SAVANNA GRASS QUALITY

REMOTE SENSING ESTIMATION FROM LOCAL TO REGIONAL SCALE

Abel Ramoelo

Examining committee:

Prof.dr. Wouter Verhoef	University of Twente
Prof.dr. Victor Jetten	University of Twente
Prof.dr. Martin Herold	University of Wageningen
Prof.dr. Paul Aplin	University of Nottingham, United Kingdom

ITC dissertation number 207
ITC, P.O. Box 6, 7500 AA Enschede, The Netherlands

ISBN 978-90-6164-331-9
Cover designed by Abel Ramoelo
Printed by ITC Printing Department

UNIVERSITY OF TWENTE.

FACULTY OF GEO-INFORMATION SCIENCE AND EARTH OBSERVATION

SAVANNA GRASS QUALITY

REMOTE SENSING ESTIMATION FROM LOCAL TO REGIONAL SCALE

DISSERTATION

to obtain
the degree of doctor at the University of Twente,
on the authority of the Rector Magnificus,
prof.dr. H. Brinksma,
on account of the decision of the graduation committee,
to be publicly defended
on Thursday 24th of May 2012 at 14.45 hrs

by

Abel Ramoelo

born on 02 November 1980

in Muduluni, South Africa

This thesis is approved by:

Prof.dr. Andrew K. Skidmore, **promoter**, University of Twente
Dr. Martin Schlerf, **assistant-promoter**, University of Twente
Dr.ir. Ignas M.A. Heitkönig, **assistant-promoter**, Wageningen University

Acknowledgements

I would like to acknowledge the Faculty of Geoinformation Science and Earth Observation (ITC), Council for Scientific and Industrial Research (CSIR), and National Research Foundation (NRF)-postgraduate development programme (PDP) in South Africa as well as Wageningen University for the funding of this PhD study. This study was part of the project called "Competing Claims on Natural Resources" led by Wageningen University. If it was not because of the concerted efforts provided by the above mentioned institutions, this PhD would have not been possible. Special thanks to CSIR-Natural resource and environment unit for providing me with the study leave for always going to the Netherlands to pursue many components of this study.

A very special and sincere gratitude goes to my promoter, Prof.dr Andrew Skidmore for his encouragement, motivation, scientific guidance and support. Andrew's optimistic approach in science instilled in me a future great scientist. Sometimes when I feel like I am not progressing, but Andrew recognized every effort I put into my work. I would like to thank my supervisors, Dr. Martin Schlerf and Dr. Ignas Heitkönig for helping me through the tough times in proposal development to the completion of the study. I also like to thank Dr. Renaud Mathieu for his supervision, support and guidance. I enjoyed several criticisms from my supervisors which always sharpen my scientific thinking. I have never taken any of your collective comments personally, but I took them as a toolkit to succeed as an independent scientist. I would also like to thank Dr.ir. Ignas Heitkönig and Dr.ir. Jan Clevers for translating the PhD summary to Dutch.

I would also like to thank the Earth Observation Group at CSIR, especially my mentor and friend Dr. Moses Cho for his commitment to build a scientific base for Africa, in the field of geoinformation science and earth observation. Moses's door is always open for many emerging young scientist (like me) in the group who wishes to be inspired and motivated by his positive attitude towards science. I would also like to extend my word of thanks to my other colleagues in Pretoria Lesley, Laven, Russell, Nobuhle, Oupa and Heidi. Thanks for your support.

I also like to thank NRS department for their support. We had very interesting interactions during Monday morning coffee, outings (November-December functions) and along corridors. Though I did not spend most of my PhD time in ITC, little time I spent with you was indeed interesting and inspiring. I would also like to thank Dr. Kees de Bie, for inspiring me through his supervision during MSc to be a hard working student. He also helped me out on the PhD proposal stage, incorporating the savanna biomass estimation section which because of time, I did not pursue. The ideas we had were really

interesting and I am willing to pursue some of these ideas during my Post-PhD activities. Our discussion during my proposal stage could form a basis for collaboration.

I also like to send my sincere gratitude to my fellow PhD students, especially in NRS who organized the tutorials, who afforded me a platform to discuss general and topical scientific issues. My former office mate; Fekerte, Baron, Dr. Jack Kuria, Xuanmei, Khammarul and my current office mates, Amjad, John, Ha, and now Li, I would also like to thank you for your support. The other ITC PhD student and colleagues to mention few, Armindo (a good friend), Anas, Atkilt, Sibu, Nthabi, Adrie, Babak, Aidin, Juan, Juan-Pablo, Mila, Sukhad, Dan (RIP), Tagel, Dawit, Vincent, Francis, Maitreyi, Priya, Anandita, Sabrina, Saleem, Ben, Dr Si Yali, Dr Mhosis Masocha, Dr Mobushir Khan and many more. I would also like to thank all PhD colleagues from the Competing Claims project, especially Armindo, Xavier and Petronella. I would also like to thank the "Competing Claims project" co-ordinators, Maja and Jens for facilitating various activities of the project.

I also like to thank people who might have contributed to my stay in ITC one way or the other: Esther, Loes, Saskia, Kim, Rebanna, Paul, Marion, Job, Benno, Dr. Tiejun and many more. I would also like to send my sincere appreciation to Mrs. Eva Skidmore for reading two of my papers, and I am sure her readings improved my writing skills.

I would like to send words of appreciation to South African National Parks (SANPARKS) for providing us a space to do research, and especially Dr. Izak Smith, Mrs. Thembi Khoza, Mrs Patricia Khoza, Mr. Adolf Manganyi, Mr. Steven Khoza, Mrs Onica Sithole, Mr Veli Ndlovu and many other guards who made my fieldwork possible. I would also like to thank Mr. Mike Glover and Dr. Jonathan Swart as well as their guards for making my fieldwork possible in Sabie Sands.

I would also like to thank traditional leaders "Induna" for Bushbuckridge area, especially for villages such as Lillydale, Huttington, Justicia, Areah, and their other neighbouring villages for accepting and allowing us to undertake successful series of fieldworks. In addition, I would like to thank Mr. Thulani Selaule, Mulalo, and Chris from Bushbuckridge area for fieldwork assistance.

I would also like to thank SADC community for always making gatherings and parties which made me always feel like I am at home. I am also indebted to the ITC Christian fellowship for providing a podium for praising and worshiping God.

I also want send my sincere gratitude to my family; my parents, siblings, aunts (especially Daisy), and grandmother for their support when I was away, taking care of my own family (my wife and kid).

I would like to send my sincere gratitude to these two important and wonderful people in my life, my wife Christinah and my son Shudufhadzo. I always left you alone, "how tough" Shudufhadzo learnt to walk, and speak while I was away…I promised to see his first mark from his first school exam. The only motivation I had was that my family understood that I need to pursue my studies to this level. Your support, love, prayers, encouragement, and motivations were the only energy I needed to finalize this work. I also thank God for protecting you while I was away. I might have not exhausted the list of the important people who made this PhD a success, but I thank you all, even if you are not mentioned above.

To my family

Table of Contents

List of figures

List of tables

List of Abbreviations

ANN	artificial neural network
ARC	agricultural research council, Nelspruit
ASD	analytical spectral devices, Inc.
ATCOR	atmospheric and topographic correction software
CR	continuum removal
CSIR	Council for Scientific and Industrial Research
DEM	digital elevation model
EVI	enhanced vegetation index
FD	first derivative
LAI	leaf area index
LOOCV	leave-one-out cross validation
KNP	Kruger National Park
MAP	mean annual precipitation
MAT	mean annual temperature
MCARI	modified chlorophyll absorption ratio index
MSAVI	modified soil-adjusted vegetation index
MTCI	MERIS terrestrial chlorophyll index
NDVI	normalized difference vegetation index
N	nitrogen
NIR	near infrared region of the reflectance
NRI	nitrogen reflectance index
OSAVI	optimized soil-adjusted vegetation index
P	phosphorus
PDP	post graduate development programme
PCA	principal component analysis
PLSR	partial least square regression
PPR	plant pigment ratio
R	reflectance
RBF-PLSR	partial least square regression with radial basis function
RDVI	renormalized difference vegetation index
REP	red edge inflation point
RMSE	root mean square error
RMSEcv	root mean square error of cross validation
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
SWIR	shortwave infrared region of the reflectance
TCARI	transformed chlorophyll absorption ratio index

TVI	triangular vegetation index
VIP	variable of importance for prediction
VIS	visible region of the reflectance
WR	water removal

Chapter 1

Introduction

1.1 Grass quality: satisfying food requirement

Spatial patterns of foliar nitrogen (N) and phosphorus (P) concentrations as indicators of grass quality are known to influence the grazing behaviour and migration patterns of wildlife and livestock in savanna landscapes (Drent and Prins, 1987; McNaughton, 1988, 1990; Seagle and McNaughton, 1992; Prins and van Langevelde, 2008). Grass quality entails attractants (e.g. minerals like N and P), and deterrents (e.g. polyphenols). Foliar N concentration is positively correlated to protein content (Clifton et al., 1994; Wang et al., 2004). Protein is one of the major nutrients for the herbivores (Prins and Beekman, 1989; Prins and van Langevelde, 2008). The importance of N supply on dry matter production as well as protein content is well documented in agricultural literature (Marschner, 1995), while grass N and P on the other hand, are important for reproduction and for lactating mammals (Agricultural Research Council, 1980; McNaughton, 1990; Prins and van Langevelde, 2008). In southern Africa, large herbivores are concentrated around nutrient rich areas e.g. termite mounds, sodic sites, or sites beneath large trees (Grant and Scholes, 2006; Treydte et al., 2007; Ludwig et al., 2008). Therefore, an accurate assessment of the spatial patterns of foliar N and P concentrations could play a vital role in the effective planning and management of savanna rangelands for sustainable livestock and wildlife grazing.

The foliar N: P ratio is postulated as one of the key indicators of nutrient limitation in several ecosystems (Koerselman and Meuleman, 1996; Ludwig et al., 2001; Prins and van Langevelde, 2008). Foliar N:P ratio reflects the balance of N and P supply which influence plants at all levels, i.e. the growth and reproduction of individual plants, plant species interactions, composition and diversity (Güsewell, 2004; Cech et al., 2008). Differences in foliar biochemistry as captured by N: P ratio in plants provide information on which nutrient is limiting and could determine the plant production or biomass which could eventually affect the feeding activity of herbivores in a particular ecosystem (Koski, 1999; Daufresne and Loreau, 2001; Prins and van Langevelde, 2008). "Smaller herbivores are known to be more N-limited than larger ones and not vice versa" (Prins and van Langevelde, 2008), they require nutrient balance for growth and to carry out daily activities. Limnetic studies hypothesized that herbivores (cladocerans) with high P concentration should be generally P limited, while herbivores with more N (e.g. copepods) should be N limited (Koski, 1999; Daufresne and Loreau, 2001). This provides a scope for testing such a hypothesis in the savanna ecosystems. This tool has been widely used in the limnetic related studies (Koerselman and Meuleman, 1996; Güsewell et al., 2003; Güsewell, 2004). A review of 40 fertilization studies discovered that the N: P ratio value <14 indicates N limitation, between 14 and 16 indicates co-limitation and >16 indicates P limitation (Koerselman and Meuleman, 1996; Ludwig et al., 2001; Güsewell, 2004). For savanna ecosystem, Cech et al. (2008) reported that the N: P values < 9 indicate N-

limitation, >10 indicate P-limitation, and values at a range 9-10 indicate co-limitation. Ludwig et al. (2001) found that below trees, N: P ratio value of 12 indicates that P is limiting, while on the open grassland the average of 6 indicates N limiting in savanna ecosystems. Information on the landscape variability of nutrient concentrations and ratios/limitation could aid in understanding the feeding patterns of herbivores.

Savanna ecosystem play a crucial role in the rural economy worldwide (James et al., 2003). Among other things, they provide grazing land important for rural livestock production (Shackleton et al., 2002; Steinfeld et al., 2006). Livestock production is one of the main sources of income in the rural areas (Shackleton et al., 2002; Steinfeld et al., 2006). Savannas are subject to land degradation as a result of overgrazing and overstocking of domestic animals (Abel and Blaikie, 1989; Du Toit and Cumming, 1999), this can be caused by lack of information about the grass condition which hampers proper management. Therefore, there is a need for sustainable utilization of the grazing land for viable livestock production, while minimizing land degradation. Spatial information about grass nutrients could be useful to guide farmers in sustainable management of their grazing land, while minimizing land degradation.

1.2 Estimation and mapping of grass quality using remote sensing

Remote sensing is widely used as a cost-effective means (Mumby et al., 1999) to estimate and map savanna grass quality at landscape level in various biomes, such as grasslands and savannas (Mutanga and Skidmore, 2004a; Mutanga et al., 2004c, 2004b; Bogrekci and Lee, 2005; Ferwerda et al., 2005; Mutanga et al., 2005; Mutanga and Kumar, 2007; Numata et al., 2008; Skidmore et al., 2010), forests (Martin and Aber, 1997; Schlerf et al., 2010) and agricultural areas (LaCapra et al., 1996; Thenkabail et al., 2000; Hansen and Schjoerring, 2003; Huang et al., 2004; Zarco-Tejada et al., 2004b; Wang et al., 2009). The remote sensing techniques used to retrieve vegetation parameters involve the use of indices (Daughtry et al., 2000; Haboudane et al., 2002; Haboudane et al., 2004), absorption features (Skidmore et al., 2010; Knox et al., 2011), full spectrum (Yoder and Pettigrew-Crosby, 1995; Asner and Martin, 2008; Ramoelo et al., 2011b), integrated modeling (Cho et al., 2009; Knox et al., 2011), which are empirical based and an inversion of physically based or radiative transfer models (Zarco-Tejada et al., 2004a; Schlerf and Atzberger, 2006; Darvishzadeh et al., 2008a). The latter approach has been successfully applied for retrieving leaf area index (LAI) and chlorophyll (Wenjiang et al., 2004; Zarco-Tejada et al., 2004a; Darvishzadeh et al., 2008a). Limited studies directly retrieved foliar N and P using inversion of radiative transfer models as they are not incorporated as one of the

parameters (Kokaly et al., 2009). An attempt to incorporate foliar N concentrations was done for the needleleaf simulations using LIBERTY model (Dawson et al., 1998), a potential was observed but estimation results showed high covariance between N and other leaf properties (Kokaly et al., 2009). Foliar N for example, could be estimated using inversion of radiative transfer models assuming a positive relationship between N and chlorophyll (Kokaly et al., 2009; Yoder and Pettigrew-Cosby, 1995), but this relationship is season specific (Wang et al., 2009; Wenjiang et al., 2004).

The conventional broadband remote sensing techniques based on the utilization of the relationship between grass quality (e.g. N and P) and spectral indices such as normalized difference vegetation index (NDVI) (Tucker, 1979), soil line concept (SLC), simple ratio (SR) (Baret and Guyot, 1991), soil-adjusted vegetation index (SAVI)(Huete, 1988) have limited applications in high grass canopy cover as they saturate (Tucker, 1977; Mutanga and Skidmore, 2004b). On the other hand, the use of spectral indices derived from the red-edge bands (700 – 750 nm) of hyperspectral or narrow-band data has been demonstrated to mitigate the saturation effect observed with broadband indices (Clevers et al., 2002; Huang et al., 2004; Cho and Skidmore, 2006; Darvishzadeh et al., 2008b; Majeke et al., 2008). The red-edge is the region of abrupt change in foliar reflectance between 680 and 780 nm (Clevers et al., 2002). Narrow-band normalized difference vegetation index and SR indices computed from red-edge bands provided more accurate estimates of foliar N compared to conventional NDVI derived from 680 and 800 nm (Mutanga and Skidmore, 2007). Studies on estimation of foliar N using vegetation indices such as red edge position, depend mainly on chlorophyll concentration (Clevers et al., 2002; Mutanga et al., 2004c; Cho and Skidmore, 2006; Numata et al., 2008), assuming a positive correlation between foliar N and chlorophyll concentrations (Vos and Bom, 1993; Yoder and Pettigrew-Crosby, 1995). This approach is limited because it is dependent on the plant phenology, meaning that the relationship will deteriorate as leaves senesce (Wang et al., 2009).

Spectral absorption features identified to be sensitive to foliar biochemical concentrations are listed in Curran (1989), Kumar et al. (2001) and Knox (2011). Most of the features used for foliar biochemical estimation dominates the near infrared (NIR) and shortwave infrared (SWIR) (Curran, 1989; Elvidge, 1990; Kokaly and Clark, 1999; Kumar et al., 2001; Cho et al., 2010; Knox et al., 2010b; Skidmore et al., 2010). For example, N has absorption features centred at 430 nm, 460 nm, 640 nm, 660 nm, 910 nm, 1510 nm, 1940 nm, 2060 nm, 2180 nm, 2300 nm, 2350 nm, dominating the SWIR region (Curran, 1989). The main leaf biochemical concentrations absorbing in the SWIR region (1000-2500 nm) include lignin, cellulose, starch and proteins (Curran, 1989; Kokaly and Clark, 1999). Specific absorption features for P have not been identified, but several studies found that the short-wave

infrared (SWIR) bands have potential for predicting foliar P concentration (Mutanga and Kumar, 2007; Cho et al., 2010; Ramoelo et al., 2011b). However, the accuracy of the estimation of foliar N and P using NIR and SWIR absorption features is highly influenced by the absorption of leaf water content, masking the subtle absorption features of other biochemical concentrations (Gao and Goetz, 1994, 1995; Fourty and Baret, 1998). Several techniques have been used to minimize the effect of leaf water content on the remote sensing of foliar biochemical estimation, including spectral transformation such as continuum-removed spectra, first derivative spectra and Log-transformed spectra (Majeke et al., 2008; Kokaly et al., 2009). These spectral data enhance absorption features of foliar biochemical concentrations, while minimizing atmospheric, soil background, and water absorption effects, as well as data redundancy (Yoder and Pettigrew-Crosby, 1995; Dawson and Curran, 1998; Cho and Skidmore, 2006). For example, Log(1/R) is preferred to reflectance because it is linearly related to absorbing components (Hruschka, 1987; Yoder and Pettigrew-Crosby, 1995). Studies such as Yoder and Pettigrew-Crosby (1995) showed a strong relationship between Log(1/R), as well as the first derivative Log(1/R)', and foliar N concentration. Fourty and Baret (1998) argued that transforming reflectance into their corresponding absorbance values improved the accuracy of estimates. Continuum removal has also been applied to enhance absorption features for foliar N and P estimations (Kokaly and Clark, 1999; Curran et al., 2001; Mutanga et al., 2005).

Water affects the absorption features for many foliar biochemicals when using fresh leaf spectra , and the removal of these effects has been recommended to increase the accuracy of foliar biochemical estimation in the SWIR (Gao and Goetz, 1994, 1995; Dawson et al., 1998; Kokaly and Clark, 1999; Mutanga and Skidmore, 2004a; Zhao et al., 2006). Absorption by these chemicals, including lignin, starch, protein, and cellulose, is not very strong (weak absorbers) and so is generally masked by water absorption in fresh leaves (Kumar et al., 2001; Zhao et al., 2006). However, in dry leaf spectra, foliar biochemical absorption is generally highly differentiated and well correlated to the concentrations of foliar biochemicals (Card et al., 1988; Elvidge, 1990). Encouraging results have been attained using spectral transformations, e.g. normalized band depth (Kokaly and Clark, 1999; Curran et al., 2001; Mutanga et al., 2004c). However, leaf water still poses a challenge when using fresh leaf spectra to estimate biochemical concentrations (Fourty and Baret, 1998; Johnson, 2001). To overcome the masking effect of leaf water, Gao and Goetz (1994, 1995) successfully removed water absorption effects from fresh leaf spectra to estimate leaf components such as lignin and cellulose. They developed a non-linear least-squares spectral matching technique that calculates a fresh leaf spectrum as a non-linear combination of a leaf water spectrum and a dry matter spectrum. In a follow-up study, Schlerf et al.

(2010) modified the technique and applied it successfully to estimate nitrogen concentrations in Norwegian spruce needles and named it the water removed approach (WR). This study was undertaken to further adapt and test this technique in estimating foliar N and P concentrations from the spectra of grass species commonly found in savanna ecosystems.

Savanna ecosystems are diverse and heterogeneous in soil and plant moisture, soil nutrients, fire regime, grazing pressures and anthropogenic activities (Ben-Shahar and Coe, 1992). This makes the estimation of grass N and P using remote sensing in this ecosystem a challenging task (Mutanga et al., 2004c; Mutanga and Kumar, 2007; He and Mui, 2010; Skidmore et al., 2010). Grass quality is influenced by geology (Bell, 1986; Ben-Shahar and Coe, 1992; Grant and Scholes, 2006), soil (Heitkönig and Owen-Smith, 1998; Cho et al., 2010), precipitation and temperature (Ben-Shahar and Coe, 1992), topography or catena position (Seagle and McNaughton, 1992; Mutanga et al., 2004a) as well as aspect (Mutanga et al., 2004a) and land use types (Zhou et al., 2002). We assume that a modeling approach that exploits the strength of environmental variables and remote sensing data could potentially improve the assessment of ecosystem state and functioning at various geographic scales (Mutanga et al., 2004a; Cho et al., 2009; Knox et al., 2011). The integrated approach could be an attempt towards estimating and mapping foliar N and P at regional scale, which according to our knowledge is yet to be done. A limited number of studies have investigated the possibility of integrating environmental and remote sensing variables to estimate foliar N and P concentrations, e.g. Cho et al. (2009) and Knox et al. (2011).

Foliar biochemical concentrations are seldom mapped at regional scale because of the limited spaceborne sensors that can sample radiance in the region of the red-edge or full spectrum. For example, conventional multispectral satellite sensors such as SPOT, Landsat, MODIS and ASTER lack the specific spectral regions such as red edge for biochemical estimations especially N and their spatial resolutions are relatively coarse. A sensor like MERIS has standard band settings which allow the computation and approximation of the red edge position (Clevers et al., 2002), but the spatial resolution is relatively coarse. The emergence of the multispectral sensors such as WorldView-2 (USA), SumbandilaSAT (South Africa) and RapidEye (Germany) with the red-edge band at very high spatial resolution (i.e. 6.5 m) could provide an opportunity for rangeland resource quality assessment at regional level. Using remote sensing data from these latest sensors, there is a need for the development of the specific vegetation indices that could be used successfully with these sensors. In this study several broad-band and hyperspectral vegetation indices were modified and new indices were developed to incorporate the red-edge band to estimate grass N concentration

at regional scale using RapidEye images of the period during peak productivity of the grass.

A challenge for using remote sensing to estimate foliar biochemicals is associated with the difficulty to disentangle the signal of biomass and foliar biochemicals, especially N (i.e. the interaction effects between N and biomass) (Skidmore et al., 2010). The effects can be minimized during peak productivity where the grass spectra have a high absorption in the red region and high scattering in the near infrared region (Plummer, 1988b; Skidmore et al., 2010). During this period, the scattering and absorption on the spectra continue to increase as captured by indices such as normalized difference vegetation index (NDVI), and the relationship between biomass and NDVI asymptotically saturate (Tucker, 1977; Mutanga and Skidmore, 2004b). At a certain critical biomass point (e.g. 3000 g/m^2), the vegetation indices are unable to estimate biomass (Mutanga and Skidmore, 2004b). That is when the foliar N can be estimated with minimal effects of the N-biomass interaction problem. Otherwise, the foliar N estimation using indices could be compromised by the variation in grass biomass. After this interaction problem is minimized, various multivariate statistical techniques are available to retrieve foliar biochemical concentrations.

Several studies have applied stepwise multiple linear regression (SMLR) (Grossman et al., 1996; Martin and Aber, 1997; Kokaly and Clark, 1999; Huang et al., 2004) to estimate N and P with hyperspectral remote sensing variables. However, SMLR operates on the assumption of normal distribution of the data, and could suffer from model overfitting and multicollinearity (Grossman et al., 1996; Huang et al., 2004). The use of partial least square regression (PLSR) has been advocated to address these issues (Geladi and Kowalski, 1986; Huang et al., 2004; Asner and Martin, 2008; Darvishzadeh et al., 2008b; Ramoelo et al., 2011b). On the other hand, conventional PLSR also makes a normality assumption about the distribution of the response variable. Input data can be normalized using mean or median centring prior to use with the conventional PLSR (Viscarra Rossel, 2008), but this does not completely address the requirement for normal distribution. To curtail this assumption, we introduce the non-linear PLSR, also known as PLSR with radial basis function neural network (RBF-PLSR) (Walczak and Massart, 1996) for foliar biochemical estimations. The advantage of the non-linear PLSR is that it is a flexible non-linear regression technique which combines the capability of the conventional PLSR, i.e., power to maximize covariance between data sets, and the non-linear nature of the RBF neural network (Walczak and Massart, 1996). The predictive models developed by non-linear PLSR have limited or no overfitting and multicollinearity problems if the optimal number of latent variables is selected (Walczak and Massart, 1996). A non-linear PLSR is also non-parametric in nature and it does not require

model input to be normally distributed. Non-parametric techniques such as artificial neural network (ANN) were also applied for retrieving foliar biochemicals using hyperspectral data (Mutanga and Skidmore, 2004a; Skidmore et al., 2010; Knox et al., 2011). The ANN is a black-box and is computational intensive (Atkinson and Tatanall, 1997; Skidmore et al., 1997), but combined with PLSR, it proved to be useful in several domains. The non-linear PLSR has been successfully applied in soil (Fidêncio et al., 2002), time series prediction (Zemouri et al., 2003), air pollution (Giering et al., 2005) and engineering related fields (Garg et al., 2010). The performance of the non-linear PLSR has not been established for extracting vegetation biochemistry in the heterogeneous savanna ecosystems.

1.3 Research Aim and Objectives

The aim of this study is to develop and improve techniques for estimating and mapping foliar biochemicals using remote sensing from local to regional scale. The specific objectives are as follows; (1) to test water removed spectra for foliar N and P estimation as compared to the existing spectral techniques, (2) to estimate foliar N: P ratio using *in situ* hyperspectral measurements or field spectroscopy, (3) to investigate the applicability of the non-linear partial least square regression in integrating *in situ* hyperspectral and environmental variables to estimate foliar N and P concentrations, and (4) to investigate the utility of the red edge band in RapidEye data for estimating foliar and canopy N at regional scale.

1.4 Research Outline

This thesis consists of six chapters including the Introduction and Synthesis. Analyses of this study were based on remote sensing measurements at greenhouse/laboratory, field and spaceborne levels. This thesis was organized to address two main issues pertaining to the estimation of foliar biochemicals; (1) water absorption effects removal and (2) regional estimation and mapping. The greenhouse/laboratory level of measurement was undertaken to test applicability of the water removal technique for estimating foliar N and P in a controlled environment or greenhouse. This level was chosen since soil nutrients and water can be easily manipulated to influence variability of foliar biochemicals. Water removal was applied in two chapters (i.e. 2 and 3). For regional estimation, we had two chapters, chapter 4, was an attempt towards regional nutrient estimation using integrated modelling approach. Chapter 4 was a step before chapter 5, demonstrating the possibility of using integrated modelling approach to estimate foliar biochemicals. Finally, chapter five focused on regional estimation of foliar and canopy N. We did not estimate P at regional scale because we concentrated only on indices derived in the visible spectrum using RapidEye which are more sensitive to foliar N than P. The chapters

were also organized following remote sensing measurement levels, and the levels are highlighted in **BOLD** below, especially for the outline of each technical chapter.

Chapter 1:
The chapter introduces research problem, aim and objectives. It introduces the importance and ecological relevance of estimating foliar N and P concentrations. The chapter further introduces progress and challenges faced by remote sensing analysts in estimating foliar biochemicals, with emphasis on water absorption effects and regional estimation of foliar biochemicals.

Chapter 2:
The objective was to test the applicability of water removed spectra in estimating foliar N and P concentrations. This chapter was based on the **greenhouse** experiments so that the environment could be easily controlled. We determined and manipulated the soil nutrients and water to influence foliar N and P variation in the plant (i.e. grass species: *Digitaria eriantha*). Spectral transformation and statistical techniques were applied to retrieve foliar N and P concentrations. Water removed spectra produced promising results.

Chapter 3
The objective was to test the utility of *in situ* hyperspectral measurements to estimate foliar N: P ratio on grass canopies. The study was conducted at **field level** applying *in situ* hyperspectral measurements using Analytical Spectral Device (ASD), FieldSpec 3®. To retrieve foliar N: P ratio, spectral transformation techniques and partial least square regression were investigated. Among the spectral transformation techniques, water removed spectra was tested. The conclusion was that hyperspectral data shows a potential to estimate foliar N: P ratio, where water removed and continuum removed spectra achieved a higher accuracy.

Chapter 4
This chapter presents an attempt towards estimating foliar N and P concentrations at regional scale, using integrated modelling approach. It was conducted at **field level** using *in situ* hyperspectral measurements collected by Analytical Spectral Device (ASD), FieldSpec 3®. We introduced a statistical technique called the non-linear PLSR for estimating foliar biochemicals. The study investigated an estimation of foliar biochemicals using the integrated modelling approach which combines the use of hyperspectral measurements and environmental or ancillary variables, which was an attempt towards regional estimation of foliar biochemicals.

Chapter 5
The objective was to investigate the utility of the red edge band embedded in the RapidEye sensor to estimate foliar and canopy N. Chapter was based on **spaceborne or satellite** remote sensing measurements using the RapidEye sensor. Vegetation indices were computed and an integrated modelling approach was applied to estimate foliar N and canopy N. The integrated modelling approach combines vegetation index and environmental or ancillary variables to estimate foliar and canopy N. This was an attempt to estimate and map grass N at regional scale using multispectral remote sensing in the savanna ecosystems.

Chapter 6
This chapter entails synthesis of various issues emanated from this thesis, including water removal, integrated modelling, biomass and foliar biochemical interaction, phenology, ecology and issues of scale. It further includes conclusions and recommendations.

Chapter 2

Grass nitrogen and phosphorus estimation using water removed spectra*

*Chapter is based on: Ramoelo, A. Skidmore, A.K. Mathieu, R. Heitkönig, I.M.A. (2011) Water removed spectra increase the retrieval accuracy when estimating savanna grass nitrogen and phosphorus concentrations, *ISPRS Journal of Photogrammetry and Remote Sensing*, 66, pp: 408-417.

Abstract

Information about the distribution of grass foliar nitrogen (N) and phosphorus (P) is important to understand rangeland vitality and to facilitate the effective management of wildlife and livestock. Water absorption effects in the near infrared (NIR) and shortwave infrared (SWIR) region pose a challenge for nutrient estimation using remote sensing. The aim of this study was to test the utility of water removed (WR) spectra in combination with partial least square regression (PLSR) and stepwise multiple linear regression (SMLR) to estimate foliar N and P, compared to spectral transformation techniques such as first derivative, continuum removal and log transformed spectra (Log(1/R)). The study was based on a greenhouse experiment with a savanna grass species (*Digitaria eriantha*). Spectral measurements were made using a spectrometer. *D. eriantha* was cut, dried and chemically analyzed for foliar N and P concentrations. WR spectra were determined by calculating the residual from the modelled leaf water spectra using the non-linear spectral matching technique and observed leaf spectra. Results indicated that WR spectra yielded a higher N retrieval accuracy than a traditional first derivative transformation (R^2=0.84, RMSE=0.28%) compared to R^2=0.59, RMSE=0.45% for PLSR. Similar trends were observed for SMLR. The highest P retrieval accuracy was derived from WR spectra using SMLR (R^2=0.64, RMSE=0.067%), while the traditional first derivative and continuum removal resulted in lower accuracy. Only when using PLSR did the first derivative result in a higher P retrieval accuracy (R^2=0.47, RMSE=0.07%) than the WR spectra (R^2=0.43, RMSE=0.070%). It was concluded that the water removal technique could be a promising technique to minimize the perturbing effect of foliar water content when estimating grass nutrient concentrations.

2.1 Introduction

Information about the distribution of grass nutrient concentration is crucial to understand rangeland health and facilitates effective management of wildlife and livestock. Grass nutrient concentrations influence the feeding patterns and distribution of wildlife and livestock species in savanna rangelands (McNaughton, 1988, 1990; McNaughton and Banyika, 1995). Large herbivores are known to concentrate in nutrient rich sites in Southern Africa, e.g. termite mounds, sodic sites, or sites beneath large trees (Owen-Smith and Danckwerts, 1997; Grant and Scholes, 2006; Treydte et al., 2007). Furthermore, herbivore diversity increases with increasing soil fertility (Ollf et al., 2002) and foliar nutrient concentration generally correlates positively with soil nutrient levels (Penning de Vries and Djiteye, 1982). Of the nutrients, foliar nitrogen (N) and phosphorus (P) concentrations are particularly important for herbivores (McNaughton, 1990). Mapping both N and P would allow computing of the N:P ratio, which is a key indicator of nutrient limitation in vegetation (Koerselman and Meuleman, 1996; Ludwig et al., 2001). Estimating N and P could therefore provide information on which nutrient is limiting for wildlife and livestock production in a particular landscape (Prins and van Langevelde, 2008).

In order to identify where foliar N and P become important, hyperspectral remote sensing has been employed in various biomes, such as grasslands and savannas (Bogrekci and Lee, 2005; Ferwerda et al., 2005; Mutanga et al., 2005; Mutanga and Kumar, 2007; Numata et al., 2008; Skidmore et al., 2010), forests (Martin and Aber, 1997; Schlerf et al., 2010) and agricultural areas (LaCapra et al., 1996; Thenkabail et al., 2000; Zarco-Tejada et al., 2004b; Wang et al., 2009). Most of the spectral absorption features that have been identified and used for N and P estimation are located in the near infrared (NIR) and shortwave infrared (SWIR). For example, N has absorption features centred at 430 nm, 460 nm, 640 nm, 660 nm, 910 nm, 1510 nm, 1940 nm, 2060 nm, 2180 nm, 2300 nm, 2350 nm, dominating in the SWIR region (Curran, 1989). The main leaf biochemicals absorbing in the SWIR region (1000-2500 nm) include lignin, cellulose, starch and proteins (Curran, 1989; Kokaly and Clark, 1999; Kumar et al., 2001). However, the accuracy of the estimation of N and P using NIR and SWIR absorption features is highly influenced by the reflectance of leaf water content, masking the subtle absorption features of other biochemicals (Gao and Goetz, 1994, 1995; Fourty and Baret, 1998).

Several techniques have been used to minimize the effect of leaf water content on the remote sensing of foliar biochemicals, including spectral transformation such as vegetation indices, continuum-removed spectra, first derivative spectra and log-transformed spectra. Studies have estimated N using vegetation indices such as red edge position, which depends mainly on chlorophyll

concentration (Clevers et al., 2002; Mutanga et al., 2004c; Cho and Skidmore, 2006; Numata et al., 2008), assuming a positive correlation between leaf N and leaf chlorophyll concentration (Vos and Bom, 1993; Yoder and Pettigrew-Crosby, 1995). This approach is limited as it depends on the leaf or plant phenology, meaning that the relationship will deteriorate during leaf senescence (Wang et al., 2009).

Derivatives, continuum-removal, and log transformed spectra (Log(1/R)) enhance absorption features of foliar biochemicals, while minimizing atmospheric, soil background, and water absorption effects, as well as data redundancy (Yoder and Pettigrew-Crosby, 1995; Dawson and Curran, 1998; Cho and Skidmore, 2006). For example, Log (1/R) is preferred to reflectance because it is linearly related to absorbing components (Hruschka, 1987; Yoder and Pettigrew-Crosby, 1995). Yoder and Pettigrew-Crosby (1995) showed a strong relationship between Log (1/R), as well as the first derivative Log (1/R)', and N concentration. Fourty and Baret (1998) argued that transforming reflectance into their corresponding absorbance values improved the accuracy of estimates. Continuum removal has also been successfully applied to enhance absorption features for foliar N and P estimations (Kokaly and Clark, 1999; Curran et al., 2001; Mutanga et al., 2005).

Water affects the absorption features for many foliar biochemicals when using fresh leaf spectra , and the removal of these effects has been recommended to increase the accuracy of foliar biochemical estimation in the SWIR (Gao and Goetz, 1994, 1995; Dawson et al., 1998; Mutanga and Skidmore, 2004a; Zhao et al., 2006) . Absorption by these biochemical concentration is not very strong and is generally masked by water absorption in fresh leaves (Kumar et al., 2001; Zhao et al., 2006). In dry leaf spectra, biochemical absorption is generally highly differentiated and well correlated to the concentrations of these chemicals (Card et al., 1988; Elvidge, 1990). Spectral transformations like normalized band depth (Kokaly and Clark, 1999; Curran et al., 2001; Mutanga et al., 2004c) improved the results, but leaf water still poses a challenge when using fresh leaf spectra to estimate biochemical concentrations (Fourty and Baret, 1998; Johnson, 2001).

To overcome the masking effect of leaf water, Gao and Goetz (1994; 1995) successfully removed water absorption effects from fresh leaf spectra to estimate leaf components such as lignin and cellulose. They developed a non-linear least-squares spectral matching technique that calculates a fresh leaf spectrum as a non-linear combination of a leaf water spectrum and a dry matter spectrum. Schlerf et al. (2010) modified this technique and applied it successfully to estimate nitrogen concentrations in Norwegian spruce needles and named it the water removed approach (WR).

Although stepwise multiple linear regression (SMLR) has been successfully used in foliar biochemical estimations, it has some limitations including multicollinearity, linear relationship assumptions, over-fitting (Curran, 1989; Martens and Naes, 2001) and difficulty in transferring the predictive models to other data sets (Grossman et al., 1996). Many studies recommend the use of partial least square regression (PLSR) (William and Norris, 1987; Hansen and Schjoerring, 2003; Cho et al., 2007b; Asner and Martin, 2008). With PLSR the spectra are decomposed into latent factors using the response variable to reduce the data dimensionality problem in the model development process (reflectance spectra typically have a large number of bands constituting as many independent variables) (Geladi and Kowalski, 1986; Geladi et al., 1999). The utility of PLSR in foliar biochemical estimation for N and P has been demonstrated, and both regression techniques (SMLR and PLSR) with various transformed spectra have been successfully applied (Huang et al., 2004; Asner and Martin, 2008). However, the performance of PLSR, SMLR and WR spectra when estimating foliar P and N remains to be established.

The main aim of this study was to test the utility of water removed spectra (WR) in combination with PLSR and SMLR for estimating foliar N and P, and compares this to other existing spectral transformation techniques such as Log(1/R), first derivative (FD), continuum removal (CR) and also the original reflectance (R). The study intended to quantify the retrieval accuracy of foliar N and P of a typical savanna grass grown under controlled conditions in a greenhouse. Reflectance spectra were collected using a visible – SWIR spectrometer. The hypothesis was that water removed spectra significantly increased the retrieval accuracy of nutrients, compared to first derivative of reflectance or other spectral transformations.

2.2 Material and Methods

A grass species (*Digitaria eriantha*) was sown in pots and grown for four (4) months in a greenhouse (Figure 2.1). Soil water content as well as N and P fertilization were adjusted to produce high variation in foliar water, N and P. Reflectance measurements were acquired with a spectrometer. Foliar N and P concentrations were analyzed in the laboratory. Various spectral transformation techniques were applied including WR spectra using PLSR and SMLR and a bootstrapping approach for validation. Figure 2.1 shows a flowchart with various sections of the study method.

2.3 Greenhouse Experiments: Set up and Sampling

The experiment was set up to produce high variation in foliar N and P concentrations, and foliar water content. A multiple factorial design was used (2x3x3) with 5 replications to ensure that each water treatment included at least 30 samples, generating a total of 90 samples (Morrison, 2001). The

selection of the grass species (*Digitaria eriantha*) was based on its wide occurrence in African savanna ecosystems and its importance as forage for livestock and wild herbivores. *D. eriantha* is a perennial grass species which grows either as a dense tussock, with or without extended stolons or as continuous stoloniferous sward. It can grow in a wide range of soil types from sands to heavy clays. Natural compost was mixed with fine red and sandy soils to form a basic stratum for sowing the grass seeds. Samples of this soil mix were taken to South Africa's Agricultural Research Council (ARC) for chemical analysis to determine its chemical composition. Then various levels of nutrients, based on Venter (1990) as well as Scholes and Walker (2004), were added. Limestone ammonia nitrate (N: 284 g/kg) and superphosphate (P: 83g/kg) were used for fertilization. Soil water levels were manipulated using different watering regimes, when the grass was fully grown and flowering. The various levels of water treatment applied were: (i) high level, where plants were watered twice a day, (ii) medium level, with plants watered once every two days and (iii) low level, with plants watered once or twice a week. This was done for two weeks to ensure contrasting levels of water content, P and N concentrations in the soil, and hence in the grass leaves. The temperature was kept between 27 and 30^0C to mimic savanna ecosystems.

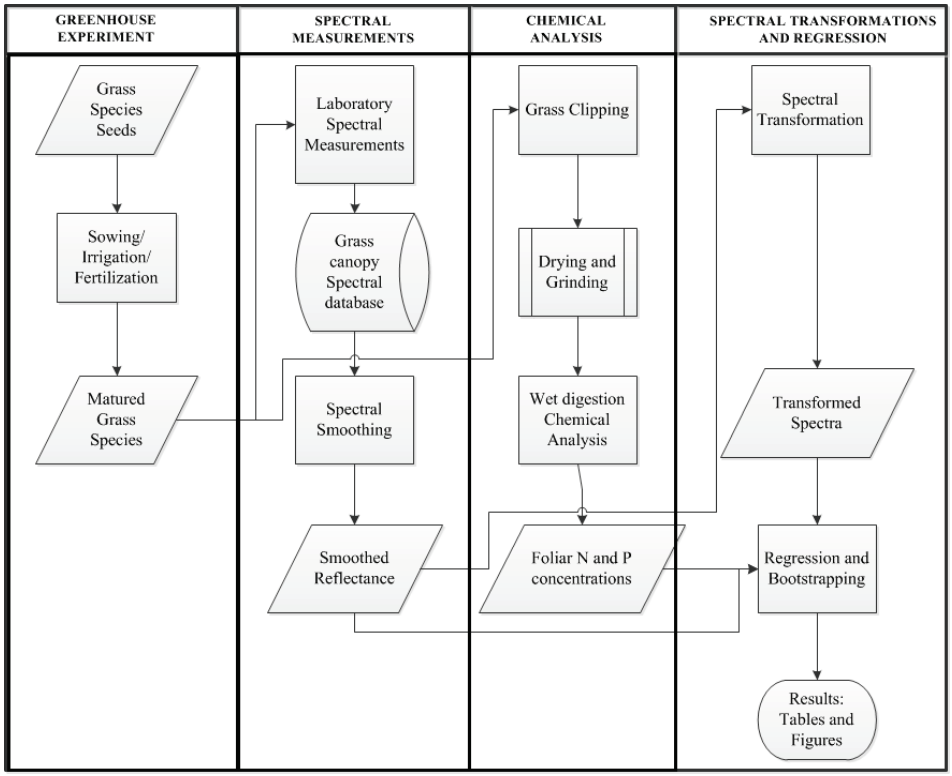

Figure 2.1: Flowchart showing sections of the methodology

2.4 Spectral Measurements

Canopy spectral measurements were taken for each pot using a FieldSpec ® 3 Portable Analytical Spectral Device (ASD®) spectrometer with a spectral range extending from 350 to 2500 nm, and a 1 nm bandwidth (www.asdi.com). Measurements were taken in a dark room to minimize wall reflections. A halogen lamp mounted on a tripod with a fixed illumination angle of 45 degrees was used as illumination source. Given the pot diameter of 15 cm, the canopy reflectance was measured by pointing the fibre optic with a field of view of 25 degrees in a nadir position, from about 33 cm above the grass canopy, to ensure that only spectral measurements of the grass canopy were taken. A white reference panel (spectralon) was used before each spectral measurement to convert spectral radiance into reflectance. Measuring followed the protocol used by e.g. Cho and Skidmore. (2006) and Mutanga et al. (2003). A single spectral measurement included an average of 10 scans. Each pot was rotated whilst 5 spectral measurements were taken and averaged to account for illumination differences and bi-directional reflectance effects (Wang et al., 2009).

2.5 Chemical Analysis

For each pot all grasses were cut at the base and oven dried at 800 C for 24 hours. The dried grasses were sent to the Agricultural Research Council's Institute for Tropical and Subtropical Crops (ARC-ITSC) for chemical analysis. N and P were analyzed using the wet or acid digestion method, using perchloric and nitric acid for P and sulphuric acid for N (Giron, 1973; Grasshoff et al., 1983). Chemically analysed N and P were henceforth referred to as observed N and P.

2.6 Data Analysis

2.6.1 Spectral transformation techniques

The spectral reflectance data were pre-processed before transformation. Spectral smoothing was performed with the commonly used Savisky-Golay filter (Savitzky and Golay, 1964), adding a second order polynomial least square function and 3-band window to remove signal noise. To compare the water removed (WR) spectra with the other spectral transformation techniques, taking some of the absorption features in the visible spectrum into account; the spectral region from 500 to 2450 nm was selected for data analysis.

Commonly used spectral transformation techniques such as log transformed spectra Log(1/R), first derivative, and continuum removal, were computed. Log(1/R) was determined by calculating a log function of the spectral reflectance's reciprocal (Hruschka, 1987; Yoder and Pettigrew-Crosby, 1995;

Fourty and Baret, 1998). The first derivative of the spectral reflectance was derived using a first-difference approach. A first-difference transformation of the reflectance spectrum calculates differences in reflectance between adjacent wavebands. More details on this can be found in Dawson and Curran (1998). The continuum removed spectra were derived by applying a convex hull or a continuum line to the reflectance spectra connecting local spectral maxima (Kokaly and Clark, 1999; Kokaly, 2001; Mutanga et al., 2004c).

To reduce water absorption effects on weak biochemical absorption, the water removed spectra (WR) were derived from a non-linear least-squares spectral matching technique calculating a fresh leaf spectrum as a non-linear combination of a leaf water spectrum and a dry matter spectrum (Gao and Goetz, 1994, 1995), modified by Schlerf et al. (2010), using the following equation:

$$R_{mod}(\lambda)=(A+B\,\lambda)exp\text{-}(C_wK_w(\lambda) + C_{dm}K_{dm}(\lambda)) \quad (2.1)$$

where Rmod(λ) is the modelled reflectance for wavelength λ, Cw the water content, Kw the absorption coefficient of water, Cdm the dry matter content, K_{dm} the absorption coefficient of dry matter content, and *A* and *B* background model coefficients. *A*, *B* and C_w were unknowns determined using the mathematical optimization procedure called Nelder-Mead simplex method (Mathews and Fink, 2004; Mathworks, 2009). It creates a generalized triangle in *N* dimensions for finding a local minimum of a function of several variables using a non-derivate method. The method tries to minimize a scalar-valued nonlinear function of *N* real variables using only function values, without any derivative information. Mathews and Fink (2004) argued that the Nelder-Mead method is efficient and computationally compact. More details can be found in Mathews and Fink (2004) and Mathworks (2009). The absorption coefficient of water and the absorption coefficient of protein or dry matter were obtained from the PROSPECT leaf model (Jacquemoud et al., 1996). The leaf water contribution to the total fresh leaf reflectance was modelled by filling in the three unknowns *A*, *B*, and C_w in equation (1), whilst setting C_{dm} at zero. Finally, the residual spectra between measured reflectance $R_{mes}(\lambda)$ and modelled reflectance $R_{mod}(\lambda)$ were computed as the water removed spectra $WR(\lambda)$:

$$WR\ (\lambda) = (R_{mes}(\lambda)- R_{mod}(\lambda))/ R_{mes}(\lambda) \quad (2.2)$$

The modifications by Schlerf et al. (2010) included the use of known and published water and protein absorption coefficients and incorporating Nelder-Mead simplex methods to determine the unknowns as indicated above. This study adopted the same technique for the savanna grass species. $WR\ (\lambda)$ was

used for analysis and compared with the other transformations of the spectra.

2.6.2 Regression analysis and bootstrapping

Two commonly used regression techniques were selected for data analysis, i.e. partial least square regression (PLSR) (Geladi and Kowalski, 1986; Naes et al., 1986; Ehsani et al., 1999; Martens and Naes, 2001; Viscarra Rossel, 2008) and stepwise multiple linear regression (SMLR) (Grossman et al., 1996; Martin and Aber, 1997; Kokaly and Clark, 1999; Huang et al., 2004; Schlerf et al., 2010). To compare the retrieval accuracy of foliar N and P using the various spectral transformation techniques, a bootstrapping approach was used (Efron, 1983). The advantage of bootstrapping is that it can be used efficiently when only a limited number of samples are available. Bootstrapping was used as an alternative to the split method since it iteratively resample the data set to be used for model development, making it a good technique for assessing model accuracy (Verbyla and Litvaitis, 1989). In this study, PLSR and SMLR were integrated with bootstrapping to derive calibrated and validated models. To integrate PLSR and bootstrapping, bagging-PLSR was implemented using the Parles 3.1 software (Viscarra Rossel, 2007, 2008). SMLR was integrated with bootstrapping using Mathworks (2009).

Using bagging-PLSR, independent or predictor variables were mean-centred to normalize them prior to further statistical analysis. The leave-one-out cross validation, as defined by the lowest root mean square error (RMSE), was used to determine the optimal number of factors or latent variables to be used for model development (Cho et al., 2007a; Darvishzadeh et al., 2008b; Viscarra Rossel, 2008). This Optimal number of factors was then used for model development and validation with the number of bootstraps equalling 1000. Bootstrapping with SLMR also used 1000 iterations and was implemented in Mathworks (2009). Only significant wavelengths were used in the model development using SMLR. Wavelengths were selected using the conventional rule for selecting independent variables in SMLR ("in" if $p<0.05$, and "out" if $p>0.01$). To associate the wavelengths selected by SMLR for N and P with the known absorption features of protein and starch by Curran (1989) and Kumar et al. (2001) respectively, the maximum difference of 30nm was chosen for the consistent comparison.

For both models (PLSR and SMLR) the retrieval accuracy was defined by the bootstrapped mean of the coefficient of determination (R2) and the RMSE. The confidence interval at a 95% confidence level was calculated for both R2 and RMSE.

2.7 Results

2.7.1 Performance of WR spectra for foliar N estimation using PLSR and SMLR

Generally, the WR technique used in combination with PLSR yielded the highest N retrieval accuracy (R2=0.84; RMSE=0.28, 17% of the mean), compared with other spectral transformation techniques (Table 2.1, Figure 2.2). The 95% confidence interval (CI) of the N retrieval accuracy (RMSE, 95% LCI=0.25 and UCI=0.33) confirmed the outperformance of the WR technique plus PLSR over other techniques. A similar trend with a slightly higher accuracy was obtained when the WR technique was combined with SMLR, producing the highest retrieval accuracy overall (R2=0.87; RMSE=0.25, 15% of the mean; Table 2.1, Figure 2.2). The second most important spectral transformation for N estimation after the WR technique was continuum removal, producing a RMSE of 0.30 (18% of the mean), and a 95% confidence interval varying from 0.24 to 0.31, when using PLSR. Similar trends were obtained when the continuum removal was combined with SMLR (R2=0.78; RMSE=0.34, 20% of the mean). The poorest performances were obtained with the first derivative, the Log(1/R), and the original reflectance spectra, with an R2 of around 0.6 and a RMSE varying between 0.43 and 0.48 (about 26% of the mean; Table 2.1, Figure 2.2).

Foliar N retrieval accuracies were not significantly (t-value=-0.18, df=8, p=0.859) different between PLSR and SMLR approaches. On average the PLSR accuracies were slightly higher than those of the SMLR, with a higher R2 and a lower RMSE. Generally, the results consistently demonstrated the high performance of the WR technique.

More bands were selected using SMLR and WR spectra (about 11 bands) to estimate N concentration, than for other spectral transformations (Table 2.2). Many of the important bands for all spectral techniques were located in the known absorption features for protein and N (Curran 1989; Kumar et al. 2001) (see Table 2.2 and Figure 2.3). For N estimation, Figure 3 depicts the PLSR weights plotted with the grass canopy spectra of the highest performing spectral transformations, namely WR and continuum removal. The original reflectance curve was added to Figure 2.3 for reference.

Table 2.1: The performance of spectral transformations for estimating N and P using PLSR and SMLR combined with bootstrapping.

Spectra		R^2	LCI 95%	UCI 95%	RMSE	LCI 95%	UCI 95%	*no. of factors or bands
CR	N *vs.* PLSR	0.81	0.68	0.94	0.30	0.26	0.35	8
	N *vs.* SMLR	0.78	0.69	0.78	0.34	0.33	0.38	4
	P *vs.* PLSR	0.40	0.38	0.41	0.08	0.06	0.08	6
	P *vs.* SMLR	0.37	0.18	0.38	0.07	0.07	0.09	2
FD	N *vs.* PLSR	0.59	0.48	0.70	0.45	0.39	0.53	3
	N *vs.* SMLR	0.59	0.48	0.70	0.45	0.39	0.53	9
	P *vs.* PLSR	0.47	0.46	0.48	0.07	0.06	0.08	7
	P *vs.* SMLR	0.25	0.12	0.27	0.08	0.08	0.09	4
Log(1/R)	N *vs.* PLSR	0.62	0.50	0.74	0.43	0.38	0.50	6
	N *vs.* SMLR	0.60	0.54	0.62	0.45	0.43	0.50	4
	P *vs.* PLSR	0.17	0.16	0.18	0.08	0.07	0.10	4
	P *vs.* SMLR	0.15	0.01	0.10	0.09	0.92	0.10	1
R	N *vs.* PLSR	0.60	0.49	0.71	0.44	0.38	0.52	4
	N *vs.* SMLR	0.55	0.48	0.56	0.48	0.46	0.57	4
	P *vs.* PLSR	0.18	0.17	0.19	0.08	0.07	0.10	4
	P *vs.* SMLR	0.10	0.01	0.11	0.06	0.09	0.10	3
WR	N *vs.* PLSR	0.84	0.71	0.97	0.28	0.25	0.33	6
	N *vs.* SMLR	0.87	0.82	0.88	0.25	0.24	0.31	11
	P *vs.* PLSR	0.43	0.42	0.44	0.07	0.06	0.08	5
	P *vs.* SMLR	0.64	0.48	0.64	0.06	0.05	0.07	9

N=Nitrogen, P=Phosphorus, CR=continuum removal, FD=first derivative, R=reflectance, WR=water removed, LCI=lower bound confidence interval, CI=upper bound confidence interval. *factors for partial least square regression (PLSR) and bands for stepwise multiple linear regression (SMLR).

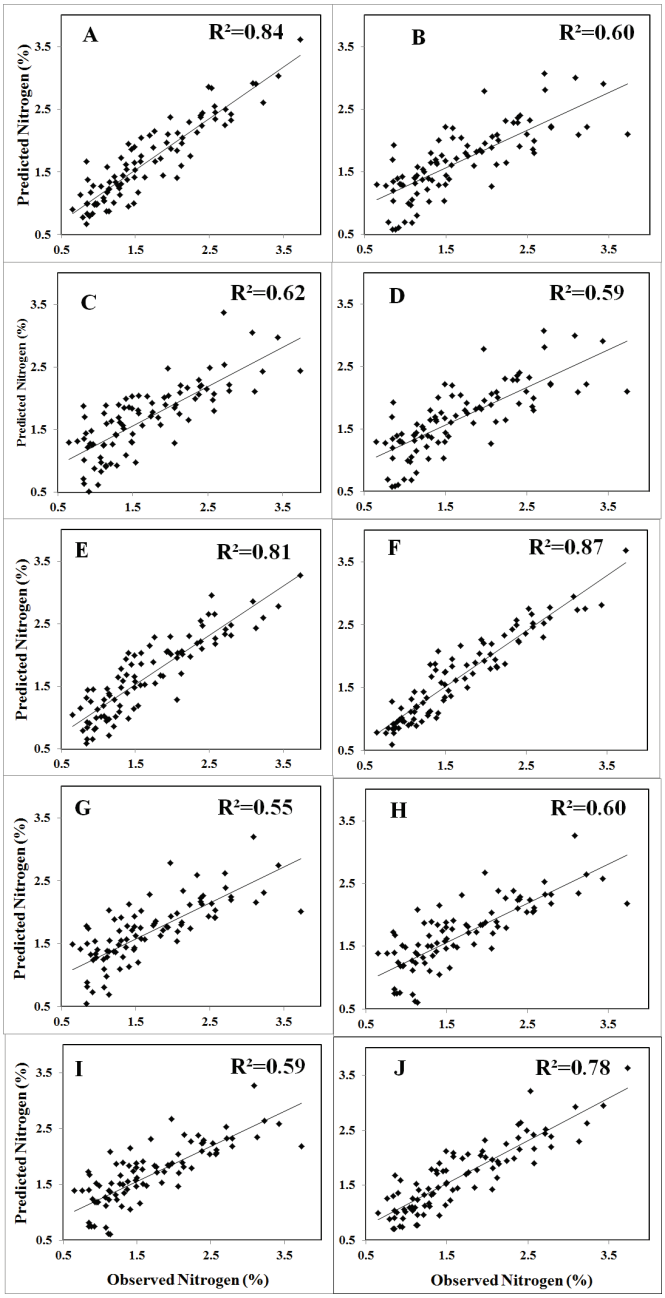

Figure 2.2**:** Predicted *vs*. observed nitrogen (N) concentrations derived from various spectral transformations and partial least square regression (PLSR, A to E) as well as stepwise multiple linear regression (SMLR, F to J); A=water removed spectra, B=original reflectance, C=Log transformed spectra, D=first derivative spectra, E=continuum removed spectra, F=water removed, G=original reflectance, H=Log transformed spectra, I=first derivative and J=continuum removed spectra.

Table 2.2: Wavelengths selected using SMLR, marking wavelengths corresponding to known absorption features.

Spectra *vs.* N/P	Selected wavelengths (nm) for N and P estimation using SMLR										
CR *vs.* N	732	1057	1836	**2129***							
CR *vs.* P	744	785									
FD *vs.* N	549	**673***	734	**1666***	1807	1908	2260	2290	**2374***		
FD *vs.* P	839	**1204***	**1667***	2450							
Log(1/R) *vs.* N	504	605	698	1396							
Log(1/R) *vs.* P	**638***	691	878								
R *vs.* N	511	604	696	1394							
R *vs.* P	640										
WR *vs.* N	522	**675***	1087	1159	1299	1360	2014	2038	**2056***	**2141***	2342
WR *vs.* P	935	1036	**1209***	**1974***	**2061***	2296	**2320***	2364	2379		

CR=continuum removal, FD=first derivative, R=original reflectance, WR=water removed spectra. *Known absorption features protein / N for nitrogen (N) and starch for phosphorus (P) (Curran 1989; Kumar et al., 2001).

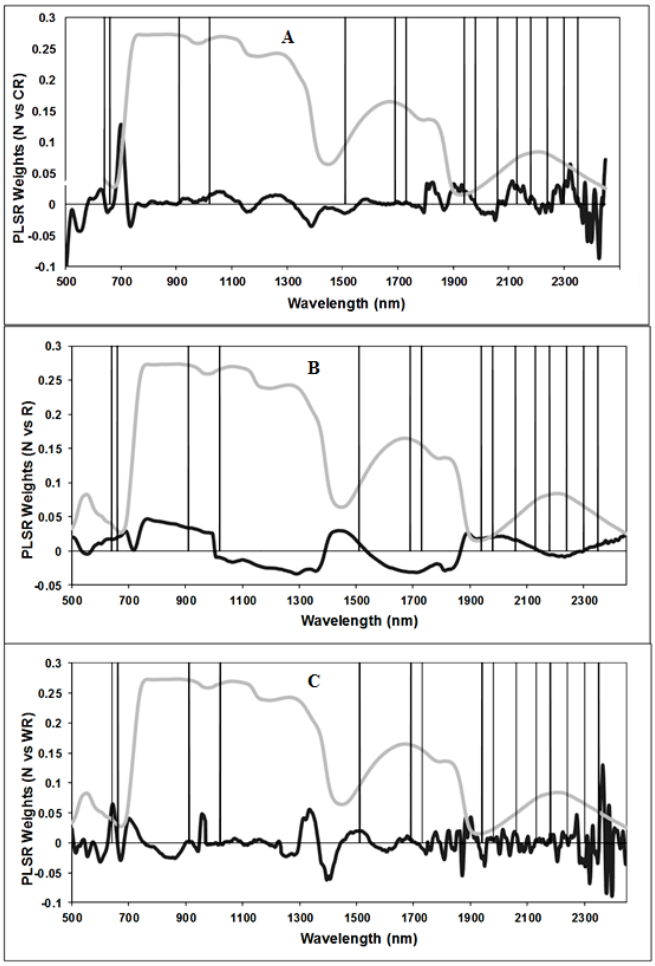

Figure 2.3: Partial least square regression (PLSR) weights showing the contribution of each wavelength in the development of models for nitrogen (N) estimation using a) continuum removal (CR) spectra (N vs. CR), b) reflectance (R) spectra (N vs. R), c) WR (water removed) spectra (N vs. WR), plotted with the grass canopy reflectance for reference. The more positive or negative the weight is on a particular wavelength, the more contribution it has towards the model development. Vertical lines indicate known N absorption features as listed in Curran (1989) and Kumar et al. (2001).

2.7.2 Performance of WR spectra for foliar P estimation using PLSR and SMLR

Generally, the WR technique yielded the highest P retrieval accuracy compared to other spectral transformation techniques. Contrary to what was observed with WR-PLSR, the WR-SMLR technique yielded a higher coefficient of determination (R^2=0.64) with a slightly higher P retrieval accuracy (RMSE=0.06, 18% of the mean; Table 2.1, Figure 2.4). This suggests that overall the WR technique minimized water masking effects on features sensitive to P. The WR technique had a higher accuracy compared to the

continuum removal technique (RMSE=0.07, 20% of the mean), which again was the second performer with an about 27% lower R^2 value (Table 2.1, Figure 2.4). Using PLSR, the first-derivative transformation produced the highest estimates of P with a RMSE of 0.068 (20% of the mean), slightly higher than the WR technique with its RMSE of 0.070 (20.6% of the mean). Continuum removal and PLSR yielded a RMSE of 0.08 (22% of the mean). The lowest performing spectral transformation techniques for estimating P with PLSR were Log(1/R) and original reflectance, with a RMSE of 0.08 (24% of the mean) and of 0.07 (20% of the mean), respectively. Again P retrieval accuracies using SMLR were not significantly different (*t-value*=-1.1095, *df*=8, *p*=0.299) from those using PLSR. Log(1/R) and original reflectance generally obtained poor results with both regression techniques.

Using the WR technique, the stepwise regression technique yielded the highest number of bands for estimating P concentrations compared to the other transformation approaches. Some of the selected bands corresponded with known absorption features of starch, as cited in Curran (1989) and Kumar et al. (2001) (Table 2.2), as was also observed in the PLSR weights noted in Figure 2.5. For P estimation, Figure 2.5 depicts the PLSR weights plotted with the grass canopy spectra of the highest performing spectral transformations, namely WR and the first derivative spectra. The original reflectance curve has been added to Figure 2.5 for reference.

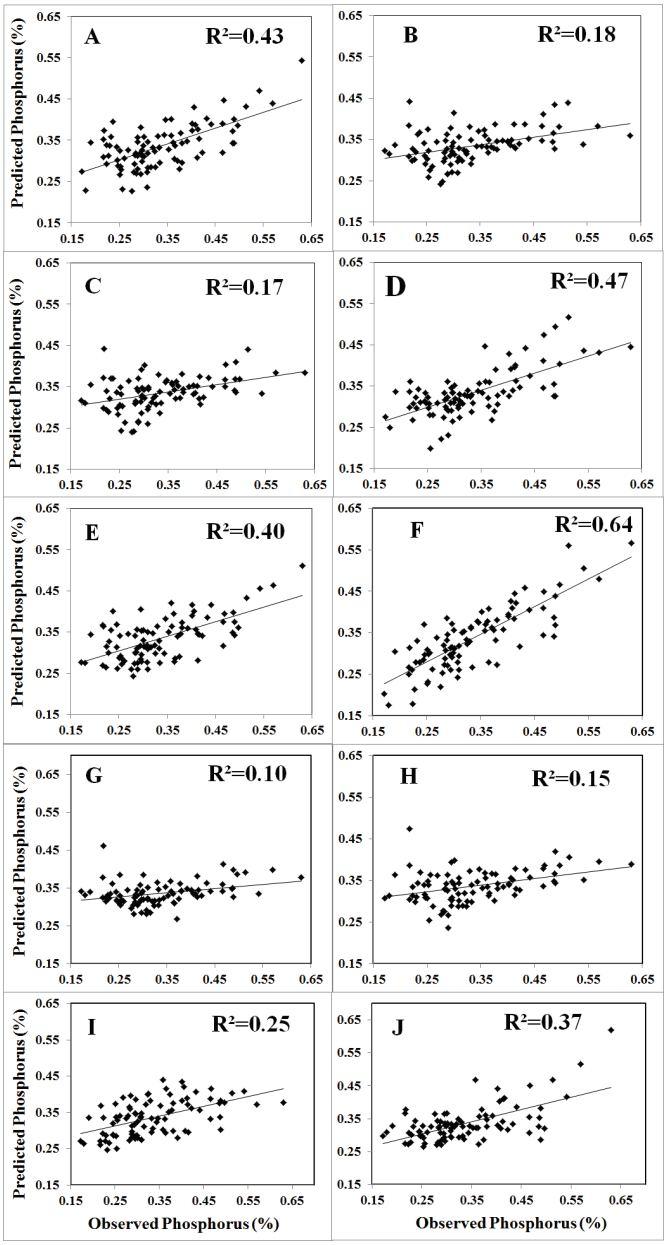

Figure 2.4: Predicted vs. observed phosphorus (P) concentrations derived from various spectral transformations and partial least square regression (PLSR, A to E) as well as stepwise multiple linear regression (SMLR, F to J); A=water removed spectra, B=original reflectance, C=Log transformed spectra, D=first derivative spectra, E=continuum removed spectra, F=water removed, G=original reflectance, H=Log transformed spectra, I=first derivative-and J=continuum-removed spectra.

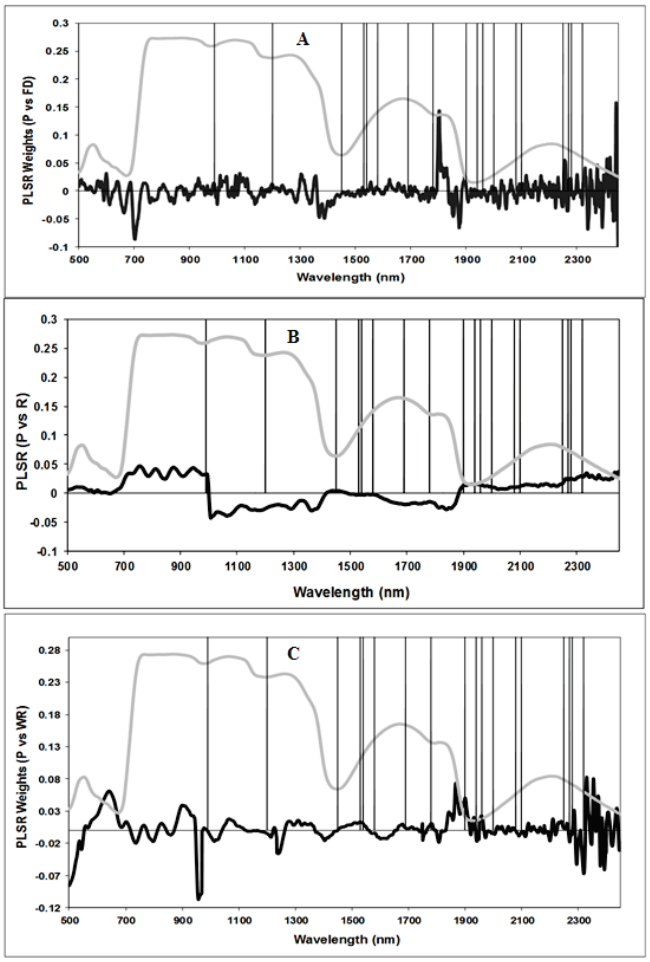

Figure 2.5: Partial least square regression (PLSR) weights showing the contribution of each wavelength in the development of models for estimating P using a) first derivative (FD) spectra (phosphorus (P) vs. FD), b) reflectance (R) spectra (P vs. R), c) water removed (WR) spectral data (P vs. WR), plotted with the grass canopy reflectance for reference. The more positive or negative the weight is on a particular wavelength, the more contribution it has towards the model development. Vertical lines indicate known starch absorption features as listed in Curran (1989) and Kumar et al. (2001).

2.7.3 Foliar N, P concentrations and spectral reflectance on grass canopies

A high variability of foliar N and P was observed in the samples of *D. eriantha* cultivated in the greenhouse and treated with various nutrient and water levels. Descriptive statistics for foliar N and P are detailed in Table 2.3; the mean foliar N and P concentrations were 1.68% and 0.34% of dry matter respectively. The distribution of both N and P foliar concentrations across all treatments were normal, as tested by the Kolmogorov-Smirnov normality test (p>0.05, 4.39 and 2.16 chi square test, respectively). Figure 2.6 shows

the descriptive statistics (minimum, maximum and mean) of the original reflectance before spectral transformation.

Table 2.3: Descriptive statistics of foliar N and P for *D. eriantha*

Measured Variables	No. of Obs.	Min	Max	Mean	StDev
Nitrogen (N%)	90	0.65	3.73	1.68	0.25
Phosphorus (P%)	90	0.17	0.63	0.34	0.07

Obs.=Observation, StDev=Standard Deviation

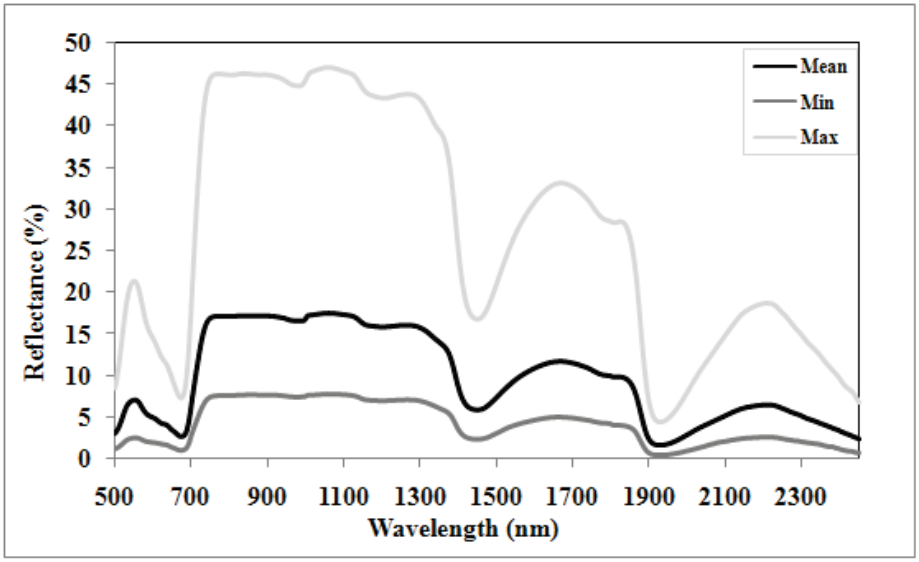

Figure 2.6: The descriptive statistics (minimum, maximum and mean) of the original spectral reflectance for the grass canopies before spectral transformations.

2.8 Discussion

2.8.1 WR spectra for estimating foliar N

This study demonstrated the potential of the WR approach as one of the spectral transformation techniques that could be used to increase the accuracy of foliar N estimation. By reducing the water effect across the fresh leaf spectra this technique enhances weak or subtle absorption features. The regions of the electromagnetic spectrum most affected by water are the NIR and SWIR, important regions for distinguishing various foliar biochemicals. This was illustrated by the selection of more bands in the SWIR region being related to foliar N concentration when using WR instead of reflectance (Table 2.2, Figure 2.3). Some of these bands correspond to known absorption features cited by Curran (1989) and Kumar et al. (2001) due to several absorption mechanisms including electron transition in the visible region, C-H stretch of the 2nd overtone mainly in the 1100-1300 nm region, and C=O, O-

H, N-H, C-O, C-H as well as C-C for the region within 1300-2380 nm (Kumar et al., 2001). Generally this study showed the applicability of the WR technique for savanna grass species, with a higher R^2 of 0.87 for N estimation compared to the results (R^2=0.52) attained by Schlerf et al. (2010). Gao and Goetz (1994; 1995) successfully implemented the WR technique for lignin and cellulose estimations, highlighting the importance of minimizing water effects on the SWIR. The WR technique is easy to implement but requires a careful and proper parameterization of the least square spectral mixture analysis model to provide reliable results (Gao and Goetz, 1994).

The continuum removal technique yielded the second highest accuracy for estimating foliar N concentrations with both PLSR and SMLR. Continuum removal enhanced the differences in absorption strength (Clark and Roush, 1984; Schmidt and Skidmore, 2001; Schmidt et al., 2004). The highest R^2 obtained in this study using continuum removal was 0.81. This is consistent with the N retrieval accuracy of forest sites reported by Kokaly and Clark (1999), where an R^2 of 0.75 to 0.94 was attained using continuum removal and continuum removal-derived indices. This study attained higher accuracy results for N based on continuum removal than a study by Mutanga et al. (2005) on *Cenchrus ciliaris* grown in the greenhouse. Estimating foliar N with Log(1/R) yielded higher retrieval accuracy than with reflectance, but not as high as with the WR and continuum removal techniques. Yoder and Pettigrew-Crosby (1995) showed Log (1/R) performed accurately estimating N concentrations, compared to reflectance. Similar results were also attained by Fourty and Baret (2001). They argued that by transforming reflectance to absorbance Log(1/R) values the accuracy of biochemical estimates was improved. Log(1/R) is likely to be used instead of the original reflectance because of the linear relation between the absorbing components and its contribution to the Log(1/R) value at the wavelength absorbed (Hruschka, 1987). However, the present study shows the performance of WR to be higher than that of Log(1/R).

2.8.2 WR spectra for estimating foliar P

The performance of the first derivative and WR techniques in terms of P retrieval accuracy highlights the importance of reducing the influence of water on the fresh leaf spectra. As shown in Table 2.2 and Figure 2.2, many bands highlighted by the models as sensitive to P correspond to known absorption features of starch. The relationship between P and starch is understood to be based on an energy molecule rich in P called *adenosine triphosphate* (ATP), which is used in starch formation (Heldt et al., 1977; Sava Stankovic, 1978; Larcher, 1980; Okita, 1992). The energy from the sun is converted into chemical energy stored as a form of ATP and then used to bond with carbon dioxide (CO_2) and hydrogen to form starch (Larcher, 1980).

To date, few studies have focused on the estimation of P using remote sensing. Bogrekci and Lee (2005) used wavelengths from 225 to 2550 nm, while Mutanga and Kumar (2007) showed that bands in the SWIR were more sensitive to P, which is similar to the findings in this study. The high precision of N and P estimation using the WR technique is evident from the lower confidence limits of the correlation coefficient and the RMSE derived from the bootstrapping technique (Table 2.1).

2.8.3 Comparing PLSR and SMLR for foliar N and P estimation

PLSR consistently performs well in estimating N. Comparative studies using both PLSR and SMLR indicated the predictive power of PLSR. For example, Bogrekci and Lee (2005) showed that PLSR had a higher P accuracy than SMLR on grass leaves. This study showed different results, with the SMLR-WR technique yielding a higher accuracy for P than the PLSR-WR did. SMLR is normally confounded with difficulties transferring the predictive models to the other data sets or other areas (Grossman et al., 1996). On average, including all other spectral transformation techniques, PSLR still showed its high predictive power. Hansen and Schjoerring (2003) concluded that PLSR can be used as an alternative univariate statistical technique. This is mainly because PLSR minimizes the multicollinearity effects by decomposing the spectral data into non-collinear latent variables. Over-fitting may be minimized using PLSR by selecting an optimal number of latent variables rather than having more redundant explanatory variables (Viscarra Rossel, 2007, 2008).

2.9 Conclusion

The estimation of N, P, and N: P at landscape level forms an important objective to facilitate investigation into the feeding patterns and distribution of wildlife and livestock in African savannas. The study tested the applicability and performance of the water removal technique in estimating P and N concentrations on grass canopies and compared this technique with other spectral transformation techniques such as first derivative, Log(1/R) and continuum removal, as well as the original reflectance. This study suggests that the water removed approach is a useful technique to retrieve foliar N and P concentrations from grass in savannas especially N which was consistently estimated with high accuracy by both SMLR and PLSR. This study focused on a single species; future studies should consider multiple species at field, airborne or satellite level to test the utility of the water removed technique at landscape level in savanna ecosystems.

2.10 Acknowledgements

The study was funded by the University of Twente, Faculty of Geoinformation and Earth Observation (ITC), the University of Wageningen, the Council for Scientific and Industrial Research (CSIR), and the South African Department

of Science and Technology. Authors also thank CSIR's Ecosystem-Earth Observation colleagues (especially Dr. Moses Cho for his technical support and advice on greenhouse setup and data analysis), ITC's Natural Resource Management and University of Wageningen's Resource Ecology colleagues (especially Prof.Dr. Herbert H.T Prins) for his scientific advices. Finally, we would also like to thank Mr. Bongani Majeke (Statistics South Africa), Mr Moses Mokoena and Mr Samuel Taola Shai (CSIR's Bioscience unit) for helping to set up greenhouse experiments and laboratory spectral measurements.

Chapter 3

Grass nitrogen and phosphorus ratio estimation using *in situ* hyperspectral measurements*

* Chapter is based on: Ramoelo A. Skidmore A.K. Schlerf M. Heitkönig I.M.A, Mathieu R. Cho M.A. Savanna grass nitrogen to phosphorus ratio estimation using *in situ* hyperspectral remote sensing data, *International Journal of Remote Sensing* (**under review**)

Abstract

Foliar nitrogen (N) and phosphorus (P) are important minerals for herbivores, and can be used as indicators for grass quality. Determining the foliar N: P ratio provides a tool for understanding nutrient limitation on plant production and consequently for herbivores. Several experimental studies in the temperate vegetation came up with critical values to determine levels of nutrient limitation, but few of this type of studies were undertaken in the savanna ecosystems. The N: P critical values are used to determine whether N or P is limiting vegetation production or biomass. According to a study on savanna ecosystems, N: P value below 9 indicates N limitation, between 9 and 10 suggests co-limitation, and a value greater than 10 indicates P limitation. In order to understand the nutrient limitation at landscape scale, remote sensing techniques offer that opportunity. The objective of this study is to investigate the utility of *in situ* hyperspectral measurements to estimate foliar N: P. Field data collection for spectral measurements using a spectroradiometer (FieldSpec 3®) was undertaken, and grass samples were collected for foliar N and P extraction. The foliar N: P prediction models were built using various spectral data and partial least square regression (PLSR). The spectral data used include the original reflectance (R), continuum removal (CR), water removal (WR), first difference derivative (FD) and Log transformation (Log(1/R)). The results showed that foliar N: P can be estimated using spectral data and PLSR. CR and WR spectra predicted foliar N: P ratio with higher accuracy as compared to FD and R spectra. The performance of CR and WR spectra were attributed to their ability to minimize sensor and water effects on the fresh leaf spectra, respectively, while enhancing the absorption features for foliar biochemicals. The study demonstrated the possibility to predict foliar N: P ratio using *in situ* hyperspectral data, and shortwave infrared (SWIR) found to be highly sensitive to foliar N: P. The study recommends the use of airborne hyperspectral data to extend the estimation of grass N: P at landscape level. At landscape level, the foliar N: P ratio information could be used by the resource managers, park managers, farmers and ecologists to understand the feeding patterns, resource selection and distribution of wild herbivores and livestock.

3.1 Introduction

Estimation of foliar biochemicals provides information that enables the assessment of ecosystem functioning, for example, nutrient cycling, gas exchange and plant productivity (Martin and Aber, 1997; Ollinger et al., 2002). Foliar biochemicals such as nitrogen (N) and phosphorus (P) are primary indicators of physiological processes such as photosynthesis, leaf respiration and growth rates (Field and Mooney, 1986; Evans, 1989; Güsewell, 2004). Foliar N and P concentrations can also be used as an indicator for grass quality (McNaughton, 1988, 1990). The foliar N concentration, for example, is known

to relate to the protein (Clifton et al., 1994) which is one of the main nutrient for the herbivores (Prins and Beekman, 1989; Prins and van Langevelde, 2008). For example, Prins and Beekman (1987) report that Buffalo require about 6.2% dietary protein (i.e. 1 % of N), while Duncan, (1992) found out that lactating equids require about 0.24% of P in their food for maintenance. The importance of N supply on dry matter production as well as protein content is well documented in agricultural literature (Marschner, 1995) while foliar P is one of the main nutrient requirements for lactating mammals (McNaughton, 1990). Therefore, grass quality information can be used to understand feeding patterns and distribution of wildlife and livestock (Owen-Smith and Cooper, 1987; Prins, 1987; McNaughton, 1988; Prins and Beekman, 1989; McNaughton, 1990; Duncan, 1992; McNaughton and Banyikwa, 1995).

To understand the nutrient limitation in vegetation, N: P ratio is a key indicator (Koerselman and Meuleman, 1996; Ludwig et al., 2001; Güsewell, 2004; Cech et al., 2008; Craine et al., 2008). The foliar N:P ratio reflects the balance of N and P supply which influence plants at all levels, i.e. the growth and reproduction of individual plants, plant species interactions, composition and diversity (Güsewell, 2004; Cech et al., 2008). Therefore, the differences in foliar biochemistry as captured by N:P ratio in plants provide information on which nutrient is limiting and could determine the plant production or biomass which could eventually affect the feeding activity of herbivores in a particular ecosystem (Koski, 1999; Daufresne and Loreau, 2001; Güsewell, 2004; Prins and van Langevelde, 2008). "Smaller herbivores are known to be more limited by N than larger ones and not vice versa" (Prins and van Langevelde, 2008), since they require more nutrients and energy for growth and to achieve their daily activities on per kilogram basis. This tool has been widely used in the limnetic related studies (Koerselman and Meuleman, 1996; Güsewell et al., 2003; Güsewell, 2004). Limnetic studies hypothesized that herbivores (cladocerans) with high P concentration should be generally P limited, while herbivores with more N (e.g. copepods) should be N limited (Koski, 1999; Daufresne and Loreau, 2001). This provides a scope for testing such a hypothesis in the savanna ecosystems. A review of 40 fertilization studies revealed that the N: P ratio value <14 indicates N limitation, between 14 and 16 indicates co-limitation and finally the >16 indicates P limitation (Koerselman and Meuleman, 1996; Ludwig et al., 2001; Güsewell, 2004). Koerselman and Meuleman (1996) critical values cannot be applied in the savanna ecosystems, because temperate vegetation is mainly C3-type, while the perennial savanna grass species are mainly C4-type (Ludwig et al., 2001; Cech et al., 2008; Craine et al., 2008). Güsewell (2004) revealed that at a vegetation level, the critical N: P ratio is <10 (N-limiting) and >20 (P-limiting) based on short term fertilization, and argued that the values might be different at individual species. Ludwig et al. (2001) found that below trees, N: P ratio value of 12 indicates that P is limiting, while on the open grassland the average of 6 indicates N

limiting in the savanna ecosystems. Cech et al. (2008) concluded that the critical values of N: P for determining the limitation of N and P range between 9 and 10, where N limitation is less than 9, co-limitation is between 9 and 10, and above 10 is P limitation, in the savanna ecosystems. There is no general consensus on which critical range of N: P values could be used, since there are limited experimental studies focusing on this, especially in the savanna ecosystems.

Using remote sensing, there is a significant progress in estimating foliar biochemicals, especially using *in situ* hyperspectral or spectroscopy approaches. A simple technique is to correlate vegetation index and a biochemical concentration of interest, e.g. N (Hansen and Schjoerring, 2003; Ferwerda et al., 2005; Jain et al., 2007; Rivero et al., 2009; Abdel-Rahman et al., 2010). Challenges for using this approach include soil exposure, atmospheric influence and saturation especially during the peak productivity (Tucker, 1977; Jackson and Huete, 1991; Mutanga and Skidmore, 2004b). There were several attempts to solve these problems using the red edge region of the spectra e.g. Mutanga and Skidmore (2004). The second approach is to use the specific regions of the spectra which are known to relate to the physical bond vibrations of the specific foliar biochemical concentrations referred to as absorption features (Darvishzadeh et al., 2008b; Knox et al., 2011). For example, bands centred at 430 nm, 640 nm, 910 nm, 1020 nm, 1420 nm, 1690 nm, 1940 nm, 2060, 2240 nm, 2300 nm, are the absorption features related to physical bond vibrations of foliar N and protein (Curran, 1989; Kumar et al., 2001; Knox et al., 2010a). The absorption features are well documented in Curran (1989) and Kumar et al. (2001). These features were successfully used for estimating N, P and chlorophyll (Darvishzadeh et al., 2008b; Knox et al., 2011). Thirdly, the use of whole spectrum became prominent following the application of statistical techniques such as stepwise regression (SMLR) and partial least square regression (PLSR). The use of the full spectrum was mainly coupled with several spectral transformation techniques such as continuum removal, Log transformed (Log(1/R)), derivatives as well as the water removal (Yoder and Pettigrew-Crosby, 1995; Dawson and Curran, 1998; Kokaly and Clark, 1999; Ramoelo et al., 2011b).

The spectral transformation techniques are mainly used to enhance absorption features of foliar biochemical concentrations, while minimizing atmospheric, soil background, and water absorption effects, as well as data redundancy (Yoder and Pettigrew-Crosby, 1995; Dawson and Curran, 1998; Cho and Skidmore, 2006). Ramoelo et al. (2011b) demonstrated that using WR and PLSR improves the estimation of foliar N and P in the controlled environment, due the capability of WR to minimize water absorption effect on the fresh leaf spectra. Continuum removal has also been successfully applied

to enhance absorption features for foliar biochemical concentrations (Kokaly and Clark, 1999; Curran et al., 2001; Mutanga et al., 2005). The Log(1/R) is preferred to reflectance because it is linearly related to absorbing components (Hruschka, 1987; Yoder and Pettigrew-Crosby, 1995). Studies such as Yoder and Pettigrew-Crosby (1995) showed a strong relationship between Log(1/R), as well as the first derivative of Log(1/R)', and foliar N concentration. Fourty and Baret (1998) argued that transforming reflectance into their corresponding absorbance values improved the accuracy of biochemical estimates. Continuum removal has also been successfully applied to enhance absorption features for foliar biochemical estimations (Kokaly and Clark, 1999; Curran et al., 2001; Mutanga et al., 2005).

Foliar N concentration has been estimated and mapped more often than P, especially using field and airborne hyperspectral data. This trend could be attributed to the following reasons;

(i) Low concentration of P in plants, normally 10 time lower than N (Stark, 1970; Meissner et al., 1999; Knox et al., 2010a).
(ii) P has few identifiable absorption features.

Nevertheless, foliar P in combination with N are crucial parameters for understanding the nutrient limitation (N: P) for both plant and herbivores. The computation of N: P relies on the accurate retrieval of both foliar N and P, but P is having the noted problems as above. Asner and Martin (2008) also argued that the retrieval of foliar P using hyperspectral remote sensing could be associated with the stoichiometry (i.e. indirect estimation based on the relationship with other foliar biochemicals) (Elser et al., 1996; Daufresne and Loreau, 2001). Predicting foliar N and P separately using *in situ* hyperspectral data and later compute foliar N: P could face error propagation, because the retrieval accuracy of N and P is not consistently similar. Therefore, to minimize these errors it is crucial to directly estimate foliar N: P using remote sensing. Studies focusing on estimating nutrient limitation especially through foliar N: P ratios in the savanna ecosystem are rare, especially at landscape level using remote sensing. The advantage of using remote sensing is the provision of the synoptic landscape view while covering larger areas which conventional field measurement cannot cover. In this study the *in situ* hyperspectral data will be tested, to demonstrate the possibility of using remote sensing to predict foliar N: P. The importance of using partial least square regression to estimate foliar biochemicals is elucidated in Ramoelo et al. (2011b) or Chapter 2. The objective of the study is to investigate the utility of *in situ* hyperspectral data in combination with partial least square regression in predicting the foliar N: P ratio in the grass layer of the savanna ecosystem.

3.2 Data collection

3.2.1 Study area

The study area is located (i.e. two corner co-ordinates; $24^0$40'0" S, $31^0$10'0" E and $25^0$0' S, $32^0$0"E) in the Lowveld savanna at the north-eastern part of South Africa (Figure, 3.1). The Lowveld landscape corresponds to the low lying area extending from the foot slopes of the Drakensberg Great Escarpment to the west and the Mozambique coastal plain to the east (Venter et al., 2003). The topography is gently undulating with flat patches in localized areas, and with an average height of 450m a.s.l. The study area covers a land use transect ranging from protected areas such as the private-owned Sabi Sands Game Reserve (SGR) and the state-owned Kruger National Park (KNP) to communal lands in the Bushbuckridge region. The western part of the transect (communal areas) receives higher mean annual rainfalls (800mm/yr.) as compared to the eastern side of the transect (580 mm/yr.) (Venter et al., 2003). The annual mean temperature is about 22^oC. The dominant geology includes granite and gneiss with local intrusions of gabbro (Venter et al., 2003). Consequently, these areas are characterized by gradients of soil moisture and nutrients. The soil fertility of gabbro areas are higher than the granitic ones (Ben-Shahar and Coe, 1992; Venter et al., 2003). The main vegetation communities include the "granitic lowveld" and the "gabbro grassy bushveld" (Mucina and Rutherford, 2006). In the gabbro patches, grass species such as *Setaria sphacelata* dominates the crest while species such as *Urochloa mosambicensis* dominates the valleys. Gabbro patches are dominated by grass species with high productive potential (e.g. *Urochloa mosambicensis*) compared to granite-derived soils (e.g. *Eragrostis rigidior and Pogonarthria squarrosa; cf.* Mutanga et al. (2004). The gabbro sites are dominated by fine leaves tree species such as *Acacia ssp* while the granite sites are dominated by broadleaves tree species such as *Combretum spp* and *Terminalia spp* (Venter et al., 2003; Ferwerda et al., 2006). Rangelands in the protected areas are grazed by wild herbivore such as impala (*Aepyceros melampus*), zebra (*Equus burchelli*), wildebeest (*Connochaetes taurinus*), buffalo (*Syncerus caffer*), while the communal rangelands support grazing of cattle (*Bos taurus*) and goats (*Capra hircus*) as well as sheep (*Ovis aries*), which determine various grazing intensities.

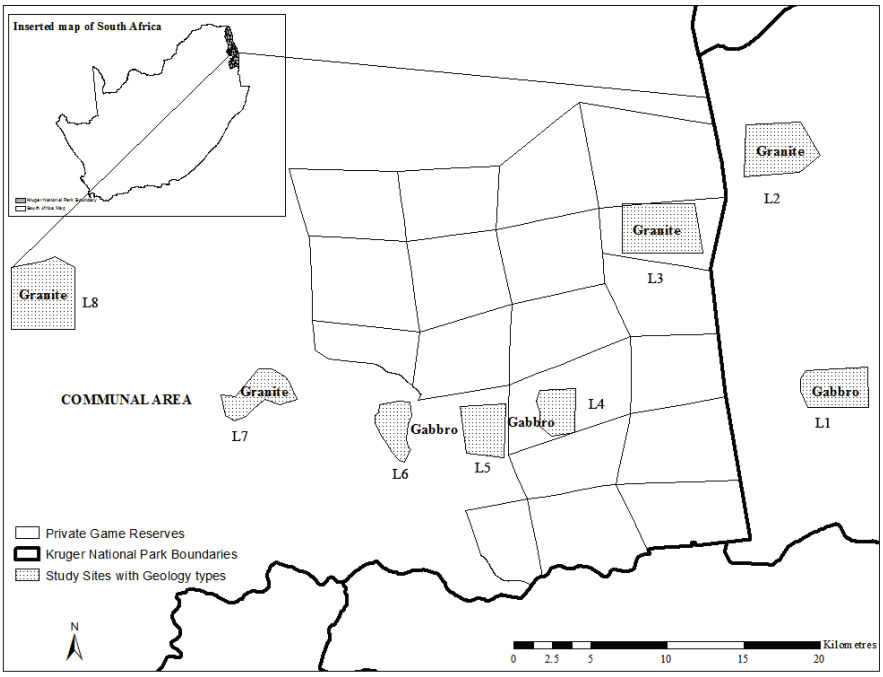

Figure 3.1: Study area map, L=Land use.

3.2.2 Sampling design

The study area comprised of eight sites which were placed along the land use gradient: two sites in KNP (L1 gabbro, L2 granite), two sites in SGR (L3 granite, L4 gabbro), and four sites in the communal areas (L5-6 gabbro, L7-8 granite) (Figure, 3.1). The sites (totalling ca. 35000ha) were demarcated using 1:250, 000 geology maps and refined using 2008 SPOT 5 images (Wessels et al., 2011). The site selection process was designed to capture the nutrient contrast from low to high in granitic-derived soils to gabbro-derived soils, respectively. A line transect sampling design was used to collect field data (Fewster et al., 2005) in each site except L3 (because of access limitations). To capture grass biomass variability, transects were laid out to sample both valley and crest land units. The topography influences the grass biomass in the savanna ecosystems with valley areas generally having higher grass biomass than crest areas. Along transects, a combination of purposive and systematic placement of sampling plot was undertaken. The distance between the plots was between 500m and 1000m depending on the accessibility and homogeneity of the area. The plot size was 30 m x 30 m. A total of 49 plots were surveyed and in each plot three to four subplots (0.5 m x 0.5m) were randomly selected to capture the plot variability. In each subplot, data on the sample location using the Leica®'s GS20 differential geographic positioning system (DGPS), dominant grass species and grass

samples were collected. Grass samples were dried at 80^0C for 24 hours and the measurements were later averaged at plot level. The DGPS points were post-processed using Leica's GeoPro software and reference GPS data from Nelspruit station to produce less than 1 m positional accuracy. The fieldwork was undertaken in March/April 2009 towards the end of the wet season, when grass biomass had achieved a maximum productivity to minimize the N/P and biomass interaction effects (Plummer, 1988b; Skidmore et al., 2010).

3.2.3 Chemical analysis

The dried grass samples were taken to the South African's Agricultural Research Council-Institute for Tropical and Subtropical Crops (ARC-ITSC)-Nelspruit for chemical analysis. Firstly, the acid digestion technique was used, where perchloric and nitric acids were used for foliar P concentration retrieval and sulphuric acid was used for retrieving foliar N concentrations (Giron, 1973; Grasshoff et al., 1983; Mutanga et al., 2004a). Secondly, the colorimetric method by auto analyser was used to measure foliar N (Technicon Industrial Method 329-74 W; Technicon Industrial Systems, Farrytown, New York). For foliar N measurements an emerald-green colour was formed by the reaction of ammonia, sodium salicylate, sodium nitroprusside and sodium hypochlorite. The ammonia-salicylate complex was read at 640 nm. For foliar P measurements, a colorimetric in which a blue colour was formed by the reaction of ortho-phosphate and the molybdate ion. The phosphomolybdenum complex was then read at 660 nm. These extraction methods were successfully used for grass foliar N and P by Mutanga et al. (2004) and Ramoelo et al. (2011b). The chemical analyzed foliar N: P ratio was hereafter referred to as observed N: P ratio.

3.2.4 Canopy Spectral measurements

The reflectance spectra were measured using an Analytical Spectral Device (ASD) spectroradiometer, Fieldspec 3®. The ASD spectral domain ranges from 350 to 2500 nm, with 1 nm band width. Within each plot, spectral measurements were made for each of the 3 to 4 randomly selected subplots. In each subplot, five spectral measurements were taken and later averaged to account for illumination and grass canopy structural differences as well as bidirectional effects (Mutanga et al., 2003; Wang et al., 2009). The measurements were taken between 10h30 and 15h00 on clear sunny days to minimize cloud effects and maximize illumination (Abdel-Rahman et al., 2010). A 25° field-of-view fibre optic was used. The fibre optic pistol was held at 1m above the ground and at nadir to cover the entire subplot. A Spectralon reference panel was utilized before each measurement to calibrate the sensor and convert spectral radiance to reflectance.

3.3 Data Analysis

3.3.1 Spectral pre-processing and transformation

The reflectance data were pre-processed before spectral transformations. Spectral data were subjected to smoothing with the commonly used Savisky-Golay filter (Savitzky and Golay, 1964), adding a second order polynomial least square function and 3-band window to remove signal noise. Commonly used spectral transformation techniques such as Log transformed spectra Log(1/R), first derivative, water removal and continuum removal, were computed. Log(1/R) was determined by calculating a log function of reciprocal of the spectral reflectance (Hruschka, 1987; Yoder and Pettigrew-Crosby, 1995; Fourty and Baret, 1998). The first derivative of the spectral reflectance was derived using a first-difference approach. A first-difference transformation of the reflectance spectrum calculates differences in reflectance between adjacent wavebands. More details on this can be found in Dawson and Curran (1998). The continuum removed spectra were derived by applying a convex hull or a continuum line to the reflectance spectra connecting local spectral maxima (Kokaly and Clark, 1999; Kokaly, 2001; Mutanga et al., 2004c). The water removal spectra (WR) were derived from a non-linear least-squares spectral matching technique calculating a fresh leaf spectrum as a non-linear combination of a leaf water spectrum and a dry matter spectrum (Gao and Goetz, 1994, 1995), modified by Schlerf et al. (2010) and Ramoelo et al. (2011b). The WR technique was implemented as per Ramoelo et al. (2011b).

3.3.2 Regression analysis and bootstrapping

Partial least square regression (PLSR) (Geladi and Kowalski, 1986; Naes et al., 1986; Ehsani et al., 1999; Martens and Naes, 2001; Viscarra Rossel, 2008) was used to predict the foliar N:P. PLSR was sought to be a robust statistical technique, and proved to reduce the problem of over-fitting (Viscarra Rossel, 2008). PLSR has been successfully used for foliar biochemical estimations (Huang et al., 2004; Asner and Martin, 2008; Ramoelo et al., 2011b). The performance of the various spectra and PLSR was measured using a bootstrapping approach (Efron, 1983). The advantage of bootstrapping is that it can be used efficiently when there is a limited sample size. Bootstrapping techniques iteratively resample the data set to be used for model development, making it an appropriate technique for assessing model accuracy (Verbyla and Litvaitis, 1989). To integrate PLSR and bootstrapping, bagging-PLSR was implemented using the Parles 3.1 software (Viscarra Rossel, 2007, 2008). The bagging PLSR has advantage of improved prediction, and derives robust models insensitive to over-fitting and provides an uncertainty measure for prediction by computing the confidence interval (Viscarra Rossel, 2008).

Using bagging-PLSR, independent or predictor variables were mean-centred to normalize them prior to further statistical analysis. The leave-one-out cross validation, as defined by the lowest root mean square error (RMSE), was used to determine the optimal number of factors or latent variables to be used for model development (Cho et al., 2007a; Darvishzadeh et al., 2008b; Viscarra Rossel, 2008). This Optimal number of factors was then used for model development and validation with 1000 bootstraps.

The retrieval accuracy of the PLSR model was defined by the bootstrapped mean of the coefficient of determination (R^2) and the RMSE. The confidence interval at a 95% confidence level was calculated for RMSE. The scores or the variable of importance for prediction (VIP) was computed to determine which bands contributed more in the foliar N:P prediction model development (Viscarra Rossel, 2008)**.** The wavebands or predictors with high scores were associated with the known absorption features by Curran (1989) and Kumar et al. (2001) respectively. Since ASD has a 1 nm band width, the maximum difference of 20 nm between the bands with high scores and known absorption features was chosen for the consistent comparison.

3.4 Results

The results showed that foliar N: P could be successfully estimated using *in situ* hyperspectral measurements and partial least square regression. This was evident by the performance of WR and CR spectra in combination with PLSR which produced a higher estimation accuracy of foliar N: P compared to the Log(1/R), R and FD spectra (Figure 3.2, Table 3.1). WR and CR spectra resulted to the similar root mean square error (RMSE) of 1.12, which equals to 14% of the mean (Table 3.1). CR spectra explained 85% of the variation of foliar N: P whilst WR explained 81% (Figure 3.2). The Log(1/R), FD and R spectra achieved foliar N: P estimation accuracy ranging between 16-19% of the mean. Additionally, Foliar N: P variation explained range between 69-76%, which is relatively lower than the ones for WR and CR (Table 3.1).

Table 3.1: The performance of each spectra and partial least square regression in estimating the foliar N: P and validated using bootstrapping.

	R^2	RMSE	95% LCI	95% UCI	No. of factors
CR	0.85	1.12	0.95	1.44	8
WR	0.81	1.12	0.93	1.40	9
R	0.76	1.22	1.01	1.52	13
Log(1/R)	0.73	1.30	1.07	1.62	12
FD	0.69	1.50	1.22	1.83	15

CR=continuum removal, WR=water removal, R=original reflectance, FD=first-difference derivative, LCI=Lower bound confidence interval at 95% confidence level. UCI=Upper bound

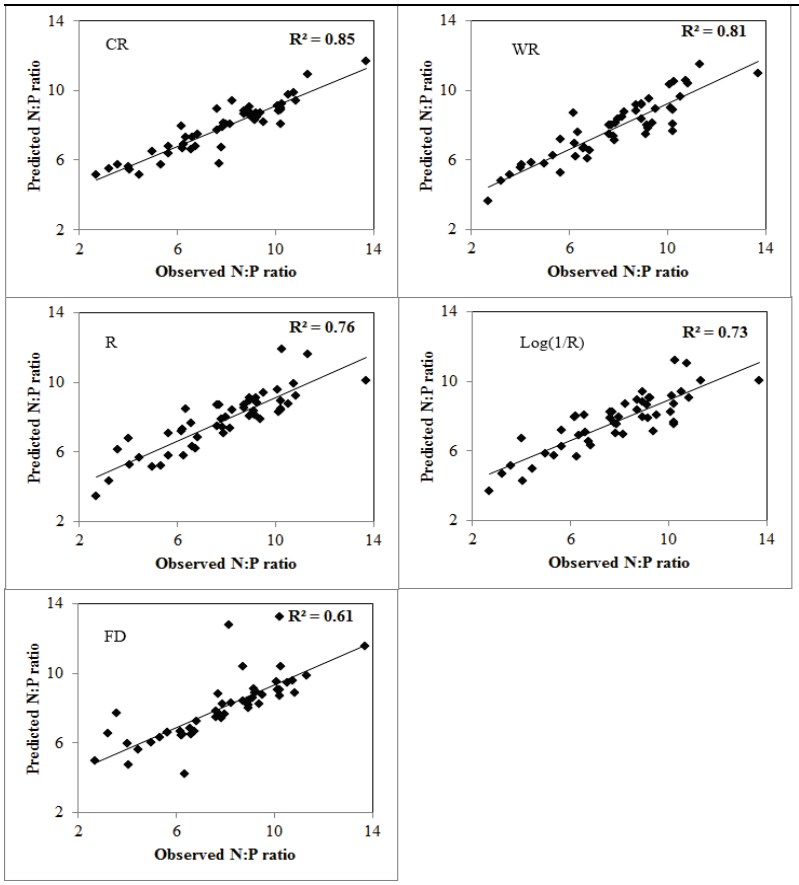

Figure 3.2: Scatterplots indicating the performance of each spectra and partial least square regression in estimating the foliar N: P validated using the bootstrapping. CR=continuum removal, WR=water removal, R=original reflectance, FD=first-difference derivative.

The results from the VIP (i.e. variable of importance for prediction) analysis derived for each spectral data and PLSR showed the importance of each band for predicting foliar N: P (Figure 3.3). Figure 3.3 shows the top 20 bands most important for predicting foliar N: P, using various spectral data. In Figure 3.3, the dark bars indicate bands associated with the known absorption features (e.g. N, protein), and grey is classified as one of the top 20 absorption features not thus far associated with known absorption features. For example, 90% of the selected bands from WR are associated with known absorption features, which 40% are that of protein and N absorption features. Sixty percent (60%) of the selected band from the reflectance spectra is associated with the known absorption features, while FD, CR as well as Log(1/R) yielded the percentage between 70-85%.

Table 3.2 shows that most of the various spectral data and PLSR models for estimating foliar N: P consistently select bands from the SWIR, for example; 1350-1360 nm, 1400-1450 nm, 1980-2000 nm, 2000-2060 nm and 2300-2360 nm. These substantiate the evidence in Figure 3.3, that most of the bands for estimating foliar N: P are located in the SWIR, and are mostly related to the known absorption features.

Table 3.2: The selected regions by each spectral data and partial least square regression, and the consistency of each spectral data to estimate foliar N: P ratio.

Spectral Data	A	B	C	D	E	F	G	H	I
WR				**X**	**X**	**X**	**X**		**X**
CR	**X**	**X**							**X**
R		**X**		**X**		**X**	**X**	**X**	
FD			**X**	**X**	**X**	**X**	**X**	**X**	**X**
Log(1/R)	**X**			**X**	**X**		**X**		**X**

WR=water removal, CR=continuum removal, R=reflectance, FD=first derivative, A=400-470nm, B=530-560nm, C=1000-1130nm, D=1350-1360nm, E=1400-1450nm, F=1980-2000, G=2000-2060, H=2200-2300nm, I=2300-2360nm.

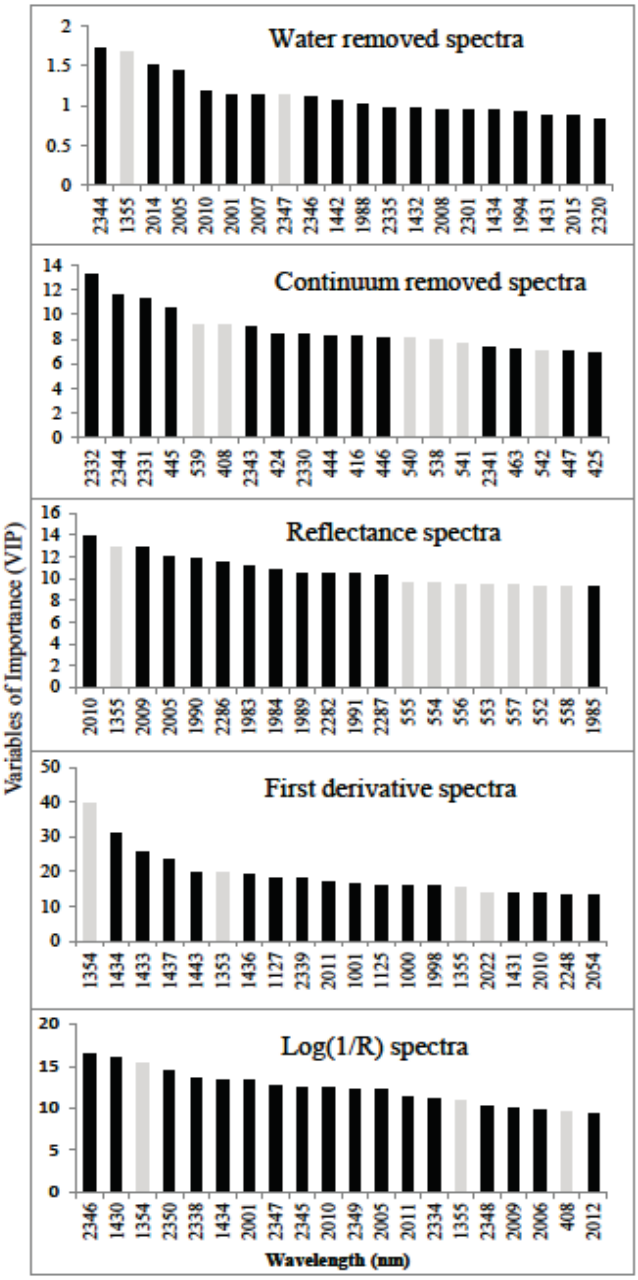

Figure 3.3: The variable of importance for prediction (VIP) i.e. important wavebands for predicting foliar N: P ratio derived from each spectral data and partial least square regression (PLSR). The displayed bars are the top 20 bands derived from the performance of each spectral data. The dark or bold bar indicates that selected bands are associated with known absorption features listed by Curran (1989) and Kumar et al. (2001) and greys are bands belonging to the top 20, but not associated with the known absorption features.

The foliar N:P data values are normally distributed as confirmed by Shapiro-Wilk normality test (W=0.983, p=0.69) (Royston, 1982). The variance of foliar N: P is relatively high indicated by 30% of the coefficient of variance (CV) and an average of 7.86 (Table 3.3). The descriptive statistics per geological and topographical types are presented in Table 3.3.

Table 3.3: Descriptive statistics for the foliar N: P, per geology and topography

Biochemicals	Min	Max	Mean	STDEV	Coefficient of Variation (CV)
Foliar N:P	2.63	14.00	7.89	2.35	0.30
N:P (Crest)	3.00	14.00	7.82	2.56	0.32
N:P (Valley)	2.60	10.45	7.13	2.37	0.33
N:P (Gabbro)	3.00	11.27	7.40	2.30	0.31
N:P (Granite)	2.66	13.59	7.67	2.58	0.34

3.6 Discussion

The study demonstrated that foliar N: P ratio can be estimated using various spectral data derived from the *in situ* hyperspectral and PLSR. The estimation of foliar N: P ratio emanated from the fact that foliar N and P were successfully estimated using remote sensing, with inconsistent retrieval accuracy. The expectation of this study was based on the assumption that the various spectral data, known absorption features and PLSR (a robust statistical technique) could be successfully used to predict foliar N: P ratio. Results showed that WR and CR spectral data achieved higher foliar N: P ratio estimation accuracy. The performance of WR is attributed to the fact that it minimizes the water absorption effects on the sensitive weak or subtle regions of the foliar biochemical concentrations (Gao and Goetz, 1994, 1995; Schlerf et al., 2010; Ramoelo et al., 2011b). Coincidentally, the regions of the reflectance spectra most affected by water absorption are regions very important for foliar biochemicals estimation. The WR technique was developed by Gao and Goetz (1994; 1995) solely to address this problem, especially on fresh leaf spectra. Gao and Goetz, (1994; 1995) tested the WR technique for estimating foliar cellulose and lignin and it was successful. Ramoelo et al. (2011b) demonstrated that WR could be used to as one of the techniques for estimating foliar biochemicals, as shown from the experimental study in the controlled environments. WR spectra are applied here for the first time to estimate foliar N: P ratio and yields promising results. The CR spectra enhances the absorption features of foliar biochemical concentrations through enhancing the differences in the absorption strength (Kokaly and Clark, 1999; Huang et al., 2004; Mutanga et al., 2004c; Mutanga et al., 2005). CR spectra have been used for estimating foliar N and P, not N: P ratio but showed a good performance. Several studies

demonstrated the applicability of CR spectra for foliar biochemicals, especially for foliar N and P (Huang et al., 2004; Ramoelo et al., 2011b). For example, Mutanga et al., (2004) successfully estimated both foliar N and P with the coefficient of determination between 43 to 80% using continuum removal derived from the *in situ* hyperspectral data. Huang et al., (2004) achieved a coefficient of determination of 0.65 for estimating foliar N using airborne hyperspectral data. The latter studies showed consistent performance of CR on foliar biochemical estimations, which are the similar trend observed by this study. The reflectance, Log(1/R) and FD spectra showed a possibility to estimate foliar N: P. In most cases, Log(1/R) and FD have higher foliar biochemicals retrieval accuracy than using the reflectance spectra (Yoder and Pettigrew-Crosby, 1995; Johnson, 2001), and this differs with the results we got in this study. This could be attributed to the fact that the foliar N: P interacts differently with spectra as compared to N or P, especially in terms of physical vibration bonds. Yoder and Pettigrew-Cosby (1995), Johnson (2001) and Fourty and Baret, (1996) found that Log(1/R) improves estimation of foliar N over reflectance, since there is a linear relationship between this foliar biochemicals and its contribution to the Log(1/R) at the absorbed wavelength (Hruschka, 1987; Kumar et al., 2001).

Selected bands in Figure 3.3 were associated with the known absorption features which are listed in Curran (1989), Knox et al. (2010) and Kumar et al. (2001), and also dominate the SWIR region (Table 3.2). This is an indication that the foliar N: P ratio is sensitive to the physical vibration bonds of foliar biochemical concentrations. The vibration mechanisms associated with these features are either, C-O, O-H, C-H or N-H (Curran, 1989; Kumar et al., 2001). Using WR spectra, 40% of the 90% of known absorption features were associated with N and protein. According to the range of values for nutrient limitation by Cech et al. (2008), our study area is limited by N and P because foliar N: P values are between 3 and 14. In this study, nutrient limitation could be associated with the contrast between low fertile soils derived from granite and high fertile soils derived from gabbro geological types (Scholes, 1990; Venter et al., 2003; Grant and Scholes, 2006). Similarly, Ludwig et al. (2001) found out that under trees in an east African savanna ecosystems, 12 is already indicating that P is limiting, which means according to the range of values of foliar N:P ratio in the study, N or P could be limiting in some areas than the other. In addition, Ludwig et al. (2001) found out that the open grassland have an average foliar N: P ratio of 6, indicating that N is limiting, with similar results observed in this study (mean foliar N: P ratio of 7.86). Craine et al. (2008) also found out that unfertilized vegetation had a mean N: P ratio of 5.8, during their experiments in KNP, which is similar to the N: P values in the open grassland found by Ludwig et al. (2001), and by Cech et al. (2008) in the tropical savanna. The difference between Koerselman and Meuleman (1996) in the wetlands or

temperate ecosystems, as well as Ludwig et al. (2001) and Cech et al. (2008) in the savanna grasses could be associated with the fact that savanna grasses have relatively lower N requirements as compared to wetlands or temperate plants. The explanation could be that, plants with C4 photosynthetic pathways have a higher nutrient efficiency than C3 plants (Wolfson and Tainton, 2000; Craine et al., 2008).

The study demonstrated that foliar N: P ratio can be estimated using *in situ* hyperspectral measurements. Since this is first study to do this, then there is a need to test the applicability of using airborne hyperspectral remote sensing data such as HyMap, to understand the landscape variability of foliar N: P ratio. The challenge for estimating foliar N: P at regional scale (up-scaling) could be associated with the fact that satellite or multispectral remote sensing data have limited bands in the SWIR region, and this study indicated that most of the bands sensitive to foliar N:P are located in the SWIR. There is a need for several experimental studies for understanding how foliar N: P ratio influence grass productivity, and hence the feeding patterns of herbivores in the savanna ecosystems. Such experimental studies could play a crucial role in determining critical values of foliar N: P for ascertaining nutrient limitation on grass productivity. The foliar N: P ratio as an indicator for nutrient limitation could be useful information to the ecologists, resource managers, farmers and park managers to understand which between N and P is limiting and how the limitation influence the resource selection (Manly et al., 1993), distribution, densities and population dynamics of herbivores (wild and livestock), at landscape level.

3.7 Acknowledgement

We would like to thank the South Africa Department of Science and Technology, South Africa's National Research Foundation (Professional Development Programme), Council for Scientific and Industrial Research, University of Twente and Wageningen University for the funding. I would also like to thank Mr. Russell Main, Mr. Laven Naidoo and Mr. Thulani Selaule for fieldwork assistance. Finally, we appreciate the involvement of South African National Parks (SANPARKS), especially Dr. Izak Smith, Mrs. Thembi Khoza and Mrs. Patricia Khoza and Mr. Adolf Manganyi for making the fieldwork possible.

Chapter 4

Grass nitrogen and phosphorus estimation using integrated modelling approach*

*Chapter is based on: Ramoelo A. Skidmore A.K. Cho M.A. Mathieu R. Heitkönig I.M.A. Dudeni-Tlhone N. Schlerf M. Prins H.H.T. Non-linear partial least square regression increases the estimation accuracy of grass nitrogen and phosphorus using in situ hyperspectral and environmental data, *ISPRS Photogrammetry and Remote Sensing* (**under review**)
Partly based on: Ramoelo A. Cho M.A. Mathieu R. Skidmore A.K. Schlerf M. Heitkönig I.M.A. Prins H.H.T. (2011). Integrating environmental and *in situ* hyperspectral remote sensing variables for grass nitrogen estimation in savanna ecosystems, *34th International Symposium on the Remote Sensing of Environment (ISRSE 2011), The GEOSS Era: Towards Operational Environmental Monitoring*, Sydney, Australia, http://www.isprs.org/proceedings/2011/ISRSE-34/index.html **(published)**.

Abstract

Grass nitrogen (N) and phosphorus (P) concentrations are direct indicators of rangeland quality and provide imperative information for sound management of wildlife and livestock grazing. It is challenging to estimate grass N and P concentrations using remote sensing in the savanna ecosystems. These areas are diverse and heterogeneous in soil and plant moisture, soil nutrients, grazing pressures and human activities. The objective of the study is to test the performance of non-linear partial least squares regression (PLSR) for predicting grass N and P concentrations through integrating *in situ* hyperspectral and environmental variables (climatic, edaphic and topographic). The data were collected along a land use gradient in the greater Kruger National Park region. The data consisted of: (i) *in situ*-measured hyperspectral spectra, ii) environmental variables and measured grass N and P concentrations. The hyperspectral variables included published starch, N and protein spectral absorption features, red edge position, narrow-band indices such as simple ratio (SR) and normalized difference vegetation index (NDVI). The results of the non-linear PLSR were compared to those of conventional linear PLSR. Integrating non-linear PLSR, integrating *in situ* hyperspectral and environmental variables yielded the highest grass N and P estimation accuracy (R^2=0.81, root mean square error (RMSE) =0.08, and R^2=0.80, RMSE=0.03, respectively) as compared to using remote sensing variables only, and conventional PLSR. The study demonstrates the importance of an integrated modelling approach for estimating grass quality which is a crucial effort towards effective management and planning of protected and communal savanna ecosystems.

4.1 Introduction

Spatial patterns of grass nitrogen (N) and phosphorus (P) are known to influence the grazing behaviour and migration patterns of wildlife and livestock in savanna landscapes (Drent and Prins, 1987; McNaughton, 1988, 1990; Seagle and McNaughton, 1992; Prins and van Langevelde, 2008). In Southern Africa, large herbivores are found in high numbers around nutrient rich areas e.g. termite mounds, sodic sites, or sites beneath large trees (Grant and Scholes, 2006; Treydte et al., 2007; Ludwig et al., 2008). Furthermore, the N:P ratio is postulated as one of the key indicators of nutrient limitation in savanna ecosystems (Koerselman and Meuleman, 1996; Ludwig et al., 2001; Prins and van Langevelde, 2008). Therefore, an accurate assessment of the spatial patterns of N and P could play a vital role in the effective planning and management of savanna rangelands for sustainable livestock and wildlife grazing production.

Communal savanna ecosystems serve as a source of livelihood for the rural community through providing valuable good and services including fuel wood

(for cooking and heating) and grazing land (for livestock production) (Shackleton et al., 2002). Sustainable livestock production depends on the quality of the grazing land. One of the causes of grazing land degradation is overgrazing resulting from poor land planning and management of grazing lands, mainly in the communal rangelands (Abel and Blaikie, 1989; Du Toit and Cumming, 1999). Therefore, information on the spatial patterns of grass quality will support sustainable rangeland management.

Remote sensing is widely used as a cost-effective means to map rangeland or savanna grass quality at landscape level (Mutanga and Skidmore, 2004a; Mutanga and Kumar, 2007; Skidmore et al., 2010; Knox et al., 2011). This notwithstanding, conventional broadband remote sensing techniques based on the utilization of the relationship between grass quality (N and P) and spectral indices such as normalized difference vegetation index (NDVI) (Tucker, 1979), soil line concept (SLC), simple ratio (SR) (Baret and Guyot, 1991), soil-adjusted vegetation index (SAVI) (Huete, 1988) have limited applications in high grass canopy environments as they saturate at high canopy cover (Tucker, 1977; Mutanga and Skidmore, 2004b). On the other hand, the use of spectral indices derived from the red-edge bands (700 – 750 nm) of hyperspectral or narrow-band data has been demonstrated to mitigate the saturation effect observed with broadband indices (Clevers et al., 2002; Huang et al., 2004; Cho and Skidmore, 2006; Darvishzadeh et al., 2008b; Majeke et al., 2008). The red-edge is the region of abrupt change in foliar reflectance between 680 and 780 nm (Clevers et al., 2002). Narrow-band normalized difference vegetation index and SR indices computed from red-edge bands provided more accurate estimates of foliar N compared to conventional NDVI derived from 680 and 800 nm (Mutanga and Skidmore, 2007). Many other studies have identified several absorption features for N and protein (Curran, 1989; Elvidge, 1990; Kokaly and Clark, 1999; Kumar et al., 2001; Cho et al., 2010; Knox et al., 2010b; Skidmore et al., 2010). Specific absorption features for P have not been identified, but several studies found that the short-wave infrared (SWIR) bands have a potential for predicting foliar P concentration (Mutanga and Kumar, 2007; Cho et al., 2010; Ramoelo et al., 2011b). Spectral transformation techniques such as water and continuum removal have been proposed to enhance nutrient absorption features (Huang et al., 2004; Mutanga et al., 2004c; Cho et al., 2010; Schlerf et al., 2010; Ramoelo et al., 2011b).

Savanna ecosystems are diverse and heterogeneous in soil and plant moisture, soil nutrients, fire regime, grazing pressures and anthropogenic activities (Ben-Shahar and Coe, 1992). Thus, making the estimation of grass N and P using remote sensing in savannas a challenging venture (Mutanga et al., 2004c; Mutanga and Kumar, 2007; He and Mui, 2010; Skidmore et al., 2010). Grass quality is influenced by geology (Ben-Shahar and Coe, 1992;

Grant and Scholes, 2006), soil (Heitkönig and Owen-Smith, 1998; Cho et al., 2010), precipitation and temperature (Ben-Shahar and Coe, 1992), topography or catena position (Seagle and McNaughton, 1992; Mutanga et al., 2004a) as well as aspect (Mutanga et al., 2004a) and land use types. We assume that a modelling approach that exploits the strength of environmental variables and remote sensing data could potentially improve the assessment of ecosystem state and functioning at various geographic scales (Mutanga et al., 2004a; Cho et al., 2009; Knox et al., 2011). The integrated modelling approach could be an attempt towards estimating and mapping foliar N and P at regional scale, which according to our knowledge is yet to be done. A limited number of studies have investigated the possibility of integrating environmental and remote sensing variables to estimate foliar N and P concentrations e.g. Cho et al. (2009) and Knox et al. (2011).

Several studies have successfully used stepwise multiple linear regression (SMLR) (Grossman et al., 1996; Martin and Aber, 1997; Kokaly and Clark, 1999; Huang et al., 2004) to estimate N and P with hyperspectral remote sensing variables. However, SMLR operates on the assumption of normal distribution of the data, and could suffer from model overfitting and multicollinearity (Grossman et al., 1996; Huang et al., 2004). The use of partial least square regression (PLSR) has been advocated to address these issues (Geladi and Kowalski, 1986; Huang et al., 2004; Asner and Martin, 2008; Darvishzadeh et al., 2008b; Ramoelo et al., 2011b). The conventional PLSR also makes a normality assumption about the distribution of the response variable. Input data can be normalized using mean or median centring prior to use with the conventional PLSR (Viscarra Rossel, 2008), but this does not completely address the requirement for normal distribution. To curtail this assumption, a non-linear PLSR has been developed, also known as non-linear PLSR with radial basis function neural network (RBF-PLSR) (Walczak and Massart, 1996). The advantage of the non-linear PLSR is that it is a flexible non-linear regression technique which combines the capability of the conventional PLSR, i.e., power to maximize covariance between data sets, and the non-linear nature of the RBF neural network (Walczak and Massart, 1996). The predictive models developed by RBF-PLSR have limited or no overfitting and multicollinearity problems if the optimal number of latent variables is selected (Walczak and Massart, 1996). RBF-PLSR is also non-parametric in nature and it does not require model input to be normally distributed. The technique has been successfully applied in soil (Fidêncio et al., 2002), time series prediction (Zemouri et al., 2003), air pollution (Giering et al., 2005) and engineering related fields (Garg et al., 2010). The performance of the non-linear PLSR has not been established for extracting vegetation biochemistry in heterogeneous savanna ecosystems.

The aim of the study was (i) to assess and compare the retrieval accuracy of grass N and P concentrations when using conventional *vs.* non-linear PLSR, and (ii) to test the performance of non-linear PLSR for integrating *in situ* hyperspectral and environmental variables (climatic, edaphic and topographic) for predicting grass N and P concentrations. The conventional and non-linear PLSR techniques were implemented with remote sensing variables only and subsequently with the integrated environmental and *in situ* hyperspectral variables. Conventional and non-linear PLSR *vs.* integrated modelling results were compared.

4.2 Material and Methods

4.2.1 Study area and sampling design

The study area is located in the Lowveld savanna at the north-eastern part of South Africa (Figure 4.1). The Lowveld landscape corresponds to the low lying area extending from the foot slopes of the Drakensberg Great Escarpment to the west and the Mozambique coastal plain to the east (Venter et al., 2003). The topography is gently undulating with flat patches in localized areas, and with an average height of 450m a.s.l. The study area covers a land use transect ranging from protected areas such as the private-owned Sabi Sands Game Reserve (SGR) and the state-owned Kruger National Park (KNP) to communal lands in the Bushbuckridge region. The western part of the transect (communal areas) receives higher mean annual rainfalls (800mm/yr.) as compared to the eastern side of the transect (580 mm/yr.) (Venter et al., 2003). The annual mean temperature is about 22^{o}C. The dominant geology includes granite and gneiss with local intrusions of gabbro (Venter et al., 2003). Consequently, these areas are characterized by gradients of soil moisture and nutrients. The soil fertility of gabbro areas are higher than the granitic ones (Ben-Shahar and Coe, 1992; Venter et al., 2003). The main vegetation communities include the "granitic lowveld" and the "gabbro grassy bushveld" (Mucina and Rutherford, 2006). In the gabbro patches, grass species such as *Setaria sphacelata* dominates the crest while species such as *Urochloa mosambicensis* dominates the valleys. Gabbro patches are dominated by grass species with high productive potential (e.g. *Urochloa mosambicensis*) compared to granite-derived soils (e.g. *Eragrostis rigidior and Pogonarthria squarrosa; cf.* Mutanga et al., (2004). The gabbro sites are dominated by fine leaves tree species such as *Acacia ssp* while the granite sites are dominated by broadleaves tree species such as *Combretum spp* and *Terminalia spp* (Venter et al., 2003; Ferwerda et al., 2006). Rangelands in the protected areas are grazed by wild herbivore such as impala (*Aepyceros melampus*), zebra (*Equus burchelli*), wildebeest (*Connochaetes taurinus*), buffalo (*Syncerus caffer*), etc., while the communal rangelands support grazing of cattle (*Bos taurus*) and goats (*Capra hircus*) as well as sheep (*Ovis aries*), which determine various grazing intensities.

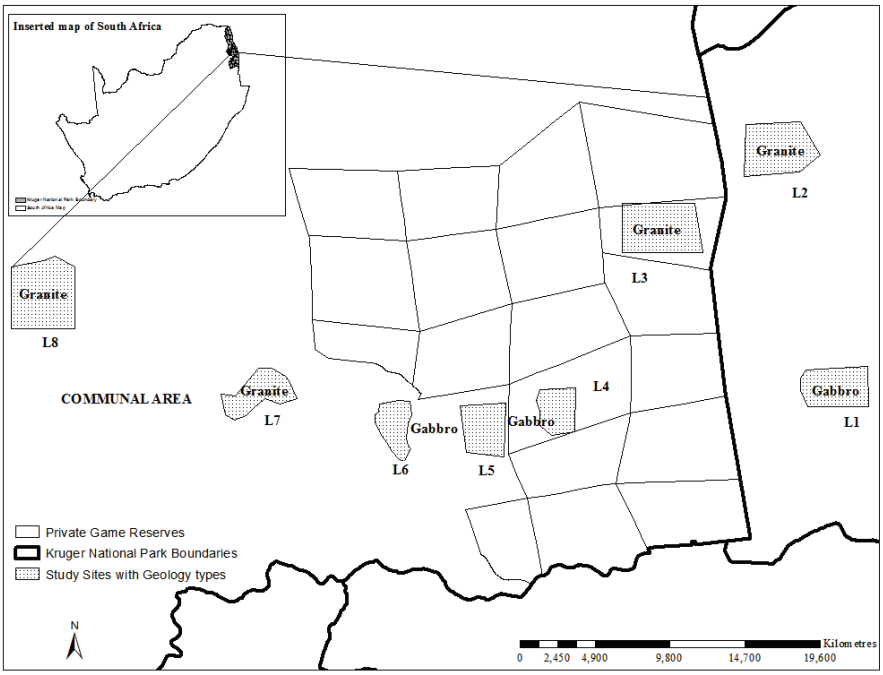

Figure 4.1: Study area map. L=Land use

The study area consisted of eight experimental sites which were placed along the land use gradient: two sites in KNP (L1 gabbro, L2 granite), two sites in SGR (L3 granite, L4 gabbro), and four sites in the communal areas (L5-6 gabbro, L7-8 granite) (Figure 4.1). The sites were demarcated using 1:250, 000 geology maps and refined using 2008 SPOT 5 images (Wessels et al., 2011). The site selection process sought to capture the nutrient contrast from low to high in granitic-derived soils to gabbro-derived soils, respectively and along the rainfall gradient. A line transect sampling design was used to collect field data (Fewster et al., 2005) in each site except L3 (because of access limitations). To better capture the grass biomass variability, transects were laid out to sample both valley and crest land units. The topography influences the grass biomass in the savanna ecosystems with valley areas generally having higher grass biomass than crest areas. Along transects, a combination of purposive and systematic placement of sampling plot was done. The distance between the plots was between 500m and 1000m depending on the accessibility and homogeneity of the area. The plot size was 30 m x 30 m. A total of 49 plots were surveyed and in each plot three to four subplots (0.5 m x 0.5m) were randomly selected to capture the plot variability. In each subplot, data on the sample location using the Leica®'s GS20 differential geographic positioning system (DGPS), dominant grass

species and grass samples were collected. Grass samples were dried at 80^0C for 24 hours and the measurements were later averaged at plot level. The DGPS points were post-processed using Leica's GeoPro software and reference GPS data from Nelspruit station to produce less than 1 m positional accuracy. The fieldwork was undertaken in March/April 2009 towards the end of the wet season, when the grass biomass was at full maximum growth or peak productivity to minimize the N/P and biomass interaction effects (Skidmore et al., 2010).

4.2.2 Chemical analysis

The dried grass samples were taken to the South African's Agricultural Research Council-Institute for Tropical and Subtropical Crops (ARC-ITSC)-Nelspruit for chemical analysis. Firstly, the acid digestion technique was used, where perchloric and nitric acids were used for foliar P concentration retrieval and sulphuric acid was used for retrieving foliar N concentrations (Giron, 1973; Grasshoff et al., 1983; Mutanga et al., 2004a). Secondly, the colorimetric method by auto analyser was used to measure foliar N (Technicon Industrial Method 329-74 W; Technicon Industrial Systems, Farrytown, New York). For foliar N measurements an emerald-green colour was formed by the reaction of ammonia, sodium salicylate, sodium nitroprusside and sodium hypochlorite. The ammonia-salicylate complex was read at 640 nm. For foliar P measurements, a colorimetric in which a blue colour was formed by the reaction of ortho-phosphate and the molybdate ion. The phosphomolybdenum complex was then read at 660 nm. These extraction methods were successfully used for foliar N and P by Mutanga et al. (2004) and Ramoelo et al. (2011b). Chemically analyzed N and P were henceforth referred to as observed N and P.

4.2.3 Canopy Spectral measurements

The reflectance spectra were measured using an Analytical Spectral Device (ASD) spectroradiometer, Fieldspec 3®. The ASD spectral domain ranges from 350 to 2500 nm, with 1 nm band width. Within each plot, spectral measurements were made for each of the 3 to 4 randomly selected subplots. In each subplot, five spectral measurements were taken and later averaged to account for illumination and grass canopy structural differences as well as bidirectional effects (Mutanga et al., 2003; Wang et al., 2009). The measurements were taken between 10h30 and 15h00 on clear sunny days to minimize cloud effects and maximize illumination (Abdel-Rahman et al., 2010). A 25° field-of-view fibre optic was used. The fibre optic pistol was held at 1m above the ground and at nadir to cover the entire subplot. A Spectralon reference panel was utilized before each measurement to calibrate the sensor and convert spectral radiance to reflectance.

4.2.4 Spectral indices and selection of absorption features

Red-edge position (REP), narrow band indices such as narrow normalized difference vegetation index, simple ratio (SR), and known absorption features of N and protein were used for N estimation and several spectral features of leaf and canopy biochemistry were selected for P estimation.

The red-edge is highly correlated with N and is less sensitive to soil background reflection (Cho and Skidmore, 2006). For this study REP was calculated using the linear extrapolation technique (Cho and Skidmore, 2006). Cho and Skidmore (2006) found out that the linear extrapolation technique achieved higher accuracy in retrieving N and chlorophyll as compared to other red-edge detection techniques. The NDVI is the most widely known vegetation index used as a surrogate for vegetation condition and health in many studies (Zhao et al., 2007). It has been reported to minimize the atmospheric effects on remote sensing data (Zarco-Tejada et al., 2004b). The narrow-band NDVI has been proposed to minimize problems of asymptotic saturation of biomass assessment particularly during the peak productivity (Mutanga and Skidmore, 2004b). A narrow band simple ratio (SR) was also computed using the red edge spectral bands. The advantages above-mentioned for narrow-band NDVI also apply for SR derived from the red-edge region. Narrow-band SR and NDVI have been successfully used for estimating vegetation parameters, e.g. chlorophyll and nitrogen concentrations, biomass and leaf area index (Mutanga and Skidmore, 2004b; Darvishzadeh et al., 2008b).

Chlorophyll, protein and N absorption features were selected for estimating foliar N concentrations (Curran, 1989; Kumar et al., 2001) (Table 4.1). Since foliar P does not have specific known absorption features, chlorophyll, protein, sugar and starch absorption features were used instead (Curran, 1989; Kumar et al., 2001). The listed absorption features for N, and protein have been successfully used for foliar N (Knox et al., 2010b; Schlerf et al., 2010; Skidmore et al., 2010), while Ramoelo et al. (2011b), Mutanga and Kumar (2007), Bogrekci and Lee (2005), and Cho et al. 2010 found that foliar P concentration is sensitive to the bands located in the shortwave infrared. Therefore, most the selected absorption features dominates the SWIR region.

Table 4.1: Absorption features used for foliar N and P estimation (Curran 1989; Kumar et al., 2001)

Nutrients	Absorption features (wavelength)
Nitrogen	430 nm, 460 nm, 640nm, 660nm, 910nm, 1510 nm, 1940 nm, 2060 nm, 2180 nm, 2300 nm, 2350 nm
Phosphorus	430 nm, 460 nm, 640nm, 660nm, 910nm, 1510 nm, 1940 nm, 2060 nm, 2180 nm, 2300 nm, 2350 nm 970 nm, 990 nm, 1200 nm, 1450 nm, 1530 nm, 1540 nm, 1580 nm, 1780 nm, 1940 nm, 2000 nm, 2080 nm, 2100 nm, 2250 nm, 2280 nm and 2320 nm

4.2.5 Environmental data

Environmental variables influence the distribution of foliar N and Including precipitation, temperature, land use, geology, soils, distance to rivers, altitude, slope and aspect (Table 4.2). Climate, topography and geologic substrate influence the distribution of the primary environmental regimes such as moisture and nutrients in soils or plants, see Skidmore et al. (In press), Pickett et al. (2003), Venter et al. (2003) and Mutanga et al (2004). Annual average precipitation and daily temperatures were acquired from the World Climate database (WorldClim) (www.WorldClim.com). The Digital Elevation Model (DEM) was produced at 50 m spatial resolution using contours and spot height data from 1:50 000 topographical maps acquired from South Africa's Department of Rural Development: Surveys and Mapping. Slope and aspect were derived from the DEM using ArcGIS 10x. The river layer was sourced from the South African National Botanical institute (SANBI)'s Beta version of vegetation data sets (Mucina and Rutherford, 2006). The distance to river variable was computed using the Spatial Analyst Tool embedded in ArcGIS 10x, where the river layer and the sample plot locations were used as an input. A soil layer was acquired from the soil and terrain database of Southern Africa (SOTERSAF) (Dijkshoorn, 2003; FAO et al., 2003; Dijkshoorn et al., 2008). The land use types (3 classes; public conservation lands, private conversation lands, and communal rangelands) were derived from the boundary layers of KNP, Sabi Sands and communal areas acquired from KNP's Geographic Information System (GIS) and remote sensing laboratory.

Table 4.2: Environmental data used for the study

Environmental Data	Type	Source	Resolution
Precipitation	Continuous	http://www.worldclim.com/	1 km
Temperature	Continuous	http://www.worldclim.com/	1 km
Land use types	Categorical	KNP	Vector layer
Geology	Categorical	Council for Geoscience	1:1000000
Altitude (DEM)	Continuous	DRDLR, South Africa	50 m
Slope	Continuous	Derived from DEM	50 m
Aspect	Continuous	Derived from DEM	50 m
Distance from rivers	Continuous	SANBI GIS data	1:1000000
Soil	Categorical	SOTERSAF database	1:1000000

DEM= digital elevation model, CSIR=Council for Scientific and Industrial Research, SANBI=South African National Botanical Institute, SOTER=Soil and Terrain of Southern Africa database, DRD=Department of Rural Development and Land Reform, KNP=Kruger National Park GIS datasets

4.2.6 Statistical analysis and modelling

Both non-linear and conventional linear partial least square regressions (PLSR) were used for data analysis. Any PLSR technique aims at

decomposing a list of independent variables into latent and uncorrelated variables to minimize the dimensionality problems associated with the raw data sets (Geladi and Kowalski, 1986; Naes et al., 1986; Martens and Naes, 2001; Viscarra Rossel, 2008). The conventional linear PLSR used in this study refers to the technique developed or used by Geladi and Kowalski (1986), Naes et al. (1986), Ehsani et al. (1999) and Viscarra Rossel (2008). The non-linear PLSR with radial basis function neural network (RBF-PLSR) is proposed for estimating foliar N and P concentrations as it is a flexible technique which can predict both non- and normally distributed response variables (Walczak and Massart, 1996; Daszykowski et al., 2007). RBF-PLSR has the mutual advantages of the non-linear nature of RBF and of the power of PLSR to maximize covariance between data sets (Walczak and Massart, 1996). The detailed theory behind RBF-PLSR can be found in Walczak and Massart (1996). The input variables were standardized or scaled to a range of [0-1] (Mutanga and Kumar, 2007; Skidmore et al., 2010; Knox et al., 2011) prior to implementing the non-linear PLSR. The implementation of the radial basis function was done by constructing a model or an activation matrix using Gaussian functions with different widths defined by their sigma values (from 0.1 to 1 with a step of 0.1). The PLSR is then applied to the activation matrix to estimate biochemicals. The scores in the activation matrix are the linear combinations of the Gaussian functions maximizing the covariance between N/P and the activation values.

The Monte-Carlo leave-one-out cross-validation technique was used to determine the optimum number of latent factors based on the lowest RMSE (Daszykowski et al., 2007), which also correspond to a particular sigma value. The Monte Carlo leave-one-out cross validation was used for validation because the available dataset (49 samples) was too small to be effectively divided into a training and test dataset. The advantage of the leave-one-out cross-validation is that it not biased, since it uses 48 samples for data calibration to predict the remaining 1 iteratively (Darvishzadeh et al., 2008b). The non-linear RBF-PLSR was implemented in the Matlab tool box for multivariate calibration techniques (TOMCAT). The software description details can be found in Daszykowski et al. (2007). The conventional PLSR weights were further analyzed and interpreted to determine whether there was a positive or negative contribution of each variable in the foliar N and P models, which is the information the non-linear PLSR technique does not provide. Correlation matrices were computed to assess the relationships between environmental variables and grass nutrient concentrations, and to help the interpretation of the integrated modelling outputs. Non-parametric spearman correlation was used because this method handles both continuous and categorical data sets irrespective of their statistical distribution, and was implemented in R programming language (Hollander and Wolfe, 1973; Lehman, 1998).

4.3 Results

4.3.1 Integrated modelling using non-linear PLSR for foliar N estimation

The results showed that non-linear PLSR with integrated in situ remote sensing and environmental variables yielded a higher foliar N estimation accuracy (R2=0.81, RMSE=0.08, 11.4% of the mean) as compared to the use of remote sensing variables only (R2 =0.66, RMSE=0.11: 15.7% of the mean) (Figure 4.2, Table 4.3). Integrating in situ hyperspectral and environmental variables with the non-linear PLSR yielded higher estimation accuracy than with the conventional PLSR. The conventional PLSR explained 64% and 58% of the variance of grass N concentration with integrated in situ hyperspectral and environmental variables, and with remote sensing variables only, respectively (Figure 4.2, Table 4.3). Generally, the non-linear PLSR led to higher estimation accuracy of grass N than the conventional PLSR, both considering remote sensing and environmental variables and remote sensing variables only (Table 4.3).

Table 4.3: Performance for foliar N and P prediction through integrating environmental and *in situ* hyperspectral variables as compared to using *in situ* hyperspectral variables only, utilizing conventional and non-linear partial least square regression

	Conventional PLSR				Non-linear PLSR			
	R^2	RMSECV	RMSE	No. of factors	R^2	RMSECV	RMSE	No. of factors
N *vs*. RS	0.58	0.11	0.12	8	0.66	0.11	0.11	10
N *vs*. RS+Env.	0.64	0.11	0.11	8	0.81	0.11	0.08	8
P *vs*. RS	0.36	0.03	0.04	8	0.44	0.04	0.04	9
P *vs*. RS+Env.	0.38	0.04	0.04	12	0.80	0.03	0.02	13

N=nitrogen, P=phosphorus, RS=remote sensing variables (all as given in Table 4.1 and in the text), Env=environmental variables (all as given in Table4. 2), PLSR=partial least square regression, RMSE=root mean square error, RMSECV= root mean square error of cross validation.

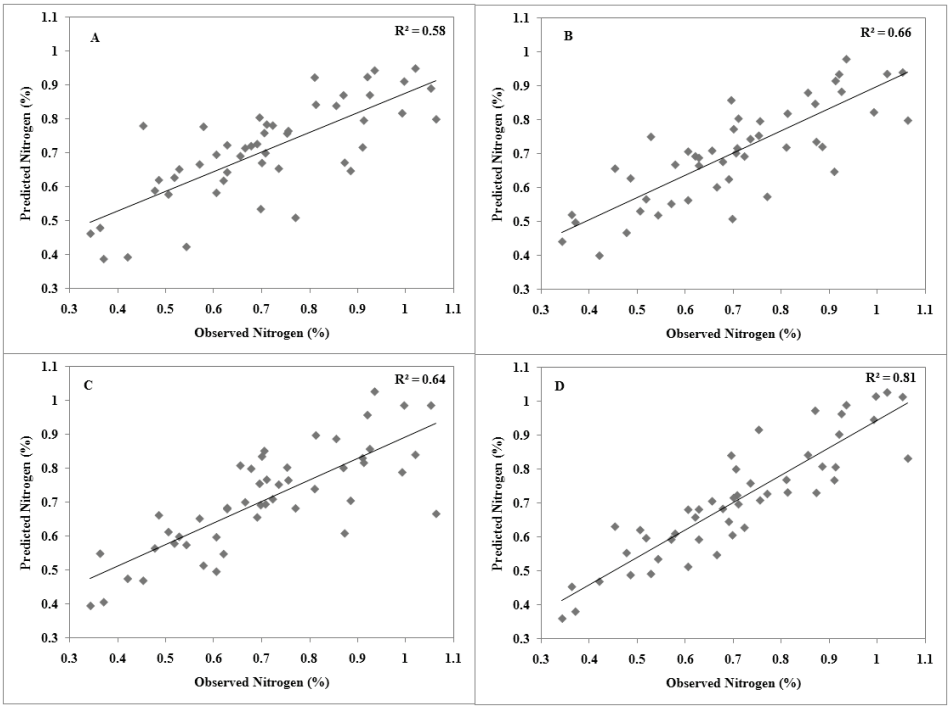

Figure 4.2: A comparison of a conventional and non-linear partial least square regression for foliar N estimation derived through Monte-Carlo leave-one-out cross validation: A= conventional PLSR vs. remote sensing variables only and B= non-linear PLSR vs. remote sensing variables only, C= conventional PLSR vs. remote sensing + environmental variables, D= non-linear PLSR vs. remote sensing + environmental variables.

Remote sensing variables such as narrow-band SR and REP positively contributed to the N model while protein absorption features at 910 nm and 1020 nm yielded the negative contribution as shown by the PLSR weights in Figure 4.4. Geology, soil types, land use, distance to rivers and temperature resulted to a positive contribution to the N prediction model, while slope contributed negatively as shown by the PLSR weights in Figure 4.4. Table 4.5 reports the non-parametric spearman correlation matrix of foliar N and environmental variables. Correlations between foliar N and environmental variables were generally not significant ($p<0.05$), with only weak relationships with slope and aspect, while a high correlation was found between precipitation and land use types or temperature, respectively (Table 4.4). The measured foliar N has a mean value of 0.7% and a coefficient of variation value of 26 %, which shows that the variability of N in grass leaves across the study area is not very high (Table 4.6).

Table 4.4: correlation matrix between N and all environmental variables

	N %	Geo	Land	Soil	DRivers	Alt	Slope	Aspect	Precip	Temp
N %	1.00									
Geo	0.03	1.00								
Land	0.06	0.09	1.00							
Soil	0.12	**0.54**	0.24	1.00						
DRivers	0.02	0.75	**-0.47**	**0.43**	1.00					
Alt	-0.03	**0.60**	**0.71**	**0.69**	0.21	1.00				
Slope	**-0.21**	**0.53**	0.32	0.31	0.26	0.48	1.00			
Aspect	**-0.20**	0.16	**-0.29**	-0.10	0.30	-0.06	0.10	1.00		
Prep	0.05	**0.44**	**0.75**	0.36	0.04	0.81	0.51	-0.11	1.00	
Temp	0.11	**-0.50**	**-0.55**	**-0.41**	-0.16	-0.81	-0.40	-0.02	-0.74	1.00

Geo=geology, Land=land use, Drivers=distance to rivers, Alt=altitude, Precip=precipitation, Temp=temperature, Bold values indicates the significance correlation at 95% significance level (p<0.05).

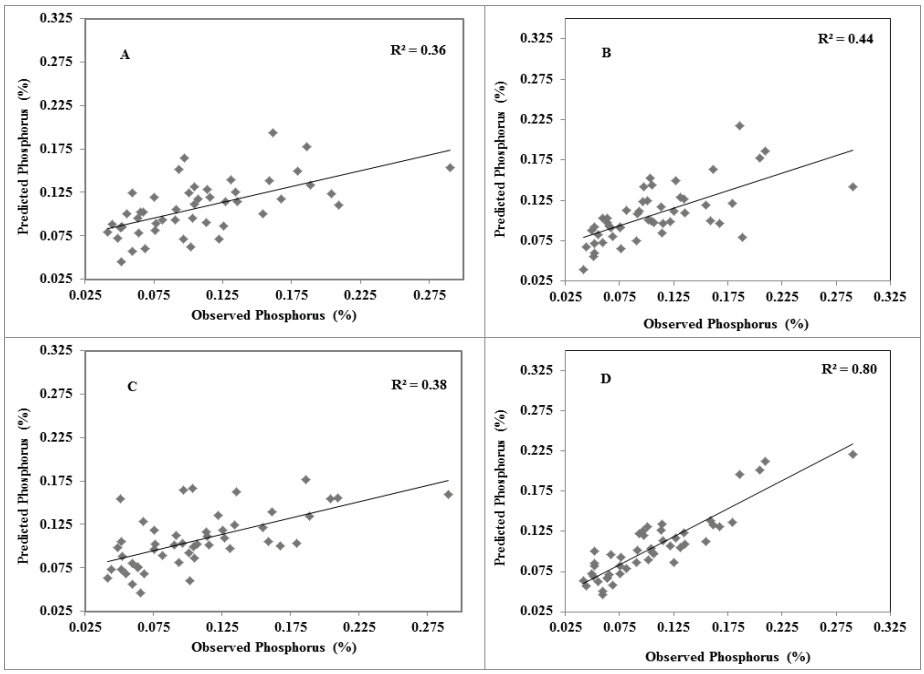

Figure 4.3: A comparison of a conventional and non-linear partial least square regression for foliar P estimation derived through Monte-Carlo leave-one-out cross validation; A= conventional PLSR vs. remote sensing variables only and B= non-linear PLSR vs. remote sensing variables only, C= conventional PLSR vs. remote sensing + environmental variables, D= non-linear PLSR vs. remote sensing + environmental variables.

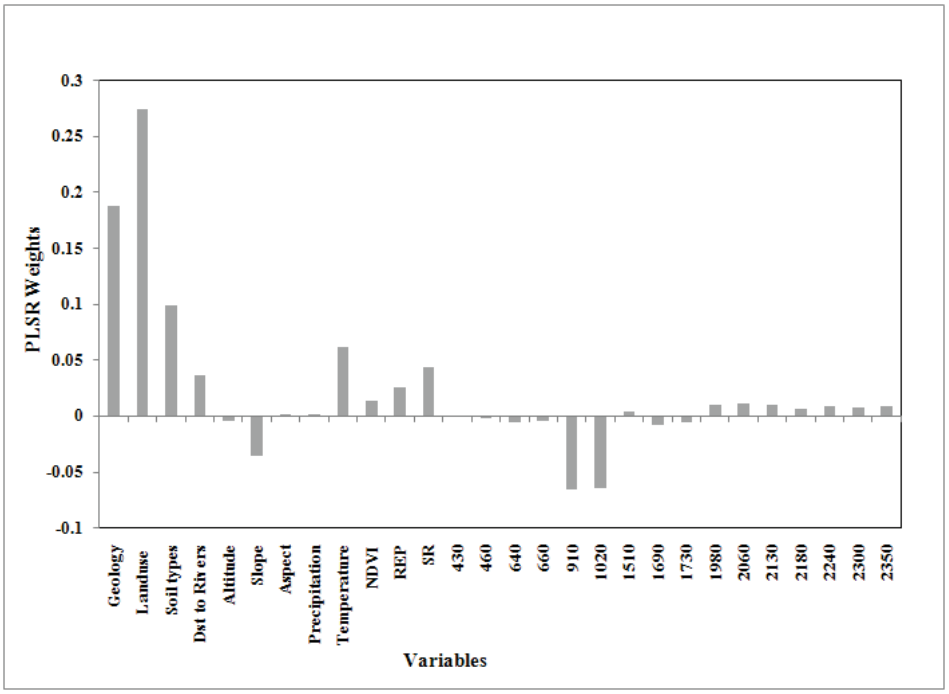

Figure 4.4: PLSR weights indicating contribution of each variable to the foliar N integrated model. Dist=Distance, NDVI=normalized difference vegetation index, REP=red edge position, SR=Simple Ratio.

4.3.2 Integrated modelling using non-linear PLSR for foliar P estimation

For the foliar P estimation, the non-linear RBF-PLSR with integrated remote sensing and environmental variables yielded a higher foliar P estimation accuracy (R^2=0.80, RMSE=0.02: 18.2% of the mean) than the non-linear RBF-PLSR model with remote sensing variables only (R^2=0.44, RMSE=0.04: 45.4% of the mean) (Figure 4.3, Table 4.3). Integrating *in situ* hyperspectral and environmental variables with the non-linear PLSR also yielded higher P estimation accuracy than with the conventional PLSR. The conventional PLSR explained 36% and 38% of the variance of grass P concentration with integrated *in situ* hyperspectral and environmental variables, and with remote sensing variables only, respectively (Figure 4.3, Table 4.3).

Considering the conventional PLSR integrated model, the narrow-band NDVI, SR, REP and several bands in the shortwave infrared showed a positive contribution to the model, while the 910 nm, 970 nm, 990 nm and 1020 nm wavebands showed a negative relationship with foliar P concentration (Figure 4.5). Geology, land use types, soils showed positive PLSR weights, while slope and temperature had negative PLSR weight or contribution to the model (Figure 4.5). The results show that the contribution of the

environmental variables to P estimation is similar to that of N (Figure 4.4 and 4.5). As for N, correlations between foliar P and environmental variables were generally not significant ($p<0.05$), with only weak relationships with slope, precipitation, and land use (Table 4.5). The measured foliar P concentration of the grass sample has a mean value of 0.11% and a coefficient of variation value of 49%, which shows that the variability of P in grass leaves across the study area is high (Table 4.6).

Table 4.5: correlation matrix between P and all environmental variables

	P %	Geo	Land	Soil	DRivers	Alt	Slope	Aspect	Prep	Temp
P %	1.00									
Geo	-0.01	1.00								
Land	-0.21	0.09	1.00							
Soil	0.04	**0.54**	**0.24**	1.00						
DRivers	0.05	**0.75**	**-0.47**	**0.43**	1.00					
Alt	-0.18	**0.60**	**0.71**	**0.69**	0.21	1.00				
Slope	-0.24	**0.53**	**0.32**	**0.31**	0.26	0.48	1.00			
Aspect	0.17	0.16	**-0.29**	-0.10	0.30	-0.06	0.10	1.00		
Prep	-0.24	**0.44**	**0.75**	0.36	0.04	0.81	0.51	-0.11	1.00	
Temp	0.04	**-0.50**	**-0.55**	**-0.41**	-0.16	-0.81	-0.40	-0.02	-0.74	1.00

Geo=geology, Land=land use, Drivers=distance to rivers, Alt=altitude, Precip=precipitation, Temp=temperature, **Bold values** indicates the significance correlation at 95% significance ($p<0.05$)

Table 4.6: Descriptive statistics of the measured foliar N and P concentrations

Nutrients (%)	Minimum	Maximum	Mean	StDev	CV
Nitrogen	0.34	1.06	0.70	0.19	0.26
Phosphorus	0.04	0.29	0.11	0.05	0.49

CV=Coefficient of variance

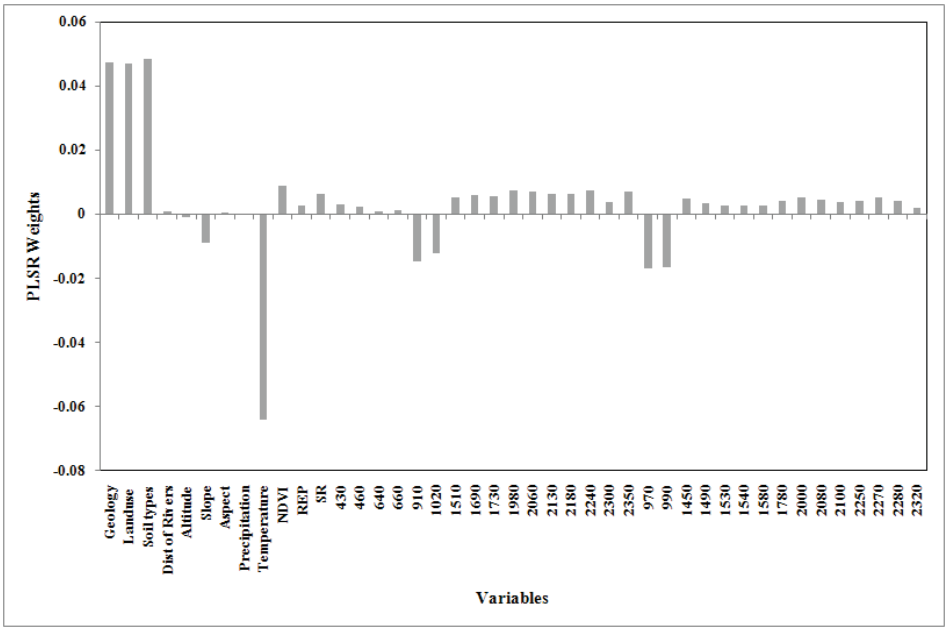

Figure 4.5: PLSR weights indicating contribution of each variable to the foliar P integrated model. Dist=Distance, NDVI=normalized difference vegetation index, REP=red edge position, SR=Simple ratio.

4.4 Discussion

The study was undertaken to address two main objectives; 1) to assess and compare the retrieval accuracy of foliar N and P concentrations when using conventional *vs*. non-linear PLSR, and 2) to test the performance of non-linear PLSR for integrating *in situ* hyperspectral and environmental variables (climatic, edaphic and topographic) for predicting grass N and P concentrations.

4.4.1 Comparing conventional and non-linear PLSR in N and P estimation

For both foliar N and P estimation, the non-linear PLSR performed with a higher accuracy than conventional PLSR. The non-linear PLSR has the mutual advantages of the linear nature of RBF (which is a neural network) and of the power of PLSR to maximize covariance between data sets, while minimizing the variance of the prediction (Walczak and Massart, 1996). Maximizing co-variance between data sets is done through decomposition of the independent variables into uncorrelated latent variables which is important for; (1) reducing the dimensionality of the data (Geladi and Kowalski, 1986; Ehsani et al., 1999; Geladi et al., 1999) and (2) minimizing the over-fitting and multicollinearity (Walczak and Massart, 1996; Huang et al., 2004), to enhance the transferability of models (Crawley, 2006). The inclusion of the

RBF model which is neural network in nature makes the non-linear PLSR to be nonparametric and can be applied without being constrained by the statistical distribution (Atkinson and Tatanall, 1997). This study demonstrated the power of the non-linear RBF-PLSR for estimating foliar N and P coupled with the integrated in situ remote sensing and environmental variables.

4.4.2 Integrated modelling for Foliar N estimation: contribution of hyperspectral variables

The results showed that integrating in situ hyperspectral and environmental variables increases the estimation accuracy of foliar N concentrations, compared to using remote sensing variables alone. Narrow-band SR, REP, and protein absorption features at 910 nm and 1020 nm significantly contributed to the prediction of foliar N concentrations. The red edge has been widely utilized because it is highly correlated to chlorophyll (Cho and Skidmore, 2006; Darvishzadeh et al., 2008b) and it minimizes soil background effects (Zarco-Tejada et al., 2004b). Positive correlation between chlorophyll and N has been reported by Yoder and Pettigrew-Crosby (1995). The results are consistent with other studies focusing on foliar N concentration using in-situ hyperspectral measurements (Gong et al., 2002; Mutanga et al., 2004c; Knox et al., 2010b). Gong et al. (2002) demonstrated the utility of blue and red edge regions for estimating foliar N estimation.

The protein absorption features at 910 nm and 1020 nm contributed to foliar N estimation model as they are influenced through various vibration mechanisms such as C-H stretch, 3rd overtone and N-H stretch (Curran, 1989; Kumar et al., 2001). The visible region of the spectra is characterized by the electron transition while the near and shortwave infrared are characterized by the various bond vibration (Curran, 1989; Kumar et al., 2001). Several studies used these absorption features not only for foliar estimation, but also for biomass and LAI estimations (Cho et al., 2007a; Darvishzadeh et al., 2008b).

This study is consistent with the initial attempts to use remote sensing and environmental or ancillary variables for vegetation mapping to improve accuracy, with variables such as slope, aspect and elevation used as a proxy for temperature and moisture conditions (Hoffer, 1975; Strahler et al., 1978). Such techniques were successfully applied for vegetation mapping in the forest environments (Strahler, 1981; Franklin et al., 1986; Skidmore, 1989).

4.4.3 Integrated modelling for foliar P estimation: contribution of hyperspectral variables

Integrating in situ hyperspectral and environmental variables improved the estimation accuracy for foliar P estimation, as compared to the use of the remote sensing variables alone. The contribution of remote sensing variables in estimating foliar P concentrations was based on several biochemical absorption features, red edge position, narrow-band NDVI and SR. Unlike the foliar N concentration with defined absorption features, the estimation of foliar P using hyperspectral remote sensing does not have specific absorption features defined. The few studies which researched the prediction of foliar P concentrations from spectral data found that most sensitive bands are located in the SWIR (Bogrekci and Lee, 2005; Mutanga and Kumar, 2007; Cho et al., 2010; Ramoelo et al., 2011b). As shown above, these regions are characterized by the various vibration mechanisms imposed by several biochemicals, e.g. O-H, C-C and N-H associated with protein, N, sugar and starch. The limited contribution of the red edge position, narrow-band NDVI and SR was expected since the bands used to calculate these indices are all located in the visible region of the spectrum. This is consistent with the results of Gong et al. (2002) who attempted to use vegetation indices derived from visible bands for estimating foliar phosphorus and generally reported low correlations.

4.4.4 Integrated modelling for foliar N and P: contribution of environmental variables

The positive PLSR weights for environmental variables such as geology, soils, distance to rivers and temperature showed that the distribution of grass nutrients are directly or indirectly linked to climatic, topographic and geologic variables (Ben-Shahar and Coe, 1992; Heitkönig and Owen-Smith, 1998; Pickett et al., 2003; Grant and Scholes, 2006; Skidmore et al., 2011). Geology (which is closely reflected into the soil layer) is a key determinant of grass nutrient concentrations in these savanna ecosystems, i.e. grass nutrient concentrations are linked to soil nutrient contents (Ben-Shahar and Coe, 1992; Pickett et al., 2003; Venter et al., 2003; Skidmore et al., 2010). In our study area, two main geological substrates contribute to the variation of grass nutrients; gabbro support highly nutritious grass (Setaria sphacelata, Digitaria eriantha, Urochloa mosambicensis), while granite support low-nutrient content grass species (e.g. Eragrostis rigidior, Sporobolus spp.) because of the clay content in the soil, which is higher in the gabbro than the granite (Ben-Shahar and Coe, 1992). The granitic substrate has low in situ clay formation potential because it is weather resistant, while gabbro is easily weathered with high in situ clay formation as basalt (Venter et al., 2003). Land use is also highlighted to contribute highly in the foliar N prediction model. Table 4.5 reports a high correlation between land use types and mean annual precipitation due to the rainfall gradient i.e.,

more precipitations fall in the communal areas than Sabi Sands and the KNP areas. Precipitation is the main source of water acts as a vector of particulate and dissolved materials which make it a primary agent of soil heterogeneity and consequence of vegetation or grass responses (Venter et al., 2003). There is also a significant contribution of mean annual temperature (Figure 4.5) in the foliar N concentration prediction which is further associated with the water availability in the soils or plants (Venter et al., 2003). Drier grasses have less photosynthetic activity due to low water content leading to lower foliar biochemical concentrations or grass quality (Prins and van Langevelde, 2008). Slope had a negative contribution to the model which generally implies that steeper slopes have lower grass nutrient contents as result of thinner and coarser-textured soil layers than the lower slopes (e.g. valleys) with relatively higher grass nutrient concentrations (Figure 4.4; 4.5, Table 4.4; 4.5). Minerals and clay particles from the crest and midslope are removed through run-off and deposited in the valley or close to the drainage line areas (Grant et al., 2000). Increased nutrient supply and water availability (through higher water retention capacity) favours highly nutritious grass species in the valley or drainage line areas (Grant et al., 2000). This trend is also confirmed by the positive contribution of distance to rivers variable, i.e. grass species found close to the rivers have high nutrients concentration than areas further away.

The significant correlations between geology, slope, distance to rivers, altitude and temperature highlight the complexity and web of the inter-relationships between geology, topography and grass nutrient concentrations determined by the availability and diversity of minerals as well as moisture in the soil (Table 4.4, 4.5) (Ben-Shahar and Coe, 1992). In this study, this complexity was addressed by applying the integrated modelling approach which accounts to the contribution of environmental variables to estimate foliar biochemicals.

The study demonstrates the importance of integrated modelling for estimating grass quality which is an imperative effort towards effective management and planning of protected and communal savanna ecosystems. Integrating environmental and remote sensing variables does increase accuracy of foliar N and P concentrations estimated, using non-linear PLSR. This integrated modelling approach is an endeavour towards mapping regional estimates of grass N and P concentrations using satellite or spaceborne remote sensing. This ultimately will provide large scale information which farmers, park or land-use managers and planners could utilize for sustainable use of protected and non-protected savanna ecosystems.

4.5 Acknowledgement

We would like to acknowledge the Council for Scientific and Industrial Research, the South African Department of Science and Technology and National Research Foundation (Professional Development Programme), the Wageningen University and the University of Twente for the funding. We also want to thank Dr. Izak Smith, Mrs. Patricia Khosa and Mrs. Thembi Khoza as well as the field guards (Mr Veli Ndlovu, Godfrey and Onica Sithole) from the South African National Parks (SANPARKS) and Mr Mike Glover as well as Jonathan Swart from the private game reserve Sabi Sands. We also would like to thank Mr (s) Laven Naidoo, Thulani Selaule and Russell Main for their field work assistance.

Chapter 5

Grass nitrogen estimation and mapping using remote sensing at regional scale[4]

[4] Chapter is based on: Ramoelo A. Skidmore A.K. Cho M.A. Schlerf M. Mathieu R. Heitkönig I.M.A. Regional estimation of savanna grass nitrogen using red-edge band of the spaceborne RapidEye sensor, *International Journal of Applied Earth Observation and Geo-information (JAG)* (**under review**)

Abstract

The regional mapping of grass nutrients is of interest in the sustainable planning and management of livestock and wildlife grazing. The objective of this study was to estimate and map foliar and canopy Nitrogen (N) at a regional scale using a recent high resolution spaceborne multispectral sensor (i.e. RapidEye) in the Kruger National Park (KNP) and its surrounding areas, South Africa. 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 for estimating foliar chlorophyll and N concentrations has been demonstrated in many previous studies, mostly using field spectroscopy. The utility of the red-edge band of the RapidEye sensor for estimating grass N was investigated in this study. A two-step approach was adopted involving (i) vegetation indices and (ii) the integration of vegetation indices with environmental or ancillary variables using a stepwise multiple linear regression (SMLR) and a non-linear spatial least squares regression (PLSR). To ensure that the estimation of grass N was not compromised by biomass variability, the field work was undertaken during peak productivity. The model involving the simple ratio (SR) index (R805/R710) defined as SR54, altitude and the interaction between SR54 and altitude (SR54*altitude) yielded the highest accuracy for canopy N estimation, while the non-linear PLSR yielded the highest accuracy for foliar N estimation through the integration of remote sensing (SR54) and environmental variables. The spatial pattern of foliar N concentrations corresponded with the soil fertility gradient induced by the geological parent material. The study demonstrated the possibility to map grass nutrients at a regional scale provided there is a spaceborne sensor encompassing the red edge waveband with a high spatial resolution. Regional maps of the grass nutrients could be used for planning and management of the savanna ecosystems by farmers, resource managers and land use planners.

5.1 Introduction

Savanna ecosystems constitute about 32.8% of the land in South Africa (Mucina and Rutherford, 2006). These ecosystems play a crucial role in the rural economy of the country, and worldwide as well (Shackleton et al., 2002; James et al., 2003). Among other things, they provide grazing resources important for livestock production, one of the main sources of income in South African rural areas (Shackleton et al., 2002). The main challenge for savannas is their sensitivity to land degradation due to overgrazing and overstocking (Abel and Blaikie, 1989; Du Toit and Cumming, 1999; Everson and Hatch, 1999). This is at least in part the result of the lack of information about grass conditions hampering proper management (Everson and Hatch, 1999). There is a need for sustainable utilization of the grazing land for viable livestock production, while minimizing land

degradation. Spatial and regional information about grass nutrients is useful to guide farmers towards sustainable management of their grazing land, thus alleviating poverty.

Regional mapping of grass nitrogen (N) provides essential information for sustainable planning and management of livestock and wildlife grazing by livestock farmers, park wardens, or game and resource managers. Grass N concentration is an indicator of grass quality as it is positively correlated to protein content (Clifton et al., 1994; Wang et al., 2004). Protein forms one of the major nutrient requirements for herbivores (Prins and Beekman, 1989; Prins and van Langevelde, 2008). Grass quality is an important parameter affecting the distribution and grazing behaviour of livestock and wildlife (McNaughton, 1990; Ben-Shahar and Coe, 1992; Heitkönig and Owen-Smith, 1998). For example, large herbivores concentrate in highly nutritious areas in southern Africa (Owen-Smith and Danckwerts, 1997; Grant and Scholes, 2006; Treydte et al., 2007) and herbivore diversity increases with increasing soil fertility levels (Olff et al., 2002). Soil fertility levels generally correlate with grass N concentrations (Ben-Shahar and Coe, 1992; Olff et al., 2002). Therefore, grass N concentrations could be used as a proxy for soil fertility levels.

Remote sensing techniques have been developed over the past decades to extract information about biophysical and biochemical parameters of vegetation such as leaf area index, chlorophyll, P, fibre, lignin, and N (LaCapra et al., 1996; Asner et al., 1998; Beeri et al., 2007; Darvishzadeh et al., 2008b; Majeke et al., 2008; Schlerf et al., 2010; Main et al., 2011; Ramoelo et al., 2011b). The conventional approach relates a specific vegetation parameter to vegetation indices derived from remote sensing data using a variety of statistical regression techniques (Hansen and Schjoerring, 2003; Haboudane et al., 2004; Darvishzadeh et al., 2008b; le Maire et al., 2008; Starks et al., 2008). For estimating foliar biochemicals (e.g. N), traditional broadband indices such as normalized difference vegetation index (NDVI) (Rouse et al., 1974), soil line concept (SLC), simple ratio (SR) (Baret and Guyot, 1991), and soil-adjusted vegetation index (SAVI) (Huete, 1988) are not conducive. These broadband vegetation indices saturate at high canopy cover (Tucker, 1977; Mutanga and Skidmore, 2004b) and are insensitive to subtle changes in the foliar N concentration.

The more recent success in estimating foliar N and chlorophyll concentrations has been possible due to the development of hyperspectral remote sensing. Studies using hyperspectral remote sensing have highlighted the utility of red-edge bands for estimating foliar N and chlorophyll concentrations (Huang et al., 2004; Cho and Skidmore, 2006; Darvishzadeh et al., 2008b). The red-edge is the region of abrupt change in leaf reflectance between 680 and 780

nm, mainly influenced by the concerted effect of spectral absorption in the red wavelengths and scattering in the near infrared region (Gates et al., 1965; Horler et al., 1983; Clevers et al., 2002). Cho and Skidmore (2006) developed a technique to compute the red-edge position (REP), which is highly sensitive to foliar chlorophyll. REP is known to be insensitive to background effects (Elvidge and Chen, 1995) and is highly correlated to foliar N (Cho and Skidmore, 2006), as chlorophyll is positively correlated to N (Yoder and Pettigrew-Crosby, 1995; Hansen and Schjoerring, 2003; Haboudane et al., 2004). Vegetation indices computed from red-edge bands, also known as narrow-band indices, have provided improved estimates of foliar N compared to conventional broad-band indices derived from red (680 nm) and near infrared (800 nm) (Hansen and Schjoerring, 2003; Mutanga and Skidmore, 2007).

Other successful hyperspectral techniques in foliar N estimation involve the use of N and protein absorption features in the visible (VIS), near infrared (NIR) and shortwave infrared (SWIR) (Huang et al., 2004; Knox et al., 2010b; Schlerf et al., 2010; Skidmore et al., 2010). Several studies argued that the use of selected absorption features surpasses the use of the full spectrum for foliar biochemical and biophysical estimation (Cho et al., 2007a; Darvishzadeh et al., 2008b), because it reduces the chance of using redundant data. The drawback to using this approach for regional estimation of foliar biochemical concentrations is that there are limited satellite sensors which sample light using the full spectral region with narrow bands adequately resolving these absorption features. Satellite sensors with strategically placed spectral bands in the red-edge region are likely to provide successful estimates of biochemical concentrations, and more specifically N. However, as these sensors are scarce, foliar N concentration is seldom mapped on a regional scale. For example, conventional multispectral satellite sensors such as SPOT, Landsat, and ASTER lack specific spectral bands in the red-edge region and their spatial resolutions are relatively coarse. The MERIS sensor has a standard band setting allowing the computation and approximation of the red-edge position (Clevers et al., 2002), but the spatial resolution is too coarse, especially for savannas, which are a complex and heterogeneous mosaic of grass and trees. The emergence of multispectral sensors such as WorldView-2 (USA), SumbandilaSAT (South Africa) and RapidEye (Germany) with red-edge bands at high spatial resolution (i.e. 6.5 m) could provide an opportunity for rangeland resource quality assessment at a regional level. There is a need for the development of specific vegetation indices that could be used successfully with these sensors. In this study several broad-band and hyperspectral vegetation indices were modified to incorporate the red-edge band of RapidEye to estimate grass N concentration at a regional scale.

A challenge when using remote sensing to estimate foliar biochemical concentrations is associated with the difficulty disentangling the signals for biomass and foliar biochemical concentrations, especially N (i.e. the interaction effects between N and biomass) (Skidmore et al., 2010). These effects can be minimized during peak productivity when the grass spectra have the highest absorption in the red region and scattering in the near infrared region (Plummer, 1988a, 1988b; Skidmore et al., 2010). During this period, the scattering and absorption processes continue to increase due to biomass production, as captured by indices such as normalized difference vegetation index (NDVI), and the relationship between biomass and NDVI asymptotically saturates (Tucker, 1977; Mutanga and Skidmore, 2004b). At a certain critical biomass point (e.g. 3000 g/m2) reached at peak productivity, the vegetation indices are unable to estimate further increase in biomass (Mutanga and Skidmore, 2004b). That is when foliar N can be estimated with minimal effect from the N-biomass interaction problem.

In addition, a few studies have highlighted the need to integrate environmental or ancillary and remote sensing variables to estimate foliar biochemical concentrations at a regional scale (Cho et al., 2009; Cho et al., 2010; Knox et al., 2011; Ramoelo et al., 2011a; Ramoelo et al., under review), which could be a crucial step towards improving regional estimation and mapping. A combination of factors such as edaphic (geology and soils), topographic (slope, aspect, and altitude), and climatic (precipitation and temperature) factors are known to influence the distribution of foliar biochemical concentrations in a very complex way (Ben-Shahar and Coe, 1992; Mutanga et al., 2004a; Ferwerda et al., 2006; Skidmore et al., 2011). Ramoelo et al. (2011a) showed that geology, slope, temperature, and land use types were the main contributing environmental variables when modeling foliar N in combination with in situ hyperspectral remote sensing variables. However, where environmental data sets are readily available at a regional scale, their resolution is relatively coarse rendering them unsuitable as sole input in the estimation of foliar biochemical concentrations. The use of remote sensing could address this issue of resolution and scale, for instance regional maps could be derived at a resolution of 5 to 10 m based on data from the newly developed spaceborne sensors. The assumption is that a modeling approach which integrates remote sensing and environmental variables potentially yields a higher foliar N estimation accuracy than approaches using either remote sensing or environmental variables (Cho et al., 2009; Cho et al., 2010; Knox et al., 2011; Ramoelo et al., 2011a; Ramoelo et al., under review). The objectives of this study were twofold; (1) to investigate the utility of the red-edge band of the RapidEye sensor for estimating grass N concentrations using various vegetation indices derived from the RapidEye data, and to determine which vegetation index correlates highly with grass foliar as well as canopy N and (2) to integrate this

vegetation index with the environmental variables to estimate and map grass foliar and canopy N at a regional scale.

5.2 Study area

The study area is located in the north-eastern part of South Africa (Figure 5.1) and covers a total area of approximately 5000 km². The area is referred to as the Lowveld landscape, which is a low lying area extending from the foot slopes of the Drakensberg Great Escarpment to the west to the Mozambique coastal plain to the east (Venter et al., 2003). Protected areas such as the privately owned Sabi Sands Game Reserve (SGR) and the state - owned Kruger National Park (KNP), as well as the communal lands in Bushbuckridge form the main land tenures. The main vegetation types are "Tshokwane-Hlane basalt lowveld", "granite lowveld", "gabbro grassy bushveld", and "Delagoa lowveld" (Mucina and Rutherford, 2006). The Tshokwane-Hlane basalt lowveld is characterized by open tree savannas with trees such as Sclerocarya birrea, Acacia nigrescens, Acacia gerrardii, Peltophorum africanum, Dichrostachys cinerea, and common grass species such as Bothriochloa radicans, Digitaria eriantha, Cenchrus ciliaris, and Urochloa mosambicensis.

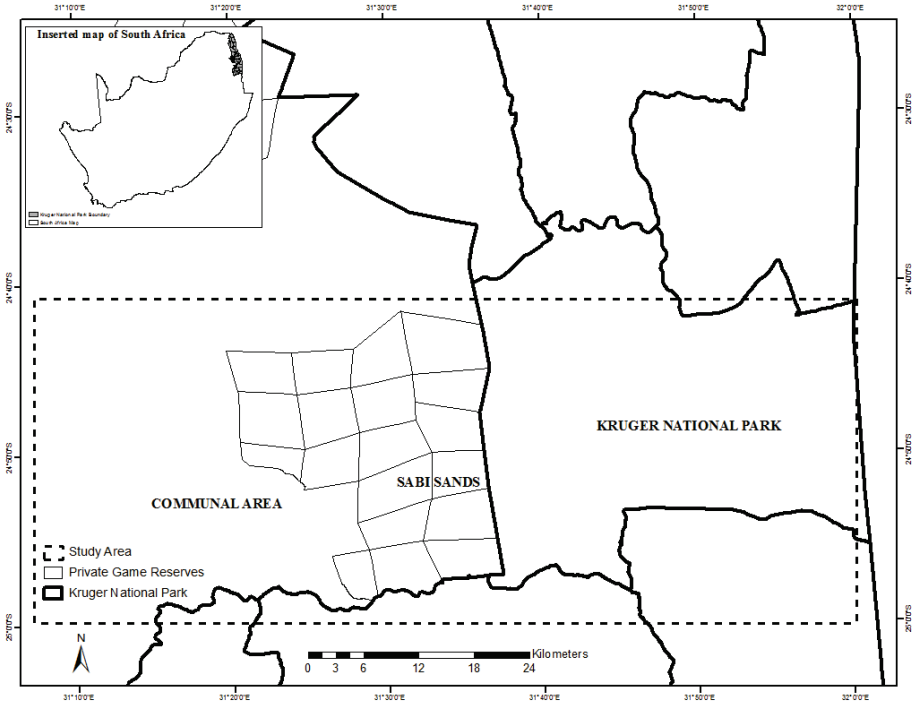

Figure 5.1: Map of the study area

This vegetation types occur in the highly fertile black, brown or red clayey soils derived from the basalt substrate. The granite lowveld comprises dense thickets dominated by trees such as several Combretum species, Dichrostachys cinerea, Grewia bicolor, and Terminalia sericea with dominant grass species being Pogonarthria squarrosa, Tracholeona monachne, and Eragrostis rigidior. The granite-derived soils are sandy in the uplands and clayey in the bottomlands, and are low in fertility compared to the basalt-derived soils. Gabbro grassy bushveld constitutes an open savanna with dense grass cover. Dominant tree species are Acacia nigrescens, Sclerocarya birrea, Bolusanthus speciosus, and Ziziphus mucronata, while common grass species are Chloris virgata, Setaria species, Themenda triandra, Bothriochloa radicans, Panicum maximum, Urochloa mosambicensis, and Eragrostis superba. Soils in this vegetation type are fertile dark vertic with 20 to 50% clay derived from the gabbro geological type (Mucina and Rutherford, 2006). The Delagoa lowveld vegetation type is characterized by dense thickets with common tree species such as Acacia welwitschii, Dichrostachys cinerea, Euclea divinorum, and Grewia bicolor and grass species such as Chloris virgata, Aristida congesta, Panicum colaratum, and Sporobolus species. This vegetation type occurs in shale and lesser sandstone layers interspersed by sheets and dykes of Jurassic dolerite (Mucina and Rutherford, 2006). The soils are rich in sodium, but the fertility is lower than in the basaltic-derived soils. There is an evident precipitation gradient from the western part (800 mm/year) to the eastern part (580 mm/year) of the study area (Venter et al., 2003). The annual mean temperature is about 22oC. Geology as mentioned above includes granite and gneiss with local intrusions of gabbro in the west and basalt as well as shale in the eastern part towards Mozambique (Venter et al., 2003). The contrasting geological substrates (and associated soil types) together with the precipitation influence, clearly define the patterns and gradients in soil moisture and nutrients. The topography is mostly undulating in the granitic sites and flat in the basalt areas, with an average height of 450 m. Rangelands in the protected areas are grazed by wild herbivores such as impala (Aepyceros melampus), zebra (Equus burchelli), wildebeest (Connochaetes taurinus), and buffalo (Syncerus caffer), while the communal rangelands support the grazing of cattle (Bos taurus), goats (Capra hircus), and sheep (Ovis aries), thus determining various grazing or land use intensities.

5.3 Data Collection

5.3.1 Field data collection

The field data were collected using a road sampling technique since deep penetration into the savanna landscape was limited by management and logistical restrictions. Field work was undertaken in April 2010, the same month the satellite imagery was collected. The areas along the main roads

covering the study area were purposively selected for the field sampling based on their underlying geological strata, both in the protected and in the communal areas. Buffers of 300 m were created on both sides of these roads using ArcGIS software (ESRI, USA). Within the buffer polygons random sample points were generated using the ArcGIS add-on called Hawth tools. All points directly on the road or on the bare areas next to the road were rejected because of the lack of grass. The plots were randomly located in areas with homogeneous grass to avoid the effect of trees on the grass signal. Each sample point (N=51) was treated as a plot of 20 m x 20 m, to account for a geometric accuracy of up to one pixel (i.e. 5 m) on the RapidEye image. In each plot, 2 subplots of 50 cm x 50 cm were used to collect information about the dominant species, the percentage cover of photosynthetic and non-photosynthetic vegetation as well as bare soil. The grass in each subplot was clipped and weighed to determine the wet biomass. The grass samples were then dried at 80ºC for 24 hours and weighed again to establish the dry biomass. Grass biomass was expressed in weight per unit area (i.e. g/m2). The biomass data were acquired to determine any interaction effects between biomass and foliar N. The field work was undertaken during peak productivity to minimize these interaction effects (Plummer, 1988b, 1988a; Skidmore et al., 2010), as discussed in the Introduction. The grass samples were dried to retrieve foliar N concentrations.

5.3.2 Chemical analysis

The dried grass samples were taken to South Africa's Agricultural Research Council Institute for Tropical and Subtropical Crops (ARC-ITSC) in Nelspruit for chemical analysis. Firstly, the acid digestion technique was used, where sulphuric acid aided the foliar N retrieval (Giron, 1973; Grasshoff et al., 1983; Mutanga et al., 2004a). Secondly, the colorimetric method by auto analyser was used to measure the foliar N (Technicon Industrial Method 329-74 W; Technicon Industrial Systems, Farrytown, New York). An emerald-green colour was formed by the reaction between ammonia, sodium salicylate, sodium nitroprusside, and sodium hypochlorite. The ammonia-salicylate complex was read at 640 nm. These two extraction methods were already successfully used for grass foliar N by Mutanga et al. (2004), Ramoelo et al. (2011b) and Ramoelo et al. (under review). Chemically analyzed N was henceforth referred to observed N.

5.3.3 Image acquisition and atmospheric corrections

The mission to collect RapidEye images was tasked in April 2010. The RapidEye sensor has a multispectral push broom imager with a spatial resolution of 6.25 m and captures data in the spectral bands: 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 RapidEye Ortho product

(Level 3A) was provided with radiometric, sensor, and geometric correction applied using the digital terrain elevation data (DTED) level 1 Shuttle Radar Terrain Mission (SRTM). The orthorectification accuracy of 1 or less pixel was achieved (RapidEye, 2010). The RapidEye Ortho product was delivered resampled to a 5m x 5m spatial resolution. To retrieve surface reflectance atmospheric correction was executed using the atmospheric and topographic correction software (ATCOR 2) implemented in the IDL Virtual Machine (Richter, 2011). ATCOR 2 models reflectance for flat surfaces, which was considered sufficient since the study area was not characterized by very rugged terrain. The advantage of ATCOR 2 is that it was developed specifically for satellite remote sensing data and includes a large database of atmospheric correction functions (look-up-tables computed with the Modtran® 5 radiative transfer code) covering a wide range of weather conditions, sun angles, and ground elevations (Richter, 2011). The Modtran® standard aerosols for "rural" were selected to compute the aerosol type, and "visibility" was computed according to Richter (2011). RapidEye metadata were used to obtain additional important information for reflectance retrieval such as satellite and solar zenith angle, satellite and solar azimuth angle, as well as relative azimuth angle. The workflow for implementing ATCOR for atmospheric correction in any terrain is well outlined in Richter (2011).

5.3.4 Environmental or ancillary variables

Several studies showed that climate, topography, and geologic substrate influence the distribution of primary environmental regimes such as moisture and nutrients in soils or plants; for details see the review by Skidmore et al. (2011), as well as Pickett et al. (2003), Venter et al. (2003) and Mutanga et al. (2004). Several environmental variables influence the distribution of grass N at different scales; these include precipitation, temperature, land use, geology, soils, distance to rivers, altitude, slope, and aspect (Table 5.1). Mean annual precipitation (MAP) and temperature (MAT) were acquired from the World Climate database (WorldClim) (www.WorldClim.com). This climatic database has been widely used for biodiversity and ecological applications (Hijmans et al., 2001) and climatic stations are spread across South Africa (Hijmans et al., 2005; Adams and Church, 2007; Saad et al., 2007). The freely available SRTM 4.1 Digital Elevation Model (DEM) with its relatively high spatial resolution of 90 m (Javis et al., 2008) was used. To make it more reliable, Javis et al. (2008) further improved the DEM by filling in the holes identified. Slope and aspect were derived from the DEM using ArcGIS 10x. The river layer was sourced from the South African National Botanical Institute (SANBI)'s Beta version of vegetation data sets (Mucina and Rutherford, 2006). The 'distance to river' variable was computed using the Spatial Analyst Tool embedded in ArcGIS 10x, where the river layer and the sample plot locations (GPS points) formed the inputs. A soil layer was acquired from the soil and terrain database of Southern Africa (SOTERSAF)

(Dijkshoorn, 2003) (Table 5.1). This soil map has been used for the Land Degradation Assessment in Drylands project, for which South Africa is one of the partners (Dijkshoorn et al., 2008). The land use types were derived from the boundary layers of KNP, Sabi Sands Game Reserve and the communal areas, acquired from KNP's Geographic Information System (GIS) and remote sensing laboratory.

Table 5.1: Environmental variables used in this study

Environmental Data	Type	Source	Resolution
Geology	Cate	Council for Geoscience	1:1000000
Soil	Cate	SOTERSAF database	1:1000000
Precipitation	Conti	http://www.worldclim.com/	1 km
Temperature	Conti	http://www.worldclim.com/	1 km
Land use types	Cate	KNP	Vector layer
Altitude (DEM)	Conti	SRTM	90 m
Slope	Conti	Derived from DEM	90 m
Aspect	Conti	Derived from DEM	90 m
Distance from rivers	Conti	SANBI GIS data	1:1000000

Cate=Categorical, Conti=Continuous, DEM= digital elevation model, CSIR=Council for Scientific and Industrial Research, SANBI=South African National Botanical Institute, SOTER=Soil and Terrain of Southern Africa database, SRTM=Shuttle Radar Topography Mission (http://srtm.csi.cgiar.org), KNP=Kruger National Park GIS datasets

5.4 Data Analysis

The reflectance data corresponding to each field plot were extracted from the image in order to perform the statistical analysis. The vegetation indices listed in Table 5.2 were computed from the extracted reflectance data. The new or modified vegetation indices were mainly developed to benefit from the inclusion of the red-edge band in the RapidEye spectral configuration. In Table 5.2, simple ratios (SRs) are written as SR53, 54, and 43, just as the normalized difference vegetation indices (NDVIs) are written as NDVI54 and so on, denoting the band combinations used. In some cases, such as plant pigment ratio (PPR), transformed chlorophyll absorption index (TCARI), and modified chlorophyll absorption index (MCARI), the indices were given new RapidEye compatible bands less than 60 nm from the original indices, to ensure that the sensitivity of the specific region of the spectrum was maintained. All the indices selected were sensitive to leaf and canopy chlorophyll (Table 5.2). For statistical analysis, the foliar N concentration was multiplied by the percentage cover of photosynthetic vegetation (PV) to derive a unit-less canopy integrated nitrogen content, denoted as N*PV, henceforth called canopy nitrogen (Wessman, 1992; He and Mui, 2010).

5.4.1 Univariate and Multivariate analysis

The univariate analysis involved bootstrapping the linear regression between biochemicals (foliar and canopy N) and vegetation indices. Subsequently, the results from that analysis were used to select the vegetation index, based on a high coefficient of determination (R2) and a low root mean square error (RMSE) (Bunke and Droge, 1984; Efron and Tibshirani, 1997; Fox, 2002; Fox and Weisberg, 2010). The multivariate analysis was undertaken using an integrated modeling approach. The vegetation index (with a high estimation accuracy resulting from bootstrapping statistics combined with environmental variables) was used to predict N concentrations. The first multivariate analysis was performed using a combination of principal component analysis (PCA) and stepwise multiple linear regression (SMLR) (Çamdevýren et al., 2005), denoted as SMLR+PCA. The aim of the PCA was to decompose the independent variables into uncorrelated components. The advantage of the PCA was that it reduced multicollinearity and overfitting (Çamdevýren et al., 2005; Jain et al., 2007). In this approach, the initial step was to run the PCA on the independent variables, i.e. the vegetation index and environmental variables. The second step was to run a forward stepwise regression to see which principal components (PC) significantly contributed to the N prediction model. The stepwise model was selected based on the lowest Akaike Information Criterion (AIC) (Sakamoto et al., 1986; An and Gu, 1989). The second multivariate analysis was performed using SMLR based on SR54 and environmental parameters (SMLR+Raw), where the model for predicting foliar N with high accuracy was selected using AIC, similar to the PCA+SMLR method. In this case the original data, i.e. highest performing vegetation index and environmental variables, were used for predicting N. Thirdly, the interaction effects between the selected variables for predicting foliar and canopy N at a SMLR+Raw stage were also tested using SMLR, and denoted SMLR+Raw+Int. Where the interaction effect between the significant variables selected according to lowest AIC improved the estimation accuracy for foliar and canopy N, this was reported, otherwise it was not reported. The final multivariate analysis was based on the non-linear partial least square regression (PLSR), and is known as PLSR with radial basis function (RBF-PLSR). This non-linear PLSR was found to achieve a higher foliar N estimation accuracy than the conventional PLSR (Ramoelo et al., under review). This was mainly attributed to the fact that the non-linear PLSR has combined capabilities of the conventional PLSR and artificial neural network, i.e. maximize covariance between data sets and non-linear model fitting (Walczak and Massart, 1996). The initial stage of applying this technique was to standardize the data sets to 0 and 1 range (Knox et al., 2011; Ramoelo et al., under review). Sigma values were specified in order to compute the activation matrix using the radial basis function. The activation matrix was then used in combination with PLSR to predict foliar N concentrations, where the number of uncorrelated latent variables or factors was specified.

Table 5.2: Vegetation indices used in this study

Index	Conventional Formulae	Modified Formulae	Reference
SR52	R_{NIR}/R_{RED}	$R_{805}/R_{557.5}$	(Jordan, 1969)
SR54		R_{805}/R_{710}	
SR43		$R_{710}/R_{657.5}$	
NDVI52	$(R_{NIR}-R_{RED})/(R_{NIR}+R_{RED})$	$(R_{805}-R_{555})/(R_{805}+R_{555})$	(Gitelson et al., 1996)
NDVI53		$(R_{805}-R_{657.5})/(R_{805}+R_{657.5})$	(Rouse et al., 1974)
NDVI54		$(R_{805}-R_{710})/(R_{805}+R_{710})$	
NDVI43		$(R_{710}-R_{657.5})/(R_{710}+R_{657.5})$	
SAVI	$((1+L)*R_{NIR}-R_{RED})/((R_{NIR}+R_{RED})+L)$	$((1+0.2)*R_{805}-R_{710})/((R_{805}+R_{710})+0.2)$	(Huete, 1988)
SAVI1		$((1+0.2)*R_{805}-R_{657.5})/((R_{805}+R_{657.5})+0.2)$	
OSAVI	$(1+0.16)*(R_{800}-R_{670})/(R_{800}+R_{670}+0.16)$	$(1+0.16)*(R_{805}-R_{710})/(R_{805}+R_{710}+0.16)$	(Rondeaux et al., 1996)
OSAVI2	$(1+0.16)*(R_{750}-R_{705})/(R_{750}+R_{705}+0.16)$	$(1+0.16)*(R_{805}-R_{657.5})/(R_{805}+R_{657.5}+0.16)$	(Wu et al., 2008)
MSAVI	$0.5*(2*R_{800}+1-SQRT((2*R_{800}+1)^2-8(R_{800}-R_{670})))$	$0.5*(2*R_{800}+1-SQRT((2*R_{800}+1)^2-8(R_{800}-R_{670})))$	(Qi et al., 1994)
TVI	$0.5*(120*(R_{750}-R_{550})-200*(R_{670}-R_{550}))$	$0.5*(120*(R_{710}-R_{555})-200*(R_{657.5}-R_{555}))$	(Broge and Leblanc, 2000)
RDVI	$(R_{800}-R_{670})/(SQRT(R_{800}+R_{670})$	$(R_{805}-R_{657.5})/(SQRT(R_{805}+R_{657.5}))$	(Roujean and Breon, 1995)
RDVI1		$(R_{805}-R_{710})/(SQRT(R_{805}+R_{657.5}))$	
MCARI	$((R_{700}-R_{670})-0.2*(R_{700}-R_{550}))*(R_{700}/R_{670})$	$((R_{710}-R_{657.5})-0.2*(R_{710}-R_{555}))*(R_{710}/R_{657.5})$	(Daughtry et al., 2000)
MTCI	$(R_{754}-R_{709})/(R_{709}-R_{681})$	$(R_{800}-R_{710})/(R_{710}-R_{657.5})$	(Dash and Curran, 2004)
PPR	$(R_{550}-R_{450})/(R_{550}+R_{450})$	$(R_{555}-R_{475})/(R_{555}+R_{475})$	(Metternicht, 2003)
NRI	$(R_{570}-R_{670})/(R_{570}+R_{670})$	$(R_{555}-R_{657.5})/(R_{555}+R_{657.5})$	(Schleicher et al., 2001)
SIPI	$(R_{800}-R_{445})/(R_{800}-R_{680})$	$(R_{805}-R_{475})/(R_{805}-R_{657.5})$	(Peñuelas et al., 1995)
SIPI1		$(R_{710}-R_{475})/(R_{710}-R_{657.5})$	
GI	(R_{554}/R_{677})	$(R_{555}/R_{657.5})$	(Smith et al., 1995)
EVI	$2.5*(R_{800}-R_{670})/R_{800}+(6*R_{670})-(7.5*R_{475})+1))$	$2.5*(R_{805}-R_{657.5})/R_{805}+(6*R_{657.5})-(7.5*R_{475})+1))$	(Huete et al., 1997)
TCARI	$3*((R_{700}-R_{670})-0.2*(R_{700}-R_{550})*(R_{700}/R_{670}))$	$3*((R_{710}-R_{657.5})-0.2*(R_{710}-R_{555})*(R_{710}/R_{657.5}))$	(Haboudane et al., 2002)

SR=Simple Ratio, NDVI=Normalized Difference Vegetation Index, SAVI=Soil Adjusted Vegetation Index, *L*=Soil Correction Factor, OSAVI=Optimized SAVI, MSAVI=Modified SAVI, TVI=Triangular Vegetation Index, RDVI=Renormalized Difference Vegetation Index, MCARI=Modified Chlorophyll Absorption Ratio Index, MTCI=MERIS Terrestrial Chlorophyll Index, PPR=Plant Pigment Ratio, NRI=Nitrogen Reflectance Index, SIPI=Structure Insensitive Pigment Index, GI=Greenness Index, EVI=Enhanced Vegetation Index, TCARI=Transformed Chlorophyll Absorption Ratio

5.4.2 Validation

Validation was performed using a bootstrapping technique because of the small sample sizes involved (Bunke and Droge, 1984; Efron and Tibshirani, 1997). Bootstrapping is an unbiased way to validate models as it has an iteration component. It samples the data a number of times, which makes it a more robust way of validating models, as well as extremely efficient when

only few samples are collected. In this study we used 1000 iterations to ensure that the bias was highly reduced. The highly accurate bootstrapped model was inverted and applied to the RapidEye image to map the predicted foliar and canopy N concentrations of the grass canopies. The validation of the non-linear PLSR was based on a Monte-Carlo cross validation, since bootstrapping was not yet incorporated in the TOMCAT software (Walczak and Massart, 1996).

5.4.3 Descriptive and exploratory analysis

One-way analysis of variance (ANOVA) was computed to test if there was any significant difference between foliar N and, firstly, geology and, secondly, soils. The Spearman's rank correlation was used to quantify the relationship between remote sensing and environmental variables, since it can be applied to both categorical and continuous data (Lehman, 1998), to quantify relationships between remote sensing and environmental variable. The descriptive statistics i.e. mean, minimum and maximum as well as standard deviation values of N and N*PV were computed using the R programming language.

5.5 Results

5.5.1 Determining the vegetation index with high N estimation accuracy

The SR54 computed with the red-edge band yielded the highest accuracy for predicting both foliar and canopy N; surpassing the results of the conventional simple ratio (i.e. SR53) (Table 5.3, Figures 5.2 and 5.3). At foliar level, the bootstrapped model resulted in R2=0.23 and RMSE=0.150%, while at canopy level, the bootstrapped model resulted in R2=0.45 and RMSE=13.506 (unit-less). Of the twenty four indices used to estimate foliar and canopy N, the inclusion of the newly embedded red-edge band in the RapidEye data improved the results especially for the top five indices, i.e. SR54, NDVI54, SAVI, OSAVI, and SIPI1 for canopy N, and SR54, NDVI54, OSAVI, SAVI, and MTCI for foliar N concentrations (Table 5.3, Figures 5.2 and 5.3). Generally, there are five indices that could be directly modified to make use of the red-edge band rather than relying on the conventional versions using red and NIR bands, namely SR, NDVI, SAVI, OSAVI, and SIPI. The least performing indices were TVI, TCARI, and MCARI with RMSEs of between 17.364 and 18.101 for canopy N and between 0.170 and 0.171% for foliar N. The variance in canopy N was explained more clearly by the vegetation indices than the variance in foliar N was, with the R2 increasing from 0.23 for foliar to 0.45 for canopy N. A similar pattern was evident in the estimation accuracy measured according to RMSE (Table 5.3).

5.5.2 Integrated modeling for grass N prediction

Integrating vegetation indices and environmental variables for estimating canopy N using SMLR (SMLR+Raw+int) yielded a significantly higher estimation accuracy (bootstrapped: R2=0.64, RMSE=11%; 17% of the mean), than the model using SR54, altitude, and SR54*altitude (Table 5.4). The non-linear PLSR (RBF-PLSR) was the second highest performer concerning the accuracy of estimating canopy N (bootstrapped: R2=0.61, RMSE=11%), after SMLR, with interaction effects from SR54 and altitude (Table 5.4). As shown in Table 5.5, altitude is significantly correlated with other environmental variables such as geology, precipitation, temperature, slope, aspect, and land use. It is evident that altitude in this study is a proxy for various other environmental variables. The last technique tested was principal component analysis and regression (SMLR+PCA), which resulted in a lower canopy N estimation accuracy (bootstrapped: R2=0.56, RMSE=12.33; 19% of the mean) than the above two techniques, with principal components (PC) 1, 3, and 9 selected (Table 5.4).

Table 5.3: Ranking by bootstrapped root mean square error for various vegetation indices in predicting (a) Canopy Nitrogen and (b) foliar Nitrogen

Indices (a)	RMSE	Indices (b)	RMSE (%)
SR54	13.506	SR54	0.150
NDVI54	13.659	NDVI54	0.151
SAVI	13.662	OSAVI	0.151
OSAVI	13.667	SAVI	0.151
SIPI1	14.409	MTCI	0.153
MTCI	14.714	SIPI1	0.153
SAVI1	15.080	GI	0.158
NDVI53	15.088	NRI	0.158
OSAVI2	15.105	NDVI53	0.158
MSAVI	15.125	OSAVI2	0.158
NRI	15.158	SAVI1	0.158
SR53	15.204	MSAVI	0.159
GI	15.232	SR53	0.159
RDVI2	15.308	RDVI2	0.160
SIPI	16.143	EVI	0.161
NDVI43	16.227	SIPI	0.163
EVI	16.232	NDVI43	0.163
NDVI52	16.309	NDVI52	0.164
SR43	16.361	SR43	0.164
RDVI	16.364	RDVI	0.165
PPR	16.451	PPR	0.166
TVI	17.661	TVI	0.170
TCARI	18.089	MCARI	0.171
MCARI	18.101	TCARI	0.171

SR=Simple Ratio, NDVI=Normalized Difference Vegetation Index, SAVI=Soil Adjusted Vegetation Index, L=Soil Correction Factor, OSAVI=Optimized SAVI, MSAVI=Modified SAVI, TVI=Triangular Vegetation Index, RDVI=Renormalized Difference Vegetation Index, MCARI=Modified Chlorophyll Absorption Ratio Index, MTCI=MERIS Terrestrial Chlorophyll Index, PPR=Plant Pigment Ratio, NRI=Nitrogen Reflectance Index, SIPI=Structure Insensitive Pigment Index, GI=Greenness Index, EVI=Enhanced Vegetation Index, TCARI=Transformed Chlorophyll Absorption Ratio.

Table 5.4: The performance of various multivariate techniques used validated by bootstrapping

	R^2	RMSE	RMSE(% of Mean)	*p*-value	Selected variables
Canopy N					
SMLR+PCA	0.56	12.33	16.50	<0.05	*PC1, PC3, PC9
SMLR+Raw	0.59	11.60	15.52	<0.05	SR54, Altitude
RBF-PLSR	0.61	11.00	14.72	<0.05	4 factors, and 0.7 Sigma value
SMLR+Raw+ Int.	0.64	11.00	14.72	<0.05	SR54, Altitude, SR54*Altitude
Foliar N		RMSE (%)			
SMLR+PCA	0.45	0.14	16.66	<0.05	*PC1-4, PC9, PC10
SMLR+Raw	0.44	0.13	15.47	<0.05	SR54, Altitude, Aspect, Dist
RBF-PLSR	0.48	0.12	14.28	<0.05	5 Factors, and 1 Sigma value

SMLR=Stepwise linear regression, PCA=principal component analysis, RBF-PLSR=partial least square regression with radial basis function, SR54=simple ratio, Dist=distance to rivers, Raw=SR54 and environmental variables used as they are. Int.=indicates a model with the interaction effects of the variables significantly selected in SMLR+Raw. *p* value at the 95% confidence level ($p<0.05$). *Using PC loadings, PC1 = (aspect, precipitation, altitude), PC2 = (aspect, precipitation, soils, altitude), PC3 = (aspect, precipitation, soils, altitude, geology), PC4 = (precipitation, soils, altitude), PC9= (SR54), PC10 = (temperature and land use).

Table 5.5: Spearman *p* correlation matrix between N and various environmental or ancillary variables

	N*PV	SR54	Geo	Soil	Prec	Tem	Asp	Alt	Slo	Lan	Dist
N*PV	1	**0.62**	-0.20	-0.09	**-0.37**	**0.42**	**0.20**	**0.56**	**-0.23**	**-0.50**	**-0.01**
SR54		1	-0.11	0.12	-0.06	0.12	0.06	-0.23	-0.15	-0.19	-0.09
Geo			1	**-0.35**	**0.57**	-0.14	**-0.36**	**0.31**	-0.05	**0.40**	**-0.60**
Soil				1	**-0.32**	-0.10	-0.09	0.05	0.07	-0.11	0.34
Prec					1	**-0.45**	**-0.32**	**0.63**	**0.27**	**0.76**	**0.40**
Tem						1	0.01	**-0.77**	**-0.47**	**-0.54**	-0.15
Asp							1	**-0.32**	0.22	**-0.31**	0.27
Alt								1	**0.44**	**0.78**	-0.18
Slo									1	**0.37**	0.07
Lan										1	-0.26
Dist											1

N*PV=Nitrogen*Photosynthetic vegetation cover, SR54=simple ratio, Geo=Geology, Prec=precipitation, Temp=temperature, Asp=Aspect, Alt=Altitude, Slo=Slope, Lan=Land use, Dist=distance to rivers. The **bold** values indicates that the correlation is significant at 95% confidence level ($p<0.05$).

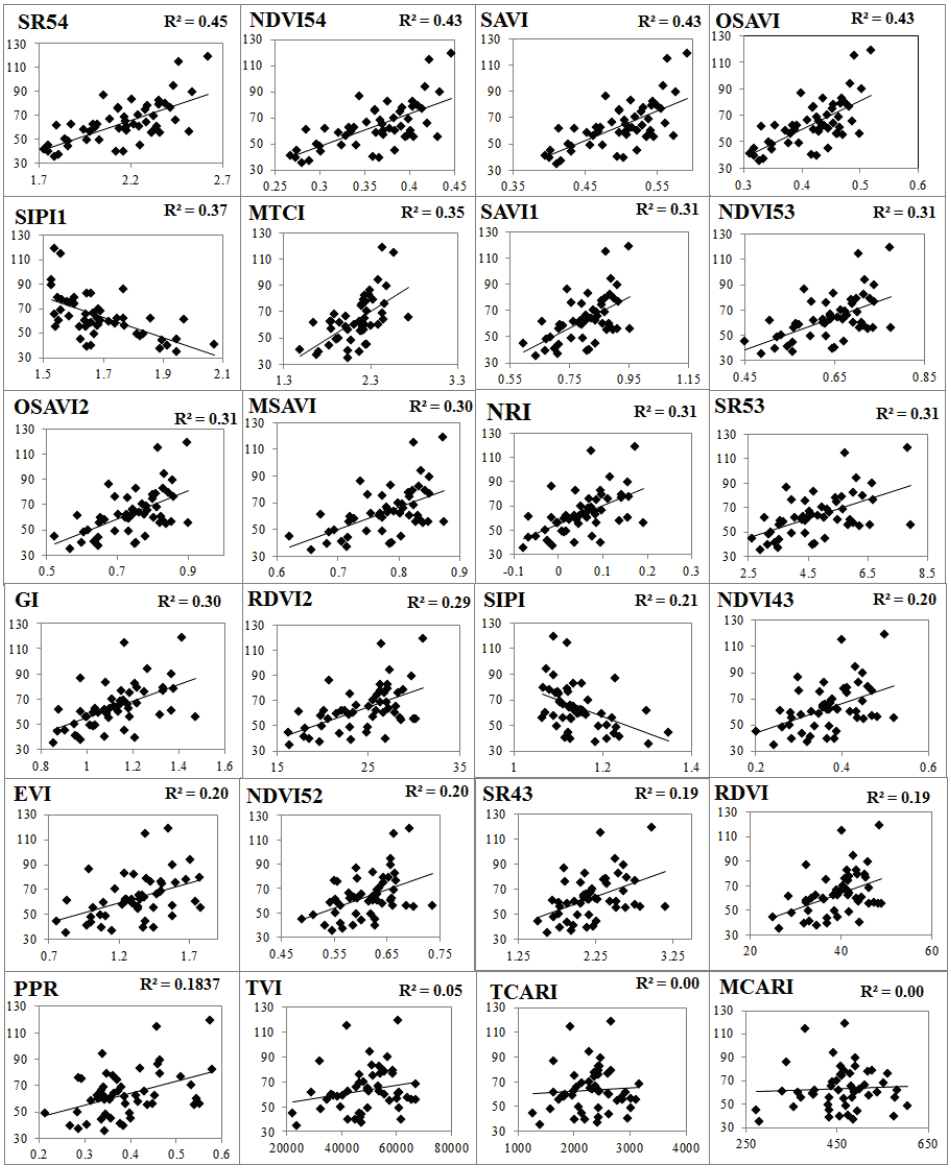

Figure 5.2: Scatterplots of canopy N (N*PV) and various vegetation indices (X-axis=vegetation index and Y-axis=N*PV). SR=Simple Ratio, NDVI=Normalized Difference Vegetation Index, SAVI=Soil Adjusted Vegetation Index, L=Soil Correction Factor, OSAVI=Optimized SAVI, MSAVI=Modified SAVI, TVI=Triangular Vegetation Index, RDVI=Renormalized Difference Vegetation Index, MCARI=Modified Chlorophyll Absorption Ratio Index, MTCI=MERIS Terrestrial Chlorophyll Index, PPR=Plant Pigment Ratio, NRI=Nitrogen Reflectance Index, SIPI=Structure Insensitive Pigment Index, GI=Greenness Index, EVI=Enhanced Vegetation Index, TCARI=Transformed Chlorophyll Absorption Ratio.

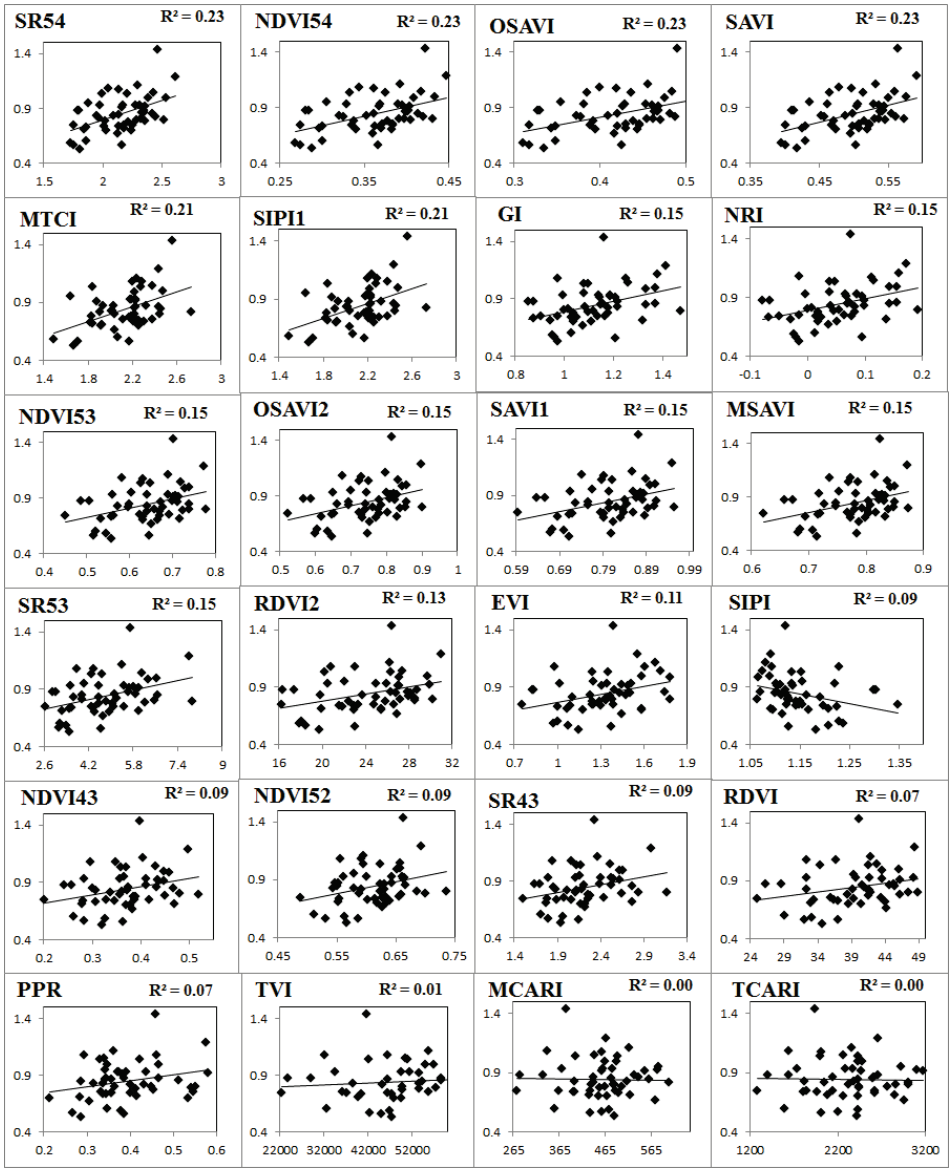

Figure 5.3: Scatterplots of foliar N (%) and various vegetation indices (X-axis=vegetation index and Y-axis=N*PV). SR=Simple Ratio, NDVI=Normalized Difference Vegetation Index, SAVI=Soil Adjusted Vegetation Index, L=Soil Correction Factor, OSAVI=Optimized SAVI, MSAVI=Modified SAVI, TVI=Triangular Vegetation Index, RDVI=Renormalized Difference Vegetation Index, MCARI=Modified Chlorophyll Absorption Ratio Index, MTCI=MERIS Terrestrial Chlorophyll Index, PPR=Plant Pigment Ratio, NRI=Nitrogen Reflectance Index, SIPI=Structure Insensitive Pigment Index, GI=Greenness Index, EVI=Enhanced Vegetation Index, TCARI=Transformed Chlorophyll Absorption Ratio.

For the estimation of foliar N, the non-linear PLSR produced a higher estimation accuracy (bootstrapped: R^2=0.48, RMSE=0.12%; 14% of the mean) than other techniques such as SMLR (Table 5.4). The SMLR+PCA yielded the second highest estimation accuracy (bootstrapped: R^2=0.45, RMSE=0.13%; 15% of the mean) and the least performing technique was the SMLR+Raw (bootstrapped: R^2=0.44, RMSE=0.14%; 17% of the mean) (Table 5.4). The interaction effects analysis of the selected variables in SMLR+Raw did not improve the results. Figure 5.4 shows the spatial distribution of foliar and canopy N at a regional scale. There is a clear N gradient between the western and the eastern part of the study area (Figure 5.4). The general pattern of foliar and canopy N follows the geological types, i.e. basalt and shale areas are characterized by more highly nutritious grass than the granitic derived grasses (Figure 5.4). The riparian zones or bottomlands in Figure 5.4 have high nutritious grass irrespective of a geological type. In Figure 5.4, part of the red dominating colour constitutes not only grass but trees, especially around the riparian zones.

5.5.3 Descriptive and exploratory statistics

The foliar N concentration across the study had a mean of 0.84%, as shown in Table 5.6. After converting the foliar N concentrations to the canopy integrated N using PV (i.e. N*PV), the recorded mean was 74.71 (Table 5.6). Grass N concentrations varied significantly according to geology (F=3.187, p=0.0322) and soil type (F=3.787, p=0.0096), as was confirmed by the ANOVA and the boxplots thereof are shown in Figure 5.5.

Table 5.6: Descriptive statistics of the data used

Variables (%)	Minimum	Maximum	Mean	Standard deviation	Coefficient of variation
Nitrogen	0.53	1.44	0.84	0.17	0.20
PV	40.00	100.00	74.71	11.42	0.15
N* PV	35.00	119.00	63.45	17.95	0.28

PV=photosynthetic vegetation cover, N*PV=Nitrogen*PV

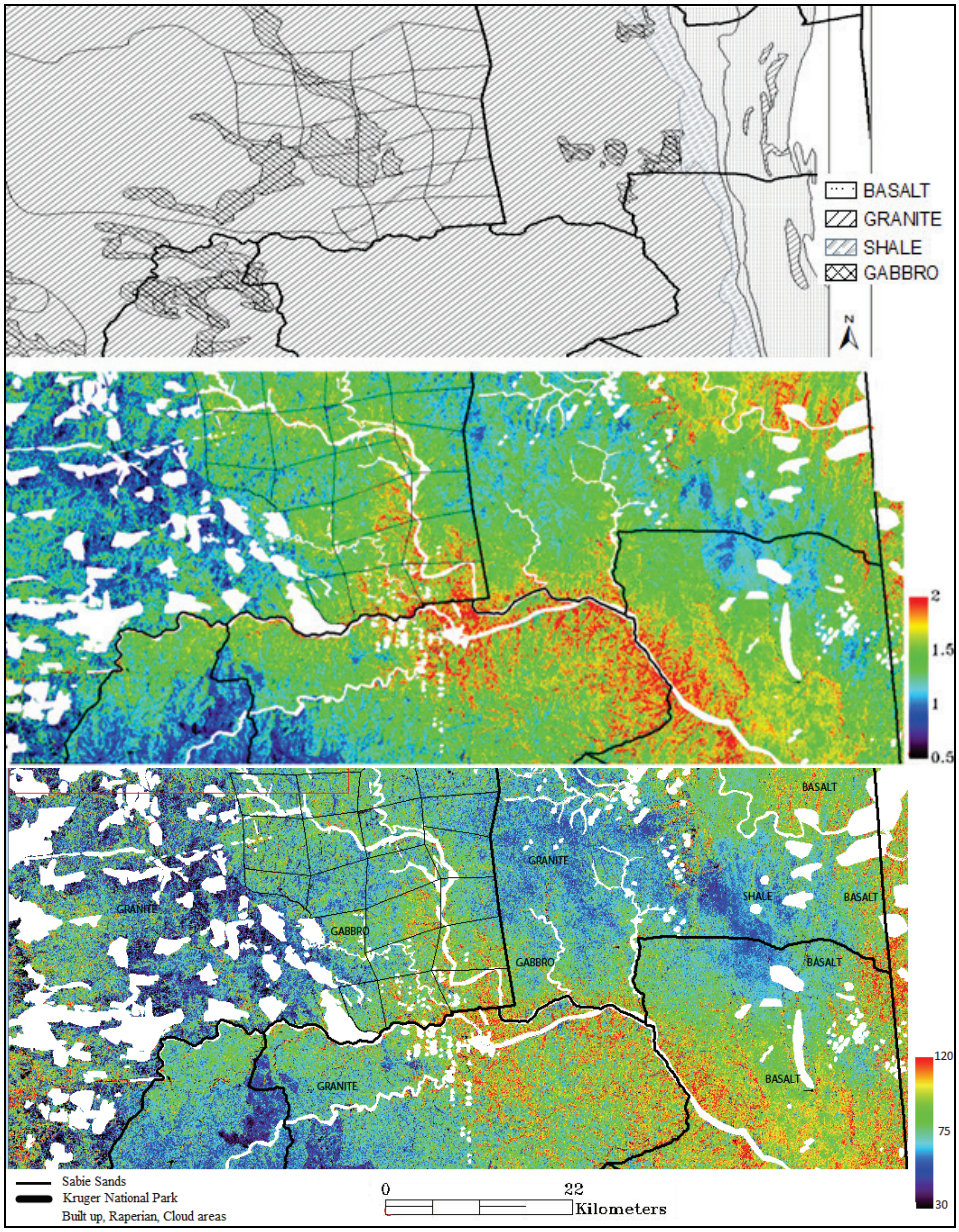

Figure 5.4: Map showing spatial distribution of the foliar N (middle) and canopy Nitrogen (N*PV) (bottom) in relation to geological types such as basalt, gabbro, granite and shale (PV=photosynthetic vegetation cover) (upper).

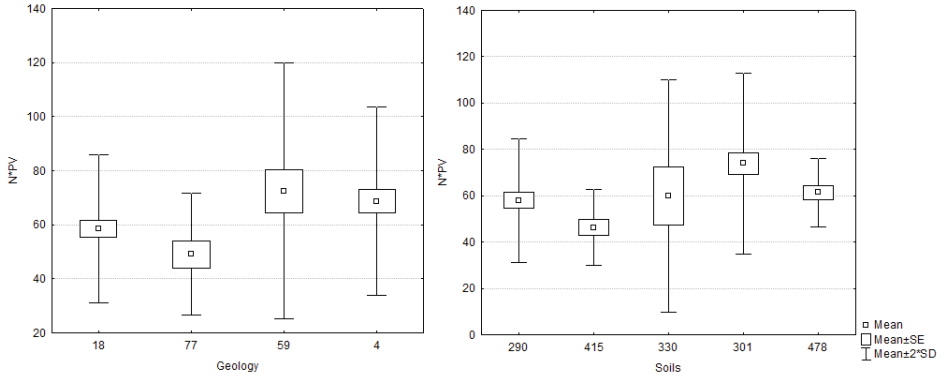

Figure 5.5: Boxplots indicating varying grass N over geological and soil types. 18=granite, 77=gabbro, 59=shale, 4=Basalt, 290=euric regosols, 415=albic arenosols, 330=calcic luvisols, 301=rhodic nitisols, 478=eutric leptic.

5.6 Discussion

The study investigated the utility of the red-edge band from the RapidEye sensor using vegetation indices, in order to determine which index correlated highly with foliar and canopy N. This index was then integrated with environmental variables to predict foliar and canopy N at a regional scale. SR54 was not only selected as the vegetation index with the highest predictive capability compared to other indices, it was also selected as a significant variable in the stepwise model successfully predicting both foliar and canopy N (Table 5.4). The performance of SR54 could be attributed to the use of red-edge waveband which contributed to the estimation of foliar N concentrations. Similar trends were observed for NDVI and SAVI, where the inclusion of the red-edge band improved the estimation results. The importance of the red-edge band is due to the fact that it is highly correlated to chlorophyll (Clevers et al., 2002; Cho and Skidmore, 2006) and insensitive to background effects (Zarco-Tejada et al., 2004b). It is known that there is a positive correlation between chlorophyll and foliar N (Vos and Bom, 1993; Yoder and Pettigrew-Crosby, 1995). This study is consistent with the *in situ* hyperspectral remote sensing studies reported by Mutanga and Skidmore (2007), Gong et al. (2002), and Cho and Skidmore (2006). Additionally, the performance of SR has not only been demonstrated with the retrieval of foliar biochemicals but also for biophysical parameters such as leaf area index (Jiang et al., 2005; Darvishzadeh et al., 2008) and biomass (Mutanga and Skidmore, 2004b).

The integrated modeling approach has produced higher grass N accuracy results compared to univariate approaches using only vegetation indices. The advantage of using an integrated modeling approach for N estimation is that both remote sensing and environmental variables are considered. The use of environmental variables is generally constrained by the lack of detail in

studies on a regional scale, rendering proper estimation of foliar N impossible. Remotely sensed imagery helps to provide the spatial detail important for characterizing foliar N in grass canopies. The combination of non-linear PLSR with environmental variables estimated foliar N with relatively high accuracy. The non-linear PLSR combined advantages of the conventional PLSR and an artificial neural network, i.e. maximizing the covariance between data sets and non-linear model fitting (Walczak and Massart, 1996). Additionally, the non-linear PLSR can be used with non-normal data.

Estimation of canopy N using SMLR integrating remote sensing (SR54) and environmental variables resulted in the highest estimation accuracy. SMLR selected SR54 and altitude, which predicted N with the lowest AIC value. Table 5.5 shows altitude to be significantly correlated with other environmental variables such as geology, mean annual precipitation, mean annual temperature, slope, aspect, and land use types. Soil and geology are also cited as factors which influence the distribution and concentrations of nutrients in grass (Venter et al., 2003; Mucina and Rutherford, 2006). Soils developed in basalts are generally high in nutrients, while the granitic soils are associated with low nutrient concentrations (Scholes et al., 2003; Venter et al., 2003). Grasses such as *Bothriochloa radicans*, *Urochloa mosambicensis*, and *Digitaria eriantha* are found dominating the basaltic-derived soils because of these high nutrient concentrations. These species also produce bigger leaves than the species usually found in the granitic-derived soils such as *Eragostris rigidior* and *Sporobolus* species. Table 5.5 shows a negative correlation between foliar N and precipitation. The western part of the study area is characterized by high precipitation and lower soil fertility-granite-derived soils, while the eastern part experiences low precipitation on high soil fertility-basaltic-derived soils. Precipitation plays a crucial role in dissolving organic matter for the uptake of minerals by plants (Pickett et al., 2003). Land use type, giving an indication of the practices or activities taking place in the study area, is important as it is related to mean annual precipitation. Land use types are characterized by a pronounced rainfall gradient, with the communal areas receiving more rainfall than the protected areas (SGR and KNP). In addition, land use activities generally affect the grass's response to differences in precipitation (Zhou et al., 2002). Altitude, aspect and slope influence the distribution of nutrient concentrations in grass through their effect on soil temperature and water run-off (Roberts, 1987). Steeper slopes normally have higher run-off leading to thin soil layers supporting less nutritious grass (Mutanga et al., 2004a). While valleys or bottomlands, characterized by deep soils, are the recipients of run-off from the steeper slopes, enabling the growth of high quality grasses (Scholes, 1990; Scholes et al., 2003).

Foliar N estimation results were low compared to the results for canopy N, for all methods. This is an indication that foliar N is not readily estimated by image spectra, which are largely dependent on canopy cover and properties (e.g. leaf area index). Canopy N, which can be accurately retrieved by image spectra, includes information about foliar N and structure or canopy productivity. The poorer results for foliar N, in comparison to the canopy N estimation, are consistent with other vegetation biochemical studies, including the ones focusing on foliar and canopy chlorophyll (Yoder and Pettigrew-Crosby, 1995; Asner, 1998; Asner et al., 1998; Asner and Martin, 2008; Darvishzadeh et al., 2008b).

In this study, the interaction effect between foliar N and biomass was minimized by conducting fieldwork and acquiring the RapidEye image during peak productivity in wet season. During this period, the relationship between biomass and vegetation indices saturates, portraying a difficulty of estimating biomass (Figure 5.6). The amount of light that can be absorbed in the red region of the spectrum plateaus during peak productivity (Tucker, 1977; Thenkabail et al., 2000; Mutanga and Skidmore, 2004b). Additionally, the NIR reflectance continues to increase, because addition of new leaves influences the multiple scattering (Kumar et al., 2001). The vegetation index (e.g. NDVI) change slightly, while causing a poor relationship between the index and biomass (Figure 5.6). Therefore, it is assumed that foliar N can be successfully estimated with minimum interaction effects (Skidmore et al., 2010).

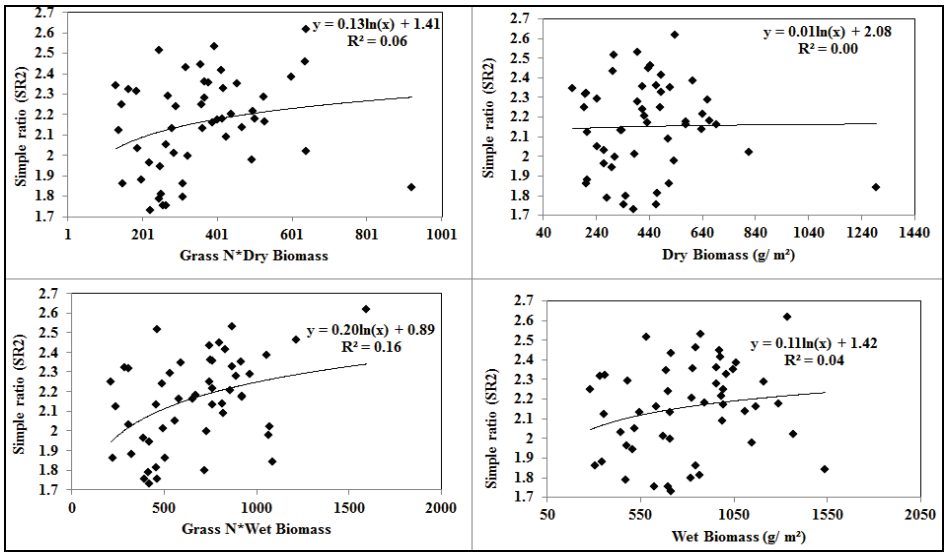

Figure 5.6: Figure shows the saturation or no relationship between a vegetation index and with the interaction between N and biomass (Top Left), dry biomass (Top Right), interaction between foliar N and wet biomass (Bottom Left), and wet biomass (Bottom Right).

In this study 60% of the variance of canopy N is attained using multispectral remote sensing data (i.e. RapidEye), which is comparable to some of the hyperspectral studies. The performance of the RapidEye data in estimating foliar and canopy N is associated with the presence of the red-edge band. The hyperspectral studies demonstrated the use of the red-edge position to estimate chlorophyll and N (Cho and Skidmore, 2006; Darvishzadeh et al., 2008b). Foliar and canopy N were estimated because of the positive correlation between chlorophyll and N (Yoder and Pettigrew-Crosby, 1995). Previous studies using hyperspectral data (*in situ* or airborne) achieved high foliar N retrieval accuracies. Using airborne systems, foliar N estimation was reported to achieve accuracies of 48 to 80% (Huang et al., 2004; Mutanga and Skidmore, 2004a; Skidmore et al., 2010; Knox et al., 2011). An explained variance of 48% was obtained during the dry season, while 80% or more was obtained during the wet season. This shows the importance of seasonality or plant phenology in the estimation of foliar biochemical levels.

The results of this study demonstrated that foliar and canopy N can be mapped at a regional scale using spaceborne multispectral remote sensing data during times of peak productivity. The red-edge band of RapidEye was found to be important in achieving this goal (compared to traditional multispectral sensors such as SPOT and Landsat). Foliar N is an indicator of crude protein (Clifton et al., 1994; Wang et al., 2004), which forms a main nutrient requirement (Prins and van Langevelde, 2008), and could be used for understanding the distribution, densities and population dynamics of herbivores in protected and communal areas (McNaughton, 1988, 1990; Ben-Shahar and Coe, 1992; Heitkönig and Owen-Smith, 1998; Mutanga et al., 2003). Photosynthetic vegetation cover is one of the canopy parameters determining key ecosystem functions, e.g. rate of carbon and nutrient intake (Guerschman et al., 2009). In addition, grass canopy N has a structural component as it was derived in combination with photosynthetic vegetation cover. The grass structure is generally defined by biochemistry, architecture, morphology and species composition (Burke, 1997; Drescher et al., 2006a; Drescher et al., 2006b). The grass structure affects grazing behavior of the herbivores (Drescher et al., 2006b). Drescher et al., (2006b) postulated that grass structure affects cattle grazing behavior in the South African savanna. Therefore, canopy N may outperform foliar N when aiming to understand the distribution of herbivores, since it can be estimated and mapped at a higher accuracy. The study further demonstrated the use of integrated modeling for grass N estimation. Regional nutrient maps could provide useful information to farmers, resource managers and park stewardships for sound planning and management of savanna ecosystems.

5.7 Acknowledgement

We would like to thank the Council of Scientific and Industrial Research (CSIR), the University of Twente Faculty of Geoinformation Science and Earth Observation (UT-ITC), and Wageningen University for the funding provided. We would also like to acknowledge the South African Department of Science and Technology as well as National Research Foundation (NRF)'s Professional Development Programme (PDP) for the funding. We appreciated the field work assistance by Mr Thulani Selaule, as well as the involvement of South African National Parks (SANPARKS), especially Dr Izak Smith, Mrs Thembi Khoza, Mrs Patricia Khoza, Mr Adolf Manganyi, and Mrs Onica Sithole for making the fieldwork possible.

Chapter 6

Synthesis: Remote sensing estimation of grass quality from local to regional Scale

6.1 Introduction

The importance of grass foliar nutrient concentrations as an indicator of grass quality has been well described in ecology related studies, and is crucial for planning and management of the livestock and wildlife grazing (Drent and Prins, 1987; McNaughton, 1988; Ben-Shahar and Coe, 1992; Meissner et al., 1999). The main indicators of grass quality are foliar nitrogen (N) and phosphorus (P) concentrations (Meissner et al., 1999). The importance of foliar N and P concentrations extend to influencing the physiological processes such as photosynthesis, leaf respiration and growth rates (Mooney, 1986; Martin and Aber, 1997; Ollinger et al., 2002). Grass quality in this context is defined as a major resource for attracting herbivores. A resource can be defined as "a usable energy or any biotic or abiotic substance directly exploited by an organism, which include food, nutrients, water, atmospheric gas compounds, as well as light, and the use of which can lead to the (temporally) exhaustion of that resource (i.e. depletable and non-depletable resource)" (Prins and van Langevelde, 2008). Studies indicated that foliar N and P concentrations are one of the factors affecting the feeding patterns, distribution, densities, movements of herbivores (wild and livestock) (Drent and Prins, 1987; McNaughton, 1988, 1990; Olff et al., 2002). Foliar N concentration generally correlates with protein (Clifton et al., 1994; Wang et al., 2004), which is the main nutrient requirement of the herbivores. On the other hand, the foliar P concentration is the main requirement for reproduction, and for lactating herbivores (McNaughton, 1990; Duncan, 1992). A combination of foliar N and P estimation is important for understanding the nutrient limitation through a tool known as foliar N: P ratio (Koerselman and Meuleman, 1996; Ludwig et al., 2001; Güsewell, 2004). The foliar N:P ratio reflects the balance of N and P supply which influences the plants at all levels, i.e. the growth and reproduction of individual plants, plant species interactions, composition and diversity (Güsewell, 2004). The differences in foliar biochemistry as captured by foliar N: P ratio in plants provide information on which nutrient is limiting plant growth or reproduction and could influence feeding activity or patterns of herbivores. Therefore, spatial information about grass nutrients could be useful to guide farmers, resource managers and land use planners in sustainable management of their grazing land.

Estimation and mapping of grass nutrients has been successful using laboratory, *in situ* and airborne hyperspectral measurements, also known as spectroscopy (Kumar et al., 2001; Majeke et al., 2008; Kokaly et al., 2009; Skidmore et al., 2010; Knox et al., 2011). In the past decades scientist developed remote sensing techniques to extract vegetation biophysical (e.g. leaf area index, biomass) and biochemical parameters (e.g. chlorophyll, fibre, lignin, P, protein and N) (Huang et al., 2004; Darvishzadeh et al., 2008b; Cho and Skidmore, 2009; Skidmore et al., 2010; Knox et al., 2011). The

techniques for extracting biochemical concentrations were successfully developed and applied in several biomes, e.g. grassland (Darvishzadeh et al., 2008b), savanna (Mutanga and Skidmore, 2004a; Skidmore et al., 2010; Knox et al., 2011), agriculture (Zarco-Tejada et al., 2004b), forestry (Asner et al., 2008; Majeke et al., 2008; Schlerf et al., 2010). Remote sensing techniques to extract vegetation parameters are as follows; use of vegetation indices (Daughtry et al., 2000; Haboudane et al., 2004), full spectrum analysis (Ramoelo et al., 2011b), absorption features (Skidmore et al., 2010; Knox et al., 2011) as well as integrated modeling approach (remote sensing and environmental variables) (Cho et al., 2009; Knox et al., 2011; Ramoelo et al., *under review*) and inversion of radiative transfer models (Darvishzadeh et al., 2008a). All these techniques have their own challenges (Kumar et al., 2001; Majeke et al., 2008; Kokaly et al., 2009). The vegetation indices are known to saturate during peak biomass productivity when used to estimate vegetation parameters (Mutanga and Skidmore, 2004b). The relationship between foliar biochemicals (e.g. N) and chlorophyll deteriorates as the leaf senesce (Wang et al., 2009), which limits the use of indices for estimating foliar biochemical concentrations. The use of absorption features especially to the fresh leaf spectra, water absorption effects are the main challenge. Few studies have tested the applicability of the integrated modeling, i.e. the use of remote sensing and environmental or ancillary variables (Cho et al., 2010; Knox et al., 2011; Ramoelo et al., *under review*) to estimate foliar biochemical concentrations. There is a need to test this approach for regional estimation of foliar biochemical concentrations. A challenge in estimating foliar biochemical concentrations, for example N, using indices is the interaction effects between N and biomass (Skidmore et al., 2010). Studies argue that if foliar biochemical estimation is not biomass corrected, then the interaction between N and biomass is likely to compromise the estimation of N (Plummer, 1988b; Skidmore et al., 2010). Thus, this challenge has been addressed in this study. Finally, this study estimated and mapped grass N at a regional scale for the first time in savanna ecosystem.

The aim of this study was to develop and improve techniques for estimating and mapping foliar biochemical concentrations from local to regional scale. The specific objectives were as follows; (1) to test water removed spectra for foliar N and P estimation as compared to the existing spectral techniques, (2) To estimate foliar N: P ratio using *in situ* hyperspectral measurements or field spectroscopy, (3) to investigate the applicability of the non-linear partial least square regression for integrating *in situ* hyperspectral meausrements and environmental variables to estimate foliar N and P and (4) to investigate the utility of the red edge band in RapidEye data for estimating foliar and canopy N at regional scale. Based on achieving these objectives, synthesis chapter will cover the following aspects; water removal, integrated modelling, issues of scale, phenology as well as key conclusions and future research.

6.2 Water removal spectra to minimize water absorption effects

The plant reflectance as measured by remote sensing device is influenced by various factors, including atmospheric, soil background, water absorption and sensor related effects (Kumar et al., 2001; Majeke et al., 2008; Kokaly et al., 2009). Of importance, the plant reflectance is also influenced by the foliar biochemical concentrations (Kumar et al., 2001). The wavelength regions of the leaf reflectance are affected differently by the foliar compounds (Kumar et al., 2001; Majeke et al., 2008). The visible-near infrared region is more sensitive to the absorption by foliar pigments (i.e. chlorophyll) (Ustin et al., 2009) and leaf area index (Darvishzadeh et al., 2008b). The shortwave infrared (SWIR) is characterized strong water absorption and subtle absorption features of biochemical concentrations such protein, nitrogen, cellulose, lignin and starch, especial on the fresh leaf spectra (Curran, 1989; Kumar et al., 2001; Majeke et al., 2008; Kokaly et al., 2009). Chlorophyll absorptions in the visible region are due to the electron transitions, and features are centred around 430 nm and 630 nm (for chlorophyll a) and 460nm and 640nm (for chlorophyll b) (Curran, 1989; Kumar et al., 2001; Majeke et al., 2008; Kokaly et al., 2009). The absorption features in the near infrared (NIR) to SWIR are influenced by harmonics and overtones of the main stretching frequencies such as C-H, N-H and O-H bonds (Curran, 1989; Kumar et al., 2001; Majeke et al., 2008; Kokaly et al., 2009). "Proteins, which are normally in the form of D-ribulose 1-5-diphosphate carboxylase, are the most abundant N bearing compounds in green leaves" (Kumar et al., 2001). Protein constitutes about 30 to 50% of N of fresh leaves (Elvidge, 1990). The absorption features of protein and N are centred around 1020nm, 1510nm, 1690nm, 1730nm, 1940nm, 1980nm, 2060nm, 2180nm, 2240nm, 2300nm, 2350nm (Curran, 1989; Kumar et al., 2001).

Coincidentally, the region of the fresh leaf spectrum is dominated by foliar biochemical concentrations and water absorption (Fourty and Baret, 1998; Kumar et al., 2001; Majeke et al., 2008; Ramoelo et al., 2011b). Absorptions by these foliar biochemical concentrations are not very strong and are generally masked by water absorption in fresh leaves (Gao and Goetz, 1994, 1995; Dawson and Curran, 1998; Kumar et al., 2001; Zhao et al., 2006; Majeke et al., 2008). In dry leaf spectra, biochemical absorption is generally highly differentiated and well correlated to the concentrations of these biochemicals (Card et al., 1988; Elvidge, 1990). Nevertheless, water absorption effects are still cited as a challenging problem in foliar biochemical estimation (Kumar et al., 2001; Mutanga and Skidmore, 2004a; Majeke et al., 2008; Kokaly et al., 2009). To estimate foliar biochemicals, derivatives, continuum-removal, and log transformed spectra (Log(1/R)) are known to enhance absorption features of foliar biochemical concentrations, while

minimizing atmospheric, soil background, and water absorption effects, as well as data redundancy (Yoder and Pettigrew-Crosby, 1995; Dawson and Curran, 1998; Zarco-Tejada et al., 2004a; Ramoelo et al., 2011b). Nonetheless, several studies further recommended the removal of water absorption effects, and postulated an increase in the estimation accuracy of the foliar biochemical concentrations (Mutanga and Skidmore, 2004a; Majeke et al., 2008).

We addressed the issue of water absorption effects and indeed improved the estimation of foliar N and P at a Laboratory level (from Greenhouse experiments) (Chapter 2) and foliar N: P at field level (Chapter 3). To reduce water absorption effects on weak biochemical absorption features, the water removed spectra (WR) were derived from a non-linear least-squares spectral matching technique calculating a fresh leaf spectrum as a non-linear combination of a leaf water spectrum and a dry matter spectrum, Eq. 2.1 and 2.2 (Schlerf et al., 2010; Ramoelo et al., 2011b). Figure 2.2 and 2.4, show that the water removal spectra achieved high performance for estimating foliar N and P, as compared to the derivatives, continuum removal and log (1/R). By removing the water effect across the fresh leaf spectra, WR enhances weak or subtle absorption features especially on NIR and SWIR which are the regions on the electromagnetic spectrum most affected by water absorption. The importance of the NIR and SWIR are illustrated in Chapter 2: Figure 2.3 and Figure 2.5 where more bands sensitive to foliar biochemical concentrations are selected using WR instead of reflectance and other spectral data. The selected bands are also related to the regions of spectrum influenced by the vibrational mechanisms through harmonics and overtones (Curran, 1989; Kumar et al., 2001). Generally this study showed the applicability of the WR technique for savanna grass species, with a higher R2 of 0.87 for N estimation. Gao and Goetz (1994, 1995) successfully implemented the WR technique for lignin and cellulose estimations, highlighting the importance of minimizing water effects on the SWIR.

In Chapter 2, the continuum removal technique achieved second highest accuracy for estimating foliar N concentrations with both PLSR and SMLR. Continuum removal enhanced the differences in absorption strength (Schmidt and Skidmore, 2001; Schmidt et al., 2004). The highest R2 obtained in this study using continuum removal was 0.81. This is consistent with the N retrieval accuracy of forest sites reported by Kokaly and Clark (1999), where an R2 of 0.75 to 0.94 was attained using continuum removal and continuum removal-derived indices. This study attained higher accuracy results for N based on continuum removal than a study by Mutanga et al. (2005) on Cenchrus ciliaris grown in the greenhouse. Estimating N with Log(1/R) yielded higher retrieval accuracy than with reflectance, but not as high as with the WR and continuum removal techniques. Yoder and Pettigrew-Crosby

(1995) showed Log(1/R) performed accurately estimating N concentrations, compared to reflectance. Similar results were also attained by Fourty and Baret (2001). They argued that by transforming reflectance to absorbance Log(1/R) values the accuracy of biochemical estimates was improved. Log(1/R) is likely to be used instead of the original reflectance because of the linear relation between the absorbing components and its contribution to the Log(1/R) value at the wavelength absorbed (Hruschka, 1987). However, the present study shows the performance of WR to be higher than that of Log(1/R).

The P estimation accuracy as demonstrated by the performance of WR and first derivative highlights the importance of reducing the influence of water on the fresh leaf spectra. As shown in Table 2.2 and Figure 2.5, many bands highlighted by the models as sensitive to P correspond to known absorption features in the SWIR. Asner and Martin (2008) also argued that the retrieval of foliar P using hyperspectral remote sensing could be associated with the stoichiometry (i.e. indirect estimation based on the relationship with other foliar biochemicals) (Elser et al., 1996). Currently, few studies have focused on the estimation of P using field spectroscopy (Bogrekci and Lee, 2005; Ramoelo et al., 2011b) and airborne (Mutanga and Kumar, 2007; Knox et al., 2011). Bogrekci and Lee (2005) used wavelengths from 225 to 2550 nm, while Mutanga and Kumar (2007) showed that bands in the SWIR were more sensitive to P, which is similar to the findings in Chapter 2. The high precision of foliar N and P estimation using the WR technique is evident from the lower confidence limits of the correlation coefficient and the RMSE derived from the bootstrapping technique (Table 2.1).

The usefulness of WR spectra were further demonstrated in estimating foliar N: P ratio towards predicting grass foliar nutrient limitation at field level (Chapter 3). The results showed that foliar N: P can be estimated using spectral data and PLSR. CR and WR spectra predicted foliar N: P ratio with higher accuracy as compared to FD, Log (1/R) and reflectance spectra. The performance of CR and WR spectra were attributed to their ability to minimize sensor and water effects on the fresh leaf spectra, respectively, while enhancing absorption features for foliar biochemical concentrations. Chapter 2 demonstrated that WR could be used to as one of the techniques for estimating foliar biochemical concentrations, as shown from the experimental study in the controlled environments. WR spectra are applied for the first time to estimate foliar N: P and produced promising results. The CR spectra enhance the absorption features of foliar biochemical concentrations through enhancing the differences in the absorption strength (Kokaly and Clark, 1999; Mutanga et al., 2004c). CR spectra have been used for estimating foliar N and P, not N: P but showed a good performance. Several studies demonstrated the applicability of CR spectra for foliar

biochemicals, especially for foliar N and P (Huang et al., 2004; Ramoelo et al., 2011b). The study demonstrated the possibility to predict foliar N: P ratio using in situ hyperspectral data, and SWIR found to be highly sensitive to foliar N: P (Figure 3.2 and 3.3). The main recommendation of the studies in Chapter 2 and 3, were to test the applicability of WR to estimate foliar N, P as well as N: P at airborne (e.g. Hymap using NIR and SWIR) and regional scale through satellite remote sensing measurements (challenge is the availability of satellite sensors with SWIR). Nevertheless, the possibility to estimate foliar biochemical concentrations at regional scale could be successful using the integrated modeling approach (combining remote sensing and ancillary or environmental variables).

6.3 Integrated modeling approach: Towards and regional estimation of foliar biochemicals

The developments of an integrated modeling approach for estimating foliar biochemical concentrations emanated from the need to estimate foliar biochemicals at regional scale. Progress for estimating foliar biochemical concentrations were achieved using laboratory (Mutanga et al., 2003), field (Yoder and Pettigrew-Crosby, 1995; Mutanga et al., 2004c; Abdel-Rahman et al., 2010) and airborne remote sensing measurements (Huang et al., 2004; Skidmore et al., 2010; Knox et al., 2011). Estimation and mapping of foliar biochemical concentrations at regional scale is often hindered by unavailability of satellite sensors with a spectral and spatial resolution adequate to retrieve foliar biochemical concentrations. Knox, et al. (2010a) and Cho et al. (2009) demonstrated that combining remote sensing and environmental variables, referred to as "integrated modeling approach" in this study, and improves estimation of foliar biochemical concentrations compared to using remote sensing data only. Savanna ecosystems are diverse and heterogeneous in soil and plant moisture, soil nutrients, fire regime, grazing pressures and anthropogenic activities (Ben-Shahar and Coe, 1992). This makes the estimation of grass N and P using remote sensing in savannas a challenging task (Mutanga and Kumar, 2007; He and Mui, 2010). Grass quality is influenced by geology (Ben-Shahar and Coe, 1992; Grant and Scholes, 2006), soil (Heitkönig and Owen-Smith, 1998; Cho et al., 2010), precipitation and temperature (Ben-Shahar and Coe, 1992), topography or catena position (Seagle and McNaughton, 1992; Mutanga et al., 2004a) as well as aspect (Mutanga et al., 2004a) and land use types (Zhou et al., 2002). The question is: could an integrated approaching involving remote sensing and environmental variables improve the assessment of grass quality as opposed to remote sensing variables only? We assumed that a modeling approach that exploits the strength of remote sensing and environmental variables could potentially improve the

assessment of ecosystem state and functioning at various geographic scales (Mutanga et al., 2004a; Cho et al., 2009; Knox et al., 2011).

The study in chapter 4 (at field level) and 5 (at regional scale) demonstrated that integrating remote sensing data and environmental variables improves foliar N and/or P estimations, as compared to using remote sensing only. In addition, using the environmental variables only might not be feasible because of their availability, scale and resolution. Therefore an ideal modeling approach has been seen as integrated modeling approach, especially for regional estimation of foliar biochemical concentrations. In Chapter 4, foliar N estimation is significantly improved when using integrated modeling, remote sensing variables such as narrow-band SR, REP, and protein absorption features centred at 910 nm and 1020 nm. In Chapter 5: Figure 5.2 and 5.3, high N estimation accuracy was achieved through the inclusion of red edge band from RapidEye image to the vegetation indices. In particular, the SR index based on band 5 (805 nm) and band 4 (710 nm) resulted to higher estimation accuracy of both foliar and canopy N. The above results were mainly attributed to the fact that red edge is highly correlated to chlorophyll (Cho and Skidmore, 2006; Darvishzadeh et al., 2008b) and it minimizes soil background effects (Zarco-Tejada et al., 2004b). Positive correlation between chlorophyll and N has been reported by Yoder and Pettigrew-Crosby (1995). The results are consistent with studies focusing on foliar N concentration using in-situ hyperspectral measurements (Gong et al., 2002; Knox et al., 2010b). Gong et al. (2002) demonstrated the utility of blue and red edge regions for estimating foliar N concentrations. The protein absorption features at 910 nm and 1020 nm contributed to foliar N estimation model as they are influenced by various vibration mechanisms such as C-H stretch, 3rd overtone and N-H stretch (Curran, 1989; Kumar et al., 2001). The visible region of the spectra is characterized by the electron transition while the near and shortwave infrared are characterized by the various bond vibration (Curran, 1989; Kumar et al., 2001). Several studies used these absorption features not only for foliar estimation, but also for biomass and LAI estimations (Cho et al., 2007a; Darvishzadeh et al., 2008b).

For P estimation, the contribution of in situ hyperspectral variables in estimating foliar P concentrations was based on several biochemical absorption features, red edge position, narrow-band NDVI and SR (Chapter 4). Unlike the foliar N concentration with defined absorption features, foliar P does not have specific absorption features. The absorption features likely to be linked to P, are coincidentally linked to other foliar biochemical concentrations such as protein, N, water, starch, cellulose and lignin (Kokaly et al., 2009). Few studies on prediction of foliar P concentrations using hyperspectral remote sensing found that most sensitive bands are located in the SWIR (Bogrekci and Lee, 2005; Mutanga and Kumar, 2007; Cho et al.,

2010; Ramoelo et al., 2011b). As shown above, these regions are characterized by the various vibration mechanisms imposed by several biochemical concentrations associated with protein, N, sugar and starch. Limited contribution of the red edge position, narrow-band NDVI and SR was expected since the bands used to calculate these indices are all located in the visible region of the spectrum. This is consistent with the results of Gong et al. (2002) who attempted to use vegetation indices derived from visible bands for estimating foliar phosphorus and generally reported low correlations. A recent study by Knox, et al. (2011) demonstrated that visible and near-infrared region can be used to estimate foliar Phosphorus in the dry season, where 57% of variance of foliar P concentrations was explained. This was possible because of the inclusion of narrow-band absorption features and environmental or ancillary variables.

In Chapter 5, foliar and canopy N were estimated using two approaches; firstly using modified vegetation indices and secondly; through integrated modeling approach. The integrated modeling approach has produced higher grass N accuracy results compared to univariate approach using vegetation indices only. The advantage of using an integrated modeling approach for N estimation is the use of both remote sensing and environmental variables. Environmental variables are constrained by a lack of details and might not be used alone to estimate foliar N. Therefore, remotely sensed imagery helps to provide the spatial details which are useful to characterize the foliar N. Environmental parameters selected in Chapter 4 for N estimation are geology, soil types, land use types, distance to rivers and temperature, while for foliar P estimation geology, land use types, soil types, slope, and temperature. In Chapter 5, altitude was selected to estimate foliar and canopy N. For foliar N estimation distance to rivers and aspect were significant. Soil and geology are also cited as one of the factors which influence the distribution and concentrations of nutrients in the grass (Venter et al., 2003; Mucina and Rutherford, 2006; Skidmore et al., 2010; Skidmore et al., 2011). Soils developed in basalts are generally associated with highly nutritious grasses, while the granitic ones are associated with low nutritious grass (Scholes et al., 2003; Venter et al., 2003). Grasses such as Bothriochloa radicans, Urochloa mosambicensis and Digitaria eriantha dominate the basalt-derived soils with high nutrient concentrations. These species also produce bigger leaves than the species usually found in the granite-derived soils such as Eragrostis rigidior and Sporobolus species. Chapter 5: Table 5.5 showed that there is a negative correlation between foliar N and precipitation. The western part of the study area is characterized by high precipitation with lower soil fertility-granite-derived soils, and the eastern part experience low precipitation with high soil fertility-basaltic-derived soils. Precipitation plays a crucial role in dissolving organic matter for uptake of minerals by the plants (Pickett et al., 2003). Land use types which

are the indication of the practices or activities taking place in the study area are important as they relate to mean annual precipitation. Land use types are characterized by a pronounced gradient of rainfall, where the communal areas are characterized by high rainfall compared with the protected areas (SGR and KNP). In addition, land use activities generally affect the grass response associated with differences in precipitation (Zhou et al., 2002). These results are consistent with (Ramoelo et al., under review). Altitude, aspect and slope influence the distribution of nutrient concentrations in the grass, through their effects on soil temperature and water run-off (Roberts, 1987). Steeper slopes normally have high run-off leading to thin soil layer which support less nutritious grass (Mutanga et al., 2004a). While valleys or bottomlands are characterized by deep soils and are a recipient of the run-off materials from the steeper slopes, they support high quality grasses (Scholes et al., 2003). Several statistical techniques were used for integrated modeling to estimate foliar biochemical concentrations in this study.

Progress on the extraction of foliar biochemical concentrations benefited from various statistical multivariate techniques. A commonly used technique is SMLR (Grossman et al., 1996; Martin and Aber, 1997; Kokaly and Clark, 1999; Huang et al., 2004; Kokaly et al., 2009), which has major demerits pertaining to multicollinearity and overfitting (Curran, 1989; Grossman et al., 1996). PLSR was used as an alternative to address the demerits of SMLR for extracting foliar biochemicals (Huang et al., 2004). PLSR decompose the independent variables into un-correlated latent variables or factors (Geladi and Kowalski, 1986; Geladi et al., 1999) which can be used to precisely estimate foliar biochemical concentrations. In Chapter 2, PLSR was successfully applied and on average it out-performed SMLR in estimating foliar N and P. This is mainly because PLSR minimizes the multicollinearity effects by decomposing the spectral data into non-collinear latent variables (Walczak and Massart, 1996; Huang et al., 2004). Over-fitting may be minimized using PLSR by selecting an optimal number of latent variables rather than having more redundant explanatory variables (Viscarra Rossel, 2007, 2008). In Chapter 4 and 5, we proposed a newly applied technique called radial basis function PLSR (i.e. non-linear PLSR) for extracting the foliar biochemical concentration. For both foliar N and P estimation, the non-linear PLSR achieved higher accuracy than conventional PLSR (Chapter 4). The non-linear PLSR has the mutual advantages of the linear nature of RBF (which is a neural network) and of the power of PLSR to maximize covariance between data sets, while minimizing the variance of the prediction (Walczak and Massart, 1996). Maximizing co-variance between data sets is done through decomposition of the independent variables into uncorrelated latent variables which is important for; (1) reducing the dimensionality of the data (Ehsani et al., 1999; Geladi et al., 1999) and (2) minimizing the over-fitting and multicollinearity (Walczak and Massart, 1996; Huang et al., 2004), to

enhance the transferability of models (Crawley, 2006). The inclusion of the RBF model, which is a neural network in nature, makes the non-linear PLSR to be nonparametric and can be applied without being constrained by the statistical distribution (Atkinson and Tatanall, 1997). We demonstrated the power of the non-linear PLSR for estimating foliar N and P coupled with the integrated in situ as well as spaceborne remote sensing and environmental variables, during peak biomass productivity.

6.3.1 Minimizing biomass and foliar biochemical interaction

A challenge in estimating foliar biochemical concentrations is the interaction effects of foliar biochemicals and biomass (Plummer, 1988b; Skidmore et al., 2010). In this study, we sought to minimize the interaction effects between foliar N and biomass by conducting fieldwork and acquiring the field spectra (Chapter 3 and 4) and RapidEye image acquisition (Chapter 5) during the peak productivity in wet season. During this period, the relationship between biomass and vegetation indices saturates, as portrayed in Chapter 5: Figure 5.6. The amount of light that can be absorbed at the red region of the spectrum reaches a plateau during the peak productivity (Tucker, 1977; Mutanga and Skidmore, 2004b). Additionally, the NIR reflectance continue to increase, because addition of new leaves influence the multiple scattering (Kumar et al., 2001). The vegetation index (e.g. NDVI or SR) changes slightly, while causing a poor relationship with biomass (Figure 5.6). It is during this period that we can successfully estimate foliar N, with less interaction effects imposed by biomass (Skidmore et al., 2010). This was demonstrated in Chapter 5: Figure 5.2 and 5.3 show significant relationships between canopy and foliar N, respectively, and vegetation indices.

6.4 Issues of scale

To date, estimation of foliar biochemicals were successful done using hyperspectral remote sensing at laboratory, field and airborne scale (Huang et al., 2004; Mutanga and Skidmore, 2004a; Mutanga et al., 2004c; Majeke et al., 2008; Kokaly et al., 2009; Skidmore et al., 2010; Knox et al., 2011; Ramoelo et al., 2011b). Scaling foliar biochemical estimation from laboratory, field, airborne sought to be challenging (Asner and Martin, 2008; Knox et al., 2010a). The plant biochemical absorption features are not stable when scaling from leaf to canopy levels (Knox et al., 2010a). This trend was also observed using the integrated modeling approach, in Chapter 4 and 5, where even the selection of absorption features, vegetation indices and environmental variables are different. Studies for foliar biochemical estimation at regional scale is hindered by a limited number of spaceborne hyperspectral systems with SWIR wavebands and high spatial resolution (<10 m). Recent spaceborne sensors such as RapidEye and WorldView II with the red edge band provide an opportunity to estimate foliar biochemical

concentrations, especially foliar N at regional scale. This is because foliar or canopy N could be estimated using visible region of the reflectance (Chapter 5). The advantage of using RapidEye sensor is that each scene covers approximately 5000 km2 (RapidEye, 2010). Studies using hyperspectral remote sensing demonstrated a successful application of the red edge region for estimating chlorophyll and N mainly at field and airborne measurements level (Horler et al., 1983; Mutanga and Skidmore, 2004a). Chapter 5 investigated the utility of red edge band embedded in the RapidEye sensor for the estimation of foliar biochemicals, especially N at regional scale. In this chapter, modified indices (Table 5.2) such as simple ratio denoted as SR54 and altitude were significant in estimating both foliar and canopy biochemicals. In Chapter 5: Figure 5.4, we mapped foliar and canopy N at regional scale for the first time in the South African savanna. What is interesting is that the predicted patterns of foliar and canopy N corroborates with geological types. For example, highly nutritious grasses were expected on the basalt-derived soils, while granite generally supports low fertility soil, hence low nutritious grasses (Ben-Shahar and Coe, 1992; Venter et al., 2003).

Generally, canopy biochemical estimation accuracies are higher than foliar biochemical concentrations using hyperspectral measurements (Townsend et al., 2003; Asner and Martin, 2008; Darvishzadeh et al., 2008b; Majeke et al., 2008) and this thesis (Chapter 5). Canopy biochemicals are often derived from a product of foliar biochemicals and canopy structural parameter, such as leaf area index (Townsend et al., 2003; Darvishzadeh et al., 2008b; He and Mui, 2010) or photosynthetic vegetation cover (Chapter 5). The photosynthetic cover is one of the canopy parameters which determine key ecosystem functions, e.g. rate of carbon and nutrient intake (Guerschman et al., 2009). Therefore, canopy N is highly estimated by the canopy image spectra, which largely dependent on the interaction between foliar N and photosynthetic vegetation cover. In addition, vegetation spectra result from a combination of physical components of the plant and structure (Wessman, 1992; Majeke et al., 2008). Poor results for foliar N as compared to canopy N estimation are consistent with other vegetation biochemicals studies including the ones focusing on foliar and canopy chlorophyll (Yoder and Pettigrew-Crosby, 1995; Asner, 1998; Asner and Martin, 2008; Darvishzadeh et al., 2008b; Majeke et al., 2008; He and Mui, 2010). These discrepancies of foliar and canopy biochemicals results were attributed to a poor propagation of light or signal from leaf to canopy (Yoder and Pettigrew-Crosby, 1995; Asner, 1998; Asner and Martin, 2008; Darvishzadeh et al., 2008b; Majeke et al., 2008; He and Mui, 2010).

In Chapter 5, 60% of the variation of grass N is explained using multispectral remote sensing data (i.e. RapidEye) which is comparable to some of the

hyperspectral studies (Mutanga and Skidmore, 2004a). The performance of the RapidEye data to estimate foliar and canopy N is associated with the presence of the red edge band centred at 710 nm. The hyperspectral studies demonstrated the use of red edge position to estimate foliar chlorophyll and N (Yoder and Pettigrew-Crosby, 1995; Cho and Skidmore, 2006; Darvishzadeh et al., 2008b). Foliar and canopy N are estimated through the positive correlation between chlorophyll and N (Yoder and Pettigrew-Crosby, 1995). Previous studies using hyperspectral data (in situ or with airborne) achieved high foliar N retrieval accuracies as compared to the results obtained in this study. At airborne hyperspectral remote sensing level, foliar N estimation was reported to explain accuracies between 48 to 80 % (Huang et al., 2004; Mutanga and Skidmore, 2004a; Skidmore et al., 2010; Knox et al., 2011). The results attained in Chapter 5 are within the latter range of explained variance of N, which shows progress in the estimation of biochemicals from local scale (field and airborne hyperspectral systems) to regional scale (new spaceborne multispectral systems).

Regional foliar and canopy N can be used for understanding feeding or foraging patterns of herbivores in protected and communal areas (Prins and Beekman, 1989; McNaughton, 1990; Duncan, 1992; Owen-Smith and Danckwerts, 1997; Drescher et al., 2006b). The foliar N concentration maps can be used to determine if foliar biochemical concentrations in a particular area are below minimum requirements for the herbivores (Grant and Scholes, 2006). For example, 1% of N is a minimum requirement of herbivores such as buffalo (Prins, 1987), and horses require a minimum of about 0.23 % of foliar P concentration (Duncan, 1992). The canopy N information could be used to determine how grass canopy structure influence foraging or feeding patterns of herbivores at landscape level (Drescher et al., 2006b). Drescher et al. (2006) found out that not only foliar biochemicals influence foraging or feeding patterns of herbivores but also grass structure plays a crucial role.

6.5 Phenology: determinant for foliar biochemicals estimation

The estimation of foliar biochemical concentrations is influenced by phenology, mainly linked to the leaf greenness and senesce condition, i.e. plant age or developmental stages (Wenjiang et al., 2004; Knox et al., 2010a). Wenjiang et al. (2004) reported that the regression coefficients vary depending on the developmental stages of the plant. The developmental stages of a plant are coupled with changes in cell structure, water content and functions, which then influence the reflectance, hence the absorption features of the foliar biochemicals over a period of plant development (Knox et al., 2010a). During wet season absorption features especially for N are in the visible/near infrared as well as the shortwave infrared (Ustin et al.,

2009), while during the dry season active N features are identifiable at the shortwave infrared region (Asner, 1998). Therefore, using vegetation indices, the relationship with foliar biochemicals deteriorates as a leaf senesces. Consequently, using vegetation indices only to estimate foliar biochemical concentrations might not be plausible, during dry season. Notwithstanding, vegetation indices play a crucial role in minimizing external factors such as wavelength numbers, illumination and atmospheric conditions (Jackson and Huete, 1991). Knox et al. (2010a) found out that foliar biochemical concentrations could be estimated during dry and wet season, using a combination of absorption features and environmental or ancillary variables. The challenge is that the variables for estimating foliar biochemical concentrations are not consistently selected in dry and wet season, using multivariate statistical techniques. Therefore, this warrants for the continuous use of integrated modeling, combining remote sensing (vegetation indices and absorption features) and environmental variables, in various seasons. This approach was successfully applied in wet season in Chapter 4 and 5. Nevertheless, seasonal estimation of biochemicals is also important for ecological studies focusing on seasonal migration of wild animals and livestock (McNaughton, 1990).

6.6 Ecological implications of the study

This study was based on the premises that foliar biochemical products derived from remote sensing can be used to understand and facilitate planning and management of the savanna rangelands. Foliar biochemicals in this case are indicators of grass quality. Grass quality indicators are not only the attractants such as protein; they also include deterrents such as polyphenols, which are more prominent in the trees than grasses. Traditional techniques for determining grass quality are point based, and do not cover larger areas (McNaughton, 1988). This study demonstrated that regional estimation of nutrients can be done using remote sensing. The foliar biochemical models, especially for foliar and canopy N developed in this study can be extended to other areas, such as the Great Limpopo Transfontier Park, provided there is RapidEye like remote sensing data, acquired during maximum biomass production and ancillary datasets. Nevertheless, several studies have demonstrated the importance of grass nutrients in understanding grazing behavior and migration patterns of wildlife and livestock in the savanna ecosystems (Drent and Prins, 1987; McNaughton, 1988, 1990; Seagle and McNaughton, 1992; Boone et al., 2008; Fryxell, 2008). Most of these studies were based on point observations, not covering spatial and continuous distribution of nutrients on the landscape. The products developed using remote sensing can cover this gap. For example, these products can be used as input in ecological studies or landscape models focusing on foraging behavior and population densities, forage selection, modeling animal movement (McNaughton, 1988; Prins and Beekman, 1989;

McNaughton, 1990; Manly et al., 1993; Owen-Smith and Danckwerts, 1997; Grant et al., 2000; Owen-Smith, 2005; Grant and Scholes, 2006; Boone et al., 2008; Fryxell, 2008). Remote sensing products show nutrients hotspots in a spatially explicit and continuous manner. The models can be developed to investigate how nutrient hotspots relate or influence distribution or feeding patterns of wild herbivores and livestock (Grant and Scholes, 2006). Nutrient hotspots can also help in modeling animal population dynamics (Owen-Smith, 2005; Grant and Scholes, 2006). For example, in the south of Kruger National Park, there is a concern of the decrease in population size of Sable antelope (Owen-Smith, personal communication). The cause of this is not known, but foliar biochemical analysis could help to understand whether changes in population size are associated with nutrient availability or limitations. Finally, foliar biochemical products can be used to understand feeding patterns along the fence-line contrast between protected and communal areas, for example, the area in and around Kruger National Park. This could be important input into for modeling probability of animal contact or risk models, i.e. livestock and wild animals, which can play a crucial role in understanding transmission of diseases e.g. tuberculosis and foot and mouth diseases (Bengis et al., 2003; Chaminuka et al., 2011). Spread of diseases between these animals could have high implications on livestock production and wild animal loss.

The societal relevance of foliar biochemical products relates to sustainable management of grazing lands or rangeland. In order to manage rangelands, one of the important parameters could be determining a state of the resource to be managed. After determining the state, then a decision on how to use the resource prevails, and thereafter, how can the resource exploitation or usage be managed. Southern African rural communities are challenged by the levels of land degradation taking place, which affect their grazing lands (Du Toit and Cumming, 1999). Central to this is the management of communal rangelands. The foliar biochemical products provide spatial information that could guide farmers on how to place their livestock in the rangelands. The foliar biochemical products are generally presented in a map, and can easily be understood by the farmers. Sustainably managed pastoral lands provide high livestock production returns (Klug et al., 2000).

Foliar biochemical products can also be used to understand the competing claims on natural resources and the extent to which they are occurring. Competing claims refers to conflicting entities like livestock, wildlife and people on natural resources. This situation normally occurs in areas where these three entities co-exist in the same ecosystem. To study and understand such kind of interactions in southern Africa, the Great Limpopo Transfontier Park (GLTP), which entails area of south-east lowveld (Gonarezhou area) in Zimbabwe and Limpopo National park in Mozambique can be an example

(Chaminuka, 2012). This is because people, livestock and wildlife live and interact together. Kruger National Park is a fenced part of GLTP, separating villages (people) and livestock from wild animals. People require fertile land for cultivation, wild animals require nutritious grazing land and finally livestock require a fertile land for grazing. In order to understand such a complex triangular situation, ecological models entailing issues like co-existence and competition can be used (Chaminuka et al.,2010; Chaminuka et al., 2011), which are going to be driven by resource quality. In this case, foliar biochemicals can be used as an indicator of soil fertility (Bell, 1986; Ben-Shahar and Coe, 1992). The competition arises because high nutrients areas have high densities of animals (Scholes, 1990; Owen-Smith and Danckwerts, 1997) and people (Barnes and Lahm, 1997). Studies indicated that if livestock, human activities (land use) and wild animals co-exist, livestock and human activities are likely to put wild animal in distress or even face total exclusion (Fritz et al., 2003). Therefore, understanding of the implications of competition and co-existence could be important in the planning and management of savanna ecosystems. This could be done to ensure that people their activities sustainably, minimizing land degradation and ensuring that livestock production does not compromise existence of wild animals.

6.7 Conclusions and future research

The key conclusions are;

- Water removal spectra can be used to minimize water absorption effects, enhancing the absorption features of foliar biochemicals, hence increasing the estimation accuracy, especially for foliar N and P concentrations (Chapter 2).
- Foliar N: P could be estimated using a combination of *in situ* hyperspectral remote sensing and partial least square regression. Water removed and continuum-removed spectra could be used for estimating foliar N: P at the high accuracy as compared to original reflectance and first derivative. SWIR bands are highly sensitive to the foliar N: P (Chapter 3).
- The integrated modeling approach, i.e. combining remote sensing and ancillary or environmental variables could be an alternative approach towards the estimation of foliar biochemicals at local and regional scale (Chapter 4 and 5).
- Regional scale estimation of foliar N is plausible, given the availability of red edge band embedded in the spaceborne sensor (e.g. RapidEye) (Chapter 5).
- The radial basis function-partial least square regression (RBF-PLSR) also referred to as non-linear PLSR could be used as alternative multivariate for estimating foliar biochemical concentrations, due to its combined capability of conventional PLSR and artificial neural network (Chapter 4).

- The biomass and foliar biochemicals interaction could be minimized by acquiring remote sensing and field data during maximum or peak productivity (Chapter 4 and 5).
- Finally, foliar biochemical maps thus mapped could enable the park manager, policy makers, farmers and ranchers to facilitate the decision making process concerning rangeland management, conservation planning and stocking or carrying capacity estimation.

This study recommends further research on testing the water removal technique at airborne level, using imagery such as HyMap with SWIR. Test the estimation of biochemical ratios such as N:P:K and carbon and N (C:N), which are also indicators of grass nutrient limitations, more especially C:N which also reflect the palatability of the grass. Seasonal monitoring of foliar biochemical concentrations using integrated modeling approach, which could be useful to understand the seasonal movements and population dynamics of the herbivores. The type of study could be done by using the GPS collars mounted on the animals and simultaneously acquiring remote sensing imagery, which could later be analyzed for landscape distribution of the foliar biochemical concentrations. The regional foliar biochemical maps developed from this study could be used for ecological studies at regional scales, i.e. to understand the densities of animals, both livestock and wildlife. The main hypothesis here could be to test if there is a significant relationship between foliar biochemicals and animal densities, or that herbivore normally concentrates on high nutritious grass area.

Bibliography

ABDEL-RAHMAN, E. M., AHMED, F. B. & VAN DEN BERG, M. 2010. Estimation of sugercane leaf nitrogen concentration using in situ spectroscopy. *International Journal of Applied Earth Observation and Geoinformation,* 12 (Suplement 1), S52-S57.

ABEL, N. O. J. & BLAIKIE, P. M. 1989. Land degradation, stocking rates and conservation policies in the communal rangelands of Botswana and Zimbabwe. *Land Degradation and Rehabilitation,* 1 (2), 101-123.

ADAMS, D. C. & CHURCH, J. O. 2007. Amphibians do not follow Bergmann's rule. *Evolution,* 62 (2), 413-420.

AGRICULTURAL RESEARCH COUNCIL 1980. *The Nutrient Requirements of Ruminant Livestock,* Slough, England, Commonwealth Agricultural Bureaux.

AN, H. & GU, L. 1989. Fast stepwise procedures of selection of variables by using AIC and BIC criteria. *Acta Mathematicae Applicatae Sinica (English Series),* 5 (1), 60-67.

ASNER, G. P. 1998. Biophysical and biochemical sources of variability in canopy reflectance. *Remote Sensing of Environment,* 64 (3), 234-253.

ASNER, G. P., JONES, M. O., MARTIN, R. E., KNAPP, D. E. & HUGHES, R. F. 2008. Remote sensing of native and invasive species in Hawaiian forests. *Remote Sensing of Environment,* 112 (5), 1912-1926.

ASNER, G. P. & MARTIN, R. E. 2008. Spectral and chemical analysis of tropical forests: Scaling from leaf to canopy levels. *Remote Sensing of Environment,* 112 (10), 3958-3970.

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.

ATKINSON, P. M. & TATANALL, A. R. L. 1997. Neural networks in remote sensing-Introduction. *International Journal of Remote Sensing,* 18, 699-709.

BARET, F. & GUYOT, G. 1991. Potentials and limits of vegetation indices for LAI and APAR assessment. *Remote Sensing of Environment,* 35 (2-3), 161-173.

BARNES, R. F. W. & LAHM, S. A. 1997. An ecological perspective on human densities in the Central African forests. *Journal of Applied Ecology,* 34, 245-260.

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.

BELL, R. H. V. 1986. Soil-Plant-Herbivore Interactions. *In:* BELL, R. H. V. & MCSHANE-CALUZI, E. (eds.) *Conservation and Wildlife Management in Africa.* US: US Peace Corps, Chapter 8, 107-130.

BEN-SHAHAR, R. & COE, M. J. 1992. The relationships between soil factors, grass nutrients and the foraging behaviour of wildebeest and zebra. *Oecologia,* 90 (3), 422-428.

BENGIS, R. G., GRANT, C. C. & DE VOS, V. 2003. Wildlife Diseases and Veterinary Controls: A Savanna Ecosystem Perspective. *In:* DU TOIT, J. T., ROGERS, K. H. & BIGGS, H. C. (eds.) *The Kruger Experience: Ecology and Management of Savanna Heterogeneity.* London: Island Press.

BOGREKCI, I. & LEE, W. S. 2005. Spectral phosphorus mapping using diffuse reflectance of soils and grass. *Biosystems Engineering,* 91 (3), 305-312.

BOONE, R. B., BURNSILVER, S. B., WORDEN, J. S., GALVIN, K. A. & HOBBS, N. T. 2008. Large-Scale Movement of Large Herbivores. *In:* PRINS, H. H. T. & VAN LANGEVELDE, F. (eds.) *Resource Ecology: Spatial and Temporal Dynamics of Foraging.* Dordrecht: Springer, 187-207.

BROGE, N. H. & LEBLANC, E. 2000. Comparing prediction power and stability of broadband and hyperspectral vegetation indices for estimation of green leaf area index and canopy chlorophyll density. *Remote Sensing of Environment,* 76 (2), 156-172.

BUNKE, O. & DROGE, B. 1984. Bootstrap and cross-validation estimates of the prediction error for linear regression models. *The Annals of Statistics,* 12 (4), 1400-1424.

BURKE, A. 1997. The impact of large herbivores on floral composition and vegetation structure in the Naukluft Mountains, Namibia. *Biodiversity and Conservation,* 6 (9), 1203-1217.

ÇAMDEVÝREN, H., DEMÝR, N., KANIK, A. & KESKÝN, S. 2005. Use of principal component scores in multiple linear regression models for prediction of Chlorophyll-*a* in reservoirs. *Ecological Modelling,* 181 (4), 581-589.

CARD, D. H., PETERSON, D. L., MATSON, P. A. & ABER, J. D. 1988. Prediction of leaf chemistry by the use of visible and near infrared reflectance spectroscopy. *Remote Sensing of Environment,* 26 (2), 123-147.

CECH, P., KUSTER, T., EDWARDS, P. & OLDE VENTERINK, H. 2008. Effects of herbivory, fire and N_2-fixation on nutrient limitation in a humid African savanna. *Ecosystems,* 11 (6), 991-1004.

CHAMINUKA, P. 2012. *Evaluating land use options at wildlife/livestock interface: An integrated spatial landuse analysis.* PhD Thesis, Wageningen University.

CHAMINUKA, P., MCCRINDLE, C. M. E. & UDO, H. 2011. Cattle farming at the wildlife/livestock interface: assessment of costs and benefits Adjacent to Kruger National Park, South Africa. *Society and Natural Resources,* DOI:10.1080/08941920.2011.5804171-16.

CHAMINUKA, P., UDO, H., EILERS, K. & VAN DER ZIJPP, A. 2010. Livestock systems and competing claims for land at the wildlife-based tourism/livestock interface. Applied Animal Husbandry and Rural Development Journal, 1 (3), 5-9.

CHO, M. & SKIDMORE, A. K. 2009. Hyperspectral predictor for monitoring biomass production in Mediterranean mountain grasslands:Majella National Park, Italy. *International Journal of Remote Sensing,* 30 (2), 499-515.

CHO, M., SKIDMORE, A. K., CORSI, F., VAN WIEREN, S. & SOBHAN, I. 2007a. Estimation of green grass/herb biomass from airborne hyperspectral imagery using spectral indices and partial least square regressions. *International Journal of Applied Earth Observation and Geoinformation,* 9 (4), 414-424.

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., SKIDMORE, A. K. P. & VAN WIEREN, S. E. P. 2007b. *Hyperspectral remote sensing of biochemical and biophysical parameters : the derivative red - edge "double - peak feature" : a nuisance or an opportunity?* , Wageningen University.

CHO, M. A., VAN AARDT, J., MAIN, R., MAJEKE, B., RAMOELO, A., MATHIEU, R., NORRIS-ROGERS, M. & DU PLESSIS, R. 2009. Integrating remote sensing and ancillary data for regional ecosystem assessment: Eucalyptus grandis agrosystem in Kwazulu Natal, South Africa. *IEEE 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. N. & ROUSH, T. L. 1984. Reflectance spectroscopy: quantitative analysis techniques for remote sensing applications. *journal of Geophysical Research,* 89 (B7), 6329-6340.

CLEVERS, J. G. P. W., 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.

CRAINE, J. M., MORROW, C. & STOCK, W. D. 2008. Nutrient concentration ratios and co-limitation in South African grasslands. *New Phytologist,* 179 (3), 829-836.

CRAWLEY, N. 2006. *Statistical Computing: An introduction to Data Analysis using S-Plus,* London, John Wiley & Sons.

CURRAN, P. J. 1989. Remote sensing of foliar chemistry. *Remote Sensing of Environment,* 30 (3), 271-278.

CURRAN, P. J., DUNGAN, J. L. & PETERSON, D. L. 2001. Estimating the foliar biochemical concentration of leaves with reflectance spectrometry: Testing the Kokaly and Clark methodologies. *Remote Sensing of Environment,* 76 (3), 349-359.

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

DARVISHZADEH, R., SKIDMORE, A., SCHLERF, M., ATZBERGER, C., CORSI, F. & CHO, M. 2008b. LAI and chlorophyll estimation for a heterogeneous grassland using hyperspectral measurements. *ISPRS Journal of Photogrammetry and Remote Sensing,* 63 (4), 409-426.

DASH, J. & CURRAN, P. J. 2004. The MERIS terrestrial chlorophyll index. *International Journal of Remote Sensing,* 25 (23), 5403 - 5413.

DASZYKOWSKI, M., SERNEELS, S., KACZMAREK, K., VAN ESPEN, P., CROUX, C. & WALCZAK, B. 2007. TOMCAT: A MATLAB toolbox for multivariate calibration techniques. *Chemometrics and Intelligent Laboratory Systems,* 85 (2), 269-277.

DAUFRESNE, T. & LOREAU, M. 2001. Plant–herbivore interactions and ecological stoichiometry: when do herbivores determine plant nutrient limitation? *Ecology Letters,* 4 (3), 196-206.

DAUGHTRY, C. S. T., WALTHALL, C. L., KIM, M. S., DE COLSTOUN, E. B. & MCMURTREY, J. E. 2000. Estimating corn Leaf chlorophyll concentration from leaf and canopy reflectance. *Remote Sensing of Environment,* 74 (2), 229-239.

DAWSON, T. P. & CURRAN, P. J. 1998. Technical note a new technique for interpolating the reflectance red edge position. *International Journal of Remote Sensing,* 19 (11), 2133 - 2139.

DAWSON, T. P., CURRAN, P. J. & PLUMMER, S. E. 1998. LIBERTY--Modeling the effects of leaf biochemical concentration on reflectance spectra. *Remote Sensing of Environment,* 65 (1), 50-60.

DIJKSHOORN, K. 2003. SOTER database for Southern Africa (SOTERSAF): Technical Report. Wageningen: Internatioal Institute for Soil Reference and Information Centre.

DIJKSHOORN, K., VAN ENGELEN, V. & HUTING, J. 2008. Global Assessment of Land Degradation: Soil and Landform propertes for Land Degradation Assessment in Drylands (LADA) partner countries (Argentina, China, Cuba, Senegal and the Gambia, South Africa and Tunisia). *ISRIC Report 2008/06 and GLADA Report 2008/03.*

Wageningen: ISRIC-Word Soil Information ; FAO-Food and Agriculture Organization of the United Nations.

DRENT, R. H. & PRINS, H. H. T. 1987. The herbivore as prisoner of its food supply. *In:* ANDEL, J. V. & BAKKER, J. (eds.) *Disturbance in Grasslands: Species and Population Responses.* Dordrecht: Dr. W. Junk Publishing Company.

DRESCHER, M., HEITKÖNIG, I. M. A., RAATS, J. G. & PRINS, H. H. T. 2006a. The role of grass stems as structural foraging deterrents and their effects on the foraging behaviour of cattle. *Applied Animal Behaviour Science,* 101 (1-2), 10-26.

DRESCHER, M., HEITKÖNIG, I. M. A., VAN DEN BRINK, P. J. & PRINS, H. H. T. 2006b. Effects of sward structure on herbivore foraging behaviour in a South African savanna: An investigation of the forage maturation hypothesis. *Austral Ecology,* 31 (1), 76-87.

DU TOIT, J. T. & CUMMING, D. H. M. 1999. Functional significance of ungulate diversity in African savannas and the ecological implications of the spread of pastoralism. *Biodiversity and Conservation,* 8, 1643-1661.

DUNCAN, P. B. 1992. *Horses and grasses: the nutritional ecology of equids and their impact on the Camargue,* NY, Springer-Verlag.

EFRON, B. 1983. Estimating the error rate of a prediction rule: improvement on cross-validation. *Journal of the American Statistical Association,* 78 (382), 316-331.

EFRON, B. & TIBSHIRANI, R. 1997. Improvements on cross-validation: The .632+ Bootstrap Method. *Journal of the American Statistical Association,* 92 (438), 548-560.

EHSANI, M. R., UPADHYAYA, S. K., SLAUGHTER, D., SHAFII, S. & PELLETIER, M. 1999. A NIR technique for rapid determination of soil mineral nitrogen. *Precision Agriculture,* 1, 217-234.

ELSER, J. J., DOBBERFUHL, D. R., MACKAY, N. A. & SCHAMPEL, J. H. 1996. Organism size, life history, and N:P stoichiometry. *BioScience,* 46 (9), 674-684.

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

ELVIDGE, C. D. & CHEN, Z. 1995. Comparison of broad-band and narrow band red and near-infrared vegetation indices. *Remote Sensing of Environment,* 54 (1), 38-48.

EVANS, J. R. 1989. Photosynthesis and nitrogen relationships in leaves of C_3 plants. *Oecologia,* 78 (1), 9-19.

EVERSON, T. M. & HATCH, G. P. 1999. Managing Veld (Rangelands) in the Communal Areas of Southern Africa. *In:* TAINTON, N. M. (ed.) *Veld Management in South Africa.* Pietermaritzburg: University of Natal Press.

FAO, ISRIC & UNEP 2003. Soil and Terrain Database for Southern Africa. *In:* (FAO), F. A. A. O. (ed.). Rome: Http://www.fao.org/ag/agl/lwdms.stm ; Http://isric.org/checkregistration.aspx?dataset=27.

FERWERDA, J. G., SIDERIUS, W., VAN WIEREN, S. E., GRANT, C. C., PEEL, M., SKIDMORE, A. K. & PRINS, H. H. T. 2006. Parent material and fire as principle drivers of foliage quality in woody plants. *Forest Ecology and Management,* 231 (1-3), 178-183.

FERWERDA, J. G., SKIDMORE, A. K. & MUTANGA, O. 2005. Nitrogen detection with hyperspectral normalized ratio indices across multiple plant species. *International Journal of Remote Sensing,* 26 (18), 4083 - 4095.

FEWSTER, R. M., LAAKE, J. L. & BUCKLAND, S. T. 2005. Line transect in small and large regions. *Biometrics,* 61, 856-859.

FIDÊNCIO, P. H., POPPI, R. J. & DE ANDRADE, J. C. 2002. Determination of organic matter in soils using radial basis function networks and near infrared spectroscopy. *Analytica Chimica Acta,* 453 (1), 125-134.

FIELD, C. & MOONEY, H. A. 1986. The Photosynthesis-Nitrogen Relationship in Wild Plants. *In:* GIVNISH, T. J. (ed.) *On the Economy of Plant Form and Function.* Cambridge, Cambridge Univesity Express (pp.25-55).

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

FOX, J. 2002. Bootstrapping regression models. *Appendix to An R and S-PLUS Companion to Applied Regression,* http://cran.r-project.org/doc/contrib/Fox-Companion/appendix-bootstrapping.pdf , Accessed January 2011.

FOX, J. & WEISBERG, S. 2010. Bootstrapping Regression Models in R. *An Appendix to An R Companion to Applied Regression, Second Edition,* http://socserv.mcmaster.ca/jfox/Books/Companion/appendix/Appendix-Bootstrapping.pdf , Accessed May 2011.

FRANKLIN, J., LOGAN, T. L. & WOODCOCK, C. E. 1986. Coniferous forest classification and inventory using Landsat and digital terrain data. *IEEE Transactions on Geoscience and Remote Sensing,* GE-24 (1), 139-149.

FRITZ, H., SAID, S., RENAUD, P., MUTAKE, S., COID, C. & MONICAT, F. 2003. The effects of agricultural fields and human settlements on the use of rivers by wildlife in the mid-Zambezi valley, Zimbabwe. *Landscape Ecology,* 18, 293-302.

FRYXELL, J. M. 2008. Predictive Modelling of Patch Use by Terrestrial Herbivores. *In:* PRINS, H. H. T. & VAN LANGEVELDE, F. (eds.) *Resouce Ecology: Spatial and Temporal Dynamics of Foraging.* Dordrecht: Springer, 105-204.

GAO, B.-C. & GOETZ, A. F. H. 1994. Extraction of dry leaf spectral features from reflectance spectra of 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 AVIRIS data. *Remote Sensing of Environment,* 52 (3), 155-162.

GARG, S., PATRA, K., KHETRAPAL, V., PAL, S. K. & CHAKRABORTY, D. 2010. Genetically evolved radial basis function network based prediction of drill flank wear. *Engineering Applications of Artificial Intelligence,* 23 (7), 1112-1120.

GATES, D. M., KEEGAN, H. J., SCHLETER, J. C. & WEIDNER, V. R. 1965. Spectral properties of plants. *Applied Optics,* 4 (1), 11-20.

GELADI, P., HADJIISKI, L. & HOPKE, P. 1999. Multiple regression for environmental data: nonlinearities and prediction bias. *Chemometrics and Intelligent Laboratory Systems,* 47 (2), 165-173.

GELADI, P. & KOWALSKI, B. R. 1986. Partial least-squares regression: a tutorial. *Analytica Chimica Acta,* 185, 1-17.

GIERING, R., KAMINSKI, T. & SLAWIG, T. 2005. Generating efficient derivative code with TAF: adjoint and tangent linear euler flow around an airfoil. *Future Generation Computer System,* 21 (8), 1345.

GIRON, H. C. 1973. Comparison between dry ashing and wet digestion in preparation of plant material for atomic absorption analysis. *Atomic Absorption Newsletter,* 12 (1), 28-29.

GITELSON, A. A., KAUFMAN, Y. J. & MERZLYAK, M. N. 1996. Use of a green channel in remote sensing of global vegetation from EOS-MODIS. *Remote Sensing of Environment,* 58 (3), 289-298.

GONG, P., PU, R. & HEALD, R. C. 2002. Analysis of in situ hyperspectral data for nutrient estimation of giant sequoia. *International Journal of Remote Sensing,* 23 (9), 1827 - 1850.

GRANT, C. C., PEEL, M., ZAMBATIS, N. & VAN RYSSEN, J. B. J. 2000. Nitrogen and phosphorus concentration in faeces: an indicator of range quality as a practical adjunct to existing range evaluation methods. *African Journal of Range and Forage Science,* 17, 81-92.

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.

GRASSHOFF, K., ERHARDT, M. & KREMLING, K. (eds.) 1983. *Methods of Seawater Analysis,* Weinhein, Germany: Verlag Cheime.

GROSSMAN, Y. L., USTIN, S. L., JACQUEMOUD, S., SANDERSON, E. W., SCHMUCK, G. & VERDEBOUT, J. 1996. Critique of stepwise multiple linear regression for the extraction of leaf biochemistry information from leaf reflectance data. *Remote Sensing of Environment,* 56 (3), 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 savanna region upscaling the EO-1 Hyperion and MODIS sensors. *Remote Sensing of Environment,* 113 (5), 928-945.

GÜSEWELL, S. 2004. N : P ratios in terrestrial plants: variation and functional significance. *New Phytologist,* 164 (2), 243-266.

GÜSEWELL, S., KOERSELMAN, W. & VERHOEVEN, J. T. A. 2003. Biomass N:P Ratios as Indicators of Nutrient Limitation for Plant Populations in Wetlands. *Ecological Applications,* 13 (2), 372-384.

HABOUDANE, D., JOHN, R., MILLERA, J. R., TREMBLAY, N., ZARCO-TEJADA, 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 (2-3), 416-426.

HABOUDANE, D., MILLER, J. R., PATTEY, E., ZARCO-TEJADA, P. J. & STRACHAN, I. B. 2004. Hyperspectral vegetation indices and novel algorithms for predicting green LAI of crop canopies: Modeling and validation in the context of precision agriculture. *Remote Sensing of Environment,* 90 (3), 337-352.

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.

HE, Y. & MUI, A. 2010. Scaling up Semi-Arid Grassland Biochemical Content from the Leaf to the Canopy Level: Challenges and Opportunities. *Sensors,* 10 (12), 11072-11087.

HEITKÖNIG, I. M. A. & OWEN-SMITH, N. 1998. Seasonal selection of soil types and grass sward by roan antelope in a South African savanna. *African Journal of Ecology,* 36 (1), 57-70.

HELDT, H. W., CHON, C. J., MARONDE, D., HEROLD, A., STANKOVIC, Z. S., WALKER, D. A., KRAMINER, A., KIRK, M. R. & HEBER, U. 1977. Role of orthophosphate and other factors in the regulation of starch formation in leaves and isolated chloroplasts. *Journal of Plant Physiology,* 59 (6), 1146-1155.

HIJMANS, R. J., CAMERON, S. E., PARRA, J. L., JONES, P. G. & JARVIS, A. 2005. Very High Resolution Interpolated Climate Surfaces for Global Land Areas. *International Journal of Climatology,* 25, 1965-1978.

HIJMANS, R. J., GUARINO, L., CRUZ, M. & ROJAS, E. 2001. Computer tools fro spatial anaylsis of plant genetic resources data: 1. DIVA-GIS. *Plant Genetic Resources Newletter,* 2001 (127), 15-19.

HOFFER, R. M. 1975. Natural resource mapping in mountanious terrain by computer analysis of ERTS-1 satellite data. *LARS Research Bulletin* 919, Perdue University.

HOLLANDER, M. & WOLFE, D. A. 1973. *Nonparametric Statistical Methods,* New York, John Wiley & Sons.

HORLER, D. N. H., DOCKRAY, M. & BARBER, J. 1983. The red egde of plant leaf reflectance. *International Journal of Remote Sensing,* 4 (2), 273-288.

HRUSCHKA, W. R. 1987. Data Analysis: Wavelength Selection Methods. . *In:* WILLIAM, P. & NORRIS, K. (eds.) *Near Infrared Technology in the Agricultural and Food Industries.* St Paul, MN: American Association of Cereal Chemists.

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.

HUETE, A. R. 1988. A soil-adjusted vegetation index (SAVI). *Remote Sensing of Environment,* 25, 295-309.

HUETE, A. R., LIU, H., Q., BATCHILY, K. & VAN LEEWEN, W. 1997. A comparison of vegetation indices global set pf TM images for EOS-MODIS. *Remote Sensing of Environment,* 59 (3), 440-451.

JACKSON, R. D. & HUETE, A. R. 1991. Interpreting vegetation indices. *Preventive Veterinary Medicine,* 11, 185-200.

JACQUEMOUD, S., USTIN, S. L., VERDEBOUT, J., SCHMUCK, G., ANDREOLI, G. & HOSGOOD, B. 1996. Estimating leaf biochemistry using the PROSPECT leaf optical properties model. *Remote Sensing of Environment,* 56 (3), 194-202.

JAIN, N., RAY, S., SINGH, J. & PANIGRAHY, S. 2007. Use of hyperspectral data to assess the effects of different nitrogen applications on a potato crop. *Precision Agriculture,* 8 (4), 225-239.

JAMES, L. F., YOUNG, J. A. & SANDERS, K. 2003. A new approach to monitoring rangelands. *Arid Land Research and Management,* 17 (4), 319-318.

JAVIS, A., REUTER, H. I., NELSON, A. & GUEVARA, E. 2008. The Holle-filled SRTM for the globe Version 4. *Available from the CGIAR-CSI SRTM 90 m database (http://srtm.csi.cgiar.org)*.

JORDAN, C. F. 1969. Derivation of leaf area index from quality of light on the floor. *Ecology,* 50, 663-666.

JOHNSON, L. F. 2001. Nitrogen influence on fresh-leaf NIR spectra. *Remote Sensing of Environment,* 78 (3), 314-320.

KLUG, J. R., VAN DER HEERDEN, J. M. & LISHMAN, A. W. 2000. Fodder Production Planning and Livestock Production Systems. *In:* TAINTON, N. M. (ed.) *Pasture Management in South Africa.* Pietermaritzburg: University of Natal Press.

KNOX, N. M., SKIDMORE, A. K. & PRINS, H. H. T. 2010a. *Observing temporal and spatial variability of forage quality.* 174, University of Twente Faculty of Geo-Information and Earth Observation ITC.

KNOX, N. M., SKIDMORE, A. K., PRINS, H. H. T., ASNER, G. P., VAN DER WERFF, 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 savanna forage quality, using the hyperspectral Carnegie Airborne Observatory sensor. *Remote Sensing of Environment,* 115 (6), 1478-1488.

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. 2010b. Nitrogen prediction in grasses: effect of bandwidth and plant material state on absorption feature selection. *International Journal of Remote Sensing,* 31 (3), 691-704.

KOERSELMAN, W. & MEULEMAN, A. F. M. 1996. The Vegetation N:P ratio: a new tool to detect the nature of nutrient limitation. *Journal of Applied Ecology,* 33 (6), 1441-1450.

KOKALY, R. F. 2001. Investigating a physical basis for spectroscopic estimates of leaf nitrogen concentration. *Remote Sensing of Environment,* 75 (2), 153-161.

KOKALY, R. F., ASNER, G. P., OLLINGER, S. V., MARTIN, M. E. & WESSMAN, C. A. 2009. Characterizing canopy biochemistry from imaging spectroscopy and its application to ecosystem studies. *Remote Sensing of Environment,* 113 (Supplement 1), S78-S91.

KOKALY, R. F. & CLARK, R. N. 1999. Spectroscopic determination of leaf biochemistry using band-depth analysis of absorption features and stepwise multiple linear regression. *Remote Sensing of Environment,* 67 (3), 267-287.

KOSKI, M. 1999. Carbon: nitrogen ratios of Baltic Sea copepods-indication of nutrient limitation. *Journal of Plankton Research,* 12, 1239-1249.

KUMAR, L., SCHMIDT, K. S., DURY, S. & SKIDMORE, A. K. 2001. Imaging Spectroscopy and Vegetation Science. *In:* VAN DER MEER, F. D. & DE JONG, S. M. (eds.) *Image Spectroscopy.* Dordrecht: Kluver Academc Publishers.

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.

LARCHER, W. 1980. *Physiological Plant Ecology,* Berlin and New York, Springer-Verlag.

LE MAIRE, G., FRANCOIS, C., SOUDANI, K., BERVEILLER, D., PONTAILLER, J. Y., BREDA, N., GENET, H., DAVI, H. & DUFRENE, E. 2008. Calibration and validation of hyperspectral indices for the estimation of broadleaved forest leaf chlorophyll content, leaf mass per area, leaf area index and leaf canopy biomass. *Remote Sensing of Environment,* 112 (10), 3846-3864.

LEHMAN, E. 1998. *Non-parametrics: Statistical Methods Based on Ranks,* Upper Saddle River, Prentice-Hall.

LUDWIG, F., DE KROON, H. & PRINS, H. H. T. 2008. Impacts of savanna trees on forage quality for large African herbivore. *Oecologia,* 155, 487-496.

LUDWIG, F., DE KROON, H., PRINS, H. H. T. & BERENDSE, F. 2001. Effects of nutrients and shade on tree-grass interactions in an east African savanna. *Journal of Vegetation Science,* 12 (4), 579-588.

MAIN, R., CHO, M. A., MATHIEU, R., O'KENNEDY, M. M., RAMOELO, A. & KOCH, S. 2011. An investigation into robust spectral indices for leaf chlorophyll estimation. *ISPRS Journal of Photogrammetry and Remote Sensing,* 66 (6), 751-761.

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.

MANLY, B., MCDONALD, L. & THOMAS, D. 1993. *Resource Selection by Animals: Statistical Design and Analysis for Field Studies,* Madras, Chapman and Hall.

MARSCHNER, H. 1995. *Mineral Nutrition of Higher Plants, 2nd ed,* London, Academic Press.

MARTENS, H. & NAES, T. 2001. Multivariate calibration by data compression. *In:* WILLIAMS, P. & NORRIS, K. (eds.) *Near Infrared Technology in the Agricultural and Food Industries.* 2nd ed. Minessota, USA: American Association of Cereal Chemists.

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.

MATHEWS & FINK 2004. *Numerical Methods using Matlab,* New Jersey, USA, Prentice-Hall Inc.

MATHWORKS 2009. *Matlab R200a: The Language of Technical Computing,* USA, The Mathworks, Inc.

MCNAUGHTON, S. J. 1988. Mineral nutrition and spatial concentrations of African ungulates. *Nature,* 334, 343-345.

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

MCNAUGHTON, S. J. & BANYIKA, F. F. (eds.) 1995. *Plant Communities and Herbivory,* Chicago: Chicago Press.

MCNAUGHTON, S. J. & BANYIKWA, F. F. 1995. Plant communities and Herbivory. *In:* SINCLAIR, A. R. E. & ARCESE, P. (eds.) *Serengeti II-dynamics, management, and conservation of an ecosystem.* Chicago: Chicago Press.

MEISSNER, H. H., ZACHARIAS, P. J. K. & O'REAGAIN, P. J. 1999. Forage Quality (Feed Value). *In:* TAINTON, N. M. (ed.) *Veld Management in South Africa.* Pietermaritzburg: University of Natal Press.

METTERNICHT, G. 2003. Vegetation indices derived from high-resolution airborne videography for precision crop management. *International Journal of Remote Sensing,* 24 (14), 2855-2877.

MOONEY, H. A. (ed.) 1986. *Photosynthesis. Plant ecology,* Oxford, UK: Blackwell Scientific Publication.

MORRISON, M. L. 2001. *Wildlife study design,* Berlin etc., Springer.

MUCINA, L. & RUTHERFORD, M. C. 2006. *The Vegetation of South Africa, Lesotho and Swaziland,* Cape Town, Strelitzia.

MUMBY, P. J., GREEN, E. P., EDWARDS, A. J. & CLARK, C. D. 1999. The cost-effectiveness of remote sensing for tropical coastal resources assessment and management. *Journal of Environmental Management,* 55 (3), 157-166.

MUTANGA, O. & KUMAR, L. 2007. Estimating and mapping grass phosphorus concentration in an African savanna using hyperspectral image data. *International Journal of Remote Sensing,* 28 (21), 4897 - 4911.

MUTANGA, O., PRINS, H. H. T., SKIDMORE, A. K., VAN WIEREN, S., HUIZING, H., GRANT, R., PEEL, M. & BIGGS, H. 2004a. Explaining grass-nutrient patterns in a savanna rangeland of southern Africa. *Journal of Biogeography,* 31 (5), 819-829.

MUTANGA, O. & SKIDMORE, A. K. 2004a. Integrating imaging spectroscopy and neural networks to map grass quality in the Kruger National Park, South Africa. *Remote Sensing of Environment,* 90 (1), 104-115.

MUTANGA, O. & SKIDMORE, A. K. 2004b. Narrow band vegetation indices overcome the saturation problem in biomass estimation. *International Journal of Remote Sensing,* 25 (19), 3999 - 4014.

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

MUTANGA, O., SKIDMORE, A. K., KUMAR, L. & FERWERDA, J. 2005. Estimating tropical pasture quality at canopy level using band depth analysis with continuum removal in the visible domain. *International Journal of Remote Sensing,* 26 (6), 1093 - 1108.

MUTANGA, O., SKIDMORE, A. K. & PRINS, H. H. T. 2004b. Discriminating sodium concentration in a mixed grass species environment of the Kruger National Park using field spectrometry. *International Journal of Remote Sensing,* 25 (20), 4191 - 4201.

MUTANGA, O., SKIDMORE, A. K. & PRINS, H. H. T. 2004c. 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., SKIDMORE, A. K. & VAN WIEREN, S. E. 2003. Descriminating tropical grass (*Cenchrus ciliaris*) canopies grown under different nitrogen treatments using spectroradiometry. *ISPRS Journal of Photogrammetry and Remote Sensing,* 57 (3), 263-272.

NAES, T., IRGENS, C. & MARTENS, H. 1986. Comparison of linear statistical methods for calibration of NIR instruments. *Journal of the Royal Statistical Society. Series C (Applied Statistics),* 35 (2), 195-206.

NUMATA, I., ROBERTS, D. A., CHADWICK, O. A., SCHIMEL, J. P., GALVAO, L. S. & SOARES, J. V. 2008. Evaluation of hyperspectral data for pasture estimate in the Brazilian Amazon using field and imaging spectrometers. *Remote Sensing of Environment,* 112 (4), 1569-1583.

OKITA, T. W. 1992. Is there an alternative pathway for starch synthesis? *Journal of Plant Physiology,* 100 (2), 560-564.

OLFF, H., RITCHIE, M. E. & PRINS, H. H. T. 2002. Global environmental controls of diversity in large herbivores. *Nature,* 415 (6874), 901-904.

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 (2), 339-355.

OWEN-SMITH, N. 2005. *Adaptive Herbivore Ecology : from Resources to Populations in Variable Environments,* Johannesburg, Wits University Press.

OWEN-SMITH, N. & COOPER, S. M. 1987. Palatabilty of woody plants to browsing ruminants in a South African savanna. *In: Ecology, 68(1987)2, pp. 319-331*.

OWEN-SMITH, N. & DANCKWERTS, J., E. (eds.) 1997. *Herbivory,* Cambridge: Cambridge University Press.

PENNING DE VRIES, F. W. T. & DJITEYE, M. A. 1982. *La productivité des paturages sahéliens. Une étude des sols, des végétations et de l'exploitation de cette ressource naturelle.* Pudoc, Pudoc, Wageningen.

PEÑUELAS, J., FILLELA, I., LLOLET, P., MUNOZ, F. & VILAJELIU, M. 1995. Reflectance assessment of mite effects on apple trees. *International Journal of Remote Sensing,* 16 (14), 2727-2733.

PICKETT, S. T. A., GADENASSO, M. L. & BENNING, T. L. (eds.) 2003. *Biotic and Abiotic Variability as Key Determinants of Savanna Heterogeneity at Spatiotemporal Scales,* London: Island Press.

PLUMMER, S. E. 1988a. Exploring the relationships between leaf nitrogen content, biomass and the near-infrared/red reflectance ratio. *International Journal of Remote Sensing,* 9 (1), 177-183.

PLUMMER, S. E. 1988b. Relationships Between The Nitrogen Content Of Grass And Reflectance. *Geoscience and Remote Sensing Symposium.,* IGARSS '88. Remote Sensing: Moving Toward the 21st Century., International, 265-267.

PRINS, H. H. T. 1987. *The buffalo of Manyara: the individual in the context of herd life in a seasonal environment of East Africa,* PhD thesis, University of Groningen.

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:* PRINS, H. H. T. & VAN LANGEVELDE, F. (eds.) *Resource Ecology: Spatial and Temporal Dynamics of Foraging.* Netherlands: Springer.

QI, J., CHEHBOUNI, A., HUETE, A. R., KERR, Y. H. & SOROOSHIAN, S. 1994. A modified soil adjusted vegetation index. *Remote Sensing of Environment,* 48 (2), 119-126.

RAMOELO, A., CHO, M. A., MATHIEU, R., SKIDMORE, A. K., SCHLERF, M., HEITKÖNIG, I. M. A. & PRINS, H. H. T. 2011a. Integrating environmental and in situ hyperspectral remote sensing variables for grass nitrogen estimation in savanna ecosystems. *34th International Symposium on the Remote Sensing of Environment (ISRSE 2011), The GEOSS Era: Towards Operational Environmental Monitoring.* Sydney, Australia, http://www.isprs.org/proceedings/2011/ISRSE-34/index.html , Accessed: August 2011.

RAMOELO, A., SKIDMORE, A. K., CHO, M. A., MATHIEU, R., HEITKÖNIG, I. M. A., DUDENI-THLONE, N., SCHLERF, M. & PRINS, H. H. T. *under review*. 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*.

RAMOELO, A., SKIDMORE, A. K., SCHLERF, M., MATHIEU, R. & HEITKÖNIG, I. M. A. 2011b. Water-removed spectra increase the retrieval accuracy when estimating savanna grass nitrogen and phosphorus concentrations. *ISPRS Journal of Photogrammetry and Remote Sensing,* 66 (4), 408-417.

RAPIDEYE 2010. RapidEye Standard Image Product Specification, Version 3.0, Germany, www.rapideye.de accessed on April 2010.

RICHTER, R. 2011. *Atmospheric/Topographic fro Satellite Imagery (ATCOR 2/3 User Guide, Version 8),* Wessling, Germany, DLR-German Aerospace Center.

RIVERO, R. G., GRUNWALD, S., BINFORD, M. W. & OSBORNE, T. Z. 2009. Integrating spectral indices into prediction models of soil phosphorus in a subtropical wetland. *Remote Sensing of Environment,* 113 (11), 2389-2402.

ROBERTS, B. R. 1987. The Availability of Herbage. *In:* HARKER, J. B. & TERNOUTH, J. H. (eds.) *The Nutrition of Herbivores.* London: Academic Press.

RONDEAUX, G., STEVEN, M. & BARET, F. 1996. Optimized of soil-adjusted vegetation indices. *Remote Sensing of Environment,* 55 (2), 95-107.

ROUJEAN, J.-L. & BREON, F.-M. 1995. Estimating PAR absorbed by vegetation from bidirectional reflectance measurements. *Remote Sensing of Environment,* 51 (3), 375-384.

ROUSE, J. W., HAAS, R. H., SCHELL, J. A., DEERING, D. W. & HARLAN, J. C. 1974. Monitoring the vernal advancement and retrogradation of natural vegetation. *NASA/GSFC, Type III Final Report.* M.D. Greenbelt, 371.

ROYSTON, P. 1982. Algorithm AS 181: the *W* test for normality. *Applied Statistics,* 31 (176-180).

SAAD, A., ADAMS, D. C. & WICKNICK, J. A. 2007. Bioclimatic modelling, morphology, and behaviour reveal alternative mechanisms regulating the distribution of two parapatric salamander species. *Evolutionary Ecology Research,* 9, 843-854.

SAKAMOTO, Y., ISHIGURO, M. & KITAGAWA, G. 1986. *Akaike Information Criteria Statistics,* US, D. Reidel Publishing Company.

SAVA STANKOVIC, Z. 1978. Starch formation in isolated pea chloroplasts. *Plant Science Letters,* 12 (3-4), 371-377.

SAVITZKY, A. & GOLAY, M. J. E. 1964. Smoothing and differentiation of data by simplified least-squares procedures. *Analytical Chemistry,* 36 (8), 1627−1639.

SCHLERF, M. & ATZBERGER, C. 2006. Inversion of a forest reflectance model to estimate structural canopy variables from hyperspectral remote sensing data. *Remote Sensing of Environment,* 100 (3), 281-294.

SCHLERF, M., ATZBERGER, C., HILL, J., BUDDENBAUM, H., WERNER, W. & SCHÜLER, G. 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.

SCHLEICHER, T. D., BAUSCH, W. C., DELGADO, J. A. & AYERS, P. D. 2001. Evaluation and refinement of the nitrogen reflectance index (NRI) for site-specific fertilizer management. *Paper number 011151, ASAE Annual Meeting, St. Joseph, Michigan.*

SCHMIDT, K. S. & SKIDMORE, A. K. 2001. Exploring spectral discrimination of grass species in African rangelands. *International Journal of Remote Sensing,* 22 (17), 3421-3434.

SCHMIDT, K. S., SKIDMORE, A. K., KLOOSTERMAN, E. H., VAN OOSTEN, H., KUMAR, L. & JANSSEN, J. A. M. 2004. Mapping coastal vegetation using an expert system and hyperspectral imagery. *ISPRS Journal of Photogrammetric Engineering and Remote Sensing,* 70 (6), 703-715.

SCHOLES, R. J. 1990. The influence of soil fertility on the ecology of Southern African dry savannas. *Journal of Biogeography,* 17, 415-419.

SCHOLES, R. J., BOND, W. J. & ECKHARDT, H. C. 2003. Vegetation Dynamics in the Kruger Ecosystem. *In:* DU TOIT, J. T., ROGERS, K. H. & BIGGS, H. C. (eds.) *The Kruger Experience : Ecology and Management of Savanna Heterogeneity.* London: Island Press.

SCHOLES, R.J., & WALKER, B.H. 2004. An African Savanna: Synthesis of the Nylsvley Study, Cambridge University Press, Cambridge

SEAGLE, S. W. & MCNAUGHTON, S. J. 1992. Spatial variation in forage nutrient concentrations and the distribution of serengeti ungulates. *Landscape Ecology,* 7 (4), 229-241.

SHACKLETON, S. E., SHACKLETON, C. M., NETSHILUVHI, P. R., GEACH, B. S., BALLANCE, A. & FAIRBANKS, D. H. K. 2002. Use patterns and value of savanna resources in three rural villages in South Africa. *Economic Botany,* 56 (2), 130-146.

SKIDMORE, A. K. 1989. An expert system classifies Eucalupt forest types using thematic mapper data and a digital terrain model. *Photogrammetric Engineering and Remote Sensing,* 55 (10), 1449-1464.

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.

SKIDMORE, A. K., FRANKLIN, J., DAWSON, T. P. & PILEJSO, P. 2011. Geospatial tools address emerging issues in spatial ecology: a review and commentary on the Special Issue. *International Journal of Geographical Information Science,* 25 (3), 337-365.

SKIDMORE, A. K., TURNER, B. J., BRINKHOF, W. & KNOWLES, E. 1997. Perfomance of a neural network: Mapping forests using GIS and remotely sensed data. *Photogrammetric Engineering and Remote Sensing,* 63 (5), 501-514.

SMITH, R. C. G., ADAMS, J., STEPHENS, D. J. & HICK, P. T. 1995. Forecasting wheat yield in a Mediterranean type of environment from the NOAA satellite. *Australian Journal of Agricultural Research,* 46 (1), 113-125.

STARK, N. 1970. The Nutrient Content of Plants and Soils from Brazil and Surinam. *Biotropica,* 2 (1), 51-60.

STARKS, P. J., ZHAO, D. & BROWN, M. A. 2008. Estimation of nitrogen concentration and in vitro dry matter digestibility of herbage of warm-season grass pastures from canopy hyperspectral reflectance measurements. *Grass and Forage Science,* 63 (2), 168-178.

STEINFELD, H., GERBER, P., WASSENAAR, T., CASTEL, V., ROSALES, M. & DE HAAN, C. 2006. *Livestock's Long Shadow: Environmental Issues and Options,* Rome, Italy, Food Agriculture Organization of the United Nations (FAO).

STRAHLER, A. H. 1981. Stratification of natural vegetation for forest and rangeland inventory using landsat digital imagery and collateral data. *International Journal of Remote Sensing,* 2 (1), 15-41.

STRAHLER, A. H., LOGAN, T. L. & BRYANT, N. A. 1978. Improving forest cover classification from Landsat by incoporating topographic information. *In: Proceedings of the 12th International Symposium on Remote Sensing of Environment,* Ann Arbor, MI, 1541-1557.

THENKABAIL, P. S., SMITH, R. B. & DE PAUW, E. 2000. Hyperspectral vegetation indices and their relationships with agricultural crop characteristics. *Remote Sensing of Environment,* 71 (2), 158-182.

TOWNSEND, P. A., FOSTER, J. R., CHASTAIN, R. A., JR. & CURRIE, W. S. 2003. Application of imaging spectroscopy to mapping canopy nitrogen in the forests of the central Appalachian Mountains using Hyperion and AVIRIS. *Geoscience and Remote Sensing, IEEE Transactions on,* 41 (6), 1347-1354.

TREYDTE, A. C., HEITKÖNIG, I. M. A., PRINS, H. H. T. & LUDWIG, F. 2007. Trees improve grass quality for herbivores in African savannas. *Perspectives in Plant Ecology, Evolution and Systematics,* 8 (4), 197-205.

TUCKER, C. J. 1977. Asymptotic nature of grass canopy spectral reflectance. *Applied Optics,* 16 (57-1151).

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

USTIN, S. L., GITELSON, A. A., JACQUEMOUD, S., SCHAEPMAN, M. E., ASNER, G. P., GAMON, J. A. & ZARCO-TEJADA, P. J. 2009. Retrieval of foliar information about plant pigment systems from high resolution spectroscopy. *Remote Sensing of Environment,* 113, S67-S77.

VENTER, F. 1990. A classification for land management planning in Kruger National Park. PhD Thesis, University of South Africa, South Africa.

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

VERBYLA, D., L & LITVAITIS, J., A 1989. Resampling methods for evaluating classification accuracy of wildlife habitat models. *Environmental Management,* 13 (6), 783-787.

VISCARRA ROSSEL, R. A. 2007. Robust modelling of soil diffuse reflectance spectra by "bagging-partial least square regression". *Journal of Near Infrared Spectroscopy,* 15 (1), 39-47.

VISCARRA ROSSEL, R. A. 2008. ParLeS: Software for chemometric analysis of spectroscopic data. *Chemometrics and Intelligent Laboratory Systems,* 90 (1), 72-83.

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.

WALCZAK, B. & MASSART, D. L. 1996. The Radial Basis Functions -- Partial Least Squares approach as a flexible non-linear regression technique. *Analytica Chimica Acta,* 331 (3), 177-185.

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

WANG, Z. J., WANG, J. H., LIU, L. Y., HUANG, W. J., ZHAO, 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.

WENJIANG, H., JIHUA, W., ZHIJIE, W., JIANG, Z., LIANGYUN, L. & JINDI, W. 2004. Inversion of foliar biochemical parameters at various physiological stages and grain quality indicators of winter wheat with canopy reflectance. *International Journal of Remote Sensing,* 25 (12), 2409-2419.

WESSELS, K. J., MATHIEU, R., ERASMUS, B. F. N., ASNER, G. P., SMITH, I. P. J., VAN AARDT, J., MAIN, R., FISHER, J., MARAIS, W., KENNEDY-BOWDOIN, T., KNAPP, D. E., EMERSON, R. & JACOBSON, J. 2011. Impact of communal land use and conservation on woody vegetation structure in the lowveld savannas of South Africa. *Forest Ecology and Management,* 261, 19-29.

WESSMAN, C. A. 1992. Imaging spectroscopy for remote sensing of ecosystem processes. *Advances in Space Research,* 12 (7), 361-368.

WILLIAM, P. C. & NORRIS, K. H. 1987. *Near-Infrared Technology in the Agricultural and Food Industries.,* St Paul, American Association of Cereal Chemists, USA.

WOLFSON, M. M. & TAINTON, N. M. 2000. The morphology and physiology of the major forage plants. *In:* TAINTON, N. M. (ed.) *Pasture Management in South Africa.* Pietermaritzburg: University of Natal Press.

WU, C., NIU, Z., TANG, Q. & HUANG, W. 2008. Estimating chlorophyll content from hyperspectral vegetation indices: Modeling and validation. *Agricultural and Forest Meteorology,* 148 (8-9), 1230-1241.

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.

ZARCO-TEJADA, P. J., MILLER, J. R., HARRON, J., HU, B., NOLAND, T. L., GOEL, N., MOHAMMED, G. H. & SAMPSON, P. 2004a. Needle chlorophyll content estimation through model inversion using hyperspectral data from boreal conifer forest canopies. *Remote Sensing of Environment,* 89 (2), 189-199.

ZARCO-TEJADA, P. J., MILLER, J. R., MORALES, A., BERJON, A. & AGUERA, J. 2004b. Hyperspectral indices and model simulation for chlorophyll estimation in open-canopy tree crops. *Remote Sensing of Environment,* 90 (4), 463-476.

ZEMOURI, R., RACOCEANU, D. & ZERHOUNI, N. 2003. Recurrent radial basis function network for time-series prediction. *Engineering Applications of Artificial Intelligence,* 16 (5-6), 453-463.

ZHAO, C. J., WANG, J. H. & ZHOU, Q. F. 2006. Relationship of 2100-2300 nm spectral charcteristics of wheat canopy to leaf area index and leaf N as affected by leaf water content. *Pedosphere,* 16 (3), 333-338.

ZHAO, D., HUANG, L., LI, J. & QI, J. 2007. A comparative analysis of broadband and narrowband derived vegetation indices in predicting LAI and CCD of a cotton canopy. *ISPRS Journal of Photogrammetry and Remote Sensing,* 62 (1), 25-33.

ZHOU, G., WANG, Y. & WANG, S. 2002. Responses of grassland ecosystems to precipitation and land use along the Northeast China Transect. *Journal of Vegetation Science,* 13 (3), 361-368.

Summary

Information about the distribution of grass foliar nitrogen (N) and phosphorus (P) is important to understand rangeland vitality and to facilitate effective management of wildlife and livestock. Grass N and P concentrations are direct indicators of rangeland quality, and they vary over space. Foliar N concentration is known to relate to the protein content. Protein is major nutrient requirement for the herbivores. On the other hand, foliar P concentration is a crucial requirement for reproduction and lactating animals. Successful estimation of foliar N and P concentrations could facilitate the computation of the key indicator of nutrient limitation, known as N: P ratio. Understanding the nutrient limitation could equally help the ecologists, farmers and resource manager to understand the feeding patterns, distribution and densities of herbivores both in protected and communal areas. Landscape view of the nutrient distribution and limitation as an interest to planners and managers could be achieved through remote sensing measurement and analysis. This study was undertaken in the north-eastern part of South Africa, in the savanna ecosystem. The study area was purposively selected as it covers the rangelands in the communal areas, Sabi Sands private game reserve and Kruger National Park (KNP). This study area offers experimental sites for foliar biochemical estimation, because of a pronounced contrast in soil fertility induced by various geological types, i.e. basalt and gabbro associated with high soil fertility and granite associated with low soil fertility. The main aim of the study was to develop and improve estimation of grass quality using remote sensing measurements from local to regional scale. The objectives were (1) to test water removed spectra for foliar N and P estimation as compared to the existing spectral techniques, (2) To estimate foliar N: P ratio using field spectroscopy or *in situ* hyperspectral data, (3) to investigate the applicability of the non-linear partial least square regression in integrating *in situ* hyperspectral remote sensing and environmental variables to estimate foliar N and P concentrations, and (4) to investigate the utility of the red edge band in RapidEye data for estimating foliar and canopy N at regional scale.

The success in remote sensing estimation of foliar biochemical is faced with various challenges which are yet to be addressed. The noted success is mainly based on the use of hyperspectral remote sensing. One of the challenges is the water absorption effects in short-wave infrared which mask weak or subtle foliar biochemical concentrations. At laboratory level using *in situ* hyperspectral, there were several attempts to address this challenge by drying leaf samples and measure the reflectance. The main problem associated with this approach was upscaling from laboratory (leaf) to canopy level. In Chapter 2, we proposed a technique that could be applied to minimize water absorption effects when estimating foliar biochemical using

hyperspectral remote sensing data. This study was based on the greenhouse experiment. The aim of this study was to test the utility of water removed (WR) spectra in combination with partial least square regression (PLSR) and stepwise multiple linear regression (SMLR) to estimate foliar N and P concentrations, compared to spectral transformation techniques such as first derivative, continuum removal and log transformed spectra (Log(1/R)). The savanna grass species (*Digitaria eriantha*) was sown in the greenhouse. Spectral measurements were made using a spectrometer. *D. eriantha* was cut, dried and chemically analyzed for foliar N and P concentrations. WR spectra were determined by calculating the residual from the modelled leaf water spectra using the non-linear spectral matching technique and observed leaf spectra. It was concluded that the water removal technique could be a promising technique to minimize the perturbing effect of foliar water content when estimating grass nutrient concentrations. The performance of WR spectra was also evident in Chapter 3 on estimation of foliar N: P using *in situ* hyperspectral remote sensing data. The objective of Chapter 3 study was to investigate the utility of *in situ* hyperspectral remote sensing to estimate foliar N: P, in combination with PLSR. The results showed that foliar N: P can be highly estimated by water removed and continuum removed spectra. This was undertaken at field level using ASD FieldSpec 3®, which shows a potential of this technique across various remote sensing measurement levels. The crucial level at which this technique could be tested is at airborne hyperspectral level. Generally, Chapter 3 demonstrated that foliar N: P ratio could be estimated using remote sensing.

The other challenge in estimating nutrients is the diverse and heterogeneous nature of the savanna ecosystems, in terms of soil and plant moisture, soil nutrients, fire regime, grazing pressure, species composition and anthropogenic activities. This makes remote sensing estimation of foliar biochemical a challenging venture. In Chapter 4, we investigated the use of remote sensing and environmental variables to estimate foliar N and P concentration at field level, using ASD FieldSpec 3 measurements. The objective of the study was to test the performance of non-linear PLSR for predicting grass foliar N and P concentrations through integrating *in situ* hyperspectral remote sensing and environmental variables (climatic, edaphic and topographic), named as integrated modeling approach. The data consisted of: (i) *in situ*-measured hyperspectral spectra, ii) environmental variables and measured grass N and P concentrations. The hyperspectral variables included published starch, N and protein spectral absorption features, red edge position, narrow-band indices such as simple ratio (SR) and normalized difference vegetation index (NDVI). The results of the non-linear PLSR were compared to those of conventional linear PLSR. Integrating *in situ* hyperspectral and environmental variables yielded highest foliar N and P estimation accuracy using non-linear PLSR as compared to using remote

sensing variables only, and conventional PLSR. This study demonstrated the possibility to use integrated modeling approach and non-linear PLSR in estimating foliar N and P concentrations.

In Chapter 5, we focused on the regional estimation of foliar and canopy N using spaceborne remote sensing measurements, which is new for the savanna ecosystems. The objective of this study was to estimate and map foliar and canopy Nitrogen (N) at a regional scale using a recent high resolution spaceborne multispectral sensor (i.e. RapidEye). RapidEye sensor contains five spectral bands in the visible-to-near infrared (VNIR), including a red-edge band centred at 710 nm. The importance of the red-edge band for estimating foliar chlorophyll and N concentrations has been demonstrated in many previous studies, mostly using *in situ* hyperspectral remote sensing data. The utility of the red-edge band of the RapidEye sensor for estimating grass N was investigated in this study. A two-step approach was adopted involving (i) vegetation indices and (ii) integration of vegetation indices and environmental or ancillary variables using a SMLR and non-linear PLSR. To ensure that grass N estimation is not comprised by biomass variability, the field work was undertaken when the grass has reached maximum productivity. The model involving the simple ratio (SR) index (R_{805}/R_{710}) defined as SR54, altitude and the interaction between SR54 and altitude (SR54*altitude) provided the highest accuracy for canopy N estimation, while the non-linear PLSR yielded the highest foliar N concentration estimation accuracy through integration of remote sensing (SR54) and environmental variables. The spatial pattern of foliar N concentrations thus mapped, corroborated with the soil fertility gradient induced by the geological parent material.

In conclusion, this study demonstrated that water removed spectra can be used to improve estimation of foliar biochemical concentrations. The study also revealed that foliar N: P ratio can be estimated using field spectroscopy or *in situ* hyperspectral remote sensing. The integrated modeling approach using non-linear PLSR showed to improve the estimation of foliar biochemical concentrations. For the first time, the regional estimation and mapping of savanna grass N was successfully done using the red-edge band embedded in the RapidEye multispectral sensor. The estimation and mapping of grass nitrogen could be used for understanding feeding patterns and changes in densities of wild and livestock (herbivores) at a regional scale.

Samenvatting

Gegevens over de ruimtelijke verdeling van stikstof (N) en fosfaat (P) in grasbladeren is van belang voor een beter begrip van de vitaliteit van weidegronden, en voor een effectief beheer van wilde grazers en vee. Concentraties van N en P in graslanden vertonen ruimtelijke variabiliteiten zijn een directe indicator voor de graskwaliteit. Concentratie van N in grasbladeren reflecteert het eiwitgehalte. Eiwit is een belangrijke voedingsbron voor herbivoren. Fosfor daarentegen is naast een belangrijke voedingsbron ook van cruciaal belang voor de voortplanting en voor zogende dieren. Een accurate schatting van N en P concentraties in de bladeren zou het eenvoudiger maken om de N: P verhouding te berekenen, een indicator voor nutriëntentekort in herbivoren. Een beter begrip van de nutriëntenbeperking stelt ecologen, veehouders en natuurbeheerders in staat om de voedselpatronen, verspreiding en dichtheden van herbivoren in beschermde en gemeenschappelijke graasgronden te begrijpen. Het gebruik van aardobservatie om gegevens te verzamelen en te analyseren kan de interesse bij gebruikers van graasgronden opwekken in een landschapsbenadering van nutriëntenverspreiding en –beperking. Deze studie vond plaats in het savanne ecosysteem in het noord-oostelijk deel van Zuid-Afrika. Dit studiegebied was doelgericht gekozen omdat het graasgronden omvat in gemeenschappelijke gebieden, het Sabi Sands privé natuurreservaat, en het Kruger Nationaal Park (KNP). Deze graasgronden bieden de gelegenheden voor empirisch onderzoek naar de biochemische samenstelling van grasbladeren, vanwege de sterke contrasten in bodemvruchtbaarheid, veroorzaakt door geologische verschillen, t.w. het vruchtbare basalt en gabbro, tegenover het weinig vruchtbare graniet. Het hoofddoel van deze studie was om een methode te ontwikkelen en te verbeteren om de graskwaliteit te bepalen aan de hand van aardobservatiegegevens, voor zowel locale als regionale schaal. De doelstellingen waren (1) om te bepalen hoe water-gecorrigeerde spectra voor N en P bepaling in bladeren zich verhouden tot gangbare spectrale technieken, (2) om de N: P verhouding in bladeren te bepalen met behulp van veld-spectroscopie, (3) om de toepasbaarheid van de niet-lineaire partiële kleinste kwadraten regressietechniek te bepalen bij de geïntegreerde in situ bepaling van aardobservatie- en omgevingsvariabelen voor het schatten van N en P concentraties in bladeren, en (4) om te bepalen hoe bruikbaar de zogenaamde ´red edge´ band is in de RapidEye gegevens voor de bepaling van N in bladeren en in grasland op regionale schaal.

Er zijn nog een aantal hindernissen te nemen op weg naar een succesvolle bepaling van concentraties van biochemische stoffen in bladeren. De tot nu toe geboekte vooruitgang is voornamelijk te danken aan hyperspectrale of spectroscopische bepalingen. Een van de uitdagingen is het corrigeren voor

de effecten van waterabsorptie in het kortegolf infrarode spectrum, die lage hoeveelheden van of subtiele verschillen in concentraties van biochemische stoffen vertroebelen. Gebruik makend van spectroscopische technieken zijn op laboratoriumschaal reeds enkele pogingen ondernomen om deze uitdaging aan te gaan, door bladeren van water te ontdoen (drogen) en dan de reflectie te meten. Het opschalen van laboratoriumexperimenten naar grassen in het veld vormde het grootste obstakel. In Hoofdstuk 2 stellen wij een techniek voor waarmee de spectrale effecten van waterabsorptie kunnen worden geminimaliseerd, ten gunste van de bepaling van concentraties van biochemische stoffen met behulp van hyperspectrale aardobservatie-gegevens. Deze studie is gebaseerd op een kasexperiment. Het doel hiervan was om te bepalen hoe bruikbaar water-gecorrigeerde (WR) spectra zijn, in combinatie met partiële kleinste kwadraten regressie (PLSR) en stapsgewijze multipele lineaire regressie (SMLR), voor de bepaling van N en P concentraties in bladeren, in vergelijking met spectrale transformatietechnieken zoals eerste afgeleide, ´continuum removal´ en logaritmisch getransformeerde (log(1/R)) spectra. Met behulp van een spectrometer werden spectrale bepalingen uitgevoerd op de grassoort Digitaria eriantha in een kas, waarna de bladeren werden geoogst, gedroogd, en chemisch geanalyseerd om de N en P concentraties te bepalen. Water-gecorrigeerde spectra werden gemodelleerd op basis van geobserveerde bladspectra met behulp van de niet-lineaire spectrale koppelingstechniek. De residuen hiervan werden berekend en daaruit werd het water-gecorrigeerde bladspectrum vastgesteld. Hieruit kon worden vastgesteld dat de correctie voor water een veelbelovende techniek is om het verstorende effect van bladwater op de bepaling van nutriëntengehalten in grassen te minimaliseren.

De bruikbaarheid van water-gecorrigeerde spectra bleek in Hoofdstuk 3 ook uit de bepaling van de N: P verhouding met behulp van hyperspectrale aardobservatie-gegevens in het veld. Het doel van de studie in Hoofdstuk 3 was om de bruikbaarheid te onderzoeken van de in situ toepassing van hyperspectrale aardobservatie om de N: P verhouding in grassen te bepalen, door gebruik te maken van partiële kleinste kwadraten regressieanalyse (PLSR). De resultaten lieten zien dat de N: P verhouding in grasbladeren goed bepaald kan worden uit water-gecorrigeerde en ´continuum removed´ spectra. Hiervoor werd in het veld gebruik gemaakt van de ASD FieldSpec 3® spectroradiometer, waarmee de toepasbaarheid van dit instrument over een breed scala van meetniveau´s is aangegeven. De volgende stap die nu te nemen is, is de toepassing van deze hyperspectrale techniek in vliegtuig-observaties van het aardoppervlak. De belangrijkste boodschap van Hoofdstuk 3 is dat de N: P verhouding in het veld geschat kan worden met behulp van aardobservatie.

Een andere uitdaging in de bepaling van nutrienten is gelegen in de uiteenlopende en heterogene savanne-ecosystemen op het gebied van vocht en bodem in planten, nutriënten in de bodem, brandregime, graasdruk, soortensamenstelling en menselijke activiteiten. Dit maakt de bepaling van de hoeveelheden van de biochemische stoffen in bladeren een complexe taak. In Hoofdstuk 4 onderzochten we het gebruik van aardobservatie- en omgevingsvariabelen om de N en P concentraties in bladeren te bepalen op veldniveau, wederom gebruik makend van ASD FieldSpec 3 gegevens. Het doel was om na te gaan of niet-lineaire partiële kleinste kwadraten regressieanalyse (PLSR) voldeed bij het voorspellen van N en P concentraties in in situ grasbladeren binnen een geïntegreerde modelmatige benadering, gebruik makend van hyperspectrale aardobservatie- en omgevingsvariabelen (klimatologisch, edafisch en topografisch). De data betroffen (i) in situ gemeten hyperspectrale spectra (ii) omgevingsvariabelen en (iii) chemisch bepaalde N en P concentraties. De hyperspectrale variabelen omvatten gepubliceerde zetmeelwaarden, N en eiwit spectrale absorptiekenmerken, ´red edge´ positie, indices op basis van smalle banden zoals de eenvoudige ratio (´simple ratio´, SR) en de ´normalized difference vegetation index´ (NDVI). De resultaten van de niet-lineaire partiële kleinste kwadraten regressieanalyse werden vergeleken met die van een conventionele (lineaire) partiële kleinste kwadraten regressieanalyse. Het integreren van in situ hyperspectrale en omgevingsvariabelen gaven de hoogste accuraatheid van de N en P schattingen op basis van niet-lineaire partiële kleinste kwadraten regressieanalyse (PLSR), in vergelijking met uitsluitend aardobservatie-variabelen en met conventionele partiële kleinste kwadraten regressieanalyse. De mogelijkheid om een geïntegreerde modelbenadering uit te voeren en niet-lineaire partiële kleinste kwadraten regressieanalyse toe te passen om N en P concentraties in bladeren te schatten is in deze studie duidelijk naar voren gekomen.

In Hoofdstuk 5 richtten we ons op de schatting van N in bladeren en graslanden met behulp van satelliet-gebaseerde aardobservatie-metingen. Dit is niet eerder toegepast in savanne-ecosystemen. Het doel van deze studie was het schatten en het karteren van het N gehalte in bladeren en graslanden op een regionaal schaalniveau op basis van een recent in gebruik genomen multi-spectrale, satelliet-sensor, de RapidEye. De RapidEye sensor heeft vijf spectrale banden in het zichtbare gebied tot het nabij-infrarood (VNIR), waaronder een ´red edge´ band rondom 710 nm. Uit veel, met name veldpectroscopisch onderzoek, is het belang van de ´red edge´ band voor het schatten van chlorofyl en N concentraties in bladeren naar voren gekomen. In deze studie is onderzocht in hoeverre met behulp van de ´red edge´ band van de RapidEye sensor het schatten van N in grassen mogelijk was. In een tweetrapsbenadering werden (i) alleen vegetatie-indices en (ii) geïntegreerde vegetatie-indices en omgevingsvariabelen of toegevoegde

variabelen gebruikt en vervolgens geanalyseerd met behulp van stapsgewijze multipele lineaire regressie (SMLR) en niet-lineaire partiële kleinste kwadraten regressieanalyse (PLSR). Door het veldwerk uit te voeren gedurende de piek van het groeiseizoen werd voorkomen dat de N bepalingen in gras werden beïnvloed door variatie in de biomassa. Op basis van alleen vegetatie-indices en omgevingsvariabelen gaven de modellen met de eenvoudige ratio (SR) index (R_{805}/R_{710}), hier gedefiniëerd als SR54, het model met hoogte, en de interactie tussen SR54 en hoogte (SR54*hoogte) de beste accuraatheid voor N gehalte in het grasland. Op basis van geïntegreerde vegetatie-indices (SR54) en omgevingsvariabelen gaf niet-lineaire partiële kleinste kwadraten regressieanalyse (PLSR) de beste accuraatheid voor N gehalte in het grasland. Het aldus in kaart gebrachte ruimtelijk patroon van N concentraties in grasbladeren kwam overeen met de bodemvruchtbaarheidsgradient die het onderliggende geologische moedermateriaal reflecteert.

Concluderend kunnen we stellen dat deze studie duidelijk heeft gemaakt dat water-gecorrigeerde spectra gebruikt kunnen worden om de schatting van biochemische concentraties van biochemische stoffen te verbeteren. De studie gaf ook aan dat de N: P verhouding in bladeren geschat kan worden met behulp van veldspectroscopie. Door geïntegreerde aardobservatie-variabelen en omgevings-variabelen te analyseren met behulp van niet-lineaire partiële kleinste kwadraten regressieanalyse (PLSR), verbeterden de schattingen van de concentraties van de biochemische stoffen. Deze studie was de eerste die een regionale schatting en kartering van biochemische bladkenmerken uitvoerde op basis van de ´red edge´ band in de RapidEye multispectrale satelliet-sensor. De schatting en kartering van stikstof in grassen kan gebruikt worden voor het analyseren van voedselzoekpatronen in wilde en gedomesticeerde grazers, en voor het analyseren van hun dichtheidsveranderingen op regionale schaal.

Publications

Scientific papers

Ramoelo A., Skidmore A.K. Schlerf M., Mathieu R, Heitkonig I.M.A. (2011). Water-removed spectra increase the retrieval accuracy when estimating savannah grass nitrogen and phosphorus concentrations, *ISPRS Journal of Photogrammetry and Remote Sensing (66),* pp. 408-417

Ramoelo A., Skidmore A.K., Schlerf M., Heitkonig I.M.A, Cho M.A. (under review) , Savanna grass nitrogen to phosphorus ratio estimation using *in situ* hyperspectral remote sensing data, *International Journal of Remote Sensing*

Ramoelo A, Skidmore A.K., Cho M.A., Mathieu R., Heitkonig I.M.A., Dudeni-Tlhone-Dudeni N, Schlerf M., Prins H.H.T (under review) 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*

Ramoelo A, Skidmore A.K, Cho M.A. Schlerf M. Mathieu R. Heitkönig I.M.A. (under review) Regional estimation of savannah grass nitrogen using red edge band of the RapidEye sensor, *International Journal of Applied Earth Observation and Geo-information (JAG)*

Main R., Cho M.A., Mathieu R., O'Kennedy M.M., **Ramoelo A**., Koch S. (2011). An investigation into robust spectral indices for leaf chlorophyll estimation, *ISPRS Journal of Photogrammetry and Remote Sensing (66),* pp. 751-761

Ramoelo A and Van Aardt J (Accepted with minor revision) An overview of ground-based light detection and ranging (LiDAR) remote sensing for system structural assessment, *South African Journal of Science*

Conference Proceedings Papers

Ramoelo A., Cho M. A., Mathieu R., Skidmore A. K., Schlerf M., Heitkönig I. M. A. Prins H. H. T. (2011) Integrating environmental and in situ hyperspectral remote sensing variables for grass nitrogen estimation in savanna ecosystems. *34th International Symposium on the Remote Sensing of Environment* (ISRSE 2011), The GEOSS Era: Towards Operational Environmental Monitoring. Sydney, Australia, http://www.isprs.org/proceedings/2011/ISRSE-34/index.html

Dudeni-Tlhone N, **Ramoelo A**, Debba P, Cho MA;, Mathieu R, 2011, Modelling grass nutrients using remote sensing and geostatistical tools, *South African Statistical Association Conference* 2011 (SASA2011), Pretoria, South Africa, 31 October-4 November 2011

Dudeni-Tlhone N, **Ramoelo A**, Cho MA; Debba P, Mathieu R, 2010, Modelling nutrient concentration to determine the environmental factors influencing grass quality, *South African Statistical Association Conference* 2010 (SASA2010), Potchefstroom, South Africa, 8-12 November 2010

Cho M, Van Aardt J, Main R, Majeke B, **Ramoelo A**, Mathieu R, Norris-Rogers M, Du Plessis R (2009) Integrating remote sensing and ancillary data for regional ecosystem assessment: Eucalyptus Grandis agro-system in KwaZulu-Natal, South Africa. *IEEE International Geoscience and Remote Sensing Symposium* (IGARSS), proceedings, Cape Town, South Africa, July 12-17, 2009, Pages: 264-267

Cho MA, Debba P, Mathieu R, Van Aardt J, Asner G, Naidoo L, Main R, **Ramoelo A**, Majeke B. (2009) Spectral variability within species and its effects on savanna tree species discrimination, Conference Attendance *IEEE International Geoscience and Remote Sensing Symposium* (IGARSS), Cape Town, South Africa, 12-17 July 2009, Pages: 191-193

Biography

Abel Ramoelo was born in Muduluni village, South Africa on the 02 November 1980. He completed his primary level school in 1993 at Madaheni primary school, Makhitha village. In 1998, he completed matric certificate from Luvhivhini secondary school, Maebani village. Given his passion on mathematics, physical sciences (Physics and chemistry), biology and geography, he pursued a Bachelor of Environmental Science degree at University of Venda (South Africa). He graduated top of the school of environmental sciences in 2003, scooping several awards. In 2003, he registered an honours degree in environmental science, at University of Venda. During his honours study, he developed a passion on using remote sensing and geographical information system tools to address and model environmental related issues. In May 2004 he started working at Human Science Research Council (HSRC) as GIS intern. After 3 months, August 2004, He joined the Council for Scientific and Industrial Research (CSIR) as junior researcher (remote sensing and GIS). After a year, he started a Master of Science in Geo-information Science and Earth Observation for Environmental Management and Modelling in four European universities (University of Southampton-UK, Lund University-Sweden, University for Warsaw-Poland, University of Twente-ITC, The Netherlands), funded by Erasmus Mundus Scholarship. The title of the thesis was "An innovative method to map land cover changes at country level utilizing hyper-temporal satellite images". Since he is an aspiring scientist, six months after acquiring his MSc., he took-up a PhD position under the supervision of prof.dr. Andrew Skidmore, which resulted in this thesis. In 2011, he was awarded "Emerging Researcher Award" during the "Excellence Award" for the Natural Resource and Environment Unit of CSIR.

ITC Dissertation List

http://www.itc.nl/Pub/research/Graduate-programme/Graduate-programme-PhD_Graduates.html