DRY SEASON BIOMASS ESTIMATION AS AN INDICATOR OF RANGELAND QUANTITY USING MULTI-SCALE REMOTE SENSING DATA

Abel Ramoelo[1,2], Moses A. Cho[1]

1. Earth Observation Group, Natural Resource and Environment, Council for Scientific and Industrial Research (CSIR), South Africa. abel.ramoelo@gmail.com
2. Risk and Vulnerability Assessment Center, University of Limpopo, Sovenga, South Africa

KEYWORDS: Biomass, RapidEye, Landsat, remote sensing, rangeland quantity

ABSTRACT

For grazing, biomass is the main indicator of rangeland quantity, which is crucial to determine the amount of food available for animals (grazers), including livestock. Livestock production in the rural communities of the world, including Africa, is the main source of income and hence livelihood. Biomass information during dry season is not only important for grazing but also for determining the fuel load for fire risk. During dry season, grazers are mainly limited by grass quantity than quality. Therefore, it is important to quantify the variability of biomass during dry season to inform decision makers on planning and management of the grazing systems. Remote sensing provides opportunity to successfully estimate biomass in natural and agricultural areas. The conventional approach makes use of the vegetation indices such as the normalized difference vegetation index (NDVI), which is a measure of vegetation greenness. The use of vegetation indices has been successful during wet periods where vegetation is green and photosynthetic active. During dry season, biomass estimation is always not plausible using vegetation indices. The aim of this study is to estimate dry biomass using the multi-scale remote sensing data in the savanna ecosystem. Field data was collected in August 2013, and concerted to the acquisition of the satellite image from RapidEye and Landsat 8. Random forest algorithm (RF) was used to predict biomass using the band reflectance data, from RapidEye and Landsat 8 respectively. The results show that RF combined with RapidEye explained over 85% of biomass variation, as compared to 81% explained by RF with Landsat 8 data. For regional assessment of biomass as an indicator of rangeland quantity, high spatial resolution data can be used for calibration and validation. This study demonstrates that dry season biomass can be estimated using remote sensing, and it is important for understanding grazing and feeding patterns of animals, including livestock and wildlife.

INTRODUCTION

Biomass is the main indicator of rangeland quantity, which is crucial to determine the amount of food available for animals (grazers), including livestock. Livestock production in the rural communities of the world, including Africa, is the main source of income and hence livelihood (Shackleton et al. 2002). Rangeland quantity influences the feeding patterns and movements of grazers. Biomass information in dry season is not only important for grazing but also for determining the fuel load for fire risk. During dry season, grazers or grazing animals are mainly limited by grass quantity than quality. Therefore, it is important to quantify the variability of biomass during dry season to inform decision makers on planning and management of the grazing systems. Nowadays, global change including land cover/ land use and climate change due to new or existing tenure systems, drought and erratic rainfall influence land degradation. In addition, poor planning of land use is one of the major drivers of land degradation. Information on the spatial distribution of grass biomass is crucial for proper planning and management of the rangeland systems.

For the past three to four decades, remote sensing provided opportunity to successfully estimate biomass in natural and agricultural areas (Tucker 1977, Tucker and Sellers 1986, Xu et al. 2014). The main approach was the use of the vegetation indices such as the normalized difference vegetation index (NDVI) (Rouse et al. 1974, Todd et al. 1998), which is a measure of vegetation greenness. Biomass estimation using vegetation indices such as NDVI was achieved by using empirical models. The empirical models are simple to implement, but site, season and data specific. The empirical models for predicting biomass using vegetation indices has been successful during wet periods where vegetation is green and photosynthetic active (Mutanga et al. 2012, Ramoelo et al. 2012). Mutanga et al. (2012) used WorldView-2 data and random forest algorithm to reduce the saturation problem in estimating biomass during peak productivity. Ramoelo et al. (2012) successfully estimated biomass using the integrated modelling approach, which combines remote sensing indices and environmental variables to minimize the saturation problem during peak productivity or wet season. During peak productivity, the saturation problem occurs when the amount of light that can be absorbed in the red region of the spectrum reaches a plateau (Tucker 1977), while the NIR reflectance continue to increase because addition of new leaves which influences the multiple scattering within the canopy (Kumar et al. 2001). During dry season, saturation might not a problem for biomass estimation because of the limited dependence on vegetation indices. Using hyperspectral data during dry season, biomass estimation can be achieved using short-wave infrared index such as Cellulose Absorption Index (CAI) (Nagler et al. 2003, Xu et al. 2014). Specific spectral bands used for CAI computations do not exist in most of the satellite sensors including RapidEye and Landsat 8. In addition, conventional vegetation indices are difficult to use in the dry season, for grass cover and biomass prediction (Xu et al., 2014). In this study, original reflectance data in combination with the random forest statistical technique were used to predict dry season biomass.

MATERIAL AND METHODS

Study area and data collection

The study area is located in the north-east of South Africa and covers part of the Kruger National Park (KNP), SabiSands and Bushbuckridge communal rangelands. Field work was undertaken in August 2013, the same month with the acquisition of the RapidEye satellite image. Sites along the main roads covering the study area were purposively selected for the field sampling, also considering the underlying geological strata. The road sampling technique was preferred since penetration into the savanna landscape was constrained by management and logistical restrictions. Buffers of 300 m were created on both sides of these roads using ArcGIS software (ESRI, USA). Within the buffer polygons random sample points were generated using the ArcGIS. Once in the field location, 61 plots were placed in relatively large areas with homogeneous grass to avoid the possible contamination of the tree signal on the grass signal. In each plot of 20m x20 m, two subplots of 0.5m x 0.5m were randomly placed and the grass samples were cut and weighed to determine green or wet biomass (g/m^2). Grass materials were dried at 80^0C for 24 hours and weighed, henceforth referred to as biomass.

The RapidEye and Landsat 8 data were acquired in August 2013. The sensor is a multispectral push broom imager with a spatial resolution of 6.25 m and captures data in the spectral bands: blue (440-550 nm), green (520-590 nm), red (630-685 nm), red edge (690-730 nm), and near infrared (760-850 nm). Surface reflectance data were retrieved using the atmospheric and topographic correction software (ATCOR 2) implemented in the IDL Virtual Machine (Richter 2011). ATCOR 2 models reflectance for flat surfaces was considered sufficient because the study area is not characterized by very rugged terrain. Since the modules for Landsat 8 were not implemented in the ATCOT 2, Quick Atmospheric Correction implemented in ENVI was used. Before QUICK atmospheric correction was implemented, digital number (DN) values were converted to top-of-atmosphere (TOA) radiance. First seven Landsat bands were used for further analysis.

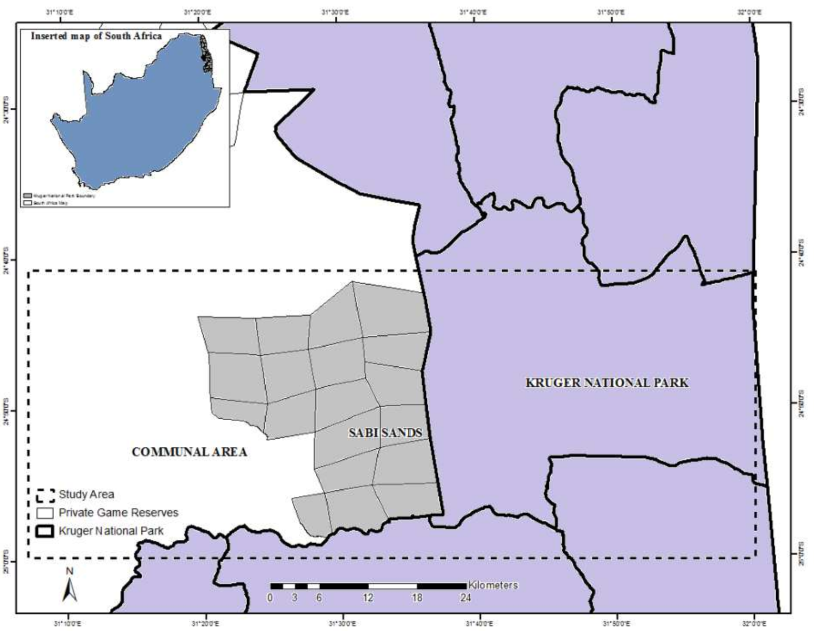

Figure 1: the performance of random forest in biomass estimation using RapidEye data

Data analysis

Reflectance data from the RapidEye and Landsat were extracted corresponding to each sampling point where the field measured biomass data was done. The first set of analysis for predicting biomass was done using Random forest and RapidEye reflectance data. The second of analysis was performed by using RF and Landsat reflectance data. Random forest was implemented from the random forest package programmed in the R statistical environment (Liaw and Wiener 2002). This technique was successfully used with remote sensing to predict wetland species biomass (Mutanga et al. 2012), plant water content (Ismail and Mutanga 2010) and for species classification (Ham et al. 2005, Adam et al. 2012). RF is a machine learning method developed to improve the classification and regression trees method (CART) by using 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 data sets. For more details about the optimization and implementation of Random forest for prediction of vegetation parameter, see Mutanga et al. (2012). The validation of the model was done using leave-one-out cross validation (LOOCV). In cross validation, samples are 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. Advantages of the cross-validation are the capability to detect outliers and provide unbiased assessment of the prediction error (Efron and Gong 1983). The statistic measure of precision and accuracy such as the coefficient of determination (R^2) and root mean square error (RMSE) were determined.

RESULTS AND DISCUSSIONS

There is a moderate variation of grass biomass with coefficient variation equalling 46%. The highest biomass value was recorded to be 344 g/m^2 and 52 g/m^2 as the lowest. The mean biomass was 126 g/m^2 (Table 1). The dry season measurement of biomass is significantly lower than the wet season. The distribution of biomass in the savanna ecosystem is influenced several edaphic and biophysical factors (Venter et al. 2003). Among other factors, fire, rainfall and geology influence the biomass distribution. For example, basalt geology is known to have high biomass as compared to the granite (Ramoelo et al. 2012).

Table 1: Descriptive statistics for grass biomass

Min (g/m²)	**Max (g/m²)**	**STDEV**	**Mean (g/m²)**	**CV (%)**
52	344	58.28	126.46	46.09

CV = coefficient of variance, STDEV = standard deviation

The predictive models based on random forest (RF), Landsat and RapidEye significantly explained over 80% of biomass variation in the study area. The results further indicate that higher accuracies for predicting biomass have been achieved using RapidEye data as compared to the Landsat (Table 2, Figure 2-3). Using RapidEye data, prediction accuracy attained was 13.42 g/m^2 which is about 10.61% of the mean, as compared to the 15.79 g/m^2 (12.49% of the mean) of Landsat data. Estimation of biomass in dry season has proved to be challenging using remote sensing, but this study proved that using RF and remote sensing biomass can be estimated with acceptable accuracy. The success of estimating biomass in dry season was achieved using CAI (Nagler et al. 2003). CAI is computed based on the premises that the dead material absorption around 2100 nm is different from bare soil, then providing an opportunity to estimate biomass.

Table 2: prediction capability of grass biomass using RapidEye and Landsat

	R^2	*RMSE(g/m²)*	*RRMSE(%)*	*F-Stats*	*P*
Rapideye	0.86	13.42	10.61	355.70	< 0.05
Landsat	0.81	15.79	12.49	264.20	< 0.05

RMSE = root mean square error, RMSE = relative RSME

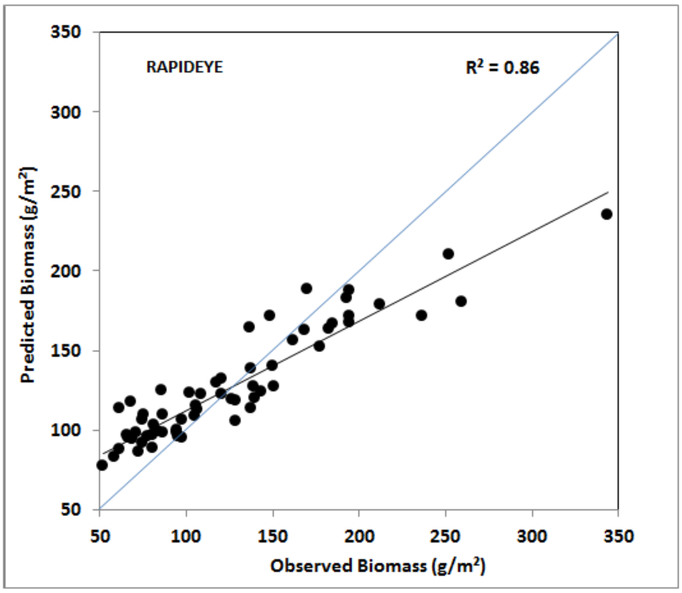

Figure 2: the performance of random forest in biomass estimation using RapidEye data

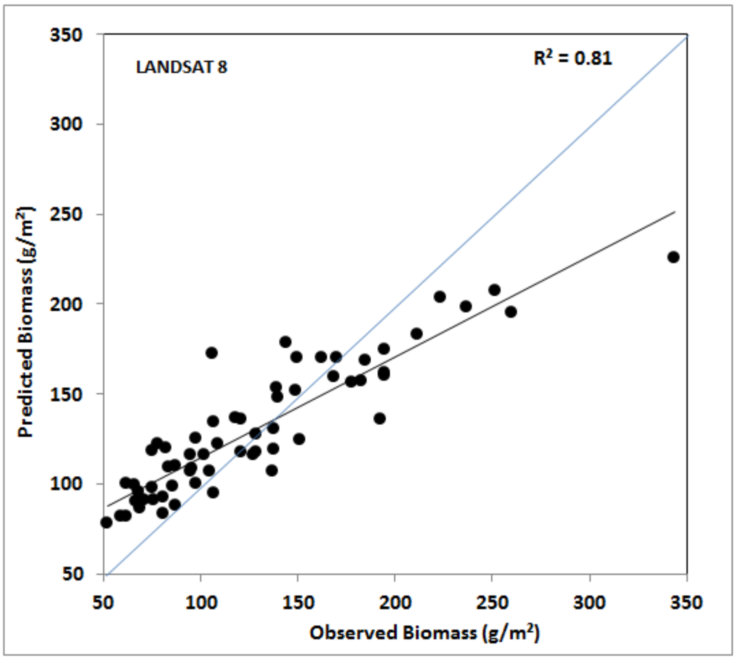

Figure 3: the performance of random forest in biomass estimation using Landsat 8 data

Variables of importance in the analysis for RF and RapidEye showed that bands centred around 710 nm (red edge band), 475 nm (blue) and 555 nm (green) were important for predicting biomass. While using Landsat data, bands centred at 560nm (green), 440 nm (deep blue), 860 nm (near infrared) and 2220 nm (shortwave infrared) were most important for predicting biomass. Among selected bands from the RapidEye, the red edge band was important, because it is known to relate to biomass, even during peak productivity, and rarely explored for dry season grasses. Hyperspectral studies demonstrated the importance of narrow band red edge based indices for estimating biomass (Mutanga and Skidmore 2004). For the Landsat 8 data analysis, short-wave infrared (SWIR) was considered important, and is sensitive to plant water stress. In the SWIR region, there is absorption of cellulose and nitrogen, which are related to biomass. Xu et al. (2014) used Landsat spectral bands in the SWIR region to derive the normalized difference water index (NDWI), and found that there was a relationship between NDWI and dead vegetation cover. SWIR is a crucial region for the estimation of biomass in dry season, because the vegetation is less green and dead.

RF model is robust and has been used successfully for regression and classification. RF can be used independent of the data distribution, and minimize overfitting and multicollinearity. Nevertheless, Figure 2 and 3 show that high biomass values are underestimated using RF. Mutanga et al. (2012) identified similar problem for biomass estimation in wetland vegetation using RF. This problem needs to be further investigated. Biomass estimation is a challenging activity, and several authors reported those challenges in literature (Todd et al. 1998, Lu 2006, Ahamed et al. 2011).

The current launch of Landsat 8 and free or open data dissemination policy provide opportunity to monitor grass biomass in the savanna ecosystem. Similar to this study, several studies demonstrated the use of Landsat to estimate grass biomass (Friedl et al. 1994, Todd et al. 1998, Xu et al. 2014). Studies in the semi or arid environments showed indices such as the modified soil adjusted vegetation index (MSAVI) and SAVI are crucial to vegetation growth and productivity assessment, especially using the coarse spatial resolution data (Qi et al. 1994, Rondeaux et al. 1996, Wang et al. 2006). This study demonstrated that high spatial resolution data such as RapidEye achieved high biomass estimation accuracy, as compared to Landsat 8. Nevertheless, the multiscale approach is crucial for the development of the biomass monitoring system, where high spatial resolution data can be used for calibration and validation of the Landsat 8 derived biomass.

CONCLUSIONS

The study demonstrated that biomass can be estimated using high and medium spatial resolution remote sensing data. High spatial resolution such as RapidEye yielded high biomass prediction accuracy as compared to the medium resolution data such as Landsat. RapidEye and Landsat can be used complimentarily for the development of the biomass monitoring system. Biomass monitoring system is crucial for planning and management of savanna ecosystems, in terms of grazing and fire control. For grazing, biomass is a key indicator of rangeland condition, and can be used as an input in the livestock carrying capacity determination.

ACKNOWLEGEMENT

We would like to thank the Council for Scientific and Industrial Research (CSIR) and the National Research Foundation (NRF)'s Thuthuka Post PhD programme for the funding. I would like to thank Mr. Lucas Mashele for fieldwork assistance and Patricia Khoza, Dr. Izak Smit and colleagues from the South African National Parks (SANPARKS)for arranging access to the Kruger National Park for field work.

REFERENCES

Adam, E. M., et al. 2012. "Discriminating the papyrus vegetation (Cyperus papyrus L.) and its co-existent species using random forest and Hyperspectral data resampled to HYMAP." *International Journal of Remote Sensing* 33(2): 552-569.

Ahamed, T., et al. 2011. "A review of remote sensing methods for biomass feedstock production." *Biomass and Bioenergy* 35(7): 2455-2469.

Efron, B. and G. Gong 1983. "A Leisurely Look at the Bootstrap, the Jackknife, and Cross-Validation." *The American Statistician* 37(1): 36-48.

Friedl, M. A., et al. 1994. "Estimating grassland biomass and leaf area index using ground and satellite data." *International Journal of Remote Sensing* 15(7): 1401 - 1420.

Ham, J., et al. 2005. "Investigation of the random forest framework for classification of hyperspectral data." *IEEE Transactions on Geoscience and Remote Sensing* 43(3): 492-501.

Ismail, R. and O. Mutanga 2010. "A comparison of regression tree ensembles: Predicting Sirex noctilio induced water stress in Pinus patula forests of KwaZulu-Natal, South Africa." *International Journal of Applied Earth Observation and Geoinformation* 12, Supplement 1(0): S45-S51.

Kumar, L., et al. 2001. Imaging Spectroscopy and Vegetation Science. *Image Spectroscopy*. F. D. Van Der Meer and S. M. De Jong. Dordrecht, Kluver Academc Publishers**:** 111-154.

Liaw, A. and M. Wiener 2002. "Classification and regression by Random Forest." *R News* 2(18-22).

Lu, D. 2006. "The potential and challenge of remote sensing-based biomass estimation." *International Journal of Remote Sensing* 27(7): 1297-1328.

Mutanga, O., et al. 2012. "High density biomass estimation for wetland vegetation using WorldView-2 imagery and random forest regression algorithm." *International Journal of Applied Earth Observation and Geoinformation* 18(0): 399-406.

Mutanga, O. and A. K. Skidmore 2004. "Narrow band vegetation indices overcome the saturation problem in biomass estimation." *International Journal of Remote Sensing* 25(19): 3999 - 4014.

Nagler, P. L., et al. 2003. "Cellulose absorption index (CAI) to quantify mixed soil-plant litter scenes." *Remote Sensing of Environment* 87(2-3): 310-325.

Qi, J., et al. 1994. "A modified soil adjusted vegetation index." *Remote Sensing of Environment* 48(2): 119-126.

Ramoelo, A., et al. 2012. Estimating grass nutrients and biomass as an indicator of rangeland quality and quantity using remote sensing in savanna ecosystems. *African Association on Remote Sensing of Environment (AARSE), Earth Observation and Geoinformation Sciences for Environment and Development in Africa: Global Vision and Local Action Synergy*. El Jadida, Morocco, October 29-November 2.

Rondeaux, G., et al. 1996. "Optimized of Soil-adjusted vegetation indices." *Remote Sensing of Environment* 55(2): 95-107.

Rouse, J. W., et al. 1974. Monitoring the vernal advancement and retrogradation of natural vegetation. *NASA/GSFC, Type III Final Report*. M.D. Greenbelt, 371. 371.

Shackleton, S. E., et al. 2002. "Use patterns and value of savanna resources in three rural villages in South Africa." *Economic Botany* 56(2): 130-146.

Todd, S. W., et al. 1998. "Biomass estimation on grazed and ungrazed rangelands using spectral indices." *International Journal of Remote Sensing* 19(3): 427 - 438.

Tucker, C. J. 1977. "Asymptotic nature of grass canopy spectral reflectance." *Applied Optics* 16(57-1151).
Tucker, C. J. 1977. "Spectral estimation of grass canopy variables." *Remote Sensing of Environment* 6(1): 11-26.
Tucker, C. J. and P. J. Sellers 1986. "Satellite remote sensing of primary production." *International Journal of Remote Sensing* 7(11): 1395 - 1416.
Venter, F. J., et al., Eds. 2003. Abiotic template and its associated vegetation pattern. The Kruger Experience: Ecology and Management of Savanna Heterogeneity. London, The Island Press.
Wang, H., et al. 2006. "[Variability of vegetation growth season in different latitudinal zones of North China: a monitoring by NOAA NDVI and MSAVI]." *Ying yong sheng tai xue bao = The journal of applied ecology / Zhongguo sheng tai xue xue hui, Zhongguo ke xue yuan Shenyang ying yong sheng tai yan jiu suo zhu ban* 17(12): 2236-2240.
Xu, D., et al. 2014. "Measuring the dead component of mixed grassland with Landsat imagery." *Remote Sensing of Environment* 142(0): 33-43.