

DISSERTATION

WOODY COVER IN AFRICAN SAVANNAS: MAPPING STRATEGIES AND  
ECOLOGICAL INSIGHTS AT REGIONAL AND CONTINENTAL SCALES

Submitted by

Gabriela Bucini

Graduate Degree Program in Ecology

In partial fulfillment of the requirements

For the degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Summer 2010

COLORADO STATE UNIVERSITY

May 7, 2010

WE HEREBY RECOMMEND THAT THE DISSERTATION PREPARED UNDER OUR SUPERVISION BY GABRIELA BUCINI ENTITLED WOODY COVER IN AFRICAN SAVANNAS: MAPPING STRATEGIES AND ECOLOGICAL INSIGHTS AT REGIONAL AND CONTINENTAL SCALES BE ACCEPTED AS FULFILLING IN PART REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY.

Committee on Graduate Work

---

Randall B. Boone

---

Michael A. Lefsky

---

Gerhard Dangelmayr

---

Advisor: Niall P. Hanan

---

Director: N. LeRoy Poff

## ABSTRACT OF DISSERTATION

### WOODY COVER IN AFRICAN SAVANNAS: MAPPING STRATEGIES AND ECOLOGICAL INSIGHTS AT REGIONAL AND CONTINENTAL SCALES

Savanna ecosystems are characterized by the coexistence of woody and herbaceous vegetation. They are recognized as highly heterogeneous, for their diversity of growth forms and woody plant spatial arrangements. The relative fraction of woody versus herbaceous cover is particularly important in determining ecosystem functions such as water and biogeochemical cycles and energy fluxes, availability of graze and browse resources for wild and domestic herbivores, and availability of fuel-wood and other savanna products for human societies.

This dissertation research focused on woody cover in tropical African savannas, with two main objectives, i) to map woody cover at regional to continental scales across Africa, and ii) to model its dependence on biotic and abiotic factors, at landscape, regional and continental scales. Among the most important outcomes are the creation of woody cover maps for Kruger National Park (South Africa) and the African continent using combined optical and radar imagery, and the development of ecological models that provided empirical evidence for resource-competition and disturbance mechanisms.

The two-scale approach allowed the identification of relationships between woody cover and spectral predictors which can successfully be scaled up to predict the continental distributions of woody vegetation across the full gradient from deserts, through grasslands and savannas, to the dense tropical forests. The ecological models identified mean annual precipitation (MAP) as the main determinant of woody cover at the continental level. Regional variations of this MAP-driven woody cover arose from dynamics dependent on perturbations such as fire frequency, herbivory, and anthropogenic activities combined with soil characteristics.

Gabriela Bucini  
Graduate Degree Program in Ecology  
Colorado State University  
Fort Collins, CO 80523  
Summer 2010

## ACKNOWLEDGEMENTS

Many people have helped me in the process of completing this dissertation. I would like to express my gratitude for their assistance, inspiration and encouragement. First of all, I would like to thank my advisor, Dr. Niall Hanan, for his attentive and challenging guidance. He has been a strong and respectful advisor. His care for scientific accuracy, his investigative and skeptical mind pushed me to ambitious research work that significantly contributed to ecology. He showed faith in my ability to carry out scientific work especially when I was hesitant and moving slowly through my data. His objective feedbacks helped me developing a better writing style. He has been a courageous advocate for me and encouraged me to pursue independent work and to stand up for my ideas and results.

My Committee Members, Randall B. Boone, Gerhard Dangelmayr and Michael A. Lefsky have highly contributed to this dissertation by always being active in discussing my project ideas and approaches. They shared their scientific knowledge providing opportunities to work together and a space where I could develop as a scientist. In particular, I am grateful to Michael A. Lefsky who has provided me with insights and high-level knowledge in remote-sensing and IDL. I thank him for challenging me to own my PhD work. I enjoyed collaborating with Randall B. Boone because he treated me as a peer and was open to innovation. The presence of Gerhard Dangelmayr on my committee helped me maintaining scientific rigor in my work. I would like to mention Indy Burke

who served on my Committee for many years. She was an example of both a faculty and a scientist for her commitment to promote ecological science.

The Director, N. LeRoy Poff, the faculty and Jerry Morgan of the Graduate Degree Program of Ecology (GDPE) who have provided me with top level graduate education: they have taught me how to think about ecological questions; they have created an environment where students could grow and feel inspired to pursue their scientific goals.

I would like to recognize my research group, Niall P. Hanan, Lara Prihodko, Mahesh Sankaran, Jayashree Ratnam, William Sea, Chris Williams, Andrew Tredennick, Alexandra Sutton, Justin Dohn and Alexandro Solórzano for being together with me in intellectual debates, field work and in the wavy weeks before my defense. I thank Lara for her warm and practical support. She is an example of grace integrated with scientific sensitivity.

I gratefully acknowledge many colleagues, collaborators and assistants for providing assistance and spending hours with me brain-storming on ideas, problems and data collections. Particularly, I thank Sassan Saatchi for introducing me to radar science, and Izak Smit for our work in Kruger National Park.

A special thank to the NREL scientists and staff. This is an incredibly active and bubbling environment where the evolution of ecology takes place. I have been fortunate to be a student at NREL. I have had the opportunity to interact with the best ecologists and be supported by their enthusiasm for research. It was great to share my years with the other NREL students some of which will always be good friends.

I thank my friends mainly for reminding me that a vibrating life is about thinking out of the box and pushing my boundaries. They have had good, shaking, shocking and

peaceful words that helped me growing and making wise decisions. In particular, I would like to thank Erandi for her wisdom that has been a guiding reference especially in time of insecurity. I will have the best memory of her supporting strength from the night before my defense. Joana has been an inspiring open-minded and inventive friend. Thank you for shaking my walls and for believing in me. Bobbie and Lynn have had the courage to see beyond my appearance. It was nice to have Stacy finishing her dissertation at the same and to give support to each other looking forward to the joy of our accomplishments. Francesca has been an amazing force that pushed me to use my skills. Thank you for our animated laughs! A special thought goes to Uffe whose exceptional sensitivity changed my worldview and opened the immense world of Nature to me.

Finally, I thank my family for growing with me in this PhD experience and for believing in my capacities. They have been patient, proud and happy to cross the ocean to affirm our bond.

## TABLE OF CONTENTS

INTRODUCTION.....	1
The ecology of savanna woody vegetation.....	1
<a href="#">1.1 Savanna ecology and research.....</a>	<a href="#">4</a>
<a href="#">1.2 References.....</a>	<a href="#">12</a>
A continental-scale analysis of tree cover in African savannas.....	15
<a href="#">2.1 Abstract.....</a>	<a href="#">16</a>
<a href="#">2.2 Introduction.....</a>	<a href="#">18</a>
<a href="#">2.3 Methods.....</a>	<a href="#">21</a>
<a href="#">2.3.1 Conceptual framework .....</a>	<a href="#">21</a>
<a href="#">2.3.2 Models.....</a>	<a href="#">22</a>
<a href="#">2.3.3 Statistical framework.....</a>	<a href="#">25</a>
<a href="#">2.3.4 Implementation.....</a>	<a href="#">26</a>
<a href="#">2.4 Results.....</a>	<a href="#">28</a>
<a href="#">2.4.1 Perturbation impacts across rainfall zones.....</a>	<a href="#">30</a>
<a href="#">2.4.2 Model robustness.....</a>	<a href="#">31</a>
<a href="#">2.5 Discussion.....</a>	<a href="#">33</a>
<a href="#">2.6 Conclusion.....</a>	<a href="#">39</a>
<a href="#">2.7 Acknowledgments.....</a>	<a href="#">40</a>
<a href="#">2.8 References.....</a>	<a href="#">41</a>
Woody Fractional Cover in Kruger National Park, South Africa: remote-sensing-based maps and ecological insights.....	58
<a href="#">3.1 Introduction.....</a>	<a href="#">59</a>
<a href="#">3.2 Study Area.....</a>	<a href="#">61</a>
<a href="#">3.3 Data and Methods.....</a>	<a href="#">62</a>
<a href="#">3.3.1 Data.....</a>	<a href="#">62</a>
<a href="#">3.3.2 Analysis and modeling .....</a>	<a href="#">65</a>
<a href="#">3.4 Results .....</a>	<a href="#">68</a>
<a href="#">3.4.1 Woody Cover Explanatory Model .....</a>	<a href="#">69</a>
<a href="#">3.5 Discussion .....</a>	<a href="#">70</a>
<a href="#">3.5.1 Woody Cover Map .....</a>	<a href="#">70</a>
<a href="#">3.5.2 Ecological Model .....</a>	<a href="#">72</a>
<a href="#">3.5.3 Research and Management Implications .....</a>	<a href="#">73</a>

<a href="#">3.6 Acknowledgements .....</a>	<a href="#">74</a>
<a href="#">3.7 References .....</a>	<a href="#">75</a>
Woody cover mapping in Africa: an approach with combined optical and radar remote sensing.....	87
<a href="#">4.1 Introduction .....</a>	<a href="#">88</a>
<a href="#">4.2 Materials and Methods.....</a>	<a href="#">94</a>
<a href="#">4.2.1 Woody cover training data (response variable).....</a>	<a href="#">94</a>
<a href="#">4.2.2 Remote-sensing data (predictive variables).....</a>	<a href="#">97</a>
<a href="#">4.2.3 Modeling approach.....</a>	<a href="#">101</a>
<a href="#">4.3 Results.....</a>	<a href="#">106</a>
<a href="#">4.3.1 MODIS VCF comparison.....</a>	<a href="#">114</a>
<a href="#">4.3.2 Comparison with KNP and Turkana woody cover maps .....</a>	<a href="#">117</a>
<a href="#">4.3.3 Woody cover comparison with ground photos in Namibia .....</a>	<a href="#">122</a>
<a href="#">4.4 Discussion .....</a>	<a href="#">123</a>
<a href="#">4.4.1 Development and applications.....</a>	<a href="#">128</a>
<a href="#">4.5 Conclusions.....</a>	<a href="#">129</a>
<a href="#">4.6 References.....</a>	<a href="#">131</a>
SUMMARY AND CONCLUSIONS.....	136
<a href="#">5.1 Ecological models.....</a>	<a href="#">139</a>
<a href="#">5.2 Woody cover predictive models and mapping .....</a>	<a href="#">142</a>
<a href="#">5.3 Conclusions.....</a>	<a href="#">144</a>
<a href="#">5.4 References.....</a>	<a href="#">146</a>

## LIST OF TABLES

<p>Table 2-1: Datasets used in analysis of African savanna tree cover dynamics. The dataset consists of ten layers. The last column relates each variable to equations 1 and 3. The lower case letter defines the variable symbol. The letter U stands for the “undisturbed tree cover” term and M for the “modifications by perturbation” term. Summary statistics, mean and range, are reported.....</p>	45
<p>Table 2-2: Four models with highest posterior model probabilities based on the Bayesian information criterion (BIC). All these models are structured with the sigmoid function. Presence of a predictor in a model is marked by a dot (r=mean annual precipitation, g=growing season length, s=soil texture, n=soil nitrogen, f=fire frequency, c=cultivation intensity, p=human population, and l=cattle density). K is number of free parameters; for each model i, <math>\Delta_i = BIC_i - BIC_{min}</math> and <math>w_i</math> is the derived BIC weight (model posterior probability). MAP, mean annual rainfall; GSL, growing season length.....</p>	47
<p>Table 2-3: Summary statistics for parameter estimates using the sigmoid relationship and MAP. The indices xf, xc and xm for fine are for coarse and medium texture classes. Parameters for the effects of soil nitrogen, fire frequency, cultivation, human population and cattle density are also shown. Values for parameter estimates, unconditional standard error (Eq. 6) and 95% confidence interval (eq. 7) are retrieved from model averaging. ..</p>	48
<p>Table 2-4: Models frequencies <math>\pi_i</math> obtained from the ten samples compared to the weights w obtained from the pilot sample. Models 1 to 4 are listed like in Table 2; models 5 and 6 are the two new best models. K is the model size.....</p>	49
<p>Table 3-5. : Environmental and ecological variables used in the explanatory model. The KNP source reference is the South African National Parks (SANParks) database (<a href="http://www.sanparks.org/parks/kruger/conservation/scientific/gis/">http://www.sanparks.org/parks/kruger/conservation/scientific/gis/</a>). Most of the rainfall gauges had more than 20 years of data except for 3 gauges with 12 years of data. We kept elephants as a separate group because they can affect woody vegetation not only as mixed-feeders but also through direct mortality of trees. Abbreviations: RSE=residual standard error; sd=standard deviation; “ = same as above.....</p>	78
<p>Table 3-6. Woody cover regression model (columns 2, 3, 4). Coefficient estimates, standard error, and p-value for the selected variables. The multiple <math>R^2 = 0.61</math>, residual standard error= 8.9 and <math>p &lt; 0.0001</math>. Jack-knife estimates of the model coefficients for the predictors (columns 5, 6, 7): mean, standard error and bias (Efron, 1993).....</p>	81
<p>Table 4-7: Remote sensing layers and metrics from MODIS and Q-SCAT (Quick Scatterometer) sensors.....</p>	99

Table 4-8: The predictive variable set: continuous variables including linear and non-linear (transformed) terms and categorical variable (stratification classes).....105

Table 4-9: The first five ranked models. For each predictive model, the table reports R<sup>2</sup>, Residual Standard Error (RSE), the number of model parameters N, ΔBIC and the model probabilities w (left side). Right side: R<sup>2</sup>, RSE, slope and intercept of the regression line between observed vs. predicted cover on the validation dataset.....107

Table 4-10: Summary statistics for the parameter estimates of the selected continental model: standard deviation, p-value and LMG metric for variable relative importance. Parameters for the main effects (exp\_qscatmean, NDVI\_MEAN, NDVI\_GREEN\_BROWN), the 3 categorical classes and for the interactions between the classes and exp\_qscatmean. Significance levels are also coded as: p<0.0001 = '\*\*\*', <0.001 = '\*\*', <0.01 = '\*', <0.05 = '.' .....109

Table 4-11 Bootstrap statistics for the parameter estimates of the selected continental model on 10000 bootstrap replications: original value, bias and standard error. Average R<sup>2</sup>=0.589.....109

Table 4-12: Different sources of error that could potentially affect the woody cover map. Note that other sources of error occur.....125

## LIST OF FIGURES

<p>Figure 1-1: Top left: distribution of fertile and infertile savannas (Scholes and Walker 1993). Dark shaded areas are nutrient-poor savanna; striped areas are nutrient-rich savannas. Top right: Mean Annual Rainfall (ANU-CRES, Centre for Resource and Environmental Studies, Australia). Bottom: woody cover % (Bucini et al., in preparation).....</p> <p>Figure 1-2: Left: Fire return over an 8-year period (Barbosa et al. 1999). Right: cattle density (World Resources Institute - PAGE, 2000).....</p> <p>Figure 2-3: Hypothesized functional responses of tree-cover to moisture. The independent variable (q) could be either MAP or GSL.....</p> <p>Figure 2-4: Tree-cover percent versus MAP: MODIS-based tree-cover% (dots), predicted rainfall-driven tree-cover on three calls of soil texture, fine, medium, and coarse (solid lines), and piecewise linear relationship marking the upper bound on tree-cover found by Sankaran et al. (2005).....</p> <p>Figure 2-5: Perturbation and Soil Nitrogen multiplicative effects on potential tree-cover. Larger symbols mark the response interval for the interquartile range of a factor.....</p> <p>Figure 2-6: Observed vs. predicted tree-cover%: linear fit. ....</p> <p>Figure 2-7: Averaged perturbation effects in relation to MAP. The top four panels a), b), c) and d) report the perturbation values and their corresponding effect as the fraction of loss (fire frequency, human population and cattle density) or gain with respect to the environmental tree-cover. The graph e) groups the perturbation effects and the graph f) reports the averaged environmental tree-cover and the modified tree-cover under the influence of the averaged summed effects of the perturbations.....</p> <p>Figure 2-8: Residuals' distributions by bioregion. The line inside the box: Median; Plus symbol marker: mean; Lower and upper edges of the box: 25th and 75th percentiles; endpoints of lower and upper whiskers: minimum and maximum. ....</p> <p>Figure 2-9: Tree-cover percent in African savannas. Left: model averaged predictions. Right: retrieved from MODIS sensed data.....</p> <p>Figure 3-10: Study area: KNP and the distribution of the field plots across the MAP gradient.....</p> <p>Figure 3-11: Model evaluation: observed (densiometer) versus predicted woody cover percent points (circles) and fitted line (solid, <math>R^2 = 0.61</math> and <math>RSE = 8.6</math>, <math>p &lt; 0.0001</math>). Model validation: LiDAR estimates versus predicted woody cover percent points (stars) and fitted line (dashed, <math>R^2 = 0.79</math>, <math>RSE = 8.9</math>, <math>p &lt; 0.0001</math>). ....</p>	<p>6</p> <p>7</p> <p>51</p> <p>52</p> <p>53</p> <p>54</p> <p>55</p> <p>56</p> <p>57</p> <p>82</p> <p>83</p>
--	---

Figure 3-12: Top left: KNP woody cover percent map (90-m nominal resolution). Top right: heterogeneity map for Kruger National Park with focal resolution of 1 km. Bottom: woody cover map zoom on Sabie river. Bottom left: false color infrared-composite of a 4-m resolution IKONOS image (4-28-2001) (photosynthetically active vegetation in red shades). Bottom right: woody cover percent estimates. ....84

Figure 3-13: Ecological model (regression tree) with end node values being woody cover percent. Each split refers to the condition for the left child branch. The values at the end nodes are the mean response for that stratum. Node variables: basalt/granite bedrocks, fire frequency (fire freq), elevation (DEM), slope, mean annual precipitation (map), mopane (*Colophospermum Mopane* species), sanveld community species (sanveld comm.) and elephant biomass density (el). ....85

Figure 3-14: Variation partitioning for the CART models. Variability explained by the ecological-spatial model: 61.5=39.1+20.3+2.1 %. Unexplained variability: 38.5 %.....86

Figure 4-15: Field measured (Sankaran et al. 2005) vs. 500-m MODIS VCF (Hansen et al. 2003) woody cover percent in 850 sites across African savannas: scatter plot with regression line. ....92

Figure 4-16: Coefficient of variation for woody cover vs. number of grid cells (n = 1, ..., 16). For each fixed number n of cells, the coefficient of variations is calculated from the set of all possible combinations (the group of 16 cells taken n at a time). ....97

Figure 4-17: Spectral spaces: 2-D spectral scatter plots of the remote-sensing variables used in this study. Each row is related to a different remote-sensing system: QuickSCAT (radar) and MODIS (optical). On the left: spectral space for the remote-sensing data for the continent. On the right: spectral space for the training dataset. The training data are symbol-coded by land cover class from the MODIS land cover product with the IGBP legend (<http://www-modis.bu.edu/landcover/index.html>). Our training data have a good representative range on the spectral ranges except for areas with very low Q-SCAT backscatter (Quick-scat\_mean < - 1500) mainly corresponding to deserts and water bodies.....100

Figure 4-18: Upper left: ISODATA unsupervised classification (4 classes) derived from the remote-sensing layers. There are affinities with known large-scale vegetation patterns. The very arid areas (class 3) were separated from the grasslands/shrublands (class 5) characterized by high annual NDVI variability and low woody biomass. Class 4 appears to include taller savannas, seasonal and moist tropical forests. Upper right: box plot of woody cover in the three unsupervised classes, excluding water. Lower left: unsupervised classes in the radar spectral space. Lower right: unsupervised classes in the optical spectral space.....104

Figure 4-19: Scatter plots of woody cover (training data) vs. the four spectral metrics selected as predictors for the modeling analysis (the Pearson correlation coefficients between NDVI\_green, NDVI\_max and NDVI\_mean were > 0.9 and NDVI\_mean had the highest correlation with woody cover therefore we only selected NDVI\_mean for the analysis). The graphs report the best-fit univariate models (coefficients, R2 and p-value). ....104

Figure 4-20: Histogram of Woody cover percent for the training dataset (832 points).. 106

Figure 4-21: Continental model performance: observed (measured) versus predicted woody cover % for the training dataset (left) and for the validation dataset (right). For our

purposes, model output values < 0 % were considered 0 % and values > 100 % were considered 100 %.....110

Figure 4-22: Woody cover percent map extrapolated from the continental model.....113

Figure 4-23: Top: our continental model and the MODIS VCF woody cover % maps. Bottom: Difference map between our model and VCF woody cover %. Red shades for our model < MODIS VCF woody cover %; blue shades for our model > MODIS VCF woody cover %. Graph: Regression line between MODIS VCF and our continental model woody cover on the training dataset. ....116

Figure 4-24: Top: woody cover % maps: subset from the continental map and KNP study map (Bucini et al. 2010). Bottom: difference map between the continental and the KNP study woody cover % maps. Graph: regression line between woody cover predictions from our continental model and the KNP study map on 100 random points across the KNP.....118

Figure 4-25: Turkana district, Kenya. Left: Ellis et al. (1987) map. Color codes: light yellow, cover < 2%; yellow, cover ~ 2-20%; light green, cover ~2 to >20%; dark green, cover >20%; red, > 20% to closed canopy. Right: continental woody cover map, zoom on the district (Turkana Lake is visible combining 0 % cover (black) and NoData (white). .....120

Figure 4-26: Zoom of continental woody cover map on Namibia. The letter S, T, X, Y and Z symbols indicate the area where the photos were taken. Predicted woody cover: ~8% at S, 15-20% at T, 15-30% at X, 15-25% at Y and 20-30% at Z.....121

Figure 4-27: Top: Field measured (Sankaran et al. 2005) vs. predicted woody cover. Bottom: Field measured (validation set) vs. predicted woody cover. Left: x-axis represents woody cover predicted by the continental model; right: x-axis represents woody cover predicted by MODIS VCF resampled at 1 km resolution. In black: data and regression lines; in red: one-to-one line. ....128

# **CHAPTER 1**

## **INTRODUCTION**

### **The ecology of savanna woody vegetation**

The central theme of this work is woody cover in tropical African savannas. Defined as the vertical projection of a woody plant crown, woody cover is a fundamental variable to characterize and quantify vegetation structure at the ecosystem level along with height, basal area and crown size. Savannas' uniqueness is the coexistence of woody and herbaceous vegetation. They are acknowledged as highly heterogeneous ecosystems for their diversity of growth forms and woody plants spatial arrangements (Scholes and Archer 1997, Rietkerk et al. 2002). The relative fraction of woody versus herbaceous cover is particularly important in determining ecosystem functions such as water and biogeochemical cycles and energy fluxes (Hanan 2001, Baldocchi et al. 2004, Ratnam et al. 2008, Williams et al. 2008). Woody cover fraction is an indicator of how resources (water, light and nutrients) are partitioned and how net primary production is distributed between woody and herbaceous plants and spatially distributed in the horizontal dimensions. Studies of woody cover therefore strongly contribute to the understanding of savanna morphology and functioning.

The assessment of woody cover can be done both by using field and remote-sensing techniques. Remote-sensing has the advantage to provide information at multiple scales and resolutions. It can reach areas that would be inaccessible from the ground and it covers spatial and temporal extents that would be unaffordable with field surveys. Most of the savanna studies have been focused on local to landscape scales to unravel the mechanisms underlying woody-grass coexistence. However, the more recent ecological questions related to global climate change and ecosystem functioning, have solicited savanna scientists to investigate how local processes are transferred across scales and how regional to global vegetation patterns and processes influence and are influenced by

large-scale dynamics and energy exchanges. There are a number of questions and issues pertaining to our current ability to both assess and model woody cover at broad scales:

- There are two tree cover maps at the continental level available at the moment: the Vegetation Continuous Field (VCF) (Hansen et al. 2003, Hansen et al. 2006) and the tree cover map developed by Rokhmatuloh, Nitto et al. (2005) at the Center for Environmental Remote Sensing, Japan. The mapping strategies used for these two maps are not optimal to accurately discriminate and assess woody vegetation in savannas: the tree cover calibration data underestimate woody plants (Bucini and Hanan 2007), and the predictive remote-sensing layers (from the Moderate Resolution Imaging Spectroradiometer, MODIS) used for the up-scaling are in the optical range where deciduous trees and grasses have low separability.
- There exists a large body of knowledge about ecological mechanisms driving savanna tree-grass coexistence (Walter 1971, Scholes and Archer 1997, Higgins et al. 2000, Jeltsch et al. 2000, House et al. 2003, van Langevelde et al. 2003, Ogle and Reynolds 2004, Sankaran et al. 2004, Lehmann and Hanan 2010) and several models have been developed (Walker et al. 1981, Walker and Langridge 1996, Higgins et al. 2000, Jeltsch et al. 2000, van Wijk and Rodriguez-Iturbe 2002, Bond et al. 2005, Wiegand et al. 2006, Hanan et al. 2008). Field experiments have been carried out to test and improve these models (Trollope 1982, Skarpe 1991, Weltzin and McPherson 1997, Skarpe et al. 2000, Gillson 2004, Otieno et al. 2005, Archibald and Scholes 2007, Higgins et al. 2007) but they have been limited to local and landscape scales or focused on controlling for a single determinant. There is a need for a comprehensive

model with empirical evidence that explains the coexistence and the relative abundance of woody and herbaceous vegetation (Sankaran et al. 2004).

- Humans and savannas have coevolved and people have adapted to live with climate variability and sparse resources. Still now most of the savanna regions are inhabited and managed. However over the past 50 years, drylands have undergone a persistent decrease in productivity and other ecosystem services and there is a need for an accurate evaluation of the human-disturbance interactions especially in face of climate change and new market policies (Scholes 2009, Verstraete et al. 2009). This phenomenon needs to be addressed at all scales and requires information about the relative importance of human activities versus natural disturbance in shaping, maintaining and modifying savanna vegetation cover.

My work focuses on the regional and continental scales and takes a semi-empirical approach to tackle two main issues: (i) to examine and improve the available spatial information on woody plant distribution in African savannas using remote sensing and (ii) to investigate ecological, environmental and anthropogenic drivers and their relationship with the mapped patterns of woody cover.

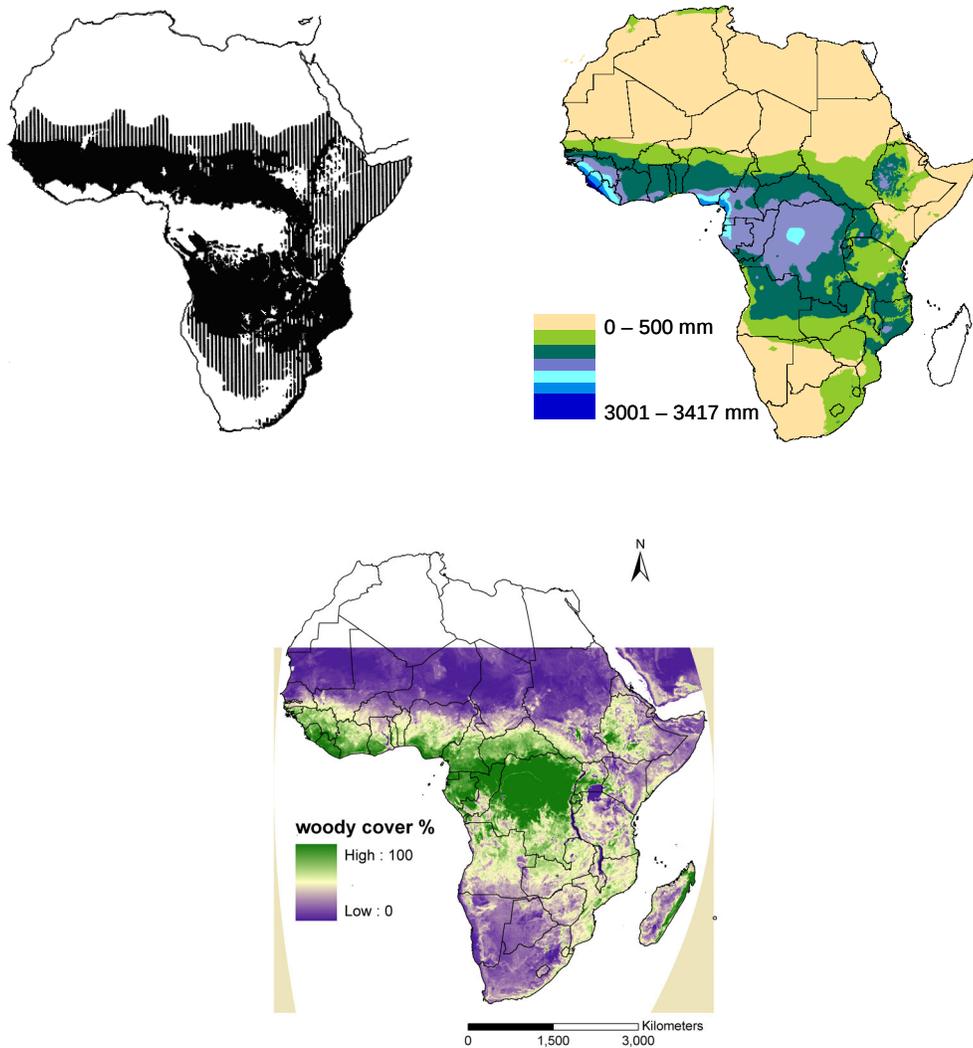
## **1.1 Savanna ecology and research**

Savannas are known to be highly complex ecosystems with strong geographic variability in vegetation structure that results from the dynamics of biotic and abiotic factors. The most prominent characteristic of savannas is the coexistence of woody and herbaceous plant functional types (Scholes and Archer 1997). Their relative proportion defines different savanna types including shrublands, open-canopy savannas with a dominating grass layer and few scattered trees, mosaic savannas with tree clumps,

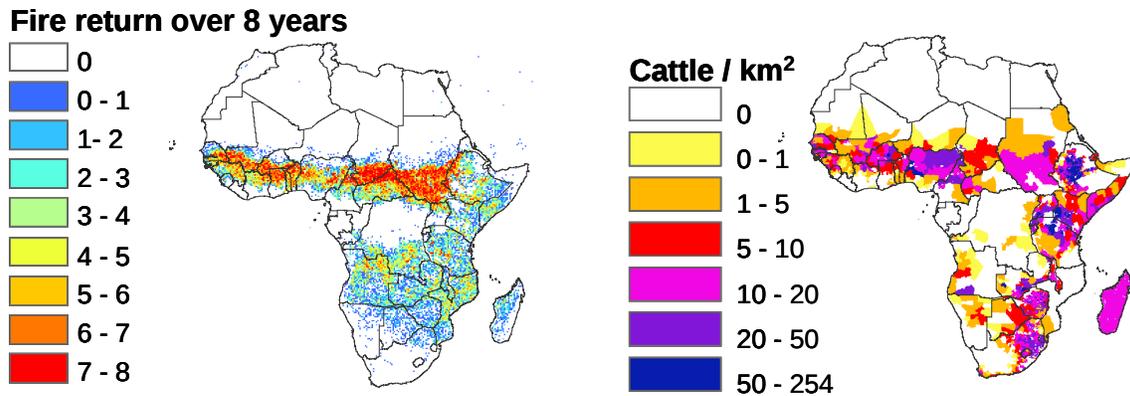
savannas with a mixture of grass, shrubs and trees, closed canopy savannas and woodlands (Scholes and Walker 1993, Solbrig et al. 1996). In the African savannas considered in this work, woody cover spans from about 5 to 70 % (House et al. 2003). Structural heterogeneity emerges at local to continental scales and is both a driver and an effect of ecological and environmental mechanisms (Rietkerk et al. 2002, Caylor et al. 2004, Levin 2005).

Tropical savannas are water-limited ecosystems because the evaporative demand (potential evapotranspiration PET) is higher than the precipitation for 4 to 11 months in the year (Solbrig 1996). The rain occurs during the summer months and therefore the growing season is associated with water availability and high temperatures. As an adaptive strategy to drought, most of the woody species are deciduous. Woody vegetation (shrubs and trees) has organs able to store water or access deep soil water (long vertical roots) and some species are found to green-up before the onset of the rain in response to longer day lengths (Shackleton 1999, Do et al. 2005, Archibald and Scholes 2007). On the other hand, the herbaceous species (grass and forbs) are dependent on rain to resprout or germinate (Scanlon et al. 2002). Similarly, woody vegetation has a longer green period and often senesces after the grasses. This temporal phenology difference is an important feature that can be exploited in remote-sensing analyses to discriminate woody plants from the herbaceous layer. Tropical savanna regions of Africa can be nutrient-limited, too. At the continental level, the infertile soils lie on old and weathered parent material in the sub-humid regions and are distinct from the more fertile soils on younger parent material in the arid and semi-arid regions (north, east and southern Africa; Scholes and Walker 1993). A visual comparison of mean annual rainfall (MAP), fertility and woody

cover suggests that MAP has a stronger association than fertility with cover (Figure 1-1). Geomorphologic characteristics become more important at local and regional level in determining water flows, soil type and water holding capacity and therefore vegetation patterns.



**Figure 1-1:** Top left: distribution of fertile and infertile savannas (Scholes and Walker 1993). Dark shaded areas are nutrient-poor savanna; striped areas are nutrient-rich savannas. Top right: Mean Annual Rainfall (ANU-CRES, Centre for Resource and Environmental Studies, Australia). Bottom: woody cover % (Bucini et al., in preparation).



**Figure 1-2:** Left: Fire return over an 8-year period (Barbosa et al. 1999). Right: cattle density (World Resources Institute - PAGE, 2000)

Fire and herbivory (Figure 1-2) are natural perturbations considered important in shaping savanna vegetation. Fires are mainly carried by the herbaceous layer and therefore their frequency depends on grass production. In general, mesic savannas present higher fire frequency than arid savannas. Humans manage fire principally to maintain grass in the system and support pastoral activities as well as agriculture.

This highly complex dynamics feeds what is the overarching research question for savanna ecologists: which are the biotic and abiotic factors and processes that give rise to and maintain coexistence and the woody-herbaceous vegetation mixture? Field experiments have allowed us to study interactions between vegetation and both environmental (climate, soil characteristics and topography) and perturbation (fire, herbivory, human activity) factors. From the theoretical point of view, two distinct paradigms exist: (i) the resource competition paradigm allows coexistence through root

niche partitioning of water and resources (Walter 1971) and (ii) the demographic bottleneck paradigm where coexistence results from fire, herbivory and water variability limiting woody vegetation establishment and growth (Higgins et al. 2000). Despite accumulated evidence for both models, they singularly are not able to explain all the properties found in savanna regions across the globe. The first attempt to integrate the two theoretical frameworks came from Sankaran et al. (2004) and was very recently extended by Hanan and Lehmann (2010). The latter effort proposes a conceptual framework where a bottleneck on seed establishment is set by water limitation and variability determining an upper bound on woody cover in more arid savannas. In more mesic savannas, water limitation is less but seedlings competition with adult trees and grasses restricts woody cover. In combination with resource-based dynamics, disturbance such as fire, herbivory and human activities, determine the observable woody cover. At the continental scale, the empirical evidence for a unified paradigm was given first by a metadata analysis (Sankaran et al. 2005) based on 850 sites distributed across the African savannas. Their semi-empirical model showed that mean annual rainfall is the main limiting resource factor in maximum cover of woody plant in the arid and semi-arid savannas (MAP 100-650 mm) while perturbations are key to inhibit tree establishment and potential canopy closure in wet savannas (MAP 650-1200mm). Both MAP and perturbations act at all levels but their role changes across the rainfall gradient.

The plan for this dissertation is to pair efforts on the mapping aspect of woody cover with efforts on woody cover modeling. The broad aim is to provide a better quantification and understanding of savannas woody cover at regional and continental levels. The variety of radiometric ranges available from more recent remote-sensing

instruments opened up the possibility to investigate the benefit of combining optical and radar systems. The resulting maps of woody cover were compared and validated to build a robust and new source of information for applications requiring woody cover estimates. Empirical models were also included in form of exploring complex relationships between woody cover, climate, soil properties, fire, herbivory and human activities. This dissertation is organized in five chapters summarized in the following paragraphs:

### **Chapter 1: Introduction**

The introduction gives an overview on the ecology of savanna vegetation with particular emphasis on woody vegetation and woody cover.

### **Chapter 2: A continental-scale analysis of tree cover in African savannas**

(Bucini, G., and N. P. Hanan. 2007. *Global Ecology and Biogeography* 16:593-605).

*How biotic and abiotic factors interact to promote and modify tree cover at the continental level?*

This is a semi-empirical analysis of continental woody-cover patterns in relation to climate, soil properties, fire, herbivory and human activities comparing and ranking alternative non-linear models according to the Bayesian information criterion (BIC). This chapter also highlights the inaccuracy of the VCF tree cover product on savannas and the necessity to investigate new approaches to map woody cover in savannas.

**Chapter 3: Woody fractional cover in Kruger National Park, South Africa: remote-sensing-based maps and ecological insights.** (Bucini, G., N. P. Hanan, R. B. Boone, I. P. J. Smit, S. Saatchi, M. A. Lefsky, and G. P. Asner. 2010. In M. J. Hill and N. P. Hanan, editors. Ecosystem function in savannas: measurement and modeling at landscape to global scales. CRC/Taylor and Francis; in press).

1. *Can the combination of optical and radar remote-sensing provide an improved approach to map woody cover (including small trees and shrubs) in savannas?*
2. *How biotic and abiotic factors interact to give rise to spatial woody vegetation patterns at landscape to regional scales?*

This is a complete study in mapping and ecological modeling focused at the regional scale of Kruger National Park. This study was also used as a pilot study to develop experience and lay out strategies for the continental woody cover mapping work.

**Chapter 4: Mapping woody cover percent in Africa: an empirical approach with radar and optical remote-sensing.** (Bucini, G., N. P. Hanan, S. Saatchi, M. Sankaran M., Lefsky, in preparation).

1. *Is the combination of optical and radar imagery a good approach to up-scale and map woody cover at the scale of continental Africa?*
2. *Are radar and optical remote-sensing systems suited for mapping woody cover across the highly variable range of woody cover (desert, shrubland, open canopy savanna with scattered or clumped trees, closed canopy savanna, dense forest)?*

- a. *Can these remote-sensing systems discriminate woody vegetation from grass and/or bare ground?*
- b. *Are remote-sensing systems sensitive to the different types of woody plants (deciduous or evergreen, shrubs, short trees, multi-stemmed trees, tall trees) in African ecosystems?*

This chapter examines different methodologies based on the combination of optical and radar imagery for woody cover mapping at the continental scale with particular emphasis on the ability to detect woody plants in arid and semi-arid system. It shows that the larger spectral domain of optical and radar bands that are sensitive to greenness (chlorophyll content), brightness, woody material, water content and structure, responds to the observed variability in woody cover and captures about 70% of it by use of multiple linear regression models with data stratification. In this work, limitations and inaccuracies come from both the woody cover training data (plot scale, spatial distribution and assessments methods) and the remote-sensing information (signal saturation, spectral separability, topography, soil moisture and roughness).

## **Chapter 5: Conclusions**

## 1.2 References

- Archibald, S., and R. J. Scholes. 2007. Leaf green-up in a semi-arid African savanna - separating tree and grass responses to environmental cues. *Journal of Vegetation Science* **18**:583-594.
- Baldocchi, D. D., L. K. Xu, and N. Kiang. 2004. How plant functional-type, weather, seasonal drought, and soil physical properties alter water and energy fluxes of an oak-grass savanna and an annual grassland. *Agricultural and Forest Meteorology* **123**:13-39.
- Barbosa, P. M., D. Stroppiana, J. M. Gregoire, and J. M. C. Pereira. 1999. An assessment of vegetation fire in Africa (1981-1991): Burned areas, burned biomass, and atmospheric emissions. *Global Biogeochemical Cycles* **13**:933-950.
- Bond, W. J., F. I. Woodward, and G. F. Midgley. 2005. The global distribution of ecosystems in a world without fire. *New Phytologist* **165**:525-537.
- Bucini, G., and N. P. Hanan. 2007. A continental-scale analysis of tree cover in African savannas. *Global Ecology and Biogeography* **16**:593-605.
- Caylor, K. K., P. R. Dowty, H. H. Shugart, and S. Ringrose. 2004. Relationship between small-scale structural variability and simulated vegetation productivity across a regional moisture gradient in southern Africa. *Global Change Biology* **10**:374-382.
- Do, F. C., V. A. Goudiaby, O. Gimenez, A. L. Diagne, M. Diouf, A. Rocheteau, and L. E. Akpo. 2005. Environmental influence on canopy phenology in the dry tropics. *Forest Ecology and Management* **215**:319-328.
- Gillson, L. 2004. Evidence of hierarchical patch dynamics in an east African savanna? *Landscape Ecology* **19**:883-894.
- Hanan, N. P. 2001. Enhanced two-layer radiative transfer scheme for a land surface model with a discontinuous upper canopy. *Agricultural and Forest Meteorology* **109**:265-281.
- Hanan, N. P. and C. E. R. Lehmann. 2010. Tree-grass interactions in savannas: paradigms, contradictions and conceptual models. *in* M. J. Hill and N. P. Hanan, editors. *Ecosystem function in savannas: measurement and modeling at landscape to global scales*. CRC/Taylor and Francis.
- Hanan, N. P., W. B. Sea, G. Dangelmayr, and N. Govender. 2008. Do fires in savannas consume woody biomass? A comment on approaches to modeling savanna dynamics. *American Naturalist* **171**:851-856.
- Hansen, M. C., R. DeFries, J. R. G. Townshend, M. Carroll, C. Dimiceli, and R. Sohlberg. 2006. *Vegetation Continuous Fields MOD44B, 2001 Percent Tree Cover, Collection 4*. University of Maryland, College Park, Maryland, 2001.
- Hansen, M. C., R. S. DeFries, J. R. G. Townshend, M. Carroll, C. Dimiceli, and R. A. Sohlberg. 2003. *Global Percent Tree Cover at a Spatial Resolution of 500 Meters: First Results of the MODIS Vegetation Continuous Fields Algorithm*. Earth Interactions.
- Higgins, S. I., W. J. Bond, E. C. February, A. Bronn, D. I. W. Euston-Brown, B. Enslin, N. Govender, L. Rademan, S. O'Regan, A. L. F. Potgieter, S. Scheiter, R. Sowry, L. Trollope, and W. S. W. Trollope. 2007. Effects of four decades of fire manipulation on woody vegetation structure in savanna. *Ecology* **88**:1119-1125.

- Higgins, S. I., W. J. Bond, and W. S. W. Trollope. 2000. Fire, resprouting and variability: a recipe for grass-tree coexistence in savanna. *Journal of Ecology* **88**:213-229.
- House, J. I., S. Archer, D. D. Breshears, and R. J. Scholes. 2003. Conundrums in mixed woody-herbaceous plant systems. *Journal of Biogeography* **30**:1763-1777.
- Jeltsch, F., G. E. Weber, and V. Grimm. 2000. Ecological buffering mechanisms in savannas: A unifying theory of long-term tree-grass coexistence. *Plant Ecology* **150**:161-171.
- Levin, S. A. 2005. Self-organization and the emergence of complexity in ecological systems. *Bioscience* **55**:1075-1079.
- Ogle, K., and J. F. Reynolds. 2004. Plant responses to precipitation in desert ecosystems: integrating functional types, pulses, thresholds, and delays. *Oecologia* **141**:282-294.
- Otieno, D. O., M. W. T. Schmidt, J. I. Kinyamario, and J. Tenhunen. 2005. Responses of *Acacia tortilis* and *Acacia xanthophloea* to seasonal changes in soil water availability in the savanna region of Kenya. *Journal of Arid Environments* **62**:377-400.
- Ratnam, J., M. Sankaran, N. P. Hanan, R. C. Grant, and N. Zambatis. 2008. Nutrient resorption patterns of plant functional groups in a tropical savanna: variation and functional significance. *Oecologia* **157**:141-151.
- Rietkerk, M., M. C. Boerlijst, F. van Langevelde, R. HilleRisLambers, J. van de Koppel, L. Kumar, H. H. T. Prins, and A. M. de Roos. 2002. Self-organization of vegetation in arid ecosystems. *American Naturalist* **160**:524-530.
- Rokhmatuloh, D. Nitto, H. Al Bilbisi, and R. Tateishi. 2005. Percent tree cover estimation using regression tree method: a case study of Africa with very-high resolution QuickBird images as training data. Pages 2157-2160 in *Geoscience and Remote Sensing Symposium, 2005. IGARSS '05. Proceedings. 2005 IEEE International*.
- Sankaran, M., N. P. Hanan, R. J. Scholes, J. Ratnam, D. J. Augustine, B. S. Cade, J. Gignoux, S. I. Higgins, X. Le Roux, F. Ludwig, J. Ardo, F. Banyikwa, A. Bronn, G. Bucini, K. K. Caylor, M. B. Coughenour, A. Diouf, W. Ekaya, C. J. Feral, E. C. February, P. G. H. Frost, P. Hiernaux, H. Hrabar, K. L. Metzger, H. H. T. Prins, S. Ringrose, W. Sea, J. Tews, J. Worden, and N. Zambatis. 2005. Determinants of woody cover in African savannas. *Nature* **438**:846-849.
- Sankaran, M., J. Ratnam, and N. P. Hanan. 2004. Tree-grass coexistence in savannas revisited - insights from an examination of assumptions and mechanisms invoked in existing models. *Ecology Letters* **7**:480-490.
- Scanlon, T. M., J. D. Albertson, K. K. Caylor, and C. A. Williams. 2002. Determining land surface fractional cover from NDVI and rainfall time series for a savanna ecosystem. *Remote Sensing of Environment* **82**:376-388.
- Scholes, R. J. 2009. Syndromes of dryland degradation in southern Africa. *African Journal of Range & Forage Science* **26**:113-125.
- Scholes, R. J., and S. R. Archer. 1997. Tree-grass interactions in savannas. *Annu. Rev. Ecol. Syst.* **28**:517-544.
- Scholes, R. J., and B. H. Walker. 1993. *An African savanna: synthesis of the Nylsvley study*. Cambridge University Press, Cambridge.
- Shackleton, C. M. 1999. Rainfall and topo-edaphic influences on woody community phenology in South African savannas. *Global Ecology and Biogeography* **8**:125-136.
- Skarpe, C. 1991. Impact of Grazing in Savanna Ecosystems. *Ambio* **20**:351-356.

- Skarpe, C., R. Bergstrom, A. L. Braten, and K. Danell. 2000. Browsing in a heterogeneous savanna. *Ecography* **23**:632-640.
- Solbrig, O. T. 1996. The diversity of the savanna ecosystem. Pages 1-27 in O. T. Solbrig, E. Medina, and J. F. Silva, editors. *Biodiversity and Savanna Ecosystem Processes - a Global Perspective*. Springer-Verlag Berlin, Berlin 33.
- Solbrig, O. T., E. Medina, and J. F. Silva. 1996. *Biodiversity and Savanna Ecosystem Processes - A Global Perspective*. Springer, Berlin
- Trollope, W. S. W. 1982. Ecological Effects of Fire in South-African Savannas. *Ecological Studies* **42**:292-306.
- van Langevelde, F., C. van de Vijver, L. Kumar, J. van de Koppel, N. de Ridder, J. van Andel, A. K. Skidmore, J. W. Hearne, L. Stroosnijder, W. J. Bond, H. H. T. Prins, and M. Rietkerk. 2003. Effects of fire and herbivory on the stability of savanna ecosystems. *Ecology* **84**:337-350.
- van Wijk, M. T., and I. Rodriguez-Iturbe. 2002. Tree-grass competition in space and time: Insights from a simple cellular automata model based on ecohydrological dynamics. *Water Resources Research* **38**:15.
- Verstraete, M. M., R. J. Scholes, and M. S. Smith. 2009. Climate and desertification: looking at an old problem through new lenses. *Frontiers in Ecology and the Environment* **7**:421-428.
- Walker, B. H., and J. L. Langridge. 1996. Modelling plant and soil water dynamics in semi-arid ecosystems with limited site data. *Ecological Modelling* **87**:153-167.
- Walker, B. H., D. Ludwig, C. S. Holling, and R. M. Peterman. 1981. Stability of Semi-Arid Savanna Grazing Systems. *Journal of Ecology* **69**:473-498.
- Walter, H. 1971. *Ecology of Tropical and Subtropical Vegetation*. Oliver & Boyd, Edinburgh, UK.
- Weltzin, J. F., and G. R. McPherson. 1997. Spatial and temporal soil moisture resource partitioning by trees and grasses in a temperate savanna, Arizona, USA. *Oecologia* **112**:156-164.
- Wiegand, K., D. Saitz, and D. Ward. 2006. A patch-dynamics approach to savanna dynamics and woody plant encroachment - Insights from an arid savanna. *Perspectives in Plant Ecology Evolution and Systematics* **7**:229-242.
- Williams, C. A., N. P. Hanan, I. Baker, G. J. Collatz, J. Berry, and A. S. Denning. 2008. Interannual variability of photosynthesis across Africa and its attribution. *Journal of Geophysical Research-Biogeosciences* **113**:15.

## **CHAPTER 2**

### **A continental-scale analysis of tree cover in African savannas**

**Gabriela Bucini<sup>1</sup> and Niall P. Hanan<sup>1</sup>**

<sup>1</sup>Natural Resource Ecology Laboratory, Colorado State University, Fort Collins, CO 80523-1499,  
USA.

**PUBLISHED:** *Global Ecology and Biogeography* (2007). 16:593-605.

## 2.1 Abstract

**Aim** We present a continental-scale analysis that explores the processes controlling woody community structure in tropical savannas. We analyse how biotic and abiotic factors interact to promote and modify tree cover, examine alternative ecological hypotheses and quantify disturbance effects using satellite estimates of tree cover.

**Location** African savannas.

**Methods** Tree cover is represented as a resource-driven potential cover related to rainfall and soil characteristics perturbed by natural and human factors such as fire, cattle grazing, human population and cultivation. Within this framework our approach combines semi-empirical modelling and information theory to identify the best models.

**Results** Woody community structure across African savannas is best represented by a sigmoidal response of tree cover to mean annual rainfall (MAP), with a dependency on soil texture, which is modified by the separate effects of fire, domestic livestock, human population density and cultivation intensity. This model explains ~66% of the variance in tree cover and appears consistent across the savanna regions of Africa.

**Main conclusions** The analysis provides new understanding of the importance and interaction of environmental and disturbance factors that create the broad spatial patterns of tree cover observed in African savannas. Woody cover increases with rainfall, but is modified by disturbances. These “perturbation” effects depend on MAP regimes: in arid savannas (MAP<400 mm), they are generally small (<1% decrease in cover), while in semi-arid and mesic savannas (400-1600 mm), perturbations result in an average 2% (400 mm) to 23% (1600 mm) decrease in cover; fire frequency and human population have more influence than cattle, and cultivation appears, on average, to lead to small

increases in woody cover. Wet savannas (1600-2200 mm) are controlled by perturbations that inhibit canopy closure and reduce tree cover by, on average, 24-34%. Full understanding of the processes determining savanna structure requires consideration of resource limitation and disturbance dynamics.

**Keywords:** Africa, disturbance, fire, savanna, tree cover, tree-grass interactions

## 2.2 Introduction

The vegetation of climatically similar regions of the world often presents similar structural and functional attributes, independent of the evolutionary history of their flora and fauna. Such ecological convergence implies that global patterns of vegetation are broadly predictable from environmental variables, primarily precipitation and temperature. In addition to climate, however, other factors such as soil characteristics, fire, herbivory and human activities are important forces in the development and control of vegetation structure and function.

Savannas are defined on the basis of both ecological characteristics and climatic attributes. The common characteristic of all savannas is the coexistence of woody and herbaceous vegetation in regions where seasonality is controlled by distinct dry seasons, rather than by cold (Scholes & Archer, 1997). Ecologists have long been interested in the mechanisms that create and maintain the coexistence of trees and grasses in savanna systems. The factors responsible for the coexistence of these two very different vegetation forms are varied: early hypotheses centred on inter-specific competition and vertical niche separation between the roots of woody and herbaceous species in the soil (Walker & Noymeir, 1982; Scholes & Archer, 1997). More recent work has highlighted the importance of disturbances such as fire, herbivory and human activities (Higgins *et al.*, 2000; van Langevelde *et al.*, 2003; Sankaran *et al.*, 2004; Sankaran *et al.*, 2005).

The critical importance of water is well recognized for savanna vegetation (Scholes & Archer, 1997; Walker & Langridge, 1997; Sankaran *et al.*, 2004). Mean tree cover among sites, and maximum observed tree cover increase with mean annual rainfall, but substantial variation in tree cover occurs with disturbance and climate variability

(Fernandez-Illescas & Rodriguez-Iturbe, 2004; Wiegand *et al.*, 2006). The role of soil substrate is inseparable from the role of water because it acts as a temporary store for precipitation inputs and as a regulator for the major outflows through evapotranspiration and deep percolation (Noy-Meir, 1973). The inverse texture hypothesis (Noy-Meir, 1973) has generally been interpreted with respect to primary productivity, with fine texture soils able to support more net primary production (NPP) than coarse soils at higher rainfall (above MAP~400 mm) and coarse soils supporting higher NPP at low rainfall (Sala *et al.*, 1988). However, in his 1973 paper, Noy-Meir also relates the inverse texture hypothesis to the prominence of perennial vegetation, stating that in wetter climates fine texture soils will support “taller and denser perennial vegetation” than coarse texture soils, and vice versa for dry climates. This prediction with respect to the impact of soil texture and rainfall on the success of perennial versus annual plants, or indeed woody versus herbaceous plants, has not, to our knowledge, been tested with observational or experimental data.

Woody community structure in savannas is also strongly affected by fire (Hochberg *et al.*, 1994; Gignoux *et al.*, 1997; Higgins *et al.*, 2000; Bond & Keeley, 2005), herbivory (Ellis & Swift, 1988; Skarpe, 1991; Van de Koppel & Prins, 1998), and other perturbations linked to human land use (Belsky, 1987; Ellis & Galvin, 1994; Higgins *et al.*, 1999; Laris, 2002). Fire and herbivory alter the mix of plant growth-forms, and the competitive interactions between them, through direct consumption of living and dead plant material. Savannas are the most frequently burnt ecosystems (Barbosa *et al.*, 1999; Dwyer *et al.*, 2000) and fire is considered to have a large regulatory influence on emergent vegetation structure (Scholes & Archer, 1997; Jeltsch *et al.*, 2000; House *et al.*,

2003; Sankaran *et al.*, 2004). In particular, fire frequency can control the probability of tree seedling escape from the flame zone and survival to mature size classes. Fires can thereby suppress woody cover and contribute to tree–grass coexistence in mesic and wet savannas (Higgins *et al.*, 2000). Browsers can have analogous effects, suppressing seedlings by browsing (Prins & Van der Jeugd, 1993; Van de Koppel & Prins, 1998), whereas grazers (particularly domestic cattle) are generally considered to suppress grass biomass and thereby release woody plants from competition and reduce the frequency or intensity of fires, leading to increased woody cover (Bond *et al.*, 2005; Metzger *et al.*, 2005). Humans also interact directly with the system using land for pasture, agriculture, fuel and timber, and causing alterations in tree cover proportions (Sinclair & Fryxell, 1985; Homewood *et al.*, 2001; Shackleton *et al.*, 2005).

Sankaran *et al.* (2004) suggest that the integration of resource-based and disturbance-based models is required to explain tree–grass coexistence and their relative abundance in savannas. A landscape-scale study conducted in an arid savanna of Namibia by Wiegand *et al.* (2006) concluded that the tree–grass ratio is influenced both by primary (rain and nutrients) and secondary (fire and herbivory) determinants. Significant advances in understanding savanna ecology have been made over many years of research, but published studies often base ecological conclusions, and mathematical models, on data from a small number of sites with particular climatic, edaphic and disturbance characteristics, thus limiting the broad relevance of the conclusions. Comprehensive analyses aimed at separating the role of resources and competition from the role of disturbances across broad-scale climatic and environmental gradients are relatively few and far between. In this paper we develop and test a series of empirical, but functionally

meaningful, models to investigate variability in tree cover across all African savannas. Our aim is to identify the processes that create and modify tree cover at these large continental-scales where climate, soil type and disturbances by fire, herbivores and humans vary substantially. We examine the extent to which variations in annual precipitation, soil nitrogen and texture, fire frequency, cultivation intensity, cattle density, and human population may contribute to observed emergent properties of tree cover across Africa.

In the following sections we develop a conceptual model for the ways in which climate and other factors interact to control woody community structure in savanna systems. We translate these concepts into quantitative models and test the alternative hypotheses using satellite-derived tree cover data. Achieving a broad understanding of resource use and perturbation influences requires describing the processes that drive variation in the dynamics of tree cover development and persistence. A general model predicting tree cover must discriminate the action of different factors and evaluate their effects in relation to their intensity. We use our models to illuminate possible dynamics that are common among African savannas and to quantify the importance and role of climate, soil and disturbance in controlling observed tree cover across the savannas of Africa.

## **2.3 Methods**

### ***2.3.1 Conceptual framework***

We postulate that the coexistence of trees and grasses in savannas results from competition between the two vegetation forms as well as their complex interactions with

climate, soil biochemistry, fire, herbivory, and human activities. These factors can lead to the direct mortality of trees or grasses, or change the competitive interactions between them (Sankaran *et al.*, 2004). To explain the tree cover fraction in African savannas, we propose to partition the variation of tree cover into two primary constituents: the potential cover configuration in undisturbed systems ( $U$ ) and a modification of  $U$  dependent on perturbation interactions ( $M$ ). We relate each constituent to a different set of explanatory variables: the first consisting of climatological and soil descriptors, and the second consisting of natural and human-related perturbations. In this approach, we build models using separate multiplicative terms  $U$  and  $M$ :

$$tc = U \cdot M$$

**Equation 2-1**

where  $tc$  is tree cover (%),  $U$  is composed of functions  $u$  that depend on variables related to undisturbed tree growth, and  $M$  is composed of functions  $m$  that depend on variables related to perturbation. By partitioning, we seek to resolve the relative importance of: (1) environmental components that determine the vegetation structure in the absence of perturbations; and (2) the perturbation components that create a departure from the “potential” (i.e. undisturbed) state.

### **2.3.2 Models**

Given the conceptual framework, we now choose explanatory variables and structure alternative models. Each model comprises a set of predictors and functions that describe ecological or heuristic relationships of tree cover with the predictors. Table 2-1 shows the variables used for the analysis, their relation to the potential-perturbation framework and their notation. The response variable is tree cover percent ( $tc$ ) estimated

using the MODerate-resolution Imaging Spectroradiometer (MODIS) sensor onboard NASA's Terra satellite (Hansen *et al.*, 2003).

The independent variables for  $U$  (Equation 2-1) are mean annual precipitation (MAP) or, alternatively, growing season length (GSL). GSL was computed from rainfall and potential evapotranspiration estimates (Allen *et al.*, 1998) to explore whether other aspects of climate (that take into account both rainfall and evaporative demand) need to be considered in addition to rainfall. Soil texture and soil nitrogen effects were also included in  $U$  to represent the possible impacts of these on water and nutrient relationships that might alter the climate-driven woody community structure. The term  $M$  depends on the perturbation variables fire frequency, cattle density, human population density and cultivation intensity. Pearson correlation coefficients between predictor variables are less than 0.5 indicating that the variables are not strongly correlated.

In structuring the models, we consider four alternative functions  $u$  (Figure 2-3) representative of hypotheses concerning the response of tree cover to moisture (using either MAP or GSL):

Linear. The linear response represents the case in which tree cover increases proportionally with water across the rainfall gradient. This can be considered as the null model and the following alternative functions are attempts to account for more complex responses of tree cover to moisture regimes.

Michaelis-Menton. The shape of this function reflects a saturation hypothesis where at relatively high moisture the response levels out as canopy closure approaches and other factors become limiting.

Sigmoid. This curve corresponds to the hypothesis that a minimum in moisture is required for woody plants to establish and survive and that tree cover then increases to an asymptote as canopy closure approaches.

Piecewise linear. This line results from quantile regression of tree cover on rainfall determined by Sankaran *et al.* (2005) using independent field measurements. We used the piecewise linear equation with fixed parameters as reported by Sankaran *et al.* to provide an estimate of the moisture-driven potential tree cover from which we attempted to fit the perturbation effects.

The full term  $U$  relates tree cover to MAP or GSL using one of these four curves with an adjustment for different soil texture classes, multiplied by a linear function of soil nitrogen. Conceptually,  $U$  represents tree cover in undisturbed systems but it is recognized that such conditions rarely occur in reality because most savannas are perturbed to a greater or lesser degree. The relationships described by  $U$  will be referred to as climate-driven tree cover and can be considered the mean climate response around which low-perturbation points lie.

The term  $M$  (Equation 2-1) is based on parametric functions with a rational form:

$$m(v) = 1 - \alpha \cdot v^\beta \quad \text{Equation 2-2}$$

with  $v$  a perturbation variable in the interval  $[0,1]$  and  $\alpha$  and  $\beta$  parameters ( $\beta > 0$ ). Each function  $m$  depends on one perturbation variable and decreases or increases the value of tree cover predicted by the term  $U$  by the fraction  $\alpha v^\beta$  depending on the nature of the interactions and the patterns embedded in the observations. This mathematical structure permits both suppression and facilitation of moisture-driven tree cover by perturbations,

with convex, quasi-linear and concave forms possible depending on fitted parameter values.

In composing different models of tree cover, a moisture function was always selected, with or without soil nitrogen and texture functions, to form  $U$ . The  $U$  term was then tested with all possible combinations of the perturbation functions  $m$  (including models with, and without, perturbations), making a total of 512 candidate models.

### **2.3.3 Statistical framework**

A critical aspect of inference in environmental science involves the search for an appropriate approximating model supported by empirical data. Information theory approaches allow evaluation of the weight of evidence for multiple hypotheses represented by a set of possible models, and selection of a single model or a subset of models (where there is ambiguity) that provide the “best”, and most parsimonious, representation of the data (Burnham & Anderson, 2002). Models that are poor at explaining the information in the data, which have errors in their structure and/or predictor choice, receive high model-selection uncertainty.

The Bayesian information criterion (BIC) (Schwarz, 1978) offers a way to provide balance between accuracy and parsimony. The expression of BIC entails terms representing lack of fit based on the maximized likelihood associated with the model parameters, a bias correction factor related to model complexity and a penalty term dependent on sample size. This criterion was developed within the Bayesian framework (Hoeting *et al.*, 1999; Hobbs & Hilborn, 2006) and provides a method for approximating model posterior probabilities that is accurate when the sample size is large, i.e. more than 20 times the number of predictors (Kass & Raftery, 1995). The application of BIC

assumes that all models and parameters are initially equally likely (Hoeting *et al.*, 1999; Wintle *et al.*, 2003); this is the ‘default’ approach when there is insufficient information to define prior distributions.

In traditional (frequentist) statistics there is an absence of formal methods for deriving and incorporating model uncertainty conditional to the data. Frequentist approaches and the Akaike information criterion (AIC) (Akaike, 1973) do not account for sample size and, when a data set is large, they would favour complex models that better fit variance but may fail to identify the factors of importance (Kass & Raftery, 1995; Zucchini, 2000; Johnson & Omland, 2004).

In this paper, we were interested in finding one or more models that describe tree cover patterns in order to better understand which factors and processes give rise to it. We used BIC to evaluate objectively and rank models of differing complexity by extracting a sample of data points from the continental dataset that was large compared to the model dimensions. Relative differences among model BIC values enable the derivation of posterior probabilities called “weights” ( $w$ ) that allow the candidate models to be ranked and quantify model selection uncertainty (Burnham & Anderson, 2002; Hobbs & Hilborn, 2006).

### **2.3.4 Implementation**

We defined the savanna regions of Africa using a bioregion map (Table 2-1) and, within these, we selected areas with MAP less than 2200 mm. We randomly sampled 0.02% from the 500-metre tree cover dataset obtaining 13,416 points for the subsequent analysis. This sample size represents a trade-off between a sample that is representative of the tree cover variability across the continent while also being manageable for

statistical analyses. The difference between the population mean ( $15.96 \pm 0.09\%$ ) and sample mean ( $15.86 \pm 0.31\%$ ) was not statistically significant.

We extracted data for all the other variables (Table 2-1) at the sample locations. In many cases the data available with continental coverage for Africa are at much coarser spatial resolution than the tree cover dataset (see, for example, soil texture and cultivation intensity in Table 2-1). This scale mismatch means that some fine-scale responses of tree cover to potential drivers cannot be represented in the following analysis. However, we decided against degrading the response variable (*tc*) to the coarsest resolution since there is real information in *tc* at these scales, and we are interested in the extent to which fine-grain patterns respond to the fine scale drivers that are available. In this situation, the coarse resolution datasets provide our “best estimate” of conditions at the *tc* locations, with an expectation that the resulting relationships will be weaker than would be expected if all data were available at high resolution.

The values of the perturbation variables and soil nitrogen were rescaled in the interval [0,1] between minimum and maximum to make their effects comparable and enhance the numerical stability of the statistical models. Log and square-root transformations were applied to soil nitrogen and human population variables, respectively, adjusting for skewness. We implemented the models using SAS software (SAS, 2000) and fitted them using PROC NLMIXED with the optimization technique based on a dual quasi-Newton algorithm. The SAS procedure provides BIC values from which we calculated the weights *w*.

## 2.4 Results

Among the 512 competing models, 96% satisfied convergence criteria but only 35% of the fitted solutions had both high likelihood ranking and, when examined more closely, fitted parameter values that were physically and biologically meaningful. Table 2-2 lists the four models with the highest BIC weights. Note that the summed BIC weights of the first four models are 0.999, indicating almost no empirical support in the data for the 508 models not shown in Table 2-2. The model for African savanna woody cover with best support ( $w=0.88$ ) from this analysis is:

$$tc = \left( t_0 + \frac{\alpha_s \cdot MAP^{\beta_s}}{\gamma^{\beta_s} + MAP^{\beta_s}} \right) \cdot (1 + \alpha_n n) \cdot (1 - \alpha_f f^{\beta_f}) \cdot (1 - \alpha_p p^{\beta_p}) \cdot (1 - \alpha_c c^{\beta_c}) \cdot (1 - \alpha_l l^{\beta_l}).$$

### Equation 2-3

This model accounts for 66% of the variability in the sample dataset and included all the predictors and 18 parameters. Table 2-3 reports the parameter estimates and their 95% confidence intervals. The parameters  $t_0$ ,  $\alpha_s$ ,  $\beta_s$ , and  $\gamma$  define the shape of the sigmoidal rainfall-driven tree cover response. The parameter  $t_0$  represents the minimum tree cover percent,  $\alpha_s$  gives the tree cover range,  $\beta_s$  determines the curvature and  $\gamma$  is the MAP value at  $tc = \alpha_s/2$ . The parameters  $\alpha_s$  and  $\beta_s$  showed a dependence on soil texture categories ( $s$ ). The parameter  $\alpha_n$  defines a linear effect of soil nitrogen, and the parameters  $\alpha_j$  and  $\beta_j$  control the perturbation effects, with  $j=f, c, p$  and  $l$  being indices for fire frequency, cultivation intensity, human population, and cattle density, respectively.

The sigmoid function was selected in all the top-ranked models and hence has the greatest support in the data. There is effectively no support for models structured with linear, piece-wise linear or Michaelis-Menten functions. Associated with the sigmoid

shape, MAP is a better predictor than GSL. The soil texture variable, soil nitrogen and all the perturbation variables, fire frequency, human population, cultivation and cattle density appear in the best model with virtually no uncertainty as to the significance of their effects.

The fitted sigmoid lines represent predicted tree cover on fine, medium and coarse soils for locations with little or no perturbation and low soil nitrogen (Figure 2-4). Fine texture soils appear to have slightly higher tree cover than coarse soils across all rainfall zones with MAP>300 mm. The medium soil texture class is similar to the fine texture soils at low rainfall, but at high rainfall it appears to have lower cover than even the coarse soils (perhaps related to the relatively small number of savanna points in the very high rainfall zones above 1700 mm MAP).

The perturbation effects are shown in Table 2-3 and Figure 2-5. The relationships are plotted with larger symbols to represent the interquartile range to give an idea of the distribution of the perturbation variables across the intensity range. The contribution of soil nitrogen to rainfall-driven tree cover appears to be positive, but for 75% of the points this effect is relatively weak ( $\alpha_n < 0.13$  in Equation 2-3). More frequent fires, higher cattle density, and larger human population densities all tend to depress rainfall-driven tree cover. Cattle normalized values do not exceed 0.04 (10 cattle per km<sup>2</sup>) for 75% of the locations and rainfall-driven tree cover is reduced at the maximum by a factor -0.16 (term  $-\alpha_l l^{bl}$  in Equation 2-3). The interquartile ranges show that suppression of tree cover by fire and human population is often important ( $\sim -0.24$  at their 75<sup>th</sup> percentiles). At the highest extremes, fire and cattle appear to reduce potential tree cover by a factor -0.4. However, these are less strong modifiers than human population that, at its highest levels

in urban areas (8900 people per km<sup>2</sup>), can effectively eliminate tree cover. An unexpected result concerns the positive effect of cultivation: for most of the pixels cultivation appears to have relatively minimal effect on woody cover (<0.002), but the fitted response curve suggests that where cultivation intensity is >0.75 (more than 75% cultivated land) this can result in an increase from rainfall-driven woody cover by a factor 0.1-0.3.

To validate our results, we used the best model (Equation 2-3 and Table 2-1) to predict tree cover on a new random sample that was ten times larger than the pilot sample. The model predictions fit the MODIS observations with a slope of 0.98 (p-value=0.002) and an intercept of -0.03 (p-value=0.5; R<sup>2</sup>=0.65), and show moderate dispersion (root mean square error RMSE = 10.9) (Figure 2-6). The residuals from all predictions were normally distributed with homoschedastic variance, though with some tendency to overestimate tree cover in dry systems and underestimate it at the wetter end of the rainfall gradient (Figure 2-6).

#### **2.4.1 *Perturbation impacts across rainfall zones***

At the continental scale, MAP defines the functional base line for tree cover, while the perturbations increase or, more generally, decrease that tree cover depending on the intensity of the various perturbation factors. We analyzed the average effect of the perturbations at different MAP levels. For this purpose, we divided the MAP range (0-2200 mm) into classes of 100 mm and, for each one, calculated the average values of the perturbations and their modifying effects with respect to the rainfall-driven tree cover  $U$ . Figure 2-7 shows the trends for each perturbation and their summed effects. Maximum fire frequencies occur in the semi-arid and mesic savannas (800-1500 mm; Figure 2-7a) and lead, on average, to 20-25% (-0.2 to -0.25) suppression of the rainfall-driven tree

cover. Human population (Figure 2-7c) in savannas does not exceed average density (averaged within each rainfall interval) of more than about 30 people per km<sup>2</sup> and its largest negative effect (-0.17 to -0.23) on tree cover occurs in areas with MAP higher than 600 mm. Cattle density (Figure 2-7d) reaches its highest averaged value of 16 cattle per km<sup>2</sup> at 800 mm MAP, where it decreases the rainfall-driven tree cover by about -0.15. For MAP < 700 mm, cattle density appears to have a stronger negative effect than fire, but both effects are relatively limited in the drier savannas. The suppressive action of fire increases between 800 and 1500 mm MAP where fire becomes a more important factor in reducing zone-averaged tree cover than human population (Figure 2-7e). The positive effect of cultivation (Figure 2-7b) is generally small, never surpassing a proportion of 0.024 of the rainfall-driven tree cover, which corresponds to a 0.4% increase in actual tree cover at 700 mm. Figure 2-7f compares the rainfall-driven tree cover to the modified tree cover and highlights the fact that the perturbing agents have a relatively minor influence when MAP < 400mm, but they become increasingly important in determining actual tree cover as MAP increases above 600 mm in both relative (reduction with respect to the climate-driven tree cover) and absolute (percent reduction in tree cover) terms.

#### **2.4.2 Model robustness**

Statistical analysis and model fitting should be relatively insensitive to the sample data used for the analysis and model parameterization. We performed the model selection analysis on ten independent random samples of 13,416 points (same number of points as the pilot sample) from the continental dataset. We tallied the proportion  $\pi_i$  of samples for which model  $i$  was selected as the best model among the ten datasets. The first four models selected using the pilot sample (Table 2-2) were also the top ranked

models across all ten independent datasets. The same model (Equation 2-3) was ranked first with a frequency  $\pi_i=0.5$  and weights  $w>0.76$ . Models two and three were each selected twice as “best” and model four was selected once. The parameter estimates of the best models were in general not significantly different from the estimates obtained for the pilot sample. On two occasions the cattle-density effect indicated a weaker grazing effect than in the pilot dataset. These results suggest that, while there is some sensitivity to the sample dataset in the model selection procedure, the balance of evidence supports a model containing all the perturbations, while indicating lower relative importance of the soil nitrogen and cattle density predictors.

We studied the residuals from the predictions with respect to the savanna bioregions to assess whether model performance varies among the different savanna types. The comparison of residuals suggests that the model performs very well across biomes, with little bias among regions, suggesting that the processes controlling tree cover in these diverse savanna types are similar.

Figure 2-8 shows predictions of woody cover for the whole of the African savanna region based on the best model parameters obtained using the pilot sample alongside the MODIS tree cover map. These maps demonstrate the ability of the fitted model to reproduce the broad-scale patterns of tree cover in Africa, but also show that the fine-scale variability is often missed in the model, reflecting in large part, the lack of fine-spatial resolution information on the long-term intensity and impact of many of the important drivers.

## 2.5 Discussion

Savannas are highly dynamic systems, where stochastic events (e.g. wet or dry years, more or less frequent fires) can have considerable impact on woody community demographics, from seed production to seedling emergence, seedling and sapling survival, and adult recruitment. Thus woody cover in any single savanna location is strongly contingent on prior conditions, over one or two years for seedlings and small trees, but over decades for adult trees. Such historical contingency could be used to support the non-equilibrium view of savannas (Ellis & Swift, 1988). That is, savanna structure fluctuates widely, and with no preferred directionality, based on multi-year history of rainfall and disturbance variability. On the other hand, the analysis of Sankaran *et al.* (2005) showed that maximum woody cover is rainfall-limited in arid and semi-arid areas (<650 mm MAP) but not in wetter savannas (>650 mm; Figure 2-4). These results were interpreted to suggest that the drier savannas of Africa are climatically ‘stable’ (canopy closure is prevented by insufficient mean rainfall), while the wetter savannas are ‘unstable’ (canopy closure and grass exclusion are possible). However, in almost all cases, woody cover is well below the mean climate potential woody cover because of climate variability and disturbance events. Thus, most savanna locations are in a ‘disequilibrium’ state. During periods of near average climate and low disturbance, such savanna systems presumably relax, with directional preference, towards the climatic potential cover. However, since climatic variability and disturbance are so prevalent in savannas, the probability of a savanna location actually achieving the woody canopy cover corresponding to the mean climate is low.

In this analysis we asked the question: “what are the processes determining tree cover in African savannas? “. We concentrated on mean rainfall and disturbance statistics across the continent as primary drivers of actual tree cover. We recognize, however, that temporal variability (the historical contingency discussed above) in both rainfall and disturbance can have strong effects on local woody community dynamics. The 500-metre tree cover data product based on MODIS observations (Hansen *et al.*, 2003) provided data at an appropriate scale and resolution for discriminating the main features of tree cover in the African savannas. We tested and applied empirical curves that: (1) reflect hypotheses relating the nature of tree-grass interactions in drought-limited systems, and the impact of disturbances such as fire, herbivory and humans, on tree community structure; and (2) were flexible enough to allow the relationships inherent in the data to determine the shape, direction and magnitude of the response functions. Inherent in our approach is the idea that the disequilibrium dynamics and variability discussed above, and the complexity of the driving factors, are such that the processes are difficult to discern using data from a small number of sites. However, by extracting a very large dataset from the whole of Africa, representing the full range of climatic, soil and disturbance conditions, we hope to identify and quantify both the mean resource (i.e. rainfall) response and the more stochastic and highly variable disturbance-based responses.

The model with the highest support in the data included a sigmoidal increase in tree cover with MAP and strongly supported all the predictor variables. Thus, at the continental scale, rainfall is a primary driver, but it is not sufficient to explain the observations without accounting for the effect of other factors. The selected model

supports a complex behaviour of woody vegetation from which some important general behaviours for the ensemble of African savannas can be extracted.

Three major types of savannas can be defined based on the sigmoidal response to MAP: arid savannas ( $MAP < 400$  mm), where tree cover is low and is relatively insensitive to increasing rainfall; semi-arid/mesic savannas ( $400 < MAP < 1600$  mm), where potential and average tree cover increase rapidly with MAP; and finally mesic/wet savannas ( $MAP > 1600$  mm), where tree cover is high on average, relatively insensitive to MAP and limited by other agents. We obtained a weak response of tree cover to rainfall when  $MAP < 400$  mm. Previous research has shown that the pulse-reserve hypothesis may be particularly relevant for arid systems: woody establishment is driven by non mean-field precipitation characteristics, such as rainfall amounts in particularly wet years, or the seasonal distribution or size of individual rain events (Higgins *et al.*, 2000; Fernandez-Illescas & Rodriguez-Iturbe, 2004; Reynolds *et al.*, 2004). In semi-arid areas, MAP sets a tree cover line around which a large variability is created by perturbations (Figure 2-4). In mesic/wet savannas, we observed a weakening of MAP control on tree cover, where the systems would tend to converge to a wooded state and disturbances are essential for tree-grass coexistence (Sankaran *et al.*, 2005).

The influence on the climatic relationship from environmental factors such as soil nitrogen and soil texture, while statistically significant, is relatively small. Fine soils appear to support higher tree cover across the rainfall gradient but we were not able to detect the differential benefit of soil texture on woody cover postulated by the “inverse texture hypothesis” (Noy-Meir, 1973). Furthermore, Sankaran *et al.* (2005) found a weak positive relationship between soil sand content and tree cover in semi-arid and mesic

savannas (>350 mm MAP), which appears to contradict the inverse texture hypothesis and our results. Thus, while the soil water availability considerations of the inverse texture hypothesis may have direct impacts on vegetation productivity (Sala *et al.*, 1988), they appear less directly linked to the competitive interactions, demographic and disturbance dynamics that control woody cover in African savannas. Soil nitrogen appears to have a positive but relatively weak effect on tree cover, increasing it by a few percent above the moisture-driven level.

The model analysis revealed several relationships between tree cover and the disturbance factors that change with the MAP regime (Figure 2-7). Given the nature of the data set and our analyses, we cannot assess causality, but the results provide (correlative) support for hypotheses regarding the role of different processes in controlling vegetation structure. In arid systems, perturbations have little impact on tree cover. In semi-arid and mesic savannas, MAP controls maximum tree cover and perturbations act as modifiers. The cattle density effect is negative and generally less strong than fire and human population effects. A small positive contribution to tree cover comes in association with cultivation, but overall the combination of perturbation factors causes a decrease from the rainfall-driven tree cover. In mesic and wet savannas (MAP>1600 mm), precipitation no longer strongly limits tree cover and thus the combined effects of disturbance factors are most important in determining tree cover, with fire being more effective than humans and cattle in the emergent tree cover fraction.

The effect of cattle reducing rainfall-driven tree cover was not expected since grazing is commonly considered to be a major cause of woody encroachment (Jeltsch *et al.*, 1997; Roques *et al.*, 2001). This is mostly observed in areas affected by overgrazing

(Scholes & Archer, 1997), but the coarse spatial resolution of the cattle data available at continental scales for Africa tends to spatially homogenize density, with only 0.2% of the points having cattle density higher than  $50 \text{ km}^{-2}$ . Hence the statistical relationship is mainly driven by low-density values (Figure 2-5 and Figure 2-7). Moreover, the absence of shrubs in MODIS tree cover product (Hansen *et al.*, 2003) is likely to make it somewhat insensitive to shrub encroachment *per se*. The important role of wild herbivores in structuring savanna vegetation (Cumming *et al.*, 1997; Scholes & Archer, 1997) could not be analyzed in this study because comprehensive data on wild herbivore biomass outside of the major parks is not available. The increase in woody cover with agricultural intensification (Figure 2-5) is also counter-intuitive but may correspond to the cultivation and protection of shade and fruit trees on the margins of arable fields. Equally important, particularly in the semi-arid regions where cereal crops (millet, sorghum and maize) and fallow rotations are prevalent, is the possibility that rapid resprouting and recolonization of woody species adapted to cultivated areas (e.g. *Guiera senegalensis* in west Africa) may lead to overall small increases in the woody cover of the agricultural areas. Direct cultivation of economically important tree crops such as karite (*Butyrospermum parkii*) and gum Arabic (*Acacia senegal*) in the drier (e.g. Sahelian and Sudanian) regions, or fruit and beverage plantations in the wetter regions (>1800 mm MAP), may also play a role in this relationship.

The MODIS tree cover dataset was created using a supervised regression tree approach (Hansen *et al.*, 2002) and therefore contains spatial and numerical noise and error. The reported standard error of estimates varies from 8 to 18%, which is similar to the residuals of our best models. A particular limitation of the MODIS tree cover dataset,

however, is that woody plants less than 5 m in height were not included in the ground calibration datasets and hence the product tends to underestimate woody cover in savannas where shrubs are common. This may explain the presence of zero or near-zero tree cover in the MODIS dataset in areas where we might expect woody plants to be present in greater abundance (e.g. southern Kalahari and eastern Namibia). This omission may explain why our predicted tree cover values have a tendency to overestimation at the low range of MAP. We also tend to underestimate tree cover, relative to the satellite observations, in high rainfall areas surrounding the moist tropical forests of the Congo Basin (MAP~2000 mm; Figure 2-6 and Figure 2-8). It appears that our model is less effective in these forest–savanna transition zones, possibly because of changing relationships with the various biotic and abiotic factors.

This work relates to the recent study by Sankaran *et al.* (2005), who used data from 850 field sites (~0.25-0.5 ha) across the African continent. We included their quantile regression line as one of the candidate models for this analysis, but it provided overall model fits that were far less efficient in representing the data than the sigmoid curves. As discussed above, the selected sigmoid curves do not predict a maximum, but rather predict the MAP-dependent cover. In fine scale datasets, such as the Sankaran *et al.* study, local heterogeneity (dense or sparse locations) is likely to be more important than in the relatively coarse scale (500 metre = 25 ha pixels) of the MODIS data used here. Nevertheless, the MODIS data do show a small number of locations with tree cover matching the maximum MAP line of the Sankaran *et al.* study (Figure 2-4), suggesting underlying agreement between the datasets despite the scale mismatch.

This study confirms the necessity to look at savannas as systems with multiple dynamic behaviours that only a combination of the resource- and perturbation-based theories can interpret. However, due to the statistical nature of this analysis, we cannot resolve specific mechanisms behind the data trends. The unexplained 44% variability in the data is probably related to data quality issues, and our failure to include certain processes and mechanisms that occur at very fine (patch-landscape) scales. As noted by other authors (Gillson, 2004; Wiegand *et al.*, 2006), research at multiple scales may provide a better understanding of the roles of rainfall (amount and variability), landscape scale processes and both local and broad scale perturbations in determining tree cover. It seems very likely that the complexity of tree cover patterns observable at different scales is an emergent property of different factors, processes and interactions in the progression from plant and patch scales, through landscape to regional and continental scales.

## **2.6 Conclusion**

Savannas ecosystems comprise coexisting herbaceous and woody vegetation forming a continuum between grassland with few scattered shrubs and trees to closed canopy woodland with few grasses. Recognition of the determinants, interactions and dynamics creating and maintaining tree-grass mixtures has been a challenge for ecologists that have developed varied theories. The conceptual framework and the character of the models developed here (intermediate between mechanistic and phenomenological) offer an integrative attempt to build our understanding of vegetation structure in savanna at very coarse (continental) scales.

The selected model was able to identify some of the controls and the processes that create coarse-scale patterns of tree cover in African savannas and explained 66% of

observed variance. Given relatively coarse spatial resolution of some datasets, the predictions were not able to capture all of the tree cover heterogeneity at finer scales seen in the MODIS tree cover (Figure 2-8). Furthermore, our analysis inherits any errors or biases inherent in the tree cover data. We cannot improve on that dataset with our models but we can use the MODIS data to provide a very large sample of data from across the continent and examine relationships and patterns that are difficult, if not impossible, to discern within one region or across a small number of sites.

Our work confirms the need for a synthetic approach to explain tree–grass coexistence based on both resource limitation and disturbance dynamics. Together with the large literature on savanna dynamics and tree–grass interactions at site scale, this research provides some new understanding of the relative importance and interaction of environmental and disturbance factors that create the broad spatial patterns of tree cover observed in savannas across Africa. Whether these patterns will hold for other tropical savannas, or the mixed tree-grass systems of temperate latitudes, remains open to further research.

## **2.7 Acknowledgments**

We wish to thank J. Ratnam, M. Sankaran and W. Sea for their support in the work and helpful discussions on savanna ecology. We are grateful to M. A. Lefsky for his technical collaboration and generous advice on data processing. This research was funded by the National Science Foundation support (Biocomplexity in the Environment project EAR 0120630), with additional support from NASA, Terrestrial Ecology Program, and NOAA, Office of Global Programs.

## 2.8 References

- Akaike, H. (1973) Information theory as an extension of the maximum likelihood principle. Proceedings of the Second International Symposium on Information Theory (ed. by B.N. Petrov and F. Caski), pp. 267-281. Akademiai Kiado, Budapest.
- Allen, R. G., Pereira, L.S., Raes, D. & Smith, M. (1998) Crop evapotranspiration – Guidelines for computing crop water requirements. *FAO Irrigation and drainage paper*, vol. **56**.
- Barbosa, P. M., Stroppiana, D., Gregoire, J.M. & Pereira, J.M.C. (1999) An assessment of vegetation fire in Africa (1981-1991): Burned areas, burned biomass, and atmospheric emissions. *Global Biogeochemical Cycles* **13**, 933-950.
- Belsky, A. J. (1987) Revegetation of natural and human-caused disturbances in the Serengeti National Park, Tanzania. *Vegetatio* **70**, 51-60.
- Bond, W. J. & Keeley, J.E. (2005) Fire as a global 'herbivore': the ecology and evolution of flammable ecosystems. *Trends in Ecology & Evolution*, **20**, 387-394.
- Bond, W. J., Woodward, F.I. & Midgley, G.F. (2005) The global distribution of ecosystems in a world without fire. *New Phytologist*, **165**, 525-537.
- Burnham, K. P. & Anderson, D.R. (2002) *Model selection and inference: a practical information theoretic approach*. New York, Springer-Verlag.
- Cumming, D. H. M., Fenton, M.B., Rautenbach, I.L., Taylor, R.D., Cumming, G.S., Cumming, M.S., Dunlop, J.M., Ford, A.G., Hovorka, M.D., Johnston, D.S., Kalcounis, M., Mahlangu, Z. & Portfors, C.V.R. (1997) Elephants, woodlands and biodiversity in southern Africa. *South African Journal of Science*, **93**, 231-236.
- Dwyer, E., Pereira, J.M.C., Gregoire, J.M. & DaCamara, C.C. (2000) Characterization of the spatio-temporal patterns of global fire activity using satellite imagery for the period April 1992 to March 1993. *Journal of Biogeography*, **27**, 57-69.
- Ellis, J. & Galvin, K.A. (1994) Climate patterns and land-use practices in the dry zones of Africa. *BioScience* **44**, 340-349.
- Ellis, J. E. & Swift, D.M. (1988) Stability of African pastoral ecosystems - alternate paradigms and implications for development. *Journal of Range Management*, **41**, 450-459.
- Fernandez-Illescas, C. P. & Rodriguez-Iturbe, I. (2004) The impact of interannual rainfall variability on the spatial and temporal patterns of vegetation in a water-limited ecosystem. *Advances in Water Resources*, **27**, 83-95.
- Gignoux, J., Clobert, J. & Menaut, J.-C. (1997) Alternative fire resistance strategies in savanna trees. *Oecologia*, **110**, 576-583.
- Gillson, L. (2004) Evidence of hierarchical patch dynamics in an east African savanna? *Landscape Ecology* **19**, 883-894.
- Global Soil Data Task Group (2000) *Global Gridded Surfaces of Selected Soil Characteristics (IGBP-DIS)*. [*Global Gridded Surfaces of Selected Soil Characteristics (International Geosphere-Biosphere Programme - Data and Information System)*]. Data set. Available on-line [<http://www.daac.ornl.gov>] from Oak Ridge National Laboratory Distributed Active Archive Center, Oak Ridge, Tennessee, U.S.A.

- Hansen, M. C., DeFries, R., Townshend, J.R., Carroll, M., Dimiceli, C. & Sohlberg, R. (2003) *Vegetation Continuous Fields MOD44B, 2001 Percent Tree Cover, Collection 3*. College Park, Maryland, 2001, University of Maryland.
- Hansen, M. C., DeFries, R.S, Townshend, J.R.G., Sohlberg, R., Dimiceli, C. & Carroll, M. (2002) Towards an operational MODIS continuous field of percent tree cover algorithm: examples using AVHRR and MODIS data. *Remote Sensing of Environment*, **83**, 303-319.
- Higgins, S. I., Shackleton, C.M. & Robinson, E.R. (1999) Changes in woody community structure and composition under contrasting landuse systems in a semi-arid savanna, South Africa. *Journal of Biogeography*, **26**, 619-627.
- Higgins, S. I., Bond, W.J. & Trollope, W.S.W. (2000) Fire, resprouting and variability: a recipe for grass-tree coexistence in savanna. *Journal of Ecology*, **88**, 213-229.
- Hobbs, N. T. & Hilborn, R. (2006) Alternatives to statistical hypothesis testing in ecology: A guide to self teaching. *Ecological Applications*, **16**, 5-19.
- Hochberg, M. E., Menaut, J.C. & Gignoux, J. (1994) Influences of tree biology and fire in the spatial structure of the West African savannah. *Journal of Ecology*, **82**, 217-226.
- Hoeting, J. A., Madigan, D., Raftery, A.E. & Volinsky, C.T. (1999) Bayesian model averaging: A tutorial. *Statistical Science*, **14**, 382-401.
- Homewood, K., Lambin, E.F., Coast, E., Kariuki, A., Kikula, I., Kivulia, J., Said, M., Serneels, S. & Thompson, M. (2001) Long-term changes in Serengeti-Mara wildebeest and land cover: Pastoralism, population, or policies? *Proceedings of the National Academy of Sciences of the United States of America*, **98**, 12544-12549.
- House, J. I., Archer, S., Breshears, D. & Scholes, R.J. (2003) Conundrums in mixed woody-herbaceous plant systems. *Journal of Biogeography*, **30**, 1763-1777.
- Hutchinson, M. F., Nix, H.A., McMahon, J.P. & Ord, K.D. (1996) The development of a topographic and climate database for Africa. *Proceedings of the Third International Conference/Workshop on Integrating GIS and Environmental Modeling, NCGIA, Santa Barbara, California*. Publisher?
- Jeltsch, F., Milton, S.J., Dean, W.R.J. & Van Rooyen, N. (1997) Analysing shrub encroachment in the southern Kalahari: a grid-based modelling approach. *Journal of Applied Ecology*, **34**, 1497-1508.
- Jeltsch, F., Weber, G.E. & Grimm, V. (2000) Ecological buffering mechanisms in savannas: A unifying theory of long-term tree-grass coexistence. *Plant Ecology*, **150**, 161-171.
- Johnson, J. B. & Omland, K.S. (2004) Model selection in ecology and evolution. *Trends in Ecology & Evolution*, **19**, 101-108.
- Kass, R. E. & Raftery, A.E. (1995) Bayes Factors. *Journal of the American Statistical Association* **90**, 773-795.
- Laris, P. (2002) Burning the seasonal mosaic: Preventative burning strategies in the wooded savanna of southern Mali. *Human Ecology*, **30**, 155-186.
- Matthews, E. (1983) Global vegetation and land use: new high resolution data bases for climate studies. *Journal of Climate and Applied Meteorology*, **22**, 474-487.

- Metzger, K. L., Coughenour, M.B., Reich, R.M. & Boone, R.B. (2005) Effects of seasonal grazing on plant species diversity and vegetation structure in a semi-arid ecosystem. *Journal of Arid Environments*, **61**, 147-160.
- Noy-Meir, I. (1973) Desert ecosystems: environment and producers. *Annual Review of Ecology and Systematics*, **4**, 25-51.
- Prins, H. H. T. & H. P. Van der Jeugd (1993) Herbivore population crashes and woodland structure in East Africa. *Journal of Ecology*, **81**, 305-314.
- Reynolds, J. F., Kemp, P.R., Ogle, K. & Fernández, R.J. (2004) Modifying the 'pulse-reserve' paradigm for deserts of North America: precipitation pulses, soil water, and plant responses. *Oecologia*, **141**, 194-210.
- Roques, K. G., O'Connor, T.G. & Watkinson, A.R. (2001) Dynamics of shrub encroachment in an African savanna: relative influences of fire, herbivory, rainfall and density dependence. *Journal of Applied Ecology*, **38**, 268-280.
- Sala, O. E., Parton, W.J., Joyce, L.A. & Lauenroth, W.K. (1988) Primary Production of the Central Grassland Region of the United-States. *Ecology*, **69**, 40-45.
- Sankaran, M., Ratnam, J. & Hanan, N.P. (2004) Tree-grass coexistence in savannas revisited – insights from an examination of assumptions and mechanisms invoked in existing models. *Ecology Letters*, **7**, 480-490.
- Sankaran, M., Hanan, N.P., Scholes, R.J., Ratnam, J., Augustine, D.J., Cade, B.S., Gignoux, J., Higgins, S.I., Le Roux, X., Ludwig, F., Ardo, J., Banyikwa, F., Bronn, A., Bucini, G., Caylor, K.K., Coughenour, M.B., Diouf, A., Ekaya, W., Feral, C.J., February, E.C., Frost, P.G.H., Hiernaux, P., Hrabar, H., Metzger, K.L., Prins, H.H.T., Ringrose, S., Sea, W., Tews, J. Worden, J. & Zambatis, N. (2005) Determinants of woody cover in African savannas. *Nature*, **438**, 846-849.
- SAS (2000) *SAS/STAT users guide*. Cary, NC, USA, SAS Institute.
- Scholes, R. J. & Archer, S.R. (1997) Tree-grass interactions in savannas. *Annu. Rev. Ecol. Syst.*, **28**, 517-544.
- Schwarz, G. (1978) Estimating dimension of a model. *Annals of Statistics* **6**, 461-464.
- Shackleton, C. M., Guthrie, G. & Main, R. (2005) Estimating the potential role of commercial over-harvesting in resource viability: A case study of five useful tree species in South Africa. *Land Degradation & Development*, **16**, 273-286.
- Sinclair, A. R. E. & Fryxell, J.M. (1985) The Sahel of Africa – ecology of a disaster. *Canadian Journal of Zoology-Revue Canadienne De Zoologie*, **63**, 987-994.
- Skarpe, C. (1991) Impact of grazing in savanna ecosystems. *Ambio*, **20**, 351-356.
- Udvardy, M. D. F. (1975) *A classification of the biogeographical provinces of the world*. IUCN Occasional Paper No. 18, prepared as a contribution to UNESCO's Man and the Biosphere (MAB) Program, Project No. 8. International Union for the Conservation of Nature and Natural Resources, Morges (now Gland), Switzerland.
- Van de Koppel, J. & Prins, H.H.T. (1998) The importance of herbivore interactions for the dynamics of African savanna woodlands: an hypothesis. *Journal of Tropical Ecology*, **14**, 565-576.
- van Langevelde, F., van de Vijver, C., Kumar, L., van de Koppel, J., de Ridder, N., van Andel, J., Skidmore, A.K., Hearne, J.W., Stroosnijder, L., Bond, W.J., Prins, H.H.T. & Rietkerk, M. (2003) Effects of fire and herbivory on the stability of savanna ecosystems. *Ecology*, **84**, 337-350.

- Walker, B. H. & Langridge, J.L. (1997) Predicting savanna vegetation structure on the basis of plant available moisture (PAM) and plant available nutrients (PAN): a case study from Australia. *Journal of Biogeography*, **24**, 813-825.
- Walker, B. H. & Noymeir, I. (1982) Aspects of the stability and resilience of savanna ecosystems. *Ecological Studies*, **42**, 557-590.
- Wiegand, K., Saitz, D. & Ward, D. (2006) A patch-dynamics approach to savanna dynamics and woody plant encroachment – insights from an arid savanna. *Perspectives in Plant Ecology Evolution and Systematics*, **7**, 229-242.
- Wintle, B. A., McCarthy, M.A., Volinsky, C.T. & Kavanagh, R.P. (2003) The use of Bayesian model averaging to better represent uncertainty in ecological models. *Conservation Biology*, **17**, 1579-1590.
- Zobler, L. (1986) A world soil file for global climate modeling. *NASA Tech. Memo. 87802*, NASA., *place of publication?*
- Zucchini, W. (2000) An introduction to model selection. *Journal of Mathematical Psychology*, **44**, 41-61.

## TABLES AND FIGURES

**Table 2-1:** Datasets used in analysis of African savanna tree cover dynamics. The dataset consists of ten layers. The last column relates each variable to equations 1 and 3. The lower case letter defines the variable symbol. The letter *U* stands for the “undisturbed tree cover” term and *M* for the “modifications by perturbation” term. Summary statistics, mean and range, are reported.

<b>Data layer</b>	<b>Origin</b>	<b>Description</b>	<b>Symbol and summary statistics</b>
<b>Tree-cover percent</b>	500m MODIS Vegetation Continuous Fields (Hansen, 2003)	Proportional estimates of tree-cover derived from monthly composites of MODIS data of 2002. Resolution: 500 m; Units: percent	$t$ (response variable) $\bar{t} = 16\%$ min=0%, max= 89%
<b>Mean annual precipitation (MAP)</b>	ANU-CRES fitted climatic grids (Hutchinson, 1996)	Total monthly rainfall values averaged over 1951-1995 period. Resolution: 0.05 degrees; Units: mm	$r; U$ $\bar{r} = 769$ mm min=11mm, max= 2192 mm
<b>Growing season length (GSL)</b>	- ANU-CRES fitted climatic grids (Hutchinson, 1996) - Potential Evapotranspiration (Allen, 1998)	GSL = (average monthly rainfall /PET) * (days per month) (Sankaran, unpublished data) Resolution: 0.05 degree; Units= days	$g; U$ $\bar{g} = 127.5$ days min=2 days, max= 337 days
<b>Soil texture</b>	Zobler soil datasets (texture layer) Source: UNEP/GRID (Zobler, 1986)	Global distribution of soil texture classes Resolution: 1 degree	$s; U$ Classes: fine, medium and coarse
<b>Soil Nitrogen</b>	Global Gridded Surfaces of Selected Soil Characteristics. Source: IGBP-DIS (Group, 2000).	Total nitrogen density. Resolution: 0.0833 degrees; Units: g/m <sup>2</sup>	$n; U$ $\bar{n} = 832$ g/m <sup>2</sup> min=0 g/m <sup>2</sup> , max= 4875 g/m <sup>2</sup>
<b>Fire frequency</b>	GVM, Joint Research Center, Italy	Burned areas in 1981-1991 derived from	$f; M$

	(Barbosa, 1999 #139; Barbosa, 1999)	NOAA-AVHRR GAC data over the period 1981-1991. Resolution: 8km; Units: fire return over 8 year	$\bar{f}$ =2 years min=0 years, max= 1 year
<b>Human population density for 2002</b>	Oak Ridge National Laboratory (ORNL) LandScan Global Population Database 2002: <a href="http://www.ornl.gov/sci/gist/landscan/">http://www.ornl.gov/sci/gist/landscan/</a>	Census counts (at sub-national level). Resolution: 30" X 30" latitude/longitude grid; Units: number of people per cell	$p$ ; $M$ $\bar{p}$ =19 km <sup>-2</sup> min=0 km <sup>-2</sup> , max=8906 km <sup>-2</sup>
<b>Cultivation intensity</b>	International Livestock Research Institute (ILRI); UNEP-GRID & GISS (Matthews, 1983)	Percentage that is under cultivation, versus the percentage of natural vegetation, including five classes. Resolution: 1 degree; Units: percent	$c$ ; $M$ $\bar{c}$ =16% min=0%, max=100%
<b>Cattle density</b>	World Resources Institute - PAGE, 2000	Density of cattle in Africa compiled by the International Livestock Research Institute (ILRI). Resolution: 0.090597 degrees; Units: Cattle per km <sup>2</sup>	$l$ ; $M$ $\bar{l}$ = 7 km <sup>-2</sup> min= 0 km <sup>-2</sup> , max= 242 km <sup>-2</sup>
<b>Biome</b>	IUCN, as digitized by UNEP/GRID in 1986. Created by (Udvardy, 1975).	Biogeographical provinces of the world, Africo-tropical Realm: Biogeographical provinces are defined as ecosystematic or biotic subdivisions of the realms (floral "regions" and faunal "provinces").	

**Table 2-2:** Four models with highest posterior model probabilities based on the Bayesian information criterion (BIC). All these models are structured with the sigmoid function. Presence of a predictor in a model is marked by a dot ( $r$ =mean annual precipitation,  $g$ =growing season length,  $s$ =soil texture,  $n$ =soil nitrogen,  $f$ =fire frequency,  $c$ =cultivation intensity,  $p$ =human population, and  $l$ =cattle density).  $K$  is number of free parameters; for each model  $i$ ,  $\Delta_i = \text{BIC}_i - \text{BIC}_{\min}$  and  $w_i$  is the derived BIC weight (model posterior probability). MAP, mean annual rainfall; GSL, growing season length.

Model no.	Climate pot.		Soil		Perturbations				Statistics			BIC	AIC
	$r$	$g$	$x$	$n$	$f$	$c$	$p$	$l$	$K$	$\Delta_i$	$L(\theta_i Y)$	$w_{BIC}$	$w_{AIC}$
1	•		•	•	•	•	•	•	18	0	1.0000	0.8786	0.9999
2	•		•	•	•		•	•	16	4	0.1353	0.1189	7.5E-05
3	•		•		•	•	•	•	17	12	0.0025	0.0022	7.5E-05
4	•		•		•		•	•	15	16	0.0003	0.0003	3.4E-09
$W_+$	1.0	0.0	1.0	0.998	1.0	0.881	1.0	1.0					

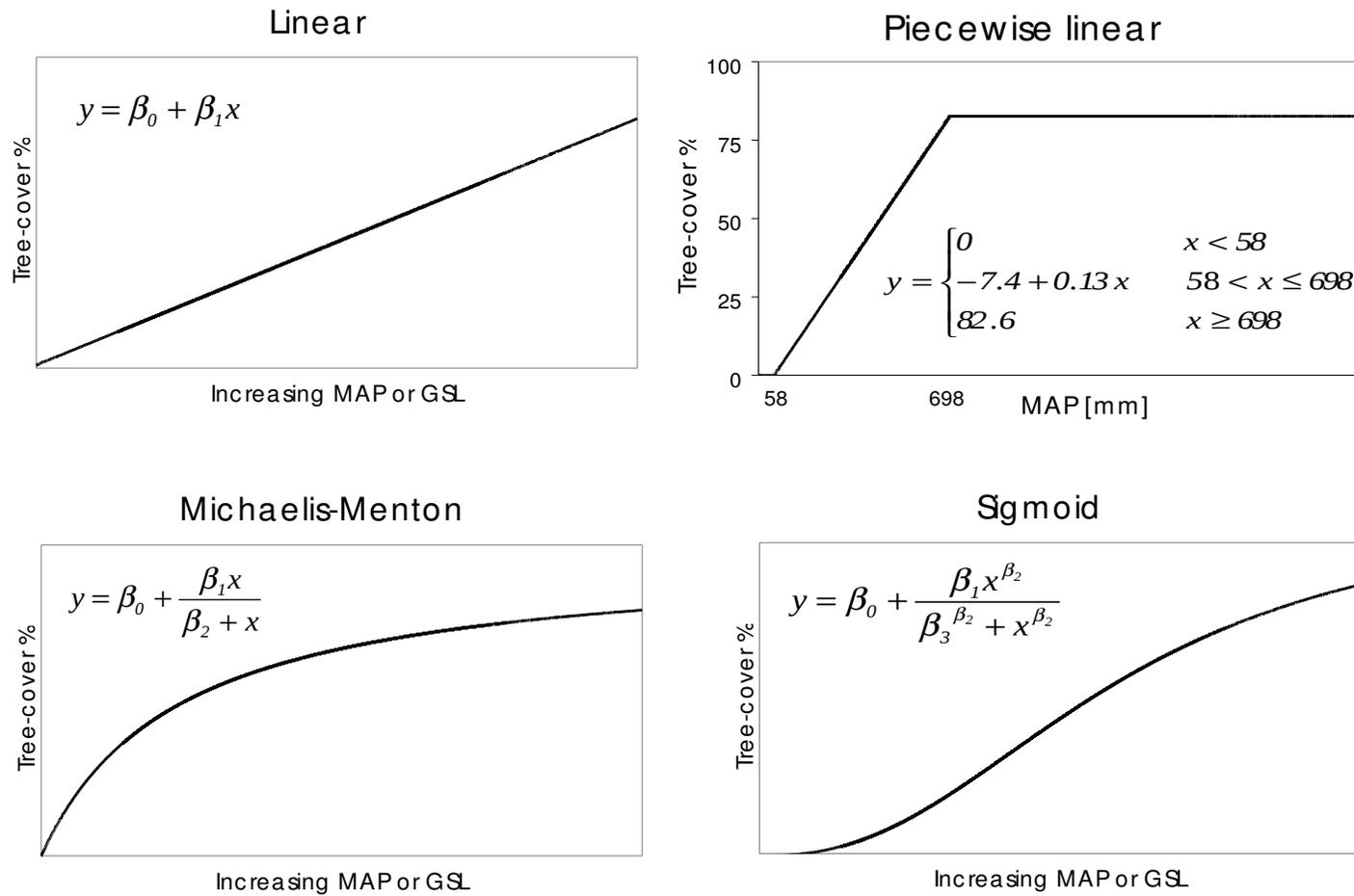
**Table 2-3:** Summary statistics for parameter estimates using the sigmoid relationship and MAP. The indices  $xf$ ,  $xc$  and  $xm$  for fine are for coarse and medium texture classes. Parameters for the effects of soil nitrogen, fire frequency, cultivation, human population and cattle density are also shown. Values for parameter estimates, unconditional standard error (Eq. 6) and 95% confidence interval (eq. 7) are retrieved from model averaging.

Variable	Par	Estimate	Uncond. s.e.	Uncond. 95% C.I.	
MAP	$t_0$	-0.89	0.25	-1.39	-0.39
	$\alpha_{xf}$	117.89	6.92	104.33	131.46
	$\alpha_{xc}$	110.65	7.09	96.76	124.54
	$\alpha_{xm}$	110.92	5.99	99.17	122.67
	$\beta_{xf}$	2.25	0.09	2.06	2.43
	$\beta_{xc}$	2.58	0.10	2.39	2.76
	$\beta_{xm}$	1.96	0.06	1.84	2.08
	$\gamma$	1435.77	65.70	1306.97	1564.57
Soil Nitrogen	$\alpha_n$	0.27	0.07	0.14	0.40
Fire frequency	$\alpha_f$	0.38	0.01	0.36	0.40
	$\beta_f$	0.54	0.04	0.46	0.61
Cultivation intensity	$\alpha_c$	-0.40	0.10	-0.61	-0.20
	$\beta_c$	3.59	0.63	2.35	4.83
Human population	$\alpha_p$	1.01	0.05	0.91	1.10
	$\beta_p$	1.24	0.06	1.12	1.36
Cattle density	$\alpha_l$	0.45	0.05	0.36	0.54
	$\beta_l$	0.33	0.04	0.26	0.40

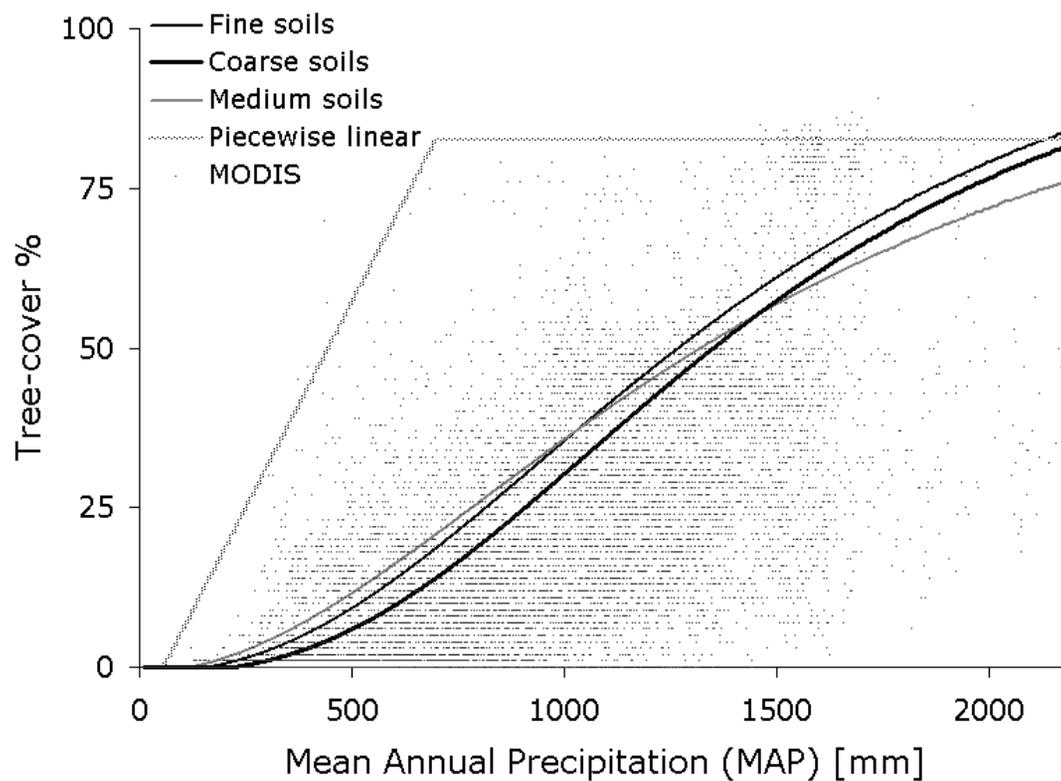
**Table 2-4:** Models frequencies  $\pi_i$  obtained from the ten samples compared to the weights  $w$  obtained from the pilot sample. Models 1 to 4 are listed like in Table 2; models 5 and 6 are the two new best models.  $K$  is the model size.

Model no.	$K$	$\pi_{AIC}$	$w_{AIC}$	$\pi_{BIC}$	$w_{BIC}$
1	18	0.7	0.9999	0.5	0.8786
2	16	0.1	7.5E-05	0.1	0.1189
3	17	0.1	7.5E-05	0.2	0.0022
4	15		3.4E-09	0.1	0.0003
5	14			0.1	
6	15	0.1			

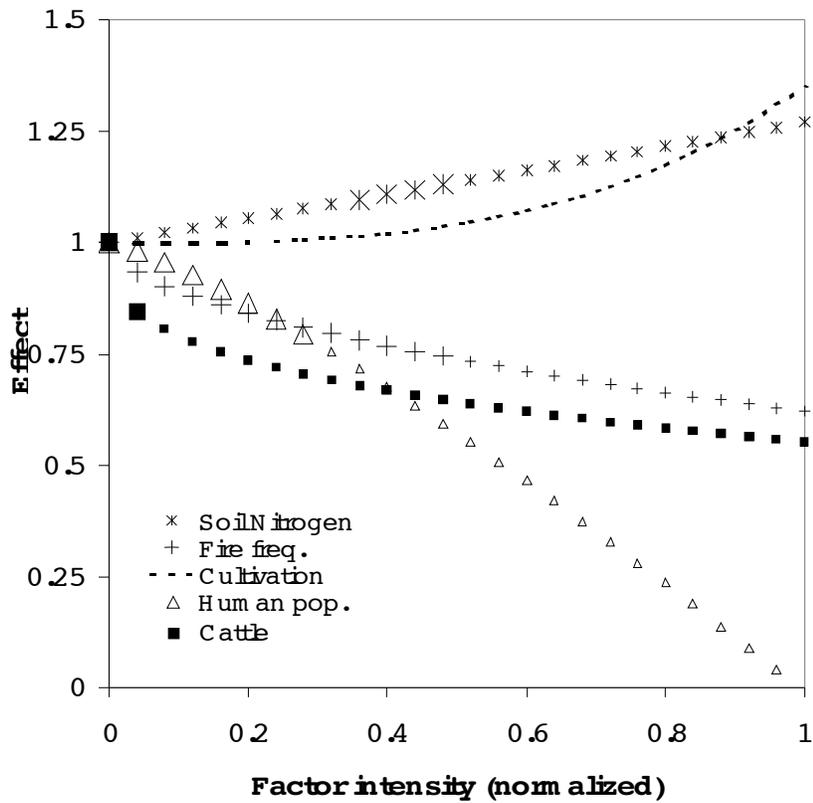




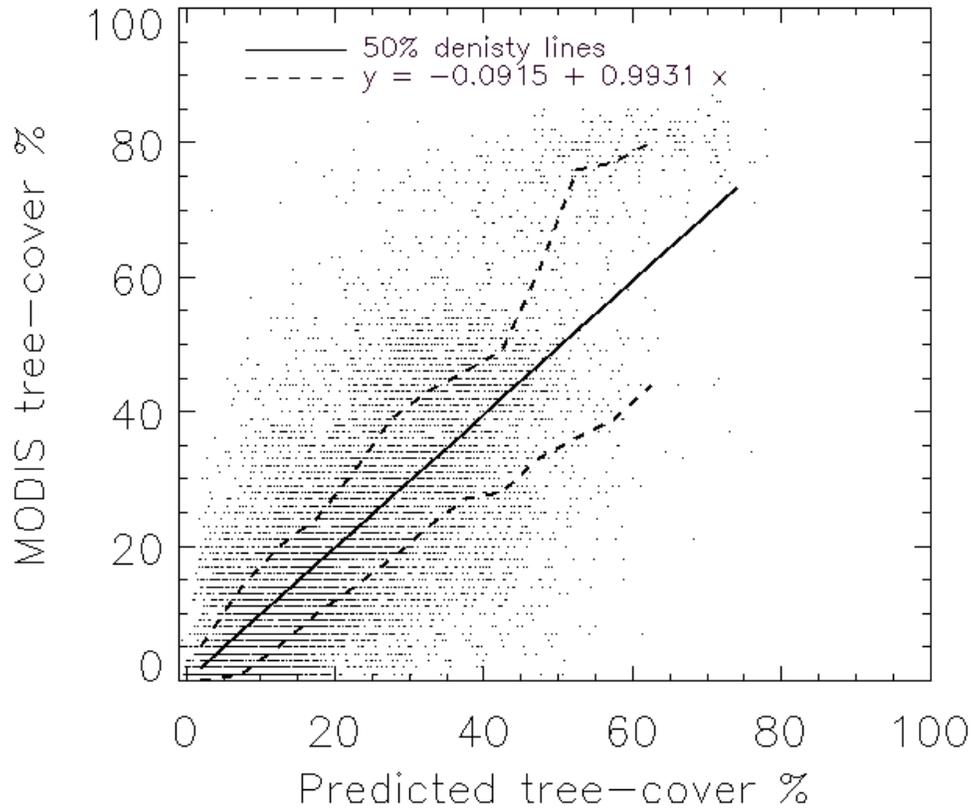
**Figure 2-3:** Hypothesized functional responses of tree-cover to moisture. The independent variable ( $q$ ) could be either MAP or GSL.



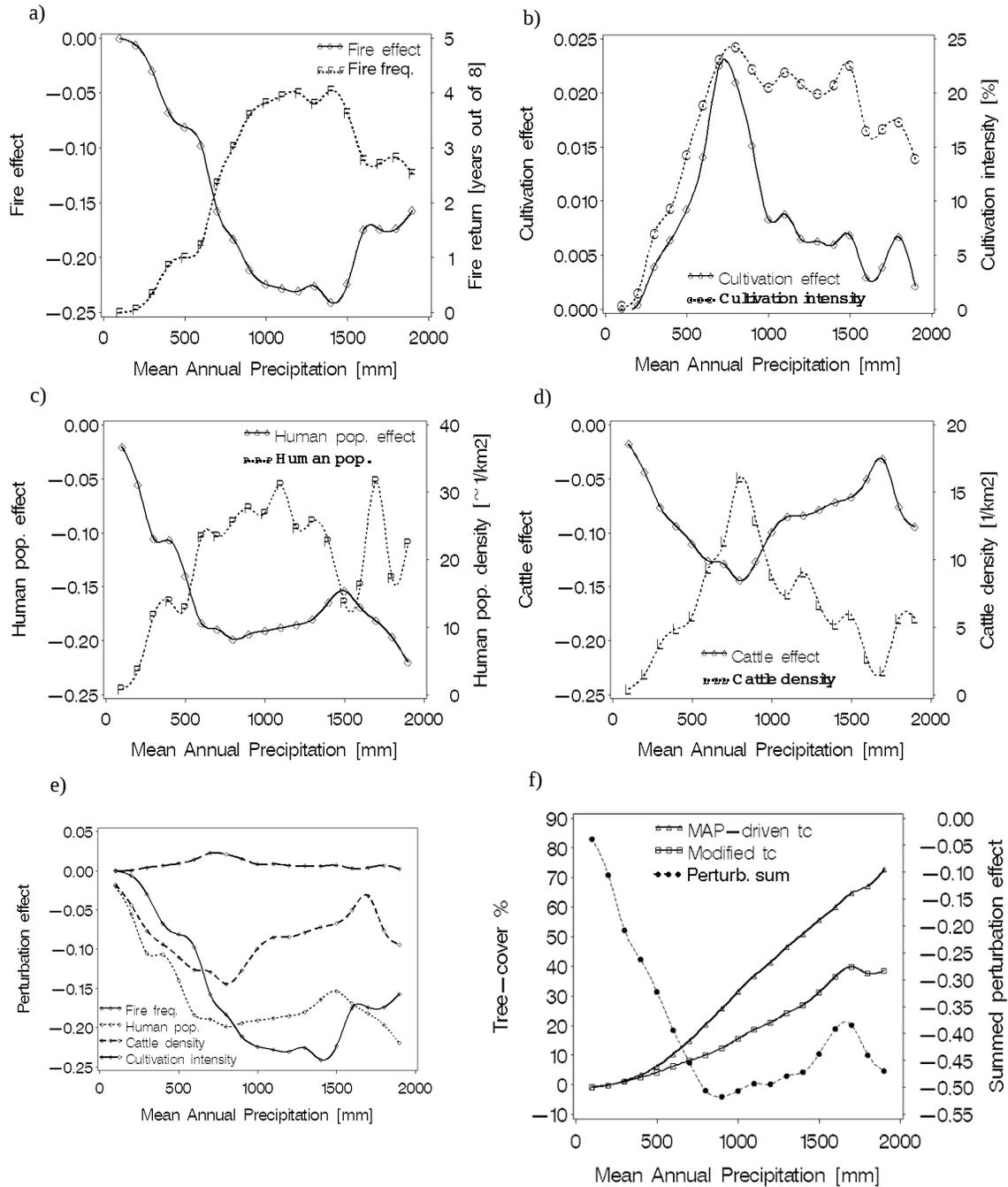
**Figure 2-4:** Tree-cover percent versus MAP: MODIS-based tree-cover% (dots), predicted rainfall-driven tree-cover on three calls of soil texture, fine, medium, and coarse (solid lines), and piecewise linear relationship marking the upper bound on tree-cover found by Sankaran et al. (2005).



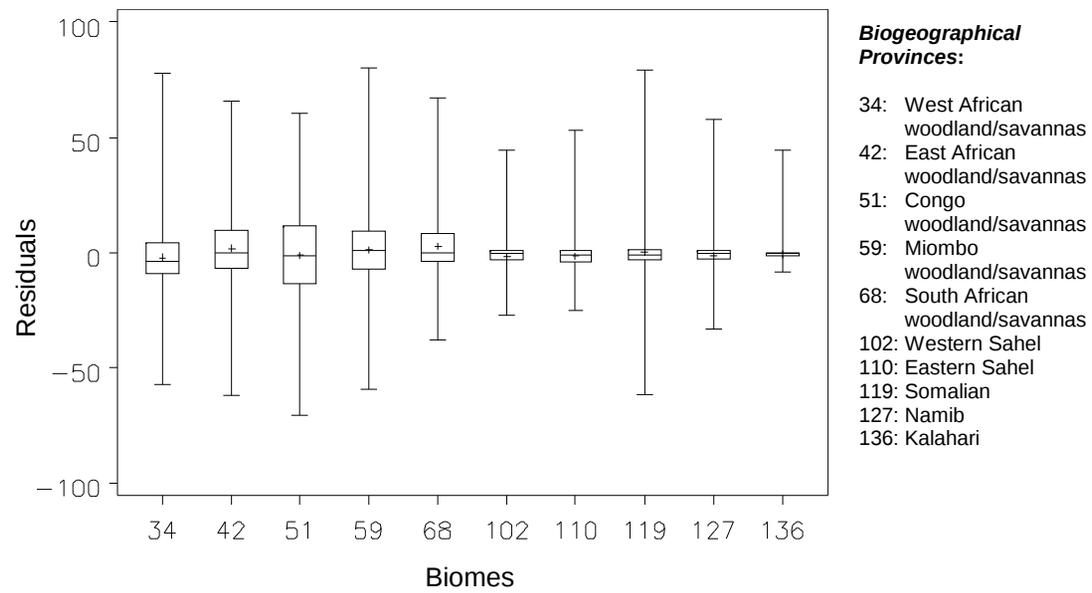
**Figure 2-5:** Perturbation and Soil Nitrogen multiplicative effects on potential tree-cover. Larger symbols mark the response interval for the interquartile range of a factor.



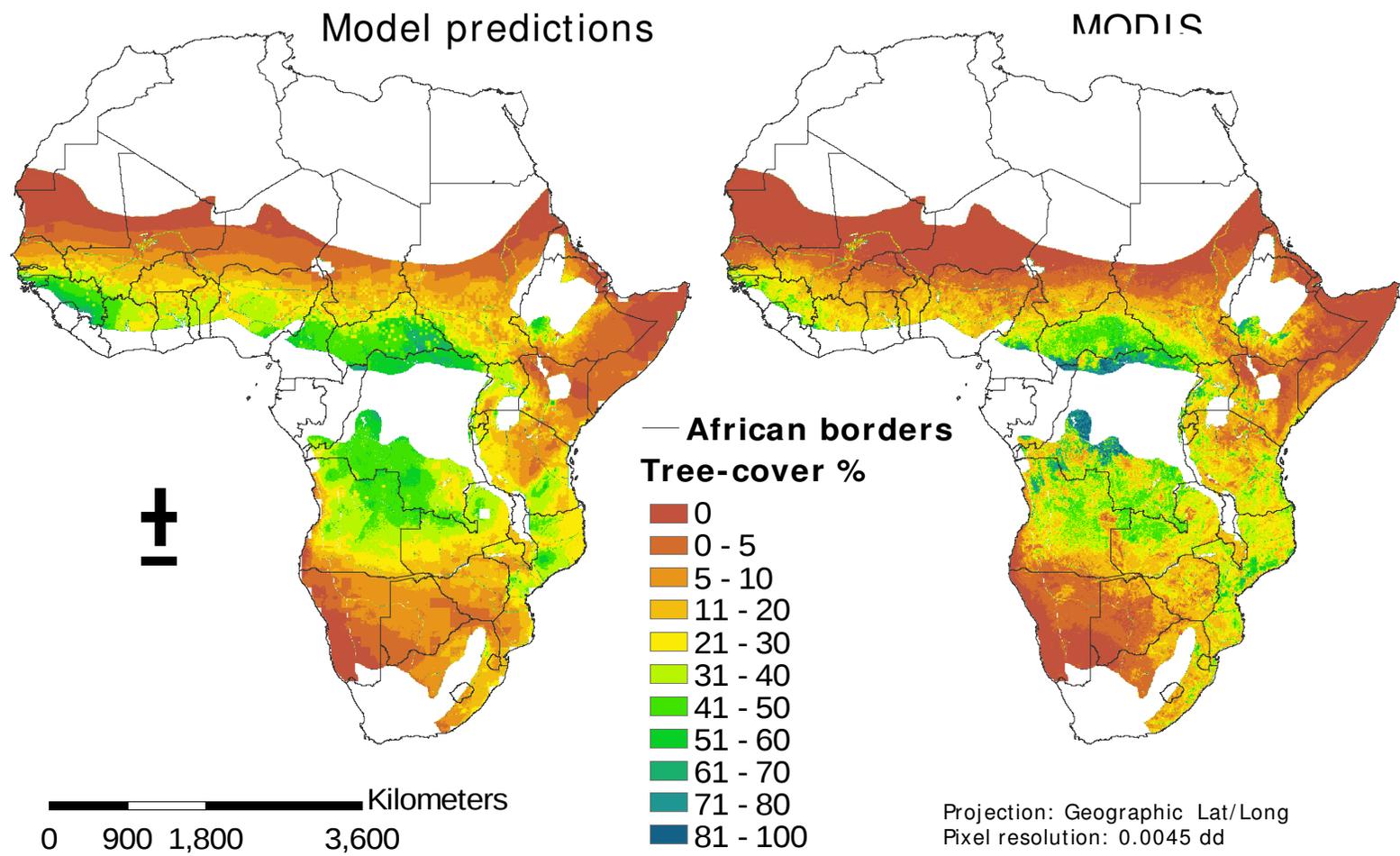
**Figure 2-6:** Observed vs. predicted tree-cover%: linear fit.



**Figure 2-7:** Averaged perturbation effects in relation to MAP. The top four panels a), b), c) and d) report the perturbation values and their corresponding effect as the fraction of loss (fire frequency, human population and cattle density) or gain with respect to the environmental tree-cover. The graph e) groups the perturbation effects and the graph f) reports the averaged environmental tree-cover and the modified tree-cover under the influence of the averaged summed effects of the perturbations.



**Figure 2-8:** Residuals' distributions by bioregion. The line inside the box: Median; Plus symbol marker: mean; Lower and upper edges of the box: 25<sup>th</sup> and 75<sup>th</sup> percentiles; endpoints of lower and upper whiskers: minimum and maximum.



**Figure 2-9:** Tree-cover percent in African savannas. Left: model averaged predictions. Right: retrieved from MODIS sensed data.

## CHAPTER 3

### **Woody Fractional Cover in Kruger National Park, South Africa: remote-sensing-based maps and ecological insights**

**Gabriela Bucini<sup>1</sup>, Niall P. Hanan<sup>1</sup>, Randall B. Boone<sup>1</sup>, Izak P.J. Smit<sup>2</sup>, Sassan  
Saatchi<sup>3</sup> Michael A. Lefsky<sup>4</sup> and Gregory P. Asner<sup>5</sup>**

1. Natural Resource Ecology Laboratory, Colorado State University, Fort Collins, CO 80523-1499, U.S.A.
2. South African National Parks, Scientific Services, ZA-1350 Skukuza, South Africa.
3. Jet Propulsion Laboratory, California Institute of Technology, Mail Stop 300-319 Pasadena, CA 91109, U.S.A.
4. Department of Forest, Rangeland and Watershed Stewardship, Colorado State University, Forestry Building, Fort Collins, Colorado 80523, U.S.A.
5. Department of Global Ecology, Carnegie Institution for Science, 260 Panama Street Stanford, California 94304, U.S.A.

**PUBLISHED:** Ecosystem function in savannas: measurement and modeling at landscape to global scales (2010) .M. J. Hill and N. P. Hanan, CRC/Taylor and Francis.

### 3.1 Introduction

The variability in woody vegetation structure characterizes different types of savannas and is usually described by quantitative variables such as fractional cover, height and biomass. Woody cover has been shown to directly affect important ecological processes of savanna ecosystems. It influences biomass production, fire regimes, herbivory, nutrient cycling, hydrology and soil erosion (Rietkerk et al., 1997; Scholes and Archer, 1997; Archibald et al., 2009). Savannas are characterized by varied spatial combinations and relative proportions of woody and herbaceous vegetation. This heterogeneity results in diversified habitats and resources determining the large number of species that can establish and survive in these ecosystems. The ability to monitor woody fractional cover is therefore fundamental to understand savanna dynamics.

Mean annual precipitation (MAP) determines a potential maximum cover for woody plants, which however is rarely reached because other regional and local factors intervene (Sankaran et al., 2005; Bucini and Hanan, 2007). The variability in observed woody cover at the regional scale can be the result of inter-annual variability in rainfall, as well as local variability in soil properties and geomorphology, disturbance history, and patch dynamics, all of which impact tree-grass interactions and woody plant survival and recruitment (Ogle and Reynolds, 2004; Wiegand et al., 2006; Groen et al., 2008; Sankaran et al., 2008).

Kruger National Park (KNP) has a tradition of science-informed management that has recently focused on maintenance of heterogeneity as a means to create resilience. Woody cover assessments have direct impacts on the Park management decisions in relation to the extent they reflect and control fire effects, herbivore utilization, elephant

damage, animal habitat suitability and hydrology. The most recent study based on aerial photography covering the KNP southern-central part, Eckhardt et al. (2000) found that between 1940-1998, woody cover slowly increased on granite bedrocks and decreased on basalt bedrocks. The most compelling result was a significant decline in trees > 5 m in all sampled transects. Elephant numbers and interactions between elephant damage and fire were found the principal drivers of this decrease. The scale of woody cover dynamics in KNP has a combined local- and regional nature but the Park still lacks high-resolution wall-to-wall woody cover maps that allow a thorough cover assessment.

Current techniques to estimate woody canopy cover at the regional level are based on remote-sensing scaling-up approaches. At the fine scale, field instruments, aerial photographs and high-resolution images provide cover estimates that can be used to calibrate remote-sensing images with coarser resolution and full coverage of the study area. The main difficulty in woody cover detection in arid and semi-arid systems is to separate woody vegetation from the background of grass and soil. Research on woody cover mapping has tended to concentrate either on optical or radar-based techniques. While optical reflectances and derived vegetation indices are very sensitive to photosynthetically active vegetation (e.g., green leaf area), they are not optimal to differentiate woody from herbaceous attributes (Glenn et al., 2008). Synthetic aperture radar (SAR) sensors, operating in the microwave C, L and P bands, are instead sensitive to woody structure and biomass, including in low-density savanna situations (Santos et al., 2002; Lucas et al., 2006; Saatchi et al., 2007). The radar backscatter signal, however, is affected by spatial and temporal variability in soil and canopy moisture, as well as surface roughness. The potential for combined optical and SAR systems should to be

explored, given the limitations and strengths in discriminating unique woody vegetation signatures using each technology separately.

In this case study, optical and radar imagery are combined to produce a woody cover map for the whole KNP at medium resolution (~90 m) and a map that quantifies woody-cover heterogeneity as a potential metric for biodiversity studies (Rogers, 2003). Then the mapped woody cover patterns are analyzed in relation to environmental and ecological factors using a classification and regression tree (CART) approach to construct an explanatory model of woody cover distribution (e.g. Breiman et al., 1984; De'Ath and Fabricius, 2000).

### **3.2 Study Area**

KNP is located in the lowveld semi-arid savanna of northeastern South Africa (22.3-25.5° S, 30.8-32° E; Figure 3-10). The northern part of the park receives 300-500 mm MAP and is classified as arid bushveld (Venter et al., 2003). The southern part receives 500-700 mm MAP and is classified as lowveld bushveld. The long-term MAP is  $506.6 \pm 144$  (mean  $\pm$  standard deviation) and rainfall distribution has a monomodal pattern where the rainy season occurs mainly between October and March (austral summer). KNP exhibits a longitudinal geomorphological division between underlying basalt bedrock (east) that forms soils rich in clay minerals and granite bedrock (west) that produces coarser sandy soils. The Park's terrain is fairly level (average slope of  $1.6 \pm 2.5^\circ$ ) except for the presence of few geological formations, varying from rugged outcrops and gorges to floodplains, in the north, in the southwestern corner and along the eastern border.

The Park is subjected to natural disturbance, including drought cycles, fire and herbivory. The physiognomy of the woody plants varies and reflects species genetic characteristics as well as phenotypic adaptations to disturbance. Woody plant heights fall in the 2-5 m range and can assume the forms of trees and shrubs with single or multi-stemmed physiognomy. Based on 1955-2004 data, fires occur with an overall average return of 4.4 years (frequency  $\sim 0.22$ ). Fires are mostly surface fires and they can suppress seedlings and saplings. Adult plants taller than 3-m generally survive and have the capacity to resprout from roots after a dieback. Large browsers (elephants, giraffes and rhinos) have high potential to alter structure and compositions of plant communities (Shannon et al., 2008; Asner et al., 2009).

### **3.3 Data and Methods**

#### **3.3.1 Data**

##### *3.3.1.1 Field measurements*

The field campaign was carried out during April-May 2006. The 2005-2006 rainy season was above long-term average and the trees still had their green canopy. We measured canopy cover using a handheld spherical densiometer (Lemmon, 1956), a simple and quick method comprising a convex mirror etched with a grid of 24 squares within which the observer estimates cover. Our measurements included woody plants taller than 1.3 m and we subtracted canopy gaps present inside the crown. Because of the 60° angle of view, densiometer measurements have been shown to overestimating cover (Korhonen et al., 2006; Ko et al., 2009). However, savanna woody plants are generally scattered and we conjectured that the bias would not be significant.

Field plots were distributed to sample the woody cover range in KNP as well as the rainfall and soil gradients but did not include riparian areas. We designated 73 plots of 250x250 m, with 63 plots laid out in the Experimental Burnt Plots (EBPs) and 10 plots in other areas. The EBPs have been maintained at different fire season and frequency treatments since 1954 (Biggs et al., 2003) and provide representative cover range levels in adjacent plots. In each plot, we collected 30 densiometer measurements on a regular 6x5 point grid with 50-m pace and averaged them to obtain the plot woody cover estimate. The sampled data ranged from 0.4-58 % cover with an average of 23.8 %  $\pm$ 13.7 percentage point (pp).

### *3.3.1.2 Remote Sensing Data and Processing*

We sought medium resolution imagery with good quality and no cloud cover. We also wanted dry-season images to minimize the moisture contamination in the radar imagery and maximize the contrast between photosynthetically active woody vegetation and senescent grass. When we started working (2006), the imagery available for the KNP with these characteristics included Landsat Enhanced Thematic Mapper (ETM+) and JERS-1 Synthetic Aperture Radar (SAR).

The three orthorectified ETM+ images (source: Global Land Cover Facility, <http://glcfapp.umiacs.umd.edu/index.shtml>) were acquired between May 21 and 30, 2001 and covered the entire Park except for a small area (1%), for which we found an image from June 12, 2000. The images were mosaicked into a single scene at 28.5-m resolution using ENVI (<http://www.itvis.com>) color adjustment, a method that uses a least-squares adjustment to ensure radiometric consistency among the scenes. Using the visible, near-infrared and short-wave infrared bands, we calculated the Normalized Difference

Vegetation Index (NDVI), the Soil Adjusted Vegetation Index (SAVI; Huete, 1988), the Modified Soil Adjusted Vegetation Index (MSAVI; Qi et al., 1994) and the first three principal components (PCA). The 25-m resolution JERS-1 SAR imagery (source: Alaska Satellite Facility <http://www.asf.alaska.edu/>) included 11 georeferenced scenes of backscatter amplitude acquired in the L-band (1275 MHz, 23.5 cm) with HH polarization and at 35-degree off-nadir angle. Eight scenes were acquired between April 12 and August 20, 1996 and three scenes on March 12, 1995. In the two weeks previous to image acquisition, the average gauge rainfall was  $20 \pm 19$  mm in 1996 and  $7 \pm 8$  mm in 1995, which should not significantly contaminate the backscatter. We applied a Lee filter (standard deviation based filter) to smooth speckles in the radar scenes but no radiometric corrections were made. Interaction with surface roughness produced some high backscatter areas unrelated to vegetation information. We corrected this effect by creating a mask for pixels with backscatter amplitude values  $>45$  DN (Digital Number) which were generally facing east (aspect  $45^\circ$ - $135^\circ$ ) on a slope  $>2.5^\circ$  and assigning them a reduced amplitude value of 45 DN. We mosaicked the radar images and coregistered it to the georectified Landsat mosaic (rms  $<1$  pixel, UTM 36 S projection), resampling at a common pixel resolution of 28.5 m. We then ran a focal analysis over all the remote-sensing layers to smooth image coregistration errors and worked at 85.5 m nominal resolution.

In between data acquisition dates, local woody cover change likely occurred because of elephants and other large browsers' plant utilization, also in combination with fire. We decided to run the analysis while recognizing that the data could record local changes and give rise to some prediction instability. We ultimately validated our model predictions

using an independent woody cover dataset obtained from the Carnegie Airborne Observatory (CAO, <http://cao.stanford.edu>) that integrates high-fidelity imaging spectrometers and waveform LiDAR (light detecting and ranging) sensors (Asner et al., 2007). The measurements were collected at 50 kHz laser pulse repetition frequency from an average flying altitude of 2,000 m above ground level, with a + 17 degree scan angle (after 2 degree cut-off) and 50% overlap between adjacent flight lines. This resulted in LiDAR data with 1.0 m laser spot spacing and 1.12 m beam diameter.

### *3.3.1.3 Geospatial input data for CART*

The geospatial data included climate, geomorphology, soils, vegetation, herbivory and fire data (Table 3-5). Most of the original data were provided without error estimates. We estimated errors for the variables derived from the original data through modeling. Data with a temporal dimension (rainfall, fire, herbivore census) were compiled into long term averages, discarding years with missing records. All data coverages were reprojected to a common format (UTM 36 S) and 1-km pixel spacing.

### **3.3.2 Analysis and modeling**

Woody-cover was predicted using field measurements and remote-sensing data. The explanatory CART model (Brieman et al., 1984) was derived using the environmental and ecological databases (Table 3-5). The regression tree (CART) can be used to explain variation in a response variable such as woody cover, by inputting a range of explanatory variables to split the response variable into more homogeneous groups. Models may be selected by: a) iterative comparison based on a hypothesis test; or b) selected on the basis of the minimum error which can be assessed by a penalty for

complexity using Akaike's Information Criteria (AIC; Akaike, 1973) or a cross validation/bootstrapping approach (De'Ath and Frabricsius, 2000). All analysis and modeling was conducted using the software package R (version 8.0.2; <http://www.r-project.org>).

### *3.3.2.1 Woody cover and heterogeneity maps*

A relationship to calibrate the optical and radar data (predictors) was developed using the measured woody cover at our 73 field sites (response variable). The initial remote-sensing dataset was reduced to eight covariates after eliminating one of a variable pair with Pearson correlations coefficient  $>0.8$ . Then multiple linear regressions were run and the AIC was used to select an optimal model. A jack-knife procedure was used to check the stability of the model parameter estimates. The best model was extrapolated to estimate woody cover percent over the entire KNP. We also quantified the relative contribution of the predictive variables to the total explanatory value of the model using the LMG metric (Gromping, 2006). From the map, the woody cover standard deviation was calculated as a metric for heterogeneity on a 1-km scale to identify where woody cover presents high variability hot spots. This is a scale at which many birds and mammals might perceive heterogeneity in the landscape.

### *3.3.2.2 Explanatory ecological model*

We studied the mapped woody-cover in association with ecological and environmental conditions (Table 3-5). We randomly sampled 1000 points (1-km square grid cells) with a minimum 2-km reciprocal distance (training set) and applied the regression tree analysis (CART, Breiman et al., 1984). Minimum cross-validation error

was used to retrieve the optimal tree size. To verify the stability of the optimal tree, we created 10 independent random samples and grew 10 separate regression trees. After pruning, we compared them to the regression tree grown on the training set to check that there was no, or minimal, deviation in tree structure and variable selection. A random forest analysis (Breiman, 2001; a combination of bootstrap aggregation and random selection of variables) was run to retrieve variable relative importance.

In order to validate the model, we applied the regression tree on 10 independent test samples of 1000 points and compared their predictions to the observed (mapped) woody cover. We evaluated the relative weight of the ecological/environmental vs. spatial component in the woody cover patterns following the procedure described by Boone and Krohn (2000) based on Borcard et al. (1992). The total variance is portioned into independent components: pure environmental and ecological variation  $v_e$ ; pure spatial variation  $v_s$ ; environmental/ecological variation with spatial structuring  $v_{es}$  and; unexplained variation  $v_u$ . Accordingly, we created three types of woody cover models (CART):

1) ecological (with ecological and environmental independent variables) explaining the variation;

$$VE = v_e + v_{es} \quad (\text{Equation 1})$$

2) spatial (with UTM x and y coordinates as independent variables) explaining the variation;

$$VS = v_s + v_{es} \quad (\text{Equation 2})$$

3) and ecological-spatial (with both ecological/environmental and spatial independent variables) explaining the variation;

$$\text{VES} = \text{ve} + \text{vs} + \text{ves}. \quad (\text{Equation 3})$$

The single variation components can then be retrieved knowing that:

$$\text{ves} = \text{VE} + \text{VS} - \text{VES} \quad (\text{Equation 4})$$

and

$$\text{vu} = 1 - (\text{ve} + \text{vs} + \text{ves}) \quad (\text{Equation 5})$$

### 3.4 Results

#### Woody Cover Percent and Heterogeneity Maps

The optimal regression model (Table 3-6) that links field measurements to the remote sensing variables explains 61% of the variability ( $p < 0.0001$ ) with a residual standard error (RSE) of 8.9 %. The residuals were near-normally distributed with a slight positive skew. The JERS-1 backscatter intensity and the Landsat green band together contributed 0.58 in the total  $R^2$ . According to the LMG metric for relative variable importance (normalized to sum 1), the JERS-1 backscatter intensity ( $\text{LMG} = 0.43 \pm 0.14$ ) and the Landsat green band ( $\text{LMG} = 0.38 \pm 0.11$ ) were the two most important variables to predict woody cover. The jack-knife analysis (Table 2) indicated model stability and absence of points in the dataset with high leverage that could create biased estimates. The model performance (Figure 3-11) is good but underestimations occur for some points having relatively dense woody cover ( $> 50\%$ ). The validation of the model predictions against the LiDAR cover estimates gives strong agreement (Figure 3-11,  $p < 0.0001$ ) and the residual standard error ( $\text{RSE} = 8.9\%$ ) was in the order of the model error.

The KNP woody cover map generated from the model is shown in Figure 3-12. The park is characterized by low to medium cover ( $\text{WC} = 34.7\% \pm 14.6 \text{ pp}$ ) with the granite landscapes (west) supporting higher woody cover ( $\text{WC} = 40.2\% \pm 11.8 \text{ pp}$ ) than

basalt landscapes (WC=24.4 % ± 14.4 pp). Two emerging features are the sedimentary rocks (WC=41.5 % ± 15.3 pp) separating the granitic and basaltic bedrocks in the south and the intrusive gabbro sills in the granites with relatively low cover (WC=19.7 % ± 9.5 pp). The northern mountainous areas have the highest average cover (WC= 45.5 % ± 13.4 pp). The heterogeneity map (Figure 3-11) brings to light the transitions between the savanna domain and the riparian areas (std. dev.>14 %) and the outcrops (std. dev. 8-14 %). Fairly large homogeneous patches occur on the basalt bedrocks and generally correspond to relatively flat upland areas far from drainage lines. 9

#### **3.4.1 Woody Cover Explanatory Model**

The best explanatory model (Figure 3-13) derived from the CART analysis explained 59.4 % of the variability in the mapped woody cover. The regression tree structure was stable and the predictions calculated for 10 test samples matched the mapped woody cover with an averaged  $R^2=0.54$  ( $p<0.0001$ ) and root mean squared error of 8 %. The random forest function assigned the highest relative importance to the fire frequency, slope, MAP and “basalt-granite” variables. The first splits in the regression tree occurred for the binary geological variable “basalt- granite”. The two branches were not characterized by differences in woody cover value ranges but rather by two separate ecological behaviors occurring on the two different parent material layers. On the basalt rocks, the areas with lowest woody cover (WC=17 %) are located where fire recurs with frequency greater than once every five years (0.2), slope  $<1.1^\circ$  and MAP  $<567$  mm. Less than five-year fire return appears necessary for covers  $>32$  %. The granite branch is first split in the northern and the eastern blocks. The CART did not further resolve the cover variability in the northern zone, predicting overall WC=49 %. In the east, areas with

MAP < 518 mm present less woody cover with crests (DEM >391 m a.s.l.) having the lowest cover (WC=28 %). If MAP >518 mm, elephants density >866 kg/km<sup>2</sup> correlates with reduced woody cover.

Partitioning of variation among ecological, spatial and combined models (Figure 3-14) showed that the variation in woody cover distribution was predominantly explained by ecological and environmental variables, since the pure spatial component explained only 2 % of the variability. Two thirds (~40 %) of the variation explained by the ecological model was associated with the shared ecological/spatial component. The other third (20 %) has a pure ecological/environmental origin coming from variable interactions.

### **3.5 Discussion**

Woody fractional cover and woody cover heterogeneity are important quantitative descriptors of savanna structure that capture biophysical processes driven by geomorphology, climate and disturbances such as fire and herbivory. We produced the first KNP woody cover map at medium-high (90 m) resolution. We also derived a woody cover heterogeneity map and a woody cover ecological model.

#### **3.5.1 *Woody Cover Map***

It is helpful to combine optical and radar imagery for woody cover up-scaling and mapping: the JERS-1 L band and the Landsat ETM+ green band were the most important variables correlated with woody fractional cover. The strong correlation ( $r=0.67$ ) between woody cover percent and radar L-band backscatter is consistent with the results found in the savannas of Queensland, Australia (Lucas et al., 2006). In general, high

backscattering is expected from tree boles because of their size similar to that of the L-band wavelength (23.5 cm). The signal return results from both direct backscatter and from stem-ground double bounce interaction (Durden et al., 1989). Topography contributes to woody plant establishment by creating local favorable conditions related to energy balance, higher soil moisture or protection from fire and/or herbivory. However, woody cover overestimates could arise from enhanced backscatter due to aggregated structural information from both vegetation and topography (van Zyl, 1993). Our corrective approach, based on applying a mask (with a fixed intensity value) on topographic areas, was a rudimentary effort that can be substituted with robust radiometric corrections. The negative correlation ( $r=-0.63$ ) between the ETM+ green band and woody cover is in agreement with research showing that in arid and semi-arid systems, increasing vegetation cover corresponds to an increase in visible light absorption by the canopy leaves (mostly in blue and red bands) and a decrease in the background brightness from dry soil and senescent grasses (Yang and Prince, 1997; Xu et al., 2003). This behavior however can be non-linear especially when accounting for shadows and complex backgrounds. We initially tested simple polynomial relationships but the increased model complexity penalized model performance gains.

The overall good agreement with the woody cover estimates retrieved from LiDAR (Figure 3-11) means that our map gives an accurate woody cover representation at the KNP scale. Our approach however has limitations that influenced both the amount of model unexplained variability and model error. They can be related to woody cover temporal variations occurred during the 11-year (1995-2006) span of data acquisition dates, field measurement errors and remote-sensing signal variability (shadows and the

background contamination effects; Spanner et al., 1990). Disturbances and vegetation demographic changes caused cover fluctuations during the data acquisition years. Our mapped woody cover is a model-combined value of different year signals and therefore it is not a point-in-time representation. Local spatial variability carries also a temporal component. The expected cover bias (overestimate) due to the densiometer design did not emerge in the validation comparison (Figure 3-11). The exclusion of canopy gaps in the densiometer measurements could have compensated the bias introduced by the densiometer angular view and lead to a good match with LiDAR estimates that are closer to crown cover (gaps included in cover at the operated resolution). Savanna sparse vegetation probably reduced the densiometer angular effect, too. Given the densiometer problematic aspects, other field methods such as the line intercept or even a modified spherical densiometer would be better suited for their unbiased sampling and accuracy (Korhonen et al., 2006; Ko et al., 2009). The LiDAR cover deviations from the map predictions in the low woody cover range (Figure 3-11) are instead caused by the LiDAR sampling approach used for this dataset. It tends to miss plants  $\leq 2$  m and hence underestimate cover in the more arid areas dominated by short woody plants (Bucini et al., unpublished manuscript).

### **3.5.2 Ecological Model**

Both the regression tree and the variance partitioning analyses highlighted a strong spatial character of the woody plant patterns that arises from interactions with spatially structured environmental factors. The analysis was in accord with pieces of previous KNP research and combined them into an integrated model. Soil substrates and water availability are strong factors in determining woody cover. Excluding the

mountainous areas, woody plants can reach fractional covers >30 % with MAP>575 mm on basalt and with MAP>518 mm on granite where moisture is more available because of clay soil types. Water is more easily found in valleys and depressions. The two tree branches “slope>1°” and “DEM<391m” mainly correspond to this location types. In particular, “slope>1°” separates both low elevation lands and some outcrops in the basalt side. Despite the longstanding concern on elephant impact (Eckhardt et al., 2000; Guldmond and Van Aarde, 2008), our model highlights only an area on the granites where elephant density is relatively high and woody cover is 30 % although MAP >518 mm. Contrary to our expectations, the variables representing seasonal or annual rainfall variability only appeared at lower unstable nodes indicating either an effect at smaller scales or sensitivity to dataset idiosyncrasies. The 40 % unexplained variability is likely to be linked to stochasticity in spatial and temporal drivers removed in the temporal means and spatially gridded data, and resulting population dynamics (Wiegand et al., 2006) that the ecological and environmental predictors do not capture. To resolve some finer-scale variability, we also run a similar model analysis maintaining the layers “distance to water”, “elevation”, “slope” and “aspect” at 90-m resolution but no improvement was reached.

### ***3.5.3 Research and Management Implications***

The procedures used in this work are of interest for implementing a Park monitoring program that will collect woody-cover data for long-term monitoring and ecological assessments. In particular, heterogeneity has emerged in KNP as a management goal in association with maintenance of biodiversity (Rogers, 2003; Venter et al., 2008). The derived woody cover and heterogeneity maps can provide information

at a scale equivalent to the watershed and landscape scales at which the Park is managed. Yet, given the KNP low-to-medium woody cover, the current 8.9 % model error needs to be reduced. Very high-resolution imagery and LiDAR have the potential to provide a better spatial and temporal sampling of the KNP complex vegetation structure and hence to improve up-scaling remote-sensing calibration models.

### **3.6 Acknowledgements**

The authors thank Dr. R. Scholes for the valuable discussions on remote sensing and savanna ecology and Sandra MacFadyen, Navashni Govender and Johan Baloyi from the Kruger Park Scientific Services for their help with the data collection. This research was funded by support from the National Science Foundation (NSF) on the project Biocomplexity in African Savannas (EAR 0120630) and the National Aeronautics and Space Administration (NASA), Terrestrial Ecology Program on the project African Carbon Exchange. The Carnegie Airborne Observatory is funded by the W.M. Keck Foundation, William Hearst III, the Andrew Mellon Foundation, and the Gordon and Betty Moore Foundation.

### 3.7 References

- Akaike, H. 1973. *Information theory as an extension of the maximum likelihood principle*. Second International Symposium on Information Theory, Akademiai Kiado, Budapest.
- Archibald, S., D. P. Roy, B. W. van Wilgen and R. J. Scholes 2009. What limits fire? An examination of drivers of burnt area in Southern Africa. *Global Change Biology* 15, 613-630.
- Asner, G. P., D. E. Knapp, T. Kennedy-Bowdoin, et al. 2007. Carnegie Airborne Observatory: in-flight fusion of hyperspectral imaging and waveform light detection and ranging (wLiDAR) for three-dimensional studies of ecosystems. *Journal of Applied Remote Sensing* 1, 21
- Asner, G. P., S. R. Levick, T. Kennedy-Bowdoin, et al. 2009. Large-scale impacts of herbivores on the structural diversity of African savannas. *Proceedings of the National Academy of Sciences of the United States of America* 106, 4947-4952.
- Biggs, R., H. C. Biggs, T. T. Dunne, N. Govender and A. L. F. Potgieter 2003. Experimental burn plot trial in the Kruger National Park: history, experimental design and suggestions for data analysis. *Koedoe* 46, 1-15.
- Boone, R. B. and W. B. Krohn 2000. Partitioning sources of variation in vertebrate species richness. *Journal of Biogeography* 27, 457-470.
- Borcard, D., P. Legendre and P. Drapeau 1992. Partialling out the Spatial Component of Ecological Variation. *Ecology* 73, 1045-1055.
- Breiman, L. 2001. Random forests. *Machine Learning* 45, 5-32.
- Breiman, L., J. Friedman, R. Olshen and C. Stone 1984. *Classification and Regression Trees*. Belmont, California, Wadsworth.
- Bucini, G. and N. P. Hanan 2007. A continental-scale analysis of tree cover in African savannas. *Global Ecology and Biogeography* 16, 593-605.
- De'Ath, G. and K. E. Frabricsius 2000. Classification and regression trees: a powerful yet simple technique for ecological data analysis. *Ecology* 81,3178-3192.
- Durden, S. L., J. J. van Zyl and H. A. Zebker 1989. Modeling and observation of the radar polarization signature of forested areas. *IEEE Transactions on Geoscience and Remote Sensing* 27, 290-301.
- Eckhardt, H. C., B. W. van Wilgen and H. C. Biggs 2000. Trends in woody vegetation cover in the Kruger National Park, South Africa, between 1940 and 1998. *African Journal of Ecology* 38, 108-115.
- Efron, B. and R. J. Tibshirani 1993. *An Introduction to the Bootstrap*. New York and London, Chapman & Hall/CRC.
- Gertenbach, W. P. D. 1983. Landscapes of the Kruger National Park. *Koedoe* 26, 9-121.
- Glenn, E. P., A. R. Huete, P. L. Nagler and S. G. Nelson 2008. Relationship between remotely-sensed vegetation indices, canopy attributes and plant physiological processes: What vegetation indices can and cannot tell us about the landscape. *Sensors* 8, 2136-2160.
- Groen, T. A., F. van Langevelde, C. de Vijver, N. Govender and H. H. T. Prins 2008. Soil clay content and fire frequency affect clustering in trees in South African savannas. *Journal of Tropical Ecology* 24, 269-279.

- Gromping, U. 2006. Relative importance for linear regression in R: The package relaimpo." *Journal of Statistical Software* 17, 1-27.
- Guldemon, R. and R. Van Aarde 2008. A meta-analysis of the impact of African elephants on savanna vegetation. *Journal of Wildlife Management* 72, 892-899.
- Huete, A. R. 1988. A Soil-Adjusted Vegetation Index (SAVI). *Remote Sensing of Environment* 25, 295-309.
- Ko, D., N. Bristow, D. Greenwood and P. Weisberg 2009. Canopy cover estimation in semiarid woodlands: comparison of field-based and remote sensing methods. *Forest Science* 55, 132-141.
- Korhonen, L., K. T. Korhonen, M. Rautiainen and P. Stenberg 2006. Estimation of forest canopy cover: A comparison of field measurement techniques. *Silva Fennica* 40, 577-588.
- Lemmon, P. E. 1956. A spherical densiometer for estimating forest overstory density. *Forest Science* 2, 314-320.
- Lucas, R. M., A. C. Lee and M. L. Williams 2006. Enhanced simulation of radar backscatter from forests using LiDAR and optical data. *IEEE Transactions on Geoscience and Remote Sensing* 44, 2736-2754.
- Ogle, K. and J. F. Reynolds 2004. Plant responses to precipitation in desert ecosystems: integrating functional types, pulses, thresholds, and delays. *Oecologia* 141, 282-294.
- Qi, J., A. Chehbouni, A. R. Huete, Y. H. Kerr and S. Sorooshian 1994. A modified soil adjusted vegetation index. *Remote Sensing of Environment* 48, 119-126.
- Rietkerk, M., F. vandenBosch and J. vandeKoppel 1997. Site-specific properties and irreversible vegetation changes in semi-arid grazing systems. *Oikos* 80, 241-252.
- Rogers, K. H. 2003. Adopting a Heterogeneity Paradigm: Implications for Management of Protected Savannas. In, *The Kruger Experience: Ecology and Management of Savanna Heterogeneity*, ed. J. T. du Toit, K. H. Rogers and H. C. Biggs, 41-58. Washington, Island Press.
- Saatchi, S., K. Halligan, D. G. Despain and R. L. Crabtree 2007. Estimation of forest fuel load from radar remote sensing. *IEEE Transactions on Geoscience and Remote Sensing* 45, 1726-1740.
- Sankaran, M., N. P. Hanan, R. J. Scholes, et al. 2005. Determinants of woody cover in African savannas. *Nature* 438, 846-849.
- Sankaran, M., J. Ratnam and N. Hanan 2008. Woody cover in African savannas: the role of resources, fire and herbivory. *Global Ecology and Biogeography* 17, 236-245.
- Santos, J. R., M. S. P. Lacruz, L. S. Araujo and M. Keil 2002. Savanna and tropical rainforest biomass estimation and spatialization using JERS-1 data. *International Journal of Remote Sensing* 23, 1217-1229.
- Scholes, R. J. and S. R. Archer 1997. Tree-grass interactions in savannas. *Annual Review of Ecology and Systematics* 28, 517-544.
- Shannon, G., D. J. Druce, B. R. Page, H. C. Eckhardt, R. Grant and R. Slotow 2008. The utilization of large savanna trees by elephant in southern Kruger National Park. *Journal of Tropical Ecology* 24, 281-289.
- Spanner, M. A., L. L. Pierce, D. L. Peterson and S. W. Running 1990. Remote-sensing of temperate coniferous forest leaf-area index – the influence of canopy closure, understory vegetation and background reflectance. *International Journal of Remote Sensing* 11, 95-111.

- van Zyl, J. J. 1993. The effect of topography on radar scattering from vegetated areas. *IEEE Transactions on Geoscience and Remote Sensing* 31, 153-160.
- Venter, F. J. 1990. A classification of land for management planning in the Kruger National Park. PhD.-Dissertation. University of South Africa., University of South Africa.
- Venter, F. J., R. J. Naiman, H. C. Biggs and D. J. Pienaar 2008. The evolution of conservation management philosophy: Science, environmental change and social adjustments in kruger national park. *Ecosystems* 11, 173-192.
- Venter, F. J., R. J. Scholes and H. C. Eckhardt 2003. The Abiotic Template and Its Associated Vegetation Pattern. *The Kruger Experience: Ecology and Management of Savanna eterogeneity*. Editors: J. T. du Toit, Rogers, K. H. and Biggs H. C. 83-129. Washington, Island Press.
- Wiegand, K., D. Saitz and D. Ward 2006. A patch-dynamics approach to savanna dynamics and woody plant encroachment - Insights from an arid savanna. *Perspectives in Plant Ecology Evolution and Systematics* 7, 229-242.
- Xu, B., P. Gong and R. L. Pu 2003. Crown closure estimation of oak savannah in a dry season with Landsat TM imagery: comparison of various indices through correlation analysis. *International Journal of Remote Sensing* 24, 1811-1822.
- Yang, J. L. and S. D. Prince 1997. A theoretical assessment of the relation between woody canopy cover and red reflectance. *Remote Sensing of Environment* 59, 428-439.

## TABLES AND FIGURES

**Table 3-5.** : Environmental and ecological variables used in the explanatory model. The KNP source reference is the South African National Parks (SANParks) database (<http://www.sanparks.org/parks/kruger/conservation/scientific/gis/>). Most of the rainfall gauges had more than 20 years of data except for 3 gauges with 12 years of data. We kept elephants as a separate group because they can affect woody vegetation not only as mixed-feeders but also through direct mortality of trees. Abbreviations: RSE=residual standard error; sd=standard deviation; “ = same as above.

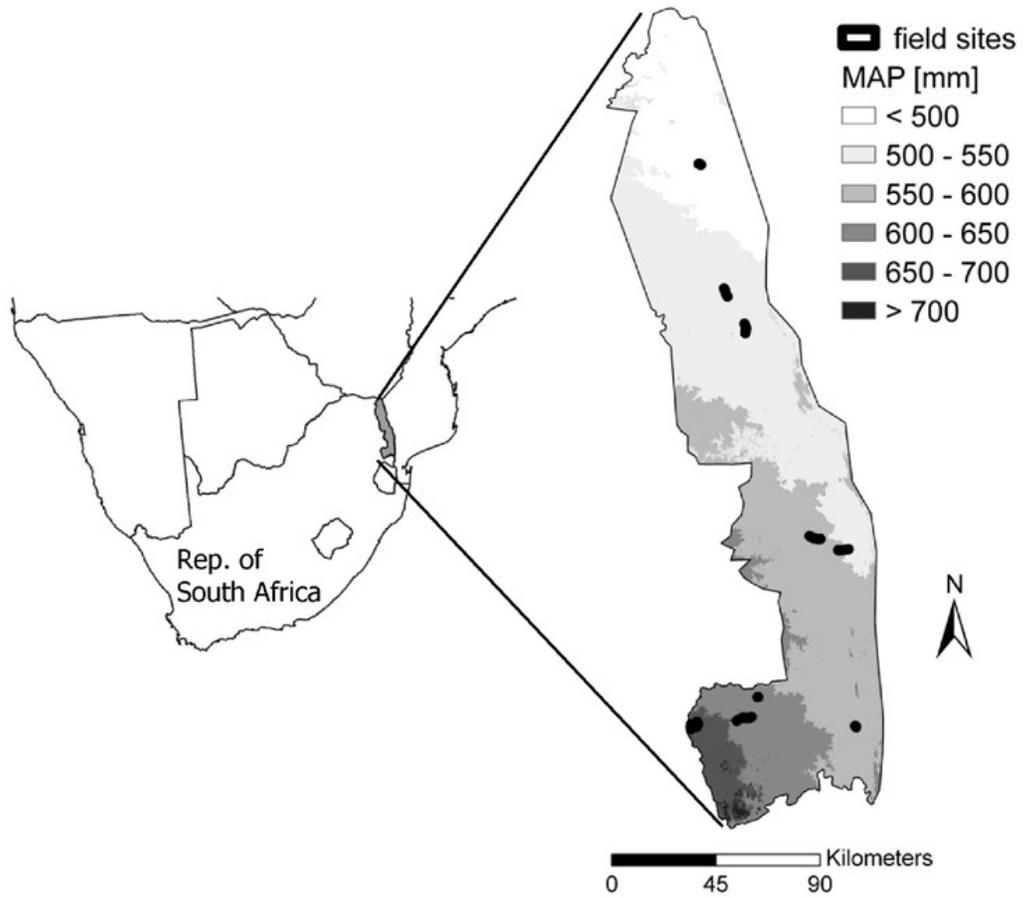
Variable name and statistic	Notes
<b><u>Climate</u></b>	KNP database; 26 rain gauges, 1970-2004 Daily rainfall
Mean annual precipitation (mm)	Calculated: generalized regression model with x, y coordinates and elevation predictors (R <sup>2</sup> =0.75, p<0.0001, RSE= 38.8)
Variability of mean annual precipitation variability – coefficient of variation (CV)	Calculated: generalized regression model with x, y coordinates and elevation predictors (R <sup>2</sup> =0.31, p=0.01, RSE=0.06)
Mean annual precipitation in the dry season (mm)	Calculated: generalized regression model with x, y coordinates predictors (R <sup>2</sup> =0.63, p<0.0001, RSE=10.9)
Variability of mean annual precipitation in the dry season (CV)	Calculated: generalized regression model with x, y coordinates predictors (R <sup>2</sup> =0.3, p=0.004, RSE=0.08)
Average length of dry spells in the wet season (No. days)	Calculated: generalized regression model with x, y coordinates, elevation and slope predictors (R <sup>2</sup> =0.8, p<0.0001, RSE=0.65)
Number of dry days in the wet season	Calculated: Interpolation with regularized spline method (ESRI, <a href="http://www.esri.com/">http://www.esri.com/</a> , ArcGIS 9.1)

<p><b><u>Fire</u></b></p> <p>Fire frequency (number/yr)</p>	<p>KNP database; annual burnt areas (polygon file), 1955-2004</p> <p>Calculated: sum of annual burnt area layers divided by 50-year time period.</p>
<p><b><u>Herbivory</u></b></p> <p>Elephant biomass density (mean kg/km<sup>2</sup>)</p> <p>Browsers and mixed-feeders biomass density (mean kg/km<sup>2</sup>)</p> <p>Grazers and mixed-feeders biomass density (mean kg/km<sup>2</sup>)</p>	<p>KNP database; annual dry-season aerial census, GPS points of herbivore counts, 1981-2005</p> <p>Calculated: For each census point and for each year, species counts were scaled to biomass using elephant average weight. An inverse-distance density function (ESRI, ArcGIS 9.1) with a radius of 40 km was applied to distribute point observations and create continuous elephant biomass maps. The annual mean of biomass maps gave the final map.</p> <p>Calculated: For each census point and for each year, species counts were scaled to biomass using browser and mixed-feeder species-specific average weights. Half of the mixed-feeders biomass was kept with the browser group and half put with the grazer group. An inverse-distance density function with a radius of 20 km was applied to distribute point observations and create continuous browser biomass maps. The annual mean of biomass maps gave the final map.</p> <p>Calculated: same as browser and mixed-feeders map but with gazer and mixed-feeder species-specific average weights.</p>
<p><b><u>Geomorphology and Soils</u></b></p> <p>Elevation (m asl)</p>	<p>KNP database; vectorized from the Chief Directorate Survey's and Mapping (CDSM) 20m contour lines</p> <p>Statistics: min=88 m, max=1172.3 m, mean=376.1 m, sd =124.3 m.</p>

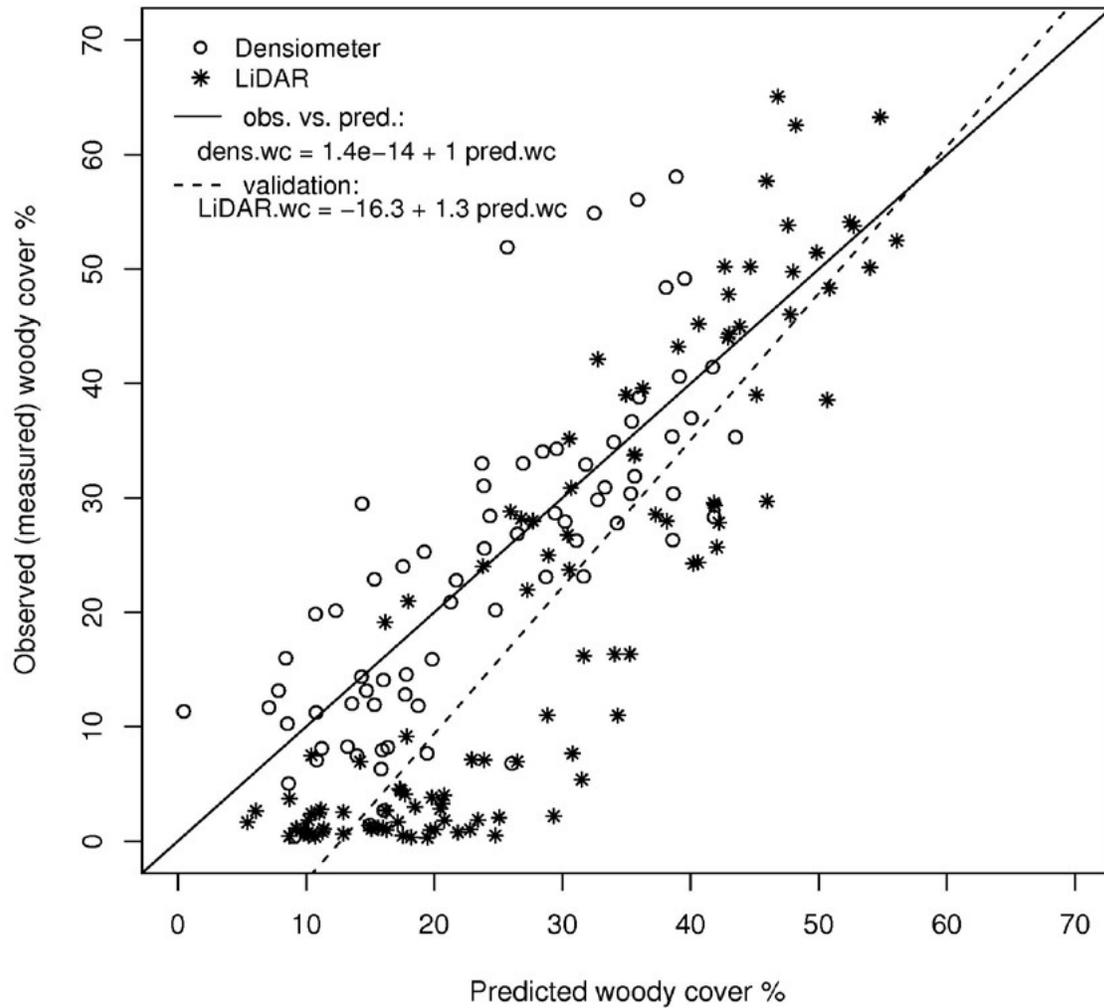
Slope (degree)	Calculated (ESRI, ArcGIS 9.1); Statistics:min=0°, max=59.8°, mean=2.1°, sd=3.2°.
Aspect (categorical)	Calculated (ESRI, ArcGIS 9.1): originally in degrees and then assigned to classes: north: (0°-44° and 315°-359°), east (45°-134°), south (135°-224°) and west (225°-314°).
Basalt/granite bedrocks	
Soil texture (g/kg)	Subset on KNP
Total soil nitrogen (categorical)	Subset on KNP
<b><u>Water</u></b>	KNP database; rivers and water points maps
Distance to water (m)	Calculated (ESRI, ArcGIS 9.1): Euclidean distance function
<b><u>Soil/Vegetation Complex</u></b>	Polygon vegetation map (Gertenbach, 1983) and polygon bedrock map (Venter, 1990)
soil-vegetation	Calculated: Vegetation map was reclassified into three classes: mopane ( <i>Colophospermum mopane</i> ), non-mopane species, sandveld communities. The intersection between the reclassified vegetation map and the bedrock map produced a five-class final map: mopane on basalt, mopane on granite, non-mopane on basalt, non-mopane on granite, sandveld communities on granite.

**Table 3-6.** Woody cover regression model (columns 2, 3, 4). Coefficient estimates, standard error, and p-value for the selected variables. The multiple  $R^2 = 0.61$ , residual standard error = 8.9 and  $p < 0.0001$ . Jack-knife estimates of the model coefficients for the predictors (columns 5, 6, 7): mean, standard error and bias (Efron, 1993).

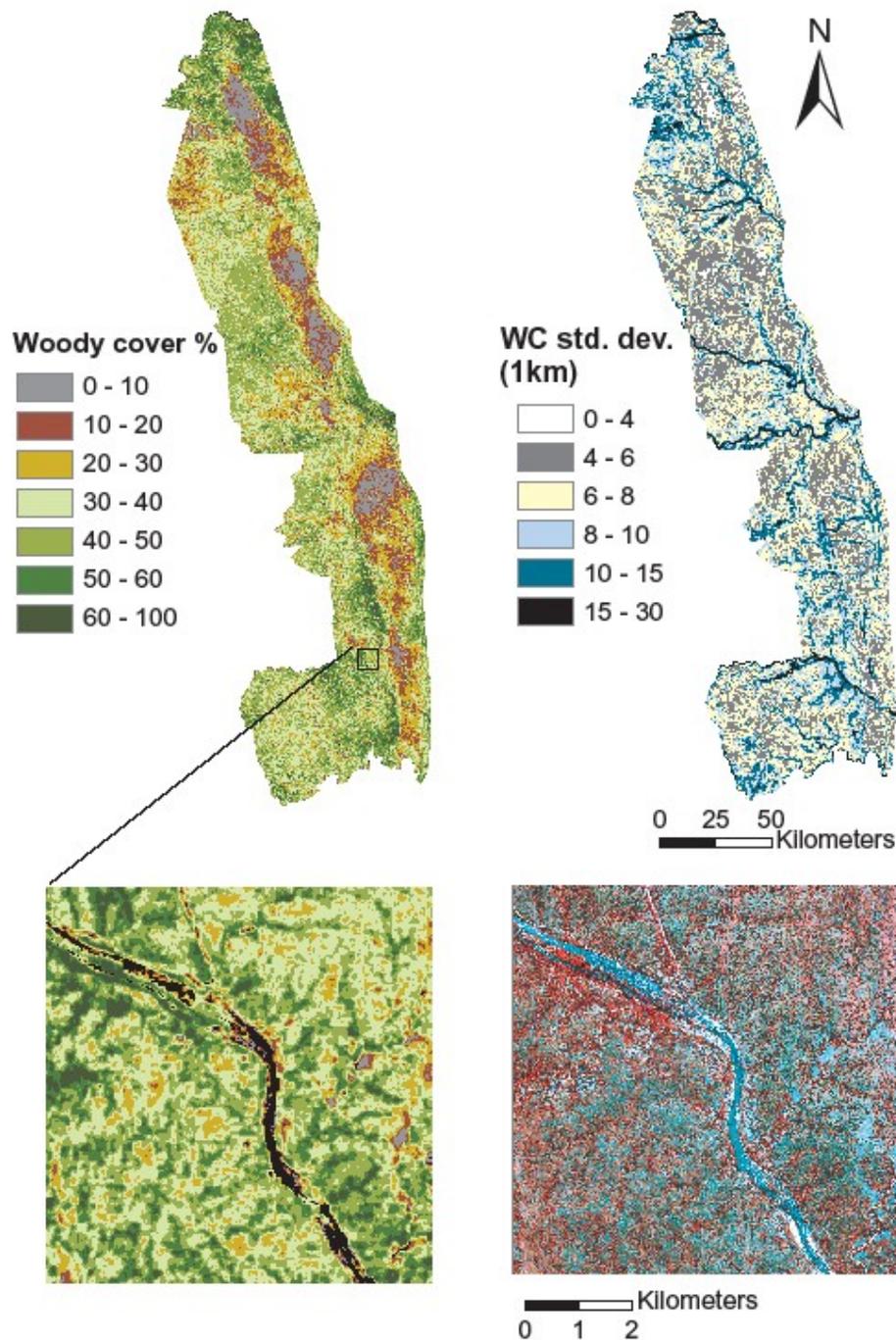
	Woody cover predictive model			Jack-knife estimates		
	Estimate	Std. Error	p-value	Mean	Std. Error	Bias
<i>Intercept</i>	2.97	39.7	0.9404	3.00	33.88	1.62
<b>JERS-1</b>	1.35	0.4	0.0003	1.36	0.29	0.03
<b>Landsat band2</b>	-6.55	1.9	0.0013	-6.55	1.71	0.06
<b>Landsat PCA3</b>	-5.39	2.4	0.0255	-5.39	2.08	0.09
<b>Landsat band5</b>	3.01	1.3	0.0222	3.02	1.11	-0.06
<b>SAVI</b>	322.21	141.5	0.0259	322.12	123.57	-6.54



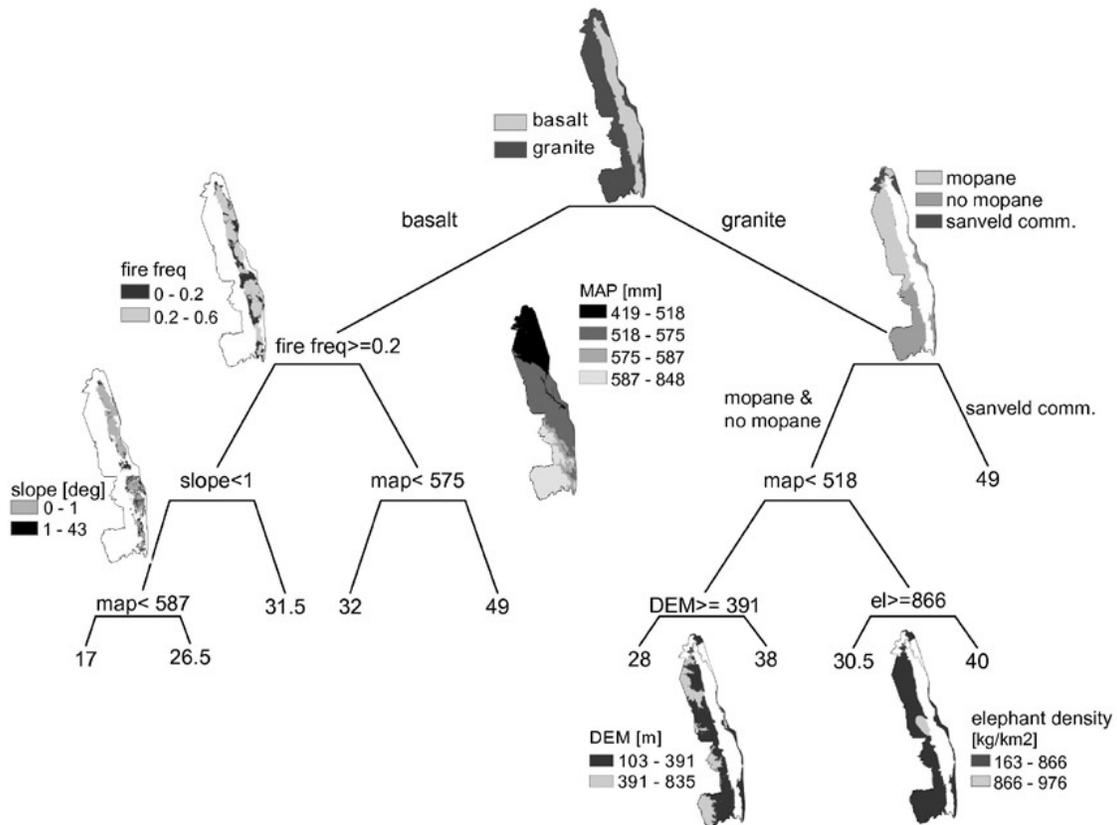
**Figure 3-10:** Study area: KNP and the distribution of the field plots across the MAP gradient.



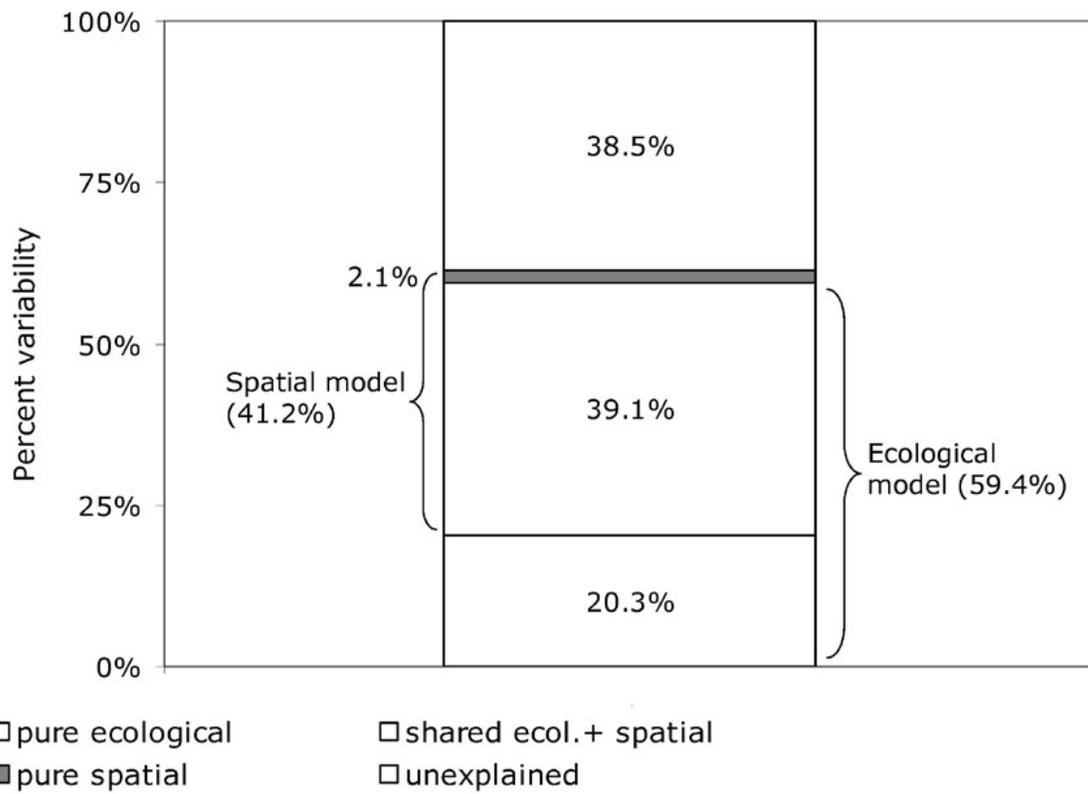
**Figure 3-11:** Model evaluation: observed (densiometer) versus predicted woody cover percent points (circles) and fitted line (solid,  $R^2 = 0.61$  and  $RSE = 8.6$ ,  $p < 0.0001$ ). Model validation: LiDAR estimates versus predicted woody cover percent points (stars) and fitted line (dashed,  $R^2 = 0.79$ ,  $RSE = 8.9$ ,  $p < 0.0001$ ).



**Figure 3-12:** Top left: KNP woody cover percent map (90-m nominal resolution). Top right: heterogeneity map for Kruger National Park with focal resolution of 1 km. Bottom: woody cover map zoom on Sabie river. Bottom left: false color infrared-composite of a 4-m resolution IKONOS image (4-28-2001) (photosynthetically active vegetation in red shades). Bottom right: woody cover percent estimates.



**Figure 3-13:** Ecological model (regression tree) with end node values being woody cover percent. Each split refers to the condition for the left child branch. The values at the end nodes are the mean response for that stratum. Node variables: basalt/granite bedrocks, fire frequency (fire freq), elevation (DEM), slope, mean annual precipitation (map), mopane (*Colophospermum Mopane* species), sanveld community species (sanveld comm.) and elephant biomass density (el).



**Figure 3-14:** Variation partitioning for the CART models. Variability explained by the ecological-spatial model:  $61.5 = 39.1 + 20.3 + 2.1$  %. Unexplained variability: 38.5 %.

## **CHAPTER 4**

### **Woody cover mapping in Africa: an approach with combined optical and radar remote sensing**

**Authors: Gabriela Bucini<sup>1</sup>, Niall P. Hanan<sup>1</sup>, Sassan Saatchi<sup>2</sup>, Michael A. Lefsky<sup>3</sup>,  
Edward Mitchard<sup>4</sup> and Leon-Jacques Theron<sup>5</sup>**

1. Natural Resource Ecology Laboratory, Colorado State University, Fort Collins, CO 80523-1499, U.S.A.

2. Jet Propulsion Laboratory, California Institute of Technology, Mail Stop 300-319 Pasadena, CA 91109, U.S.A.

3. Department of Forest, Rangeland and Watershed Stewardship, Colorado State University, Forestry Building, Fort Collins, Colorado 80523, U.S.A.

4. School of Geosciences, University of Edinburgh, Drummond St, EH8 9XP UK.

5. The Cirrus Group, A Climate Change and Natural Resources Practice, Cape Town, South Africa.

## 4.1 Introduction

The extent, the spatial arrangement and the structure of woody vegetation influence important ecosystem processes including carbon storage, respiration, energy transformations, biogeochemical and biogeophysical cycles (Belsky et al. 1993, Hanan 2001, Baldocchi et al. 2004, Ratnam et al. 2008, Williams et al. 2008). For the African continent in particular, there is a strong need for further research on carbon dynamics, especially in savanna systems that appear to be highly sensitive to climate variations (Williams et al. 2007, Williams et al. 2008, Scheiter and Higgins 2009). Woody plants regulate runoff and erosion (Mougin et al. 2009) and are of considerable significance for biodiversity as their structural characteristics influence flora (Ludwig et al. 2008) and fauna (du Toit 2003, Owen-Smith 2004) survival. Humans also depend on woody plants to obtain fruits, firewood, construction material and medicinal products. In the African continent, especially in rural and poorer urban households, fuel wood is a major resource for cooking, lighting and heating (Mahiri and Howorth 2001, Arthur et al. 2010). In contrast, woody encroachment and thickening in grasslands and savannas has long been a concern because it decreases herbaceous production and hence negatively impacts livestock-dependent livelihoods (Shackleton and Gambiza 2008, Wigley et al. 2009). The importance of information on woody cover to understand ecosystem functioning has grown in the recent decades also in response to scientific questions related to global earth-atmosphere dynamics, pressing climate change issues, biodiversity and natural resource management. The scientific community has since then expanded its spatial scopes and aimed at integrating plot- with large-scale data and models. Technological

advances have been pivotal in this process with remote-sensing instruments and image processing techniques providing multi-scale data for earth observing and monitoring.

In this work, we develop an approach that quantitatively relates woody plant radiative and structural properties to create a woody cover map. Over the last two decades, the large-scale mapping approach has mainly changed in three ways: (i) with a transition from discrete vegetation categories to continuous fields (DeFries et al. 1999); (ii) with the use of multi-temporal spectral data (Hansen et al. 2000, Hansen et al. 2002) and (iii) with the inclusion of more detailed woody cover information for model calibration. The continuous fields approach moved away from rigid class boundaries and assigned a continuous value of fractional cover to each pixel bringing to light gradients and heterogeneous spatial patterns in the landscape (DeFries et al. 1999).

The use of multi-temporal data takes advantage of phenological information for the mapping process. In general, different life-forms (annual and perennial herbaceous vegetation and deciduous and evergreen woody vegetation) exhibit different temporal behaviors in their seasonality (green-up, maturity and senescence) that can be exploited to extract vegetation attributes from remote-sensing imagery. Annual metrics derived from multi-temporal data such as annual mean, maximum, minimum and range of reflectance and/or vegetation indices provide quantities with possible biophysical significance (DeFries et al. 1995, DeFries et al. 1997, Hansen et al. 2002). These metrics are detached from a specific date and hence are useful for analyses over large areas where phenological cycles have different timings or there is opposite hemispherical seasonality. When derived from optical spectral bands, these metrics are related to vegetation photosynthetic activity and they indicate either the phase or changes in vegetation

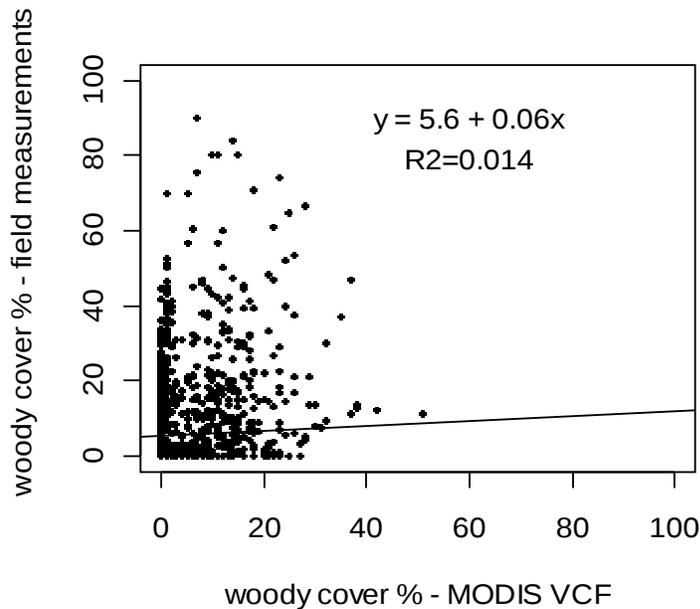
greenness over the year. In the case of dense evergreen forests, the phenological patterns are relatively simple with an average high greenness maintained throughout the year (Xiao et al. 2005, Prasad et al. 2007). Drought seasonal grasslands, on the other hand, can be successfully distinguished by relatively high greenness ranges (Butterfield and Malmstrom 2009) originating from the temporally limited season with green standing grass. The application of annual metrics for savanna or woodland characterizations is instead more complicated because of the phenological complexity arising from their mixture of herbaceous, deciduous and semi-deciduous life-forms (Scanlon et al. 2002, Do et al. 2005, Archibald and Scholes 2007). Results from Asner et al. (1998) in a South Texas savanna show that areas with mixed woody and herbaceous vegetation have lower Fraction of Absorbed Photosynthetically Active Radiation (fAPAR) than dense woodland but higher than areas dominated by grass. This ecological information combined with multi-temporal remote-sensing data is useful for identifying savannas among others ecosystems. However, the ability to quantitatively separate the woody vegetation contribution to the pixel reflectance remains poor because of low contrast between green woody and herbaceous reflectance.

The capacity of all-weather condition Synthetic Aperture Radar (SAR) systems is advantageous for vegetation remote-sensing (Waring, 1995) and mapping purposes (Santos et al. 2002, Lucas et al. 2006, Lucas and Armston 2007, Lucas et al. 2007, Saatchi et al. 2007, Viergever et al. 2007, Ribeiro et al. 2008, Mitchard et al. 2009). Radar waves not only can penetrate clouds but also the vegetation canopy (Waring et al. 1995). In general, the radar wave interacts more strongly with objects that have a size comparable to the IR wavelength. Differently from optical sensors, radars are sensitive to

the woody (nonphotosynthetic) components of vegetation. Shorter wavelengths (K-, X-, C-band) are more sensitive to the elements of the canopy surface (leaves and small branches), leaf water content and moisture (Waring et al. 1995, Kasischke et al. 1997). In particular, shrub-grass separation can be problematic in the optical domain (DeFries et al. 1999), but it could become possible in the radar domain since shrubs and grasses have different canopy structures and/or water contents. Important factors that may contaminate the vegetation backscatter are topological features, soil roughness and soil moisture (van Zyl 1993). Conceptually, if we can combine the information contained in the radiative responses of both optical and radar systems, the approach to woody cover mapping can result in improved predictions with respect to approaches based on one system alone. It is therefore very important to examine the benefits of combining the different and complementary data from these two systems.

The most current woody cover map products that include the whole African continent are two: the Vegetation Continuous Fields (VCF, Hansen et al. 2003) and the tree cover map developed by Rokhmatuloh et al. (2005). Both studies derived the percent tree cover from optical features. The training data for the VCF product are built on 250 Landsat land cover classifications (De Fries et al. 1998, Hansen et al. 2002) that were then aggregated to four cover strata (0%, 25%, 50% and 80%). Averaging the cover strata to MODIS 500-m cells provided a continuous cover data set. However, small trees and shrubs were eliminated from the field calibration set. That is, vegetation classes including bushland and shrubland with an approximate height of mature woody vegetation  $< 5$  m were included in the zero tree stratum even if the estimated woody cover was  $> 40$  %. The consequence is a strong underestimation of woody cover in savanna ecosystems

(Figure 4-15) that comprise almost half of the African land surface. Rokhmatuloh et al. (2005) potentially improved the woody cover estimates for savannas by classifying 11 very high-resolution (0.6 m) pan-sharpened IKONOS images. The fraction of high-resolution pixels classified as “tree” provided a continuous woody cover value in a MODIS 1-km resolution cell. The authors stratified their sampling by selecting one IKONOS image for each class of the Global Land Cover 2000 map (GLC2000, Bartholome and Belward 2005). This approach provides a limited representation of the within-class cover variability across the continent. The positive side of these two studies is that the training datasets had a very large number of points (pixels) because they were generated from remote-sensing images.



**Figure 4-15:** Field measured (Sankaran et al. 2005) vs. 500-m MODIS VCF (Hansen et al. 2003) woody cover percent in 850 sites across African savannas: scatter plot with regression line.

In this work, we aim at providing a more complete assessment of woody cover for Africa giving particular attention to include shrubs and small trees that represent a large component of woody vegetation in savannas regions. We believe that we can attain this by including two particular elements in our analysis: (i) our training data are direct measurements of woody cover comprising shrubs, small and large trees and (ii) we combine microwave and optical remote-sensing to include data more sensitive to woody material that we wish to detect. The rationale for our approach builds on earlier research. Bucini et al. (2010) combined optical and radar imagery to map woody cover at the regional scale of Kruger National Park (KNP, South Africa) using field measurements. This savanna area spans arid and semi-arid climate conditions with vegetation composed by shrubs and trees in diverse proportions. Our woody cover model explained 61% of the variability and its predictions were comparable with LiDAR cover estimates. The present work is an extension to the African continent. Other studies in Australia, Amazon, Belize and Mozambique (Lucas et al. 2000, Saatchi et al. 2007, Viergever et al. 2007, Ribeiro et al. 2008) showed the benefits of combined radar and optical remote-sensing for biomass and vegetation structure mapping at large scales. Mitchard et al. (2009) showed that the radar sensitivity to woody structures creates a consistent baseline to detect woody biomass across African savannas and woodlands. While recognizing that different radar wavelengths interact with different vegetation elements and depths, studies show consensus on microwave sensitivity to woody biomass (Lucas et al. 2004). We hypothesize that this property can be exploited to create direct relationships between woody cover and microwave backscatter over continental gradients. Specifically, our research questions are the following:

1. What are the relationships between woody cover and remote sensing data and which remote-sensing indicators are the most suitable and important to predict woody cover?
2. What are the characteristics of these relationships: do overall continental relationships exist or is it useful/necessary to introduce regional data stratifications?
3. Does the combination of microwave and optical data increase the predictive power of ordinary least squares models compared to models based only on optical or microwave data?

We predicted woody cover using an empirical modeling approach built on a sample of direct woody cover measurements. We sought a direct relationship to tie woody cover to spectral features in the optical and shortwave spectra. More specifically, we made use of MODIS, Quick-Scat and SRTM satellite products to derive spectral metrics. We tested different ways to model (calibrate) remote-sensing data and validated them against an independent woody cover validation dataset. We estimated the error components originating from the training data and from the regression model.

## **4.2 Materials and Methods**

### ***4.2.1 Woody cover training data (response variable)***

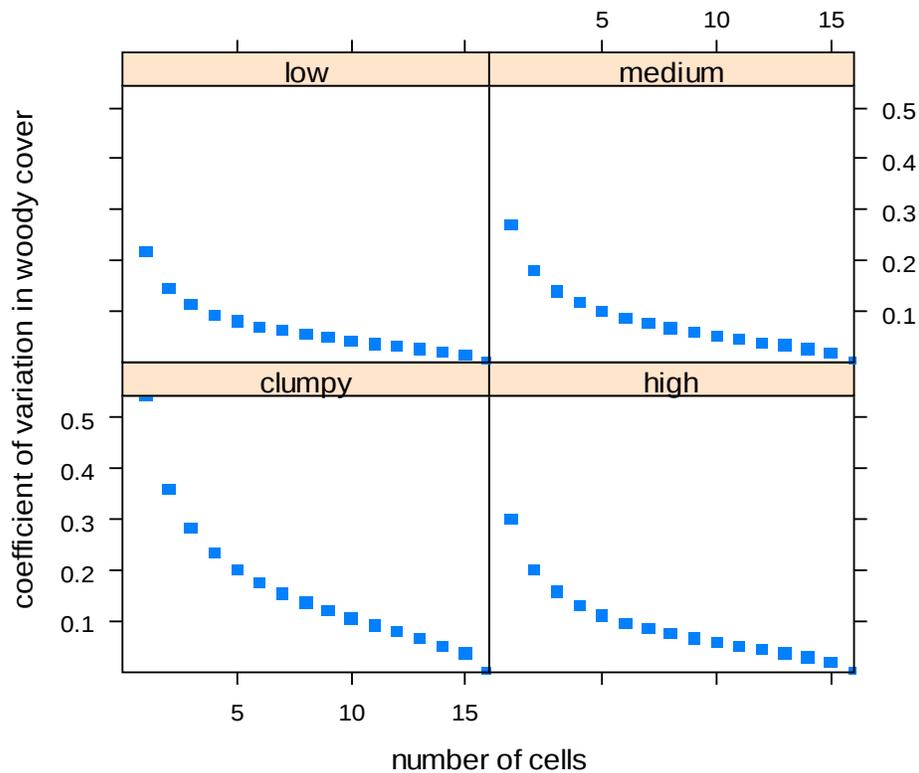
The woody cover training data play a critical role because they provide data for model calibration. They should represent the potential variability that can be captured by the modeling process. Our sampling strategy aimed at building a representative dataset of woody cover (%) across the range found across African ecosystems. In collecting our

data, we took into account the observable variation in cover as well as in environmental and climatic conditions that could affect remote-sensing measurements. We worked with a dataset of 982 woody cover assessments from two different sources, field measurements and high-resolution remote-sensing image estimates.

Most of our field measurements (850) were compiled in 2000-2003 at sites across African savannas (see Sankaran et al. (2005) for a complete description of the dataset). These data provide a unique source of information over a range of woody cover between 0-90%, with 75% of the points between 0-20%. A drawback of this dataset is the relatively small plot size (~ 1 ha) that might cause scale incompatibility with the 1x1 km remote-sensing cells. However, while the scale mismatch may increase variability in the calibration data relative to the remote-sensing data, the large sample size means that it should not contribute any bias. Densimeter measurements from three field surveys were also included: (i) 73 data points from a 2006 campaign conducted by Bucini et al. (2010) in Kruger National Park, South Africa (see also Chapter 3), (ii) 12 data from a 2008 campaign conducted by Mitchard E. in 11 transects (20m x 200m) and one 2-ha plot in the Budongo Forest, Uganda, and (iii) 49 data points from a 2009 campaign conducted by Theron L.-J. 250x250 m plot in Zambia. We averaged the values from plots falling in the same remote-sensing 1-km pixel and obtained 803 points.

We also used high-resolution imagery available in Google-Earth® (GE, <http://earth.google.com>) to estimate woody cover in non-savanna biomes. A total of 173 sample sites were identified within forest and seasonal woodland ecosystems (Miombo woodlands, dry and wet forests), deserts, water bodies, urban and agricultural areas across Africa. The high-resolution images consist of IKONOS and Quickbird (Digital

Globe, <http://worldview2.digitalglobe.com/>) true color composites (blue, green, red bands), with resolutions between 0.6 and 4 m. The Satellite Imaging Corporation ([http://www.satimagingcorp.com/google\\_earth.html](http://www.satimagingcorp.com/google_earth.html)) showed that the actual image resolution on GE might not match the detail and quality of the original images. Regardless, the GE high-resolution imagery was appropriate to resolve woody plant crowns in forested areas. The selected GE images dated from 2000 to 2007 with the majority of them being close to 2000-2002, to minimize temporal discrepancies between sample and large-scale remote-sensing data. For each sample site, we built a 1x1 km georectified grid in ArcGIS 9.1 (Environmental Systems Research Institute, CA, USA) representing a remote-sensing pixel. Each grid was composed of 16 250x250 m cells and within each cell we deployed an 8x8 point sub-grid. We estimated fractional cover based on presence and absence of woody plant crowns (pinpoint assessment) at each of the 8x8 digital points of a cell. While conducting the pinpoint assessments, we also recorded the number of points with uncertainty in crown identification. The error estimated from this uncertainty amounted to an average 3.7 % cover across all the high-resolution sites. The woody cover percent for the entire 1x1 km pixel was calculated averaging cover values from six random cells in the grid. This criterion was decided after conducting a sensitivity analysis on the number of grid cells assessments needed to obtain a cover value with a coefficient of variation (CV) < 0.2, accounting for all possible cell combinations (Figure 4-16). The sensitivity analysis was run in four separate sites characterized by low, medium, high and clumped woody cover and we found that with 6 cells, we met the requisite in all four cover types.



**Figure 4-16:** Coefficient of variation for woody cover vs. number of grid cells ( $n = 1, \dots, 16$ ). For each fixed number  $n$  of cells, the coefficient of variations is calculated from the set of all possible combinations (the group of 16 cells taken  $n$  at a time).

#### 4.2.2 Remote-sensing data (predictive variables)

We compiled a set of remote-sensing data from optical and microwave sensors. We derived metrics that could potentially be linked to physiological and structural attributes of woody vegetation (Table 4-7). To quantify greenness patterns, we used optical multispectral data derived from NASA MODIS products. We used the 32-day MODIS composite reflectances over five years (2000 – 2004) to develop our average

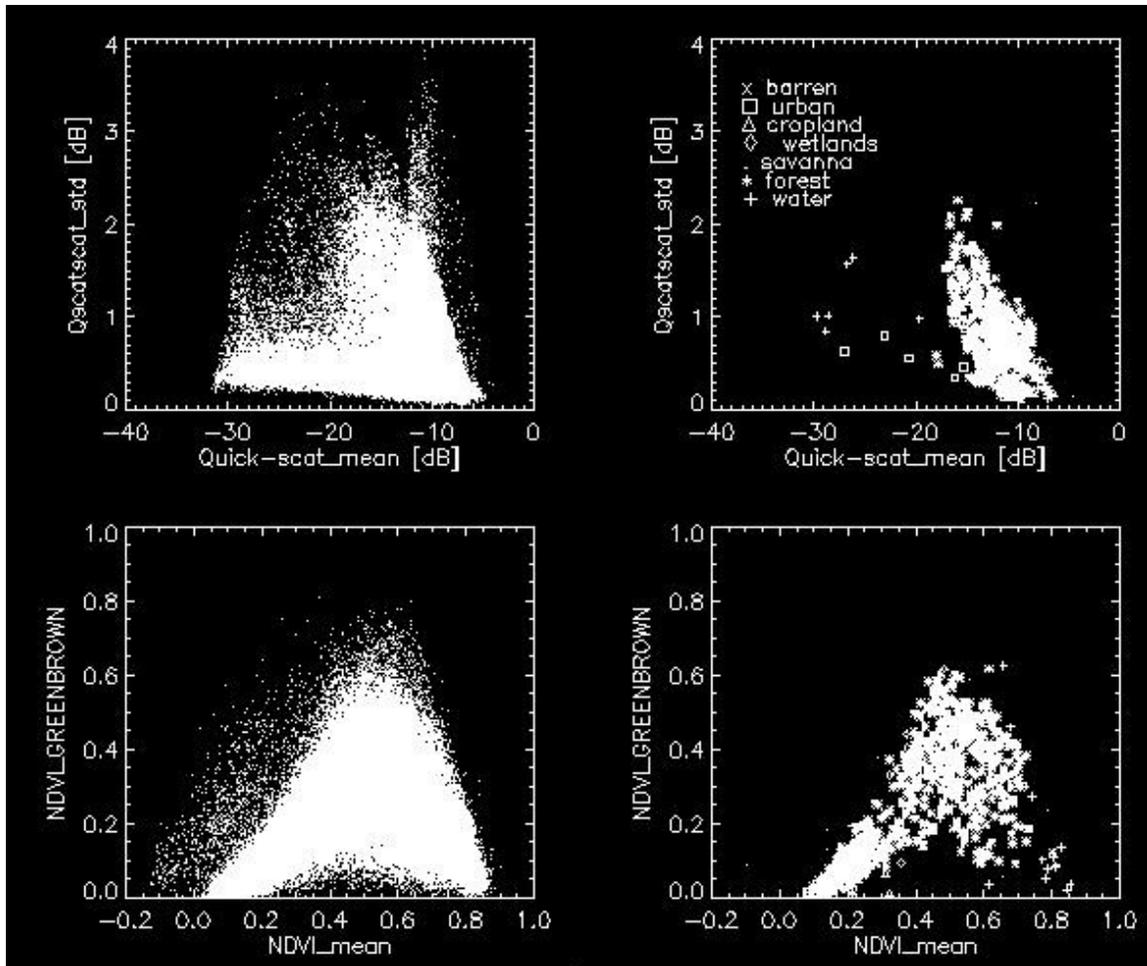
spectral metrics. The processes of compositing and averaging over multiple years improved the data quality by minimizing the cloud and noise contamination in the reflectance values. From this process, we developed a one-year averaged monthly reflectance data at 1-km resolution and subsequently generated the normalized difference vegetation index (NDVI), and phenological metrics of mean, maximum, green (average of the greenest quarter related to maximum NDVI) and green\_brown (the difference between NDVI of maximum and minimum quarters) NDVI.

Imagery from Q-SCAT (Quick Scatterometer) available in three-day composite at 2.25 km resolution over 5 years (2000-2004) were processed into average monthly composite and resampled (rebinned) at 1-km resolution. Q-SCAT is an active microwave sensor operating in the Ku band (12 GHz frequency,  $\sim 2$  cm wavelength) at both HH and VV polarizations and it is sensitive to canopy roughness, moisture and leaf water content (Saatchi et al. 2007, Saatchi et al. 2008). From these data, we produced two metrics, the averaged annual mean and standard deviation in the HH polarization (assuming a high correlation between HH and VV polarizations over vegetation). In ecosystems with low woody plant density, these metrics are correlated with aboveground biomass (Long et al., 2001, Saatchi et al. 2007, 2008). Given Q-SCAT short wavelength, the backscatter measurements should be sensitive to small branches and leaves and therefore to the presence of shrubs and small trees. However, backscatter from the herbaceous background may also be included in the signal (Hill M. 2005, Lucas 2006). By taking the backscatter annual mean and standard deviation, we hoped to detect the consistent signal from woody vegetation and roughness variations due to structural differences between grass and branches. In forested areas with tall trees and large canopies, the Q-SCAT

metrics respond to top of the canopy roughness and moisture and should be directly related to the canopy size and hence cover. In summary, we created six remote-sensing layers at 1 km resolution in sinusoidal projection.

<b>Spectral variable</b>	<b>Remote-sensing instrument</b>	<b>Biophysical variable</b>	<b>Derived metrics</b>
Monthly NDVI (2000-2004)	MODIS	Photosynthetic activity	NDVI_mean: annual mean NDVI NDVI_max: annual maximum NDVI NDVI_green: average of the greenest quarter related to maximum NDVI NDVI_green_brown: difference between NDVI of maximum and minimum quarters
Radar backscatter (HH) monthly mean (2000-2004)	Quick-Scatterometer	Canopy roughness and moisture, wood density	QSCAT_mean: annual mean (HH Polarization) QSCAT_std: annual standard deviation (HH Polarization)

**Table 4-7:** Remote sensing layers and metrics from MODIS and Q-SCAT (Quick Scatterometer) sensors.



**Figure 4-17:** Spectral spaces: 2-D spectral scatter plots of the remote-sensing variables used in this study. Each row is related to a different remote-sensing system: QuickSCAT (radar) and MODIS (optical). On the left: spectral space for the remote-sensing data for the continent. On the right: spectral space for the training dataset. The training data are symbol-coded by land cover class from the MODIS land cover product with the IGBP legend (<http://www-modis.bu.edu/landcover/index.html>). Our training data have a good representative range on the spectral ranges except for areas with very low Q-SCAT backscatter (Quick-scat\_mean < - 1500) mainly corresponding to deserts and water bodies.

### 4.2.3 Modeling approach

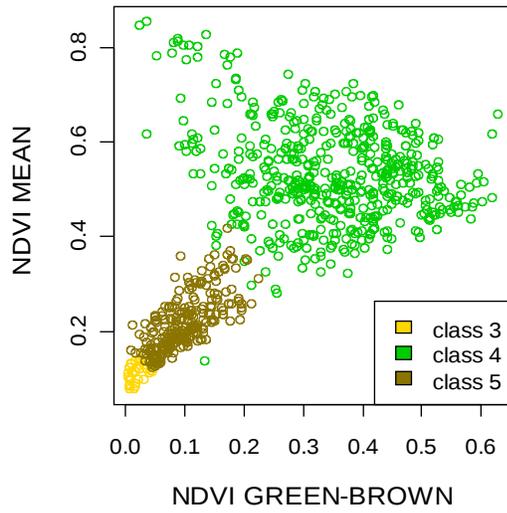
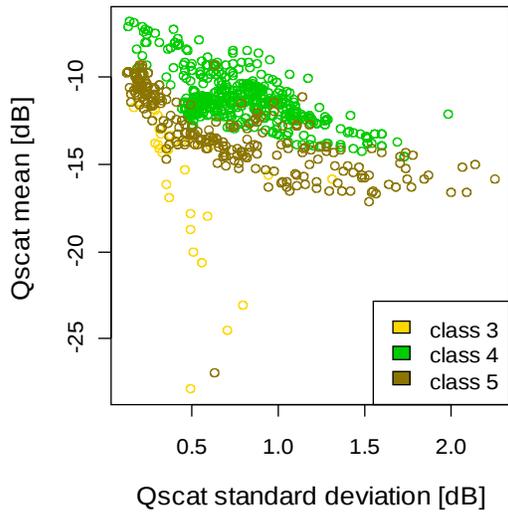
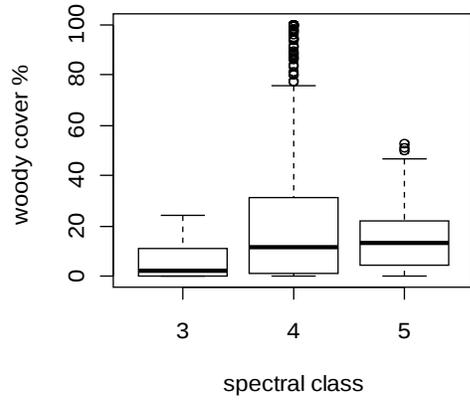
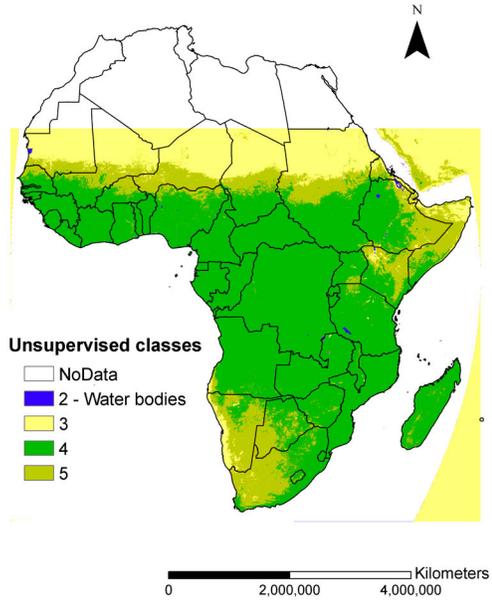
The ultimate goal of our analysis is to find a predictive model of woody cover using our training sample and to extrapolate it over the whole African continent. In our modeling approach, we sought quantitative relationships that predict cover as a function of remote-sensing spectral features (Figure 4-17) and for this purpose we used multiple linear regression analysis. We accounted for non-linearity by applying transformations to the predictor variables and including them in a model selection procedure using the Bayesian information criterion (BIC, Schwarz 1978). BIC favors parsimonious models by penalizing both for the number of model parameters and number of observations. The optimal model, contingent to the data at hand, is the one with the lowest BIC value. We did not include points (20) from the classes “water”, “city” and “wetland” because they are not representative of the main land covers and their spectral position is distant from the rest of the data such as to affect the regression parameters.

In structuring our models, we also investigated the practical benefit of including a class variable (categorical variable) to stratify the data. The stratification was based on an initial ISODATA unsupervised classification (*ENVI ITT Visual Information Solutions* <http://www.itvis.com>) of the spectral variables (Figure 4-18). By including the resulting classes in the predictor set, we sought to test whether there exist significantly different spectral behaviors at continental scales. Two sets of models are hence considered in the analysis:

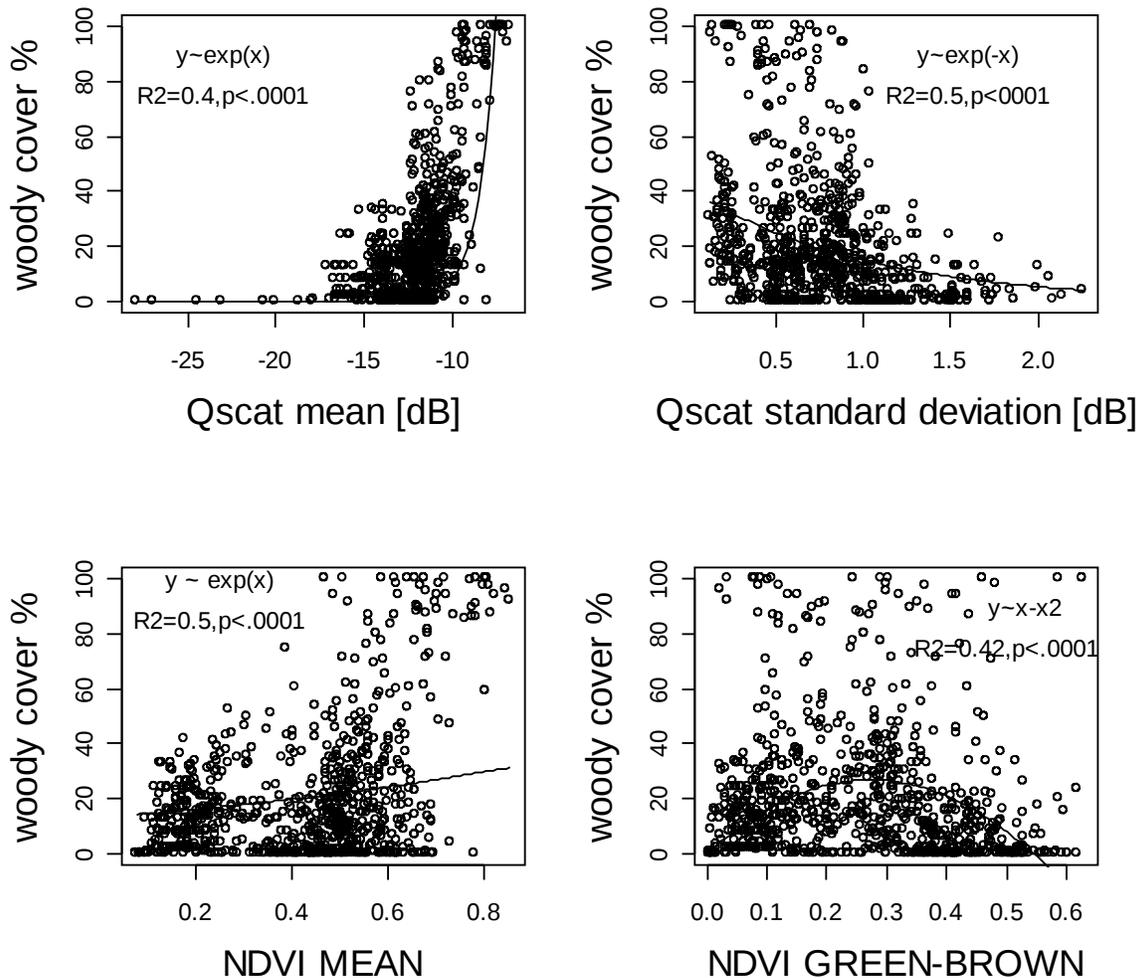
1. Dependent on spectral variables only (no stratification)
2. Dependent on the set of spectral variables and a class variable (stratified). This stratification reflects a natural spectral clustering of our remote-sensing

information for Africa and it is independent from any bias based on vegetation cover type.

Before running the regressions, we discarded NDVI\_max and NDVI\_green variables because of high collinearity with NDVI\_mean (Pearson correlation coefficient  $> 0.9$ ) and we kept all the other variables, which had Pearson correlation coefficients  $< 0.7$ . We examined the trends of woody cover vs. each predictor (Figure 4-19) to check for non-linearity and selected the appropriate non-linear transformations. We also rescaled the Q-SCAT variable between  $(-10, 0)$  to have more comparable input ranges for the models. The full set of predictors (12) is reported in Table 4-8. We split our sample points into a training (90 %, Figure 4-20) and a validation (10%) sets. After selecting an optimal model, we tested it on the validation dataset and created a woody cover map that we compared to various other information sources of woody cover at different spatial scales.



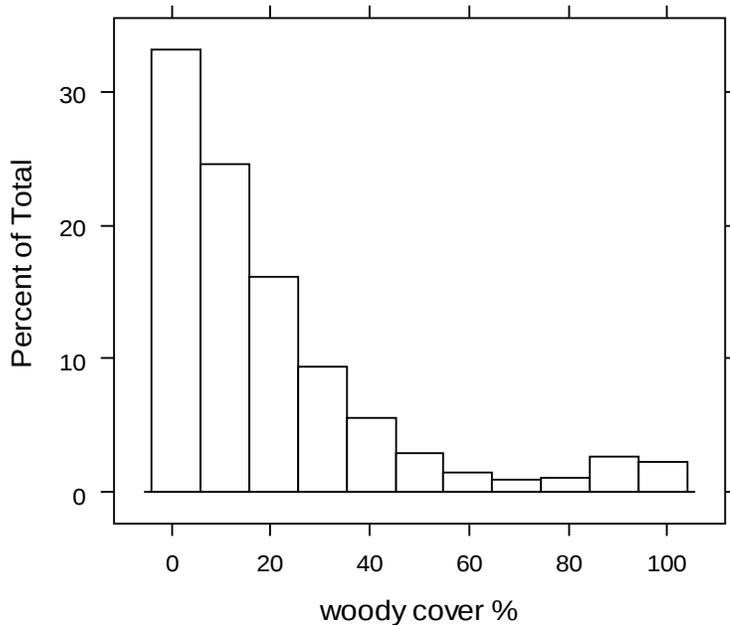
**Figure 4-18:** Upper left: ISODATA unsupervised classification (4 classes) derived from the remote-sensing layers. There are affinities with known large-scale vegetation patterns. The very arid areas (class 3) were separated from the grasslands/shrublands (class 5) characterized by high annual NDVI variability and low woody biomass. Class 4 appears to include taller savannas, seasonal and moist tropical forests. Upper right: box plot of woody cover in the three unsupervised classes, excluding water. Lower left: unsupervised classes in the radar spectral space. Lower right: unsupervised classes in the optical spectral space.



**Figure 4-19:** Scatter plots of woody cover (training data) vs. the four spectral metrics selected as predictors for the modeling analysis (the Pearson correlation coefficients between NDVI\_green, NDVI\_max and NDVI\_mean were  $> 0.9$  and NDVI\_mean had the highest correlation with woody cover therefore we only selected NDVI\_mean for the analysis). The graphs report the best-fit univariate models (coefficients,  $R^2$  and p-value).

Sprectral predictor	Variable	Transformation	Transformed variable
QuickSCAT mean	Qscat_mean	$\exp(x)$	exp_qscatmean
QuickSCAT standard deviation	Qscat_std	$\exp(-x)$	exp_qscatstd
NDVI MEAN	NDVI_mean	$\exp(x)$	exp_ndvimean
NDVI GREEN-BROWN	NDVI_green_brown	$-x^2$	sqr_ndvigrbr
Stratification classes (categorical variable)	class		

**Table 4-8:** The predictive variable set: continuous variables including linear and non-linear (transformed) terms and categorical variable (stratification classes).



**Figure 4-20:** Histogram of Woody cover percent for the training dataset (832 points).

### 4.3 Results

We fitted woody cover percent data (832 points) with models that depended on all possible combinations of spectral metrics with and without data stratification. Each model was ranked according to its BIC value. Table 4-9 reports the first five ranked models (the other models had essentially no empirical support according to BIC). The weight of evidence  $w$  for the selected best model (1) relative to model 2 and 3 is weak. This suggests that the model selection uncertainty is likely to be high and the best model could vary depending on the sample (Burnham and Anderson 2002). When we examined other performance parameters, we noticed that the model 1 has comparable  $R^2$ , residual standard error with respect to the other two models. For the three models, residuals are near-normally distributed with ranges between  $-66$  and  $66$  % cover, median  $\sim -1.5$  % and

1<sup>st</sup>-3<sup>rd</sup> quartile ranges between -9 and 7 %. On the independent validation dataset, the regressions between observed and predicted woody cover values show that these three models perform similarly (Table 4-9).

MODEL	STRATIF.	PREDICTIVE MODEL					VALIDATION			
		R <sup>2</sup>	RSE	N	ΔBIC	w	R <sup>2</sup>	RSE	slope	intercept
1	YES	0.589	15.13	8	0	0.417	0.666	19	0.96	-0.8
2	YES	0.592	15.08	9	1	0.264	0.659	19.2	0.96	-0.9
3	YES	0.595	15.04	10	2	0.190	0.655	19.3	0.96	-0.8
4	YES	0.591	15.11	9	4	0.056	0.665	19.1	0.96	-0.8
5	YES	0.590	15.12	9	5	0.029	0.666	19	0.6	-0.8

**Table 4-9:** The first five ranked models. For each predictive model, the table reports R<sup>2</sup>, Residual Standard Error (RSE), the number of model parameters N, ΔBIC and the model probabilities w (left side). Right side: R<sup>2</sup>, RSE, slope and intercept of the regression line between observed vs. predicted cover on the validation dataset.

We selected model 1 (Table 4-10) and refer to it as the continental model. The significance of the stratification (class variable) is related to: (1) separating the woody cover behavior in more arid systems (classes 3 and 5 intercept) and (2) achieving a strong positive exponential relationship between cover and the Q-SCAT annual backscatter mean in class 4. The relationship with exp\_qscatmean is positive in all the classes. Woody cover is also positively correlated with NDVI\_mean as expected. Model 1 includes the linear term of NDVI\_green\_brown that might capture the lower temporal

variability of greenness in dense evergreen forests. A bootstrap analysis (Table 4-11) showed that the model is stable as the parameter estimates had relatively small bias and standard error. The LMG metric (Gromping 2007) is derived from decompositions of the model  $R^2$  and provides a measure of variable relative importance scaled between 0 and 1. The metric ranked Qscat\_mean as the most important variable to predict woody cover (LMG=0.59) followed by NDVI\_mean (LMG=0.23) and NDVI\_green\_brown (LMG=0.18). Model 2 contains the additional quadratic term of NDVI\_green\_brown describing an unexpected positive parabola. Model 3 contains the NDVI\_green\_brown quadratic term and the QuickSCAT standard deviation.

Predictor	Model parameter	Std. error	p-value	Variable Importance (LMG)
Intercept:				
class 3	-5.9	9.2	0.51	
class 4	-92.6	10.9	< 0.0001 ***	
class 5	-19	10.3	0.2	
exp_qscatmean:				0.59
class 3	44.5	39.1	0.2	
class 4	357	42.9	< 0.0001 ***	
class 5	117.8	42.7	0.09 .	
NDVI_MEAN	27.7	7.2	0.0001 ***	0.23
NDVI_GREEN_BROWN	-42.7	5.8	< 0.0001 ***	0.18

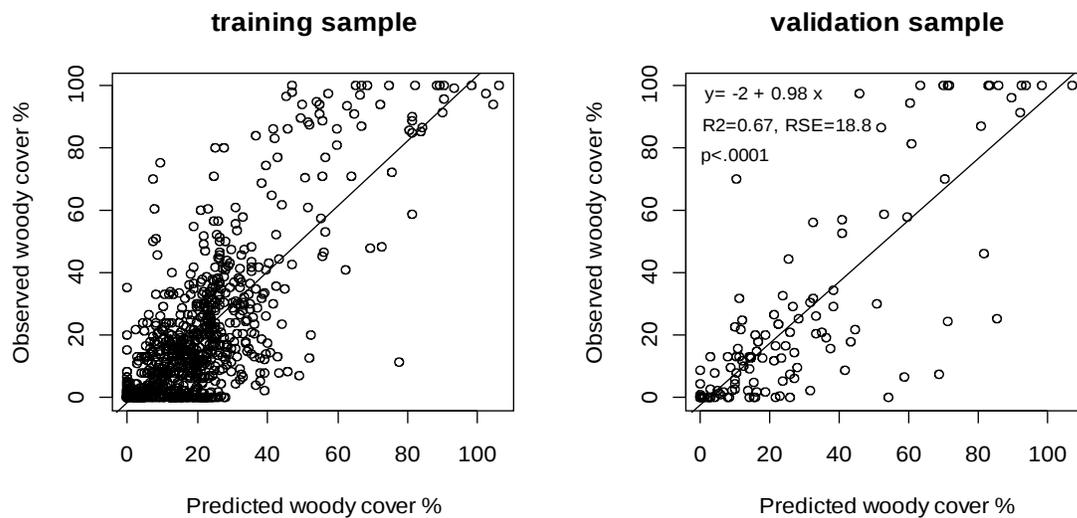
**Table 4-10:** Summary statistics for the parameter estimates of the selected continental model: standard deviation, p-value and LMG metric for variable relative importance. Parameters for the main effects (exp\_qscatmean, NDVI\_MEAN, NDVI\_GREEN\_BROWN), the 3 categorical classes and for the interactions between the classes and exp\_qscatmean. Significance levels are also coded as:  $p < 0.0001 = '***'$ ,  $< 0.001 = '**'$ ,  $< 0.01 = '*'$ ,  $< 0.05 = '.'$ .

Predictor	Original model parameter	Bias	Std error
Intercept:			
class 3	-5.9	0.09	3.2
class 4	-92.6	0.3	8.8
class 5	-19	-0.11	4.4
exp_qscatmean:			
class 3	44.5	-0.42	14.3
class 4	357	-0.71	29.2
class 5	117.8	0.47	18.5
NDVI_MEAN	27.7	0.14	9
NDVI_GREEN_BROWN	-42.7	-0.34	7.5

**Table 4-11** Bootstrap statistics for the parameter estimates of the selected continental model on 10000 bootstrap replications: original value, bias and standard error. Average  $R^2=0.589$ .

The model validation graph (Figure 4-21) shows a tendency to overestimate in the low cover range possibly because of soil and/or grass contamination. Points in middle cover range exhibit the highest scatter. We believe that this is due to the relatively little availability of training data for this range. High woody cover points show instead some

underestimation. In our modeling analysis, we eliminated one outlier. It was collected on Google Earth in a region of forest-savanna mosaic with spatial heterogeneity comparable to or finer than the remote-sensing data pixel resolution. Coregistration errors probably caused the association of this low cover site to a neighbor remote-sensing pixel carrying the spectral information of a densely wooded patch.



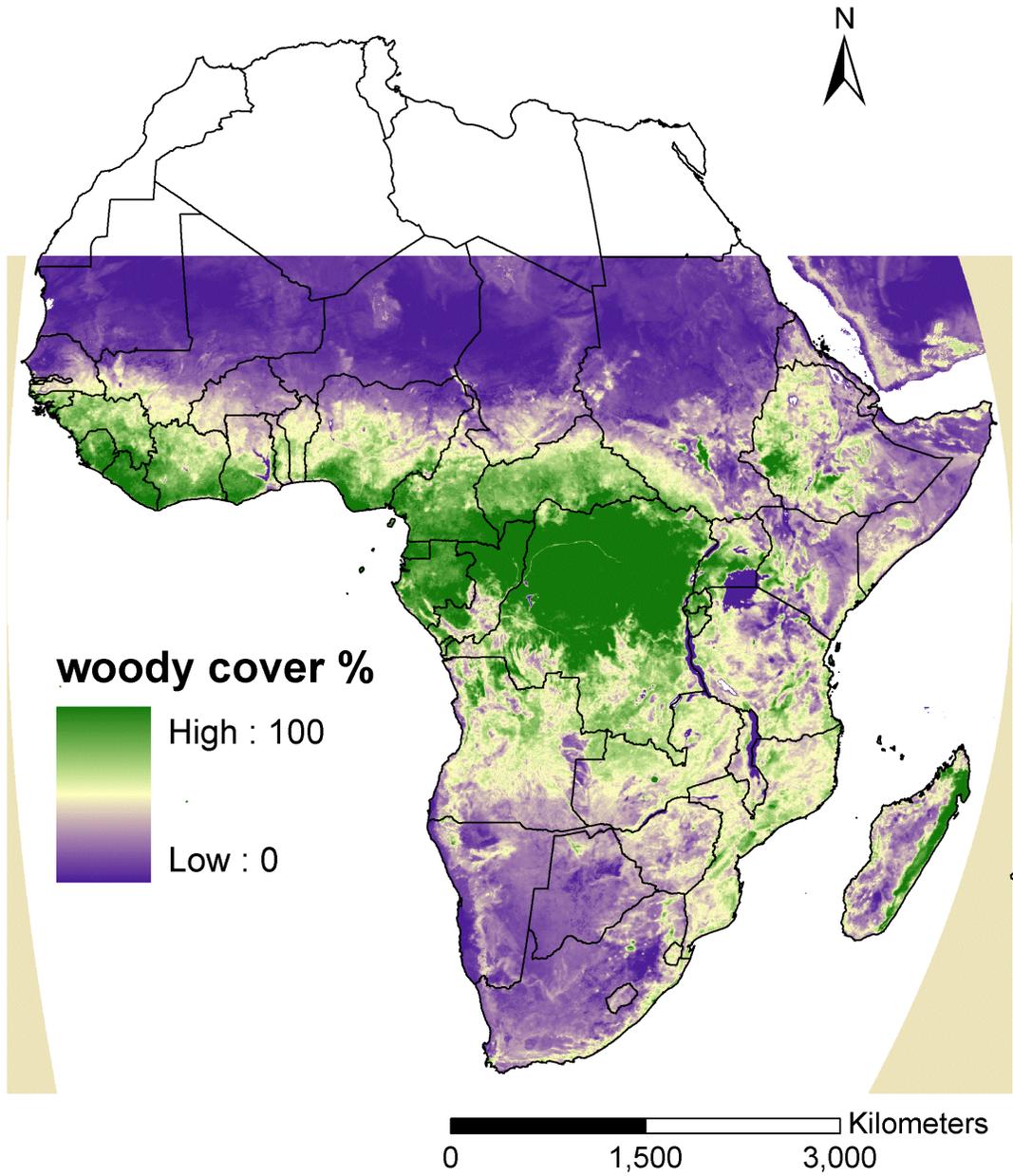
**Figure 4-21:** Continental model performance: observed (measured) versus predicted woody cover % for the training dataset (left) and for the validation dataset (right). For our purposes, model output values  $< 0\%$  were considered  $0\%$  and values  $> 100\%$  were considered  $100\%$ .

The woody cover percent map is shown in Figure 4-22. A visible characteristic of this map is the high spatial heterogeneity at several scales deriving from the combined radar, optical and altitude features. The expected continental-scale trend reflecting the mean annual rainfall gradients is well represented. Both the radar and the optical imagery are usually able to capture this trend at least in relative terms. The challenge is to both

attain accurate absolute cover values and to reproduce real smaller-scale patterns. Deserts have covers ranging from 0-10 %. A closer look at our training data revealed some points with relatively high cover in class 3 that might leverage it to slightly high levels. The desert highlands (Chad and Niger) and the escarpment in south Namibia stand out for unrealistically high cover. In these locations, surface roughness and topography result in high radar backscatter (van Zyl 1993) that the model translates into a false presence of woody cover (up to 40% overestimation). The dynamic radar range at low to medium cover levels (0-40%) is well reflected in gradients and regional patterns present in arid and mesic savannas. In the tropical wet forests of central Africa, cover has values > 80% as expected. Signal saturation in both NDVI (Huete et al. 1999) and radar backscatter (Lucas et al. 2000, Santos et al. 2002) in densely vegetated systems decreases sensitivity and the ability to detect cover variability. In our study, saturation was not pronounced and only limited to woody cover close to 100 % (Figure 4-19 and Figure 4-21). Around the northern and eastern sides of the Congo basin, we expected to find higher cover contrast between the Central African Republic and the Cameroon-Gabon areas. The central and northern parts of the Central African Republic (800-1300 mm mean annual rainfall) are characterized by humid and sub-humid savannas while the south and south-west Cameroon, Equatorial Guinea and northern Gabon (1300-2500 mm mean annual rainfall) are classified as dense moist forests (Eerens et al., <http://www.spot-vegetation.com/vegetationprogramme/Pages/vgtprep/vgt2000/eerens.pdf>). This effect on the map results from NDVI\_mean values generally lower (0.3-0.6) in the west than in the Central African Republic savannas (0.5-0.7). Likely, the low NDVI values are caused by the persistent stratiform-cloud-cover contamination (Le Moigne et al. 2002) that is

generated on the coast by the cold Atlantic waters and can spread far inland (Giresse 2007). The Q-SCAT backscatter, which is not affected by cloud cover, shows higher values for the western regions than the central savannas. The Rift Valley of East Africa is a very well recognizable topographic feature in the map and is characterized by low cover on higher altitudes and regional higher cover in valleys around lakes. Finally, agricultural areas are more difficult to evaluate for their variable woody cover dependent on the type of cultivation. For example, the area situated along the Senegalese coast north of Dakar is cultivated for vegetables (Sall and Vanclooster 2009) and results in low cover in our map. On the other hand, in the intensely cultivated areas in the Lake Victoria catchment in Uganda, our model predicts woody cover as high as the dense dry forests (both mean Q-SCAT and NDVI\_MEAN values similar to forested areas). These values could be realistic and are confirmed by observations of high tree density and biomass associated with indigenous agroforestry systems (Isabirye et al. 2008).

To further investigate qualities and weaknesses of our map, we compared it to the MODIS VCF product (continental scale), the optical-radar woody cover map produced for KNP (regional scale, Bucini et al. 2010), the vegetation-biomass map of the Turkana District, Kenya (regional scale, Ellis et al. 1987) and to photos taken on the ground in various parts of Namibia (local scale).



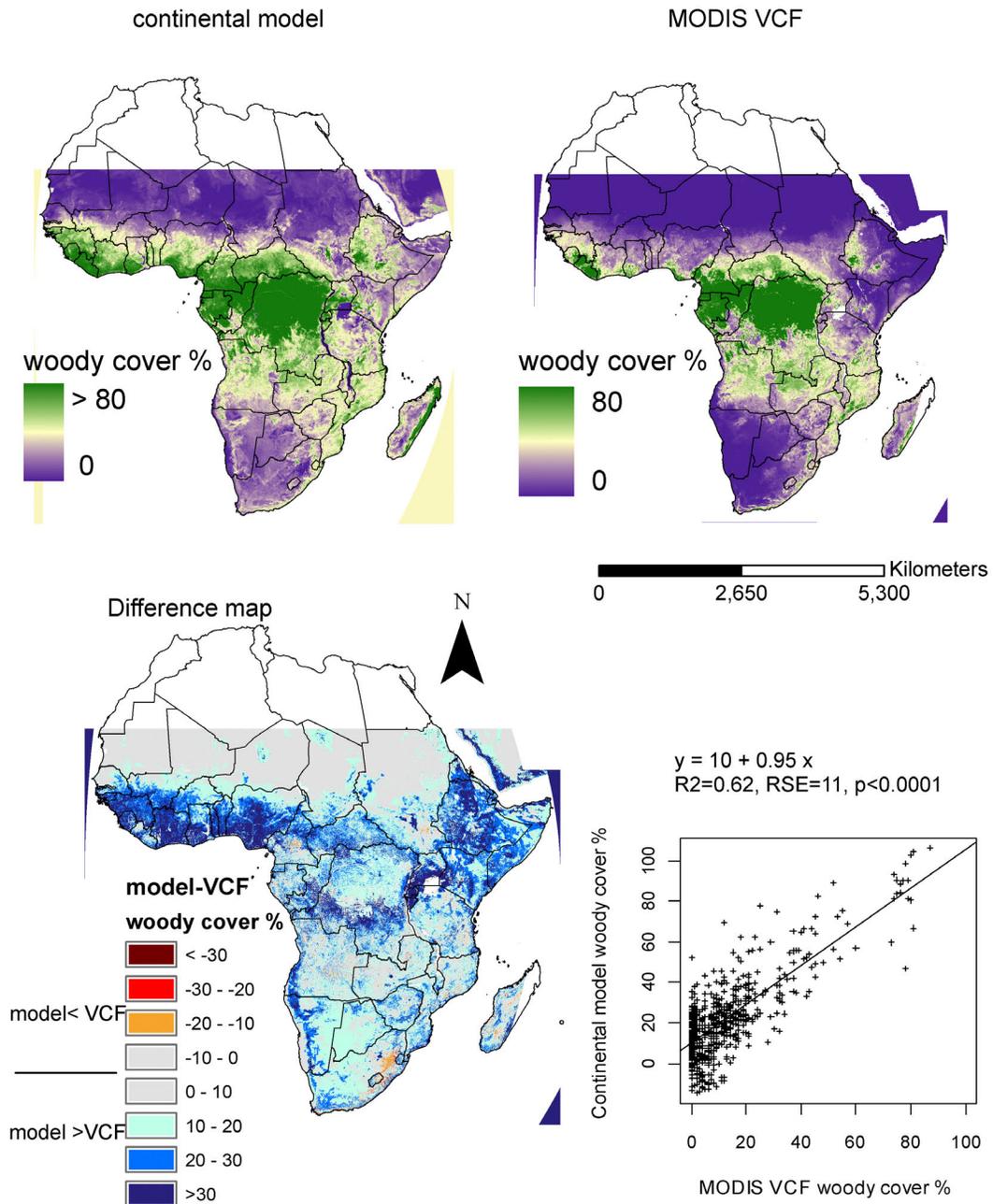
**Figure 4-22:** Woody cover percent map extrapolated from the continental model.

#### 4.3.1 MODIS VCF comparison

The regression line between VCF and our continental model woody cover % (Figure 4-23) shows an overall (continental scale) tendency of the VCF product to predict lower cover than the our model. We mapped the difference  $\Delta = \text{VCF} - \text{continental model cover \%}$  (Figure 4-23). The results from this comparison analysis are the following:

- In some desert areas, our model produces overestimates with  $2 < \text{cover} < 10\%$ . The VCF product correctly predicts 0 % cover.
- Arid savannas characterized by shrubs and thickets (White 1983) have near 0 % cover in the VCF map. In our map, they have up to 20% more cover (see in southern Africa: southeast Namibia, south Botswana, northeast South Africa; in west Africa: south Senegal, south Mali, Burkina Faso, southwest Chad and south east Sudan; in the Eastern Horn).
- The seasonal woodlands have comparable values in the two maps ( $-10 < \Delta < 10$ )
- The Guineo-Congolian areas along the African west coast (from Guinea Bissau to Nigeria) and the regions surrounding the dense tropical forests (Cameroon, Central African Republic, west Uganda and southern part of the Republic of Congo) are predicted with up to 30 % more cover with respect to VCF estimates. Here the mixture of deciduous, semi-deciduous and evergreen vegetation creates complex behaviors in annual metrics that could be difficult to model. The signal from woody structures is more consistently captured in the radar backscatter and could be the reason behind the prevalent higher cover. However, there are cases where Google Earth shows open areas with little or no woody plants where our map assigns nonzero cover.

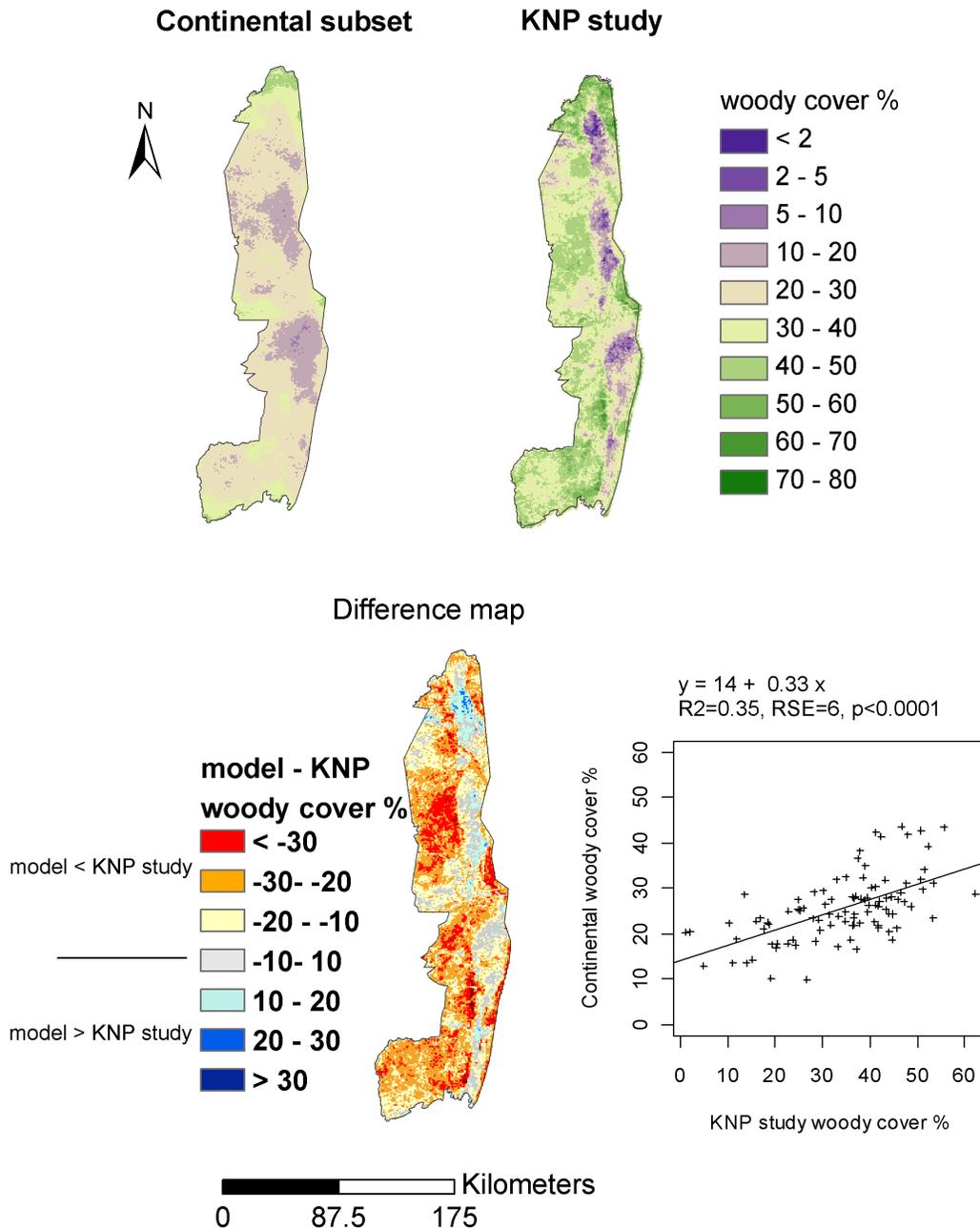
- The dense wet forests in the Congo basin have higher predicted woody cover in our map because the VCF product has a maximum 80 % cover value imposed by the cover classes used for training data.
- Our map is less affected by the persistent cloud cover in the central west areas (Gabon, Cameroon, Central Guinea) .
- One large area of higher MODIS VCF predictions is evident (orange-red shades) in eastern South Africa. This is a very intensely cultivated with no plantations. The predicted cover in the VCF map (0-20%) may be related to high greenness signal during the growing season.



**Figure 4-23:** Top: our continental model and the MODIS VCF woody cover % maps. Bottom: Difference map between our model and VCF woody cover %. Red shades for our model < MODIS VCF woody cover %; blue shades for our model > MODIS VCF woody cover %. Graph: Regression line between MODIS VCF and our continental model woody cover on the training dataset.

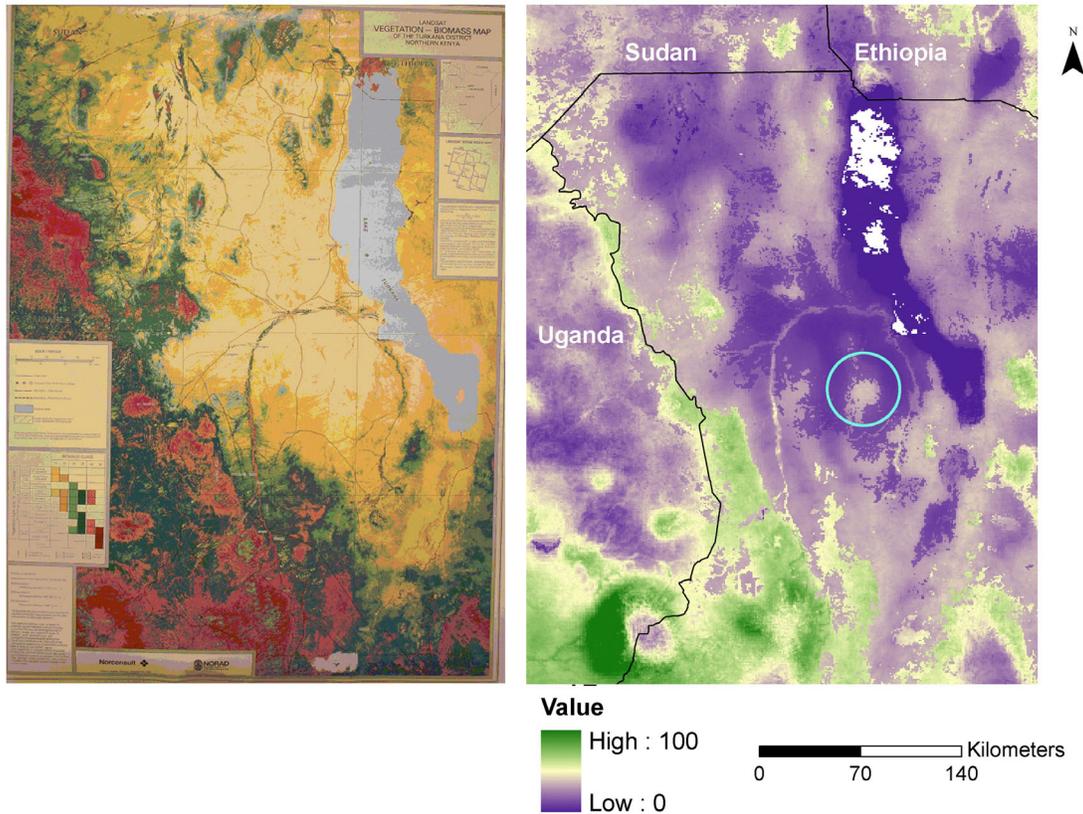
#### ***4.3.2 Comparison with KNP and Turkana woody cover maps***

At the regional scale of Kruger National Park, Figure 4-24 depicts the subset from the continental map, the KNP woody cover map (Bucini et al. 2010) resampled at 1 km resolution and their relative difference map. The continental map tends to predict lower woody cover % in the denser areas (west) but up to + 20 % cover on the eastern open-canopy areas. It broadly captures few lower-cover patches (east) and the higher cover areas in the northern and southern hills. Cross-validation with LiDAR cover estimates, gave good confidence about the KNP map estimates (Bucini et al. 2010). Therefore, we can say that in KNP, the coarse resolution of the remote-sensing layers used to build the continental layers homogenizes the cover patterns towards the average cover value but that the average spatial patterns are well represented.

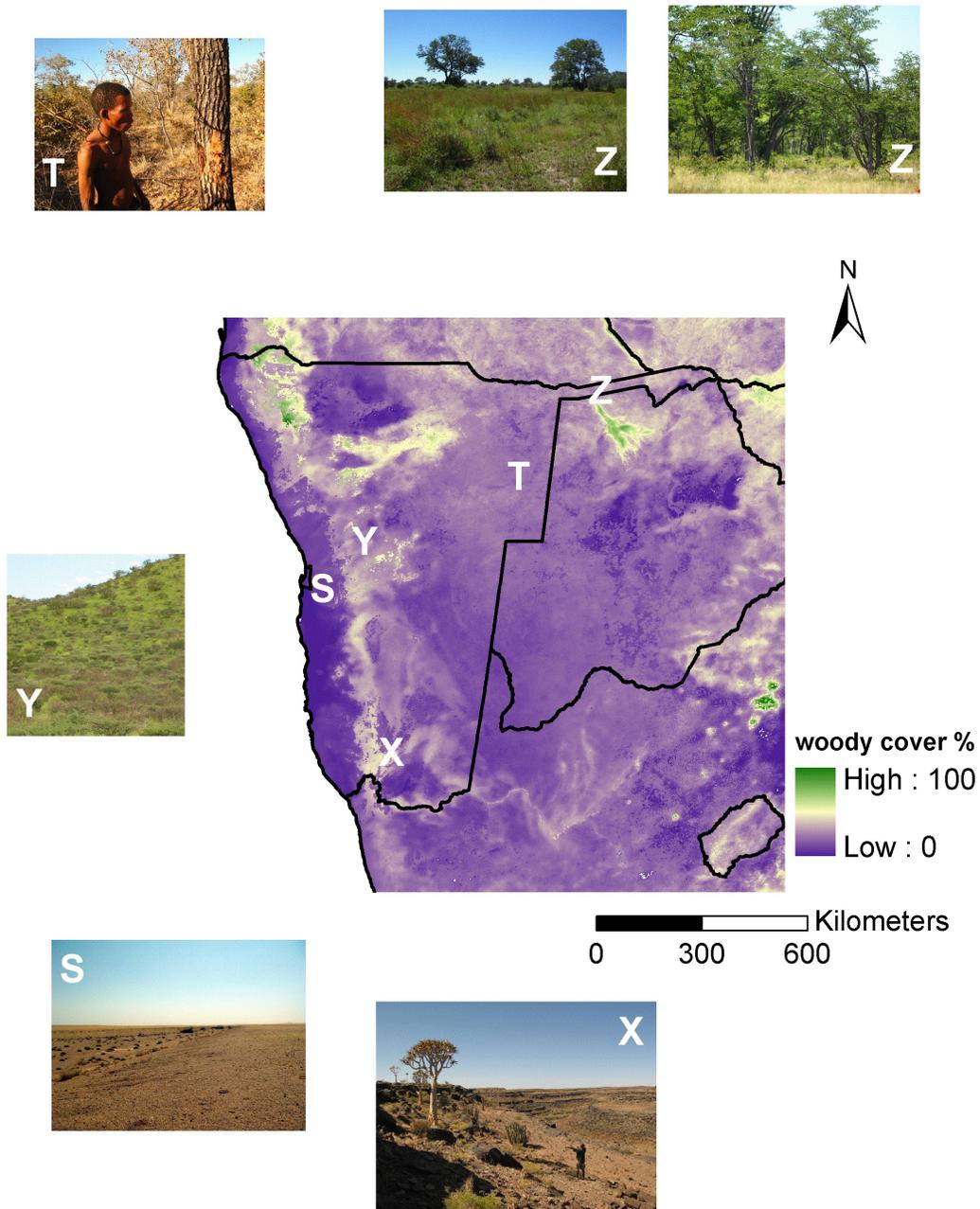


**Figure 4-24:** Top: woody cover % maps: subset from the continental map and KNP study map (Bucini et al. 2010). Bottom: difference map between the continental and the KNP study woody cover % maps. Graph: regression line between woody cover predictions from our continental model and the KNP study map on 100 random points across the KNP.

Moving to the Turkana district (Kenya), we compared our map (Figure 4-25) with a non-digital vegetation map (Ellis et al. 1987) reporting information about % canopy cover in four classes (<2%, 2-20%, >20%, closed canopy) derived from Landsat MSS imagery (1975 - 1979). The high variability in woody cover with patches of higher vegetation at the landscape scale is real and matches the patterns delineated by Ellis et al. in terms of both cover values and spatial location. Our map nicely separates the riparian vegetation along the Turkwel river (center-bottom Figure 4-25) and the gradient from Turkana Lake grasslands to the upland forests along the border with Uganda defined as closed canopy by Ellis et al. (50-60 % woody cover on our map). Our map features a high-cover patch in the southern part of this study area (blue circle) missing in Ellis et al. map but visible in Google Earth.



**Figure 4-25:** Turkana district, Kenya. Left: Ellis et al. (1987) map. Color codes: light yellow, cover < 2%; yellow, cover ~ 2-20%; light green, cover ~2 to >20%; dark green, cover >20%; red, > 20% to closed canopy. Right: continental woody cover map, zoom on the district (Turkana Lake is visible combining 0 % cover (black) and NoData (white)).



**Figure 4-26:** Zoom of continental woody cover map on Namibia. The letter S, T, X, Y and Z symbols indicate the area where the photos were taken. Predicted woody cover: ~8% at S, 15-20% at T, 15-30% at X, 15-25% at Y and 20-30% at Z.

### ***4.3.3 Woody cover comparison with ground photos in Namibia***

Ground pictures can also provide a qualitative way to evaluate the map outputs. We received six photos taken across the woody cover range of Namibia (Figure 4-26, courtesy of Theo Wassenaar, African Wilderness Restoration, Windhoek, Namibia). The far north east of the country receives about 600 mm of annual rain and can be densely wooded. The area around point Z has variable woody cover depending on fire and elephant interactions and comprises sites with very dense scrubs. Similarly to the KNP case, our continental map tends to miss the high woody cover spots and to predict a fairly homogeneous average cover around 20-30 %. The point T picture suggests low to medium cover similar or slightly higher than our map suggests (15-20% cover). The area of point Y appears dominated by fairly dense shrubs and some short scattered trees. Our continental map depicts variable cover (15-25%) but it is hard to evaluate our predictions from this photo. Point S lies in the central Namibia plains that are crossed by several large ephemeral rivers often supporting large trees. The continental map shows coarse patterns of river drainages (20-25% cover) surrounded by low cover landscapes. The Fish River Canyon (point X) is raviney, mountainous and sparsely vegetated. On the continental map, we find up to 35% cover overestimates. These overestimation zones are recognizable by their high backscatter on the mean Q-SCAT layer (van Zyl 1993) along the escarpment. The continental map appears fairly consistent with the predominant cover patterns found in the photographed areas except for the rough topographic zone in the country south.

#### 4.4 Discussion

Satellites provide a means for estimating ecologically relevant vegetation parameters such as cover, biomass, height, phenology and canopy structure that influence ecosystem functioning. In this paper, we used statistical model-fitting techniques to retrieve woody cover percent from optical and SAR (radar) systems. This is the first attempt to combine the two systems at the continental scale of Africa for woody cover predictions. The interest for this work mainly spawned from the need for a woody resource map that better represents savannas.

Because we were interested in cover, we sought radiative signatures related to canopy. We used data in the shortwave radar domain (Ku band,  $\lambda \sim 2\text{cm}$ ) where the contribution to backscatter mainly comes from two vegetation components: the season-specific leaf component and the small branch woody component (Saatchi et al. 2007, Saatchi et al. 2008). We combined the radar information with annual NDVI metrics reflecting greenness dynamics. We found that the Q-SCAT annual mean backscatter and mean annual NDVI were the most important variables to predict woody cover with the data we had at hand. Q-scat backscatter intensity was related to cover with an exponential relationship indicating non-linearity and higher dynamic range in the low-to-medium backscatter intensities (Table 4-10). Importantly, we found significant large-scale relationships denoting a consistent spectral response to woody cover across ecosystems (Table 4-10). From these results, we can answer our third research question and conclude that a synergistic approach with radar and optical observations provides a framework to

overcome limitations of single sensors and exploit their complementary sensitivities to woody cover.

We found a relatively high model selection uncertainty with three competing models presenting substantial level of support in the data. Two extra predictors could serve nearly equally well in modeling the data information. Model averaging can in this case provide a more stabilized inference and better precision (Burnham and Anderson 2002). For the manuscript, we plan to create a new woody cover map based on model averaged parameters weighted by their BIC probability ( $w$ ) using the set of the three best models.

The sensitivity of a specific sensor or a spectral band to woody vegetation can change across environments. We noticed for example that Q-SCAT backscatter maintained higher spectral sensitivity than NDVI in arid areas and allowed to predict shrub and small tree cover. On the other hand, bare surfaces such as outcrops or cliffs can exhibit strong backscatter (Vanzyl et al. 1993) resulting from surface instead of vegetation roughness. This special effect in the radar data is absent in optical radiometry. In our work, this resulted in significant overestimates of woody cover in some mountainous areas (e.g., Namibia escarpment and Saharan highlands). In final preparation of this chapter for journal submission, we will examine ways to remove or correct for these topographic effects. Optical data combined with elevation information could provide a heuristic toolbox for terrain corrections on radar backscatter.

In our modeling procedure, several sources of errors affect model predictive power (Table 4-11). The variability in spectral metrics mainly arises from diversity in plant growth-forms, phenology, spatial arrangement and density as well as surface

backgrounds and topography. Our training data captured a fairly good proportion of this variation but did not represent the entire spectral space (Figure 4-17). The extrapolated map however presented reasonable values with some areas that need further investigation and/or increased sampling such as the deserts and topographically variable sites and the region around the Cameroon-Gabon border. We believe that the relatively high unexplained variability (41%) and model error (15% cover on training data and 19% cover on validation data) also originated from the scale mismatch between field and remote-sensing data (Table 4-12). The heterogeneity found at the plot scale (~1 ha) did not directly translate in spectral variability and hence could not be accounted for by the model.

<b>Error origin</b>		<b>Mean error</b>
Sample data	Field data	11.6% cover
	Google Earth estimations	3.7 % cover
Model fit RSE		15.1 % cover
Image co-registration		3 pixels

**Table 4-12:** Different sources of error that could potentially affect the woody cover map. Note that other sources of error occur.

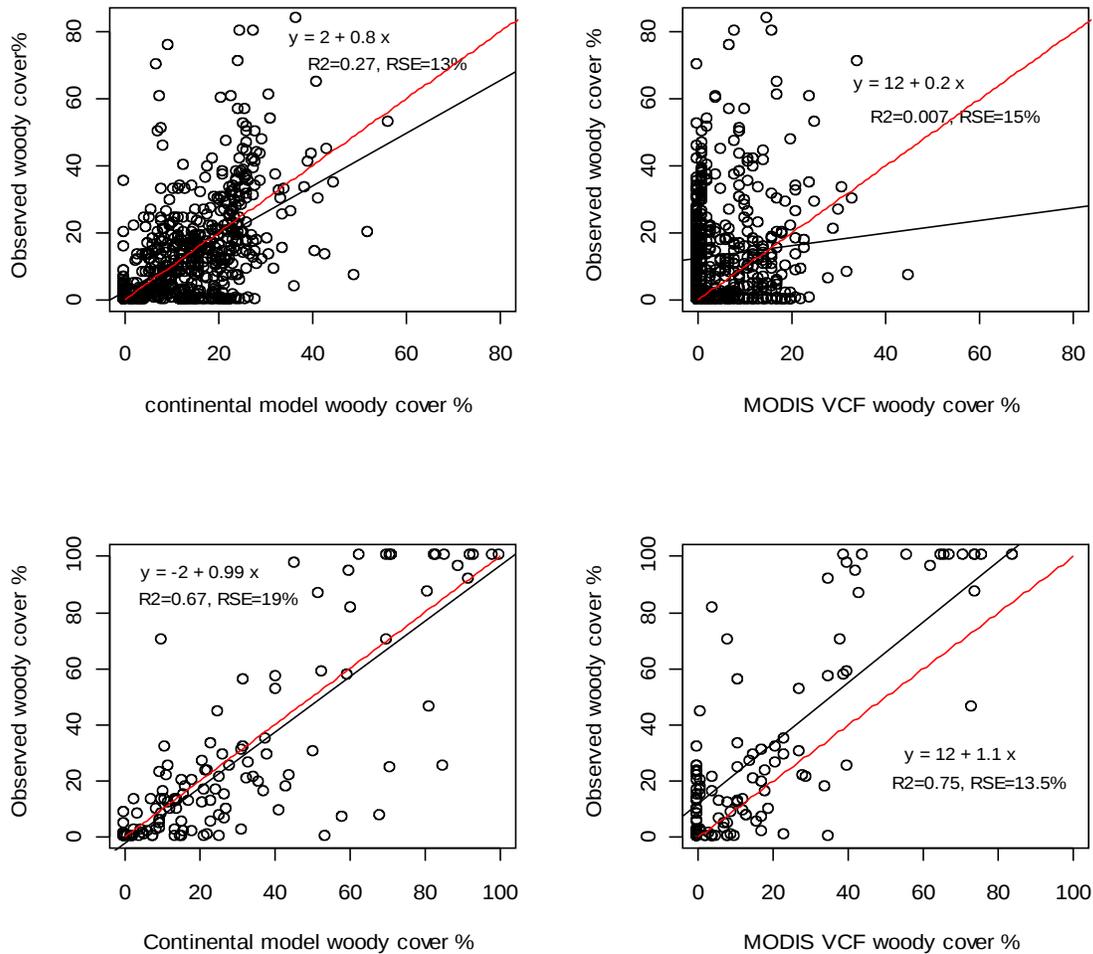
A possible way to improve model predictions is to put more effort in the training data collection. Very high-resolution imagery ( $\leq 1$  m) provides an appropriate base to detect woody plants including shrubs and trees but assessments can be labor-intensive. Object-oriented analysis coupled with rule-based classification algorithms can be partially automated and offer an improved approach for crown detection (Laliberte et al.

2004, Bunting and Lucas 2006, Laliberte et al. 2007). Investing in a large high-resolution image sample has a high informative potential: its comprehensive radiometric properties can be modeled to determine relationships with woody cover. These relationships can then be confidently extrapolated and repeatedly applied over time on new remote-sensing data acquisitions.

The adoption of SAR imagery with higher spatial resolution and different wavelengths than Q-SCAT (2.25 km) could also improve mapping of woody cover. For example, the ENVISAT satellite operating at C-band with original resolution of 30 m could be considered for future mapping projects. Furthermore, longwave SAR (L- and P-bands) interacting with deeper vegetation layers can be adopted to better discriminate woody vegetation. The Advanced Land Observing Satellite (ALOS) launched in 2006 is an important resource to explore for the high data quality that it provides and its multi-temporal observations at global scale (Shimada et al. 2009). The data acquired by the Phase Array L-band SAR (PALSAR) on-board ALOS have already been used in regional-scale savanna to study vegetation biomass and structure (Lucas and Armston 2007, Mitchard et al. 2009). In particular, simulations based on data from central Queensland, Australia, (Lucas et al. 2006) suggest that in medium dense vegetation with stem heights about 2-5 m, the integration of L-band HH and HV backscatter brings information from both stem-ground double bounce and branch volume scattering, respectively. In mixed vegetation systems, this extra information can enhance the ability to detect the presence of small to medium size woody plants.

Our work has demonstrated that it is possible to retrieve reliable woody cover estimates for savanna systems despite their spectral complexity arising from the presence

of coexisting woody and herbaceous vegetation. Field measurements from savanna sites evidence that our map has improved the woody cover estimates currently provided by the VCF product built on optical remote-sensing alone (Figure 4-27). If we consider the full range of woody cover values (0-100%), MODIS VCF performs better in terms of  $R^2$  and residual standard error (RSE) than the continental model (Figure 4-27). A plausible reason for the higher point dispersion related the continental model is the scale mismatch between many of the field plots (~1 ha) and the 1-km remote-sensing pixel: the higher small-scale woody cover variability is smoothed in the coarser remote-sensing pixel signals and remains unexplained by the continental model. On the other hand, the sampling method based on aggregating Landsat cover estimates, results in tree cover values spatially more concordant to the remote-sensing MODIS pixels (500 m). For this reason the relationship between predicted and observed cover is tighter in the case of MODIS VCF. However, the bias introduced in the sample tree cover value by assigning 0% cover to arid systems and a 80% tree cover upper threshold, results in a significant intercept of 12 % cover (p-value < 0.0001) and a general woody cover underestimation.



**Figure 4-27:** Top: Field measured (Sankaran et al. 2005) vs. predicted woody cover. Bottom: Field measured (validation set) vs. predicted woody cover. Left: x-axis represents woody cover predicted by the continental model; right: x-axis represents woody cover predicted by MODIS VCF resampled at 1 km resolution. In black: data and regression lines; in red: one-to-one line.

#### 4.4.1 Development and applications

Variations and temporal trends in vegetation properties are critical to characterize and monitor ecosystems. The African continent has been undergoing extensive land cover

change and experiencing land degradation especially in arid systems (Verstraete et al. 2009). Limited research (Williams et al. 2008, Scheiter and Higgins 2009) has been conducted at the continental scale to understand consequences of climate change on natural resources and ecosystem services derived from woody plants. One of the reasons for a delay in scientific understanding of continental vegetation dynamics is a lack of appropriate data and adequate dynamic global vegetation models especially for woodland-savanna systems (Scheiter and Higgins 2009). Our modeling approach has the potential to create improved maps and hence contribute to scientific progress. A pertinent direction for future research is the creation of an algorithm for regular temporal assessments of woody cover at continental scale.

#### **4.5 Conclusions**

The use of combined optical and shortwave radar remote sensing enhanced the ability to relate woody cover data to remote-sensing observations and to build strong predictive models across African eco-regions. Large training datasets are necessary to represent woody cover variability in space across the continent and in the spectral space of remote-sensing data.

We improved the current continental information on woody cover particularly in savannas by accounting for shrubs and short trees that were ignored in previous continental-scale analyses. Our mapping results are consistent with other woody cover information sources from continental to regional scale. This research can provide the necessary map accuracy for robust ecological inference. Because African ecosystems endure increasing land-use pressures, remote sensing analyses focused on quantifying

vegetation structural characteristics and their change can significantly contribute to understanding change in the continent and the role of Africa in global carbon, water and biogeochemical cycles.

## 4.6 References

- Archibald, S., and R. J. Scholes. 2007. Leaf green-up in a semi-arid African savanna - separating tree and grass responses to environmental cues. *Journal of Vegetation Science* **18**:583-594.
- Asner, G. P., C. A. Wessman, and D. S. Schimel. 1998. Heterogeneity of savanna canopy structure and function from imaging spectrometry and inverse modeling. *Ecological Applications* **8**:1022-1036.
- Arthur, M. d. F. S. R., S. Zahran, and G. Bucini. 2010. On the adoption of electricity as a domestic source by Mozambican households. *Journal of Energy Policy*. (accepted for publication)
- Baldocchi, D. D., L. K. Xu, and N. Kiang. 2004. How plant functional-type, weather, seasonal drought, and soil physical properties alter water and energy fluxes of an oak-grass savanna and an annual grassland. *Agricultural and Forest Meteorology* **123**:13-39.
- Bartholome, E., and A. S. Belward. 2005. GLC2000: a new approach to global land cover mapping from Earth observation data. *International Journal of Remote Sensing* **26**:1959-1977.
- Belsky, A. J., S. M. Mwonga, R. G. Amundson, J. M. Duxbury, and A. R. Ali. 1993. Comparative Effects of Isolated Trees on Their Undercanopy Environments in High-Rainfall and Low-Rainfall Savannas. *Journal of Applied Ecology* **30**:143-155.
- Bucini, G., N. P. Hanan, R. B. Boone, I. P. J. Smit, S. Saatchi, M. A. Lefsky, and G. P. Asner. 2010. Woody fractional cover in Kruger National Park, South Africa: remote-sensing-based maps and ecological insights. *in* M. J. Hill and N. P. Hanan, editors. *Ecosystem function in savannas: measurement and modeling at landscape to global scales*. CRC/Taylor and Francis.
- Bunting, P., and R. Lucas. 2006. The delineation of tree crowns in Australian mixed species forests using hyperspectral Compact Airborne Spectrographic Imager (CASI) data. *Remote Sensing of Environment* **101**:230-248.
- Burnham, K. P., and D. R. Anderson. 2002. *Model selection and inference: a practical information theoretic approach*. 2nd edition. Springer-Verlag, New York.
- Butterfield, H. S., and C. M. Malmstrom. 2009. The effects of phenology on indirect measures of aboveground biomass in annual grasses. *International Journal of Remote Sensing* **30**:3133-3146.
- De Fries, R. S., M. Hansen, J. R. G. Townshend, and R. Sohlberg. 1998. Global land cover classifications at 8 km spatial resolution: the use of training data derived from Landsat imagery in decision tree classifiers. *International Journal of Remote Sensing* **19**:3141-3168.
- DeFries, R., M. Hansen, M. Steininger, R. Dubayah, R. Sohlberg, and J. Townshend. 1997. Subpixel forest cover in central Africa from multisensor, multitemporal data. *Remote Sensing of Environment* **60**:228-246.
- DeFries, R., M. Hansen, and J. Townshend. 1995. Global discrimination of land cover types from metrics derived from AVHRR pathfinder data. *Remote Sensing of Environment* **54**:209-222.

- DeFries, R. S., J. R. G. Townshend, and M. C. Hansen. 1999. Continuous fields of vegetation characteristics at the global scale at 1-km resolution. *Journal of Geophysical Research-Atmospheres* **104**:16911-16923.
- Do, F. C., V. A. Goudiaby, O. Gimenez, A. L. Diagne, M. Diouf, A. Rocheteau, and L. E. Akpo. 2005. Environmental influence on canopy phenology in the dry tropics. *Forest Ecology and Management* **215**:319-328.
- du Toit, J. T. 2003. Large Herbivores and Savanna Heterogeneity. Pages 292-309 in J. T. du Toit, Rogers, K. H. and Biggs H. C. , editor. *The Kruger Experience: Ecology and Management of Savanna Heterogeneity*. Island Press, Washington.
- Ellis, J. K., K. Galvin, J. T. McCabe, and D. M. Swift. 1987. Pastoralism and drought in Turkana District, Kenya. NORAD, Nairobi.
- Giresse, P. 2007. TROPICAL AND SUB-TROPICAL WEST AFRICA - MARINE AND CONTINENTAL CHANGES DURING THE LATE QUATERNARY. Elsevier Science
- Gromping, U. 2007. Estimators of relative importance in linear regression based on variance decomposition. *American Statistician* **61**:139-147.
- Hanan, N. P. 2001. Enhanced two-layer radiative transfer scheme for a land surface model with a discontinuous upper canopy. *Agricultural and Forest Meteorology* **109**:265-281.
- Hansen, M. C., R. DeFries, J. R. Townshend, M. Carroll, C. Dimiceli, and R. Sohlberg. 2003. Vegetation Continuous Fields MOD44B, 2001 Percent Tree Cover, Collection 3. University of Maryland, College Park, Maryland, 2001.
- Hansen, M. C., R. S. Defries, J. R. G. Townshend, and R. Sohlberg. 2000. Global land cover classification at 1km spatial resolution using a classification tree approach. *International Journal of Remote Sensing* **21**:1331-1364.
- Hansen, M. C., R. S. DeFries, J. R. G. Townshend, R. Sohlberg, C. Dimiceli, and M. Carroll. 2002. Towards an operational MODIS continuous field of percent tree cover algorithm: examples using AVHRR and MODIS data. *Remote Sensing of Environment* **83**:303-319.
- Huete, A., K. Didan, W. van Leeuwen, and E. Vermote. 1999. Global-scale analysis of vegetation indices for moderate resolution monitoring of terrestrial vegetation. Pages 141-151 in G. Cecchi, E. T. Engman, and E. Zilioli, editors. *Remote Sensing for Earth Science, Ocean, and Sea Ice Applications*. Spie-Int Soc Optical Engineering, Bellingham.
- Isabirye, M., B. Verbist, M. K. Magunda, J. Poesen, and J. Deckers. 2008. Tree density and biomass assessment in agricultural systems around Lake Victoria, Uganda. *African Journal of Ecology* **46**:59-65.
- Kasischke, E. S., J. M. Melack, and M. C. Dobson. 1997. The use of imaging radars for ecological applications - A review. *Remote Sensing of Environment* **59**:141-156.
- Laliberte, A. S., E. L. Fredrickson, and A. Rango. 2007. Combining decision trees with hierarchical object-oriented image analysis for mapping arid rangelands. *Photogrammetric Engineering and Remote Sensing* **73**:197-207.
- Laliberte, A. S., A. Rango, K. M. Havstad, J. F. Paris, R. F. Beck, R. McNeely, and A. L. Gonzalez. 2004. Object-oriented image analysis for mapping shrub encroachment from 1937 to 2003 in southern New Mexico. *Remote Sensing of Environment* **93**:198-210.

- Le Moigne, J., N. Laporte, and N. S. Netanyahu. 2002. Enhancement of tropical land cover mapping with wavelet-based fusion and unsupervised clustering of SAR and landsat image data. Pages 190-198 in S. B. Serpico, editor. *Image and Signal Processing for Remote Sensing VII*.
- Lucas, R. M., and J. D. Armston. 2007. ALOS PALSAR for characterizing wooded savannas in northern Australia. Pages 3610-3613 *Igarss: 2007 IEEE International Geoscience and Remote Sensing Symposium, Vols 1-12 - Sensing and Understanding Our Planet*. IEEE, New York.
- Lucas, R. M., N. Cronin, M. Moghaddam, A. Lee, J. Armston, P. Bunting, and C. Witte. 2006. Integration of radar and Landsat-derived foliage projected cover for woody regrowth mapping, Queensland, Australia. *Remote Sensing of Environment* **100**:388-406.
- Lucas, R. M., A. K. Milne, N. Cronin, C. Witte, and R. Denham. 2000. The potential of synthetic aperture radar (SAR) for quantifying the biomass of Australia's woodlands. *Rangeland Journal* **22**:124-140.
- Lucas, R. M., A. L. Mitchell, A. Rosenqvist, C. Proisy, A. Melius, and C. Ticehurst. 2007. The potential of L-band SAR for quantifying mangrove characteristics and change: case studies from the tropics. *Aquatic Conservation-Marine and Freshwater Ecosystems* **17**:245-264.
- Lucas, R. M., M. Moghaddam, and N. Cronin. 2004. Microwave scattering from mixed-species forests, Queensland, Australia. *IEEE Transactions on Geoscience and Remote Sensing* **42**:2142-2159.
- Ludwig, F., H. De Kroon, and H. H. T. Prins. 2008. Impacts of savanna trees on forage quality for a large African herbivore. *Oecologia* **155**:487-496.
- Mahiri, I., and C. Howorth. 2001. Twenty years of resolving the irresolvable: Approaches to the fuelwood problem in Kenya. *Land Degradation & Development* **12**:205-215.
- Mitchard, E. T. A., S. S. Saatchi, I. H. Woodhouse, G. Nangendo, N. S. Ribeiro, M. Williams, C. M. Ryan, S. L. Lewis, T. R. Feldpausch, and P. Meir. 2009. Using satellite radar backscatter to predict above-ground woody biomass: A consistent relationship across four different African landscapes. *Geophysical Research Letters* **36**:6.
- Mougin, E., P. Hiernaux, L. Kergoat, M. Grippa, P. de Rosnay, F. Timouk, V. Le Dantec, V. Demarez, F. Lavenue, M. Arjounin, T. Lebel, N. Soumaguel, E. Ceschia, B. Mougnot, F. Baup, F. Frappart, P. L. Frison, J. Gardelle, C. Gruhier, L. Jarlan, S. Mangiarotti, B. Sanou, Y. Tracol, F. Guichard, V. Trichon, L. Diarra, A. Soumare, M. Koite, F. Dembele, C. Lloyd, N. P. Hanan, C. Damesin, C. Delon, D. Serca, C. Galy-Lacaux, J. Seghieri, S. Becerra, H. Dia, F. Gangneron, and P. Mazzega. 2009. The AMMA-CATCH Gourma observatory site in Mali: Relating climatic variations to changes in vegetation, surface hydrology, fluxes and natural resources. *Journal of Hydrology* **375**:14-33.
- Owen-Smith, N. 2004. Functional heterogeneity in resources within landscapes and herbivore population dynamics. *Landscape Ecology* **19**:761-771.
- Prasad, V. K., K. V. S. Badarinath, and A. Eaturu. 2007. Spatial patterns of vegetation phenology metrics and related climatic controls of eight contrasting forest types in India - analysis from remote sensing datasets. *Theoretical and Applied Climatology* **89**:95-107.

- Ratnam, J., M. Sankaran, N. P. Hanan, R. C. Grant, and N. Zambatis. 2008. Nutrient resorption patterns of plant functional groups in a tropical savanna: variation and functional significance. *Oecologia* **157**:141-151.
- Ribeiro, N. S., S. S. Saatchi, H. H. Shugart, and R. A. Washington-Allen. 2008. Aboveground biomass and leaf area index (LAI) mapping for Niassa Reserve, northern Mozambique. *Journal of Geophysical Research-Biogeosciences* **113**.
- Rokhmatuloh, D. Nitto, H. Al Bilbisi, and R. Tateishi. 2005. Percent tree cover estimation using regression tree method: a case study of Africa with very-high resolution QuickBird images as training data. Pages 2157-2160 *in* Geoscience and Remote Sensing Symposium, 2005. IGARSS '05. Proceedings. 2005 IEEE International.
- Saatchi, S., W. Buermann, H. Ter Steege, S. Mori, and T. B. Smith. 2008. Modeling distribution of Amazonian tree species and diversity using remote sensing measurements. *Remote Sensing of Environment* **112**:2000-2017.
- Saatchi, S. S., R. A. Houghton, R. Alvala, J. V. Soares, and Y. Yu. 2007. Distribution of aboveground live biomass in the Amazon basin. *Global Change Biology* **13**:816-837.
- Sall, M., and M. Vanclouster. 2009. Assessing the well water pollution problem by nitrates in the small scale farming systems of the Niayes region, Senegal. *Agricultural Water Management* **96**:1360-1368.
- Sankaran, M., N. P. Hanan, R. J. Scholes, J. Ratnam, D. J. Augustine, B. S. Cade, J. Gignoux, S. I. Higgins, X. Le Roux, F. Ludwig, J. Ardo, F. Banyikwa, A. Bronn, G. Bucini, K. K. Caylor, M. B. Coughenour, A. Diouf, W. Ekaya, C. J. Feral, E. C. February, P. G. H. Frost, P. Hiernaux, H. Hrabar, K. L. Metzger, H. H. T. Prins, S. Ringrose, W. Sea, J. Tews, J. Worden, and N. Zambatis. 2005. Determinants of woody cover in African savannas. *Nature* **438**:846-849.
- Santos, J. R., M. S. P. Lacruz, L. S. Araujo, and M. Keil. 2002. Savanna and tropical rainforest biomass estimation and spatialization using JERS-1 data. *International Journal of Remote Sensing* **23**:1217-1229.
- Scanlon, T. M., J. D. Albertson, K. K. Caylor, and C. A. Williams. 2002. Determining land surface fractional cover from NDVI and rainfall time series for a savanna ecosystem. *Remote Sensing of Environment* **82**:376-388.
- Scheiter, S., and S. I. Higgins. 2009. Impacts of climate change on the vegetation of Africa: an adaptive dynamic vegetation modelling approach. *Global Change Biology* **15**:2224-2246.
- Schwarz, G. 1978. Estimating Dimension of a Model. *Annals of Statistics* **6**:461-464.
- Shackleton, C. M., and J. Gambiza. 2008. Social and ecological trade offs in combating land degradation: The case of invasion by a woody shrub (*Euryops floribundus*) at Macubeni, South Africa. *Land Degradation & Development* **19**:454-464.
- Shimada, M., R. Touzi, T. Tadono, and J. A. Smith. 2009. Foreword to the Special Issue on Calibration and Validation of ALOS Sensors (PALSAR, AVNIR-2, and PRISM) and Their Use for Bio- and Geophysical Parameter Retrievals. *Ieee Transactions on Geoscience and Remote Sensing* **47**:3911-3913.
- van Zyl, J. J. 1993. The effect of topography on radar scattering from vegetated areas. *Geoscience and Remote Sensing, IEEE Transactions on* **31**:153-160.

- Vanzyl, J. J., B. D. Chapman, P. Dubois, and J. C. Shi. 1993. The Effect of Topography on Sar Calibration. *Ieee Transactions on Geoscience and Remote Sensing* **31**:1036-1043.
- Verstraete, M. M., R. J. Scholes, and M. S. Smith. 2009. Climate and desertification: looking at an old problem through new lenses. *Frontiers in Ecology and the Environment* **7**:421-428.
- Viergever, K. M., I. H. Woodhouse, and N. Stuart. 2007. backscatter and interferometry for estimating above- ground biomass in tropical savanna woodland. Pages 2346-2349 in *Geoscience and Remote Sensing Symposium, 2007. IGARSS 2007. IEEE International Geoscience and Remote Sensing Symposium, 2007. IGARSS 2007. IEEE International VO -*.
- Waring, R. H., J. B. Way, E. R. Hunt, L. Morrissey, K. J. Ranson, J. F. Weishampel, R. Oren, and S. E. Franklin. 1995. Biologists Toolbox - Imaging Radar for Ecosystem Studies. *Bioscience* **45**:715-723.
- White, F. 1983. *Vegetation of Africa - a descriptive memoir to accompany the Unesco/AETFAT/UNSO vegetation map of Africa*. U. N. Educational, Scientific and Cultural Organization, Place de Fontenoy, 75700 Paris, France.
- Wigley, B. J., W. J. Bond, and M. T. Hoffman. 2009. Bush encroachment under three contrasting land-use practices in a mesic South African savanna. *African Journal of Ecology* **47**:62-70.
- Williams, C., N. Hanan, J. Neff, R. Scholes, J. Berry, A. S. Denning, and D. Baker. 2007. Africa and the global carbon cycle. *Carbon Balance and Management* **2**:3.
- Williams, C. A., N. P. Hanan, I. Baker, G. J. Collatz, J. Berry, and A. S. Denning. 2008. Interannual variability of photosynthesis across Africa and its attribution. *Journal of Geophysical Research-Biogeosciences* **113**:15.
- Xiao, X. M., Q. Y. Zhang, S. Saleska, L. Hutyyra, P. De Camargo, S. Wofsy, S. Frolking, S. Boles, M. Keller, and B. Moore. 2005. Satellite-based modeling of gross primary production in a seasonally moist tropical evergreen forest. *Remote Sensing of Environment* **94**:105-122.

## **CHAPTER 5**

### **SUMMARY AND CONCLUSIONS**

The overall objectives of this dissertation, to map woody cover and model its dependence on biotic and abiotic factors, were addressed at two spatial scales, regional and continental. The research focused on African savannas where woody plants coexist with herbaceous vegetation and explored the interface between ecosystem ecology and remote-sensing science. These two disciplines have long blended because ecosystem ecologists recognized remote-sensing as a unique and powerful means to expand research to larger scales with consistent long-term observations.

A thread underlying my research is an interest in scale. Ecosystem ecology deals with interactions between organisms and their environment as a whole system. It recognizes that we observe a phenomenon/pattern at a specific scale but this phenomenon/pattern is the result of processes taking place also at finer and broader scales (spatial and temporal). One of the most fascinating objectives of ecosystem ecology is to integrate multi-scale dynamics to understand ecosystem functioning. I see research as an active effort to both test and integrate insights from three different sources: theoretical (mechanistic) work, experiments and empirical correlations. A critical place to start is to measure the variable of interest and to define its spatial detail and extent.

Under the overarching question, what creates and maintains the coexistence of herbaceous and woody vegetation in savannas, I have taken a heuristic approach to explore how complex interactions among biotic and abiotic factors manifest as patterns in woody vegetation cover and I sought interrelations between regional and continental scales. Many ecological theories (Sankaran et al. 2004) have been developed to explain tree-grass coexistence and tested with experiments. From this body of work, I have learned that coexistence mechanisms involve direct interactions between the two

vegetation functional types for resource access as well as interactions with disturbance factors such as fire, herbivory and anthropogenic activity. These interactions always need to be contextualized to account for environmental and physical conditions.

My research work developed in response to the need for accurate woody cover data and for empirical evidence in relation to conceptual models of tree-grass coexistence:

- The ecological model (chapter 2) developed to explain continental woody cover patterns derived from MODIS sensor (Hansen et al. 2006) revealed a multi-scale behavior. Continental gradients of woody cover dependent on mean annual rainfall (MAP) were found to be broken down into finer-scale patterns controlled by a set of different variables including perturbations and soil factors.
- The Kruger National Park (KNP) case study provided the opportunity to directly explore regional interactions. Using a woody cover map that we created for the Park, we found that the underlying parent material was the main driver in shaping woody cover across the Park and that mean annual rainfall was correlated with some landscape-scale patterns (chapter 3). We also found significant effects from fire frequency, elephant density, elevation and slope. Other tested variables such as rainfall variability, grazer density, browser density, distance to water and aspect did not significantly influence the ecological model.
- The continental study (chapter 2) also brought the attention to the inadequacy of the MODIS VCF tree cover percent map for savanna research. This map does not include trees < 5 m height, which are dominant in savanna systems and therefore

provides underestimated cover values in systems where small trees and shrubs are common.

- The regional-scale mapping process (chapter 3) provided accurate and more detailed information of woody cover for the savannas of KNP. At the same time, it was a learning ground for sampling strategies and for testing the suitability of remote-sensing data types for woody cover mapping. Combined optical and radar remote-sensing were found to significantly increase the ability to predict woody cover. This result directed the design of the continental mapping project (chapter 4).
- The continental mapping project confirmed that woody cover models greatly benefit from including predictors derived from both optical and radar data. The strong relationship found at regional scale between woody cover and optical and radar variables holds at larger scales from shrublands across savannas to seasonal woodlands and dense forests. Our approach improved mapping with more realistic cover estimates in savannas where radar remained sensitive to woody structures and able to differentiate their signal from the grass signal. In general, our woody cover estimates across the continent are higher compared to the MODIS VCF estimates.

## **5.1 Ecological models**

The ecological/environmental models developed for my work have a semi-empirical (continent) and empirical (regional) nature. Statistical models must be taken with prudence because correlations can also result from chance or from biases of modeling methods. However, several statistical techniques have been developed to quantify model strength and I tried to use the most up-to-date and validated approaches. I see empirical models contributing to ecological research in three ways: (i) they can

provide initial suggestions for mechanisms, (ii) given a theoretical model, they can provide empirical evidence for hypothesized mechanisms and (iii) they can find correlations in the data that were not expected but are related to real ecological phenomena.

The models tested in this dissertation quantitatively described how woody cover is organized in relation to biotic and abiotic factors. The continental model was structured to test several possible relationships between rainfall and woody cover accounting for disturbance (fire frequency, cattle density, cultivation intensity and human population density) and soil nitrogen and texture. The results showed that a sigmoidal (non-linear) relationship best describes how woody cover responds to mean annual rainfall (MAP) and that this relationship represents a potential woody cover level in the absence of the disturbance factors (suppressors). With more and better continental data on factors that could create deviations from the MAP-driven woody cover line such as browsers and grazers densities, rainfall variability, fire intensity, cultivation types, we would expect the MAP-driven sigmoidal curve to probably approach an upper envelope line. However, the envelope line for our data may not only be the result of MAP in absence of perturbations. Woody cover could exhibit higher values than the MAP-driven potential cover at local/regional scales in response to promoting factors. For example, our model found that agriculture, most likely in the form of agroforestry or plantations, can lead to regionally high woody cover values not expected by the average climate conditions (and elimination of fire and herbivory disturbance).

Our results were harder to interpret in relation to soil nitrogen, also found as a promoting factor, and to soil texture. In arid and semi-arid areas (MAP < 500 mm), when

the soils get sandier (coarser texture), woody cover tends to increase while nutrient availability decreases (Walker and Langridge 1997). Our model did not distinguish nitrogen content effect by soil texture type. The positive effect of nitrogen was generally small for most of the points. The model indicates that fine textured soils support higher woody cover than coarse textured soils across the whole MAP gradient. However, several studies have found that there should be a soil texture inversion going from arid (MAP < 500 mm) to humid climates (Noy-Meir 1973, Walker and Langridge 1997); in more arid climates, coarser textured soils generally support taller and denser vegetation than finer soils for the same amount of rain. The opposite happens in humid climates. Our model shows agreement with the textural observations in the humid sites but it does not show the hypothesized inverse soil texture effect moving to the arid sites. We could not think of any biological data or statistical reason to explain these contradictory findings.

An important contribution of this work is the relative quantification of disturbance effects taking into account their prominent dependence on climate regimes. For example, fire frequency increases with MAP until about 1500 mm and its effect is generally negative on woody cover with increasing suppressive intensity going from arid to mesic savannas. Cattle density instead has a relatively high suppressing effect only in semi-arid and mesic savannas where pastoral activity is concentrated.

The model was also an important contribution towards the effort to integrate resource competition- and perturbation-based theoretical frameworks. It gave empirical support for the integrated model proposed by Sankaran et al. (2004, 2005): in stable savannas, tree-grass coexistence occurs because woody vegetation is driven and limited by MAP whereas, in unstable savannas, MAP does not limit woody cover that could

hence potentially outcompete grass but it is suppressed by perturbations. Tree-on-tree competition is another mechanism that could limit woody cover at medium to high cover levels.

The regional-scale ecological/environmental model clearly showed that, when we zoom in, we lose direct sense of the continental MAP driving trend and we see that finer woody cover patterns emerge in relation to other possible determinants. In our KNP case study, the basalt and granite underlying rocks explained the broad trend in woody cover. Smaller scale patterns on each rock type were related to perturbations such as fire frequency and herbivory in interplay with rainfall regimes and topographic features. For example, the lowest woody cover areas in the Park are located in the open savannas on the basalt rocks in association with fires recurring more often than once every five years, flat areas and relatively lower MAP. This work also emphasized that the variability in woody cover explained by the model had a strong ecological/environmental origin. Basically all the spatial structure of woody cover explained by the model derived from the spatial structure of the ecological/environmental determinants.

## **5.2 Woody cover predictive models and mapping**

The mapping work was essentially motivated by the need to obtain more accurate woody cover estimates for savanna systems than the ones provided by existing products and to use them to improve ecological inference. These products will also be made available in digital form for additional research (e.g. managers and scientists in KNP are using the KNP woody cover map for modeling habitat selection by elephants (de Knegt et al. 2010), net primary production, effects of climate change on vegetation and

biodiversity characterization, and the continental map will be of interest in the study of woody resource dynamics and carbon research).

My work has demonstrated the potential of combined SAR and optical remote-sensing for prediction of woody cover through scaling-up plot- and high-resolution-based measurements to regional and continental extents. The advantage of using combined optical and radar data comes from a wider set of spectral information available to distinguish woody from herbaceous vegetation and the background. In particular, optical remote-sensing is sensitive to photosynthetically active vegetation but does not provide sufficient discrimination of woody and herbaceous vegetation. This puts serious limits in woody plant detection in savanna systems. Microwave data, on the other hand, are sensitive both to woody branches and leaf moisture content and their backscatter provides distinctive information on woody vegetation.

The approach was successful at both scales and multiple linear regression appeared appropriate for the task producing models that explained 61 % of the variability in Kruger National Park and 58% at continental level. At the regional level, we worked at 90-m resolution with imagery original resolutions of 25 m from JERS-1 and 28.5 m from Landsat ETM+. At the continental scale we used remote-sensing imagery with resolutions of 2.25 km (QuickScatterometer) and 500 m (MODIS). The output map for KNP captured woody cover patterns recognizable by local experts to the scale of ~ 100 m and the absolute values of woody cover in the more mesic areas were in agreement with LiDAR assessments. The continental map also reproduced continental to landscape patterns consistent with average cover values from independent regional to landscape

sources. In particular, it improved predictions for arid and semi-arid systems that show up to 20% cover values.

A very important finding from these two mapping studies is that the strong predictive ability of radar backscatter intensity and optical greenness scaled-up from regional to continental scale. Although the up-scaling involved different sensors and predictive metrics (radar L-band vs. Ku-band for KNP vs. continental map and single-date vs. annual metrics for the KNP vs. continental map, respectively) and differing environmental and vegetation characteristics, our analyses show that the significant positive correlation of cover with radar backscatter and optical greenness found in the KNP savannas was transferable across scale and ecosystems, accounting for some non-linearity. We recognized that radar sensitivity to surface roughness was a common problem in both these maps. It created significant cover overestimates especially evident in bare rocky and rugged areas. This problem was partially corrected in the KNP woody cover map but not in the continental map. Terrain corrections for radar imagery usually require fairly complex geometric models that account for both radar incidence angle and terrain slope. We will probably use a more heuristic approach in which we will identify problematic areas by high backscatter, terrain roughness (high elevation variations) and low NDVI variations. We could then assign corrected radar backscatter values from neighbor pixels.

### **5.3 Conclusions**

Following the progress in savannas ecology theoretical and experimental work and in remote-sensing, my work generated empirical evidence for a savanna unified theory including resource- and disturbance-based mechanisms, it elucidated some scale

dependencies and tested a new approach for woody cover mapping with combined radar and optical imagery that has enhanced sensitivity to woody vegetation. The impact of the new woody cover maps is significant for savanna research that uses woody cover as explanatory or input variable such as in modeling work at different scales to predict vegetation distribution (Dynamic Global Vegetation Model, DGVM), climate (Global Climate Model, GCM), carbon stocks and sequestration, biogeochemical cycles, animal habitat, population dynamics, fire spread and other ecological processes and quantities that can inform both scientists and managers.

## 5.4 References

- de Knegt, H. J., F. van Langevelde, A. K. Skidmore, A. Delsink, R. Slotow, S. Henley, G. Bucini, W. F. de Boer, M. B. Coughenour, C. C. Grant, I. M. A. Heitkönig, M. Henley, E. M. Knox, N. M. Kohi, E. Mwakiwa, B. Page, M. Peel, Y. Pretorius, S. E. van Wieren, and H. H. T. Prins. 2010. The spatial scaling of habitat selection by African elephants. *Journal of Animal Ecology*. (*accepted for publication*)
- Hansen, M. C., R. DeFries, J. R. G. Townshend, M. Carroll, C. Dimiceli, and R. Sohlberg. 2006. Vegetation Continuous Fields MOD44B, 2001 Percent Tree Cover, Collection 4. University of Maryland, College Park, Maryland, 2001.
- Noy-Meir, I. 1973. Desert ecosystems: environment and producers. *Annual Review of Ecology and Systematics* 4:25-51.
- Sankaran, M., N. P. Hanan, R. J. Scholes, J. Ratnam, D. J. Augustine, B. S. Cade, J. Gignoux, S. I. Higgins, X. Le Roux, F. Ludwig, J. Ardo, F. Banyikwa, A. Bronn, G. Bucini, K. K. Caylor, M. B. Coughenour, A. Diouf, W. Ekaya, C. J. Feral, E. C. February, P. G. H. Frost, P. Hiernaux, H. Hrabar, K. L. Metzger, H. H. T. Prins, S. Ringrose, W. Sea, J. Tews, J. Worden, and N. Zambatis. 2005. Determinants of woody cover in African savannas. *Nature* 438:846-849.
- Sankaran, M., J. Ratnam, and N. P. Hanan. 2004. Tree-grass coexistence in savannas revisited - insights from an examination of assumptions and mechanisms invoked in existing models. *Ecology Letters* 7:480-490.
- Walker, B. H., and J. L. Langridge. 1997. Predicting savanna vegetation structure on the basis of plant available moisture (PAM) and plant available nutrients (PAN): a case study from Australia. *Journal of Biogeography* 24:813-825.