

Integrating Environmental and *in situ* Hyperspectral Remote Sensing Variables for Grass Nitrogen Estimation in Savannah Ecosystems

A. Ramoelo^{a, b*}, M. Cho^a, R. Mathieu^a, A.K. Skidmore^b, M. Schlerf^b, I.M.A. Heitkonig^c, H.H.T. Prins^c

^aNatural Resource and Environment, Ecosystem-Earth Observation, Council for Scientific and Industrial Research (CSIR), Meiring Naude Rd, Brummeria, Pretoria, 0184, Republic of South Africa (aramoelo, mcho, rmathieu)@csir.co.za

^bUniversity of Twente, Faculty of Geoinformation Science and Earth Observation (UT-ITC), Hengelosestraat 99, 7500 AE, Enschede, The Netherlands (ramo14741, skidmore, schlerf)@itc.nl

^cUniversity of Wageningen, Resource Ecology Group, Droevendaalsesteeg 3a, 6708 PB Wageningen, The Netherlands (herbert.prins, ignas.heitkonig)@wur.nl

Abstract-Information about the distribution of grass nitrogen (N) concentration is crucial in understanding rangeland vitality and facilitates effective management of wildlife and livestock. A challenge in estimating grass N concentration using remote sensing in savannah ecosystems is that these areas are characterised by heterogeneity in edaphic, topographic and climatic factors. The objective is to test the utility of integrating environmental variables and *in situ* hyperspectral remote sensing variables for predicting grass N concentration along a land use gradient in the greater Kruger National Park. Data used include i) environmental variables, ii) measured grass N concentration and iii) *in situ* measured hyperspectral spectra. Non-linear partial least square regression was used. Results showed that several environmental variables were important for N estimation. Integrating environmental variables with *in situ* hyperspectral variables increased grass N estimation accuracy. The study demonstrated the importance of integrated modelling for savannah ecosystem state assessment.

Keywords: Hyperspectral, Savannah ecosystem, Resources, Foliar nitrogen concentration

1. INTRODUCTION

Knowledge of the distribution of grass nitrogen (N) concentration is crucial in understanding rangeland vitality for effective planning and management of wildlife and livestock grazing. Studies indicated the importance of N concentration in grass to influence the distribution of livestock and wildlife (McNaughton 1988; Seagle and McNaughton 1992). For assessing the distribution of the grass quality at landscape level, remote sensing can play a pivotal role as it provides a synoptic view of the landscape.

Conventional broadband indices e.g. the normalized difference vegetation index (NDVI) (Tucker 1979), are inadequate for estimating subtle changes in leaf N concentrations. The above indices do saturate at high canopy cover and are insensitive to subtle changes in leaf N concentration. On the other hand, the advent of hyperspectral remote sensing has highlighted the utility of red-edge bands (700 – 750 nm) for estimating foliar N and chlorophyll content (Huang et al. 2004; Cho and Skidmore 2006; Majeke et al. 2008; Darvishzadeh et al. 2008a). The red-edge is the region of abrupt change in leaf reflectance between 680 and 780 nm. Several studies also identified and successfully used the absorption features for N and protein on leaf spectra for foliar N concentrations (Curran 1989; Kumar et al. 2001). Normalized difference or ratio indices computed from red-edge

bands have provided more accurate estimates of leaf N compared to conventional NDVI derived from 680 and 800 nm (Mutanga and Skidmore 2007). Further, innovative spectral transformation technique such as continuum removal was used to enhance N and protein absorption features for N estimation (Huang et al. 2004; Mutanga et al. 2004a; Cho et al. 2010).

A challenge in estimating grass N concentration using remote sensing in savannah is that these ecosystems are characterised by heterogeneity in soil and plant moisture, soil nutrients, grazing pressures and anthropogenic activities (Ben-Shahar and Coe 1992). Geology, soil, precipitation, fire and temperature, topography or catenal position, aspect and land use types influence the grass biochemical concentrations. Therefore, to develop robust models for foliar N estimation, techniques which integrate hyperspectral remote sensing and environmental (edaphic, topographic and climatic) variables could be necessary. Limited studies investigated the use of environmental and remote sensing variables to estimate grass N. The objective is to test the utility of integrating environmental variables (climatic, edaphic and topographic) and *in situ* hyperspectral remote sensing variables for predicting grass N concentrations in savannahs.

2. MATERIAL AND METHODS

2.1. Study area and sampling design

The study was undertaken in the greater Kruger National Park (KNP) including public and privately-owned protected areas such as KNP and SabiSands (SGR), as well as communal rangelands (Bushbuckridge area), north-eastern part of South Africa. These areas are characterized by gradients of moisture and nutrients. The communal areas on the western side receive more rainfall than eastern part of the study area (KNP). The gradient of nutrients is influenced by the domination of granite and gabbro underlying strata. The gabbro areas are characterized by higher soil fertility than the granitic ones (Ben-Shahar and Coe 1992). The topography is mainly undulating throughout the study area with some flat patches in localized areas. In the gabbro side we find dominant grass such as *Setaria sphacelata* dominating the crest and *Urochloa mosambicensis* dominating the valleys. There is more high grass biomass in the gabbro (e.g. *Urochloa mosambicensis*) than on the granitic soils (e.g. *Eragrostis rigidior*). The protected areas have various grazing activities by wild herbivore such impala, elephant, kudu, rhinoceros, etc. while the communal rangelands support grazing of mainly cattle and goats as well as sheep, which determine various rangeland or land use intensities.

*Corresponding author: Abel Ramoelo, email: aramoelo@csir.co.za or ramo1474@itc.nl, Tel: +27128413840, Fax: +2712 841 3909. We would like to thank the Council for Scientific and Industrial Research, South Africa's Department of Science and Technology (DST) and National Research Foundation, University of Wageningen and University of Twente for the research funding.

Eight research sites were determined using the 1:250000 geology map and 2008 SPOT 5 images. Two sites were located in KNP (1 in each, gabbro and granite), 2 in SGR (1 in each, gabbro and granite), 4 in the communal lands (2 in each, gabbro and granite). The sites were selected to capture the gradient of grass nutrients. In each site, a line transect was purposively placed covering the main topographical classes (crest and valleys) because biomass levels vary with a topographical position. Along the transects, homogeneous grass plots of 30 m x 30 m were located 500m to 1000m from each other depending on the accessibility and homogeneity of the patch. In each plot, 3 to 4 subplots with the size of 0.5 m x 0.5 m were randomly selected making a total of 189 grass sample points. The fieldwork was conducted during the rainy season of March/April 2009.

2.2. Canopy Spectral measurements

The spectral reflectance data was collected for each subplots using Analytical Spectral Device (ASD) spectroradiometer, Fieldspec 3®. The ASD has spectral bands ranging from 350 to 2500, with 1 nm band width. For each subplot, five (5) spectral measurements were taken and later averaged to account for grass canopy heterogeneity, illumination differences, and bidirectional effects. The spectral measurements were taken between 10h30 and 15h00 on clear sunny days, to minimize cloud effects. A 25⁰ fibre optic was used and it was held at nadir and kept at a height of 1m to cover the entire subplot. A white reference panel (Spectralon) was used after each measurement to convert radiance to reflectance. After spectral measurements, for each subplot, grass samples were collected and dried at 80⁰C for 24 hours.

2.3. Chemical analysis

The dried grass samples were taken to the Agricultural Research Council-Institute of Tropical and Subtropical Crops-(ARC-ITSC), Nelspruit, South Africa for analysis of N concentration. The acid digestion technique (utilizing sulphuric acid) was used for foliar N retrieval on grass samples (Giron 1973; Grasshoff et al. 1983).

2.4. Spectral indices and selection of known absorption features

Two vegetation indices were selected for this study, namely the red-edge position and the normalized difference vegetation index (NDVI). The red-edge position is known to have a high correlation with N and is less sensitive to soil background (Cho and Skidmore 2006). For this study the linear extrapolation technique was used to compute the red edge position (Cho and Skidmore 2006). Cho and Skidmore (2006) found out that linear extrapolation technique achieved higher accuracy as compared to other red-edge detection techniques in retrieving N and chlorophyll. More details about the linear extrapolation can be found in Cho and Skidmore (2006).

The NDVI is a known vegetation index used in many studies as a surrogate for vegetation condition and health (Tucker 1979). It is known to minimize the atmospheric effects on the image (Zarco-Tejada et al. 2004). The narrow-band NDVI has been proposed to minimize problems of saturation generally observed with broadband-NDVI (Mutanga and Skidmore 2004). As a result, bands at 700 and 750nm were selected for the red and near-infrared reflectance values respectively.

The known absorption features for chlorophyll, protein and N, ranging from the visible, near infrared and shortwave infrared wavelength were selected centred at 430nm, 460nm, 640nm, 660nm, 910nm, 1510nm, 1940nm, 2060nm, 2180nm, 2300nm and 2350nm (Curran 1989) and the reflectance values for this bands were extracted and used for further analysis.

2.5. Environmental data

Precipitation, temperature, land use, geology, altitude, slope and aspect were used as potential environmental variables influencing the accuracy of N prediction. Slope and Aspect were computed from digital elevation model (DEM) which was derived using ArcGIS 10x at a resolution of 50m. The DEM was derived using 1:50000 topographical maps. Mean annual precipitation and temperatures were acquired from the Global Climatic datasets at a resolution of 1 km (Hijmans et al. 2005). The Land use types were derived from the boundary layers of KNP, SabiSands and communal areas acquired from Kruger National Park's Geographic Information System (GIS) and remote sensing laboratory. The 1:250000 geological data was acquired from South Africa's Council of Geoscience.

2.6. Statistical analysis

Multi-variate statistical technique such as nonlinear partial least square (PLSR) regressions was used. The non-linear PLSR known as PLSR with radial basis function was used for foliar N estimation (Walczak and Massart 1996; Daszykowski et al. 2007). The non-linear PLSR technique is a flexible technique which can predict both non-and normally distributed response variables, with the mutual advantage of the non-linear nature of the radial basis function and the capability of PLSR to maximize covariance between datasets. It couples the power of neural network and PLSR. It is implemented like any other PLSR techniques. The objective is to decompose a list of independent variables into latent and uncorrelated variables. Monte-Carlo leave-one-out cross-validation was used to determine the optimum number of factors based on the lowest RMSE. The weight loadings were analysed to determine the contribution of each independent variables. A conventional PLSR was used to show the PLSR weights assigned to each independent variable sensitive to foliar N, which the non-linear PLSR does not provide. For independent testing of the foliar N estimation models, the data was randomly split into 70% training and 30% test data sets. Training data sets (132 samples) were used for model development or calibration and test data sets (57 samples) for validation.

3. RESULT AND DISCUSSION

The results shows that integrating environmental variables with *in situ* hyperspectral variables increased grass N estimation precision and accuracy (Training: $R^2=0.77$, root mean square error of cross-validation (RMSECV) =0.11%, Test: $R^2=0.49$, root mean square error of prediction (RMSEP) = 0.14%) compared to using remote sensing variables only (Training: $R^2=0.50$, RMSECV=0.13%, Test: $R^2=0.41$, RMSEP=0.16%) (Figure 1). The incorporation of environmental variables with *in situ* hyperspectral data produced a more robust model of N which better accounts for the heterogeneity and complexity of the savannah ecosystem.

Results showed that generally environmental variables had a higher influence on the model outputs than the spectral variables. Amongst the environmental variable geology (GEO) and temperature (TEM) had the highest PLSR weights followed

by land use types (LAN) and slope (SLO) (Figure 2). The diversity and availability of minerals in the soil determines the complex web of the inter relationships between geology, topography and grass N concentrations. Geology was significantly important in estimating grass N concentration since there are evidence of the correlation between foliar N and soil fertility levels (Ben-Shahar and Coe 1992). In the study area, the gabbro underlying geology influence the higher soil nutrient (dark and humus rich) which in turn influence higher grass N concentration and generally higher grass biomass for species such as *Setaria sphacelata*, *Urochloa mosambicensis*, *Digiraria eriantha*, *Themenda triandra* and *Panicum maximum* (Ben-Shahar and Coe 1992). In contrary, the granitic area supports low foliar N concentrations with relatively lower biomass of species such as *Eragrostis rigidior*, *Sporobolus spp*, *Tracheolena monache*. The slope also determines the occurrence of high quality grasses mainly along the valleys while the crests have relatively lower N concentrations. This is due to the fact that run-off of nutrients and fine clay particles from the crest or along the slope are deposited into the valley or depressions along the riparian zones.

Amongst the spectral variables REP was found to have a higher impact on the model, as compared to the narrow band NDVI, while that for the absorption bands 910 and 1020nm had the highest contribution to the model. The red edge position is known to minimize the soil background effects, and it has been widely utilized and highly correlated to chlorophyll (Cho and Skidmore 2006; Darvishzadeh et al. 2008a). Studies also found out a positive correlation between chlorophyll and N (Yoder and Pettigrew-Crosby 1995). The results found here are in line with other studies focusing on foliar N concentration using *in*

situ hyperspectral remote sensing (Gong et al. 2002; Mutanga et al. 2004a). Gong et al. (2002) showed the importance of using the blue and red edge for N estimation. While Mutanga and Skidmore (2004a) find out that the use of continuum removal indices also improve the estimation of N, because it enhance the absorption features.

The selected absorption features are sensitive to foliar N through various vibration mechanisms such as electron transition or bond vibrations (Curran 1989). Absorptions in the visible region of the spectra are produced by electron transition while in the near and short wave infrared they are produced by bond vibration; see Curran (1989) for details. Several studies used N and protein absorption features not only for foliar N estimation, but also for biomass and LAI estimations (Cho et al. 2007; Darvishzadeh et al. 2008a). PLSR has so far proved to be a robust multivariate approach for biochemical estimation as compared to univariate analysis. The multivariate techniques yield higher precision and accuracy for biochemical estimation than the univariate ones, which use narrow band NDVI or red edge position (Gong et al. 2002; Darvishzadeh et al. 2008a). However, integrating NDVI, red edge, known absorption features for protein and N as well as environmental variables using multivariate techniques increase the accuracy of foliar N estimations.

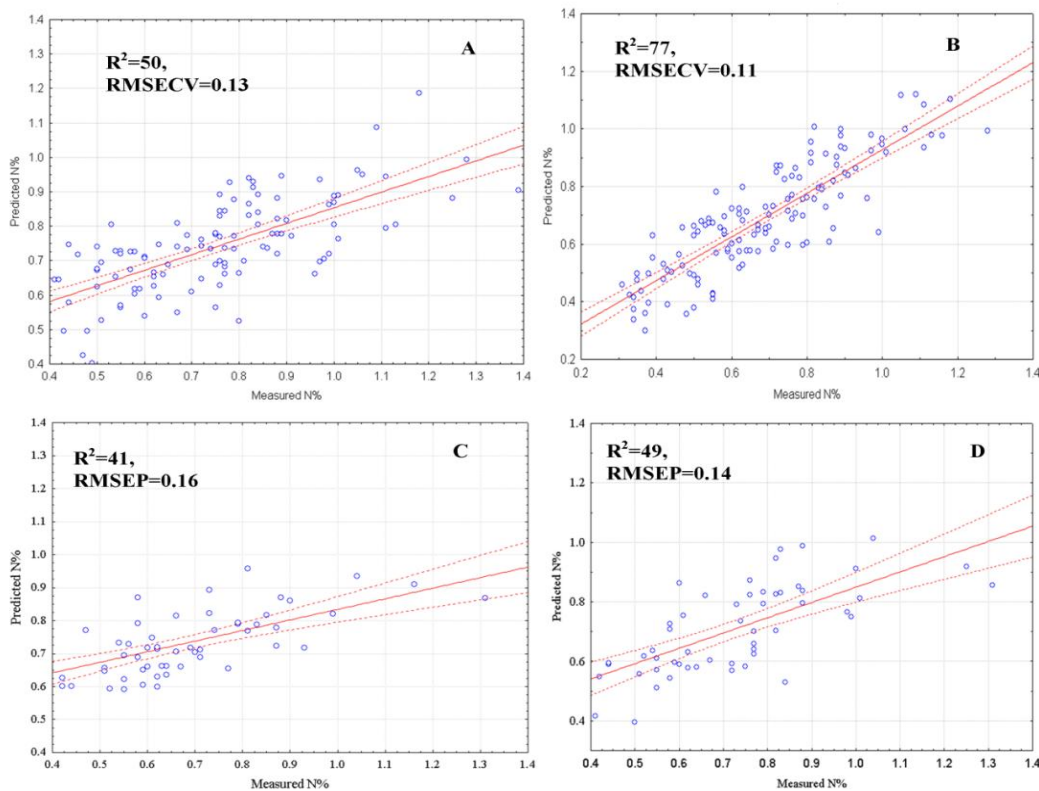


Figure 1: Foliar N estimation A) calibration model with remote sensing variables only B) calibration with the remote sensing and environmental variables, C) validation for remote sensing derived model and D) validation for remote sensing and environmental variables integrated. RMSECV=root mean square error of cross validation, RMSEP=root mean square error of prediction.

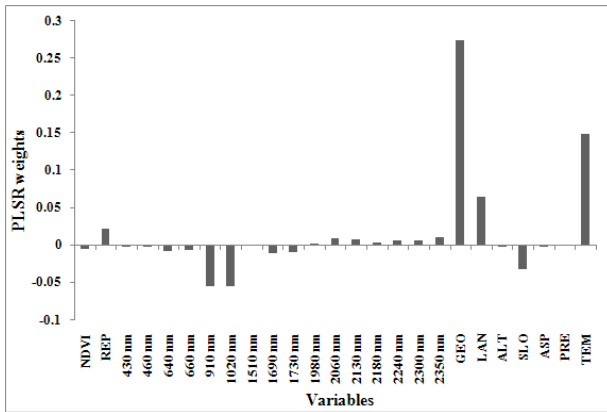


Figure 2: Partial least square regression (PLSR) weights for remote sensing and environmental variables in grass N estimation (GEO=Geology, LAN=Land use, Alt=altitude, ASP=aspect, PRE=precipitation, REP=red edge position).

4. CONCLUSIONS

The study demonstrated the importance of integrated modelling for ecosystem state assessment which is imperative for planning and management of protected and communal savannah rangelands. Environmental variables could play a crucial role in improving the accuracy of foliar N estimation. The use of environmental variable could be imperative for upscaling N estimation models to regional scale to understand distribution of livestock and wildlife densities for planning and management purpose of protected and communal savannah ecosystems.

5. REFERENCE:

Ben-Shahar, R. and Coe, M. J. The relationships between soil factors, grass nutrients and the foraging behaviour of wildebeest and zebra. *Oecologia* 90(3): 422-428. 1992.

Cho, M., Skidmore, A. K., Corsi, F., van Wieren, S. and Sobhan, I. Estimation of green grass/herb biomass from airborne hyperspectral imagery using spectral indices and partial least square regressions. *Int. J. App. Earth Obs.* 9: 414-424. 2007.

Cho, M. A. and Skidmore, A. K. A new technique for extracting the red edge position from hyperspectral data: The linear extrapolation method. *Remote Sens. Environ.* 101(2): 181-193. 2006.

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

Curran, P. J. Remote sensing of foliar chemistry. *Remote Sens. Environ.* 30(3): 271-278. 1989.

Darvishzadeh, R., Skidmore, A., Schlerf, M., Atzberger, C., Corsi, F. and Cho, M. LAI and chlorophyll estimation for a heterogeneous grassland using hyperspectral measurements. *Int. Soc. Photogramme.* 63(4): 409-426. 2008a.

Daszykowski, M., Serneels, S., Kaczmarek, K., Van Espen, P., Croux, C. and Walczak, B. TOMCAT: A MATLAB toolbox for multivariate calibration techniques. *Chemometr Intell. Lab.* 85(2): 269-277. 2007.

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

Gong, P., Pu, R. and Heald, R. C. Analysis of in situ hyperspectral data for nutrient estimation of giant sequoia. *Int. J. Remote Sens.* 23(9): 1827 - 1850. 2002.

Grasshoff, K., Erhardt, M. and Kremling, K., Eds. (1983). *Methods of seawater analysis.* Weinheim, Germany, Verlag Chemie.

Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G. and Jarvis, A. Very High Resolution Interpolated Climate Surfaces for Global Land Areas. *Int. J. Climatol.* 25: 1965-1978. 2005.

Huang, Z., Turner, B. J., Dury, S. J., Wallis, I. R. and Foley, W. J. Estimating foliage nitrogen concentration from HYMAP data using continuum removal analysis. *Remote Sens. Environ.* 93: 18-29. 2004.

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

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

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

Mutanga, O. and Skidmore, A. K. Narrow band vegetation indices overcome the saturation problem in biomass estimation. *Int. J. Remote Sens.* 25(19): 3999 - 4014. 2004.

Mutanga, O. and Skidmore, A. K. Red edge shift and biochemical content in grass canopies. *Int. Soc. Photogramme.* 62: 34-42. 2007.

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

Seagle, S. W. and McNaughton, S. J. Spatial variation in forage nutrient concentrations and the distribution of serengeti ungulates. *Landscape Ecol.* 7(4): 229-241. 1992.

Tucker, C. J. Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sens. Environ.* 8: 50-127. 1979.

Walczak, B. and Massart, D. L. The Radial Basis Functions -- Partial Least Squares approach as a flexible non-linear regression technique. *Anal. Chim. Acta* 331(3): 177-185. 1996.

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

Zarco-Tejada, P. J., Miller, J. R., Morales, A., Berjon, A. and Aguera, J. Hyperspectral indices and model simulation for chlorophyll estimation in open-canopy tree crops. *Remote Sens. Environ.* 90(4): 463-476. 2004.

6. ACKNOWLEDGEMENT

We would like to thank my colleagues Mr (s) Laven Naidoo and Russell Main, as well as Thulani from Bushbuckridge for their field work assistance.