

EVALUATING THE ECOGEOGRAPHICAL EFFECTS OF EARTH'S LARGEST
TERRESTRIAL HERBIVORE

By

Ryan Lee Nagelkirk

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Geography—Doctor of Philosophy

2020

ABSTRACT

EVALUATING THE ECOGEOGRAPHICAL EFFECTS OF EARTH'S LARGEST TERRESTRIAL HERBIVORE

By

Ryan Lee Nagelkirk

Savannas cover a fifth of Earth's land surface, are home to over a half billion people, and disproportionately affect interannual variability of the global carbon cycle. In Africa, these open, grassy and sparsely wooded habitats support pastoralist cultures, the world's largest array of megafauna and thriving tourist economies. However, savannas and their uses are under threat: woody plant encroachment linked to increases in atmospheric CO₂ concentrations is reducing grassy cover required by both domestic livestock and wildlife, while also encroaching on the open views of wildlife critical to tourism. Yet, understanding of the determinants of savanna woody cover (bushes and trees) is limited. To this end, a growing number of site-specific studies have found that tree mortality rates in protected areas are principally controlled by African savanna elephants (*Loxodonta africana*). However, it is not known whether these impacts significantly affect the total woody cover of the larger landscapes and region.

This dissertation focuses on quantifying the relationship between elephant densities and savanna woody cover in protected areas across elephants' Eastern African range. Research questions are addressed in three self-contained chapters. Chapter 1 tests multiple approaches to mapping savanna woody cover fractions across 12 protected areas (PAs) using Landsat imagery, and presents a novel approach to reference data generation. The results show a machine learning approach, Random Forests, produces the most accurate maps and demonstrates that accurate maps

do not require more than a single annual image – which is advantageous given the general image scarcity in these areas.

In Chapter 2, the most accurate mapping approach from Chapter 1 is used to produce over 30 years of savanna woody cover data. These data are then used to assess whether there is a relationship between woody cover and elephant densities across the 12 PAs, as well as for specific landscapes within the PAs. Results point toward climate, principally wet season precipitation, being the major determinant of woody cover across the PAs ($R^2 = 0.38$, $p = 0.03$), though elephants were linked to increased woody cover on hill slopes far from permanent water bodies ($R^2 = 0.41$, $p = 0.03$). In addition, areas near water contain consistently low levels of woody cover unexplained by any of the variables considered. Last, Chapter 3 presents a meta-analysis of the literature comparing the relative impacts of elephants and fire on woody cover. The majority of studies (80.3%) find elephants to be the primary disturbance, with the relative dominance of the two disturbances linked to climate. The coolest and wettest savannas are more likely to be dominated by fire, while elephants are most likely to dominate across a comparatively broad set of environmental conditions.

Overall, the evidence presented here suggests (1) both overall woody cover and the relative impacts of elephants and fire are principally regulated by climate; (2) elephants, perhaps through the dispersal of seeds and nutrients, increase woody cover on hillslopes far from permanent water bodies; (3) areas near water are in a long-term state of low woody cover, potentially driven by disturbances, and (4) given the dominant role of elephants as a disturbance in the majority of sites and climates, conservationists should consider increasing elephant populations as a means of mitigating the woody encroachment threatening Africa's savannas, wildlife and pastoralist cultures.

Copyright by
RYAN LEE NAGELKIRK
2020

This thesis is dedicated to my elementary school teachers:
Ms. Willis, Ms. Bohs, Ms. Linhard and Ms. Merriman.
I have always looked up to you.

ACKNOWLEDGMENTS

First, I want to acknowledge the Michigan State University (MSU) Graduate School Fellowship that funded my degree, granted to me by the prior Dean of the Graduate School, Dr. Karen Klomparens. Without Dr. Klomparens' actions, which came at a time when I was at an academic crossroads, none of this would have been possible. Similarly, I want to thank Dr. Jennifer Olson who, before and after I found an academic department to call home, opened her door to me and was my mentor about all things Africa. Along with Dr. Olson, Drs. Nathan Moore and Gary Roloff formed a guidance committee that was responsive, challenging and fun. Thank you. Finally, I want to thank Dr. Kyla Dahlin, my advisor, who played the biggest role in all of this, welcoming me into her lab and helping me chart my own course. Dr. Dahlin's tactful mentorship and guidance always seemed to lead me to the right place and her example will be hard to live up to.

This thesis project also involved many hours of assistance from Logan Brissette, who tackled the unenviable challenge of creating reference data for the maps developed in Chapter 1, along with reviewing the majority of the literature used in Chapter 3. I also want to thank Dr. Catherine Yansa, whose course inspired what would become Chapter 3, and Dr. Elikana Kalumanga, who grounded me in the research of African elephants.

This dissertation was also supported by an MSU Graduate School Dissertation Completion Fellowship and US National Science Foundation (NSF) Award No. 1900108, the latter of which funded fieldwork that has been delayed by COVID-19.

Last, I want to thank my parents and grandparents for their years of continual support, inspiration and involvement. You made it possible for me to take a road less traveled and it has, indeed, made all the difference.

TABLE OF CONTENTS

LIST OF TABLES	ix
LIST OF FIGURES	x
KEY TO ABBREVIATIONS	xiii
INTRODUCTION	1
Research Context.....	1
The Impacts of Megafauna: A Changing Perspective.....	2
Dissertation Focus and Organization	4
CHAPTER 1. WOODY COVER FRACTIONS IN AFRICAN SAVANNAS FROM LANDSAT AND HIGH-RESOLUTION IMAGERY	6
Introduction	6
Materials and Methods	9
Study Sites	9
Reference Data	11
Landsat Image Collection and Processing.....	14
Mapping Woody Cover	19
Linear Regression.....	20
Spectral Unmixing.....	21
Regression Trees.....	25
Accuracy Assessment.....	26
Post-processing.....	28
Results and Discussion.....	29
Evaluation of accuracy measures	29
Best approaches and sub-approaches	29
Evaluation of seasonal images.....	32
Evaluation of protected area accuracies	34
Evaluation of variables	34
Best models.....	35
Map comparisons.....	39
Caveats & Concerns	42
Conclusions	43
CHAPTER 2. PRECIPITATION, NOT ELEPHANTS, BEST EXPLAINS SAVANNA WOODY COVER ACROSS EASTERN AFRICAN PROTECTED AREAS	44
Introduction	44
Materials and Methods	47
Study sites.....	47
Data sources.....	48
Woody Cover	48
Elephant Density	51

Precipitation	51
Temperature	52
Drought	52
Fire	53
Human Population Density	53
Latitude	54
Elevation	54
Data processing.....	54
Analysis	56
Results	57
PA-wide woody cover	57
Analyzing woody cover based on slope and proximity to water.....	59
Heterogeneity.....	62
Discussion	64
Primary and secondary roles of precipitation and elephants	64
Effects on woody cover heterogeneity	67
Habitat degradation by elephants	68
Conclusion.....	69
CHAPTER 3. FIRE OR FORCE: A META-ANALYSIS OF THE RELATIVE IMPACTS OF FIRE AND ELEPHANTS ON WOODY VEGETATION IN AFRICAN SAVANNAS	70
Introduction	70
Methods	73
Literature search	73
Analysis	74
Results	77
Discussion	83
The dominance of elephants	83
Management Implications	85
Caveats and concerns.....	86
Conclusion.....	88
CONCLUSIONS	90
Summary and Opportunities for Future Research.....	90
Broader Impacts	92
APPENDICES	93
APPENDIX A: Chapter 1 Supplementary Material.....	94
APPENDIX B: Chapter 2 Supplementary Material.....	111
BIBLIOGRAPHY.....	129

LIST OF TABLES

Table 1. Protected Area Names, Abbreviations and Attributes	11
Table 2. Index Equations	14
Table 3. Names and Equations of Accuracy Measures.....	27
Table 4. Best Seasons, Approaches & Models as Measured by VECv (%).	38
Table 5. Protected Area Names and Attributes.....	48
Table 6. Descriptions of Variables Used in Analysis	75
Table 7. Results of Univariate and Bivariate Models Produced Using MLR.....	81
Table A.1. Variables input to random forests and regression models	95
Table B.1. Ranked accuracies of models listed by the individual and composite images used to produce woody cover maps	112
Table B.2. Seasonal imagery available for woody cover mapping.....	113
Table B.3. Pearson's correlation matrix of independent variables used in study	117

LIST OF FIGURES

Figure 1. Conceptual Model of the Controls of Woody Cover in African Savannas	2
Figure 2. Areas of the World with Woody Cover Unexplained by Temperature and Precipitation	3
Figure 3. The Twelve Protected Areas Used in this Study	10
Figure 4. Example of Reference Data Generation	12
Figure 5. Conceptual Diagram of Methodology	17
Figure 6. Approach and Sub-Approach Accuracies	31
Figure 7. Heat Map of Models that Performed Best Across PAs, Seasons and Approaches	33
Figure 8. Our Best Maps and VCF Compared to Testing Data	40
Figure 9. Comparison of the Best Woody Cover Maps, VCF and VHR Satellite Imagery	41
Figure 10. The Twelve Protected Areas Used in this Study	47
Figure 11. Woody Cover Fractions and Elephant Densities.....	56
Figure 12. Relationship between PA Mean Woody Cover and a Subset of Variables.....	59
Figure 13. Relationships in the Four Within-PA Areas Analyzed.....	60
Figure 14. Relationships between Hill Slopes and Flat Areas.....	62
Figure 15. Relationships with Woody Cover Heterogeneity Near and Far from Water.....	64
Figure 16. Conceptual Diagram of Changes in Woody Cover in Relation to Elephant Densities and Precipitation	67
Figure 17. Maps of the Studies Comparing Elephant and Fire Disturbances.....	78
Figure 18. Boxplots of the Variables Used in Analysis by Disturbance Class.....	79
Figure 19. Probability Surfaces Generated in Multinomial Logistic Regression	82
Figure A.1. Ranked map accuracies	98

Figure A.2. Relationship across seasons between best PA map accuracies and PA average woody cover, mean annual precipitation (MAP) and precipitation seasonality	99
Figure A.3. Relationship between best PA map accuracies and PA average woody cover, mean annual precipitation, and precipitation seasonality	100
Figure A.4. Accuracy of maps produced using each of 143 variables in regression.....	101
Figure A.5. The best map produced for Chobe National Park.....	102
Figure A.6. The best maps produced for Kruger and Limpopo National Parks	103
Figure A.7. The best map produced for Mpala Research Center.....	104
Figure A.8. The best maps produced for North and South Luangwa National Parks.....	105
Figure A.9. The best map produced for Queen Elizabeth National Park	106
Figure A.10. The best map produced for Ruaha National Park.....	107
Figure A.11. The best map produced for Selous Game Reserve.....	108
Figure A.12. The best map produced for the combined Serengeti National Park and Maasai Mara National Reserve	109
Figure A.13. The best map produced for Murchison Falls National Park.....	110
Figure B.1. Woody cover fraction (black) and elephant density (red) data from the 12 PAs used in this study	114
Figure B.2. Temporal extents of the different data used in this study	115
Figure B.3. Relationship between PA mean woody cover and the 17 variables used in this study	116
Figure B.4. Relationship between mean woody cover in flat areas near water and the 17 variables used in this study	118
Figure B.5. Relationship between mean woody cover in hilly areas near water and the 17 variables used in this study	119
Figure B.6. Relationship between mean woody cover in flat areas far from water and the 17 variables used in this study	120
Figure B.7. Relationship between mean woody cover in hilly areas far from water and the 17 variables used in this study	121
Figure B.8. Relationship between woody cover differences between hills and flat areas near water and the 17 variables used in this study.....	122

Figure B.9. Relationship between woody cover differences between hills and flat areas far from water and the 17 variables used in this study.....	123
Figure B.10. Relationship between woody cover differences between flat areas far from water and those near water and the 17 variables used in this study.....	124
Figure B.11. Relationship between woody cover differences between hilly areas far from water and those near water and the 17 variables used in this study.....	125
Figure B.12. Relationships between all variables used in this study and whole-PA woody cover heterogeneity.....	126
Figure B.13. Relationships between all variables used in this study and woody cover heterogeneity in flat areas near water	127
Figure B.14. Relationships between all variables used in this study and woody cover heterogeneity in flat areas far from water	128

KEY TO ABBREVIATIONS

BC	Brightness Context
BurnFrac	Burned Fraction
Burn _{mult}	Fraction of Protected Area Burned Multiple Times per Year
Burn ₂₅	Number of Years >25% of Protected Area Burned
CHO	Chobe National Park, Botswana
DWP _{ann}	Annual Days with Precipitation
DWP _{dry}	Dry Season Days with Precipitation
DWP _{wet}	Wet Season Days with Precipitation
ELE	Elephant density
E ₁	Legates and McCabe's accuracy measure
GEE	Google Earth Engine
GRVI	Green-Red Vegetation Index
KRU	Kruger National Park, South Africa
LIM	Limpopo National Park, Mozambique
MAP	Mean Annual Precipitation
MAR	Maasai Mara National Reserve, Kenya
MAT	Mean Annual Temperature
MPA	Mpala Research Center, Kenya
MSAVI2	Modified Soil Adjusted Vegetation Index 2
MUR	Murchison Falls National Park, Uganda
NDVI	Normalized Difference Vegetation Index

NLU	North Luangwa National Park, Zambia
PA	Protected area
P _{dry}	Dry Season Precipitation
PDSI	Palmer Drought Severity Index
PET	Potential Evapotranspiration
POP	Human Population Density
P _{wet}	Wet Season Precipitation
QUE	Queen Elizabeth National Park, Uganda
RMSE	Root Mean Square Error
RUA	Ruaha National Park, Tanzania
R ²	Coefficient of Determination
SATVI	Soil Adjusted Total Vegetation Index
SEL	Selous Game Reserve, Tanzania
SER	Serengeti National Park, Tanzania
SLU	South Luangwa National Park, Zambia
SNDI	Soil Normalized Difference Index
SPEI	Standardized Precipitation-Evapotranspiration Index
T _{max}	Maximum Temperature
T _{min}	Minimum Temperature
VEcv	Variance Explained by Cross Validation

INTRODUCTION

Research Context

Savannas cover a fifth of Earth's land surface (Channan *et al.*, 2014), are home to over a half billion people (Nagelkirk and Dahlin, 2020), and disproportionately affect interannual variability of the global carbon cycle (Ahlström *et al.*, 2015). In Africa, the open, grassy and sparsely wooded habitats that define savannas (Scholes and Archer, 1997) support local pastoralists (Reid, 2012), burgeoning tourism industries (Balmford *et al.*, 2015; Naidoo *et al.*, 2016) and the world's largest array of megafauna (Malhi *et al.*, 2016). Over the past several centuries most savannas have been converted to agriculture, and the savannas that remain are under threat: woody plant encroachment linked to increases in atmospheric CO₂ concentrations (Stevens *et al.*, 2017, 2016) is reducing grassy cover required by both domestic livestock and wildlife (Smit and Prins, 2015), while also encroaching on the open views of wildlife critical to tourism (Gray and Bond, 2013). Yet, our understanding of the determinants of savanna woody cover is limited (Staver, 2018). To this end, researchers are finding that megaherbivores (>1,000 kg), whose numbers are falling across the continent (Ripple *et al.*, 2015), might play a significant role in regulating woody cover dynamics (e.g., Guldmond & van Aarde, 2008). In particular, a growing number of site-specific studies have found that tree mortality rates in protected areas are principally controlled by African savanna elephants (*Loxodonta africana*; Figure 1), who ring-bark, delimb and knock over trees up to 12 meters tall to consume the bark and leaves (Asner *et al.*, 2016; Malhi *et al.*, 2016; Shannon *et al.*, 2008). However, whether these impacts translate to the total woody cover of the larger landscapes and region is unknown.

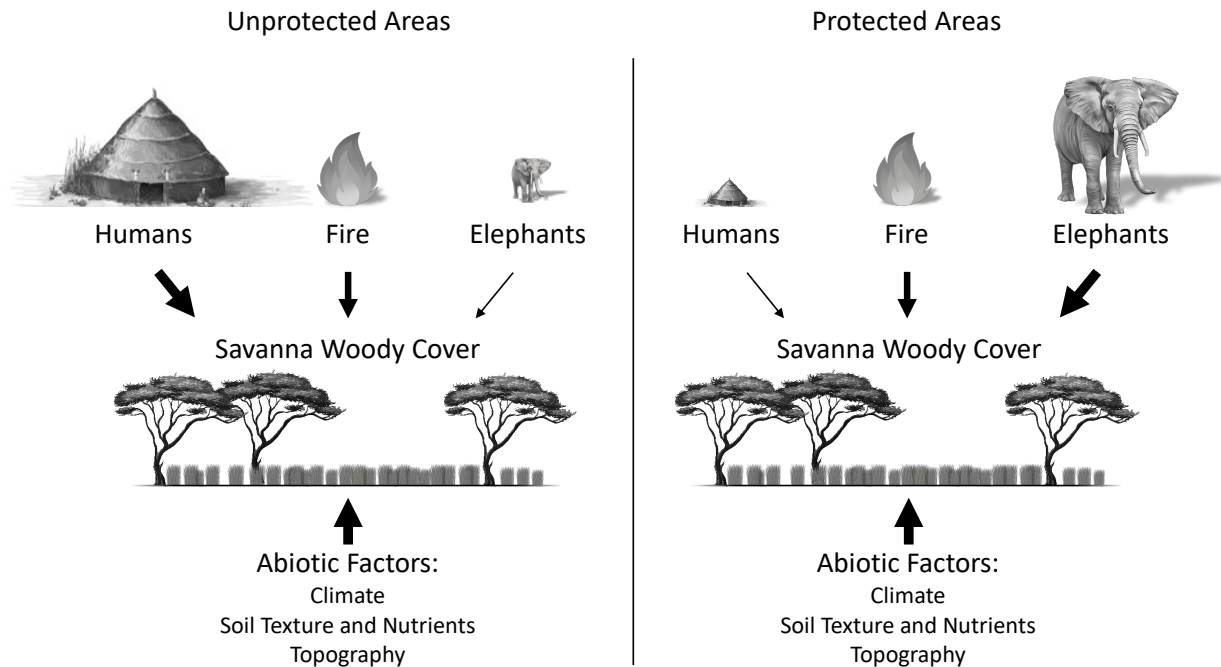


Figure 1. Conceptual Model of the Controls of Woody Cover in African Savannas. Abiotic factors provide the resources necessary for savannas and/or wooded landscapes, while humans, fire and elephants prevent establishment of woodlands, maintaining a savanna state. In unprotected areas (left), humans are the chief disturbance of woody cover (denoted by arrow size). In protected savannas, where human impacts are restricted, elephants dominate. In both, fire – whether set by humans or lightning – plays a secondary role.

The Impacts of Megafauna: A Changing Perspective

Understanding the geographic distribution of plants has been a focus of geographers since the days of von Humboldt and Bonpland’s explorations of South America (von Humboldt and Bonpland, 1807). While much of the focus of Humboldt’s work, and others, has been on the relationship between vegetation distributions and climate, recent work, particularly in savanna-type ecosystems, has offered a new focus. In 2005, Bond (2005) challenged the widely-accepted idea that climate alone determined vegetation distributions on land (Polis, 1999), arguing that herbivores and fire suppress woody biomass across large parts of the world, including savannas (Figure 2). In particular, Bond (2005) found it “puzzling that large mammals, especially the megafauna, have not been more central in the trophic control literature.”

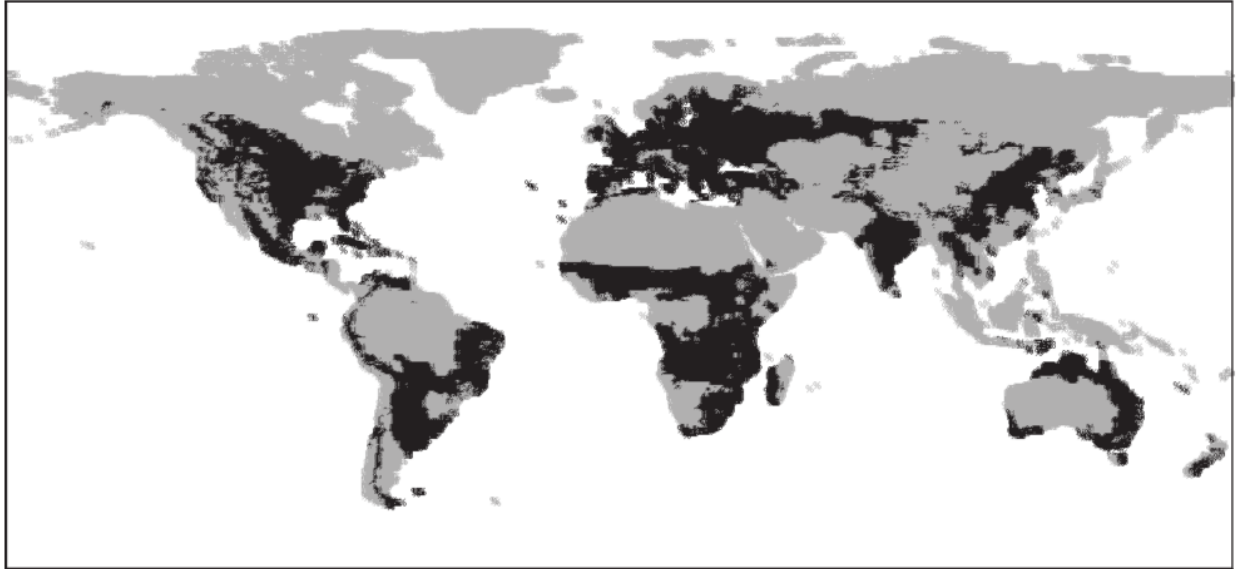


Figure 2. Areas of the World with Woody Cover Unexplained by Temperature and Precipitation. *Termed “ecosystems uncertain” by Whittaker (1975), the black areas represent those where disturbances such as herbivores and fire are thought to reduce woody cover below its potential levels. Figure from Bond (2005).*

Since Bond (2005), ecogeographic studies have demonstrated how megafauna alter regional vegetation distributions through nutrient dispersal (Doughty, 2017; Doughty et al., 2015, 2013; Wolf et al., 2013) and/or woody cover suppression (Bakker et al., 2016; Barnosky et al., 2015; Doughty et al., 2016; Jia et al., 2018; Malhi et al., 2016). Others have taken this work a step further, demonstrating how megafauna’s impacts might have played a significant role in regional-to-global climate (Doughty et al., 2010). However, studies of large-scale impacts tend to focus on *extinct* megafauna (Doughty *et al.*, 2010, 2013, 2015; Barnosky *et al.*, 2015; Doughty, 2017). Meanwhile, small-scale studies (Jia et al., 2018) and local extinctions (Daskin et al., 2016; Stevens et al., 2016) of extant species demonstrate that these species also suppress local vegetative biomass, but the large-scale consequences – particularly for woody cover – remain unclear (Staver, 2018). This dissertation attempts to quantify the relationship between Earth’s largest terrestrial megaherbivore – the African savanna elephant – and savanna woody cover in protected areas across Eastern Africa.

Dissertation Focus and Organization

The central objective of this dissertation is to establish where elephants rank with respect to other factors influencing woody cover (e.g., precipitation, humans and fire) through the compilation and analysis of remote sensing data, elephant census counts and existing literature. Chapters 1 through 3 are self-contained studies respectively addressing three interrelated research questions:

- (1) What is an accurate approach to mapping savanna woody cover through the use of satellite imagery, and what amount and timing of imagery best serves this purpose?
- (2) Across African protected savannas, what is the relationship between elephant densities and woody cover?
- (3) Of the two major disturbances in protected savannas – elephants and fire – which is the primary disturbance related to woody cover?

Chapter 1 addresses a key deficiency in savanna research to this point: the lack of accurate maps of savanna woody cover. While there are instances of individual protected areas being mapped, the data are rarely comparable across sites due to the use of different units describing woody cover. Chapter 1 develops a semi-automated approach capable of accurately mapping woody cover fractions at 30-meter resolution across protected areas. The chapter also outlines the development of a particularly helpful index describing an area's brightness in the context of its surroundings; the importance of imagery captured between the wet and dry season, during the senescence of grasses; and the development of a new method to quickly map reference data using red-green-blue (RGB) imagery.

Chapter 2 uses the approach developed in Chapter 1 to map woody cover from the years 1984 to 2016 in 12 protected areas. The data is then used to analyze the relationship between average

levels of woody cover and elephant densities, along with climate, fire and human population densities. The chapter assesses woody cover at the level of the protected areas, along with areas within the protected areas related to their proximity to permanent water sources and hill topography. At all levels, climate variables have the strongest relationships with woody cover. However, unexplainedly low woody cover near permanent water sources, along with unexpected positive relationships between elephant densities and woody cover on hill slopes far from permanent water sources, indicate elephants might be affecting woody cover across major portions of the landscape.

Chapter 3 analyses the existing literature in a meta-analysis seeking to establish whether, amongst each other, elephants or fire are the primary disturbance of woody cover. The results reveal elephants are the primary disturbance across a broad swath of environmental conditions, with fire dominating in wetter and cooler areas. The results also reveal a regional bias in the literature, pointing toward deficiencies in central and western Africa, as well as an overall lack of research objectively comparing the impacts of elephants and fire. Overall, these findings imply that woody encroachment – currently a primary concern in savanna management – might be partially addressed by increasing elephant numbers, while also suggesting that the wide-scale loss of elephant populations partially enabled woody encroachment.

CHAPTER 1. WOODY COVER FRACTIONS IN AFRICAN SAVANNAS FROM LANDSAT AND HIGH-RESOLUTION IMAGERY

Citation: Nagelkirk, R.L. & Dahlin, K.M. (2020). Woody cover fractions in African savannas from Landsat and high-resolution imagery. *Remote Sensing*. DOI: 10.3390/rs12050813.

Introduction

Savannas cover a fifth of Earth's land surface (Channan et al., 2014) and are home to over 500 million people (estimate derived by intersecting The Nature Conservancy's savanna and shrubland ecoregions (Olson et al., 2001) with population estimates for 2015 (CIESIN, 2016)). Defined as mixed tree-grass communities (Scholes and Archer, 1997), savannas support pastoralist cultures (Reid, 2012), the world's largest, functionally complete set of terrestrial megafauna (Malhi et al., 2016) and thriving tourist economies (Balmford et al., 2015; Naidoo et al., 2016). Savannas also play a disproportionate role in the interannual variability of the global carbon cycle (Ahlström et al., 2015; Poulter et al., 2014). In all of these, the woody cover of savannas plays an important role: pastoralism, tourism and grazing wildlife all rely on sparse woody cover and are threatened by woody encroachment (Gray and Bond, 2013; Reid, 2012; Smit and Prins, 2015), while savanna carbon dynamics are disproportionately affected by woody vegetation (Belsky et al., 1989; Scholes and Archer, 1997). Yet, our understanding of the factors controlling savanna woody cover, in comparison to other biomes (i.e. grasslands and forests), is relatively limited.

Unlike the climate-determined woody cover of forests and grasslands, savanna woody cover exists in a climate-indeterminate state (Bond, 2005). In savannas, disturbances work individually or synergistically to maintain low woody cover in areas that might otherwise transition to forest or grassland (Sankaran et al., 2008, 2005), a characteristic termed "multiple stable states" (Holling, 1973; May, 1976). Chief among the disturbances are fire (Bond and Keeley, 2005; Hantson et al., 2017; Smit et al., 2010), drought (Good and Caylor, 2011; Porensky et al., 2013b; Van Der Waal

et al., 2009) and herbivory (Asner et al., 2016; Holdo et al., 2009b; Staver and Bond, 2014; Traore et al., 2015). However, despite efforts to understand how disturbances interact with each other and climate across broad spatial scales (Lehmann et al., 2014; Sankaran et al., 2008), our ability to predict savanna woody cover to any significant degree is still limited (Staver, 2018).

Part of the challenge of understanding savannas has likely come from the data used in prior studies. Several influential studies of savannas and/or multiple stable states (Favier et al., 2012; Hirota et al., 2011; Murphy and Bowman, 2012; Ratajczak and Nippert, 2012; Scheffer et al., 2012; Staver et al., 2011) centered on analyses of Vegetation Continuous Fields (VCF) global tree cover data (Hansen et al., 2003) – data now known to have significant inaccuracies in savannas (Hanan et al., 2015, 2014; Staver and Hansen, 2015). Consequently, the producers of the dataset cautioned against the use of VCF in areas with less than 20-30% tree cover (Staver and Hansen, 2015), effectively ruling out savannas given their characteristic ~20% mean VCF tree cover (Hirota et al., 2011). Additionally, VCF, which was primarily developed to monitor forests, not savannas, defines trees as woody vegetation > 5 meters tall (Hansen et al., 2003). While a fair threshold for forests, the majority of savanna woody vegetation likely falls below it (Levick et al., 2009), thus under-representing the woody component of these systems. More importantly, disturbances predominately act upon woody vegetation recruits (< ~5-6 meters) (Asner et al., 2016, 2009b; Asner and Levick, 2012; Levick et al., 2009; Levick and Asner, 2013; Smit et al., 2010), meaning VCF is unlikely to detect the direct impacts of disturbances, i.e. the major determinants of woody cover. Combined, these factors make the use of VCF in any savanna problematic, potentially undermining our ability to understand these ecosystems.

In lieu of VCF, researchers often make maps of their own. However, making large-scale maps of savanna woody cover is a challenge because, unlike in forests, savanna imagery contains a high

proportion of pixels with a combination of woody vegetation, herbaceous vegetation and bare soil – referred to as mixed pixels (Lawton and Sylvestre, 1971; Settle and Drake, 1993). Unmixing pixels requires knowledge of the spectral characteristics of all the materials within the pixel. Meeting that requirement across large areas has been a long-standing challenge (Choodarathnakara et al., 2012) because vegetation and soil spectral properties change across space and time, particularly in savannas (Dawelbait and Morari, 2008; Hansen et al., 2003; Ringrose et al., 1989; Staver and Hansen, 2015).

To limit the spectral variability of cover types, researchers often produce small-scale, site-specific maps using a range of approaches, including simple linear regressions (Poitras et al., 2018), regression trees (Yang and Crews, 2019), cluster analysis (Marston et al., 2017) and spectral unmixing (Asner and Lobell, 2000). Meanwhile, some of those who have attempted to produce maps across larger areas or across several sites abandoned such approaches, instead manually classifying a high number small areas (~0.5 hectare) using very high-resolution (VHR; < 1 meter) imagery (Bastin et al., 2017; Messina et al., 2018) or, in an effort to lower VCF error, resampling the data to coarser resolutions, thereby abandoning fine scale analysis (Favier et al., 2012). Still others use commercially produced national landcover maps (Skowno et al., 2017). The different data sources and methodologies mean maps are rarely easily comparable across studies. That, combined with an overall shortage of these maps, limits large-scale studies of savanna dynamics, while also limiting our ability to compare the approaches used to generate them. Further, studies rarely map woody cover across time, despite the demonstrated insights from temporal data (Ratajczak and Nippert, 2012; Ward et al., 2014; Western and Maitumo, 2004).

Here, we develop an accurate, replicable method for mapping total woody cover across 12 protected areas in eastern Africa, from Uganda to South Africa, using Landsat 8 imagery. While

we recognize separate maps of shrub cover and tree cover would be ideal when studying savannas and could be possible using ancillary data (e.g. LiDAR-derived digital surface models), we sought a method that could be applied through time, back to the launch of Landsat 4 (1982) – a span that extends beyond most ancillary data. Further, we chose to not use ancillary data that could be assumed to be constant across the temporal record (e.g. topography) to avoid circularity in future studies of landscape-scale controls on savanna distributions and woody cover. Last, given the marked decline in Landsat image availability going back to the 1990s and 1980s, especially in rural Africa (Wulder et al., 2016), we limited ourselves to three images per site: one each from the wet and dry seasons, along with the transition between them. The result was the development of 30-meter resolution maps of woody cover across all the sites; site-specific rankings of seasonal imagery; a novel approach for reference data generation; and a clear designation of the best mapping approach.

Materials and Methods

Study Sites

We mapped woody cover in 12 sites across eastern and southern Africa (Figure 3; Table 1). All of our sites are protected areas (International Union for Conservation of Nature protected status II-IV). We selected protected areas (PAs) because, as part of a larger project studying the drivers of savanna processes, we sought to avoid modern anthropogenic impacts to the extent possible, though we recognize that humans have long played a role in savannas (Archibald et al., 2013, 2012; Bowman et al., 2011). The PAs cover a broad range of mean annual precipitation (MAP) (~500-1250 mm) and size (~200-44,800 km²). Initial inspection using satellite imagery suggested the PAs also cover a broad range of woody cover, with wetter PAs appearing significantly woodier than drier PAs.

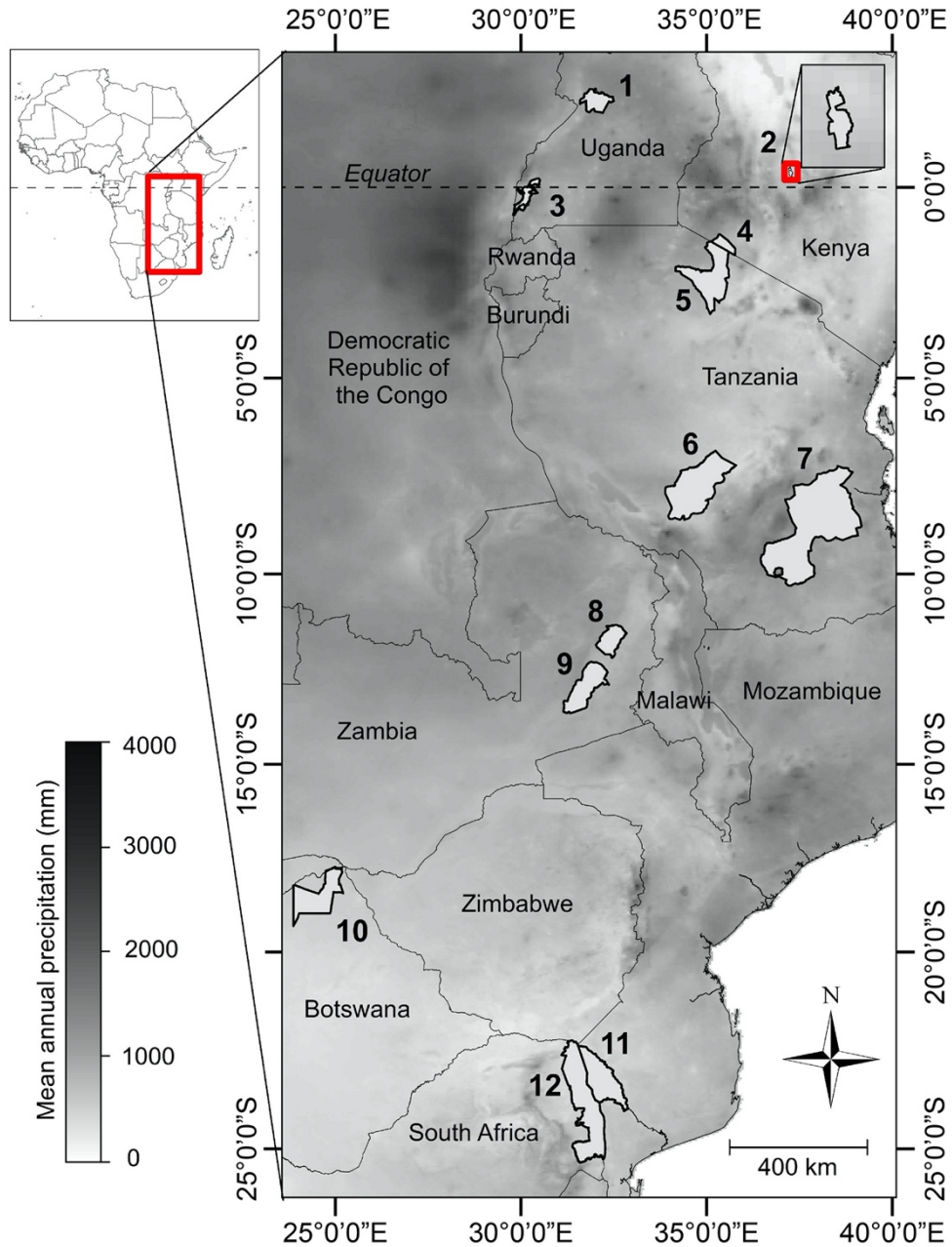


Figure 3. The Twelve Protected Areas Used in this Study. Protected areas are shown in gray. Red boxes denote inset map borders. The background represents the mean annual precipitation from 1988-2017 (Climate Hazards Infrared Precipitation with Stations (CHIRPS); Funk et al., 2015). See Table 1 for corresponding protected area names and attributes. Mpala Research Center (2) is ~11 km at its widest point (no scale in inset).

Table 1. Protected Area Names, Abbreviations and Attributes. PA numbering corresponds to Figure 3.

PA #	Name & Country	Abbr.	Latitude (degrees)	Elev. (m)	MAP (mm)	Area (km ²)
1	Murchison Falls National Park, Uganda	MUR	2.27	846	1262	3877
2	Mpala Research Center, Kenya	MPA	0.40	1694	601	194
3	Queen Elizabeth National Park, Uganda	QUE	-0.25	977	998	7395
4	Maasai Mara National Reserve, Kenya	MAR	-1.50	1624	950	1510
5	Serengeti National Park, Tanzania	SER	-2.27	1546	850	14763
6	Ruaha National Park, Tanzania	RUA	-7.80	1168	700	20226
7	Selous Game Reserve, Tanzania	SEL	-8.86	396	1121	44800
8	North Luangwa National Park, Zambia	NLU	-11.88	752	904	4636
9	South Luangwa National Park, Zambia	SLU	-13.09	623	917	9050
10	Chobe National Park, Botswana	CHO	-18.56	968	532	11000
11	Limpopo National Park, Mozambique	LIM	-23.32	246	534	10000
12	Kruger National Park, South Africa	KRU	-23.93	342	511	19175

Reference Data

To train our approaches and assess the accuracy of the resulting maps, we generated woody cover reference data using VHR imagery available through Google Earth (Figure 4) – a practice that is increasingly common in broad-scale studies (Bastin et al., 2017; Hansen et al., 2013; Messina et al., 2018; Michishita et al., 2012; Pengra et al., 2015; Skowno et al., 2017). We note that in our maps and reference data, we mapped crown cover (the vertically projected area occupied by a tree crown), not canopy cover (crown cover minus intra-canopy skylight). We assume the globally derived relationship between canopy cover and crown cover (canopy cover = 0.8*crown cover) (Hansen et al., 2003) holds in our study sites, allowing us to convert when necessary (e.g., VCF uses canopy cover instead of crown cover).

Previous studies mapped reference point landcover using a range of techniques, such as visual estimation (Michishita et al., 2012; Skowno et al., 2017), decision tree algorithms (Pengra et al.,

2015) and augmented visual interpretation using software such as CollectEarth (Bastin et al., 2017; Bey et al., 2016; Messina et al., 2018). In particular, CollectEarth uses VHR imagery from Google Earth and Bing Maps to compute helpful metrics such as the Normalized Difference Vegetation Index (NDVI), then uses tens of user-classified sampling points within the reference image to estimate the spatial extent of each cover type.

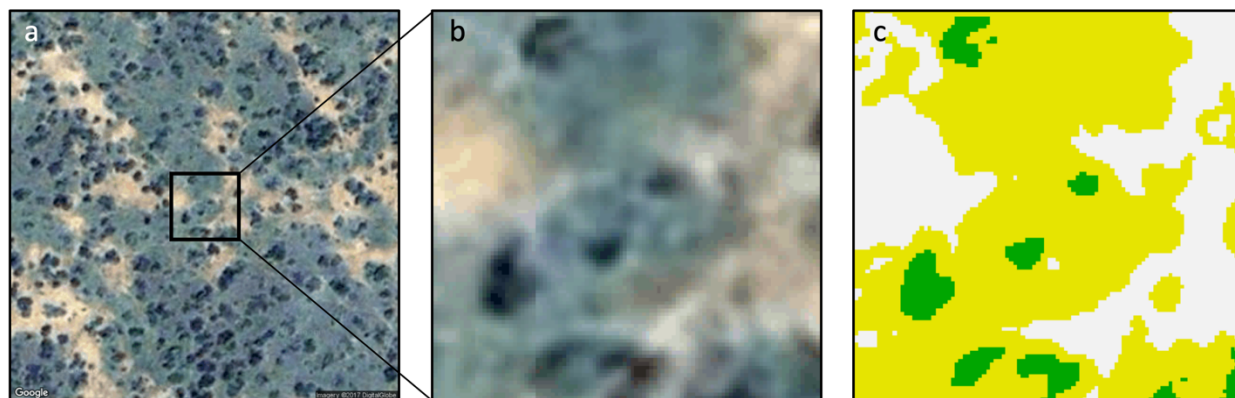


Figure 4. Example of Reference Data Generation. For each reference point, we downloaded a 180x180-meter scene from Google Earth imagery (a) centered on the 30x30-meter Landsat pixel (inset of a; b). We then mapped the woody cover (green), grass (yellow) and soil (white) of the pixel (c). The percent woody cover was then extracted to generate the reference data. All reference data is available online: (Nagelkirk and Dahlin, 2019).

Similar to CollectEarth, we developed a fast, semi-automated approach using VHR imagery from Google Earth. However, our approach mapped, rather than sampled, the complete extent of each cover type (code available in the online dataset: (Nagelkirk and Dahlin, 2019). We downloaded VHR true color imagery from Google Earth using the RgoogleMaps package (Loecher and Ropkins, 2015) in R (R Core Team, 2018), retrieving 180x180-meter areas centered on one 30x30-meter Landsat pixel (the reference point). We then mapped our perceived extent of woody cover, herbaceous cover (simply “grass”, hereafter) and soil cover of the 30x30-meter reference point by manually thresholding image-derived metrics.

In our mapping, we took advantage of the fact that woody cover is predominantly darker than both soil and grass, thresholding pixel brightness (the sum of RGB values) to classify woody cover. When necessary, we adjusted thresholds to account for shadows that would have otherwise inflated woody cover values. In addition, when visually distinguishing woody cover from grass was a challenge, we assumed that objects with shadows were woody cover. We also used any on-the-ground photographs available through Google Earth, along with our own on-the-ground experience in African savannas, to further inform our mapping. If we could not confidently distinguish a reference point's woody cover, or if brightness thresholding failed to do the same – both of which were common when defoliated woody cover was present – a new reference point in a different location was automatically generated.

After woody cover, the remaining grass and soil were classified using pixel brightness in conjunction with the Green-Red Vegetation Index (GRVI; Table 2) (Motohka et al., 2010), an index that exploits the red and green spectral differences between green vegetation and soil. However, we found that the GRVI threshold could often be left static (ca. -0.1), with only brightness threshold adjustment needed. We mapped a minimum of 100 reference points in each PA, increasing the number when densities fell below one point per 200 km², altogether mapping 1,330 reference points. All the points were in cloud-free positions in the Landsat imagery. We then randomly subset the data, splitting it into PA-specific training (70% of points) and testing data (30%).

Table 2. Index Equations. Color names refer to the corresponding image band.

Index	Equation*
Green-Red Vegetation Index (GRVI)	$\frac{Green - Red}{Green + Red}$
Normalized Difference Vegetation Index (NDVI)	$\frac{NIR - Red}{NIR + Red}$
Soil Normalized Difference Index (SNDI)	$\frac{Red - (NIR + SWIR1)}{Red + (NIR + SWIR1)}$
Soil Adjusted Total Vegetation Index (SATVI)	$\frac{SWIR1 - Red}{SWIR1 + Red + L} (1 + L) - \frac{SWIR2}{2}$
Modified Soil Adjusted Vegetation Index (MSAVI2)	$\frac{2NIR + 1 - \sqrt{(2NIR + 1)^2 - 8(NIR - Red)}}{2}$

* Spectra names (*Red*, *Green*, *NIR* etc.): Landsat reflectance bands. *L*: the soil adjustment factor – a constant set at 0.5 (Marsett et al., 2006; Qi et al., 1994).

Landsat Image Collection and Processing

We collected Landsat 8 Tier 1 Surface Reflectance imagery for 2016 using Google Earth Engine (GEE; Figure 5) (Gorelick et al., 2017). GEE utilizes the computational capabilities of Google to enable researchers to access and process the Landsat archive. We used GEE to select, preprocess and download imagery, as well as carry out a subset of our mapping approaches (described below). However, we limited our exposure to any potential change or loss of services from GEE, such as Google’s 2018 decision to phase out their “Fusion Table” data type (Google Fusion Tables Team, 2019), by carrying out most of our mapping outside of GEE. We selected relatively cloud-free imagery (< 30% cloud cover over PAs in wetter regions (SEL, SER, MAR, RUA, MUR, QUE); < 10% in drier regions (CHO, MPA, KRU, LIM, NLU, SLU) (Table 1)), and required all PA images within the same Landsat path to be from the same date. We then cloud masked the imagery and selected images we estimated to correspond with the wet, dry and transition seasons. For each PA, the wet season image (Wet) was characterized as that with the highest mean NDVI (Table 2) (Tucker, 1979); the dry season image (Dry) as that with the lowest

NDVI; and the transition (Tran) as that with its mean NDVI closest to the midpoint between the Wet and Dry NDVIs, falling chronologically after the wet season and before the dry season. If no image met our requirements for the Tran image, we took the image closest to the NDVI midpoint, regardless of where it fell in the year.

The purpose of using three images was three-fold. First, by using the images individually, we aimed to identify the time of year that led to the best maps of woody cover. Past studies attempting to map woody cover and/or biomass have used the dry season (Brandt et al., 2017; Marston et al., 2017), while others have used the transition (Bucini et al., 2009; Gizachew et al., 2016) – sometimes at the same sites (Bucini et al., 2009; Marston et al., 2017) – with no consensus on the most accurate approach. The dry season is often selected because many woody species remain green, enhancing the difference between their spectral signatures and that of brown, senesced herbaceous plants, thereby aiding mapping (De Bie et al., 1998; Horion et al., 2014; Wagenseil and Samimi, 2007). However, drier tropical and subtropical sites generally have more trees that drop all or a fraction of their leaves during the peak of the dry season (drought deciduous or raingreen) (Murphy and Lugo, 1986; Santiago et al., 2004). Therefore, we expected Tran images to outperform Dry images in the drier of our sites. Finally, we expected the Wet images, captured when both woody and herbaceous plants are photosynthetically active and therefore most spectrally similar, to yield the least accurate maps. The second reason we used three images was that, when stacked together into a single image or used to create summative statistics (see Section 2.2.3), we expected to increase map accuracies by capturing the phenological differences of woody and herbaceous vegetation, as done in similar work (Brandt et al., 2016; Gasparri et al., 2010; Hansen and Loveland, 2012; Skowno et al., 2017). Third, we limited ourselves to three images to simulate the shortage we would likely encounter when mapping historical imagery.

Because many of the PAs spanned multiple Landsat paths and each additional path added significant time investment, paths that covered less than 10% of a PA were excluded. Adjacent PAs in the same path with imagery from the same dates were treated as a single PA (requirements only met by SER and MAR, hereafter together referred to as ‘SMR’).

In all, 105 individual Landsat scenes were selected. Images in the same path were subsequently mosaicked, yielding 45 images. We then masked all the images within each PA using the same PA-specific mask. Each PAs mask was generated using the “pixel_qa” band provided with the Landsat data, and all but the pixels marked as “clear” were removed by the masks. The masks were combined across seasonal images, meaning masked areas in one image were masked in the others. This meant that sometimes large image fractions (approx. 15-20%) were removed, which is not ideal when creating a map. However, our objective was not to create contiguous images but was to produce images with identical data to test which seasons and approaches best predicted woody cover.

Finally, we clipped the images to 20-km buffers around each PA boundary, with the exception of the smallest PA, MPA, which we clipped to a 50-km buffer. The buffers allowed us to capture areas with 100% crown cover – areas often only available outside PAs and required in a subset of our approaches. All images were then downloaded from GEE (code available in the online dataset: Nagelkirk and Dahlin, 2019)

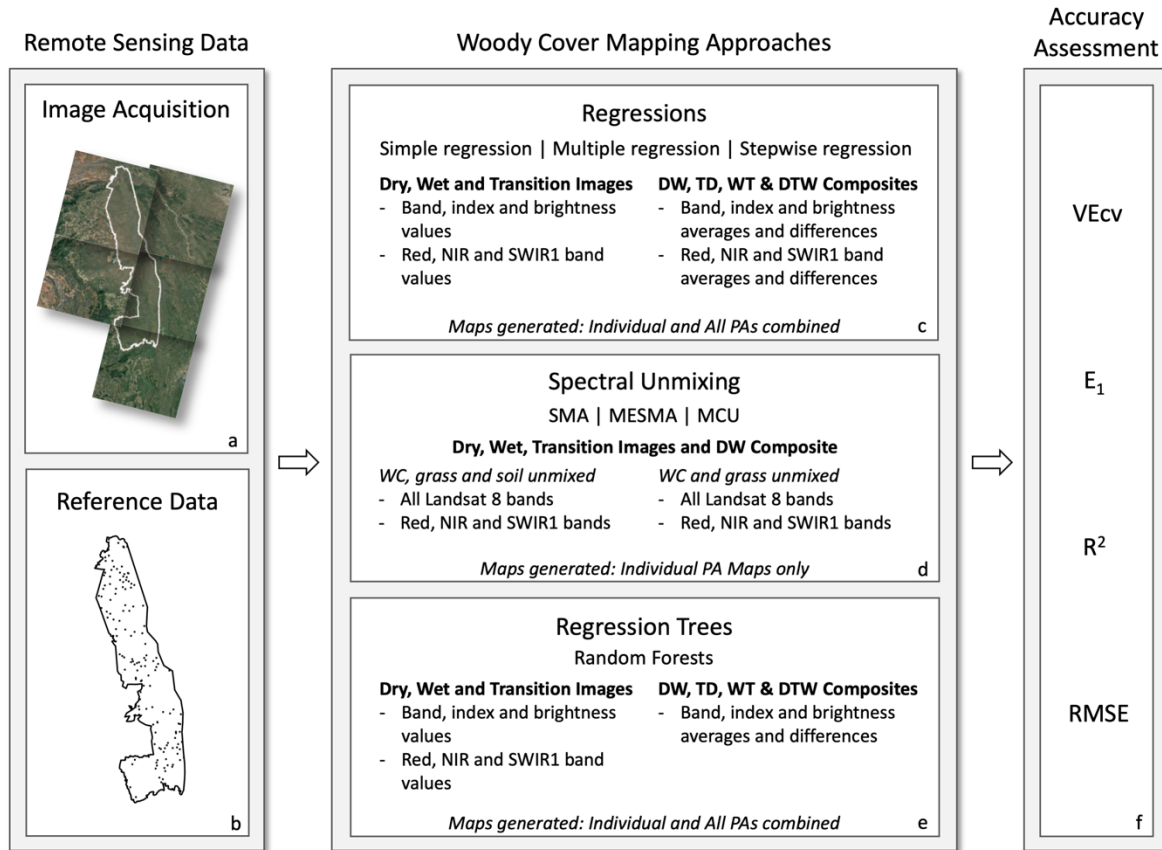


Figure 5. Conceptual Diagram of Methodology. For each PA, we collected, processed and then downloaded Landsat imagery from the dry, wet and transition seasons using GEE (a). We then extracted band and index values at reference points throughout each PA and used very high-resolution (VHR) imagery available through Google Earth to map the woody cover at each point (b). These data, which varied across single and combined season imagery, were used to create woody cover maps of the individual PAs using regression, spectral unmixing and regression trees (c-e). In addition, the reference data from all the PAs (ALL) were pooled to generate another series of Random Forests- and regression-derived maps. We assessed the accuracy of the maps using VEc_v, E_1 , R^2 and RMSE (f). SMA: Spectral Mixture Analysis; MESMA: Multiple Endmember Spectral Mixture Analysis; MCU: Monte Carlo Unmixing.

For the 70% of reference points serving as training data, we extracted Landsat reflectance values to serve as predictors of woody cover. For each point, we supplemented reflectance values with indices known or likely to have relationships with vegetation in semi-arid environments (Poitras et al., 2018): the Soil Normalized Difference Index (SNDI) (Poitras et al., 2018), Soil Adjusted Total Vegetation Index (SATVI) (Marsett et al., 2006), and the updated Modified Soil

Adjusted Vegetation Index (MSAVI2) (Qi et al., 1994) and NDVI (Tucker, 1979) (Table 2). We also computed the visual brightness (mean of red, green and blue bands; hereafter simply “brightness”) of each reference point. We chose brightness because of the strong relationship it had with woody cover during reference point generation, as mentioned in Section 2.2.2 above. We then created an additional variable to contextualize the brightness within the landscape because, when we were creating the reference data, views of the larger landscape helped the user distinguish between grass and woody cover. In particular, a completely wooded pixel was often only distinguishable from a grass-dominated pixel when we viewed the larger landscape and saw all the brighter, grass-dominated pixels. To generate the contextualized variable, we computed the normalized difference between reference point brightness and PA mean reference point brightness. We refer to this as the brightness context (BC; Table A.1).

In addition to the variables generated from single images, we created multi-image composite variables to incorporate phenology – a practice not novel to vegetation mapping (Hansen et al., 2003). The specific image composites were: Dry, Tran and Wet (DTW); Dry and Wet (DW); Tran and Dry (TD) and Wet and Tran (WT). The values for each composite’s variables were calculated by taking the mean and normalized difference of each single-image variable across images (testing showed the normalized difference had stronger relationships with woody cover than the simple difference). This doubled the number of variables in the composites compared to the single images. For example, while the Dry and Wet images each generated a single NDVI variable, the DW composite had two NDVI variables: the 1) mean and 2) normalized difference of the Dry and Wet NDVIs. However, the DTW composite presented a unique challenge: calculating the difference of three values (e.g., Dry NDVI, Wet NDVI and Tran NDVI). We therefore computed the DTW differences by subtracting the normalized differences of the WT composite from the DW (Table

A.1). This metric had a distinct phenological meaning. Assuming the date of the Tran image was between the Wet and Dry, a value of zero meant all changes in the variable occurred early (between the Wet and Tran images) – likely representative of the early senescence of a grass-dominated pixel. As more of the total change occurred after the Tran image (i.e. the late season senescence of woody cover), the value exponentially approached 1 or -1, depending on the direction of variable change moving from the wet to the dry season (e.g., for NDVI, which decreases across the seasons, the DTW difference metric approaches -1). For simplicity, we use “difference” to describe the difference metrics of all the composite images, even though all but DTW use the normalized difference.

Across the single images and composites, these steps generated 132 variables (Table A.1). Hereafter, we refer to these data collectively as the “reflectance data”.

Mapping Woody Cover

We mapped woody cover using seven techniques falling under three general approaches: linear regression, spectral unmixing, and regression trees (Figure 5). To incorporate their nested structure, we refer to these as “sub-approaches” and “approaches”, respectively. Because all the approaches were not available in a single software package, we worked across multiple programs. Further, not all approaches used all the possible variables when those variables did not improve results and/or caused inputs to exceed the capacity of most modern computer systems, thereby limiting their adoption (see *Spectral Unmixing*). The programs and approaches are described below, with detailed procedures and code provided in the online dataset (see Nagelkirk and Dahlin, 2019).

Linear Regression

We used linear regressions (Neter and Wasserman, 1974) to identify any consistent relationships between the reflectance data and woody cover. All regressions were carried out in R (R Core Team, 2018). We expected maps generated using linear regression would form baseline accuracies compared to the other, more complex, approaches.

We conducted regressions using three sub-approaches. In the first, all 132 reflectance data variables were regressed independently against woody cover (simple linear regression). In the second, we used the red, near infrared (NIR) and first short-wave infrared (SWIR1) bands together in multiple regression (for the composite images, we did separate regressions using the means (1) and differences (2) of the red, NIR and SWIR bands). We refer to these as the “RNS” bands/regressions. We chose the RNS bands because they capture the spectral signature of vegetation, hence their widespread use in vegetation indices (Marsett et al., 2006; Poitras et al., 2018; Qi et al., 1994; Tucker, 1979). Third, we conducted forward stepwise regressions (STEP) to identify novel relationships that might warrant further investigation. To avoid overfitting the models and multicollinearity between variables, we only added variables with high levels of significance within the new model (p-values below 0.05) and low collinearity ($R^2 < 0.6$).

We carried out the regressions in individual PAs to derive PA-specific relationships, then with all PAs combined (“ALL”) to derive regional relationships. Altogether, the simple, RNS and STEP regressions yielded 1,800 regression models. We applied the models to their respective PAs, then applied the ALL-derived, regional models to the individual PAs to quantify the local accuracy of the regional models.

Spectral Unmixing

Spectral unmixing, which some have shown outperforms linear regression in savanna woody cover mapping (Yang et al., 2012), is based on the premise that pixel reflectance is an area-weighted, linear combination of the landcovers within each pixel (Adams et al., 1986; Smith et al., 1985). Spectral unmixing, then, uses the reflectance values of each individual cover type – values referred to as “endmembers” – to unmix, or back-calculate, the fractional coverage of each land cover type within a pixel. A unique feature of this approach, compared to other approaches used here, is that field-based measurements are not required – all values can come from the image being unmixed (Roberts et al., 2007). This means unmixing can be done on historical imagery where no ground truth or reference data is available, which made it particularly attractive for our goals. The challenge, however, is selecting endmembers that 1) have only the target land cover in the pixel and 2) best summarize a land cover’s spectral variability.

Multiple unmixing sub-approaches have been developed with their own method of endmember selection. We selected three common sub-approaches and used each to unmix woody cover, grass and soil fractions in each image. The first sub-approach, spectral mixture analysis (SMA) (Adams et al., 1986; Roberts et al., 1993; Settle and Drake, 1993; Smith et al., 1985), relies entirely on the user to identify the best endmembers. We selected endmembers with the minimum, maximum, mean, and median brightness values (sum of all bands). Separate SMAs were run for each endmember in GEE using the ‘unmix’ function (Gorelick et al., 2017).

The second sub-approach, Multiple Endmember SMA (MESMA) (Roberts et al., 1998), builds upon SMA by allowing multiple endmembers within a single class, instead of only one, affording the classifier more flexibility and greater mapping accuracies (Fernández-Manso et al., 2012). MESMA selects endmembers by running individual SMAs with every possible

combination of endmembers, while also allowing the number of classes to vary (Roberts et al., 1998). It then retains, on a per-pixel basis, only those results from the best-fit model. In other words, a MESMA result is a mosaic of the best-fit SMAs. The MESMAs were run in ENVI (Harris Geospatial Solutions, Boulder, CO) using the ViperTools add-on (Roberts et al., 2007).

Even though MESMA can take as input all potential endmembers, the evaluation of all the models (>200,000 in many cases) required significant processing time. To minimize this, Roberts et al. (Roberts et al., 2007) recommends screening out spectrally similar endmembers by examining whether the endmembers formed clusters, then selecting the endmembers that best represented each cluster. To do this, we used metrics built into VIPER Tools: Endmember Average Root Mean Square Error (EAR) (Dennison and Roberts, 2003), Count-based Endmember Selection (COB) (Roberts et al., 2003) and Minimum Average Spectral Angle (MASA) (Dennison et al., 2004). We used the metrics according to the methodology suggested by Roberts et al. (Roberts et al., 2007) (p.44), selecting endmembers with the lowest EAR within each COB-identified cluster, or the lowest MASA when COB failed to identify any clusters. We referred to this as the EMC (EAR-MASA-COB) selection method. The EMC method yielded 1-5 endmembers for each class.

In a separate method, after testing multiple selection criteria, we found selecting endmembers from each class using the minimum, maximum and mean brightness values (sum of all bands) yielded the best results. We referred to this as the HML (high-middle-low) selection method.

Altogether, these endmember selection methods yielded ca. 20-40 models per MESMA – a significant reduction from the thousands the full suite of endmembers generated. To ensure the reduction in models was not having a detrimental effect on the accuracy of the MESMA results,

we ran preliminary tests using the full suite of endmembers versus our subsets and found no significant difference.

The third unmixing sub-approach, Monte Carlo Unmixing (MCU) (Asner et al., 2003; Asner and Lobell, 2000), is similar to MESMA in that it allows multiple endmembers within a class. However, whereas MESMA selects results based on best-fit models, MCU iteratively draws random selections from the pool of endmembers, running an SMA for each. MCU then reports the mean and standard deviation of all the results as the final result. However, the final result is sensitive to the number of iterations; with too few iterations, different MCUs can yield very different results. Therefore, the stability of the final result depends on the number of iterations. Others achieved stability after 30 iterations (Asner et al., 2009a). Because we had several sites with multiple images per site, all of which might achieve stability at a different number of iterations, we set the number of iterations at 300 – a number we assumed would achieve stability in all locations. We wrote and ran our own MCU function in GEE utilizing the ‘unmix’ function. The code for the MCU function is available in the online dataset (Nagelkirk and Dahlin, 2019). Unlike the regressions, which used the reflectance data described in Section 2.2.3 (i.e., point data) to train and test models, unmixing used entire Landsat images. This, in combination with the fact that MESMA was not automatable, led to marked increases in user input and computational demands. Consequently, we limited the number of images we unmixed, along with the number of bands/variables in each image. We created DW composites with the Dry and Wet image bands stacked into a single image without any of the additional bands/variables described in Section 2.2.3. The same applied to the Dry, Wet and Tran images: they only contained the original Landsat bands. However, unlike the regressions, the unmixing approaches utilized the full suite of bands available from Landsat 8: bands 1-7 (visible, near infrared and short-wave infrared bands) & 10-

11 (thermal bands). While earlier Landsat satellites do not have bands 1, 10 & 11 we found that including these bands substantially increased accuracies, while bands representing the vegetation indices did not. We did not test all the possible additional bands and combinations due to the amount of time this would have required.

Because linear mixture models can produce results outside the range of 0-1, all the unmixing methods allow the user to constrain the results to stay within the 0-1 range. Depending on the software, we found that constraining the results caused pixels to be removed (ViperTools) or forced to zero or one (GEE) – something we could do independently in post-processing. Therefore, we chose not to constrain the values. In addition, ViperTools had the option to constrain candidate models based on their RMSE and residuals, and GEE to constrain the unmixed values to sum to one. We set the RMSE constraint to 0.1 (effectively unconstrained) and left the other two unconstrained to avoid differences across unmixing methods.

As an alternative to constraining results in the individual software, we extracted mean mapped woody cover values from the woody cover ($\overline{WC}_{WC}$) and grass ($\overline{WC}_{grass}$) endmember pixel locations, which were meant to have 100% and 0% woody cover, respectively. We then normalized the unmixed values (WC_i) following Eq. (1):

$$WC_{normalized} = \frac{WC_i - \overline{WC}_{grass}}{\overline{WC}_{WC} - \overline{WC}_{grass}}, \quad (1)$$

This was meant to preserve more of the relationship between the unmixed values and the reference data – a relationship that was lost using the software constraints described above if all unmixed values were negative or over 1, which was not uncommon. However, normalization still left some mapped values outside the 0-1 range, and it was at this point that we set negative values to zero and high values (>1) to one.

For each approach, in addition to unmixing woody cover, grass and soil, we also unmixed only woody cover and grass to test if accuracies improved when unmixing only these, the two most similar cover types. As in the regressions, we also ran a subset of approaches using only the RNS bands. However, unlike the regressions, we did not combine all the PAs and unmix them as a single PA, nor use such a model to map individual PAs.

Regression Trees

Regression trees, unlike the other methods used here, are capable of handling non-linear relationships between landcover types and their reflectance – a situation which is particularly common when mapping at regional to global scales (Staver and Hansen, 2015). Accordingly, during the development of VCF, which used a coupled, regression tree and linear regression approach, Hansen et al. (Hansen et al., 2002) found their approach outperformed spectral mixture analysis. We expected our regression tree approach to do the same.

However, we did not attempt to replicate the VCF approach, which used stepwise regressions at the regression tree nodes to smooth the outputs, along with variables derived from MODIS bands not available in Landsat imagery. Instead, we used another regression tree approach, Random Forest (RF) (Breiman, 2001). RF has become a popular and effective procedure for mapping woody cover in savannas (Gessner et al., 2013; Naidoo et al., 2012; Symeonakis et al., 2016), along with regional-to-global scale land cover (Rodriguez-Galiano et al., 2012; Vogeler et al., 2016; Wulder et al., 2018; Zhang and Roy, 2017). RF draws random samples from a dataset of predictor and response variables using bootstrap aggregation to initiate a regression tree. RF then divides the data based on variance, creating branches with the smallest possible intra-subset variance at each split. However, instead of considering the entire subset at each split, RF uses a random selection of predictor variables to determine the split. This process is repeated to generate

a “forest” of different regression trees. Once all trees are grown, a predicted value is calculated as the mean prediction of all the regression trees.

We performed the RFs in R using the randomForest package (Liaw and Wiener, 2002) and the reflectance data described in Section 2.2.3. The RFs used the entire set of predictor variables available for each image. Each RF was set to grow 500 trees and a set number of variables were randomly selected at each split: seven (out of 12 available variables) for the single-season images and seven (out of 24) for the composite images. We set the number of variables after we ran an optimization process across all the PAs that showed including more variables did not significantly improve accuracy – a threshold we sought in order to avoid overfitting the models. Like the linear regressions, the RFs were trained and applied to their respective PAs (including ALL), with the ALL-derived models also applied to the individual PAs.

Accuracy Assessment

We assessed the accuracy of the woody cover results from each mapping method using the variance explained by predictive models based on cross-validation (VEcv; Table 3) (Li, 2017). VEc and equivalent measures such as the G-statistic (Schloeder and Jacobs, 2001) and Nash-Sutcliffe efficiency (Nash and Sutcliffe, 1970) closely resemble the coefficient of determination (R^2 ; Table 3). R^2 measures the proportion of the observed variance that is predictable from an independent variable (in this case the predicted values). However, in model validation, we are not interested in the ability of predicted values to *predict* observed values; we are interested in how well predicted values *match* observed values. VEc was developed for the latter case. It does this by directly comparing observed values to predicted values, i.e., a 1:1 line; rather than compare values to a fitted regression line, as R^2 does. In this way, VEc is also similar to root mean square error (RMSE; Table 3). Beyond combining the utilities of R^2 or RMSE, VEc values below or

equal to zero correspond to instances where the mean of the observed values better predicts the observed values than the model being evaluated – something demarcated by neither R^2 nor RMSE. This means that while VEcv generates negative values that can appear meaningless, it is important to remember that the model producing a negative score could have a R^2 of 1.0 if it exhibits errors that can be perfectly predicted using the observed data. In such a case, RMSE would be required to show the model was flawed. Therefore, when a single metric is needed to rank performance across models, VEcv is superior and we use it for that reason. However, we recognize that most studies have used R^2 and RMSE and we include R^2 and RMSE values where we felt it would aid comparisons with other studies.

Table 3. Names and Equations of Accuracy Measures.

Accuracy measure	Equation*
Variance Explained (VEcv)	$\left(1 - \frac{\sum_1^n (O_i - P_i)^2}{\sum_1^n (O_i - \bar{O})^2}\right) 100 (\%)$
Coefficient of Determination (R^2)	$\left(1 - \frac{\sum_1^n (O_i - F_i)^2}{\sum_1^n (O_i - \bar{O})^2}\right) 100 (\%)$
Legates and McCabe's (E_i)	$\left(1 - \frac{\sum_1^n O_i - P_i }{\sum_1^n O_i - \bar{O} }\right) 100 (\%)$
Root Mean Square Error (RMSE)	$\sqrt{\frac{\sum_1^n (O_i - P_i)^2}{n}}$

* O_i : the observed value; P_i : the predicted value; F_i : the fitted value from regressing predicted and observed values; n : the number of observations.

In addition to VEcv, we used Legates and McCabe's efficiency (E_1 ; Table 3) (Legates and McCabe, 2013) as a secondary measure of model accuracy. Whereas model errors are squared in the VEcv equation, E_1 takes their absolute value, thereby quantifying the percentage of the sum of absolute differences explained by the model. Like VEcv, E_1 reports accuracy as percentages, with 100% corresponding to an exact match and values $< 0\%$ denoting situations where the mean of the observed values is a better predictor of the observed values than the model. However,

because VEcv uses squared errors, it is the more sensitive measure. For this reason, we primarily used VEcv and only utilized E1 if we needed to differentiate models with similar VEcv scores, as suggested by Li (Li, 2017).

We assessed the accuracy of the maps using the withheld testing data/points (30% of reference points). The same PA-respective testing points were used across all maps. We compared accuracies across approaches and sub-approaches, then across seasons and variables, and finally present the best approaches both overall and for each PA. We used ANOVAs and post-hoc Tukey honest significant different (HSD) tests to compare accuracies across approaches and sub-approaches. To compare seasons and variables, we used repeated measures ANOVAs and HSD. Throughout, we also evaluate the relationship between the Landsat-rescaled VCF tree cover dataset (Sexton et al., 2013) and woody cover. We refer to both VCF (Hansen et al., 2003) and the Landsat-rescaled VCF (Sexton et al., 2013) as “VCF” and assume that for our purposes there are no meaningful differences between the two. This is not an accuracy assessment of VCF, which monitors woody vegetation >5 meters, while our reference data monitor all woody vegetation, regardless of height. However, we included VCF to demonstrate the amount of woody vegetation excluded by VCF in these systems and assess whether VCF might be used to predict woody cover (i.e. whether an adjustment factor can convert VCF tree cover to woody cover).

Post-processing

Some post-processing of the accuracy assessment data was necessary. Overall, models trained and applied to individual PAs and all the PAs pooled together created 4842 woody cover maps with accuracies ranging from a high VEcv of 91.1% down to scores below -1000%. While negative VEcv scores are to be expected, some appeared so low they might bias our final results. Normal outlier analysis removed more points than we felt appropriate, so instead we ranked and plotted all

4842 map accuracies in search of an inflection point, i.e., where gains in accuracies from map to map become relatively consistent. We found this point around a VEcv of -500% (Figure A.1). We removed maps below this threshold from further analysis, eliminating 157 maps (~ 3% of the total number of maps). The majority (147) of these were produced using spectral unmixing. The remaining were from regressions.

Results and Discussion

Evaluation of accuracy measures

Because VEcv is not a commonly used measure, we first compared approach accuracies using VEcv, E_1 , RMSE and R^2 (Figure 6a-d). While all the measures found RF significantly outperformed the others, R^2 gave VCF scores significantly higher than the regressions and unmixing approaches, while VEcv, E_1 , and RMSE did not (Figure 6d). As outlined in Section 2.4, this is likely due to the fact that R^2 is not a measure of model accuracy while the other measures are, including RMSE (Table 3) (Li, 2017). Because VEcv produced results similar to those using E_1 and RMSE, and is the most sensitive measure of accuracy, we used it for the remaining analyses.

Best approaches and sub-approaches

Across both the approaches and sub-approaches, RF significantly ($p < 0.001$) outperformed the others, rarely scoring below zero, with a VEcv mean and standard deviation of $49.0 \pm 30.8\%$ (Figure 6a,e).

Spectral unmixing underperformed the other approaches (VEcv mean and standard deviation of $-148.8 \pm 123.0\%$; Figure 6a). We expect spectral unmixing was limited due to the fundamental challenge of choosing endmembers. While it might be possible to accurately define endmembers across a relatively small area, when the study area expands, so does the spectral variability within each landcover type, making endmember definition more of a challenge. For this reason,

successful applications of spectral unmixing often include a regional component in endmember selection (Powell et al., 2007). We decided adding a regional component to our endmembers was infeasible: candidate endmembers were often scarce within the PAs, let alone areas within the PA. Others successful applications of spectral unmixing have been aided by additional data, such as the increased number of bands and finer spectral resolution of hyperspectral imagery and/or airborne lidar which help differentiate cover types (Degerickx et al., 2019). Here we were limited to Landsat's relatively coarse spectral resolution (9 bands). Though including variables identical to those available to RF and the regressions might have improved accuracies, this was not possible due to computational limitations. Further, this would have required the removal of bands 1, 10 and 11 – bands whose testing showed improved results while the vegetation indices did not.

Among the unmixing sub-approaches, MESMA significantly outperformed MCU and SMA. We attributed this to the fact that MESMA uses a more advanced approach to selecting the best endmembers for each pixel. Because of the overall poor performance of the spectral unmixing approach, we did not evaluate the effects of the different settings within the models (e.g., unmixing woody cover and grass versus unmixing woody cover, grass and soil).

Regressions significantly ($p < 0.05$) outperformed the spectral unmixing sub-approaches (Figure 6a). While no regression sub-approach significantly outperformed the others, stepwise regression, with its ability to add significant predictor variables, performed better on average (Figure 6e). The regressions did not significantly outperform VCF in an HSD test, likely due to the limited number of VCF data points ($n = 11$).

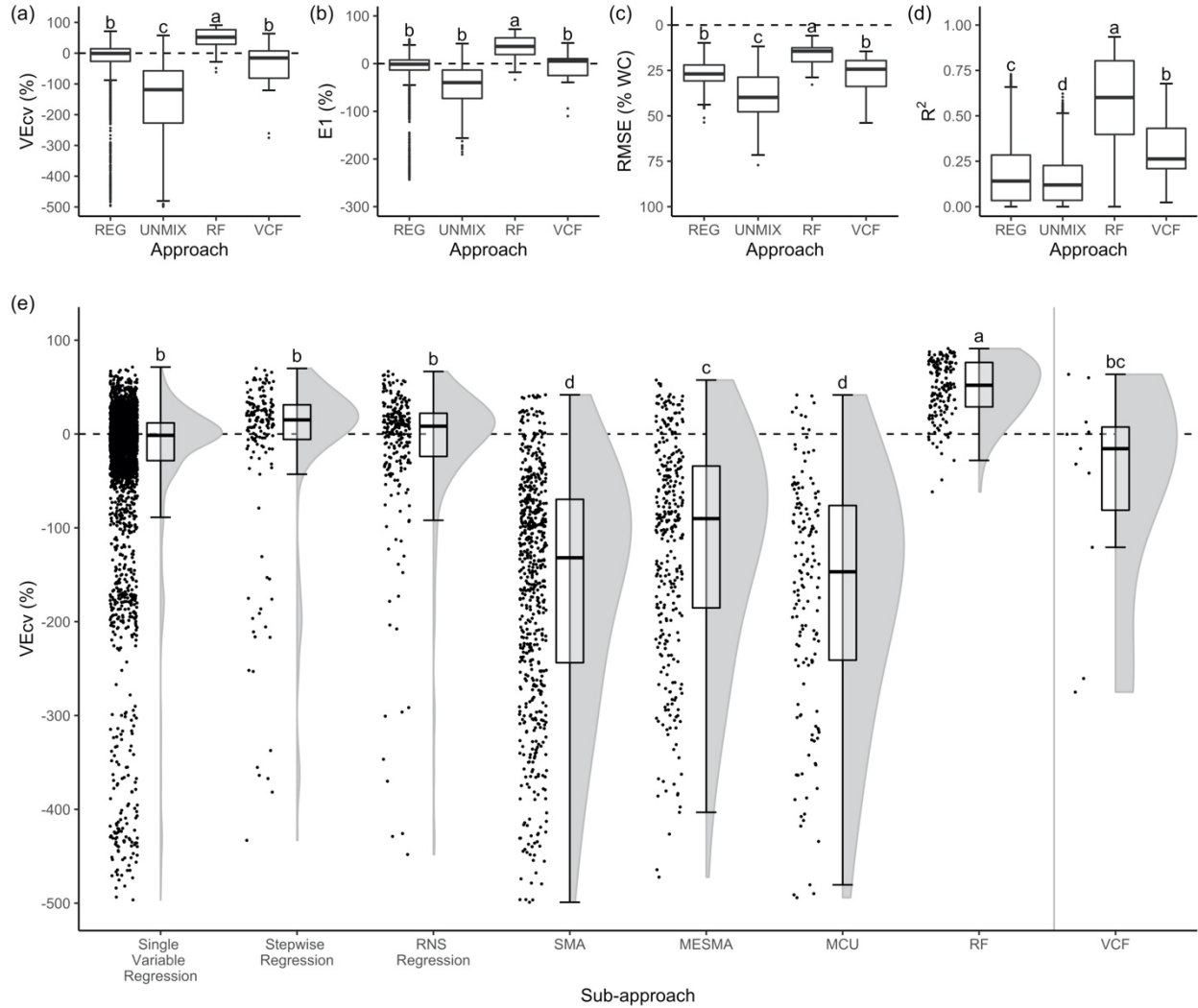


Figure 6. Approach and Sub-Approach Accuracies. Across all accuracy measures, RF significantly outperformed the other approaches (a-d). However, we note that of the accuracy measures, R^2 was the only one to give VCF a median score higher than the regressions (REG), demonstrating the discrepancies created when using R^2 as a measure of accuracy. Among the sub-approaches (e), RF remained the best performer, significantly outperforming the others. In all plots, letters signify approaches whose accuracies were not significantly different ($p < 0.05$). We tested for significance using ANOVAs and post-hoc Tukey honest significant different tests. In the boxplots, the bold centerline represents the median score, the box encompasses the 2nd and 3rd quartiles, and the top and bottom whiskers respectively represent the largest and smallest values within 1.5 times the interquartile range. Values outside that range are marked as outliers. In the sub-approach plot (e), to the left of the boxplots, points representing the accuracies of individual maps are scattered horizontally to limit overlap. The points illustrate both the number and distribution of map scores. To the right of the boxplots, smoothed histograms further illustrate the distribution of each sub-approach's scores. The gray line separates VCF from the other results as a reminder that this was not a true accuracy assessment of VCF (VCF does not incorporate all woody cover as our maps do).

Evaluation of seasonal images

Unlike the approaches and sub-approaches, both within and across PAs, no season significantly outperformed all the others (Figure 7; Table 4). DTW did the best on average (VEcv mean and SD of $16.6 \pm 23.8\%$) and significantly outperformed the Dry, Wet and DW images, with the Wet image performing the worst ($6.5 \pm 24.9\%$). DTW did not significantly outperform TD, WT or Tran. This confirmed our expectation that the image with the most information (DTW) would do the best, while the Wet season image, captured when woody cover and grass are most spectrally similar, would do the worst. Meanwhile, the Tran image was included in all the best performing images (DTW, TD, WT and Tran). We also found that Tran was the best image when all PAs were combined (see “ALL” in Table 4). These findings suggest the Tran image is critical to mapping woody cover. Indeed, all the images with the lowest scores did not contain Tran (Dry, Wet and DW) and Tran significantly outperformed both Dry and Wet, making Tran the best option if a user were going to use only a single image in a new site.

The best image for each PA varied and, like above, no image significantly outperformed all the others within a single PA (Table 4). However, unlike above, most of the best images in the PAs were something other than DTW (Table 4). Instead, Tran was most often selected as the best image and was included in the majority (9 of 12) of the PAs’ best images (i.e. DTW includes Tran). This provided further evidence that the Tran image is critical to mapping savanna woody cover, likely capturing woody cover and grass at their most spectrally dissimilar point, when woody cover is still green and grasses have senesced.

We tested whether we could generalize the best image for a PA based on the PA’s average woody cover, MAP or precipitation seasonality (WorldClim bioclimatic variables #12 & 15, respectively (Fick and Hijmans, 2017)). For instance, we expected one image might do best in

drier PAs and another in wetter PAs. However, MAP and precipitation seasonality had no relationship with accuracies. Meanwhile, woody cover had a positive relationship with accuracies, but the relationship existed across all the images (Figure A.2). While this meant woodier PAs are likely to have the most accurate maps, it also meant that, like MAP and seasonality, the woody cover of a PA cannot be used to determine the best image.

While no image across or within PAs significantly outperformed the others and the PA characteristics we tested could not be used to predict a best image, it was also encouraging, the findings also suggest one could use any of the seasonal images – whichever available. The notable exception to this was the Wet image. The Wet image had the lowest average score across the PAs, never appeared as a PA’s best image, nor was it ever used alone in any of the best models (more on models in Section 3.7; Table 4).

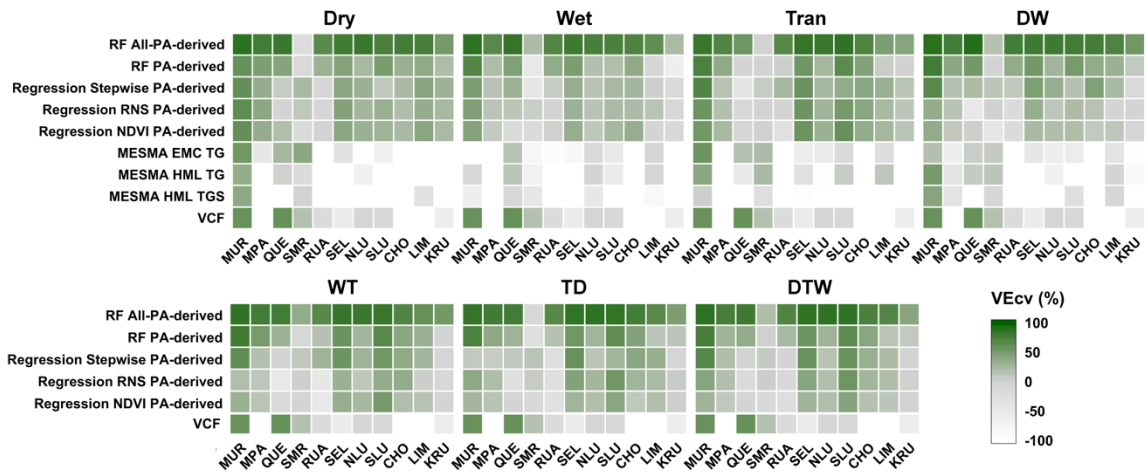


Figure 7. Heat Map of Models that Performed Best Across PAs, Seasons and Approaches. To produce this figure, we took up to three model types from each approach (RF and VCF only had two and one, respectively) with the highest average VECv scores across all the data. We then separated model performances by season and PA. Because the unmixing approach did not use the DTW, TD or WT composites, those spaces are not shown. Higher scores are darker and colors with no green coloring representing accuracies below zero. The contrast between approaches is apparent, as are the lack of any clear improvement in accuracy across the images and the consistently high scores of MUR. For display purposes only, scores less than -100 were set to -100 and VCF results were repeated across seasons.

Evaluation of protected area accuracies

Among the individual PAs, MUR had significantly ($p < 0.05$) higher scores than all the other PAs (VEcv mean and SD of $41.7 \pm 23.6\%$; Figure 7). RUA, KRU and QUE all produced the least accurate maps (average accuracies amongst the three were not significantly different). Of them, QUE had the lowest average score ($-9.1 \pm 33.7\%$). Like the images, we tested for relationships between PA accuracies and MAP, precipitation seasonality or average woody cover. Like before, we found a positive relationship (significant) with PA woody cover (Figure A.3) but no relationship with MAP or seasonality. Again, this suggests woodier PAs are easier to map while a similar relationship does not extend to MAP or seasonality.

Evaluation of variables

Across PAs and seasons, we found that regression models using RNS, NDVI and BC performed the best (average accuracies amongst the three were not significantly different), with VEcv means and standard deviations of $23.7 \pm 20.5\%$, $22.9 \pm 23.9\%$ and $21.5 \pm 21.5\%$, respectively. The variables significantly outperformed all the other variables, except Band 4 (red), which NDVI and BC did not significantly outperform. The variables with the lowest scores were SNDI, SATVI and Band 5 ($-4.7 \pm 12.1\%$, $-0.6 \pm 32.5\%$ and $-0.4 \pm 12.4\%$, respectively).

We also tested for the image-specific variable that best predicted woody cover across all PAs. We found this to be the mean NDVI of the TD composite (Figure A.4; VEcv: $29.1 \pm 25\%$). While it did not have an average score significantly higher than all the other variables, it was the only variable to significantly outperform some of the other variables (47 out of 143). Most of these (37 of 47) were composite image variables derived using the normalized difference, suggesting the mean was the better statistic to use for these images. We tested this and found the mean significantly ($p < 0.001$) outperformed the variables derived using the normalized difference, with

respective scores of $13.0 \pm 23.3\%$ and $-5.2 \pm 22.7\%$. Therefore, while the variables using the normalized difference might add some explanatory power to woody cover models, it is likely minimal.

Stepwise regression revealed no novel, consistently strong ($VE_{cv} > 50\%$) relationships between any combination of variables and woody cover across the PAs. However, NDVI and BC were the most commonly selected variables across all 84 total stepwise regressions: NDVI was selected in 29, and BC in 17. When we pooled all PA data together as if they were single PA, NDVI was not selected in any of the regressions. Instead, the stepwise regressions selected BC alone or in combination with other variables in five of the seven regressions (one regression per image). Accuracies ranged from 18.2 to 24.9%. These findings suggest that as the mapped area expands, relationships between NDVI and woody cover break down, while for BC, which simply quantifies how bright the pixel is in relation to its surroundings, they remain relatively robust.

Best models

The main objective of our work was to develop a single, accurate model for mapping woody cover. We first evaluated the best models trained and tested within the individual PAs ('Best Locally Derived Model' in Table 4). Across the 12 PAs (including ALL), most of the best models (11 of 12) were either regression- or RF-based, and several utilized NDVI. However, there was no clear pattern between the models. We then expanded the pool of models to include those trained using data from all the PAs combined ('Best Overall Model' in Table 4). Only one of the locally derived models was the best overall model for the PA: the SMR-derived MESMA that unmixed only woody cover and grass (TG) using the EMC endmember selection method and the Dry image (MESMA EMC TG – Dry). For all the other PAs besides SMR, the best overall models were RF models trained using data from all the PAs combined. When we expanded the comparison to

include all RF models (not just the best), those trained using data from all the PAs and tested in the individual PAs significantly ($p < 0.001$) outperformed those trained and tested in a single PA (VEcv: $67.7 \pm 23.3\%$ versus $30.5 \pm 27.4\%$; R^2 : 0.76 ± 0.16 versus 0.42 ± 0.19 ; RMSE: $12.5 \pm 2.6\%$ woody cover versus $19.5 \pm 5.2\%$ woody cover). This implies researchers can use training data from many savanna sites, even those hundreds of kilometers away, to improve models for a single site. This might be particularly helpful for researchers with limited data for their site. This finding also suggests that PAs share woody cover-reflectance relationships that a single model could use to accurately map woody cover across space and time.

We also tested for the model that produced the highest average accuracy across PAs ('Best General Model' in Table 4). This model was RF-ALL-DW: an RF model using training data from all the parks (ALL) combined and the DW image. Compared to the best overall models for each PA, RF-ALL-DW did not cause a significant decrease in accuracy: average accuracy across the parks fell 4.9 percentage points, from VECv values of $76.5 \pm 15.4\%$ to $71.6 \pm 20.8\%$ ($p = 0.063$). We note that when all the PAs were combined (i.e. the testing data from the different PAs was combined) and mapped as a single PA ('ALL' – last row of Table 4), the RF-ALL-DW model did not produce the best results – instead, the best model was RF-ALL-DTW. The difference is attributable to the fact that we ranked each candidate model based on its average accuracy across all the PAs, not based on its ability to map all PAs as one. (Like R^2 , the average VECv of different datasets will not equal the VECv when those data are pooled and evaluated as one.)

Last, we tested which model performed best on average across all the PAs when only using training data from the individual PAs. In other words, we wanted to know the model that was most likely to perform well when mapping woody cover somewhere outside our list of sites using only

training data from the new site. Like above, this model proved to be a RF using the DW image (RF-DW).

Both within and across PAs, while the best models did significantly outperform many of the other models, they did not significantly outperform them all (Table 4). This, in combination with the similar finding for the images, suggests that when attempting to map woody cover through the Landsat archive, researchers can be flexible in their year-to-year image and model selections. For example, if a researcher wanted to produce a map for SEL, the best model to use would be the RF-ALL-DTW (VEcv of 86.2%). But if the Wet image was not available, then the next best model, RF-ALL-TD (VEcv of 85.4%), would result in what is still an accurate map with an accuracy that is not significantly lower than RF-ALL-DTW. The same would apply if only the Tran image were available: RF-ALL-Tran would produce an accuracy of 84.5% (see the online dataset for a complete list of model accuracies). Besides providing flexibility year-to-year, these findings suggest areas with cloud cover in one image could be filled using woody cover values derived from the next best cloud-free image – something we did not attempt to do here.

Table 4. Best Seasons, Approaches & Models as Measured by VEcv (%). The best locally derived model refers to those trained using reference data from the respective PA only. The best overall model is that with the absolute highest accuracy, regardless of the source of the training data. The best general model is the single model that did best on average across all the PAs. Model naming structure: Approach - source of training data - image season - variables used in model. We list the source of the training data only when it was ALL – the combination of all the PAs in this study – otherwise the source is the PA itself. Similarly, only models that used individual variables (i.e., regressions; “REG”) list those variables. Standard deviations are given where accuracies are mean values; otherwise, values are the accuracy of the single map produced by the model. Because we did not apply the unmixing approach to all image combinations (only DW), we excluded unmixing results from the best season evaluation, resulting in higher than otherwise expected mean accuracies.

PA abbr.	Best Season		Best Approach		Best Locally Derived Model		Best Overall Model		Best General Model [†]	VCF
	Season	VEcv	Approach	VEcv	Model	VEcv	Model	VEcv	VEcv	VEcv
MUR	DTW	46.8 ± 24.5	RF*	79.3 ± 8.6	RF - WT	78.4	RF - ALL - DW	88.6	88.6	60.0
MPA	Dry	15.6 ± 16.6	RF*	55.7 ± 21	RF - WT	50.1	RF - ALL - TD	78.1	76.5	-120.7
QUE	Dry	-1.7 ± 20.3	RF*	55.4 ± 27.6	RF - DW	52.0	RF - ALL - DW	91.1	91.1	63.6
SMR	DTW	10.8 ± 18.5	RF*	-2.8 ± 21.7	MESMA EMC TG - Dry	42.1	MESMA EMC TG - Dry	42.1	15.9	12.7
RUA	WT	1.2 ± 23.9	RF*	46.1 ± 25.4	RF - DW	35.6	RF - ALL - DW	76.3	76.3	-0.2
SEL	Tran	37.1 ± 19.5	RF*	67.3 ± 16.3	REG - TD - Mean NDVI & Band 5 Normalized Difference	58.3	RF - ALL - DTW	86.2	80.2	-32.1
NLU	Tran	20.6 ± 10.4	RF*	53.7 ± 30.1	REG - TD - Mean NDVI	40.1	RF - ALL - DTW	87.1	79.7	-2.1
SLU	Tran	35.5 ± 19.6	RF*	67.3 ± 18.6	RF - WT	65.8	RF - ALL - DTW	86.6	77.6	-15.6
CHO	Tran	29.9 ± 14.1	RF*	56.2 ± 18.5	REG - DTW - Mean NDVI	46.1	RF - ALL - DW	76.8	76.8	-260.6
LIM	TD	23.2 ± 11	RF*	41 ± 27.2	REG - Dry - Brightness & NDVI	40.8	RF - ALL - Dry	75.4	71.5	-275.1
KRU	Dry	0.7 ± 18.9	RF*	21.1 ± 31.2	REG - Dry - MSAVI2	29.4	RF - ALL - DW	53.4	53.4	-41.6
ALL	Tran	15.3 ± 10.5	RF*	46.2 ± 4.7	RF - ALL - DTW	51.1	RF - ALL - DTW	51.1	49.5	-7.9

* Seasons, approaches and models that significantly ($p < 0.05$) outperformed their counterparts. Significance testing was carried out using two different measures depending on the comparison: Repeated Measures ANOVA (best season) and simple ANOVA (all others).

[†] RF - ALL - DW was the best general model.

Map comparisons

We compared the accuracies of both the maps produced by the best overall models ('Best Overall Model' in Table 4; Figures A.5-A.13) and VCF (Figures 6 & 7). For VCF, in one respect this was a simple demonstration of the difference between tree cover and woody cover, but it was also a true measure of VCF error where VCF exceeded woody cover (woody cover includes tree cover, therefore VCF values that exceed woody cover are true overestimates).

Overall, our maps had a significant ($p < 0.001$) and strong relationship with woody cover (combined $VE_{cv} = 87.0\%$, $R^2 = 0.881$; Figure 8b), while VCF had a significant ($p < 0.001$) but relatively weak relationship with woody cover (combined $VE_{cv} = 1.36\%$, $R^2 = 0.324$; Figure 8d). Our maps overpredicted woody cover in areas of sparse woody cover and underpredicted in woodier areas (Figure 8b), while VCF was mostly below woody cover but exceeded woody cover (true errors) at 139 of 421 testing points, mostly in sparsely wooded areas ($< \sim 10\%$ woody cover; Figure 8d). Wald tests showed these biases were significant (the regressions between woody cover and our maps and VCF differed significantly from 1:1 lines). However, when we used one-sample t-tests to test for overall biases, i.e., whether errors were significantly different than zero, we found our maps combined did not have a significant bias (mean and SD of errors: $0.95 \pm 10.9\%$ woody cover; $p = 0.074$), while VCF did (mean and SD: $-16.7 \pm 25.1\%$ woody cover; $p = 2.5e-35$). The size and significance of VCF's bias presented a potential adjustment factor. However, adding 16.7 percentage points (the mean error across PAs) to VCF was not enough to move the average accuracy across PAs above a VE_{cv} of 0% (average accuracy increased from a VE_{cv} of -57.1% to -12.8%). Further, when looking at PA-specific errors, our maps' errors generally center around zero (Figure 8a), while VCF's did not (Figure 8c). Combined, we interpreted these to mean VCF cannot be used to represent woody cover generally, nor can it be easily adjusted for that purpose

at the scale of a single PA. Visual assessment of our maps and VCF underscored these findings. Our maps accurately represented woody cover, even along ecotones and gradients, while VCF, whether due to error or chance (tree cover and woody cover might have no relationship), did not represent any of these well (Figure 9).

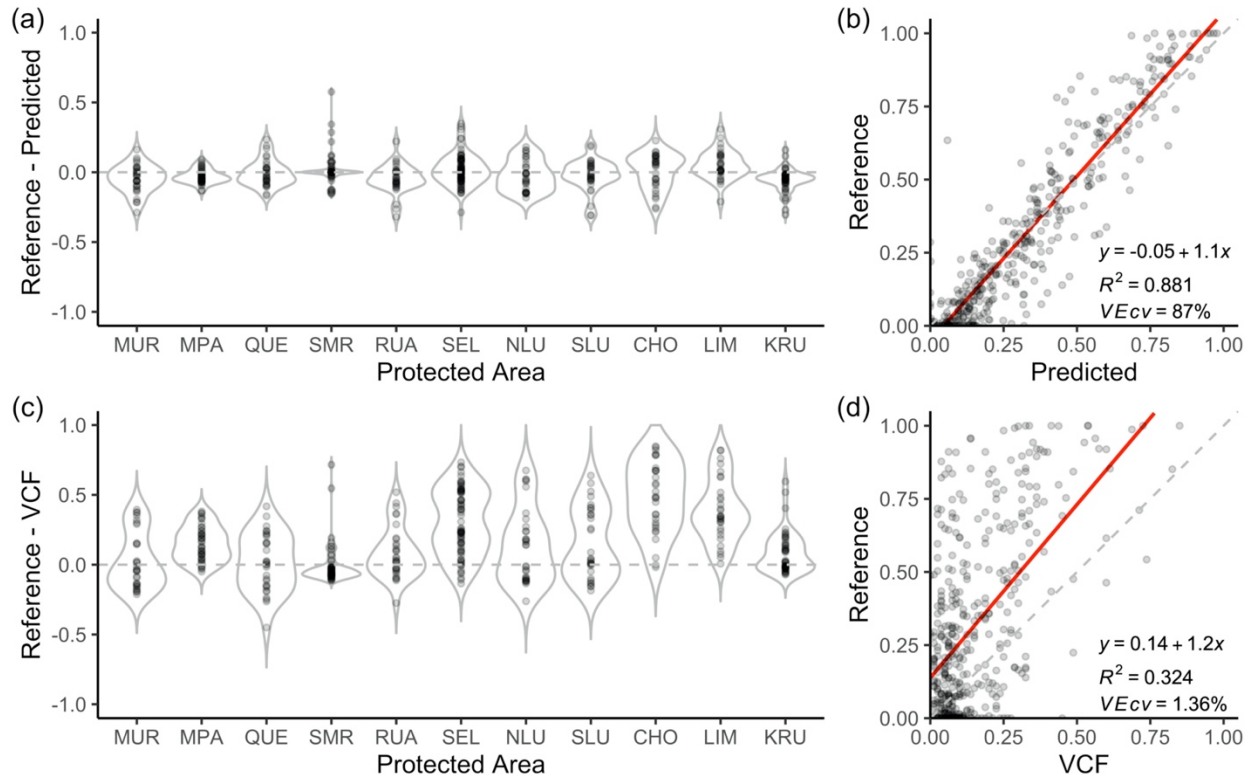


Figure 8. Our Best Maps and VCF Compared to Testing Data. Using the testing data ('Reference') from each PA, we plotted the errors of our best maps ('Best Overall Model' in Table 4) and the differences of VCF tree cover (a & c, respectively). We then pooled the PAs and compared the reference data to our best maps ('Predicted') and VCF (b & d, respectively). Overall, our best maps have errors that are not significantly different than zero (a) and a strong relationship with the reference data ($R^2 = 0.881$, $VEcv = 87.0\%$) (b). Meanwhile, differences between VCF and woody cover are significantly different than zero (c) and VCF has a weak relationship with woody cover ($R^2 = 0.324$, $VEcv = 1.36\%$) (d). Both regression lines (red) are significant ($p < 0.001$). Individual PA maps, scatter plots and accuracy metrics ($VEcv$ & R^2) are available in Figures A.5-A.13.



Figure 9. Comparison of the Best Woody Cover Maps, VCF and VHR Satellite Imagery. Maps correspond to those produced using the respective 'Best Overall Model' in Table 4. All inset maps have diameters of 275 meters. Our maps ('WC') for KRU & LIM (a), SMR (b) and CHO (c) capture gradients in woody cover (KRU/LIM inset), the complete absence of woody cover (SMR inset) and ecotones (CHO inset) better than VCF. CHO and LIM both show data missing at their eastern and western boundaries, respectively, from excluded Landsat paths that did not cover >10% of the PA. LIM, SMR and CHO also all show some areas where cloud masks removed data. Our maps and VCF have a resolution of 30m. The VHR imagery has a resolution < 1m. This figure was produced using Esri ArcGIS. Satellite imagery from ArcGIS base map. Source: Esri, DigitalGlobe, GeoEye, Earthstar Geographics, CNES/Airbus DS, USDA, USGS, AEX, Getmapping, Aerogrid, IGN, IGP, swisstopo, and the GIS User Community.

Caveats & Concerns

Because our best maps use RF, they are susceptible to the same criticism as VCF when it comes to detecting multiple stable states: the regression tree approach might create artificial discontinuities in the woody cover data that could be misinterpreted as support for multiple stable states (Hanan et al., 2014). However, we argue that by taking the average prediction of hundreds of separate regression trees, RF should minimize the risk of any artificial discontinuities in the data. Further, the sheer inaccuracy of the other approaches would also be a barrier to such studies. However, to be cautious, we advise duplicate analyses – one using the RF-based maps and the other using the best regression-based. Any significant difference in the results should be investigated accordingly.

Of our study areas, SMR was an unusual challenge to map. While SMR did not have the lowest average scores across all the PAs, its best model had the lowest score and even VCF nearly outperformed our best general model ('Best General Model' in Table 4). However, VCF also had its largest proportion of true errors in SMR, with overpredictions at 76% of its testing points, making it appear the PA is a challenge to map in general. The fact that VCF overestimated the tree cover of SMR, when it appears to have mostly underestimated it in the other PAs, suggests to us that the unusually fertile plains of SMR might make woody cover and grass spectral differences less distinct, thereby confounding models. This was further supported by the fact that, in addition to the VCF overpredictions, the largest underpredictions (and largest errors overall) of our maps were also in SMR (Figure 7). However, like the other PAs, RF models trained using data from all the PAs did better in SMR than RF models only using SMR's training data, suggesting additional training data is likely to improve results.

Conclusions

The main objective of this study was to find a procedure we could use to accurately map woody cover in both current and historical imagery using a limited number of images per year. We found that RF clearly outperformed the other approaches, while no season or specific model significantly outperformed all the others. This suggests that having limited historical imagery available should not significantly affect map accuracies. However, special consideration should be given to the Tran image, which appeared in all the best image composites and significantly outperformed the Wet and Dry images. Further, NDVI, BC and the RNS bands had the strongest relationships with woody cover, suggesting they should be included in future mapping efforts. Using training data from all the sites led to the models that performed best in all but one of the PAs, suggesting mapping efforts at one site can be aided by training data from other sites. Finally, while the best models varied by PA, the best general model across all the PAs did not significantly decrease accuracies.

Savanna ecosystems play a significant role in the global carbon cycle (Ahlström et al., 2015; Poulter et al., 2014), are critically important wildlife habitat (Malhi et al., 2016), and support a large fraction of the global human population (Dahlin et al., 2017). As such, it is essential that we develop and refine tools for mapping and understanding the spatiotemporal patterns of vegetation change in these systems. The high accuracies of our general model across our sites suggest that with proper sampling of heterogeneity, a single model could accurately map woody cover across space and time, opening the door to critical discoveries in these crucial ecosystems.

CHAPTER 2. PRECIPITATION, NOT ELEPHANTS, BEST EXPLAINS SAVANNA WOODY COVER ACROSS EASTERN AFRICAN PROTECTED AREAS

Introduction

Plants play a fundamental role in our planet's climate and biogeochemical processes (Frei et al., 2009; Kopp et al., 2005; Planavsky et al., 2014). Understanding the factors that affect the geographic distribution of plants is therefore critical to our ability to predict future changes in the Earth system. For centuries, studies have focused on the bottom-up control (i.e. nutrients and climate) of plant distributions (Forster, 1778; Holdridge, 1947; Polis, 1999; von Humboldt and Bonpland, 1807). However, recent work has offered a new focus on top-down controls, indicating that disturbances maintain woody cover below its climate-regulated potential across broad swaths of Earth – particularly in the open, grassy and sparsely wooded habitats of savannas (Bond, 2005; Good and Caylor, 2011; Sankaran et al., 2008; Scholes and Archer, 1997; Staver et al., 2011).

Understanding the processes that drive savanna vegetation patterns is important because savannas cover a fifth of Earth's land surface (Channan et al., 2014), are home to over a half billion people (Nagelkirk and Dahlin, 2020), and disproportionately affect interannual variability of the global carbon cycle (Ahlström et al., 2015). However, understanding savannas is a challenge because they appear to have multiple stable states (Hirota et al., 2011; Holling, 1973), meaning they can transition to grassland or forest states when environmental conditions or disturbance regimes are altered, making the study of disturbances critical to predicting the future of these systems.

The primary disturbances in savannas are fire (Bond and Keeley, 2005; Devine et al., 2015; Smit et al., 2010), drought (Porensky et al., 2013b; Skarpe, 1992), herbivores (Asner et al., 2016; Bond, 2008; Sankaran et al., 2013a) and humans (Bolognesi et al., 2015; Reid, 2012). Of these,

the large-scale effects of herbivores, especially megaherbivores (>100 kg), have received increasing attention. Recent findings suggest megaherbivores are capable of altering both regional vegetation distributions and biogeochemical cycles through myriad processes (Ripple et al., 2015; Schmitz et al., 2018), including nutrient and seed dispersal (Doughty, 2017; Doughty et al., 2015; Ripple et al., 2015; Wolf et al., 2013) and woody cover suppression (Bakker et al., 2016; Daskin et al., 2016; Guldemon and van Aarde, 2008; Jia et al., 2018; Malhi et al., 2016). The continued loss of megaherbivores is expected to have major impacts on ecosystem function (Enquist et al., 2020; Malhi et al., 2016; Ripple et al., 2015). Of the megaherbivores, the largest and that with perhaps the most dramatic local effects on vegetation is the African savanna elephant (*Loxodonta africana*).

Savanna elephants, as part of their normal browsing behavior, ring-bark, delimb and topple bushes and trees up to 12 m tall (Asner et al., 2016; Malhi et al., 2016; Shannon et al., 2008). Capable of increasing background treefall rates 6-fold (Asner and Levick, 2012), elephants fell trees at twice the rate of some local human communities (Mograbi et al., 2017). Further, elephants limit aboveground carbon accumulation (Davies and Asner, 2019) and the survival of woody species across all size classes (Asner et al., 2016; Dublin et al., 1990), helping to maintain areas of low tree cover while also limiting woody encroachment (Daskin et al., 2016; Stevens et al., 2016). These findings suggest that elephants could be a regional top-down control responsible for both the creation and maintenance of savannas. If they are, this would have implications not only for megaherbivore and savanna conservation, but also for global vegetation models and understanding the consequences of prior megafauna extinctions.

However, regional studies examining relationships between elephants and vegetation abundance and/or woody cover lack consensus, likely due to methodologies that rely on coarse

satellite imagery (km resolution), gross estimates of vegetation greenness, relatively short records (~10 years) and/or small-scale field plots (Duffy and Pettorelli, 2012; Hayward and Zawadzka, 2010; Sankaran et al., 2008). Meanwhile, meta-analyses largely agree, finding that elephants are related to reductions in the abundance and structural homogeneity of woody cover (Guldemand and van Aarde, 2008; Guldemand et al., 2017). This contradiction between regional studies and meta-analyses is likely due to both site selection and scale. Elephants generally stay and browse within 10 km of water during the dry season (Loarie et al., 2009a; Roever et al., 2012), when browsing rates on woody cover are highest (Codron et al., 2013; Loarie et al., 2009b), and avoid slopes greater than $\sim 4^\circ$ (Edkins et al., 2008; Roever et al., 2012). Therefore, elephant impacts on woody cover should be greatest in relatively flat areas near permanent water and we would expect disparate results between small scale studies that are likely restricted to these areas and regional studies that extend beyond them.

We tested for these landscape-level impacts of elephants across 12 protected areas (PAs) from Uganda to South Africa using a 33-year record of satellite-derived, 30-m resolution maps of woody cover. We evaluated relationships across protected areas with respect to 1) PA-wide woody cover, 2) woody cover heterogeneity, 3) proximity to water and 4) slope. By looking across a broad climatic gradient and including other variables known to affect woody cover (e.g., climate, humans and fire), we rank elephant impacts relative to these other variables. Our results have direct implications for the understanding and conservation of elephant populations, savanna ecosystems and human livelihoods.

Materials and Methods

Study sites

We selected 12 PAs located in eastern and southern Africa (Figure 10). We used PAs because most elephant populations are currently limited to PAs in eastern and southern Africa (Thouless et al., 2016). In addition, elephant census data are largely restricted to PAs. We selected PAs with a range of elephant densities and precipitation (Table 5). Further, to help ensure changes in woody cover were not the direct result of human activity, we avoided PAs known to have large amounts of charcoal production or other human activities within their borders.

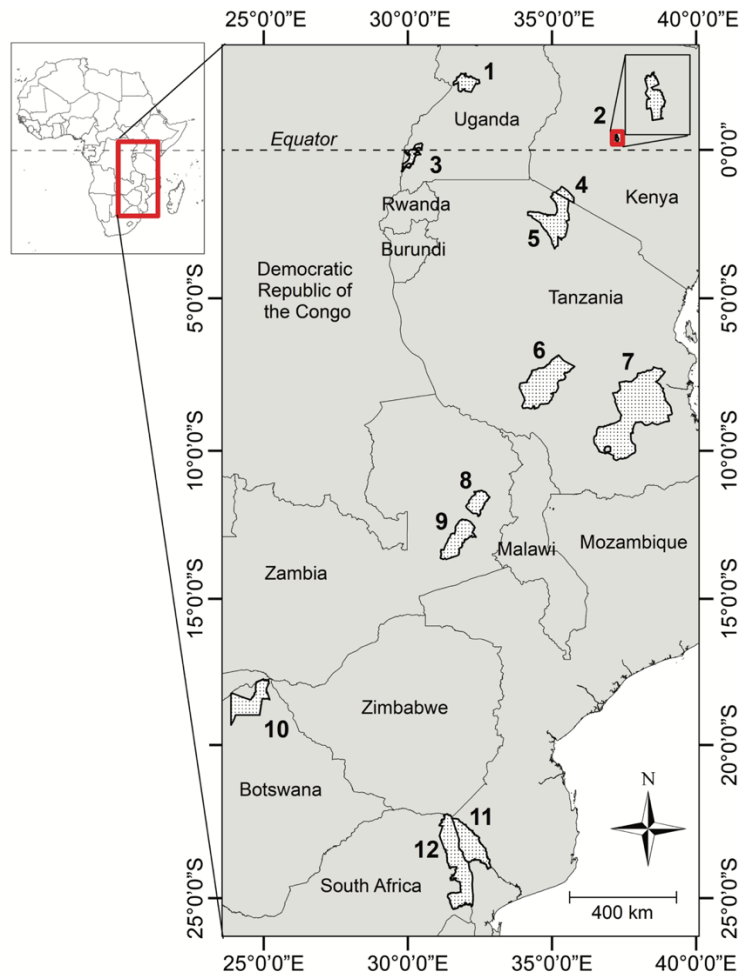


Figure 10. *The Twelve Protected Areas Used in this Study. Site numbers align with those in Table 5.*

Table 5. Protected Area Names and Attributes. Elephant densities are averages based on elephant census data since 1980. Mean annual precipitation (MAP) was calculated using the Climate Hazards Group Infrared Precipitation with Station Data (Funk et al., 2015) from 1986-2015. PA numbering corresponds to Figure 10.

PA #	Name & Country	Abbr.	Elephants per km ²	MAP (mm)	Area (km ²)
1	Murchison Falls National Park, Uganda	MUR	0.14	1262	3877
2	Mpala Research Center, Kenya	MPA	0.36	601	194
3	Queen Elizabeth National Park, Uganda	QUE	1.30	998	7395
4	Maasai Mara National Reserve, Kenya	MAR	0.98	950	1510
5	Serengeti National Park, Tanzania	SER	0.13	850	14763
6	Ruaha National Park, Tanzania	RUA	0.47	700	20226
7	Selous Game Reserve, Tanzania	SEL	0.41	1121	44800
8	North Luangwa National Park, Zambia	NLU	0.74	904	4636
9	South Luangwa National Park, Zambia	SLU	0.52	917	9050
10	Chobe National Park, Botswana	CHO	2.44	532	11000
11	Limpopo National Park, Mozambique	LIM	0.07	534	10000
12	Kruger National Park, South Africa	KRU	0.60	511	19175

Data sources

For our analysis, we compared mapped woody cover to 17 variables describing climate, fire and elephant and human population densities. The sources and methodologies for these variables are addressed in the subsequent sections.

Woody Cover

We generated 30 m resolution maps of woody cover fractions spanning 1984-2016 using Landsat TM, ETM+ and OLI surface reflectance data downloaded from GEE (Gorelick et al., 2017). All imagery was harmonized across satellites using the coefficients provided by Roy et al. (2016). We used imagery from three times of year: the wet (W) and dry (D) seasons, and the

transition (T) between the two, going from wet to dry. The images were used separately and together as composite images, yielding seven images per year: D, W, T, DW, TD, WT & DTW.

To select the seasonal images (D, W and T), we first used GEE to export the dates and average Normalized Difference Vegetation Index (NDVI) values of all available Landsat images overlaying the PAs. Each PA's area included a 20 km buffer. We then used the NDVI values to identify the dates of the D, W and T images. If fewer than three images were available in a given year, we took those images while ensuring they represented the season they had been assigned. We did this by comparing a given image's NDVI values to those of the 2016 season, along with the standard deviation of the NDVI values for each season within the respective PA. We then determined the breaks between the data using the formulae:

$$Dry_{2016} - DrySD \text{ to } \frac{Dry_{2016} + Tran_{2016}}{2} = Dry \text{ Season NDVI range} \quad (1)$$

$$\frac{Dry_{2016} + Tran_{2016}}{2} \text{ to } \frac{Tran_{2016} + Wet_{2016}}{2} = Tran \text{ Season NDVI range} \quad (2)$$

$$\frac{Tran_{2016} + Wet_{2016}}{2} \text{ to } Wet_{2016} + WetSD = Wet \text{ Season NDVI range} \quad (3)$$

where Dry₂₀₁₆, Tran₂₀₁₆ and Wet₂₀₁₆ represent the NDVI value of the 2016 Dry, Tran and Wet season images, respectively, and DrySD and WetSD represent the NDVI standard deviation of the seasonal images before filtering. Filtering the data this way still allowed some variation in the definition of the seasons while also ensuring, for example, that a dry season image from 1984 was similar – at least in terms of its NDVI – to the dry season image from 2016. This was critical, given that the RF models used to generate the maps were developed (i.e., “trained”) using the 2016 imagery. If a seasonal image did not fall within its range, the image was either discarded, or assigned to a different season if it both fell within that season's range and no other image had already been selected for the season.

For each PA, we then used the most accurate Random Forest (RF; Breiman, 2001) models developed by Nagelkirk and Dahlin (2020) to create an initial series of maps spanning 1984-2016. The models used the visible, NIR and SWIR bands shared across Landsat satellites as input (i.e., the coastal aerosol band of OLI was not used). All of the models were trained using reference data from 2016 sourced from all the PAs. When tested using separate testing data from 2016, these models had R^2 values of 0.829 ± 0.103 .

Like others who have used models developed using a single year of data to map back through time (Vogeler et al., 2018), we adopted a “space-for-time” design (Pickett, 1989). That is, since the models were developed using data from all the PAs and were consistently accurate across PAs, which each cover hundreds to thousands of square kilometers, are up to ~2700 km apart and use imagery from different dates within the year 2016, the models should incorporate enough spatial (along with some temporal) variation in their relationships to hold through time as well as space.

We used the most accurate models to create an initial series of maps spanning 1984-2016. Clouded areas in these maps were sequentially filled using the next best RF models (Table B.1). For example, a map generated using a dry season image could have its clouded areas filled using a wet season image, each using their own season-specific RF model. This was repeated to fill as many clouded areas as possible. Overall, when tested using data from 2016, the combined models had R^2 values of 0.748 ± 0.157 .

In addition to data missing due to clouds, no single PA had imagery available every season or year of the entire 1984-2016 record (Table B.2), which limited the total images available for cloud filling, along with the final number of maps. Even with these limitations, enough data were available to produce 250 maps of woody cover across the 12 PAs, or an average of ~20 woody cover maps per PA.

Elephant Density

We compiled elephant census data from multiple sources, most made available by the African Elephant Database (AED). Since 1987, the AED has periodically (4-6 years) published the best available, PA-specific elephant census data for the African continent (Blanc et al., 2007, 2003; Burrill and Douglas-Hamilton, 1987; Said et al., 1999, 1995; Thouless et al., 2016). We also collected census data not reported by the AED (e.g., Dublin & Douglas-Hamilton, 1987; Chase *et al.*, 2016), resulting in a record spanning 1984-2016, with individual census record lengths varying by PA (CHO had the shortest record: 1994-2014; Figure B.1).

The quality of census data varies across sites and years. The majority of censuses are high quality, i.e., aerial surveys covering a portion or all of a PA. The lowest quality data, and the least common (<2% since 1984), represent informed estimates from park wardens. Further, censuses represent a short-term assessment (collected over a few days) of the elephant population in a park. In an attempt to account for uncertainties, 95% confidence intervals are often reported by census teams, though they are less common in older censuses.

Elephant census teams typically conduct counts during the dry season, when elephants are most visible and, incidentally, most likely to be browsing and felling woody vegetation (Birkett and Stevens-Wood, 2005; Codron et al., 2011). Elephant ranges also constrict during the dry season (Western and Lindsay, 1984; Young et al., 2009), hence surveys should be a reasonably accurate representation of elephant populations and their browsing pressure within each PA.

Precipitation

We derived yearly annual, wet- and dry-season precipitation for each PA using the 0.05° (~5.5km) resolution Climate Hazards Group Infrared Precipitation with Station Data daily dataset

(CHIRPS; Funk *et al.*, 2015) available in GEE. All pixels with centroids within the PA boundary were included.

We also calculated the mean number of days with precipitation (DWP). Previous work demonstrated that mean wet season precipitation (WSP) and mean depth of wet season rainfall events (MDE) together affect the woody cover of African savannas (Good and Caylor, 2011). We combined these variables into a single metric, DWP, by dividing WSP by MDE ($DWP = WSP / MDE$). In doing so, we assumed that no more than one rainfall event occurred per day. We computed DWP for each year, wet season and dry season of our study in GEE using CHIRPS daily data. Because we needed to describe DWP at the PA level, we defined DWP as days where more than half of the PA received rain.

Temperature

We used 0.05° (~5.5km) resolution TerraClimate mean monthly minimum and maximum temperature data (Abatzoglou *et al.*, 2018) available in GEE to calculate mean annual temperature, along with average minimum and maximum temperatures.

Drought

We quantified drought conditions in the PAs using the 0.5° spatial resolution Standardized Precipitation-Evapotranspiration Index (SPEI; Beguería *et al.*, 2010). SPEI calculates water balance (precipitation minus potential evapotranspiration) over 1 to 48-month time periods, then uses a standardized Gaussian variate with mean and standard deviation of zero and one, respectively, to compare the values to those across a record spanning the years 1901-2015. The use of a long-term climate record helps place drought events in context, while the multi-scalar nature allows users to assess droughts across varying time scales. Because our analysis used variables with annual timesteps, we used the 12-month SPEI data.

Fire

To assess the various spatial and temporal aspects of fire and associated effects on woody cover, we calculated three metrics using the MODIS Collection 6 MCD64A1 global burned area product (Giglio et al., 2015) in GEE: the average annual burned fraction of each PA ('burned fraction'), the average annual fraction of each PA burned multiple times ($\text{burn}_{\text{mult}}$) and the total number of years >25% of the PA burned (burn_{25}). We generated the latter two variables as estimates of fire return intervals, which are known to have strong relationships with woody cover (Sankaran et al., 2008). The MCD64A1 product used to generate these variables uses 1km MODIS active fire observations coupled with a burn sensitive vegetation index based on 500m MODIS Surface Reflectance data to estimate the burn dates. These data are then compiled to produce monthly, 500m resolution maps of burned area available from November 2000 to near-present (continuously updated).

Human Population Density

We used GEE to calculate the mean estimated human population density within a 20 km buffer around each PA for years 1975, 1990, 2000 and 2015 (the only years available) using the 250m resolution Global Human Settlement Layers (JRC and CIESIN, 2015). In addition to direct impacts on woody cover, human population densities over 15.2 persons/km² deter elephants (Hoare and Du Toit, 1999). Therefore, relatively high human population densities in areas surrounding PAs might confine elephants to the PAs, thereby increasing elephant impacts on woody cover and creating an additional, indirect, negative relationship between human population densities and woody cover.

Latitude

We included latitude (PA centroid) as an independent variable because it has a significant relationship with many global and regional processes, including climate, day length, and biodiversity (e.g., Kerkhoff *et al.*, 2014). However, we did not expect latitude to have a significant relationship with woody cover in savannas because climate, especially precipitation, varies even along similar lines of latitude in eastern Africa (Nicholson, 2017), where most of the PAs are located. However, in any regional study of vegetation, latitude could play a larger role than anticipated (e.g., Dahlin *et al.*, 2017) and should be tested rather than ignored.

Elevation

We included PA mean elevation from the Shuttle Range and Topography Mission Version 3 dataset (SRTM; 30-m resolution) in our analysis because, like latitude, elevation has long been recognized as a major determinant of plant and biome distributions (von Humboldt and Bonpland, 1807).

Data processing

In addition to quantifying woody cover at the level of the PA, we divided each PA into four areas based on slope and distance from water: flat areas ($\leq 4^\circ$) near water ($<10\text{km}$); flat areas far from water ($10\text{-}40\text{km}$); hilly ($>4^\circ$) areas near water; and hilly areas far from water. To delineate these areas, we first calculated slope using 30-m resolution SRTM v3 data (Farr *et al.*, 2007) in GEE. We then calculated distance to water using the 30-meter resolution Global Surface Water dataset (Pekel *et al.*, 2016). The dataset maps monthly surface water presence across the globe. We downloaded the data for each PA, including a 50km buffer, from GEE and defined water bodies as those with water present year-round from 1984-2015.

We then calculated PA-specific annual values for both mean woody cover and the semivariance of woody cover at the level of the entire PA and within each of the four areas. Semivariance quantifies the heterogeneity of a surface, calculated as half the variance of the differences between all point values a given distance apart. We calculated semivariance at 100-meter intervals, starting at 50 meters and going to 950. This allowed us to assess woody cover heterogeneity at small (50 m) to large (950 m) distances. Hereafter, we use semivariance and heterogeneity interchangeably.

Average woody cover in 10 of our 12 sites did not show a significant trend over time, while seven did not show significant elephant density trends (Figure 11, Figure B.1). In addition, the producers of elephant census data caution against using the data in temporal analyses given that changes between years could be due to changed methodologies as well as actual changes in elephant populations, with only more recent reports attempting to account for these changes (Blanc et al., 2007). These factors, combined with gaps in the timeseries (e.g., Figure 11), led us to conduct an analysis using the temporally weighted timeseries mean of the annual values for both woody cover and elephant density, along with all other variables. In practice this consisted of a gap-filling procedure interpolating values between measured data points (Figure B.2), followed by taking the timeseries' mean, creating a single value per PA for all the variables considered.

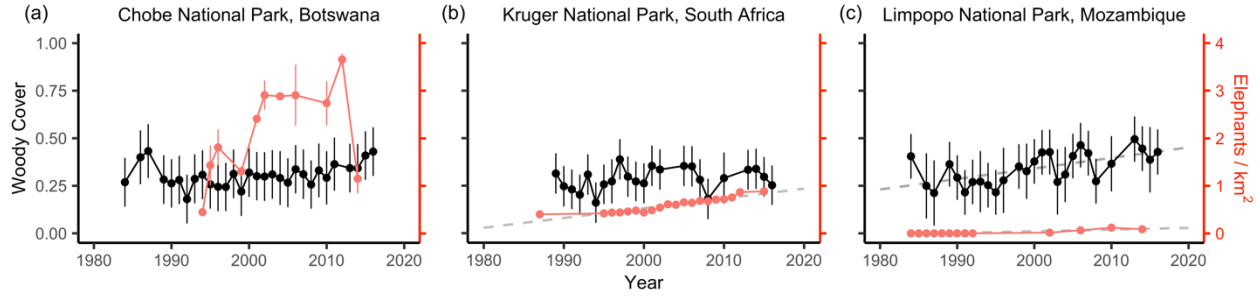


Figure 11. Woody Cover Fractions and Elephant Densities. Woody cover fraction (black) and elephant density (red) data from three of the 12 PAs used in this study (a-c). Dashed lines denote a significant trend ($p \leq 0.05$) in the data. Error bars on the woody cover data denote the spatially weighted RMSE of the models used to generate the woody cover maps (multiple models were used to generate each year's map). Error bars for the elephant data denote 95% confidence intervals. In cases where confidence intervals are absent, it is either due to values not being reported, or values were reported as zero – a common practice when an aerial census covered the entire PA. For data from all PAs, see Figure B.1.

Analysis

Our analysis consisted of two main components. In the first, we analyzed mean woody cover across all the areas using simple linear regression. For each regression, we used PA-specific means ($n=12$) from across the data record, as described above. However, for variables that did not span the dates of the full woody cover record (1984-2016), the means for both the woody cover and the variable were calculated using data from overlapping years only. For example, given that fire data was only available from 2001-2016, we calculated the mean using fire and woody cover values from 2001-2016.

In the second, at each of the ranges used to calculate semivariance, we regressed the PA-specific semivariance values ($n=12$) against the individual variables. This allowed us to assess the relationship between each variable and woody cover heterogeneity at each range. Like above, we matched the temporal record of the datasets before calculating the mean PA-specific semivariance and variable values used in each regression.

Given the limited number of data points ($n=12$) used in all the regressions, we were careful to attend to the effects of outliers. To detect potential outliers, we utilized Tukey's rule, designating values outside of 1.5 times the interquartile range as outliers. However, we did not remove all outliers, particularly when they made regression relationships weaker. Instead, we took two different approaches to removing outliers for the mean woody cover and heterogeneity regressions. For the mean woody cover regressions, we kept all woody cover values and only assessed independent variable outliers. Because we favored a conservative approach that reduced the chances of a type I error and relationships that applied for all the data, not a subset, we removed groups of outliers (two or more) that made relationships stronger while keeping those that made relationships weaker. However, we did remove lone outliers, given the disproportionate leverage they exerted on the regressions. For the heterogeneity regressions, we did remove all outliers in both the woody cover data and independent variables. We did this to generate plots using the same number of PAs in the regressions at each range, something not achieved when removing independent variable outliers only.

Results

PA-wide woody cover

Of the 17 possible variables to explain woody cover that we considered, wet season precipitation ($R^2 = 0.38$, $p = 0.03$; P_{wet}) and mean monthly minimum temperature ($R^2 = 0.43$, $p = 0.02$; T_{min}) had the strongest relationships with the average woody cover of the PAs (Figure 12), followed by the number of days with precipitation in the wet season ($R^2 = 0.36$, $p = 0.04$; DWP_{wet}) and mean annual precipitation ($R^2 = 0.36$, $p = 0.04$; MAP). Meanwhile, mean annual temperature (MAT) had a near significant relationship with woody cover ($R^2 = 0.31$, $p = 0.06$) (Figure B.3). However, T_{min} and P_{wet} were correlated ($r = 0.62$, $p = 0.03$; Table B.3), meaning their effects are

unlikely additive. Meanwhile, $T_{\min}$ was not correlated with MAP ($r = 0.51$, $p = 0.09$) nor DWP_{wet} ($r = 0.46$, $p = 0.13$), while P_{wet} was not correlated with MAT ($r = 0.47$, $p = 0.12$), suggesting these variable pairs could have additive effects.

All other variables, including those describing elephant densities, fire, drought and human population densities, did not have significant relationships with woody cover (Figure 12, Figure B.3).

Across subsequent sections analyzing areas in relation to slope and proximity to water (below), variables describing precipitation had the strongest relationships with woody cover except in a singular case where elephants had the stronger relationship. Of the precipitation variables, P_{wet} had the highest number of significant relationships (though not always the strongest) with woody cover. We therefore only show results for P_{wet} and elephant densities in the majority of the remaining sections (for all results, see Figures B.3-B.11).

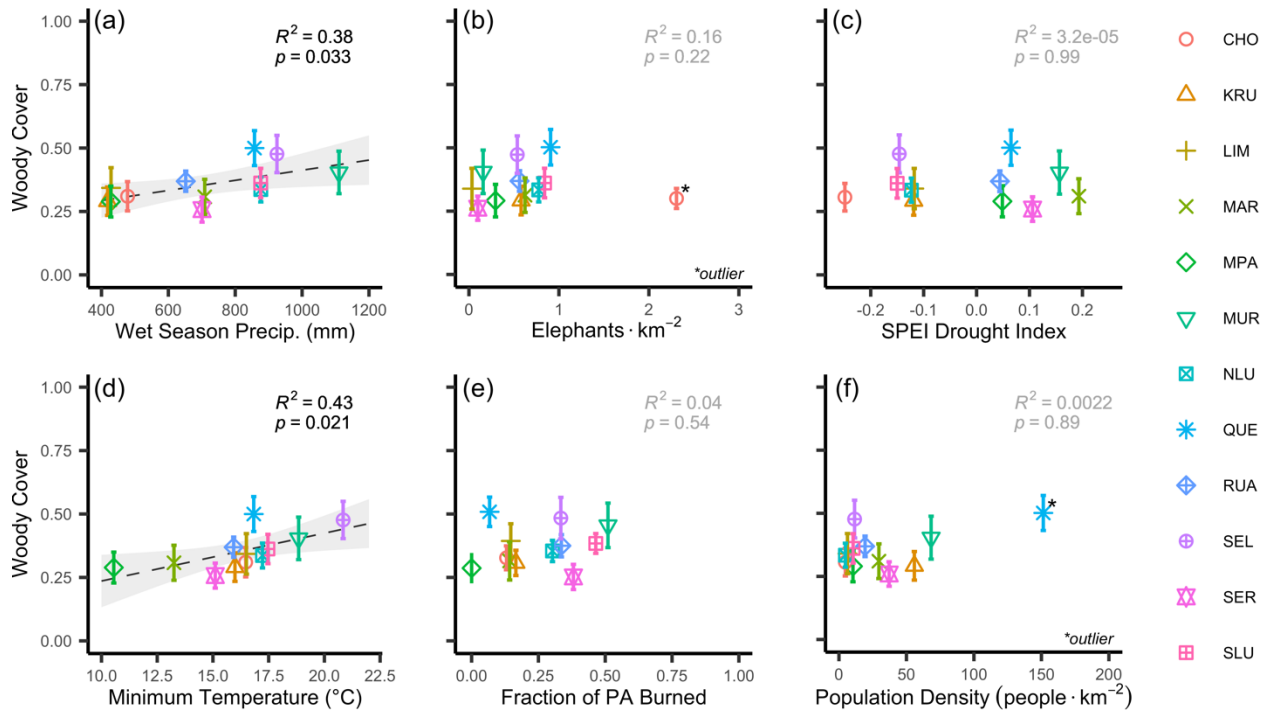


Figure 12. Relationship between PA Mean Woody Cover and a Subset of Variables. Variables related to precipitation and temperature were related to woody cover, with wet season precipitation (a) and mean monthly minimum temperatures (d) having the strongest relationships. All variables outside of precipitation and temperatures did not have a relationship with woody cover, including elephant densities (b,c,e,f). Gray text denotes non-significant R^2 values. Significant relationships have dashed regression lines with shading indicating the 95% CI. Outliers are labeled with asterisks. Relationships for all variables used in this study are available in Figure B.3. See Table 5 for PA names corresponding to legend abbreviations.

Analyzing woody cover based on slope and proximity to water

Near water, in flat and hilly areas, none of the variables had significant relationships with woody cover (Figure 13a,b,e,f; Figures B.4 & B.5).

Far from water, woody cover in flat areas was related to P_{wet} ($R^2 = 0.37$, $p = 0.036$), MAP ($R^2 = 0.36$, $p = 0.038$), DWP_{wet} ($R^2 = 0.46$, $p = 0.016$) and T_{min} ($R^2 = 0.38$, $p = 0.033$) (Figure 13c,d; Figure B.6). On hills far from water, woody cover was related to P_{wet} ($R^2 = 0.53$, $p = 0.008$), DWP_{wet} ($R^2 = 0.57$, $p = 0.005$), MAP ($R^2 = 0.42$, $p = 0.02$) and T_{min} ($R^2 = 0.40$, $p = 0.03$) (Fig 4g,h; Appendix B.10). However, elephant densities also had a significant, positive relationship with

woody cover on hills far from water ($R^2 = 0.41$, $p = 0.03$). None of the other variables had significant relationships with woody cover in these areas (Figures B.4-B.7).

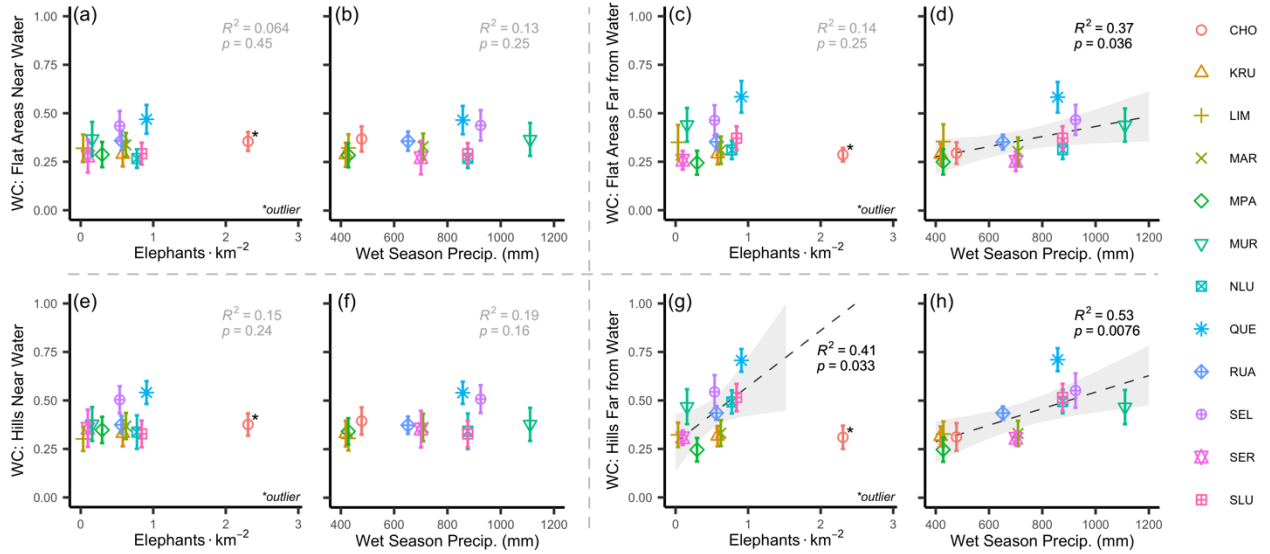


Figure 13. Relationships in the Four Within-PA Areas Analyzed. Relationships between both elephant densities and wet season precipitation (P_{wet}) and mean woody cover (WC) in the four within-PA areas analyzed: flat areas near (a,b) and far from water (c,d), and hilly areas near (e,f) and far from water (g,h). Gray text denotes non-significant R^2 values. Significant relationships have dashed regression lines with shading indicating the 95% CI. Outliers are labeled with asterisks. Relationships for all variables used in this study are available in Figures B.4-B.7. See Table 5 for PA names corresponding to legend abbreviations.

To highlight differences between sloped and flat areas near and far from water, not just absolute values, we also tested relationships between predictor variables and these differences. Overall, hilly areas had higher levels of woody cover than flat areas (Figure 14a-d). The only exception to this was Limpopo National Park (LIM), which also had the lowest elephant densities of the PAs. Near water, the difference in woody cover between flat and hilly areas was not attributable to any of the variables tested (Figure 14a,b; Figure B.8). Far from water, the differences between hills and flat areas were explained by only two variables: elephant densities ($R^2 = 0.59$, $p = 0.006$) and P_{wet} ($R^2 = 0.36$, $p = 0.038$) (Figure 14c,d; Figure B.9).

When comparing areas near and far from water, relationships were more complex. Differences between flat areas were explained by P_{wet} ($R^2 = 0.44$, $p = 0.019$), MAP ($R^2 = 0.36$, $p = 0.039$) and DWP_{wet} ($R^2 = 0.47$, $p = 0.015$) (Fig 5f, Figure B.10). However, drier PAs had higher woody cover in flat areas near water, while wetter PAs were woodier in flat areas far from water. The transition occurred near P_{wet} values of ~ 600 mm, MAP values of ~ 650 mm and DWP_{wet} values of ~ 50 days (Figure 14f, see Figure B.10).

Differences between hilly areas near and far from water were related to elephant density ($R^2 = 0.38$, $p = 0.043$), DWP_{wet} ($R^2 = 0.50$, $p = 0.01$) and P_{wet} ($R^2 = 0.48$, $p = 0.013$) (Figure 14g,h; Figure B.11). Once again, no other variables were colinear with elephant densities, but the two precipitation-related variables were colinear ($r = 0.84$, < 0.005 ; Table B.3). Like the flat areas, hills near water in drier PAs had more woody cover than those farther from water, while wetter PAs had the opposite. Also like the flat areas, the transition happened at P_{wet} values of ~ 600 mm and DWP_{wet} values of ~ 50 days.

Unlike the flat areas, elephant densities also had a relationship with the woody cover differences between hilly areas ($R^2 = 0.38$, $p = 0.043$). In PAs with lower elephant densities, hills near water had higher levels of woody cover than those far from water, while PAs with higher elephant densities displayed the opposite (Figure 14g). The transition occurred at elephant densities of ~ 0.25 elephants km^{-2} .

These positive relationships with precipitation and elephant densities were driven by increases in woody cover far from water, while woody cover near water remained relatively constant regardless of precipitation and elephant densities (Figure 13e-h).

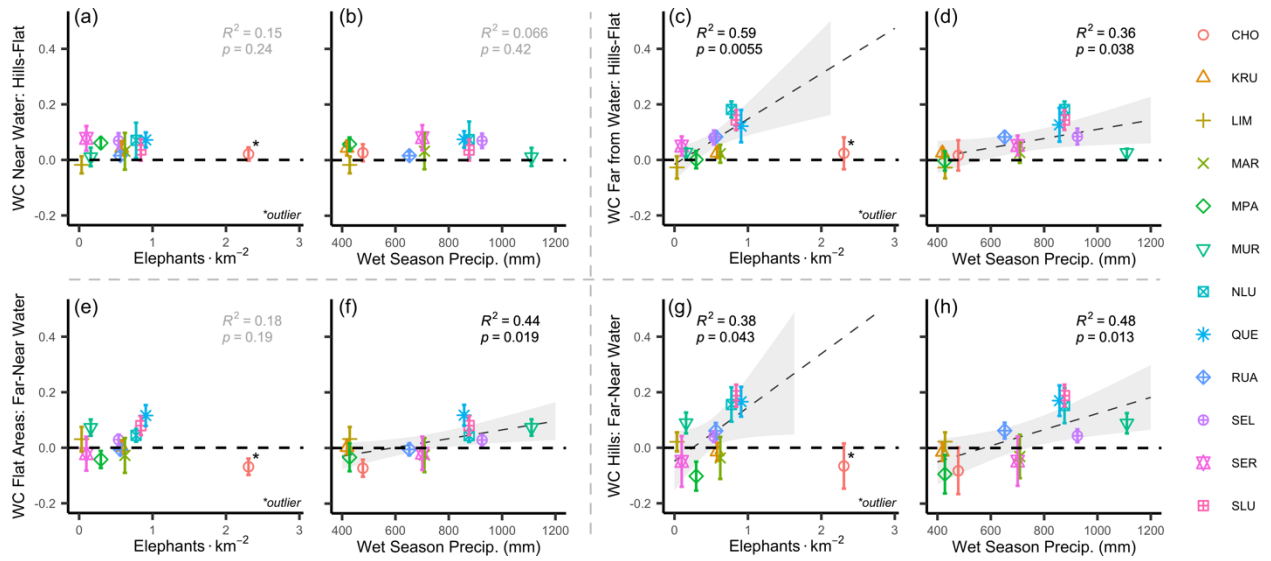


Figure 14. Relationships between Hill Slopes and Flat Areas. Relationships between both elephant densities and wet season precipitation (P_{wet}) and woody cover (WC) differences between areas. Gray text denotes non-significant R^2 values. Significant relationships have gray dashed regression lines with shading indicating the 95% CI. Black dashed lines indicate zero on the y-axis. Outliers are labeled with asterisks. Relationships for all variables used in this study are available in Figures B.8-B.11. See Table 5 for PA names corresponding to legend abbreviations.

Heterogeneity

To assess whether variables had a relationship with the heterogeneity of woody cover within and among the PAs, we calculated the semivariance of woody cover across multiple ranges using the mean woody cover maps for each PA. As before, we assessed woody cover both at the level of the PA and within areas related to slope and proximity to water. However, for areas within PAs, we only assessed woody cover heterogeneity in flat areas because hilly areas were not generally contiguous, nor large enough to assess heterogeneity at large scales.

We found that in all areas of the PAs elephant densities had no significant relationships with woody cover heterogeneity (Figure 15f,h; Figures B.12-B.14). Instead, like with average woody cover, in all areas the majority of the strongest relationships with woody cover heterogeneity were attributable to P_{wet} (Figure 15a,c; Figures B.12-B.14). P_{wet} was also the only variable to have

significant relationships with heterogeneity in all areas and at all distances measured. Similarly, MAP and DWP_{wet} had some of the strongest relationships with heterogeneity, with significant relationships at all ranges but one (50 meters). Meanwhile the number of days with precipitation in the dry season (DWP_{dry}), MAT, T_{min} , latitude, elevation and $burn_{mult}$ also had significant relationships at a smaller number of ranges (Figure 15b,e; Figures B.12-B.14).

The strength of the relationships between variables and woody cover heterogeneity were not uniform across ranges, nor in relation to water proximity, though all relationships were positive. In areas near water, P_{wet} and $burn_{mult}$ had their strongest relationships with woody cover heterogeneity at larger ranges, while T_{min} had the opposite relationship (Figure 15a,b,e). In areas far from water, P_{wet} had its strongest relationships at shorter ranges (Figure 15c). T_{min} and $burn_{mult}$ had no significant relationships with heterogeneity in areas far from water (Figure 15g,h). The only other variables to have relationships in areas far from water were other precipitation-related variables: DWP_{wet} and MAP (Figure B.14).

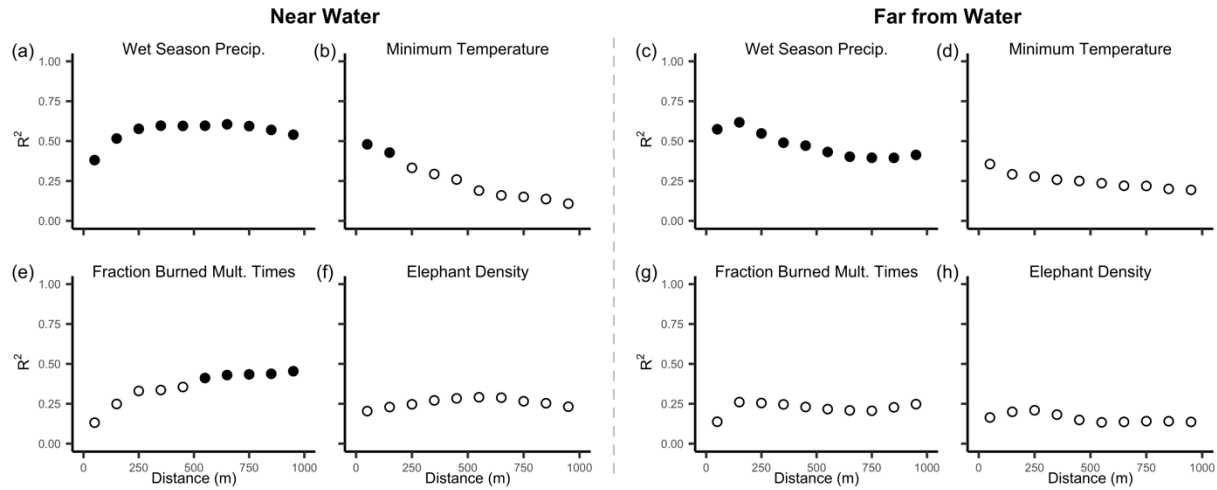


Figure 15. Relationships with Woody Cover Heterogeneity Near and Far from Water. Relationships between P_{wet} , T_{min} , $burn_{mult}$, and elephant densities and woody cover heterogeneity in areas near and far from water. Filled circles denote significant relationships ($p \leq 0.05$) between variables and the semivariance (i.e., heterogeneity) of woody cover at each range from 50 to 950 meters. P_{wet} was the only variable to have significant relationships with heterogeneity at all ranges, while the other variables had some or none depending on the area's proximity to water. Relationships for all variables used in this study are available in Figures B.12-B.14.

Discussion

Primary and secondary roles of precipitation and elephants

Decades of research into the relative importance of climate versus disturbance in maintaining savanna ecosystems has led to conflicting results across spatial resolutions and extents (e.g., Sankaran *et al.*, 2008; Asner *et al.*, 2016). Here we asked whether landscape-scale patterns differed across a regional extent. Our results demonstrate that total woody cover and heterogeneity at the landscape scale across a large latitudinal gradient are predominantly controlled by climatological variables – principally precipitation – not elephants. Specifically, our findings emphasize the importance of wet season precipitation: P_{wet} and DWP_{wet} were repeatedly among the best predictors of woody cover and heterogeneity. Consequently, these findings suggest that it should be possible to model savanna woody cover at coarse spatial resolutions without including the impacts of herbivores and fire, despite calls for their inclusion (Bond, 2005).

Precipitation is not, however, the only factor shaping woody cover within these landscapes. When we divided the landscape by slope and distance to water, our results suggest elephants are shaping woody cover in novel, unexpected ways.

While elephant densities did not have our hypothesized relationship with woody cover in flat areas near water, they did have a positive relationship with woody cover on hills far from water (Figure 13g). We propose two possible explanations for this relationship. First, previous studies have shown elephants and other megaherbivores transport nutrients and seeds upslope (Doughty, 2017; Doughty et al., 2015; Ripple et al., 2015; Wolf et al., 2013), which could be aiding the establishment of woody cover on hills in PAs with higher elephant densities. Second, because elephants switch from browsing to primarily grazing in the wet season, the same time of year they are known to range farther from water (Codron et al., 2012; Loarie et al., 2009b), they may increasingly reduce fire occurrence on slopes by reducing fuel loads (grasses) in areas near water overall – a known impact of grazers on savanna landscapes more generally (Ripple et al., 2015). Indeed, elephant density had a near significant negative relationship with $\text{burn}_{\text{mult}}$ ($r = -0.52$, $p = 0.1$; Table B.3). However, none of the variables quantifying fire had relationships with woody cover in any of the areas tested, which supports other studies finding fire impacts to be less than those of elephants and precipitation (Asner et al., 2016; Morrison et al., 2016; Shannon et al., 2011). Untangling the relationship between elephant densities, fire and woody cover would require high spatial resolution information about fire frequency and extent. This could be estimated from Landsat, however, Landsat's 16-day revisit rate and lack of data collected in Africa in the 20th century (Wulder et al., 2016) make fire frequency and extent difficult to assess across the temporal extent of this study (Hawbaker et al., 2017).

Our within-landscape results also suggest areas near water may be in a stable state. Both flat and hilly areas near water had no relationship with any of the 17 variables we tested. Instead, woody cover near water remained largely unchanged between the PAs (Figure 13a,b,e,f and Figure 16). Further, while woody cover near permanent water was highest in drier PAs - potentially due to groundwater supporting the woody cover nearer to surface water – wetter PAs displayed the opposite (higher woody cover far from water). Instead of increasing alongside the areas far from water, areas near water showed no change in woody cover. This suggests that in wetter PAs something limits woody cover near water – potentially the top down control of disturbances such as fire and elephants, much like proposed by Bond (Bond, 2005) and supported by others (Asner et al., 2016; Axelsson and Hanan, 2018; Dublin et al., 1990; Stevens et al., 2016). This could also explain why neither elephants nor fire had relationships with woody cover in these areas: the areas are in a state that does not change with relatively small fluxes in either disturbance levels or climate variables – a central tenant of multiple stable states theory (Holling, 1973). If these areas are in fact in a stable state, then disturbance regimes would need to be substantially altered to detect their impacts. Indeed, experiments using elephant exclosures or studying the effects of local extinctions find significant increases in woody cover after elephant removal (Asner et al., 2016; Stevens et al., 2016). Our work shows something similar: in PAs with sufficient rainfall to support high levels of woody cover, areas with less browsing pressure from elephants (areas far from water) have higher woody cover than areas that experience high levels of browsing pressure (areas near water; Figure 16b,d).

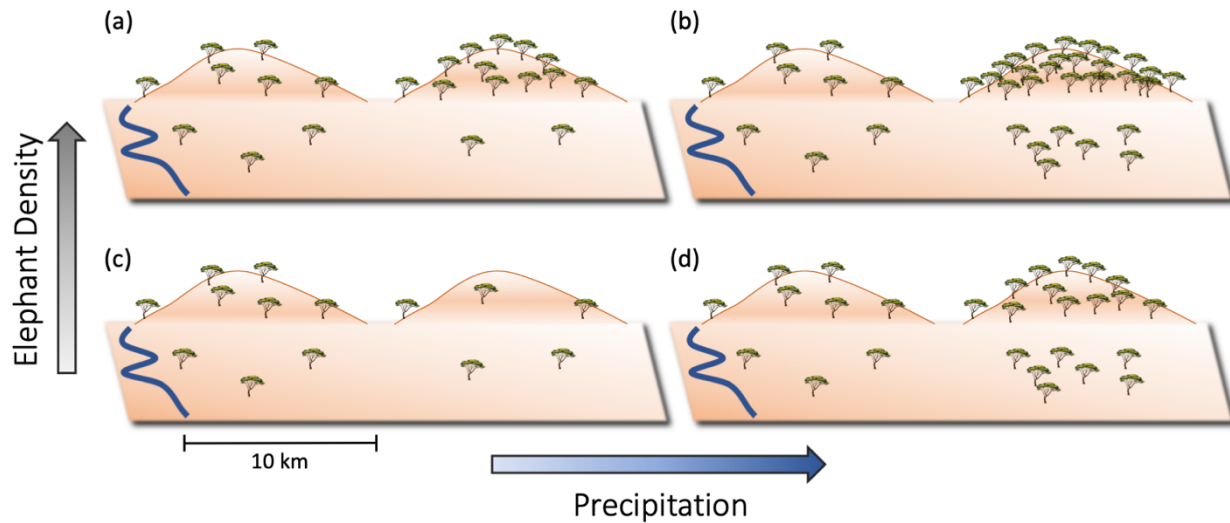


Figure 16. Conceptual Diagram of Changes in Woody Cover in Relation to Elephant Densities and Precipitation. Both flat and hilly areas near water do not change with elephant densities nor precipitation, leading us to conclude these areas are in a stable state (left side of a-d). However, in areas far from water (right side of a-d), woody cover on hills increased with respect to both elephant densities and precipitation. Flat areas far from water also had a positive relationship with precipitation, but no relationship with elephant densities.

Effects on woody cover heterogeneity

Like overall woody cover, heterogeneity was primarily associated with precipitation, principally P_{wet} (Figure 15, Figures B.12-B.14). Further, P_{wet} was highly correlated with heterogeneity at all scales, suggesting it is the major driver of the spatial patterns of woody cover at both large and small scales. Meanwhile, we found no relationship between elephants and woody cover heterogeneity at any range or in any area. In addition to precipitation, fire was related to heterogeneity at larger ranges, similar to other studies (Holdo et al., 2009a; Smit et al., 2010). However, fire first became a significant predictor of heterogeneity after a range of 500 meters, which matches the resolution of the MODIS fire product. This likely means that the relationship between fires and heterogeneity at large scales is due to the fire product predominately detecting large-scale fires (i.e. $\geq 500 \text{ m}^2$). Therefore, again, the use of a higher resolution fire product might find relationships at a smaller scale, if one were available. Last, temperatures (T_{min} and MAT;

Figure 15b, Figure B.12) were only related to heterogeneity at the smallest scales in areas near water, suggesting increased temperatures cause relatively uniform increases in landscape-level woody cover but increased small-scale patchiness in these areas.

Habitat degradation by elephants

Elephants do not appear to degrade landscapes via the unsustainable removal of woody cover or a reduction in heterogeneity. This finding corroborates other work focused on single PAs or smaller regions (Kalwij et al., 2010; Owen-Smith et al., 2006; Skarpe et al., 2004). This does not mean elephants are not suppressing woody cover, particularly in areas near water; rather, any potential effects are uniform across the elephant densities considered here. However, the different components of woody cover could be changing while woody cover itself is not; many savannas are undergoing high levels of bush encroachment while losing their mature trees (Owen-Smith et al., 2006). Differentiating between bushes and trees, however, was not possible using our methods (Nagelkirk & Dahlin 2020).

The PA that has been central to the debate of whether elephants degrade ecosystems is Chobe National Park, Botswana (CHO). CHO has some of the highest elephant densities of any PA in Africa (Thouless et al., 2016), along with relatively low woody cover levels and heterogeneity. Our results highlight these characteristics, while also providing a climatological context that helps explain its low woody cover and heterogeneity: both CHO's woody cover and heterogeneity fall in line with values predicted by P_{wet} (Figure 12a, Figure 13d,h and Figure 15a,c). Therefore, while CHO's elephant population is an anomaly, its woody cover is near that expected based on its precipitation levels, demonstrating the importance of including climate in any assessment of savanna woody cover.

Conclusion

The overall amount and heterogeneity of woody cover in our sites was principally controlled by precipitation, not elephants or fire. This suggests that current modeling frameworks that exclude the impacts of fire and herbivores should be able to accurately predict regional woody cover at coarse spatial and temporal extents. However, our analysis did not include the different components of woody cover, which might not be individually predictable. Within PAs, elephants do appear to significantly increase woody cover on slopes far from water, perhaps through fire suppression and/or nutrient and seed dispersal, suggesting that within landscapes elephants are impacting the distribution of woody cover. In addition, across our study extent, areas near permanent water appear to be in a stable state potentially maintained by disturbances. Yet, our work provides no compelling evidence of elephants degrading landscapes, even at high densities, suggesting elephant culling is an unnecessary measure in most PAs while reinforcing the importance of bottom-up controls in savannas.

CHAPTER 3. FIRE OR FORCE: A META-ANALYSIS OF THE RELATIVE IMPACTS OF FIRE AND ELEPHANTS ON WOODY VEGETATION IN AFRICAN SAVANNAS

Introduction

Despite covering over a fifth of the Earth's surface (Channan et al., 2014), being home to an estimated 500 million people (Nagelkirk and Dahlin, 2020) and driving intra-annual variability in global carbon cycle (Ahlström et al., 2015), the factors shaping the woody cover of the world's savannas are not well understood, contributing to the challenges of predicting changes in these ecosystems (Baudena et al., 2015; Dahlin et al., 2015). At the same time, widespread woody encroachment is threatening savannas and human livelihoods by reducing the grassy cover required by both domestic livestock and wildlife (Smit and Prins, 2015), while also encroaching on the open views of wildlife critical to the tourist industry (Gray and Bond, 2013).

Understanding the factors that shape savannas has been a challenge because many savannas exist in a climate-indeterminate state, containing woody cover below their potential (Hirota et al., 2011; Sankaran et al., 2005; Staver et al., 2011). A growing consensus has identified fire and herbivory as the primary disturbances responsible for the characteristically low woody cover of savannas (Bond, 2005; Lehmann et al., 2014; Sankaran et al., 2005; Staver et al., 2009), with megaherbivores (>100 kg) playing an outsized role among herbivores (Ripple et al., 2015).

Globally, the impact of megaherbivores on savannas and other biomes is thought to have been especially significant in the past. In the Cretaceous period (145-66 million years ago), for example, terrestrial megaherbivores appear to have increased regional nutrient availability and dispersal (Doughty, 2017). More recently, in the Pleistocene (2.6 ma to 11,700 years ago) it is thought that megaherbivores performed a similar role, with megafaunal extinctions reducing global nutrient transport by animals by over 90% (Doughty et al., 2015). The extinctions also likely led to

increases in regional woody cover as high as 40% in South America and Europe (Doughty et al., 2016; Sandom et al., 2014), with similar changes in Siberia and Beringia causing decreases in albedo and subsequent regional warming near 1°C (Doughty et al., 2010). Meanwhile, extant megafauna are still thought to significantly influence the global carbon cycle (Schmitz et al., 2018) and suppress local vegetative biomass in several regions (Daskin et al., 2016; Jia et al., 2018; Stevens et al., 2016), but the large-scale consequences – particularly for woody vegetation – remain unclear (Staver, 2018).

Some of the last relics of the Pleistocene are in Africa (Malhi et al., 2016; Ripple et al., 2015), where the world's largest terrestrial megaherbivore, the savanna elephant (*Loxodonta africana*), de-limbs, ring-barks and topples woody cover up to 12 meters tall (Shannon et al., 2008). Elephants can drive landscape-level annual woody vegetation mortality rates of nearly 8% (Asner et al., 2016; Birkett, 2002; Mograbi et al., 2017), thereby limiting woody cover and its encroachment into grasslands (Marston et al., 2017; Skowno et al., 2017; Stevens et al., 2016).

In addition to elephants, fire limits the survival of savanna woody vegetation, but unlike elephants, fire principally limits woody cover recruitment (Higgins et al., 2007; Smit et al., 2010). However, intense fires can still kill mature woody vegetation (Bond and Keeley, 2005; Cochard and Edwards, 2011; Smit et al., 2016). Fire intensity depends on the amount and condition of grassy biomass (fuel) available, which is in turn controlled by season (dry season fires are hotter), annual rainfall (increases fuel loads), grazing herbivores (decrease fuel loads) and management (prescribed burns lead to fewer high-intensity fires) (Bond and Keeley, 2005; Govender et al., 2006; Midgley et al., 2010; Pachzelt et al., 2015; Smit et al., 2010; Van Langevelde et al., 2003).

Together, elephants and fire control woody cover in many savanna ecosystems (Sankaran et al., 2008; Shannon et al., 2011; Vanak et al., 2012). However, our ability to model and conserve

savannas relies on our understanding of the relative impacts of elephants and fire. This has led to decades of work aimed at determining whether elephants or fire play the primary role in controlling savanna woody cover (e.g., Buechner and Dawkins, 1961; Pellegrini et al., 2017; Smart et al., 1985).

Determining the relative impacts of fire and elephants is a challenge because of their complex interactions. For example, by knocking over or damaging the bark of trees, elephants make them more susceptible to fire (Moncrieff et al., 2008; Smit et al., 2016; van Wilgen et al., 2008; Vanak et al., 2012). Further, by opening up otherwise impenetrable stands of woody cover, elephants allow grasses – and therefore fire – to invade (Buechner and Dawkins, 1961; Dublin et al., 1990; Eltringham, 1979; Van Langevelde et al., 2003). Complicating the issue, elephants also contribute to the fire-limiting effects of other grazing herbivores (Bond and Keeley, 2005; Smit and Archibald, 2019) given grasses comprise ~50% of their diet (Codron et al., 2006). Meanwhile, fire-damaged trees are more likely to be toppled by elephants (Chafota and Owen-Smith, 2009). Elephants and fire also interact with drought which can exacerbate the impacts of both disturbances (Allen et al., 2010; Birkett and Stevens-Wood, 2005). For example, when droughts kill trees, they add to the fire fuel load. Droughts also cause grasses to senesce early, which causes elephants to switch to browsing woody cover for more of the year (Birkett and Stevens-Wood, 2005; Chafota and Owen-Smith, 2009), increasing the trees felled by elephants while once again creating additional fuel for fires.

Here, we build upon prior meta-analyses focused on the impacts of elephants (Guldmond and van Aarde, 2008; Guldmond et al., 2017), fire (Archibald et al., 2010; Pellegrini et al., 2018) and global changes in savanna woody cover (Stevens et al., 2017) by focusing on studies that specifically test the relative impacts of fire and elephants on woody cover, while accounting for

the abiotic factors that lead to their dominance. The results affect our understanding of the forces shaping today's savannas, along with the potential consequences of megafaunal extinctions in the past.

Methods

Literature search

We searched the literature for all peer-reviewed articles comparing the impacts of fire and elephants on woody vegetation using Google Scholar and Web of Science. In each, we performed keyword searches using “elephants” and “fire” together in combination with “savanna”, “savannah”, “tree cover”, “woody cover” and “woody vegetation” individually (i.e. five total combinations). We also limited articles to those published in English and available online, making our search a lower bound of the available literature. The last search was conducted on July 23rd, 2020. We used titles and abstracts to select 167 articles appearing to compare the impacts of elephants and fire. We reviewed these articles, along with relevant citations within their text, bringing the total article count to 227. All records were entered into a shared spreadsheet, then downloaded and kept in CSV format for analysis.

From the 227 articles, we recorded only those articles with findings on the relative impacts of elephants and fire, classifying the findings as either quantified or descriptive. We defined “quantified” findings as those that explicitly measured the effects of elephants and fire, making comparisons possible, even if the study itself was not explicitly designed to test the relative impacts of fire and elephants. Articles that drew their conclusions primarily based on anecdotal evidence and/or quantified the effects of only one disturbance while making assumptions about the effects of the other were categorized as descriptive.

Given that both elephant populations and fire regimes have been heavily managed throughout Africa (Reid, 2012; Thouless et al., 2016), we did not exclude studies conducted within areas under such management, so long as the management was not responsible for the complete removal of either disturbance at the time of the study (unless the removal itself was part of the experiment, such as in exclosure experiments). We also included studies whose sites naturally had little to no fire but we required the presence of elephants in all studies. This means that our results only apply to elephant-inhabited savannas, which today, due to habitat loss, hunting and poaching, are only a fraction of Africa's savannas (Thouless et al., 2016).

When we encountered studies where fire and elephants were interacting, we designated the disturbance that facilitated the other disturbance as primary. For example, unless otherwise concluded by the authors of the study, if elephant damage was causing fire-tolerant trees to become susceptible to fire (e.g., Okula & Sise, 1986), we identified elephants as the primary disturbance. When authors reported findings from multiple sites in the same general area, the disturbance that was dominant in the majority of sites was recorded as the dominant disturbance for the study. Last, we did not include studies that inextricably grouped the impacts of elephants with those of other herbivores. For example, Staver *et al.* (2009) used exclosures that removed all herbivores from small plots of land, meaning any potential effects of elephant removal could not be separated from those of the other excluded herbivores.

Analysis

Given that all of the studies recorded had some or all of their field sites within PAs, combined with the challenge of integrating studies with various spatial extents, units of measure and methodologies, we chose to generalize the study locations and findings to the PAs. To test for the abiotic and biotic differences between fire-dominated and elephant-dominated protected areas

(PAs), we collected PA-specific variables describing climate, fire occurrence, elephant densities and human population densities (Table 6), all of which are known to affect the extent and degree of elephant and/or fire impacts (Bond et al., 2005; Guldemon et al., 2017; Hoare and Du Toit, 1999).

Table 6. Descriptions of Variables Used in Analysis.

Variable(s)	Descriptions
ELE	The number of elephants km ⁻² in the study site, when reported by the authors. Otherwise, we calculated elephant density of the PA based on census data collected nearest the study's publication date.
MAP & P _{wet}	The average annual precipitation (MAP) and wet season precipitation (P _{wet}) from 1990-2019 calculated using the 0.05° (~5.5km) resolution Climate Hazards Infrared Precipitation with Stations daily data (CHIRPS; Funk et al., 2015).
MAT, T _{min} & T _{max}	The average annual temperature (MAT), minimum and maximum temperatures (T _{min} & T _{max}) from 1990-2019 calculated using the monthly, 0.05° (~5.5km) resolution TerraClimate dataset (Abatzoglou et al., 2018).
BurnFrac	The average fraction of the PA burned annually from 2001-2019 calculated using the 500 m resolution MODIS Collection 6 MCD64A1 global burned area product (data available from 2001-2019; Giglio <i>et al.</i> , 2015).
POP	The average human population density within a 20-km buffer of each PA from 1990-2015 using the 250m resolution Global Human Settlement Layers product (data available to 2015; JRC & CIESIN, 2015).
PET	The average annual potential evapotranspiration from 1990-2019 calculated using the monthly, 0.05° (~5.5km) resolution TerraClimate dataset (Abatzoglou et al., 2018).
PDSI	The average annual Palmer Drought Severity Index from 1990-2019 calculated using the monthly, 0.05° (~5.5km) resolution TerraClimate dataset (Abatzoglou et al., 2018).

We used Google Earth Engine (Gorelick et al., 2017) to compile all data except the elephant densities, which we compiled from the individual studies. When elephant densities were not reported in the studies themselves, we used census data published by the African Elephant Database (Blanc et al., 2007, 2003; Burrill and Douglas-Hamilton, 1987; Said et al., 1999, 1995; Thouless et al., 2016) and other sources (e.g., Dublin & Douglas-Hamilton, 1987; Chase *et al.*, 2016).

To avoid pseudo-replication, we filtered the results in the same PA to the dominant finding. For example, if four studies were conducted within the same PA with three identifying elephants

as dominant, then elephants were recorded as dominant. When this resulted in a tie, we recorded it as such.

To test for abiotic and biotic differences between elephant- and fire-dominated PAs, we used Welch's t-test, which is more reliable than Student's t-test when using small sample sizes with unequal variances and/or unequal within-class sample sizes (Ruxton, 2006). To test for differences between all PA types (elephant dominated, fire dominated and ties), we used ANOVA with post-hoc Tukey honest significant difference (HSD) tests. We then used multinomial logistic regression (MLR) to assess potential relationships between variables and the dominant disturbances. MLR models the relationship between categorical dependent variables and categorical and/or continuous independent variables (Campling et al., 2002). We tested both single-variable relationships, along with all possible paired variable combinations, excluding pairs with significant correlations ($p < 0.05$). We then assessed the models using the Akaike information criterion (AIC), accuracy, Cohen's kappa statistic (kappa) and their significance. Like accuracy, kappa assesses the fraction of observations correctly classified, but then penalizes for the number of correct classifications that could be due to random chance (MacGarigal et al., 2000), making it a better measure of accuracy when the proportion of observations in each class varies significantly, such as in our case. Unlike accuracy and kappa, the AIC assesses the relative amount of information not included in a model, making lower scores more desirable (Chatterjee and Hadi, 2012). Models were deemed significant if all variable coefficients had p-values less than or equal to 0.05. Using the most accurate model, we then constructed probability surfaces illustrating the probability of dominance the model would assign to each disturbance across the range of observed inputs. All analyses were carried out in R (R Core Team, 2018), with the MLR conducted using the nnet package (Venables and Ripley, 2002).

Results

In all, from our original 227 studies we recorded 66 studies comparing the impacts of elephants and fire. The studies spanned 13 countries and 25 PAs (Figure 17a). Of the 66 studies, 53 (80.3%) reported elephants as the dominant disturbance and 13 reported fire. However, only 39 of the studies made quantified comparisons, reducing the findings to 8 countries and 16 PAs (Figure 17b). Of these studies, 30 (76.9%) found elephants to be dominant. Filtering the data to a single finding per PA further reduced the data to 16 data points (one per PA). Of these, 12 (75%) were dominated by elephants, two by fire and two produced a tie (Figure 17c).

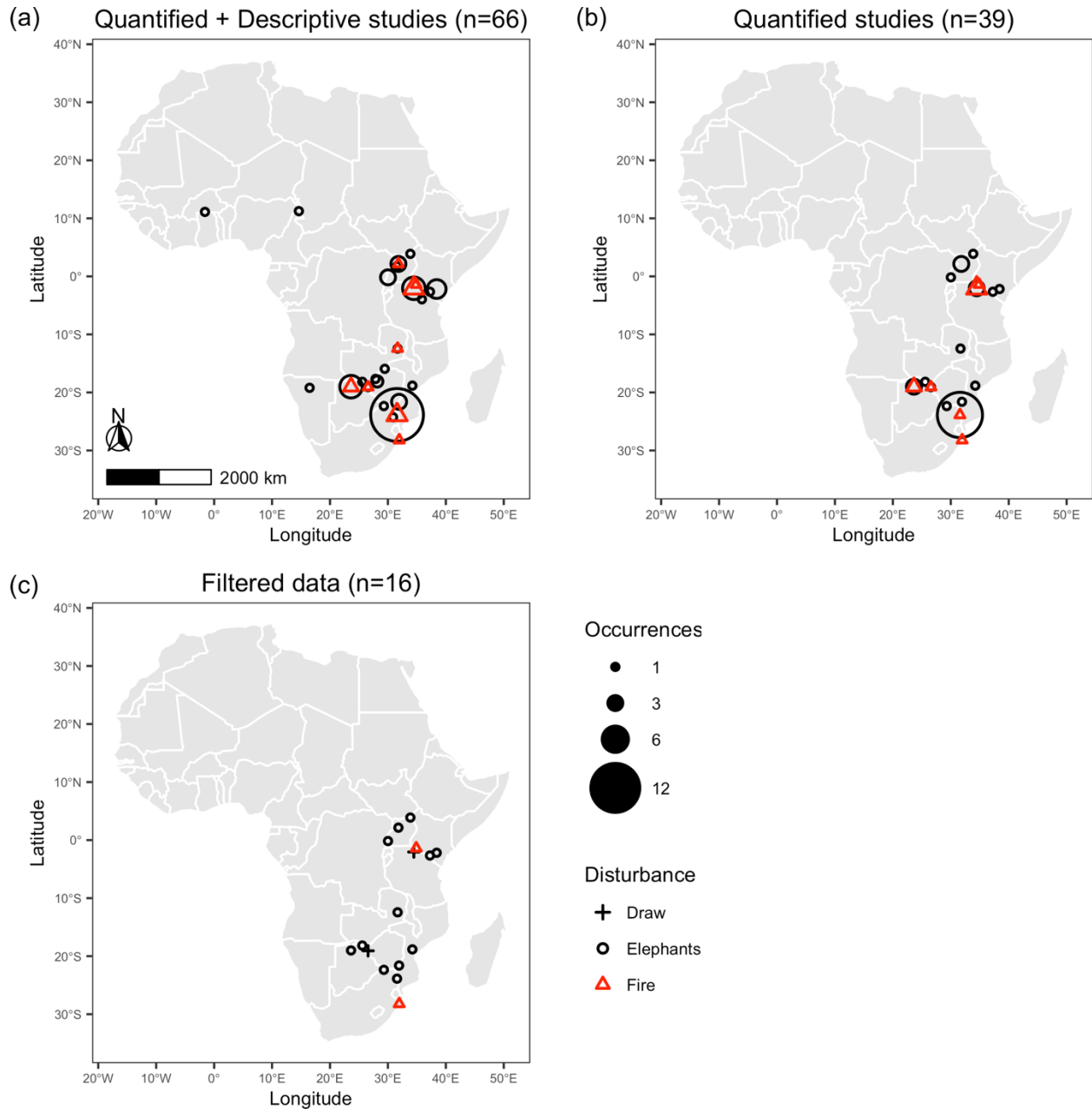


Figure 17. Maps of the Studies Comparing Elephant and Fire Disturbances. Descriptive and quantified comparisons were found in 66 studies (a), 39 of which were quantified (b). Filtering the data to the dominant findings in each PA resulted in 16 data points to be used in analysis (c).

The final set of 16 PAs covered a broad range of variable values, with MAP ranging from *ca.* 300 to 1,300 mm, ELE from 0.14 to 2.84 elephants/km² and burned fraction from *ca.* 0% to 48% of the PA burned annually (Figure 18). Welch's t-tests revealed elephant- and fire-dominated sites varied significantly in their levels of potential evapotranspiration (PET; $p = 0.045$) and MAP ($p =$

0.015), with elephant-dominated sites having higher average PET (1323 vs. 1227 mm) and lower average MAP (730 vs. 1014 mm). No other variables had significant differences between the sites. However, when the two sites with ties between elephant- and fire-dominant findings were included, ANOVA using Tukey HSD post-hoc tests revealed no significant differences in variable values across the sites (Figure 18).

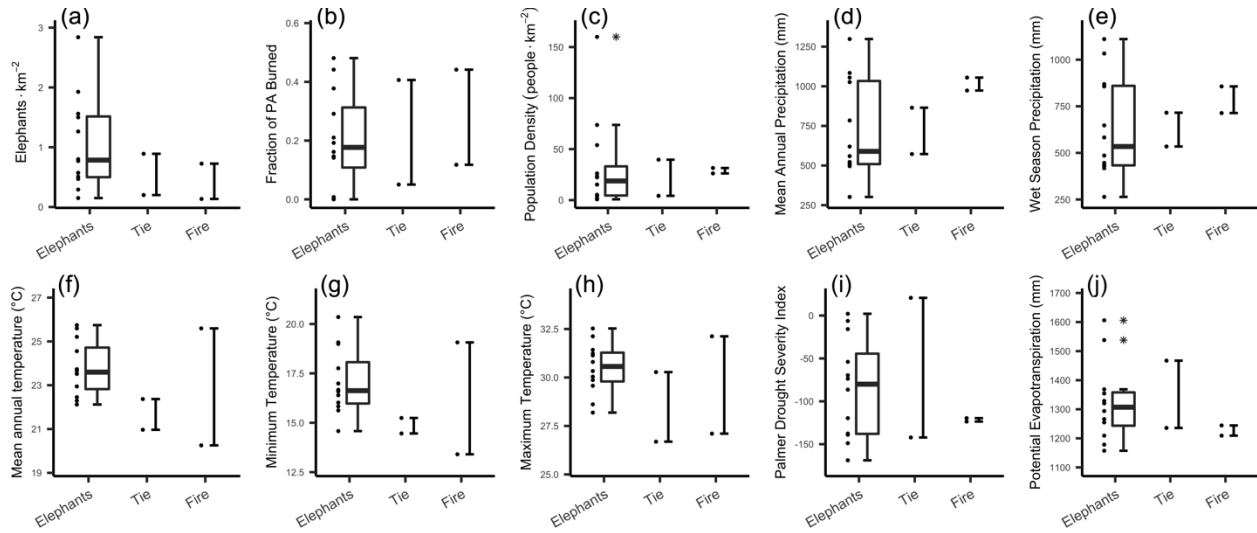


Figure 18. Boxplots of the Variables Used in Analysis by Disturbance Class. While elephant and fire classes did significantly vary in their mean annual precipitation (d) and potential evapotranspiration (j) values when evaluated using Welch's *t*-test, ANOVA using Tukey's honest significant different post-hoc tests showed no difference between all three classes (elephants, fire and tie). Bold centerlines in the boxes represent the median scores, with the boxes encompassing the 2nd and 3rd quartiles and the top and bottom whiskers representing 1.5 times the interquartile range (IQR). Outliers are labeled with asterisks. In classes where only two values existed (i.e. fire and ties), the whiskers represent the minimum and maximum values. Points to the left of each represent individual data values.

Using all 16 locations (including ties), we quantified the importance of variables and their combinations to the primary disturbances in a site using MLR. When we used single variables in MLR, only PET yielded a model with a significant coefficient (Table 7). We then tested all possible two-variable combinations. We limited the models to two variables because, with only 16 sites, we felt that using additional variables in combination would overfit the models. Correlated

variables were not used in combination. The variable combinations produced 14 models that tested significant ($p < 0.05$). Of these, we identified the best model as that using PET and maximum temperatures ($T_{\max}$). The model had the single lowest AIC (26.24) and matched both the second highest Kappa (0.35) and the maximum accuracy achieved by the models (0.81; Table 7).

We calculated and plotted probability surfaces produced using the PET+ $T_{\max}$ model (Figure 19). Across the range of PET and $T_{\max}$ values observed, elephants had the highest probabilities of dominance across the largest area within the two-variable space. This area spanned from PET and $T_{\max}$ values over 1250 mm and 26°C, respectively. Ties were most likely for temperatures above 28°C and PET below 1250 mm. Meanwhile, fire occupied the smallest space, being most probable when temperatures and PET were below approximately 28°C and 1250 mm, respectively (Figure 19).

Table 7. Results of Univariate and Bivariate Models Produced Using MLR. Only significant bivariable models are shown. The fraction of correctly classified sites in each class (Elephants, Fire and Tie) is also listed.

Model	AIC	Accuracy	Kappa	Elephants	Fire	Tie	Significance
ELE	28.886	0.75	0	1	0	0	NS
BurnFrac	31.167	0.75	0	1	0	0	NS
POP	31.379	0.75	0	1	0	0	NS
MAP	29.631	0.75	0	1	0	0	NS
Pwet	30.798	0.75	0	1	0	0	NS
MAT	27.993	0.8125	0.4285	1	0	0.5	NS
T _{min}	28.396	0.75	0.1351	1	0	0	NS
T _{max}	28.962	0.8125	0.3513	1	0	0.5	NS
PDSI	30.304	0.75	0	1	0	0	NS
PET	29.839	0.75	0	1	0	0	S
T _{max} + PET	26.236	0.8125	0.3513	1	0	0.5	S
MAT + PET	28.845	0.75	0.1351	1	0	0	S
ELE + MAP	28.935	0.75	0.2380	0.9167	0.5	0	S
MAT + MAP	29.197	0.8125	0.4286	1	0	0.5	S
T _{min} + MAP	29.253	0.6875	0.0476	0.9167	0	0	S
T _{min} + PET	30.496	0.75	0.1351	1	0	0	S
T _{max} + MAP	30.690	0.8125	0.3513	1	0	0.5	S
ELE + PET	30.817	0.75	0	1	0	0	S
MAT + P _{wet}	30.953	0.8125	0.4286	1	0	0.5	S
T _{max} + P _{wet}	32.129	0.8125	0.3513	1	0	0.5	S
MAP + PET	32.300	0.75	0	1	0	0	S
POP + PET	32.395	0.75	0	1	0	0	S
PET + P _{wet}	33.366	0.75	0	1	0	0	S
BurnFrac + PET	33.597	0.75	0	1	0	0	S

Notes: MLR, multinomial logistic regression; AIC, Akaike information criterion. Models whose variable coefficients have p-values less than 0.05 are classified as significant (S).

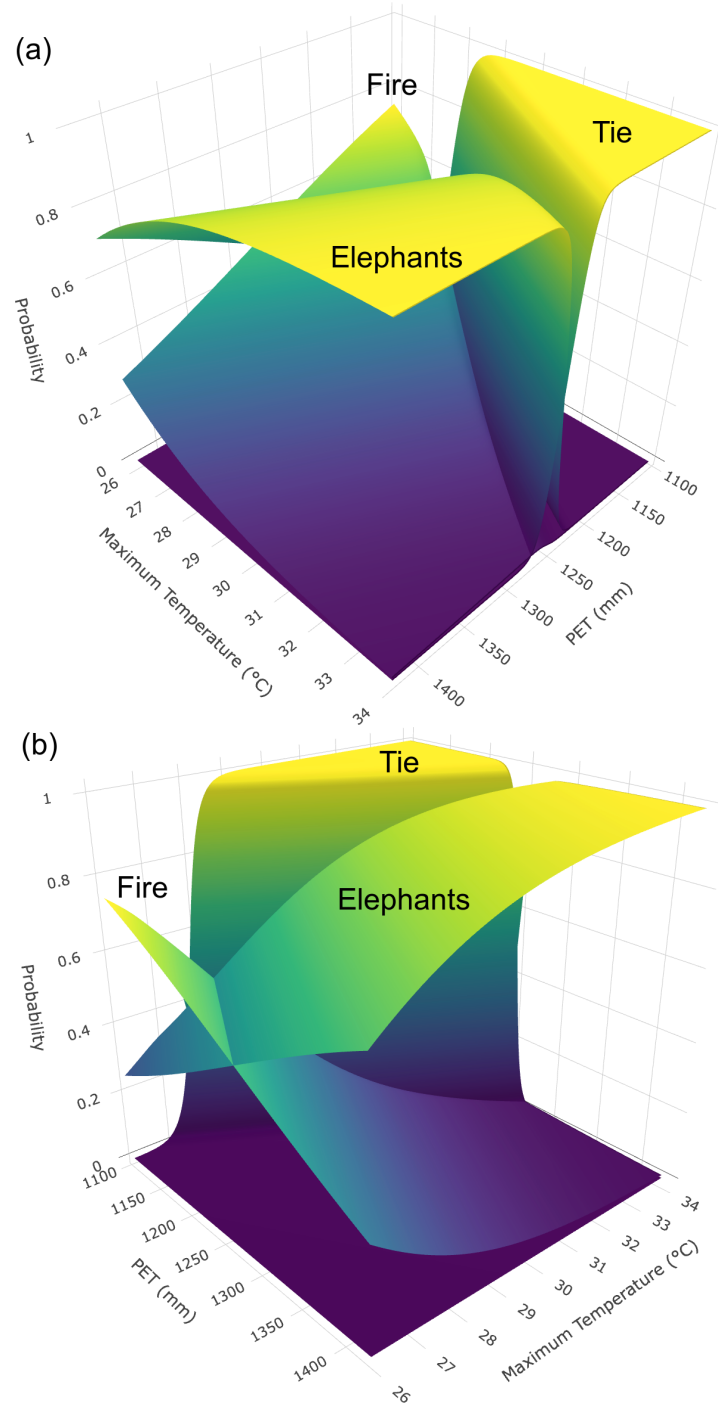


Figure 19. Probability Surfaces Generated in Multinomial Logistic Regression. Surfaces generated using PET and T_{max} in multinomial logistic regression. Each surface depicts the model's assigned probability of dominance for each disturbance (or tie) at each PET and T_{max} combination. Probabilities range from 0.0 (purple) to 1.0 (yellow). Elephants were predicted to be dominant in the largest portion of the variable space, while fire was only dominant at low values of both PET and T_{max} . Plots (a) and (b) depict the same surfaces from different viewing angles to aid interpretation.

Discussion

The dominance of elephants

Our literature search found elephants to be the dominant disturbance across the majority of studies and sites. Our analysis suggests that elephants dominate across a broad swath of environmental conditions (Figure 19), while fire dominates in cooler, wetter savannas (high MAP, low PET, low $T_{\max}$; Figures 18 & 19). Further, differences between elephant- and fire-dominated PAs were best explained by climate variables, not those describing elephant density (ELE) or the average fraction of the PA burned annually (BurnFrac), suggesting climate predisposes savannas to be more or less susceptible to a particular disturbance, regardless of the actual density of elephants or prevalence of fire (Table 7).

These findings are in line with the work of others who have suggested mesic savannas are more likely to produce the fuel loads and the subsequent high-intensity fires required to affect larger woody vegetation (Cochard and Edwards, 2011; Smit et al., 2016; Staver et al., 2011). In more arid savannas, fires predominantly affect trees in the understory (Higgins et al., 2007; Smit et al., 2010), while elephants limit the survivorship of trees and bushes across all height classes (Asner et al., 2016; Guldmond and van Aarde, 2008; Shannon et al., 2011, 2008).

The finding that climate plays a larger role than the amount of fire and elephant density in determining the dominant disturbance is partially supported by other work finding that increases in elephant densities do not lead to significant changes in the total woody cover and/or degradation of the landscape (Guldmond et al., 2017; Owen-Smith et al., 2006). The more likely explanation, however, is that the impacts of elephants and fires are themselves climate-dependent: elephants browse more during the dry season and droughts (Birkett and Stevens-Wood, 2005; Chafota and Owen-Smith, 2009; Laws, 1970), while fires, as mentioned above, burn most intensely in areas with enough precipitation to support high grass biomass. In this way, two PAs with the same levels

of elephants and fire, at least as defined here, would be affected by each differently depending on climate.

Our analysis also showed that elephants are most likely to be dominant across a much broader swath of the environmental gradient than fire (Figure 19). This could be due to the fact that, unlike fire, elephants are capable of adjusting to changes in climate to remain abundant on a landscape. At the same time, protected areas generally promote wildlife while limiting the severity of fires (Reid, 2012), and while elephants were previously thought to enhance fires by opening up bushland (Buechner and Dawkins, 1961; Eltringham, 1979), new evidence suggests they, like other grazing herbivores, overall limit fire through the reduction of grassy fuels (Smit and Archibald, 2019). Compounding these effects, other browsers such as reticulated giraffes (*Giraffa camelopardalis reticulata*) enable elephants by limiting the growth of taller trees and keeping them in height classes more vulnerable to elephant damage, while shorter browsers such as black rhinoceroses (*Diceros bicornis*) compete with elephants for the shorter woody vegetation, likely encouraging them to feed on larger trees and bushes (Birkett, 2002). Altogether, these factors likely contribute to elephants being the dominant disturbance of woody cover (compared to fire) in many of today's protected savannas.

Our findings were similar to those of Sankaran *et al.* (2008), an influential study that used data from 161 field sites across Africa to assess the determinants savanna woody cover. They also found climate played the central role in the myriad factors shaping savanna woody cover. However, unlike us, they found that between elephants and fire, fire had the stronger relationship with woody cover. Unfortunately, because the field sites themselves were not comparing elephant and fire impacts at the level of each site, they could not be included in our analysis. That methodological difference, or their use of a more comprehensive dataset (161 versus 16 data points), might account

for the differences between their results and ours. It could also be that individual field studies comparing the impacts of elephants and fire are not adequately capturing the number of woody cover recruits suppressed by fires. However, it is also possible that fire is the main disturbance across most areas of savanna, except in those areas with sufficiently high elephant populations – areas that may have been disproportionately selected for in the studies we analyze here. These concerns could be addressed by increasing the number and longevity of exclosure experiments that directly manipulate the presence of elephants and fires (e.g., Young *et al.*, 1998; Kraaij & Milton, 2006; Levick *et al.*, 2009).

Management Implications

The sometimes startling impacts of elephants became a central focus of savanna conservationists in the 1960s when many protected areas were experiencing increased elephant populations accompanied by marked declines in woodlands (e.g., Croze, 1974; Lamprey *et al.*, 1967; Napier Bax and Sheldrick, 1963). Concerns of habitat degradation led to annual elephant culls (Bell, 1983; Cumming, 1981, 1983; Owen-Smith *et al.*, 2006; Rodgers & Lobo, 1980; Whyte *et al.*, 1998). For example, management at Kruger National Park in South Africa culled over 16,000 elephants between 1967 and 1994 (Owen-Smith *et al.*, 2006). However, culling was halted in 1994 when there was no strong evidence showing elephants, and not fires, were responsible for the marked changes in woody cover, or that the changes in woody cover qualified as habitat degradation (Owen-Smith *et al.*, 2006).

Our results indicate that elephants likely were responsible for the widespread loss of woodlands, though whether those changes should be considered habitat degradation is likely to remain contentious. However, today, woody encroachment appears to be the bigger threat to savanna habitats (e.g., Gray & Bond, 2013; Smit & Prins, 2015), presenting a potential need for

more elephants, not fewer. Indeed, areas with elephants have already been found to experience reduced rates of encroachment compared to those without (Marston et al., 2017; Skowno et al., 2017; Stevens et al., 2016). This also suggests that prior culling likely made savannas more susceptible to woody encroachment. Therefore, we suggest that in arid to semi-arid savannas – where elephants are more likely to be the dominant disturbance – elephants might be a more effective control of woody encroachment than prescribed burns, providing incentive to elephant conservation efforts. In mesic savannas, fire is likely to be the best control. In neither scenario does there appear to be a need to reduce or eliminate either disturbance, though we acknowledge that the preservation of some woodlands might require the exclusion of fire, elephants, or both. Though not studied here, these results are likely to extend to other herbivores, both wild and domestic, given their similar ability to limit woody cover recruitment and growth rates (Porensky et al., 2013a; Sankaran et al., 2013b, 2008; Styles and Skinner, 2000).

Caveats and concerns

Our meta-analysis, like others (Guldemon and van Aarde, 2008; Guldemon et al., 2017), was limited by the lack of studies, the generally limited data provided by relevant studies, and the unavoidable challenge of comparing studies with different methodologies. Our approach of generalizing findings to the PA might be improved upon by treating each study's field sites as individual data points. For example, Ben-Shahar (1993) provided 16 analyzable points in a single study. However, most studies lacked one or more of the explanatory variables we used here, particularly elephant densities. Filling these gaps in the data using coarser resolution data would likely cause pseudo-replication. For example, if PA-wide elephant densities were used to fill gaps in the data, all points within the same PA using those values would become pseudo-replicates and would need to be reduced to a single point in analysis, similar to what we did here. Further, studies

like Ben-Shahar (1993) are generally confined to a relatively small geographic area and would likely count for a disproportionate amount of the data, thereby biasing the results.

A substantial regional bias also existed in the data. While our initial set of 66 studies included sites in western and central Africa (Figure 17a), the majority of studies, and all the quantified studies, were in eastern and southern Africa. This is likely largely due to the region being home to many of the last major populations of elephants occurring in Africa (Thouless et al., 2016), but there is a clear need for more research outside the region. For example, both Waza National Park in Cameroon and the Nazinga Ecosystem in Burkina Faso are located outside the region and are already the sites of descriptive comparisons between elephant and fire impacts (Hema et al., 2017; Okula and Sise, 1986). Further, both PAs appear to be undergoing changes in their elephant populations that could present natural experiments: the Nazinga Ecosystem's population increased by an estimated 53% from the years 2003 to 2012, while Waza NP likely lost many of its elephants while, from 2014-2016, it was occupied by Boko Haram – a terrorist organization who has been known to use ivory sales to fund its activities (Thouless et al., 2016).

In addition to the regional bias, much of the research on disturbances is likely to be biased toward selecting sites in areas where elephants are known to be active or where there are significant management concerns of complete tree cover removal by elephants (e.g., Rodgers & Lobo, 1980; Bell, 1983; Whyte *et al.*, 1998). Meanwhile, areas with more precipitation and less risk of losing their tree cover – areas also more likely to be dominated by fire – might have attracted less research. The influence of site selection was demonstrated by Ben-Shahar (1993), who found that within the same protected area the dominant disturbance can switch from elephants to fire across relatively short distances, the maximum of which was 40 km. Their work demonstrates the problematic

nature of assuming site-specific findings can be generalized to an entire PA, as we do here due to the constraints of comparing multiple studies employing different methodologies.

Despite these limitations, our analysis yielded results that are in line with much of the previous research of savanna disturbances, demonstrating a growing consensus that in drier savannas without the high vegetative biomass needed to support high-intensity fires, elephants are the primary disturbance to woody vegetation. However, contradictions with studies such as Sankaran *et al.* (2008) will likely require similarly large-scale studies that manipulate elephant and fire occurrence (i.e. exclosures), or remote sensing studies that adequately capture both the spatial resource use of elephants (i.e. through the use of GPS collars) and high spatial and temporal resolution fire and woody cover data – data that is likely to become increasingly available via a growing number of very high-resolution satellites. Such data could also differentiate the impacts to bushes and trees – something we did not account for given the limited number of studies available for analysis.

Conclusion

Savannas and the factors shaping them have significant impacts on the global carbon cycle (Ahlström *et al.*, 2015), along with the millions of people and wildlife that rely on them (Reid, 2012; Ripple *et al.*, 2015). The goal of this meta-analysis was to compare the relative impacts of elephants and fire on the woody cover of Africa's savannas. We found elephants to be the dominant disturbance in the majority of quantified studies (30 of 39; 76.9%) and sites (12 of 16; 75.0%). Further, rather than being driven by elephant densities or the extent of fires, dominance was most closely linked to climate, with cooler and wetter sites more likely to be dominated by fire. Our findings indicate that the management of woody encroachment should be based on climate, with fire likely being the most effective control agent in mesic savannas and elephants in semi-arid to

arid savannas. However, our work was significantly hampered by the low number of quantified comparisons of elephant and fire disturbances, and future research should focus on increasing the use and number of exclosure experiments, particularly in central and western Africa. By untangling the relative roles of elephants and fire, while also better understanding the variables that impact them individually, we will be able to both better manage and predict changes in these globally relevant, dynamic ecosystems while also gaining insight into the potential consequences of the continued loss of the world's megaherbivores.

CONCLUSIONS

Summary and Opportunities for Future Research

For decades, researchers have clearly documented the local effects of elephants on woody cover (Asner et al., 2016; Malhi et al., 2016; Shannon et al., 2008), while the large-scale effects have remained unclear (Duffy and Pettorelli, 2012; Guldemon et al., 2017; Hayward and Zawadzka, 2010; Sankaran et al., 2008). This dissertation presents several new findings that advance both the methods used to study savannas and our understanding of the factors that shape them.

Chapter 1 presents an advance in our ability to map savanna woody cover, finding that approaches using Random Forests outperform those using simple linear regression or spectral unmixing. The chapter also demonstrates that while the use of multiple Landsat images per year is ideal, even single images can be used to produce accurate maps – a key finding in an area of the world where cloud cover and limited data availability often limit the amount of imagery available to mapping efforts. Chapter 1 also establishes the transition between the wet and dry season as the ideal time of year for imagery used in mapping woody cover, while strong relationships with woody cover demonstrate that NDVI, BC and the RNS bands should be included in any additional mapping efforts. Last, given that the use of training data from all the sites created the most accurate models and the best general model across all the PAs did not significantly decrease accuracies, it is likely that our approach and training data can be applied to accurately map woody cover in additional sites. However, ideally an approach would be able to map all of Africa's savannas and future work should focus on the logistical and technical challenges to this effort.

Chapter 2 uses the methodology developed in Chapter 1 to get at the crux of this dissertation, directly assessing whether elephant densities relate to woody cover across the PAs. Instead of

elephants, the findings largely support the argument that climate is the primary determinant of a biome's woody cover (Polis, 1999): wet season precipitation had the strongest relationship with woody cover in all areas studied. However, the unexplained low woody cover near permanent water sources, along with relationships between elephant densities and increased woody cover on slopes far from water sources, suggest further research might reveal elephants, or other disturbances, determine the woody cover in these still expansive areas. Further, because the analysis did not include areas without elephants, it is possible that the woody cover of the PAs studied, while having a strong relationship with climate, nonetheless had woody cover universally lower woody cover than areas free of disturbances – one of the key arguments made by Bond (2005). This could be addressed by including disturbance-free sites in a similar analysis, or perhaps through the analysis of errors from an uncalibrated Earth systems model (ESM), with the assumption that the ESM will overpredict woody cover in savannas.

Chapter 3 reveals that while climate appears to be the main determinant of woody cover in savannas, elephants are the primary disturbance across a broad swath of environmental conditions, while fire dominates in wetter and cooler climates. This suggests that if disturbances are creating the areas of low woody cover near water found in Chapter 2, elephants would likely be the responsible disturbance under most climate conditions. The findings also imply that woody encroachment might be partially addressed by increasing elephant numbers, while the continued wide-scale loss of elephant populations enables woody encroachment. However, the analysis was limited by the overall lack of quantitative studies, particularly outside of eastern and southern Africa – gaps that might be best addressed through the additional use of exclosure experiments or future remote sensing studies using detailed information describing the locations of both fires and elephants – data which is especially lacking for elephants.

Broader Impacts

The world's megafauna are disappearing (Ripple et al., 2015). As they do, the ecosystems, national revenues and local livelihoods they support face detrimental consequences (Gray and Bond, 2013; Ripple et al., 2015; Smit and Prins, 2015). In the case of elephants, the consequences are especially dire due to a potential positive feedback loop: as elephants disappear, increasingly woody PAs will become less suitable for wildlife viewing, reducing the tourism revenue that helps fund elephant conservation. That, in turn, may exacerbate the decline in elephants, causing increased declines in open savanna and savanna-dependent wildlife. Accordingly, the decline and potential extinction of elephants might precipitate the extinction of dependent species – much like what is thought to have happened when the Pleistocene megafauna went extinct (Owen-Smith, 1987). This dissertation suggests that while elephants are the dominant disturbance across many of today's protected savannas and likely play key roles in limiting woody cover, particularly near water sources, climate is the primary determinate of both woody cover and the relative roles of disturbances in African savannas. If true, then the woody cover of savannas and other climatically determined ecosystems might be more resilient to the impacts of megafauna than previously thought. In turn, this suggests megafauna are unlikely to mitigate habitat loss due to a changing climate.

APPENDICES

APPENDIX A

Chapter 1 Supplementary Material

Table A.1. Variables input to random forests and regression models. Each variable name and the respective equation used in the models. All equations used bands from Landsat imagery.

Variable	Definition
dryB2	Dry season's Landsat 8 Band 2 value / 10,000
dryB3	Dry season's Landsat 8 Band 3 value / 10,000
dryB4	Dry season's Landsat 8 Band 4 value / 10,000
dryB5	Dry season's Landsat 8 Band 5 value / 10,000
dryB6	Dry season's Landsat 8 Band 6 value / 10,000
dryB7	Dry season's Landsat 8 Band 7 value / 10,000
wetB2	Wet season's Landsat 8 Band 2 value / 10,000
wetB3	Wet season's Landsat 8 Band 3 value / 10,000
wetB4	Wet season's Landsat 8 Band 4 value / 10,000
wetB5	Wet season's Landsat 8 Band 5 value / 10,000
wetB6	Wet season's Landsat 8 Band 6 value / 10,000
wetB7	Wet season's Landsat 8 Band 7 value / 10,000
tranB2	Transition season's Landsat 8 Band 2 value / 10,000
tranB3	Transition season's Landsat 8 Band 3 value / 10,000
tranB4	Transition season's Landsat 8 Band 4 value / 10,000
tranB5	Transition season's Landsat 8 Band 5 value / 10,000
tranB6	Transition season's Landsat 8 Band 6 value / 10,000
tranB7	Transition season's Landsat 8 Band 7 value / 10,000
dwtB2mean	$(\text{dryB2} + \text{wetB2} + \text{tranB2}) / 3$
dwB2mean	$(\text{dryB2} + \text{wetB2}) / 2$
wtB2mean	$(\text{wetB2} + \text{tranB2}) / 2$
tdB2mean	$(\text{dryB2} + \text{tranB2}) / 2$
dwtB3mean	$(\text{dryB3} + \text{wetB3} + \text{tranB3}) / 3$
dwB3mean	$(\text{dryB3} + \text{wetB3}) / 2$
wtB3mean	$(\text{wetB3} + \text{tranB3}) / 2$
tdB3mean	$(\text{dryB3} + \text{tranB3}) / 2$
dwtB4mean	$(\text{dryB4} + \text{wetB4} + \text{tranB4}) / 3$
dwB4mean	$(\text{dryB4} + \text{wetB4}) / 2$
wtB4mean	$(\text{wetB4} + \text{tranB4}) / 2$
tdB4mean	$(\text{dryB4} + \text{tranB4}) / 2$
dwtB5mean	$(\text{dryB5} + \text{wetB5} + \text{tranB5}) / 3$
dwB5mean	$(\text{dryB5} + \text{wetB5}) / 2$
wtB5mean	$(\text{wetB5} + \text{tranB5}) / 2$
tdB5mean	$(\text{dryB5} + \text{tranB5}) / 2$
dwtB6mean	$(\text{dryB6} + \text{wetB6} + \text{tranB6}) / 3$
dwB6mean	$(\text{dryB6} + \text{wetB6}) / 2$
wtB6mean	$(\text{wetB6} + \text{tranB6}) / 2$

Table A.1. (cont'd)

tdB6mean	$(\text{dryB6} + \text{tranB6}) / 2$
dwtB7mean	$(\text{dryB7} + \text{wetB7} + \text{tranB7}) / 3$
dwB7mean	$(\text{dryB7} + \text{wetB7}) / 2$
wtB7mean	$(\text{wetB7} + \text{tranB7}) / 2$
tdB7mean	$(\text{dryB7} + \text{tranB7}) / 2$
dwB2diff	$(\text{dryB2} - \text{wetB2}) / (\text{dryB2} + \text{wetB2})$
wtB2diff	$(\text{tranB2} - \text{wetB2}) / (\text{tranB2} + \text{wetB2})$
tdB2diff	$(\text{dryB2} - \text{tranB2}) / (\text{dryB2} + \text{tranB2})$
dtwB2diff	$\text{dwB2diff} - \text{wtB2diff}$
dwB3diff	$(\text{dryB3} - \text{wetB3}) / (\text{dryB3} + \text{wetB3})$
wtB3diff	$(\text{tranB3} - \text{wetB3}) / (\text{tranB3} + \text{wetB3})$
tdB3diff	$(\text{dryB3} - \text{tranB3}) / (\text{dryB3} + \text{tranB3})$
dtwB3diff	$\text{dwB3diff} - \text{wtB3diff}$
dwB4diff	$(\text{dryB4} - \text{wetB4}) / (\text{dryB4} + \text{wetB4})$
wtB4diff	$(\text{tranB4} - \text{wetB4}) / (\text{tranB4} + \text{wetB4})$
tdB4diff	$(\text{dryB4} - \text{tranB4}) / (\text{dryB4} + \text{tranB4})$
dtwB4diff	$\text{dwB4diff} - \text{wtB4diff}$
dwB5diff	$(\text{dryB5} - \text{wetB5}) / (\text{dryB5} + \text{wetB5})$
wtB5diff	$(\text{tranB5} - \text{wetB5}) / (\text{tranB5} + \text{wetB5})$
tdB5diff	$(\text{dryB5} - \text{tranB5}) / (\text{dryB5} + \text{tranB5})$
dtwB5diff	$\text{dwB5diff} - \text{wtB5diff}$
dwB6diff	$(\text{dryB6} - \text{wetB6}) / (\text{dryB6} + \text{wetB6})$
wtB6diff	$(\text{tranB6} - \text{wetB6}) / (\text{tranB6} + \text{wetB6})$
tdB6diff	$(\text{dryB6} - \text{tranB6}) / (\text{dryB6} + \text{tranB6})$
dtwB6diff	$\text{dwB6diff} - \text{wtB6diff}$
dwB7diff	$(\text{dryB7} - \text{wetB7}) / (\text{dryB7} + \text{wetB7})$
wtB7diff	$(\text{tranB7} - \text{wetB7}) / (\text{tranB7} + \text{wetB7})$
tdB7diff	$(\text{dryB7} - \text{tranB7}) / (\text{dryB7} + \text{tranB7})$
dtwB7diff	$\text{dwB7diff} - \text{wtB7diff}$
dryNDVI	Dry season image's NDVI*
wetNDVI	Wet season image's NDVI*
tranNDVI	Transition season's NDVI*
dryMSAVI2	Dry season image's MSAVI2*
wetMSAVI2	Wet season image's MSAVI2*
tranMSAVI2	Transition season's MSAVI2*
drySNDI	Dry season image's SNDI*
wetSNDI	Wet season image's SNDI*
tranSNDI	Transition season's SNDI*
drySATVI	Dry season image's SATVI*

Table A.1. (cont'd)

wetSATVI	Wet season image's SATVI*
tranSATVI	Transition season's SATVI*
dtwNDVI _{mean}	$(\text{dryNDVI} + \text{wetNDVI} + \text{tranNDVI}) / 3$
dwNDVI _{mean}	$(\text{dryNDVI} + \text{wetNDVI}) / 2$
wtNDVI _{mean}	$(\text{wetNDVI} + \text{tranNDVI}) / 2$
tdNDVI _{mean}	$(\text{dryNDVI} + \text{tranNDVI}) / 2$
dtwMSAVI2 _{mean}	$(\text{dryMSAVI2} + \text{wetMSAVI2} + \text{tranMSAVI2}) / 3$
dwMSAVI2 _{mean}	$(\text{dryMSAVI2} + \text{wetMSAVI2}) / 2$
wtMSAVI2 _{mean}	$(\text{wetMSAVI2} + \text{tranMSAVI2}) / 2$
tdMSAVI2 _{mean}	$(\text{dryMSAVI2} + \text{tranMSAVI2}) / 2$
dtwSNDI _{mean}	$(\text{drySNDI} + \text{wetSNDI} + \text{tranSNDI}) / 3$
dwSNDI _{mean}	$(\text{drySNDI} + \text{wetSNDI}) / 2$
wtSNDI _{mean}	$(\text{wetSNDI} + \text{tranSNDI}) / 2$
tdSNDI _{mean}	$(\text{drySNDI} + \text{tranSNDI}) / 2$
dtwSATVI _{mean}	$(\text{drySATVI} + \text{wetSATVI} + \text{tranSATVI}) / 3$
dwSATVI _{mean}	$(\text{drySATVI} + \text{wetSATVI}) / 2$
wtSATVI _{mean}	$(\text{wetSATVI} + \text{tranSATVI}) / 2$
tdSATVI _{mean}	$(\text{drySATVI} + \text{tranSATVI}) / 2$
dwNDVIDiff	$(\text{wetNDVI} - \text{dryNDVI}) / (\text{wetNDVI} + \text{dryNDVI})$
wtNDVIDiff	$(\text{wetNDVI} - \text{tranNDVI}) / (\text{wetNDVI} + \text{tranNDVI})$
tdNDVIDiff	$(\text{tranNDVI} - \text{dryNDVI}) / (\text{tranNDVI} + \text{dryNDVI})$
dtwNDVIDiff	$\text{dwNDVIDiff} - \text{wtNDVIDiff}$
dwMSAVI2diff	$(\text{wetMSAVI2} - \text{dryMSAVI2}) / (\text{wetMSAVI2} + \text{dryMSAVI2})$
wtMSAVI2diff	$(\text{wetMSAVI2} - \text{tranMSAVI2}) / (\text{wetMSAVI2} + \text{tranMSAVI2})$
tdMSAVI2diff	$(\text{tranMSAVI2} - \text{dryMSAVI2}) / (\text{tranMSAVI2} + \text{dryMSAVI2})$
dtwMSAVI2diff	$\text{dwMSAVI2diff} - \text{wtMSAVI2diff}$
dwSNDIdiff	$(\text{wetSNDI} - \text{drySNDI}) / (\text{wetSNDI} + \text{drySNDI})$
wtSNDIdiff	$(\text{wetSNDI} - \text{tranSNDI}) / (\text{wetSNDI} + \text{tranSNDI})$
tdSNDIdiff	$(\text{tranSNDI} - \text{drySNDI}) / (\text{tranSNDI} + \text{drySNDI})$
dtwSNDIdiff	$\text{dwSNDIdiff} - \text{wtSNDIdiff}$
dwSATVIDiff	$(\text{wetSATVI} - \text{drySATVI}) / (\text{wetSATVI} + \text{drySATVI})$
wtSATVIDiff	$(\text{wetSATVI} - \text{tranSATVI}) / (\text{wetSATVI} + \text{tranSATVI})$
tdSATVIDiff	$(\text{tranSATVI} - \text{drySATVI}) / (\text{tranSATVI} + \text{drySATVI})$
dtwSATVIDiff	$\text{dwSATVIDiff} - \text{wtSATVIDiff}$
dryBright	$(\text{dryB2} + \text{dryB3} + \text{dryB4}) / 3$
wetBright	$(\text{wetB2} + \text{wetB3} + \text{wetB4}) / 3$
tranBright	$(\text{tranB2} + \text{tranB3} + \text{tranB4}) / 3$
dtwBright _{mean}	$(\text{dryBright} + \text{wetBright} + \text{tranBright}) / 3$
dwBright _{mean}	$(\text{dryBright} + \text{wetBright}) / 2$

Table A.1. (cont'd)

wtBrightmean	$(\text{wetBright} + \text{tranBright}) / 2$
tdBrightmean	$(\text{dryBright} + \text{tranBright}) / 2$
dwBrightdiff	$(\text{dryBright} - \text{wetBright}) / (\text{dryBright} + \text{wetBright})$
wtBrightdiff	$(\text{tranBright} - \text{wetBright}) / (\text{tranBright} + \text{wetBright})$
tdBrightdiff	$(\text{dryBright} - \text{tranBright}) / (\text{dryBright} + \text{tranBright})$
dtwBrightdiff	$\text{dwBrightdiff} - \text{wtBrightdiff}$
dryBC	$(\text{mean}(\text{dryBright}_{\text{all}}) - \text{dryBright}_i) / (\text{mean}(\text{dryBright}_{\text{all}}) + \text{dryBright}_i)$
wetBC	$(\text{mean}(\text{wetBright}_{\text{all}}) - \text{wetBright}_i) / (\text{mean}(\text{wetBright}_{\text{all}}) + \text{wetBright}_i)$
tranBC	$(\text{mean}(\text{tranBright}_{\text{all}}) - \text{tranBright}_i) / (\text{mean}(\text{tranBright}_{\text{all}}) + \text{tranBright}_i)$
dtwBCmean	$(\text{dryBC} + \text{wetBC} + \text{tranBC}) / 3$
dwBCmean	$(\text{dryBC} + \text{wetBC}) / 2$
wtBCmean	$(\text{wetBC} + \text{tranBC}) / 2$
tdBCmean	$(\text{dryBC} + \text{tranBC}) / 2$
dwBCdiff	$(\text{wetBC} - \text{dryBC}) / (\text{wetBC} + \text{dryBC})$
wtBCdiff	$(\text{wetBC} - \text{tranBC}) / (\text{wetBC} + \text{tranBC})$
tdBCdiff	$(\text{tranBC} - \text{dryBC}) / (\text{tranBC} + \text{dryBC})$
dtwBCdiff	$\text{dwBCdiff} - \text{wtBCdiff}$

* Index formulas are listed in Table 2.

Figure A.1. Ranked map accuracies. We ranked and plotted all 4842 map accuracies in search of an inflection point, i.e., where gains in accuracies from map to map become relatively consistent. We found this point around a VEcv of -500% and removed all the maps with accuracies below that value from the analysis.

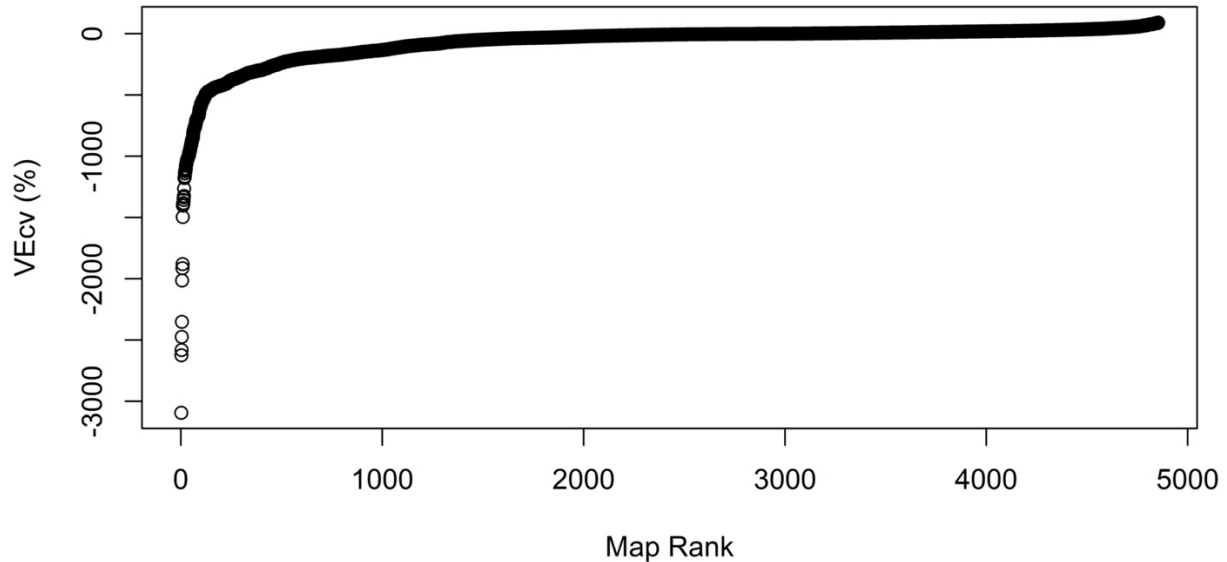


Figure A.2. Relationship across seasons between best PA map accuracies and PA average woody cover, mean annual precipitation (MAP) and precipitation seasonality. Rather than one season's image doing well in woodier, drier or more seasonal of PAs and another season doing the opposite, we found that nearly universally woodier PAs yielded the more accurate maps. More generally, neither of the three variables predict which seasonal image will produce the best maps in a PA. Only significant relationships have regression lines plotted. Numbers in the plots correspond to the PA order in Table 1: MUR (1), MPA (2), QUE (3), SMR (4), RUA (5), SEL (6), NLU (7), SLU (8), CHO (9), LIM (10), KRU (11).

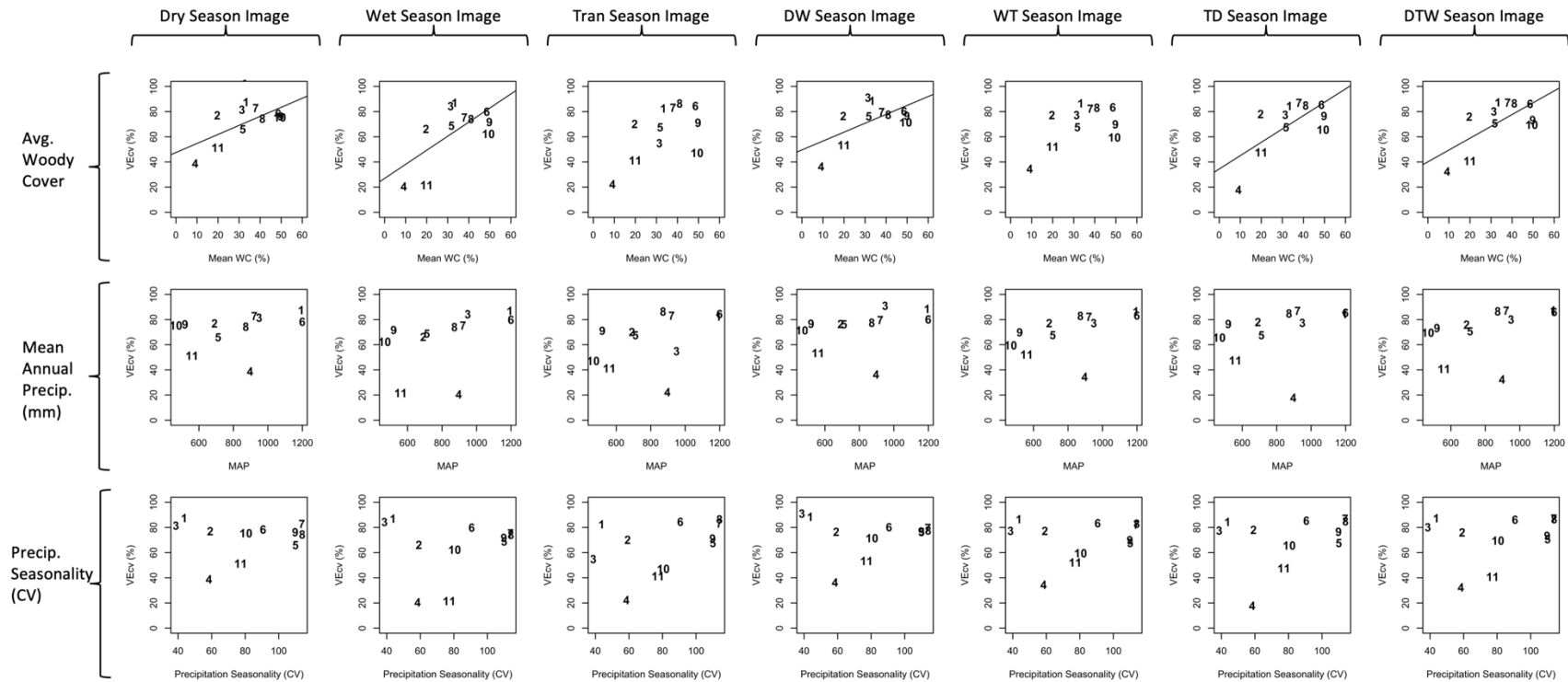


Figure A.3. Relationship between best PA map accuracies and PA average woody cover, mean annual precipitation, and precipitation seasonality. Only woody cover was a significant predictor of accuracy ($p < 0.05$), meaning woodier PAs are generally easier to map. Numbers in the plots correspond to the PA order in Table 1: MUR (1), MPA (2), QUE (3), SMR (4), RUA (5), SEL (6), NLU (7), SLU (8), CHO (9), LIM (10), KRU (11).

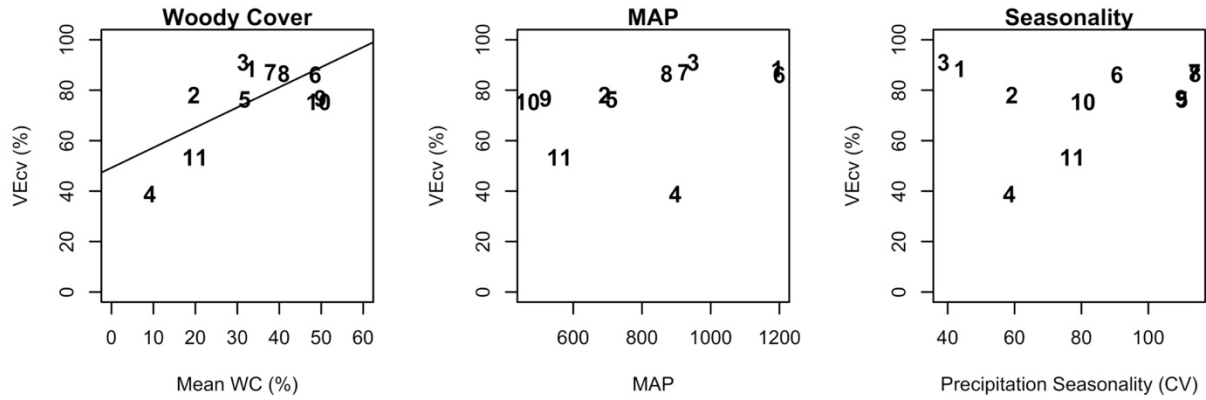


Figure A.4. Accuracy of maps produced using each of 143 variables in regression. The scores are from maps of all the protected areas (PAs), including the PA formed by combining all the other PAs into one. This means that each box represents 12 scores, one for each of the PAs. The mean NDVI of the Tran and Dry composite (TD mean NDVI) performed the best and was the only variable to significantly ($p < 0.05$) outperform some (58) of the 143 total variables. In the plot, the bold centerline represents the median score, the box encompasses the 2nd and 3rd quartiles, and the top and bottom whiskers respectively represent the largest and smallest values within 1.5 times the interquartile range. Values outside that range are marked as outliers. RNS = multiple regression using bands 4, 5 and 6 (red, near infrared, first shortwave infrared band).

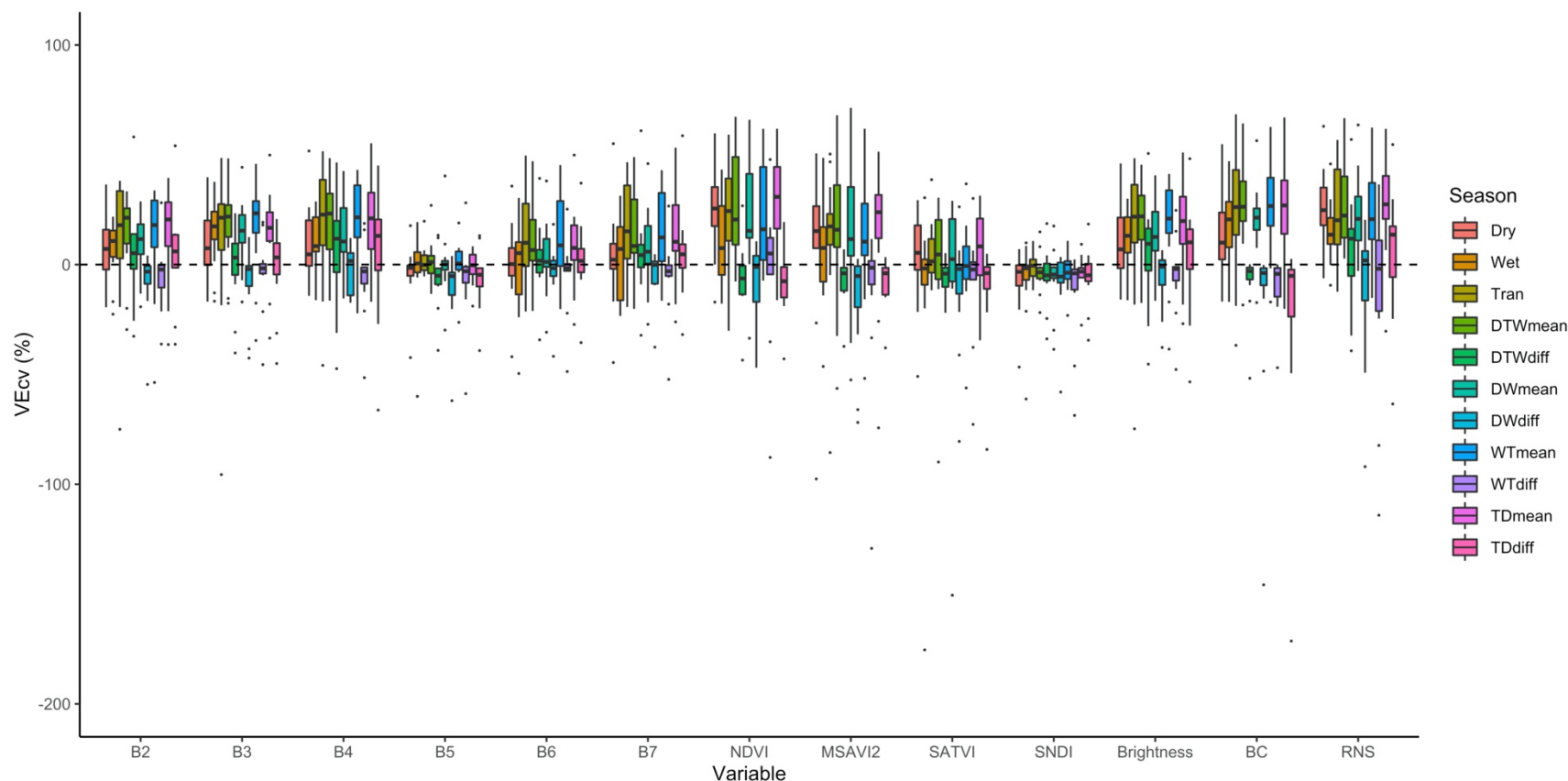


Figure A.5. The best map produced for Chobe National Park. The map was produced using a random forest model trained using training data from all the PAs and variables from the dry and wet season composite (RF-ALL-DW; “Best Overall Model” column in Table 4 in Chapter 1). The map had a VEcv of 76.8%. The TIFF file of this map can be found in the online dataset (see Nagelkirk and Dahlin, 2019). Gray areas are clouds (top) and sections of the PA not mapped because they fell within a path that did not cover more than 10% of the PA (eastern edge). The scatterplot shows the regression line (solid red line) and equation relating the training data (“Reference”) to the mapped values (“Predicted”), along with R^2 and VEcv metrics. The dashed line represents the 1:1 line.

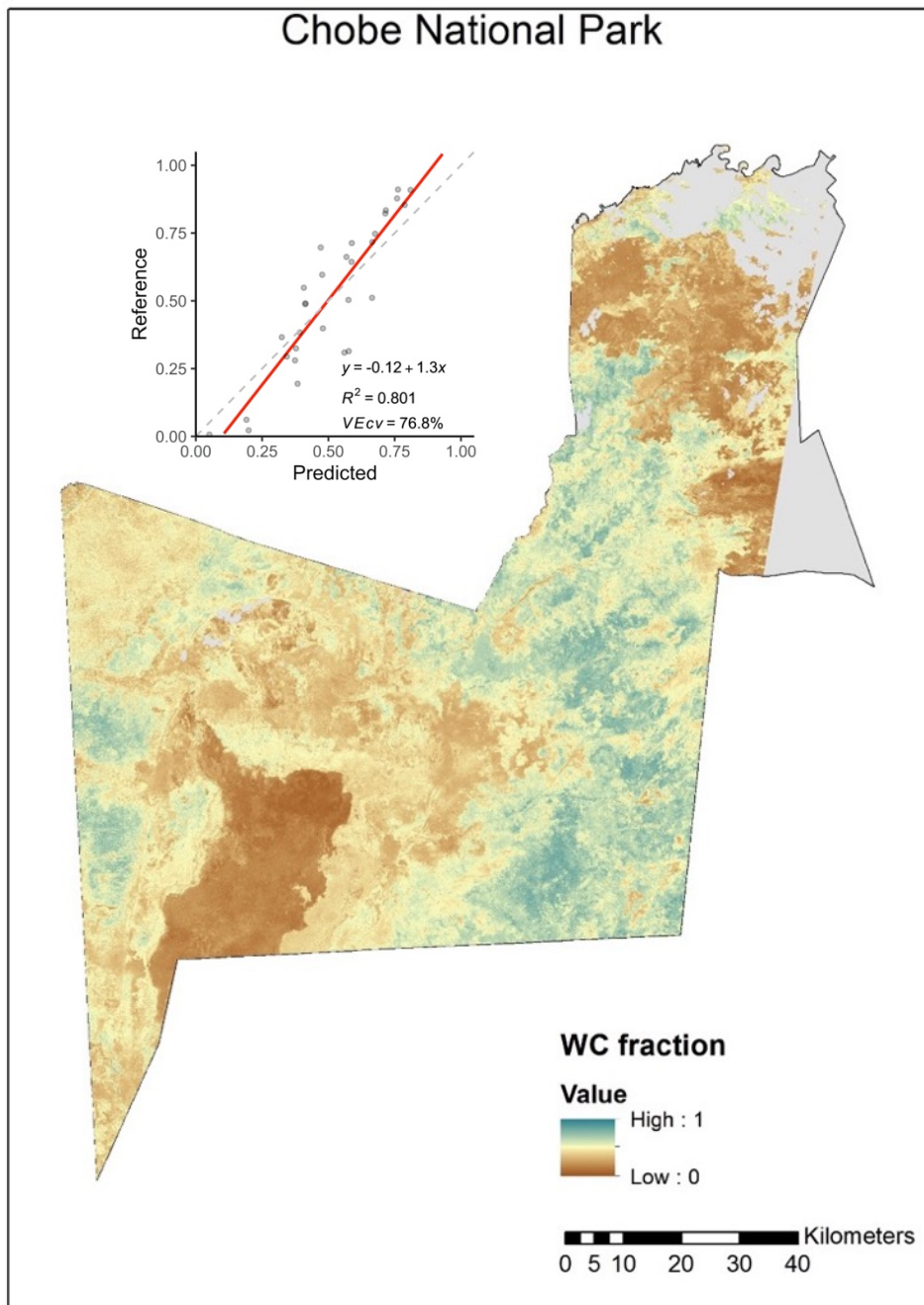


Figure A.6. The best maps produced for Kruger and Limpopo National Parks. The maps of Kruger (left) and Limpopo (right) were produced using random forest models trained using training data from all the PAs and variables from the dry and wet season composite (Kruger) and the dry season (Limpopo) (RF-ALL-DW and RF-ALL-Dry, respectively; “Best Overall Model” column in Table 4 in Chapter 1). The maps had a VECv of 53.4% and 75.4%, respectively. The TIFF files of these maps can be found in the online dataset (see Nagelkirk and Dahlin, 2019). Gray areas are clouds. The scatterplots show the regression lines (solid red lines) and equations relating the training data (“Reference”) to the mapped values (“Predicted”), along with R^2 and VECv metrics. The dashed lines represent the 1:1 line.

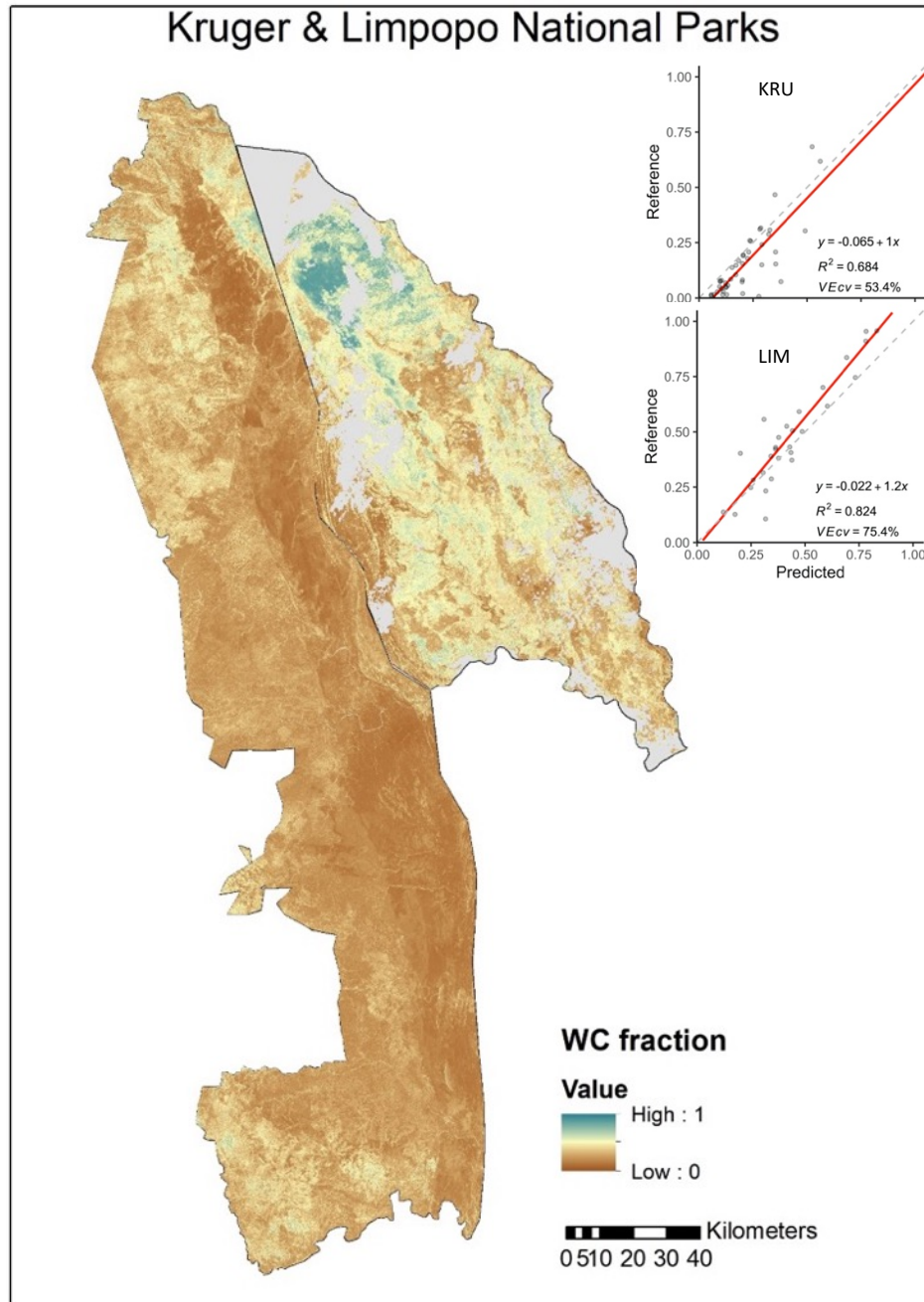


Figure A.7. The best map produced for Mpala Research Center. The map was produced using a random forest model trained using training data from all the PAs and variables from the transition and dry season composite (RF-ALL-TD; “Best Overall Model” column in Table 4 in Chapter 1). The map had a VECv of 78.1%. The TIFF file of this map can be found in the online dataset (see Nagelkirk and Dahlin, 2019). Gray areas are clouds. The scatterplot shows the regression line (solid red line) and equation relating the training data (“Reference”) to the mapped values (“Predicted”), along with R^2 and VECv metrics. The dashed line represents the 1:1 line.

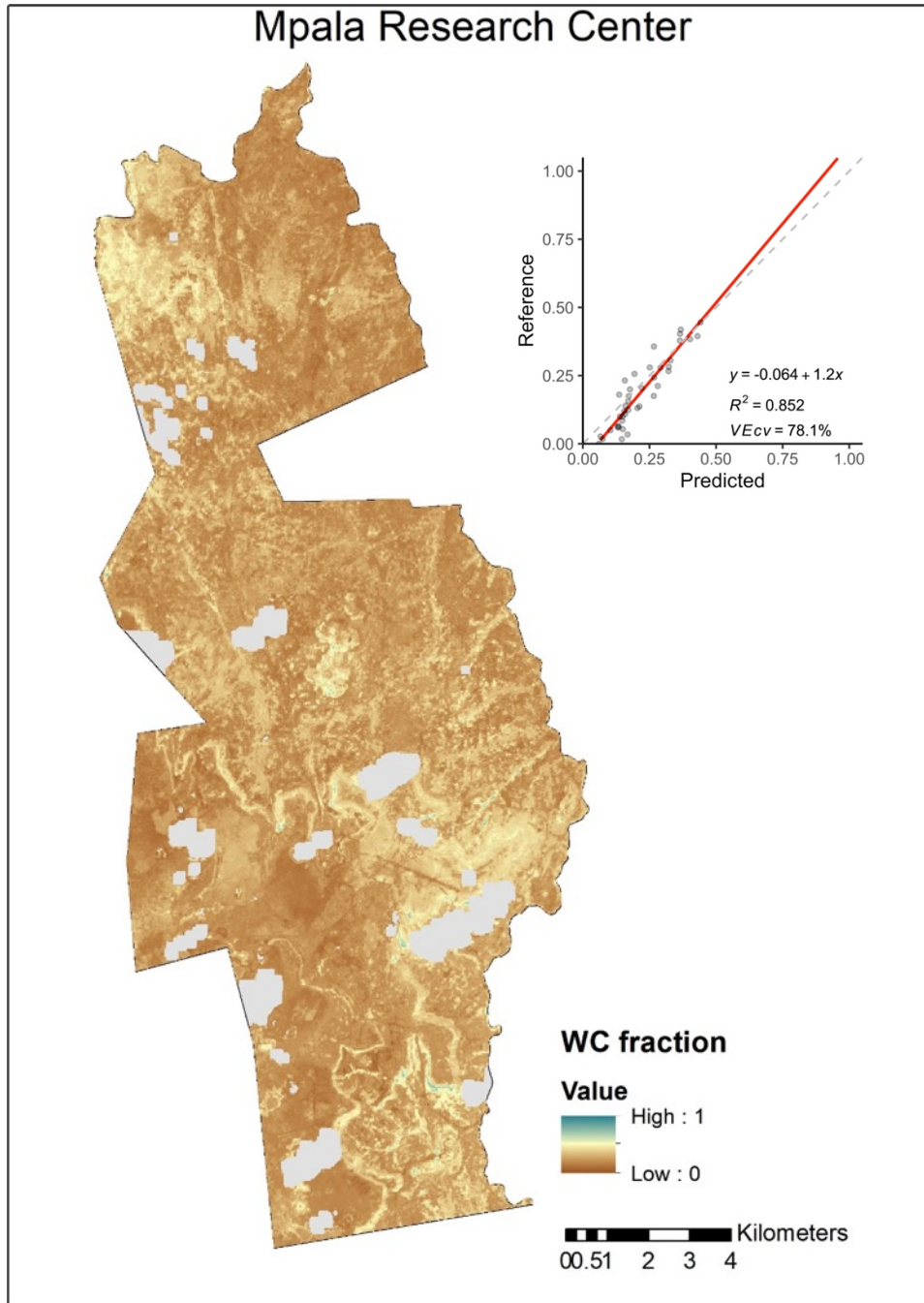


Figure A.8. The best maps produced for North and South Luangwa National Parks. Both maps were produced using random forest models trained using training data from all the PAs and variables from the dry, transition and wet season composite (RF-ALL-DTW; “Best Overall Model” column in Table 4 in Chapter 1). The maps had a VECv of 87.1% and 86.6%, respectively. The TIFF files of these maps can be found in the online dataset (see Nagelkirk and Dahlin, 2019). Gray areas are clouds. The scatterplots show the regression lines (solid red lines) and equations relating the training data (“Reference”) to the mapped values (“Predicted”), along with R^2 and VECv metrics. The dashed lines represent the 1:1 line.

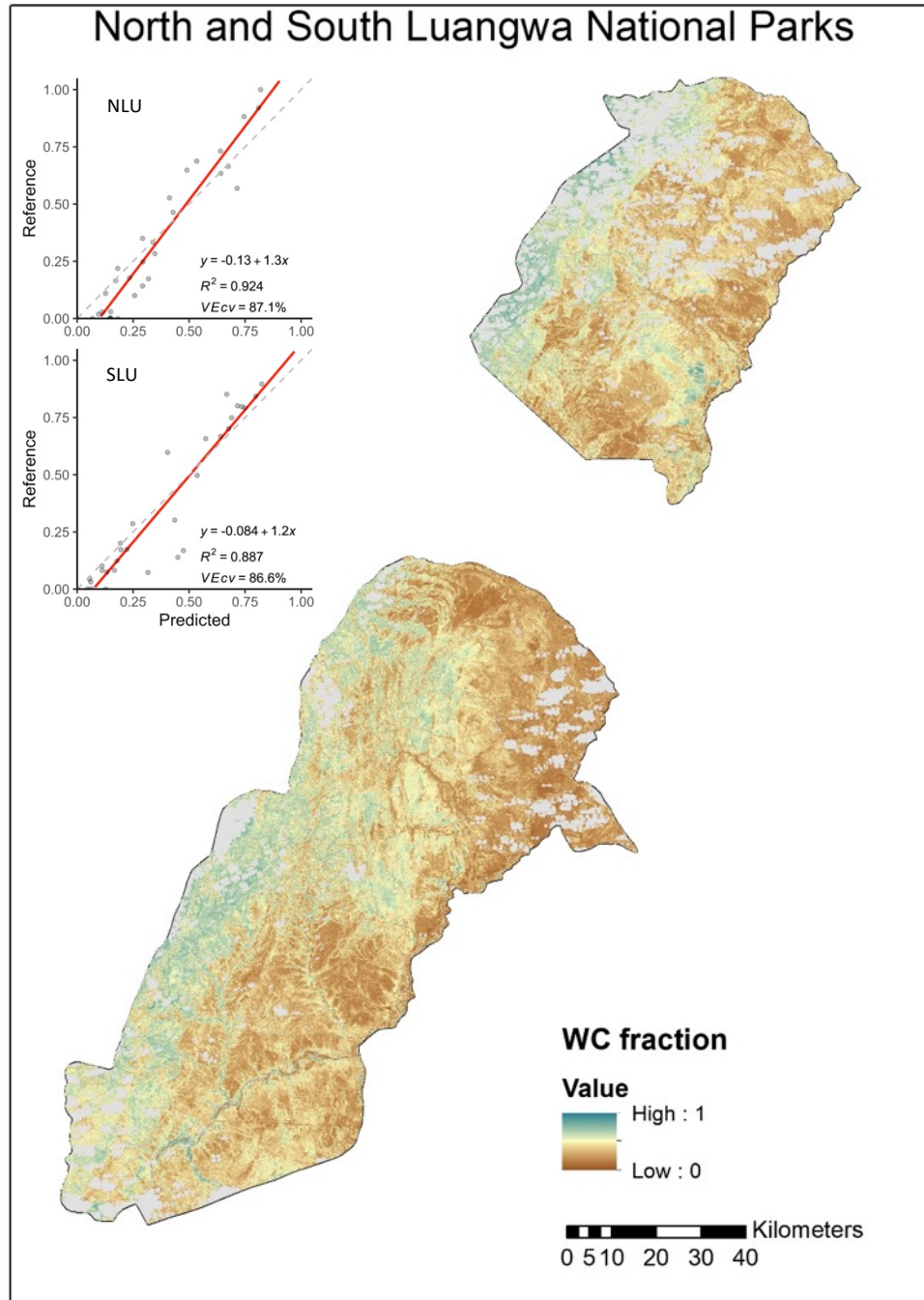


Figure A.9. The best map produced for Queen Elizabeth National Park. The map was produced using a random forest model trained using training data from all the PAs and variables from the dry and wet season composite (RF-ALL-DW; “Best Overall Model” column in Table 4 in Chapter 1). The map had a VEcv of 91.1%. The TIFF file of this map can be found in the online dataset (see Nagelkirk and Dahlin, 2019). Gray areas are clouds. The scatterplot shows the regression line (solid red line) and equation relating the training data (“Reference”) to the mapped values (“Predicted”), along with R^2 and VEcv metrics. The dashed line represents the 1:1 line.

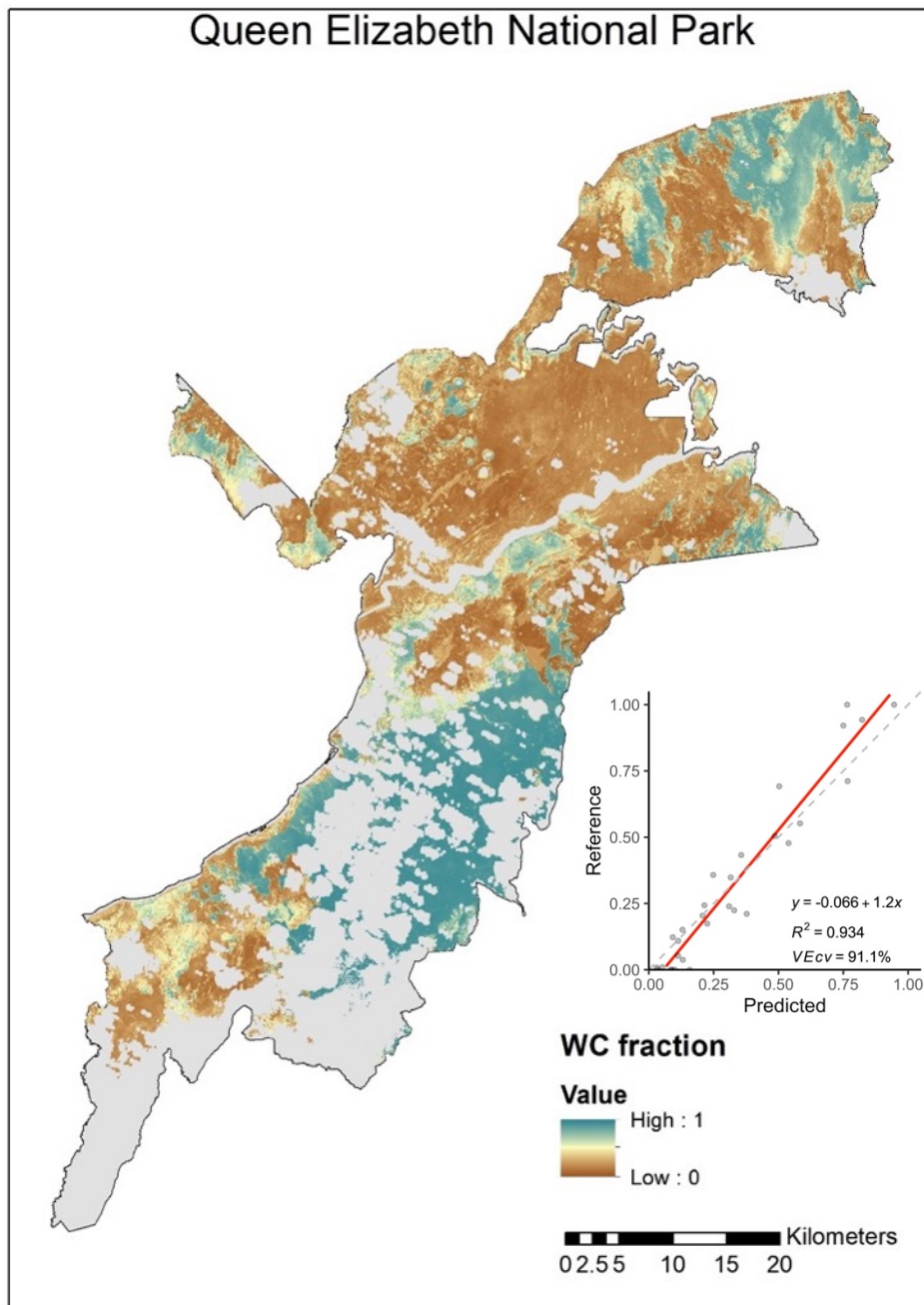


Figure A.10. The best map produced for Ruaha National Park. The map was produced using a random forest model trained using training data from all the PAs and variables from the dry and wet season composite (RF-ALL-DW; “Best Overall Model” column in Table 4 in Chapter 1). The map had a VECv of 76.3%. The TIFF file of this map can be found in the online dataset (see Nagelkirk and Dahlin, 2019). Gray areas are clouds. The scatterplot shows the regression line (solid red line) and equation relating the training data (“Reference”) to the mapped values (“Predicted”), along with R^2 and VECv metrics. The dashed line represents the 1:1 line.

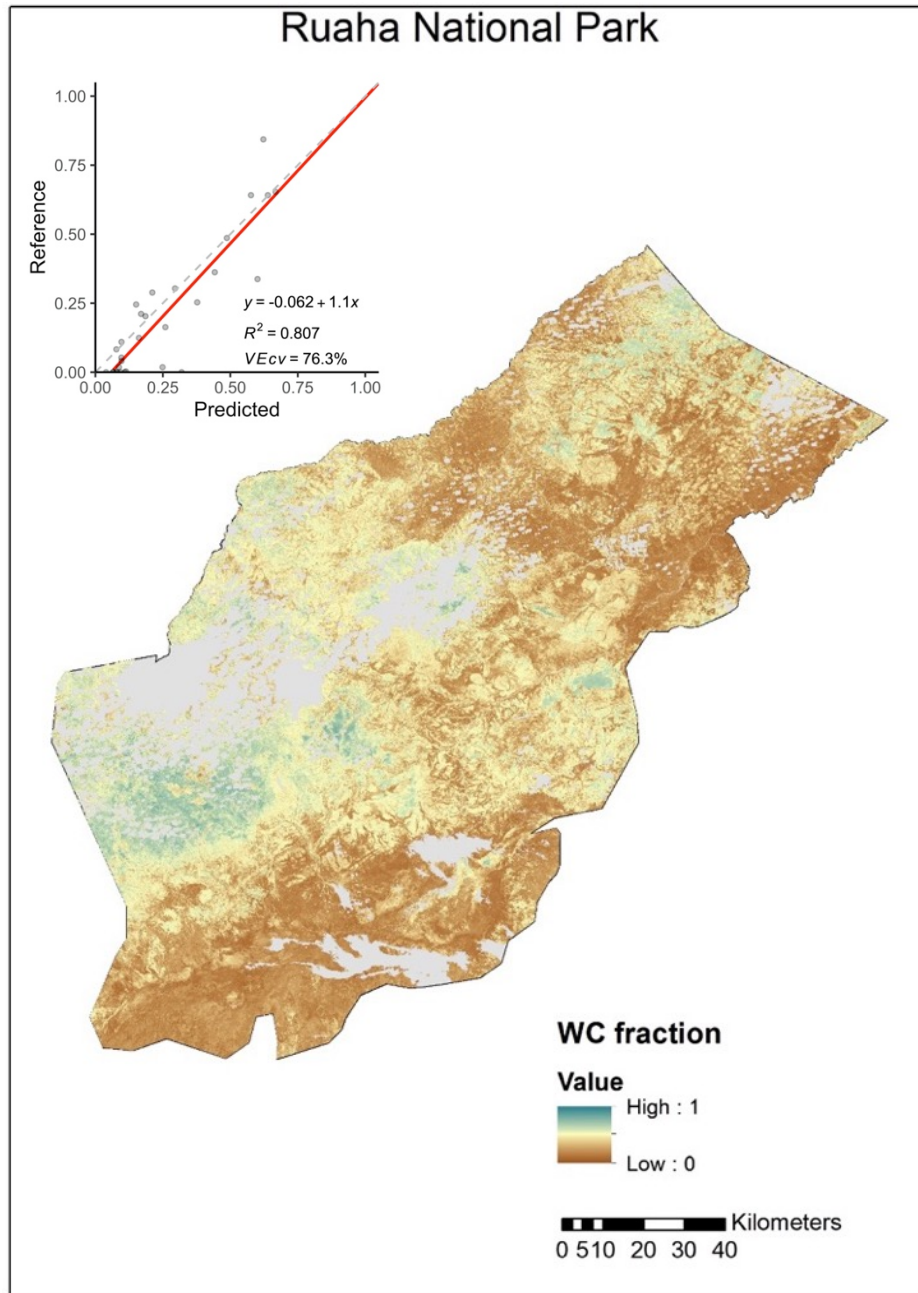


Figure A.11. The best map produced for Selous Game Reserve. The map was produced using a random forest model trained using training data from all the PAs and variables from the dry, transition and wet season composite (RF-ALL-DTW; “Best Overall Model” column in Table 4 in Chapter 1). The map had a VEcv of 86.2%. The TIFF file of this map can be found in the online dataset (see Nagelkirk and Dahlin, 2019). Gray areas are clouds. The scatterplot shows the regression line (solid red line) and equation relating the training data (“Reference”) to the mapped values (“Predicted”), along with R^2 and VEcv metrics. The dashed line represents the 1:1 line.

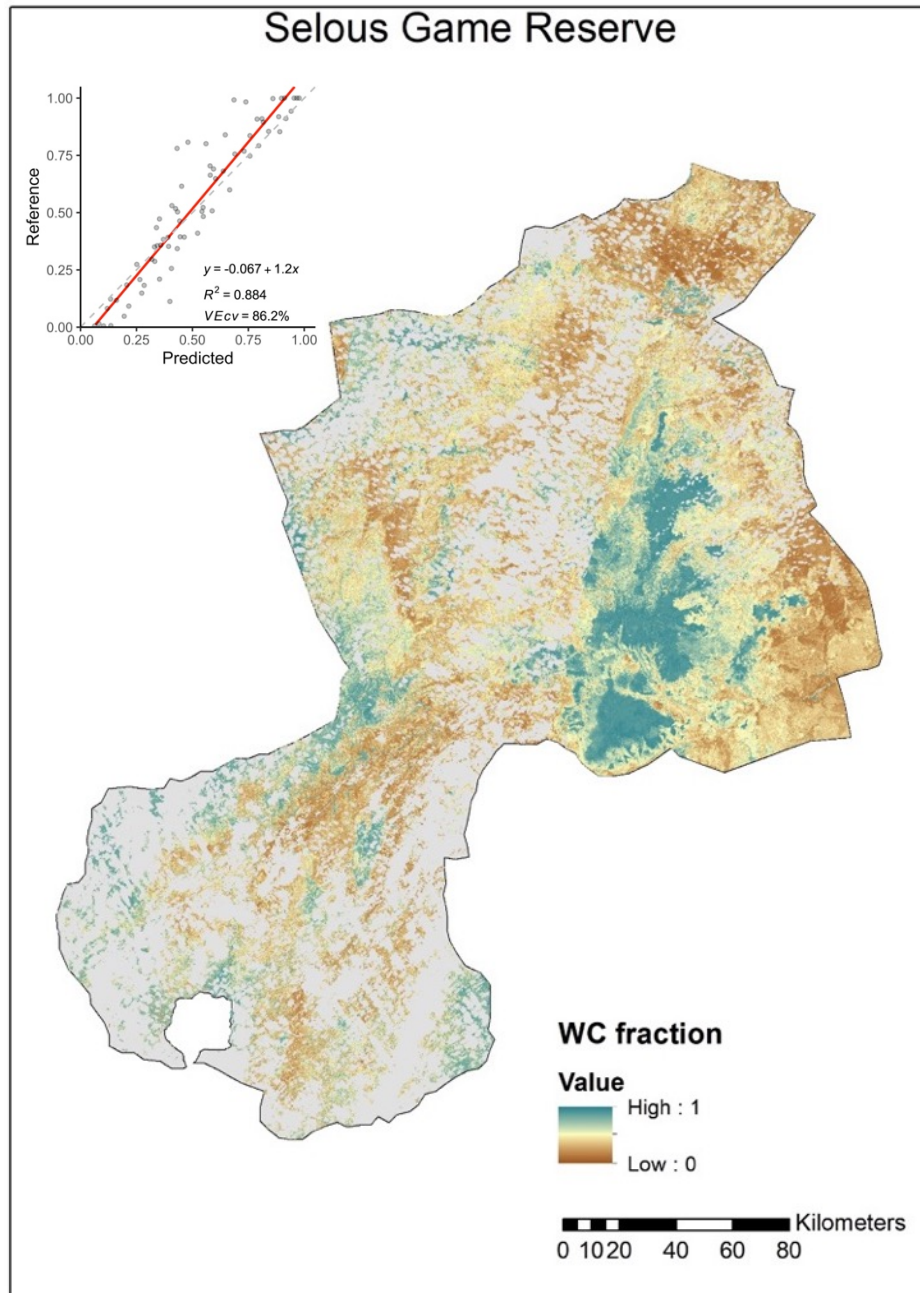


Figure A.12. The best map produced for the combined Serengeti National Park and Maasai Mara National Reserve. The map was produced using a Multiple Endmember Spectral Mixture Analysis model that unmixed only woody cover and grass and was trained using the dry season image and the EMC endmember selection method (MESMA EMC TG - Dry; “Best Overall Model” column in Table 4 in Chapter 1). The map had a VECv of 38.9%. The TIFF file of this map can be found in the online dataset (see Nagelkirk and Dahlin, 2019). Gray areas are clouds. The scatterplot shows the regression line (solid red line) and equation relating the training data (“Reference”) to the mapped values (“Predicted”), along with R^2 and VECv metrics. The dashed line represents the 1:1 line.

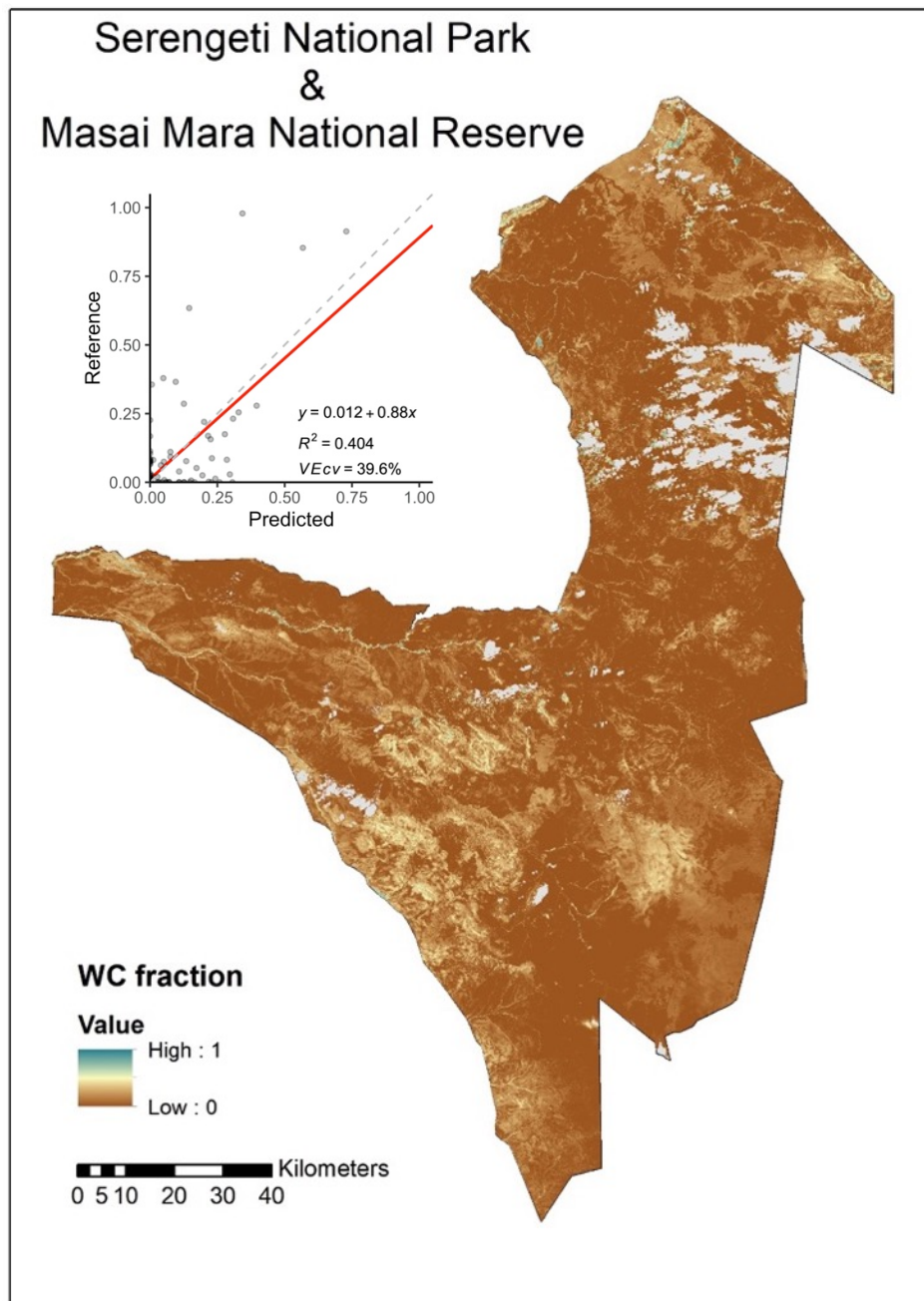
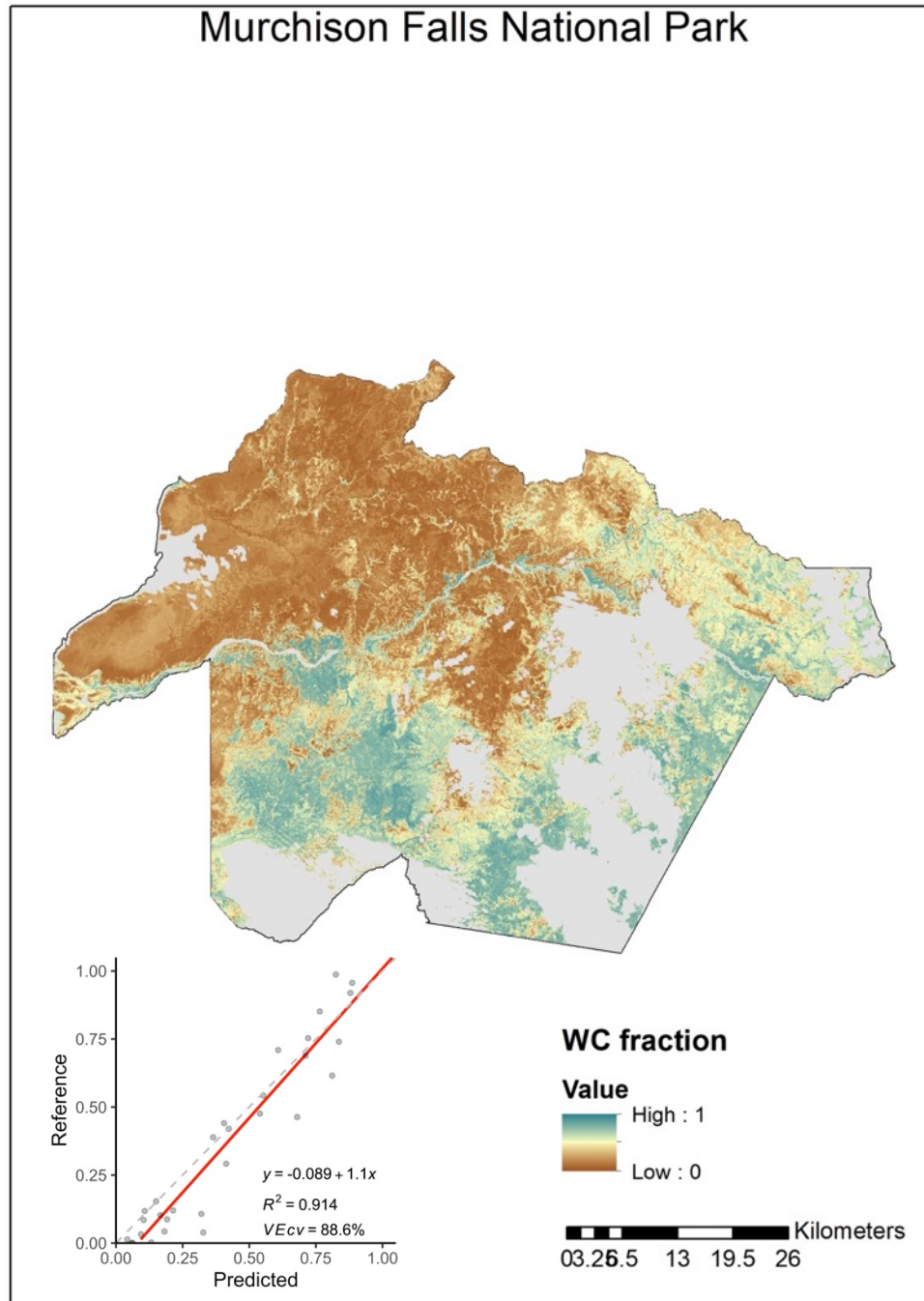


Figure A.13: The best map produced for Murchison Falls National Park. The map was produced using a random forest model trained using training data from all the PAs and variables from the dry and wet season composite (RF-ALL-DW; “Best Overall Model” column in Table 4 in Chapter 1). The map had a VEcv of 88.6%. The TIFF file of this map can be found in the online dataset (see Nagelkirk and Dahlin, 2019). Gray areas are clouds. The scatterplot shows the regression line (solid red line) and equation relating the training data (“Reference”) to the mapped values (“Predicted”), along with R^2 and VEcv metrics. The dashed line represents the 1:1 line.



APPENDIX B

Chapter 2 Supplementary Material

Table B.1. Ranked accuracies of models listed by the individual and composite images used to produce woody cover maps. In the parentheses, “ALL” denotes that training data from all the PAs was used in developing the model – otherwise only training data from the PA itself was used. While all possible combinations of images and training data sources yield 14 combinations, we list only the first seven here because subsequent listings most likely use a repeated image or image composite, meaning they would not be capable of filling any cloud gaps (this was the case for the two models using SMR training data).

PA	1st	2nd	3rd	4th	5th	6th	7th
CHO	DW (ALL)	TD (ALL)	D (ALL)	DTW (ALL)	W (ALL)	T (ALL)	WT (ALL)
KRU	DW (ALL)	WT (ALL)	D (ALL)	TD (ALL)	T (ALL)	DTW (ALL)	W (ALL)
LIM	D (ALL)	DW (ALL)	DTW (ALL)	TD (ALL)	W (ALL)	WT (ALL)	T (ALL)
MPA	TD (ALL)	WT (ALL)	D (ALL)	DW (ALL)	DTW (ALL)	T (ALL)	W (ALL)
MUR	DW (ALL)	D (ALL)	DTW (ALL)	W (ALL)	WT (ALL)	TD (ALL)	T (ALL)
NLU	DTW (ALL)	TD (ALL)	T (ALL)	D (ALL)	WT (ALL)	DW (ALL)	W (ALL)
QUE	DW (ALL)	W (ALL)	D (ALL)	DTW (ALL)	TD (ALL)	WT (ALL)	T (ALL)
RUA	DW (ALL)	DTW (ALL)	W (ALL)	WT (ALL)	T (ALL)	TD (ALL)	D (ALL)
SEL	DTW (ALL)	TD (ALL)	T (ALL)	WT (ALL)	DW (ALL)	W (ALL)	D (ALL)
SMR	WT (ALL)	W (ALL)	DTW (ALL)	DW (ALL)	DTW (SMR)	T (ALL)	T (SMR)
SLU	DTW (ALL)	T (ALL)	TD (ALL)	WT (ALL)	DW (ALL)	D (ALL)	W (ALL)

Table B.2. Seasonal imagery available for woody cover mapping. Image names are listed in the order they were used in mapping (i.e. the first image and corresponding model were used to generate the base woody cover map whose clouded areas were filled using the subsequent images). *D* = dry season; *W* = wet season and *T* = transition season. Seasonal combinations (e.g. DTW) represent composite images.

Year	CHO	KRU	LIM	MPA	MUR	NLU	QUE	RUA	SEL	SER	SLU	MAR
1984	D		D	TD, D, T	T	TD, T, D	D, TD, T			D	T, TD, D	D
1985				D								
1986	T		T	WT, T, W	T	TD, T, D	D, TD, T			T, TD, D	DW, D, W	T, TD, D
1987	T		T	TD, WT, D, DW, T, W			D, TD, T					
1988				T								
1989	TD, D, T	D	D			D	T				T, TD, D	
1990	DW, D, W	D, TD, T	D, TD, T		T	D, DW, W	T	T, TD, D	T	T	DTW, T, TD, WT, DW, D, W	T
1991	DW, D, W	DW, WT, D, TD, T, W	D, TD, T			TD, T, D		T			T, TD, D	
1992	D	D	D	D							D	
1993	TD, D, T	D	D, TD, T			T					T	
1994	TD, D, T	D, TD, T	D, TD, T	TD, D, T	T	T		T		D	T, TD, D	D
1995	TD, D, T	DW, WT, D, TD, T, W	D, DW, TD, W, WT, T	TD, D, T	T	TD, T, D	D	W	T		D	
1996	DW, D, W	T	T			T		T		D	T, TD, D	
1997	DW, TD, D, W, T	D				D		T			T	
1998	TD, D, T	D	D			TD, T, D		T, TD, D	D		DTW, T, TD, WT, DW, D, W	
1999	D	DW, D, W	W	D	T	D	W, WT, T	T	T	D	D	D
2000	TD, D, T	DW, D, W	D, DW, W		T	TD, T, D	D, TD, T	W	T, WT, W	T, TD, D	DTW, T, TD, WT, DW, D, W	T, TD, D
2001	DW, TD, D, W, T	D	D	T	W	D, DW, W	DW, W, D, TD, WT, T		D	W	DW, D, W	W
2002	TD, D, T	D, TD, T	D, TD, T	D, DW, W	T	TD, T, D	DW, W, D, TD, WT, T		W	W, DW, D	DTW, T, TD, WT, DW, D, W	W, DW, D
2003	TD, D, T		T	D, DW, W	D		D, TD, T			WT, W, T	W	WT, W, T
2004	T		W			TD, T, D		D			DTW, T, TD, WT, DW, D, W	
2005	D	D	D			D		W, WT, T			T, TD, D	
2006	DW, TD, D, W, T	D	D, DW, TD, W, WT, T			TD, T, D			W		DTW, T, TD, WT, DW, D, W	
2007	TD, D, T	WT, T, W	D			T		DW, W, D			T	
2008	DW, TD, D, W, T	D	D, TD, T	D	T	TD, T, D	D, TD, T		W		T, TD, D	
2009	T					T	T	T		T	T	
2010	W, T	W	W	TD, D, T	D		T			T		T
2011	W			TD, D, T	D, TD, T		D			T, TD, D	W	T, TD, D
2012												
2013	TD, D, T	DW, WT, D, TD, T, W	D, DW, TD, W, WT, T	WT, T, W	W, WT, T	TD, T, D	DW, W, D, TD, WT, T	DW, W, WT, T, TD, D		WT, W, DW, T, TD, D	DTW, T, TD, WT, DW, D, W	WT, W, DW, T, TD, D
2014	DW, D, W	D	D, DW, TD, W, WT, T	TD, WT, D, DW, T, W	DW, D, W, WT, TD, T	DTW, TD, T, D, WT, DW, W	DW, W, D	DW, W, WT, T, TD, D	DW, W, D	WT, W, T	DTW, T, TD, WT, DW, D, W	WT, W, T
2015	DW, TD, D, W, T	DW, D, W	T	WT, T, W	DW, D, W, WT, TD, T	DTW, TD, T, D, WT, DW, W	DW, W, D, TD, WT, T	DW, W, WT, T, TD, D	TD, T, D	WT, W, DW, T, TD, D	DTW, T, TD, WT, DW, D, W	WT, W, DW, T, TD, D
2016	TD, D, T	DW, WT, D, TD, T, W	D, DW, TD, W, WT, T	TD, WT, D, DW, T, W	DW, D, W, WT, TD, T	DTW, TD, T, D, WT, DW, W	DW, W, D, TD, WT, T	DW, W, WT, T, TD, D	DTW, TD, T, WT, DW, W, D	WT, W, DW, T, TD, D	DTW, T, TD, WT, DW, D, W	WT, W, DW, T, TD, D

Figure B.1. Woody cover fraction (black) and elephant density (red) data from the 12 PAs used in this study. Dashed lines denote a significant trend ($p \leq 0.05$) in the data. Error bars on the woody cover data denote the spatially weighted RMSE of the models used to generate the maps (multiple models were used to generate each year's map). Error bars for the elephant data denote 95% confidence intervals. In cases where confidence intervals are absent, it is either due to the values not being reported, or the values were reported as zero – a common practice when an aerial census covered the entire PA. PA acronyms correspond to Table 5 in the main text. Consistent declines in elephant densities across census years, such as in SEL and RUA, are often the result of poaching, whereas marked declines such as that in CHO are more often the result of methodological changes between census years.

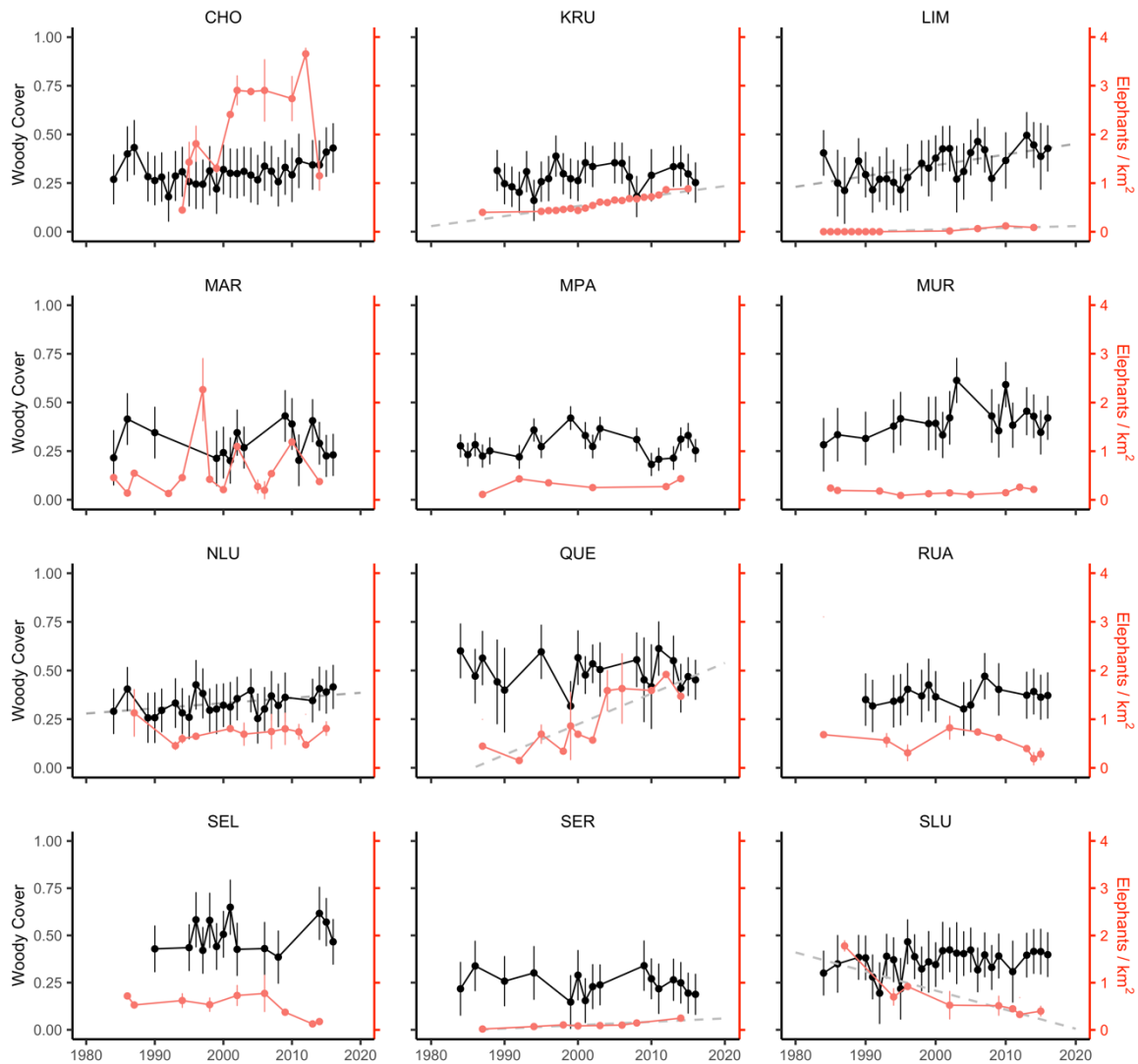


Figure B.2. Temporal extents of the different data used in this study. These seven variables were used to generate the full set of 17 variables. Gray points represent values interpolated using the original data (black). Some of the original data was from before 1984 and is not shown (e.g. the earliest human population density data is from 1975).

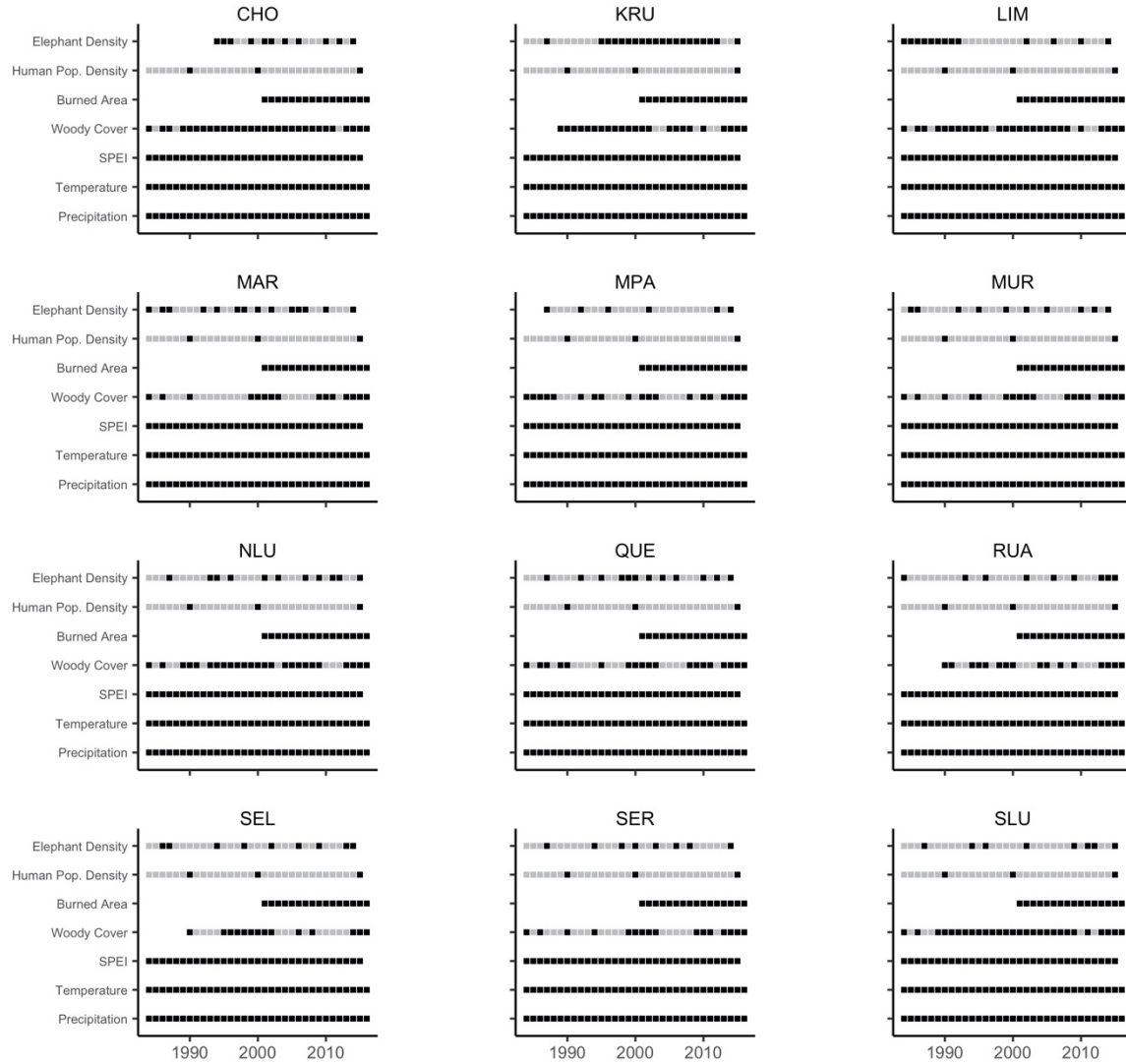


Figure B.3. Relationship between PA mean woody cover and the 17 variables used in this study. Significant relationships have dashed regression lines with shading indicating the 95% CI. Outliers are labeled with asterisks (see Methods for outlier analysis). Symbols correspond to those used in the main text (Figures 9-11).

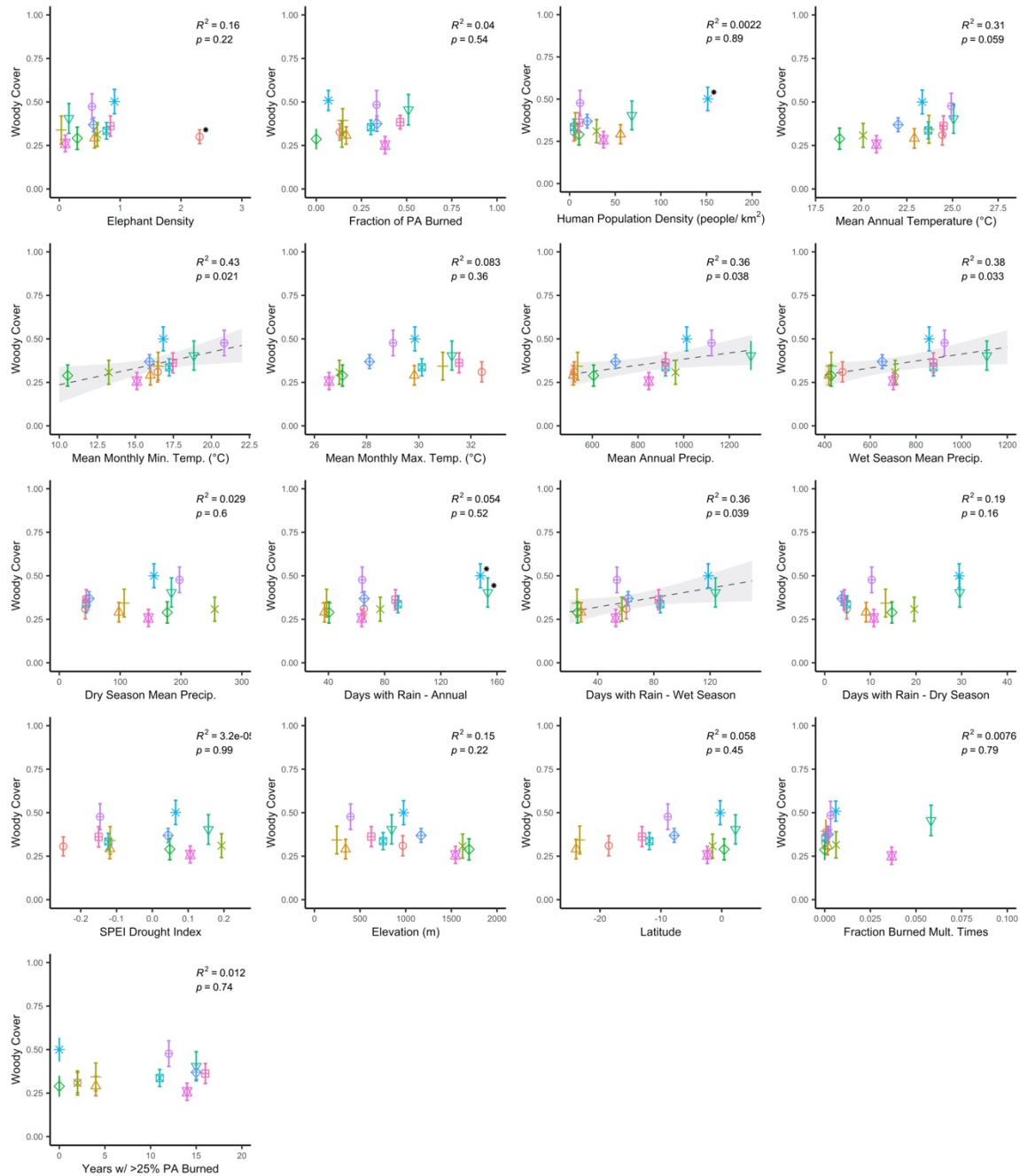


Table B.3. Pearson's correlation matrix of independent variables used in study. Correlation coefficients are reported with p-values in parentheses.

	Elephant density	Burned fraction	Human density	MAT	T _{min}	T _{max}	MAP	P _{wet}	P _{dry}	DWP _{ann}	DWP _{wet}	DWP _{dry}	SPEI	Elevation	Latitude	Burn _{mult}	Burn ₂₅
Elephant density	1 (0)	-0.08 (0.82)	-0.28 (0.42)	0.18 (0.59)	0.16 (0.64)	0.18 (0.59)	0.19 (0.57)	0.29 (0.39)	-0.26 (0.44)	0.3 (0.37)	0.38 (0.24)	-0.14 (0.68)	-0.25 (0.46)	-0.1 (0.77)	-0.02 (0.96)	-0.52 (0.1)	-0.07 (0.83)
Burned fraction	-0.08 (0.82)	1 (0)	0.35 (0.29)	0.49 (0.11)	0.61 (0.04)	0.2 (0.54)	0.58 (0.05)	0.7 (0.01)	-0.18 (0.58)	0.35 (0.26)	0.45 (0.15)	-0.13 (0.68)	0.06 (0.86)	-0.22 (0.48)	0.12 (0.7)	0.58 (0.05)	0.96 (0)
Human density	-0.28 (0.42)	0.35 (0.29)	1 (0)	0.01 (0.97)	0.09 (0.79)	-0.09 (0.79)	0.35 (0.3)	0.26 (0.43)	0.37 (0.26)	0.44 (0.18)	0.32 (0.34)	0.63 (0.04)	0.55 (0.08)	0.03 (0.93)	0.28 (0.41)	0.74 (0.01)	0.21 (0.55)
MAT	0.18 (0.59)	0.49 (0.11)	0.01 (0.97)	1 (0)	0.92 (0)	0.85 (0)	0.32 (0.32)	0.47 (0.12)	-0.38 (0.23)	0.39 (0.21)	0.47 (0.12)	-0.04 (0.9)	-0.56 (0.06)	-0.81 (0)	-0.4 (0.2)	0.12 (0.72)	0.37 (0.23)
T _{min}	0.16 (0.64)	0.61 (0.04)	0.09 (0.79)	0.92 (0)	1 (0)	0.58 (0.05)	0.51 (0.09)	0.62 (0.03)	-0.19 (0.56)	0.4 (0.2)	0.46 (0.13)	-0.01 (0.97)	-0.39 (0.21)	-0.73 (0.01)	-0.22 (0.49)	0.21 (0.51)	0.53 (0.07)
T _{max}	0.18 (0.59)	0.2 (0.54)	-0.09 (0.79)	0.85 (0)	0.58 (0.05)	1 (0)	-0.02 (0.95)	0.15 (0.65)	-0.54 (0.07)	0.29 (0.36)	0.36 (0.25)	-0.07 (0.84)	-0.63 (0.03)	-0.7 (0.01)	-0.52 (0.08)	-0.04 (0.91)	0.06 (0.85)
MAP	0.19 (0.57)	0.58 (0.05)	0.35 (0.3)	0.32 (0.32)	0.51 (0.09)	-0.02 (0.95)	1 (0)	0.96 (0)	0.47 (0.12)	0.78 (0)	0.76 (0)	0.54 (0.07)	0.44 (0.15)	0.03 (0.92)	0.65 (0.02)	0.58 (0.05)	0.47 (0.13)
P _{wet}	0.29 (0.39)	0.7 (0.01)	0.26 (0.43)	0.47 (0.12)	0.62 (0.03)	0.15 (0.65)	0.96 (0)	1 (0)	0.21 (0.51)	0.81 (0)	0.84 (0)	0.38 (0.23)	0.28 (0.37)	-0.08 (0.8)	0.53 (0.07)	0.53 (0.08)	0.6 (0.04)
P _{dry}	-0.26 (0.44)	-0.18 (0.58)	0.37 (0.26)	-0.38 (0.23)	-0.19 (0.56)	-0.54 (0.07)	0.47 (0.12)	0.21 (0.51)	1 (0)	0.18 (0.59)	0.01 (0.97)	0.7 (0.01)	0.67 (0.02)	0.38 (0.23)	0.62 (0.03)	0.35 (0.26)	-0.27 (0.4)
DWP _{ann}	0.3 (0.37)	0.35 (0.26)	0.44 (0.18)	0.39 (0.21)	0.4 (0.2)	0.29 (0.36)	0.78 (0)	0.81 (0)	0.18 (0.59)	1 (0)	0.98 (0)	0.67 (0.02)	0.4 (0.2)	0.04 (0.9)	0.57 (0.05)	0.54 (0.07)	0.2 (0.52)
DWP _{wet}	0.38 (0.24)	0.45 (0.15)	0.32 (0.34)	0.47 (0.12)	0.46 (0.13)	0.36 (0.25)	0.76 (0)	0.84 (0)	0.01 (0.97)	0.98 (0)	1 (0)	0.51 (0.09)	0.28 (0.38)	0 (0.99)	0.49 (0.1)	0.48 (0.11)	0.32 (0.31)
DWP _{dry}	-0.14 (0.68)	-0.13 (0.68)	0.63 (0.04)	-0.04 (0.9)	-0.01 (0.97)	-0.07 (0.84)	0.54 (0.07)	0.38 (0.23)	0.7 (0.01)	0.67 (0.02)	0.51 (0.09)	1 (0)	0.67 (0.02)	0.18 (0.57)	0.61 (0.04)	0.54 (0.07)	-0.29 (0.36)
SPEI	-0.25 (0.46)	0.06 (0.86)	0.55 (0.08)	-0.56 (0.06)	-0.39 (0.21)	-0.63 (0.03)	0.44 (0.15)	0.28 (0.37)	0.67 (0.02)	0.4 (0.2)	0.28 (0.38)	0.67 (0.02)	1 (0)	0.67 (0.02)	0.81 (0)	0.57 (0.05)	0.02 (0.94)
Elevation	-0.1 (0.77)	-0.22 (0.48)	0.03 (0.93)	-0.81 (0)	-0.73 (0.01)	-0.7 (0.01)	0.03 (0.92)	-0.08 (0.8)	0.38 (0.23)	0.04 (0.9)	0 (0.99)	0.18 (0.57)	0.67 (0.02)	1 (0)	0.71 (0.01)	0.19 (0.56)	-0.17 (0.59)
Latitude	-0.02 (0.96)	0.12 (0.7)	0.28 (0.41)	-0.4 (0.2)	-0.22 (0.49)	-0.52 (0.08)	0.65 (0.02)	0.53 (0.07)	0.62 (0.03)	0.57 (0.05)	0.49 (0.1)	0.61 (0.04)	0.81 (0)	0.71 (0.01)	1 (0)	0.52 (0.08)	0.11 (0.74)
Burn _{mult}	-0.52 (0.1)	0.58 (0.05)	0.74 (0.01)	0.12 (0.72)	0.21 (0.51)	-0.04 (0.91)	0.58 (0.05)	0.53 (0.08)	0.35 (0.26)	0.54 (0.07)	0.48 (0.11)	0.54 (0.07)	0.57 (0.05)	0.19 (0.56)	0.52 (0.08)	1 (0)	0.45 (0.15)
Burn ₂₅	-0.07 (0.83)	0.96 (0)	0.21 (0.55)	0.37 (0.23)	0.53 (0.07)	0.06 (0.85)	0.47 (0.13)	0.6 (0.04)	-0.27 (0.4)	0.2 (0.52)	0.32 (0.31)	-0.29 (0.36)	0.02 (0.94)	-0.17 (0.59)	0.11 (0.74)	0.45 (0.15)	1 (0)

Figure B.4. Relationship between mean woody cover in flat areas near water and the 17 variables used in this study. None of these relationships were significant. Outliers are labeled with asterisks (see Methods for outlier analysis). Symbols correspond to those used in the main text (Figures 9-11).

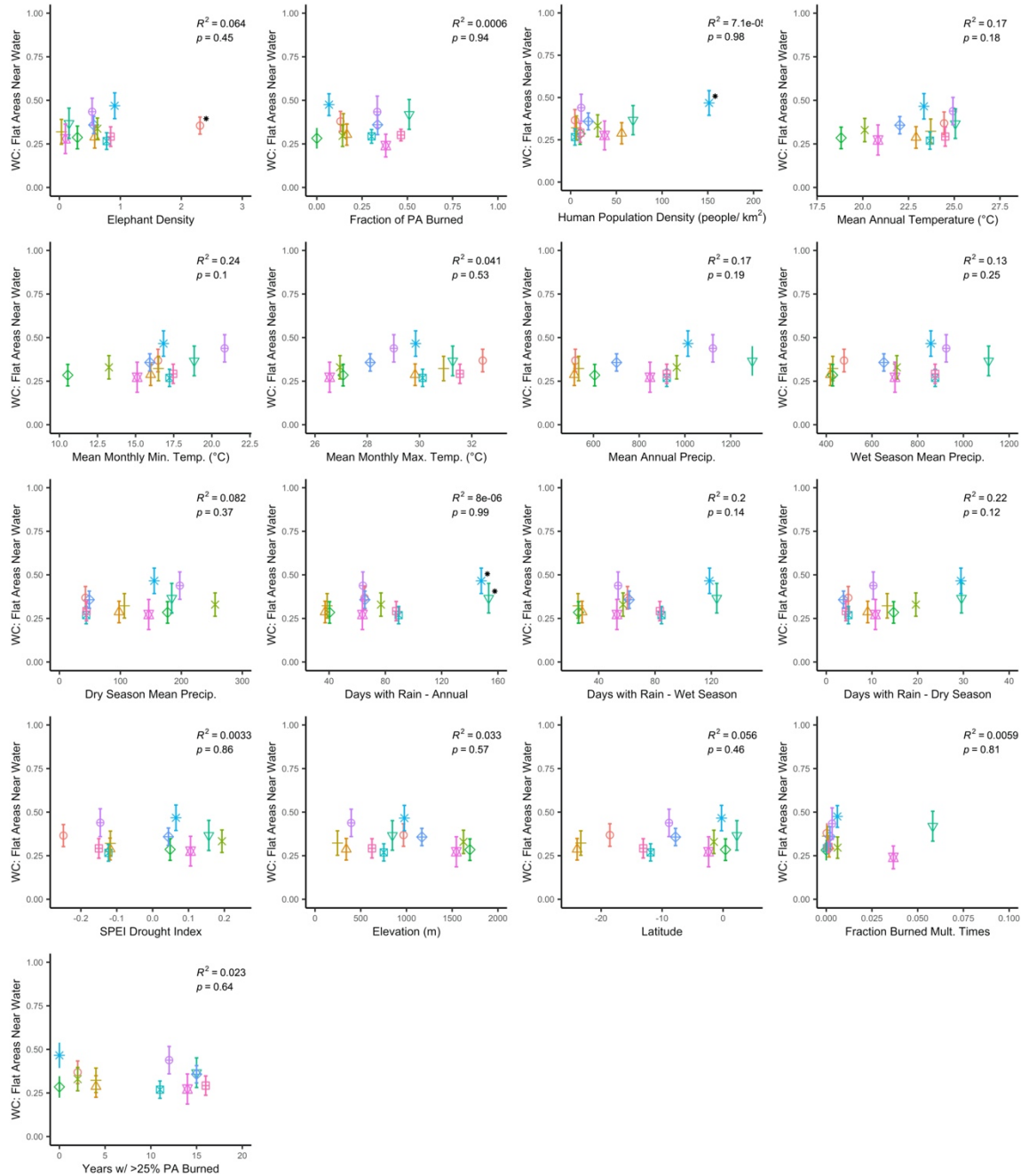


Figure B.5. Relationship between mean woody cover in hilly areas near water and the 17 variables used in this study. None of these relationships were significant. Outliers are labeled with asterisks (see Methods for outlier analysis). Symbols correspond to those used in the main text (Figures 9-11).

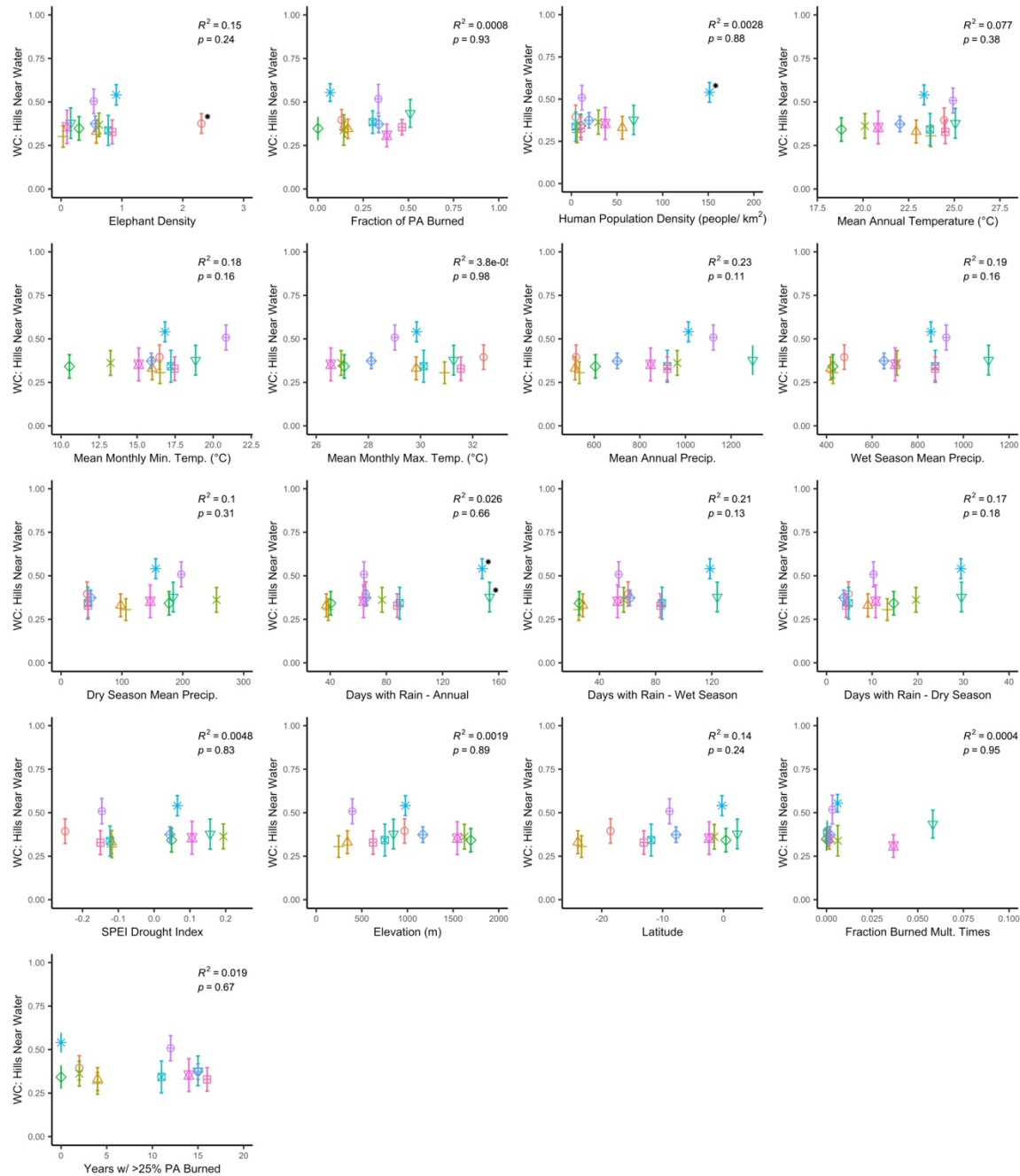


Figure B.6. Relationship between mean woody cover in flat areas far from water and the 17 variables used in this study. Significant relationships have dashed regression lines with shading indicating the 95% CI. Outliers are labeled with asterisks (see Methods for outlier analysis). Symbols correspond to those used in the main text (Figures 9-11).

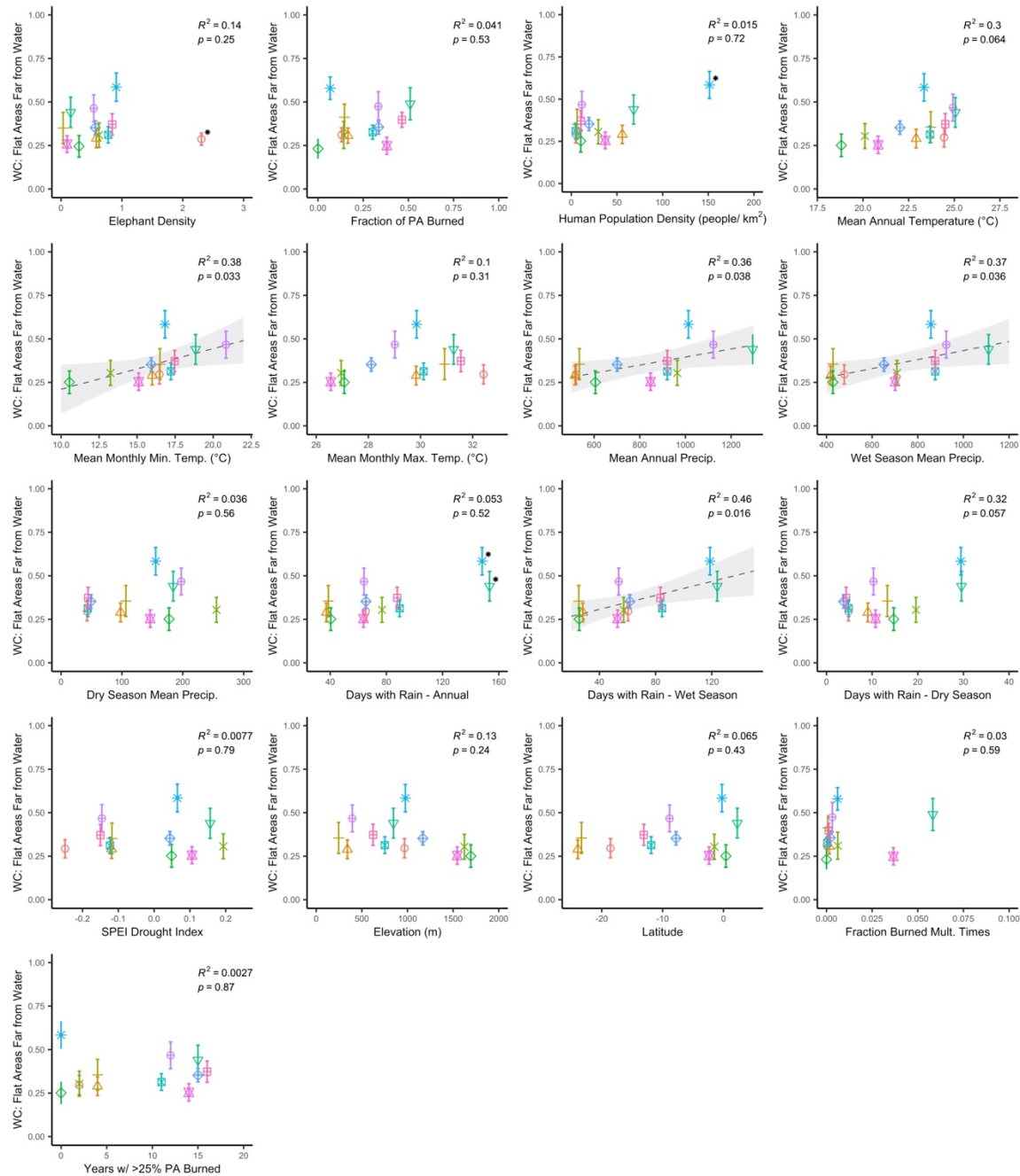


Figure B.7. Relationship between mean woody cover in hilly areas far from water and the 17 variables used in this study. Significant relationships have dashed regression lines with shading indicating the 95% CI. Outliers are labeled with asterisks (see Methods for outlier analysis). Symbols correspond to those used in the main text (Figures 9-11).

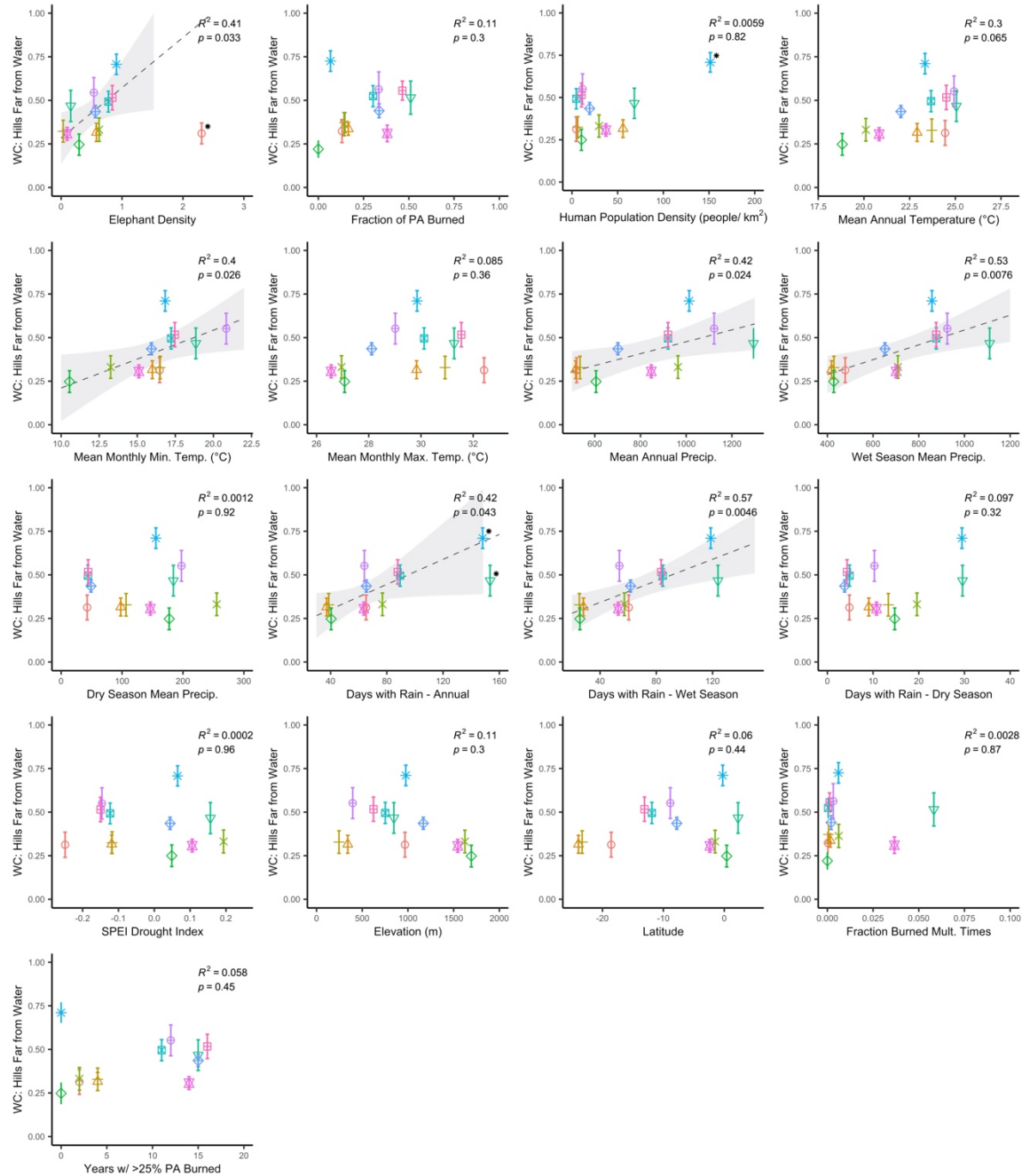


Figure B.8. Relationship between woody cover differences between hills and flat areas near water and the 17 variables used in this study. None of these relationships were significant. Outliers are labeled with asterisks (see Methods for outlier analysis). Symbols correspond to those used in the main text (Figures 9-11).

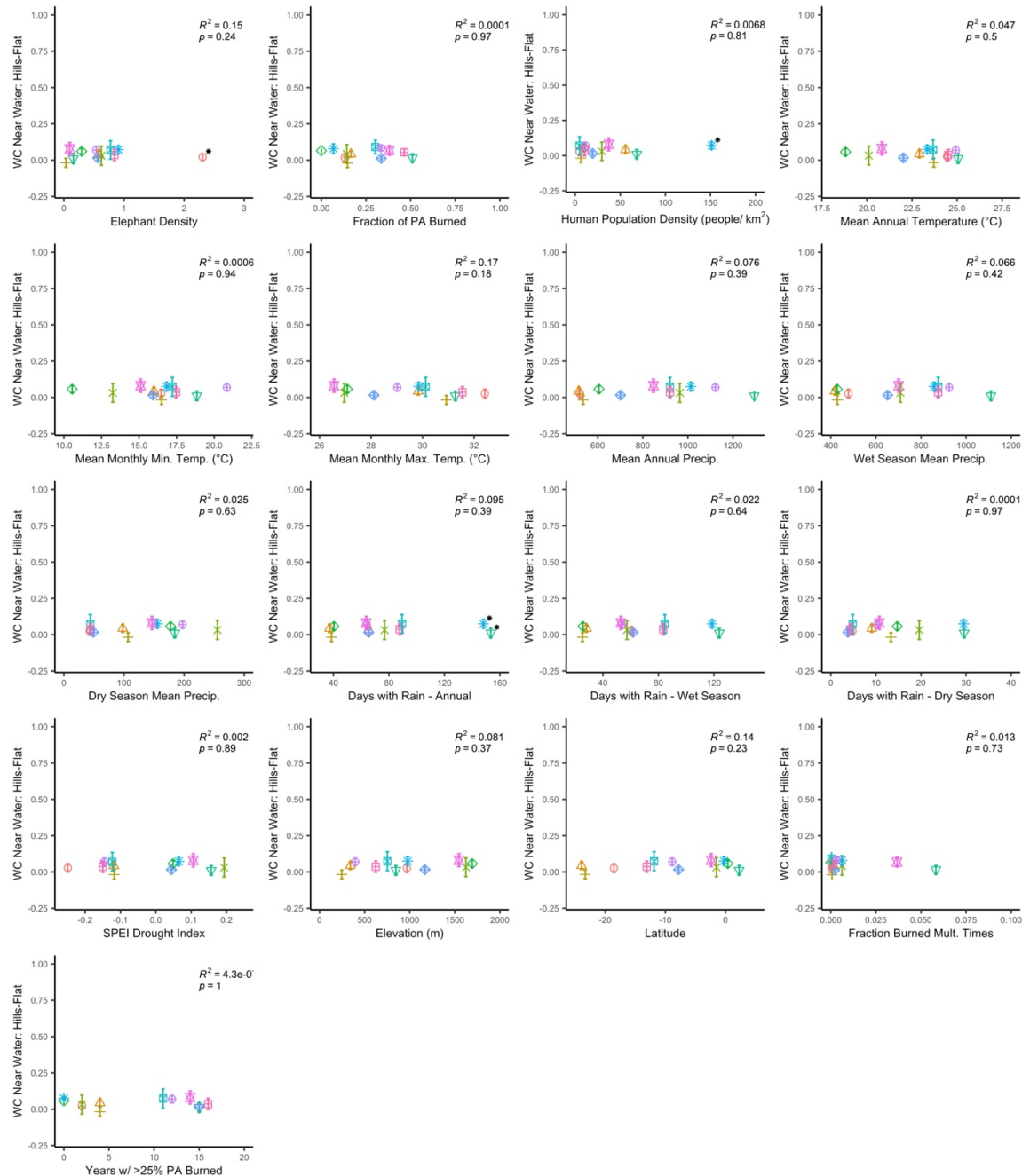


Figure B.9. Relationship between woody cover differences between hills and flat areas far from water and the 17 variables used in this study. Significant relationships have dashed regression lines with shading indicating the 95% CI. Outliers are labeled with asterisks (see Methods for outlier analysis). Symbols correspond to those used in the main text (Figures 9-11).

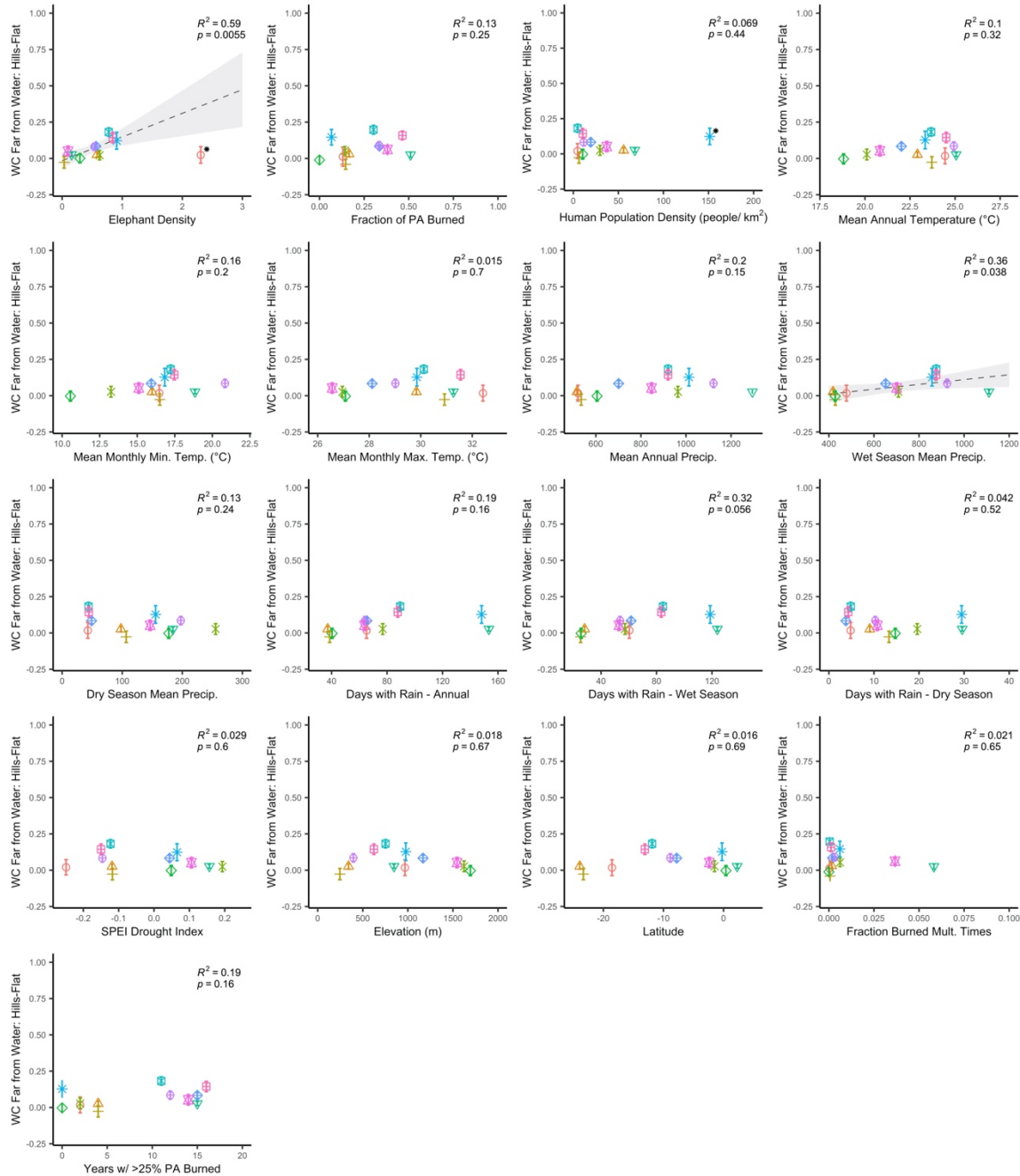


Figure B.10. Relationship between woody cover differences between flat areas far from water and those near water and the 17 variables used in this study. Significant relationships have dashed regression lines with shading indicating the 95% CI. Outliers are labeled with asterisks (see Methods for outlier analysis). Symbols correspond to those used in the main text (Figures 9-11).

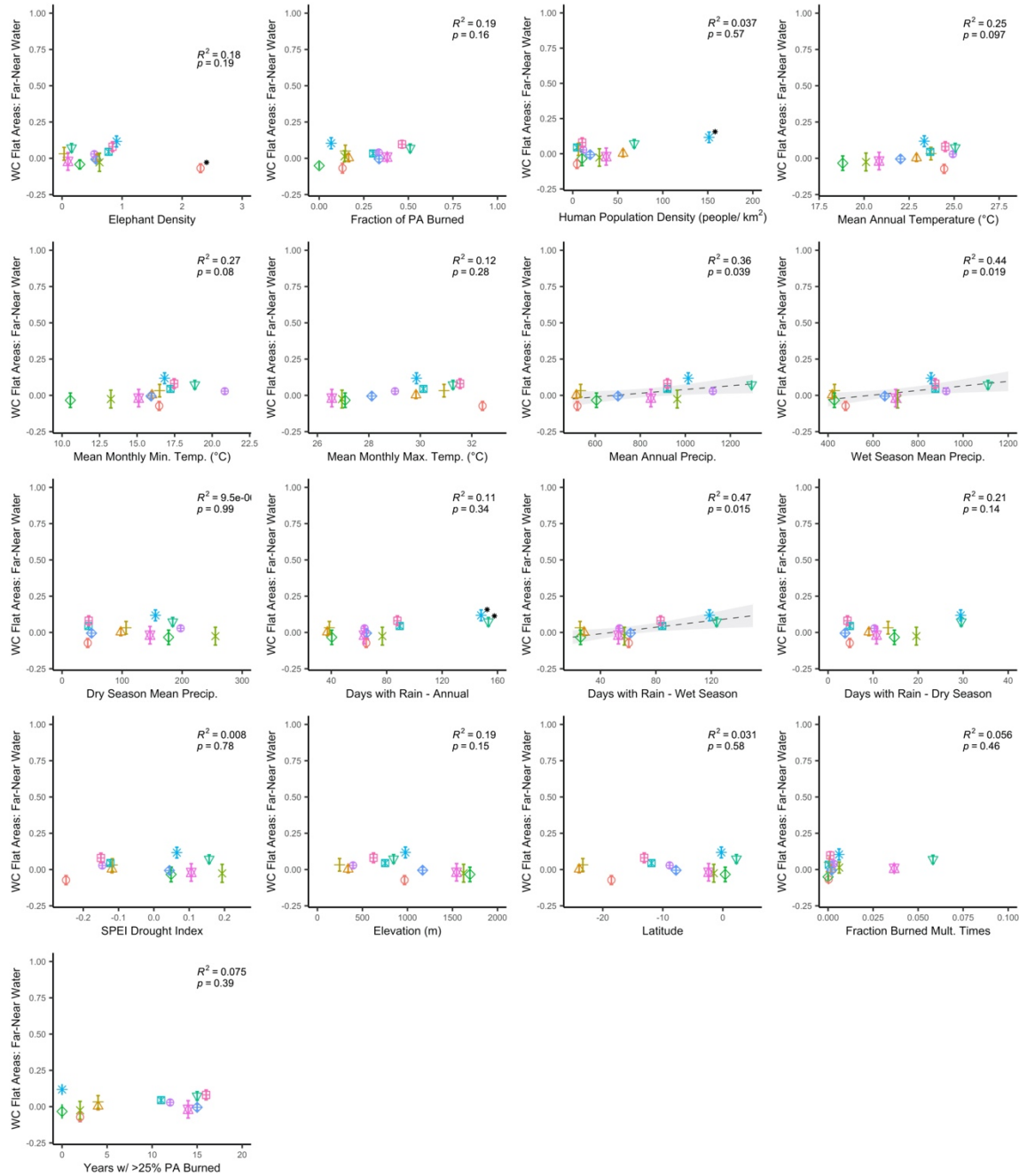


Figure B.11. Relationship between woody cover differences between hilly areas far from water and those near water and the 17 variables used in this study. Significant relationships have dashed regression lines with shading indicating the 95% CI. Outliers are labeled with asterisks (see Methods for outlier analysis). Symbols correspond to those used in the main text (Figures 9-11).

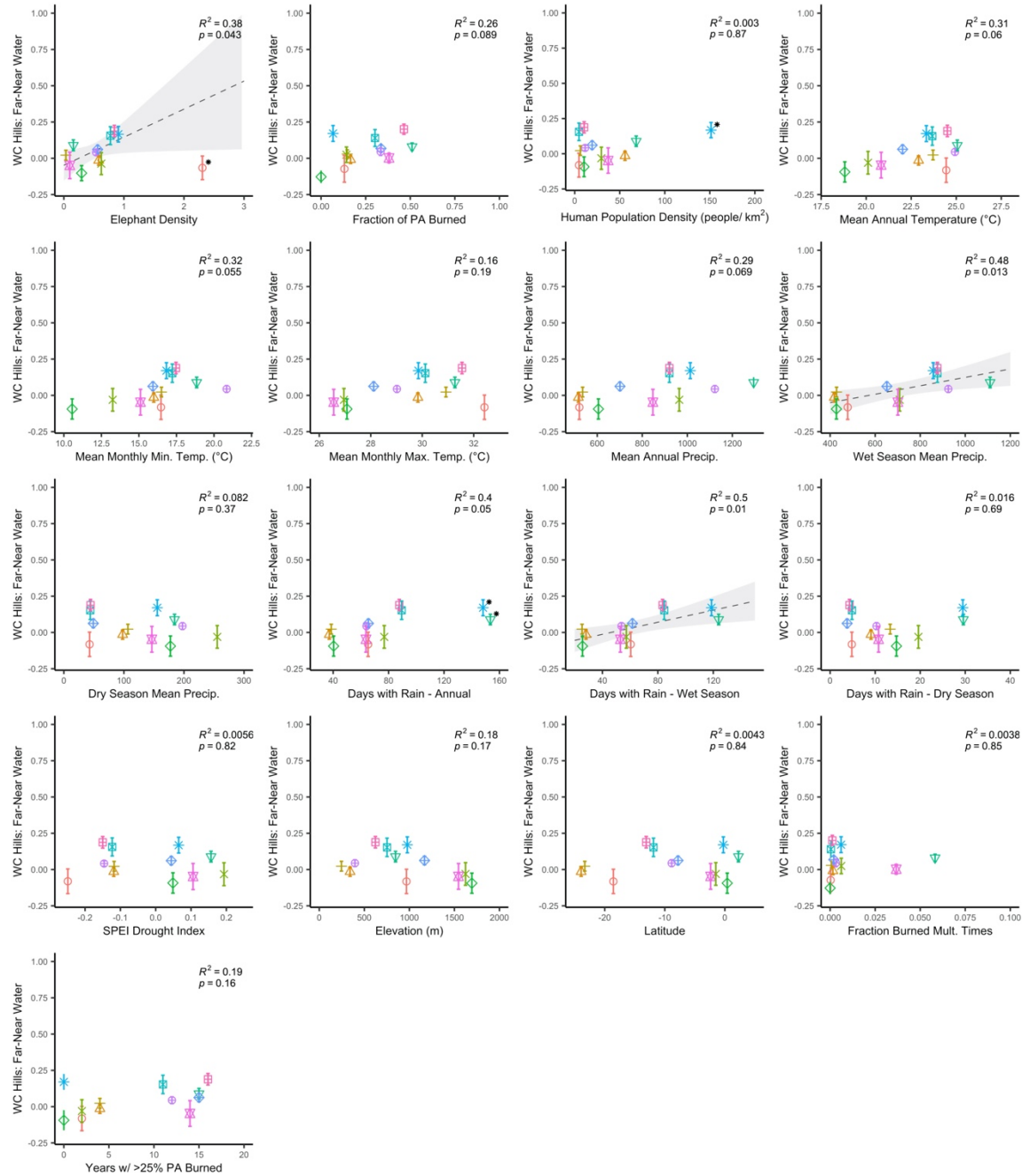


Figure B.12. Relationships between all variables used in this study and whole-PA woody cover heterogeneity. Filled circles denote significant relationships ($p \leq 0.05$) between variables and the semivariance (i.e., heterogeneity) of woody cover at each range from 50 to 950 meters.

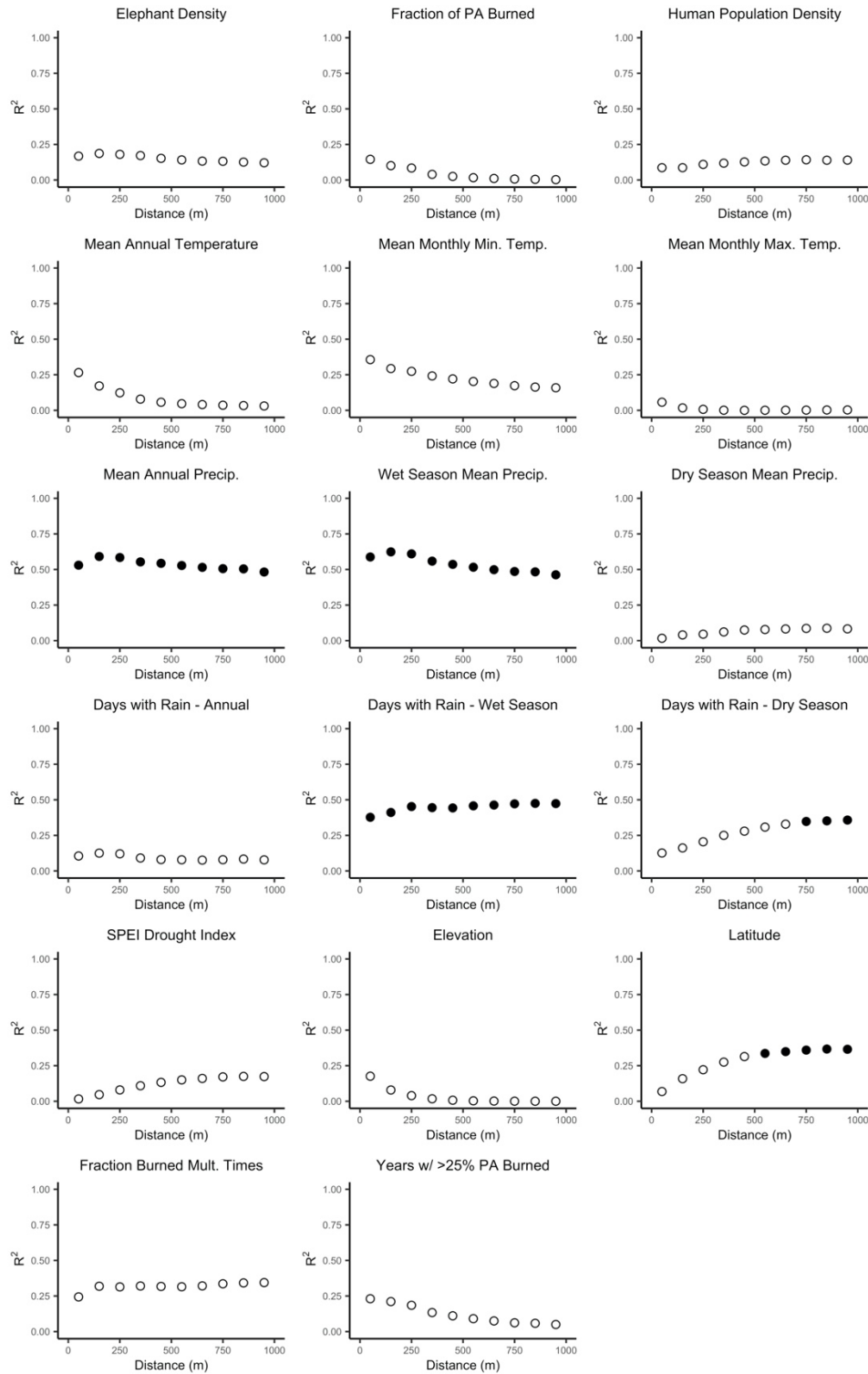


Figure B.13. Relationships between all variables used in this study and woody cover heterogeneity in flat areas near water. Filled circles denote significant relationships ($p \leq 0.05$) between variables and the semivariance (i.e., heterogeneity) of woody cover at each range from 50 to 950 meters.

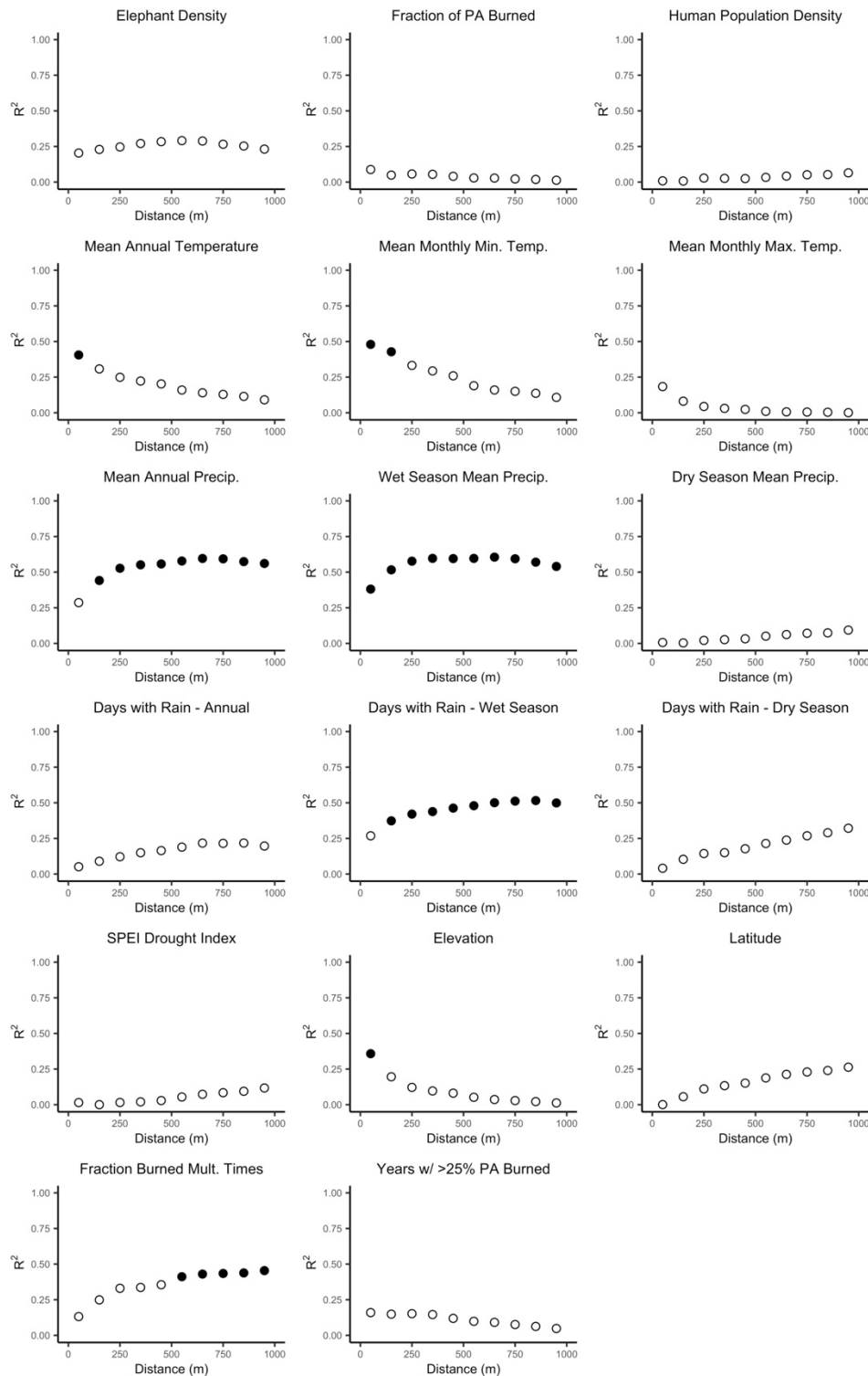
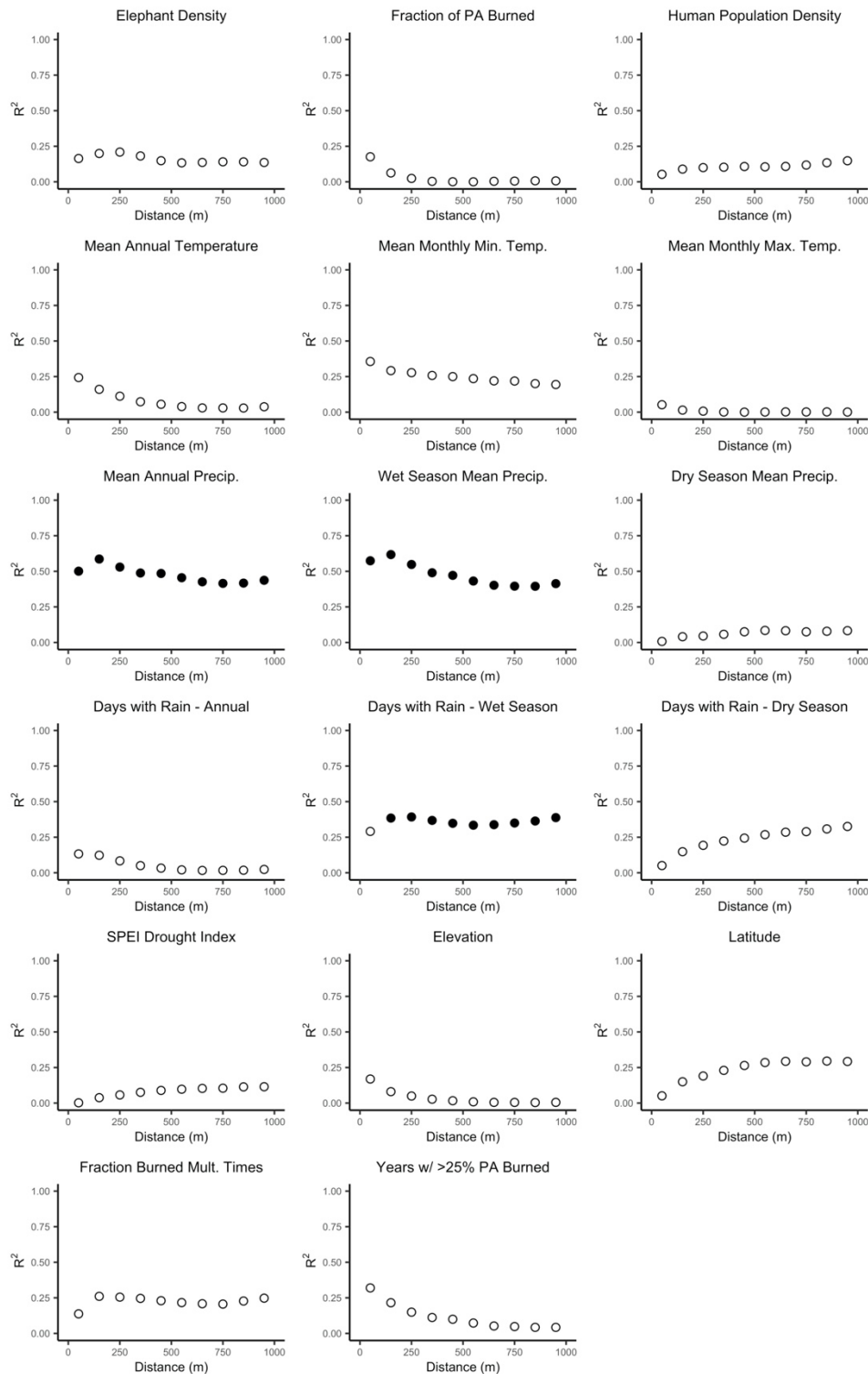


Figure B.14. Relationships between all variables used in this study and woody cover heterogeneity in flat areas far from water. Filled circles denote significant relationships ($p \leq 0.05$) between variables and the semivariance (i.e., heterogeneity) of woody cover at each range from 50 to 950 meters.



BIBLIOGRAPHY

BIBLIOGRAPHY

- Abatzoglou, J.T., Dobrowski, S.Z., Parks, S.A., Hegewisch, K.C., 2018. TerraClimate, a high-resolution global dataset of monthly climate and climatic water balance from 1958-2015. *Sci. Data* 5, 1–12. <https://doi.org/10.1038/sdata.2017.191>
- Adams, J.B., Smith, M.O., Johnson, P.E., 1986. Spectral mixture modeling: A new analysis of rock and soil types at the Viking Lander 1 Site. *J. Geophys. Res. Solid Earth* 91, 8098–8112. <https://doi.org/10.1029/JB091iB08p08098>
- Ahlström, A., Raupach, M.R., Schurgers, G., Smith, B., Arneth, A., Jung, M., Reichstein, M., Canadell, J.G., Friedlingstein, P., Jain, A.K., Kato, E., Poulter, B., Sitch, S., Stocker, B.D., Viovy, N., Wang, Y.P., Wiltshire, A., Zaehle, S., Zeng, N., 2015. The dominant role of semi-arid ecosystems in the trend and variability of the land CO₂ sink. *Science* (80-.). 348, 895–899. <https://doi.org/10.1002/2015JA021022>
- Archibald, S., Lehmann, C.E.R., Gómez-dans, J.L., Bradstock, R.A., 2013. Defining pyromes and global syndromes of fire regimes. *Proc. Natl. Acad. Sci.* 110, 6442–6447. [https://doi.org/10.1073/pnas.1211466110/-](https://doi.org/10.1073/pnas.1211466110/-/DCSupplemental)
[/DCSupplemental.www.pnas.org/cgi/doi/10.1073/pnas.1211466110](https://doi.org/10.1073/pnas.1211466110)
- Archibald, S., Nickless, A., Govender, N., Scholes, R.J., Lehsten, V., 2010. Climate and the inter-annual variability of fire in southern Africa: A meta-analysis using long-term field data and satellite-derived burnt area data. *Glob. Ecol. Biogeogr.* 19, 794–809. <https://doi.org/10.1111/j.1466-8238.2010.00568.x>
- Archibald, S., Staver, A.C., Levin, S.A., 2012. Evolution of human-driven fire regimes in Africa. *Proc. Natl. Acad. Sci.* 109, 847–852. <https://doi.org/10.1073/pnas.1118648109>
- Asner, G.P., Bustamante, M.M.C., Townsend, A.R., 2003. Scale dependence of biophysical structure in deforested areas bordering the Tapajós National Forest, Central Amazon. *Remote Sens. Environ.* 87, 507–520. <https://doi.org/10.1016/j.rse.2003.03.001>
- Asner, G.P., Knapp, D.E., Balaji, A., Paez-Acosta, G., 2009a. Automated mapping of tropical deforestation and forest degradation: CLASlite. *J. Appl. Remote Sens.* 3, 33543. <https://doi.org/10.1117/1.3223675>
- Asner, G.P., Levick, S.R., 2012. Landscape-scale effects of herbivores on treefall in African savannas. *Ecol. Lett.* 15, 1211–1217. <https://doi.org/10.1111/j.1461-0248.2012.01842.x>
- Asner, G.P., Levick, S.R., Kennedy-Bowdoin, T., Knapp, D.E., Emerson, R., Jacobson, J., Colgan, M.S., Martin, R.E., 2009b. Large-scale impacts of herbivores on the structural diversity of African savannas. *Proc. Natl. Acad. Sci. U. S. A.* 106, 4947–4952. <https://doi.org/10.1073/pnas.0810637106>

- Asner, G.P., Lobell, D.B., 2000. A biogeophysical approach for automated SWIR unmixing of soils and vegetation. *Remote Sens. Environ.* 74, 99–112. [https://doi.org/10.1016/S0034-4257\(00\)00126-7](https://doi.org/10.1016/S0034-4257(00)00126-7)
- Asner, G.P., Vaughn, N., Smit, I.P.J., Levick, S., 2016. Ecosystem-scale effects of megafauna in African savannas. *Ecography (Cop.)*. 39, 240–252. <https://doi.org/10.1111/ecog.01640>
- Axelsson, C.R., Hanan, N.P., 2018. Rates of woody encroachment in African savannas reflect water constraints and fire disturbance. *J. Biogeogr.* 45, 1209–1218. <https://doi.org/10.1111/jbi.13221>
- Bakker, E.S., Gill, J.L., Johnson, C.N., Vera, F.W.M., Sandom, C.J., Asner, G.P., Svenning, J.-C., 2016. Combining paleo-data and modern exclosure experiments to assess the impact of megafauna extinctions on woody vegetation. *Proc. Natl. Acad. Sci.* 113, 847–855. <https://doi.org/10.1073/pnas.1502545112>
- Balmford, A., Green, J.M.H., Anderson, M., Beresford, J., Huang, C., Naidoo, R., Walpole, M., Manica, A., 2015. Walk on the Wild Side: Estimating the Global Magnitude of Visits to Protected Areas. *PLoS Biol.* 13, 1–6. <https://doi.org/10.1371/journal.pbio.1002074>
- Barnosky, A.D., Lindsey, E.L., Villavicencio, N.A., Bostelmann, E., Hadly, E.A., Wanket, J., Marshall, C.R., 2015. Variable impact of late-Quaternary megafaunal extinction in causing ecological state shifts in North and South America. *Proc. Natl. Acad. Sci.* 113, 1–6. <https://doi.org/10.1073/pnas.1505295112>
- Bastin, J., Berrahmouni, N., Grainger, A., Maniatis, D., Mollicone, D., Moore, R., Patriarca, C., Picard, N., Sparrow, B., Abraham, E.M., Aloui, K., Atesoglu, A., Attore, F., Bassüllü, Ç., Bey, A., Garzuglia, M., García-Montero, L.G., Groot, N., Guerin, G., Laestadius, L., Lowe, A.J., Mamane, B., Marchi, G., Patterson, P., Rezende, M., Ricci, S., Salcedo, I., Diaz, A.S.-P., Stolle, F., Surappaeva, V., Castro, R., 2017. The extent of forest in dryland biomes. *Science (80-.)*. 356, 635–638. <https://doi.org/10.1126/science.aam6527>
- Baudena, M., Dekker, S.C., Van Bodegom, P.M., Cuesta, B., Higgins, S.I., Lehsten, V., Reick, C.H., Rietkerk, M., Scheiter, S., Yin, Z., Zavala, M.A., Brovkin, V., 2015. Forests, savannas, and grasslands: Bridging the knowledge gap between ecology and Dynamic Global Vegetation Models. *Biogeosciences* 12, 1833–1848. <https://doi.org/10.5194/bg-12-1833-2015>
- Beguería, S., Vicente-Serrano, S.M., Angulo-Martínez, M., Beguería, S., Vicente-Serrano, S.M., Angulo-Martínez, M., 2010. A Multiscalar Global Drought Dataset: The SPEIbase: A New Gridded Product for the Analysis of Drought Variability and Impacts. *Bull. Am. Meteorol. Soc.* 91, 1351–1356. <https://doi.org/10.1175/2010BAMS2988.1>
- Bell, R.H.V., 1983. Decision-making in wildlife management with reference to problems of overpopulation , Pretoria ., in: Owen-Smith, R.N. (Ed.), *Management of Large Mammals in African Conservation Areas*. HAUM, Pretoria, pp. 145–172.

- Belsky, A.J., Amundson, R.G., Duxbury, J.M., Riha, S.J., Ali, A.R., Mwongat, S.M., 1989. The Effects of Trees on Their Physical , Chemical and Biological Environments in a Semi-Arid Savanna in Kenya. *J. Appl. Ecol.* 26, 1005–1024.
- Ben-Shahar, R., 1993. Patterns of elephant damage to vegetation in northern Botswana. *Biol. Conserv.* 65, 249–256. [https://doi.org/10.1016/0006-3207\(93\)90057-8](https://doi.org/10.1016/0006-3207(93)90057-8)
- Bey, A., Díaz, A.S.P., Maniatis, D., Marchi, G., Mollicone, D., Ricci, S., Bastin, J.F., Moore, R., Federici, S., Rezende, M., Patriarca, C., Turia, R., Gamoga, G., Abe, H., Kaidong, E., Miceli, G., 2016. Collect earth: Land use and land cover assessment through augmented visual interpretation. *Remote Sens.* 8, 1–24. <https://doi.org/10.3390/rs8100807>
- Birkett, A., 2002. The impact of giraffe, rhino and elephant on the habitat of a black rhino sanctuary in Kenya. *Afr. J. Ecol.* 40, 276–282. <https://doi.org/10.1046/j.1365-2028.2002.00373.x>
- Birkett, A., Stevens-Wood, B., 2005. Effect of low rainfall and browsing by large herbivores on an enclosed savannah habitat in Kenya. *Afr. J. Ecol.* 43, 123–130. <https://doi.org/10.1111/j.1365-2028.2005.00555.x>
- Blanc, J.J., Barnes, R.F.W., Craig, G.C., Dublin, H.T., Thouless, C.R., Douglas-Hamilton, I., Hart, J. a., 2007. African elephant status report 2007: an update from the African elephant database, Occasional paper series of the IUCN Species Survival Commission, no. 33. IUCN/SSC African Elephant Specialist Group. IUCN, Gland, Switzerland. <https://doi.org/10.2305/IUCN.CH.2007.SSC-OP.33.en>
- Blanc, J.J., Thouless, C.R., Hart, J. a, Dublin, H.T., Douglas-Hamilton, I., Craig, G.C., Barnes, R.F.W., 2003. African elephant status report 2002: An update from the African Elephant Database. <https://doi.org/10.1021/np010647c>
- Bolognesi, M., Vrieling, A., Rembold, F., Gadain, H., 2015. Rapid mapping and impact estimation of illegal charcoal production in southern Somalia based on WorldView-1 imagery. *Energy Sustain. Dev.* 25, 40–49. <https://doi.org/10.1016/j.esd.2014.12.008>
- Bond, W., Keeley, J., 2005. Fire as a global ‘herbivore’: the ecology and evolution of flammable ecosystems. *Trends Ecol. Evol.* 20, 387–394. <https://doi.org/10.1016/j.tree.2005.04.025>
- Bond, W.J., 2008. What limits trees in C4 grasslands and savannas? *Annu. Rev. Ecol. Evol. Syst.* 39, 641–659.
- Bond, W.J., 2005. Large parts of the world are brown or black : A different view on the ‘ Green World ’ hypothesis. *J. Veg. Sci.* 16, 261–266. [https://doi.org/10.1658/1100-9233\(2005\)016](https://doi.org/10.1658/1100-9233(2005)016)
- Bond, W.J., Woodward, F.I., Midgley, G.F., 2005. The global distribution of ecosystems in a world without fire. *New Phytol.* 165, 525–538. <https://doi.org/10.1111/j.1469-8137.2004.01252.x>

- Bowman, D.M.J.S., Balch, J., Artaxo, P., Bond, W.J., Cochrane, M. a., D'Antonio, C.M., DeFries, R., Johnston, F.H., Keeley, J.E., Krawchuk, M. a., Kull, C. a., Mack, M., Moritz, M. a., Pyne, S., Roos, C.I., Scott, A.C., Sodhi, N.S., Swetnam, T.W., 2011. The human dimension of fire regimes on Earth. *J. Biogeogr.* 38, 2223–2236.
<https://doi.org/10.1111/j.1365-2699.2011.02595.x>
- Brandt, M., Hiernaux, P., Tagesson, T., Verger, A., Rasmussen, K., Diouf, A.A., Mbow, C., Mougin, E., Fensholt, R., 2016. Woody plant cover estimation in drylands from Earth Observation based seasonal metrics. *Remote Sens. Environ.* 172, 28–38.
<https://doi.org/10.1016/j.rse.2015.10.036>
- Brandt, M., Tappan, G., Diouf, A.A., Beye, G., Mbow, C., Fensholt, R., 2017. Woody vegetation die off and regeneration in response to rainfall variability in the west african sahel. *Remote Sens.* 9. <https://doi.org/10.3390/rs9010039>
- Breiman, L., 2001. Random forests. *Mach. Learn.* 45, 5–32.
<https://doi.org/10.1023/A:1010933404324>
- Bucini, G., Saatchi, S., Hanan, N., Boone, R.B., Smit, I., 2009. Woody cover and heterogeneity in the savannas of the Kruger National Park, South Africa, in: *International Geoscience and Remote Sensing Symposium (IGARSS)*. pp. 334–337.
<https://doi.org/10.1109/IGARSS.2009.5417381>
- Buechner, H.K., Dawkins, H.C., 1961. Vegetation Change Induced by Elephants and Fire in Murchison Falls National Park, Uganda. *Ecology* 42, 752–766.
- Burrill, A., Douglas-Hamilton, I., 1987. African Elephant Database Project: Final Report.
- Campling, P., Gobin, A., Feyen, J., 2002. Logistic Modeling to Spatially Predict the Probability of Soil Drainage Classes. *Soil Sci. Soc. Am. J.* 66, 1390–1401.
<https://doi.org/10.2136/sssaj2002.1390>
- Chafota, J., Owen-Smith, N., 2009. Episodic severe damage to canopy trees by elephants: interactions with fire, frost and rain. *J. Trop. Ecol.* 25, 341.
<https://doi.org/10.1017/S0266467409006051>
- Channan, S., Collins, K., Emanuel, W.R., 2014. Global mosaics of the standard MODIS land cover type data.
- Chase, M.J., Schlossberg, S., Griffin, C.R., Bouché, P.J.C., Djene, S.W., Elkan, P.W., Ferreira, S., Grossman, F., Kohi, E.M., Landen, K., Omondi, P., 2016. Continent-wide survey reveals massive decline in African savannah elephants. *PeerJ* 1–24.
<https://doi.org/10.7717/peerj.2354>
- Chatterjee, A., Hadi, A.S., 2012. Regression analysis by example. Wiley & Sons Inc, Hoboken, New Jersey, USA. <https://doi.org/https://doi.org/10.1136/bmj.f4488>

- Choodarathnakara, A.L., Kumar, T.A., Koliwad, S., 2012. Mixed Pixels: A Challenge in Remote Sensing Data Classification for Improving Performance. *Int. J. Adv. Res. Comput. Eng. Technol.* 1, 2278–1323.
- CIESIN, 2016. Gridded Population of the World, Version 4 (GPWv4): Population Count.
- Cochard, R., Edwards, P.J., 2011. Tree dieback and regeneration in secondary *Acacia zanzibarica* woodlands on an abandoned cattle ranch in coastal Tanzania. *J. Veg. Sci.* 22, 490–502. <https://doi.org/10.1111/j.1654-1103.2011.01276.x>
- Codron, J., Codron, D., Lee-Thorp, J. a., Sponheimer, M., Kirkman, K., Duffy, K.J., Sealy, J., 2011. Landscape-scale feeding patterns of African elephant inferred from carbon isotope analysis of feces. *Oecologia* 165, 89–99. <https://doi.org/10.1007/s00442-010-1835-6>
- Codron, J., Codron, D., Sponheimer, M., Kirkman, K., Duffy, K.J., Raubenheimer, E.J., Melice, J.-L., Grant, R., Clauss, M., Lee-Thorp, J.A., 2012. Stable isotope series from elephant ivory reveal lifetime histories of a true dietary generalist. *Proc. R. Soc. B Biol. Sci.* 279, 2433–2441. <https://doi.org/10.1098/rspb.2011.2472>
- Codron, J., Kirkman, K., Duffy, K.J., Sponheimer, M., Lee-thorp, J. a, Ganswindt, A., Clauss, M., Codron, D., 2013. Stable isotope turnover and variability in tail hairs of captive and free-ranging African elephants (*Loxodonta africana*) reveal dietary niche differences within populations. *Can. J. Zool.* 91, 124–134. <https://doi.org/10.1139/cjz-2012-0155>
- Codron, J., Lee-Thorp, J. a., Sponheimer, M., Codron, D., Grant, R.C., de Ruiter, D.J., 2006. Elephant (*Loxodonta Africana*) Diets in Kruger National Park, South Africa: Spatial and Landscape Differences. *J. Mammal.* 87, 27–34. <https://doi.org/10.1644/05-MAMM-A-017R1.1>
- Croze, H., 1974. The Seronera bull problem: I. the elephants. *East African Wildl. J.* 12, 1–27.
- Dahlin, K.M., Fisher, R.A., Lawrence, P.J., 2015. Environmental drivers of drought deciduous phenology in the Community Land Model. *Biogeosciences* 12, 5061–5074. <https://doi.org/10.5194/bgd-12-5803-2015>
- Dahlin, K.M., Ponte, D. Del, Setlock, E., Nagelkirk, R., 2017. Global patterns of drought deciduous phenology in semi-arid and savanna-type ecosystems. *Ecography (Cop.)*. 40, 314–323. <https://doi.org/10.1111/ecog.02443>
- Daskin, J.H., Stalmans, M., Pringle, R.M., 2016. Ecological legacies of civil war: 35-year increase in savanna tree cover following wholesale large-mammal declines. *J. Ecol.* 104, 79–89. <https://doi.org/10.1111/1365-2745.12483>
- Davies, A.B., Asner, G.P., 2019. Elephants limit aboveground carbon gains in African savannas. *Glob. Chang. Biol.* 25, 1368–1382. <https://doi.org/10.1111/gcb.14585>
- Dawelbait, M., Morari, F., 2008. Limits and potentialities of studying dryland vegetation using

- the optical remote sensing. *Ital. J. Agron.* 3, 97–106. <https://doi.org/10.4081/ija.2008.97>
- De Bie, S., Ketner, P., Paasse, M., Geerling, C., 1998. Woody plant phenology in the West Africa savanna. *J. Biogeogr.* 25, 883–900. <https://doi.org/doi:10.1046/j.1365-2699.1998.00229.x>
- Degerickx, J., Roberts, D.A., Somers, B., 2019. Enhancing the performance of Multiple Endmember Spectral Mixture Analysis (MESMA) for urban land cover mapping using airborne lidar data and band selection. *Remote Sens. Environ.* 221, 260–273. <https://doi.org/10.1016/j.rse.2018.11.026>
- Dennison, P.E., Halligan, K.Q., Roberts, D.A., 2004. A comparison of error metrics and constraints for multiple endmember spectral mixture analysis and spectral angle mapper. *Remote Sens. Environ.* 93, 359–367. <https://doi.org/10.1016/j.rse.2004.07.013>
- Dennison, P.E., Roberts, D.A., 2003. Endmember selection for multiple endmember spectral mixture analysis using endmember average RMSE. *Remote Sens. Environ.* 87, 123–135. [https://doi.org/10.1016/S0034-4257\(03\)00135-4](https://doi.org/10.1016/S0034-4257(03)00135-4)
- Devine, A.P., Stott, I., McDonald, R.A., Maclean, I.M.D., 2015. Woody cover in wet and dry African savannas after six decades of experimental fires. *J. Ecol.* 103, 473–478. <https://doi.org/10.1111/1365-2745.12367>
- Doughty, C.E., 2017. Herbivores increase the global availability of nutrients over millions of years. *Nat. Ecol. Evol.* 1, 1820–1827. <https://doi.org/10.1038/s41559-017-0341-1>
- Doughty, C.E., Faurby, S., Svenning, J.C., 2016. The impact of the megafauna extinctions on savanna woody cover in South America. *Ecography (Cop.)*. 39, 213–222. <https://doi.org/10.1111/ecog.01593>
- Doughty, C.E., Roman, J., Faurby, S., Wolf, A., Haque, A., Bakker, E.S., Malhi, Y., Dunning, J.B., Svenning, J.-C., 2015. Global nutrient transport in a world of giants. *Proc. Natl. Acad. Sci.* 1–6. <https://doi.org/10.1073/pnas.1502549112>
- Doughty, C.E., Wolf, A., Field, C.B., 2010. Biophysical feedbacks between the Pleistocene megafauna extinction and climate: The first human-induced global warming? *Geophys. Res. Lett.* 37, 1–5. <https://doi.org/10.1029/2010GL043985>
- Doughty, C.E., Wolf, A., Malhi, Y., 2013. The legacy of the Pleistocene megafauna extinctions on nutrient availability in Amazonia. *Nat. Geosci.* 6, 761–764. <https://doi.org/10.1038/ngeo1895>
- Dublin, H.T., Douglas-Hamilton, I., 1987. Status and trends of elephants in the Serengeti-Mara ecosystem. *Afr. J. Ecol.* 25, 19–33. <https://doi.org/10.1111/j.1365-2028.1987.tb01087.x>
- Dublin, H.T., Sinclair, A.R.E., McGlade, J., 1990. Elephants and Fire as Causes of Multiple Stable States in the Serengeti-Mara Woodlands. *J. Anim. Ecol.* 59, 1147–1164.

- Duffy, J.P., Pettorelli, N., 2012. Exploring the relationship between NDVI and African elephant population density in protected areas. *Afr. J. Ecol.* 50, 455–463.
<https://doi.org/10.1111/j.1365-2028.2012.01340.x>
- Edkins, M.T., Krugar, L.M., Harris, K., Midgley, J.J., 2008. Baobabs and elephants in Krugar National Park: nowhere to hide. *Afr. J. Ecol.* 46, 119–125.
- Eltringham, S.K., 1979. The ecology and conservation of large African mammals. Macmillan, London.
- Enquist, B.J., Abraham, A.J., Harfoot, M.B.J., Malhi, Y., Doughty, C.E., 2020. The megabiota are disproportionately important for biosphere functioning. *Nat. Commun.* 11, 1–11.
<https://doi.org/10.1038/s41467-020-14369-y>
- Farr, T.G., Rosen, P.A., Caro, E., Crippen, R., Duren, R., Hensley, S., Kobrick, M., Paller, M., Rodriguez, E., Roth, L., Seal, D., Shaffer, S., Shimada, J., Umland, J., Werner, M., Oskin, M., Burbank, D., Alsdorf, D., 2007. The Shuttle Radar Topography Mission. *Rev. Geophys.* 45, RG2004. <https://doi.org/10.1029/2005RG000183>
- Favier, C., Aleman, J., Bremond, L., Dubois, M.A., Freycon, V., Yangakola, J.M., 2012. Abrupt shifts in African savanna tree cover along a climatic gradient. *Glob. Ecol. Biogeogr.* 21, 787–797. <https://doi.org/10.1111/j.1466-8238.2011.00725.x>
- Fernández-Manso, A., Quintano, C., Roberts, D., 2012. Evaluation of potential of multiple endmember spectral mixture analysis (MESMA) for surface coal mining affected area mapping in different world forest ecosystems. *Remote Sens. Environ.* 127, 181–193.
<https://doi.org/10.1016/j.rse.2012.08.028>
- Fick, S.E., Hijmans, R.J., 2017. Worldclim 2: New 1-km spatial resolution climate surfaces for global land areas. *Int. J. Climatol.*
- Forster, J.R., 1778. Observations made during a voyage round the world, on physical geography, natural history and ethic philosophy. G. Robinson., London.
- Frei, R., Gaucher, C., Poulton, S.W., Canfield, D.E., 2009. Fluctuations in Precambrian atmospheric oxygenation recorded by chromium isotopes. *Nature* 461, 250–253.
<https://doi.org/10.1038/nature08266>
- Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., Husak, G., Rowland, J., Harrison, L., Hoell, A., Michaelsen, J., 2015. The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes. *Sci. Data* 2, 150066.
<https://doi.org/10.1038/sdata.2015.66>
- Gasparri, N.I., Parmuchi, M.G., Bono, J., Karszenbaum, H., Montenegro, C.L., 2010. Assessing multi-temporal Landsat 7 ETM+ images for estimating above-ground biomass in subtropical dry forests of Argentina. *J. Arid Environ.* 74, 1262–1270.
<https://doi.org/10.1016/j.jaridenv.2010.04.007>

- Gessner, U., Machwitz, M., Conrad, C., Dech, S., 2013. Estimating the fractional cover of growth forms and bare surface in savannas. A multi-resolution approach based on regression tree ensembles. *Remote Sens. Environ.* 129, 90–102.
<https://doi.org/10.1016/j.rse.2012.10.026>
- Giglio, L., Justice, C., Boschetti, L., Roy, D., 2015. MCD64A1 MODIS/Terra+Aqua Burned Area Monthly L3 Global 500m SIN Grid V006 [Dataset] [WWW Document]. NASA EOSDIS L. Process. DAAC. URL <https://doi.org/10.5067/MODIS/MCD64A1.006> (accessed 1.13.20).
- Gizachew, B., Solberg, S., Næsset, E., Gobakken, T., Bollandsås, O.M., Breidenbach, J., Zahabu, E., Mauya, E.W., 2016. Mapping and estimating the total living biomass and carbon in low-biomass woodlands using Landsat 8 CDR data. *Carbon Balance Manag.* 11.
<https://doi.org/10.1186/s13021-016-0055-8>
- Good, S.P., Caylor, K.K., 2011. Climatological determinants of woody cover in Africa. *Proc. Natl. Acad. Sci.* 108, 4902–4907. <https://doi.org/10.1073/pnas.1013100108>
- Google Fusion Tables Team, 2019. Notice : Google Fusion Tables Turndown [WWW Document]. URL <https://support.google.com/fusiontables/answer/9185417?hl=en>
- Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., Moore, R., 2017. Google Earth Engine: Planetary-scale geospatial analysis for everyone.
<https://doi.org/10.1016/j.rse.2017.06.031>
- Govender, N., Trollope, W.S.W., Van Wilgen, B.W., 2006. The effect of fire season, fire frequency, rainfall and management on fire intensity in savanna vegetation in South Africa. *J. Appl. Ecol.* 43, 748–758. <https://doi.org/10.1111/j.1365-2664.2006.01184.x>
- Gray, E.F., Bond, W.J., 2013. Will woody plant encroachment impact the visitor experience and economy of conservation areas? *Koedoe* 55, 1–9.
<https://doi.org/10.4102/koedoe.v55i1.1106>
- Guldemon, R., van Aarde, R., 2008. A Meta-Analysis of the Impact of African Elephants on Savanna Vegetation. *J. Wildl. Manage.* 72, 892–899. <https://doi.org/10.2193/2007-072>
- Guldemon, R.A.R., Purdon, A., van Aarde, R.J., 2017. A systematic review of elephant impact across Africa. *PLoS One* 12, 1–12. <https://doi.org/10.1371/journal.pone.0178935>
- Hanan, N.P., Tredennick, A.T., Prihodko, L., Bucini, G., Dohn, J., 2015. Analysis of stable states in global savannas - A response to Staver and Hansen. *Glob. Ecol. Biogeogr.* 24, 988–989.
<https://doi.org/10.1111/geb>
- Hanan, N.P., Tredennick, A.T., Prihodko, L., Bucini, G., Dohn, J., 2014. Analysis of stable states in global savannas : is the CART pulling the horse? *Glob. Ecol. Biogeogr.* 23, 259–263.
<https://doi.org/10.1111/geb.12122>

- Hansen, M.C., DeFries, R.S., Townshend, J.R.G., Carroll, M., Dimiceli, C., Sohlberg, R.A., 2003. Global Percent Tree Cover at a Spatial Resolution of 500 Meters: First Results of the MODIS Vegetation Continuous Fields Algorithm. *Earth Interact.* 7, 1–15. [https://doi.org/10.1175/1087-3562\(2003\)007<0001:GPTCAA>2.0.CO;2](https://doi.org/10.1175/1087-3562(2003)007<0001:GPTCAA>2.0.CO;2)
- 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 Sens. Environ.* 83, 303–319. [https://doi.org/10.1016/S0034-4257\(02\)00079-2](https://doi.org/10.1016/S0034-4257(02)00079-2)
- Hansen, M.C., Loveland, T.R., 2012. A review of large area monitoring of land cover change using Landsat data. *Remote Sens. Environ.* 122, 66–74. <https://doi.org/10.1016/j.rse.2011.08.024>
- Hansen, M.C., Potapov, P. V., Moore, R., Hancher, M., Turubanova, S.A., Tyukavina, A., Thau, D., Stehman, S. V., Goetz, S.J., Loveland, T.R., Kommareddy, A., Egorov, A., Chini, L., Justice, C.O., Townshend, J.R.G., 2013. High-Resolution Global Maps of 21st-Century Forest Cover Change. *Science* (80-.). 342, 850–854. <https://doi.org/10.1126/science.1244693>
- Hantson, S., Scheffer, M., Pueyo, S., Xu, C., Lasslop, G., Van Nes, E.H., Holmgren, M., Mendelsohn, J., 2017. Rare, Intense, Big fires dominate the global tropics under drier conditions. *Sci. Rep.* 7, 7–11. <https://doi.org/10.1038/s41598-017-14654-9>
- Hawbaker, T.J., Vanderhoof, M.K., Beal, Y.J., Takacs, J.D., Schmidt, G.L., Falgout, J.T., Williams, B., Fairaux, N.M., Caldwell, M.K., Picotte, J.J., Howard, S.M., Stitt, S., Dwyer, J.L., 2017. Mapping burned areas using dense time-series of Landsat data. *Remote Sens. Environ.* 198, 504–522. <https://doi.org/10.1016/j.rse.2017.06.027>
- Hayward, M.W., Zawadzka, B., 2010. Increasing elephant *Loxodonta africana* density is a more important driver of change in vegetation condition than rainfall. *Acta Theriol. (Warsz)*. 55, 289–299. <https://doi.org/10.4098/j.at.0001-7051.076.2009>
- Hema, E.M., Barnes, R.F.W., Di Vittorio, M., Luiselli, L., Guenda, W., 2017. Selective disturbance by elephants (*Loxodonta africana*) on eight tree species in a West African savannah. *Ecol. Res.* 32, 205–214. <https://doi.org/10.1007/s11284-016-1431-2>
- Higgins, S.I., Bond, W.J., February, E.C., Bronn, A., Euston-Brown, D.I.W., Enslin, B., Govender, N., Rademan, L., O'Regan, S., Potgieter, A.L.F., Scheiter, S., Sowry, R., Trollope, L., Trollope, W.S.W., 2007. Effects of four decades of fire manipulation on woody vegetation structure in savanna. *Ecology* 88, 1119–1125. <https://doi.org/10.1890/09-0929.1>
- Hirota, M., Holmgren, M., Van Nes, E.H., Scheffer, M., 2011. Global Resilience of Tropical Forest and Savanna to Critical Transitions. *Science* (80-.). 334, 232–235. <https://doi.org/10.1126/science.1210657>

- Hoare, R.E., Du Toit, J.T., 1999. Coexistence between people and elephants in African savannas. *Conserv. Biol.* 13, 633–639. <https://doi.org/10.1046/j.1523-1739.1999.98035.x>
- Holdo, R.M., Holt, R.D., Fryxell, J.M., 2009a. Grazers, browsers, and fire influence the extent and spatial pattern of tree cover in the Serengeti. *Ecol. Appl.* 19, 95–109. <https://doi.org/10.1890/07-1954.1>
- Holdo, R.M., Sinclair, A.R.E., Dobson, A.P., Metzger, K.L., Bolker, B.M., Ritchie, M.E., Holt, R.D., 2009b. A disease-mediated trophic cascade in the Serengeti and its implications for ecosystem C. *PLoS Biol.* 7. <https://doi.org/10.1371/journal.pbio.1000210>
- Holdridge, L.R., 1947. Determination of World Plant Formations from Simple Climatic Data. *Science* (80-.). 105, 367–368.
- Holling, C.S., 1973. Resilience and stability of ecological systems. *Annu. Rev. Ecol. Syst.* 4, 1–23. <https://doi.org/10.1146/annurev.es.04.110173.000245>
- Horion, S., Fensholt, R., Tagesson, T., Ehammer, A., 2014. Using earth observation-based dry season NDVI trends for assessment of changes in tree cover in the Sahel. *Int. J. Remote Sens.* 35, 2493–2515. <https://doi.org/10.1080/01431161.2014.883104>
- Jia, S., Wang, X., Yuan, Z., Lin, F., Ye, J., Hao, Z., Luskin, M.S., 2018. Global signal of top-down control of terrestrial plant communities by herbivores. *Proc. Natl. Acad. Sci. U. S. A.* 1–6. <https://doi.org/10.1073/pnas.1707984115>
- JRC, CIESIN, 2015. GHS-POP R2015A - GHS population grid, derived from GPW4, multitemporal (1975, 1990, 2000, 2015) [WWW Document]. Eur. Comm. Jt. Res. Cent. URL http://data.europa.eu/88u/dataset/jrc-ghsl-ghs_pop_gpw4_globe_r2015a
- Kalwij, J.M., De Boer, W.F., Mucina, L., Prins, H.H.T., Skarpe, C., Winterbach, C., 2010. Tree cover and biomass increase in a southern African savanna despite growing elephant population. *Ecol. Appl.* 20, 222–233. <https://doi.org/10.1890/09-0541.1>
- Kerkhoff, A.J., Moriarty, P.E., Weiser, M.D., 2014. The latitudinal species richness gradient in New World woody angiosperms is consistent with the tropical conservatism hypothesis. *Proc. Natl. Acad. Sci.* 111, 8125–8130. <https://doi.org/10.1073/pnas.1308932111>
- Kopp, R.E., Kirschvink, J.L., Hilburn, I.A., Nash, C.Z., 2005. The Paleoproterozoic snowball Earth: A climate disaster triggered by the evolution of oxygenic photosynthesis. *Proc. Natl. Acad. Sci.* 102, 11131–11136. <https://doi.org/10.1073/pnas.0504878102>
- Kraaij, T., Milton, S.J., 2006. Vegetation changes (1995-2004) in semi-arid Karoo shrubland, South Africa: Effects of rainfall, wild herbivores and change in land use. *J. Arid Environ.* 64, 174–192. <https://doi.org/10.1016/j.jaridenv.2005.04.009>
- Lamprey, H.F., Glover, P.E., Turner, M.I.M., Bell, R.H.V., 1967. Invasion of the Serengeti National Park by Elephants. *East African Wildl. J.* 5, 151–166.

- Laws, R.M., 1970. Elephants as agents of habitat and landscape change in East Africa. *Oikos* 21, 1–15. <https://doi.org/10.2307/3543832>
- Lawton, W.H., Sylvestre, E.A., 1971. Self Modeling Curve Resolution. *Technometrics* 13, 617–633.
- Legates, D.R., McCabe, G.J., 2013. A refined index of model performance: A rejoinder. *Int. J. Climatol.* 33, 1053–1056. <https://doi.org/10.1002/joc.3487>
- Lehmann, C.E.R., Anderson, T.M., Sankaran, M., Higgins, S.I., Archibald, S., Hoffmann, W.A., Hanan, N.P., Williams, R.J., Fensham, R.J., Felfili, J., Hutley, L.B., Ratnam, J., Jose, J.S., Montes, R., Franklin, D., Russell-smith, J., Ryan, C.M., Durigan, G., Hiernaux, P., Haidar, R., Bowman, D.M.J.S., Bond, W.J., 2014. Savanna Vegetation-Fire-Climate Relationships Differ Among Continents. *Science* (80-.). 343, 548–553. <https://doi.org/10.1126/science.1247355>
- Levick, S.R., Asner, G.P., 2013. The rate and spatial pattern of treefall in a savanna landscape. *Biol. Conserv.* 157, 121–127. <https://doi.org/10.1016/j.biocon.2012.07.009>
- Levick, S.R., Asner, G.P., Kennedy-Bowdoin, T., Knapp, D.E., 2009. The relative influence of fire and herbivory on savanna three-dimensional vegetation structure. *Biol. Conserv.* 142, 1693–1700. <https://doi.org/10.1016/j.biocon.2009.03.004>
- Li, J., 2017. Assessing the accuracy of predictive models for numerical data: Not r nor r^2 , why not? Then what? *PLoS One* 12, 1–16. <https://doi.org/10.1371/journal.pone.0183250>
- Liaw, A., Wiener, M., 2002. Classification and Regression by randomForest. *R News* 2, 18–22.
- Loarie, S.R., Aarde, R.J. Van, Pimm, S.L., 2009a. Fences and artificial water affect African savannah elephant movement patterns. *Biol. Conserv.* 142, 3086–3098. <https://doi.org/10.1016/j.biocon.2009.08.008>
- Loarie, S.R., van Aarde, R.J., Pimm, S.L., 2009b. Elephant seasonal vegetation preferences across dry and wet savannas. *Biol. Conserv.* 142, 3099–3107. <https://doi.org/10.1016/j.biocon.2009.08.021>
- Loecher, M., Ropkins, K., 2015. RgoogleMaps and loa: Unleashing R Graphics Power on Map Tiles. *J. Stat. Softw.* 63, 1–18.
- MacGarigal, K., Cushman, S., Stafford, S., 2000. Multivariate statistics for wildlife and ecology research. Springer Science + Business Media, LLC, New York, New York, USA. <https://doi.org/https://doi.org/10.1007/978-1-4612-1288-1>
- Malhi, Y., Doughty, C.E., Galetti, M., Smith, F.A., Svenning, J., Terborgh, J.W., 2016. Megafauna and ecosystem function from the Pleistocene to the Anthropocene. *Proc. Natl. Acad. Sci.* 113, 838–846. <https://doi.org/10.1073/pnas.1502540113>
- Marsett, R.C., Qi, J., Heilman, P., Biedenbender, S.H., Watson, M.C., Amer, S., Weltz, M.,

- Goodrich, D., Marsett, R., 2006. Remote sensing for grassland management in the arid Southwest. *Rangel. Ecol. Manag.* 59, 530–540. <https://doi.org/10.2111/05-201R.1>
- Marston, C., Aplin, P., Wilkinson, D., Field, R., O'Regan, H., 2017. Scrubbing Up: Multi-Scale Investigation of Woody Encroachment in a Southern African Savannah. *Remote Sens.* 9, 419. <https://doi.org/10.3390/rs9050419>
- May, R.M., 1976. Thresholds and breakpoints in ecosystems with a multiplicity of stable states. *Nature* 260, 471–477. <https://doi.org/10.1038/269471a0>
- Messina, M., Cunliffe, R., Farcomeni, A., Malatesta, L., Smit, I.P.J., Testolin, R., Ribeiro, N.S., Nhancale, B., Vitale, M., Attorre, F., 2018. An innovative approach to disentangling the effect of management and environment on tree cover and density of protected areas in African savanna. *For. Ecol. Manage.* 419–420, 1–9. <https://doi.org/10.1016/j.foreco.2018.03.019>
- Michishita, R., Jiang, Z., Xu, B., 2012. Monitoring two decades of urbanization in the Poyang Lake area, China through spectral unmixing. *Remote Sens. Environ.* 117, 3–18. <https://doi.org/10.1016/j.rse.2011.06.021>
- Midgley, J.J., Lawes, M.J., Chamail-Jammes, S., 2010. Turner review No. 19. Savanna woody plant dynamics: The role of fire and herbivory, separately and synergistically. *Aust. J. Bot.* 58, 1–11. <https://doi.org/10.1071/BT09034>
- Mograbi, P.J., Asner, G.P., Witkowski, E.T.F., Erasmus, B.F.N., Wessels, K.J., Mathieu, R., Vaughn, N.R., 2017. Humans and elephants as treefall drivers in African savannas. *Ecography (Cop.)*. 1274–1284. <https://doi.org/10.1111/ecog.02549>
- Moncrieff, G.R., Kruger, L.M., Midgley, J.J., 2008. Stem mortality of *Acacia nigrescens* induced by the synergistic effects of elephants and fire in Kruger National Park, South Africa. *J. Trop. Ecol.* 24, 655–662. <https://doi.org/10.1017/S0266467408005476>
- Morrison, T.A., Holdo, R.M., Anderson, T.M., 2016. Elephant damage, not fire or rainfall, explains mortality of overstorey trees in Serengeti. *J. Ecol.* 3, 409–418. <https://doi.org/10.1111/1365-2745.12517>
- Motohka, T., Nasahara, K.N., Oguma, H., Tsuchida, S., 2010. Applicability of Green-Red Vegetation Index for remote sensing of vegetation phenology. *Remote Sens.* 2, 2369–2387. <https://doi.org/10.3390/rs2102369>
- Murphy, B.P., Bowman, D.M.J.S., 2012. What controls the distribution of tropical forest and savanna? *Ecol. Lett.* 15, 748–758. <https://doi.org/10.1111/j.1461-0248.2012.01771.x>
- Murphy, P.G., Lugo, A.E., 1986. Ecology of Tropical Dry Forest. *Annu. Rev. Ecol. Syst.* 17, 67–88. <https://doi.org/10.1146/annurev.es.17.110186.000435>
- Nagelkirk, R.L., Dahlin, K.M., 2020. Woody cover fractions in African savannas from Landsat

- and high-resolution imagery. *Remote Sens.* 12, 1–25. <https://doi.org/10.3390/rs12050813>
- Nagelkirk, R.L., Dahlin, K.M., 2019. Data from: Woody cover fractions in African savannas from Landsat and high-resolution imagery [WWW Document]. Mendeley Data. <https://doi.org/10.17632/26djkgjzhf.1>
- Naidoo, L., Cho, M.A., Mathieu, R., Asner, G., 2012. Classification of savanna tree species, in the Greater Kruger National Park region, by integrating hyperspectral and LiDAR data in a Random Forest data mining environment. *ISPRS J. Photogramm. Remote Sens.* 69, 167–179. <https://doi.org/10.1016/j.isprsjprs.2012.03.005>
- Naidoo, R., Fisher, B., Manica, A., Balmford, A., 2016. Estimating economic losses to tourism in Africa from the illegal killing of elephants. *Nat. Commun.* 7, 13379. <https://doi.org/10.1038/ncomms13379>
- Napier Bax, P., Sheldrick, D.L.W., 1963. Some preliminary observations on the food of elephant in the Tsavo Royal National Park (East) of Kenya. *East African Wildl. J.* 1, 40–53.
- Nash, E., Sutcliffe, V., 1970. River flow forecasting through conceptual models Part I - A discussion of principles. *J. Hydrol.* 10, 282–290.
- Neter, J., Wasserman, W., 1974. *Applied linear statistical models: regression, analysis of variance, and experimental designs*, 1st ed. Richard Irwin.
- Nicholson, S.E., 2017. Climate and Climatic Variability of Rainfall over Eastern Africa. *Rev. Geophys.* <https://doi.org/10.1002/2016RG000544>
- Okula, J.P., Sise, W.R., 1986. Effects of Elephant *Loxodonta-Africana* Browsing on *Acacia-Seyal* in Waza National Park Cameroon. *Afr. J. Ecol.* 24, 1–6. <https://doi.org/10.1111/j.1365-2028.1986.tb00335.x>
- Olson, D.M., Dinerstein, E., Wikramanayake, E.D., Burgess, N.D., Powell, G.V.N., Underwood, E.C., D'amico, J. a., Itoua, I., Strand, H.E., Morrison, J.C., Loucks, C.J., Allnutt, T.F., Ricketts, T.H., Kura, Y., Lamoreux, J.F., Wettengel, W.W., Hedao, P., Kassem, K.R., 2001. *Terrestrial Ecoregions of the World: A New Map of Life on Earth*. *Bioscience* 51, 933. [https://doi.org/10.1641/0006-3568\(2001\)051\[0933:TEOTWA\]2.0.CO;2](https://doi.org/10.1641/0006-3568(2001)051[0933:TEOTWA]2.0.CO;2)
- Owen-Smith, N., 1987. Pleistocene Extinctions : The Pivotal Role of Megaherbivores. *Paleobiology* 13, 351–362.
- Owen-Smith, N., Kerley, G.I.H., Page, B., Slotow, R., Van Aarde, R.J., 2006. A scientific perspective on the management of elephants in the Kruger National Park and elsewhere. *S. Afr. J. Sci.* 102, 389–394. <https://doi.org/10.1016/j.actbio.2010.06.025>
- Pachzelt, A., Forrest, M., Rammig, A., Higgins, S.I., Hickler, T., 2015. Potential impact of large ungulate grazers on African vegetation, carbon storage and fire regimes. *Glob. Ecol. Biogeogr.* 24, 991–1002. <https://doi.org/10.1111/geb.12313>

- Pekel, J.-F., Cottam, A., Gorelick, N., Belward, A.S., 2016. High-resolution mapping of global surface water and its long-term changes. *Nature* 540, 1–19. <https://doi.org/10.1038/nature20584>
- Pellegrini, A.F.A., Ahlström, A., Hobbie, S.E., Reich, P.B., Nieradzik, L.P., Staver, A.C., Scharenbroch, B.C., Jumpponen, A., Anderegg, W.R.L., Randerson, J.T., Jackson, R.B., 2018. Fire frequency drives decadal changes in soil carbon and nitrogen and ecosystem productivity. *Nature* 553, 194–198. <https://doi.org/10.1038/nature24668>
- Pellegrini, A.F.A., Pringle, R.M., Govender, N., Hedin, L.O., 2017. Woody plant biomass and carbon exchange depend on elephant-fire interactions across a productivity gradient in African savanna. *J. Ecol.* 105, 111–121. <https://doi.org/10.1111/1365-2745.12668>
- Pengra, B., Long, J., Dahal, D., Stehman, S. V., Loveland, T.R., 2015. A global reference database from very high resolution commercial satellite data and methodology for application to Landsat derived 30m continuous field tree cover data. *Remote Sens. Environ.* 165, 234–248. <https://doi.org/10.1016/j.rse.2015.01.018>
- Pickett, S.T.A., 1989. Space-for-Time Substitution as an Alternative to Long-Term Studies, in: Likens, G.E. (Ed.), *Long-Term Studies in Ecology*. Springer, New York, New York, USA.
- Planavsky, N.J., Asael, D., Hofmann, A., Reinhard, C.T., Lalonde, S. V., Knudsen, A., Wang, X., Ossa Ossa, F., Pecoits, E., Smith, A.J.B., Beukes, N.J., Bekker, A., Johnson, T.M., Konhauser, K.O., Lyons, T.W., Rouxel, O.J., 2014. Evidence for oxygenic photosynthesis half a billion years before the Great Oxidation Event. *Nat. Geosci.* 7, 283–286. <https://doi.org/10.1038/ngeo2122>
- Poitras, T.B., Villarreal, M.L., Waller, E.K., Nauman, T.W., Miller, M.E., Duniway, M.C., 2018. Identifying optimal remotely-sensed variables for ecosystem monitoring in Colorado Plateau drylands. *J. Arid Environ.* 153, 76–87. <https://doi.org/10.1016/j.jaridenv.2017.12.008>
- Polis, G.A., 1999. Why are parts of the world green? Multiple factors control productivity and the distribution of biomass. *Oikos* 86, 3. <https://doi.org/10.2307/3546565>
- Porensky, L.M., Bucher, S.F., Veblen, K.E., Treydte, A.C., Young, T.P., 2013a. Megaherbivores and cattle alter edge effects around ecosystem hotspots in an African savanna. *J. Arid Environ.* 96, 55–63. <https://doi.org/10.1016/j.jaridenv.2013.04.003>
- Porensky, L.M., Wittman, S.E., Riginos, C., Young, T.P., 2013b. Herbivory and drought interact to enhance spatial patterning and diversity in a savanna understory. *Oecologia* 173, 591–602. <https://doi.org/10.1007/s00442-013-2637-4>
- Poulter, B., Frank, D., Ciais, P., Myneni, R.B., Andela, N., Bi, J., Broquet, G., Canadell, J.G., Chevallier, F., Liu, Y.Y., Running, S.W., Sitch, S., van der Werf, G.R., 2014. Contribution of semi-arid ecosystems to interannual variability of the global carbon cycle. *Nature* 509, 600–603. <https://doi.org/10.1038/nature13376>

- Powell, R.L., Roberts, D.A., Dennison, P.E., Hess, L.L., 2007. Sub-pixel mapping of urban land cover using multiple endmember spectral mixture analysis: Manaus, Brazil. *Remote Sens. Environ.* 106, 253–267. <https://doi.org/10.1016/j.rse.2006.09.005>
- Qi, J., Chehbouni, A., Huete, A.R., Kerr, Y.H., Sorooshian, S., 1994. A modified soil adjusted vegetation index. *Remote Sens. Environ.* 48, 119–126. [https://doi.org/10.1016/0034-4257\(94\)90134-1](https://doi.org/10.1016/0034-4257(94)90134-1)
- R Core Team, 2018. R: A language and environment for statistical computing. R Foundation for Statistical Computing.
- Ratajczak, Z., Nippert, J.B., 2012. Comment on “ Global Resilience of to Critical Transitions .” *Science* (80-.). 336, 541c-541d. <https://doi.org/10.1126/science.1219346>
- Reid, R., 2012. *Savannas of Our Birth: People, Wildlife, and Change in East Africa*, 1st ed. University of California Press.
- Ringrose, S., Matheson, W., Mogotsi, B., Tempest, F., 1989. The darkening effect in drought affected savanna woodland environments relative to soil reflectance in Landsat and SPOT wavebands. *Remote Sens. Environ.* 30, 1–19. [https://doi.org/10.1016/0034-4257\(89\)90043-6](https://doi.org/10.1016/0034-4257(89)90043-6)
- Ripple, W.J., Newsome, T.M., Wolf, C., Dirzo, R., Everatt, K.T., Galetti, M., Hayward, M.W., Kerley, G.I.H., Levi, T., Lindsey, P. a, Macdonald, D.W., Malhi, Y., Painter, L.E., Sandom, C.J., Terborgh, J., Van Valkenburgh, B., 2015. Collapse of the world’s largest herbivores. *Sci. Adv.* <https://doi.org/10.1126/sciadv.1400103>
- Roberts, D., Halligan, K., Dennison, P., 2007. VIPER Tools User Manual. V1.5. 1–91.
- Roberts, D.A., Dennison, P.E., Gardner, M.E., Hetzel, Y., Ustin, S.L., Lee, C.T., 2003. Evaluation of the potential of Hyperion for fire danger assessment by comparison to the airborne visible/infrared imaging spectrometer. *IEEE Trans. Geosci. Remote Sens.* 41, 1297–1310. <https://doi.org/10.1109/TGRS.2003.812904>
- Roberts, D.A., Gardner, M., Church, R., Ustin, S., Scheer, G., Green, R.O., 1998. Mapping chaparral in the Santa Monica Mountains using multiple endmember spectral mixture models. *Remote Sens. Environ.* 65, 267–279. [https://doi.org/10.1016/S0034-4257\(98\)00037-6](https://doi.org/10.1016/S0034-4257(98)00037-6)
- Roberts, D.A., Smith, M.O., Adams, J.B., 1993. Green vegetation, nonphotosynthetic vegetation, and soils in AVIRIS data. *Remote Sens. Environ.* 44, 255–269. [https://doi.org/10.1016/0034-4257\(93\)90020-X](https://doi.org/10.1016/0034-4257(93)90020-X)
- Rodgers, W.A., Lobo, J.D., 1980. Elephant control and legal exploitation: 1920 to 1976. *Tanzan. Notes Rec.* 84/85, 25–54.
- Rodriguez-Galiano, V.F., Ghimire, B., Rogan, J., Chica-Olmo, M., Rigol-Sanchez, J.P., 2012.

- An assessment of the effectiveness of a random forest classifier for land-cover classification. *ISPRS J. Photogramm. Remote Sens.* 67, 93–104.
<https://doi.org/10.1016/j.isprsjprs.2011.11.002>
- Roever, C.L., van Aarde, R.J., Leggett, K., 2012. Functional responses in the habitat selection of a generalist mega-herbivore, the African savannah elephant. *Ecography (Cop.)*. 35, 972–982. <https://doi.org/10.1111/j.1600-0587.2012.07359.x>
- Roy, D.P., Kovalskyy, V., Zhang, H.K., Vermote, E.F., Yan, L., Kumar, S.S., Egorov, A., 2016. Characterization of Landsat-7 to Landsat-8 reflective wavelength and normalized difference vegetation index continuity. *Remote Sens. Environ.* 185, 57–70.
<https://doi.org/10.1016/j.rse.2015.12.024>
- Ruxton, G.D., 2006. The unequal variance t-test is an underused alternative to Student's t-test and the Mann-Whitney U test. *Behav. Ecol.* 17, 688–690.
<https://doi.org/10.1093/beheco/ark016>
- Said, M.Y., Chunge, R.N., Craig, G.C., Thouless, C.R., Barnes, R.F.W., Dublin, H.T., 1999. African Elephant Database 1998.
- Said, M.Y., Chunge, R.N., Craig, G.C., Thouless, C.R., Barnes, R.F.W., Dublin, H.T., 1995. African Elephant Database 1995 225.
- Sandom, C.J., Ejrnæs, R., Hansen, M.D.D., Svenning, J.-C., 2014. High herbivore density associated with vegetation diversity in interglacial ecosystems. *Proc. Natl. Acad. Sci. U. S. A.* 111, 4162–7. <https://doi.org/10.1073/pnas.1311014111>
- Sankaran, M., Augustine, D.J., Ratnam, J., 2013a. Native ungulates of diverse body sizes collectively regulate long-term woody plant demography and structure of a semi-arid savanna. *J. Ecol.* 101, 1389–1399. <https://doi.org/10.1111/1365-2745.12147>
- Sankaran, M., Augustine, D.J., Ratnam, J., 2013b. Native ungulates of diverse body sizes collectively regulate long-term woody plant demography and structure of a semi-arid savanna. *J. Ecol.* 101, 1389–1399. <https://doi.org/10.1111/1365-2745.12147>
- 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. <https://doi.org/10.1038/nature04070>
- Sankaran, M., Ratnam, J., Hanan, N., 2008. Woody cover in African savannas: the role of resources, fire and herbivory. *Glob. Ecol. Biogeogr.* 17, 236–245.
<https://doi.org/10.1111/j.1466-8238.2007.00360.x>
- Santiago, L.S., Kitajima, K., Wright, S.J., Mulkey, S.S., 2004. Coordinated changes in

- photosynthesis, water relations and leaf nutritional traits of canopy trees along a precipitation gradient in lowland tropical forest. *Oecologia* 139, 495–502.
<https://doi.org/10.1007/s00442-004-1542-2>
- Scheffer, M., Hirota, M., Holmgren, M., Van Nes, E.H., Chapin, F.S., 2012. Thresholds for boreal biome transitions. *Proc. Natl. Acad. Sci.* 109, 21384–21389.
<https://doi.org/10.1073/pnas.1219844110>
- Schloeder, C.A., Jacobs, N., 2001. Comparison of methods for interpolating soil properties using limited data. *Soil Sci. Soc. Am. J.* 65, 470–479.
- Schmitz, O.J., Wilmers, C.C., Leroux, S.J., Doughty, C.E., Atwood, T.B., Galetti, M., Davies, A.B., Goetz, S.J., 2018. Animals and the zoogeochemistry of the carbon cycle. *Science* (80-.). 362, 1–10. <https://doi.org/10.1126/science.aar3213>
- Scholes, R.J., Archer, S.R., 1997. Tree-Grass Interactions in Savannas. *For. Sci.* 517–544.
- Settle, J.J., Drake, N.A., 1993. Linear mixing and the estimation of ground cover proportions. *Int. J. Remote Sens.* 14, 1159–1177. <https://doi.org/10.1080/01431169308904402>
- Sexton, J.O., Song, X.-P., Feng, M., Noojipady, P., Anand, A., Huang, C., Kim, D.-H., Collins, K.M., Channan, S., DiMiceli, C., Townshend, J.R., 2013. Global, 30-m resolution continuous fields of tree cover: Landsat-based rescaling of MODIS vegetation continuous fields with lidar-based estimates of error. *Int. J. Digit. Earth* 6, 427–448.
<https://doi.org/10.1080/17538947.2013.786146>
- Shannon, G., Druce, D.J., Page, B.R., Eckhardt, H.C., Grant, R., Slotow, R., 2008. The utilization of large savanna trees by elephant in southern Kruger National Park. *J. Trop. Ecol.* 24, 281–289. <https://doi.org/10.1017/S0266467408004951>
- Shannon, G., Thaker, M., Vanak, A.T., Page, B.R., Grant, R., Slotow, R., 2011. Relative Impacts of Elephant and Fire on Large Trees in a Savanna Ecosystem. *Ecosystems* 14, 1372–1381.
<https://doi.org/10.1007/s10021-011-9485-z>
- Skarpe, C., 1992. Dynamics of savanna ecosystems. *J. Veg. Sci.* 3, 293–300.
<https://doi.org/10.2307/3235754>
- Skarpe, C., Aarrestad, P.A., Andreassen, H.P., Dhillion, S.S., Dimakatso, T., du Toit, J.T., Duncan, Halley, J., Hytteborn, H., Makhabu, S., Mari, M., Marokane, W., Masunga, G., Ditshoswane, M., Moe, S.R., Mojaphoko, R., Mosugelo, D., Motsumi, S., Neo-Mahupeleng, G., Ramotadima, M., Rutina, L., Sechele, L., Sejoe, T.B., Stokke, S., Swenson, J.E., Taolo, C., Vandewalle, M., Wegge, P., 2004. The return of the giants: ecological effects of an increasing elephant population. *Ambio* 33, 276–282.
<https://doi.org/10.2307/4315497>
- Skowno, A.L., Thompson, M.W., Hiestermann, J., Ripley, B., West, A.G., Bond, W.J., 2017. Woodland expansion in South African grassy biomes based on satellite observations (1990–

- 2013): general patterns and potential drivers. *Glob. Chang. Biol.* 23, 2358–2369. <https://doi.org/10.1111/gcb.13529>
- Smart, N.O.E., Hatton, J.C., Spence, D.H.N., 1985. The effect of long-term exclusion of large herbivores on vegetation in Murchison Falls National Park, Uganda. *Biol. Conserv.* 33, 229–245. [https://doi.org/10.1016/0006-3207\(85\)90015-1](https://doi.org/10.1016/0006-3207(85)90015-1)
- Smit, I.P.J., Archibald, S., 2019. Herbivore culling influences spatio-temporal patterns of fire in a semiarid savanna. *J. Appl. Ecol.* 56, 711–721. <https://doi.org/10.1111/1365-2664.13312>
- Smit, I.P.J., Asner, G.P., Govender, N., Kennedy-Bowdoin, T., Knapp, D.E., Jacobson, J., 2010. Effects of fire on woody vegetation structure in African savanna. *Ecol. Appl.* 20, 1865–1875. <https://doi.org/10.1890/09-0929.1>
- Smit, I.P.J., Asner, G.P., Govender, N., Vaughn, N.R., van Wilgen, B.W., 2016. An examination of the potential efficacy of high-intensity fires for reversing woody encroachment in savannas. *J. Appl. Ecol.* 53, 1623–1633. <https://doi.org/10.1111/1365-2664.12738>
- Smit, I.P.J., Prins, H.H.T., 2015. Predicting the Effects of Woody Encroachment on Mammal Communities, Grazing Biomass and Fire Frequency in African Savannas. *PLoS One* 10, e0137857. <https://doi.org/10.1371/journal.pone.0137857>
- Smith, M.O., Adams, J.B., Johnson, P.E., 1985. Quantitative determination of mineral types and abundances from reflectance spectra using principal components analysis. *J. Geophys. Res.* 90, C797–C804. <https://doi.org/10.1029/JB090iS02p0C797>
- Staver, A.C., 2018. Prediction and scale in savanna ecosystems. *New Phytol.* 219, 52–57. <https://doi.org/10.1111/nph.14829>
- Staver, A.C., Archibald, S., Levin, S.A., 2011. The Global Extent and Determinants of Savanna and Forest as Alternative Biome States. *Science* (80-.). 334, 230–232. <https://doi.org/10.1126/science.1210465>
- Staver, A.C., Bond, W.J., 2014. Is there a “browse trap”? Dynamics of herbivore impacts on trees and grasses in an African savanna. *J. Ecol.* 102, 595–602. <https://doi.org/10.1111/1365-2745.12230>
- Staver, A.C., Bond, W.J., Stock, W.D., Rensburg, S.J. Van, Waldram, M.S., 2009. Browsing and fire interact to suppress tree density in an African savanna. *Ecol. Appl.* 19, 1909–1919. <https://doi.org/10.1890/08-1907.1>
- Staver, A.C., Hansen, M.C., 2015. Analysis of stable states in global savannas: Is the CART pulling the horse? - a comment. *Glob. Ecol. Biogeogr.* <https://doi.org/10.1111/geb.12285>
- Stevens, N., Erasmus, B.F.N., Archibald, S., Bond, W.J., 2016. Woody encroachment over 70 years in South African savannas: Overgrazing, global change or extinction aftershock? *Philos. Trans. R. Soc. B Biol. Sci.* 371. <https://doi.org/10.1098/rstb.2015.0437>

- Stevens, N., Lehmann, C.E.R., Murphy, B.P., Durigan, G., 2017. Savanna woody encroachment is widespread across three continents. *Glob. Chang. Biol.* 23, 235–244. <https://doi.org/10.1111/gcb.13409>
- Styles, C. V., Skinner, J.D., 2000. The influence of large mammalian herbivores on growth form and utilization of mopane trees, *Colophospermum mopane*, in Botswana's Northern Tuli Game Reserve. *Afr. J. Ecol.* 38, 95–101. <https://doi.org/10.1046/j.1365-2028.2000.00216.x>
- Symeonakis, E., Petroulaki, K., Higginbottom, T., 2016. Landsat-based woody vegetation cover monitoring in Southern African savannahs. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. - ISPRS Arch.* 41, 563–567. <https://doi.org/10.5194/isprsarchives-XLI-B7-563-2016>
- Thouless, C.R., Dublin, H.T., Blanc, J.J., Skinner, D.P., Daniel, T.E., Taylor, R.D., Maisels, F., Frederick, H.L., Bouche, P., 2016. African Elephant Status Report 2016: an update from the African Elephant Database. Gland, Switzerland.
- Traore, S., Tigabu, M., Jouquet, P., Ouedraogo, S.J., Guinko, S., Lepage, M., 2015. Long-term effects of *Macrotermes* termites, herbivores and annual early fire on woody undergrowth community in Sudanian woodland, Burkina Faso. *Flora Morphol. Distrib. Funct. Ecol. Plants* 211, 40–50. <https://doi.org/10.1016/j.flora.2014.12.004>
- Tucker, C.J., 1979. Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sens. Environ.* 8, 127–150. [https://doi.org/10.1016/0034-4257\(79\)90013-0](https://doi.org/10.1016/0034-4257(79)90013-0)
- Van Der Waal, C., De Kroon, H., De Boer, W.F., Heitkönig, I.M.A., Skidmore, A.K., De Knegt, H.J., Van Langevelde, F., Van Wieren, S.E., Grant, R.C., Page, B.R., Slotow, R., Kohi, E.M., Mwakiwa, E., Prins, H.H.T., 2009. Water and nutrients alter herbaceous competitive effects on tree seedlings in a semi-arid savanna. *J. Ecol.* 97, 430–439. <https://doi.org/10.1111/j.1365-2745.2009.01498.x>
- Van Langevelde, F., Van De Vijver, C. a D.M., 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. [https://doi.org/10.1890/0012-9658\(2003\)084\[0337:EOFAHO\]2.0.CO;2](https://doi.org/10.1890/0012-9658(2003)084[0337:EOFAHO]2.0.CO;2)
- van Wilgen, B.W., Govender, N., MacFadyen, S., 2008. An assessment of the implementation and outcomes of recent changes to fire management in the Kruger National Park. *Koedoe* 50, 22–31.
- Vanak, A.T., Shannon, G., Thaker, M., Page, B., Grant, R., Slotow, R., 2012. Biocomplexity in large tree mortality: Interactions between elephant, fire and landscape in an African savanna. *Ecography (Cop.)*. 35, 315–321. <https://doi.org/10.1111/j.1600-0587.2011.07213.x>
- Venables, W.N., Ripley, B.D., 2002. Modern applied statistics with S, Fourth edi. ed. Springer, New York, New York, USA.

- Vogeler, J.C., Braaten, J.D., Slesak, R.A., Falkowski, M.J., 2018. Extracting the full value of the Landsat archive: Inter-sensor harmonization for the mapping of Minnesota forest canopy cover (1973–2015). *Remote Sens. Environ.* 209, 363–374.
<https://doi.org/10.1016/j.rse.2018.02.046>
- Vogeler, J.C., Yang, Z., Cohen, W.B., 2016. Mapping post-fire habitat characteristics through the fusion of remote sensing tools. *Remote Sens. Environ.* 173, 294–303.
<https://doi.org/10.1016/j.rse.2015.08.011>
- von Humboldt, A., Bonpland, A., 1807. *Essay on the Geography of Plants*. Fr. Schoell and J.G. Cotta, Paris. <https://doi.org/10.1017/CBO9781107415324.004>
- Wagenseil, H., Samimi, C., 2007. Woody vegetation cover in Namibian savannahs: a modelling approach based on remote sensing. *Erdkunde* 61, 325–334.
- Ward, D., Hoffman, M.T., Collocott, S.J., 2014. A century of woody plant encroachment in the dry Kimberley savanna of South Africa. *African J. Range Forage Sci.* 31, 107–121.
<https://doi.org/10.2989/10220119.2014.914974>
- Western, D., Lindsay, W.K., 1984. Seasonal herd dynamics of a savanna elephant population. *Afr. J. Ecol.* 22, 229–244. <https://doi.org/10.1111/j.1365-2028.1984.tb00699.x>
- Western, D., Maitumo, D., 2004. Woodland loss and restoration in a savanna park: a 20-year experiment. *Afr. J. Ecol.* 42, 111–121.
- Whyte, I., van Aarde, R., Pimm, S.L., 1998. Managing the elephants of Kruger National Park. *Anim. Conserv.* 1, 77–83. <https://doi.org/10.1111/j.1469-1795.1998.tb00014.x>
- Wolf, A., Doughty, C.E., Malhi, Y., 2013. Lateral Diffusion of Nutrients by Mammalian Herbivores in Terrestrial Ecosystems. *PLoS One* 8, 1–10.
<https://doi.org/10.1371/journal.pone.0071352>
- Wulder, M.A., Coops, N.C., Roy, D.P., White, J.C., Hermosilla, T., 2018. Land cover 2.0. *Int. J. Remote Sens.* 39, 4254–4284. <https://doi.org/10.1080/01431161.2018.1452075>
- Wulder, M.A., White, J.C., Loveland, T.R., Woodcock, C.E., Belward, A.S., Cohen, W.B., Fosnight, E.A., Shaw, J., Masek, J.G., Roy, D.P., 2016. The global Landsat archive: Status, consolidation, and direction. *Remote Sens. Environ.* 185, 271–283.
<https://doi.org/10.1016/j.rse.2015.11.032>
- Yang, J., Weisberg, P.J., Bristow, N.A., 2012. Landsat remote sensing approaches for monitoring long-term tree cover dynamics in semi-arid woodlands: Comparison of vegetation indices and spectral mixture analysis. *Remote Sens. Environ.* 119, 62–71.
<https://doi.org/10.1016/j.rse.2011.12.004>
- Yang, X., Crews, K., 2019. Fractional Woody Cover Mapping of Texas Savanna at Landsat Scale. *Land* 8, 9. <https://doi.org/10.3390/land8010009>

- Young, K.D., Ferreira, S.M., Van Aarde, R.J., 2009. Elephant spatial use in wet and dry savannas of southern Africa. *J. Zool.* 278, 189–205. <https://doi.org/10.1111/j.1469-7998.2009.00568.x>
- Young, T.P., Okello, B., Kinyua, D., Palmer, T.M., 1998. KLEE: a long-term multi-species herbivore exclusion experiment in Laikipia, Kenya. *African J. Range Forage Sci.* 104, 92–104. <https://doi.org/10.1080/10220119.1997.9647929>
- Zhang, H.K., Roy, D.P., 2017. Using the 500 m MODIS land cover product to derive a consistent continental scale 30 m Landsat land cover classification. *Remote Sens. Environ.* 197, 15–34. <https://doi.org/10.1016/j.rse.2017.05.024>