# Assessment of foliar nitrogen as an indicator of vegetation stress using remote sensing: The case study of Waterberg region, Limpopo Province.

**By**

**Enoch Khomotšo Manyashi**

submitted in accordance with the requirements

for the degree of

**MASTER OF SCIENCE**

in the subject of

**ENVIRONMENTAL MANAGEMENT**

at the

**UNIVERSITY OF SOUTH AFRICA**

**SUPERVISOR: DR. ABEL RAMOELO**

**CO-SUPERVISOR: MR. MAARTEN JORDAAN**

**JUNE 2015**

Student number: 44649088

# DECLARATION

I declare that "Assessment of foliar nitrogen as an indicator of vegetation stress using Remote Sensing: The case study of Waterberg region, Limpopo Province**"** is my own work and all the sources that I have used or quoted have been indicated and acknowledged by means of complete references.

……………………………                                        ……………………………………..

SIGNATURE                                        DATE

(Mr. Enoch Khomotšo Manyashi).

**ACKNOWLEDGEMENT**

This work was possible with the help provided by the following:

- CSIR for granting me with the resources such as the research data to complete this research
- For lessons and patience from my supervisor Dr A Ramoelo, he dedicated and invested a lot to ensure that I learn and apply the skills he holds.
- University of South Africa (UNISA) M & D Research Bursary department for offering me a bursary which financed my studies since 2012.
- The Department of Environmental Sciences with the great help of my co-supervisor Mr M Jordaan.
- The support from my partner and family (immediate and extended), the inspiration I got from my daughter, friends and more importantly the Almighty. I have learnt a lot from this journey since 2012, and want to impart these skills to bettering our economy and the ecosystem.

## ABSTRACT

Vegetation status is a key indicator of the ecosystem condition in a particular area. The study objective was about the estimation of leaf nitrogen (N) as an indicator of vegetation water stress using vegetation indices especially the red edge based ones, and how leaf N concentration is influenced by various environmental factors. Leaf nitrogen was estimated using univariate and multivariate regression techniques of stepwise multiple linear regression (SMLR) and random forest. The effects of environmental parameters on leaf nitrogen distribution were tested through univariate regression and analysis of variance (ANOVA). Vegetation indices were evaluated derived from the analytical spectral device (ASD) data, resampled to RapidEye. The multivariate models were also developed to predict leaf N. The best model was chosen based on the lowest root mean square error (RMSE) and higher coefficient of determination ($R^2$) values. Univariate results showed that red edge based vegetation index called MERRIS Terrestrial Chlorophyll Index (MTCI) yielded higher leaf N estimation accuracy as compared to other vegetation indices. Simple ratio (SR) based on the bands red and near-infrared was found to be the best vegetation index for leaf N estimation with exclusion of red edge band for stepwise multiple linear regression (SMLR) method. Simple ratio (SR3) was the best vegetation index when red edge was included for stepwise linear regression (SMLR) method. Random forest prediction model achieved the highest leaf N estimation accuracy, the best vegetation index was Red Green Index (RGI1) based on all bands with red green index when including the red edge band. When red edge band was excluded the best vegetation index for random forest was Difference Vegetation Index (DVI1). The results for univariate and multivariate results indicated that the inclusion of the red edge band provides opportunity to accurately estimate leaf N. Analysis of variance results showed that vegetation and soil types have a significant effect on leaf N distribution with *p-values*<0.05. Red edge based indices provides opportunity to assess vegetation health using remote sensing techniques.



Key words: foliar nitrogen, remote sensing, red edge, vegetation index, leaf N estimation, univariate regression, multivariate regression, indicator, vegetation stress, leaf N map.

## TABLE OF CONTENT

## LIST OF FIGURES



## LIST OF TABLES



## LIST OF ABBREVIATION

ARC= Agricultural Research Council

ANOVA= Analysis of variance

ASD= analytical spectral devices, Inc.

AZa 7= Subtropical Alluvial Vegetation

ATCOR= atmospheric and topographic correction software

BGI= Blue Green Pigment Index

CR= continuum removal

CSIR= Council for Scientific and Industrial Research

DEM= digital elevation model

EVI= enhanced vegetation index

GI= Green index

GRR= Green/Red Index

LAI= leaf area index

MTCI= MERIS terrestrial chlorophyll index

NDVI= normalized difference vegetation index

N= nitrogen

NIR= Near infrared region of the reflectance

NGRR= Nitrogen Green/Red Ratio

nm= Nano meter

NRI= Nitrogen reflectance index

OSAVI= Optimized soil-adjusted vegetation index

$R^2$= coefficient of determination

RGI= Red/ Green Index

RMSE= Root mean square error

RE_NDVI= Red Edge_ Normalised Difference Vegetation Index

RE_SR= Red Edge_ Simple Ratio

RS= remote sensing

SAVI= Soil-adjusted vegetation index

SIPI= Structure insensitive pigment index

SMLR= Stepwise multiple linear regression

SLC= Soil line concept

SOTERSAF= Soil and terrain of Southern Africa database

SR= Simple ratio

SRTM= Shuttle radar terrain mission

SVcb 12= Central Sandy Bushveld

SVcb 16= Western Sandy Bushveld

SVcb 18= Roodeberg Bushveld

SVcb 19= stands for Limpopo Sweet Bushveld

SWIR= Shortwave infrared region of the reflectance

VIS= Visible region of the reflectance

# CHAPTER 1: BACKGROUND

## 1.1. Introduction and background

Vegetation health is dependent on leaf nitrogen (N) concentrations. Plant nutrient such as Leaf N concentration, leaf water content and plant pigments such as chlorophyll concentration can be used as indicators of water stress for vegetation (Field and Mooney, 1986; Ollinger *et al*, 2002). Factors such as increasing temperature and inadequate water supply as a results of erratic rainfall or unfavourable climate, affect the health or condition of vegetation. Vegetation also experience stress due to suboptimal conditions, leading to plant physiological functions such as light and dark photosynthesis decline from their physiological standard (Logan *et al*, 2003; Ninements, 2010).

Human action has altered the land surface significantly since the beginning of industrial revolution, as more than half of accessible water is used by man. More nitrogen is fixed through anthropogenic means than any other way; this is a clear indication of the impact of man on the environment (Vitousek *et al*, 1997). A practical example of man-made activities include burning of fossil fuels which affect the environment including vegetation negatively through emissions of toxins such as carbon monoxide, sulphur dioxide and heavy metals (Kampa & Castanas, 2008). Impacts of anthropogenic (man-made) activities on vegetation health include agricultural activities, mining and urbanization. It is evident that the impact of human beings on the environment is increasingly strenuous, hence the need to study how vegetation reacts to stress.

The Waterberg area has developments which impact on the natural vegetation, such as mining, agriculture, and urban development. For an example the residential development in the town of Lephalale, the construction of the second power station named Matimba B and the commissioning of additional open cast coal mines are likely to cause an increase in water requirements as well as effluent to process in the catchment (Ramoelo *et al*, 2014). Hydrological transfer schemes to the Mogol River will seem to increase the water availability and planned rising of the Mokolo Dam wall will further change hydrological conditions to which riparian vegetation will be exposed to. Monitoring the condition of natural resource base and its ecosystem services can enable management intervention and conservation planning of ecosystems.

Ramoelo *et al*, (2013) studied vegetation nutrients in a context of sustainable livestock and wildlife grazing, whereby amongst other objectives of the study was to estimate and map foliar and canopy N at a regional scale using high resolution spaceborne multispectral sensor (i.e. RapidEye). The RapidEye sensor contains five spectral bands in the visible-to-near infrared (VNIR), including a red edge band centered at 710 nm. The importance of the red-edge band has been widely demonstrated in many studies for estimation of foliar chlorophyll and leaf N concentration, especially through field spectroscopy (Cho & Skidmore, 2006; Darvishzadeh *et al*, 2008; Huang *et al*, 2004). It is known that grass N concentration is an indicator of grass quality as it is positively correlated to protein content (Clifton *et al*, 1994; Wang, 2004). Recently, Ramoelo *et al,* (2014) demonstrated that remote sensing tools can be used to assess plant water stress, using leaf water potential and leaf N as indicators. The study did not assess the relationship between leaf N distribution and environmental parameters, including vegetation, soil and geological types, so this study aims to fill such a gap.

This study intends to estimate leaf N as an indicator of vegetation water stress using remote sensing techniques. The study will further assess how leaf N is influenced by other environmental factors. Success in estimating leaf N was possible because of the development of hyperspectral remote sensing. Hyperspectral remote sensing have displayed the utility of red edge bands to estimate leaf N and chlorophyll concentrations (Cho & Skidmore, 2006; Darvishzadeh *et al*, 2008; Huang *et al*, 2004). In this study, vegetation indices computed from red edge bands, also known as red edge based broadband indices derived from red (710 nm) and near infrared (800 nm) (Hansen and Schjoerring, 2003; Mutanga and Skidmore, 2007; Ramoelo, 2014) were tested to assess a potential to predict vegetation water stress.

### 1.2. Problem statement

The problem researched in this study is whether leaf N concentration in vegetation, acts as an indicator of water stress or not. This is due to land use impacts around Lephalale area which calls for the study of the current hydrological conditions effects. These include the second power station construction (Matimba B); and the commissioning of another open cast coal mines which increases water needs and effluent to process in the catchment. Estimating leaf N as an indicator of vegetation water stress will help to understand environmental impact of the land use in Lephalale area.

Remote sensing indices help in estimating vegetation condition at sub-regional level for monitoring purposes. Therefore several vegetation indices are investigated to predict leaf nutrients during wet periods. Vegetation indices computed from red edge bands which are also called narrow-band indices, improve estimation of leaf N better than conventional broad-band indices derived from red (680 nm) and near infrared (800 nm) (Hansen and Schjoerring, 2003; Mutanga & Skidmore, 2007). There are a few systems to monitor natural vegetation condition over a large area. The purpose of this study is to estimate leaf N as an indicator of vegetation condition using vegetation indices and various statistical techniques.

More innovation in this study is the development of leaf N estimation models based on ASD measured data re-sampled to RapidEye spectral band configurations. Eventually, the best models will be inverted on the actual RapidEye image to estimate the spatial distribution of leaf N.

## 1.3. Aims and research objectives

The study aims to asses water stress on vegetation using leaf Nitrogen (N) concentration as an indicator:

Specific objective:

- To estimate leaf N concentrations using vegetation indices
- To determine if leaf N concentration vary across different vegetative land cover or vegetation types.
- To determine if foliar concentration vary across different soil types, slope and aspect.

## 1.4. Research hypotheses

Hypotheses

- Alternative hypothesis: The inclusion of red edge band in the vegetation indices improves the estimation accuracy of leaf N.
    - The null hypothesis: The inclusion of red edge band in the vegetation indices does not improve the estimation accuracy of leaf Nitrogen.
- The alternative hypothesis: leaf N varies across different vegetative land cover or vegetation types.

- The null hypothesis: Leaf N does not vary across different vegetative land cover or vegetation types.

### 1.5 Motivation of the study

Water is a limited resource in the semi-arid environments. High proportion of water in the semi-arid environments are used for various developmental activities such mining, agriculture and domestic use. Often, these activities out-compete natural vegetation through water use, which induces vegetation stress. The vegetation stress or condition has high implications to the conservation of biodiversity, which could eventually lead to the loss of species and their habitats. Land use by man alters the entire ecosystem (Vitousek *et al,* 1997), and the area of Lephalale is under serious developments of mining, agriculture, urbanization.

Therefore, there is a need to develop and assess spatially explicit tools to monitor the condition of the vegetation. Remote sensing has proved to be an alternative tool to assess the status of vegetation. This technique collects a lot of environmental impact data in the most scientific and cost effective way through satellite imagery (Vitousek *et al,* 1997). Old methods of collecting environmental data cannot achieve what remote sensing accomplishes. This will enable the natural resource and environmental planners to take informed decisions to preserve or conserve biodiversity.

### 1.6. Study area

The study was conducted in Waterberg region, Limpopo Province (see Figure 4.1). The area is semi-arid with a general shortage of water. Several land cover types such as agriculture, private game reserves, power stations, built up (residential, industrial and commercial) and natural vegetation occur in this region (WDEMF-Draft Report, 2010). The erection of new power station in this area might exert more pressure on water availability and use. It is therefore imperative to understand the distribution of leaf N to know the vegetation stress levels. Below are the details of the geology, landscape, climate and hydrology of the study area.

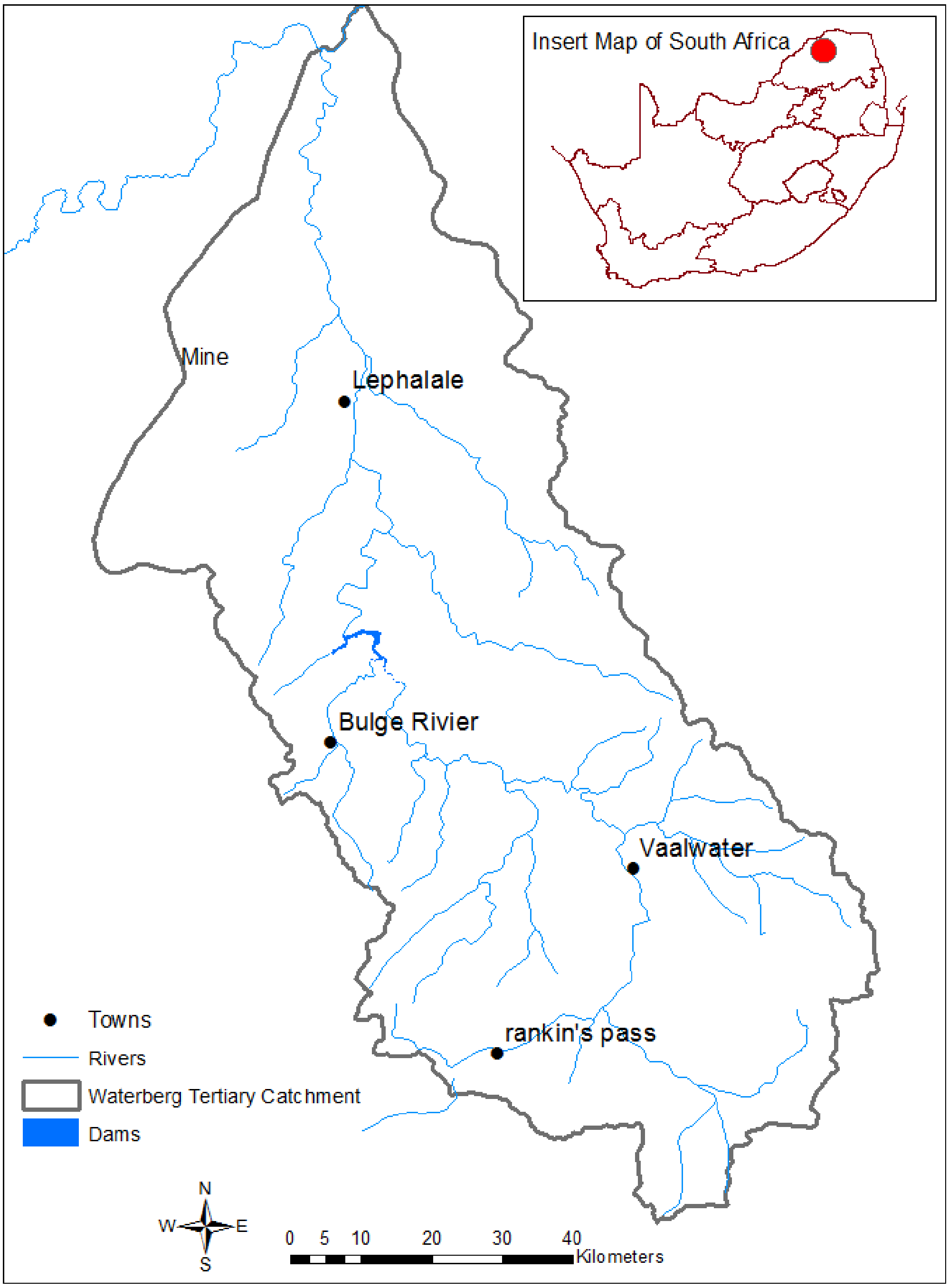


Figure 4.1: Demonstration of the study area map with the rivers, dams and developments such as the mine in the Lephalale area. The insert map shows the red dot which represents where the study area is geographically located in the northern province of South Africa called Limpopo province.

**Geology:** The Waterberg District can be classified into five geological types which are Transvaal Super Group, Waterberg Group, Bushveld Igneous Complex, and the Archaean Granite/Gneiss and the Swazian Complex. Important sources of platinum and chromium are found in the Bushveld Ingneous Complex, while Karoo Super group contains coal deposits. Transvaal Super Group contains iron ore deposits. The Waterberg District Lithology studies show that there are 26 main rock types (WDEMF-Draft Report, 2010).

**Landscape:** The district has unique landscape features distinguishing Lephalale from the rest in the country. It consists of four main landscapes which are Waterberg Plateau, the Transvaal Plateau Basin, the Pietersburg Plain and the Limpopo Depression. The Waterberg Escarpment character is an asset and should be well protected. The key selling point employed by the tourism sector for marketing strategy is the wide open bushveld plains of the Limpopo Peneplain which represent a special South African bushveld character. The slopes are steep and are inherently sensitive to change. The soil types of the area are diverse, and the major soil associations include weakly developed soils on mountains catchments, dystrophic, red and yellow, plinthic upland duplex and paraduplex soils on undulating middleveld, rugged terrain and uplands and rocky areas (WDEMF-Draft Report, 2010).

**Climate:** The mean circulation of the atmosphere over southern Africa is anticyclonic throughout the year. Air circulation has a direct effect of dispersing air pollution and that is because of various reasons. The northern and western regions of the area have a hot and semi-arid climate. The Waterberg District Municipality Air Quality Management Plan provided the information for the atmospheric conditions and wind. There was no measurable evidence of global warming or climate change from the information, due to significant natural fluctuations (WDEMF-Draft Report, 2010).

**Hydrology:** The district is covered by the Limpopo water management area and the Crocodile (West) and Marico Water Management Area. There are five catchments within Waterberg District boundaries which are: Lower Crocodile River Sub-catchment; Mokolo (or Mogol) River Catchment; Lephalale River Catchment; Mogalakwena River Catchment; and a small-portion of the Olifants River Catchment. Most rivers drain in the north-westerly direction to the Limpopo River. The main dams in the Waterberg District Municipality are Mokolo Dam, the Doorndraai Dam, and the Glen Alpine Dam. Rivers are in a fair condition and groundwater is limited, and remains an important resource in the area (WDEMF-Draft Report, 2010).

### 1.7 Guide to chapters

There are six chapters in this dissertation which are chapter 1 with background of the study including problem statement, aims and objectives, research hypotheses, motivation, the study area details and guide to chapters. Chapter 2 is all about literature review, followed by chapter 3 with the methodology of the research. The fourth chapter presents results, then the discussion is under chapter 5, and lastly is chapter 6 elaborates on conclusion and recommendations of the study based on findings made.

# CHAPTER 2: LITERATURE REVIEW

## 2.1. Introduction

### 2.1.1. Importance of N in ecosystem

Global change including climate change and land cover-land use change are postulated to be the drivers of change in vegetation quality and quantity. Vegetation quality and quantity are signals of health (nutrient and pigment content) and productivity of the ecosystems. Plant nutrients such as N concentrations and pigments such as chlorophyll concentration, can act as indicators of condition and water stress in vegetation. Water content, plant nutrient and pigments influence the rate of photosynthetic activities. Litter decomposition, leaf respiration, growth rates and nutrient cycling also act as indicators for ecosystem condition (Field & Mooney, 1986, Ollinger *et al*, 2002). Leaf nitrogen concentration relates to net photosynthesis across various plant species and functional groups, thus represents a link between terrestrial cycles of nitrogen and carbon cycles (Field & Mooney, 1986, Reich *et al*, 1992). It is important to quantify the leaf N concentration to understand the stress levels or the condition of vegetation.

Nitrogen remains one of the most crucial and vital biochemicals that vegetation needs as a major part of proteins and nucleic acids, also helps as a regulator of carbon assimilation in the carbon cycle (Wright *et al*, 2004; Field & Mooney, 1986; Ollinger *et al*, 2002). Studies have been conducted, such as those assessing the availability of N as a key constraint of carbon cycling in terrestrial ecosystem, to consider the role of N in the earth's climate system. Leaf N is a key variable for photosynthesis, and if it is limited plants are stressed.

### 2.1.2. Vegetation stress assessment

Vegetation stress is due to suboptimal conditions which could be caused by global change, leading to the decline in plant physiological functions such as light and dark photosynthesis (Logan *et al*, 2003; Ninements, 2010). Leaf nitrogen (N) levels are associated with photosynthetic capacity and aboveground net primary production (ANPP), thus representation of a simple and meaningful link between terrestrial cycles of N and Carbon cycles. Understanding the spatial variability of leaf N on a landscape, suitable tools and techniques are required. Remote sensing tools can aid in understanding the distribution of N, which can assist in the ecosystem such as assessments of wildlife. Importance of foliar chemical composition cannot be overemphasised, since data about ecosystem processes can be acquired through remote sensing (Wessman *et al,* 1988, Abber & Federrer, 1992). Leaf N relates to maximum photosynthetic rate (Field & Mooney, 1986, Reich *et al*, 1992), and to water availability. Foliar chemistry remote sensing is important as it helps in large scale, spatially explicit estimates of ecosystem's role and vitality.

Leaf N concentration and protein content are related (Clifton et al., 1994; Wang *et al,* 2004) as proteins are some of the major nutrients for the herbivores, so estimating leaf N helps to inform and advise in fields such as agriculture (Prins & Beekman, 1989; Prins & van Langevelde, 2008). Assessing the spatial patterns of leaf N concentrations can assist in effective planning and management of savannah rangelands for sustainable livestock and wildlife grazing. Studies suggest that the availability of leaf N affects canopy spectral reflectance measurements. Nitrogen deficiency leads to decreased chlorophyll content (Moorby & Besford, 1983, Peñuelas *et al*, 1994). Severe N limitation causes plants to reflect more on the spectral region due to lower chlorophyll content. In other words N stress (chlorosis) has an impact on vegetation indices, which influences changes in soil cover, plant density and vegetation colour (Steven *et al,* 1990).

### 2.2. Conventional means for estimating leaf N

The old methods of estimating nitrogen among others are chemical tests like the Kjedahl method, SPAD meter device, and leaf colour chart (LCC). SPAD meter is a hand held and non-destructive device which measures chlorophyll concentration, through leaf transmittance in the red and near-infrared electromagnetic spectrum. It measures chlorophyll by producing transmittance values proportional to leaf chlorophyll amount (Uddling *et al*, 2007), calibration curves are used to convert SPAD meter readings to absolute chlorophyll values (Markwell *et al*, 1995).

Chemical methods include Kjedahl method invented by a chemist named Johan Kjeldal in the year 1883. It helps to quantitatively determine nitrogen in chemical substances. The method was designed to study proteins in malt production, especially for quick and accurate determination of nitrogen content. It is a three steps method which are called: (1) digestion in which nitrogen is decomposed in organic sample by boiling in sulfuric acid to form ammonium sulphate solution, (2) distillation step involves addition of excess base to the acid mixture in which ammonium ($NH_4$) will be converted to ammonia ($NH_3$) and lastly (3)

Titration process is when the amount of ammonia is quantified in the receiving solution, the nitrogen content will be then calculated from ammonia ions (LABCONCO, 2012).

The other method is Leaf colour chart (LCC) which is a colour chart with shades of leaf colour from light green to dark green (King-Brink & Sebranek, 1993). LCC tool is used as an indicator of leaf colour, and has 6 different colour shades from light yellowish green in the first chart to the dark green on the last chart. This tool takes the reading from two weeks after transplanting to initiation of flowering, whereby the colour of the leaf is measured by comparing the leaf colour with the colour of the shades of LCC. The colour of each leaf is measured by holding the tool on the leaf and comparing the colour of the chart with that of the leaf. If the colour of the leaf is between the two charts then the mean of the two values will be calculated. A reading below 4.0 means there is leaf N deficiency while a reading above that means that the leaf or vegetation has N (Regmi, 2006). The other technique is called Dumas method. It measures the total nitrogen gas through combustion, using automated instrument, and has good precision and high throughput of samples (King-Brink and Sebranek, 1993).

### 2.2.1. Point based assessment of leaf N

Plant nutrients such as N were estimated and mapped by time consuming field data collection methods, on small or localized scale basis. Remote sensing combined with ecosystem models, employ one approach to estimate forest ecosystem function on a regional scale (Martin & Aber, 1997). One of the cost effective methods which was applied in estimation of leaf area and feeding damage (herbivory), is the desktop scanner. It is used to estimate the leaf area removed from the low, medium, or high degree of simulated leaf feeding.

This leaf area meter unfortunately overestimate low levels of simulated feeding injury. The method is used with the aid of a desktop scanner and requires two steps: firstly creation of a digital image files, secondly calculating the area represented by the image. Time required to measure leaf impact is shorter than with leaf area meter. It is a less complex and cost-effective method of estimating leaf area and feeding damage. It also helps in some experiments where pre-feeding measurement of the leaf is either challenging or undesirable, or when there are low amounts of feeding (O’neal *et al*, 2002).

### 2.2.2. Chemical analysis, using several methods

Chemical analysis methods used for foliar estimations, are slow, expensive and challenging to map or estimate over large geographic areas (Curran, 1989). Nutrient cycling measurement and process such as photosynthesis are important for assessment of exchange of greenhouse gasses between soil, vegetation and atmosphere (Mooney *et al*, 1987, Steudler *et al*, 1989, Worfsy *et al*, 1993). The Kjedahl method is one of the most accurate and unfortunately cumbersome, as it may take a week for a total of 72 samples to be analysed. A method such as this is not suitable if the aim is to rapidly estimate nitrogen concentration in large areas.

The other chemical based method is the SPAD meter (also called chlorophyll meter) for chlorophyll measurement, using colour chart with shades, with leaf colours from light green to dark green (Auearunyawat *et al*, 2012). The common thing between SPAD meter and LCC is that they help to adjust fertilizers N when plants have a deficiency of N (Balasubramanian *et al*, 1999). SPAD meter and LCC are not used simultaneously as each device can be used independently for similar or different purposes. Islam *et al*, 2009 correlated the results from both devices to study effect of change and impact of environmental parameter. Radiochemical methods are also the other N estimation techniques, used for analysis of nitrogen and protein, but require costly instrumentation (Pomeranz & Moore, 1975).

Old methods can be exclusive and specific for a particular point of interest, especially in the cases where the aim of the study is not to cover larger areas. These methods are labour intensive and consume time, but results obtained can help in small scale studies that are not intended for large areas. These methods can give reliable results which can also be comparable and give similar outcomes just like advanced methods, as long as correct procedures, sampling, tests, estimations or quantifications are conducted. The downside of methods such as the chemical based, are that they can be laborious and destructive at times to natural flora. The other factor is the time consuming nature of the conventional methods which often negatively affects the time to make an informed decision. For large scale application for assessing leaf N, the conventional methods are just not suitable, as sampling, because of the above factors.

## 2.3. Estimation of leaf N using remote sensing

### 2.3.1. Use of vegetation indices

Vegetation index can be defined as an indicator that describes the greenness, thus the relative density and health of vegetation for each picture element/pixel in satellite (Tucker, 1979). The common technique for estimating vegetation parameters is the use of vegetation indices as predictors. NDVI and SR indices based on hyperspectral data computed from red-edge bands, provide accurate estimates of leaf N compared to conventional NDVI derived from 680nm and 800nm (Mutanga and Skidmore, 2007; Ramoelo *et al*, 2012). NDVI is a technique that employs numerical indicator through the aid of visible and near infrared bands of the electromagnetic spectrum to observe the greenness of the target (Tucker, 1979).

Successful estimation of leaf N using vegetation indices such as red edge position, depends mainly on chlorophyll concentration (Clevers *et al*, 2002; Mutanga *et al*, 2004; Cho and Skidmore, 2006; Numata *et al*, 2008), this assumes a positive correlation between leaf N and chlorophyll concentrations (Vos and Bom, 1993; Yoder and Pettigrew-Crosby, 1995). This method is limited because the dependence on the plant phenology, meaning that the relationship will deteriorate as leaves senesce (Wang et al., 2009).

### 2.3.2. Use of known absorption features to estimate leaf N

Electromagnetic spectra have absorption features which are known specific regions, linked to electron transition or physical bond vibrations of the specific foliar biochemical concentrations (Darvishzadeh *et al*, 2008; Knox *et al*, 2011). The electron transitions in chlorophyll (400-700 nm) and O-H bond in water stretch & bend for absorption features to be realised (Osborne & Fearn, 1986; Williams & Norris, 1987). A better understanding of absorption features lies in the biochemistry of plants as they are made up of hydrogen (H), carbon (C), oxygen (O), nitrogen (N). In other words plants absorption bands exist because of vibrations of bonds that are presented in this manner: C-O, O-H, C-H and N-H together with other vibrations and overtones (Curran, 1989). Absorption features are studied through hyperspectral remote sensing or imaging spectroscopy and are closely linked to plant nutrients such as leaf N.

Determining the wavelength and spectral features related to biochemicals of interest helps to achieve known absorption features. Hyperspectral remote sensing technique estimates leaf N by using spectral absorption features situated in the near-infrared (NIR) and shortwave infrared (SWIR). Absorption features are studied through near infra-red spectroscopy (NIRS) by selecting wavelengths and spectral features, which relates to any biochemical of interest such as leaf-N (Card *et al*, 1988, Curran *et al*, 1992, Grossman *et al*, 1996). For this study absorption features of leaf N are centred at wavelengths: 430 nm, 460 nm, 640 nm, 660 nm, 910 nm, 1510 nm, 1940 nm, 2060 nm, 2180 nm, 2300 nm and dominate in the SWIR (Curran, 1989).

Methods that only use red and near-infrared bands are only sensitive to leaf pigments that have strong absorption differences at either of the two sides of the red edge/710 nm band. The disadvantage is that the technique becomes limited especially when estimating concentrations of leaf nutrients, as they have many absorption bands lying outside the red and near-infrared region (Dixit & Ram, 1985; Shah *et. al*, 1990; Tsai & Phillip, 1998; Wessman, 1989). Leaf N estimation using NIR and SWIR may be inaccurate because of reflectance of leaf water content which masks the absorption features of biochemicals (Gao & Goetz, 1994, 1995; Fourty & Baret, 1998).

### 2.3.3. Use of full spectra to estimating leaf N

High spectral resolution sensors can help to study the detection and mapping of foliar chemistry and vegetation stress. These sensors are based on resolution specifications such as narrow channels sensors, thus fewer than 2 nm bandwidths. Spectral resolution is the manner in which spectral spaces are divided in the number and range of wavelength, spectral breath of each wavelength sample, including the number and contiguous nature of sampled wavelengths (Mutanga *et al*, 2009). Full spectral based leaf N estimation can be achieved by using hundreds of bands in hyperspectral study, using techniques like stepwise multiple linear regressions. Such techniques can lessen the dimensionality of the data, to maximize leaf N estimation. The limiting factor of full spectrum, when combined with SMLR is overfitting and multicollinearity (Curran, 1989; Martens & Naes, 2001), these two are always higher in dimensionality of the full spectrum data (Curran, 1989).

The other technique which minimises limiting factors such as labour and time consumption is Partial least square regression (PLSR). This method combines the most useful information from hundreds of bands into the first several factors, while less important factors may likely include background effects (Bolster *et al*, 1996). The method also reduces background effects and avoids the potential over-fitting challenges associated with SMLR. Technical challenges such as different scattering effects occur because of sample differences such as additive offsets (baseline shifts) and multiplicative effects (Datt, 1998). These can be accounted for and corrected before any statistical models are used to reduce the background effect (Bolster *et al*, 1996). Ramoelo (2012) used full spectrum technique to estimate the nitrogen (N) to phosphorus (P) ratio using multivariate technique of partial least squares regression (PLSR), coupled with continuum removal technique. The findings revealed that N: P ratio was successfully estimated using field spectra and partial least square regression.

### 2.3.4. Use of integrated modelling approach (combining RS and environmental parameters) to estimate leaf N

Integrated modelling approach estimates biochemical concentration by combining environmental variables such as climate, topography with *in situ* hyperspectral variables (Ramoelo *et al*, 2011). A study by Ramoelo *et al*, (2011) tested the performance of the non-linear partial least squares regression (PLSR) to predict grass N and P concentrations. The study highlighted that when non-linear partial least squares regression (PLSR), is integrated *in situ* hyperspectral with environmental variables; there is improvement in grass nitrogen (N) and Phosphorus (P) estimation accuracy. This is better than only using remote sensing variables or conventional PLSR.

The other study entails two-step method first, using vegetation indices and second integration of vegetation indices with environmental variables through SMLR and non-linear partial least squares regression PLSR. This research was pursued because there were fewer studies focusing on the leaf biochemical concentration estimation at a regional scale, using integration of environmental and remote sensing variables (Cho *et al,* 2009; Cho *et al,* 2010; Knox *et al,* 2011, Ramoelo *et al,* 2011; Ramoelo, 2012). Ramoelo, (2012) found that altitude combined with red-edge based vegetation indices were significant in estimating leaf N. Knox *et al,* (2012) demonstrated that combining absorption features and ecological or environmental parameters improves estimation of leaf N. The challenge in using this technique is that environmental variables are often unavailable in usable accuracies.

## 2.4 Statistical methods

### 2.4.1. Univariate statistical analysis based on vegetation indices

Simple linear regression or univariate is used to study the relationship between two variables, one being the dependant variable while the other is an independent or predictor variable. This type of regression fits a straight line plotted on a graph called scatter plot, to predict the outcome between the dependant (leaf N) and independent variable (either wave-bands or vegetation indices) (Dowdy *et al*, 2004). Use of vegetation indices were mainly based on simple regression, which determines the relationship between leaf N and indices or wave bands, for this study leaf N is the dependant variable while either waveband or vegetation indices are independent variables.

### 2.4.2 Multivariate estimation of leaf N

Researchers usually have to choose an appropriate statistical technique when conducting quantitative research. The correct choice depends on an accurate and appropriate research question (Metler, & Vanata, 2002). Simultaneous analysis of independent variables with dependent variables with the help of matrix algebra describes multivariate regression in simple terms (Dowdy *et al*, 2004). Multivariate statistical method is applied when several measurements are made on each individual in one or more samples. These techniques have been applied in fields such as biological sciences, geology, mining and many others. For example SMLR has been successful in biochemical estimation such as leaf N; however it has disadvantages such as over-fitting and multicolinearity (Curran 1989; Martens & Naes 2001) and has challenges when transferring predictive models to other data sets, or site (Grossman *et al*, 1996).

**Stepwise multiple linear regression (SMLR) and partial least square regression (PLSR)**

Foliar biochemical concentrations are important indicators of the ecosystem processes. Studies have shown that remote sensing arguably offers the only practical solution when compared to chemical methods used for foliar estimations. Such methods are slow, expensive and pose challenges when mapping or estimating over large geographical areas (Curran, 1989). In other words data collected through remote sensing provides hundreds or thousands of bands within visible to near-infrared wavelengths to identify many subtle absorption features attributable to a wide range of chemicals. Leaf N is an important indicator of photosynthetic rate and overall nutritional status (Curran, 1989; Field & Mooney, 1986). This has been observed after many spectroscopic studies.

SMLR is usually used to estimate biochemical concentrations although it has disadvantages of over-fitting and multicolinearity (Curran, 1989; Grossman *et al*, 1996). PLSR method on the other hand combines the most useful information from hundreds of bands into the first several factors, while less important factors may likely include background effects (Bolster *et al*, 1996; Atzberger *et al*, 2003). PLSR reduces background effects and avoids the potential over-fitting challenges associated with SMLR. Technical challenges including different scattering effects occurring because of sample differences such as additive offsets (baseline shifts) and multiplicative effects (Datt, 1998), are considered and corrected before using any statistical models. PLSR appears to be a good technique to estimate N, most probably due to its predictive power (Ramoelo, 2012; Huang et al., 2004). Several studies used multivariate techniques for estimation of leaf N with success (Ramoelo *et al*, 2011; 2013; Huang *et al*, 2004). Examples of these studies include leaf N estimation in the savanna environments (Ramoelo *et al.*, 2012; 2013); leaf N estimation in forest (Cho *et al.*, 2010); and leaf N estimation in crops (Huang *et al.*, 2004; Habounde *et al.*, 2002)

**The use of machine learning techniques (Random forests)**

Random forest is an ensemble (combination of results from different models) classification method that uses many tree models, in a regression or classification mode (Breiman, 2001). Mutanga *et al*, 2012 used random forest regression and SMLR to predict biomass estimation for wetland vegetation through WorldView-2 imagery. Random forest regression the better predictor of wetland biomass with a root mean square error prediction (RMSEP) of 0.0441

$kg/m^2$ (a better observed mean biomass of 12.9 %), compared to that of SMLR which was root mean square error prediction (RMSEP) of 0.5465 $kg/m^2$ (a poor observed mean biomass of 15.9 %). The other study which used random forest to estimate leaf area index (LAI) was conducted by Vuolo *et al*, (2013). LAI was investigated from the two agricultural areas one in Italy while the other was in Austria. The random forest regression mode results were as follows: the Italian agricultural area (RMSE= 0.502 and $R^{2=}$ 0.82) which had lower errors when compared with the Austrian agricultural area (RMSE= 0.860 and $R^{2=}$ 0.017).

Random forest regression mode was again used to estimate canopy height in French Guiana with ICESst/GLAS data. The result revealed the random forest regressions were better compared with linear models. The relationship between GLAS metrics and canopy heights is not really linear; this might have affected the linear models results. Random forest regressions had RMSE of 3.4, thus the best configuration for canopy height estimation at all metrics used; and also showed a slight improvement in canopy height estimation with RMSE of 3.6 (Fayad *et al*, 2014).

The last study was by Abdel-Rahman *et al*, (2013) using random forest regression and spectral band selection to estimate sugarcane leaf nitrogen concentration. In this study random forest regression algorithm was tested for a potential of selecting necessary spectral features in hyperspectral data to predict leaf N concentration. The findings of random forest were not so good compared to those of SMLR with R= 0.67, root mean square error validation RMSEV= 0.67 and a mean of 8.44%. This could have been affected by the lower parameters settings as the data here was for site specific applications which can help in the field of precision farming (Abdel-Rahman *et al*, 2013).

# CHAPTER 3: METHODOLOGY

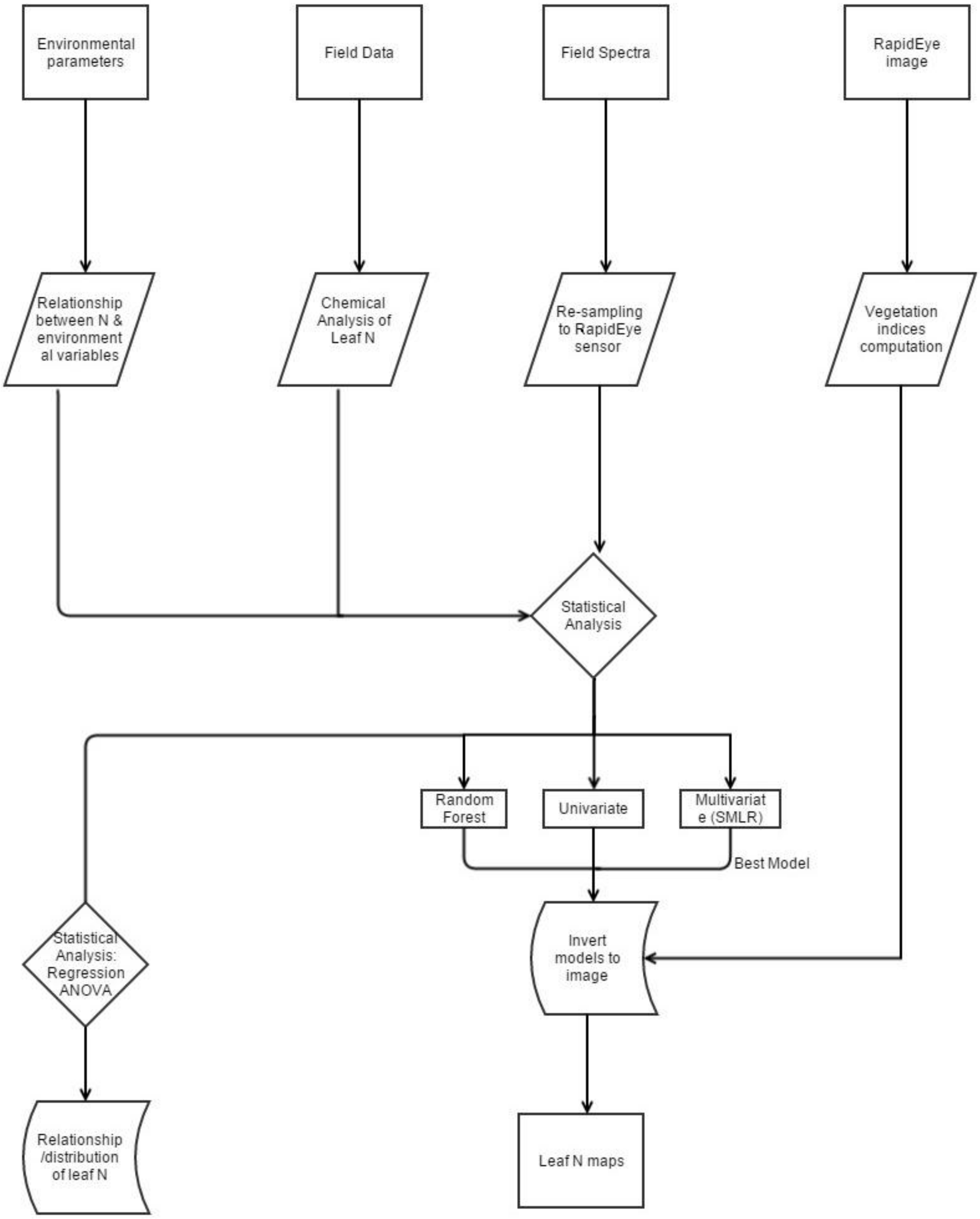


Figure 3.1: Flow chart of the entire sections of the study process.

## 3.1 Data Requirements

### 3.1.1. Field data collection

- **Sample collection using road sampling**

The field data was collected in December 2011 with the satellite overpass. It was mainly road sampling or road side sampling, as this was the accessible site to collect field data in the study area. The sampling approach was purposive in nature, since the random method could not be suitable because of the restricted access due to impermeable fences. The limitation of this approach could relate to the number of sample to be collected and the variability of nitrogen. Since the roads were prior selected covering various slopes and geological types, the desired variability was expected to be achieved. About 5 leaves around the canopy of the tree were collected, to ensure full canopy coverage. For the grass samples, a plot of 20 x 20m was used. Two to three subplots of 0.5m x 0.5m were randomly placed and within each subplot, grass samples were cut.

- **Spectral measurements**

An Analytical Spectral Device (ASD) (FieldSpec 3) was used for spectral measurement in the field for each point visited, and this was done with satellite overpass. Leaf samples collected or harvested from the trees, spectral measurements were collected, and later averaged for each tree canopy. For the grass, canopy reflectance was measured in each subplot, and five measurements were made and averaged at a later stage. Using the spectral response function for RapidEye, the data collected was re-sampled to RapidEye spectral configurations. The re-sampled spectra were used for further analysis.

- **Chemical analysis**

Leaf samples were dried at 80 $^{o}$C for at least 24 hours at the laboratory to remove moisture and water content, while preserving the nutrient content. The samples were then taken to Bemlab laboratories for chemical analysis, and leaf N values were extracted using a Leco FP528 nitrogen analyser (Horneck and Miller, 1998). The leaf N concentration was chemically analysed, and the unit of measurement is percentage of dry matter (%).

### 3.1.2. Satellite data – RapidEye

The RapidEye multispectral data was already collected and available for use on this study. This satellite has a sensor type described as a multi-spectral push broom imager. Spectral bands of the satellite are as follows: sensor contains five spectral bands in the visible to near-infrared (VNIR) including red edge centred at 710 nm. It is a multispectral imager with spatial resolution of 6.25 m, and samples light in the spectral bands which are: blue (440-550 nm), green (520-590 nm), red (630-685 nm), red edge (690-730 nm), and near infrared (760-850 nm) (RapidEye, 2010). The ground sampling distance is 6.5 m, with a pixel size of 5 m, and has a swath width of 77 km, while the camera dynamic range of 12 bit and an image capture capacity of 5 million square km/day (RapidEye, 2009; Blackbridge, 2013). Images of RapidEye satellite were collected in December 2011, in order to predict the vegetation status at tree scale. Tree canopy was captured through the RapidEye Ortho product which was acquired at 5 m x 5 m re-sampled spatial resolution.

### 3.1.3. Environmental data or variables

Spatial variability and distribution of leaf N on a landscape and its interaction with environmental data was analysed. The environmental data included soil, aspect, vegetation type, Digital Elevation Model (DEM) (altitude) and slope.

- **Vegetation type**

The vegetation map was acquired from South African National Botanical Institute (SANBI), and the vegetation types used for understanding the distribution of leaf N were: Limpopo Sweet Bushveld, Subtropical Alluvial Vegetation, Central Sandy Bushveld, Waterberg Mountain Bushveld, Western Sandy Bushveld and Roodeberg Bushveld. Table 4.1 below describes the vegetation types studied.

Table 3.1: Vegetation types's description information sourced from Mucina & Rutherford, 2010

| **Vegetation type** | **Vegetation & landscape features** | **Geology & soils** | **Climate, distribution & taxa** |
|---|---|---|---|
| Limpopo Sweet Bushveld (SVcb 19) | The landscape of vegetation are plains, sometimes undulating or irregular, also on riparian areas such as tributaries of Limpopo river, and vegetation described as short open woodland. | Geology is mainly gneisses, metasediments, metavolcanics, basalts, sandstone, siltstone and mudstone. Soils have calcrete & surface layers with clayey-loamy form and black clayey soils. | Hot and wet season from November to April. Maximum temperature +/- 38 $^{o}$C. Common taxa are *Acacia robusta* (tall trees), *Acacia tenuispina* (low shrubs). |
| Subtropical Alluvial Vegetation (AZa 7) | The vegetation and landscape features described as riverine (relate to a river) terraces which supports intricate complex or macrophytic vegetation (in river flowing channels and river-fed pans). Includes highly flooded grassland, short lived herb land and riverine thickets. | Found in deep fine structured sandy to loamy soils, usually water logged and prone to floods during rainy seasons. Has higher salt accumulation due to higher evaporation. This vegetation types are mainly found on channels of flowing river or river-fed pans, and in areas were water flows slowly. | Hot and wet season from November to April. Subtropical seasonal summer rainfall climate conditions with temperatures up to 22 $^{o}$C. Common taxa: *Acacia Natalitia* (small trees); *Justicia flava* (low shrubs); *Salvadora angustifolia* (tall shrubs). |

| | | | |
|---|---|---|---|
| Central Sandy Bushveld (SVcb 12) | Common in low undulating areas, sandy plains, supports tall and deciduous woodland on deep sandy soils. Also on lower slopes on eutrophic sands, and less sandy soils. Dominated by grassy herbaceous layer and low basal cover on dystrophic (inadequate nutrition disorder) sands. | In areas underlain by granite and granophyres (fine grained rock) rocks. Soils dominating are sandstone, conglomerate and siltstone. | Hot and wet season from November to April. Maximum temperature +/- 35.3 $^{o}$C. Taxa such as *Combretum Hereroense* (tall shrubs), *Acacia burkei* (tall trees). |
| Waterberg Mountain Bushveld (SVcb 17) | Commonly found in rocky mountainous areas. Vegetation characterised by bushveld on higher slopes & broad leaved deciduous bushveld. Grass layer is either moderately or well developed. | Dominated by sandstone, subordinate conglomerate, siltstone and shale including medium to coarse-grained sandstone. Common soils features are sandy, loamy to gravely and dystrophic. | Hot and wet season from November to April. Maximum temperature +/- 35.3 $^{o}$C. Common taxa are *Acacia robusta* (tall trees), *Acacia robusta* (small trees) |
| Western Sandy Bushveld (SVcb 16) | Vegetation type varying from tall open woodland to low woodland, broad-leaved which includes microphylous tree species. Some vegetation found in shallow soil of gravelly upland sites and | Geology of sandstone and mud-stone; siltstone and shale. Soils are mainly plinthic catena, eutrophic, red-yellow apedal which are freely drained with | Hot and wet season from November to April. Maximum temperature +/- 35.3 $^{o}$C. Taxa such as *Combretum apiculatum* (small trees), *Terminalia* |

| | | | |
|---|---|---|---|
| | deep sands occurring in slightly undulating plains. | high base status. | *sericea, Acacia burkei* (tall trees) are common under this vegetation type. |
| Roodeberg Bushveld (SVcb 18) | Found in landscapes such as plains, low hills, with closed woodland to tall open woodland. It is characterised by vegetation of tall open woodland and poorly developed grass layer. | Geology is composed of sandstone, conglomerate, siltstone and shale, while the vegetation is found in sandy soil with red yellow apedal status. | Hot and wet season from November to April. Maximum temperature +/- 37.1 $^{o}$C. Common taxa include *Acacia burkei* (tall trees); *Dichostachys cineria* (tall shrubs); *Commiphora africana* (low shrubs). |

- **Soils**

The soil types studied map was acquired from ARC through SOTER database and these soil types were: Rhodic Lixisols, Rubic Arenosols, Chromic Acrisols, Ferric Luvisols and Eutric Arenosols. Each soil type was assessed for their effect on leaf N distribution, using one-way Analysis of Variance (ANOVA). This was to test whether the soil types affects leaf N distribution by checking if the mean values of leaf N among soil types are equal or not. To determine if the soil types varied significantly in spatial distribution of leaf N distribution a criteria was used. The criteria was the p-value < 0.05 means that leaf N distribution varies significantly among the soil types, thus soil types affects the distribution of leaf N. The p-value > 0.05 means leaf N distribution does not vary significantly across soil types meaning leaf N distribution is not affected by the soil types. A box-plot (see Figure 5.4) of all the soil types helped to find out if the mean of soil types varied or was similar.

**DEM – SRTM 90 m**

Digital Elevation Model – Shuttle Radar Topography Mission with a resolution of 90 meters (DEM – STRM 90 m) was also used to determine whether altitude influences leaf N distribution or not. DEM is a digital representation of cartographic (the practice of making maps) information in a raster form, this model has elevations for a number of ground positions at regularly spaced intervals (Nijmeijer *et al*, 2001). DEM is called a model because computer language uses the topographical data (map) to automatically analyze the study area in 3 dimensions, instead of the cumbersome and laborious human interpretation (Das, 2013). The STRM data was downloaded from Http://glovis.usgs.gov.

- **Slope**

Slope which is the gradient (steepness) of a unit terrain was computed from the DEM using ArcGIS 10.2.1 software (in degrees). Slope data was used to find out if it affects leaf N distribution or not. DEM in a raster format has elevation of individual cells, and slope as a first derivative of DEM represents elevation change. Calculation of slope was done in ArcGIS (through ArcToolbox or Spatial Analyst toolbar) as a raster with slope value for every cell presented in degrees for this study (Burrough & McDonell, 1998; ArcGIS 10.2.1, 1999-2013).

For each cell the slope tool in ArcGIS computes the slope by calculating the maximum rate of change in value from that cell to its neighbours. The maximum change in elevation over the distance between the cell and its eight neighbours will identify the steepest downhill descent from the cell. The calculation is done by fitting a plane on the data points called z-values of a 3 x 3 cell neighbourhood around the centre cell. An average maximum technique is utilised whereby; for every individual cell in the centre of 3 x 3 windows, the slope value calculation is based on the rate of change of the surface horizontally and vertically around the centre cell. For this study the output slope raster range of slope values was in degrees which are between 0 and 90 degrees (Burrough & McDonell, 1998; ArcGIS 10.2.1, 1999-2013).

- **Aspect**

Aspect describes the direction of the slope and was computed from the DEM, as well. In other words the value of the output raster here is purely about the direction in which the slope faces in a clockwise direction. The input raster is either mosaicked or clipped or otherwise-prepared DEM, for example a hillside facing east has an eastern aspect. Calculating aspect is based on a concept of the down slope direction as a result of maximum rate of change of value from each cell to its neighbours. A plane is fitted on data points called z-values of 3 x 3 cell neighbourhood around the centre cell, so the direction that the plane faces will represent the aspect, while flat areas are given a value of -1. The values of aspect in this study are measured in degrees from 0 to 360 (Burrough & McDonell, 1998; Das, 2013; ArcGIS 10.2.1, 1999-2013). Aspect was tested for its impact on Leaf N distribution and concentration.

### 3.2 Data Analysis

For each point visited in the field, the reflectance on the Rapid Eye image was extracted for data analysis. NDVI, SAVI and SR were derived using the red and near infrared bands. The aim of selecting the best predictive model is to invert it to the RapidEye imagery to derive a leaf N map. The statistics was done using SPSS and R programming language.

#### 3.2.1 Univariate – leaf N vs Vegetation indices

Vegetation indices are useful to assess plant stress or health, and also to enhance vegetation greenness signal, while minimizing solar irradiance and soil background effects. Vegetation indices offers a better option than solar irradiance which depends on time and atmospheric conditions, because a simple light reflection measurement of objects is insufficient to accurately estimate biochemical such as leaf N (Jackson & Huette, 1991). A combination of data from two or more spectral bands creates a vegetation index. For best results the use of vegetation indices needs knowledge of input variable units to form indices including external environment, architectural aspects of vegetation canopy effects (Jackson & Huette, 1991). Univariate or simple regression technique was used to determine which vegetation indices correlated with leaf N well. The other important point to note is that the univariate regression displayed the effect of red edge in the estimation of leaf N, thus to see if the estimation accuracy is improved or not.

#### 3.2.2 Multivariate analysis

- **SMLR**

Leaf N prediction models were studied using multivariate regression technique known as stepwise multiple linear regressions (SMLR). To test the applicability of all five bands in combination with best performing vegetation index, and to test effect of red edge band whether it improves leaf N estimation accuracy or not. The selection of the best model with specific important variables to estimate leaf N was done using the lowest Akaike Information Criterion (AIC) (Sakamoto *et al*, 1986). The process was implemented in R statistical programming language or technique.

- **Random Forest (RF) – implemented in rattle/ R**

Random forest was implemented in rattle software for leaf N prediction models as a multivariate regression technique, to test the applicability of all five bands in combination with best performing vegetation index and also to test effect of the red edge band. Breiman, (2001) developed random forest to improve regression trees by combining a large set of decision trees, whereby each tree is built by selecting a random set of variables and a random sample from training dataset.

This is the workflow used in this study: different training data subsets are selected (about 2 or 3 subsets) with replacement to train each tree; the remainder of the training data is used for estimation of error and variable importance. The number of trees from all trees and for regression will then makes class assignment, and then the average of the results is used. Randomly selected subset of variables are used to split every individual node whereby the user decides the number of variables to be used (Breiman, 2001; Horning, 2010). One of the characteristics of variables subsets is that small subsets produces less correlation thus lower error rate, and at the same time low predictive power has high error rate as well, so preferable value range is often wider (Horning, 2010). For this study number of trees (*ntree*) was a default value of 500, and the number of different predictors tested at each node (*mtry*) was a default value of 1.

Random forest has common variables such as number of trees (*ntree*), input data (predictor and response), and number of variables to use at each split, error calculation & variable significance information, sampling with or without replacement (Breiman, 2001; Horning, 2010). One of the benefits from random forest is the measurement of the frequency of the unique pairs of training samples will subsequently be in the terminal mode, this is called proximity. Proximity is important to fill in lost or missing data and also for calculation of outliers. Advantages of random forest are quite profound they include easy setting of parameters, less sensitivity to outliers in training data, automatic generation accuracy and variable importance and the fact that overfitting is not an issue at all (Horning, 2010). The disadvantages exist as well, firstly regression is unable to predict beyond range in the training data, secondly in the case of extreme values of regression prediction is not desirably accurate as higher values are underestimated and lower values are overestimated. Although there are shortfalls and benefits, random forest remains vital in earth observation applications as it is utilised in regression (Horning, 2010).

- **Selection of the best model and validation**

The models were validated by bootstrapping statistical method implemented in R statistical programming language. Bootstrapping is an unbiased way of validating models by drawing many independent bootstrap samples to evaluate corresponding bootstrapped replications and most importantly estimating the standard error. It is described as a non-parametric method which repetitively re-samples the sample data in order to validate specific characteristics of a population (Fox, 1997). It is named bootstrapping derived from the "expression of pulling oneself by the boot straps", the analogy used is that: "the population is to sample as the sample is to the bootstrap samples" (Efron, 1979). In other words the method creates a sort of a pseudo-data from the sample at hand, to explore the regression parameters variability so that the uncertainty in the estimated standard errors can be calculated (Freedman & Peters, 1984).

Models were bootstrapped by calculating the RMSE between the predicted and the measured values of the regression results. It is implemented in a computer through non-parametric or parametric maximum likelihood, by allowing computation of maximum likelihood estimates of standard errors (Efron & Tibshirani, 1993). The bootstrapping results validated the uncertainty of the root mean square errors from the study which was not much, meaning that the error rate of regression results was acceptable. The best prediction model was therefore chosen based on high coefficient of determination ($R^2$) and a lowest root mean square error (RMSE).

### 3.2.3 Influence of environmental variables on leaf N distribution

To investigate how leaf N distribution is influenced by the environmental parameters such as soils, vegetation types, digital elevation model (DEM), slope and aspect, basic statistical analysis were employed. For the categorical variables such as soil and vegetation types, one way analysis of variance (ANOVA) was used, based on the 95% confidence level ($p$-$value$<0.05). For continuous variables such as altitude, slope and aspect, simple regression was used and leaf N was always put as a dependant variable, and it was also done with 95% confidence level ($p$<0.05).

# CHAPTER 4: RESULTS

This section present results of the univariate analysis based on simple regression and the multivariate analysis based on stepwise multiple linear regression (SMLR) and random forest. The results also include the impact of environmental parameters on the spatial distribution of leaf N, including leaf N maps created from the best regression models.

### 4.1. Descriptive statistics

Leaf N in plants showed to be relatively higher with a mean of 1.78 %. The minimum and maximum values of leaf N were 0.93 and 4.18 % respectively; the lowest values can be associated with grass and the higher ones to trees. The variability of leaf N was high as demonstrated by the coefficient of variance (CV) of 33.91 %. This is because of the grass and tree leaf N values are combined, which present an interesting variation.

### 4.2. Univariate analysis results: leaf N vs various vegetation indices

The results of univariate analysis are presented in Table 4.1. Univariate statistical method showed that the red edge based vegetation index called MTCI yielded the best results in predicting leaf N ($R^2$=0.1454, RMSE= 0.5625). This highlighted the importance of red edge on improving leaf N estimation, and the second best predictor was red edge or 710 nm band. The third best predictor of N was SR4 computed by bands such as NIR and the red band while the fourth was RE_NDVI (based on band 710 nm and 805 nm). The last one was SAVI (based on these bands, 805 nm near infrared band, 657 nm band, and L which is the soil brightness correction factor) (Qi *et al*, 1994). These top five bands demonstrates the positive effect of red edge in estimating leaf N. Figure 4.1 illustrates the results of simple regression as, with scatter plots of the top five vegetation indices and wave bands for estimating leaf N, thus the best vegetation indices and wave bands correlating with leaf N well. Generally, the relationship between leaf N and vegetation indices is poor ($R^2$ <0.20). All top five performing vegetation indices are arranged in order of the best performing to the least performing vegetation index (1 as the best and 5 as the least perfoming vegetation index).

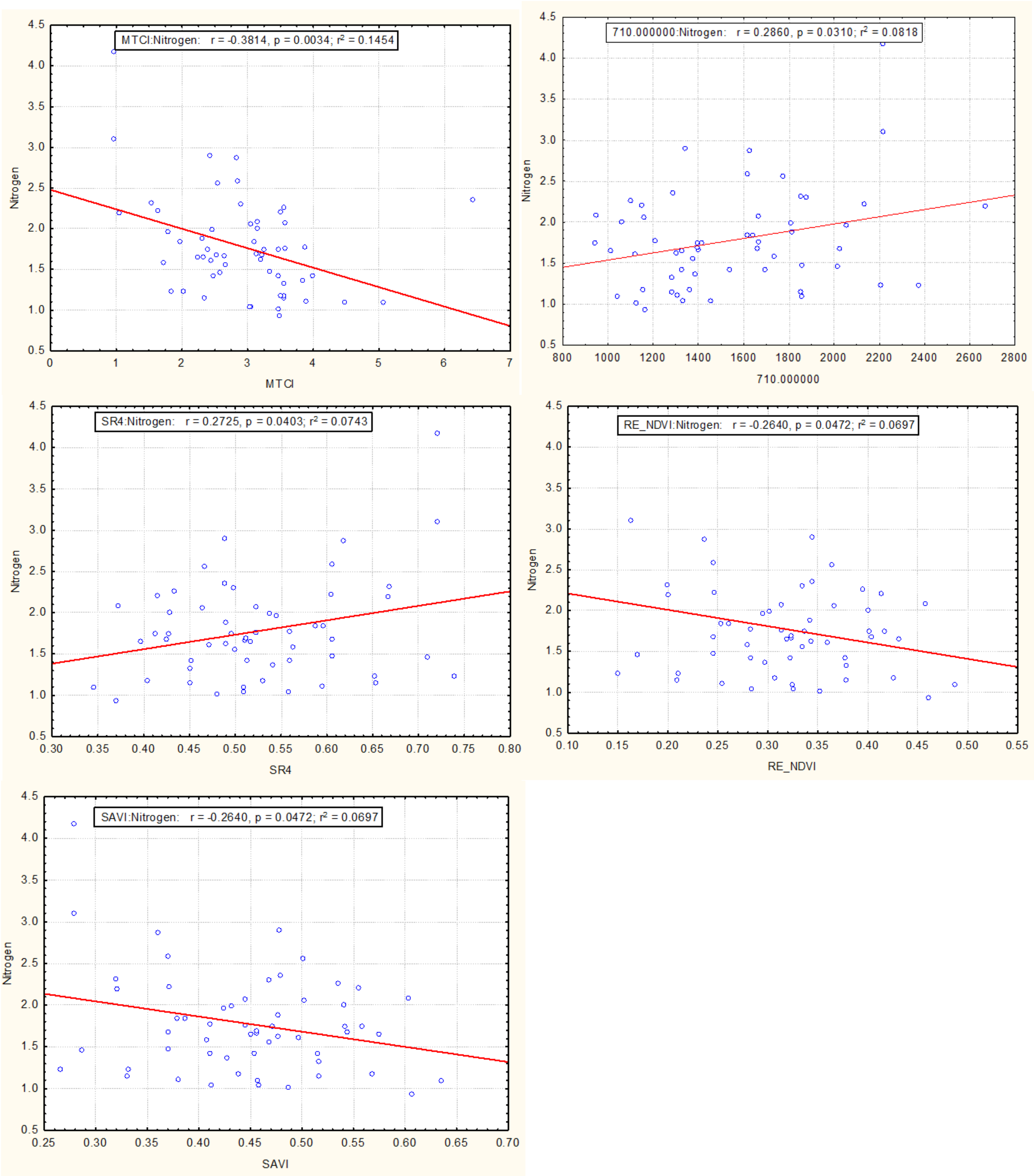


Figure 4.1: Scatter plots of the top five vegetation indices (X-axis) against leaf N (Y-axis) when using univariate or simple linear regression to determine vegetation indices which correlate with leaf N, based on the highest coefficient of determination ($R^2$) criteria. These top five vegetation indices/ wave bands are: 1. MTCI= MERIS Terrestrial Chlorophyll Index,

2. 710 nm or red edge wave band, 3. SR4= Simple Ratio, 4. RE_NDVI= Red Edge_ Normalised Difference Vegetation Index and 5. SAVI= Soil Adjusted Vegetation Index.

Table 4.1: Tabulated results of univariate regression for each band or vegetation index against leaf Nitrogen, a value of $R^2$ that is closer to 1 means the N-estimation model is good while a value closer to 0 means the model is not good, while a p-value that is less than 0.05 also means the N-estimation is good which is termed 95 % significance level. The vegetation indices and wave bands are: 475= Blue, 555= Green, 657= Red, 710 nm= Red-edge 805= Near Infra-Red, NDVI= Normalised Difference Vegetation Index, RE_NDVI= Red Edge_ Normalised Difference Vegetation Index, SR= Simple Ratio, RE_SR= Red Edge_ Simple Ratio, MTCI= MERIS Terrestrial Chlorophyll Index, GI= Green index, RGI= Red/ Green Index, RGI1= Red/ Green Index, BGI= Blue Green Pigment Index, BRI= Blue Red Pigment Index, GRR= Green-Red Ratio, NGRR= Normalized Green-Red Reflectance, NGRR1= Normalized Green-Red Ratio, SR3= Simple Ratio, SR4= Simple Ratio, DVI= Difference Vegetation Index, DVI1= Difference Vegetation Index, SIPI= Structural Insensitive Pigment Index, SIPI1= Structural Insensitive Pigment Index, EVI= Enhanced Vegetation Index, NRI= Nitrogen Reflectance Index, SAVI= Soil Adjusted Vegetation Index, and SAVI1= Soil Adjusted Vegetation Index.

| Variable | $R^2$ | P-value |
|---|---:|---:|
| **MTCI** | 0.14542 | 0.003424 |
| **710** | 0.0818 | 0.031035 |
| **SR4** | 0.0743 | 0.040277 |
| **RE_NDVI** | 0.0697 | 0.047232 |
| **SAVI** | 0.0697 | 0.47234 |
| **RE_SR** | 0.0622 | 0.061359 |
| **SIPI1** | 0.0444 | 0.115372 |
| **555** | 0.0394 | 0.138657 |
| **DVI** | 0.0229 | 0.261675 |
| **805** | 0.017 | 0.333792 |
| **475** | 0.0163 | 0.342947 |
| **RGI** | 0.0144 | 0.373657 |

| | | |
|---|---|---|
| **RGI1** | 0.0133 | 0.391837 |
| **NGRR** | 0.012 | 0.416855 |
| **NRI** | 0.012 | 0.416855 |
| **BRI** | 0.0096 | 0.462457 |
| **SRPI** | 0.0096 | 0.462457 |
| **GI** | 0.0086 | 0.494438 |
| **GRR** | 0.0085 | 0.494438 |
| **657** | 0.0073 | 0.527907 |
| **DVI1** | 0.0058 | 0.574216 |
| **SR3** | 0.0046 | 0.874325 |
| **NGRR1** | 0.0041 | 0.637267 |
| **SR** | 0.004 | 0.6407 |
| **SIPI** | 0.0021 | 0.736624 |
| **BGI** | 0.0017 | 0.757789 |
| **NDVI** | 0.0013 | 0.7924 |
| **SAVI1** | 0.0013 | 0.792488 |
| **EVI** | 0.0006 | 0.862078 |

## 4.3. Multivariate analysis

### 4.3.1. Leaf N estimation based on stepwise multiple linear regressions

The multivariate regression technique of SMLR was used to predict leaf N models and to test the applicability of all five bands in combination with various vegetation indices. Red edge effect was tested through SMLR, by combining all five bands with various vegetation indices including, and excluding red edge band to estimate leaf N separately (Table 4.2 and 4.3). Studies showed the effect of red edge improves estimation of leaf N (Clevers *et al*, 2002, Ramoelo, 2012). The effect of red edge was not as significant when using the SMLR compared to univariate regression, as the red edge (710 nm) band did not improve leaf N estimation as expected through narrow band vegetation indices tested. The top five indices which performed well when using SMLR were chosen based on the higher coefficient of variance ($R^2$), and a lower root mean square error (RMSE). These top five vegetation indices were with the inclusion of red edge band (710 nm) were: BRI, SIPI, BGI, NGRR1 & NDVI, these results are showed in Table 4.3. The top five leaf N models with exclusion of 710 nm prediction were therefore based on SR4, SAVI, SR, DVI and RGI1. All top five performing vegetation indices are arranged in order of the best performing to the least performing vegetation index (1 as the best and 5 as the least performing vegetation index). They are all displayed in Table 4.2 according to a higher coefficient of variance ($R^2$), and a lower root mean square error (RMSE).

Table 4.2: Multivariate regression (SMLR) results of all four bands except for 710 nm band against vegetation index, each index showing their coefficient of determination ($R^2$), root mean square error (RMSE) and the probability-value (p-value). The vegetation indices and wave bands are: 475= Blue, 555= Green, 657= Red, 710= Red edge 805= Near Infra-Red, NDVI= Normalised Difference Vegetation Index, RE_NDVI= Red Edge_ Normalised Difference Vegetation Index, SR= Simple Ratio, RE_SR= Red Edge_ Simple Ratio, MTCI= MERIS Terrestrial Chlorophyll Index, GI= Green index, RGI= Red/ Green Index, RGI1= Red/ Green Index, BGI= Blue Green Pigment Index, BRI= Blue Red Pigment Index, GRR= Green-Red Reflectance, NGRR= Normalized Green-Red Ratio, NGRR1= Normalized Green-Red Reflectance, SR3= Simple Ratio, SR4= Simple Ratio, DVI= Difference Vegetation Index, DVI1= Difference Vegetation Index, SIPI= Structural Insensitive Pigment Index, SIPI1= Structural Insensitive Pigment Index, EVI= Enhanced Vegetation Index, NRI= Nitrogen Reflectance Index, SAVI= Soil Adjusted Vegetation Index, and SAVI1= Soil Adjusted Vegetation Index.

| **Variable (without 710/RE)** | **$R^2$** | **RMSE** | **P-value** |
|---|---|---|---|
| **All bands + SIPI** | 0.48261 | 0.610504 | 0.623117 |
| **All bands + NGRR1** | 0.339359 | 0.615081 | 0.767098 |
| **All bands + SR** | 0.301544 | 0.528098 | 0.002077 |
| **All bands + SIPI1** | 0.178344 | 0.572784 | 0.06705 |
| **All bands + BRI** | 0.148669 | 0.588036 | 0.133395 |
| **All bands + SRPI** | 0.148669 | 0.583036 | 0.133395 |

| | | | |
|---|---|---|---|
| **All bands + SR4** | 0.12732 | 0.579056 | 0.03357 |
| **All bands + SAVI** | 0.12447 | 0.580001 | 0.068465 |
| **All bands + NGRR** | 0.093912 | 0.601494 | 0.394945 |
| **All bands + NRI** | 0.093912 | 0.601494 | 0.394945 |
| **All bands + RGI** | 0.092553 | 0.601945 | 0.404276 |
| **All bands + GI** | 0.091374 | 0.602336 | 0.412487 |
| **All bands + GRR** | 0.091374 | 0.602336 | 0.412487 |
| **All bands + DVI** | 0.085612 | 0.592732 | 0.188027 |
| **All bands + SR3** | 0.07981 | 0.606156 | 0.49822 |
| **All bands + SAVI1** | 0.076383 | 0.607284 | 0.525292 |
| **All bands + NDVI** | 0.076383 | 0.607284 | 0.525293 |
| **All bands + RGI1** | 0.047101 | 0.599456 | 0.271808 |
| **All bands + BGI** | 0.042893 | 0.612223 | 0.676598 |

Table 4.3: Multivariate regression (SMLR) results of all five bands including 710 nm band against each vegetation index, showing their coefficient of determination, root mean square error (RMSE) and the probability-value (p-value) of each index. The vegetation indices and wave bands are: 475= Blue, 555= Green, 657= Red, 805= Near Infra-Red, NDVI= Normalised Difference Vegetation Index, RE_NDVI= Red Edge_ Normalised Difference Vegetation Index, SR= Simple Ratio, RE_SR= Red Edge_ Simple Ratio, MTCI= MERIS Terrestrial Chlorophyll Index, GI= Green index, RGI= Red/ Green Index, RGI1= Red/ Green Index, BGI= Blue Green Pigment Index, BRI= Blue Red Pigment Index, GRR= Green-Red Reflectance, NGRR= Normalized Green-Red Reflectance, NGRR1= Normalized Green-Red Ratio, SR3= Simple Ratio, SR4= Simple Ratio, DVI= Difference Vegetation Index, DVI1= Difference Vegetation Index, SIPI= Structural Insensitive Pigment Index, SIPI1= Structural Insensitive Pigment Index, EVI= Enhanced Vegetation Index, NRI= Nitrogen Reflectance Index, SAVI= Soil Adjusted Vegetation Index, and SAVI1= Soil Adjusted Vegetation Index.

| **Variable (with 710/RE)** | **$R^2$** | **RMSE** | **P-value** |
|---|---|---|---|
| **All bands + SR3** | 0.306522 | 0.53145 | 0.00416 |
| **All bands + SIPI** | 0.237839 | 0.551657 | 0.014125 |

| | | | |
|---|---|---|---|
| **All bands + RGI** | 0.221826 | 0.557422 | 0.021967 |
| **All bands + SIPI1** | 0.21687 | 0.559195 | 0.025101 |
| **All bands + SR** | 0.215741 | 0.559598 | 0.02587 |
| **All bands + MTCI** | 0.214634 | 0.554582 | 0.012307 |
| **All bands + NGRR** | 0.193532 | 0.561982 | 0.022476 |
| **All bands + NRI** | 0.193532 | 0.561982 | 0.022476 |
| **All bands + GI** | 0.185995 | 0.564603 | 0.027696 |
| **All bands + GRR** | 0.185995 | 0.564603 | 0.027696 |
| **All bands + SAVI1** | 0.174216 | 0.568673 | 0.038126 |
| **All bands + BGI** | 0.138814 | 0.57523 | 0.04614 |
| **All bands + NGRR1** | 0.133219 | 0.577096 | 0.053889 |
| **All bands + BRI** | 0.130322 | 0.578059 | 0.058361 |
| **All bands + SRPI** | 0.130322 | 0.578059 | 0.058361 |
| **All bands + RE_NDVI** | 0.125666 | 0.585151 | 0.129948 |
| **All bands + SAVI** | 0.125665 | 0.584091 | 0.129948 |
| **All bands + RE_SR** | 0.124467 | 0.585552 | 0.133677 |
| **All bands + SR4** | 0.122046 | 0.586361 | 0.141492 |
| **All bands + RGI1** | 0.103925 | 0.586767 | 0.118127 |
| **All bands + NDVI** | 0.072698 | 0.602615 | 0.406138 |

#### 4.3.2. Leaf N estimation based on Random forest regression

Random forest was used to estimate leaf N and to test the applicability of all five bands in combination with best performing vegetation indices. Red edge effect on estimation accuracy was tested, by computing all five bands with narrow-band indices (red-edge based indices) separately (Figures 4.2 and 4.3) and Tables 4.4 and 4.5. The top five vegetation indices selected according to a higher $R^2$ value, and when red edge was included with all the bands were RGI1, BGI, MTCI, BRI and SRPI; these vegetation indices were top 5 better predictors of leaf N. On the other hand, the top five vegetation indices when red edge was excluded were DVI1, SIPI, EVI, BGI and SR. All top five performing vegetation indices are arranged in order of the best performing to the least performing vegetation index (1 as the best and 5 as the least performing vegetation index).

Table 4.4: Multivariate regression results using random forest, of four bands excluding 710 nm band against vegetation index showing their coefficient of determination, root mean square error (RMSE) and the probability-value (p-value) of each index. The vegetation indices and wave bands are: 475= Blue, 555= Green, 657= Red, 805= Near Infra-Red, NDVI= Normalised Difference Vegetation Index, RE_NDVI= Red Edge_ Normalised Difference Vegetation Index, SR= Simple Ratio, RE_SR= Red Edge_ Simple Ratio, MTCI= MERIS Terrestrial Chlorophyll Index, GI= Green index, RGI= Red/ Green Index, RGI1= Red/ Green Index, BGI= Blue Green Pigment Index, BRI= Blue Red Pigment Index, GRR= Green-Red Reflectance, NGRR= Normalized Green-Red Reflectance, NGRR1= Normalized Green-Red Ratio, SR3= Simple Ratio, SR4= Simple Ratio, DVI= Difference Vegetation Index, DVI1= Difference Vegetation Index, SIPI= Structural Insensitive Pigment Index, SIPI1= Structural Insensitive Pigment Index, EVI= Enhanced Vegetation Index, NRI= Nitrogen Reflectance Index, SAVI= Soil Adjusted Vegetation Index, and SAVI1= Soil Adjusted Vegetation Index.

| **Variable (without 710/RE)** | **$R^2$** | **RMSE** | **P-value** |
|---|---|---|---|
| **All bands + DVI1** | 0.883913 | 0.20732 | <0.05 |
| **All bands + SIPI** | 0.883818 | 0.207405 | <0.05 |
| **All bands + EVI** | 0.877829 | 0.212684 | <0.05 |
| **All bands + SR** | 0.875684 | 0.214543 | <0.05 |
| **All bands + NDVI** | 0.875679 | 0.214547 | <0.05 |
| **All bands + SAVI1** | 0.875679 | 0.214547 | <0.05 |
| **All bands + BGI** | 0.875544 | 0.212739 | <0.05 |
| **All bands + RGI1** | 0.875349 | 0.214831 | <0.05 |
| **All bands + SRPI** | 0.872632 | 0.21716 | <0.05 |

| | | | |
|---|---|---|---|
| **All bands + NGRR1** | 0.871613 | 0.218027 | <0.05 |
| **All bands + BRI** | 0.870316 | 0.21716 | <0.05 |
| **All bands + SIPI1** | 0.869711 | 0.219636 | <0.05 |
| **All bands + GI** | 0.864256 | 0.224187 | <0.05 |
| **All bands + GRR** | 0.864256 | 0.224187 | <0.05 |
| **All bands + NRI** | 0.864228 | 0.22421 | <0.05 |
| **All bands + RGI** | 0.861557 | 0.226404 | <0.05 |
| **All bands + NGRR** | 0.861336 | 0.226585 | <0.05 |
| **All bands + DVI** | 0.858082 | 0.38013 | <0.05 |
| **All bands + SR3** | 0.852036 | 0.23406 | <0.05 |
| **All bands + SR4** | 0.852036 | 0.23406 | <0.05 |
| **All bands + SAVI** | 0.845705 | 0.239015 | <0.05 |

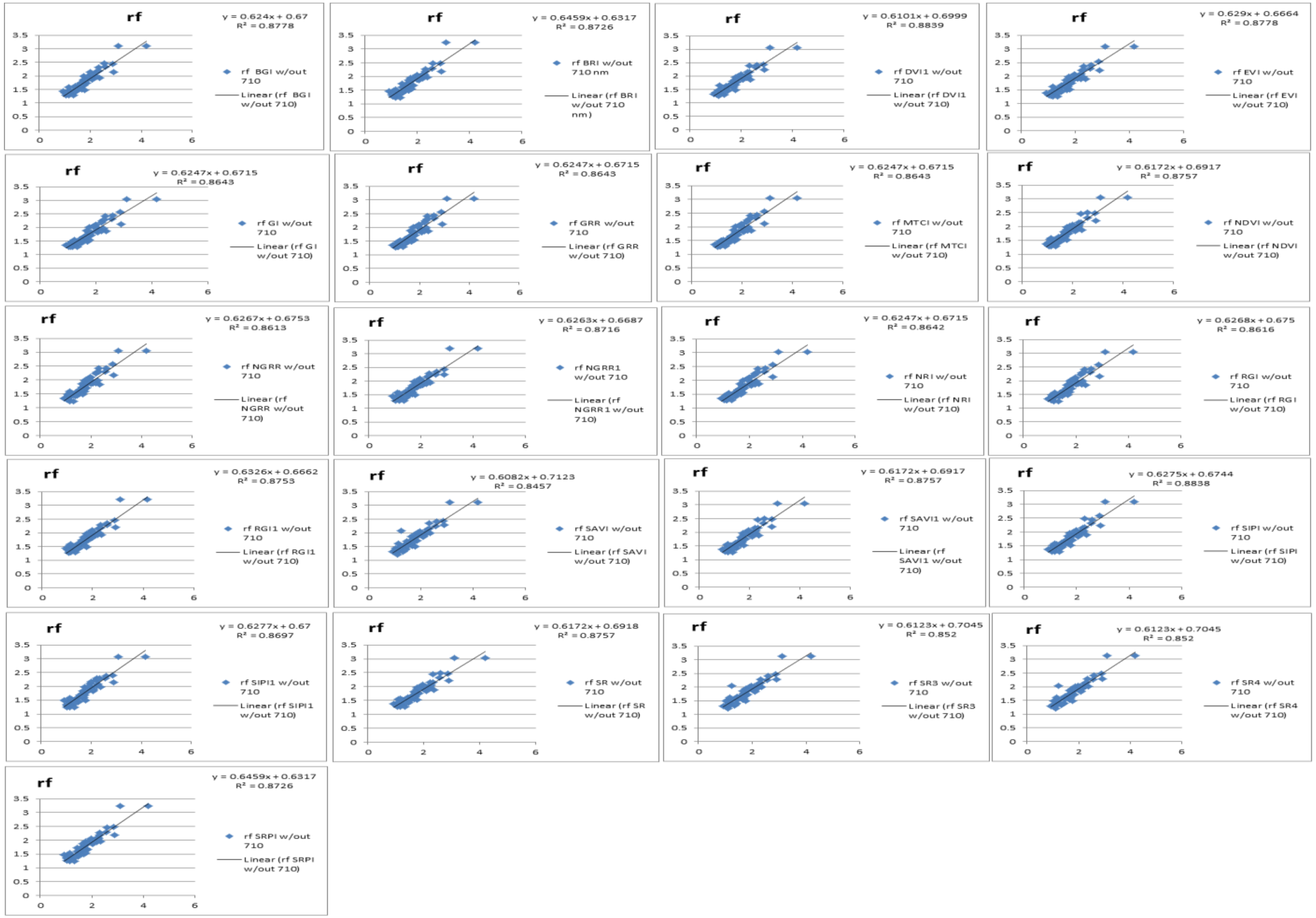


Figure 4.2: Scatter plots of leaf N model prediction calculated by random forest excluding the red edge/ 710 nm band, it is a combination of 4 bands and a vegetation index. The vegetation indices include: SR, SR4, SIPI1, SAVI, BRI, SRPI, DVI, RGI1, NGRR, NRI, RGI, GI, GRR, SR3, SAVI1, NDVI, SIPI, BGI and NGRR1. Each vegetation index displays a coefficient of determination ($R^2$), used to select the best model.

Table 4.5: Random forest results of all bands including 710 nm band against vegetation of each index showing their coefficient of determination, root mean square error (RMSE) and the probability-value (p-value) of each index. The wave bands and vegetation indices and are: 475= Blue, 555= Green, 657= Red, 805= Near Infra-Red, NDVI= Normalised Difference Vegetation Index, RE_NDVI= Red Edge_ Normalised Difference Vegetation Index, NGRR= Normalized Green-Red Ratio SR= Simple Ratio, RE_SR= Red Edge_ Simple Ratio, MTCI= MERIS Terrestrial Chlorophyll Index, GI= Green index, RGI= Red/ Green Index, RGI1= Red/ Green Index, BGI= Blue Green Pigment Index, GRR= Green-Red Reflectance, SAVI= Soil Adjusted Vegetation Index.

| Variable(with710/RE) | $R^2$ | RMSE | P-value |
|---|---|---|---|
| **All bands + RGI1** | 0.867168 | 0.221769 | <0.05 |
| **All bands + BGI** | 0.863504 | 0.224807 | <0.05 |
| **All bands + MTCI** | 0.863003 | 0.225219 | <0.05 |
| **All bands + BRI** | 0.862431 | 0.225689 | <0.05 |
| **All bands + SRPI** | 0.862431 | 0.225689 | <0.05 |
| **All bands + NGRR1** | 0.859455 | 0.228117 | <0.05 |
| **All bands + DVI1** | 0.857402 | 0.229777 | <0.05 |
| **All bands + DVI** | 0.856448 | 0.230545 | <0.05 |
| **All bands + EVI** | 0.856423 | 0.230565 | <0.05 |
| **All bands + SIPI** | 0.854168 | 0.232368 | <0.05 |
| **All bands + NDVI** | 0.852401 | 0.233771 | <0.05 |
| **All bands + SAVI1** | 0.852401 | 0.233771 | <0.05 |
| **All bands + SR** | 0.852369 | 0.233797 | <0.05 |
| **All bands + SR3** | 0.847295 | 0.237781 | <0.05 |
| **All bands + SIPI1** | 0.847122 | 0.237915 | <0.05 |
| **All bands + RE_NDVI** | 0.841567 | 0.242199 | <0.05 |
| **All bands + SAVI** | 0.841567 | 0.242199 | <0.05 |

| | | | |
|---|---|---|---|
| **All bands + RE_SR** | 0.841566 | 0.24219 | <0.05 |
| **All bands + GI** | 0.838899 | 0.24423 | <0.05 |
| **All bands + GRR** | 0.838899 | 0.24423 | <0.05 |
| **All bands + NRI** | 0.838839 | 0.244276 | <0.05 |
| **All bands + RGI** | 0.838177 | 0.244776 | <0.05 |
| **All bands + NGRR** | 0.838121 | 0.244819 | <0.05 |
| **All bands + SR4** | 0.835712 | 0.246634 | <0.05 |

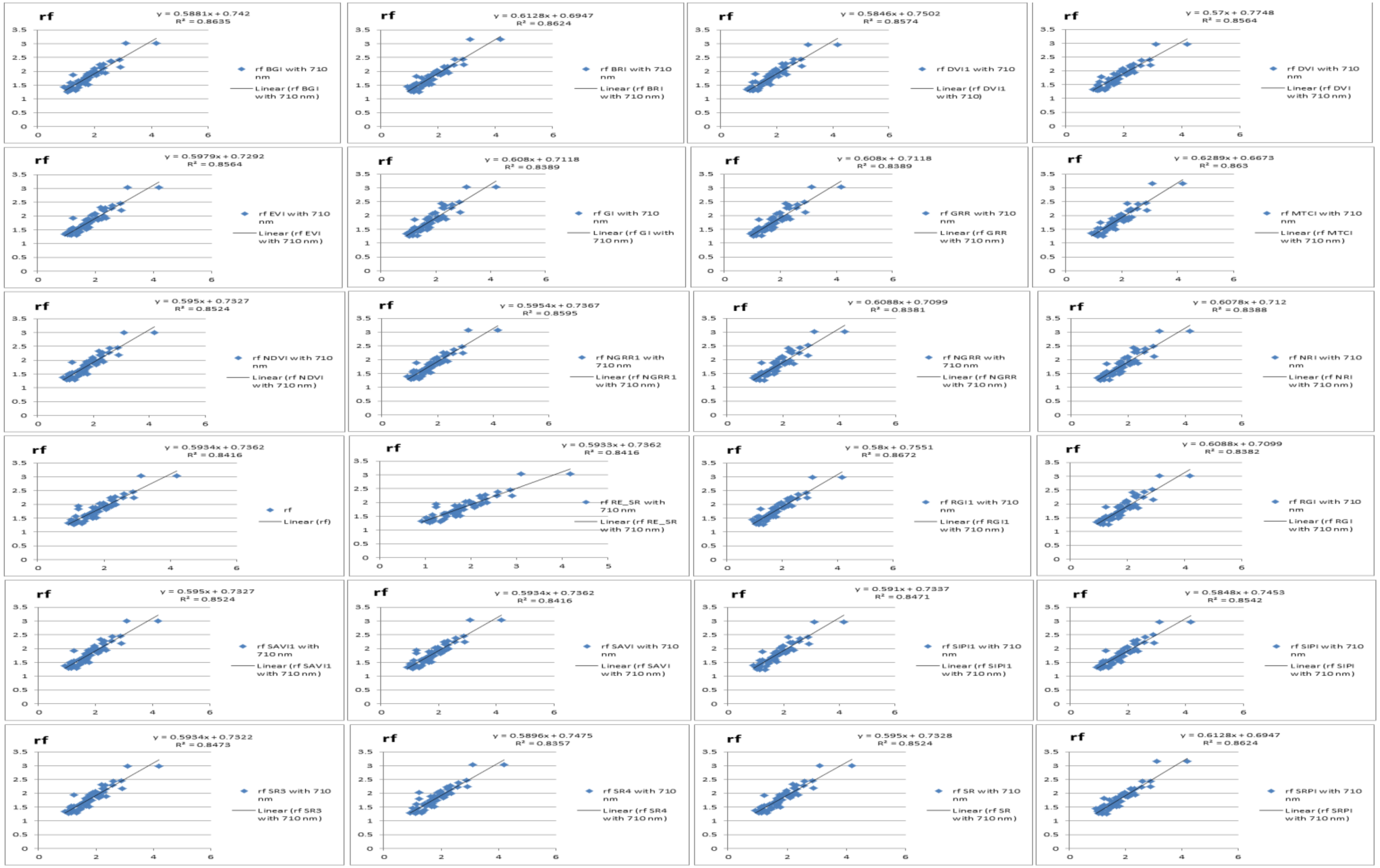

Figure 4.3: Scatter plots of foliar N model prediction calculated by random forest excluding the red edge/ 710 nm band, so it is a combination of 4 bands and a vegetation index. The vegetation indices include: SR, SR4, SIPI1, SAVI, BRI, SRPI, DVI, RGI1, NGRR, NRI, RGI, GI, GRR, SR3, SAVI1, NDVI, SIPI, BGI and NGRR1. Each vegetation index displays a coefficient of determination ($R^2$), used to select the best model.

## 4.4. Determination of factors influencing leaf N distribution as an indicator of plant stress

The vegetation health depends on leaf N, thus the lower concentration of leaf N actually leads to a poor vegetation health. Factors such as increasing temperature, inadequate water supply or unfavourable climatic conditions also contribute to vegetation stress (Field and Mooney, 1986; Ollinger *et al*, 2002). Ecological hypothesis testing of vegetation stress has a lot to do with the guidance and allocation of resources for stress mitigation and control (Pontius *et al*, 2005; Wulder *et al*, 2005). For example there is a close link between soil and vegetation types in dry and savanna areas than in high rainfall areas, for areas with low rainfall water becomes a limiting growth factor which affects the vegetation composition (Mucina & Rutherford, 2010). To disentangle and study the major influence of leaf N concentration, various environmental variables were related to leaf N.

- **Leaf N vs DEM**

The results showed that DEM or altitude does not significantly influence the concentration and distribution of leaf N ($R^2 = 0.0045168$, and p-value = 0.119280). This actually means that there was no significant relationship between leaf N and DEM, with the $R^2$ of 0. The null hypothesis was accepted with $p>0.05$.

- **Leaf N vs aspect**

For leaf N vs aspect, the results showed that aspect does not significantly influence the concentration and distribution of leaf N with the $R^2$ of 0.13 (p-value = 0.403149). The p-value showed that the effect of aspect on leaf N distribution was not significant, and the null hypothesis was accepted with $p>0.05$.

- **Leaf N vs. slope**

Slope has a remarkable effect on the distribution of plant nutrients, for example steep slopes have more run off which can lead to water stress to plants. Less organic matter is lost due to erosion from steep slopes (ARC, 2009). This simply means that the steeper the slope the lesser the nutrients for vegetation compared less steeper slopes habitats. The effect of slope on the distribution of leaf N concentration was determined through univariate regression ($R^2$ = 0.046965, and p-value =0.112010), which shows no significant relationship. The null hypothesis was then accepted with $p>0.05$.

- **Leaf N vs. vegetation types**

Vegetation types have the ability to produce various types of soil organic matter, thus any change on vegetation such as seasonal change affecting water supply in vegetation, can alter the pattern of accumulation of organic matter of soil. Organic matter of forest soil for example is mainly from fallen leaves (ARC, 2009). For understanding the effects of vegetation types on leaf N concentration and distribution, one way ANOVA was used to test this. The results showed that leaf N distribution significantly differ from various vegetation types ($p<0.05$). This indicates that, vegetation types influence the leaf N distribution, which could be further linked to the soil types (see Figure 4.4). It simply means spatial distribution of vegetation stress displayed the central areas with lower stress levels irrigation of agricultural crops. The northern parts of the area showed higher stress levels as there is extensive land use including mining. In this case, the null hypothesis was rejected, with $p<0.05$.

- **Leaf N vs. soils**

The ability of soil to supply vegetation with necessities such as nitrogen depends on parent material thickness, texture and mineral content. The structure of soil is basically in such a way that soils from the parent material formed by hard rock has less plant growth than the deeper soils. Unfertilized sandy soils are less likely to be fertile while clay soils are more fertile (ARC, 2009). Environmental parameters such as steep slopes are typical of Lephalale area, and soil types are also diverse. These soil types are formed in: weakly developed areas, mountainous catchment, uplands and rocky areas to mention but a few (WDEMF-Draft Report, 2010).

Different soil types were also studied to see if there will be any effect of soil on the leaf N distribution and concentration. One way ANOVA statistical method was used, to see if there is a significant difference across various soil type ($p<0.05$) and can be depicted in Figure 5.4, showing spatial distribution of leaf N significant variation across different soil types. This means spatial distribution of plant stress was as follows: the central areas had lower stress levels because of irrigated agricultural crops. The northern parts of Lephalale are dominated by extensive land use such as mining showed significant stress levels. The null hypothesis was rejected with $p<0.05$.

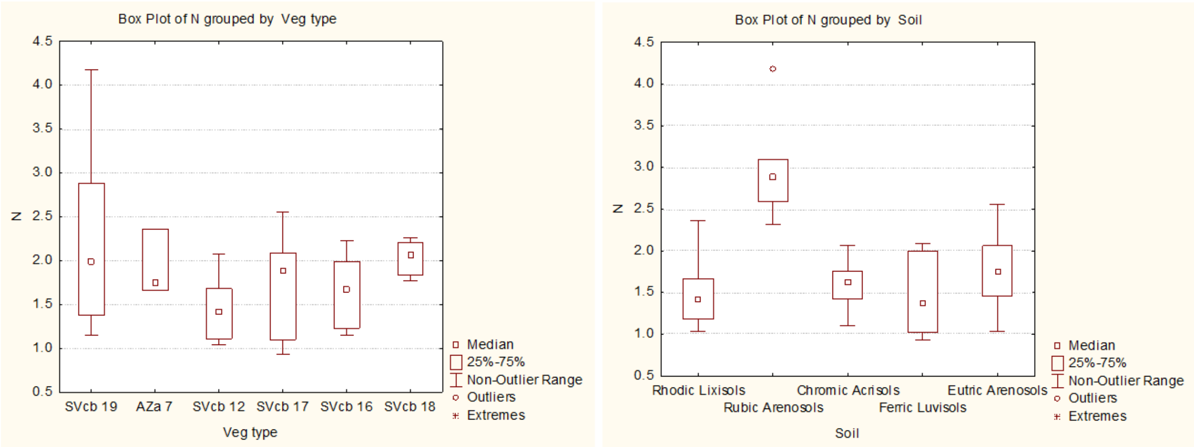


Figure 4.4: Box plots of vegetation (veg) types and soil types against N, demonstrating the effect of soil and vegetation types on the distribution of leaf N. The names of the types of vegetation SVcb 19 stands for Limpopo Sweet Bushveld, AZa 7 is for Subtropical Alluvial Vegetation, SVcb 12 is for Central Sandy Bushveld, SVcb 17 stands for Waterberg Mountain Bushveld, SVcb 16 is Western Sandy Bushveld, while SVcb 18 stands for Roodeberg Bushveld.

### 4.5 Leaf N maps – stress levels in the Waterberg region

The spatial distribution of leaf N was demonstrated in Figure 4.5 and 4.6. Figure 4.5 shows the general vegetation greenness based on the NDVI, while Figure 4.6 shows the spatial distribution of leaf N (%). The predicted leaf N values range between 0.01 to 3%. High leaf N values are found in the riparian zones and mostly on the irrigated agricultural areas towards the south. There is a general consensus between leaf N map and NDVI on the northern part of the region, which is relatively stressed than other regions. The unstressed areas in the riparian zones are well-depicted by the leaf N. Sharper and more informative stress levels are depicted by leaf N maps.

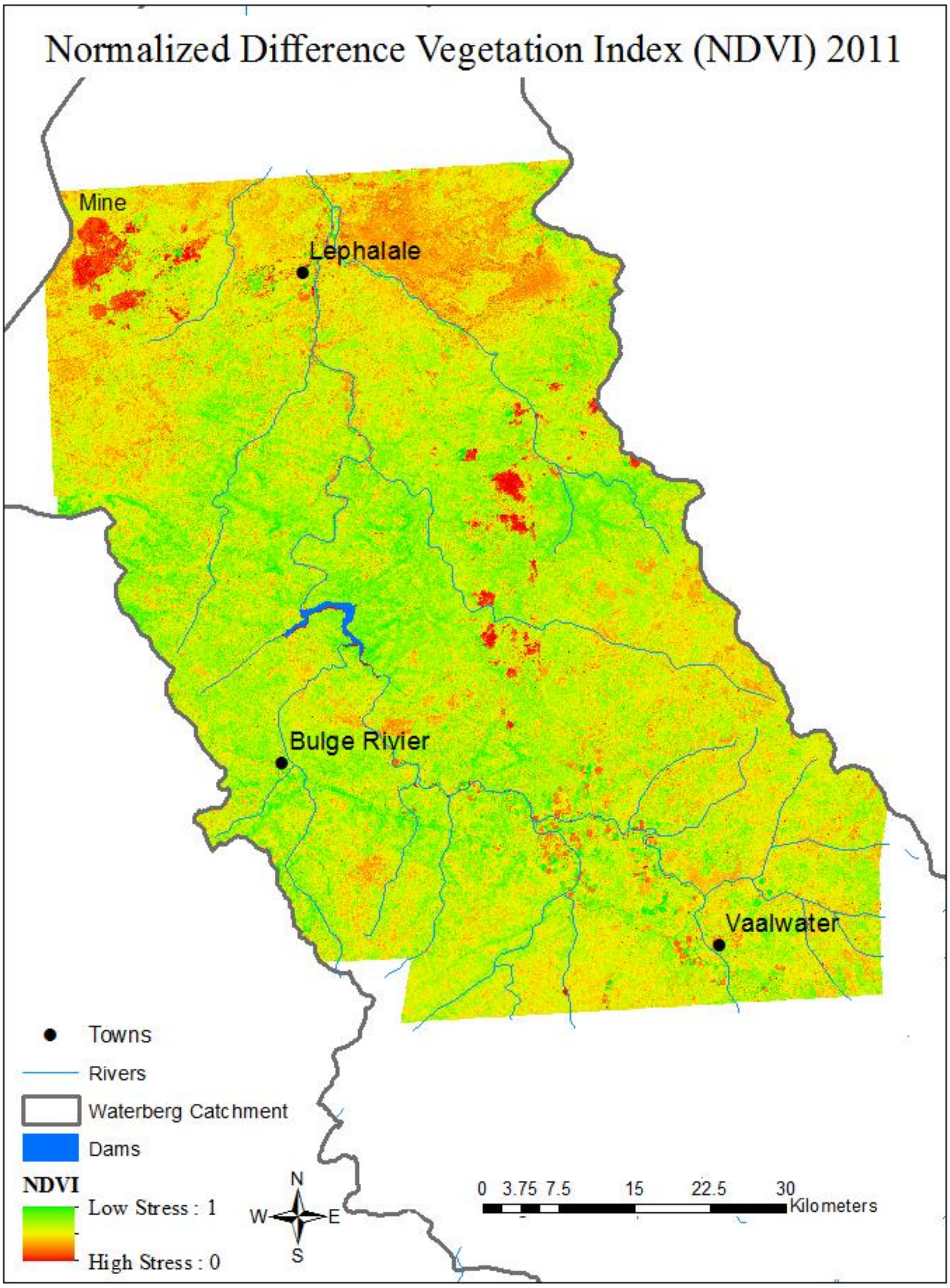


Figure 4.5: Vegetation index based on RapidEye image, vegetation (NDVI showing the level of greenness in the Waterberg region.

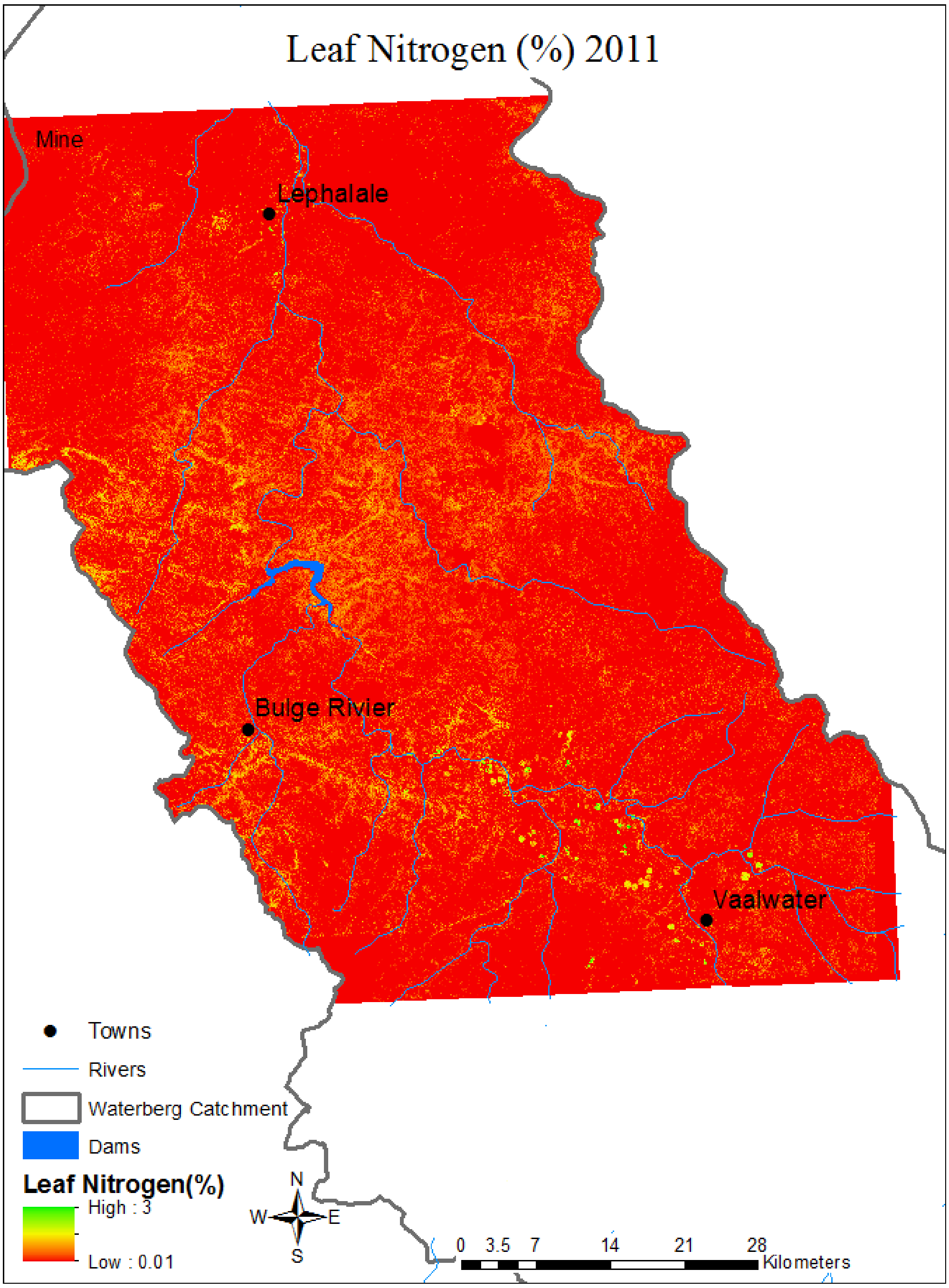


Figure 4.6: Leaf N distribution mapped through one of the best multivariate models which shows the leaf N stressed northern part of Lepahalale area mine associated with more water usage. The middle to the Southern site is less stressed with more leaf N concentration where there are rivers which contribute to vegetation hydration.

# CHAPTER 5: DISCUSSION

Nutrients information is critical in understanding the condition of the vegetation. The findings showed that leaf N correlated with vegetation indices, for example MTCI and 710 nm wave band were the top two best predictors of leaf N when using simple linear regression/univariate analysis. Red-edge band (710nm) and the MTCI (based on the red edge band) are important as they are highly correlated with chlorophyll (Clevers *et al*, 2002; Cho & Skidmore, 2006). Furthermore the red edge position is described as the inflection point on the slope which connects the red and NIR regions (Mutanga & Skidmore, 2007; Pu *et al,* 2003); the steep increases in reflectance influences the chlorophyll absorption feature in the red region. This is between the photosynthesis region and the high reflectance value region of NIR, this is also where plant cell structure and leaves layers are usually affected. The red edge position provides a better understanding of vegetation health (Herrman *et al,* 2010). Univariate results highlighted the positive effect of red edge for accurate leaf N estimation, implying that leaf N estimation is better when red edge band was included, compared to when it is excluded as demonstrated in Figure 4.1.

SMLR results were used to test the applicability of all five bands in combination with best performing vegetation index and also to test the effect of red edge band. The top five performing vegetation indices, according to the highest coefficient of determination ($R^2$) when red edge was included were: BRI, SIPI, BGI, NGRR1 & NDVI while when red edge was excluded the top five indices were: SR4, SAVI, SR, DVI & RGI1. All the top five performing vegetation indices are presented in the order of the best performing to the least performing vegetation index. This further highlights the contribution of the red edge band. Red edge band inclusion displayed improvement of leaf N estimation accuracy.

For random forest method top five performing vegetation indices, according to the highest coefficient of determination ($R^2$) when red edge was used were: RGI1, BGI, MTCI, BRI and SR, while when red edge was not used top five indices were: DVI1, SIPI, EVI, SR, and NDVI. All the top five performing vegetation indices are presented in the order of the best performing to the least performing vegetation index, thus RGI1 is the best when red edge was used while DVI1 was the best vegetation index to estimate leaf N when red edge was excluded. MTCI has been a better vegetation index to predict leaf N for both univariate (the best vegetation index) and random forest (third best vegetation index) results, which means there is correlation between leaf N and MTCI. Red-edge band has once again demonstrated its capability to improve leaf N estimation and modelling of leaf N, as the red edge computed MTCI was the third best vegetation index for random forest when red edge was included. The random forest results were generally higher than that of SMLR. Random forest has the capability of solving overfitting and multicollinearity as compared to SMLR, this may explain the better performance of random forest than SMLR.

**5.1 Leaf N estimation using resampled ASD to RapidEye**

In this study, the ASD measured reflectance was resampled to RapidEye spectral configuration. The models were developed from ASD resampled data and eventually the best model was inverted on the actual image to map the spatial distribution of leaf N. Mutanga *et al*, 2015 also resampled field spectra data using ASD to develop models and testing them on an actual WorldView-2 (a type of satellite imagery used) image. The study used random forest regression model and normalised vegetation indices (NDI) to predict leaf N concentration in grassland environment. It was also mentioned that there were no other similar study studies which employed this approach according to the records obtained by the researchers. The results of the study revealed that prediction of leaf N was successful proving that resampling field spectra data has a potential to provide earth observation field with reliable information.

The approach of resampling ASD data provides future opportunity to estimate leaf N on atmospherically corrected images without going to the field. This could be achieved by developing a spectral library for various tree and grass species, and including crops with their corresponding leaf N values. Therefore, robust machine learning techniques such as random forest can be used to estimate leaf N without extensive field work, given that there is an existing spectral library. This study presents an innovative approach in estimating leaf N concentrations as an indicator of vegetation status or quality. The advantage of remote sensing compared to other methods is highlighted by this type of approach which displays a sustainable way of conducting scientific research without harming the ecosystem.

## 5.2 Use of leaf N for photosynthesis

Leaf N is important for photosynthesis as it determines major functions of ecosystem such as the rate of nutrient and carbon intake (Guerschman *et al*, 2009). Water content, plant nutrient and pigments influence the rate of photosynthetic activities, litter decomposition, leaf respiration, growth rates as well as nutrient cycling; thus acting as indicators for ecosystem condition (Field and Mooney, 1986; Ollinger *et al*, 2002). The correlation between chlorophyll and N is an indication that leaf N has a role to play in photosynthesis. In other words chlorophyll which is important for photosynthesis process is directly proportional to leaf N and therefore an indicator of vegetation health or status. The concentration of leaf N has been proven to be related to the net photosynthesis across various plant species and functional groups, thus there is a profound link between terrestrial cycles and carbon cycles. Vegetation may experience stress due to unfavourable conditions which leads to plant physiological functions such as light and dark photosynthesis declining from their optimal physiological standards (Logan *et al*, 2003; Ninements, 2010).

## 5.3. Factors influencing leaf N over the landscape

### 5.3.1. Soil as a factor influencing distribution of leaf N distribution

Different soil forms and World Reference Base soil groups such as Rhodic lixisols, Ferric luvisols, Eufric arenosols, Rubic arenosols, Chromic Acrisols exist in Lephalale area and there are unique properties which distinguishes each group or soil type (Fey, 2010). The boxplots in Figure 4.4 shows the graphical demonstration of the effect of various soil and vegetation types on leaf N distribution. This is supported by the fact that leaf nutrients are influenced by soil (Mutanga *et al*, 2004). The effect of soil types on leaf N distribution is as follows:

(a) Rubic arenosols is associated with high leaf N distribution according to Figure 4.4, they fall under oxidic category which may be of either Xanthirhodic (yellow-brown apedal B over red apedal B horizon) or Xanthic-hydromorphic (yellow-brown apedal B over unspecified material with signs of wetness) soil formation and reddish on the landscape. They are said to be enriched with clay, high water retention and having a good shrink-swell potential (Fey, 2010). It is therefore not surprising that Rubic arenosols is associated with high leaf N distribution compared to others as depicted on the box plot in Figure 4.4. The soil type has nutrients and retains water which is good for vegetation health.

(b) Eutric arenosols, are a type of soil which are sandy in nature, easy to till, unable to store water and are permeable (FAO, 1993). Arenosols are a product of weathering of rock, generally this type of soils are hard to explore so nitrogen percentage are usually at lower concentration in many of the performed analysis, although there is a wide variation into the Arenosols, they may also have the less concentration because of their exploitation challenges (van Englen & Dijkshoorn, 2013; FAO, 1993). So Eutric arenosols had a fairly good distribution of leaf N in Figure 4.4 shown as the second best leaf N distribution. One of the reasons why there seems to be a positive relationship with leaf N might be because the data was collected during wet season. The types of vegetation studied during data collection also play a role thus if there is a need for deep soil or requires sufficient water supply or not.

(c) Chromic acrisols are a type of Plinthic soils which are distinguished by a hard soil formation. Plinthic soils are characterised by segregation and concentration of Iron oxides with marked spots and particles binding together. It should be noted that they are not found in higher or lower rainfall regions thus they are largely absent in most arid or humid regions. Soils horizons of Plinthic nature may act as water barriers for vegetation, the formation of the soil may be soft or hard. Plinthic soils are usually found in neither higher nor lower rainfall regions (Fey, 2010), and characterised by low activity clays in argic (horizon with higher clay content) subsurface horizon, crops cultivated on acrisols are dependent on fertilizers and need to be supplemented by water (rain or irrigation) (van Englen & Dijkshoorn J. A, 2013). Chromic acrisols showed moderate levels of leaf N distribution, when compared to other soil types which may be due to the moderate fertile nature of the soil.

(d) Rhodic Lixisols, strongly weathered soils characterised by luvial horizon that has been washed away by clay down to a horizon called argic (subsurface horizon with a distinctly higher clay content than the overlying horizon) that has low clay activity with a moderate to high base saturation level. Lixisols are strongly weathered and leached with fine texture, covered or overlain by sandy and coarser textures of soil throughout. Erosion on the slopes affects the nature of lixisols and they need to be supplemented by fertilisers (van Englen &

Dijkshoorn, 2013). Lixisols are not so fertile, as the box plot (see Figure 4.4) shows lower levels of leaf N distribution.

(e) Ferric luvisols are fertile soils which are suitable for vegetation growth; they are unfortunately prone to deterioration when tilted. Luvisols are commonly found on the slopes and are sensitive to processes such as erosion (van Englen & Dijkshoorn, 2013), as shown in Figure 4.4, Ferric luvisols appears to display varied (different concentrations) leaf N distribution which may be due to land use or erosion which affects the nutrient uptake by plants and the retention of water.

### 5.3.2. Vegetation types as a factor influencing leaf N distribution

It is worth mentioning that under extreme water stress conditions such as low rainfall which leads to high evapotranspiration, trees and shrubs will need deep soil to survive. On the other hand in case there is optimal water supply, shallower soils are able to support the growth of trees and shrubs. Leaf nutrients can be closely related to soil texture (Mucina & Rutherford, 2010). Leaf N distribution variation in vegetation types were as follows as well:

(a) Roodeberg bushveld described the vegetation type to be the one that grows in a sandstone conglomerate siltstone, mostly found in sandy high base status and vegetation features include short closed woodland to tall open woodland and poorly developed grass layer while trees of this vegetation type are not limited to hills. The species are classified as the least threatened and there are attempts for preservation are said to be fairly successful (Mucina & Rutherford, 2010). Results for Roodeberg bushveld showed good leaf N distribution levels according to the box plot as displayed in Figure 4.4. In other words the leaf N concentration is expected not to be so great due to factors such as sandstone and metavolcanic stone, so leaf N distribution level is also affected by malnourished soil and the inability of soil to retain water during the wet and hot season.

(b) Limpopo sweet Bushveld, has a landscape features such as irregular plains and vegetation features include short open woodland. They survive in various types of conditions such as clayey-loamy soil (for example black clayey soil), surface limestone layers, brownish sandy soils. They are also the least threatened of the vegetation types in terms of conservation (Mucina & Rutherford, 2010). According to the boxplot in Figure 4.4 leaf N distribution levels in this vegetation type was varied with lower and higher concentration of leaf N for this vegetation type. This may be due to the fact that this vegetation type survives under different types of environmental condition such as fertile (for example black clayey soil) or less fertile soil (such as sandy soils).

(c) Western sandy Bushveld are found in various species such as tall, open to low woodland, broad-leaved or even microphylous trees. Typical habitats include shallow soils of gravelly upland sites and deep sands (Mucina & Rutherford, 2010). Leaf N distribution in Western sandy Bushveld as demonstrated in Figure 4.4 appears to be the second lowest due to factors such as deep sand and shallow soils which contribute to vegetation stress or low vegetation cover.

(d) Central sandy Bushveld, are typically found in between mountainous areas, sandy plains, deep sandy soils and low rocky or gravelly soils. Other area common habitats include lower slopes on eutrophic sands and including soils that are less sandy. This vegetation type survives the hot and wet season as well, although they are described as vulnerable species in terms of conservation (Mucina & Rutherford, 2010). The boxplot in Figure 4.4 depicts Central Sandy Bushvelds to have variation in the level of leaf N distribution than other vegetation types (thus lower, moderate and high leaf N concentrations). This can be owed to the environmental conditions such as sandy or rocky soil type which usually have limited nutrients and lower water retention, thus contribution to the vulnerability such as water scarcity. The vegetation type survives under strenuous conditions, hence there is leaf N concentration detected.

(e) Waterberg Mountain Bushveld is mainly found on mountains, higher slopes, also in rocky mid and foot-slopes habitats and characterised by broader leaves. The vegetation type is typically found in lower-lying valleys including deeper sands of the plateaus. The vegetation mainly grows on sandstone, siltstone & shale, and also medium to coarse grained sandstone. So the vegetation is subjected to acidic, sandy, loamy to gravelly soil due to the nature of the environment they are found in (Mucina & Rutherford, 2010). Waterberg Mountain Bushveld had a varied (lower to high nitrogen concentrations) leaf N distribution this may be brought by various environmental conditions which affects the concentration of nitrogen. For example acidic, sandy, loamy to gravelly soil and the mountainous environment affect the nutrition and water supply.

(f) Subtropical Alluvial vegetation are typically found in soils which are sandy to loamy, water logged, and prone to floods during rainy season and has higher salt accumulation due to higher evaporation. This vegetation types are mainly found on channels of flowing river or river-fed pans, and also in areas were water flows slowly (Mucina & Rutherford, 2010). According to the box plot in Figure 4.4 the Subtropical Alluvial vegetation had a skewed leaf N distribution levels thus leaf N concentration is not spread evenly. It is supported by the fact that water supply was high since it was a rainy season (typically the vegetation is found in water logged habitat) and the soil type is being sandy-loamy which can retain water and nutrients supply including nitrogen is also fair. Basically the soil type that can retain water is usually fertile meaning it has sufficient nitrogen to supply to vegetation.

### 5.3.3. Topographical features (DEM, Aspect and Slope) as factors influencing leaf N distribution

Road sampling was the approach employed to collect leaf field data, and the results showed that leaf N distribution does not show correlation and significance with DEM, aspect and slope, thus topographical features did not have an influence on leaf N distribution. This may be because these environmental parameters might not be the best demonstration of leaf N distribution on the road compared to sampling the entire field beyond the fences. Environmental parameters such as soil and topographic factors affect distribution and concentration of plant nutrients (Venter, *et al*, 2003, Mucina & Rutherford, 2006). To simplify the meaning of DEM is to describe it as a digital format of earth's surface either wholly or as a part of it (Bolstad & Stowe, 1994). DEM therefore includes slope and aspect (Das, 2013), so the effect of slope and aspect on leaf N distribution also affects DEM in the similar manner.

Steep slopes often lead to vulnerability such as nutrients being washed away by run-off, and soil erosion which threatens vegetation health. So the slope type has an effect on leaf N distribution. The slope direction or aspect also depends on factors such as the amount of water supply and sunlight exposure to the vegetation. In geographical terms the direction of aspect is measured in degrees towards the downslope of maximum rate of change from the north in a clockwise direction (Das, 2013). For example if the slope direction is exposed to excessive sunlight the vegetation experiences higher evapotranspiration compared to when it is not, this affects the leaf N distribution. A lower leaf N concentration will typically be because vegetation experiences extreme environmental conditions such higher exposure to sunlight, site quality, plant and animal behaviuor and even an unfavourable drainage type whereby vegetation has lower water uptake (Das, 2013) as the soil on such aspect cannot retain sufficient water. The data was collected during summer (December 2011) which is a hot and rainy, it is therefore not strange to have findings such as the insignificance of aspect on leaf N distribution. It is therefore this underlying analogy which explains the reasons that lead to these topographical features not to significantly influence leaf N distribution.

### 5.3.4 Leaf N distribution map

The distribution of leaf N as demonstrated on the maps (see Figures 4.5 and 4.6) created through the best models can be interpreted as follows: the northern part of Lephalele is stressed with leaf N whereas the middle to the south is less stressed.. Land use of any kind whether it is agriculture, or mines changes how the ecosystem relates to the atmosphere and land including the natural ecosystem structure and functioning (Kampa & Castanas, 2008; Vitousek *et al*, 1997). These types of land use are due to human enterprise which interacts seriously with the global environmental change components (Vitousek *et al*, 1997). Anthropogenic activities have a direct impact on vegetation stress in the area, as common scarce resources such as water are used by plants as well. The distribution of leaf N in the study area highlights how the human activity induces vegetation stress, because where there is land usage there is a higher vegetation stress as well.

Leaf N spatial distribution clearly displays the fact that the vegetation in the periphery of sufficient water supply, such as rivers and dams is not water stressed. The vegetation around the mine is more stressed which may be due to the land use such as mine which demands higher amount of water. It is from this observation that leaf N concentration is higher where there is sufficient water supply, and lesser where there is insufficient water supply such as the northern part of the study area. Leaf N can therefore act as an indicator of water stress, as it is directly proportional to water availability.

According to Ramoelo *et al*, (2014) who studied the potential of monitoring plant stress using remote sensing for dry and wet season in Lephalale area. The study revealed that the spatial distribution of plant stress displayed the central areas had lower stress levels due to the fact that there are irrigated agricultural crops, while the northern parts of the area is dominated by extensive land use showed significant leaf N stress levels. The study therefore supports the findings of this research which demonstrates similar pattern of plant stress of leaf N within the area.

# CHAPTER 6 :CONCLUSIONS AND RECOMMENDATIONS

Assessment of vegetation condition or stress is possible with the use of remote sensing, especially, new sensors such as RapidEye with the red edge band. The effect of red edge on the estimation of leaf N has been highlighted through the methods of univariate regression, multivariate methods such as random forest. The study further demonstrated that red edge (710 nm) improves leaf N estimation results as compared to when it is not used. Vegetation indices computed from red edge should therefore be used for leaf N estimation, as red edge wave band is not sensitive to background effects and produces better and accurate results for biochemical estimation in order to study vegetation health.

The objectives of the study were reached as leaf N was successfully estimated using vegetation indices through univariate and multivariate regression methods, except for the insignificance of DEM, slope and aspect on leaf N distribution patterns. Leaf N concentration was also proven to be varied across different vegetation types. The distribution of leaf N also varied across soil types studied, however slope and aspect had no significance on leaf N distribution. Insignificance of DEM, slope and aspect on leaf N distribution can be due to various reasons including unfavourable or even extreme environmental conditions. Unfavourable environmental conditions include high sunlight exposure, poor site quality, and drainage conditions or erosion. Topographical features such as slope and aspect did not influence the distribution of leaf N significantly. This could be attributed to the fact that the sampling procedure was based on the road, due to access restrictions anthropogenic sites such as agricultural sites and mines. The detailed and overall findings of the study area would have been told if there were no restrictions of sampling, meaning a random sampling rather than purposive sampling is preferential for future studies. However the road sampling also paints a picture of lower leaf N concentration due to unfavourable environmental parameters impact, experienced by the vegetation.

Soil and vegetation types play a crucial role in the understanding of the distribution of leaf N as they significantly affect and influence the concentration of leaf N. The estimation of leaf N is plausible, and can be used as an indicator of vegetation stress, the information can be used for deriving baseline information for biodiversity and conservation purposes.

Red edge inclusion has proved to improve leaf N estimation and therefore future studies can utilise this useful waveband to accurately study vegetation health or condition. The study further recommends that additional samples points need to be collected through further engagement with the protected, mining and agricultural areas which resulted due to limited restrictions of access to have additional sampling points. This could enable a detailed understanding of the variability of leaf N and its drivers. Further recommendations includes that the approach of resampling ASD measured reflectance data into satellite imagery should be employed for future studies as the findings highlighted how this approach produces credible results. Although fewer studies have resampled ASD data, this innovative approach has a great potential in scientific research and development of earth observation. This study contributes profoundly to environmental impact studies, to measure the effect of anthropogenic activities to the ecosystem.

## REFERENCES


Abdel-Rahman, E.M., Ahmed F.B., & Ismail, R. (2013). Random forest regression and spectral band selection for estimating sugarcane leaf nitrogen concentration using EO-1 Hyperion hyperspectral data. *International Journal of Remote Sensing* Vol. 31, No 2, 712-728.

Aber, J. D., & Federer, C. A. (1992). A generalized, lumped-parameter model of photosynthesis, evapotranspiration, and net primary production in temperate and boreal forest ecosystems. *Oecologia* 92, 463-474.

Agricultural Research Council. (2009). *Soil Science: Training of Subject Advisors for the Department of Education*. Pretoria: ARC.

ArcGIS 10.2.1 for Desktop: Help files. 1999-2013. Esri Inc.

Asner, G.P., Wessman, C.A., Schimel, D.S., & Archer, S. (1998). Variability in leaf and litter optical properties: implications for BRDF model inversions using AVHRR, MODIS, and MISR. *Remote Sensing of Environment* 63(3), 243-257.

Asner, G.P. (1998). Biophysical and Biochemical Sources of Variability in Canopy Reflectance. *Remote Sensing of Environment* 64 (3), 234-253.

Atzberger, T., Jarmer, T., Schlerf, M, Kotz, B., & Werner, W. (2003). Spectroradiometric determination of wheat bio-physical variables: comparison of different empirical statistical approaches. *Remote Sensing in Transition, Proc 23rd EARSeL Symposium, Belgium,* pp 463-470.

Auearunyatwat, P., Kasetkasem, T., Wongmaneeroj, A., Nishihara A., & Keinprasit R. (2012). An Automatic Nitrogen Estimation Method in Sugarcane Leaves Using Image Processing Techniques. *International Conference on Agricultural, Environment and Biological Sciences (ICAEBS'2012) May 26-27, Phuket.*

Balasubramanian, V., Morales, A.C., Cruz, R.T., Abdulrachman, S. (1999). On farm adaptation of knowledge-intensive nitrogen management technologies for rice systems. *Nutrient Cycling Agroecosystems* 53(1), 59-69.

Baret, F., Guyot, G. (1991). Potentials and limits of vegetation indices for LAI and APAR assessments. *Remote Sensing of Environment,* 35(2-3), 161-173.

Beeri, O., Phillips, R., Hendrickson, J., Frank, A.B., Kronberg, S. (2007). Estimating forage quantity and quality using aerial hyperspectral imagery for northern mixed-grass prairie. *Remote Sensing of Environment* 110 (2), 216-225.

BlackBridge: Satellite Imagery Product Specification, version 6.0. (2013). [http://www.blackbridge.com/rapideye/about/resources.htm?tab=7 [Date accessed 20 September 2014]

Bolster K., Martin, M. E., & Aber J, D., (1996). Determination of Carbon fraction and Nitrogen concentration in tree foliage by near infrared reflectance: A comparison of statistical methods. *Canadian Journal of Forest Research,* 26, 590-600.

Bolton, J.F. (2004). Full spectral imaging: a revisited approach to remote sensing. *The International Society for Optical Engeneering* vol: 5234, 243-251.

Bolstad, P.V., & Smith, J.L. (1994). An evaluation of DEM Accuracy: Elevation, Slope Aspects. *Photographic Engeneering and Remote Sensing,* Vol: 60, No 11, 1327-1332.

Breiman, L. (2001). Random forest. *Machine learning*, 45(1), 5-32.

Bunke, O., & Droge, B. (1984). Bootstrap and cross-validation estimates of the prediction error for linear regression models. *The Annals of Statistics,* 76 (2), 156-172.

Burrough, P.A., & McDonell, R.A. (1998). *Principles of Geographical Information Systems.* New York: Oxford University Press.

California Fertilizer Foundation. (2011). *Natural Resource Fact Sheet: Plant nutrients Nitrogen*. Information compiled by California Fertilizer Foundation.

Card, D.H., Peterson, D.L., Matson, P.A., & Aber, J.D. (1988). Prediction of leaf chemistry by the use of visible and near infrared reflectance spectroscopy. *Remote Sensing of Environment*, 26, 123-147.

Chapelle, E.W., Kim, M.S., & McMurtrey, J.E. (1992). Ratio analysis of reflectance spectra (RARS): An algorithm for the remote estimation of the concentrations of Chlorophyll A, Chlorophyll B, and Carotenoids in soybean leaves. *Remote Sensing of Environment*, 39, 239-247.

Cho, M.A., & Skidmore, A.K. (2006). A new technique for extracting the red edge position from hyperspectral data; the linear extrapolation method. *Remote Sensing of Environment* 101 (2), 181-193.

Cho, M.A., Van Aardt, J.A.N., Main R., Majeke B., Ramoelo A., Mathieu, R., Norris-Roggers, M., & Duplessis R. (2009). Integrating Remote Sensing and ancillary data for regional ecosystem assessment: *Eucalyptus grandis* agrosystem in Kwazulu Natal, South Africa. *IEE International Geoscience and Remote Sensing Symposium (IGARSS).* Cape Town, South Africa.

Cho, M.A., Van Aardt, J.A.N., Main, R., & Majeke, B. (2010). Evaluating variations of physiology-based hyperspectral features along a soil water gradient in a Eucalyptus grandis plantation*.* *International Journal of Remote Sensing*, 31 (16), 4507 - 4507.

Clark, R., & Roush, T.L. (1984). Reflectance spectroscopy: Quantitative analysis techniques for remote sensing applications. *Journal Geophysical Research*, 89, 63929-6340.

Cleugh, H.A., Leuning, R., Mu, Q., Running, S.W. (2007). Regional evaporation estimates from flux tower and MODIS satellite data. *Remote Sensing of Environment.* 106, 285-304.

Clevers, J.G.P., De Jong, S.M., Epema, G.F., Van Der Meer, F.D., Bakker, W.H., Skidmore, A, K., Scholte, K. H. (2002). Derivation of the red edge index using the MERIS standard band setting. *International Journal of Remote Sensing* 23 (16), 3169-3184.

Clifton, K.E., Bradbury, J.W., & Vehrencamp, S.L. (1994). The fine-scale mapping of grassland protein densities. *Grass and Forage Science*, 49 (1), 1-8.

Curran, P.J. (1989). Remote sensing of foliar chemistry. *Remote Sensing of Environment*, 30, 271-278.

Curran, P., Dungan, J., Macler, B., & Plummer, S. (1992). Reflectance spectroscopy of fresh whole leaves for estimation of chemical concentration. *Remote Sensing of Environment*, 39, 153-166

Das, P. (2013). Digital Elevation Model is a Tool for Terrain Analysis: Implication and Interpretation with Reference to Kuya River Basin. *International Journal of Engeneering Science and Innovative Technology (IJESIT),* vol 2, issue 1.

Datt, B. (1998). Remote sensing of Chlorophyll a, Chlorophyll b, Chlorophyll a + b, and total Carotenoid content in Eucalyptus leaves. *Remote Sensing of Environment*, 66, 111-121.

Darvishzadeh, R., Skidmore, A., Schlef, M., Atzberger, C., Corsi, C., Cho, M. (2008). LAI and chlorophyll estimation for a heterogeneous grassland using hyperspectral measurements. *ISPRS Journal of Photogammetry and Remote Sensing* 63(4), 409-426.

Dixit, L., & Ram, S. (1985). Quantitative analysis by derivative electronic spectroscopy. *Applied Spectroscopy Review,* 21, 311-418.

Dowdy, S, Weardon, S. C., & Chilko, D. (Eds.). (2004). *Statistics for Research* (3rd ed.). New Jersey: John Wiley & Sons, Inc.

Efron, B. (1979). Bootstrap Methods: Another Look at the Jacknife. *Annual of statistics*, 7, 1-26.

Efron, B., & Tibshirani, R. (1993). *An Introduction of the Bootstrap.* Chapmanans Hall/CRC: New York.

Efron, B., & Tibshirani, R. (1997). Improvements on cross-validation: The 632+ Bootstrap method. *Journal of the American Statistical Association,* 91 (438), 548-560.

Elvidge, C.D. (1990). Visible and near infrared reflectance characteristics of dry plant materials. *International Journal of Remote Sensing*, 11 (10), 1775-1795.

Fayad, I., Baghdadi, N., Bailly, J., Barbier, N., Gond, V., Hajj, M.E., Fabre, F., & Bourgine, B. (2014). Canopy Height Estimation in French Guiana with LiDAR ICESat/GLAS Data Using Principal Component Analysis and Random Forest Regressions. *Remote Sensing*, 2014, 6, 11883-11914.

Fey, M. (2010). *Soils of South Africa*. Cape town: Cambridge University Press.

Field, C., & Mooney, H. A. (1986). The photosynthesis-nitrogen relationship in wild plants. In: T. J. Givinish, (Ed.), *On the economy of plant and function* (pp 25-55). Cambridge: Cambridge University Press.

Food and Agriculture Organisation of the United Nations. (1993). *World Soil Resources: An explanatory note on the FAO World Soil Resources Map at 1:25 000 000 scale*. Rome.

Fourty, T. H. & Baret, F. (1998). On spectral estimates of fresh leaf biochemistry. *International Journal of Remote Sensing*, 19 (7), 1283-1297.

Fox, J. (1997). *Applied regression analysis, linear models and related models.* London: Sage Publications.

Fox, J. (2002). Bootstrapping regression models. *Appendix to An R and S-PLUS Comparison to Applied Regression,* http://cran.r-project.org/doc/contrib/Fox-Companion/appendix-bootstrapping.pdf [Date accessed 10 September 2013].

Fox, J & Weisberg, S. (2010). *Bootstrapping Regression Models in R. An Appendix to An R Companion to Applied Regression*, (2nd ed). http://socserv.mcmaster.ca/jfox/Books/Companion/appendix/Appeendix-Bootstrapping.pdf [Date accessed 20 September 2013]

Freedman, J.H., & Peters, S.C. (1984). Bootstrapping a Regression Equation: Some Empirical Results. *Journal of the American Statistical Association,* 76 (385), 97-270.

Gao, B. C., & Goetz, A.F.H. (1994). Extraction of dry leaf spectral features from reflectance spectra green vegetation. *Remote Sensing of Environment*, 47(3), 369-374.

Gao, B.C., & Goetz, A.F.H. (1995). Retrieval of equivalent water thickness and information related to biochemical components of vegetation canopies from AVRIS data. *Remote Sensing of Environment,* 52(3), 155-162.

Geladi, P., & Kowalski, B.R. (1986). Partial Least Squares regression: a tutorial. *Annalytica Chemica Acta*, vol. 185, 1-17.

Grant, C.C., & Scholes, M.C. (2006). The importance of nutrient hot-spots in the conservation and management of large wild mammalian herbivores in semi-arid savannas. *Biological Conservation*, 130 (3), 426-437.

Grossman, Y.L., Ustin, S.L., Jacquemoud, S., Sanderson, E.W., Schmuch, G., & Verdebout, J. (1996). Critique of Stepwise mulitiple linear regression for the extraction of leaf biochemistry information for leaf reflectance data. *Remote Sensing of Environment,* 56, 182-193.

Guerschman, J.P., Hill, M,J., Renzullo, L.J., Barrett, D.J., Marks A.S., & Botha E.J. (2009). Estimating fractional cover of photosynthetic vegetation, non-photosynthetic vegetation and bare soil in the Australian tropical savannah region upscaling the EO-1 Hyperion and MODIS sensors. *Remote Sensing of Environment,* 113 (5), 928-945.

Habounde, D., Miller, J.R., Tremlay, N., Zarco, P.J., & Dextraze, L. (2002). Integrated narrow-band vegetation indices for prediction of crop chlorophyll content for application to precision agriculture. *Remote Sensing of Environment* 81, 416-426.

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

Herrman, I., Pimstein, A., Karnieli, A., Cohen, Y, Alchanatis, V., & Bonfil, J.D. (2010). Utilizing VEμS Red-Edge Bands For Assessing LAI in Crop Fields. *ISPRS Arcchive,* Vol. XXXVIII, Part 4-8-2-W9, 34-39.

Horneck, D.A., & Miller R.O. (1998). Determination of total nitrogen in plant tissue. In: Karla, Y.P. (Ed), *Handbook of Reference Methods for Plant Analysis.* CRC Press, New York, pp 75-83.

Horning, N. (2010). Introduction to decision trees and random forests. American Meseum of Natural History's Center for Biodiversity and Conservation.

Huang, Z., Turner, B.J., Dury, S.J., Wallis, I.R. & Foley, W.J. (2004). Estimating foliage nitrogen concentration from HYMAP data using continuum removal analysis. *Remote Sensing of Environment,* 93 (1-2), 18-29.

Http://glovis.usgs.gov

Islam, M.S., Bhuiya, M.S.U., Rahman S., & Husain M.M. (2009). Evaluation of SPAD and LCC based Nitrogen Management in Rice (*Oryza sativa* L.). *Bangladesh Journal of Agricultural Research,* 34(4): 661-672.

Jago, R.A., Cuttler, M.E.J., & Curran, P. J. (1999). Estimating canopy chlorophyll concentration from field and airborne spectra. *Remote Sensing of Environment,* 68, 217-224.

Jackson, R.D., & Huette, A.R. (1991). Interpreting vegetation indices. *Preventive Veterinary Medicine*, 11: 185-200.

Kampa, M., & Castanas, E. (2008). Human effects of air pollution. *Environmental Pollution*, vol 151, issue 2, 362-367.

Kancheva, R., & Borisova, D. (2007). Vegetation stress indicators derived from multispectral and multitemporal data. *Space Technology,* vol 26, 1-8. Lister Scence, Great Britain.

Kerle N., Jansen L.L.F., & Gerrit, C. (Eds). (2004). Principles of Remote Sensing (ITC Educational Textbook Series; 2) (3rd ed). ITC Enschede, The Netherlands.

Knox, N.M., Skidmore, A.K., Schlerf, M., De Boer, W.F., Van Wieren, S.E., Van Der Waal, C., Prins, H.H.T., & Slotow, R. (2010). Nitrogen prediction in grasses: effect of bandwidth and plant material state on absorption feature selection. *International Journal of Remote Sensing,* 31 (3), 691-704.

Knox, N. M., Skidmore, A.K., Prins, H.H., Asner, G.P., Van der Werf, H.M.A., De Boer, W. F., Van der Waal, C., De Knegt, H.J., Kohi E.M., Slotow, R., & Grant, R.C. (2011). Dry season mapping of savannah forage quality, using the hyperspectral Carnegie Airborne Observatory sensor. *Remote Sensing of Environment,* 115 (6), 1478-1488.

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.

Kumar, L., Schmidt, K. S., Dury, S., & Skidmore, A. K. (2001). Imaging Spectroscopy and Vegetation Science. In: F. D., Van Der Meer, & S. M De Jong, (Eds.) *Image Spectroscopy*. Dordrecht: Kluver Academic Publishers.

LABCONCO. (2012). *A Guide To Kjedahl Nitrogen Determination Methods and Apparatus: An industry Service Publication*.

LaCapra, V.C., Melack, J.M., Gastil, M., & Valeriano, D. (1996). Remote sensing of foliar chemistry of inundated rice with imaging spectrometry. *Remote Sensing of Environment* 55(1), 50-58.

Lang, S., & Blaschke T. (2006). Bridging remote sensing and GIS-What are the main supportive pillars?. *Paper presented at first International Conference on Object based Image*

*Analysis (OBIA 2006)*. http://www. Commission$.isprs.org/obia06 [Date accessed 20 April 2014].

Ludwig, F., De Kroon, H., & Prins, H.H.T. (2008). Impacts of savannah trees on forage quality for large African herbivore. *Oecologia,* 155,487-496.

Majeke, B., van Aardt, J.A.N., Cho, M.A. (2008). Imaging spectroscopy of foliar biochemistry in forestry environments. *Southern Forests,* 70(3), 275-285.

Markwell, J., Ostermann, J.C., Mitchell, J.L. (1995). Calibration of the Minolta SPAD-502 leaf chlorophyll meter. *Photosynthesis Research,* 46:467-471.

Martin, M.E., & Aber, J.D. (1995). Analysis of forest foliage III. Determining Nitrogen, Lignin, and Cellulose fresh leaves using near-infrared reflectance data. *Journal of Near-Infrared Spectroscopy* 2, 25-32.

Martin, M.E., & Aber, J. D. (1997). High spectral resolution remote sensing of forest canopy lignin, nitrogen, and ecosystem processes. *Ecological Applications,* 7 (2), 431-443.

McNaughton, S. J. (1990). Mineral nutrition and seasonal movements of African migratory ungulates. *Nature,* 345, 613-615.

Metler, C.A., & Vannata, R, A. (2002). *Advanced and Multivariate Statistical Methods: Practical Application and Interpretation, 2nd ed*. Pyrzak Publishing, LA.

Mitchel, J. (2010). Application in Hyperspectral and Lidar Remote Sensing to improve the Characterisation of Low Height Sparse Vegetation Ecosystem. Idaho State University.

Mooney, H. A., Vitousek, P. M., & Matson P. A. (1987). Exchange of materials between terrestrial ecosystems and the atmosphere. *Science* 238, 926-932.

Moorby, J., & Besford, R.T. (1983). Mineral nutrition & growth. In: A. Lauchi & R.L Bieleski (Eds), *Inorganic plant nutrition. Encyclopedia of Plant Physiology New Series* (pp 481-529). Berlin: Springer-Verlag.

Morrison, D. F. (1990). *Multivariate statistical methods, 3rd ed.* McGraw Hill, Inc, New York.

Mucina, L., & Rutherford, R.C. (2006). *The Vegetation of South Africa, Lesotho and Swaziland*. Cape town: Strelitzia.

Mucina, L. & Rutherford, R.C. (Eds). (2006-2010). (CD set). *The Vegetation of South Africa, Lesotho and Swaziland*. Pretoria: SANBI.

Mutanga, O., Skidmore, A.K., & Prins, H.H.T. (2004). Predicting in situ pasture quality in the Kruger National Park, South Africa, using continuum-removed absorption features. *Remote Sensing of Environment*, 89 (3), 393-408.

Mutanga, O., Prins, H.H.T., Skidmore, A.K., Van Wieren, S., Huizing, H., Grant, R., Peel, M., & Biggs H. (2004). Explaining grass nutrient patterns in Savanna rangeland of southern Africa. *Journal of Biogeography* 31, 819-829.

Mutanga, O., & Skimore, A. K. (2007). Red edge shift and biochemical content in grass canopies. *ISPRS Journal of Photogrammetry and Remote Sensing,* 62, 34-42.

Mutanga, O., van Aardt, J., & Kumar, L. (2009). Imaging spectroscopy (hyperspectral remote sensing) in the southern Africa: an overview. *South African Journal of Science* 105, 193 - 198.

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(2012), 399-406.

Mutanga, O., Adam E., Adjorloo, C., Abdel-Rahman, E.M. (2015). Evaluating the robustness of models developed from field spectral data in predicting African grass foliar nitrogen concentration using WorldView-2 image as an independent test dataset. *International Journal of Applied Earth Observation and Geoinformation* 34(2015), 178-187.

Nijmeijer, R., de Haas A., & Dost, R. (2001). *ILWS 3.0 Academic: User's guide*. International Institute for Aerospace Survey & Eearth Sciences (ITC).

Niinemets Ü. (2010). Responses of forest trees to single and multiple environmental stresses from seedling to mature plants: past stress history, stress interactions, tolerance and acclimation. *Forest Ecology and Management*,–260, 1623-1639.

Ollinger, S.V., Smith, M.L., Martin, M.E., Hallett, R.A., Goodale, C.L., & Aber, J.D. (2002). Regional variation in foliar chemistry and N cycling among forests of diverse history and composition. *Ecology,* 83, 339-355.

Ollinger, S.V., Richardson A.D., Martin M.E., Hollinger D.Y., Frolking, S.E, Reich P.B., Plourde, L.C., Katul G.G., Munger J.W., Oren R, Smith M.L., Paw K.T., Bolstad P.V., Cook B.D., Day M.C., Martin T.A., Monson R.K., & Schmid, H.P. (2008). Canopy Nitrogen, Carbon assimilation, and albedo in temperate and boreal forests: Functional relations and potential climate feedbacks. *PNAS*, vol 105. No 49, 19336-19341.

O'neal, M. E., Landis, D. A., & Isaacs, R. (2002). An Inexpensive, Accurate method for Measuring Leaf Area and Defoliation Through Digital Image Analysis. *Journal of Economic Entomology,* 95(6), 1190 – 1194.

Osborne, B.G., & Fean, T. (1986). *Near Infrared Spectroscopy in Food Analysis*. Longman, London.

Peñ-uelas, J., Gamon A., Merino F. J., & Field C. (1994). Reflectance indices associated with physiological changes in N-and water limited sunflower leaves. *Remote Sensing of Environment* 46:100-118.

Pomeranz, Y., & Moore, R.B. (1975). Reliability of several methods for protein determination in wheat. *Baker's Digest,* 49: 44-58.

Pontius, J., Hallet, R., Martin, M. (2005). Using AVIRIS to access hemlock abundance and early decline in the Catskills, New York. *Remote Sensing of the Environment, 97I*. 163- 173.

Prins, H. H. T., & Beekman, J. (1989). A balanced diet as a goal for grazing: The food of the manyara buffalo. *African Journal of Ecology,* 27, 241-259.

Prins, H.H.T., & Van Langevelde, F. (2008). Assembling diet from different places. In: H.H.T, Prins, & F, Van Langevelde (Eds.) *Resource Ecology: Spatial and Temporal Dynamics of Foraging.* Netherlands: Springer.

Pu, R.L., Gong, P., Biging, G.S., & Larrieu, M.R. (2003). Extraction of red edge optical parameters from Hyperion data for estimation of forest leaf area index. *IEEE Transaction on Geoscience and Remote Sensing,* 41(4), 916-921.

Qi, J., Chehbouni., Huete, A.R., Kerr, Y.H., & Sorooshian, S. (1994). A modified Soil Adjusted Vegetation Index. *Remote Sensing of the Environment, 48: 119 – 126.*

Ramoelo, A., Cho, M.A., Mathieu, R., Skidmore, A.K, Sclerf, M., Heitkönig, I.M.A., & Prins, H.H.T. (2011). Integrating environmental and in situ hyperspectral remote sensing variables for the grass nitrogen estimation in savannah ecosystems. *34th International Symposium on the Remote Sensing of Environment (ISRSE 2011), The GEOSS Era: Towards Operational Environmental Monitoring. Sydney, Australia*.

Ramoelo, A. (2012). *Savanna Grass Quality Remote Sensing Estimation from Local to Regional Scale*. University of Twente, ITC dissertation no. 270.

Ramoelo, A., Skidmore, A.K., Cho, M.A., Mathieu, R., Heitkonig, I.M.A., Dudeni-Thlone, N., Schlerf, M. & Prins, H.H.T. (2013). Non-linear partial least square regression increases the estimation accuracy of grass nitrogen and phosphorus using *in situ* hyperspectral and environmental data. *ISPRS Journal of Photogrammetry and Remote Sensing*,82, 27-40.

Ramoelo, A., Dzikiti, S., van Deventer, H., Maherry, A., Cho, M.A., & Gush M. (2014). Potential to monitor plant stress using remote sensing tools. *Journal of Arid Environments,* 113(2015) 134-144.

Rapideye. (2010). *RapidEye Standard Image Product Specification, Version 3.0*. Germany. www.rapideye.de [Date accessed 20 July 2014].

Regmi, A.P. (2006). *Optimize Nitrogen Application with a simple tool (Leaf clour chat)*. Bhairahawa: International Rice Research Institute (IRRI).

Reich, P.B., Walters M.B., & Ellsworth D.S. (1992). Leaf life-span in relation to leaf, plant, and stand characteristics among diverse ecosystems. *Ecological Monographs*, 63, 365-392.

Reich P,B., Ellsworth D.S., Walters M.B., Vose J.M., Gresham C., Volin J.C., & Bowman W.D. (1999). Generality of leaf trait relationships: a test across six biomes. *Ecology* 80, 1955-1969.

Sakamoto, Y., Ishiguro, M., & Kitagawa, G. (1986). *Akaike information criterion statistics*. The Netherlands: D. Reidel Publishing Company, Dordrecht.

Schlerf, M., Atzberger, C., Hill, J. (2010). Retrieval of chlorophyll and Nitrogen in Norway spruce (*Picea abies* L. Karst.) using imaging spectroscopy. *International Journal of Applied Earth Observation and Geoinformation* 12(1), 17-26.

Shah, T.H., Steven, M.D., & Clark, J.A. (1990). High resolution derivative spectra in remote sensing. *Remote Sensing of Environment,* 33, 55-64.

Shaw, G.A., & Burke, H.K. (2003). Spectral Imaging for Remote Sensing. *Lincoln Laboratory Journal,* vol 14:1, 1-28.

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 (1), 64-72.

Steudler, P.A., Bowden, R.D., Melillo J.M., & Aber J.D. (1989). Influence of Nitrogen fertilization on methane uptake in temperate forest soils. *Nature* 341, 314-316.

Steven M.D., Malthus, T.J., Demetriades-Shah, T.H., Danson, F.M., & Clarck, J.A. (1990). High-spectral resolution indices for crop stress. In: M,D Steven & J,A Clarck (Eds). *Applications of remote sensing in agriculture.* London: Butterworths.

Treydte, A.C., Heitkonig, I.M.A., Prins, H.H.T., & Ludwig, F. (2007). Trees improve grass quality for herbivores in African savannahs. *Perspectives in Plant Ecology, Evolution and Systematics*, 8 (4), 197-205.

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

Tucker, C.J. (1979). Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sensing of Environment*, 8**,** 50-127.

Turner, D.P., Cohen, W.B., Kennedy, R.E., Fassnacht, K.S., & Briggs, J.M. (1999). Relationships between leaf area index and Landsat TM spectral vegetation indices across three temperate zone sites. *Remote Sensing of Environment*, *70*, 52-68.

Uddling J., Gelang-Alfredsson, J., Piikki K., Pleijel, H. (2007). Evaluating the relationship between leaf chlorophyll concentration and SPAD-502 chlorophyll meter readings. *Photosynthesis Research,* 91: 37-46.

van Englen V.W.P., & Dijkshoorn J.A. (2013). Global and National Soils and Terrain Database (SOTER): Procedure Manual, version 2.0. *ISRIC Report 2013/04*.

Venter, F.J., Scholes, R.J. & Eckhardt, H.C. (eds.). (2003). *Abiotic template and its associated vegetation pattern,* London: The Island Press.

Vitousek, P.M., Mooney, H.A., Lubchenco, J., Mellilo, J.M. (1997). Human Domination of Earth's Ecosystems. *Science,* vol 277, 493-499.

Vos, J. & Bom, M. (1993). Hand-held chlorophyll meter: a promising tool to assess the nitrogen status of potato foliage. *Potato Research,* 36 (4), 301-308.

Vuolo, F., Neugebauer, N., Bolognesi, S.L., Atzberger, C., & D'Urso, G. (2013). Estimation of Leaf Area Index DEIMOS-1 Data: Application and Transferability of a Semi-Empirical Relationship between two Agricultural Areas. *Remote Sensing*, 2013, 5, 1274-1291.

Wang, Z.J., Wang, J.H., Liu, L.Y., Huang, W.J., Zao, C.J. & Wang, C.Z. (2004). The prediction of grain protein in winter wheat (*Triticum aestivum*) using plant pigment ratio (PPR). *Field Crops Research,* 90, 311-321.

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

Waterberg District Environmental Management Framework(WDEMF) -Draft EMF Report. (2010). Environomics Environmental Consultants, NRM Consulting, Metro GIS.

Well, K. V. (2010). *Object-based Segmentation and Classification of One Meter Imagery for Use in Forest Management Plans: A thesis submitted in partial fulfillment of requirements of MSc degree in Geography*. Logan, Utah: Utah State University.

Wessman, C. A. (1989). Evaluation of canopy biochemistry. In: R.J Hobbs, & H.A Mooney (Eds), *Remote Sensing of biosphere functioning* (pp 135 - 156). New York: Springer-Verlag.

Wessman, C. A., Aber J. D., Peterson, D. L., & Melilo J.M. (1988). Remote sensing of canopy chemistry and Nitrogen cycling in temperate forest ecosystems. *Nature* 335, 154-156.

Williams, P., & Norris, K (Eds). (1987). *Near-Infrared Technology in Agricultural and Food Industries*. St. Paul: American Association of Cereal Chemists.

Wright, I.J., Reich, P.B., Westoby, M., Ackerly, D.D., Baruch, Z., Bongers, F., Cavender-Bares, J., Chapin, T., Cornelissen, J.H.C., Diemer, M., Flexas, J., Garnier, E., Groom, P.K, Gulias, J., Hikosaka, K., Lamont, B.B., Lee, T., Lee, W., Lusk, C., Midgley, J.J., Navas, M., Niinemets, Ü., Oleksyn, J., Osada, N., Poorter, H., Poot, P., Prior, L., Pyankov, V.I., Roumet, C., Thomas, S.C., Tjoelker, M.G., Veneklaas, E.J & Villar, J. (2004). The worldwide leaf economics spectrum. *Nature* 428:821–827.

Wulder, M.A., Dymond, C.C., White, J.C., Leckie, D.G., & Carroll, A.L. (2005). Surveying mountain pine beetle damage of forests: a review of remote sensing opportunities. *Forest Ecology and Management*, 221, 27- 41.

Yoder, B.J. & Pettigrew-Crosby, R.E. (1995). Predicting nitrogen and chlorophyll content and concentrations from reflectance spectra (400-2500 nm) at leaf and canopy scales. *Remote Sensing of Environment*, 53 (3), 199-211.

Zheng, G., Tian, Q.J., Chen M.J., Ju, W.M., & Xia, X.Q. (2006). Combining remote sensing imagery and forest age inventory for biomass mapping. *International Journal of Remote Sensing*,10, 932-940.

Zheng, G., & Moskal, M.L. (2009). Retrieving Leaf Area Index (LAI) using Remote Sensing Theories, Methods & Sensors. *Sensors*, 2009, 9, 2719-2745.