Geographic Information Systems and Remote Sensing modelling of tree species diversity in the woodlands of Zimbabwe

Godfrey Mutowo (R028894A)

Thesis submitted to the Department of Geography and Environmental Sciences, University of Zimbabwe in fulfilment of the requirements for the Master of Philosophy degree in Geography

September 2010
Dedication

To my mom Mavis and my siblings Pamela, Raymond, Tatenda, and Nyasha.

To my late father Stanley Garai Mutowo, and my late sister Patron: you passed on too soon before you could reap the fruits of your wise words of counsel! May your dear souls rest in eternal peace!!

To my wife Rodie and our son Carl for being my strength to soldier on!
Declaration 1: Originality

I hereby declare that this thesis submitted for the Master of Philosophy degree at the University of Zimbabwe is my original work and has not been previously submitted to any other institution of higher education. I further declare that all sources cited or quoted are indicated by means of a comprehensive list of references.

Mutowo Godfrey
copyright@ University of Zimbabwe, 2010
Declaration Statement: Publication in Preparation

Details that form part and/ or include research presented in this thesis (include publications in submitted, as well as in preparation.

Publication submitted and accepted with revision 1: Mutowo G¹ and Murwira A² (accepted at the International Journal of Remote Sensing) Relationship between remotely-sensed variables and tree species diversity in savanna woodlands of Southern Africa. The work was conducted by the first author under the guidance and supervision of the second author.

¹, ² University of Zimbabwe, Department of Geography and Environmental Science, P.O. Box, MP167, Mount Pleasant, Harare, Zimbabwe.

Publication 2: Mutowo G¹ and Murwira A² (accepted at with minor revisions at Geocarto International Journal) GIS and remote sensing based spatial modelling of tree species diversity in the savanna woodlands of Southern Africa. The work was conducted by the first author under guidance and supervision of the second author.

¹, ² University of Zimbabwe, Department of Geography and Environmental Science, P.O. Box, MP167, Mount Pleasant, Harare, Zimbabwe.

Certification by Supervisor (Dr Amon Murwira)

----------------------------------------------------------------------------------
Acknowledgements

I would like to express my heartfelt gratitude to Dr Amon Murwira for his unwavering support throughout my tenure as a graduate student. It was a tough road through and through but your encouragement and counsel kept me on the go. Special mention also goes to, Mbulisi, Searchmore, Emmerson, Timothy, Tawanda, George and Gift, some of my fellow MPhil students. Together colleagues we kept each other going strong. Also, the jokes you provided made the journey shorter and worthwhile undertaking. To I Gwitira, M.D. Shekede and Dr M Masocha, I say thank you for the guidance that you provided from time to time. It really was a pleasure working with all of you. On a parting note, I would also like to thank Xavier Poshiwa, Patience Zisadza, Inos Dhau and Thulani Majoni for their help in collecting field data in Mabalauta and Kyle respectively.

It was my pleasure to have all of you around!!!!
Abstract

The continued loss of biological diversity has prompted managers and conservationists to find ways of monitoring ecosystems for the purposes of protecting and preventing further loss of biological diversity. Remote sensing has been suggested as a way of monitoring ecosystems because of its ability to provide spatial data which can be updated relatively faster and at a relatively cheaper than conventional methods. The relationship between tree species diversity and satellite derived indices is well documented. However, not much work has been done in mapping tree species using the obtained relationships, more so in savanna woodlands. Thus, the main objective of this thesis was to investigate whether and to what extent remote sensing can be used to measure tree species diversity in savanna woodlands. In addition the thesis also tested the extent to which linear and spatial regression can be used within a Geographic Information System to predict the spatial distribution of tree species diversity. We based our study on the hypotheses that 1) variance in the reflectance within remotely-sensed images is directly related with tree species diversity in savanna woodlands, and 2) that a remote sensing index of biomass is related to tree species diversity. The former is based on the Spectral Variation Hypothesis, which predicts that species richness can be estimated from habitat heterogeneity while the latter is based on the biomass-diversity or productivity-diversity hypothesis. The biomass-diversity hypothesis predicts that optimum biodiversity is found in ecosystems which have intermediate biomass, while low species diversity is characteristic of ecosystems which have low or high biomass. Specifically, we used regression analysis to test whether and in what way tree species diversity is related to the standard deviation of the Near Infrared (NIR) radiance (a measure of spectral variation) and tree biomass estimated via the Soil Adjusted Vegetation Index (SAVI) in three selected savanna woodlands of Zimbabwe. We also tested the wider applicability of the derived regression models by applying them in a study site away from where the models were developed. Our results showed that tree species diversity has a significant (p< 0.05) hump-shaped response to variations in the standard deviation of NIR radiance and SAVI. Furthermore, results show that the combination of the standard deviation of NIR and SAVI explained between 59% and 73% of the variance in tree species diversity in the study sites. Also, we found that our regression models can be used to spatially predict tree species diversity in landscapes with comparable physical characteristics. Overall, we
conclude that remote sensing and GIS can be used to successfully estimate tree species diversity in savanna woodlands.
<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ASTER</td>
<td>Advanced Spaceborne Thermal Emission and Reflection Radiometer</td>
</tr>
<tr>
<td>CBD</td>
<td>Convention on Biological Diversity</td>
</tr>
<tr>
<td>GIS</td>
<td>Geographic Information Systems</td>
</tr>
<tr>
<td>NIR</td>
<td>Near Infra-red</td>
</tr>
<tr>
<td>SADC</td>
<td>Southern African Development Community</td>
</tr>
<tr>
<td>SAVI</td>
<td>Soil Adjusted Vegetation Index</td>
</tr>
</tbody>
</table>
## Table of Contents

Dedication ........................................................................................................................ ii  
Declaration 1: Originality ........................................................................................................ iii  
Declaration Statement: Publication in Preparation .............................................................. iv  
Acknowledgements ........................................................................................................... v  
Abstract .............................................................................................................................. vi  
List of Acronyms ................................................................................................................ viii  
List of Characters................................................................................................................... ix  
List of figures ........................................................................................................................ xi  
List of Tables ........................................................................................................................ xii  

1.0: Introduction .................................................................................................................. 1  
1.1 General Background ................................................................................................ 1  
1.2 Statement of the Problem ....................................................................................... 4  
1.3: Thesis objectives ..................................................................................................... 5  
1.5 Justification of the study ......................................................................................... 6  
1.7. Organisation of the thesis ....................................................................................... 7  

Chapter 2: Determinants of and consequences of loss of tree diversity in savanna woodlands ................................................................. 8  
2.1. Major determinants of woody species distribution in the Savanna .................. 8  
2.2 The importance of tree species diversity .............................................................. 9  
2.3 Factors contributing to tree diversity loss ............................................................. 10  
2.4 Consequences of tree species diversity loss ......................................................... 11  
2.5 Remote sensing methods of predicting tree species diversity ......................... 11  

Chapter 3: Methods ........................................................................................................... 14  
3.1: Study sites ............................................................................................................. 14  
3.2. Remotely sensed data ......................................................................................... 17  
3.3. Tree species data ................................................................................................. 21  
3.3.1 Sampling and tree species data collection ..................................................... 21  
3.3.2. Tree diversity indices .................................................................................... 23  
3.4. Relating tree species diversity indices to standard deviation of NIR and SAVI... 24  
3.5. Mapping the spatial variations in tree species diversity ................................. 26  
3.6: Comparison of the linear and spatial regression predictions ............................ 27  

Results and Discussion ..................................................................................................... 29  

Chapter 4: The relationship between tree species diversity and satellite derived indices 30  
4.2: Discussion ............................................................................................................ 38  

Chapter 5: GIS and remote sensing based spatial modelling of tree species diversity .... 43  
5.1: Results .................................................................................................................... 43  
5.2: Discussion ............................................................................................................. 51  

Chapter 6: Conclusions ..................................................................................................... 53  
6.1: The relationship between tree species diversity and satellite imagery derived satellite indices .............................................................................................................. 53  
6.2: Remote sensing based spatial modelling of tree species diversity .................... 53  

References ......................................................................................................................... 55
List of figures

Figure 3.1: The location of the four woodland study sites in Zimbabwe. ......................... 15

Figure 3.2: ASTER satellite images for a) Mukuvisi, b) Kyle and c) Mabalauta study sites showing the boundary of the study site (red line) as well as the sampling points (red diamonds). Satellite images d) and e) cover part of the Sebungwe study site (red diamonds show the sampling points). Vegetated areas are shown by the green colour, and the map units are Geographic coordinates. ................................................................. 18

Figure 3.3: Distribution of the sampling points in relation to roads in Sebungwe and Mabalauta study sites. ........................................................................................................ 22

Figure 4.1: Relationships between tree canopy cover and the Soil Adjusted Vegetation Index (SAVI) in a) Mukuvisi, b) Kyle and c) Mabalauta study sites of Zimbabwe. ........... 30

Figure 4.2: Relationships between the standard deviation of NIR radiance and a) Simpson’s 1-D and b) Shannon’s H in 1) Mukuvisi, 2) Kyle and 3) Mabalauta study sites. The standard deviation of the NIR radiance was calculated at 45m. ........................................ 32

Figure 4.3: Relationships between SAVI and a) Simpson’s 1-D, and b) Shannon’s H in 1) Mukuvisi, 2) Kyle and 3) Mabalauta study sites. SAVI was averaged over a 45 m window. ............................................................................................................................. 33

Figure 4.4: Relationships between tree species diversity estimated by a) Simpson’s 1-D and b) Shannon’s H and a combination of the standard deviation of the NIR radiance and estimated by SAVI in 1) Mukuvisi, 2) Kyle and 3) Mabalauta study sites. The graphs show decreasing tree species diversity from highest diversity (deep red) to lowest diversity (dark green). ................................................................. 35

Figure 5.2: Relationship between Shannon’s H and Simpson’s 1-D indices of diversity in a) Mukuvisi, b) Kyle and c) Mabalauta study sites. ................................................................. 44

Figure 5.3: Spatial variations in tree species diversity (estimated by Simpson’s 1-D) mapped using a linear regression model in Kyle Game Reserve. ........................................ 45

Figure 5.4: Spatial variations of tree species diversity (estimated Simpson’s 1-D) predicted using the a) linear and b) spatial regression models in Mukuvisi study site. The ellipse shows the area of greatest contrast in prediction between this (spatial) model and the linear regression model. ........................................................................................................ 45

Figure 5.5: The spatial variations in tree species diversity estimated by Simpson’s 1-D index in Mabalauta study site using a) linear and b) spatial regression models. ............ 46

Figure 5.6: The relationship between tree species diversity estimated by the linear and the spatial (Conditional Auto) regression models in a) Mabalauta and b) Mukuvisi study sites. ........................................................................................................ 48
Figure 5.7: The spatial variations in tree species diversity predicted using a) linear and b) spatial regression models in Sebungwe study site. Here, Simpson’s 1-D values show areas of lowest diversity (red) through yellow to areas of highest diversity (green). The white oblique line in the circled area is a gap left between the two satellite images.

List of Tables

Table 3.1: A description of the four study sites.

Table 4.1: The distribution of tree species and the associated diversity and satellite imagery derived indices in some of the sampling plots.

Table 5.1: Results of the test for spatial autocorrelation in tree species diversity estimated by Simpson’s 1-D in the three study sites.

Table 5.2: The spatial regression equations used to relate tree species diversity to a combination of the standard deviation of NIR radiance and SAVI in Mabalauta and Mukuvisi study sites.

Table 5.3: The magnitude of error (root mean square error) resulting from the estimation of tree species diversity (Simpson’s 1-D and Shannon’s H indices) in the three study sites.

Table 5.4: The magnitude of root mean square error (RMSE) resulting from the estimation of tree species diversity (Simpson’s 1-D index) using the linear and spatial regression models in Sebungwe district.

Table 5.5: Regression models that were generated in Mabalauta study site, and were used in the prediction of tree species diversity in Mopane woodlands in Sebungwe district.
1.0: Introduction

1.1 General Background

During the last few decades, biodiversity has increasingly come under pressure as a result of the acceleration of human induced habitat destruction (Wilson 1988; Kondratyev 1998; Gillespie et al., 2008). This has called for the need to collect data on species diversity for the purposes of conservation. The most common way to collect data on species diversity and their distributions are intensive ground surveys or inventories of species in the field. The expense and time required for such studies make it difficult or even impractical to expect that all potential sites for biodiversity conservation can be timely inventoried (Moerman and Estabrook 2006; Pautasso and McKinney 2007). In addition, consistency of the results of ground-based biodiversity surveys is often difficult to achieve because different field data sources can lead to dissimilar maps of species distributions and diversity due to subjectivity, even in relatively well-studied areas (Graham and Hijmans 2006). Consequently, it has become imperative to adopt robust and rapid methods to understand the spatial distribution of biodiversity, as well as its changes through time (Whittaker et al., 2001; Lomolino et al., 2004) for the purposes of biodiversity conservation.

To this end, remote sensing has long been proposed as a potential tool for use in understanding the distribution of vegetation species diversity because it covers large spatial extents, and has a large spectral coverage which enables the differentiation of different vegetation types (Muldavin et al., 2001; Duro et al., 2007). Nevertheless, the level of success in the application of remote sensing for understanding ecological patterns such as tree species diversity largely depends on the spatial resolution of the images, as
well as the appropriateness of the remotely-sensed indices used to relate to ground measurements of biological diversity indicators. In addition, the ecological relevance of the hypotheses that drive the extraction of these remotely-sensed indicators has also been identified as an important factor determining the level of success of the remote sensing methods (Nagendra 2001; Turner et al., 2003; Rocchini 2007). This is because potentially chance relationships, which only apply in the study site they were derived and therefore do not allow for generalisations, may exist between remotely-sensed indices and biological diversity.

Two hypothetical frameworks that have potential in improving the application of remotely-sensed indices to determine biodiversity in ecosystems are the biomass-biodiversity hypothesis and the spectral variation hypothesis (SVH) hypothesis. The biomass-biodiversity hypothesis predicts that in the presence of abundant and reliable resources, species biomass is high and become more specialised, allowing more species per unit area (Guo 2006). In this case, a hump–shaped relationship is predicted where low levels of biomass coincide with low species diversity, intermediate biomass levels are associated with high species diversity while high levels of biomass coincide with low species diversity (van Ruijven and Berendse 2005; Fargione et al., 2007). In savanna ecosystems, low biomass ecosystems are expected to be associated with low diversity, while at intermediate biomass diversity is highest. At very high biomass, tree species diversity is expected to be low because such areas are associated with intense competition for light leading to dominance of usually a single species (Huston 1997; Oindo and Skidmore 2002). This means that more species will be found where there is very little competition for resources amongst the species. The hump-shaped response has mainly been explained by the complementarity effect (Guo 2006; Fargione et al. 2007). Complementarity results when ecological differences among species lead to more
complete utilisation of resources (Qin et al., 2003). Vegetation indices such as the normalised difference vegetation index (NDVI) or the soil adjusted vegetation index (SAVI) have been proven to be indicators of biomass (Huete 1988; Huete et al., 1997; Oindo and Skidmore 2002). Thus, the biomass-biodiversity hypothesis can be substituted with a SAVI-biodiversity relationship. However, this general hypothesis has been untested especially in savanna woodlands. Although vegetation indices such as NDVI or SAVI have been used as a proxy for mapping diversity in landscapes, there is little evidence that their application has been conducted within the productivity-biodiversity hypothetical framework (Oindo and Skidmore 2002).

Recent applications of remote sensing have also proposed the Spectral Variation Hypothesis (SVH) as a hypothetical framework that could be used to apply remotely-sensed data in estimating diversity in ecosystems. This hypothesis states that spectral heterogeneity in remotely-sensed imagery is related to habitat heterogeneity and thus to species diversity in that ecosystem (Palmer et al., 2002). Specifically, the SVH predicts that higher species diversity is associated with areas which are more heterogeneous (Palmer et al. 2002; Rocchini et al., 2007). Potentially, the SVH can be used to map tree species diversity at different observational scales. For example, the high reflectance in the near infrared is linked to scattering processes at the leaf scale in such a way that different types of vegetation show distinctive variability that can be distinguished on satellite imagery (Rocchini et al. 2007). At the landscape scale, the spectral heterogeneity must be related to the complexity and structure of the landscape, properties which are related to habitat heterogeneity and therefore to species diversity (Oindo and Skidmore 2002; Oldeland et al., 2010). Thus, in this context, it is reasonable to hypothesise that variations in the spectral reflectance, measured by the standard deviation of the NIR, should be
significantly correlated to the diversity of tree species. To the best of our knowledge, the SVH hypothesis has not been widely statistically tested especially in savanna landscapes.

The standard deviation of NIR radiance and SAVI can potentially be correlated. This is because they are both derived from the near infrared band. Ecologically the amount of variation in the NIR band could be related to tree canopy cover and therefore to the amount of biomass present. However, in this study we hypothesise that the standard deviation in the NIR radiance is a function of the variety in the reflecting leaf surfaces, which do not necessarily have to be correlated to the amount of leaf matter (canopy cover) in a particular sampling plot. Therefore, it is reasonable to hypothesise that a combination of spectral heterogeneity indices and biomass indices derived from satellite imagery could improve the estimation of tree species diversity in ecosystems. However, a review of literature shows that species diversity as a function of both biomass and heterogeneity has not been largely tested.

1.2 Statement of the Problem

Although there has been increasing attempts to map vegetation species diversity within landscapes from remote sensing (Sanchez-Azofeifa et al., 2003; Turner et al. 2003; Goodwin et al., 2005; Cayuela et al., 2006; Saatchi et al., 2008), these have largely been based on unclear hypothetical frameworks thereby limiting prospects for the wider adoption of these methods in vegetation species diversity studies. For example, previous studies that used the standard deviation of NDVI as a measure of spatial heterogeneity (Foody and Cutler 2003; Feeley et al., 2005) and remotely sensed indices of biomass such as SAVI and NDVI (Oindo and Skidmore 2002; Gould 2003; Kerr and Ostrovsky 2003) to measure tree diversity made little reference to the spectral variation hypothesis and the biomass-diversity hypothesis respectively. Yet, the standard deviation of remotely sensed indices can be used within the framework of the spectral variation
hypothesis while SAVI and NDVI can be used within the framework of the biomass-diversity hypothesis.

In addition, the relationship between species diversity and satellite derived indices is well documented but there is little evidence to show that these relationships have been used for mapping biological diversity within landscapes (Nagendra 2001; Oindo and Skidmore 2002; Palmer et al. 2002; Michalet et al., 2006; Fargione et al. 2007; Rocchini et al. 2007; Gillespie et al. 2008). This has been either because of the weak relationships obtained (Oindo and Skidmore 2002; Feeley et al. 2005; Rocchini et al. 2007) or because linear regression, which is often used to describe these relationships, assumes randomness in the data yet most ecological phenomenon exhibit spatial autocorrelation (Anselin 2002). Particular attention has thus been paid on kriging and its various forms as mapping methods (Boschetti et al., 2007), although there is no clear evidence to suggest that any one approach is universally optimal (Dubrule 1984; Knotters et al., 1995). In this regard, the performance of linear regression as a mapping tool has largely been untested.

Also, it has not been widely statistically tested whether the results of a linear regression model are significantly different from those obtained using other methods such as spatial regression which take into account the spatial autocorrelation of a data set.

1.3: Thesis objectives

The objectives of this thesis were to (1) test whether and in what way tree species diversity is related to the standard deviation of the near infrared radiance and tree canopy cover estimated via the soil adjusted vegetation index (SAVI) in three selected savanna woodlands of Zimbabwe (2) model the spatial distribution of tree species diversity in savanna woodlands; and (3) compare the performance of linear and spatial regression methods for predicting tree species diversity. We based our study on the hypothesis that
variance in remotely-sensed images is directly related with tree species diversity in savanna woodlands.

1.5 Justification of the study
Natural woodlands in many parts of rural Africa provide many products and services essential for the well being of different communities. They provide wood for energy, timber and fencing, among a variety of uses. They also fulfil a number of ecological functions which include provision of wildlife habitat, fertilising the soil and cleaning the air by absorbing carbon dioxide and releasing oxygen (Chenje et al., 1998). The importance of woody forest resources therefore warrants the need to obtain a large amount of spatially defined data for the purposes of conservation.

Biodiversity inventorying and monitoring are at the core of the conservation, as well as sustainable use of forests. These are provided for under Article 17 of the Convention on Biological Diversity (CBD), which spells out the need to identify and monitor various components of biodiversity. Article 7b of the Convention on Biological Diversity (CBD) also calls for the monitoring of components of biological diversity. In this study, we therefore seek to develop relatively cheaper and faster techniques which can be used for the effective monitoring of tree species diversity. In addition, article 8c calls for the regulation and management of biological resources which are important for the maintenance of biological diversity either within or outside protected areas with a view to provide for their sustainable utilization. There is therefore need to provide quantitative data which will assist in the sustainable conservation and utilisation of biological resources. This study aims at providing a remote sensing method which can be used for the monitoring of tree species diversity at a landscape level. Remote sensing, despite its potential advantages, has not been extensively used for the monitoring of tree species diversity.
1.7. Organisation of the thesis

This thesis is divided into six chapters. Chapter 1 is an outline of the main objectives and hypotheses pursued in this study. It also provides a background to the tree species measuring methods.

Chapter 2 is a review of literature on the major determinants of tree species diversity in the savanna, the major causes of its loss and the consequences of such a loss. It also gives an overview of the methods used in the inventorying and measuring of tree species diversity.

Chapter 3 is an outline of the methods used in the collection and analysis of data. This chapter also explains how we used regression analyses to relate tree species diversity to satellite derived indices. It further explains how the generated relationships were used to spatially predict tree species diversity in a landscape from satellite imagery.

Chapter 4 presents the results of the relationship between tree species diversity and the satellite imagery derived indices, as well as a discussion of these results.

Chapter 5 is a presentation of the results and discussion on the mapping of tree species diversity using least squares linear regression, as well as conditional autoregressive models.

Chapter 6 is a general conclusion that is drawn from both the relationship between tree species diversity and the satellite imagery derived indices, as well as the mapping of tree species using the relationships derived using regression analyses.
Chapter 2: Determinants of and consequences of loss of tree diversity in savanna woodlands

2.1. Major determinants of woody species distribution in the Savanna

Savannas occur where trees and grasses interact to create a biome that is neither grassland nor forest. Thus the common ecological usage of the term savanna denotes landscapes with a continuous grass layer and scattered trees (Scholes and Archer 1997). Soil moisture availability and soil nutrient status are the key factors affecting both the balance between grasses and woody plants and the patterns of primary production and plant quality. These in turn influence the kinds and extent of herbivory, associated animal impacts, and the frequency and intensity of fires (Campbell et al., 1996). As a result, moisture and nutrient availability are largely treated as the primary determinants of savanna ecosystems, while disturbance regimes (fire and herbivory) are treated as secondary factors (Scholes and Walker 1993). However, the observed spatial patterns and relative abundances of grasses and woody plants in most savannas are dictated by the complex and dynamic interactions among climate, topography, soils, geomorphology, herbivory, and fire (Scholes and Archer 1997; Sankaran et al., 2005). These interactions may be synergistic or antagonistic and may reflect stochastic variation or positive feedbacks (Scholes and Archer 1997).

In addition savanna vegetation has been influenced by prehistoric, historic, or recent human activities. Humans can affect savanna structure either directly, as wood-cutters and cultivators, or indirectly, through the ability to manipulate fire and to influence herbivore numbers and distribution, both by hunting and by introducing and managing domestic animals (Frost et al., 1985; Cole 1986; Scholes and Archer 1997; Sankaran et al. 2005). Since the turn of the 20th century, humans have developed the capacity to bring
about rapid and considerable change in savanna structure through mechanical and chemical means. Thus, the interactions between “natural” and anthropogenic factors make it difficult to identify, isolate, or quantify the key determinants of savanna structure.

2.2 The importance of tree species diversity

The major motivation for worrying about tree species diversity is its centrality to economic as well as ecological activities (Campbell et al. 1996). Tree species supply entire suites of environmental services that can be defined as functional attributes of natural ecosystems that are beneficial to humankind (Daily et al., 1997). These include generating and maintaining soils, converting solar energy into plant tissue, sustaining hydrological cycles, storing and cycling essential nutrients, supplying clean air and water, and running biogeochemical cycles (Miller 1991).

Furthermore, the presence of tree species diversity provides an ecosystem the resilience necessary to cope with periodic stresses of the environment such as drought and climate change (Qin et al. 2003). While the links between genetic, species and ecosystem diversity and ecological function are seldom obvious and are not well understood, it is known that species/community losses at one level can induce serious cascading effects which compromises the ecosystem’s ability to provide services.

A number of livelihoods and economic activities depend on free and open access to a great variety of biological resources for food, fuel, medicines, housing materials and economic security (Campbell et al. 1996). Due to the fact that the protection of biodiversity in general and tree species diversity in particular is necessary for the maintenance of the biological resource base, tree species diversity conservation is understood by rural communities in developing nations.
2.3 Factors contributing to tree diversity loss

Several factors explain the loss of tree species diversity. These factors can be classified into two broad groups: natural causes and anthropogenic causes. The natural causes include scale insects, die-backs and climate variability and global climate change (Chenje et al. 1998; MMET 1998). Scale insects such as *Aspidoproctus glaber* Lindergrer (*Homptera margarodidae*) (Chenje et al. 1998; MMET 1998) attach themselves to tree stems and camouflage themselves by mimicking bark colour and form; and in some instances the spine of Acacia species. These sap sucking insects attack twigs, branches and the main stems of a number of miombo trees with preferred ones being *Brachystegia*, *Julbernadia* and *Acacia* species. Affected trees show defoliation followed by tip die back, bark splitting, die back of branches and stems and eventual death of the whole tree (Chenje et al. 1998; MMET 1998).

In the face of climate change, the changing environmental activities call for a change in the characteristics of the tree species, an example where genetic diversity becomes very crucial. This is important since it determines whether organisms will adapt to changing environmental conditions and consequently evolve into new forms or they go extinct.

Until the turn of the 20th century, human activity had little impact on the forests, with the earlier cultures engaging mainly in hunting and gathering (Archibald 1995). Living in widely dispersed groups, these societies did little to alter these forest ecosystems. However, the contemporary rapid population growth has led to the increase in the pressure on land, as the communities are mainly subsistence farmers. The opening up of forestland for agricultural expansion is the major reason for the loss of forest biodiversity as well as habitat fragmentation (Campbell et al. 1996; Clarke et al., 1996; Chenje et al. 1998). The remaining species rich, semi-natural woodland islands are usually considered
uneconomic and left fragmented. Furthermore, depletion of tree species diversity in these woodlands occurs because they would be lacking in management.

2.4 Consequences of tree species diversity loss

Tree species are both an economic asset at country level, as well as a source of livelihood at community level, and their loss entails problems for human survival. At community level, losses in tree species diversity result in a corresponding decline in the quality and quantity of services such as fuel-wood, thatching grass, fruit, honey and mushrooms. Also, trees provide habitats for a number of other animal and insect species, therefore a loss in tree diversity results in a loss of habitat diversity (Clarke et al. 1996; Chenje et al. 1998). At the national level and beyond, tree species present future areas of research for, chief among them, the drug and chemical industry. A loss in tree species diversity thus represents a loss in potential areas of research for a number of industries. Climatically, trees play a crucial role in carbon sequestration (Burroughs 2001), and their loss will result in a continual build up of carbon dioxide (CO₂), a greenhouse gas, in the atmosphere. The removal of trees implies a removal of the major sink of carbon dioxide hence the gradual build up of the gas in the atmosphere.

2.5 Remote sensing methods of predicting tree species diversity

The availability of satellite images at various spatial resolutions has resulted in a number of methods designed to determine spatial variations in species diversity (Sanchez-Azofel FA et al. 2003; Turner et al. 2003; Goodwin et al. 2005; Cayuela et al. 2006; Rocchini et al. 2007; Saatchi et al. 2008). Three main methods that use satellite images to determine tree species diversity can be recognised from literature, namely the image-based classification, the habitat-based mapping and the use of satellite indices as surrogates of diversity. The image-based classification method is regarded as a direct
mapping of species (Nagendra 2001; Gillespie et al. 2008) and it utilises mainly hyperspectral satellite imagery (Schmidt 2003; Boschetti et al. 2007; Rocchini et al. 2007). The method has been found to be mostly applicable in ecosystems that consist of relatively homogeneous vegetation stands (Nagendra 2001; Schmidt 2003). Studies that have employed the direct method to determine vegetation species distribution have had relatively high accuracy levels of between 65% and 80%. However, apart from the difficulties experienced in distinguishing species which have similar growth forms, the classification method often yields inaccurate results in heterogeneous environments (Bradley and Fleishman 2008). In addition, the high cost of acquiring hyperspectral data makes it very expensive to continuously monitor changes in tree species distribution at any scale, over shorter time periods.

Next, the habitat-based method determines species diversity based on their habitats (Nagendra 2001; Turner et al. 2003). However, one main limitation of the habitat method of determining species diversity is the failure to accurately represent diversity especially where species composition varies both within a habitat and between habitats. Thus, the determination of species diversity using the habitat method has often been limited to scales of tens of kilometres (Roughgarden et al., 1991; Nagendra 2001; Turner et al. 2003; Rocchini et al. 2007; Gillespie et al. 2008).

There is now a growing interest among ecologists to use remotely-sensed vegetation indices when modelling species distributions in the landscape (Nagendra 2001; Turner et al. 2003; Seto et al., 2004; Gillespie et al. 2008). To this end, remotely-sensed measures such as the standard deviation or the coefficient of variation in NDVI have been related to underlying species diversity in the landscape (Nagendra 2001; Oindo and Skidmore 2002; Levin et al., 2007; Gillespie et al. 2008).
2.6: Mapping tree diversity patterns using linear and spatial regression analyses

There have been very few attempts that have been made to map the spatial variations in tree species diversity using linear regression models, even though significant relationships existed between diversity indices and satellite information (Roughgarden et al. 1991; Stoms and Estes 1993; Nagendra 2001; Oindo and Skidmore 2002; Turner et al. 2003; Rocchini et al. 2007; Gillespie et al. 2008). These studies were mainly focussed on describing the nature of the relationship that exists between tree species diversity and indices calculated from satellite images. Furthermore, linear regression analysis has not been widely used for mapping ecological data. Ecological data, which is mainly spatially auto-correlated, violates the assumption of independence, a precondition for the application of linear regression analysis (Fortin and Dale 2003; Dormann et al., 2007; Kissling and Carl 2008). Studies have thus mainly been dependent on direct interpolation methods. The major problem with these methods is that they assume that distance from observed values is the most important predictor of the variable of interest, and these methods have very limited capabilities to include environmental covariates as predictors (Fortin and Dale 2003; Franklin 2009). In addition, there has been no direct comparison of the performance of linear and spatial autocorrelation-corrected methods in predicting tree species diversity patterns.
Chapter 3: Methods

3.1: Study sites

The study was carried out in four study sites: Mukuvisi Woodlands, Mabalauta, which is part of Gonarezhou National Park, Sebungwe Region and Kyle Game Reserve, all in Zimbabwe (fig. 1). We chose these study areas because they follow a gradient of rainfall intensity from high (800 mm in Mukuvisi) through medium (635 mm in Kyle) to low (400 mm in Mabalauta). The Sebungwe study site was chosen because it has physical characteristics that are comparable to those from Mabalauta study site. Tree species diversity data from Mukuvisi, Mabalauta and Kyle study sites were used for developing regression models while data from the Sebungwe study site data was used for validating the regression models that were developed in Mabalauta study site.
Figure 3.1: The location of the four woodland study sites in Zimbabwe.
Table 3.1: A description of the four study sites. The soil data was extracted from (Nyamapfene 1991)

<table>
<thead>
<tr>
<th>Study site</th>
<th>Climate data</th>
<th>Soils</th>
<th>Vegetation data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mukuvisi woodlands (265 ha). It lies between 17°49'S and 17°51'S, as well as longitudes 31°04'E-31°06'E.</td>
<td>Mean annual temperature of 18°C and between 800 and 1000 mm of rainfall (Mukuvisi rainfall data)</td>
<td>Ferallic cambisols</td>
<td>Miombo woodlands dominated by Julbernadia globiflora, with Burkea Africana, Parinari curatellifolia and Brachystegia spiciformis as co-dominants</td>
</tr>
<tr>
<td>Kyle Game reserve (7 600 ha). It lies between 20°04' S and 20°14' S, and 31°07' E and 30°50' E.</td>
<td>Mean annual temperature of 20°C and around 635 mm of rainfall (Masocha 2009)</td>
<td>Chromic-leptic lixisols and chromic luvisols</td>
<td>Miombo woodlands dominated by J. globiflora and B. spiciformis, bushland dominated by Peltophorum africanum and Acacia karoo. Grasslands are dominated by Hyparrhenia fillipendula, Themeda triandra and Hyperthelia dissolute (Vincent and Thomas 1960).</td>
</tr>
<tr>
<td>Mabalauta (87 200ha). lies between 21°44'S and 21°43' S, and 31°18'E and 31°19' E</td>
<td>Mean annual temperature of 27°C and between 300 and 500 mm of rainfall (Goodwin 1997)</td>
<td>Ferralic arenosols and chromic luvisols</td>
<td>Colophospermum mopane in association with Combretum and Terminalia species (<a href="http://www.zimparks.org">www.zimparks.org</a> )</td>
</tr>
<tr>
<td>Sebungwe (1500000 ha of which 341 000 ha are mopane woodlands) 16°44' S and 18°11' S, and 27°21' E and 29°08' E</td>
<td>Mean annual temperature of 27°C and between 500 to 600 mm rainfall</td>
<td>Rhodic nitisols, gleic solonetz and haplic arenesols</td>
<td>Woodlands are divided into C. mopane, Faidherbia albida and J. globiflora and B. spiciformis dominated. Grasslands are dominated by Setaria incrassata, Ischaemum afrum and Dicathium papillosum (Timberlake et al., 1993).</td>
</tr>
</tbody>
</table>
3.2. Remotely sensed data

We used Advanced Spaceborne Thermal Emmission and Reflection Radiometer (ASTER) imagery, as well as IKONOS imagery in this study (fig 3.2). The former were obtained from an online source www.glovis.usgs.gov, while the latter, with an exception from Mabalauta, was based on Google Earth (www.earth.google.com). In Mabalauta, an IKONOS image was obtained from the online database www.landinfo.com. ASTER satellite imagery was used to (1) classify Kyle and Mabalauta study sites, (2) estimate tree cover via a Soil Adjusted Vegetation Index (SAVI) in Kyle, Mabalauta, Mukuvisi and Sebungwe study sites and (3) calculate an index of diversity of satellite imagery. Using the maximum likelihood classifier algorithm, we classified Kyle and Mabalauta study sites into different land cover types that include grassland, water, wooded grassland and woodland at an accuracy of 85 % and 80 % respectively. For our remote sensing indices of diversity, we calculated the standard deviation of NIR radiance in Mukuvisi, Kyle and Mabalauta study sites, as well as the coefficient of variation of NIR radiance in Mabalauta and Sebungwe study sites. IKONOS images were used to (1) classify, by visual interpretation, Mukuvisi study site into different grassland, wooded grassland and woodland, as well as (2) for estimating the amount of tree canopy cover as a percentage of ground covered by aerial projection of trees in Mukuvisi, Kyle and Mabalauta study sites.
Figure 3.2: ASTER satellite images for a) Mukuvisi, b) Kyle and c) Mabalauta study sites showing the boundary of the study site (red line) as well as the sampling points (red diamonds). Satellite images d) and e) cover part of the Sebungwe study site (red diamonds show the sampling points). Vegetated areas are shown by the green colour, and the map units are Geographic coordinates.
ASTER imagery used in this study was acquired on the 11th of August 2006 for Kyle Game reserve, the 10th of August 2006 for Mukuvisi Woodlands, the 5th of August 2006 for Mabalauta, Gonarezhou National Park, and 23rd August 2006 for Sebungwe region. The ASTER images were georeferenced using at least 10 ground control points from scanned topographic maps, to an accuracy of at most 0.500 pixels (7.5 m). These ground control points included points such as road junctions and river intersections. The ASTER images, apart from being readily available, were also appropriate because during this time of the year the tree species would still be in leaf while grass would have senesced for all the study sites (Chidumayo 1993). The IKONOS images for Mukuvisi and Kyle were based on google earth (www.earth.google.com) and were acquired in 2007. In Mabalauta, the IKONOS image used was acquired on the 22nd of April 2006, and was georeferenced to an accuracy of 0.5 pixels (2m) using ground control points that included road junctions as well as river intersections. The IKONOS image for Mabalauta, as well as ASTER imagery for the four study sites were first radiometrically corrected by a dark object subtraction, which involves subtracting a constant digital number (DN) value from the entire image to reduce atmospheric effects (Chavez 1988). Absolute image correction is recommended, especially for image mosaicking (Cihlar 2000) or change detection (Oetter et al., 2001; Du et al., 2002), but these were not relevant in this study.

To estimate a remotely-sensed diversity index, we calculated the standard deviation of radiance values of the Near Infrared (stdev NIR) band for the Kyle, Mabalauta and Mukuvisi ASTER images, as well as the coefficient of variation of radiance for Mabalauta and Sebungwe images in ILWIS GIS, version 3.3. Both the standard deviation and the coefficient of variation were calculated at a window size of 45 m by 45 m, the smallest size at which we can calculate diversity indices from an ASTER satellite image. We selected the NIR band because green vegetation returns the strongest signal in this
band and is useful for differentiating species that differ in their foliage content (Nixon et al., 1985; Taylor 1993). We chose the standard deviation because it has been used as an index for estimating diversity using remotely-sensed data (Nagendra 2001; Oindo and Skidmore 2002; Feeley et al. 2005).

Next, we estimated the amount of tree canopy cover by first digitising the area covered by tree canopies from an IKONOS image of 2007 in Mukurvisi Woodlands, an IKONOS image of 2007 in Kyle, both based on Google Earth (Knorn et al., 2009) and the downloaded IKONOS image of 22nd April 2006 in Mabalauta. We then estimated tree canopy cover by calculating a ratio of the area of the digitised tree canopies to the area of the sampling plot (900 m² for all the three study sites). Next, we calculated the Soil Adjusted Vegetation Index (SAVI), as a proxy for tree cover, from ASTER as follows:

\[
SAVI = \left( \frac{NIR - R}{NIR + R + L} \right) \times 1 + L
\]

where \( NIR \) is radiance in the Near Infrared band, \( R \) is radiance in the Red band, \( L \) is a soil correction factor. In our case, we used a value of \( L= 0.5 \), which is the value taken to be appropriate for areas that have medium vegetation such as the Southern African savanna ecosystems (Huete 1988; Gilabert et al., 2002).

We used SAVI values averaged at 45 m to match the scale at which both the standard deviation and the coefficient of variation of NIR radiance were calculated. SAVI has been documented as a measure of biomass and vegetation cover which allows for the correction of the influence of the soil background (Huete 1988; Huete et al. 1997). In this study we evaluated whether we can confidently use SAVI as a proxy for tree cover by correlating SAVI with tree canopy cover that we estimated using IKONOS high spatial resolution imagery. Since our data was normally distributed (based on the Kolmogorov-
Smirnov test of normality), we used Pearson’s correlation coefficient to test for the strength of the relationship between SAVI and canopy cover.

3.3. Tree species data

3.3.1 Sampling and tree species data collection
We selected areas of woodland in a GIS from the classified images for each study site (see section 2.2), except for Sebungwe where stratification was based on a 1996 floral map obtained from the Forestry Commission, Zimbabwe. Next, we created random points in the woodlands of each of the study sites using the random point generator in ArcView GIS (ESRI, 1992-1998). We generated 80, 60, 80 and 80 random points in Mukuvisi, Kyle, Mabalauta and Sebungwe study sites respectively. These points were at least 70 m apart to avoid the problem of overlapping sampling points. In the field, we managed to sample 39, 30, 37 and 33 random points in Mukuvisi, Mabalauta, Kyle and Sebungwe study sites respectively. Mukuvisi study site is made up of undulating terrain and therefore the random points were easily accessible. In addition, Mukuvisi was the smallest of the four study sites therefore the total cost of accessing the random points was lowest. In Mabalauta and Sebungwe study sites, we sampled the least number of random points because the other points were inaccessible. The sparse road network, coupled with the rugged terrain in Mabalauta made some of the generated random points inaccessible. In Sebungwe study site, the sparse road network also rendered a number of the random points inaccessible (fig 3.3). Also, in terms of area, Mabalauta and Sebungwe study sites are the largest of the four study sites therefore the total cost of accessing all the points was very high.
Figure 3.3: Distribution of the sampling points in relation to roads in Sebungwe and Mabalauta study sites.

Tree species data were collected in March 2007 for Mukvisi, September 2007 for Kyle, late October to early November 2007 for Mabalauta and April to May 2001 for Sebungwe study sites. During these different times of the year, tree species were in leaf (in the case of Mukvisi, Sebungwe and Kyle study sites) or had already flushed (in the case of Mabalauta study site) and therefore were easily identifiable. The local or vernacular names for the tree species were recorded, and the corresponding botanical names were inferred from www.zimbabweflora.co.zw, a website containing floral species found in Zimbabwe. Tree species whose names were not immediately available were taken for identification by a botanist in the Biological Sciences Department at the University of Zimbabwe. In each case, the generated sample point locations became the centre of the sampling plots. In the field, we navigated to the generated point location using a handheld Garmin Etrex global positioning system (GPS) receiver at an accuracy of 10 m (Garmin Inc., 2000-2003). We then defined a sampling plot of 30 m by 30 m in all the study sites around each point location using a tape measure. Previous researches have shown that sampling plot sizes most used are 25-400 m² in tall shrub communities.
and 400-25000 m$^2$ for trees in woods and forests (Sutherland 1996), and our plot size of 900 m$^2$ falls within this range.

In each sampling plot we recorded tree species present which had a minimum diameter at breast height of 5 cm and a minimum height of 3 m in accordance with FAO (2005) recommendations for classifying trees. In addition, the tree species with a minimum of 5 cm diameter at breast height (dbh) and at least 3 m of height were observed to be dominating the amount of canopy cover and were therefore expected to also dominate the spectral signature recorded in each sampling plot.

### 3.3.2. Tree diversity indices

Simpson’s index of diversity (1-D) and the Shannon Weiner, commonly referred to as the Shannon Weaver’s index of diversity (H') were used as measures of tree species diversity in this study. Previous studies used species richness as a measure of diversity and related it to satellite derived indices (Roughgarden et al. 1991; Nagendra and Gadgil 1999; Oindo and Skidmore 2002; Seto et al. 2004; Rocchini et al. 2007). Species richness is simply the number of species in a plot, while evenness measures the proportions of these species in that plot. Diversity indices combine both species richness and evenness into a single value. Thus, diversity indices are preferred measures of describing an ecosystem’s diversity.

The Simpson’s Index, 1-D, gives the probability that two individuals drawn at random from a sample belong to different species. The index ranges from 0 (a single species) to 1 (perfectly evenly distributed species). A high value (closer to 1) of the index denotes a
greater probability of picking two different species, and therefore high species diversity.

The Simpson’s (1-D) index of diversity is calculated by:

\[ 1 - D = \sum_{i=1}^{n} P_i^2 \]  

(2)

where \( P_i \) is the proportion of the \( i^{th} \) species in the sampling plot, given by:

\[ P_i = \frac{n_i}{N}, \]  

(3)

In this equation \( n_i \) is the number of stems in each sampling plot belonging to the \( i^{th} \) species and \( N \) is the total number of stems in the sampling plot.

The Shannon Weaver Index (\( H' \)) is a measure of the average ‘uncertainty’ of predicting to what species an individual chosen at random from a collection of species belong. This average uncertainty increases as the number of species increases, and as the distribution of individuals among the species become more even. The Shannon Weaver index, \( H = 0 \) if and only if there is one species, and reaches a maximum when all species present are represented by the same number of individuals (perfect even distribution). It is calculated using the formula:

\[ H = \sum_{i=1}^{N} -p_i \log p_i \]  

(4)

In equation 4 \( p_i \) is the proportional abundance of the \( i^{th} \) species as calculated in Equation (3)

3.4. Relating tree species diversity indices to standard deviation of NIR and SAVI

We used linear regression analysis to relate the standard deviation of NIR (stdev NIR), as well as SAVI to tree species diversity. However, as a preamble to implementing linear regression analysis, we randomly split our data into two sets, i.e., 20% of the sample
points were set aside for validating the model by calculating the root mean square error (RMSE) in each of the study sites (8, 6 and 8 points in Kyle, Mabalauta and Mukuvisi study sites respectively). The remaining 80% was then used in regression analysis, which was accomplished in several steps.

Firstly, in order to fulfil the requirements of linear regression analysis, we tested whether our data followed a normal distribution using the Kolmogorov-Smirnov (K-S) test, as well as tested for spatial autocorrelation using Moran’s I. Thus the K-S test for normality was applied to SAVI, NIR and tree diversity data based on the 80% sample set aside for analysis. Moran’s I was used to investigate for spatial autocorrelation in Simpson’s 1-D index. The results of the normality tests showed that tree species diversity data from Mabalauta and Kyle study sites did not significantly (K-S statistic = 0.176, p = 0.106 for Mabalauta, K-S statistic = 0.173, p = 0.200 for Kyle) deviate from a normal distribution while tree species diversity data from Mukuvisi study site significantly (K-S statistic = 0.115, p = 0.000) deviated from a normal distribution. Thus, for the Mukuvisi data, we had to apply a log transformation to attain normality. Also, results from Moran’s I shows that the data was not spatially autocorrelated (Mukuvisi, I = -0.115, z = 0.501, p = 0.617; Kyle, I = -0.323, z = -1.178, p = 0.239; Mabalauta, I = 0.181, z = 1.092, p = 0.275).

Secondly in order to test whether (1) the standard deviation of the NIR (stdev NIR) band, (2) SAVI, and (3) SAVI together with the stdev NIR are significantly related with tree species diversity, we fitted the least squares linear regression model. However, prior to implementing the multiple regression model of SAVI and stdev NIR on tree species diversity, we tested whether the stdev NIR and SAVI were independent from each other using Pearson’s correlation. That predictor variables are independent is a key assumption of multiple linear regression.
Next, we checked whether the distribution pattern of residuals was random as a requirement of regression analysis. For this, we inspected a residual plot. Results showed that the residuals were randomly distributed. Thus, in this case, we concluded that the assumptions of linear regression were not violated.

3.5. Mapping the spatial variations in tree species diversity

For the purposes of investigating the utility of mapping tree species diversity as a function of remotely-sensed indices, we adopted two interpolation methods; least squares linear regression and spatial regression. We applied the linear regression models produced from the random samples that we generated for all the three study sites (see section 3.4). On the other hand, we also applied spatial regression models for those study sites where sampled tree species data showed evidence of spatial autocorrelation after test samples once excluded in section 3.4 were replaced. Linear regression assumes that samples are random while spatial regression is based on spatially correlated data. To test whether the tree diversity data was random or spatially autocorrelated, we calculated Moran’s index of spatial autocorrelation, $I$ (Moran 1950). The index, $I$, computes the degree of correlation between the values of a variable as a function of spatial lags. Moran’s $I$ varies from -1 (negative autocorrelation) to 1 (positive autocorrelation), with an expected value close to zero showing the absence of spatial autocorrelation (Moran 1950; Fortin and Dale 2003).

Spatial regression models based on the Conditional Autoregressive model (CAR) were developed and applied for Mukuvisi and Mabalauta study sites only. The CAR was adopted because it accounts for both broad scale trend (via inclusion of trend surface
terms) and fine scale autocorrelation (via the correlated error structure). In this regard, the CAR has generally been the recommended spatial modelling technique (Lichstein et al., 2002).

3.6: Comparison of the linear and spatial regression predictions

We investigated the utility of remote sensing indices in mapping tree species diversity in three ways. Firstly, we determined the accuracy of mapping tree species diversity using the linear regression model by calculating the root mean square error in the three study sites (Moore and McCabe 1998). A test data set already selected (see Section 3.4) was used for calculating the root mean square error in Mukuvisi, Kyle and Mabalauta study sites. In this test, a small value of the root mean square error would mean higher accuracy compared with a large value of the root mean square error.

Secondly, we tested whether the linear and spatial regression models can be successfully used in mapping tree species diversity in an independent study site with similar physical characteristics. But first, we had to develop a model that can be successfully transferred from one study site to another. Since the effects of the atmosphere were not completely corrected for in this study (Section 3.2), we adopted the coefficient of variation as a measure of variation in NIR radiance for Mabalauta. This enabled us to develop a model that can be used to predict tree species diversity in Sebungwe. The co-efficient of variation normalises deviations in NIR radiance in each of the two data sets, thus making them comparable against each other (Moore and McCabe 1998). Using the co-efficient of variation of NIR radiance and SAVI images from Sebungwe, we used both linear and spatial regression models to map tree species diversity in the Sebungwe region. Finally, after extracting spatial and linear regression mapped diversity values from our sample points, we calculated the root mean square error to compare the prediction accuracies of
the linear regression and spatial regression models. We also used the Student’s t-test (Moore and McCabe 1998) to test whether there were any significant differences in tree species diversity values mapped by the linear and spatial regression models in Sebungwe study site.

Finally, we also used Pearson’s correlation coefficient to check for consistency in the tree species diversity mapped by the linear and spatial regression models. In a GIS, we randomly generated 100 points in Mabalauta and 50 points in Mukuvisi study sites. We then extracted tree species diversity values from the maps produced by linear as well as spatial regression models.
Results and Discussion
Chapter 4: The relationship between tree species diversity and satellite derived indices

4.1: Results

We found a significant ($p < 0.001$) relationship between tree canopy cover and SAVI (fig. 4.1). Thus, we concluded that SAVI can be used as a surrogate for tree canopy cover or as an estimate of biomass in the three study sites.

Figure 4.1: Relationships between tree canopy cover and the Soil Adjusted Vegetation Index (SAVI) in a) Mukuvisi, b) Kyle and c) Mabalauta study sites of Zimbabwe.

Also, we observed that the standard deviation of NIR radiance and SAVI were not significantly ($r = -0.194, p = 0.943$ for Mukuvisi, $r = 0.233, p = 0.330$ for Mabalauta and $r = 0.314, p = 0.177$) correlated. This means that the two variables could be used in the same regression model without introducing redundancy.
Figure 4.2 illustrates that there is a significant ($p < 0.05$) relationship between tree species diversity (estimated by both Simpson’s $D$ and Shannon’s $H$ indices) and the standard deviation of NIR radiance in Mukuvisi, Kyle, as well as Mabalauta study sites. We chose the better fitting models, as well as those that returned significant relationships between tree species diversity and the satellite derived indices. The standard deviation of NIR radiance significantly explained at most 35%, 38% and 54% of the variations in tree species diversity in Mukuvisi, Kyle and Mabalauta study sites respectively (fig. 4.2). In Kyle and Mabalauta study sites, tree species diversity decreases with an increase in the standard deviation of the NIR (stdev NIR) band. In these two study sites, high tree species diversity values coincide with low values of the stdev NIR radiance, whilst low diversity values occur at high values of the stdev NIR radiance (fig. 4.2). In addition, in Kyle study site, tree species diversity starts to increase at very high values of the stdev NIR radiance. In Mukuvisi woodlands a hump-shaped relationship exists between tree species diversity and the standard deviation of the NIR band. In this study site, we observe that low diversity values coincide with both low and high values of the stdev NIR radiance, while the highest values of diversity occur at intermediate values of the stdev NIR radiance (fig. 4.2).
Figure 4.2: Relationships between the standard deviation of NIR radiance and a) Simpson’s 1-D and b) Shannon’s H in 1) Mukuvisi, 2) Kyle and 3) Mabalauta study sites. The standard deviation of the NIR radiance was calculated at 45m.

Tree species diversity and SAVI show a hump-shaped relationship in Mukuvisi and Kyle study sites, and a negative linear relationship in Mabalauta study site (fig 4.3). In Mukuvisi and Kyle study sites, diversity values initially increase with an increase in SAVI reach a peak at intermediate SAVI and decline with a further increase in SAVI. In Mabalauta, tree species diversity shows a decrease with an increase in SAVI.
**Figure 4.3:** Relationships between SAVI and a) Simpson’s 1-D, and b) Shannon’s H in 1) Mukuvisi, 2) Kyle and 3) Mabalauta study sites. SAVI was averaged over a 45 m window.

Overall, in all the three study sites, low tree species diversity occurs at high values of SAVI, while the highest diversity occurs at low to intermediate SAVI values.

Also, tree species diversity is a significant function of the standard deviation of NIR radiance and SAVI in Mukuvisi, Kyle, as well as Mabalauta study sites (fig. 4.4). We
observe that in combination, the standard deviation of NIR radiance and SAVI values explain between 56% and 70% of the variation in tree species diversity.
Figure 4.4: Relationships between tree species diversity estimated by a) Simpson’s 1-D and b) Shannon’s H and a combination of the standard deviation of the NIR radiance and estimated by SAVI in 1) Mukuvisi, 2) Kyle and 3) Mabalauta study sites. The graphs
show decreasing tree species diversity from highest diversity (deep red) to lowest diversity (dark green).

Moreover, several observations can be made on the way tree species diversity relates to the standard deviation of NIR radiance and SAVI in all the three study sites. Firstly, it can be observed that high diversity values coincide with low or high values of the stdev NIR radiance coinciding with intermediate SAVI values (fig. 4.4). Secondly, we observe that intermediate diversity values occur when medium to high standard deviation of NIR radiance values interdigitate with intermediate SAVI values. Finally, we observe that low tree species diversity occurs where low values of stdev NIR radiance are in association with high or low SAVI values.

Sample plots that have low or high values of the standard deviation of NIR radiance in combination with intermediate SAVI values do not show dominance of any of the main species, but are associated with high tree species diversity in the three study sites (Table 4.1). For example in Kyle study site these sampling plots are characterised by high tree species diversity that show less dominance of the main species, i.e., *Brachystegia spiciformis* and *Julbernardia globiflora* (Table 4.1a). In Mukuvisi study site, the sampling plots are also characterised by the co-occurrence of several species that include *Parinari curatellifolia, Peltophorum africanum, Burkea Africana, Terminalia sericea* (Table 4.1c). In Mabalauta study site, the sampling plots have several species such as *Terminalia sericea, Pterocarpus angolensis,* and *Guibourtia conjugata,* resulting in high tree species diversity (Table 4.1b).
Table 4.1: The distribution of tree species and the associated diversity and satellite imagery derived indices in some of the sampling plots.

<table>
<thead>
<tr>
<th>Species</th>
<th>PI</th>
<th>SI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Uapaca kirkiana</td>
<td>0.003</td>
<td></td>
</tr>
<tr>
<td>Julbernardia globiflora</td>
<td>0.594</td>
<td>0.926</td>
</tr>
<tr>
<td>Brachystegia spiciformis</td>
<td>0.307</td>
<td>0.055</td>
</tr>
<tr>
<td>Burkea africana</td>
<td>0.255</td>
<td>0.012</td>
</tr>
<tr>
<td>Erythrina abyssinica</td>
<td>0.106</td>
<td>0.002</td>
</tr>
<tr>
<td>Parinari curatellifolia</td>
<td>0.106</td>
<td>0.002</td>
</tr>
<tr>
<td>Celtis africana</td>
<td>0.005</td>
<td></td>
</tr>
<tr>
<td>Combretum apiculatum</td>
<td>0.064</td>
<td>0.024</td>
</tr>
<tr>
<td>Dombeya rotundifolia</td>
<td>0.064</td>
<td>0.005</td>
</tr>
<tr>
<td>Lannea discolor</td>
<td>0.002</td>
<td></td>
</tr>
<tr>
<td>Ochna pulchra</td>
<td>0.004</td>
<td></td>
</tr>
<tr>
<td>Strychnos spinosa</td>
<td>0.043</td>
<td>0.002</td>
</tr>
<tr>
<td>Peltophorum africanum</td>
<td>0.021</td>
<td></td>
</tr>
<tr>
<td>Cussonia arborea</td>
<td>0.043</td>
<td></td>
</tr>
<tr>
<td>Dichrostachys cinerea</td>
<td>0.064</td>
<td></td>
</tr>
<tr>
<td>Terminalia sericia</td>
<td>0.191</td>
<td></td>
</tr>
<tr>
<td>Vitex payos</td>
<td>0.021</td>
<td></td>
</tr>
<tr>
<td>Ziziphus mucronata</td>
<td>0.021</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Species</th>
<th>PI</th>
<th>SI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Combretum fragrans</td>
<td>0.111</td>
<td></td>
</tr>
<tr>
<td>Pterocarpus angolensis</td>
<td>0.127</td>
<td>0.024</td>
</tr>
<tr>
<td>Guibourtia conjugata</td>
<td>0.143</td>
<td></td>
</tr>
<tr>
<td>Xeroderris stuhlmannii</td>
<td>0.063</td>
<td></td>
</tr>
<tr>
<td>Terminalia sericea</td>
<td>0.175</td>
<td>0.087</td>
</tr>
<tr>
<td>Strychnos madagascariensis</td>
<td>0.381</td>
<td>0.405</td>
</tr>
<tr>
<td>Combretum molle</td>
<td>0.444</td>
<td></td>
</tr>
<tr>
<td>Combretum apiculatum</td>
<td>0.04</td>
<td></td>
</tr>
<tr>
<td>Colophospermum mopane</td>
<td>0.732</td>
<td></td>
</tr>
<tr>
<td>Combretum hereoense</td>
<td>0.152</td>
<td></td>
</tr>
<tr>
<td>Flueggea virosa</td>
<td>0.03</td>
<td></td>
</tr>
<tr>
<td>Acacia sp</td>
<td>0.037</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Species</th>
<th>PI</th>
<th>SI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lippia javanica</td>
<td>0.029</td>
<td></td>
</tr>
<tr>
<td>Pavetta schummaniana</td>
<td>0.029</td>
<td></td>
</tr>
<tr>
<td>Rhoicissus tridentata</td>
<td>0.029</td>
<td></td>
</tr>
<tr>
<td>Vitex mombassae</td>
<td>0.029</td>
<td></td>
</tr>
<tr>
<td>Psydrax livida</td>
<td>0.057</td>
<td></td>
</tr>
<tr>
<td>Sclerocarya birrea</td>
<td>0.057</td>
<td></td>
</tr>
<tr>
<td>Grewia monticola</td>
<td>0.086</td>
<td></td>
</tr>
<tr>
<td>Dichrostachys cinerea</td>
<td>0.229</td>
<td></td>
</tr>
<tr>
<td>Combretum molle</td>
<td>0.029</td>
<td>0.083</td>
</tr>
<tr>
<td>Julbernardia globiflora</td>
<td>0.429</td>
<td>0.4</td>
</tr>
<tr>
<td>Brachylaena rotundifolia</td>
<td>0.042</td>
<td></td>
</tr>
<tr>
<td>Lannea discolor</td>
<td>0.042</td>
<td></td>
</tr>
<tr>
<td>Brachystegia spiciformis</td>
<td>0.433</td>
<td></td>
</tr>
<tr>
<td>Pavetta schummaniana</td>
<td>0.133</td>
<td></td>
</tr>
<tr>
<td>Ozoroa reticulata</td>
<td>0.033</td>
<td></td>
</tr>
</tbody>
</table>
Moreover, it can also be observed that sample plots that have medium to high stdev NIR radiance values coinciding with intermediate SAVI values are marked by co-dominance of around 2 of the major tree species (Table 4.1). For instance, in Kyle and Mukuvisi study sites, intermediate values of diversity occur where there is a co-dominance of *Julbernardia globiflora* and *Brachystegia spiciformis* (Tables 4.1a and 4.1c). The same pattern is observed in Mabalauta study site where sampling plots that have intermediate tree species diversity show a co-dominance of *Colophospermum mopane* and *Combretum imberbe* or *Strychnos madagascariensis* and *Combretum molle* (Table 4.1b).

Results also show that sample plots that show low standard deviation of NIR values and high SAVI values and vice versa are associated with the dominance of a single tree species in all the three study sites (Table 4.1). For instance, in Kyle and Mukuvisi study sites, low tree species diversity occurs where there is dominance of *Julbernardia globiflora* (Tables 4.1a and 4.1c) while in Mabalauta, study site, low tree species diversity occurs where there is a dominance of *Colophospermum mopane* (Table 4.1b).

### 4.2: Discussion

In this chapter, we address the question of whether and in what form remotely-sensed NIR standard deviation, as well as SAVI are related with tree species diversity. In other words, we try to understand whether the standard deviation of NIR can be used to explain tree species diversity within the SVH hypothesis, while on the other hand testing whether SAVI can be used to explain tree species diversity within the biomass-diversity hypothetical framework. Our results show that tree species diversity has a hump-shaped response to SAVI, an indicator of biomass (Huete 1988). Separate studies based on experimental plots have also documented a humped relationship between species
diversity and biomass in the long term, while a positive asymptotic linear relationship has also been noted to exist between species diversity and biomass in newly seeded grassland communities (van Ruijven and Berendse 2005; Fargione et al. 2007). The humped relationship between species diversity and biomass has mainly been attributed to mature vegetation types because in these ecosystems the complementarity effect would be very effective (Fargione et al. 2007). Thus, we deduce that our results are consistent with the biomass-biodiversity hump-shaped relationship hypothesis.

The hump-shaped relationship between SAVI and tree species diversity found in this study is consistent with the predictions of the complementarity hypothesis. This hypothesis predicts that at low biomass there is less complementarity which results in low diversity while at intermediate biomass, there is high diversity resulting from the complementarity interactions (Michalet et al. 2006; Fargione et al. 2007). Complementarity results when ecological differences among species lead to more complete utilisation of the resources, and therefore allows for different species to co-exist hence high diversity (Qin et al. 2003). In other words, intermediate biomass allows niche differentiation (high complementarity) amongst species, leading to high species diversity. At high biomass, species diversity is low due to competitive exclusion, or because the environment is limiting. Based on our results, we therefore deduce that inclusion of remotely-sensed SAVI, an indicator of canopy cover and biomass, improves the prediction of tree species diversity in savanna woodlands. It is however important to note that in Mabalauta, there is a linear relationship between biomass and tree species diversity. This might be due to a sampling anomaly where we would have sampled only that part of the ecosystem with higher biomass. This could also have been avoided if we
had further stratified the study sites according to the amount of biomass, i.e. high, medium and low biomass.

Our results also indicate that the standard deviation of NIR is significantly related with tree species diversity. This is consistent with the Spectral Variation Hypothesis (SVH) which states that spectral heterogeneity is related to spatial (ecosystem) heterogeneity and thus to species richness and diversity (Palmer et al. 2002). However, in this case, high species diversity is found at low values of the standard deviation, while low species diversity is found at high standard deviation of NIR radiance. The results contradict the findings of other studies that found a significant positive relationships between spectral heterogeneity and plant species richness (Oindo and Skidmore 2002; Rocchini et al. 2007).

Analysis of the results of our model indicate that tree species diversity in savanna woodlands can be successfully predicted by remote sensing methods when a combination of the standard deviation of NIR radiance and tree canopy cover estimated using SAVI is employed (fig. 4.4). In this relationship tree species diversity is high at intermediate canopy cover and decreases towards both low and high canopy cover. In their work on the Kenyan savanna landscape, Oindo and Skidmore (2002) also found a humped relationship between NDVI and plant species richness in Kenya with an R² equal 0.54. However, the main improvement in this study is that by considering tree cover estimated via SAVI, and the standard deviation of NIR as an additional variable, we were able to improve the predictive power of the standard deviation of NIR on tree species diversity as evidenced by higher R² values ranging from 0.59 in Mukuvisi study site to 0.72 in Mabalauta study site. We therefore make a claim that the inclusion of a biomass indicator
such as SAVI in a model that includes the standard deviation of NIR improves prediction of tree species diversity in savanna landscapes. In addition, this model has also proved to be reliable by way of a low root mean square resulting from the comparison of test sites with the regression model predictions.

Overall, where our study differs from previous studies is in the use of the spectral variation hypothesis as well as the biomass-productivity hypothesis to improve our predictions of tree species diversity. In other words, we have combined a remotely-sensed satellite index of diversity (the standard deviation of NIR) with a remotely-sensed biomass indicator (SAVI) to improve our estimation of tree species diversity. However, we have to caution that in this study we used imagery of a single season which may result in our model being season dependent. But, the use of multiple date imagery for relating tree species diversity to satellite indices could be tested to find out whether successful remote sensing of tree species diversity is season dependent. In addition, future sampling should stratify according to the amount of biomass present, as well as the intensity of the variation (measured by the standard deviation) in order to sample across the full spectrum of the two indices.

Potentially, this method might not do very well in areas that have more canopy closure than the savannas. This is because vegetation indices such as SAVI and NDVI tend to saturate in areas of dense vegetation (Huete 1988). In this case, assuming uniform canopy cover, any variations in tree species diversity should be well explained by the standard deviation of the NIR radiance. In ecosystems of comparable vegetation types, the method should perform as well as it did in the three study sites. However, in areas of sparse vegetation such as arid deserts, the method might not do well because the soil
background will tend to dominate radiance in the NIR band, therefore distorting the true value of the standard deviation of the NIR band.
Chapter 5: GIS and remote sensing based spatial modelling of tree species diversity

5.1: Results

Our results of Moran’s I show that tree species data in Mukuvisi and Mabalauta study sites is spatially auto-correlated (Table 5.1).

Table 5.1: Results of the test for spatial autocorrelation in tree species diversity estimated by Simpson’s 1-D in the three study sites.

<table>
<thead>
<tr>
<th>Study Site</th>
<th>Moran’s I</th>
<th>Z</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mukuvisi</td>
<td>0.2473</td>
<td>2.715</td>
<td>0.0066</td>
</tr>
<tr>
<td>Kyle</td>
<td>-0.1427</td>
<td>-0.717</td>
<td>0.4731</td>
</tr>
<tr>
<td>Mabalauta</td>
<td>0.3567</td>
<td>2.508</td>
<td>0.0122</td>
</tr>
</tbody>
</table>

Table 5.2 shows the spatial regression models that were used to relate tree species diversity in Mabalauta and Mukuvisi study sites.

Table 5.2: The spatial regression equations used to relate tree species diversity to a combination of the standard deviation of NIR radiance and SAVI in Mabalauta and Mukuvisi study sites.

<table>
<thead>
<tr>
<th>Study site</th>
<th>Regression equation model</th>
<th>$R^2$</th>
<th>Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mabalauta</td>
<td>$z = 1.31 - 0.761x - 1.93y + 0.15x^2 + 1.81y^2 + 0.68xy$</td>
<td></td>
<td>Spatial</td>
</tr>
<tr>
<td>Mukuvisi</td>
<td>$z = -1.0 - 0.05x + 3.70y - 0.01x^2 - 13.51y^2 + 0.69xy$</td>
<td></td>
<td>Spatial</td>
</tr>
</tbody>
</table>

In all the cases, x is the standard deviation of NIR radiance, y is SAVI. In Mabalauta study site, $z$ is Simpson’s 1-D index, while in Mukuvisi study site, $z$ is the natural logarithm of Simpson’s 1-D index.

Our results also show that Simpson’s 1-D and Shannon Weaver’s H indices of diversity are positively correlated (fig. 5.2) thus for mapping tree species diversity, we used only Simpson’s 1-D because we conclude that the two indices would return similar tree species diversity patterns.
The spatial regression equations in table 5.2, as well as the linear regression equations for Simpson’s 1-D in figure 4.4 were used to map the spatial variations in tree species diversity in the three study sites.

Figure 5.2: Relationship between Shannon’s H and Simpson’s 1-D indices of diversity in a) Mukuvisi, b) Kyle and c) Mabalauta study sites.

Figure 5.3 shows the spatial variations in tree species diversity in Kyle study site mapped using a linear regression model.
Figure 5.3: Spatial variations in tree species diversity (estimated by Simpson’s 1-D) mapped using a linear regression model in Kyle Game Reserve.

Figure 5.4: Spatial variations of tree species diversity (estimated Simpson’s 1-D) predicted using the a) linear and b) spatial regression models in Mukuvisi study site. The ellipse shows the area of greatest contrast in prediction between this (spatial) model and the linear regression model.
Figures 5.4 and 5.5 show the spatial variations in tree species diversity in Mukuvisi and Mabalauta study sites mapped using both linear and spatial regression models.

![Maps showing tree species diversity](image)

**Figure 5.5:** The spatial variations in tree species diversity estimated by Simpson’s 1-D index in Mabalauta study site using a) linear and b) spatial regression models.

Also, we observe that the predictions from our linear regression models only slightly deviate from the tree species diversity for the test plots (Table 5.3). The lowest error magnitude of 1.28 % occurs when we predict tree species diversity using Simpson’s Index in Kyle Game Reserve, while the highest error occurs when we predict diversity in Mukuvisi Woodlands using Shannon’s index (Table 5.3).
Table 5.3: The magnitude of error (root mean square error) resulting from the estimation of tree species diversity (Simpson’s 1-D and Shannon’s H indices) in the three study sites.

<table>
<thead>
<tr>
<th>Study site</th>
<th>Simpson's Index</th>
<th>Shannon’s Index</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mabalauta</td>
<td>0.142</td>
<td>0.774</td>
</tr>
<tr>
<td>Kyle Recreational Park</td>
<td>0.128</td>
<td>0.575</td>
</tr>
<tr>
<td>Mukuzivi woodlands</td>
<td>0.327</td>
<td>1.001</td>
</tr>
</tbody>
</table>

Our results also show that the overall patterns of tree species diversity predicted by the linear and spatial regression models are similar, though some differences can be noted (figs 5.4 and 5.5). For example, in Mukuzivi woodlands, the north-eastern part shows patches of low tree species diversity when mapped using the spatial regression model. These patches are absent from the diversity map generated by a linear regression model (fig 5.4). In Mabalauta study site, tree species diversity mapped by the linear regression model shows higher values than that mapped by the spatial regression model (Fig 5.5). However, the differences in tree species diversity mapped by the two models are not significant (p > 0.05). In addition, a significant relationship exists between tree species diversity estimated using spatial and linear regression models (fig 5.6). In other words, the two regression models are consistent in their prediction of tree species diversity.
Figure 5.6: The relationship between tree species diversity estimated by the linear and the spatial (Conditional Auto) regression models in a) Mabalauta and b) Mukuvisi study sites.

Figure 5.7 shows the spatial variations in tree species diversity in Sebungwe district mapped using the regression models generated in Mabalauta study site. The maps produced using the linear regression and spatial regression models do not show significant differences in the patterns of tree species diversity. Table 5.4 shows the magnitude of the error (RMSE) that resulted from estimating tree species diversity in Sebungwe district using regression models that were generated in Mabalauta study site. The regression models that were used to predict tree species diversity from a combination of the standard deviation of NIR radiance and SAVI are shown in Table 5.5.

Table 5.4: The magnitude of root mean square error (RMSE) resulting from the estimation of tree species diversity (Simpson’s 1-D index) using the linear and spatial regression models in Sebungwe district.

<table>
<thead>
<tr>
<th>Model</th>
<th>Root mean square error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear regression</td>
<td>0.727</td>
</tr>
<tr>
<td>Spatial regression</td>
<td>0.703</td>
</tr>
</tbody>
</table>
Table 5.5: Regression models that were generated in Mabalauta study site, and were used in the prediction of tree species diversity in Mopane woodlands in Sebungwe district.

<table>
<thead>
<tr>
<th>Type of model</th>
<th>Equation model</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear</td>
<td>$z = 0.41 + 1.80y + 0.04x - 3.50y^2 - 0.01x^2 - 0.12xy$</td>
<td>0.92</td>
</tr>
<tr>
<td>Spatial</td>
<td>$z = 0.43 + 1.77y + 0.04x - 3.50y^2 - 0.01x^2 - 0.11xy$</td>
<td></td>
</tr>
</tbody>
</table>

$x$ is the co-efficient of variation of NIR radiance, $y$ is SAVI and $z$ is Simpson’s 1-D index.
Figure 5.7: The spatial variations in tree species diversity predicted using a) linear and b) spatial regression models in Sebungwe study site. Here, Simpson’s 1-D values show areas of lowest diversity (red) through yellow to areas of highest diversity (green). The white oblique line in the circled area is a gap left between the two satellite images.
5.2: Discussion

In this chapter we tested whether we can map the spatial variations in tree species diversity using linear and spatial regression models. Our results indicate that regression analysis can successfully be used in mapping tree species diversity. Also, we found no significant ($p > 0.05$) differences between linear regression and spatial regression predicted tree species diversity values. Our results differ with Dormann (2007) who found a significant difference between linear regression predicted data and spatially predicted data.

Although there were significant differences in the $R^2$ of linear and spatial regression models, several studies found no spatial autocorrelation in the model residuals, thereby resulting in no need to include a spatial auto correlation correction (Dormann 2007; Kissling and Carl 2008). This contradicts the general hypothesis that conventional statistics such as linear regression may not be suitable for analysing ecological data which is spatially auto-correlated (Anselin 2002).

Furthermore, our results also indicate that the spatial structure of tree diversity predicted using linear regression does not significantly differ from that of predictor variables, in this case the standard deviation of NIR radiance. This is not surprising as the predictor variables used in the linear regression model are raster images which already show spatial autocorrelation. Further to this point, when linear regression is used on data that is spatially auto correlated, the spatial structure of the data is transferred to the predicted variable. Previous studies involving spatially auto-correlated data mainly focused on interpolation methods such as kriging and inverse distance weighting which make use of the spatial structure of the sampled data (Lichstein et al. 2002; Sclocco and Di Marzio...
However, the accuracy of these interpolated maps has often been observed to depend on the character of the data set used, as well as the intensity of field sampling (Briggs et al., 1997; Lichstein et al. 2002). We claim that until a comprehensive study is done to assess the utility of including the spatial component of the data, linear regression analysis remains one of the methods that can be used to map tree species diversity.

Unlike previous studies which only described the nature of the relationship between tree species diversity and satellite imagery derived indices (Muldavin et al. 2001; Oindo and Skidmore 2002; Rocchini et al. 2007; Bradley and Fleishman 2008), in this study, we used these relationships to map tree species diversity within the landscape where the models were developed. Furthermore, we used the remote sensing based models to map tree species diversity in an ecosystem where the model was not developed.

We further make a claim that remote sensing based models developed in one ecosystem have the potential to characterise tree species diversity in other ecosystems that have similar physical characteristics. Overall, we showed the utility of remote sensing in the mapping of tree species diversity within the landscape. However, we have to caution that these generalisations were drawn from single season satellite imagery. Further studies may need to be conducted to find out if the utility of remotely-sensed imagery indices is only season dependent.
Chapter 6: Conclusions

6.1: The relationship between tree species diversity and satellite imagery derived satellite indices

Our first objective was to determine whether a significant relationship exists between tree species diversity and satellite imagery derived indices. We found that the standard deviation of NIR radiance derived from ASTER images, as well as, tree canopy cover estimated by SAVI, are significantly (p < 0.05) related to tree species diversity in savanna woodlands. Also, we found that our prediction of tree species diversity is improved when we use the standard deviation of NIR radiance and SAVI in combination. We therefore conclude that using the SVH and biomass-diversity hypothetical frameworks, we can successfully use remotely-sensed data particularly, the standard deviation of NIR radiance an indicator of diversity, as well as SAVI, an indicator of biomass to predict tree species diversity in a savanna landscape.

6.2: Remote sensing based spatial modelling of tree species diversity

The second objective was to determine whether the relationships between tree species diversity and satellite imagery derived indices can be used for mapping tree species diversity in savanna woodlands. We found that the relationship between tree species diversity and a combination of the standard deviation of NIR radiance, as well as, tree canopy cover estimated by SAVI calculated from an ASTER satellite image can be used to map the spatial variations in tree species diversity in savanna woodlands. Also, we found that our remote sensing and GIS based models can be used to predict tree species diversity outside of the study sites in which they were developed. In addition we also found out that both linear and spatial regression models are useful in the mapping of tree
species diversity using the standard deviation of NIR radiance and SAVI. Furthermore, we could not find any significant differences between tree species diversity mapped by the two models. We therefore conclude that remote sensing and GIS can be used to map tree species diversity patterns in a savanna landscape.

Overall, results of this study demonstrate that remote sensing and GIS based models can successfully be used to predict tree species diversity in the savanna woodlands when based on ecologically plausible hypothetical frameworks. However, to understand the temporal variations in the relationship with tree species diversity, we recommend more research to include the use of multi-date imagery. This, we hypothesise would be an important step towards producing a generalised model for predicting tree species diversity using remote sensing.
References


Masocha, M. (2009). Savanna aliens. Faculty of Geoinformation science and earth observation. Enschede, University of Twente. PhD.


