12-2021

A remote sensing perspective: mapping the human footprint in the Zambezi region of Namibia.

Ariel E Weaver
University of Louisville

A REMOTE SENSING PERSPECTIVE: MAPPING THE HUMAN FOOTPRINT IN THE ZAMBEZI REGION OF NAMIBIA

By

Ariel E. Weaver
B.S., University of Louisville, 2015

A Thesis
Submitted to the Faculty of the
College of Arts and Sciences of the University of Louisville
in Partial Fulfillment of the Requirements
for the Degree of'

Master of Science
in Applied Geography

Department of Geography and Environmental Sciences
University of Louisville
Louisville, Kentucky

December 2021

A REMOTE SENSING PERSPECTIVE: MAPPING THE HUMAN FOOTPRINT IN THE ZAMBEZI REGION OF NAMIBIA

By Ariel E. Weaver
B.S., University of Louisville, 2015

A Thesis Approved on:

November 19, 2021

By the following Thesis Committee:

____
Thesis Advisor
Andrea E. Gaughan, PhD

____
Forrest R. Stevens, PhD

____
Narcisa Pricope, PhD UNCW

ACKNOWLEDGEMENTS

There are so many people and organizations that helped me in the process of completing this thesis. I would like to thank Dr. Andrea Gaughan for introducing me to southern Africa and the research projects that I have been associated with while studying at the University of Louisville. Her mentorship has been invaluable throughout my research. I would also like to thank the rest of my committee, Dr. Forrest Stevens and Dr. Narcisa Pricope.

Additionally, I would not have been able to complete this project without the support I received from the US Department of State and the Fulbright Program.

Thank you to my colleagues in the Wildlife Management Department at the University of Namibia for the unique experience of working alongside others in the field, especially Dr. Ekkehard Klingelhoeffer, Dr. Jesaya Nakanyala, Dr. Ezequiel Fabiano, Mr. Jim Kairu, and Ms. Selma Khosmas for their support and friendship.

I would also like to thank Lise Hanssen from the Kwando Carnivore Project, Courtney Dehn-Gurbacki, Kimmy Eide, Christoph Rocher, and Carol Murphy for their friendship and help during various stages of research in the field.

I would like to thank my friends and family for the encouragement and love they have shown me, particularly when I was struggling. I would like to mention specifically the following people who helped make this possible: Chris Deatrick, Elspeth Bloodgood, Sion Weaver, Dana Pattillo, Maia Weaver, Leila Williams, Eden Williams, Oliver Williams,

Brenda Mahler, Matt Horne and Sarah Benton. I am so lucky to have such a strong support network. I love you all.

And most importantly, I would like to thank the people of the Zambezi Region for welcoming me to such a wonderful place.

TABLE OF CONTENTS

ABSTRACT

A REMOTE SENSING PERSPECTIVE: MAPPING THE HUMAN FOOTPRINT IN THE ZAMBEZI REGION OF NAMIBIA

Ariel Weaver

November 19, 2021

The “human footprint” can be used as a general proxy to estimate human activities across the landscape. The human footprint in the Zambezi Region of Namibia is critically important for regional management of conservation efforts and land use planning. The land covers in the Zambezi Region are characteristically difficult to separate spectrally, due to a highly heterogeneous savanna landscape. Object Based Image Analysis (OBIA) and Random Forest (RF) methods are notable for their ability to improve classification accuracies of remotely sensed imagery. In this study, I investigate the extent of the human footprint in the Zambezi Region of Namibia, using OBIA, RF, and a hybrid Object-based Random Forest approach. Results highlight that Object-based approaches score 5-10% better than a pixel-based RF approach in overall accuracy. Further investigation into the human footprint of the Zambezi Region is necessary for regional and local conservation and sustainable development.

LIST OF FIGURES

LIST OF TABLES

LIST OF EQUATIONS

INTRODUCTION

The expansion of settlements and agriculture, as major drivers of land transformations, is a major concern for conservation of terrestrial ecosystems worldwide (Lambin and Turner 2007). Human activities are driving environmental change at an alarming rate across the globe, putting pressure on ecosystems and wildlife populations. As humans encroach further into habitats, human-wildlife conflict becomes more prevalent, as crop predation, injuries and deaths may occur. Few studies investigate the impacts of human activities on ecosystems at a regional scale (Yang et al. 2014) despite much of the work in conservation and sustainable development occurring at regional and local scales. Global scale datasets often lack the accuracy that is needed for regional and local management for conservation and sustainable development. It is necessary to monitor and manage settlement and agricultural expansion for sustainable development and conservation of ecosystems at regional and local levels.

Remote Sensing (RS) provides the tools for scientists to investigate the scope of human land use through the quantification of land cover. As developments in computing capabilities, the increasing availability of medium and high-resolution satellite- imagery, as well as the development of computer vision and machine learning techniques continue, scientists are increasingly able to model land use and land cover at a variety of spatial and temporal scales. RS allows for the investigation of the human footprint using satellite imagery from a regional perspective. Object Based Image Analysis (OBIA) and Random Forest (RF) are two different

RS methodologies that offer a solution to the problem of classifying heterogeneous savanna landscapes, through computer vision and machine learning technologies.

This study seeks to identify a baseline for settlement and agricultural land cover extent through remote sensing analysis for the eastern Zambezi Region of Namibia to aid sustainable development efforts along a central wildlife corridor for a transboundary conservation area in southern Africa. In doing so, the following research objectives are addressed:1) quantify the human footprint for a semi-arid savanna region in Southern Africa where conservation and development needs require a regionally specific land cover product and 2) identify the most robust RS-technique to produce such a map through modeling settlements and agriculture through an OBIA approach, a RF approach, and a hybrid Object-based/RF approach. The resulting dataset will be a snapshot of the rural human footprint in the year 2017 for the eastern Zambezi Region to serve as a baseline for future studies and will be helpful to land managers and planners in the region.

Literature Review

Remote Sensing for Agriculture

Agricultural expansion remains one of many threats to conservation globally (Nagendra et al. 2013). RS data can provide insight into human pressure on the landscape in terms of LULC, management, and other disturbances in and around protected areas (Nagendra et al. 2013). RS has long been used to monitor and analyze agriculture (Shanmugapriya et al. 2019), as it is a cheap alternative for assessing large areas of cropland without destruction of crops or costly and time-consuming surveys (Nabil et al. 2020). Spatial distribution and extent of croplands areas are critical for effective management of crop production and forecasting (Nabil et al. 2020). Other applications of RS include the detection of stress and disturbances to crop fields, forecasting yields, pest and disease management, and monitoring atmospheric dynamics (Wojtowicz, Wojtowicz & Pikarezyk 2016; Shanmugapriya et al. 2019). In Sub-Saharan Africa, the landscapes are complex, and farms are usually small (<2 hectares) (Mohammed et al. 2020). Large uncertainties and discrepancies among different cropland/land cover products have been found over Africa largely due to high heterogeneity of the landscape, frequent cloud cover, and small field size (Nabil et al. 2020).

The Human Footprint

Biodiversity loss, particularly over the last century, is substantially driven by human-driven land transformation (Lambin et al., 2001; Foley et al., 2005; Lambin & Meyfroidt, 2011). The conversion of lands for settlement and agriculture remains one of the biggest threats to natural ecosystems (Lambin & Meyfroidt 2011). Remotely sensed imagery is often used for quantifying and analyzing of land use and land cover (LULC) to inform land managers and policymakers to meet conservation and development goals (Turner et al. 2003). Remote Sensing is an important tool for quantifying LULCC as this multi-scalar lens provides robust ways to analyze how shifting land use and land cover patterns impact or potentially will impact coupled human-environment systems. Human activities and consumption patterns are having an unprecedented effect on Earth's biological, chemical, and geomorphological systems, and with continuous and rapid expansion of human populations across the globe, it is critical to examine these human-driven impacts to move forward with more effective strategies for sustainable conservation and development.

The term "human footprint" denotes the impact of human activities across a landscape

(Sanderson et al. 2002). The ways in which humans alter the biophysical world through LULC are many and inherently complex (e.g., habitat fragmentation, invasive species introductions, and alterations of nitrogen cycles; Vitousek et al. 1996; Lambin et al. 2001; Grimm et al. 2006; Lambin & Meyfroidt 2011). RS provides a myriad of tools and approaches to looking at LULC at a variety of scales. Settlement areas are important for modelling accessibility to various types of resources, exposure to risks, and impacts after events (e.g., floods, fires, mudslides, volcanos, etc.) at local, regional, and global scales. As such, mapping the human footprint and different types of land-cover defined "settlement" classes has been a focus for many researchers, notably at the global scale for creating consistent and comparable data products (Klotz et al. 2016).

The Last of the Wild/Global Human Footprint (GHF) was the first dataset to model the human footprint at a global scale. The most recent global scale model of the human footprint was released in 2018 (Venter et al 2018) and produced human footprint datasets for the years 1993 and 2009. The model was developed using the following indicators for human impacts 1) population density, 2) railroads score, 3) road score, 4) navigable rivers, 5) coastlines, 6) stable lights at night, 7) urban vector layers, and 8) land cover classes (Sanderson et al. 2002). The scale of this dataset in particular may be inappropriate for analyses designed to aid in conservation and development on the ground at more local scales—due to the low-resolution (1km) of the imagery that was used to develop these datasets. GHF limited the scope of agriculture to intensive farming operations, which would exclude the primarily subsistence level farming that is ubiquitous in rural southern Africa. They also noted that agriculture may be confounded with natural woodlands and savanna systems (Sanderson et al. 2018). Many settlement areas in rural southern Africa still do not have electricity, and night-time lights will underestimate settlement areas in areas where electricity is still scarce.

Other global datasets attempt to delineate urban footprints or settlement extents such as Global Urban Footprint (GUF) (Esch et al. 2013), World Settlement Footprint (WSF, formerly GUF+) (Esch et al. 2017; Marconcini et al. 2020), Global Human Settlement Layer (GHSL) (Pesaresi et al. 2016) and the artificial surface layer from the GLOBELAND30 (GLC30) project (Chen et al. 2014). Table 1 shows the spatial resolution and other information pertaining to these global datasets.

Table 1

Description of global scale human footprint models available

Dataset	Citations	Years	Output Resolution	Input Imagery
GHF	Sanderson et al. 2018	1993, 2009	1 km	Defense Meteorological Satellite Program Operational Line Scanner (DMSP-OLS, 30 arc seconds, ~1km at equator)
GUF	Esch et al. 2017	2011-2012 (2013/2014 to fill gaps)	12 m	Terra SAR-X and TanDEM-X
WSF	Marconcini et al. 2020	2014-.2015	30 m	Landsat Sentinel-1A Sentinel-1B
GHSL 2016	Pesaresi et al. 2016	1975, 1990, 2000, 2013/2014 (ad hoc)	30 m	Landsat
GLC30 artificial surfaces layer	Chen et al. 2014	2010	30 m	Landsat

Many scientists advocate for an ecoregional (as opposed to a global) approach to understanding the human footprint for improving conservation outcomes and preserving biodiversity for local and regional interventions (Sanderson et al. 2002; Woolmer et al.,2008). While the GUF dataset is high resolution, ~12 meters, the cost of the commercial imagery they used to create the model is cost prohibitive for many and it is designed to show only built-up areas. While improvements have been made to global settlement datasets, the resolution of the global-scale datasets underestimate settlements in rural and peri-urban areas because they are calibrated to highlight built areas at a global scale and do not capture traditional structures like those found in rural southern Africa (Esch et al. 2013; Johannes 2017).

WSF utilizes multitemporal radar and optical imagery to classify these datasets with an ensemble of Support Vector Machines and post-classification verification (Marconcini et al. 2020). The multitemporal data is used because settlements behave differently than other LULC over time. GHSL uses a data mining technique called Symbolic Machine Learning in a supervised classification approach to leverage large, heterogeneous ancillary datasets on settlements from around the globe (Pesaresi et al. 2016). GUF an unsupervised classification method based on advanced Support Vector Data Description (Esch et al. 2013). The GLC30 approach uses a mixed pixel-based and object-based approach, using pixel-based metrics and different segmentation methods, in a hierarchical classification scheme (Chen et al. 2015).

Some rural traditional buildings and informal settlement structures may be less than 2m (Esch et al. 2017), which makes detection of structures difficult without fine scale imagery due to the size of the pixels of the satellite imagery. GHSL and WSF use Landsat, but are parameterized to locate built areas, excluding agricultural lands and all under-represent small settlement areas. GLC30 also utilizes Landsat imagery and has an agricultural layer, but the artificial surfaces layer is not sufficient to capture the built environment of the Zambezi Region.

Bare ground is often misclassified as settlement in datasets that use multispectral datasets, whereas terrain that is more complex or forested may be misclassified as settlements in models that primarily use radar (Marconcini et al. 2020). GUF uses only the TerraSAR-

X/TanDEM-X satellites, and its' data may be encumbered by the effects of complex terrain or forests. Whereas WSF utilizes Sentinel 1-a and 1-b Synthetic Aperture Radar (SAR) data in addition to Landsat to minimize the errors associated with spectral and radar satellite data in classifying built classes across the globe (Marconcini et al. 2020).

For local land managers, a more detailed map of the extents of settlement and agricultural areas would allow for more sustainable and strategic planning to meet conservation and development goals. The goal of this study is to have a picture of the human footprint as it relates to where people spend the majority of their time. In the Zambezi Region, this predominantly relates to settlement and agricultural areas. This would be beneficial for those making decisions about land use and conservation, particularly because it would provide insight into areas that are critical for maintaining open wildlife corridors or implementing mitigation techniques for reducing human-wildlife conflict.

Savanna Mosaics and Remote Sensing

Savannas are complex, heterogeneous and dynamic ecosystem of trees, grasses, shrubs, and bare ground (Sankaran et al. 2004; Naidoo et al. 2012). Tropical savanna ecosystems account for 1/8th of the global land surface and are home to a significant and growing number of human populations, in addition to livestock grazing and rangelands (Scholes & Archer, 1997). The heterogeneous, semi-arid savannas of southern Africa are highly influenced by a number of variables, including climate, topography, soils, geomorphology, herbivory, fire, and human-driven activities that results in a heterogeneous mixture of grasses, trees, shrubs, and bare ground (Scholes & Archer, 1997). The heterogeneity across the landscape makes it challenging to quantify land cover/land use based on spectral signatures alone (Hurskainen et al. 2019). Additionally, this heterogeneity stems from both the use of natural material for much of the building materials, unpaved roads and walking paths, grazing of wildlife and domestic livestock (cattle and goats, primarily).

Advancements in remote sensing and geographic information systems (GIS) have become increasingly important tools for analyzing landscape patterns and processes (Newton et. al. 2009). Previous attempts at mapping savanna landscapes in southern Africa include mapping small holder croplands in southern Zambia using Spectral Mixture Analysis and Logistic Regression using logit models to significantly improve the classification of cropland accurately (Sweeney et al. 2015). Caixeta (2016) completed a Multiple Endmember Spectral Mixture Analysis of Mayuni Conservancy in Namibia where they were able to capture photosynthetic vegetation and non-photosynthetic vegetation using spectral characteristics to separate the land cover classes. Gibbes et al. (2010) completed an object-based analysis of tree crown coverage using high resolution imagery of the Zambezi Region of Namibia, formerly known as the Caprivi, with an overall accuracy of 84%. Kamwi et al. (2017) time-change analysis of LULC in Zambezi region of Namibia using pixel-based methods with an overall accuracy of 81%. Hurskainen et al. (2019) produced an object-based image analysis (OBIA) classification using Random Forest algorithm for the foothills of Mt. Kilimanjaro with 60% overall accuracy but improved with the inclusion of auxillary datasets by 6-16%. These approaches all vary in technique, spectral and spatial considerations, and result in a varying level of success at capturing the intended landscape patterns.

Wildlife conservation requires larger geographic ranges for management and planning purposes. As communities develop inside major wildlife corridors, a multi-scale approach is necessary for understanding the patterns that could affect people and wildlife populations in these regions. There are different approaches to quantifying landscape through remote sensing, pixel-based and object-based approaches, and both have benefits and limitations. Remote sensing technologies are becoming more powerful as computing technologies improve, in parallel to the rapid developments being made in machine learning (e.g. Support Vector Machines, Random Forest, Neural Networks, etc.) and computer vision (Object Based Image Analysis, OBIA) that allow for nonparametric statistical and/or knowledge-driven analyses and data mining techniques that allow for the incorporation of larger and more complicated datasets to model the Earth's surface (Blaschke et al., 2014; Lary et al., 2016; Chen et al., 2018).

Historical Context of Conservation in southern Africa

Arid and semi-arid ecosystems are complex systems that are comprised of savannas, grasslands, woodlands, and shrublands which are capable of alternatives states based on land cover and species composition, that can shift based on a variety of interactions and non-linear processes occurring at multiple scales (Peters et al. 2006). Shifts between these land cover states is often as a result of the complex interactions and processes, including the historical states and disturbances, environmental factors such as (weather, climate modes, natural and anthropogenic disturbances), soil characteristics and geomorphic processes, transport vectors (aeolian, fluvial, or animal), and the redistribution of resources across land areas.

Over the last few decades, the arid and semi-arid savannas of southern Africa shifted from so-called "fortress conservation" towards Community Based Conservation (CBC) (Büscher & Dietz, 2005). Fortress conservation, in which centralized islands of land set aside for rigid conservation, were codified and enforced by colonial regimes in southern

Africa. The historical displacement and disenfranchisement of southern African communities at the hands of conservation persisted until the fall of the colonial governments in Angola and Mozambique in the 1970s, and during the post-apartheid era in Namibia, Zimbabwe, and South Africa in the 1980s and early 1990s (Büscher & Dietz 2005) shifting towards CBC practices for conservation.

CBC is premised on the idea that communities can and should manage natural resources and be able to benefit from them while at the same time conserve and maintain biodiversity. The popularity of CBCs proliferated throughout southern Africa, with Community Based Natural Resource Management (CBNRM) initiatives sweeping though the region through organizations like Communal Areas Management for Indigenous Resources (CAMPFIRE) in Zimbabwe (Childs 1996), the development of conservancies in Namibia (Mosimane & Silva 2015), and the Peace Parks Foundation (PPF) (Peace Parks 2010) which created numerous Trans-Frontier Conservation Areas (TFCA), including the Kavango-Zambezi TFCA (which contains parts of Angola, Namibia, Botswana, Zambia, and Zimbabwe) and Limpopo TFCA (which lies between Mozambique and South Africa).

The biodiversity of southern Africa is increasingly being relied upon for economic growth, particularly under the frameworks of CBNRM which began being implemented in the 1990s in southern Africa. In neighboring Zimbabwe, the CBNRM framework is CAMPFIRE (Communal Areas Management Programme for Indigenous Resources). In Zambia, CBNRM is managed through Game Management Areas, and in Botswana, they have Community Based Organizations (CBOs). CBNRM varies to a certain degree in each of these countries, however the basic framework behind CBNRM seeks to devolve management of natural resources and land to communities in order to allow these communities to derive benefits from natural resources and ecotourism directly through various economic activities.

Conservancies in Namibia manage communal areas so that communities can form enterprises that benefit the community financially, particularly through hunting and ecotourism (e.g. recreation and photography based tourism) (Naidoo et al. 2016). These community conservation areas are often situated directly within or in close proximity to protected areas, and are important for maintaining connectivity between protected areas, particularly for preserving migration corridors, reducing impacts on wildlife populations and human-wildlife conflict. CBNRM aims to alleviate the pressures of poverty and human-wildlife conflict on these growing communities who are simultaneously trying to balance the needs of wildlife conservation with expansion and development of settlement infrastructure (Craigie et al., 2010; Hoare, 1999).

Kavango Zambezi Transfrontier Conservation Area

In southern Africa, the Kavango-Zambezi Transfrontier Conservation Area (KAZA) remains the largest terrestrial managed conservation globally (520,000 km^2). KAZA spans through parts of Angola, Botswana, Namibia, Zambia, and Zimbabwe, and is considered to be a biodiversity hotspot (Kamwi et al. 2017). KAZA's savanna ecosystems are patchwork mosaic of a variety of land tenures including communal land areas, protected areas, commercial enterprises, and subsistence farming with varied approaches to land management practices, attitudes towards conservation, and legislation across the five member states (Cumming 2008; Stoldt et al. 2020). The goal of KAZA is to manage wildlife and natural resources by bringing together a diverse group of stakeholders at local, national, and regional levels to manage conservation and development at a bioregional level (Stoldt et al. 2020).

This conservation area is also home to many charismatic, threatened and critically endangered species, including the African elephant (*Loxodonta africana*), wild dogs (*Lycaon pictus*), lions (*Panthera leo*), and more (IUCN Red List, 2018). CBNRM in Namibia contributed to a strong recovery of wildlife through social and economic incentives to promote conservation on communal lands (Naidoo et al. 2016). Additionally, it has been found that in Namibian conservancies richer biodiversity is linked to higher incomes and other social/economic benefits and allows for wildlife tourism and conservation as competitive forms of land use that acts as an alternative to expanding agricultural activities (Naidoo et al. 2011). The areas inside and adjacent to protected areas are of particular importance for understanding where to prioritize "large gain/small loss" (where large gain corresponds with areas that are of concentrated gains for conservation and afford a small loss in terms of human land use (DeFries et al. 2005). A regional dataset of the human footprint is necessary for greater understanding of human impacts on habitat quality, ecosystem functioning, and wildlife population dynamics for this region.

In the context of the study region, the predominantly rural, heterogeneous, semi-arid savannas of southern Africa, the distinctions between and within the land cover types are more difficult to disentangle due to the overlapping spectral characteristics in medium resolution imagery, such as that provided by Landsat platforms (Laris, 2005; Gibbes et al., 2010; Mutanga et al. 2012; Nagendra et al. 2013). For the study region, Figure 1, the eastern portion of the Zambezi Region of Namibia among the more rural areas within the study area, structures are built primarily from locally sourced reed, muds, and timber poles. Moreover, the areas surrounding homes are predominately exposed soils with sporadic grasses, shrubs, and trees. The heterogeneity of vegetation and ground cover across and between land cover types of both human-dominated and surrounding savanna landscape makes it difficult to separate land cover type using more traditional, pixel-based remote sensing techniques.

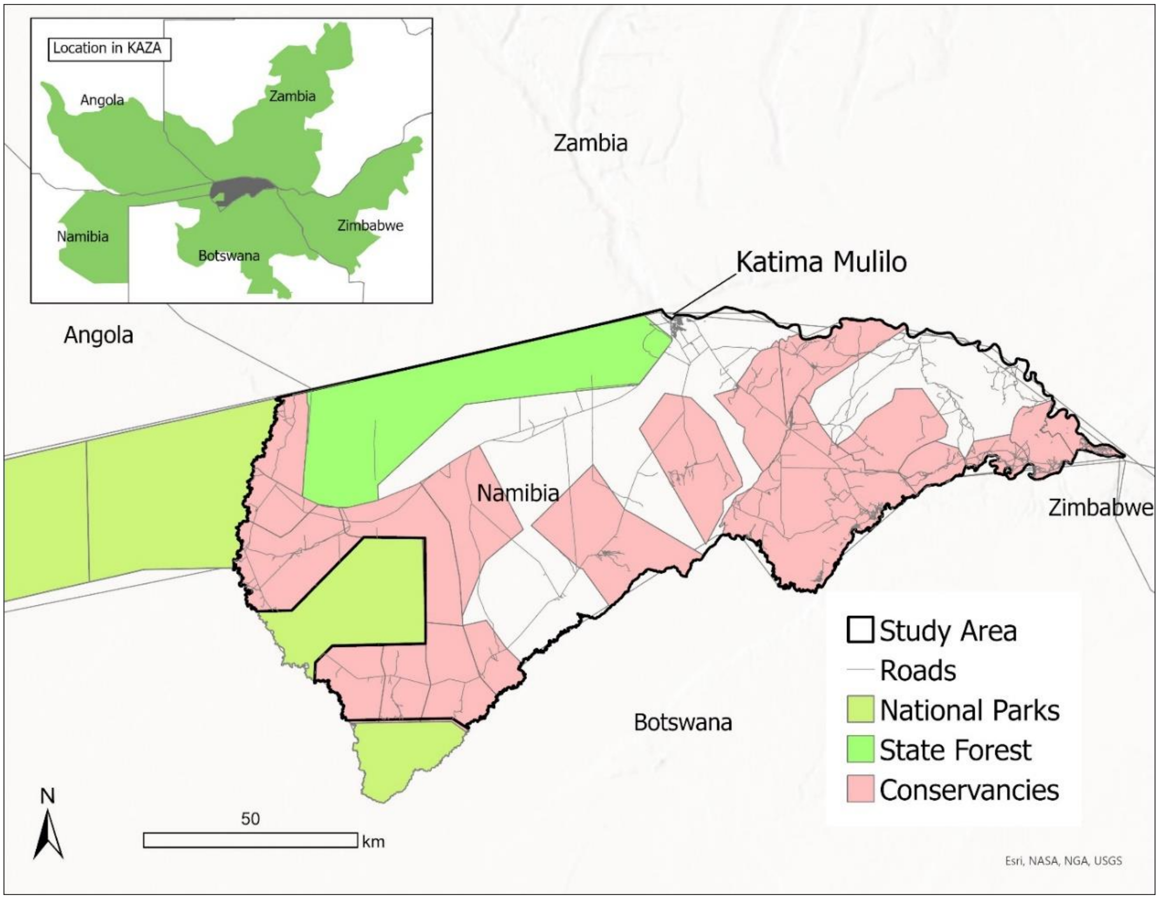


Figure 1. Map of the study region, the eastern Zambezi Region, excluding the national parks.

Object Based Image Analysis (OBIA)

Since the early 2000s, Object-based Image Analysis (OBIA), which is also referred to in the literature as Geographic Object-Based Image Analysis (GEOBIA) has grown in popularity with geospatial analysts for a variety of applications in land cover classifications and analyses (Blaschke, 2010; Blaschke et al., 2014; Castilla & Hay, 2008). For the purpose of this study, the term “OBIA” will be used for the remainder of the paper in accordance with the majority of the literature on the subject. OBIA departs from the traditional pixel-based approaches in remote sensing and merges concepts from both computer vision (i.e. image segmentation) with Geographic Information Science (GIS) in order to automate and improve land cover classification accuracies for remotely-sensed satellite imagery. The explicitly spatial nature of remote sensing imagery makes it a distinct and separate science from that of general computer vision, biomedical imaging, and bioinformatics (Castilla and Hay, 2008). OBIA is a technique that incorporates image segmentation and allows for the incorporation of ancillary data (e.g. thematic shapefiles, Digital Elevation Models, point clouds, spectral information, geometric information, and texture) to be included to enhance the segmentation and classification processes for the purposes of classifying distinct land cover types in a given remotely-sensed image (Blaschke 2010). Land surfaces are spatially heterogeneous, complex, and variable over time and OBIA approaches are designed to use knowledge-based systems to automate and improve the image classification process for geographic image datasets.

Object Based Image Analysis (OBIA) since the early 2000s has become more prominent and has even been proclaimed as a major “paradigm shift” away from more traditional pixel-based remote sensing techniques (Castilla and Hays 2008; Blaschke 2010; Blaschke et al. 2014). OBIA provides a framework that allows for the processing of data using a wide variety of spatial, spectral, and textural measures to enhance the separation of images into image objects, and has led to a revolution in computer vision in various fields, from the very small (e.g. biomedical) to larger geographic scales. OBIA allows for the incorporation of large image datasets, in addition to Digital Elevation Models, Point clouds, as well as vector data (Blaschke 2014). Objects may or may not fit to "real world objects" as objects are often made of various materials, and segmentation is relative to the scale that is set within the parameters of the segmentation algorithm. At times, "oversegmentation" may be a necessity, particularly in terms of the scale of the analysis being performed and the land cover features of interest, particularly in regional-scale analyses.

Random Forest

Random Forest (RF) is a statistical machine learning approach developed by Leo Breiman (2001), and has become more popular with remote sensing scientists for utilizing big data in connection with increased computing compacity in order to create highly accurate classifications (Pal, 2005; Rodriguez-Galiano et al., 2012). RF is an ensemble decision tree algorithm that allows for the classification of non-parametric and non-linear data and large datasets. Single Decision Trees are often prone to overfitting of model to the input data and can be highly susceptible to outliers and noise. Ensemble methods of decision tree allow for weighting or voting for classes based on a number of decision trees.

RF combines bagging (also referred to as bootstrap aggregation) (Breiman 1994) and random subspace methods (Ho 1995) for classification and regression of datasets. Bagging is a process where each tree is grown from random selections (without replacement) from the training data. Tin Kam Ho's Random Decision Trees (1995) introduces a novel method of randomly sampling features (with replacement) instead of the entire set to determine where to split the nodes of the decision tree. Random Forest combines bagging and random subspace methods to improve classifier accuracy, minimize noise, and reduce correlation between decision trees (Pal 2005, Breiman 2001, Liaw & Wiener 2002).

The RF approach allows users to grow many decision trees, with each tree voting for class membership. The training data is split into two groups, with a portion of data used to grow the decision trees, and the remainder is set aside for out-of-bag (OOB) error estimates. The large number of decision trees helps to avoid overfitting of data, which is highly problematic in many single decision tree methods and is robust to outliers and noise, unlike ensemble learning methodologies. RF has rapidly become popular for statistical approaches to image classification and have been widely used in a variety of applications for statistical analyses from biomedical imaging to remote sensing classifications (Diaz-Uriarte & De Andres 2006, Pal 2005).

Hybrid Object-based Random Forest Approaches

In a comparison of pixel-based and object-based approaches, Duro, Franklin & Dube (2012) found that classification accuracy may not actually be meaningfully different. They argue that user's preferences towards generalizability and contiguous land cover areas may perhaps be a greater factor in the preference towards object-based and pixel-based approaches. Stumpf and Kerle (2011) argue that variability within classes may be a challenge in lower-resolution imagery and this is where object-based techniques can help improve accuracy and interpretability of classification results. Many studies have used a hybrid OBIA-RF style approach for various geospatial applications from landslide detection to savanna vegetation morphologies (Stumpf & Kerle, 2011; Mishra & Crews 2013; Puissant, Rougier & Stumpf, 2014).

Hybrid approaches have been used in a number of studies (Belgiu & Dragut 2016). RF has been shown to be biased in cases where there is a major class imbalance issue (He &

Garcia 2009, Stumpf & Kerle 2011). Stumpf and Kerle (2011) used a hybrid RF and OBIA method to detect landslides in order to reduce the effects of this imbalance amongst classes to improve accuracy. Stefanski et al. (2013) found that segmentation overall improves the accuracy of classification results, and that random forest is robust even with smaller numbers of training samples. The hybrid approach offers the ability to reduce effects that may occur from the small number of training sites for the size of the study area, as well as provides the ability to address the extreme complexity of savanna land cover classification in this region. For these reasons, an investigation of the above methodologies is being included in this study in order to explore their applicability for addressing the heterogeneity of the landscape and the constraints imposed by data collection in the present study.

METHODS

Study Area

For the purposes of this study, we will be focusing on the eastern portion of the Zambezi Region in Namibia, as mentioned previously in Figure 1 above. The eastern Zambezi Region ($10{,}336.5\ km^2$) is bounded by four perennial rivers, the Zambezi River to the north/east, Chobe and Linyanti Rivers to the south, and the Kwando River to the west. We have excluded the National Parks core areas that lie within the study area from the analysis, due to the prohibition of settlements in the protected core conservation areas of these national parks. However, Bwabwata National Park has settlements inside the Multiple-Use Zone. Khwe (San) people have lived in Bwabwata since the 1800s, centered around the town of Chetto (Nkengbeza, Kangumu, & Sibuku 2021) and some newer settlers are coming from the Kavango East region, however constraints on data collection forced the exclusion of Bwabwata National Park entirely from the study. The Forest Reserve was retained in the study, as there are a number of settlements within and adjacent to the forest reserve, as well as a high profile clearing of timber for a tobacco farm within the State Forest which are of great interest to conservationists in the region (Ministry of Environment 2014).

The eastern Zambezi Region is a dryland savanna system with rich biodiversity that lies central to the Kavango-Zambezi Transfrontier Conservation Area (KAZA), the largest terrestrial conservation management zone in the world with an area of $440{,}000\ km^2$ covering parts of Angola, Botswana, Namibia, Zambia, and Zimbabwe (Peace Parks 2010). The eastern Zambezi Region is adjacent to several national parks, including Bwabwata National Park, Mudumu National Park and Nkasa Rupara National Park in Namibia, and Chobe National Park in Botswana. The soils of the eastern Zambezi Region vary in terms of mixtures of sand, clay, and loam.

The Zambezi Region experiences a unimodal wet and dry season climate, with higher average temperatures in the rainy season (Rainfall amounts decline north to south and east to west in the Zambezi Region (Mendelsohn & Roberts 1997). While most of Namibia receives less than 250 mm of rainfall per year, the Zambezi Region experiences highly variable rainfall inter-annually and intra-annually (Gaughan et al. 2015). The Zambezi Region sees anywhere from 300 mm to over 1000 mm per year in annual rainfall, while average rainfall totals range from 500 mm to 750 mm, and modal totals range from 400 mm to 550 mm per year depending on the gradient of rainfall patterns in the region (Mendelsohn & Roberts 1997). Yearly precipitation averages for the study area are displayed in Figure 2 below, from CHIRPS satellite derived precipitation datasets for the years 1981-2019 (Funk et al. 2015). This also shows a trend of increasing precipitation moving eastward in the Zambezi Region, but also some areas of greater precipitation to the south along the Kwando and Linyanti River systems.

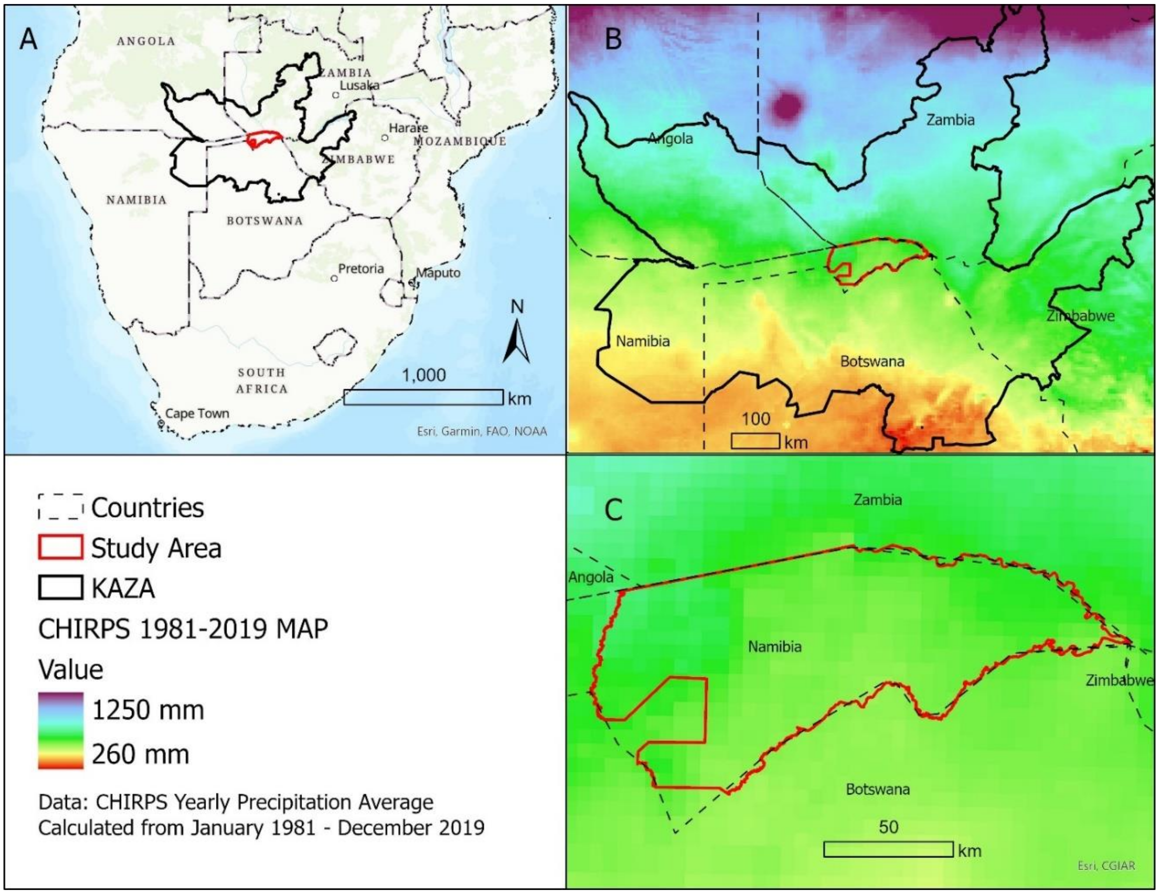


Figure 2. A) Location of study area in southern Africa, B) rainfall gradient throughout the KAZA region, and C) the rainfall gradient in the eastern Zambezi Region. Rainfall gradient was computed using data from the years 1981 through 2019, using CHIRPS satellite derived precipitation data (Funk et al. 2015).

Within the Zambezi Region land use designations vary in terms of management regimes and tenures. Protected areas and communal lands represent the main two land tenures; however, there are variations in the management and tenure designations within each of these categories. In protected areas, core areas of protected lands prohibit settlements, whereas multiple-use zones allow for settlements, croplands, and other infrastructure. In communal lands, land use, land tenure, and land transference are managed through the Traditional Authority system and moderated through the Ministry of Land Reform. In the Zambezi Region, the Traditional Authority is based on a chieftaincy system (Kamwi et al. 2017), with

Chiefs and area "*induna*" (local Traditional Authority leaders) as the main decision makers on land rights in the region.

Some communal lands have been given special conservation status in Namibia with the introduction of Community Based Natural Resource Management (CBNRM), beginning in the late 1990s and 2000s (Cumming 2008). Conservancies have been lawfully designated as the legal vehicle for communities to implement CBNRM in their local areas, and this legal designation allows communities to manage and derive benefit from the land through wildlife tourism, hunting, and harvesting of natural resources, such as *Harpagophytum zeyheri* (commonly known as Devil's Claw, a highly valuable plant used as a pain reliever in local traditional and western homeopathic medicine) and timber resources (Cumming 2008, IRLUP 2015).

Data Collection and Preparation

Global Land Cover Datasets

A comparison of the GHF, GLC30, GUF, and WSF datasets was created. Each of the datasets was mapped in ArcGIS Pro (Version 2.8.2 , ESRI 2021). The GHF dataset was displayed using a stretch using Standard Deviation of 2, with a range of 0 to 50, with 50 denoting higher level of pressure from the human footprint. GUF Dataset was displayed using unique values corresponding to "Non-Urban" and "Urban." GLC30 was displayed using unique values for "Cropland" and "Built". The WSF dataset was mapped using unique values corresponding to "Settlement", and no data represented as black.

Land Cover Classes

The land classes that will be considered for this are built settlement areas, fallow agricultural land, bare ground, grasses, trees, shrubs, and water. These vegetation classes, water, and bare soil lands are combined for the purposes of highlighting settlement and agricultural lands, for areas for an approximation of the human footprint within the Zambezi Region.

Training Samples

Training sample collection was designed to occur during the dry season, so that distinctions between fallow agricultural fields and surrounding vegetation would be most distinct, because the crop residues in the fallow fields would spectrally be easier to distinguish from photosynthetic vegetation in surrounding areas. Additionally, I selected the dry season imagery due to the availability of cloud-free data which is unavailable for the rainy season. Training sample locations were randomly generated within a 500-meter buffer zone to roads in ArcGIS (ESRI, version 10.3) and located using a Trimble Juno GPS. I visited 137 sites distributed throughout the Zambezi Region and collected in situ data points using a Samsung A5 cellphone, a Trimble JUNO, EpiCollect Plus for filling out training site forms (Appendix A), and Google Cloud Services to host pictures and information collected in the field. Data was uploaded daily to Google Cloud Service. At each training site I collected photos of the ground, the canopy and a 360-degree panoramic photo (Figure 3) to give context to the training sample forms. Training data were additionally supported by creating another 350 ex situ points in google earth using in-field experience and expert knowledge of the study area with the fine scale imagery available for 2017. Google Earth points were selected by randomly generating points, and then supplementing with additional points to balance the classes. The training sample points were classified as agriculture, built, or other. Table 2 shows the breakdown for the training samples and validation points.

Table 2

Breakdown of classes by land cover for training samples and validation points.

Land Cover Classes	Training Samples	Validation Points	Total
Agriculture	80	52	132
Built	77	46	123
Other	145	87	232
Total	302	185	487

Figure 3. Three panoramic photos showing a gradient of grass, trees, shrubs, bare ground, and crop residues from a harvested field from three different field sites.

Image Pre-Processing

For this study, I acquired four Landsat scene tiles, as described in Table 3 and shown in Figure 4. Each Landsat image was collected at the L1T level, meaning that it was preprocessed data that comes geo-rectified and radiometrically corrected, suitable for geographic analysis. Once downloaded, each scene was then atmospherically corrected using the Fast Line-of-Sight Atmospheric Analysis of Spectral Hypercubes (FLAASH) calibrated (Anderson et al., 2002) to surface reflectance values in ENVI (Harris Geospatial, vers. 5.3). The scenes were mosaicked together using ENVI's mosaicking tool and subset to the study region as previously described. Normalized Difference Vegetation Index (NDVI) and Tasseled Cap Transformations (TCT) were derived from the Landsat Imagery and were layer-stacked with the Landsat OLI multispectral bands. NDVI was calculated in ENVI. TCT layers were created in using the programming language R with the RSToolbox package (Leutner & Horning, 2017).

Table 3

Information of Layer-stacked Images used for the image classifications

Product	**Path/Row**	**Date Acquired**	**Bands**
Landsat OLI	174, 072	15 July 2017	Coastal (0.435 - 0.451µm)
LT1	174, 073	15 July 2017	Blue (0.452 - 0.512 µm)
	175, 072	06 July 2017	Green (0.533 - 0.590 µm)
	175, 073	06 July 2017	Red (0.636 - 0.673 µm)
			NIR (0.851 - 0.879 µm)
			SWIR 1 (1.566 - 1.651 µm)
			SWIR 2 (2.107 - 2.294 µm)
NDVI			NDVI
Tasseled Cap			Brightness
			Greenness
			Wetness

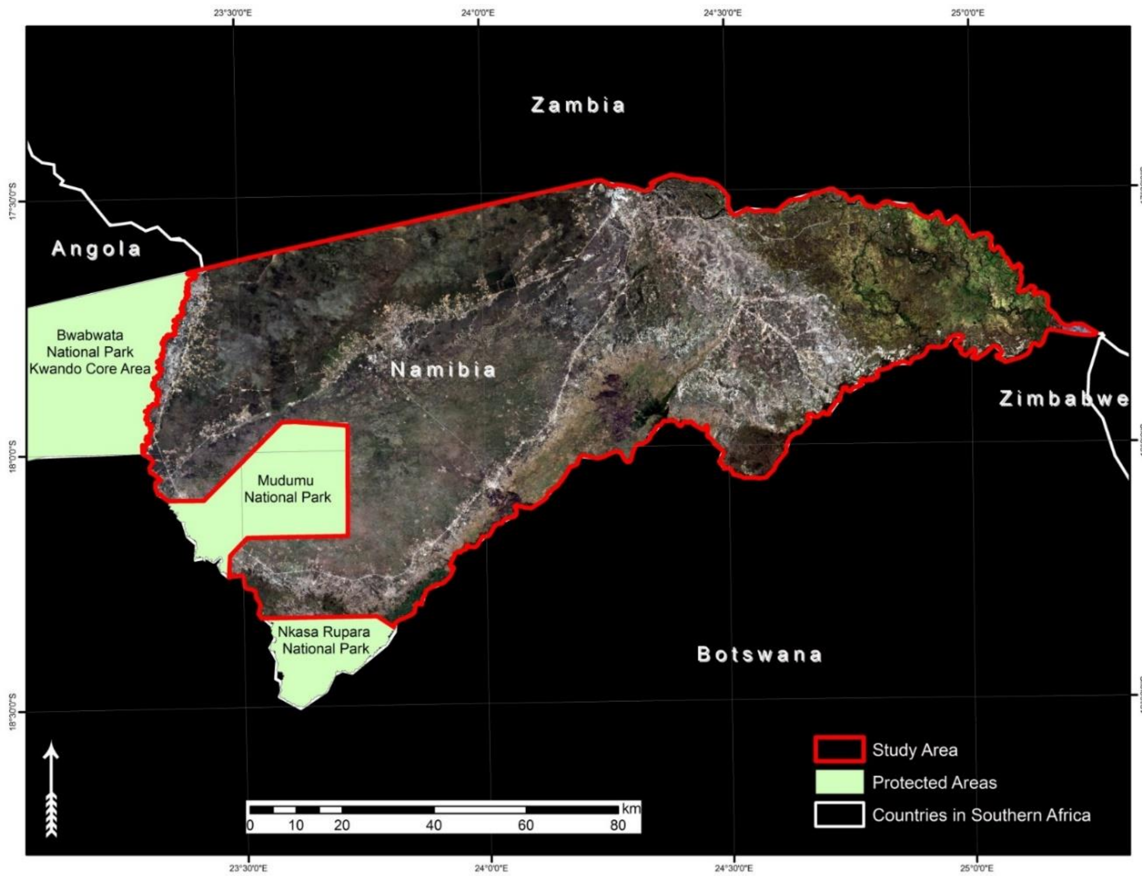


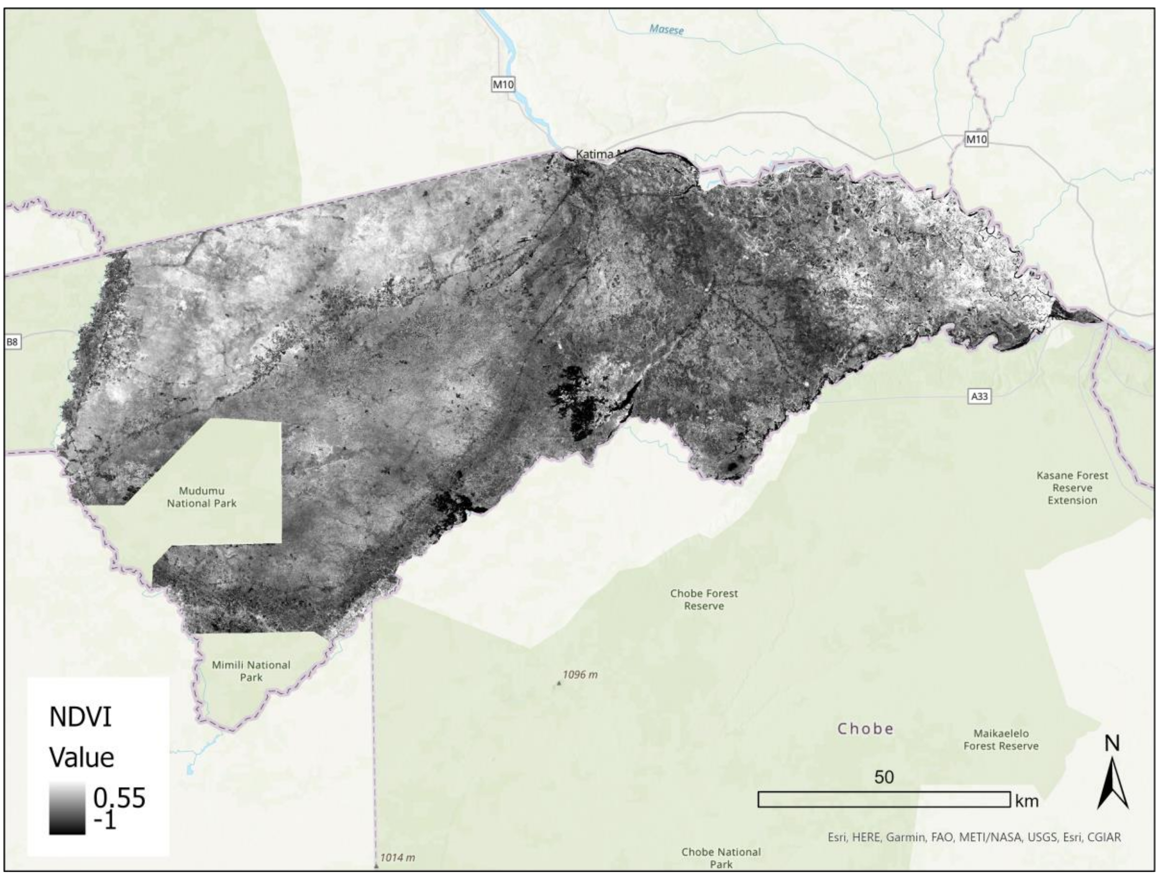
Figure 4. Landsat OLI Mosaicked Image.

Spectral indices and transformations of band information are widely used in remote sensing to aid in the visual and computational analysis of earth imagery to aid in the analysis of the physical properties of land cover and exploit this information for classification and analysis of landscape dynamics within the imagery (Jensen 2015).

NDVI

NDVI is used widely in remote sensing as a proxy measure of photosynthetically active vegetation for a wide variety of research applications (Myneni et al. 1995). NDVI was first used by Rouse et al. (1973) to look at vegetation dynamics in the Great Plains of the United States. NDVI is based on the physical properties of the role of light in photosynthetic plants, which absorb blue and red light, reflecting a higher percentage of green light (which is why healthy plants appear green to our eyes), and photosynthesizing vegetation reflects infrared wavelengths. The equation for NDVI (Eq. 1, Figure 5) exploits this physical basis to model vegetation health and photosynthetic activity occurring across the landscape.

$$NDVI = \frac{NIR - Red}{NIR + Red} \qquad \textit{Equation (1)}$$



<u>Figure 5.</u> NDVI calculated for the study region.

The range of NDVI values fall between -1 and 1, with values near zero and below indicating no green leaf photosynthetic vegetation. Similarly, values greater than 0.8 indicate highly saturated photosynthetic vegetation on the landscape. NDVI was included in the layer-stacked imagery used for this analysis to aid particularly in distinguishing settlement areas from the surrounding natural savanna vegetation.

Tasseled Cap Transformation

Kauth and Thomas (1976) first developed the Tasseled Cap Transformation (TCT) to leverage spectral information for analysis of croplands, but TCT has been used for revealing insight into forest disturbance (Healy et al., 2005), and distinctions between dryland land covers (Zanchetta, Bitelli & Karnieli, 2016; Ouedogo et al., 2016). Initially, the TCT was designed to work with Landsat MSS imagery, but has been adapted for later Landsat sensors as well. TCT is an orthogonal transformation that uses spectral information from multiple wavelengths ranging from the visible and infrared bands of Landsat to increase interpretability of remotely sensed imagery and reduce data. The most commonly used TCT metrics are soil brightness, greenness, and canopy and soil moisture (wetness). The rationale for employing TCT for this study is that it helps distinguish the patterns around settlement areas, where soil brightness is high, particularly in rural areas and agricultural areas, and the mostly bare areas present in the floodplains. To derive TCT from the Landsat OLI imagery, I used the RSToolbox package in R to produce brightness, greenness, and wetness (Appendix B) as depicted in Figure 6. The RSToolbox package uses the coefficients for developing the Tasseled Cap Transformation as developed by Baig et al. (2014).

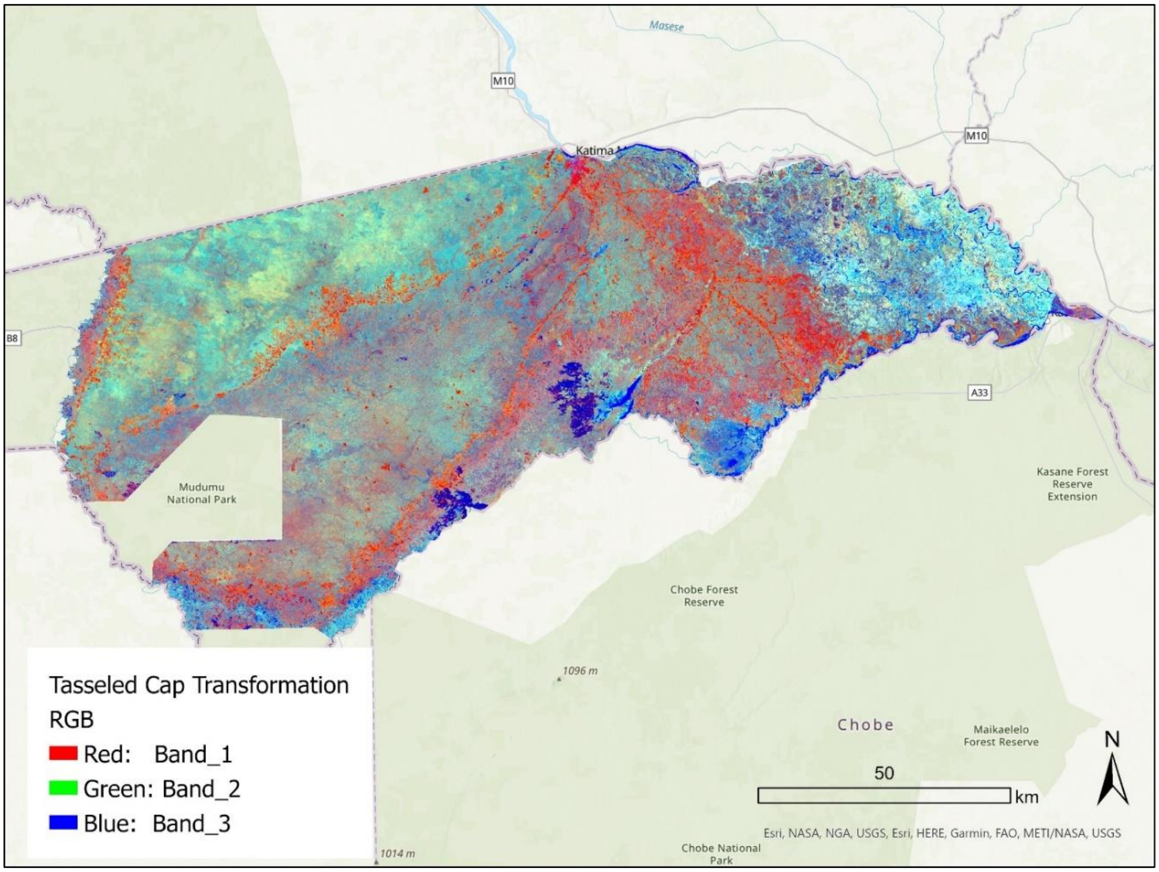


Figure 6. Visualization of the Tasseled Cap Transformation for the Landsat OLI 2017 mosaicked image subset to the study area, the bands that are visualized with red (TCT brightness), green (TCT greenness) and blue (TCT wetness). Settlement and agricultural areas are distinctly saturated in red/orange and distinct in terms of hue and texture from bare areas near the floodplains.

Segmentation

Segmentation is a process that divides images into groups of neighboring pixels using digital numbers, texture, and shape characteristics (Vieira et al., 2012; Meinel & Neubert, 2004). There are many kinds of segmentation algorithms for image segmentation (Meinel & Neubert, 2004; Baatz & Schape, 2010), as well as many different approaches to segmentation refinement (Drăguţ, Tiede & Levick, 2010; Benz et al. 2004; Tian & Chen, 2010; Ming, Wang & Zhang, 2015).

The most commonly used for satellite imagery is the multiresolution segmentation algorithm. The multi-resolution segmentation algorithm weights size, shape, scale and spectral information to create homogenous regions within an `image into image objects. Larger objects are also allowed for when they meet a minimum threshold for similarity between neighboring pixels or regions within an object. Baatz and Schape (2008) offer a review of various segmentation algorithms and offers an approach towards optimizing multi-resolution segmentation parameters. Under and over-segmentation can reduce the efficacy of the classification, and also renders the multi-resolution algorithm somewhat irrelevant. Thus, it is important to segment the image to meaningfully delineate between objects, while reducing over- and under-segmentation (Viera et al., 2010). Table 4 gives details to the segmentation parameters that were utilized for this study.

Table 4

Segmentation Algorithm Parameters for each of the approaches.

Approach	Segmentation Algorithm	Scale	Compactness	Shape	Ancillary Data
Object-based	Multiresolution Segmentation	60	0.3	0.5	2017 Updated OSM Roads Layer
Object-based Random Forest	Multiresolution Segmentation	60	0.3	0.5	2017 Updated OSM Roads Layer
Random Forest	Chessboard Segmentation	9	n/a	n/a	n/a

Chessboard segmentation was used for the Random Forest Model in eCognition, to apply the region grown training sample shapefile to a region of comparable size in the imagery. The scale parameter sets the number of pixels for each object; in this case a segmentation scale of 9 produces a 9x9 (270 meter) pixel object. This was to expand the selection of training data for the Random Forest approaches. In applying the training samples to the objects, training samples are only applied to the object if a majority of the shapefile overlaid the object, or in the case of minimal overlap a majority of the pixels in the object were similar to the pixels overlapping the shapefile. The 270-meter extent was chosen because there is likely to be similar land cover in the adjoining pixels to the sample points.

OBIA Variables

Texture

The eCognition software can compute a variety of texture layer and texture is commonly included to aid in classification for object-based approaches (Castilla and Hay, 2008). For this study, texture was computed based on Haralick's (1973) Gray Level Co-Occurrence Matrix (GLCM) statistics. The statistics give information about gray-level pixel values and the relationships between pixels in a region. To reduce the computing time, texture layers were chosen based on information about correlations between texture statistics found in previous literature (Clausi, 2002; Laliberté & Rango, 2009). GLCM Entropy (Eq. 2) , GCLM Contrast (Eq. 3), GLCM Angular Second Moment (ASM, Eq. 4), were used for this analysis, and are not directionally biased as we used the average of the statistics in all four directions. Laliberté & Rango (2009) found that entropy was the least correlated with other texture layers, so it was included. In previous studies, dissimilarity has been found to be highly correlated to contrast (Clausi, 2002; Laliberté & Rango, 2009), so dissimilarity was excluded from this analysis. Entropy measures the amount of disorder within an area. Contrast and ASM measure smoothness of texture. Laliberté (2009) found that entropy has the highest variable importance and in analyses where limited texture features can be run and gives the greatest amount of information for classification purposes.

eCognition calculates the Gray Level Co-Occurrence Matrix (GLCM) and subsequent measures of image texture. The GLCM is a matrix of the number of rows and occurring in a given image. The equations for GLCM texture layers used in this study are as follows:

$$Entropy = \sum_{i,j=0}^{n-1} P_{i,j} \, (-lnP_{i,j}) \qquad \textit{Equation (2)}$$

$$Contrast = \sum_{i,j=0}^{n-1} P_{i,j} (i - j)^2 \qquad \textit{Equation (3)}$$

$$ASM = \sum_{i,j=0}^{n-1} P_{i,j} \qquad \textit{Equation (4)}$$

Where, $P_{i,j}$ is a matrix element of pixels within the image that are being evaluated. The texture values were added to the RF-OBIA classifier and used in differentiating land cover types in the rule-based, OBIA approach as well.

Spectral Variables

Minimum, maximum and mean spectral values were calculated for each of the imagery, including the OLI multispectral layers, NDVI, Brightness, Greenness, and Wetness.

Edge Contrast of Neighbor Pixels. Additionally, we calculated several additional spectral indices, including the Normalized Difference Infrared Index (NDII, Eq. 5), Normalized Difference Built Index (NDBI, Eq. 8), Normalized Difference Water Index (NDWI, Eq. 6), Modified Normalized Difference Water Index (MNDWI, Eq. 7). NDII is used for a variety of classification purposes from moisture content to aiding in classification of burned areas across a landscape, and this has led to a proliferation in the names for this band combination. NDII is also referred to variously across scientific literature as the normalized difference water index (distinct from the two that are also used in this study), as well as the Normalized Burn Ratio (Ji, Wylie, and Rover 2011). For the purpose of this study, NDII will be used to distinguish it from the other normalized water indices. NDWI and MNDWI are included for distinguishing moisture content across the various land cover types (Gao et al. 1996). NDBI are included to help aid in the detection of settlement areas (Zha, Gao and Ni 2003). The equations for the additional spectral indices (Table 5) used for the object-based analyses are listed as follows:

$$NDII = \frac{NIR-SWIR2}{NIR+SWIR2} \quad \textit{Equation (5)}$$

$$NDWI = \frac{NIR-SWIR1}{NIR+SWIR1} \quad \textit{Equation (6)}$$

$$MNDWI = \frac{Green-NIR}{Green+NIR} \quad \textit{Equation (7)}$$

$$NDBI = \frac{SWIR1-NIR}{SWIR1+NIR} \quad \textit{Equation (8)}$$

Table 5

List of Spectral Indices used in the OBIA rulesets

Spectral Index	Equations	Citations
NDII	(NIR-SWIR2)/(NIR+SWIR2)	Sriwongatanon et al. 2015
NDWI	(NIR-SWIR1)/(NIR+SWIR1)	Gao 1996
MNDWI	(Green-NIR)/(Green+NIR)	Xu 2006
NDBI	(SWIR1-NIR)/(SWIR1+NIR)	Zha, Gao & Ni 2003
NDVI	(NIR-Red)/(NIR+Red)	Rouse et al. 1973

Geometric Variables

A variety of shape/size metrics were selected to inform the analysis about the nature of the image objects (Table 6), Area, Length, Border length, Number of pixels, Length/Width, Compactness, Density, Elliptic fit, Rectangular fit, Roundness, Shape index, Number of segments, Maximum branch length, Degree of skeleton branching, average area represented by segments were all selected for attributes that would distinguish different objects appropriately, particularly in the case of overlapping spectral characteristics.

Table 6

List of Variables by type used for the OBIA approach.

Types of Variables	Variables
Shape	Compactness, Density, Elliptic fit, rectangular fit, Roundness, Shape index, Number of segments, Maximum branch length, Degree of skeleton branching, Average area represented by segments
Size	Area, Length, Border length, Number of pixels, Length/Width
Texture	GLCM Entropy, GLCM Contrast, GLCM Angular Second Moment
Spectral	Mean, Minimum, Maximum metrics for Multispectral image layers, NDVI, TCT, MNDWI, NDWI, NDBI, NDII

Random Forest

The Random Forest algorithm builds an ensemble of decision trees based on the input variables (Input Image data) and random subsets of the training data. The algorithm combines bagging and bootstrap sampling of the data to form ensemble decision trees to reduce overfitting of model to the sample data, as well as reduce susceptibility to noise and outliers (Breiman 2001, Pal 2005, Liaw & Wiener 2002). To construct a Random Forest Model, sample data is split into training data and verification data (or as Breiman refers to it "out-of-bag" (OOB) error estimation). Usually about 1/3 of the dataset is set aside for OOB

Error Estimates. Decision trees are created by randomly sampling from the training data, at each node splits randomly. Trees are not pruned.

Region Growing is a method that allows for the expansion of training sample areas in order to increase the size and therefore the number of pixels to be included in the training data, this was particularly important for the Random Forest model, which relies on a large dataset in order to train the classifier model with an ample amount of variation within and between classes to improve classification outcomes. The training sample points were region-grown using a 270-meter buffer in ArcGIS 10.3. These were then applied to image objects in eCognition (version 5.2.0, Trimble) after the segmentation process in the random forest and hybrid object-based random forest approaches.

eCognition versions 9.2 and on can train and classify random forest models, the algorithm is called “Random Trees”. While the technology is cloaked in a black box, it was designed based on Leo Breiman’s Random Forest and provides a Graphic User Interface (GUI) interface to perform a Random Forest analysis, however the technology is nascent and lacks the ability for one to assess variable importance in the model and perform an accuracy assessment, as you can when you are programming a random forest model in R. This study wanted to explore the functionality of the Random Trees algorithm in eCognition to compare them to the other approaches in eCognition.

To train the Random Forest Classifier in eCognition, we used a maximum tree number of 500, and terminated decision tree creation once the maximum number was reached or by terminating with a forest accuracy of 0.01 (the default setting in eCognition). Once the RF classifier was trained, the RF classifier was applied to the input imagery data.

Hybrid Approach: Object-Based Random Forest

To apply the Object based approach to a Random Forest Model, we used eCognition's Random Forest algorithm "Random Trees" to segment the image and then train and apply the random forest model. The same variables used for the OBIA variables as for the OBIA-RF approach, as detailed above. Once segmentation is complete, you begin by applying training samples to image objects. We then proceeded by creating the RF classifier. The maximum number of trees and the termination of the classifier were same as RF classifier detailed above. The eCognition community on the Trimble website has a plethora of tutorials and sample rulesets to learn from, and is free to join, while the software remains quite expensive to acquire.

Accuracy Assessment

In order to assess the correctness of the image classifications, a confusion matrix to derive several measures of accuracy using training sample data that had previously been set aside for verification purposes. These accuracy metrics include the overall accuracy, indicating the percent of cases correctly allocated (Foody 2002), the user's accuracy which is a measure of the probability that a pixel classified in the image correctly reflects the land cover it represents on the ground (Story and Congalton 1986), and the producer's accuracy which measures the likelihood of commission errors, or pixels that were incorrectly assigned to a class (Story and Congalton 1986).

The Confusion matrices were computed using R Caret Package (Kuhn et al. 2008) in RStudio using the validation data points generated from the field and through google earth as previously discussed (n=362). The confusion matrix report gives several measures of accuracy based on. The no information rate is the percentage of the largest class and is used as an indicator of whether the overall accuracy is greater than the rate of the class with the greatest number. Because the largest class by far was the combined vegetation, bare, water class, it will be important to compare the overall rate to this class when we are considering the results. Overall Accuracy is the total number of correctly classified points divided by the total number of validation points. User's Accuracy (eq. 9) refers to the vertical agreement between correctly classified and the total number of classified in the reference columns, also referred to as User's Accuracy. Table 7 provides the reference for the confusion matrix and how it is used in the following accuracy metrics. Producer's Accuracy, as detailed in eq. 10, refers to the horizontal agreement between the reference and the predicted classifications, or producer's accuracy. Balanced Agreement (BA, Eq. 16) is an average of the accuracies for each of the classes in the analysis. In addition, the Kappa Statistic (Eq. 17) was calculated which is a common form of accuracy assessment that relies on the entire confusion matrix to determine the percent of correctly classified pixels for a given image (Fitzgerald and Lees 1994). The Kappa Statistic employs a multivariate technique in assessing the accuracy of a classification (Congalton 1991).   For a given matrix as depicted in Table 5, the following equations are the foundation of the "confusionMatrix" command in R to generate the accuracy assessment report.

Table 7

Sample Confusion Matrix adapted from Kuhn et al. (2008)

	Reference	
Predicted	Type 1	Type 2
Type 1	A	b
Type 2	C	d

$$UA = \frac{a}{a+c}$$ Equation (9)

$$PA = \frac{d}{b+d}$$ Equation (10)

$$BA = \frac{UA+PA}{2}$$ Equation (11)

$$K = \frac{N\sum_{i=1}^{r} x_{ii} - \sum_{i=1}^{r}(x_{i+} * x+i)}{N^2 \sum_{i=1}^{r}(x_{i+} * x+i)}$$ Equation (12)

RESULTS

For comparison, Figure 7 shows the extent of GHF, GHSL and WSF datasets over Google Earth imagery for two areas in Zambezi Region of Namibia, a mostly rural area in southern Africa. The inconsistency in the detection of rural settlement areas in the GUF and GHSL make these datasets incomplete for the purposes of understanding local human-environment interactions at the local level. In particular, a more detailed and current map could be used to identify zones of land use that may be directly competing with conservation efforts or for identifying areas that should be protected such as critical habitat areas for threatened and endangered species.

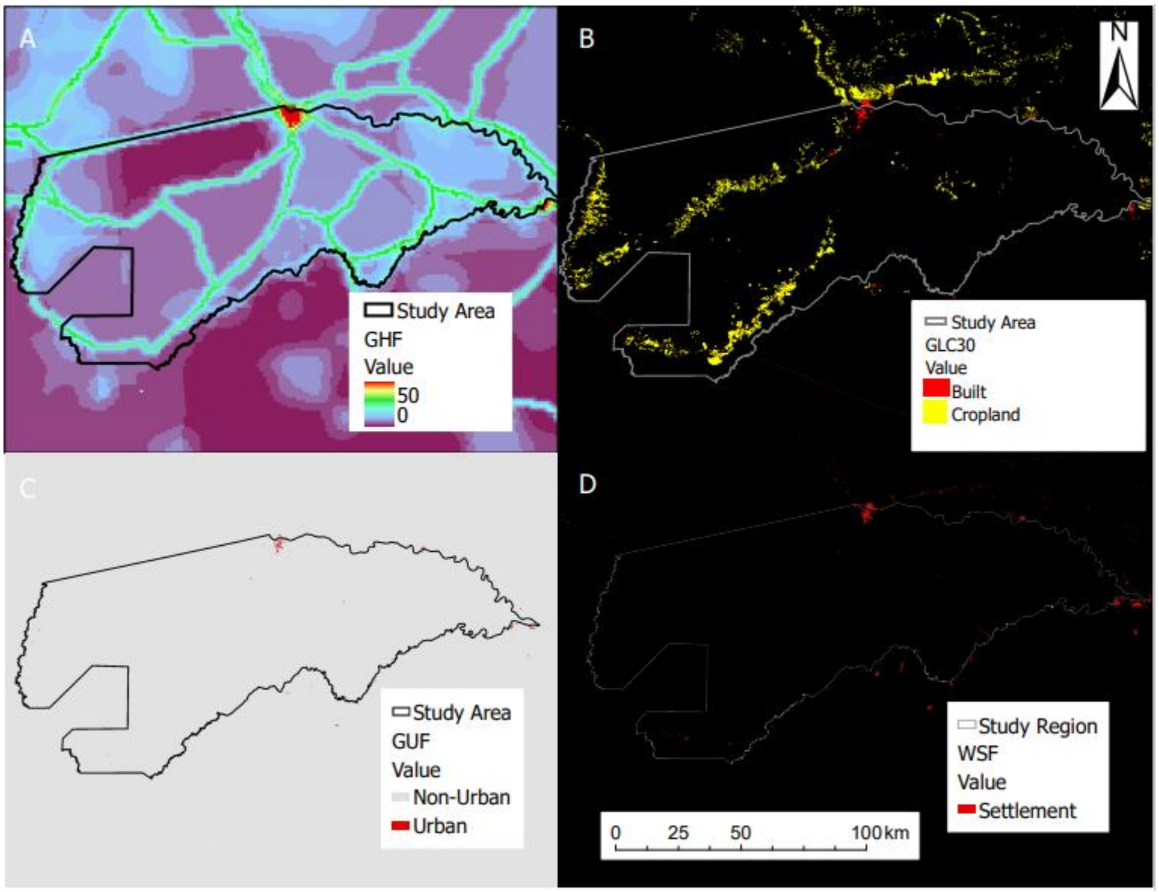


Figure 7. Image depicting global datasets A) GHF human pressure (2009), B) GLC30 cropland and built layers (2010), C) GUF urban and non-urban (2011-2012), and D) WSF settlement areas (2014-2015) for Zambezi Region

Segmentation

Figure 8 depicts a) multi-resolution segmentation for the study region and b) chessboard segmentation used for the random trees (based off of the Random Forest algorithm) in eCognition. The multiresolution segmentation parameters were scale of 60, compactness of 0.3, and 0.5 for shape. The resulting chessboard segmentation resulted in a 9X9 pixel extent to region grow for the Random Forest to increase the amount of pixels being used in the training data, due to the low number of training data sites available.

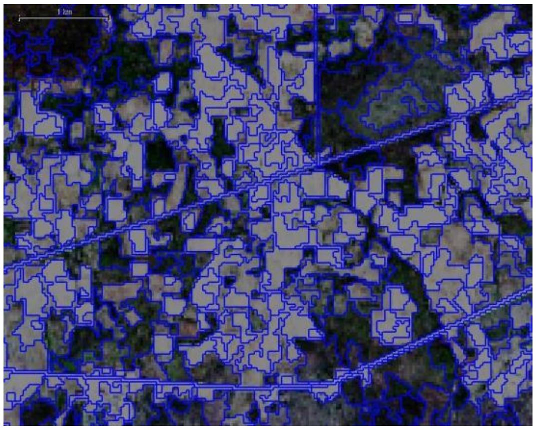
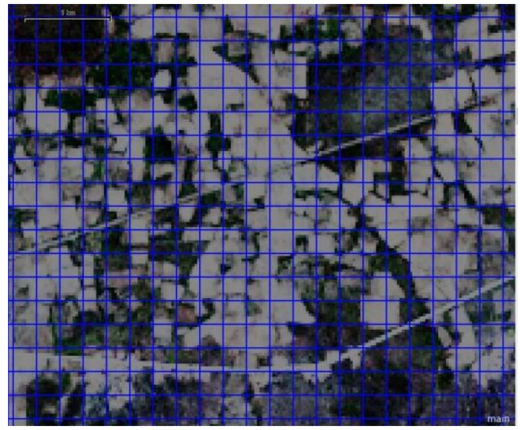
A  B 

Figure 8. Multi-resolution segmentation (A) and Chessboard Segmentation (B) boundaries overlying a true color composite image created in eCognition 9.2.0.

Land Cover Results

Total land cover areas for each class in each classification are represented in Table 8. Agriculture accounts for 10.74%, 17%, and 11.92% of total land area of the Zambezi East for the OBIA, RF-OBIA, and RF classifications, respectively. Built up areas account for 11.89%, 4.08%, and 9.68%, respectively. Combined background classes account for ~78% in each of the classifications.

Table 8

Total Land Cover Areas by Class for OBIA, RF-OBIA, and RF classifications.

Classification	Land Cover Class	Square Km	Pct Area
OBIA	Agriculture	1110.00	10.74%
	Built	1228.94	11.89%
	Vegetation	7997.54	77.37%
Hybrid	Agriculture	1757.20	17.00%
	Built	421.29	4.08%
	Vegetation	8158.01	78.92%
RF	Agriculture	1231.77	11.92%
	Built	1000.87	9.68%
	Vegetation	8103.85	78.40%
Total Area		10336.5	100%

The classification results are depicted in Figure 9 and Table 9 displays the information relating to the accuracy assessment results. The overall accuracy for each of the approaches were as follows: 79% (OBIA), 76% (RFOBIA), and 70% (RF). The Kappa Statistics for each of the approaches were 0.67 (OBIA), 0.60 (RFOBIA), and 0.52 (RF).

Class-wise, agriculture was the most difficult to distinguish and held the most errors of User's Accuracy (UA) and Producer's Accuracy (PA) in all three methods, in the OBIA results we find that the UA is 52%; whereas, in the RFOBIA UA is 64% and RF UA is 57%. In terms of PA for the agricultural class, the results include 94.16% for OBIA, 87.23% for RF-OBIA, and 86.13% for RF.

In terms of the built class, OBIA attained 94.67% (UA) and 86.36% (PA); whereas RF-OBIA scored 71% (UA) and 93% (PA), and RF scored 63.16% (UA) and 86% (PA). Balanced Accuracy for the built classes favors the OBIA classification with a score of 90%.

Balanced Accuracy for RF-OBIA and RF are respectively Producer's Accuracy, 82% and 74%.

OBIA performs better in terms of UA and PA. However, RF-OBIA scores higher Balanced Accuracy. Overall, the RF approach scores consistently lower across the board for each class. Whereas RF-OBIA outperforms the OBIA classification in PA.

As for the combined vegetation/other class, the UA were respectively 86% (OBIA), 83% (RF-OBIA), 79% (RF). PA for the vegetation/other class were 89% (OBIA), 81 % (RF-OBIA), and 82% (RF).

Table 9

Accuracy Assessment results for OBIA, RF-OBIA, and RF classifications.

	Land Cover	UA%	PA%	BA%	OA%	K
OBIA	Agriculture	52.3	94.2	73.2	79.8	0.67
	Built	94.7	86.4	90.6		
	Other	86.4	89	87.7		
Hybrid	Agriculture	63.6	87.2	75.4	76	0.60
	Built	71	93	82		
	Other	83.3	80.5	82		
RF	Agriculture	56.8	86.1	71.5	70.4	0.51
	Built	63.2	86	74.6		
	Other	79.2	82.3	80.8		

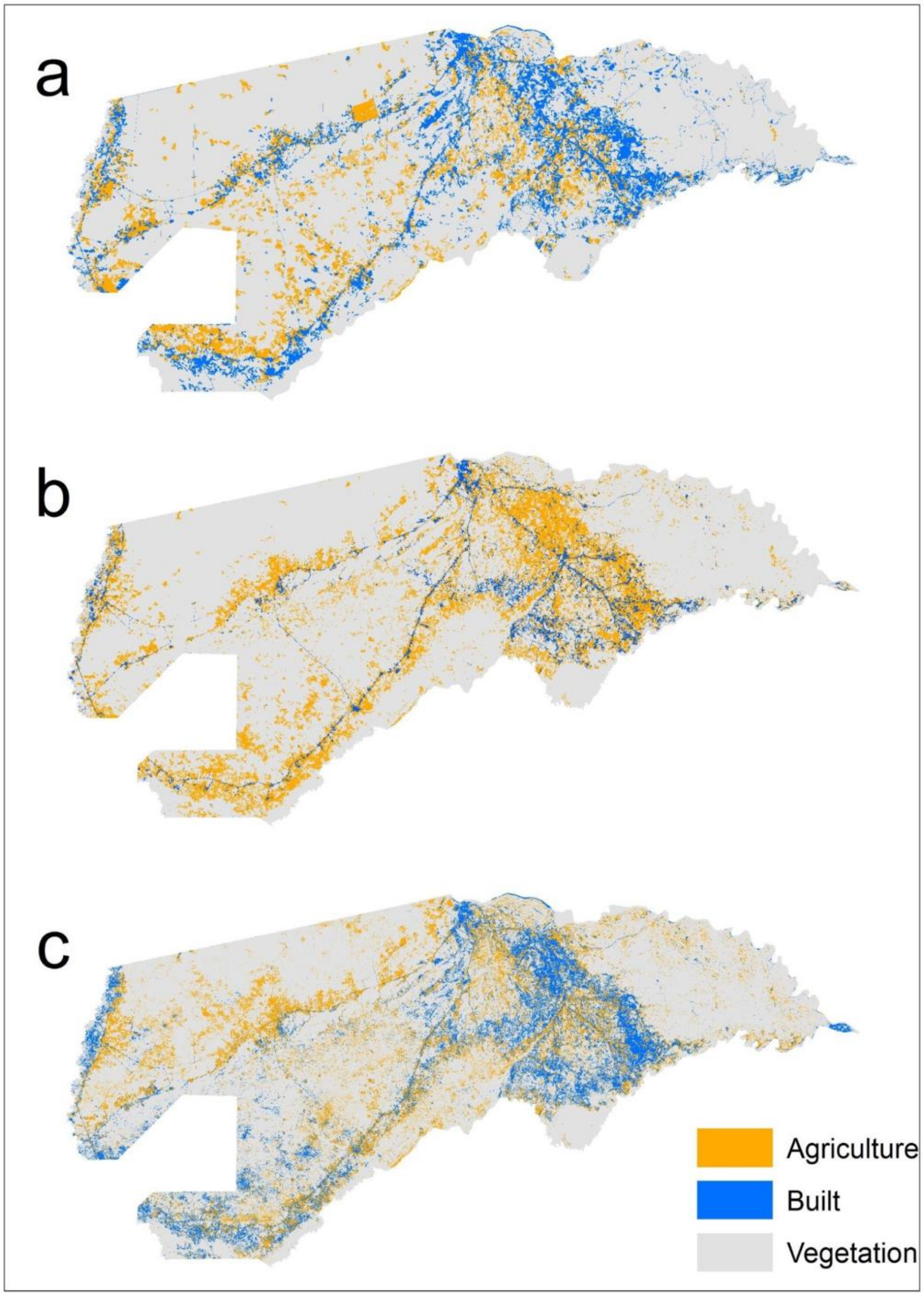


Figure 9. Results from a) OBIA b) the hybrid Object-Based/Random Forest approach, and c) the Random Forest approach.

In comparing OBIA, RF-OBIA, and RF for both Agricultural and Built classes (Figure 10), we can see several patterns. Built areas are more conservatively allocated in the RF-OBIA classification. However, this is balanced by an over-classification of agriculture. For the Random Forest classifications, both agriculture and built land covers are overestimated. In the OBIA classification, the built areas are over-represented, and agriculture is misallocated across both built and the combined vegetation areas.

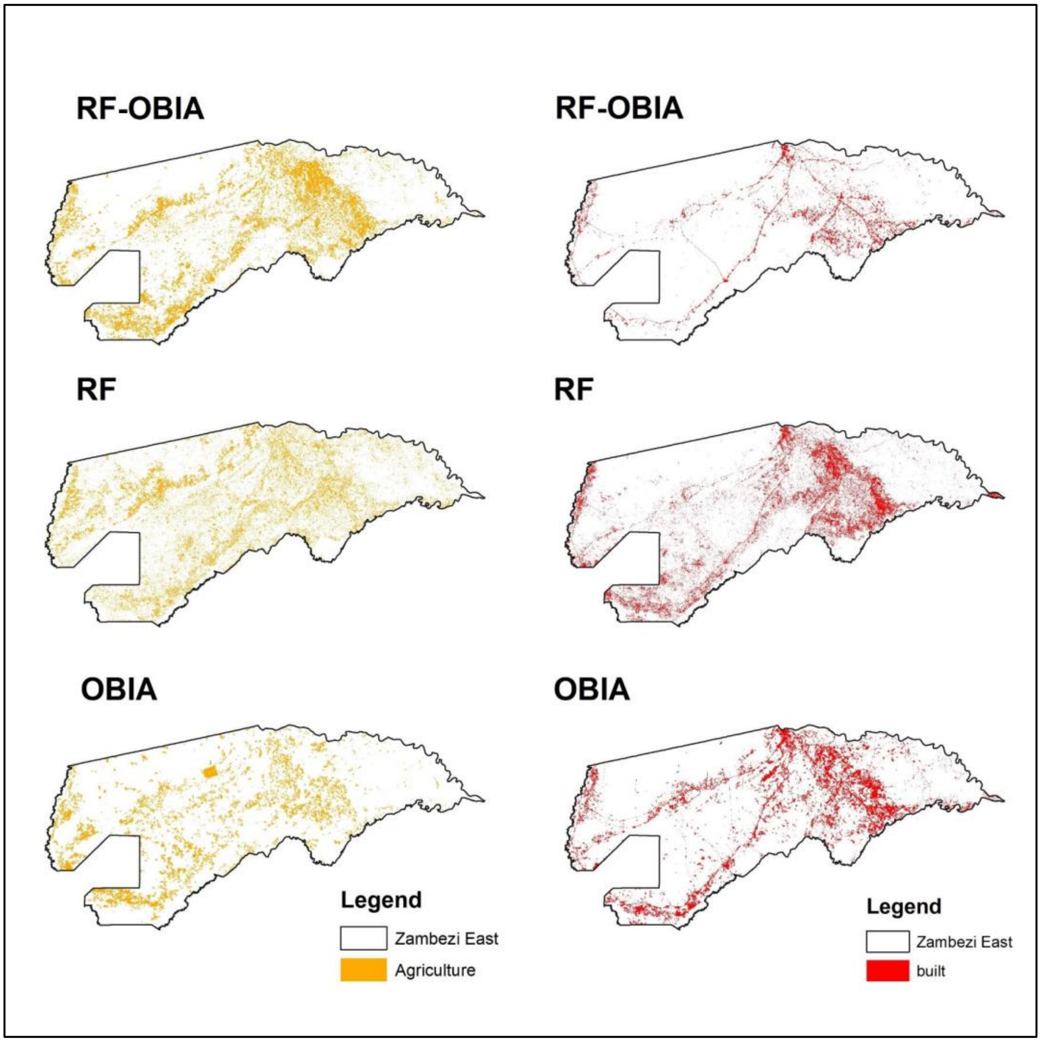


Figure 10. Comparison of Agriculture and Built classes across OBIA, RF-OBIA, and RF classifications for the study region.

DISCUSSION

This study evaluated the performance of three remote sensing techniques (OBIA, RF,and a hybrid approach) for classifying highly heterogenous savanna land covers for delineating the human footprint in Zambezi East in Namibia. While the OBIA classification had the highest agreement based on my accuracy assessments, there is a high amount of misclassification with agriculture, between both “built” and “other” classes in all three classifications which could propagate errors through subsequent analyses. Additionally, the same can often be said for the vegetation and bare areas, where the metrics used may still leave a lot to be desired in terms of separating bare ground and sparse vegetation from agricultural areas, particularly as the imagery includes fallow, harvested crop fields that carry similar characteristics and spectral information. The OBIA method proved to be time consuming, due to the iterative nature of setting up the parameters. Further optimization of the parameters may yield better results. Notably, I would include more ancillary datasets and spectral indices to improve the capture of the intended target classes. The addition of radar and multispectral wet season data could be an avenue for improving the detection of settlements and agriculture. The global datasets are insufficient for estimating the human footprint in the Zambezi Region. GLC30 visually looks the closest to the classification from this study, particularly with the agriculture layer, but the artificial surfaces layer does not adequately represent the totality of all settlements in the region. The GLC30 dataset is processed with a mixed pixel-based and object-based hierarchical classification approach. The other global datasets lack the detail necessary for estimating the human footprint. GUF, GHSL, and WSF all focus primarily on built up areas, which are not capturing the rural structures of the Zambezi Region. The GHF dataset estimates the human footprint in the Zambezi Region, but it relies heavily on land cover types and road networks to estimate the human footprint, and the detail is not sufficient for on the ground land use planning and conservation efforts.

Other considerations for improvement include additional band indices to the layer-stacked imagery may improve the classifications, such as Normalized Difference Middle Infrared Index (NDMIR) using bands 6 and 7 of Landsat OLI. NDMIR leverages the difference between the mid-infrared bands, which is helpful for highlighting built structures (Marconcini et al. 2020). The settlements in the Zambezi Region reflect highly in the Mid-Infrared range, so it may be beneficial to include this particular spectral index.

Additionally, the relative size of the number of training data samples should be large (Belgiu & Dragut 2016), and due to the constraints, the number of training samples in this study may have not been sufficiently large enough for the RF classifier. Classes that occupy larger land areas need more samples than those with relatively smaller land areas (Belgiu & Dragut 2016). The size of settlement and agriculture is relatively smaller in land area than the classes that represent “other” in this study. Increasing the number of samples in classes that were categorized as “other” may substantially improve the results of the RF classifiers. Moreover, the Random Trees algorithm in eCognition does not produce estimates for feature importance. It might prove beneficial to run the model through RandomForest in R, in order to assess the importance of the variables that are used for the model. This could help reduce the dimensionality of the dataset, by refining the data that is incorporated into the model in order to improve the results of the RF classifier models. The accuracy values of the RF models are higher when the classification using an object-based approach, rather than a pixel-based approach, as is consistent with the findings of Chetan, Dornick, and Urdea (2017).

Duro, Franklin and Dube (2012) argued that pixel-based and object-based approaches may not be meaningfully different in terms of classification accuracy and that user's preference towards generalizability and contiguous land cover areas may be why many choose object-based approaches over pixel-based approaches. I found that object-based approaches were 5-10% higher in overall accuracy score. The preference may in fact be appropriate in savanna landscapes, where there is a continuous gradient of trees, grasses, and shrubs interspersed across different LULC and generalizability is necessary.

Resource collection might also be affecting the accuracy of the three models. This region is highly dependent on natural resources such as foods, medicine, timber poles, reeds, and grasses for construction and maintenance of housing structures (Kamwi et al. 2015; Bailey et al. 2020; Woodward et al. 2021). Figure X shows a collection of timber poles cut and Figure X shows thatching grasses for sale on the side of the road in the Zambezi Region. Areas where collection of these natural resources takes place may be confused in the classifications with crop fields or settlement areas. Incorporating information about resource collection areas would be an important factor for future attempts at modeling the human footprint for this region, particularly as resource extraction is such a widespread practice and a major source of supplemental income for households in this region (Bailey et al. 2015, Salerno et al. 2021; Woodward et al. 2021).

Figure 11. Timber poles from Impalila Island, Zambezi Region of Namibia.

Figure 12. Thatching for sale on the side of the road in the Zambezi Region of Namibia

Additionally, the misclassification of agriculture particularly in areas that are vegetated could be the result of high cattle densities confusing the classifications. Intensification of grazing and clearing in and around conservancies are putting pressure on lands within the Zambezi Region (Colpaert. Matengu & Polojarvi 2013). While the influences of cattle grazing and cattle populations were not taken into consideration for this analysis, more effort should be made to analyze the spatial analysis of cattle grazing across the Zambezi East region. According to the Integrated Rural Land Use Plan (2015), there are approximately 136,221 head of cattle and approximately 10,000 goats. Overstocking and overgrazing cattle may be linked to bush encroachment and land degradation and has led to shifts in the bush-to-grass ratio of Namibia's arid savannas (Routhage 2007; Nagendra et al. 2004). Approximately 69% of the population of the Zambezi Region lives in rural areas, where they raise cattle and practice subsistence farming (IRLUP 2015). Grazing is unrestricted across the communal lands, and rights are moderated through the traditional authorities of the area. According to Mendelsohn (2006), 43% of households own 30 cattle or less, and 15% of the population, owns more than 30 cattle. Incorporating information about cattle densities into these models may improve our understanding of human footprint in the eastern Zambezi Region.

CONCLUSIONS

The spatial extent of varying LULC is important at regional scales for balancing conservation objectives and regional development and human needs derived from the landscape (DeFries et al. 2005, Hurskainen et al. 2019). I employed object-based and random forest methods, as well as a hybrid approach to quantifying the human footprint for the savanna of the eastern Zambezi Region of Namibia, using the land cover classes of built, cropland, and other (containing vegetation, bare ground). While the OBIA classification performed strongly across most of the accuracy assessment metrics, it is much more tedious to build the classification. There was poor agreement of the Kappa statistic for all three classifications (<66.86%), with the object-based approach as the highest Kappa score. The heterogeneity across the landscape may render around 70% kappa scores. The mixed pixel problem may persist in 30-meter Landsat 8 imagery when used to quantify land cover with a range of shrubs, trees, grasses, and bare ground interspersed between built structures, particularly in areas where built contains more traditional structures, such as those found in the study area ubiquitously.

According to Hurskainen et al. (2019), multispectral single-date satellite data by itself is insufficient in heterogeneous landscapes with extensive human-induced LULC disturbance. This is likely a major factor in the poor results of the models within this study. The lack of fine-scale population data, soil data, multitemporal wet season data, radar other ancillary datasets at the time of this study may have contributed to the overall lack of accuracy across the object-based and random forest models. Including multitemporal wet season imagery and radar could help improve the detection of agricultural fields and settlement areas (Haack & Bechdol 2000) and reduce the errors associated with both multispectral data and radar data (Marconcini et al. 2020).

Due to constraints of this study, I was unable to include the Multiple Use Area within Bwabwata National Park, where there are settlements of predominately Khwe San people. There have been reports of resettlement of Mbukhushu speaking people from neighboring region of Kavango into the Multiple Use Area of Bwabwata National Park, as well as an influx of cattle within the park (Personal Communication, Fidi Alpers 2017) warrant further study on human footprint within the Bwabwata National Park.

Expansion of tarred roads and extension of roadways are a major concern for conservationists in the region, but as pressure for more modern infrastructure mounts to make it easier for people to travel the Zambezi Region, increasing development of settlement and agricultural lands have taken place linearly alongside the development of road infrastructure. Millions of Namibian Dollars have been spent on extending roads (New Era 2015). With the construction of new roads and the paving of older roads, cars and trucks are traveling much faster, which may result in increased risk for wildlife traveling across these roadways. Additionally, formerly inaccessible or remote areas are becoming more attractive for settlement and agricultural development, which could further reduce the connectivity of wildlife corridors

The results of this study call for further investigation into the human footprint, land cover change, land use decisions. The human footprint, environmental change, and climate change could pose threats to livelihoods, particularly those dependent on rainfed agriculture, natural resource collection, livestock, and ecotourism. Ecotourism and hunting revenues are dependent on healthy wildlife populations, and it is important to prioritize building new settlements and agriculture in areas that do not further fragment wildlife habitats or block major migration corridors. Moreover, increasing habitat fragmentation can lead to increasing threats from human wildlife conflict, and human lives and livelihoods.

REFERENCES

Anderson, G. P., et al. 2002. MODTRAN4-based atmospheric correction algorithm: FLAASH (Fast Line-of-sight Atmospheric Analysis of Spectral Hypercubes). *Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery VIII*, International Society for Optics and Photonics.

ArcGIS Pro, Version 2.8.2. ESRI. 2021.

Baig, Zhang, Shuai, and Tong. 2014. “Derivation of a Tasseled Cap Transformation Based on Landsat 8 At-Satellite Reflectance.” *Remote Sensing Letters* 5 (5): 423–31. doi:10.1080/2150704X.2014.915434.

Bailey, K.M., Drake, M.D., Salerno, J., Cassidy, L., Gaughan, A.E., Stevens, F.R., Pricope, N.G., Woodward, K.D., Luwaya, H.M. and Hartter, J., 2020. Mapping natural resource collection areas from household survey data in Southern Africa. Applied Geography, 125, p.102326.

Baatz, M. and A. Schäpe. 2000. Multiresolution segmentation: an optimization approach for high quality multi-scale image segmentation. 2000.

Benz, U.C., Hofmann, P., Willhauck, G., Lingenfelder, I. and M. Heynen. 2004. Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information. *ISPRS Journal of photogrammetry and remote sensing*, *58*(3-4), pp.239-258.

Blaschke, Thomas. 2010. “Object Based Image Analysis for Remote Sensing.” *ISPRS Journal of Photogrammetry and Remote Sensing* 65 (1): 2–16. doi:10.1016/j.isprsjprs.2009.06.004.

Blaschke, T., G.J. Hay, M. Kelly, S. Lang, P. Hofmann, E. Addick, R. Q. Feitosa, et al. 2014. “Geographic Object-Based Image Analysis - Towards a New Paradigm.” ISPRS Journal of Photogrammetry and Remote Sensing 87: 180–91. doi:10.1016/j.isprsjprs.2013.09.014.

Caixeta, Fernando Ferreira. 2016."An evaluation of Multiple Endmember Spectral Mixture Analysis applied to Landsat 8 OLI images for mapping land cover in southern Africa's Savanna." (2016). Electronic Theses and Dissertations.Paper 2419. https://doi.org/10.18297/etd/2419

Chen, J., Chen, J., Liao, A., Cao, X., Chen, L., Chen, X., He, C., Han, G., Peng, S., Lu, M. and Zhang, W., 2015. Global land cover mapping at 30 m resolution: A POK-based operational approach. ISPRS Journal of Photogrammetry and Remote Sensing, 103, pp.7-27.

Cheţan, M.A., Dornik, A. and Urdea, P., 2017. Comparison of object and pixel-based land cover classification through three supervised methods. Journal of Geodaesy, Geoinformation and Land Management.

Clausi, D.A., 2002. An analysis of co-occurrence texture statistics as a function of grey level quantization. *Canadian Journal of remote sensing*, *28*(1), pp.45-62.

Craigie, I.D., Baillie, J.E., Balmford, A., Carbone, C., Collen, B., Green, R.E. and Hutton, J.M., 2010. Large mammal population declines in Africa's protected areas. Biological Conservation, 143(9), pp.2221-2228.

Crist, E. P., & R.C. Cicone 1984. A physically-based transformation of Thematic Mapper data---The TM Tasseled Cap. *IEEE Transactions on Geoscience and Remote sensing*, (3), 256-263.

Cumming, D.H., 2008. Large scale conservation planning and priorities for the Kavango-Zambezi Transfrontier Conservation Area. Unpublished report commissioned by Conservation International.

Esch, T., Marconcini, M., Felbier, A., Roth, A., Heldens, W., Huber, M., Schwinger, M., Taubenböck, H., Müller, A. and Dech, S., 2013. Urban footprint processor—Fully automated processing chain generating settlement masks from global data of the TanDEM-X mission. *IEEE Geoscience and Remote Sensing Letters*, *10*(6), pp.1617 - 1621.

Esch, T., W. Heldens, A. Hirner, M. Keil, M. Marconcini, A. Roth, J. Zeidler, S. Dech, and E. Strano. 2017. Breaking new ground in mapping human settlements from space – The Global Urban Footprint. *ISPRS Journal of Photogrammetry and Remote Sensing* (134), pp. 30–42.

FAO, IFAD, UNICEF, WFP and WHO. 2017. The State of Food Security and Nutrition in the World 2017. *Building resilience for peace and food security.* Rome, FAO. Retrieved from http://www.fao.org/3/a-I7695e.pdf. Accessed September 5, 2018

Foley, J.A., DeFries, R., Asner, G.P., Barford, C., Bonan, G., Carpenter, S.R., Chapin, F.S., Coe, M.T., Daily, G.C., Gibbs, H.K. and Helkowski, J.H., 2005. Global consequences of land use. *science*, *309*(5734), pp.570-574.

Funk, C., Peterson, P., Landsfeld, M. et al. (2015) The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes. *Science Data*. (2). https://doi.org/10.1038/sdata.2015.66

Gao, B.C., 1996. NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space. Remote sensing of environment, 58(3), pp.257-266.

Gaughan, A.E., Staub, C.G., Hoell, A., Weaver, A. and Waylen, P.R., 2016. Inter-and Intra-annual precipitation variability and associated relationships to ENSO and the IOD in southern Africa. International Journal of Climatology, 36(4), pp.1643-1656.

Gibbes, C., Adhikari, S., Rostant, L., Southworth, J. and Qiu, Y., 2010. Application of object based classification and high resolution satellite imagery for savanna ecosystem analysis. *Remote Sensing*, *2*(12), pp.2748-2772.

Gorelick, N., Marconcini, M., Üreyen, S., Zeidler, J., Svaton, V., and T. Esch. 2017. "Mapping the Urban Side of the Earth- the new GUF+ Layer", vol. 2017.

Grimm, N.B., Faeth, S.H., Golubiewski, N.E., Redman, C.L., Wu, J., Bai, X. and J.M. Briggs. 2008. Global change and the ecology of cities. *Science*, 319(5864), pp.756-760.

Haack, B. and Bechdol, M., 2000. Integrating multisensor data and RADAR texture measures for land cover mapping. Computers & Geosciences, 26(4), pp.411-421.

Haralick, R.M., 1979. Statistical and structural approaches to texture. *Proceedings of the IEEE*, *67*(5), pp.786-804.

Haralick, R, K. Shanmugan, and I. Dinstein. 1973. "Textural features for image classification," *IEEE Trans. Syst., Man, Cybern*., vol. SMC-3, no. 1, pp. 610–621, Nov. 1973.

Hay, G.J. and G. Castilla. 2008. Geographic Object-Based Image Analysis (GEOBIA): A new name for a new discipline. In *Object-based image analysis* (pp. 75-89). Springer, Berlin, Heidelberg.

Healey, S.P., Cohen, W.B., Zhiqiang, Y. and O.N. Krankina. 2005. Comparison of Tasseled Cap-based Landsat data structures for use in forest disturbance detection. *Remote Sensing of Environment*, *97*(3), pp.301-310.

Ho, T.K., 1995, August. Random decision forests. In Document analysis and recognition, 1995., proceedings of the third international conference on (Vol. 1, pp. 278-282). IEEE.

Hurskainen, P., Adhikari, H., Siljander, M., Pellikka, P.K.E. and Hemp, A., 2019. Auxiliary datasets improve accuracy of object-based land use/land cover classification in heterogeneous savanna landscapes. *Remote sensing of environment*, *233*, p.111354.

IUCN Red List of Threatened Species. Version 2017-3. Website. Accessed online 16 March 2018 at: www.iucnredlist.org.

Ji, L., Zhang, L., Wylie, B.K. and Rover, J., 2011. On the terminology of the spectral vegetation index (NIR− SWIR)/(NIR+ SWIR). International journal of remote sensing, 32(21), pp.6901-6909.

Johannes, H., 2017. Assessing the accuracy of multi-temporal built-up land layers across rural-urban trajectories in the United States. *Remote sensing of environment.*

Kamwi, J.M., Chirwa, P.W., Manda, S.O., Graz, P.F. and Kätsch, C., 2015. Livelihoods, land use and land cover change in the Zambezi Region, Namibia. Population and Environment, 37(2), pp.207-230.Kauth, R. J., & G.S. Thomas. 1976. The tasseled cap--a graphic description of the spectral-temporal development of agricultural crops as seen by Landsat. *In LARS Symposia (p. 159).*

Klotz, M., Kemper, T., Geiß, C., Esch, T. and Taubenböck, H., 2016. How good is the map? A multi-scale cross-comparison framework for global settlement layers: Evidence from Central Europe. *Remote Sensing of Environment,* 178, pp.191-212.

Lambin, E.F. and Meyfroidt, P., 2011. Global land use change, economic globalization, and the looming land scarcity. *Proceedings of the National Academy of Sciences,* 108(9), pp.3465-3472.

Lambin, E.F., Turner, B.L., Geist, H.J., Agbola, S.B., Angelsen, A., Bruce, J.W., Coomes, O.T., Dirzo, R., Fischer, G., Folke, C. and P. George. 2001. The causes of land-use and land-cover change: moving beyond the myths. *Global environmental change*, *11*(4), pp.261-269.

Laris, P.S., 2005. Spatiotemporal problems with detecting and mapping mosaic fire regimes with coarse-resolution satellite data in savanna environments. *Remote Sensing of Environment*, 99(4), pp.412-424.

Lary, D.J., Alavi, A.H., Gandomi, A.H. and A.L. Walker. 2016. Machine learning in geosciences and remote sensing. *Geoscience Frontiers*, 7(1), pp.3-10.

Leitao, A. B., & Ahern, J. (2002). Applying landscape ecological concepts and metrics in sustainable landscape planning. Landscape and urban planning, 59(2), 65-93.

Marconcini, Mattia; Metz-Marconcini, Annekatrin; Üreyen, Soner; Palacios-Lopez, Daniela; Hanke, Wiebke; Bachofer, Felix; et al. (2020): World Settlement Footprint (WSF) 2015. figshare. Dataset. https://doi.org/10.6084/m9.figshare.10048412.v1

Meinel, G. and Neubert, M., 2004. A comparison of segmentation programs for high resolution remote sensing data. *International Archives of Photogrammetry and Remote Sensing*, *35*(Part B), pp.1097-1105.

Mendelsohn J (2006). Farming Systems in Namibia. NNFU: Windhoek, Namibia.

Mendelsohn, J. and C. Roberts. 1997. Chapter 2: General description of Caprivi. *An Environmental Profile and Atlas of Caprivi.* Directorate of Environmental Affairs. Namibia. pp 4-7. Ministry of Environment, Forestry, and Tourism. 2014. Environment clears Tobacco Project. 11 April 2014. https://www.met.na/new/130/Environment-clears-tobacco-project/ . Accessed 8/30/21.

Ministry of Lands and Resettlement, 2015. Integrated Rural Land Use Plan for the Zambezi Region of Namibia. March 2015.

Mishra, N.B. and Crews, K.A., 2014. Mapping vegetation morphology types in a dry savanna ecosystem: integrating hierarchical object-based image analysis with Random Forest. *International Journal of Remote Sensing*, *35*(3), pp.1175-1198.

Mohammed, I., Marshall, M., de Bie, K., Estes, L. and Nelson, A., 2020. A blended census and multiscale remote sensing approach to probabilistic cropland mapping in complex landscapes. ISPRS journal of photogrammetry and remote sensing, 161, pp.233-245.

Mutanga, O., Van Aardt, J. and L. Kumar. 2009. Imaging spectroscopy (hyperspectral remote sensing) in southern Africa: an overview. *South African Journal of Science*, 105(5-6), pp.193-198.

Murwira, A., A. K. Skidmore. 2006. Monitoring change in the spatial heterogeneity of vegetation cover in an African savanna, International Journal of Remote Sensing, 27:11, 2255-2269, DOI: 10.1080/01431160500396683

Myint, S.W., Gober, P., Brazel, A., Grossman-Clarke, S. and Q. Weng. 2011. Per-pixel vs. object-based classification of urban land cover extraction using high spatial resolution imagery. *Remote sensing of environment*, *115*(5), pp.1145-1161

Myneni, R.B., Hall, F.G., Sellers, P.J. and Marshak, A.L., 1995. The interpretation of spectral vegetation indexes. IEEE Transactions on Geoscience and Remote Sensing, 33(2), pp.481-486.

Nabil, M., Zhang, M., Bofana, J., Wu, B., Stein, A., Dong, T., Zeng, H. and Shang, J., 2020. Assessing factors impacting the spatial discrepancy of remote sensing based cropland products: A case study in Africa. International Journal of Applied Earth Observation and Geoinformation, 85, p.102010.

Naidoo, R., Weaver, L.C., Stuart-Hill, G. and Tagg, J., 2011. Effect of biodiversity on economic benefits from communal lands in Namibia. *Journal of Applied Ecology*, *48*(2), pp.310-316.

Naidoo, L., Cho, M.A., Mathieu, R. and 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 journal of Photogrammetry and Remote Sensing*, *69*, pp.167-179.

Naidoo, R., Weaver, L.C., Diggle, R.W., Matongo, G., Stuart-Hill, G. and Thouless, C., 2016. Complementary benefits of tourism and hunting to communal conservancies in Namibia. *Conservation Biology*, *30*(3), pp.628-638.

Nagendra, H., Munroe, D.K. and Southworth, J., 2004. From pattern to process: landscape fragmentation and the analysis of land use/land cover change.

Nagendra, H., Lucas, R., Honrado, J.P., Jongman, R.H., Tarantino, C., Adamo, M. andMairota, P., 2013. Remote sensing for conservation monitoring: Assessing protected areas, habitat extent, habitat condition, species diversity, and threats. *Ecological Indicators,* 33, pp.45-59.

Newton, A.C., Hill, R.A., Echeverría, C., Golicher, D., Rey Benayas, J.M., Cayuela, L. and Hinsley, S.A., 2009. Remote sensing and the future of landscape ecology. *Progress in Physical Geography*, *33*(4), pp.528-546.

Nkengbeza, D., Kangumu, B., Sibuku, C. 2021. The Perceptions of the Kwe (San) Community at Bwabwata on the Importance and Influence of Formal Education. International Journal of Education, Learning, and Development 9 (5): 1-19.

Ouedraogo, Issa, Jennie Barron, Siza D Tumbo, and Frederic C Kahimba. 2016. "Land Cover Transition in Northern Tanzania." Land Degradation & Development 27 (3): 682–92. doi:10.1002/ldr.2461.

Peace Parks Foundation, The Global Solution, South Africa, 2010. Accessed online September 24, 2017 at: https://www.peaceparks.org.

Pesaresi, M., D. Ehrlich, S. Ferri, A. Florczyk, F. S. M. Carneiro, S. Halkia, M. Andreea, T. Kemper, P. Soille, V. Syrris. 2016. Operating Procedure for the Production of the Global Human Settlement Layer from Landsat data of the epochs 1975, 1990, 2000, and 2014. Publications Office of the European Union. JRC97705

Peters, D., BT Bestelmeyer, JE Herrick, EL Fredrickson, HC Monger, KM Havstad. 2006. Disentangling Complex Landscapes: New Insights into Arid and Semiarid System Dynamics, BioScience 56 (6), pp. 491–501, https://doi.org/10.1641/0006-3568(2006)56[491:DCLNII]2.0.CO;2

Petit, C., Scudder, T. and Lambin, E., 2001. Quantifying processes of land-cover change by remote sensing: resettlement and rapid land-cover changes in south-eastern Zambia. *International Journal of Remote Sensing*, *22*(17), pp.3435-3456.

Puissant, A., Rougier, S. and Stumpf, A., 2014. Object-oriented mapping of urban trees using Random Forest classifiers. *International Journal of Applied Earth Observation and Geoinformation*, *26*, pp.235-245.

Raichlen DA, Gordon AD, Harcourt-Smith WEH, Foster AD, W. Haas. 2010. Laetoli Footprints Preserve Earliest Direct Evidence of Human-Like Bipedal Biomechanics. PLoS ONE 5(3): e9769. https://doi.org/10.1371/journal.pone.0009769

Reardon, T., Barrett, C., Kelly, V. and Savadogo, K., 1999. Policy reforms and sustainable agricultural intensification in Africa. Development policy review, 17(4), pp.375-395.

Rodriguez-Galiano, V.F., Ghimire, B., Rogan, J., Chica-Olmo, M. and Rigol-Sanchez, J.P., 2012. An assessment of the effectiveness of a random forest classifier for land-cover classification. ISPRS Journal of Photogrammetry and Remote Sensing, 67, pp.93-104.

Rothauge, A., 2007. Some principles of sustainable rangeland management in Namibia. Agricola, 17, pp.7-15.

Rouse, J. W., Haas, R. H., Schell, J. A., and D.W. Deering. 1973. Monitoring vegetation systems in the Great Plains with ERTS. In Third ERTS Symposium, Vol. I.

SADC. 2018. Website: https://www.sadc.int/issues/environment-sustainable-development/ . Accessed 10/28/2018.

Salerno, J., Stevens, F., Gaughan, A., Hilton, T., Bailey, K., Bowles, T., Cassidy, L., Mupeta-Muyamwa, P., Biggs, D., Pricope, N. and Mosimane, A.W., et al. 2021. Wildlife impacts and changing climate pose compounding threats to human food security. Current Biology 31, pp. 1-9.

Sanderson, E.W., Jaiteh, M., Levy, M.A., Redford, K.H., Wannebo, A.V. and G. Woolmer. 2002. The human footprint and the last of the wild: the human footprint is a global map of human influence on the land surface, which suggests that human beings are stewards of nature, whether we like it or not. BioScience, 52(10), pp.891-904.

Sankaran, M., Ratnam, J. and N.P. Hanan. 2004. Tree–grass coexistence in savannas revisited–insights from an examination of assumptions and mechanisms invoked in existing models. Ecology letters, 7(6), pp.480-490.

Scholes, R.J. and S.R Archer., 1997. Tree-grass interactions in savannas. Annual review of Ecology and Systematics, 28(1), pp.517-544.

Shanmugapriya, P., Rathika, S., Ramesh, T. and Janaki, P., 2019. Applications of remote sensing in agriculture-A Review. International Journal of Current Microbiology and Applied Sciences, 8(01), pp.2270-2283.

Sriwongsitanon, N., Gao, H., Savenije, H.H.G., Maekan, E., Saengsawan, S. and Thianpopirug, S., 2015. The Normalized Difference Infrared Index (NDII) as a proxy for soil moisture storage in hydrological modelling. Hydrology & Earth System Sciences Discussions, 12(8).

Stumpf, A. and N. Kerle. 2011. Object-oriented mapping of landslides using RandomForests. Remote sensing of environment, 115(10), pp.2564-2577.

Sweeney, S., Ruseva, T., Estes, L. and Evans, T., 2015. Mapping cropland in smallholder-dominated savannas: integrating remote sensing techniques and probabilistic modeling. *Remote Sensing*, *7*(11), pp.15295-15317.

Thakkar, Amee K, Venkappayya R Desai, Ajay Patel, and Madhukar B Potdar. 2016. “An Effective Hybrid Classification Approach Using Tasseled Cap Transformation (Tct) for Improving Classification of Land Use/Land Cover (Lu/Lc) in Semi-Arid Region: A Case Study of Morva-Hadaf Watershed, Gujarat, India.” *Arabian Journal of Geosciences*9 (3): 1–13. doi:10.1007/s12517-015-2267-8.

Tian, J. and D.M. Chen. 2007. Optimization in multi-scale segmentation of high-resolution satellite images for artificial feature recognition. *International Journal of Remote Sensing*, *28*(20), pp.4625-4644.

Turner, W., Spector, S., Gardiner, N., Fladeland, M., Sterling, E. and Steininger, M., 2003. Remote sensing for biodiversity science and conservation. Trends in ecology & evolution, 18(6), pp.306-314.\\

United Nations General Assembly, 2015. Sustainable development goals. *SDGs Transform Our World*, *2030*.

Venter, O., Sanderson, E.W., Magrach, A., Allan, J.R., Beher, J., Jones, K.R., Possingham, H.P., Laurance, W.F., Wood, P., Fekete, B.M. and Levy, M.A., 2016. Sixteen years of change in the global terrestrial human footprint and implications for biodiversity conservation. Nature Communications, 7, p.12558.

Venter, O., E. W. Sanderson, A. Magrach, J. R. Allan, J. Beher, K. R. Jones, H. P. Possingham, W. F. Laurance, P. Wood, B. M. Fekete, M. A. Levy, and J. E. Watson. 2018. Last of the Wild Project, Version 3 (LWP-3): 1993 Human Footprint, 2018 Release. Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC). https://doi.org/10.7927/H4H9938Z. Accessed 5 November 2018.

Vitousek, P.M., Antonio, C.M., Loope, L.L. and Westbrooks, R., 1996. Biological invasions as global environmental change. American scientist, 84(5), p.468.

Wildlife Conservation Society - WCS, and Center for International Earth Science Information Network - CIESIN - Columbia University. 2005. Last of the Wild Project, Version 2, 2005 (LWP-2): Global Human Footprint Dataset (Geographic). Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC). [http://dx.doi.org/10.7927/H4M61H5F. Accessed 12/4/2018](http://dx.doi.org/10.7927/H4M61H5F)

Wingate, V.R., Phinn, S.R., Kuhn, N., Bloemertz, L. and Dhanjal-Adams, K.L., 2016. Mapping decadal land cover changes in the woodlands of north eastern Namibia from 1975 to 2014 using the Landsat satellite archived data. *Remote Sensing*, *8*(8), p.681.

Wójtowicz, M., Wójtowicz, A. and Piekarczyk, J., 2016. Application of remote sensing methods in agriculture. Communications in Biometry and Crop Science, 11(1), pp.31-50.

Woodward, K.D., Pricope, N.G., Stevens, F.R., Gaughan, A.E., Kolarik, N.E., Drake, M.D., Salerno, J., Cassidy, L., Hartter, J., Bailey, K.M. and Luwaya, H.M., 2021. Modeling Community-Scale Natural Resource Use in a Transboundary Southern African Landscape: IntegratingRemote Sensing and Participatory Mapping. Remote Sensing, 13(4), p.631.

Woolmer, G., Trombulak, S.C., Ray, J.C., Doran, P.J., Anderson, M.G., Baldwin, R.F., Morgan, A. and E.W. Sanderson. 2008. Rescaling the human footprint: a tool for conservation planning at an ecoregional scale. *Landscape and Urban Planning*, 87(1), pp.42-53.

Wójtowicz, M., Wójtowicz, A. and J. Piekarczyk. 2016. Application of remote sensing methods in agriculture. Communications in Biometry and Crop Science, 11(1), pp.31-50.

Xu, H., 2006. Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery. International journal of remote sensing.27(14), pp.3025-3033.

Yang, F., Matsushita, B., Yang, W. and Fukushima, T., 2014. Mapping the human footprint from satellite measurements in Japan. ISPRS journal of Photogrammetry and Remote Sensing, 88, pp.80-90.

Zanchetta, Anna, Gabriele Bitelli, and Arnon Karnieli. 2016. "Monitoring desertification by remote sensing using the Tasselled Cap transform for long-term change detection." *Natural Hazards: Journal of the International Society for the Prevention and Mitigation of Natural Hazards* **83**(1): 223-237.

Zha, Y., Gao, J. and S. Ni. 2003. Use of normalized difference built-up index in automatically mapping urban areas from TM imagery. International Journal of Remote Sensing, 24(3), pp.583-594.

APPENDIX A

Training Site Sample

```
<h:html xmlns="http://www.w3.org/2002/xforms"
xmlns:h="http://www.w3.org/1999/xhtml"
xmlns:ev="http://www.w3.org/2001/xml-events"
xmlns:xsd="http://www.w3.org/2001/XMLSchema"
xmlns:jr="http://openrosa.org/javarosa">
  <h:head>
    <h:title>Zambezi_HF_Form2017</h:title>
    <model>
      <instance>
        <data id="build_Zambezi-HF-Form2017_1500460427">
          <meta>
            <instanceID/>
          </meta>
          <Hum_act/>
          <LogInfo>
            <TypeSam/>
            <VilName/>
            <AREA/>
            <Region>
              Zambezi Region
            </Region>
            <Date/>
            <Time/>
            <Loc/>
            <Loc2/>
            <Collector/>
            <ObsType/>
          </LogInfo>
          <RefSamp>
            <Ref_ID/>
            <Sam_Typ/>
            <ObsTyp3/>
          </RefSamp>
          <Canopy/>
          <PanPic/>
          <Ground_P/>
          <RefDesc>
            <AccCLS/>
            <RdEst/>
            <Topo/>
            <Slope/>
            <Azimuth/>
            <RivDist/>
            <Weather/>
          </RefDesc>
          <LC1>
            <LCType/>
            <B_INT/>
            <Veg_Type/>
            <undstry/>
            <Disturb_Reg/>
            <infras/>
```

```
    <Ag_Type/>
    <WetType/>
    <Animal/>
  </LC1>
  <C_Info>
  </C_Info>
  <PctCano/>
  <GCEst>
    <PctHerb/>
    <PctLit/>
    <PctSoil/>
    <PctRock/>
    <GType/>
  </GCEst>
  <Canop1>
    <CanClos/>
    <avgCan/>
    <HtEmTr/>
    <NEmTree/>
    <DBHAvT/>
    <AvEmTre/>
    <VisUND/>
    <PreSeed/>
    <PresSap/>
  </Canop1>
  <SpecVeg>
    <GenusN/>
    <EnBush/>
    <VgTrend/>
    <LUHist/>
  </SpecVeg>
  <untitled54/>
  <DomSpec/>
  <BushEn/>
  <GenObs/>
  <untitled62/>
  <Ph_Set/>
  <endtime/>
  </>
  <untitled66/>
</data>
```

APPENDIX B

```
##### Ariel
Weaver#####################################################################
##### Tasseled Cap Transformation for Landsat 8 Imagery in R
                    #
##### December 5, 2017
        #
##############################################################################
###

##LOAD WORKSPACE

setwd = "c:\\tmp\\Tassled Cap\\output\\"

#install.packages(RStoolbox)
#install.packages(raster)
#install.packages(sp)
#install.packages(rgdal)

library(raster)
library(sp)
library(rgdal)
library(RStoolbox)

### load raster as rasterbrick to prepare for tassled cap
zr <- stack(readGDAL("C:\\tmp\\Tassled Cap\\JUL
2017\\input\\OLI_2017_JUL_ZR_EAST_MOS_FLAASH_SUB.dat"))
zr

attributes(zr)

# Create an RGB (654 bands, band 1 ommited from layerstacked image, if you need to adapt
that if your layerstack includes band 1)
# image from the raster stack

lsat_tc <- tasseledCap(zr[[c(1:6)]], sat = "landsat8oli")
lsat_tc

names(lsat_tc) = c("brightness", "greeness", "wetness")
plot(lsat_tc)

tc_data <- plotRGB(lsat_tc, 3, 2, 1, stretch = "hist", colNA = "black")

writeRaster(lsat_tc, "tcZAMBEZI2017_jul_oli_sub_mos",
format="ENVI",  options="INTERLEAVE=BIP", overwrite=TRUE)
```

APPENDIX C

LIST OF ABBREVIATIONS

ASM – Angular Second Moment
CAMPFIRE – Communal Areas Management for Indigenous Resources
CBC – Community Based Conservation
CBNRM – Community Based Natural Resource Management
CBO – Community Based Organizations
DR – Detection Rate
DRP – Detection Rate Prevalence
GEOBIA – Geographic Object Based Image Analysis
GHF – Global Human Footprint
GIS – Geographic Information Systems
GLC30 – GLOBELAND 30
GPS – Global Positioning System
GUI – Graphic User Interface
GUF – Global Urban Footprint
GHSL – Global Human Settlement Layer
GLCM – Grey Level Co-Occurrence Matrix
IRDNC – Integrated Rural Development and Nature Conservation (NGO)
KAZA – Kavango-Zambezi Trans-Frontier Conservation Area
LULC – Land Use Land Cover
MNDWI – Modified Normalized Difference Water Index
NDBI – Normalized Difference Built Index
NDII – Normalized Difference Infrared Index
NDMIR – Normalized Difference Mid Infrared Index
NDVI – Normalized Difference Vegetation Index
NDWI – Normalized Difference Water Index
NGO – Non-Governmental Organization
NNF – Namibian Nature Fund (NGO)
NPV- Negative Predictive Value
OBIA – Object-based Image Analysis
OOB – Out-of-Bag
PA- Producer's Accuracy
PPF – Peace Parks Foundation
PPV – Positive Prediction Value
RF – Random Forest
RS – Remote Sensing
SADC – Southern African Development Community (Transnational Organization)
TCT – Tasseled Cap Transformation
TFCA – Trans Frontier Conservation Area
UA – User's Accuracy
WSF – World Settlement Footprint
WWF – World Wildlife Fund (NGO)

CURRICULUM VITA

Name: Ariel E. Weaver

Address: 1445 Goddard Avenue
Louisville, Kentucky 40204

DOB: Tulsa, Oklahoma – November 5, 1986

EDUCATION
& TRAINING: B.S. in Applied Geography
University of Louisville
2015

M.S. in Applied Geography
University of Louisville
2021

AWARDS: **Fulbright Research Grant** 2017 *($20,000 + Travel)*

Leonard Pardue Scholarship 2015 *($500)*

Wilma Shaw Roseman Scholarship 2014 *($500)*

Rachel Davis Scholarship 2014 *($5000)*

Kentucky Academic Commonwealth Scholarship 2014

Critical Language Scholarship 2007 *($5000+ Travel)*

PROFESSIONAL SOCIETIES: American Association of Geographers
Kentucky Association of Mapping Professionals

PUBLICATIONS: *Gaughan, A. E., Staub, C. G., Hoell, A., Weaver, A., & Waylen, P. R. (2015). Inter-and Intra-annual precipitation variability and associated relationships to ENSO and the IOD in southern Africa. International Journal of Climatology.*

Salerno, J., Stevens, F.R., Gaughan, A.E., Hilton, T., Bailey, K., Bowles, T., Cassidy, L., Mupeta-Muyamwa, P., Biggs, D., Pricope, N. and Mosimane, A.W., AE Weaver, 2021. Wildlife impacts and changing climate pose compounding threats to human food security. *Current Biology*.

Weaver, A.E., Williams, P., Wright, N., Woolsey, A. and Yenney, E., 2015. Social Sustainability Guidelines and Metrics for Transportation in Louisville: A Proposal for TARC's APTA Sustainability Commitment.

Stevens, F.R., Gaughan, A.E., Pricope, N., Salerno, J., Hartter, J., Cassidy, L., Drake, M., Weaver, A., Kolarik, N. and Woodward, K., 2019, August. Unmixing our methods to link land functions and environmental change to food insecurity in a southern African savanna context. In *2019 ESA Annual Meeting (August 11--16)*. ESA.

PRESENTATIONS: *Inter-and Intra-annual precipitation variability and associated relationships to ENSO and the IOD in southern Africa*. **Illustrated paper** presented at the American Association of Geographer's Conference*, 2015 (*also chaired the presentation panel at the AAG conference).