Available online at www.sciencedirect.com

SCIENCE DIRECT

Remote Sensing of Environment 94 (2005) 491–507

Remote Sensing of Environment

www.elsevier.com/locate/rse

Mapping tropical forest structure in southeastern Madagascar using remote sensing and artificial neural networks

J. Carter Ingram*, Terence P. Dawson[1], Robert J. Whittaker

Biodiversity Research Group, School of Geography and the Environment, University of Oxford, England

Received 12 July 2004; received in revised form 30 November 2004; accepted 5 December 2004

Abstract

Tropical forest condition has important implications for biodiversity, climate change and human needs. Structural features of forests can serve as useful indicators of forest condition and have the potential to be assessed with remotely sensed imagery, which can provide quantitative information on forest ecosystems at high temporal and spatial resolutions. Herein, we investigate the utility of remote sensing for assessing, predicting and mapping two important forest structural features, stem density and basal area, in tropical, littoral forests in southeastern Madagascar. We analysed the relationships of basal area and stem density measurements to the Normalised Difference Vegetation Index (NDVI) and radiance measurements in bands 3, 4, 5 and 7 from the Landsat Enhanced Thematic Mapper Plus (ETM+). Strong relationships were identified among all of the individual bands and field based measurements of basal area (p<0.01) while there were weak and insignificant relationships among spectral response and stem density measurements. NDVI was not significantly correlated with basal area but was strongly and significantly correlated with stem density (r=−0.69, p<0.01) when using a subset of the data, which represented extreme values. We used an artificial neural network (ANN) to predict basal area from radiance values in bands 3, 4, 5 and 7 and to produce a predictive map of basal area for the entire forest landscape. The ANNs produced strong and significant relationships between predicted and actual measures of basal area using a jackknife method (r=0.79, p<0.01) and when using a larger data set (r=0.82, p<0.01). The map of predicted basal area produced by the ANN was assessed in relation to a pre-existing map of forest condition derived from a semi-quantitative field assessment. The predictive map of basal area provided finer detail on stand structural heterogeneity, captured known climatic influences on forest structure and displayed trends of basal area associated with degree of human accessibility. These findings demonstrate the utility of ANNs for integrating satellite data from the Landsat ETM+ spectral bands 3, 4, 5 and 7 with limited field survey data to assess patterns in basal area at the landscape scale.

Keywords: Remote sensing; Forest structure; Artificial neural networks; Madagascar; Basal area

1. Introduction

Tropical forests play crucial roles in the functioning of our planet and the maintenance of life (Myers, 1996). They serve as regulators of global and regional climate systems (Gedney & Valdes, 2000), act as carbon sinks (Grace et al., 1995), are rich in biodiversity, containing over half of the planet's life forms (Wilson, 1988), provide valuable ecosystem services, and serve as vital resources for human populations (Laurance, 1999). Thus, monitoring the state and condition of tropical forests can also provide indications of the health of our planet and its inhabitants. A considerable amount of research has investigated the use of satellite imagery for measuring tropical deforestation, a process which can be readily observed by comparing pixels that have changed from forest to non-forest in images collected on different dates (Green & Sussman, 1990; Skole & Tucker, 1993). However, this crisp, binary approach is

* Corresponding author. The Earth Institute, Columbia University, 1200 Amsterdam Avenue, 10th Floor Schermerhorn Extension, New York, NY 10027-5557, United Kingdom. Tel.: +44 212 854 0051; fax: +44 212 854 8188.

E-mail address: jci2102@columbia.edu (J.C. Ingram).

[1] Now at: The Centre for Environmental Change and Sustainability, School of Geosciences, University of Edinburgh, Scotland.

0034-4257/$ - see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.rse.2004.12.001

unable to capture and describe the great variety of processes that reduce or alter forest cover without eliminating it (Sgrenzaroli et al., 2002; Stone & Lefebvre, 1998). This more subtle process of 'forest degradation' refers to the temporary or permanent decrease in the density, biomass and/or overall structure of vegetation cover and/or its species composition (Grainger, 1993; Sgrenzaroli et al., 2002). Because forest degradation involves internal or vertical change within the forest, it is more difficult to detect with remote sensing than forest clearance. It is important to investigate methods for monitoring these subtle changes since forest modification is generally a more prevalent process than deforestation (Lambin, 1999; Nepstad et al., 1999).

A degraded or modified forest may also be classified as a secondary forest. Tropical secondary forests are excluded from many forest assessments and conservation plans despite the fact that they constitute approximately 40% of tropical forest cover (Brown & Lugo, 1990). Conservation and management of these forests are crucial since they often provide valuable resources to human communities, retain significant amounts of biodiversity and may relieve pressure on primary forests (Cadotte et al., 2002). Sustainable management of tropical secondary forests will require more scientific knowledge and a better understanding of human impact on these forests (Brown & Lugo, 1990; Moran et al., 1996).

Multi-spectral satellite images available at high spatial and temporal resolutions can provide a useful means for monitoring and assessing forest condition, forest modification and 'top down' secondary forest formation (i.e., the conversion of old growth forest to secondary forest through continual degradation processes). A key step towards using remotely sensed images for this purpose is to determine the relationship between spectral information contained within an image and forest structural properties that are indicative of forest condition. The spectral response of a forest is determined by the structure of the canopy through its relationships with leaf area index or canopy cover (Danson, 1995), which controls the amount of understory vegetation, leaf litter and soil that are visible to the sensor (Franklin, 1986). The spectral response is indirectly determined by features that shape the structure of the canopy such as biomass, age, density, mean tree height and basal area (Lee & Nakane, 1996; Peterson et al., 1987; Rock et al., 1986). Many of these features, such as basal area and stem density, are also indicators of forest condition. For example, basal area and density of large stems has been shown to be higher in protected areas and old growth forests and tend to decrease with increasing levels of disturbance (Bhat et al., 2000; Bhuyan et al., 2002; Chittibabu & Parthasarathy, 2000; Macedo & Anderson, 1993). Reduction in basal area or tree biomass due to human disturbance, such as selective logging, may be attributed to a preference for trees of larger size classes for construction purposes (Medley, 1993; Vermeulen, 1996), although these patterns may vary across sites. Similarly, most tropical secondary forests are characterized as having a high density of trees <10 cm diameter at breast height (dbh), short trees with small diameters, low overall basal area and a high leaf area index (Brown & Lugo, 1990). The distinct structural characteristics of disturbed and secondary forests and the correlative relationships with canopy reflectance support the utility of remote sensing as a useful tool for assessing forest condition and forest disturbances, which are often patchily distributed across a landscape both spatially and temporally (Cannon et al., 1994).

Multiple studies have compared the application of different satellite sensors for monitoring forest structural features (Brockhaus & Khorram, 1992; Hyyppa et al., 2000; Lefsky et al., 2001). These studies have shown that the Landsat Thematic Mapper (TM) provides comparable and, in some cases, stronger predictions of certain forest structural features, such as basal area, when compared to radar satellite systems (Hyyppa et al., 2000; Lefsky et al., 2001) or other optical sensors of similar spectral and spatial resolution (Brockhaus & Khorram, 1992). The Landsat data have clear practical advantages over the spectrally comparable SPOT imagery, which include lower costs (Hyyppa et al., 2000). In comparison to hyperspectral or hyperspatial resolution sensors, the Landsat data are less expensive, have lower storage requirements, higher spatial coverage and comparative ease of processing, which is aided by a substantial body of published literature concerning Landsat image processing methods. Due to the strong results derived from Landsat imagery for monitoring forest structural features as shown in previous studies combined with the practical advantages of the sensor, this research utilized Landsat ETM+ imagery for the assessment of tropical forest basal area and stem density in southeastern Madagascar. The practical advantages of a sensor are especially important to consider when determining suitable imagery to use for research or monitoring in developing countries where a high percentage of tropical forests are located and where resources for conservation and environmental management are often limited.

Previous studies have found significant relationships among spectral information within Landsat TM or ETM+ imagery and variables such as forest age, successional status, basal area, height, biomass, density and volume (Brockhaus & Khorram, 1992; Foody et al., 2001, 2003; Jakubauskas, 1996; Olsson, 1994; Puhr & Donoghue, 2000; Steininger, 2000). Research on remote sensing of tropical forest attributes has indicated that the optimal spectral bands for estimating forest structural features may vary across studies and across sites within the same study (Foody et al., 2001, 2003; Steininger, 2000). These issues are at the crux of a central problem in remote sensing applications, which is the inability to generalize across studies in both space and time (Woodcock et al., 2001).

To overcome these obstacles, more knowledge is needed on the relationships among spectral response and forest structural characteristics at different sites. Additionally, there is a need to apply and test similar methodologies so as to increase comparability of results from separate investigations.

Various methods exist for utilizing and analyzing spectral information to assess vegetation or forest structural parameters. One widely used approach is to combine spectral information from multiple bands into a composite value known as a spectral vegetation index (Cohen & Goward, 2004). Of the many vegetation indices that exist, the Normalised Difference Vegetation Index (NDVI) is among the most common in remote sensing studies and provides an estimate of vegetation greenness or biomass per pixel (Goward et al., 1985). However, several factors limit the applications of NDVI in tropical forest studies. One limitation of the NDVI is that vegetation greenness within a pixel saturates at a threshold level, beyond which NDVI values are insensitive to increasing vegetation amount (Ripple, 1985). Furthermore, NDVI provides measures of vegetation greenness and soil reflectance, which may be more sensitive to topographic variation than to actual soil or vegetation properties (Cohen & Goward, 2004). An additional disadvantage of using the NDVI alone is that it utilizes a limited amount of the total spectral information available within an image (Foody et al., 2001). Methods that integrate a broader range of spectral data may provide more information on vegetation cover than possible with the use of a single vegetation index.

The statistical analyses used for understanding the relationships among spectral data and forest attributes should accommodate for the possibility that these relationships may be non-linear and complex. Regression and correlation analyses have commonly been used within remote sensing studies (Jensen et al., 1999; Lawrence & Ripple, 1998). However, these approaches typically assume linear relationships among variables of interest while plant biophysical characteristics often do not conform to these criteria (Jensen et al., 1999). For this reason, nonparametric statistical methods may be more useful for describing the relationship between remotely sensed imagery and environmental variables since these tests make no a priori assumptions about the data. An artificial neural network (ANN) offers a powerful method for analysing complex relationships among variables without making assumptions about the data. ANNs are capable of handling non-normality, nonlinearity and collinearity in a system (Haykin, 1994). This capability is a major advantage of ANNs for assessing the relationships between forest structural attributes and spectral reflectance values, which are frequently non-linear and complex and, in turn, may vary across the different wavebands.

An ANN is defined by an assemblage of "neurons," a protocol for the way the neurons are networked, organized, weighted and connected, and a learning rule (Baret, 1995). An ANN is typically composed of an input layer, one output layer and one or more hidden layers (Jensen et al., 1999). The system 'learns' by predicting output data from patterns learned from a set of input training data (Pearson et al., 2002). By comparing the current output layer to a desired output response, the difference between the two can be obtained and used to adjust weights within the network. The goal is to achieve a set of weights that produce results that closely resemble the target output. This adaptive learning process is repeated until the difference between predicted and training values drop below a predetermined threshold of user-defined accuracy (Jensen et al., 1999). Once constructed and after patterns in the data have been learned, the ANN can be used to estimate or predict values for similar but unexperienced instances of the data (Carvalho, 2001).

ANNs have recently been shown to provide useful alternatives to traditional statistical analyses in forest remote sensing research. Jensen et al. (1999) observed stronger relationships among the age of coniferous forests in Brazil and spectral data from bands 3,4 and 5 of the Landsat TM when using artificial neural networks versus multiple regression. They suggest that this reflects the ability of ANNs to handle collinearity among the spectral bands, which may degrade the predictive power of the multiple regression. ANNs have also been shown to provide stronger relationships between actual and predicted estimates of biomass when using ground data and Landsat TM data of moist tropical forests in Brazil, Malaysia and Thailand when compared to predictions derived from multiple regression or vegetation indices (Foody et al., 2003). In a comparison of the predictive power of neural networks to multiple vegetation indices, including NDVI, Foody et al. (2001) found that the neural networks provided the strongest relationship between predicted and actual estimates of above ground biomass (derived from dbh, basal area and tree height) using the 6 non-thermal bands of the Landsat TM in a tropical rain forest in Borneo.

In this study, we used Landsat ETM+ imagery to assess structural attributes of coastal forests in southeast Madagascar. Madagascar is one of the world's foremost biodiversity hotspots due to its combination of exceptional amounts of endemic species and high estimates of forest loss (Myers et al., 2000). The island's forests are also important for human livelihoods since approximately 80% of the island's population is rural and villagers are often largely dependent upon forests for ecosystem services and resources such as fuel wood and construction materials (Shyamsundar & Kramer, 1997; World Bank, 2003). Despite considerable research on the clearance and extent of Madagascar's rain forests (Green & Sussman, 1990; Mayaux et al., 2000; Nelson & Horning, 1993), there has been surprisingly little research on the application of remote sensing for estimating forest structural features on the island. An initial assessment of the utility of remote sensing data for estimating forest structure in Madagascar is a necessary first step towards developing a program for monitoring forest degradation. Thus, the aims of this study are (i) to assess the utility of using remotely sensed data from the Landsat ETM+ for the assessment of forest structural features in southeast Madagascar, (ii) to assess patterns in forest structure at a landscape scale and (iii) to determine if remote sensing can provide a useful quantitative alternative to commonly used semi-quantitative surveys of forest condition. In this study, basal area and stem density are the primary structural features assessed because they have been shown to be useful indicators of forest disturbance (Bhat et al., 2000; Bhuyan et al., 2002; Chittibabu & Parthasarathy, 2000; Macedo & Anderson, 1993) and because basal area can be directly related to biomass and collected with relative ease (Salvador, 2000).

2. Methods

2.1. Site description

The study site is comprised of approximately 2800 ha of tropical littoral rain forest in southeastern Madagascar and is distributed across three sites: Ste. Luce, Mandena and Petriky (Fig. 1). The forests of St. Luce and Mandena exist as a mosaic of littoral forest fragments while Petriky, the southernmost of the three forests, exists predominately as one large littoral forest parcel. These forests are the last intact vestiges of littoral forest located on sand and have been characterized as floristically and structurally distinct from other similar forests in the region due to their low height and low dbh values (Dumetz, 1999). The three forests have been classified as a subtype of rain forest within a broader class of regional lowland rain forest, which, collectively, are among the world's most biodiverse forest regions for plants due to high tree species richness and a high proportion of endemics (Dumetz, 1999). There is a distinct climatic gradient from north to south, with conditions becoming drier and hotter to the south (Goodman et al., 1997). The wet season occurs from November to May with annual average rainfall amounts averaging 2400 mm in Mandena and Ste. Luce and 1200 mm in Petriky (QMM, 2001). The forests are located at elevations less than 50 m and comprise a relatively narrow band of coastal plain and adjacent foothills averaging approximately 7 km in width and extending from 24°35′ S to 25°08′ S latitude (Lewis Environmental Consultants, 1992).

The forests in the three areas are used by local people for subsistence purposes, which include fuel wood, construction materials, food and medicine. Due to long-term human pressure, the forests are considered as degraded or secondary forests. The future persistence of the forests is under pressure not only from local use but also from the potential establishment of a large mining operation, which will progressively exploit each of the three forest blocks for ilmenite deposits.

These forests have been mapped and preclassified by degradation level as determined from a semi-quantitative ground assessment of canopy closure conducted from 1999 to 2001 as part of an Environmental Impact Assessment of mining in the region (QMM, 2001, Fig. 1). Forest stands were classified as belonging to one of five classes of forest condition: very good condition, good condition, moderate degradation, strong degradation, extreme degradation. Classification was based on percent canopy cover as determined from visual assessments made by at least two observers walking transect(s) throughout forest parcels. If a forest block possessed a high degree of heterogeneity in canopy closure, then more than one class could be assigned to a single stand. This classification system was reviewed by Missouri Botanical Gardens, Royal Botanic Gardens Kew and the Centre National de Recherche Appliquée au Développement Rural (FOFIA) in Madagascar and was deemed valid for the assessment and mapping of canopy degradation in littoral forests (Lowry et al., 1999). However, the reviewers noted that an investigation into the use of canopy degradation as a proxy for overall levels of littoral forest disturbance was needed. The reviewers also suggested that a quantitative study in selected parcels representing each of the forest condition classes be undertaken to provide information on patterns of forest degradation, such as the spatial correlation of tree cutting to villages, roads and paths (Lowry et al., 1999). Forest resource use may intensify differentially across the landscape when the mining proceeds as there would be progressively less forest available to local people. Thus, because such pressure will directly impact biodiversity and resource needs of local communities, it is important to identify and map patterns in forest structure before widespread land use change begins.

2.2. Field surveys

In November 2001, twenty-one belt transects were surveyed in the three study sites. Samples were taken within six forest stands belonging to four of the five different degradation classes (2 strongly degraded, 2 moderately degraded, 1 in good condition, 1 in very good condition). The more degraded stands were smaller in size and, for this reason, two stands were sampled within the strongly degraded and moderately degraded classes. No samples were collected in the extreme degradation class since there was often no intact forest at these sites. Each transect was 100 m by 4m and the site of each transect was randomly selected within each stand and sited at least 100 m from the forest edge when stand size permitted. The cardinal direction of each transect was randomly selected unless the selected direction of the transect would extend the transect into a land cover type other than forest (i.e., swamp or matrix). At each site, geographic coordinates were recorded with a Geographical Positioning System (GPS) allowing geo-referencing to the satellite data. Within each transect, all trees with a dbh (dbh taken at 1.3 m from the forest floor) >5 cm within the 4 m by 100 m transect were identified and the dbh recorded. Where there was more than one stem, all stems >5 cm dbh were recorded, enabling basal area and stem density to be calculated.

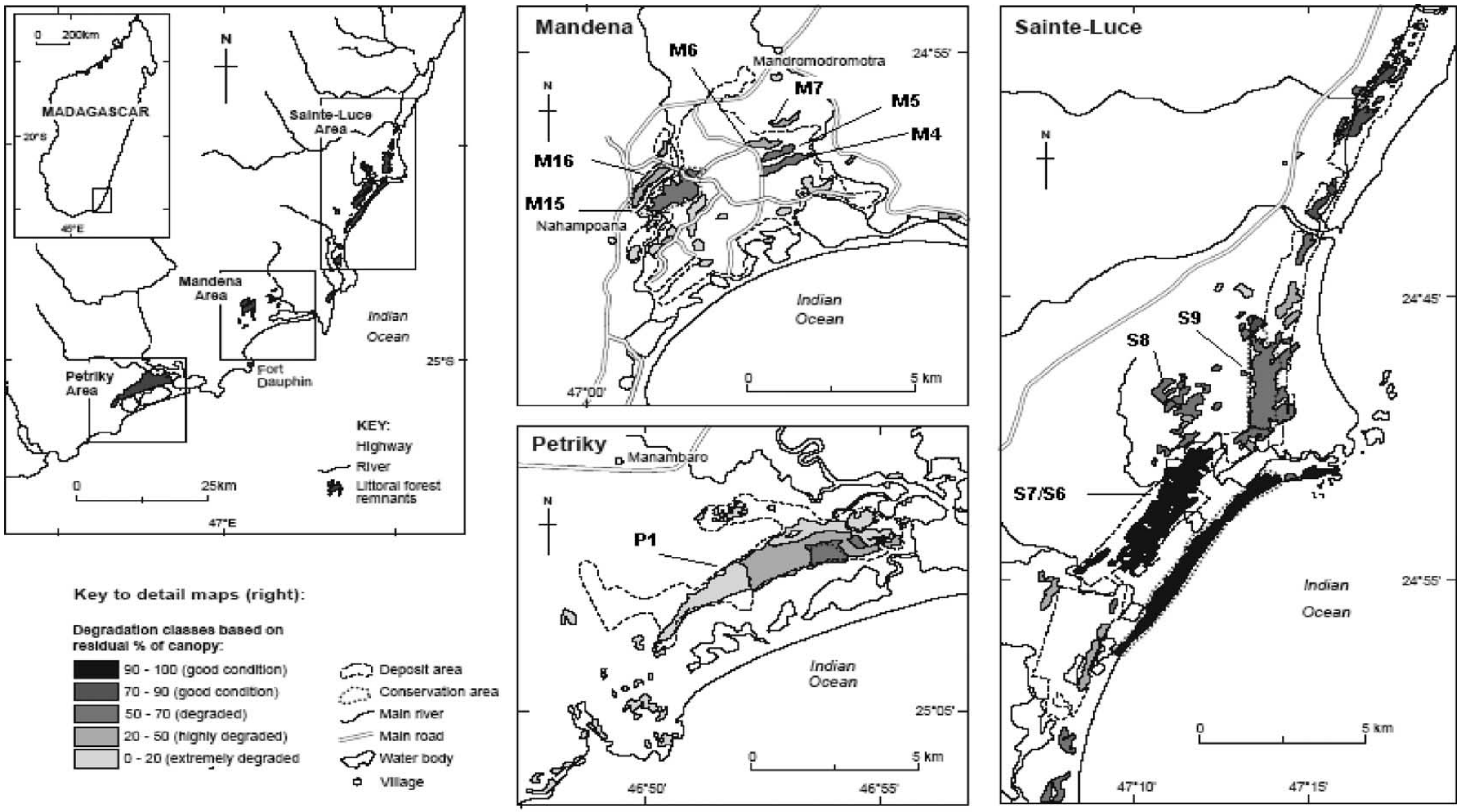

Fig. 1. The map of forest condition created by QMM of the three study sites, St. Luce, Mandena and Petriky, derived from the ground-based semi-quantitative assessment technique. The shades of black and grey represent different classes of forest condition: black represents forest in very good condition, dark grey represents forest in good condition, grey represents forest of moderate degradation, light grey represents forests of strong degradation and very light grey represents areas of extreme degradation. Map adapted by A. Allen (2004) from a map provided by Rio Tinto Iron and Titanium (personal communication, Martin Theberge, 2003).

An earlier field campaign was conducted in the region during July 2000 and measurements of stem density and dbh within 23 transects were collected. This data set was used to test the predictive power of the neural networks when trained with the ground data collected in 2001. However, the 18-month time lag between the collection of the 2000 test data and satellite image acquisition in January 2002 is less than ideal since these forests are used daily by local people and, thus, there could easily have been significant change in the forest structure in the interim. This is a problem characteristic of many remote sensing studies in which it is often difficult to obtain appropriate satellite imagery and ground data from identical time periods (Foody & Curran, 1994). Seven points were removed from the data set because significant forest change had occurred at these sites as indicated by ground observations, visual analysis of aerial photography of the study region taken in December 2000 and a comparison of Landsat TM satellite images of the area from November 1999 and January 2002. Thus, the final set of testing data was comprised of 16 samples. The 2000 field data were collected using a variable area transect (VAT) which is a rapid assessment method for measuring forest structure (Askew, 2001). Transect width was identical to the field surveys conducted in 2001, although the length of the VAT varied as a function of the amount of trees counted within each of three size classes (5–10 cm dbh, 10–15 cm dbh, and >15 cm dbh). The length of the transect terminated when the quota of forty trees of each of three size classes had been filled or when the length of the transect reached 100 m. These transects were located in 6 forest stands (5 in Mandena and 1 in St. Luce) and were selected to represent the five different forest condition classes (2 stands were necessary to obtain an adequate number of samples within the extreme degradation class). The forest stands selected in the 2000 survey were based on classifications of forest condition made in 1998 and were derived using the same semi-quantitative assessment method (QMM, 1998). Transects within each stand were carefully sited so as to reflect the overall classification status of the forest stand.

To determine the compatibility between the measurements of total basal area/m^2 and stem density derived from the different methods, the ground data collected in 2001 were transformed into the VAT data format. This transformation was done by dividing the data collected in 2001 into the three size classes designated in 2000 and recalculating the basal area and stem density. The relationship between the estimates of forest structure using the VAT method of 2000 and the fixed transect method of 2001 was very strong for basal area (r=0.98, n=21, p<0.01) and strong for stem density (r=0.79, n=21, p<0.01). Thus, the potential error associated with the use of these slightly different methods is negligible for basal area but more of a concern for stem density.

2.3. Landsat ETM+ data

A Landsat ETM+ scene (path 158, row 077) covering the study site was obtained for January 2002. Preprocessing requirements of satellite imagery may include atmospheric, radiometric, topographic and geometric correction. Atmospheric and radiometric corrections are necessary when comparing multiple images from different locations or time periods. However, atmospheric corrections were not performed in this study since single date imagery was used, all of the sites of interest are located within the same image (Song et al., 2001), the image was free from cloud cover, and the topography of the study site is relatively uniform (Jensen et al., 1999). For the same reasons, it was not necessary to convert radiance values into reflectance values (Jensen et al., 1999). The image was received in the UTM/WGS 84 coordinate system and reprojected to the Oblique Mercator Hotine projection system using a nearest neighbour resampling method and a polynomial approximation. The root-mean-square error (RMSE) was 0.058 pixels. Pixel size was 28.5 m. All transect geographical coordinates were converted into the Oblique Mercator Hotine projection system and then located on the image. All of the image preprocessing and spectral extraction operations were conducted in Erdas Imagine version 8.4.

The radiance of the red (band 3), near-infrared (NIR)(band 4), mid-infrared (MIR) (band 5) and MIR (band 7) bands and standard deviations (SD) for each band at each transect site were extracted using a 3×3 pixel window. Each window was centered on the GPS recorded position of each transect in order to minimize any potential mislocation errors (similar to Asner et al., 2002; Foody et al., 2001). An NDVI image was created from the red and NIR bands and, in the same manner, the NDVI value and SD of each window was extracted. Bands 3, 4, 5 and 7 of the Landsat ETM+ were selected for this study since these bands have distinct spectral responses to vegetation, soil and litter and have successfully been used within similar studies to predict forest structural features such as age and basal area (Jensen et al., 1999; Steininger, 2000). In band 3 (red), vegetation appears dark and soils and surface litter appear brighter than vegetation (Richardson & Wiegand, 1990) due to chlorophyll absorption of red wavelengths. In band 4 (NIR), vegetation appears bright due to high reflectance and multiscattering of photons as a function of the internal structure of the leaves and the number of layers in the canopy (Lillesand & Kiefer, 1994) while soils and litter appear darker (Asner, 1998). The mid-infrared (MIR) bands are less studied than the red and NIR bands and are commonly known as the ‘water absorption bands’ because water in the leaf strongly absorbs radiation at these wavelengths, meaning reflectance in these bands is inversely related to the total water present in the leaf (Lillesand & Kiefer, 1994). The inclusion of the MIR bands on the Landsat Thematic and Enhanced Thematic sensors, the first Earth Observing satellite system to provide data at these wavelengths, has enhanced the amount and type of forestry information that can be derived from satellite imagery (Cohen & Goward, 2004). Steininger (2000) found that the MIR bands were the most useful Thematic Mapper bands for estimating age and biomass of Amazonian forests. Nemani et al. (1993) found that MIR response decreases with increasing canopy closure. Open canopy sites have a relatively high MIR response compared to closed canopy sites due to the lower water content of understory vegetation and bare soil. In turn, among open canopy stands, sites with understory vegetation had a lower MIR response compared to those with bare soil (Nemani et al., 1993). Horler and Ahern (1986) suggested that the Landsat TM MIR bands are also sensitive to shadowing and vegetation density, making them useful for monitoring clear cuts and regeneration. The two MIR bands on the Landsat TM and ETM+ are band 5 (mid-infrared centered at 1.65 μm) in which vegetation appears moderately dark while soils and litter appear very bright, and band 7 (mid infrared at 2.22 μm) in which vegetation appears extremely dark while soils and litter are extremely bright (Asner et al., 2002).

The standard deviation of the spectral response of each band can provide a textural assessment of canopy structure (Hudak & Wessman, 1998). Canopy roughness is known to increase with the age and/or successional status of a forest (Whitmore, 1990) as a canopy changes from a homogenous canopy of a few light loving pioneers to a more diverse, uneven canopy composed of many mature, slow growing species of various ages (Swaine & Hall, 1983). Mature forests have natural intercrown shadowing and treefall gaps that give rise to background reflectance and textural values (Asner et al., 2002).

2.4. Statistical analyses

2.4.1. Correlation analyses

Preliminary analyses of the structural, spectral reflectance and NDVI data were conducted within the SPSS statistical package using descriptive statistics and correlation analyses. A Kolmogorov–Smirnov test was conducted to determine the normality of the data (Dytham, 1999). All of the bands and NDVI statistically conformed to normal distributions, as did the forest structural parameters of stem density and basal area. However, the degree of normality was variable, there were potential outliers in the dataset(s) and the relationships of interest were likely to be non-linear and complex. Therefore, Spearman's rank correlations were used in this analysis (Catlow, 1993; Zou et al., 2003). Spearman's tests were conducted to determine the correlations of the spectral information in each band with the field data. The data sets collected in 2000 and 2001 were analysed separately in order to assess the influence of the time lag and slightly different ground survey techniques on the relationships among spectral response and forest structural features.

2.4.2. Artificial neural networks

Artificial neural networks (ANNs) were used in this research (i) to determine the relationship of Landsat ETM+ bands 3, 4, 5 and 7 and textural information in those bands to basal area and stem density and (ii) based on this relationship, to use the spectral data within the image to predict structural attributes of pixels comprising littoral forest where no ground data had been collected. NDVI was excluded from the ANN analysis because of the weak correlations observed among NDVI and structural measures. Multi Layer Perceptrons (MLPs) are the most common type of ANN used for remote sensing studies (Carvalho, 2001) and have proved to be effective in comparable studies (Foody et al., 2001). MLP networks were applied in this research and were trained with a back propagation learning principle and one hidden layer. The back propagation learning principle has been demonstrated in other studies to be the best learning principle for modelling non-linear relationships (Jensen et al., 1999). The number of neurons, also known as processing elements, is data dependent and so the following general formula was applied:

$$N = 2i + 1 \tag{1}$$

where N is the number of neurons and i is equal to the number of inputs in the input layer. There were eight inputs consisting of the spectral response values for each of the four spectral bands and the SD values of each band. The desired output was either the basal area or stem density. These structural parameters were tested separately in relation to the spectral information but the same architecture was used for each network: one input layer, one hidden layer and one output layer. For all of the networks created, a sigmoid axon transfer function, which is a non-linear transfer function, was used with a momentum learning rule, a step size of 1.00 and a momentum set to 0.700. It is recommended to standardize input variables to values between 0 and 1 (Jensen et al., 1999; Lloyd, 1996) so the data were standardized using the formula:

$$\mathrm{SV} = (R - \min)/(\max - \min) \tag{2}$$

where SV is the standardized value, R is the real value, *min* is the minimum value in the training data and *max* is the maximum value of the training data.

Two different approaches were used for testing the relationship between spectral response and forest structure within the neural networks, although the network architecture and structure was identical for both approaches.

2.4.2.1. The ANN using ground data from 2001 and 2000: the 2001/2000 ANN. In the 2001/2000 ANN, the 21 transects collected in November of 2001 were used as the training data and the 16 transects collected in July of 2000 were used as the testing data. After creating and training the network, the model was tested and actual measures of basal area and stem density were plotted against the values for each feature as predicted by the network for each transect site. The predictive power of the model was tested using a linear regression and a Bland–Altman test (Bland & Altman, 1986). The linear regression gives information on the extent to which the results of two methods are related while the Bland–Altman test examines whether values derived from two methods agree sufficiently as to be used interchangeably. In this case, we compared the values of basal area derived from two methods: (i) the remote sensing combined with the ANN and (ii) the field surveys. Using a Bland–Altman test, the lack of agreement between two methods can be summarized by calculating the bias, or the mean difference between methods, and the standard deviation of the differences using the equation:

$$\mathrm{LA} = d \pm 2\mathrm{SD} \tag{3}$$

where d is the mean difference, SD is the standard deviation of the differences and LA represents the upper and lower limits of agreement. The degree of difference deemed suitable to permit the substitution of one method for another must be judged by the researcher and will depend on the specific applications of the method.

2.4.2.2. The ANN using the 2001 ground data: 2001 Jackknife ANN. Due to concern over (a) the possibility of structural change between the collection of the validation data in July 2000 and acquisition of the satellite image in January 2002 and (b) potential error associated with differences in the ground survey methodologies, a jackknifing technique was also used to test the predictive power of the network. This analysis utilized only the ground data collected in November 2001. In this approach, all but one of the 21 transects from 2001 were used as training data and the excluded transect was used as the testing data. This process was repeated 21 times so that all values of stem density and basal area were used as testing data once, which produced a predicted value of stem density and basal area for each of the 21 transects. The relationship between predicted and actual values of stem density/ basal area were assessed using a linear regression and the Bland–Altman test.

2.5. Predictive map of basal area

The littoral forests were isolated from other vegetation types by creating a littoral forest mask developed from previous land cover classifications of the region. Using the basal area data set collected from the field survey conducted in 2001 as the training data, the neural network was used to predict the basal area for all unsurveyed littoral forest pixels from the spectral information in bands 3, 4, 5 and 7. The predictions of basal area made by the neural network were spatially displayed across the littoral landscape. The resultant predictive map of basal area was compared to the map of forest condition derived from the semi-quantitative ground survey.

3. Results

3.1. Relationship of forest structure to spectral reflectance values using correlation analyses

The results showed that the 2001 ground data were strongly and significantly correlated with spectral response in all wavebands while the relationships varied in strength and significance when using the ground data collected in 2000 (Table 1). The correlations among basal area measured in 2001 and spectral values were (in decreasing order of strength) MIR band 5 (r=−0.77, p<0.01), MIR band 7 (r=−0.76, p<0.01), NIR band 4 (r=−0.66, p<0.01) and the red band 3 (r=−0.61, p<0.01). The basal area measurements from the data collected in 2000 were most strongly and significantly correlated with the NIR band 4 (r=−0.67, p<0.01), followed by reasonable and significant relationships for the red band (r=−0.55, p<0.05) and MIR band 5 (r=−0.54, p<0.05). The relationship between basal area from the year 2000 data and MIR band 7 (r=−0.45, p>0.05) was not strong or significant. In all cases, the correlation between basal area and the spectral response was negative. NDVI was not strongly or significantly correlated to basal area in either of the data sets.

Stem density was not strongly related to spectral response in any of the individual bands and demonstrated widely divergent relationships with NDVI between the two data sets. Stem density from the 2000 data was strongly and significantly related to NDVI (r=−0.69, p<0.01) but there was no strong or significant relationship between the 2001 stem density data and NDVI (r=−0.09, p>0.05). To test if the difference in the relationship between the two data sets was due to the different ground survey methods, the 2001 data were transformed into the VAT format and then reanalysed with a Spearman's test in relation to NDVI. The results showed no improvement in the strength of the relationship (r=0.03, p>0.05) suggesting that the methodological difference does not account for the variability in the correlations between the two years of ground data. A plot of the negative relationship between NDVI and stem density measurements from the year 2000 revealed a trend in which

Table 1
Spearman's rank correlation coefficients for bands 3, 4, 5 and 7 and NDVI from January 2002 Landsat ETM+ image with ground measurements of stem density and basal area collected in November 2001 (n=21) and July 2000 (n=16)

	2001 Ground data n=21		2000 Ground data n=16	
	Stem density	Basal area	Stem density	Basal area
Red (band 3)	−0.21	−0.61**	0.34	−0.55*
NIR (band 4)	−0.27	−0.66**	0.08	−0.67**
MIR(band 5)	−0.33	−0.77**	0.39	−0.54*
MIR (band 7)	−0.37	−0.76**	0.41	−0.45
NDVI	−0.09	0.14	−0.69**	−0.26

* Correlation is significant at the 0.05 level.

** Correlation is significant at the 0.01 level.

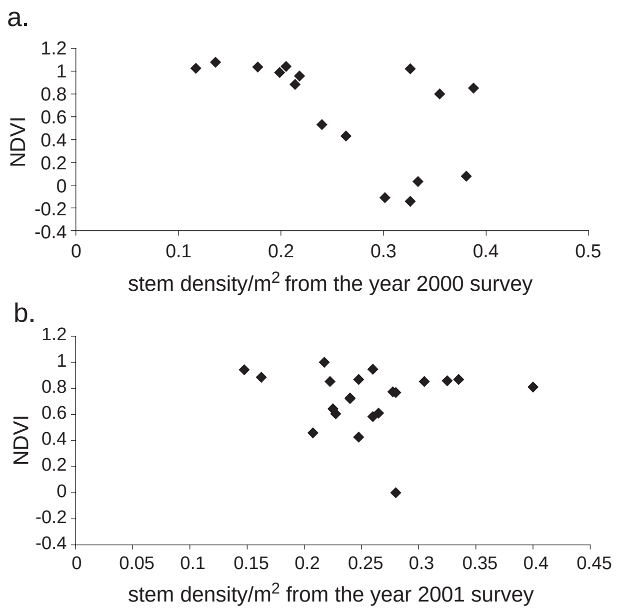

Fig. 2. The relationship between NDVI and stem density. (a) NDVI plotted against stem density measures collected in the year 2000 survey. (b) NDVI plotted against stem density measures collected in the year 2001 survey.

very high values of NDVI were tightly clustered at sites of low stem density and scattered values of NDVI were observed at sites of mid to high stem density (Fig. 2a). A comparison with the stem density data collected during the 2001 survey showed less stratification in stem density and NDVI values (Fig. 2b).

3.2. Results from the artificial neural networks

3.2.1. The 2001 Jackknife ANN

The 2001 Jackknife ANN produced stronger predictions for basal area than for stem density. The relationship between actual and predicted values of stem density was weak and insignificant (r=0.02, n=21, p>0.05). The values of basal area predicted by the 2001 Jackknife ANN were strongly and significantly related to actual measurements of basal area (r=0.79, n=21, p<0.01; Fig. 3a). The regression line for this method was closely aligned with the 1:1 line. The results of the Bland–Altman test (Fig. 3b) revealed that the mean difference between the actual measures and predicted measures of basal area was low, at −0.25. The upper limit of agreement was 9.41 and the lower limit of

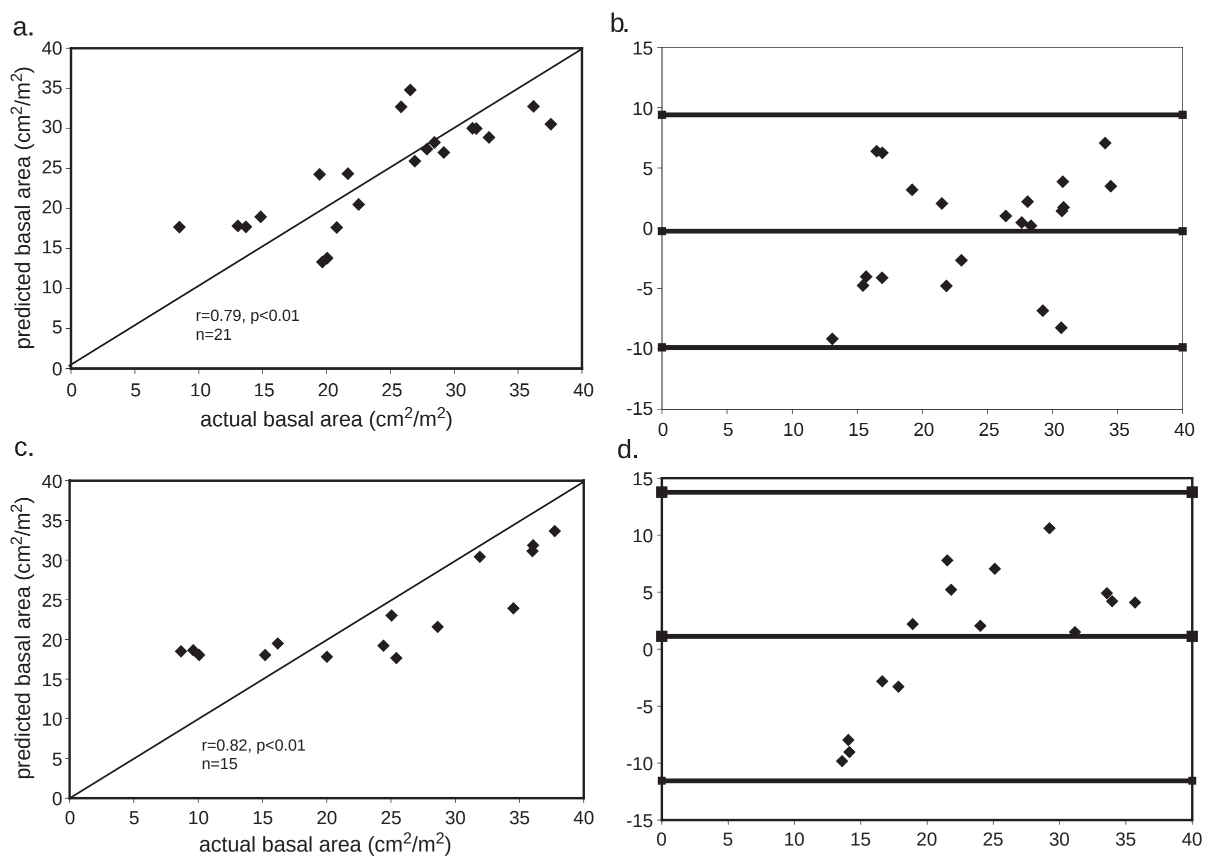

Fig. 3. Relationship of neural network predicted values of basal area to ground measurements of basal area using linear regression and the Bland–Altman test of agreement. (a) The relationship between predicted and actual values of basal area derived from the 2001 Jackknife ANN. (b) The difference between actual and predicted values of basal area from Fig. 3a. using the Bland–Altman test. The horizontal lines represent the mean difference between the values (1.11) and the upper (13.78) and lower (−11.56) limits of agreement. (c) The relationship between actual and predicted values of basal area derived from the 2001/2000 ANN. (d) The difference between actual and predicted values of basal area from Fig. 3c. using the Bland–Altman test. The horizontal lines represent the mean difference between the values (−0.25) and the upper (9.41) and lower limits of agreement −9.90).

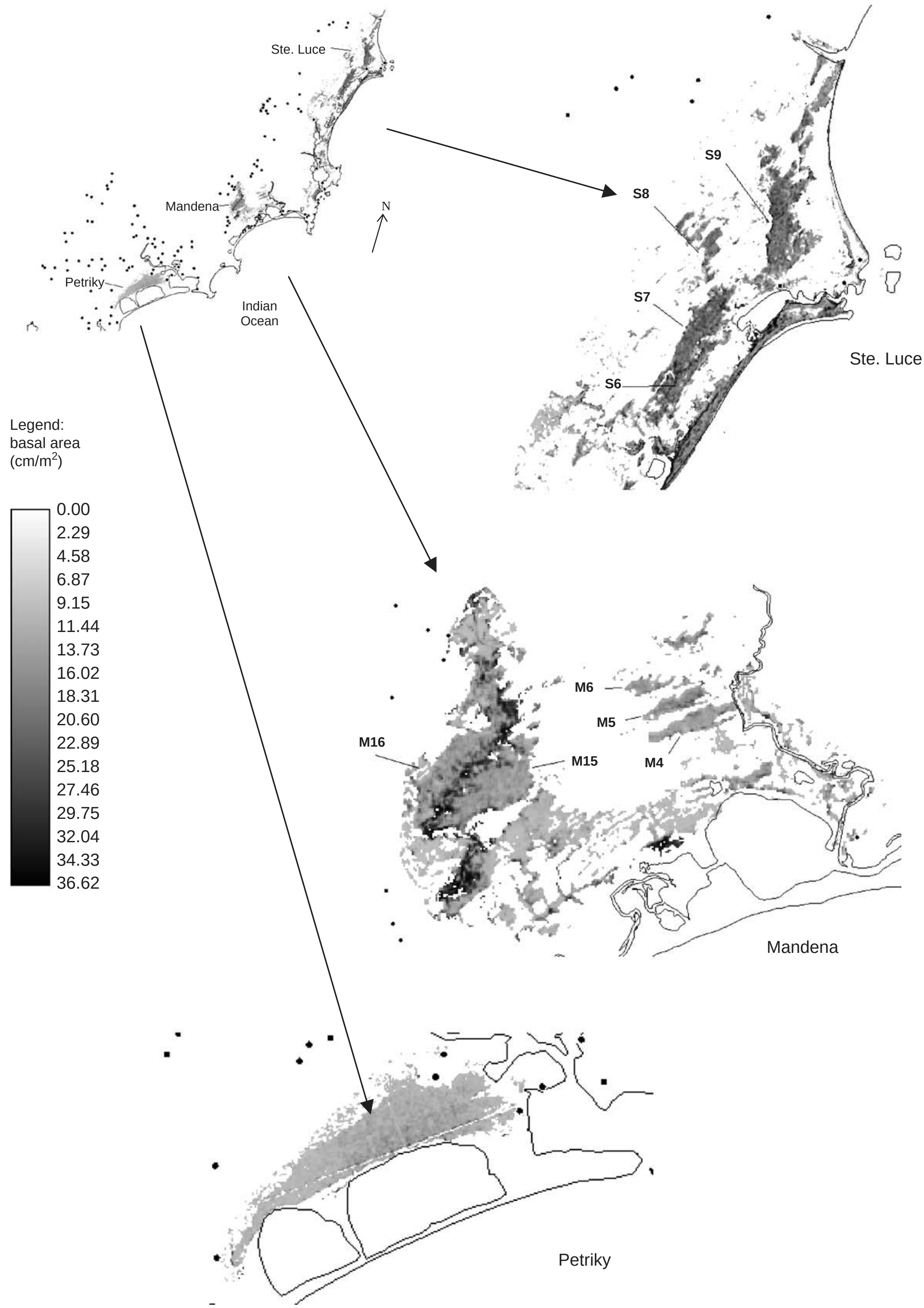

Fig. 4. Map of forest basal area for the study of sites as predicted by the ANN using spectral information from Landsat ETM+ bands 3, 4, 5 and 7 and ground data from the year 2001. Points represent villages and lines represent roads, rivers, lakes and coastal boundaries. Legend presents basal area in cm^2/m^2 values.

agreement was −9.90. All of the measurements occur within 2 SD of the mean difference.

3.2.2. The 2001/2000 ANN

The 2001/2000 ANN proved to be more powerful for predicting basal area than for predicting stem density. The relationship between predictions of stem density and actual values of stem density was weak and insignificant (r=0.04, n=16, p>0.05). The 2001/2000 ANN produced a strong and significant relationship between actual and predicted values of basal area (r=0.69, n=16, p<0.01). There was one outlier for which the predicted values of basal area produced by the 2001/2000 ANN were considerably lower than the actual measure of basal area. A plot of the residuals identified this as an outlier, having a residual value greater than two standard deviations from the mean residual value. It is likely that this low prediction could be due to tree removal, which has occurred at this transect since the ground data were collected in July of 2000 and which was not detectable in the visual comparison of aerial photography from December 2000 and satellite imagery from 2002 and 1999. This sampling station was located in a forest fragment that is close to the main village in St. Luce (S9) and is frequently accessed by villagers for resources. Once this outlier was removed, the relationship improved (r=0.82, n=15, p<0.01; Fig. 3c). A comparison of the regression line with the 1:1 line demonstrates that, although the relationship is strong between predicted and actual values of basal area, the relationship substantially deviates from an exact relationship. This is shown in Fig. 3d through the results of the Bland–Altman test: the mean difference between actual and predicted measures of basal area/m^2 was 1.11. The upper limit of agreement was 13.78 and the lower limit of agreement was −11.56. The 15 data points fell within these limits, which represent a 2 SD envelope around the mean difference.

3.2.3. Textural information

The contribution of textural values to basal area predictions was tested by running each network, the 2001/2000 ANN and the 2001 Jackknife ANN, with the radiance values within each band but without the textural information. The relationship between predicted and actual basal area did not change significantly when the textural data were removed from the input data within either the 2001/2000 ANN (r=0.82, n=15, p<0.01) or the 2001 Jackknife ANN (r=0.78, n=21, p<0.01).

3.2.4. The predictive map of basal area

The predictive map of basal area (Fig. 4) portrays climatic influences on forest structure and trends in basal area that appear to be associated with human access and distance from forest edge. The model predicts Petriky to have an overall low basal area when compared to St. Luce. Within each forest site, where climate and other physical conditions are generally uniform, distinctive trends of basal area that appear to be linked to human access and distance from forest edge can be observed.

In Petriky, pixels predicted to have the highest basal area occur within the middle of the forest fragment. Although these high predictions of basal area occur within pixels proximate to the road, they were at a maximal distance from the three villages located on the periphery of the forest.

A high degree of heterogeneity in the basal area predictions across the Mandena site. Two fragments, M15 and M16, have been designated by the mining company as conservation zones and were predicted to have a relatively high basal area when compared to the other forest stands in the site. Forest fragments that lie proximate to roads, such as M13 and M3, are predicted to have very low basal area. These predictions of low basal area have been confirmed by ground observations within these forest parcels (JCI, the author's observations). In comparison, fragments M4, M5, M6 and M7, which are set slightly back from the road and are bordered by a river on the northern boundaries of the fragments, were predicted to have more inter-stand heterogeneity in basal area values and a higher abundance of high basal area pixels when compared to unprotected fragments within Mandena. The pixels predicted to have the lowest basal area within M4, M5, M6 and M7 fragments were located predominantly along the edge of these fragments. Multiple pixels across the Mandena site were predicted to have very high basal area values (>20 cm/m^2), notably in the area between stands M15 and M16. However, ground observations have revealed that most of these areas are wetland-like forest habitats (JCI and TD, authors' observations). These swampy areas are characterized by a high abundance of palm-like trees, with large leaves, which may have a spectral signal similar to dense vegetation.

In St. Luce, there is one main road that passes by the three surveyed forest parcels (S9, S8 and S7). Three villages are located towards the end of this road, one of which is extremely close to the forest stand of S9. The forest pixels located closest to this village were predicted to have a relatively lower basal area than pixels at greater distances from the village and the road, most likely due to the ease of access at these locations. The core areas of the forest parcels were generally comprised of pixels with higher predicted values of basal area. Within the three larger fragments, S9 and S6/S7 (which is bisected by a stream into S6 and S7), there was a higher overall basal area when compared to the smaller fragments such as S8 and other forest stands included in the region.

4. Discussion

The relationships among spectral values and basal area derived from the correlation analyses and those between predicted and actual measures of basal area derived from the ANNs demonstrate that the selected wavebands from the

Landsat ETM+ are useful for estimating forest basal area for the study site. The correlation analyses provided descriptive information on the strength of each band in relation to stem density and basal area while the ANNs made use of these relationships to predict basal area across the landscape with reasonable accuracy. The spatial depiction of basal area across the landscape displayed climatic impacts on basal area as well as trends associated with distance from human population centres and accessibility to forest fragments. The findings in this study support observations from similar studies and contribute to the growing body of knowledge on the application of remote sensing for assessing tropical forest condition.

4.1. Correlations of NDVI to stem density and basal area

There were inconsistent relationships between stem density measures and NDVI. The stronger relationship between the stem density measurements collected in 2000 and NDVI in comparison to the very weak relationship between stem density measures collected in 2001 and NDVI could be due to the careful pre-stratification of sites within each forest fragment during the 2000 field survey. During the 2000 field survey, each transect site was chosen to represent the deterioration status of the forest fragment. In contrast, during the 2001 survey, transect sites were randomly selected within each fragment in order to assess the range of internal fragment structural diversity. Because of the stratification during the 2000 survey, more extreme differences in stem density among fragments were captured when compared to the 2001 data on stem density (as seen in Fig. 2). This stratification of sites and fragments resulted in an extreme separation between the high and low-to-moderate stem densities and, thus, NDVI values. NDVI may be sensitive to detecting extremes in forest condition (i.e., old growth forest versus highly degraded forest), but is less able to predict structural values of forests characterized by moderate condition. The negative relationship observed here is similar to the negative, but weak, relationship observed by Foody and Curran (1994) using Landsat TM derived NDVI to tree density in a West African tropical forest.

There was no significant relationship between NDVI and basal area using data from either of the two field campaigns. Previous studies in tropical forests have also reported weak and insignificant relationships between Landsat TM-derived NDVI values and biomass (Foody et al., 2003). The use of the NDVI is problematic for estimating structural features of mixed, broad leaf tropical forests because it loses sensitivity at high vegetation amounts (Foody et al., 1996; Sader, 1989). For this reason, one of the most valuable applications of the NDVI for tropical forest assessments continues to be for estimation of vegetation loss through clearance or detecting extreme differences in forest condition rather than estimating values of forest structural features or detecting subtle landscape variations in these features.

4.2. Correlations of spectral response with stem density and basal area

The lack of strong and significant relationships of individual bands with stem density is similar to findings in other studies, which have failed to identify a strong relationship between stem density and spectral response (e.g., Puhr & Donoghue, 2000). From these results, stem density does not seem to be a structural feature that can be accurately assessed using our methods.

All correlations of the spectral values in the individual wavebands with basal area were negative. Negative correlations of TM bands 3, 4, 5 and 7 with forest structural parameters have also been observed in other studies investigating canopy spectral reflectance (Jensen et al., 1999; Steininger, 2000). The negative relationship of basal area to NIR reflectance is unexpected since green vegetation reflects strongly in the NIR, which means that the relationship between NIR reflectance and increasing LAI of vegetation should be positive. However, the effects of increasing shadow of the canopy, which may accompany an increase in LAI, will counteract the expected positive relationships of NIR with measured forest structural attributes (Danson, 1995). Within the correlation analyses using the 2001 ground data, reflectance values in the MIR band 5 and MIR band 7 demonstrated the strongest relationships with basal area and the weakest correlations with the red band 3. These findings are consistent with other studies, which have identified strong relationships among the MIR region and forest structural features (Foody et al., 2001; Horler & Ahern, 1986; Steininger, 2000). Foody et al. (2001) found that the data acquired in MIR band 5 was the most useful and red band 3 was the least useful for predicting biomass of tropical rain forests in Borneo using ANNs. Similar to our results, basal area of Bolivian tropical secondary forests has been correlated with the following bands in decreasing order of strength: MIR band 5, MIR band 7, NIR band 4 and red band 3 (Steininger, 2000). Our study contributes to a growing body of research that has demonstrated the usefulness of these bands for discriminating structural features such as basal area of tropical secondary forests.

The strength of the relationship between basal area collected in 2000 and spectral response varied from the correlations observed from the analysis of the 2001 ground data. The decrease in the strength of the relationship between basal area and spectral response in MIR band 5, MIR band 7 and the red band 3 could be due to physical alterations to the forest vegetation at sites surveyed in 2000 and the sensitivity of these bands to forest change. The MIR band 5 has been shown to have high sensitivity to canopy changes during tropical forest succession (Asner et al., 2002; Steininger, 2000). MIR band 7 has also been strongly correlated with vegetation condition, forest stand basal area and height in previous studies in temperate regions (Ahern et al., 1991; Jakubauskas, 1996; Puhr & Donoghue, 2000) and should show a strong contrast between forest canopy and gaps (Asner et al., 2002). The red band is highly sensitive to variability across the forest canopy (Foody & Curran, 1994) and to changes in forest cover associated with selective logging and regrowth (Asner et al., 2002). Asner et al. (2002) found that the relationships among Landsat ETM+ bands 3 and 5 and canopy gap fraction of Amazonian forests were significantly weakened after 1.5 years, similar to the time lag between the July 2000 ground survey and the January 2002 satellite image in this study. Thus, these bands would be expected to be sensitive to any subtle ground-based changes. The high degree of human activity in these forests means that it is very plausible that physical changes to the forest canopy, basal area and/or height due to the removal or regrowth of trees have occurred since the ground data were collected in 2000. This would act to weaken the relationship between ground measures of basal area and spectral response within these bands.

The NIR band was the only band in which the relationship between spectral response and basal area did not change significantly when using the ground data collected in 2000. A plot of the NIR values versus basal area measurements from the 2000 survey revealed a tight cluster of relatively high NIR values at sites with low values of basal area (Fig. 5). There was a wider variation and scatter among the lower NIR values at sites where moderate to high basal area was recorded in 2000. The strength of the relationship between NIR and basal area measurements collected in the year 2000 is most likely because NIR response increases in open stands due to the high reflectance of understory vegetation and other highly reflective backgrounds (Nemani et al., 1993). Thus, if there had been growth or succession at the sites recorded as having low basal area in the year 2000, NIR reflectance would have increased at these locations due to an increase in understory vegetation. Such a scenario is likely since sites with very low basal area would not be preferred for further cutting and, thus, temporarily, would be abandoned and left to regenerate. The tight clustering of these high values at sites with low basal area in comparison to the scatter of the relatively lower NIR values contributes to this observed, but unexpected, strong relationship.

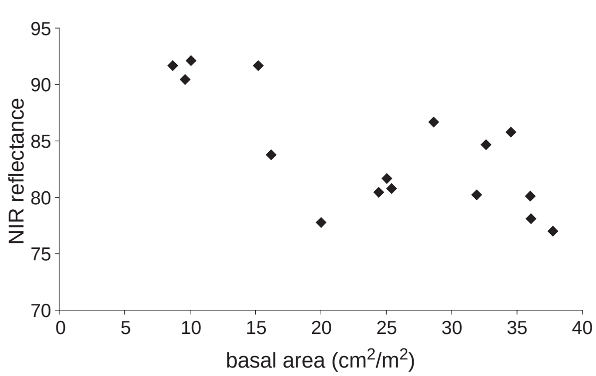

Fig. 5. Near-infrared radiance values plotted against basal area measurements collected in the year 2000 ground survey.

4.3. Neural networks for predicting stem density and basal area

Although the correlation analyses showed that the correlations of the MIR bands 5 and 7 with basal area were very strong and, thus, potentially useful individually for predictive purposes, the ANN approach was robust across the two data sets and allowed inclusion of multiple wavebands also known to provide useful information on forest structure and, thus, providing more predictive power than one or two bands used alone. Using a neural network, the radiance values within the four spectral bands produced predictions of forest basal area, which were strongly and significantly correlated with the actual measures of basal area. The strong and significant relationships between predicted and actual measures of basal area using both the 2001/2000 ANN and 2001 Jackknife ANN support the utility of neural networks for predicting patterns of basal area from the spectral bands used in this study. However, the two networks did differ in the prediction tendencies. The predictions of basal area using the 2001 Jackknife ANN were fairly close to a 1:1 relationship and were consistent in variability with respect to the field measurements. The 2001/2000 ANN produced a strong and significant relationship between actual and predicted measures of basal area but, in contrast to the 2001 Jackknife ANN, the relationship deviated substantially from the 1:1 line as indicated from the results of the Bland–Altman test. The model derived from the 2001/2000 ANN had a tendency to underpredict values of basal area at sites with medium to high actual values of basal area and to overpredict values of basal area at sites with relatively low actual values of basal area. This trend can be attributed to changes in basal area since the data were recorded in July 2000, which would alter the relationship between actual and predicted estimates of basal area derived from the 2002 satellite image. The nature of the model's systematic error, due largely to observed physical changes occurring in this forest, supports the utility of the method. However, the results from the two ANN models demonstrate the importance of obtaining a satellite image as close as possible to the date of the ground survey for more accurate estimations of basal area.

Remote sensing alone may not be able to produce exact values of basal area, but when used in combination with ground data, the bias and limits of agreement can be calculated to give an indication of the deviation from a 1:1 agreement and, thus, the error associated with the predictions. The strong relationship between actual and predicted values of basal area for both ANNs demonstrate the usefulness of remote sensing for assessing patterns in basal area across a landscape and its potential as a tool for first-order assessments of forest condition when ground survey data are not available.

4.4. Contribution of textural information for predicting basal area

Textural information appears to be insignificant in this study since the relationships between predicted and actual measures of basal area did not change significantly when textural data were removed from the network. Textural measures derived from the standard deviation of pixels within a 3×3 window have been useful for discriminating heterogeneity in vegetation structure using imagery of higher spatial resolutions than Landsat TM or ETM+ images (Hudak & Wessman, 1998). The 30-m pixel size of the Landsat ETM+ image may be too coarse for detecting subtle differences in basal area or stem density across a forest fragment. However, textural measures may be useful at this spatial resolution for assessing more extreme differences in land cover types (such as forest versus non-forest) or for assessment of structural differences across less disturbed, old growth forests where canopy roughness is at a maximum. The lower, more homogenous canopy of the forests in these study sites, such as the Petriky forest, make textural analyses less useful for discriminating variations in basal area within these forest stands.

4.5. Predictive map of basal area

The map derived from the ANN predictions of basal area displayed patterns of forest condition comparable to those presented within the semi-quantitative map, captured inter-site differences in basal area associated with environmental variability and portrayed intra-site trends in basal area associated with human impact. The ANN predictive map provided finer spatial detail on the heterogeneity of basal area within a forest fragment and displayed trends in basal area that were lacking within the semi-quantitative preexistent map of forest condition (Fig. 1).

In the absence of humans, spatial patterns of forest basal area will be determined by environmental variables (Salvador, 2000). The environmental factors that influence tropical rain forest structure and biomass include soil type, soil nutrients, climate, disturbance regime and topography (see Clark & Clark, 2000, for a review). Soil nutrients and climate are known to vary across the three study sites (Lewis Environmental Consultants, 1992) and influence forest basal area. Rainfall decreases progressively from north to south, which has discernible impacts on Petriky, the southernmost of the forest sites. Petriky is a drier, more nutrient-poor zone, similar to a scrub woodland, and for this reason has been classified as a distinct subtype from St. Luce and Mandena (Dumetz, 1999). The ANN predictive map captured these differences, most notably between the extremes of St. Luce and Petriky, by predicting low basal area values throughout Petriky and higher basal area values throughout St. Luce. However, structural differences were not as evident between Mandena and Petriky. Mandena represents a transition area amid the two climatic extremes and it has been extremely impacted by human pressure, such as charcoal making (QMM, 2001). Thus, basal area and forest structure at this site may have been shaped by both anthropogenic factors and environmental factors such as climate. Additionally, it should be noted that there are fewer villages within the St. Luce region, which may also contribute to the higher overall basal area predictions and ground measurements, when compared to Mandena. The qualitative classification of forest condition employed by QMM (Fig. 1) suggests forest structure and condition at these sites is primarily determined by human impact. In contrast, the predictive map of basal area produced by the ANN provides numerical values related to forest condition and, thus, avoids the use of value laden categorisations (i.e., very good condition, good condition, moderate deterioration, high deterioration, extreme deterioration), which deemphasize the role of climate and other environmental factors in shaping forest structure.

Within each of the three forest sites where environmental variables are relatively constant, predicted patterns of basal area appear to be associated with centres of human population combined with ease of accessibility to the forests. Changes in forest structure and decreases in basal area have been observed with distance from human settlements or roads and represent the impact of human pressure on forests (Chittibabu & Parthasarathy, 2000; Medley, 1993; Vermeulen, 1996). In this region, specifically, degree of stand isolation and human accessibility has been shown to impact community composition and biodiversity of trees and shrubs (Cadotte et al., 2002). The predictive map of basal area portrayed such trends by displaying gradients of increasing basal area with increasing distance away from forest edges, villages and roads. This trend is especially noticeable in the forest of Petriky, where the basal area is highest in the middle of the forest parcel, at a maximal distance from the three villages, but, interestingly, proximate to the road, which bisects the forest parcel. Deforestation along roads has been observed in other areas both globally and in Madagascar due to increased ease of access (Kistler & Spack, 2003), especially when using mechanized transport for exporting trees from the forest. However, at this site, all tree removal occurs by the manual labour of local people, few of whom, if any, have access to any form of ground transport other than walking. This limits the distance people can travel while carrying heavy timber. Additionally, the ‘road’ is a very coarse grained sand path, which makes walking long distances difficult, especially during the heat of the day. Such findings underline the fact that local context is crucial when considering trends in forest cover and land use (Kull, 2000). The high basal area predicted for the conservation zones in Mandena, M15 and M16, are attributed to protection status of the fragments since the stands were predicted to have the highest overall basal area values for the site despite their very close proximity to villages and roads. Within St. Luce and Mandena, unprotected stands that were bisected by roads and located close to villages were predicted by the model to have lower basal area values when compared to fragments that were slightly more removed from the road, villages and/or protected by river or water boundaries. The ANN was also able to capture edge effects by predicting low basal area pixels along stand edges. This was especially evident for the larger parcels comprising S8 in St. Luce. The ability of the ANN to portray trends of basal area in relation to environmental and anthropogenic factors, known to influence forest systems from other studies, demonstrates the utility of this method for broad assessments of forest structural patterns at a landscape level.

4.6. The application of this method for monitoring forest condition

The relationship between spectral information and ground-collected forest structural data can be combined with other indicators of forest change detectable throughout a time series of satellite images to monitor forest modification (Lambin, 1999). The results found in this study provide a foundation and crucial first step for establishing such a monitoring program of forest degradation in southeastern Madagascar. Such research could provide a useful supplement to traditional ground-based forest inventories, which can be expensive and time consuming (Hyyppa et al., 2000) and are limited in spatial extent. The ability to derive temporal and spatial generalizations of the relationships among spectral data and forest structural features awaits more research (Foody et al., 2003) and should be pursued to produce a method useful for long-term monitoring.

5. Conclusion

This study has demonstrated the potential for using a limited amount of ground-collected data and Landsat ETM+ spectral information from bands 3, 4, 5 and 7 for estimating the basal area of tropical forests in southeastern Madagascar. The strong and significant relationships between predicted and actual measurements of basal area derived from ANNs and the compatibility of the ANN-derived map of basal area (Fig. 4) with a preexistent map of forest condition (Fig. 1) support the use of the methods presented here for assessments of basal area across a forest landscape. Future research should build upon these relationships and investigate how the predictive relationships between spectral information and basal area observed here can be generalized temporally and, thus, incorporated into long term monitoring of forest condition for the region.

Acknowledgements

This study was supported by the DEFRA (Department for Environment Food and Rural Affairs) Darwin Initiative (project number 162/9/006). The authors would like to thank QIT–Madagascar Minerals for geographical data and in country support and assistance. We thank Rose Askew for providing us with supplementary field data from the year 2000. The most recent field survey would not have been possible without the hard work and assistance of Christian Stamm, Ramesy Edmonds, Lalaina Andriamiharisoa, Andry Rabemanantsoa and Emmanuel Randriambinintsoa. Thank you to Martin Theberge of Rio Tinto for providing the original map of forest condition and cartographer Ailsa Allen for re-drawing this map in black and white for publication. The authors would also like to thank Richard Pearson who provided valuable assistance and expertise on the application of the artificial neural networks and James Watson for comments an earlier draft of this manuscript.

References

Ahern, F. J., Erdle, T., Maclean, D. A., & Kneppeck, I. D. (1991). A quantitative relationship between forest growth rates and Thematic Mapper reflectance measurements. *International Journal of Remote Sensing*, *12*, 387–400.

Askew, R. (2001). An investigation into plant composition and structure of highly fragmented degraded forests of southeastern Madagascar. Unpublished BS Final Honours Project thesis. Department of Zoology, University of Oxford.

Asner, G. P. (1998). Biophysical and biochemical sources of variability in canopy reflectance. *Remote Sensing of Environment*, *64*, 234–253.

Asner, G. P., Keller, M., Pereira, R., & Zweede, J. C. (2002). Remote sensing of selective logging in Amazonia: Assessing limitations based on detailed field observations, Landsat ETM+, and textural analysis. *Remote Sensing of Environment*, *80*, 483–496.

Baret, F. (1995). Use of spectral reflectance variation to retrieve canopy biophysical characteristics. In F. M. Danson, & S. E. Plummer (Eds.), *Advances in environmental remote sensing* (pp. 33–51). Chichester, UK: John Wiley & Sons.

Bhat, D. M., Naik, M. B., Patagar, S. G., Hegde, G. T., Kanade, Y. G., Hegde, G. N., et al. (2000). Forest dynamics in tropical rain forests of Uttara Kannada district in Western Ghats, India. *Current Science*, *79*, 975–985.

Bhuyan, P., Khan, M. L., & Tripathi, R. S. (2002). Regeneration status and population structure of Rudraksh (*Elaeocarpus ganitrus* Roxb.) in relation to cultural disturbances in tropical wet evergreen forest of Arunachal Pradesh. *Current Science*, *83*, 1391–1394.

Bland, J., & Altman, D. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*, *i*, 307–310.

Brockhaus, J. A., & Khorram, S. (1992). A comparison of spot and Landsat-TM data for use in conducting inventories of forest resources. *International Journal of Remote Sensing*, *13*, 3035–3043.

Brown, S., & Lugo, A. E. (1990). Tropical secondary forests. *Journal of Tropical Ecology*, *6*, 1–32.

Cadotte, M. W., Lovett-Doust, J., Franck, R., & Reza, L. (2002). Tree and shrub abundance and diversity in fragmented coastal tropical forest of southeastern Madagascar. *Biodiversity and Conservation*, *11*, 1417–1436.

Cannon, C., Peart, D., Leighton, M., & Kartawinata, K. (1994). The structure of lowland rainforest after selective logging in West Kalimantan, Indonesia. *Forest Ecology and Management*, *67*, 49–68.

Carvalho, L. M. T. (2001). Mapping and monitoring forest remnants: A multiscale analysis of spatio-temporal data. Doctoral thesis. Wageningen University, Wageningen, The Netherlands.

Catlow, B. (1983). *Quantitative methods for computing students*. Guernsey, UK: Guernsey Press Co.

Chittibabu, C. V., & Parthasarathy, N. (2000). Attenuated tree species diversity in human-impacted tropical evergreen forest sites at Kolli hills, Eastern Ghats, India. *Biodiversity and Conservation*, *9*, 1493–1519.

Clark, D. B., & Clark, D. A. (2000). Landscape scale variation in forest structure and biomass in a tropical rain forest. *Forest Ecology and Management*, *137*, 185–198.

Cohen, W. C., & Goward, S. N. (2004). Landsat's role in ecological applications of remote sensing. *BioScience*, *54*, 535–546.

Danson, F. M. (1995). Developments in the remote sensing of forest canopy structure. In F. M. Danson, & S. E. Plummer (Eds.), *Advances in environmental remote sensing* (pp. 53–69). Chichester, UK: John Wiley & Sons.

Dumetz, N. (1999). High plant density of lowland rainforest vestiges in eastern Madagascar. *Biodiversity and Conservation*, *8*, 273–315.

Dytham, C. (1999). *Choosing and using statistics: A biologist's guide*. Oxford, UK: Blackwell Science.

Foody, G. M., Boyd, D. S., & Curran, P. J. (1996). Relations between tropical forest biophysical properties and data acquired in AVHRR channels 1–5. *International Journal of Remote Sensing*, *17*, 1341–1355.

Foody, G. M., Boyd, D. S., & Cutler, M. E. J. (2003). Predictive relations of tropical forest biomass from Landsat TM data and their transferability between regions. *Remote Sensing of Environment*, *85*, 463–474.

Foody, G. M., & Curran, P. J. (1994). Estimation of tropical forest extent and regenerative stage using remotely sensed data. *Journal of Biogeography*, *21*, 223–244.

Foody, G. M., Cutler, M. E., McMorrow, J., Pelz, D., Tangki, H., et al. (2001). Mapping the biomass of Bornean tropical rain forest from remotely sensed data. *Global Ecology and Biogeography*, *10*, 379–387.

Franklin, J. (1986). Thematic Mapper analysis of coniferous forest structure and composition. *International Journal of Remote Sensing*, *7*, 1287–1301.

Gedney, N., & Valdes, P. J. (2000). The effect of deforestation on the northern hemisphere circulation and climate. *Geophysical Research Letters*, *27*, 3053–3056.

Goodman, S. M., Pidgeon, M., Hawkins, A. F. A., & Schulenberg, T. S. (1997). *The Birds of southeastern Madagascar*. Chicago, IL: Field Museum of Natural History.

Goward, S. N., Tucker, C. J., & Dye, D. G. (1985). North American vegetation patterns observed with the NOAA-7 advanced very high-resolution radiometer. *Vegetatio*, *64*, 3–14.

Grace, J., Lloyd, J., McIntyre, J., Miranda, A. C., Meir, P., Miranda, H. S., et al. (1995). Carbon-dioxide uptake by an undisturbed tropical rain-forest in southwest Amazonia, 1992 to 1993. *Science*, *270*, 778–780.

Grainger, A. (1993). *Controlling tropical deforestation*. London, UK: Earthscan Publications.

Green, G. M., & Sussman, R. W. (1990). Deforestation history of the eastern rain forests of Madagascar from satellite images. *Science*, *248*, 212–215.

Haykin, S. (1994). *Neural networks: A comprehensive foundation*. Upper Saddle River, NJ: Prentice Hall.

Horler, D. N. H., & Ahern, F. J. (1986). Forestry information content of Thematic Mapper data. *International Journal of Remote Sensing*, *7*, 405–428.

Hudak, A., & Wessman, C. (1998). Textural analysis of historical aerial photography to characterize woody plant encroachment in South Africa savanna. *Remote Sensing of Environment*, *66*, 317–330.

Hyyppa, J., Hyyppa, H., Inkinen, M., Engdahl, M., Linko, S., & Zhu, Y. H. (2000). Accuracy comparison of various remote sensing data sources in the retrieval of forest stand attributes. *Forest Ecology and Management*, *128*, 109–120.

Jakubauskas, M. E. (1996). Thematic Mapper characterization of lodgepole pine seral stages in Yellowstone National Park, USA. *Remote Sensing of Environment*, *56*, 118–132.

Jensen, J. R., Qiu, F., & Ji, M. H. (1999). Predictive modelling of coniferous forest age using statistical and artificial neural network approaches applied to remote sensor data. *International Journal of Remote Sensing*, *20*, 2805–2822.

Kistler, P., & Spack, S. (2003). Comparing agricultural systems in two areas of Madagascar. In S. M. Goodman, & J. P. Benstead (Eds.), *The natural history of Madagascar* (pp. 123–134). Chicago, IL: University of Chicago Press.

Kull, C. (2000). Deforestation, erosion, and fire: degradation myths in the environmental history of Madagascar. *Environment and History*, *6*, 423–450.

Lambin, E. F. (1999). Monitoring forest degradation in tropical regions by remote sensing: some methodological issues. *Global Ecology and Biogeography*, *8*, 191–198.

Laurance, W. F. (1999). Reflections on the tropical deforestation crisis. *Biological Conservation*, *91*, 109–117.

Lawrence, R. L., & Ripple, W. J. (1998). Comparisons among vegetation indices and bandwise regression in a highly disturbed heterogenous landscape: Mount St. Helens, Washington. *Remote Sensing of Environment*, *64*, 91–102.

Lee, N. J., & Nakane, K. (1996). Forest vegetation classification and biomass estimation based on Landsat TM data in a mountainous region of west Japan. In H. L. Gholz, K. Nakane, & H. Shimoda (Eds.), *The use of remote sensing in the modeling of forest productivity* (pp. 159–171). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Lefsky, M. A., Cohen, W. B., & Spies, T. A. (2001). An evaluation of alternate remote sensing products for forest inventory, monitoring, and mapping of Douglas-fir forests in western Oregon. *Canadian Journal of Forest Research*, *31*, 78–87.

Lewis Environmental Consultants (1992). Environmental impact assessment study: Part I. Natural environment. Canada: QIT-Madagascar Minerals Environmental Impact Assessment.

Lillesand, T., & Kiefer, R. W. (1994). *Remote sensing and image interpretation*. New York: John Wiley and Sons.

Lloyd, R. (1996). *Spatial cognition, geographic environments*. Dordrecht, Netherlands: Kluwer.

Lowry, P., Smith, P., & Rabevohitra, R. (1999). *Review of MIR Teledetection Inc. Deforestation Study in the region of Fort-Dauphin (Tolagnaro), Madagascar.* QIT–Madagascar Minerals, Contract No. MMC-0120.

Macedo, D. S., & Anderson, A. B. (1993). Early ecological changes associated with logging in an Amazon floodplain. *Biotropica*, *25*, 151–163.

Mayaux, P., Gond, V., & Bartholome, E. (2000). A near-real time forest cover map of Madagascar derived from Spot-4 VEGETATION data. *International Journal of Remote Sensing*, *21*, 3139–3144.

Medley, K. (1993). Extractive forest resources of the Tana River National Primate Reserve. *Economic Botany*, *47*, 171–183.

Moran, E. F., Packer, A., Brondizio, E., & Tucker, J. (1996). Restoration of vegetation cover in the eastern Amazon. *Ecological Economics*, *18*, 41–54.

Myers, N. (1996). The world's forests: Problems and potentials. *Environmental Conservation*, *23*, 156–168.

Myers, N., Mittermeier, R. A., Mittermeier, C., Fonseca, G., & Kent, J. (2000). Biodiversity hotspots for conservation priorities. *Nature*, *403*, 853–859.

Nelson, R., & Horning, N. (1993). AVHRR-LAC estimates of forest cover area in Madagascar. *International Journal of Remote Sensing*, *14*, 1463–1475.

Nemani, R., Pierce, L. L., Running, S. W., & Band, L. (1993). Forest ecosystem processes at the watershed scale: Sensitivity to remotely-sensed index estimates. *International Journal of Remote Sensing*, *14*, 2519–2534.

Nepstad, D. C., Verissimo, A., Alencar, A., Nobre, C., Lima, E., Lefebvre, P., et al. (1999). Large-scale impoverishment of Amazonian forests by logging and fire. *Nature*, *398*, 505–508.

Olsson, H. (1994). Changes in satellite-measured reflectances caused by thinning cuttings in boreal forest. *Remote Sensing of Environment*, *50*, 221–230.

Pearson, R. G., Dawson, T. P., Berry, P. M., & Harrison, P. A. (2002). SPECIES: A spatial evaluation of climate impact on the envelope of species. *Ecological Modelling*, *154*, 289–300.

Peterson, D. L., Spanner, M. A., Running, S. W., & Tueber, K. B. (1987). Relationship of Thematic Mapper simulator data to leaf area index of temperate coniferous forests. *Remote Sensing of Environment*, *22*, 323–341.

Puhr, C. B., & Donoghue, D. N. M. (2000). Remote sensing of upland conifer plantations using Landsat TM data: A case study from Galloway, south-west Scotland. *International Journal of Remote Sensing*, *21*, 633–646.

QMM, QIT–Madagascar Minerals, S.A. (1998). Project QMM. Social and environmental reports. Montreal, Canada: QIT–Madagascar Minerals.

QMM, QIT-Madagascar Minerals S.A. (2001). Social and environmental impact assessment. Montreal, Canada: QIT–Madagascar Minerals.

Richardson, A. J., & Wiegand, C. L. (1990). Comparison of two models for simulating the soil-vegetation reflectance of a developing cotton canopy. *International Journal of Remote Sensing*, *11*, 447–458.

Ripple, W. J. (1985). Asymptotic relationship characteristics of grass vegetation. *Photogrammetric Engineering and Remote Sensing*, *51*, 1915–1921.

Rock, B. N., Vogelmann, J. E., Williams, D. L., Vogelmann, A. F., & Hoshizaki, T. (1986). Remote detection of forest damage. *Bioscience*, *36*, 439–445.

Sader, S. A. (1989). Tropical forest biomass and successional age class relationships to a vegetation index derived from Landsat TM data. *Remote Sensing of Environment*, *28*, 143–156.

Salvador, R. (2000). An assessment of the spatial variability of basal area in a terrain covered by Mediterranean woodlands. *Agriculture, Ecosystems & Environment*, *81*, 17–28.

Sgrenzaroli, M., Baraldi, A., Eva, H., DeGrandi, G., & Achard, F. (2002). Contextual clustering for image labeling: An application to degraded forest assessment in Landsat TM images of the Brazilian Amazon. *IEEE Transactions on Geoscience and Remote Sensing*, *40*, 1833–1848.

Shyamsundar, P., & Kramer, R. (1997). Biodiversity conservation—at what cost? A study of households in the vicinity of Madagascar's Mantadia National Park. *Ambio*, *26*, 180–184.

Skole, D., & Tucker, C. (1993). Tropical deforestation and habitat fragmentation in the Amazon—satellite data from 1978 to 1988. *Science*, *260*, 1905–1910.

Song, C., Woodcock, C. E., Seto, K. C., Lenney, M. P., & Macomber, S. A. (2001). Classification and change detection using Landsat TM data: When and how to correct atmospheric effects? *Remote Sensing of Environment*, *75*, 230–244.

Steininger, M. K. (2000). Satellite estimation of tropical secondary forest above-ground biomass: Data from Brazil and Bolivia. *International Journal of Remote Sensing*, *21*, 1139–1157.

Stone, T. A., & Lefebvre, P. (1998). Using multi-temporal satellite data to evaluate selective logging in Para, Brazil. *International Journal of Remote Sensing*, *19*, 2517–2526.

Swaine, M. D., & Hall, J. B. (1983). Early succession in tropical forests. *Journal of Ecology*, *71*, 601–627.

Vermeulen, S. J. (1996). Cutting of trees by local residents in a communal area and an adjacent state forest in Zimbabwe. *Forest Ecology and Management*, *81*, 101–111.

Whitmore, T. C. (1990). *An introduction to tropical rain forests*. Oxford, UK: Clarendon Press.

Wilson, E. O. (1988). *Biodiversity*. Washington, DC: National Academy of Sciences.

Woodcock, C. E., Macomber, S. A., Pax-Lenney, M., & Cohen, W. B. (2001). Monitoring large areas for forest change using Landsat: Generalization across space, time and Landsat sensors. *Remote Sensing of Environment*, *78*, 194–203.

World Bank (2003). http://www.worldbank.org

Zou, K., Tuncali, K., & Silverman, S. (2003). Correlation and simple linear regression. *Radiology*, *227*, 617–622.