli, A., Rodriguez, D., Veselin, V., Strobl, P., Burger, A., Soille P., Achard, F., Eva, H., Stibig, H.J. & Beuchle, R. (2017). Sentinel-2 web platform for

REDD+ monitoring, online web platform for browsing and processing Sentinel-2 data for forest cover monitoring over the Tropics. EUR 28658; doi:10.2760/790249.

Skidmore, A.K., Ferwerda, J.G., Mutanga, O., Van Wieren, S.E., Peel, M., Grant, R.C., Prins, H.H.T., Balcik, F.B. & Venus, V. (2010). Forage quality of savannas - simultaneously mapping foliar protein and polyphenols for trees and grass using hyperspectral imagery. *Remote Sensing of Environment*, 114: 64-72.

Sterckx, S., Knaeps, E., Adriaensen, S., Reisen, I., De Keukelaere, L. Hunter, P., Giardino, E. & Odermatt, D. (2015). OPERA: An atmospheric correction for land and water. In Proceedings of the ESA Sentinel-3 for Science Workshop, Venice, Italy, 2–5 June 2015.

Stevens, A. & Ramirez-Lopez, L. (2015). Miscellaneous functions for processing and sample selection of vis-NIR diffuse reflectance data. Retrieved from <https://github.com/antoinestevens/prospectr> [accessed: 26 September, 2017].

Stuart, B. (2004). *Infrared spectroscopy: fundamentals and applications*. John Wiley & Sons Ltd, London.

Tegegne, E.D. (2008). Importance of *Ficus thonningii* blume in soil fertility improvement and animal nutrition in Gondar Zuria, Ethiopia. MSc Thesis. University of Natural Resources and Applied Life Science, Vienna.

van Deventer, H., Cho, M.A., 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.

van Maarschalkerweerd, M. & Husted, S. (2015). Recent developments in fast spectroscopy for plant mineral analysis. *Frontiers in Plant Science* (6), 169.

Van Wyk, P. & Van Wyk, B. (2013). *Field guide to trees of Southern Africa*. Struik Nature, Cape Town.

Vlaams Instituut Voor Technologisch Onderzoek (VITO). (2017). ICOR plugin for snap toolbox software user manual version 1.0., VITO Remote Sensing Unit, Boeretang, Belgium.

https://blog.vito.be/remotesensing/icor_available [accessed: 17 January, 2018].

Ward, A., Nielsen, A.L. & Møller, H. (2011). Rapid assessment of mineral concentration in meadow grasses by near infrared reflectance spectroscopy *Sensors*, 11: 4830-4839.

Wold, S., Sjöström, M. & Eriksson, L. (2001). PLS-regression: a basic tool of chemometrics. *Chemometrics and intelligent laboratory systems*, 58: 109–130.

World Agroforestry Centre. (2015). Annual Report 2014-2015: Trees, landscapes, climate, food. Nairobi: World Agroforestry Centre.

Yarce, C.J. & Rojas, G. (2012). Near infrared spectroscopy for the analysis of macro and micro nutrients in sugarcane leaves. *Sugar Industry*, 137: 707–710.

<http://pza.sanbi.org/> [accessed: 22 September, 2016].

<https://earth.esa.int/web/sentinel/user-guides/sentinel-2-msi/resolutions>, [accessed: 30 September, 2017]

http://www.gdal.org/frmt_sentinel2.html [accessed: 15 October, 2017]

<http://www.sentinel-hub.com/eotaxonomy/indices> [accessed: 26 October, 2017]

<https://earthexplorer.usgs.gov/> [accessed, 12 September, 2017]

<http://glovis.usgs.gov/> [accessed, 12 September, 2017]

<http://www.dionex-france.com/library/manuals/software/34941-10.pdf> [accessed: 09 March, 2017]

<https://www.spectro.com/products/icp-oes-aes-spectrometers/genesis-icp-analysis> [accessed: 06 April, 2017]

<https://www.asdi.com/about-us/news/asd-introduces-fieldspec-3-spectroradiometer> [accessed: 09 March 2017]

<https://www.intechopen.com/books/developments-in-near-infrared-spectroscopy/near-infrared-spectroscopy-combined-with-multivariate-tools-for-analysis-of-trace-metals-in-environm> (accessed: 10 February 2018).

<http://www.sthda.com/english/wiki/kruskal-wallis-test-in-r>
[accessed: 12 April 2018]

APPENDICES

Appendix A01: Mean nutrient composition (g/100g DM \pm s.d.) for the nine tree species

Species	Season	N	Nitrogen	Phosphorus	Potassium	Calcium
<i>Bauhinia galpinii</i>	Dry	5	1.836 \pm 0.285	0.126 \pm 0.073	0.426 \pm 0.075	1.340 \pm 0.306
	Wet	5	2.562 \pm 0.180	0.164 \pm 0.032	0.884 \pm 0.085	0.812 \pm 0.148
	All	10	2.199 \pm 0.444	0.145 \pm 0.034	0.655 \pm 0.253	1.076 \pm 0.359
<i>Philenoptera violacea</i>	Dry	5	2.280 \pm 0.146	0.112 \pm 0.013	0.806 \pm 0.193	1.640 \pm 0.395
	Wet	5	3.224 \pm 0.242	0.162 \pm 0.029	1.210 \pm 0.294	0.860 \pm 0.250
	All	10	2.752 \pm 0.532	0.137 \pm 0.034	1.008 \pm 0.317	1.250 \pm 0.516
<i>Schotia brachypetala</i>	Dry	5	1.998 \pm 0.152	0.130 \pm 0.053	0.794 \pm 0.151	1.690 \pm 0.342
	Wet	5	2.176 \pm 0.214	0.162 \pm 0.045	1.022 \pm 0.435	1.044 \pm 0.567
	All	10	2.087 \pm 0.198	0.146 \pm 0.049	0.908 \pm 0.330	1.367 \pm 0.557
<i>Peltophorum africanum</i>	Dry	5	1.436 \pm 0.355	0.088 \pm 0.034	0.588 \pm 0.228	1.216 \pm 0.582
	Wet	5	2.056 \pm 0.470	0.084 \pm 0.013	0.692 \pm 0.150	0.882 \pm 0.356
	All	10	1.746 \pm 0.511	0.086 \pm 0.025	0.640 \pm 0.190	1.049 \pm 0.488
<i>Dichrostachys cinerea</i>	Dry	5	2.028 \pm 0.170	0.098 \pm 0.008	0.568 \pm 0.048	2.268 \pm 0.679
	Wet	5	2.646 \pm 0.433	0.144 \pm 0.062	0.960 \pm 0.312	0.954 \pm 0.199
	All	10	2.337 \pm 0.450	0.121 \pm 0.048	0.766 \pm 0.295	1.611 \pm 0.838
<i>Vachellia gerrardii</i>	Dry	5	1.668 \pm 0.072	0.074 \pm 0.013	0.700 \pm 0.173	2.300 \pm 0.663
	Wet	5	2.644 \pm 0.215	0.134 \pm 0.022	0.892 \pm 0.158	0.812 \pm 0.176
	All	10	2.156 \pm 0.536	0.104 \pm 0.036	0.796 \pm 0.186	1.556 \pm 0.908

<i>Combretum apiculatum</i>	Dry	5	1.772 ± 0.314	0.092 ± 0.015	0.570 ± 0.154	1.150 ± 0.361
	Wet	5	2.404 ± 0.393	0.142 ± 0.054	0.878 ± 0.476	1.066 ± 0.518
	All	10	2.088 ± 0.473	0.117 ± 0.045	0.724 ± 0.371	1.108 ± 0.423
<i>Terminalia sericea</i>	Dry	5	1.004 ± 0.209	0.078 ± 0.016	0.590 ± 0.110	1.086 ± 0.296
	Wet	5	1.956 ± 0.176	0.116 ± 0.013	0.754 ± 0.117	0.690 ± 0.170
	All	10	1.480 ± 0.534	0.097 ± 0.025	0.672 ± 0.138	0.888 ± 0.309
<i>Euclea natalensis</i>	Dry	5	1.756 ± 0.118	0.086 ± 0.005	0.532 ± 0.105	1.372 ± 0.382
	Wet	5	1.684 ± 0.107	0.110 ± 0.029	0.948 ± 0.162	0.550 ± 0.120
	All	10	1.720 ± 0.113	0.098 ± 0.023	0.740 ± 0.254	0.961 ± 0.509

DM = dry matter; s.d. = standard deviation

Appendix A02: Statistical analysis for chemical data

Summary Statistics

```
> tapply(Nitrogen, list(Species, Season), sd)
```

	summer	winter
A. gerrardii	0.2154762	0.07190271
B. galpinii	0.1797776	0.28465769
C. apiculatum	0.3932302	0.31427695
D. cinerea	0.4331628	0.17035258
E. natalensis	0.1071448	0.11844830
P. africanum	0.4698191	0.35514786
P. violacea	0.2419297	0.14645819
S. brachypetala	0.2138457	0.15155857
T. sericea	0.1758693	0.20863844

```
> tapply(Nitrogen, list(Season), sd)
```

	D	W
	0.4049457	0.5142831

statistical differences between leaf samples

```
aov1=aov(Nitrogen~Species*Season, data=leaf)
```

```
> summary(aov1)
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
Species	8	10.893	1.362	18.361	1.76e-14	***
Season	1	12.947	12.947	174.577	< 2e-16	***
Species:Season	8	1.649	0.206	2.779	0.00995	**
Residuals	70	5.191	0.074			

```
TukeyHSD(aov1)
```

Tukey multiple comparisons of means

	p adj
B. galpinii:summer-A. gerrardii:summer	1.0000000
C. apiculatum:summer-A. gerrardii:summer	0.9914483
D. cinerea:summer-A. gerrardii:summer	1.0000000
E. natalensis:summer-A. gerrardii:summer	0.0000228
P. africanum:summer-A. gerrardii:summer	0.0594210
P. violacea:summer-A. gerrardii:summer	0.0677239
S. brachypetala:summer-A. gerrardii:summer	0.3181456
T. sericea:summer-A. gerrardii:summer	0.0096906
C. apiculatum:summer-B. galpinii:summer	0.9999506
D. cinerea:summer-B. galpinii:summer	1.0000000
E. natalensis:summer-B. galpinii:summer	0.0001596
P. africanum:summer-B. galpinii:summer	0.2008099
P. violacea:summer-B. galpinii:summer	0.0159846
S. brachypetala:summer-B. galpinii:summer	0.6566069
T. sericea:summer-B. galpinii:summer	0.0438994
D. cinerea:summer-C. apiculatum:summer	0.9906704
E. natalensis:summer-C. apiculatum:summer	0.0051100
P. africanum:summer-C. apiculatum:summer	0.8045875
P. violacea:summer-C. apiculatum:summer	0.0005993
S. brachypetala:summer-C. apiculatum:summer	0.9950827
T. sericea:summer-C. apiculatum:summer	0.3932368

```

E. natalensis:summer-D. cinerea:summer 0.0000217
P. africanum:summer-D. cinerea:summer 0.0574883
P. violacea:summer-D. cinerea:summer 0.0699488
S. brachypetala:summer-D. cinerea:summer 0.3111064
T. sericea:summer-D. cinerea:summer 0.0093177
P. africanum:summer-E. natalensis:summer 0.7145580
P. violacea:summer-E. natalensis:summer 0.0000000
S. brachypetala:summer-E. natalensis:summer 0.2399240
T. sericea:summer-E. natalensis:summer 0.9705805
P. violacea:summer-P. africanum:summer 0.0000001
S. brachypetala:summer-P. africanum:summer 0.9999991
T. sericea:summer-P. africanum:summer 0.9999999
S. brachypetala:summer-P. violacea:summer 0.0000026
T. sericea:summer-P. violacea:summer 0.0000000
T. sericea:summer-S. brachypetala:summer 0.9967047
B. galpinii:winter-A. gerrardii:winter 0.9998859
C. apiculatum:winter-A. gerrardii:winter 0.9999999
D. cinerea:winter-A. gerrardii:winter 0.7613438
E. natalensis:winter-A. gerrardii:winter 1.0000000
P. africanum:winter-A. gerrardii:winter 0.9940506
P. violacea:winter-A. gerrardii:winter 0.0395859
S. brachypetala:winter-A. gerrardii:winter 0.8613782
T. sericea:winter-A. gerrardii:winter 0.0153911
C. apiculatum:winter-B. galpinii:winter 1.0000000
D. cinerea:winter-B. galpinii:winter 0.9993532
E. natalensis:winter-B. galpinii:winter 1.0000000
P. africanum:winter-B. galpinii:winter 0.5964289
P. violacea:winter-B. galpinii:winter 0.4091811
S. brachypetala:winter-B. galpinii:winter 0.9999303
T. sericea:winter-B. galpinii:winter 0.0004577
D. cinerea:winter-C. apiculatum:winter 0.9834835
E. natalensis:winter-C. apiculatum:winter 1.0000000
P. africanum:winter-C. apiculatum:winter 0.8436100
P. violacea:winter-C. apiculatum:winter 0.1956176
S. brachypetala:winter-C. apiculatum:winter 0.9955400
T. sericea:winter-C. apiculatum:winter 0.0018723
E. natalensis:winter-D. cinerea:winter 0.9705805
P. africanum:winter-D. cinerea:winter 0.0556104
P. violacea:winter-D. cinerea:winter 0.9858772
S. brachypetala:winter-D. cinerea:winter 1.0000000
T. sericea:winter-D. cinerea:winter 0.0000048
P. africanum:winter-E. natalensis:winter 0.8882676
P. violacea:winter-E. natalensis:winter 0.1575672
S. brachypetala:winter-E. natalensis:winter 0.9906704
T. sericea:winter-E. natalensis:winter 0.0026308
P. violacea:winter-P. africanum:winter 0.0003488
S. brachypetala:winter-P. africanum:winter 0.0900913
T. sericea:winter-P. africanum:winter 0.4585160
S. brachypetala:winter-P. violacea:winter 0.9593250
T. sericea:winter-P. violacea:winter 0.0000000
T. sericea:winter-S. brachypetala:winter 0.0000100

```

```
> aov4=aov(Nitrogen~Type:Season,data=leaf)
```

```
> summary(aov4)
```

```

      Df Sum Sq Mean Sq F value    Pr(>F)
Type:Season  5   13.95   2.7906   17.32 9.49e-12 ***
Residuals  84   13.53   0.1611
---

```

```
TukeyHSD(aov4)
```

```
$`Type:Season`
```

```

              diff            lwr            upr            p adj
B:summer-A:summer -0.20533333 -0.6327480  0.222081286  0.7262355

```

```
C:summer-A:summer -0.63933333 -1.0667480 -0.211918714 0.0005067
C:summer-B:summer -0.43400000 -0.8614146 -0.006585381 0.0444353
B:winter-A:winter -0.32733333 -0.7547480 0.100081286 0.2336736
C:winter-A:winter -0.52733333 -0.9547480 -0.099918714 0.0069275
C:winter-B:winter -0.20000000 -0.6274146 0.227414619 0.7477100
```

```
> kruskal.test(Phosphorus~Type,data=leaf)
```

```
kruskal-wallis rank sum test
```

```
data: Phosphorus by Type
Kruskal-wallis chi-squared = 23.438, df = 2, p-value = 8.137e-06
```

```
kruskal.test(Potassium~Type,data=leaf)
```

```
kruskal-wallis rank sum test
```

```
data: Potassium by Type
Kruskal-wallis chi-squared = 4.1592, df = 2, p-value = 0.125
```

```
> kruskal.test(Calcium~Type,data=leaf)
```

```
kruskal-wallis rank sum test
```

```
data: Calcium by Type
Kruskal-wallis chi-squared = 5.1574, df = 2, p-value = 0.07587
```

```
> kruskal.test(Phosphorus~Season,data=leaf)
```

```
kruskal-wallis rank sum test
```

```
data: Phosphorus by Season
Kruskal-wallis chi-squared = 21.032, df = 1, p-value = 4.517e-06
```

```
> kruskal.test(Potassium~Season,data=leaf)
```

```
kruskal-wallis rank sum test
```

```
data: Potassium by Season
Kruskal-wallis chi-squared = 26.648, df = 1, p-value = 2.441e-07
```

```
> kruskal.test(Calcium~Season,data=leaf)
```

```
kruskal-wallis rank sum test
```

```
data: Calcium by Season
Kruskal-wallis chi-squared = 36.686, df = 1, p-value = 1.388e-09
```

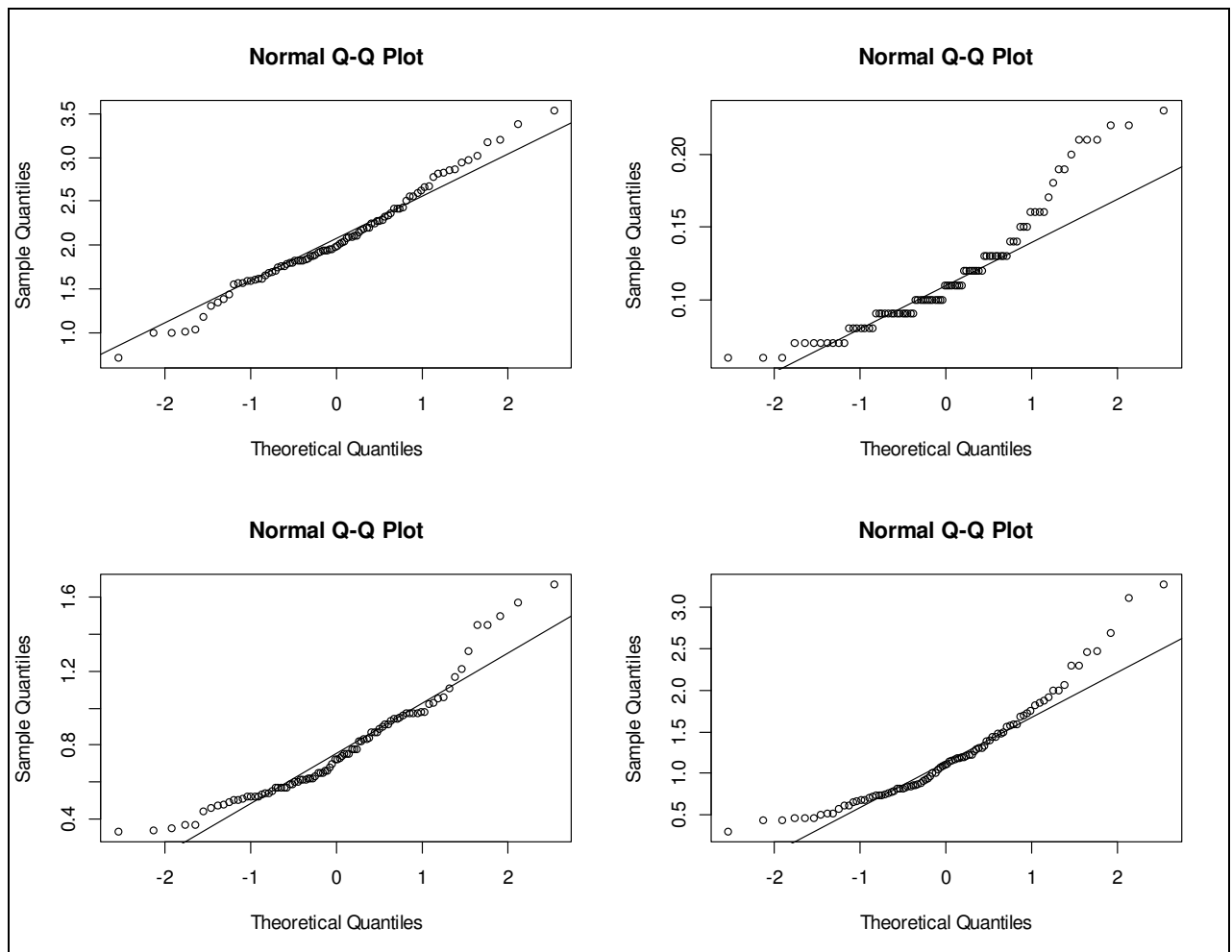
```
> pairwise.wilcox.test(leaf$Phosphorus, leaf$Type,
+                       p.adjust.method = "BH")
```

```
Pairwise comparisons using wilcoxon rank sum test
```

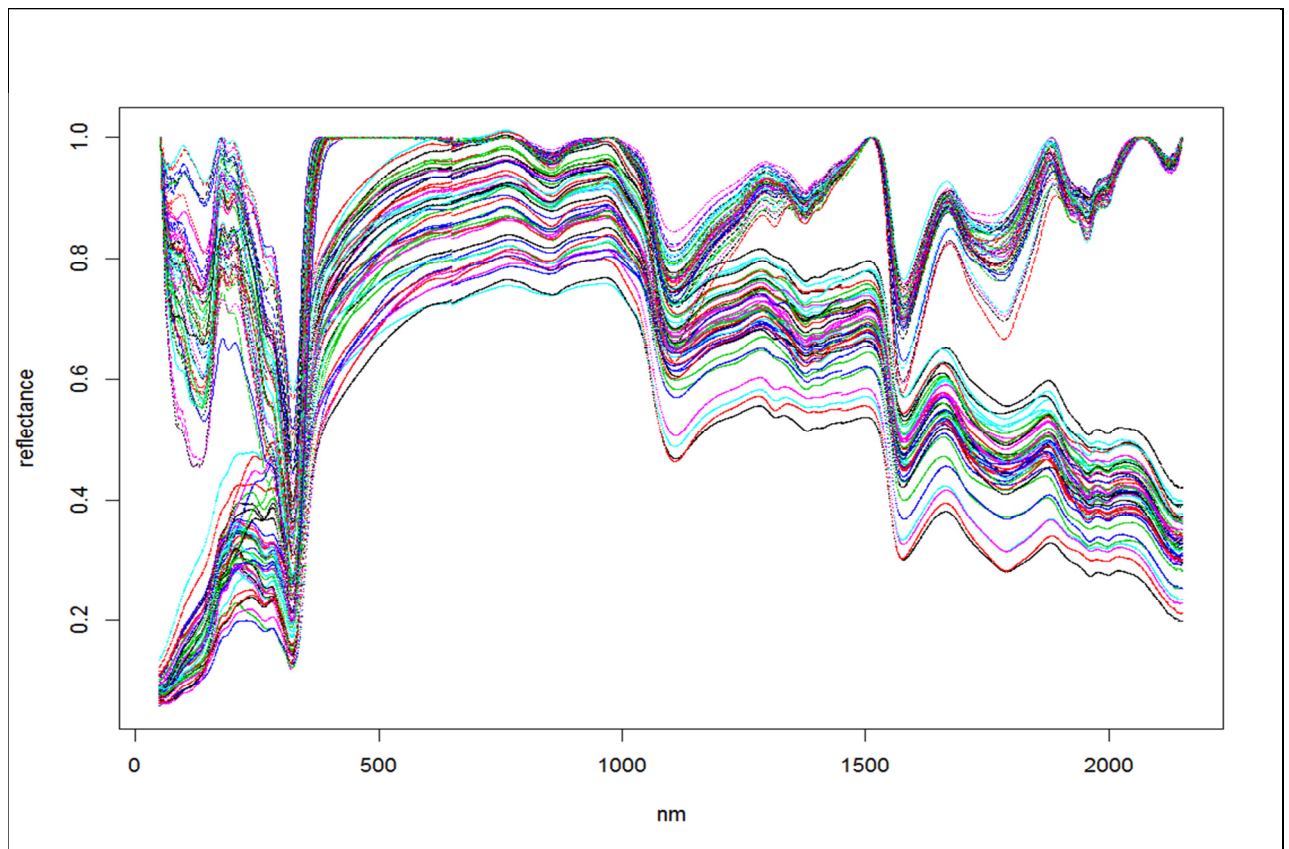
```
data: leaf$Phosphorus and leaf$Type
```

```
      BLL      NLL
NLL 6.7e-05 -
NOL 6.4e-05 0.64
```

Appendix A03: Q-Q Plots for nitrogen, phosphorus, potassium and calcium



Appendix A04: The continuum-removed spectra and spectral analysis.



CALL:

```
Leafdata=read.csv("winterLeaves.csv",header = TRUE, row.names = 1)
head(Leafdata)

ALLSPECTRA<-Leafdata[,51:2151]
ALLSPECTRA<-as.matrix(ALLSPECTRA)

str(ALLSPECTRA)

Nitrogen<-Leafdata[,2152]
Phosphorus<-Leafdata[,2153]
Potassium<-Leafdata[,2154]
Calcium<-Leafdata[,2155]

Chemical<-Leafdata[,2152:2155]
Chemical<-as.matrix(Chemical)

Leafdata<-data.frame(I(ALLSPECTRA),I(Chemical))

names(Leafdata)

wave_allspectra<-seq(51,2151, by=1)

windows(7,7)
matplot(wave_allspectra,t(Leafdata$ALLSPECTRA),lty=1,pch=".",xlab="nm",ylab="reflectance")

#####

crALLSPECTRA <- continuumRemoval(Leafdata$ALLSPECTRA,wave_allspectra,type='R')

matlines(wave_allspectra,t(crALLSPECTRA[1:45,]))
```

```
#####

leafTrain <- Leafdata[1:45,]

leafChem2 <- pls(Chemical ~ crALLSPECTRA, ncomp = 10, data = leafTrain, validation = "LOO")

summary(leafChem2)
R2(leafChem2)

windows(7,7)
plot(RMSEP(leafChem2), legendpos="topright", main = "RMSE for Summer Model")

windows(7,7)
plot(R2(leafChem2), legendpos="topright")

NleafChem <- pls(Nitrogen ~ crALLSPECTRA, ncomp = 10, data = leafTrain, validation = "LOO")
PleafChem <- pls(Phosphorus ~ crALLSPECTRA, ncomp = 10, data = leafTrain, validation = "LOO")
KleafChem <- pls(Potassium ~ crALLSPECTRA, ncomp = 10, data = leafTrain, validation = "LOO")
CaleafChem <- pls(Calcium ~ crALLSPECTRA, ncomp = 10, data = leafTrain, validation = "LOO")

summary(NleafChem)
summary(PleafChem)
summary(KleafChem)
summary(CaleafChem)

R2(NleafChem)
R2(PleafChem)
R2(KleafChem)
R2(CaleafChem)

windows(7,7)
par(mfrow=c(2,2), pch=16)
plot(RMSEP(NleafChem), legendpos="topleft")
plot(RMSEP(PleafChem), legendpos="topleft")
plot(RMSEP(KleafChem), legendpos="topleft")
plot(RMSEP(CaleafChem), legendpos="topleft")

par(mfrow=c(2,2), pch=1)
plot(NleafChem, ncomp = 6, asp = 1, line = TRUE)
plot(PleafChem, ncomp = 6, asp = 1, line = TRUE)
plot(KleafChem, ncomp = 7, asp = 1, line = TRUE)
plot(CaleafChem, ncomp = 5, asp = 1, line = TRUE)

> leafChem <- pls(Chemical ~ crALLSPECTRA, ncomp = 10, data = leafTrain, validation = "LOO")
>summary(leafChem).
```

Appendix A05: Spectral band and indices used from Sentinel-2 image.

Plots	Band2	Band3	Band4	Band5	Band6	Band7	Band8	Band9	Band10	Band11	NDVI	NDVI705	mNDVI705	RE NDVI	PSRI-NIR	NDVI-GRE	CRI2	CHL RE	N	P	K	Ca
K101	0.04404	0.064367	0.061992	0.10728	0.226904	0.27721	0.292548	0.293712	0.211255	0.135423	0.650295	0.357959	0.486071	0.126371	0.061366	0.041858	13.38536	0.366709	2.6	0.13	0.87	0.78
K102	0.03125	0.056883	0.048638	0.118451	0.308488	0.380663	0.368677	0.403325	0.240904	0.123395	0.766901	0.445114	0.521448	0.088883	0.047164	0.043623	23.55822	0.321288	2.32	0.14	0.83	0.67
K103	0.03125	0.056883	0.048638	0.118451	0.308488	0.380663	0.368677	0.403325	0.240904	0.123395	0.766901	0.445114	0.521448	0.088883	0.047164	0.043623	23.55822	0.321288	2.56	0.16	0.82	1.05
K104	0.032379	0.058003	0.043513	0.126521	0.28917	0.334011	0.343432	0.359125	0.23973	0.150028	0.775097	0.391274	0.463476	0.085776	0.032419	0.044958	22.98048	0.368402	2.51	0.18	0.87	0.72
K301	0.052212	0.085946	0.07675	0.144138	0.314689	0.371232	0.38985	0.392966	0.284238	0.191131	0.671024	0.37171	0.481232	0.106681	0.062944	0.057672	12.21509	0.369727	2.41	0.2	1.5	0.44
K302	0.059517	0.089885	0.074595	0.143794	0.249973	0.299061	0.325706	0.327745	0.301588	0.214747	0.627308	0.269651	0.386482	0.131554	0.046292	0.056385	9.847548	0.441483	2.19	0.22	1.45	0.79
K303	0.054999	0.094096	0.10361	0.13955	0.319692	0.381372	0.405677	0.403401	0.241623	0.146335	0.593116	0.39226	0.515806	0.11854	0.119828	0.05581	11.0163	0.343992	1.94	0.13	0.61	1.82
K304	0.080201	0.116732	0.130895	0.131827	0.305962	0.352673	0.378638	0.376242	0.258152	0.181114	0.486217	0.397759	0.627766	0.106159	0.133885	0.056757	4.883037	0.348162	2.36	0.14	0.94	0.73
K305	0.119678	0.183402	0.235529	0.229207	0.307269	0.358334	0.390948	0.368648	0.342335	0.260871	0.248085	0.145508	0.262727	0.119847	0.296333	0.045499	3.992922	0.586286	1.98	0.12	0.61	1.44
A101	0.061847	0.079748	0.116609	0.139459	0.167024	0.178972	0.194178	0.205644	0.287548	0.225591	0.249592	0.089939	0.150801	0.075178	0.282016	0.019904	8.998356	0.718202	1.88	0.16	0.34	1.31
A102	0.07014	0.103068	0.167552	0.150408	0.174027	0.204419	0.259278	0.220989	0.305911	0.229138	0.214902	0.072799	0.128254	0.196747	0.375703	0.022149	7.608633	0.580103	2.1	0.13	0.5	1.23
A103	0.07014	0.103068	0.167552	0.150408	0.174027	0.204419	0.259278	0.220989	0.305911	0.229138	0.214902	0.072799	0.128254	0.196747	0.375703	0.022149	7.608633	0.580103	1.91	0.12	0.35	1.87
A104	0.048387	0.066423	0.102336	0.151886	0.182367	0.209907	0.197263	0.227604	0.308855	0.224581	0.316847	0.091191	0.128353	0.039238	0.273486	0.021046	14.08269	0.769968	1.35	0.09	0.48	1.1
A105	0.063828	0.092616	0.104226	0.146947	0.183418	0.197528	0.248942	0.248272	0.339214	0.294379	0.409766	0.110394	0.179915	0.151551	0.162278	0.037951	8.861935	0.590286	1.94	0.13	0.46	1.19
A401	0.063861	0.084824	0.123907	0.175835	0.201318	0.231307	0.234854	0.256291	0.3555	0.263471	0.309251	0.067567	0.102165	0.076889	0.255674	0.026232	9.971896	0.748696	1.85	0.08	0.75	1.91
A402	0.046239	0.066229	0.092123	0.165517	0.199657	0.226545	0.203394	0.250779	0.349496	0.275511	0.37653	0.093489	0.125193	0.009272	0.225594	0.024937	15.58528	0.813776	2.04	0.11	0.6	2.06
A403	0.046239	0.066229	0.092123	0.165517	0.199657	0.226545	0.203394	0.250779	0.349496	0.275511	0.37653	0.093489	0.125193	0.009272	0.225594	0.024937	15.58528	0.813776	1.83	0.12	0.78	1.72
A404	0.056639	0.081162	0.118364	0.141569	0.181672	0.189955	0.21976	0.22738	0.300086	0.212591	0.29988	0.124064	0.190999	0.094881	0.280873	0.024339	10.59199	0.644199	2.1	0.12	0.82	1.59
A405	0.045392	0.062811	0.079305	0.107724	0.158674	0.181657	0.211187	0.209148	0.233914	0.157494	0.453995	0.191258	0.29013	0.14198	0.160583	0.028516	12.74729	0.510086	2.17	0.22	1.02	1.17