Potential of Sentinel-2 spectral configuration to assess rangeland quality

Abel Ramoelo,[a,b,*] **Moses Cho,**[a,c] **Renaud Mathieu,**[a,d] **and Andrew K. Skidmore**[e]

[a]Council for Scientific and Industrial Research, Earth Observation Group, Natural Resources and Environment, P.O. Box 395, Pretoria 0001, South Africa

[b]University of Limpopo, Risk and Vulnerability Assessment Centre, Private Bag X1106, Sovenga 0727, South Africa

[c]University of KwaZulu-Natal, School of Agricultural, Earth and Environmental Sciences, Private Bag X01, Scottsville, Pietermaritzburg 3209, South Africa

[d]University of Pretoria, Department of Geography, Geoinformatics and Meteorology, Private Bag X20 Hatfield, Pretoria 0028, South Africa

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

Abstract. Sentinel-2 is intended to improve vegetation assessment at local to global scales. Today, estimation of leaf nitrogen (N) as an indicator of rangeland quality is possible using hyperspectral systems. However, few studies based on commercial imageries have shown a potential of the red-edge band to accurately predict leaf N at the broad landscape scale. We intend to investigate the utility of Sentinel-2 for estimating leaf N concentration in the African savanna. Grass canopy reflectance was measured using the analytical spectral device (ASD) in concert with leaf sample collections for leaf N chemical analysis. ASD reflectance data were resampled to the spectral bands of Sentinel-2 using published spectral response functions. Random forest (RF), partial least square regression (PLSR), and stepwise multiple linear regression (SMLR) were used to predict leaf N using all 13 bands. Using leave-one-out cross validation, the RF model explained 90% of leaf N variation, with a root mean square error of 0.04 (6% of the mean), which is higher than that of PLSR and SMLR. Using RF, spectral bands centered at 705 nm (red edge) and two shortwave infrared bands centered at 2190 and 1610 nm were found to be the most important bands in predicting leaf N. © *2015 Society of Photo-Optical Instrumentation Engineers (SPIE)* [DOI: 10.1117/1.JRS.9.094096]

Keywords: Sentinel-2; leaf nitrogen; WorldView-2; RapidEye; red-edge band; random forest.

Paper 15185P received Mar. 9, 2015; accepted for publication Jul. 8, 2015; published online Aug. 7, 2015.

1 Introduction

Copernicus, previously known as the Global Monitoring for Environment and Security, is a European Union initiative established to facilitate the European capacity to provision and use operational monitoring information for the environment and security. Within the Copernicus program, the European Space Agency (ESA) is responsible for the development of the "Space Component," a fully operational space-based capability to supply earth-observation data to sustain environmental information services.[1] ESA has embarked on the development of the Sentinel constellation, called the Sentinels (1 to 5). Among the Sentinels, Sentinel-2 comprises a high-resolution optical sensor to provide SPOT or Landsat type data for services such as land management, including for the agricultural industry (food security), forestry, disaster control, and humanitarian relief operations.[2]

It is expected that this constellation will contribute to monitoring essential climate[3] and biodiversity variables.[4] For land surface, the essential climate and biodiversity variables include leaf

*Address all correspondence to: Abel Ramoelo, E-mail: aramoelo@csir.co.za

area index, fraction of photosynthetically active radiation, biomass, fire disturbance (burned area), land cover, fraction of vegetation cover, vegetation types, leaf chlorophyll, and leaf water content.[1,5] This list of variables is not only important for climate and biodiversity related monitoring, but also for local and global vegetation assessment. Sentinel-2 can achieve these because of its spectral band configurations and improved spatial resolution. The spectral configuration of Sentinel-2 is comparable with some commercial satellite data such as WorldView-2 and RapidEye because of the presence of the red-edge band and is made even better by incorporating the short-wave infrared bands. The common characteristic of Sentinel-2 and the commercial satellites is only in the inclusion of the red-edge band, which is important for vegetation assessment and monitoring.

Leaf nitrogen (N) is one of the interesting variables for understanding the vitality and quality of vegetation. Leaf N is known to relate to protein content in the plants,[6,7] which is a major nutrient requirement for animals.[8] Leaf N is an indicator of rangeland quality and is crucial for the planning and management of grazing areas, e.g., determination of carrying capacity, spatial zoning, and grazing camps. Rangeland quality information is important for understanding ecosystem function and services. Well managed and improved grazing areas lead to high livestock production, which is a pillar of the rural economy and livelihoods. Proper management of the communal grazing lands in the rural areas is more crucial than relying on costly improved or irrigated grasslands for grazing. Conventional means of assessing leaf N over wider areas is costly in terms of labor and it is time consuming, for example, to cover every square meter of the ground.[9] Remote sensing provides a proper and more efficient option for wide area assessment of leaf N.

Attempts to estimate leaf N have been successful using field and imaging spectroscopy or airborne hyperspectral data.[10–12] Using hyperspectral technologies, leaf N has been estimated with acceptable accuracies using statistical analysis based on (1) vegetation indices, (2) full spectra, (3) absorption features, and (4) integrated modeling (combining indices and environmental variables).[11–14] The challenge is that the hyperspectral imagery is only available at a local scale and is costly. Regional assessment of leaf N has been lagging behind because of the paucity of satellite data with spectral configurations which are appropriate to detect subtle leaf N variation. Sensors such as WorldView-2 and RapidEye are among the first operational satellites with a red-edge band. The red-edge band is known to have a positive relationship with chlorophyll and nitrogen.[15] The use of the red-edge band for the computation of vegetation indices is feasible for estimating leaf N in peak productivity, when the interaction between biomass and leaf N is minimal.[16,17] On the other hand, Sentinel-2 has a combination of the red edge and shortwave infrared bands and can accurately estimate leaf N during peak productivity (wet season).

Several statistical analysis techniques were used to estimate vegetation parameters such as leaf N. Common and simple statistical techniques for estimating leaf N are simple and include stepwise multiple linear regression (SMLR).[16] The latter methods suffer from overfitting and multicollinearity.[18,19] To date, there are a series of machine learning techniques such as random forest (RF)[20] and artificial neural network (ANN) that are also applicable.[13,17] These latter techniques were found to be robust and they circumvent the overfitting and multicollinearity problem when estimating vegetation parameters. Traditionally, machine learning techniques such as RF and ANN were applied more for classification than for regression analysis.[21] RF was found to be more robust than other parametric regression techniques for estimating vegetation parameters.[20] This study intends to test the applicability of all Sentinel-2 bands combined with the conventional and machine learning techniques such as RF to estimate leaf N. The study focused on the use of full Sentinel-2 bands, because vegetation indices only use the visible and near-infrared red regions and exclude short-wave infrared. To our knowledge, Sentinel-2 spectral configurations were not previously assessed to estimate leaf N as an indicator of rangeland quality. The objective of the study was to assess a potential for the spectral configuration of Sentinel-2 in estimating leaf N as an indicator of rangeland quality in the savanna ecosystem.

2 Materials and Methods

2.1 *Study Area*

The study area is in the northeastern part of South Africa in the Lowveld savanna (Fig. 1). The Lowveld landscape is a low-lying area extending from the foot slopes of the Drakensberg Great

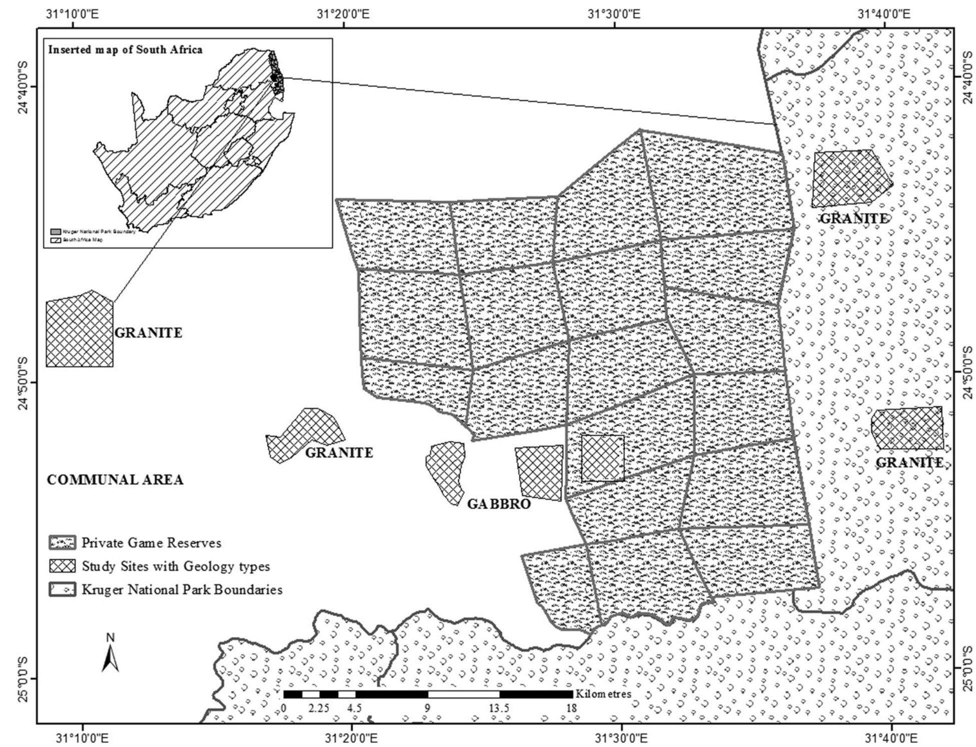

Fig. 1 Study area locality map with the indication of sites and their geological types.

Escarpment to the west and the Mozambique coastal plain to the east.[22] The topography is gently undulating with flat patches in localized areas and with an average height of 450 m above sea level. The study area covers a land use transect ranging from protected areas such as the privately 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 (800 mm/year.) as compared with the eastern side of the transect (580 mm/year).[22] The annual mean temperature is about 22°C. The dominant geology includes granite and gneiss with local intrusions of gabbro.[22]

The gradients of soil moisture and nutrients are important in this area. The soil fertility of the gabbro areas is higher than that of the granitic ones.[22,23] The main vegetation communities include the "granitic lowveld" and the "gabbro grassy bushveld."[24] In the gabbro patches, grass species such as *Setaria sphacelata* dominate the crest while species such as *Urochloa mosambicensis* dominate the valleys. The gabbro patches are dominated by grass species with a high productive potential (e.g., *U. mosambicensis*) compared with the granite-derived soils (e.g., *Eragrostis rigidior* and *Pogonarthria squarrosa*). The gabbro sites are dominated by fine-leaved tree species such as *Acacia* spp., while the granite sites are dominated by broadleaf tree species such as *Combretum* spp. and *Terminalia* spp.[22,25]

2.2 *Field Data Collection*

The study area consisted of seven experimental sites which were placed along the land use gradient: two sites in KNP (granite), one site in SGR (gabbro), and four sites in the communal areas (gabbro, granite) (Fig. 1). The demarcation of the sites was done using 1 : 250,000 geology maps and 2008 SPOT 5 images.[26] The site selection process captured the nutrient contrast from low to high between granitic-derived soils to gabbro-derived soils, respectively, and along the rainfall gradient. A line transect sampling design was used to collect field data[27] in each site. To further ensure grass biomass variability, transects were laid out to sample both valley and crest land units. Grass biomass in the savanna ecosystems is also influenced by topography with valley areas generally having higher grass biomass than crest areas. Along transects, a combination of purposeful and systematic placement of the sampling plots was done. The distance between the

Table 1 Days for field sampling and spectral measurements.

Sample name	Date	Geology types	Land use	Measurements[a]
L1 (G1-5)	April 17, 2009	Granite	Protected area	X
L2 (G1-4)	April 01, 2009	Granite	Protected area	X
L2 (G4-5)	April 08, 2009	Granite	Protected area	X
L4 (G1-6)	April 15, 2009	Gabbro	Communal area	X
L4 (G7-10)	April 16, 2009	Gabbro	Communal area	X
L5 (G1-5)	April 18, 2009	Gabbro	Communal area	X
L6 (G1-5)	April 07, 2009	Gabbro	Communal area	X
L7 (G1-6)	April 06, 2009	Granite	Communal area	X
L8 (G1-4)	April 02, 2009	Granite	Communal area	X
L8 (G5-6)	April 03, 2009	Granite	Communal area	X
L8 (G7-9)	April 04, 2009	Granite	Communal area	X

[a]Field sampling and spectral measurements were done concurrently.

plots was between 500 and 1000 m, depending on the accessibility and homogeneity of the area. The plot size was 30 m × 30 m. A total of 49 plots was surveyed and in each plot three to four subplots (0.5 m × 0.5 m) were randomly selected to capture the plot's variability. In each subplot, data on the sample location were collected using the Leica®'s GS20 differential geographic positioning system (DGPS), as well as the dominant grass species and grass samples. Grass samples were dried at 80°C for 24 h. The DGPS points were postprocessed using Leica's GeoPro software and reference GPS data from the Nelspruit station to produce a less than 1 m positional accuracy. The fieldwork was undertaken in April 2009 toward the end of the wet season (Table 1), when the grass biomass was at full maximum growth or peak productivity to minimize the N and biomass interaction effects.[17]

2.3 *Chemical Analysis*

The dried grass samples were chemically analyzed at the South Africa's Agricultural Research Council-Institute for Tropical and Subtropical Crops-Nelspruit. First, the acid digestion technique was used, where sulfuric acid was used for retrieving foliar N concentrations.[28–30] Second, the colorimetric method by autoanalyzer 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 color was formed by the reaction of ammonia, sodium salicylate, sodium nitroprusside, and sodium hypochlorite. The ammonia-salicylate complex was read at 640 nm. These extraction methods were successfully used for leaf N retrieval.[11,12]

2.4 *Spectral Measurements*

The reflectance spectra were measured in the field using an Analytical Spectral Device (ASD) spectroradiometer, Fieldspec 3®. The ASD spectral domain ranges from 350 to 2500 nm, with a 1-nm band width. Within each plot, spectral measurements were made for each of the three to four 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.[31,32] The measurements were taken between 10:30 a.m. and 3:00 p.m. on clear sunny days to minimize cloud effects and maximize illumination.[33] A 25-deg field-of-view fiber optic was used. The fiber optic pistol was held at 1 m 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 the spectral radiance to reflectance.[11]

2.5 *Data Analysis*

A test for normality of leaf N was done using the Shapiro–Wilk normality test implemented in R.[34] The Shapiro–Wilk normality test was selected because it is commonly used.[11] The spectral measurement data acquired using the ASD were convolved to the spectral band configuration of Sentinel-2 using the expected spectral response function of Sentinel-2.[35] Band centers and widths are presented in Table 2.

To predict leaf N and assess the Sentinel-2 data, several regression techniques were used, RF,[36] SMLR, principal component analysis (PCA), and partial least square regression (PLSR). SMLR, PCA, and PLSR have been widely used for predicting vegetation parameters, including leaf N.[12,14,16] RF was implemented from the RF package programmed in the R statistical environment.[37] RF technique was successfully used with remote sensing to predict wetland species biomass,[20] plant water content,[38] and for species classification.[21,39,40] RF is a machine learning method developed to improve the classification and regression trees method (CART) by using a large set of decision trees. RF builds each tree by using a deterministic algorithm selecting a random set of variables and a random sample from the calibration datasets. There are three main variables that need to be parameterized: *ntree*, number of regression trees grown based on a bootstrap sample of observation (the default value is 500 trees); *mtry*, number of predictors tested at each node (default is the square root of the total number of variables); and *nodesize*, minimal size of the terminal nodes of the trees (the default value is one). These variables need to be well defined.[20] In this study, RF parameterizations were as follows: *ntree* used was a default value = 500, *nodesize* was a default value = 1, and *mtry* = 1. The optimization of the number of variables required to predict leaf N was determined using a recursive feature selection based on leave-one-out cross validation (LOOCV)[41] and the root mean square error (RMSE). The conventional application of RF is outlined by Breiman[36] and Mutanga et al.[20] For RF, the variable of importance was used to determine the importance of each predictor by measuring the percentage increase in the mean square error when the out-of-bag data for each variable is permuted, while the others are unchanged.[20] These variable importance values are then used to rank the predictors in terms of their relationship to response variables. The higher the variable of an importance score or value, the higher is the importance of that particular variable in the model. For PLSR, the optimum numbers of factors or latent variables used for predicting leaf N were achieved through LOOCV, depicted with the lowest RMSE.[12] For SMLR, significant variables were

Table 2 Band centers and spectral width for the Sentinel-2 sensor.

Bands	Center λ_{center} (nm)	Spectral width $\Delta\lambda$ (nm)
1	443	20
2	490	65
3	560	35
4	665	30
5	705	15
6	740	15
7	783	20
8	842	115
8a	865	20
9	945	20
10	1375	30
11	1610	90
12	2190	180

selected using the Akaike Information Criterion (AIC), based on the lowest AIC value.[16] For PCA, 10 principal components were derived from 13 wavelengths because they explained the high variance of the data, and the first three components explained the most variance. Finally the SMLR process was undertaken to estimate leaf N, as per Ramoelo et al.[16]

The RF, SMLR, PCA, and PLSR model validations were done using LOOCV because of the small sample size. In cross validation, prediction of individual samples is estimated by the remaining samples. For example, if there are 20 samples, each sample will be predicted by 19 samples iteratively to determine the performance of the model. Merits of the cross validation are the capability to detect outliers and the ability to provide unbiased assessment of the prediction error.[42] The statistical measures of precision and accuracy such as the coefficient of determination (R^2) and RMSE were determined.

3 Results and Discussions

The variability of leaf N is relatively low, with a coefficient of variation (CV) equal to 26%. The average leaf N is 0.7, which is relatively low for grasses (see Table 3). The normality test was done using the Shapiro–Wilk normality test implemented in R, and the results indicated that leaf N is not normally distributed ($W = 0.9816$, $p = 0.6347$). Leaf N distribution is influenced by a number of factors, including geology and soil types, topography, fire regimes, and climatic factors.[14,22,43] For example, rainfall gradient from east (low) to west (high) could have played a major role in the distribution of leaf N, and hence, the high variability of leaf N.

The Sentinel-2 convolved reflectance data and the RF model explained 90% of the variation of leaf N with an RMSE of 0.04%, which is 6% of the mean (Fig. 2). The SMLR model explained about 61% of leaf N variation, which is better than the PCA and PSLR models. RF is a nonparametric model which makes no assumption about the data distribution. By the nature of the model, a well-optimized RF minimizes overfitting and multicollinearity when predicting variables. These advantages make RF a predictor of superior performance and popularity than the parametric predicting models. The performance of RF was tested for predicting several vegetation parameters and showed superior results, e.g., wetland species biomass[20] and plant water content.[38] The study used LOOCV as means for validating the models because of the limited number of samples. Several studies used cross-validation techniques to evaluate models and these are said to be an unbiased means of validation.[41] Figure 2 also shows that lower values were overestimated, while higher values were underestimated. Mutanga et al.[20] discovered similar patterns especially for higher values of biomass and indicated that there is a need to further investigate this phenomenon. It is indeed true that further investigation is required to solve this problem. The underperformance of PCA and PLSR could be due to the fact that only 13 bands were used. Though almost 50% of the factors were optimally selected, the limited number of bands could have influenced the results. PLSR and PCA techniques are developed for the analyses with a high dimensionality of data, e.g., field and imaging spectroscopy. Ramoelo et al.[11] and Huang et al.[10] found out that PLSR yielded a higher prediction capability as compared with SMLR. In this study, SMLR yielded a higher prediction accuracy as compared with PLSR solely because of the low dimensionality of data, which reduces the chances of overfitting and multicollinearity.

Two of the most important variables for predicting leaf N is the red-edge band centered at 705 nm and the SWIR bands centered at 2190 and 1610 nm based on the RF model (Fig. 3). SMLR picked similar bands excluding 1610 nm, including the second red-edge band. Hyperspectral studies demonstrated the utility of the red-edge inflation point for estimating leaf biochemicals.[15,44] This information is captured by the red-edge band featured in the multispectral

Table 3 Descriptive analysis of leaf nitrogen.

	Min (%)	Max (%)	Average (%)	SD	CV (%)
Leaf nitrogen	0.34	1.06	0.7	0.19	26

SD = standard deviation; CV = coefficient of variation.

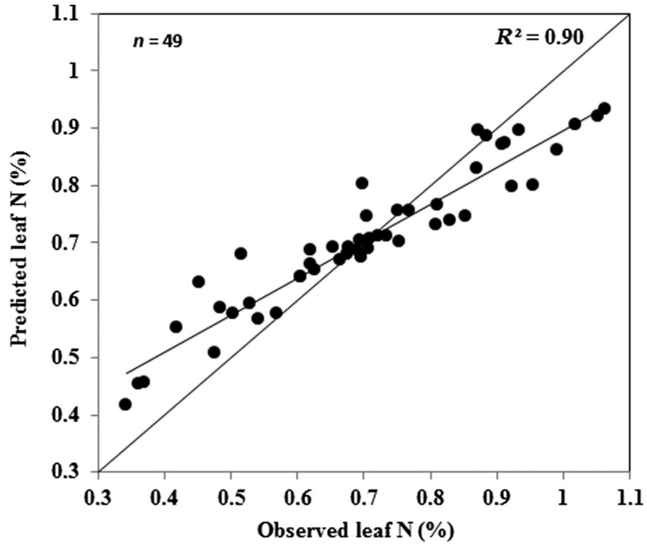

Fig. 2 Leaf N prediction capability of Sentinel-2 and random forest (RF), which is the best performing model.

sensors. In the multispectral sensors such as RapidEye, the red-edge band was crucial for assessing leaf N at a regional scale.[16] The red-edge band is related to leaf chlorophyll.[15] On the other hand, leaf chlorophyll is related to leaf N.[15,45] The second red-edge band centered at 740 nm was the least important variable for predicting leaf N. The first red-edge band (705 nm) is positioned at the onset of the high reflectivity portion of the vegetation response and is crucial for plant health estimation (e.g., leaf N or chlorophyll), while the second red-edge band is influenced by the combined effects of plant health (e.g., leaf N or chlorophyll) and the vegetation structure (e.g., leaf area index and biomass).[15,46]

Using RF, two Sentinel-2 shortwave infrared bands centered at 1610 and 2190 nm, were selected as important bands (Fig. 3). The 2190 nm band was also selected using SMLR (Table 4). According to the list of the known absorption features in Curran,[18] 1610 nm is close to the feature around 1690 nm, which is related to lignin, starch, and protein, while 2190 nm is close to the feature around 2180 nm, which is associated with protein or nitrogen. The 2190 nm band is more important than 1610 nm band as per ranking in Fig. 3, because the 2180 nm feature is mainly associated with protein and nitrogen. Leaf N is positively related to protein.[47] These absorption features have been successful in the estimation of leaf N using field and imaging spectroscopy.[11,17,19]

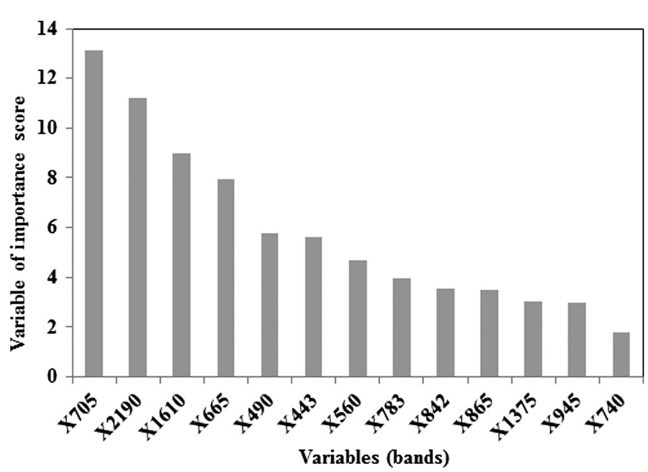

Fig. 3 Important variables in predicting leaf N using Sentinel-2 simulated data and RF.

Table 4 Performance of various statistical techniques with Sentinel-2 bands in estimating leaf N.

	R^2	RMSE (%)	RRMSE (%)	$P < 0.05$	No. of factors or bands (nm)
RF	0.90	0.04	6	Yes	705; 1610; 2190
SMLR	0.61	0.12	17	Yes	665; 705; 740;783; 842; 2190
PCA + SMLR	0.49	0.14	20	Yes	7
PLSR	0.46	0.14	20	Yes	5

RF = random forest; SMLR = stepwise multiple linear regression; PCA = principal component analysis; PLSR = partial least square regression.

One to two decades of field and imaging spectroscopy or hyperspectral studies led to the emergence of new sensors such as the planned Sentinel-2. These sensors (constellation, planned 2 satellites) are unique because of the strategic placement and configurations of the wavelengths to characterize or sample the radiation. As with the commercial sensors on-board RapidEye and WorldView-2, Sentinel-2 is planned to have two red-edge bands. In addition to this, two strategically placed bands in the short-wave infrared make it unique as compared with commercial sensors such as RapidEye and WorldView-2. Interestingly, the strategically placed bands are crucial for the assessment of the rangeland quality. Assessment of leaf N with high accuracy also proves that it could be useful for crop health assessment. According to ESA, Sentinel-2 is likely to be free for Africa, and this will provide opportunities to assess leaf N from regional to global levels.

4 Conclusions

This study demonstrated a potential for simulated Sentinel-2 reflectance data to estimate leaf N using RF. Cross-validated statistics from RF showed that Sentinel-2 simulated data can explain about 90% of leaf N variation. SMLR achieved the second best results in the estimation of leaf N using Sentinel-2 data. PCA and PLSR yielded the lowest performances mainly because of the low dimensionality of their data. PCA and PLSR perform better when exploratory variables are in their tens to hundreds and even thousands, especially using spectroscopy or hyperspectral data. Using RF and SMLR, both red-edge bands and SWIR were selected, showing the importance of these bands in the multispectral data. Currently, red-edge bands are only available in commercial sensors such as WorldView-2 and RapidEye, but Sentinel-2 shall provide ease of access of these data for vegetation assessment. Sentinel-2 will provide data with similar spectral configurations as WorldView-2 and RapidEye, which are cutting edge technology, and this will facilitate the estimation of leaf N as an indicator of rangeland quality at a wider scale and in a cost effective manner. The emergence of Sentinel-2 data shall provide an opportunity to assess grass quality for the planning and management of rangeland systems.

Acknowledgments

Authors would like to thank the Council for Scientific and Industrial Research, Department of Science and Technology (DST), National Research Foundation's Thuthuka program and University of Twente's Faculty of Geo-information Science and Earth Observation (UT-ITC), and Wageningen University for funding. We would like to thank the South African National Parks (SANPARKS) for facilitating our field work, especially Dr. Izak Smith, Mrs. Patricia Khoza, and field guards Mr. Veli Ndlovu and Mrs. Onica Sithole. We are also grateful to Mr. Mike Glover for facilitating field work in Sabi Sands. Finally, we are grateful to Mr. Laven Naidoo, Mr. Thulani Selaule, and Mr. Russell Main for the field work assistance.

References

1. M. Drusch et al., "Sentinel-2: ESA's optical high-resolution mission for GMES operational services," *Remote Sens. Environ.* **120**, 25–36 (2012).

2. J. Aschbacher and M. P. Milagro-Pérez, "The European Earth monitoring (GMES) programme: status and perspectives," *Remote Sens. Environ.* **120**, 3–8 (2012).
3. GCOS, "Implementation plan for the global observing system for climate in support of the UNFCC," WMO [Vol.GCOS-138 (GOOS-184, GTOS-76, WMO-TD/No.1523)], http://www.wmo.int/pages/prog/gcos/publications/gcos-138.pdf (20 July 2014).
4. H. M. Pereira et al., "Essential biodiversity variables," *Science* **339**(6117), 277–278 (2013).
5. Z. Malenovský et al., "Sentinels for science: potential of Sentinel-1, -2, and -3 missions for scientific observations of ocean, cryosphere, and land," *Remote Sens. Environ.* **120**, 91–101 (2012).
6. K. E. Clifton, J. W. Bradbury, and S. L. Vehrencamp, "The fine-scale mapping of grassland protein densities," *Grass Forage Sci.* **49**(1), 1–8 (1994).
7. Z. J. Wang et al., "The prediction of grain protein in winter wheat (Triticum aestivum) using plant pigment ratio (PPR)," *Field Crop Res.* **90**, 311–321 (2004).
8. H. H. T. Prins and J. Beekman, "A balanced diet as a goal for grazing: the food of the Manyara buffalo," *Afr. J. Ecol.* **27**, 241–259 (1989).
9. R. R. Pullanagari et al., "Remote sensing of pasture quality," in *Proc. 22nd International Grasslands Congress 2013*, Sydney, Australia, 15–19 September, pp 6 (2013).
10. Z. Huang et al., "Estimating foliage nitrogen concentration from HYMAP data using continuum removal analysis," *Remote Sens. Environ.* **93**(1–2), 18–29 (2004).
11. A. Ramoelo et al., "Non-linear partial least square regression increases the estimation accuracy of grass nitrogen and phosphorus using in situ hyperspectral and environmental data," *ISPRS J. Photogramm.* **82**, 27–40 (2013).
12. A. Ramoelo et al., "Water-removed spectra increase the retrieval accuracy when estimating savanna grass nitrogen and phosphorus concentrations," *ISPRS J. Photogramm.* **66**(4), 408–417 (2011).
13. N. M. Knox et al., "Dry season mapping of savanna forage quality, using the hyperspectral Carnegie Airborne Observatory sensor," *Remote Sens. Environ.* **115**(6), 1478–1488 (2011).
14. N. M. Knox et al., "Remote sensing of forage nutrients: combining ecological and spectral absorption feature data," *ISPRS J. Photogramm.* **72**, 27–35 (2012).
15. M. A. Cho and A. K. Skidmore, "A new technique for extracting the red edge position from hyperspectral data: the linear extrapolation method," *Remote Sens. Environ.* **101**(2), 181–193 (2006).
16. A. Ramoelo et al., "Regional estimation of savanna grass nitrogen using the red-edge band of the spaceborne RapidEye sensor," *Int. J. Earth Obs.* **19**, 151–162 (2012).
17. A. K. Skidmore et al., "Forage quality of savannas—simultaneously mapping foliar protein and polyphenols for trees and grass using hyperspectral imagery," *Remote Sens. Environ.* **114**(1), 64–72 (2010).
18. P. J. Curran, "Remote sensing of foliar chemistry," *Remote Sens. Environ.* **30**(3), 271–278 (1989).
19. R. F. Kokaly et al., "Characterizing canopy biochemistry from imaging spectroscopy and its application to ecosystem studies," *Remote Sens. Environ.* **113**(Suppl. 1), S78–S91 (2009).
20. O. Mutanga, E. Adam, and M. A. Cho, "High density biomass estimation for wetland vegetation using WorldView-2 imagery and random forest regression algorithm," *Int. J. Earth Obs.* **18**, 399–406 (2012).
21. E. M. Adam et al., "Discriminating the papyrus vegetation (Cyperus papyrus L.) and its co-existent species using random forest and hyperspectral data resampled to HYMAP," *Int. J. Remote Sens.* **33**(2), 552–569 (2012).
22. F. J. Venter, R. J. Scholes, and H. C. Eckhardt, "Abiotic template and its associated vegetation pattern," in *The Kruger Experience: Ecology and Management of Savanna Heterogeneity*, J. T. Du Toit, H. R. Kevin, and H. C. Biggs, Eds., Vol. 83–129, The Island Press, London (2003).
23. R. Ben-Shahar and M. J. Coe, "The relationships between soil factors, grass nutrients and the foraging behaviour of wildebeest and zebra," *Oecologia* **90**(3), 422–428 (1992).
24. L. Mucina and M. C. Rutherford, *The Vegetation of South Africa, Lesotho and Swaziland*, South African National Biodiversity Institute (SANBI), Cape Town, Strelitzia (2006).

25. J. G. Ferwerda et al., "Parent material and fire as principle drivers of foliage quality in woody plants," *Forest Ecol. Manage.* **231**(1–3), 178–183 (2006).
26. K. J. Wessels et al., "Impact of communal land use and conservation on woody vegetation structure in the lowveld savannas of South Africa," *Forest Ecol. Manage.* **261**, 19–29 (2011).
27. R. M. Fewster, J. L. Laake, and S. T. Buckland, "Line transect in small and large regions," *Biometrics* **61**, 856–859 (2005).
28. H. C. Giron, "Comparison between dry ashing and wet digestion in preparation of plant material for atomic absorption analysis," *At. Absorpt. Newsl.* **12**(1), 28–29 (1973).
29. K. Grasshoff, M. Erhardt, and K. Kremling, Eds., *Methods of Seawater Analysis*, Verlag Cheime, Weinhein, Germany (1983).
30. O. Mutanga et al., "Explaining grass-nutrient patterns in a savanna rangeland of southern Africa," *J. Biogeogr.* **31**(5), 819–829 (2004).
31. O. Mutanga, A. K. Skidmore, and S. E. Van Wieren, "Descriminating tropical grass (*Cenchrus ciliaris*) canopies grown under different nitrogen treatments using spectroradiometry," *ISPRS J. Photogramm.* **57**(3), 263–272 (2003).
32. Y. Wang et al., "Validation of artificial neural network techniques in the estimation of nitrogen concentration in rape using canopy hyperspectral reflectance data," *Int. J. Remote Sens.* **30**(17), 4493–4505 (2009).
33. E. M. Abdel-Rahman, F. B. Ahmed, and M. Van den Berg, "Estimation of sugarcane leaf nitrogen concentration using in situ spectroscopy," *Int. J. Earth Obs.* **12**(Suppl. 1), S52–S57 (2010).
34. P. Royston, "Algorithm AS 181: the W test for normality," *Appl. Stat.* **31**, 176–180 (1982).
35. ESA, "Sentinel-2 MSI Technical Guide (European Space Agency)" [cited 2014 08-04-2014], https://sentinel.esa.int/web/sentinel/sentinel-2-msi-wiki (2014).
36. L. Breiman, "Random forests," *Mach. Learn.* **45**, 5–32 (2001).
37. A. Liaw and M. Wiener, "Classification and regression by random forest," *R News* **2**, 18–22 (2002).
38. R. Ismail and O. Mutanga, "A comparison of regression tree ensembles: predicting Sirex noctilio induced water stress in *Pinus patula* forests of KwaZulu-Natal, South Africa," *Int. J. Earth Obs.* **12**(Suppl. 1), S45–S51 (2010).
39. J. Ham et al., "Investigation of the random forest framework for classification of hyperspectral data," *IEEE Trans. Geosci. Remote Sens.* **43**(3), 492–501 (2005).
40. L. Naidoo et al., "Classification of savanna tree species, in the Greater Kruger National Park region, by integrating hyperspectral and LiDAR data in a Random Forest data mining environment," *ISPRS J. Photogramm. Remote Sens.* **69**, 167–179 (2012).
41. R. Diaz-Uriarte and S. Alvarez de Andres, "Gene selection and classification of microarray data using random forest," *BMC Bioinf.* **7**(1), 3 (2006).
42. B. Efron and G. Gong, "A leisurely look at the bootstrap, the jackknife, and cross-validation," *Am. Stat.* **37**(1), 36–48 (1983).
43. C. C. Grant et al., "Nitrogen and phosphorus concentration in faeces: an indicator of range quality as a practical adjunct to existing range evaluation methods," *Afr. J. Range Forage Sci.* **17**, 81–92 (2000).
44. P. J. Curran et al., "The effect of a red leaf pigment on the relationship between red edge and chlorophyll concentration," *Remote Sens. Environ.* **35**(1), 69–76 (1991).
45. B. J. Yoder and R. E. Pettigrew-Crosby, "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).
46. M. Cho and A. K. Skidmore, "Hyperspectral predictor for monitoring biomass production in Mediterranean mountain grasslands, Majella National Park, Italy," *Int. J. Remote Sens.* **30**(2), 499–515 (2009).
47. Z. J. Wang et al., "The prediction of grain protein in winter wheat (Triticum aestivum) using plant pigment ratio (PPR)," *Field Crops Res.* **90**, 311–321 (2004).

Abel Ramoelo is a senior remote sensing researcher at the Council for Scientific and Industrial Research (CSIR) and is a research associate at the University of Limpopo, South Africa. He has a

PhD in remote sensing completed in 2012 at the University of Twente—Faculty of Geo-Information Science and Earth Observation (ITC). His work focuses on development of remote sensing-based techniques to estimate biochemical (e.g., leaf nitrogen and chlorophyll) and biophysical (e.g., biomass and leaf area index) parameters at various spatial and temporal scales.

Moses Cho is a principal research scientist at the CSIR and a research fellow at the University of KwaZulu Natal South Africa. He completed his PhD in hyperspectral remote sensing from Wageningen University, The Netherlands, and the International Institute for Geo-information Science and Earth Observation, Enschede, The Netherlands, in 2007. His research is centered on the development of remote sensing algorithms for modeling and mapping vegetation biochemistry and biophysical properties and for discriminating vegetation types or species.

Renaud Mathieu is a principal researcher with the Council for Scientific and Industrial Research, Pretoria, South Africa, and leads the Earth Observation Research Group in Natural Resources and Environment. He received his PhD in geographic information sciences from Université de Marne-la-Vallée, Paris, France. His research interests focus on the application of remote sensing technologies to soil and water conservation, natural resource and wildlife management, environment, and agriculture.

Andrew K. Skidmore is head of the Department of Natural Resources at the Faculty of Geo-Information Science and Earth Observation (ITC) of the University of Twente, The Netherlands. He has a PhD in remote sensing and GIS at the Australian National University. Vegetation mapping and monitoring have been his main ongoing research theme, while recent research includes wildlife habitat assessment in East Africa, hyperspectral remote sensing, all techniques for handling geoinformation, and accuracy assessment.