Modelling Canopy Density Variations from Remotely Sensed Data:
Implication on Monitoring Floristic and Macro-benthic Properties of
Mangrove Ecosystems

Dante Dias Torio
March, 2007
<table>
<thead>
<tr>
<th><strong>Course Title:</strong></th>
<th>Geo-Information Science and Earth Observation for Environmental Modelling and Management</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Level:</strong></td>
<td>Master of Science (M.Sc.)</td>
</tr>
<tr>
<td><strong>Course Duration:</strong></td>
<td>September 2005 - March 2007</td>
</tr>
</tbody>
</table>
| **Consortium partners:** | University of Southampton (UK)  
Lund University (Sweden)  
University of Warsaw (Poland)  
International Institute for Geo-Information Science And Earth Observation (ITC) (The Netherlands) |
| **GEM thesis number:** | 2005-23                                                                                   |
Modelling Canopy Density Variations from Remotely Sensed Data: Implications on Monitoring Floristic and Macro-benthic Properties of Mangrove Ecosystems

By

Dante Dias Torio

Thesis submitted to the International Institute for Geo-information Science and Earth Observation in partial fulfilment of the requirements for the degree of Master of Science in Geo-information Science and Earth Observation, Specialisation: (Environmental Modelling and Management)

Thesis Assessment Board

Chairperson: Prof. Alfred de Gier, ITC
External Examiner: Prof. Peter Pilesjo, Lunds Universitet, Sweden
Primary Supervisor: Dr. Jan de Leeuw, ITC
Member: M.Sc. Andre Kooiman

International Institute for Geo-Information Science and Earth Observation, Enschede, The Netherlands
Disclaimer

This document describes work undertaken as part of a programme of study at the International Institute for Geo-information Science and Earth Observation. All views and opinions expressed therein remain the sole responsibility of the author, and do not necessarily represent those of the institute.

I certify that although I have conferred with others in preparing for this assignment, and drawn upon a range of sources cited in this work, the content of this thesis is my original work.

Dante D. Torio
Abstract

Mangroves are biologically diverse and fragile coastal ecosystems that are severely threatened due to human activities in the coastal zone. These have depleted and degraded large mangrove areas resulting in the decline of wild fishery stocks, biodiversity loss and increased susceptibility of coastal areas to natural hazards. Because of the important ecological functions of mangroves and its continued degradation, there is a growing and urgent need to manage, restore and rehabilitate remaining mangrove areas. But with insufficient information on the current biophysical conditions of as well as the knowledge gap on ecological interactions in mangrove ecosystems, management are either short-sighted or severely constrained.

This study investigated the potential of Generalized Additive Models, Generalized Linear Model-based Huisman-Olff-Fresco Models, Multiple Linear Regression and Logistic Regression models in predictive modelling of canopy density variations from reflectance and vegetation indices derived from ASTER data. Macro-benthic and lower plant species response to the observed and modelled canopy density gradient were also investigated using logistic regression models. The result showed that soil vegetation indices like TSAVI and SAVI, and NDVI are the ideal indices for modelling mangrove canopy densities considering differences in environment. Multiple Linear Regression models improved prediction with the addition of Band 1, 2 and SWIR bands 5 and 6 ($\alpha-R^2 = 80\%$, Residual SE= 15%). Also the result of the modelled macro-benthic and lower plant species distribution is consistent with the observed and simulated canopy densities where the best results were obtained from the linear model with the greatest number of variables.

The two major results implies a promising utility of remote sensing and modelling techniques to determine biophysical changes in mangrove ecosystem based on easily measurable variables like canopy densities. With these, the goal of developing a reliable, rapid and cost-effective method of mapping mangrove biophysical indicators which can be related to the quality of mangroves both aboveground and understorey is not far from being reached.
Acknowledgements

In God nothing is impossible and with the compassion, charity and wisdom of his son Jesus Christ, made the Author comprehend the splendour and elegance of the Universe and the Ecosystem.

My sincere thanks to the institutions, and individuals, who, in the course of the M.Sc. Programme have played an important part towards the fulfilment of this exceptional degree on Environmental Modelling and Management, first; The European Commission and its Erasmus Mundus Programme for providing the financial support to pursue an excellent higher education programme in four European Countries. To the university consortium who provided not only excellent teaching methods and facilities but also their most dedicated professors and teachers. Specifically to Prof. Peter Atkinson and Mr. Gary Llewelyn of the University of Southampton-UK, Prof. Peter Pilejsjo and Prof. Karin Larsen of Lund University-Sweden, Prof. Katarzyna Dabrowska of Warsaw University-Poland, and Prof. Dr. Andrew Skidmore of ITC.

My deep gratitude to Prof. Jan de Leeuw, for his untiring support, critical comments and encouragement; Prof. Iris van Duren for her support for this thesis topic and Mr. Andre Kooiman who accommodated helpful discussions even in the most difficult times. Also to Mr. Gerard Reinink for providing access to ASTER images.

To the M.Sc. GEM Secretariat for their untiring efforts, professionalism and dedication especially in giving valuable practical and academic advices on how to make life easier for a foreign M.Sc. student in a European culture; particularly to Ms. Jorien Terlouw, Ms. Lyande Eelderink and Ms. Stef Webb.

Prof Jadunandan Dash, Prof. Micael Runnstrom, Prof. Simon Wood, Prof. Mike Austin and Prof. Jari Oksanen have also contributed their time by giving useful comments and technical advises. To the Provincial Agriculture Office-Pangasinan, the local governments of Bani, Bolinao, Alaminos and Anda, SAGIP-Lingayen Gulf Project and KAISAKA for providing an unlimited access and guide to the mangrove sites.

And lastly, to Matt, Irvin, Abel, Maina, Idham, Juan and Sally and the rest of the M.Sc. G.E.M Class of 2005 and to the ITC ‘Filipino Community’ for the discussions and friendship.
Dedicated to the two most influential women in my life,

Yanyan and Inang.
# Table of contents

1. Introduction ............................................................................................ 1  
   1.1. Background and significance........................................................ 1  
   1.2. Problem Statement ...................................................................... 6  
   1.3. General objectives ....................................................................... 7  
      1.3.1. Specific objectives .............................................................. 7  
   1.4. Assumptions ............................................................................... 8  
   1.5. Research Hypothesis ................................................................... 8  
      1.5.1. Linearity vs. Non-linearity Testing with GAM ................... 8  
      1.5.2. Huisman-Olff-Fresco Modelling ....................................... 8  
      1.5.3. Multiple Linear Regression Modelling ............................ 8  
      1.5.4. Logistic Regression Modelling of Species Distribution .... 9  
   1.6. Research Questions .................................................................... 9  
   1.7. Research Approach ................................................................. 10  

2. Materials and Methods ......................................................................... 11  
   2.1. Study Area .................................................................................. 11  
   2.2. Remote Sensing Data Acquisition and Image Pre-processing ...... 12  
   2.3. Unsupervised Classification ....................................................... 14  
   2.4. Percent Canopy Cover and Species Distribution Data ............... 14  
   2.5. Data Analysis and Modelling ..................................................... 15  
   2.6. Model Evaluation and Accuracy Assessment ............................. 17  

3. Results .................................................................................................. 19  
   3.1. Vegetation Properties and Canopy Density Modelling .............. 19  
      3.1.1. Data Characteristics ........................................................... 19  
      3.1.2. Non-parametric modelling with GAM ............................ 21  
      3.1.3. Parametric modelling with HOF .................................... 25  
      3.1.4. Multiple Linear Regression Modelling ............................ 27  
      3.1.5. Visualization in GIS ......................................................... 28  
   3.2. Understorey macro-benthos and cryptogams .............................. 31  

4. Discussion ............................................................................................ 35  
   4.1. Site biophysical characteristics and canopy cover ...................... 35  
   4.2. Relationship between band reflectance, vegetation index and  
        canopy cover ............................................................................. 37
## List of figures

Figure 1 Causal Loop Diagram of stable mangrove system (R) means reinforcing loop and (B) means balancing loop. Adapted with simplification from Lugo and Snedaker (1974). ................................................................. 3

Figure 2 ASTER Bands 231 composite of the study Area, located in Tambac Bay, Lingayen Gulf, Pangasinan, North-western Philippines.................. 11

Figure 3 (A) Canopy cover dendrogram (in decimals) and its (B) Frequency distribution ................................................................. 19

Figure 4 Three different types of mangrove habitats A) Seaward B) Middleward and C) Riverine mangroves in ASTER (bands 231) composite with corresponding photos. ................................................................. 20

Figure 5 Average spectral reflectance (in percentage) of the different mangrove communities. ................................................................. 21

Figure 6 Partial residual plots of the smooth components of the fitted GAM Reflectance model................................................................. 23

Figure 7 Partial residual plots of the smooth components of the fitted GAM vegetation index models................................................................. 24

Figure 8 Comparison between HOF and GAM fitted curves for the selected parsimonious models................................................................. 26

Figure 9 Residual and scatter plots of the fitted Multiple Linear Regression (MLR) models fitted with the unbiased estimator (α=0; and β=1). ............. 28

Figure 10 Prediction curve of the HOF-fitted TSAVI and SAVI models for canopy density................................................................. 29

Figure 11 Comparison of the predicted canopy density maps (A) TSAVI-HOF II, (B) SAVI-HOF II (C) MLR Model 1 and (D) MLR Model 2 ........ 30

Figure 12 Comparison of the models with ground observation in terms of mean canopy densities and standard error. ............................................. 30

Figure 13 Cryptogams and Macro-benthic species found in mangroves ........................................................................................................... 31

Figure 14 HOF fitted response curve of the species distribution on observed and modelled canopy density gradients ........................................... 33
List of tables

Table 1 Reflectance-based vegetation Indices ............................................. 13
Table 2 Huisman-Olff-Fresco Model Hierarchies ........................................ 17
Table 3 Statistical summary of the fitted GAM models for band reflectance and vegetation indices. Enclosed e.d.f. values are results from individual model fitting. Enclosed a-R^2 are linear model values ................................. 22
Table 4 Statistics summary of selected parsimonious HOF fitted models... 25
Table 5 Comparison of the logistic model R^2 for the four species distribution in the four canopy density models and the observed canopy density, NS indicate non-significance................................................................. 32
Table 6 Statistical summary of the logistic regression models for species distribution. .................................................................................................................. 32
1. Introduction

1.1. Background and significance

Mangroves, an important and fragile tropical coastal ecosystem are in need of conservation and management. These ecosystems are not only biologically and ecologically significant (Mumby, 2006) but also an important natural buffer (Dahdouh-Guebas et al., 2005, Danielsen et al., 2005, Kathiresan and Rajendran, 2005). Many coastal communities benefit from mangroves in terms timber and non-timber forest products, food from fisheries and protection against violent storm surges, tsunamis, floods and typhoons. In fishery alone, it was estimated that for each hectare of mangrove, about 600 kg of fish and 600 kg of shrimp could be sustained and captured annually (Naylor et al., 2000). This makes mangroves a valuable national asset. However, the increasing population in coastal areas coupled with the increasing demand for coastal resources have resulted in the depletion of large areas of mangroves through overexploitation, modification and pollution (Lotze et al., 2006, Naylor et al., 2000).

One of the most adverse forms of modification of mangrove habitats especially in Asia is conversion to aquaculture for shrimps and milkfish. In the Philippines for instance, only a fifth of the 500,000 hectare (ha) mangroves remains at the present while brackish water ponds have increased dramatically from 61,000 ha in 1940 to 230,000 ha at present (Primavera, 2005) all in the expense of mangrove habitats. As a consequence this does not only reduced the areas of mangroves but also undermined its ecological functions as nursery habitats, nutrient sink, and coastline protection (Lotze et al., 2006). It was estimated that more than 100 kg of on-site fish biomass was lost for every hectare of converted mangroves (Naylor et al., 2000, Primavera, 2005). Mangrove destruction in the long term does not only lead to decline in fisheries but also to negative impacts on adjacent sea grasses and coral reef because of their interconnection (Yáñez-Arancibia et al., 1993). Although mangroves are one of the most important tropical ecosystems, it receives less publicity (Valiela et al., 2001). Recently, however, as the issues on climate change, biodiversity, tsunamis and sea level rise become sensational the need for mangrove
conservation and management in tropical areas were being seriously reconsidered (Dahdouh-Guebas et al., 2005, Naylor et al., 2000).

The continued depletion and degradation of mangroves has gained international concern because of its significance on fisheries and coastal ecology. Hence, there is an ongoing worldwide effort to conserve and sustainably manage remaining mangrove areas. Some of these programs are spearheaded by the Convention on Wetlands of International Importance (CWII-Ramsar) and United Nation (UN)-Man and Biosphere Programme (MAB). As of August 2006 there are 182 protected mangrove sites all over the world under CWII protection (Ramsar, 2007). Government, non-government and community initiated mangrove conservation efforts are also contributing to the effort with some remarkable success (Salmo et al., 2004). On the contrary, the availability of baseline data to base the formulation of management plans and as a benchmark for monitoring remains a major hindrance.

There are major challenges in relation to mangrove monitoring. Diaz and Blackburn (2003) argued that reliable indicators or quantitative evidence to measure the susceptibility of mangroves to natural and man-made disturbances are lacking as well as the tools to do so. Aside from these, the accessibility of mangrove areas makes direct measurements a formidable technical challenge. Like for instance the onsite measurement of canopy cover, leaf area index (LAI) and biomass (Green et al., 1997a, Benfield et al., 2005) are limited by tidal regimes and weather. Even more difficult, yet fundamental, is relating understorey conditions to the changes in vegetation properties which is an important ecological feature in mangroves. The spatial and temporal aspects of monitoring are also of primary concern. One possible solution to this problem is the use of remotely sensed vegetation data in a modelling environment.

Ecologically, there are two related aspects of mangroves which are important to monitor (Fig. 1); the aboveground and the understorey (Lugo and Snedaker, 1974). The above ground includes the canopy cover or density and forest structure which are related to photosynthesis and primary production. Canopy cover or density for instance is considered as a surrogate for stand density, biomass, crown bulk density and an inferential measure of productivity, nutritional value, stand health and stand micro-climate (Singkran and Sudara, 2005, Hall et al., 2006). Canopy cover is one of the easiest measurable biophysical properties of mangroves. In relation to remote sensing, canopy cover is a key spatial variable that governs scene brightness and controls the fraction of over-storey and understorey visible to the sensor (Nemani et
MODELLING MANGROVE CANOPY DENSITY VARIATIONS

al., 1993). Lowman and Wittman (1996) also discovered that in most natural forests, canopy cover is directly linked with understory species diversity. As such, any changes in the canopy could directly and strongly affect the understory condition. Though very obvious in mangroves, studies to prove the significance of the canopy-understorey relationships are insufficient.

![Causal Loop Diagram of stable mangrove system](image)

Figure 1 Causal Loop Diagram of stable mangrove system (R) means reinforcing loop and (B) means balancing loop. Adapted with simplification from Lugo and Snedaker (1974).

The understorey components of the mangroves are equally important and their relationships with the aboveground components are profound. Biomass produced in the canopy are converted into detritus which provide nourishment to the understory macro and micro-organism and even export to adjacent habitats like sea-grass and coral reefs (Yáñez-Aracibia et al., 1993). Aside from food production, the ideal micro-climate provided by dense mangroves encourages organism recruitment from other ecosystems to shelter in mangroves during their juvenile or delicate life stages (Blasco et al., 1996, Field et al., 1998). Ideal moisture condition provided by adequate mangrove canopies also promotes the growth of lower plant species such as lichen, algae and epiphytes adding to the variability of related species. Because of these, mangroves are regarded as one of the most diverse ecotones in the tropics.

The implication of a changing above ground properties to the below ground conditions in mangroves are enormous. These changes whether natural and man-made could determine the future resilience, functioning and structure of the ecosystem (Lugo and Snedaker, 1974). Disturbances such as deforestation and
habitat modification have resulted in discontinuities, patchiness and increased canopy openings which not only reduced stand densities and regeneration but also impacted stand micro-climate (Berger et al., 2006, Kovacs et al., 2001, Walters, 2005). Specifically, thinning of canopy cause heat influx and increase salinity leading to desiccation of understory flora and fauna (Debenay et al., 2002). Lugo and Snedakker (1974) observed that thinning causes a significant reduction in evapotranspiration and an associated 40% reduction in the net primary productivity of the system. Likewise, mangrove habitat modification to aquaculture ponds and other land uses has resulted in significant degradation of coastal water quality and the health of fish habitats, causing the decreases in wild fish stock abundance and species richness (Singkran and Sudara, 2005). Considering these impacts any monitoring attempt for management must address both the spatial and ecological aspects of aboveground and understory mangrove components simultaneously. This is only possible if there is a direct variable that could be measured by sensors and could link the above ground properties to the understory. Modelling canopy properties using sensor derived data are some prospects.

The use of remote sensing in spatial prediction and modelling of vegetation biophysical properties are prominent in terrestrial forest (Jing and Cihlar, 1996, Foody et al., 2001, Muukkonen and Heiskanen, 2005, Falkowski et al., 2005, Xu et al., 2003) but very limited in mangroves (Green et al., 1997a, Diaz and Blackburn, 2003, Vaiphasa et al., 2006). Vegetation indices resulting from differencing, ratioing and orthogonalization of different sensor bands could provide a surrogate and quantitative measure of vegetation properties such as “greenness” and of leaf area index (Green et al., Elvidge and Chen, 1995). Some of these commonly used vegetation indices include, Advance Vegetation Index (AVI) by Rikimaru and Miyake (2004), Normalized Difference Vegetation Index (NDVI), Ratio Vegetation Index or Simple Ratio (RVI or SR), Difference Vegetation Index (DVI) and Soil Adjusted Vegetation Index (SAVI), modified RVI or RSR, modified SAVI (SAVI2) and Transformed SAVI (TSAVI) (Rouse, 1974, Jordan, 1969, Huete, 1988, Tucker, 1979, Major et al., 1990, Baret et al., 1989, Richardson, 1977, Brown et al., 2000).

Accordingly, two of the most important vegetation properties commonly modelled using remotely sensed data are Leaf Area Index (LAI) (Green et al., 1997b) and canopy density or canopy cover. These are relevant in many forestry applications such as quantifying crown fuel layer development (Falkowski et al., 2005), biomass estimation (Muukkonen and Heiskanen, 2005, Heiskanen, 2006) and mapping
invasive plant species (Joshi et al., 2006). Empirical modelling using ground data with the different reflectance or vegetation indices as inputs are the usual approaches in most vegetation studies in temperate forests (Falkowski et al., 2005, Heiskanen, 2006). In mangroves there are limited studies reported. Green et al.(1997a) used NDVI in an empirical model for predictive modelling of mangrove percent canopy cover for Landsat and Diaz and Blackburn (2003) found that SAVI, DVI, and PVI have an high affinity to LAI in a laboratory conditions. Not all vegetation indices however are useful because they performed differently in different vegetation and background conditions (Huete, 1988). As such, the indices ideal for mangroves are largely unknown.

Predictive modelling using remotely sensed data offers robust alternative to the rigorous field observation techniques. On the contrary, aside from the problem as to which vegetation indices or satellite information are useful, the appropriate models that could deliver the optimum information both in spatial and ecological context remain unresolved. Further, there are major considerations before the potential of predictive models are fully utilized as different models depend on several statistical and ecological assumptions. In ordinary least square regression for instance; the errors are assumed to be identically and independently distributed with constant variance of the response variable across observations; the errors are assumed to follow a normal distribution or Gaussian; and the regression function is linear in predictors (Guisan et al., 2002). Most often when dealing with ecological data these assumptions are violated (Austin, 2002).

Modelling an ecological system requires the consideration of both statistical and ecological assumptions. This means treating the relationship between covariates and response in the light of existing ecological theories. The neglect of ecological knowledge is argued to be one of the major limitation of predictive modelling of species response or distribution along an environmental gradient (Austin, 2002). In vegetation, there are two prominent conflicting theories explaining distribution. The continuum theory (Gauch and Whittaker, 1972) views vegetation as a continuum with a changing species composition along an environmental gradient, while community unit theory regard vegetation as communities of naturally co-evolved species forming a homogenous, discrete and recognizable units (Austin, 1985, Austin and Smith, 1989). Vegetation continuum is related to the environmental space while community is with geographical space (Austin and Smith, 1989).
The implication in predictive modelling of the two theories is that when vegetation is assumed as a continuum one could generalize the response for the whole environmental space and apply in other regions provided that the correlation between growth influencing variables and gradients do not change (Austin and Austin, 1980, Austin, 1980). While for community, each of the different groups (geographical space) or communities are modelled separately and can only be relevant to a particular landscape (Austin and Smith, 1989). Aside from this, variability of response from continuum is to be expected when generalization is applied while a more uniform response is to be expected from a community. In mangroves both cases are possible and their implication could shed light in finding the optimum modelling approach for mangrove vegetation properties and belowground species response.

1.2. Problem Statement

The question as to what satellite information or vegetation indices are relevant in modelling vegetation properties in mangroves remains unresolved. Specifically, there is a need for a reliable, cost-effective method for predicting and mapping mangrove canopy density variations consistent with the existing ecological theories of vegetation. From the usual location-based or site specific canopy observations, extrapolation over the complete spatial extent is needed if the requirement of constant monitoring and the problem of accessibility and practicality is to be addressed. The use of remote sensing data for this is straightforward, but still needs further illumination. Alongside with this, relating aboveground changes with understorey biota is inseparable to the objective of comprehensive monitoring as mangrove vegetation is strongly linked with the presence and distribution of understorey species.

Predictive modelling is powerful tool; however, its utility in ecosystems like mangroves is defined by the ecological condition of the modelling space, the characteristics of the variables along with its statistical assumptions and the type and appropriateness of the model architecture. Ordinary least squares regressions are the norm in many forest modelling studies using field and remotely sensed data, but these models have limitations in terms of their assumptions. Generalised models on the other hand, have the potential to offer more robustness and ecological sense in determining the relevant satellite information or vegetation indices as well as the nature of response between the variables with ground data. So it is relevant in this study to determine if generalized models are of use in determining the optimum
satellite information on vegetation that could be used in predictive modelling of both canopy variations and understorey species distribution.

1.3. **General objectives**

This study was aimed at testing the performance and usefulness of both least squares regression and generalized models for predictive modelling of canopy density and species distribution from remotely sensed data. Two set of results were expected, one for canopy density and the other for the species response. Particular focus were on testing Generalized Additive Models (GAM) (Hastie and Tibshirani, 1987, Wood and Augustin, 2002), Generalized Linear Model (GLM)-based Huismann-Off-Fresco (HOF) (Nelder and Wedderburn, 1972, Huisman *et al.*, 1993, Guisan and Zimmermann, 2000, Guisan *et al.*, 2002, Oksanen and Minchin, 2002) Models and Multiple Linear Regression (MLR) to determine the actual relationship of canopy cover with vegetation reflectance. On the other hand, logistic regression models (Nicholls, 1989) and HOF models were used to test the different response of macro-benthic and cryptogam species in the observed and modelled canopy densities.

1.3.1. **Specific objectives**

For the canopy density/cover, the first objective was to evaluate the nature of the canopy cover data whether it satisfies assumptions on normality, linearity and constant variance. Secondly, explore its actual relationship with remotely sensed vegetation reflectance and indices using non-parametric GAM models and test for linearity and non-linearity. The focus of the exploratory analysis was to determine the shape of the relationship curves and identify relevant indices and band reflectance that are useful in parameterization with HOF. Thirdly, determine the most parsimonious and relevant HOF model hierarchies for canopy density prediction. Using the result from HOF models, a stepwise MLR was implemented to determine if prediction would improve by the addition of other satellite information. Ultimately, use both the linear and HOF model algorithm in a GIS to create a predicted canopy density maps.

For the macro-benthic species models, a stepwise logistic regression was implemented to test the relationship between presence or absence data and the modelled and observed canopy densities. The resulting significant models were used in HOF model to determine the shape of the species response curve to canopy density gradient and compare the resulting curves whether they are consistent. Finally, the shape of the response curve was explained in the context of the species autecology.
1.4. **Assumptions**

- Mangroves are a continuum of vegetation with a changing canopy structure in response to changing indirect and resource gradients along their fundamental niche. These gradients therefore define the expression of the vegetation characteristics in terms of reflectance and canopy cover.
- Canopy density is a continuous variable.
- Canopy cover encompasses different angles of view.

1.5. **Research Hypothesis**

1.5.1. **Linearity vs. Non-linearity Testing with GAM**

Ho: The relationship between canopy density with vegetation index and reflectance is linear (e.d.f. = 1)

Ha: The relationship between canopy density with vegetation index and reflectance is non-linear (e.d.f. > 1).

1.5.2. **Huisman-Olff-Fresco Modelling**

Ho: The change in $R^2$ or residual deviance between model $M_i$ with parameters $p_i$ and model $M_j$ with parameter $(p_i - 1)$ with difference in the number of their parameters (K) is not significant ($F_{0.05}$, K, n-p)

Ha: The change in $R^2$ or residual deviance between model $M_i$ with parameters $p_i$ and model $M_j$ with parameter $(p_i - 1)$ with difference in the number of their parameters (K) is significant ($F_{0.05}$, K, n-p)

1.5.3. **Multiple Linear Regression Modelling**

Ho: The addition of other satellite information to the vegetation indices determined by HOF models does not significantly improve prediction.

Ha: The addition of other satellite information to the vegetation indices determined by HOF models significantly improves prediction.
1.5.4. **Logistic Regression Modelling of Species Distribution**

Ho: The negative log-likelihood difference of the relationships between modelled canopy densities, reflectance and vegetation indices with species presence probabilities is not significant ($X^2_{0.05, n}$)

Ha: The negative log-likelihood difference of the relationships between modelled canopy densities, reflectance and vegetation indices with presence probabilities is significant ($X^2_{0.05, n}$).

1.6. **Research Questions**

1. What is the characteristic of the field observation data on canopy cover in terms of normality? Does this reflect the ecology of the site surveyed? What does it imply on the succeeding modelling approach?

2. What are the relevant indices or bands that showed strong relationship with the canopy cover data in GAM and HOF? Are the modelled relationship statistically significant and what confidence?

3. What is the optimum or parsimonious HOF model and what particular vegetation index or bands are relevant?

4. Does the use of these indices with other bands in a Multiple Linear Regression improve canopy density prediction?

5. Are there any consistent and significant relationship between the species distribution and the observed and modelled canopy density? At which model and what does it implies? What are the shapes of the individual species response curve and are these relationships consistent with the species autecology?
1.7. **Research Approach**

### Process for Modelling Mangrove Canopy Density

<table>
<thead>
<tr>
<th>Initial stage</th>
<th>Non-Parametric</th>
<th>Parametric</th>
<th>Evaluation and GIS Integration</th>
<th>Final Output</th>
</tr>
</thead>
<tbody>
<tr>
<td>Image pre-processing</td>
<td>Unsupervised Classification</td>
<td>aR-squared or Deviance explained</td>
<td>Mangrove/Non-mangrove Mask</td>
<td>Mangrove Canopy Density Map</td>
</tr>
<tr>
<td>Reflectance and Vegetation Indices</td>
<td>Generalized Additive Modelling</td>
<td>HOF Modelling</td>
<td>GIS Implementation of Models</td>
<td>Final Document</td>
</tr>
<tr>
<td>Field Data</td>
<td>Statistical Test and Accuracy Assessment</td>
<td>Stepwise Multiple Linear Regression</td>
<td>Discussion and conclusion</td>
<td></td>
</tr>
<tr>
<td>Literature Review</td>
<td>Ecological Theory, Models, Vegetation Remote Sensing, Species Autecology</td>
<td>GAM Modelling Results</td>
<td>Linear Regression Results</td>
<td></td>
</tr>
</tbody>
</table>
2. **Materials and Methods**

2.1. **Study Area**

The study area (Fig. 2) is located between 15.90-16.49 degrees latitude and 119.72 – 120.15 degrees longitude in the North-western part of Philippines. The Gulf itself is biologically, geologically and economically significant because of mangroves, coral reef and sea grass communities in various conservation statuses. Along with this, several marine protected areas are established and managed in the area. The Gulf itself is a major source of regional income from fishing, and eco-tourism.

![Figure 2 ASTER Bands 231 composite of the study Area, located in Tambac Bay, Lingayen Gulf, Pangasinan, North-western Philippines](image-url)
The climate of the areas is characterized by pronounced dry season from November to April and wet season during the rest of the year. The average annual temperature is 28°C. A maximum temperature of 35°C is usually recorded in April, and a minimum of 18°C in January. Rainfall patterns vary significantly throughout the year in which the average is 2,500 mm, with a peak of 8,000 mm of rain falling in August and as low as 1 mm in January (Synthesis Report, 2000).

Although the mangrove areas in the study site were severely reduced (<10% remaining) there were still remarkable number of species found on small patches (<50 hectares). These areas, though small, were still significant habitats of birds and fishes, crustaceans and gastropods. Based on the previous surveys there are 15 mangrove species found along, seaward, middleward and riverine areas. Some of the dominant species includes: *Avicennia sp.*, *Ceriops decandra*, *Sonneratia alba*, and *Rhizophora sp.* (See Appendix in 7.1)

2.2. Remote Sensing Data Acquisition and Image Pre-processing

Four satellite images from the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) (Abrams, 2000) with each bands calibrated at radiance-at-sensor units (L1B) were obtained for the study area. The images were taken in separate dates; two images were taken on March 14, 2004 and the other two on April 15, 2004. These images were further geometrically corrected using Ground Control Points (GCPs) that comes with the images and to local projection system. For this the Universal Transverse Mercator (UTM) with Clarke 1866 Spheroid on Luzon Geographic Zone 50 was used. Only the first 3 (VNIR) bands and 9 SWIR bands were used in the analysis. Re-sampling to 15 meter from the original 30 meters spatial resolution of the SWIR bands were performed before stacking them with the 15 meter resolution VNIR bands into one single file. This was applied to the rest of the ASTER level 1B images which were used in the analysis. Since the study area was widespread the four ASTER images were mosaic into a single image and then subsetted.

The study dealt with several vegetation indices so atmospheric and radiometric calibration was implemented to eliminate possible atmospheric perturbations. The output was a scaled 0-100 % surface reflectance. This was implemented in ATCOR2 software using the latest ASTER calibration file (Geosystems, 2007). After radiometric and atmospheric calibration the image was imported into a desktop and Arcpad Mobile GIS platform for further processing and for the field surveys, respectively.
Nine (9) vegetation indices (Table 1) were calculated and the corresponding image maps were produced. The values of the different vegetation indices and reflectance at the location where canopy cover was measured were extracted using nearest neighbourhood sampling. For the regression analysis, vegetation indices and bands formed the covariate variables (X’s) and percent canopy cover was the response (Y).

### Table 1 Reflectance-based vegetation Indices

<table>
<thead>
<tr>
<th>Abbr.</th>
<th>Description</th>
<th>Equation</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>AVI</td>
<td>Advanced Vegetation Index</td>
<td>$B_{43} = B_4 - B_3$ after normalization of the data range.</td>
<td>(Rikimaru and Miyatake, 1997)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$f(AVI) = \begin{cases} 0 &amp; (</td>
<td>AI</td>
</tr>
<tr>
<td>DVI</td>
<td>Difference Vegetation Index</td>
<td>$NIR - RED$</td>
<td>(Tucker, 1979)</td>
</tr>
<tr>
<td>NDVI</td>
<td>Normalized Differented Vegetation Index</td>
<td>$\frac{NIR - RED}{NIR + RED}$</td>
<td>(Rouse, 1974)</td>
</tr>
<tr>
<td>PVI</td>
<td>Perpendicular Vegetation Index</td>
<td>$\frac{NIR - a RED - b}{\sqrt{1 + a^2}}$; $a$ and $b$ are intercept and slope values derived from soil baseline: $a = 3.69, b = 1.17$</td>
<td>(Richardson, 1977)</td>
</tr>
<tr>
<td>RVI</td>
<td>Ratio vegetation Index</td>
<td>$\frac{NIR}{RED}$</td>
<td>(Jordan, 1969)</td>
</tr>
<tr>
<td>RSR</td>
<td>Modified RVI</td>
<td>$RV_{\left(1 - \frac{SWIR_{\text{max}} - SWIR_{\text{min}}}{SWIR_{\text{max}} - SWIR_{\text{min}}}\right)}$; SWIR, corresponds to Band 4 in Aster and min and max are the observed reflectance values from the same Bands in the field points</td>
<td>(Brown et al., 2000)</td>
</tr>
<tr>
<td>SAVI2</td>
<td>Modified SAVI</td>
<td>$\frac{NIR}{RED + a/b}$; $a$ and $b$ have the same values as in PVI</td>
<td>(Major et al., 1990)</td>
</tr>
<tr>
<td>SAVI</td>
<td>Soil Adjusted Vegetation Index</td>
<td>$(1 + L) \frac{NIR - RED}{NIR + RED + L}; L = 0.5$</td>
<td>(Huete, 1988)</td>
</tr>
<tr>
<td>TSAVI</td>
<td>Transformed Soil Adjusted Vegetation Index</td>
<td>$a(NIR - a RED - b) / RED + a(NIR - ab)$; $a$ and $b$ values are same as in PVI</td>
<td>(Baret et al., 1989)</td>
</tr>
</tbody>
</table>
2.3. **Unsupervised Classification**

An unsupervised classification using 20 iterations in Isoclsuteralgorithm (Campbell, 2002) was performed using the first 6 ASTER bands to create a mangrove/non-mangrove image map. This image was then reclassified into a Boolean raster image (1 and 0) representing mangrove and non-mangrove areas and used as a mask. This is to limit the output map only to mangrove areas.

2.4. **Percent Canopy Cover and Species Distribution Data**

Fieldwork was undertaken from September 25, 2006 to October 25, 2006. The two main purposes of the field activities were 1) to measure percentage mangrove canopy at different mangrove sites and 2) collection of macro-benthic and algae species data in the same spot as the canopy measurements.

Ten mangrove sites around the study area were visited and surveyed for vegetation and macro-benthos distribution. Cluster sampling was implemented in an irregular transect/line-intercept method (Fiala et al., 2006) moving from seaward to inland direction and in riverine areas. All measurement was conducted during low tide where the soil surface was often exposed. One transect were established per site, so there are 10 transects in total. In each sites and transect, at least 5 observation points which were approximately 50 meters away from each other were identified and marked with GPS. In each observation points 4 spots were designated including the centre of the observation station which was approximately 20 meters equidistant. The 20 meters distance of point from the centroid corresponds to resolution of the image (15mX15m) +5 meters accuracy of the GPS.

Canopy cover measurement was made using spherical densiometer (Fiala et al., 2006) in each of the 4 spots and in the four cardinal directions, these were then averaged to calculate the estimated the mean % canopy cover in that point. In total there were 4 sets of average canopy closure readings. The readings are further grouped into two and averaged such that there are 2 sets of data with the same number of samples (n=84). The first set was used to calibrate the model and the second set was used for validation. The averaged percent canopy closure data were then arcsine transformed (eq.1) to minimized heterosedasticity and systematic error as suggested by Huisman et al (1993):

\[ tY = \arcsin(\sqrt{\frac{y}{100}}) \]  \hspace{1cm} (1)

Where \( tY \) is the transformed response variable and \( y \) is the original response variable.
In the case of Macro-benthos and algae, target species were identified by the local communities. Any species occurring within a roughly estimated 10 meter radius plot from the centre of the observation station were recorded as binaries (presence=1, absence=0).

2.5. **Data Analysis and Modelling**

Generalized Additive Model (GAM) (Hastie and Tibshirani, 1987) and GLM (Nelder and Wedderburn, 1972) were developed to address non-linear relationships common to ecological data. In this study, GAM was used in a regression analysis to explore the relationship between percent canopy cover, band reflectance and vegetation indices. An individual GAM model was fitted with each variable separately. In GAM the predictor fits the data by using a smooth function (Yee and Mitchell, 1991) in this case the penalized regression spline (Wood and Augustin, 2002).

\[
g(\mu) = \alpha + \sum_{j=1}^{p} f_j(x_j)
\]  

(2)

Where \( x \) is the mean of the response variable; \( g \) is a pre-specified function called link function, \( \alpha \) is the intercept or constant term and \( f_j \) is the penalized regression spline calculated using Generalized Cross Validation through iteration (R Documentation, 2007). Model fitting was implemented in R using the GAM function in the MGCV library.

For reason of parsimony; parametric fits are preferred whenever a parametric curve is statistically allowable (Yee and Mitchell, 1991). For parametric modelling, both the Multiple Linear Regression (MLR) and GLM were used. MLR is expressed as:

\[
Y = a + b_0X1 + b_1X2 + \ldots + b_iX_i + \varepsilon
\]  

(3)

where \( Y \) is the predicted response variable, \( a \) is the intercept, \( b_i \) is the linear predictor or slope of independent variable \( X_i \) and \( \varepsilon \) is the error term.

In general, GLM modelling (eq. 4) combines the predictor variable \( X_j \) (j=1…p) to produce a linear predictor \( LP \) which is related to the expected value \( \mu = E(Y) \) of the response variable \( Y \) through different link function \( g() \) that is related to the actual
distribution of the response variable Guisan and Zimmermann (Guisan and Zimmermann, 2000, Yee and Mitchell, 1991).

\[ g(\mu) = LP = \alpha + \beta^T x = \alpha + \sum_{j=1}^{n} \beta_j x_j \]  

(4)

Where \( x \) is the mean of the response variable; \( g \) is a pre-specified function called link function (i.e. logit, probit, Gaussian, Poisson), \( \alpha \) is the intercept or constant term; \( \beta = [\beta_1, \beta_2, \ldots, \beta_p]^T \) is a vector of regression coefficients.

Specifically for GLM, HOF models were used. Huisman-Olff-Fresco (HOF) are GLM-based parametric models that uses logistic regression to estimate maximum likelihood parameters for non-linear relationships by maximising log-likelihood or minimizing squared residuals (Huisman et al., 1993, Oksanen and Minchin, 2002). HOF models are based on the principle of GLM and GAM where different fitting is done iteratively but unlike in GAM which uses smooth function to fit the data, HOF instead uses hierarchical set of non-linear logistic regression models (eq.4) with increasing number of parameters (Table 2).

\[ | Ey | = \frac{e^{(a+bX)}}{1+e^{a+bX}} \]  

(5)

The five hierarchical set of models behaves sensitively with respect to the data. HOF models were fitted in the same way as GAM using R and a stand alone library package called GRAVY developed by Oksanen and Minchin (2002). All statistical analysis was performed in R© Statistics.
Table 2 Huisman-Olff-Fresco Model Hierarchies

<table>
<thead>
<tr>
<th>Model Type</th>
<th>Characteristics</th>
<th>Model Equation</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>No significant trend in Space or Time; Flat Response Type</td>
<td>( y = M \frac{1}{1 + e^a} )</td>
</tr>
<tr>
<td>II</td>
<td>An increasing or decreasing trend where the maximum is equal to the Upper bound M; Monotonic Response Type</td>
<td>( y = M \frac{1}{1 + e^{ax}} )</td>
</tr>
<tr>
<td>III</td>
<td>An increasing of decreasing trend where the maximum is below the upper bound M; Plateau response type</td>
<td>( y = M \frac{1}{1 + e^{ax+b}} )</td>
</tr>
<tr>
<td>IV</td>
<td>Increase and decrease by the same rate; Symmetric (Bell-shaped) response type</td>
<td>( y = M \frac{1}{1 + e^{ax+bx}} )</td>
</tr>
<tr>
<td>V</td>
<td>Increase and decrease by different rates; Skewed response</td>
<td>( y = M \frac{1}{1 + e^{ax+bx}} )</td>
</tr>
</tbody>
</table>

Adopted from: Huisman, et al., 1993

Where \( y \) and \( x \) are response and explanatory variable respectively; \( M \) is a constant values equal to the maximal value in frequencies or percentages (1 and 100 respectively).

2.6 Model Evaluation and Accuracy Assessment

Evaluation of fitted GAM model in terms of linearity and prediction was based on the explained deviance, adjusted coefficient of determination (a-R\(^2\)) and the effective degrees of freedom (e.d.f). The deviance explained is the proportion of null deviance explained by the model, analogous to coefficient of determination (R\(^2\)) in ordinary least squares. Fitting was tested against F-statistics. The e.d.f on the other hand is a measure of the number of “knots” that the model takes in fitting a smooth curve to the data, similar to polynomial terms. It is also an indication whether the relationship are linear or non-linear, which an e.d.f equal or close to 1 means a linear relationship (Wood and Augustin, 2002). Only those GAM models with realistic fits and high R\(^2\) or explained deviance were chosen for HOF modelling. Goodness of fit is measured using either the transformed maximized likelihood deviance- or R\(^2\) which follows an asymptotic chi-squared and F-distribution, respectively. Both are applicable for continuous data. In HOF, models are ranked according to increasing complexity and model selection is based on backward elimination. Backward elimination evaluates the fitted models through cross-checking of R\(^2\) between models starting with the least number of parameters (Table 2 Model Type I). The resulting values from the cross-checking (eq.6) procedure are then compared to F-critical values with K, N-K.
degrees of freedom. Comparisons are then tested using F-test. The main purpose of evaluating the hierarchical sets of models is to determine whether an increase in the number of parameter yields a significantly better fit and identify which model is the optimum. The optimum model is chosen based on F-statistics and parsimony (Huisman et al., 1993).

\[
F_{K,N-J-K} = \frac{(R^2_{\text{new}} - R^2_{\text{old}}) N - J - K}{K}
\]

where \( N \) is the total number of observations, \( R^2_{\text{old}} \) is the \( R^2 \) of a model with \( J \) estimated parameters and \( R^2_{\text{new}} \) represents \( R^2 \) of a model with \( J+K \) parameters (Huisman et al., 1993).

For binary data such as species distribution fitted with logistic regression, the residual deviance is used. Residual deviance is calculated as twice the negative of the difference between the log-likelihood deviance of the model with \( p \) parameters and the null model. This follows an asymptotic chi-squared distribution thus a significant model has a residual deviance greater than the chi critical value at a certain confidence with \( n \) degrees of freedom (\( X^2_{0.05, 1, \text{d.f.}} \)).

\[
-2\times(\text{LL} (N)-\text{LL} (0)); \text{where LL=} \sum Y \ln (p)+(1-Y) \ln (1-p)
\]

For evaluating predictive performance of the HOF models, the accuracy statistics were applied using the validation data set. Both the relative Root Mean Square Error (RMSEr) and the Root Mean Square Error (RMSE) was computed (Heiskanen, 2006, Makela and Pekkarinen, 2004, Muukkonen and Heiskanen, 2005).

\[
RMSE_r = \frac{RMSE}{\bar{y}} \times 100
\]

\[
RMSE = \sqrt{\frac{\sum_{i=1}^{n} (\hat{y}_i - y_i)^2}{n}}
\]

where \( \hat{y}_i \) is the modelled value, \( y_i \) is the observed value, \( \bar{y} \) is the mean of the observed values and \( n \) is the number of observation.
3. Results

3.1. Vegetation Properties and Canopy Density Modelling

3.1.1. Data Characteristics

Ten different mangrove sites were surveyed for percentage canopy cover. A dendrogram in figure 3A illustrates the similarities and differences of these sites in terms of canopy cover. It can be noted that there is an increasing trend towards a denser canopy and each sites overlaps with the other. Beside this, there were sites which were structurally similar, for instance, site 10 and 2, sites 3 and 4, and sites 5 and 9. Based on the field data, sites 5 and 9 are mangrove areas in the seaward zones characterized by stunted mangroves and thin canopy cover resulting from constant tidal inundation and erosion of substrate. The second group consist of sites 3 and 4, and sites 10 and 2, which are found on middle-ward mudflats on more sheltered estuaries. These consist of mature mangrove which showed some indication of human disturbance showing some canopy openings. The majority of the sites (7, 6, 8, and 1) which are found both in mudflats further inland or in coves and riverine areas are more pristine and consist of more intact and denser stand.
Statistical inspection of the percent canopy cover from these sites reveals that the frequency distribution is not normal (Shapiro’s W = 0.7442, p-value = 8.609e-11) and there is a significant positive skewness (Fig. 3B) even after transformation (median = 68%, mean=49.49%, n=84). The arcsine transformation however seems to minimize heteroscedasticity by pooling the observations towards the median.

![Figure 4 Three different types of mangrove habitats A) Seaward B) Middleward and C) Riverine mangroves in ASTER (bands 231) composite with corresponding photos.](image)

Variation in canopy cover could also be related to the difference in species composition aside from the influence of geographic location. Some of the common species recorded in the Sea-ward areas (Fig 4A) consists of *Avicennia marina*, and *Sonerratia alba*. Inland and Middle-ward mangroves (Fig. 4B) consist of mixture of *Sonneratia*, *Avicennia* and *Rhizophora*. The same species compositions were observed in riverine mangroves (Fig. 4C). These species have different canopy shapes and leaf structures which could influence canopy cover readings from site to site.

In relation to canopy cover, there are spectral differences among the three different types of mangroves communities (Fig. 5) which is very prominent in the NIR region. Seaward mangroves which have the thinnest canopy cover and stunted vegetation
have the lowest reflectance of about 18%. Likewise, the highest reflectance about 21% is recorded in Riverine mangroves. In between are the mangroves which are located in embayment or middle-ward zones which covers the whole 9-band spectrum. Overall, the 9 bands of ASTER have high affinity on mangrove vegetation characteristics specially the NIR band. From these features, canopy density could be theoretically modelled from in situ percent canopy cover and the combination of the VNIR and SWIR bands.

![Figure 5 Average spectral reflectance (in percentage) of the different mangrove communities.](image)

### 3.1.2. Non-parametric modelling with GAM

Non-parametric test are more appropriate for non-normally distributed data. In this case GAM models are used as exploratory in a regression model to test the relationship between canopy cover and canopy reflectance. Table 3 provides a statistical summary of the different models and Figure 6 show the results of the fitting. Four out of 9 bands showed significant relationship with canopy cover (F_{0.05}, 1, 82 d.f). Except for band 4 there was a significant non-linear relationship of the first 3 bands with percent canopy cover. This is indicated in the effective degrees of freedom (e.d.f) which are greater than 1 resembling a higher order polynomial. The resulting parametric terms are the same and are significantly different from 0 (p <0.001). Among the 4 bands, the NIR band has the highest correlation with percent canopy cover as indicated by the adjusted R^2 (a-R^2) and explained deviance (D^2) (69.4% and 70.5%, respectively), this is followed by the Red band (D^2=51% and a-R^2=49%). The SWIR band has the weakest correlation with canopy cover (a-R^2=32%). When the four bands are combined together in a single GAM model, the
a-R^2 and D^2 increase significantly making additive interaction explained about 86.4% (a-R^2) and 88.7% (D^2) of the variation in canopy cover. Although the smooth function for Band 4 is not significant, dropping it from the model causes 1 unit decrease in D^2 (87.7%).

Table 3 Statistical summary of the fitted GAM models for band reflectance and vegetation indices. Enclosed e.d.f. values are results from individual model fitting. Enclosed a-R^2 are linear model values.

<table>
<thead>
<tr>
<th>Bands</th>
<th>p-terms</th>
<th>R^2-adj</th>
<th>Deviance explained</th>
<th>GCV score</th>
<th>e.d.f</th>
<th>std. error</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Band1</td>
<td>0.495</td>
<td>0.414</td>
<td>43.2</td>
<td>0.0617</td>
<td>2.515 (3.447)</td>
<td>0.0279</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Band2</td>
<td>0.495</td>
<td>0.494</td>
<td>51.1</td>
<td>0.0591</td>
<td>2.828 (4.913)</td>
<td>0.0259</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Band3</td>
<td>0.495</td>
<td>0.694</td>
<td>70.5</td>
<td>0.0359</td>
<td>3.065 (5.191)</td>
<td>0.0202</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Band4</td>
<td>0.495</td>
<td>0.326</td>
<td>37.2</td>
<td>0.0816</td>
<td>5.625 (1.000)</td>
<td>0.0299</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Indices</th>
<th>p-terms</th>
<th>R^2-adj</th>
<th>Deviance explained</th>
<th>GCV score</th>
<th>e.d.f</th>
<th>std. error</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>PVI</td>
<td>0.495</td>
<td>0.777</td>
<td>78.7</td>
<td>0.0262</td>
<td>3.651</td>
<td>0.0172</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>SAVI</td>
<td>0.495</td>
<td>0.836</td>
<td>(0.777)</td>
<td>84.4</td>
<td>0.0195</td>
<td>4.102</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>TSAVI</td>
<td>0.495</td>
<td>0.828</td>
<td>(0.731)</td>
<td>83.7</td>
<td>0.0205</td>
<td>4.180</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>NDVI</td>
<td>0.495</td>
<td>0.834</td>
<td></td>
<td>84.2</td>
<td>0.0197</td>
<td>4.018</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>SAVI2</td>
<td>0.495</td>
<td>0.824</td>
<td></td>
<td>83.9</td>
<td>0.0216</td>
<td>6.962</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>RSR</td>
<td>0.495</td>
<td>0.552</td>
<td></td>
<td>57.7</td>
<td>0.0244</td>
<td>4.632</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>SR</td>
<td>0.495</td>
<td>0.748</td>
<td></td>
<td>76.3</td>
<td>0.0304</td>
<td>5.167</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>AVI</td>
<td>0.495</td>
<td>0.856</td>
<td></td>
<td>86.6</td>
<td>0.0173</td>
<td>5.445</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>DVI</td>
<td>0.495</td>
<td>0.855</td>
<td></td>
<td>(0.710)</td>
<td>86.6</td>
<td>0.0176</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>
GAM models are used in the same way for the vegetation indices. All of the vegetation indices have a significant positive non-linear relationship with the percent canopy cover (e.d.f. > 1) in terms of coefficient and the shape of the response curve (Table 3 and Fig.7). This implies that an increase in vegetation indices corresponds to increase in percent canopy cover. In general almost all the models have a good fit based on \( a-R^2 \) and \( D^2 \), but soil vegetation indices gave a better results and more realistic shapes. In particular PVI, SAVI and TSAVI have consistent monotonic shape with their corresponding \( a-R^2 \) greater 75\% (\( a-R^2=78.7\%, 84.4\%, 83.7\% \) respectively). NDVI also has a consistent monotonic response with high prediction power (\( a-R^2=83.4\% \)). Other indices like SAVI2, RSR and SR have high \( a-R^2 \) values but the shapes indicate a saturation problem as indicated by the plateau response. AVI and DVI, although they have high prediction power, the curve from these models are too complicated gives an indication of overfitting with the data which might be the reason for their high \( aR^2 \) values.

GAM has identified Band 3 and vegetation indices NDVI, SAVI, TSAVI and PVI as the most relevant in canopy cover modelling. It is also noticeable that this result is better than linear models in terms of \( a-R^2 \). However this is a non-parametric modelling, so next question is to test whether this relationship is consistent under a parametric non-linear and linear models.
Figure 7 Partial residual plots of the smooth components of the fitted GAM vegetation index models.
3.1.3. **Parametric modelling with HOF**

The result of the HOF models indicates that the most prominent and parsimonious type is Model II and III, which consists of 2 to 3 parametric estimates (Table 4), supporting the results generated from GAM. In general, the soil vegetation indices have better fitting which are very much consistent GAM results. DVI explained the highest variation in canopy density about 86%. Band 3 on the other hand explained only 71% of the canopy variations. Alternatively, TSAVI which explained about 84% of the canopy density variations is almost similar to NDVI (84.7%) in terms of predicting power. SAVI has also a high predicting power and PVI have the lowest predicting among the soil vegetation indices. Aside from individual a-R², models differ in terms or RMSE’s. HOF models II have higher relative RMSE (RMSEr) > 30% canopy cover than Model III.

Table 4 Statistics summary of selected parsimonious HOF fitted models.

<table>
<thead>
<tr>
<th>Band/Index</th>
<th>HOF/Model Type</th>
<th>Estimated parameters</th>
<th>RMSEr (%)</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>NDVI</td>
<td>II</td>
<td>a = 4.02, b = -5.49, c = NA, d = NA</td>
<td>0.802</td>
<td>33.86</td>
</tr>
<tr>
<td></td>
<td>III</td>
<td>a = 6.847, b = -12.330, c = -1.095, d = NA</td>
<td>0.847</td>
<td>2.457</td>
</tr>
<tr>
<td>TSAVI</td>
<td>II</td>
<td>a = 6.876, b = -8.525, c = NA, d = NA</td>
<td>0.844</td>
<td>19.438</td>
</tr>
<tr>
<td></td>
<td>III</td>
<td>a = 10.669, b = -14.485, c = -1.230, d = NA</td>
<td>0.824</td>
<td>3.762</td>
</tr>
<tr>
<td>SAVI</td>
<td>II</td>
<td>a = 3.97, b = -5.96, c = NA, d = NA</td>
<td>0.804</td>
<td>34.421</td>
</tr>
<tr>
<td></td>
<td>III</td>
<td>a = 6.882, b = -23.700, c = -0.945, d = NA</td>
<td>0.859</td>
<td>4.708</td>
</tr>
<tr>
<td>DVI</td>
<td>III</td>
<td>a = 10.669, b = -14.485, c = -1.230, d = NA</td>
<td>0.824</td>
<td>3.762</td>
</tr>
<tr>
<td></td>
<td>III</td>
<td>a = 6.882, b = -23.700, c = -0.945, d = NA</td>
<td>0.859</td>
<td>4.708</td>
</tr>
<tr>
<td></td>
<td>III</td>
<td>a = 3.97, b = -5.96, c = NA, d = NA</td>
<td>0.804</td>
<td>34.421</td>
</tr>
<tr>
<td></td>
<td>III</td>
<td>a = 6.882, b = -23.700, c = -0.945, d = NA</td>
<td>0.859</td>
<td>4.708</td>
</tr>
<tr>
<td>B3</td>
<td>III</td>
<td>a = 4.303, b = -11.162, c = -1.069, d = NA</td>
<td>0.716</td>
<td>6.370</td>
</tr>
<tr>
<td>PVI</td>
<td>III</td>
<td>a = 5.114, b = -12.860, c = -1.087, d = NA</td>
<td>0.801</td>
<td>1.949</td>
</tr>
</tbody>
</table>

Aside from the individual a-R² of the models, evaluation also considered the shape of the curves produced by the models in relation to the curve produced in GAM (Fig. 9). In all of the fitted HOF models the generalised functional shape of HOF Model II and III is consistent with the shape of GAM models. The prominent curve shape is monotonic. Although DVI has the highest predicting power it has a saturation problem as indicated by the extreme plateau of its response curve. PVI and Band 3 also suffered the same problem but less prominent. TSAVI, SAVI and NDVI gave the best results in terms prediction and of the shape of the response curve. However, a close inspection of the response curve of best fitted models indicates an underestimation of canopy density values. For instance Model III in the best models
(TSAVI, SAVI and NDVI) could only predict up to a maximum value of 70% which is unrealistic. The model II however could predict higher but with larger errors. Nonetheless, TSAVI, SAVI and NDVI produced the best fitted HOF Models. This implies that soil adjusted vegetation indices are the most useful indices in mangroves.

Figure 8 Comparison between HOF and GAM fitted curves for the selected parsimonious models.

In summary, the indices and Bands identified in GAM has a consistent relational behaviour to the % canopy cover under the parametric model. TSAVI, SAVI and NDVI are identified by HOF as the most relevant indices relating to the canopy cover. Also the relationship between canopy cover and satellite information is clearly non-linear. However there is a problem of prediction in terms of underestimation and errors when these models are used in predictive modelling directly, so the next
question is to test whether the addition of other satellite information in a Multiple Linear Regression improves prediction.

3.1.4. Multiple Linear Regression Modelling

Although the results from the GAM modelling indicate that majority of the relationship are non-linear with non-normal distribution of the response variable, a Multiple Linear Regression was still fitted under the assumption of Central Limit Theorem vis-à-vis vegetation continuum theory. This is to test whether the addition of other sensor variables improves linear relationship as well as the prediction. The results indicate that two significant models are possible. Using the result from HOF model, it was found that a combination of Bands 1, 2, 5, 6 with TSAVI could explain about 79.95/78.67% ($a-R^2/R^2$) of the variation in canopy density. The model coefficients are highly significant ($F=62.22$, 5, 78 D.F. $p<0.001$), with reduced errors (Residual SE=15.42, RMSEr = 7%, $e = 10.29$).

**Model 1:** $Y=13.040+57.256$(TSAVI) +6.470 (B1) -8.106(B2)-6.670 (B5) +10.071(B6)

On the other hand, when SAVI is combined with Band 5 and 6, 80.38%/79.65% ($R^2$ /$a-R^2$) of canopy density variations could be explained. All the coefficient of the second model are also highly significant ($F=109.3$, 3, 80 DF, $p<0.001$, residual SE=14.94, $e=3.644$).

**Model 2:** $Y = 4.914 + 59.264$(SAVI) -5.967(B5) + 7.939(B6)

NDVI has the same results as in SAVI so they could be used interchangeably in this particular case. Even when the models are highly significant there is still an evidence of non-constant variance and dispersion as indicated by the residual plots (Fig 8). Also the scatter plot indicates that the prediction cluster towards high values with less prediction in the lower values. The results do improve prediction significantly not necessarily in terms of increased $a-R^2$ but reduced errors significantly. So the addition of other satellite information to the indices identified in HOF in multiple linear regression models improved overall prediction.
3.1.5. Visualization in GIS

In creating the final maps, type II SAVI and TSAVI based HOF models were chosen together with MLR Models 1 and 2 for comparison. Types II HOF models are almost similar to linear models are the most parsimonious models and that they could not under predict but they have large prediction errors (outside 95% confidence interval) (Figure 10). SAVI map has an RMSeL equals to 33.86% and TSAVI has an RMSeL of 19.44%. TSAVI and SAVI map appears differently (Fig. 11A and 11B). The main difference is in the mapped canopy density variations. The TSAVI-based map offers more variability than the SAVI map. Besides the SAVI map is too generalized and overestimates sparse canopies. The TSAVI map on the other hand displays the gradual change in canopy density and sparse canopies are well mapped, hence it
offers sufficient generalization of canopy density. However, the predicted values of
the TSAVI models are lower hence they should not be used for quantification
studies, limiting their utility only to visual interpretations.

On the other hand the resulting maps of the two linear models are almost the same in
terms of canopy variations (Fig. 11C and Fig. 11D). It can be noted however that
sparse canopies are more distinct in the Model 2, whereas Model 1 provides more
information on denser canopies. Both model 1 and 2 have lower residual errors
compared to the previous HOF models (15.42\% and 14.94\%, respectively) and lower
RMSEr of 7\%. Overall the linear models are better than the 2 previous HOF models.

A comparison of the predictive capability of the models with the observed canopy
density indicates that the linear model prediction are closer to the observed canopy
density and the two HOF models either overestimate or underestimate prediction
(Fig. 12). All the standard error of the models is significant but HOF-TSAVI and
MLR model 1 have the lowest standard errors.
Figure 11 Comparison of the predicted canopy density maps (A) TSAVI-HOF II, (B) SAVI-HOF II (C) MLR Model 1 and (D) MLR Model 2

Figure 12 Comparison of the models with ground observation in terms of mean canopy densities and standard error.
3.2. **Understorey macro-benthos and cryptogams**

The mangrove understorey in the study area is characterized by the presence of algae, lichens, gastropods and bivalves. Some of the gastropod and bivalve species are important local food sources. A commonly observed characteristic of these species is their abundance and presence in dense and shady mangrove areas. So it is interesting to relate whether there is a dependency with respect to canopy cover. Two dominant lower plant species were observed, an unknown species of green lichen (Fig. 12A) and a species of a bottom alga common belonging to the genus *Chaetomorpha* (Fig. 12B). These 2 species are frequent in old growth and dense mangroves. For gastropods and bivalves four dominant species are observed; these are (Fig. 12C) Mangrove murex (*Chicarus capucinus*), (D) Sulcate cerith (*Terebralia sulcata*), (E) Coffee bean snail (*Melampus coffeus*) and (F) Common nerite (*Neritina communis*).

![Figure 13 Cryptogams and Macro-benthic and species found in mangroves](image)

The result of the logistic regression models indicated that the two species of cryptogams and 2 species of gastropods have a significant and consistent relationship with the observed and modelled canopy densities (F-test, p <0.001). A comparison of the $R^2$ of the various models indicates that the canopy density derived from linear model 1 gave the best results for all of the species (Table 5). So this model was used to compare the final species response curved for the four species with the observed canopy density. The resulting final model indicates that the relationship between the
modelled and observed canopy cover to the species distribution is highly significant except for the unknown lichen species (Table 6).

Table 5 Comparison of the logistic model $R^2$ for the four species distribution in the four canopy density models and the observed canopy density, NS indicate non-significance

<table>
<thead>
<tr>
<th>Species</th>
<th>Variable: Canopy Density</th>
<th>a</th>
<th>b</th>
<th>$R^2$ (%)</th>
<th>d.f.</th>
<th>$-2\times[LL(N)-LL(0)]$</th>
</tr>
</thead>
<tbody>
<tr>
<td>T. sulcata</td>
<td>Observed</td>
<td>-2.806</td>
<td>0.063</td>
<td>44.9</td>
<td>1</td>
<td>50.929</td>
</tr>
<tr>
<td></td>
<td>Modeled</td>
<td>-2.426</td>
<td>0.058</td>
<td>33.2</td>
<td>1</td>
<td>37.649</td>
</tr>
<tr>
<td>M. coffeus</td>
<td>Observed</td>
<td>-6.307</td>
<td>0.073</td>
<td>19.4</td>
<td>1</td>
<td>15.290</td>
</tr>
<tr>
<td></td>
<td>Modeled</td>
<td>-5.426</td>
<td>0.063</td>
<td>19.1</td>
<td>1</td>
<td>15.048</td>
</tr>
<tr>
<td>C. gracilis</td>
<td>Observed</td>
<td>-6.508</td>
<td>0.092</td>
<td>32.5</td>
<td>1</td>
<td>33.778</td>
</tr>
<tr>
<td></td>
<td>Modeled</td>
<td>-6.710</td>
<td>0.091</td>
<td>29.4</td>
<td>1</td>
<td>30.598</td>
</tr>
<tr>
<td>Lichen X</td>
<td>Observed</td>
<td>NS</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Modeled</td>
<td>-9.230</td>
<td>0.103</td>
<td>25.3</td>
<td>1</td>
<td>13.348</td>
</tr>
</tbody>
</table>

Figures 13 illustrates the simulated shape of the species response curve both to the modelled and observed canopy densities. It can be observed that the shapes of the response of the modelled species are consistent with the shape of the distribution in
the observed canopy density. For the two gastropod species two shapes are possible; a positive monotonic and a near symmetric and skewed bell-shaped curve. For the algae and the lichens the most prominent and consistent shape is positive monotonic where an increase in canopy density corresponds to high species presence. The addition of satellite information like band reflectance to the model also improved the model in terms of R2 squared however most of these bands are already used in modelling canopy density, thereby increasing the chance of autocorrelation aside from confounding the results so they are not further used in model which logically made sense because Canopy density model 2 was already derived from linear combination of TSAVI and some bands like 1, 2, 5 and 6.

Figure 14 HOF fitted response curve of the species distribution on observed and modelled canopy density gradients

The modelled distribution and presence probabilities of the lower plant species and macro-benthos in mangroves are consistent with ground observation. In addition, the canopy density from linear model with the most number of satellite variables provided the most significant improvement in the species models. And when it comes to shapes of the species distribution, the monotonic and skewed bell-shaped response are prominent.
4. Discussion

The result demonstrated that generalized models like GAM and HOF are very useful in simulating the true relationship between the observed canopy cover and satellite derived reflectance or indices. In effect, it gives sufficient ecological and statistical backing as to what information is relevant in predictive modelling of canopy density variations and species distribution. There are two clear advantages of using these models. First, anyone does not have to implement computationally rigorous methods like radiative transfer models or geometric optical models in finding the appropriate satellite information that could model canopy variation of a particular forest specially when highly accurate ground data are insufficient and secondly, the selection does not rely on just random selection of variables in conducting the regression analysis specially when no a priori information on the specifically relevant band or indices are available. A more informed modelling approach is more realistic and therefore more sensible.

GAM and HOF models provided a useful complement to multiple linear regression models particularly in selecting the ideal indices in the regression modelling process. The methods and results are promising however there are important implications of which are further discussed in the light of the research questions, here as outlined. Firstly, the nature of the environment, data, its accompanying assumptions and the constraint imposed in modelling approach; secondly, the performance of non-parametric and parametric, linear and non-linear models in depicting actual relationships. Thirdly, the essentiality of ecological knowledge in relating species distribution to environmental or vegetation variables and finally, the awareness of the advantages and limitations of the model and modelling technique especially with regards to its transferability or applicability to other areas and the future directions in using results from predictive models in mangrove monitoring.

4.1. Site biophysical characteristics and canopy cover

There is sufficient evidence why most vegetation properties are non-normally distributed. Site condition affects the biophysical feature of mangrove forest such as stand structure, species composition and canopy properties. In the three types of mangrove habitat surveyed, the seaward areas have sparser canopies and stunted vegetation compared to more sheltered areas in the riverine and middleward zones.
This is consistent with the observation of (Lugo and Snedaker, 1974) who reported that mangroves prefer areas with frequent seawater flushing and high nutrient concentration but extreme salinity as can be found in seaward areas. These factors also explain the classic zonation pattern (Smith and Huston, 1989) observed in the mangroves which accordingly have an important implication with respect to the position of mangroves along environmental gradients. Further, this phenomena can be explained as a fundamental niche response of a species (Austin and Smith, 1989). Mangroves exhibit zonation patterns, that when viewed at the species level corresponds to the community concept. However, they could also exist in wide environmental range as a continuum and in turn influences their observed physiological characteristics or response such as canopy cover. The range might be a combination of favourable indirect and resource gradients such as distance to shoreline and soil availability. Most of the surveyed areas are located in coves, middleward zones and riverine areas. As such the canopy cover is expected to shift towards high values causing significant skewness in the data distribution. In addition even within these sites there is variability due to species composition causing non-normality in the data. Non-normality and non-linearity of ecological data relationships are not artefacts.

Zonation, species composition, age class and forest structures, drive variability in canopy density which occurs continuously with no distinct boundaries a phenomenon consistent with the continuum theory of vegetation. Statistically, this variation causes the data distribution to deviate from normality resulting in a significant skewness even after transformation or standardization. Though controversial, this is a common characteristics of vegetation and conformed with the observation of Austin (2007) and Guisan et al (2002). Two issues arise in relation this, first, what types of model are appropriate for ecological data with non-normal distribution and what satellite information is relevant as input in the model. When the data, statistically speaking, have non-normal distribution and non-constant variance, the assumptions of ordinary least squares modelling are violated and renders the use of linear modelling approaches questionable. On one hand, there is no vegetation index for mangroves ideal for mangroves.
4.2. **Relationship between band reflectance, vegetation index and canopy cover**

If linearity is imposed without exploratory analysis of the actual relationship between the variables then the chance of losing valuable information is high. Also it is a clear violation of the established assumption of linearity knowing that the data are non-normal in the first place. In this instance, non parametric models like GAM and parametric non-linear models like HOF are useful. By applying a non-parametric smooth function (Wood and Augustin, 2002) GAM does not constrain the relationship to linearity hence it was able to detect the actual relationships of canopy cover to the different satellite information. In HOF a non-linear logistic regression fits a curvilinear function to the data. This is the reason why GAM is successful in several ecological and forestry models (Moisen and Frescino, 2002) which are recently developed for integration with GIS (Lehmann, 1998, Lehmann et al., 2002, Austin et al., 2006). HOF on the other hand was proven to be useful in modelling species response. Applying GAM and HOF as the exploratory model for determining the relationship of canopy cover and vegetation reflectance provided useful insights as to what band and indices are relevant.

4.2.1. **GAM Modelling**

The result of GAM allowed the rejection of the null hypothesis in favour of the alternative hypothesis which implies that there is enough evidence to conclude that the relationship between canopy cover and vegetation reflectance is indeed non-linear. The nonlinearity between stand attributes and reflectance values was reported by Song et al.(2002) and lately by Hall (2006) and Heiskanen (2006). Both authors used radiative transfer/geometric optical models and exponential regression techniques. Among the bands, GAM has detected that the NIR band (Band 3) have the strongest positive correlation with canopy cover ($a-R^2=70\%$) which implies that this band can be use to model canopy density. Similarly, the NIR band is related to canopy vigour, health, productivity and canopy woody tissues like mesophyll. Mangroves like any other salt water plants are characterized by thick layer of spongy mesophyll which could contain large quantities of water does making them highly reflective in the NIR band.

The affinity of NIR band to canopy cover was founded by studies in mixed temperate forest by Hall et al. (2006) and boreal forest by Falkowski et al. (2005) using linear regression for Landsat and ASTER derived reflectance, respectively. Other bands were also highly correlated with canopy cover in a negative relationship like for
instance Band 1 and 2. This implies strong chlorophyll absorption and that the vegetation at the time of image acquisition is photosynthetically active. Xu et al. (2003) found that Landsat Band 3 (Red) is highly positively correlated with canopy cover of savannah oak during its senescent stage. In another study by Rocchio et al. (2003), Landsat band 5 and 7 are found to be a significant predictor of forest parameter when combined with different indices. In this study, a weak correlation is observed for the SWIR bands. Overall, the result from GAM is consistent with other studies on vegetation cover employing more sophisticated methods.

All of the vegetation indices fitted with GAM also showed significant curvilinear relationship with percent canopy cover. But in general, the soil vegetation index like TSAVI, SAVI and PVI fitted better than the indexes based on simple differencing and ratioing, with the exception of NDVI. This relationship is consistent with the result obtained by (McDonald et al., 1998) in conifer forests using geometric optical model and Monte-Carlo ray-tracing and Diaz and Blackburn (2003) in a simulated laboratory condition for LAI measurement. NDVI, though has a problem of saturation (Huete et al., 1997) in other studies, did not show here. This may imply that mangrove NDVI in these areas were well within the effective NDVI range. The non-linear trajectories are also believed to be caused by vegetation succession and the presence of understorey. Understorey is an important factor in modelling canopy properties from reflectance because of its non-uniformity as Hu et al., (2000) observed. In mangroves, the presence of understorey vegetation is prominent especially in sea ward areas and within the forest edges because of regeneration.

Another important cause of variability and influencing factor for the canopy cover and vegetation reflectance is the effect of background. McDonald et al. (1998) elaborated that SAVI, TSAVI and PVI are less sensitive to background perturbation in medium to dense vegetations, thus no saturation is observed in the resulting curve for these indices. The importance of background in modelling canopy cover was emphasized by (Chen and Cihlar, 1996) in modelling LAI and canopy closure in boreal forest. SR and SAVI which showed extreme saturation effect by having the most prominent plateau response is consistent with Diaz and Blackburn’s (2003) results. They further elaborated that DVI is the least affected index. However, their result in this aspect seems inconclusive and inconsistent because it is on laboratory condition and that they used a different type of plant (Gardenia jasminoides) which is not a mangrove species. The result from this research showed that DVI exhibit saturation and variability to some extent. Background effects are prominent in
mangroves where the substrate could vary from sandy to clay loam so it is not surprising that the soil based vegetation indices like SAVI, TSAVI and PVI are the most useful and realistic models generated from GAM because these indices has almost constant sensitivity to background (Diaz and Blackburn, 2003) and have a large dynamic ranges (McDonald et al., 1998).

The result in GAM identified six possible vegetation models from initial 13 indices and bands, which are better fitting to the canopy cover data. These are TSAVI, SAVI, PVI, NDVI, DVI and Band 3. Yee and Mitchell (1991) suggested that whenever possible a parametric model is preferred over their non-parametric counterpart. Parameterization is necessary for simplicity, generality and easy integration in a GIS. GAM being a non-parametric and data driven model could result in overfitting, limiting its applicability in other areas at other vegetation conditions. HOF model offer both the flexibility and generality of GAM as well its capacity to calculate parametric terms by fitting several non-linear logistic regression models with the data. The result from GAM was used in HOF models.

4.2.2. HOF Modelling

Of the six covariates fitted with HOF the soil vegetation indices gave the best results together with the NDVI. The prominent model types are II (monotonic) and III (plateau). When interpreted in ecological terms, this is realistic and consistent with vegetation limiting factor response with optima (Clymo, 1995, Austin and Austin, 1980) and fundamental niche response (Austin and Smith, 1989). Similarly, canopy density cannot have values lower than 0% and higher than 100%, hence, the relationships are much closer to reality. Although the adjusted coefficient of determination ($a-R^2$) of the different models is high, not all of them have realistic fit in terms of shape of the relationship curve. DVI, PVI and Band 3 still showed saturation consistent with GAM. This phenomenon is also reported in other vegetation studies (Johnson, 2001). The result of the HOF exploratory analysis indicates that NDVI, TSAVI and SAVI are the most useful indices that fit well with the canopy cover data in the type II and type III HOF models which are the most parsimonious types. Both models have high coefficients of determination. TSAVI and SAVI are consistently better both in GAM and HOF models.

It is possible to use the models for direct prediction of canopy density for the whole image. However, there are limitations for their direct usage. First, even with parsimony and high $a-R^2$, type II HOF models could underestimate and overestimate prediction. Type II SAVI and TSAVI models have these features, respectively. This
is a natural tendency of data driven models because the prediction depends very much on the accuracy of the field data. Type II models also have high relative errors and are outside the 95% confidence interval of the GAM models. Type III models on the other hand although with lesser errors and are within the 95% confidence tend to underestimate prediction to an optimum of 70% canopy density which is not realistic based on the field data. The reasons for these inconsistencies are not on the models per se but most probably on the accuracy of the ground data (Chen and Cihlar, 1996). Fiala et al. (2006) implied that using spherical densiometer instrument for canopy cover has a questionable accuracy and the readings could subjective. Nevertheless, the models and method are equally robust for identifying relevant vegetation indices closely related to canopy cover.

4.2.3. Multiple Linear Regression Modelling

Fitting a linear model to a nor-normal data is only possible under the assumption of Central Limit Theorem and Continuum Theory. Many studies that implemented linear regression (Hu et al., 2000, Chen and Cihlar, 1996, Moskal and Franklin, 2004, Rocchio et al., 2003) have obtained improved results for crown closure modelling but made no explicit mention of the normality of the data. The addition of other satellite band reflectance to the SAVI and TSAVI indices in a multiple linear regression produced 2 significantly improved models with almost similar predictive power but differ in number of parameters. The first model which includes TSAVI was improved when the green and red band are added together with the 2 short infrared bands 5 and 6. Bands green and red are chlorophyll absorption bands which at highly photosynthetic vegetation provides strong absorption signal, on the other hand SWIR bands 5 and 6 are related to canopy moisture content and soil organic carbon (Stephens, 2007) which is also related to the vegetation, therefore their contribution to the overall prediction of canopy density is significant. This is also valid in the in the case of the second model involving SAVI.

In terms of parsimony and visual quality of the resulting model and map, the SAVI-based model seems to be the optimum for canopy density prediction. Further testing using NDVI reveals that the result is similar to SAVI so they can be used interchangeably in this case. Overall, it was clear that the addition of other satellite information improved linearity and prediction of the model and there is an informed basis for imposing linear relationship and selecting the ideal indices. On the other hand, the use of ASTER has proven to be effective in predicting mangrove canopy density having a multi-spectral resolution especially in the SWIR regions.
4.3. **Species Response to canopy density gradients**

The result from the logistic regression confirmed existing ecological theories that the symmetric bell-shape response curve of species to different environmental gradients are not universal. All of the modelled responses are significantly correlated with the canopy densities from the both actual observation and modelled canopy density. Here, two significant results are highlighted with their important implications. First, the consistency of the resulting model for both the observed canopy density and modelled implies that macro-benthic species and cryptogam distribution could be monitored using satellite derived canopy densities. In this case, the presences of both species are optimized at around 50% canopy cover. So any variation in canopy cover beyond this value is expected impact species distribution either positively or negatively. This, to a limited extent, is an innovative approach in addressing the issue of difficult monitoring of canopy densities and species distribution in mangroves which was never attempted before. Second, most of the improvement in the species distribution model in terms of coefficient of determination, log likelihood deviance and shape of species response curve was accrued to the modelled canopy density from model 1, the model that has the most number of parameters. This implies that the two models have different usefulness. The first model is important in relating to species distribution and the second model solely for predicting and mapping canopy density. Thus, both models are equally important.

The result from the model when relating to species distribution must be interpreted with caution as James and McCulloch (1990) emphasized that high correlation does necessarily imply causation. As such the knowledge on the ecological behaviour of the species is important.

In the case of the gastropods, the presence and distribution of *Terebralia sulcata’s* and *Melampus coffeus* could be describe by two models; a monotonic response and a skewed bell-shaped response. A study by (Houbrick, 1991) noted that the Asian population of *T. sulcata* can occupy wide range of habitats within the coastal zone and are resistant to desiccation so dependency on mangroves for shade is ruled out. However, the species are highly dependent on algae and mangrove biomass for food. This might explain the monotonic response of the species that is positively correlated with increasing canopy density. A related species *T. palustris* found in Kenyan mangroves are more susceptible to desiccation, so they are more sensitive to changes in the canopy. *Terebralia* species play a significant role in mangrove leaf litter removal and processing (Slim et al., 1997, Fratini et al., 2001).
The second model implies a decreasing presence with increasing vegetation. Some research indicates that this relationship is highly possible. Matthijs et al (1999) and Kryger and Lee (1996) observed an increase in the accumulation of hydrogen sulphide, methane, ammonia and soil acidity in dense mangroves. These compounds in excess amounts are found to be detrimental to some gastropod species as recorded by Hiong Kum et al. (2004) regarding the response of a mangrove-related bivalve species Polymesoda expansa. On the other hand, M. coffeus is observed to have preference on shady and muddy mangrove areas (Smithsonian Institute, 2007). This species is a climber species which can stay above the mangrove forest floor during high tide, thus, it is unlikely to be affected by acidity and sulphide. As such, the monotonic response model is appropriate for this species, however, no sufficient information on the species autecology exist to verify this.

The shape of the species response curve of the algae and lichen are consistent monotonic which is positively correlated with increasing vegetation. Algae are susceptible to drying out, so they proliferate in dense vegetation where the moisture condition is ideal. To some extent they require sufficient amount of light especially for photosynthesis. There are no existing studies on the autecology of these species. Although, Chaetomorpha gracilis, a relative species was observed in Asia and Mediterranean, studied by Nabivailo et al. (2005) and Menendez (2005) is not mangrove related. Lichen is more resistant to desiccation so they can tolerate various kinds of canopy light and moisture conditions. The observed positive correlation with canopy cover might be a symbiosis in terms of physical support where they are common even in partially opened canopies but they are most frequent in older stands.

4.4. **Model Uncertainties and sources of errors**

Although the results of the models are promising and ecologically sensible, the existences of error both in the field and satellite data, analysis and the methods cannot be discounted nor their implication be taken for granted. One source of error is the 1 year lag and seasonal difference between the image acquisition and data gathering because fieldwork was conducted during the rainy season. The accuracy of the spherical densiometer is also of concern, since reading was difficult under a dense canopy and there is a tendency to overestimate canopy cover and underestimate undergrowth. Reading could also become variable from person to person and easily affected by external factors. When it comes to macro-benthic data, there is no systematic sampling method. Lastly, there is no accurate validation data to the results of the model.
5. Conclusion and Recommendation

Integration of remote sensing data and ecological knowledge is possible in the context of ecological modelling of vegetation properties and its components. This offers a productive and synergistic effects as well as sensibility in dealing with complex ecosystems monitoring such as mangroves. Because ecological knowledge is an important part of management and conservation of mangroves, ecological modelling provides an opportunity of coupling together ecological principles with spatial data. This gives managers and conservationist an edge over the impracticality and uncertainty imposed by the reliance on field measurement, thus, resulting to apt generation of useful mangrove information.

The modelling activity accomplished by this research has generated some outcomes with insights into the usefulness and applicability of spatial-ecological modelling particularly for mangroves. First, negatively skewness and non-normality is a characteristic of mangrove canopy density which conforms to existing ecological theories of vegetation as a continuum. Second, the relationship between canopy density and vegetation reflectance is non-linear and models like GAM and HOF are effective compared to linear models in detecting these relationships. With these, TSAVI, SAVI and NDVI were determined to be the useful mangrove vegetation indices. Thirdly, linear models improved prediction by taking into account additional variables such as bands to 1, 2, 5 and 6 in addition to the indices identified by HOF and GAM. Fourth, the spatial and spectral properties of ASTER are important especially with fragmented mangrove areas and highly variable canopy reflectance. Finally, species response curves of mangrove associated cryptogams and macro-benthos varies and is significantly correlated with canopy density changes. This has an important implication especially to the current management approach in mangrove reforestation which is biased towards monoculture resulting in uniform species, stand structure and canopy densities. It is strongly recommended that this outcome be tested using different sensors or more accurate field instruments in different settings to see if there are any divergences and improvements.

Ultimately, an important lesson in modelling particularly when dealing with ecological systems is that correlation does not necessarily mean causation so determining the actual shape of species response curve and the expected result from an ecological model requires knowledge on the species autecology and of ecological theory.
6. References


7. Appendices

7.1. Some mangrove species in the area

- *Rhizophora mucronata*
- *Avicennia marina*
- *Ceriops decandra*
- *Sonneratia alba*
7.2. HOF Model Evaluation

<table>
<thead>
<tr>
<th>Index</th>
<th>NDVI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Model</th>
<th>R²</th>
<th>a-R²</th>
<th>d.f</th>
<th>RMSEr</th>
<th>F-critical</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>0</td>
<td>0.000</td>
<td>83</td>
<td>7.001</td>
<td>0.05</td>
<td>84</td>
</tr>
<tr>
<td>II</td>
<td>0.803</td>
<td>0.714</td>
<td>82</td>
<td>33.859</td>
<td></td>
<td></td>
</tr>
<tr>
<td>III</td>
<td>0.85</td>
<td>0.848</td>
<td>81</td>
<td>2.457</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IV</td>
<td>0.844</td>
<td>0.769</td>
<td>81</td>
<td>9.626</td>
<td></td>
<td></td>
</tr>
<tr>
<td>V</td>
<td>0.85</td>
<td>0.852</td>
<td>80</td>
<td>2.002</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>MODEL</th>
<th>Tested Against</th>
<th>F-comp</th>
<th>K</th>
<th>d.f</th>
<th>F-tab</th>
<th>S/NS</th>
</tr>
</thead>
<tbody>
<tr>
<td>IV</td>
<td>I</td>
<td>219.115</td>
<td>2</td>
<td>81</td>
<td>3.109</td>
<td>S</td>
</tr>
<tr>
<td>IV</td>
<td>II</td>
<td>21.026</td>
<td>1</td>
<td>81</td>
<td>3.959</td>
<td>S</td>
</tr>
<tr>
<td>IV</td>
<td>III</td>
<td>0.000</td>
<td>0</td>
<td>0</td>
<td>0.000</td>
<td>NA</td>
</tr>
<tr>
<td>V</td>
<td>IV</td>
<td>3.280</td>
<td>1</td>
<td>80</td>
<td>3.960</td>
<td>S</td>
</tr>
<tr>
<td>V</td>
<td>III</td>
<td>0.000</td>
<td>1</td>
<td>80</td>
<td>3.960</td>
<td>S</td>
</tr>
<tr>
<td>III</td>
<td>II</td>
<td>25.693</td>
<td>1</td>
<td>81</td>
<td>3.959</td>
<td>S</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Model</th>
<th>DF</th>
<th>Deviance</th>
<th>Chi</th>
<th>S/NS</th>
</tr>
</thead>
<tbody>
<tr>
<td>V</td>
<td>4</td>
<td>1.486</td>
<td>9.49</td>
<td>NS</td>
</tr>
<tr>
<td>IV</td>
<td>3</td>
<td>1.552</td>
<td>7.81</td>
<td>NS</td>
</tr>
<tr>
<td>III</td>
<td>3</td>
<td>1.486</td>
<td>7.81</td>
<td>NS</td>
</tr>
<tr>
<td>II</td>
<td>2</td>
<td>1.931</td>
<td>5.99</td>
<td>NS</td>
</tr>
<tr>
<td>I</td>
<td>1</td>
<td>9.257</td>
<td>3.84</td>
<td>S</td>
</tr>
<tr>
<td>Model</td>
<td>SAVI</td>
<td>Index</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td>------</td>
<td>-------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I</td>
<td>0</td>
<td>0.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>II</td>
<td>0.804</td>
<td>0.802</td>
<td></td>
<td></td>
</tr>
<tr>
<td>III</td>
<td>0.853</td>
<td>0.850</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IV</td>
<td>0.847</td>
<td>0.843</td>
<td></td>
<td></td>
</tr>
<tr>
<td>V</td>
<td>0.853</td>
<td>0.848</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Model</th>
<th>Testes</th>
<th>Against</th>
<th>F-comp</th>
<th>K</th>
<th>d.f</th>
<th>F-tab</th>
<th>S/NS</th>
</tr>
</thead>
<tbody>
<tr>
<td>IV</td>
<td>I</td>
<td>224.206</td>
<td>2</td>
<td>81</td>
<td>3.109</td>
<td>S</td>
<td></td>
</tr>
<tr>
<td>IV</td>
<td>II</td>
<td>22.484</td>
<td>1</td>
<td>81</td>
<td>3.959</td>
<td>S</td>
<td></td>
</tr>
<tr>
<td>IV</td>
<td>III</td>
<td>0.000</td>
<td>0</td>
<td>0</td>
<td>0.000</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>V</td>
<td>IV</td>
<td>3.347</td>
<td>1</td>
<td>80</td>
<td>3.960</td>
<td>NS</td>
<td></td>
</tr>
<tr>
<td>V</td>
<td>III</td>
<td>0.000</td>
<td>1</td>
<td>80</td>
<td>3.960</td>
<td>NS</td>
<td></td>
</tr>
<tr>
<td>III</td>
<td>II</td>
<td>27.333</td>
<td>1</td>
<td>81</td>
<td>3.959</td>
<td>S</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Model</th>
<th>DF</th>
<th>Deviance</th>
<th>chi</th>
<th>S/NS</th>
</tr>
</thead>
<tbody>
<tr>
<td>V</td>
<td>4</td>
<td>1.462</td>
<td>9.49</td>
<td>S</td>
</tr>
<tr>
<td>IV</td>
<td>3</td>
<td>1.535</td>
<td>7.81</td>
<td>S</td>
</tr>
<tr>
<td>III</td>
<td>3</td>
<td>1.461</td>
<td>7.81</td>
<td>S</td>
</tr>
<tr>
<td>II</td>
<td>2</td>
<td>1.922</td>
<td>5.99</td>
<td>S</td>
</tr>
<tr>
<td>I</td>
<td>1</td>
<td>9.257</td>
<td>3.84</td>
<td>NS</td>
</tr>
<tr>
<td>Index</td>
<td>TSAVI</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td>-------</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Model</th>
<th>R²</th>
<th>a-R²</th>
<th>d.f</th>
<th>RMSEr</th>
<th>F-critical</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>0.000</td>
<td>0</td>
<td>83</td>
<td>7.001</td>
<td>0.05</td>
<td>84</td>
</tr>
<tr>
<td>II</td>
<td>0.826</td>
<td>0.824</td>
<td>82</td>
<td>19.438</td>
<td></td>
<td></td>
</tr>
<tr>
<td>III</td>
<td>0.847</td>
<td>0.844</td>
<td>81</td>
<td>3.762</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IV</td>
<td>0.849</td>
<td>0.846</td>
<td>81</td>
<td>0.927</td>
<td></td>
<td></td>
</tr>
<tr>
<td>V</td>
<td>0.850</td>
<td>0.844</td>
<td>80</td>
<td>2.042</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>MODEL</th>
<th>Tested Against F-comp</th>
<th>K</th>
<th>d.f</th>
<th>F-tab</th>
<th>S/NS</th>
</tr>
</thead>
<tbody>
<tr>
<td>IV</td>
<td>I</td>
<td>227.712</td>
<td>2</td>
<td>81</td>
<td>3.109</td>
</tr>
<tr>
<td>IV</td>
<td>II</td>
<td>12.185</td>
<td>1</td>
<td>81</td>
<td>3.959</td>
</tr>
<tr>
<td>IV</td>
<td>III</td>
<td>0.000</td>
<td>0</td>
<td>0</td>
<td>0.000</td>
</tr>
<tr>
<td>V</td>
<td>IV</td>
<td>0.547</td>
<td>1</td>
<td>80</td>
<td>3.960</td>
</tr>
<tr>
<td>V</td>
<td>III</td>
<td>1.640</td>
<td>1</td>
<td>80</td>
<td>3.960</td>
</tr>
<tr>
<td>III</td>
<td>II</td>
<td>11.255</td>
<td>1</td>
<td>81</td>
<td>3.959</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Model</th>
<th>DF</th>
<th>Deviance</th>
<th>chi</th>
<th>S/NS</th>
</tr>
</thead>
<tbody>
<tr>
<td>V</td>
<td>4</td>
<td>1.5</td>
<td>9.49</td>
<td>S</td>
</tr>
<tr>
<td>IV</td>
<td>3</td>
<td>1.505</td>
<td>7.81</td>
<td>S</td>
</tr>
<tr>
<td>III</td>
<td>3</td>
<td>1.5098</td>
<td>7.81</td>
<td>S</td>
</tr>
<tr>
<td>II</td>
<td>2</td>
<td>1.212</td>
<td>5.99</td>
<td>S</td>
</tr>
<tr>
<td>I</td>
<td>1</td>
<td>9.257</td>
<td>3.84</td>
<td>NS</td>
</tr>
</tbody>
</table>
7.3. **R Script**

-------------
### Task: Canopy Density Modelling
### Period: November-February, 2007
### Creator: Dante Torio
-------------
## Load data sets
# Canopy Cover, Band Reflectance and Vegetation Indices
canopy <- read.csv("D:\THESISf\Analysis \Statistics\newdata\nnN.csv")
# Species Data
grasp <- read.csv("D:\THESISf\Analysis \Statistics\newdata\logress.csv")
# Canopy cover average
hc <- read.csv("D:\THESISf\Analysis\Statistics\newdata\hc.csv")

## Attach data sheets
attach(canopy)
attract(gradsp)
attach(hc)

## Load required libraries
library(mgcv)  ## GAM (Woods and Augustin, 2002)
library(gravy)  ## HOF Model (Oksanen and Minchin, 2002)
library(gplots)  ## for Graphics

## Data Summaries
summary(canopy)
summary(bands)

## Linearity Test
hist(tCANOPY, breaks=30, freq=T,
     main="Percent canopy cover frequency distribution",
     xlab="Canopy Cover(%)")
lines(frequency(tCANOPY))
rug(tCANOPY)
shapiro.test(tCANOPY)
plot(ecdf(tCANOPY))

## Construct Dendrogram
hc1 <- hclust(dist(hc), "ave")
(dend1 <- as.dendrogram(hc1)) # "print()" method
plot(dend1, nodePar=list(pch = 2:1, cex=.4*2:1, col = 2:3), horiz=TRUE,
     xlab="Average Distance",
     ylab="Observation Sites", main= "Site Canopy Similarity")
### GAM Modelling

#### 1. Band 4321 reflectance vs. canopy percentage

```r
gam1 <- gam(tCANOPY~s(NDVI))
gam2 <- gam(tCANOPY~s(SAVI))
gamband4 <- gam(tCANOPY~s(TSAVI))
gamband1234 <- gam(tCANOPY~s(B3))
summary(gamband**)
plot(gamband**)```

#### HOF Modelling

#### 1. NDVI, DVI, SAVI, TSAVI, PVI,

```r
hofavi <- HOF(tCANOPY, TSAVI,1, family=gaussian)
hofavi
ftavi1 <- fitted(hofavi, model=5)
ftavi2 <- fitted(hofavi, model=4)
ftavi3 <- fitted(hofavi, model=3)
ftavi4 <- fitted(hofavi, model=2)
ftavi5 <- fitted(hofavi, model=1)
plot(hofavi)
```

```
RMSE5 <- sqrt(sum(ftavi5-vCANOPY)^2/84)
RMSE4 <- sqrt(sum(ftavi4-vCANOPY)^2/84)
RMSE3 <- sqrt(sum(ftavi3-vCANOPY)^2/84)
RMSE2 <- sqrt(sum(ftavi2-vCANOPY)^2/84)
RMSE1 <- sqrt(sum(ftavi1-vCANOPY)^2/84)
```

```
rRMSE5 <- (RMSE5/mean(vCANOPY))*100
rRMSE4 <- (RMSE4/mean(vCANOPY))*100
rRMSE3 <- (RMSE3/mean(vCANOPY))*100
rRMSE2 <- (RMSE2/mean(vCANOPY))*100
rRMSE1 <- (RMSE1/mean(vCANOPY))*100
```

```
MA5 <- (1-RMSE5/mean(vCANOPY))*100
MA4 <- (1-RMSE4/mean(vCANOPY))*100
MA3 <- (1-RMSE3/mean(vCANOPY))*100
MA2 <- (1-RMSE2/mean(vCANOPY))*100
MA1 <- (1-RMSE1/mean(vCANOPY))*100
```

```
Rs5 <- sum((vCANOPY-ftavi5)^2)/(sum((vCANOPY-mean(ftavi5))^2))
Rs4 <- sum((vCANOPY-ftavi4)^2)/(sum((vCANOPY-mean(ftavi4))^2))
Rs3 <- sum((vCANOPY-ftavi3)^2)/(sum((vCANOPY-mean(ftavi3))^2))
Rs2 <- sum((vCANOPY-ftavi2)^2)/(sum((vCANOPY-mean(ftavi2))^2))
Rs1 <- sum((vCANOPY-ftavi1)^2)/(sum((vCANOPY-mean(ftavi1))^2))
```

```
bias5 <- sum(ftavi5-vCANOPY)/100
bias4 <- sum(ftavi4-vCANOPY)/100
bias3 <- sum(ftavi3-vCANOPY)/100
bias2 <- sum(ftavi2-vCANOPY)/100
bias1 <- sum(ftavi1-vCANOPY)/100
```

```
R5 <- (1-Rs5)
R4 <- (1-Rs4)
R3 <- (1-Rs3)
R2 <- (1-Rs2)
R1 <- (1-Rs1)
```

```
aR5 <- (1-{(100-1)/(100-4)})*(1-R5)
```
MODELLING MANGROVE CANOPY DENSITY VARIATIONS

\[ aR4 \leftarrow (1-\frac{(100-1)}{(100-3)})(1-R4) \]
\[ aR3 \leftarrow (1-\frac{(100-1)}{(100-3)})(1-R3) \]
\[ aR2 \leftarrow (1-\frac{(100-1)}{(100-2)})(1-R2) \]
\[ aR1 \leftarrow (1-\frac{(100-1)}{(100-1)})(1-R1) \]

#### Multiple Linear Regression Model

```r
par(mfrow=c(2,2))
## Model 1
cML1 <- lm(CANOPY~TSAVI+B2+B5+B6)
summary(cML1)
anova(cML1)
step(cML1, direction="backward")
plot(cML1$fit, cML1$res, main = "MLR Model 1", xlab = "Fitted", ylab="Residuals")
abline(h=0)
abline(v=0)
plot(cML1$fit, abs(cML1$res),xlab="Fitted", ylab="|Residuals|")
summary(lm(abs(cML1$res)-cML1$fit))
pred<- predict.lm(cML1, interval="confidence")
```

### Plotting the Results

```r
plot(tCANOPY~SAVI, main="Modelled Canopy Density with SAVI")
gamband2 <- gam(tCANOPY~s(SAVI))
fit3 <- predict.gam(gamband2, se.fit=T)
lwr <- fit3$fit-fit3$se.fit
uppr <- fit3$fit+fit3$se.fit
lines(spline(ftavi2~SAVI), col="blue", lty=3, type="l", lwd=3)
lines(spline(fit3$fit~SAVI), col="blue", lty=1, type="l")
lines(spline(lwr~SAVI), col="red", lty=2)
lines(spline(uppr~SAVI), col="red", lty=2)
lines(spline(fit3$fit~SAVI), col="green", lty=1, type="l", lwd=3)
```

```r
rug(tCANOPY)
smartlegend(x="left", y="top",
c("HOF II", "GAM", "Upper 95% Confidence", "Lower 95% Confidence", "HOF III"),
col=c("blue", "blue", "red", "red", "green"),
linetype=c(3,1,2,2,1), lwd=2)
```

```r
plot(tCANOPY~TSAVI, main="Modelled Canopy Density with TSAVI")
gamband2 <- gam(tCANOPY~s(TSAVI))
fit3 <- predict.gam(gamband2, se.fit=T)
lwr <- fit3$fit-fit3$se.fit
uppr <- fit3$fit+fit3$se.fit
lines(spline(ftavi2~TSAVI), col="blue", lty=3, type="l", lwd=3)
lines(spline(fit3$fit~TSAVI), col="blue", lty=1, type="l")
lines(spline(lwr~TSAVI), col="red", lty=2)
lines(spline(uppr~TSAVI), col="red", lty=2)
lines(spline(fit3$fit~TSAVI), col="green", lty=1, type="l", lwd=3)
```

```r
rug(tCANOPY)
smartlegend(x="left", y="top",
c("HOF II", "GAM", "Upper 95% Confidence", "Lower 95% Confidence", "HOF III"),
col=c("blue", "blue", "red", "red", "green"),
linetype=c(3,1,2,2,1), lwd=2)
```
MODELLING MANGROVE CANOPY DENSITY VARIATIONS

-----
Species Modelling
-----
sp1 <- HOF(TSULC, CANOPY, 1, family=gaussian)
sp1
tfavi11 <- fitted(sp1, model=5)
tfavi12 <- fitted(sp1, model=4)
tfavi13 <- fitted(sp1, model=3)
tfavi14 <- fitted(sp1, model=2)
tfavi15 <- fitted(sp1, model=1)
sp2 <- HOF(MLCOF, CANOPY, 1, family=gaussian)
sp2
tfavi21 <- fitted(sp2, model=5)
tfavi22 <- fitted(sp2, model=4)
tfavi23 <- fitted(sp2, model=3)
tfavi24 <- fitted(sp2, model=2)
tfavi25 <- fitted(sp2, model=1)
sp3 <- HOF(ALGA1, CANOPY, 1, family=gaussian)
sp3
tfavi31 <- fitted(sp3, model=5)
tfavi32 <- fitted(sp3, model=4)
tfavi33 <- fitted(sp3, model=3)
tfavi34 <- fitted(sp3, model=2)
tfavi35 <- fitted(sp3, model=1)
sp4 <- HOF(LCN, CANOPY, 1, family=gaussian)
sp4
tfavi41 <- fitted(sp4, model=5)
tfavi42 <- fitted(sp4, model=4)
tfavi43 <- fitted(sp4, model=3)
tfavi44 <- fitted(sp4, model=2)
tfavi45 <- fitted(sp4, model=1)
plot(sp1)
plot(sp2)
plot(sp3)
plot(sp4)

par(mfrow=c(2,1))
plot(TSULC~CANOPY, xlab="Observed Canopy Density (%)",
     ylab="Presence Probability", main="Terebralia sulcata")
lines(spline(tfavi21~CANOPY), col="blue", lty=1, type="l", lwd=3)
lines(spline(tfavi41~CANOPY), col="red", lty=2, type="l", lwd=3)
smartlegend(x="left", y="top",
c("HOF II", "HOF IV"),
    col=c("blue", "red"), lty=c(1,2), lwd=1)

plot(MLCOF~CANOPY, xlab="Observed Canopy Density (%)",
     ylab="Presence Probability", main="Melampus coffeus")
lines(spline(tfavi22~CANOPY), col="blue", lty=1, type="l", lwd=3)
lines(spline(tfavi42~CANOPY), col="red", lty=2, type="l", lwd=3)
smartlegend(x="left", y="top",
c("HOF II", "HOF IV"),
    col=c("blue", "red"), lty=c(1,2), lwd=1)

plot(ALGA1~CANOPY, xlab="Observed Canopy Density (%)",
     ylab="Presence Probability", main="Chaetomorpha gracilis")
MODELLING MANGROVE CANOPY DENSITY VARIATIONS

```r
lines(spline(ftavi23~CANOPY), col="blue", lty=1, type="l", lwd=3)
smartlegend(x="left", y="top",
c("HOF II"),
col=c("blue"), lty=c(1), lwd=1)

plot(LCN~CANOPY, xlab="Observed Canopy Density (%)", ylab="Presence Probability",
main="X Lichen")
lines(spline(ftavi24~CANOPY), col="blue", lty=1, type="l", lwd=3)
lines(spline(ftavi44~CANOPY), col="red", lty=2, type="l", lwd=3)
smartlegend(x="left", y="top",
c("HOF II", "HOF IV"),
col=c("blue", "red"), lty=c(1,2), lwd=1)
```
7.4. Worldwide Web Sources

1 UNESCO-Man and Biosphere Programme
   Available at:
   Access period: February, 2007

2 Ramsar: Under-represented wetland types in the Ramsar: List of Wetlands of International Importance, Available at: http://www.ramsar.org/types_mangroves.htm
   Access period: October, 2007

   Available at: http://www.seafdec.org.ph/pew/jesteban.html
   Access period: January, 2007

   Access period: September-February, 2007

5 ATCOR Calibration files for ASTER.
   Available at Geosystems: http://www.geosystems.de/atcor/sensor_updates/index.html
   Access period: September, 2007

6 Rikimaru and Miyatake (1997). Development of Forest Canopy Density Mapping and Monitoring Model using Indices of Vegetation, Bare soil and Shadow
   Available at: http://www.gisdevelopment.net/aars/acrs/1997/ts5/ts5006.asp

7 R Documentation
   Available at: http://www.r-project.org/
   Access period: September-February, 2007
8 Johnson, Lee F. Vineyard Canopy Density Mapping with IKONOS Satellite Imagery
Available at: http://geo.arc.nasa.gov/sge/vintage/ERIM_reprint.pdf
Access period: January, 2007

http://extnasa.usu.edu/link_pages/projects_rsc.html
Access period: December, 2007

10 Smithsonian Institution (2007) Field guide to Indian River Lagoon Florida: Melampus Coffeus:
Available at: http://www.sms.si.edu/IRLFieldGuide/Coffee_melamp.htm
Access period: February, 2007