Non-linear partial least square regression increases the estimation accuracy of grass nitrogen and phosphorus using *in situ* hyperspectral and environmental data

Ramoelo A[a,b,*]. Skidmore A.K[b]. Cho M.A.[a], Mathieu R[a]. Heitkönig I.M.A[c]. Dudeni-Tlhone N[d]. Schlerf M[b]. Prins H.H.T[c]

[a]Earth Observation Research Group, Natural Resource and the Environment Unit, Council for Scientific and Industrial Research (CSIR), P.O.Box 395, Pretoria, 0001, South Africa

[b]Faculty of Geoinformation Science and Earth Observation, University of Twente (UT-ITC), P.O.Box 217, Enschede, 7500 AE, The Netherlands

[c]Resource Ecology Group, Wageningen University, Droevendaalsesteeg 3a, 6708 PB Wageningen, The Netherlands

[d]Statistical Analysis and Modelling Research Group, Logistics and Quantitative Methods, Built Environment Unit, Council for Scientific and Industrial Research (CSIR), P.O.Box 395, Pretoria, 0001, South Africa

*Corresponding author: Dr Abel Ramoelo, **Tel**: +27 12 841 3968, **Fax**: +27 12 841 3909, **Email**: ramo14741@itc.nl or aramoelo@csir.co.za

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. 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 remote sensing 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. Using 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.

Keywords: *in situ* hyperspectral remote sensing, ecosystem, partial least square regression, radial basis neural network, nitrogen concentrations, phosphorus concentrations

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; Prins and van Langevelde, 2008; Seagle and McNaughton, 1992). 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; Ludwig et al., 2008; Treydte et al., 2007). 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.

The 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, and thus contribute to poverty alleviation in rural areas

Remote sensing is widely used as a cost-effective means (Mumby et al., 1999) to estimate and map plant condition or quality at landscape level in various biomes, such as grasslands and savannas (Bogrekci and Lee, 2005; Ferwerda et al., 2005; Mutanga and Kumar, 2007; Mutanga and Skidmore, 2004a; Mutanga et al., 2005; Mutanga et al., 2004b, c; Numata et al., 2008; Skidmore et al., 2010), forests (Martin and Aber, 1997; Schlerf et al., 2010) and agricultural areas (Hansen and Schjoerring, 2003;

Huang et al., 2004; LaCapra et al., 1996; Thenkabail et al., 2000; Wang et al., 2009; Zarco-Tejada et al., 2004). The 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), and soil-adjusted vegetation index (SAVI) (Huete, 1988) have limited applications in high grass canopy environments as they saturate at high canopy cover (Mutanga and Skidmore, 2004b; Tucker, 1977). 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 (Cho and Skidmore, 2006; Clevers et al., 2002; Darvishzadeh et al., 2008; Huang et al., 2004; 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 (Cho et al., 2010; Curran, 1989; Elvidge, 1990; Knox et al., 2010; Kokaly and Clark, 1999; Kumar et al., 2001; 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 (Cho et al., 2010; Mutanga and Kumar, 2007; Ramoelo et al., 2011). Spectral transformation techniques such as water and continuum removal have been proposed to enhance nutrient absorption features (Cho et al., 2010; Huang et al., 2004; Mutanga et al., 2004c; Ramoelo et al., 2011; Schlerf et al., 2010).

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 (He and Mui, 2010; Mutanga and Kumar, 2007; Mutanga et al., 2004c; Skidmore et al., 2010). Grass quality is influenced by geology (Ben-Shahar and Coe, 1992; Grant and Scholes, 2006), soil (Cho et al., 2010; Heitkönig and Owen-Smith, 1998), precipitation and temperature (Ben-Shahar and Coe, 1992), topography or catena position (Mutanga et al., 2004a; Seagle and McNaughton, 1992) as well as aspect (Mutanga et al., 2004a), and land use types. 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 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 (Cho et al., 2009; Knox et al., 2011; Mutanga et al., 2004a). 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., (2009b) and Knox et al., (2011).

Several studies have successfully used stepwise multiple linear regression (SMLR) (Grossman et al., 1996; Huang et al., 2004; Kokaly and Clark, 1999; Martin and Aber, 1997) 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 multicolinearity (Grossman et al., 1996; Huang et al., 2004). The use of partial least square regression (PLSR) has been advocated to address these issues (Asner and Martin, 2008; Darvishzadeh et al., 2008; Geladi and Kowalski, 1986; Huang et al., 2004; Ramoelo et al., 2011). 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.

In the conventional linear PLSR model, the centred data matrices $\boldsymbol{X}$ and $\boldsymbol{Y}$ are projected onto the low-dimensional score matrices, $\boldsymbol{T}$ and $\boldsymbol{U}$, respectively (Martens and Naes, 2001; Viscarra Rossel 2008; William and Norris, 1987), as well outlined by Walczak and Massart, (1996),

$$\boldsymbol{X}=\boldsymbol{TP'}+\boldsymbol{C} \quad (1)$$

$$\boldsymbol{Y}=\boldsymbol{UC'}+\boldsymbol{F} \quad (2)$$

where $\boldsymbol{P}$ and $\boldsymbol{C}$ are the regression coefficients (loadings).

T and ***U*** in PLS are developed the same way as the principal component analysis (PCA) (Geladi and Kowalski, 1986), but the difference is that PLSR uses both dependent and independent variables to decompose the input data into latent variables (Geladi and Kowalski, 1986).

When the weights are not normalized, the linear relation between the scores matrices T and U can be represented as

$$\boldsymbol{U=T+H} \quad (3)$$

and then,

$$\boldsymbol{Y=TC'+F^{*}}, \quad (4)$$

where matrices ***E***, ***F***, ***F**** and ***H*** contain residuals.

For the RBF-PLS, the activation matrix ***A*** should be constructed using a Gaussian function which is normally characterized by two parameters, namely, center and width (Walczak and Massart, 1996). The RBF can be considered as a 3-layer net containing input, hidden and output layers, similar to any neural network procedure (Walczak and Massart, 1996). When the width and centres are specified, the input value to each output node is a weighted sum of all outputs of the hidden nodes (i.e. corresponds to the dimension of the input data). The final RBF model has the following form:

$$\boldsymbol{Y=A \times w}$$

where ***w*** are weights which are normally adjusted to minimize the mean square error of the net output.

Therefore, the PLS procedure can be applied to model matrices *A* and *Y*. In this case, the centred *A* and *Y* construct a linear PLS model:

$$\boldsymbol{Y=TC'+F},$$

with $\boldsymbol{T}$ representing the score matrix of $\boldsymbol{A}$.

Scores $\boldsymbol{A}$ are the linear combinations of the Gaussian maximizing the covariance between $\boldsymbol{A}$ and $\boldsymbol{Y}$ (Walczak and Massart, 1996). In essence, the nonlinear relation is transformed to the problem in linear algebra (Walczak and Massart, 1996). A key thing is to construct the activation matrix, which is then used with PLSR to predict foliar biochemical.

The non-linear PLSR as described above is also known as 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 multicolinearity problems if the optimal number of latent variables are 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 the 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 remote sensing 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 remote sensing variables. Conventional and non-linear PLSR *vs.* integrated modelling results were compared.

2. Material and Methods

2.1. Study area and sampling design

The study area is located in the Lowveld savanna at the north-eastern part of South Africa (Fig. 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.) 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 Pogonathria 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* (Ferwerda et al., 2006; Venter et al., 2003). 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 primigenius*) and goats (*Capra hircus*) as well as sheep (*Ovis aries*), which determine various grazing intensities.

(Figure 1)

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) (Fig. 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 sought to capture the nutrient gradient 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, grass cover (%) 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 from 31 March to 17th 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 (Ramoelo et al., 2012; Skidmore et al., 2010).

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 grass foliar N and P by Mutanga et al., (2004) and Ramoelo et al., (2011).

2.3. Canopy Spectral measurements

The reflectance spectra were measured using an Analytical Spectral Device (ASD) spectroradiometer, Fieldspec 3®. The full-width-half-maximum (FWHM) spectral resolution of the ASD is 3 nm for the region 350 -1000 nm and 10 nm for the region 1000 - 2500 nm. 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.

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.

2.4.1. Spectral indices

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., 2004). 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 (Darvishzadeh et al., 2008; Mutanga and Skidmore, 2004b).

2.4.2. Selection of absorption features

Chlorophyll, protein, and N absorption features were selected for estimating foliar N concentrations (Curran, 1989; Kumar et al., 2001) (Table 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., 2010; Schlerf et al., 2010; Skidmore et al., 2010), while Ramoelo et al., (2011), 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 1)

2.5. Environmental data

Environmental variables used in this study include precipitation, temperature, land use, geology, soils, distance to rivers, altitude, slope, and aspect (Table 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., (2011), Pickett et al., (2003), Venter et al., (2003), and Mutanga et al., (2004). Details for each environmental variable are mentioned below;

- Annual average precipitation and 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. The unit for the distance were measured in kilometres (km).
- Geology data was acquired from the council for Geoscience in South Africa. Major classes used in this study are granite and gabbro. Granite is associated with high soil fertility, while gabbro is associated with relatively high soil fertility.

- A soil layer was acquired from the soil and terrain database of Southern Africa (SOTERSAF) (Dijkshoorn, 2003; Dijkshoorn et al., 2008; FAO et al., 2003). Three major classes of soils such as albic arenosols, calvic vertisol and eutric regosols occurs in the study area. More details about these types of can be found in FAO et al., (2003) and Dijkshoorn, (2003).
- 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.

For all the data layers, ArcGIS spatial analyst tool was used to extract the values and classes corresponding to the sampling points, performing "Extract values to points" for rasters and "Overlay" for vectors. The extracted values and classes were used to create a database for statistical analysis.

(Table 2)

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; Martens and Naes, 2001; Naes et al., 1986; 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 (Daszykowski et al., 2007; Walczak and Massart, 1996). 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] (Knox et al., 2011; Mutanga and Kumar, 2007; Skidmore et al., 2010) 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 biochemical. 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 is not biased, since it uses 48 samples for data calibration to predict the remaining 1 iteratively (Darvishzadeh et al., 2008). 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, non-linear PLSR technique does not provide this information. 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).

3. Results

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 (R^2=0.81, RMSE=0.08, 11.4% of the mean) as compared to the use of remote sensing variables only (R^2 =0.66, RMSE=0.11: 15.7% of the mean) (Fig. 2, Table 3). Integrating *in situ* hyperspectral remote sensing 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 (Fig. 2, Table 3). Generally, the non-linear PLSR achieved higher estimation accuracy of grass N than the conventional PLSR, both considering remote sensing and environmental variables and remote sensing variables only (Table 3). The predictive capability of foliar N concentrations using environmental variables alone is low, as compared to using remote sensing variables as well as integrated modelling approach (Table 3).

(Table 3)

(Figure 2)

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 Fig. 4. Geology, soil types, land use, distance to rivers and temperature resulted to a positive contribution to the N prediction model, while slope and grass cover contributed negatively as shown by PLSR weights in Fig. 4. Table 4 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). Table 6 shows that REP is the only variable achieved high correlation with foliar N. 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 8).

(Table 4)

(Figure 3)

(Figure 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) (Fig. 3, Table 3). Integrating *in situ* hyperspectral remote sensing 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 (Fig. 3, Table 3). The predictive capability of foliar P concentrations using environmental variables alone is low, as compared to using remote sensing variables as well as integrated modelling approach (Table 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, 970, 990, and 1020 nm wavebands showed a negative relationship with foliar P concentration (Fig. 5). Geology, land use types, and soils showed positive PLSR weights, while slope, grass cover, and temperature had negative PLSR weight or contribution to the model (Fig. 5). The results show that the contribution of the environmental variables to P estimation is similar to that of N (Figs. 4 and 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 5). Table 7 shows low correlation between remote sensing variables and foliar P. 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 8).

(Table 5)

(Table 6)

(Table 7)

(Table 8)

(Figure 5)

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 remote sensing and environmental variables (climatic, edaphic and topographic) for predicting grass N and P concentrations.

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 (Ehsani et al., 1999; Geladi et al., 1999; Geladi and Kowalski, 1986) and (2) minimizing the over-fitting and multicolinearity (Huang et al., 2004; Walczak and Massart, 1996), 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.

4.2. Integrated modelling for Foliar N estimation: contribution of remote sensing variables

The results showed that integrating *in situ* hyperspectral remote sensing 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., 2008) and it minimizes soil background effects (Zarco-Tejada et al., 2004). 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 remote sensing (Gong et al., 2002; Knox et al., 2010; Mutanga et al., 2004c). 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., 2007; Darvishzadeh et al., 2008).

This study is in 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 (Franklin et al., 1986; Skidmore, 1989; Strahler, 1981).

4.3. Integrated modelling for foliar P estimation: contribution of remote sensing 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. Few studies on prediction of foliar P concentrations from spectral data found that most sensitive bands are located in the SWIR (Bogrekci and Lee, 2005; Cho et al., 2010; Mutanga and Kumar, 2007; Ramoelo et al., 2011). 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 in 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. 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; Grant and Scholes, 2006; Heitkönig and Owen-Smith, 1998; Pickett et al., 2003; 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 is linked to soil nutrient contents (Ben-Shahar and Coe, 1992; Pickett et al., 2003; Skidmore et al., 2010; Venter et al., 2003). 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. *Erogrostis rigidior, Sporobolus spp.*) because of the clay content in the soil, which is higher in the gabbro than the granite. 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 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 and 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 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 (Figs. 4; 5, Tables 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. A negative correlation between slope and grass cover also shows that high vegetation cover with high nutrient concentrations is along the drainage lines or valleys than the crest (Fig. 6). Fig. 6 shows various selected subplots where the grass samples were collected, with different grass cover.

(Figure 6)

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, 5) (Ben-Shahar and Coe, 1992). In systems like savannas, the gradient of soil nutrients and moisture is influenced by an interaction of local parent material, topography, climate, and living organisms which occur in a complex way (Venter et al., 2003). A main disadvantage of using environmental parameters in this type of study is that they are not often available in relatively high resolutions and scales, but integrated with remote sensing variables the effects of this challenge are minimized. In addition, Table 3 shows that predictive capability of foliar N and P concentrations using environmental variables is low, as compared to using remote sensing variables as well as integrated modelling approach. This might be due to the resolutions and scales of environmental variables, which are coarse in most cases.

5. Conclusions

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 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 communal savanna ecosystems.

6. 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 Dr. 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. Special thanks to the anonymous reviewers for their valuable comments.

7. Reference

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.

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.

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.

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.

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

Cho, M., Skidmore, A.K., Corsi, F., van Wieren, S., Sobhan, I., 2007. 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., 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, pp. 264-267.

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.

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.

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

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

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

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.

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

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

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. Dr. W. Junk Publishing Company, Dordrecht, pp. 133-149.

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.

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.

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

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

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.

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.

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.

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.

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., 1983. Methods of Seawater Analysis. Verlag Cheime, Weinhein, Germany.

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.

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.

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. John Wiley & Sons, New York.

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(3), 295-309.

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., 2010. Nitrogen prediction in grasses: effect of

bandwidth and plant material state on absorption feature selection. International Journal of Remote Sensing 31(3), 691-704.

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., 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.

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. Kluver Academc Publishers, Dordrecht, pp. 111-154.

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.

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

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.

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.

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. American Association of Cereal Chemists, Minessota, USA, pp. 59-100.

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

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.

Mucina, L., Rutherford, M.C., 2006. The Vegetation of South Africa, Lesotho and Swaziland. Strelitzia, Cape Town.

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.

Pickett, S.T.A., Gadenasso, M.L., Benning, T.L., 2003. Biotic and Abiotic Variability as Key Determinants of Savanna Heterogeneity at Spatiotemporal Scales, In: Du Toit,

J.T., Rogers, K.H., Biggs, H.C. (Eds.), The Kruger Experience: Ecology and Management of Savanna Heterogeneity. Island Press, London, pp. 22-40.

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. Springer, Netherlands, pp. 129-154.

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

Ramoelo, A., Skidmore, A.K., Schlerf, M., 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(4), 408-417.

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.

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. Photogramm. Eng. Remote Sens. 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.

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.

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.

Venter, F.J., Scholes, R.J., Eckhardt, H.C., 2003. Abiotic template and its associated vegetation pattern, In: Du Toit, J.T., Kevin, H.R., Biggs, H.C. (Eds.), The Kruger Experience: Ecology and Management of Savanna Heterogeneity. The Island Press, London.

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

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.

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.

Zarco-Tejada, P.J., Miller, J.R., Morales, A., Berjon, A., Aguera, J., 2004. 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, 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.

1

2

3

4

5

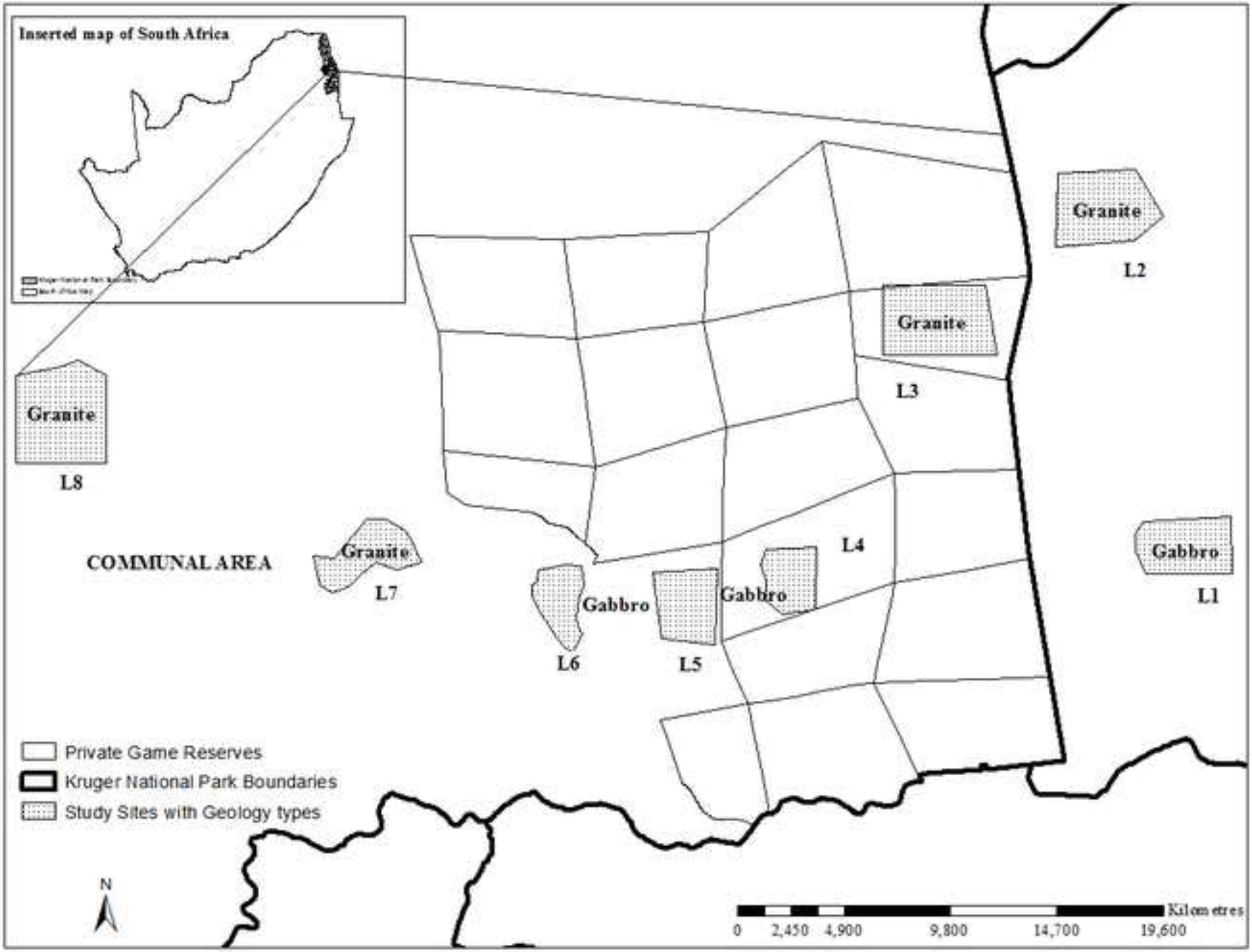
6

7

8

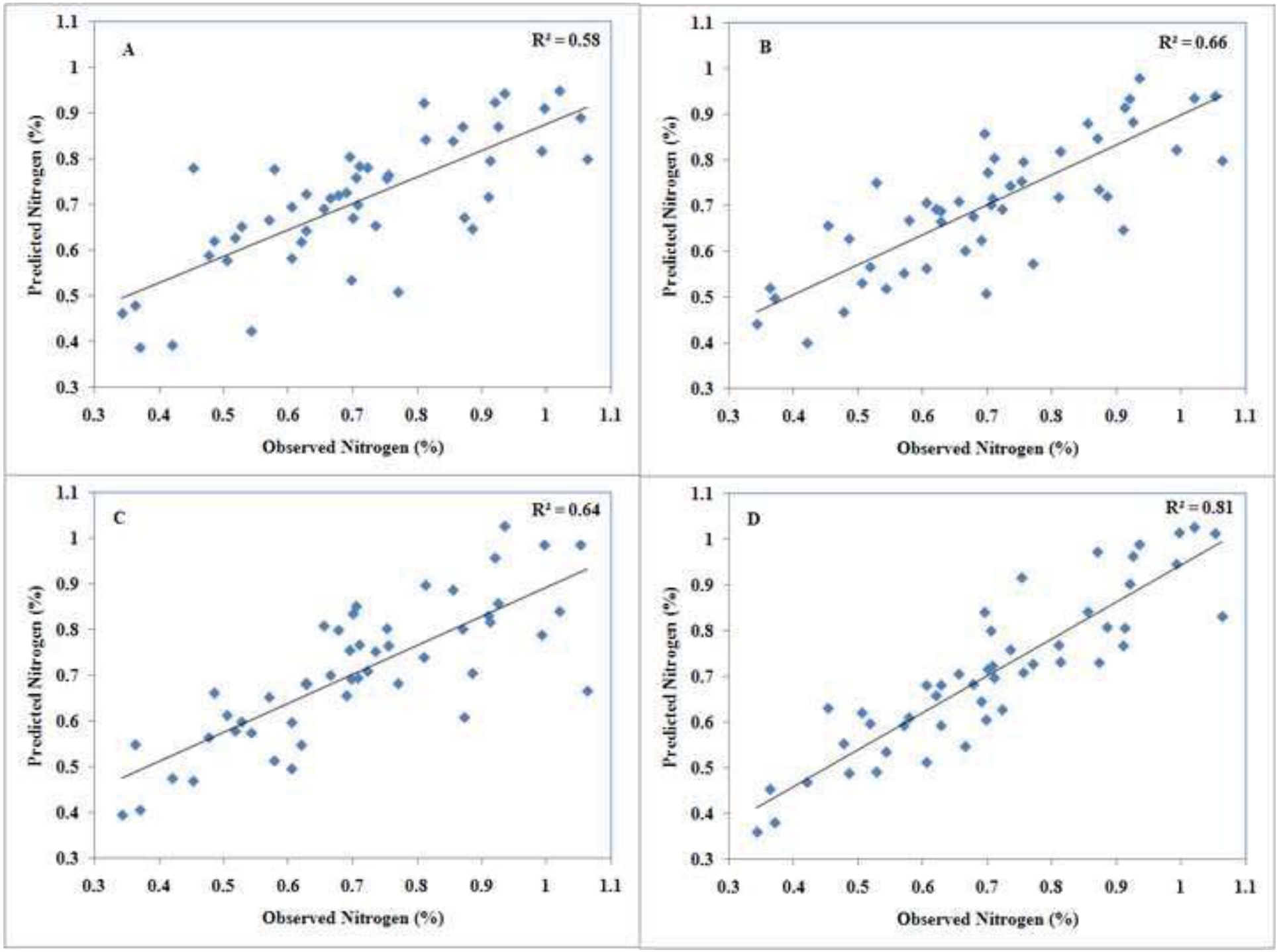
9

Figure(1)
Click here to download high resolution image

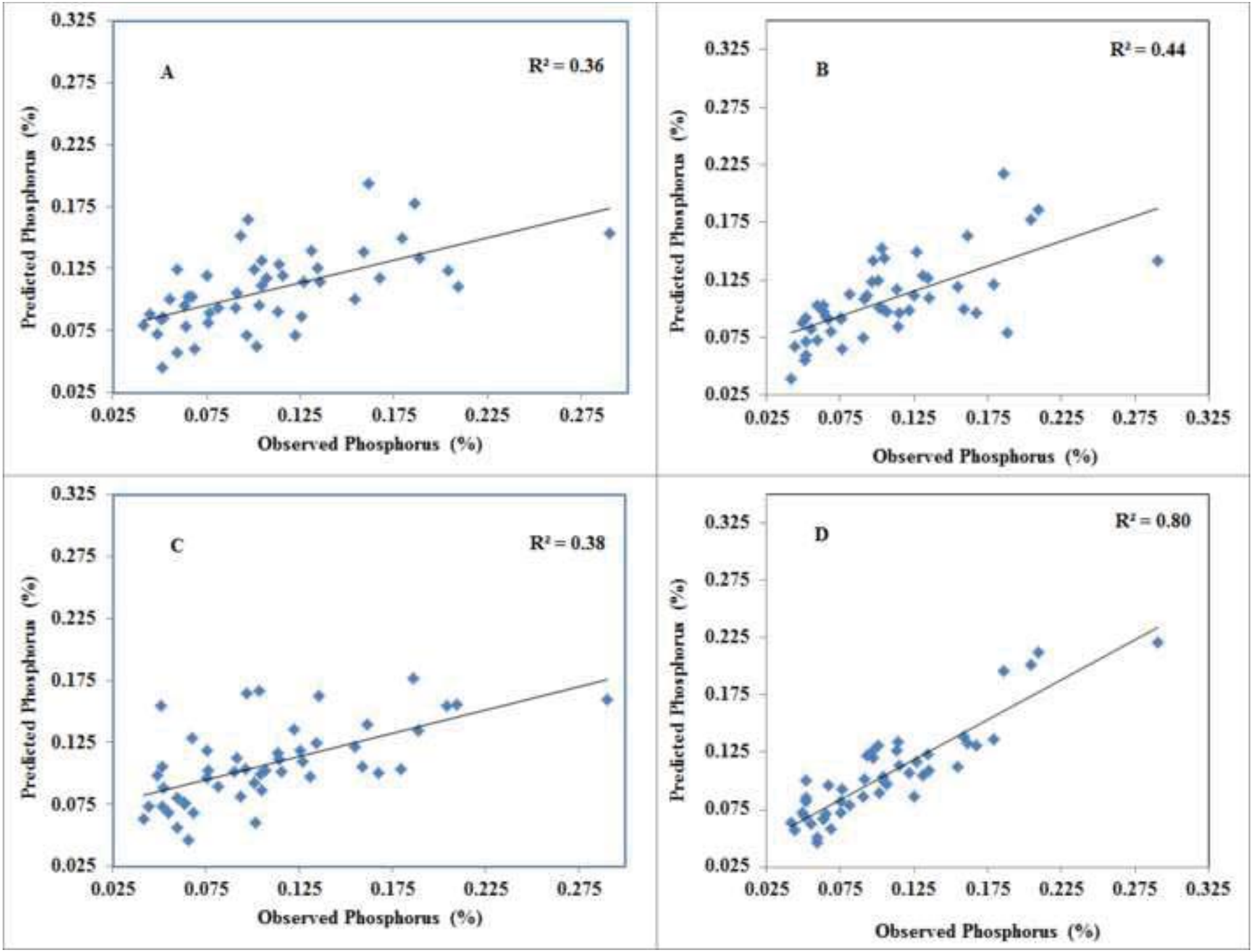


Figure(2)

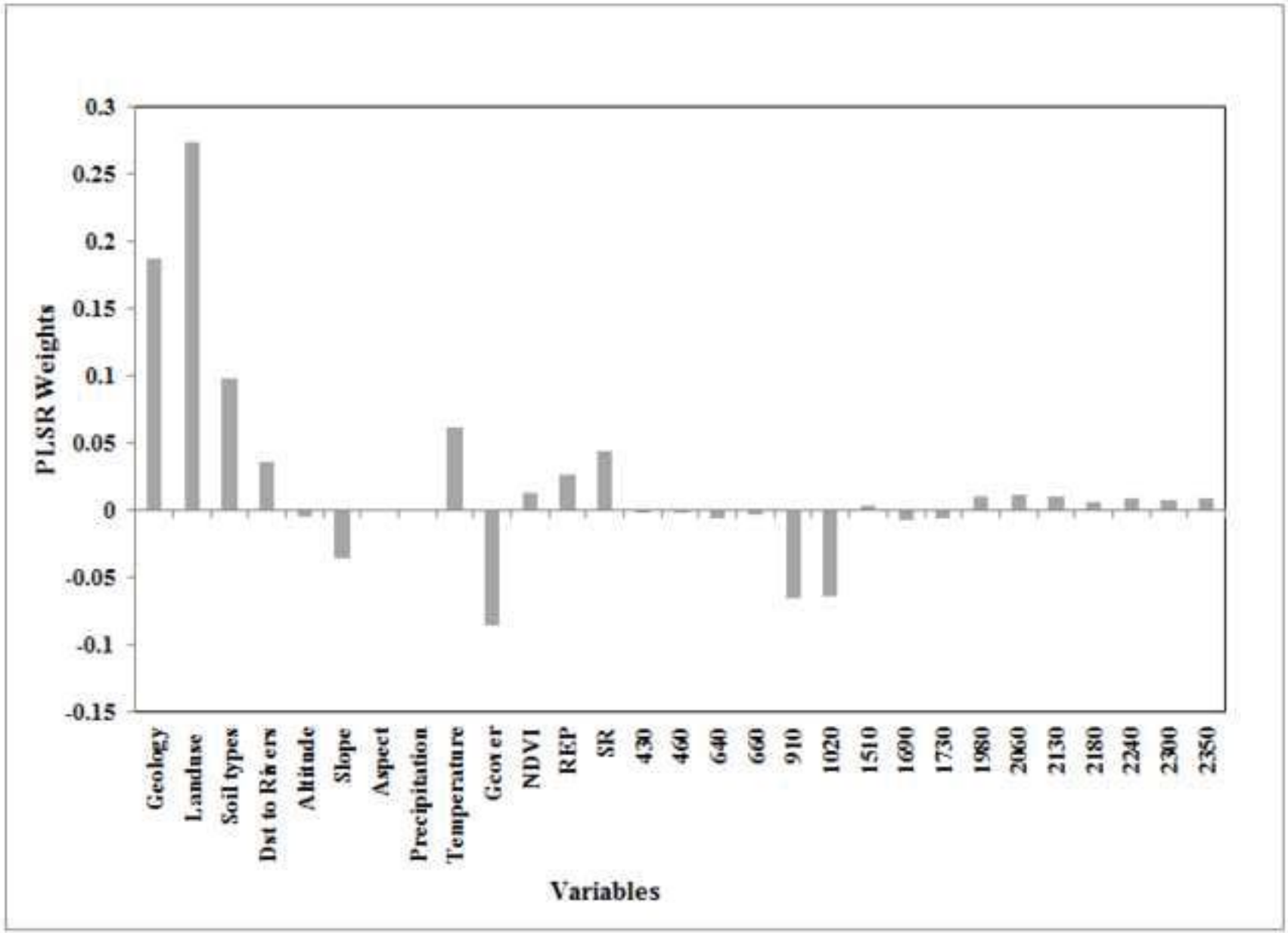
Click here to download high resolution image



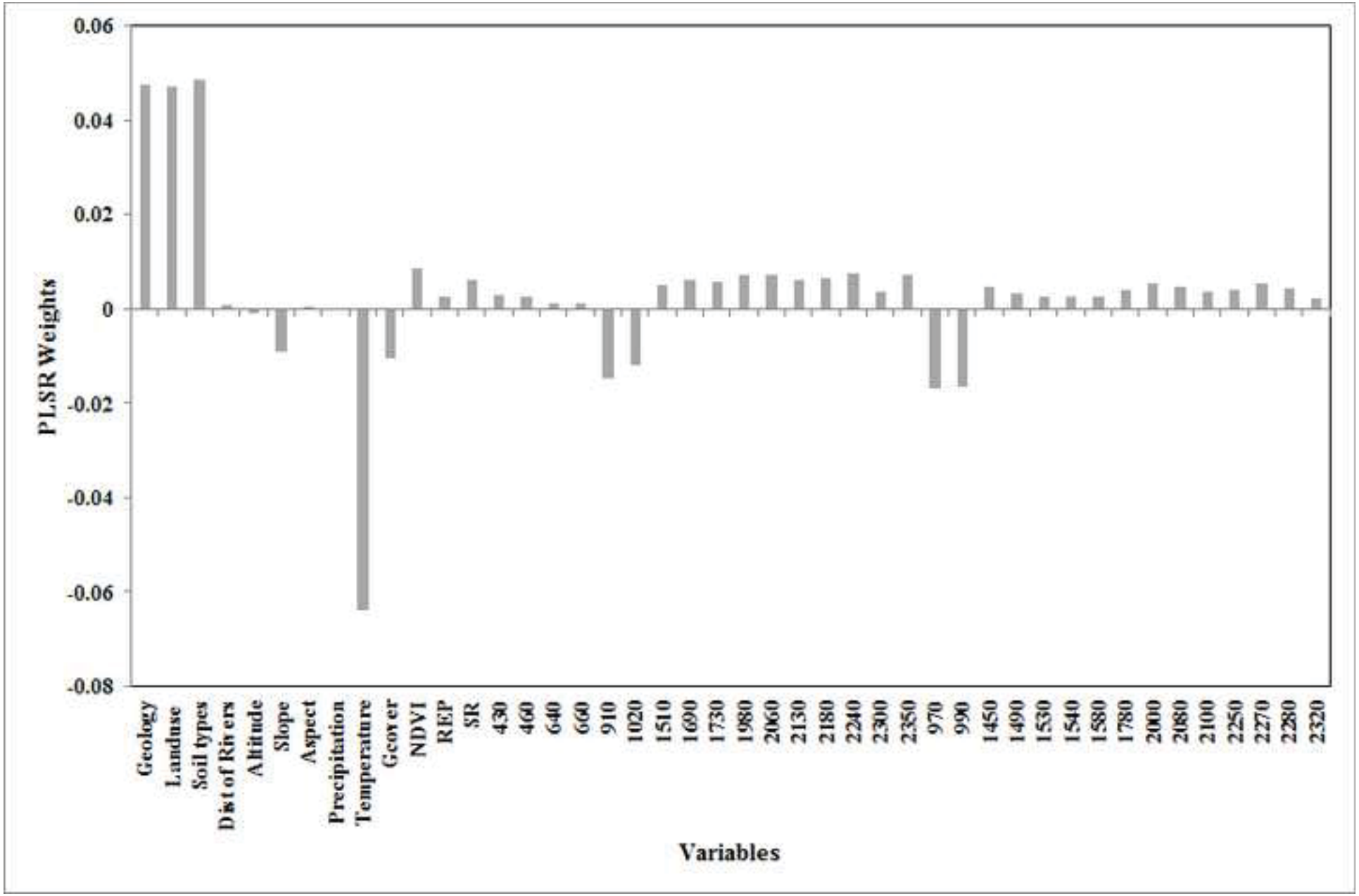
Figure(3)
Click here to download high resolution image



Figure(4)
Click here to download high resolution image



Figure(5)
Click here to download high resolution image



Figure(s)

Click here to download high resolution image

Figure Captions

Figure 1: Study area map: *L=Land use.*

Figure 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.

Figure 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: 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, Gcover=grass cover.*

Figure 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, Gcover=grass cover.*

Figure 6: Shows pictures of the selected subplots (50 cm x 50 cm quadrant) with various grass cover levels, for example; *Top Left* (green: 95% and 5% dry), *Top Right* (green: 90% and 10% dry), *Bottom Left* (green: 85%%, dry: 10% and 5% bare) and *Bottom Right* (green: 85%, dry:15%).

Tables

Table 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

Table 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

Table 3: Performance for foliar N and P prediction through integrating environmental and *in situ* hyperspectral remote sensing variables as compared to using *in situ* hyperspectral remote sensing 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
N vs. Env	0.16	0.17	0.17	5	0.38	0.11	0.15	5
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[a]	0.80	0.03	0.02	13[ψ]
P vs. Env	0.13	0.05	0.05	5	0.23	0.04	0.05	6

N=nitrogen, P=phosphorus, RS=remote sensing variables (all as given in Table 1 and in the text), Env=environmental variables (all as given in Table 2), PLSR=partial least square regression, RMSE=root mean square error, RMSECV= root mean square error of cross validation. For the selected latent factors above 10, RMSECV of the first four factors are listed,=0.1325, 0.1237, 0.1122, 0.1129, [a] =0.0443, 0.0443, 0.0438, 0.0442 and [ψ]=0.0360, 0.0373, 0.0442, 0.0387.*

Table 4: correlation matrix between foliar N and all environmental variables

	N %	Geo	Land	Soil	DRivers	Alt	Slope	Aspect	Precip	Temp	Gcover
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	
Gcover	**0.22**	-0.16	-0.09	-0.13	-0.12	-0.12	**-0.27**	-0.10	-0.06	0.02	1.00

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

Table 5: correlation matrix between foliar P and all environmental variables

	P %	Geo	Land	Soil	DRivers	Alt	Slope	Aspect	Prep	Temp	Gcover
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						
Altitude	-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	
Gcover	0.19	-0.16	-0.09	-0.13	-0.12	-0.12	**-0.27**	-0.10	-0.06	0.02	1.00

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

Table 6: correlation matrix between foliar N and remote sensing variables such as vegetation indices and absorption features in nanometers (nm)

	N	**NDVI**	**REP**	**SR**	**430**	**460**	**640**	**660**	**910**	**1020**	**1510**	**1690**	**1730**	**1980**	**2060**	**2130**	**2180**	**2240**	**2300**	**2350**
N	1.00																			
NDVI	-0.05	1.00																		
REP	**0.51**	0.06	1.00																	
SR	-0.13	**-0.36**	**-0.73**	1.00																
430	0.00	-0.27	0.07	0.28	1.00															
460	-0.04	**-0.32**	0.05	**0.30**	**0.99**	1.00														
640	-0.21	**-0.36**	-0.19	**0.45**	**0.85**	**0.91**	1.00													
660	-0.18	**-0.39**	-0.23	**0.54**	**0.83**	**0.89**	**0.99**	1.00												
910	-0.13	-0.17	**0.45**	**-0.42**	**0.49**	**0.54**	**0.51**	**0.40**	1.00											
1020	-0.14	-0.22	**0.41**	**-0.36**	**0.54**	**0.59**	**0.57**	**0.47**	**0.99**	1.00										
1510	-0.06	**-0.35**	-0.16	**0.54**	**0.79**	**0.84**	**0.93**	**0.95**	**0.35**	**0.42**	1.00									
1690	-0.07	**-0.37**	0.00	**0.33**	**0.82**	**0.88**	**0.94**	**0.93**	**0.59**	**0.66**	**0.95**	1.00								
1730	-0.07	**-0.35**	-0.02	**0.35**	**0.82**	**0.88**	**0.94**	**0.93**	**0.57**	**0.64**	**0.96**	**1.00**	1.00							
1980	-0.03	-0.23	-0.18	**0.51**	**0.70**	**0.72**	**0.75**	**0.79**	0.16	0.22	**0.80**	**0.74**	**0.74**	1.00						
2060	-0.03	-0.26	-0.23	**0.61**	**0.75**	**0.79**	**0.88**	**0.91**	0.18	0.25	**0.98**	**0.88**	**0.90**	**0.83**	1.00					
2130	0.00	-0.26	-0.17	**0.56**	**0.77**	**0.82**	**0.89**	**0.92**	0.26	**0.32**	**0.99**	**0.91**	**0.92**	**0.81**	**0.99**	1.00				
2180	-0.03	**-0.30**	-0.15	**0.52**	**0.80**	**0.85**	**0.92**	**0.94**	**0.34**	**0.41**	**0.99**	**0.95**	**0.96**	**0.82**	**0.98**	**0.99**	1.00			
2240	-0.02	**-0.29**	-0.15	**0.53**	**0.80**	**0.84**	**0.91**	**0.94**	**0.32**	**0.39**	**0.99**	**0.94**	**0.95**	**0.83**	**0.99**	**0.99**	**1.00**	1.00		
2300	-0.01	-0.25	-0.17	**0.55**	**0.77**	**0.82**	**0.89**	**0.92**	0.27	**0.33**	**0.98**	**0.91**	**0.92**	**0.81**	**0.99**	**1.00**	**0.99**	**0.99**	1.00	
2350	0.00	-0.21	-0.13	**0.48**	**0.79**	**0.82**	**0.85**	**0.87**	0.27	**0.33**	**0.91**	**0.86**	**0.87**	**0.96**	**0.93**	**0.92**	**0.93**	**0.94**	**0.93**	1.00

N=Nitrogen, NDVI=normalized difference vegetation index, REP=red edge position, SR=simple ratio

Table 7: correlation matrix between foliar P and remote sensing variables such as vegetation indices and absorption features in nanometers (nm)

	P	NDVI	REP	SR	430	460	640	660	910	1020	1510	1690	1730	1980	2060	2130	2180	2240	2300	2350	970	990	1450	1490	1530	1540	1580	1780	2000	2080	2100	2250	2270	2280	2320
P	1.00																																	
NDVI	-0.18	1.00																																
REP	0.14	0.06	1.00																															
SR	-0.02	**-0.36**	**-0.73**	1.00																														
430	0.22	-0.27	0.07	0.28	1.00																													
460	0.19	**-0.32**	0.05	**0.30**	**0.99**	1.00																												
640	0.09	**-0.36**	-0.19	**0.45**	**0.85**	**0.91**	1.00																											
660	0.08	**-0.39**	-0.23	**0.54**	**0.83**	**0.89**	**0.99**	1.00																										
910	0.05	-0.17	**0.45**	**-0.42**	**0.49**	**0.54**	**0.51**	**0.40**	1.00																									
1020	0.06	-0.22	**0.41**	**-0.36**	**0.54**	**0.59**	**0.57**	**0.47**	**0.99**	1.00																								
1510	0.08	**-0.35**	-0.16	**0.54**	**0.79**	**0.84**	**0.93**	**0.95**	**0.35**	**0.42**	1.00																							
1690	0.11	**-0.37**	0.00	**0.33**	**0.82**	**0.88**	**0.94**	**0.93**	**0.59**	**0.66**	**0.95**	1.00																						
1730	0.10	**-0.35**	-0.02	**0.35**	**0.82**	**0.88**	**0.94**	**0.93**	**0.57**	**0.64**	**0.96**	**1.00**	1.00																					
1980	0.04	-0.23	-0.18	**0.51**	**0.70**	**0.72**	**0.75**	**0.79**	0.16	0.22	**0.80**	**0.74**	**0.74**	1.00																				
2060	0.07	-0.26	-0.23	**0.61**	**0.75**	**0.79**	**0.88**	**0.91**	0.18	0.25	**0.98**	**0.88**	**0.90**	**0.83**	1.00																			
2130	0.07	-0.26	-0.17	**0.56**	**0.77**	**0.82**	**0.89**	**0.92**	0.26	**0.32**	**0.99**	**0.91**	**0.92**	**0.81**	**0.99**	1.00																		
2180	0.09	**-0.30**	-0.15	**0.52**	**0.80**	**0.85**	**0.92**	**0.94**	**0.34**	**0.41**	**0.99**	**0.95**	**0.96**	**0.82**	**0.98**	**0.99**	1.00																	
2240	0.09	**-0.29**	-0.15	**0.53**	**0.80**	**0.84**	**0.91**	**0.94**	**0.32**	**0.39**	**0.99**	**0.94**	**0.95**	**0.83**	**0.99**	**0.99**	**1.00**	1.00																
2300	0.05	-0.25	-0.17	**0.55**	**0.77**	**0.82**	**0.89**	**0.92**	0.27	**0.33**	**0.98**	**0.91**	**0.92**	**0.81**	**0.99**	**1.00**	**0.99**	0.99	1.00															
2350	0.07	-0.21	-0.13	**0.48**	**0.79**	**0.82**	**0.85**	**0.87**	0.27	**0.33**	**0.91**	**0.86**	**0.87**	**0.96**	**0.93**	**0.92**	**0.93**	0.94	**0.93**	1.00														
970	0.04	-0.22	**0.40**	**-0.35**	**0.52**	**0.58**	**0.56**	**0.46**	**0.99**	**0.99**	**0.41**	**0.64**	**0.62**	0.21	0.24	**0.31**	**0.39**	0.38	**0.32**	**0.32**	1.00													
990	0.05	-0.23	**0.40**	**-0.35**	**0.53**	**0.58**	**0.56**	**0.46**	**0.99**	**0.99**	**0.41**	**0.65**	**0.63**	0.21	0.24	**0.31**	**0.39**	0.38	**0.32**	**0.32**	**1.00**	1.00												
1450	0.06	**-0.31**	-0.21	**0.57**	**0.77**	**0.82**	**0.91**	**0.94**	**0.30**	**0.38**	**0.99**	**0.93**	**0.94**	**0.81**	**0.98**	**0.98**	**0.99**	0.99	**0.97**	**0.90**	**0.37**	**0.37**	1.00											
1490	0.05	**-0.32**	-0.20	**0.57**	**0.77**	**0.82**	**0.91**	**0.94**	**0.31**	**0.38**	**0.99**	**0.93**	**0.94**	**0.81**	**0.98**	**0.99**	**0.99**	0.99	**0.98**	**0.91**	**0.37**	**0.38**	**1.00**	1.00										
1530	0.06	**-0.35**	-0.15	**0.52**	**0.79**	**0.84**	**0.93**	**0.94**	**0.39**	**0.46**	**0.99**	**0.96**	**0.97**	**0.79**	**0.96**	**0.97**	**0.99**	0.98	**0.97**	**0.90**	**0.45**	**0.45**	**0.99**	**1.00**	1.00									
1540	0.06	**-0.35**	-0.14	**0.50**	**0.80**	**0.85**	**0.93**	**0.94**	**0.41**	**0.48**	**0.99**	**0.97**	**0.97**	**0.78**	**0.96**	**0.97**	**0.99**	0.98	**0.97**	**0.89**	**0.48**	**0.48**	**0.99**	**0.99**	**1.00**	1.00								
1580	0.07	**-0.35**	-0.08	**0.43**	**0.81**	**0.87**	**0.93**	**0.93**	**0.50**	**0.57**	**0.98**	**0.99**	**0.99**	**0.76**	**0.92**	**0.95**	**0.97**	0.97	**0.94**	**0.88**	**0.56**	**0.57**	**0.97**	**0.97**	**0.99**	**0.99**	1.00							
1780	0.08	**-0.33**	-0.08	**0.42**	**0.81**	**0.87**	**0.94**	**0.94**	**0.51**	**0.57**	**0.97**	**0.99**	**0.99**	**0.77**	**0.92**	**0.94**	**0.97**	0.97	**0.94**	**0.88**	**0.57**	**0.57**	**0.97**	**0.97**	**0.99**	**0.99**	**1.00**	1.00						
2000	0.02	-0.23	-0.20	**0.53**	**0.71**	**0.73**	**0.77**	**0.80**	0.16	0.22	**0.81**	**0.75**	**0.75**	**1.00**	**0.85**	**0.82**	**0.84**	0.84	**0.82**	**0.96**	0.22	0.22	**0.83**	**0.83**	**0.81**	**0.80**	**0.78**	**0.79**	1.00					
2080	0.04	-0.24	-0.25	**0.62**	**0.73**	**0.77**	**0.85**	**0.89**	0.17	0.24	**0.97**	**0.86**	**0.88**	**0.81**	**0.99**	**0.99**	**0.98**	0.98	**0.99**	**0.91**	0.23	0.23	**0.98**	**0.98**	**0.96**	**0.96**	**0.92**	**0.92**	**0.83**	1.00				
2100	0.04	-0.24	-0.22	**0.59**	**0.75**	**0.79**	**0.87**	**0.90**	0.21	0.28	**0.97**	**0.88**	**0.90**	**0.82**	**0.99**	**0.99**	**0.98**	0.98	**0.99**	**0.92**	0.27	0.27	**0.98**	**0.98**	**0.97**	**0.96**	**0.93**	**0.93**	**0.84**	**1.00**	1.00			
2250	0.05	-0.26	-0.17	**0.54**	**0.79**	**0.83**	**0.90**	**0.92**	**0.30**	**0.37**	**0.98**	**0.92**	**0.94**	**0.83**	**0.98**	**0.99**	**0.99**	0.99	**0.99**	**0.93**	**0.36**	**0.36**	**0.99**	**0.99**	**0.99**	**0.98**	**0.96**	**0.96**	**0.84**	**0.99**	**0.99**	1.00		
2270	0.06	-0.24	-0.17	**0.53**	**0.80**	**0.83**	**0.89**	**0.92**	**0.30**	**0.36**	**0.98**	**0.92**	**0.93**	**0.85**	**0.98**	**0.99**	**0.99**	0.99	**0.99**	**0.95**	**0.36**	**0.36**	**0.98**	**0.98**	**0.98**	**0.97**	**0.96**	**0.96**	**0.87**	**0.98**	**0.99**	**1.00**	1.00	
2280	0.04	-0.24	-0.17	**0.53**	**0.80**	**0.84**	**0.90**	**0.92**	**0.30**	**0.36**	**0.97**	**0.92**	**0.93**	**0.86**	**0.98**	**0.98**	**0.99**	0.99	**0.98**	**0.95**	**0.36**	**0.36**	**0.98**	**0.98**	**0.98**	**0.97**	**0.95**	**0.96**	**0.88**	**0.98**	**0.99**	**1.00**	**1.00**	1.00
2320	0.03	-0.23	-0.18	**0.55**	**0.77**	**0.81**	**0.88**	**0.90**	0.26	**0.32**	**0.96**	**0.89**	**0.90**	**0.86**	**0.98**	**0.98**	**0.98**	0.98	**0.99**	**0.95**	**0.32**	**0.32**	**0.97**	**0.97**	**0.96**	**0.96**	**0.93**	**0.93**	**0.88**	**0.98**	**0.99**	**0.99**	**0.99**	**0.99**

N=Nitrogen, NDVI=normalized difference vegetation index, REP=red edge position, SR=simple ratio

Table 8: Descriptive statistics of the measured foliar N and P concentrations

Nutrients (%)	Minimum	Maximum	Mean	Standard deviation	Coefficient of variation
Nitrogen	0.34	1.06	0.70	0.19	0.26
Phosphorus	0.04	0.29	0.11	0.05	0.49